SLAC-373
LBL-30097
CALT-68-1697
T/E

PROCEEDINGS of the WORKSHOP on PHYSICS and DETECTOR ISSUES for a HIGH-LUMINOSITY ASYMMETRIC B FACTORY at SLAC

January-June, 1990

Work supported in part by the Department of Energy
under contracts DE-AC03-76SF00515, DE-AC03-76SF00098 and DE-AC03-81-ER40050

Printed in the United States of America. Available from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, Virginia 22161
Price: Printed Copy A12, Micofiche A01.

TABLE OF CONTENTS

Interaction Region

Vertex Detection

Charged Particle Tracking

Particle Identification

Electromagnetic Calorimetry

Computing

Trigger and Data Acquisition

Other Projects

erator, which has resulted in a Conceptual Design Report for such a machine (SLAC-372), as well as the continuation and refinement of the theoretical and experimental work in SLAC-353. This work was carried out in the context of a *Workshop on Physics and Detector Issues for a High-Luminosity Asymmetric B Factory at SLAC* which held four plenary meetings from January through June of 1990. The first two two-day sessions were held at SLAC, the third two-day meeting at Nevis Laboratories of Columbia University and a week-long summary week was held at SLAC. At each of these meetings there were reports from other efforts aimed at establishing a B Factory elsewhere; the number of such efforts attests to the substantial scientific interest in the detailed study of the conistency of the Kobayashi-Maskawa matrix.

Nine Working Groups were formed at the Workshop. These groups and their Coordinators were:

Interaction Region	H. DeStaebler
Vertex Detection	V. Lüth/M. Witherell
Tracking	A. Boyarski/P. Burchat
Particle Identification	B. Ratcliff/N. Roe
Electromagnetic Calorimetry	G. Godfrey/M. Tuts
Computing	F. Porter
Trigger and Data Acquisition	A. Lankford
Two-Photon Physics	D. Bauer
Physics and Simulation	S. Komamiya/Y. Nir/A. Snyder

These groups held regular meetings both at the plenary sessions and between them, and reported to the Workshop at the plenary sessions. The goals of the Physics and Simulation Group were to explore new methods of measuring CP violation in B meson decay, in order to extend the range of measurement and to reduce the integrated luminosity required to establish the effect; in this they succeeded admirably. This group also simulated a number of decay modes to establish the effective sensitivity of various techniques. The groups concerned with specific detector subsystems explored the requirements placed by physics concerns on detector performance and investigated in detail the potential of specific experimental techniques in meeting these requirements. In this context, several interesting R&D projects were spawned. This *Proceedings* is comprised mainly of reports from the Working Groups. The format of the reports varies. In some instances, the entire group produced a unified summary, while in other cases the summary is accompanied by individually-authored contributions which deal with specific questions in further detail. No attempt was made to design a single detector; this task will be undertaken at a new Workshop, the *Workshop on the Design of a Detector for a High-Luminosity Asymmetric B Factory*, which is now underway. The study of theoretical issues and their relation to detector design will also continue in this new Workshop.

The participants at the *Workshop on Physics and Detector Issues for a High-Luminosity Asymmetric B Factory at SLAC*, together with their institutional affiliations, are listed below. Thanks are due to the many people who have helped with the organizational aspects of this series of four Workshops.I would particularly like to thank Sharron Lankford, Anamaria Pacheco, Bill Wisniewski, and the staff of Nevis Labs. Ray Cowan produced the TEXformatting macros for this document as well as for SLAC-353.

Pasadena, California *David Hitlin*

WORKSHOP PARTICIPANTS

Abe, Kazuo
KEK

Albrecht, Hartwig
DESY

Aleksan, Roy
Saclay

Avery, Paul
University of Florida

Bartelt, John
Vanderbilt University

Bauer, Daniel A.
UC, Santa Barbara

Bingham, Harry H.
Lawrence Berkeley Laboratory

Bjorken, James
Stanford Linear Accelerator Center

Bloom, Elliott
Stanford Linear Accelerator Center

Bolton, Tim
Columbia University

Boyarski, Adam M.
SLAC

Breakstone, Alan M.
University of Hawaii

Briggs, Donald D.
SLAC

Browder, Tom
Cornell University

Brown, Karl L.
SLAC

Buchanan, Charles D.
UC, Los Angeles

Bulos, Fatin
SLAC

Burchat, Patricia
UC, Santa Cruz

Byers, Nina
UC, Los Angeles

Caldwell, David
UC, Santa Barbara

Cavalli-Sforza, Matteo
UC, Santa Cruz

Chauveau, Jacques
LPNHE - Univ. Paris VI - VII

Cords, Dieter
SLAC

Coupal, David P.
SSC Lab

Cousins, Robert
UC, Los Angeles

Coyle, Paschal
UC, Santa Cruz

Daoudi, Mourad
SLAC

DeStaebler, Hobey
SLAC

Dib, Claudio
UCLA

Dorfan, Jonathan
SLAC

Dubois, Richard
SLAC

Dunietz, Isard
CERN-EP Division

Dunwoodie, William
SLAC

Eidelman, Simon
Inst. of Nucl. Physics, Novosibirsk

Eigen, Gerald
CALTECH

Eisenberg, Yehuda
Weizmann Institute

Eisner, Alan M.
UC, IIRPA

Erbacher, Robin
UC, San Diego

Fox, John D.
SLAC

Franzini, Paolo
Columbia University

Glanzman, Thomas
SLAC

Godfrey, Gary L.
SLAC

Goldhaber, Gerson
Lawrence Berkeley Laboratory

Harris, Frederick A.
University of Hawaii

Hearty, Christopher
Lawrence Berkeley Laboratory

Hertzbach, Stanley S.
Univ. of Mass., Amherst

Hitlin, David G.
CALTECH

Innes, Walter
SLAC

Jaros, John A.
SLAC

Johnson, Tony
Boston Univ.

Jones, Lawrence A.
CALTECH

Jung, Chang-Kee
SLAC

Karshon, Uri
Weizmann Institute

Kayser, Boris
National Science Foundation

Kenney, Christopher
University of Hawaii

Kim, Peter C.
University of Rochester

King, Mary Elizabeth
UC, Santa Cruz

Klaisner, Lowell
SLAC

Komamiya, Sachio
SLAC

Knopfle, Karl Tasso
SLAC

Koetke, Dale S.
SLAC

Kofler, Richard R.
Univ. of Mass., Amherst

Kozanecki, Witold
CERN-PPE Division

Krizan, Peter
University of Ljubljana

Lankford, Andrew
UC, Irvine

LeDiberder, Francois
Orsay

Lee-Franzini, Juliet
State University of New York

Leith, David
SLAC

Lipkin, Harry
Weizmann Institute

Lockman, William
UC, Santa Cruz

London, David
University of Montreal

Lou, Xinchou
CERN-EP Division

Lu, Adolph
UC, Santa Barbara

Luth, Vera G.
SLAC

Lynch, Gerald
Lawrence Berkeley Laboratory

MacFarlane, David
McGill University

Mandelkern, Mark
UC, Irvine

Marsiske, Helmut
SLAC

Meadows, Brian
University of Cincinnati

Messner, Robert
SLAC

Mistry, Nariman
Cornell University

Morrison, Rollin
UC, Santa Barbara

Mtingwa, Sezaki
Argonne National Laboratory

Munger, Charles T.
SLAC

Nir, Yosef
Weizmann Institute

Nozaki, Tadao
KEK

Oddone, Pier
Lawrence Berkeley Laboratory

Odian, Allen C.
SLAC

Ogg, Michael
Carleton University

Ozaki, Hitoshi
KEK

Pellett, David
UC, Davis

Petradza, Assimina G.
SLAC

Pitman, Dale
Stanford Linear Accelerator Center

Porter, Frank
CALTECH

Prepost, Richard
Univ. of Wisconsin

Quinn, Helen
SLAC

Ratcliff, Blair N.
SLAC

Rehak, Pavel
Brookhaven National Laboratory

Roe, Natalie
Lawrence Berkeley Laboratory

Ronan, Michael
Lawrence Berkeley Laboratory

Rowson, Peter
Columbia University

Rubin, Phil
Syracuse University

Sagawa, Hiroyuki
KEK

Schlein, Peter
UCLA

Schultz, Jonas
UC, Irvine

Selove, Walter
Univ. of Pennsylvania

Shaevitz, Michael
Columbia University

Shapiro, Gilbert
Lawrence Berkeley Lab

Shapiro, Stephen L.
SLAC

Shen, Benjamin C.
UC, Riverside

Silverman, Dennis, J.
UC, Irvine

Simopoulos, Constantine
SLAC

Smith, Andrew
UC, Irvine

Snyder, Arthur E.
SLAC

Spahn, Martin
DESY

Spengler, Joachim
DESY

Spieler, Helmut
Lawrence Berkely Laboratory

Steiner, Robert
Adelphi Univ.

Stone, Sheldon
Syracuse University

Stoker, David P.
SLAC

Sullivan, Michael
UC, IIRPA

Thaler, Jon J.
Univ. of Illinois

Tichonov, Yury
Inst. of Nucl. Physics, Novosibirsk

Toge, Nobukazu
SLAC

Toki, Walter H.
SLAC

Tuts, Michael
Columbia University

Va'Vra, Jaroslav
SLAC

Vernon, Wayne
UC, San Diego

Wagner, Steven
SLAC

Wang, Yao-Xun
UC, IIRPA

Ward, B. F. L.
University of Tennessee

Weinstein, Alan
CALTECH

Williams, David A.
SLAC

Wisniewski, William
CALTECH

Witherell, Michael
UC, Santa Barbara

Wormser, Guy
Orsay

Ypsilantis, Tom
CERN-EP Division

Yurko, Michael
Univ. of Indiana

Zapalac, Geordie H.
SLAC

INTRODUCTION

DAVID HITLIN

Despite more than a quarter century of investigation, the phenomenon of CP violation remains a tantalizing mystery. One of the deepest attributes of weak interactions (it is responsible for the predominance of matter over antimatter in the universe), it has been directly observed only in the K meson system. No unique explanation of the effect presently exists. Ongoing experiments in the K system promise to attain sufficient precision to tell us whether the superweak model of Wolfenstein is ruled out, but are unlikely to further enlighten us as to the origin of the phenomenon. The six-quark Standard Model naturally incorporates CP violation through the phase of the Cabibbo-Kobayashi-Maskawa quark-mixing matrix, but we know very little about the imaginary part of the matrix.

The B meson system can serve as a laboratory both for deepening our understanding of CP violation and, potentially, for performing quite stringent consistency tests of the Standard Model, by providing an overconstrained set of measurements of the Cabibbo-Kobayashi-Maskawa matrix elements. The CP-violating asymmetries in the B system are large (tens of percent), and for those decays involving CP-eigenstate final states, are thought to be cleanly interpretable in terms of CKM matrix elements. B meson decays to CP eigenstates typically have small branching ratios, however, making it necessary to produce of the order of 10^8 B mesons to make the measurement. The best way to make such a measurement appears to be to produce $B\bar{B}$ pairs in the decay of the $\Upsilon(4S)$ resonance at an *asymmetric* e^+e^- collider. Sufficient design work on such an asymmetric storage ring has now been done (see SLAC-372, for example) that there is reasonable confidence that such a machine is practical, and likely to produce the 30 fb^{-1} per year needed to cover the allowed range of CP-violating asymmetries. The purpose of this Workshop is therefore to explore both the outstanding physics issues in this area and the technologies available to us to build a detector capable of doing the physics.

In the area of physics, there are a number of interesting questions, many of which are addressed in these *Proceedings*:

1. Are there modes other than the "classical" ones such as $B^0 \to J/\psi K_S^0$ and $B^0 \to \pi^+\pi^-$, which are useful for measuring the angles of the CKM Unitarity Triangle through measurements of CP-violating asymmetries?
2. If so, what techniques are required to exploit these new modes?
3. Can these new modes, if they exist, be used to reduce the integrated luminosity required to establish a non-zero CP-violating asymmetry?
4. How can different decay modes, which purport to measure the *same* angle of the Unitarity Triangle, be used to test for breakdown of specific Standard Model assumptions?
5. Can our assumptions about the direct connection between measured asymmetries and angles of the unitarity triangle be tested? That is, can we be sure that the decays of interest are dominated by a single weak phase?

Much more work is required to provide detailed answers to these and other physics questions, but some of the answers are now clear:

1. There are several other useful modes, which fall into distinct categories:
   (a) Other CP eigenstate modes, such as D^+D^-, are measurable.
   (b) Decays to states which are not CP eigenstates themselves, but which decay to CP eigenstates, such as $J/\psi K^{*0} \to J/\psi K_S^0 \pi^0$ are analyzable.
2. Moments analysis and transversity analysis can be used to separate decays of the B^0 to two vector particles into partial waves with definite CP, making them useful for asymmetry measurements.
3. Decays to non-CP-eigenstates which are nonetheless self-conjugate collections of quarks can be extremely useful for asymmetry measurements.

4. "Penguin pollution", that is, the potential presence of a second weak amplitude, can be tested for, both through comparison of isospin-related final states and through detailed analysis of time development.

Simulations of several of these modes are discussed herein. Others are the subject of current investigation.

One of the central objectives of the Workshop, albeit one which is not treated explicitly here, is the development of improved Monte Carlo tools, both for the purpose of improving physics simulations, and for studying specific detector designs. This development is, and will be, ongoing.

A B Factory, while it is primarily aimed at the study of CP violation in the B meson system, is also a rich source of many other types of physics. Studies of charmed particles and τ leptons, of transitions between the Υ resonances and of two-photon physics are all possible at new levels of precision. Many of these capabilites are treated in SLAC-353, the precursor to this document. These *Proceedings* do, however, contain a study of the capability of an asymmetric B Factory in the area of two-photon physics.

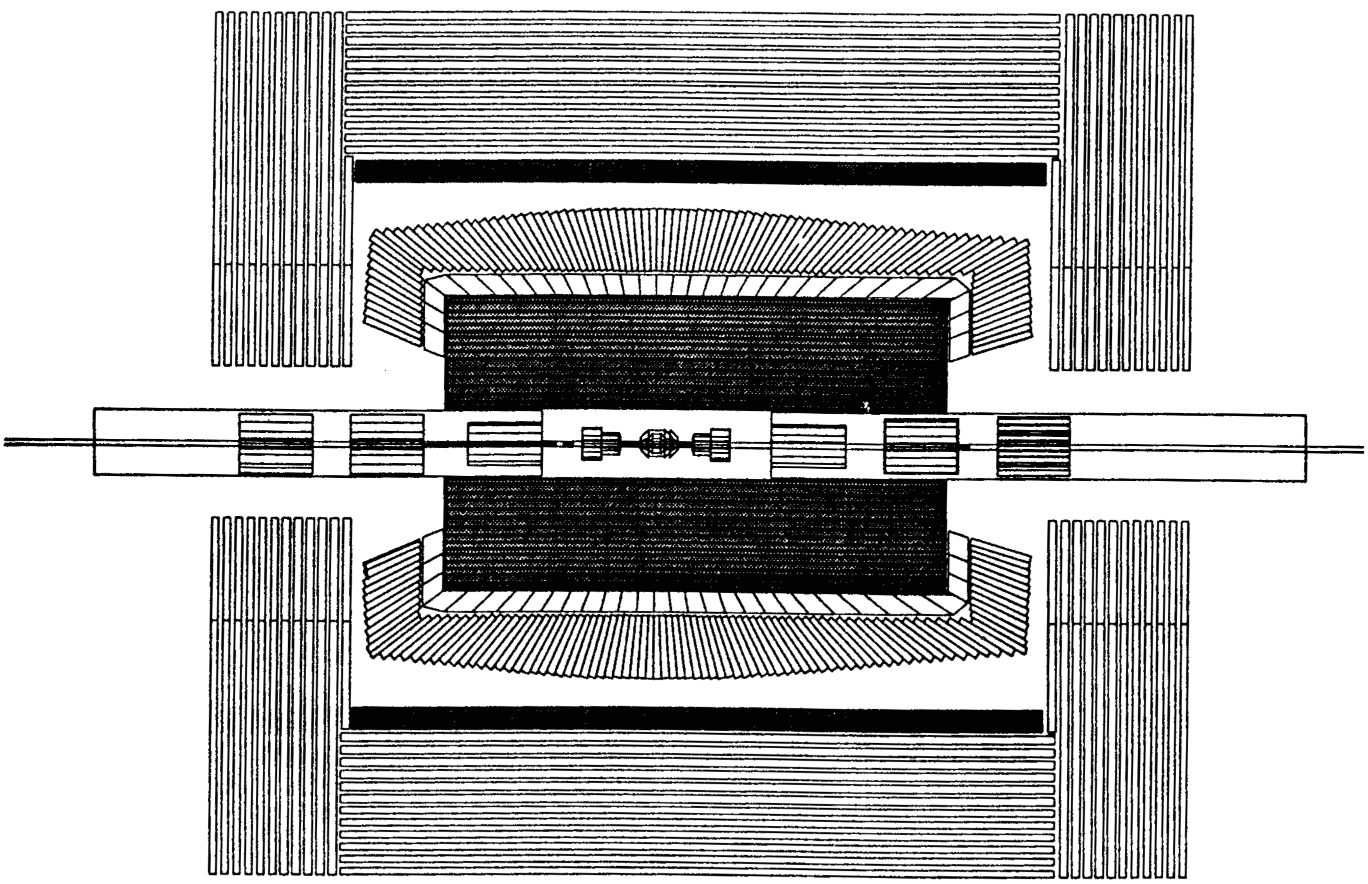
A high-luminosity, asymmetric e^+e^- storage ring presents new and challenging detector design questions. Trigger rates are such that pipelined data acquisition architectures are required. Masking of the detector against synchrotron radiation and lost beam particles is of crucial significance. Radiation damage to vertex detectors, drift chambers and calorimeters becomes a serious issue. Very large data samples will have to be acquired, stored and analyzed. These issues are discussed in some detail in these *Proceedings*.

High-precision vertex measurements along the beam direction, are central to the measurement of CP-violating asymmetries. Both single-and double-sided silicon strip detectors and the new pixel devices currently under development are attractive candidates for this function. It will be necessary to balance concerns about detector thickness, the amount of dead material, power consumption, resolution, robustness of pattern recognition and radiation hardness in reaching a decision about the final configuration.

Tracking chambers provide the primary momentum measurement for charged particles. In choosing a magnetic field value, it is necessary to balance momentum resolution at the highest momenta against efficiency for low p_t tracks. Both jet cell and small cell drift chambers have been considered in some detail. Total radiation dose to the chamber is an issue at a B Factory; margins seem adequate, but not luxurious. A major advance in the context of the Workshop has been the development of helium-based gases with six ties the radiation length of conventional argon-based drift chamber gases.These afford a substantial reduction in the multiple scattering contribution to momentum resolution.

Given the wider range of laboratory momenta due to the moving center-of-mass, and the importance of kaon and lepton tagging in CP asymmetry measurements, particle identification takes on increased importance for a B Factory detector. We have studied dE/dx, time-of-flight and both ring-imaging and aerogel threshold Čerenkov techniques for particle identification, spawning several ongoing R&D projects in the area of Čerenkov counters. A ring-imaging counter provides the best coverage, but is complicated and expensive. There are also questions about the amount of material which can be tolerated in front of the high-quality electromagnetic calorimeter which is also a crucial part of the detector. Aerogel counters potentially provide adequate coverage with less material, if it can be demonstrated that they yield a sufficient number of detectable photons.

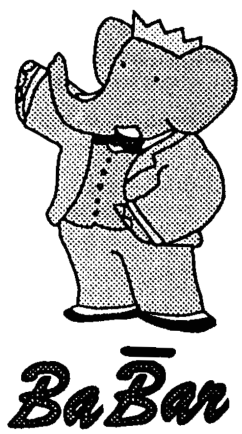
A high-resolution, highly segmented electromagnetic calorimeter, efficient for low-energy photons, is likely to be the single most expensive component of a new detector. Of the scintillating crystals, CsI(Tl) appears to be the best choice. A readout system for the crystals in a magnetic field which preserves the low-energy photon resolution and efficiency is the subject of several ongoing investigations. It appears that substantial improvement over current capabilites can be achieved. Some questions as to the radiation-hardness of CsI(Tl) remain, although much progress has been made in this area. A liquid krypton calorimeter has also been investigated. This device provides advantages of stability, ease of calibration and



radiation hardness, but the inevitable material in the dewar seems to present a problem for low-energy photons.

It may be useful to segment the iron flux return of the solenoidal magnet to allow identification of low- momentum muons by range. A substantial extension of the lower momentum cutoff appears possible.

While it was not the purpose of this Workshop to make final decisions as to the design of a detector for the SLAC asymmetric B Factory, it has proved useful to provide a 'straw man" design to focus the study of tradeoffs, coverage, cost, *etc.* This design in shown in the accompanying figure. In order to maximize laboratory solid angle coverage at minimum cost, the detector is asymmetric, emphasizing the boost direction. The scale can be ascertained from the drift chamber outer radius, which is 80 cm. Further progress on detector design is ongoing in the *Workshop on the Design of a Detector for a High-Luminosity Asymmetric B Factory* which is currently underway



PHYSICS

and SIMULATION

How to Extract *CP*-Violating Asymmetries from Angular Correlations

I. Dunietz, H. J. Lipkin, H. Quinn,
A. Snyder and W. Toki

1. Introduction

Our ability to study *CP* violation in neutral B decays[1] is considerably broadened if we can include modes in which different partial waves contribute to the decay with different *CP* parities.[2–6] Many such modes exist; for example, those where the B^0 decays to two particles with spin, such as ψK^{*0} or $D^{*+}D^{*-}$. The asymmetry in the total rate from such a channel suffers from a partial cancellation or dilution of the asymmetry from the two different *CP* contributions. Hence such modes require an angular analysis of the decays of the spinning particles to separate out definite *CP* contributions and thus obtain asymmetry measurements that probe the basic Standard Model predictions.[7] Of course, if nature is kind and a single *CP* channel dominates the decay, then the *CP* asymmetry may be approximately measured without any angular analysis. However, in these cases, an angular analysis can be performed without any loss in statistical accuracy and without any error from the small opposite *CP* contribution to yield a more precise measurement of the *CP* asymmetry.

Many modes can be used to measure CP violation and extract fundamental model parameters.

The particle content of all the modes discussed here is such that one can construct *CP* eigenstates from a superposition of helicity states, without invoking a different particle content. Thus, for example, the modes $\psi K_S^0 \pi^0$ and $D^{*+}D^{*-}$ are considered here, but not modes such as $D^{*+}\rho^-$. This report presents several different approaches to the angular analysis. All are based on standard helicity formalism.[8–10] The merits of the various approaches depend on a number of factors, many of which are not yet known, such as the relative strengths of the different helicity amplitudes. By the time sufficient data is accumulated to attempt any of these analyses, a great deal more will be known about these factors.

For any given channel the preferred method will be clear. We present here four approaches and briefly discuss the merits of each.

Transversity provides the simplest angular analysis capable of isolating definite CP contributions.

The first approach analyzes events in terms of a quantity we call transversity, which characterizes the spin projections of a three body intermediate state in a direction transverse to the plane of the three body system.[9] This approach requires the minimum amount of angular analysis to arrive at definite *CP* quantities. We show that, in certain cases, moments of the data with respect to a single polar angle can achieve the required separation. This method has the advantage that it allows us to sum resonant and non–resonant contributions to certain final states, whereas the more detailed angular analysis requires reconstruction of a specific two–body parent system for the three–body state. For another simple method applicable for some modes, see Ref. 4.

The second method uses a more complete angular analysis and forms all possible independent angular moments of the data. This allows the study of additional channels not amenable to the transversity treatment. Like the transversity moment analysis, it allows asymmetries to be extracted without *a priori* knowledge of the relative strengths of the different helicity contributions. In both cases, this can be done by combining results from both B^0 and $\bar{B}^0$ decays.

When data from isospin–related channels is available maximum likelihood fits to all channels and to angular distribution provide the most accurate CP–asymmetry measurement.

The remaining two methods use a maximum–likelihood fit to the angular structure of the *CP*–violating decay and to a set of isospin–related channels which are not influenced by *CP*-violating effects. This can be done with either the transversity polar angle distributions or the full angular distributions. For a transversity analysis of this type one needs to know the relative strengths of the contributions for each possible absolute value of the transversity. This can frequently be determined from isospin–related channels.[11] The full angular analysis requires determination of the full set of helicity amplitudes and their relative strong–interaction phases. Data from isospin–related channels may make this possible, providing the most accurate measure of the asymmetry for those modes where sufficient data is available to well-determine all the necessary quantities.

The plan of this report is as follows. In Section 2 we introduce some general notation and review the dilution of asymmetry that occurs when two different *CP* channels contribute to a given final state. Section 3 presents a discussion of transversity analysis, Section 4 reviews the many channels for which it can be used, and Section 5 presents as an example the transversity analysis for the case of two spin one particles. Results from more complete angular analyses are discussed in Section 6. Section 7 reviews the accuracy of the asymmetry measurements obtained by each of the methods and discusses the relative advantage of maximum-likelihood methods compared to moment analyses. Section 8 contains some concluding observations. Appendix A contains a proof that transversity is a projector for definite *CP*, Appendix B contains the details of the full angular analysis, and Appendix C presents a summary of an analysis of the sensitivity of results to various measurement errors.

2. PRELIMINARIES—DILUTION OF *CP* VIOLATION

This section introduces some notation and discusses the dilution in the *CP* violating asymmetry when the final state is a mixture of different angular momenta which contribute with different parity and hence different *CP*. One can most readily treat these processes using the helicity formalism, which gives a correct relativistic analysis of the angular momentum in the decay process. This is a well established formalism which provides the basis for analysis of angular structure in the subsequent decays of the two spinning particles.[2,8,10]

We begin by discussing the results obtained for such processes without any angular analysis. We show that the asymmetries thus measured depend on the ratio of *CP*-even to *CP*-odd contributions and are diluted, that is, reduced in magnitude, relative to the asymmetry of a pure *CP* state. We denote a time-evolved, initially pure B^0 as $B^0_{\rm phys}$. Any rate difference between the process, $B^0_{\rm phys} \to f$, and the *CP* conjugated process, $\bar{B}^0_{\rm phys} \to \bar{f}$, signals *CP* violation. The rate difference comes about because the processes have each two interfering contributions to each partial wave or helicity amplitude, see Figure 1.

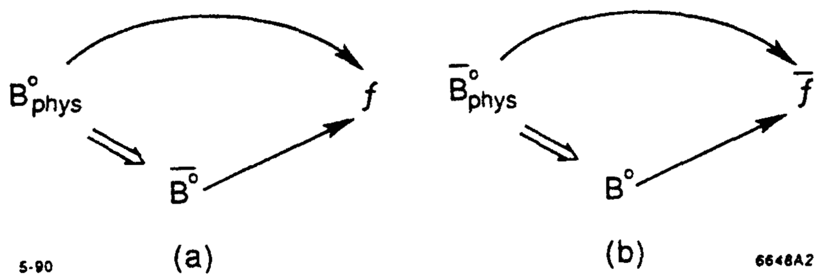


Figure 1. *Schematic representation of two paths (a) from B^0 to the final state f, direct or via mixing to the $\bar{B}^0$ followed by decay, and (b) from $\bar{B}^0$ to the final state $\bar{f}$, the CP conjugate of f, via direct decay or via mixing to B_0 followed by decay.*

The *CP* violating interference term is denoted by $\mathrm{Im}\,\lambda_{KM}$. The rate of a B^0 to f is

$$\Gamma(B^0_{\mathrm{phys}} \to f) = \Gamma_+(1+a) + \Gamma_-(1-a) \ , \tag{2.1}$$

and for the $\bar{B}^0$ to $\bar{f}$ is

$$\Gamma(\bar{B}^0_{\mathrm{phys}} \to \bar{f}) = \Gamma_+(1-a) + \Gamma_-(1+a) \ . \tag{2.2}$$

The *CP*–even and *CP*–odd rates are parametrized by the widths Γ_+ and Γ_-, respectively. The parameter a is proportional to $\mathrm{Im}\,\lambda_{KM}$, and would be the asymmetry if the *CP*–even state dominated. The rates of Eqs. (2.1)–(2.2) could be time dependent or time integrated. In the former case[1]

$$a = -\mathrm{Im}\,\lambda_{KM} \sin(\Delta m\, t) \ , \tag{2.3}$$

and Γ_+ and Γ_- contain a factor $e^{-\Gamma t}$, where Γ is the width of the B^0.⋆ In this case the analysis of angular distributions must be made for each time bin separately, since the asymmetry is different at different times. Because the angular dependence and the time dependence factorize this introduces no particular complication for the extraction of the *CP* asymmetry

⋆ We here make the approximation that the heavy and light mass eigenstates of the $B^0\bar{B}^0$ system have the same widths, expected to be correct to within 10^{-3}.

from the angular information; the method is the same for every time–bin data set. For an experiment which measures a time–integrated asymmetry the prediction is

$$a = \frac{-x \operatorname{Im} \lambda_{KM}}{1+x^2} \,, \tag{2.4}$$

where $x \equiv \Delta m / \Gamma$, and $\Gamma_\pm$ denote time–integrated quantities.

The asymmetry measured without angular analysis may be diluted by cancellation between Γ_+ and Γ_- contributions.

The measured asymmetry is

$$\text{Asym} \equiv \frac{\Gamma(B^0_{\text{phys}} \to f) - \Gamma(\bar{B}^0_{\text{phys}} \to \bar{f})}{\Gamma(B^0_{\text{phys}} \to f) + \Gamma(\bar{B}^0_{\text{phys}} \to \bar{f})} = a \, \frac{\Gamma_+ - \Gamma_-}{\Gamma_+ + \Gamma_-} \,. \tag{2.5}$$

The last factor gives a dilution when the final state f is an admixture of *CP*–even and *CP*–odd parities. No information currently exists on the ratio Γ_+/Γ_- and large dilutions could occur. Study of angular distributions allows us to avoid such dilutions regardless of the Γ_+/Γ_- ratio.

3. Transverse Projection as an Analyzer of *CP* Parities

Consider the decay of a spinless neutral particle B^0 into unstable particles A and C. (We require A and C to be unstable so that spin information can be learned from their subsequent decay.) All the subsequent discussion holds equally for decays of any neutral spin zero particle, in particular for B_s and D^0 which we will discuss later. Let the particles A and C have spins s_a, s_c, helicities λ_a, λ_c, and intrinsic (reflection) parities π_a, π_c, respectively. We consider cases where C is seen in a two body mode $C \to C_1 C_2$, with spins s_1 and s_2.

$$\begin{aligned} B^0 \to A&C \\ &\hookrightarrow C_1 C_2 \end{aligned} \tag{3.1}$$

A simple example to keep in mind is the case $A = \psi$, $C = K^{*0}$, $C_1 = K^0_S$, $C_2 = \pi^0$.

Let us define the transverse axis as the normal to the plane containing the three particles, AC_1C_2, in either the

Transversity measures the spin component along the transverse axis in the rest frame of the particle.

B^0 or C rest frame (or in the A rest frame where the plane is defined by the particles C_1 and C_2). The CP-parity eigenstates of the mode AC_1C_2 can be classified by the spin projection of the particles along this transverse axis, which we call the transversity. The state of transversity τ_i of each particle is defined as that linear combination of helicity states which represents a spin component τ_i along the transverse axis in the rest frame of particle i. This definition is invariant with respect to boosts between the C rest frame, in which we analyse C decay, and the particle's own rest frame, which will be used to analyse its decay.

In Appendix A we prove, using the helicity formalism, that projection onto states of definite transversity $\tau = \tau_a + \tau_1 + \tau_2$ projects out quantities of definite CP. The following argument gives a more intuitive understanding of this result. First consider three particles moving in a plane, and let the y-axis be chosen transverse to the plane. A reflection about that plane can be written as a product of a space inversion P and a 180° rotation about the transverse axis.

$$R_P \equiv P e^{i\pi J_y} = P_{int} \cdot e^{i\pi\tau} \tag{3.2}$$

where P_{int} denotes the total intrinsic parity of the three particle system, J_y denotes the projection of the total angular momentum of the three particle state on the y axis and τ denotes sum of the transversities of the three particles.[12] The three body state can be viewed as a product of three one-body plane wave states, the reflection acts independently on each particle.

The operator R_p has been used extensively in applying the consequences of space-time symmetries to four-point functions; i.e. processes characterized[13] by three independent momenta. In a relativistic group-theoretical description, the operator R_p is seen to be the generator of the "coplanar little group"; i.e. the subgroup of the inhomogeneous Lorentz group which leaves three momenta invariant.[14]

To evaluate the action of the reflection operation on any one particle we can use the fact that reflection commutes with boosts in the plane. We thus go to the rest frame of the particle under consideration. In that frame one readily sees from the definition of the reflection operator that its eigenvalue for the particle j is the product of intrinsic parity times $e^{i\pi\tau_j}$, where τ_j is defined as the spin projection along the transverse axis. Equation (3.2) simply combines the result for each of the three particles to give the eigenvalue for a three particle state of definite transversity.

In decays like $B_d \to J/\psi K_S^0\pi^0$ and $B_d \to \eta_c K_S^0\pi^0$, each of the three particles in the final state has a definite intrinsic CP. For such cases one can define the operation of the product of charge conjugation and the reflection in the plane.

Because the system has zero total angular momentum, the quantity R_{CP} is equal to CP.

$$R_{CP} \equiv CR_P \equiv CPe^{i\pi J_y} = CP = (CP)_{int}e^{i\pi\tau} \qquad (3.3)$$

where $(CP)_{int}$ denotes the product of the intrinsic CP of the three particles. The first equality of Eq. (3.3) is true since the initial state has spin zero, hence the final state must also have spin zero and be invariant under rotations in the center of mass system, and the second equality follows from Eq. (3.2).

For example, any $J = 0$ state of the type $\left|(c\bar{c})K_S^0\pi^0; J = 0\right\rangle$ in which the $(c\bar{c})$ has a definite intrinsic CP and is in an eigenstate of transversity τ can be shown to be a CP eigenstate with CP-parity given by the relation

$$CP\left|(c\bar{c})K_S^0\pi^0; J = 0\right\rangle = R_{CP}\left|(c\bar{c})K_S^0\pi^0; J = 0\right\rangle$$

$$= (CP)_{int}(-1)^\tau\left|(c\bar{c})K_S^0\pi^0; J = 0\right\rangle \qquad (3.4)$$

The relation (3.4) applies to any three-body system with a well defined intrinsic CP for each particle. It also applies if particle A does not have definite intrinsic CP but decays to a state of definite CP, for example $D^0 \to \pi^+\pi^-$. In this case we define the intrinsic CP of particle A to be the CP of its decay channel. This allows a considerable extension of the class of channels that can be used for CP analysis. Modes such as $\eta_c K_S^0\pi^0$, with three spinless particles have $\tau = 0$ and the CP is the intrinsic CP of the three particles. For the final state $J/\psi K_S^0\pi^0$, for which the intrinsic CP is odd, the total CP is odd if τ is zero and even if τ is ± 1. Note also that similar

Polar angular distributions about the transverse axis can be used to isolate $|\tau_i|$.

results apply also to all radial excitations of the charmonium states.

For each of the three particles, the polar angular distribution of its decay with respect to the transverse axis can be used to separate contributions for each $|\tau_i|$, integrated over all other decay angles. From each set of $|\tau_i|$ one can then extract a measurement of the undiluted asymmetry. These measurements can be combined to give an improved value but their errors are highly correlated and must be treated correctly, as is discussed in Section 7.

When particles C_1 and C_2 are spinless, two further classes of decays can be analysed using transversity. Table 1 summarizes the situation, similar results for the full angular analysis have been tabulated by Dell'Aquila and Nelson.[2] The first column of Table 1 defines the classes of decays of a spinless neutral particle that can be analysed for CP asymmetries using transversity projections. For each class, Table 1 defines the quantity ξ such that the CP is given by

$$CP = \xi(-1)^{\tau} \tag{3.5}$$

Examples for each class of decay defined in Table 1 are shown in Tables 2–4. Whenever decays of the spinning particles allow projection of the magnitude of the transverse spin, the data can be separated into definite CP classes. The errors on the various transversity projections are correlated, so care must be taken when combining results.

Table 1. *The CP parity for the mode AC_1C_2 with transversity τ. The first column defines possible classes. The symbol $X \leftrightarrow \bar{X}$ denotes that particle X is either a CP eigenstate or is observed via its decay into a CP eigenstate. $\eta(X)$ denotes the intrinsic CP-parity of particle X.*

CLASS		Example AC_1C_2	CP Parity $\equiv \xi\ (-1)^{\tau}$
(1)	$A \leftrightarrow \bar{A}$ $C_1 \leftrightarrow \bar{C}_1 \quad C_2 \leftrightarrow \bar{C}_2$	$\psi K_S^0 \pi^0$	$\eta(A)\eta(C_1)\eta(C_2)(-1)^{\tau}$
(2)	$A \leftrightarrow \bar{A} \quad C \leftrightarrow \bar{C}$ $C_1 = \bar{C}_2$ $S_{C_1} = S_{C_2} = 0$	$D^{*0}(\pi^+\pi^-)_\rho$	$\eta(A)(-1)^{s_c+\tau}$
(3)	$A = \bar{C}$ $S_{C_1} = S_{C_2} = 0$	$D^{*+}D^{*-}$	$(-1)^{s_c+\tau}$

For Class 1, the CP does not depend on the spin of particle C. Thus, it is not necessary to determine that C_1C_2 arise from the decay of a well-defined particle C. Hence, in this class of decays, the resonant and non-resonant production of C_1C_2 can be combined in the data sample, since all events of a given τ contribute with the same CP. This may allow the transversity analysis for such a channel in cases where the full angular analysis cannot be reliably used because of wrong spin backgrounds. This will probably be the most useful application of transversity analysis. In Class 1 the particles C_1 and C_2 may have any integer spin as long as their subsequent decay allows reconstruction of their transversity.

Resonant and non-resonant production of the system C_1C_2 can be combined for Class 1 processes analysed using transversity.

In Class 2, C_1C_2 must have a well-measured total spin (modulo 2), but not necessarily a unique parent particle C. In this situation the helicities of particles C_1 and C_2 are interchanged as well as sign-reversed under CP. Hence, we must require that particles C_1 and C_2 have spin zero in order to form definite-CP quantities using transversity projections.

For Class 3, particle C must be identified as the antiparticle of A, and again the transversity analysis can only be applied when both C_1 and C_2 have spin zero.

One last comment on Table 1. Whereas $X \leftrightarrow \bar{X}$ demands that X is seen in a CP eigenmode, $Y = \bar{X}$ puts no constraints on the decay products of either X or Y. For example, Class 3 allows any decay mode for particle A provided it allows transversity projections to be made, and requires only that C decay to two spinless particles.

4. Some Modes which can be Analysed Using Transversity

Equipped with Table 1 and its interpretation, we can increase the number of modes that can be used for CP violation studies, without dependence on any specific model. The particle content of all the modes discussed here is such that one can construct CP eigenstates from a superposition of helicity states, without invoking a different particle content. Thus, for example, the modes $\psi K_S^0\pi^0$ and $D^{*+}D^{*-}$ are considered here, but not modes such as $D^{*+}\rho^-$. The pure CP

When a final state D^0 is present it must decay to a definite CP state in order to use these modes, but see Table 4.

eigenmodes of $\bar{B}_d$, such as ψK_S^0, D^+D^-, $D^0\rho^0$, and $D^0\pi^0$, can now be augmented by the many modes given in Table 2. This Table is not exhaustive, the reader will see obvious extensions of the list presented here.

In the Standard Model with three generations of quarks, we can study the three angles of the unitarity triangle, see Figure 2. Modes of $\bar{B}_d$ driven by the quark subprocesses

$$b \to s + \bar{q}q, \qquad c + \bar{c}s, \qquad c + \bar{c}d, \qquad c + \bar{u}d \tag{4.1}$$

are all governed by $\sin(2\beta)$. The $b \to s$ transition via a penguin is denoted by $b \to s + \bar{q}q$. The interference term is $\sin(2\alpha)$ for processes governed by the $b \to u + \bar{u}d$ quark subprocess. Again, several modes can be analysed. However, for this quark subprocess, because only light quarks occur in the final state, there may be competing contributions from penguin amplitudes which have similar CKM strength but different CKM phases.[15] These must be considered in assessing the Standard Model prediction.

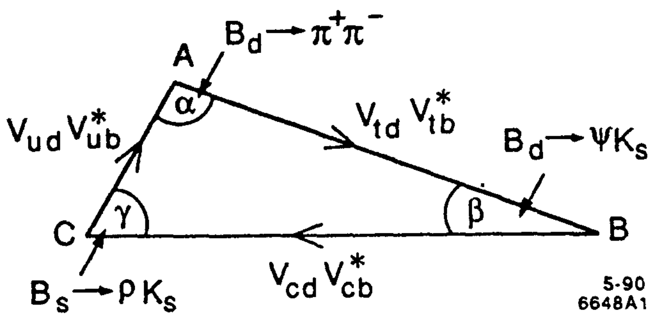


Figure 2. *The unitarity triangle for the three generation standard model, showing the definitions of the angles α, β and γ and some processes that could be used to measure each angle.*

For the Class 1 processes, it is irrelevant whether the $K_S^0\pi^0$ arises from K^{*0} or non-resonant production, as discussed above. In fact, for any three-body mode of Class 1, AC_1C_2, there is no need to find a pseudo-two-body mode AC. The C_1C_2 could come from non-resonant as well as resonant production.

To use these modes one must have good separation of γ and π^0 contributions.

For Class 2, the D^{*0} of the mode $D^{*0}\rho^0$ must be seen in a *CP* eigenmode. Either of two decay chains qualify:

$$D^{*0} \to \gamma f_D \qquad D^{*0} \to \pi^0 f_D$$

where f_D denotes any *CP* eigenstate produced from D^0 decay.

Both processes occur through $L = 1$, because of parity conservation. Note, however, that it is important in such cases to be able to distinguish between the photon and the π^0 as these have opposite intrinsic *CP*, and hence give opposite *CP* contributions for the same transversity.[6]

Quark subprocess	Class (1)	Class (2)	Class (3)	Full Angular Analysis
$b \to c\bar{c}s$	$\psi K_S^0 \pi^0$, $\psi'' K_S^0 \pi^0$	$(K_S^0\pi^0)_{\bar{K}^*}(D\bar{D})_{\psi''}$		$\psi(\rho K_S^0)_{\bar{K}_1}$
$b \to c\bar{u}d$	$\omega\pi^0 f_D$, $\rho^0\pi^0 f_D$ $\omega\omega f_D$, $\rho^0\rho^0 f_D$ $\omega\pi^0 f_{D^*}$, $\rho^0\pi^0 f_{D^*}$	$f_{D^*}(\pi^+\pi^-)_{\rho^0}$		$\omega(\gamma f_D)_{D^*}$ $a_1(\gamma f_D)_{D^*}$
$b \to c\bar{c}d$	$\psi\rho^0\omega$ $\psi\rho^0\pi^0$ $\psi\omega\pi^0$	$\psi(\pi^+\pi^-)_{\rho^0}$	$(\pi D)_{D^*}(\pi\bar{D})_{\bar{D}^*}$ $(\gamma\bar{D})_{\bar{D}^*}(\pi D)_{D^*}$ $(\gamma D)_{D^*}(\pi\bar{D})_{\bar{D}^*}$	$(\gamma D)_{D^*}(\gamma\bar{D})_{\bar{D}^*}$ $\psi\omega$ ψa_1
$b \to u\bar{u}d$	$\omega\omega\rho^0$, $\omega\rho^0\pi^0$ $\omega\omega\pi^0$, $\omega\omega\omega$	$\omega(\pi^+\pi^-)_{\rho^0}$ $a_1^0(\pi^+\pi^-)_{\rho^0}$	$\rho^+\rho^-$ $\rho^0\rho^0$	$\omega\omega$ $a_1^+a_1^-$, $a_1^0a_1^0$ ωa_1^0, $\Delta\bar{\Delta}$

Table 2. *Examples of $\overline{B}_d$ modes, which are mixtures of CP eigenstates, that can be studied with an angular analysis. Here f_D denotes any CP eigenmode of D^0 and $(C_1C_2)_C$ denotes particles C_1 and C_2 coming from a parent particle C.*

Some further comments on the processes listed in Table 2 follow:

In the Class 3 processes $D^{*+}D^{*-}$, $\bar{D}^{*0}D^{*0}$ the D^*'s can be studied in all decay modes. We do not require the neutral D^0, which could arise in the decays $D^* \to \pi D^0$ or $D^* \to \gamma D^0$, to decay to a *CP* eigenmode. However, we do need at least one of the D^*'s to decay to two spin-zero particles (usually πD).

The final column of Table 2 lists a few of the many additional modes that can be analysed using full angular analysis, which we discuss in Section 6. The modes listed here are not accessible via transversity analysis alone. In contrast, any quasi–two– body mode that can be analysed using transversity can also be treated by the more complete angular analysis which we will discuss later.

Table 3 presents a similar list for the decays of the $\bar{B}_s$. For all modes of B_s of the type studied here that are driven by the quark subprocesses of Eq. (4.1), the *CP* asymmetries are predicted to be tiny in the Standard Model. In contrast, modes mediated by the $b \to u\bar{u}d$ subprocess have a *CP* asymmetry proportional to $\sin(2\gamma)$ which could be large. Any modes of the type $X^0Y^0K_s$ or $X^0Y^0(K^0_S\pi^0)_{K^*}$ belong to Class 1 of Table 1 and can be used to study *CP* violating asymmetries. Here X^0Y^0 is any pair of light neutral mesons of zero total strangeness which decay in such a way that transversity can be reconstructed. The transversity analysis can thus also considerably enrich the possibilities for a measurement of $\sin(2\gamma)$. Here again, however, the contributions of penguin amplitudes may complicate the theoretical predictions.

Table 3. *Examples of $\bar{B}_s$ modes which are admixtures of CP eigenstates that can be studied with an angular analysis. Here $f_{D^{*0}}(f_D)$ denotes any CP eigenmode of $D^{*0}(D^0)$.*

Quark–subprocess	Class 1	Class 2	Class 3	Full Angular Analysis
$b \to c\bar{c}s, s$	$\psi\phi\pi^0$	$\psi(K^+K^-)_\phi, \psi''\phi, \ldots$	$\phi\phi$	$(\gamma D_s^+)_{D_s^{*+}}(\gamma D_s^-)_{D_s^{*-}}$
$b \to c\bar{u}d$	$f_{D^{*0}}\pi^0K^0_S$			$(\gamma f_D)_{D^{*0}}(\rho^0K^0_S)_{K_1}$
$b \to c\bar{c}d$	$\psi\pi^0K^0_S$, $\psi''\pi^0K^0_S$			$\psi(\rho^0K^0_S)_{K_1}$
$b \to u\bar{u}d$	$\omega\pi^0K^0_S$, $\rho^0\pi^0K^0_S$, $a_1\pi^0K^0_S$ $\omega\omega K^0_S, \omega\rho^0K^0_S$			$\omega(\rho^0K^0_S)_{K_1}, \rho^0K_1, a_1K_1, \ldots$

Consider now $\bar{B}^0$ decays which are driven by $b \to c\bar{u}d$ and produce a neutral D. Such modes can be used for *CP* violation studies[1,6] when this neutral D decays into a *CP* eigenstate. It is therefore advantageous to increase the data sample for D^0 decays into *CP* eigenstates. Hence in Table 4 we list modes that can be analysed by applying the same type of transversity analysis to the D^0 decay itself. This may in turn allow significant increase in the analyzable data sample of B decays. The Mark III collaboration has already determined that the $D^0 \to \rho^0K^{*0}$ is dominated by the S– and D–waves. That means that this mode is dominated by a single *CP* when the K^* decays to $K^0_S\pi^0$, and hence this mode can be readily treated with this analysis without significant loss of statistical accuracy compared to a pure *CP* channel.

Class 1	Class 2	Class 3	Full Angular Analysis
$\omega\pi^0 K^0_S$, $\rho^0\pi^0 K^0_S$ $\rho^0(K^0_S\pi^0)_{K^{*0}}$	$\phi(\pi^+\pi^-)_{\rho^0}$, $\rho^0(K^+K^-)_\phi$, $\omega(K^+K^-)_\phi$	$K^{*0}\bar{K}^{*0}$, $K^{*-}K^{*+}$, $\rho^+\rho^-$ $\rho^0\rho^0$	$\omega\omega$

Table 4. *D^0 Modes which are admixtures of CP eigenstates that can be separated by angular analysis.*

5. Extraction of Definite *CP* Quantities from Transversity

We now turn to the transversity analysis which we present for the case of spin one for particle A. For higher spins the method is similar; the separation of each $|\tau|$ can always be made from the polar angle distribution about the transversity axis. If particles C_1 or C_2 have spin a similar analysis is needed also for their decays.

Group	Example	$r_1(\theta)$	$r_0(\theta)$
(a) $1 \to 0+0$ or $1^- \to 3(0^-)$	$D^* \to \pi D$ $\omega \to \pi^+\pi^-\pi^0$	$\frac{3}{4}\sin^2\theta$	$\frac{3}{2}\cos^2\theta$
(b) $1^+ \to 3(0^-)$ or $1 \to \gamma + 0$ or $1 \to \frac{1}{2}+\frac{1}{2}$	$f_1 \to \eta\pi\pi$ $D^* \to \gamma D$ $\psi \to e^-e^+$	$\frac{3}{8}(1+\cos^2\theta)$	$\frac{3}{4}\sin^2\theta$

Table 5. *Angular structure as a function of the polar angle about the transverse axis.*

To analyse the decay of A we go to the A rest frame. In Table 5 we present the results. The first column defines two readily analysed groups of possible decays. Group (a) includes all decays of a spin 1 particle to two spinless particles and also decays of a vector particle to three pseudoscalar ones. Group (b) includes the decay of an axial-vector particle to three pseudoscalars, the decay of any spin 1 particle to a photon plus a spinless particle, and the decay of a spin one particle to a pair of negligible mass spin 1/2 particles via a vector or axial-vector coupling. The second column presents examples for decays of particle A. (We implicitly assume that this decay proceeds through parity conserving interactions.) Columns 3 and 4 present the angular distributions for each $|\tau|$, $r_\tau(\theta)$, normalized so that

$$\int_{-1}^{1} d\cos\theta\, r_\tau(\theta) = 1 \; . \tag{5.1}$$

Here, θ is an angle that describes the angular distribution of the decaying particle A, in the rest frame of A, relative to the transverse axis. When A decays into two particles the angle θ is the polar angle for one of the particles. When A decays to three spinless particles the angle θ is the polar angle of the normal to the plane containing the three decay products.[9] In all cases all other decay angles have been integrated out.

Using the angular distributions $r_\tau(\theta)$ one can then define the quantities

$$r_\pm(\theta) = r_0(\theta)(1 \pm \xi)/2 + r_1(\theta)(1 \mp \xi)/2 \tag{5.2}$$

where ξ is given in Table 1. The rate for a B^0_{phys} decaying to f_θ can be written as

$$\Gamma(\theta) \equiv \Gamma(B^0_{\text{phys}} \to f_\theta) = \Gamma_+(1+a)r_+(\theta) + \Gamma_-(1-a)r_-(\theta) \; , \tag{5.3}$$

where a and $\Gamma_\pm$ may be time–dependent or time–integrated quantities (see Eqs. (2.3)–(2.4)). Let us now define the weighted integrals,

$$M_0 \equiv \int_{-1}^{1} d\cos\theta\, \Gamma(\theta) = \Gamma_+(1+a) + \Gamma_-(1-a) \; , \tag{5.4}$$

and

$$M_2 \equiv \int_{-1}^{1} d\cos\theta\, P_2(\cos\theta)\, \Gamma(\theta) = \Gamma_+(1+a)\omega_+ + \Gamma_-(1-a)\omega_- \; . \tag{5.5}$$

where $\omega_\pm$ are defined by

$$\omega_\pm \equiv \int_{-1}^{1} d\cos\theta \, P_2(\cos\theta) \, r_\pm(\theta) \; . \tag{5.6}$$

Similarly the rate for a $\bar{B}^0_{\rm phys}$ to $\bar{f}_{\bar{\theta}}$ is

$$\bar{\Gamma}(\bar{\theta}) \equiv \Gamma(\bar{B}^0_{\rm phys} \to \bar{f}_{\bar{\theta}}) = \Gamma_+(1-a)r_+(\theta) + \Gamma_-(1+a)r_-(\theta) \; , \tag{5.7}$$

The state $|\bar{f}_{\bar{\theta}}\rangle$ means the state $CP\,|f_\theta\rangle$. Hence in Eq. (5.7) the quantity $\bar{\theta}$ is sometimes $\pi - \theta$ and sometimes θ depending on the particle content of the state. Since the angular dependence is such that $r_\tau(\pi - \theta) = r_\tau(\theta)$ this introduces no complication in the analysis. Thus we can extract

$$\bar{M}_0 \equiv \int_{-1}^{1} d\cos\theta \, \bar{\Gamma}(\theta) = \Gamma_+(1-a) + \Gamma_-(1+a) \; , \tag{5.8}$$

and

$$\bar{M}_2 \equiv \int_{-1}^{1} d\cos\theta \, P_2(\cos\theta) \, \bar{\Gamma}(\theta) = \Gamma_+(1-a)\omega_+ + \Gamma_-(1+a)\omega_- \; . \tag{5.9}$$

The moments $M_0, \bar{M}_0, M_2, \bar{M}_2$ derived from both the B^0 and $\bar{B}^0$ data samples, can be combined in many different ways to construct ratios which each give an undiluted measurement of the *CP*–violating asymmetry a. First construct the combinations

$$W_\pm = \Gamma_\pm(1 \pm a) = \mp \frac{\omega_\mp M_0 - M_2}{\omega_+ - \omega_-} \tag{5.10}$$

from the B data and the similar quantities $\bar{W}_\pm$ obtained from the $\bar{B}$ data. These then allow two determinations of the asymmetry a,

$$a_\pm = \frac{\pm(W_\pm - \bar{W}_\pm)}{W_\pm + \bar{W}_\pm} \; . \tag{5.11}$$

Note that neither of these results requires prior knowledge of the ratio of Γ_+ to Γ_-. Furthermore, each measures the intrinsic *CP* asymmetry of the underlying quark process without dilution. To obtain the most accurate value of asymmetry from this analysis one takes a linear combination

$$a = \alpha a_+ + (1 - \alpha) a_- \tag{5.12}$$

with α chosen to minimize the error on the result. The optimal choice of α does depend on the actual values of the Γ's. We will return to a discussion of the best choice of α in Section 7.

With a limited amount of data one could alternatively begin by dividing the angular distribution into two angular bins, commonly called polar and equatorial, with a cut at some appropriate angle. Let us cut at $\theta = \pi/3$ where $\cos\theta = 1/2$ and define the equatorial and polar components E and P by the relations

$$E = 2 \int_0^{1/2} d\cos\theta \; \Gamma(\theta) = \Gamma_+(1+a)e_+ + \Gamma_-(1-a)e_- \tag{5.13}$$

$$P = 2 \int_{1/2}^{1} d\cos\theta \; \Gamma(\theta) = \Gamma_+(1+a)p_+ + \Gamma_-(1-a)p_- \; , \tag{5.14}$$

where the numbers $e_\pm$ and $p_\pm$ are defined by

$$e_\pm \equiv 2 \int_0^{1/2} d\cos\theta \; r_\pm(\theta) \tag{5.15}$$

$$p_\pm \equiv 2 \int_{1/2}^{1} d\cos\theta \, r_\pm(\theta) \; . \tag{5.16}$$

The quantities $W_\pm$ and $\bar{W}_\pm$ can now be extracted using E and P and the corresponding quantities $\bar{E}$ and $\bar{P}$ constructed from the $\bar{B}^0$ data sample, just as was done above using M_0, M_2, $\bar{M}_0$ and $\bar{M}_2$. The asymmetries $a_\pm$ can then be determined as before. The only differences between the two procedures will be in the errors on the estimates of a, which will be reduced by the moment treatment. However, the simpler binning procedure could be used for a preliminary look. If a *CP*-violating effect is present, it should show up as a statistically significant nonvanishing value for a even in this simpler analysis. Once such an asymmetry is found, the treatment of the data can be refined.

In either of these analyses the result for the asymmetry a does not depend on Γ_+ or Γ_-, or, in principle, on the choice of α. However the error on a can be minimized by choosing α in a way that depends on Γ_+ and Γ_-. The values of Γ_+ and Γ_- can be determined from examination of channels related to the channel under study by isospin symmetry. In many cases these channels will provide much more data than the channel used for the *CP* analysis and so the errors on Γ_+/Γ_- will have little effect on the error on the asymmetry.

6. Full Angular Analysis

For classes (2) and (3) a full angular analysis will usually prove superior to the simple moment treatments described above, since the error on the asymmetry measurement from a given set of data can be reduced by more fully exploiting the known angular structure to extract several measurements of the asymmetry a with different correlations among their errors. Such analysis also allows study of modes for which the transversity treatment is not applicable, for example modes where neither particle A nor C decay to two spin zero particles.

Appendix B presents the general helicity formalism and develops a method based on using the Y_{lm} functions to perform angular projections. The treatment of the case of $AC = \psi K^{*0}$ is given as an example. The results for $D^*\bar{D}^*$ are also tabulated. The angular analysis of the decay of each particle is most simply carried out in the rest frame of that particle. One needs to specify a choice of coordinates for each decay of a spinning particle. Once this is done, the angular projections can be used to separate out the combinations of helicity amplitudes that have a definite *CP* and hence to measure asymmetries.[2] As in the case of the transversity analysis, this can be achieved by combining B^0 and $\bar{B}^0$ data, without any *a priori* knowledge of the various amplitudes involved. One can obtain an asymmetry measurement for each possible combination of helicities. These results can then be combined for an improved measurement as discussed in Section 7.

This analysis can be applied for any of the modes discussed previously, provided the system $(C_1\, C_2)$ has a well defined total angular momentum. In addition, modes where the transversity analysis cannot readily be used can also be analysed; some such modes are listed in Tables 2–4. For example consider the case

$$B^0 \to \Delta\bar{\Delta}$$

where each Δ subsequently decays to a proton and a pion. Although the proton spin cannot be measured it is still possible to use the angular projections of these decays to extract quantities of definite *CP*. This analysis is also presented in Appendix B. We find for this case that two definite *CP* quantities $Re\mathcal{G}_{3/2+}\mathcal{G}^*_{1/2+}$ and $Re\mathcal{G}_{3/2-}\mathcal{G}^*_{1/2-}$ where $\mathcal{G}_{\lambda\pm} = \frac{A_\lambda \pm A_{-\lambda}}{\sqrt{2}}$ can be isolated using angular projections. Each of these provides a possible measurement of the intrinsic *CP* asymmetry. This result applies for any pair of spin 3/2 resonances, both of which decay to $p\pi$ (or $\bar{p}\pi$). For the case of two spin 1/2 resonances which both decay to $p\pi$ (or $\bar{p}\pi$) the averaging over the proton spins removes all possibility of separating the different *CP* contributions by angular analysis as only a uniform angular distribution survives.

7. Minimizing the Error on the Measured Asymmetry

The methods described above each give more than one measurement of the asymmetry. With the transversity analysis we had $a_\pm$ or in the more general case one measurement for each set of $|\tau_i|$. Consider, for example, integer-spin particles A and C, the full angular analysis gives effectively $(s+1)^2$ measurements, one from the square of each of the $2s+1$ definite *CP* combinations of helicity amplitudes and one from each interference term between any two such amplitudes with the same CP. Here $s = \min(s_a, s_c)$. Interference terms between opposite CP contributions depend only quadratically on the asymmetry and hence do not provide as sensitive a measurement. Furthermore, as can be seen from the example of

$$\begin{aligned}\mathrm{Im}\,\mathcal{G}_{1+}\mathcal{G}_{1-}^* = e^{-\Gamma t}\Big[&\mathrm{Re}(G_{1+}G_{1-}^*)\eta\,\mathrm{Re}\lambda_{KM}\,\sin\,\Delta mt \\ &+\mathrm{Im}(G_{1+}G_{1-}^*)\,\cos\,\Delta mt\Big]\end{aligned} \tag{7.1}$$

derived in Appendix B the separation of the weak phase dependence from the strong phases is not as clean in this case. For the pure CP terms such as

$$\mathcal{G}_{1+}\mathcal{G}_{1+}^* = |G_{1+}|^2[1-\eta\,\mathrm{Im}(\lambda_{KM})\sin(\Delta mt)]\;e^{-\Gamma t}$$

the time and asymmetry dependence is much simpler. From such a term one can readily form the combination

$$a_{++} = \frac{(\mathcal{G}_{1+}\mathcal{G}_{1+}^* - \bar{\mathcal{G}}_{1+}\bar{\mathcal{G}}_{1+}^*)}{(\mathcal{G}_{1+}\mathcal{G}_{1+}^* + \bar{\mathcal{G}}_{1+}\bar{\mathcal{G}}_{1+}^*)} = -\eta\mathrm{Im}(\lambda_{KM})\sin(\Delta mt) \tag{7.2}$$

In either analysis the errors on the various measurements of the asymmetry are correlated, and these correlations must be treated properly in determining the error on any value of the asymmetry extracted by combining them. All this is standard statistical analysis, we review it briefly here for completeness.

Consider first the case where we have only the two measurements $a_{\pm}$ extracted from the single moment transversity analysis. We choose*

$$a = \alpha a_+ + (1-\alpha)a_-\ . \tag{7.3}$$

Minimizing the χ^2 with respect to α yields, for small asym-

⋆ In Eq. (7.3) we have restricted the possible value of α to lie between zero and one. This restriction actually excludes the "best" value when a single CP contribution dominates. Since the errors on $a_{\pm}$ are anticorrelated the fit actually prefers to overshoot and choose α less than zero or greater than one in these cases. However the value thus chosen is extremely sensitive to a precise knowledge of the ratio of Γ_+ to Γ_-, and hence unreliable.

metry,

$$\alpha = \frac{\sigma_+^2 + \rho\sigma_+\sigma_-}{\sigma_-^2 + \sigma_+^2 + \rho\sigma_+\sigma_-}, \tag{7.4}$$

where $\sigma_\pm$ are the standard errors on $W_\pm$ and ρ measures the correlation of these errors. The result of this treatment is shown as the solid curve in Figure 3 as a function of

$$\epsilon = \frac{\Gamma_+ - \Gamma_-}{\Gamma_+ + \Gamma_-} \tag{7.5}$$

for the case of two spin one particles, one of which decays to two spin zero particles while the other decays to an e^+e^- pair e.g.ψK^{*0}. We plot the ratio of the expected error from this analysis to that obtainable with an equal number of decays to a pure *CP* state.[16] For comparison, we also show the errors obtained for a fixed value $\alpha = \frac{1}{2}$, this gives the upper curve on Figure 3. One sees that, in the worst case, where Γ_+ and Γ_- are equal, this analysis requires approximately nine times more data than a pure *CP* channel to achieve equal accuracy for the asymmetry measurement.

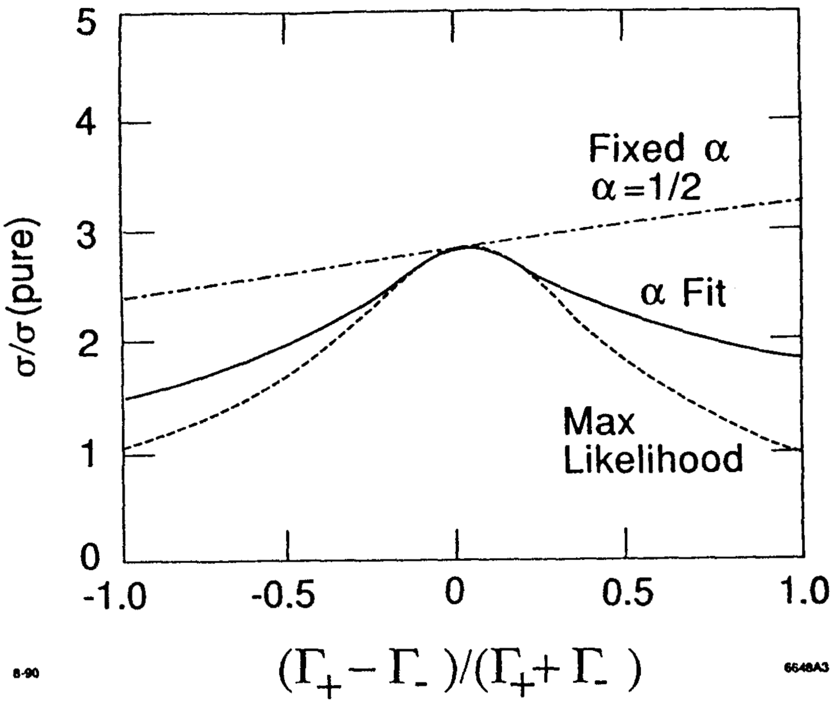


Figure 3. *The ratio of the expected error on the CP-violating asymmetry extracted using transversity from the mixed CP state ψK^* to that obtainable with an equal number of decays to a pure CP state. The curve labelled "$\alpha = 1/2$" is based on equal weightings of $a_\pm$, while that labelled "α fit" uses the optimal choice for each ϵ. The lowest curve is obtained from a maximum likelihood fit to the asymmetry, assuming Γ_+/Γ_- is known.*

This situation can be improved by making a maximum-likelihood fit for the asymmetry using expression for the θ–angular distribution given by Eq. (5.3). This analysis requires further parameters, namely the quantities $\Gamma_\pm$ which we assume can be extracted from isospin–related channels. The result of this treatment for the error on the asymmetry is shown as the dashed curve in Figure 3.

Another way to improve the result is by making a more complete angular analysis. Again we have two options, an analysis based on moments that does not require knowledge of the relative strengths and phases of the various amplitudes, and a maximum-likelihood fit to the full set of parameters. Where sufficient information is available, the latter method is superior. Figure 4 shows the result for the errors on the asymmetry from using a maximum-likelihood fit to the angular dependence of the data where it is assumed that the quantities $G_{1\pm}$ and G_0 , defined in Appendix B, are all known

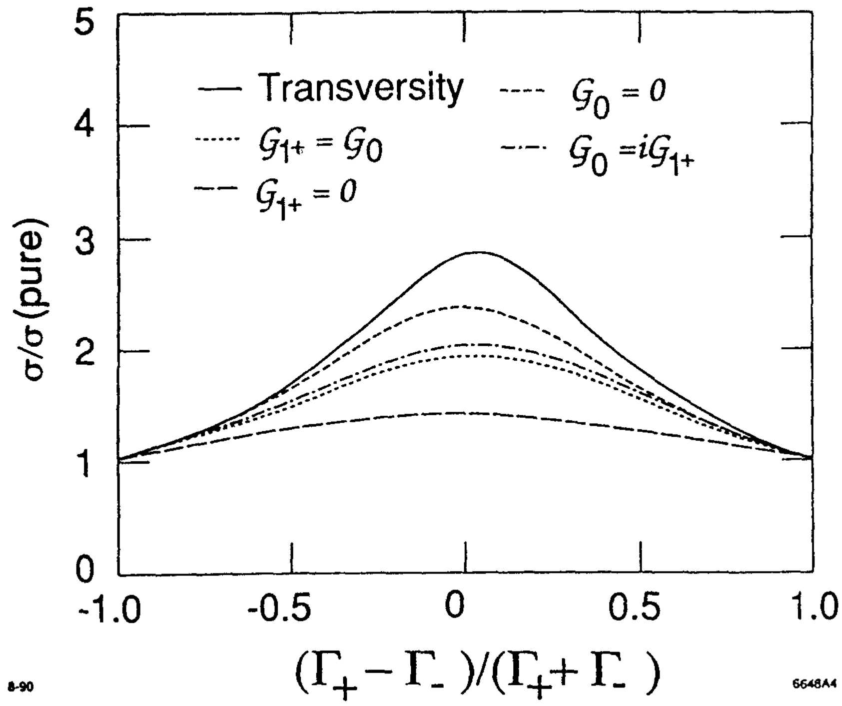


Figure 4. *The ratio of expected error on asymmetries obtained using maximum liklihood fits for a mixed CP (ψK^*) channel to that for a pure CP channel with equal number of decays. The top curve is the transversity based result, shown also on Figure 3. The remaining curves represent different assumed values for $\mathcal{G}_{1+}/\mathcal{G}_0$, with $\mathcal{G}_{1+}$ and $\mathcal{G}_{1-}$ taken to be relatively real.*

from measurements on isospin related channels. For simplicity, we assumed $a \ll 1$ in making this error analysis. For comparison, we plot the result against the same combination of variables as were available in the transversity analysis. The various cases shown are chosen to indicate the range of possibilities. It can be seen that even in the worst case that we studied this type of analysis provides a more accurate value for the asymmetry than the best transversity treatment. Figure 4 also shows that in the fortunate situation where a single CP contribution dominates either treatment gives accuracy comparable to that obtained for a pure CP channel.

We have also carried out a study of the sensitivity of the asymmetry measurement in a maximum-likelihood fit procedure to errors in the estimated values and phases of the various amplitudes. This analysis is summarized in Appendix C. The results are encouraging, the asymmetry measurement errors will most likely be dominated by the statistics of the channel for which the asymmetry measurement is made and is relatively insensitive to small errors in the amplitude values or phases. However, if these quantities are poorly determined, one can fall back to the moment analyses to extract asymmetry measurements that do not depend on them.

To summarize the situation, we again emphasize that the value of the transversity analysis is greatest in the channels described by Class 1 of Table 1, namely three CP-self-conjugate particles, where it allows combination of resonant and non-resonant production of the particles C_1 and C_2. It also has the feature of being particularly insensitive to the non-CP-violating asymmetries of the amplitudes, that is to asymmetries between B and $\bar{B}$ data that arise from interference between the CP-odd and CP-even contributions. However, whenever there is not a single dominant CP contribution the most accurate results for asymmetries will come from the use of a maximum-likelihood fit to the parameters that define the angular distributions, rather than any of the analyses which depend on projecting out specific moments to identify definite CP contributions. Whenever possible, such a treatment will include isospin-related channels in the fitting procedure. Since the isospin-related channels typically have higher rates than the CP-eigenstate channels the additional parameters needed for this type of analysis will be well-measured for many modes by the time one has sufficient data to measure the asymmetry, so this method will be the one used for most channels.

8. Summary and Conclusions

We have shown that there are many channels for which one can study CP violations in B^0 decays if one uses angular analysis to separate the different CP contributions which arise from different helicity terms. Some of these modes will compete in accuracy with the modes with unique CP which have already been much discussed. In general, to carry out the angular analysis accurately somewhat more data is needed for these modes than for the unique CP modes; in the worst case that we have analysed this requires approximately ten times as much data for an equally accurate measurement of asymmetry, the degradation will possibly be even greater for higher spin channels.

We have presented several different approaches to the angular analysis, each of which has merit in different situations. To summarize:

Transversity analysis will be most useful for modes with three self-conjugate particles.

Transversity analysis is most useful in the case of decays to three self-conjugate particles, Class 1 of Table 1, where it allows the combination of resonant and non-resonant production of the particles C_1 and C_2. If the relative strengths of the two CP contributions ($\Gamma_\pm$) are not known, then a moment analysis of the type described in Section 5 should be used. Whenever the values of $\Gamma_\pm$ can be determined using data from isospin related channels then a maximum-likelihood fit to the transversity polar angle distributions will provide a more accurate result.

In other modes the most acurate results will come from maximum-likelihood fits to angular distributions and to isospin-related channels.

For all other modes, including those listed as Class 2 and 3 in Table 1 which could be analysed using the transversity method, the full angular distribution analysis will prove superior. Again there are two choices, a moment analysis of the type described in Appendix B or a maximum-likelihood fit to the full angular distributions. Wherever sufficient information on the various helicity amplitudes can be extracted from data on isospin-related channels, the latter method will again give more accurate results. Clearly, what this means is that in such cases one should make a global fit of all parameters, the helicity amplitudes and the asymmetry, to the data from all related channels, to obtain the most accurate asymmetry measurement.

This discussion makes no distinction between a time-integrated experiment or one that measures time dependencies of the B and $\bar{B}$ decays. In the latter case, the angular structure and the time dependence factorize in a simple way, as demonstrated in Appendix B. In a time-dependent experiment one simply performs the angular analysis for each time bin separately, or in a maximum-likelihood fitting procedure one includes the predicted time dependence in the fitting prescription, and treats the data as a function of time as well as angles.

However, the angular analysis so enriches the sample of modes to study that we expect it will play an important role in the extraction of the CP-violating physics at a B factory. Among the many channels listed in Tables 2 and 3 there well may be some that provide as accurate or more accurate asymmetry measurements as the unique CP modes. Since we do not yet have much information on branching fractions to the various modes it is too early to be certain which of the many modes will provide the best measurements. Hence, we have simply presented summary tables of some of the modes which, according to the Standard Model, will measure the various angles of the Unitarity Triangle. We have not found any one mode for which the currently measured branching fractions suggest it would be markedly superior to the unique CP modes, but several may be comparable, especially in the fortunate circumstance that a single CP channel dominates the process. Our knowledge of these branching fractions will certainly be much better by the time any B factory capable of measuring CP asymmetries is built, so at that time it will be obvious which of these modes is most readily used, and which method of analysis to apply to it.

APPENDIX A

HELICITY AND TRANSVERSITY

This Appendix gives the general proof that transversity projections can be used to select definite CP quantities. When a spin zero particle B, at rest, decays into two particles (A and C),they must have equal helicities (λ). Now we consider the case where the particle C decays to two integer spin particles C_1 and C_2, which have spins s_1 and s_2 and helicities λ_1 and λ_2 . In the rest frame of particle C we can write the state formed by the decay of B as

$$|f(\theta)\rangle = \sum_{\lambda,\lambda_1,\lambda_2} \sqrt{\frac{2s_c+1}{4\pi}} A_{\lambda,\lambda_1,\lambda_2} D^{s_c^*}_{\lambda,\lambda_1-\lambda_2}(R_C)\,|\lambda,\lambda_1,\lambda_2;\theta\rangle\,, \tag{A.1}$$

where we define

$$|\lambda,\lambda_1,\lambda_2;\theta\rangle = |A(\lambda;0,0),C_1(\lambda_1;\theta,0)C_2(\lambda_2;\pi-\theta,\pi)\rangle\,. \tag{A.2}$$

We use Jackson conventions to define our angles and axes.[8] In addition we have chosen to define the C–decay angles so that $\phi_{C_1} = 0$, thus $R_C = (0,\theta,0)$. The choice $\phi_{C_1} = 0$ is made event by event without any loss of generality. It is a convenient choice for the transversity discussion since it identifies the y–axis of these coordinates with the direction transverse to the plane. In Appendix B we will use a different convention for ϕ in the full angular analysis. These choices are, of course, merely a matter of convenience for each analysis and have no physical content.

It is important to note that for three self–conjugate particles we can here avoid the assumption of a specific particle C and simply sum over all possible angular momenta for the system C_1C_2 in its rest frame, in which case λ is the projection of this angular momentum along the direction opposite to particle A. Then Eq.(A .1) generalizes to

$$|f(\theta)\rangle = \sum_{\lambda,\lambda_1,\lambda_2} \sum_{J} \sqrt{\frac{2J+1}{4\pi}} A^{J}_{\lambda,\lambda_1,\lambda_2} D^{J^*}_{\lambda,\lambda_1-\lambda_2}(R_C)\,|\lambda,\lambda_1,\lambda_2;\theta\rangle\,. \tag{A.3}$$

We now wish to use the decays of A,C_1 and C_2 to analyse their transversity. To do this for each particle, we go to its rest frame. The transversity for each particle is defined as the spin component along the y—axis in the particle's rest frame. With the choice $\phi_{C_1} = 0$ all three y–axes are parallel. We must, however, choose the same direction for the definition of transversity for all three particles, so that we can define the total transversity as the sum of the three projections. We will fix this as the direction of the positive y–axis for the decay of particle C. Then we can relate transversity states to helicity states for a particle of integer spin s by

$$\begin{aligned} |s,\tau\rangle_{y_c} &= \sum_{\lambda} D^s_{\lambda,\tau}(\kappa\pi/2,\pi/2,0)\,|s,\lambda\rangle_z \\ &= \sum_{\lambda\geq 0}[1-\tfrac{1}{2}\delta_{\lambda,0}]D^s_{\lambda,\tau}(\kappa\pi/2,\pi/2,0) \qquad \text{(A.4)} \\ &\quad \times \left\{|s,\lambda\rangle_z + (-1)^{s-\lambda-\tau}\,|s,-\lambda\rangle_z\right\}\,. \end{aligned}$$

The rotations are defined with respect to the axes just described, and $\kappa = 1$ for particle C_1 and $\kappa = -1$ for particles A and C_2 in order to achieve the required matching of positive transversity direction. The phase factor on the negative helicity term arises from redefining the D–function for $-\lambda$ in terms of that for λ.

Now let us first consider decays in which the three particles A, C_1 and C_2 all are neutral bosons self–conjugate under CP (Class 1 in Table 1). Then

$$CP\,|\lambda,\lambda_1,\lambda_2;\theta\rangle = \xi(-1)^{s_a+s_1+s_2-\lambda-\lambda_1-\lambda_2}\,|-\lambda,-\lambda_1,-\lambda_2;\theta\rangle\,, \qquad \text{(A.5)}$$

where ξ is the product of the intrinsic CP–parities of the three particles. Notice that the angle θ is unchanged under CP, since it is defined to be the angle between particle C_1 and the direction opposite particle A in the C rest frame, and hence is unaltered by the reversal of all momenta. Now we use the property of the D–function

$$D^s_{\lambda,\mu}(0,\theta,0) = d^s_{\lambda,\mu}(\theta) = (-1)^{\lambda-\mu}d^s_{-\lambda,-\mu}(\theta) \qquad \text{(A.6)}$$

to rewrite Eq. (A.3) as

$$|f(\theta)\rangle = \frac{1}{\sqrt{2}} \sum_{\lambda,\lambda_1,\lambda_2 \geq 0} [1-\tfrac{1}{2}\delta_{\lambda,0}][1-\tfrac{1}{2}\delta_{\lambda_1,0}][1-\tfrac{1}{2}\delta_{\lambda_2,0}]$$

$$\left\{ \left|\mathcal{S}^{+}_{\lambda,\lambda_1,\lambda_2}\right\rangle + \left|\mathcal{S}^{+}_{\lambda,\lambda_1,-\lambda_2}\right\rangle + \left|\mathcal{S}^{+}_{\lambda,-\lambda_1,\lambda_2}\right\rangle + \left|\mathcal{S}^{+}_{\lambda,-\lambda_1,-\lambda_2}\right\rangle \right.$$

$$\left. + \left|\mathcal{S}^{-}_{\lambda,\lambda_1,\lambda_2}\right\rangle + \left|\mathcal{S}^{-}_{\lambda,\lambda_1,-\lambda_2}\right\rangle + \left|\mathcal{S}^{-}_{\lambda,-\lambda_1,\lambda_2}\right\rangle + \left|\mathcal{S}^{-}_{\lambda,-\lambda_1,-\lambda_2}\right\rangle \right\}, \tag{A.7}$$

where

$$\left|\mathcal{S}^{\pm}_{\lambda,\lambda_1,\lambda_2}\right\rangle = \sum_J \sqrt{\frac{2J+1}{4\pi}} d^J_{\lambda,\lambda_1-\lambda_2}(\theta) \times$$

$$\mathcal{G}^{J\pm}_{\lambda,\lambda_1,\lambda_2} \{ |\lambda,\lambda_1,\lambda_2;\theta\rangle \pm (-1)^{\lambda+\lambda_1+\lambda_2} |-\lambda,-\lambda_1,-\lambda_2;\theta\rangle \} \tag{A.8}$$

and we have introduced the amplitudes

$$\mathcal{G}^{J\pm}_{\lambda,\lambda_1,\lambda_2} = \frac{A^J_{\lambda,\lambda_1,\lambda_2} \pm A^J_{-\lambda,-\lambda_1,-\lambda_2}}{\sqrt{2}} \tag{A.9}$$

which correspond to definite *CP* contributions. Under *CP*

$$CP \left|\mathcal{S}^{\pm}_{\lambda,\lambda_1,\lambda_2}\right\rangle = \pm\xi(-1)^{s_a+s_1+s_2} \left|\mathcal{S}^{\pm}_{\lambda,\lambda_1,\lambda_2}\right\rangle, \tag{A.10}$$

where we have used Eq. (A .5).

Now let us project out the contribution obtained by requiring the transversities τ_a, τ_1, τ_2 for the particles A, C_1, C_2. We can write the result in the form

$$\begin{aligned} A(\tau_a,\tau_1,\tau_2) &= \langle \tau_a,\tau_1,\tau_2;\theta|f(\theta)\rangle \\ &= (1+(-1)^{S+\tau})\mathcal{G}^+ + (1-(-1)^{S+\tau})\mathcal{G}^-, \end{aligned} \tag{A.11}$$

where $S = s_a + s_1 + s_2$, $\tau = \tau_a + \tau_1 + \tau_2$, and

$$
\begin{aligned}
\mathcal{G}^{\pm} = \sum_{\lambda,\lambda_1,\lambda_2 \geq 0} \rho_{\lambda,\lambda_1,\lambda_2} \sum_j \sqrt{\frac{2j+1}{4\pi}} \\
\Big\{ d^j_{\lambda,\lambda_1-\lambda_2}(\theta)\, \mathcal{G}^{j\pm}_{\lambda,\lambda_1,\lambda_2} \\
+ (-1)^{s_2-\lambda_2-\tau_2} d^j_{\lambda,\lambda_1+\lambda_2}(\theta)\, \mathcal{G}^{j\pm}_{\lambda,\lambda_1,-\lambda_2} \\
+ (-1)^{s_1-\lambda_1-\tau_1} d^j_{\lambda,-\lambda_1-\lambda_2}(\theta)\, \mathcal{G}^{j\pm}_{\lambda,-\lambda_1,\lambda_2} \\
+ (-1)^{s_a+\lambda-\tau_a} d^j_{-\lambda,\lambda_1-\lambda_2}(\theta)\, \mathcal{G}^{j\pm}_{-\lambda,\lambda_1,\lambda_2} \Big\} .
\end{aligned}
\tag{A.12}
$$

Here ρ is

$$
\begin{aligned}
\rho_{\lambda,\lambda_1,\lambda_2} = \frac{1}{\sqrt{2}} \{[1 - \tfrac{1}{2}\delta_{\lambda,0}][1 - \tfrac{1}{2}\delta_{\lambda_1,0}][1 - \tfrac{1}{2}\delta_{\lambda_2,0}]\}^2 [1 + \delta_{0\lambda}\delta_{0\lambda_1}\delta_{0\lambda_2}] \\
(-1)^{\lambda+\lambda_2}(i)^{\lambda+\lambda_1+\lambda_2} d^{s_a}_{\lambda,\tau_a}(\pi/2) d^{s_1}_{\lambda_1,\tau_1}(\pi/2) d^{s_2}_{\lambda_2,\tau_2}(\pi/2) .
\end{aligned}
\tag{A.13}
$$

In Eq. (A .11) we have used the fact that

$$d^s_{\lambda,\tau}(\pi/2) = (-1)^{s-\tau} d^s_{-\lambda,\tau}(\pi/2). \tag{A.14}$$

Equation (A .11) shows that each τ projects out a definite *CP* contribution. Combining Eq. (A .12) and Eq. (A .10), we see that the non-vanishing contributions all have *CP* parity $\xi(-1)^\tau$.

Examination of Eq. (A .11) clearly shows that only the absolute values of τ, τ_1, τ_2 need be definite, since they are all integers. This then indicates that the simplest experimental procedure to separate definite *CP* quantities will be to integrate over the azimuthal dependence of the decays with respect to the transverse axis and to project for definite $|\tau_i|$ using the polar angle distribution about this axis. We thus need only take non-trivial moments of a single angular dependence for each particle to reconstruct the magnitude of its transversity. We then can combine B and $\bar{B}$ data to obtain a measurement of the *CP* asymmetry for each set of

$|\tau|, |\tau_1|, |\tau_2|$, as discussed for the example in Section 5. These results can then be combined to yield an improved estimate as discussed in Section 7.

If C_1 and C_2 are not self–conjugate particles, as in classes 2 and 3 of Table 1, then Eq. (A .5) does not apply since CP interchanges particles. However, if we require both C_1 and C_2 to be spin zero particles, then the transversity of particle A will again allow separation of CP–odd and CP–even contributions. The proof can readily be seen from the case discussed above, with the sums over J reduced to the single term $J = s_c$ and with $s_1 = \lambda_1 = 0$ and $s_2 = \lambda_2 = 0$.

APPENDIX B

FULL ANGULAR ANALYSIS AND TIME DEPENDENCE

In this Appendix we will present a method for using the full angular distribution to define a set of moments from which all measurable combinations of helicity amplitudes can be extracted. The method is a standard helicity analysis which we present here for completeness. We analyse here the B^0 decays into two spin-one particles, one of which decays to two spin zero particles and the other to an e^+e^- pair, for example the mode ψK^* where $\psi \to e^+e^-$ and $K^* \to K_S^0\pi^0$. A similar analysis for B^0 decays into two spin 1 particles which each subsequently decay to two spin zero particles is also presented. An analysis for the case of two spin 3/2 particles is also briefly discussed. We further present here the explicit structure of the time dependence of the various quantities that can be measured and discuss the extraction of time–dependent CP asymmetries. For more details on this analysis, see W. Toki.[18]

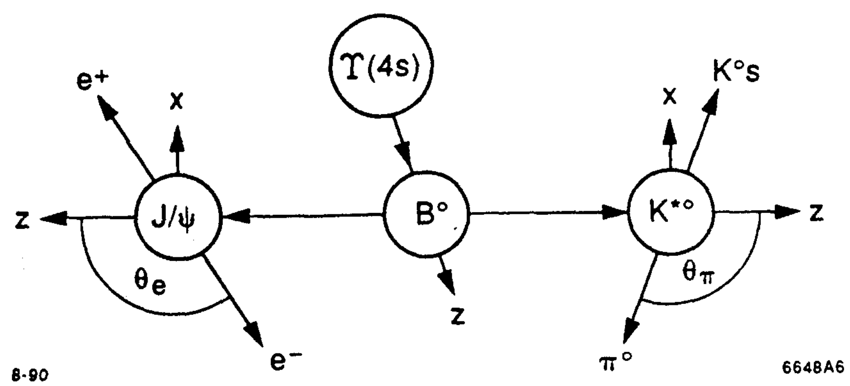


Figure 5. *Schematic drawing of the kinematics of B^0 production and decay showing definitions of the various axes and angles. Each decay is considered in the rest–frame of the parent particle.*

The first step in this analysis requires the definition of some conventions. We use here the conventions of Jackson for the definition of the rotation D–functions. The decay angles for the process $B \to \psi K^{*0}$, $\psi \to e^+e^-$, $K^{*0} \to K_S^0\pi^0$ are shown in Fig 5. We assume the B, $\Upsilon(4s)$, J/ψ and the K^* are in the plane of the paper. The Z axis in the respective helicity frames are opposite to the parent particle. The Y axes are in the direction of the cross product of the Z axis of the parent and the Z axis of the helicity frame. This causes the Y axis of the K^* to be out of the paper and the Y axis of the ψ to be into the paper. Hence the X axes are both pointing upward. This will cause the ϕ angle of the e^- and the π^0 to be going in opposite directions such that their sum will yield the relative angle between the two decay planes. In this drawing neither the e^+, e^-, π^0 nor the K^0 need to lie in the plane of the paper.

The matrix element of the decay of $B \to \psi K^{*0}$, $K^{*0} \to K_S^0\pi^0$ can be written using the helicity formalism as

$$|M|^2 = \sum_{\alpha=\pm 1}\left|\sum_{\lambda=0,\pm 1}\sqrt{\frac{2J_\psi+1}{4\pi}}\sqrt{\frac{2J_K+1}{4\pi}} \times A_\lambda D^{J_\psi *}_{\lambda,\alpha}(R_\psi)D^{J_K *}_{\lambda,0}(R_K)\right|^2 . \tag{B.1}$$

The amplitudes A_λ in (B .1) contain implicit time dependence which we will discuss later. The important point is that the time dependence and the angular dependence factorize in this way, so one can perform the angular analysis for each time bin and thus extract time–dependent asymmetries. For the Jackson convention $R = (\phi, \theta, 0)$, which differs from the Jacob–Wick[8] convention in which $R = (\phi, \theta, -\phi)$. Expanding Eq. (B .1) gives

$$|M|^2 = \frac{2J_\psi+1}{4\pi}\,\frac{2J_K+1}{4\pi}\sum_{\alpha=\pm 1}\;\sum_{\lambda,\lambda'=0,\pm 1} A_\lambda A^*_{\lambda'} D^{J_\psi *}_{\lambda,\alpha}(R_\psi)D^{J_\psi}_{\lambda',\alpha}(R_\psi)D^{J_K *}_{\lambda,0}(R_K)D^{J_K}_{\lambda',0}(R_K) \tag{B.2}$$

Changing the charge conjugate to real

$$D^{J*}_{M'M}(R) = (-1)^{M'-M}D^{J}_{-M'-M}(R)$$

and inserting the double D summations

$$D^{J_1}_{M_1'M_1}(R)\,D^{J_2}_{M_2'M_2}(R) =$$

$$\sum_{J_3=|J_1-J_2|}^{J_1+J_2} \langle J_1M_1, J_2M_2 \,|\, J_3M_3\rangle \, \langle J_1M_1', J_2M_2' \,|\, J_3M_3'\rangle \, D^{J_3}_{M_3'M_3}(R) \tag{B.3}$$

gives

$$\begin{aligned}|M|^2 = {} & \frac{2J_\psi+1}{4\pi}\,\frac{2J_K+1}{4\pi} \\ & \sum_{\lambda,\lambda'=0,\pm1} A_\lambda A^*_{\lambda'} \sum_{\alpha=\pm1}(-1)^\alpha \sum_{J_L,J_R=0,1,2} \\ & \langle 1\alpha, 1-\alpha \,|\, J_L0\rangle \, \langle 1\lambda, 1-\lambda' \,|\, J_LM_L'\rangle \, D^{J_L}_{-M_L',0}(R_\psi) \\ & \times \langle 10, 10 \,|\, J_R0\rangle \, \langle 1\lambda, 1-\lambda' \,|\, J_RM_R'\rangle \, D^{J_R}_{-M_R',0}(R_K)\end{aligned} \tag{B.4}$$

After a little rearranging, we have

$$\begin{aligned}|M|^2 = {} & -2\left(\frac{3}{4\pi}\right)^2 \sum_{\lambda,\lambda'=0,\pm1} A_\lambda A^*_{\lambda'} \sum_{J_L,J_R=0,2} \\ & \langle 11, 1-1 \,|\, J_L0\rangle \, \langle 1\lambda, 1-\lambda' \,|\, J_L\lambda-\lambda'\rangle \, D^{J_L}_{\lambda'-\lambda,0}(R_\psi) \\ & \times \langle 10, 10 \,|\, J_R0\rangle \,|\, \langle 1\lambda, 1-\lambda' \,|\, J_R\lambda-\lambda'\rangle \, D^{J_R}_{\lambda'-\lambda,0}(R_K)\end{aligned} \tag{B.5}$$

The $J_L = 1$ terms vanish because of the sum over α on the Clebsch-Gordan coefficient $\langle 1\alpha 1-\alpha \,|\, J_L0\rangle$ and the $J_R = 1$ terms vanish because of the coefficient $\langle 1010 \,|\, J_R0\rangle$. We now simplify with the following relations

$$D^L_{M,0}(\phi,\theta,\chi) = \sqrt{\frac{4\pi}{2L+1}}\; Y^*_{LM}(\theta,\phi)$$

where $Y^*_{LM} = (-1)^M Y_{L-M}$ and $\int Y^*_{\ell m}(\Omega) Y_{\ell' m'}(\Omega)\, d\Omega =$

$\delta_{\ell\ell'}\delta_{mm'}$, to obtain

$$
\begin{aligned}
|M|^2 = -2\left(\frac{3}{4\pi}\right)^2 & \sum_{\lambda,\lambda'=0,\pm1} \\
& \times A_\lambda A^*_{\lambda'} \sum_{J_L,J_R=0,2} \frac{4\pi}{\sqrt{2J_L+1}\sqrt{2J_R+1}} \\
& \times \langle 11,1-1\,|\,J_L0\rangle \; \left\langle 1\lambda,1-\lambda'\,|\,J_L\lambda-\lambda'\right\rangle \, Y^*_{J_L,\lambda'-\lambda}(\Omega_\psi) \\
& \times \langle 10,10\,|\,J_R0\rangle \; \left\langle 1\lambda,1-\lambda'\,|\,J_R\lambda-\lambda'\right\rangle \, Y^*_{J_R,\lambda'-\lambda}(\Omega_K)\,.
\end{aligned}
\tag{B.6}
$$

Let us now define the moments

$$T_{J_LJ_RM} = \int\!\!\int |M|^2 Y_{J_L,M}(\Omega_\psi)\,Y_{J_R,M}(\Omega_K)\,d\Omega_K\,d\Omega_\psi \tag{B.7}$$

and thus

$$|M|^2 = \sum_{J_L=0,2}\;\sum_{J_R=0,2}\;\sum_{M=0,\pm1,\pm2} T_{J_LJ_RM}\,Y^*_{J_L,M}(\Omega_\psi)\,Y^*_{J_R,M}(\Omega_K)\,. \tag{B.8}$$

and $T_{J_LJ_R-M} = T^*_{J_LJ_RM}$. The relation between the helicity amplitudes and the moments is

$$
\begin{aligned}
T_{J_LJ_R-M} = \frac{-9}{2\pi}\,\frac{1}{\sqrt{2J_L+1}}\,\frac{1}{\sqrt{2J_R+1}} & \sum_{\lambda,\lambda'=0,\pm1} \\
& \langle 11,1-1\,|\,J_L0\rangle \; \left\langle 1\lambda,1-\lambda'\,|\,J_LM\right\rangle \\
& \times \langle 10,10\,|\,J_R0\rangle \; \left\langle 1\lambda,1-\lambda'\,|\,J_RM\right\rangle \, A_\lambda A^*_{\lambda'}\,.
\end{aligned}
\tag{B.9}
$$

J_L J_R M	$4\pi T_{J_L J_R M}$	$4\pi T_{J_L J_R M}$
0 0 0	$2(A_1A_1^* + A_{-1}A_{-1}^* + A_0A_0^*)$	$2(\mathcal{G}_{1+}\mathcal{G}_{1+}^* + \mathcal{G}_{1-}\mathcal{G}_{1-}^* + \mathcal{G}_0\mathcal{G}_0^*)$
0 2 0	$\frac{2}{\sqrt{5}}(2A_0A_0^* - A_1A_1^* - A_{-1}A_{-1}^*)$	$\frac{2}{\sqrt{5}}(2\mathcal{G}_0\mathcal{G}_0^* - \mathcal{G}_{1+}\mathcal{G}_{1+}^* - \mathcal{G}_{1-}\mathcal{G}_{1-}^*)$
2 0 0	$\frac{1}{\sqrt{5}}(A_1A_1^* + A_{-1}A_{-1}^* - 2A_0A_0^*)$	$\frac{1}{\sqrt{5}}(\mathcal{G}_{1+}\mathcal{G}_{1+}^* + \mathcal{G}_{1-}\mathcal{G}_{1-}^* - 2\mathcal{G}_0\mathcal{G}_0^*)$
2 2 0	$-\frac{1}{5}(A_1A_1^* + A_{-1}A_{-1}^* + 4A_0A_0^*)$	$-\frac{1}{5}(\mathcal{G}_{1+}\mathcal{G}_{1+}^* + \mathcal{G}_{1-}\mathcal{G}_{1-}^* + 4\mathcal{G}_0\mathcal{G}_0^*)$
2 2 −1	$-\frac{3}{5}(A_0A_{-1}^* + A_1A_0^*)$	$-\frac{3\sqrt{2}}{5}(\mathrm{Re}\,\mathcal{G}_{1+}\mathcal{G}_0^* + i\mathrm{Im}\,\mathcal{G}_{1-}\mathcal{G}_0^*)$
2 2 −2	$-\frac{6}{5}A_1A_{-1}^*$	$-\frac{3}{5}\left\{\lvert\mathcal{G}_{1+}\rvert^2 - \lvert\mathcal{G}_{1-}\rvert^2 + 2i\mathrm{Im}\,(\mathcal{G}_{1-}\mathcal{G}_{1+}^*)\right\}$

Table 6. *Moments for $B^0 \to (e^+e^-)_\psi(K_S^0\pi^0)_{K^*}$ in terms of helicity amplitudes.*

Depending on the relative strengths of different CP contributions the various moments $T_{J_L J_R M}$ will show different asymmetries. Linear combinations of moments can always be found which give undiluted asymmetry measurements. The various moments are given in Table 6, where we have defined the definite CP quantities $\mathcal{G}_{\lambda\pm} = (A_\lambda \pm A_{-\lambda})/\sqrt{2}$ and $\mathcal{G}_0 = \mathcal{G}_{0+}/\sqrt{2} = A_0$. Table 7 presents the results of a similar analysis for the decay into two spin 1 particles which each in turn decay to two spin zero particles; for example, the mode $D^{*+}D^{*-}$, where both D^*'s decay to $D\pi$. Clearly in either case we can extract the quantities

$$\mathcal{G}_{1+}\mathcal{G}_{1+}^*, \quad \mathcal{G}_{1-}\mathcal{G}_{1-}^*, \quad \mathcal{G}_0\mathcal{G}_0^*, \quad \mathrm{Re}\,\mathcal{G}_{1+}\mathcal{G}_0^*,$$

and

$$\mathrm{Im}\,\mathcal{G}_{1+}\mathcal{G}_{1-}^* \text{ and } \mathrm{Im}\,\mathcal{G}_{1-}\mathcal{G}_0^*.$$

The first four of these are each definite CP quantities, combining B and $\overline{B}$ data they can each be used to give an asymmetry measurement. The last two quantities represent interference terms between CP odd and CP even amplitudes, which have a more complicated time dependence. They depend only quadratically on the CP asymmetry and so are less sensitive for small asymmetries.

J_L	J_R	M	$4\pi T_{J_L J_R M}$	$4\pi T_{J_L J_R M}$
0	0	0	$(A_1 A_1^* + A_{-1} A_{-1}^* + A_0 A_0^*)$	$(\mathcal{G}_{1+}\mathcal{G}_{1+}^* + \mathcal{G}_{1-}\mathcal{G}_{1-}^* + \mathcal{G}_0\mathcal{G}_0^*)$
0	2	0	$\frac{1}{\sqrt{5}}(2A_0A_0^* - A_1A_1^* - A_{-1}A_{-1}^*)$	$\frac{1}{\sqrt{5}}(2\mathcal{G}_0\mathcal{G}_0^* - \mathcal{G}_{1+}\mathcal{G}_{1+}^* - \mathcal{G}_{1-}\mathcal{G}_{1-}^*)$
2	0	0	$\frac{1}{\sqrt{5}}(2A_0A_0^* - A_1A_1^* - A_{-1}A_{-1}^*)$	$\frac{1}{\sqrt{5}}(2\mathcal{G}_0\mathcal{G}_0^* - \mathcal{G}_{1+}\mathcal{G}_{1+}^* - \mathcal{G}_{1-}\mathcal{G}_{1-}^*)$
2	2	0	$\frac{1}{5}(A_1A_1^* + A_{-1}A_{-1}^* + 4A_0A_0^*)$	$\frac{1}{5}(\mathcal{G}_{1+}\mathcal{G}_{1+}^* + \mathcal{G}_{1-}\mathcal{G}_{1-}^* + 4\mathcal{G}_0\mathcal{G}_0^*)$
2	2	-1	$\frac{3}{5}(A_0A_{-1}^* + A_1A_0^*)$	$\frac{3\sqrt{2}}{5}(\mathrm{Re}\,\mathcal{G}_{1+}\mathcal{G}_0^* + i\mathrm{Im}\,\mathcal{G}_{1-}\mathcal{G}_0^*)$
2	2	-2	$\frac{6}{5}\, A_1A_{-1}^*$	$\frac{3}{5}\,\{\lvert\mathcal{G}_{1+}\rvert^2 - \lvert\mathcal{G}_{1-}\rvert^2 + 2i\mathrm{Im}\,(\mathcal{G}_{1-}\mathcal{G}_{1+}^*)\}$

Table 7. *Moments for the B^o decay to two spin one particles, each of which subsequently decays to two spin zero particles.*

To display the time–dependent phase structure explicitly we introduce the parametrization

$$\mathcal{G}_{\lambda\pm}(t) = g\, G_{\lambda\pm}\, e^{-imt}\, e^{-\Gamma t/2}\left[\cos\frac{\Delta mt}{2} \pm i\eta\,\lambda_{KM}\,\sin\frac{\Delta mt}{2}\right] \tag{B.10}$$

where $\eta = \xi\,(-1)^s$ and ξ is given in Table 1. For the mode $\psi K_S^0\pi^0$ we have $\eta = -\xi = +1$ The quantity g is the phase of the CKM matrix elements and the quantity G_j contains any phases from final-state interactions and other strong-interaction effects as well as the magnitude of the time–zero amplitude. The CP–violating quantities are contained in the λ_{KM}, in the standard model $\lambda_{KM} = e^{2i\phi}$ where $\phi = -\beta$ or α or $-\gamma$ is one of the angles of the unitarity triangle, see Figure 2.

The equivalent quantities for the $\overline{B}_{\text{phys}}$ decays are

$$\begin{aligned}\overline{\mathcal{G}}_{\lambda\pm}(t) &= \pm\eta g^*\, G_{\lambda\pm}\, e^{-imt}\\ &\quad\times e^{-\Gamma t/2}\left[\cos\frac{\Delta mt}{2} \pm i\eta\,\lambda_{KM}^{-1}\,\sin\frac{\Delta mt}{2}\right]\end{aligned} \tag{B.11}$$

One then sees that

$$\lvert\mathcal{G}_{\lambda\pm}\rvert^2 = \lvert G_{\lambda\pm}\rvert^2\, e^{-\Gamma t}\,(1 \mp \eta\,\mathrm{Im}\,\lambda_{KM}\,\sin\Delta mt) \tag{B.12}$$

and the equivalent quantity extracted from $\overline{B}$–decays give a simple time dependent asymmetry. For example

$$a(1+,1+) = \frac{\mathcal{G}_{1+}\mathcal{G}^*_{1+} - \overline{\mathcal{G}}_{1+}\overline{\mathcal{G}}^*_{1+}}{\mathcal{G}_{1+}\mathcal{G}^*_{1+} + \overline{\mathcal{G}}_{1+}\overline{\mathcal{G}}^*_{1+}} = -\eta\,\mathrm{Im}\,\lambda_{KM}\,\sin\Delta mt \tag{B.13}$$

Similarly, the interference term between two same-CP amplitudes gives a direct asymmetry measurement, for example

$$\mathrm{Re}\mathcal{G}_{1+}\mathcal{G}^*_0 = \mathrm{Re}(G_{1+}G^*_0)(1-\eta\,\mathrm{Im}\,\lambda_{KM}\,\sin\Delta mt)\,e^{-\Gamma t}\ . \tag{B.14}$$

Thus we have several asymmetry measurements, one for each possible pair of same–CP contributions. A best asymmetry can be obtained by minimizing the error on an arbitrary linear sum of these values. This requires some knowledge of the relative sizes of the various $\mathcal{G}$'s. The even-odd interference terms are less readily used. We find

$$\begin{aligned}\mathrm{Im}\,\mathcal{G}_{1+}\mathcal{G}^*_{1-} = e^{-\Gamma t}\Big[&\mathrm{Re}(G_{1+}G^*_{1-})\eta\,\mathrm{Re}\lambda_{KM}\,\sin\,\Delta mt\\ &+\mathrm{Im}\,(G_{1+}G^*_{1-})\,\cos\,\Delta mt\Big]\end{aligned} \tag{B.15}$$

which is not particularly useful for extracting the value of λ_{KM}.

Experimentally, we obtain the moments by weighting the experimental events by the $Y_{\ell m}$. For example, the T_{222} moment is extracted from the data by calculating

$$T_{222} \cong \frac{1}{N_{\mathrm{evt}}}\sum_{i=1}^{N_{\mathrm{evt}}} Y_{2,2}(\Omega^i_\psi)\,Y_{2,2}(\Omega^i_K)\ . \tag{B.16}$$

The $M = 1,2$ terms will have a $\phi_\psi + \phi_K$ dependence, with our definition of axes this is the phase between the planes of the two two–particle decay states in the B rest frame. To predict the time dependence of the moments one needs to substitute Eq. (B .10) in Eq. (B .9). The relevant time–dependent expression has the form

$$
\begin{aligned}
A_\lambda A^*_{\lambda'} = \tfrac{1}{2} e^{-\Gamma t} \Big\{ & [G_{\lambda +} G^*_{\lambda' +} + G_{\lambda -} G^*_{\lambda' -}] \\
& + [G_{\lambda +} G^*_{\lambda' -} + G_{\lambda -} G^*_{\lambda' +}] \cos(\Delta m t) \\
& + i\eta [G_{\lambda +} G^*_{\lambda' -} - G_{\lambda -} G^*_{\lambda' +}] \mathrm{Re}(\lambda_{KM}) \sin(\Delta m t) \\
& - \eta\, [G_{\lambda +} G^*_{\lambda' +} - G_{\lambda -} G^*_{\lambda' -}] \mathrm{Im}\,(\lambda_{KM}) \sin(\Delta m t) \Big\},
\end{aligned} \tag{B.17}
$$

where $G_{0+} = \sqrt{2} G_0$. Thus, we see that the general moment has three terms with distinct time dependent behaviors $e^{-\Gamma t}$, $e^{-\Gamma t} \cos(\Delta m t)$, and $e^{-\Gamma t} \sin(\Delta m t)$. Extracting the moments requires convolution with the relevant resolution function.

A similar analysis for the decay

$$
\begin{array}{llll}
B^0 \rightarrow & \Delta & \overline{\Delta} & \\
 & \big| & \hookrightarrow & \overline{p}\pi \\
 & \hookrightarrow & p\pi &
\end{array}
\quad .
$$

can be carried out. In this case the proton (or antiproton) helicity is not observed. Summing over the possible helicity values once again eliminates odd values for J_L and J_R. The distinct moments are given in Table 8, in addition $T_{200} = T_{020}$ and $T_{220} = \frac{1}{5} T_{000}$. From this one can identify the definite *CP* quantities

$$
\mathrm{Re}\,(\mathcal{G}_{(3/2)+} \mathcal{G}^*_{(1/2)+}) \quad \text{and} \quad \mathrm{Re}\,(\mathcal{G}_{(3/2)-} \mathcal{G}^*_{(1/2)-})
$$

and thus this mode can be used to measure the *CP* violating asymmetry. A similar analysis can be applied to any spin 3/2

$$
\begin{array}{llll}
B \rightarrow \Lambda & \overline{\Lambda} & & \\
\hookrightarrow p\pi & \hookrightarrow \overline{p}\pi &
\end{array}
\quad \text{or} \quad
\begin{array}{ll}
\Sigma & \overline{\Sigma} \\
\hookrightarrow p\pi & \hookrightarrow \overline{p}\pi
\end{array}
$$

channels. For two spin 1/2 particles which each decay strongly to a nucleon and a spin zero meson only T_{000} survives after summing over nucleon and antinucleon spins, hence one cannot construct undiluted *CP* asymmetries in these cases. We have not studied the situation for weak decays of such particles.

J_L	J_R	M	$\frac{4\pi}{4} T_{J_L J_R M}$
0	0	0	$A_{3/2}A^*_{3/2} + A_{-3/2}A^*_{-3/2} + A_{1/2}A^*_{1/2} + A_{-1/2}A^*_{-1/2}$
0	2	0	$\frac{1}{\sqrt{5}}(-A_{3/2}A^*_{3/2} - A_{-3/2}A^*_{-3/2} + A_{1/2}A^*_{1/2} + A_{-1/2}A^*_{-1/2})$
2	2	1	$\frac{2}{5}\,(A_{-3/2}A^*_{-1/2} + A_{1/2}A^*_{3/2})$
2	2	2	$\frac{2}{5}\,(A_{-3/2}A^*_{+1/2} + A_{-1/2}A^*_{3/2})$
0	0	0	$\lvert\mathcal{G}_{3/2+}\rvert^2 + \lvert\mathcal{G}_{3/2-}\rvert^2 + \lvert\mathcal{G}_{1/2+}\rvert^2 + \lvert\mathcal{G}_{1/2-}\rvert^2$
0	2	0	$\frac{1}{\sqrt{5}}\,(-\lvert\mathcal{G}_{3/2+}\rvert^2 - \lvert\mathcal{G}_{3/2-}\rvert^2 + \lvert\mathcal{G}_{1/2+}\rvert^2 + \lvert\mathcal{G}_{1/2-}\rvert^2)$
2	2	1	$\frac{2}{5}\Big[\mathrm{Re}\,(\mathcal{G}_{1/2-}\mathcal{G}^*_{3/2-}) + \mathrm{Re}\,(\mathcal{G}_{3/2+}\mathcal{G}^*_{1/2+}) + i\,\mathrm{Im}(\mathcal{G}_{1/2-}\mathcal{G}^*_{3/2+}) - i\,\mathrm{Im}\,(\mathcal{G}_{3/2-}\mathcal{G}^*_{1/2+})\Big]$
2	2	2	$\frac{2}{5}\Big[\mathrm{Re}\,(\mathcal{G}_{3/2+}\mathcal{G}^*_{1/2+}) + i\,\mathrm{Im}\,(\mathcal{G}_{3/2+}\mathcal{G}^*_{1/2-}) + i\,\mathrm{Im}\,(\mathcal{G}_{1/2+}\mathcal{G}^*_{3/2-}) - \mathrm{Re}\,(\mathcal{G}_{3/2-}\mathcal{G}^*_{1/2-})\Big]$

Table 8. *Moments for $B^0 \to (p\pi)_\Delta(\overline{p}\pi)_{\overline{\Delta}}$ in terms of helicity amplitudes.*

APPENDIX C

SENSITIVITY ANALYSIS

For the maximum-likelihood fits, we have assumed that the amplitudes of the decay, except for the *CP*-violating part, are understood from study of the untagged and isospin related channels. The question of the sensitivity of the results to this assumption naturally arises; specifically an error on the even to odd ratio (Γ_+/Γ_-) in the transversity case or the amplitudes $(\mathcal{G}_{1+}, \mathcal{G}_0, \mathcal{G}_{1-})$ in the case of the full angular-distribution analysis will lead to an error in the measured value of asymmetry.

The transversity analysis is relatively simple. We parameterize the error as follows:

$$\delta a/a = (1/a)(da/d\epsilon) \times \delta\epsilon \tag{C.1}$$

where $\epsilon = (\Gamma_+ - \Gamma_-)/(\Gamma_+ + \Gamma_-)$. Evaluating the derivative numerically we find $|(1/a)(da/d\epsilon)| < 1.2$ for all ϵ.

For ψK^*, one can estimate that the data sample which will be available for evaluating ϵ from untagged and isospin-related channels will be ≈ 20 times larger than the tagged sample. This implies that typical errors on ϵ will be $\approx \sqrt{20}$ smaller than the error on a. Thus considering Eq. (C .1), it is clear that the effect of a typical $\delta\epsilon$ on δA is negligible.

In the case of the full angular analysis the situation is more complex. Three amplitudes $(\mathcal{G}_{1+}, \mathcal{G}_0, \mathcal{G}_{1-})$ are needed. Two of the amplitudes ($\mathcal{G}_{1+}$ and $\mathcal{G}_0$) are *CP* even and the third ($\mathcal{G}_{1-}$) is *CP* odd. We parameterize the errors on these amplitudes by rotations between the magnitudes of two amplitudes and by errors on the relative phases. For example, our estimate for the magnitudes of the $\mathcal{G}_{1+}$ and $\mathcal{G}_0$ amplitudes might be related to the true values as follows:

$$|\widehat{\mathcal{G}}_{1+}| = |\mathcal{G}_{+1}|\cos(x) + |\mathcal{G}_0|\sin(x) \tag{C.2}$$

$$|\widehat{\mathcal{G}}_0| = |\mathcal{G}_0|\cos(x) - |\mathcal{G}_{1+}|\sin(x)\,, \tag{C.3}$$

where the $\widehat{\mathcal{G}}$'s represent the estimated values and the plain $\mathcal{G}$'s represent the true values. Similar relationships can be used to quantify the possible experimental confusion between $|\mathcal{G}_{1+}|$ and $|\mathcal{G}_{1-}|$, and between $|\mathcal{G}_0|$ and $|\mathcal{G}_{1-}|$. Figure 6 shows $(1/a)(da/dx)$ for each possible angle of confusion. Note that confusion between the two *CP*-even states ($\mathcal{G}_{1+}$ and $\mathcal{G}_0$) has little effect but that confusion between either *CP*–even and the *CP*–odd amplitudes typically produces noticeable effects. Thus it appears that the overall *CP*–even to *CP*–odd ratio is the most sensitive parameter. As seen above it should be possible to determine this parameter to an accuracy much better than needed using the untagged and isospin-related channels. We have also investigated the effect of phase errors in $\mathcal{G}_{\lambda\pm}$ and we find them to be small. For example, a phase error of 30 degrees changes the asymmetry by only 0.003 when the true asymmetry is 0.15 and $\Gamma_+ = \Gamma_-$ (the worst case).

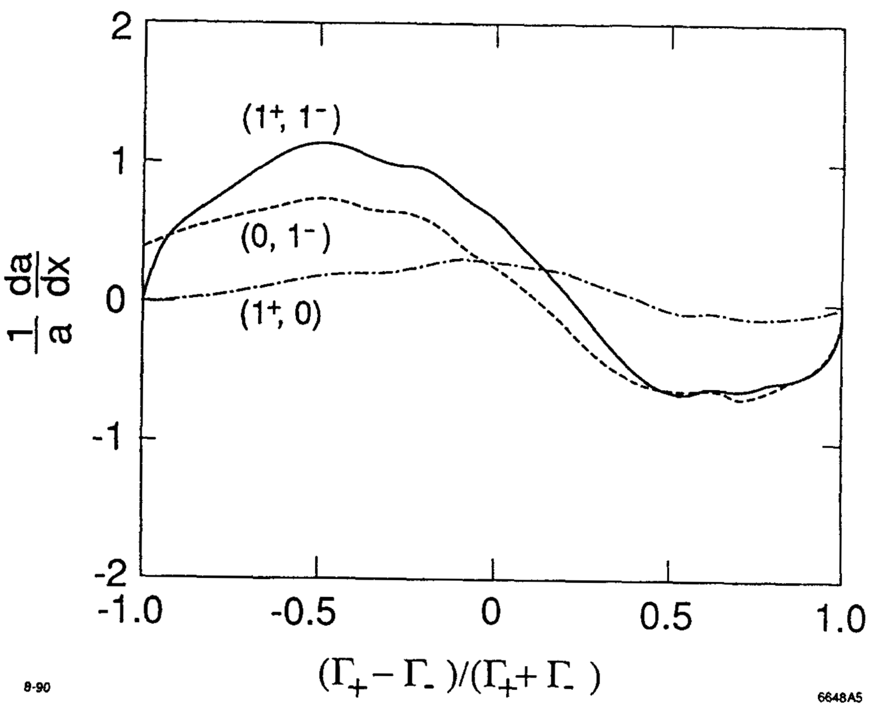


Figure 6. *The fractional derivative of the asymmetry with respect to an angle, (x), which describes the confusion between the amplitudes. The derivative was obtained numerically.*

The final analysis will probably be a maximum-likelihood fit of all the parameters (the three amplitudes and the *CP*-violating asymmetry) to all the data samples (tagged, untagged, isospin-related). This analysis of sensitivity of the measured asymmetry to assumed values of the parameters indicates that the resulting errors will be only marginally worse than single-parameter analysis used in this paper for illustrative purposes.

REFERENCES

1. For a recent review of the topic of *CP* violation in the *B* system, see, for example, I. I. Bigi, V. Khoze, N. G. Uraltsev and A. I. Sanda in "*CP Violation*", edited by C. Jarlskog (World Scientific), pp 175–248.
2. C. A. Nelson, *Phys. Rev.* **D30**, 1937 (1984). J. R. Dell'Aquila and C. A. Nelson, *Phys. Rev.* **D33**, 80 (1986) and *Phys. Rev.* **D33**, 101 (1986).
3. J. D. Bjorken, *Nucl. Phys.* **B** (Proc. Suppl.) **11**, 325 (1989).
4. B. Kayser in *Proceedings of Workshop on B Factories and Related Physics Issues*, (Blois, France, 1989). B. Kayser, M. Kuroda, R. D. Peccei and A. I. Sanda, *Phys. Lett.* **B237**, 508 (1990).
5. G. Valencia, *Phys. Rev.* **D39**, 3339 (1989).
6. I. Dunietz and A. Snyder, SLAC–PUB–5234.
7. M. Kobayashi and T. Maskawa, *Prog. Theor. Phys.* **49**, 652 (1973). P. Krawczyk, D. London, R. D. Peccei and H. Steger, *Nucl. Phys.* **B307**, 19 (1988). C.O. Dib, I. Dunietz, F.J. Gilman and Y. Nir, *Phys. Rev.* **D41**, 1522 (1990). C.S. Kim, J.L. Rosner and C.P. Yuan, *Phys. Rev.* **D42**, 96 (1990).
8. See, for example, M. Jacob and G. C. Wick, *Ann. Phys.* **7**, 404 (1959). We use the conventions defined by J. D. Jackson in *Proceedings of High Energy Physics*, edited by. C. deWitt and M. Jacob (Les Houches Lectures, 1965).
9. S. M. Berman and M. Jacob, *Phys. Rev.* **B139**, 1023 (1965).
10. For a pedagogical review of helicity formalism, see J. D. Richman, Ph.D. Thesis (1985), Caltech Report CALT-68–1231. M. Jacob and S. M. Berman, SLAC Report–43 (1965).
11. For the example of ψK_S^0, this is discussed in H. J. Lipkin and A. I. Sanda, *Phys. Lett.* **201B**, 541 (1988).
12. H.J. Lipkin and S. Meshkov, *Phys. Rev.* **143**, 1269 (1966).

13. A. Bohr, *Nucl. Phys.* **10**, 486 (1959).

14. H.J. Lipkin, *Phys. Reports* **8C**, 173 (1973).

15. D. London and R. D. Peccei, *Phys. Lett.* **B223**, 257 (1989); M. Gronau, *Phys. Rev. Lett.* **63**, 1451, (1989); B. Grinstein, *Phys. Lett.* **B229**, 280 (1990).

16. T. Nakada in *Heavy Quark Physics*, edited by P. Drell and D. Rubin, *AIP Conference Proceedings 196* (1989); F. Le Diberder, *Ba$\overline{B}$ar* Note #33.

17. Non-*CP*-violating asymmetries arise between the rates $B^)$ and $\overline{B^0}$ into identical final states, both in their particle content and in their angular dependence. This non-*CP*-violating asymmetry occurs because the rate of $B^0_{\rm phys} \rightarrow f$ is not compared to its *CP* conjugate process $\overline{B^0}_{\rm phys} \rightarrow \overline{f}$, but rather to the process $\overline{B^0}_{\rm phys} \rightarrow f$, and is due to an inteerference between *CP*-even and *CP*-odd amplitudes. Here the final state f is defined by particle content and by a set of angles and $barf$ is the *CP*-conjugated state with the set of *CP*-conjugated angles.

18. W. Toki, *Ba$\overline{B}$ar* Note # 53.

Getting the Maximum Information from B Decays into CP Eigenstates

Harry J. Lipkin

Abstract

B decays into CP eigenstates have low branching ratios and low efficiencies for complete reconstruction of final states. Statistics can be improved by using as many different decay modes as possible and by using incompletely reconstructed decays. We discuss such improvement by partial-wave analyses and missing-mass techniques.

1. The Search for Usable CP Eigenstates

Many of the proposed methods for finding CP-violation effects in B decays involve the use of final states which are CP eigenstates.[1] Since the dominant decay of the b quark is to charm, the dominant decay modes of B mesons into CP eigenstates must contain a charmed quark-antiquark pair. The Cabibbo-favored vertex for creating the additional charmed antiquark also creates a strange quark. Thus dominant CP eigenstates are produced by the following weak transitions at the quark level:

$$b \to c + W^- \to c + \bar{c} + s \tag{1.1a}$$

$$\bar{b} \to \bar{c} + W^+ \to c + \bar{c} + \bar{s} \tag{1.1b}$$

These weak vertices have been noted to be $\Delta I = 0$ transitions, which satisfy isospin relations[2] useful in the analysis of data.

For the B_d and $\bar{B}_d$ decays, these transitions (1.1) lead to a strange final state, which can only be a CP eigenstate if it contains a neutral kaon. Thus the first state considered[1] was $J/\psi K_S^0$. However, since this state and all similar states have very low branching ratios, it is advantageous to consider accumulating data on decays into all possible CP eigenstates.

There are also the nonstrange final states resulting from the Cabibbo-suppressed transitions:

$$b \to c + W^- \to c + \bar{c} + d \tag{1.1c}$$

$$\bar{b} \to \bar{c} + W^+ \to c + \bar{c} + \bar{d} \tag{1.1d}$$

These can be *CP* eigenstates without the requirement of containing a neutral kaon.

The charmed quark-antiquark pair can be present either in a single charmonium state or in a pair of charmed hadrons. Since the pair is produced at a very short range on the scale of the W mass, only S-wave charmonium states are expected to be appreciably produced; namely the J/ψ and η_c and their radial excitations. The other orbitally excited states have wave functions that vanish at short distances because of the centrifugal barrier.

Although many of the decays into final states containing these other S-wave charmonium states may have branching ratios comparable to those of the decays into corresponding states containing the J/ψ, the detection efficiency for these other states is much smaller, and there is probably little to be gained over the use of $J/\psi K_S^0$ unless some tricks are used to enhance the detection efficiency of these other decays. One such trick is to use missing-mass kinematics to include decays which are not completely reconstructed. This is discussed in detail below.

Another possibility to be considered is the use of decay modes containing K_L^0. For every *CP* eigenstate which contains a K_S^0 the corresponding state with the K_S^0 replaced by K_L^0 has the same branching ratio and the opposite *CP* eigenvalue. In an experiment where both types of states can be detected and distinguished, the *CP* asymmetries will be equal and opposite and the effective statistics will be doubled.

2. Getting Information From Incompletely Reconstructed Decays

2.1. Additional Decay Modes into CP Eigenstates

Inclusive K_S^0 spectroscopy has been suggested[3,4] for detecting and using quasi-two-body decay modes of the form $B_d \to XK_S^0$, where X denotes a narrow state like a charmonium or charmed meson state which is not easily detected directly because of low branching ratios to observable final states but which can be identified by missing mass. In these quasi-two-body decays the momentum of the K_S^0 is uniquely defined in the rest frame of the decaying B and rough estimates indicate that if it is produced in the decay of a slowly moving B from Υ(4S) decay, the smearing of the momentum may not be excessive. Therefore such decays might be observable as a signal in the inclusive K_S^0 spectrum.

Some particular decays suggested for CP tests are:

$$B_d \to J/\psi K_S^0 \tag{2.1a}$$

$$B_d \to \psi' K_S^0 \to \psi\pi\pi K_S^0 \tag{2.1b}$$

$$B_d \to \psi'' K_S^0 \to D\bar{D}K_S^0 \tag{2.1c}$$

$$B_d \to \eta_c K_S^0 \tag{2.1d}$$

$$B_d \to \eta_c' K_S^0 \to \eta_c\pi\pi K_S^0 \tag{2.1e}$$

$$B^0 \to D^o K_S^0 \tag{2.1f}$$

The three-body decay

$$B_d \to K_S^0 + \pi^0 + \eta_c \tag{2.2a}$$

which includes the case where the $K_S^0 - \pi^0$ can be in a K^* also leads to a CP eigenstate and can be observed by the inclusive method, if only the K_S^0 and π^0 are detected and the η_c is identified by missing mass. This decay mode can be expected to have a branching ratio comparable to that of the corresponding decay

$$B_d \to K_S^0 + \pi^0 + J/\psi \tag{2.2b}$$

which has been suggested[5] as a candidate for CP violation experiments. This decay mode can also be observed by the inclusive method if only the K_S^0 and π^0 are detected and the J/ψ is identified by missing mass. The η_c decay mode (2.2a) has the advantage over the J/ψ decay mode (2.2b) in that the three-pseudoscalar state is a CP eigenstate with the same eigenvalue for all partial waves and can therefore be used directly in experiments without further analysis. The vector-pseudoscalar-pseudoscalar state has partial waves with both eigenvalues of CP, and some partial wave analysis is needed for its use in CP experiments.[5]

Any inclusive experiment which looks at final states containing charmonium can also look for η_c as well as J/ψ. The CP eigenvalue is the same[3,4] for the $\eta_c K_S^0$ decay as in $J/\psi K_S^0$ and also for all correspopnding states containing radially excited charmonium states. Although the η_c and J/ψ have opposite intrinsic CP eigenvalues, the $\eta_c K_S^0$ decay is S-wave while the $J/\psi K_S^0$ is P-wave, thus giving the same overall CP eigenvalue for both states; namely odd. Thus the $B_d \to \eta_c K_S^0$ decays can also be used in inclusive decays where the η_c is not observed, and data from all the decays $B_d \to (c\bar{c})K_S^0$ can be combined to improve statistics in an experiment which tests CP by observing asymmetries in the decays of a tagged B_d into a CP eigenstate. It may not even be necessary to resolve the missing mass distribution to separate different charmonium states if the background is not too high.

The inclusive spectrum technique can be applied to all the decays (2.1) with narrow states recoiling against the kaon. The $B^0 \to \psi' K_S^0$ signal might be sharpened by observing the two charged pions emitted in the decay $\psi' \to J/\psi\pi^+\pi^-$ which has a branching ratio of 33% and placing the appropriate constraints on the missing mass of the unobserved J/ψ. For the case of the $B^0 \to \eta_c K_S^0$, where direct detection of the η_c is difficult, this method might be the best way to either detect this mode or to put a reliable upper limit on the branching ratio.

The decay $B^0 \to D^0 K^0_S$ may be more easily detected in this way than directly because of the higher energy of the K^0_S and the consequently lower background in the inclusive spectrum. If this decay mode turns out to be appreciable, it may offer interesting possibilities since it is also a mixed state. It could be used in exclusive spectroscopy with events selected for D^0 decays into CP eigenstates.

The decays

$$B \to \psi'' K \to D\bar{D}K \tag{2.3a}$$

have not yet been observed but may have a significant branching ratio, since[6]

$$B\{B^+ \to \psi(2S)K^+\} = 0.22 \pm 0.17 \times 10^{-2} \tag{2.3b}$$

This implies by isospin[2]

$$B\{B^0 \to \psi(2S)K^0_S\} = \frac{\tau(B^0)}{\tau(B^+)} \cdot 0.11 \pm 0.08 \times 10^{-2} \tag{2.3c}$$

However, the branching ratios for detecting D's would make the direct observation of this mode with fully reconstructed events extremely difficult. The question is whether the $\psi'' K^0_S \to D\bar{D}K^0_S$ decay mode can be detected without fully reconstructing the event because of the unique momentum of the kaon emitted from the decay of a B at rest, which will hopefully not be too smeared by the initial B momentum, and because of the large branching ratio for the D's into strange particles. The inclusive branching ratios[6] for $D \to KX$ are $24 \pm 8\%$ and $16 \pm 5\%$ for $D^\pm \to K^0_S X$ and $D^o \to K^0_S X$ respectively, and $47 \pm 9\%$ and $66 \pm 8\%$ respectively, when charged kaons are included as well as K^0_S but K^0_L decays are not detected. This gives a good probability for observing a $D\bar{D}K^0_S$ event as $KKK^0_S X$. One can look at the kaon momentum spectrum in events containing two or three kaons and see if there is a signal at the appropriate momentum corresponding to $\psi'' K^0_S$ decay.

The same considerations apply to the case where there an additional π^0 is observed; i.e. to decays of the form $B_d \to XK_S^0\pi^0$, where X is identified by missing mass as charmonium. This would include the $B_d \to \eta_c K_S^0 \pi^0$ and $B_d \to J/\psi K_S^0 \pi^0$ decays.

2.2. *The Signal-to-Noise Ratio*

The crucial point which determines the feasibility of this method is the signal-to-noise ratio. Using the inclusive spectrum clearly increases both signal and noise. The signal is improved because it includes all events with the desired final state recoiling against the kaon, with no loss due to branching ratio of the detectable signature, detection efficiencies, etc. The noise level is increased because of the presence of other decay modes which can produce a K_S^0 in this momentum range, and because of the smearing of the momentum due to the initial momentum of the decaying B^0. The question is whether the increase in noise can be controlled sufficiently by judicious cuts to eliminate background so that the overall effect is an improvement in signal to noise.

Consider an experiment which tests CP violation by measuring the asymmetry in the decay of a B^0 - $\bar{B}^0$ pair where one decays into $J/\psi K_S^0$ and the other into a mode with either a positive or a negative lepton. Let $N \pm \xi$ denote the numbers of events in a particular experiment in which this decay occurs with a positive or a negative lepton, respectively.

We compare two possible ways to determine the asymmetry parameter ξ which measures the extent of CP violation. The conventional method uses only events with a fully determined J/ψ decay. Let B denote the branching ratio to the observed final state and η the detection efficiency for this state. The difference denoted by D_{exc} between the number of exclusive events observed with positive and negative leptons is then given by

$$D_{exc} = 2\eta \cdot B \cdot \xi \pm \sqrt{2\eta \cdot B \cdot N} = 2\eta \cdot B \cdot \xi \left[1 \pm \sqrt{N/(2\eta \cdot B \cdot \xi)} \right]. \qquad (2.4a)$$

The fractional statistical error is therefore

$$\delta D_{exc} = \sqrt{N/(2\eta \cdot B \cdot \xi)} \tag{2.4b}$$

However, one can also look at the inclusive K_S^0 spectrum in the momentum bin which contains all the $B^0 \to J/\psi K_S^0$ decays and determine the difference between the number of events observed with positive and negative leptons. Here the efficiency is 100% for detecting these decays, but there is also a background denoted by B so that the number of events observed in the inclusive spectrum bin with positive and negative leptons is now $N + B \pm \xi$. For this case

$$D_{inc} = 2\xi \pm \sqrt{2(N+B)} = 2\xi \left[1 \pm \sqrt{(N+B)/2\xi}\right], \tag{2.5a}$$

and the fractional statistical error is

$$\delta D_{inc} = \sqrt{(N+B)/2\xi]}\,. \tag{2.5b}$$

Which method gives a better result depends upon the ratio

$$\frac{\delta D_{inc}}{\delta D_{exc}} = \frac{\sqrt{\eta \cdot B}}{\sqrt{N/(N+B)}}\,. \tag{2.6}$$

The question is, therefore, which is greater, the detection efficiency $\times$ branching ratio $\eta \cdot B$ in the conventional exclusive method or the signal to background ratio $\frac{N}{(N+B)}$ in the inclusive method. In the case where $\eta \cdot B \approx 1/10$, a reasonable value for the detection of the J/ψ via the leptonic decay mode, the inclusive technique will be competitive even if the background is ten times bigger than the signal. Combining the exclusive and inclusive data can give improved statistics as well as a consistency check.

One can also examine inclusive J/ψ spectroscopy where the J/ψ is completely determined and look at the momentum range for the J/ψ relevant to the $J/\psi K_S^0$ decay. Here one loses the product of detection efficiency $\times$ branching ratio $\eta \cdot B$ for the J/ψ in comparison with the case of inclusive K_S^0 spectroscopy, but gains both the efficiency of K_S^0 detection and a factor of four because these events include both the

K_L^0 events and the $K^\pm$ events from decays of $B^\pm$ produced equally with B^0 in $\Upsilon(4S)$ decays. This gives good statistics for the determination of the branching ratio. However, the determination of *CP* asymmetries without observing the kaon is probably not feasible, since the K_S^0 and K_L^0 events have opposite *CP* and give opposite *CP* asymmetries when used to tag another B decay, while decays tagged by the $K^\pm$ events should show no *CP* asymmetries.

2.3. Using Cuts to reduce Background

All possible cuts for reducing background should be investigated, with particular attention paid to the search for better cuts possible in future experiments with improved technologies.

A principal contribution to the background of inclusive K_S^0 events comes from decays into charmed mesons which then decay into a mode containing a K_S^0. Possible cuts to eliminate these events are:

1. Semileptonic decays of either a B or D could be eliminated by cutting out all decays with a lepton present in addition to the K_S^0, while keeping those with a lepton pair from J/ψ decay.

2. The $B \to D\pi$ decay should give a monoenergetic high energy pion which can be used as a signature for background to be eliminated.

3. Background events from charm can be identified if all the particles from the D decay are detected. These can be found by searching for events in which the invariant mass of some combination of particles together with the K_S^0 is equal to the D mass and corresponds to a decay mode of the D.

4. Charm events can also be identified if there is sufficiently accurate time information from vertex detectors indicating that the K_S^0 was not emitted from the prompt B decay vertex but from a secondary vertex of charm decay.

5. Charm events can also be identified if all the particles emitted originally before the D decay are detected. These can be found by searching for events in which the missing mass

recoiling against some combination of particles is equal to the D mass.

Events which escape such detection must have at least two unidentified particles in the final state, with one missing from the original $B \to D + X$ decay and the other from the charm decay. Monte Carlos seem to indicate that decays into charm which produce a K_S^0 in the momentum region investigated involve low multiplicity B decays or D decays. In such low multiplicity events, the probability of having a large number of missing particles may not be so great, and the cuts suggested above may help reduce background.

The remaining events with more unidentified particles might be examined to look for differences in signatures between these events and the events for the desired modes; e.g. in their multiplicities, missing masses, missing energies and angular distributions.

In the case, for example, where a K_S^0 is observed with a momentum corresponding to the $B_d \to \eta_c K_S^0$ decay, one can examine the remaining particles in the rest frame of the system (hopefully η_c) recoiling against the K_S^0. In this frame the energy of the K_S^0 is 3146 MeV, while the total energy of the remaining particles is just the η_c mass of 2980 MeV distributed among several particles. The angular distribution of the η_c decay must be isotropic having no correlation with the kaon momentum. In background events in which the K_S^0 comes from a charm decay, the remaining particles can be expected to be correlated with the kaon.

We also note that for multi-pion decay modes of the J/ψ, the number of neutral pions is required to be odd by G-parity, while there is no such restriction for the charm background. One might expect the dominant unreconstructable background from charm decay would have one π^0 emitted in the original B decay and one π^0 emitted in charm decay. This might lead to a very different missing-mass and missing-energy spectrum from the dominant unreconstructable J/ψ decay which would have three undetected π^0's.

Additional cuts can be based on time information obtainable with vertex detectors having a high resolution. In experiments where the $\Upsilon(4S)$ is produced in flight by $e^+ - e^-$

annihilation with unequal energies, it is in principle possible to resolve the vertices for the decays of the two mesons and also the secondary vertices for the decays of any charmed particles produced in the primary B decay.

Since charmonium decays are either strong or electromagnetic, all B decays of the form $B^0 \to (c\bar{c})K_S^0$ where $(c\bar{c})$ denotes a bound charmonium state below the naked-charm threshold are prompt and all observed hadrons are emitted from a single vertex. Decays to a bound charmonium state above the naked-charm threshold will have two additional vertices corresponding to the decays of the two secondary charmed particles. Thus, in principle, B decays can be separated into three groups depending upon the number of charmed particles in the final state.

1. Decays with no charmed particles in the final state include mainly decays into charmonium, and a small number of charmless B decays which probably will not produce kaons.

2. Decays with a single charmed particle in the final state are the main source of background for the detection of K_S^0-charmonium final states via the inclusive spectrum. If these can be separated out by the presence of the secondary vertex, the background will be reduced considerably.

3. Decays with two charmed particles in the final state will include events where charmonium is produced above the $D\bar{D}$ threshold and will be interesting in their own right.

3. USE OF THE $(c\bar{c})K_S^0\pi^0$, $(c\bar{c})\pi^+\pi^-$ AND $D^*\bar{D}^*$ DECAY MODES

3.1. The Transversity Analysis

A difficulty arises in considering more complicated states containing either additional particles or replacing the pseudoscalar kaon by a strange resonance which has nonzero spin. Most of these decays have more than one partial wave allowed. Thus, even though all the final-state particles are eigenstates of charge conjugation, states of both orbital parities can be allowed, giving rise to states having components with both CP eigenvalues. The different partial waves can be separated by analysis of angular distributions.[5]

A general treatment of the problem of using symmetries and angular distributions to separate quasi-three-body final states containing several partial waves into their components which are CP eigenstates is given elsewhere in these proceedings. We note here a general property of a B decay into a three-body or quasi-three-body final state which can be used in several ways. The three final-state momenta define a plane and are invariant under the transformation of reflection in the plane. The reflection can be written as a product of a space inversion P and a 180° rotation about an axis normal to the plane, which we choose as the xy plane.[7]

$$R_{xy} \equiv Pe^{i\pi J_z} = P_{int} \cdot e^{i\pi S_z} \tag{3.1}$$

where P_{int} denotes the total intrinsic parity of the system, J_z denotes the projection of the total angular momentum of the three particle state on the z axis and S_z denotes the projection of the total spin angular momentum of the three particle state on the z axis. This projection of the total spin on the axis normal to the plane is conventionally called transversity, by analogy with helicity which is the projection on an axis in the direction of momentum. The equality follows since all three momenta in the center of mass system remain invariant under R_{xy} and therefore it acts only on the internal degrees of freedom.

Since the initial state has spin zero, the final state must also have spin zero and be invariant under rotations in the center of mass system. For any $J = 0$ state $|J = 0\rangle$

$$R_{xy}\,|J=0\rangle = Pe^{i\pi J_z}\,|J=0\rangle = \\ P_{int}\cdot e^{i\pi S_z}\,|J=0\rangle = P\,|J=0\rangle \tag{3.2a}$$

Thus for a state in which all particles have well-defined intrinsic parities, a measurement of the transversity S_z determines the parity of the state; *i.e.* whether the state is 0^- or 0^+ coming respectively from parity-conserving or parity-violating weak decays.

In decays like $B_d \to J/\psi K_S^0 \pi^0$ and $B_d \to \eta_c K_S^0 \pi^0$, where each of the three particles in the final state has a definite intrinsic parity and intrinsic charge conjugation the CP eigenvalue of the state is given by

$$CR_{xy} \left| J=0 \right\rangle = CPe^{i\pi J_z} \left| J=0 \right\rangle = \\ CP_{int} \cdot e^{i\pi S_z} \left| J=0 \right\rangle = CP \left| J=0 \right\rangle \qquad (3.2b)$$

and a measurement of the transversity S_z determines the CP eigenvalue of the state.

3.2. The Decay $B_d \to (c\bar{c}) K_S^0 \pi^0$

The relation (3.2b) applies to any three-body or quasi-three-body state with a well defined $(CP)_{int}$. For the final state $\eta_c K_S^0 \pi^0$, where $(CP)_{int}$ is even and S_x is always zero for three spinless particles, we find that CP is always even for all possible partial waves in the final state, including all the states where the $K - \pi$ system is in any K^* resonance. This can also be seen by examining all allowed partial waves. The K_S^0 has even intrinsic CP, the η_c and π^0 are both odd (even C, odd P). There are two independent relative orbital angular momenta for this three-body system, but they must both be equal in order to couple to the spin zero required by angular momentum conservation. Thus although all values of this equal orbital angular momentum are allowed, and there can be many partial waves, all have even orbital parity and even intrinsic CP. They are therefore all even CP. This holds also for all radial excitations of the η_c.

For the final state $J/\psi K_S^0 \pi^0$, where $(CP)_{int}$ is odd, the total CP is odd if S_z is zero and even if S_z is ± 1. Note also that these results apply also to all radial excitations of these charmonium states.

Further applications of the result (3.2b) to decays into states containing several partial waves with different CP are discussed elsewhere.

3.3. *The Decay $B_d \to D^{*+}D^{*-}$*

The final state $D^{*+}D^{*-}$ contains several partial waves which have both CP eigenvalues. However, a transversity measurement can be used to separate the different CP eigenstates.

Consider a $D^{*+}D^{*-}$ state in which the D^{*+} has momentum $\vec{k}$ and the D^{*-} has momentum $-\vec{k}$ and the spins of the two mesons are coupled to total spin S with $S_z = M$

$$\begin{aligned}&\left|D^{*+}(\vec{k})D^{*-}(-\vec{k}); S, M\right\rangle \\ &\qquad = \sum_{m_1,m_2} (11m_1m_2|11SM)\left|D^{*+}_{m_1}(\vec{k})D^{*-}_{m_2}(-\vec{k})\right\rangle\end{aligned} \tag{3.3}$$

where $D^{*\pm}_m(\vec{k})$ denotes the state of a $D^{*\pm}$ meson with momentum $(\vec{k})$ and spin projection $S_z = m$ and $(11m_1m_2|11SM)$ is a Clebsch-Gordan coefficient. Since CP reverses the charge and the momentum direction of a D^* meson while leaving the spin unchanged, the net effect of CP on the wave function (3.3) is a spin exchange. Thus

$$\begin{aligned}&CP\left|D^{*+}(\vec{k})D^{*-}(-\vec{k}); S, M\right\rangle \\ &\quad = \sum_{m_1,m_2} (11m_1m_2|11SM)\left|D^{*-}_{m_1}(-\vec{k})D^{*+}_{m_2}(\vec{k})\right\rangle, \qquad (3.4a) \\ &\quad = (-1)^S\left|D^{*+}(\vec{k})D^{*-}(-\vec{k}); S, M\right\rangle\end{aligned}$$

where we have used the symmetry of the Clebsch-Gordan coefficients under spin exchange. The final state in B decay has total angular momentum $J = 0$ and therefore orbital angular momentum $L = S$. Thus

$$\begin{aligned}&CP\left|D^{*+}D^{*-}; L = S, J = 0\right\rangle \\ &\quad = (-1)^S\left|D^{*+}D^{*-}; L = S, J = 0\right\rangle \\ &\quad = (-1)^L\left|D^{*+}D^{*-}; L = S, J = 0\right\rangle \\ &\quad = \quad P\left|D^{*+}D^{*-}; L = S, J = 0\right\rangle\end{aligned} \tag{3.4b}$$

The CP eigenvalue for this state is therefore equal to the parity eigenvalue P. The parity eigenvalue can be obtained by the relation (3.2a) where we consider the plane defined in the center-of-mass system by the momenta of the D^{*+} and the D and π coming from the decay of the D^{*-}. Since the intrinsic parity of the $D^{*+} - D - \pi$ system is negative, eq. (3.2a) shows that the total CP is odd if S_z is zero for the D^{*+} and even if S_z is ± 1. The same result is obtained from the correponding analysis using the spin of the D^{*-} and the transversity plane defined by the momenta of the D^{*-} and the D and π coming from the decay of the D^{*+}.

4. B Decay Modes with an η_c in the Final State

Consider the decays

$$B_d \to K_S^0 + \eta_c \to K_S^0 + \pi + \pi + \eta \tag{4.1a}$$

$$B_d \to K_S^0 + \pi^0 + \eta_c \to K_S^0 + \pi^0 + \pi + \pi + \eta \tag{4.1b}$$

where the K_S^0 and all the pions are observed but the η is not. This is analogous to the decay

$$B_d \to K_S^0 + \psi' \to K_S^0 + \pi + \pi + J/\psi \tag{4.2}$$

where the K_S^0 and the two pions are observed but the ψ is not. In all these cases an additional kinematic constraint is imposed if the initial B meson is at rest by the condition that the missing mass recoiling against the $K_S^0 + n\pi$ system must have a definite value, either the mass of the η or the mass of the J/ψ.

Consider the principal decay modes of the η_c of this type which can be detected with no more than one missing particle. The total branching ratio into such modes[6] is about 15% . These include:

$$B(\eta_c \to \eta\pi\pi) = 5.0 \pm 1.1 \tag{4.3a}$$

$$B(\eta_c \to \eta'(958)\pi\pi) = 4.1 \pm 1.7 \tag{4.3b}$$

$$B(\eta_c \to K\bar{K}\pi) = 5.5 \pm 0.8 \tag{4.3c}$$

$$B(\eta_c \to K^+K^-\pi^+\pi^-) = 2.0 \pm 0.3 \qquad (4.3d)$$

$$B(\eta_c \to 2\pi^+2\pi^-) = 1.2 \pm 0.3 \qquad (4.3e)$$

The decays (4.3a) and (4.3b) are already in the category of decays with only one missing particle when the η or η', respectively, are not observed at all. Additional constraints can be used to reduce background since many decay modes of the η or η' produced in these decays can either be detected completely, or detected with only one missing particle. In particular,

$$B(\eta' \to \eta\pi^+\pi^-) = 44.1 \pm 1.6 \qquad (4.4a)$$

$$B(\eta' \to \rho^o\gamma) = 30.1 \pm 1.4\,. \qquad (4.4b)$$

These can give an additional kinematic constraint on the events. This suggests that it will be useful to investigate B-decay modes involving the η_c as well as the J/ψ.

It is therefore of interest to investigate B decays into CP eigenstates containing a K_S^0 and an η_c. This can proceed as follows:

1. Detection and determination of the branching ratios of these modes. This can be most easily achieved by the use of the corresponding charged decays modes, with charged B's and charged kaons. These are uniquely related to the neutral decays by isospin, and will have a higher branching ratio into detectable particles. The values of these branching ratios are in themselves of physical interest as the ratio of the branching ratios into corresponding decay modes involving J/ψ and η_c are predicted in a simple way by theoretical models as the ratio depends only upon helicity factors and the magnitudes of the charmonium wave functions at the origin.

2. Investigation of the use of decay modes with one missing particle. This requires examination of background either from real data or Monte Carlo simulations.

5. Review of Flavor Oscillations and Isospin Relations

5.1. Description of Oscillations in the B_d system

We review for the the record the following description of $B_d - \bar{B}_d$ oscillations for general reference,[1,8,9] using the notation of ref. 9:

$$|\langle B_d|B_d(t)\rangle|^2 = |\langle \bar{B}_d|\bar{B}_d(t)\rangle|^2 = \frac{e^{-\Gamma t}}{2}[1+\cos(\Delta m t)] \tag{5.1a}$$

$$|\langle \bar{B}_d|B_d(t)\rangle|^2 = |\langle B_d|\bar{B}_d(t)\rangle|^2 = \frac{e^{-\Gamma t}}{2}[1-\cos(\Delta m t)] \tag{5.1b}$$

$$|\langle B_2|B_d(t)\rangle|^2 = |\langle B_1|\bar{B}_d(t)\rangle|^2 = \frac{e^{-\Gamma t}}{2}[1-\sin\theta\sin(\Delta m t)] \tag{5.1c}$$

$$|\langle B_1|B_d(t)\rangle|^2 = |\langle B_2|\bar{B}_d(t)\rangle|^2 = \frac{e^{-\Gamma t}}{2}[1+\sin\theta\sin(\Delta m t)] \tag{5.1d}$$

where $|B_d(t)\rangle$ and $|\bar{B}_d(t)\rangle$ denote the states at time t of states which were respectively $|B_d\rangle$ and $|\bar{B}_d\rangle$ at time $t=0$, $|B_1\rangle$ and $|\bar{B}_2\rangle$ denote the CP eigenstates, Γ denotes the mean decay width of the mass eigenstates and Δm the mass difference between them.

These expressions (5.1) can be seen to be equivalent to those given in the standard treatments[1,8] for the case where the difference between the widths of the two eigenstates $\Delta\Gamma$ is set equal to zero. Our CP-violation parameter θ is given in terms of the parameters q, p and ϵ in the conventional notation by the relation

$$\frac{q}{p} = \frac{1-\epsilon}{1+\epsilon} = e^{i\theta} \tag{5.2a}$$

This is consistent only if

$$\left|\frac{q}{p}\right| = 1. \tag{5.2b}$$

The consistency condition (5.2b) is seen in conventional treatments to follow from the requirement that the two states $|B^0(t)\rangle$ and $|\bar{B}^0(t)\rangle$ remain orthogonal for all times t.

5.2. *Isospin Considerations*

The branching ratios for the decays $B_d(\bar{b}d) \rightarrow (c\bar{c})K^o\pi^0$ and $\bar{B}_d(b\bar{d}) \rightarrow (c\bar{c})\bar{K}^o\pi^0$ to neutral final states are related to corresponding decays to charged final states by isospin considerations,[2] because these decays are dominated by the $\Delta I = 0$ transitions (1.1a-b). The $K - \pi$ system must be in a state of isospin 1/2, not only for final states where the $K - \pi$ system is in a known $I = 1/2$ K^* resonance, but also for nonresonant backgrounds.

$$B\{B_d \rightarrow (c\bar{c})K^o\pi^0\} = (1/2)B\{B_d \rightarrow (c\bar{c})K^+\pi^-\} \quad (5.3a)$$

$$B\{\bar{B}_d \rightarrow (c\bar{c})\bar{K}^o\pi^0\} = (1/2)B\{\bar{B}_d \rightarrow (c\bar{c})K^-\pi^+\} . \quad (5.3b)$$

Since the K^o is usually detected experimentally in the $K^0_S \rightarrow \pi^+\pi^-$ decay mode, it is convenient to write

$$B\{B_d \rightarrow (c\bar{c})K^0_S\pi^0 \rightarrow (c\bar{c})\pi^+\pi^-\pi^0\} = (1/6)B\{B_d \rightarrow (c\bar{c})K^+\pi^-\} \quad (5.4a)$$

$$B\{\bar{B}_d \rightarrow (c\bar{c})K^0_S\pi^0 \rightarrow (c\bar{c})\pi^+\pi^-\pi^0\} = (1/6)B\{\bar{B}_d \rightarrow (c\bar{c})K^-\pi^+\} , \quad (5.4b)$$

where we have taken the branching ratio $B(K^0_S \rightarrow \pi^+\pi-) = 2/3$. The factor of six in the relations (5.4) can be used to improve the statistics in any analysis of these decays where the neutral decay is of interest by including charged-mode data together with the data for the decay in the neutral mode.

We now apply the general treatment of ref. 2 to the pair of charge-conjugate final states $(c\bar{c})K^+\pi^-$ and $(c\bar{c})K^-\pi^+$. The B_d decays only to $(c\bar{c})K^+\pi^-$ and not to $(c\bar{c})K^-\pi^+$ and vice versa for the $\bar{B}_d$. Then we can write

$$B[B_d \to (c\bar{c})K^+\pi^-] = \kappa B[B_d \to (c\bar{c})K_S^0\pi^0] \propto \kappa \mid A \mid^2 \tag{5.5a}$$

$$B[\bar{B}_d \to (c\bar{c})K^+\pi^-] = 0 \tag{5.5b}$$

$$B[\bar{B}_d \to (c\bar{c})K^-\pi^+] = \kappa B[\bar{B}_d \to (c\bar{c})K_S^0\pi^0] \propto \kappa \mid \bar{A} \mid^2 \tag{5.5c}$$

$$B[B_d \to (c\bar{c})K^-\pi^+] = 0 \tag{5.5d}$$

where κ denotes the ratio of the observed branching ratios into the decays of the relevant B_d states into the two decay modes $(c\bar{c})K^+\pi^-$ and $(c\bar{c})K_S^0\pi^0$. This includes the factor six of eqs. (5.4) and any corrections due to additional differences in detection efficiencies for the charged and neutral final states.

For the decays of the oscillating $B_d - \bar{B}_d$ system into these modes

$$B[B_d(t) \to (c\bar{c})K^+\pi^-] = \kappa F e^{-\Gamma t}(1 + cos\Delta mt) \mid A \mid^2 \tag{5.6a}$$

$$B[B_d(t) \to (c\bar{c})K^-\pi^+] = \kappa F e^{-\Gamma t}(1 - cos\Delta mt) \mid \bar{A} \mid^2 \tag{5.6b}$$

$$B[\bar{B}_d(t) \to (c\bar{c})K^-\pi^+] = \kappa F e^{-\Gamma t}(1 + cos\Delta mt) \mid \bar{A} \mid^2 \tag{5.6c}$$

$$B[\bar{B}_d(t) \to (c\bar{c})K^+\pi^-] = \kappa F e^{-\Gamma t}(1 - cos\Delta mt) \mid A \mid^2 \tag{5.6d}$$

We note that a function of the decays to the neutral final states $(c\bar{c})K^+\pi^-$ and $(c\bar{c})K^-\pi^+$ has the same time dependence as the CP-violating interference terms in eqs. (5.1),

$$B_{B(c\bar{c})K^+\pi^-}(t) =$$

$$= \sqrt{\{B[B_d(t) \to (c\bar{c})K^+\pi^-] + B[\bar{B}_d(t) \to (c\bar{c})K^-\pi^+]\}}\cdot$$

$$\cdot\sqrt{\{B[B_d(t) \to (c\bar{c})K^-\pi^+] + B[\bar{B}_d(t) \to (c\bar{c})K^+\pi^-]\}} =$$

$$= \kappa F e^{-\Gamma t} \mid \sin(\Delta mt) \mid (\mid A \mid^2 + \mid \bar{A} \mid^2) \tag{5.7a}$$

Thus

$$\frac{B[\bar{B}_d(t) \to (c\bar{c})K_S^0\pi^0] - B[B_d(t) \to (c\bar{c})K_S^0\pi^0]}{B_{B_d(c\bar{c})K^+\pi^-}(t)} = \frac{4}{\kappa(1+\mid \rho_f \mid^2)} \cdot \frac{\sin(\Delta mt)}{\mid \sin(\Delta mt) \mid} \cdot \mathrm{Im}(\frac{p}{q}\rho_f) \tag{5.7b}$$

This relation gives the *CP* violation parameter $\mathrm{Im}(\frac{p}{q}\rho_f)$ as a time-independent constant function of the experimentally determined decays to the final states $(c\bar{c})K_S^0\pi^0$, $(c\bar{c})K^+\pi^-$ and $(c\bar{c})K^-\pi^+$ which are all themselves individually functions of time. A plot of the quantity (5.7b) as a function of time then gives the *CP* violation parameter $\mathrm{Im}(\frac{p}{q}\rho_f)$ by finding the best constant fit, without the necessity of fitting with a given time dependence and knowledge of the value of Δm. We also note that

$$\frac{\int g(t)B[\bar{B}_d(t) \to (c\bar{c})K_S^0\pi^0]dt - \int g(t)B[B_d(t) \to (c\bar{c})K_S^0\pi^0]dt}{\int g(t)B_{B(c\bar{c})K^+\pi^-}(t)dt}$$

$$= \frac{4}{\kappa(1+\mid \rho_f \mid^2)} \cdot \frac{\sin(\Delta mt)}{\mid \sin(\Delta mt) \mid} \cdot \mathrm{Im}(\frac{p}{q}\rho_f) \tag{5.7c}$$

where g(t) is any arbitrary weighting factor and the integrals are taken over an interval in which $\sin(\Delta mt)$ does not change sign. Thus in any experiment limited by statistics in which large time bins are used in accumulating data, and in which there may be variations in acceptances as a function of time, the value of $\mathrm{Im}(\frac{p}{q}\rho_f)$ is still obtained directly from eq. (5.7c) provided that the binning and acceptance variations are the same for all measurements. It is also possible to use weighting factors g(t) to optimize signal to noise, since the signal has a sinusoidal variation with time, which is absent in the noise.

References

1. I. I. Bigi and A. I. Sanda, *Nucl. Phys.* **B281**, 41 (1987).
2. Harry J. Lipkin and A. I. Sanda, *Physics Letters* **B201**, 541 (1988).
3. Harry J. Lipkin, *Nucl. Phys.* **B** (Proc. Suppl.) **13**, 491 (1990).
4. Harry J. Lipkin, in *Proceedings of the Workshop Towards Establishing a B Factory*, Edited by M. Goldberg and S. Stone, Syracuse University, Syracuse, New York, September 6-9 (1989), p. 149.
5. B. Kayser, M. Kuroda, R. D. Peccei and A. I. Sanda, *Phys. Lett.* **B237**, 508 (1990).
6. Particle Data Group, *Phys. Lett.* **B239**, 1 (1990).
7. A. Bohr, *Nucl. Phys.* **10**, 486 (1959).
8. A. Pais and S. B. Treiman, *Phys. Rev.* **D12** 2744, (1975).
9. Harry J. Lipkin, *Phys. Lett.* **B219** 474, (1989).

CP Violation Using non-*CP* Eigenstate Decays of Neutral *B* Mesons

R.Aleksan, I.Dunietz, B.Kayser and F.Le Diberder

Abstract

Neutral B meson decays into non-CP eigenstates are considered as a means for observing CP nonconservation. We show that sizable CP-violating asymmetries and clean measurements of the phases of the CKM matrix elements can be obtained using these decays, provided that they satisfy certain experimentally viable conditions. It is also argued that two meson final states, in which the quark content of one meson is the CP conjugate of the quark content of the second one, fulfill these requirements and are a particularly fertile and promising ground for studying CP violation. Two interesting examples, $B \to \rho^{\pm}\pi^{\mp}$ and $B \to a_1^{\pm}\pi^{\mp}$, are studied in detail. Their reconstruction efficiencies and their backgrounds are discussed for a generic asymmetric B Factory detector. It is found that, within this framework, those decay modes can give a better opportunity for observing CP violation and measuring the CKM phases than the $B \to \pi^+\pi^-$ decays.

1. Introduction

Observing CP violation in the b sector and confronting it with our present understanding[1] of this phenomenon is going to be one of the most challenging issues in particle physics during this decade. The experimental problem is compounded by the fact that CP-violating effects are small for those decay modes which have relatively large branching ratios, while channels in which potentially large effects are to be seen have small branching ratios[2]. Among this second class of channels, CP eigenstate final states such as $B^\circ \to \psi K_s^0$ [3] or $B^\circ \to \pi^+\pi^-$ [4] are the most promising but the expected number of such events at the future B factories running at the Υ(4S) is not very large. It is therefore very important to study as many B decay modes as possible in order to get the best chances to observe CP violation.

In this paper we shall concentrate on specific neutral B decays into final states that are not CP eigenstates. Section 2 describes how CP violation can be observed using these decay modes. In Section 3, we shall discuss what are the particular conditions that these final states should fulfill in order to be useful. The discussion will be illustrated with a typical and interesting example, $B^0 \to \rho^{\pm}\pi^{\mp}$. Finally in Section 4, the reconstruction efficiencies of this particular final state as well as for $B^0 \to a_1^{\pm}\pi^{\mp}$ and their backgrounds will be studied for a generic B factory detector.

2. CP VIOLATION USING B^0 DECAYS INTO NON-CP EIGENSTATES

In the standard model, every CP-violating effect is the result of an interference between different amplitudes for the same process. When a final state f can come both from a pure B^0 and a pure $\overline{B^0}$, the amplitudes for the direct decay $B^0 \to f$ and the mixing-induced sequence $B^0 \to \overline{B^0} \to f$ interfere. A time-dependent CP-violating effect can thus appear [5]. We shall focus in the following on B meson decay modes that can be reached from both a B^0 and $\overline{B^0}$ meson and that are not CP eigenstates. Examples of decays to such final states are given in the next section where they are studied in more detail. Let us for the moment write any such states as f, and suppose that n amplitudes contribute in the decay of a B^0 ($\overline{B^0}$) into f. The overall transition amplitude can be written as :

$$\langle f \mid T \mid \overline{B^0} \rangle \;=\; \sum_{j=1}^{n} e^{i\Phi_j}\, e^{i\alpha_j}\, M_j \;\equiv\; e^{i\beta}\, M \qquad (1.a)$$

$$\langle f \mid T \mid B^0 \rangle \;=\; \sum_{j=1}^{n} e^{-i\Phi'_j}\, e^{i\alpha'_j}\, M'_j \;\equiv\; e^{-i\beta'}\, M' \qquad (1.b)$$

Here, Φ_j and Φ'_j are the phases coming from the Cabibbo-Kobayashi-Maskawa (CKM) matrix and are at the origin of CP violation in the Standard Model, and α_j and α'_j are the phases due to the strong final state interactions. Finally, M_j

and M'_j are the magnitudes of the decay amplitudes. In equations (1), the terms with different j are meant to denote amplitudes with **different** CKM phases. Thus, distinct Feynman diagrams involving the same CKM phase are summed into a single amplitude. It should be noted that one could easily recover the case where f is a CP eigenstate by writing $\Phi'_j = \Phi_j$, $M'_j = M_j$ and $\alpha'_j = \alpha_j$ or $\alpha'_j = \alpha_j + \pi$ according to the nature of f (CP-even or CP-odd). We have used the convention $CP|B^0\rangle = +|\overline{B^0}\rangle$.

The time-dependent amplitudes for the decay of a neutral B meson created at the time $t_0 = 0$ as a pure $\overline{B^0}$ or as a pure B^0 are, respectively,

$$\mathcal{A}\left(\overline{B^0} \to f\right) =$$
$$e^{-\frac{\Gamma t}{2}} e^{-imt} \left[\cos\frac{\Delta mt}{2} \left\langle f|T|\overline{B^0}\right\rangle + i \sin\frac{\Delta mt}{2} e^{-2i\Phi_M} \langle f|T|B^0\rangle\right]$$

and

$$\mathcal{A}\left(B^0 \to f\right) =$$
$$e^{-\frac{\Gamma t}{2}} e^{-imt} \left[\cos\frac{\Delta mt}{2} \langle f|T|B^0\rangle + i \sin\frac{\Delta mt}{2} e^{2i\Phi_M} \left\langle f|T|\overline{B^0}\right\rangle\right] \quad (2)$$

We have defined :

$$\Gamma \equiv \frac{\Gamma_1 + \Gamma_2}{2} \quad , \quad m \equiv \frac{m_1 + m_2}{2} \quad , \quad \Delta m \equiv m_1 - m_2$$

where Γ_i and m_i are the width and the mass of the neutral B mass eigenstates, B_i. Φ_M is the CKM phase that appears in the mixing mechanism. We have assumed $\Delta\Gamma = \Gamma_1 - \Gamma_2 \ll \Gamma$ and $\Delta\Gamma \ll \Delta m$.

It is straightforward to derive the time-dependent decay probabilities using formulae (1) and (2) :

$$Pr\left(\overset{(-)}{B^0} \to f\right) \propto \left(\frac{A^2}{2}\right) e^{-\Gamma t}$$
$$\times \left\{1 \overset{(+)}{-} R \times \cos\Delta mt \overset{(+)}{-} D \sin\left(2\Phi_M + \beta + \beta'\right) \times \sin\Delta mt\right\}$$
$$(3)$$

where

$$R = \frac{\rho^2 - 1}{\rho^2 + 1}, \; D = \frac{2\rho}{\rho^2 + 1} = \sqrt{1 - R^2},$$

$$A^2 = M^2 + M'^2 \text{ and } \rho = \frac{M}{M'}.$$

Let us now focus on the CP conjugate final state, $\overline{f}$. The decay amplitudes for $B^0 \to \overline{f}$ and $\overline{B^0} \to \overline{f}$ are :

$$\langle \overline{f} \mid T \mid B^0 \rangle = \sum_{j=1}^{n} e^{-i\Phi_j} e^{i\alpha_j} M_j \equiv e^{-i\overline{\beta}} \, \overline{M} \qquad (1'.\text{a})$$

$$\langle \overline{f} \mid T \mid \overline{B^0} \rangle = \sum_{j=1}^{n} e^{i\Phi'_j} e^{i\alpha'_j} M'_j \equiv e^{i\overline{\beta}'} \, \overline{M}' . \qquad (1'.\text{b})$$

These are obtained from the amplitudes for $\overline{B^0} \to f$ and $B^0 \to f$ by applying a CP transformation, under which only the signs in front of Φ_j and Φ'_j change, since CP is conserved in the strong interactions. The decay probabilities are :

$$Pr\left(\overset{(-)}{B^0} \to \overline{f}\right) \propto \left(\frac{\overline{A}^2}{2}\right) e^{-\Gamma t}$$

$$\times \left\{1 \overset{(-)}{+} \overline{R} \times \cos\Delta mt \overset{(+)}{-} \overline{D} \sin\left(2\Phi_M + \overline{\beta} + \overline{\beta}'\right) \times \sin\Delta mt\right\} \qquad (3')$$

where

$$\overline{R} = \frac{\overline{\rho}^2 - 1}{\overline{\rho}^2 + 1}, \; \overline{D} = \frac{2\overline{\rho}}{\overline{\rho}^2 + 1} = \sqrt{1 - \overline{R}^2},$$

$$\overline{A}^2 = \overline{M}^2 + \overline{M}'^2, \text{and } \overline{\rho} = \frac{\overline{M}}{\overline{M}'}.$$

CP invariance demands that the probabilities of CP conjugate processes be identical. Thus, if CP invariance holds, we must have $Pr\left(\overline{B^0} \to \overline{f}\right) = Pr\left(B^0 \to f\right)$ and $Pr\left(\overline{B^0} \to f\right) =$

$Pr\left(B^0 \to \overline{f}\right)$. These relations must hold for the time-dependent rates and thus for the number of events corresponding to the time-integrated rates (where, for simplicity, equal production of B^0 and $\overline{B^0}$ is assumed) : $N\left(\overline{B^0} \to \overline{f}\right) = N\left(B^0 \to f\right)$ and $N\left(\overline{B^0} \to f\right) = N\left(B^0 \to \overline{f}\right)$. Thus, CP conservation requires :

$$\overline{M} = M \tag{4.a}$$

$$\overline{M}' = M' \tag{4.b}$$

$$S \equiv \sin(2\Phi_M + \beta + \beta') = -\overline{S} \equiv -\sin(2\Phi_M + \overline{\beta} + \overline{\beta}') \tag{4.c}$$

CP violation occurs if **any** of these three equations is not satisfied. Violation of either of the first two equations is known in the literature as direct CP violation [6]. The third equation tests CP violation generated by the interference of the direct decay $B^0 \to f$ and the mixing-induced decay $B^0 \to \overline{B^0} \to f$. From the ratios $N\left(\overline{B^0} \to f\right)/N\left(B^0 \to f\right)$ and $N\left(\overline{B^0} \to \overline{f}\right)/N\left(B^0 \to \overline{f}\right)$, and from the shapes of the four time-dependent decay rates given in equations (3) and (3′), one fits for the four unknowns, ρ, $\overline{\rho}$, S and $\overline{S}$. In addition, the ratio $A^2/\overline{A}^2$ can be extracted from the time-integrated measurements. This ratio is nothing else but $[N\left(\overline{B^0} \to f\right) + N\left(B^0 \to f\right)]/[N\left(\overline{B^0} \to \overline{f}\right) + N\left(B^0 \to \overline{f}\right)]$. Since $M'^2/\overline{M}'^2 = A^2(1+\overline{\rho}^2)/\overline{A}^2(1+\rho^2)$ and $M^2/\overline{M}^2 = M'^2/\overline{M}'^2 \times \rho^2/\overline{\rho}^2$, both equations (4.a) and (4.b) can be verified experimentally. Different values for ρ and $\overline{\rho}$ and/or different values for A^2 and $\overline{A}^2$ would show direct CP violation [7]. It would also mean that several amplitudes with different CKM phases and with different strong interaction phases have comparable magnitudes in the sums over j in equations (1) and (1′). This would complicate the measurement of the CKM phases and thereby obscure the origin of CP violation. It should be noted that finding that $\rho = \overline{\rho}$ and $A^2 = \overline{A}^2$ does not exclude the possibility (which, however, is unlikely) of having several competing amplitudes with different CKM phases but with almost equal strong interaction phases (see Appendix A).

However, let us now suppose that **one amplitude dominates** in the unmixed $\overline{B^0} \to f$ decay, and one in $B^0 \to f$ decay. In equations (1), we may then keep, say, the term with $j = 1$. By applying a CP transformation, we conclude that the same amplitudes will dominate in the CP conjugate decays. In such a scenario, one gets the following equalities :

$$\overline{M} = M \ , \ \overline{M}' = M' \tag{5.a}$$

$$\beta = \Phi_f + \alpha \ , \ \beta' = \Phi'_f - \alpha' \tag{5.b}$$

$$\overline{\beta} = \Phi_f - \alpha \ , \ \overline{\beta}' = \Phi'_f + \alpha' \tag{5.c}$$

Here, we have defined $\Phi_f \equiv \Phi_1$, $\alpha \equiv \alpha_1$, $\Phi'_f \equiv \Phi'_1$, and $\alpha' \equiv \alpha'_1$.

The four previous unknowns have now become only three: $\rho = \overline{\rho} = M/M'$, $S = \sin(2\Phi_M + \Phi_f + \Phi'_f + \Delta\alpha)$ and $\overline{S} = \sin(2\Phi_M + \Phi_f + \Phi'_f - \Delta\alpha)$, where $\Delta\alpha = \alpha - \alpha'$. The sine of the CP violating phase $2\Phi \equiv 2\Phi_M + \Phi_f + \Phi'_f$ can be extracted from the measurements of S and $\overline{S}$, up to a four-fold ambiguity, using :

$$\sin^2 2\Phi = \frac{1}{2}\left[1 + S\overline{S} \pm \sqrt{(1 - S^2)(1 - \overline{S}^2)}\right] \tag{6}$$

Here, one of the signs on the right hand side gives the true $\sin^2 2\Phi$, while the other gives $\cos^2 \Delta\alpha$. How do we resolve this ambiguity? Consider modes where a theoretical bias for $\Delta\alpha = 0$ or π exists, and where one solution is measured to be compatible with 1. Then, one could attempt to resolve the ambiguity by identifying $\cos^2 \Delta\alpha$ with that solution. Furthermore, one can always use other decay modes (such as CP eigenstates) which yield perhaps a less precise but unambigu-

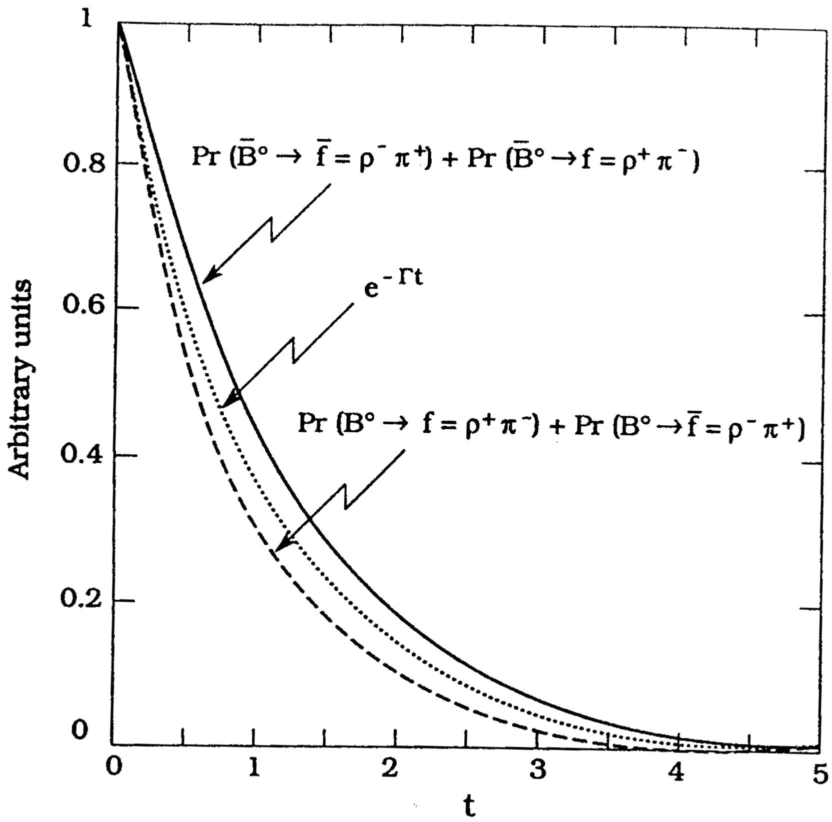


Figure 1. *Evolution, in proper time units, of the decay probabilities for the sum of the reactions $B^o \to f$ and $B^o \to \bar{f}$ on one hand, and reactions $\overline{B^o} \to \bar{f}$ and $\overline{B^o} \to f$ on the other hand. Here $f = \rho^+\pi^-$ and $\bar{f} = \rho^-\pi^+$ are chosen as an example.*

ous value for $\sin^2 2\Phi$. Different time-dependent distributions for the $B^o \to f$ decay and its *CP* conjugate $\overline{B^o} \to \bar{f}$ decay, or for the $B^o \to \bar{f}$ decay and its *CP* conjugate $\overline{B^o} \to f$ decay, would demonstrate in a model independent way that *CP* is not conserved in the *B* meson decays. *CP* violation can be simply exhibited experimentally by adding both final states, f and $\bar{f}$. From (3) and (3′) one gets (when one amplitude dominates) :

$$Pr\left(\overset{(-)}{B^o} \to f\right) + Pr\left(\overset{(-)}{B^o} \to \bar{f}\right) \propto A^2 e^{-\Gamma t}$$

$$\times \left\{1 \overset{(+)}{-} D \cos\Delta\alpha \sin\left(2\Phi_M + \Phi_f + \Phi'_f\right) \sin\Delta m t\right\} \quad (7)$$

Since the $\sin\Delta m t$ term in these formulae distinguishes between B^o and $\overline{B}^o$ decays, this term violates *CP* conservation. We show in Figure 1 this *CP* violating asymmetry for a particular set of values for ρ, $\sin(2\Phi_M + \Phi_f + \Phi'_f)$ and $\Delta\alpha$. At this stage, one can make the following remarks.

1. The asymmetry between B^0 and $\overline{B^0}$ decay probabilities will be larger as M and M' tend to be equal. However the asymmetry is not diluted dramatically as the ratio M/M' differs sensibly from 1 (see Figure 2).

2. If the B mesons are pair-produced at the $\Upsilon(4S)$, t has to be replaced, in $\sin\Delta mt$ and $\cos\Delta mt$, by $\Delta t = t_2 - t_1$ and by $\Sigma t = t_1 + t_2$ in $e^{-\Gamma t}$ (note that, after integration over Σt, the exponential term is simply $e^{-\Gamma|\Delta t|}$). Here, t_1 and t_2 are the times at which the B^0 and the $\overline{B^0}$ are decaying. One of these decays is the final state f of $\bar{f}$, while the other one is used to tag the flavor of the parent B. It is therefore necessary to know, at least, which of the B mesons has decayed first in order to observe the *CP* violation effect exhibited by equations (7) (*i.e.* through the $B^0\bar{B}^0$ mixing). This can be achieved at an asymmetric B factory [3,8]. Furthermore, after integrating over Σt and ignoring the time order (i.e. using $|\Delta t|$ instead of Δt) equations (3) and (3′) yield

$$Pr\left(\overset{(-)}{B^0}\rightarrow f\right) \propto \left(\frac{A^2}{2}\right) e^{-\Gamma|\Delta t|} \times \left\{1 \overset{(+)}{-} R \times \cos\Delta m|\Delta t|\right\} \tag{8}$$

$$Pr\left(\overset{(-)}{B^0}\rightarrow \bar{f}\right) \propto \left(\frac{\overline{A}^2}{2}\right) e^{-\Gamma|\Delta t|} \times \left\{1 \overset{(-)}{+} \overline{R} \times \cos\Delta m|\Delta t|\right\} \tag{8'}$$

Thus without the assumption that one amplitude dominates, one can probe direct *CP* violation in a simple way by measuring independently ρ and $\overline{\rho}$ either from the shapes and/or the relative normalisations of equations (8) and (8′) respectively (Figure 3).

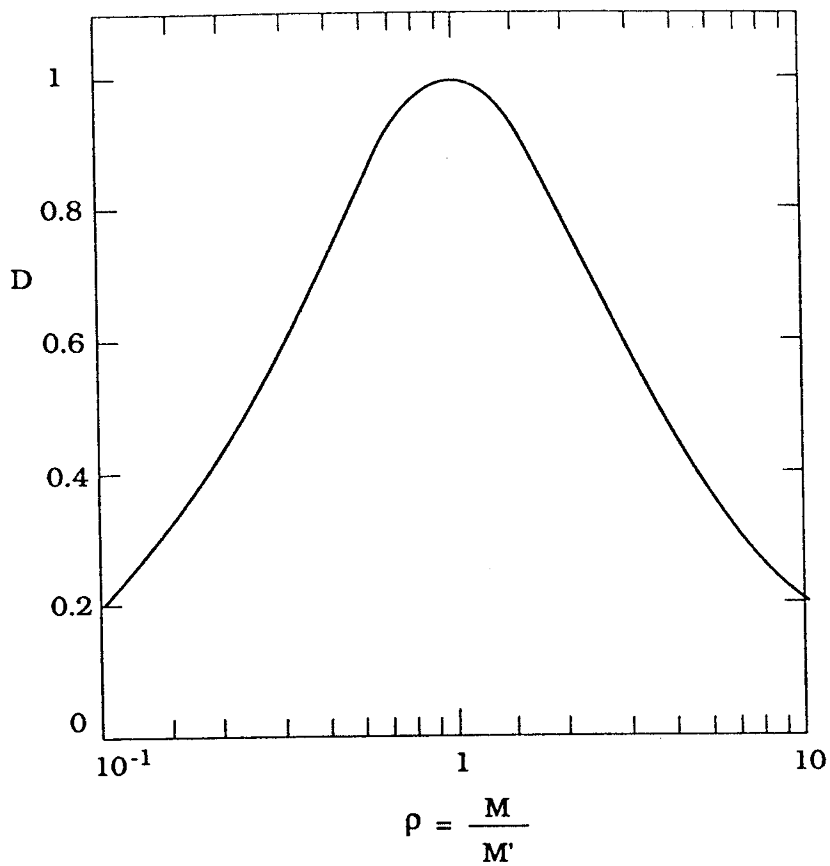


Figure 2. *The CP violating asymmetry dilution factor D as a function of ρ.*

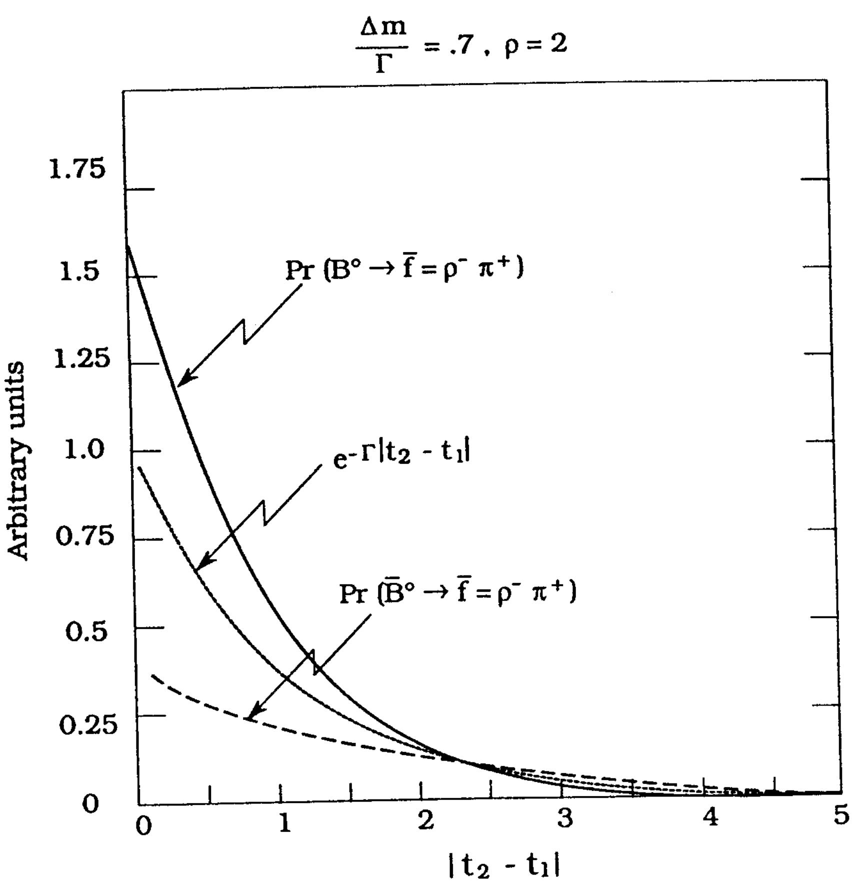


Figure 3. *Evolution, in proper time units, of the decay probabilities for the reactions* $\overline{B^o} \rightarrow \bar{f}$ *and* $B^o \rightarrow \bar{f}$*. Here,* $\bar{f} = \rho^- \pi^+$ *is chosen as an example.*

3. Decays to CP-Self-Conjugate Collections of Quarks

Suppose that the final state f is a pair of mesons a_1 and $\overline{a}_2$ in which the two mesons differ in J^{PC} or mass, so that $a_1\overline{a}_2$ is not a CP eigenstate. Suppose, however, that at the quark level, $a_1\overline{a}_2 = \left(q_x\overline{q}_y\right)\left(\overline{q}_x q_y\right)$, so that f is made up of a CP-self-conjugate collection of quarks and antiquarks (for example $f = D^{*+}D^-, \rho^+\pi^-, a_1^+\pi^-, K^{*+}K^-$ ...) . Then, as we shall argue, f is sufficiently close to being a CP eigenstate that CP violation in $\overset{(-)}{B}{}^o \to f$ is expected to be large.

As has been said in Section 2, CP violation is the result of several interfering amplitudes. If the interference is to be a large effect, then the magnitudes M and M' of the interfering $\overline{B^0}$ and B^0 decay amplitudes in equations (2) must be of comparable size. The quantity D in the decay rates will then be near its maximum possible value. Now as pointed out some time ago[5,9], when f is a CP eigenstate and $\left|\langle f|T|\,B^0\rangle\right|$ is dominated by one CKM factor, M and M' are of equal size. This is one of the reasons why decays to CP eigenstates are of particular interest.

When f is not a CP eigenstate, but is a final state of the type $a_1\overline{a}_2 = \left(q_x\overline{q}_y\right)\left(\overline{q}_x q_y\right)$, the CP-self-conjugacy of the quark content has the consequence that for each diagram for $B^0 \to a_1\overline{a}_2$, there is a related one for $\overline{B^0} \to a_1\overline{a}_2$ which is just the CP conjugate of the $B^0 \to a_1\overline{a}_2$ diagram except for the way in which the final quarks are joined up to make hadrons. This is illustrated in Figures 4 and 5. Because the $B^0 \to a_1\overline{a}_2$ diagram and its $\overline{B^0} \to a_1\overline{a}_2$ partner are almost CP conjugates, one expects that in general these two diagrams are of comparable size. If some single diagram dominates $B^0 \to a_1\overline{a}_2$, then its "almost-$CP$-conjugate" partner is expected to dominate $\overline{B^0} \to a_1\overline{a}_2$, and the amplitudes for these two decays should be of comparable size, as desired*. As discussed earlier, when

* Using the heavy quark techniques[10,11] in addition to factorization, it has been shown[12,13] that the diagram which dominates $B^0 \to D^{*-}D^+$ and the "almost-CP-conjugate" one which dominates $\overline{B^0} \to D^{*-}D^+$ are of equal size. Furthermore, since only one universal form factor governs the $\bar{B}_d \to D^*$ and $\bar{B}_d \to D$

one diagram dominates $B^o \to f$ and one dominates $\overline{B^o} \to f$, equations (5) apply, and the weak phase $2\Phi = 2\Phi_M + \Phi_f + \Phi'_f$ can be extracted from the data using equation (6). Since the "almost-*CP*-conjugate" diagrams dominating $B^o \to a_1\overline{a}_2$ and $\overline{B^o} \to a_1\overline{a}_2$ have equal and opposite weak phases (see, for example, Figure 5) $\Phi'_f = \Phi_f$.

A promising example of the final states of the type $(q_x\overline{q}_y)(\overline{q}_x q_y)$ is $\rho^{\pm}\pi^{\mp}$. The CKM phase involved in this final state is the same as the one involved in the $\pi^+\pi^-$ mode. This latter mode was considered, up to now, as the practical candidate for measuring this particular phase. Since the branching ratio for $B^0 \to \pi^+\pi^-$ is expected to be small ($\sim 2 \times 10^{-5}$),

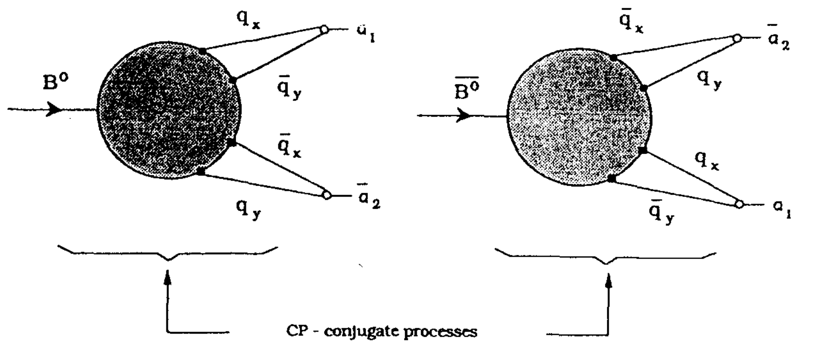


Figure 4. *Related diagrams for $B^o \to a_1\overline{a_2}$ and $\overline{B^o} \to a_1\overline{a_2}$. Except for the way in which the final quarks are joined up to make hadrons, the diagrams are CP conjugates of each other.*

it is important to find other competitive channels. This is one of the reasons why we are focussing our study on $\rho^{\pm}\pi^{\mp}$ (and $a_1^{\pm}\pi^{\mp}$, which is a similar mode) in the following. The diagrams expected to dominate the decays into $\rho^+\pi^-$ are the $\overset{(-)}{b}$ quark decay diagrams shown in the top row of Figure 5 (comments on the complications that arise from other diagrams will be made in Appendix A). Henceforth, we assume that the $\overset{(-)}{b}$ decay diagram dominates and that the rescattering processes such as $B_d^o \to D^{*-}D^+ \to \rho^-\pi^+$ are negligible. Let us discuss the ρ parameter. Under the previous assumptions, the factorization approximation[14] ought to hold for the $\rho^{\pm}\pi^{\mp}$ and $\pi^+\pi^-$ decay modes. One predicts that

$$\frac{\Gamma(\overline{B}_d^o \to \rho^-\pi^+)}{\Gamma(\overline{B}_d^o \to \pi^+\pi^-)} \approx \left|\frac{F_\rho}{f_\pi}\right|^2 \times \left|\frac{f_+(m_\rho^2)}{f_+(m_\pi^2)}\right|^2$$
$$= \left|\frac{F_\rho}{f_\pi}\right|^2 = \left(\frac{190MeV}{132MeV}\right)^2 \simeq 2$$

transitions and since the D and D^* decay constants are related, no difference in final state phases is expected, i.e. $\cos^2\Delta\alpha = 1$. Those modes, then, would measure the same CKM phase as D^+D^- and ψK_s, and complement them.

Here, the ρ^- meson is made from the virtual W boson, F_ρ and f_π are the decay constants of the ρ and the pion, and the relevant vector form factor for the $\overline{B}_d^o \to \pi^+$ transition is denoted by f_+. Only factorization is required to obtain the quoted numerical value for this ratio, since the decay constant F_ρ is measured from $\tau \to \rho\nu_\tau$ decay, and the vector form factor f_+ ought to be the same for the two light scales $q^2 = m_\rho^2$ and $q^2 = m_\pi^2$.

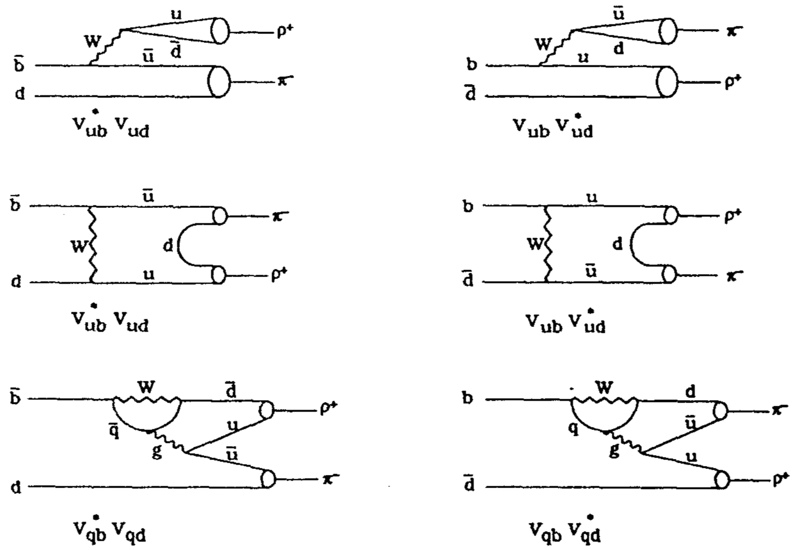


Figure 5. . *Diagrams for $B^0 \to \rho^+\pi^-$ and $\overline{B}^0 \to \rho^+\pi^-$. In the penguin diagrams (bottom row), q runs over all the positively charged quarks, t, c, and u. The product of CKM elements to which each diagram is proportional is written beneath it. All the products in this figure are of order λ^3, where λ is the Wolfenstein parameter. Replacing the quark $u(\overline{u})$ by the quark $c(\overline{c})$ wherever it is written explicitly, leads to another interesting final state, namely $D^{*+}D^-$.*

In contrast, at present we cannot predict reliably the ratio $\dfrac{\Gamma(\overline{B}_d^o \to \rho^+\pi^-)}{\Gamma(\overline{B}_d^o \to \pi^+\pi^-)}$ and therefore $\dfrac{\Gamma(\overline{B}_d^o \to \rho^+\pi^-)}{\Gamma(\overline{B}_d^o \to \rho^-\pi^+)} = \dfrac{M^2}{M'^2} = \rho^2$ will be kept as unknown in the following. However, using many additional assumptions, we estimate the ratio $\Gamma(\overline{B}_d^o \to \rho^+\pi^-)/\Gamma(\overline{B}_d^o \to \pi^+\pi^-)$ as being either 10 ± 4 or 31 ± 7 in Appendix B.

While we have allowed for final state phases, we believe that there are no such phases involved in the decays of a heavy neutral B meson into two charged light particles. In the factorization approximation, the $q\overline{q}$ system issued from the virtual W is created at the same space-time point. Because of its pointlike creation, the $q\overline{q}$ pair starts out as a color singlet without any color dipole moment. It is a light system which flies off with a large Lorentz boost factor γ [15]. Only after having flown a distance of a few Fermis from its creation point does the $q\overline{q}$ system develop a sizeable color dipole moment through which strong interaction could have been felt. But by that time the $q\bar{q}$ system is too far away from the other quark pair to interact appreciably with this other light particle. Furthermore, this latter particle has also travelled a sizeable distance in the opposite direction. In short, no final state interactions are expected and, thus, **no final state interaction phases.**

We have seen that the ratio of rates

$$\overline{B^0} \to \rho^-\pi^+/\overline{B^0} \to \pi^+\pi^- = B^0 \to \rho^+\pi^-/B^0 \to \pi^+\pi^- \simeq 2$$

is theoretically well-established and not very model-dependent. With this in mind, let us study the error on the measurement of sin2Φ using the decay mode $B \to \rho^\pm\pi^\mp$ as compared to $B \to \pi^+\pi^-$, where B stands for the sum of the B^0 and $\overline{B^0}$ mesons. This error, $\sigma[\sin 2\Phi]$, depends on the value of the CP violating asymmetry (cf. equation 7) and the total number of available events. Therefore, neglecting for the moment the relative selection efficiencies and backgrounds, one gets :

$$\sigma[\sin 2\Phi]_{\rho^\pm\pi^\mp}/\sigma[\sin 2\Phi]_{\pi^+\pi^-} = \sqrt{N_{\pi^+\pi^-}/N_{\rho^\pm\pi^\mp}} \times 1/d \quad (10)$$

where $N_{\pi^+\pi^-}$ and $N_{\rho^\pm\pi^\mp}$ are the numbers of observed $B \to \pi^+\pi^-$ and $B \to \rho^\pm\pi^\mp$ decays, respectively, and $d = [2\rho/(\rho^2 + 1)] \times \cos\Delta\alpha$ is the dilution factor for $B \to \rho^\pm\pi^\mp$ (this factor is 1 for $B \to \pi^+\pi^-$). If we suppose, as we have argued previously, that $\Delta\alpha = 0$ and acknowledging that

$$N_{\rho^\pm\pi^\mp} \;\propto\; A^2 = M'^2(\rho^2 + 1) \propto$$

$$\Gamma(\overline{B^0} \to \rho^-\pi^+)(\Gamma(\overline{B^0} \to \rho^+\pi^-)/\Gamma(\overline{B^0} \to \rho^-\pi^+) + 1) \quad (11)$$

we obtain :

$$\begin{aligned}\sigma[\sin 2\Phi]_{\rho^\pm\pi^\mp}/\sigma[\sin 2\Phi]_{\pi^+\pi^-} &= \sqrt{(\rho^2+1)/4\rho^2} \times \\ &\sqrt{\Gamma(\overline{B^0} \to \pi^+\pi^-)/\Gamma(\overline{B^0} \to \rho^-\pi^+)} \\ &\simeq \sqrt{(\rho^2+1)/8\rho^2} \qquad (12)\end{aligned}$$

We show this ratio as a function of the value of ρ^2 in Figure 6. It can be seen from this figure that the error on

$\sin 2\Phi$ as measured in the $B \to \rho^{\pm}\pi^{\mp}$ channel is smaller than the one obtained from the $B \to \pi^+\pi^-$ channel as long as $\rho^2 \equiv \Gamma(\overline{B^0} \to \rho^+\pi^-)/\Gamma(\overline{B^0} \to \rho^-\pi^+)$ is larger than 1/7.

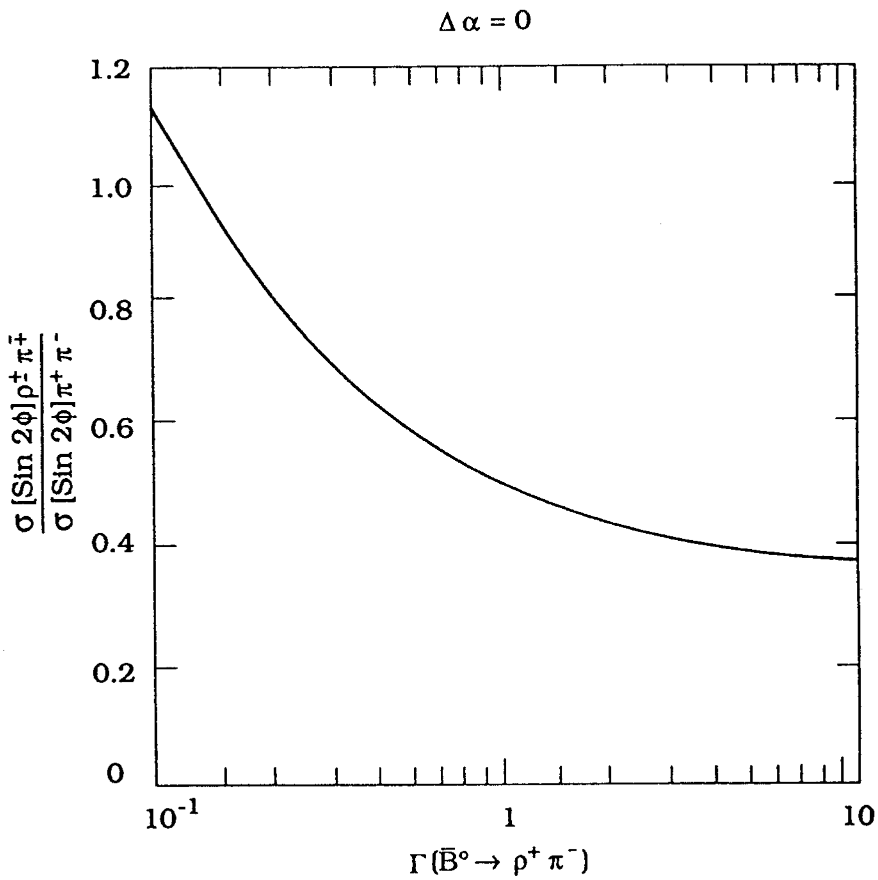


Figure 6. *The expected error on $\sin 2\Phi$ using the four reactions $B^0(\overline{B^0}) \to \rho^{\pm}\pi^{\mp}$ relative to the error obtained using the two reactions $B^0(\overline{B^0}) \to \pi^+\pi^-$ as a function of $\Gamma(\overline{B^0} \to \rho^+\pi^-)/\Gamma(\overline{B^0} \to \rho^-\pi^+)$.*

4. Analysis

We now address the issue of the experimental feasibility of using the $\rho^{\pm}\pi^{\mp}$ and the $a_1^{\pm}\pi^{\mp}$ decays of the B mesons to study CP violation at an asymmetric e^+e^- collider whose center of mass energy is set to the $\Upsilon(4S)$ mass.

The $\Upsilon(4S)$ decay into a $B_d^0\overline{B_d^0}$ pair involves a P-wave final state which ensures a definite correlation between the decay times of the two mesons[2]. Having observed at time t_1 the decay of one of the B mesons into a final state tagging the flavor of a $\overset{(-)}{B}{}^o$, and the decay of the second B into the final state f at time t_2, the time difference $\Delta t = t_2 - t_1$ follows the distribution given by equations 3, where t is replaced by Δt in the sine and cosine terms and by $|\Delta t|$ in the exponential term. (We assume that $t_1 + t_2$ is not measured.)

The use of an asymmetric e^+e^- collider is prompted by the need of a moving $\Upsilon(4S)$ system in the laboratory frame in order to facilitate the measurement of Δt. More precisely the determination of sin2Φ is performed on the basis of the event distribution with respect to Δz, the distance along the beam axis between the vertex of the $B_d^0(\overline{B_d^0}) \rightarrow f$ disintegration and the decay vertex of the tagged B_d^0 or $\overline{B_d^0}$. The two quantities are related according to

$$\Delta z = \gamma\gamma^{cm}\beta c\Delta t + \gamma\gamma^{cm}\beta^{cm}c\cos\theta_2^{cm}(t_2 + t_1) \qquad (13)$$

where c is the velocity of light and

1. β and γ define the boost which links the $\Upsilon(4S)$ rest frame to the laboratory frame. For the high energy beam of the machine we assume in the following $E_b(e^-) = 9\ GeV$, which gives $\beta = 0.49$ and $\gamma = 1.14$.
2. $\beta^{cm} = 0.066$ and $\gamma^{cm} \simeq 1$ define the boost which links the B rest frames to the $\Upsilon(4S)$ rest frame.
3. θ_2^{cm} is the polar angle of the B decaying at time t_2 in the $\Upsilon(4S)$ rest frame.

Although the second term in equation (13) spoils the ideal one-to-one correspondence between Δz and Δt that one would obtain with $\beta^{cm} = 0$, its impact on the sin2Φ measurement turns out to be very small. This results from $< \cos\theta_2^{cm} >= 0$ and from the $\sin^2\theta_2^{cm}$ angular distribution which favors small values of $\cos\theta_2^{cm}$. Moreover, the f final state being completely reconstructed, the value of θ_2^{cm} is available and can be used to limit even further the dilution of the information. A detailed analysis shows that the increase on the resolution σ[sin2Φ] induced by the non zero value of β^{cm} is given by the factor

$$\Sigma_{\beta^{cm}} = 1 + (\frac{\beta^{cm}}{\beta})^2 \frac{2x^2+1}{20} \qquad (14)$$

for $\beta^{cm}x \ll \beta$, where $x = \Delta m/\Gamma$ is the B_d^0 mixing parameter. For $x = 0.7$ one gets the negligible correction $\Sigma_{\beta^{cm}} = 1.002$. Hence, one can safely ignore the non zero value of β^{cm} and use the approximate relation $\Delta z = \beta\gamma c\Delta t$.

Using $\tau_d = 1.18 \times 10^{-12}$s for the B_d lifetime, the average longitudinal distance between the two B decays is $L_{lab} \leq \Delta z >= \beta\gamma c\tau_d \simeq 200\ \mu$m. For a typical[3] detector resolution $\sigma[\Delta z] = 55\ \mu$m, the inequality $L_{lab} >> \sigma[\Delta z]$ is not satisfactorily satisfied. However, the quantity which governs the measurability of the CP-violating effect is not L_{lab} but the wavelength of the CP violating component of the Δz-distribution, $\lambda_{CP} = \frac{2\pi}{x} L_{lab} \simeq 1800\ \mu$m. For a 9 GeV high energy beam (but also for a smaller machine asymmetry) the relevant inequality $\lambda_{CP} >> \sigma[\Delta z]$ is satisfied. Hence, the quality of the vertex resolution of the detector is not critical for an experiment at an asymmetric collider. Accordingly, the algorithm used to measured Δz will not be discussed (details can be found in Ref. 3).

For small values of sin2Φ (≤ 0.5) and assuming that $\cos\Delta\alpha \approx 1$, equation (6) simplifies to :

$$\sin^2 2\Phi = \left((S + \overline{S})/2\cos\Delta\alpha\right)^2 \tag{15}$$

with $\quad \cos\Delta\alpha \simeq \sqrt{1 - \left((S - \overline{S})/2\right)^2}.$

The measurement of sin2Φ is obtained from the shape of the four distributions (3 and 3′) using the likelihood method. The overall resolution is :

$$\sigma^2[\sin 2\Phi] \simeq \frac{1+4x^2}{2x^2}\, d^{-2}\, \frac{1+r\eta_z}{N}\, (1-2\omega)^{-2} \times e^{\sigma_t^2 (x^2 + \frac{1}{2})} \times [1 + \mathcal{O}(\sin^2 2\Phi)] \tag{16}$$

where

- $d = \dfrac{2\rho}{\rho^2+1} \cos\Delta\alpha$ is the dilution factor defined in section three.
- N is the total number of selected $B_d^0 \overline{B_d^0}$ events.
- r is the ratio of the number of selected background events over the number of selected signal events.
- η_z is an effective rejection factor resulting from the shape differences between the Δz distributions of background and signal events. The η_z value depends on the

resolution $\sigma[\Delta z]$. For example, with $\sigma[\Delta z] = 55\ \mu$m, one has $\eta_z \simeq 0.20$ for the continuum $q\overline{q}$ background. The conservative value $\eta_z = 0.5$ is used below.

- ω represents the fraction of $B_d^0\overline{B_d^0}$ events where the flavor of the tagged B is wrongly identified.
- The exponential term accounts for the loss in resolution due to the finite vertex resolution, expressed in unit of L_{lab} : $\sigma_t = \sigma[\Delta z]/L_{lab} \simeq 0.3$, with our numerical values. For $x = 0.7$, the vertex resolution correction term on $\sigma[\sin 2\Phi]$ is then very small ; 4%.

The contribution to $\sigma^2[\sin 2\Phi]$ of the measurement error on ρ enters in the $\mathcal{O}(\sin^2 2\Phi)$ term of equation (16). For $x = 0.7$, $\sigma[\rho]/\rho$ goes from $\sim 4\%$ when $\rho = 1$ to $\sim 12\%$ when $\rho = 5$ (for $N = 10^3$ events).

The study presented below was carried out using the Monte-Carlo program simulating a generic B detector as described in Ref. 4. The most relevant detector characteristics assumed in the simulation are the following. The tracking device consists of a drift chamber completed with a silicon vertex detector. The transverse momentum resolution is $\sigma(p_\perp)/p_\perp^2 = \left(0.23^2 + 0.27^2/(p\beta_c)^2(p_z/p_\perp)^3\right)^{1/2}$ (% GeV^{-1}), where $p_\perp$, p, and p_z are the transverse, total and longitudinal momenta of the charged particle and β_c is its velocity. The vertex resolution in the longitudinal direction is $\sigma = \left(17^2 + 25^2/(p\beta_c)^2(p_z/p_\perp)^5\right)^{1/2}$ (μm). The calorimeter is made of 18 radiation-length CsI crystals. The angular and energy resolutions for the electromagnetic showers are $\sigma(\theta) = \sigma(\phi) = 10$ mrad and $\sigma(E)/E = \left((0.02/E^{1/4})^2 + 0.01^2\right)^{1/2}$ (E in GeV), respectively. Both the tracking device and the calorimeter cover the polar angle region $|\cos\theta| < 0.95$. For the sake of simplicity the particle identification capabilities are assumed to be perfect in this analysis. More details can be found in Ref. 4.

Although the analysis implies the use of $\rho^+\pi^-$ together with $\rho^-\pi^+$ events and $a_1^+\pi^-$ together with $a_1^-\pi^+$ events, as a matter of convention we refer to these final states as $\rho^-\pi^+$ and $a_1^-\pi^+$ in the following. In distinction to the notation used in the previous sections, we will denote the generic final states

$\rho^\pm\pi^\mp$ and $a_1^\pm\pi^\mp$ by f in this section. In both final states, the observed signal consists of 4 particles; 2 charged pions and 2 photons for the $\rho^-\pi^+$ mode, 4 charged pions for the $a_1^-\pi^+$ mode. In the latter case, we have restricted the analysis to the charged-pion decay of the a_1 ($a_1^- \to \pi^-\rho^0 \to \pi^-\pi^+\pi^-$) because events from the single charged-pion decay of the a_1 involve two π^0 and are much more difficult to extract experimentally. Both final states involve a π^+ with a well defined high-momentum in the Υ rest frame : $P_{\pi^+} \simeq M_{\Upsilon(4S)}/4$. Because of this particularity, the background coming from B decays is very effectively suppressed. The most important background to these final states is due to continuum $q\overline{q}$ events, which, accordingly, are the only ones considered below.

In addition to the previously quoted values, the following numbers have been used in our analysis :

- $q\overline{q}$ continuum cross-section at the $\Upsilon(4S)$ center-of-mass energy : $\sigma_{q\overline{q}} = 2.7\ nb$.
- $e^+e^- \to \Upsilon(4S)$ cross-section : $\sigma_{\Upsilon(4S)} = 1.2\ nb$.
- Branching ratios
  1) $\Upsilon(4S)$ into $B_d^0\overline{B_d^0}$ pairs : $B(B_d^0) = 50\%$.
  2) B_d^0 into $\pi^-\pi^+$: $B(\pi^-\pi^+) = 2 \times 10^{-5}$.
  3) B_d^0 into $\rho^-\pi^+$ or $\rho^+\pi^-$: $B(\rho^\mp\pi^\pm) = 6 \times 10^{-5}$.
  4) B_d^0 into $a_1^-\pi^+$ or $a_1^+\pi^-$: $B(a_1^\mp\pi^\pm) = 6 \times 10^{-5}$.
  5) $a_1^\mp$ into $\rho^0\pi^\mp$: 50%.
- Dilution factor : $d = 0.94$ ($\rho = 1/\sqrt{2}$ and $\Delta\alpha = 0$).

The quoted values for $B(\rho^\mp\pi^\pm)$ and $B(a_1^\mp\pi^\pm)$ are chosen as an example. The dependence of $\sigma[\sin 2\Phi]$ on these branching ratios is presented at the end of this section.

The following quantities are used to select the $\rho^-\pi^+$ and $a_1^-\pi^+$ final states:

- $\delta W_f = \mid W_f - M_B \mid$, where W_f is the invariant mass formed by the four particle 4-momenta, and M_B is the B mass. The experimental resolutions on W_f are $\sigma[W_f] = 31\ MeV$ for the $\rho^\pm\pi^\mp$ events and $\sigma[W_f] = 17\ MeV$ for the $a_1^\pm\pi^\mp$ events. The larger $\rho^\pm\pi^\mp$ resolution results from the angular and energy resolutions attached to the two photons of the π^0 decay which are significantly poorer than the ones of the charged pions.

- $\delta\overline{W}_f = \mid \overline{W}_f - M_B \mid$ with $\overline{W}_f^2 = M^2_{\Upsilon(4S)}/4 - \mathrm{p}^2_{cm}$ and where $\mathbf{p}_{cm}$ is the total 3-momentum of the four particles in the $\Upsilon(4S)$ rest frame. The experimental resolutions on $\overline{W}_f$ are essentially determined by the beam energy spread of the machine (assumed to be $\sigma[E_b(e^\pm)] = 7 \times 10^{-4} \times E_b(e^\pm)$). The resolutions are therefore almost the same, $\sigma[\overline{W}_f] \simeq 3\ MeV$, for both the $\rho^\pm\pi^\mp$ events and $a_1^\pm\pi^\mp$ events.

- $\delta\rho = \mid W_{\pi\pi} - m_\rho \mid$, where m_ρ is the ρ mass and where $W_{\pi\pi}$ is the invariant mass of the two pions which should originate from a ρ decay : $\rho^- \rightarrow \pi^-\pi^0$, for the $\rho^\pm\pi^\mp$ events and $\rho^0 \rightarrow \pi^+\pi^-$, for the $a_1^\pm\pi^\mp$ events.

- $\delta a_1 = \mid W_{\pi^-\rho^0} - m_{a_1} \mid$ (defined for the $a_1^-\pi^+$ final state), where m_{a_1} is the a_1 mass and where $W_{\pi^-\rho^0}$ is the invariant formed by the π^- and a pair of charged pions attributed to a ρ^0 decay.

- $\delta\pi^0 = \mid W_{\gamma\gamma} - m_{\pi^0} \mid$ (defined for the $\rho^-\pi^+$ final state), where m_{π^0} is the mass of the π^0 and where $W_{\gamma\gamma}$ is the invariant mass formed by a photon pair, candidate for a π^0 decay. The experimental resolution is $\sigma[W_{\gamma\gamma}] = 7.5\ MeV$. The typical value of the smallest of the two photon energies is $E^\gamma_{min} \simeq 0.4\ GeV$ (cf Figure 7) and is well within the expected capabilities of future B factory detectors.

- P^v_f the χ^2 probability of the fit constraining the charged particles of the final state f to share the same vertex. This quantity is particularly useful in the case of the $a_1^\pm\pi^\mp$ final state, where four charged particles are involved. For example, the cut $P^v_f > 0.05$ rejects 45% of the $q\bar{q}$ background to this final state.

- N_c the number of charged particles detected in addition to the charged particles of the final state f.

- $\cos\theta^{cm}_+ = \mid \vec{u}^+ . \vec{u}_{sph} \mid$, where $\vec{u}^+$ is the unit vector along the direction of the high-energy π^+ and where $\vec{u}_{sph}$ is a unit vector pointing along the sphericity axis of the final state (the four particles of the final state f are excluded). For $q\bar{q}$ background events, the high energy π^+ direction tends to be close to the sphericity axis ($\cos\theta_+^{cm} \simeq 1$), while the two directions are uncorrelated for the signal events.

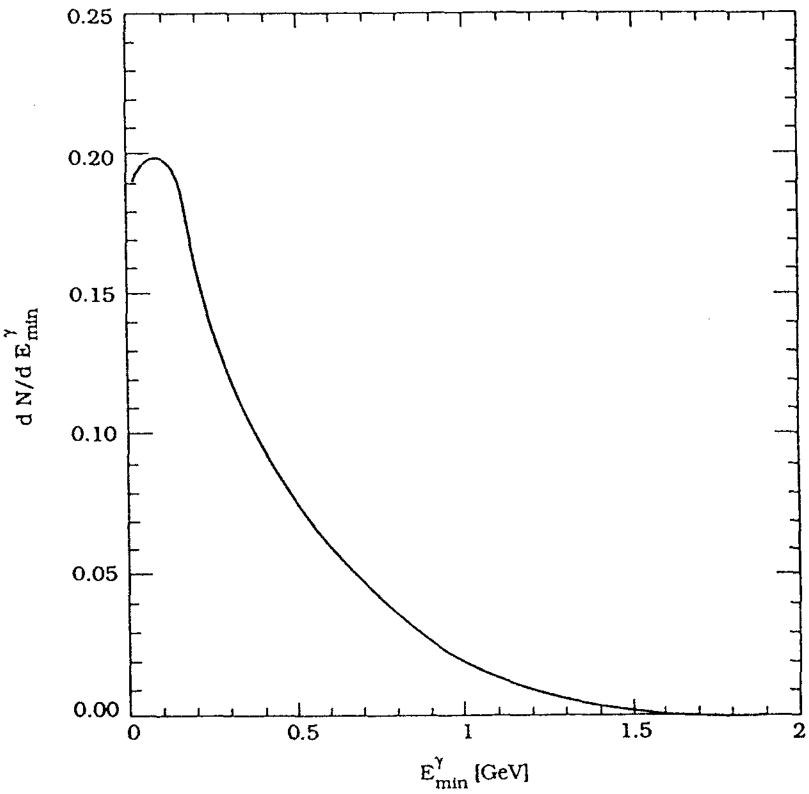


Figure 7. *Energy spectrum of the softest photon issued from the cascade decay $B^0(\overline{B^0}) \to \rho^\pm\pi^\mp \to \pi^+\pi^-\pi^o \to \pi^+\pi^-\gamma\gamma$.*

The preselection is obtained by applying the following cuts :

$\rho^-\pi^+$:

$$\delta W_f < 100 \text{ MeV} \;,\; \delta\overline{W}_f < 10 \text{ MeV} \;,\; \delta_\rho < 300 \text{ MeV}$$
$$\delta_{\pi^0} < 20 \text{ MeV} \;,\; P_f^v > 0.05 \;,\; N_c \geq 3$$

$a_1^-\pi^+$:

$$\delta W_f < 50 \text{ MeV} \;,\; \delta\overline{W}_f < 10 \text{ MeV} \;,\; \delta_\rho < 300 \text{ MeV}$$
$$\delta_{a_1^-} < 660 \text{ MeV} \;,\; P_f^v > 0.05 \;,\; N_c \geq 3$$

The overall efficiencies of these cuts are $\epsilon_{\rho\pi} = (58.5 \pm 0.5)\%$ and $\epsilon_{a_1\pi} = (30 \pm 0.3)\%$, respectively. The latter value includes the 50% reduction factor coming from the $a_1^- \to \rho^0\pi^-$ branching ratio. The rejection factors on the $q\bar{q}$ background events are $\eta_{\rho\pi} = (1.7 \pm 0.1) \times 10^{-4}$ and $\eta_{a_1\pi} = (1.8 \pm 0.1) \times 10^{-4}$, respectively. The errors quoted above reflect the limited Monte-Carlo statistics used (2×10^6 generated $q\overline{q}$ Monte-Carlo events). At this stage of the preselection, the detectable signal events have been almost completely preserved.

In addition, the remaining N_c charged particles not associated with the final state f must allow a $B_d^0/\overline{B_d^0}$ tagging. This tagging is done by requiring the presence of either one high-energy lepton (electron or muon) or one charged Kaon. The minimum lepton energy in the $\Upsilon(4S)$ rest frame is taken as $E_{min}^l = 1.4\ GeV$. The lepton-tagging efficiency is $\epsilon^l = 11\%$. The fraction of the lepton-tagged B's which are wrongly identified is $\omega^l = 4\%$. The rejection factor on the $q\overline{q}$ background associated with a lepton-tagging is taken as $\eta^l = 1\%$, although, in the limit of a perfect particle identification, the Monte-Carlo simulation gives $\eta^l = (0.3 \pm 0.1)\%$. The corresponding tagging efficiency using kaons, including the losses due to the decays in flight, is $\epsilon^K = 35\%$. The fraction of the kaon-tagged B's which are wrongly identified is $\omega^K = 7\%$. The background rejection factor associated with the kaon-tagging is $\eta^K = 25\%$. Depending on the lepton or kaon-tagging mode of the B_d^0, the background/signal ratios are $r_{\rho\pi}^l = 1$, $r_{\rho\pi}^K = 8$, $r_{a_1\pi}^l = 2$ and $r_{a_1\pi}^K = 16$.

To optimize the $\sigma[\sin 2\Phi]$ resolution one must improve the rejection factors and, inevitably, one must accept also a reduction of the selection efficiencies. For a given tagging mode and a given final state, let us define ϵ_* and η_* as the efficiency and the background rejection factor associated with the final cuts of the selection, and r_o as the background/signal ratio obtained at the preselection level, before these final cuts are applied. In order to optimize $\sigma[\sin 2\Phi]$, it can be seen, from equation (16), that the final cuts must minimize the quality factor :

$$Q = \frac{\epsilon_* + r_o \eta_z \eta_*}{\epsilon_*^2}$$

Five distributions are available for this purpose, the $\cos\theta_+^{cm}$ distribution and the δ invariant mass distributions. The shapes of the latter distributions and the correlations between the variables have not been used in the preselection. However, for most of the selected $q\bar{q}$ events, at least one of the δ-invariant masses is close to the preselection cut. This is in contrast with the signal events, for which all the δ-invariant masses tend to be close to zero. This information can be used to reduce the rejection factor. In order to perform easily the optimization of Q, we express the 5-dimension distribution in term of a single variable distribution as follows. For a given event i, let us define f^i, the fraction of Monte-Carlo signal events for which **all** the variables $\cos\theta_+^{cm}$, δW_f, $\delta \overline{W}_f$ $\delta\rho$ and $\delta\pi^0$ (or δa_1) have a larger value than the corresponding ones of the i-event. Having calculated f^i, we define for the same i-event, the R^i_{Rarity} variable as being the probability, for a Monte-Carlo signal event, to be associated with a **smaller** f^i. Hence, the smaller the R^i_{Rarity}, the more likely the i-event is to be a background event. In other words, the R_{Rarity} provides a definition of the probability for an event to be a signal event. The R_{Rarity} distribution of the $q\bar{q}$ background events for the $\rho^-\pi^+$ channel is shown in Figure 8 (lepton- and kaon-tagged events are combined). The dashed line indicates the signal-event distribution; it is flat by definition of the R_{Rarity} variable. Both background and signal distributions correspond to an integrated luminosity of 50 fb^{-1} (970 selected $\rho^\pm\pi^\mp$ signal events). The shape of the R_{Rarity} distribution of the $q\bar{q}$ background for $a_1^\pm\pi^\mp$ events turns out to be very similar.

Depending on the tagging mode of the B_d^0, the minimal values of Q are found to be : $Q^l(\rho\pi) = 1.2$, $Q^K(\rho\pi) = 1.9$, $Q^l(a_1\pi) = 1.4$ and $Q^K(\rho\pi) = 2.5$. The corresponding sin2Φ resolutions, for an integrated luminosity of 50 fb^{-1}, are σ[sin2Φ] =0.15, 0.11, 0.23 and 0.18, respectively. The various numerical values obtained in our analysis are summarized in Tables 1 to 4. We also give in these tables the corresponding values for the $\pi^+\pi^-$ mode for comparison.

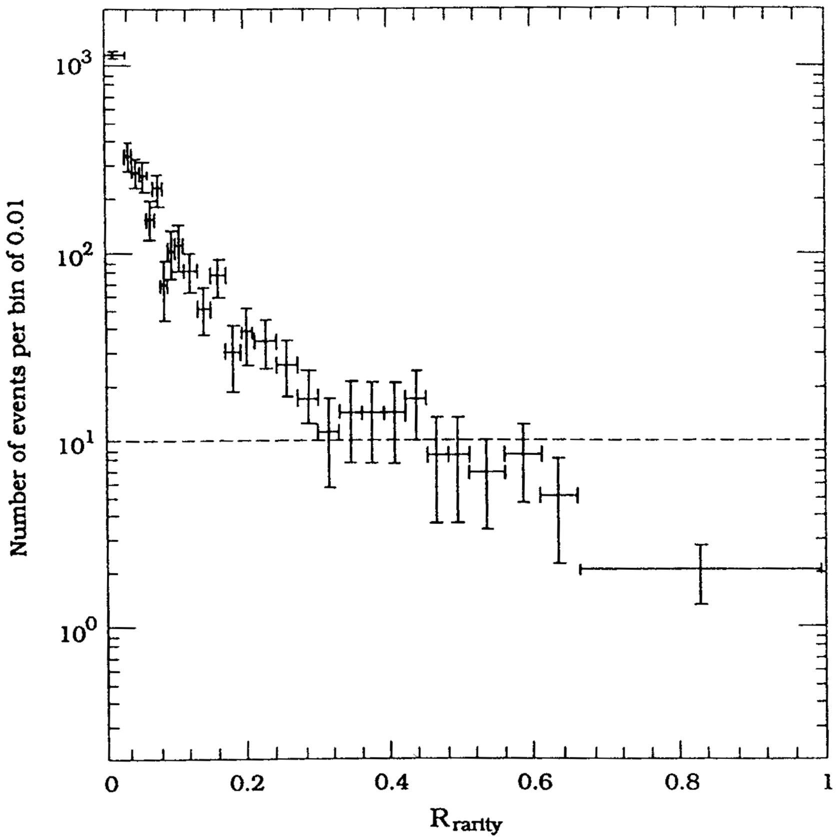


Figure 8. *The distribution of the $q\bar{q}$ background events contaminating the $\rho^{\pm}\pi^{\mp}$ channels as a function of the R_{Rarity} variable (see text for definition). Events tagged with leptons and kaons are combined. The dashed line shows the distribution for the signal-events and is flat by definition of the R_{Rarity} variable. Both background and signal distributions correspond to an integrated luminosity of* 50 fb^{-1} *(970 selected $\rho^{\pm}\pi^{\mp}$ signal events).*

	$\rho^{\pm}\pi^{\mp}$	$a_1^{\pm}\pi^{\mp}$	$\pi^+\pi^-$
ϵ	58%	30%	43%
η	1.7×10^{-4}	1.8×10^{-4}	10^{-6}

Table 1: *Preselection efficiencies and background rejection factors for the three decay modes.*

	Lepton-tagging	kaon-tagging
ϵ	11%	35%
η	1%	25%
ω	4%	7%

Table 2: *Lepton and kaon tagging efficiencies, background rejection factors and wrong identification probabilities.*

Table 3: *Background/signal ratios at the preselection level(r_0), effective background/signal ratios ($r_{eff} = r_0\eta_Z\eta_*/\epsilon_*$), selection efficiencies of the final stage of the selection (ϵ_*) and increase on $\sigma[\sin 2\Phi]$ due to the background ($\sqrt{Q}-1$).*

	$\rho^\pm\pi^\mp$		$a_1^\pm\pi^\mp$		$\pi^+\pi^-$	
	lepton	kaon	lepton	kaon	lepton	kaon
r_0	1	8	2	16	0.02	0.18
r_{eff}	0.13	0.42	0.27	0.57	0.01	0.09
ϵ_*	0.93	0.76	0.89	0.62	1	1
$\sqrt{Q}-1$	10%	40%	20%	60%	0.6%	5%

Table 4: *Overall selection efficiencies (ϵ_{tot}), assumed branching ratios (B), numbers of signal events for a 50 fb^{-1} integrated luminosity (N_{tot}) and the corresponding resolutions $\sigma[\sin 2\Phi]$.*

	$\rho^\pm\pi^\mp$	$a_1^\pm\pi^\mp$	$\pi^+\pi^-$
ϵ_{tot}	21.5%	9.5%	20%
B	6×10^{-5}	6×10^{-5}	2×10^{-5}
N_{tot}	775	340	240
$\sigma[\sin 2\Phi]$	0.09	0.14	0.14

Combining the measurements using the lepton- and kaon-tagging modes, one gets $\sigma[\sin 2\Phi](\rho\pi) = 0.09$ and $\sigma[\sin 2\Phi](a_1\pi) = 0.14$. These results should be compared with the resolution expected with the $\pi^+\pi^-$ events. The $\pi^+\pi^-$ event selection is similar to the one described in Ref. 4, and close to the $\rho^\pm\pi^\mp$ and $a_1^\pm\pi^\mp$ selection. It yields $\epsilon_{\pi\pi} = 0.43$ and $\eta_{\pi\pi} = 10^{-6}$, which give $\sigma[\sin 2\Phi](\pi\pi) = 0.14$. A much smaller reduction factor $\eta_{\pi\pi}$ can be achieved at the price of a significantly smaller selection efficiency. However, the quality factor, which has not been optimized here, is close to one and would increase with the use of more stringent cuts.

Hence, for the particular values of the branching ratios assumed above, the measurement of $\sin 2\Phi$ using the $\rho^\pm\pi^\mp$ final states is more accurate than the one using the $\pi^+\pi^-$ mode. Combining the above resolutions, one gets

$$\sigma[\sin 2\Phi](\pi^+\pi^- + \rho^\pm\pi^\mp + a_1^\pm\pi^\mp) = 0.07 \quad . \tag{17}$$

Compared to the $\pi^+\pi^-$ resolution only, it represents a factor of four improvement on the luminosity needed to reach a given level of precision on sin2Φ. The expected resolutions on sin2Φ as a function of the branching ratios $B(\rho^{\pm}\pi^{\mp})$ and $B(a_1^{\pm}\pi^{\mp})$ are shown in Figure 9. The dotted lines indicated the resolution that would be achieved in the absence of background ($Q = 1$).

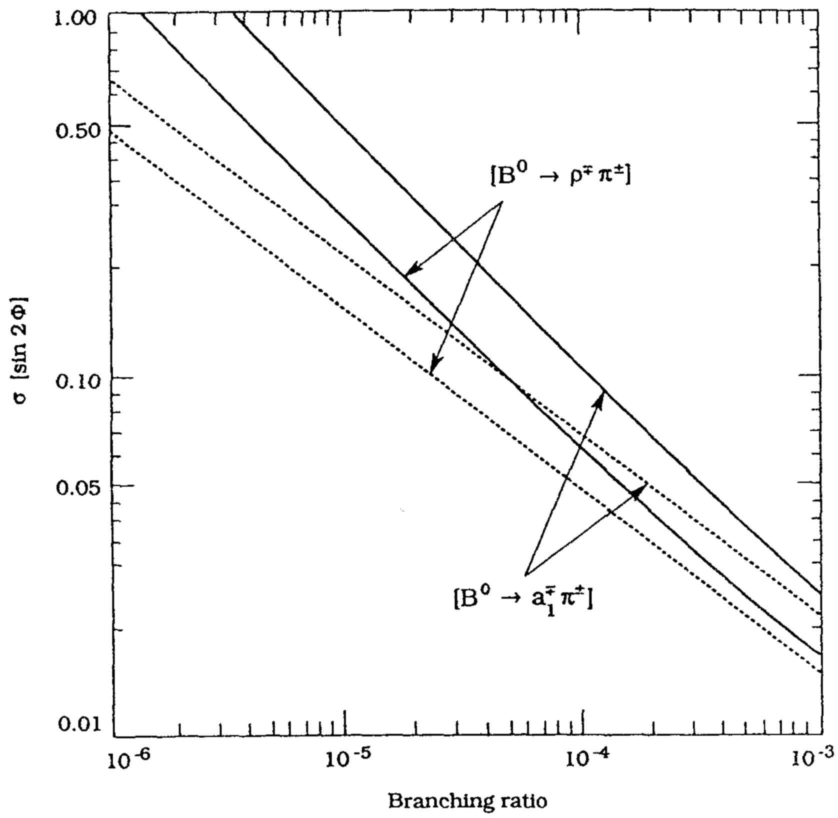


Figure 9. *The expected experimental resolutions on* sin2Φ *as a function of the branching ratios* $B(\rho^{\mp}\pi^{\pm})$ *and* $B(a_1^{\mp}\pi^{\pm})$ *for* $\rho = 1/\sqrt{2}$ *and for an integrated luminosity of* 50 fb^{-1}. *The dotted lines indicated the resolution that would be achieved in the absence of background.*

5. Conclusion

We have shown that neutral B mesons decaying into non-CP eigenstates are very interesting reactions for studying CP violation in particular when these final states are made out of two mesons for which the quark content of one meson is the CP conjugate of the quark content of the second one. Provided this condition is fulfilled, sizeable CP-violating asymmetries can be expected and the measurement of the CKM phases is possible. Two interesting examples of such decays, $B \to \rho^{\pm}\pi^{\mp}$ and $B \to a_1^{\pm}\pi^{\mp}$, have been studied in detail in terms of their reconstruction efficiencies and their background using a generic asymmetric B Factory detector. It is found that these two modes can indeed be efficiently used for probing CP non-conservation and for measuring the CKM phases. Furthermore, these reactions can turn out to be better candidates than the final state $B \to \pi^+\pi^-$.

APPENDIX A

In the main body of this paper, we neglected rescattering and penguin diagrams (see the bottom row of Figure 5). This appendix discusses some physical effects arising when those diagrams are not negligible. When the magnitudes of these diagrams are comparable to the magnitude of the $\overset{(-)}{b}$ quark decay one, we would expect that direct CP violation occurs, i.e. $M \neq \overline{M}$ and $M' \neq \overline{M}'$. The reason is that different CKM phases are involved, and there is no reason to believe that the final-state phases of the rescattering and penguin diagrams and the $\overset{(-)}{b}$ decay one should be the same. In Section 3, we argued that the final state phase of the $\overset{(-)}{b}$ quark decay diagram vanishes, but this argument by no means suggests that the same is true of other diagrams. However, if for some reason the final state phases of the rescattering and penguin diagrams do vanish, then $\alpha_j = \alpha'_j = 0$ for all j in equations (1) and (1′) and we have $M = M'$ and $\overline{M} = \overline{M}'$. Nevertheless, since the various contributing diagrams have different CKM phases, when one measures the time-dependent asymmetry for, say, $B^0 \to \rho^+\pi^-$ versus $\overline{B^0} \to \rho^-\pi^+$, one does not measure $\sin 2\alpha$* but rather a different, unknown combination of CKM elements. To guard against such a wrong interpretation of the measurements (that is in terms of $\sin 2\alpha$), one must be able to disentangle penguin and rescattering contributions from the $\overset{(-)}{b}$ quark decay ones. Such efforts have recently begun to appear in the literature [17].

⋆ Where α is one of the angles of the unitarity triangle [16].

APPENDIX B

We now estimate the decay rate for $\Gamma(\overline{B_d^0} \to \rho^+\pi^-)$ relative to the $\Gamma(\overline{B_d^0} \to \pi^+\pi^-)$. Neglecting penguin and rescattering contributions, factorization ought to hold for color-allowed two-body B decays into light mesons [14]. With this assumption, the ratio of rates is given in terms of form factors.

$$\frac{\Gamma(\overline{B_d^0} \to \rho^+\pi^-)}{\Gamma(\overline{B_d^0} \to \pi^+\pi^-)} \approx \frac{|a+f|^2}{4m_\rho^2|f_+|^2} \tag{B.1}$$

where the π^- is made from the virtual W boson. Here the form factors are defined by :

$$\langle\rho^+|\mathcal{J}_\mu|\overline{B^0}\rangle = ig\epsilon_{\mu\alpha\beta\gamma}\epsilon^{*\alpha}(P_B+P_\rho)^\beta(P_B-P_\rho)^\gamma - f\epsilon^*_\mu$$
$$-a_+\epsilon^*.P_B(P_B+P_\rho)_\mu - a_-\epsilon^*.P_B(P_B-P_\rho)_\mu \tag{B.2}$$

and

$$a \equiv a_+(m_B^2 - m_\rho^2) + a_- m_\pi^2 \approx a_+(m_B^2 - m_\rho^2) \tag{B.3}$$

The 4-momenta of the B and ρ mesons are denoted by P_B and P_ρ respectively. The numerator of equation (B.1) depends upon axial form factors of the transition $\overline{B_d^0} \to \rho^+$ whereas the denominator depends on the vector form factor for $\overline{B_d^0} \to \pi^+$. All the form factors are evaluated at very low $q^2 = m_\pi^2$. Many assumptions and indirect pieces of information are used to estimate this ratio to be 10 ± 4 or 31 ± 7. Let us first list them and then discuss how they are used to obtain our estimate.

1. We use the work of Isgur and Wise where they show how the unknown form factors of $B \to K^*$ can be related to the measured ones of $D \to K^*$, when the 4-velocity transfers of the two transitions are the same [18].
2. A universal q^2-dependence is used for the form factors in all the $B \to X$ transitions[14], where X denotes any light meson ρ, π, K^*, K.

$$form\ factor^{B\to X}(q^2) = \frac{form\ factor^{B\to X}(q^2=0)}{1-\frac{q^2}{m_B^2}} \tag{B.4}$$

3. Factorization is assumed valid for the decays $B \to K^*\psi$

and $B \to K\psi$. That yields information on the form factors for $B \to K^*$ and $B \to K$ at $q^2 = m_\psi^2$[14].

4. ARGUS found that, for large ψ momenta, the decay $B \to \psi X$ proceeds mainly through its helicity zero amplitude[19].
5. SU(3) invariance holds when relating $B \to K^*$ to $B \to \rho$ form factors, if the two transitions have the same 4-velocity transfer :

$$form\ factor^{B\to K^*} = form\ factor^{B\to\rho} \qquad \text{(B.5)}$$

Whereas the vector form factors of $D \to K$ and $D \to K^*$ are in agreement with theory, the axial-vector form factors of $D \to K^*$ are not. Experiment E691 measures amongst others the axial-vector form factors, f and a_+, and finds them to be lower than what existing theoretical models predict [20].

Our objective is to numerically estimate the ratio in equation (B.1). The denominator depends on the vector form factor $f_+^{B\to\pi}$. As no inconsistencies between theory and experiment have yet been found for the vector form factors, we use the theoretical estimate of Ref. 14. Conditional support for this choice comes from the $B \to K\psi$ decay. Under assumption (3), only one form factor -that of the vector $f_+^{B\to K}$- is involved in the $B \to K\psi$ decay. The form factor $f_+^{B\to K}$ can thus be estimated from the $B \to K\psi$ rate[21] and agrees well with the predictions of Ref. 14. What is even more fortunate is that we measure $f_+^{B\to K}$ at a rather low $q^2 = m_\psi^2$. This $B \to K$ form factor is closely related to the $f_+^{B\to\pi}$ at $q^2 = 0$, which must be known in order to estimate the denominator in the ratio of equation (B.1).Thus, we have some confidence in using the result of Ref. 14 for this vector form factor and use $f_+^{B\to\pi}(q^2 = 0) = 0.333$.

In contrast, the numerator depends on the axial form factors $a + f$, where experiment and theory disagree. Thus, to estimate the form factors in the numerator, we wish to maximize experimental and minimize theoretical input. We start by utilizing assumption (1) and relate f and g for the $B \to K^*$ transition at $q_0^2 = 16.5\ \text{GeV}^2$ to the $D \to K^*$ one at $q^2 = 0$. This step is fairly sound[22]. However, scaling the form factors

f and g from q_0^2 down to $q^2 = m_\psi^2$, with hypothesis (2), is riskier, but not absurd as the value $q_0^2/m_B^2 \simeq 0.6$ and is not too close to the poles.

Assuming factorization (assumption (3)) the rate measurement of $B \to K^*\psi$ involves all the three form factors f, g, and a_+. Only a_+ is unknown, and is obtained by solving a quadratic equation, and thus two solutions are obtained[23]. Having a_+, f, and g at $q^2 = m_\psi^2$, hypothesis (2) scales them down to $q^2 = 0$. From assumption (5), i.e. SU(3) invariance, we obtain the form factor $a + f$ for the $B \to \rho$ transition. Hypothesis (2) yields this form factor at the desired $q^2 = 0$ scale.

This analysis assumes that the errors are gaussians and uses the inputs below:

The $D \to K^*$ measurements[20] :

$$V(0) = 0.9\pm0.3\pm0.1\,, \quad A_1(0) = 0.46\pm0.06\pm0.03 \quad \text{(B.6)}$$

where V and A_1 are proportional to g and f respectively, up to different mass factors.

The $B \to K^*\psi$ rate and the B meson lifetime[24] :

$$B(B \to K^*\psi) = 0.11 \times 10^{-2} \pm 35\% \qquad \text{(B.7)}$$

$$|V_{cb}|^2 \tau_B = 3.6 \times 10^9 \text{ GeV}^{-1} \pm 40\% \qquad \text{(B.8)}$$

Without incorporating assumption (4), we get for the ratio of equation (B.1) either 4 ± 4 or 20 ± 9. This prediction allows tiny values for this ratio and is not much of a prediction at all.

We can do better than that by incorporating the information from ARGUS (assumption (4)). Even though the $B \to K_S^0\psi$ component always has $\lambda = 0$, we still note that the rates for $B \to K^*\psi$ are larger than $B \to K_S^0\psi$ and take the finding of ARGUS as indicating that $\lambda = 0$ dominates $B \to K^*\psi$ decays. Incorporating this fact in a standard way, one gets the numerical estimates 10 ± 4 or 31 ± 7.

Some of the assumptions involve larger uncertainties than others. We cannot justify the use of hypothesis (2), as many poles ought to be included. However, a recent calculation by J.L.Rosner hints that when one fits the $b \to c$ transitions to one universal pole, this universal pole is lower than expected[12]. That is the reason why our pole was taken as m_B^2 and not at a higher mass. As for hypothesis (3), rescattering from color-allowed processes, which have a much larger branching rate, such as $B_d^0 \to D_s^+ D^{*-} \to K^{*o}\psi$ can invalidate it. Assumption (5) probably holds to 20%, that is supposedly the level where SU(3) breaking terms become important[25].

It should be noted that were the K and K^* treated as heavy mesons, the Isgur-Wise symmetry limit[11] applied to the $B \to K^{(*)}$ transition and were assumption (3) made, one would predict :

$$B(B_d^0 \to K^{*o}\psi)/B(B_d^0 \to K^0\psi) \approx 3 \qquad \text{(B.9)}$$

and

$$\frac{\Gamma[B_d^0 \to K^{*o}(\lambda = 0)\ \psi(\lambda = 0)]}{\Gamma[B_d^0 \to K^{*o}\psi]} = 0.43 \qquad \text{(B.10)}$$

Since equation (B.10) disagrees with the ARGUS measurement (assumption (4)), this approach was not used in estimating the ratio of equation (B.1).

References

1. M. Kobayashi and T. Maskawa, *Prog. Theor. Phys.* **42**, 652 (1973).
2. K.J.Foley *et al.*, *Proceedings of the Workshop on Experiments, Detectors and Experimental Areas for the Supercollider,* 1987, Berkeley, California, ed. by R. Donaldson and M.G.D. Gilchriese (World Scientific, Singapore, 1988), pp. 701-727.
3. R. Aleksan, J. Bartelt, P.R. Burchat and A. Seiden, *Phys. Rev.* **D39**, 1283 (1989).
4. *The Physics Program of a High-Luminosity Asymmetric B Factory at SLAC*, Preprint SLAC-PUB-353, LBL-27856, CALT-68-1588, edited by D. Hitlin, October 1989.
5. I. Dunietz and J.L. Rosner, *Phys. Rev.* **D34**, 1404 (1986).
6. M. Bander, D. Silverman and A. Soni, *Phys. Rev. Lett.* **43**, 242 (1979).
7. M. Gronau, *Phys. Lett.* **B233**, 479 (1989).
8. G. Feldman *et al.,* SLAC-PUB-4838, CNLS 89/884, LBL-26790 (1989).
9. I.I. Bigi, V.A. Khoze, N.G. Uraltsev and A.I. Sanda, *Advanced Series on Directions in High Energy Physics* - Vol.3 p.175, *CP* Violation, ed. C. Jarlskog, published by World Scientific, Singapore, 1989.
10. N. Isgur and M.B. Wise, *Phys. Lett.* **B232**, 113 (1989).
11. N. Isgur and M.B. Wise, *Phys. Lett.* **B237**, 527 (1990).
12. J.L. Rosner, *Phys. Rev.* **D42**, 3107 (1990).
13. T. Mannel, W. Roberts and Z. Ryzak, *Phys. Lett.* **B248**, 392 (1990); H. Georgi, private communication.
14. M. Bauer, B. Stech and M. Wirbel, *Z. Phys.* **C34**, 103 (1987). For a more recent discussion see also M. Bauer, and M. Wirbel, *Z. Phys.* **C42**, 671 (1989).
15. J.D. Bjorken, *Nucl. Phys.* **B11**, 325 (1989).

16. We use the notation of Reference [4], wherein the reader can find earlier references.

17. M. Gronau and D. London, *Phys. Rev. Lett.* **65**, 3381 (1990).

18. N. Isgur and M.B. Wise, *Phys. Rev.* **D42**, 2388 (1990).

19. ARGUS Collaboration, reported by H. Schroeder at the Singapore conference, September 1990.

20. J. C. Anjos et al., *Phys. Rev. Lett.* **65**, 2630 (1990).

21. Our calculation uses $|a_2| = 0.2$ for this color-suppressed diagram, see H. Albrecht *et al.*, DESY 90-046 (1990): N. G. Deshpande and J. Trampetic, OITS 421 (1989).

22. Here we have chosen $m_b/m_c = 5/1.5$ and the QCD correction factor is 1.1 in equations (11) of Reference [18].

23. This quadratic equation depends, among other things, on $Re(fa_+^*)$. For an exclusive transition, such as the $B \to K^*$ one considered here, the relative phase between f and a_+ is probably 0 or π.

24. We thank S. Stone for providing us the latest CLEO, ARGUS and ALEPH results which were used here.

25. We thank Mark Wise for informing us about his educated guess.

CP Asymmetries in B^0 Decays Beyond the Standard Model

C. Dib, D. London and Y. Nir

Abstract

Of the many ingredients of the Standard Model that are relevant for the analysis of CP asymmetries in B^0 decays, some are likely to hold even beyond the Standard Model while other are sensitive to new physics. Consequently, certain predictions are maintained while other may show dramatic deviations from the Standard Model. Many classes of models may show clear signatures when the asymmetries are measured: four quark generations, Z-mediated flavor-changing neutral currents, supersymmetry and "real superweak" models. On the other hand, models of left-right symmetry and multi-Higgs sectors with natural flavor conservation are unlikely to modify the Standard Model predictions.

1. Introduction

Measurements of CP asymmetries in B^0 decays into CP eigenstates[1] are guaranteed to provide us with most valuable information. They will address three fundamental questions:

(*i*) **Is the Kobayashi-Maskawa phase of the three generation Standard Model (SM) the only source of CP violation?**

So far, CP violation has been clearly observed only in the measurement of the ϵ-parameter in the K^0 system. While the experimental value of ϵ can be accommodated in the SM, it does not by itself test this model. CP asymmetries in B^0 decays will provide us with an observation of CP violation in a different system and are subject to a clean theoretical interpretation. Thus, they will clearly test whether the single phase of the CKM matrix is the only source of CP violation.

The measurement of the CP-violating ϵ parameter in K^0 decay does not provide a test of the Standard Model

(ii) What are the exact values of the CKM parameters?

CP asymmetries in the B^0 system are largely free of hadronic uncertainites

The parameters of the CKM matrix are important physical quantities that merit careful measurement. The determination of V_{ub} and V_{td} or, equivalently, of s_{13} and δ in the standard parametrization, is limited in accuracy due to theoretical uncertainties in modeling $b \to u$ transitions and in hadronic matrix elements (f_B). *CP* asymmetries in B^0 decays provide us with a unique way to measure the CKM parameters. They measure relative phases between various combinations of CKM elements. As these asymmetries are expected to be sizable, systematic errors will probably not obscure the signal. Various consistency checks will further reduce such errors. Most important, theoretically, *CP* asymmetries in B^0 decays are free of hadronic uncertainties.

We emphasize that even in the case that the answers to the above two questions are consistent with the SM, we will still gain most valuable information. However, in this work we are interested in the following question:

(iii) Is there new physics in the quark sector?

CP asymmetries in B^0 decays test those aspects of the quark sector which are most sensitive to the possible existence of new physics: *CP* violation, mixing in the neutral meson systems, and unitarity of the CKM matrix.

We first describe the ingredients of the SM which are relevant to the analysis of *CP* asymmetries in B^0 decays. We study the prospects of these ingredients being modified in the presence of new physics. In the second part, we list those classes of asymmetries which can be cleanly interpreted, and give the SM predictions for these asymmetries. We then explain which of the predictions are likely to be modified with new physics and which are maintained. Finall, we survey specific models: Four quark generations; Z-mediated flavor-changing neutral currents (FCNC); Left-Right Symmetry (LRS); Supersymmetry (SUSY); Multi-Higgs Doublets with natural flavor conservation (NFC); Real superweak contributions to B mixing. We investigate for each model whether it is likely to modify the SM predictions and discuss whether these modifications have unique properties.

Specific extensions of the Standard Model produce specific patterns of CP violation in B^0 decay

Previous general discussions of CP asymmetries in B^0 decays beyond the SM can be found in refs. 2–4. References to studies of specific models will be given in Section 3. A comprehensive analysis of $B - \bar{B}$ mixing beyond the SM is given in ref. 5. The present status of the SM predictions for CP asymmetries in B^0 decays is described in refs. 6–7.

2. The Standard Model Assumptions

The CP asymmetry in neutral B decay, A^{CP}, is the ratio

$$A^{CP}(t) = \frac{\Gamma(\bar{B}^0_{\text{phys}}(t) \to f_{CP}) - \Gamma(B^0_{\text{phys}}(t) \to f_{CP})}{\Gamma(\bar{B}^0_{\text{phys}}(t) \to f_{CP}) + \Gamma(B^0_{\text{phys}}(t) \to f_{CP})}. \quad (2.1)$$

$B^{(0)}_{\text{phys}}$ $[\bar{B}^0_{\text{phys}}(t)]$ is a is a time-evolving initially pure B^0 $[\bar{B}^0]$ state. f_{CP} is a final CP eigenstate. $\Gamma(t)$ is the time-dependent decay rate. Two ingredients of the SM are essential for any clean interpretation of a measurement of A^{CP}. (By "clean interpretation" we mean that the measured value of the asymmetry can be translated into a value of a basic parameter of the electroweak sector or its extension with no significant hadronic or other uncertainties.)

- In the neutral B system the difference in width between the two mass eigenstates is much smaller than the difference in mass, $\Gamma_{12} \ll M_{12}$.
- The direct decay is dominated by a single combination of CKM parameters (or by a single strong phase). This means that the asymmetry is a result of the interference between a direct decay $B \to f_{CP}$ and a process that involves mixing $B \to \bar{B} \to f_{CP}$.

A clean connection between measured CP asymmetries in B^0 decay and parameters of the electroweak sector depends on the direct decay being dominated by a single strong phase, an assumption which can be tested

Under these two conditions, A^{CP} depends on only two properties of the decay process: The type of the decaying neutral B, whether B_d or B_s, and the quark sub-process involved in the direct decays. Therefore, we denote an asymmetry in a B_q decay through a quark sub-process i by A^{CP}_{iq}. Under the same conditions, the asymmetries have a simple time-dependence:

$$A_{iq}^{CP}(t) = \text{Im}\ \lambda_{iq} \sin(\Delta m_q\ t). \tag{2.2}$$

Δm_q is the mass difference in the B_q system, defined to be positive:

$$\Delta m_q \equiv m[B_q(\text{heavy})] - m[B_q(\text{light})]. \tag{2.3}$$

Im λ_{iq} is the amplitude of the sinusoidal time-oscillation, to be determined by experiment; λ_{iq} is a pure phase. To show what ingredients of the SM are further used to calculate λ_{iq}, we derive the specific prediction:

$$\text{Im}\ \lambda(B_s^0 \to \rho K_S^0) = -\sin 2\gamma, \tag{2.4}$$

where γ is an angle in the unitarity triangle (see Figure 1). The list of ingredients goes as follows:

- The direct decay $\bar{b} \to \bar{u}u\bar{d}$ is dominated by the W-mediated tree level diagram. This gives:

$$\lambda \propto \left(\frac{X}{X^*}\right); \quad X_{\bar{b}\to\bar{u}u\bar{d}} = V_{ub}V_{ud}^*. \tag{2.5}$$

- The mixing in the B_s^0 system is dominated by a box diagram with virtual t-quarks. This gives:

$$\lambda \propto \left(\frac{Y}{Y^*}\right); \quad Y_s = V_{tb}^*V_{ts}. \tag{2.6}$$

- The mixing in the K^0 system is dominated by box diagrams with virtual c-quarks. (As B_s^0 produces a $\bar{K}^0$ and $\bar{B}_s^0$ produces a K^0, interference is possible only with $K - \bar{K}$ mixing.) This gives:

$$\lambda \propto \left(\frac{Z}{Z^*}\right); \quad Z_K = V_{cd}V_{cs}^*. \tag{2.7}$$

The result is:

$$\lambda(B_s^0 \to \rho K_S^0) = \left(\frac{V_{ub}V_{ud}^*}{V_{ub}^*V_{ud}}\right)\left(\frac{V_{tb}^*V_{ts}}{V_{tb}V_{ts}^*}\right)\left(\frac{V_{cd}V_{cs}^*}{V_{cd}^*V_{cs}}\right). \tag{2.8}$$

- The following unitarity constraint holds:

$$\mathcal{U}_{sb} \equiv (V_{us}^* V_{ub}) + V_{cs}^* V_{cb} + V_{ts}^* V_{tb} = 0. \tag{2.9}$$

(We put the first term in parentheses since, based on experimental information, we can safely neglect it.) This allows a simplification of (1.8):

$$\lambda(B_s^0 \to \rho K_S^0) = \left(\frac{V_{ub}V_{ud}^*}{V_{ub}^* V_{ud}}\right)\left(\frac{V_{cb}^* V_{cd}}{V_{cb}V_{cd}^*}\right). \tag{2.10}$$

- The following unitarity constraint holds:

$$\mathcal{U}_{db} \equiv V_{ud}^* V_{ub} + V_{cd}^* V_{cb} + V_{td}^* V_{tb} = 0. \tag{2.11}$$

Geometrically, this relation can be represented by a triangle (see Figure 1) with angles α, β, γ that are given by:

$$\alpha \equiv \arg\left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right]; \quad \beta \equiv \arg\left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right];$$

$$\gamma \equiv \arg\left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right]. \tag{2.12}$$

Thus, we derive the prediction of eq. (2.4):

$$\mathrm{Im}\ \lambda(B_s^0 \to \rho K_S^0) = -\sin 2\gamma. \tag{2.13}$$

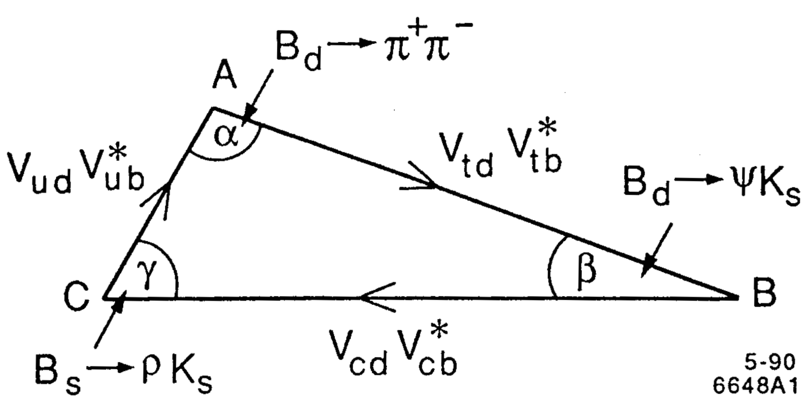


Figure 1. *The unitarity triangle. Relevant classes of CP asymmetries are indicated for each angle (see Tables 1 and 2).*

The l.h.s of this equation will be experimentally determined: it is the amplitude of the sinusoidal oscillation in time of A^{CP}. As for the r.h.s, the SM allowed range for $\sin 2\gamma$ will be determined using information on the length of the sides of the triangle. Thus we need to assume, for example:

- Mixing in the B_d^0 system is dominated by a box diagram with virtual t quarks and, therefore, proportional to $|V_{td}V_{tb}|$.

There are five distinct groups of assumptions within the context of the Standard Model

For any of the asymmetries that we discuss, the analysis goes along similar lines to those presented above. To discuss the sensitivity of the analysis to new physics, we divide the various ingredients into five groups, and comment on each of them in turn.

a. In neutral B^0 systems, $\Gamma_{12} \ll M_{12}$.

Within the SM, one can explicitly calculate the two relevant quantities (assuming that a quark-level description is appropriate):

$$\frac{\Gamma_{12}}{M_{12}} = \frac{3\pi}{2}\frac{1}{f_2(y_t)}\frac{m_b^2}{m_t^2} \sim 10^{-2}. \tag{2.14}$$

$f_2(y_t)$ is a slowly varying function of $y_t \equiv m_t^2/M_W^2$, which assumes values in the range $\{3/4, 1/4\}$ for y_t in the range $\{1, \infty\}$. However, it seems that the order of magnitude estimate holds far beyond the SM[3]. For Γ_{12} to be enhanced, one needs a new decay mechanism which significantly dominates over the W mediated decay. This is most unlikely; there seems to be no viable model that suggests such a situation. Therefore, a ratio Γ_{12}/M_{12} significantly higher than in the SM is possible only in models where M_{12} is significantly suppressed. This requires fine-tuning to cancel the known top contribution with some new physics mechanism. Again, we know of no model where a cancellation to two orders of magnitude is predicted. The argument is particularly solid for the B_d^0 system, as it is supported by experimental evidence: $\Delta M/\Gamma \sim 0.7$, while (upper limits on) branching ratios into states which contribute to Γ_{12} are at the level of 10^{-3}.

b. The relevant decay processes are dominated by the SM W-mediated amplitudes.

Within the SM, there are contributions from penguin diagrams as well. If the matrix elements for the penguin operator are not significantly enhanced, then these amplitudes are suppressed by a factor of $(\alpha_s/12\pi)\ln(m_t^2/m_b^2)$ compared to the tree-level amplitudes. The situation is particularly promising in the $\bar{b} \to \bar{c}c\bar{s}$ processes, where the CKM combinations for the W-mediated and penguin amplitudes carry the same phase. It seems reasonable that for other processes (except for $\bar{b} \to \bar{u}u\bar{s}$ which we do not consider here) the effect is within 10% or less.[8–10]

In the standard Model, penguin amplitudes are unlikely to be larger than 10% of the corresponding tree-level amplitudes

In models beyond the SM, violation of this SM assumption is possible if there is a new decay mechanism which competes with the W-mediated tree-level decay. Unlike our discussion of Γ_{12}, the effect will be important even if it is comparable to the SM diagram (and not necessarily dominating over it). However, experimental measurements of rare processes (*e.g.* $B - \bar{B}$ mixing or $B \to X\ell^+\ell^-$ decays) typically constrain the couplings or the scale of the new physics in a way which renders the contribution from the new physics to tree-level processes very small. For example, amplitudes from new physics at the 1 TeV scale typically give $\lesssim 1\%$ of the SM amplitude.

c. $K-\bar{K}$ mixing is dominated by box-diagrams with virtual c quarks.

Even within the SM there is a non-negligible long-distance contribution. The important ingredient is that the relevant CKM combination is $\arg(Z) = \arg(V_{cd}^* V_{cs})$. The validity of this assumption holds far beyond the SM: Although Z may be modified with new physics, $\arg(Z)$ is not.[4] Consider the condition on mixing in the K system from the measurement of the ϵ parameter:

M_{12} and Γ_{12} for K^0 mesons have the same phase in most models

$$\arg(M_{12}/\Gamma_{12})(\text{mod } \pi) = 6.6 \times 10^{-3}. \tag{2.15}$$

Therefore, to an excellent approximation, M_{12} and Γ_{12} carry the same phase (mod π). Assuming the $K \to 2\pi$ amplitude is proportional to $V_{ud}^* V_{us}$, we may use $\arg(Z) = \arg(V_{ud}^* V_{us})$

independent of the model for mixing. A new mixing mechanism will not be revealed through CP asymmetries in B^0 decays as long as

$$\arg(V_{ud}^* V_{us}) = \arg(V_{cd}^* V_{cs}) (\text{mod } \pi). \tag{2.16}$$

Within any three generation model, eq. (1.16) holds to an excellent approximation due to unitarity constraints. Even within extended models, eq. (1.16) is likely to hold, but with contrived models it could be violated.

d. $B-\bar{B}$ **mixing is dominated by box diagrams with virtual t-quarks.**

$B^0 - \bar{B}^0$ mixing is sensitive to new physics contributions

The SM box diagram is suppressed by being fourth order in the weak coupling and by small mixing angles (the GIM mechanism). Thus, it is not unlikely that new physics contributions, even when suppressed by a high energy scale, will compete with or even dominate over the SM diagram. Indeed, in many models, a new such mechanism for mixing of neutral B's is suggested. If this is the case, there are two possibilities:

(*i*) The phase of the new mixing mechanism is the same as that of the SM mechanism. Consequently, the SM predictions for A^{CP} will not be violated, even though there is new physics in the relevant processes.

A new mixing mechanism could have a different phase than the CKM amplitude, producing dramatic effects in CP-violating asymmetries

(*ii*) The phase of the new mixing mechanism is different from the SM mechanism. Consequently, CP asymmetries in B^0 decays may be very different from the SM predictions. They no longer measure the relative phase between the CKM combinations that determine the decay and the mixing. Instead they measure the relative phase between the CKM combination that determines the decay and the phase from new physics that determines the mixing. As these new phases have no experimental constraints, their effect could be rather dramatic, *e.g.* give maximal asymmetry where the SM predicts zero asymmetry.

e. **The three generation CKM matrix is unitary.**

The relevant unitarity constraints are:

$$\mathcal{U}_{qb} \equiv \sum_{k=1}^{3} V_{kb}V_{kq}^{*} = 0; \quad q = d, s. \tag{2.17}$$

We would like to argue that there is a connection between the three generation unitarity constraints and mixing in the B system. More specifically, if $\mathcal{U}_{sb} \neq 0$ [$\mathcal{U}_{db} \neq 0$], there will be significant contributions from beyond the SM to $B_s^0 - \bar{B}_s^0$ [$B_d^0 - \bar{B}_d^0$] mixing.[4] If the full spectrum of colored fermions consists of the three known generations of quarks, the 3×3 CKM matrix is unitary, and all the constraints hold. There are two basic ways to extend the quark sector, thus allowing a violation of the unitarity constraints:

i. Adding sequential quarks, namely left-handed doublets and right-handed singlets. With n generations, the CKM matrix is a sub-matrix of an $n \times n$ unitary mixing matrix. The relevant unitarity constraints of eq. (1.17) are replaced by:

$$\mathcal{U}_{qb} = -\sum_{k=4}^{n} V_{kb}V_{kq}^{*}. \tag{2.18}$$

At the same time, the u_k quark contributes to $B_q^0 - \bar{B}_q^0$ mixing through box-diagrams proportionally to $(V_{kb}V_{kq}^{*})^2$. This contribution is enhanced by m_k^2/m_t^2.

ii. Adding non-sequential quarks. The charged-current mixing matrix is non unitary; consequently there are flavor-changing neutral currents. The best-known example is the model with an $SU(2)_L$ singlet of charge -1/3 quark. In this case, the unitarity constraints are modified to

$$\mathcal{U}_{qb} = U_{qb}, \tag{2.19}$$

where U_{qb} is a flavor-changing coupling of the Z^0 gauge boson. At the same time, there is a contribution to $B_q^0 - \bar{B}_q^0$ mixing from tree-level Z-mediated diagrams, proportional to $(U_{qb})^2$. This contribution is enhanced because it appears at tree-level.

Small violations of the unitarity constraints can give rise to large changes in the Standard Model predictions for CP-violating asymmetries

The conclusion is that a small violation of the unitarity constraints usually gives a significant new contribution to $B^0 - \bar{B}^0$ mixing. For CP asymmetries in B^0 decays, this second effect is the one that may give substantial deviations from the SM predictions.

To summarize: When we survey models of new physics for possible violation of the SM predictions for CP asymmetries in B^0 decays, the main questions to be asked are:

1. Is there a possibility of a new mechanism for mixing of neutral B's?
2. Does this mechanism carry new phases?

We also check the following aspects:

3. Is unitarity of the 3×3 CKM matrix violated?
4. Are there significant new contributions to the direct $B \to f_{CP}$ decays?

3. Modifications of the SM Predictions

Our study involves those classes of asymmetries for which, within the SM, the direct decay is expected to be dominated by a single combination of CKM parameters. The asymmetries are denoted by Im λ_{iq}. The index $i = 1, \ldots, 5$ denotes the quark sub-process. The index $q = d, s$ denotes the type of decaying meson, B_q. In Tables 1 and 2 we list CP asymmetries in B_d^0 and B_s^0 decays, respectively. The list of hadronic final states gives examples only; other states may be more favorable experimentally. We always quote the CP asymmetry for CP-even states, regardless of the specific hadronic state listed.

From our general analysis in the previous section, it follows that in most models of new physics:

- λ_{iq} is of the form ($|\lambda_{iq}| = 1$):

$$\lambda_{iq} = \left(\frac{X_i}{X_i^*}\right)\left(\frac{Y_q}{Y_q^*}\right)\left(\frac{Z_{iq}}{Z_{iq}^*}\right). \tag{3.1}$$

The X_i-factor depends on the quark sub-process amplitude. The Y_q-factor depends on the mixing amplitude of the decaying meson. The Z_{iq}-factor (which differs from 1 only for final states with an odd number of neutral kaons) depends on the $K-\bar{K}$ mixing amplitude.

- The X_i factor is given by

$$\begin{aligned} X_1 &\equiv X(\bar{b}\to\bar{c}c\bar{s}) = V_{cb}V_{cs}^*, \\ X_2 &\equiv X(\bar{b}\to\bar{c}c\bar{d}) = V_{cb}V_{cd}^*, \\ X_3 &\equiv X(\bar{b}\to\bar{u}u\bar{d}) = V_{ub}V_{ud}^*. \end{aligned} \tag{3.2}$$

Class (iq)	Quark sub-process	Final state (example)	SM prediction
$1d$	$\bar{b}\to\bar{c}c\bar{s}$	ψK_S^0	$-\sin 2\beta$
$2d$	$\bar{b}\to\bar{c}c\bar{d}$	D^+D^-	$-\sin 2\beta$
$3d$	$\bar{b}\to\bar{u}u\bar{d}$	$\pi^+\pi^-$	$\sin 2\alpha$
$4d$	$\bar{b}\to\bar{s}s\bar{s}$	ϕK_S^0	$-\sin 2\beta$
$5d$	$\bar{b}\to\bar{s}s\bar{d}$	$K_S^0 K_S^0$	0

Table 1. *CP Asymmetries in B_d^0 Decays*

Class (iq)	Quark sub-process	Final state (example)	SM prediction
$1s$	$\bar{b}\to\bar{c}c\bar{s}$	$D_s^+D_s^-$	0
$2s$	$\bar{b}\to\bar{c}c\bar{d}$	ψK_S^0	0
$3s$	$\bar{b}\to\bar{u}u\bar{d}$	ρK_S^0	$-\sin 2\gamma$
$4s$	$\bar{b}\to\bar{s}s\bar{s}$	$\eta'\eta'$	0
$5s$	$\bar{b}\to\bar{s}s\bar{d}$	ϕK_S^0	$\sin 2\beta$

Table 2. *CP Asymmetries in B_s^0 Decays*

For classes $i=4,5$, the dominant direct decay mechanism within the SM is the penguin amplitude. We include them in the tables for completeness, but will not discuss them in detail. A detailed analysis is given in ref. 11.

- The Z_{iq} factor is given by:

$$\begin{aligned} Z_{2d} &= Z_{3d} = Z_{5d} = Z_{1s} = Z_{4s} = 1, \\ Z_{1d} &= Z_{4d} = Z_{2s}^* = Z_{3s}^* = Z_{5s}^* = V_{ud}^*V_{us}. \end{aligned} \tag{3.3}$$

On the other hand:

◊ $\arg[Y_d]$ and $\arg[Y_s]$ may differ significantly from the SM values, if there are new contributions to the mixing of neutral B's, and if these contributions carry new phases.

◊ The unitarity constraints on $\mathcal{U}_{db}$ and $\mathcal{U}_{sb}$ may be significantly violated in models of extended quark sector.

Through the effect on mixing, failure of unitarity affects CP-violating asymmetries primarily because the asymmetries than measure different angles

Within the SM, the asymmetries measure angles in the complex plane between various combinations of the charged current mixing matrix, as those determine both b decays and $B_q^0 - \bar{B}_q^0$ mixing. These angles are calculated within the SM on the basis of direct measurements and unitarity of the CKM matrix. Within models of new physics, unitarity of the charged current mixing matrix may be lost, but this is not the main reason for the asymmetries being modified. The reason is rather that, when $B_q^0 - \bar{B}_q^0$ mixing has significant contributions from new physics, the asymmetries measure different quantities, namely angles between combinations of elements of the charged current mixing matrix determining b decays and elements of mixing matrices in sectors of new physics (squarks, multi-Higgs, etc.) which determine $B_q^0 - \bar{B}_q^0$ mixing.

In view of these observations, let us examine which of the predictions of Tables I and II are likely to hold and which may be violated with new physics.[4,11]

The predictions

$$\text{Im } \lambda_{1d} = \text{Im } \lambda_{2d}, \quad \text{Im } \lambda_{1s} = \text{Im } \lambda_{2s}, \tag{3.4}$$

do not depend on the mixing mechanism for neutral $B's$. Instead, they depend only on the mechanism for tree-level decays and for $K^0 - \bar{K}^0$ mixing. They will hold as long as $\arg[X_1] + \arg[Z_{1q}] = \arg[X_2] + \arg[Z_{2q}]$. As explained above, this relation will hold in all but some very contrived models with both new mechanism for $K^0 - \bar{K}^0$ mixing and extended quark sector.

The predictions

$$\text{Im } \lambda_{1d} = \text{Im } \lambda_{4d}, \quad \text{Im } \lambda_{1s} = \text{Im } \lambda_{4s}, \tag{3.5}$$

do not depend on the mixing mechanism for neutral $B's$. Instead, they depend only on the mechanism for direct decays and the unitarity constraint $\mathcal{U}_{sb} = 0$. They are likely to be violated in any model with $\mathcal{U}_{sb} \neq 0$. Similarly, certain relations between asymmetries in classes $i = 2, 3$ and $i = 5$ will be violated if $\mathcal{U}_{db} \neq 0$.

The prediction

$$\mathrm{Im}\ \lambda_{1s} = 0 \tag{3.6}$$

depends on the mechanism for tree-level decays, on the unitarity constraint $\mathcal{U}_{sb} = 0$ and on the mechanism for B_s^0 mixing. It is likely to be violated in models with new phases in $B_s^0 - \bar{B}_s^0$ mixing.

The predictions

$$\mathrm{Im}\ \lambda_{2d} = -\sin(2\beta), \quad \mathrm{Im}\ \lambda_{3d} = \sin(2\alpha), \tag{3.7}$$

depend on the mechanism for tree-level decays, on the unitarity constraint $\mathcal{U}_{db} = 0$ and on the mechanism for B_d mixing. They are likely to be violated in models with new phases in $B_d^0 - \bar{B}_d^0$ mixing.

Finally, we note that the three angles deduced from measurements of the Im λ_{1d}, Im λ_{3d} and Im λ_{3s} will sum up to 180^o whenever the amplitude for $B_s^0 - \bar{B}_s^0$ mixing is real.[4] This is independent of whether they correspond to the angles of the unitarity triangle or not.

4. Models of New Physics

We now briefly survey relevant models of new physics. As explained in previous sections, we look for violation of the unitarity constraints:

$$\mathcal{U}_{db} = 0; \quad \mathcal{U}_{sb} = 0, \tag{4.1}$$

and, more important, for new contributions to $B_q^0 - \bar{B}_q^0$ mixing which are at least comparable to the SM contribution:

$$M_{12}^t(B_q^0) = \frac{G_F^2}{12\pi^2}\eta M_B(B_B f_B^2) M_W^2 y_t f_2(y_t)(V_{tb}^* V_{tq})^2. \tag{4.2}$$

1. **Four quark generations**[12–15]:

There are no new tree-level contributions to b decays. Thus, Γ_{12} remains unmodified and the direct tree-level decays are still dominated by the W-mediated diagrams. Unitarity of the CKM matrix is violated:

$$\mathcal{U}_{qb} = -V_{t'b}V_{t'q}^*. \tag{4.3}$$

A fourth quark generation could significantly alter B mixing

There could be significant new contributions to $B_q^0 - \bar{B}_q^0$ mixing. For example, a box diagram with virtual t' quarks contributes:

$$M_{12}^{t'}(B_q^0) = \frac{G_F^2}{12\pi^2}\eta M_B(B_B f_B^2)M_W^2 y_{t'} f_2(y_{t'})(V_{t'b}^* V_{t'q})^2. \tag{4.4}$$

The full (4×4) mixing matrix has three independent phases, which could appear in M_{12}.

2. **Z-mediated flavor-changing neutral currents (FCNC)**[16]:

Left-right symmetric models are unlikely to modify the pattern of CP violation

There are tree-level Z-mediated contributions to b decays. Experimental constraints imply that they are below 10% of the W diagram for $i = 1$, but could be as large as 20% for $i = 2, 3$. Although Γ_{12} has new contributions from Z mediated diagrams, it is not expected to be enhanced. The direct decays are still dominated by the W-mediated diagrams, but the theoretical analysis of $b \to d$ processes may be less solid. Unitarity of the CKM matrix is violated:

$$\mathcal{U}_{qb} = U_{qb}, \tag{4.5}$$

where U_{qb} is a non-diagonal Z-coupling. There could be significant new contributions to $B_q^0 - \bar{B}_q^0$ mixing from tree-level diagrams:

$$M_{12}^Z(B_q^0) = \frac{\sqrt{2}G_F}{12}\eta M_B(B_B f_B^2)(U_{qb}^*)^2. \tag{4.6}$$

There are new independent phases in the neutral current mixing matrix which could appear in M_{12}.

3. **Multi-Higgs doublets with natural flavor conservation (NFC):**

There are tree-level ϕ^+-mediated contributions to b decays. Experimental limits on the mass of the charged Higgs imply that they are negligible. Thus, there is no significant effect on Γ_{12} and on the direct decays. Unitarity of the CKM matrix is maintained. There could be significant new contributions to $B_q^0 - \bar{B}_q^0$ mixing from box-diagrams with charged Higgs. In a general n-doublet model with NFC, the couplings of the physical charged scalars to quarks are given by:[17]

$$\mathcal{L} = \sum_{k=2}^{n} \frac{g_2 \phi_k^+}{2\sqrt{2} M_W} \bar{U}[-(Y_{1k}/Y_{11}) M_u V(1-\gamma_5)$$

$$+(Y_{2k}/Y_{21}) V M_d (1+\gamma_5)] D + h.c. \qquad (4.7)$$

Y is the matrix that rotates the mass eigenstates of charged scalars to the interaction eigenbasis. Without loss of generality we took ϕ_1^+ to be the Goldstone boson. The Y-matrix introduces new phases which are not related to those of V_{CKM}. However, the leading contribution from ϕ_k^+-exchange diagrams to $B_q^0 - \bar{B}_q^0$ mixing comes from the term proportional to m_t. This gives $(Y_{1k} V_{td})(Y_{1k} V_{tb})^*$, and has exactly the same phase as the SM W-exchange contribution. Consequently, $\arg(M_{12})$ remains unmodified.

It is amusing to note that in the multi-scalar models with NFC and with spontaneous *CP* violation (SCPV), where $\delta_{KM} = 0$ (so that the unitarity triangle becomes a line), *CP* asymmetries in classes $i = 1, 2, 3$ all vanish. (This was shown in detail for $B_d^0 \to \psi K_S^0$ in ref. 18. A more general discussion of the $\delta_{KM} = 0$ case is given in ref. 2.) However, it seems that with the new limits on scalar masses from LEP, this class of models is phenomenologically excluded.

4. **Left-Right Symmetry (LRS)**[19–20]:

There are tree-level W_R-mediated contributions to b decays. Experimental limits on the mass of W_R imply that they are negligible. Thus, there is no significant effect on Γ_{12} and on the direct decays. Unitarity of the CKM matrix is maintained. The experimental limits on $M(W_R)$ from $K^0 - \bar{K}^0$

mixing and the relations between the mixing matrices for W_L and W_R interactions imply that there could be no significant new contributions to $B_q^0 - \bar{B}_q^0$ mixing. The only way to evade these conclusions is by giving up the left-right symmetry (namely, a model of $SU(2)_L \times SU(2)_R \times U(1)_{B-L}$ gauge symmetry but no discrete $L \leftrightarrow R$ symmetry), and even then one needs to fine tune the quark sector parameters.

5. **Supersymmetry (SUSY)**[21]:

There are no new tree-level contributions to b decays. Thus, Γ_{12} remains unmodified and the direct tree-level decays are still dominated by the W-mediated diagrams. Unitarity of the CKM matrix is maintained. There could be significant new contributions to $B_q^0 - \bar{B}_q^0$ mixing from box diagrams with intermediate gluinos and squarks. Whether these box diagrams carry phases that are different from those of the SM box diagrams depends on the specific SUSY model. In the *minimal* SUSY model, only left-handed squarks (namely, superpartners of left-handed quarks) contribute. The couplings $\tilde{g}\tilde{d}_{Li}\tilde{d}_{Lj}$ are proportional to the CKM element V_{ij} and thus no new phases are introduced:

$$M_{12}^{\tilde{g}}(B_q^0) = \frac{\alpha_s^2}{27m_{\tilde{g}}^2} B f_B^2 M_B (V_{td}V_{tb}^*)^2 \Delta S_t(m_{\tilde{d}}, m_{\tilde{b}}, m_{\tilde{g}}). \quad (4.8)$$

The function ΔS_t can be found, for example, in ref. 5. Thus *CP* asymmetries are not modified in minimal SUSY models. However, in less restrictive SUSY models, there are contributions from box diagrams with right-handed squarks as well. The mixing matrices are not related to V_{CKM} and carry, in general, new phases.[22] We emphasize that (unlike our discussion of LRS models), the difference between minimal and extended SUSY models is only in simplicity and predictive power, but not in the basic theoretical principles, and thus extended models are not less motivated than the minimal ones.

Extended SUSY models can modify $B^0 - \bar{B}^0$ mixing

6. **"Real Superweak" models**[2]:

This generic framework assumes that $\Delta B = 1$ processes are dominated by the SM amplitudes, but $\Delta B = 2$ processes may have significant new contributions. The only assumption additional to our discussion in Section 1 is that these

Although in superweak models $B_d^0 \to \psi K_S^0$ and $B_d^0 \to \pi\pi$ no longer measure α and β, the angles deduced from the measurements still sum with γ (from $B_s^0 \to \rho K_S^0$) to 180°

new contributions are real. This means that the phases from the direct decays ($\arg X$) remain the same as in the SM. As for the mixing, while the phase in B_s^0 mixing ($\arg Y_s$) remains the same, the phase in B_d^0 mixing ($\arg Y_d$) is reduced. Consequently, this model predicts no modification of the SM prediction for asymmetries in B_s^0 decays; a reduction in the asymmetry in $B_d^0 \to \psi K_S^0$; and a modification (in either direction) of the asymmetry in $B_d^0 \to \pi^+\pi^-$. This model demonstrates a general feature noted in ref. 4: Even though the measurements of $B_d^0 \to \psi K_S^0$ and $B_d^0 \to \pi^+\pi^-$ do not measure β and α anymore, the angles deduced from these measurements will sum up with γ (deduced correctly from $B_s^0 \to \rho K_S^0$) to 180^o. This is guaranteed by the B_s^0 mixing amplitude being real.

A summary of our conclusions is given in Table 3. The second column describes, for each model, whether unitarity of the three generation CKM matrix is maintained (a triangle) or violated (a quadrangle). The third column gives an example of a new contribution to $B_q^0 - \bar{B}_q^0$ mixing. Unless otherwise mentioned, the contribution could be large and carry new phases.

The measurement of CP asymmetries in B^0 decays should constitute a whole program: the more classes of asymmetries measured, the better we understand the detailed nature of new physics which may account for deviations from the SM predictions.

ACKNOWLEDGMENTS

We thank Ikaros Bigi, Helen Quinn and Lincoln Wolfenstein for useful discussions and communications.

Model	CKM Unitarity	B - $\bar{B}$ Mixing	SM Predictions for A^{CP}
SM		t W W t	
Four Quark Generations	$V_{t'd}V^*_{t'b}$	t′	Modified
Multi-Scalar with NFC (General)		ϕ^+	Unmodified
(+ SCPV)	——	No New Phases	All Asymmetries Vanish
Z-Mediated FCNC	U^*_{db}	Z	Modified
LRS		W_R Small	Unmodified
SUSY (General)		$\tilde{g}$ $\tilde{q}_L,\tilde{q}_R$ $\tilde{q}_L,\tilde{q}_R$ $\tilde{g}$	Modified
(Minimal)		$\tilde{g}$ $\tilde{q}_L$ $\tilde{q}_L$ $\tilde{g}$ No New Phases	Unmodified
"Real Superweak"		Real	Modified for B_d Unmodified for B_s

8-90 6708A1

Table 3. *Effects of new physics on CP asymmetries*

REFERENCES

1. A.B. Carter and A.I. Sanda, *Phys. Rev. Lett.* **45**, 952 (1980); *Phys. Rev.* **D23**, 1567 (1981);
I.I. Bigi and A.I. Sanda, *Nucl. Phys.* **B193**, 85 (1981); **B281**, 41 (1987).
2. J. Liu and L. Wolfenstein, *Phys. Lett.* **B197**, 536 (1987).
3. I.I. Bigi, V.A. Khoze, N.G. Uraltsev and A.I. Sanda, in: *CP Violation*, ed. C. Jarlskog (World Scientific, Singapore, 1989), p. 175.
4. Y. Nir and D. Silverman, *Nucl. Phys.* **B345**, 301 (1990).
5. P.J. Franzini, *Phys. Reports* **173**, 1 (1989).
6. P. Krawczyk, D. London, R.D. Peccei and H. Steger, *Nucl. Phys.* **B307**, 19 (1988).
7. C.O. Dib, I. Dunietz, F.J. Gilman and Y. Nir, *Phys. Rev.* **D41**, 1522 (1990).
8. D. London and R.D. Peccei, *Phys. Lett.* **B223**, 257 (1989).
9. M. Gronau, *Phys. Rev. Lett.* **63**, 1451 (1989).
10. B. Grinstein, *Phys. Lett.* **B229**, (1989) 280.
11. Y. Nir and H.R. Quinn, *Phys. Rev.* D42, 1473 (1990).
12. I.I. Bigi and S. Wakaizumi, *Phys. Lett.* **B188**, 501 (1987).
13. M. Tanimoto, Y. Suetaka and K. Senba, *Z. Phys.* **C40**, 539 (1988).
14. T. Hasuike, T. Hattori, T. Hayashi and S. Wakaizumi, Mod. *Phys. Lett.* **A4**, 2465 (1989); *Phys. Rev.* **D41**, 1691 (1990).
15. D. London, *Phys. Lett.* **B234**, 354 (1990).
16. Y. Nir and D. Silverman, *Phys. Rev.* **D42**, 1477 (1990).
17. See *e.g.* G. Boyd, A.K. Gupta, S.P. Trivedi and M.B. Wise, *Phys. Lett.* **B241**, 584 (1990).

18. J.F. Donoghue and E. Golowich, *Phys. Rev.* **D37**, 2543 (1988).
19. G. Ecker and W. Grimus, *Z. Phys.* **C30**, 293 (1986).
20. D. London and D. Wyler, *Phys. Lett.* **B232**, 503 (1989).
21. I.I. Bigi and F. Gabbiani, UNDHEP-90-BIG03 (1990).
22. See *e.g.* Y. Nir, *Nucl. Phys.* **B273**, 567 (1986).

DYNAMICAL MODELS OF $B \to J/\psi\, K^*$

B.F.L. WARD

ABSTRACT

An issue of some interest in the general theory of *CP* violation in B meson decay is the relative admixture of even and odd final *CP* eigenstates in $B^0 \to J/\psi K^{*0}$. For example, dominance of one sign of *CP* in this decay would increase the available modes for measuring the *CP* violation in the respective B meson decays in a B Factory by a factor ~ 2.5 and, hence, would decrease the luminosity requirements for such a Factory from the *CP* violation perspective. Here, we therefore study this relative admixture in several dynamical models of the respective decay. We are encouraged by our findings.

1. INTRODUCTION

There is currently considerable interest[1] in exploring the phenomenon of *CP* violation in the B meson system. For the respective *CP*-violating asymmetries are expected to be large and, hence, present a unique window on the possible dynamical origins of the phenomenon itself, for example. Further, the decay mode $B^0 \to J/\psi K_S^0$ has already been observed and, based on the observed BR, one may guesstimate the required number of B's in a B Factory if one wants to probe *CP* violation in this mode via the standard asymmetries. The result of this exercise is well-known to be $\sim 10^8$ B's.

Such a number of B's, if it is to be obtained in a reasonable ($\sim$ 3 year) period would then necessitate luminosities of $\sim 3 \times 10^{33} - 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, well-beyond those achieved at any previous colliding beam device. Accordingly, it is imperative to explore other possible B physics avenues which may help to reduce the luminosity required to probe *CP* violation in the B system. Thus, in this paper, we explore the potential of the $B^0 \to J/\psi K^{*0}$ decay for the measurement of *CP* violating asymmetries in B decay. Indeed, elementary arguments[2] would suggest that its BR is ~ 3 times that for $B_d^0 \to J/\psi K_S^0$

and measurements[3] in $B^+ \to J/\psi K^{*+}$ are consistent with this expectation. Hence, it is indeed appropriate to explore such a mode as $B_d^0 \to J/\psi K^{*0}$ as a potential additional mode in which to study CP violation in B decays.

Specifically, the primary issue[4] can be summarized as follows. In $B_d^0 \to J/\psi K^{*0}$, there are three possible orbital angular momentum quantum numbers, $L = 0, 1$ and 2. The CP eigenvalue of the final state is $(-1)^L(-1)^2(1) = (-1)^L$. Thus, the even values of L have positive (even) CP and the odd values of L have negative (odd) CP. Hence, the expected CP violating asymmetries are opposite in sign for the even and odd L cases. Thus, the important question is what is the relative amount of even or odd L in the decay $B_d^0 \to J/\psi K^{*0}$: if either type of L dominates, the mode can be used as another avenue to CP violation in B decay without further qualification. If, on the other hand, the ratio of L even to L odd contributions in the respective final states is near 1, a detailed moments analysis of the various subsequent decay distributions[5] will be necessary to isolate the respective CP violating asymmetries. Such analysis is discussed in detail in Ref. 5. Hence, in what follows, we wish to explore the expected ratio of CP odd to CP even final states in $B_d^0 \to J/\psi K^{*0}$ in dynamical models of this decay.

The models with which we shall work are the model of Wirbel, Stech and Bauer (BSW)[2], the model of Isgur, *et al.* (ISGW),[6] the methods of Lepage and Brodsky and the recent heavy quark limit methods of Isgur, Wise and Bjorken.[7] These models and methods are, of course, not a complete list of all published approaches to the decay under consideration. Rather, they represent a reasonable span in the types of assumptions which are involved in the various approaches so that we can get a quantitative estimate for the effect of these various assumptions on any conclusions we may wish draw by considering these four examples.

Our work is organized as follows in this paper. In the next Section, we present the expectation of $B^0 \to J/\psi K^{*0}$ in the BSW model and, at the same time, we set our notational conventions. In Section 3, we consider the method of Lepage and Brodsky in our present context. In Section 4, we consider the predictions of the ISGW model for our decay. In Section 5, we discuss this decay from the standpoint of the recent heavy quark mass limits methods of Wise, Isgur and Bjorken. Section 6 contains some summary remarks.

2. $B_d^0 \to J/\psi K^{*0}$ IN THE BSW MODEL

The prediction of the BSW model for the process $B_d^0 \to J/\psi K^{*0}$ as to the ratio of CP odd to CP even final states is presented in this Section. We begin by setting our notational and kinematical conventions. The process of interest to us is illustrated in Figure 1.

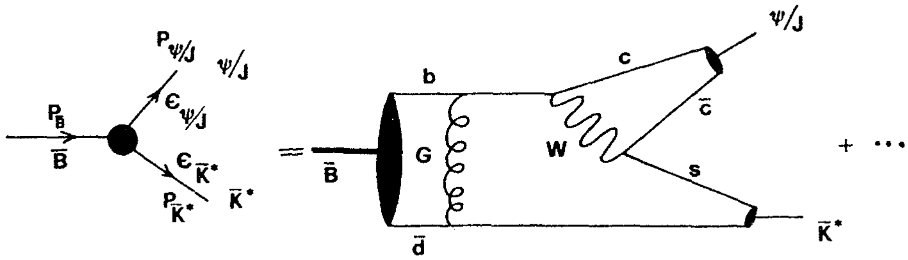


Figure 1. *The process $\bar{B} \to J/\psi \bar{K}^*$ to lowest order in the $SU(2)_L \times U_1$ electroweak theory. P_A is the 4-momentum of A, $A = \bar{B}, J/\psi, \bar{K}^*$, and $E_A \equiv P_A^0$.*

We then follow the theory of BSW, working, as they do, in the point-like limit for the W propagator; for m_b^2/M_W^2 is indeed small compared to 1. Taking into account the standard QCD corrections, BSW identify, then, the effective interaction Lagrangian

$$\mathcal{L}_{eff} = \frac{G_F}{\sqrt{2}}[a_1 \bar{c}\gamma_\mu(1-\gamma_5)s'\bar{b}'\gamma^\mu(1-\gamma_5)c + a_2\bar{s}'\gamma_\mu(1-\gamma_5)b'\bar{c}\gamma^\mu(1-\gamma_5)c + h.c.] \qquad (1)$$

where a_1 and a_2 are determined by the standard QCD short-distance corrections, G_F is the Fermi constant, h.c. denotes hermitian conjugation and b' and s' are the usual CKM rotated mixtures of the d, s, b mass eigenstates. The values of a_1 and a_2 are well-known if one takes them from the perturbative QCD theory (in conjunction with the partial differential equations of QCD); here we leave a_1 and a_2 as parameters; for, we will not need their values in what follows.

To proceed, BSW then invoke the current-field identity (and the general neglect of final state interactions between J/ψ and K^{*0}) to arrive at the amplitude

$$\mathcal{M}(\bar{B} \to J/\psi K^{*0}) = \frac{iG_F}{\sqrt{2}} a_2 (2\pi)^4 \delta(P_{\bar{B}} - P_{J/\psi} - P_{K^{*0}}) \times$$

$$\langle K^{*0} | J_\mu^{(b's')}(0) | \bar{B} \rangle \varepsilon^{*\mu}_{J/\psi} F_{J/\psi} m_{J/\psi} / (2E_{J/\psi}(2\pi)^3)^{1/2} \qquad (2)$$

where we have now introduced the further notations

$$\langle 0 | \bar{c}(0) \gamma_\mu c(0) | \psi / J \rangle = \varepsilon_{J/\psi\mu} F_{J/\psi} m_{J/\psi} / (2E_{J/\psi}(2\pi)^3)^{1/2} \quad (3)$$

and

$$J_\mu^{(b's')} = \bar{b}' \gamma_\mu (1 - \gamma_5) s' \qquad (4)$$

The complete prediction for the process in Figure 1 then requires a systematic analysis of the transition matrix element $\langle \bar{K}^* | J_\mu^{(b's')}(0) | \bar{B} \rangle$. For this, in general, a model is necessary at the present state of our knowledge of the strong interaction.

Indeed, a model-independent decomposition of the respective transition matrix element is afforded by writing (following BSW and suppressing trivial kinematical factors)

$$\tilde{\mathcal{M}}_\mu \equiv \langle K^{*0} | J_\mu^{(b's')}(0) | \bar{B} \rangle = \frac{2}{(m_B + m_{K^*})} \varepsilon_{\mu\nu\rho\sigma} \varepsilon^{*\nu}_{K^{*0}} P^\rho_{\bar{B}} P^\sigma_{K^{*0}} V(q^2)$$

$$+ i\{\varepsilon^*_{K^{*0}\mu}(m_B + m_{K^{*0}}) A_1(q^2) - \varepsilon^*_{K^{*0}} q (P_{\bar{B}} + P_{K^{*0}})_\mu A_2(q^2) / m_{\bar{B}} + m_{K^*}$$

$$\frac{-\varepsilon^*_{K^{*0}} q}{q^2} (2m_{K^{*0}}) q_\mu A_3(q^2)\} + i(\varepsilon^*_{K^{*0}} q / q^2)(2m_{K^{*0}}) q_\mu A_0(q^2) \qquad (5)$$

for some invariant functions V, A_i of $q^2 = m^2_{J/\psi}$ where $q = P_{\bar{B}} - P_{\bar{K}^{*0}}$. The strategy is to make a pole approximation to the V and A_i, of the form

$$P(q^2) = P_0 / (1 - q^2/m^2) \qquad (6)$$

for masses and residues (m, P_0) which are, in the BSW model, determined from a relativistic harmonic oscillator potential model with a gaussian transverse momentum distribution characterized by a scale $\omega \sim .35$ GeV. In this way, then, one may arrive at a prediction for the ratio

$$r \equiv \Gamma(CP = odd)/(\Gamma(CP = even) + \Gamma(CP = odd)) \qquad (7)$$

for the process in Figure 1, where we now define $\Gamma(CP = s)$ to be the total width of $\bar{B}^0 \to J/\psi \bar{K}^{*0}$ into final states with $CP = s$, $s =$ even, odd.

Indeed, specializing to the specific results for the V and A_i of the BSW model, we first observe that the respective $J/\psi - \bar{K}^*$ helicity amplitudes are

$$\mathcal{M}(0,0) = \left(-\frac{G_F}{\sqrt{2}}\right) \frac{a_2 F_{J/\psi} m_{J/\psi}}{\sqrt{2E_{J/\psi}(2\pi)^3}} \left\{ \frac{(\vec{P}^{\,2} + E_{J/\psi} E_{\bar{K}^*})}{m_{\bar{K}^*} m_{J/\psi}} \times \right.$$

$$\left. (m_{\bar{B}} + m_{\bar{K}^*}) A_1(q^2) - 2m_{\bar{B}}^2 \vec{P}^2 A_2(q^2)/(m_{\bar{K}^*} m_{J/\psi}(m_{\bar{B}} + m_{\bar{K}^*})) \right\} \qquad (8)$$

$$\mathcal{M}\left(\begin{smallmatrix} + \\ (-) \end{smallmatrix}, \begin{smallmatrix} + \\ (-) \end{smallmatrix}\right) = \left(-\frac{G_F}{\sqrt{2}}\right) \frac{a_2 F_{J/\psi} m_{J/\psi}}{\sqrt{2E_{J/\psi}(2\pi)^3}} \times$$

$$\left[\begin{smallmatrix} + \\ (-) \end{smallmatrix} \left(\frac{2m_{\bar{B}}}{m_{\bar{B}} + m_{\bar{K}^*}}\right) |\vec{P}| V(q^2) + (m_{\bar{B}} + m_{\bar{K}^*}) A_1(q^2) \right] \qquad (9)$$

where $\vec{P} \equiv (\vec{P}_{\bar{K}^+} - \vec{P}_{J/\psi})/2$ in the $\bar{B}$ rest frame. Standard methods may now be used to conclude that

$$r = 2(4m_{\bar{B}}^2/(m_{\bar{B}} + m_{\bar{K}^*})^2)\vec{P}^2|V|^2 \Bigg/$$

$$\left[\left| \left(\frac{\vec{P}^2 + E_{\bar{K}^*} E_{J/\psi}}{m_{\bar{K}^*} m_{J/\psi}}\right) (m_{\bar{B}} + m_{\bar{K}^*}) A_1 \right. \right.$$

$$\left. -2m_{\bar{B}}^2 \vec{P}^{\,2} A_2/(m_{\bar{K}^*} m_{J/\psi}(m_{\bar{B}} + m_{\bar{K}^*})) \right|^2$$

$$\left. +2(4m_{\bar{B}}^2/(m_{\bar{B}}+m_{\bar{K}^*})^2)\vec{P}^2|V|^2 + 2(m_{\bar{K}^*}+m_{\bar{B}})^2|A_1|^2 \right] . \qquad (10)$$

On substituting in (10) for the A_i and V from the BSW relativistic harmonic oscillator model, we find

$$r \simeq .091 \tag{11}$$

This is encouraging. We note also that $\Gamma(00)/\Gamma(All) \simeq .57$; the latter result means that the original idea of Kayser[4] to focus on $\Gamma(00)$ in the K^* decay is not completely excluded. We understand, however, that such an isolation of $\Gamma(00)$ is difficult. ($\Gamma(00)$ is the rate associated with $\mathcal{M}(0,0)$.)

Nonetheless, it is necessary to recall that the BSW framework is a model. The question naturally arisises as to the sensitivity of (11) to the details of the model. Here, we address this question by considering other models. Thus, we turn next to the methods of Lepage and Brodsky for the analysis of r.

3. Lepage-Brodsky Approach to $\bar{B}^0 \to J/\psi \bar{K}^0$

In this Section, we present the prediction of the methods of Lepage and Brodsky[7] for the process $\bar{B}^0 \to J/\psi \bar{K}^0$ insofar as the quantity r is concerned. We emphasize that these methods are relativistically invariant and represent perturbative QCD corrections to the spectator model of the respective transition; in fact, this accounts for the response of the would be spectator the decay of the b quark. We begin by recalling the elements of the Lepage-Brodsky theory.

Specifically the method of Lepage and Brodsky is explicitly illustrated by Figure 1 if we use for example their distribution amplitudes for the incoming and outgoing hadrons and follow their calculational methods. The transition will be handled in the framework of the effective Lagrangian in (1); this means that we continue to take the point-like limit of the W propagator in Figure 1. We will always do this. Hence, we can interpret the prediction of the Lepage-Brodsky theory for our process in Figure 1 in terms of its implications for V, A_1 and A_2 in (5). On doing this here, we simply adapt our results in the Lepage-Brodsky theory for the $D \to K^*$ transition in Ref. 9 to $B^0 \to K^*$; thus we find the distribution amplitudes

$$\begin{aligned}
\phi_{K^0_{''}}(y_1, y_2) &= \sqrt{3} f_{K^*} y_2(-.344 y_1 + 2.69 y_1^2) \\
\phi_{K^0_{\perp}}(y_1, y_2) &= \sqrt{3} f_{K^*} y_2(-.609 y_1 + 2.76 y_1^2) \\
f_{K^*} &\simeq .175\text{GeV} \\
\phi_B(x_1, x_2) &= (f_B/\sqrt{12})\delta(x_2 - x_B),
\end{aligned}$$

where

$$\begin{aligned}
x_B &= (m_d - (m_b/(m_b + m_d))(m_b + m_d - m_B))/m_B \\
&\simeq .036 \qquad (12)
\end{aligned}$$

and f_B is the B meson decay constant (its precise value is not important for present purposes, but, for definiteness, we note that potential model estimates would suggest $f_B \simeq .136$ GeV). In arriving at x_B, the masses $m_u = m_d \simeq .33$ GeV, $m_b \simeq 5.1$ GeV and $m_B = 5.28$ GeV have been used. On repeating the steps in our analysis of $D \to K^*$ in Ref. 9 with the results (12) for the $B^0 \to K^*$ transition, we get the predictions (the superscript on A_i and V denote the K^* polarization)

$$\begin{aligned}
A_1^{\perp('')} &= (m_B g_A/(m_B + m_{K^*}))F^{(1)}_{\perp('')} \\
A_2^{''} &= (m_B + m_{K^*}) m_B g_A F^{(3)}_{''} \\
V^{\perp} &= \frac{1}{2} m_B g_V (m_B + m_{K^*}) F^{(2)}_{\perp} \qquad (13)
\end{aligned}$$

where the functions $F^{(i)}_H$, $H = \perp, ''$, are the generalizations to $B^0 \to K^*$ of the corresponding functions in $D \to K^*$ in Ref. 9 and will be presented in detail elsewhere.[10] Here, the $g_{A,V}$ are the usual axial-vector and vector form factors which we treat with the standard pole approximation from the BSW model for example. In this way, we get the predictions

$$r \simeq .083 \qquad (14)$$

and

$$\Gamma(00)/\Gamma(ALL) \simeq .834\ . \qquad (15)$$

Again, the value which we find for r is encouraging and we see further that $\Gamma(00)$ is substantially enhanced in the Lepage-

Brodsky theory. We turn next to the potential model of Isgur, *et al.*

4. Model of Isgur, et.al.

Here, we consider the Cornell-type potential model of Isgur, *et al.*[6] with the idea of illustrating what one expects in a non-relativistic potential model for our ratio r. Again, we make contact with the model by focussing on its predictions for the form-factors A_i and V in (5).

Specifically following the methods in Ref. 7, we may identify the form factors A_i and V as

$$A_1 = -2\tilde{m}_B F_3/(m_B + m_{K^*}) ,$$

$$A_2 = -\frac{(m_B + m_{K^*})}{2\tilde{m}_{K^*})} F_3 \left[1 + \frac{m_d}{m_b}\left(\frac{\beta_B^2 - \beta_{K^*}^2}{\beta_B^2 + \beta_{K^*}^2}\right) - \frac{m_d^2}{4\mu_- \tilde{m}_B} \frac{\beta_{K^*}^4}{\beta_{BK^*}^4}\right] ,$$

and

$$V = \frac{-(m_B + m_{K^*})}{2} F_3 \left[\frac{1}{m_s} - \frac{1}{2\mu_-} \frac{m_d}{\tilde{m}_{K^*}} \frac{\beta_B^2}{\beta_{BK^*}^2}\right] , \tag{16}$$

where

$$F_3(t) = (\tilde{m}_{K^*}/\tilde{m}_B)(\beta_B \beta_{K^*}/\beta_{BK^*}^2)^{1/2} \times \exp\left[-\left(\frac{m_d^2}{4\tilde{m}_B \tilde{m}_{K^*}}\right)\left(\frac{t_m - t)}{\kappa^2 \beta_{BK^*}^2}\right)\right] \tag{17}$$

with

$$\begin{aligned} &\mu_- = (1/m_s - 1/m_b)^{-1}, \ \kappa = .7, \\ &\tilde{m}_B = m_b + m_d, \ \tilde{m}_{K^*} = m_s + m_d , \\ &\beta_B = .41, \beta_{K^*} = .34, \ \beta_{BK^*}^2 = \frac{1}{2}(\beta_B^2 + \beta_{K^*}^2) , \\ &m_b = 5.12 \text{ GeV}, m_s = .55 \text{ GeV}, m_d = .33 \text{ GeV} . \end{aligned} \tag{18}$$

Here,

$$t = m_{J/\psi}^2 \text{ and } t_m = (m_B - m_{K^*})^2 .$$

In this way, we get

$$r = \Gamma(CP = odd)/\Gamma(ALL) = .516 \tag{19}$$

and

$$\Gamma(00)/\Gamma(ALL) \simeq .064 \tag{20}$$

Thus, not unexpectedly, the non-relativistic model disagrees with the relativistically invariant methods of BSW and Lepage and Brodsky in the ratio of $\Gamma(CP = odd)$ to Γ(all). Indeed, the dominance of the CP even modes means that the D-wave and relativistic corrections to the S-wave are indeed important: the non-relativisitic model will naively tend to omit both of these effects. Here, we have corrected for the kinematical aspects of the relativistic corrections via the manifestly covariant representation in (5). The dynamical aspect of these corrections then is still significant. In another way of looking at this matter, we may say that even when we use a model which suppresses the kind of large values of t as we clearly have in $B^0 \rightarrow J/\psi K^*$, we still find that $\Gamma(CP = odd)/\Gamma(ALL) \sim .52$. Hence, there is a strong possibility that the further relativisitic effects, which may enhance the S and D waves for example, but, due to the invariance of $\varepsilon(\pm)$ to boosts along $\bar{P}_{K^*}$, which should have a reduced effect on the P-wave, do indeed suppress $\Gamma(CP = odd)$ compared to $\Gamma(CP = even)$.

To look more deeply into the expected value of r, we now turn to the more model independent methods of Isgur and Wise and Bjorken. This we do in the next Section.

5. WISGUR-B_j Theory

The basic idea of this approach to B decay is to work in the limit of infinite heavy quark masses. Here, this would mean infinite values of m_b and m_s for fixed velocities $\vec{v}_b$ and $\vec{v}_s$. Already one has to ask an immediate question: How far is m_s from infinity?

The situation is not entirely hopeless if one expands our amplitude in the constitutent quark masses; in this case, the dominant corrections of size $\Lambda_{QCD}/m_s \sim 100$ MeV/500 MeV $= 20\%$ are expected. Thus, we may expect that WISGUR$-B_j$ theory might work at the 20% level in the amplitudes for $B^0 \rightarrow J/\psi K^{*0}$. Let us keep this expectation in mind. It means that we cannot hope to prove anything in precise detail for our particular B decay but that 20% amplitude statements may be possible. We will now proceed with the analysis of the WISGUR-B_j theory prediction for our ratio r.

First, we note that from the standpoint of the Lepage-Brodsky theory, the diagrams in Figure 2 allow us derive the

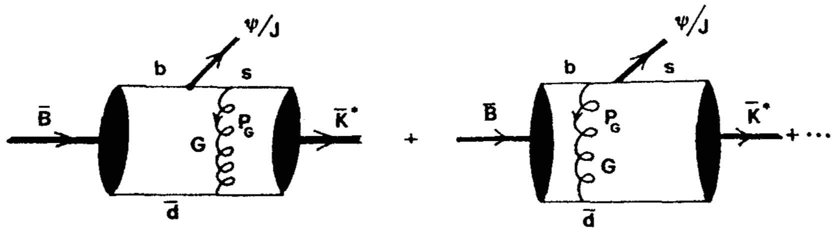


Figure 2. *Lepage-Brodsky approach to $\bar{B} \to J/\psi \bar{K}^*$. G is a QCD gluon of four momentum P_G.*

result of Bjorken for our decay. Specifically, in the WISGUR-B_j limit, so long as $\vec{v}_b$ and $\vec{v}_s$ are such that the gluon 4-momentum transfer squared in Figure 2, P_G^2, satisfies

$$-P_G^2 \lesssim 1/r_{hadron}^2 \tag{21}$$

where $r_{hadron} \sim 1$ fm is the typical hadron size, that particular gluon is already included in the Lepage-Brodsky distribution amplitudes, as we show in Figure 3, due to the presence of the collinear projection operator $\mathcal{P}_Q$ as we illustrate in Figure 3. Such small momentum transfer glue is all collected by $\mathcal{P}_Q$ to form the respective Lepage-Brodsky distribution amplitude $\phi_{K^*}(x_1, x_2, Q)$ at scale $Q \sim \{m_B, m_{J/\psi}\}$ in our transition. Thus from the Lepage-Brodsky equations for the distribution amplitudes ϕ_B and ϕ_{K^*} we see that, in the WISGUR-B_j limit, the matrix element $\mathcal{M}_\mu$ in (5) becomes, when $|P_G^2| \leq 1/r_{hadron}^2$,

$$\tilde{\mathcal{M}}_\mu = \int d^2x \int d^2y \delta(1-x_1-x_2)\delta(1-y_1-y_2)\phi_{K^*}(y)\phi_B(x)\times$$

$$tr\left[\frac{\not{\epsilon}^*_{K^*}(\not{P}_{K^*}+m_{K^*})}{\sqrt{2}}(\gamma_\mu g_V - \gamma_\mu\gamma_s g_A)\frac{\gamma_5(\not{P}_B - m_B)}{\sqrt{2}}\right] \tag{22}$$

which is precisely the result of Bjorken (and Georgi)[8] for such matrix elements. The advantage of (22) is that it already includes QCD corrections in ϕ_{K^*} and ϕ_B. The issue of whether or not P_G^2 satisfies (21) is of course delicate. Specifically, we have

$$(.1973)^2 \mathrm{GeV}^2 \gtrsim -P_G^2 \simeq 2x_2 y_2 P_B \cdot P_{K^*} = .687 y_2 \mathrm{GeV}^2$$

or

$$y_2 \lesssim .057 \tag{23}$$

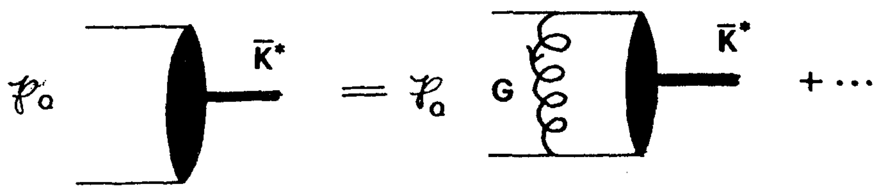


Figure 3. *Lepage-Brodsky equation for the $\bar{K}^*$ distribution amplitude. $\mathcal{P}_Q$ is the collinear projection operator for scale Q.*

From (12) it seems that such values of y_2 are not favored. In fact, when P_G^2 violates (23), the gluon propagator in Figure 2 has the behaviour $1/P_G^2$ so that the region $y_2 \leq .057$ would correspond to $\sim$ 20% of the total contribution of the y integration in (12) for both $\phi_{K^0_{\parallel}}$ and $\phi_{K^0_{\perp}}$. Thus we expect that (22) may have sizable (bound-state) QCD corrections.

Indeed, when we compute the values of A_i and V which follow from (22) we get

$$V = \left(\frac{m_{K^*} + m_B}{2}\right) g_V, \quad A_1 = -\frac{(E_{K^*} + m_{K^*})}{m_B + m_{K^*}} m_B g_A\,,$$

and

$$A_2 = -(m_B + m_{K^*}) g_A / 2 \tag{24}$$

so that we find

$$r = \Gamma(CP = odd)/\Gamma(ALL) \simeq .196 \tag{25}$$

and

$$\Gamma(00)/\Gamma(ALL) \simeq .398 \tag{26}$$

Thus, the comparison of (25) and (26) with (14) and (15) is consistent with the existence of sizeable QCD corrections to the WISGUR-B_j theory for the $B^0 \to K^*$ transition. This is consistent with general expectations.

Nonetheless, it is reassuring that in spite of the relatively large QCD-binding corrections, the results (25) and (26) still indicate that Γ(CP=odd) is a small fraction of Γ(ALL). Thus, they support the use of $B_d^0 \to J/\psi K^{*0}$ as a mode for B Factory exploitation in *CP* violation experimentation.

6. Conclusions

In this presentation we have explored dynamical models of the process $B^0 \to J/\psi K^{*0}$ from the standpoint of *CP* violation experiments at a B Factory type e^+e^- colliding-beam device. We have been encouraged by what we have found, in general.

Specifially, we have considered four different views of the dynamics of such $B_d^0 \to J/\psi K^{*0}$ decays: the model of Wirbel, Stech and Bauer; the model of Isgur, *et al.*; the methods of Lepage and Brodsky; and, the recent approach of Wise, Isgur and Bjorken (WISGUR-B_j theory). As we expected, the relativistic nature of the K^{*0} in the transition causes the Isgur, *et al.* model to deviate substantially from the other 3 approaches in its prediction for $\Gamma(CP = odd)/\Gamma(ALL) \equiv r$ for the decay in question. The other 3 approaches all give $r < 0.2$, whereas the non-relativistic Isgur, *et al.* model gives $r \simeq 0.5$. Thus, we feel we have found strong support for the use of $B_d^0 \to J/\psi K^{*0}$ decays as an additional model for exploration of *CP* violation in B Factory type colliding-beam devices. Such modes may apparently be used to reduce the luminosity requirement for such machines by a factor of ~ 2.5 relative to estimates based on the use of the mode $B_d^0 \to J/\psi K_S^0$ alone.[11] This, then, is the primary conclusion of our theoretical analysis.

Recently[12] we have learned that the Argus Collaboration has measured the polarization of the J/ψ in the decays $B_d^0 \to J/\psi K^{*0}$. What they find is completely consistent with our result that $\Gamma(CP = even)$ dominates the decay. This then gives experimental support to the main conclusion of our work.

ACKNOWLEDGMENTS

The author is grateful to A. Snyder and to the other members of the SLAC B Factory Physics and Simulation Working Group for many useful discussions. The author also thanks Prof. Abdus Salam for the kind hospitality and support of the ICTP where this manuscript was completed.

References

1. See, for example, SLAC-353, these *Proceedings* and references therein.
2. See, for example W. Wirbel *et al.*, *Z. Phys.* **C29**, 637 (1985).
3. Particle Data Group, *Phys. Lett.* **B204** (1988).
4. See for example, B. Kayser, NSF preprints, 1990 and references therein.
5. See, for example, I. Dunietz, *et al.*, these *Proceedings* and references therein.
6. N. Isgur, *et al.*, *Phys. Rev.* **D39** , 799 (1989).
7. G.P. Lepage and S. Brodsky, *Phys. Rev.* **D22**, 2157 (1980).
8. N. Isgur and M. Wise, *Phys. Rev. Lett.* **66**, 1130 (1991); J.D. Bjorken, *Proceedings of the 18th SLAC Summer Institute on Particle Physics* (1990).
9. B.F.L. Ward, *Il Nuovo Cimento* **98A**, 401 (1987).
10. B.F.L. Ward, to appear.
11. H. Schroder, talk presented at the 1990 International Conference on High Energy Physics (Singapore, 1990).
12. A. Snyder, private communication, 1990.

Monte Carlo study of the $B^0 \to \psi K_S^0$ Mode and $B^0/\bar{B}^0$ Tagging

S. Komamiya

1. Introduction

A Monte Carlo study of the ψK_S^0 channel and B-tagging has been performed using the ASLUND Monte Carlo program for event generation and detector simulation. This study is totally independent of the previous study as described in SLAC-353.[1] The conclusions of the previous study are mostly confirmed.

2. Event Selection

Assuming a branching fraction B($B^0 \to \psi K_S^0$) $= 4 \times 10^{-4}$, the number of observed events of the type

$$e^+e^- \to \Upsilon(4S) \to B^0\bar{B}^0 \to \psi K_S^0 + X \quad (\psi \to \ell^+\ell^-),$$

for a luminosity of 30 fb^{-1}, is

$$\begin{aligned}
N_{\psi K_S^0} \times \epsilon_{\psi K_S^0} & \\
&= (\int L dt) \times (\sigma_{B\bar{B}}) \times \frac{B^0\bar{B}^0}{B\bar{B}} \times \frac{B^0 + \bar{B}^0}{B^0\bar{B}^0} \\
&\quad \times Br(B^0 \to \psi K_S^0) \times Br(\psi \to \ell^+\ell^-) \times \epsilon_{\psi K_S^0} \\
&= (30 \times 10^6 \text{ nb}^{-1}) \times 1.2 \text{ nb} \times 0.50 \times 2 \\
&\quad \times (4 \times 10^{-4}) \times 0.14 \times \epsilon_{\psi K_S^0} \\
&= 2016 \times \epsilon_{\psi K_S^0}
\end{aligned}$$

where ℓ is e or μ and $\epsilon_{\psi K_S^0}$ is the detection efficiency of the ψK_S^0 events.

The Monte Carlo simulation is based on the ASLUND package. We assume asymmetric energy collisions at $\sqrt{s} = M_{\Upsilon(4S)}$ with the higher beam energy of 9 GeV. The polar angle acceptance cut for charged particles and photons is

$|\cos\theta| < 0.95$. Energy resolution, momentum resolution or position resolution of each detector component are assumed to be just the same as those used in SLAC-353 (*pp.* 36-49), unless the values are explicitly given.*

The selection criteria for events with $B \to \psi K^0_S$, followed by short comments on each criterion, are listed below.

(1) $\psi \to \ell^+\ell^-$ are selected by the cut $|M_{\ell^+\ell^-} - M_\psi| < 0.05$ GeV. The detection efficiency of $\psi \to \ell^+\ell^-$, taking into account lepton identification probabilities, is 78%. The M_ψ resolution is 25 MeV (FWHM) (see Figure 1).

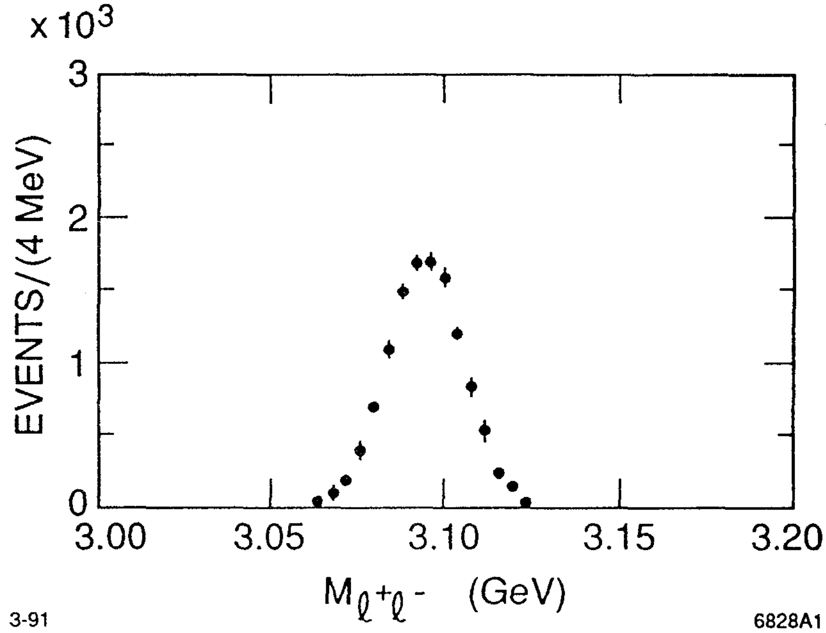


Figure 1. *J/ψ mass resolution for the $J/\psi \to \ell^+\ell^-$ decay mode.*

(2) $K^0_S \to \pi^+\pi^-$ are selected by the cut $|M_{\pi^+\pi^-} - M_{K^0_S}| < 0.02$ GeV (see Figure 2), $l_{xy} > 0.002$ m (see Figures 3a and 3b) and $|z^+ - z^-| < 0.005$ m (see Figures 4a and 4b), where l_{xy} is the reconstructed K^0_S decay length projected onto the r-ϕ plane and $z^\pm$ is the reconstructed z-coordinate for $\pi^\pm$ at the crossing point of the two tracks. The $M_{K^0_S}$ resolution is 5 MeV (FWHM) and the background is negligible after the cuts (nuclear interactions in the beam pipe and surrounding material are not simulated, but they are estimated to be low).

* The analysis of the ψK^0_S mode in SLAC-353 is based on the Monte Carlo generator and detector simulation described in Ref. 2 , not on the ASLUND package which was used for the other studies in SLAC-353.

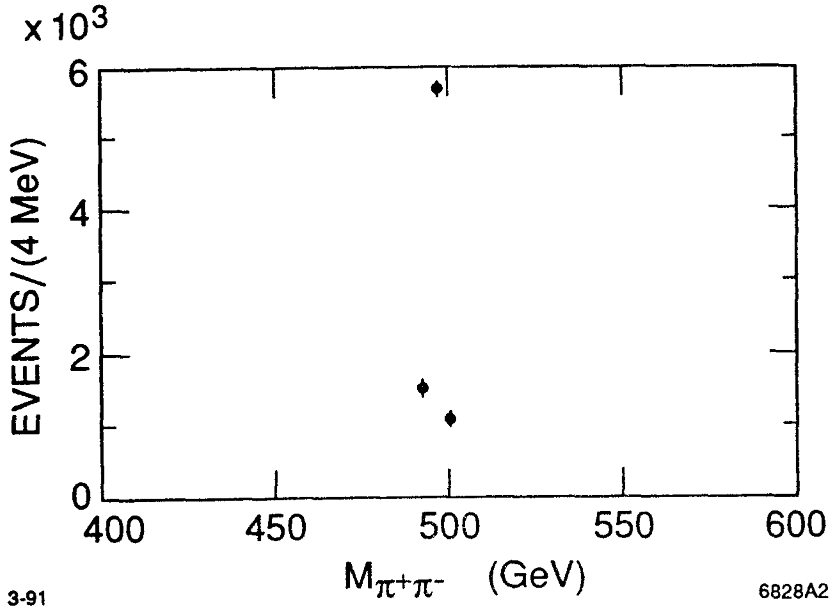


Figure 2. *K_S^0 mass resolution for the $K_S^0 \to \pi^+\pi^-$ decay mode.*

(2′) $K_S^0 \to \pi^0\pi^0$ are selected by the cut $|M_{\pi^0\pi^0} - M_{K_S^0}| < 0.03$ GeV, where π^0's are selected by $|M_{\gamma\gamma} - M_{\pi^0}| < 0.02$ GeV. To reduce combinatorial background, $p_t(K_S^0) > 0.25$ GeV is required, where $p_t(K_S^0)$ is measured with respect to the beam direction. The M_{π^0} resolution is 11 MeV (FWHM) and the signal/background ratio in the accepted mass range of $M_{\gamma\gamma}$ is 1.1 (see Figure 5). The $M_{K_S^0}$ resolution is 25 MeV (FWHM) and the signal-to-background ratio in the accepted mass range of $M_{\pi^0\pi^0}$ is 0.73 (see Figure 6). The K_S^0 peak is on the slope of the relatively high combinatorial background.

(3) For the charged decay mode of K_S^0 ($K_S^0 \to \pi^+\pi^-$), B^0 (or $\bar{B}^0$) $\to \psi K_S^0$ are selected by the cut $|M_{\psi K_S^0} - M_{B^0}| < 0.05$ GeV (see Figure 7a) and 0.22 GeV $< |\vec{p}^{\,*}_{\psi K_S^0}| < 0.45$ GeV (see Figure 7b), where $\vec{p}^{\,*}_{\psi K_S^0}$ is the momentum of ψK_S^0 system boosted back to the Υ(4S) CMS. The M_{B^0} resolution is 35 MeV (FWHM). After this cut, 0.3% of accepted events have more than one ψK_S^0 combinations. The overall efficiency of $B^0 \to \psi K_S^0 \to \ell^+\ell^- + \pi^+\pi^-$ is 55%.

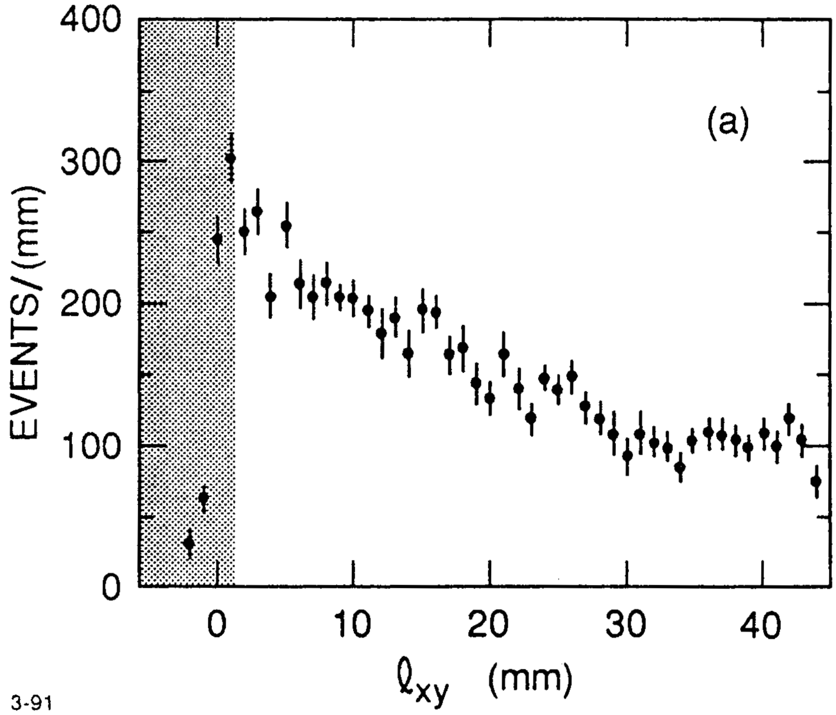


Figure 3a. *Distribution of the projection of the K_S^0 decay length in the transverse plane for the $K_S^0 \to \pi^+\pi^-$ decay mode signal, with the cut at $l_{xy} > 0.002$ m indicated.*

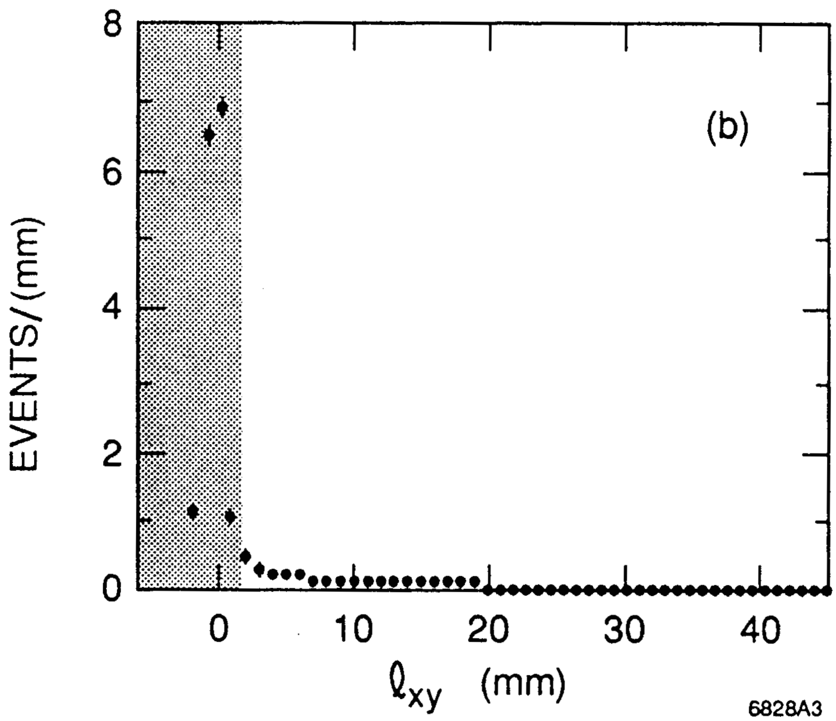


Figure 3b. *Distribution of the projection of the K_S^0 decay length in the transverse plane for the $K_S^0 \to \pi^+\pi^-$ decay mode background, with the cut at $l_{xy} > 0.002$ m indicated.*

(3′) For the neutral decay mode of K_S^0 ($K_S^0 \to \pi^0\pi^0$), B^0 (or $\bar{B}^0$) $\to \psi K_S^0$ are selected by the cut $|M_{\psi K_S^0} - M_{B^0}| < 0.15$ GeV (see Figure 8a] and $0.22 \text{ GeV} < |\vec{p}^{\,*}_{\psi K_S^0}| < 0.45$ GeV (see Figure 8b), where $\vec{p}^{\,*}_{\psi K_S^0}$ is the momentum of the ψK_S^0 system boosted back to $\Upsilon(4S)$ CMS. The M_{B^0} resolution is 66 MeV (FWHM). After this cut, 19% of accepted events have more than one ψK_S^0 combination. This is because a ψ is combined with several combinatorial K_S^0 and it falls near M_{B^0}. Since this background is due to combinatorial including particles from real ψK_S^0, non-combinatorial background is actually very low. The overall efficiency of $B^0 \to \psi K_S^0 \to \ell^+\ell^- + \pi^0\pi^0$ is 54%.

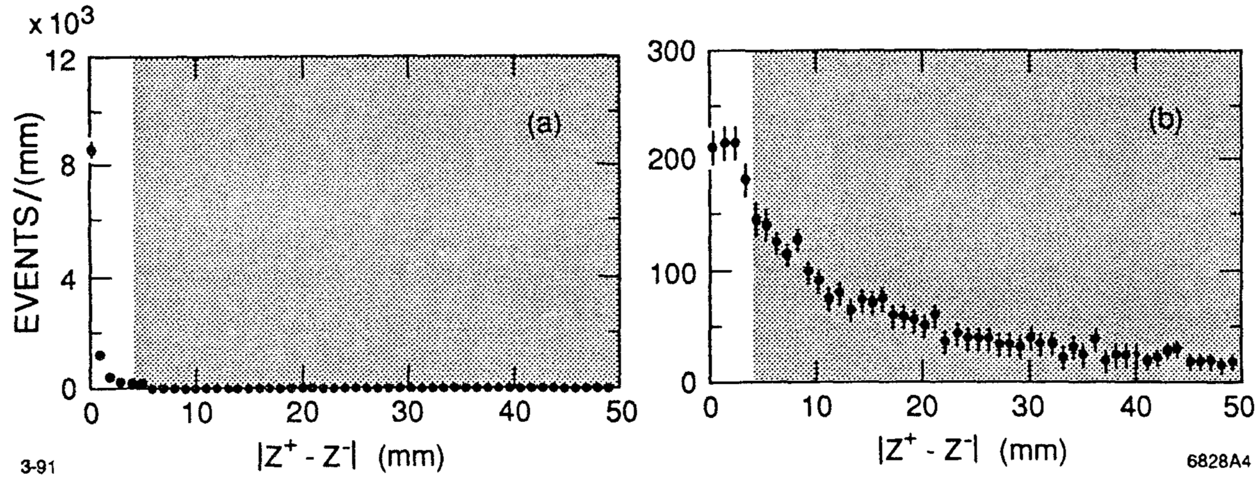


Figure 4
a. *Distribution of the z coordinate differences for the $K_S^0 \to \pi^+\pi^-$ decay mode signal, with the cut at $|z^+ - z^-| < 0.005$ m indicated.*
b. *Distribution of the z coordinate differences for the $K_S^0 \to \pi^+\pi^-$ decay mode background, with the cut at $|z^+ - z^-| < 0.005$ m indicated.*

The combined overall efficiency of $B^0 \to \psi K_S^0$ is 55%. The main background comes from ψK^{*0} with a decay of $K^{*0} \to K_S\pi^0$. $0.30 \pm 0.04\%$ of $B^0 \to \psi K^{*0}$ (with $\psi \to \ell^+\ell^-$) survive after all the cuts. More than 80% of this background is for the $K_S^0 \to \pi^0\pi^0$ decay, and hence background for the $K_S^0 \to \pi^+\pi^-$ decay is very low. Other background sources like $B^+B^- \to \psi K_S^0 + X$ are estimated to be small. No background events are found in the sample of 2×10^6 $B^0\bar{B}^0$ plus B^+B^- events (excluding events with ψK_S^0) or in the sample of 2×10^6 $u\bar{u} + d\bar{d} + s\bar{s} + c\bar{c}$ events.

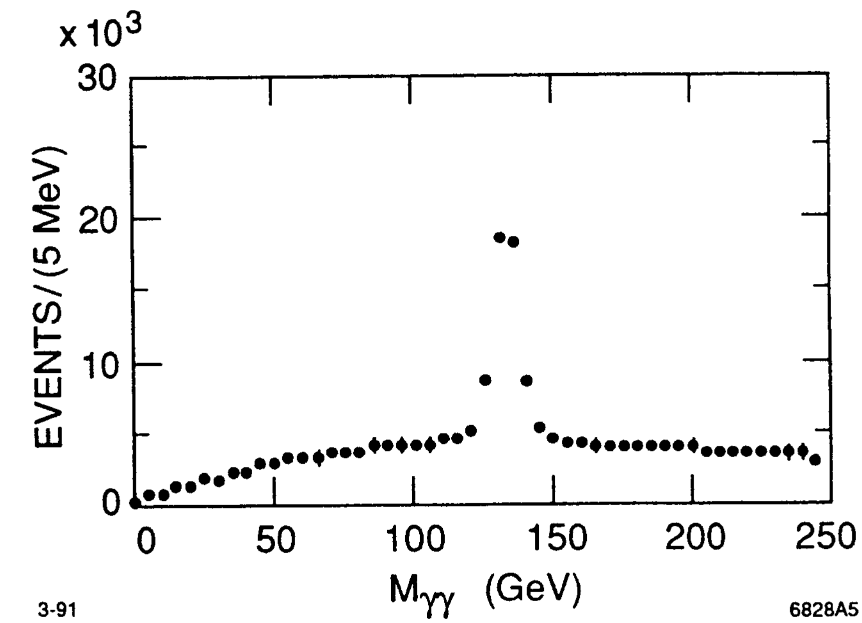


Figure 5. *π^0 mass resolution for π^0's from $K_S^0 \to \pi^0\pi^0$ decay.*

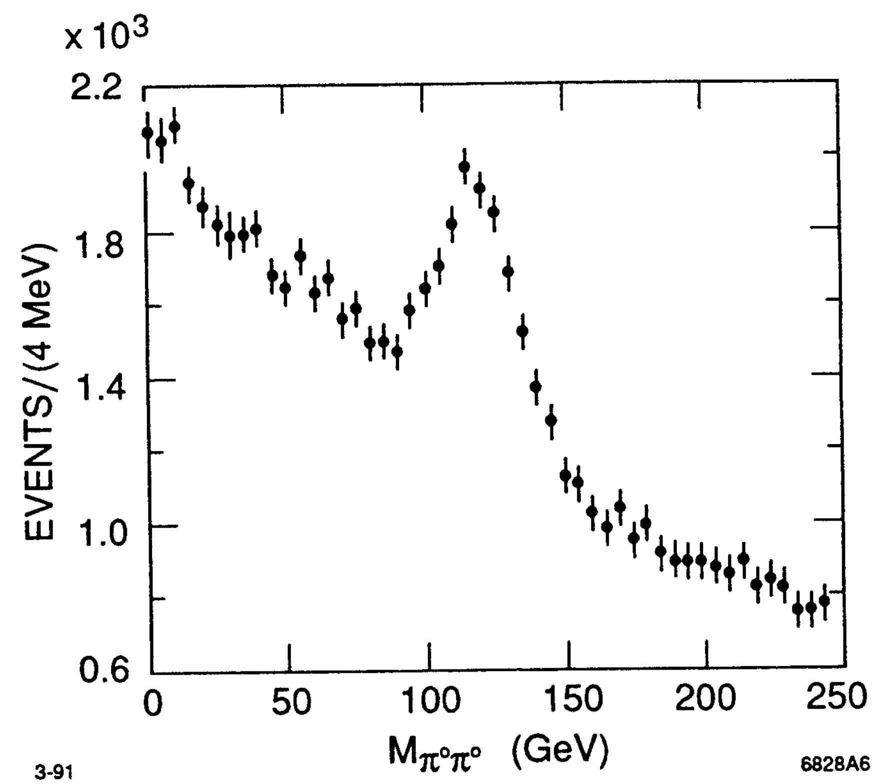


Figure 6. *K_S^0 mass resolution in $K_S^0 \to \pi^0\pi^0$ decay.*

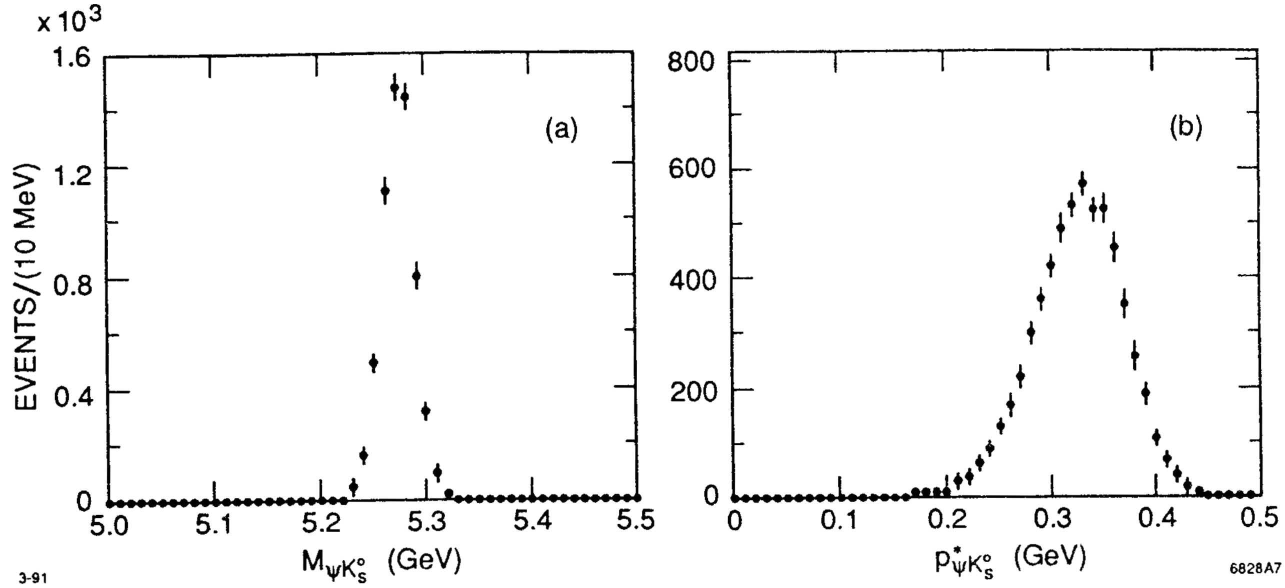


Figure 7.
a. *ψK_S^0 mass resolution, and*
b. *ψK_S^0 momentum in the $\Upsilon(4S)$ center-of-mass for $K_S^0 \to \pi^+\pi^-$ decay.*

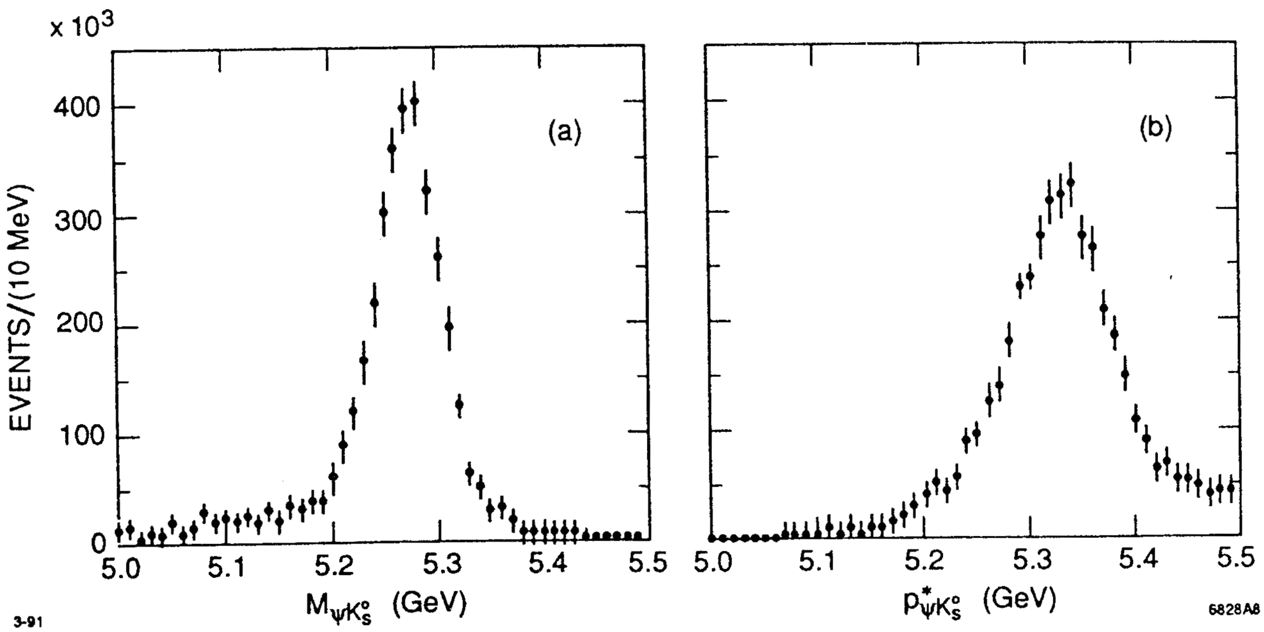


Figure 8.
a. *ψK_S^0 mass resolution, and*
b. *ψK_S^0 momentum in the $\Upsilon(4S)$ center-of-mass for $K_S^0 \to \pi^0\pi^0$ decay.*

The efficiency is a function of the boost factor and the detector acceptance. In Figure 9a, the overall efficiency is plotted as a function of higher beam energy and the maximum $|\cos\theta|$ of the detector acceptance (both for charged tracks and for photons). Similar plot for the partial efficiency of the charged mode

$$\frac{\text{Number of accepted events with } \psi K_S^0 \to \ell^+\ell^-\pi^+\pi^-}{\text{Number of generated events with } \psi K_S^0 \to \ell^+\ell^-\pi^+\pi^- \text{ or } \pi^0\pi^0}$$

is plotted in Figure 9b. As can be seen from the figure, we definitely need the coverage down to $|\cos\theta| = 0.95$, if the higher beam energy is more than 9 GeV.

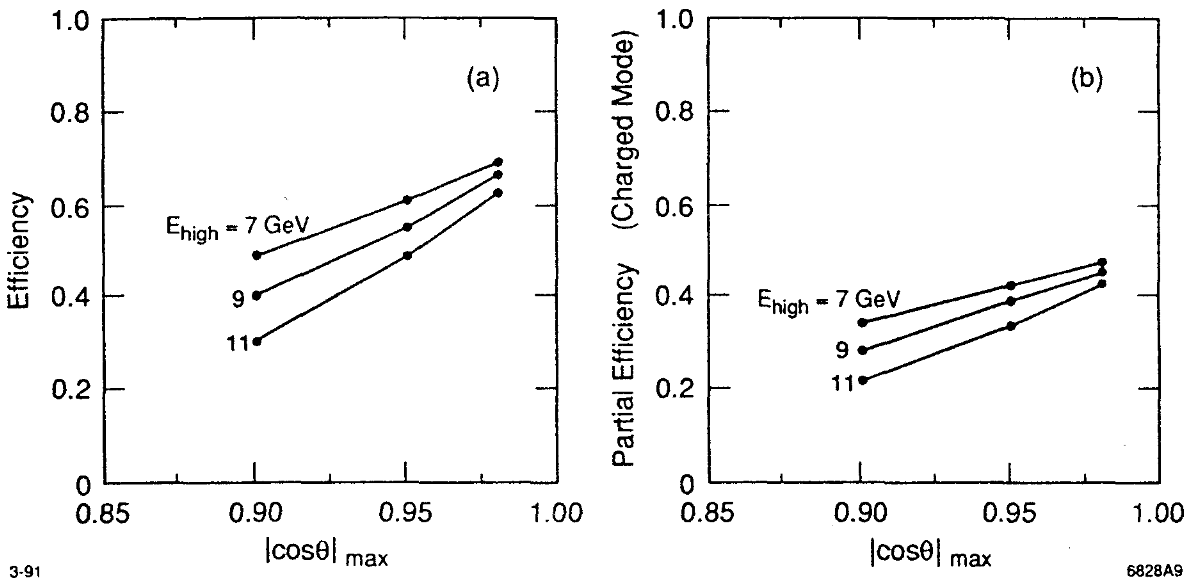


Figure 9.
a. *Overall ψK_S^0 detection efficiency as a function of $|\cos\theta|_{max}$ for three asymmetric configurations, and*
b. *ψK_S^0 detection efficiency for the charged pion mode only as a function of $|\cos\theta|_{max}$ for three asymmetric configurations.*

3. $B^0/\bar{B}^0$ TAGGING

All the particles except for the particles from the ψK_S^0 (*CP* eigenstate) are coming from the other B meson. High momentum leptons as well as charged kaons are used for sign determination of the tagged B. This problem is independent of the particular decay mode of *CP* eigenstate.

3.1. Lepton Tag

Leptons ($e^{\pm}$ or $\mu^{\pm}$) with momentum larger than 1.4 GeV in the $\Upsilon(4S)$ CMS are used for the tag. In Figures 10a, b and c, the momentum spectrum of leptons in the $\Upsilon(4S)$ CMS ($p^*_{\ell\pm}$) is plotted for primary leptons from B decay, for secondary

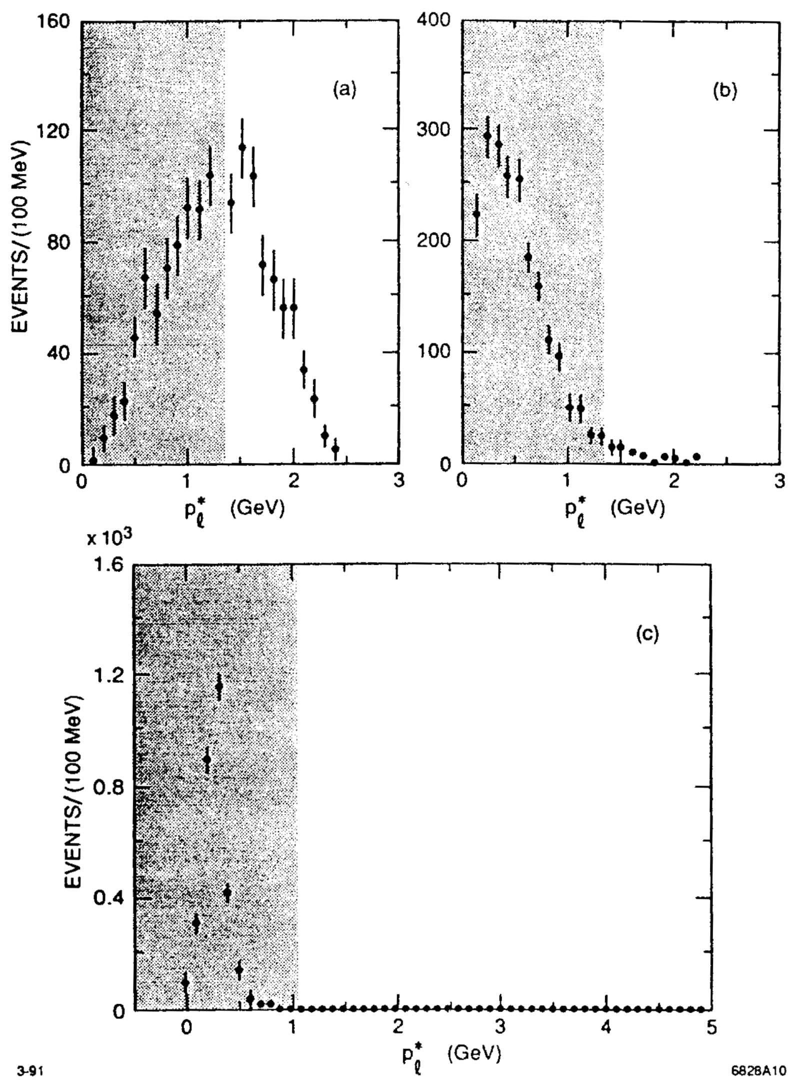


Figure 10. *Lepton spectra in the $\Upsilon(4S)$ CMS.*
a. *Primary leptons from B decay,*
b. *secondary leptons, and*
c. *leptons which do not come from B decay.*

leptons, and for leptons which do not come from B decay, respectively. Tagging efficiency (ϵ) for an event to have only one such lepton is 11.8%. Wrong-sign-tag probability ($w \equiv$ [number of wrong-sign-tag events]/[number of all the lepton-tag events]) is 4.4%. In SLAC-353, the lepton tag efficiencyis larger (14%) but the wrong tag probability is also larger (6%). The difference is partially because they assumed larger solid angle ($|\cos\theta| < 0.98$) and partially because they accept low momentum leptons (0.8 GeV$< p_\ell^* <$ 1.4 GeV) if there are more than one reconstructed vertices in an event. We stick to high momentum leptons because we can statistically gain very little if we also use the low momentum leptons, and we prefer simpler cuts in order to reduce the bias in the Δz measurement.

3.2. Kaon Tag

The sign of B is tagged if there is only one charged kaon or two same sign charged kaons in an event. For perfect kaon identification with an ideal detector, the tagging efficiencies after cuts are listed in Table 1.

Table 1. *The kaon tagging efficiency after cuts.*

Cut	Cut Definition	$\epsilon(K^\pm + K^\pm K^\pm)$	$w(K^\pm + K^\pm K^\pm)$
1	no cuts	0.473	0.065
2	$\lvert\cos\theta\rvert < 0.95$	0.428	0.078
3	$p_K > 0.2$ GeV	0.416	0.086
4	After K decay	0.340	0.100

We have also studied the tagging efficiency and wrong tag probability with a slightly more realistic detector simulation. For example, we use a combination of CRID (with index of refraction 1.277 corresponding to the kaon momentum threshold of 0.63 GeV) and dE/dx in the tracking chamber (assuming $\frac{\sigma_{dE/dx}}{dE/dx} = 0.075$). The following two criteria are used to identify charged kaons from other charged particles.

(1) If $p_K > 0.63$ GeV, CRID is used. The χ^2 probability to be kaon relative to the sum of χ^2 probabilities to be

$e^{\pm}$, $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$, p or $\bar{p}$ is larger than 0.5:

$$\frac{P_{CRID}(K)}{\sum_{i=e,\mu,\pi,K,p,\bar{p}} P_{CRID}(i)} > 0.5.$$

(2) If $p_K < 0.63$ GeV, and if no signal is seen in the CRID, a similar condition as for (1) is required for the dE/dx measurement:

$$\frac{P_{dE/dx}(K)}{\sum_{i=e,\mu,\pi,K,p,\bar{p}} P_{dE/dx}(i)} > 0.5.$$

The tagging probability after the above requirement is shown in the table with or without taking into account $K^{\pm}$ decay within the tracking detector. Each kaon is weighted by the decay probability within the outer wall of the tracking chamber. This is a rather conservative estimation since it might be possible to identify some fraction of the tracks with kinks due to the kaon decays in the chamber.

Modes	ϵ	w
$K^{\pm}$	0.399	0.083
$K^{\pm}K^{\pm}$	0.013	≤ 0.001
Sum	0.412	0.080

Table 2a. *The kaon tagging efficiency using CRID and dE/dx, when K decay is NOT taken into account.*

Modes	ϵ	w
$K^{\pm}$	0.316	0.101
$K^{\pm}K^{\pm}$	0.011	0.051
Sum	0.327	0.100

Table 2b. *The kaon tagging efficiency using CRID and dE/dx with K decay taken into account.*

The simulation of $K^{\pm}$ identification by CRID might be optimistic since the χ^2 is calculated from the resolution of the radius of the Čerenkov rings but background hits near the rings, overlap of the rings or dead space (boundaries) are not considered.

The tagging efficiency using kaons in SLAC-353 is 41%; the wrong tag probability is 6% (kaon decays are not considered).

In combining the lepton-tag and the kaon-tag, we used the lepton tag if it was available, and the kaon tag for the rest. Efficiencies are listed in Table 2c.

Table 2c. *The tagging efficiency using CRID and dE/dx The lepton tag is used if available. K decay is taken into account.*

Modes	ϵ	w
$\ell^{\pm}$ tag	0.118	0.044
$K^{\pm}$ and not $\ell - tag$	0.278	0.104
$K^{\pm}K^{\pm}$ and not $\ell - tag$	0.011	0.051
Sum	0.407	0.088

The overall tagging efficiency in SLAC-353 is 48% with a wrong tag probability of 8% (kaon decay is not taken into account), which is consistent with our numbers. The main difference is that charged kaon decay and low momentum leptons are not used in this analysis.

4. Measurement of Δz

The difference between the z coordinate of the decay vertex of the CP eigenstate (z_{CP}) and that of tag-B (z_{tag}) is approximately proportional to $t_{CP} - t_{tag}$, where t_{CP} and t_{tag} are the decay time of the CP eigenstate and the tag-B in the $\Upsilon(4S)$ CMS. The decay point of the CP eigenstate is reconstructed using the decay point of $\psi \to \ell^+\ell^-$. The

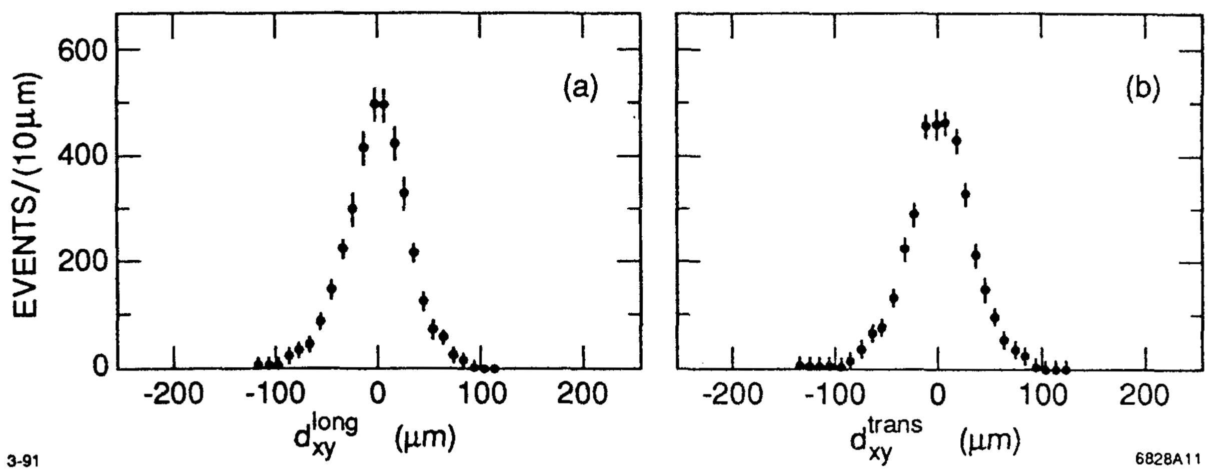


Figure 11. *The difference of the reconstructed ψ decay point and the decay point of the CP eigenstate in the $r-\phi$ plane,($\vec{d}_{xy}^{\,\ell^+\ell^-}$).*
a. *Component of $\vec{d}_{xy}^{\,\ell^+\ell^-}$ parallel to $\vec{p}_t^{\,\ell^+\ell^-}$, and*
b. *transverse component.*

difference of the reconstructed ψ decay point and the decay point of the CP eigenstate in the $r-\phi$ plane ($\vec{d}_{xy}^{\,\ell^+\ell^-}$) is shown in Figure 11. In the figure, the component of $\vec{d}_{xy}^{\,\ell^+\ell^-}$ parallel to $\vec{p}_t^{\,\ell^+\ell^-}$ is plotted in Figure 11a and the transverse component is plotted in Figure 11b. The distribution of the transverse component is wider because the opening angle of $\ell^+\ell^-$ is typically larger than 90° in the $r-\phi$ plane. For the z-component, the width of $\delta z_{CP} \equiv z_{\ell^+\ell^-}(reconstructed) - z_{CP}(true)$ is 55 μm FWHM ($\rightarrow$ $\sigma(\delta z_{CP}) = 24$ μm) as shown in Figure 12.

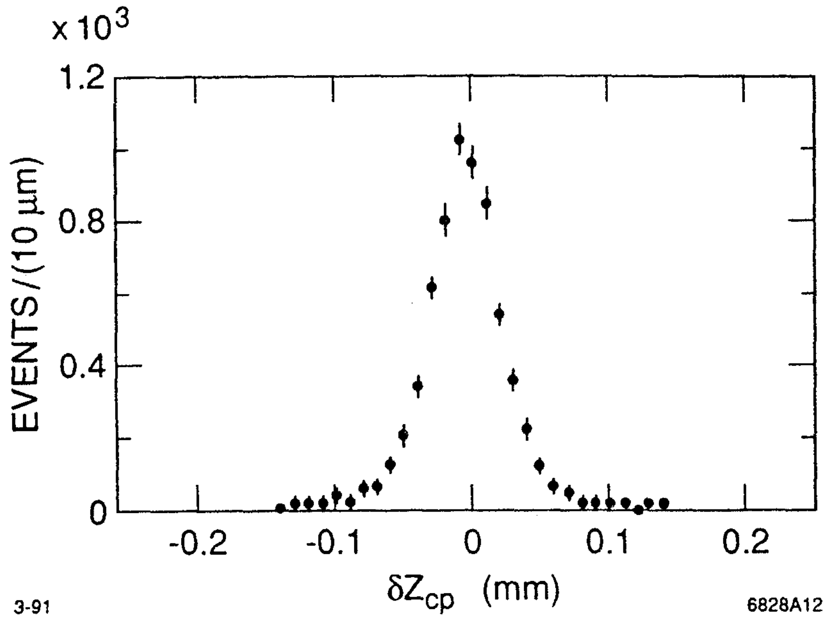


Figure 12. *The resolution of the z component of the reconstructed* J/ψ *vertex:* $\sigma = 24$ μm.

The tag-B vertex is reconstructed by fitting all the charged particle tracks with momentum larger than 0.3 GeV, except for 1) the particles coming from the CP eigenstate and 2) tracks with impact parameter larger than 800 μm from the reconstructed ψ point in $r-\phi$ plane. The impact parameter cut is necessary to get rid of tracks from K_S^0-, Λ- or $\bar{\Lambda}$-decays, which shift the tag-B vertex towards the boost direction (direction of the high energy beam). Since the $B^0\bar{B}^0$ oscillation time scale is rather slow, D-meson decay does not significantly affect the Δz ($\equiv z_{CP} - z_{tag}$) measurement. The width of $\delta\Delta z(\equiv \Delta z(reconstructed) - \Delta z(true))$ is 147 μm for FWHM ($\rightarrow$ $\sigma(\delta\Delta z) = 64$ μm). The distribution of $\delta\Delta z$ is shown in Figure 13. Note that we treat lepton-tag events and kaon-tag events in the same way in the calculation of Δz.

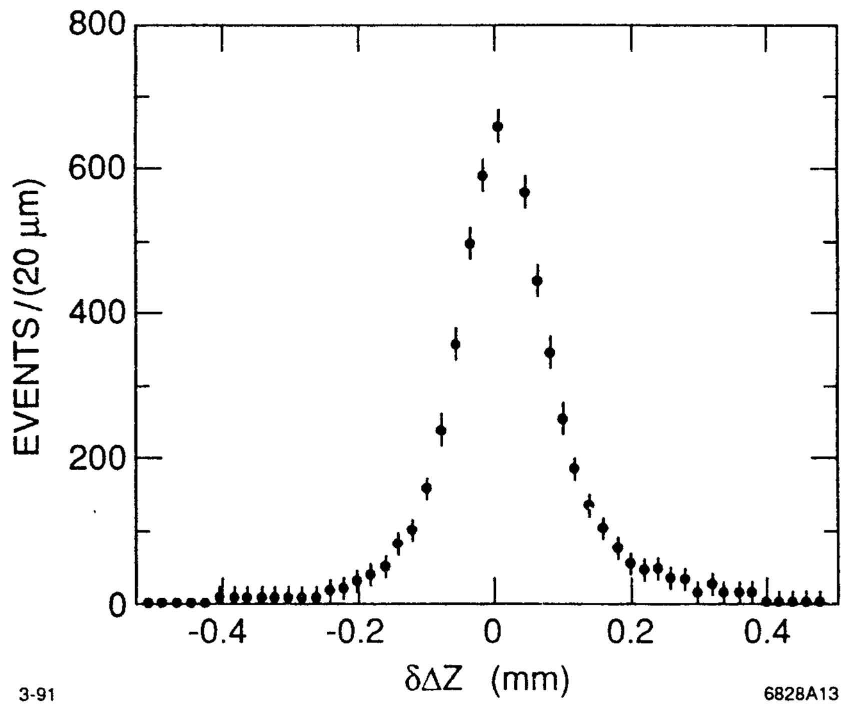


Figure 13. *The resolution in the difference of the reconstructed z coordinates of the tag and the CP eigenstate: $\sigma = 64\mu m$.*

5. Detector effects for ψK_S^0 mode

To study the allowed parameter range of the detector components, parameters are varied from the default values and the effects are investigated for

- detection efficiency of ψK_S^0,
- widths of δz_{CP} and $\delta\Delta z$,
- fraction of events with $|\delta\Delta z| > 400$ μm or > 800 μm.

5.1. *Vertex detector*

We assumed three layers of silicon strip vertex detector each with a thickness of 300 μm. The thickness was varied from 300 μm to 600, 1500 and 3000 μm in order to study the effects. The detection efficiency of ψK_S^0 does not significantly depend on the thickness as shown in Figure 14a. Effects for δz_{CP} and $\delta\Delta z$ are mild as shown in Figure 14b. The fraction of events with $|\delta\Delta z| > 400$ μm or > 800 μm increases rather rapidly with increasing the thickness due to multiple scattering (see Figure 14c).

The intrinsic resolution of the Si strip detector is assumed to be 10 μm (σ) for the default value. The effects due to the variation of the resolution are shown in Figure 15a, b and c. The width of $\delta\Delta z$ increases rapidly with increasing Si resolution, on the other hand the fraction of the tails is rather stable.

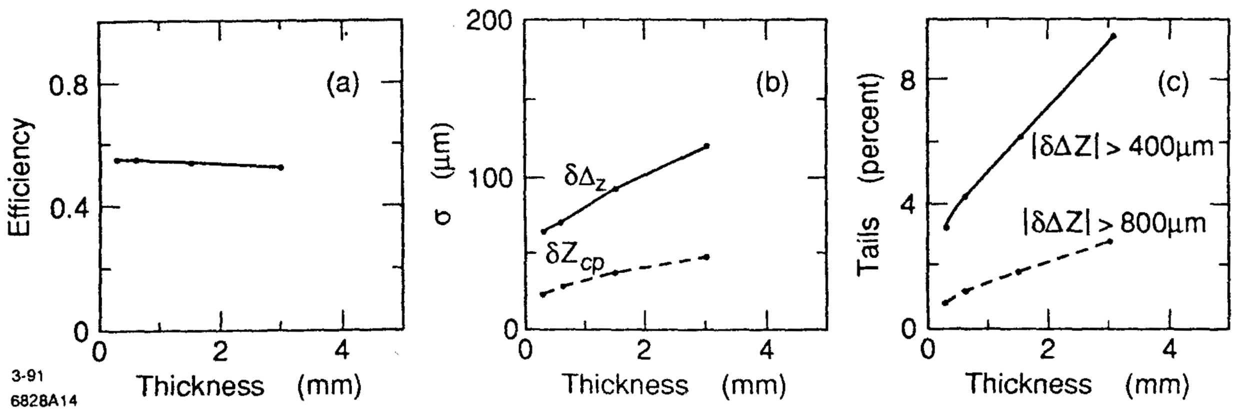


Figure 14. *Variation of vertex detector-related measurements with Si detector thickness.*
a. *efficiency,*
b. *$\delta\Delta z$, and*
c. *the percentage of tails.*

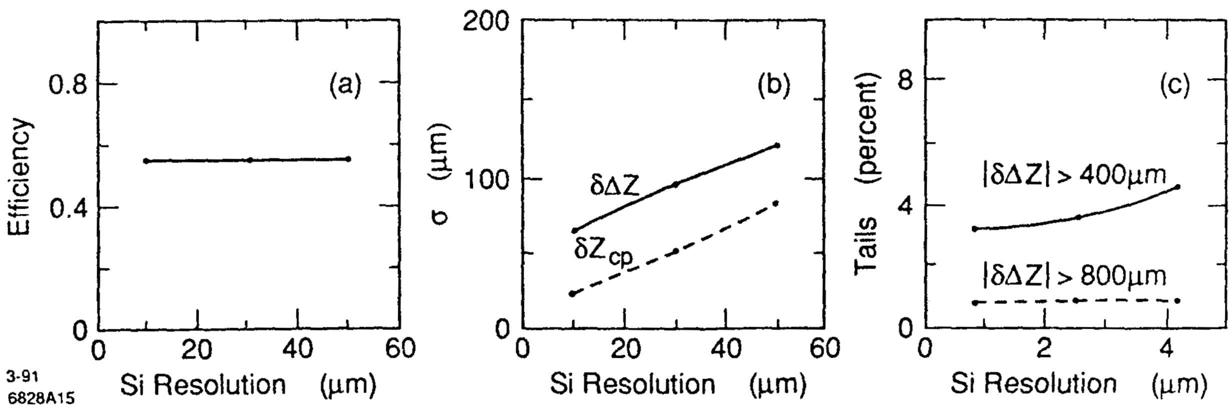


Figure 15. *Variation of vertex detector-related measurements with intrinsic resolution of the Si detector.*
a. *efficiency,*
b. *$\delta\Delta z$, and*
c. *the percentage of tails.*

5.2. *Drift Chamber*

Detection efficiency of the ψK_S^0 mode and Δz measurement are insensitive to both the intrinsic resolution of each wire (from 150 μm to 300 μm) and the radiation length of the gas (600 m/X_0 to 100 m/X_0), as shown in Figure 16.

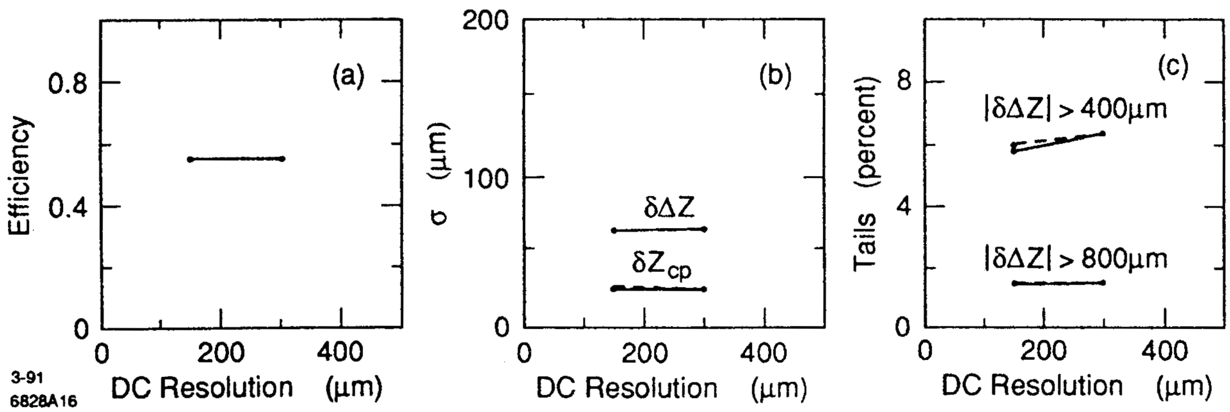


Figure 16. *Variation of* **a.** *ψK_S^0 detection efficiency and* **b.** *Δz with intrinsic wire resolution of the drift chamber* **c.** *variation of detection efficiency for $\psi K_S^0 \to \ell^+\ell^-\pi^0\pi^0$ with calorimeter resolution.*

5.3. *EM Calorimeter*

The partial efficiency of $\psi K_S^0 \to \ell^+\ell^-\pi^0\pi^0$ as a function of calorimeter resolution is plotted in Figure 16c. The horizontal axis of the plot is $\sigma_{intrinsic}$. Here the calorimeter resolution is assumed to be

$$\frac{\sigma_E}{E} = \sqrt{(\frac{\sigma_{intrinsic}}{E^{1/4}})^2 + (0.01)^2},$$

where E is the photon energy in GeV.

5.4. Particle Identification

We have studied several options for particle identification, especially for the kaon tag. We considered TOF counters, dE/dx measurement in the tracking chamber and CRID.

If we want to identify $K^\pm$ by TOF counters alone,* we need resolution better than $\approx$ 100 ps. For TOF counter of 75 ps resolution,† we can identify tag-$K^\pm$'s with more than 90% efficiency‡ with negligible background. If the resolution is 150 ps, the $K^\pm$-purity is only 60% for 90% kaon efficiency.

We cannot use dE/dx alone for the kaon identification. If the resolution is 7.5%, the $K^\pm$-purity is less than 60% for 90% kaon efficiency. dE/dx is useful for identifying low-energy kaons; hence it should be combined with other devices which can identify high-energy kaons, like CRID.

Our simulation of CRID is slightly optimistic, because effects of noise hits or dead spaces are not taken into account. We assume liquid radiator (C_6F_{14}) with index of reflection 1.277 (just as SLD CRID). Except for near the threshold region ($p_K \approx 0.63$ GeV), kaon identification and rejection of pions are almost perfect.

We have also studied the aerogel threshold Čerenkov detector. Indices of refraction between 1.0008 and 1.13 are now available, corresponding to threshold kaon momenta of 12.3 GeV and 1.01 GeV, respectively. The threshold momenta for π, K and p in the $B^0 \to \psi K_S^0$ events are plotted as a function of refractive index in Figure 17. The maximum kaon momentum is about 4 GeV with our energy asymmetry and number of high momentum protons are very small. We can cover the required momentum range for kaons with two indices: 1.008 and 1.06, which allows us to identify kaons with

---

* Here, we have not used dE/dx measurement.

† The averaged beam-beam crossing timing and the averaged position as well as the fluctuations must be known, since the timing must be corrected for the actual decay point of the tagged B vertex. The beam has $\sigma_z \simeq$ 0.7 cm, which adds (in quadrature) an uncertainty of 25 ps to the time-of-flight resolution.

‡ This is just for the kaon identification efficiency and the geometrical acceptance is not included.

momentum between 0.4 GeV and 4.0 GeV. The aerogel detector can be combined with dE/dx measurement to identify kaons over the entire momentum range.

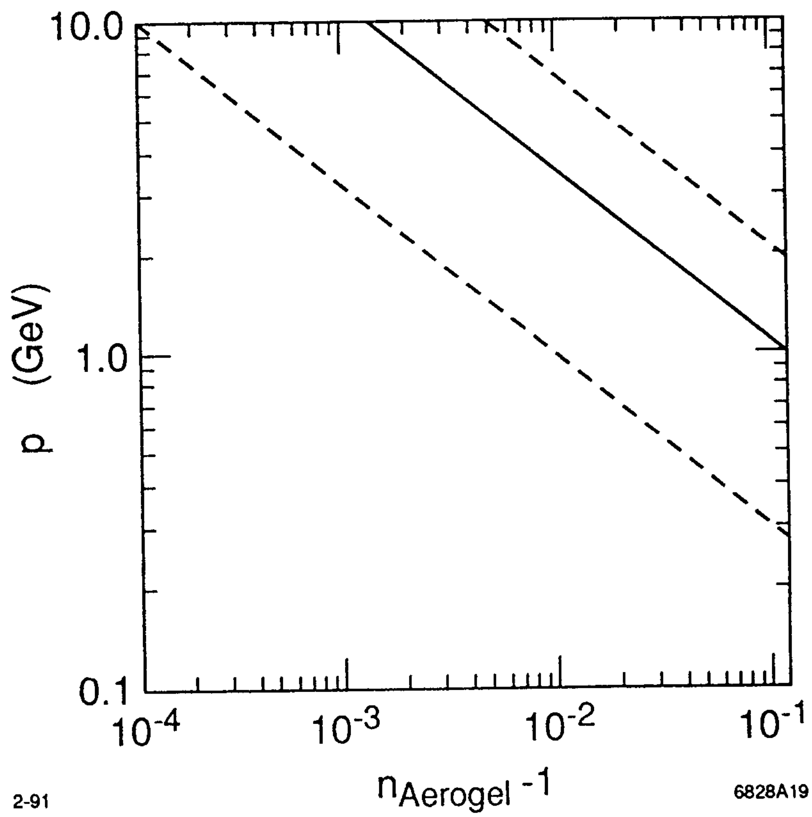


Figure 17. *Threshold momentum for π, K and p as a function of refractive index.*

References

1. *The Physics Program of a High-Luminosity Asymmetric B-Factory at SLAC*, edited by D. Hitlin, SLAC-PUB-353, (1989).
2. R. Aleksan, J. Bartelt, P.R. Burchat and A. Seiden, *Phys. Rev.*, **D39** 1283, (1989).

Reconstruction, Vertexing and Background for $B^0 \to D^+D^-$

A. Snyder and S. Wagner

1. Introduction

We have investigated the detector efficiency, backgrounds and vertex resolution for the *CP* eigenstate $B^0 \to D^+D^-$ in the decay modes

$(K^-\pi^+\pi^+)_{D^+}(K^+\pi^-\pi^-)_{D^-}$ (mode I)

and

$(K^-\pi^+\pi^+)_{D^+}(K^+\pi^-\pi^-\pi^0)_{D^-}$ (mode II).

Vertex resolution was investigated for mode I and for $(K^-\pi^+\pi^+\pi^0)_{D^+}(K^+\pi^-\pi^-\pi^0)_{D^-}$ (mode III). We have also studied how variations in the detector such as those described in $Ba\overline{B}ar$ Note # 40 effect these decay channels.[1] The *CP* asymmetry in $B^0 \to D^+D^-$ decay is proportional to $\sin 2\beta$; this mode can therefore potentially be used either to reduce the amount of integrated luminosity required to establish *CP* violation in the *B* system, or to explore extensions of the Standard Model, should it be found that the asymmetry found here was different from that found in the ψK_S^0 final state.

2. Expected Rates

The branching fraction for the decay $B^0 \to D^+D^-$ can be estimated from the branching fraction for $B^0 \to D^+D_s^-$ measured by CLEO[2] since the two differ only by the Cabibbo angle. We find

$$B(D^+D^-) = B(D^+D_s^-) \times \sin^2\theta_c \approx 6 \times 10^{-4} \times F \qquad (1)$$

where $F = 0.02/B(D_s \to \phi\pi)$ corrects for the assumption of a 2% branching fraction for $D_s \to \phi\pi$ used by CLEO. Using[3] $F \approx 2/3.5$ yields about 10300 $B^0 \to D^+D^-$ decays in a 30 fb^{-1} sample.

$\sim 10K$ $B^0 \to D^+D^-$ decays can be expected in a 30 fb^{-1} sample.

Table 1 gives branching fractions for decay modes of the D^+ with no more than one π^0. The column labeled "visible" gives the branching fraction when only the $\pi^+\pi^-$ decay mode of the K^0_S is used to detect $\overline{K^0}$. The efficiencies are estimated using the formula

Table 1. D^+ *decay modes which are useful in this analysis.*

Decay Mode	Branching Fraction	Visible	× efficiency	× vertex
$K^-\pi^+\pi^+$	7.8%	7.8%	6.0%	6.0%
$K^-K^+\pi^+$	0.96%	0.96%	0.74%	0.74%
$K^-\pi^+\pi^+\pi^0$	3.7%	3.7%	2.3%	2.3%
$\overline{K^0}\pi^+$	2.8%	0.93%	0.70%	0.0%
$\overline{K^0}\pi^+\pi^+\pi^-$	7.0%	2.3%	1.4%	1.4%
$\overline{K^0}\pi^+\pi^0$	8.3%	2.8%	1.6%	0.0%
$\overline{K^0}\pi^+\pi^+\pi^-\pi^0$	4.4%	1.5%	0.77%	0.77%
Totals	34.0%	19.0%	13.5%	11.2%

$$\epsilon_D = (0.93)^{N_{ch}}(0.80)^{N_{\pi^0}}(0.98)^{N_{vert}}(0.98)^{N_{cut}} \qquad (2)$$

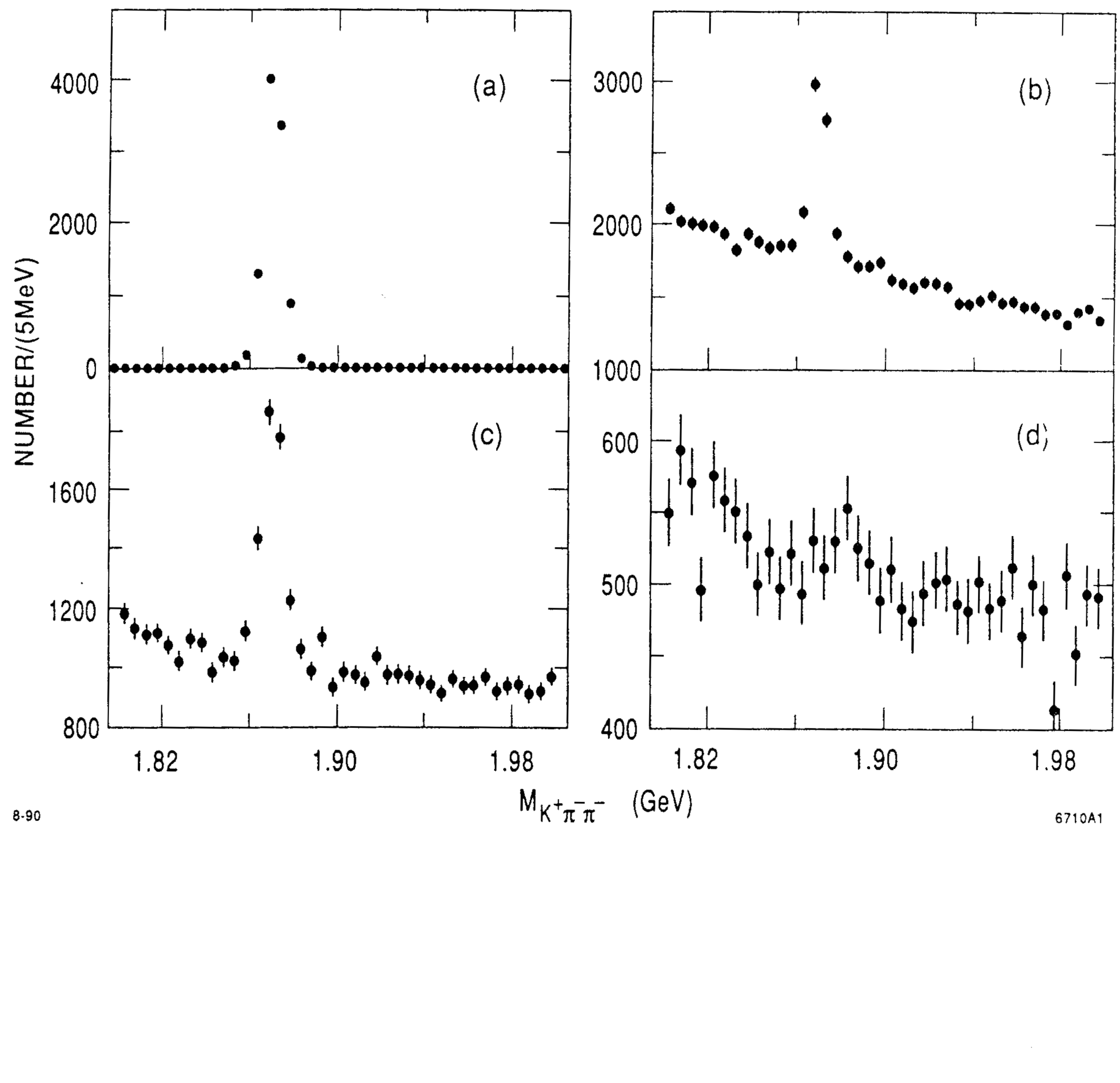
which has been adjusted using $K\pi\pi$ and $K\pi\pi\pi^0$ channels for the cuts described in the next section. The last column (labeled "×vertex") gives the result if we assume that the vertex is not reconstructible for modes that do not have at least two charged tracks coming from the D decay vertex. When the D^+ and D^- are combined to make a B^0 there is additional efficiency factor of

$$\epsilon_B = 0.89 \times 0.98 \times 0.93 \qquad (3)$$

155 B^0s can be reconstructed in modes where at least one of the D vertices can be measured.

where the first factor is due to a cut on the momentum in the $\Upsilon(4S)$ rest frame (see Section 3) and the second and third are due to the vertex cut and the mass cut. Assuming that the decay modes in which at least one of the vertices can be reconstructed are useful leads to an efficiency of about 1.5%. If both D's must be verticized, the efficiency is reduced to 1.2%. This leads to 155 reconstructed $B^0 \to D^+D^-$ decays in the first case and 113 reconstructed decays in the second case for a 30 fb^{-1} sample. We will study in detail reconstruction and background for the B decay channels $(K^-\pi^+\pi^+)_{D^+}(K^+\pi^-\pi^-)_{D^-}$ and $(K^-\pi^+\pi^+)_{D^+}(K^+\pi^-\pi^-\pi^0)_{D^-}$. If only these two channel are used the reconstruction efficiency is 0.6% yielding to 61 reconstructed events in 30 fb^{-1}. Only vertex reconstruction will be studied for mode III.

Figure 1. *$K^+\pi^-\pi^-$ mass distribution for*
a. *signal,*
b. *unbiased $B\overline{B}$ events,*
c. *$c\bar{c}$ events and*
d. *uds events.*



3. Analysis

For efficiency and background studies we focus on the two decays channels:

$$(K^-\pi^+\pi^+)_{D^+}(K^+\pi^-\pi^-)_{D^-} \quad \text{(mode I)},$$

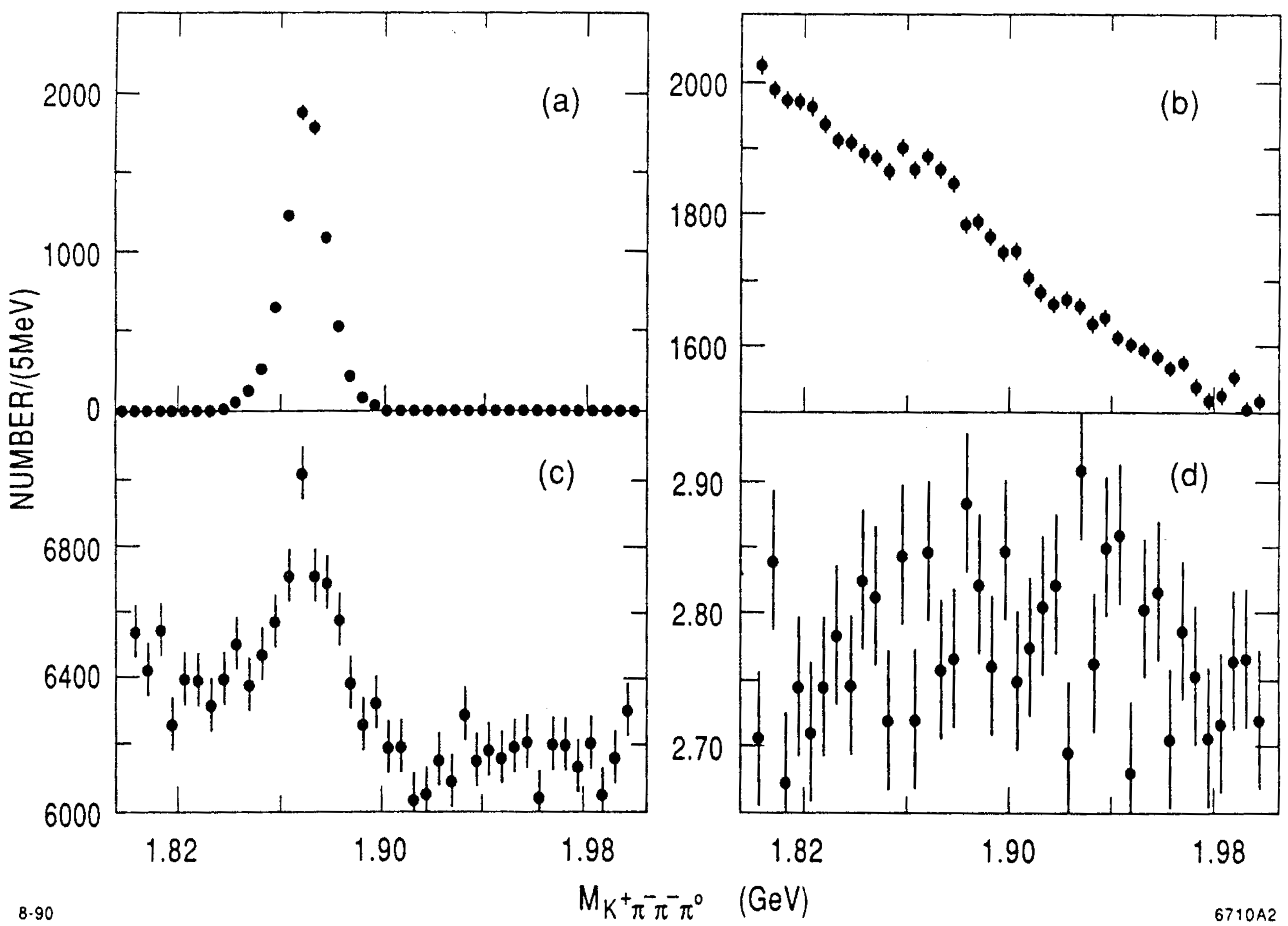
and

$$(K^-\pi^+\pi^+)_{D^+}(K^+\pi^-\pi^-\pi^0)_{D^-} \quad \text{(mode II)}.$$

To search for these decays, we create list of tracks which are consistent with the relevant particle ID hypothesis (namely $K/\pi/\pi^0$). For the charged particles we will compare results obtained using perfect particle ID, CRID + dE/dx, and aerogel + dE/dx. π^0s are obtained by taking all $\gamma\gamma$ pairs with effective mass within 15 MeV of the π^0 mass. We then loop through all possible combinations of three charged tracks or three charged tracks and a π^0. Combinations which fall within 10 MeV (20 MeV when a π^0 is used) of the $D^\pm$ mass are entered in a lists of D^+ and D^- candidates. The $K\pi\pi$ and $K\pi\pi\pi^0$ mass distributions for signal events are given in Figures 1a and 2a. D^+D^- pairs that are within 20 MeV of the B^0 mass are considered B^0 candidates. The D^+D^- mass distribution is shown in Figure 3a for mode I. We assume that by using constraints provide by the π^0 mass in our calculation of the $D^\pm$ mass, we can achieve a resolution on M_B in mode II equal to the resolution for mode I. This assumption has not been verified by using a kinematic fitting package, but it is reasonable given three constraints and only one 'badly' measured track - the π^0. Finally the momentum of the B^0 candidate in the $\Upsilon(4S)$ rest frame is required to be between 0.22 and 0.40 GeV. This distribution is shown in Figure 4a.

The first step in the vertexing process is to fit the three charged tracks from a D decay candidate to a common vertex using SECVTX. The χ^2-probability of this fit is required to be greater than 1%. Next the two D vertices are used to form a B^0 vertex. This is done with a program especially written for this purpose called DVERT. DVERT uses the approximation that the directions of the D mesons are perfectly known; only errors on the position of the D vertices are used. Figure 5 shows the χ^2-probability plot from DVERT for mode I, which is flat demonstrating that neglecting the errors on the direction of the D's was a good approximation. All D^+D^- vertices with $P(\chi^2) > 2.0\%$ are accepted. For mode III the probability plot is not quite flat (see Figure 5b) which can be attributed to the large directional error when the D reconstruction contains a π^0. We again require $P(\chi^2) > 2\%$. In addition, for mode II we reject events with more than one $B^0 \to D^+D^-$ candidate.

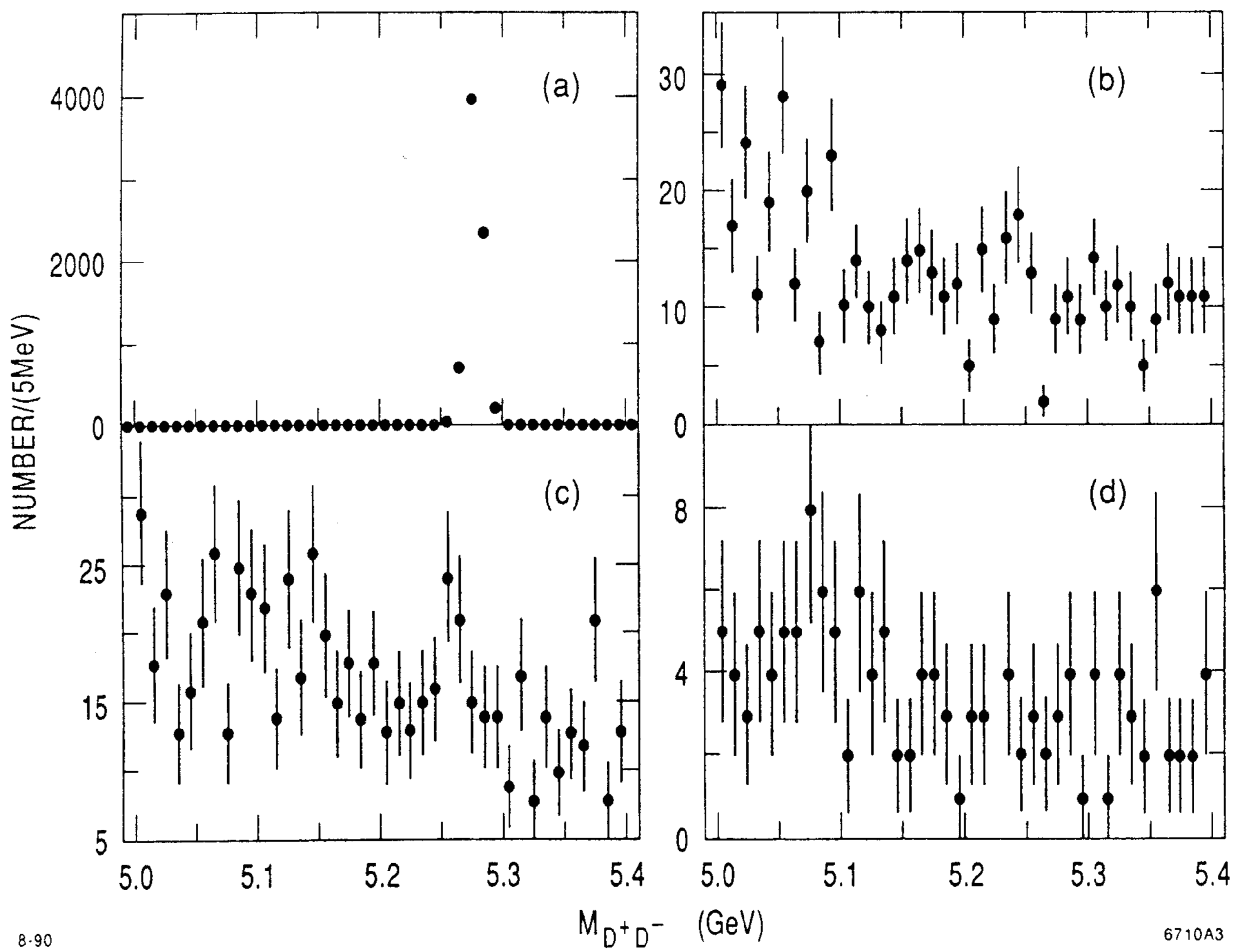
Figure 2. *$K^+\pi^-\pi^-\pi^0$ mass distribution for*
a. *signal,*
b. *unbiased $B\overline{B}$ events,*
c. *$c\bar{c}$ events and*
d. *uds events.*



Before going on to the efficiency obtained we summarize the analysis procedure:

- Select K, π and π^0 tracks using specific particle ID techniques.
- Find all combinations within 10 MeV (20 MeV when a π^0 is involved) of the D^+ mass.
- Find all D^+D^- pairs within 20 MeV of the B^0 mass.
- Require the momentum (p_{cm}) in the Υ(4S) rest frame of the D^+D^- pair to be between 0.22 and 0.4 GeV.
- Require that each D form a vertex with $P(\chi^2) > 2.0\%$.
- Require that the D^+D^- form a vertex with $P(\chi^2) >$ 2.0%.

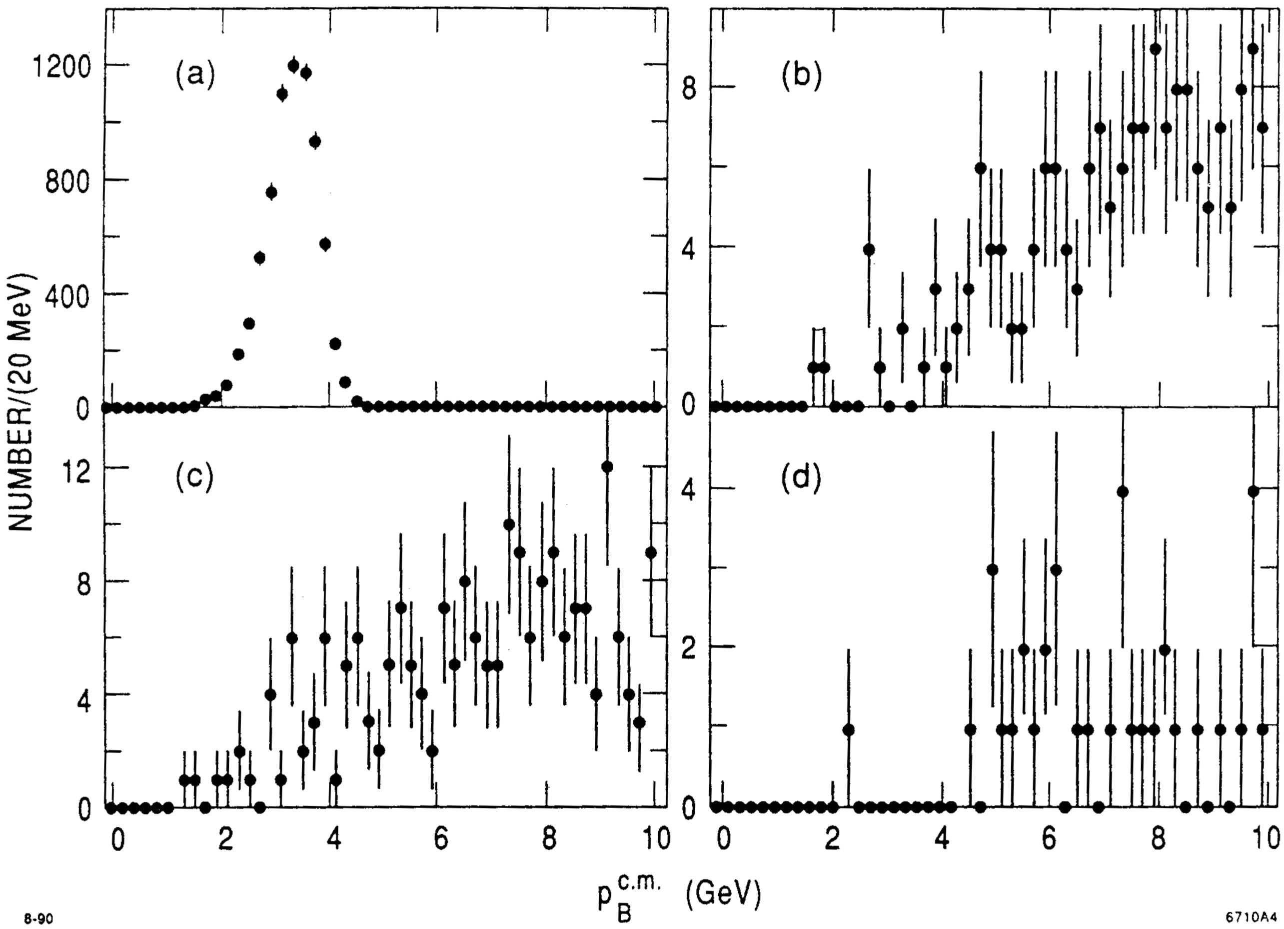
Figure 3. *D^+D^- mass distribution (for mode I) for*
a. *signal,*
b. *unbiased $B\overline{B}$ events,*
c. *$c\bar{c}$ events and*
d. *uds events.*



4. Efficiency

The efficiency of this procedure for selecting D^+D^- decays is 51% for the all charged $K\pi\pi$ mode (mode I) and 41% for the mode II using the default ASLUND detector.[4] For the all charged modes, the kinematic cuts alone yield an efficiency of 54% and the vertexing accepts 94% of the decays kinematically reconstructed. For mode II, the comparable numbers are 44% and 94%.

Figure 4. *Momentum of the D^+D^- (for mode I) in the $\Upsilon(4S)$ rest frame for*
a. *signal,*
b. *unbiased $B\overline{B}$ events,*
c. *$c\bar{c}$ events and*
d. *uds events.*



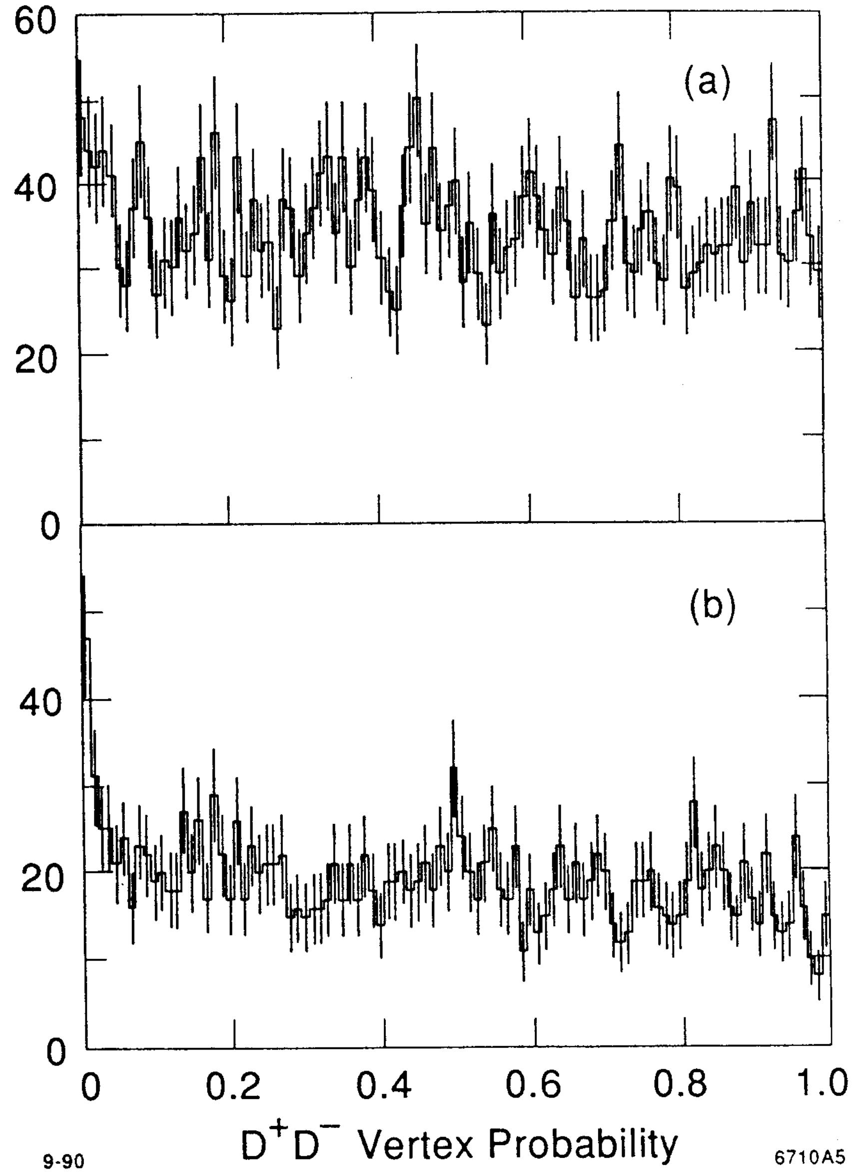


Figure 5. *A χ^2-probability distributions for*
a. *mode I and*
b. *mode III.*

Table 2a. shows the effect of different detector configurations on the efficiency for mode I. We varied the beam pipe radius, beam pipe and silicon thickness, particle ID technique

(3σ cuts are used[5]), drift chamber resolution and calorimeter resolution; the variations (up to three) are described in Table 2c. The second variation for the calorimeter is to put 0.3 radiation lengths of material in front of it, without changing intrinsic resolution. The numbers shown are relative the value of 51% obtained with the default ASLUND detector configuration. Table 2b gives the same information for mode II.

Table 2a. *Effect of detector variations on efficiency for decay mode I*
Table 2b. *Effect of detector variations on efficiency for decay mode II*
Table 2c. *Specification of detector variations*

Property varied	Variation I	Variation II	Variation III
Beam pipe radius	1.0	1.0	–
Beam pipe and Si thickness	1.0	1.0	–
Particle ID	1.0	0.4	1.0
Drift chamber resolution	1.0	1.0	1.0
Calorimeter resolution	1.0	–	–

Property varied	Variation I	Variation II	Variation III
Beam pipe radius	1.0	1.0	–
Beam pipe and Si thickness	1.0	1.0	–
Particle ID	1.0	0.4	1.0
Drift Chamber resolution	1.0	1.0	1.0
Calorimeter resolution	1.0	–	–

Property varied	Variation I	Variation II	Variation III
Beam pipe radius	1.0 cm	3.0 cm	4.0 cm
Beam pipe and Si thickness	2mm/600μm	–	–
Particle ID	CRID+dE/dx	Aerogel+dE/dx	TOF+dE/dx
Drift chamber resolution	300μm	$L_{RAD} = 100M$	–
Calorimeter resolution	$4.0/E^{1/4}$	0.3 r.l.	–

Efficiencies are nearly independent of detector details.

The overall conclusion to be drawn from this study is that the reconstruction efficiencies are nearly independent of the detector details as long as the basic geometry is kept fixed. Note that mass cuts have been widened to keep the efficiency constant when the resolution is degraded; this will result in larger backgrounds.

The one exception to the detector independence of the efficiency concerns the use of aerogel for particle ID. The aerogel detectors suffer from a high inefficiency due to the small number of photoelectrons detected. The standard ASLUND detector assumes that a $\beta = 1$ particle produces 10 detected photoelectrons in the aerogel with index of refraction of 1.0267 and 5 detected photoelectrons in the aerogel with index of refraction of 1.008. This corresponds to a detection efficiency of about 12.5% for electrons. A detection system substantially more efficient than this will be required if aerogel is to be a viable particle ID alternative. In Figure 6 the D^+D^- detection efficiency obtained using aerogel relative to the efficiency with perfect particle ID is plotted versus the photo-electron efficiency of the aerogel counters. A photo-electron efficiency of 35% will increase D^+D^- efficiency by about a factor of 2.

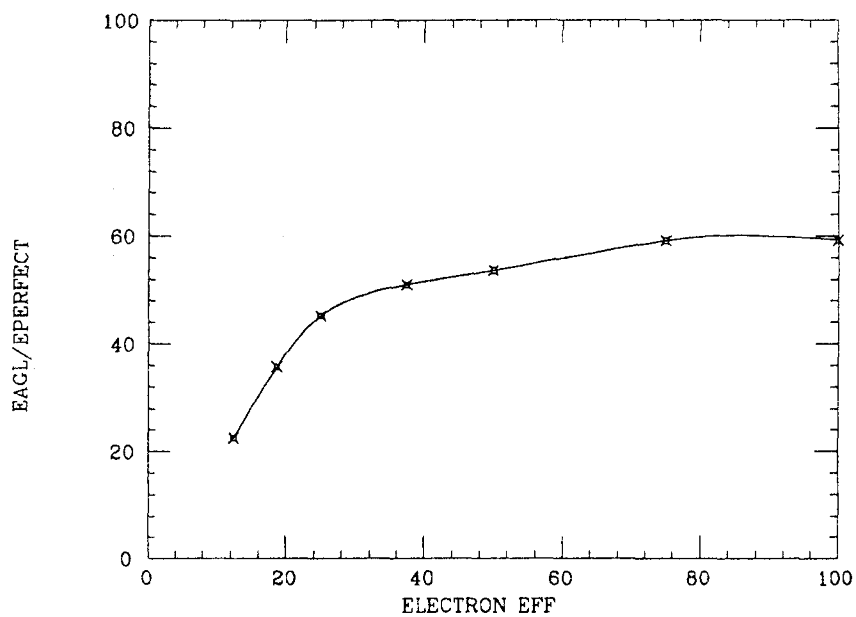


Figure 6. *Efficiency obtained using aerogel for particle ID divided by efficiency obtained with perfect particle ID as a function of assumed photo-electron efficiency of aerogel counters.*

5. Vertex Resolution

Figure 7 shows a plot of the difference ($z_{found} - z_{thrown}$) between the reconstructed z for $B^0 \to D^+D^-$ decays (in mode I) and the original z of the B^0 decay generated by the Monte Carlo.[6] The width is $\approx 40\mu$m. For comparison the same distribution for $B^0 \to \pi^+\pi^-$ is 27μm wide. The difference is not due to multiple scattering or due to the extrapolation along the flight paths of the D mesons, but due the loss of information on the position of the B^0 decay along the flight paths of the D mesons caused by the non-zero $D^\pm$ decay length. This can be demonstrated by examining the vertex resolution for a sample of $B^0 \to D^+D^-$ decays generated with $D^\pm$ lifetime set to zero. If SECVTX is used to reconstruct the vertex (using the information that all six tracks come from a common vertex) a resolution of about 25 μm is obtained from the D^0 lifetime sample. If DVERT is used (explicitly ignoring the fact that all six tracks come from the same vertex) then the same resolution is obtained with $\tau_D = 0$ (40 μm) as was obtained with realistic D lifetimes. Thus the loss of resolution is not cause by multiple scattering (that loss is made up by having more tracks) or by the extrapolation along the D mesons flight paths, but by the loss of the constraint that all six charged tracks come from a common vertex. The lost information comes out in the fit as two additional measured quantities, the decay lengths of the two D mesons. It is conceivable that these two decay lengths could be used to the reject certain types of background (with some accompanying loss of efficiency, of course), but this will not be investigated here.

The $z_{found} - z_{thrown}$ distribution for decay mode III is 55μm wide - only marginally worse than mode I. Mode II (with only one π^0) will be narrower than mode III.

The resolution on δz between the tag and a CP eigenstate such as $\pi^+\pi^-$ is about 80μm; it is dominated by the resolution on the tag.[7] The marginally worse resolution for D^+D^- decays should have almost no effect on our ability to extract CP violations from this mode.

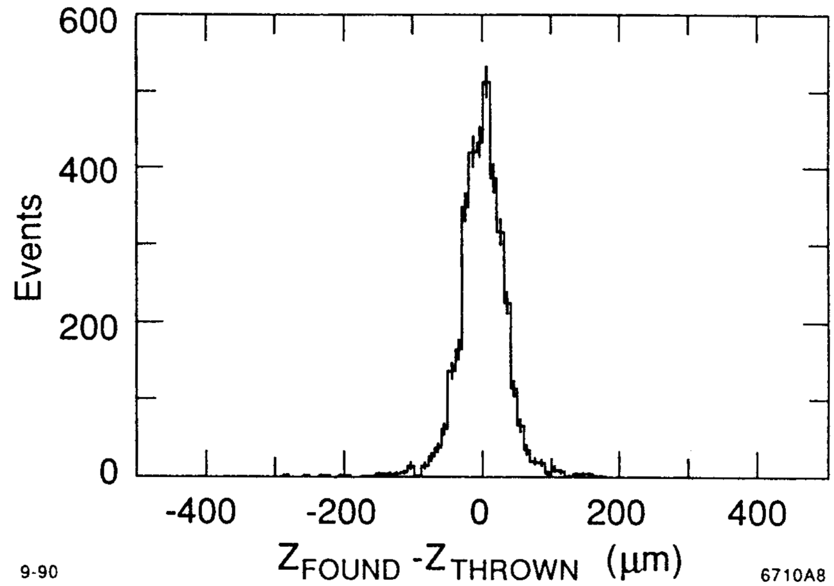


Figure 7. *Distribution of $\delta z = z_{found} - z_{thrown}$ for mode I.*

The degradation in vertex resolution with the beam pipe radius increased to 4 cm does not destroy our sensitivity to CP violation in $B^0 \to D^+D^-$ decays.

An important issue is the sensitivity of the vertex resolution to changes in the beam pipe radius and to changes in the amount of material in the beam pipe and vertex detector. The results of these studies are summarized in Figure 8 for variations described in Table 2c. In no case is the degradation serious enough to seriously threaten our ability to measure the CP-violating asymmetry in these modes.

6. Backgrounds

To estimate backgrounds we widen the mass cuts so as to reduce the statistics in unbiased Monte Carlo events required to see the effect. We increase the $D^{\pm}$ mass cuts to 100 MeV and the B^0 mass cut to 200 MeV. We then scale down by the ratio of the true cut (e.g. 10 MeV for D mesons) to these cuts to estimate the background. This increases the statistical power of a given sample by a factor of 1000 for purely combinatorial background and by factor of 100 for background in which either the D^+ or the D^- is real. The backgrounds studied here are purely combinatorial. We did not have enough statistics to study the background due to one real D; preliminary indications are that it is comparable to the combinatorial background. We generated $\approx 200K$ events each for uds, $c\bar{c}$ and $B\overline{B}$ production. Figures 1b, 2b, 3b and 4b show the mass and P_{BCM} distribution for the unbiased $B\overline{B}$ events. The same distributions for $c\bar{c}$ and uds are shown in portions c and d of Figures 1-4. The results are summarized in Table 3.

Mode	Signal	uds	$c\bar{c}$	$B\overline{B}$
I	36	0.13	2.2	0.6
II	23	0.7	4.4	3.3

Table 3. *Signal and combinatorial background expected in 30 fb^{-1} for $B^0 \to D^+D^-$ decays.*

Mode I has little background. Mode II has substantial background but not so much as to make this channel unusable. For mode II we have rejected events which contain more than one $B^0 \to D^+D^-$ candidate; this selection retains 80% of the signal while reducing the background by a factor of two. It is worth noting that employing the π^0 mass constraint to reduce the mass resolution for $(K^-\pi^+\pi^+\pi^0)_D$ to the same level that for $(K^-\pi^+\pi^+)_D$ would reduce these backgrounds by a factor of about four. Note also that K-tagging favors signal over background by about a factor of two.

Backgrounds are low for mode I. For mode II background are substantial, but not overwhelming.

Tables 4a and 4b show the effect of detector variations on backgrounds for mode I and mode II respectively. The detector variations are the same ones specified in Table 2c. The numbers given in Tables 4b and 4c are relative to the background estimates in Table 3 for the standard ASLUND detector with perfect particle ID. The backgrounds are very sensitive to detector variations. It is particularly important from the background point of view to keep the radiation length of the drift chamber long and to have high quality particle ID such as provided by a CRID.

Property varied	Variation I	Variation II	Variation III
Beam pipe radius	1.0	1.0	1.0
Beam pipe and Si thickness	1.0	1.0	–
Particle ID	1.4	0.3	2.6
Drift chamber resolution	1.2	5.4	–
Calorimeter resolution	1.0	1.0	–

Property varied	Variation I	Variation II	Variation III
Beam pipe Radius	1.0	1.0	1.0
Beam pipe and Si thickness	1.0	1.0	–
Particle ID	1.6	0.9	3.0
Drift Chamber resolution	1.1	2.7	–
Calorimeter resolution	2.3	1.0	–

Table 4a. *Effect of detector variations on background for decay mode I*

Table 4b. *Effect of detector variations on background for decay mode II*

7. Conclusions

It should be possible to extract *CP* violating asymmetries from $B^0 \to D^+D^-$ decays. The vertex resolution is sufficient and the background is acceptable, at least in the all charged track channel $(K^-\pi^+\pi^+)_{D^+}(K^+\pi^-\pi^-)_{D^-}$ (mode I). Our experience with $(K^-\pi^+\pi^+)_{D^-}(K^+\pi^-\pi^-\pi^0)_{D^-}$ (mode II) leads us to suspect that backgrounds may be unacceptably large for channels with more than three particles coming from both of the D decays (e.g. mode III). Kinematic fits constraining to the D masses as well as the π^0 mass may help reduce these backgrounds to a workable level. Because of the low branching fraction and the small number of decay modes that can realistically be reconstructed, the $B^0 \to D^+D^-$ channel is not as sensitive as the $B^0 \to \psi K^0_S$ channel. It is worth pursuing, however, since in extensions of the Standard Model, the asymmetry in $B^0 \to D^+D^-$ need not be the same as that for $B^0 \to \psi K^0_S$.[8]

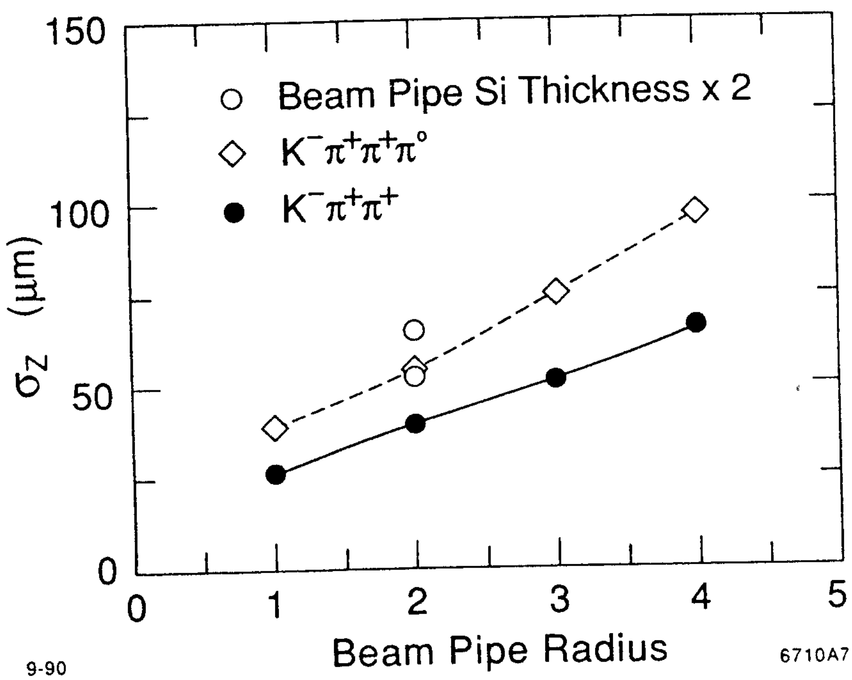


Figure 8. *Dependence of δz resolution on beam pipe radius. Dashed curve is for Mode III and solid curve is for Mode I. The two circles at 2 cm show the effect of doubling the thickness of the beam pipe and silicon detector (1 mm Be $\rightarrow$ 2 mm Be and 300μm Si $\rightarrow$ 600μm Si) at fixed radius.*

References

1. J. Dorfan, S. Komamiya and A. Snyder, *Filling in the Detector Variation Matrix*, *Ba$\overline{B}$ar* Note # 40.

2. D. Bortoletto, *et al.*, *Phys. Rev. Lett.* **64**, 2117 (1990).

3. J. Anjos *et al.*, *Phys. Rev. Lett.* **64**, 2885 (1990).

4. *The Physics Program of a High-Luminosity Asymmetric B Factory at SLAC*, Edited by D. Hitlin. SLAC-353, LBL-27856, CALT-68-1558, October 1989.

5. We work from list of tracks that give signal consistent within 3σ with the hypothesis of interest in all the particle ID devices being used. In the case of aerogel the counters that fire or fail to fire must be consistent with the mass hypotheses. No attempt is made to resolve ambiguous identifications.

6. In the studies of vertex resolution, we have removed all mass and vertex probability cuts, and have only reconstructed the true signal by using Monte Carlo traceback (the tracks used for these studies have been put though the full detector simulation package, of course). This means that our calculated resolutions will be slightly *worse* than those that would have been measured using mass and vertex probability cuts. This occurs because the combinations in the tails of the mass and vertex probability distributions are highly correlated with those in the tails of the vertex position distributions, and we have included all these poorly measured decays in the vertex resolution studies.

7. See section on $B^0 \to \pi^+\pi^-$ in [3]; Also F. LeDiberder, *Ba$\overline{B}$ar* Notes # 33 and # 34.

8. C. Dib, D. London and Y. Nir, *CP* Asymmetries Beyond the Standard Model, SLAC-PUB-5323 and these *Proceedings*; Y. Nir and H. Quinn, *Phys. Rev.* **D42**, 1473(1990).

TWO-PHOTON PHYSICS AT A B FACTORY

DANIEL A. BAUER

1. INTRODUCTION

In the past decade, the study of two-photon interactions at e^+e^-storage rings has yielded significant new results on the properties of C=+1 mesons, QCD and the nature of virtual photon interactions. However, there is still much that can be learned with larger data sets, including searches for new C=+1 mesons and exotic particles at masses > 2 GeV and the systematic testing of reliable QCD predictions at large two-photon center of mass energies. While several second generation experiments are currently collecting data, the much higher luminosities available at a B Factory should make it an ideal place to pursue two-photon physics, as well as the important B physics objectives. This section demonstrates the two-photon physics which could be done at a B Factory and the resulting possible impacts on the machine and detector design.

A B factory will provide the necessary increase in luminosity required to extend two-photon physics into higher mass regions

2. THE TWO-PHOTON REACTION

The two-photon reaction, shown in Figure 1, involves the emission of virtual photons by both $e^+$and e^-.

These photons have space-like four-momenta, q_i , and the distribution of q_i^2 is strongly peaked near 0. Such nearly-real photons couple to hadrons like vector mesons and vector meson dominance (VDM) adequately explains the kinematics of the outgoing hadrons. However, occasionally one or both of the photons is quite virtual ($q_i^2 << 0$) and such photons couple directly to quark loops in ways calculable by QCD. Detection of one of the scattered e^+ or e^- 'tags' the corresponding virtual photon and determines $q_i^2 = 4E_i\ E_i'\ \sin^2\ \theta_i/2$. Since the flux of virtual photons falls like $1/q_i^2$, such tagging must be done at small angles with respect to the beam direction to preserve any rate. Single-tagging determines $Q^2 = max(-q_1^2, -q_2^2)$ and greatly reduces backgrounds from e^+e^-

Detection of the scattered e^+or e^- 'tags' the corresponding virtual photon, giving kinematic information and suppressing annihilation background annihilation at both trigger and analysis levels. Double-tagging gives an even greater background suppression and allows the direct determination of the final state kinematics without the need to reconstruct any final state hadrons.

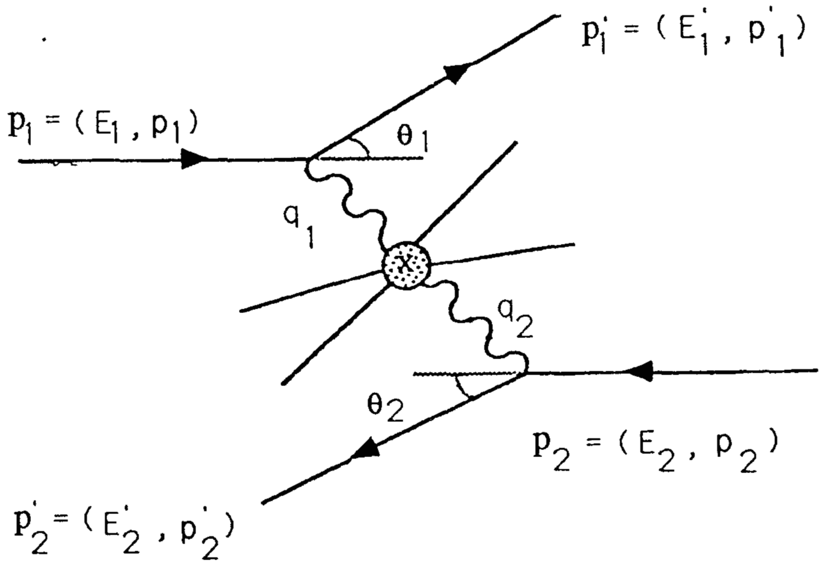


Figure 1. *Feynman diagram of the two-photon reaction*

There are numerous difficulties involved in the experimental study of the two-photon reaction. First and foremost is the rapid fall-off of the virtual photon flux with Q^2 and W, the two-photon center of mass energy. This makes it very difficult to probe even moderate mass regions without achieving enormous integrated luminosities. Compounding the problem of low rates is the fact that the $\gamma\gamma$ system is moving in the lab frame; this tends to force tags and final state particles into the forward parts of the detector where coverage has often been incomplete. Another problem is the generally low multiplicity of two-photon produced final states. Typically there are 2-4 final state particles, and such signatures can be very hard to trigger on in the face of significant backgrounds, particularly those from radiative Bhabha scattering and beam-gas interactions. These effects often make it difficult to reconstruct two-photon induced events with good efficiency.

The two-photon reaction rate is concentrated at small Q^2 and masses, and results in low multiplicity final states

3. Two-Photon Physics

Two-photon reactions provide a clean method for the study of C=+1 mesons and exotic mesons

Two-photon physics [1] is now one of the primary sources of information about meson spectroscopy and QCD, as well as a unique laboratory for identifying exotic mesons. Charge conjugation C=+1 states are formed which are only indirectly accessible in e^+e^- annihilation. Furthermore, the two-photon reaction probes the quark content of hadrons in a different manner than in e^+e^- annihilation; the two-photon reaction rate is proportional to e_q^4, instead of e_q^2. This enhances the production of mesons containing u and c quarks relative to those containing s quarks and suppresses the production of gg. Two photons can also couple to exotic states such as $q\bar{q}q\bar{q}$ and $q\bar{q}g$. The quantum numbers of resonances are usually relatively easy to detect given sufficient statistics in two-photon reactions primarily because continuum backgrounds are small. While hadron colliders and fixed target experiments can produce both C=+1 mesons and exotic states, the experimental detection is difficult and the presence of spectator hadrons complicates the analysis; thus huge sample sizes and partial-wave analyses are required to even verify that the appropriate resonance has been seen. In the special case of spin-1 particles, two-photon reactions offer a unique method for determining spin/parity, as first demonstrated by the TPC/2γ experiment;[2] Yang's theorem implies that a resonance detected in single-tagged reactions and not in untagged reactions must have spin-1. This is especially interesting because, if one can then identify the parity of the resonance, one can distinguish between the $1^{++}q\bar{q}$ nonet and possible 1^{-+} hybrid states.

Single-tagging facilitates the determination of spin/parity of spin-1 resonances at low statistics, allowing sensitive searches for hybrids

In addition to the study of resonances, the measurement of exclusive and inclusive hadron production in two-photon reactions allows access to many aspects of QCD which are difficult to probe in other ways. For example, by measuring the reaction rates and kinematic distributions for the two-photon production of single hadrons and hadron pairs one can determine the distribution of quarks and gluons within hadrons and perhaps begin to understand the non-perturbative realm of QCD. At high Q^2 and high W, where VDM is no longer important, there are reliable calculations from perturbative

The combination of exclusive and inclusive hadron production in two-photon reactions gives a powerful tool for the study of QCD

QCD with which one can compare. The study of two-photon inclusive reactions allows a measurement of the total two-photon cross section, which can shed light on the hadronic nature of photons. One can also determine the photon structure functions and study jet production; at high Q^2, such inclusive measurements should provide a definitive test of perturbative QCD. These examples serve to illustrate the richness of the physics which can be done using the two-photon reaction.

4. Present Status of Two-Photon Physics

First generation two-photon experiments have yielded many results from small data samples

Many excellent reviews of two-photon physics results exist in the literature.[1] In this summary, I will mention only a few measurements to illustrate the present status and to establish the context for what remains to be studied. Most of the experiments with two-photon results to date have been done with detectors optimized for e^+e^- annihilation events and thus have had relatively large systematic errors on their two-photon measurements. In addition, most of the experiments have achieved data samples of only $100 - 300\text{pb}^{-1}$, and thus statistical errors are also large. Finally, very few of the detectors have been equipped with the necessary high-resolution, low-angle calorimetry to allow tagging. Nevertheless, there has been a large physics output from these initial measurements.

Untagged two-photon resonance studies have yielded $\Gamma_{R\gamma\gamma}$ values for many low mass mesons

The emphasis in resonance studies has been on the measurement of the two-photon widths, $\Gamma_{R\gamma\gamma}$, of low-mass, spin-0 and spin-2 mesons since this quantitity is often predictable and gives information about the quark content of the meson. Figure 2 shows the π^0, η, and η' as seen by the Crystal Ball[3] and the observation by CLEO[4] of the $J^{PC} = 0^{-+}$ charmonium resonance η_c in untagged two-photon reactions. Note the very sharp fall-off of statistics with increasing mass; this limits the comparison to theoretical predictions to masses < 1.5 GeV.

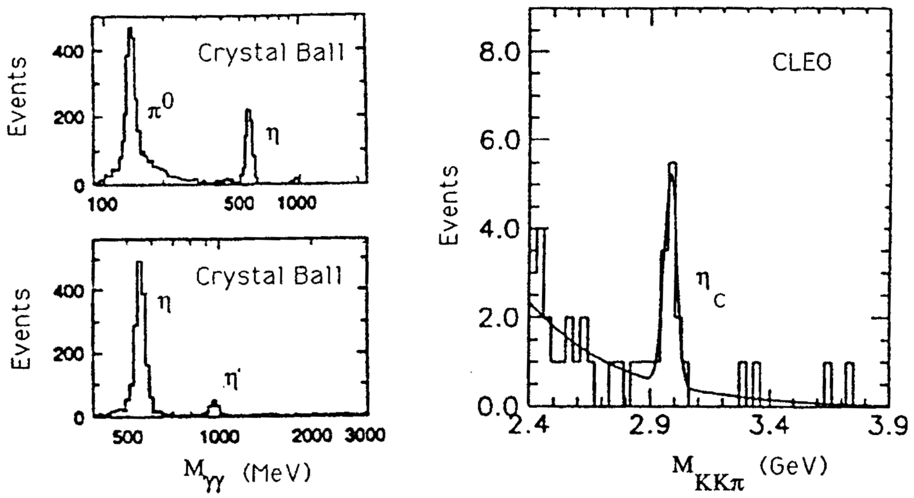


Figure 2. *$\gamma\gamma$ and $K^{\pm}K_S^0\pi^{\mp}$ mass spectra from untagged two-photon reactions as seen by the Crystal Ball and CLEO collaborations, respectively*

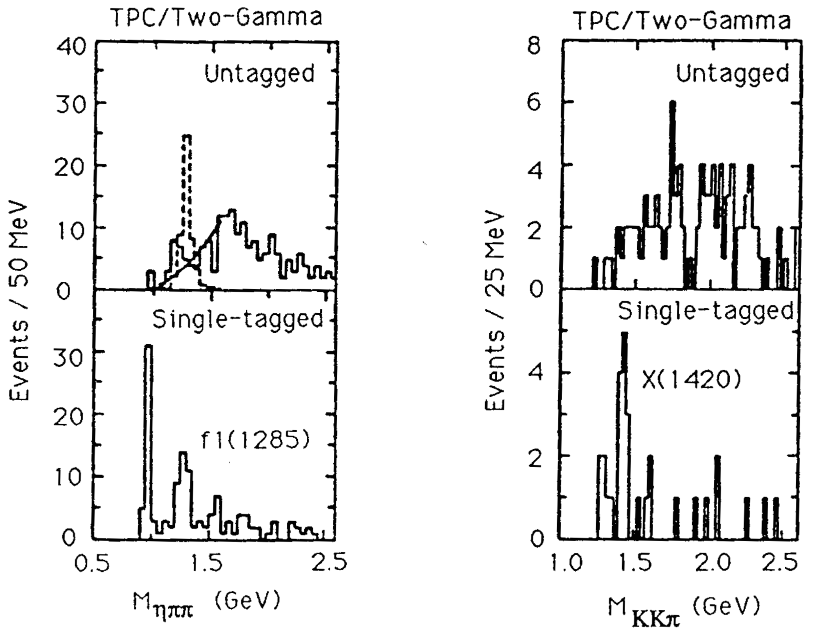


Figure 3. *$K^{\pm}K_S^0\pi^{\mp}$ mass spectra from two-photon reactions as seen by the TPC/2γcollaboration*

Figure 3 shows measurements of the $J^{PC} = 1^{++} f_1(1285)$ meson and the $J^{PC} = 1^{?+}$ X(1420) particle in tagged reactions by TPC/2γ.[5] Despite the paucity of events, the spin 1 identification is solidly based on the Yang's theorem suppression in untagged two-photon reactions. Recent workshops[6] have highlighted the present experimental status of the lowest-lying meson nonets available in the two-photon reaction and illustrate how few mesons have been experimentally studied. Only a few of the lowest-lying mesons have been seen with

Spin-1 resonances are seen in tagged, but not in untagged, two-photon reactions

The goals of two-photon resonance studies are to complete the $q\bar{q}$ nonets and to find exotic particles

sufficient statistics such that the statistical errors are smaller than the systematic uncertainties. Upper limits have been established for the production of a few candidates for exotic states, but they are as yet insufficient to really challenge predictions. Although the focus of resonance studies has recently been concentrated on finding exotic states, it has been pointed out[7] that one cannot expect to determine whether one is seeing such states without a complete understanding of the conventional meson spectrum.

Figure 4 shows recent results from the TPC/2γ collaboration on the exclusive production of $\pi^+\pi^-, K^+, K^-$ and $p\bar{p}$ in two-photon reactions.[8,9]

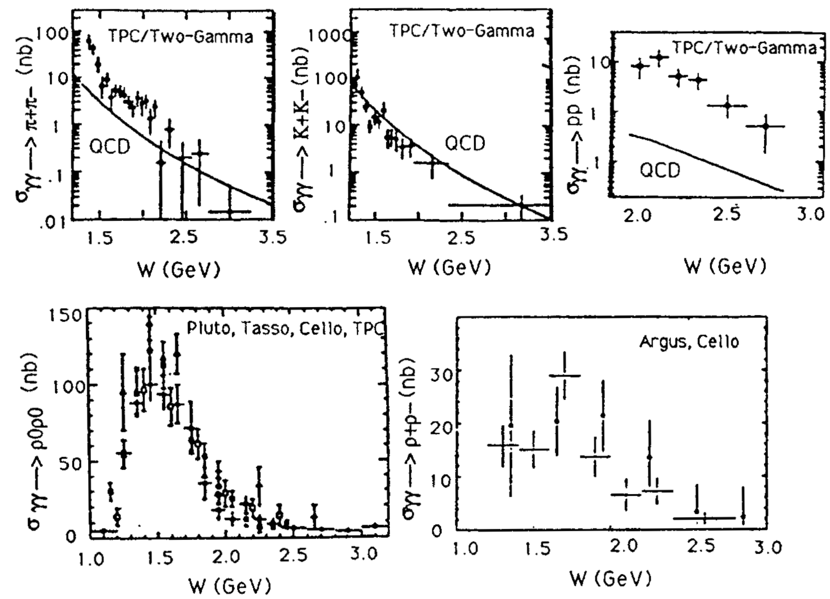


Figure 4. *Hadron pair mass spectra from untagged two-photon reactions*

Exclusive hadron production in two-photon reactions has not yet reached the mass region where it can be reliably compared to perturbative QCD predictions

Note that, although the QCD predictions fall considerably below the data in the $\pi^+\pi^-$ and $p\bar{p}$ channels, they fit much better in K^+K^-. The calculations are not really expected to describe the data at such low values of W, but the agreement should become much better for $W > 5$ GeV if QCD is correct. There are many outstanding puzzles in exclusive hadron pair production such as the enhancement in the cross section for $\gamma\gamma \rightarrow \rho^0\rho^0$ which is not seen in $\gamma\gamma \rightarrow \rho^+\rho^-$, as shown in Figure 4.[10] It is vital to obtain high statistics data samples to understand these discrepancies and to reach the realm of applicability of perturbative QCD. Meanwhile, the present data serves to help refine non-perturbative QCD techniques.

There have been numerous measurements of the photon structure functions and two-photon total cross section. However, much of the data has been in kinematic regions where its interpretation involves the use of vector dominance ideas more than QCD. Another area of considerable interest in inclusive analyses is the study of jet structure in two-photon physics. Several experiments agree that there is an excess of two-photon events containing jets with intermediate values of transverse momentum. The best present hypothesis to explain this involves higher-order QCD diagrams. Thus further data should help to clarify the QCD calculations. Experiments in the near future at PEP and LEP will likely concentrate on the area of inclusive two-photon physics and may answer many of the present questions well before a B Factory could be operational.

Currently operating experiments may suffice to understand two-photon inclusive reactions, which are difficult to study at a B Factory

5. The Future of Two-Photon Physics

Several second-generation experiments capable of doing two-photon physics are now taking data. These include a new version of TPC/2γ at PEP, ARGUS at DESY, CLEO II at CESR and the experiments at LEP and TRISTAN. Most of these experiments expect to achieve much larger data samples within the next 3-5 years than have been obtained in the past 10 years, due partly to improved storage ring performance and partly to improved detector efficiency. Current projections would give the following approximate integrated luminosities before 1995: 1000pb^{-1} for TPC/2γ, 700pb^{-1} for ARGUS, 3000pb^{-1} for CLEO II, 300pb^{-1} for LEP and for TRISTAN. These projections assume only modest improvements in peak luminosities but substantial running time (>6 months/year). It is likely that, for differing reasons, all of these experiments will have concluded their present programs before a B Factory could be built. All of these experiments use detectors which are much better suited to the difficult two-photon kinematics than those of the first generation. ARGUS and CLEO II have particularly good neutral energy detection while TPC/2γ has emphasized particle identification and high

Upcoming two-photon experiments should take about an order of magnitude more data than now exists, with more suitable and complementary detectors resolution, low-angle tagging. Correspondingly, TPC/2γ will likely concentrate on tagged resonance production and tagged inclusive studies, with a particular emphasis on its unique double-tagging capability. ARGUS and CLEO II will explore resonances whose decays involve neutrals and also hadron pair production, primarily in untagged reactions. The higher energy facilities at LEP and TRISTAN may explore inclusive reactions at very high Q^2 and W.

A Monte Carlo generator approach with simple geometric cuts serves to give crude rate estimates; realism is supplied by normalization to measured reaction rates

In order to predict which of the physics questions might be already settled before a B Factory could be online, we have developed some Monte Carlo tools to estimate event rates for some of these second-generation experiments. As examples, we have chosen to simulate CLEO II, TPC/2γ and a generic LEP experiment. The basic tool for predicting the event rates is a Monte Carlo event generator which does the complete calculation of the two-photon fluxes and cross sections, as well as decay kinematics for resonances. Simple cuts are used to simulate the effects of detector acceptance. These cuts are primarily in polar angle for final state particles and in polar angles and energies for tags. Except where noted, the final state particles were required to be at polar angles greater than 0.3 radians, which is an average detection angle for the experiments considered. Tagging angles and energies were chosen as follows: $0.045 < \theta(radians) < 0.2$, $E > 4$ GeV for TPC/2γ; $0.22 < \theta < 0.63$, $E > 0.5$ for CLEO II and $0.03 < \theta < \pi$, $E > 5$ for LEP.* The resulting answers were then normalized by applying a 'detector/analysis' efficiency which attempts to account for remaining detector and analysis effects. These efficiencies were determined from present data of TPC/2γ and CLEO; they vary somewhat from reaction to reaction but are typically around 2% for single-tagged resonance production, 5% for untagged resonance production and $10-20\%$ for inclusive processes and hadron pair production. Note that these estimates are likely to be lower limits for the second-generation detectors since they were obtained using the results from older detectors where the acceptance was limited. We wish to emphasize here that estimates obtained

* It is very unclear whether LEP can do tagging at such low angles in the face of backgrounds.

using Monte Carlo generators such as these are necessarily crude and should be used only as guides to event rates for purposes of developing the accelerator and detectors. True event rate expectations will require much more detailed detector, and particularly trigger, simulations which can only be done once a specific design is available.

We use Monte Carlo tools to study untagged and single-tagged resonance production and exclusive hadron production

Using these Monte Carlo tools, we have studied three representative reactions to determine how well experiments might do. The first of these is untagged $\gamma\gamma \to R(J^{PC} = 0^{-+}) \to \pi^+\pi^-\pi^+\pi^-$ where we have assumed a branching ratio of 1%, $\Gamma_{R\gamma\gamma} = 5$ keV and a total width of 10 MeV. This is a measured decay mode of η_c and a likely decay mode of radially-excited pseudoscalars or pseudoscalar exotic states.** It is also representative of untagged two-photon production of even-spin resonances in general. Other decay modes have been studied in less detail but give rates of the same order. Figure 5 shows the expected event samples. There is a substantial advantage in event rate for a given integrated luminosity at LEP and TPC/2γ as compared to CLEO II, due to the increase of the virtual photon flux with center of mass energy for a given W. However, the expected integrated luminosities turn the picture around, leading to the conclusion that the LEP experiments will not be competitive with either TPC/2γ or CLEO II, except possibly at the very largest resonant masses. Clearly the low-mass region, and particularly the region $1.5 < M_R < 2.0$ GeV which is poorly measured now, will be understood much better in a few years. However, none of the experiments will likely have more than a small handful of events for resonant masses greater than 2 GeV, where many of the glueball and exotic states are predicted to lie.

Second generation experiments will use untagged two-photon reactions to probe for resonances up to at least 2 GeV

** It is likely that radially-excited mesons and most of the exotic states will have smaller branching ratios for such decay modes and also smaller two-photon widths than the standard $q\bar{q}$ states.

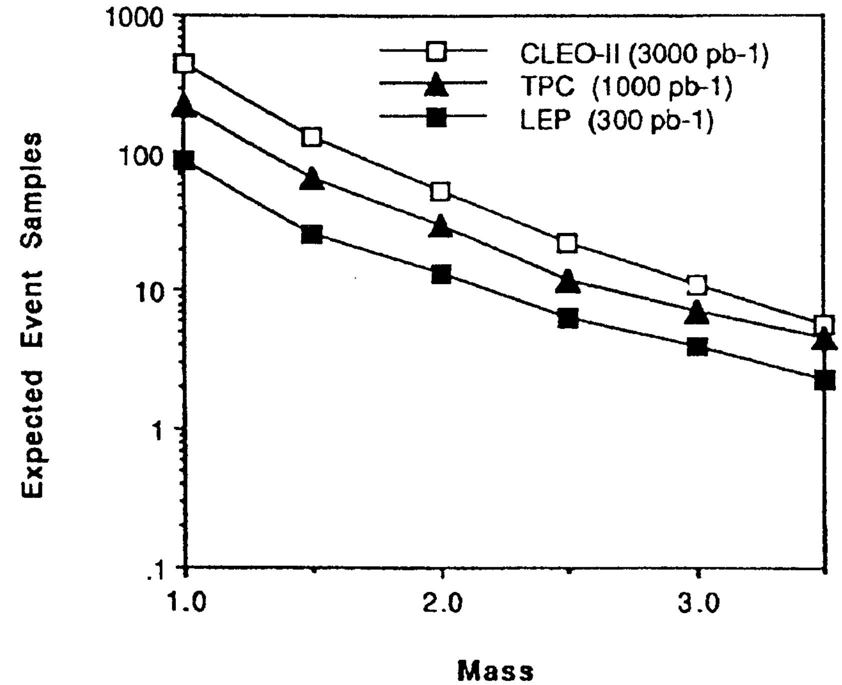


Figure 5. *Expected event samples by 1995 at several different experiments for the untagged process $\gamma\gamma \to R(J^{PC} = 0^{-+}) \to \pi^+\pi^-\pi^+\pi^-$*

Single-tagged two-photon reactions may identify spin 1 mesons up to about 2 GeV in the next generation of experiments

The second reaction studied is the tagged reaction $\gamma\gamma^* \to R(J^{PC} = 1^{++}) \to K^\pm K^0_S \pi^\mp$ where we have assumed a branching ratio of 1.0,*** a coupling parameter $\tilde{\Gamma}_{R\gamma\gamma} = 1.3$ and a total width of 50 MeV, which are the measured values for the X(1420) as seen by TPC/2γ.[5] Since little is known about other spin 1 resonances, it is not clear how representative these parameters might be. A ρ form factor has been assumed for these studies, although not established by the experiment. In contrast to the untagged case, single-tagged event rates are quite sensitive to the form factor chosen. The use of a ρ form factor will give 'worst case' rate estimates, since other choices would favor higher Q^2 which has higher detection efficiency. Again, the use of other decay modes does not substantially alter the rates. Figure 6 shows the expected data samples. The tagged rates are of course considerably smaller than those of untagged production; however, it seems likely that spin 1 states up to 2 GeV might be detected, and the lower mass states such as the X(1420) should be produced in sufficient numbers to determine the parity. Clearly the low-angle tagging of TPC/2γ gives an advantage over the other experiments, although the expected larger integrated luminosity of CLEO II keeps it competitive. LEP rates are suppressed because the large center of mass energy causes the tags to lie

*** Higher mass states and exotics may have smaller individual channel branching ratios.

at very small angles with respect to the beam, which makes their detection very difficult.

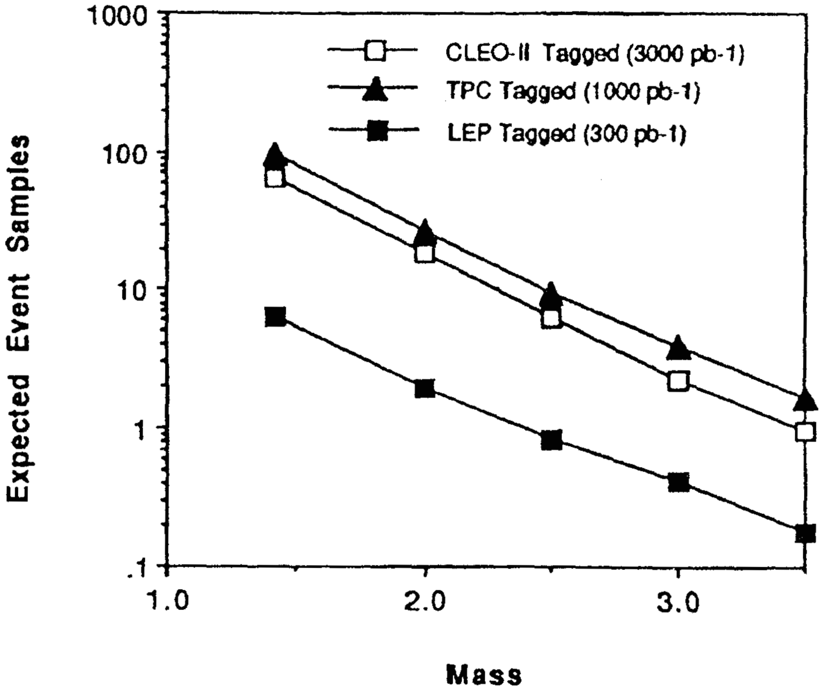


Figure 6. *Expected event samples by 1995 at several different experiments for the tagged reaction* $\gamma\gamma^* \to R(J^{PC} = 1^{++}) \to K^{\pm}K^0_S\pi^{\mp}$

Finally, we have studied the exclusive reaction $\gamma\gamma \to \pi^+\pi^-$ as a typical example of an interesting hadron pair production process. The W dependence of the cross section was taken to match present data[8] in the limited range $1.75 < W < 2.5$ GeV and the Monte Carlo generator was used to extrapolate this to larger values of W. The Q^2 dependence was taken to be that of a ρ pole form factor. Note that, while the extrapolation and form factor assumptions are rather uncertain given present data, they lead to conservatively low event rate estimates at higher W and Q^2. It is expected that higher masses will be associated with a 'flatter' form factor, which puts more events at higher Q^2 where they are easier to detect. Given the luminosity estimates above, Figure 7 details the expected event samples. All of the present experiments will give roughly comparable statistics in the interesting high W region but none of them will extend the data out to the $W > 5$ region where perturbative QCD theorists are completely comfortable.

The next generation of experiments should begin to study perturbative QCD with exclusive two-photon hadron production

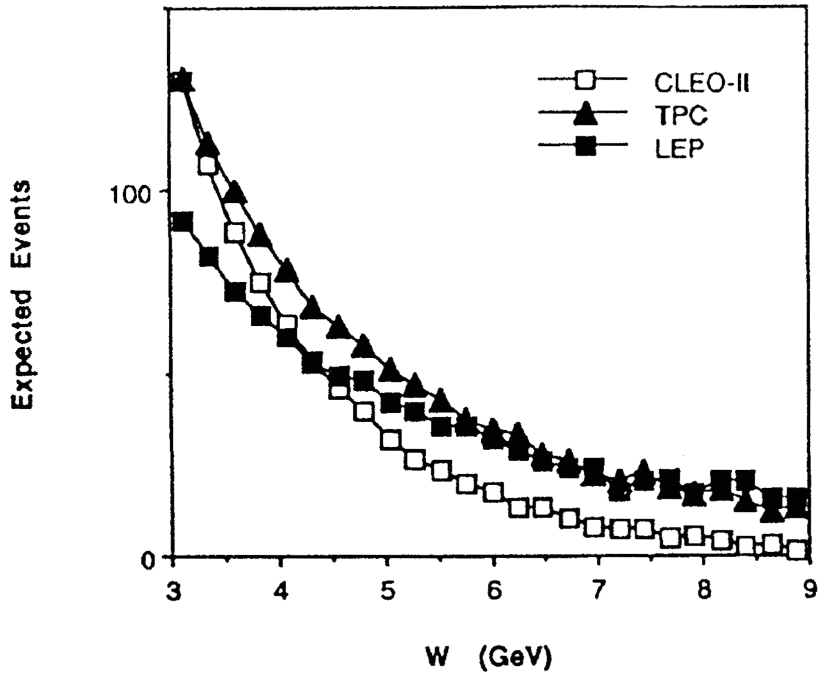


Figure 7. *Expected event samples by 1995 at several different experiments for the untagged process $\gamma\gamma \to X \to \pi^+\pi^-$*

Thus, when the second generation of two-photon experiments has ended and the time for a B Factory draws nigh, we will be left with much to be learned. It is likely that the C=+1 meson resonances will have been adequately explored only up to masses of about 2 GeV, short of the prime hunting ground for glueballs and exotics. While the charmonium region may have been examined somewhat better than now, many interesting questions will remain which will require much higher statistics, such as spin/parity analyses and measurements of the two-photon widths to the level that can confront potential models. There may be new candidates for exotics, especially in the spin-1 sector, but statistics will likely be too low to definitively separate these from the quark model nonets. Exclusive hadron pair production will have barely reached the kinematic region where QCD predictions should apply, and there will undoubtedly still be puzzles such as the current $\rho^0\rho^0$ and $p\bar{p}$ excesses. Higher statistics would allow study of interesting channels such as $\psi\rho$ and $D\bar{D}$. There is also the very interesting process $\gamma^*\gamma^* \to$ pseudoscalar mesons for which there are definite QCD predictions[11] at high Q^2, requiring good statistics in single-tagged and/or double-tagged two-photon reactions. Inclusive studies will probably have probed the QCD region reasonably well by then, with good

Much will be left to do in two-photon experiments by the time a B Factory is ready, especially in resonance production and exclusive hadron production measurements of the photon structure functions and perhaps a better determination of $\Lambda_{\overline{MS}}$. In particular, experiments at TPC/2γ, LEP, TRISTAN and even HERA should have extensively studied the high Q^2 regions, which are not easily accessible at low center of mass energy machines.

There are two possibilities for extending the statistics in two-photon physics to higher W, and thus larger resonance masses. The first is simply (!) to obtain much larger e^+e^- integrated luminosities, which will require machines with larger peak luminosities if we wish to do the physics in a finite time. The logical next step beyond current machines is the proposed *B* Factory. The alternative is a true two-photon factory, where real photon beams are produced from the backscattering of laser light from linear collider electron beams; these intense photon beams are then collided, and the resulting two-photon luminosities are nearly equal to e^+e^- luminosities up to $\sqrt{s}$.[12] This latter scheme will be the only practical method to study the two-photon coupling to heavy quark (b,t) states and to search for heavy exotic resonances which couple to two photons. However, present estimates for the completion of such a machine put it at least a decade away. Thus the *B* Factory is the best hope for two-photon physics in the latter half of the 1990's.

Although backscattered laser beams at a linear collider may be the ultimate two-photon facility, there is a clear need for a high-luminosity e^+e^- facility like the B Factory in the 1990's

6. Two-Photon Physics at a *B* Factory

The obvious question about doing two-photon physics at a *B* Factory is "Can anything worthwhile be done?" After all, the facility will run at low center of mass energy and will clearly be designed primarily to study CP violation in *B* decays. However, the primary need for two-photon studies is a high-quality detector at a very high luminosity machine. The proposed detectors needed to do *B* physics all have good polar angle coverage at least down to 300 mrad with high resolution electromagnetic calorimetry, tracking and particle identification. While this is a much larger angle for detecting scattered e^+ or e^- tags than is usual in two-photon experiments (more typical minimum angles are 20-30 mrad), the low center of mass energy serves to keep the single-tagged rates reasonable. The energy/momentum range covered by the proposed

A B factory will be a good place to study two-photon resonance and exclusive hadron production detectors extends from 50 MeV up to several GeV. Triggers are being considered which will efficiently keep events with 2 or more particles, some of which may be neutral, and simple extensions of logic to recognize Bhabha scattering events will allow tagged two-photon triggering. These requirements are just what one desires for the study of resonances and exclusive hadron production in at least the untagged and single-tagged modes from the two-photon reaction.

Event rate predictions are given using simple geometric cuts but with normalization determined from previous experiments to add realism. An integrated luminosity of 100 fb^{-1} is assumed

To determine whether the event rates will be adequate, we have used the same Monte Carlo tools described in the previous section and assumed a B Factory operating at the Υ(4S), although there is little sensitivity to the actual energy used within the Υ system. The geometric acceptance of the detector is described by the simple cuts $\theta_{min} = 0.3$ radians and $E_{min} = 0.5$ GeV. Despite the simplicity of this 'detector simulation', the normalization of the Monte Carlo predictions to results from first-generation experiments, which were not optimized for two-photon events, makes it likely that these are conservative estimates of the event rates. Still, since two-photon rates are especially sensitive to details of the low-angle coverage, precise expected rates cannot be determined until the detector (and trigger) are better determined. These Monte Carlo rates serve as a guide to detector requirements and physics possibilities. We have assumed that the B factory luminosity will be 100 times present peak luminosities (i.e. 1×10^{34}) and our event rate estimates are based on an integrated luminosity of 100 fb^{-1}. Obviously peak luminosities will start lower and it may take several years to accumulate such integrated luminosities. Most of the two-photon physics topics we have outlined above will benefit already from luminosities at the 10 fb^{-1} level.

Untagged two-photon resonance production of even spin C=+1 resonances will extend up to 5-6 GeV at B factories

Figs. 8-10 give expected event rates at 2 different B factories [symmetric 5.3 x 5.3 GeV and asymmetric 3.1 GeV (e^+) x 9 GeV (e^-)] for the three representative processes discussed in the previous section. Clearly the quantum jump in integrated luminosity provided by the B Factory significantly extends the mass range one can explore, even though the event rates per unit of integrated luminosity tend to be lower than at larger $\sqrt{s}$. As seen in Figure 8, the charmonium region will now be easily accessible in the untagged mode and searches for even-spin, C=+1 resonances will be possible up to masses of 5-6 GeV. The increased statistics at lower masses will allow the detailed kinematic analyses needed to determine spin/parities and production helicity fractions. This information is vital in separating standard $q\bar{q}$ mesons from exotic states, which will obviously be a major focus of the program. Note that the symmetric B Factory gives higher rates at large mass than the asymmetric design for untagged resonance production.

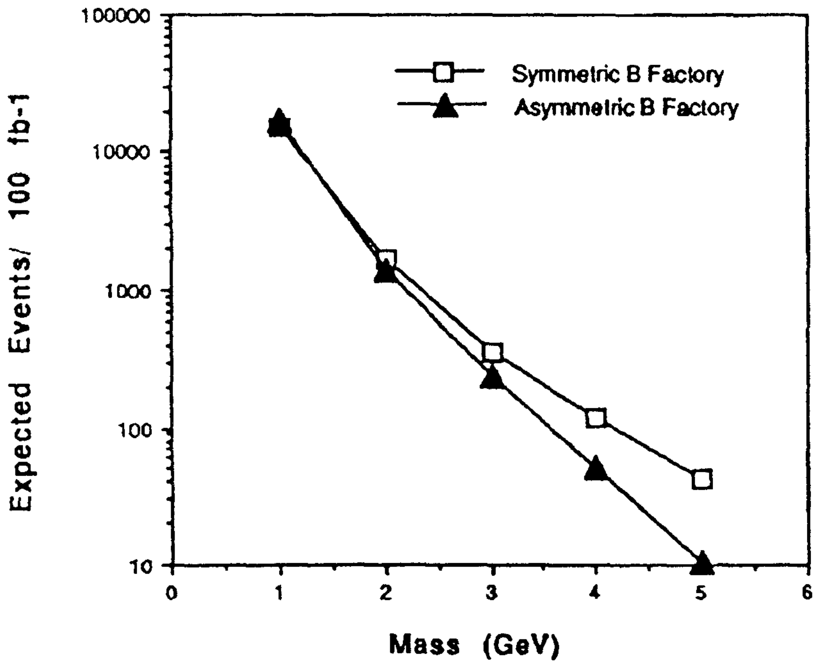


Figure 8. *Expected event samples at symmetric and asymmetric B factories for the untagged process $\gamma\gamma \rightarrow R(J^{PC} = 0^{-+}) \rightarrow \pi^+\pi^-\pi^+\pi^-$*

This is because the higher mass states tend to be produced nearly at rest in the center of mass frame of the colliding $e^+$and e^-(symmetric case) and the decay products are then boosted at the asymmetric machine so that some of them are lost in the hole around the beampipe of the forward end. However, the difference in rates between the two designs in the charmonium region is less than a factor of 2.

Tagged two-photon production of spin-1 resonances will be explored up through the charmonium region at B factories

Figure 9 shows that the production of spin 1 mesons using single-tagging will be extended to masses of around 4 GeV, sufficient to measure the $\chi_1(3510)$ and perhaps $c\bar{c}$ radial excitations. The spin 1 sector may well be the best place to find hybrid $q\bar{q}g$ states at lower masses.

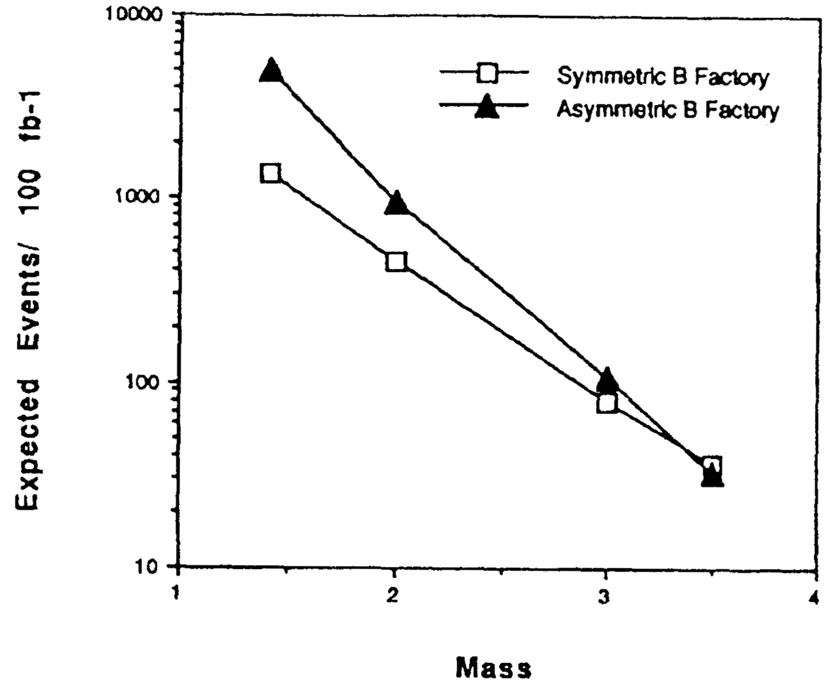


Figure 9. *Expected events at symmetric and asymmetric B factories for the tagged reaction* $\gamma\gamma^* \to R(J^{PC} = 1^{++}) \to K^{\pm}K^0_S\pi^{\mp}$

Predicted spin-1 rates are probably low since form factors for higher mass objects will likely be flatter than the assumed ρ *-pole*

In the single-tagged case, the asymmetric B Factory has an advantage at low masses (about a factor of 5 at 1.42 GeV) over the symmetric B Factory because the boost tends to 'pull' the scattered e^+ tag out into the detector on the backward (low-energy) end, where the single-tagging would occur. Another way of saying this is that a lab angle of 300 mrad at the asymmetric B Factory corresponds to a center of mass frame angle of about 170 mrad. Of course, the forward (e^-or high-energy) tags will not easily be detected at an asymmetric B Factory since they would be boosted into the detector hole, but this straight factor of 2 loss is more than made up for by the rapid rise in the virtual photon flux accessible on the low-energy end. Again it should be emphasized that these predictions may be conservatively low for higher-mass particles because of the use of a ρ form factor. The opposite extreme of a flat form factor will boost the predicted rates by a factor of 20, and greatly lessen the difference between symmetric and asymmetric B factories. Although no detailed studies have been made, other single-tagged processes should be measurable as well. Of particular interest will be $\gamma\gamma* \to \pi^0, \eta, \eta'$, for which there are detailed QCD predictions of the Q^2 dependence which will allow the direct determination of the parton distributions within hadrons.[11]

Finally, Figure 10 shows that the exclusive hadron pair production reactions will easily reach into the W > 5 GeV region, where the perturbative QCD predictions are reliable. As before, we emphasize that these estimates for exclusive hadron production are based on conservative extrapolations of present data and the event rates at high W (and high Q^2) may be considerably larger.

A B Factory will be a testing ground for QCD using two-photon production of hadron pairs at high W

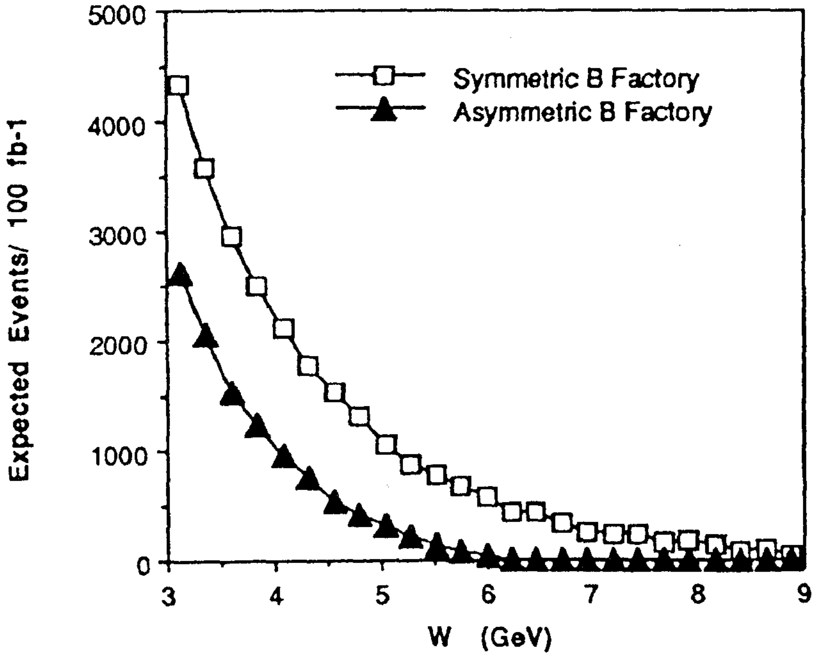


Figure 10. *Expected event samples at asymmetric and symmetric B factories versus W for the untagged process* $\gamma\gamma \rightarrow \pi^+\pi^-$

It is fair to ask what parts of two-photon physics will NOT be readily accessible at B factories. Our Monte Carlo studies indicate that double-tagging will likely be unprofitable at either type of B Factory due to the low rates for detecting both tags at these relatively large scattering angles. It also appears that two-photon inclusive processes will not be easily studied at B factories, since the low center of mass energy tends to push the two-photon and annihilation regions close together, as well as limiting the W and Q^2 ranges available. It is also unclear what the remaining physics questions will be in the inclusive area by the time a B Factory is built.

A B factory will likely not be a good place to pursue two-photon double-tagged and inclusive studies

We have done some specific studies of how the requirements of two-photon physics affect the design of a B Factory and its detectors. Our event rates to this point assumed that all produced hadrons had a minimum polar angle of 0.3 radians. Figure 11 shows the dependence of one of the processes, *The two-photon event rates are not very sensitive to the minimum detection angle for final state hadrons.*

the tagged X(1420) production, on this selection angle. While there is an increase in rate if one pushes the acceptance to smaller angles, it is not sufficiently large in and of itself to warrant heroic efforts in that regard. This relative insensitivity to final state selection angle is characteristic of two-photon exclusive processes.

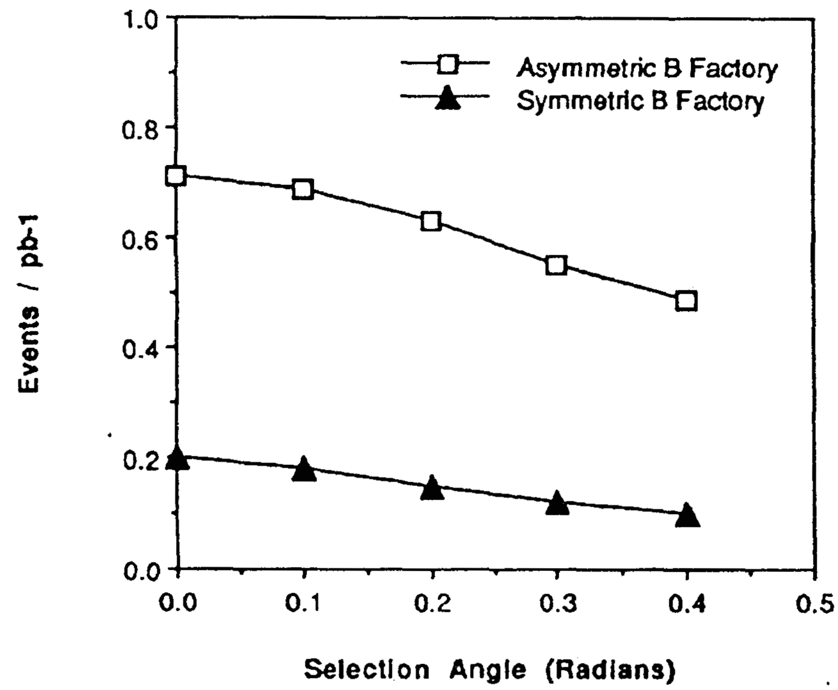


Figure 11. *Events/pb^{-1} at B factories for the tagged reaction $\gamma\gamma^* \to R(J^{PC} = 1^{++}) \to K^{\pm}K^0_S\pm^{\mp}$ versus minimum polar angle for detection of the final state particles*

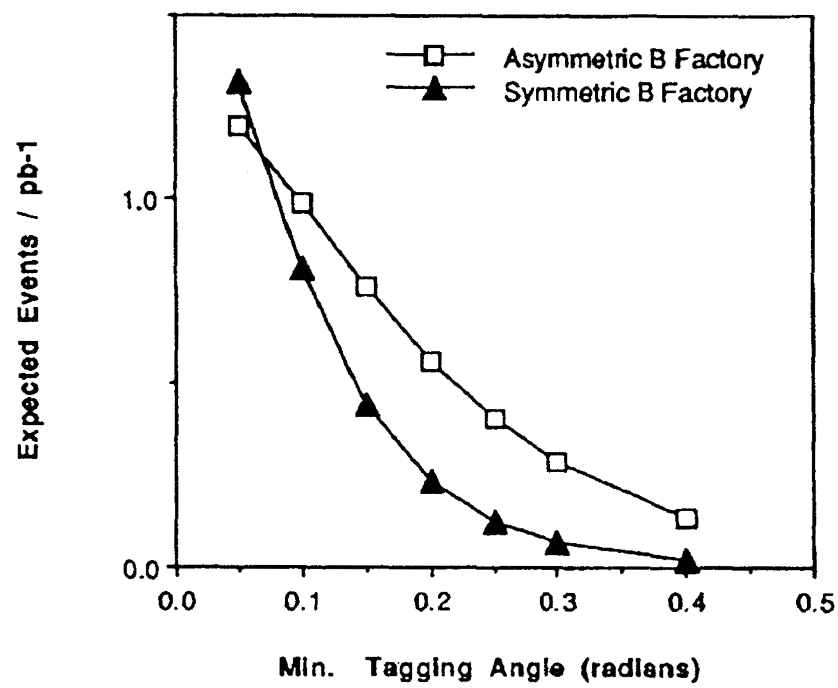


Figure 12. *Events/pb^{-1} at B factories for the tagged reaction $\gamma\gamma^* \to R(J^{PC} = 1^{++}) \to K^{\pm}K^0_S\pi^{\mp}$ versus minimum polar angle for detection of a single tag (scattered e^+or e^-)*

Of much more importance is the minimum tagging angle, for which we also have used a default of 0.3 radians. Figure 12 shows that the rates for tagged X(1420) are quite sensitive to this angle. An even more important reason for preserving some tagging capability at low angles is to maintain acceptance in the low Q^2 region, $0.1 < Q^2 < 0.5$ GeV^2, where the separation of spin-1 from even-spin mesons, and the distinction between positive and negative parities, is cleanest. The Q^2 distributions from the process $\gamma\gamma^* \to R(J^{PC} = 1^{++}) \to K^\pm K^0_S \pi^\mp$ are shown in Figure 13 with $\theta > 300$ and in Figure 14 as a function of the minimum tagging angle. Clearly, there will be almost no data at low Q^2 unless angles less than 300 mrad are possible. Unfortunately, the B Factory machine designs make coverage at small angles very difficult as shown in Figure 15 for one of the most favorable asymmetric configurations. Present detector designs assume a minimum angle of about 300 mrad largely due to these machine constraints and in fact are studying the possibility of shielding against particles at smaller angles to minimize backgrounds from stray beam particles. Obviously two-photon physics needs would be best served by extending the endcap calorimetry coverage to as low angles as possible, even if the energy resolution begins to be degraded by additional material. However, it might also be possible to gain some rate at very low Q^2 by installing, at small angles near the interaction point, a compact (large radiation length) and radiation resistant detector with good resolution for electromagnetic showers caused by the scattered $e^\pm$. At an asymmetric collider, this would be installed only on the low-energy end of the detector. This would require detailed studies as the 'real estate' is very tight in that region and the backgrounds might be prohibitively large. Presumably such a detector would consist either of BGO or one of the newer technologies such as BaF or scintillating fibers, with the additional possibility of including a proportional tube layer in front to discriminate e's from γ's and perhaps a silicon layer at shower maximum to help with e/π discrimination. This would be a very small and very challenging detector to build, especially if was also required to supply information for tagged triggering.

It is important for spin-1 studies to preserve tagging at low angles and thus low Q^2

It might be possible to install a small detector near the interaction point to supply low-angle tagging for spin-1 studies

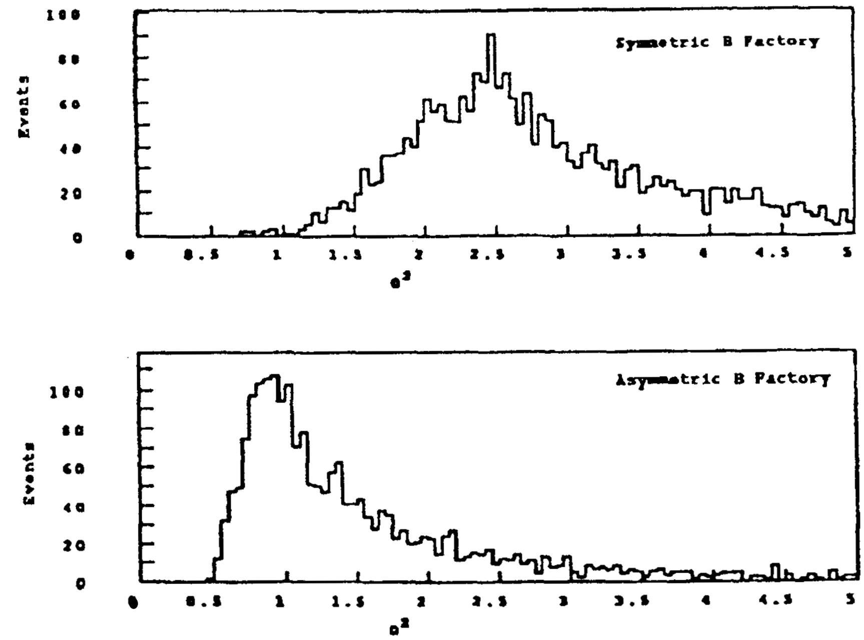


Figure 13. *Q^2 distributions for the process $\gamma\gamma^* \to R(J^{PC} = 1^{++}) \to K^{\pm}K^0_S\pi^{\mp}$ for $\theta > 300$ mrad at a symmetric and an asymmetric B Factory*

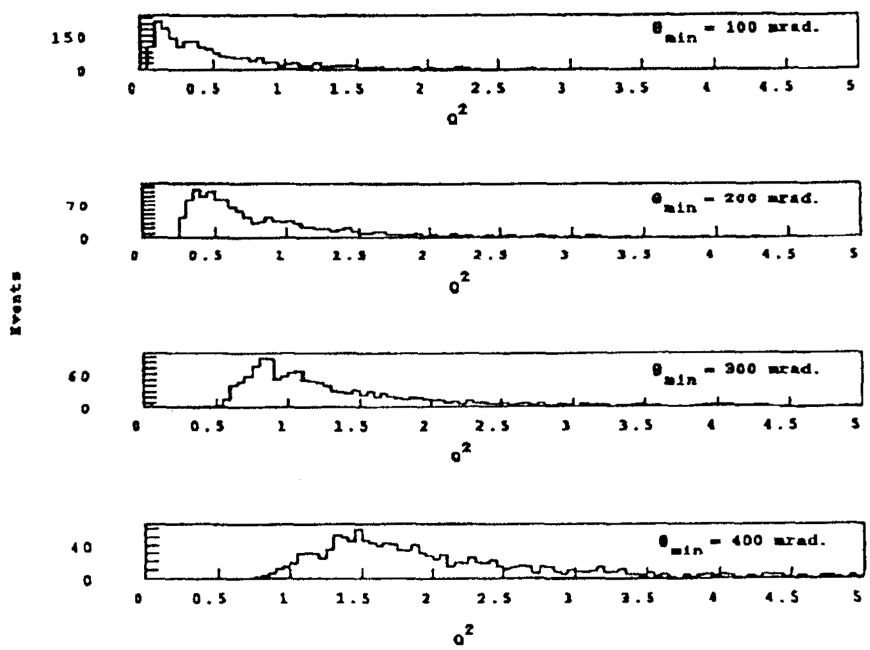


Figure 14. *Q^2 distributions for the process $\gamma\gamma^* \to R(J^{PC} = 1^{++}) \to K^{\pm}K^0_S\pi^{\mp}$ as a function of minimum tagging angle at an asymmetric B Factory*

Whether or not it is feasible to construct such a detector, further study needs to go into the possibility of determining spin/parity of produced resonances using angular distributions or by separating the contributions of longitudinal and transverse photons. It should be emphasized that only the production of spin-1 resonances in the single-tagged mode is affected by the lack of low Q^2 coverage; untagged resonance studies and single-tagged form factor and W dependence measurements for comparison to QCD are not significantly impacted.

Lack of very low Q^2 coverage will not harm the bulk of two-photon physics to be done at a B Factory

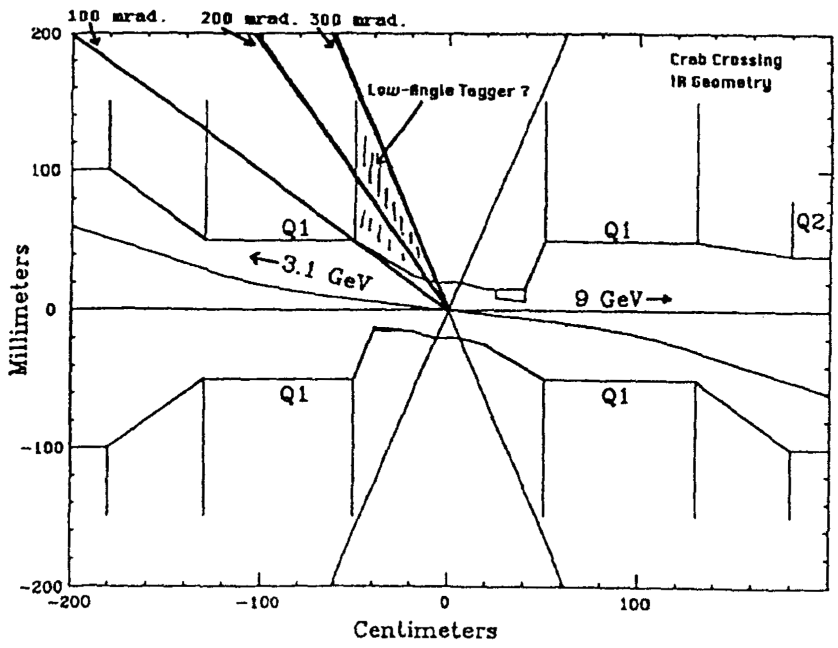


Figure 15. *Interaction region geometry for a crab-crossing asymmetric B Factory design at SLAC*

It is clearly desirable to be able to detect tags at low energy since Q^2 is linearly dependent on tag energy. However, as seen in Figure 16, the two-photon production rates for tagged processes will not be substantially affected unless the minimum tag energies need to be larger than about 2 GeV. Since the calorimetry is designed with good resolution down to at least 0.5 GeV, only the presence of large backgrounds at low angles could necessitate such a large tag energy threshold and such backgrounds will likely also cause problems for B physics. Note that the asymmetric B Factory is more sensitive to the tag energy threshold because tagging occurs primarily on the low energy end (3 GeV in this case).

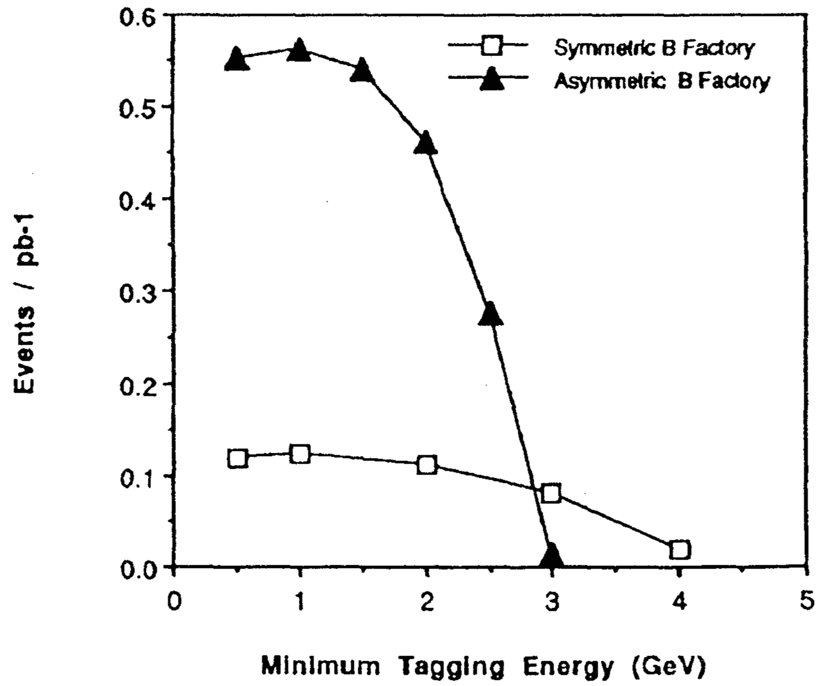


Figure 16. *Events/pb^{-1} at B factories for the process $\gamma\gamma^* \to R(J^{PC} = 1^{++}) \to K^{\pm}K^0_S\pi^{\mp}$ as a function of minimum tagging energy*

The cross section of two-photon hadron production for masses > 1 GeV is small compared to annihilation cross sections at B Factory energies

The triggering schemes envisioned for B physics should be flexible enough to accomodate two-photon physics as well

Extension of the trigger logic needed to reject Bhabha scattering events will allow tagged two-photon triggering

Another important issue for two-photon physics at B factories is the definition of triggering. The total cross section for $e^+e^- \to e^+e^- hadrons$ at the $\Upsilon(4S)$ is about 1nb, as compared to the other physics cross sections $e^+e^- \to$ hadrons of 4 nb, $e^+e^- \to$ leptons of 3 nb and $e^+e^- \to e^+e^-$ of nearly 32 nb! Thus the overall trigger rate will not be substantially increased by triggering on two-photon events, as long as there is a restriction to masses (transverse momenta) well above the QED range ($\approx$ 1 GeV). However, as mentioned before, two-photon events tend to involve a small number of low momentum and low polar angle tracks. For example, it is desirable to be able to trigger efficiently on such processes as $\gamma\gamma \to \pi^+\pi^-$ in the untagged mode and $\gamma\gamma* \to \gamma\gamma$ in the single-tagged mode. These are very challenging topologies for triggering, especially given the usual backgrounds resulting from beam-gas interactions and radiative Bhabha scattering. Recent ideas for triggering at B factories have evolved along the lines of a multi-staged pipelined system. At level 1 (typically about 1μs), one might require either a total energy from the calorimetry of 2 GeV, or the presence of at least 2 calorimeter 'supertowers' each with at least 1/2 minimum ionizing (about 300 MeV) or perhaps a single 'track' from the inner part of the central drift chamber with at least p_t

Preserving triggering to small angles is vital for two-photon physics

of 50-100 MeV. The primary rejection here would be of cosmics and beam-gas backgrounds. The precise definition of calorimeter supertowers is not yet clear but suggestions have ranged from arrays of 5 x 5 crystals to overlapping hemicylinders of the whole calorimeter. Level 2 triggering (at the 10 μs time scale) would then impose tighter requirements such as either 2 tracks in the central drift chamber(one of which could be at low angles) or a total energy of at least 3 GeV in the calorimetry or the presence of 2 calorimeter energy clusters of some minimum energy. The requirement of localized energy cluster finding is essential at this stage to allow recognition and prescaling of the enormous rate of Bhabha events. The same logic that does this can, with simple extensions, be used for tagged two-photon triggering. The final stage of triggering would then be a microprocessor farm which directs the various topologies into different streams, which can be preserved or prescaled as data acquisition rates demand. We believe, based on experience with present detectors, that such a relatively loose trigger scheme should efficiently catch most of the two-photon events of interest, IF it can be made to function in the high-rate B Factory environment. Primarily, the impact of two-photon physics here is to push very hard to preserve the endcap calorimetry and low-angle drift chamber regions in the trigger, especially on the low-energy end at an asymmetric B Factory, where single-tagging will be important. There are other issues which need serious study, such as the effects of the detector magnetic field which might curl up low-momentum tracks before they reach central detector triggering elements. This may necessitate the use of chambers between the central drift chamber and the vertex chamber to supply trigger information. To suppress beam-gas backgrounds, it would be desirable to make use of z information in the trigger. However, one must be careful with such small radius and z trigger requirements so as not to throw away events with K^0_S. Such problems are more serious for untagged two-photon studies, where one is not aided by the relatively high energy of the detected tag in forming a trigger. While it seems premature to try to evaluate trigger efficiencies, we have done some preliminary studies on the effects of the calorimeter 'topological' triggers by studying the trans-

Calorimeter 'topological' triggers should function relatively well for two-photon events verse momentum and azimuthal angle differences between final state particles in a untagged process like $\gamma\gamma \to \pi^+\pi^-$. Although there is a significant impact at low mass, it appears that the higher-mass regions which will be of most interest at a B Factory should survive relatively unscathed. Note that the event rates we have quoted are normalized by results from present experiments and thus include some of the effects of trigger efficiencies already .

Two-photon physics is especially susceptible to backgrounds in low-angle regions

The whole question of backgrounds from such sources as synchrotron radiation and beam-gas interactions is central to the design of the B Factory. It appears that masking solutions exist which can keep synchrotron radiation to acceptable levels in the detector in present B Factory designs. Particles from beam-gas interactions pose a much more serious hazard to the detector, both in terms of occupancy and radiation damage, as well as for triggering. Monte Carlo studies at SLAC have already shown a substantial rate of γ's entering the central drift chamber from beam-gas scattering in the high-energy beam. The particular sensitivity of two-photon physics to such backgrounds comes in two areas. First, as already discussed, if the backgrounds of low-energy particles are too great, the trigger will have to be made more stringent and this may adversely affect some of the two-photon processes of interest. Secondly, if the flux of background particles is too large in the endcap calorimetry (or in any low-angle calorimetry which might be proposed), it could limit the ability to do single-tagged two-photon physics. Fortunately, at an asymmetric B Factory, the tagging would be done on the low-energy end of the detector and it is expected that the flux of particles there will be much less. Further studies of these impacts await the refinement of the background calculations and perhaps actual data from CLEO II.

Single-tagging on the low energy end of an asymmetric collider should be reasonably insensitive to beam backgrounds

As for the impact of two-photon physics on the rest of the B Factory detector components, only particle identification might pose a problem above and beyond what is required to do B physics. Figure 17 gives the momentum spectrum for a typical process, that of untagged production of a low-mass resonance decaying into 4 charged particles.

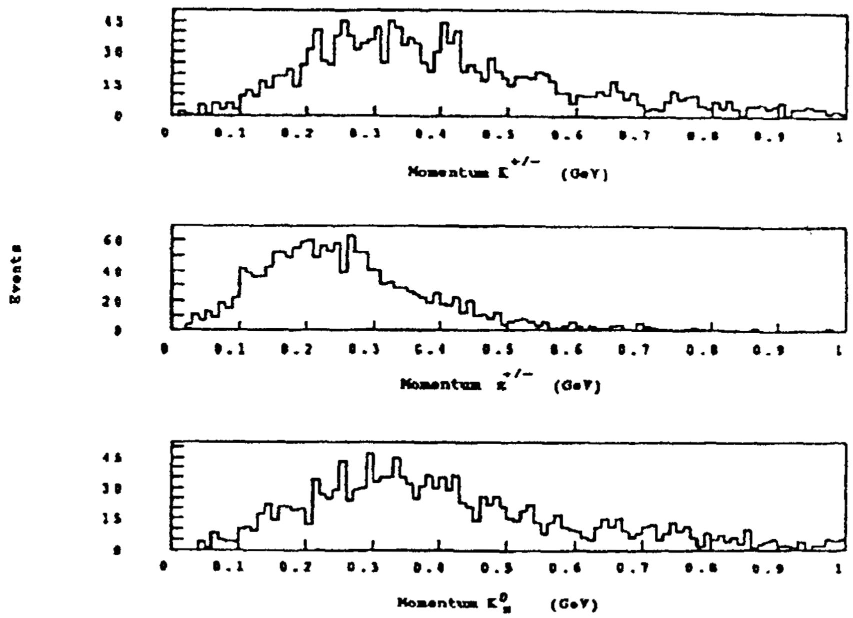


Figure 17. *Momentum of final state particles from the process $\gamma\gamma \rightarrow \iota(1450) \rightarrow K^{\pm}K^0_S\pi^{\mp}$*

Final state hadron momenta are mainly in the few-hundred MeV range, where dE/dx from the central drift chamber should be quite adequate. The upper end of the range, where dE/dx overlaps would occur, would be readily handled by any of the proposed Cerenkov techniques needed for B physics. The other particle identification need for two-photon physics is to distinguish the scattered e^+ tags from pions and from photons. The high-resolution central and endcap calorimetry should provide plenty of e/π discrimination, especially in the important 2-3 GeV region and the central drift chamber should allow e/γ discrimination in the same range. However, both of these are contingent on having full detector coverage to somewhat below the minimum tagging angle to prevent loss of resolution due to leakage effects. Furthermore, there must not be increasing amounts of material at small angles which could degrade tagging recognition. Subject to these caveats, it appears that two-photon physics will have very minimal impacts on the detector design.

Two-photon physics requires moderately good hadron particle i.d. over the full angular range and good $e\pi$ and $e\gamma$ discrimination for tagging

7. Conclusions

Two-photon physics is now a maturing field of investigation which promises a wealth of information concerning the fundamental theory of strong interactions, QCD. Measurements on subjects ranging from meson spectroscopy to hadron pair production will provide a many-pronged challenge to the theorists and may well uncover new physics. Such measurements will, however, require orders of magnitude larger event samples than are available today or even in the next few years. The high luminosities available at a B Factory would make such event samples obtainable. We believe that the two-photon production of resonances and exclusive hadron production can be accomplished at such a facility with only small impacts on the detector coverage, calorimetry and triggering and with no compromise in the important B physics to be done. Thus, a facility built for the very interesting study of CP violation should serve to expand greatly our knowledge of more conventional hadron physics.

Acknowledgements

I would like to thank the other members of the Two-Photon Physics subgroup at the B Factory Workshop at SLAC, Jan-June 1990, including D. Caldwell (UCSB), A. Eisner (UC IIRPA), U. Karshon (Weizmann Inst.), M. Sullivan (UC IIRPA), Y. Wang (UC IIRPA), and S. Yellin (UCSB). I would also like to thank T. Jensen (Ohio State Univ.) for his input from the Cornell effort. It is also a pleasure to acknowledge the help and advice of Stan Brodsky. Finally, I would like to thank the workshop organizers, David Hitlin and Jonathan Dorfan, for all of their prodigious efforts in running such an extended workshop.

REFERENCES

1. Several excellent summaries of the theoretical and experimental status of two-photon physics exist; see *e.g.* S. Cooper, *Ann. Rev. Nucl. Part. Phys.* **38**, 705 (1988); Ch. Berger and W. Wagner, *Phys. Rep.* **146**, 1 (1987); also see proceedings of the VIIIth International Workshop on Photon-Photon Collisions, Jerusalem Hills, Israel, April 1988; ed. U. Karshon (World Scientific, Singapore, 1988).
2. H. Aihara, *et al.*, *Phys. Rev. Lett.* **57**, 2500 (1986).
3. D.A. Williams, *et al.*, *Phys. Rev.* **D38**, 1365 (1988).
4. T. Jensen, CLEO collaboration, DOE/ER/01545-429, talk presented at the 1989 International Symposium on Heavy Quark Physics, Cornell, June 1989.
5. H. Aihara, *et al.*, *Phys. Rev.* **D38**, 1 (1988).
6. G.Gidal, *Proceedings of the VIIIth International Workshop on Photon-Photon Collisions, ibid.*, p. 182.
7. M. Chanowitz, *Proceedings of the VIIIth International Workshop on Photon-Photon Collisions, ibid.*, p.205.
8. H. Aihara, *et al.*, *Phys. Rev. Lett.* **57**, 404 (1986).
9. H. Aihara, *et al.*, *Phys. Rev.* **D36**, 3506 (1987).
10. A. Nilsson, *Proceedings of the VIIIth International Workshop on Photon-Photon Collisions, ibid.*, p. 261.
11. S. Brodsky, *Proceedings of the VIIIth International Workshop on Photon-Photon Collisions, ibid.*, p. 455.
12. See, for example: J.C. Sens, preprint CERN-EP/88-99 and *Proceedings of the VIIIth International Workshop on Photon-Photon Collisions, ibid.*, p. 143.

Interaction Region Considerations for a *B* Factory

H. DeStaebler

Abstract

Machine-detector interface issues at an asymmetric *B* Factory are discussed, with an emphasis on detector backgrounds. A collection of useful formulae is included.

1. Introduction

The goal of the *B* Factory project is to observe *CP* violation in the $\bar{B}B$ system. In order to achieve this goal, the machine must truly operate as a factory for high energy physics, not an R&D project for accelerator physics. (The necessary R&D is supposed to be done before the machine is built.)

A *B* Factory operates at unprecendentedly high beam currents, placing a greater than usual burden on the proper design of masks to shield against synchrotron radiation and energy deposition by lost beam particles. There are a number of interrelated design issues arising from the different desires of the detector and the machine, some of which are listed below.

A number of background and beampipe issues are mentioned. The emphasis is on calculations. Any satisfactory design will combine measurements on existing machines with calculations pertaining to the measurement conditions as well as to the proposed machine.

This article will serve as a general introduction to the problemes of masking the interaction region at a high-luminosity asymmetric-energy *B* Factory. The following articles, by M. Sullivan on detector backgounds from synchrotron radiation and by D. Coupal and C. Hearty on detector backgrounds from scattered beam particles, treat the specific case of the SLAC *B* Factory in more detail.

2. Detector Requirements

It is the limits of the various subsystems of the detector as to occupancy and radiation dose which place the most stingent requirements on the interaction region design.

- Many events are required, which implies high luminosity (this means high $\mathcal{L}_{\text{ave}}$, not only high $\mathcal{L}_{\text{peak}}$), which, in turn, implies high current and small spots. High current is achieved with customary bunch population but much closer bunch spacing. Small spots imply small β^* which, in turn, implies short bunches and fairly large IP angles.
- Good vertex resolution is required, which implies a small, thin beampipe. The IP beampipe will be the smallest physical aperture in the machine.
- Luminosity requirements are reduced by having a moving center of mass, which implies unequal beam energies, which requires two rings. Even with equal energies, two rings are necessary to eliminate the effects of parasitic bunch crossings.
- There is competition between the detector and machine for the scarce real estate near the IP; cables and other services require space.
- The detector will have a solenoidal field of $\approx$1 T extending over $\pm$2 m.
- The detector must experience acceptable backgrounds during luminosity running. For the detector, this means a design relatively insensitive to backgrounds. For the machine, this means a number of masks (both near and far from the IP), an appropriate optics design that minimizes background problems, and a pressure profile that reduces backgrounds.
- Frequent and rapid injection is required to keep $\mathcal{L}_{\text{ave}}$ high. This obviously constrains the machine. The detector must be insensitive to radiation damage during injection, and it must go quickly from data taking to injection and back again. Some kind of rapidly insertable and removable shielding to protect the detector against injection and poor machine performance might be useful.

- Radiation damage to the detector from commissioning and machine physics work should be small compared with the inevitable amounts during luminosity running. The detector will undoubtedly be absent during initial commissioning, but it will be present during the final stages of commissioning, since it is the best instrument to measure backgrounds, and it will be present during recommissioning following shutdowns.
- The design of the detector/machine should be flexible, for example to accomodate different head-on or crossing angle geometries at the IP, or changes in the energy asymmetry.
- Special IR instrumentation is needed. For tuning, whether by operator or by computer, prompt signals proportional to background and to luminosity are needed. Radiation detectors near masks and limiting apertures would be useful in identifying sources of background, as would detectors that were sensitive to only one beam. Since some bunches might contribute more background than others, it might be very educational to be able to identify individual bunches, or at least synchronize the background detectors to the revolution frequency. Possibly special BPMs and corrector magnets should be added.

3. Synchrotron Radiation Background

Bends and quads near the IP are the main sources of synchrotron radiation that cause background problems. For head-on collisions with unequal energy beams, bends are needed near the IP to separate the beams to avoid parasitic bunch collisions. Bend magnets are required to get the beams into the arcs; the final bending should be done at low field to reduce the characteristic energy of the SR.

Masks shield the IP beampipe from direct SR as well as from scattered SR. Only the higher energy photons that eventually can cause problems in the detector are of interest. The SR background can always be reduced by increasing the inner radius of the IP beampipe, but at the expense of degrading detector resolution. Almost all the SR photons go through the IP beampipe without hitting any masks, but they all eventually interact somewhere.

Appendix A contains a collection of useful synchrotron radiation formulas.

3.1. Sources of SR

A bend magnet produces a fan of radiation with the extreme rays being the incoming and outgoing beam trajectories. Usually the bend angle is large enough so it determines the width of the SR swath. Perpendicular to the bend plane the height of the swath is determined by the size and angular divergence of the radiating beam plus the intrinsic angular distribution of the SR photons relative to the radiating trajectory.

A quadrupole produces a much more complicated beam of SR than a bend magnet. A program commonly in use at SLAC that calculates radiation from quadrupoles, and traces the photons through a series of masks, QSRAD,[1] makes various approximations: (a) each ray of the beam emits a flat fan, (b) a single value of B characterizes the SR, derived from an average offset, and (c) the intrinsic angular spread in the beam is neglected. The four-dimensional integral over the beam distribution (x, $x\prime$, y, $y\prime$) is reduced to an integral over only x, y. Evenly spaced rays are traced through the system and the photons from each ray are weighted by the chosen beam density at that x, y. Or an external ray distribution file from an optics calculation may be used as input to QSRAD. A partial check on the approximations can be made by dividing physical quadrupoles into a number of shorter quads for computational purposes.

The distribution of photons from a quad depends on the transverse distribution of the beam, which often becomes more poorly known the farther one gets from the beam axis (see Section 6). This is especially true for the outer part of the photon beam which is most likely to hit masks and cause detector background.

3.2. *Masking and Mask Reradiation*

Masks shadow the detector beampipe from photons coming directly from the magnets. Unfortunately, there is no such thing as a perfectly black mask; that is, every photon hitting a mask has some probability of reradiation, depending on energy, angle, material, and geometry. So frequently secondary masks shadow primary masks, and so on. Table 1 gives some representative reradiation probabilities calculated with EGS4[2] for forward scattering from a mask. A photon beam with $k_c = 15\text{keV}$ and width 1 cm is incident on a rectangular mask, starting from the edge. All photons are scored that scatter out with $\theta < 11.5°$. K-shell fluorescent radiation is included. All emerging photons had $k > 30$ keV.

Material	$\frac{1}{n_{\text{in}}} \frac{dn_{\text{out}}}{d\Omega}$ (sr^{-1})
Ta	3×10^{-6}
Cu	4×10^{-5}
Al	7×10^{-4}

Table 1. ***Representative reradiation probabilities for forward scattering from a mask***

High Z masks are better because all the cross sections per atom increase with Z, and in addition the major absorption cross section (photoelectric effect) grows more rapidly with Z than the scattering cross sections. The probability of fluorescence radiation following photoelectric absorption increases with Z. Fortunately, it is not as severe for the softer spectra characteristic of B Factories as it is at linear colliders, for example. Often, masks are coated with thin layers of lower Z materials on top of higher Z to optimize the competition between absorption and reradiation. This applies as well to the inside of the IP beampipe.

4. Background From Lost Beam Particles

Beam particles hitting the masks and beampipe near the detector will send degraded shower debris into the detector. As is well known, there are no black masks for high-energy beam particles. IR masks honor a beam-stay-clear that is supposed to keep beam tail particles from hitting them. This means that a distant mask system shadows the masks close to the IP. However, beam-gas interactions relatively close to the IR may cause beam particles to hit the inner masks depending on details of optics, masking, and residual pressure.

4.1. Beam-gas Bremsstrahlung

The cross section for fractional energy loss u by radiation is approximately[3]

$$\frac{d\sigma}{du} = 4\alpha\ r_e^2\ Z\,(Z+1)\ \frac{4}{3u}(1-u+.75u^2)\ell n\left(\frac{183}{Z^{1/3}}\right) \quad (4.1)$$

$$u\ =\ \frac{k}{E}\ . \quad (4.1a)$$

The $Z+1$ takes approximate account of radiation from the atomic electrons.[4] Note that the radiated photons themselves may be a noticeable source of background, even though their average energy is only a fraction of the energy of the beam. The angular distribution of the radiation process is usually neglected in this application. The angular distribution has characteristic angle $1/\gamma$ (that is, the transverse momentum is about mc).[5]

4.2. Beam-gas Nuclear Coulomb Scattering

The cross section for Rutherford scattering at polar angle θ (taken much less than 1) is[3]

$$\frac{d\sigma}{d\Omega} = 4r_e^2\ Z^2 \frac{\left(\frac{m}{p}\right)^2}{(\theta^2+\theta_1^2)^2}\ , \quad (4.2)$$

$$\theta_1 = \alpha Z^{1/3}\left(\frac{m}{p}\right)\ . \quad (4.2a)$$

The screening of the atomic electrons is accounted for by the angle θ_1. Any nuclear form factor effects are neglected, which requires $q \approx E\theta < q_{\max} = 137\ \mathrm{m}/A^{1/3}$. The energy lost by the beam particle is $q^2/2A$ which can safely be neglected.

4.3. Coulomb Scattering from Atomic Electrons

This is Rutherford scattering of the beam particles from free electrons. Changing Z^2 to $Z(Z+1)$ in Eq. (4.2) will roughly take account of this. One might worry that the fractional energy loss on a light target, which is approximately

$$u = \frac{q^2}{2m}\frac{1}{E} = \frac{\theta^2\gamma}{2}\,, \tag{4.3}$$

might be a concern. However, calculations show that energy losses greater than the natural energy spread correspond to a small scattering cross section.

4.4. Number of Beam Particles Hitting Masks

The products of beam-gas interactions are transported through the optical system to well beyond the IP. It is convenient to use the program DECAY TURTLE[6] to track the beam-gas interaction products through a system of optical elements and masks. Note that in SLC calculations it was found that including sextupoles affected the tracking results, presumably because large amplitude particles are important.[7] The source probability is weighted according to the pressure profile and the composition of the residual gas.

The rate of particles hitting masks can be estimated as follows: Take 6.25×10^{12} beam particles (corresponding to 1 A for 1 μs) traversing 10 m of CO (37.42 g/cm^2 radiation length)[8] at a pressure of 10^{-8} Torr, which comes to 4.1×10^{-13} radiation lengths. Consider all bremsstrahlung collisions that radiate more energy than 10 times the natural energy spread in the beam, taken as 10^{-3}. The rate is $(6.25 \times 10^{12}) \times (4.1 \times 10^{-13}) \times 4/3 \times \ell n(100) = 16/\text{A} - \mu\text{s}$.

4.5. Reradiation of Shower Debris into the Detector

The reradiation probability into the detector is greater for particles that hit near the IP, but distant sources must be evaluated numerically. Reradiation is also greater for particles that hit the face or near the edge of a mask. The shower debris also has to be transported through the lattice, the masks, the detector magnetic field and into the detector. EGS4[2] can be used to follow the shower debris through the beampipe into the detector.

4.6. Lost Electrons from Interactions with Trapped Ions

The beams ionize the residual gas. The attractive field of the e^- beam traps positive ions, concentrating them near the beam orbit. The e^+ beam repels the ions; the ionization electrons are too light to be trapped.[9] Techniques are available to un-trap the ions.[10] For background purposes, the beam-ion interactions are similar to the beam-atom interactions discussed above.

4.7. Lost Particles from Beam Interactions with Thermal Photons

This effect is small for B Factory conditions,[11] but is measurable at LEP. [12]

5. Other Sources of Background

5.1. Injection Shield

Radiation damage during luminosity running is likely to be significant, so it is important to reduce damage during injection as much as possible. An attractive idea is a massive shield that can be quickly inserted inside the main drift chamber to provide protection to all the detector elements except the silicon vertex detector. This would supplement any other possible protective measures.

5.2. *Multiple Reflections of Synchrotron Radiation*

The calculations in Section 3 typically take account of 2 or 3 photon reflections. One always worries that there is some efficient mechanism involving multiple reflections for transporting SR over long distances to the IP. I do not know of such a process. Total external reflection[13] requires exceedingly smooth surfaces and only occurs for photon energies that are quite low by our standards. Multiple forward Rayleigh scattering is diminished by the competition with photoelectric absorption; any fluorescence reradiation is isotropic. X-ray diffraction scattering from the polycrystalline wall of the beampipe is also in competition with photoelectric absorption.

5.3. *Gas Interactions near the IP*

Section 4 dealt with beam-gas interactions fairly far from the IP which caused beam particles to hit masks near the IP. There are also interactions near the IP (within the z acceptance of the detector) that send background particles directly into the detector. Consideration of these processes may set a restrictive limit on the IP pressure.

(a) Some convenient FORTRAN programs calculate a number of e and γ interactions on nuclei, including quasi-elastic and inelastic electron scattering and various photopion reactions.[14] These are a useful supplement to rates measured with random triggers on existing machines. The recoil proton cross sections agree with measurements.[15,16]

(b) SR photons can scatter into the detector by Compton or Rayleigh scattering; photoelectric fluorescence is not a problem from C and O because the K edges are below 1 keV. For $x \equiv k/m < 1$, approximate values[17] for Compton scattering per free electron are shown in Table 2:

Table 2. *Approximate Compton scattering probability per free electron*

Scattered Angle (degrees)	$d\sigma/d\Omega$ (r_e^2/sr)	k'/k
0	1	1
90	$0.5/(1+2x)$	$1/(1+x)$
180	$1/(1+4x)$	$1/(1+2x)$

Rayleigh scattering in our energy range is a more complicated function of k and Z. Useful fits to the cross sections based on a Fermi-Thomas atom[18] are

$$\sigma_{\mathrm{tot}} = \frac{8\pi}{3}\, Z^2 r_e^2 \left[1 + \left(\frac{B}{1.394}\right)^{1.162}\right]^{-1.628} \tag{5.1}$$

$$B = \frac{2k}{\alpha m Z^{1/3}} = \frac{k\,(\mathrm{keV})}{1.865 Z^{1/3}}\ . \tag{5.1a}$$

and

$$\frac{d\sigma}{d\Omega} = Z^2 r_e^2\, \frac{1+\cos^2\theta}{2} \left[1 + \left(\frac{U}{1.186}\right)^{1.199}\right]^{-2.436} \tag{5.2}$$

$$U = B \sin\frac{\theta}{2}\ . \tag{5.2a}$$

The fit ranges are $0 < B < 10$ and $0 < U < 40$. The Fermi-Thomas model is pretty good for high Z. But, for example, Eq. (5.2) overestimates the cross section for C and O by about a factor of 2 at $U = 12$, and a factor of 5 at $U = 40$.

As an order-of-magnitude estimate of this background, consider a beam of 1 A for 1 μs with 1 photon per electron, a pressure of 10^{-8} Torr N_2, an acceptance of 1 m and 2π sr, and evaluate the cross sections at 90°. The results are shown in Table 3. A proper calculation would integrate the SR fluxes from both beams over the actual cross sections and acceptances, and include absorption in the beampipe.

k (keV)	Number of Scattered Photons		
	Rayleigh	Compton	Total
5	0.9	0.7	1.6
10	0.3	0.7	1.0
20	0.06	0.7	0.8

Table 3. *Rayleigh- and Compton-scattered photons at three different energies, under conditions specified in the text*

5.4. *Synchrotron Radiation-Beam Interactions*

The bunches of synchrotron radiation photons and the charged beam bunches collide at the IP and at $s_b/2$. Most of the interactions are Compton scattering, like a back-scattered laser beam,[19] although some of the highest energy photons will make pairs.[20] The interaction rates are not high, and the reaction products have low p_t and make small angles with the beam axis.

5.5. *Photon Radiation from a Transverse Crab Cavity*

Both synchrotron radiation from the transverse kick, and Compton scattering from the RF photons in the cavity[21] are very weak.

5.6. *Background from RF Cavities*

At PEP, the DELCO detector experienced background from the RF cavities in a nearby straight section, apparently from field emitted electrons that were accelerated in the cavities. These were eliminated by putting a weak magnetic bump on the IP side of the closest cavity.[22] DELCO was an open detector with little self-shielding; the other PEP detectors with adjacent RF, MkII and MAC, had no such problem.

5.7. *Photo-Hadrons and Photo-Muons*

Is it possible that the effects of hadronic/muonic debris from lost particles hitting near the detector could be more severe than the electromagnetic debris, possibly in causing triggers? It doesn't seem likely, but should be checked.

6. Beam Shape

The beam distribution near the IP affects both the SR and lost particle backgrounds. The beam size in the final quads affects the distribution of SR photons in number, energy, and spatial extent that the masks are designed to cope with. The lost particle rate depends on the beam distribution through over-focusing of low-energy particles in the final quads. It is useful, although somewhat artificial, to divide the beam into a central Gaussian core plus a halo or tail extending to many sigma.

6.1. *Gaussian Core*

For a single beam, the core size is set by the emittance (SR fluctuations) and the optics. However, the beam-beam interaction increases the core size, especially in the vertical direction (for flat beams), and this is seen in the luminosity.[23] A simulation of a PEP-like machine gave an increase in σ of 5% in the horizontal and 10% in the vertical.[24]

6.2. *Halo or Tail*

The halo is generated by beam-gas interactions, the beam-beam interaction, nonlinear aspects of the optics encountered at large excursions, and the resonant and tune structure. (It seems to be the conventional wisdom that when beams are first brought into collision after a fill, the backgounds get worse, implying that gas scattering alone does not set the halo.) The halo distribution is in dynamic equilibrium between the processes tending to kick particles out, and radiation damping tending to bring them back. Beam lifetime is related to the distribution near the limiting aperture.[25]

There are several approaches to arriving at a model to use for the halo distribution, but, to my mind, none is completely satisfactory. This is unfortunate, because a bold SR masking scheme would depend critically on the halo distribution.

(a) Computer Tracking Simulation

One might think that the halo could be calculated since all the processes are known, with the possible exception of non-linearities at large radius. Simulations for the core seem relatively satisfactory, but the present beam-beam codes are not designed to accurately predict the small population ($\sim 10^{-5}$) in the halo.

(b) Fit a Model to Measurements

Measurements at CESR of the vertical beam distribution clearly showed a tail.[26] This was fit to the following forms

$$\frac{1}{n}\frac{dn}{dy} = \frac{1}{\sigma\sqrt{2\pi}}\left[\exp\left\{\frac{-y^2}{2\sigma^2}\right\}\right],$$

or

$$3.7 \times 10^{-6} \exp\left\{-1.2\left(\frac{y}{\sigma} - 5\right)\right\}, \tag{6.1}$$

depending whether y/σ is less than or greater than 5, and used for subsequent SR calculations.

Background measurements at PEP, assumed to come entirely from SR, were used to adjust the parameters of an assumed Gaussian tail,[27] which was used for subsequent SR calculations.[28]

The problem with the first approach is the basic assumption that the vertical distribution will be the same in the new machine of interest. But the beam-beam simulations indicate significant sensitivity to various machine parameters. To use this approach with confidence, one should demonstrate scaling.

The second approach suffers the same shortcomings as the first, and in addition, it is only an integral measurement.

(c) Semiquantitative (Qualitative?) Approach

Ritson argues that the relative population in the halo should be roughly the ratio of damping time to beam lifetime, and it should fall off relatively slowly, say, as a power ~ 4 or so, rather than as an exponential or Gaussian.[29]

(d) Conservative Approach

Assume a flat background out to the limiting aperture with a population larger than is implied by beam lifetime. Since presumably this is a worst case, it is useful to at least check a mask design against it. What to do if the worst-case backgrounds are too high is another question.

6.3. *Limiting Apertures*

As masks get closer to the IP, their size in sigma units should increase. The limiting apertures should be far from the IP, designed to shadow the IP region.

7. Heating and Cooling

Heat loads on various beampipes, masks, and surfaces need to be specified so that adequate cooling can be provided. Possible problem areas are cooling the IP beampipe, which will decrease the IP resolution, and high SR power densities. Allowable temperature rises need to be established and the consequences of thermal expansion investigated. The final temperature of an object depends on the relative rates of heating and cooling.

7.1. *SR Heating*

SR heating of masks near the IP is usually small, since a significant heat load would be an intolerable background source ($1W = 6.25\ 10^{15}\,\mathrm{keV/s}$).

Machines with head-on collisions and small bunch separations (1–2 m) produce dozens of kW of SR from the bend magnets initiating the orbit separation and in beams off-axis in common quadrupoles. (The irreducible SR from the quad focusing is roughly 10 times less.) This power must be conducted to a water-cooled dump, possibly first going through a very thin window in the vacuum pipe. The transverse power density can be high, and the initial rate of energy absorption is also high.

7.2. *Image Current Heating*

All beampipes are heated on the inside by image currents flowing in the skin depth. This is basically a bunched beam pulse heating. The appropriate formula for a Gaussian beam is[30]

$$\frac{dP}{dz} = \frac{\Gamma(3/4)}{4\pi^2 a} \frac{s_b \langle I \rangle^2}{\sigma_z^{3/2}} \sqrt{\frac{\mu Z_0}{2\,\sigma(Z)}} \tag{7.1}$$

$$= 2.75 \left(\frac{W}{m}\right) \left(\frac{2cm}{a}\right) \left(\frac{1cm}{\sigma_z}\right)^{3/2} \times \left(\frac{s_b}{1m}\right) \left(\frac{\langle I \rangle}{1A}\right)^2 \left[\frac{\mu\,\sigma(Cu)}{\sigma(Z)}\right]^{1/2} . \tag{7.1a}$$

7.3. *HOM Heating*

A more serious source of heating comes from the RF power radiated by the beams as they traverse changes in size and shape of the beampipe. This is frequently referred to as higher order mode power (HOM) because the RF cavities are a major discontinuity in the beampipe and the radiated energy typically has higher frequencies than the cavity fundamental. There are two fairly separate parts to the problem: how much HOM power is radiated, and where is it absorbed. All the HOM power is absorbed somewhere. It's just a question of providing enough cooling at the right places.

The energy radiated when a bunch passes a discontinuity is

$$U = k\, q^2 , \tag{7.2}$$

where q is the bunch charge, and k is a loss parameter depending on the geometry of the discontinuity and on the bunch length (frequency spectrum). k is usually given in V/pC. The power radiated is

$$P = f_b\, U = k\, \frac{s_b}{c}\, \langle I \rangle^2 \; , \tag{7.3}$$

$$f_b = \frac{c}{s_b} \; , \tag{7.3a}$$

$$\langle I \rangle = f_b\, q \; . \tag{7.3b}$$

Analytic expressions for k are available for simple geometries. For more complicated (realistic) geometries, computer codes are available (for example, MAFIA, TBCI, URMEL).[31] These computer calculations are tending to replace the experimental determination of k values.

Consider a cylindrical pipe of radius a that abruptly increases to radius b for a distance g before returning to a. Approximate expressions for k are [32,33]

$$\begin{aligned} k &= \Gamma\left(\frac{1}{4}\right) \frac{Z_0 c}{4\pi^{5/2}\, a} \sqrt{\frac{g}{\sigma_z}} \\ &= 1.023\, \frac{Z_0 c}{2\pi^2\, a} \sqrt{\frac{g}{\sigma_z}} \; , \quad g \, < \, g_c \end{aligned} \tag{7.4}$$

$$k = \frac{Z_0 c}{2\pi^{3/2}\, \sigma_z}\, \ell n\, \frac{b}{a} \; , \qquad g \, > \, g_c \tag{7.5}$$

$$g_c \approx \frac{(b-a)^2}{2\sigma_z} \; . \tag{7.6}$$

Equation (7.4) corresponds to a pillbox, and Eq. (7.5) to a step down in radius (there is very little loss for a step up in radius). Reference 32finds quite good agreement between Eqs. (7.4)-(7.6) and the code TBCI. Tapering a transition reduces k, but not by more than a factor of 2. [34] The length of the taper need not exceed $2g_c$. [34]

These results are for a single bunch traversing a single cavity. For high enough Q and short enough bunch spacing, interference effects may become important. [32,35,34]

7.4. *Absorption of HOM*

All the HOM energy is absorbed somewhere, in a few skin depths on the inside surface of the beampipe. The absorption is more complicated than the generation, since it depends on the mode and frequency distribution of the energy. High frequencies can propagate down the beampipe. Low frequencies are trapped in the cavities or are rapidly attenuated in the beampipe. The critical wavelength is comparable with the diameter of the beampipe.[36,37] The propagating energy is absorbed with a characteristic $1/e$ length of roughly

$$\lambda_e \sim (100 - 300 \text{ m}) \left[\frac{\sigma(Z)}{\mu\sigma(Cu)}\right]^{1/2} \left(\frac{\text{radius}}{3 \text{ cm}}\right) , \qquad (7.7)$$

which depends on mode and also on frequency relative to cutoff.

The small IP beampipe could have the highest cutoff frequency, so it might absorb HOM generated far away. One might contemplate isolating the IP with a lossy section of ferromagnetic stainless steel.[38]

Billing has discussed HOM generation and absorption in the context of CESR.[39] He reports $k = 0.09\,\text{V/pC}$ for a 2 inch ID SR mask in a 4 inch ID pipe with 27° tapers, and $k = 0.014\,\text{V/pC}$ for about 3 m of IP beampipe with gently tapered (2°–5°) transitions. He makes the interesting point that configurations with large k scale roughly as σ_z^{-1}, whereas those with small k scale roughly as $\sigma_z^{-(2-4)}$. Presumably as the scale of the geometrical irregularities approaches σ_z, the dependency on σ_z becomes stronger; however, for irregularities much smaller than σ_z, one would expect k to approach zero.

7.5. Acceptable Temperature Rise

Preliminary calculations indicate the necessity of active cooling of the beampipe at and near the IP.[40] Once this big headache is accepted, it's just a question of deciding how much cooling is needed and how to supply it. It is possible that very little of the beampipe anywhere can be adequately cooled simply by convection to ambient air.

Thermal expansion and stresses set limits to temperature rise. Also, thermal desorption increases with temperature (see Section 9 on vacuum).

8. Acceptable Detector Backgrounds

The effects of background on the detector elements are usually divided into three categories: radiation damage, extra hits (occupancy) which confuse tracking and pattern recognition, and false triggers.

The first is cumulative; the second two accumulate over resolving times of the order of a microsecond, and depend on the details of the detector. For radiation damage, one might design for a useful life of five years (of luminosity running plus injection and machine physics) with some safety factor added.

8.1. Silicon Vertex Detector

The detector elements themselves are relatively insensitive to radiation damage with acceptable levels of the order of Mrad.[41] However, the associated electronics, which is mounted on or near the detector elements, is more sensitive by a factor of 10 or more.[42] This is presently a field of active research for SSC applications and one can hope for increased radiation hardness on a time scale of interest to a B Factory.[43]

There are so many channels in a pixel detector that occupancy is not the limit, and even in a strip detector occupancy is less of a limit than radiation damage. To see this, consider a strip 25 μm wide by 20 cm long, uniformly irradiated by charged tracks. The flux that produces 0.1 Mrad/10^7s is 3.1 10^5 tracks/cm^2-s corresponding to an occupancy of $0.016/\mu$s.

8.2. Main Drift Chamber

Avalanches at the sense wires cause the accumulation of deposits that degrade performance. This only occurs when the HV is on, so this is mainly a concern during luminosity running, assuming that a fast HV ramp is provided for injection. The degradation is proportional to the integrated charge density on the wire. Present limits are around 1 C/cm,[45] and encouraging progress is being made in identifying the role of trace impurities,[44,46] so I believe one may reasonably take this as a design value for a B Factory. Note that 1 C/cm spread over a 1 m wire for five years of 10^7s each corresponds to 2 μA average current.

First, compare radiation damage and occupancy for charged tracks. Assume that a track at normal incidence gives 0.8 pC at the wire (for example, 100 ion pairs with gain 5×10^4; with 30 eV/ip, this corresponds to 3 keV deposited per track). Then 2 μA corresponds to 2.5×10^6 tracks/s or an occupancy of 250% per μs. Occupancy thus sets a much more severe limit than radiation damage. This result is independent of gas and cell size through the assumption of constant charge per hit. Inclined tracks give more radiation damage for the same occupancy.

For SR photons, the relationship between radiation damage and occupancy is similar to that for charged particles, but there tends to be more radiation damage per unit occupancy. This arises from the energy spectrum of the photons, which interact mainly by photoelectric effect and, at higher k, Compton scattering. Low-energy interactions always produce damage, but may not trip a discriminator for a hit. High-energy interactions can produce much more pulse height (hence damage) than necessary for a hit. Note that for a given SR flux, a He-based gas will have many fewer interactions than an Ar-based gas.

Compton scattering at low energies is not very effective at transferring energy to the recoil electron. The average kinetic energy is approximately[17]

$$\langle T \rangle \; = \; k \, \frac{x}{1 + \frac{11}{5}\, x} \, , \qquad x \; = \; \frac{k}{m} \, . \tag{8.1}$$

8.3. Calorimeter and CRID

One must consider background effects in other detector elements. For example, CsI, a frequently considered material for an electromagnetic calorimeter, seems to be especially sensitive to radiation damage.

8.4. Rare Earth Permanent Magnets

For relatively radiation hard material (Sm_2Co_{17}), the tolerable exposure is roughly 10^{10} rad.[47]

8.5. Triggers

Triggers require one or two fairly high energy tracks and/or a significant energy deposition. These probably come only from lost particles. One needs a fairly detailed model of the detector to estimate the rate.

9. Beampipe Pressure/Vacuum

Beam-gas interactions in the IR set limits on acceptable pressure. Residual gases are mostly hydrogen with a quarter to a half of CO_2 and CO. The principal uncertainty associated with vacuum design for a storage ring is related to desorption of gas by synchrotron radiation.

9.1. Some Formulas

In a system, the rate of gas flow per unit pressure drop defines the conductance.[48,49] (Note: the quantity of gas is measured in Torr-liter, with 3.3 10^{19} molecules/Torr-liter at 20° C.)

$$\text{Throughput} = \text{Pressure drop} \times \text{Conductance} \tag{9.1}$$

$$\dot{Q}\ (\text{Torr-l/s}) = \Delta P\ (\text{Torr}) \times F\ (\text{l/s}) \tag{9.1a}$$

Two simple but useful examples of conductance follow. For a small hole with area A

$$F = \frac{A\bar{v}}{4} \quad \text{or} \quad F_0 = \frac{F}{A} = \frac{\bar{v}}{4}, \tag{9.2}$$

where $\bar{v}$ is the average speed of the molecules. For a Maxwellian distribution

$$F_0 = \frac{1}{4}\sqrt{\frac{8}{\pi}\frac{kT}{M}} = 11.4\left(\frac{\mathrm{l/s}}{\mathrm{cm}^2}\right)\sqrt{\frac{T}{293\mathrm{K}}\frac{28}{M}}. \tag{9.3}$$

For a round pipe with diameter d and length L

$$F = F_0\,\frac{\pi}{4}\,d^2\;\;\frac{4}{3}\,\frac{d}{L}. \tag{9.4}$$

The effective conductance of a pump (which conducts the gas to a land of no return) is called its pumping speed, S (liter/s), so a pump has associated with it a pressure drop

$$S = \frac{\dot{Q}}{\Delta P}. \tag{9.5}$$

To get the pressure at a particular point, add up the values of ΔP from the pump to the point.

9.2. Sources of Gas

(a) Thermal Desorption

The desorption rate depends on the material, how it has been cleaned, how clean it has been kept, and the temperature. A reasonable, ballpark value (at room temperature) is

$$q = 10^{-11}\ \text{Torr-l/s-cm}^2. \tag{9.6}$$

This follows an Arrhenius temperature dependence. The effective binding energy can range from 0.1 to several eV,[49] which corresponds to an increase in desorption per 10° C from 15% to a factor of 50. In a mixed Al/SS system coming off bake, I observed a factor of 1.6, corresponding to 0.35 eV.

(b) Photodesorption by Synchrotron Radiation

SR photons desorb gas from masks and beampipes. The yield of molecules/photon depends on photon energy[50] but usual practice seems to simply use an average value, for example,[51]

$$\eta = 5 \times 10^{-6} \text{ molecule/photon,} \tag{9.7}$$

where only photons with $k > 5$–10 eV are included.

The value of η depends on the angle of incidence of the photons (90° is perpendicular incidence). The angular variation of relative η measured on a clean, lightly scrubbed Al surface using SR with $k_c = 3$ keV is given in Table 4.

Table 4. *Relative photdesorption coefficients at several angles*

Angle of Incidence	Relative η
90°	1.0
16°	2.0
5°	4.0
19 mrad	5.6
11 mrad	7.0

Continued irradiation by SR photons ("scrubbing") reduces η approximately as the 2/3 power of the accumulated exposure.[51] Surfaces irradiated by high-flux primary SR scrub faster than surfaces irradiated by low-flux scattered SR. Hence, eventually scattered SR can contribute a significant gas load.

Compact radial ion pumps using the detector field may be useful inside the detector. [52]

10. Other Issues

10.1. Safety Factor

What degree of conservatism is appropriate in designing the IP region? How much insurance should be provided against tolerances, misalignment, our imperfect understanding, possible machine upgrades, and the vagaries of the real world? Should one take seriously the goal of a turnkey B Factory, which implies a brief detector commissioning period, as apparently has happened at LEP? Or should one anticipate a period of development following first operation, with the possibility of significant modifications, as has often happened in the past?

10.2. Approximations

Approximations enter in two ways, and we need to be sure they are adequate. Approximations are made in calculating a particular background process, although all the basic cross sections are well known. We also make approximations in deciding which are the dominant processes, and which can be neglected.

10.3. Comparison with Actual Experience

It is valuable, and possibly essential for a successful design, to compare our calculational techniques and procedures with data from a real detector at a real storage ring, to check whether our understanding is in tune with nature. Acceptable agreement does not assure success at a B Factory, of course, because scaling from one machine to another is imperfectly understood. But disagreement should surely cause hard thinking and lost sleep.

APPENDIX A

FORMULAS FOR SYNCHROTRON RADIATION

11.1. Bend Magnets

These formulas pertain to electrons in circular motion.[53,36,54] The average number of photons radiated in path length ds is

$$\frac{dn}{ds} = \frac{5}{2\sqrt{3}} \frac{\alpha\gamma}{\rho} , \tag{A.1}$$

$$\begin{aligned} n &= \langle n \rangle = 20.6 \text{ E(GeV)}\phi \text{ (rad)} \\ &= 0.618 \; B \; L(\text{kG}-\text{m}) \; . \end{aligned} \tag{A.1a}$$

The spectrum is a universal function of the characteristic energy

$$k_c = \frac{3}{2}\gamma^3 \frac{\hbar c}{\rho} = \frac{3}{2} \frac{\gamma^3}{\alpha} \frac{r_e}{\rho} mc^2 , \tag{A.2}$$

$$= 2.22(\text{keV}) \left(\frac{E}{10 \text{ GeV}}\right)^3 \left(\frac{1 \text{ km}}{\rho}\right) , \tag{A.2a}$$

$$= 6.66(\text{keV}) \left(\frac{E}{10 \text{ GeV}}\right)^2 \left(\frac{B}{1 \text{ kG}}\right) . \tag{A.2b}$$

The normalized photon energy is

$$v = \frac{k}{k_c} . \tag{A.3}$$

The normalized number distribution of the photons is

$$\frac{1}{n}\frac{dn}{dv} = \frac{3}{5\pi} \int_v^\infty K_{5/3}(y)\, dy , \tag{A.4}$$

$$\approx 0.4105\, v^{-2/3}(1 - 0.8438\, v^{2/3} + 0.0 v^{4/3} + \ldots) \quad v << 1 ,$$

$$\approx \frac{3}{5\sqrt{2\pi}} \frac{e^{-v}}{\sqrt{v}} \left(1 + \frac{55}{72}\frac{1}{v} - \frac{0.9791}{v^2} + \ldots\right) \qquad v >> 1 .$$

Half of the energy is carried above $v = 1$, by only 8.7% of the photons. Half of the photons are above $v = 0.078$.

For convenience, Table A.1 lists some values for the spectrum (A.4) and its integrals.[55] f_N is the fraction of the number of photons above v, and f_E is the fraction of the energy carried by photons above v.

v	$dn/n\ dv$	f_N	f_E
0.01	8.50	0.7381	0.9979
0.03	3.91	0.6277	0.9912
0.1	1.562	0.4628	0.9502
0.2	0.863	0.3483	0.9052
0.3	0.584	0.2775	0.8485
0.5	0.333	0.1896	0.7369
1.	0.1244	0.08677	0.5000
2.	0.0288	0.02326	0.2150
3.	0.00818	0.00703	0.0886
5.	0.00081	0.000737	0.0142

Table A.1. *Photon energy spectrum v for electrons in circular motion, normalized to the critical energy. f_N is the number of photons above v and f_E is the fraction of energy carried by photons above v*

The energy loss per electron has average value

$$\langle U \rangle = \langle n \rangle \langle k \rangle \tag{A.5}$$

$$= 1.267\ (\text{keV}) \left(\frac{E}{10\ \text{GeV}}\right)^2 \left(\frac{B}{1\ \text{kG}}\right)^2 L(\text{m}) \tag{A.5a}$$

$$= 140.8\ (\text{keV}) \left(\frac{E}{10\ \text{GeV}}\right)^4 \left(\frac{1\ \text{km}}{\rho}\right) \phi(\text{rad}) \tag{A.5b}$$

and variance

$$\text{var}(U) = \langle (U - \langle U \rangle)^2 \rangle = \langle n \rangle \langle k^2 \rangle \tag{A.6}$$

$$= 11.17 (\text{keV}^2) \left(\frac{E}{10\ \text{GeV}}\right)^4 \left(\frac{B}{1\ \text{kG}}\right)^3 L(\text{m}) \tag{A.6a}$$

$$= 414\ (\text{keV}^2) \left(\frac{E}{10\ \text{GeV}}\right)^7 \left(\frac{1\ \text{km}}{\rho}\right)^2 \phi(\text{rad}) \tag{A.6b}$$

where $\langle v \rangle = 0.3079$ and $\langle v^2 \rangle = 0.4074$ have been used.[53] Equation $(A.6)$ arises because $\mathrm{var}(U)$ depends on fluctuations in n as well as in k; a Poisson distribution for n has been used. The angular distribution of the radiated energy integrated over all k is

$$\frac{1}{U}\frac{dU}{dw} = \frac{21}{32}\frac{1+(12/7)w^2}{(1+w^2)^{7/2}}\ , \qquad -\infty \leq w \leq \infty\ , \quad (A.7)$$

$$w = \gamma\psi\ , \qquad (A.7a)$$

where ψ is the angle perpendicular to the bend plane. For some cases of heating by bend SR, it is useful to have an approximate expression for the full double differential distribution. For fixed v, approximate the w dependence with a Gaussian. Then

$$\frac{1}{U}\frac{dU}{dv\,dw} \approx \frac{1}{U}\frac{dU}{dv}\cdot\frac{1}{\sqrt{2\pi}\sigma_w}\exp\left\{\frac{-w^2}{2\sigma_w^2}\right\}\ , \qquad (A.8)$$

$$\sigma_w \approx \left(\frac{0.363}{v}\right)^{0.44}\ . \qquad (A.9)$$

This approximation is reasonable at the 10–20% level for $0.1 < v < 3$. Outside this range, Eq. $(A.9)$ overestimates the effective σ.

11.2. Quadrupoles

A photon spectrum integrated over a quadrupole field may be derived from $(A.4)$.[56] This spectrum is not very useful for background calculations because only the spectrum hitting the mask is interesting, not the spectrum going down the beampipe. The moments of a quad spectrum may be interesting for power reasons. For a Gaussian beam, the field scale is B_σ, the quad field at 1 σ of the beam. The number of radiated photons and the radiated energy in units of the values for a bend magnet with B_σ are, with b the beam centroid offset in σ units:

Gaussian Beam	Photon Number $\langle n\rangle/n_\sigma$	Radiated Energy U/U_σ
1-D	$\sqrt{(2/\pi)}+b$	$1+b^2$
2-D (round)	$\sqrt{(\pi/2)}$ ($b=0$ only)	$2+b^2$

Table A.2. *Number of radiated photons and the radiated energy in units of the values for a bend magnet with B_σ, with b the beam centroid offset in σ units.*

11.3. Random Sampling from the Synchrotron Radiation Distribution

The nicest routine I know for random sampling from the spectrum Eq. (A.4) is due to Yokoya.[57] RN is a random number uniform between 0 and 1 representing the integral number distribution between v and infinity. The returned values of v are within 0.05% of Mack's values,[1] at least for $0.001 < v < 12$.

```
DATA YA1/0.5352 /, YA2/0.3053 /, YA3/0.1418 /, YA4/0.4184 /,
% YB0/0.01192 /, YB1/0.2065 /, YB2/-0.3281 /,
% YC0/0.003314 /, YC1/0.1927 /, YC2/0.8877 /,
% YD0/148.3 /, YD1/675.0 /, YE0/-692.2 /, YE1/-225.5 /

  IF(RN.GT..342)THEN
P1=1.0-RN
P2=P1*P1
V=(((YA4*P2+YA3)*P2+YA2)*P2+YA1)*P2*P1
  ELSEIF(RN.GT.0.0297)THEN
V=((YB2*RN+YB1)*RN+YB0)/(((RN+YC2)*RN+YC1)*RN+YC0)
  ELSE
T1=-LOG(RN)
V=T1+(YD1*T1+YD0)/((T1+YE1)*T1+YE0)
```

11.4. Short-Bend Radiation

A magnet in which the bend angle is less than $1/\gamma$ is called a short magnet; the spectrum does not follow Eq. $(A.4)$ and depends on the z variation of B. The amount of energy is roughly the same as given in Section A.2 above, but the scale or characteristic energy is greater,[54] where with $k_{c-\text{long}}$ given by Eq. $(A.2)$,

$$k_{c-\text{short}} \approx k_{c-\text{long}} \frac{\frac{2}{\gamma}}{\phi} . \tag{A.10}$$

11.5. General SR Spectrum

The spectrum in Section A.1 is for $k_c \ll E$. The general case is[58,59,60]

$$\frac{1}{n}\frac{dn}{dy} = \frac{3}{5\pi}\frac{1}{(1+\xi y)^2} \times \left[\int_y^{\infty} K_{5/3}(x)\,dx + \frac{\xi^2 y^2}{1+\xi y} K_{2/3}(y)\right] , \tag{A.11}$$

$$y = \frac{\frac{k}{k_c}}{1-\frac{k}{E}} = \frac{v}{1-u} = v(1+\xi y) , \tag{A.12}$$

$$\xi = \frac{k_c}{E} = \frac{3}{2}\Upsilon , \tag{A.13}$$

where n is given by $(A.1)$, and v only enters in the combination y. For $\xi = 0$, $(A.11)$ reduces to $(A.4)$. Υ is frequently used as a measure of k_c/E rather than ξ.[59,60] For Monte Carlo sampling of $(A.11)$ see Reference 57.

Appendix B
NOTATION

a	Radius
b	Radius
c	Velocity of light
f_b	Bunch crossing frequency, $f_b = c/s_b$
k	HOM loss parameter, usually pV/C
k	Photon energy, usually keV
k'	Scattered photon energy
k_c	Characteristic energy of synchrotron radiation
m	Electron mass (energy or momentum, i.e., missing factors of c)
n	Number of radiated SR photons
p	Momentum
q	Bunch charge
q^2	(Momentum transfer)^2
r_e	Classical radius of electron, 2.82×10^{-13} cm
s	Path length along orbit
s_b	Bunch spacing
u	Energy loss normalized to E, e.g., k/E
v	Normalized photon energy, k/k_c
w	Angle nomalized to $1/\gamma$, $w = \gamma \times$ angle
z	Distance along beam axis
A	Atomic weight
B	Magnetic field, usually kG
E	Beam energy
I	Beam current
L	Magnet length, usually m
SR	Synchrotron radiation
T	Kinetic energy
U	Energy radiated
Z	Atomic number
Z_o	377 ohms
α	1/137
γ	E/m
δ	Skin depth, usually μm
μ	RF magnetic permeability relative to vacuum
ϕ	Bend angle
σ	Cross section
σ_z	Rms bunch length, Gaussian parameter
$\sigma(Z)$	DC electrical conductivity of material of atomic number Z
θ	Scattering angle
ρ	Radius of curvature

References

1. A. R. Clark (LBL) wrote QSRAD in the early 1970s for use with PEP. Many versions exist now. There is no write-up, as far as I know.
2. W. R. Nelson, H. Hirayama, and D. W. O. Rogers, *The EGS4 Code System,* SLAC Report–265, December 1985. The only approximations in EGS4 that I know of that might affect B Factory calculations are no L-shell fluorescence radiation, no elastic nuclear form factors, and no quasi-elastic or inelastic nucleon/nuclear scattering of any kind.
3. B. Rossi, *High-Energy Particles*, Prentice-Hall, 1952. These formulas may be found in many other places.
4. H. Bethe and J. Ashkin in *Experimental Nuclear Physics,* Vol. 1, ed. E. Segre (Wiley, NY, 1953).
5. H. W. Koch and J. W. Motz, Bremsstrahlung Cross-Section Formulas and Related Data, *Rev. Mod. Phys.* **31**, 920 (1959).
6. D. C. Carey, K. L. Brown, and Ch. Iselin, *DECAY TURTLE,* SLAC–246/UC–28/Fermilab PM–31, March 1982. For beam-gas interactions, we use a version of DECAY TURTLE as modified by W. Kozanecki at SLAC.
7. W. Kozanecki (SLAC), private communication, March 1990.
8. Y.-S. Tsai, Pair Production and Bremsstrahlung of Charged Leptons, *Rev. Mod. Phys.* **46**, 825 (1974); Errata, **49**, 421 (1977).
9. Y. Baconnier, Neutralization of Accelerator Beams by Ionization of the Residual Gas, CERN 85-15, CERN Accelerator School, Gif-sur-Yvette, 1984.
10. C.J. Bocchetta and A. Wrulich, The Trappping and Clearing of Ions in Elletra, *Nucl. Inst. Meth.* **A278**, 807 (1989).
11. V.I. Telnov, Scattering of Electrons on Thermal Radiation Photons in e^+e^- Storage Rings, *Nucl. Inst. Meth.* **A260**, 304 (1987).
12. B. Dehning, *et al.*, Scattering of High Energy Photons off Thermal Photons, *Phys. Lett.* **2349B**, 145 (1990).

13. L. W. Jones and T. O. Dershem, p. 183, *12th Int. Conf. on High-Energy Accelerators*, Fermi National Accelerator Laboratory, 1983.

14. J. W. Lightbody, Jr. and J. S. O'Connel, Modeling single-arm electron scattering and nucleon production from nuclei by GeV electrons, *Computers in Physics*, May/June 1988, p. 57.

15. K. V. Alanakyan *et al.*, On the Angular Dependence of Photoprotons from Nuclei Irradiated with γ-Quanta with Maximum Energy 4.5 GeV, *Nucl. Phys.* **A367**, 429 (1981).

16. K. W. Chen *et al.*, Electroproduction of Protons at 1 and 4 BeV, *Phys. Rev.* **135**, B1030 (1964). These authors deduced an equivalent radiator of about 0.025 radiation lengths.

17. R. D. Evans, *Encyclopedia of Physics*, Vol. XXXIV (Springer-Verlag, Berlin, 1958).

18. A. T. Nelms and I. Oppenheim, Data on the Atomic Form factor: Computation and Survey, *J. Res. Nat. Bureau of Standards* **55**, No. 1, July 1955, pp. 53–62.

19. I. F. Ginzburg, G. L. Kotkin, V. G. Serbo, and V. I. Telnov, Colliding γe and $\gamma\gamma$ Beams Based on the Single-Pass ee Colliders (VLEPP Type), *Nucl. Inst. Meth.* **205**, 47 (1983).

20. J. W. Motz, H. A. Olsen, and H. W. Koch, Pair Production by Photons, *Rev. Mod. Phys.* **41**, 581 (1969). This gives the γe cross section.

21. S. A. Heifets, Photon Beam at CEBAF, CEBAF TN/90/202, January 1990.

22. J. Kirkby (CERN), private communication, April 1990.

23. J. T. Seeman, Observation of the Beam-Beam Interaction, Nonlinear Dynamics Aspects of Particle Accelerators, *Lecture Notes in Physics*, Vol. 247 (Springer-Verlag, Berlin, 1986).

24. YongHo Chin (LBL), private communication, April 1990. The simulation used a code due to K. Yokoya (KEK) and applies to a *PEP*-like machine as described in Feasibility Study for an Asymmetric *B* Factory Based on *PEP*, LBL PUB-5244/SLAC-352/CALT-68-1589, October 1989.

25. M. Sands, The Physics of Electron Storage Rings: An Introduction, SLAC–121, November 1970.

26. G. Decker and R. Talman, Measurement of the Transverse Particle Distribution in the Presence of the Beam-Beam Interaction, *IEEE Trans. Nuc. Sci.*, **NS–30** (4), 2188 (1983).

27. M. K. Sullivan, Synchrotron Radiation Background in the MAC and TPC Vertex Chambers, Univ. Calif. Intercampus Inst. for Research at Particle Accelerators, c/o SLAC, UC–IIRPA–88–01, May 1988.

28. See Feasibility Study for an Asymmetric *B* Factory Based on *PEP*, LBL PUB–5244/SLAC–352/CALT–68–1589, October 1989.

29. G. von Holtey and D. Ritson, Mini Beam Pipe at the DELPHI Interaction Region, CERN LEP NOTE 614, October 1988.

30. P. Morton and P. Wilson, Energy Loss and Wall Heating for a Gaussian Bunch in a Cylindrical Pipe, SLAC–AATF/79/15, November 1979.

31. A compendium of computer codes useful in accelerator physics is given in *AIP Conf. Proc. 184*, Vol. 2, NY, 1989.

32. J. J. Bisognano, S. A. Heifets, and B. C. Yunn, Loss Parameters for Very Short Pulses, CEBAF–PR–88–005, July 1988.

33. P. B. Wilson, Introduction to Wakefields and Wake Potentials, Physics of Particle Accelerators (Fermilab, 1987; Cornell, 1988), *AIP Conf. Proc. 184,* Vol. 1, NY, 1989.

34. S.A. Heifetz and S.A. Keifets, Coupling Impedance in Modern Accelerators, SLAC-PUB-5297, September 1990. Extensive summary of mainly analytic results. These authors measure k in cm^{-1}. To convert to V/pC multiply by $Z_0c/4\pi = 1/4\pi\epsilon_0 = 1/1.11V\text{cm}/pC$.

35. P. B. Wilson, High Energy Linacs: Applications to Storage Ring RF Systems and Linear Colliders, Physics of Particle Accelerators (Fermilab 1981), *AIP Conf. Proc. 87,* NY, 1982.

36. J. D. Jackson, *Classical Electrodynamics*, 2nd ed. (Wiley, 1975). Note: Jackson's k_c is two times larger than Schwinger's.

37. A. F. Harvey, *Microwave Engineering* (Academic Press, NY, 1963).

38. At 10 GHz, type 430 SS has μ about 10. J. Haimson (Haimson Research Corp.), private communication, April 1990.

39. M. Billing, Power Dissipation in the CLEO II Beam Pipe from Beam Heating, unpublished CESR internal note CON 87-6, May 1987; Issues Affecting Single Bunch Current Limits, *B* Factory Workshop Proc., Syracuse University, September 1989.

40. J. Kent, *Procedings of the Workshop on High Luminosity Asymmetric Storage Rings for B Physics*, Report CALT-68-1552, Caltech, April 1989.

41. See articles in Proc. of 5th European Sym. on Semiconductor Detectors, *Nucl. Inst. Meth.* **A288** (1), 1-292 (1990).

42. J. G. Jernigan *et al.*, Performance Measurement of Hybrid PIN Diode Arrays, SLAC-PUB-5211, May 1990; Intl. Industrial Sym. on SSC, Miami Beach, March 1990.

43. S. L. Shapiro *et al.*, Progress Report on Use of Hybrid Silicon PIN Diode Arrays in High Energy Physics, SLAC-PUB-5212, May 1990; *Vth Intl. Conf. on Instrumentation for Colliding Beam Physics*, Novosibirsk, March 1990.

44. J. V. Allaby *et al.*, The MAC Detector, *Nucl. Inst. Meth.* **A281**, 291 (1989).

45. J. Kadyk *et al.*, Anode Wire Aging Tests with Selected Gases, *IEEE Trans. Nuc. Sci.* **37** (2), 478 (1990).

46. J. Va'vra, Aging in Gaseous Detectors, SLAC-PUB-5207, March 1990; *Vth Intl. Conf. on Instrumentation for Colliding Beam Physics*, Novosibirsk, March 1990.

47. H. B. Luna *et al.*, Bremsstrahlung Radiation Effects on Rare Earth Permanent Magnets, *Nucl. Inst. Meth.* **A285**, 349 (1989). (Note typo in Table 3: for 10^8 cm^{-2}, read 10^{18}.)

48. S. Dushman, *Scientific Foundations of Vacuum Technique,* 2nd edition, eds. S. Dushman and J. M. Lafferty (Wiley, NY, 1962).

49. N. B. Mistry, Ultrahigh Vacuum Systems for Storage Rings and Accelerators, Physics of Particle Accelerators (SLAC 1985), *AIP Conf. Proc. 153,* Vol. 2, NY, 1987.

50. O. Groebner *et al.*, Neutral Gas Desorption and Photoelectric Emission from Aluminum Alloy Vacuum Chambers Exposed to Synchrotron Radiation, *J. Vac. Sci. Technol.* **A7** (2), 223 (1989).
51. B. A. Trickett, The ESRF Vacuum System, *Vacuum* **39**, 607 (1988).
52. H. Hartwig and J.S. Koupsidis, A New Approach for Computing Sputter-Ion Pump Characteristics, *J. Vac. Sci. Technol.* 11, 1154 (1974).
53. J. Schwinger, *Phys. Rev.* **75**, 1912 (1949). Only a few of the numerous references dealing with SR are listed.
54. R. Chrien, A. Hofmann, and A. Molinari, *Phys. Reports* **64**, 249 (1980).
55. R. A. Mack, Cambridge Electron Accelerator Report CEAL–1027, February 1966. Although it is not especially difficult to evaluate these SR formulas using modern computation facilities, Mack's report was quite an accomplishment in 1966.
56. E. Keil, Synchrotron Radiation from a Large Electron-Positron Storage Ring, CERN/ISR–LTD/76–23, June 1976.
57. K. Yokoya, A Computer Simulation Code for the Beam-Beam Interaction in Linear Colliders, KEK–Report 85–9, October 1985.
58. K. Yokoya and P. Chen, Electron Energy Spectrum and Maximum Disruption Angle under Multi-Photon Beamstrahlung, SLAC–PUB–4935, March 1989, *IEEE Particle Accelerator Conf.*, Chicago, March 1989.
59. A. A. Sokolov and I. M. Ternov, *Synchrotron Radiation* (Pergamon Press, NY, 1968); *Radiation from Relativistic Electrons* (American Inst. Phys., NY, 1986).
60. T. Erber, High-Energy Conversion Processes in Intense Magnetic Fields, *Rev. Mod. Phys.* **38**, 626 (1966). The spectrum here uses a different, but equivalent, combination of Bessel functions in Eq. ($A.11$).

Detector Backgrounds from Synchrotron Radiation

M. Sullivan

1. Introduction

The primary goal of any synchrotron radiation masking scheme is to prevent radiation from directly hitting the detector beam pipe. It is equally important to protect the beam pipe from secondary radiation. There are several sources of secondary radiation:

1) Photons which forward-scatter from the tip of a mask onto the detector beam pipe (tip scattering).
2) Photons which scatter from a surface upstream of the Interaction Region (IP) and onto the detector beam pipe (forward scattering).
3) Photons which backscatter from masks and surfaces downstream of the IP and have a chance of striking the detector beam pipe (back scattering).

Forward scattering (Case 2) can usually be avoided by ensuring that all upstream surfaces which intercept synchrotron radiation have no solid angle for the scattered photons to reach the detector beam pipe. The remaining two problems (tip scattering and back scattering) are generally studied together, since a mask which generates tip scattering in one direction will usually generate back scattering in the other direction.

2. Overview

In the SLAC *APIARY* 6 design, both the incoming and outgoing beams have off-axis trajectories in the focusing quadrupoles near the IP. This is in contrast to the Cornell design, which has off-axis trajectories for only the outgoing beams. In particular, for *APIARY* 6, the low-energy beam

(LEB) is off-axis in the Q1 and Q3 magnets and the high-energy beam (HEB) is off-axis in the Q2 magnets (see Figure 1). The geometry of the *APIARY* 6 design attempts to minimize the amount of off-axis induced or 'bend' radiation which strikes surfaces near the IP. The LEB bend radiation from Q1 and B1 passes through the IP, as does the HEB bend radiation coming from Q2 and B1. However, the bend radiation generated by the LEB off-axis trajectory in Q3 strikes the AB mask shown in Figure 1.

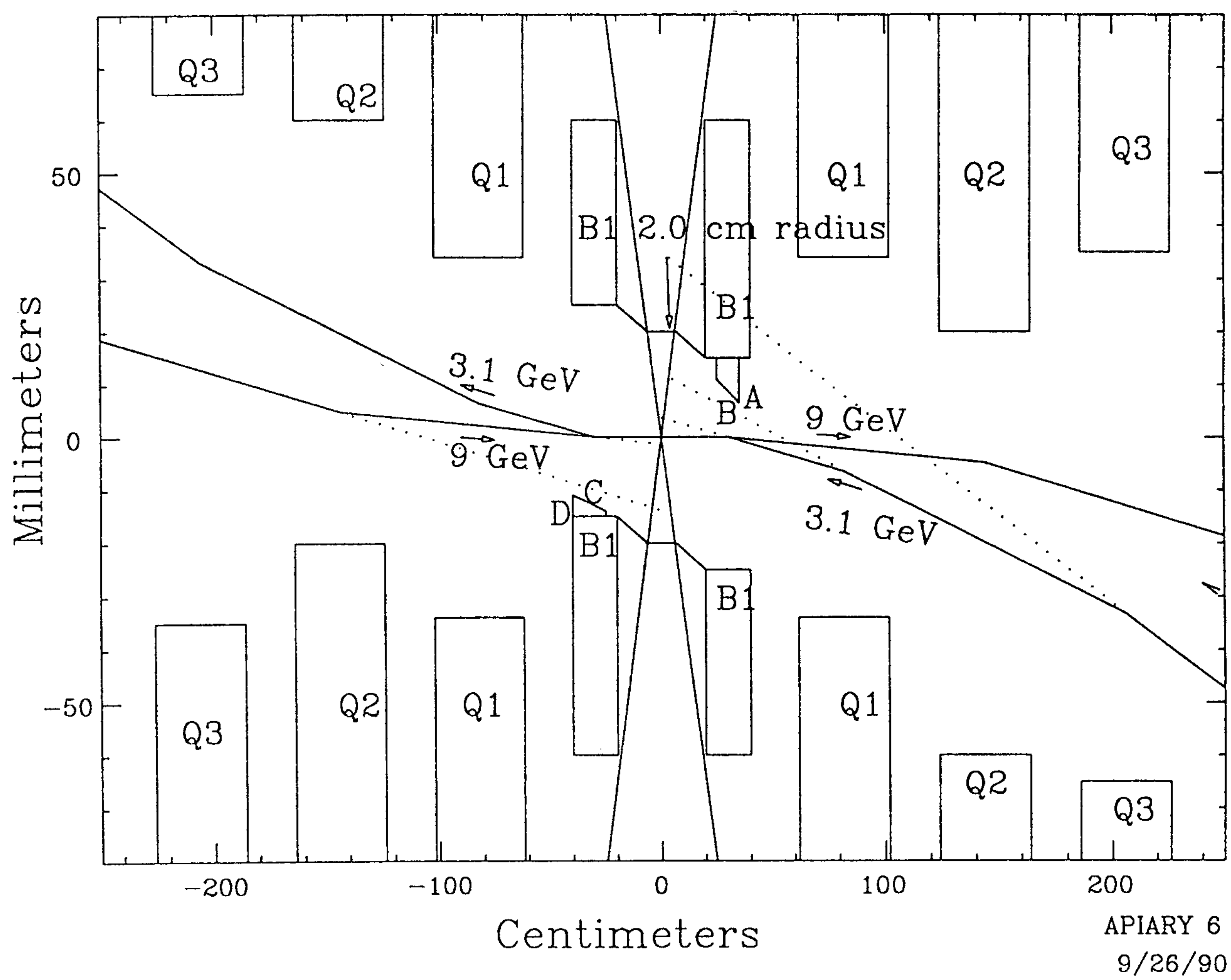


Figure 1. *Arrangement of magnets and synchrotron radiation masks in the APIARY 6 design.*

The intensity of radiation striking mask AB is mitigated by the fact that Q3 is the weakest of the three magnets (Q1-Q3). Mask CD in Figure 1 is required to prevent the HEB synchrotron radiation from directly striking the detector beam pipe, which is assumed in this example to have a 2 cm radius. The primary source of radiation striking this mask has its origin in Q4 and Q5, the last major focusing elements of the HEB.

Three meters from the IP, the beams are fully separated and heading into separate beam pipes. The surface formed by the intersection of the two beam pipes intercepts a significant portion of the generated synchrotron radiation. Cooling must be provided for this "crotch" region. The radiation from the HEB bend in Q2 strikes the surface on one side of the IP and the radiation from the LEB in Q1 strikes the surface on the other side of the IP.

3. Calculations

The synchrotron radiation calculations start with a program called MAGBENDS, which computes the beamline geometry and generates information on beam trajectories and displacements. This information is used as input to the main program SYNC_BKG, which is a descendent of QSRAD, a program originated by A. Clark of LBL to study synchrotron radiation at *PEP*. The program traces particle rays through a series of quadrupole magnets and generates fans of synchrotron radiation from the focussing and defocussing bends in the magnets. These fans of radiation are then tallied as a function of photon energy on the various user supplied masks and apertures. QSRAD uses a gaussian distribution for the beam profile.

The following improvements to QSRAD have been incorporated into SYNC_BKG:

1) Nongaussian beam tails. Beams develop nongaussian tails when they collide. There is as yet no definite understanding of what this tail distribution should look like. The population density of this beam tail can be important. The fact that the particles are out at large distances from the beam core generally gives them a better chance of directing photons straight onto the detector beampipe. Present background studies are based on nongaussian tails that simulate detector experience

at *PEP*. The *APIARY* 6 design has proven to be insensitive to the exact tail distribution. This implies that the bulk of the background synchrotron radiation in this design originates from the core of the beam.

2) Offset magnets and magnet segmentation. Magnets can be offset with respect to the beam axis. This improvement was made by S. Hertzbach in another version of QSRAD developed to study backgrounds in the SLC. In addition, the magnet elements can be broken up into sections, permitting an improved calculation of the synchrotron radiation fans.

3) Nonzero beam emittance. Most versions of QSRAD integrate the beam profile over only two dimensions, x and y. This is equivalent to assigning one slope to one beam position or collapsing the phase ellipse of the beam to a line. Incorporating finite emittance beams allows the program to integrate the beam distribution over four dimensions x, x', y, y', although this greatly increases computing time. The correction to the backgrounds from finite emittance is generally small but can become important when beam particle tracing involves four or more quadrupoles.

The energy spectrum produced by SYNC_BKG is used as input by an EGS (Electron Gamma Shower) interface program called MASKING. This program is used to calculate photon reflection, transmission, or absorption coefficients for various selections of masks, beam pipes and detector elements. The EGS package includes K-shell fluoresence and Rayleigh scattering cross-sections. At present, there is no provision in EGS for L-shell fluoresence, which can be important when elements with a high Z value are used. The low average critical energies of the synchrotron radiation (≤ 10 keV) make high-Z elements the materials of choice for masks. For the present study, the masks are assumed to be coated with gold.

Photons which back-scatter from a surface have a generally isotropic angular distribution. The solid angle fraction of the detector beam pipe as seen from the source of back-scattered radiation can thus be used to predict the number of

Photons/ Crossing	Incident on Be pipe	Absorbed in 1st Si layer	Absorbed in 2nd Si layer	Absorbed in 3rd Si layer	Incident on DC
		3.1 GeV beam			
$4 < E_\gamma < 100$ keV					
Number of photons	0.47	2.7×10^{-3}	1.2×10^{-5}	5.8×10^{-7}	8.6×10^{-9}
Energy (keV)	3.61	0.023	1.0×10^{-4}	1.5×10^{-5}	1.3×10^{-7}
$4 < E_\gamma < 20$ keV					
Number of photons	0.47	2.7×10^{-3}	1.2×10^{-5}	5.8×10^{-7}	8.6×10^{-9}
Energy (keV)	3.61	0.023	1.0×10^{-4}	1.5×10^{-5}	1.3×10^{-7}
		9.0 GeV beam			
$4 < E_\gamma < 100$ keV					
Number of photons	0.55	0.014	1.0×10^{-3}	1.0×10^{-3}	1.2×10^{-4}
Energy (keV)	8.41	0.32	0.02	0.02	6.3×10^{-3}
$4 < E_\gamma < 20$ keV					
Number of photons	0.43	8.3×10^{-3}	4.7×10^{-5}	4.8×10^{-6}	1.4×10^{-7}
Energy (keV)	4.37	0.071	4.6×10^{-4}	8.0×10^{-5}	2.6×10^{-6}
		Totals			
$4 < E_\gamma < 100$ keV					
Number of photons	1.02	0.017	1.0×10^{-3}	1.0×10^{-3}	1.2×10^{-4}
Energy (keV)	12.0	0.34	0.02	0.02	6.3×10^{-3}
$4 < E_\gamma < 20$ keV					
Number of photons	0.90	0.011	5.9×10^{-5}	5.4×10^{-6}	1.5×10^{-7}
Energy (keV)	7.98	0.094	5.6×10^{-4}	1.0×10^{-4}	2.7×10^{-6}

Table 1. *Synchrotron radiation background predictions (per beam crossing) for APIARY 6. Multiply by 2.38×10^8 to get photons/sec. The energy refers to the total energy of the indicated photons. The beam pipe materials are 25 μm Cu on 1 mm Be. The beam pipe radius is 2 cm with a 14 cm length for the Be section. The average angle of incidence is 100 mr for the radiation striking the pipe. The silicon layers are 300 μm thick.*

photons incident on the beam pipe from that source. However, when the source of radiation is the tip of a mask, some of the incident photons will forward-scatter through the tip. Inspection of angular distributions reveals the increase in photon rate from forward scattering to be a factor of 2 or 3 above an isotropic distribution. Present studies increase the number of tip-scattered photons by a factor of 2.5. Radiation hitting the beam pipe from a mask tip or other surface has a range of incident angles on the pipe. An average angle of incidence of 100 mr is chosen in order to calculate the transmission coefficient for the beam pipe. The detector beam pipe is a cylinder 14 cm long with a 2 cm radius. The pipe materials are 1 mm Be with a 25 μm inside layer of Cu.

The *APIARY* 6 background estimates for synchrotron radiation are shown in Table 1.

Figures 2 and 3 display the energy spectra of the photons that scatter from the masks, are incident on the detector beam pipe and transmit through the beam pipe and silcon detector layers. Figure 2 shows the photon spectra from the LEB synchrotron radiation and Figure 3 displays the spectra from the HEB radiation.

Investigation of the synchrotron radiation which strikes the beam pipe at 3 meters has just begun. There is no detector beam pipe solid angle from the 3 meter surfaces on either side of the IP. However, the intensity of the radiation striking these surfaces is very high, which raises the possibility of back-scattered photons bouncing off surfaces near the detector beam pipe and then directly striking the pipe (double bouncing). In addition, the high power levels produce significant heating at these surfaces. Present calculations estimate the maximum power density on the surface which intercepts the HEB radiation at about 200 watts/mm^2. These issues must be more thoroughly investigated.

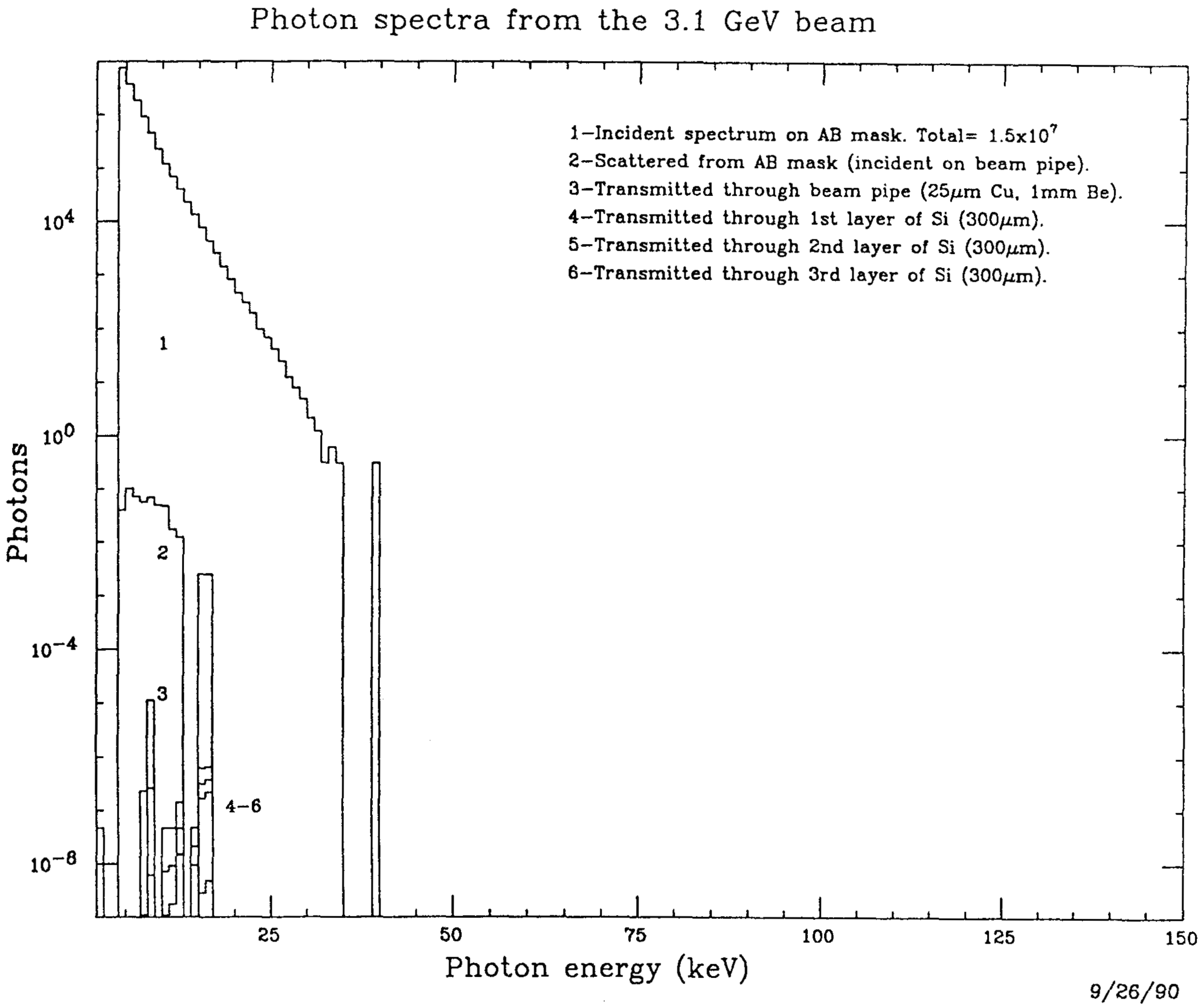


Figure 2. *Photon energy spectra (photons/keV) from the low-energy beam.*

For all present B Factory designs under study, the total amount of synchrotron radiation power generated near the IP is significant; all of this power must be accounted for in any design of absorbing masks and vacuum beam pipes. These high power levels imply high photon fluxes and the scattering and backscattering of photons from various surfaces has to be carefully traced in order to insure an adequate understanding of synchrotron radiation backgrounds.

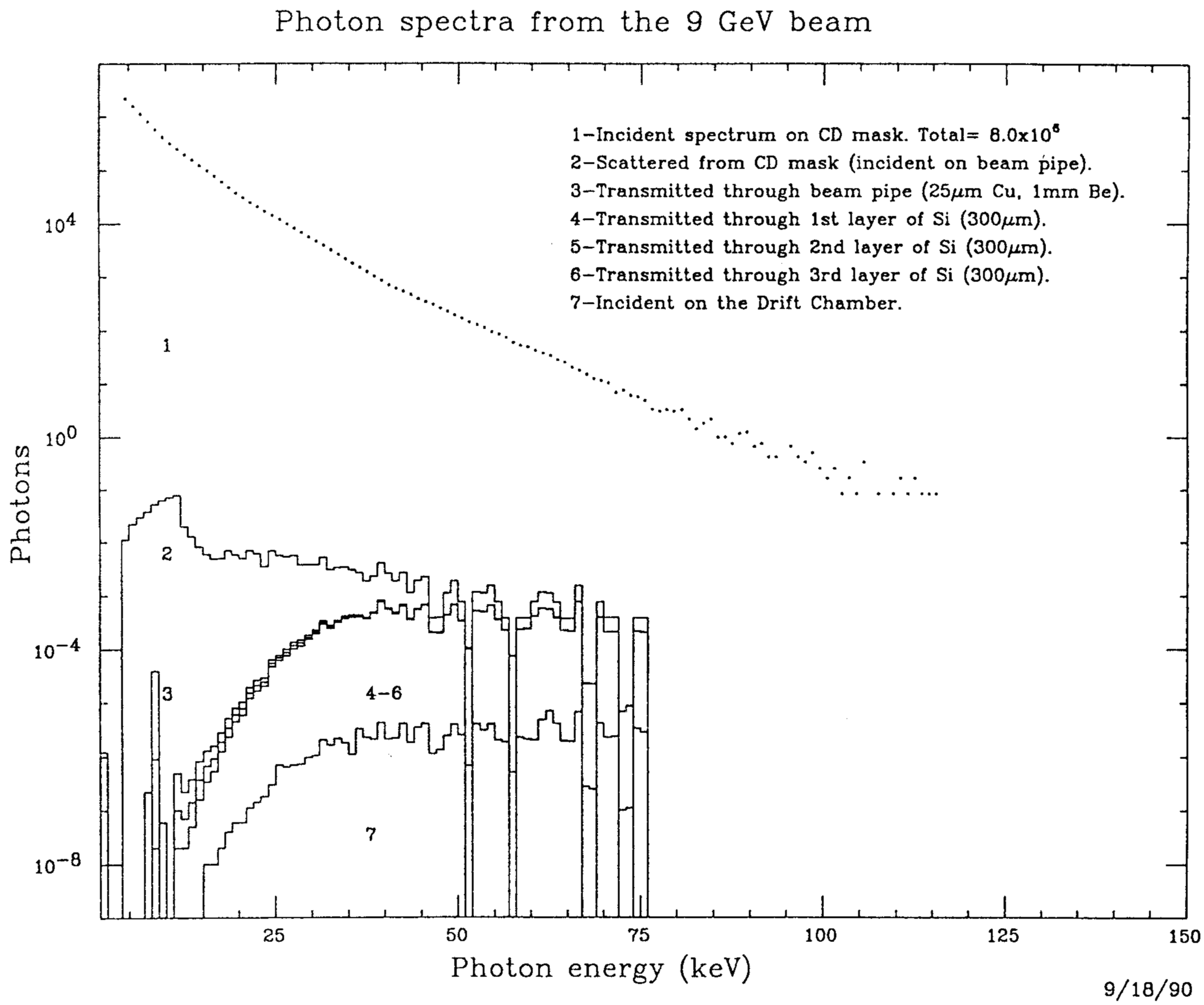


Figure 3. *Photon energy spectra (photons/keV) from the high energy beam*

Detector Backgrounds from Scattered Beam Particles

D. P. Coupal and C. Hearty

1. Introduction

The B Factory design calls for colliding e^+e^- beams at asymmetric energies of 9 and 3.1 GeV at a minimum luminosity of $3 \times 10^{33} \text{cm}^{-2}\text{s}^{-1}$. The design uses the *PEP* tunnel for both the high and low-energy beams. Beam currents in the range 1–2 amperes are required to achieve this high luminosity. A major concern with these high currents is the detector backgrounds caused by beam particles scattering from residual gas in the vacuum pipe. Coulomb scattering and bremsstrahlung create off-axis and off-energy beam particles and, in the case of bremsstrahlung, photons, which can hit the beampipe or synchrotron radiation masks near the interaction point (IP). The resulting electromagnetic showers contribute background in the detector. This report presents some preliminary results on a simulation of this background in the SLAC B Factory design.

As of this writing, the design is still in flux; this report is thus a review of work in progress. We present results for the high-energy beam of the *APIARY* 6.0 design, with partial modelling of the lattice around the interaction region.

The next section describes the tools and method of modelling the beam-gas scattering backgrounds. The B Factory lattice used in this study is described in the following section. We then present the results for particle rates into the masks and beampipe around the IP for the high-energy beam. The next section uses these results to predict the rates in the different elements of a generic detector. The final section concludes with a list of the limitations of this study and a discussion of open issues.

2. Tools

In addition to the high currents, a high luminosity B Factory will require relatively strong focussing at the interaction point, resulting in a large flux of synchrotron radiation generated in the quadrupoles near the IP. The first task with any given lattice is to design masks to shield the beampipe around the IP from this flux of photons. Given the interaction region (IR) magnets, beampipe and masks, we evaluate the rate of scattered particles into these apertures. This study is done using DECAY TURTLE, [1] a ray-tracing program based on TRANSPORT, [2] modified by W. Kozanecki to model Coulomb scattering and bremsstrahlung from gas molecules along a beamline. This program requires a lattice and apertures specified in TRANSPORT/DECAY TURTLE format, the number of rays that should be scattered and traced through the lattice, the assumed pressure and composition of the residual gas in the beampipe, and the number of particles per bunch. It returns the number of rays that strike each of the apertures and a weighting factor which is used to convert the number of traced rays into the number of particles per bunch which strike each aperture.

To evaluate the resulting background in a generic detector, the particles striking apertures around the IP are fed to an EGS simulation (see Section 5).

3. Lattices

The beams must be in separate rings due to the asymmetric energies. Two schemes are being considered to bring the beams together at the IP and separate them again after the collision. The first method — the basis of the *APIARY* designs [3] — uses a bend magnet 20 cm from the IP, which bends the different energy beams by different amounts. The second method, called *crab crossing*, collides the beams at an angle, using transverse cavities before and after the IP to rotate the bunches so that they effectively collide head-on. A more detailed description of the crossing geometries can be found in reference 3. The crab crossing geometry appears to have lower backgrounds from synchrotron radiation and beam-gas scattering, but since this idea is untested, the SLAC design has concentrated on head-on collision geometry; the results presented here are for head-on geometry only. Table 1 shows some of the parameters for the different lattices being considered for the *B* Factory design.

Parameter			*APIARY* 5.1	*APIARY* 6.0	CRAB (Oide)	CRAB (Raubenheimer)
Beam current	3.1 GeV		2.2	2.2	3.0	2.2
(amps)	9.0 GeV		1.5	1.5	3.0	1.5
Number of bunches			1746	1746	1746	1746
Bunch separation (m)			1.27	1.27	1.27	1.27
β^* (cm)	3.1	x	60	37.5	50	37.5
		y	1.5	1.5	1	1.5
	9.0	x	120	75.0	50	75.0
		y	3.0	3.0	1	3.0
Emittance (nm-rad)	3.1	x	98	92.0	39	92.0
		y	2.5	3.6	2	3.6
	9.0	x	49	46.0	39	46.0
		y	1.2	1.8	2	1.8
σ_x:σ_y			40:1	25:1	50:1	25:1
Tune shift			.03	.03	.03	.03
Luminosity			3×10^{33}	3×10^{33}	10^{34}	3×10^{33}
B1	Strength	(T)	.5	.64	–	–
	Length	(cm)	20	20	–	–
	z (IP end)	(cm)	20	20	–	–
Q1	Strength	(K)	-5.57	-5.46	-2.63	-2.11
	Length	(cm)	40	40	30	80
	z (IP end)	(cm)	62	62	160	50

Table 1. *B Factory machine parameters.*

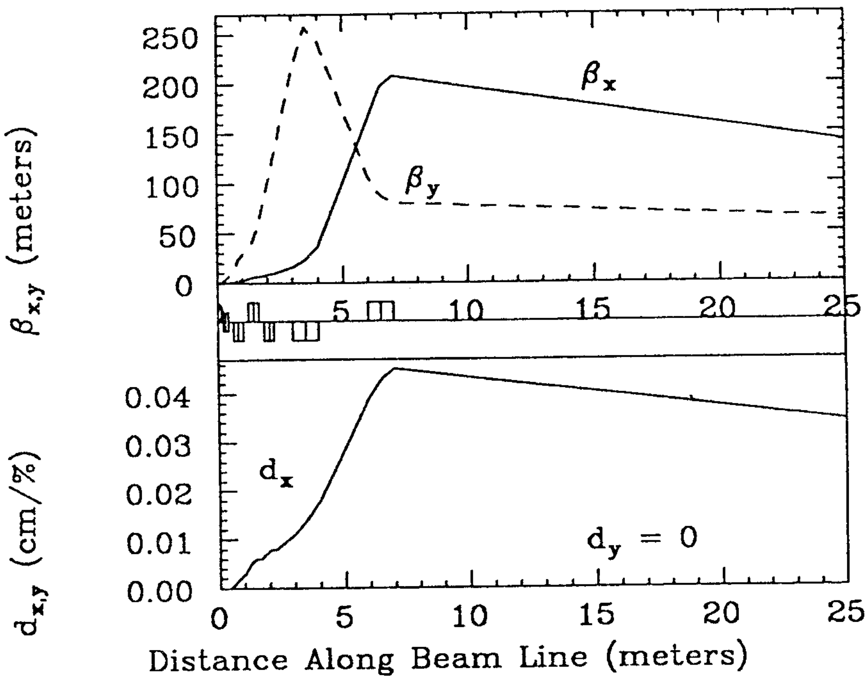


Figure 1. *The β functions and dispersion near the IP for the 9.0 GeV beam in the APIARY 6.0 B Factory design.*

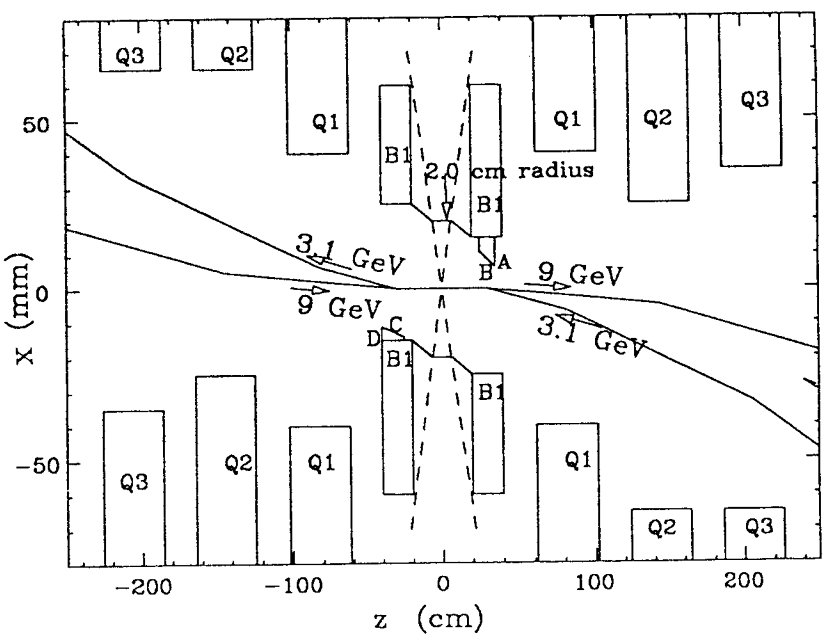


Figure 2. *An expanded view of the IP region for the APIARY 6.1 high energy beam. The mask CD is positioned to shield the IP beampipe from SR radiation from the high-energy beam, while AB is the low-energy ring SR mask.*

The lattice used in this study is the high-energy ring of *APIARY* 6.0 — the low-energy lattice did not exist at the time of the study. The rates for the low energy beam tend to be lower than the high-energy beam for reasons that will be explained below.

Figure 1 shows the beta and dispersion functions for the high-energy ring of *APIARY* 6.0 out to 25 meters upstream of the IP. The arc starts at 60 meters from the IP. In this design there are no vertical bends in the high-energy beam. The region close to the IP is shown in Figure 2. Both incoming and outgoing beams go through the three quadrupoles in the region within 3 meters of the IP and the bend magnet 20 cm from the IP. Quadrupoles Q1 and Q3 are centered on the high-energy beam; Q2 is centered on the low-energy beam.

4. Scattered Particle Rates

We generate Coulomb scattering and bremsstrahlung events in the region from the IP to 60 meters upstream of the IP. The vacuum pressure is taken to be 10^{-8} Torr. The scattered rays are propagated through the lattice until they either strike an aperture or reach the end of the defined lattice.

Figure 3a shows an x-y plot of beam particles Coulomb-scattered in the range 0.2 to 50.0 mrad at the upstream end of the SR mask in B1. Figure 3b shows the same plot but for bremsstrahlung with an energy loss in the range $0.005 < \delta E/E < 0.90$. Table 2 shows the rate of particles hitting the various apertures around the IP. The results are presented as a rate per μs, since many detector components integrate over this kind of time interval.

Scattered beam particles hitting the beampipe very close to the IP, which are potentially serious, come entirely from bremsstrahlung. Figure 4 shows the energy loss, $\Delta E/E$, and the z position of the scatter for rays hitting within 10 cm of the IP. As there are no bend magnets between Q2 and the exit of the arc 60 meters upstream of the IP, the beam particles which lose energy through bremsstrahlung — together with the resulting photons — continue along the beam direction until they hit the off-axis quadrupole Q2 and B1. The beam particles which have lost energy through bremsstrahlung are overbent by Q2 and B1 and hit the downstream mask or the IP beampipe. This tail of overbent particles is seen clearly in Figure 3b. Although most of the photons pass through the IP, a significant fraction still hits the upstream mask.

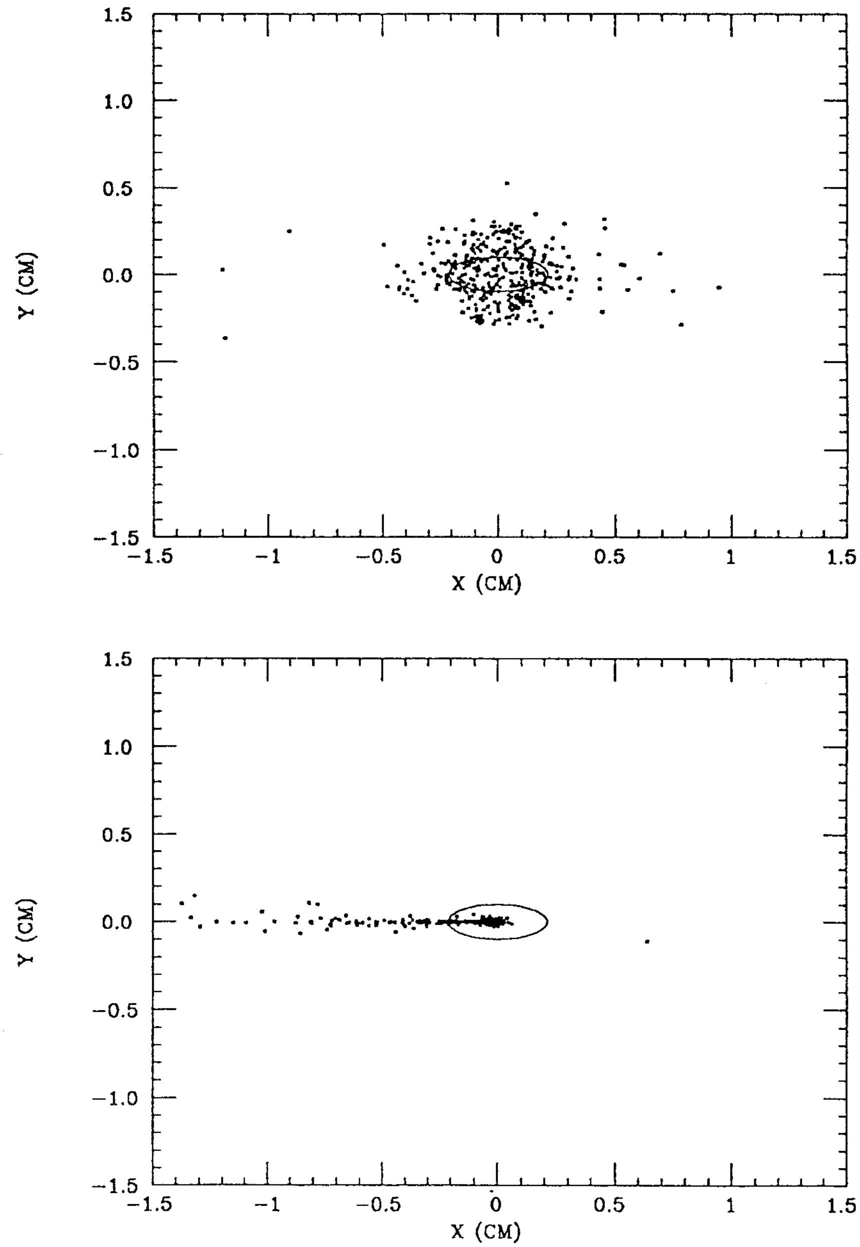


Figure 3. *Location of* **a.** *Coulomb-scattered and* **b.** *bremsstrahlung electron rays at the entrance to B1. The solid line shows the beam 10σ points.*

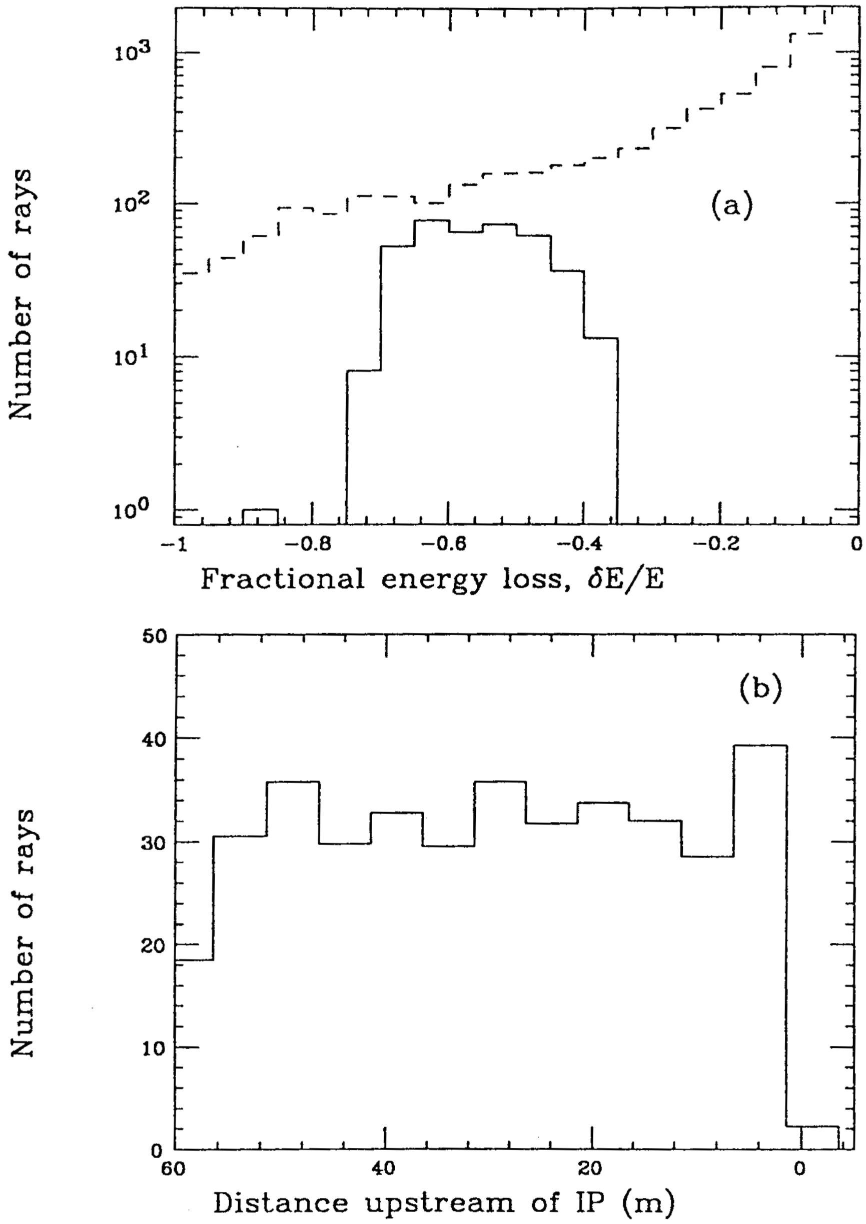


Figure 4 a. *Fractional energy loss $\Delta E/E$ and* **b.** *distance from IP of the production point of bremsstrahlung electrons which hit the IP beampipe. The dashed curve in* **a** *is the generated distribution.*

One way to reduce this background is to interrupt this long straight section with bend magnets, thereby creating a high-dispersion point at which the low-energy particles can be removed. This is the case in the *APIARY* designs for the low-energy beam; vertical dipole magnets 4 meters from the IP bend the low-energy beam up out of the plane of the high energy beam. The addition of bend magnets to the high-energy beam is being studied; this can potentially reduce the rate by as much as a factor of five. Scatters which occur between these bend magnets and the IP are still a problem. Pressures of 10^{-9} Torr in the IR may be feasible; this would reduce the background by another factor of ten. The vacuum pump capacity required depends on the amount of photodesorption caused by synchrotron radiation.

Aperture	z (cm)	Coulomb $e^{\pm}/\mu s$	bremsstrahlung $e^{\pm}/\mu s$	bremsstrahlung $\gamma/\mu s$	Total $e^{\pm}/\mu s$
Q3	-226 – -186	0.0	5.7	0.0	5.7
Q2	-164 – -124	0.0	1.0	0.0	1.0
Q1	-102 – -62	0.0	2.7	0.1	2.7
B1 Mask (upstr.)	-40 – -25	0.1	0.0	42.5	0.1
IP beampipe (2 cm rad)	-7.5 – 7.5	0.0	6.8	0.0	6.8
B1 Mask (downstr.)	25 – 40	0.7	19.2	0.0	19.9

Table 2. *Rates for scattered beam particles hitting apertures near the IP for the high-energy beam of APIARY 6.0. A vacuum of 10^{-8} Torr is assumed. Entries of 0.0 indicate that the rate is less than 0.05 per μs.*

5. Backgrounds in the Detector

The effects of the lost particles are calculated using EGS, run at double precision (to eliminate problems with particles stepping repeatedly across boundaries) with 10 kev cutoffs on photon energies. The magnetic fields of the detector solenoid, B1 and Q1–Q3 are included. The geometry is approximated entirely by cylinders parallel to and centered on the z axis. Thus, every object — detector component, shielding, or magnet — is defined by its inner and outer radius, its extent in z and its material. The objects included for the background calculations for the *APIARY* 6.0 high-energy ring are listed in Table 3. Several objects are not simulated by the correct materials; for example, the silicon strips are made of aluminum, the CRID is vacuum, and the CsI is lead. The latter approximation is incorrect to the extent that the leakage from CsI is different from that of lead. This model includes a substantial amount of beampipe and magnet shielding (the magnets themselves are approximately the density of nickel and are not particularly good shielding). The *APIARY* 6.01 IR design includes longer magnets and an 18 cm radius support tube, which are inconsistent with some aspects of this design.

The DECAY TURTLE output for this calculation is a file of approximately 9100 rays, each characterized by a weight equal to its probability of occurrance per beam crossing. Rays are randomly selected to be propagated through the detector on the basis of this weight. A typical job, simulating 35 μs of running time ($\sim$ 2800 rays), requires 6200 CPU seconds on the SLAC IBM 3090.

The actual planned layout of the IR, and the geometry used by TRANSPORT, is not cylindrically symmetric about the z axis. This is particularly true of the masks. The TRANSPORT rays must, therefore, by mapped from the correct geometry to EGS geometry. This is done to preserve the incident polar angle of the particle with respect to the surface, the azimuthal angle of the point of impact, and the distance of the impact point in from the edge of the struck mask. Figure 5 is an example of this mapping. The calculated backgrounds from the two cases in Figure 5 can be different by a factor of two, but the typical error is less.

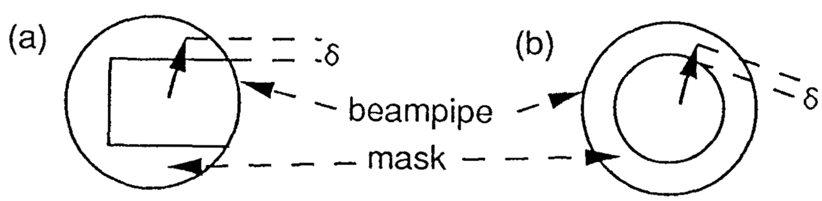


Figure 5. *A ray striking a mask in DECAY TURTLE* (**a**), *is represented in EGS by a ray striking a cylinder* (**b**). *The distance δ is the same in both cases.*

The EGS calculation finds the energy deposited in, and the number of photons and electrons entering each object. Electrons are counted each time they loop through an object. The results are summarized in Table 3. Blanks indicate that the object was not scored, not that there were no backgrounds. (For "DC first 2 cm", photons entering are not counted, but photons interacting are included as electrons.)

The background photons in the DC and calorimeter generally come from the B1 masks, where most of the rays strike. The background in the silicon vertex detector, which subtends a small solid angle from these masks, is caused primarily by

Object	Material	z (cm)	r (cm)	Backgrounds per μs		
				Energy (MeV)	γ	$e^{\pm}$
Q3	Sm_2Co_{17}	186 – 226	2.1 – 9.0	1936 155	–	–
Q2	Sm_2Co_{17}	124 – 164	2.1 – 9.0	767 473	–	–
Q1	Sm_2Co_{17}	62 – 102	2.1 – 9.0	4091 4351	–	–
B1	Sm_2Co_{17}	20 – 40	2.1 – 6.0	7251 30447	–	–
BP II	Al	20 – 900	2.0 – 2.1	1339 1152	6444 8609	726 539
EC Shield I	Lead	161 – 171	50. – 100.	106 3	–	–
EC Shld II	Lead	150 – 161	40. – 50.	66 362	–	–
Shield I	Lead	30 – 62	6.0 – 9.0	709 3237	–	–
Shield II	W	35 – 164	9.0 – 11.0	132 1086	–	–
Shield III	Lead	102 – 124	2.1 – 4.1	155 541	–	–
Shield IV	Lead	164 – 186	2.1 – 4.1	1410 152	–	–
Shield V	Lead	226 – 900	2.1 – 4.1	4757 277	–	–
Mask A/C Shld	W	40 – 50	2.1 – 6.0	8684 4299	–	–
Mask A	Ta	25 – 47	1.0 – 2.0	87794	–	–
Mask C	Ta	-47 – -25	1.3 – 2.0	40394	–	–
Beampipe	Be	-20 – 20	2.0 – 2.1	208	470	121
Endcap	Lead	150 – 161	50. – 100.	32 568	39 191	< 1 1
Si Layer 1	Al	-4.5 – 4.5	2.3 – 2.33	23	70	26
Si Layer 2		-4.5 – 4.5	4.5 – 4.53	10	105	16
Si Layer 3		-4.5 – 4.5	6.9 – 6.93	6	106	10
Forward TPC		4.5 – 20	2.3 – 10.0	0	556	64
DC Wall	Be	-150 – 150	18.0 – 18.1	28	–	–
DC Layer 1	He/CO_2/Iso	-150 – 150	18.1 – 20.1	1	–	8
DC		-150 – 150	20.1 – 79.5	7	1773	4
DC Out Wall	Al	-150 – 150	79.5 – 80.0	75	–	–
CRID		-150 – 150	80. – 94.	0	1482	19
Barrel Calor	Lead	-150 – 150	94. – 105.	1968	1374	12

Table 3. *Geometry and calculated backgrounds of simulated objects. The 9 GeV electrons travel in the direction of increasing z. Double entries in the "Backgrounds" columns indicate that the object appears with $z < 0$ and $z > 0$. For example, there is a Q3 with $-226 < z < -186$, which receives 1936 MeV/μs, and another at $186 < z < 226$, which receives 155 MeV/μs.*

bremsstrahlung electrons hitting the beampipe. There are few electrons at the larger radius objects because of the detector solenoid field.

Figure 6 shows the energy deposited in the middle layer of the vertex detector and the drift chamber; the average energy of background particles is low. For example, 1945 MeV of energy per μs is deposited in the barrel calorimeter by $\sim$ 1400 photons of average energy 1.4 MeV. The largest radiation dose in the CsI is in the downstream endcap. The first 8 *r.l.* of the innermost 10 cm ($50 < r < 60$ cm) receives a dose of 190 rad in 10^7 s, almost entirely through the front face (there is shielding behind and below the endcap). This value, like all others presented here, does not include the effects of the low-energy ring or injection. As mentioned earlier, backgrounds could be reduced by the addition of bend magnets or by lower pressures.

6. Open Issues

Consistent high-energy and low-energy lattices, with acceptable synchrotron radiation levels, are required to calculate accurate background levels. These lattices may include additional bend magnets and upstream masks to isolate the IR from upstream scatters. In the longer term, a more accurate detector simulation may be required, including showers that start more than two meters from the IP.

The effects of injection on the CsI and the silicon and the methods of dealing with this radiation require additional work. Extra injection shielding may be needed.

The reliability of the methods used in this section will be tested by comparing calculated backgrounds in the *PEP* experiment TPC with the observed rates.

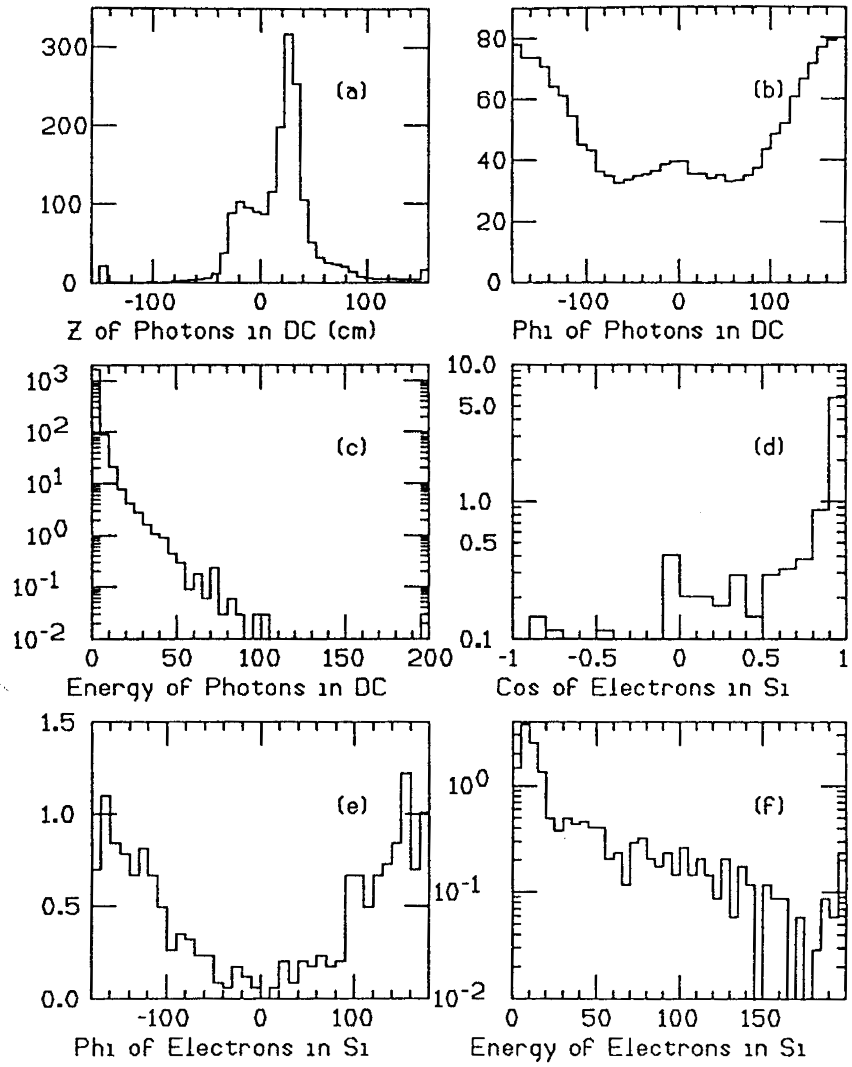


Figure 6. *The distribution of photons entering the drift chamber in* **a.** z *and* **b.** *phi of impact point, and* **c.** *energy (MeV), and the distributions in* **d.** $\cos\theta$ *of particle direction* **e.** *phi of impact point and* **f.** *energy (MeV) for electrons entering the middle layer of the silicon strips. Vertical scales are in particles per* μs. *The peak in* z *in the DC is due to masking elsewhere; the peak in phi in the Silicon is due to the asymmetry of electrons hitting the beampipe (Figure* **3 b***).*

References

1. D. C. Carey, *et al.*, *DECAY TURTLE*, SLAC-246 (1982).
2. K. L. Brown, *et al.*, *TRANSPORT*, SLAC-91 (1977).
3. *Feasibility Study for an Asymmetric B Factory Based on PEP*, SLAC-352 (1989).
4. W. R. Nelson, H. Hirayama and D. W. O. Rogers, *The EGS4 Code System*, SLAC-265 (1985).

Report of the Vertex Detector Group

V. Lüth, A. Breakstone, J. Chauveau, D. Cords,
R. Erbacher, W. Innes, C.K. Jung, D. MacFarlane,
R. Morrison, S. Shapiro, A. Smith, D. Stoker,
J. Va'vra, W. Vernon, M. Witherell,
G. Wormser and M. Yurko

1. Introduction

A finely-segmented silicon vertex detector can provide a high-precision measurement of the impact parameter, b_z parallel and b_{xy} transverse to the direction of the beam, and the angles, the azimuth ϕ and the polar angle θ, of charged particle tracks close to the beam-beam interaction point, and can thereby complement the angle and momentum measurement in the main tracking chamber.

At a high luminosity B Factory, such a system will allow

- the measurement of decay distance between the B and $\bar{B}$ and thereby permit the measurement of time-dependent effects such as CP asymmetries,
- a reduction in the background for lepton and kaon tags which determine the flavor of the second B^0 or $\bar{B}^0$ in the event,
- separation of charm and beauty decays from light quark background events,
- a substantial reduction in combinatorial background in B and D decay, and
- a reduction in the hadronic and Bhabha background in τ decays.

This report summarizes a study of many aspects of the physics program that impact on the design and performance of a vertex detector. Specification for resolution, solid angle coverage and lay-out, and segmentation are derived from a variety of analyses of different physics topics that can be addressed with such a machine (Section 2). Given the specification for position resolution and stereo read-out, it becomes obvious that a vertex detector built from finely segmented silicon diodes (Section 3) represents a very suitable, if not the only viable solution. Different lay-outs and strip and pixel segmentations have been investigated (Section 4). With the help of software tools (Section 5) developed during this Workshop, the resolutions in angle and impact parameter have been studied for different lay-outs, as a function of the number of layers and their spacing, the radius and the thickness of the vacuum pipe and the detectors, etc.. for particles of different momentum and angles (Section 6). No conclusion has yet been reached as to the necessity of an intermediate wire chamber between the silicon vertex detector and the main drift chamber (Section 7), Such a chamber is not expected to improve the resolution for tracks measured in the two other tracking devices, it may, however, enhance the efficiency for the reconstruction of low momentum tracks. Since the performance of the vertex detector depends critically on the radius and thickness of the vacuum pipe, its specifications and design are also discussed (Section 8). Tolerance to background radiation is a major concerned for all detector elements at a machine that operates at high currents. Synchrotron radiation and electro-magnetic showers caused by lost beam particles increase the detector occupancy and can cause radiation damage in the silicon detectors and the associated VLSI electronics (Section 9). Electronics for signal amplification, sparse read-out, and triggering have not been studied in detail; only the basic principles and options are discussed (Section 10). Many groups planning experiments at the SSC are involved in R&D projects addressing very similar performance criteria. Estimates of trigger rates for multi-layer coincidences in double-sided strip or pixel detectors look promising for use in a secondary trigger (Section 11).

Appended to this report are a number of independent contributions on pin diode arrays (S. Shapiro), on the use of microstrip avalanche chambers for the forward direction (J. Va'vra), on the use of a TPC as a forward detector (G. Wormser), on detailed studies of single track measurements (D. Stoker), on studies of intrinsic resolution in silicon for tracks with large angles of incidence (V. Lüth), and on ideas for beam pipe cooling (R. Erbacher and W. Vernon).

2. Design Specifications

The performance of the vertex detector is central to many measurements that can be performed at an asymmetric B Factory, and a variety of studies have been carried out to set the specifications for this device.

2.1. Position Resolution

The primary task of a vertex detector at an asymmetric B Factory is the measurement of the time evolution of the $B^0 - \bar{B}^0$ asymmetry in decays to CP eigenstates, caused by interference between the mixed and unmixed B mesons. The relevant time is the difference in proper time between the two decays, which is proportional to the difference in decay distances. The resolution in measuring this difference in decay lengths Δz should be good enough as not to significantly degrade the measurement of this asymmetry.

A number of studies have been done to investigate the specifications for the vertex detector resolution, simulating the process $B^0 \to f_1$ and $\bar{B}^0 \to f_2$, where f_1 represents a CP eigenstate, like ψK_S^0, and f_2 is a decay that tags the identity of the B meson, like the semileptonic decay to $D^+ l^- \nu$. In the presence of CP violation, the decay rate of a B^0 or $\bar{B}^0$ to a CP eigenstate f_1 is

$$R \propto e^{-\Gamma|\Delta t|}[1 \pm \sin 2\phi \sin(x\Gamma\Delta t)]G_m(\sigma_{\Delta t}),$$

where $x = \Delta m/\Gamma$ is the mixing parameter, and Δm is the difference in mass of the two neutral B meson mass eigenstates, Γ is the inverse of the average B^0 decay time, ϕ is the CP phase for the particular decay mode. Δt is the time difference between the decay of the two B's. G_m describes the resolution of the detector in the measurement of Δt, which is assumed to be described by a single Gaussian resolution function of width $\sigma_{\Delta t}$. Thus the CP-violating rate asymmetry is represented by a constant term $\sin 2\phi$ that is modulated by a sine function of period $T = 2\pi/x$ and convoluted by the Gaussian resolution function, G_m. The attenuation of the amplitude $\sin 2\phi$ remains small as long as $\sigma_{\Delta t} \ll T$. This observation has been verified by analytical calculations.[1,2] and Monte Carlo simulation.

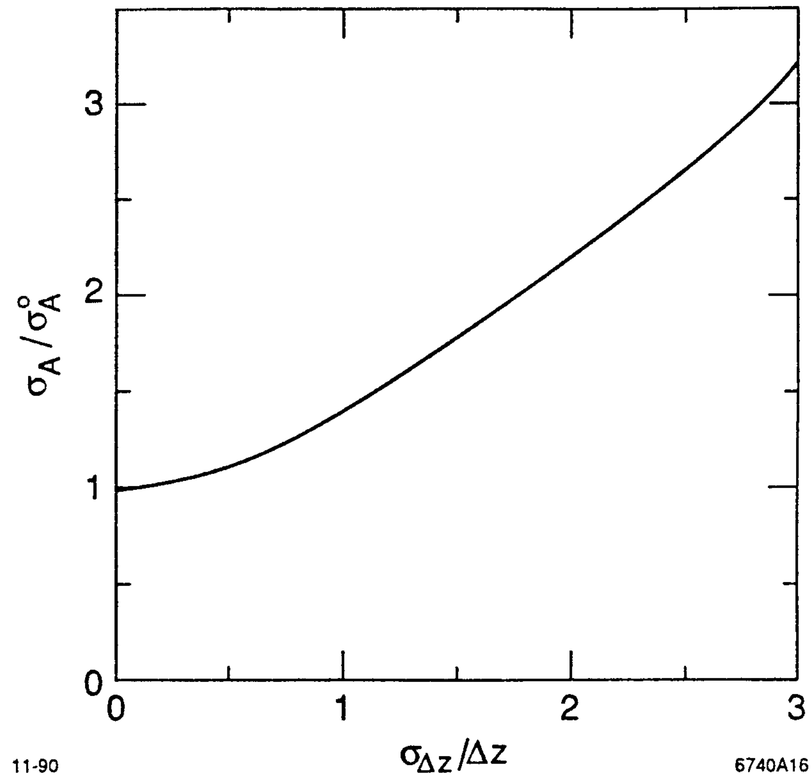


Figure 1. *Dilution of the measurement of the CP asymmetry as a function of $\sigma_{\Delta z}$, the resolution in the decay distance of the two B^0 mesons.*

Figure 1 shows that the error in the *CP* term $\sin 2\phi$ depends only weakly on $\sigma_{\Delta t}$, which is simply related to $\sigma_{\Delta z}$, the error on the decay distance $\Delta z = z_{CP} - z_{tag}$:

$$\sigma_{\Delta t}/ < \Delta t > \approx \sigma_{\Delta z}/\beta\gamma c\tau.$$

In an asymmetric storage ring with beam energies of 3.1 GeV and 9.0 GeV the Υ(4S) resonance is boosted with $\beta\gamma$=0.56 along the direction of the high energy beam. The distances between the decay points of the two B mesons are shown in Figure 2 : the averages are $<\Delta z>=$ 180μm in the direction along the beam, and $<\Delta xy>=$ 32μm in the plane transverse to the beam. Thus $\sigma_{\Delta z} < 0.5\beta\gamma c\tau$ translates to $\sigma_{\Delta z} < 90\mu$m and - compared to a detector with perfect resolution - causes a dilution of the measurement of the *CP* asymmetry in B^0 decays by less than 10%. Monte Carlo simulations have also shown that the resolution $\sigma_{\Delta z}$ does not depend very critically on the radius of the beam pipe and the innermost layer of the silicon detector for decays to simple *CP* eigenstates like

$B_d^0 \to \psi K_S^0$, and a lepton from semi-leptonic decay as a flavor tag. Best estimates for $\sigma_{\Delta z}$ are of the order of 65μm for a beam pipe radius of 20mm, a three layer silicon detector, and an intrinsic position resolution of $\sigma_i = 10\mu$m. Consequently there is some margin of safety for the measurement of the time dependent asymmetry.

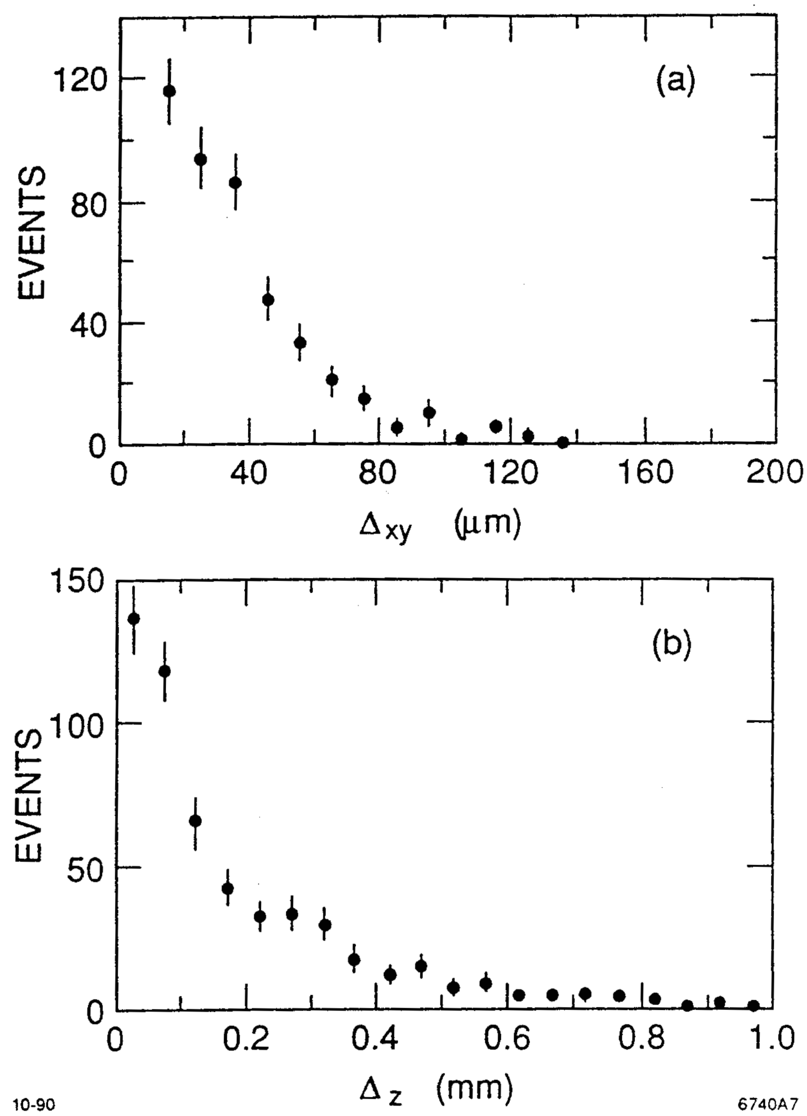


Figure 2. *Distance between the two decay points of the two B mesons (a) in the plane transverse to the beam and (b) along the direction of the beam.*

Good vertex resolution is important for many other measurements, in most cases it helps to separate charm, beauty and $\tau^+\tau^-$ events from light quark background and reduces combinatorial background in mass distributions since the origin of most charged particles can be determined.

2.2. Solid Angle Coverage

To obtain high efficiency for complete event reconstruction, it is essential to cover more than 90% of the solid angle in the center-of-mass system, i.e. the rest frame of the $\Upsilon(4S)$. In the laboratory, this corresponds to $-0.75 < \cos\theta < 0.97$.

Detection efficiencies have been studied for a number of different production and decay processes, namely $B\bar{B}$ events with specific decays like $B_d^0 \to \psi K_S^0$ and $\bar{B}^0 \to l^- +$anything

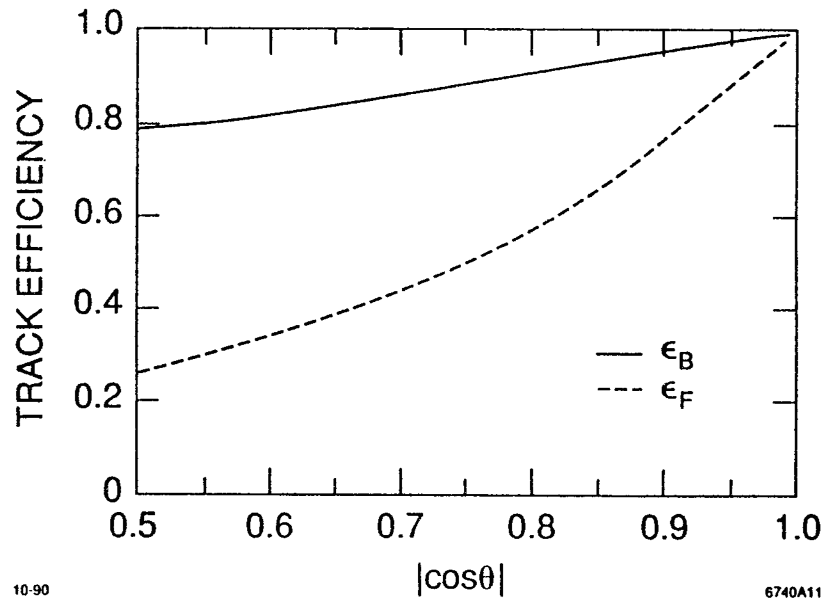


Figure 3. *Single track detection efficiency as a function of the maximum polar angle θ in the forward (F) and backward (B) direction for tracks from B decays (generic mixture)*

or with a generic mixture of B decays.[3] Figure 3 illustrates how the acceptance for single tracks from generic decays of $B\bar{B}$ events drops as the polar angle coverage in the forward and backward direction is restricted. For a cut-off at $|\cos\theta| < 0.9$ about 25% (5%) of the tracks are lost in the forward (backward) direction. Good solid angle coverage is more critical for the reconstruction of complete B, D, or τ decays. For example, the geometric efficiency for detecting all tracks from the decay $B_d^0 \to \psi K_S^0 \to e^+e^-\pi^+\pi^-$ is 77% for a detector extending up to $\cos\theta \leq 0.95$. (compare Figure 4a). An additional cut in the minimum transverse momentum reduces the efficiency further, though for $p_\perp^{min} = 100$ MeV/c the impact is rather modest. Figure 4b illustrates this reduction in efficiency. For decays of higher multiplicity the effects of limited solid angle coverage are more severe, for example, the decay $B_s^0 \to D_s^+ D_s^- \to \phi\pi^+\phi\pi^- \to K^+K^-\pi^+K^+K^-\pi^-$ has an efficiency of 66% for a coverage of $|\cos\theta| \leq 0.95$.

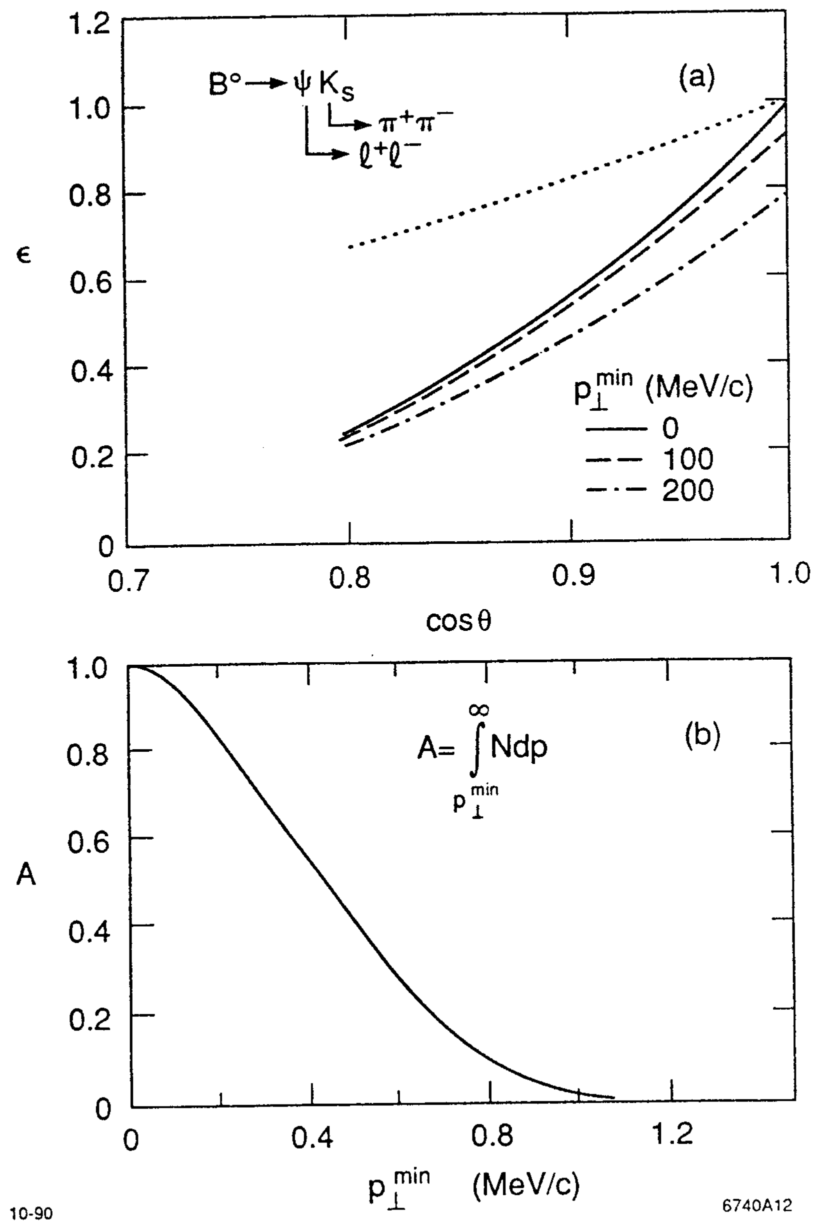


Figure 4. *Detection efficiency for the decay $B_d^0 \to \psi K_S^0 \to e^+e^-\pi^+\pi^-$:* **a.** *$\epsilon$ as a function of the forward polar angle coverage and for different cuts on the minimum track momentum transverse to the beam. The dotted l ine gives ϵ as a function of the backward angle coverage.* **b.** *ϵ for full angular coverage as a function of the cut-off in the minimum transverse momentum $p_\perp^{min}$*

2.3. Segmentation

The average charged particle multiplicity in the decay of an Υ resonance is $\leq$ 12, and apart from the effect of the small boost along the beam direction, particles are distributed rather uniformly. Such events do not require a tracking device with high granularity. The distribution in opening angle for the decay $\phi \rightarrow K^+K^-$ peaks at $\Delta\phi = 0.200$ rad and does not set any specific requirement for detector segmentation.

The reaction $e^+e^- \rightarrow \tau^+\tau^-$ produces charged particle multiplicities, angular and momentum distributions which differ substantially from the kinematics of the Υ(4S) decays. Studies of three different $\tau^+\tau^-$ decay topologies were performed,[4] one charged track from the decay of one τ and one, three, or five charged tracks from the other τ decay. The distributions in Figure 5 lead to conclusions similar to those drawn from the study of Υ decays; the geometric efficiency drops rapidly with polar angle coverage and final state multiplicity, and the minimum momenta peak at rather low values, in particular for decays of higher multiplicity. A cut in the minimum detectable transverse momentum at 200 MeV/c would reduce the number of fully reconstructed events to less than 50%. The minimum angle between two tracks peaks at about 100 mrad, large compared to the two track resolution of 10 mrad which can be obtained with a silicon detector of 50μm pitch.

2.4. Summary

Performance requirements for the vertex detector are thus well-matched to a silicon diode array with an intrinsic resolution of $\sim 10 - 20\mu$m, requiring a pitch of $\sim 50\mu$m, for strips several centimeters long. The requirement that the B decay vertex be well-measured in the direction along the beam demands silicon detectors with coordinate read-out transverse to the beam direction, in contrast to all present applications

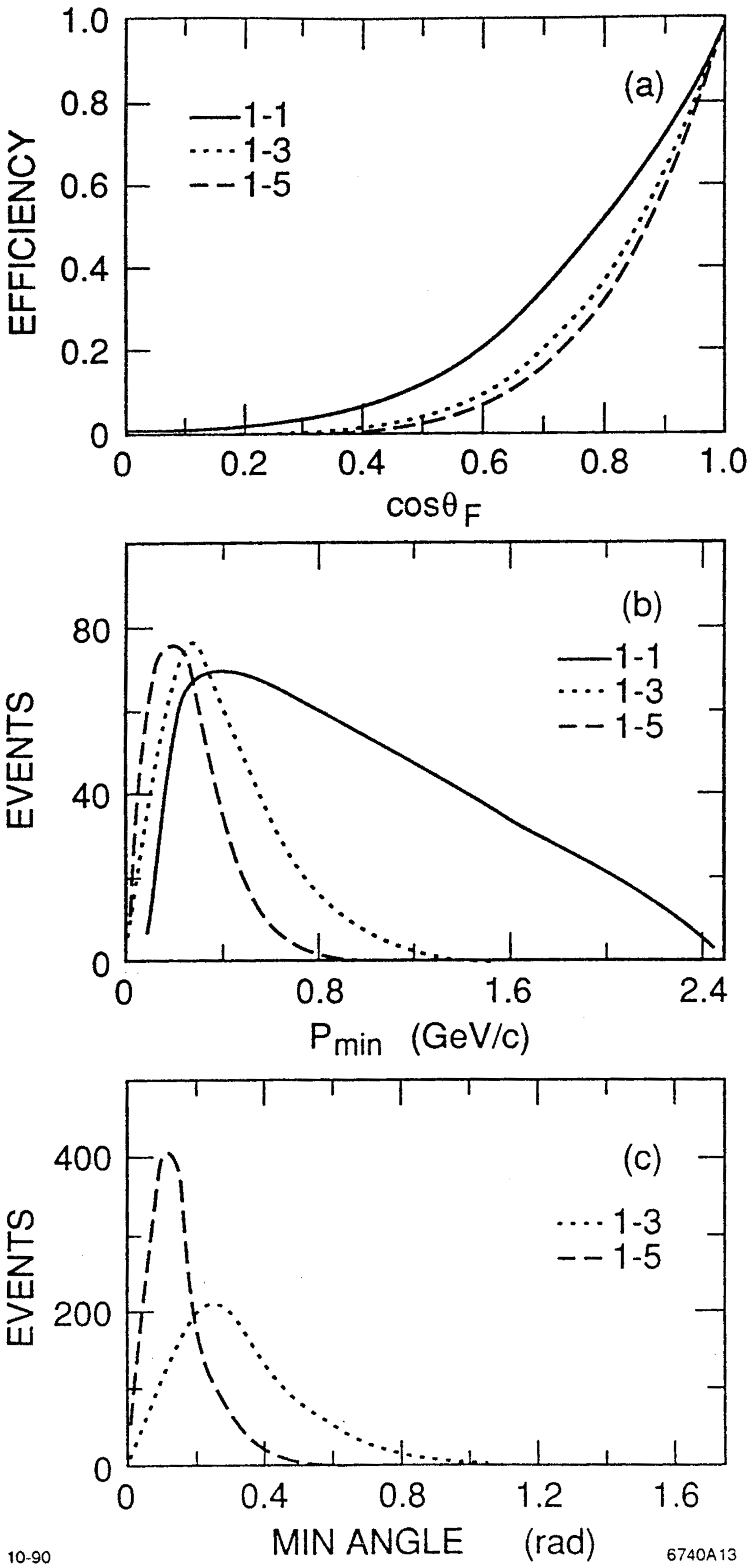


Figure 5. *Study of $\tau^+\tau^-$ production and decay with different topologies of the charged tracks, 1-1, 1-3, and 1-5 (the relative normalization of the curves is arbitrary):* **a.** *geometric efficiency for detecting all tracks as a function of the polar angle coverage in the forward direction,* **b.** *spectra of the minimum track momentum per event, and* **c.** *distribution of the minimum opening angle of any two tracks in a given event*

of such devices at colliding beam machines. While vertex detection along the beam axis is mandatory, good precision in the $x-y$ plane is also required. Thus silicon detectors with read-out in two orthogonal coordinates are necessary. This can be done using pairs of single-sided detectors, double-sided strip detectors or pixel arrays with direct read-out.

As will be shown in the studies of track resolution the performance limits of a multi-layer silicon vertex detector are not set by its intrinsic resolution, but by the multiple scattering in the material of the beam pipe, the detector support, and the detectors themselves. The segmentation requirements are not set by the multiplicity of the decay particles, but by the additional signals generated by beam related background. The choice of segmentation is also influenced by the maximum input capacitance per channel that is compatible with the design of low noise and low power amplifier circuits. The solid angle coverage is obviously an important parameter for all detector elements. Tracks emitted at small angles to the beam undergo larger multiple scattering and in many cases do not contribute to the determination of the vertex. They may, however, in many cases aid in the event reconstruction or background suppression.

3. Silicon Detectors

Silicon detectors with finely segmented electrodes are an outgrowth of the semiconductor counters that have been used in nuclear physics for more than twenty years. In high energy physics, their use as high resolution tracking devices was pioneered by two groups at CERN.[5,6] Since then, many other experimenters have built on this experience and have employed commercially available silicon detectors. At present, several silicon vertex detectors are operating in e^+e^- experiments at SLC and LEP. Future detectors of this kind will greatly benefit from the experience that has been gained in their construction and operation.

Silicon strip detector vary greatly in their dimensions, some have up to one thousand strips. The minimum pitch realized so far is 12.5μm; the maximum length is 90mm, limited by the 4" diameter of the silicon wafer. Digital and analog read-out has been used, depending on the spatial resolution and two-track separation required. Most experiments have used analog pulse height information and determined the exact position of the particle trajectory by a pulseheight weighted mean of the strip coordinates. The best resolution achieved so far is 2.5μm.

Silicon detectors have high intrinsic position resolution because of their precision fabrication, the short range of electrons, the limited charge diffusion, and large ionization statistics in silicon.

The advantages of silicon detectors are the following:

- good localization of the ionization charges due to short range of electrons and photon in the material, limited diffusion of the drifting charges;
- large ionization statistics due to the narrow band gap and low ionization potential, 80 electron-hole pairs per μm;
- good mechanical rigidity allowing for low mass support structures;
- VLSI technology guaranteeing high precision in the fabrication of the strip diodes, and a potential of full integration of detector elements with the circuitry necessary for amplification and read-out.

The main drawback of finely-segmented arrays of silicon micro-strip detectors is the multiple scattering in the detectors that are typically 300μm thick. The thickness of the detectors determines the amount of charge produced by ionization and thereby the size of the signal. For a thickness of 300 μm and a strip pitch of 25 μm and length of 8 cm, typical signal to noise ratios of 20:1 have been obtained for minimum ionizing particles at normal incidence. Given the expected improvements in electronics, substantially thinner detectors may be feasible in the near future. Thinner depletion depths will also prevent the degradation of the resolution for tracks of non-normal incidence, which is caused by the spread of the ionization charges over many of the narrow read-out strips. The thinner detectors will reduce the multiple scattering, but will also reduce the mechanical strength of the detectors which may result in the need for external support.

3.1. Strip Detectors

Single-sided silicon strip detectors have been used as vertex detectors in many fixed target and colliding beam experiments.

Single-sided silicon detectors have strip diodes on one side of the high resistivity wafer while the other side is covered by a uniform ohmic contact.[7] Double-sided strip detector are fabricated by subdividing the ohmic side into strips, placing a p^+ blocking strip between every two n^+ read-out strips. The p^+ blocking strips interrupt the conducting n-type channel and thus increase the interstrip resistance so that full depletion can be obtained when the reverse bias voltage is applied (see Figure 6).[8] Based on this idea, double-sided detectors have been built up to size of $5 \times 5\text{cm}^2$ and are presently installed in the ALEPH detector at LEP. The strip pitch is 25μm on the junction side (x coordinate) and 50μm on the ohmic side (z coordinate) allowing for the insertion of the p^+ blocking in between the two contiguous n^+ strips On both sides a bonding pad connects every fourth strip to the read-out. The floating strips are coupled capacitively inducing a signal on the read-out strip and allowing for interpolation of the coordinate in between. First results show that the devices work, though they are substantially more difficult to fabricate and operate. The position resolution was measured in a test beam to be $\sigma_x = 12.8\mu$m on diode side and $\sigma_z = 18.7\mu$m on the ohmic side.[9] The ratio of the total signal for a minimum ionizing particle to the single channel noise has been measured to be roughly 12:1, somewhat less than the typical value of 20:1 for single-sided strips of similar dimensions.

Double sided strip detectors are presently under study in many laboratories, but at present detectors that meet the specifications required are not commercially available.

Recently, Hamamatsu Co. has fabricated detectors of similar design, they were successfully tested in a beam at FNAL.[10] With a double-sided detector one measures two coordinates per detector, but ambiguities due to multiple hits are not necessarily eliminated. While for many applications strips with small stereo angles ($\leq 20°$) may be adequate because the vertex reconstruction is performed in the $r-\phi$ plane, for an asymmetric B Factory the measurement of both coordinates is equally important and therefore only strip detectors with 90° stereo angle are considered here.

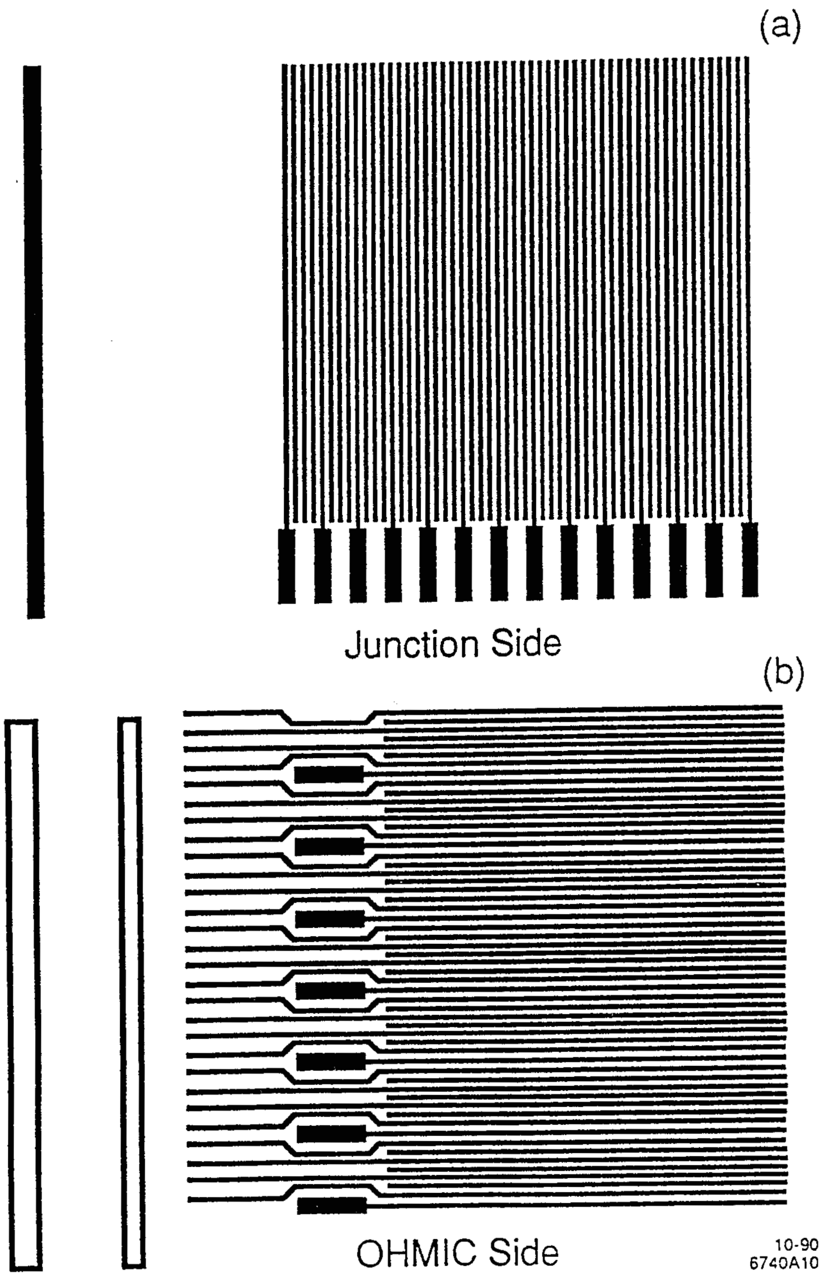


Figure 6. *Details of the strip lay-out for the ALEPH double sided silicon detectors, for*
a. *the junction and*
b. *the ohmic side.*

3.2. *Pixel Detectors*

Pixel devices which provide a space point per layer are presently under development.

Alternatives to the concept of strip read-out are pixel devices with truly two-dimensional coordinate read-out. Charged Coupled Devices (CCD) which store the ionization in two-dimensional potential wells are commercially available and have been used successfully in a fixed target experiment at CERN. [11] Typical pixel sizes are 22 μm x 22 μm in arrays of several cm^2, the position resolution is of the order of 5μm. CCD's are inappropriate for a high rate machine like the B Factory for a variety of reasons. First, the serial read-out results in a typical read-out time of 10ms for a device with 10^5 pixels. Secondly, the shallow depletion depth of 12μm results in only about 1000 charge carriers per minimum ionizing particle and makes cooling of these devices mandatory; it also makes these devices extremely susceptible to degradation by radiation.

A number of researchers are presently working and have succeeded in producing prototypes of monolithically integrated pixel or strip detectors, i.e. devices that have the detector diodes and the amplifier and read-out circuit fabricated on the same silicon wafer.[12,13] Such devices not only avoid large numbers of wirebonds and connections, their small capacitance also leads to a simplification of the read-out electronics and event buffering.

An alternative to this approach is a hybrid structure, a silicon PIN diode array bonded to a mating array of amplifier read-out circuits.[14,15]

Indium bump bonding provides the connections between the detector pixels and the VLSI circuit via an array of aligned indium metal bumps that are cold welded under pressure to form ohmic contact.[16] Figure 7 gives a schematic representation of such a hybrid detector. The advantage to this approach is that the detector and the electronics are designed and constructed as two separate chips, each optimized for its specific function. Thinning of the electronics chip seems quite feasible, thus avoiding large build-up of material.

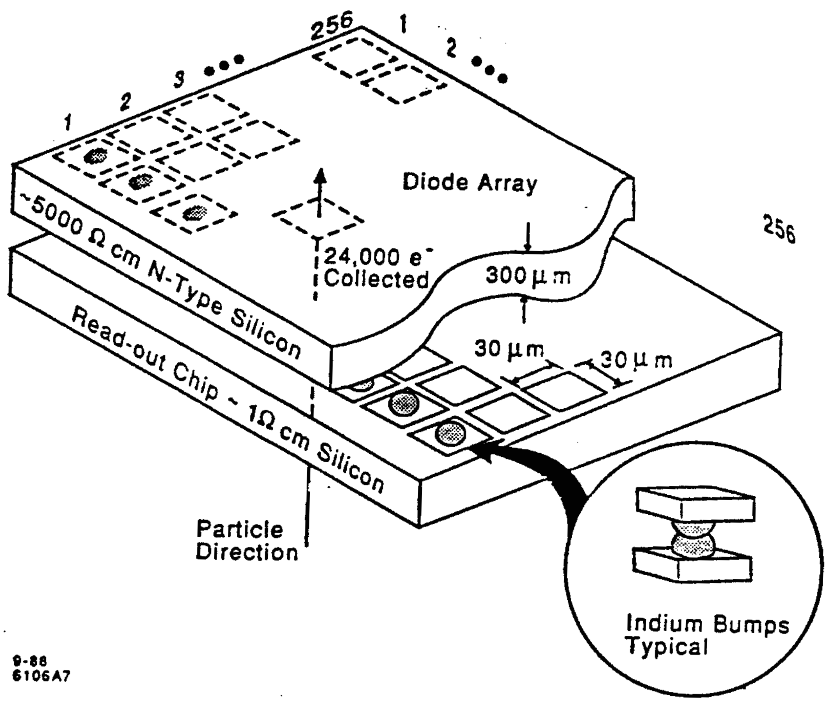


Figure 7. *Schematic of a hybrid pixel detector made of a PIN diode array and a bump bonded array of read-out circuits*

3.3. *Silicon Drift Chambers*

Another very challenging development that aims at full two-dimensional read-out are silicon drift chambers, in which the internal field in the silicon forces the charge created by an ionizing particle to drift over long distances parallel to the detector surface. The transit time of the electrons inside the detector measures the distance of the incident particle from the anode. The anode is divided into short segments to measure the coordinate perpendicular to the drift direction as illustrated in Figure 8a. In a drift field of 500V/cm a typical velocity is $v = 4\mu$m/ns.[17] In beam tests position resolutions of better than 4μm have been obtained (see Figure 8b).[18] The advantages of a silicon drift chamber over a strip detector are its higher granularity, the continuous read-out, and the lower number of read-out channels. The anode capacitance is much smaller than for a strip detector of the same dimension resulting in a much smaller amplifier noise, which is the principal reason for its excellent position resolution. Its drawbacks are the more elaborate detector fabrication, the need for space-time calibration, and the novelty of the device. Also, this device has the same left-right ambiguity as ordinary drift chambers, and the two particle resolution is expected to be limited. The drift velocity in silicon depends on the absolute temperature as $v \propto T^{-2.4}$, and consequently the coordinate measurement is sensitive to temperature variations of $0.1°C$. While it is clearly impractical to stabilize the detector temperature to this level, calibration of the electron drift velocity can be obtained by measuring the drift time of a test charge injected from an electrode on the detector.

4. Layout and Segmentation

With several years of experience with silicon vertex detectors a number of design rules have been developed, most of which have been incorporated into several layout schemes which are presented below. The principal rules are:

- two or more layers, with the first layer as close to the interaction point as practical,
- maximum solid angle coverage,
- full coverage per layer, and small areas of overlap of modules that form a layer,
- modularity of detectors and electronics, as imposed by fabrication and assembly cost,
- a minimum amount of material between the interaction point and the first layer, and
- a minimum support structure inside the active area.

The choice of the segmentation of a strip detector is driven by the following considerations:

- the pitch should be compatible with the required position resolution of about $10\mu m$,

- the expected occupancy per strip should not exceed a few percent of all read-out channels,
- the area of an individual strip should be compatible with the signal/noise and power dissipation characteristics of the amplifier circuit,
- the dark currents in the presence of radiation damage should not exceed a typical limit of a few nA/channel.
- the density of the amplifier circuits should be compatible with the layout for support and connections.

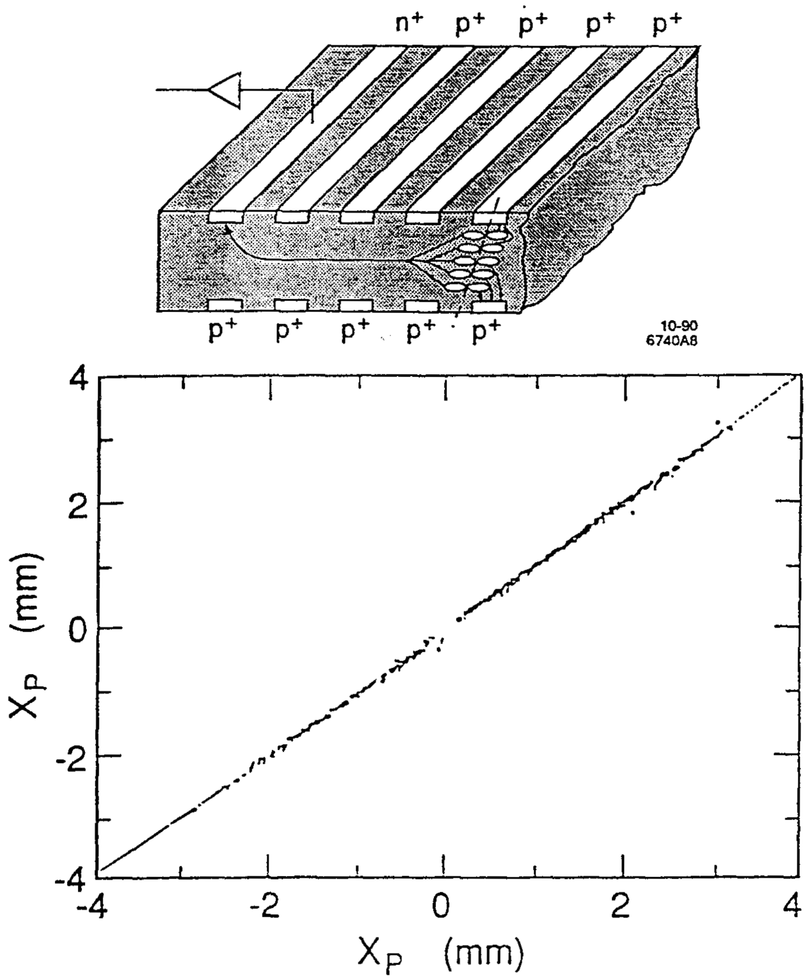


Figure 8. *A silicon drift chamber:* **a.** *schematic and* **b.** *comparison of the measured and predicted position in a test beam*

Schematic layouts for two silicon strip vertex detector are given in Figures 9 and 10 . A typical layout for a pixel detector is shown in Figure 11 . A summary of the dimensions of the detector elements and their segmentation are given in Tables 1, 2, and 3.

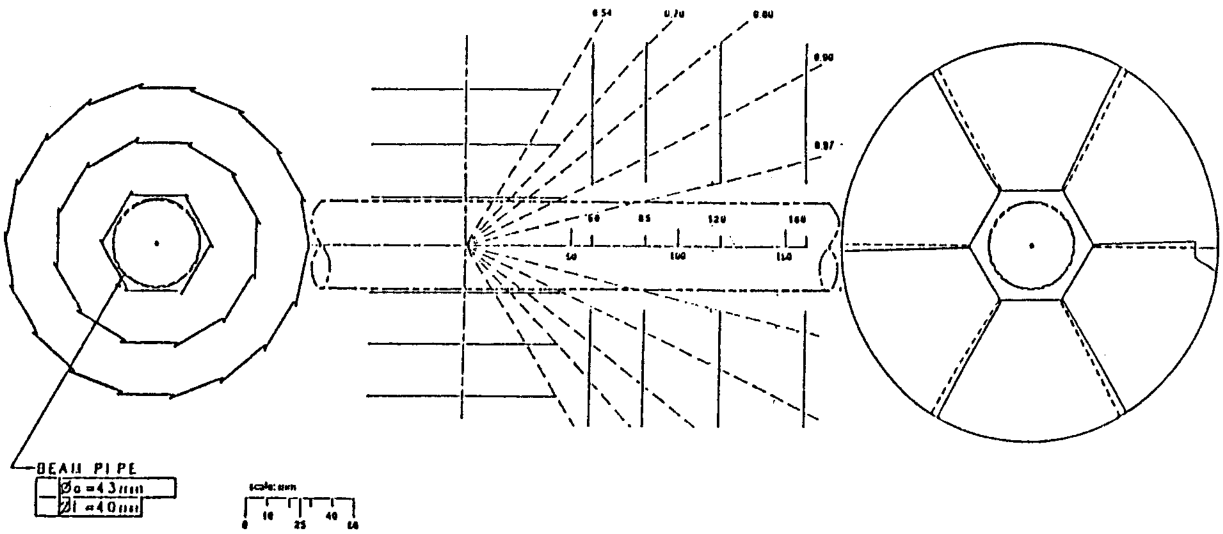


Figure 9. *Layout of a silicon strip detector, Example I. All dimensions are given in mm*

Parameters	Layer 1	Layer 2	Layer 3
Central Detector			
radius of layer (mm)	23	46	69
active length of module (mm)	89.6	178	178
active width of module (mm)	25.6	25.6	25.6
size of detectors (mm^2)	25.6×89.6	25.6×89.6	25.6×89.6
no. modules/layer	6	12	18
no. detectors/module	1	2	2
no. strips/detector	$2 \times 512 + 1792$	$2 \times 512 + 1792$	$2 \times 512 + 1792$
Forward Detector			
distance from IP (mm)	110	140	
inner radius of detector (mm)	28	28	
outer radius of detector (mm)	85.6	85.6	
no. detectors/layer	6	6	
no. strips/detector	$2 \times 512 + 1024$	$2 \times 512 + 1024$	

Table 1. *Dimensions and segmentation of a silicon vertex detector (Example I)*

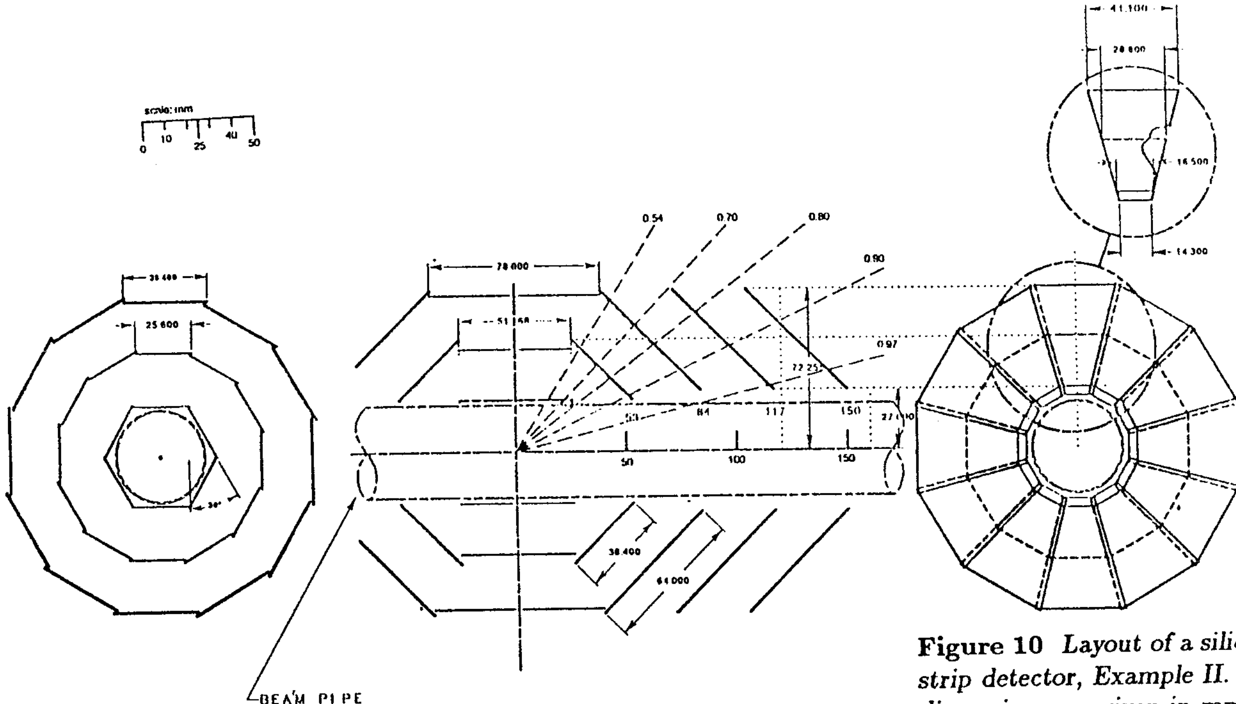


Figure 10 *Layout of a silicon strip detector, Example II. All dimensions are given in mm*

Parameters	Layer 1	Layer 2	Layer 3	Layer 4
Central Detector				
radius of layer (mm)	23	46	69	
active length of detector (mm)	51.2	51.2	76.8	
active width of detector (mm)	25.6	25.6	38.4	
No. detectors/layer	6	12	12	
No. strips/detector	$2 \times 512 + 1024$	$2 \times 512 + 1024$	$2 \times 768 + 1536$	
Forward Detector				
distance from IP (mm)	50	85	120	155
height of detector (mm)	38.4	64.0	64.0	64.0
base of detector (mm)	14.3	16.5	16.5	16.5
top of detector (mm)	28.8	41.1	41.1	41.1
no. detectors/layer	12	12	12	12
no. strips/detector	$2 \times 512 + 768$	$2 \times 512 + 1280$	$2 \times 512 + 1280$	$2 \times 512 + 1280$

Table 2. *Dimensions and segmentation of the silicon vertex detector (Example II)*

All three layouts are cylindrically symmetric about the beam direction. The polar angle coverage extends from $\cos\theta = -0.75$ to $\cos\theta = 0.97$, corresponding to $|\cos\theta^*| = 0.91$. Any particle emitted from the interaction region will traverse three planes, each measuring two orthogonal coordinates. The detector sections are composed of individual detector modules of rectangular or trapezoidal shape. This modularity is imposed by the detector fabrication and assembly costs. For the strip detectors, the overall dimensions of the individual detector modules are chosen as not to exceed the boundary given by a 4 inch silicon wafer. The channel count corresponds to a segmentation with 50μm pitch in both coordinates; strip length varies between 25mm and 45mm. The expected position resolution for 50μm pitch is better than 10μm for tracks of normal incidence.

The strip detectors are divided into a central section and endcaps. The two layouts in Figures 9 and 10 differ primarily in the design of the endcap sections. In the central section the detectors are of rectangular shape. They are arranged in polygons of 6, 12, and 18 sides. The endcaps are made of sections that are arranged to cover the surface of a circular disc (example I) or form a twelve sided pyramid (example II). For example I, there are only two types of detector modules, the rectangular elements in the central section, and the 60° pie sections forming the endcap discs. Example II has two sizes of rectangular modules for the central section, and two sizes of modules of trapezoid shape for the endcaps. The rectangular modules have orthogonal read-out strips, the endcap module have strips measuring azimuthal angles ϕ_i and the radial distances r_i. Example I is somewhat simpler in concept, and the central section resembles existing vertex detectors at SLC, LEP and the Tevatron. There is room for VLSI readout circuits on both ends of the rectangular modules, and for a thin endplate to support the modules. The endcap modules could be supported from a ring at the outer radius and/or inner ring.

The potential problem with this layout is the amount of multiple scattering caused by the material at the end of the cylindrical central section. Also, complete coverage in the transition region between the central and endcap sections requires a rather large radial extent of the first two endcap discs. Furthermore, the angle of incidence of tracks on the detector planes for intermediate polar angles is rather large, leading to somewhat reduced track resolution and larger multiple scattering. Example II represents an attempt to improve these short comings of the simpler layout given in example I.

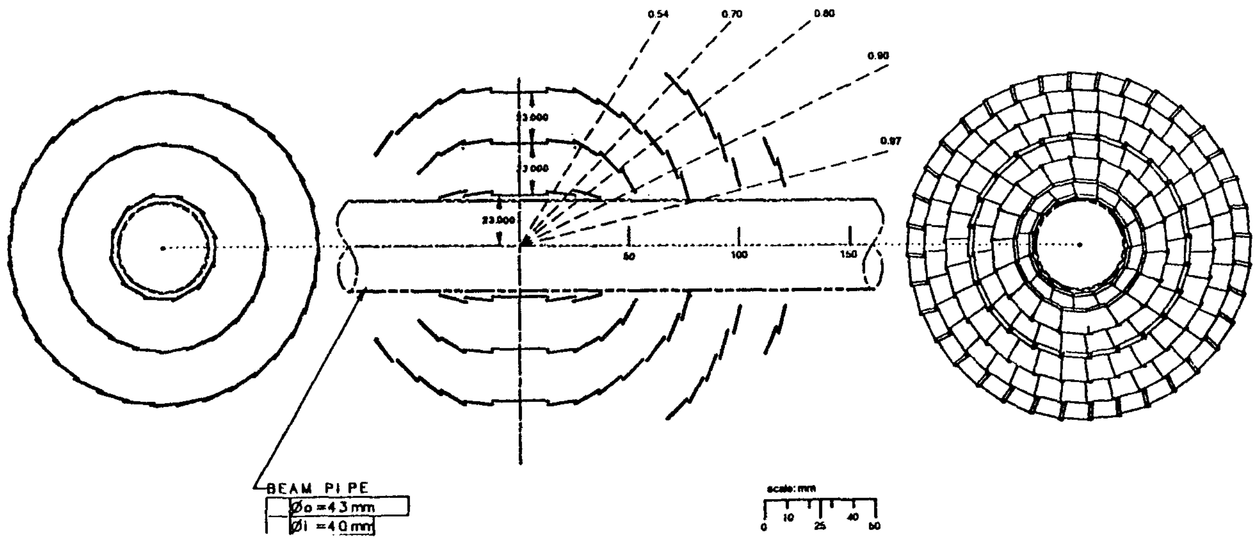


Figure 11. *Schematic layout of a Silicon Pixel Detector, Example III*

In the pixel detector, the small quadratic detector modules are arranged to form layers of ellipsoidal contour. Each pixel module holds 256×256 pixels of $50 \times 50 \mu m^2$ dimension, thus there are 40,000 pixels per cm^2. This is to be compared with a segmentation of $50 - 100/cm^2$ for a strip detector with a pitch of $50\mu m$ and a strip length of $2-4$cm. There are a total of about 900 individual pixel modules containing about 60 million pixels. A summary of the dimensions of the detector elements and their segmentation is given in Table 3.

Parameters	Layer 1	Layer 2	Layer 3	Layer 4
min. radius of layer (mm)	23	46	69	92
No. detectors/layer	72	154	276	88
No. pixels/detector	256 x 256	256 x 256	256 x 256	256 x 256
No. pixels/layer	4.7×10^6	10.1×10^6	18.1×10^6	5.8×10^6

Table 3. *Dimensions and segmentation of a pixel vertex detector (Example III)*

5. Software Tools

A number of programs have been developed to study the performance of vertex detectors of different segmentation and layout. The most commonly used programs are ASLUND, adapted by A. Weinstein and updated by A. Snyder,[19] GEANT, adapted by D. Aston and A. Breakstone,[20] TRK-SIM, written by D. Stoker,[21] and SITRK and SISIM, written by V. Lüth.[22,23] In addition, several groups have adopted software written for other e^+e^- experiments for study of detection efficencies and vertex resolution.

5.1. *ASLUND - Generator for $B\bar{B}$ events*

This set of subroutines serves as a front end to the LUND Monte Carlo program, it generates the $B^0\bar{B}^0, B^+B^-$, and $B_s^0\bar{B}_s^0$ events, allows for mixing, CP violation, etc.. The detector response is simulated by a set of simple simulation routines for tracking of charged and neutral particles. Average resolution functions are used for parameter smearing, efficiencies are based on typical values extrapolated from present experience. For the vertex detector, a three layer cylindrical detector of infinite length and constant resolution is assumed. Multiple scattering is included in the evaluation of tracking errors, energy loss in the beam pipe or wall material is not yet incorporated. Vertex fitting routines have been adopted, they take into account the full error matrices from the single track reconstruction.

5.2. *GEANT Simulation*

In the future, more detailed detector simulation will be required for pattern recognition, trigger, and tracking studies. In preparation, the GEANT Monte Carlo, developed at CERN and used world-wide, has been transferred to SLAC. Two versions of GEANT (versions 311 and 313) now exist (libraries GEANT311 and GEANT313 on the BFACT 192 disk) for use with the VM operating system on the IBM cluster. Both versions use the ZEBRA memory management system and GEISHA, the hadron interaction package. Since the CERN versions use TV and GKS graphics routines (for versions 311 and 313, respectively) David Aston translated the TV routines to unified graphics (UG) routines for use at SLAC, and Alan Breakstone did the same for the GKS routines. HBOOK calls have been replaced by the corresponding HANDYPAK routines.

Currently, the detector simulation is set-up for the beam pipe and a preliminary version of the vertex silicon detector. The simulation of the drift chamber is in preparation. More work needs to be done to interface the GEANT detector simulation with event generators and track reconstruction programs.

5.3. Special Vertex Detector Software

A number of stand-alone programs have been written to obtain a fast evaluation of specific performance data. TRKSIM and SITRK project individual tracks of a given momentum and angle of incidence on to a silicon detector of a proposed geometry and segmentation. They evaluate the resolution in measured track angles and impact parameters. TRKSIM calculates exact track helices in a magnetic field, and takes into account the impact angles of the tracks onto the exact detectors geometry, as well as multiple scattering and the variation of the intrinsic resolution as a function of angle. The results are obtained from averages over many tracks. SITRK assumes strictly Gaussian error functions for the measurement and multiple scattering in the beam pipe, the detector walls and the detectors themselves. All are of exact cylindrical geometry. The addition of external tracking devices, like the main drift chamber and a possible intermediate vertex chamber is is provided for, taking into account stereo angles and different resolutions.

SISIM examines the intrinsic position resolution of a single silicon detector as a function of the detector thickness, segmentation, the signal/noise ratio of the read-out electronics, and the angle of incidence of a minimum ionizing particle. It assumes a Landau distribution for the charge deposition, including additional contributions from large delta ray losses. It also permits the study of different algorithms for centroid finding with analog or digital read-out.

6. Resolution

The track reconstruction errors have been studied for a three layer silicon detector of cylindrical geometry as a function of detector spacing, detector and beam pipe thickness, intrinsic resolution, *etc.*[22] The results were obtained from a fit to the track parameters taking into account multiple scattering and intrinsic position errors, assuming cylindrical detector geometry. The track trajectories are described by a helix with an axis parallel to the B field (z-axis). We chose to parameterize the tracks by the following standard parameters $\alpha_i = (\kappa, b_{xy}, \phi, b_z, s)$, where b_{xy} and b_z are the components of the closest distance of approach to the origin, $\phi = \tan^{-1}(p_y/p_x)$ is the azimuthal angle, s is related to the polar angle θ as $s = \cot\theta = p_z/p_{xy}$, and κ is the curvature $\kappa = -0.03Bq/p_{xy}$. In the limit of small impact parameters, the predicted coordinate measurement at a radius r_i can be expressed as:

$$d_i^o = (b_{xy} + r_i\phi + r_i^2\kappa)\cos\mu_i + (b_z + r_i s)\sin\mu_i$$

For n measurements d_i at radii r_i with stereo angles μ_i and position errors σ_i the track parameters are obtained by a χ^2 fit with

$$\chi^2 = \sum_{i,j}(d_i - d_i^o)W_{ij}(d_j - d_j^o)$$

where W_{ij} is the inverse of the measurement error matrix, $W_{ij}^{-1} =< \partial d_i \partial d_j >$. For m scattering planes at radii r_k, $k = 1, m$ the added contributions to the measurement error are:

$$< \partial d_i \partial d_j >= \sigma_i^2 \partial_{ij} + (\cos^2\mu_i + \sin^2\mu_i \sin^{-2}\theta)\sin^{-2}\theta$$

$$\times \sum_{k<i,j} < \xi_k^2 > (r_i - r_k)(r_j - r_k)$$

$< \xi_k^2 >$ refers to the rms of the multiple scattering angle in layer k. The sum runs over all scattering planes inside the measurement plane.

For the study of the three layer vertex detector, the track momentum is assumed to be measured in the outer tracking chamber and not varied in the fit. The detectors are equally spaced in radius, $r_1 = 23$mm, $r_2 = 46$mm, $r_3 = 69$mm. They are assumed to be 300μm thick, have a constant position resolution of 10μm in the two orthogonal coordinates $d_i = r_i\phi$ and $z_i = r_i \cot\theta$.

This study ignores the degradation of the position resolution for non-normal incidence. This effect is discussed in a separate note contributed to this workshop.[23] The calculations assume cylindrical detector planes of unlimited length. Thus the results can be improved by the introduction of end-cap detectors that reduce the effects of multiple scattering and non-normal incidence. Some of these issues are addressed in a separate note by D. Stoker.[21] Also not included are errors in the placement and alignment of the detector modules, instabilities in their position with time and possible effects of temperature changes during operation. Uncertainties in the position of the beam-beam interaction point are also ignored though, as stated earlier, for the measurement of the time dependent *CP* asymmetries in *B* decay knowledge of the interaction point is not required.

Resolution studies are presented for detectors with a fixed intrinsic resolution of 10μm.

6.1. *Momentum and angular dependence*

Figure 12 shows the errors on the impact parameters b_{xy} and b_z, and on the angles ϕ and θ. In the $x-y$ plane, a track of 1 GeV/c momentum has an impact parameter of 31μm, and an angular error of 1.2mrad. The errors due to multiple scattering decrease with momentum as $1/p$, and decrease with the polar angle θ: $\Delta b_{xy} \propto \sin^{-3/2}\theta$, $\Delta\phi \propto \sin^{-3/2}\theta$, $\Delta b_z \propto \sin^{-5/2}\theta$, and $\Delta\theta \propto \sin^{-1/2}\theta$.

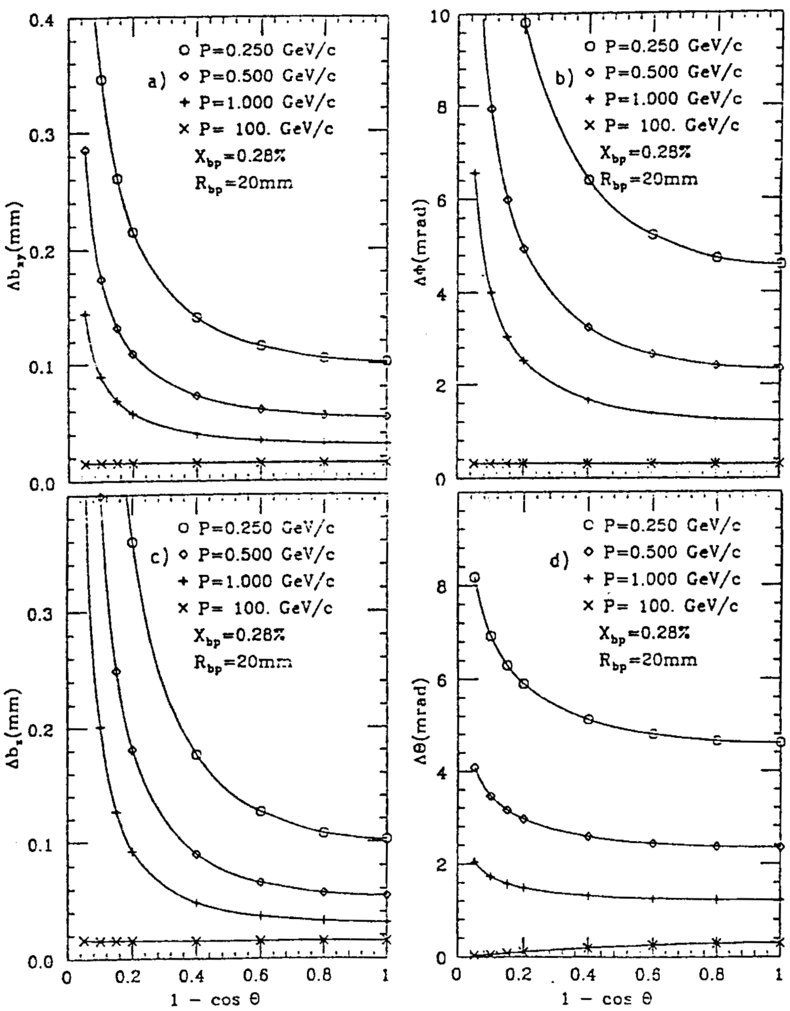
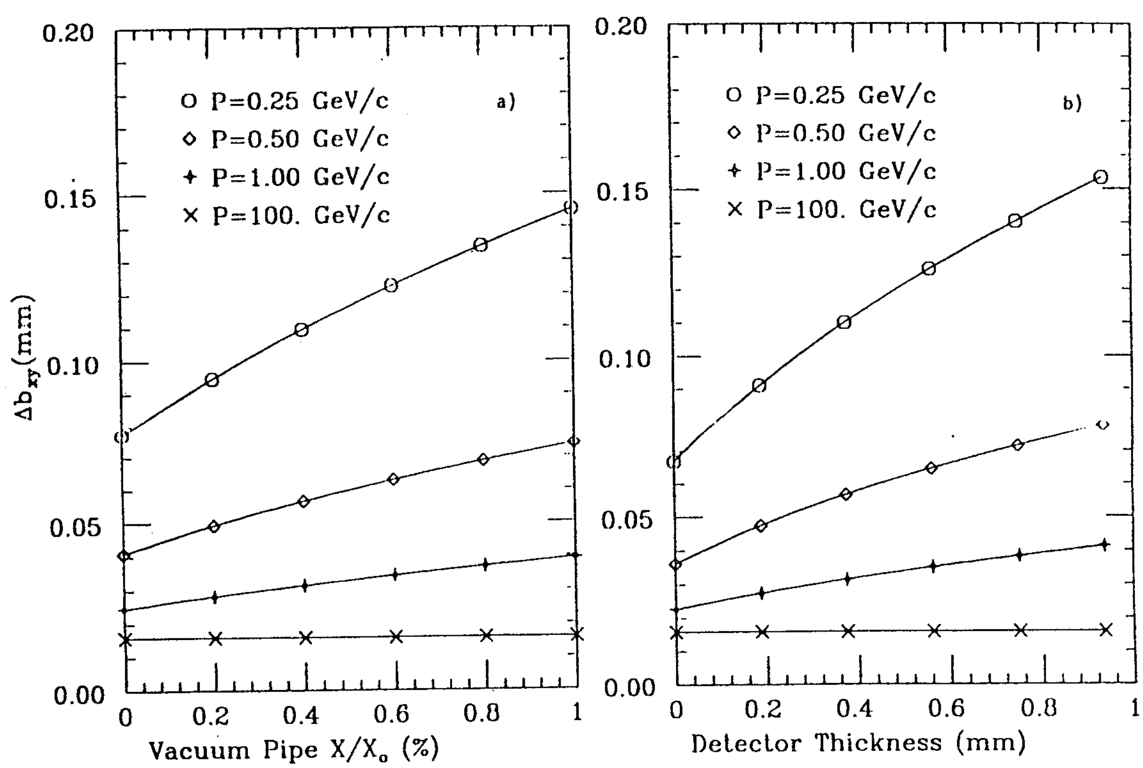


Figure 12. *Track measurement errors for a three layer silicon vertex detector as a function of the particle momentum p and the polar angle θ:*
a. *impact parameter b_{xy},*
b. *azimuth angle ϕ,*
c. *impact parameter b_z, and*
d. *polar angle θ*

6.2. Beam pipe and detector thickness

We have chosen to place the detector outside of the vacuum chamber because this will avoid problems of assembly, access, vacuum feedthroughs, etc. and will not substantially increase the amount of multiple scattering so critical here. We have assumed a double walled pipe made of 2x0.5mm of beryllium (0.28%X_o). This will allow for a cool gas or fluid to be pumped in the annulus between the two walls to carry the heat load from the beam. Such a pipe will also have the skin depth needed to shield the detectors and their electronics from the beam. In practice, it is necessary to add a lining of a few μm of copper on the inside, to absorb soft photons from synchrotron radiation.

Figure 13. *Resolution in the impact parameter as a function of the thickness of*
a. *the beam pipe X_{bp} = X/X_o(%) (with 300μm detectors) and*
b. *the silicon detectors (with a 1mm thick beryllium beampipe) measured in the x-y plane where $\Delta b_z = \Delta b_{xy}$ and $\Delta\phi = \Delta\theta$.*



The track resolution is dominated by multiple scattering in the beam pipe.

Figure 13 shows the effect of variable beam pipe and detector thickness on the error in the impact parameter b_z. Over this limited range the resolution varies roughly linearly with the beampipe thickness from a minimum of 27μm to 40μm for a beampipe of 1.0% of a radiation length. Doubling the detector thickness from 300μm to 600μm, increases the resolution from 31μm to 36μm at normal incidence, and from 93μm to 115μm at cos $\theta = 0.8$.

6.3. Radial spacing of layers

The detector planes are arranged such that any particle emitted from the beam center will traverse three planes, each measuring two coordinates. In principle, one only needs two planes to measure positions and angles. In practice, there are inefficiencies due to dead channels, edges, and possibly support structures so that the addition of a third layer reduces detection losses. Also, pattern recognition in the presence of background requires at least one redundant measurement.

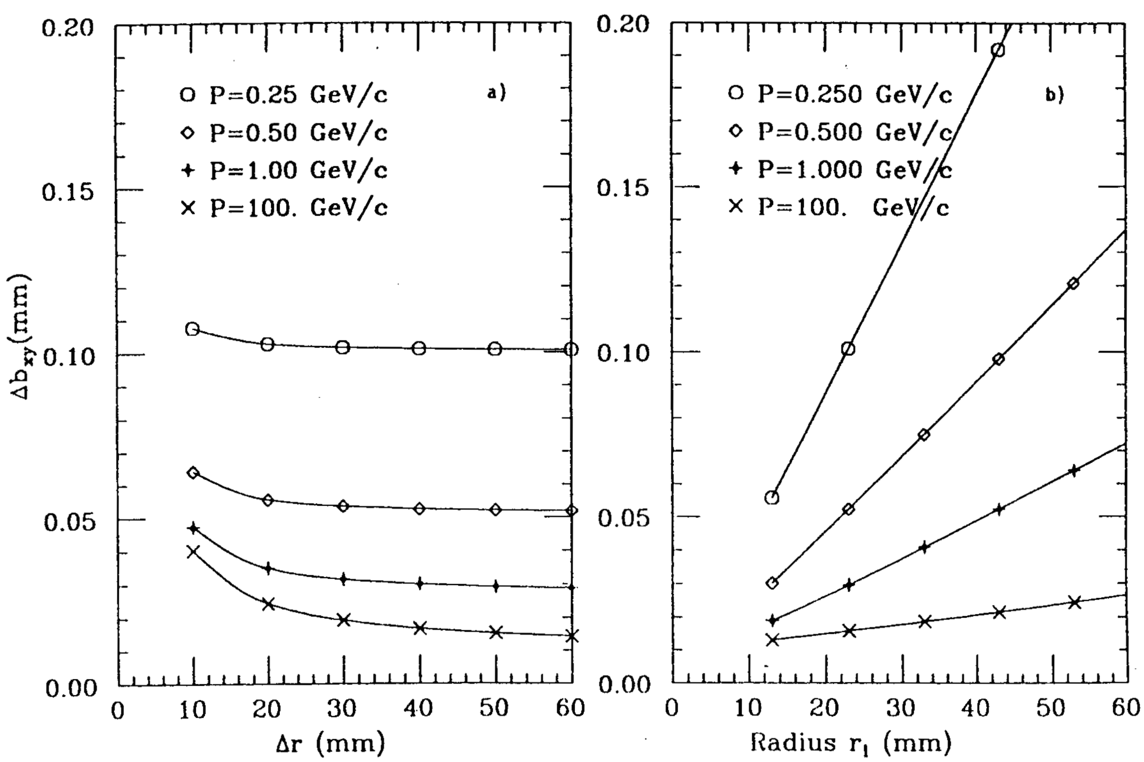


Figure 14. *Resolution in the impact parameter as a function of* **a.** *the spacing of the layers of a two layer silicon detector and* **b.** *the radius r_1 of the innermost layer for a silicon detector with two layers spaced by $\Delta r = 46mm$. The beam pipe radius is kept 3 mm smaller than r_1. The beam pipe thickness and the detector thickness are fixed.*

Figure 14a shows the variation of the resolution in the impact parameter b_z for different layer spacings with the radius of the beam pipe and inner layer (r_1=23mm) fixed. Except for the highest momenta, the impact parameter error and angular resolution are not very sensitive to the layer spacing. For a track of 1 GeV/c momentum a reduction in the layer spacing from 46mm to 23mm degrades the impact parameter (angular) resolution by 13% (10%). This is less than one would expect because again multiple scattering dominates the resolution.

6.4. Number of Layers

While one expects the resolution to improve with the number of measurements like $1/\sqrt{n}$, we observe that due to the presence of multiple scattering this improvement is only observable for the highest momenta. In Figure 15 the resolution for 1 GeV/c tracks is plotted as a function of the number of detector layers for different spacings between the layers, where the position of the inner layer and the detector

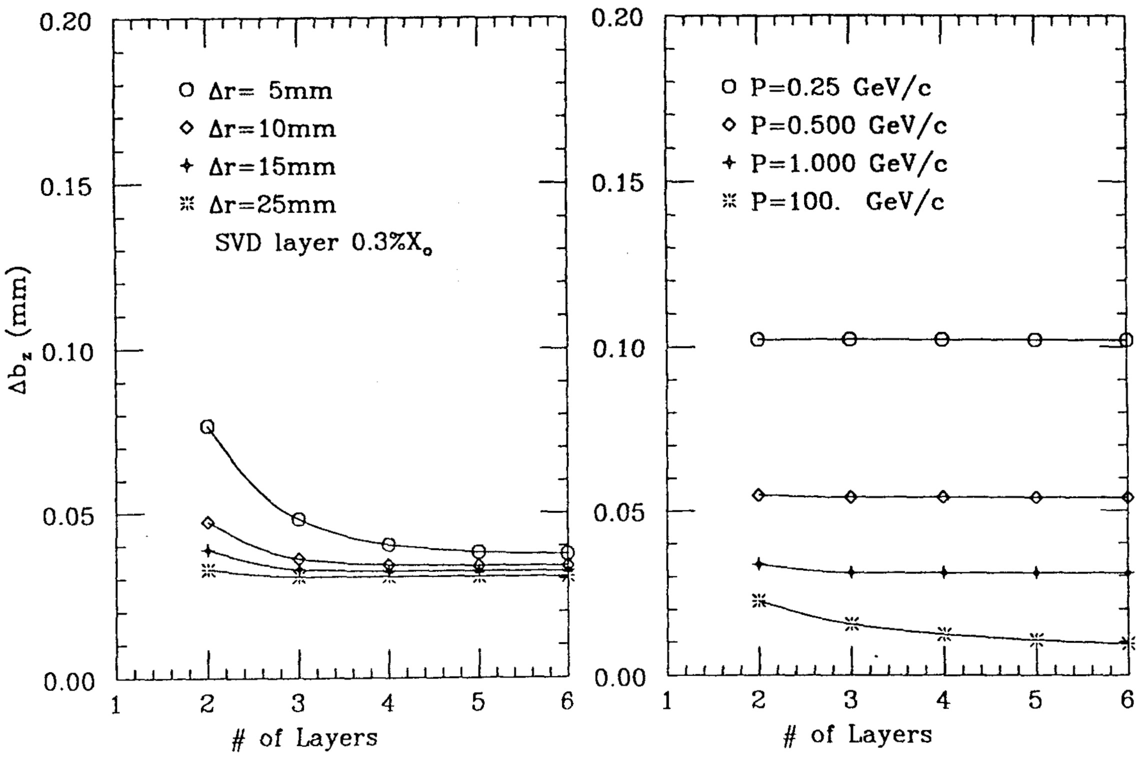


Figure 15. *Resolution in impact parameter b_z as a function of the number of silicon detector layers* **a.** *for different spacing between the layers and a fixed momentum of 1 GeV/c and* **b.** *for a fixed spacing of Δr = 23mm and different momenta. The curves are calculated for normal incidence, i.e.* $\cos\ \theta = 0$, *and detector of 300μm thickness.*

thickness are kept fixed.[24] Thus, for a given spacing the measurement error is determined by the number of measurements, the track length and the multiple scattering, for three layers and a spacing of 20-25mm the resolution is optimum at all momenta. On the other hand additional layers do not hurt, even at very low momentum. However, these calculations only include the effects of multiple scattering; energy loss, interactions or absorption, delta rays, etc. are not included and deserve separate consideration.

At least 3 detector layers are required, additional layers do not improve the track resolution.

6.5. *Beam pipe radius*

Figure 14b shows the resolution as a function of the radius of the inner layer r_1 (the inner beam pipe radius is taken to be 3mm smaller than r_1) for a fixed spacing of the two layers of $\Delta r = 46$mm. It is obvious that the impact parameter measurement is most sensitive to the radius of the innermost layer and the beam pipe inside. A reduction of the inner radius by 10mm would improve the impact parameter resolution from 30μm to 18μm. Furthermore, for $r_1 = 13$mm (23mm) tracks with momenta as low as 280 (510) MeV/c would have a resolution of better than 50μm in the $x-y$ plane, and 1 GeV/c tracks could be measured to this accuracy for $\cos\theta < 0.7 (< 0.54)$. Here the beam pipe thickness has been held constant, even though, from the point of vacuum safety a smaller beam pipe radius would allow for reduced thickness. However, the heat load per unit area increases.

The track resolution is most sensitive to the radius of the innermost detector layer.

6.6. *Intrinsic resolution*

The resolution in the impact parameter and angles is expected to depend linearly on the single point measurement error. However, as we see in Figure 16 and have noted previously, for a track of 1 GeV/c momentum, multiple scattering is dominant. For perpendicular incidence, the difference between a detector of $10\mu m$ resolution and a perfect detector corresponds a change in the impact parameter error from $25\mu m$ to $31\mu m$. Clearly, for a thicker beam pipe or larger angles the difference is even smaller. A $10\mu m$ intrinsic resolution appears to be a good match to the layout scheme described here.

A $10 - 15\mu m$ intrinsic position resolution is a good match to overall limitations of a silicon detector system.

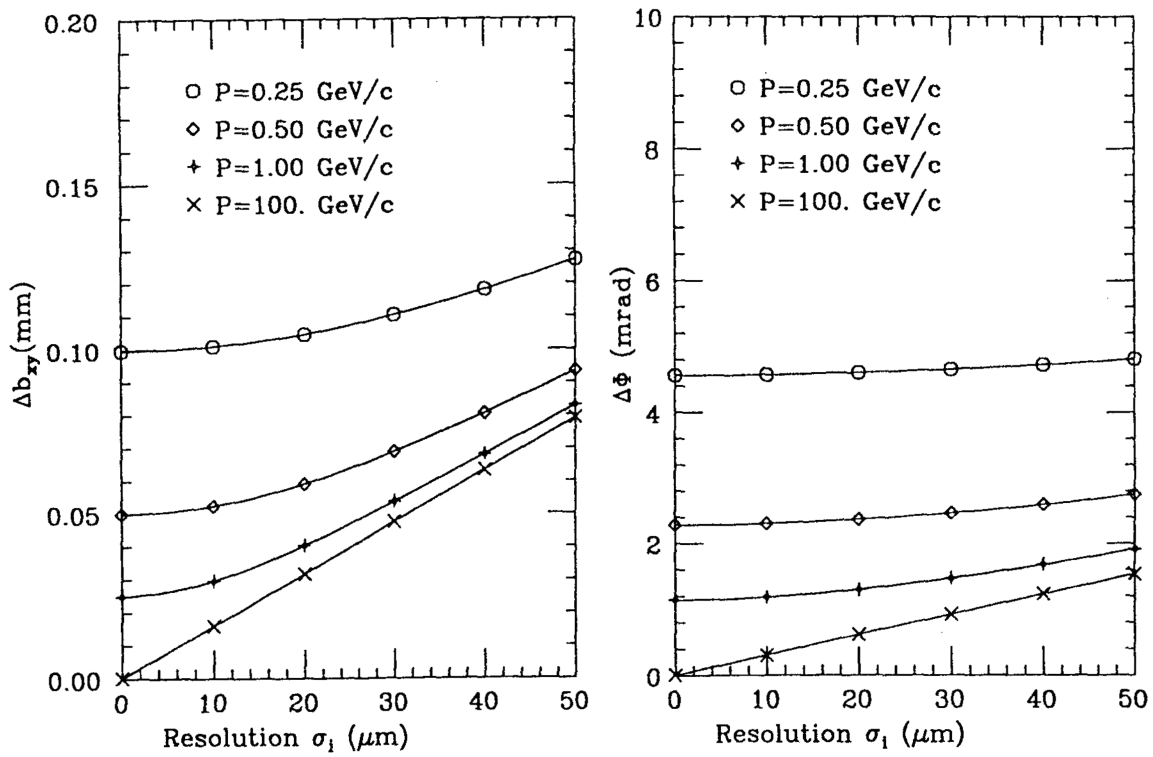


Figure 16. *Resolution in*
a. b_{xy} *and*
b. ϕ *as a function of detector resolution* σ_i *for tracks of different momentum, where* $\Delta b_{xy} = \Delta b_z$ *and* $\Delta\phi = \Delta\theta$.

6.7. Conclusions

The performance limits of a multi-layer silicon vertex detector are not set by its intrinsic resolution, but by multiple scattering in the material of the vacuum pipe, the detector support structure, and the detectors themselves. The most important parameter is the distance of the innermost detector from the beam and this will be determined by the design of the beam and the machine interaction region including the shielding to protect against radiation background. Based on the present lay-out of the IR region with a vacuum pipe made of 1mm of beryllium with an inner radius of 20mm, a silicon vertex detector can measure impact parameters of charged particles of 1 GeV/c momentum to an accuracy of 50μm or better over a large fraction of the solid angle. The angular resolution is of the order of 2 mrad or better in azimuth and polar angle. This can be achieved with double-sided strip read-out or pixel devices of 50μm strip pitch or pixel size. The development of low power, low noise, high density amplifiers and read-out circuits needs to be pursued so as to fit the detectors into the limited space without loss of solid angle coverage, and without a substantial increase in multiple scattering. Precision assembly and alignment techniques need to be studied.

7. Hybrid Vertex Detectors

Do we need an intermediate tracking chamber?

In almost all applications of silicon vertex detectors at SLC, LEP and the Tevatron a precision wire drift chamber or TPC was added to fill the radial space up to the large central tracking chamber (CDC). In most cases the intermediate tracking devices were added as a back-up for the then new silicon technology or the silicon devices were added to upgrade the performance of existing vertex detectors. Now that silicon strip detectors have proven to operate well at e^+e^- storage rings, one should question the need and purpose of an intermediate tracking device. In discussions which drew from the experience at ARGUS, CLEO and Mark II, and from previous workshops at SIN (J. Chauveau) and Syracuse (D. MacFarlane) it was concluded that there are several aspects of tracking and triggering that would benefit from an intermediate tracking chamber - intermediate in resolution and intermediate in radial position relative to the SVD on the inside and the CDC on the outside. These benefits are:

- an improvement in the acceptance and trackfinding ability for particles of low transverse momentum,
- a significant improvement in the measurement of the track angle and position for particles decaying outside the SVD, *e.g.* K^0_S,
- an improvement in the z and θ measurement for particles emitted at small angles to the beam,

— an improvement in the matching of track segments measured in the SVD and the CDC.

— an improvement in the rejection of beam-generated background especially in the $r - z$ plane, and

— an improvement in the efficiency of triggering on charged particles produced at small angles to the beam.

For the equivalent thickness of a single layer of silicon a drift chamber with 10-15 layers could be installed.

Probably none of these reasons is strong enough on its own to justify the construction of an additional device, but it was felt that all six reasons together make a much stronger case and certainly justify further study. While the first two items on the list have been addressed by the physics analysis group, the improvements in track resolution were studied for an intermediate vertex drift chamber as a function of the wires stereo angle. Very little information is presently available on triggering and pattern recognition. Such studies require a very detailed Monte Carlo generator and a track reconstruction program, both of which are not available at this time.

7.1. *Measurement of low p_t tracks*

Probably the most important physics argument for an intermediate tracking chamber is the gain in acceptance and resolution for low momentum tracks, driven principally by the need to detect the slow spectator pion in the decay $D^{*+} \to D^0\pi^+$.

Figure 17 gives the transverse momentum spectrum of low energy pions from the decay of the D^{*+} in the semi-leptonic decay $B^0 \to D^{*+}l^-\nu$. Also shown is $1 - \epsilon$, the fraction of decays with a transverse momentum below a cut-off value p_t. Of all spectator pions, 14% are below 50 MeV/c. Taking into account that pions of less than 38 MeV/c total momentum range out in the beampipe and first silicon layer, a cut at $p_t = 50$ MeV/c does not cause a large additional loss. In a magnetic field of 1 T, particles with 50 MeV/c transverse momentum curl at a distance $2\rho(\text{cm}) = 0.667p(\text{MeV/c})/B(\text{T}) = 33$ cm in the tracking chamber. Thus in a CDC with jet cells, such a track will traverse two radial and two stereo cells and consequently should be measurable in both the SVD and CDC. On the other hand, additional measurements in a VDC would enhance the track finding efficiency of these slow tracks.

An intermediate tracking chamber would increase the efficiency for low p_t tracks.

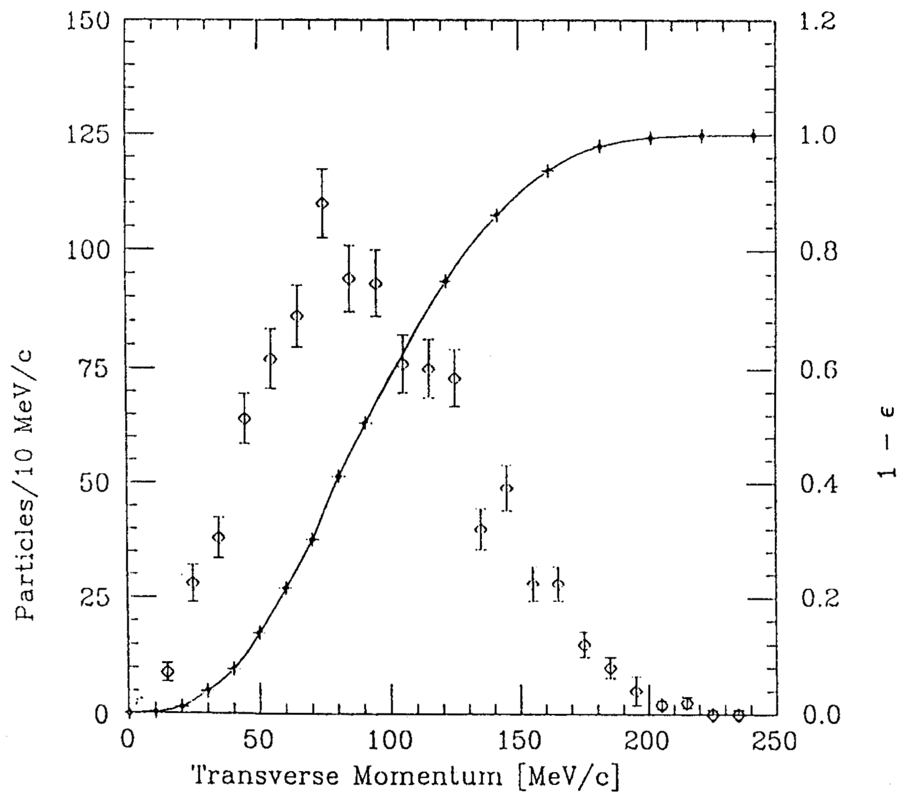


Figure 17. *Transverse momentum spectrum of charged pions from the decay $D^{*+} \rightarrow D^0\pi^-$ for $|\cos\theta| < 0.95$. The distribution is normalised to a total of 1000 decays. The curve gives the fraction of pions with transverse momentum below the cut-off value p_t.*

7.2. Detection of K_S^0 decays

Since many decays of B mesons produce kaons via the cascade to charm mesons, the detection of secondary vertices from charged and neutral kaons is important. The distribution of transverse decay lengths R_{xy} of K_S^0's from the decay $B_d^0 \rightarrow \psi K_S^0$ is given in Figure 18, both in differential form as the number of decays per mm, and in integral form.[25] We observe that 22% of all K_S^0 decay beyond the outer layer of the proposed silicon detector at 8 cm. For a track of 1 GeV/c momentum that is measured only in the CDC, the resolution in the $x - y$ plane would degrade by a factor of 6 in impact parameter and 1.3 in azimuth angle ϕ, and in the $r - z$ plane by a factor of 30 in impact parameter and 3 in polar angle θ. On the average the tracks of the decay pions would have an impact parameter resolution Δb_z in excess of 1mm, their angular resolution would of the order of 2mrad, their momentum error $\Delta p/p^2$ in excess of 0.7%. Present Monte Carlo studies do not take these effect into account, but it is expected that with appropriate resolution-dependent cuts on the mass and angle of the $\pi^+\pi^-$ pair, efficiency losses can be held to a few percent. Background rejection may, however, get more difficult, in particular for decays such as semileptonic modes, which include a missing neutrino.

22% of all K_S^0 from the decay $B_d^0 \to \psi K_S^0$ decay outside a radius of 8cm.

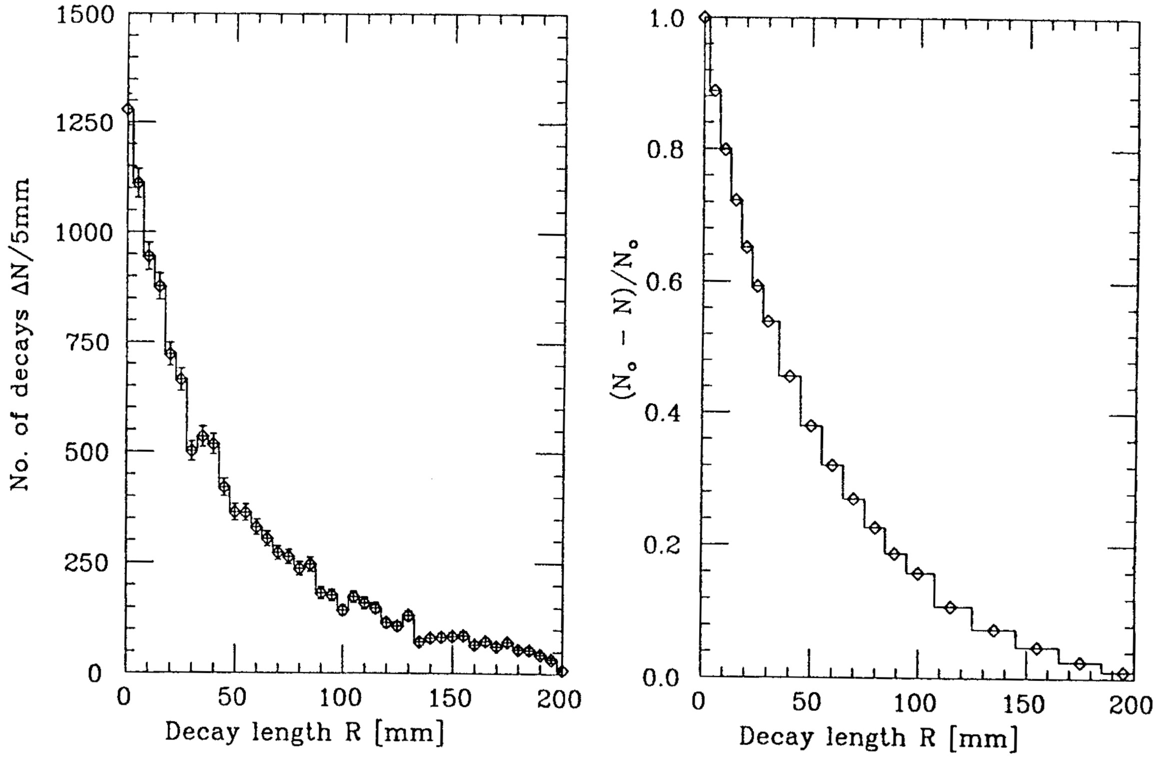


Figure 18. *The differential and integral distribution of decay distances of K_S^0 from the decay $B_d^0 \to \psi K_S^0$. measured in the plane transverse to the beam.*

7.3. Layout of a Vertex Drift Chamber

The proposed Silicon Vertex Detector (SVD) extends from the beam vacuum pipe to an outer radius of about 7 – 8cm, and along the beam direction to about ±15 cm. Thus there is a radial gap up to the Central Drift Chamber (CDC) of about 10 cm that could accommodate another tracking chamber. Figure 19 shows a schematic layout of such a chamber, which could cover a polar angle range up to $|\cos\theta| = 0.97$. While there is an additional radial gap of 2 cm on the outside, the placement of the first bending magnet at $z = 20$ cm may be somewhat difficult, in particular if one takes into account the additional space needed for endplates, preamplifiers, cables, *A vertex drift chamber with full polar angle coverage is not compatible with the placement of dipole magnets at a distance of 20cm from the IP.*

gas lines, *etc.* While the inner wall can probably kept very thin, the mechanical rigidity has to be provided by the end-plate and the outer wall (4mm of C or Be). Figure 19 shows only the forward half of the VDC, it may not be necessary to extend the chamber to the same angle in the backward direction. Also, most of the connections and supplies could be located on that side.

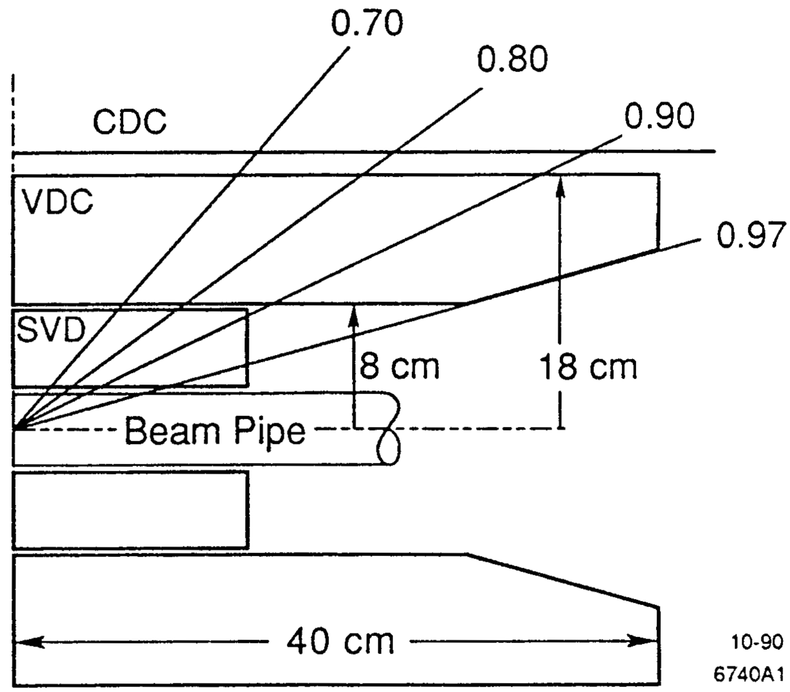


Figure 19. *Layout of a VDC between the SVD on the inside and the CDC on the outside.*

7.4. *Resolution of a Hybrid Vertex Detector*

Resolution for the track parameters $\alpha_i = (1/p, b_{xy}, b_z, \phi, \theta)$, for momentum p, impact parameter b_{xy} and b_z, and angles ϕ and θ, has been studied for individual charged particles measured in a combination of tracking detectors.[26] The central drift chamber (CDC) has 39 layers, 4° stereo, and 150μm position resolution, and extends from 20cm to 80cm in radius. The silicon detector (SVD) is assumed to have 3 layers, at radii of 23, 46, and 69 mm radius, 10μm resolution in both xy and z. The vertex drift chamber (VDC), with 10 layers and stereo layers of $\pm 45°$ alternating, has a position resolution of 50μm. The CDC is filled with a gas of 450m radiation length, the VDC has a denser gas of 120m radiation length. The outer (inner) wall of the VDC is assumed to have a thickness of 2% (0.3%) of a radiation length. All other detector

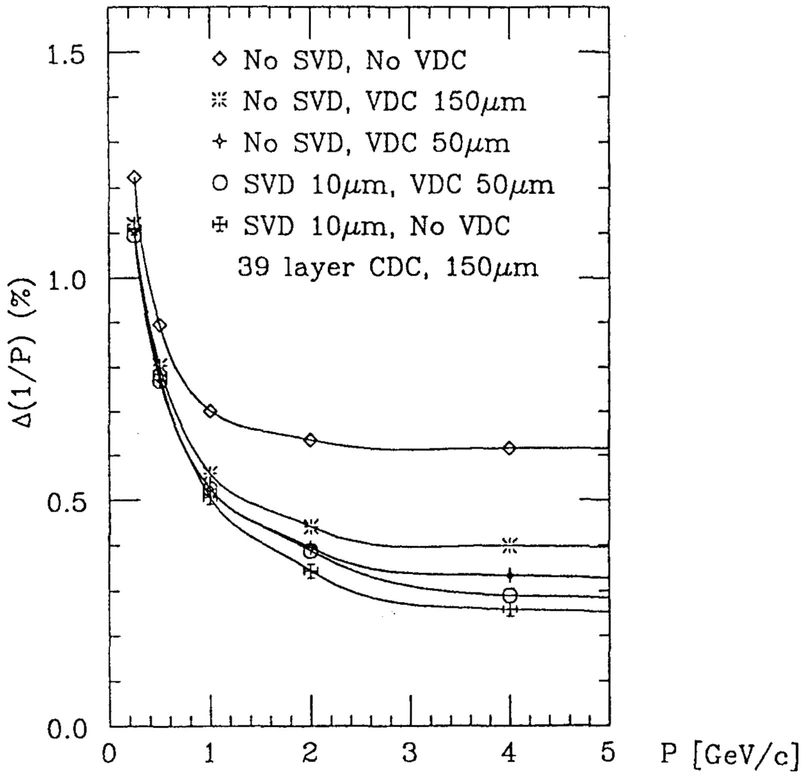


Figure 20. *Resolution in track momentum for charged particles measured in either the CDC alone, the CDC and VDC, or the CDC and the silicon vertex detector (SVD).*

walls are less than $0.3\% X_o$. Figures 20 , 21 and 22 show the errors in the track parameters measured in a combination of tracking devices. The error on a track parameter α is usually expressed as a sum of two terms that are to be added in quadrature:

$$\Delta\alpha = A_\alpha \oplus C_\alpha / p[GeV/c].$$

The first term A_α describes the geometric resolution and is a function of most of the following parameters, the intrinsic resolution per layer, the number of layers, the stereo angle of the layers, the total length of the tracking device, and the polar angle of the track. In particular, $A_{1/p} \sim L^{-2}$, where L is the track length. The second term represents the contribution from multiple scattering, and is dominated by the amount of material in the beam pipe and the detector walls. It depends critically on the polar angle, i.e. $C_\alpha \sim \sin^{-n/2}\theta$, where n=5 for C_z (impact parameter b_z), n=3 for C_{xy} (impact parameter b_{xy}) and C_ϕ (azimuth angle ϕ), and n=1 for C_θ (polar angle θ). Table 4 lists the parameters A_α for combinations of different tracking devices. The last row gives the multiple scattering terms C_α for the CDC combined with the SVD.

Device	measurement	stereo angle	radial length	$A_{1/p}$ [%]	A_ϕ [mrad]	A_{xy} [μm]	A_θ [mrad]	A_z [μm]
CDC	39 $\times 150\mu$m	$0, \pm 4°$	60 cm	0.60	0.91	206	3.52	1744
CDC VDC	39 $\times 150\mu$m 10 $\times 50\mu$m	$0, \pm 4°$ $\pm 45°$	70 cm	0.31	0.36	46	0.99	124
CDC VDC SVD	39 $\times 150\mu$m 10 $\times 50\mu$m 6 $\times 10\mu$m	$0, \pm 4°$ $\pm 45°$ $0, 90°$	76 cm	0.18	0.18	10	0.24	13
CDC SVD	39 $\times 150\mu$m 6 $\times 10\mu$m	$0, \pm 4°$ $0, 90°$	76 cm	0.20	0.20	11	0.30	15
m.sc.			76 cm	0.46	1.21	29	1.20	28

Table 4. *Charged particle track resolution for combinations of different devices: the CDC (Central Drift Chamber), the VDC (Vertex Drift Chamber), and the SVD (Silicon Vertex Detector). The last line lists the contributions from multiple scattering for a track of 1* GeV/c *and* $\cos\theta = 0$.

The measurement of the track momentum p and angle ϕ relies primarily on the CDC. The addition of the VDC with 50μm resolution improves the momentum resolution $\Delta p/p^2$ from 0.6% to 0.3%, half of this improvement is due to the extension of the track length from 60 cm to 70 cm. The insertion of the silicon vertex detector further improves the momentum resolution to 0.2%, it extends the track length to 76 cm. The angular resolution, ~1.2mrad/p[GeV/c], is dominated by multiple scattering. The most dramatic improvement achieved by the addition of the vertex detectors is the resolution in the polar angle θ and the impact parameter b_z. This is due to the much larger stereo angles assumed for the VDC. The VDC in addition to the SVD adds more multiple scattering and does not enhance the resolution in the relevant momentum range below 5 GeV/c. In the study on track resolution, we have assumed that the VDC has 10 layers of signal wires of alternating $+45°$ and $-45°$ stereo angle. Such a chamber was built and installed by the ARGUS group and has achieved a position resolution of better than 40 μm over half the drift cell of 5.3 mm width.[27] A very large stereo angle requires an unconventional wire suspension technique.

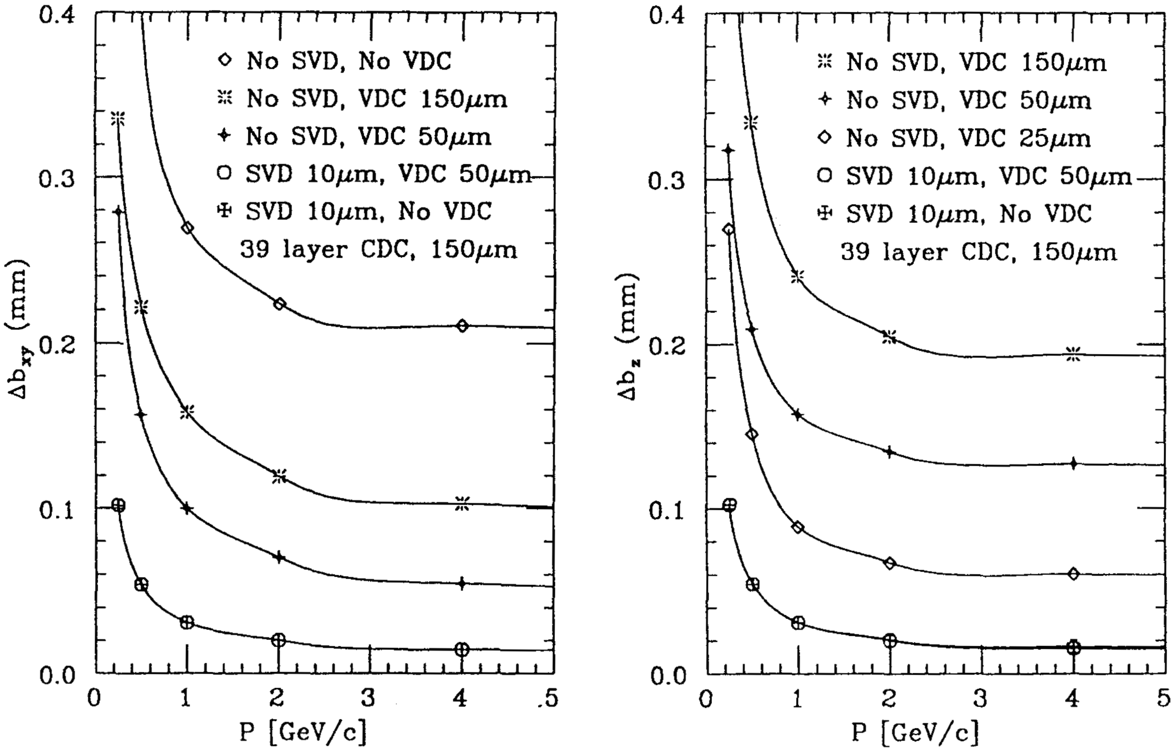


Figure 21. *Resolution in impact parameter for charged particles measured in either the CDC alone, the CDC and VDC, or the CDC and the silicon vertex detector (SVD).*

We have studied the resolution of a tracking system consisting of the standard CDC with $\pm 4°$ stereo angles in 18 of 39 layers, and a 10 layer VDC described above. If the information from the SVD is incomplete, then the VDC noticeably improves the track measurement. In this case, the resolution in the impact parameter b_z and the polar angle θ depends critically on the stereo angle of the wires in the VDC.[26] Figure 23 shows the resolution in impact parameter b_z and polar angle θ for different sequences of stereo layers with angles $\pm\mu$ in the VDC. For a track measured only in the VDC with alternating stereo layers the error in the impact parameter varies as $\Delta b_z = 86\mu\text{m}/\sin\mu$ and the polar angle error is $\Delta\theta = 0.7\text{mrad}/\sin\mu$. If the stereo layers are grouped such that the first 5 layers are placed at an angle $+\mu$ and the outer 5 layers have a stereo angle $-\mu$ the resolution degrades by a factor of two. In conclusion, if one wanted to reconstruct track elements in the VDC independent of the other tracking devices, it would be necessary to have wires placed at two relatively large stereo angles, *i.e.* $|\mu| \geq 0.2$mrad. If we

In the absence of a good measurement in the SVD, the addition of a VDC improves the track resolution.

combine the VDC and the CDC measurements, the dependence on the VDC stereo angle is moderated. For a typical stereo angle of $\mu = \pm 0.1$rad, the impact parameter resolution is $\Delta b_z = 0.3$ mm compared to $\Delta b_z = 1.75$ mm for the CDC

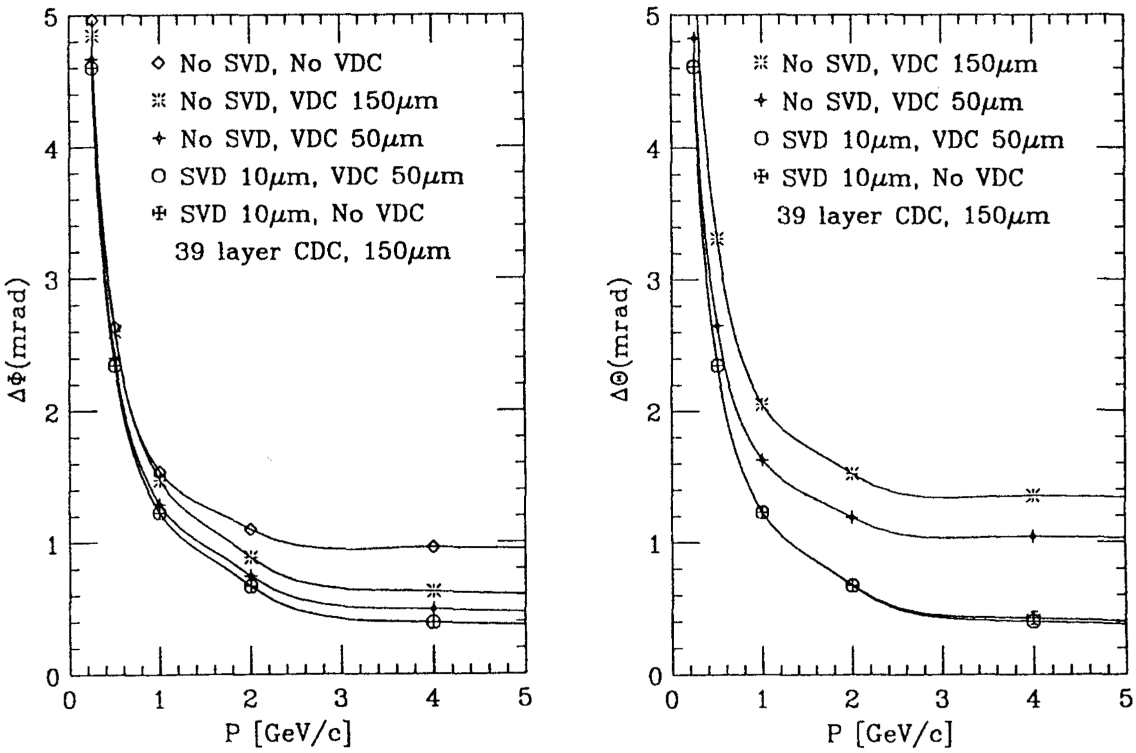


Figure 22. *Resolution in angles for charged particles measured in either the CDC alone, the CDC and VDC, or the CDC and the silicon vertex detector (SVD).*

and a VDC with axial layers or $\Delta b_z = 0.86$ mm for a measurement with a VDC alone. An improvement that will clearly help in the association of CDC and SVD track elements. Alternating layers with stereo angle $\pm\mu$ require no more radial space than a chamber with all layers of the same orientation $+\mu$. The change in the radial position of a stereo wire over its length L at a radius r_i is given by the simple geometric relation: $\Delta r = r_i(1 - \cos\alpha)$ with $\tan\alpha = L\tan\mu/2r_i$. If one adjusts the wire length L_i as a function of its placement radius r_i by a conical endplate then Δr can be held to roughly 6mm for all radii r_i.

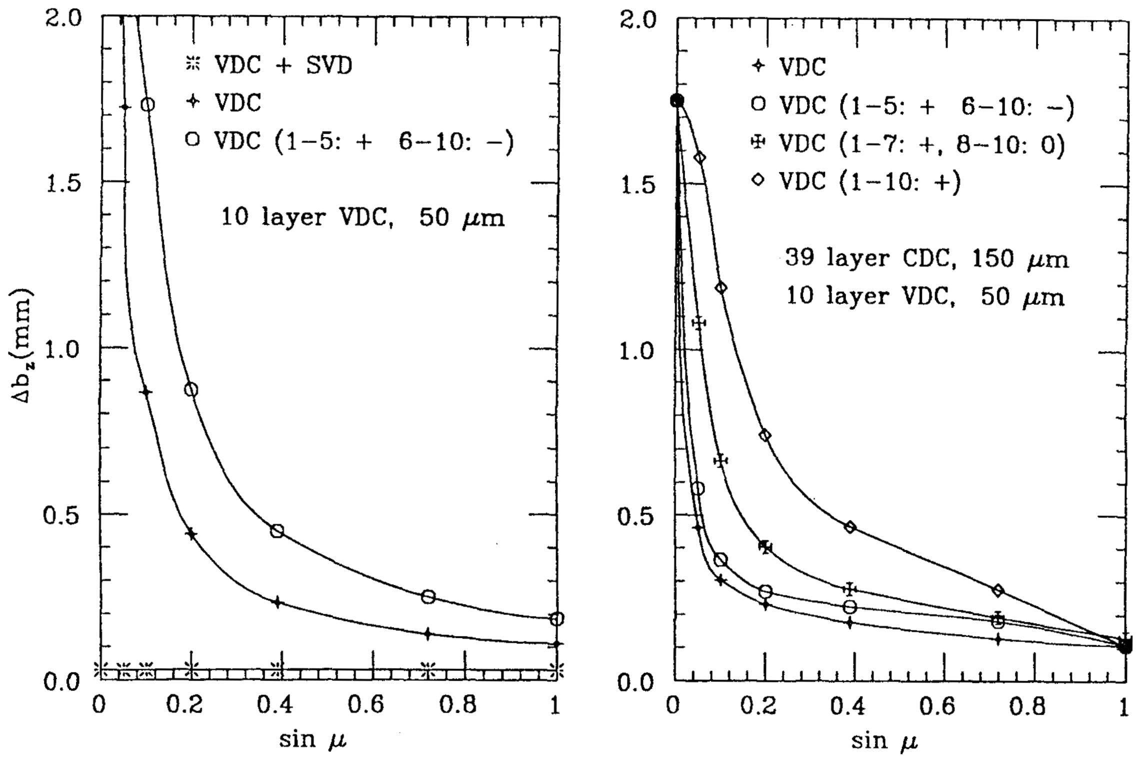


Figure 23. *Resolution in the impact parameter b_z and polar angle θ for tracks of 1 GeV/c momentum and cos θ=0 as a function of the stereo angle μ of the wires in the VDC,* **a.** *for the VDC measurement alone, and* **b.** *for a combined tracking system including the CDC. The different curves refer to different sequences of stereo angles ($\pm\mu$) in the 10 layers.*

7.5. Forward Time Projection Chamber

At small polar angles, track reconstruction may become difficult due to a reduced number of hits and the increase in multiple scattering in the silicon. One approach to this problem is to replace the forward silicon vertex detector by a gas filled time projection chamber (TPC).[28]

The proposed layout is presented in Figure 24. The TPC forms an annulus that extends from $z_{min} = 3$ cm to $z_{max} = 23$ cm, has an inner radius $r_{min} = 2.5$ cm and an outer radius $r_{max} = 12.5$ cm. The TPC is surrounded by a single layer of silicon strip detectors which are assumed to have a position resolution of 10 μm. The TPC is operated at atmospheric pressure. The internal position resolution is given by a sum of three terms added in quadrature, (a) the error due to the longitudinal diffusion $\sigma_d = C_d\sqrt{L}$, with diffusion coefficient of $C_d = 3\text{mm}/\sqrt{\text{m}}$, (b) the contribution from the electronics which is assumed to be roughly 100μm for 100MHz

FADC, and (c) a contribution due to the gap between the wires and the endplate which is assumed to be 1mm wide. The parameters quoted here are somewhat optimistic extrapolations based on the experience with the DELPHI detector at CERN, and would result in a point resolution of approximately $\sigma_z = \sigma_{xy} = 200 - 300\mu$m. There are 10 z measurements per cm radial distance. The entrance windows to the chamber are kept to a thickness of 0.1% of a radiation length.

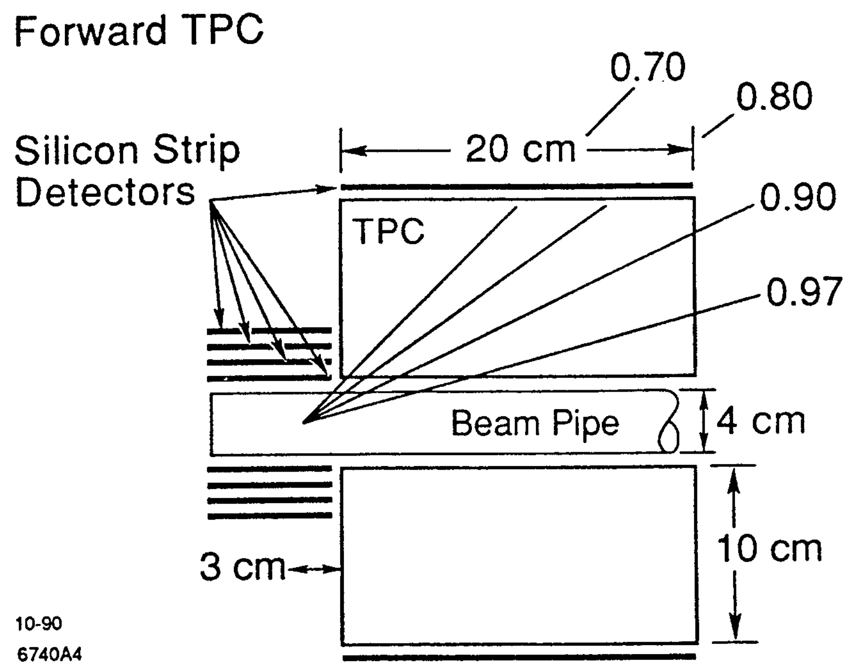


Figure 24. *Layout of a vertex detector with a central silicon barrel and a forward TPC*

Figure 25 compares the resolution of different detectors covering the forward direction. The resolution is improved by $15-20\mu$m by the silicon strip detectors on the outside, though this layer covers only up to polar angles of $\cos\theta = 0.80$. At larger angles the resolution would be improved if the length of the chamber could be extended beyond $z = 23$ cm. Also, a thinner entrance window, a higher electron yield and smaller diffusion would improve the resolution. Beyond $\cos\theta = 0.90$ the TPC performance is comparable to a three-layer silicon strip device of purely cylindrical layout, with radii of 23, 46, and 69 mm as presented above. Substantially better impact parameter and angular resolution can be obtained with a silicon strip detector with endcaps of either disc or conical shape.[24] In these lay-outs the resolution is totally dominated by the multiple scattering in the beam pipe, while the contribution from the detectors is reduced and largely independent of the angle θ.

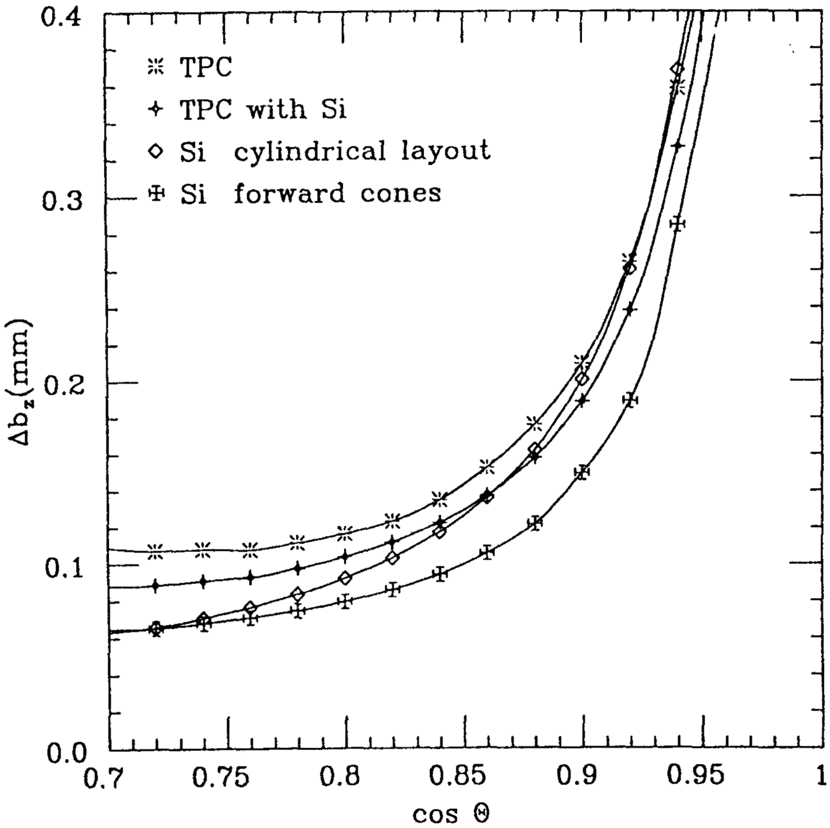


Figure 25. *Comparison of the resolution in the impact parameter b_z at small angles θ relative to the beam for a variety of detector options.*

7.6. Forward Microstrip Avalanche Chamber

Another device that could improve the tracking efficiency and resolution in the forward direction, is a forward avalanche drift chamber as described by J. Va'vra.[29] Its basic layout is shown in Figure 26. Such a device is designed to complement the central silicon vertex detector in the forward angular range from $\cos\theta = 0.85$ to 0.97. It potentially could measure $50 - 100$ points per track with a resolution of $40 - 50\mu$m per point. The device operates in proportional mode and collects and amplifies charge on micro-strips that are vacuum evaporated onto a smooth cylindrical insulator surface.[30] This relatively novel idea needs considerable R&D, engineering and testing. As for the TPC, one expects that it is the multiple scattering in the beam pipe and not the intrinsic resolution that determines the track resolution in the forward direction.

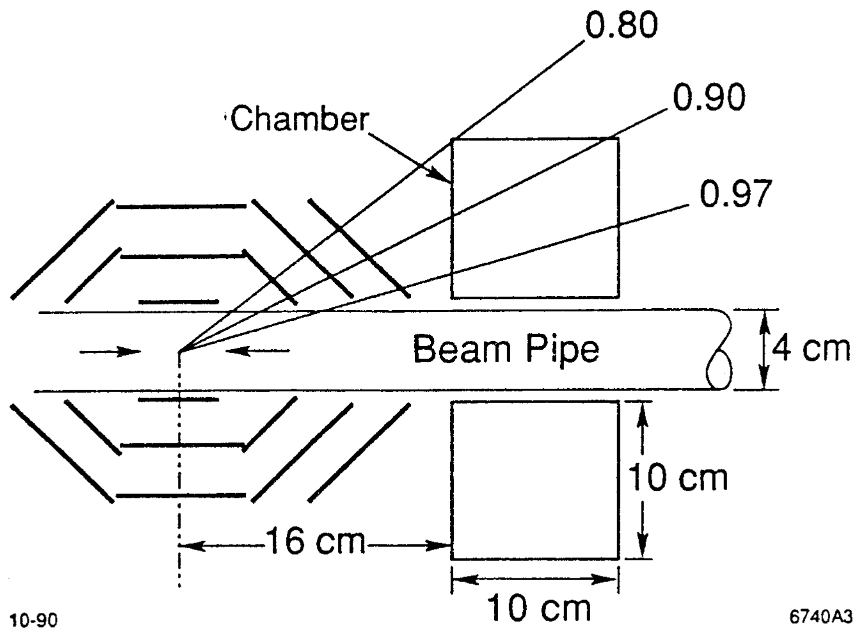


Figure 26. *Schematic layout of a micro-strip avalanche chamber that complements the central silicon vertex detector in the extreme forward direction.*

7.7. *Conclusions*

The selection of the most suitable device will depend on how important the detection of tracks in the forward direction will be for the physics analysis. The VDC as well as the proposed forward devices are not compatible with the present layout of the interaction region, which places the first dipole at a distance of 20cm from the interaction point. Also, background simulations show that the rate of photons and electrons from lost beam particles is of the order of $550\,\gamma/\mu s$ and $65e^{\pm}/\mu s$ in the volume of the TPC.

8. Design of the Vacuum Pipe

The performance of a precision vertex detector depends critically on the design of the vacuum pipe. A vacuum pipe must be designed not only to have the mechanical strength to support the vacuum load; it must also maintain mechanical and thermal stability in the presence of the expected heat load due to ohmic losses and higher order mode rf losses generated by the fields of the beams. Furthermore, it must shield the detectors and their electronics from synchrotron radiation photons and the rf fields of the circulating particle bunches. Compared to the relatively simple metallic structures that are presently in use at storage rings, the B Factory places much higher demands on the cooling and shielding because of the considerably higher beam currents.

From the point of view of the vertex detector that is located immediately outside the vacuum pipe and might need to be supported on the vacuum pipe, it is important to

- minimize its thickness in terms of radiation length,
- minimize its total radial thickness, and
- control its temperature and avoid large temperature gradients on its surface.

8.1. *Specifications*

The dimensions of the pipe are given by the constraints imposed by the design of the beam interaction region. As a working number we have assumed a value of $r_i = 20$ mm for the inner radius and a length of $L = 40$ cm. A maximum radiation length equivalent of 1.5 mm of beryllium is allowed for all material placed between the collision point and the first detector layer. A smaller value is preferred, and most of the detector performance studies have used the equivalent of 1 mm Be. Also, the inside wall of the vacuum pipe is usually coated with a few μm of a heavy metal like copper or titanium to absorb soft synchotron photons. It is desirable to have the material distributed as uniformly as possible to maintain a radial and axial symmetry for the whole detector. The radial thickness of the vacuum pipe should be held to a few mm, we have assumed that 3 mm would be sufficient.

It is expected that engineering studies will provide practical constraints for such parameters as temperature stability during operation at variable heat loads, maximum allowable gradients, maximum velocities and pressures of the coolant, etc. No specifications have been set for the maximum and minimum temperature of the surface. Ideally, one would like to keep the absolute temperature close to $20°C$, since this is the most suitable temperature for the assembly of detector components and would thus keep temperature effects on the stability of the detector elements to a minimum. Likewise temperature gradients should be minimized, though temperature differences of up to $10°C$ may have to be accommodated.

Several studies have been pursued to attain a working model for a beam pipe that can take a heat load of up to 1kW. With this level of power absorbed in the pipe heat removal at the pipe ends is not practical, given the temperature gradients specified here.

8.2. The JPL design

A generic investigation of beam pipe cooling alternatives was carried out at JPL.[31] and engineered further by A. Lisin of SLAC.[32] (see Figure 27). The design is largely based on the specification listed above. In this design the vacuum pipe is made of a double walled beryllium tube with a coolant pumped from the ends through the gap in between. A 0.5 mm inner wall would hold a pressure of up to 10 atm, the outer wall could be substantially thinner and hold the same pressure. The wall thickness t depends linearly on the radius of the pipe and varies with pressure as $P^{1/3}$. The maintenance of the gap over the length of the tube is obviously a concern, a beryllium wire of diameter equal to the gap wrapped around the outside of the inner tube could serve as a spacer and might also separate coolants flowing in opposite direction. The major problem is the temperature differential between the inner and outer tube. It will be necessary to support the tube on bellows joints to accommodate the thermal expansion.

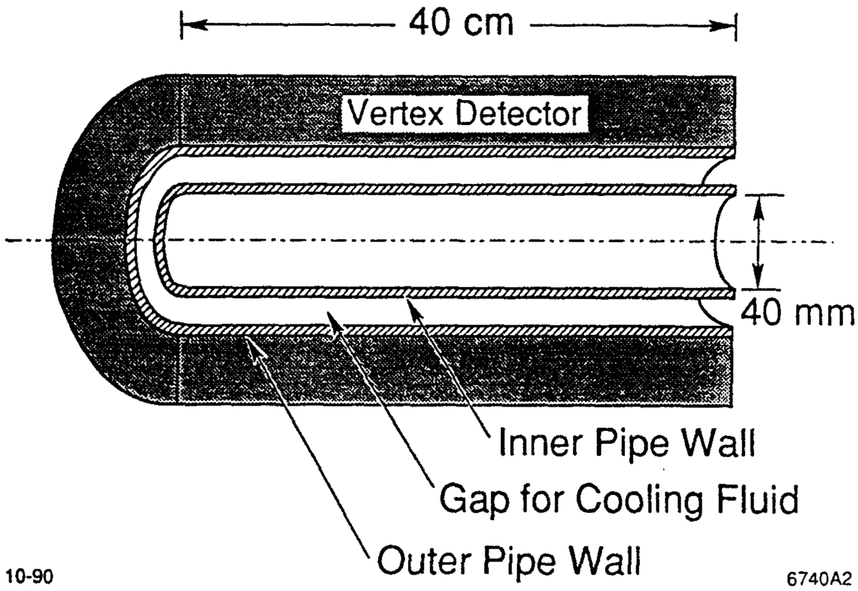


Figure 27. *A sketch showing the key features and dimensions of the vacuum pipe designed by JPL.*

The two methods recommended in the JPL study are water cooling and high-pressure helium or hydrogen cooling. Water cooling is the preferred method, because a water cooling circuit can be designed that meets the temperature stability and gradient requirements for heat loads of several kW. A gap of 0.5 mm is sufficient. Such a design represents a proven and simple approach and while avoiding excessive flow or pressure gradients it provides excellent growth potential to higher heat loads and larger pipe diameters. A potential problem is the corrosion of the beryllium surfaces by the water coolant; a very thin coating of gold and the use of very high purity de-ionized water could avoid this problem.

A double-walled Be pipe with water or pressurized gas as a coolant can take heat loads of more than one kilowatt.

A gas such as helium or hydrogen at 5 to 15 atm. pressure is an alternate choice. Helium reduces the amount of material substantially and it obviously avoids corrosion problems provided the gas is of ultra-high purity. For pressures of 5 atm. or less the limiting parameter is the gas flow rate, up to 500m/s. Such high velocities are not practical because they reduce the mass flow rate due to compressibility effects. Also the radial gap has to be of the order of 5 – 10mm, *i.e.*, a factor of 10 larger than for water, to accommodate the required mass flow.

8.3. Expansion Cooling

While the JPL study provides us with a conservative design based on proven technology several discussions have focussed on novel ideas on how to reduce the total amount of material between the beam and the vertex detector, to increase the conductivity, and thereby reduce the source of heating, and to lower the skin depth and improve the rf shielding capability of the beam pipe. R. Erbacher and W. Vernon[33] have suggested filling the gap between the two thin metallic surfaces with an extrusion of a porous ceramic called cordierite. Nitrogen gas would be entered at high pressure (20 atm) at the inner radius and undergo rapid expansion and thereby undergo a temperature change of roughly $-0.5°C/\mathrm{atm}$. While a few laboratory tests are underway, many technical problems, like the containment of the vacuum and the gas coolant, the maximum pressure load sustainable by the proposed structure, the machinability of the material, the attachment of extremely thin metallic skins, etc. remain unspecified and unsolved.

8.4. Cryogenic Beampipe

The total heat load due to the absorption of higher order rf modes and ohmic heating could be substantially reduced in a cryogenically cooled vacuum pipe.[14] At lower temperatures the resistivity decreases and substantially reduces the skin depth. Thus the metallic walls could be replaced by thin foils that provide sufficient skin depth and serve as a vacuum and gas seal. The mechanical strength would be provided by a layer of boron carbide (B_4C) or silicon carbide (SiC) foam[34] which fills the gap between the metal foils. B_4C is a material known for its extremely high specific stiffness:

$$E/\rho = 448.2[GPa]/2.52[g/\mathrm{cm}^3] = 178\,.$$

B_4C is machinable: the support structure for individual vertex detector modules could be built of foam of less than 5% of the nominal density. SiC has a thermal coefficient of expansion that is similar to silicon. The coolant, cold helium or hydrogen gas, would be circulated through the pores of the foam. The idea that fosters this scheme is that solid state detectors and circuits perform at substantially lower noise levels at low temperatures. However, this advantage may be offset by the difficulty of building a detector hybrid and support structure that can survive temperature cycles of more than $100°C$.[35] Experience in infrared astronomy with large arrays of solid state detectors may be useful for this approach.

8.5. Conclusions

In summary, substantial engineering and R&D would be required to allow for an evaluation of either of the two ideas for composite beam pipe structures and cooling. A double walled beryllium pipe with water cooling at room temperature represents a solution based on well understood engineering principles.

9. Effects of Beam Background

An issue of primary importance for the operation of vertex detectors is the rate of background that can be tolerated. As outlined in the chapter on machine backgrounds, there are two major sources of background:

- synchrotron radiation from dipole and quadrupole magnets near the interactions point, and
- photons and beam particles created by bremsstrahlung or Coulomb scattering on the residual gas upstream of the interaction region. These sources produce photons and electrons that interact in the beam pipe or the surrounding beam line components or detector elements.

Two upper limits are relevant for the operation of a silicon detector in the presence of these backgrounds: one is set by the maximum acceptable occupancy of the detector read-out channels, and the other by the maximum radiation dose to the detector and the associated VLSI electronics.

9.1. Radiation Damage in Silicon

Radiation damage effects are caused by displacement of atoms in the crystal lattice and the accumulation of charges in and on the surface of insulating layers.

There are two basic mechanisms that are responsible for radiation damage in silicon detectors and VLSI circuitry:

- displacement of atoms from the lattice (so-called bulk damage), and
- electron-hole pair creation in the insulator layers by ionizing radiation (so-called surface damage).

VLSI read-out circuits are most susceptable to surface damage.

The most important surface damage effects are the build-up of fixed, positive charges in the oxide and the generation of trapped negative charges at the $\mathrm{Si} - \mathrm{SiO_2}$ interface. These are the principal causes of radiation damage in MOS devices. Silicon detectors are relatively insensitive to oxide surface charges because the ionization charge is transported normal to the surface. However, the increase in surface charge can cause strip interconnection, increase the leakage current and lower breakdown voltages. Passive components such as punchthrough and accumulation resistors will also be affected.

Presently available silicon detectors are radiation hard to doses of several MRad.

Bulk damage arises mainly from the displacement of atoms from their lattice site due to collisions with high energy charged particles or neutrons. This can lead to an increase in leakage current and the creation of charge defects. In high purity silicon most of the leakage current originates from excitations of electrons or holes across the bandgap. Such transitions are very sensitive to the presence of energy levels within the band gap. Displacement damage creates new levels and thereby enhances the leakage current. Some of these levels also act as traps for mobile charge carriers in the silicon. Since the field inside a depleted n-type silicon detector is generated by positive charges of the dopant, the dominantly negatively charged defect structures produced by irradiation tend to compensate these doping charges. Such effects are expected to change the charge collection time, worsen the signal-to-noise ratio, and could also degrade the position resolution, in particular in the cases of non-uniform irradiation. Changes in the effective carrier density lead to an increase in the bulk resistivity and change the depletion thickness. Extensive studies are presently under way by groups preparing large silicon strip tracking systems for the SSC.[36]

Photons and electrons contribute primarily to the build-up of interface charges, a process that depends critically on the details of the device processing and geometry, in particular, the effects increase quadratically with the thickness of the oxide or insulating layers. VLSI circuits are more sensitive to surface damage, one observes threshold shifts in transistors and increased noise in analog circuits.[37] Amplifier and read-out circuits that are presently in operation have not been specifically designed to tolerate large amounts of ionizing radiation. They fail at doses of 50 kRad or less. Much higher radiation tolerance is expected from specially designed CMOS or JFET circuits.

Extensive studies of radiation effects are being pursued in preparation for SSC experiments.

Silicon detectors have operated in particle beams up to total doses of several MRad, and for non-uniform irradiation there is clear evidence for field distortions leading to degradation of in the position resolution.[38]

Increases in detector leakage current have been measured by several groups during exposure to neutrons, protons and photons. The damage coefficient α is usually defined as

$$\alpha \cdot \phi = \Delta I / A \cdot t,$$

where ΔI is the change in current due to irradiation, $A \cdot t$ is the depleted volume, and $\phi = N/A$ measures the particle flux per cm^2. α not only depends on the type and energy of the particles, but also on the specification of the silicon, the operating temperature, etc. Typical values are for protons of several GeV momentum $\alpha_p = 4 \times 10^{-7}$nA/cm, for slow neutrons (1MeV) $\alpha_n = 2 \times 10^{-7}$nA/cm, and for 20 keV photons $\alpha_\gamma = 8 \times 10^{-11}$nA/cm. Thus over an area of 1 cm^2 one expects an increase in the detector leakage current of 1 nA or a flux of about 10^8 protons/cm^2 or 4×10^{11} γ/cm^2. AC coupling of the detector output to the preamplifier avoids saturation effects due to this dark current increase.

9.2. Estimated Radiation Dose

For the purpose of deriving a maximum background rate during data taking, we assume a dose limit of 100 kRad/year and operation over a period of five years. This rather conservative limit is based on the expectation that the effort that is presently directed towards the development of radiation hard analog circuits for the SSC will lead to a substantial increase in the dose tolerance of circuits and detector during the next few years. The annual dose limit also takes into account that in addition to the radiation absorbed during data taking substantial doses are expected to be accumulated during injection and machine studies.

For a flux of photons incident on a silicon detector the received dose rate, D, can be expressed as the energy weighted integral over the incident photon spectrum, $f(E)$:[39]

$$D = c_o N < E > / A\rho t$$

with

$$< E >= \int f(E) E (1 - e^{-t/\lambda(E)}) dE,$$

where N is the photon flux, $c_o = 1\text{kRad}/6.24 \times 10^{13}\text{keV}/g$, $\rho = 2.33\ g/\text{cm}^3$ is the density, $t = 0.03$ cm the thickness, and A the area of the silicon detector. $\lambda(E)$ is the photon absorption length of silicon, which can be parameterized as $\lambda(E) = 0.13\text{mm}(E/10\ \text{keV})^{7/2}$.

The 9 GeV beam (APIARY 6.0 lattice) produces a synchrotron radiation spectrum incident on the first layer of silicon that is approximately flat from 15 keV to about 70 keV, it has a large peak at 8 keV from fluorescence in the copper lining of the vacuum pipe.[40] About 50% of the photons transmitted through the beam pipe are absorbed in the first layer; the absorption rate in the second and third layer is substantially smaller. The rates and absorbed energies averaged over the total detector layer and normalized per area are listed in Table 5. Clearly, the first layer of the vertex detector acts as an active absorber for synchrotron photons! In this layer the photon flux is 0.04 $\gamma/\text{cm}^2/\mu\text{s}$ resulting in an absorbed annual dose of 1 kRad and an occupancy of less than 0.1%.

About 50% of the synchrotron photons transmitted through the beam pipe get absorbed in the first layer of the silicon detector.

Source	Layer	N_γ $[\text{cm}^{-2}\mu s^{-1}]$	N_e $[\text{cm}^{-2}\mu s^{-1}]$	E $[\text{keVcm}^{-2}\mu s^{-1}]$	Dose $[\text{kRad}/10^7 s]$	Occupancy $[\%/\mu s]$
SR	1	$42 \cdot 10^{-3}$	–	$400 \cdot 10^{-3}$	1.0	0.08
	2	$11 \cdot 10^{-3}$	–	$13 \cdot 10^{-3}$	0.03	0.005
	3	$7 \cdot 10^{-3}$	–	$8.3 \cdot 10^{-3}$	0.02	0.003
LP	1	0.54	0.20	177	407	1.6
	2	0.40	0.061	39	89	0.5
	3	0.27	0.026	15	35	0.2

Table 5. *Background rates, absorbed energies, annual doses and occupancies in a three layer silicon vertex detector for synchrotron photons (SR) and for photons and electrons ($e^{\pm}$) from showers generated by lost beam particles (LP)*

The average energy absorbed per photon incident on the first silicon layer is approximately $< E >= 10$ keV, thus on the basis of the calculated synchrotron spectrum the maximum tolerable flux of photons transmitted through the vacuum pipe is

$$N_{max} = 3.7 \cdot 10^7 D_{max} t / < E >= 4 \cdot 10^6 \text{cm}^{-2}\text{s}^{-1}.$$

Photons and electrons from lost beam particles showering near the interaction region have rather steeply falling spectra, with most of the particles in the range of a few MeV.[41] Less than 0.3% of these photons convert and their energy deposition remains small compared to that for electrons. Electrons traverse the silicon at rather large angles, curl in the magnetic field, and lose large amounts of energy. Every crossing of a spiraling electron is counted as a separate particle. Due to the axial magnetic field the rate of electrons decreases rapidly with the radial distance from the interaction point, the number of photons remains roughly constant per solid angle.

The result of an EGS simulation are given in Table 5. There is a rather large energy deposition by the electrons in the silicon. As a result, the annual dose rate is well above the stated limit of 100 kRad/y. Assuming that the EGS shower calculations and incident spectra are correct, the maximum tolerable flux of electrons from showers is

The dominant source of background is lost beam particles.

$$N_{max} = 1.3 \cdot 10^5 \text{cm}^{-2}\text{s}^{-1}.$$

(n.b. this limit is a factor of 10 lower than the number one obtains under the assumptions that the energy loss per electron is equivalent to minimum ionization at normal incidence.) While these doses are very high, one should keep in mind that the storage ring lattice used for these calculations is still preliminary in nature, and one way to reduce this background at the source is to introduce bend magnets in the long straight sections outside the interaction region. It is estimated that this would reduce the rate of scattered beam particles by at most a factor of 5. An improvement in the vacuum pressure of 10^{-8} Torr may also be feasible over a limited length.

Special measures need to be taken to reduce beam particle losses in the straight sections outside the IR.

9.3. *Occupancy Estimates*

Synchrotron radiation as well as photons and electrons generated from lost beam particles showering near the interaction region can cause signals above threshold in the silicon detector. Such signals may influence the position resolution if a hit is close to a particle track or if hits in different layers line up to fake a track segment and contribute to the trigger or affect pattern recognition. Both of these effects become severe at a level of strip occupancy of about 1% per trigger.

The silicon strip detector modules are typically segmented into 4 cm long read-out strips of 50μm pitch, thus the area per strip is $S = 4 \times 0.005 = 0.02\text{cm}^2$. It is assumed that the electronics provides analog pulseheight information for every strip and has (depending on the mode of read-out) a sensitive time of up to $\tau = 1\mu$s. This time can be reduced by a factor of 10 if necessary, probably at the expense of higher power dissipation. Two hits overlap in space if they are separated by less than 4 strip widths (for large angles of incidence the pulseheight spreads over more strips and this limit is too tight). Thus the occupancy generated by a particle flux $N[\text{cm}^{-2}\mu\text{s}^{-1}]$ can be defined as

$$O[\%] = 4S\tau\epsilon N = 8 \cdot \epsilon N \,,$$

where ϵ is the efficiency for detecting a photon or electron.

Most of the background hits in the silicon are due to low energy $e^{\pm}$ generated by lost beam particles showering near the IP.

For electrons, we take $\epsilon = 1.0$, even though one could lower ϵ by using pulseheight information to discriminate against large energy clusters from particles that range out. The observed fluxes translate directly into occupancies of 1.6% or less (Table 5).

For photons from showers generated by lost beam particles that hit the silicon layers the conversion probability is small, of the order of 0.3%, translating to an increase in the occupancy rates by 0.01% or less.

Synchrotron photons are of lower energy and have a much higher probability of converting in the silicon. About 1/2 of the photons incident on the first layer convert, roughly 1/4 of them fall below a typical pulseheight threshold of 10 keV, thus the effective detection efficiency is about $\epsilon = 0.25$. So synchrotron photons cause an occupancy of about 0.08% in the first layer, and much less in the others.

9.4. Summary and Comments

From the background calculations we conclude that the effects that are most likely limiting the operation of a silicon strip vertex detector in the presence of beam background are

- the occupancy in the first detector layer, and
- the surface damage in the VLSI electronics.

At this time, the radiation damage appears to be the more severe limit though there is a very active R&D program underway to improve the radiation hardness of analog circuits. The occupancy limits are given for a rather long sensitive time of $1\mu s$, leaving room for substantial improvement in case the strip occupancy should become the more severe limit.

The average occupancy in a silicon strip detector is expected to be below than 1%, even if we assume a sensitive time of up to $1\mu s$.

While rates have been calculated for strip detectors of medium length and segmentation, limiting doses and particle fluxes are much less stringent (more than an order of magnitude) for pixel devices that have finer segmentation and possibly reduced thickness.

The estimates derived from the background simulations and the evaluation of their effect on the performance of a silicon vertex detector are relatively crude and preliminary. The numbers presented here should only be used for a first evaluation of the expected and acceptable background levels, to be followed up by more detailed analyses with more realistic detector specifications and simulations. To calculate realistic limits on occupancy requires a complete pattern recognition algorithm and a Monte Carlo event simulation to which background hits can be added, with rates based on measurements. The effect on a charged particle trigger cannot be estimated without a conceptual design of the read-out and trigger system, the detector design and segmentation. In addition, one must take into account collective effects, e.g., beam-gas background is likely to generate a single shower concentrated in a limited area of the detector.

In summary, background simulations based on the first design of the interaction region indicate that the dominant limitation is not due to synchrotron radiation but due to energy deposition of charged particles generated in showers caused by lost beam particles. Future lattice designs will have to take this limitation more seriously. Furthermore, it is important to remember that a large fraction of the total radiation dose is likely to be delivered during injection and machine studies.

10. Electronics

The silicon vertex detector discussed here places about 10^5 read-out channels in a very limited space imposing severe restrictions on the volume occupied by the read-out electronics, cables and connectors, and on the power consumption. A typical front end electronic circuit for each read-out channel includes a preamplifier, a comparator, analog storage cell, and digital circuitry for time tagging and read-out. A block diagram of such a circuit is shown in Figure 28. Typical specifications for operation at a B Factory are:

- A signal/noise ratio exceeding $S/N = 30 : 1$, here the signal S is taken to be equivalent to the charge collected for a minimum ionizing particle at normal incidence and the noise N corresponds to the single channel rms noise. This ratio determines the relative setting of the threshold, and is particularly important for tracks at non-normal incidence causing the signal charge to spread over many channels. A larger S/N also eases a radiation resistant design and operation at high absorbed doses, because it permits a higher threshold setting.

- The signal rise time should be in the range of 20-50 ns to allow for an integration time of the order of 100 ns. This would limit the sensitive time of the detector and thereby reduce noise and background hits.
- The read-out should be selective, i.e. only transfer information for channels that have pulseheights exceeding a preset threshold.

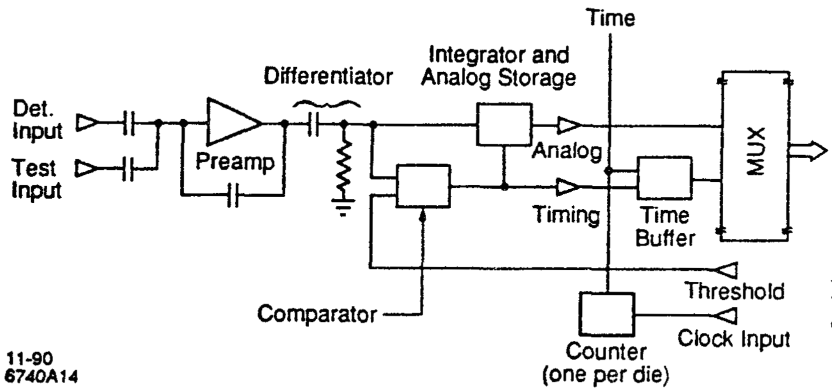


Figure 28. *Block diagram of silicon detector read-out.*

- For fast track finding at the secondary trigger level bit patterns for groups of 32 or 64 strips could be extremely useful.
- The power dissipation should be as low as practical, preferably less than 1 mW/strip or 20 μW/pixel, to ease the specifications for the cooling system.
- The circuit architecture must have radiation resistance adequate for five years of operation, preferentially longer, including periods of machine testing and beam injection. The effects of reverse bias current in the detector need to be accommodated either by capacitive coupling of the input signal or specific signal shaping.
- For analog read-out and comparator threshold setting signal calibration is necessary.
- The amount of material and its placement need to be optimized to reduce the multiple scattering and evenly distribute the material inside the active volume. Given the large number of channels and the goal of complete geometric coverage for each detector layer, packaging of

the amplifier and read-out circuits and the associated cable connections will require special design efforts.

10.1. Preamplifier Circuit

The design of the amplifier circuit depends critically on the input capacitance of the channel which by-and-large is determined by the input capacitance. For a strip detector with strips of up to 4cm in length, 50μm pitch, and 300μm thickness this capacitance is typically 5 pF; for a pixel detector stray capacitances associated with signal leads represent the dominant fraction of the total input capacitance. A minimum ionizing particle generates 80 electron-hole pairs per μm of silicon, so for a 300μm thick detector we expect a signal of 4 fC. Thus the front end electronics needs to provide integration and charge amplification by about a factor of 1000. Elements of the circuit that is schematically laid out in Figure 28 are under study for a vertex detector at the SSC.[42] The high gain, charge sensitive capacitive feedback preamplifier is similar to CMOS circuits presently in use at SLC, LEP, and CDF, however, care has to be taken to assure that the low noise performance is compatible with low current operation. A slow DC restoration loop could be used to return the preamplifier to its quiescent state within 10μs of a hit. This technique makes the amplifier less sensitive to high frequency noise. Also, no external clock signals are necessary for operation of the amplifier circuit, i.e. there is no switching activity until a signal arrives. Each channel has an auto-reset that is activated after a maximum anticipated time to generate a trigger. Given an integration time of 100 ns and a typical detector leakage current of 1-2 nA per microstrip the integrated leakage is 0.2 fC, roughly 1/20 of the signal charge. Taking into account that the leakage current can vary from strip to strip and also increases with exposure to radiation it is necessary to take special precaution. First of all, it is advisable to couple the input capacitively to the amplifier input. Second, one can either employ a multiple sample and hold technique as in the SVX or Microplex chips or use a passive CR shaper at the output of the preamplifier with a time constant of at least 100 ns.

Analog gain can be calibrated with the help of test pulses that are selectively injected into the circuit input via individual calibration capacitors.

10.2. Sparse Readout

At the shaper output the signal splits into a fast channel with the comparator and a slow channel for analog storage. The amplified pulse is compared to a preset threshold, and a latch is set for channels above threshold, the neighboring channels are also latched. Following a secondary trigger the digital control circuitry on each chip of 64 or 128 channels switches sequentially to those channels which are latched. For each latched channel the stored analog voltage is connected to a single bus, the channel address is strobed simultaneously to a digital bus. With this on-chip sparsification feature the read-out time of the detector is set by the occupancy rather than the total number of channels. Typical read-out times of $1\mu s$ per channel have been achieved. The detector occupancies are expected to be dominated by charge deposition from beam related background, occupancies at the level of a 1/100 are acceptable.

At a B Factory with a beam frequency of 250 MHz it may be necessary to mark the latches by an on-chip clock to distinguish background pulses from beam crossings that are more than 100ns apart. The clock information and bit patterns could then be stored in a digital pipeline. In the event of a primary trigger the stored latches could be scanned and for each time bin a fast track search derived from groups of 32 or 64 channels in several layers could be performed. In the event of a secondary trigger, the read-out would select the channels with signals recorded in the corresponding time bins and transfer the pulseheight to the read-out bus.

10.3. Radiation Hardness

Several efforts are presently underway towards the design, fabrication and testing of low noise, low power analog circuits that can operate at total radiation doses of more than 10Mrad. One is based on 1.2μm radiation hard CMOS process developed and offered by UTMC of Colorado, the other is based on JFET's, n-type Junction Field Transistors. JFET's are majority carrier devices and are therefore not affected by the creation of minority carriers by irradiation. They also are capable of high switching speeds. The only disadvantage of JFET's appears to be the difficulty to fabricate low noise JFET's and MOSFET's on the same wafer.

Another way to further reduce the effects of radiation and to increase the potential speed of the circuit, is to use a bulk insulator to separate devices on the same substrate, thus replacing the field SiO_2 by a gap of insulator, so-called silicon-on-insulator, SOI, or silicon-on-sapphire, SOS. Due to the absence of the SiO_2 the effect of so-called oxide charges trapped during exposure to ionizing radiation is expected to be much less severe.

10.4. Cables and Packaging

Methods for interconnection of the detector strips and of the electronics to control and read-out lines are closely related to the details of the lay-out and the mechanical design. For each detector element there will be one connection to the amplifier chip. In addition, individual chips need to be connected to power, clocks, and signal distribution busses carried on a mounting substrate or the detector surface, and connected via cable feeders to external control and DAQ systems.

The large number of channels make these interconnections a major fabrication and quality control issue. Traditional wirebonds for the detector inputs may be precluded and could possibly be replaced by indium bump bonds with the electronics chip placed on top of the detector surface. In such a hybrid assembly the read-out circuit would be located in the active volume of the detector and efforts are under way to thin the read-out chip to 100μm or less and thereby reduce the total material.

New interconnection techniques which might apply to silicon tracking systems are under development in industry. These include high-density tape bonding, and polimide interconnect films that carry traces and contact pads to provide power and signal bussing for read-out chips.

10.5. Summary

A very active program for the development of amplifier and read-out circuits is under way in several laboratories. These efforts are supported as part of the detector R&D program for the SSC. While some of the specifications for silicon vertex detectors at a B Factory are different, most of the problems are similar. We can expect that much will be learned in the next few years. This is particularly true for the basic low noise, low power amplifier and comparator circuit design and problems related to a radiation hardness. Likewise, problems of signal connection and distribution, as well as heat transfer are similar, and the B Factory design can expect to benefit from the SSC R&D program.

11. Study of a Silicon Detector Trigger

11.1. Introduction

We examine the possibility of using the silicon detector in the second level trigger.

A charged particle trigger based solely on the central drift chamber may have problems due to the lack of z information and the rather large distance between the IP and the chamber's first layer. The use of the silicon vertex detector in a charged particle trigger has been explored,[43] both as an adjunct to the central drift chamber and on its own. It is intended to be part of the second level trigger because it appears unlikely that the vertex detector information can be accessed fast enough for a primary trigger. The tolerable primary rate is 10^5 Hz and the secondary rate should not exceed 100 Hz by very much. Thus the secondary trigger needs to reduce the trigger rate by a factor of order 10^3.

No attempt has been made so far to describe such a trigger in detail or to model the detector background. Instead, formulae have been derived for the trigger rate in the presence of a uniform, uncorrelated background (salt and pepper) for several schemes of using the silicon vertex detector in a trigger. Trigger rates are calculated for different levels of background and detector resolutions.

11.2. Segmentation and Singles Rate

Channels are grouped to form towers for trigger purposes.

In the following we assume a three layer silicon vertex detector. Towers are formed from detector segments in different layers which have the same angles ϕ or θ to the IP. Towers may be true $\phi - \theta$ towers or they could be projected towers in ϕ or θ. Each tower is made of up to three elements in the three different layers. These elements cover the same solid angle and therefore experience the same rate in the presence of a uniform background. Since the background and the signal tracks are expected to peak strongly in the direction of the higher energy beam, endcap planes have been added to the central vertex detector in increase the segmentation and achieve a more uniform occupancy. The occupancy O of a tower element is given as

$$O = (\Omega/4\pi) b\tau_s,$$

where Ω is the solid angle of a tower, b is the total background rate per layer, and τ_s is the time resolution of the silicon as seen by the trigger electronics. In the following, plausible values for Ω, b, and τ_s are discussed.

The useful tower size is determined by the beam size and stability, resulting in an effective spot size of about 2 mm in the x and y, and 5 mm in z projection. This translates into a segmentation of approximately 100 projected towers, in ϕ or θ, or $4\pi/\Omega$=10,000 "true" $\theta - \phi$ towers.

The dominant background is

The anticipated sources of background are synchrotron radiation and electromagnetic showers due to beam particles which have interacted with residual gas. Recent calculations by Sullivan[40] (synchrotron radiation) and Coupal and Hearty[41] (off-energy beam particles) indicate that the main source is expected to be shower particles that hit the silicon detector at a rate of approximately 25 MHz in the first layer, and less in the two outer layers. This rate estimate may be too low because shower particles are not followed below 3 MeV. On the other hand, the present design has a total radiation dose that exceeds the tolerable limit of 100 krad/y and therefore techniques for lowering this background must be pursued.

expected to be caused by lost beam particles.

Since the dominant background is caused by charged particles, the assumption of "randomness" might be called into question. However most of these background tracks are of low momentum and originate at large radius or z and thus are not likely to produce hits in more than one element of a tower. Still, there remains a concern because the trigger rate estimates rely heavily on absence of correlation. Without a detailed simulation we can only establish necessary but not sufficient conditions for the trigger to work.

We assume that the background hits are uncorrelated.

The time τ_s is the precision to which the time of a hit can be recorded. τ_s is mainly determined by the design of the amplifier and read-out circuit which may be driven by limits on the power consumption and input capacitance. While $\tau_s =$ 100 ns may be possible for a strip detector, $\tau_s = 30$ ns should certainly be achievable for pixel detectors. $\tau_s = 100$ ns is well matched to the time resolution obtainable in the central drift chamber. The time resolution for the primary trigger τ_p is also assumed to be $100ns$ or less.

A time resolution of 100 ns is assumed in most examples.

11.3. Coincidence Rates for Towers

The probability that two out three elements of a tower will be hit is

$$P_{2/3} = 3(\Omega/4\pi)^2 b^2 \tau^2 \kappa$$

if $O = (\Omega/4\pi)b\tau_s \ll 1$. The factor κ takes into account correlations and tower overlaps, one expects $\kappa = 2 - 5$. If the detection efficiency for minimum ionizing tracks is very high one could reduce the coincidence rate for tracks by a factor 3 by requiring that two out of two elements were hit. The probability that any tower coincidence is satisfied for a primary trigger is then

$$3(\Omega/4\pi)b^2\tau_s\tau_p\kappa \qquad \tau_p \geq \tau_s.$$

For a tower size of $10^{-2}\,4\pi$, a background rate of $b = 10^7$ Hz, and time resolutions of $\tau = 10^{-7}$ s, the ratio of secondary to primary trigger is 3%, provided the primary trigger is uncorrelated with the silicon detector.

This first estimate is encouraging. Next, we look at a strip detector with two coordinate read-out segmented into modules. For every ϕ tower coincidence we search for a θ tower coincidence in the same module. If $b\tau \leq 1$ the rate for an accidental ϕ and θ tower coincidence is small and the ratio of secondary to primary rate is

$$R_T = 3(\Omega_\phi/4\pi)(\Omega_\theta/4\pi)b^2\tau_s\tau_p \quad \tau_p \geq \tau_s.$$

By using both θ and ϕ information, we get a trigger rate reduction of 10^4 per track.

For 100 ϕ towers and 100 θ towers we have $R_T = 3 \times 10^{-4}$. A pixel detector would give the same result.

This estimate is valid for a single track. If we do not permit a hit to be used in more than one tower coincidence, the two track coincidence rate is obtained by simply squaring the single track rate.

11.4. Associating SVD Towers with CDC Tracks

The probability of associating a tower with a track in the central drift chamber which only provides ϕ information can be obtained by simple extension of the formulae derived above, and we obtain

$$3N(\Omega_\phi/4\pi)^2b^2\tau_s\tau_p = 3N \cdot 10^{-4}.$$

Here N is the number tracks in the CDC, and we have assumed that there are $\Omega_\phi/4\pi = 100$ towers in ϕ in the SVD. Furthermore, we take $b = 10^7$, and $\tau_s = \tau_p = 10^{-7}$. The probability that a CDC track will match a ϕ tower which is associated with one of $\Omega_\theta/4\pi = 100$ θ towers is

$$3N(\Omega_\phi/4\pi)^2(\Omega_\theta/4\pi)b^2\tau_s\tau_p = 3N \times 10^{-6}.$$

To obtain the rejection factor for a two-track trigger, we square these results, replacing N by $N(N-1)$. Thus for the most effective trigger, two CDC tracks matched in ϕ and associated with a θ tower in the SVD, we have

$$18(\Omega_\phi/4\pi)^4(\Omega_\theta/4\pi)^2b^4\tau_s^2\tau_p^2 \times \kappa^2 = 1.8 \cdot 10^{-11}$$

This is clearly a very encouraging result! For a higher background rate, $b = 10^8$, the rejection factor is $1.8 \cdot 10^{-7}$. If in addition the tower segmentation is coarser in both ϕ and θ, 50×50 towers, and also the time resolutions degrades to $\tau_s = \tau_p = 200$ ns, then the rejection is 2×10^{-4}. Since we are trying to reduce a primary trigger rate of 10^5 Hz to 10^2 Hz, this appears to be entirely adequate.

Good rejection factors are obtained even with conservative assumptions.

11.5. Correlations

If we wish to reject tracks which originate from interactions in the beam pipe wall we need to define towers which point to only a small fraction of the central beam pipe. Examination of the geometry of the SVD in the $r - \phi$ plane shows that the number of towers must be greater than 60 to ensure that most of the beam pipe will not be seen by towers formed from the outer two layers.

Assuming this number of towers we can examine the momentum acceptance of towers of this size. For three detector layers at radii r_i with $r_1 : r_2 : r_3 = 1 : 2 : 3$ and $r_3 = 69$mm, operating in a magnetic field of 1 Tesla the minimum transverse momentum p_t is 12 MeV/c. Tracks originating in the beam pipe will not be able to form a tower if they do not reach at least one of the two outer layers. The p_t cut off is 3.3 MeV/c for the second, and 6.6 MeV/c for the third layer.

Correlations due to multiple tracks from the same shower can only be studied by detailed simulation, and they have been ignored here.

11.6. Conclusion

Based on presently available estimates of background rates and the assumption that they will dominantly generate uncorrelated signals in a three layer silicon vertex detector, we conclude that this device would be a very powerful addition to a charged particle secondary trigger. It can be used to effectively reject tracks originating from outside the beam-beam interaction region. The major impediment to its implementation may be the additional electronics and connections necessary to transfer the coordinates of the hits or towers to a trigger processor in a few μs.

The use of the silicon detector in the trigger appears to be a practical problem rather then one of principle.

REFERENCES

1. F. Lediberder, *BaB̄ar* Note #34.
2. F. Lediberder, W. Toki, M. Witherell, *BaB̄ar* Note #45.
3. D. Cords, contribution to this workshop.
4. M. Yurko, contribution to this workshop.
5. G. Bellini, *et al.*, *Phys. Reports* **84**, 9 (1982).
6. E. Belau, *et al.*, *Nucl. Inst. Meth.* **217**, 224 (1983).
7. J. Kemmer, *Nucl. Instr. Meth.* **225**, 89 (1984).
8. E. Focardi, *IEEE Trans. Nucl. Science*, (1986).
9. A. Schwarz, private communication.
10. J. Alexander, *et al.*, CLNS-90-1034 (1990).
11. S. Barlag, *et al.*, *Phys. Lett.* **218B**, 374 (1989)
12. S. Parker, *Nucl. Instr. Meth.* **A275**, 494 (1989).
13. S. Holland and H. Spieler, *IEEE Trans. Nucl. Science*, San Francisco, 1989.
14. S.L. Shapiro, contribution to this Workshop.
15. S. L. Shapiro, et al., *Nucl. Instr. Meth.* **275**, 580 (1989).
16. S. Gaalema, *IEEE Trans. Nucl. Science,* **NS-32**, 417 (1985).
17. E. Gatti and P. Rehak, *Nucl. Inst. Meth.* **A235**, 365 (1985), E. Gatti, *et al.*, *Nucl. Inst. Meth.* **A273**, 865 (1988) and **A274**, 469 (1989).
18. P. Rehak, contribution to this Workshop.
19. A. Weinstein, ASLUND documentation.
20. A. Breakstone, private communication.
21. D. Stoker, contribution to this Workshop.
22. V. Lüth, *BaB̄ar* Note #55.
23. V. Lüth, *BaB̄ar* Note #48.
24. V. Lüth, contribution to this Workshop.
25. S. Komamiya, contribution to this Workshop.

26. V. Lüth and D. MacFarlane, contribution to the DPF Summer Study on High Energy Physics in the Nineties, Snowmass (1990). See also SLAC-PUB-5419.

27. E. Michel, *et al.*, *Nucl. Inst. Meth.* **A283**, 554 (1989).

28. G. Wormser, *Ba$\bar{B}$ar* Note #39.

29. J. Va'vra, Contribution to this Workshop.

30. F. Angelini, *et al.*, , *IEEE Trans. Nucl. Science*, 69 (1990).

31. H. Schember and P. Bhandari, An Investigation of Beam Pipe Cooling Techniques for a Proposed Particle Accelerator, JPL D-7151 (1990).

32. A. Lisin, private communication.

33. R. Erbacher and W. Vernon, contribution to this Workshop.

34. Energy Research and Generation, Inc., Oakland, CA.

35. C. Damerell and F. Wickens, private communication.

36. T. Ohsugi, *et al.*, *Nucl. Instr. Meth.* **A265**, 105 (1988); M. Hagesawa, *et al.*, *Nucl. Instr. Meth.* **A277**, 395 (1989); G. Lindström, *et al.*, DESY preprint 89-105; H. Sadrozinski, SCIPP 90-02 (1990).

37. P. Dauncey, *et al.*, Contribution to the IEEE Nucl. Science Symposium, San Francisco, 1987, SLAC-PUB-4462 (1987).

38. H. Dietl, *et al.*, Contribution to the European Conference on High Energy Physics, Bari (Italy) 1985.

39. J. Alexander and P. Drell, Synchrotron Radiation Rate Limits for Tracking Detectors, Cornell Internal Report (1990).

40. M. Sullivan, contribution to this Workshop.

41. D. P. Coupal and C. Hearty, contribution to this Workshop.

42. S. Holland, *et al.*, Short Strip Detector Systems, Proposal submitted to the SSC Detector R&D Committee, 1990.

43. W. Innes and C.K. Jung, contribution to this Workshop.

Single-Track Resolution Studies in Silicon Vertex Detectors

D. P. Stoker

1. Introduction

Monte Carlo simulations of single-track resolution in silicon vertex detectors, which are of use in designing a silicon vertex detector for an asymmetric B Factory are presented. A stand-alone program was written to propagate helical tracks through an arbitrary geometry of cylinders and planes parallel to the z-axis, and planes perpendicular to the z-axis. Multiple scattering, in the Gaussian approximation, was performed as the tracks passed through the beam-pipe, detector planes, and the intervening air. The resulting "hits" in the silicon detector were smeared by the assumed single-layer position measurement resolutions. The program then fitted the tracks to helical track parameters using the iterative Newton's method. The error matrix used in the fit included both the correlated errors due to multiple scattering in the beam-pipe and detector planes, and the single-layer position measurement resolutions. The fit assumes that the vertex detector is enclosed by a drift chamber which measures the sign of the particle's charge and its momentum (GeV/c) with a resolution $\sigma_p/p = [(0.014)^2 + (0.0040p_t)^2]^{1/2}$, similar to that of the Mark II Upgrade CDC, but provides no other information. Unless otherwise stated, the tracks simulate 1.0 GeV/c muons in a 1.0 T magnetic field parallel to the beam direction.

At an asymmetric B Factory, where the $\Upsilon(4S)$ is produced with a considerable boost in the direction of the higher energy beam, the distance between the decay vertices of the $B\overline{B}$ pair is increased in the beam direction relative to that at a symmetric B Factory. The primary concern of this study is therefore the z resolution of silicon vertex detectors. The generated tracks originated from $(r, z) = (0, 0)$. After fitting each track, the z position of the track's closest approach to $r = 0$ was determined. Each z resolution shown in the following figures is the standard deviation of these z positions for 500 tracks.

2. Detector Geometry

Figures 1 and 2 show a "default" detector geometry about which several variations were made as described later. In the default geometry, the beam-pipe has an inner radius of 20 mm and a thickness of 2.1 mm, and consists of beryllium with an inner copper coating and internal water cooling. This corresponds to 0.66% of a radiation length for particles at normal incidence. The x-y view (Figure 2) of the three silicon barrel layers shows 6-, 12-, and 18-sided polygons with minimum distances from the beam axis of 2.3 mm, 4.6 mm and 6.9 mm, respectively. The barrel layers cover $|z| < 6$ cm. The silicon end planes have inner and outer radii of 3 cm and 9 cm and are located at ± 7.5 cm, ± 10 cm, ± 13 cm, and ± 16 cm. In a real detector, the region between the beam-pipe and the inner edge of the end planes may be partially occupied by support structures and cables which are not modelled here. Similarly, some additional material would extend in z beyond the ends of the barrel planes shown in Figure 1. This geometry provides coverage by at least three detector planes over the angular range $|\cos\theta| < 0.95$. The silicon layers are assumed to be 200 μm thick (0.21% radiation lengths) and to have position measurement resolutions of 11 μm in each of two orthogonal directions on the silicon surface for normal incidence particles. The dependence of the single-layer position measurement resolution on silicon thickness and track incidence angle has been calculated (see the following article by Vera Lüth); the results were used in this study.

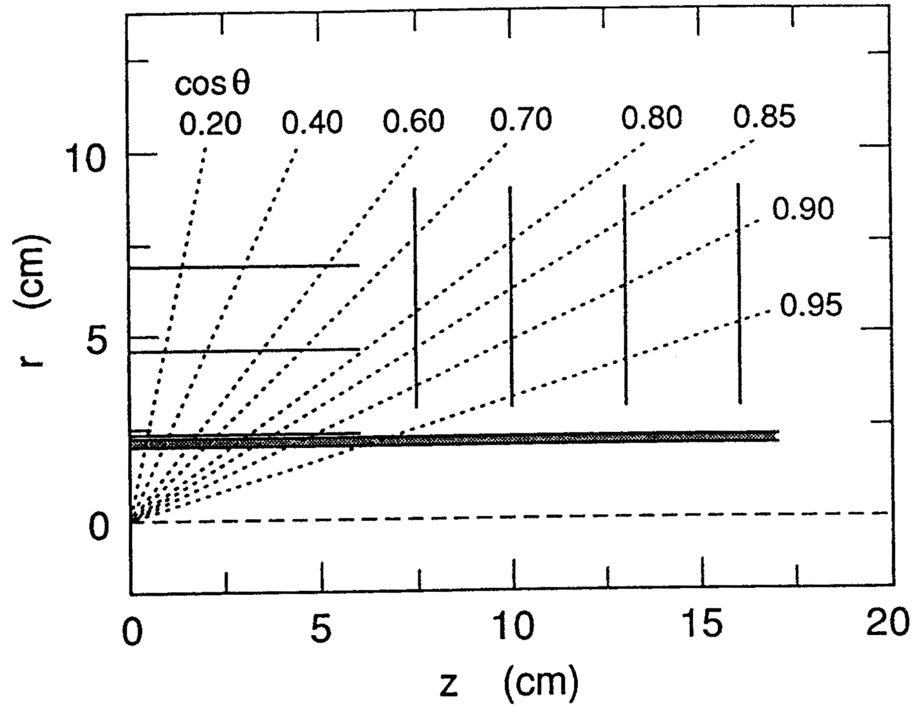


Figure 1. *rz view of the vertex detector used in this study*

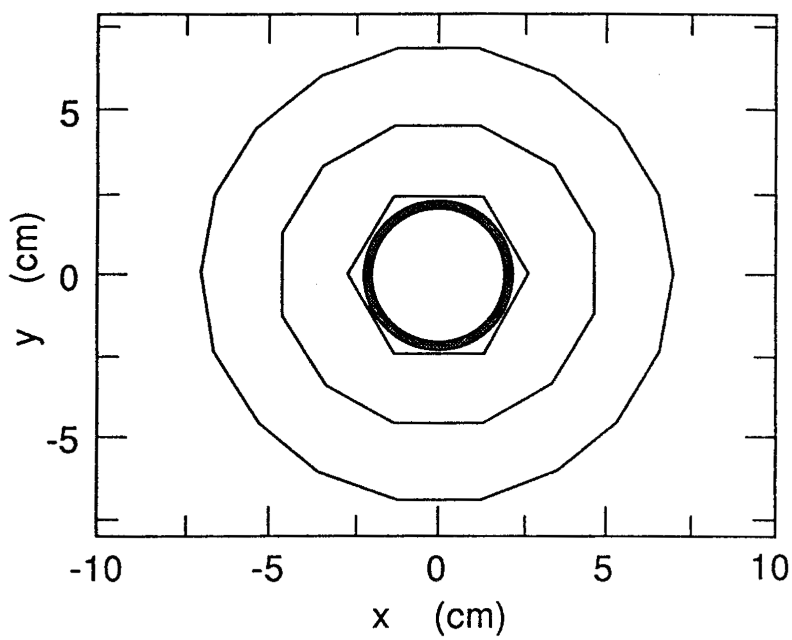


Figure 2. *xy view of the vertex detector used in this study*

3. Study Results

The results of the z resolution studies are presented in Figures 3–10 as a function of $\cos\theta$. Since the detector geometry described in the previous section is symmetric under $z \rightarrow -z$ one can take $\cos\theta$ to mean $|\cos\theta|$ in the following discussion.

Figure 3 shows the z resolution for the default geometry described above (crosses) as a function of $\cos\theta$. Cases where the silicon thickness is doubled to 400 μm (0.43% radiation lengths) and/or the beam-pipe is changed to 1.0 mm thick beryllium (0.28% radiation lengths) are also shown. In the thin silicon/thick beam-pipe case (crosses), multiple scattering in the beam-pipe is dominant, while in the thick silicon/thin beam-pipe case (squares), multiple scattering in the silicon is the larger. The cases with the thinner beam-pipe have the better resolution for $\cos\theta \lesssim 0.25$, while those with the thinner silicon layers have better resolution for $\cos\theta \gtrsim 0.25$. A slight kink in the curves for the 400 μm silicon cases near $\cos\theta = 0.8$ corresponds to the angle at which the second measurement is made by the first end plane instead of the second barrel layer.

Figure 4 repeats the study in Figure 3 for a modified geometry in which the end planes are removed and the three barrel layers are lengthened to provide coverage out to $\cos\theta = 0.95$. The results are similar to those in Figure 3 except at $\cos\theta = 0.95$ where the "long barrel" case has significantly worse resolution. As $\cos\theta$ increases, multiple scattering effects become more severe in the barrel layers and less so in the end planes. However, as $\cos\theta$ increases, the barrel layer z measurement resolution worsens rather the corresponding θ measurement resolution, while the increasing track length between barrel planes aids measurement of the track direction. The advantage of the default geometry over the long barrel geometry is probably more in its shorter length and less silicon area than in its better resolution at $\cos\theta = 0.95$.

Figure 5 repeats the study in Figure 3 for a modified geometry in which the barrel planes are shortened from $|z| < 6$ cm to $|z| < 3.5$ cm and end planes are added at $z = \pm 5$ cm.

The outer radius of the end planes is increased to 11 cm to retain coverage by at least 3 layers. The "wiggles" between $\cos\theta = 0.6$ and $\cos\theta = 0.85$ correspond to the first two measurement planes changing from being two barrel planes to one barrel plane and one end plane, and then to two end planes. Comparison with Figure 3 shows better resolution at $\cos\theta = 0.85$, 0.90, which are outside the coverage of the first barrel plane, and worse resolution at $\cos\theta = 0.70$, 0.80.

The resolutions obtained in Figures 3–5 for the three geometries are quite similar, perhaps surprisingly so. No clear advantage over the default geometry was obtained in either of the other cases.

Figure 6 shows the z resolution for the default geometry of Figures 1 and 2 (200 μm Si and 0.66% radiation length beam-pipe) and for the cases where the beam-pipe radius is decreased to 1.5 cm and increased to 2.5 cm. The distances of the silicon layers from the beam-pipe are kept constant. The resolution is best for the smallest beam-pipe radius because in this case, the measurements are made closest to the beam axis.

Figure 7 shows the z resolution for the barrel layers as the distance of the second layer from the beam axis is varied. This distance is shown as the ratio of the distance between the second and inner layers to the distance between the outer and inner layers. This ratio is 0.5 for the equally spaced barrel layers shown in Figures 1 and 2. The z resolution is worse when the second layer is closer to the inner layer because these layers then give a poorer measurement of the track direction as it emerges from the beam-pipe. It should be noted that this effect is less apparent if the single-layer position measurement resolution is improved.

In Figure 8, the dependence of the z resolution on the thickness of the silicon layers is shown for the default geometry of Figures 1 and 2. At normal incidence in the barrel layers ($\cos\theta = 0$), the thinnest silicon (100 μm) gives the worst position measurement resolution, and therefore the worst z resolution. At larger $\cos\theta$, the thinnest silicon gives the best resolution not only because it has the least multiple scattering, but also because the position measurement resolution is best, due to having the least track length component parallel to its surface.

Figure 9 shows the z resolution for single-layer position measurement resolutions of 0 μm, 10 μm, and 20 μm (at normal incidence) in the default geometry. As expected, the z resolution improves with better position measurement resolution.

Finally, Figure 10 shows the z resolution for 250 MeV, 500 MeV, 1.0 GeV, and 2.0 GeV muons for the default geometry. The improvement in the z resolution with increasing momentum is primarily due to decreasing multiple scattering which scales as $(p\beta)^{-1}$ where p and β are the track momentum and velocity, respectively.

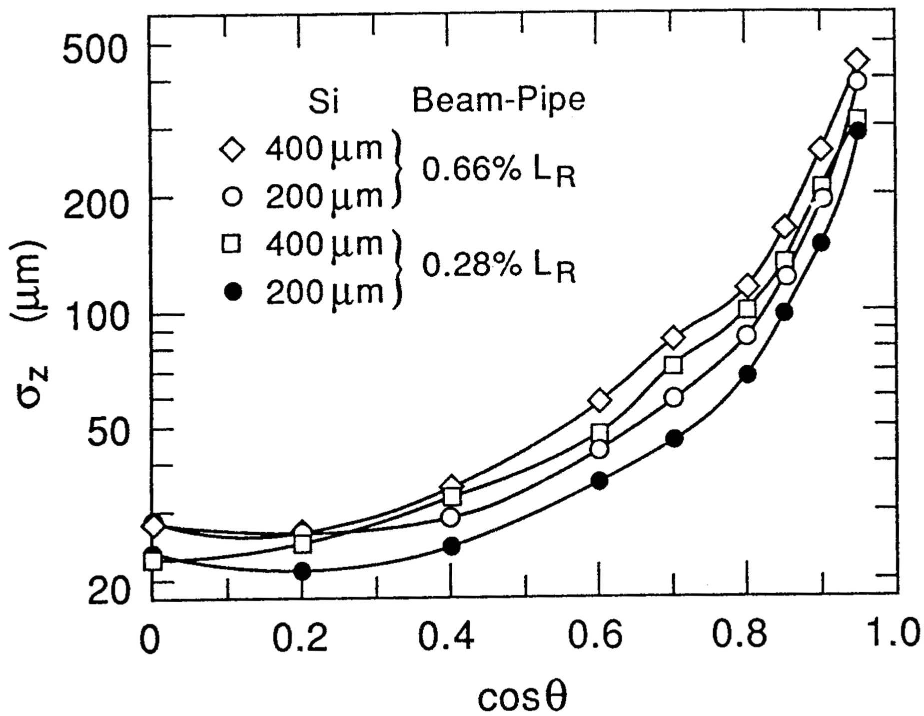


Figure 3. *z resolution for barrel with end planes.*

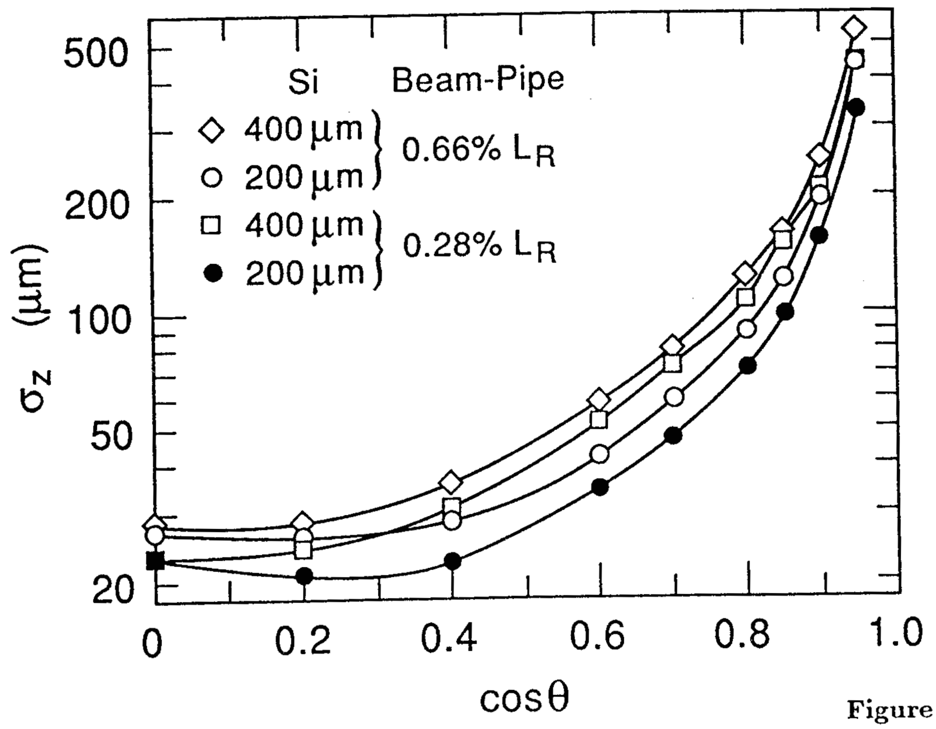


Figure 4. *z resolution for long barrel planes.*

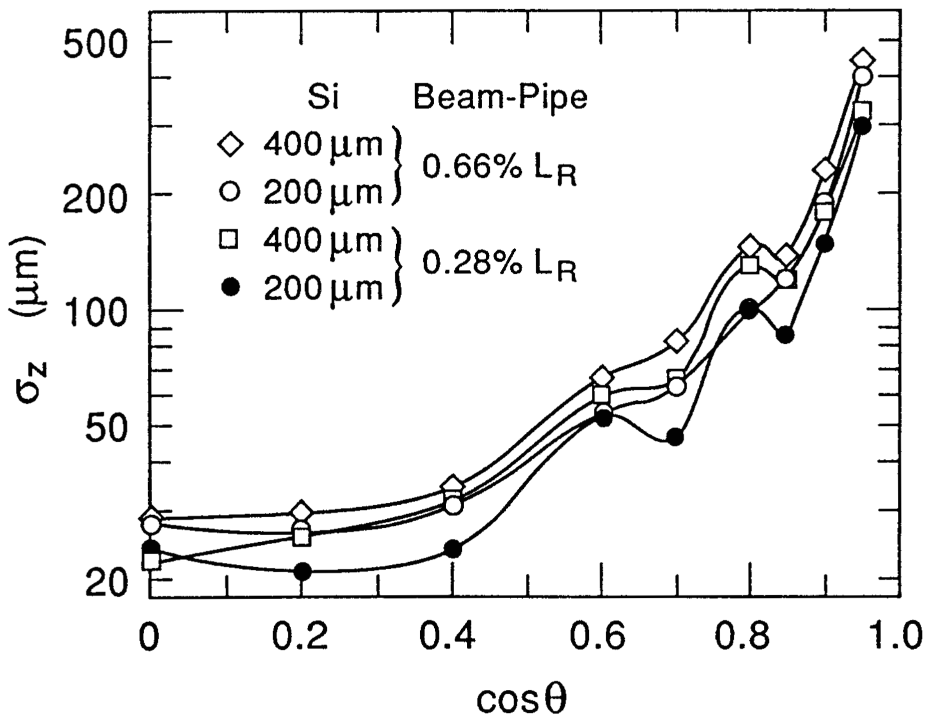


Figure 5. *z resolution for short barrel with end planes.*

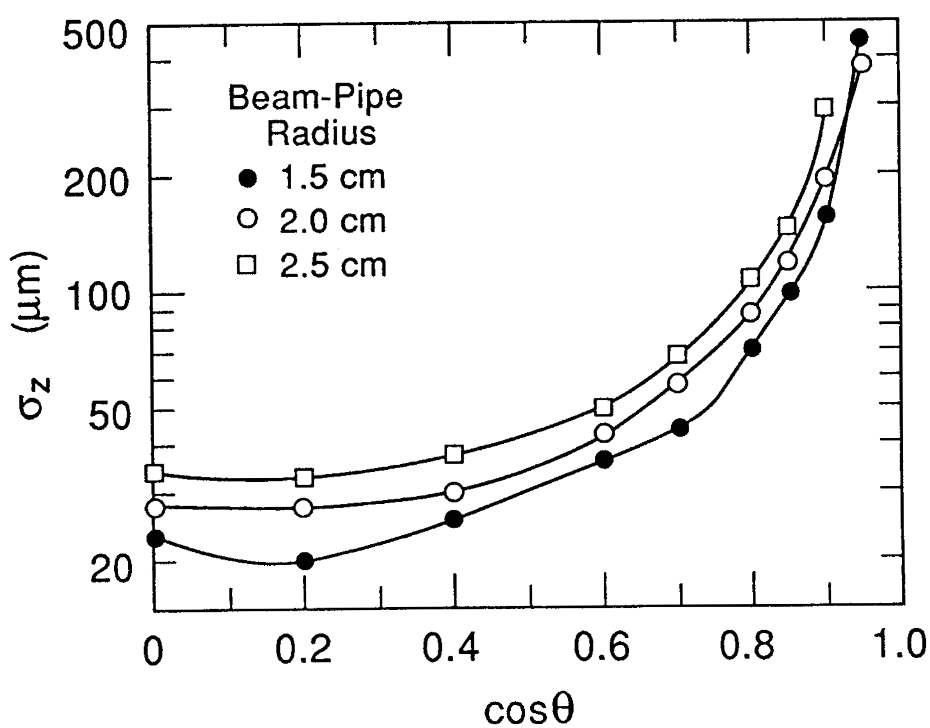


Figure 6. *Beam pipe radius dependence of z resolution.*

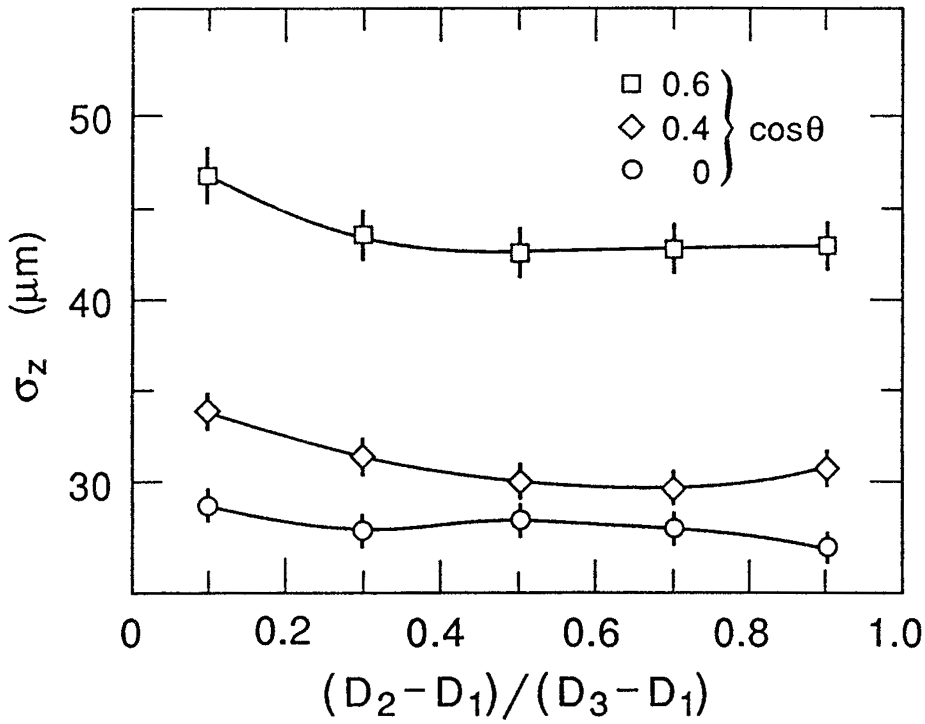


Figure 7. *z resolution vs. second barrel plane position.*

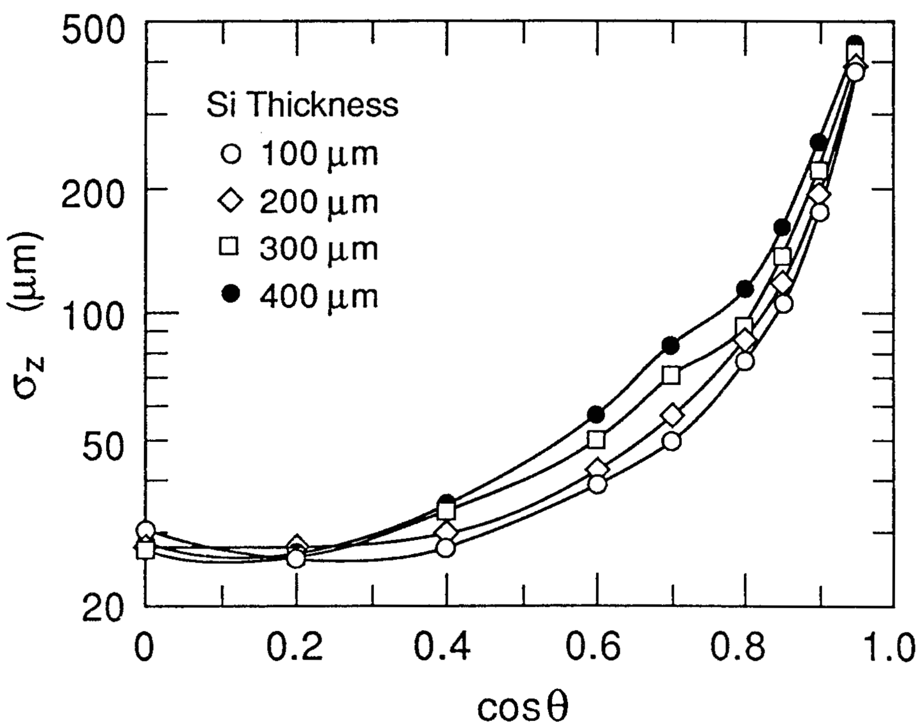


Figure 8. *Dependence of z resolution on silicon thickness.*

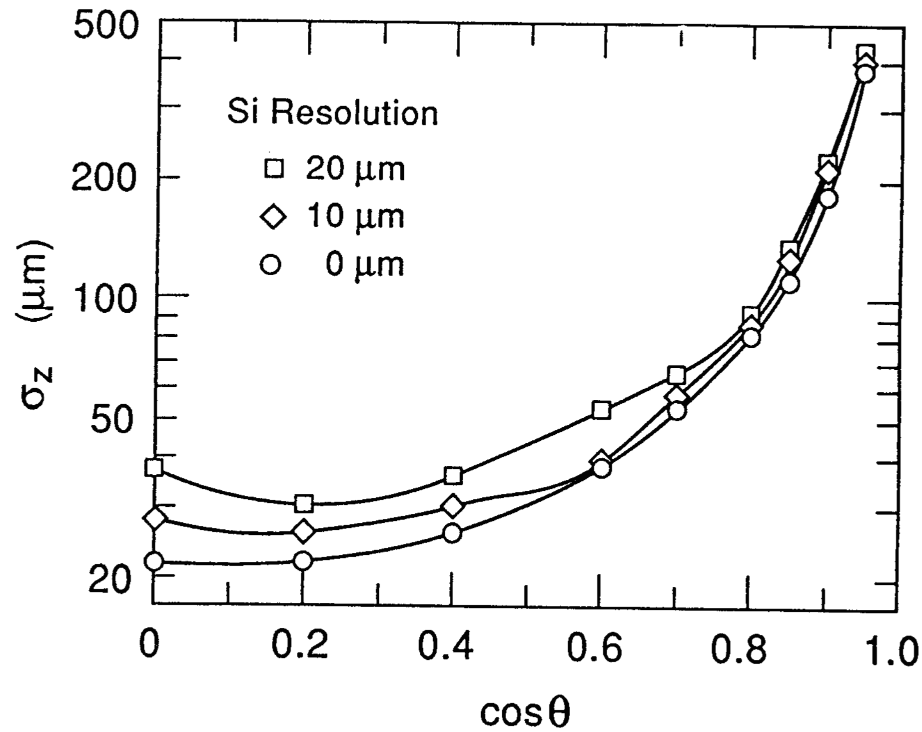


Figure 9. *Dependence of z resolution on silicon detector resolution.*

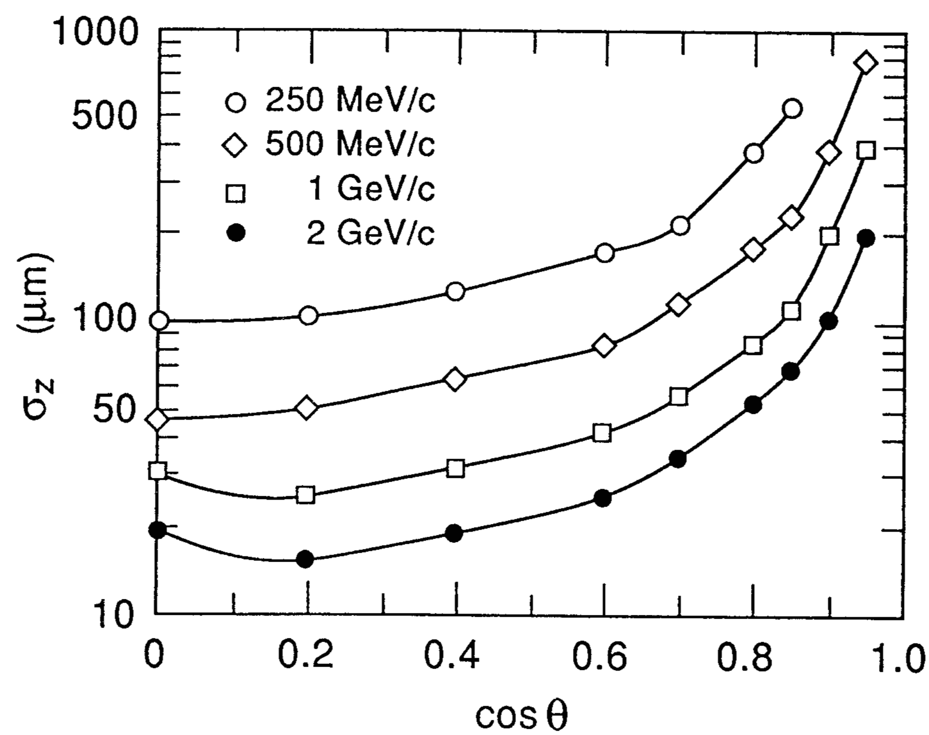


Figure 10. *Momentum dependence of z resolution.*

Spatial Resolution of Silicon Detectors
A Monte Carlo Study

V. Lüth

Abstract

The intrinsic position resolution for minimum ionizing particles traversing a fully depleted silicon strip detector is studied as a function of the angle of incidence, the read-out pitch, the thickness of the counter, the signal to noise ratio, and the pulse height cluster definition.

1. Introduction

In high energy physics experiments Silicon strip detectors have up to now been placed in such a way that the charged particles traverse the detector either almost perpendicular to the surface and parallel to the internal electric field, or such that the projections of the tracks are roughly parallel to the charge collecting strips. The charge generated by ionization along the track in the bulk silicon drifts parallel to the field. Due to transverse diffusion and delta rays this charge distribution widens to $5 - 10\mu$m and the signal spreads over one or two strip electrodes on the surface. Capacitive coupling between neighboring strips also causes the signal to spread to adjacent strips.

In future applications at storage rings particles will impact at larger angles. This is particularly the case for an asymmetric B Factory because a large fraction of the secondary particles from B decay are emitted at smaller angles around the direction of the high energy beam. For non-perpendicular incidence, the charge is generated along the track and then drifts parallel to the field, thus projecting the track onto the strip surface and resulting in a spreading of the collected charge over several strips.

This note presents a study of the effects of this charge spread by a simple Monte Carlo simulation of the charge collection, as a function of read-out pitch, detector thickness, and signal-to-noise ratio for a given amplifier and read-out circuit.

2. Simulation

The simulation is based on a very simple model of a silicon strip detector as illustrated in Figure 1.[1] All parameters of the Si detectors are assumed to be of ideal uniformity, possible edge and surface effects are neglected, and the depletion region is taken to be constant over the whole device. The detector surface is segmented into strip diodes, their pitch P is typically 50μm. The diode strips are parallel to the x axis. A minimum ionizing particle traverses the silicon detector at an angle α. α is measured in the plane defined by the normal to the detector surface (y axis) and a line perpendicular to the diode strips (z axis). Along the particle track on the average 80 e^-–hole pairs are generated per μm of track length.

For simplicity of simulation, the bulk of the Si detector is divided into N slices of thickness D/N, where D refers to the total thickness, D = 200 - 300 μm and $N = D[\mu\text{m}]/10\mu\text{m}/\alpha$. In each slice, a cluster of charge is generated at a random position along the track, the pulseheight proportional to the pathlength and it is chosen according to the Landau distribution function. The width of the initial charge distribution is assumed to be zero, except for delta electrons that range out in the silicon. The electrons and holes drift parallel to the field that is applied transverse to the detector, typically 2-3 kV/cm at the surface, falling linearly with the distance from the surface. The charge cluster generated in each slice diffuses during the drift, the width of the Gaussian charge distribution is proportional to $\sqrt{t}$, where t is the total drift time, $\sigma_i = \sqrt{2D_e t} = D\sqrt{2 \times 0.026/V}$. The total charge collected at the detector surface is calculated as a sum of N Gaussian distributions of different widths σ_i and different center positions z_i. The charge collected on each strip is obtained by integration of the sum of the individual charge distributions over the width of each strip. Random noise is added to the integrated signal on each strip, and the effect of strip to strip amplifier gain variations can be taken into account. Also, the primary charge per strip is smeared due to capacitive coupling between adjacent strips.

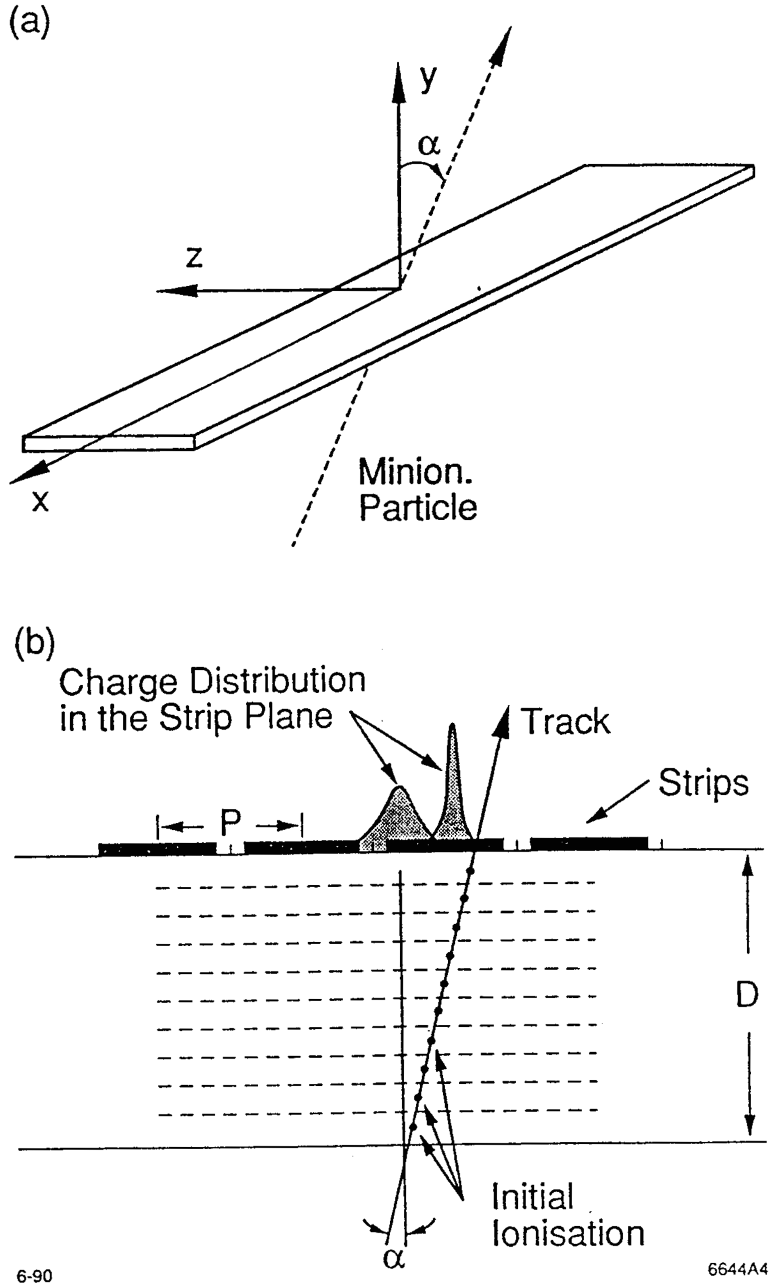


Figure 1. *Geometry of a silicon strip detector:*
a. *overall view, and*
b. *cross section, illustrating the charge generation by a charged track ionizing the silicon.*

The simulated spatial distribution of the charge collected on the Silicon detector surface is shown in Figure 2. The *rms* width is 6.7 μm, though the distribution is clearly non-Gaussian. Figure 3 shows the pulse height distribution for minimum ionizing tracks, it is normalized to peak at 100.

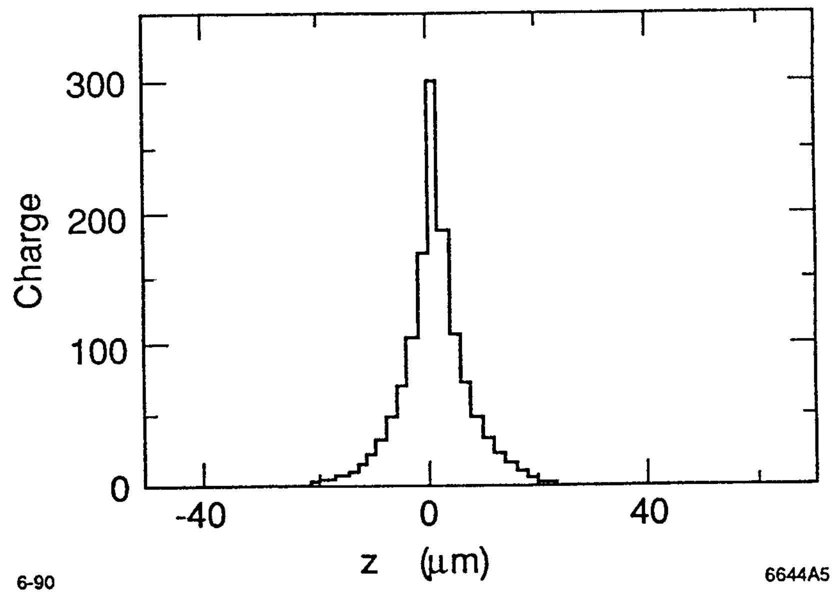


Figure 2. *Charge distribution on the surface of a Silicon detector for a minimum ionizing particles traversing a counter of thickness $D = 300\mu m$ at normal incidence.*

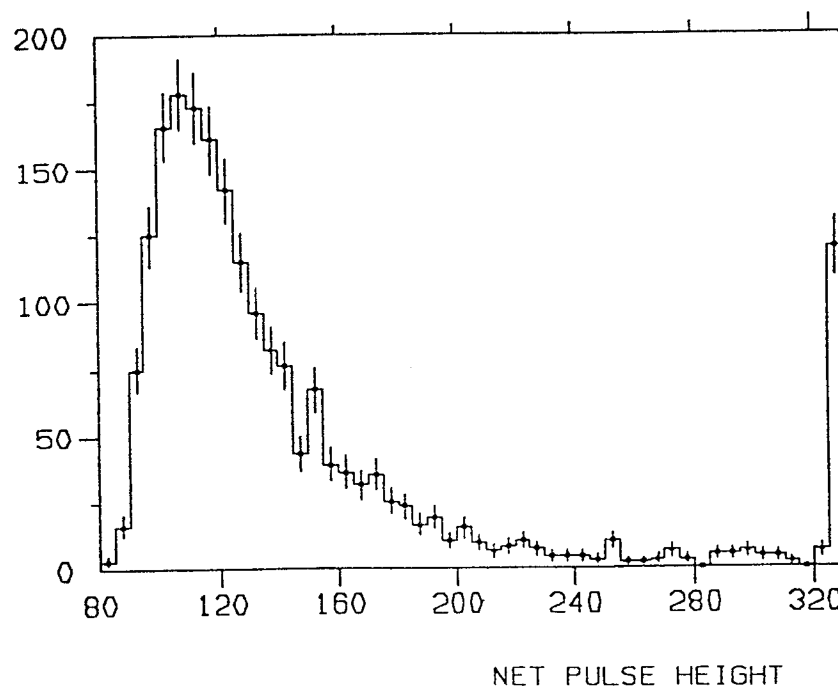


Figure 3. *Distribution in pulse height for minimum ionizing tracks at normal incidence.*

3. Centroid Calculation

To determine the position of impact of a track traversing a silicon strip detector, the observed charge distribution in several adjacent strips is investigated. A number of different algorithms have been used to determine the centroid of the charge distribution. In the following, the results of three algorithms are given.

- The standard algorithm represents an analog read-out: One or several adjacent strips define a cluster if each of them have a pulseheight (proportional to the collected charge) exceeding a given threshold, that is typically set at 2-4 times the rms noise above the zero reading of the channel. The centroid for the cluster is then defined as the pulseheight weighted mean of the strip coordinates:

  $$< x_m >= \frac{\sum_{i=1}^{n} z_i PH_i}{\sum_{i=1}^{n} PH_i}$$

  where the sum is formed over all strips in a cluster, z_i is the coordinate of each strip center, and PH_i is the recorded pulseheight on each strip.

- The second algorithm represents digital read-out: A cluster is defined as above, but the centroid is calculated as the average coordinate of all strips in the cluster, ignoring the pulseheight information.

- The third algorithm is designed to improve the centroid estimate for inclined tracks, it assumes that the angle of incidence is known, and therefore n, the number of strips in the pulseheight cluster, is also known for a detector of a known thickness D and strip pitch P.

  $$n = max\{3 \ , \ D \ \tan\alpha / P\},$$

  α is the angle of incidence. The cluster is then chosen as the group of n strips with the maximum pulseheight sum, $S = \sum PH_i$. A threshold cut of $S > 0.5 PH_{max}$ is applied, here PH_{max} refers to the most probable pulseheight for normal incidence. The centroid for the cluster is then defined as the pulseheight weighted average coordinate of the strips, as cited above.

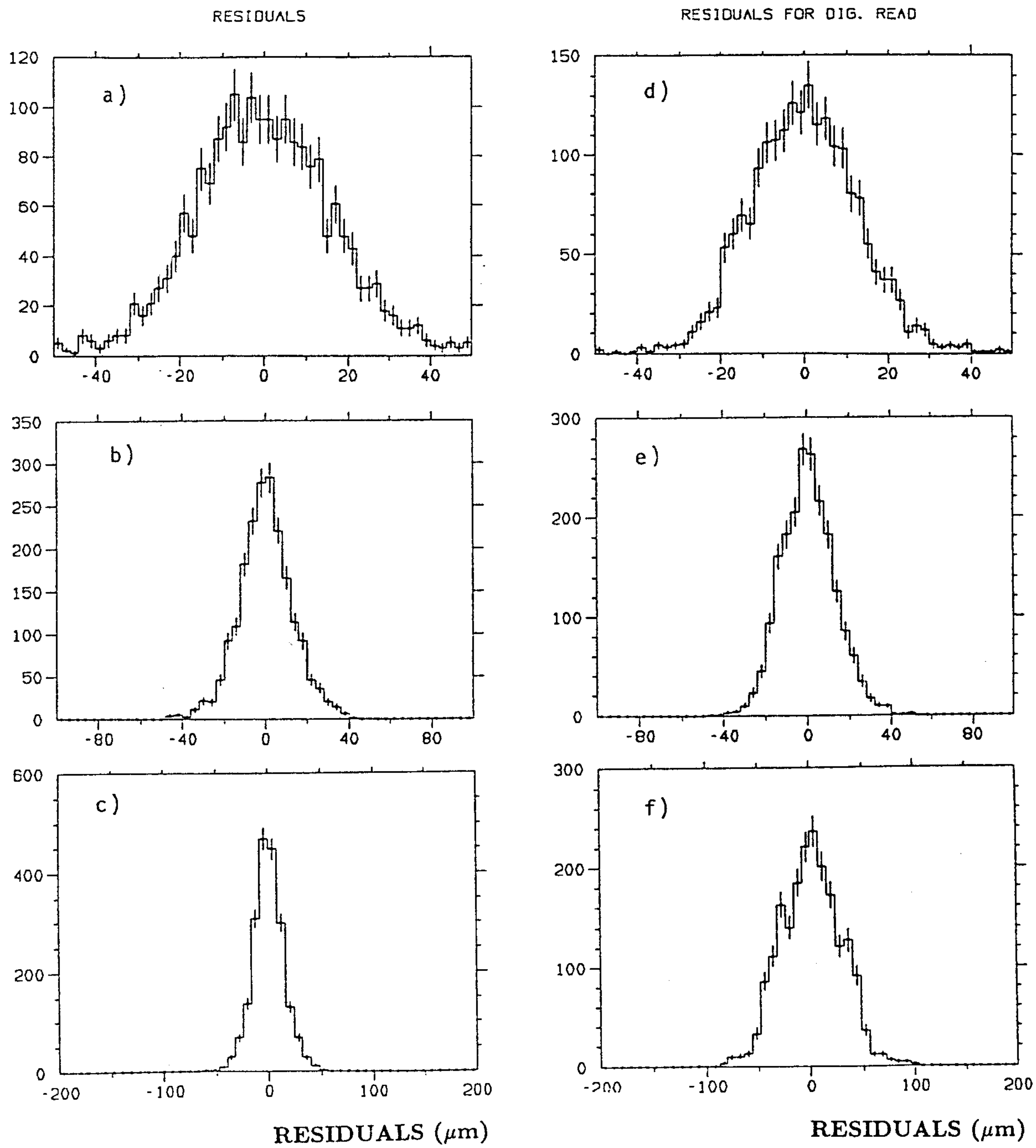


Figure 4. *Residual distribution for tracks at an angle $\alpha = 0.8$ impacting on a detector of thickness $D = 200\mu m$ with a pitch $P = 25, 50, 100\mu m$. The residuals are given for both analog (a-c) and digital (d-f) read-out.*

In Figure 4 residual distributions are presented for tracks impacting at an angle $\alpha = 0.8$ on a detector with thickness D=200μm. Clusters of adjacent channels are selected by a threshold cut of 0.15 PH_{max} on each channel, i.e. 3 σ above zero. for a signal to noise ratio of 20/1. The residual of a track measurement is defined as the difference between the reconstructed centroid and the actual position of impact of the track. The distributions given represent two algorithms described above, a-c) analog read-out with a threshold per channel of 0.15 PH_{max}, d-f) digital read-out with the same threshold per channel. For a signal to noise ratio of 20/1 this threshold corresponds to a level of 3 rms width of the noise above zero. The distributions are remarkably similar in width, though for larger pitch the analog read-out gives substantially better resolution.

One needs to emphasize at this point that the resolution quoted here is derived by a straight forward rms calculation, bin-by-bin, including the non-Gaussian tails. Truncation of these tails or reduced widths that measure only the center of the error distribution will lead to a much smaller increase of the width of the error distributions with angle.

4. Intrinsic Resolution

In the following a number of distributions illustrate the dependence of the error in the centroid determination on various parameters.

4.1. *Strip Pitch*

At close to normal incidence the resolution is roughly proportional to the pitch because the charge spread is small compared to the strip pitch of 50μm or more. Thus the resolution improves with increasing angle because the charge spreads over several strips (see Figure 5a). Above $\alpha = 0.6$ the resolution begins to degrade because the channel to channel pulseheight variations due to Landau fluctuations and single channel noise affect the centroid calculation. Below $\alpha = 0.7$ a resolution of better than 15μm can be obtained with a pitch of 50 μm or less. In Figure 5b the intrinsic resolution for a pitch of 50μm is shown for signal/noise ratios varying from 10 to 1000. It illustrates that a signal/noise ratio of 20 : 1 is adequate. All curves in Figure 5 have been obtained with the third centroid algorithm, i.e. assuming that the angle of incidence is known to 3–5 degrees.

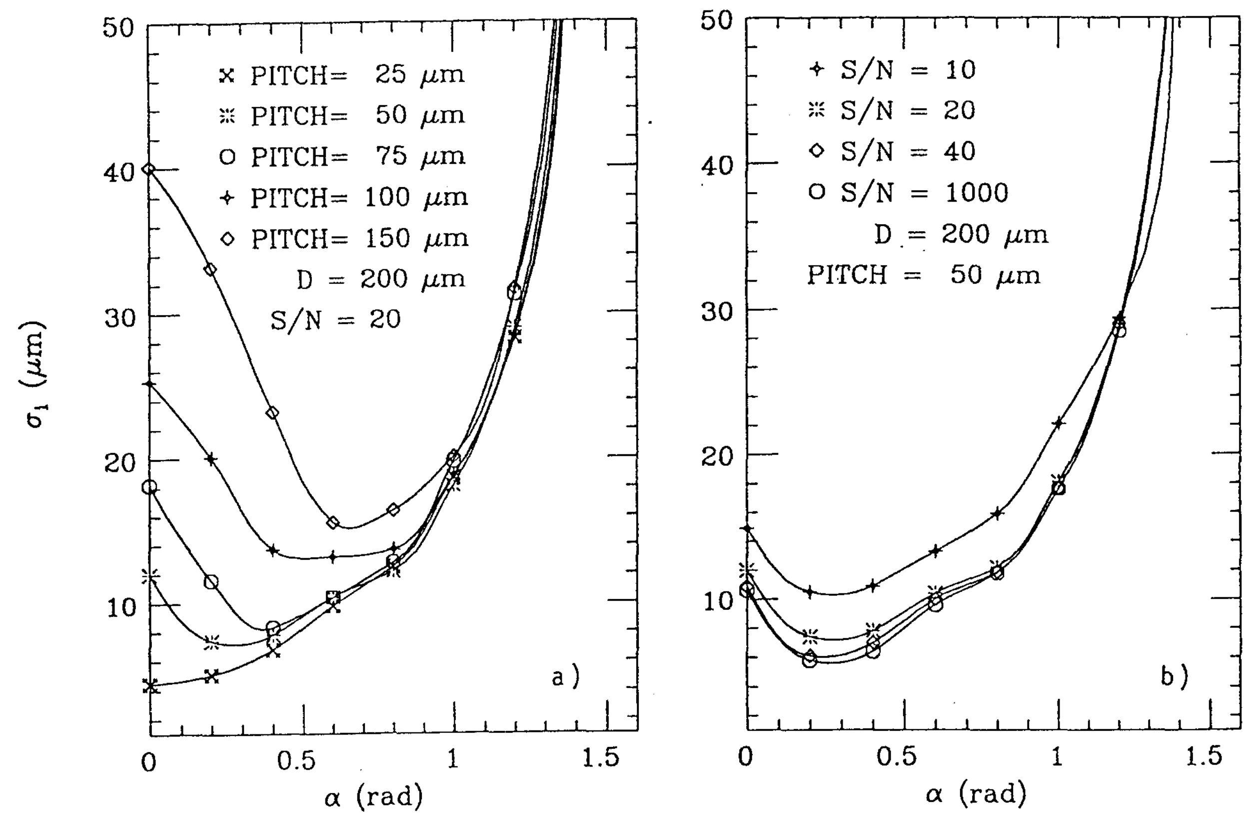
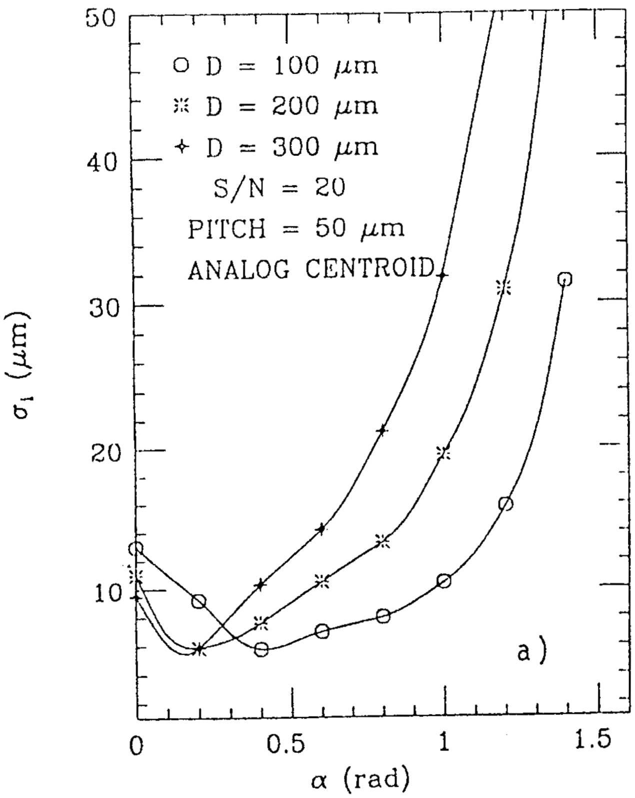


Figure 5. *Intrinsic resolution as a function of α in a detector of thickness $D = 200\mu m$ for* **a.** *different read-out pitch and* **b.** *different signal/noise ratios.*

4.2. Detector Thickness

Silicon strip detectors have been fabricated from 3 and 4 inch diameter wafers, the typical wafer thickness is 300μm. For double-sided detector it may be difficult to reduce the thickness of the material and maintain a reasonable yield. While the signal size increases linearly with the thickness of the depletion region, the charge spread increases also linearly. Figure 6 shows the dependence of the resolution on the detector thickness for different centroid algorithms. Above $\alpha = 0.4$ the resolution degrades linearly with the detector thickness D, and since the number of channels per cluster increases proportionally with tanα, the resolution degrades rapidly above α=0.7.



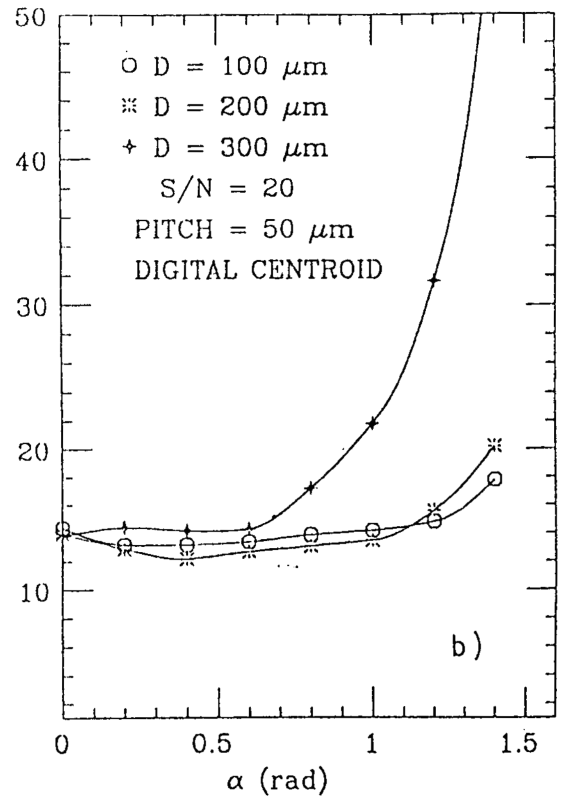
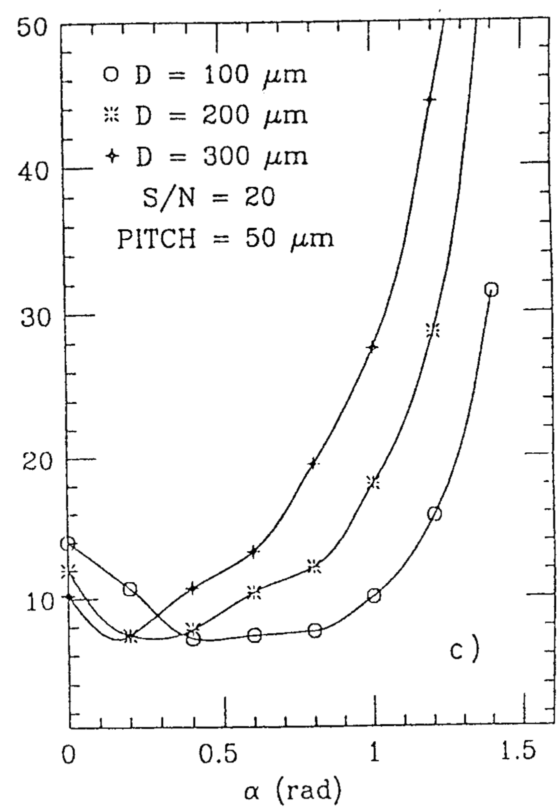


Figure 6. *Intrinsic position resolution as a function of D the thickness of the depletion region of the detector of a fixed signal/noise ratio $S/N = 20$, a pitch $P = 50\mu m$ as a function of the angle of incidence α for*
a. *analog read-out,*
b. *digital read-out and*
c. *definite cluster width.*



4.3. *Centroid Algorithms*

The cluster finding and centroid algorithm can of course be tailored to the specific application. Analog information leads to better resolution at nearly perpendicular incidence of the tracks, digital read-out shows very little dependence of the resolution with angle α for thicknesses $D = 200\mu\mathrm{m}$ and below. Analog read-out leaves room for fine tuning of the algorithms, though the two examples used here give similar results. For large angles α Landau fluctuations affect the centroid finding and lead to large non-Gaussian tails. This causes the increase in the measured rms width of the residual distribution. No attempt has been made to truncate these tails or diminish their effect on the determination of the rms width. The digital algorithm shows much less dependence on angle for depletion thicknesses of 200 μm and below and signal-to-noise ratio of 20, larger thickness or larger noise contributions introduce a very similar results as the analog centroid calculation.

5. Summary

In summary, the effect of non-perpendicular incidence on the position resolution for minimum ionizing track traversing a Silicon strip detector has been studied by Monte Carlo simulation. A considerable rise on the rms width and the tails of the residual distribution is observed. Up to angles of 1.0rad a resolution of 20 μm is obtainable, at larger angles simple cluster finding algorithms may become unreliable in the presence of background, and better signal/noise ratios become essential. In most layouts, this region coincides with larger track fitting errors due to an increase in multiple scattering, and therefore a degradation in position resolution can be tolerated.

REFERENCES

1. M. Turala, Mark II SSVD Internal Note (1986).

Silicon PIN Diode Array Hybrids as Building Blocks for a Vertex Detector at an Asymmetric *B* Factory

S. L. Shapiro

Abstract

Silicon PIN diode hybrid arrays are proposed as the ideal building blocks for a vertex detector at an asymmetric B Factory. The two-dimensional nature of the detector segmentation allows for the maximum in confusion elimination. Fine spatial resolution, on the order of 10μm per layer, is more than adequate to resolve the displaced vertices of beauty and charm decays. A high signal-to-noise ratio allows for the thinning of the detectors, reducing multiple scattering. Time tagging within the detector permits higher background levels than could otherwise be tolerated, and on-board electronics which includes zero suppression and ghost elimination, eases downstream data handling and analysis.

1. Introduction

Pixel devices, in particular silicon PIN diode arrays, are a natural choice for vertex detectors. These devices provide three-dimensional coordinate information with spatial resolution of a few microns, and so provide efficient trackfinding with a minimum number of layers.

An architecture which is appropriate for high-energy charged particle detection is that of a hybrid.[1,2] The charged particle detector, a silicon PIN diode array, and the readout electronics are constructed as two separate silicon chips, each optimized for its specific function. The two chips, indium bump bonded together, form the array hybrid.

The indium bump-bonding process is one in which each diode of the detector array is bonded to an independent amplifier readout circuit on a mating VLSI chip via an array of aligned indium metal bumps that cold weld under pressure to form ohmic contact. This process allows flexibility in the detector and the readout electronics specification. In the case of the *B* Factory, this permits the eventual thinning of both the detector array and the electronics without a change in their design. Figure 1 is a schematic representation of a hybrid detector.

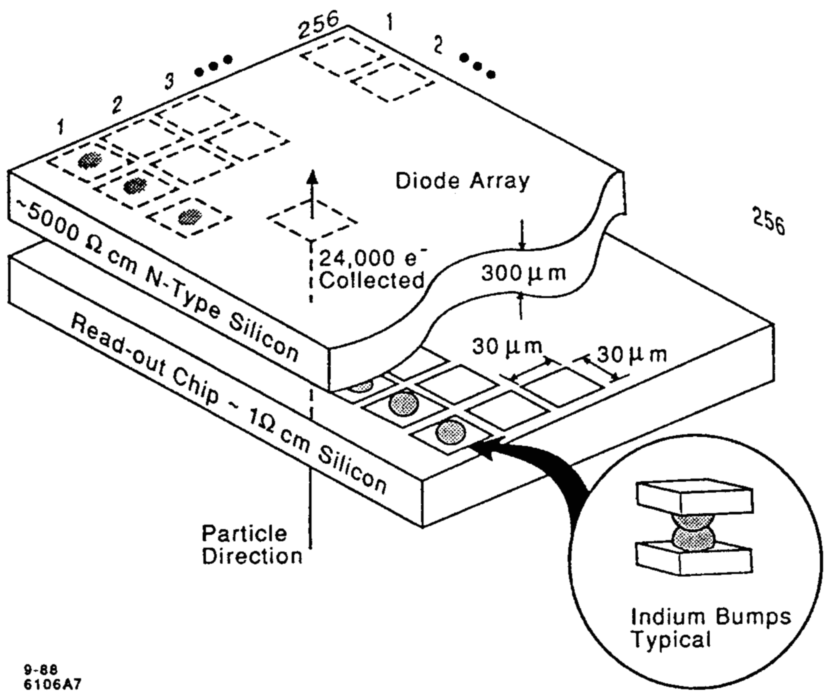


Figure 1. *Schematic representation of a hybrid detector showing the two separate silicon chips and their bump bond interconnects.*

2. Existing Arrays

Development of hybrid arrays has been a goal of the author since 1984. To this end, three hybrid arrays have been designed and fabricated. The high resistivity silicon diode arrays were fabricated by Micron Semiconductor Inc., and the readout arrays by the Hughes Aircraft Company. The indium bump-bonding was also done by Hughes Aircraft Company with bumps measuring less than 15 μm in diameter.

A program of laboratory and high-energy beam line testing of the arrays has been carried out, and is nearing completion. Support for this project has been primarily from SLAC and from generic detector development funds provided by the SSC Laboratory.

The first array, a 10 $\times$ 64 array, having pixels 120 μm on a side, has ten readout channels per array, one for each column. The readout structure allows random access to any pixel. There is no power necessary during the time data is being detected and stored by the array; however, during the read cycle 10 mW (1 mW per channel) is necessary. The readout electronics is an NMOS VLSI circuit which is radiation resistant at the level of 1 MRad of ^{60}Co gamma rays.

The second array, a 256×256 array, having pixels 30 μm on a side, has two readout channels per array. Random access readout is available via row and column shift registers. Data stored in the pixel represented by the intersection of row and column address is presented to the readout node. No power is necessary for storing data, and only 2 mW is necessary for readout. The readout electronics is an NMOS VLSI circuit which is not particularly radiation-hard.

The third array is similar to the second, except that the readout electronics, though identical to the second array, is implemented in a PMOS VLSI circuit. This variation was found necessary to preserve the option of using these arrays at room temperature. The PMOS array is more appropriate to collect the holes rather than the electrons generated by the passage of the charged particle through the silicon. The PMOS circuit has a much larger dynamic range of operation, and dark current integrated over the time necessary to read the 32,768 pixels on each readout node no longer develops voltages that drive the circuit to its supply rails.

A description of the hardware necessary to read out these arrays is published elsewhere,[3,4] as are the results of preliminary laboratory tests.[5]. In brief, alpha and beta sources have been used on the first two of the above arrays, confirming their ability to detect charged particles. The noise level has been measured to be less than 300 electrons rms at room temperature, resulting in a signal-to-noise ratio of about 80:1. The spatial resolution for those particles which share their charge between two pixels has been found to be less than 2 μm.

In August 1990, arrays of the third type were placed in a 250 GeV/c pion beam at Fermilab. These results, unpublished at this time due to the preliminary nature of the data, show clearly the detection of these high energy pions, with a signal-to-noise ratio in excess of 50:1 at room temperature.

Figure 2 is a three-dimensional plot of a single frame of data showing a number of particles traversing the array.

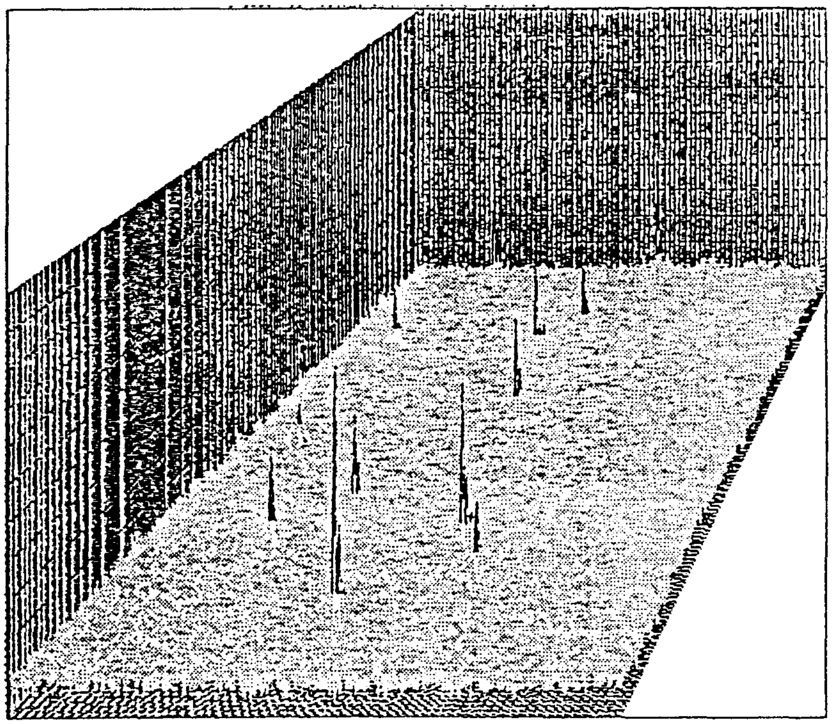


Figure 2. *A three-dimensional plot of a number of minimum ionizing particles from the Fermilab test beam incident on a detector array, demonstrating excellent signal-to-noise and the power of two-dimensional arrays to eliminate confusion in complex events.*

3. The Proposed Array

The hybrid array being designed and developed for use at the SSC has many features which are useful and exciting for the asymmetric B Factory. A summary of the design goals is presented in Table 1.

There are actually three phases to the design effort. The first phase requires the implementation of the essential elements of the design in a FORESIGHT chip. FORESIGHT is a fabrication capability similar to MOSIS offered by the Orbit company, which allows circuits to be fabricated on a community wafer for quick turnaround. The actual pixel design, shown in Figure 3, is incorporated into a 32×64 array and only minimal external support circuitry is included on the chip itself. The goal is to confirm that the time between the arrival of the charge, and the DETECT signal arriving at the periphery of the chip, is about 50 ns. This feature (called time stamping) requires a discriminator within each pixel, and is one of the more exciting features of this design. One will also be able to test details of the analog circuitry of the pixel, such as linearity, a noise level of less than 300 electrons rms, and pixel-to-pixel variations. All array operations will be under external contro, rather than being generated internally, but confirmation of the random access nature of the Read/Write signals can be achieved. The FORESIGHT chip was received on September 20, 1990, and is undergoing tests at this time.

SPECIFICATION	PROTOTYPE	FINAL GOAL
Pixel size	50 μm × 150 μm	50 μm × 50 μm
Array size	128 × 64	256 × 256
Noise	$< 200\ e^-$	$< 200\ e^-$
Time stamping	YES	YES
Time resolution	50 ns	16 ns
Ghost elimination	NA	YES
Nearest neighbor read	NA	YES
Readout time	10 μs	1.5 μs
Radiation hardness	NO	10 MRad
Power	NA	0.1-1.0 W/cm^2
Technology	0.8-1.2 μm	0.5 μm SOS/CMOS
Clock speed	NA	1 GHz

Table 1. *Parameters of proposed PIN hybrid array*

After successful operation and testing of the FORESIGHT chip, the PROTOTYPE chip design will proceed. The pixel cell will be identical to that in the FORESIGHT chip, differing only in that the feature allowing electronic testing of the pixel will be removed: to test the PROTOTYPE chip one will have to hybridize it to an array of PIN diodes. The support circuitry will be implemented so that one can read out only interesting data. The self-timing operation will be tested, confirming the time-stamping and the time resolution with the chip internally generating all of its control signals. The reading of interesting pixels will be controlled by a single TRIGGER command. The design of the support circuitry is complete and is described elsewhere.[6]

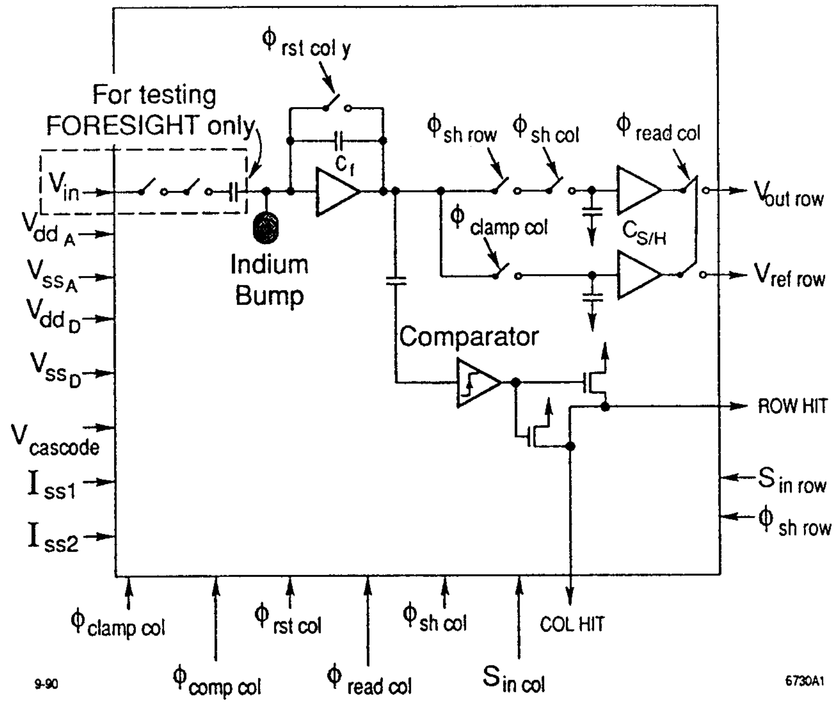


Figure 3. *Schematic block diagram of the pixel design employed in the FORESIGHT chip, including the input test circuit.*

The third phase of the program is the reduction in size of the pixel to 50 μm × 50 μm by the use of 0.5 μm minimum feature size SOS/CMOS technology. The FORESIGHT chip employed 1.2 μm double-metal technology; however, the reduction in minimum feature size does not necessarily guarantee a smaller pixel, as many features of the circuit (such as capacitors and metal line width) do not scale with this parameter. Lastly, the implementation of radiation hardening techniques, such as thinning of the gate oxide to achieve the 10 MRad specification, will be implemented. Additionally, changes in circuitry to further harden the circuit will be attempted at this stage. The radiation hardness specification is certainly not the driving issue for the *B* Factory that it is for the SSC, and this step can possibly be relaxed to 1 MRad in this case.

A few details are noteworthy concerning the pixel design in the FORESIGHT chip. It has 21 transistors, 5 capacitors, and 17 lines comprised of power rails, clocks, biases, and outputs per pixel. Theoretically, the noise will be below the 300 electron limit, the power consumption is less than 20 μW per pixel, dead time is about 400 ns, the feedback capacitor is designed to be about 10 fF and the open-loop gain of the first stage is about 200. These parameters are being put to the test this month. A calculation of the fill factor—the area of the chip containing pixels compared to the total area of the chip—for the final 56 × 256 array is 94%.

4. Mechanical Design Issues

Pixel array hybrids differ from microstrips in that they are not self-supporting as are microstrip detectors, but must be supported on a substrate. The substrate must provide support, not only for the arrays, but for the their support circuitry such as traces, preamps, ADCs, output multiplexers, and DSPs (if necessary). Two candidates for this substrate material are foams made of either silicon carbide or boron carbide. Boron carbide has the longer radiation length, 20.8 cm compared to 10.1 cm for silicon carbide. Both materials can be foamed to 3% of their density while retaining machineability. They are both electrical insulators. Silicon carbide has superior thermal properties when compared to silicon, thus allowing one to consider other than room temperature operation. A number of mechanical design studies are under way[6] investigating the proper use of these foams as support material for SSC vertex detectors. These studies also deal with cooling requirements at the level of 1 W/cm^2.

An embryonic mechanical design study of a silicon vertex detector for the B Factory has been reported.[7] The pixel detectors described in this note and the use of either of the foam materials mentioned would provide the basic material for such a vertex detector. However, much mechanical design work remains to be done.

5. A Design Challenge

A number of colleagues have reached the conclusion that in the measurement of CP asymmetry parameters, one should not expend heroic efforts to design a vertex detector having the smallest possible inner radius.[8] The gain in overall precision when one takes into account large-angle tracks and the multiple scattering in material interior to the first detector array does not warrant the effort. This is true, up to a point.

However, there are interesting decay channels that contain soft pions that will be unavailable to analysis unless these soft pions and the decay vertices from which they originate are recognized. This can be accomplished if we place the vertex detector within the beam vacuum. Pixel detectors can sustain orders of magnitude more background radiation than can microstrip detectors, due to their inherent two-dimentional nature, the small size of an individual pixel, and the time stamping of each hit with its time of arrival, and they can be fabricated radiation hard to 10 MRad. If one designs a Faraday cage between the vertex detector and the beam, to provide a continuous path for the beam image charge, and if this Faraday cage has less material than a conventional beam pipe, then one can press the vertex detector inward radially to the limits set by the machine designers. This would open up new physics channels for study and increase the resolution in impact parameter as well. Philosophically one is doing the best job possible of finding displaced vertices, while separating this function from that of tracking by placing the actual beam pipe between the vertex detector and the tracking detector.

In designing the Faraday cage, one can combine the cooling problems presented by both the vertex detector electronics and the ohmic and RF heating of the Faraday cage by the beam. The heating of a beam pipe by the beam is proportional to the resistivity of the beam pipe material. Thus, cryogenically cooling the Faraday cage would reduce its resistivity, and remove this source of heating at its source.

The placement of a vertex detector within the beam vacuum may not be the most prudent first step in the design of a new machine. However, this design takes full advantage of the features of pixel arrays which have not heretofore been available, and as such is deserving of further professional study.

6. Conclusions

Silicon PIN diode hybrid arrays are serious candidates as building blocks of a vertex detector for an asymmetric B Factory. The two-dimensional nature of the detectors provide the optimum in confusion elimination and ease in downstream software analysis. The high signal-to-noise ratio allows for the thinning of the detector arrays reducing multiple scattering. Time stamping allows the separation of background-induced hits from those related to the event of interest. Much additional work is needed to understand completely the properties of these detectors and to generate a viable vertex detector design based on them.

Acknowledgments

The testing of the hybrid arrays was done in collaboration with J. Arens and G. Jernigan of the University of California Space Sciences Laboratory. The design and fabrication of the SSC arrays is being done in collaboration with the Hughes Aircraft Company and physicists forming the Pixel Detector Development Collaboration for work at the SSC. Useful and interesting discussions with Walter Toki and Andrew Hutton are also acknowledged with thanks.

References

1. S. Shapiro and T. Walker, "The Microdiode Array—A New Hybrid Detector," SLD—New Detector Note 122, (1984).
2. S. Gaalema, *IEEE Trans. Nucl. Sci.* **NS–32**, 417 (1985).
3. S. Shapiro, W. Dunwoodie, J Arens, J. G. Jernigan and S. Gaalema, "Silicon PIN Diode Array Hybrids for Charged Particle Detection," *Nucl. Instr. Meth.* **A275**, 580 (1989).
4. S. Gaalema, G. Kramer, S. L. Shapiro, W. Dunwoodie, J. Arens, and J. G. Jernigan, "Silicon PIN Diode Hybrid Arrays for Charged Particle Detection: Building Blocks for Vertex Detectors at the SSC," *Proc. Int. Industrial Symposium on the Supercollider*, New Orleans, LA, 1989, p. 173; SLAC–PUB–4942.
5. S. L. Shapiro, J. G. Jernigan, J. F. Arens, "Progress Report on the Use of Hybrid Silicon PIN Diode Arrays in High Energy Physics," Invited talk presented at the Vth Int. Conf. on Instrumentation for Colliding Beam Physics, Novosibirsk, USSR, 1990; SLAC–PUB–5212.
6. G. Kramer, D. Nygren, E. Arens, G. Jernigan, and S. Shapiro, Summary Report for FY 90 of the Pixel Detector Development Collaboration submitted to the SSC Laboratory on September 1, 1990. Contributions have been prepared for the October 15, 1990 meeting of the SSC Laboratory in Fort Worth, Texas which detail the design of the circuits and the results of the beam line testing of August 1990.
7. Vera Lüth, "Design of a Silicon Vertex Detector," $Ba\bar{B}ar$ Note #55.
8. F. Le Diberder, W. Toki, and M. Witherell, "The Effect of Beam Pipe Radius on the Measurement of *CP*-violating asymmetry at an asymmetric B Factory," $Ba\bar{B}ar$ Note #45.

B Factory Forward-Backward Vertex Chamber Based on the Microstrip Gas Avalanche Chamber

J. Va'vra

Conventional vertex chambers may have difficulty in providing good measurements for tracks at small angles. For *B* physics, it is important to measure even tracks emitted close to a beam pipe. Obviously, the silicon vertex chamber will have a few layers placed in the forward-backward region,[1] but perhaps measuring 50–100 points in a gaseous device with 40–60 μm resolution per point has some competitive edge compared to fewer, more precise, measurements in a silicon detector. The device discussed herein is based on the microstrip idea[2–4] with strips deposited on glass by a vacuum evaporation technique. The device operates in the proportional mode. The proposed device would improve the measurement of the polar angle at small angles, constrain the extrapolated z coordinate to better than 200 μm at the vertex point, and the azimuthal coordinate to a precision of several hundred microns, by measuring a charge induced on strips. The polar angle coverage is between 5 and 30 degrees, assuming that the device is placed 20 cm away from the interaction point, the beam pipe has 2 cm radius, the chamber is 5 cm long in z direction, and that it has about 10 cm drift in radial direction. The device could also measure dE/dx in the nonrelativistic region,[5] which may help to identify slow recoil particles. This device would require an R&D effort to develop appropriate vacuum deposition techniques on curved surfaces, and to study the operation of microstrip chambers, which have not yet been tested in actual experiments.

A. Oed has shown[2] that one can amplify on thin 10 μm wide strips evaporated on a glass substrate. The thin strips are surrounded by two nearby cathode strips which are biased to control the gain on the thin strip, i.e., most of the field lines are between the thin anode strip and the cathode strips (see Figure 1). In this way, the positive ions travel only a short distance and do not enter the main drift volume.

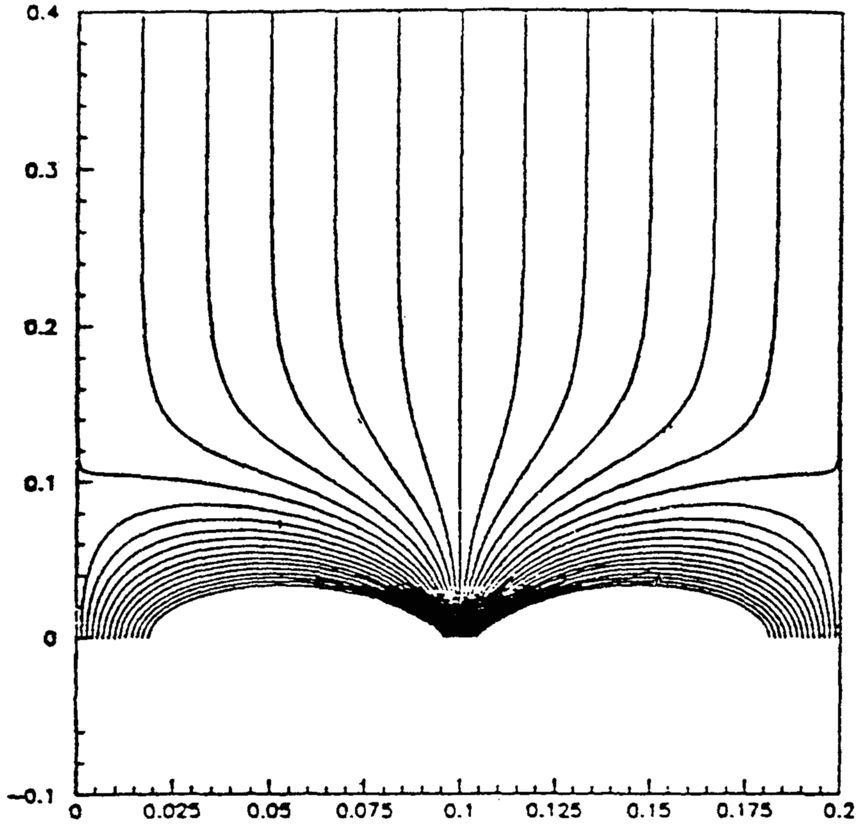


Figure 1. *Field lines in the Oed chamber, with* $V_{drift} = -1000\ V$, $V_{cathode} = 0\ V$, $V_{anode} = +60\ V$ *and* $V_{back} = -500\ V$. *The horizontal scale is millimeters.*

Such a device is, in principle, able to operate at a high rate. However, two effects can spoil this beautiful idea. First, the positive ions can land on the glass and alter the field. This can be prevented by small gaps between strips and appropriate biasing of the "back" electrode (which can be used for the second coordinate readout) to force the field lines away from the glass. The second problem has to do with a very high surface gradient on the glass (higher than 10 kV/mm!) which tends to limit the operating wire gain to 10^4 or so. It remains to be proven how these gradients can be mantained in a real device under all sorts of conditions. Nevertheless, it has been shown by others[3,4] that the idea is worth pursuing. A microstrip chamber was installed in a test beam at CERN, using as reference the spatial coordiantes provided by a pair of silicon strip detectors. The plane of the device was perpendicular to the beam. A resolution of about 40 μm was achieved with a track sample size of 200 μm per anode strip in a 90% Xe + 10% DME gas mixture. In this chamber, all anodes were connected together and the cathode strips were connected to charge-sensitive amplifiers. The back electrode was not made of strips. This chamber demonstrated high rate capability, good cathode signals, and a good ^{55}Fe signal resolution (18% FWHM) (see Figure 2).

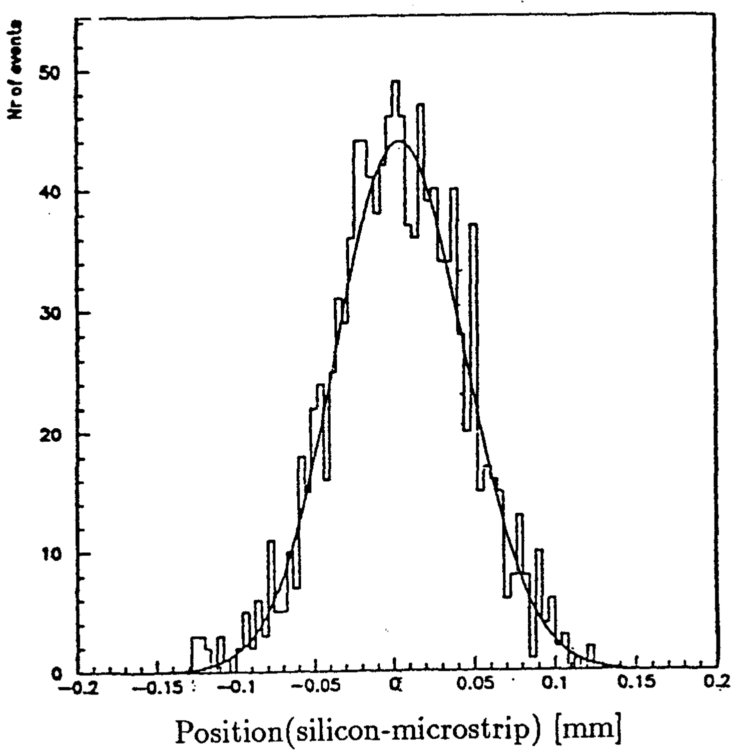


Figure 2. *Position resolution obtained with the 90% Xe + 10% DME gas mixture. The resolution is $\sigma = 42\ \mu m$.*

One possible application of this idea is for the intermediate central region tracker placed between the silicon vertex detector and the main drift chamber. Reconstructing induced charges on the cathode strips, one could achieve good resolution over some fraction of the solid angle. Another possibility is as a forward-backward vertex chamber. In this application, one would measure not only the cathode strip charges, but also drift time to the anode strips. To achieve a resolution 0f $\sim 60\mu m$, it is necessary to increase the sample size to about 1 mm and to use a gas with a large electron yield, such as 90% Xe + 10% DME.

Figure 3 shows our idea of a possible application for the forward–backward vertex chamber geometry to the *B* Factory experiment. One builds the chamber on a glass cylinder of about 2 cm radius and 15–20 cm length. This cylinder is supported by the beryllium beam pipe. Although there are several alternatives construction methods, we mention only one possibility. The first operation is to deposit the back strips on a curved glass surface. This would be done in a vacuum tank with a fixed metal source and a rotating glass cylinder with a stainless steel mask attached to it. The back strips are evaporated in vacuum by heating the metal. The next operation is to deposit a layer of glass over the back strips, again by the vacuum evaporation technique, working on the rotating glass cylinder without any mask except at one end to allow future contacts to the strips. The exact thickness of this layer has to be determined experimentally, but thinner layers will yield larger induced signals on the back strips. We next deposit 10 μm wide anode and 50–100 μm wide nearby cathode strips. These strips are deposited with the help of a mask fixed with respect to the rotating cylinder: the rotation mechanism has to be very precise in order to achieve a few micron accuracy, and must function in a very good vacuum. The pattern of these strips and a voltage choice define the track sample size of about 0.5-1 mm per anode measurement. This choice must be studied in detail from the point of view of efficiency, drift arrival tail, *etc.* The strips must be cleanly deposited, with no residual deposition between strips, in view of the danger from very high surface gradients; this may be the biggest challenge of this type of construction. To make the metal stick to the glass surface, one has first to clean it. This is done in the same tank by argon ion bombardment. The chamber needs end flanges with a strip pattern and an outer cylinder to shape the field, aswell as a feedthrough electrode to collect the anode charge and to define voltages on the cathode strips. The contact to these strips must also be developed; one possibility is to use a conducting epoxy. The small size of the device lends itself to work under a microscope. The feedthrough electrode must have field shaping strips to define the drift field properly over a total drift distance of about 10 cm.

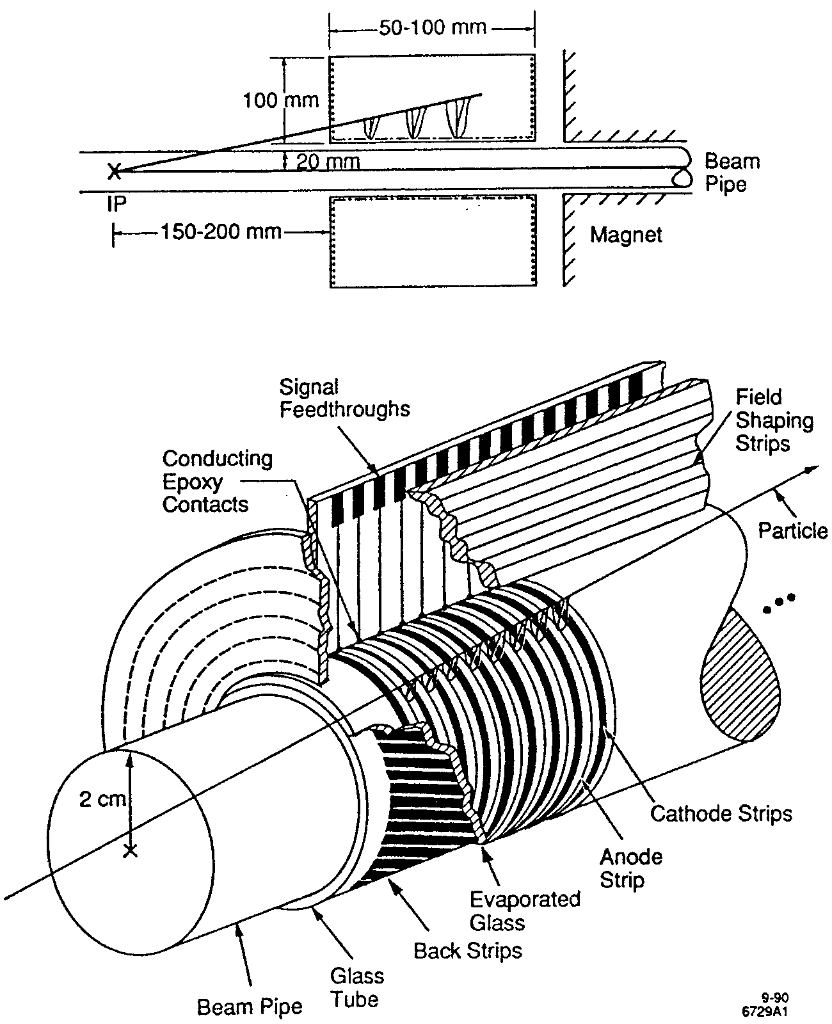


Figure 3. *Conceptual design of a forward-backward microstrip gas avalanche chamber.*

There has recently been an attempt to develop amplifying structures on a Kapton film,[6] which would be advantageous. Considerable R&D effort is required to determine whether such plastic materials are suitable.

We have not pushed the design of this chamber to its full limit because one has to first establish a clear need for such a device from the viewpoint of physics.

References

1. V. Lüth, *Ba$\bar{B}$ar* Note #55.
2. A. Oed, *Nucl. Instr. Meth.* **A263**, 351 (1988).
3. F. Angelini, *et. al.*, *Nucl. Instr. Meth.* **A283**, 755 (1989)
4. F. Sauli, Novosibirsk Instrumentation Conference, 1990.
5. J. Va'vra, see Particle Identification section in these Proceedings.
6. D. Mattern, M. C. S. Williams, and A. Zichichi, CERN/EF 90–4.

A Forward TPC for an Asymmetric B Factory Detector

Guy Wormser

Abstract

A small TPC surrounded by a layer of silicon strip detectors is proposed to improve the z impact parameter resolution in the forward region of an asymmetic B Factory detector. Substantial gain can be achieved over an extended barrel-type Si layout provided the TPC entrance window is less than 0.1% r.l. thick.

1. Introduction

In the forward region, the z impact parameter resolution is dominated by multiple scattering in the beam pipe and in the first layer of the microvertex detector. Therefore, it may be advantageous to replace the Si vertex detector by a gaseous detector with an extra-thin entrance window. Furthermore, in the case of the cylindrical geometry, the position resolution in the silicon degrades at small angle and the angular resolution becomes of the order of 1 mrad. A gaseous detector can improve this resolution with only a 10 cm lever arm. We propose here to use a small Time Projection Chamber (TPC) in conjunction with a layer of Si detectors on the outside. The gaseous detector is expected to be resistant to the high radiation level foreseen in this area.

2. Set-Up of the Forward Hybrid Spectrometer

The system geometry (Figure 1) is described by five parameters:

- Beam pipe radius: r_{beam}
- Inner radius of the TPC: r_{in}
- Outer radius of the TPC: r_{TPC}
- Coordinate along the beam of the TPC HV plane: z_{TPC}, and
- TPC length: l_{TPC}.

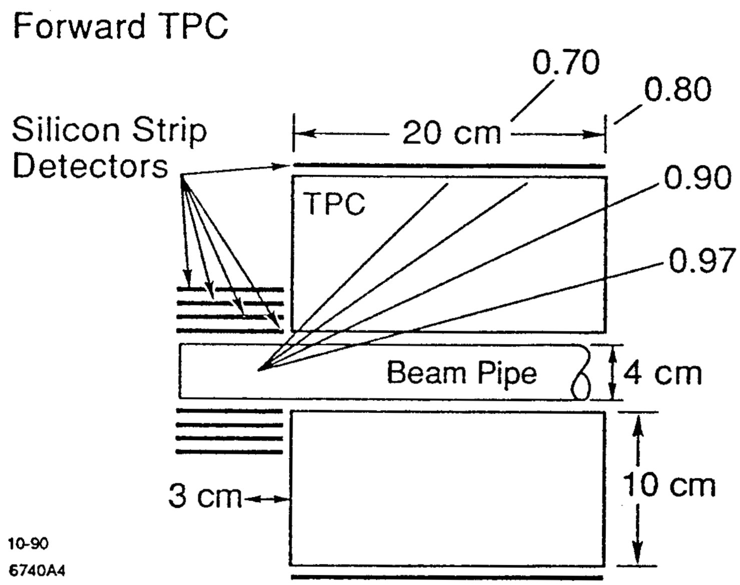


Figure 1. *Schematic layout of the forward TPC.*

We choose the following values: r_{beam} = 2 cm, r_{in} = 2.5 *cm*, and $z_{TPC} = 3$ cm. The beam pipe thickness has been assumed to be 0.1% *r.l.*. (Should this value not be reached, the resolution will be degraded such that the usefulness of covering this region with a microvertex detector is doubtful.) The parameter z_{TPC} determines the angular range covered by the TPC. It does not have a significant impact on the resolution. With the value chosen here, the TPC coverage begins at 40° or $\cos\theta = 0.8$.

The radius r_{TPC} has to be large enough to assure a sufficient lever arm and small enough to ensure that the impact parameter resolution is not dominated by the extrapolation error from the precise point measured in the outer Si layer. As will be shown in the next section, the optimum is around r_{TPC} = 10 cm. l_{TPC} determines the smallest polar angle θ covered by the angular acceptance of the detector. It cannot be extended too far for two reasons: the presence of magnetic elements along the beam pipe causing magnetic field inhomogeneities and the increase of the contribution to the resolution of the longitudinal diffusion. However, the diffusion varies as $r(l_{TPC})^{1/2}$. We have chosen l_{TPC} = 18 cm, to cover up to $\cos\theta = 0.98$.

3. The TPC parameters and resolution

We will describe below the TPC internal parameters from which we will estimate our resolution. We will discuss here only z or θ resolution as measured using the drift time information by an array of wires in the endplate. There will be of course pad information available from the cathode plane which will be very useful to provide information in the $r - \phi$ plane.

The entrance window thickness, t_{TPC}, is a crucial parameter. From a purely mechanical view point, it can be chosen very small since a small depression in the chamber will inflate the entrance window. The tension of this layer will be ensured by a small rigid ring situated at the corner of the chamber. However, this inner layer has to ensure the uniformity of the electric drift field and hence some conductor has to be used. Possible choices may include graphited mylar, silver painted mylar, etc. The study of the electrostatic properties of such a simplified cage have to be studied in detail. We have used a simple model to describe the TPC resolution on each wire. We assumed it is the quadratic sum of three terms:

- Longitudinal diffusion: σ_{drift}
- Error on the track position due to the gap width: σ_{gap}
- Contribution of the drift time measurement: σ_{el}

There are therefore several important parameters:

- Longitudinal diffusion coefficient: C_d
- Number of ionized electrons per mm: N_{e^-}
- Gap between wires: Gap

The error on the z coordinate is given by the error in the drift time

$$\sigma_{\mathrm{drift}} = C_d \times l_{\mathrm{drift}}^{1/2} \quad ,$$

where l_{drift} is the drift distance on each wire

$$\sigma_{\mathrm{gap}} = 1/2 \times \frac{\mathrm{Gap}^{1/2} \times \cos\theta}{N_{e^-}^{1/2} \times \sin^{1/2}\theta} \quad .$$

We assumed in the above formula that with sufficiently fast electronics, the time arrival could be measured for the first electrons and for the last electrons to gain a factor of two compared to the time-integrated resolution. With fast sampling (50 or 100 MHz FADC), the error contribution of the electronics, σ_{el}, can be reduced to 100 μm.

We have studied the influence of these parameters, taking as starting point values obtained in the DELPHI TPC. Our basic parameters are: $C_d = 3$ mm/$\sqrt{m}$; $N_{e^-} = 7.5$ e^-/mm; Gap = 1 mm; $t_{TPC} = 0.1\%$ *r.l.*. For C_d, the DELPHI value is 4 mm/$\sqrt{m}$, but by changing the gas and/or raising the drift field, a value of 3 mm/$\sqrt{m}$ can be reached rather easily.

3.1. *Effect of r_{TPC}*

Figure 2 shows the effect of the outer radius of the TPC on the resolution of the impact parameter b_z. As expected, for $r_{TPC} = 8$ cm, the angular resolution is not good enough, while going to 14 cm slightly improves the resolution.

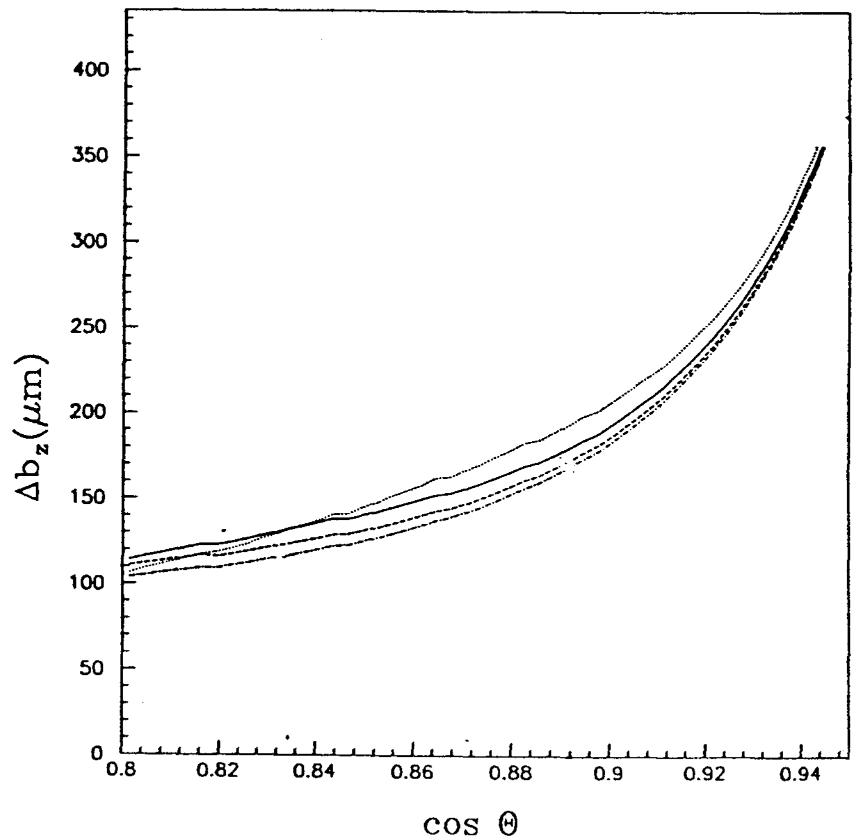


Figure 2. *Effect of the TPC outer radius r_{TPC}: r_{TPC} = 8 cm: dotted line; r_{TPC} = 10 cm: solid line; r_{TPC} = 12 cm: dashed line; r_{TPC} = 14 cm: dash-dotted line.*

3.2. Effect of the gap width

Figure 3 shows the deterioration of the resolution with a 1.5 mm gap compared to a 1 mm gap or a 0.75 mm gap. For mechanical and electrical reasons, going below 1 mm gap width seems unrealistic.

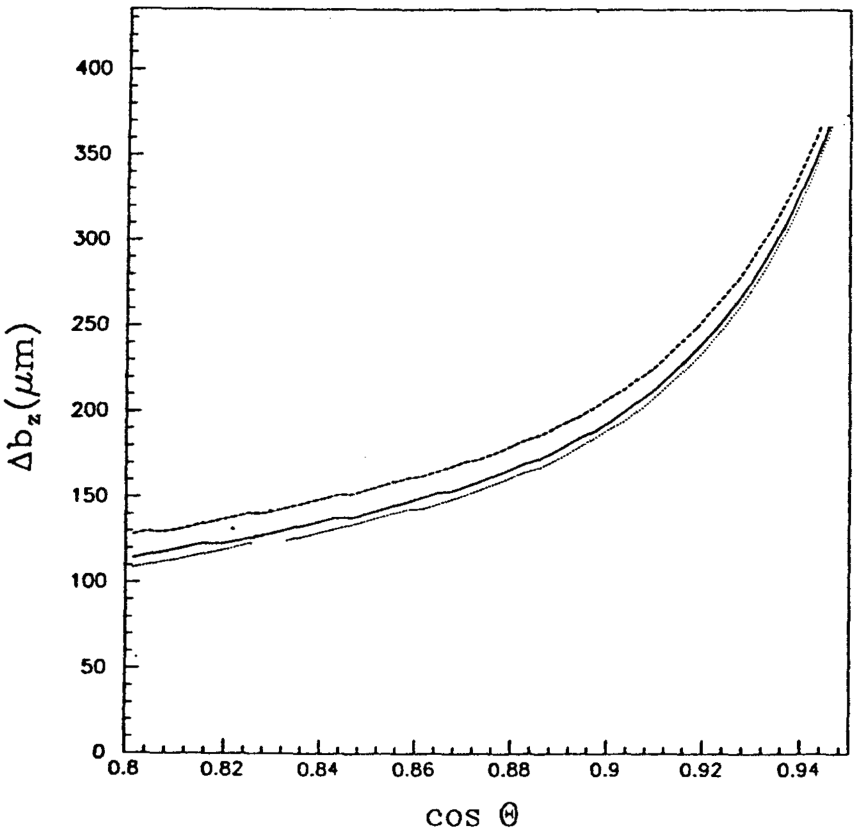


Figure 3. *Effect of the gap width gap = 1.5 mm: dashed line; gap = 1.0 mm: solid line; gap = 0.75 mm: dotted line.*

3.3. Effect of the gas properties N_{e^-} or C_d

Between $\cos\theta = 0.8$ and 0.9, some improvement can be achieved by using a higher electron yield (going to a pressurized gas for example). Doubling the e^- yield can improve the resolution from 110 μm to 90 μm. Some work has to be done to investigate the different methods to get a higher electron yield in the least disturbing way.

Figure 4 shows the effect of the diffusion coefficient that can be varied by changing the gas or the drift field E. (C_d varies roughly as $1/E^{1/2}$.) The solid line corresponds to $C_d =$

4.0, the value used in the DELPHI TPC. It is clearly preferable to work at $C_d = 3.0$ or even 2.0, but this value will be difficult to reach.

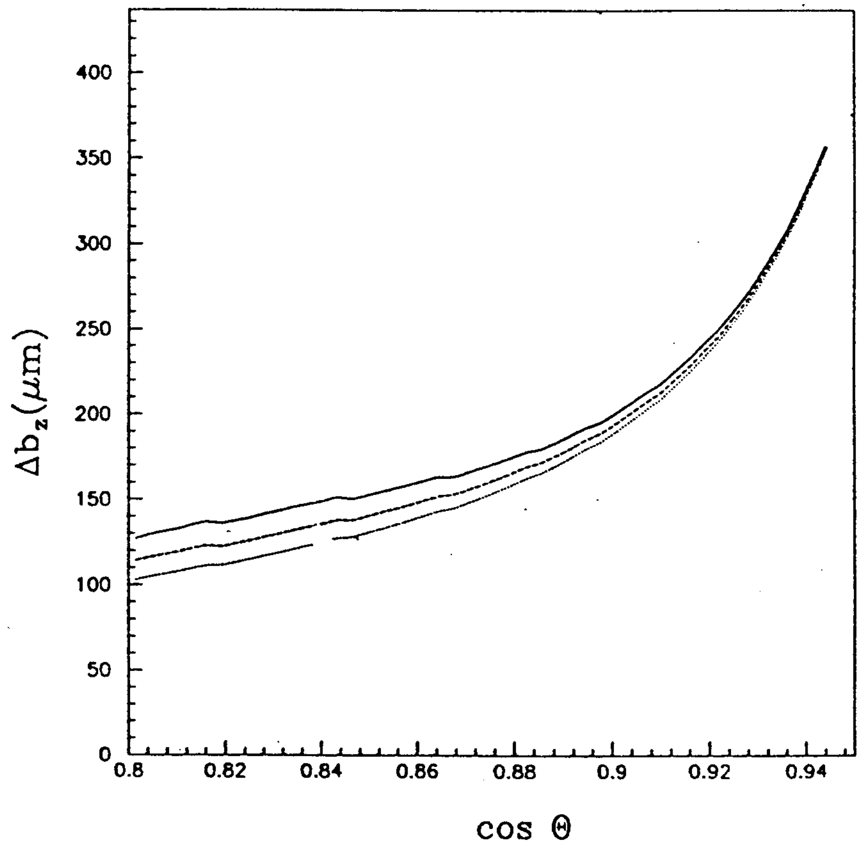


Figure 4. *Effect of the diffusion coefficient C_d: $C_d = 4$: solid line; $C_d = 3$: dashed line; $C_d = 2$: dotted line.*

3.4. Effect of entrance window thickness

Figure 5 shows the comparison between the pure Si layout, the 0.1% *r.l.* case and the 0.05% *r.l.* case. All three set-ups give comparable resolution at $\cos\theta = 0.8$, but at $\cos\theta = 0.92$, the resolutions are respectively 270 μm, 240 μm, and 210 μm. The dotted line corresponds to 0.05% *r.l.* case with high electron yield, $N_{e^-} = 15$. This shows that even though the beam pipe thickness has been taken of 0.1%, there is considerable interest to go well below this number for the TPC entrance window and if one succeeds to go below 0.1% *r.l.*, the resolution of a pure Si layout can be improved by a factor of 1.4. It is very important to note that the benefit for the physics is a highly non-linear function of the resolution improvement. For instance, the number of events surviving a 3 σ cut would be enhanced by a factor of three.

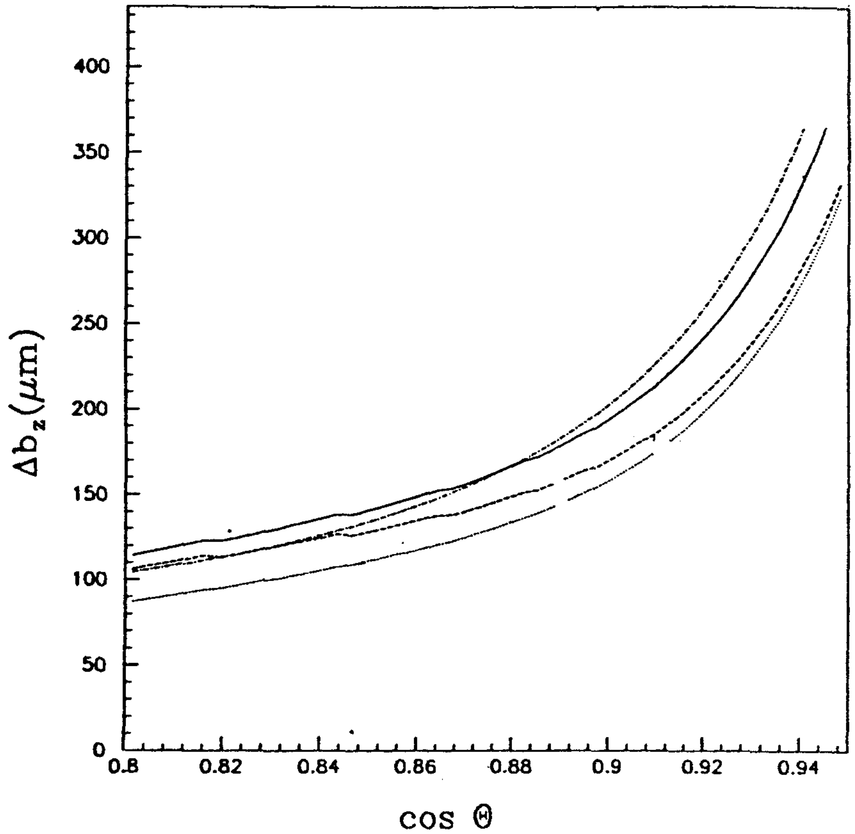


Figure 5. *Effect of the entrance window thickness:* t_{TPC} = 0.1%*: solid line;* t_{TPC} = 0.05*: dashed line;* t_{TPC} = 0.05 *with* N_{e-} = 15*: dotted line. The dash-dotted line corresponds to the pure Si layout, with a Si thickness of* 0.15% *r.l.*.

3.5. *Effect of the Si outer layer*

Finally we have studied the effect of the resolution of the Si outer layer. The TPC does not provide enough accuracy at large angles, and the outer Si layer is essential in that domain. However, its intrinsic resolution is not critical and no significant effect was observed when it was varied from 15 μm to 30 μm.

4. Conclusion

A small forward TPC combined with an outer layer of Si has the capability to significantly improve the resolution in the forward region, provided the entrance window of the TPC can be made thinner than 0.1%, and there is hope to get a position resolution of 200 μm down to angles of $\cos\theta =$ 0.92. We also assume that the beam pipe can be reduced to

0.1% *r.l.*. We believe that the assumptions used in the TPC modeling are realistic, but R&D work has to be performed in several areas: the electrostatic properties of a TPC with a simple and thin inner layer, improvement of the electron yield and drift coefficient, and fast electronics to help resolve the individual arrival times for electrons.

Innovative Beam Pipe Technology

R. Erbacher and W. Vernon

1. Introduction

Extremely good vertex resolution is the hallmark of the B Factory. Although for some purposes there may not be a need to set complicated standards on vertex reconstruction abilities, failing to do as good a job as possible will perhaps rule out future physics that otherwise could have been done on the same machine. There is always room to improve a measurement, and a need to do more physics. Each new machine should be of optimal usefulness, and the path to future improvements should be explored in a reasonably thorough fashion. Even if the ultimate vertex resolution is not incorporated in the initial design, there are many reasons to investigate new and future technology to assist in making some early design choices so that options are kept open for later upgrades.

In this workshop, it has been shown that CP violation ought to be observable in the very clean ψK^0_S channel without spectacular vertex resolution. Yet there may be other sources of such violations from events which require full reconstruction. If vertex resolution is improved significantly, it would mean a large increase in the number of fully reconstructed events. It could also lead to more interesting physics – more kinds of reconstructed events – in the future.

Well-resolved vertices allow many more events to be fully reconstructed

Ideally, the total thickness of the beampipe in the beam interaction region should allow for 10 μm (or less) spatial resolution in vertex reconstruction.[1] With the need for both structural integrity and cooling capability, is it possible to achieve such a thin pipe ($< 10^{-3}$ radiation lengths)? By choosing both the material and the structure design carefully, we believe that the effective thickness can be optimized to achieve a spatial resolution better than is commonly thought possible. We have tested one sample material/structure combination and are in the process of testing another, hoping to find a way to meet all the criteria of this complicated beam pipe structure.

2. Previous Study

In the initial investigation on the feasibility of an asymmetric *B* Factory, the Jet Propulsion Laboratory (JPL) in Pasadena did a study[2] on cooling possibilities for the beam interaction region, assuming a need for the expulsion of about 2 kilowatts of heat from a 50 cm pipe of radius 1 cm. The sources of the generated heat were resistive losses due to induced currents on the pipe's inner surface (ohmic heating), which come from the absorption of higher order RF modes in the pipe, and image currents from the beam. They concluded that the optimal structure for cooling was two concentric pipes with an annular region for water flow cooling. The pipes were designed to be 0.5 mm of beryllium each. The water region contributed about 0.5 mm Be equivalent thickness, making about 1.5mm Be total equivalent thickness. Since we desire about 1/8 of this, we decided to study gas flow cooling more closely, using different materials than JPL had used. We were also interested in cooling the heat absorbing gas by having it expand into the region where the electronics surrounds the beam pipe. Expansion cooling ought to allow us to keep the electronics temperature as low as possible while reducing the thermal gradients seen by the detector support, perhaps to only one or two $^\circ C$ over the length of the pipe.

Clever placement of expansion cooling holes can control thermal gradients

3. New Approaches

The first material we tested was a cylindrically-shaped extruded ceramic called cordierite, made of $2MgO\text{-}2Al_2O_2\text{-}5SiO_2$. It has an approximate Z equivalent of neon, and a radiation length of about 28.9 g/cm^2 (or 294 cm extruded, 70.5 cm in bulk). The extrusions form a grid of squares on the cross section and long, rectangular tunnels axially, of wall thickness 0.16mm. Since the structure walls are thin and assumed to be porous, our first test was to see if the diffusion rate of gas through the material would allow for enough gas flow to cool the structure. The idea was to send gas through both ends of the structure in the inner tunnels and allow the gas to diffuse toward the outside of the cylinder, carrying the heat with it. As shown in Figure 1, the cordierite would make

Low density ceramics have very long radiation lengths up the annular region of the pipe referred to in the JPL study, and the ceramic would be covered by metallic skins, capable of holding a vacuum for the beam interaction region. We used nitrogen gas for our test, and we assumed a 50 cm pipe of radius 2 cm, carrying a 1 kW heat load. To properly cool the structure, a flow rate of 4000 cc/s, or 2000 cc/s on each end, would be needed. Unfortunately, our measurements showed a mere 100 cc/s diffusion rate, not nearly enough for our cooling purposes.

An alternative to pure diffusion exists, however. If we were to drill tiny holes in the walls of the material to provide for a higher gas flow rate, not only would we be able to obtain the needed gas flow for 1 kW of heat, but we would also get a temperature gradient from the inner to the outer region of the pipe due to Joule-Thomson gas expansion cooling. The pressure gradient from the inner pressurized region to the vacuum in the outer region would allow the gas to cool as it expands. Thus, the gas that travels to the region of the detector electronics will be at a lower temperature than it would have been with pure diffusion cooling. In this case, the use of helium can be ruled out, because for practical temperatures, helium gas warms, rather than cools upon expansion.

To find the actual temperature change achieved by the expansion of nitrogen in going from 300 psi to near vacuum, we interpolated a curve of the second virial expansion coefficient[3] to find the value for nitrogen at room temperature. The change in temperature with respect to pressure is about 0.456 deg/atm for nitrogen. Hence, for a pressure drop of 20 atm (300 psi), we would get a temperature drop of about 9 $^{\circ}C$. This is a definite plus, as it will help minimize differential expansion and it will help keep the delicate electronics of the detector cooler.

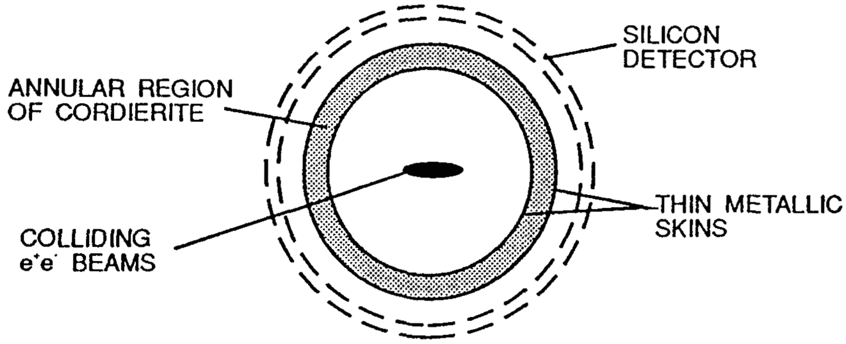


Figure 1. *Cross-section of the pipe, showing relative positions of the pieces, but not to scale. The detector is schematic only.*

There are still tests to be done before considering this material for beam pipe use. We must determine if the needed gas flow can be achieved by punching holes in the tunnel walls, and if so, if the walls of the cordierite can withstand pressure drops of about 6.7 atm. The material must also be tested to see if it can withstand high radiation loads, and if it is structurally sound. In addition, the content of the metallic inner skins must be determined so as to minimize ohmic heating and thermal expansion while maximizing electric field and x-ray attenuation.

3.1. Work In Progress

The next material, which is currently being studied, is a form of reticulated vitreous carbon. The material is available in several different porosities, and is fairly machinable. Boron carbide and silicon carbide can be made in the same form, but for study purposes, RVC is cheaper and more easily obtained. We have made an aluminum holder for our sample of RVC that allows gas flow through a channel in it. We plan to cover the sample with a thin aluminum skin so as to emulate a possible pipe assembly. We are in the process of testing adhesive for the metal skin, and hope to find one that will hold under turbulent temperature, pressure, and radiation conditions. We will measure several properties, including gas flow rate through the material and through a channel in the material, diffusion rate, buckling under pressure, and amount of pressure it can withstand.

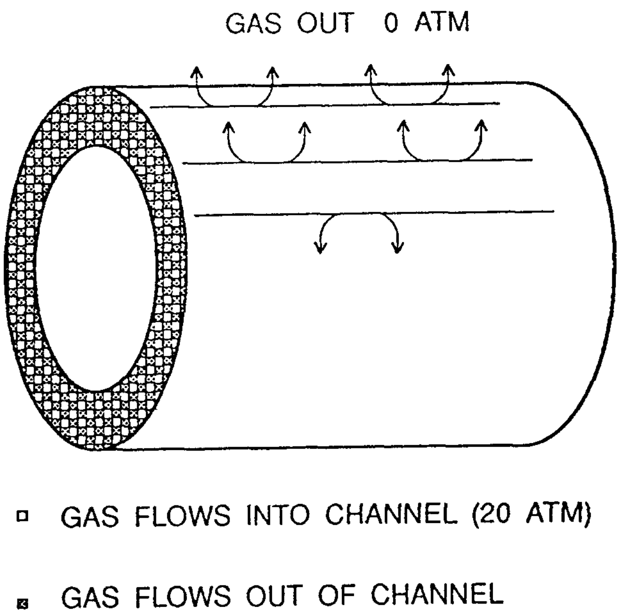


Figure 2. *End view of the beam pipe with indicators of how gas flow would be arranged*

4. Conclusion

It is our hope that either one of these materials, or something along these same lines, will give us our desired combination of cooling, thinness, and strength. It may be, however, that the thin wall pipe structure, which must be matched by a very thin detector, will then have to run at low temperatures to allow the use of CCD structures. This would also improve the electric field shielding problem by reducing the RF skin depth on the inside of the pipe. In this case, cryogenic problems will come into play, and we will need to test the materials for other conditions, such as the behavior of glue joints at liquid nitrogen temperatures.

Tektronix is selling an evaluation CCD, 10 micron thin, for $750

References

1. R. Erbacher and W. Vernon, The Physics Program of a High-Luminosity Asymmetric *B* Factory at SLAC (Appendix), SLAC-353.
2. H. Schember and P. Bhandari, Jet Propulsion Laboratory, California Institute of Technology, JPL D-7151, Feb. 1990.
3. The Theory of Thermodynamics – Fig. 14.4, J.R. Waldram, Cambridge University Press 1985.

Report of the Charged Particle Tracking Group

A. Boyarski, P. Burchat, M. King and A. Weinstein

1. Introduction

The central tracking device is located in the radial region outside the vertex detector and inside the particle identification system and electromagnetic calorimeter. The role of this device is to measure the position of charged particles emanating from the interaction point. The tracking chamber is placed in a magnetic field in order to extract the momentum of the particle from the position measurements by measuring the curvature of the charged-particle trajectory.

The role of the central tracking device is to measure charged particle momenta.

It is expected that the rate of beam-related background signals in the tracking device could be high. The tracking chamber should, therefore, have ample segmentation and sufficient position measurements per track to allow robust pattern recognition of the real tracks among the background signals. A large number of position measurements per track also increases the precision of the measured track parameters, improves the track-finding efficiency at small forward or backward angles where charged particles traverse only the inner part of the central tracker, and provides some particle identification from a measurement of energy loss per unit distance (dE/dx) through ionization. On the other hand, the number of samples per track cannot be too large since the radial size of the device increases with the number of samples thus increasing the size and cost of the components outside the central tracker. Finally, this device should have as little mass as possible in order to reduce the uncertainty on the measured track parameters of low momentum particles due to multiple-Coulomb-scattering, and also to prevent degradation in the measurement of photon energy in the electromagnetic calorimeter.

2. Tracking Resolution and Physics Issues

The central tracking device provides a measurement of charged track parameters by measuring the position of the particle at approximately 40 points along the trajectory. The precision of the measurement of the track parameters is determined by two factors: the intrinsic resolution with which the track position is measured at each of the 40 points, and the amount of multiple-Coulomb-scattering which the particle undergoes while traversing the tracking device. See the vertex detector chapter in this report for a thorough discussion of the track parameter resolution for various vertex detector/central detector configurations.

To evaluate the importance of each of these effects on actual physics measurements, we considered three scenarios for tracking resolution and multiple scattering and compared the results of four representative physics analyses. We assumed that the track parameters are measured by the central tracker plus three layers of 50 micron square pixels at radii of 2.5, 4.5 and 6.5 cm. This inner silicon detector provides most of the precision for all track parameters except the track curvature.

Three detector scenarios having a nominal resolution, poorer resolution, or more material were used in physics studies.

The three scenarios which we considered for the central tracking resolution are the following:

- 150 μm intrinsic position resolution for each measurement and a gas with a radiation length of 600 m, a typical radiation length for a helium-based gas.
- 300 μm intrinsic position resolution for each measurement and a gas with a radiation length of 600 m.
- 150 μm intrinsic position resolution for each measurement and a gas with a radiation length of 100 m. These parameters are typical for an argon-based gas.

The scenario in which the position resolution is 150 μm and the radiation length of the gas is 600 m is considered to be the nominal scenario. All of the physics analyses discussed in this report assume this scenario. In each case, we assumed that the wires are made of a low-Z material such as aluminum or magnesium.

Four physics analyses were chosen which provide a wide range of typical charged-particle momentum spectra in order to study the effects of intrinsic position resolution, which dominates in high-momentum tracks, and the multiple-scattering contribution, which dominates in low-momentum tracks.

Four representative physics examples were studied with each of the detector scenerios.

1. At the high-momentum end, we consider the decay $B \rightarrow \pi^+\pi^-$, the primary mode for measuring the angle α in the unitarity triangle. The momentum spectrum of the pions as measured in the laboratory is flat between about 1.5 and 4.5 GeV. The figure of merit for this mode is the $\pi^+\pi^-$ mass resolution which determines the signal-to-background ratio.
2. At the low-momentum end, we consider the decays $\tau \rightarrow 5\pi^{\pm}\nu_\tau$ and $\tau \rightarrow KK\pi\nu_\tau$. These modes are of interest for measuring the ν_τ mass. The figure of merit for this analysis is the lowest upper limit one could place on the ν_τ mass, assuming the ν_τ mass is zero.
3. We also compare the reconstruction efficiency and the resolution for measuring the distance between B decay vertices for the benchmark mode $B \rightarrow J/\psi K^0_S$.
4. Finally, we compare the signal-to-noise ratio, for constant reconstruction efficiency, for the mode $B \rightarrow D^+D^-$ when both D's decay to $K\pi\pi$. Like the mode $B \rightarrow J/\psi K^0_S$, this mode is sensitive to the angle β in the unitarity triangle.

The results of the comparison are shown in Table 1. In each physics analysis, the degradation in the mass resolution or signal-to-background ratio, compared to the nominal case, is worse when the radiation length of the gas is decreased (from 600 m to 100 m) than when the intrinsic position-measurement resolution is degraded (from 150 μm to 300 μm). This is true even for the decay $B \rightarrow \pi^+\pi^-$ in which the pion momentum is relatively high. Note that the $B \rightarrow J/\psi K^0_S$ analysis is so robust that changing the tracking resolution has little influence on the reconstruction efficiency. The resolution of the separation of B decays is dominated by the silicon detector and hence is not strongly influenced by the resolution of the central tracking device. For the analysis of $B \rightarrow D^+D^-$, with both D's decaying to $K\pi\pi$, the selection criteria (cuts on reconstructed mass) were modified under each scenario to keep the overall reconstruction efficiency the same. The resulting signal-to-background ratio is degraded by a factor of *six* when the radiation length of the material is decreased by a factor of six.

Table 1. *Influence of intrinsic wire resolution and multiple scattering in the gas on various physics analyses.*

Intrinsic Wire Resolution:	150 μm	300 μm	150 μm
Radiation Length:	600 m (He)	600 m (He)	100 m (Ar)
ν_τ mass limit			
$\tau \to 5\pi^{\pm}\nu_\tau$:	3.8 MeV	4.1 MeV	5.3 MeV
$\tau \to KK\pi\nu_\tau$:	4.9 MeV	5.9 MeV	7.0 MeV
$B \to \pi^+\pi^-$			
Invariant mass resolution:	23 MeV	30 MeV	35 MeV
$B \to J/\psi K_S$			
Reconstruction efficiency:	44%	42%	40%
Δz resolution:	59 μm	62 μm	64 μm
$B \to D^+D^-$, $D \to K\pi\pi$			
D mass resolution:	4.6 MeV	4.9 MeV	8.1 MeV
B mass resolution:	6.6 MeV	7.5 MeV	12 MeV
Signal-to-noise ratio:	40	30	7

Physics analyses are more sensitive to multiple scattering in the gas than to intrinsic wire resolution.

In conclusion, analyses in which reconstructed-mass resolution is important (in particular, multibody final states) benefit greatly from the lower multiple scattering of a helium-based gas. Even if the intrinsic position resolution of a helium-based gas is a factor of two worse than an argon-based gas, all of the representative physics analyses considered for this study benefit from the lower multilple scattering contribution from the gas. Many analyses are not strongly dependent on the resolution of the central tracking chamber because the precision of the measurement of most of the track parameters (with the exception of curvature) is dominated by the three layers of silicon close to the beam pipe.

3. Options for Drift Chamber Gases

Traditionally, the noble gas used in tracking devices has been argon because of its abundance, relatively low atomic number (Z=18), low ionization potential and high gain. In the "factory" era of high-luminosity, low-energy facilities, the small statistical errors drive us to decrease systematic errors to unprecedentedly low values. We must therfore consider lower-Z noble gases to reduce the multiple scattering contribution to the position resolution of tracking devices. As discussed in the previous section, the benefits of using a gas based on a low-Z atom such as helium are substantial. Additives are necessary to "quench" the gas (absorb ultraviolet and x-ray photons) and decrease the characteristic electron energy so that the diffusion component of the resolution will not be large. Isobutane is known to be a good quencher and CO_2 is a "cool" gas (low characteristic electron energy).

We have used a computer program[1] to investigate the electron transport properties of mixtures of helium, isobutane and CO_2. In the program, the Boltzmann transport equations are solved and elastic and inelastic electron scattering cross sections for components of the gas mixture are used to predict the drift velocity, Lorentz angle and diffusion coefficient for a particular gas mixture. The predictions of this program are in good agreement with the measured properties of many gas mixtures. Table 2 lists the properties of an argon-based gas ("HRS gas") and a helium-based gas (78% He, 15% CO_2, 7% isobutane) predicted by this program for an electric field of 1 kV/cm and a magnetic field of 1 Tesla. The Lorentz angle for the helium-based gas is half that of the argon-based gas; this is particularly attractive for a small-cell drift chamber.

Theoretical studies indicate that a 78% helium-15% CO_2-7% isobutane mixture could be a viable choice for a drift chamber gas.

The small Lorentz angle is primarily due to the large CO_2 content; the electron drift velocity is relatively low and is not saturated. The electron drift velocity is plotted versus electric field in Figure 1. The radiation length is ~725 m, compared to 110 m for pure argon.

Table 2. *Expected properties of a helium-based and an argon-based gas. The electron transport properties (Lorentz angle and drift velocity) are predicted by a computer program [1] for an electric field of 1 kV/cm and a magnetic field is 1 Tesla. The predicted number of ion pairs pruduced per cm for each gas mixture is based on a simple weighting of the expected number for the pure gases.*

Gas Content:	89% argon 10% CO_2 1% Methane	78% helium 15% CO_2 7% Isobutane
Lorentz Angle	33°	14°
Drift Velocity	45 μm/ns	24 μm/ns
Primary ions/cm	29	13
Total ions/cm	93	34

The ionization potential of helium is significantly greater than that of argon. A simple calculation of the expected number of ion pairs per unit length leads to a prediction of 93 total ion pairs per cm for HRS gas and only 34 total ion pairs per cm for the helium-based gas. Consequently, we expect the mean pulse height for the helium-based gas to be 36% of that of HRS gas. The number of primary ion pairs/cm is expected to be 42% of that of HRS gas, so the intrinsic resolution will not be as good as the HRS gas, although it may be possible to improve the resolution by running at a higher voltage.

A detailed study of this helium-based gas is being carried out at SLAC with a prototype drift chamber with jet cell geometry. The initial results indicate that the mean pulse height for the helium-based gas described above is about a factor of 2.5 lower than HRS gas at the same high voltage setting. When we increase the high voltage on the chamber by 9%, the pulse height increases to about 2 times that of the HRS gas, and the intrinsic wire resolution of the helium-based gas is comparable to that of the argon-based gas. We also find that the maximum operable high voltage setting for the helium-based gas is about 7% higher than that of the argon-based gas. Overall, these studies indicate that we can be optimistic about using a helium-based gas.

A pressurized TPC has more material; space-charge effects may be severe; readout time is prohibitively long.

distortions in the trajectories of the electrons drifting to the readout pads. Finally, a TPC with a 1.5 meter drift path filled with a gas with an electron drift velocity of 20 mm/μs has a readout time of about 75 μs. With the pipeline triggering scheme being considered, this would require very long pipeline buffers for all detector components so that the TPC data could be used in the trigger. For these reasons, we did not study a TPC for the B Factory. On the other hand, we did not investigate the above problems in detail, and solutions may exist for a TPC in a B Factory environment.

A drift chamber does not have the above problems. It operates at atmospheric pressure, is less prone to space charge degradation, and has a short readout time (1 - 2 μs). Although a single wire in a drift chamber only provides two dimensional coordinates, there are ways to achieve three dimensional tracking. Such chambers are widely used in storage ring experiments, albeit at much lower beam intensities than will exist in a B Factory. Attention will have to be given to the background levels expected in a drift chamber, as well as aging effects from the collection of ions at the wire surfaces. These concerns are discussed in a later section.

5. General Design Parameters

A wire drift chamber with axial and stereo layers is our choice for a tracking device.

In this study, we assume a magnetic field strength of 1.0 Tesla. The full length of the drift chamber is 3.0 m; the inner radius is 18 cm and the outer radius is 80 cm. The drift chamber consists of concentric layers of cells. In a drift chamber, only two space coordinates are determined per measurement. However, the third (z) coordinate can be determined by one of several methods, namely charge division (measuring pulse heights at each end of a sense wire), time difference (measure pulse timing at each end), or layers of wires which are not parallel to the cylindrical axis of the chamber (stereo layers). The latter does not provide three-dimensional space points, but allows tracks to be reconstructed in three dimensions. We prefer the stereo layer approach, since the other two methods require twice the electronics to read out both ends of each wire. In our design, pairs of stereo layers are located between

A pressurized TPC has more material; space-charge effects may be severe; readout time is prohibitively long.

distortions in the trajectories of the electrons drifting to the readout pads. Finally, a *TPC* with a 1.5 meter drift path filled with a gas with an electron drift velocity of 20 mm/μs has a readout time of about 75 μs. With the pipeline triggering scheme being considered, this would require very long pipeline buffers for all detector components so that the *TPC* data could be used in the trigger. For these reasons, we did not study a *TPC* for the *B* Factory. On the other hand, we did not investigate the above problems in detail, and solutions may exist for a *TPC* in a *B* Factory environment.

A drift chamber does not have the above problems. It operates at atmospheric pressure, is less prone to space charge degradation, and has a short readout time (1 - 2 μs). Although a single wire in a drift chamber only provides two dimensional coordinates, there are ways to achieve three dimensional tracking. Such chambers are widely used in storage ring experiments, albeit at much lower beam intensities than will exist in a *B* Factory. Attention will have to be given to the background levels expected in a drift chamber, as well as aging effects from the collection of ions at the wire surfaces. These concerns are discussed in a later section.

5. General Design Parameters

A wire drift chamber with axial and stereo layers is our choice for a tracking device.

In this study, we assume a magnetic field strength of 1.0 Tesla. The full length of the drift chamber is 3.0 m; the inner radius is 18 cm and the outer radius is 80 cm. The drift chamber consists of concentric layers of cells. In a drift chamber, only two space coordinates are determined per measurement. However, the third (z) coordinate can be determined by one of several methods, namely charge division (measuring pulse heights at each end of a sense wire), time difference (measure pulse timing at each end), or layers of wires which are not parallel to the cylindrical axis of the chamber (stereo layers). The latter does not provide three-dimensional space points, but allows tracks to be reconstructed in three dimensions. We prefer the stereo layer approach, since the other two methods require twice the electronics to read out both ends of each wire. In our design, pairs of stereo layers are located between axial layers. The angle of the stereo wires is $\pm 4.0^\circ$ with respect to the axial direction. Wire chambers also provide some particle identification by dE/dx when the pulse heights of the signals are measured.

6. Cell Design

There are several choices for the cell design of a wire chamber: small cells as used by ARGUS and CLEO II; large jet cells as used by OPAL; and segmented jet cells as used in the Mark II, SLD, and CDF detectors. However, the large jet cell could suffer from diffusion effects at the larger drift distances, so we avoid considering the large jet cell. This leaves us with the small cell and the segmented jet cell, both of which appear to be possible candidates.

A small cell or segmented jet cell are both viable, and both are studied in this report.

An electrostatic-simulation program was run for these cell configurations. The program accepts as input the positions, radii, and voltages for a configuration of wires, and calculates the resulting charges and electric fields on the wires. For a 20 μm diameter sense wire, and a design goal of 5×10^4 avalanche gain, the necessary charge on the sense wire is about 13 picocoulombs per meter in an argon-based gas. We have used this same value for the helium-based gas (although initial prototype studies indicate $\approx$ 10% larger value may be required). A criteria in the design of the cells is that the electric field on the surface of the field wires should not exceed 20 kV/cm in order to prevent growth of deposits on the field wires.

Small Cell Drift Chamber In this design, several cylindrical layers of cells are grouped in superlayers, with 3 to 6 layers per superlayer as shown in Figure 2. As the radius increases, the superlayers alternate between an axial layer and a pair of stereo layers at angles of $+4.0^\circ$ and -4.0°. For a cell size of 15 mm along the radial direction, there is room for a total of 10 superlayers (4 axial and 3 pairs of stereo layers). All wires within a superlayer are parallel to each other except the very inner most layer of field wires, which must be shared from the neighbouring superlayer at the smaller radius. Within this edge cell, the relative position of the sense wire and five of the eight field wires remain fixed along the length of the chamber, but the position of the the three field wire can vary along the length of the chamber. This produces a change in the charge density on the sense wire along its length, and therefore a gain change along the wire. However, these layers still provide useful tracking information, and the dE/dx information may be useful if z-dependent corrections are applied.

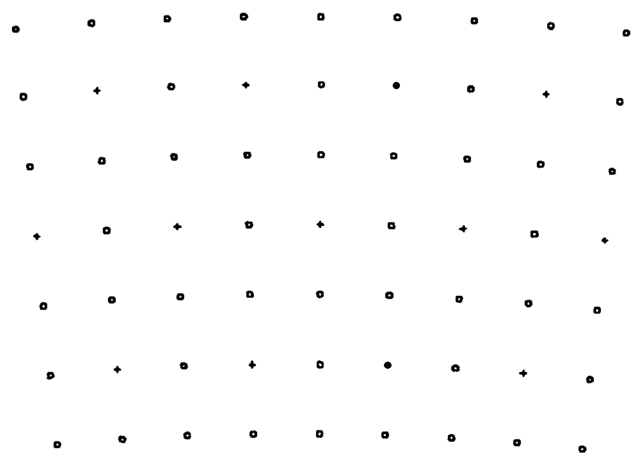


Figure 2. *Small Cell Configuration. Sense wires are marked by +, field wires are circles. Three layers of cells are shown.*

The small cell design consists of 15 mm square cells, 39 samples per track, grouped into 10 super layers

A square cell size of fullwidth 15 mm is chosen. Making it smaller quickly increases the number of wires, while making it much larger leads to undesirable distortions in the electron drift trajectory due to spiraling of the electrons in the magnetic field as they drift to the sense wires.

The radial size of the chamber allows 39 layers of cells with the 15 mm cell size. Figure 3 shows the wire locations at the end plate for a quadrant of the small cell configuration. The axial (stereo) layers are labelled with an A (U,V) and the number of cells per superlayer is given in parentheses. Within a superlayer, each cell layer has the same number of cells, and the sense wires in each cell layer are staggered by a half cell with respect to the neighboring layer to allow resolution of the left-right ambiguity of a track within the superlayer. This geometry forces the cell widths to increase in size from the inner layers to the outer layers within a superlayer, and is most pronounced at the innermost superlayer. The innermost superlayer has only four layers to avoid too large a variation in cell width. Extra radial space has been added at the 3 axial-stereo boundaries to accommodate the stereo waist (reduced radius) for the wires at the mid-length of the chamber.

The field wires at the very inner and outer radii are replaced by cathode pad layers to provide measurements of the z coordinate of tracks at these radii. The cathode pads would be made of printed circuit planes along the inner side of the cylindrical walls of the chamber, with conducting strips running azimuthally around the walls. The best segmentation and width of these strips has not been explored in this study.

The 20 kV/cm limit on the electric field at the surface of the field wires constrains the field wires to be at least 80 microns in diameter, and the voltage on the field wires to be -1830 V with the sense wires at 0 volts.

The properties of these cells were studied with a simulation program which calculates the trajectory of drifting electrons at any point in the cell. In a magnetic field, electrons drift along trajectories which are tilted by the "Lorentz" angle, the angle between the drift direction and the electric field direction. A Lorentz angle of 14° was used for the simulation. Trajectories of drifting electrons starting at points spaced a millimeter apart along two sample particle tracks are shown in Figure 4. Spiraling trajectories can be seen, but the spiraling is not severe due to the relatively small Lorentz angle expected for a helium-based gas. The boundaries of the cells are also seen to be "S" shaped lines rather than straight lines. Tracks passing near the edge of one cell will produce a signal in the neighbouring cell.

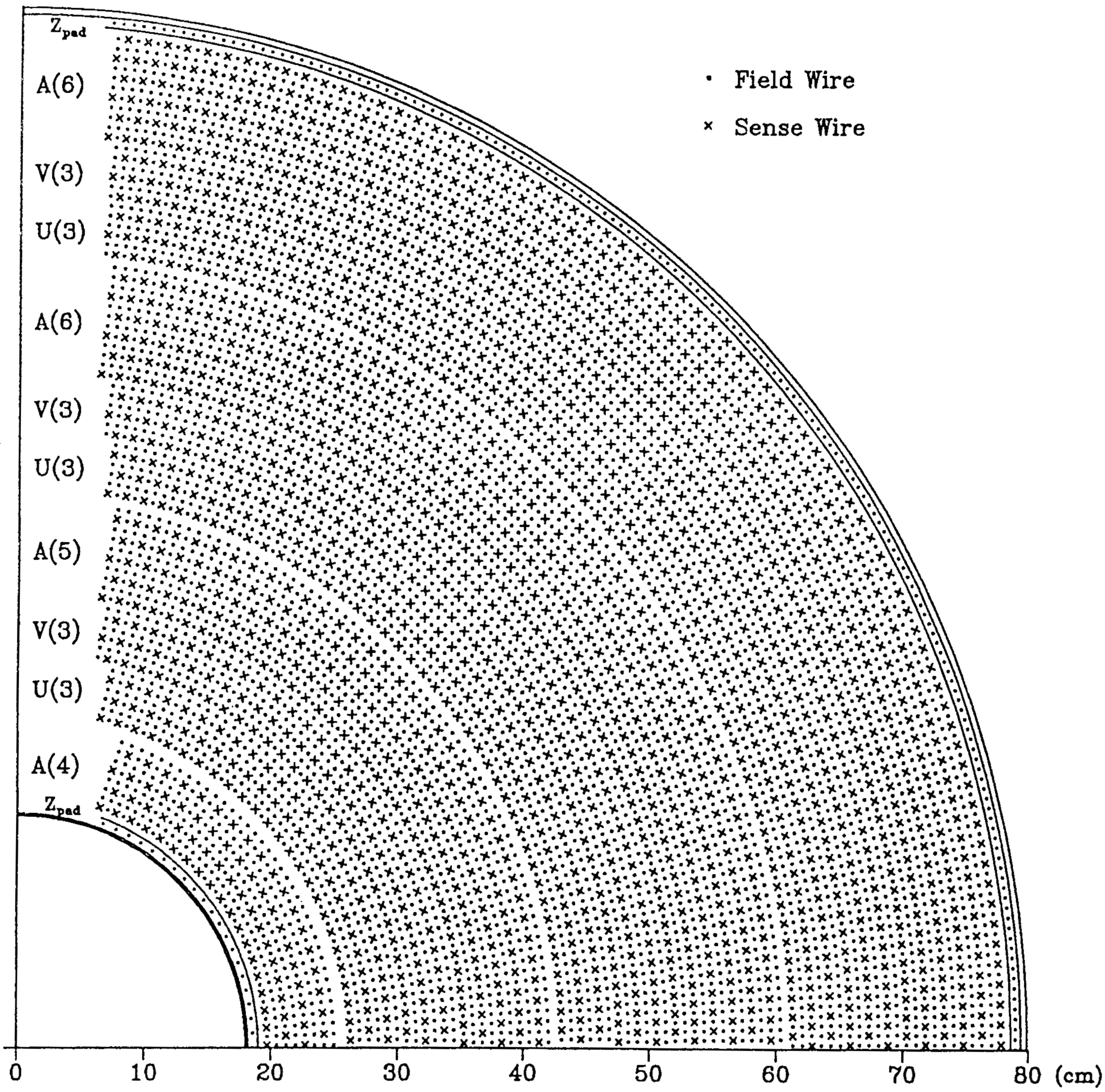


Figure 3. *Quadrant of a small cell chamber, showing the axial (A) and stereo (U,V) super layers and the z pad layers.*

The sample size across the cell has also been studied in the simulation. The sample size is the length of the track segment from which electrons drift to the sense wire in the same cell. A uniform sample size is required for good dE/dx performance. Figure 5 shows the sample size for straight tracks passing along the radial direction as a function of the track distance from the sense wire. The sample size is quite uniform for this configuration.

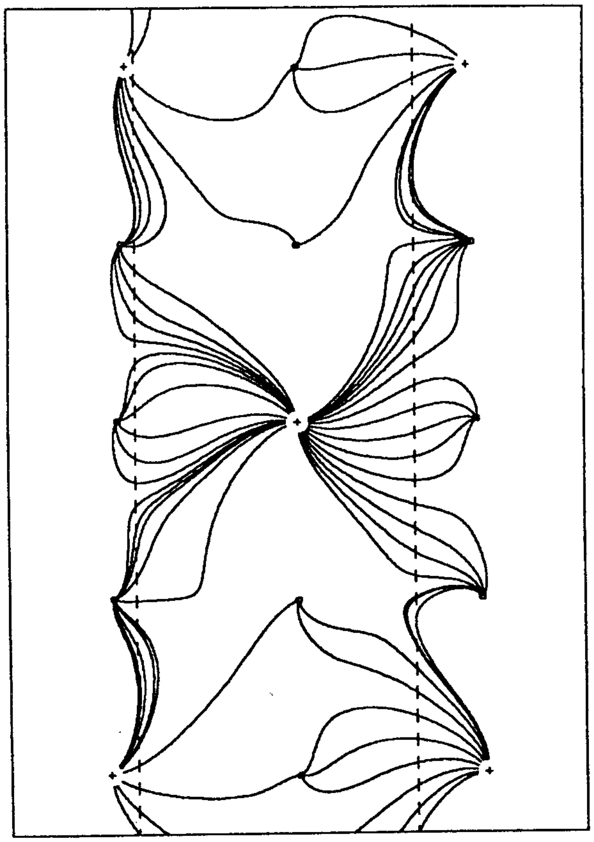


Figure 4. *Electron drift trajectories in a small cell, for a Lorentz angle of 14° in a 1 Tesla magnetic field. Trajectories passing through points every mm along two fiducial straight lines are shown.*

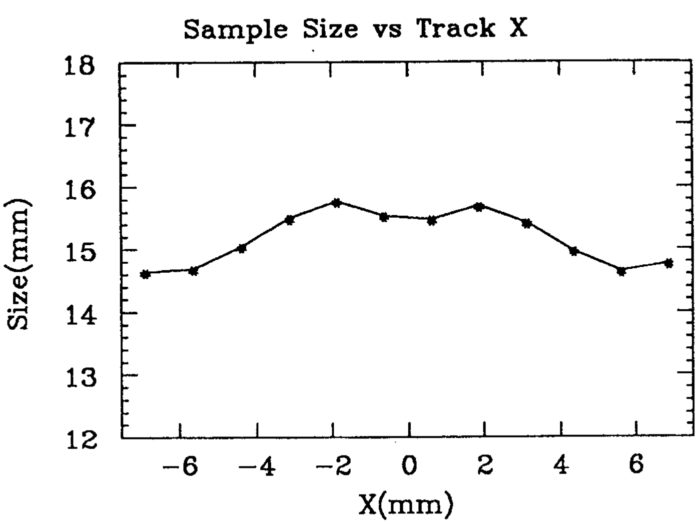


Figure 5. *Sample size, or length of track segment collected by the sense wire, as a function of the track distance from the sense wire for the small cell configuration.*

6.1. Segmented Jet Cell Drift Chamber

The jet cell design discussed here is a modified version of that used in the Mark II chamber. With the smaller radial size of this chamber, there is room for only 5 sense wires per cell, and 8 layers of cells. A smaller cell width (25 mm halfwidth) is chosen to avoid loss of resolution due to diffusion in the helium-based gas. The spacing between sense wires is 9.0 mm. The sense wires are staggered alternately by a few hundered microns to resolve the left-right ambiguity. Potential wires are placed midway between the sense wires to reduce the signal coupling between the neighbouring sense wires. A schematic of one jet cell is shown in Figure 6.

There are 4 axial and 4 stereo superlayers. All wires within a superlayer are parallel to each other. Figure 7 shows one quadrant of the jet-cell chamber. In order to minimize the space between superlayers, while keeping the coupling between superlayers at a small level, a single shielding layer of wires is placed between superlayers. This shield layer reduces the variation in charge density, and therefore gain, along a wire due to the change in the relative position of wires from one superlayer to the next at a stereo-axial boundary. A criteria of 1% maximum change in the charge on the sense wire was used in determining the minimum radial separation between adjacent super layers. This minimum is found to be 15 mm. A 1% change in charge results in $< 20\%$ change in gain.

For a 20-micron diameter sense wire, a design gain of 5×10^4 can be achieved with a cathode potential of -3220 V at the central radius of a cell. The 20 kV/cm limit for the field at the surface of the field wires then determines the diameters of the wires: 130 microns for the field wires (200 microns for each end most field wire in a cell), 110 microns for the shield wires, and 80 microns for the potential wires.

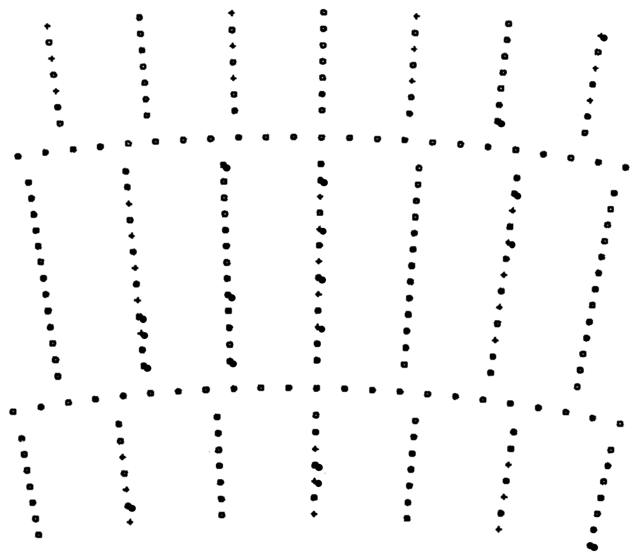


Figure 6. *Jet Cell Configuration. Each cell has 5 sense wires (+) spaced 9 mm apart. The cell half width is 25 mm at its mid radius. Circles mark the field wires, potential wires (between sense wires), and shield wires (between super layers).*

Drift trajectories and sample size (as described in the previous section) were also studied for the jet cell and are shown in Figures 8 and 9, respectively. The drift trajectories illustrate that the electric field in the cell is quite uniform for all regions except near the edgemost wire where the trjectories are bowed at the larger drift lengths.

The sample sizes, shown in Figure 9, vary as much as 20% for the edge wire, and less than 10% for all other cases.

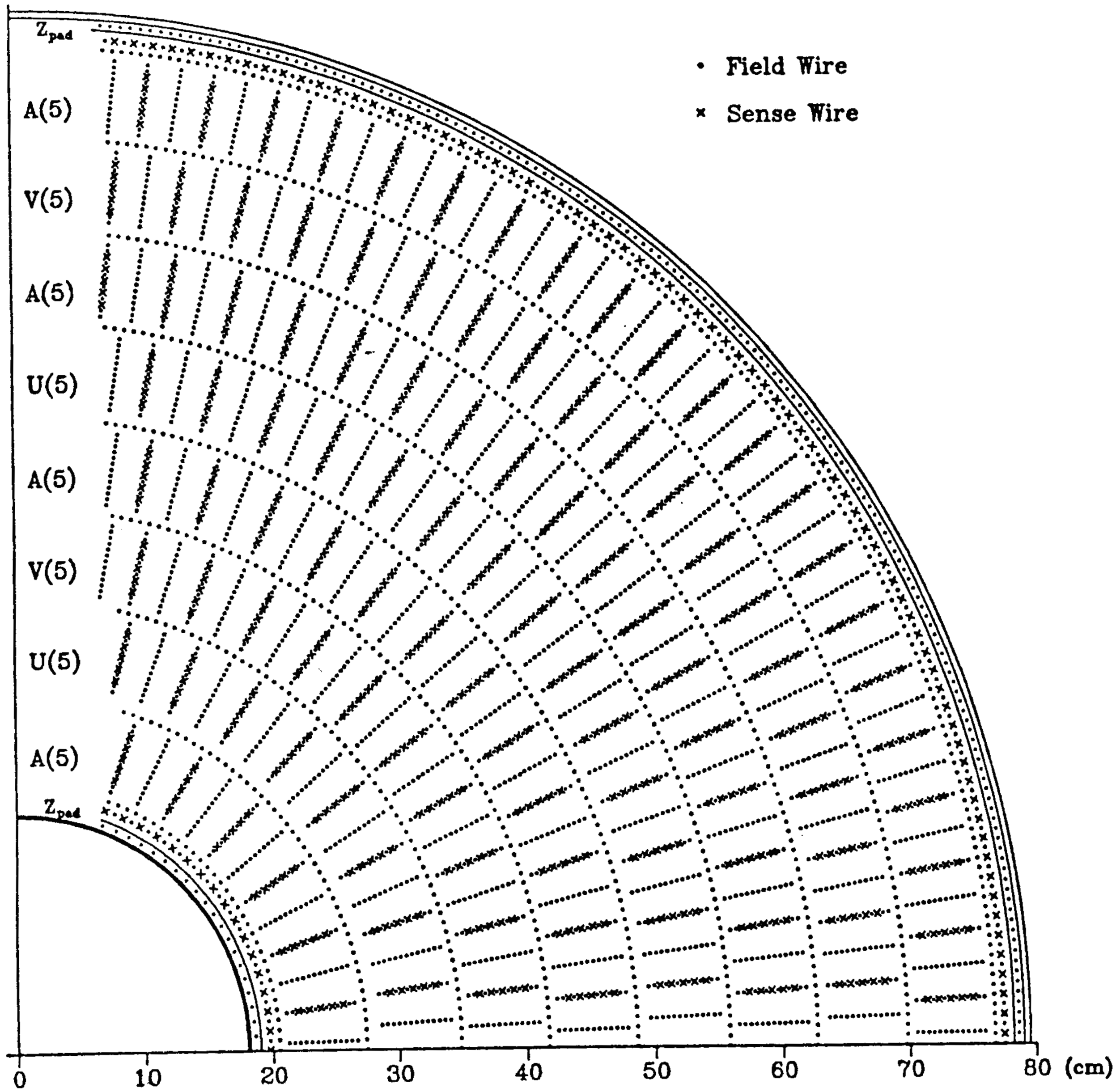


Figure 7. *Quadrant of a jet cell chamber, showing the 4 axial (A) and 4 stereo (U,V) super layers and z pad layers.*

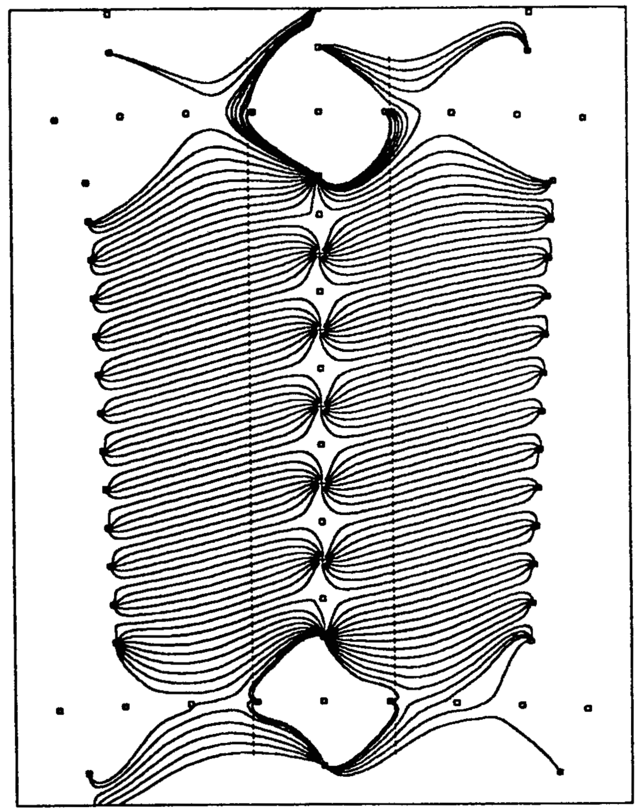


Figure 8. *Electron drift trajectories in a jet cell, for a Lorentz angle of 14° in a 1 Tesla magnetic field. Trajectories passing through points every mm along two fiducial straight lines are shown. Trajectories are quite linear for all but the end most wire.*

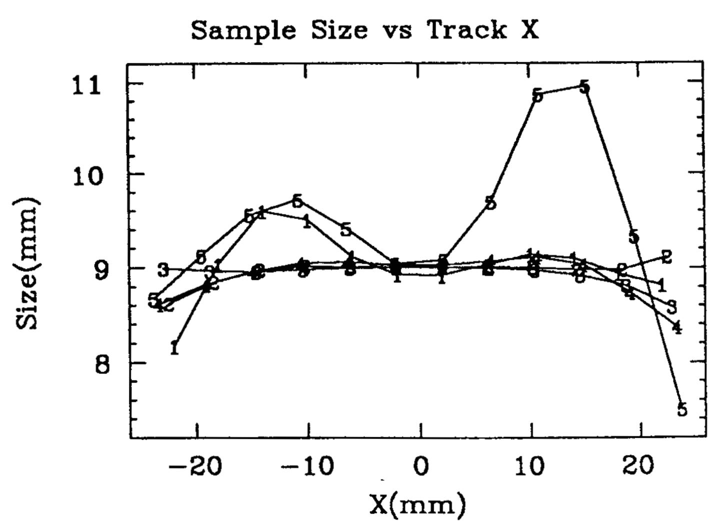


Figure 9. *Sample size across a jet cell for the five wires in a cell. The middle wires (2,3,4) are very uniform while the edge wires (1,5) show variations up to 20%.*

6.2. *Forces on Wires*

Wires in a drift chamber are subject to two kinds of transverse forces, namely the gravitational force acting on the mass in the wire, and the electrostatic forces between wires due to the charges on the wires.

The gravitational force acts perpendicular to the wire in our chamber orientation. For a wire with density ρ, diameter d, length L, tension T, and a gravitational constant g, the deflection δ as a function of distance z along the wire is given by

$$\frac{d^2\delta}{dz^2} = -\frac{\pi d^2 \rho g}{4T}$$

which has the solution

$$\delta(z) = \frac{\pi d^2 \rho g}{8T}\left(z^2 - \frac{L^2}{4}\right)$$

where the boundary conditions $\delta(-L/2) = \delta(L/2) = 0$ have been applied. The maximum deflection, at $z = 0$, is

$$\delta^g_{max} = -\frac{\pi d^2 \rho g L^2}{32T}.$$

Gravitational sag of wires is 130 μm at the highest tension allowable for tungsten.

For a 20 μm tungsten wire with tension 0.5 N (50 gram weight) and length 3.0 m, the maximum deflection is about 130 μm. If all wires have the same gravitational sag, the relative shape of the cells remain unchanged along the length of the chamber, and the tracking software can easily make z-dependent corrections to wire coordinates for the ultimate tracking precision. In order to achieve equal gravitational sag for wires of differing diameters and composition, the tension for wire f should scale with that of tungsten W as

Wire tensions scale with the wire's density and cross sectional area.

$$T_f = T_W\left(\frac{\rho_f}{\rho_W}\right)\left(\frac{d_f}{d_W}\right)^2.$$

However, we will see that electrostatic forces constrain the tension to be greater than a certain value.

The electrostatic force per unit length on wire i is given by

$$\mathbf{F}_i = Q_i \mathbf{E}_i$$

where Q_i is the charge per unit length on wire i and $\mathbf{E}_i$ is the electric field at wire i produced by the charges on all the other wires, not including wire i. For equilibrium, the transverse electrostatic force $F_\perp$ must be balanced by the transverse component of the tensing force:

$$T \frac{d^2 \delta}{dz^2} = -F_\perp$$

where δ is the transverse displacement of the wire along its length z. For a jet cell, the sense wires are staggered with respect to the centerline of the cell by a distance $\pm\delta_0$ at the end plates. The electrostatic force on a sense wire is perpendicular to the sense wire plane and under the assumption of an infinite plane of sense wires can be calculated analytically as [2]

$$F_\perp = K(\delta + \delta_0)$$

where

$$K = \frac{\pi Q^2}{4 \epsilon_0 s^2}$$

and s is the spacing between sense wires, Q is the charge per unit length on each sense wire, and $\epsilon_0 = 8.85 \times 10^{-12}$ F/m.

The solution is

$$\delta(z) = \delta_0 \left(\frac{\cos\left(\sqrt{\frac{K}{T}} z\right)}{\cos\left(\sqrt{\frac{K}{T}} \frac{L}{2}\right)} - 1 \right)$$

where the conditions $\delta(-L/2) = \delta(+L/2) = 0$ have been applied. The maximum deflection at $z = 0$ is

$$\delta^e_{max} = \delta_0 \left(\frac{1}{\cos\left(\sqrt{\frac{K}{T}} \frac{L}{2}\right)} - 1 \right)$$

For a given tension T, as the high voltage on the chamber increases, Q increases and the predicted deflection grows unbounded as the argument of the cosine in the denominator approaches $\pi/2$. Actually, as the deflection increases, the wire stretches and the tension increases, thereby prolonging the stable region to higher voltages. However, at sufficiently high voltage, the tension will exceed the ultimate strength of the wire and the wire will break. Alternately, for a given charge Q, as the tension is decreased from a high value there is a critical tension T_c given by

$$T_c = \frac{KL^2}{\pi^2}$$

Electrostatic forces require tensions above a certain critical value to avoid large and unstable deflections of wires.

at which the wire starts to becomes unstable. For safe operation, a drift chamber should be designed with wire tensions considerably larger than this critical tension T_c.

Critical tension for a 20 μm sense wire is 0.14 Newtons, or 14 g weight, in a jet cell.

For the jet cell, s is 9.0 mm. For $Q = 13 \times 10^{-9}$ C/m, we get $K = 0.15$ N/m/m and $T_c = 0.14$ N. This is well below the familiar tension of 0.5 N used on 20 μm tungsten sense wires. (However, see the end of this section for reasons to use lower tensions). The maximum deflection is $\delta^e_{max} = 0.4\delta_0$.

The above calculation was verified by the simulation program. The simulation program calculates the charges on all the wires, and so can compute the electrostatic force on any wire. The program predicts 0.14 N/m/m for K, in good agreement with the above calculation.

For the small cell configuration, it is more difficult to derive an analytic expression for the electrostatic forces, so the values given by the simulation program are used instead. In the small cell, the sense wire is approximately in the center of the cell formed by its 8 nearest field wires, so the force on the sense wire is approximately zero. However, if the wire is displaced by a small amount δ in any direction, there will be a force pointing away from the center of the cell with a magnitude proportional to δ for small values of δ.

The constant K in the jet cell formalism can again be used for the small cell case to relate the force and the electric field by $F_\perp = K\delta$. Only the value of K is needed. The simulation program determines K by calculating the change in force on the wire as it is moved by a distance δ away from the center. The critical tension can then be calculated once K is known.

The results for the small cell are: $K = 0.029$ N/m/m, and $T_c = 0.027$ N. The critical tension for the small cell is about 5 times smaller than that for the jet cell.

Critical tension in a small cell is only 0.027 Newtons, or 3 g weight, which is 5 times less than the jet cell value.

It is desirable to use a low-Z material in the field wires to reduce the multiple Coulomb scattering. Also, low density material requires less tensile force for the same gravitational sag, thereby reducing the force on the end plate of the chamber. Of all the normal wire materials, magnesium has the lowest Z (12) and the lowest density ($\rho = 1.74$ g/cm^3) making it the ideal material. Aluminum is next best, with $Z = 13$ and $\rho = 2.70$ g/cm^3, and has been used in drift chambers.

In order to match the gravitational sag of a tungsten sense wire tensed with a 50 g weight, the stress in aluminum wire would have to be 31,800 psi, while for magnesium it is 20,400 psi. The yield strength of hard drawn aluminum wire alloy 56S is rated at 48,000 psi, which meets the required stress. For magnesium however, the yield strength is only about 22,000 psi, which is too close to the required value. However, for the small cell case, the critical tension is so low that the tensions on all wires could safely be reduced by $\approx 50\%$.⋆ Also, since magnesium has only 20% the density of copper (a commonly used material for field wires) the force on the end plates from the field wires would be a factor of 5 less with magnesium than with copper. Taking advantage of both of the above reductions in tension, the total force on the endplate for a magnesium-wired, small-cell chamber would then be about 10% that of a copper-wired chamber. This would allow for much thinner end plates and inner and outer walls of the drift chamber.

The small cell configuration allows reduced wire tensions, thereby reducing the force on the end plates. Using light-weight magnesium wire provides a further reduction in the force on the end plates.

The critical tension for sense wires in a jet-cell design puts too large of a corresponding stress on magnesium wires, but not too large for aluminum wires. The force on the end plate with aluminum wires would be 30% that for copper wires.

⋆ This will increase the gravitational sag to 260 μm which is not a problem.

6.3. *Comparison of the two configurations*

A comparison of the two chamber designs is given in Table 3. Both designs have nearly the same number of samples per track although the sample size is larger in the small cell design providing slightly better dE/dx performance. The dE/dx resolution listed in Table 3 was calculated for each cell configuration with a program developed by J. Va'Vra.[3] The two designs have almost identical amounts of wire material (the small cell has more field wires, but the wire diameters can be smaller). With more superlayers in the small cell case, the minimum angle for tracks passing through at least 4 superlayers is smaller (15.5°) than in the jet cell design (17.7°).

Small cell and jet cell configurations have about the same amount of material. The jet cell design has fewer sense wires, but the small cell design will have less wire aging, and can be supported by thinner chamber walls.

The jet cell design has about half the number of wires of the small cell design, which means less electronics for the jet cell design. However, the shorter drift time in the small cell (0.3 μs, compared to 1.2 μs for the jet cell) may allow simpler electronics for the small cells. The small cell has a much lower critical tension requirement, so it allows for lower tensions on all wires, thereby reducing the thickness of the end plate and the walls of the chamber. Finally, the ionic volume collected by a sense wire in the small cell is half that of the jet cell, so the small-cell sense wires should have twice the lifetime for the same amount of beam-induced backgrounds in the chamber.

7. Beam-Related Backgrounds and Wire-Aging Issues

It is known that wire chambers can lose performance after the accumulation of about 0.1 to 1.0 Coulumbs/cm on a wire from ion collection. This process is referred to as aging. Beam-related backgrounds can contribute significantly to the aging process, so we have made estimates of the maximum tolerable rate of backgrounds in a B Factory wire chamber. The estimates conservatively assume that the maximum charge each wire can tolerate is 0.1 Coulumbs/cm. With this assumption, and based on a 3 m wire length and a five-year chamber lifetime (5×10^7 seconds), the maximum average chamber current during running is estimated to be 0.6 μA.

Operating over a 5 year span, a wire current of 0.6 μA from ion collection would deposit 0.1 C/cm on a sense wire.

Comparison of Cell Designs		
Configuration	Small Cell	Jet Cell
Number of samples	39	40
Sample size (mm)	15.0	9.0
Half cell width (mm)	7.5	25.0
$\lvert\cos\theta\rvert$ coverage	< 0.965	< 0.951
Min forward angle	15.5°	17.7°
Number of sense wires	6999	2885
Number of field wires	24997	15027
Total Wires	31996	17912
Total X_0 from,		
Cu field wires	.0027	.0032
(Mg field wires)	(.00026)	(.00031)
Tungsten sense wires	.00023	.00008
He-based gas	.00083	.00083
Total radiation lengths	.0037(.0013)	.0041(.0012)
Critical tension (N)	0.027	0.14
dE/dx		
Resolution (%)	6.2	6.9
Plateau/minimum	1.46	1.43
Separation (σ)		
e/π $p < 0.6$	6.6	6.2
$p = 1.0$	6.0	5.5
$p = 2.0$	4.3	4.0
$p = 4.0$	3.0	2.8
K/π $p < 0.6$	< 4.0	< 3.8
$p = 1.0$	0.0	0.0
$p = 2.0$	2.2	2.1
$p = 4.0$	2.3	2.1

Table 3. *Parameters for the small-cell and jet-cell drift chamber designs. X_0 refers to the number of radiation lengths seen by a particle traversing the full radial length of the chamber, not including the walls, at 90°. The dE/dx values are for tracks at 90°.*

To relate this number to existing chambers, we compare it to the observed current drawn by the Mark II drift chamber at *PEP*. The wires in the innermost layer (27 cm radius) of the Mark II chamber drew about 4 nA each. The beam current at *PEP* was ≈6 mA. If we assume that a *B* Factory detector has backgrounds similar to the Mark II at *PEP*, the chamber current would scale with the beam current. Thus, for a 1.5 A beam current, the chamber current would be 0.8 μA/wire, which slightly exceeds the 0.6 μA limit, but would be acceptable. Of course, some of the sources of backgrounds at a *B* Factory were not present at *PEP*.

Extrapolating from MarkII/PEP conditions to B Factory conditions leads to a prediction for the wire current of 0.8 μA, which is slightly above the lifetime criteria, but acceptable. However, background conditions are not the same.

We now discuss the limits on background photons and charged particles imposed by this limit of 0.6 μA per wire. We compare these limits to the background rates predicted by simulations discussed in the chapter entitled "Detector Backgrounds from Scattered Beam Particles in the SLAC *B* Factory Design" in this Proceedings. We assume a wire gain of 5×10^4.

For synchrotron photons with a typical energy of 30 keV (as seen in the Mark II drift chamber at SLC), the maximum tolerable photon rate per wire due to aging would be 0.1 MHz interacting photons, or 0.1 interacting photons/μs/wire. This is the same limitation imposed by pattern recognition (¡10% occupancy). As of this writing, we do not have the simulation information necessary to compare this limit with the expected rate.

Photons are also produced from beam-gas interactions and electromagnetic showers in detector elements near the interaction point. As discussed in the detector-background chapter, simulations predict a rate of 1770 photons/μs entering the drift chamber gas volume. A helium-based drift chamber gas is assumed in the simulation. Then the number of electrons due to showers in detector elements and photons interacting in the gas is about 10 electron tracks/μs just inside the inner wall of the drift chamber and 0.4 electrons tracks/μs about half way through the chamber. The average energy of these particles is ≈36 MeV and the average transverse momentum is ≈13 MeV/c. Because of their low momentum, these electrons will spiral tightly in the 1 Tesla magnetic field, losing all of their energy through ionization in the neighborhood of a few sense wires. The above predictions for the number of electron tracks already take into account the spiralling trajectory of the electrons (*i.e.* , a single electron can produce more than one track in the neighborhood of a wire). Assuming that each charged track results in the creation of 100 ion pairs per wire (*i.e.* , approximately 100 ion pairs per cm of path length), the limit (due to aging) on the charged track rate is about $0.8/\mu s$/wire. Therefore, for 92 sense wires in the first layer, the limit, due to aging, on the number of electron tracks just inside the drift chamber wall is 70 electrons/μs, a factor of 7 higher than the expected rate of 10 electrons/μs. This is not a large safety margin especially since the distribution of background photons is not flat in azimuthal angle. A rate of 0.1 charged particles/μs/wire is also the upper limit for pattern recognition. In general, these low-energy background electrons pose a problem warranting further consideration.

The number of beam-related background electrons in the gas predicted by Monte Carlo studies is one-seventh of the limit set by aging.

The simulations discussed in the detector-background chapter predict much lower levels for high momentum electrons ($\gtrsim$ 100 MeV) than the maximum rate of $0.8/\mu s$/wire due to aging. For efficient pattern recognition, the background rate should be less than $0.1/\mu s$ per wire (10% occupancy). To maintain a reasonable livetime (*i.e.* , a sufficiently low trigger rate), the rate of high momentum charged tracks from beam-related backgrounds in the *entire* drift chamber must be significantly less than 1 MHz. The exact limit depends on the details of the trigger design.

A number of factors can contribute to slowing the nominal aging process,[45] thereby prolonging drift chamber lifetime. Many of them are listed below. Although the understanding of these variables is generally qualitative, whenever possible and appropriate, it is suggested that they be implemented in the construction and operation of the B Factory wire chamber.

1. In general, high purity of system and gas is desirable.

2. A high gas flow rate is useful to remove impurities in the system. In particular, oxygen and hydrocarbon radicals are volatile and can be removed by sufficient gas flow.

3. The gas tubing should be made of electro-polished stainless steel or of a nylon tubing such as RISLAN. Gas tubing made of molecular chains containing halogens, such as PVC, teflon, tygon and neoprene rubber, should be avoided.

4. In general, impurities to be avoided are halogens (e.g, Freon-11), silicones (e.g., greases and oils, such as might be used in bubblers, raw G-10, etc.) and soft glues (e.g., one part RTV, urethane, etc.). To avoid silicon contamination, the use of good filters to prevent dust from entering the active volume of the wire chamber is recommended. (Silicon species are heavier than oxygen and hydrocarbon radicals and hence not easily removable by gas flow. Consequently, silicon contamination may be a more serious problem than contamination from oxygen and hydrocarbon radicals.)

5. Good impurities include a small amount of oxygen or oxygen-containing molecules such as H_2O, alcohols, ether and methylal. Oxygen may react with hydrocarbon radicals to form more stable end products that can be removed with sufficient gas flow. Oxygen-containing molecules generally have large dipole moments with large scattering cross-sections for electron/molecule scattering. This implies that the mean value of electron energy will be decreased in an avalanche. The molecules also have large cross-sections for absorption of ultraviolet photons. Also note that H_2O may increase the conductivity of deposits which are otherwise poorly conductive, thereby prolonging chamber lifetime. In this context, Va'Vra suggests that the use of nylon tubing may naturally add the appropriate amounts of H_2O.

6. Chamber wires should be gold-plated, especially if the wire is made of a complex alloy. In particular, anode wires containing aluminum or nickel are more prone to react to some oxygen-containing molecules in ways that cause gain loss. As a result, alloys like Nicotin and Stableohm should be used with caution.

7. The anode wire, where the polymer building process occurs, should be as large in diameter as feasible: assuming uniform anode deposits, the smaller the anode wire diameter, the larger the gain drop.

8. If dimethyl ether (DME) is utilized, it should be in

a purified form to reduce the content of Freon-11. Further, its solvent capabilities should be considered in wire chamber construction. Most plastics expand and swell in its presence.

9. The wire chamber high voltage should be turned down during beam injection to avoid exposing the chamber to high backgrounds unnecessarily. In a B Factory, injection as often as every 10 minutes is a possibility. Therefore, the wire chamber should be designed and built such that the high voltage can be ramped up or down very quickly (within approximately 10 seconds). In the past, drift chambers have not been built with the capability to change voltages this rapidly. However, their limitations are apparently due to design and construction, not as a result of intrinsic limitations or difficulties in technology.

References

1. P. Coyle, Lorentz Angle and Drift Velocity Program, SLAC, 1987
2. F. Sauli, CERN 77-09, 56, (1977).
3. J. Va'Vra, SLAC-PUB-2882, (1982).
4. J. Va'Vra, Aging of Gaseous Detectors. SLAC-PUB-5207 (1990).
5. J. Va'Vra, Review of Wire Chamber Aging. *Nucl. Instr. Meth.* **A252**, 547 (1986).

Report of the Particle Identification Group

P. Coyle, G. Eigen, D. Hitlin, P. Oddone
B. Ratcliff, N. Roe, J. Va'vra and T. Ypsilantis

1. Introduction

Particle identification systems are an important component of any detector at a high-luminosity, asymmetric B Factory.In particular, excellent hadron identification is rerquired to probe CP violation in B^0 decays to CP eigenstates. The particle identification systems discussed below also provide help in separating leptons from hadrons at low momenta.

We begin this chapter with a discussion of the physics motivation for providing particle identification, the inherent limitations due to interactions and decays in flight, and the requirements for hermiticity and angular coverage. A special feature of an asymmetric B Factory is the resulting asymmetry in the momentum distribution as a function of polar angle; this will also be quantified and discussed.

In the next section the three primary candidates, time-of-flight (TOF), energy loss (dE/dx), and Čerenkov counters, both ring-imaging and threshold, will be briefly described and evaluated. Following this, one of the candidates, a long-drift Čerenkov ring-imaging device, is described in detail to provide a reference design. Design considerations for a fast RICH are then described. A detailed discussion of aerogel threshold counter designs and associated R&D conclude the chapter.

2. Physics Requirements for Particle Identification

This section briefly summarizes the particle identification criteria which are dictated by the physics objectives of an asymmetric B Factory. We begin by discussing which particle species must be identified, and the momentum and angular distributions which are typical at an asymmetric machine. The limitations imposed by particle decays in flight are also considered. We will then consider a specific case, the "acid test" for any particle identification system: separating $B \to \pi\pi$ from $B \to K\pi$.[1] These rare two-body decays set the maximum momentum range. Finally, we will consider the problem of flavor tagging B's by means of a kaon tag; this will give us a more typical momentum spectrum.

In the most general case there are five particle species which one would like to be able to distinguish at a B Factory: electrons, muons, pions, kaons and protons. Electrons and muons will be identified by the electromagnetic calorimeter and the muon system. There may be some difficulty separating muons from pions at low momenta, so it would be helpful if the particle identification system could provide some μ/π separation below about 600 MeV/c. Fortunately, the low momentum range is where most devices under consideration work best and this complementarity is easily achieved. The remaining three particle species are all hadrons and require a device which can distinguish between them on the basis of mass. In general, the problem of π/K separation is more difficult than that of π/p or K/p separation. Protons are also much less frequently produced in B meson decays than are pions and kaons. We will therefore focus on π/K separation as the most important criterion for particle ID performance.

The momentum distribution of pions produced in B decays at an asymmetric machine with beam energies of 9 on 3.1 GeV is shown in Fig. 1a; the corresponding distribution for kaons is shown in Fig. 1b.

The distributions peak toward low momenta; the average value for pions is 0.56 GeV/c, and for kaons it is 0.85 GeV/c. Eighty-six percent of all pions have momenta below 1.0 GeV/c; the corresponding value for kaons is 70%.

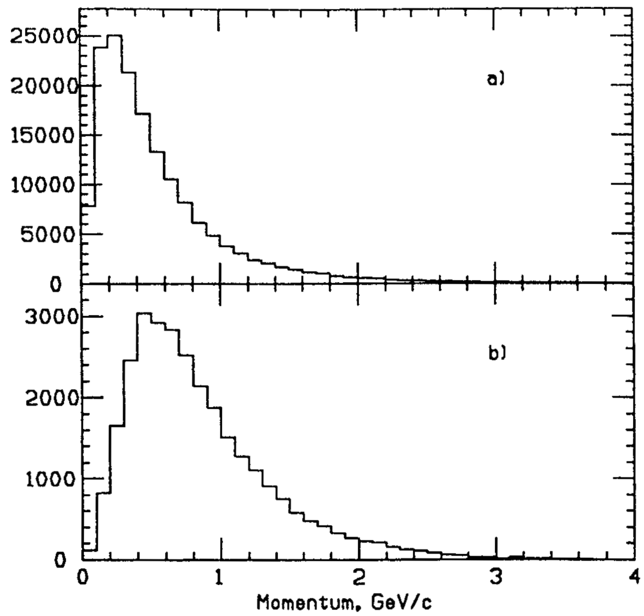


Figure 1. *The momentum distribution for* **a.** *pions and* **b.** *kaons produced in B meson decays for beam energies of 9 on 3.1 GeV.*

The angular distribution for pions is shown in Fig. 2; the angular distribution for kaons is very similar.

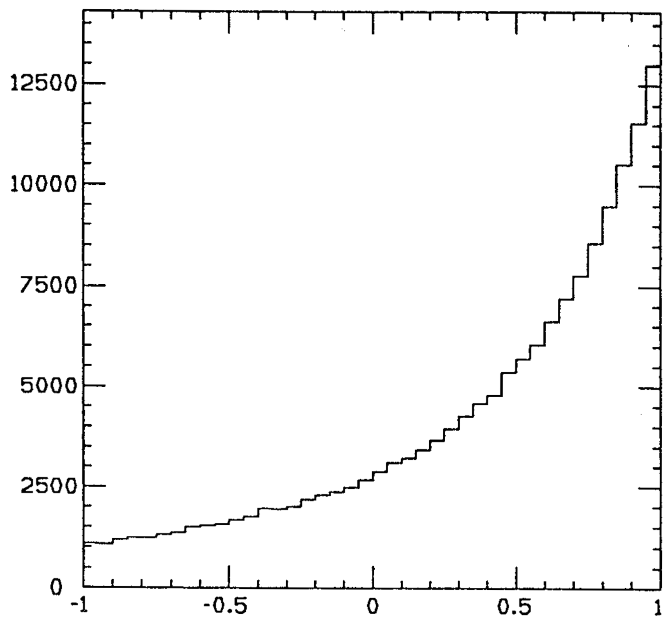


Figure 2. *The angular distribution for pions produced in B meson decays for beam energies of 9 on 3.1 GeV.*

The distribution is asymmetric and forward-peaked due to the asymmetry in the beam energies. In Fig. 3a the mean momentum as a function of $\cos\theta$ is shown for pions; Fig. 3b shows the same thing for kaons.

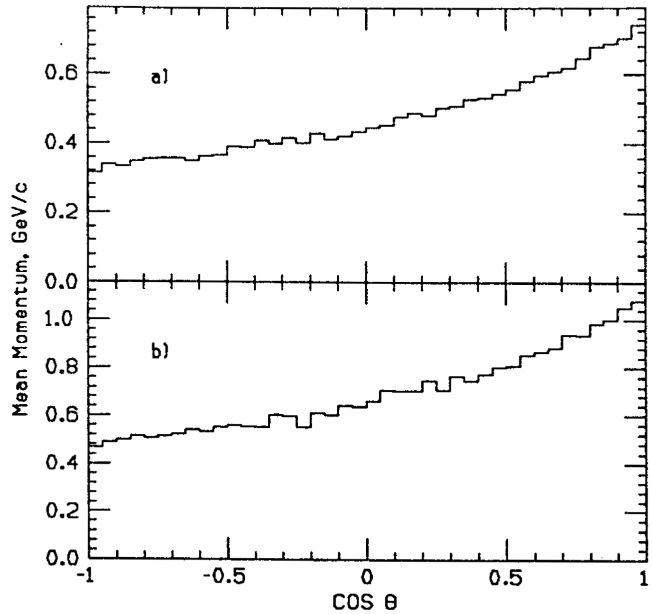


Figure 3. *The mean momentum distribution vs* $\cos\theta$ *for* **a.** *pions and* **b.** *kaons produced in* B *meson decays.*

In the forward direction the mean momentum increases; for $\cos\theta = 1$ it is more than twice that for $\cos\theta = -1$.

From Figs. 2 and 3 we see that significantly more particles are produced in the forward direction, with significantly higher momenta. Therefore, particle identification will generally be both more important and more difficult in this region. For TOF and dE/dx systems the increased path length for forward tracks relative to central tracks results in improved particle identification capability. This is especially important for the higher momentum tracks in this region.

Increased path length in the forward direction does have the disadvantage of increased particle decays. Decays in flight will prevent even the very best system from achieving perfect particle identification. The decay length distributions projected onto the plane perpendicular to the beam axis for pions and kaons produced in B meson decays at an asymmetric machine with beam energies of 9 on 3.1 GeV/c are shown in Figs. 4a and 4b.

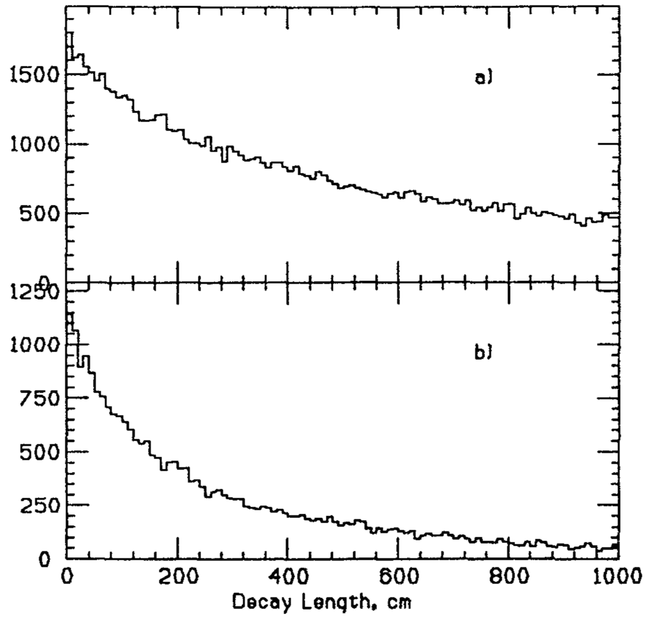


Figure 4. *The decay lengths projected onto the xy plane for* **a.** *pions and* **b.** *kaons produced in B meson decays.*

Since 7.8% of pions and 24.6% of kaons will decay inside a radius of 80 cm, a particle identification device located outside this radius will misidentify a significant fraction of kaons unless the kink can be reconstructed. Since 63.5% of charged kaons decay into a muon, this will also present a serious background for the muon identification system. If we include an endcap particle identification system located at z=150 cm, the fraction of decays is reduced to 6.7% for pions and 21.6% for kaons. Using dE/dx information from the drift chamber has the potential advantage over TOF or Čerenkov devices of reducing the misidentification probability due to decays in flight, although the precise level depends on the details of the dE/dx algorithm used.

Next let us consider the specific example of separating $B \to \pi^+\pi^-$ from $B \to K^+\pi^-$. These processes are expected to be rare, with calculated branching ratios on the order of 10^{-4} to 10^{-5}. A measurement of $B \to \pi^+\pi^-$ can be used to extract V_{ub} and, if the other B is tagged, to measure CP violation, while $B \to K^+\pi^-$ proceeds through a penguin diagram and is sensitive to the presence of heavy particles in the loop. These are important physics processes to measure, and they must be separated if the underlying physics is to be understood.

For the case of a two-body decay into $\pi^+\pi^-$ the momentum distribution of the pions is shown in Fig. 5, and the average momentum as a function of $\cos\theta$ is shown in Fig. 6.

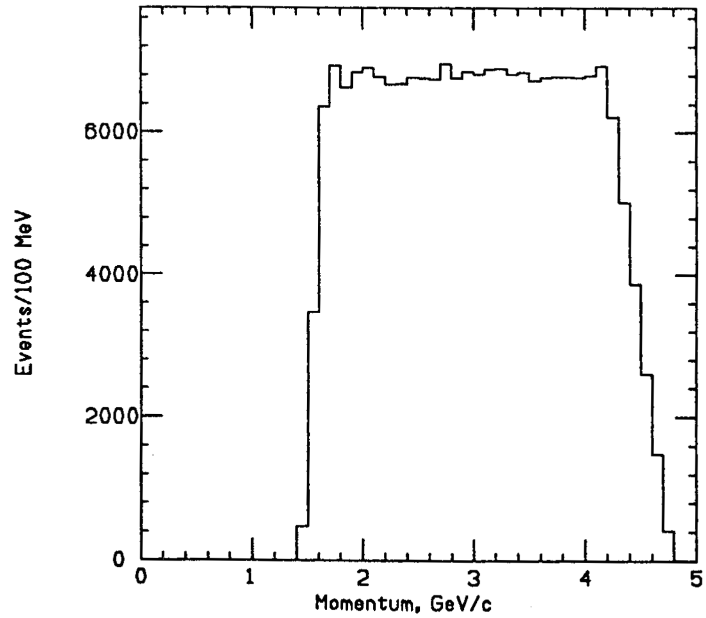


Figure 5. *The momentum distribution for pions from the decay $B \rightarrow \pi^+\pi^-$, for beam energies of 9 on 3.1 GeV.*

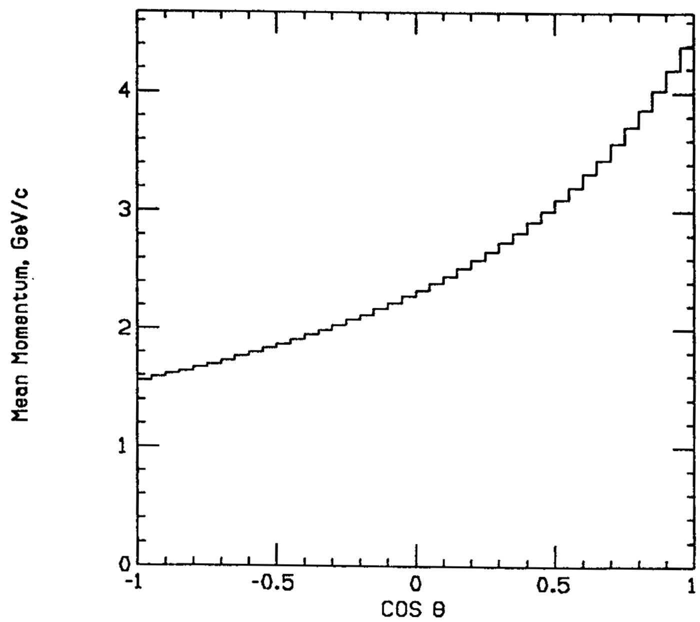


Figure 6. *The mean momentum vs. $\cos\theta$ for pions from the decay $B \rightarrow \pi^+\pi^-$, for beam energies of 9 on 3.1 GeV.*

By comparison with the corresponding distributions for generic B meson decays (Figs. 1 and 3) we see that the pions from the two-body decay have much higher momentum, and they will therefore pose a much bigger challenge to the particle identification system.

Let us first consider whether these two processes could be kinematically separated. For a resolution on transverse momentum given by $\sigma_{p_t}/p_t^2 = 0.23$, the reconstructed invariant mass of the $\pi\pi$ system has $\sigma_M = 22$ MeV/c^2. If the $K\pi$ final state is reconstructed assuming the pion mass for both particles, the mean of the resulting invariant mass distribution is 44 MeV below the nominal B meson mass and has $\sigma_M = 24$ MeV/c^2; see Fig. 7.*

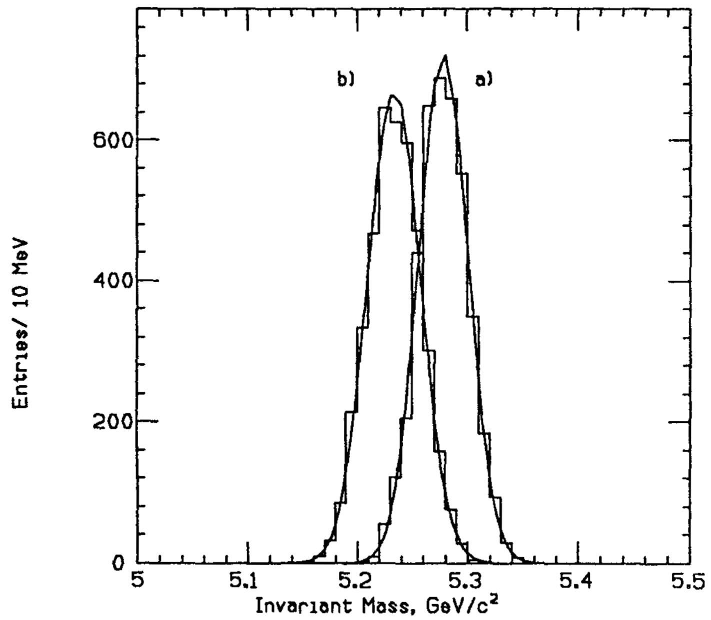


Figure 7. *The invariant mass distributions for* **a.** $B \to \pi^+\pi^-$ *and* **b.** $B \to K\pi$. *The latter was calculated assuming both particles were pions.*

It is clearly difficult to separate these processes kinematically. In principle one could impose a cut on the invariant mass requiring $M_{\pi\pi} > 5.290$ GeV/c^2, for example, which would reject 99% of the $K\pi$ events while retaining 30% of the $\pi\pi$ events. However, one loses efficiency in a channel which is rare to begin with, and furthermore one is forced to rely heavily on Monte Carlo simulation of the tails of the invariant mass distribution. For these rare decays it is desirable to have a particle identification system capable of performing $K\pi$ separation for momenta up to about 4.5 GeV/c. For this purpose it is likely that only a Čerenkov ring-imaging system will suffice.

Finally, let us consider the less stringent case of π/K separation for the purpose of kaon tagging. This technique, in which the sign of the kaon is used to tag whether a B or a $\bar{B}$ decayed, is important as a means of enhancing the tagged

* The mass resolution can be improved by an order of magnitude if one boosts back to the $\Upsilon(4S)$ rest frame and uses the beam energy constraint; however, in order to perform the boost one must know the particle masses.

sample for *CP* violation studies. The semi-leptonic decays are also useful for tagging, but have a smaller total branching fraction. Also, cuts must be applied to reduce the contribution from charmed semi-leptonic decays. Kaon tagging relies on the fact that in the cascade $b \to c \to s$, a b quark will produce a K^+ and a $\bar{b}$ quark a K^-. There is some wrong-sign contamination from Cabibbo-suppressed decays, which can be reduced by rejecting events with extra K^0_S's or multiple charged kaons.[2]

To evaluate these effects we have considered a simplified tagging scheme which requires one charged kaon and no extra K^0_S's.* Only those K^0_S's which decayed to charged pions were used as a veto. All particles were required to have $|\cos\theta| < 0.95$ and $p_t > 40$ MeV/c. With no decays and 100% particle ID, 38% of all events were tagged, with 94.0% of the tags being correct. If we turn on K decays in flight, 31% of events are tagged with 91% being correct. The increased fraction of wrong-sign tags is due to feed-down from events with multiple charged kaons. The effect of decays in flight is therefore to reduce the tagging efficiency by 18% while increasing the fraction of wrong-sign tags by 50%.

Next we will include the effect of particle identification. The momentum distribution for tagging kaons is shown in Fig. 8. It is slightly harder than the momentum distribution for all kaons; the mean is 0.92 GeV/c.

We will consider three possible particle identification systems; TOF with a resolution of 75 psec; dE/dx with a resolution of 7.5%; and a liquid CRID with the performance described in Section 1.4 below. Each device is assumed to be 100% efficient for $|\cos\theta| < 0.95$. An identified kaon must be within 3σ of the value expected for a kaon and at least 2σ away from that expected for a pion. The resultant efficiencies for a tagging kaon are 90% for TOF, 57% for dE/dx, and 99% for CRID. The dE/dx efficiency is low because the cross-over region between π's and K's is around 1 GeV/c, where π/K separation is poor but many tagging K's are found. The

* The B and D decays were simulated using the ASLUND decay table described in SLAC-353

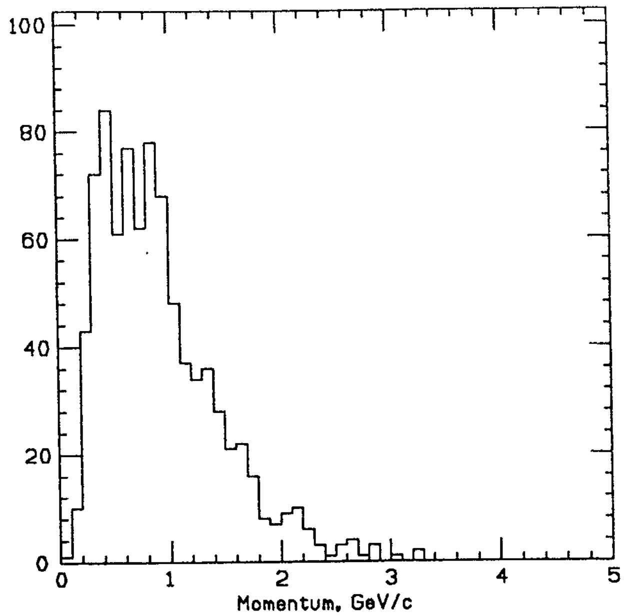


Figure 8. *The momentum distribution for tagging kaons.*

corresponding misidentification probability for π's is 0.3% for TOF, 0.9% for dE/dx and 0.04% for CRID. Given that there are approximately 6.3 π's per K in a typical B decay, the additional fraction of wrong-sign tags due to misidentified π's would be 2% for TOF, 6.4% for dE/dx, and 0.3% for CRID. (This takes into account the fact that events with two or more charged K's are not counted as tags.) Relaxing the particle ID requirements would improve the efficiency, but at the cost of increasing wrong-sign tags due to misidentified pions.

In conclusion, particle identification efficiency must be good in both the medium momentum range, for efficiency in kaon tagging, and at the high end of the momentum range, to separate rare two-body decays. A very good TOF system is adequate for the former but totally inadequate for the latter. The hypothesized dE/dx system has a disappointing performance for both, because the π/K crossover makes it inefficient for kaon tagging while the separation at high momenta is not very good. A CRID, which has the disadvantage of being the most expensive and complex device to build, gives the best performance in both cases. Other devices, such as threshold counters, or better quality dE/dx, remain to be evaluated in this model.

3. Technological Choices for Particle Identification

This section briefly summarizes the basic properties of the candidate technological choices available for particle identification at the B Factory. There appear to be only three candidates that could cover a fair portion of the momentum range with good performance; (1) time-of-flight (TOF); (2) ionization loss in gaseous detectors (dE/dx); and (3) Čerenkov counters. Moreover, the choice of technology is constrained not only by physics requirements, but by the machine environment and the properties of the remainder of the detector. Thus, a number of properties of each technology, in addition to specific performance, must be considered, including (1) mass; (2) size/space requirements; (3) decay/interaction limits to performance; (4) performance at high rates; (5) triggering considerations; (6) coverage and hermiticity; (7) construction and operating costs.

The TOF technique is clearly useful and well understood at low momentum, but above about 1.5 GeV/c requires either that the resolution reach unprecedented precision or that the flight path (and thus the whole detector) be unacceptably large. Modest quality dE/dx in a central tracker sufficient to provide good hadronic separation in the $1/\beta^2$ region and to help in low momentum e/π discrimination is likely to exist in any modern tracking device essentially for "free." However, high momentum π/K separation (above about 1.5 GeV/c) requires working in the relativistic rise region of the energy loss curves, and leads to a requirement for many samples (and therefore a large chamber), and also suggests a heavier gas (perhaps pressurized) than would be desired for the best tracking resolution. In addition, the existence of a "cross-over" between the $1/\beta^2$ and relativistic rise regions requires that it must be combined with another technology (such as TOF) if complete momentum coverage is to be attained. Such a combined system could perform the particle ID task adequately and does have a number of attractive features. The central chamber is much larger and more massive than a tracking only device, but it does "double duty." Track matching is simple and the device is very uniform geometrically. Moreover, the large dip angles (with the highest momentum tracks) also have the largest sample lengths and thus the best resolution, matching the requirements of the asymmetric design. The particle ID in such a device starts at the smallest possible radius and thus should be the least sensitive to misidentification due to particle decay. Finally, the separation attainable in the relativistic rise region is nearly independent of momentum from 2 GeV/c to well above the maximum momentum seen at the B Factory.

The third (and leading) contender for hadronic identification at the B Factory is the Čerenkov technique. As will be discussed below, either threshold or ring imaging techniques could provide adequate performance, in principle, provided that the efficiency for Čerenkov photon detection can be made sufficiently high. The primary conceptual advantages over dE/dx are that the devices (1) are generally rather thin radially (typically around 20 cm), leading to a smaller overall detector size and less expensive calorimetry; (2) are fully modular and separable from the tracking, allowing the optimization of the tracking independent of particle ID concerns; (3) have excellent nominal separations (especially for the ring imaging devices) over the phase space of the B Factory; (4) are quite robust against degradation in performance, in contrast to the requirements on dE/dx in the relativistic rise region. Conversely, the Čerenkov devices are probably even more challenging to construct and operate, may add more mass in front of the calorimeter, and are probably more difficult to build in a fully hermetic configuration.

In the following subsections, the general properties of each technology are discussed in somewhat more detail. The design realizations for a B Factory are presented in the subsequent sections.

3.1. The Time-of-Flight Technique

The time-of-flight (TOF) technique is well-known; this review will therefore be brief.[3] It has been extensively applied in earlier experiments to identify low momentum particles. In e^+e^- collider experiments, a typical system consists of a cylindrical bank of scintillators positioned so as to maximize the flight path, *e.g.*, surrounding the central tracker just inside the calorimeter. The time of passage of the particle through these counters stops a clock which was started by the beam crossing signal from the machine giving the time-of-flight (T). Photomultiplier tubes are generally included on each end so as to give two independent measures of the stop time, and a $\sqrt{2}$ improvement in resolution. The length (L) of the particle's trajectory is calculated with high accuracy from measurements in the tracking chamber. The velocity (β) of the particle can be calculated as $\beta = L/cT$.

The TOF technique using scintillation counters is simple, well-understood, and robust. The problem for the B Factory is that it is not effective at sufficiently high momentum. The fractional error on a particle of mass (M) and velocity β scales like γ^2 (*i.e.* like $1/(1-\beta^2)$) for $\beta \approx 1$. Thus, for fixed time resolution, the path length required to attain a given separation between particle species is a rapidly increasing function of momentum.[4]

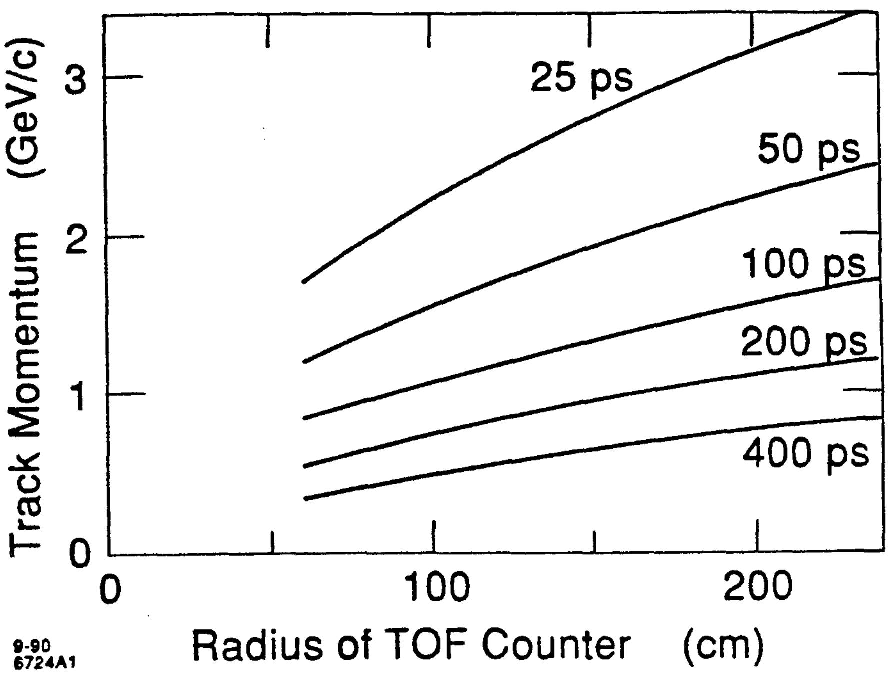


Figure 9. *Maximum momentum at which π/K separation exceeds 3σ as a function of the TOF counter radius and resolution.*

Figure 9 gives the upper limit of the momentum range over which 3σ π/K separation is possible as a function of the radius of the TOF system. The situation is shown for the shortest path length particles in the usual solenoidal geometry, and assumes a 1.0 T field, although the path lengths for fast particles are nearly independent of the field. Since the path length L increases approximately like $1/\cos\theta_d$, where θ_d is the dip angle, an asymmetric collider has a "natural" path length enhancement in that the highest momentum particles which tend to go forward get the largest path lengths. On the other hand, a large magnetic field will cause the softest particles to curl and be lost from TOF identification.

The best resolutions which can be attained with "conventional" scintillation counters in a collider environment are optimistically around σ=100 ps.[5] This limits the π/K separation to momentum below 1.1 GeV/c for a TOF counter with a 1 m radius. More typically, resolutions obtained in experiments are around 110 to 150 ps.[6] In principle, other technological approaches to fast timing, such as spark gap counters, can attain substantially better resolution.[3] However, they are difficult to work with, and have yet to be used successfully in a large 4π acceptance detector. If a TOF scintillation counter were combined with dE/dx in a tracking chamber, the radius of the counter would need to be sufficient to cover the π/K crossover region in ionization loss which extends from about 800 to 1600 MeV/c. This leads to a minimum radius counter of about 2 m, assuming 100 ps resolution. Though reasonably consistent with the radius needed for a good dE/dx device operating in the relativistic rise region, it leads to a large, expensive calorimeter, and the large path lengths before TOF ID will lead to misidentification problems for particles which decay.

There are a few other issues of B Factory TOF counter design which should be briefly mentioned. First, the resolution obtainable in a scintillation TOF counter depends on the length of the counter, mostly due to photon absorption inside the scintillator. A typical effective absorption length is about 2 m, which limits the practical counter length to be around 4 to 5 m. With a device at a radius of 1.6 m, a 4 m long counter could cover only down to 39° in the lab, necessitating end cap counters to cover the forward/backward regions. Second, the variation in flight and collection time combined in a long counter is of order 15 to 20 ns. So when the design machine bunch spacing is of this order or less, a simple, single hit indication of which bunch to associate with the track is no longer available. Third, the start time of the TOF system is determined synchronously by the beam crossing, but the event can occur at any time within the colliding beam overlap time. Thus, the longitudinal size of the B Factory beams provides a "hard" limit to the resolution attainable in the system. For example, if each beam is 1 cm long, the contribution to the resolution from the start time alone is about 50 ps. In principle, this resolution could be improved stochastically by averaging over the tracks in the event. Fourth, a TOF counter system is massive, placing about 15% of a radiation length of material in front of the calorimeter. On the other hand, the good pulse height resolution necessary to attain the best timing resolution also makes it possible to measure the energy deposited, so the low energy performance of the calorimeter should not be significantly affected.

3.2. The dE/dx Technique

In recent years, many of the large-scale detectors which have been built have incorporated a measurement of particle energy loss (dE/dx) in a gas as a central element of their design. The acronyms of some of the detectors that have succesfully implemented this technique are TPC, EPI, JADE, CLEO, ARGUS, OPAL, ALEPH, DELPHI, and MARK II; and this list is far from exhaustive.[7] The energy loss of a particle in a gas is a function of the properties of the gas and the velocity of the particle, as shown in Figure 10. It is conventional to divide this curve into a low velocity region (the $1/\beta^2$ region; the minimum ionizing region (around $(\ln(\beta\gamma) = 1)$; the relativistic rise region; and the relativistic plateau at high velocity. Since the curve of energy loss versus velocity for a particular gas depends only on particle velocity and gas composition, when coupled with a momentum measurement in the magnet, it determines particle mass.

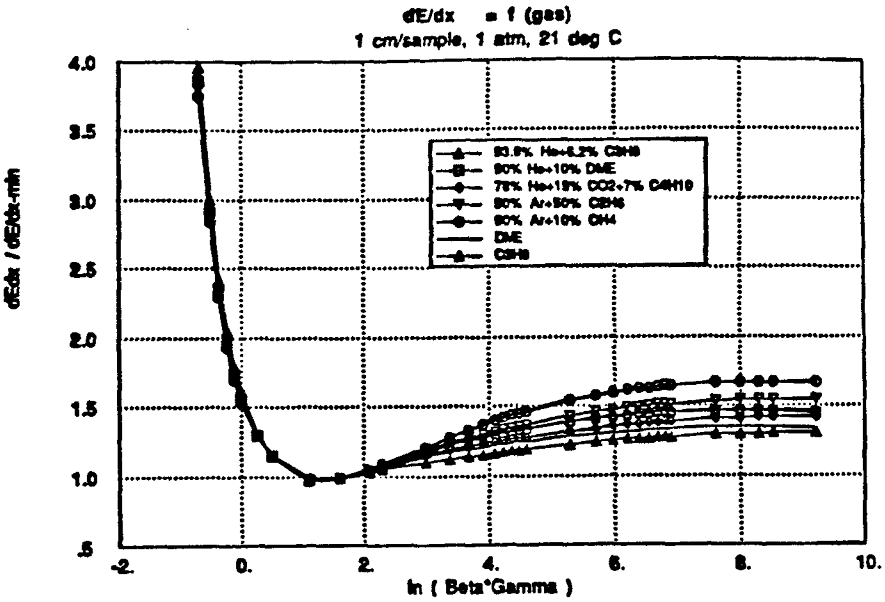


Figure 10. *Most probable energy loss (dE/dx) versus* $\ln(\beta\gamma)$ *for several different gas mixtures at 1 atm. pressure and 21°C.*

The identification is not always unique. There are ambiguities in the so called "cross-over" regions, where two particles of different mass but the same momentum produce an equal energy loss (see *e.g.*, Figure 12 below). In these regions, another method of particle identification is required. As mentioned above, a major advantage of using dE/dx for particle identification is that the detector that measures the energy loss can simultaneously measure the track of the particle, saving space and greatly reducing confusion in associating the energy loss measurements with tracks.

To achieve good hadron-hadron identification in the relativistic rise region, the device must be very large (≥ 150 cm); have many samples; operate at high gas pressure; or, most likely, all of the above. There is no "magic gas" choice which can alter these basic requirements. This can be shown either empirically,[8] or using a Monte Carlo calculation,[9] or using a phenomenological calculation.[10] This results from the fact that any single energy loss measurement on a thin gaseous sample has large Landau fluctuations, and at the same time different particles do not yield very different energy losses, with the exception of the nonrelativistic region. Examples of experiments with good dE/dx performance are TPC, ALEPH, and OPAL, and they provide some examples of different optimizations. TPC has 183 4 mm samples over a path length of 73 cm at a gas pressure of 8.5 atm.; ALEPH has 340 4 mm samples over a path length of 136 cm at a gas pressure of 1 atm.; and OPAL has 120 1.3 cm samples over a path length of 160 cm at a gas pressure of 4.0 atm.

However, dE/dx measurement can still be useful for smaller devices, only1 m long, with 60–100 samples operating at 1 atm. In this domain one can perform hadron-electron identification into the few GeV/c range, and hadron-hadron identification in the nonrelativistic region (below $\sim$ 1 GeV/c). If the nonrelativistic region is the only region of interest, it is advantageous either to use heavy gases such as propane or DME, or to increase the gas pressure, in order to improve the resolution. In the relativistic region, this trick is not as effective because one also reduces the relative height of the relativistic plateau with respect to the minimum energy loss, thus negating much of the advantage gained from smaller resolution. An example of a detector optimized for the the lower velocity region is ARGUS, which uses propane gas.

In understanding the relationship of parameter choices made in various designs, it is helpful to use the following approximate empirical scaling relationships[11] for the resolution on dE/dx:

$$\begin{aligned} dE/dx \text{ Resolution} &\sim (\text{Pressure})^{-0.32} \\ &\sim (\text{Sample Thickness})^{-0.32} \\ &\sim (\text{Total Number of Samples})^{-0.43} \end{aligned}$$

The dependence of the resolution on gas composition can be expressed approximately by the following formula (see the notes following the table below for definitions of variables):

$$dE/dx \text{ Resolution} \sim \left(\frac{\alpha t}{I\beta^2}\right)^{-0.32}$$

In the relativistic region, the size of the relativistic plateau can be affected by both the gas pressure and the gas density. Figure 10 shows that the denser gases have a smaller relativistic plateau. Similarly, the pressure dependence of the relativistic rise can be parametrized approximately as follows (see the notes following the table below for definitions of the variables):

Relativistic Rise $(E_{max}/E_{min}) \sim 1.6 \times (\text{Pressure})^{-0.09}$

The incorporation of dE/dx measurement in a chamber requires that a number of tight constraints on the tracking device design be enforced. Although these are now generally understood, they should not be underestimated. Usually they degrade the dE/dx performance achieved in practice. Therefore, one should assume that a "minimal dE/dx system" must be designed with at least 4σ of theoretical particle separation. A few areas of concern follow:

1. All mechanical tolerances must be carefully controlled and their stability ensured.
2. The high voltage power supply must be kept stable.
3. The total gas gain should be limited to $\sim 5 \times 10^4$ if a jet cell with charge division is used, and preferably to a much lower value in a TPC. This is necessary to reduce the track angle dependent corrections due to the wire gain saturation effects.
4. Wire-to-wire cross-talk in jet cells should be compensated by hardware.
5. Pulse digitization techniques (*e.g.*, flash-ADC's) should be used to resolve close tracks in the jet cells. For this reason, a chamber design with small cells may be easier to use.
6. A number of background problems and associated space charge effects, efficiency gating, etc., can distort a track and possibly affect the dE/dx sampling.
7. A dE/dx measurement must deal with many corrections, e.g., because of the gain variation due to temperature, barometric pressure, and gas composition; or other effects due to wire staggering, electrostatic and gravitational deflections, space charge effects, etc.
8. Many details of the electronic design can also affect the performance. Due consideration must be give to electronic noise, stability, base line shifts, cross talk, dynamic range, chamber termination, etc.

B Factory Design

From a tracking and calorimetric perspective, the optimized tracking device for a B Factory should be modest in size and low in density. Here, we will consider the dE/dx performance of such a chamber with 60 (1 cm) samples operating at 1 atm. pressure. The performance of helium-based gases will be compared with the performance of heavy gases like propane or DME. The calculations are done using the phenomenological method.[10] A similar method was used in the low momentum regime for the Tau-Charm Workshop.[10] Here, we extend the results up to 6 GeV/c.

Table 1 shows a brief summary of parameters for a variety of gases, including their expected dE/dx resolution for a single 1 cm sample, and the size of the relativistic plateau as calculated using the phenomenological method.[10] The dE/dx value corresponds to the most probable energy loss. The resolution is scaled to one sample after performing a "truncated mean" operation in which about 40–50% of the highest samples are thrown out to get rid of the Landau tail.

Table 1. *Parameters of various which are relevant to dE/dx measurements.*

Gas	ρ(at 21°C and 1 atm.) (g/cm3)	I (ev)	$\frac{\alpha t}{I}$	$\frac{E_{max}}{E_{min}}$	FWHM(E^{-1})
C_3H_8	1.88×10^{-3}	50.3	3.39	1.30	0.55
90% Ar + 10% CH_4	1.57×10^{-3}	191.2	0.58	1.67	0.96
50% Ar + 50% C_2H_6	1.46×10^{-3}	129.1	0.89	1.54	0.84
DME	1.89×10^{-3}	59.8	2.74	1.35	0.59
90% He + 10% DME	3.39×10^{-4}	43.8	0.64	1.47	0.94
78% He+15% CO_2+7% C_4H_{10}	5.76×10^{-4}	50.0	0.93	1.43	0.83
93.8% He+6.2%C_3H_8 [13]	2.73×10^{-4}	42.5	0.53	1.48	0.99

$E_{max}/E_{min} = \frac{dE/dx_{max}}{dE/dx_{min}}$ is the ratio of the relativistic plateau to the minimum energy loss; FWHM(E^{-1}) is the full width at half maximum of the resolution for a single 1 cm sample divided by the most probable energy loss; I is the mean ionization potential of the gas mixture; ρ is the gas density; t is the sample thickness (1 cm); and αt=0.153 (Z/A) ρt ($\frac{\text{MeV}}{\text{gm/cm}^2}$).

Figure 11 shows the predicted performance of this system for π/K separation in the $1/\beta^2$ region as a function of different gases. Dense gases clearly perform better. Above 1 GeV/c, π/K separation is hopeless for such a small dE/dx system with any gas. Perhaps there is some modest π/p separation, as can be seen in Figure 12.

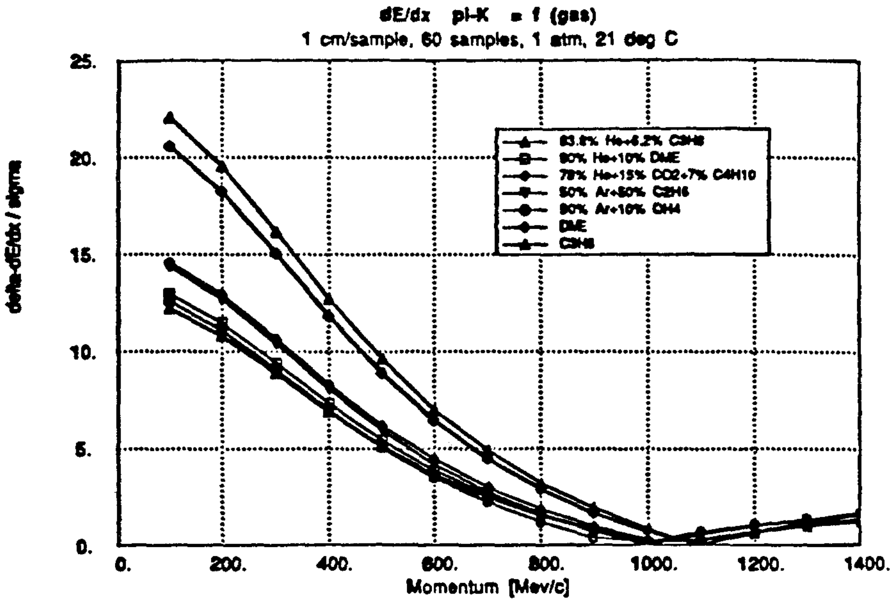


Figure 11. *Predicted π/K separation in the $1/\beta^2$ region for a variety of gases in a "small" tracking chamber (see text).*

However, as shown in Figure 13, hadron-electron identification is quite satisfactory over the entire momentum region (except for the cross-over regions), even with a relatively small dE/dx system.

One may ask how to change the tracking design in a minimal way so that the dE/dx system is capable of doing π/K particle identification above 2 GeV/c. There is no unique prescription for such a chamber, but it must have the general features discussed above. Here, we consider a device with 200 samples each 1 cm long, running at 2 atm. with helium-based gas. Figures 14 and 15 show the results.

Clearly, while hadron-electron, π/p and π/K separations are acceptable, p/K separation in the region above 2 GeV/c is not. The cross-over regions are very large and difficult to cover with other techniques. Moveover, such a large device would increase substantially the size (and the cost) of the calorimetry, and the increased mass in the pressure vessel

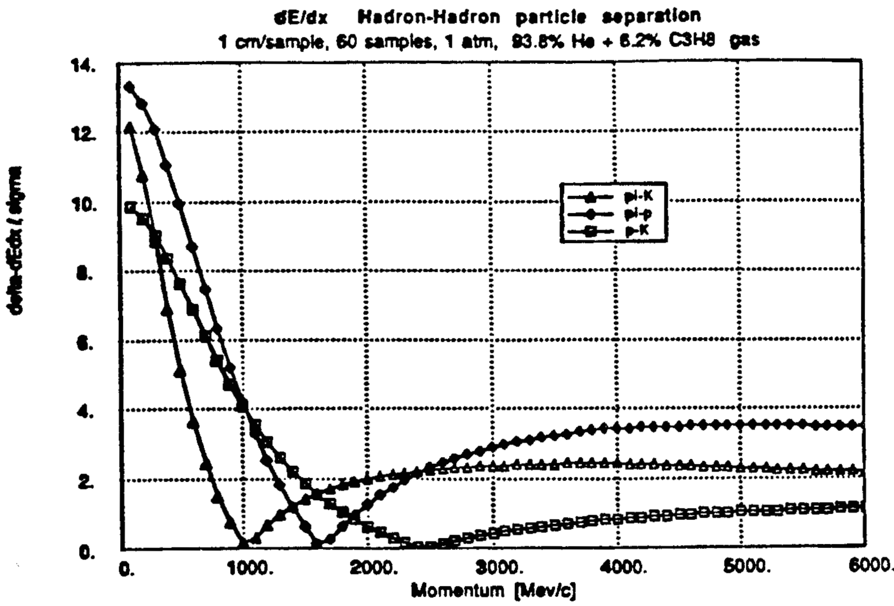


Figure 12. *Predicted hadron-hadron separation performance as a function of momentum for a "small" tracking chamber (see text).*

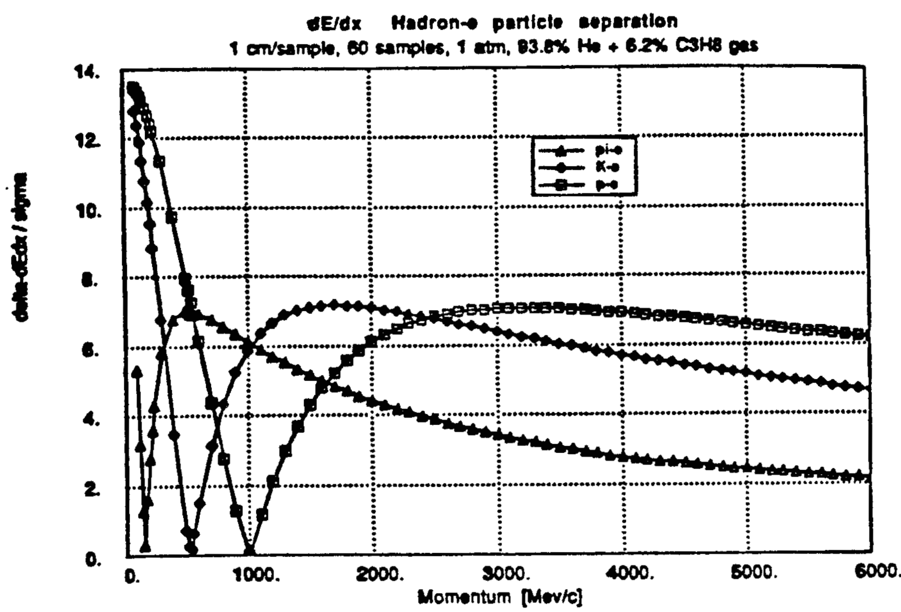


Figure 13. *Predicted hadron-electron separation performance as a function of momentum for a "small" tracking chamber (see text).*

walls would decrease the low momentum tracking resolution and the low energy photon detection efficiency and resolution.

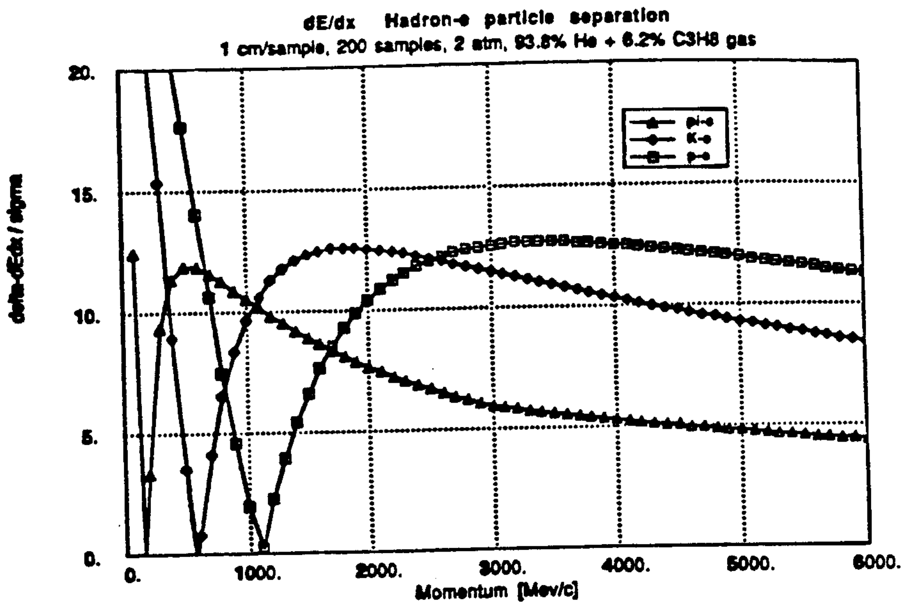


Figure 14. *Predicted hadron-electron separation performance as a function of momentum for a 200 cm B Factory chamber design (see text).*

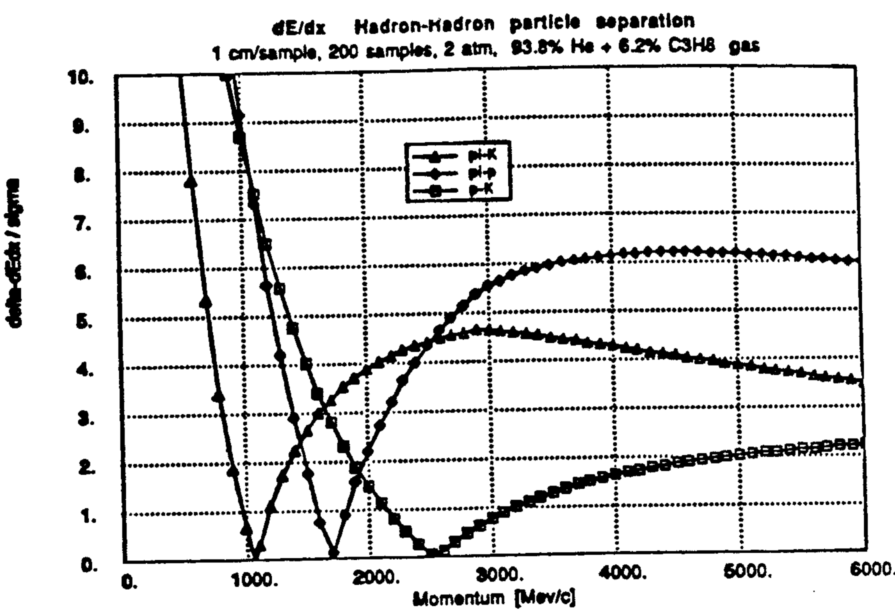


Figure 15. *Predicted hadron-hadron separation performance as a function of momentum for a 200 cm B Factory chamber design (see text).*

3.3. *The Čerenkov Technique*

Introduction

The detection of Čerenkov radiation has been an invaluable tool for measurement of charged particle velocity since its first quantitative measurements in 1934 by Čerenkov and subsequent theoretical explanation by Tamm and Frank in 1937. When combined with a momentum measurement of the charged track in a magnetic field, excellent particle identification discrimination is achievable. In this section we first briefly discuss the theory of Čerenkov radiation and collect together some equations relevant for detector design. The various choices and options available to the Čerenkov designer are then discussed and technologies relevant to B Factory designs are considered. A number of excellent reviews cover most of this material in detail.[12,13]

Some Čerenkov Theory

The passage of a charged particle of velocity $v = \beta$, through a dielectric medium of refractive index n will induce transient dipoles within the atoms of the medium. When the particle velocity is greater than the velocity of light in the medium, the radiation emitted from these dipoles adds coherently to give a detectable signal. The Čerenkov radiation is confined to a cone of angle $2\theta_c$ about the particle direction, where the cone angle is given by,

$$\cos\theta_c = 1/\beta n. \tag{1.1}$$

The emission angle is random in the azimuthal direction. As $\cos\theta_c$ can never be greater than one, Čerenkov radiation can only be emitted if

$$\beta > \frac{1}{n} \tag{1.2}$$

The Čerenkov radiation has an energy dependence (E) and is detectable between two energies (E_L and E_H). It can be shown that the number of observed photons N_D per unit radiator length (L) is given by

$$\frac{dN_D}{dL} = 370 \cdot Z^2 \int_{E_L}^{E_H} \left(1 - \frac{1}{(\beta n(E))^2}\right) \epsilon(E) dE. \qquad (1.3)$$

where Z is the charge of the particle; the energy unit is electron volts; and the length unit is centimeters (cm). The efficiency factor $\epsilon(E)$ parametrizes the overall detection efficiency and may include losses due to window and radiator transmissions, reflection at surfaces, photocathode efficiencies of the detector, etc. The integration bandwidth is essentially the range over which ϵ is finite and $\beta n(E) > 1$.

Under the approximation that the medium is nondispersive, it is conventional to clump the efficiency factor with the numerical constant to form a new constant N_0 given by

$$N_0 = 370 \int_{E_L}^{E_H} \epsilon(E) dE \qquad (1.4)$$

Thus, for a charged particle of unit charge the number of detectable photons is

$$N_D = N_0 L \left(1 - \frac{1}{(\beta n)^2}\right) = N_0 L \sin^2 \theta_c \qquad (1.5)$$

N_0 is the figure of merit for a detector and allows detector performances to be compared independently of the radiator length and refractive index. Typical N_0's for useful detectors range from 30 to 200 cm^{-1}. For example, a photomultiplier tube operating in the visible ($E_L = 2\ eV, E_H = 3.2\ eV$) with an average efficiency of 13% would yield an N_0 of $\sim$ 60 cm^{-1}. Assuming a 1 cm thick radiator of index 1.27 this would correspond to about 30 photons for a $\beta = 1$ particle. Photoionization gas detectors operating in the UV achieve comparable or larger values.

In practice, the available radiator media are all dispersive. This chromatic distortion smears the emission angle of the photons as a function of energy and becomes a significant effect for imaging Čerenkov counters operating in the near vacuum ultraviolet region. For energies beyond the vacuum ultra violet, $n(E)$ becomes less than one and no Čerenkov photons are produced.

Čerenkov Counter Type

The information from Čerenkov radiation can be utilized in two distinct ways. This has given rise to two different types of counters. The first type, called a threshold counter, is designed to measure the light emitted by charged tracks with β above threshold as given by Eq. (1.2). In its simplest form, it provides a binary flag as to whether the track velocity is above or below the threshold. In most modern counters, the number of Čerenkov photons is also measured approximately, allowing the separation range to be extended in momentum somewhat above the threshold for the heavier particle. The threshold counter is simple, and can be designed to cover large acceptance. To provide useful information for complex events, the detector must be split into many cells whose size is determined by the typical track separation. A counter with one radiator can usually only distinguish between two particle hypothesis over a rather limited momentum range. Additional counters containing radiators with different indexes of refraction can be used to extend coverage and distinguish between more than two particle types.

The second type of device, called a focusing counter, not only counts the number of Čerenkov photons, but also measures their emission angle. Before the advent of position-sensitive photodetectors, such counters were built so that light was detected only when emitted within a particular narrow angular range. These devices, called differential counters, have very small acceptance in both angle and momentum space, and are of no interest in a B Factory secondary particle detector. They were used instead to identify particles in high momentum beamlines.

The development of the gaseous photocathode coupled to a position-sensitive gas detector[14] has made the ring imaging focusing detector of the RICH (or CRID) type possible. In these detectors, both the number of photons and their emission angles are measured. These devices have very large acceptance in both angle and momentum, and excellent performance for closely spaced tracks. The effective resolution of an imaging counter in the region well above threshold is greatly superior to a threshold device as given by the derivative of Eq. (1.1),

$$\delta_\beta(\text{Imaging}) = \frac{\sigma_\beta}{\beta} = \tan\theta_c \cdot \frac{\sigma_\theta}{\sqrt{N_D}}, \tag{1.6}$$

where δ_β is the percentage error on β, σ_β is the error on β, and σ_θ is the measured angular error on each detected photon. This may be compared with the resolution of a threshold counter as calculated from Eq. (1.5)

$$\delta_\beta(\text{Threshold}) = \frac{\sigma_\beta}{\beta} = \frac{\tan^2\theta_c}{2\sqrt{N}}. \tag{1.7}$$

Thus, the resolution of the imaging counter is better than that of the threshold counter by a factor R;

$$R = \frac{\tan\theta_c}{2\sigma_\theta}. \tag{1.8}$$

The precise values of R depend on a number of factors, including radiator, detector resolution and geometry, particle momentum, etc. In principle, for an optimized gaseous detector it can exceed 250.[13] More typically, in an existing detector, the SLD CRID, $R = 140$ for the liquid radiator and 40 for the gas.[15] Thus, the separation in an imaging counter of the SLD type between a light and heavy particle extends over a momentum range from just above the high particle threshold to over five times the threshold momentum of the heavy particle. In contrast, the momentum region of reasonable

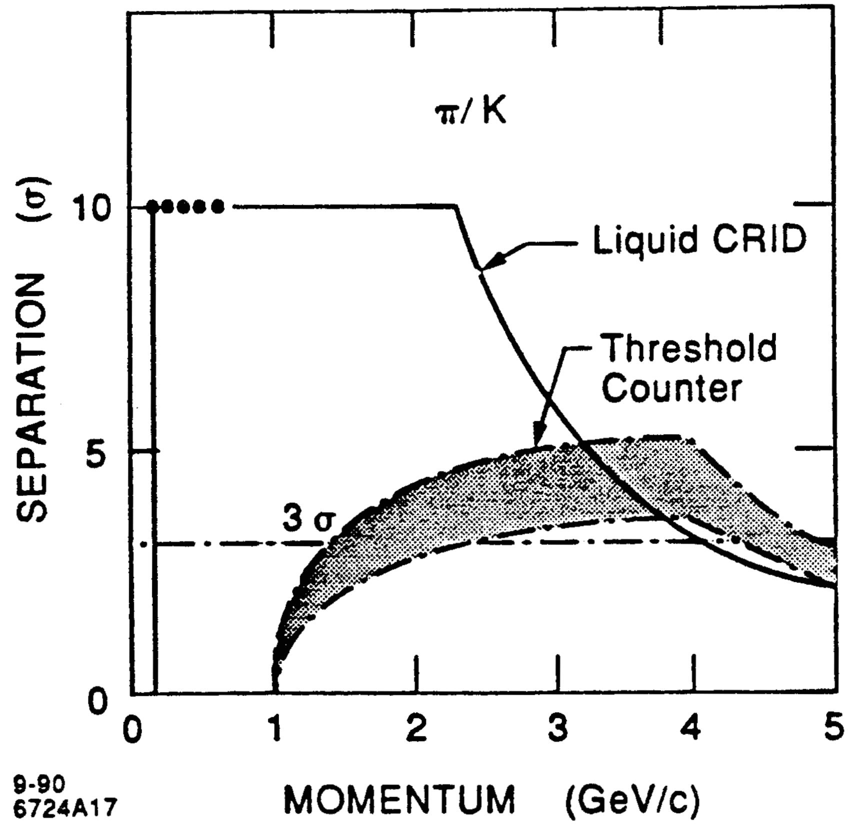


Figure 16. *The separation of a liquid CRID compared to that of a threshold counter filling the same space.The range of the threshold counter is chosen to give high momentum π/K separation.*

separation in a threshold counter extends from $\sim$ 20% above the light particle threshold to $\sim$ 20% above the heavy particle threshold. Figure 16 compares the separation expected for the liquid imaging counter (CRID) described in the following section with a threshold counter filling the same space. The index of refraction n for the threshold counter is chosen to be $n = 1.008$ so that the counter is optimized for π/K separation in the high momentum regime (1–4.5 GeV/c) at the B Factory. N_0 is taken to be 60 cm^{-1} and L is 15 cm. The number of detected photons is then 14, which is essentially the same as the standard CRID at 64°. The separation of a threshold device is very asymmetric—it depends on whether the true particle is below or above threshold—so that translating separation into the equivalent Gaussian sigma (σ) appropriate for the CRID is not well-defined. The shaded band in Figure 16 gives a spread which encompasses two different definitions; the highest is based on a likelihood test statistic and the lower on the number of standard derivations a true π is separated from the K hypothesis.

Clearly, either device can give adequate π/K separation for the higher momentum region at a B Factory, provided that the segmentation of the threshold counter is sufficient. However, the CRID also has excellent performance in the region below 1.0 GeV; can identify protons over the whole region; and e's and μ's at lower momentum as well.

Photodetectors

The development of the vacuum photomultiplier tube in the 1940's led to practical threshold Čerenkov counters by providing efficient ($\sim 20\%$) detectors of single photons in the visible region. More recent tubes have bialkali cathodes, and special window materials, giving bandwidths into the vacuum ultraviolet and higher peak efficiency. Phototubes can be very fast devices, so that non-event associated background is not usually a problem. On the other hand, conventional tubes have no position resolution (except the tube aperture) and they are very sensitive to magnetic fields. More recently, tubes have been developed which perform rather well in fields up to 1 Tesla provided that the tube is oriented properly in the field.

Solid state photodetectors have also become available recently. Typically, their efficiency is quite high ($> 25\%$) at longer wavelengths but cuts off in the UV below about 320 nm. Perhaps the most promising device for single photon detection is the avalanche diode, described below. At room temperature, their noise ratio is high. Operation for single photon detection requires low temperatures and/or coincidence techniques. The sensitive area of these devices is also very small (~ 0.1 mm^2) and expensive per unit area covered. Somewhat larger devices may become available soon but flux concentration techniques are necessary if a large area counter is to be covered with such a photon detector.

The development of gaseous photocathodes contained within conventional time projection or multiwire chambers has enabled the inexpensive instrumentation of large areas with fine position resolution (~ 1 mm), and made the ring imaging Čerenkov counter a reality. A number of high efficiency gases are known, but only those operating at the lowest photon energy [triethylamine (TEA), and tetrakis-dimethyl-amino-ethyline (TMAE)] are commonly used. Of these two, TMAE is the more widely used since it has significant efficiency ($\sim 40\%$) for photon energies below $\sim$ 7.5 eV and can therefore be used with quartz windows. TEA requires operation in the far vacuum ultraviolet which makes finding appropriate radiators difficult and demands that expensive materials, such as CaF, and MgF_2, be used for windows.

TMAE is a truly nasty gas. Not only does it react very strongly with air, but it also attacks common construction materials; and the reaction products are quite damaging to the chamber. Thus, the detector materials must be carefully chosen and the host gas must be kept very pure. TMAE has a relatively long absorption length ($\sim$ 2.5 cm) at room temperature, and the detector will therefore be rather thick and slow unless operated at high temperature. Conversely, room temperature TEA detectors can be quite thin since the absorption length is very short (0.6 mm).

Research aimed at finding suitable candidates for position-sensitive gas-filled detectors continues.[16,17] Encouraging results have recently been obtained with a CsI cathode with an adsorbed TMAE layer. This cathode has efficiencies which exceed that of TMAE gas ($> 40\%$ below 7.0 eV) and appears to be stable and robust. This work is discussed in more detail below.

Radiator Medium

The radiator medium must conform to a number of criteria. It should have the correct refractive index for optimum particle identification separation; have low chromatic dispersion, to reduce the contribution of the chromatic error to the smearing of the emission angle and the threshold momentum; have high photon transmission over the detector bandwidth; have good radiation resistance so that the transmission does not deteriorate significantly on exposure to the few kilorad dose expected during the lifetime of the experiment; have negligible scintillation light compared to the Čerenkov light emission; have small radiation length, to reduce the probability of photon conversions in the detector; and should be mechanically and chemically robust.

The choices available in practice are heavily constrained. Room temperature gases have indices ranging from 1.000035 (He) to ~ 1.0018 (*e.g.*, C_5F_{10} at 30°). The gas pressure can be varied to give somewhat higher indices but a pressure vessel is then required. The lowest index liquids (*e.g.*, C_6F_{14}) candidates have an index of ~ 1.27, and transmit to photon energies of ~ 7.0 with modest chromatic dispersion. The indexes of refraction for solids in the visible range from 1.32 (NaF) to about 6.0. Of the solids, the alkali halides (*e.g.*, NaF, LiF, CaF) have the best combination of UV transmittance, low indexes, low dispersion, and ease of handling for use in imaging counters; but they are generally quite expensive. In any case, the best performing proximity counters in the *B* Factory environment use liquid radiators primarily because the index of refraction and chromatic dispersion are less. In the limit of no dispersion, it can be shown that the measurement error on the velocity (β) scales with the index of refraction (n) as

$$\sigma_\beta = \frac{n\beta^2\sigma_\theta}{\sqrt{N_0 L}}. \tag{1.9}$$

Thus, for a fixed length radiator (L) and constant quality factor (N_0), the resolution scales with n, so that the resolution for a liquid counter with $n = 1.28$ (C_6F_{14}) exceeds that of a solid counter with $n = 1.5$ (NaF at 8 eV) by $\sim 15\%$.

Indices of refraction between the high index gases at ~ 1.0018 and the low index liquids at ~ 1.27 are quite difficult to attain. Unfortunately, the indices needed for a *B* Factory threshold counter fall in this range as shown in Table 2.

Table 2. *Threshold momenta for π's, K's and p's for several indices of refraction.*

	Threshold Momentum		
n	π	K	P
1.006	1.27	4.50	8.55
1.008	1.10	3.89	7.40
1.01	0.98	3.48	6.62
1.02	0.69	2.46	4.7
1.03	0.57	2.00	3.81

One approach to reaching these indices is to use pressurized gas counters; the other is to use a "blown" SiO_2 material called aerogel. New developments now allow aerogels to be manufactured with indices ranging between 1.0008 and 1.126 at 637.8 nm. The transmittance is thought to be high but Rayleigh scattering leads to a very short bulk scattering length and consequent randomization of the Čerenkov angle. Thus, these radiators are not useful for ring imaging devices. However, if light collection can be made efficient, they might be very useful in a threshold counter optimized for the B Factory. The design of such a device is discussed in more detail below.

4. The CRID—A Reference Design

As discussed in the preceding section, ring imaging Čerenkov technology promises superior particle identification performance.[14] Two different approaches to this technology are now under study—the long drift device (CRID or RICH)[15,18] and the Fast RICH.[19] Though they use similar principles to produce and focus the Čerenkov image, they differ in their approach to the problem of readout. Several long drift devices have now been constructed and the body of operation experience is becoming quite substantial. The Fast RICH technology is now in the prototype stage, and there is an active R&D program in place to investigate the problems associated with constructing a full scale system.

In this section, we describe a candidate ring imaging device based on the long drift, Čerenkov Ring Imaging Detector

(CRID) now being built for the SLD.[15] It is representative of the kind of particle separation performance that can be attained with a ring imaging counter, and thus can be considered as a reference study for imaging counters. Large imaging counters have operated in experiments for several years, and even though no large scale 4π acceptance devices are now in full operation, two rather similar ring imaging Čerenkov detectors are now commissioning; (the CRID for SLD and the RICH for DELPHI[18]) their performance in prototypes is well-established;[20,21] and within about a year, their operation in the colliding beam environment should be understood; so that it will soon be clear if the promise will be met in practice.

The concern is that the long drift technology may not translate easily into the high rate environment of the B Factory. This issue will be partially addressed in the following sections, but not all facets are fully solved, and further R&D is required in some cases to prove that suggested solutions are really viable, or perhaps to develop a more robust approach. A CRID is a rather complex device; it contains a significant amount of material which unavoidably affects the performance of the electromagnetic calorimetry. A few suggestions for improvement are offered, but large improvements will require major changes to the technology, such as the Fast RICH.

A schematic design for a CRID, a discussion of its performance and mass, and a brief discussion of costs are given below. For this exercise, we will make use of existing techniques developed for the SLD to the maximum extent possible. Changes in some of these features that might be desirable in a real device will be discussed as appropriate.

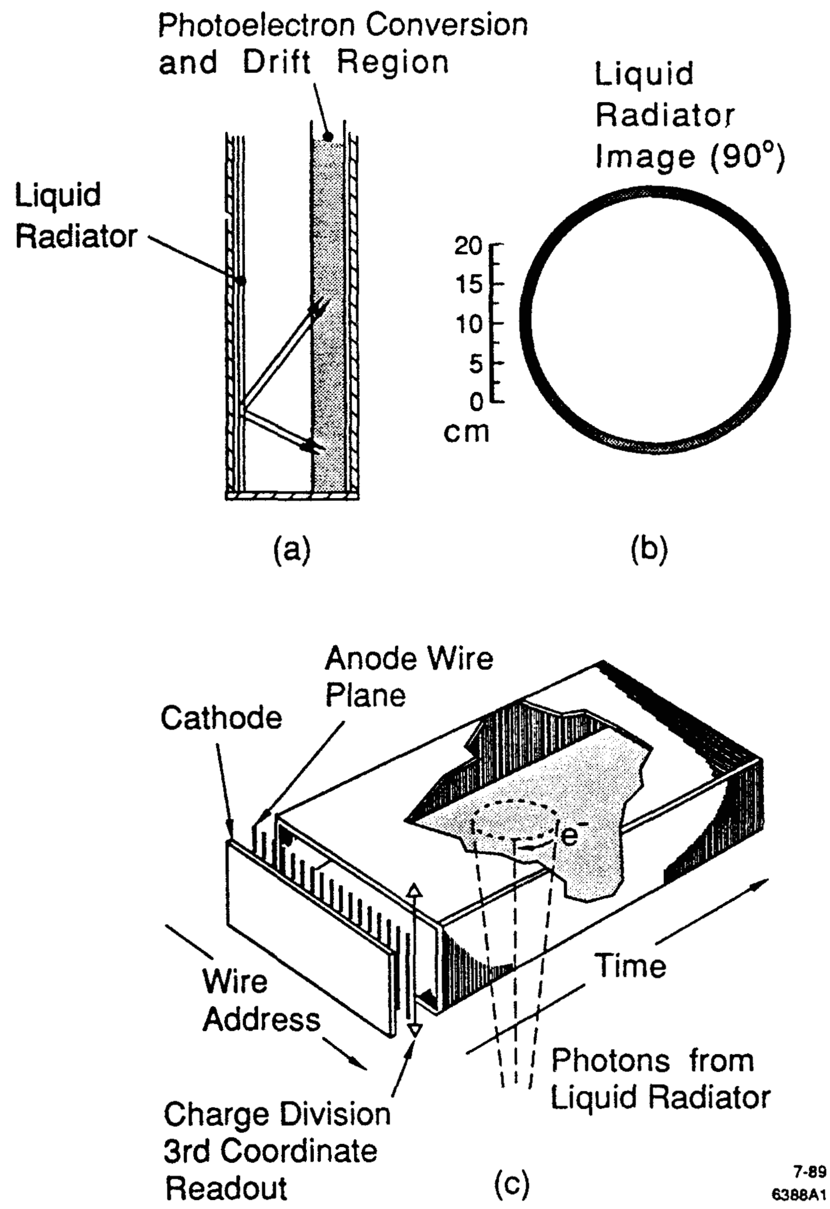


Figure 17. **a.** *The CRID principle for a liquid radiator device;* **b.** *size and thickness of a ring image produced at 90° in a liquid radiator;* **c.** *photon detector schematic.*

Principle of CRID Operation

As shown in Figure 17, when a charged particle passes perpendicularly though a thin liquid radiator, it produces a cone of light whose angle depends on the particle velocity and the index of refraction of the liquid. If this Čerenkov light is allowed to propagate some distance before hitting the photon detector, it will form an image which is said to be "proximity focused" on the detector. For our geometry, using liquid C_6F_{14} with an index of refraction $n = 1.277$, the proximity focused circle for a relativistic ($\beta = 1$) particle is around 17 cm radius. There are typically around 25 photoelectrons produced for tracks which cross the liquid radiator near the normal, but some photons are lost by total internal reflection for angles exceeding $\sim 13°$ so that, more typically, there are around 14 photoelectrons observed. The Čerenkov photons from the charged particle of interest pass through quartz windows in the liquid radiator cell and the detector box, and are converted by photo-ionization of gaseous TMAE (Tetrakis Dimethyl Amino Ethylene), which has a very high quantum efficiency in the range 170 to 220 nm. The photoelectrons drift at constant velocity in a uniform electric field along the box until they arrive at a proportional wire detector. The position of the photoelectron conversion must be measured in three-space. Two coordinates come from the drift time of the electrons and the wire address of the hit wire, while charge division on the proportional wires determines the conversion depth in the box.

Geometry

Conceptually, a CRID would sit just outside the central tracker, in front of a TOF (if any) and the calorimeter as indicated in Figure 18a. The barrel CRID can start at any radius (or at any z distance for an end cap device) determined only by the outer dimensions of the tracker. However, the thickness of the device is essentially "fixed" at around 20 cm. That is, the performance of a thinner device will be compromised. An end cap can be re-entrant to make an approximately hermetic device in polar angle. However, cracks between the liquid radiator boxes will lead to approximately 1–3% acceptance loss in azimuth.

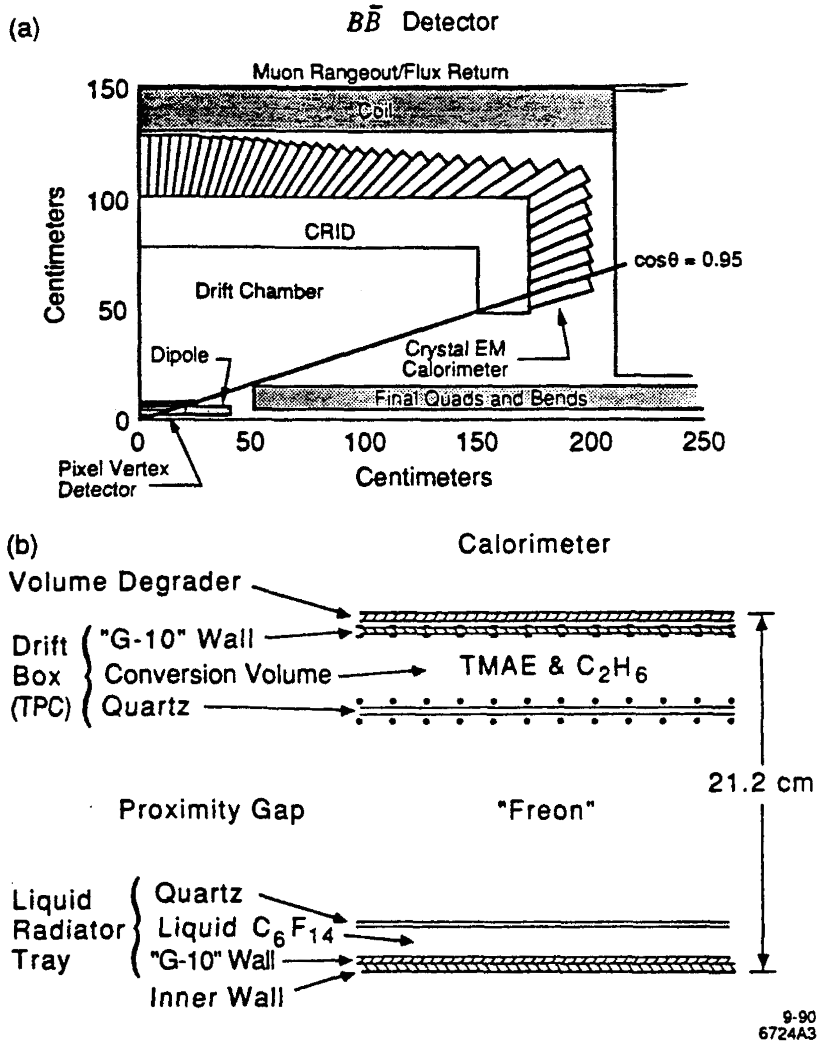


Figure 18. *Conceptual geometry for a CRID:*
a. *location in the detector;*
b. *schematic cross section.*

A cross section of the CRID is shown in Figure 18b. Starting from the inside, a particle first passes through an inner wall which in this design is a 7mm HEXCELL cylinder with aluminum skins. It then passes through the 4 mm thick G-10 back of the liquid radiator tray, through 1 cm of C_6F_{14}, and out through the 4 mm thick quartz window of the liquid radiator tray into a 13 cm drift region filled with a UV transmitting gas with good HV properties (such as a light fluoro-carbon). It then passes through a 3 mm quartz window into the 4.4 cm photon conversion region of the drift box (containing C_2H_6 plus TMAE), through the 3 mm back wall of the drift box, and finally through about 7 mm of epoxy equivalent in a HV degrader of the ALEPH/DELPHI type.[18,22]

Table 3 summarizes the properties and parameters of the proposed B Factory Čerenkov Ring Imaging Device.

SUMMARY OF CRID PARAMETERS	
1. Solid Angle Coverage (reentrant design)	92%
2. Angular Coverage (Barrel)	$37^\circ - 90^\circ$
3. Angular Coverage (End Cap)	$18^\circ - 37^\circ$
4. Radiator Material	C_6F_{14} (1 cm)
5. Index of Refraction (at 6.5 eV)	1.277
6. Window Material	Quartz
7. Photocathode Material	TMAE Gas at 20°C
8. Focusing Method	Proximity
9. Proximity Gap	13 cm
10. Čerenkov Angle (for $\beta = 1$)	672 mr
11. Radius of Čerenkov ring (for $\beta = 1$)	17 cm
12. Čerenkov Quality Factor N_0	65
13. Number of Photoelectrons (for $\beta = 1$)	25 (at 90°)
14. Momentum Threshold (for five Photoelectrons at 90°)	
e (cut off by p_t)	~1 MeV/c
μ	0.17 GeV/c
π	0.23 GeV/c
K	0.8 GeV/c
p	1.5 GeV/c
15. Particle Separation Range at 64° (3σ Level)	
e/π	0.2 to 1.200 GeV/c
μ/π	0.2 to 0.8 GeV/c
π/K	0.23 to 4.1 GeV/c
K/p	0.8 to 6.9 GeV/c

Table 3. *Summary of CRID parameters.*

Performance

The particle separation capabilities of the CRID depend primarily on the precision with which the Čerenkov angle is measured, which in turn, depends on the angular precision of each photoelectron measurement and the number of photoelectrons detected. A detailed discussion of the various sources of measurement error and photoelectron production rates is given elsewhere.[15,23]

In general, the number of photons is largest for particles that hit the radiator at normal incidence. For angles of incidence greater than $\approx 13°$, some photons get cut off by total internal reflection in the radiator. This photon loss is partially compensated by the increasing path length the particle travels in the radiator, so that the typical number is $\approx 60\%$ of the maximum. The number of photoelectrons $N_{p.e.}$ observed in the detector for a particle traveling at normal incidence is

$$N_{p.e.} = N_0 L \sin^2 \theta_c$$

where N_0 is the Čerenkov quality factor (about 65 cm^{-1}) for a C_6F_{14} radiator and a TMAE photocathode, θ_c is the Čerenkov angle, and L in the length the particle travels in the radiator. For the configuration under discussion, $N_{p.e.} = 25$ at normal incidence.

The important contributions to the total error are; (a) geometrical error, which is dominated by the relative depth of the radiator and the proximity drift space; (b) detector measurement error, which includes diffusion, measurement granularity, and nonuniformities in electron drift in the detector; (c) multiple scattering error in projecting the track measured in the central tracking chamber out to the CRID; (d) momentum smearing error due to the bending of the particle in the radiator; and (e) chromatic aberration error in the radiating liquid.

The contributions from various sources as a function of momentum at a track angle of 64° are shown in the Table 4.

		Angular Resolution per photoelectron (mrad)						
P	$\theta_C(\pi)$ (mrad)	N_{pe}	Geometry	Detector Measurement	Multiple Scattering	Momentum Smearing	Chromatic Aberration	Total Total
0.5	622.2	12.5	7.3	5.7	2.9	0.7	5.3	3.2
1.0	660.0	13.8	7.3	5.7	1.4	0.3	5.0	2.9
1.5	666.4	14.0	7.3	5.7	0.9	0.2	4.9	2.9
2.0	668.8	14.1	7.3	5.7	0.7	0.2	4.9	2.8
3.0	670.5	14.2	7.3	5.7	0.5	0.1	4.9	2.8
4.0	671.2	14.2	7.3	5.7	0.4	0.1	4.9	2.8
5.0	671.4	14.2	7.3	5.7	0.3	0.1	4.9	2.8

Table 4. *Angular resolution (mrad) per photoelectron as a function of momentum at 64°.*

The sizes of each particular contribution depend on track angle and momentum, but typically the largest single contribution ($\sim$ 7.3 mr per photoelectron at 64°) comes from the geometrical term. Typically, the combined contribution from all other terms is about the same size, so that the error per photon is $\sim$ 11 mr. The size of the geometrical term can only be reduced either by making the radiator thinner, which loses photoelectrons. or by making the drift space longer, which makes the overall counter thicker.

The particle separation capabilities of the CRID are shown in Figure 19. The Monte Carlo particles for results shown here entered the radiator at a polar angle of 64° relative to the beam line, so that these plots include the typical loss of photons due to total internal reflection that occurs when the particles enter the radiator at oblique angles. The vertical scales give the separation between different kinds of particles in "standard deviations (σ)." It is "saturated" at 10σ, since the "Gaussian" separation probabilities estimated here will surely be unreliable for very large estimated separations. The hadronic separations are generally excellent, for π/K, for example, exceeding 10σ from near threshold up to $\approx$ 2.3 GeV/c, and falling to 3σ around 4 GeV/c. The CRID also provides excellent low momentum lepton identification which complements the calorimetry and muon tracking chambers very well. At the 3σ level, e/π separation is achieved up to 1.2 GeV/c, and μ/π separation up to 0.8 GeV/c.

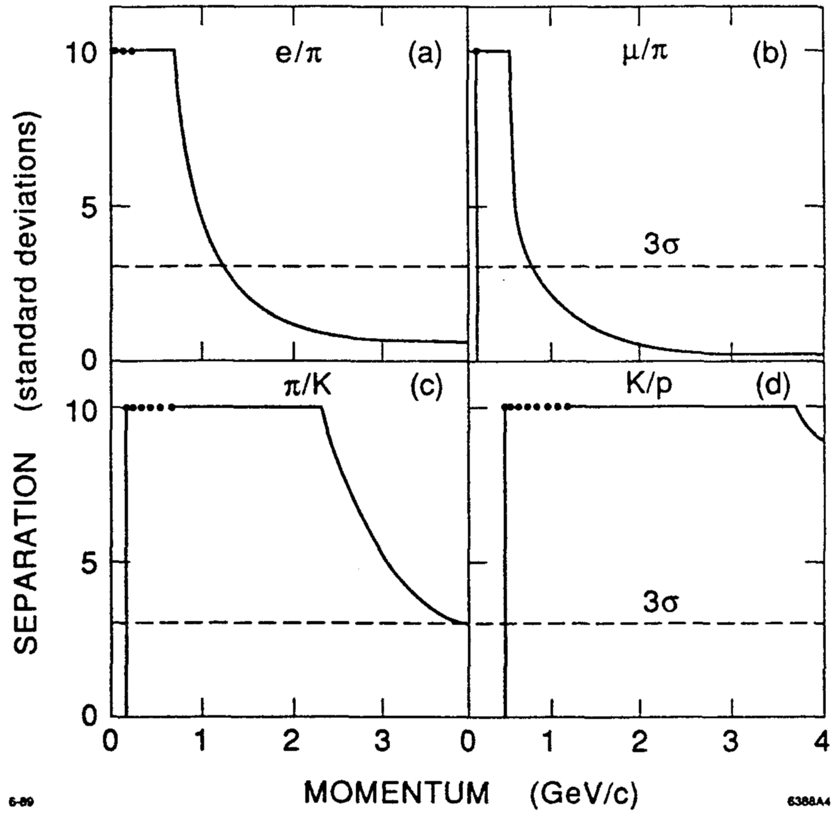


Figure 19. *Particle separation capability of the liquid CRID system at 64°. The curves are saturated at 10σ (see text). In the dotted regions of the curves, only the lighter particle emits light in the liquid. There are thresholds shown at low momentum below which the lighter particle emits too few photons in either quartz or liquid to be reliably separated (see text). The separation is shown for;* **a.** *e/π;* **b.** *μ/π;* **c.** *π/K; and* **d.** *K/p.*

There are regions at the lowest momenta where both particles are below threshold. This is a significant effect only for the K/p case. K threshold in the liquid is at 620 MeV/c. The K/p separation extends down to $\approx$ 480 MeV/c using radiation from the quartz. For lower momentum, the number of photons produced by the kaon becomes too small for reliable separation, and, in any case, below 430 MeV/c neither particle emits any photons. However, in cases where particles are too slow to be separated by the Čerenkov it is very simple to obtain dE/dx separation since one or both particles are far into the $1/\beta^2$ region where even very modest amplitude information on energy loss will provide excellent separation. There is also a low momentum cutoff due to particle curvature in the magnetic field, which depends on the CRID radius, the magnetic field, and the polar angle of the particle. For example, a device at 80 cm radius in a 10 KG field is unable to identify a particle traveling normal to the magnetic field below $\sim$ 170 MeV/c. Again, this "hole" in the CRID coverage can be filled with very modest dE/dx information from the tracking chamber in the $1/\beta^2$ region.

SLD CRID Technology

As was shown above, the performance of a CRID using SLD technology with a liquid radiator matches the particle ID requirements for $\pi/K/p$ extremely well and also gives significant help to the calorimeters for low momentum e and μ identification. The device, as described above, is essentially a liquid only version of the SLD CRID except

1. there is an outer volume degrader, which is necessary to keep the space requirements down if a long drift configuration is chosen for the detector;
2. this CRID can be run at room temperature, which makes the gas and fluid systems much simpler with little effect on performance for a liquid radiator only system;
3. a fast readout system is necessary to handle the 1 KHz trigger rates (see below).

The CRID contains approximately 20% of a radiation length of material distributed as shown in Table 5.

RADIATION LENGTH OF THE CRID COMPONENTS	
INNER WALL	0.018 (L/L_{RAD})
LIQUID RADIATOR TRAY	0.105 (L/L_{RAD})
PROXIMITY GAP	0.000 (L/L_{RAD})
DRIFT BOX	0.044 (L/L_{RAD})
OUTER WALL	0.036 (L/L_{RAD})
TOTAL	0.203 (L/L_{RAD})

Table 5. *Radiation length of the SLD CRID components.*

Significant reductions in this mass inventory are difficult so that substantial effects on precision electromagnetic calorimetry seem unavoidable in any detector which contains a CRID of the SLD type. Moreover, the device has a "bad geometry" in that much (about 50%) of the mass is located a long way from the calorimeter. A few possibilities for modifications to the technology which could both reduce the total mass in the device and improve its distribution are discussed below. However, large improvements (i.e., $\sim 50\%$ or more) seem unlikely in this technology.

A CRID is a rather expensive device to construct. In general, the component costs scale approximately with the area covered, whereas the electronic costs scale linearly with the radius. For a given detector size, costs can be scaled from actual SLD costs. If we assume a B Factory CRID just outside a central tracker of 80 cm radius by 250 cm length, with an end cap CRID starting at a 30 cm radius, the estimated costs are as given in Table 6.

CRID CONSTRUCTION COSTS			
BARREL			
	VESSEL	448	
	LIQUID RADIATORS	212	
	DRIFT BOXES	566	
	e^- DETECTORS	191	
	SPARES	97	
	E.D.&.I.	500	
BARREL TOTAL			2014
END CAP			
	VESSEL	323	
	LIQUID RADIATORS	52	
	DRIFT BOXES	138	
	e^- DETECTORS	153	
	SPARES	34	
	E.D.&.I.	360	
END CAP TOTAL			1060
FLUID SYSTEMS			560
ELECTRONICS			1440
CRID TOTAL			5074

Table 6. *Estimated cost of a B Factory CRID.*

Electronics

The SLD CRID electronics are matched to the properties of the SLC environment: for example, the trigger rate is slow (typically 1 Hz); there are ≈ 5 ms between machine pulses; and the particles are highly collimated into jets, so that high hit densities and pulse pair resolution are important issues. At the B Factory, on the other hand, the trigger rate is fast (perhaps 1 KHz); there are ≈ 5 ns between machine pulses; the typical particle densities are relatively low; and there are only large radius liquid rings so that the number of real hits per box will usually be small. The total number of hits should be dominated by background.

These differences in machine environment lead to several modifications in the electronics. Conceptually, the front-end (preamp) electronics can be just like the SLD CRID (with a different shaping time) but a different readout is required. Full wave form digitization (as in the SLD) is possible, but would be expensive at these data rates, and it is probably unnecessary. Simple considerations of hit densities expected at the B Factory suggest that four hits per wire should be sufficient (see below). At these hit densities, it is probably much cheaper to readout each wire end by a combination TDC/ADC.

A four hit version is indicated schematically in Figure 20a. The TDC portion is quite straightforward. There are a variety of ways one might implement the ADC. For example, it could use a simple modification of the SLD CDU[15] module as shown in Figures 20b and 20c. This is a simple analog storage device with parallel-in, serial-out architecture which has 4 cells per channel. The exact level of multiplexing into the ADC depends on the readout time desired and on the ADC digitizing speed. In particular, with the 32 preamp multiplexing shown in Figure 20c, an ADC with 1.5 μs digitizing speed leads to 20% deadtime at a projected 1 KHz trigger rate.

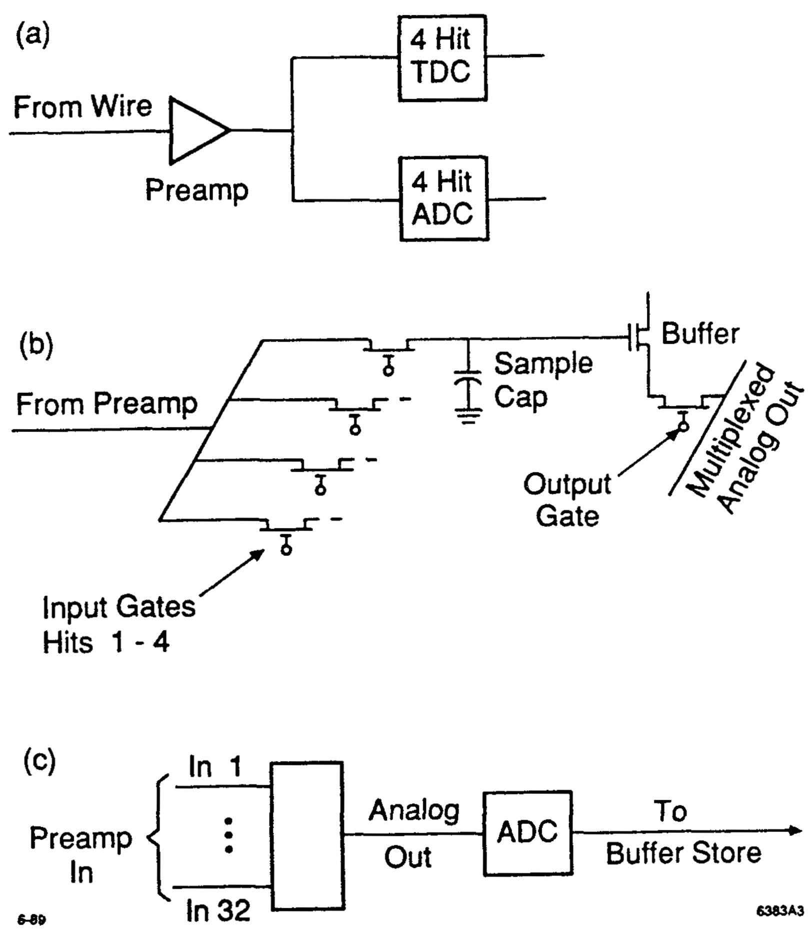


Figure 20. *Conceptual design for CRID readout:* **a.** *four-hit electronics for each wire end;* **b.** *parallel-in, serial-out analog memory unit (1 of 32 channels);* **c.** *preamp multiplexing scheme.*

Problems of the CRID Technology

There are two major difficulties associated with the CRID technology which we discuss more fully below. First, since the CRID is massive and sits between the interaction region and the calorimeter, it will affect the low energy performance of the calorimeter. These effects are discussed more fully in chapter on calorimetry. We will briefly describe below a few ways to reduce this problem.

Secondly, photoelectrons in the CRID drift a long distance (taking ≈ 25 μs). Backgrounds from beam individual crossings add during the total drift time, leading to uncorrelated backgrounds. This, in turn, leads to questions about the sensitivity of the chambers to average current, and to possible difficulties in identifying tracks.

None of these problems seem to be fatal to the CRID technology described, but they do have uncomfortable implications for triggering and operations.

CRID Mass and Its Distribution.

A CRID of the SLD type affects the low energy performance of any calorimeter which follows it because (1) the radiation length of the device is rather large (about 20%), and (2) the mass is poorly distributed in that much of it is located a substantial distance in front of the calorimeter. Thus, not only does a photon have a substantial interaction probability ($\sim$ 17%), but the conversion products can also get lost, either by scattering or by bending in the magnetic field. The scattering loss goes like L (the distance from the conversion point to the calorimeter) while the loss due to bending[24] in the the magnetic field goes like L^2.For a conventional CRID in a 1 T field, the loss from the magnetic field becomes significant at around 200 MeV, while that coming from the scattering begins around 25 MeV. Thus, the low energy performance of the calorimeter can be improved by making the CRID thinner, and by either reducing the total amount of material or improving the overall material distribution by placing the maximum amount of material close to the calorimeter. It is helpful, in addition, to either make the counter a good charged particle tracker or to place a layer of tracking between the CRID and the calorimeter, so that the energy of converting photons can be corrected, or, at worst, events with converting photons can be cut from the data sample.

A number of suggestions for possible improvements were made during the B Factory workshop which deserve further study.[24,25,26,27] A sample of these are briefly described below. However, to date, all modifications which have been suggested trade-off improved low energy photon properties in the calorimeter for degraded particle identification performance.

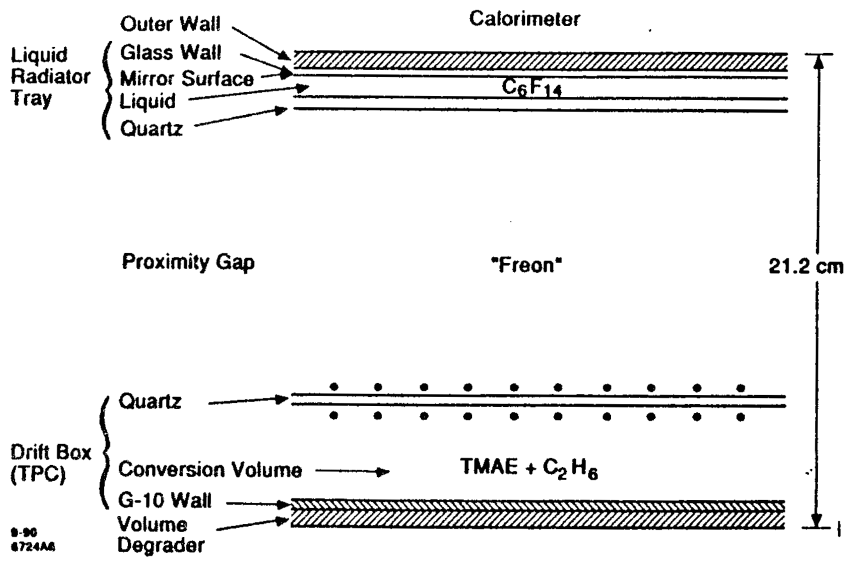


Figure 21. *cross section for a CRID with a reflected proximity gap.*

At first sight, the scheme shown in Figure 21 appears useful in improving the mass distribution. A mirrored surface in the liquid radiator tray reverses the direction of the Čerenkov photons so that the most massive element (the liquid tray) has been placed near the calorimeter. Unfortunately, in actuality this has almost no effect on the mass distribution, as can be seen from the table above. The radiator plus wall are about 12% L_{RAD} total whereas the TPC plus degrader are 9% L_{RAD}; there is little net gain from turning the system around. Moreover, it degrades the performance substantially compared to the normal configuration, as can be seen in Figure 22, because the number of Čerenkov photons is reduced by about 40% by absorption and reflection. The performance of the counter in the presence of background will degrade for the same reason.

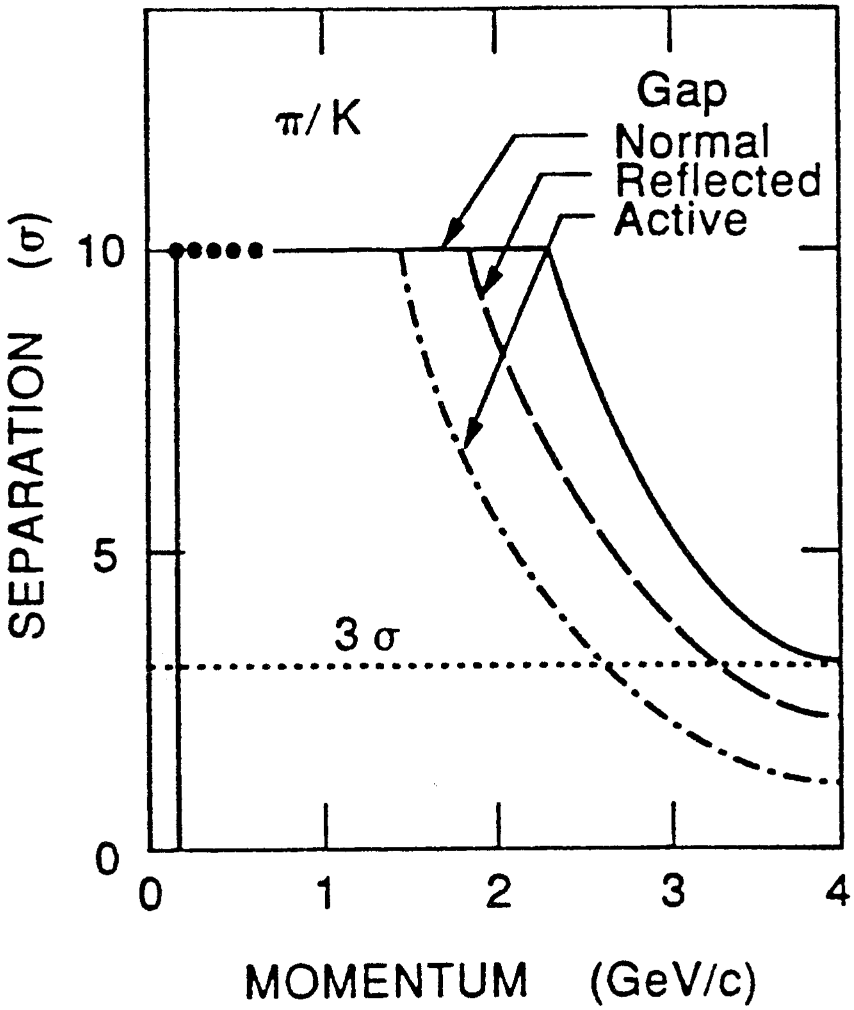


Figure 22. *Particle separation (π/K) capability of several different liquid CRID systems at 64°; solid line for the standard configuration (Fig. 18); dashed line for the reflected proximity gap counter (Fig. 21) and the "honeymoon" configuration (Fig. 23); dot-dash line for the active proximity gap configuration (Fig. 24).*

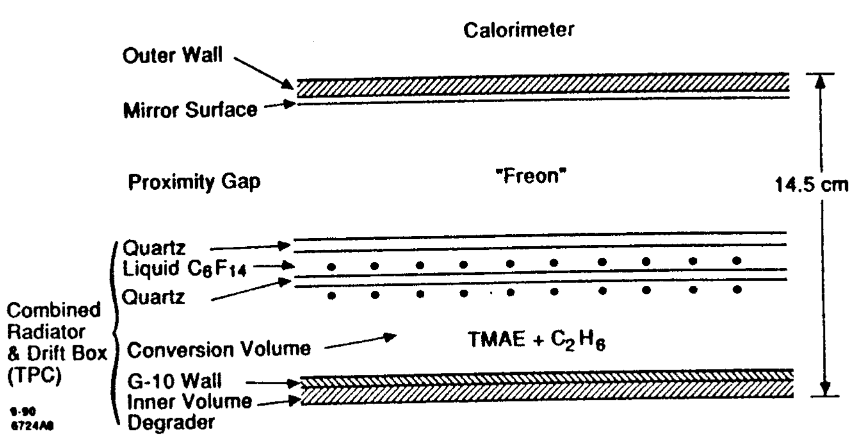


Figure 23. *Schematic cross section for a "honeymoon" CRID with a combined radiator tray and detector and a reflected proximity gap.*

A more amusing "honeymoon" configuration which marries the detector and radiator is shown in Figure 23. Its performance is essentially the same as the that of the device shown in Figure 21. The mass of one nontransmitting window can be saved, but this savings is compensated by the material needed for an independent reflecting surface. The main advantage is that the proximity gap gets "used" twice so that the overall counter becomes about 14 cm thick rather than about 21 cm. On the other hand, the distribution of mass is very bad.

As a final example, Figure 24 shows a configuration for an active proximity gap counter. The TMAE vapor pressure has been reduced to give an absorption length in the conversion gap of 20 cm, which gives near the optimum performance for this total counter thickness. One quartz window is eliminated, but the material savings are compensated for by the need for an additional volume degrader. The performance given in Fig. 22 is substantially worse than that of the normal configuration. The main advantage is that this counter is active over most of its volume so that even low energy conversion products can be reliably recognized. However, the sensitivity of the single electron detectors to background might keep this feature from full exploitation, if restrictive triggers are used for the CRID (see the following section).

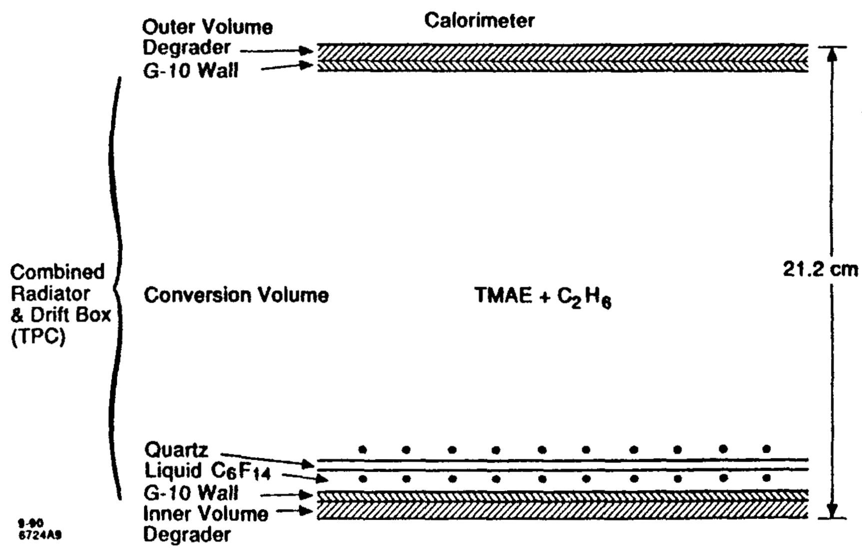


Figure 24. *Schematic cross section for an active proximity gap CRID with a combined radiator tray and detector and a long absorption length conversion volume.*

Detector Performance and Background Sensitivity

The high rate environment at a B Factory can affect the long drift TPC photon detector of the CRID in a number of ways. The primary concerns are that positive ions build-up in the drift volume will distort the field and the measured positions for the conversion Čerenkov photons; and that the single electron detectors, which in the CRID are PWCs with 7 μm carbon filament sense wires operating with ethane plus TMAE gas, will not have a lifetime commensurate with the expected time of experimental operation.

The distorting effects of positive ion build up in TPCs are well known, and generally are dealt with by a gate at the detector-drift region interface which collects the positive ions before they enter the drift region.[28] When gated closed, electrons from the drift volume are also prevented from entering the avalanche region of the chamber. Since the positive ions drift much more slowly than the electrons, the gate can be made to collect 100% of the positive ions (in principle) that are produced by avalanches of electrons which come from the TPC drift volume provided that the trigger rate is properly controlled. However, at the very high rates of a B Factory, the ions coming from electrons produced by background in the PWC itself could become troublesome, since they leak back into the TPC drift volume when the gate switches open. In this case, the avalanche gain also needs to be gated, as is discussed below.

The addition of TMAE gas to the usual PWC mixtures has been shown to reduce the lifetime of chambers by factors of at least 100 and more typically around 1000.[29] Thus, the lifetime expected for CRID PWCs in a B Factory is of great concern and was discussed extensively during the Workshop.[27,30,31] A number of approaches were attempted. For example, attempts were made to scale the TPC experience at PEP to the B Factory,[27,30] while other studies looked at backgrounds more directly.

More research in this area is needed. The tentative conclusion is that the CRID sensitivity can be made similar to or better than that of the CDC provided that the fast (~ 2 μs or better) trigger rate can be kept under 1000 HZ. With this gating ratio, the net charge collected on the sense wire begins to be dominated by avalanches coming from electrons which are initially produced on the PWC side of the gate. Thus, it is important to develop a means of lowering the gain during the time the TPC gate is off. Since the average multiplicity in events at the B Factory are rather small, another factor of 5 can be gained by gating only those TPCs in the vicinity of charged tracks. Finally, since the gain of the CRID TPCs can be quickly recovered by the wire heating technique[29] which burns off polymerized deposits on the sense wires, it would appear, in a sense, that the CRID is substantially more robust than the CDC. Of course, this does not guarantee the performance of either device.

In summary, the single issue of greatest concern in translating the SLD technology to the B Factory detector remains the single electron detection chamber lifetime. Studies to date indicate that it can be controlled if (1) care is given to PWC chamber gating; (2) the CRID gate-on trigger rate can be kept under 1000 Hz; and (3) the CRID detection PWC's are routinely rejuvenated by wire heating. However, a more robust solution would be to develop a single electron detector for the TPC which is less sensitive to ageing effects. Some potential candidates are now under development.[32,33]

Hypothesis Testing and Background Sensitivity

Both SLD and DELPHI studies have shown that the hypothesis testing method of Čerenkov analysis is quite insensitive to rather large numbers of background hits. The readout is essentially fully segmented into $\approx 10^6$ 2-D pixels so that the number of hits that can be well separated per event is very large. True signal hits are on average separated by many pixels, so that they are not confused even in the presence of large numbers of background hits. Moreover, analysis is attempted only after points are correlated with tracks found in the central chamber. No independent "pattern recognition" is attempted. Instead, points are associated with tracks; tested against all five possible hypotheses; and the best one chosen. Provided that the geometry and performance parameters of the device are well understood, and that the background can be simply modeled, the results are indeed very robust as is demonstrated briefly below. These results are discussed in more detail elsewhere.[33,34]

This study uses a Monte Carlo, which throws Čerenkov photons for the true particle and adds random background hits. A likelihood ratio test of the SLD/DELPHI type is then employed by performing, for each track, a test for each possible hypothesis. The likelihood ratios obtained can then be compared for the possible hypotheses. For simplicity, the study performed here only considers the π and K hypothesis, and log-likelihoods are used instead of likelihoods. The function calculated for each hypothesis has the form:

$$L_{HYP} = -N_{HYP} + \sum_{i=1}^{N} ln[N_{HYP} \times e^{-(\theta_i - \theta_{HYP})^2/2\sigma_i^2} + B(\theta_i)], \tag{1.10}$$

where N_{HYP} is the number of photons expected for the h ypothesis; θ_i is the measured Čerenkov angle for the i^{th} photon; θ_{HYP} is the expected Čerenkov angle for the hypothesis; σ_i is the error on the Čerenkov angle measured for the i^{th} photon; and $B(\theta_i)$ is a background function. For simplicity, the background as taken to be flat and was normalized so that a value of 1 gives 1.0 hit in a $\pm 3\sigma$ road about the true hypothesis. By way of comparison, if the background is dominated by interacting soft photons, then 1 hit is approximately equivalent to 10% occupancy in the CDC.[33,34]

The differences between the log-likelihoods for 2.5 GeV/c particles are shown in Figure 25 for a variety of different background conditions. At low values of the background, there is no overlap between the true π and true K distributions, whereas as the background gets very large the overlap becomes much more significant. A clean sample can then only be obtained by reducing the efficiency for the wanted particle. However, the highest background level given (B=40) is very large indeed and much more than 10 times the tolerable background in the CDC. The efficiency for K identification (labeling a true K as a K) is shown in Figure 26 as a function of momentum and background when the maximum π punch-through probability is limited to 1%, whereas Figure 27 shows the relationship between efficiency and background at a single momentum (3.0 GeV/c). The background is given in equivalent CDC occupancy using the scaling noted above. The identification is affected very little by background well beyond the maximum tolerable levels acceptable by the CDC.

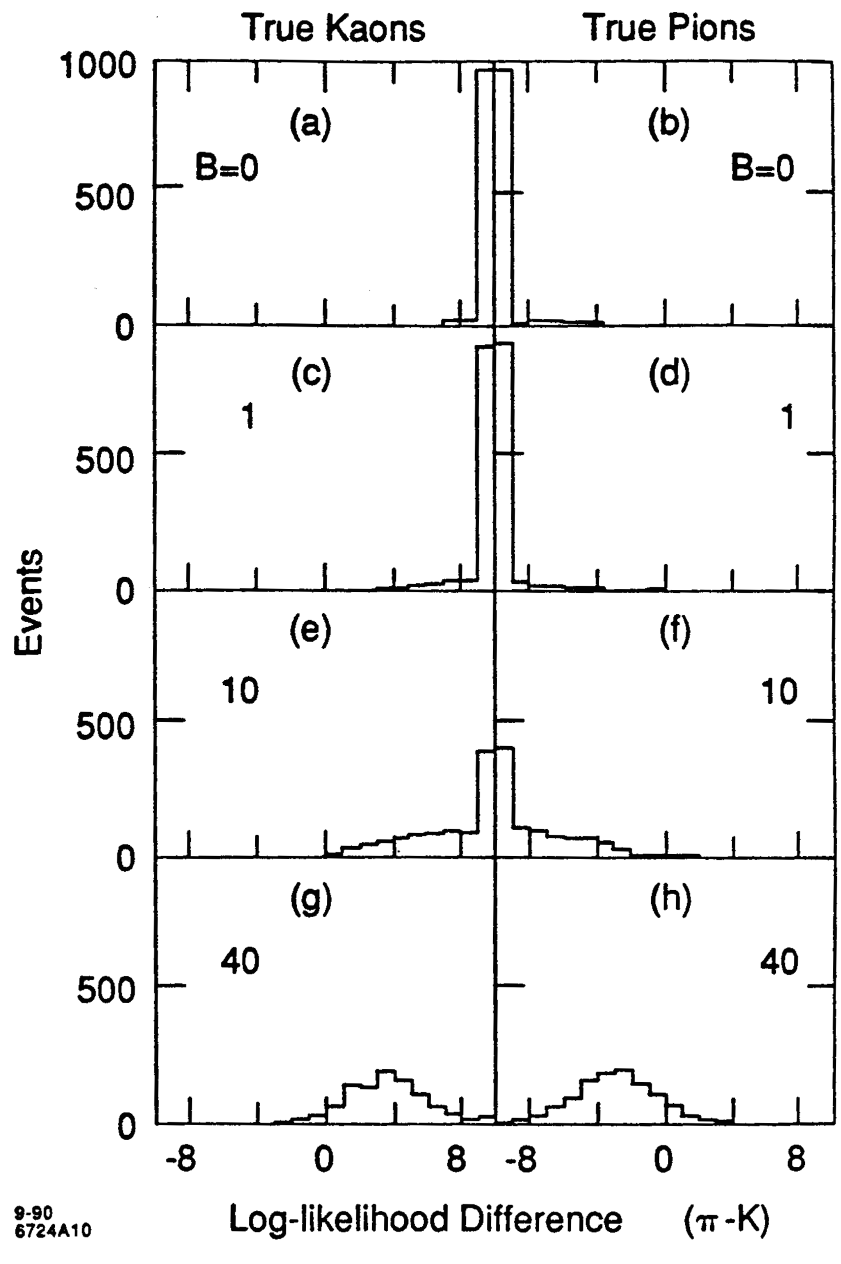


Figure 25. *Log-likelihood differences ($L_\pi - L_K$) for 1000 2.5 GeV/c particles; the end bins include the overflow;* **a.** *True kaons, B=0.0;* **b.** *True pions, B=0.0;* **c.** *True kaons, B=1.0;* **d.** *True pions, B=1.0;* **e.** *True kaons, B=10.0;* **f.** *True pions, B=10.0;* **g.** *True kaons, B=40.0;* **h.** *True pions, B=40.0.*

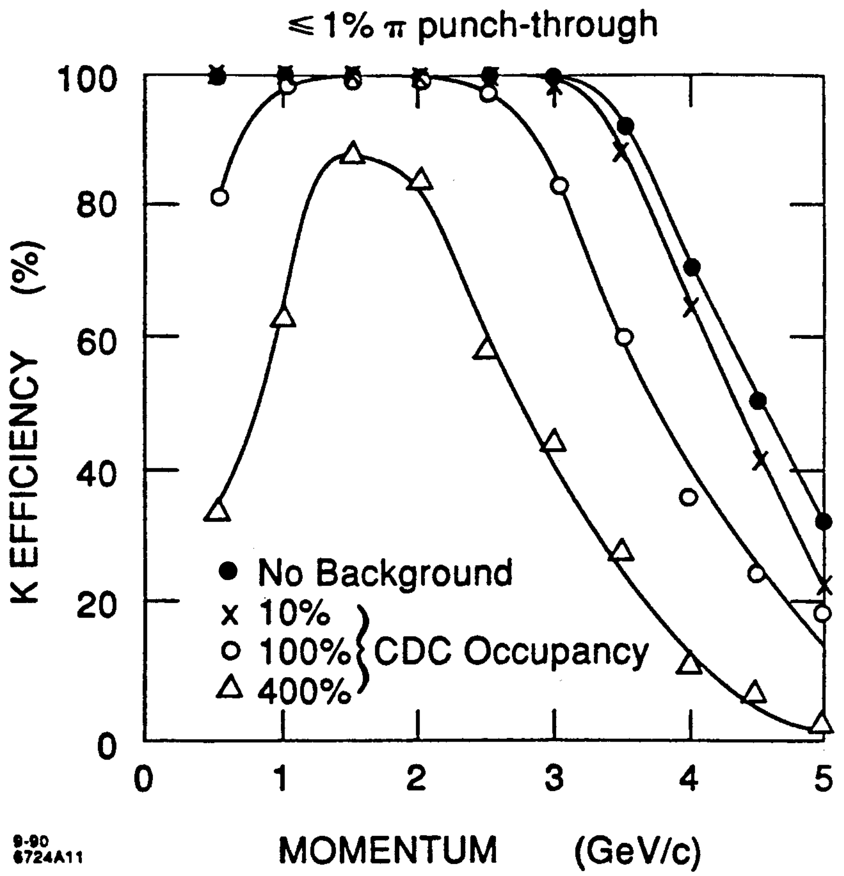


Figure 26. *Efficiency for labeling a true K as a K as a function of momentum when the maximum π punch-through is limited to 1%.*

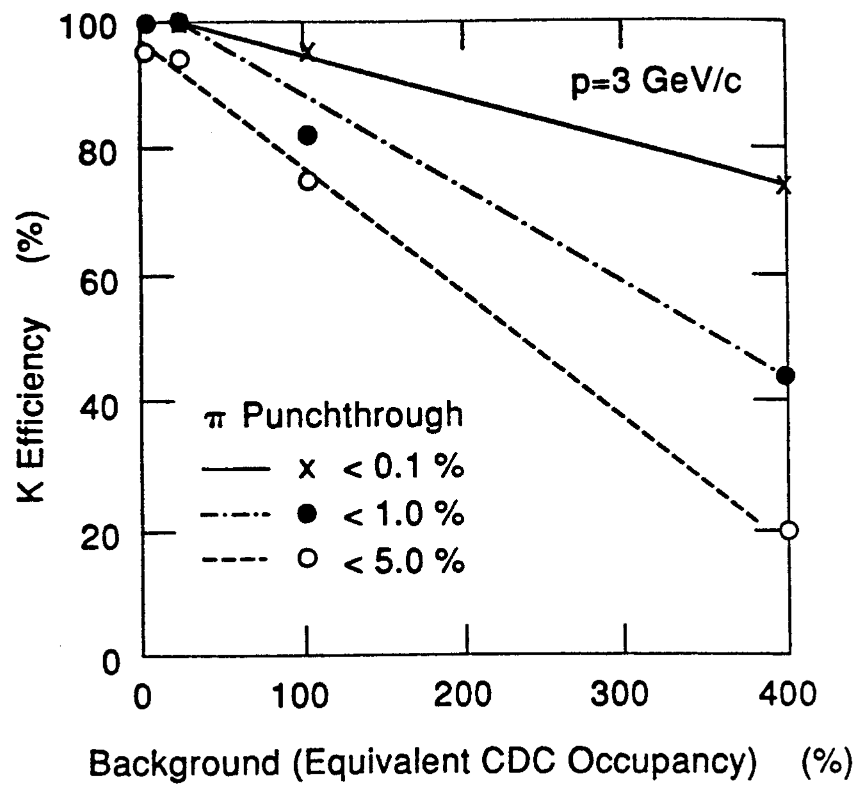


Figure 27. *Efficeincy for labeling a true K as a K as a function of background and maximum π punch-through probabilities.*

4.1. *CRID Summary*

The CRID technology described above meets the performance requirements for particle ID at a B Factory very well. In addition to providing superior π/K/p separation it also fills in the lepton ID holes which are left by the calorimetry and muon tracking techniques. However, it adds $\sim 20\%$ of a radiation length of material in front of the calorimeter and takes about 20 cm of space. The device using SLD technology is rather insensitive to backgrounds at the pattern recognition level. The integrated charge sensitivity appears to be tolerable but only with a rather restrictive fast trigger, and excellent gating efficiencies. New single electron detectors are under study which could make the hardware less sensitive to backgrounds.

5. A Design for Particle ID Using a Fast RICH

One very attractive potential solution to the problems posed by the use of a TPC in an imaging Čerenkov is a new type of imaging Čerenkov which is called the Fast RICH. It was proposed some time ago in connection with the design of detectors for the symmetric B Factory at the PSI and for LEAR.[19,35]

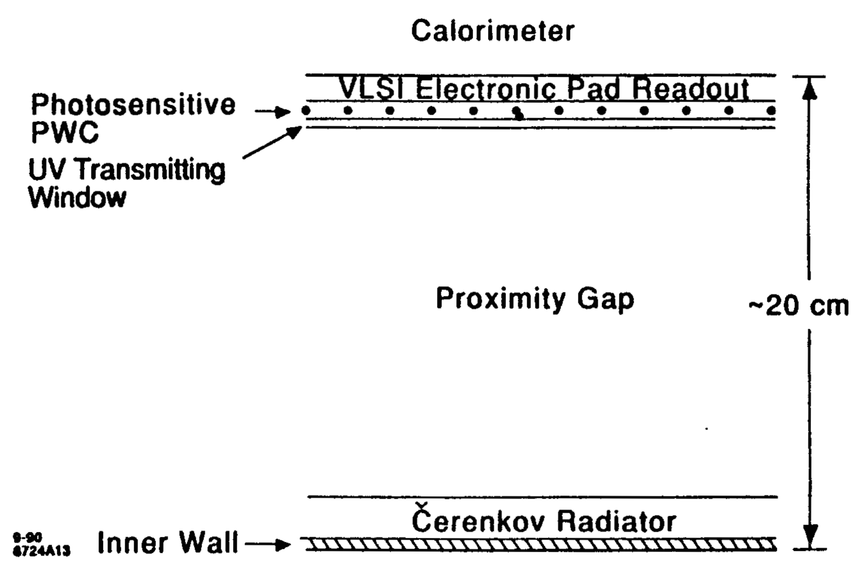


Figure 28. *Schematic cross section for a Fast RICH.*

A generic schematic of this device is shown in Figure 28. The essential difference between this device and the CRID is the replacement of the three-dimensional readout of the CRID TPC by a two-dimensional readout via a pad system. The photoelectron drift is kept very short (in the radial direction in a solenoidal geometry). The necessity for reading out three dimensions is eliminated by using a photocathode with a short absorption length. Since the photoelectron which is produced by the Čerenkov photon has a short drift length, it also has a short drift time; the device is fast; and it has been proposed for use at high luminosity hadron colliders as well as B Factories.[36]

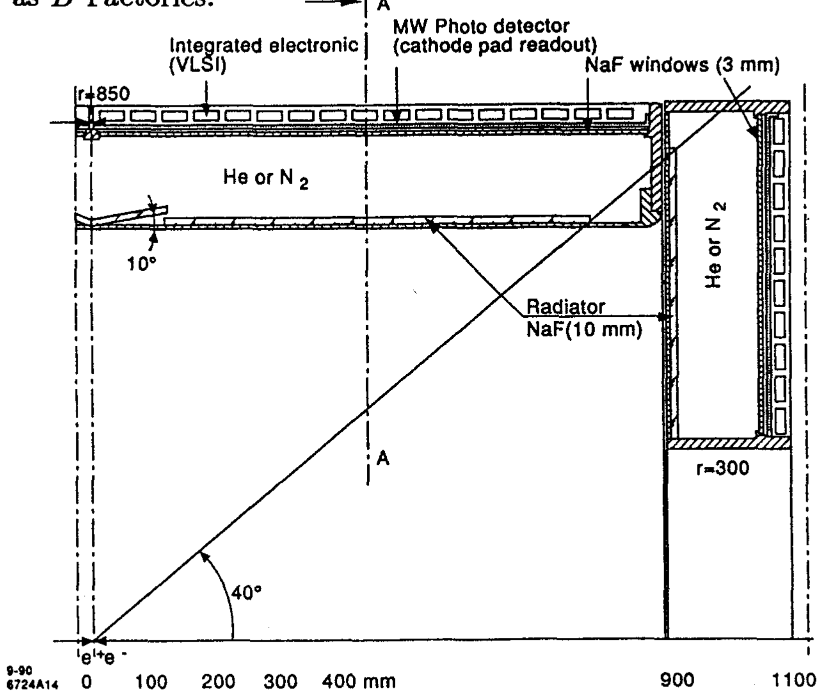


Figure 29. *Schematic cross section for the BMF (PSI) Fast RICH.*

A schematic of the device proposed for the BMF detector at PSI is shown in Figure 29. It includes a number of novel features in addition to the fast photoelectron readout which were necessary implementation details at the time it was proposed. In particular, the photocathode chosen is TEA, which has the short (0.5 mm) absorption length required at room temperature. The radiator proposed is NaF, since it transmits into the far UV range required by the TEA photo-efficiency spectrum. Because NaF has a high index of refraction in the UV region where TEA is photosensitive, Čerenkov photons emitted by fast particles traveling at nearly normal incidence are trapped in the radiator by total internal reflection. Therefore, a section of radiator in the 90° region is tilted by 10°. The device contains approximately 20% L_{RAD} of material for particles at normal incidence.

A number of issues of design and construction must be solved before such a device can be built. A major prototyping effort is now well underway. In particular, the VLSI electronics chips for reading out the pads have now been made and successfully tested; tests of TEA quantum efficiency have been made; and a full scale test of a B Factory prototype in a beam to confirm operations, electronics performance, resolution, and speed will be performed in the fall of 1990.[16]

This device originally was thought to provide good π/K separation up to 3 GeV/c.[16,19,35] Recent measurements of the refractive index of NaF crystals demonstrate that NaF is more dispersive than was predicted from the 1-pole fit extrapolation used previously, and the predicted 3σ cutoff of π/K separation is now limited to 2.2 GeV/c.[16] This is confirmed by a simulation of the performance of a NaF device, using a different detector model, by Coyle[33,37] Though such performance is marginally adequate for a symmetric collider, such as PSI, it is clearly insufficient for the asymmetric machine described in this report.

Recently, a study of reflective UV photocathodes with gas phase electron extraction has been underway as part of the LAA project at CERN.[16,17]

A CsI photocathode with an adsorbed TMAE film has been observed to have an efficiency equal to or exceeding the TMAE gas-phase region of interest as shown in Figure 30. It should be well adapted to Fast RICH counters because of its isochronous signals and room temperature operation. Further studies of photocathode lifetime, feedback, and operation in the presence of positive ion build up are necessary, but results to date are very impressive.

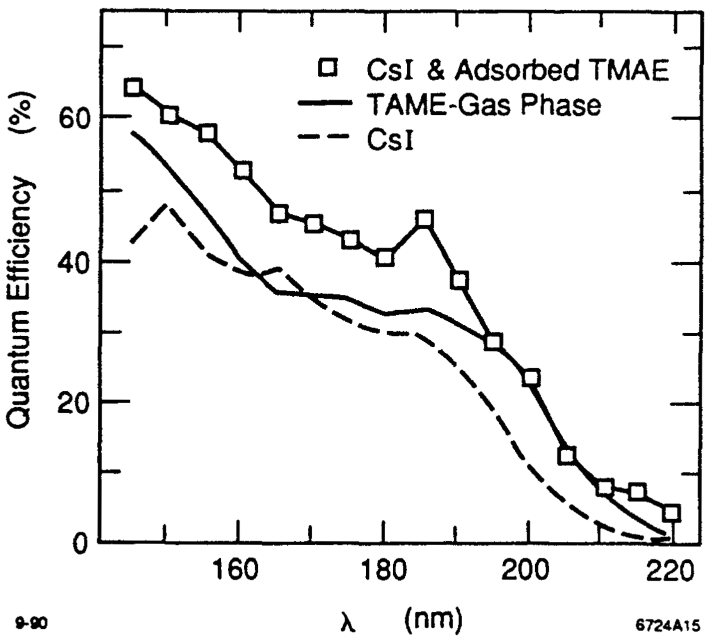


Figure 30. *The measured quantum efficiency of a 500 nm thick CsI photocathode as a function of photon wavelength λ; the solid line is the TMAE gas-phase efficiency; the dashed line is the CsI efficiency; and the open squares give the efficiency for a CsI photocathode (30° C) on which an absorbed film of TMAE has been deposited.*

The existence of this new photocathode leads to a proposal for a new design for a Fast RICH using a liquid C_5F_{12} radiator and a reflective CsI+TMAE photocathode as shown schematically in Figure 31.[16] The CsI + TMAE cathode operates in the 6–7 eV photon region where fluoro-carbon liquids and quartz windows transmit well, so that expensive, dispersive solids like NaF are not needed for radiators or windows. The VLSI readout chips developed for the BMF NaF+TEA Fast RICH discussed above can be used for pad readout. Hexane gas is added to the methane chamber atmosphere to eliminate the UV feedback photons above 7.0 eV.

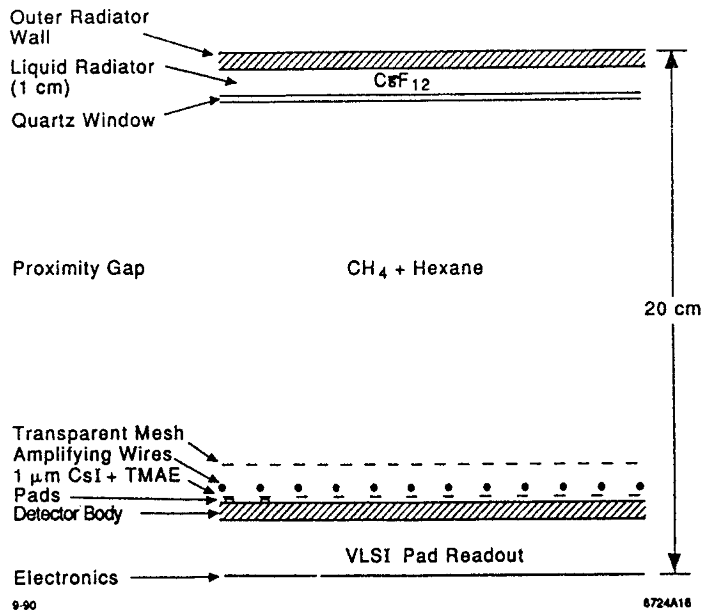


Figure 31. *Schematic cross section for a liquid radiator Fast RICH with a reflected proximity gap and a CsI plus TMAE cathode.*

The C_5F_{12} liquid is slightly superior to the C_6F_{14} used in the CRID, and the π/K identification range for this device extends well above 4.0 GeV/c (to 5.5 GeV/c)[16] as required for an asymmetric B Factory. However, the device must be run at room temperature (or slightly below) to stay below the 28°C boiling temperature of the liquid. A C_6F_{14} liquid radiator of the type discussed in the preceding section would also give adequate performance. The amount of material in this device also appears to be less than required for the CRID, and a reflected proximity gap appears to improve the material distribution. The estimate for particles at normal incidence is about $12\% L_{RAD}$ in total, and the contribution from the material farthest from the calorimeter, *i.e.*, the detector and readout chips, is about $2.5\% L_{RAD}$.

In summary, the Fast liquid RICH is a very promising idea which has the performance specifications required at the asymmetric B Factory without the background and lifetime issues that raise concerns for the CRID discussed in the preceding section. Additional R&D is now under way,[16] which should soon indicate if the promise of this approach can be met in practice.

6. Aerogel Threshold Čerenkov Counters

6.1. Introduction

Although excellent K/π separation at high momentum is the most important requirement of a particle identification system at a B Factory, such a system has to satisfy additional criteria. First, it has to operate in a magnetic field, which limits the choices of read-out systems. Second, in order to design a highly efficient and uniform detector, it is necessary to cover almost the entire solid angle hermetically. Third, the entire device should be built with low-density materials amounting to a few percent of a radiation length (r.l.), in order to maintain the good energy resolution of the electromagnetic calorimeter for all photons. Finally, the costs to built such a detector have to be acceptable.

A particle identification system at an asymmetric B Factory must provide π/K separation to high momentum.

The Čerenkov effect-based particle id systems under discussion can be divided into threshold devices and ring imaging devices. The Čerenkov ring imaging counters certainly provide the best particle identification in the required momentum range. Besides high costs, their biggest disadvantage is, however, the large amount of inactive material in front of the electromagnetic calorimeter. It amounts to $\sim 20\%$ of a r.l., thus affecting the energy resolution of $\sim 15\%$ of all photons. This has motivated us to search for alternatives. Interesting candidates are Čerenkov threshold counters using silica aerogel as radiator, because these low density devices provide good K/π separation up to high momenta. They were first proposed by E. Lorenz for the PSI and CERN B Factory projects. [38]

We will discuss a particle identification system using silica aerogel threshold Čerenkov counters. After reviewing the main properties of silica aerogel, we will propose different cell designs, discuss their light yields and efficiencies, and present the results from a Monte Carlo simulation. We also present preliminary test results and address technical issues leading to improvements. We then discuss the particle identification performance of aerogel threshold Čerenkov counters and compare it to that of CRID/RICH counters. Finally, we give conclusions and an outlook on future R&D.

Candidate Čerenkov systems are based on threshold and ring-imaging techniques

6.2. Properties of Silica Aerogel

Silica aerogel is produced by hydrolysis and condensation of silicon alkoxides

Silica aerogel is a glass-like structure which consists of amorphous SiO_2. In the classical synthesis it is produced in a single-step process by hydrolysis and condensation of silicon alkoxides: [39,40]

$$Si(OR)_4 + 4H_2O \overset{basecatalyst}{\Longleftrightarrow} Si(OH)_4 + 4ROH \quad (hydrolysis)$$

$$Si(OH)_4 \overset{basecatalyst}{\Longleftrightarrow} SiO_2 + 2H_2O \qquad (condensation)$$

where R represents an acylgroup ($CH_3, C_2H_5, ...$). Ammonia is typically used as a catalyst. During condensation, 3-20 nm size clusters of SiO_2 are formed, which then partially collapse forming a structure of colloidal connected spheres, as shown in Figure 32. The sphere sizes are typically 4-6 nm in diameter, whereas the diameters of the pores range from 30 to 50 nm. Since the density range achieved in this production process is rather limited (22 mg/cm^3-280 mg/cm^3), a new two-step process has been developed at Lawrence Livermore National Laboratory (LLL) to produce both lower-density and higher-density aerogel.[41] In the first step a partially hydrolyzed, partially condensed silica is formed. This is achieved by supplying only 2 moles H_2O per mole of $Si(OR)_4$ for the hydrolysis, controlling the reaction by an acid catalyst and replacing all produced alcohol with a non-alcoholic solvent. In the second step both the hydrolysis and condensation are completed by adding water and a base catalyst. The density is controllable by the amount of non-alcoholic solvent. Due to the acid catalyst, silica aerogel produced by the two-step process has a different morphology, as shown in Figure 33.[42] It consists of chains of clusters, where the average chain diameter is 2 nm.

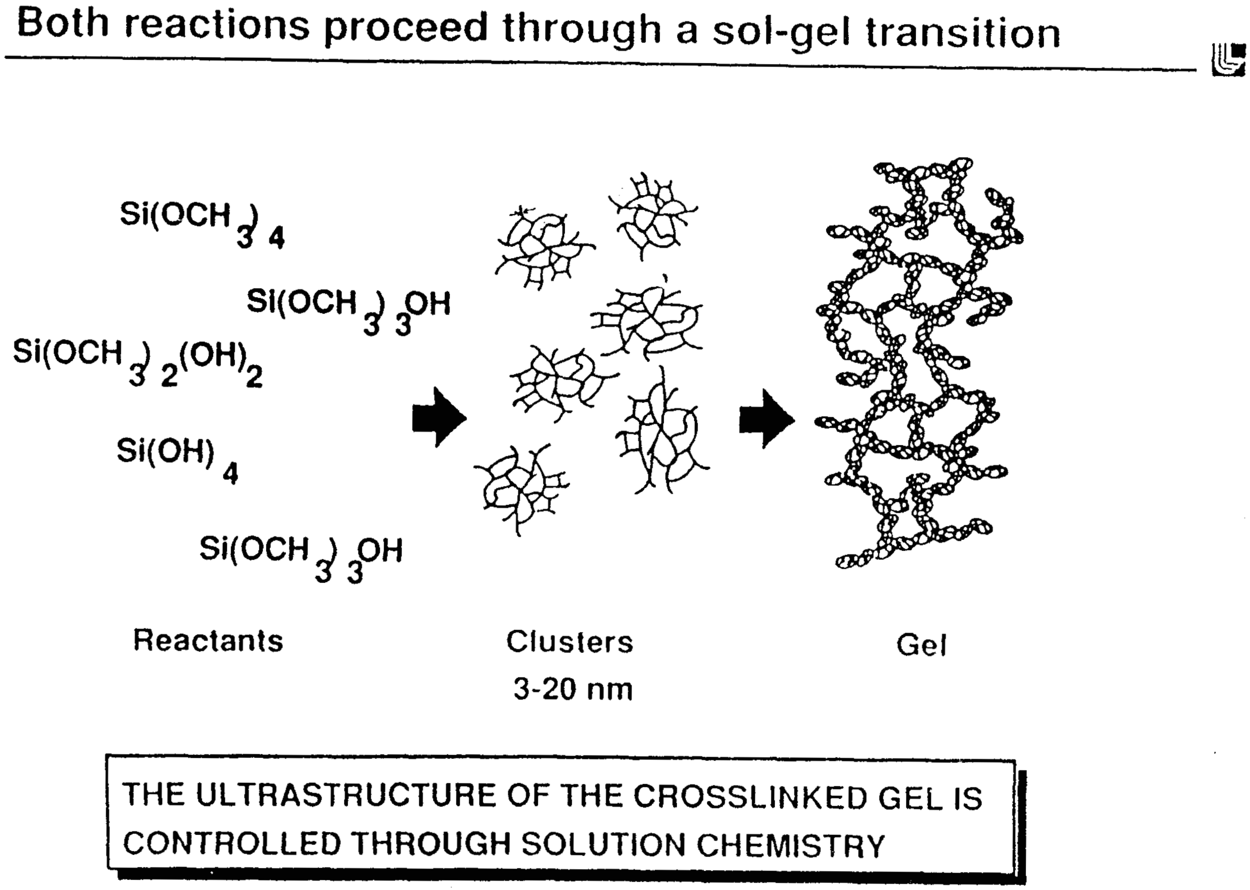
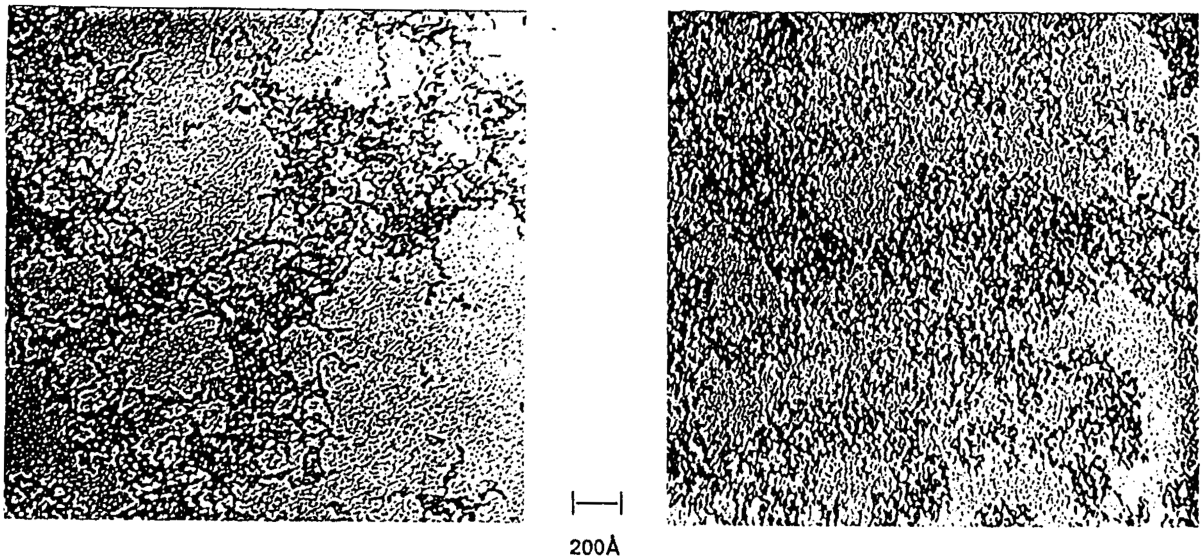


Figure 32. *Formation of silica aerogel.*

Aerogel is dried by a supercritical extraction to eliminate the effects of surface tension,[42] which tend to collapse the solid. Two different techniques are in use for the production of large aerogel monoliths. In the high-temperature technique, the solvent is evaporated at 260°C in an autoclave under a pressures of ~ 140 atm. This process, which takes up to 12 hours, produces hydrophobic aerogel. An additional heat treatment at 450° C, however, is frequently applied to remove alcohol groups still bound to the silicon. In this process, aerogel becomes hydrophyllic. In the low-temperature technique the solvent is exchanged with liquid CO_2. This procedure can take up to a week depending on the size of the monolith. The extraction is then done at a temperature of 35° C and a pressure of ~ 80 atm. The extraction, which takes 12-24 hours, produces hydrophyllic aerogel. This technique is well suited for the production of aerogels mixed with wavelength shifters.

High-temperature and low-temperature techniques exist for supercritical extraction



TEM of ULD aerogel (0.008 gm/cm^3) TEM of base/TMOS (0.04gm/cm^3)

Figure 33. *Transmission Electron Microscope (TEM) pictures of aerogels produced in a two-step process and in a single-step process.*

Silica aerogel is the lowest-density, transparent, man-made bulk material. Aerogels with densities ρ ranging from 3 mg/cm^3 to 600 mg/cm^3, can be produced. The relationship between the density and the index of refraction n is given by:

$$\rho = \frac{n-1}{0.21}\left(\frac{\mathrm{g}}{\mathrm{cm}^3}\right) \tag{1}$$

Silica aerogel with refractive indices between n=1.0006 and n=1.126 can be produced

These densities correspond to refractive indices of n=1.0006 and n=1.126 at 637.7 nm, respectively, which yield thresholds of 0.27 GeV/c to 3.47 GeV/c for pions and 0.95 GeV/c to 12.34 GeV/c for kaons. Since the production of large homogeneous blocks is straightforward, it is interesting to explore the use of these aerogels for radiators in threshold Čerenkov counters.

Density fluctuations are controllable[42] to $\Delta\rho/\rho \leq 1\%$; most inhomogeneities occur within a few mm of the edges. The uncertainty of the momentum threshold due to density fluctuations is approximately given by:

$$\frac{\Delta p}{p} =\simeq \frac{1}{2}\frac{\Delta\rho}{\rho}. \tag{2}$$

Since both the pore and sphere dimensions are close to wavelengths in the UV, aerogel scatters light like a Rayleigh

Čerenkov light is Rayleigh-scattered in aerogel

scatterer. For a homogeneous and isotropic aerogel, the Rayleigh cross section is given by:

$$\frac{d\sigma}{d\Omega} = a^6 k^4 \frac{|n^2 - 1|^2}{|n^2 + 2|^2} \frac{1}{2} \left(1 + \cos^2 \theta\right) \tag{3}$$

where θ is the scattering angle, a is the size of the scatterer and $k = 2\pi/\lambda$ is the wave number. Thus, for low-density aerogels ($n - 1 \ll 1$), the scattering cross section is basically proportional to $(n-1)^2$, λ^{-4} and $1 + \cos^2 \theta$. Since the Čerenkov spectrum falls as λ^{-2}, the original directionality of the Čerenkov photons is lost, after passing a few centimeters of aerogel. However, in the visible and infrared light region aerogel is highly transmissive as shown in Figure 34.[40]

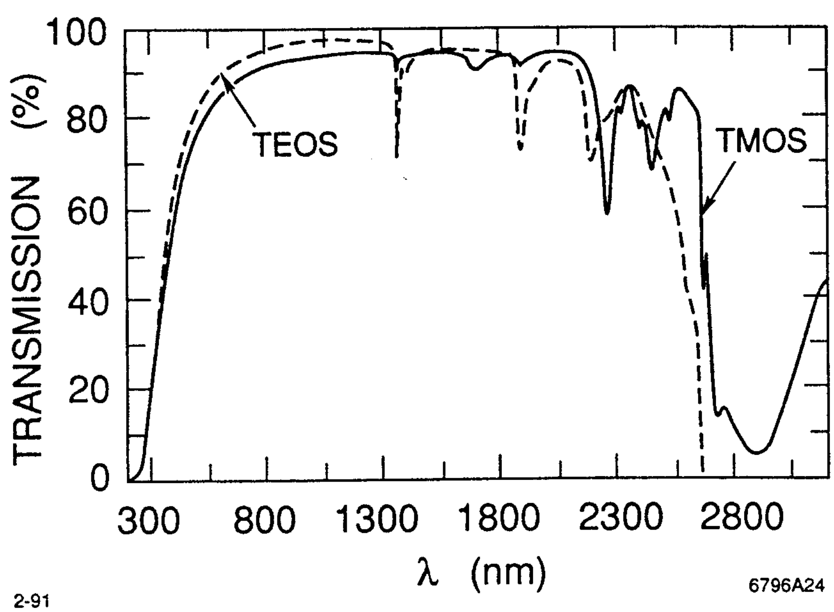


Figure 34. *Transmission spectra from a 4 mm thick sample of aerogel catalyzed from tetramethylorthosilicate (TMOS) supercritically dried at 270°C and 110 bars and from a 12 mm thick sample of aerogel catalyzed from tetraethylorthosilicate (TEOS), supercritically dried but not heat treated.*

Aerogel is very fragile; pieces easily break off at edges and corners, and cracks may develop inside the aerogel when handled, since its tensile strength is small. Near the edges, stresses caused by the mold, may build up, which eventually lead to cracking. However, an internal "frozen" stress,[42,43] which is formed in the gelation process due to an alignment of fine chains of silicon, causes no cracking even over longer periods of time, since it is small. In our application cracks and broken fragments have little effect on performance.

Aerogel is quite fragile, but can be machined

Machining aerogel is possible if done carefully. Except for lower-density material ($\rho \leq 60$ mg/cm^3), aerogel can be held in a vise with appropriate cushioning. Smooth cuts can be obtained with a diamond blade running in a mill at 4200 rpm. However, it is important to wear a mask, since fine airborne SiO_2 dust is generated. We have cut 1 inch thick pieces with a 4 inch blade. Cuts with diamond-covered wire provide a less smooth surface but may be more suitable for cutting thicker pieces. In order to catch the piece being cut, we hang a soft cotton towel behind the mill. Holes can be drilled with an ordinary drill operated at 4200 rpm. Fine polish is obtained with Kleenex tissue or soft cotton. Machining of low-density material is much more difficult, because it is so fragile. A special vacuum chuck is required to hold the aerogel. Not much experience has been gathered so far. For our application, however, machining of low-density aerogel may not be an issue, since molds of any given shape can be made of glass, anodized aluminum or silicone.

Aerogel is damaged by most gases and liquids

Aerogel is affected by most fluids. Hydrophyllic aerogel exposed to water is destroyed immediately, since H_2O molecules fill the pores. This results in a collapse of the aerogel structure which can be rather explosive. Acetone and alcohol have similar destructive effects on any aerogel, turning it into a white powder immediately. Contact with oil results in visible surface damage. Exposure of aerogel to organic liquids and gases, even in small amounts, must be avoided, since it may cause structural changes, leading to a density increase or other damage. Therefore, outgassing of paints and wavelength shifters, which cover the walls of the aerogel containers, is an important issue.[44] So far only fluid paper (green label) does not show visible surface effects. Detailed studies must be performed to find appropriate materials.

7. Conceptual Design of a Cell

We first describe a cell design similar to that in the PSI proposal.[38,45] The principle cell layout is sketched in Figure 35. Small cells of aerogel are read out by a fluorescent fiber and an avalanche photodiode. Two cells, with refractive indices of n=1.06 and n=1.008, are necessary to cover the entire momentum range from 0.4 to 4.0 GeV/c . Below 0.5 GeV/c, dE/dx is expected to provide good K/π separation. In the forward region of the detector, cells with lower refractive indices may be used. For example, the refractive index n=1.06 may be lowered to n=1.04 to increase the proton threshold such that their observation is unlikely, while the refractive index n=1.008 may be lowered to n=1.006 to increase the kaon threshold up to 4.5 GeV. For mechanical reasons, however, n=1.008 is preferred over n=1.006. In Table 7, thresholds for pions, kaons and protons are listed in addition to the Čerenkov angles for different refractive indices under consideration.

Two aerogel cells with different refractive indices suffice to cover the entire momentum range

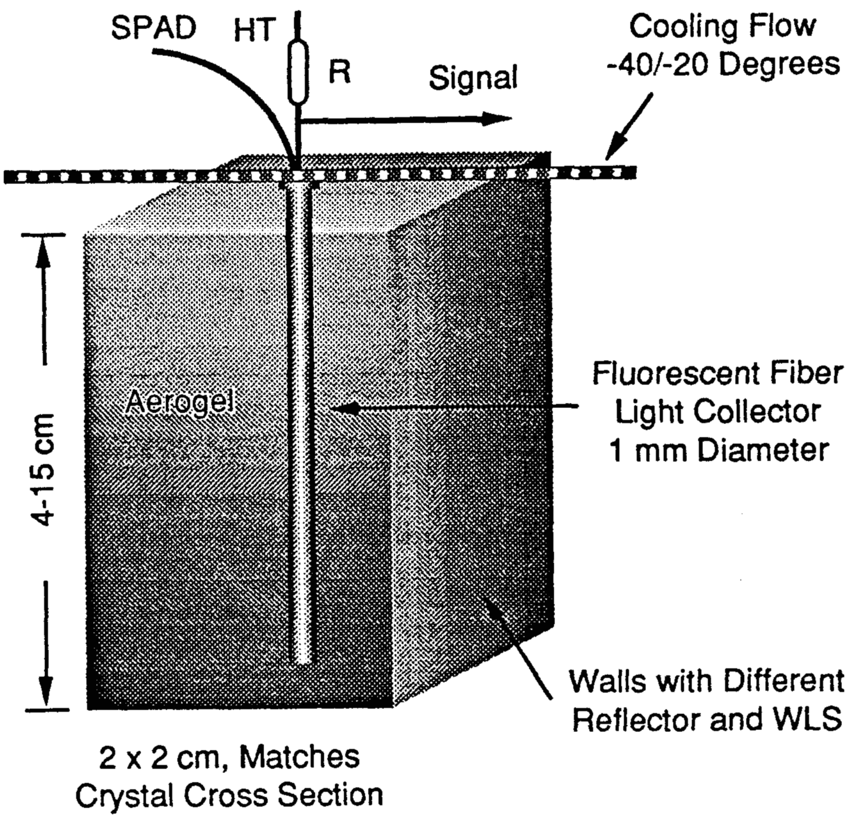


Figure 35. *Basic cell design for fluorescent fiber plus SPAD readout.*

One design uses small cells of aerogel read out with a fluorescent fiber coupled to a SPAD

Typical block sizes are $3 \times 3 \times 4$ cm^3 for n=1.06 and $2 \times 2 \times 10$ cm^3 for n=1.008. The blocks are appropriately tapered to provide projective geometry, matching the boundaries of the CsI crystals in the electromagnetic calorimeter. In order to optimize the light yield, the lower-density cell comes first, the higher-density cell second. Each block is housed in a light-tight, gas-tight, low-density container. This can be achieved with aluminized mylar, covered on one side either by black paint or black cardboard. The inner walls are sprayed with a highly reflecting white paint, which in addition may be coated with a thin layer of wavelength shifter (WLS), such as p-terphenyl, POPOP or PMP. The wall coating with WLS is most effective in a small-cell geometry, since here the probability for shifting UV light to higher wavelengths, where aerogel is highly transmissive, is larger than the probability for scattering and absorption. An alternative solution for a large-cell geometry is to mix a WLS into the aerogel. This idea will be discussed in more detail below.

Table 7. *Momentum thresholds and Čerenkov angles for different refractive indices.*

n	p_π [GeV/c]	p_K [GeV/c]	p_p [GeV/c]	θ_c
1.06	0.4	1.42	2.7	19.4°
1.0173	0.75	2.65	5.03	10.6°
1.026	0.6	2.15	3.9	12.9°
1.008	1.1	3.9	7.4	7.2°

To collect the photons from the aerogel and transport them to a photon detector, a thin fluorescent flux concentrator (FFC) of 1 mm diameter is inserted into a channel which is either drilled or made by the mold. The FFC is basically a transparent unclad plastic fiber which contains several different fluorescent laser dyes. The emission and absorption spectra of these laser dyes are matched such that light of lower wavelengths is shifted into a region where the photon detector has optimal quantum efficiency. Flux concentrations of a factor 10-100 are achievable.

For the photon detector, a single photon avalanche diode (SPAD) is used, matching the diameter of the fiber. For optimal coupling of the fiber to the SPAD, a customized holder is mounted to the SPAD after removing its window (see below). The fiber is inserted through a hole in the center, minimizing the distance between the photosensitive surface and the fiber. The SPAD, which works in a magnetic field, is operated in the Geiger mode, thus providing a high gain of $> 1 \times 10^8$.

SPAD's require quenching

The SPAD must be quenched using either an active or passive circuit, in order to prevent a continuous discharge of the photodiode. In passive quenching, the photodiode is normally kept in the non-conducting state by using a large load resistor R_L. An avalanche triggered by a photoelectron or bulk-generated electron in the depletion region discharges the photodiode from its reverse bias V_R to a voltage slightly lower than the breakdown voltage V_{BR}. Recharging it back to the operating voltage $V_R - I_{ds} \cdot R_L$, where I_{ds} is the surface dark current, has a time constant CR_L of typically 300 ns where C is the total capacitance including all stray capacitances. In active quenching a circuit drops the operating voltage below the breakdown voltage for a few 100 ns after an avalanche discharge. This allows for a collection of all remaining electrons and holes. When the operating voltage is reapplied, there are no electrons left in the depletion layer to trigger a new avalanche. Recharging through a small load capacitor or a transistor can be very rapid. The advantages of active quenching are shorter dead times and a synchronization of the detector with the beam crossing. A slight disadvantage is that the probability of after-pulsing is higher: 10% for 100 ns delay time compared to 0.6% for 300 ns.

SPADs are rather noisy. The noise rate is 10 kHz at room temperature. Cooling to $-20°$C reduces the noise by a factor of 50. Since a thermally-generated electron may trigger an avalanche, noise levels have to be reasonably low. If the SPADs are gated for 50 ns , one expects 5×10^{-4} noise counts at room temperature. Thus, in a system of 20,000 cells, 10 randomly distributed SPADS will fire accidentally per event. This is certainly tolerable. Lower noise occupancies can be achieved by cooling the SPADs. However, this requires a cooling system and water-tight sealing of the cells.

Cooling to $-20°$ C reduces noise considerably

In this design, the region around the fiber has to be discarded, because charged particles traversing the fiber or the SPAD produce Čerenkov photons at lower momenta due to larger refractive indices and thus create a signal. Although such particles can be located from the projected trajectory in the drift chamber, they cannot be identified. In order to reduce inefficiencies the diameters of the fibers and SPADs have to be small compared to the cell dimension.

Unambiguous particle identification is provided by two cels of aerogel read out by three fibers

In order to obtain a very low noise performance without cooling and to keep the fiducial volume of the entire cell for most tracks, one can use two fibers per cell in coincidence. Three rather than four fibers suffice for two cells, since the front cell and back cell can share one fiber, as shown in Figure 36. If we call the fiber in the lower-density cell A, the one in the higher-density cell B and the fiber shared by both cells C, the following physical coincidence patterns can occur:

1. AC, BC: a pion with $p_\pi \geq 1.1$ GeV/c or a kaon with $p_K \geq 3.9$ GeV/c.
2. $\overline{AC}, BC$: a pion with $0.4 \leq p_\pi \leq 1.1$ GeV/c or a kaon with $p_K \geq 1.4$ GeV/c.
3. $\overline{AC}, \overline{BC}$: a pion with $p_\pi \leq 0.4$ GeV/c , a kaon with $p_K \leq 1.4$ GeV/c, , or a proton with $p_P \leq 2.7$ GeV/c.

Since the particle trajectory and its momentum are known, the observation of a coincidence pattern together with the dE/dx information provides identification of kaons and pions of all momenta inside the fiducial detector volume, while for K/p, π/e and π/μ separation additional information is necessary.

Note that the coincidence pattern $AC, \overline{BC}$ is only observed if the track misses the front cell, SPAD B does not fire due to an inefficiency, or SPAD A and SPAD C accidentally fire. Since the accidental rate is low (1 Hz for 100 ns coincidence time), and a geometric miss can be detected, the observation of $AC, \overline{BC}$ provides a measure for inefficiencies. Since the momentum is known, this observed coincidence pattern may still be useful for particle ID: for $p > 1.1$ GeV/c, SPAD B was inefficient and the particle was very likely a pion; for $0.4 \leq p \leq 1.1$ GeV/c, SPAD B was inefficient and SPAD A was accidentally triggered. To allow for more redundancy one may use four fibers where two are shared by both cells. If one again calls them A, B, C, and D, where C and D are the shared fibers, then one can form four coincidences: AC, AD, BC and BD, and analyze patterns similar to those discussed above.

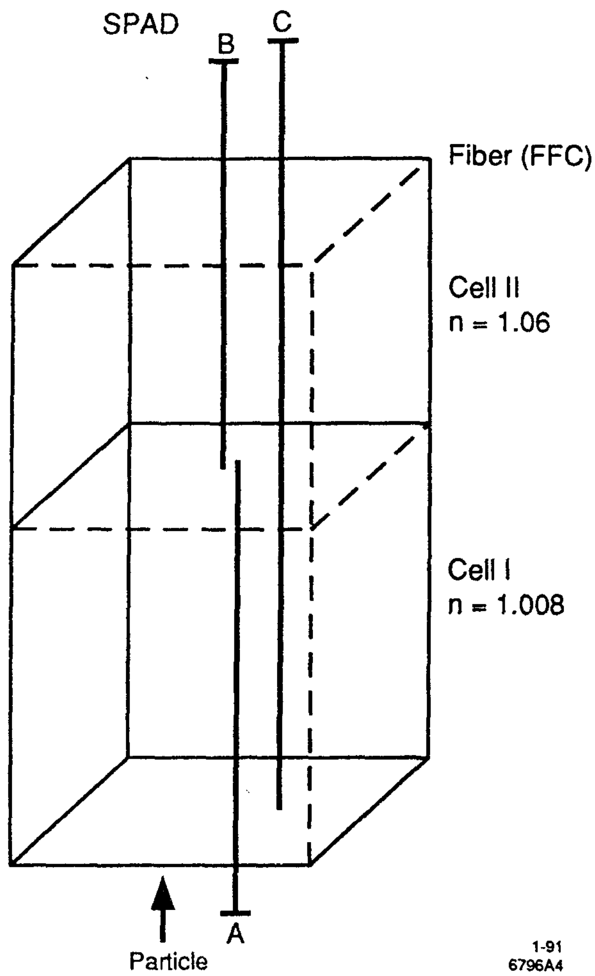


Figure 36. *A two cell design for a readout with three fluorescent fibers.*

An alternative design uses larger cells read out by vacuum phototetrodes

An alternative cell design consists of larger aerogel cells read out by a phototetrode (PT), as shown in Figure 37. Typical cell sizes are 12×12 cm^2 for n=1.008 and 13.5×13.5 cm^2 for n=1.06. As before, the minimum cell heights are 4 cm and 10 cm, but the cells are extended to allow for optimal coupling of the aerogel to the PT. To provide projective geometry, the cells are appropriately tapered. In this arrangement, each aerogel cell covers the solid angle of nine CsI crystals. Crucial for this design is the use of aerogel mixed with WLS to achieve large scattering lengths. The containers housing the cells are also fabricated from white-painted aluminized mylar. Each cell is read out by two 2 inch PT's, which are oriented parallel to the magnetic field. The PT of the front and back cells are interleaved to save space. The PT's have a maximum have a gain of 50; further amplification is thus required. Though the expected signals are small, as shown in the next section, preamp noise is tolerable for this readout scheme. The equivalent noise charge of a preamplifier for this circuit is given by:[46]

$$\overline{ENC}^2 = 8e_n^2\frac{1}{\lambda_t}[C_d + C_{FET}] \tag{4}$$

where:

e_n: *rms* spot noise of the FET ($V/\sqrt{Hz}$)
C_d: capacitance of the detector ($< 10\ pF$)
C_{FET}: capacitance of FET (on the order of $10\ pF$)
λ_t: integration time.

By integrating the pulses for a few microseconds, the ENC can be reduced to less than 30 *rms* electrons, although this is state-of-the-art performance.

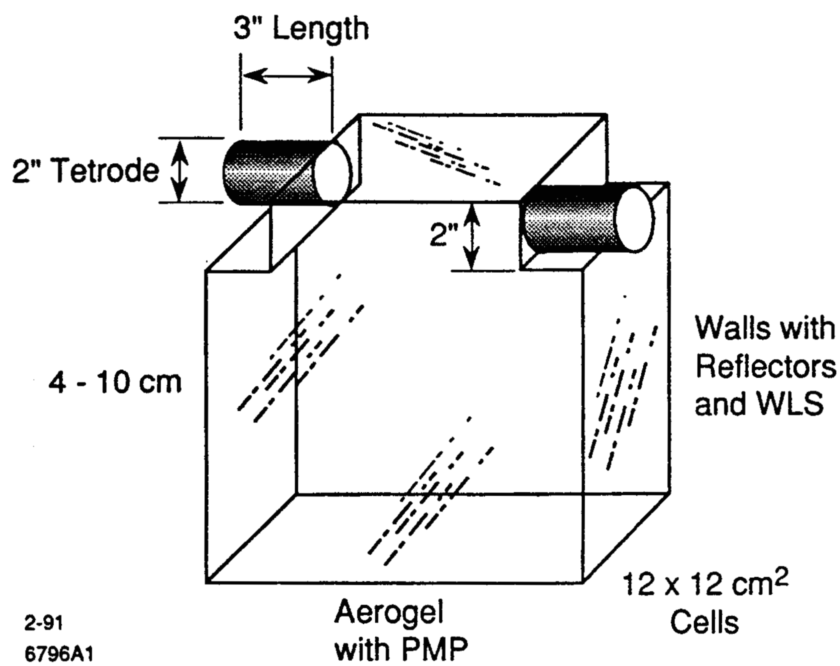


Figure 37. *Basic cell design for phototetrode readout.*

In order to keep the entire cell as fiducial volume, both PT must be put in coincidence. While the small-cell design with the FFC/SPAD readout provides high granularity, so that the probability of two tracks in an event hitting the same cell is very small, the large-cell design with PT readout reduces the number of channels by a factor of > 10. Figure 38 shows the implementation of aerogel threshold Čerenkov counters into the *B* Factory multipurpose detector. The magnified view shows the arrangement of the two cells in the presence of a cooling line.

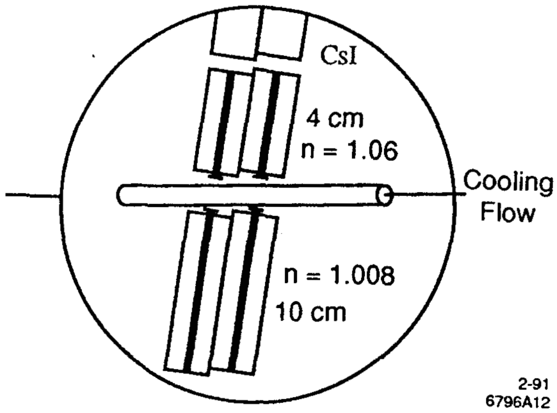


Figure 38. *Implementation of a two-cell aerogel threshold Čerenkov counter system into a general-purpose B Factory detector.*

7.1. Light Yield and Efficiency

A charged particle with a velocity β traversing a medium of refractive index n radiates Čerenkov photons under the angle $\theta_c = \cos^{-1}(1/(n\beta))$, if $\beta \geq 1/n$. Assuming that the cell height L is much larger than the wavelength λ ($\lambda \ll L$), the number of Čerenkov photons radiated per wavelength is:

$$dN = \frac{2\pi\alpha}{\lambda^2} L \sin^2\theta_c \, d\lambda \,, \tag{5}$$

where α denotes the fine structure constant. Integrating Eq. (5) over λ from 220 to 550 nm yields for a 1 cm thick block:

$$N = 1.25 \times 10^3 \sin^2\theta_c \text{ photons} \,. \tag{6}$$

An increase of the sensitivity range to 200-580 nm gives a 20% higher yield of Čerenkov photons.

We will first discuss the light collection process in the context of the FFC/SPAD-readout scheme. The Čerenkov photons have to undergo a complex process to reach the SPAD and trigger an avalanche. The photon path, consisting of several bounces off the container walls, is altered by several

Rayleigh scatterings in the aerogel. Typical values for different cell designs are given in the next section. For cell designs under consideration, the number of reflections is about 15. Before reaching the detector, the main losses are due to absorption in the aerogel and absorption on the walls. A photon reaching the fiber may be absorbed and emitted at a different angle with a higher wavelength. If $\sin\theta < 1/n_{fib}$ the photon is trapped in the fiber. Since the fiber contains several laser dyes, the photon may be reabsorbed and reemitted at a different angle and a higher wavelength. If the new angle does not satisfy the condition for total reflection, the photon might leave the fiber and eventually be absorbed in the aerogel. Several dyes are necessary to shift the photon to 800 nm, where the sensitivity of the SPAD is largest. Upon reaching the SPAD, the photon must create an electron-hole pair, which has to trigger an avalanche. Considering all these effects, the expected number of photoelectrons for the FFC/SPAD readout is given by:

$$N_{pe} = N_{ph} \cdot T_A \cdot R_W^{\bar{n}_r} \cdot Q_W \cdot A_{FFC} \cdot Q_{FFC} \cdot T_{FFC} \cdot Q_{F-D} \cdot T_D \cdot Q_D \tag{7}$$

where:

N_{ph}:	produced number of Čerenkov photons
T_A:	transmission in aerogel: $1 - \exp(-\bar{n}\bar{w}\mu_a) \sim 0.45$
$\bar{n}_r$:	average number of bounces at the wall (~ 15)
$\bar{w}$:	average path length between bounces
μ_a:	absorption coefficient, 10^{-2}cm^{-1}
R_W:	reflectivity of the walls, (> 0.95 for $\lambda < 300$ nm and > 0.98 for $300 < \lambda < 700$ nm)
Q_W:	quantum efficiency of WLS on the walls (0.9)
A_{FFC}:	photon absorption probability in the fiber (0.9)
Q_{FFC}:	quantum efficiency of dyes in the fiber (0.9)
T_{FFC}:	transmission efficiency in FFC including selfabsorption (0.2)
Q_{F-D}:	coupling efficiency of FFC and SPAD (0.25)
Q_D:	quantum efficiency of SPAD (0.65)
T_D:	efficiency for a photoelectron to trigger an avalanche (0.6)

For a 2×2 cm^2 cell, the efficiency is estimated to be $\epsilon = 1.9\%$, while for the 3×3 cm^2 cell, it is $\epsilon = 1.3\%$.

Using these estimates, Table 8 shows the number of produced photons and the photoelectron yield for different refractive indices computed for $220 < \lambda < 550$ nm and $200 < \lambda < 580$ nm.

For the alternative design of large aerogel cells with PT readout, the photoelectron yield is determined by:

$$N_{pe} = N_{ph} \cdot T_A \cdot R_W^{\bar{n}_r} \cdot T_{tet} \cdot Q_{tet}, \qquad (8)$$

where:

T_{tet}: transmission coefficient of the PT window (0.95)

Q_{tet}: quantum efficiency of the PT photocathode (0.2)

For a $15 \times 15\text{cm}^2$ cell, the transmission coefficient T_A is only 0.1, while the average number of reflections at the walls is 10. The efficiency is estimated to be $\epsilon = 1.6\%$, which is comparable to the efficiency of the FFC/SPAD readout. The yield of Čerenkov photons and photoelectrons estimated for a 2×2 cm^2 for different refractive indices is summarized in Table 8 .

The overall photon detection efficiency is estimated to be a few percent

n	L [cm]	$N_{ph}^{\dagger}$	$N_{pe}^{I\dagger}$	$N_{pe}^{II\dagger}$
1.06	4.0	552 (663)	10.5 (12.6)	8.8 (10.6)
1.0173	10.0	423 (508)	8.0 (9.7)	6.8 (8.1)
1.026	4.0	250 (300)	4.8 (5.7)	4.0 (4.8)
1.008	10.0	196 (235)	3.7 (4.5)	3.1 (3.8)

† The first set of numbers represent yields for the interval $220\ nm \leq \lambda \leq 550\ nm$, the second set for the interval $200\ nm \leq \lambda \leq 580\ nm$.

Table 8. *Number of Čerenkov photons and photoelectron yield for different refractive indices, using the calculated efficiency of a $2 \times 2\text{cm}^2$ cell.*

7.2. *Monte Carlo Simulations*

In order to obtain a realistic estimate of all factors entering Eqs. 7 and 8, it is necessary to study the photon paths in different cell designs in more detail. For this purpose a Monte Carlo program has been developed using vector algorithms. Čerenkov photons are generated uniformly in x, y, z inside a rectangular box. Assuming that the trajectory of the charged track is parallel to the z direction, which in turn is parallel

The Monte Carlo simulation includes the Čerenkov spectrum, Rayleigh scattering and wavelength-dependent absorption and wall reflection to the cell height, the direction of the Čerenkov photons is generated isotropically in ϕ and $\cos\theta$, with $\cos\theta$ limited to the interval $[1, \cos\theta_c = 1/(\beta n)]$. In addition, a wavelength is generated for each photon according to the Čerenkov spectral distribution $dN/d\lambda \propto 1/\lambda^2$. Next the path length is determined, considering diffuse reflection on the walls with wavelength-dependent absorption. Both Rayleigh scattering and wavelength-dependent absorption in the aerogel are considered. The scattered photon is generated with the correct angular distribution. If the photon hits the fiber or the PT, it is recorded as a hit. In the future, it is planned to include wavelength shifting in the aerogel, wavelength shifting on the walls plus a detailed simulation of the transport of photons through the fiber. Furthermore, the simulation of different geometric shapes, such as trapezoids and pyramids will be made available.

With the current Monte Carlo model, samples of 10,000 events have been generated both for the FFC/SPAD-readout scheme and the PT-readout scheme for different cell sizes.[47] The index of refraction has been set to $n = 1.06$. Figures 39 and 40 show the resulting distributions for the number of reflections n_r, number of Rayleigh scatterings n_{sc}, total path length l_g, and the path length between reflections l_g/n for a $2 \times 2 \times 10$ cm^3 cell, read out with a 1 mm fiber, and a $15 \times 15 \times 15$ cm^3 cell, read out with a PT, respectively. The average values for these distributions are given in Table 9 together with the corresponding averages for other geometric arrangements. Figure 41 displays the average values for $\bar{n}_r$, $\bar{n}_{sc}$, $\bar{l}_g$, and $\overline{l_g/n_r}$ as a function of the cell width, d, for cell heights, h, of 4 cm and 10 cm in the FFC/SPAD-readout scheme and for 15 cm long cells in the PT-readout scheme. For the FFC/SPAD-readout scheme, $\bar{n}_r$ first increases rapidly with d because of increasing path lengths. At about $d = 4$ cm, the path lengths are so large that due to absorption, the tails of the n_r distribution are truncated and thus $\bar{n}_r$ starts to fall. Since the fiber is parallel to the cell height, the dependence of $\bar{n}_r$ on d is weak; for example the maximum $\bar{n}_r$ is 16 for $d = 4$ cm compared to $\bar{n}_r = 14$ for $d = 10$ cm, being slightly larger for shorter cells because average path lengths between reflections are smaller and thus tails in the n_r distribution are

The efficiency depends strongly on the transverse cell size and weakly on the cell length larger. The average number of Rayleigh scatterings increases rapidly with d due to an increase in $\bar{l}_g$, whereas $\bar{n}_{sc}$ is basically independent of h. The average length between reflections rises slowly with d and is obviously larger for long cells than for short cells. The efficiency decreases rapidly with d, as expected, but is independent of h. Since for the PT-readout scheme the average path length is already rather large, tails of the n_r and n_{sc} distributions are truncated, thus yielding lower averages. Larger cell sizes do not alter this fact, such that $\bar{n}_r$ is 10 independent of d and n_{sc} rises only slowly with d. The variable $\overline{l_g/n_r}$ is about a factor of two larger than for the FFC/SPAD-readout scheme. The efficiency is rather low, but is about a factor of two larger than for similar cell sizes in the FFC/SPAD-readout design.

d [cm]	h [cm]	$\bar{n}_r$	$\bar{n}_{sc}$	$\bar{l}_g$ [cm]	$\overline{l_g/n_r}$ [cm]	ϵ %
0.5	4.0	7.5±0.1	8.3±0.1	2.32±0.03	0.366±0.004	77.9±0.9
1.0	4.0	12.4±0.2	21.9±0.4	5.6±0.09	0.58±0.005	58.1±0.8
1.5	4.0	15.2±0.3	34.9±0.8	8.68±0.17	0.62±0.007	43.4±0.7
2.0	4.0	16.3±0.3	44.0±1.2	11.0±0.3	0.72±0.01	32.7±0.6
3.0	4.0	16.9±0.4	54.2±1.8	14.8±0.5	0.9±0.01	21.4±0.5
4.0	4.0	16.0±0.5	61.0±2.7	16.3±0.5	1.02±0.02	13.4±0.4
5.0	4.0	15.8±0.6	57.3±2.6	19.5±1.0	1.2±0.03	10.4±0.3
10.0	4.0	11.7±0.9	71.7±6.1	19.2±2.0	1.4±0.03	3.3±0.2
0.5	10.0	6.9±0.1	8.5±0.1	2.3±0.03	0.395±0.004	79.2±0.9
1.0	10.0	11.5±0.2	23.6±0.5	5.5±0.09	0.54±0.006	58.3±0.8
1.5	10.0	13.8±0.2	34.2±0.8	9.0±0.2	0.70±0.009	44.0±0.7
2.0	10.0	14.1±0.3	45.0±1.2	10.8±0.3	0.82±0.01	33.8±0.6
3.0	10.0	13.8±0.4	53.3±1.7	14.1±0.5	0.99±0.02	21.4±0.5
4.0	10.0	13.7±0.5	62.4±2.6	16.1±0.7	1.16±0. 03	14.6±0.4
5.0	10.0	13.1±0.5	59.3±3.0	19.2±0.9	1.4±0.06	11.1±0.3
10.0	10.0	10.1±0.7	67.2±5.2	20.1±1.8	1.8±0.08	3.3±0.2
5.0	15.0	9.9±0.2	39.2±2.3	18.2±0.5	2.7±0.05	18.5±0.4
10.0	15.0	9.7±0.2	40.5±1.8	27.0±1.0	3.6±0.09	7.7±0.3
15.0	15.0	9.4±0.2	50.8±2.0	28.8±1.9	4.0±0.2	3.5±0.2

Table 9. *Monte Carlo results*

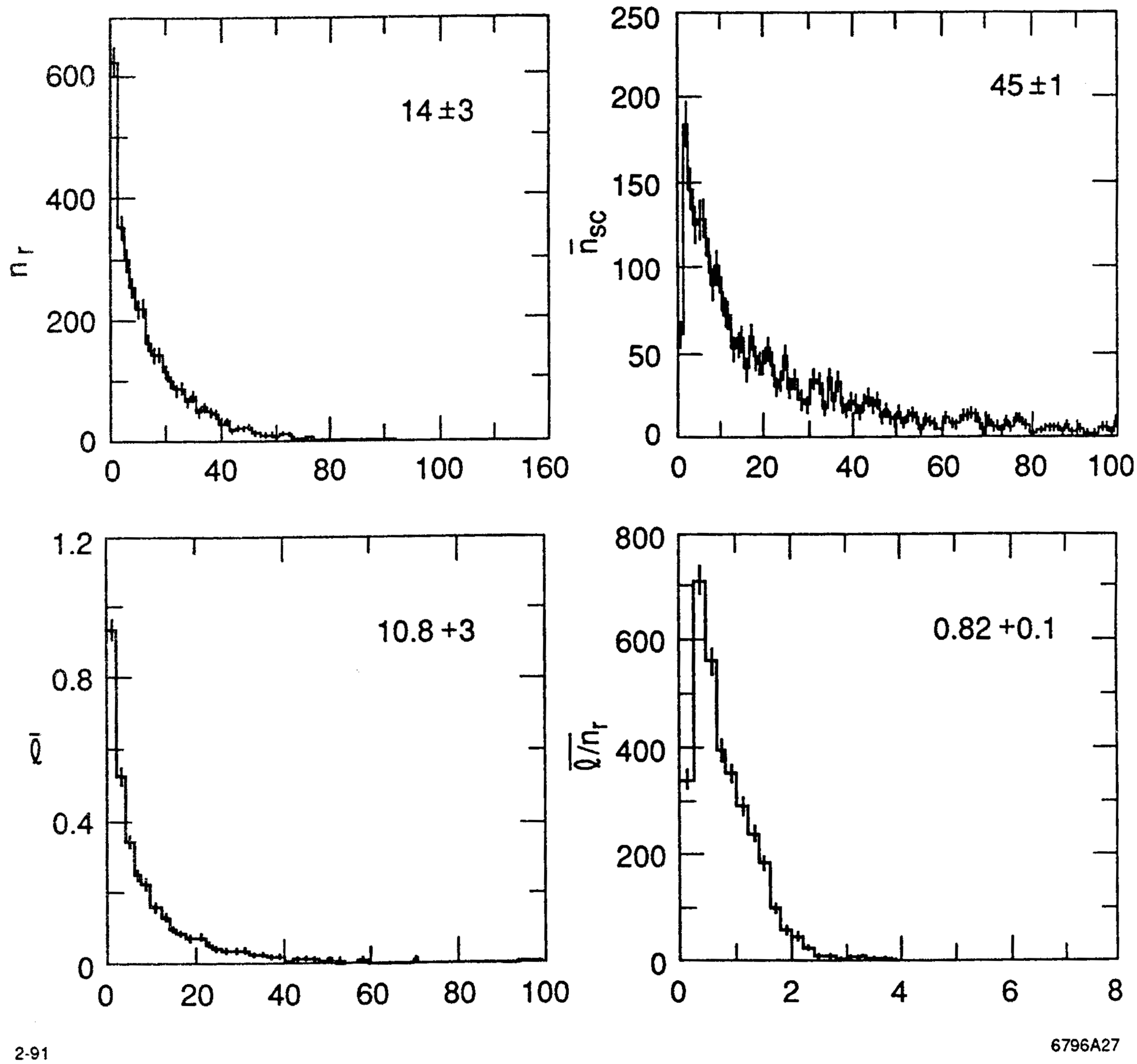
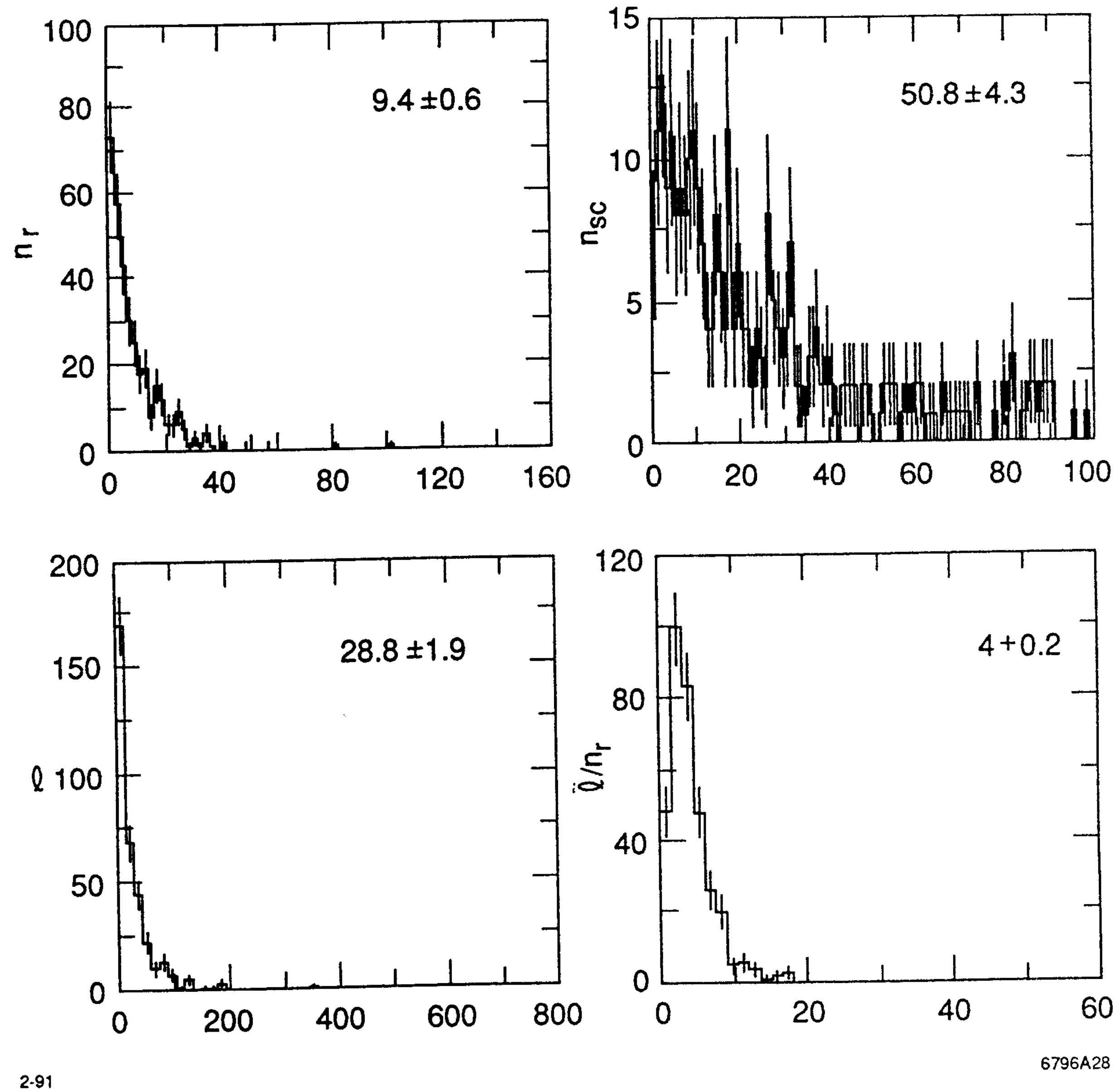


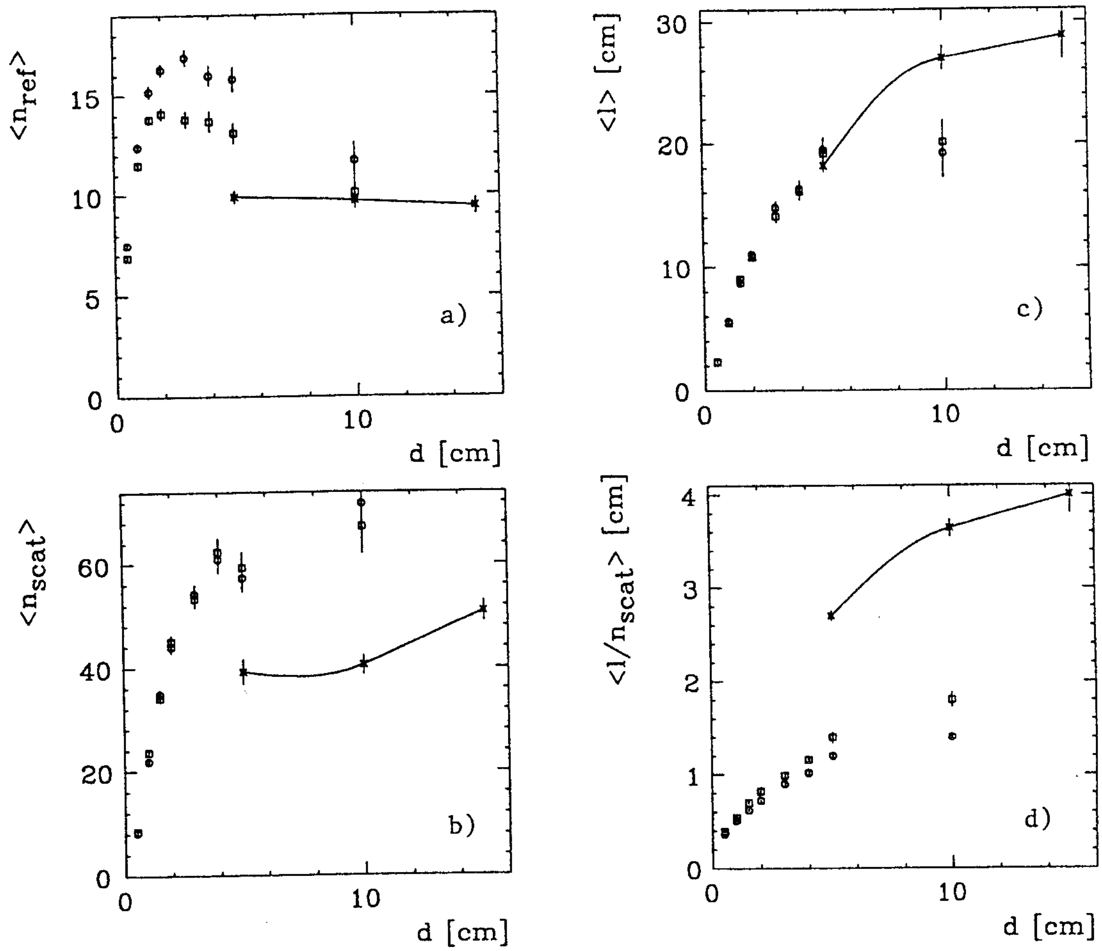
Figure 39. *Monte Carlo distributions for a 2 × 2 × 10 cm³ cell with FFC/SPAD readout:* **a.** *number of reflections,* **b.** *number of Rayleigh scatterings,* **c.** *total path length and* **d.** *path length per number of reflections.*

For both designs, the $(1 + \cos^2\theta)$ dependence of Rayleigh scattering, the simulation of the Čerenkov angle and a variation of the index of refraction have no effect on the quantities discussed above. The inclusion of wavelength-dependent scattering and absorption in the aerogel and absorption on the walls are rather important. For example, if these effects are neglected in the FFC/SPAD-readout scheme, $\bar{n}_r$ increases by about a factor of two and $\bar{l}_g$ by about a factor of four.



In conclusion, these studies show that the optimal cell size for the FFC/SPAD-readout scheme lies between 2 and 3 cm, while for the PT-readout design *d* should not be larger than 10 cm.

Figure 40. *Monte Carlo distributions for a* $15 \times 15 \times 15$ cm^3 *cell a 2 inch phototetrode readout:* **a.** *number of reflections,* **b.** *number of Rayleigh scatterings,* **c.** *total path length and* **d.** *path length per number of reflections.*



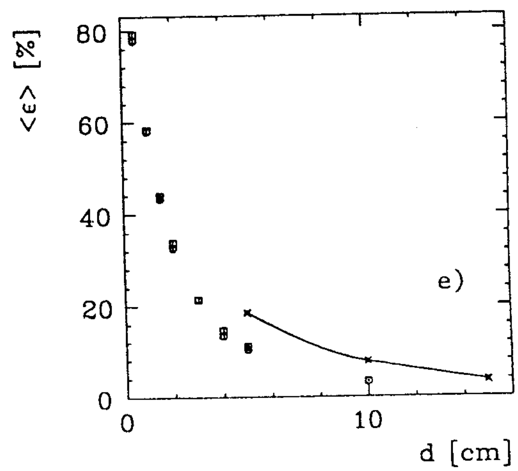


Figure 41. *Monte Carlo distributions as a function of cell radii for*
a. *the average number of reflections,*
b. *the average number of Rayleigh scatterings,*
c. *the average total path length,*
d. *the average path length per number of reflections, and*
e. *efficiency.*
Dots and squares show 4 cm and 10 cm long cells read out by a 1 mm fiber, while stars show 15 cm long cells read out with a 2 inch PT.

7.3. *Preliminary Test Results*

Our first aerogel studies consisted of transmission measurements with an HP spectrophotometer, in order to determine the absorption length Λ_a and the scattering length Λ_s. For a successful operation of aerogel counters, it is important to have large absorption lengths ($\Lambda > 40$ cm). Figure 42 shows the transmittance measured for three samples manufactured by Aerglas: a 1.8 mm thick piece with $n = 1.032$, which has been cut off a $12 \times 12 \times 2.5$ cm^3 thick block (several years old) and two recently made 3.0 cm thick pieces with n=1.025 and n=1.055. The measured curves can be fitted with an exponential function:

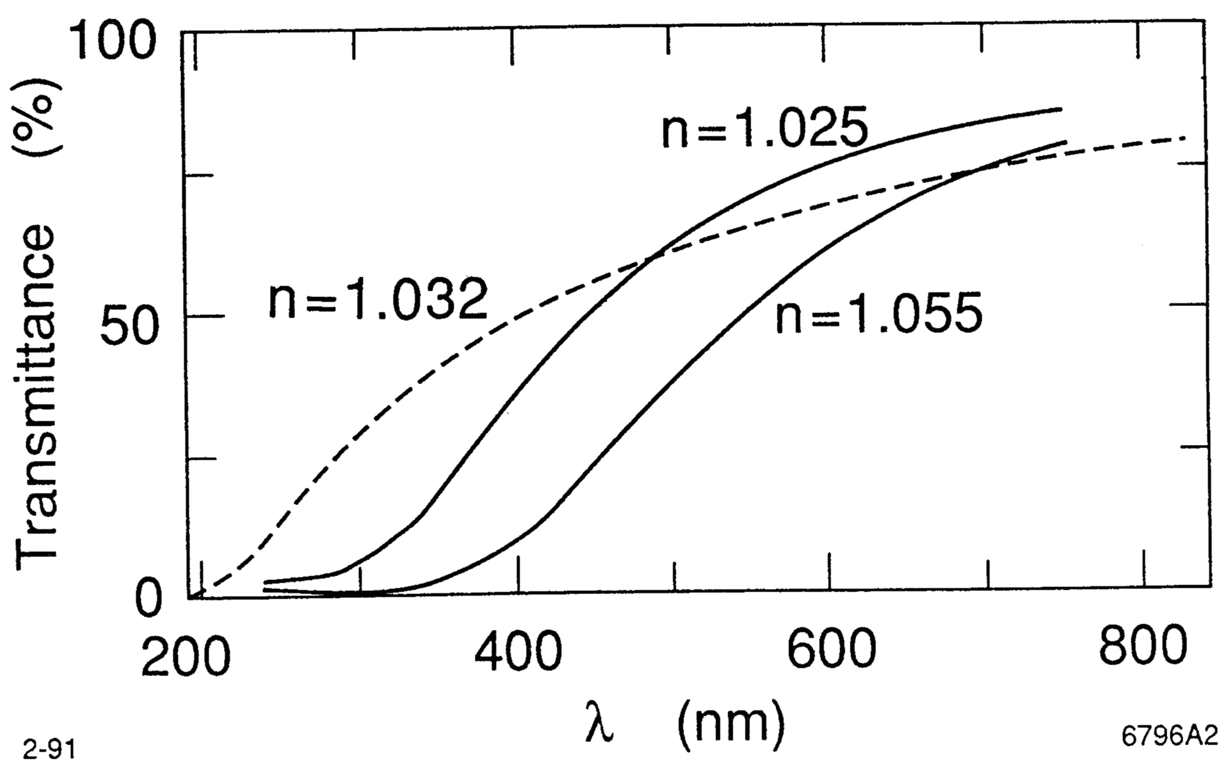
The transmission of aerogel is well-described by a λ^{-4} scattering term and a λ^{-2} absorption term

$$T = T_0 \exp(-\mu(\lambda)\, d), \tag{9}$$

where d is the thickness, T_0 is the maximal transmission accounting for reflection and $\mu(\lambda)$ is the attenuation coefficient, which can be parametrized by a Rayleigh scattering term and an absorption term:

$$\mu(\lambda) = a/\lambda^4 + b/\lambda^2. \tag{10}$$

Figure 42. *Transmission of different aerogel samples from Aerglas.*



The average scattering and absorption lengths for Čerenkov distributed light are given by:

$$\Lambda_s = \frac{1}{3a}\frac{(\lambda_{max}^3 - \lambda_{min}^3)\lambda_{max}\lambda_{min}}{\lambda_{max} - \lambda_{min}} \tag{11}$$

and

$$\Lambda_a = \frac{1}{b}\lambda_{max}\lambda_{min}. \tag{12}$$

where λ_{max} and λ_{min} are the maximum and minimum wavelengths detected by the counter.

For technical reasons, it is more convenient to determine the measured values for $\mu(\lambda)$ first and then fit these with the parametrization in Eqn. 10. The results are listed in Table 10. For comparison, measurement from other groups are also shown. Except for our aged samples, all other data are consistent. We have done several measurements with samples from the aged block, which yielded similar results. The most plausible explanation is an aging effect, as the aerogel may absorb water and gases from the air, thus changing its properties. Since similar effects have also been noticed by S. Schindler,[44] this implies that proper storage is absolutely necessary. In addition, the aerogel in a detector cell must be well isolated from humidity and a careful choice for the reflector must be made. Thus, for a well-handled sample of aerogel, the average scattering length is $\sim$ 3cm, while the average absorption length is $\sim$ 50cm for the interval $300 \leq \lambda \leq 800$ nm.

It is important to store and use aerogel under air-tight conditions

sample	thickness [cm]	n	Λ_a [cm]	Λ_s [cm]	sensitivity range
CIT	0.18	1.032	0.4	3.6	200-800
CIT [44]	3.0	1.055	50.0	3.2	300-800
CIT [44]	3.0	1.025	46.0	6.2	300-800
Bonn [48]	2.8	1.035	33.9	2.3	300-600
Novosibirsk [49]	2.5	1.035	> 100†	2-3†	-

† Estimated for pure SiO_2 aerogel

Table 10. *Absorption and scattering length in aerogel.*

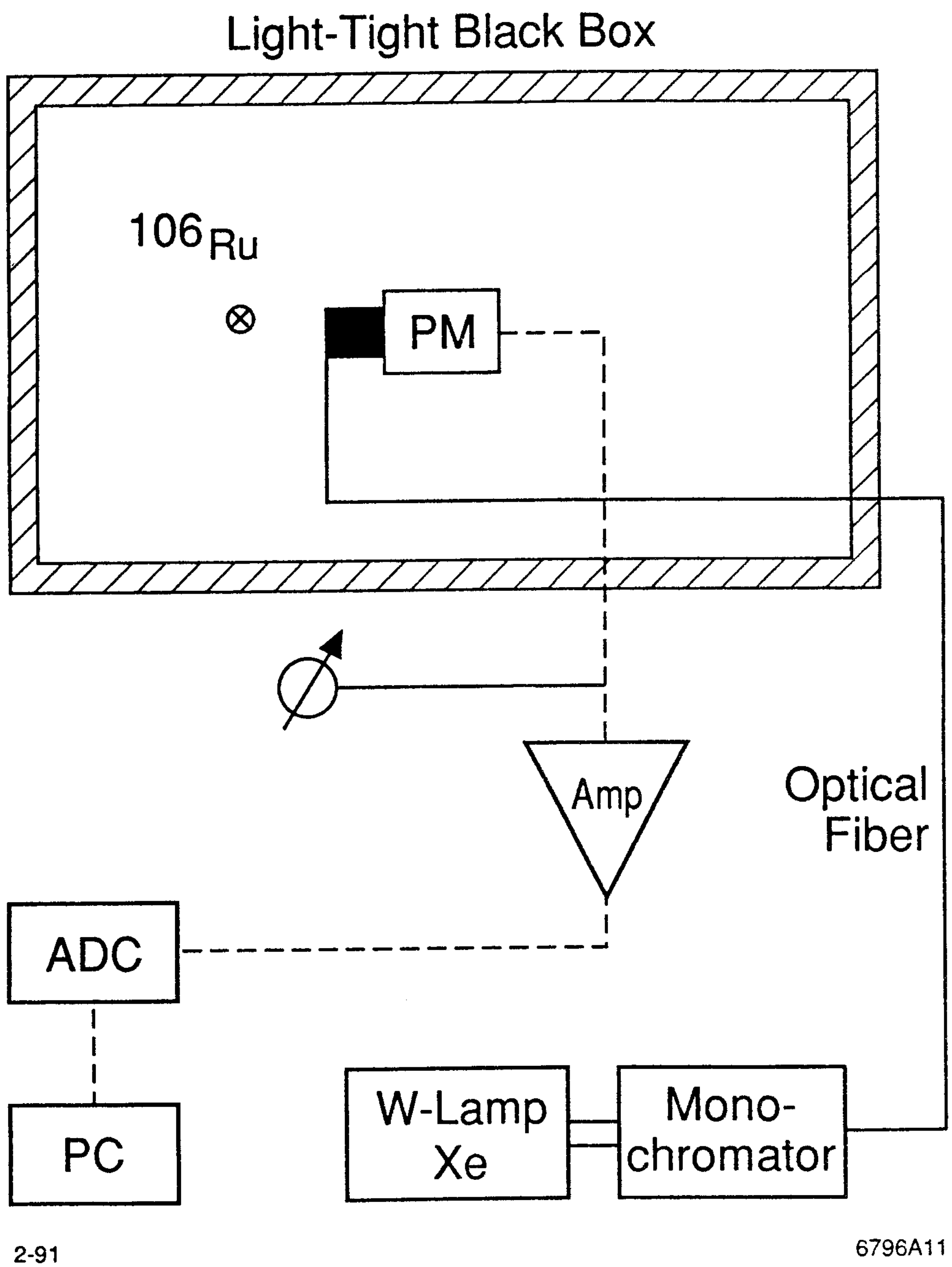


Figure 43. *Test setup for measuring the efficiency from a light source as function of wavelength.*

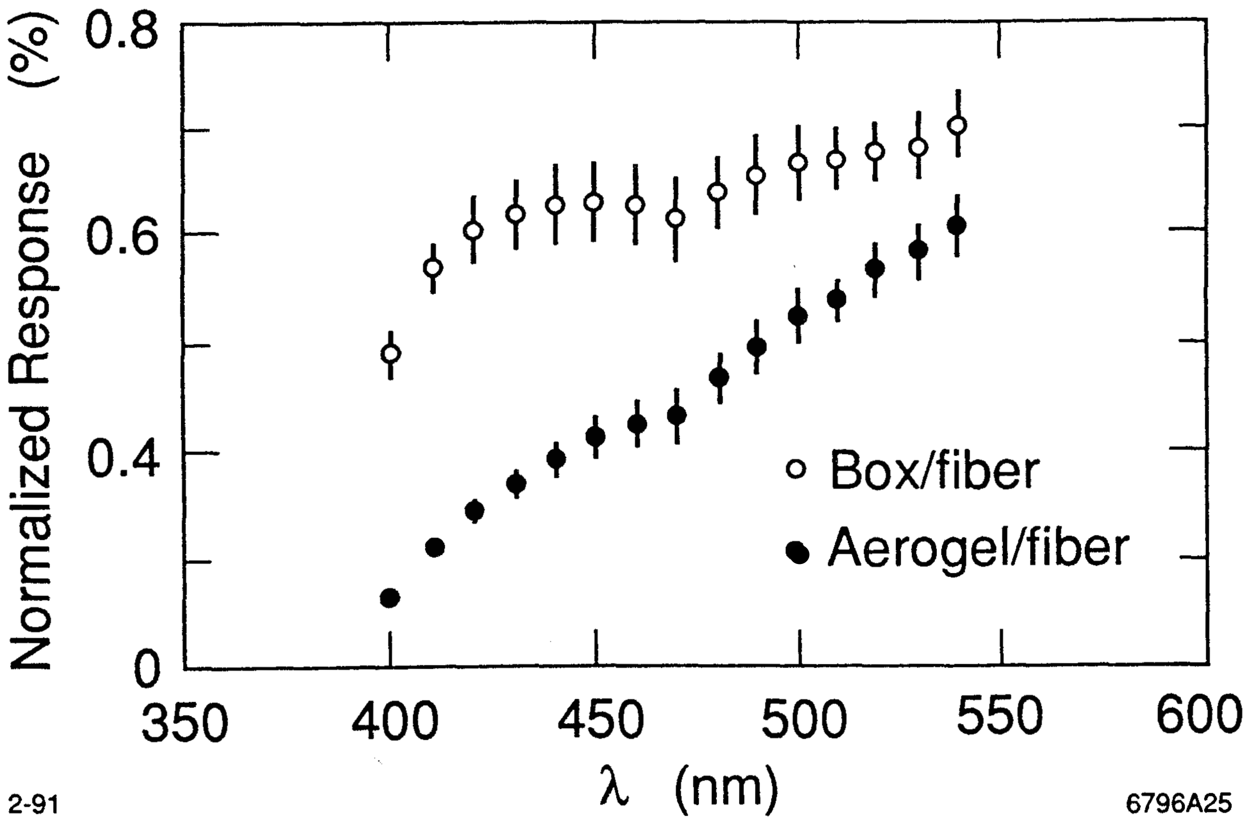


Figure 44. *Efficiency of a $3 \times 3 \times 4$ cm^3 aerogel cell in the visible-light region. Solid dots show the data for the cell with aerogel, light dots show the data for the container itself.*

In order to study the properties of individual aerogel cells and monitor their performance the test setup shown in Figure 43 was used. Pieces of aerogel housed in light-tight containers are attached to a photomultiplier, which is mounted in a black light-tight box. The containers are made of a 250 μm thick layer of black cardboard glued to a 50 μm thick layer of aluminized mylar, which is sprayed with highly-reflecting white paint (Nuclear Enterprise NE 560). For our current studies, light from either a tungsten or xenon lamp or electrons from the β decay of ^{106}Ru was used. The light, after passing through a monochrometer, was fed via a clear plastic fiber into the cell, in such a way that the direct beam could not hit the photocathode. For a selected wavelength interval, the anode current was recorded with a digital ammeter. For the measurements with the ^{106}Ru source, the anode signal was amplified (Ortec EG&G 560) and then passed into a peak-sensitive ADC (Ortec EG&G 916) operated by a PC.

Figure 44 displays the efficiency of a $3 \times 3 \times 4$ cm^3 block of aerogel cell for visible light as a function of the wavelength. For comparison, the efficiency for the reflecting box itself is also shown. For normalization, the intensity of the light beam is directly measured with the PM. The data are averages of three measurements; the errors are given by the corresponding standard deviations. In order to avoid saturation, measurements were performed at moderate high voltage and the light intensity was reduced appropriately. The data for the box without aerogel give the efficiency after $\bar{n}_r$ reflections, where for this geometry $\bar{n}_r$ is about 14. Above 420 nm, the efficiency depends rather weakly on λ, averaging $\sim 65\%$. This implies a reflectivity of $\sim 97\%$. However, for the UV region, the reflectivity is significantly lower, requiring improvement. The properties of different diffuse reflecting materials are discussed in the next section. We require a reflectivity of $\geq 95\%$ at 250 nm, increasing to 98% above 350 nm. The results for the aerogel cell show a strong wavelength dependence, as expected, due to Rayleigh scattering and absorption. These measurements will be extended to the UV-light region and visible red-light region. This technique can be used to study the dependence on geometric shapes, the effects of wavelength shifters mixed into the aerogel, and to monitor the aging of aerogel.

A reflectivity of 97% has been measured for NE560 paint in the visible region

The test setup shown in Figure 45 is used to measure the photoelectron yield from Čerenkov light in different aerogel cells. Electrons from a collimated ^{106}Ru source hit an aerogel cell perpendicular to the photomultiplier. To trigger, the electron must penetrate the aerogel and reach a plastic scintillator. A coincidence between the aerogel and the scintillator generates a gate signal for the ADC. The discriminators have an upper and lower threshold, allowing selection of specific energy intervals. The analog PM signal from the aerogel cell is amplified and fed into the Ortec ADC. In this arrangement it is unlikely for an electron itself to trigger. The electron energy spectrum is a typical β spectrum with an endpoint energy of 3.7 MeV. For $n = 1.032$ aerogel, electrons above 2 MeV produce Čerenkov photons.

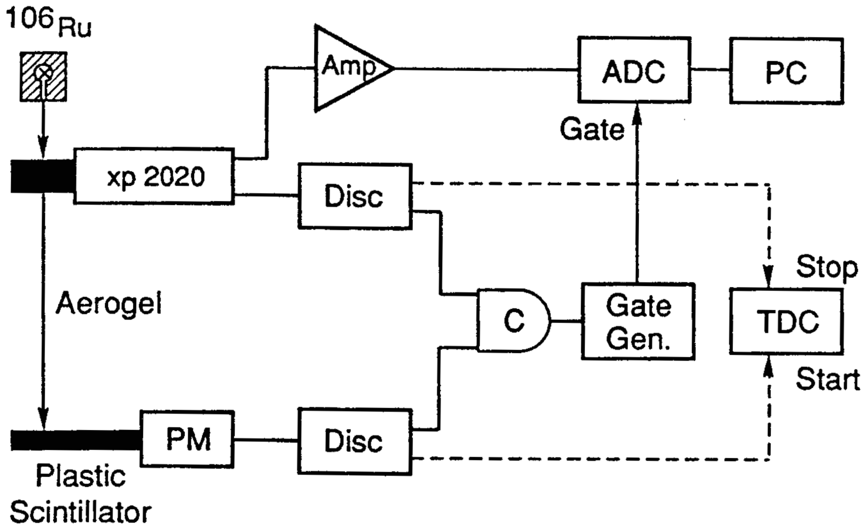


Figure 45. *Test setup for measuring the efficiency and photoelectron yield of Čerenkov photons.*

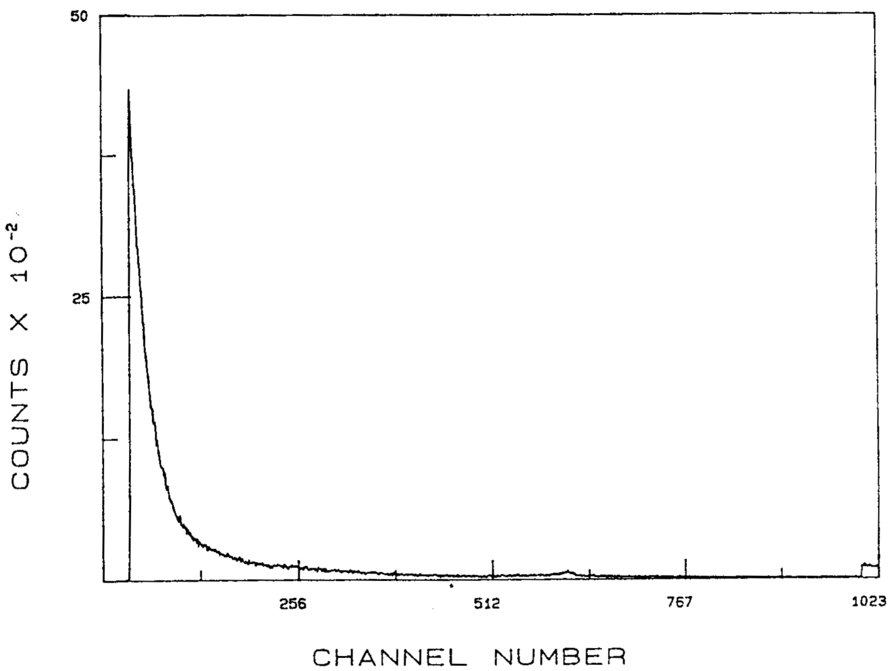


Figure 46. *Čerenkov photon spectrum observed with a ^{106}Ru source.*

A spectrum recorded with this arrangement is shown in Figure 46, indicating the observation of Čerenkov photons with a photomultiplier. Schindler has measured 5(12) p.e./cm in $n = 1.025(1.055)$ aerogel with photomultiplier readout.[44] Since the PMs were equipped with regular borosilicate windows, the sensitivity was restricted to photons above 350 nm. In order to measure timing properties of different readout systems, the logic signal from the scintillator starts a TDC which is stopped by the logic signal from the aerogel.

Initial tests to determine the efficiency of the FFC/SPAD readout were made by E. Lorenz at the Max-Planck Institute in Munich.[50] Figure 47 shows the experimental arrangement. A Y7 fiber from Kuraray was inserted into a block of plastic scintillator NE 111, which was read out with an XP 2020 PM on one end. On the other end the Y7 fiber was attached to an RCA SPAD (C 30902). The absorption spectrum of the Y7 fiber in comparison to emission spectrum of the NE 111 scintillator is shown in Figure 48. The PM signal after passing through a discriminator with a threshold set at 800 keV started a TDC. This value was chosen to trigger on the Compton edge of the 1275 keV ^{22}Na line. Sufficient light from this γ line is produced in the scintillator to trigger the TDC. Some of these photons were absorbed in the fiber and re-emitted at an angle smaller than the critical angle and therefore travelled inside the fiber. The scintillator end of the fiber contained a mirror. All photons thus eventually reached the SPAD, unless they were lost by another absorption and re-emission at an angle larger than the critical angle. The SPAD output provided the stop signal for the TDC. In addition, the SPAD signals coincident with the PM signal and the PM signals themselves were counted both with and without a SPAD dead-time veto. The efficiency, the ratio of the two numbers, is 90% at room temperature. Including a dead-time veto for the SPAD, increased the efficiency to 98% at room temperature. Additional cooling to $-20°$ improved the efficiency to 99.98%. The decay time spectrum of the Y7 fiber is shown in Figure 49. The time resolution for the TDC measurements was $\sigma = 1.3$ ns at room temperature and 0.9 ns at $-20°$ C. These results obtained in a high statistics environment, look very encouraging. However, a few deficiencies, which will affect the efficiency of the readout of a threshold Čerenkov counter, need to be remedied: a low dye concentration in the Y7 fiber, a spectral mismatch between the Y7 fiber and the SPAD (see Figures 48 and 56), a rather long decay time of 9.5 ns of the Y7 fiber, the coupling between the fiber and the SPAD, and area mismatch of the fiber and the SPAD by a factor of four.

An efficiency of 99.98% has been measured for the FFC/SPAD readout combination at $-20°$ C, after deadtime correction

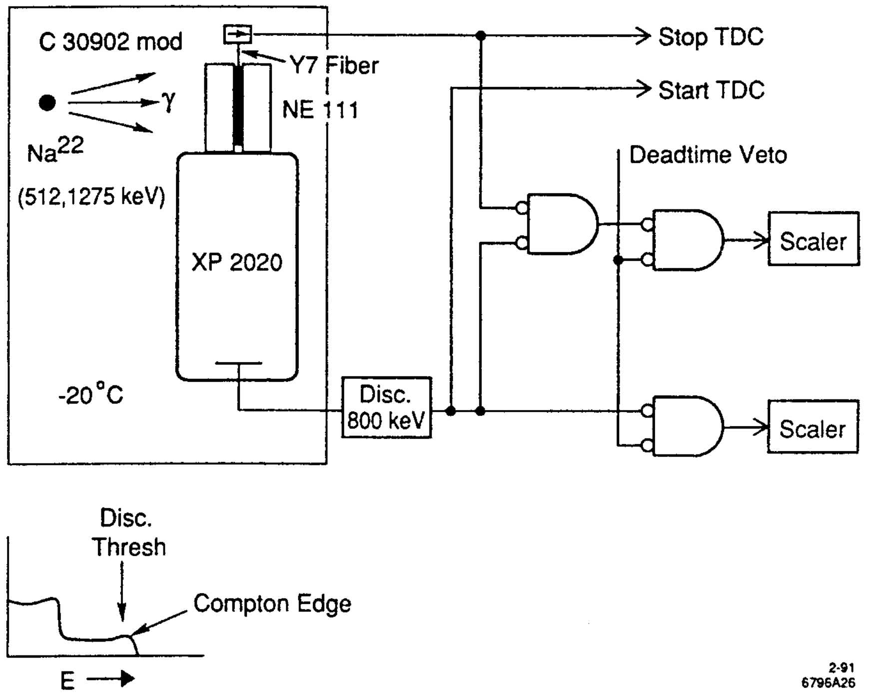


Figure 47. *Test setup for reading out an NE 111 scintillator with a Y7 fiber and a single avalanche photodiode.*

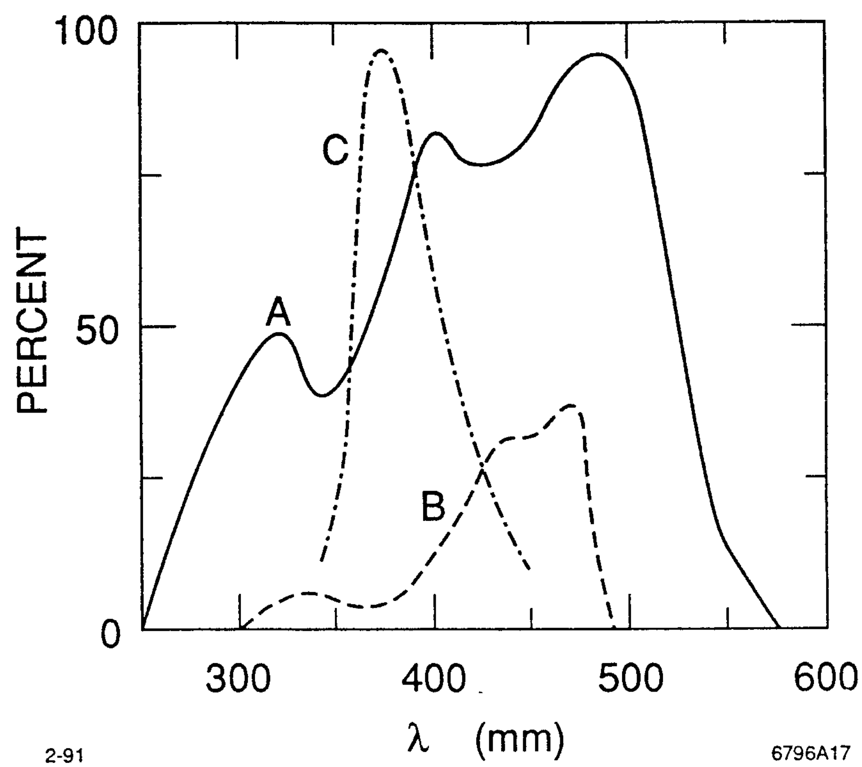


Figure 48. *Absorption spectra of the Y7 fiber and two laser dyes, BASF #84 and #241 in comparison to the NE 111 emission spectrum.*

The Y7-fiber plus SPAD readout was then used in another test, to record Čerenkov photons in a $2.5 \times 2.5 \times 2.5$ cm^3 block of aerogel with a refractive index of n=1.038. The setup is shown in Figure 50. The Čerenkov photons are produced by traversing electrons, which are emitted from a ^{106}Ru source. The aerogel was housed in an light-tight container with white walls coated with POPOP. A plastic scintillator, read out by an XP 2020 PM, was used as a reference. The discriminator threshold was set at 1.8 ± 0.2 MeV, which is the threshold for electrons in aerogel with n=1.038. Čerenkov photons were then counted by requiring a coincidence between the SPAD and the plastic scintillator. In this configuration, a coincidence rate of 3.8% was measured after correction for accidentals and direct hits. In order to build a successful device, the coincidence rate needs to be pushed to nearly 100%. Since in this first test the cell design and readout were far from being optimal, there is much room for improvement, as can be seen in Table 10. The largest gains could be obtained by matching the areas of the fiber and SPAD, increasing the wall reflectivity from 0.85 to > 0.95, increasing the sensitivity of wavelength detection and increasing the dye concentration in an unclad fiber.

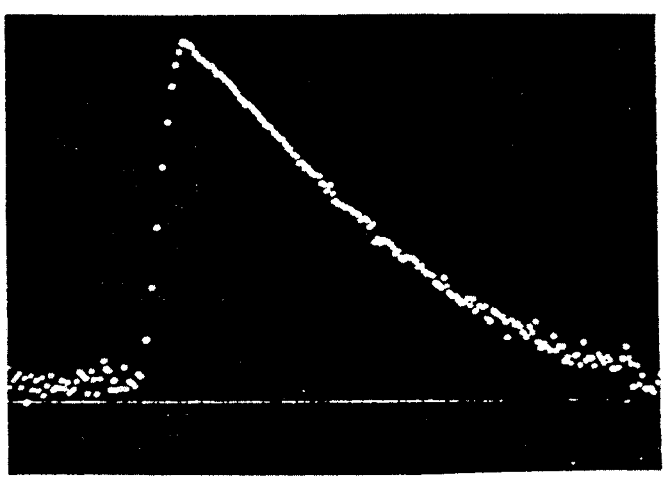


Figure 49. *Fluorescent decay-time spectrum of the Y7 fiber.*

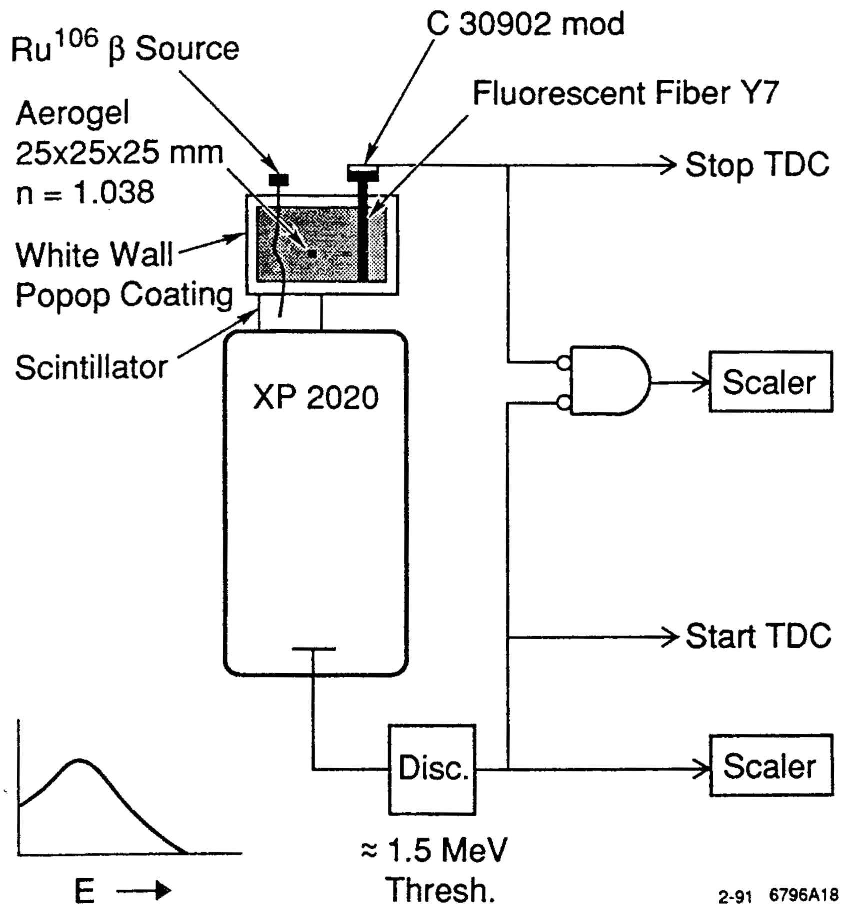


Figure 50. *Test setup for measuring the efficiency of an aerogel cell read out with a Y7 fiber and a SPAD.*

An efficiency of ~ 2% has been measured for FFC/PD readout

In order to measure the wavelength dependence of the efficiency directly for an aerogel cell read out with a fluorescent fiber, E. Lorenz illuminated a $3 \times 3 \times 4$ cm^3 block of aerogel with n=1.032 in a spectrophotometer.[50] The aerogel was housed in a light-tight container with about 90-95% of its walls covered with highly reflective Teflon. The light beam selected by a monochrometer from a Xe lamp was fed in through a 1 cm^2 hole in the container. The aerogel was read out with a 1 mm thick unclad Y7 fiber containing additional BASF dyes #84 and #241,[51] each with a concentration of 400 ppm. The fiber was read out by a 1 cm^2 Hamamatsu photodiode. For normalization, the direct photon flux onto the photodiode without the cell was measured. Figure 51 shows the resulting efficiency distribution. In the visible region the efficiency plateaus at 2.3% and drops rapidly below 370 nm to 0.2% at 250 nm. The main loss is due to light backscattered through the hole, since for lower wavelengths the first scattering occurs closer to the hole. It has been shown that a reduction of the hole size leads to a considerable increase in efficiency at lower wavelengths. Other factors include a higher absorption probability in the aerogel, because both the path length and the absorption cross section become larger, and an increased absorption probability in the Teflon due to a smaller reflectivity. The cutoff at 550 nm is caused by the absorption cutoff of the dyes in the fiber as shown in Figure 48. The efficiency estimate for this setup was 1.3%, which is consistent with the average of the measured distribution. In more recent measurements the efficiency was increased by another factor of 1.5, still leaving room for further improvements.

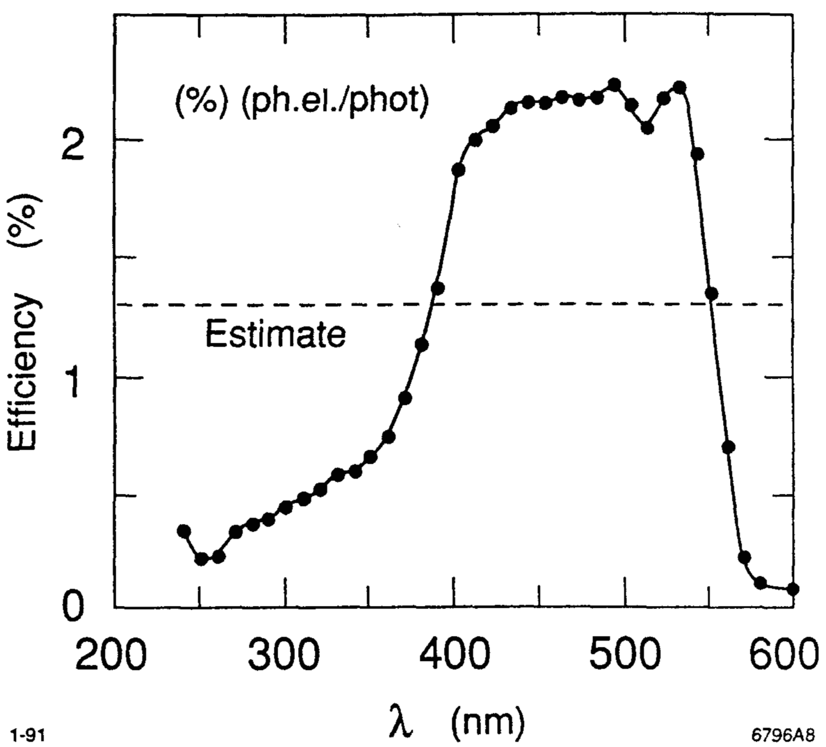


Figure 51. *Efficiency of an aerogel cell read out with a fluorescent fiber and a photodiode as a function of wavelength.*

Table 11. *Improvements for the FFC plus SPAD readout.*

item	present condition	possible improvement	gain
all reflectivity	0.85	> 0.95	2
fiber properties	1 mm clad	1 mm unclad	1.2
fiber dye concentration	none	> 400 ppm	4-6
dye self absorption	0.4	0.5	0.8
absorption in fiber	$350 < \lambda_a < 480$ nm	$240 < \lambda_a < 580$ nm	5
average emission	500 nm	600	1.3-1.7
area fiber/SPAD	0.25	1	4
distance fiber to SPAD	1 mm	0.3	~2
cooling	$-20°$ C	$-40°$ C	1.3
optical coupling	n=1.48	n=1.58	1.1
dye decay time	9.5 ns	< 5 ns	1.3

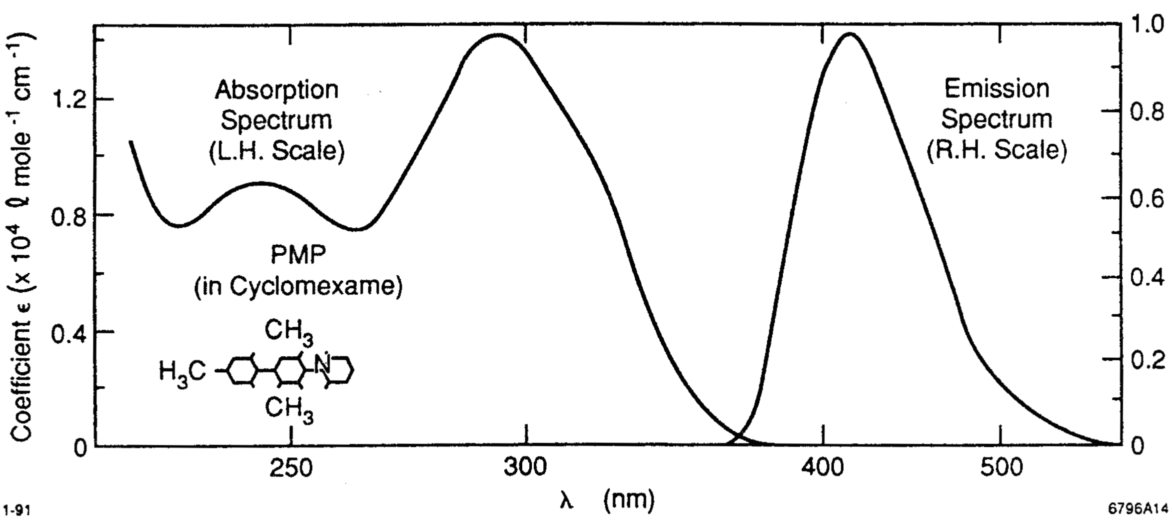
7.4. *Technical Issues*

Aerogel Mixed with Wavelength Shifters

PMP is a suitable wavelength shifter for mixing into aerogel

Mixing a wavelength shifter into the aerogel may yield a considerable efficiency increase, since both the scattering and absorption probabilities are reduced and the wavelengths are more closely matched to the absorption spectrum of the fiber. PMP seems to be a promising candidate for this purpose, since it has a rather large Stokes' shift, it dissolves in alcohol and the PMP molecules fit well into the aerogel structure (see below). POPOP is less suitable, because it has a lower Stokes' shift, does not dissolves in alcohol, which is necessary for aerogel fabrication, and crystallizes in the aerogel. The efficiency is also much lower than for PMP. The absorption and emission spectra of PMP are shown in Figure 52. The absorption range is from 220 to 370 nm with a peak at 290 nm. The shifted light is in the 380-550 nm region with a peak at 410 nm, thus lying in the most sensitive range of photomultiplier cathodes (see Figure 61). The efficiency for this process is high ($> 80\%$). In the 220-330 nm range the extinction coefficient is greater than 7.5×10^3 l mole^{-1} cm^{-1}. For a concentration of 300 ppm of PMP and a silica aerogel of density $\rho = 0.15$ g/cm^3, this yields an attenuation length of $\Lambda < 0.18$ cm. New derivatives of PMP have been developed recently;[52] some of them may actually be even more suited for our purpose than PMP because of even larger Stokes' shifts and higher efficiencies. Figure 53 shows the absorption and emission spectra of PMP derivatives, where PMP 460 is the WLS we are currently using. According to the figure, PMP 440 may yield a higher efficiency, since the emission intensity is largest.

Figure 52. *Absorption and emission spectra of PMP.*



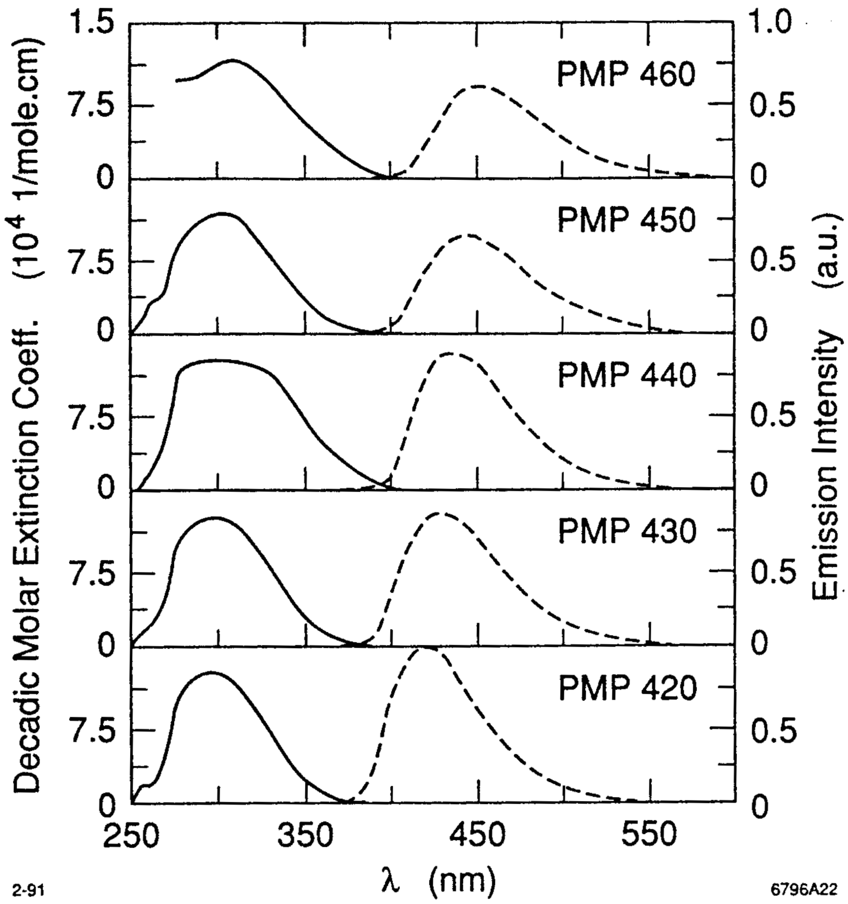


Figure 53. *Absorption and emission spectra of several PMP derivatives.*

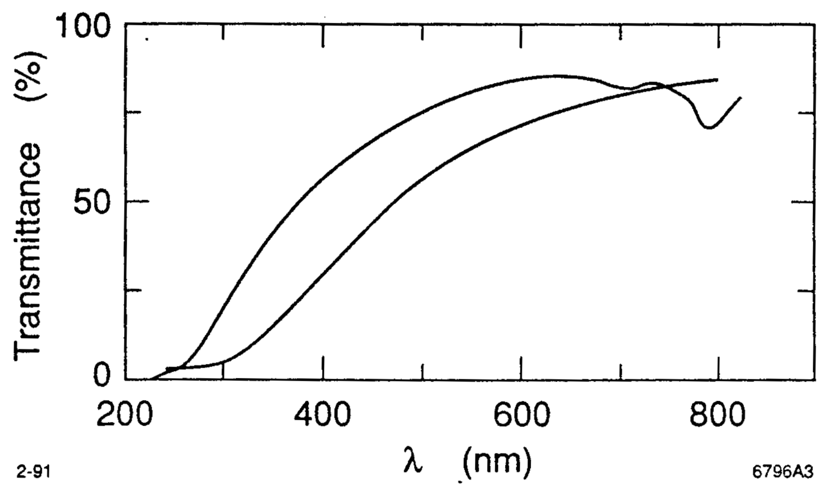


Figure 54. *Transmission spectrum of aerogel mixed with PMP in comparison to pure aerogel.*

The production of aerogel with PMP is currently underway at Lawrence Livermore Laboratory. We have obtained 4 pieces: two 0.5 cm thick discs, 4 cm in diameter, and two 3.4 cm long cylinders, 1.1 cm and 1.2 cm in diameter. Except for the smaller cylinder which has a density of 0.27 g/cm^3 all other pieces have ρ = 0.15 g/cm^3 The refractive indices for the pieces are n=1.057 and n=1.032, respectively. The PMP concentration of each piece is 300 ppm. Figure 54 shows the transmission measured with one of the discs. Comparing this to the transmission of pure aerogel shows that the aerogel doped with PMP is considerably more transparent. A $3 \times 3 \times 6$ cm^3 block of aerogel containing 300 ppm of PMP will be made in the near future.

7.5. Reflectivity of Diffuse Reflectors

It is important to have highly-reflecting container walls, since the average number of reflections is rather large, especially at low wavelengths, where the number of Čerenkov photons is enhanced as $1/\lambda$. The reflectivity has to be greater than 98% in the visible region and greater than 90-95% in the UV region to keep losses minimal. Table 12 lists the reflectivity of several materials in the 250-700 nm range. Figure 55 plots the wavelength dependence of the reflectivity for some white reflectors listed in Table 12.[44] Most of the samples do not meet our requirements, because the reflectivity is poor in the UV-light region or R< 95% in the visible-light region. The Kodak paint seems to be the best candidate, but outgassing of the organic solvent may cause problems. We have obtained a sample of the Kodak paint to perform extensive tests. For comparison, studies with white paints from Bicron and Nuclear Enterprise will be made. Another promising candidate is two layers of 6-SFC millipore paper. This paper will be tested to ascertain whether the reflectivity remains above 90% in the 200-350 nm range. The two teflon samples shown in Figure 55 do not meet our requirements, though Teflon would be ideal, since outgassing and absorption of humidity are not an issue. It may therefore be interesting to investigate other teflon films, since it is likely that some have better reflective properties. Furthermore, if the losses are due to

Kodak white paint is the best diffuse reflector, but it has problems with outgassing

Table 12. *Reflectivity of white diffuse reflectors.*
a. *Kodak paint;*
b. *2 sheets of 6-SFC millipore paper 0.45 HAWP;*
c. *millipore paper 0.45 HAWP;*
d. *$BaSO_4$ paint;* **e.** *Bicron millipore paper 6VHP 0.22;*
f. *Teflon on linol used by MACRO;*
g. *Teflon 11 mils;*
h. *one sheet of 6-SFC millipore paper 0.45 HAWP;*
i. *Floropore FGLP 0.2;*
j. *millipore 0.1 VCWP;*
k. *white paint;*
l. *millipore JVLP 0.1;*
m. *millipore 6VWP 0.22;*
n. *millipore VCWP 0.1;*
o. *2 sheets millipore HAWP 0.45;*
p. *millipore VCWP 0.1*
q. *6SFC Floropore;*
r. *6SFC Duropore;*

λ [nm]	a)	b)	c)	d)	e)	f)	g)	h)	i)
250	95.0	-	37.0	72.5	89.0	4.9	91.5	-	71.0
300	96.8	-	65.0	86.0	92.2	4.5	94.0	-	73.0
350	97.8	95.5	80.5	92.6	94.5	6.0	95.8	95.0	83.0
400	98.6	97.0	92.0	93.8	94.0	32.0	94.2	96.1	83.8
450	99.0	97.8	94.1	94.0	93.2	86.8	92.9	96.3	82.5
500	99.1	98.0	94.8	94.0	92.6	86.0	91.2	96.3	81.0
550	99.2	98.0	95.0	94.2	92.5	85.0	90.0	96.0	80.0
600	99.3	98.0	94.1	93.8	91.5	84.0	87.0	95.8	78.2
650	99.3	97.9	94.0	93.8	91.0	84.5	83.5	95.6	77.0
700	99.3	97.9	94.0	93.2	90.0	86.0	81.0	95.8	71.0
λ [nm]	j)	k)	l)	m)	n)	o)	p)	q)	r)
250	71.0	6.3	40.8	81.0	50.5	36.5	51.5	-	-
300	73.0	6.5	43.5	85.5	80.0	67.0	81.0	-	-
350	83.0	7.0	52.6	90.5	94.5	83.0	95.0	84.6	81.3
400	83.8	42.0	58.2	91.2	95.5	92.5	95.5	82.5	85.2
450	82.5	76.0	65.8	92.0	95.5	95.5	96.0	80.8	88.0
500	81.0	77.0	72.8	92.0	95.4	96.0	95.5	79.1	89.8
550	80.0	77.0	79.2	92.2	95.5	97.0	95.5	78.0	91.0
600	78.2	76.8	83.0	91.9	95.0	96.2	94.8	77.0	91.4
650	77.0	78.0	85.8	91.5	94.0	96.1	94.2	75.7	91.7
700	71.0	77.0	87.0	91.0	93.8	96.0	93.8	75.2	91.8

transmission rather than to absorption, the use of aluminized mylar behind the white reflector may increase the reflectivity. There are, however, some caveats, which may, in the end, disqualify teflon: if good reflectivity is only obtained by increasing the thickness, background levels from Čerenkov photons produced in the teflon layer may become intolerable ($\sim$ 6 photons are produced in the 220-550 nm range per 100 μm thickness).

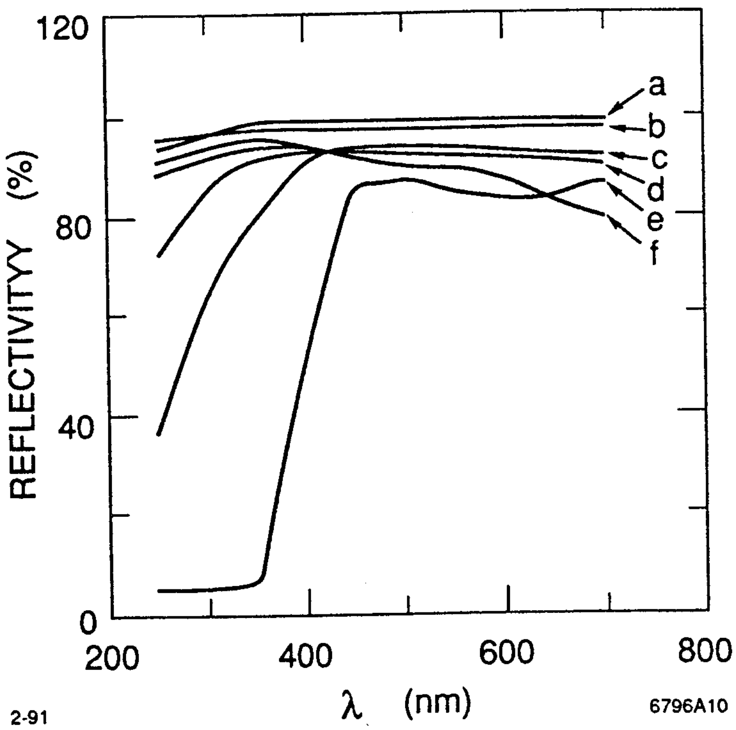


Figure 55. *Reflectivity measurements of white diffuse reflectors as a function of wavelength.*

Fluorescent Fibers

BASF laser dyes are suitable wavelength shifters for the FFC's

In order to keep the efficiency high, the properties of the fluorescent fiber need optimization. The number of Čerenkov photons increases with $1/\lambda$, while the sensitivity of the SPAD peaks at 800 nm, as shown in Figure 56. Thus, the photon wavelengths have to be shifted by several hundred nanometers, which requires several shifts. The first shift to 450 nm is accomplished by PMP mixed in the aerogel. The remaining shifts have to done with several laser dyes, since their typical Stokes' shifts are only 30-50 nm. The quantum yields of fluorescence, however, are typically greater than 95%. In Figure 17 the absorption spectra of the BASF dyes #84 and #241 are shown in comparison of the absorption spectrum of the Y7 fiber. From the figure it is evident that the dye in the Y7 fiber can be omitted if the dyes #84 and #241 are used. But it is important to add one or two more laser dyes, to increase the absorption range above 500 nm. Appropriate dyes are made by BASF.[51] These dyes are mixed into an unclad fiber made of the highly transparent and photochemically stable plastic Polymethylmethacrylate (PMMA). For cell sizes of 2-3 cm, the optimal fiber diameter is about 1 mm, since this leaves the fiducial volume large enough while keeping a moderate average path length. Dye concentrations of 300-400 ppm are probably sufficient for a 1 mm fiber diameter, as the absorption length is estimated to be of the order of a few tenths of a millimeter. A thinner fiber would require higher dye concentrations, which would increase the probability for other absorptions. Though a minimum number of absorptions is necessary to achieve the full shifting, too many absorptions result in an efficiency decrease, since photons are either lost by self-absorption or a new emission angle allows them to leave the fiber. Thus, the transport efficiency may be different in the two cells, since the fiber lengths differ by a ratio of 5 to 2. The fiber end located in the aerogel will have an end reflector made of aluminized mylar.

Single Photon Avalanche Diodes

We are curently investigating two avalanche photodiodes: the C30902S from RCA and the S2383 from Hamamatsu.[53] Their spectral response is shown in Figure 56. In order to obtain the best coupling of the FFC to the SPAD, the FFC-SPAD gap must be very small (< 1 mm). RCA SPADs have a borosilicate window which must be removed and replaced by a fiber-holder, as sketched in Figure 57. To ensure a minimal gap without damaging the surface, a stop, which rests on the non-sensitive SiO_2 surface, defines a constant gap. A cross-section of the RCA SPAD is shown in Figure 58.

The SPAD's must be modified to optimize light collection

A major current deficiency is the mismatch of the diameters of the SPAD and the fiber. Apparently RCA is having problems with fabricating larger area SPADs. Since these problems may be intrinsic to the design, it may be impossible for tem to fabricate SPAD's with a 1 mm diameter. For the Hamamatsu avalanche photodiodes, operation in the Geiger mode, which is crucial for obtaining a high gain of $> 10^8$,

A mismatch between the diameters of the fiber and the SPAD causes loss of efficiency

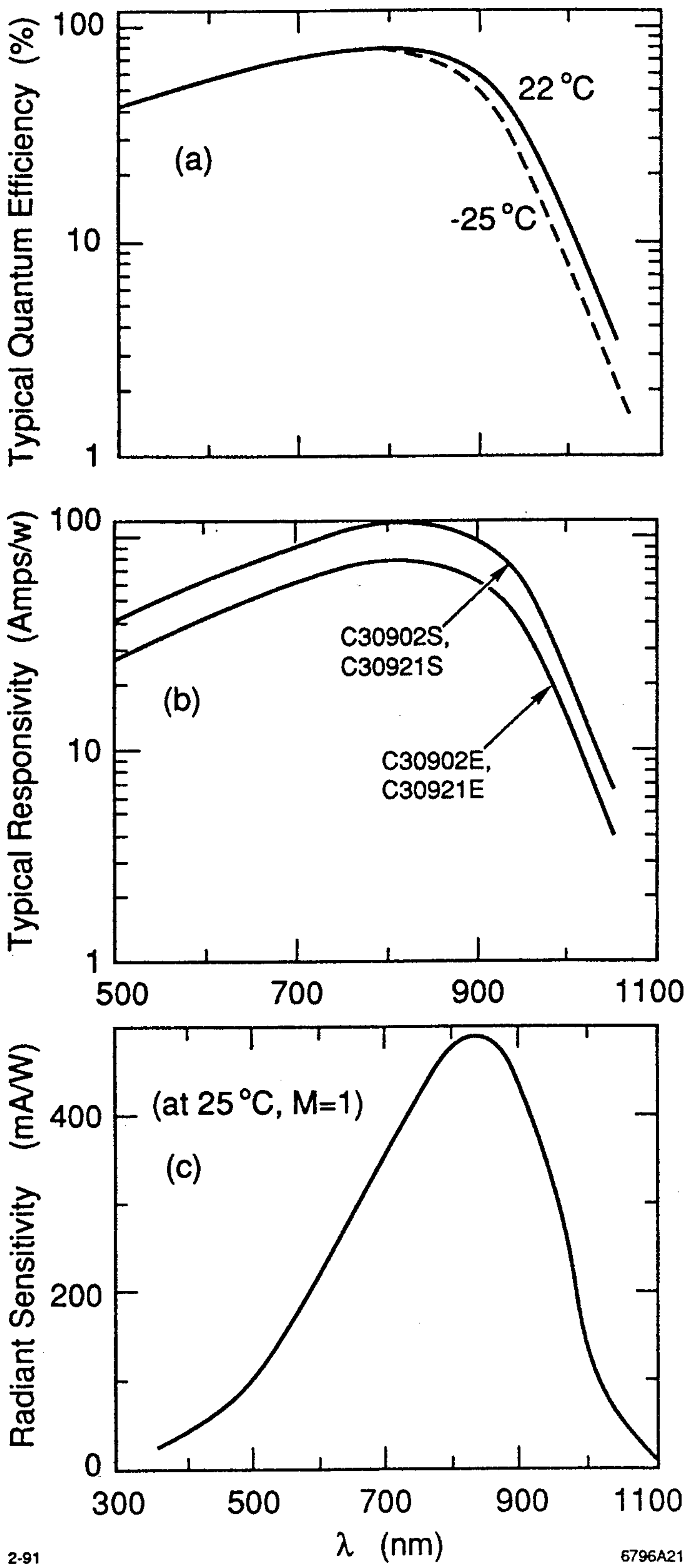


Figure 56. *Spectral sensitivities of single photon avalanche diodes.*

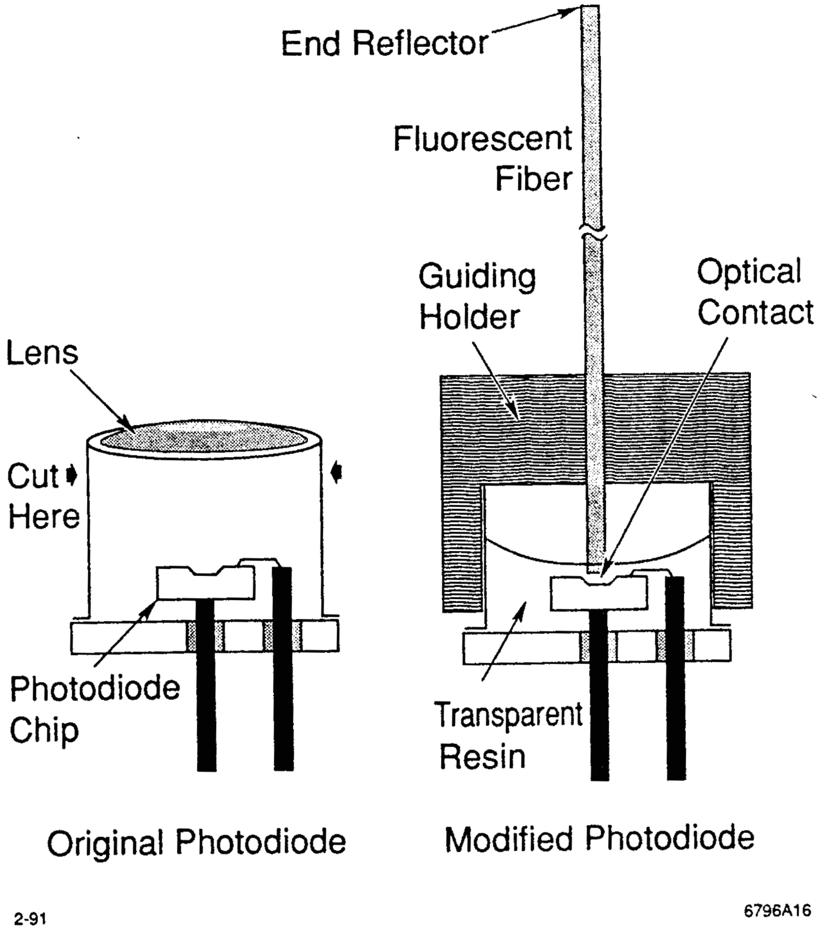


Figure 57. *Mounting of the fiber to the avalanche diode.*

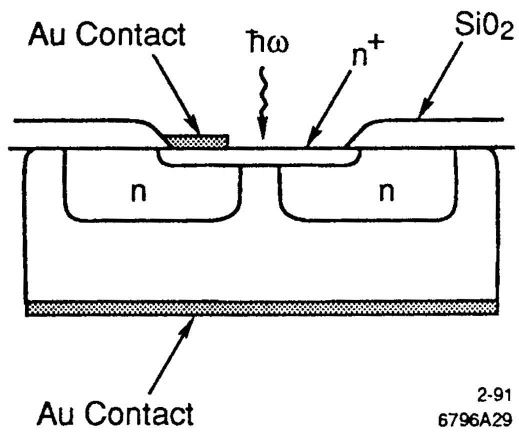


Figure 58. *Cross section of a single photon avalanche diode.*

remains to be demonstrated. If, however, a 1 mm SPAD operating in the Geiger mode is not commercially available, one has two options: one can either develop a focussing system or use four 0.5 mm thick fibers. For both cases, efficiency losses are inevitable; the second option also increases costs.

The SPADs work best if they are cooled to −20°C to −40°C. The noise rate drops from ∼12 kHz at room temperature to ∼ 500 Hz at T= −20° C, as shown in Figure 59a. Thus, for optimal performance, a cooling system is required. It is certainly not as involved as a cryogenic system, but proper insulation is needed, since moisture will condense near the cooling lines. Though the noise level of the RCA SPADs was tolerable at room temperature, one may in fact use cheaper SPADs in the real detector, which may be too noisy at 20°C. The coincidence technique does not require a cooling system. It will be necessary to demonstrate that cross-talk between the fibers is negligible. It may, in any case, be necessary to keep the humidity at a low level in order to avoid damage to the aerogel. This is achieved with a steady flow of nitrogen, which may be cooled to −20°C.

Cooling of RCA SPAD's considerably reduces noise

Improvements by Hamamatsu may allow counting of all Čerenkov photons

Due to the long deadtimes (> 300 ns) required to quench an avalanche, the RCA SPAD's can detect only a single photon. If enough Čerenkov photons are present and noise levels are low, the detection of one photoelectron is sufficient for particle identification. However, it would be helpful to have some redundancy by observing more than one photoelectron. This would be possible if the dead time is reduced to a few hundred ps. In the RCA SPADs the long dead time is needed to reduce the probability for afterpulsing. Figure 59b shows the probability for afterpulsing in the 100-300 ns range. The curve indicates that a minimum dead time must be 300 ns. Hamamatsu, however, claims that it may be possible to achieve dead times less than 1 ns.[54] Hamamatsu further claims that their avalanche photodiodes show a lower after-pulsing rate at room temperature than when cooled, while noise levels at room temperature are lower than for the RCA SPAD's. Thus the Hamamatsu PDs may be the ideal photon detectors; it is expected that they will soon become available.

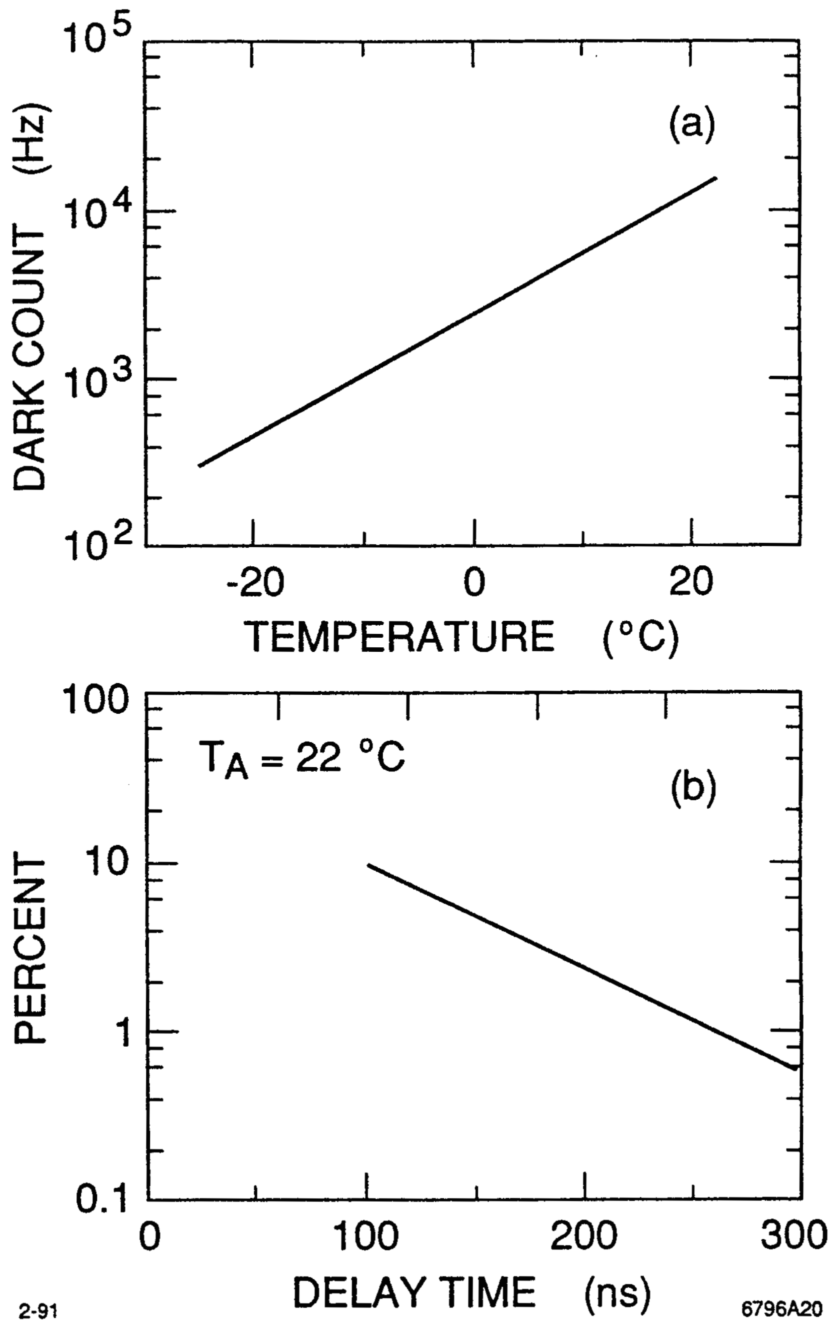


Figure 59. **a.** *Dark current as a function of temperature;* **b.** *probability of after pulsing as a function of dead time.*

Phototetrodes

A phototetrode (PT) is a 4 stage photomultiplier, which is can operate in magnetic fields up to 1 Tesla. The gain at nominal operation of 1 kV is 50 at zero field; it is 12.5 at 1 T. The gain is largest if the PT is oriented nearly parallel to the $\vec{B}$ field, while it drops to zero if the PT is perpendicular to $\vec{B}$. The gain as a function of the direction of the PT is shown in Figure 60. For a 30° angle, the gain is 75% of the maximum

Phototetrodes must be operated at angles of less than 45° to the magnetic field

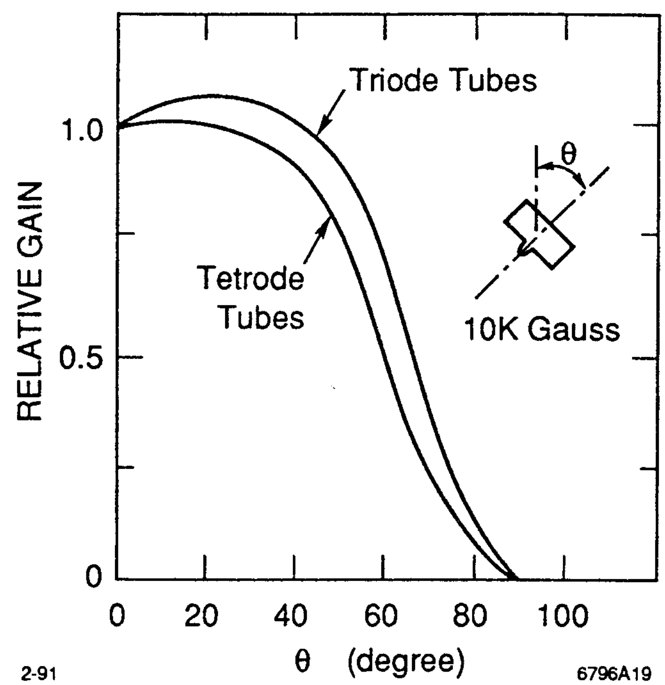


Figure 60. *Relative gain of a phototetrode versus its angle with the magnetic field.*

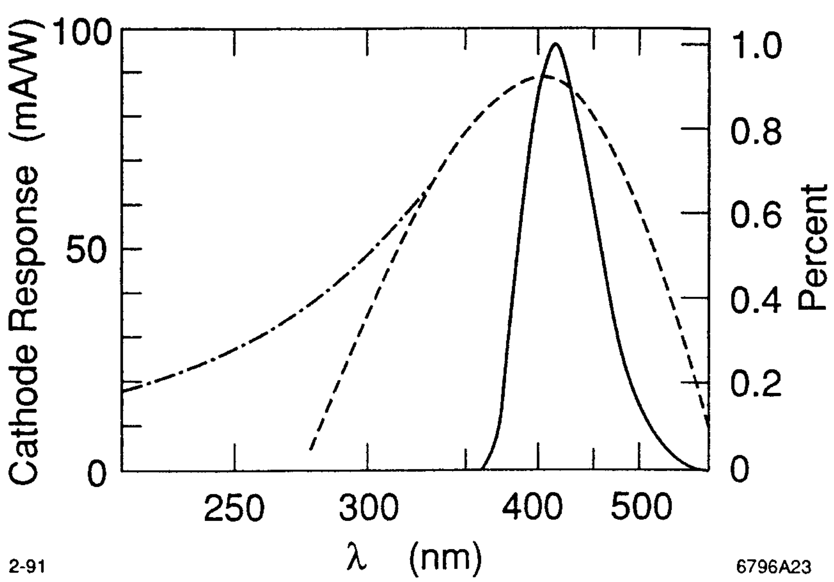


Figure 61. *Quantum efficiency of a multialkaline cathode compared to the PMP emission spectrum.*

gain. Since the gain variations are still relatively small at 30°, this seems to be a suitable operation point. Operation at 45° is desirable, even though the gain is still 0.5, as the variation of gain with angle is very large. These PTs are obtainable with a multialkali photocathode. Figure 61 shows the quantum efficiency of a multialkali cathode for a borosilicate and a quartz window in comparison to the PMP emission spectrum. If PMP is mixed into the aerogel, a multialkali cathode is important but a quartz window is not essential.

An Estimate of Inactive Material

The inactive material for a two-cell aerogel threshold Čerenkov detector amounts to 6% of a radiation length

In order to estimate the inactive material in front of the electromagnetic calorimeter resulting from an aerogel Čerenkov counter, we have used the cell design described in Section 3. Table 13 gives a detailed list of the individual components. We have assumed that the mechanical support structure consists of two layers of carbon fiber. Thus, the total inactive material amounts to about 6% of a radiation length. The estimate for white paint is based on $BaSO_4$. Using Teflon or millipore paper as a reflector reduces the inactive material by about 0.5%. The overall contribution from the fiber and SPAD is negligible. For the tetrode readout the inactive material increases by 1-2%, depending on the mechanical support structure. For both designs, this is considerably less inactive material than needed for a CRID/RICH particle identification system.

Table 13. *Inactive material of a two-cell aerogel threshold Čerenkov counter.*

Material	thickness [cm]	density [g/cm^3]	X_0 [g/cm^2]	rad. length [%]
aerogel 1.008	10	0.038	29.85	1.27
aerogel 1.06	4	0.286	29.85	3.83
mylar	4×0.005	1.39	39.95	0.07
cardboard	4×0.015	0.6	40.0	0.09
white paint	4×0.0025	4.5	9.91	0.45
mech. support	2×0.1	1.2	40.	0.6

7.6. Performance of Threshold Čerenkov Counters

In concluding this section, we will demonstrate that a two-cell threshold Čerenkov counter with refractive indices of $n = 1.06$ and $n = 1.008$ provide sufficient particle identification to achieve the physics aims of an asymmetric B Factory. The probability for observing a single photoelectron from a sample of Čerenkov photons produced by a charged particle of mass m traversing an aerogel cell with a momentum p above threshold ($\beta > \beta_t = 1/n$) is given by:

$$P_e = 1 - \exp(-\overline{N}_e) \tag{13}$$

where $\overline{N}_e$ is the average number of photoelectrons, which is determined from the maximal number of photoelectrons, N_{max}, expected at $\beta = 1$, by:

$$\overline{N}_e = N_{max}\left(1 - \frac{\beta_t\gamma_t}{\beta\gamma}\right) \tag{14}$$

For the two cells, N_{max} is assumed to be 10 and 5 photoelectrons, respectively.

Figure 62 plots the probability[55] for detecting electrons, muons, pions, kaons and protons as a function of momentum. Solid curves correspond to the $n = 1.06$ cell while dashed curves correspond to the $n = 1.008$ cell. False identifications due to noise and δ-rays produced by particles below threshold are neglected in this estimate. From our previous discussion we know that noise is small. The production of δ-rays is expected to be also small, but needs further study. If one requires an identification probability of $> 50\%$, K/π separation is provided in the 0.4-4.2 GeV/c momentum range;

For an identification probability above 50%, aerogel cells with indices of n=1.06 and 1.008 provide K/π separation between 0.4 and 4.2 GeV/c

For π/e and π/μ separation as well as K/p separation in the 1-1.4 GeV/c region and above 2.7 GeV/c, one needs additional information: lower-momentum electrons are identified by dE/dx, while higher-momentum electrons can be identified by pattern recognition in the electromagnetic calorimeter; for muon identification a hadronic calorimeter or a muon range out system has to be used; antiprotons leave a clear signature in the electromagnetic calorimeter. A residual problem for detecting protons above 2.7 GeV/c remains unsolved.

The performance of a two-cell aerogel threshold Čerenkov counter has also been calculated for typical $B\bar{B}$ events. An example is shown in Figure 63.[56] Both the B and $\bar{B}$ decay semileptonically. The final state consists of $\nu_\mu \mu^+ \pi^- K^+ \pi^-$, $\bar{\nu}_e e^- \pi^+ K^- \pi^+$.

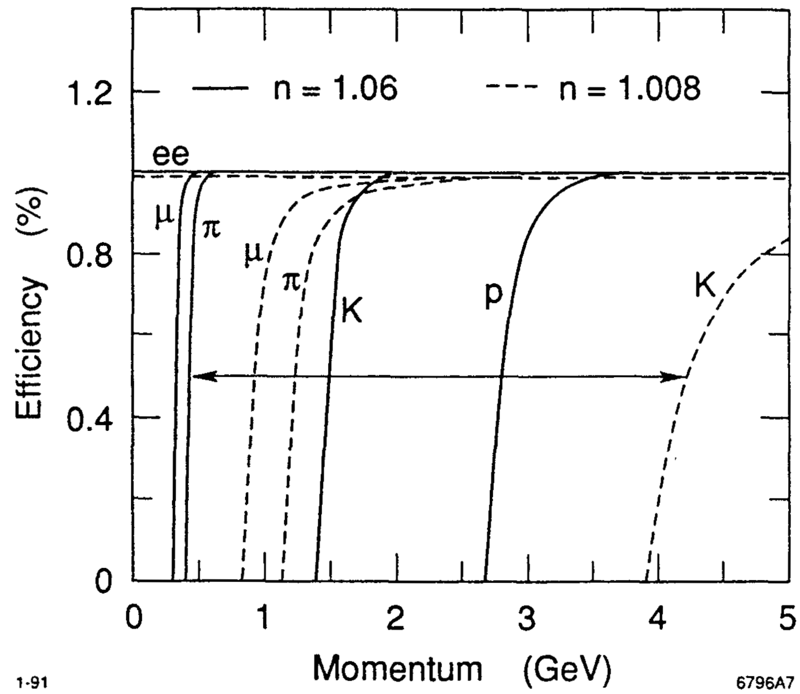


Figure 62. *Probability of detecting particles in a two-cell threshold Čerenkov counter detector.*

A two-cell aerogel Čerenkov detector, supplemented by dE/dx measurement in the drift chamber, provides excellent particle identification for an asymmetric B Factory

Figure 64 shows the performance of a dE/dx system, a TOF system, a CRID system and two-cell threshold Čerenkov counters to identify all charged particles in the above event. The dE/dx simulation was based on 32 one cm samples taken in the drift chamber operated with a He/isobutane/CO_2 gas at atmospheric pressure. The resolution was assumed to be 7.5% and a plateau to minimum ratio of 1.4. For the TOF system, a barrel and two endcaps of 5 cm thick scintillation counters have been assumed, with a resolution of 75 ps and an efficiency of 100%. The CRID simulation is based on 1 cm C_6F_{12} liquid radiator ($n = 1.277$) assuming 36.6 observed photons out of 61 produced. The CRID angle is assumed to be measured with 100% efficiency. The simulation of a two-cell threshold Čerenkov counter system is based on the parameters given above.

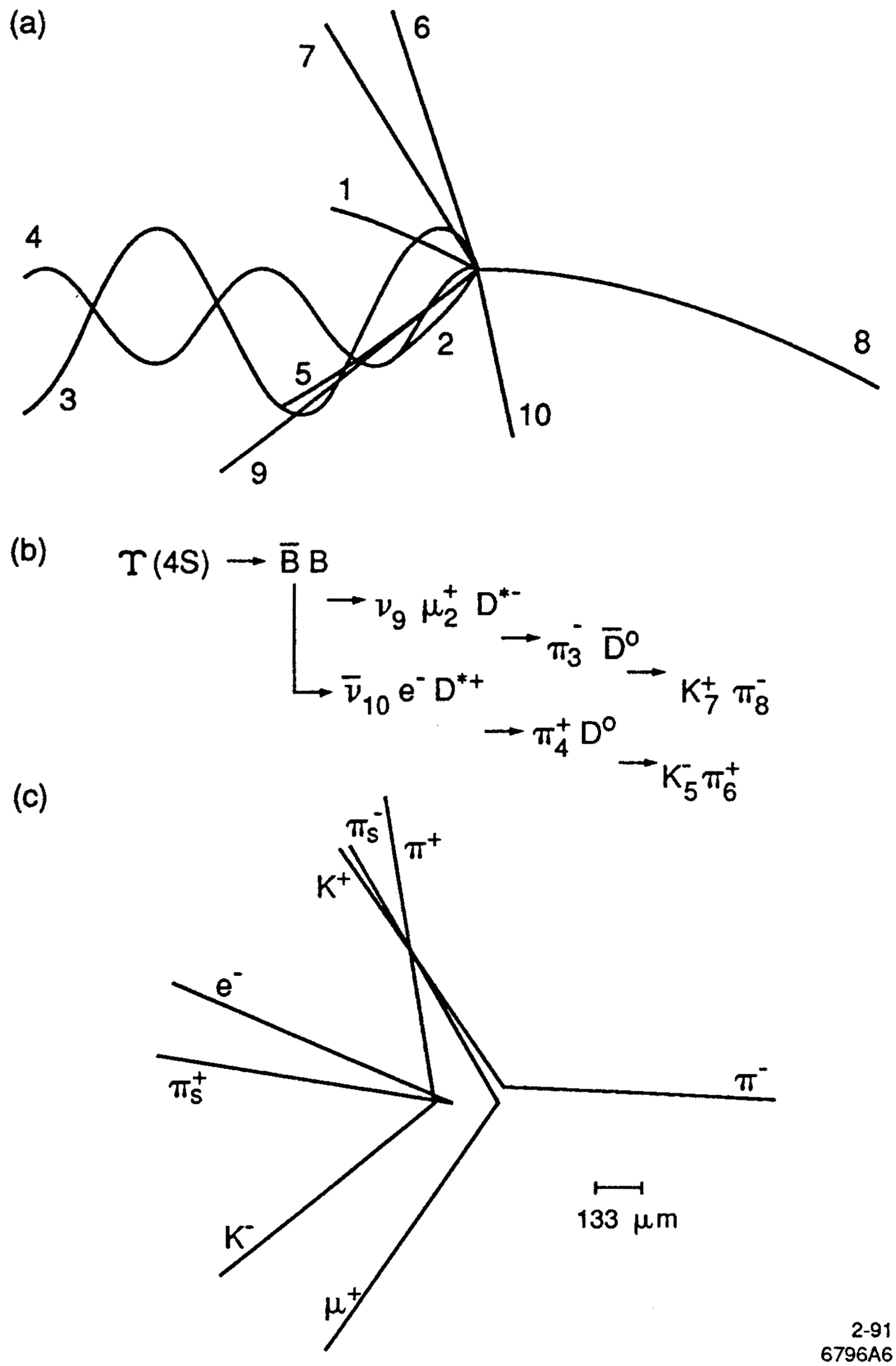


Figure 63. *A $B\bar{B}$ decay of the $\Upsilon(4S)$.*

Particles 1,3,4 are well-identified by dE/dx. Both the CRID and the two-cell threshold Čerenkov counter do a good job in identifying the remaining particles, while TOF identification is limited to momenta below 1.5 GeV/c and thus cannot classify all of the remaining particles. A dE/dx system and a two-cell threshold Čerenkov counter provide good K/π separation up to 4.2 GeV.

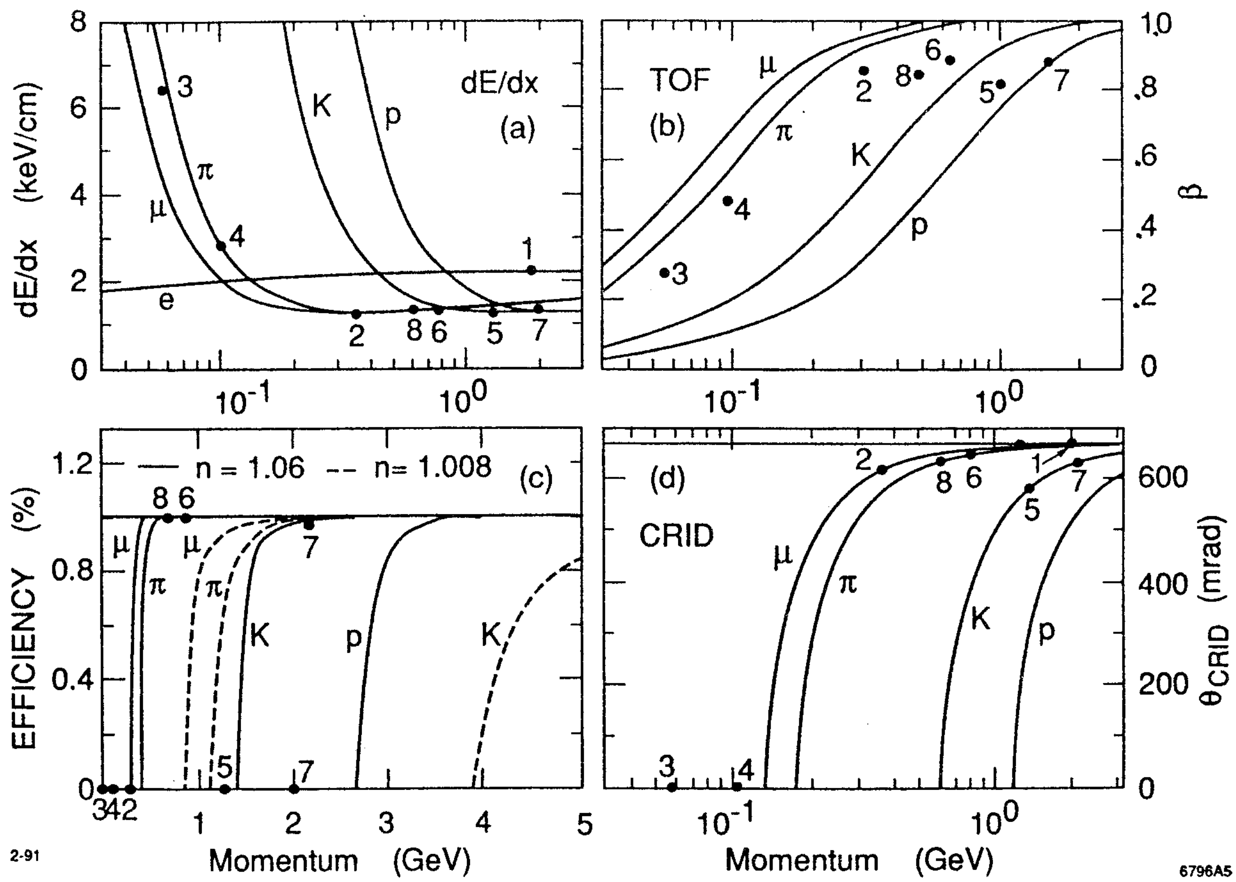
7.7. Conclusions and Outlook

Aerogel threshold counters may provide an interesting alternative to ring-imaging devices

In conclusion, threshold Čerenkov counters using two cells of aerogel with refractive indices of $n = 1.06$ and $n = 1.008$ provide a potential alternative to CRID/RICH counters. The inactive material for the entire device is only a few percent of a radiation length, thus leaving the good energy resolution in the electromagnetic calorimeter unaffected. Though the initial test results look very encouraging, a lot of work remains to be done before this idea can be turned into a large scale detector system. The most important issues are whether one actually can detect enough photons, whether the readout is sufficiently reliable and efficient to build a particle identification system and whether aging effects due to radiation, humidity and absorption of organic molecules are serious. Our main goal for the near future is to investigate each of these issues. Specific R&D items include:

- Both the absorption and scattering length have to be measured in different samples, examining the dependence on ρ, monitoring time dependence and studying the transmittance after exposure to various paints, wavelength shifters and high doses of radiation. It is important to obtain more understanding of which precautions are necessary for the construction of a large detector.

- The properties of aerogel mixed with PMP have to be studied focussing on questions like: does PMP scintillate? What is the optimal PMP concentration? Does radiation change the PMP bounds and therefore damage the aerogel? Are there other long-term aging effects? Is the PMP concentration uniformly distributed? In order to find the best radiator, aerogel containing other PMP derivates and the wavelength shifter HF3 will also be made and tested.



- In order to optimize the efficiency, it is necessary to conduct various material studies: i) measure the flux concentration in the fiber by testing different laser dyes; ii) optimize the coupling of the fiber and the SPAD; iii) measure the reflectivity of white diffuse reflectors; iv) find the optimal container; v) test different SPADs and PT.

- Both the photoelectron yield and efficiency have to be measured in several $n = 1.032$ aerogel cells with and without PMP. For readout a UV sensitive PM, a PT and a FFC plus SPAD will be used, for a direct comparison.

- Once the best readout is found, a test array for $n = 1.008$ aerogel will be built and tested. If this test is successful, we will build a two-cell prototype array consisting of $n = 1.008$ and $n = 1.06$ cells and do a beam test using pions and kaons.

Figure 64. *Performance of* **a.** *dE/dx,* **b.** *TOF,* **c.** *CRID/RICH, and* **d.** *a two-cell threshold Čerenkov counter detector for identifying the particles of the decay in Figure 63.*

We have established a collaboration with L. Hrubesh of Lawrence Livermore Laboratory, which will emphasize the production and testing of aerogel incorporating various wavelength shifters.

ACKNOWLEDGEMENTS

L. Hrubesh provided us with numerous smaples of aerogel. We would like to thank L. Hrubesh, E. Lorenz, S. Schindler and A. Vorobiov for fruitful discussions. We also wish to thank A. Weinstein for generating several plots and J. Chiu for helping with some of the aerogel measurements.

References

1. For a discussion of these final states, see SLAC–353, "The Physics Program of a High Luminosity Asymmetric B Factory at SLAC," p. 133–141, and F. Le Diberder, *Ba$\bar{B}$ar Note* # 33, "Background Study of the $B \to \pi^+\pi^-$ Channel."
2. R. Aleksan, *et al.*, *Phys. Rev.* **D39**, 1283 (1989).
3. See, for example, W. B. Atwood, in *Proceedings of the Summer Institute on Particle Physics,* (1980) SLAC–Report–239.
4. R. Stroynowski, in *Proceedings of the Tau-Charm Factory Workshop,* May 23–27, 1989, SLAC, Stanford, CA, SLAC-343 (1989).
5. See, for example, S. E. Willis, in *Proceedings of the Symposium on Particle Identification at High Luminosity Colliders,* (1989) Fermilab, Batavia, IL, April 5–7, 1989.
6. See, for example, S. Banerjee, *et al.*, *Nucl. Instr. and Meth.* **A269**, 121 (1988).
7. Properties of most of these devices are discussed in "Major Detectors in High Energy Physics," Particle Data Group, LBL–91 Supplement UC–34D (1985).
8. I. Lehraus, *Nucl. Instr. and Meth.* **217**, 43 (1983).
9. W. Alison and P. Wright, *Exp. Tech., in High Energy Physics,* Addison-Wesley Publ., 1987; G. Lynch, LBL–TPC–81.
10. J. Va'vra, *et al., Nucl. Instr. and Meth.* **203**, 109 (1982); J. Va'vra, Contribution to *Proceedings of the Tau-Charm Factory Workshop,* May 23–27, 1989, SLAC–343.
11. The dependence obtained empirically by fitting argon data presented in I. Lehraus, R. Matthewson, and W. Tejessy, *Phys. Scr.* **23**, 727 (1981).
12. See, for example, J. V. Jelly, "Čerenkov Radiation and its Applications," (1958) Permagon Press.

13. J. Seguinot and T. Ypsilantis, Ring Imaging Čerenkov Detectors (1990), manuscript in preparation.

14. J. Seguinot and T. Ypsilantis, *Nucl. Instr. Meth.* **142**, 377 (1977).

15. SLD Collaboration, SLD Design Report, SLAC–REP–273 (1984).

16. T. Ypsilantis, Working Group on Particle ID at the Workshop on Physics and Detector issues for a High Luminosity Asymmetric B Factory, June 7, 1990, SLAC; see also T. Ypsilantis, $Ba\bar{B}ar$ Note #44.

17. J. Sequinot, *et al.*, CERN EP–90–88, June 1990.

18. CERN/LEPC/83–3 (1983), CERN/LEPC/P2 (1983).

19. R. Arnold, *et al.*, *Nucl. Instr. Meth.* **A273**, 466 (1988).

20. V. Ashford, *et al.*, *Proceedings of the XXIII International Conference on High Energy Physics,* 1470 (1986).

21. R. Arnold, *et al.*, *Nucl. Instr. Meth.* **A270**, 255 (1988).

22. CERN/LEPC/84–15 (1984).

23. T. Ypsilantis, *Phys. Scr.*, **23**, 371 (1981).

24. B. Ratcliff, in $Ba\bar{B}ar$ Note #14.

25. D. Williams, Notes from B Factory Particle ID Working Group, March 9, 1990, SLAC; see also, B. Ratcliff, in $Ba\bar{B}ar$ Note #30.

26. P. Coyle, Notes from B Factory Particle ID Working Group, April 20, 1990, (Nevis meeting).

27. See also, plenary talk on Particle ID, B. Ratcliff, in $Ba\bar{B}ar$ Note #37.

28. See, for example, K. Abe, *et al.*, SLAC–PUB–5214.

29. J. Va'vra, SLAC-PUB-5207 and SLAC-PUB-4116.

30. J. Va'vra, Notes from B Factory Particle ID Working Group, April 20, 1990, (Nevis meeting).

31. B. Ratcliff, Notes from B Factory Particle ID Working Group, April 20, 1990, (Nevis meeting).

32. J. Va'vra, Notes from B Factory Particle ID Working Group, June 6, 1990, SLAC.

33. See also, plenary talk on Particle ID, B. Ratcliff, in *Ba$\bar{B}$ar* Note #44.

34. B. Ratcliff, Notes from *B* Factory Particle ID Working Group, June 6, 1990, SLAC.

35. R. Arnold, *et al.*, *Design Proposal for a B Factory Detector at PSI.*

36. J. Seguinot, *Proceedings of the Symposium of Particle Identification at High Luminosity Hadron Colliders*, April 1989, Fermilab, Batavia, IL; T. Ypsilantis, *ibid.*

37. P. Coyle, Working Group on Particle ID at the Workshop on Physics and Detector Issues for a High Luminosity Asymmetric *B* Factory, June 6, 1990, SLAC.

38. E. Lorenz, MPI-PAE/Exp. El. 184, preprint, (1987); R. A. Eichler, *et al.*, Proposal for a *B* Meson Factory at PSI, PSI-PR-88, 1988.

39. S.S. Kistler, *Nature* **127**, 741 (1931); S.S. Kistler, *J. Phys. Chem.* **34**, 52 (1932).

40. J. Fricke, *Aerogels*, Springer Proceedings in Physics **6**, (Springer Verlag, Berlin).

41. T.M. Tillotson, L.W. Hrubesh and I.M. Thomas, *Mat. Res. Soc. Symp. Proc.* **121**, 685 (1988); T.M. Tillotson and L.W. Hrubesh, UCRL-Ext.Abs 102517, preprint, (1990); L.W. Hrubesh, T.M. Tillotson and J.F. Poco, UCRL-Ext.Abs. 102518 preprint, LLL, (1990).

42. L.W. Hrubesh, private communication.

43. L.W. Hrubesh, UCRL-53794 preprint, LLL, (1987); L.W. Hrubesh and C.T. Alviso, *Mat. Res. Soc. Symp. Proc.* **121**, 703 (1988).

44. S. Schindler, private communication.

45. I. Holl, E. Lorenz and G. Mageras, MPI-PAE/Exp. El. 224, preprint, (1990).

46. W.J. Willis and V. Radeka, *Nucl. Instr. Meth.* **120**, 221 (1974); V. Radeka and E. Rescia, *Nucl. Instr. Meth.* **A265**, 228 (1988).

47. We have set the coefficients for Rayleigh scattering and for absorption to $a = 1.6 \times 10^{-18}$ cm^3 and $b = 1.6 \times 10^{-11}$ cm, which correspond to an average scattering length $\Lambda_s = 2.3$cm and an average absorption length $\Lambda_a = 100$cm in the 300-600 nm wavelength region.The reflectivity of Kodak paint has been used (see Section 7.5).
48. H.P. Lüttenberg, Bonn-IR-88-51 preprint, Bonn PhD thesis (unpublished), (1989).
49. A. Vorobiov, private communication.
50. E. Lorenz, private communication.
51. Seybold and Wagenblast, *Dyes and Pigments* **11**, 303 (1989).
52. C. Ambrosio, *et al.*, CERN-PPE-90.6, preprint, (1990).
53. K. Yamamoto, private communication.
54. K. Yamamoto *et al.*, *Nucl. Instr. Meth.* **A253**, 542 (1987).
55. A. Weinstein, private communication.
56. The Physics Program of a High-Luminosity Asymmetric B Factory, SLAC-353, (1989).

Report of the Electromagnetic Calorimetry Group

E.D. Bloom, F. Bulos, G. Eigen, A. Eisner,
P. Franzini, G. Godfrey, D. Hitlin,
J. Lee-Franzini, M. Mandelkern, H. Marsiske
and P.M. Tuts

1. Introduction

The B Factory detector will include state-of-the-art electromagnetic calorimetry with target energy resolution

$$\frac{\sigma_E}{E} = \frac{(1-2)\%}{\sqrt[4]{E[GeV]}} \oplus 1\%.$$

Resolution of this quality can be obtained only with homogenous active media such as scintillating crystals, scintillating glass or noble liquids; we have not, therefore, considered calorimeter configurations based on sampling. We will assume that the calorimeter begins outside the central tracking chamber and particle identification system at a radius of 1.0 meter, and that it extends radially outward for a minimum of 18 radiation lengths. Access to the central tracking chamber requires the calorimeter configuration to be a barrel with endcaps. As discussed in SLAC-353, an angular resolution of 10-20 mr is required in order to retain the intrinsic energy resolution by removing Doppler shifts caused by the boosted center of mass of the asymmetric collision. This is particularly important for the study of electromagnetic transitions in the Υ system. The geometrical arrangement should thus consist of projective towers pointing at the interaction point. The transverse size of all towers is made identical. This makes the sharing of shower energy between towers independent of the location within the calorimeter. The position resolution at a given photon energy is thus independent of polar angle.

2. General Considerations

2.1. Segmentation

The desired angular resolution can be obtained by partioning the entire 4π of solid angle into approximately 10,000 elements, providing an angular resolution of $\sigma_{\theta,\phi} \approx 10$ mrad. For more than 10^4 elements, the cost of electronics (assuming \$100 -\$300 per channel) becomes an unreasonably large fraction of the total calorimeter cost of approximately \$25 M. This number of elements is also well-matched to the area covered by a typical electromagnetic shower at the one meter radius of the calorimeter. For a material with a Moliere radius of 5 cm, there are 8042 (Moliere radius)^2 elements in the barrel and 1256 elements in each end cap. For comparison, this is similar to the number of CsI(Tl) crystals in CLEO II (8000), and is about 14 times the number of NaI elements in the Crystal Ball (720).[1]

2.2. Active Materials

Table 1 lists the properties of a variety of possible calorimeter materials. Of the crystal scintillators CsI(Tl) appears to be the most desirable: it has the lowest cost (for a calorimeter with inner radius of one meter), it is only weakly hygroscopic, it yields the largest number of photons/MeV, and there is demonstrated industrial capacity to produce the required number of large, clear crystals. BGO would have cost advantages if the calorimeter were built closer to the interaction point (and was then scaled down in proportion to its radiation length), but would cost more than CsI(Tl) under the constraint that the calorimeter begins at a radius of one meter. No clear 18 radiation length-long crystals of BaF_2 have been manufactured to date. NaI(Tl) has a longer radiation length, is more fragile, and is more hydroscopic than CsI(Tl). Scintillating glass was also considered. Compared to CsI(Tl), it costs more per cubic radiation length, makes a much thicker calorimeter, and has an order-of-magnitude lower light output (thus requiring a higher gain photomultiplier rather than a phototetrode or photodiode in order to operate in solenoidal magnetic field). Cryogenic liquid total absorption (non-sampling) calorimeter was also considered in some detail. Liquid krypton is the medium of choice for such a detector. Liquid xenon has a shorter radiation length, but is too expensive; liquid argon has too long a radiation length. The group therefore narrowed its considerations to a CsI(Tl) crystal calorimeter and a homogeneous liquid krypton calorimeter.

Parameter	BGO	CsI(Tl)	BaF_2	NaI(Tl)	Scint Glass	Liquid krypton
Density[g/cm^3]	7.13	4.5	4.86	3.7	4.8	2.41
Radiation length [cm]	1.13	1.86	2.05	2.6	4.35	4.7
Had abs length [cm]	21.9	34	29.9	41.1	44	60.1
Critical Energy [MeV]	8.8	10.2	12	10.7		16.2
Moliere radius [cm]	2.7	3.8	4.3	5.1		6.1
Min I dE/dX [MeV/cm]	8.07	5.1	5.72	4.13		3.44
$< \lambda_{scint} >$ [nm]	400	550	325	410	440	
Decay time constant[ns]	300	900	620	230	70	3μs/cm
Photons/MeV	8,000	50,000	10,000	40,000	300	
Photoelectrons/MeV		2500			10	
Hydroscopic?	No	Weakly	No	Very	No	No
Mechanical Stability	Good	Good	Cleaves	Cleaves	Good	Liquid
Radiation resistance	Fair	Fair	Good	Fair	Good	Excellent
Main Vol [m^3]	4.5	7.9	8.8	11.6	21.9	24.2
Cost [$/cm^3]	7-13	2-3.5	3-5	2	.25-.5	.4
Total Cost [$M]	31-58	16-28	26-44	23	5-11	10

Table 1. *Properties of candidate radiators. The main volume is taken as a cylinder 3.2 meters long, with inner radius 1.0 meter, and thickness 18 radiation lengths*

2.3. Material in Front of the Calorimeter

The physics program of the B Factory calls for a detector with excellent electromagnetic calorimetry as well as excellent particle identification. Choices for the electromagnetic calorimeter have been narrowed to either liquid krypton or CsI(Tl); particle identification options include a CRID, aerogel or high-pressure Čerenkov counters.

A major consideration in any choice of particle identification system is the amount of material it introduces in front of the electromagnetic calorimeter; this must include the material in the cryostat in the case of a liquid krypton calorimeter as well as the material in the particle identification system. To evaluate the effect on this material on the calorimeter energy resolution we have performed a GEANT-based Monte Carlo study which assumes between 5% and 40% of a radiation length (*r.l.*) of material right in front of the calorimeter. This represents any part of the particle identification system near the calorimeter, as well as a possible cryostat; the material in the beam pipe, the vertex detector and in the drift chamber has been neglected.

Samples of photons with three different energies (E_γ = 10, 100, and 1000 MeV) and two different polar angles (θ = 90° and 35°) have been generated. (Note that for the smaller polar angle the amount of material seen by a particle is increased by 74%. Figure 1 shows for some selected cases the distribution of the energy deposited by photons interacting in the material placed in front of the sensitive volume of the calorimeter.

Note that the distributions are non-gaussian. This makes it impossible to estimate the effect on the energy resolution in the usual way, that is by adding in quadrature the properly weighted sigmas of the two distributions. Note also that a significant fraction of low-energy photons are fully absorbed in the front material. This has a serious impact on the ability to reconstruct final states with multiple π^0's. Table 2 shows the fraction of photons interacting in the front material as a function of photon energy and angle, as well as the mean and *rms* energy deposited in the front material.

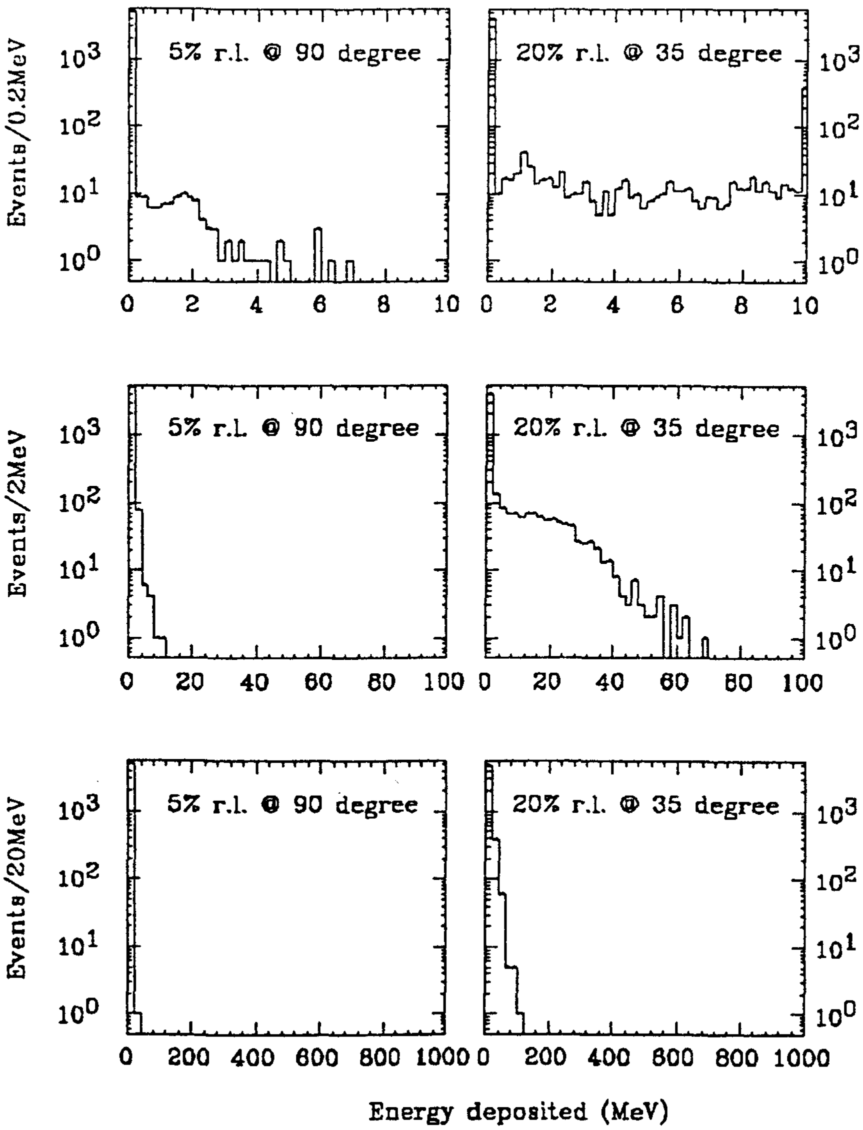


Figure 1. *Energy (in MeV) deposited by interacting photons (10, 100 and 1000 MeV) in the material in front of the calorimeter.*

Radiation Length	$E_\gamma = 10$ MeV		$E_\gamma = 100$ MeV		$E_\gamma = 1000$ MeV	
	$\theta = 90°$	$\theta = 35°$	$\theta = 90°$	$\theta = 35°$	$\theta = 90°$	$\theta = 35°$
5%	2.3%	5.4%	4.0%	8.1%	3.9%	6.5%
	1.8 ± 1.4	3.3 ± 2.8	2.1 ± 1.8	3.4 ± 3.1	2.3 ± 1.9	4.0 ± 3.0
10%	5.0%	10.3%	6.9%	14.6%	7.0%	12.5%
	3.4 ± 2.7	5.2 ± 3.4	4.0 ± 3.6	6.3 ± 6.4	4.5 ± 3.4	8.5 ± 7.7
20%	10.5%	19.7%	13.3%	25.1%	13.5%	22.5%
	5.3 ± 3.4	6.4 ± 3.6	7.9 ± 6.5	$13. \pm 12.$	9.6 ± 8.9	$18. \pm 15.$
40%	20.0%	32.6%	24.6%	42.7%	25.3%	41.1%
	6.6 ± 3.5	7.5 ± 3.3	$16. \pm 12.$	$27. \pm 22.$	$21. \pm 17.$	$41. \pm 30.$

Table 2. *The fraction of photons which interact (top entry in each box) and the mean and rms of the energy deposited in the front material (in MeV) for various photon energies, polar angles, and radiation lengths.*

Table 3 shows the mean and rms of $(E_\gamma^{generated} - E_\gamma^{measured})$ for $E_\gamma = 10,\ 100$ MeV with and without material in front of the calorimeter.

Clearly, the effect of the extraneous material on the energy resolution can be significant. The results for $E_\gamma =$ 10 MeV are accurate to $\sim 10\%$, as a 1 MeV cutoff in the energies of secondaries was used.

It is possible to detect the fact that a photon interacted in the extraneous material. In this case the mean of the energy deposited in the material in front can be added to the energy in the calorimeter proper. Since the rms of the deposited energy is almost as large as its mean, however, the energy resolution of the calorimeter can only slightly be improved by this method. One could do better if the *actual* energy deposition in the entrance material could be measured, but this is difficult to do.

In summary, our studies show that in order to achieve the required energy resolution of $\sigma_E/E = (1-2)\%/\sqrt[4]{E}$ every effort has to be made to keep the amount of material in front of the calorimeter below $10\% - 15\%$ of a radiation length. This places a stringent requirement on the material in the particle identification system. It also means that the intrinsic energy resolution of a liquid krypton ionization chamber can only be fully realized if a very thin cryogenic dewar can be designed.

Table 3. *Mean and rms of $(E_\gamma^{generated} - E_\gamma^{measured})$ (in MeV) for various photon energies, polar angles, and radiation lengths placed in front of the calorimeter*

Radiation	$E_\gamma = 10$ MeV		$E_\gamma = 100$ MeV	
Length	$\theta = 90°$	$\theta = 35°$	$\theta = 90°$	$\theta = 35°$
0%	0.02 ± 0.36		0.10 ± 0.65	
5%	0.1 ± 0.5	0.2 ± 0.9	0.2 ± 1.1	0.4 ± 1.7
10%	0.2 ± 1.1	0.5 ± 1.9	0.4 ± 1.9	1.1 ± 3.6
20%	0.6 ± 2.1	1.2 ± 3.0	1.2 ± 3.8	3.7 ± 8.7
40%	1.3 ± 3.1	2.5 ± 4.0	4.0 ± 9.1	$12. \pm 20.$

If the radiator is not directly in front of the calorimeter, its effect on efficiency is even more severe because of the magnetic field. Essentially all of the electrons from photon conversions which lie below some cutoff energy are lost. The electron and positron from a converting photon may curl up before reaching the calorimeter, or they may curve enough to strike the calorimeter sufficiently far from the main energy deposition that they are not found by the pattern recognition algorithm. For example, suppose a radiator of L_{rad} radiation lengths is 20 cm in front of the calorimeter. In a 1 Tesla field, a $p_t = 60$ MeV/c electron has a 20 cm radius of curvature. Thus any $E_\gamma \leq 120$ MeV, or $E_{\pi^0} \leq 240$ MeV is definitely lost if there is a conversion. Since the average π^0 momentum is ~ 500 MeV/c, most π^0's will be lost if one of their decay photons converts. The probability of converting a photon is $\frac{7}{9}L_{rad}$. Either photon from the π^0 may convert; the probability of losing a π^0 is thus $\frac{14}{9}L_{rad}$. The thickness of the radiator must be averaged over the polar angle, which results in a 20% increase in material in the barrel region. Thus the fraction of π^0's lost is $1.8L_{rad}$, which is 36% for a particle identification system which places 20% of a radiation length in front of the calorimeter.

2.4. Trigger

During the course of this Workshop, it has become clear that the calorimeter can provide the primary trigger and event timing for the experiment. This may, in fact, be simpler to implement than a tracking trigger (see the section on Triggering and Data Acquisition). The difficulty in triggering with a drift chamber is most easily seen by scaling the *PEP* experience with a magnetic detector. At *PEP* the charge per bunch (40 nC) is similar to the *B* Factory charge per bunch (20 nC). However, crossings at the *B* Factory (every 4 ns) will occur more frequently than crossings at *PEP* (every 2.4 μs). Where as at *PEP* there was only one crossing in the typical 1 μs drift time of the chamber, there will be 250 crossings at the *B* Factory. If masking can be done as well at the *B* Factory as at *PEP*, then the central drift chamber will have $\sim$ 125 times the single hit (soft photons) occupancy as at *PEP*!

At *PEP*, hardware road-finders for the trigger were able to deal with the typical occupancies of about 10. Unfortunately, at the B Factory, with the scaled up occupancy of 1250, the road-finders will likely find tracks on every beam crossing. We are thus motivated to investigate a calorimetric trigger. Expected rates for a total energy calorimetry trigger may be scaled from the Crystal Ball at DORIS and from CUSB at CESR.

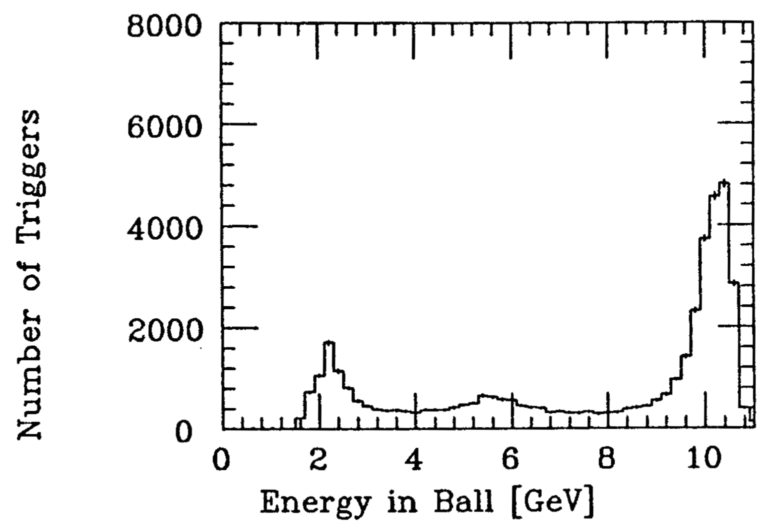


Figure 2. *Crystal Ball at DORIS. The number of triggers versus the total energy deposited in the Ball. E_{CM} = 10.5 GeV. $I_{e^-} = I_{e^+}$ = 30 ma. The trigger rate for events in this plot (1.8 GeV total energy discriminator setting) was 1 Hz.*

Figure 2 shows the energy distribution of triggers for the Crystal Ball at $\sqrt{s}$ = 10.5 GeV at DORIS, with a minimum energy requirement of 1.8 GeV. Beam gas background is just starting to be visible, producing a trigger rate of about 1 Hz. At DORIS, the beams cross every 1 μs. Scaling this to the 4 ns crossing time of the B Factory yields a trigger rate of 250 Hz for a similar 1.8 GeV total energy threshold. On the $\Upsilon(4S)$, CUSB had a trigger rate of .5 Hz for a 1.2 GeV total energy threshold (measured after the first layer of calorimeter). Scaling the 350 ns crossing time of CESR to the B Factory yields a trigger rate of 45 Hz. Therefore, all events which leave more than about 2 GeV in the calorimeter will efficiently trigger, leading to a trigger rate of $\leq$ 250 Hz. In order to trigger on events which leave less than 2 GeV in the calorimeter (for example $\mu^+\mu^-$ or $\tau^+\tau^-$ events) combinations of calorimeter topology and drift chamber roads and lower total energy thresholds ($\sim$ 200 MeV) will have to be OR'd with the primary 2 GeV total energy trigger. Clearly, the electronics of the calorimeter must be designed to be able to sum 10,000 elements without overwhelming a 200 MeV threshold with electronic noise. The greatest danger is coherent noise sources, such as clocks and reset signals from around the experiment. To minimize such coherent sources, the calorimeter electronics itself should be designed without clock pulses, and should use ADC start, latch and clear signals which occur only after the trigger. The drift chamber electronics requires clocks; grounding, cabling, and rack placement must be carefully designed to avoid introducing noise from this source into the calorimeter trigger energy sum. This also argues for phototubes rather than FET amplifiers as the first element of electronic amplification in the calorimeter, as a phototube is largely immune to low-level coherent noise pickup.

2.5. *Timing*

The earliest repeatable timing signal for an event could come from a CsI(Tl) calorimeter. The rise time of the timing siganl is essentially a property of the electronics. The total energy sum in the Crystal Ball had a rise time of $\approx$ 15 ns. With a constant-fraction discriminator, the timing resolution can be a small fraction of this rise time. We conservatively estimate σ = 10 ns for this event timing signal. With a liquid krypton calorimeter, the risetime would be the fraction ($\approx$ .1) of the electron drift time ($\approx$ 6 μs). A liquid krypton calorimeter with a 600 ns rise time will therefore not yield a time resolution as good as that of CsI(Tl), with its 15 ns rise time. Thus a CsI(Tl) calorimeter can resolve the time of an individual event to one or two beam crossings, which is quite useful for hardware road-finding in the tracking system.

3. Crystal Calorimetry

Crystal scintillators have long been used as precision electromagnetic calorimeters at e^+e^- accelerators. They can provide the segmentation, depth, and precision required for a B Factory calorimeter at an acceptable cost with a well established technology. A schematic view of such a crystal calorimeter, in relation to the full detector, is shown in Figure 3.

The goal of the electromagnetic calorimeter design is a device which can provide neutral particle resolution comparable to that expected for charged particles. This is in practice an unattainable goal, but one for which we should nevertheless strive. The weaknesses of existing detectors are particularly evident at low photon energies, where poor resolution influences the lowest usable photon energy, below which photons are no longer identified. This cutoff influences π^0 reconstruction efficiency, which is vital to the main goal of efficient B meson reconstruction.The low energy photon resolution plays an equally important role in precision spectroscopic studies.[2]

We will consider several crystal calorimeter options. We discuss the material choices, and then focus on CsI(Tl) as the crystal material. We have considered a variety of readout options; we find that in a magnetic field, phototetrode readout provides low-energy performance which is superior to conventional silicon photodiode readout. Issues of mechanical support are briefly discussed. Calibration methods for a CsI(Tl) crystal calorimeter are considered. Finally,the crucial issue of radiation resistance of pure CsI and CsI(Tl) is addressed by a survey of the literature on the subject, as well as some by specifying a program of new measurements. Whatever the radiation resistance of the crystal material, it will be important in the high-radiation environment of a B Factory to pay attention to shielding against beam losses in the IR region at injection or during accelerator studies. This can be addressed in several ways. A movable tungsten or mercury "injection shield" could be brought into position during those times when radiation losses are expected to be the highest. Alternatively, masks could be placed in the storage ring to intercept off-energy electrons far from the interaction region, thereby minimizing any additional radiation dose to the calorimeter. We have also estimated costs and time-scales for the construction of a CsI(Tl) crystal calorimeter. Our conclusion is that such a calorimeter can be built with existing technology in a timely manner at an acceptable cost.

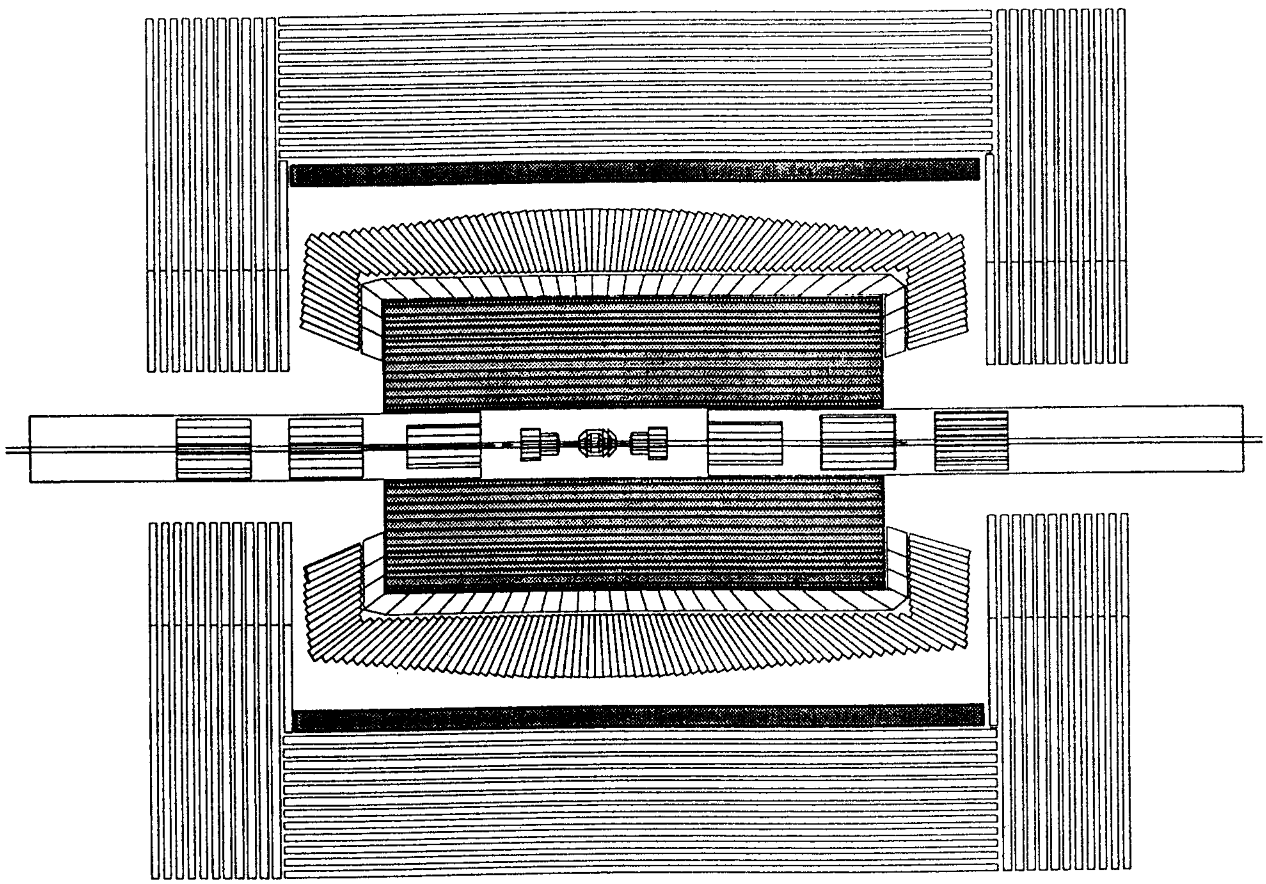


Figure 3. *The B Factory detector with a crystal calorimeter.*

3.1. Material Choice

The issue of crystal material choices have been summarized in Table 1, which leads one to the conclusion that CsI(Tl) is the best available compromise between performance (radiation length, material properties, radiation hardness, manufacturability) and cost. A large CsI(Tl) calorimeter has been built and put into operation for CLEO II at CESR. We have therefore based our crystal calorimeter studies on a CsI(Tl) crystal calorimeter. One of the most important issues to be addressed in choosing CsI(Tl) as the detector material is its radiation resistance. The radiation doses anticipated at the *B* Factory during luminosity running have been estimated and appear to be tolerable. It is important that large additonal doses not be produced during injection or machine physics studies. Were these additional doses to amount to many krad, they may be unacceptably large for CsI(Tl) unless further steps are taken to protect the crystals, either through stationary or movable shielding or the use of upstream masks in the storage rings. This problem is much less important for the liquid krypton alternative; such shielding may, howver, still be required by the other detector elements.

3.2. Shielding Requirements

It is difficult to estimate the shielding requirements for a CsI(Tl) calorimeter as neither the radiation resistance of CsI(Tl) nor the expected radiation dose at the B Factory are completely understood. There was a consensus, however, that there must be some form of radiation shielding for the detector. One approach is to have a $\approx$ 5 cm thick tungsten "gun barrel" which can close during injection and machine studies periods, when we expect the largest doses, and be withdrawn for normal data-taking. The shielding should extend beyond the intersection region, providing protection for the calorimeter from beam losses as far upstream of the IR as possible. Measurements[3] at CLEO II during two months of normal running and injection extrapolate to a dose from CESR of 4 rad/year at a radius of 100 cm. Scaling this by the ratio of injected charge at the B Factory to that at CESR predicts 500 rads/year at the B Factory. CLEO measurements on BDH, Horiba, Bicron and Harshaw CsI(Tl) crystals showed 5% to 15 % decreases in light output after 100 rad ^{60}Co irradiation at the non-readout end of a 30 cm long crystal.[3] This scheme presents serious mechanical difficulties, and as it unavoidably occupies radial space, can increase the cost of components at larger radii (cost $\propto r^2$). An alternative is therefore to explore the placement of masks far upstream of the interaction region which intercept beam particles which stray far from the central orbit. The idea is to remove such particles from the beam in such a way as to prevent them from depositing energy in the interaction region. Such a scheme has been investigated in detail at KEK [4]and appears to be practical.

3.3. Readout and Electronics

Two constraints have been imposed on the calorimeter readout: it must work in a magnetic field, and it must preserve the excellent intrinsic response of the crystal calorimeter down to the lowest possible energy. The readout scheme adopted in the CLEO II detector uses four silicon photodiodes on each crystal with a charge sensitive preamplifier. The weakness of this approach is that at low energies the energy resolution is dominated by electronic noise in the preamplifier. For CLEO II, the electronic noise has $\sigma \simeq 1$ MeV rms. This noise dominates the energy resolution below ~ 200 MeV. We have studied a variety of alternative readout schemes in an attempt to improve the low energy performance.

3.4. Readout Types

The first approach we have considered employs "flux concentrators", in which silicon photodiodes detect the light from a wavelength shifter which covers the entire rear face of the crystal. The advantage of this scheme is that one can afford to cover the entire rear face of the crystal, which would be prohibitively expensive with conventional photodiodes. The price paid is the amount of light which finally enters the photodiode.

A schematic view of this type of readout is shown in Figure 4. Recent improvements in this technique by E. Lorenz have made this approach quite attractive.[5] The second type of readout considered is a phototetrode (a two stage photomultiplier tube). A typical phototetrode is shown in Figure 5. This scheme has the distinct advantage that it employs an essentially noiseless amplifier (the phototetrode) as the first amplification element. Phototetrodes are usable in relatively high magnetic fields (see Figure 6), but their axis cannot be too steeply inclined to the field (see Figure 7).

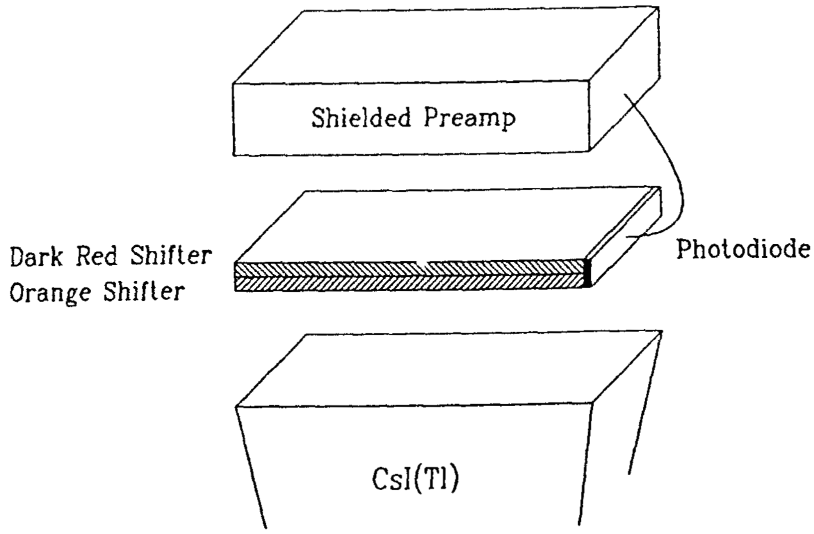


Figure 4. *Schematic view of "flux concentrator" readout.*

The phototetrode readout scheme requires a more complex mounting scheme, since those crystals which are oriented at 90° to the beam axis (and hence at 90° to the field axis) must have their phototetrodes mounted at an angle in order retain a useful gain. Several solutions to this problem have been considered; they all consist of angling the phototetrode with a plastic wedge (a *cookie*) attached to the crystal, as shown in Figure 8. In this scheme it is possible to orient all phototetrodes at angles of 45° or less to the B field axis. This solution may appear somewhat more expensive than the photodiode scheme, as phototetrodes are more expensive than silicon photodiodes, but the electronics should be somewhat simpler, and hence less expensive.

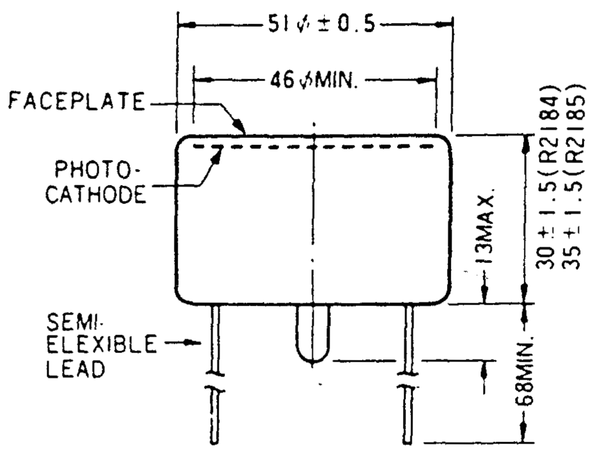


Figure 5. *Dimensions of a Hamamatsu R2185 phototetrode.*

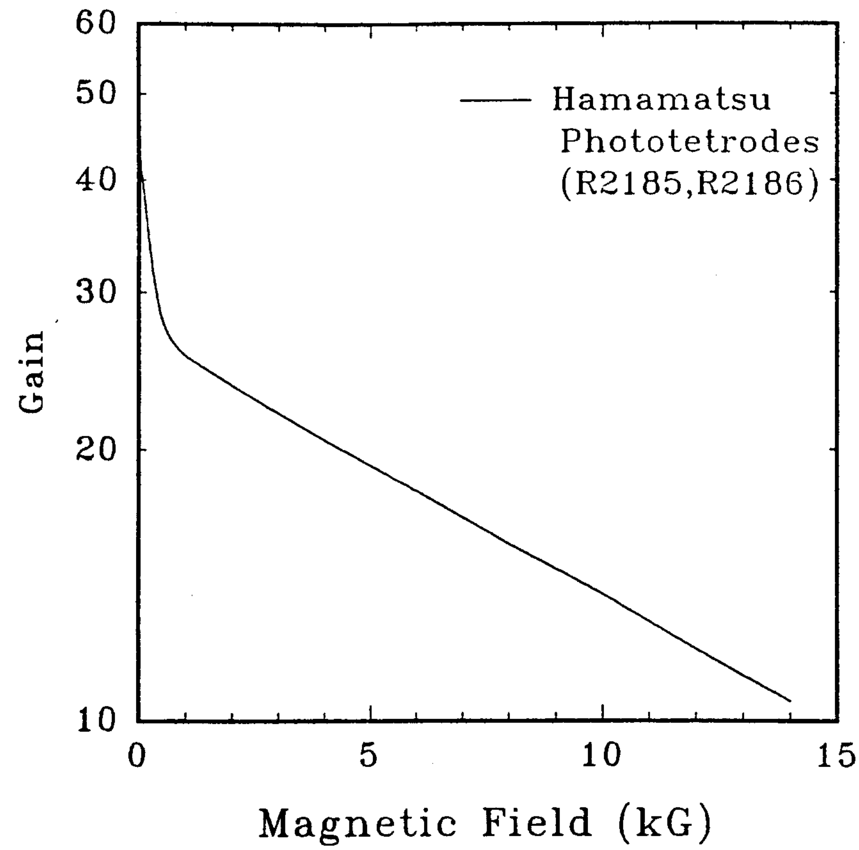


Figure 6. *Dependence of phototetrode gain on magnetic field strength. The field is aligned along the tube axis.*

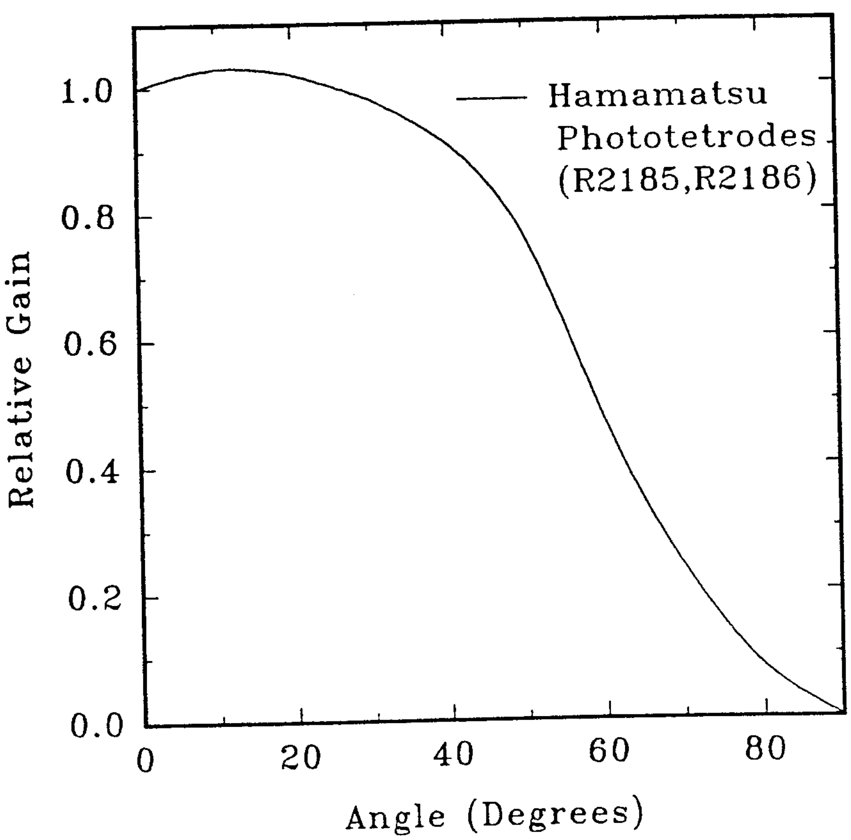


Figure 7. *Dependence of phototetrode gain on orientation of the magnetic field relative to the tube axis.*

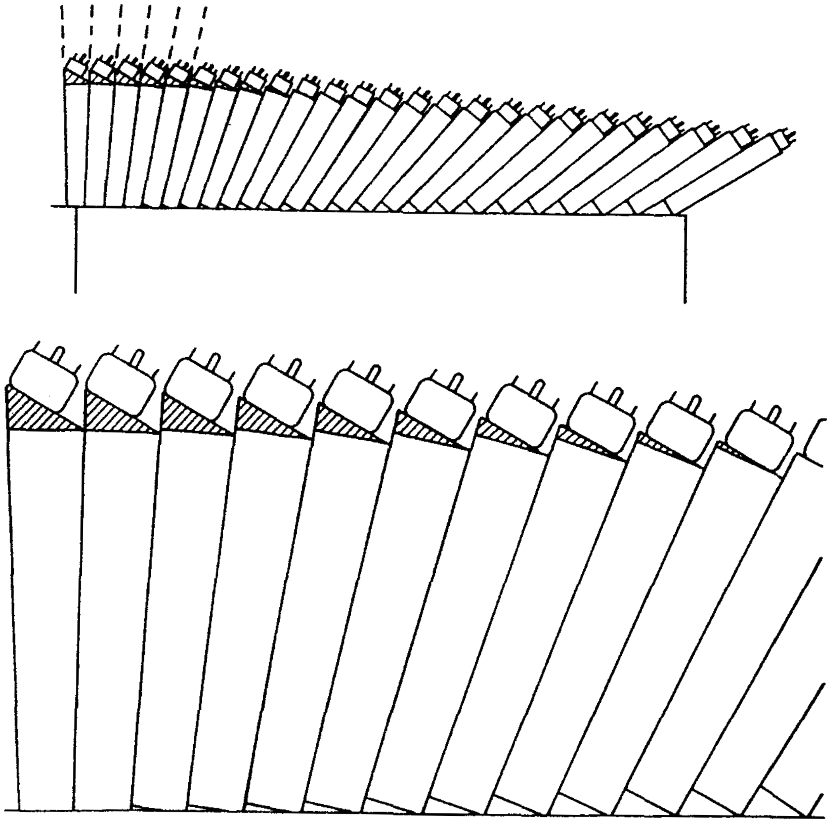


Figure 8. *Sketch of a design which tilts the phototetrodes toward the B field axis.*

3.5. *Tests*

We have carried out a number of tests of CsI(Tl) crystals. Besides studies of radiation damage with large crystals the main issues are light yield, energy resolution and homogeneity using crystals from different manufacturers and testing various readout systems. Figure 9 shows four possible readout techniques for CsI(Tl) crystals. For the B Factory detector the readout of an angled phototetrode coupled to the crystal via a prism and the readout with wavelength shifters and photodiodes are of particular interest, while the other two are studied for comparison.

Crystals

We have collected various crystals from different companies to perform these tests. Table 4 lists the different properties of the crystals. The Quartz & Silice crystals were grown using the Kyropoulos method, in which the mold is cooled from the top, allowing the crystal to grow downwards. The ingots are typically 19 inches in diameter and 3 inches thick. Depending on the size, several crystals can be cut from one ingot. The Tl doping is uniformly distributed across the crystal. The other crystals were grown using the Bridgeman method, in which the crystals are cooled from the bottom so that the crystals grow upwards. This technique introduces a gradient for impurities and the Tl doping with the highest concentration at the top.

Source	Shape	Cross section [cm^2]	Length [cm]	Comment
Quartz & Silice	block	3×4	34	$\sim$ 1000 ppm Tl
Quartz & Silice	pyramid	$5 \times 5 \rightarrow 6.5 \times 6.5$	37	high Tl concentration
Quartz & Silice	cylinder	0.79	1.0	Tl doping
Horiba	pyramid	$3 \times 3 \rightarrow 5 \times 5$	18	Tl conc. varies w. length
Horiba	block	5×5	30	Cleo crystal with 4 PD's
BDH	block	5×5	30	Tl doping
BDH	block	5×5	15	Tl doping
Bicron	block	2.5×2.5	30	pure, no Tl
Harshaw	block	3×3	12	Tl doping
Harshaw	block	3×3	21	Tl doping

Table 4. *Properties of test CsI crystals.*

The crystals were wrapped with a teflon film and aluminized black plastic foil. For the following tests the crystals were read out with a 56DVP photomultiplier (PM), coupling the crystal to the PM via a silicone rubber cookie. The anode signal was amplified by an Ortec EG&G amplifier and fed into an Ortec EG&G ADC which was operated by a PC. For most of the measurements a ^{22}Na source was used. All test were performed inside a light-tight box.

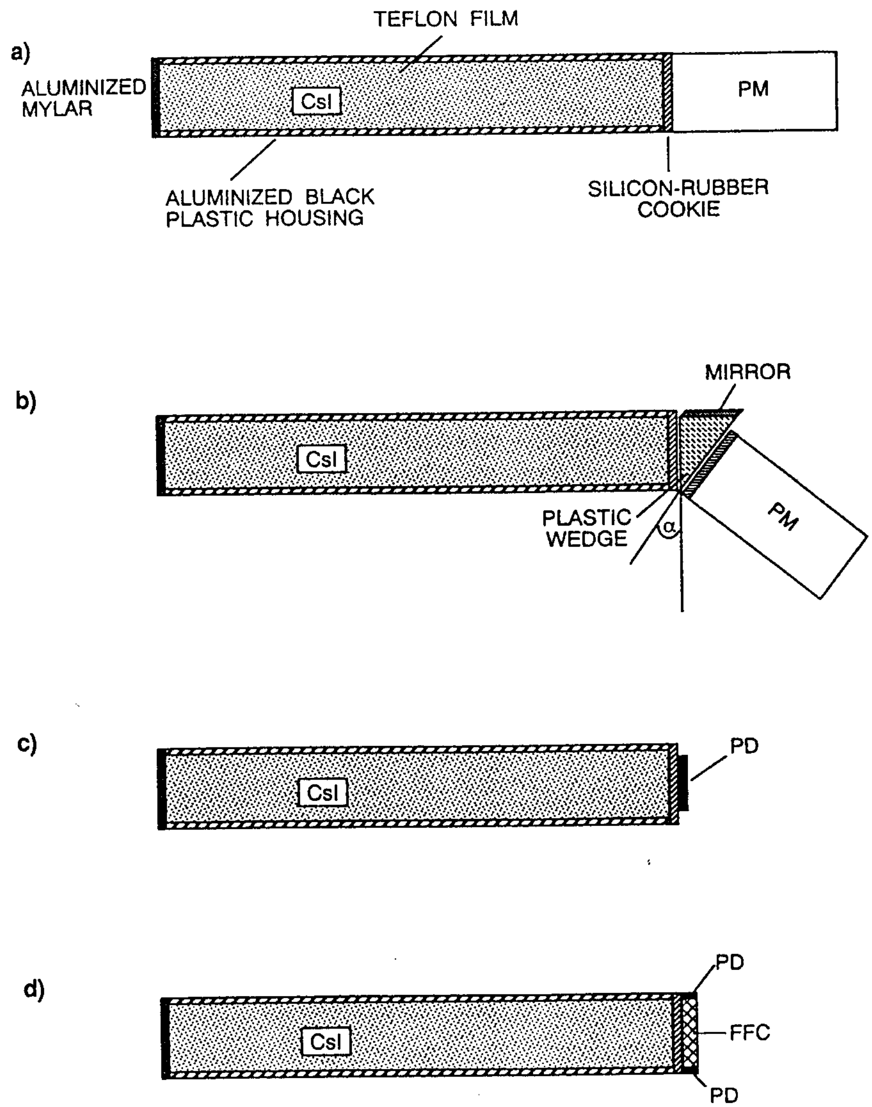


Figure 9. *CsI crystal readout techniques;* **a.** *direct coupling to photomultiplier or phototetrode;***b.** *angled coupling via a prism to a photomultiplier or a phototetrode; bf c. direct coupling to phototdiodes;* **d.** *coupling to photodiodes via fluorescent flux concentrators.*

Linearity

In order to check the linearity of our measurement technique, the system was first calibrated with the ^{22}Na lines and then tested with the ^{137}Cs line and the lower of the two ^{60}Co lines. The result, shown in Figure 10, demonstrates that the system is linear. In order not to be sensitive to gain fluctuations, which may occur over longer time periods, the ^{22}Na source was used in all measurement.

Comparison of Different Crystals

The energy resolution of the different CsI(Tl) crystals was first measured with a 56DVP photomultiplier. The results for the 0.511 MeV and 1.275 MeV ^{22}Na lines are presented in Figure 11. The energy resolution is better for the small crystal (typically 2.9% at $E_2 = 1.275$ MeV and 4.5% at $E_1 = 0.511$ MeV) than for the large crystals (4.4% at 1.275 MeV and 7.4% at 0.511 MeV). If photon statistics dominated the resolution and all photons emitted under angles larger than the critical angle reached the photocathode, one would expect a resolution of $\sim 1.5\%$ at 1.275 MeV from a crude estimate, assuming a 15% photocathode quantum efficiency.

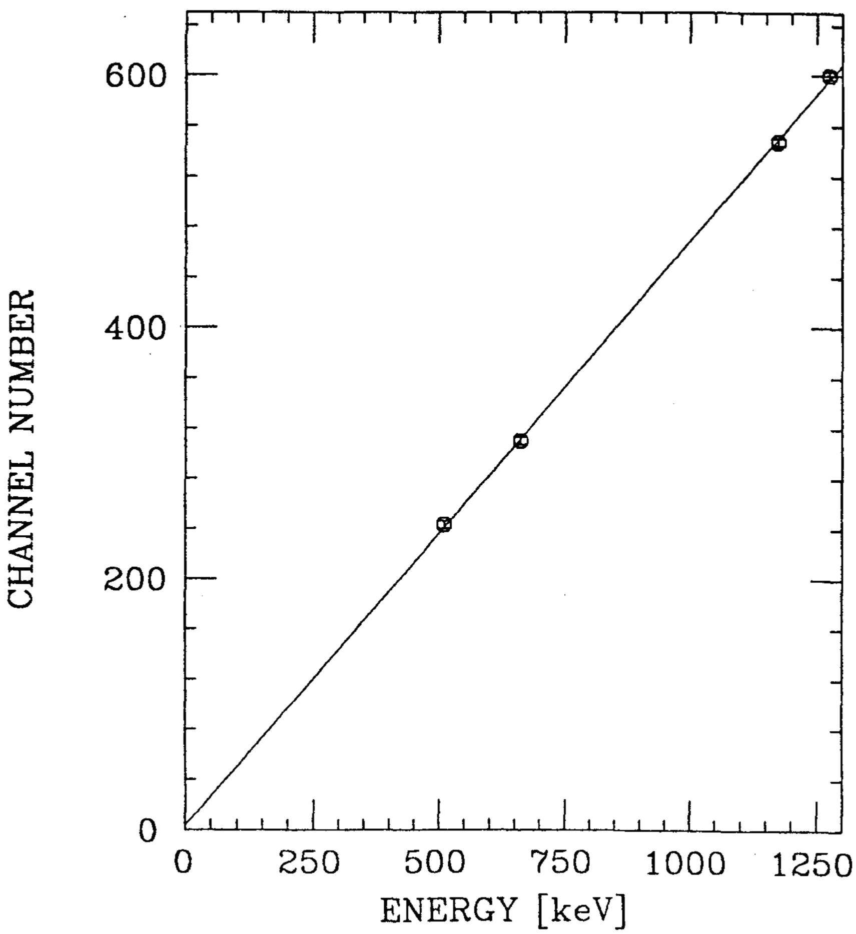


Figure 10. *Linearity of CsI readout with a photomultiplier*

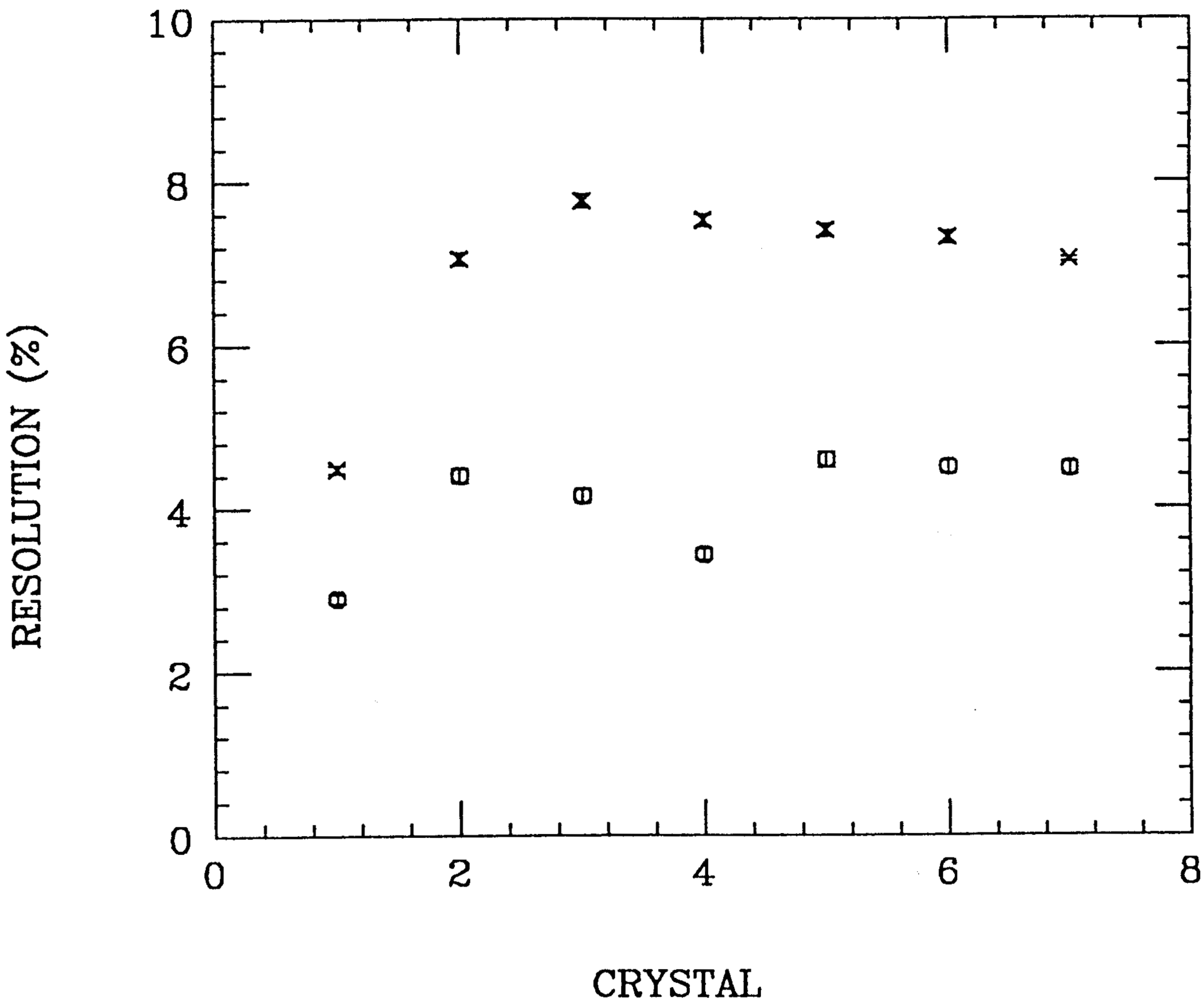


Figure 11. *Energy resolution of different CsI(TL) crystals with photomultiplier readout using a ^{22}Na source. The crosses are data for the 0.511 keV line and the squares are data for the 1.275 keV line. 1) small Quartz & Silice cylinder, 2) Quartz & Silice block, 3) Quartz & Silice pyramid, 4) BDH block, 5) Horiba pyramid, 6) long Harshaw block, 7) short Harshaw block.*

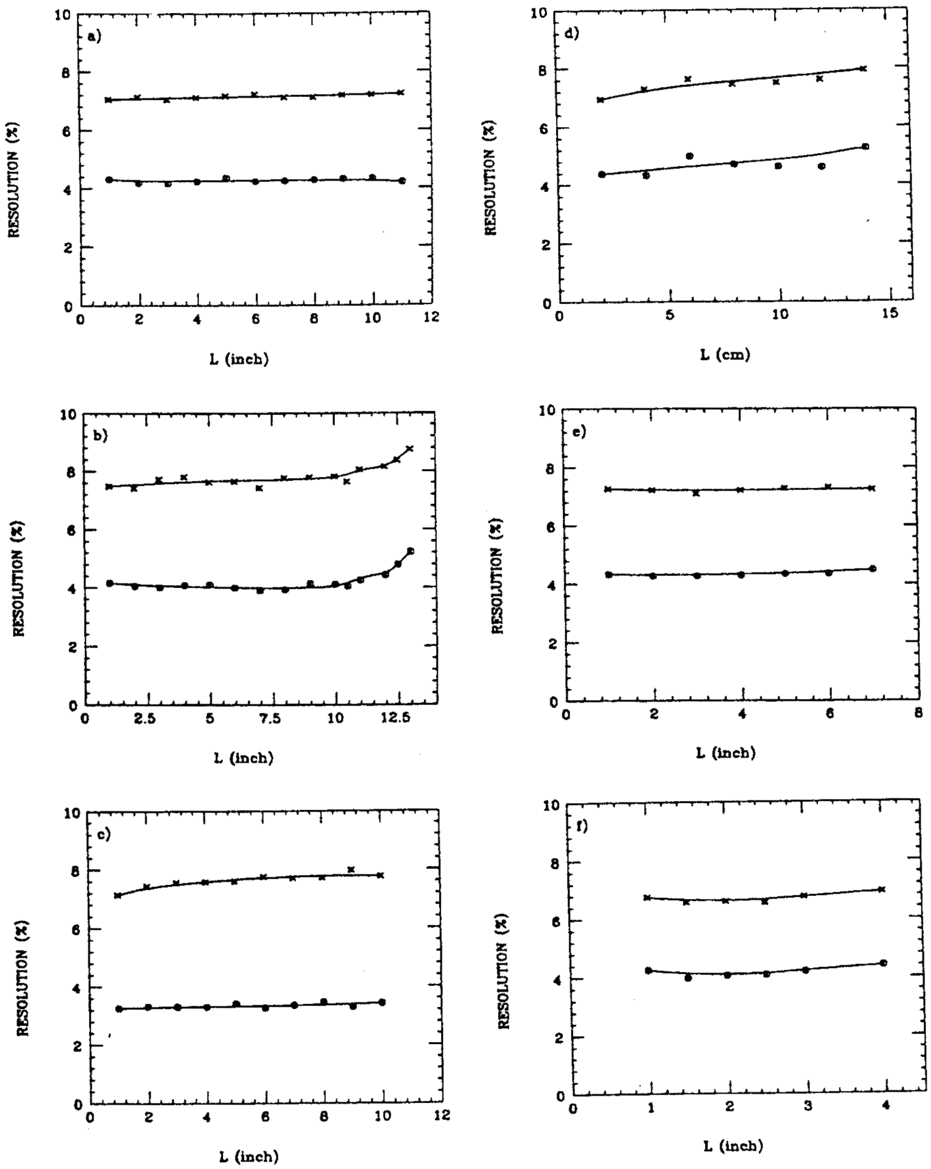


Figure 12. *Energy resolution as a function of crystal position for* **a.** *Quartz & Silice block,* **b.** *Quartz & Silice pyramid,* **c.** *BDH block,* **d.** *Horiba pyramid,* **e.** *long Harshaw block,* **f.** *short Harshaw block.*

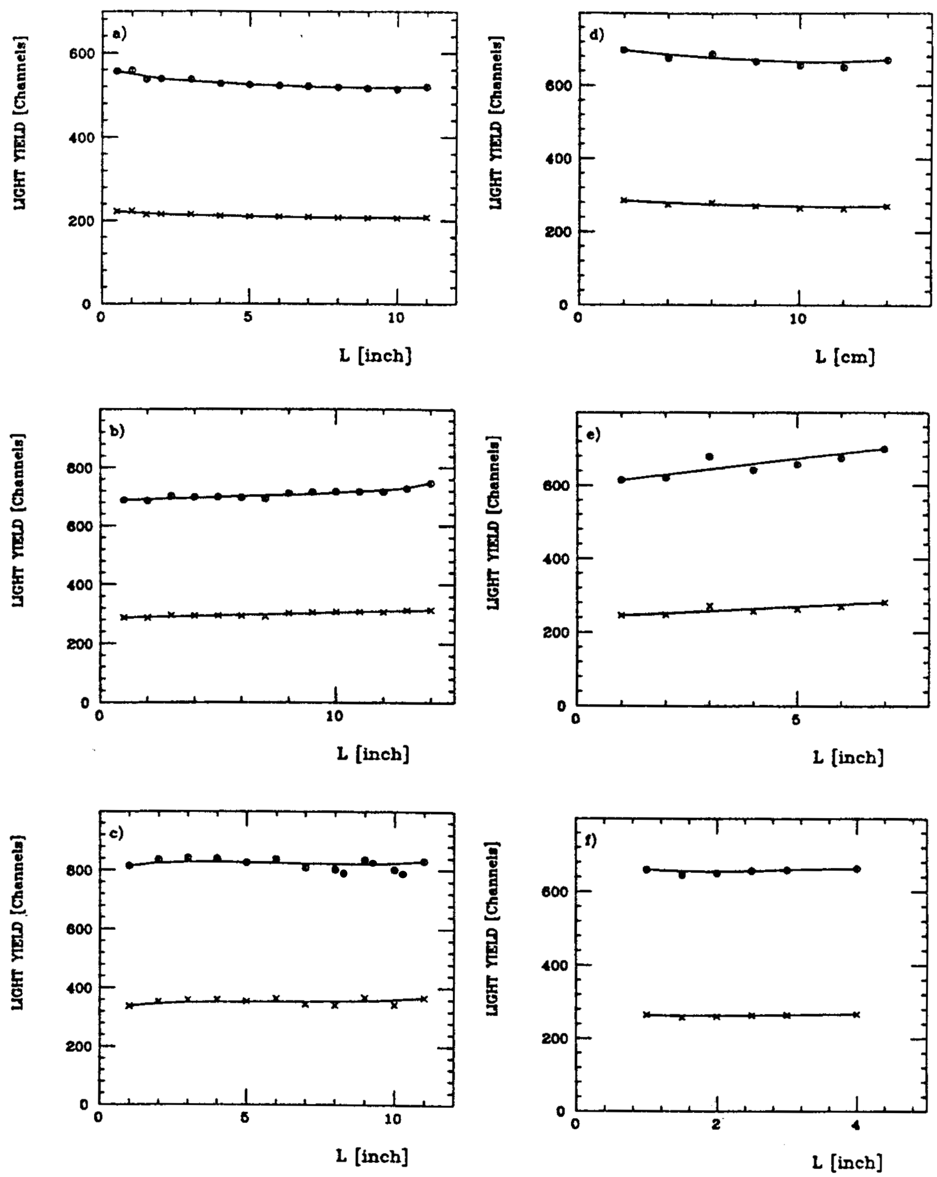


Figure 13. *Light yield resolution as a function of crystal position for* **a.** *Quartz & Silice block,* **b.** *Quartz & Silice pyramid,* **c.** *BDH block,* **d.** *Horiba pyramid,* **e.** *long Harshaw block,* **f.** *short Harshaw block.*

If all produced light reached the photocathode, the resolution would be 1%. Since for a PM read out both noise and leakage effects are small, the worsening of the resolution results from both a loss of photons and line broadening, both caused by geometric effects, surface irregularities, light attenuation and gain fluctuations. For the small crystal the ratio of resolutions of both lines scales closely with $\sqrt{E_2/E_1}$, implying that here the main source is a loss of photons. In the long crystals, however, geometric effects and surface irregularities become more important. Therefore, one expects line broadening to cause additional worsening of the resolution. One indication is that the ratio of resolutions is larger than $\sqrt{E_2/E_1}$. Further evidence supporting this observation will be provided below. For the BDH crystal, the ratio of resolutions for the two lines is larger than for any of the other crystals.

Homogeneity Measurements

Homogeneity studies have been performed for six of the large crystals. The results for the energy resolution and light yield are shown in Figures 12and 13. The four large blocks are rather homogeneous across the entire length except for edge effects. The Quartz & Silice pyramid is rather homogeneous except for the last inch near the PM, where the resolution becomes worse by about 15%, whereas the Horiba pyramid shows a slight increase in resolution of about 15% between front and back. While the Quartz and Silice crystal underwent a surface treatment to obtain uniform light output, the Horiba crystal achieves uniformity by increasing the Tl concentration by appropriate amounts.

Phototetrode Readout

The performance of the smaller BDH block has been studied with a Hamamatsu 5 cm phototetrode which was coupled directly to the crystal with Dow Corning optical grease. A large effective electron signal is produced, which corresponds to an equivalent signal of 87,000 electrons into the preamplifier for the 0.662 MeV ^{137}Cs γ line. Note this does not include the effects of an angled coupling (which loses less than a factor of 2), the effect of the field (a factor of 3), and the reduction in gain of the phototetrode caused by an off-axis B field (a factor of 2). Taking into account these additional anticipated losses, the signal into the preamp should correspond to $\approx 8,000$ electrons, while the noise can be kept below 100 rms electrons with a several μs integration time, thus allowing us to measure the response to low energy γ sources in an individual crystal. Such a radioactive source calibration scheme has many advantages, but cannot be used if the low-energy performance of the system is dominated by electronic noise. This scheme also ensures that the low photon energy resolution is dominated by photon statistics and not by electronic noise. It should be noted that although the design collision rate is extremely high (250 MHz), the singles rate into any individual detector element is not high at all, so that long integration times can readily be employed to reduce $1/f$ noise.

Photodiode Readout

The photodiode (PD) readout of CsI(Tl) crystals is used in the Cleo II detector. The rear face of the crystals, which is on average $\sim 6 \times 6$ cm^2, is read out by four 1×1 cm^2 Hamamatsu photodiodes. The measured energy resolution is $\sigma_E/E = 1.4\%$ at 5 GeV and 3.9% at 100 MeV.[6] The equivalent noise is about 700 keV. With the development of new low-noise photodiodes from Hamamatsu, significant improvements are expected. E. Lorenz has recently performed such measurements.[7] A pyramidal Quartz & Silice crystal (3.5×3.5 cm^2 front face, 4.5×4.5 cm^2 back face, 34 cm long) has been read out by an 18×18 mm^2 Hamamatsu photodiode (S3204-03) attached to the back face. The crystal was wrapped in three layers of Teflon film. Using a ^{60}Co source a yield of 9700 photoelectrons per MeV was measured, which is about a factor of two higher than previous measurements. The equivalent noise was 59 keV. The two ^{60}Co lines could be resolved. Other encouraging results with smaller CsI crystals have been obtained by another group at MPI.[8]

Another interesting readout technique developed by Lorenz consists of coupling the crystal to a thin fluorescent flux concentrator (FFC), which covers the rear face of the crystal and is read out by two rectangular low noise photodiodes located on two side faces. The other two side faces are covered by reflectors. The FFC contains several laser dyes which shift the light from 570 nm to about 800 nm, a region where the sensitivity of the PD's is largest. Light from the crystal enters the FFC through the large face. The dye concentrations are typically a few hundred ppm, thus providing a high probability for absorbing light within a few mm. The light emitted with angles larger than the critical angle ($\sim 42^o$) is trapped inside the FFC, travelling towards the side faces. With this technique, flux concentrations of 10-100 have been achieved. Currently, the Crystal Barrel detector at CERN uses such a readout system. The measured[9] energy resolution is $\sigma_E/E = 2.5\%/^4\sqrt{E}$. With the development of new low-noise photodiodes and more efficient flux concentrators, the FFC/PD system is a promising candidate for the CsI(Tl) calorimeter readout.

Study of angled Photomultiplier Readout

As readout of CsI(Tl) crystals with phototetrodes (PT) in a magnetic field requires tilting of the PT, we have studied the light yield and resolution as a function of the tilting angle. The measurements were performed with a 56DVP photomultiplier in the absence of a magnetic field. The crystal was coupled to the PM via a prism as shown in Figure 9b. To obtain good optical contact, silicone-rubber cookies were used at all boundaries; all exterior surfaces were covered with aluminum foil. Prisms with different angles were made of lucite, wavelength shifter material and plastic scintillator. The light yield and energy resolution measurements for the large Quartz & Silice block are summarized in Figures 14 and 15, respectively. The light yield drops considerably with increasing angle, as expected. While the wavelength shifter and lucite prisms yield similar results, the plastic scintillator prisms are slightly better. The resolution, however, remains unchanged, and is nearly the same for all three materials. C. Woody at BNL has found similar results with lucite prisms.[10] We have also studied the angular dependence of the small cylindrical Quartz & Silice crystal. As shown in Figure 16 for the prisms made of wavelength shifter, here the resolution increases

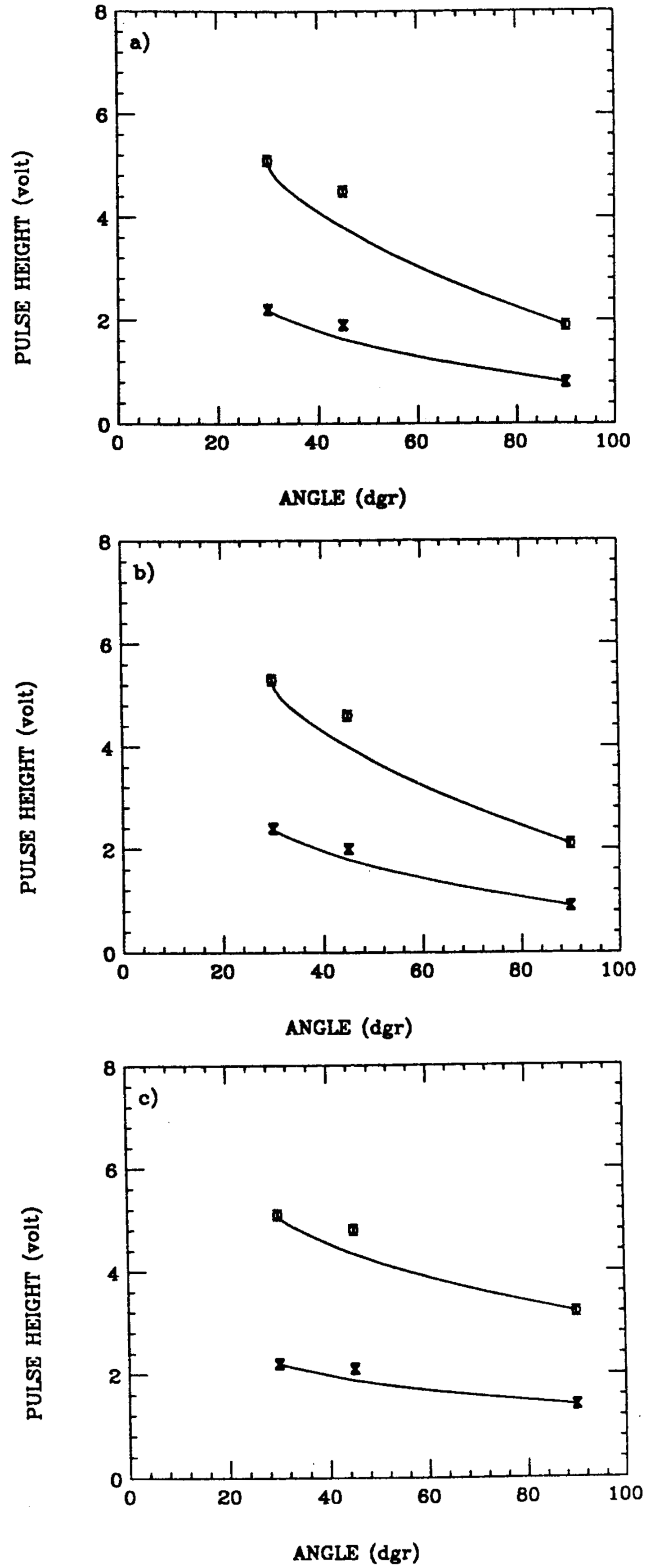


Figure 14. *Angled readout of large Quartz & Silice block showing the light yield as a function of the angle between crystal and photomultiplier for prism couplings made of* **a.** *lucite,* **b.** *green wavelength shifter,* **c.** *plastic scintillator.*

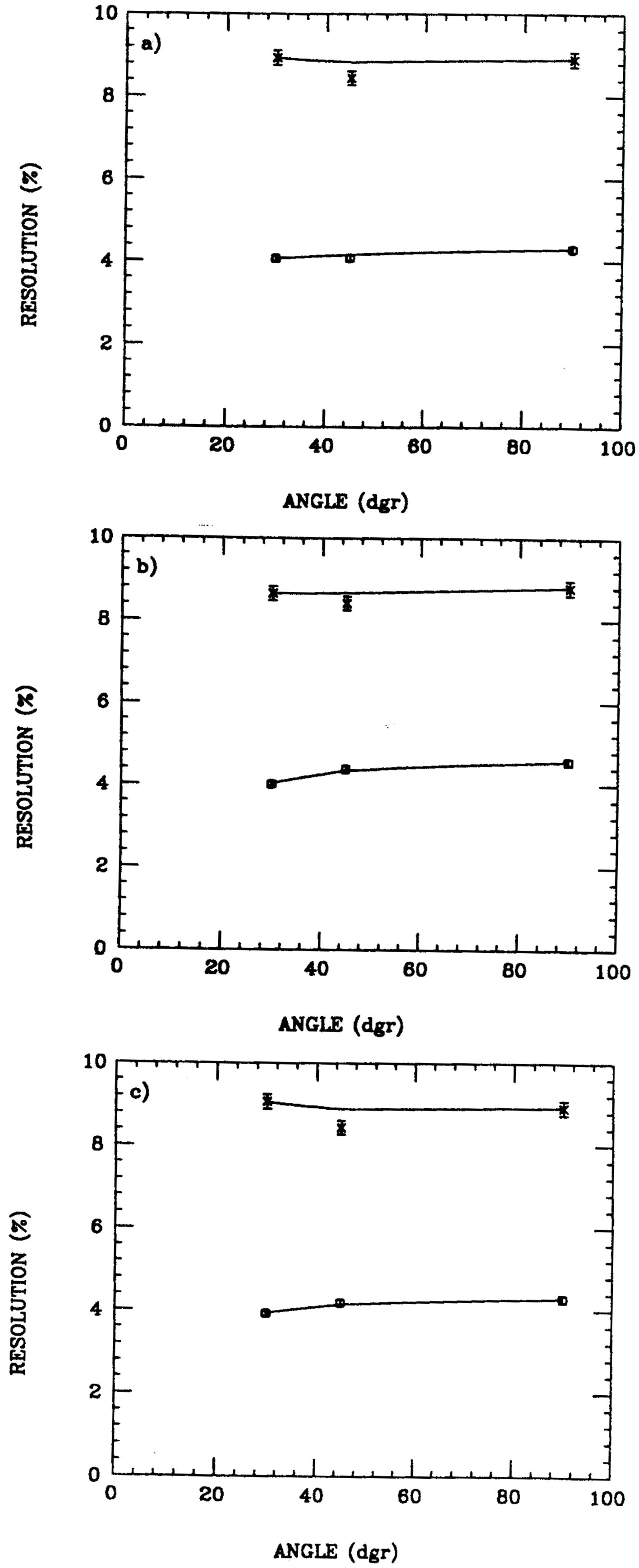


Figure 15. *Angled readout of large Quartz & Silice block showing the energy resolution as a function of the angle between crystal and photomultiplier for prism couplings made of* **a.** *lucite,* **b.** *green wavelength shifter,* **c.** *plastic scintillator.*

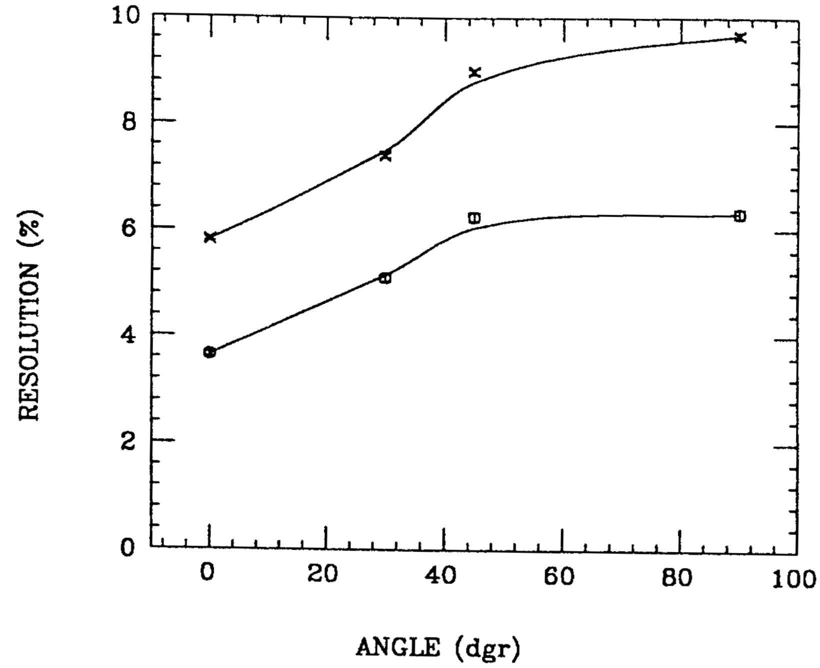
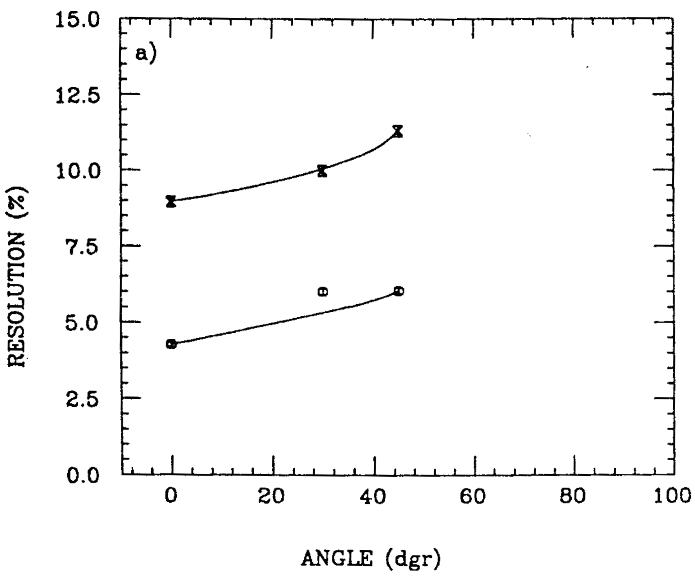
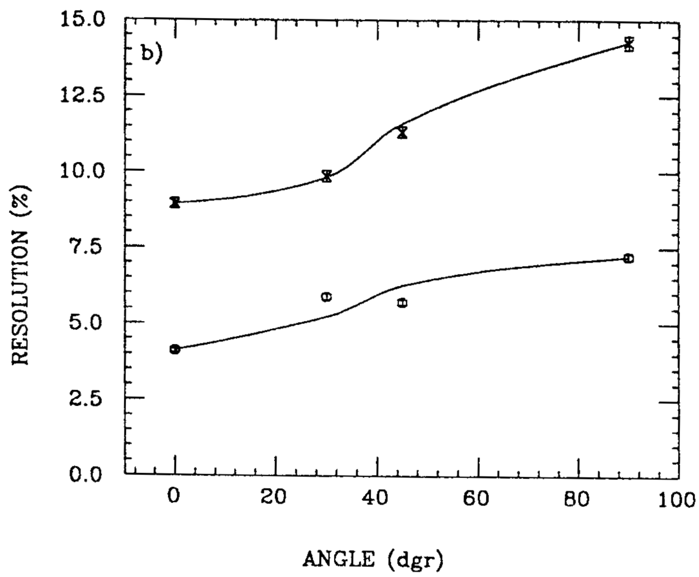


Figure 16. *Angled readout of small Quartz & Silice cylinder showing the energy resolution as a function of the angle between crystal and photomultiplier for prism couplings made of green wavelength shifter.*





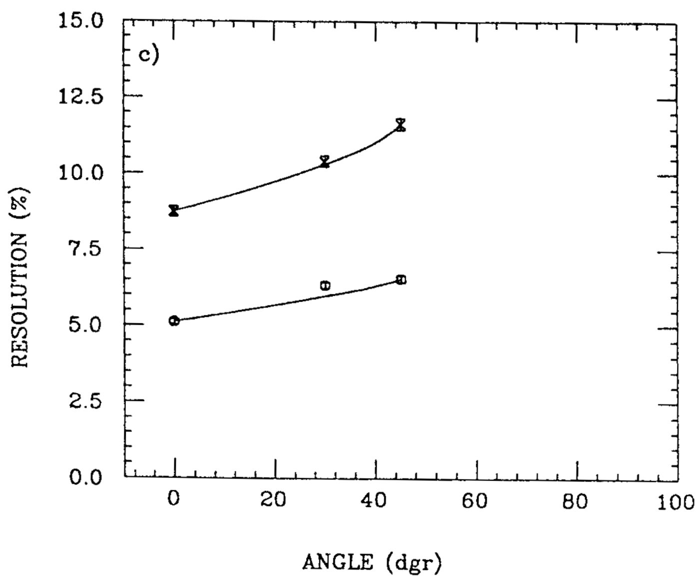


Figure 17. *Angled readout of large Quartz & Silice block showing the energy resolution as a function of the angle between crystal and photomultiplier for different couplings:* **a.** *air,* **b.** *air and mirrors,* **c.** *Fresnel lens.*

considerably. We have performed further studies in which the prism was replaced by an air gap, a mirror and a Fresnel lens. The results for the energy resolution are presented in Figure 17. Again, the resolution becomes worse with increasing angles as expected.

The most plausible explanation of all this data is that in small crystals photon statistics is the main contribution to the energy resolution, while in large crystals line broadening due to geometric effects and surface irregularities determines the resolution. Only in the case of very lossy couplings does photon statistics dominate the resolution. Therefore, the readout with a lucite prism and a phototetrode provides another promising candidate for the a readout system of CsI(Tl) crystals. We still must verify that these results hold for a phototetrode readout in the presence of a magnetic field.

3.6. Geometry

The geometry of the crystals is a tapered trapezoidal prism. An example of a typical crystal is shown in Figure 18. The entrance face is about 4×4 cm^2, the length 34 cm, and the exit face about 5.5×5.5 cm^2. For these typical sizes, we are led to a total of 8496 crystals in the barrel and 998 crystals in each endcap. The calorimeter endcaps would extend to 300 mrad on both sides (the minimum angle before being shadowed by the first dipole of the storage ring), although the detector itself would be asymmetric with respect to the electron and positron directions (reflecting the energy asymmetry of the beams).

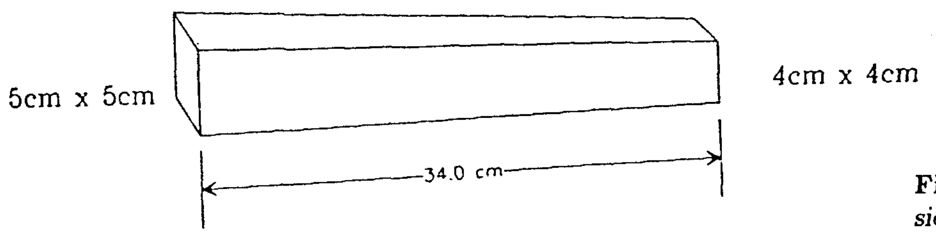


Figure 18. *Typical dimensions of CsI(Tl) crystals.*

3.7. Mechanical Support

Mechanical support of a large array of CsI(Tl) crystals presents a state-of-the-art engineering design problem, which we have not attempted to solve at this Workshop. The advent of large carbon-fiber/epoxy structures allows support of an array of 10,000 crystals with a minimum of dead material between the crystals. The superb energy resolution of which a single CsI(Tl) crystal is capable can rapidly be compromised if there is an excessive amount of inactive material between neighboring crystals which share the energy deposited by a single photon.

As a guide to what is possible, we can take the example of the BGO electromagnetic calorimeter of the L3 experiment.[11] The L3 calorimeter is supported by a carbon fiber composite structure which has walls 250μm thick. At this thickness, it is worth noting that the fiber sheet is not completely opaque to visible light. At twice this thickness, effectively complete opacity is achieved. A thickness of 250μm corresponds to 1.3×10^{-3} *r.l.* No other structural material approaches carbon fiber composites in this regard. The *B* Factory CsI(Tl) design is larger than the L3 BGO device; somewhat thicker intercrystal separators will likely be necessary. The remarkable mechanical properties of carbon fiber composites (modulus of elasticity 22,000 daN/mm^2, tensile strength 250-310 daN/mm^2 along the direction of fiber orientation) make it highly likely that they will be the primary structural material of the crystal support.

Mechanically, CsI(Tl) is plastic, rather than brittle like NaI(Tl) and BGO. It has a compressive strength less than 1% that of BGO, and even less than that of NaI.[12] This could mean that a structural design suiting the latter will not be applicable to the former. The design question comes down to whether the entire load must be taken in tension by the composite walls, or whether some of the compressive load can be taken through the walls of the individual compartments to the crystals themselves.

3.8. Calibration

After all the care which goes into designing a high precision, highly segmented calorimeter, equal care must be given to the calibration. Existing detectors have used several methods of calibration. For example, CUSB has used individual ^{60}Co sources (which have monochromatic photon lines at 1.17 and 1.33 MeV) on each crystal to maintain cross-calibration to better than 1%; the overall scale is then adjusted using Bhabha scattering events. The Crystal Ball has successfully used ^{137}Cs (662 keV), a pulser system, and Bhabhas to obtain cross calibrations good to 1% rms. The CLEO II CsI calorimeter uses Bhabha events to obtain a relatively long time scale ($\sim$days) calibration. The energy scale can then be checked against the π^0 mass. We propose to utilize all three methods to obtain and maintain the calibration of the B Factory calorimeter to better than 0.5% *rms.*.

3.9. Bhabhas

At a luminosity of 3×10^{33} cm^{-2} s^{-1}, the Bhabha rate integrated over the entire detector down to a polar angle of 300 mrad is 90 Hz . Even at 90° the rate into a single crystal is 0.005 Hz. The high luminosity of the B Factory thus makes it possible to employ Bhabha events for intercalibartion as well as for setting the overall scale. Maintaining an intercalibration error at the 0.5% systematic level using only the Bhabhas, requires ten events per crystal, which can be obtained in a relatively short time ($\approx$ 30 minutes at 90°).

3.10. Sources

Another calibration technique is the use of low-energy radioactive sources (*e.g.* ^{60}Co) on every crystal, monitoring the source peak position through a separate high gain channel. This calibration can run simultaneously with data taking in a self-triggered mode. A few minutes of running per crystal is required to obtain the desired intercalibration accuracy of 0.5%. Depending on the level of multiplexing used, this data can be obtained in a fraction of the time of a single fill, with no substantial impact on data-taking.

3.11. Light Pulser

A light pulser signal distributed to individual crystals can be used to track the cross-calibration of the crystals in the data ADC channels. Since all channels can be pulsed at once, this method provides a very fast calibration technique; it is an especially useful method to employ between runs.

3.12. Radiation Resistance of CsI(Tl) and CsI

The most serious potential problem with a crystal calorimeter at a B Factory concerns the radiation-hardness of the material. The problem of evaluating the potential damage is exacerbated by the uncertainty in the expected radiation dose at the B Factory, and even more by a less than consistent set of measurements on the radiation damage of CsI(Tl). This is an area which requires further study if the decision is made to build a crystal calorimeter. The Crystal Ball received approximately 300 rads in its tunnel modules during two years of operation at DORIS II. This resulted in a noticeable drop of a few percent in the photopeak position for a ^{137}Cs source placed at the opposite end of the 16 inch long crystal from the PMT. The Crystal Ball was opened to 1.5 meters from the beam pipe and a 1 inch thick lead shield put around the beam pipe for all injections. A 20 radiation length stack of lead was placed around the beam pipe on machine physics days. The light transmission in NaI(Tl) for a ^{137}Cs source was measured by the Crystal Ball group to be exp(-(Dose$\times$Thickness/10^5 rad-inches). The literature on radiation damage to CsI(Tl) and pure CsI spans the past six years. It is not entirely self-consistent; in addition it is difficult to compare results, since differing testing methods and varying crystal sizes were employed in the tests. We have summarized some of the salient features of the measurements where known in Table 5.

Early studies on CsI(Tl) indicated noticeable damage at modest exposures of a few hundred rad. More recent studies have indicated very good radiation damage resistance up to the hundred kilorad range (for 1 cm cubes). The differences may be due to the amount and type of impurities in the CsI(Tl) samples used. One measurement of the radiation damage to pure CsI indicates good radiation resistance up to the megarad range for small samples. The measured test results[13] are shown in Figure 19. Most of these measurements on the relative light output were done on 10 cm long crystals by monitoring the peak position of a ^{137}Cs (0.66 MeV) source, except for the Renker and the Woody measurements which were done on small samples.

In Figure 20 we have attempted to normalize the measurements to the effect on a 35 cm long crystal uniformly exposed to radiation. This was done assuming that all the damage is due to reductions in the transmission coefficient (thus in extrapolating from 10 cm to 35 cm we used $[\text{measurement}]^{3.5}$). Note that the large correction for the small 1 cm^3 crystal measurements (Renker) makes the extrapolation to 35 cm long crystals quite suspect. The problem with pure CsI, which *may* have radiation resistance superior to CsI(Tl), is that it produces about a factor of five less light than CsI(Tl).

D. Renker at PSI has recently carried out a new study on radiation hardness of CsI(Tl) crystals.[14] These results, not included in Table 4 or Figures 19 and 20, help to clarify the radiation-damage issue. A total of 28 small 1 cm^3 crystals from several manufacturers with different Tl concentrations were exposed up to doses of 5000 rad. Both the light yield and transmission of green light were monitored during exposure. All crystals were read out with photodiodes. Only four of the crystals showed effects due to radiation exposure; after 500 rads the light yield decreased by 3-20%. In one case, the sample was cut from the top of an ingot grown with the Bridgeman method. A sample cut from the bottom of the same ingot, however, showed no reduction in light yield. The relative light yield and transmission for green light (565 nm) is shown in Figures 21 and 22 for both samples. Due to the crystal growing technique, the thallium concentration differed from 260 ppm at the bottom to 980 ppm at the top. A chemical analysis also revealed higher concentrations of impurities (Fe, Rb, and Ba) in the top sample. Although specific impurities are expected to be the cause for radiation damage, a study of the other crystals allowed no conclusive answer as to the identity of the harmful impurities.

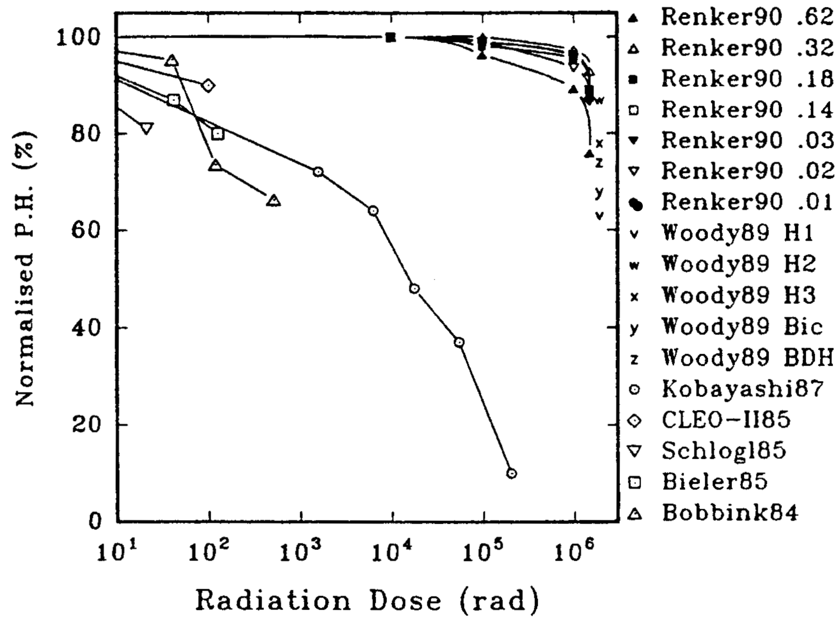


Figure 19. *Measured radiation damage to CsI crystals of various sizes.*

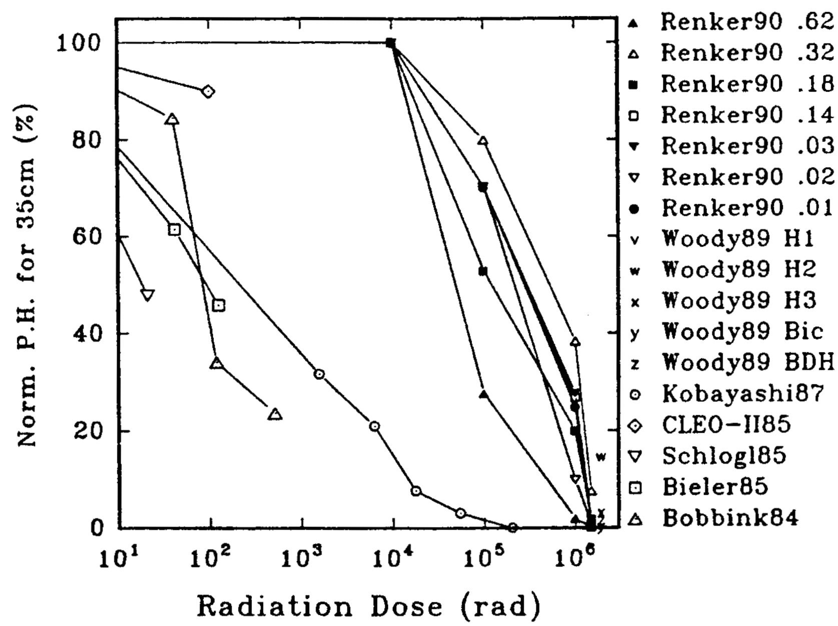


Figure 20. *CsI radiation damage measurements normalized to a 35 cm long crystal.*

	Bobbink 84	Bieler 85	Schögl 85
Size (cm^3)	$3 \times 3 \times 10$	$1.5 \times 1.5 \times 10$	$1.5 \times 1.5 \times 10$
Vendor	?	Korth	BDH
Rad. type	^{60}Co	^{137}Cs	PETRA
Rate (rad/hr)	6,000	1.8-3.5	.02
Exposure	5cm w/PMT	full w/oPMT	nonuniform
Readout	PMT (R1306)	PMT (R268)	PMT (R268)
Signal	^{137}Cs	^{137}Cs	^{207}Bi
	Kobayashi 87	CLEO II 85	
Size (cm^3)	$1 \times 1 \times 10$	$1.5 \times 5 \times 34$	
Vendor	Horiba	Horiba	
Rad. type	^{60}Co	^{60}Co	
Rate (rad/hr)	$\approx 10^4 - 10^5$	fast	
Exposure	full w/o PMT	full w/o PMT	
Readout	PMT (R594)	Photodiode	
Signal	^{137}Cs	^{137}Cs	
	Woody 89	Renker 90	
Size (cm^3)	2.54 dia $\times 1$	$1 \times 1 \times 1$	
Vendor	Hor.,Bic.,BDH	custom/MPI	
Tl doping (%)	0	.01-.62	
Rad. type	^{60}Co	^{60}Co	
Rate (rad/hr)	fast	fast	
Exposure	full	full w/ PD	
Readout	PMT (R594)	Photodiode	
Signal	^{137}Cs(?)	^{137}Cs	

Table 5. *Crystal radiation damage measurements*

In another study, several small crystals, which were grown at the MPI in Munich using the highest purity raw material, were tested. The thallium concentrations varied between 0.01% and 0.6%. The light yield and energy resolution measured with a ^{137}Cs source bebere irradiation are shown in Figure 23. The highest data point corresponds to 48,000 photons/MeV. While the light yield increases rapidly with the Tl concentration, the decay time of the scintillation light, 900 ns, is independent of concentration. All crystals were exposed to high-dose radiation at a special facility and then measured in the laboratory. The results are shown in Figure 24. For an exposure of 10^5 rad, only the crystal with the highest Tl concentration showed a 5% decrease in light yield. For a 10^6 rad exposure, slight reductions in light output were noticed in most crystals, while at 5 Mrad all crystals showed 10-20% effects.

Though these results look very encouraging, one has to keep in mind that these results were obtained with small crystals. It is therefore absolutely necessary to perform similar measurements with large crystals (15-20 *r.l.*). We are undertaking such a study.

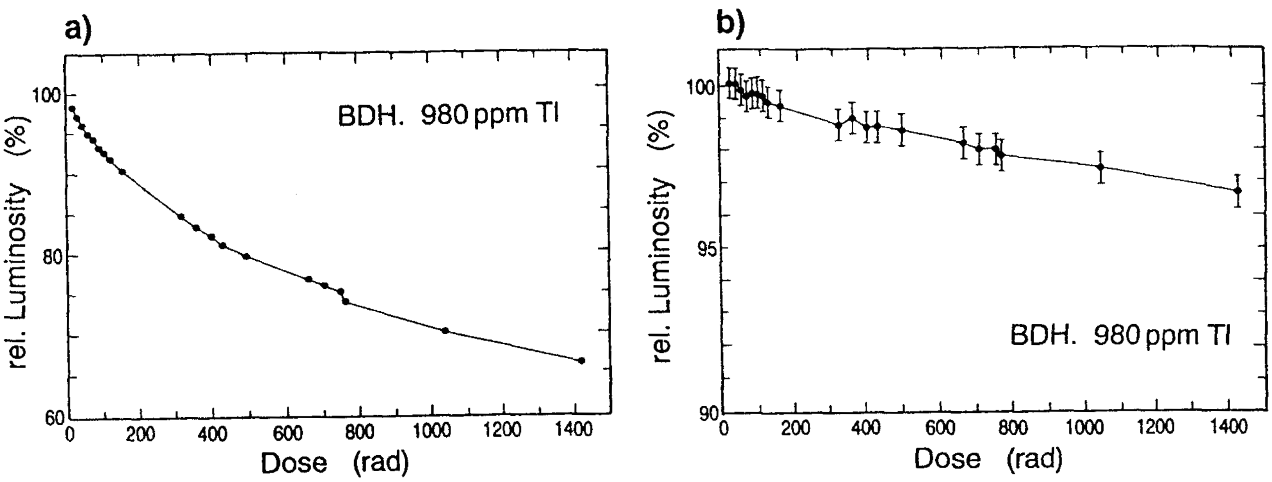


Figure 21. *Relative light yield and transmission for green light versus radiation dose for a BHD crystal cut off the top of an ingot.*

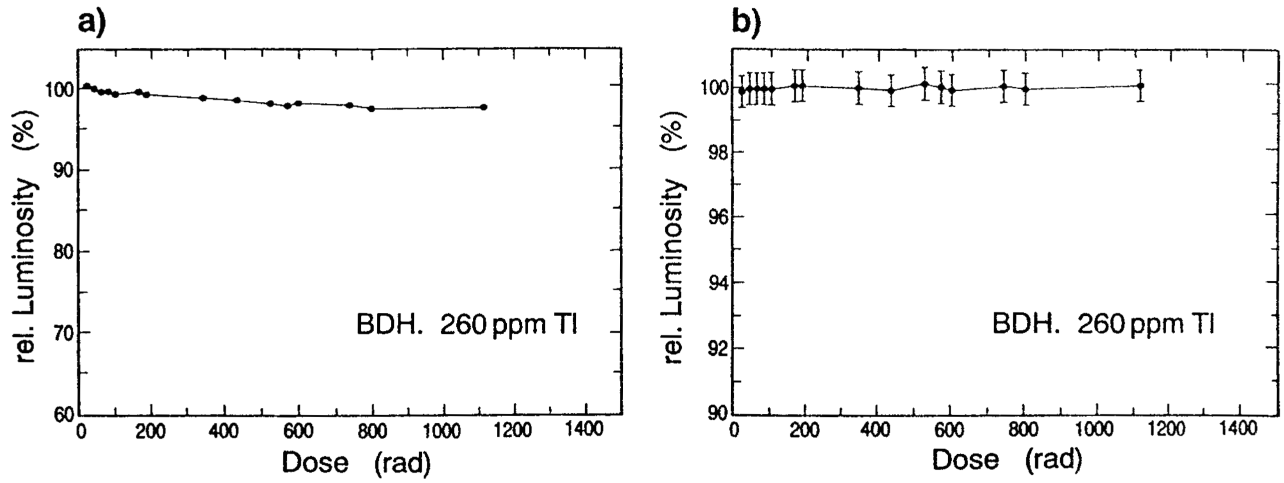


Figure 22. *Relative light yield and transmission for green light versus radiation dose for a BHD crystal cut off the bottom of the same ingot used in Figure 21.*

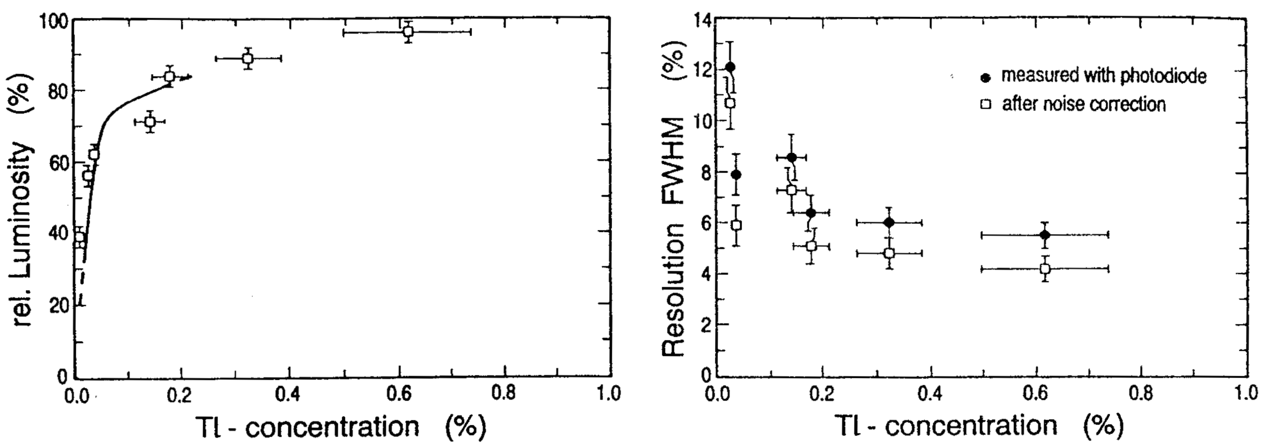


Figure 23. **a.** *Relative light yield as a function of Tl concentration for home-grown crystals;* **b.** *energy resolution as a function of Tl concentration for the home-grown crystals.*

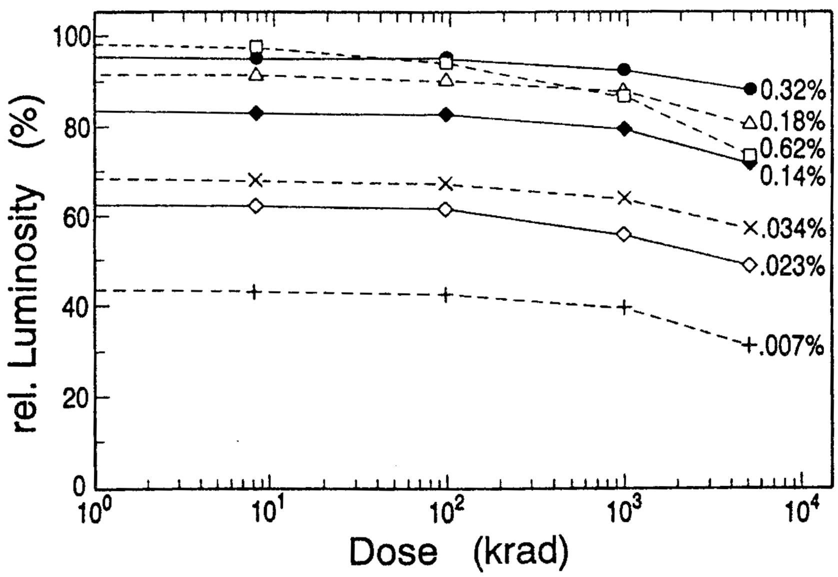


Figure 24. *Relative light yield versus radiation dose for the different home-grown crystals.*

3.13. Cost and Time Estimates

It is difficult to make firm cost estimates for a CsI(Tl) calorimeter at this stage, as the design is still fluid. A rough estimate is, however, shown in Table 6.

Table 6. *CsI(Tl) calorimeter cost estimate*

Item	Quantity	Cost
CsI(Tl)	12,000 crystals	\$20M total
Phototetrodes	12,000	\$200 each
Electronics	12,000	\$100/channnel
	TOTAL	\$24M

4. Liquid Krypton Calorimetry

Liquid krypton is being considered for the calorimeter because it may offer advantages in energy resolution, position resolution, ease of calibration, and cost compared to CsI.

It also may have disadvantages associated with the cryogenics, longer radiation length, and the electronics. Novosibirsk has built a 0.17 m^3 prototype and is now building a full size 14.6 m^3 liquid krypton calorimeter for the KEDR detector at the VEPP-5 *B* Factory.

4.1. Charge Collection

In liquid krypton 20.5 eV is required to produce an electron and a positive ion. CsI requires approximately 400 eV to produce a photoelectron. Thus, in principle, liquid krypton should have a better energy resolution due to improved statistics. Unfortunately, the velocity of the positive ions is very low (1 cm/s) compared to the electrons (Figure 25), and the positive ions cannot be collected in a reasonable integration time.

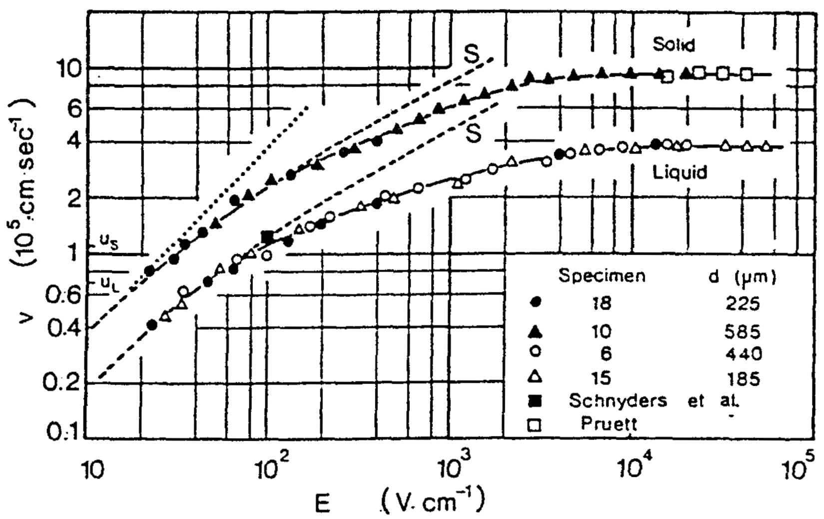


Figure 25. *Electron mobilities in Krypton (Phys. Rev.* **166**, *871 (1968))*

Assume an anode and a cathode plate with a spacing d between them. If a charge Q is deposited at some location between the plates, then the current in an external circuit connecting the two plates is

$$I = \frac{v_e Q}{d} + \frac{v_p Q}{d}$$

where the second term is negligible since $v_p << v_e$. If this charge was deposited near the cathode, then the current would flow for a time d/v_e as the charge drifted across the entire gap.

Thus the external circuit would integrate the entire charge Q. If the charge were deposited immediately in front of the anode, the current would flow for zero time and the external circuit would integrate zero charge. Thus the signal depends on the position in the gap at which the charge is deposited. Fluctuations in the location of energy deposition thus lead to a degraded energy resolution. There are two possible solutions to this problem. First, a grid (Frisch grid) can be placed at some distance such as $d/10$ away from the anode. The amplifier is connected between the grid and the anode and therefore records only the current flowing in this smaller gap. The amplifier thus sees the same signal, no matter where in the 90% gap the charge is deposited. There is still a position-dependent contribution from the 10% grid gap; this variation can be important. A second solution, which does not use a grid, is to integrate the signal only for 1/10 of the drift time. For 90% of the gap, the current is constant during the integration time, independent of where the charge was deposited in the 90% of the gap. We will consider only the shorter integration time procedure (also chosen by KEDR) as it provides faster rise times and doesn't introduce inactive grid material into the detector.

4.2. *Mechanical Design*

Figure 26 shows the central barrel calorimeter, endcaps and electronics placement for a model liquid krypton calorimeter structure.

The engineering of supports for the barrel and endcaps has not been addressed in this study. Cathodes and anodes are constructed on concentric cyclir within the barrel. An end view of three of the many cylinders comprising the radiator volume is shown in Figure 27.

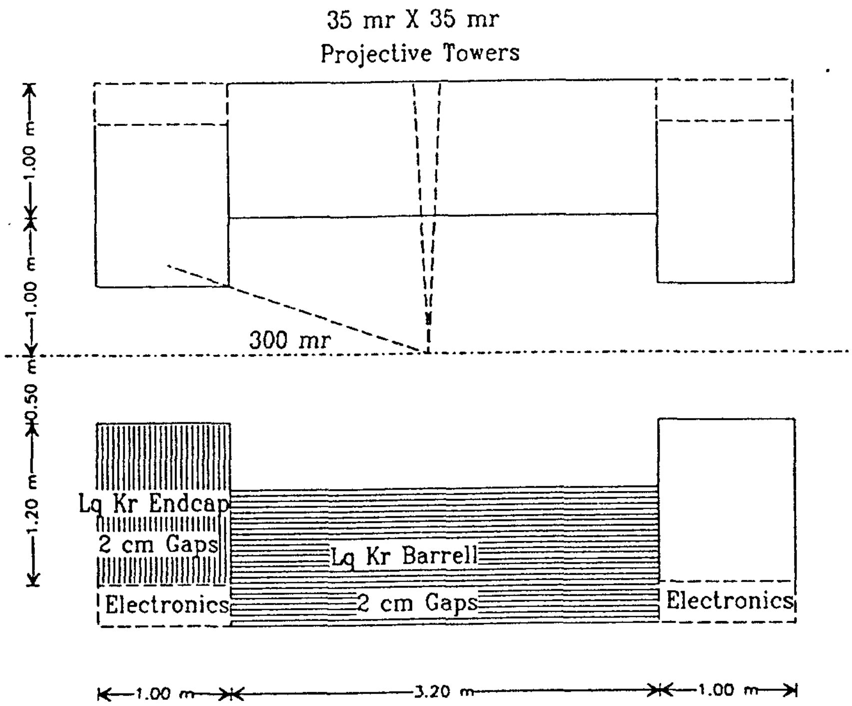


Figure 26. *Layout of the barrel calorimeter, endcaps, and electronics of a model liquid krypton calorimeter.*

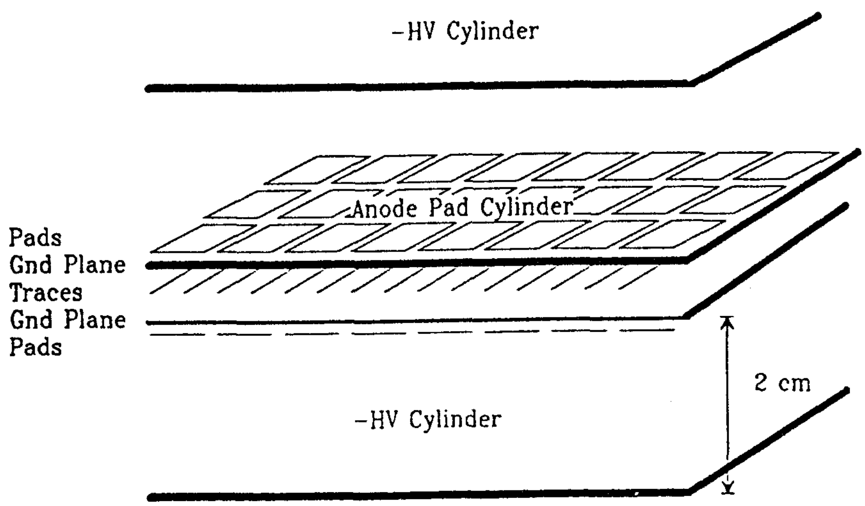


Figure 27. *End view of two cathode and one anode cylinders showing the layered structure of the anode cylinder.*

1) Every other trace is a ground.

2) Each trace connects through the ground planes to a pair of pads.

The traces from the pads on an anode cylinder run along the axis of the cylinder to the ends of the detector where the appropriate pads are connected together to form towers before going to the electronics. The pads are adjusted in size so that each tower covers the same size solid angle in the boosted center-of-mass of the asymmetric B Factory. The endcap cathodes and anodes (Figure 28) consist of a stack of disks with the pad traces buried in the anode sandwich as was done for the barrel.

In the endcap, the traces run radially to the outer circumference of each disk, where appropriate pads are connected into towers. Both the endcap signals and the barrel signals penetrate the cryostats in the region labelled electronics in Figure 26.

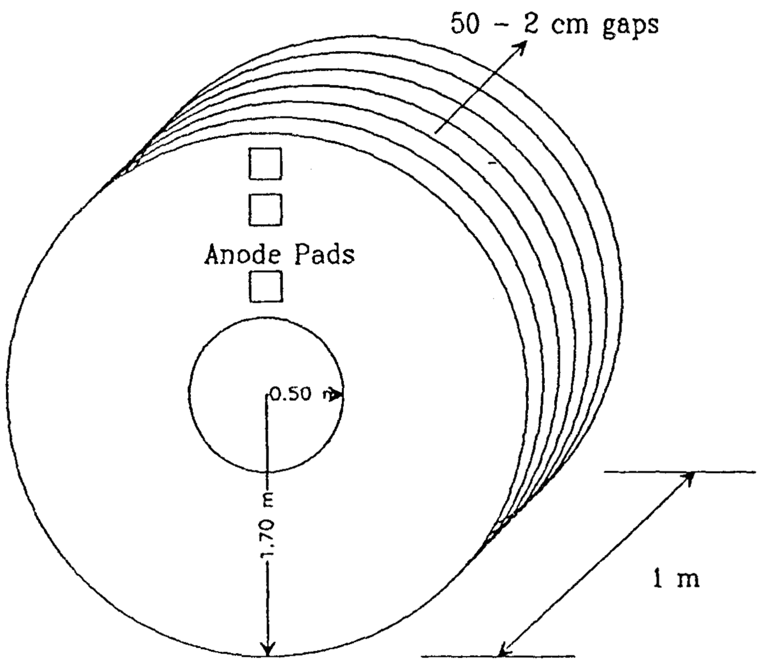


Figure 28. *A view of the endcap calorimeter.*

1) Traces come out radially under the pads.
(similar to the pad/trace sandwich of the barrel)

2) Traces are connected at the outer radius to form towers.

4.3. Cryostat

Liquid krypton must be maintained between $116°K$ (the freezing point) and $120°K$ (the boiling point). This is a problem very similar to that already solved many times for large liquid argon calorimeters. The cryostat must have a wall to contain the liquid krypton, a vacuum space insulated with many layers of aluminized mylar superinsulation, and a second wall to the atmosphere. The liquid krypton wall at a radius of 1 meter must be approximately 5/8" thick aluminum ($\sim$ 0.18 radiation lengths) to withstand the liquid krypton pressure head. The atmospheric wall at a radius of 1 meter must also be $\sim$ 5/8" thick (another 0.18 radiation lengths). This thickness is not dictated by the vacuum load, but rather by safety considerations in the event of failure of the liquid krypton wall, unless a means of very rapidly removing krypton gas from the vacuum space can be provided. We shall assume that any leak into this space will be rapidly pumped out to an alternate krypton storage tank, as it is important to recover the expensive krypton; venting to the atmosphere in an emergency is a last resort. Thus we will optimistically take the atmospheric wall to be 1/8" thick aluminum (.04 radiation lengths). The total assumed cryostat wall thickness of 3/4" (.21 radiation lengths) could be further reduced if a design using cryogenic composites were employed, an engineering task we did not undertake.

A further complication is that the cylinder containing the 72 metric tons of liquid krypton must be supported in a manner which doesn't introduce unacceptable heat leaks and which can survive the horizontal acceleration of a major earthquake. This problem was solved for the transverse degress of freedom of the liquid argon calorimeter of the SLD by cradling the cryogenic cylinder in two stainless steel slings whose ends extend out to atmosphere and connect to exterior supports. Axial restraint was provided by a group of shock-absorber-like "snubbers".

4.4. Krypton Storage

The 47 m^3 of liquid krypton is costly and must be recovered if the cryogenics of the barrel or endcaps fails. There must therefore be an alternative krypton storage vessel. The most foolproof scheme would be a large evacuated vessel (30,000 m^3) capable of storing the krypton in gaseous form at atmospheric pressure. This would occupy a cube 31 meters on a side. This is an absurdly large permanent structure (particularly if evacuated), but the emergency storage container could be a large collapsed balloon which expanded into an empty area. It would be dangerous to rely on pumps to compress the krypton into cylinders, since the cryostat failure might occur too rapidly. Another alternative would be a second cryogenic tank into which the liquid krypton could be dumped. Unfortunately, the liquid krypton could be boiling away due to a cracked cryostat, compromised vacuum in the insulating jacket, or loss of liquid nitrogen refrigerant. The gas would not recondense quickly in the second cryostat and the boiling liquid may not transfer rapidly enough. Design of a fail-safe, properly engineered capture system must be an integral part of any realistic liquid krypton calorimeter system.

4.5. Preamp Electronics

The best signal-to-noise ratio can be obtained with a charge sensitive preamplifier in which the signal is applied to the gate of an FET. The charge is stored on the feedback capacitor between the amplifier output and the FET gate. A Johnson noise voltage is produced across the channel resistance of the FET. This is seen as a charge fluctuation $\Delta Q_{noise} = C_{detector} \Delta V_{noise}$. Thus ΔQ_{noise} ($\propto$ Energy noise) is proportional to the capacitance of the detector. The anode and cathode plates are spaced 2 cm apart so as not to introduce too much inactive material into the krypton volume. The anode-cathode plate capacitance at this spacing is actually negligible compared to the capacitance to ground of the trace carrying the anode signal to the end of the cylinder. Assuming 50 gaps (25 traces) per tower leads to a capacitance of about 2 nF per tower. Using a 0.6μs shaping time and a 25 ohm channel resistance for the FET gives an rms noise of 1.6 fC, which converts to an equivalent noise of 2.7 MeV/tower. A longer shaping time could reduce the electronic noise further, but this would only increase the noise contribution from krypton radioactivity.

For a complete noise analysis of the liquid krypton calorimeter, see the following paper by P. Franzini

4.6. Radioactivity

Krypton has a radioactive isotope ^{85}Kr with a half life of 11 years, which is a β-emitter with a 0.67 MeV endpoint and an average energy of 0.25 MeV. This isotope is a fission product produced by reactors and nuclear bomb tests which now contaminates the atmospheric source of krypton. KEDR has measured 300 Decays/s-cm^3 of liquid krypton. Thus one tower ($10^{-4} \times 30$ m^3) has an rms decay noise of 0.14 MeV for a 0.6 μs integration time. A suggestion was made that South African mines are a potential source of uncontaminated krypton. It should be noted however, that in our design (and in essentially any practical design) electronic noise, not noise due to radioactivity, is the dominant contribution.

4.7. Energy Resolution

Energy deposited in:	Incident γ [MeV]		
	10	100	1000
Krypton (Uniform collection region)	.76	.89	.90
Krypton (Non-uniform collection region)	.05 ± .08	.05 ± .02	.05 ± .01
Anode+Cathode Cylinders(dead)	.05 ± .05	.05 ± .02	.05 ± .01
Cryostat Wall (3/4 inch Al)(dead)	.14 ± .14	.01 ± .01	.00 ± .00
Electronic Noise (9 T Ters=8.1 MeV)	.00 ± .81	.00 ± .08	.00 ± .01
Radioactivity (9 Towers=0.42 MeV)	.00 ± .04	.00 ± .00	.00 ± .00
Total *rms* noise	±.83	±.09	±.02
KEDR Measurement		±.05	±.02

Table 7. *The fraction of energy deposited in the various parts of the liquid krypton calorimeter*

The fraction of a photon's energy deposited in the various calorimeter media is given in Table 7 for 10, 100 and 1000

MeV photons. Also shown is the *r.m.s.* variation in the fractional deposition which, of course, contributes to the energy resolution. It is assumed that the central tower and its eight neighboring towers are summed to measure the energy of the photon shower.

4.8. *Cost*

With the dimensions shown in Figure 26, the barrel contains 30 m^3 and the endcaps contain 17 m^3 of liquid krypton. We have two examples of bulk purchases of krypton. First, SLAC purchased 20 liquid liters of Kr for \$420/liter from Liquid Air Products in 1985. The krypton is a byproduct of liquid nitrogen and oxygen production. Thus the price is very much determined by supply and demand rather than the actual cost of production. Secondly, KEDR purchased 14.6 m^3 for 490 Rubles/liquid liter (\$890/liquid liter at the official current floating exchange rate of \$1=0.55 Rubles). At \$420 /liquid liter, the 47 m^3 required would cost \$20 M. It is conceivable, though uncertain at this time, that the krypton could be obtained from the Soviet Union for less. Table 8 summarizes cost estimates for the various components of a liquid krypton calorimeter system.

Table 8. *Cost estimate for the liquid krypton calorimeter*

Item	Estimated Cost
Krypton (47 m^3 × \$42M/m^3)	\$20. M
Amps+ADCs(24,000 chan × \$100/chan)	\$ 2.4M
Cryostats + Supports	\$ 2.0 M
Emergency Krypton Storage	\$ 0.5M

5. CsI(Tl)/Liquid Krypton Comparison

Clearly, these two approaches to high-quality electromagnetic calorimetry require further detailed study. Nonetheless, at this point the virtues and problems of each approach are quite apparent. For CsI(Tl) the problems involve noise-free readout in a magnetic field, calibration and radiation hardness. For liquid krypton, they involve dewar design, reduction of electronic noise, handling of krypton and costs.

The comparison between CsI and liquid krypton is summarized in Table 9.

Topic	CsI(Tl)	Liquid Krypton
Energy resolution σ_E/E	3.9% @ 100 MeV; 2.0% @ 1 GeV	5.5% @ 100 MeV; 2.0% @ 1 GeV
Position resolution (photons)	0.5 cm	better
Entrance material	Particle ID	Particle ID + ~.2 *r.l.*
Calorimeter thickness (18 *r.l.*)	50 cm with PMTs	100 cm
Hermiticity	Good	Worse (cryogenics)
Radiation hardness	Shield desirable	Good
Calibration	Bhabhas+sources	Bhabhas+pulser
Analog energy sum for trigger	Easy	Difficult: coherent noise
Fast timing resolution	$\leq$ 10 ns	$\leq$ 600 ns
Cost	\$25M	\$25M
Cost certainty	High: CLEO II experience	Lower: Kr price

Table 10. *Comparison between CsI(Tl) and liquid krypton calorimeters*

ACKNOWLEDGEMENTS

G.E. is supported in part by a Heisenberg Fellowship of the Deutsche Forshungsgemeinschaft. Zhi-Rong Huang performed a number of the CsI test measurements.

References

1. A complete description of the Crystal Ball is given in E.D. Bloom and C. Peck, *Ann. Rev. Nucl. Part. Sci.* **33**, 143.
2. The Physics Program of a High- Luminosity Asymmetric *B* Factory at SLAC, D. Hitlin, ed., SLAC-353 (1989)
3. B. E. Blucher, *et al.*, *Nucl. Instr. Meth.* **A249**, 201 (1986).
4. S. Oide, presented at the KEK *B* Factory Workshop, October (1990)
5. E. Lorenz, private communication
6. C. Woody private communication
7. U. Kilgus, R. Kotthaus and E. Lange, MPI-PAE/ Exp.El. 228, MPI preprint (1990).
8. T. Skwarnicki, presented at the TRIUMF *B* Factory Workshop, (1991).
9. K. Braune and K. Königsmann, private communication.
10. E. Lorenz, private communication.
11. D. Renker, PSI Preprint (1990).
12. M. Lebeau, KEK Internal Report 89-19 (1989)
13. M.R. Farukhi, Presented at the Workshop on Positron Emission Tomography, Oct. 21-23, 1981, San Francisco. Quoted by Lebeau.
14. G.J.Bobbink, *et al.*, *Nucl. Instr. Meth.*, **227**, 470 (1984); Ch. Bieler, *et al.*, *Nucl. Instr. Meth.*, **A234**, 435 (1985); S. Schlogl, *et al.*, *Nucl. Instr. Meth.*,**A242**, 89 (1985); CLEO II updated proposal CLNS 85/634 (1985); M. Kobayashi, *et al,*, *Nucl. Instr. Meth.*, **A254**, 275 (1987); H. Schönbacher, CERN TIS Commission Report, TIS-RP/201 (1987); C.L.Woody, *et al.*, BNL preprint (1989); D. Renker, PSI preprint (1990) and private communication.

Signal and Noise in a Homogeneous, High Z, Liquid Calorimeter

Paolo Franzini

1. Charge and Current Signal from a Drifting Charge

1.1. Introduction

When ionization is produced in "noble" liquids in an electric field ~10 kV/cm, the freed electrons drift in the field at a few μsec/cm while the positive ions move much more slowly. The external current is given by the charge of the electrons in the gap divided by the time to drift across the whole gap.[1] While the initial value of the current is a measure of the total ionization and therefore of the energy deposit in the gap, the integral of the current in the external circuit depends on how the initial charge is distributed along the gap.[2] A measurement of the integrated external current, *i.e.* the charge, increases the sampling fluctuations. Measuring the current for a short time T, just after drift begins, reduces the fluctuations. Since only a fraction of the charge is measured, the electronics noise scales as $1/(T\sqrt{T})$, instead of the usual $1/\sqrt{T}$. In the following we derive expressions for signal and noise for this readout scheme.

1.2. Basic Relations

Consider an ion column of length l, ending at a distance d from the positive electrode, with uniform charge density and total charge Q, in a gap of width g with an appropriate field drifting the electrons with velocity v_d, as in Figure 1.

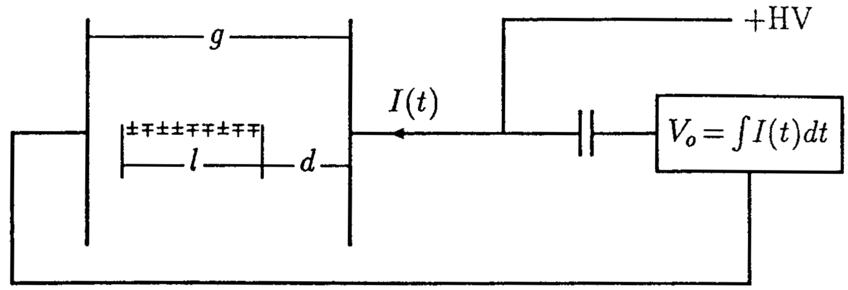


Figure 1. *Gap and ion column.*

The external current is given by:

$$I_{ext} = \frac{Q}{\tau} \qquad 0 < t < \frac{d}{v_d}$$
$$I_{ext} = \frac{Q}{\tau}(1 - \frac{tv_d}{l}) \qquad \frac{d}{v_d} < t < \frac{d+l}{v_d}$$

which we write as

$$I(t) = I_0 \left(\theta(t) - \theta(t - t_1)\frac{t - t_1}{\delta} + \theta(t - t_2)\frac{t - t_2}{\delta} \right), \quad (1.1)$$

and is shown in Figure 2, where $\theta(x) = 1$ for $x > 0$ and $= 0$ otherwise, t_1 is the time for electrons to drift a distance d, t_2 is the time to drift $d + l$ and $I_0 = Q/\tau$ with $\tau = g/v_d$, the time to drift across the whole gap and $\delta = t_2 - t_1$.

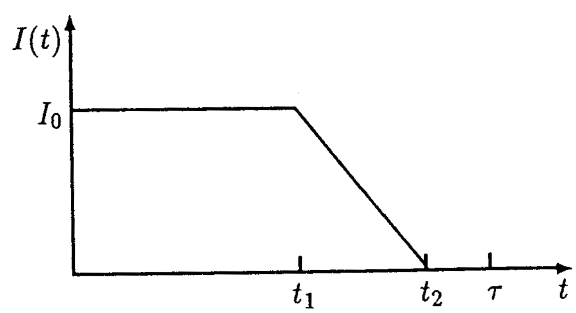


Figure 2. *Current vs time.*

While the peak current I_0 is a correct measure of the ionization and therefore of the energy deposited in the calorimeter, the charge from integrating the current is not. Depending where in the gap the energy is deposited the external integrated current, *i.e.* the charge, can vary between 0 and the freed charge Q. One should therefore measure the initial current, in practice first integrating and then differentiating. Below we derive the general form of the voltage signal *vs* time after such manipulations, in order, ultimately, to evaluate the noise. The Laplace transform of the current in equation (1.1) is:

$$I(s) = \frac{Q}{\tau}\Big(\frac{1}{s} - \frac{e^{-st_1}}{s^2\delta} + \frac{e^{-st_2}}{s^2\delta}\Big) \qquad (1.2).$$

If this current is fed to a charge sensitive amplifier, *i.e.* an ideal integrator for which $V_{out}(t) = (1/C)\int I_{in}(t)\mathrm{d}t$ or $V_{out}(s) = (1/sC)I_{in}(s)$, where C is the feedback capacitance, the output voltage signal is given by:

$$V(s) = \frac{Q}{s\tau C}\Big(\frac{1}{s} - \frac{e^{-st_1}}{s^2\delta} + \frac{e^{-st_2}}{s^2\delta}\Big). \qquad (1.3)$$

To control "parallel" and especially "series" noise a filter rejecting both low and high frequencies is necessary. Only one differentiation is necessary to remove parallel noise, but a second differentiation is necessary to measure the initial current rather than the integrated charge, which would result in an incorrect measurement of the electromagnetic shower energy. This second differentiation is best done by double sampling (also called *correlated sampling*) and taking the difference of the samples. A typical filter consists of one differentiation and two integrations, all with the same cutoff frequency a, $a = 1/rc$. The frequency response, $g(\omega)$, and the transfer function, $g(s)$, of the filter are:[1]

$$|g(\omega)|^2 = |V_{out}(\omega)/V_{in}(\omega)|^2 = \omega^2 a^4/((\omega^2 + a^2)^3),$$

$$a = 1/rc \qquad (1.4)$$

$$g(s) = V_{out}(s)/V_{in}(s) = sa^2/(s+a)^3 \qquad (1.5)$$

and the output signal, for a unit input step, is:

$$V(t) = \frac{(at)^2}{2}e\ -at, \qquad (1.6)$$

If the charge in the gap were collected in a time short with respect to $1/a$, the signal at the output of the filter would be given by equation (1.6). $V(t)$ peaks at $t = 2/a$, with a value $2\exp(-2) = 0.27$ (times the magnitude of input step). For an input current given by equation (1.1), the preamp response of equation (1.3) and the filter response of equation (1.5), the signal at the filter output is given by:

$$V_{out}(s) = \frac{Qa^2}{\tau C}\frac{1}{(s+a)^3}\left(\frac{1}{s} - \frac{e^{-st_1}}{s^2\delta} + \frac{e^{-st_2}}{s^2\delta}\right)$$

or

$$V_{out}(t) = \frac{Qa^2}{\tau C}\left(f(t) - \frac{g(t-t_1)\theta(t-t_1)}{\delta} + \frac{g(t-t_2)\theta(t-t_2)}{\delta}\right) \tag{1.7}$$

with

$$\begin{aligned} f(t) &= \frac{1}{a^3} - \left(\frac{t^2}{2a} + \frac{t}{a^2} + \frac{1}{a^3}\right)e^{-at} \\ g(t) &= \frac{t}{a^3} - \frac{3}{a^4} + \left(\frac{t^2}{2a^2} + \frac{2t}{a^3} + \frac{3}{a^4}\right)e^{-at} \end{aligned} \tag{1.8}$$

2. Noise

For the filter described above and a charge sensitive preamp using type 2SJ72–2SK147 FET's, the input equivalent noise charge is:[1]

$$Q_{noise} = 2650 \times e \times \frac{C_D}{1\ \text{nF}} \times \sqrt{\frac{1\ \mu\text{s}}{2/a}}, \tag{2.1}$$

where C_D is the detector (plus signal connection) capacitance. In krypton one electron-ion pair is produced for an energy loss of 24 eV. Thus for a 1 MeV energy loss, $Q_{(1\ \text{MeV})} = 41700\ e$, where e is the electron charge. If Q is uniformly distributed across the gap, along a line normal to the negative and positive electrodes, the measured charge is $Q/2$. If the drift time can be neglected, this is the charge to be compared with the noise in equation (2.1). Thus for $a = 2 \times 10^6\ \text{sec}^{-1}$ and C_D=1 nF, the energy noise is (1 MeV)×(2650)/(41700/2)=0.13 MeV. If we were to measure the initial current rather than the charge, we can make a guess about noise as in the following:

1. Choose the signal sampling time to be about one tenth of the total drift time, which should be large in order to keep the sampling time long and therefore the noise down.

2. The filter cutoff frequency a should be chosen to give short shaping times, say equal to the sampling time, to avoid large slopes in the signal to be sampled.

There is, of course, a practical constraint, the actual high voltage applied across the calorimeter gap. Few people like to go as high as 10 kV and those who don't are wise. Assuming 10 kV and saturated drift velocity ($E = 1$ kV/mm) gives $\tau \sim 3$ μs, in a gap of 1 cm. Choosing therefore $2/a = 0.1\tau = 0.3$ μs, increases the input equivalent charge noise by $1/\sqrt{.3} \sim 1.8$. In addition, since we are integrating the signal for ~1/10 of the time we measure about 1/10 of the charge, increasing the noise by another factor of 10/2=5. The net result is that the energy noise becomes $0.13 \times 1.8 \times 5 = 1.17$ MeV. This in fact is not so bad for 1 nF detector capacitance!

Before continuing, one should mention one way to increase the total drift time: operate at non-saturated drift velocity. This might, however, be even less wise than going to higher voltage. At low fields electron attachment rapidly becomes a very severe problem. And at contamination levels of 0.01 ppm and with non-saturated charge collection, it is not even conceivable to attempt to measure the charge loss, since it is going to be different everywhere, in a large calorimeter.

d mm	$V(t = 0.6\ \mu\text{s})$ V	$V(t = 0.6\ \mu\text{s})$ normalized
0.5	0.01625	0.503
1.5	0.03161	0.978
2.5	0.03233	1.0
$\vdots$	$\vdots$	$\vdots$
9.5	0.03233	1.0

Table 1. *Signal for 0.6 μs shaping and 3 μs drift times.*

Assuming a 10 mm gap, a preamp with a feedback capacitor C, a filter with two integrations and $a = 3.333 \times 10\ 6$ s -1, we find in Table 1 the signal at 0.6 μs for unit Q/C freed along a line 0.25 mm long, ending at d=0.5, 1.5, ... 9.5 mm from the positive electrode. (If the charge is freed at the positive electrode, the initial current is the same, but since the current lasts for zero time, no signal is measurable.)

Note that with this choice of parameters, only 10% of the gap has a severe signal loss. For uniform charge deposition, the normalized response is 0.941. More important however, is the fluctuation in the response. For an input (energy) of 1, deposited uniformly and randomly, at a point in the center of a 1 mm layer of liquid, the normalized output is 0.948, with an rms spread of 0.148. This result is not very accurate but not terribly wrong either. A better calculation requires merging EGS with the formulae above.

Finally, we can come back to noise. The noise increase, over the value 0.13 MeV above, is

$$(1/.948)(0.157/0.0323)\sqrt{1/0.6} = 6.61\,,$$

giving an rms noise of 0.86 MeV. In practical applications it is necessary to measure the baseline before the signal of interest appears, the signal, at the appropriate $t = 0.6$ μs and take the difference of the two measurements. Because of the necessary choice of the filter peaking time for an input step, the two samples are almost completely uncorrelated. The difference therefore has an rms noise of $\sqrt{2} \times 0.86 = 1.2$ MeV. An exact calculation gives about 10% more noise, which we ignore. Scaling of the noise is as $a\sqrt{a}$, because preamp noise grows as $\sqrt{a}$ and the filter output signal decreases as $1/a$, the latter for constant gap width and drift velocity.

3. A Projective EM Calorimeter

3.1. *Tower parameters*

I have derived the noise for a relative large capacitance, because I wanted to find whether a standard, many gap, projective structure could deliver reasonable resolution without too much degradation from noise. The read–out planes could be sandwiches of etched aluminized plastic and foam. To properly shield signal lines from pads, four foam layers are necessary and five plastic layers: two outer ones with charge collecting pads, two ground planes and a center layer with signal traces. For a 20 radiation length λ_0 thick liquid krypton calorimeter, the sandwiches add up to about $0.05\lambda_0$ of inert material.

We assume a structure with sandwiches carrying pads on both sides, alternating with high voltage electrodes; each pad thus views too gaps. For 0.5 mm foam thickness the capacitance to ground of a 3×3 cm^2 double pad is 31.7 pF. Adding a trace, 0.5 mm wide and 50 cm long contributes 8.8 pF, for a total of 40.5 pf per pad. From 1000/40=25, we get that 1 nF corresponds to 2×25 gaps of liquid. For 1 cm gaps and Kr, this corresponds to 10.6 radiation lengths. For a projective tower the area increases, approximately doubling. Thus, 3×3 cm^2 at the entrance becomes 4.5×4.5 cm^{2}on average. Four sections with 12×2 cm double gaps each, remain below 1 nF each and correspond to 20.4 radiation lengths of krypton. Noise values per tower are $\sqrt{4} = 2$ times larger than for the 1 nF example above. Thus the total noise due to the preamp is 2.4 MeV per tower.

3.2. *A crude estimate of the signal fluctuations.*

While, as I have mentioned, a proper estimate requires marrying EGS and equation (1.7), I will estimate the fluctuations using the last column in Table 1, assuming that energy (charge) is deposited in small (1 MeV) clumps, at random, in 1 mm liquid layers. Then, for a 100 MeV shower, I find that one measures on average 94.8 MeV, with an rms spread of 1.48 MeV or 1.6%. The average is, of course, just the sum of the weights in Table 1 and is also the measured signal for a uniform column of ionization from one electrode to the other. The estimate of the fluctuation is, I believe, an over-estimate, because at 100 MeV there are a few electrons crossing longer portions of the gap. If the energy clumps in the shower are smaller than I have chosen, the fluctuations are again smaller. Typically 1/2 of a shower energy is in electrons and photons of less than 1 MeV. Finally the fractional fluctuation scales as $1/\sqrt{E}$: at 20 MeV, the fluctuation is still only 3.3% or 0.7 MeV, quite negligible with respect to noise.

3.3. *Radioactive noise*

I use the KEDR[3] quoted value of 300 decays/s/cc. In terms of the signal V(t) produced by a randomly occurring, fixed energy deposit with average frequency f, the mean square fluctuation F^2 of the signal is given by ($\langle V(t)\rangle \equiv 0$):[1]

$$F^2 = \langle (V(t))^2 \rangle = \lim_{T\to\infty} \frac{fT\int_0^T (V(t))^2\,dt}{T} = f\int_0^\infty (V(t))^2\,dt. \tag{3.1}$$

In the equation above one must use $V(t) = V_{out}(t) - V_{out}(t - T_s)$, where T_s is the sampling interval and $V_{out}(t)$ is given in equations (1.7) and (1.8). It's quite straightforward to do a better calculation, by I just assume that all decays happen in the middle of the gap and deposit 0.33 MeV, one half of the end point of $85Kr$ β decay spectrum. The thing one must pay attention is the energy-to-voltage relation. I write:

$$\begin{aligned} \text{Signal from shower} &= K\frac{Q}{C} = K\frac{E\mathcal{R}}{C} \\ \text{Signal noise from radioactivity} &= F\frac{Q}{C} = F\frac{E\mathcal{R}}{C} \end{aligned} \tag{3.2}$$

where F is given in equation (3.1) above, K is the average response to an em shower and $\mathcal{R}$ is the responsivity of the calorimeter, in coulomb/MeV. E and Q are the energy deposit in the liquid and the charge producing the signal. In this way I find that the total noise from radioactivity for the KEDR test calorimeter should be 2.6 MeV. They give 0.6 MeV per section and therefore 1.8 to 2.5 for the total, depending how they count sections. Pretty good, since I do not know what filter they use. In Table 2, I give values of various parameters, for different shaping times, $2/a$.

Table 2. *Signal and noise contribution for a tower, at 100 MeV.*

$2/a$ μs	rms sampling fluctuation %	Electronics Noise MeV	Radioactivity Noise MeV	Max signal in units of Q/C	Average sig. in units of Q/C
0.3	0.33	6.9	0.19	0.0162	0.0160
0.6	1.5	2.4	0.26	0.0323	0.0306
0.75	1.88	1.7	0.29	0.0404	0.0376
0.9	2.2	1.3	0.29	0.0485	0.0443
$\gamma\gamma$drift time	6	< 0.26	> 0.5	0.270	0.135

4. Conclusions

I have shown that only a moderate amount of fast shaping is necessary, in order to avoid fluctuation in the signal from a homogeneous calorimeter, due to differences in drift paths, thus alleviating problems of cross talk in signal traces. The use of drift fields as low as 100 volts per mm, with the correspondingly severe demand on liquid purity can also be avoided. I have also shown that structures with good shielding do not imply large capacitances and that capacitances in the nF range result in quite acceptable noise levels, better than scintillating crystals with silicon diode readout. I conclude that it is possible to construct a conventional geometry calorimeter with projective tower and adequate resolution and noise, even adding the capacitance of the cables carrying the signals to a feedthrough port. Limitations to the performance of a homogeneous calorimeter using a cryogenic liquid are most likely to come from the wall thickness of the cryostat.

References

1. P. Franzini, Noise and Pile-up in Liquid Sampling Calorimeters, in *Proceedings of the 1987 Berkeley Workshop on Experiments, etc. for the Supercollider*, R. Donaldson and M. G. D. Gilchriese eds., World Scientific, Singapore, 1988, p. 636.
2. This was well known since the days of ionization chambers. For a cure (impractical in this case) see O. Frisch, British Atomic Energy Report BT–49 (1944), unpublished.
3. D. Hitlin, G. Godfrey, private communication.

COMPUTING

Report of the Computing Group

A. Boyarski, T. Glanzman, F. A. Harris
and F. C. Porter

1. Introduction

The triggering needs, high event rate, and large data sets at a B Factory imply significant computing requirements. The anticipated event rate is approximately two orders of magnitude higher than existing e^+e^- experiment rates. While high event rates have already been successfully dealt with in hadronic collisions, the problem at a B Factory is somewhat different, because of the large fraction of triggered events which are useful for physics. In this chapter, we make estimates for the event rate and size and the resulting CPU requirements. With these data handling requirements, we then produce a model for the computing of the B Factory experiment. We attempt to give both an "existence proof" given current technology, and to speculate on how things will evolve in the next few years. Finally, we offer some remarks concerning software issues and required development.

The computing environment at a B Factory must be capable of dealing with relatively high data rates and very large data sets.

2. Event Rates

We are interested here in the rate at which events pass trigger and online cuts and hence must be saved on permanent storage. A significant fraction of this rate will be due to e^+e^- collisions; we must therefore estimate these rates. We will then add a contribution from backgrounds, based on an estimate of the maximum acceptable background rate.

We include the following physical processes in our event rate calculation:

1. Hadronic events: We'll base this contribution on the $\Upsilon(4S)$, with a total hadronic cross section of about 4 nb.
2. Bhabha events: The Bhabha cross section, integrated up to a given value of $\pm\cos\theta^{cm}_{max}$, is given approximately

by:

$$\sigma_{e^+e^- \to e^+e^-} \approx \frac{8\pi\alpha^2}{E_{\mathrm{cm}}^2} \frac{1}{1 - \cos\theta_{\mathrm{max}}^{\mathrm{cm}}}$$

This cross section is about 100 nb for $\cos\theta_{\mathrm{max}}^{\mathrm{cm}} = 0.95$.

3. Muons and taus: The lowest order cross section for $\mu^+\mu^-$ and $\tau^+\tau^-$ pair production is $\sigma_{\mu\mu} \approx \sigma_{\tau\tau} \approx 0.8\,\mathrm{nb}$.
4. The two-photon physics rate can also be a large contribution, depending on physics goals, and hence on how one adjusts the trigger. We will assume here that this rate is adjusted so that it does not dominate the single photon hadronic rate.

We make the further assumption that it is unnecessary to record all of the Bhabha events within the detector solid angle. Instead, the Bhabhas accumulated will be prescaled, probably in a $\cos\theta$-dependent manner, so that a fairly flat distribution of events in the detector will be recorded. Thus, we will use $\sigma_{\mathrm{TOT}} \approx 10\,\mathrm{nb}$ as our working physics cross section.

A rate of at least 100 events/second to storage is a design criterion.

A machine running at $3 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1} = 3\,\mathrm{nb}^{-1}\mathrm{s}^{-1}$ has a 30 Hz trigger rate for physics under these assumptions. With backgrounds and possibly luminosity upgrades, the system should be designed to be capable of handling a data rate of at least 100 events/second to storage. To provide for further luminosity improvements, a 100 event/second design should have the flexibility to facilitate upgrades beyond this initial criterion.

3. Event Size

We have made an estimate of the typical event size, which is required to determine the data transfer rates and ultimate data set size. This estimate is based on hadronic events plus machine backgrounds, including integration over several crossings. There is, of course, a great deal of uncertainty in this – the numbers should be interpreted as a reasonable design goal. The components of this estimate are shown in Table 1.

Note that this estimate is dominated by background hits. This suggests that significant gains could be made by providing removal of background hits using elementary pattern recognition in the path from the data acquisition hardware to the permanent storage. Presumably in such a strategy one would prescale a fraction of the full events to permit checks and background studies. Since the backgrounds are difficult to predict accurately, and could turn out in reality to be worse than expected, it would be wise to be prepared with some capability of this sort.

System	Hadron (kbytes/event)	Background	Total
Vertex	1	10	11
Drift chamber	2	1	3
Particle ID (CRID)	2.5	2.5	5
Calorimeter	$\lesssim 2$	2	4
All else	2		2
Total	7.5 2	15.5	25

Table 1. *We estimate an event size of 25 kbytes.*

The analysis of the data for a particular physics topic will quite likely be aided by the use of a more condensed "mini-DST" format. For a "typical" $B\bar{B}$ event with ten charged tracks and ten photons, we may estimate the mini-DST event size as follows: For each charged track, we measure its momentum, direction, and location of closest approach to the beam line. This takes five variables. The error matrix for these measurements is, in general, another 15 quantities. In addition, there will be information associated with particle identification, energy in the calorimeter, and possibly bookkeeping such as vertex links. An estimate of a total of 30 words/charged track is plausible. Similarly, we estimate 15 words/neutral track. Thus, we have $10 \times 45 \times 4 = 1800$ bytes per event on the mini-DST. Not all events are $B\bar{B}$, of course, but it seems likely that most physics mini-DST's, with reason-

An event size of $\sim$ 2 kbytes is anticipated for mini-DST's.

ably complete event information, will still require this amount of storage.

This can lead to rather large datasets for such things as inclusive analyses. Some of our estimates in the following discussion on network bandwidth will assume the use of such a mini-DST for large sample analysis. It is possible, however, that for certain physics, more heroic measures, such as compressed storage formats and further elimination of information, will be used.

4. Data Set Size

A raw data set of the order of 25 Tbytes/year should be anticipated. The initial storage capacity goal should be 50 Tbytes.

If we adopt the now-common convention that a calendar year of data-taking corresponds to 10^7 s, then at a peak luminosity of $3 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, the 100 Hz event rate corresponds to a dataset of 10^9 events accumulated in one year at a B Factory. At 25 kbytes/event, this implies a storage requirement of 25 Tbytes.

As 25 Tbytes is a substantial amount of data, we have considered possible ways to achieve reduction without seriously affecting the information. One possibility, mentioned in the previous section, is to do sufficient pattern recognition on the "good" events to further suppress background hits. Also, one could try to extract the essential information from the raw hits into a compressed form before storage. We estimate that it may be possible to reduce the dataset size to around 10 Tbytes/year, but probably not to much less than that.

It should be remarked that the various DST's and Monte Carlo data sets will also be significant, probably contributing a few additional terabytes of data per year. For example, a "mini-DST" with 10^8 events at 2 kbyte/event requires 200 Gbytes. While Monte Carlo datasets may also be large, we do not expect them to be comparable to the actual data, since Monte Carlo distributions so subtle as to require such enormous statistics are probably not to be trusted, and hence are not worth generating (independent of any CPU considerations).

5. CPU Requirements

We make an estimate of the CPU computational power required by drawing on the experience of the Mark II experiment at *PEP*. We will assume that a *B* Factory event and a hadronic event at *PEP* impose similar computing needs. Our discussion will be in terms of MIPS (for millions of instructions per second – one MIPS is one VAX 11/780 unit of processing power, as measured with the Dhrystone test). We caution that actual comparisons of computers on high energy code may not scale according to this measure. Nevertheless, it is a convenient unit for the purposes of this analysis.

A Mark II hadronic event at *PEP* requires of order 1 second to process on a 25 MIPS CPU (one CPU of a 3090 mainframe). By "process" we mean taking the raw chamber, TOF, calorimeter, *etc.*, data, applying any corrections required, performing pattern recognition, and fitting of points to obtain reconstructed charged and neutral directions, momenta, particle identification, and vertex information (this is known as "PASS 1" in Mark II parlance). Thus, we may express the computing load as approximately 25 MIPS–s/event. Taking into account Monte Carlo generation, reprocessing, and various data analysis activities, this number must be multiplied by a factor of about three to obtain the total CPU requirement of the Mark II at PEP.

For a *B* Factory, we will also assume a 25 MIPS–s/event requirement for "PASS 1" analysis, and the same factor of three for Monte Carlo and physics analysis. In addition, during data acquisition, significant computing power is required to build and filter (Level 3 trigger) events, pass the data to tertiary storage, and for online event sampling. We'll estimate this additional load at another 25 MIPS–s/event (this is based on the desire to maintain an online sampling fraction of order 10%, even as the online event rate approaches 1 kHz). Thus, at a *B* Factory with an event rate of 100 events/second, the total CPU power required to keep up with the data is 100 events/s $\times$ 100 MIPS–s/event = 10^4 MIPS.*

We estimate a 10^4 MIPS CPU requirement.

⋆ Since the 100 events/second number is a peak rate (or, alterna-

6. I/O Bandwidth Requirements

The high event rates and large data sets suggest the need for high I/O capacity, both for transferring the original data to storage, and for later offline analysis. We have made estimates of these demands in order to evaluate the networking and computer bus requirements.

The data rate from the Level 2 trigger is estimated at 25 Mbytes/second.

With an event size of 25 kbytes and a rate of 1 kHz from the Level 2 trigger, an aggregate bandwidth of 25 Mbytes per second is required in the online system up to the level 3 trigger. Any devices between Levels 2 and 3 must be able to handle this average rate both at input and output, with a comfortable margin for peaks (possibly achieved with buffering). This will have to be done with individual or parallel channels of sufficient capacity.

The data rate from the detector to tape is estimated at 2.5 Mbytes/second.

From Level 3 to storage, the design rate is 2.5 Mbytes/-second (100 Hz at 25 kbytes/event). This is within the specification of an FDDI fiber network (2×100 Mbits/s). Assuming we keep up with the data-taking with our PASS 1 analysis, a similar additional capacity, from and to storage, will be required for this activity.

The physics analysis, Monte Carlo simulations, *etc.*, will require still further network bandwidth. The desire for fast turn-around analysis of large datasets imposes significant I/O demands. To get an idea of this requirement, let us suppose that we have a dataset with 10^8 events in "mini-DST" format, with 2 kbytes/event. We assume a distributed computing environment. We take it to be acceptable to do an analysis on such a DST with a turn-around time of one week. This translates into an average transfer rate of about 1/3 Mbyte/second for the week. Several people could be expected to be looking at such a DST at a time; this is, furthermore, only one of the various activities to be anticipated. It is thus clear that one or more FDDI-type networks will be required to handle the physics analysis demands.

---

tively, is a continuous rate for 10^7 s out of a year), this calculation is at least superficially conservative. However, we believe this is justified, because: (*i*) It is expected that the CPU load from interactive computing will continue to increase as evolution towards sophisticated graphical interfaces progresses; (*ii*) The load on the workstation farm envisioned due to the networked computing environment may be significant; (*iii*) There must be some flexibility occasionally to redo the "PASS 1" analysis of the data.

7. Comparison of Requirements with Fermilab and SSC

To place the computing needs associated with a future B Factory in perspective, we have checked on the computing experience of CDF for their previous run and their requirements for the upcoming runs.[1]

7.1. *Results of the 1988-1989 CDF Run*

The following table summarizes details of the last CDF run:

Run duration	44 weeks
Useful physics running	33 weeks
Trigger rate to tape	1 Hz
Nine-track physics tapes	5500
Events per tape	1150
Total events	6 million
Average event size (raw)	115 kbytes
Full DST event size	185 kbytes
Mini-DST event size (in process)	27 kbytes
Micro-DST event size (in process)	14 kbytes
Initial capacity (ACP's)	100 MIPS
Final capacity (ACP's + 3100's)	150 MIPS
Time to reprocess full sample	6 months
Time to reconstruct event (most time is for tracking)	250 MIPS-s
Disk space available	20 Gbytes

The raw tapes were processed on the average three times during the run to check the detector, strip off about 10% of the events for high p_t, W and Z physics, and other special purposes. Most of the published results so far have come from this 10% data sample. When the full sample was processed, seven different data sets (DST's) were created, each corresponding to different physics topics. A simple data base is used to keep track of files associated with each of the seven data sets at each analysis stage. Event processing is handled by the Production Manager, written in DCL, and is driven by the database. The Production Manager automatically updates the data base when new files (or tapes) are written. Information kept in the database includes for each file the creation date, computer used, tape, tape drive used, quality, and a large comment field. The smallest DST set consists of 500 tapes; the largest of 2500 tapes. Much use has been made of 8mm helical scan tapes, although the numbers of tapes mentioned here are in terms of the number of 9 track tapes. In the future, even the raw tapes will be 8 mm.

Most of the production running was done on ACP farms or on the DEC VAXstation 3100 farm, all at Fermilab. Much of the physics analysis is done at Fermilab, either locally or remotely, usually reading disk datasets. DST's (8 mm) have also been sent to other institutions for analysis.

7.2. Expectations for Summer 1991 CDF Run

High priority physics to tape	1 Hz
Possible general physics to tape	10 Hz
Event size	200 kbytes
High Priority data sample	1–2 Tbyte
Possible general physics	10 Tbytes
Data rate	200 kbytes/s – 2 Mbytes/s
CPU requirement – reconstruction	500–750 MIPS
Disk space requirement	200–300 Gbytes

CDF is currently putting a major effort into switching their system to UNIX, since this is where the computing expansion is to be found. The total computing capacity expected at Fermilab in 1992 is several thousand MIPS, in UNIX processor farms. Currently Fermilab has 300 MIPS and is in the process of ordering another 1100 MIPS of UNIX workstations. Silicon Graphics workstations are being purchased, but other workstations are being evaluated, since Fermilab does not want to be a single-vendor shop.

A major problem (as for a B Factory) is the size of the data sample. E769 is testing a Summus 8mm tape jukebox, which holds 55 tapes and contains two drives. A jukebox will eventually hold about 1/4 Tbyte of data and will cost under

$50k (in the future, 8 mm tapes are expected to hold about 5 Gbytes and to have transfer rates of 500 kbytes/second). Each jukebox drive will be able to process about 5 tapes per day. One possible mode of operation is to play through the DST's continuously and let people use the continuous data stream whenever they are ready.

7.3. *Comparison with B Factory*

Comparing our B Factory projections with expectations for CDF's 1991 run, we find that: (*i*) The B Factory event rate is an order of magnitude higher; (*ii*) The B Factory event size is an order of magnitude smaller; (*iii*) The B Factory dataset size is comparable for a similar calendar run period, but the total CDF storage requirement is not as large because of the relatively low running duty cycle; (*iv*) The B Factory CPU requirement is at least an order of magnitude larger. The luminosity at Fermilab will continue to improve in the future, so that the CDF needs will approach those of the B Factory in all respects before the B Factory is running. We conclude that CDF is already facing the scale of problems that will be encountered at a B Factory.

The computing requirements of a B Factory are similar to those to be faced by CDF in the near future.

7.4. *SSC Computing Projections*

It is also interesting to compare the B Factory computing requirements with projections for a major SSC experiment. Areti[2] quotes a projected CPU requirement at the SSC of one million MIPS, two orders of magnitude larger than our B Factory estimate. The SSC event size is expected to be approximately 1 Mbyte, to be compared with the B Factory estimate of 25 kbytes.

Areti proposes an online processing farm consisting of 5000 nodes of single board processors, each providing 200 MIPS of computing. Together with the estimate of 100 MIPS-second/event for online reconstruction, this implies that such a farm can handle up to 10^4 events/second, or 10^{10} bytes/-second, provided each processor can accommodate 2 Mbytes/-second transfer rates. Lankford[3] expects that the events will be filtered by the online farm so that a rate of 10-100 Hz to tape is realized. At a 10 Hz rate to tape, and a 1 Mbyte event size, this gives a transfer rate of 10 Mbytes/second, or a dataset size of 100 Tbytes in our nominal year of 10^7 seconds.

Cottrell[4] provides some estimates for the Monte Carlo simulations for the SSC. He assumes that a typical simulation takes 5000 MIPS-second/event, and generates a 2 Mbyte event. Again, due to the lower energy and simpler detector, the expectations for a B Factory are considerably smaller. Cottrell also notes that the yearly offline storage requirement at the SSC will be "several tens of Tbytes". This number is, in fact, comparable with our B Factory estimate, a reflection of the fact that most of the data rate at the SSC may be discarded as relatively uninteresting low-p_T physics, while at a B Factory a large fraction of the online triggers are worth storing.

The computing demands for an SSC experiment will be greater than those required at a B Factory.

We conclude that, with the possible exception of the data storage requirements, computing for a B Factory presents a considerably less daunting prospect than for a major SSC experiment.

8. A Possible Computing Model

The traditional mode of analyzing current e^+e^- experiments largely on a mainframe computer does not appear practical for the large CPU requirements of a B Factory experiment. In addition, the large datasets envisioned impose substantial requirements on both storage and data transfer capacity. In this section, we present a possible computing model, and make some estimates concerning practicality.

A CPU requirement of 10^4 MIPS is beyond the present or near-term extrapolations of mainframe capabilities, such as the current six processor IBM 3090-600J mainframe with around 180 MIPS. This implies that a solution using only mainframes is not affordable (the IBM 3090-600J costs nearly \$15 million,* though price declines fairly quickly for used main-

---

⋆ IBM has just announced its new mainframe series, with a top-of-the-line performance of 600 MIPS and 2.5 Gbyte/second internal bus. The list price is around \$20 million.

frames). We are thus forced to consider distributing the computing to a farm of workstations with a much higher CPU performance/price ratio. We consider the following example to set the scale in terms of current technology. The IBM RISC/6000 model 320 computer has a rating of 27.5 MIPS (and a floating-point rating of 7.4 MFLOPS). Thus, if we were to start ordering our farm today, and we decided to use these machines, we would have to order around 360 of them to meet our requirement (at least superficially). At a base list price of about $13,000, this implies a purchase price of order $5 million. The expectation is that the CPU power available in such machines will grow by a factor of order two each year for the next five years, for modest growth in price. Thus, it is reasonable to assert that the pure CPU power can be met without being the dominant cost factor.

The CPU requirement at a B Factory is too large for a mainframe model to be practical. Instead, we consider a distributed model with high-performance workstations.

Figure 1 is a block diagram of a computing model for a *B* Factory, including both online data acquisition and offline analysis. To obtain the required CPU power, it is based on a multi-workstation environment. Each of the square-cornered boxes in this picture represents a workstation. A desirable feature seems to be to have them all basically interchangeable, at least in terms of operating system, software facilities, and networking support. We note immediately that the box labeled "storage" might contain a mainframe, functioning, at a minimum, to manage the central storage.

Starting at the detector, we see lines indicating data transfer to a set of computers. These first general purpose commercial computers are connected to the data acquisition electronics at the subsystem level. Thus, we envision a computer devoted to the vertex detector subsystem, another to the drift chamber, another for the calorimeter, one for detector environmental monitoring, and so forth. These computers are responsible for transferring the subsystem data to the next higher level, possibly making corrections and reducing the data in appropriate ways in the process. In addition, these computers may be used for monitoring of subsystem performance and subsystem calibration. A design criterion is that these computers should be able to respond to subsystem data at a 1000 triggers/second rate from the Level 2 trigger.

A computer is envisioned for each detector subsystem, capable of responding at the Level 2 trigger rate.

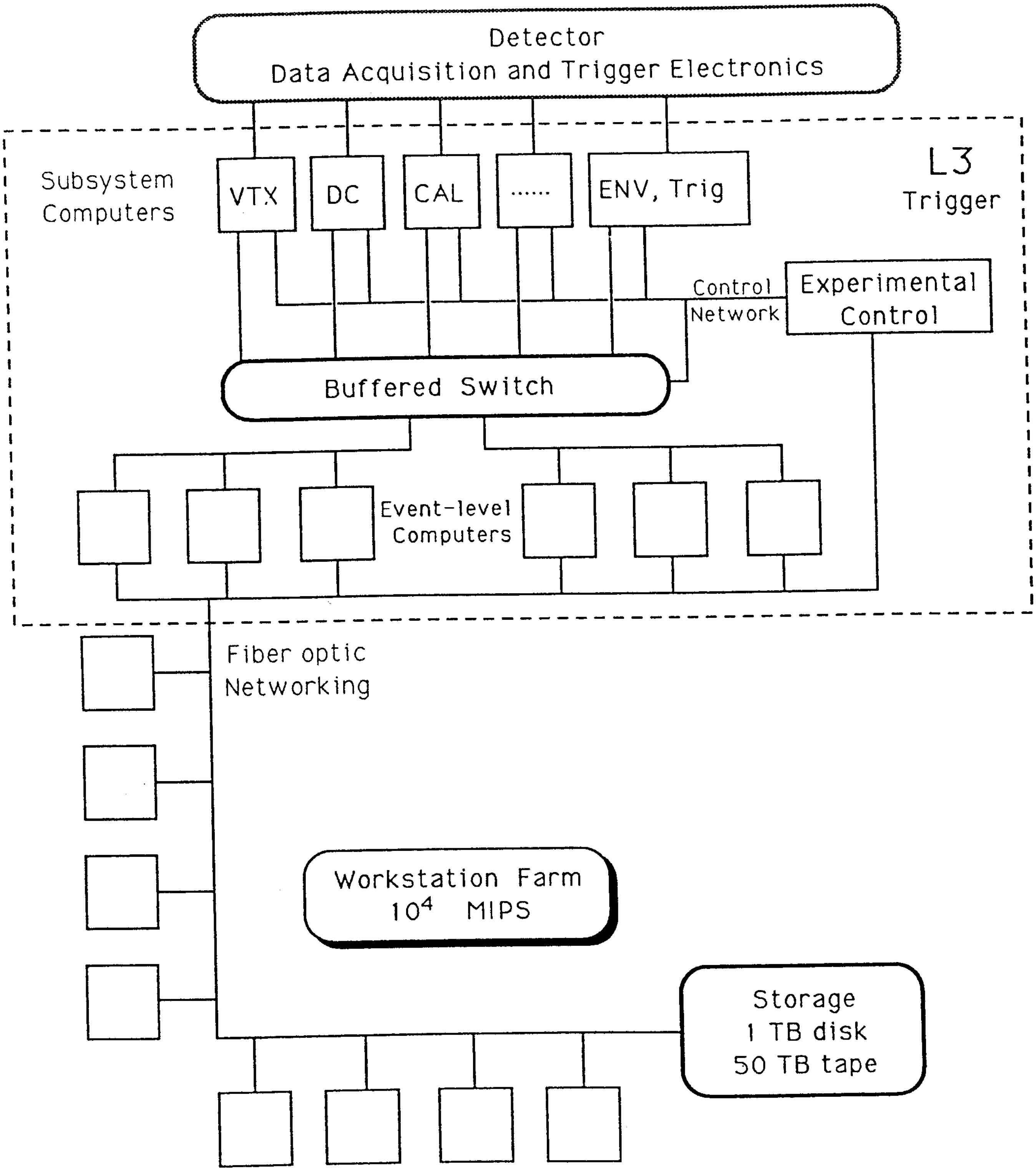


Figure 1. A possible model for the computing facilities for a B Factory.

The subsystem computers pass the data on to a "buffered switch", which could either be a specialized data switch, or possibly a computer with sufficient I/O bandwidth (Note that this is where the I/O bandwidth requirement of a single device is heaviest, with a worst case average rate of as much as 25 Mbytes/second. One should design for a peak rate that is even higher, although the peaks can be suppressed somewhat with adequate buffering). The task of this switch is to route all of the subsystem data from a particular event (time stamp) to one of several output addresses. The output address selects which of several online computers gets that event. We have not investigated specific solutions for this switch, so this is an area requiring further work (and this area of Figure 1 should be regarded as tentative – we have also come up with other possibilities which are somewhat different in detail).

The event level computers are responsible for actually building the subsystem data into whole event records. The number of computers allocated to this online activity will be flexible, determined by actual running conditions. Additional data corrections/reduction, especially anything depending on inter-subsystem information may take place here. The presence of a high trigger level, implemented in software in these computers, is envisioned in the likely case the background rates cannot be adequately suppressed in the electronic triggering algorithms. The design goal for these computers is that they will be able to handle an aggregate input of 1000 events/second (Level 2 trigger rate), and an output of 100 events/second (Level 3). In addition, these computers can be used to sample the event stream for online event displays and histograms. Our goal is to sample of order 10% of the output events, *i.e.* , an aggregate rate of 10 events/second.

The online event level computers perform the Level 3 trigger, provide online event sampling, and transfer data to permanent storage.

From the event-level computers, the data proceeds to permanent storage via fiber-optic networking (*e.g.* , FDDI). Many workstations have access to this storage, and the "PASS 1" analysis is accomplished by reading events from storage and distributing them to available workstations running the processing code. When finished with an event, the workstation transfers it back to permanent storage. The management of this distributed computing environment will require development.

Centralized permanent storage is accessed by distributed workstations over fiber optic networks.

We do not think it is practical to keep events within a "run" in any particular time-ordering, even at the initial storage stage. The difficulty is that the events are processed in parallel on different computers in the online data stream, and different events will take different amounts of time to reach the storage medium. However, it does appear to be useful to keep data from one "run" distinct from other runs. The term "run" is somewhat abstract – it at least means that periods with different experimental conditions are different runs, and, in practice, is likely to be connected to the fills of the accelerator, in which case a run could require $\gtrsim 10$ Gbytes of storage. As long as there is ample time to flush the buffers between runs, the separation of data by run should not pose a problem. If it eventually is deemed important to maintain some form of ordering even within a run, there are various possibilities beyond the nominally straightforward, but resource-intensive, one of maintaining very large buffers (on disk) and sorting to tape. For example, if we are more concerned about sequencing for searching and bookkeeping than about true time-ordering, we could simply add a sequence number to each event as it is written to tape. If, on the other hand (or in addition), it is important to have ready access to the true time-ordering, then a separate index file with this information could be maintained.

It is useful to take this model a step further and try to come up with some sort of "existence proof" with specific technology, and cost estimates as an indication of practicality. The following Table, together with the notes following it, gives some estimates in terms of current (June 1990) technology, and a projection five years into the future:

Component	Current (1990) technology	1995 projection
CPU (10^4 MIPS)	IBM RISC/6000-320 $360 \times \$13k = \$5M$	250 MIPS/CPU at $20k $40 \times \$20k = \$0.8M$
Networking ($\gtrsim 25$ MBps)	Fiber optics (*e.g.* , FDDI) $360 \times \$12k = \$4.3M$	Fiber optics $40 \times \$12k = \$0.5M$
Tape storage (50 Tbytes)	8 mm jukeboxes $200 \times \$50k = \$10M$ (3480 cartridge silos $50TB \times 0.3M\$/TB = \$15M$)	19 mm jukeboxes $4M
Disk storage (1 Tbyte)	$1M	$0.3M
Total	$20.3M	$5.6M

Table 2. *Estimates of computing costs for a B Factory*

1. Prices are estimated list prices; to the extent discounts are available they may thus be overestimates.
2. We choose the IBM RISC/6000-320 as our 1990 technology CPU, because it is one of the most economical current high-performance workstations.
3. The 1995 CPU projection is based on informal discussions with IBM representatives.
4. The current networking cost/node is from Ref. 4. We use the same cost/node in 1995 even though this cost is expected to decrease significantly with time (FDDI costs are expected to become comparable to current Ethernet costs). We do this as an expedient way to take into account the relatively greater importance of the basic network cost as the number of nodes decreases.
5. The 3480 cartridge silo prices are from Ref. 4. The 8 mm jukebox price is also in Ref. 4, and has been found to be consistent with estimates from other, informal, sources. The 19 mm jukebox price is estimated based on informal discussions with a representative of SONY. Further specifications for the 19 mm format are included in Appendix A.

6. The disk storage pricing is based on the discussion in Ref. 4, and is an exception to the stated use of list prices.

9. Software Considerations

Moving into a distributed UNIX environment poses the burden of learning a new system, as well as the opportunity for modernizing our software technology.

A broad range of software issues will face a new B Factory experiment. The architecture of the system is sufficiently different from traditional models (*e.g.* , the current SLAC environment) that many new concerns must be addressed. Many of the most basic system level software components will have to be carefully designed to fully exploit the capability of this system. Moving into a highly distributed UNIX-based system will also completely change the "look and feel" of the computing environment. This may add the burden of learning a new system to the users (physicists) while simultaneously providing an opportunity to improve efficiency and productivity by employing new software technologies such as graphical user interfaces, graphical application builders, and visualization techniques.

At the system level the single most important issue may be a scheme to distribute production analyses across many workstations. Since workstations may be distributed in offices as well as in isolated clusters (without monitors or keyboards), such software must be able to dynamically allocate CPU resources across the network as needed to support production and analysis jobs.

This new environment poses significant development challenges in terms of networked computing system management.

Another crucial element involves network file serving support for the multi-terabyte storage facility. Other system level matters include: distributed database access and management (for access to experimental data); network-based distributed system management (resource allocation, software updates, user accounts, security, hardware fault detection with diagnostics, backups and network management); on-line help and other utilities; and support for multi-vendor platforms.

At the user level new tools will need to be developed to support real-time functions and interfaces to detector components. Real-time capability is a relatively new aspect of UNIX and must be investigated and carefully evaluated. Access to modern workstations is generally via a local console or remote X-terminal equipment using various graphical user interface products. Users will require tools for editing, processing and displaying text and graphical data in a uniform, friendly environment. The choice of a primary programming language may depend upon availability and functionality; FORTRAN may not be the best choice. We will need many support procedures and programs for defining coding styles, code management, graphics (including visualization), documentation libraries and the design of off-line and interactive analysis environments.

Finally, we must, for the sake of productivity, manageability and maintainability, attempt to avoid mistakes made in past software projects and strive to consider and implement relevant modern ideas in this area (*e.g.* , see Ref. 5). In particular, software issues must be identified as early as possible to provide ample time for prototyping upon which to base even the earliest development.

A significant amount of R&D must be done to prepare for a B Factory experiment. The software challenges outweigh those of the computing hardware. Software investigations must begin soon to clarify the key issues, identify areas requiring significant development and begin to prototype that development, seek out existing solutions to common problems, monitor the progress of advances in industry and form a core of individuals with expertise.

10. Summary

The computing demands for a B Factory are considerably larger than we are used to in e^+e^- collider experiments. We have noted that the need for substantial CPU power implies a distributed computing environment. This, in turn, coupled with the large datasets, implies that a sustained and significant software development effort will be required.

The computing hardware demands are summarized in the Table below. These demands are somewhat different than those in hadron colliders, because of the larger fraction of useful events. However, the overall requirements are rather similar to those anticipated in the next few years for a hadron collider experiment, and considerably less than anticipated for a major SSC experiment. By considering the requirements and the possible technologies, we conclude that the support of a full scale B Factory experiment with currently available products is possible with considerable difficulty, but that in the appropriate five year time frame it appears quite tractable.

Table 3. *Summary of B Factory computing requirements*

Level 2 trigger rate	1 kHz
Level 3 trigger rate	100 Hz
Event size	25 kbytes
I/O bandwidth	
To level 3	25 Mbytes/second
Level 3 to tape	2.5 Mbytes/second
Offline analysis	> 10 Mbytes/second
Initial storage requirement	50 Tbytes (per year)
Large mini-DST size (10^8 events)	200 Gbytes
CPU power	10^4 MIPS

APPENDIX A

Advanced Tape Technology

An emerging technology in magnetic tape storage may answer many of the demands posed by the *B* Factory. This technology is a direct evolution of that used in the broadcast television world and is based upon helical scan principles already in wide use for both analog and digital applications. Only SONY Corporation currently markets an ultra-high capacity tape system. A specification sheet from SONY is reproduced below as Figure 2.

These units are currently available. The price is $200k–250k per drive. At the present time, fewer than a dozen evaluation units have been shipped. This product as delivered is not complete nor completely satisFactory.

The guaranteed error rate $(1 : 10^{10})$ should improve by at least three orders of magnitude to compare favorably with existing 3480 and 8 mm technology. SONY is working on a solution involving wrapping an additional error correction layer to the data.

A highly-buffered interface is required for the *B* Factory distributed computing environment. Both SONY and an unidentified third party vendor are working on solutions. The time scale for a product is unknown.

The *B* Factory demand for $\gtrsim$ 50 Terabytes of storage can only be met by the use of a robotic system capable of handling at least 500 of the large tape cartridges. SONY is investigating such a robot and expects a product (based upon an existing broadcast video device) within two years. The need for massive storage is not confined to the *B* Factory project. In addition to broadcast video, this volume of data storage is needed by other agencies of the government such as the IRS, military, census, Library of Congress, and NASA. Although SONY is the only vendor marketing a product with the capacity needed by the *B* Factory, it is expected that other vendors will appear in this growing market within the next one to two years.

SPECIFICATIONS

PERFORMANCE

Recorded tape format:	ANSI X3B.6 ID-1 Standard
Cassette tape:	19mm type D-1 Broadcast standard (Hi-Hc) Large/Medium/Small sizes
Recording capacity:	Max. 770 Gbits (L-cassette, 16μm) Max. 330 Gbits (M-cassette, 16μm) Max. 100 Gbits (S-cassette, 16μm)
User date rate: (Record/playback)	256, 128, 64, 32, 16, 10.7 Mbps

Recording time/Tape speed:

Data rate (Mbps)	Recording Time (H: hours, M: minutes) L-size	M-size	S-size	Tape Speed (mm/sec)
256	50M	20M	7M	423.8
128	1H 40M	45M	15M	211.9
64	3H 20M	1H 30M	30M	105.9
32	6H 40M	3H 00M	1H 00M	53.0
16	13H 30M	6H 00M	2H 00M	26.5
10.7	20H 20M	9H 00M	3H 00M	17.7

Bit error rate (Corrected):	1 × 10E-10
Data assurance:	Read-after-Write for Data and CTL (Data and Track Set ID)
Tape loading time:	Less than 14 seconds
Servo lock time:	Approx. 10 sec. from stop mode Approx. 4 sec. from standby mode
Fast forward/rewind time:	Less than 180 sec. with L-cassete Less than 90 sec. with M-cassette Less than 45 sec. with S-cassette

GENERAL

Power requirements:	100V ~ 120V/220V ~ 240V ±10%, 50/60Hz
Power consumption:	Max. 550W
Operating temperature:	10°C to 35°C (50°F to 95°F)
Storage temperature:	−20°C to +60°C (−4°F to 140°F)
Operating humidity:	20% to 80% (non-condensing)
Storage humidity:	10% to 90% (non-condensing)
Weight:	Approx. 67 kg (147 lb 11 oz)
Dimensions:	436(W) × 432.5(H) × 635.5(D)mm (17 1/4 × 17 1/8 × 25 1/8") Including handles and feet

INPUT/OUTPUT SIGNALS (Connector)

DATA INPUT (D25S):	8 line pairs for data (ECL, NRZ) (with clock, sync, parity)
DATA OUTPUT (D25S):	8 line pairs for data (ECL, NRZ) (with clock, sync, parity and error flag)
REF (reference) INPUT (D25S):	Clock and sync (ECL)
ANNOTATION INPUT CH-1/CH-2 (XLR 3-pin, female):	+4dBm, 600ohms, balanced
ANNOTATION OUTPUT CH-1/CH-2 (XLR 3-pin, male):	Low impedance, balanced
MIC IN (Standard jack):	For ANNOTATION CH-1
HEADPHONE OUT (Standard jack):	For ANNOTATION CH-1
AUX (auxiliary) DATA INPUT/OUTPUT (D15S):	RS-422 interface
REMOTE 1/2 (D25S):	RS-232C interface
REMOTE 3:	IEEE-488 (GPIB) interface
REMOTE 4/5 (D9S):	RS-422/485 (Primary) interface

SUPPLIED ACCESSORIES

Rack angle assemblies (2)
AC power cord (1)
Plug holder (1)
Operation and installation manual (1)

OPTIONAL ACCESSORIES

Cables
- For data input/output signals
  - Digital video cable: VCD-2D/5D/10D/30D (2/5/10/30m)
- For annotation input/output signals
  - Digital audio cable: ECD-3C/10C/30C (3/10/30m)
  - Analog audio cable: EC-5XLR2/10XLR2 (5/10m)
- For remote interface
  - RS-422/485 Remote control cable: RCC-5G/10G/30G (5/10/30m)
  - IEEE-488 (GPIB) interface cable: SMK-0032 (2m)

Rack mount kit
- Digital VTR rack mount rail: RMM-18DV (for use with Rack angle assemblies)

Tapes
- 19mm type D-1 cassette (Hi-Hc): L/M size

Design and specifications subject to change without notice.

Figure 2. SONY specifications for a high capacity tape storage system, model DIR-1000. Copyright 1990 by SONY Corporation.

References

1. Eugene Schmidt, who has done much to organize CDF computing, provided this information.
2. H. Areti, "Future Microprocessor Farms: Offline and Online", FNAL-TM-1642, January 1990. The event size and rates given are compatible with those in the Lankford reference following.
3. A. J. Lankford, "Computing and Data Handling Requirements for SSC and LHC Experiments", SLAC-PUB-5243, May 1990, Talk delivered at the 8th Conference on Computing in High Energy Physics, Santa Fe, NM, April 9-13, 1990.
4. L. Cottrell, "Thoughts on a Balanced Model for SSC Off-line Computing", prepared for the *Report of the SSC Computer Planning Committee*, SSCL-N-691, January 1990.
5. P. Kunz, *Computer Physics Communications*, **57** (1989) 191-197.

Report of the Trigger and Data Acquisition Group

T. Bolton, D.D. Briggs, C. Hearty, W.R. Innes,
C.K. Jung, A.J. Lankford, P.C. Rowson, P. Rubin,
M.H. Shaevitz and W.J. Wisniewski

1. Introduction

The Asymmetric *B* Factory will present a new environment for triggering in e^+e^- storage rings. The significant differences arise from the greatly increased bunch crossing frequency. Some of the differences from previous e^+e^- machines are outlined below:

- The first level of trigger decision must be pipelined.

  Bunch crossings will occur approximately every 4 ns., compared to every 2.3 μs as in *PEP* operation with three bunches. Even the fastest stage of the trigger decision will require much more time than exists between crossings. Consequently, the trigger must process every successive crossing even as prior crossings undergo subsequent processing steps.

  Any B Factory detector trigger must employ a pipline architecture

- The trigger must identify the time (*i.e.*: bunch crossing) of each event candidate.

  Since the time between bunch crossings is less than the response time of even the fastest detector element, the trigger processors must be capable of identifying, at least approximately, the bunch crossing which is being triggered upon. This is possible since the crossing time is not less than detector resolving times.

- The ratio of triggers on e^+e^- annihilation events to triggers on background processes must be greatly improved, with no loss in trigger efficiency.

  The rate of real physics events is large, approximately 50 Hz at the $\Upsilon(4S)$ and greater at other Υ resonances. Consequently, the trigger must select events with sufficient purity that the overall trigger rate is not many times larger than the physics rate. At the same time, the trigger must select b events with high efficiency since the rare decay modes are produced only in small numbers.

- Sources of potential background will be greatly increased.

  The currents in the storage rings of the Asymmetric B Factory are significantly higher than in previous e^+e^- colliders, approximately 3 A as compared to approximately 30 mA in *PEP*. Consequently, backgrounds from synchrotron radiation and from degraded beam particles are potentially much larger than in previous machines.

  The very frequent bunch crossings result in the detector being live for cosmic ray backgrounds at all times, rather than only every few microseconds.

Despite the above differences of the Asymmetric B Factory from previous e^+e^- machines, an important similarity is that the average time between annihilation events at the Asymmetric B Factory is still small as compared to the signal-response time of the detector and to the decision times of trigger processors. In addition, the rate of annihilation events is sufficiently low that no physics processes must be rejected by the trigger. The modest physics rate distinguishes the Asymmetric B Factory from high-luminosity hadron colliders such as the SSC, leaving event selection relatively straightforward although more challenging than in the past.

The trigger must be designed so that its performance does not limit the effective luminosity

A defining criterion for the trigger system of an experiment at the Asymmetric B Factory should be that trigger performance not limit the luminosity at which the machine can operate. This requirement is particularly relevant in this machine since demands for luminosity will drive accelerator operation to the limits of acceptable backgrounds. The trigger architecture and processing must be designed such that trigger performance degrades gracefully in response to background levels much greater than expected.

2. Overview of the Trigger

The purpose of the trigger is to select bunch crossings in which an interesting physics process occurred from the large rate of all bunch crossings. At the Asymmetric B Factory, the rate of interesting physics is in the range of tens of events per second, while the rate of bunch crossings is about 250 MHz. The challenge to the trigger is to distinguish the interesting physics events from potentially severe machine-related backgrounds. Unlike triggers in hadron colliders, the trigger is not required to distinguish interesting physics processes from uninteresting processes, since essentially all processes are of physics interest in the B Factory.

As in existing experiments, the trigger of an experiment in the Asymmetric B Factory can be expected to consist of event selection decisions in a sequence of levels. Each trigger level processes a subset of detector signals in order to reduce an input rate of event candidates to a lower output rate, thus affording the subsequent trigger level additional time for more sophisticated processing of a larger subset of detector signals.

The trigger design employs three levels

Here we refer to *Levels 1, 2,* and 3, although the choice of three levels is arbitrary and could be adjusted if appropriate. Figure 1 sketches ranges of possible trigger rates to and from each trigger level. These rates are given by the requirement that each level of the trigger on average make a decision in a time equal to one over the rate at which candidate events are input (*e.g.*: an input rate of 10^5 Hz affords a 10 μs decision time on average). This rule applies in the case of an adequately buffered data acquisition system. The complexity of the data acquisition system drives the trigger towards a design with a low output rate; whereas the complexity of the trigger system drives the trigger towards a design with a high output rate.

The values shown in Figure 1 are loosely based upon experience with trigger processing techniques in prior experiments and reflect the flexibility which is available in designing the trigger. The approximately one microsecond decision time for *Level 1* corresponds to the minimum possible time to collect trigger signals, perform rudimentary processing, and distribute strobes. More complex processing can be done by hardware in ten or more microseconds by *Level 2*. Ten microseconds processing time limits the input rate to 100 kHz from *Level 1*. Lower rates from *Level 1* would afford *Level 2* additional processing times. The lowest conceivable rate from *Level 1* (as discussed below) is 1 kHz, which would afford

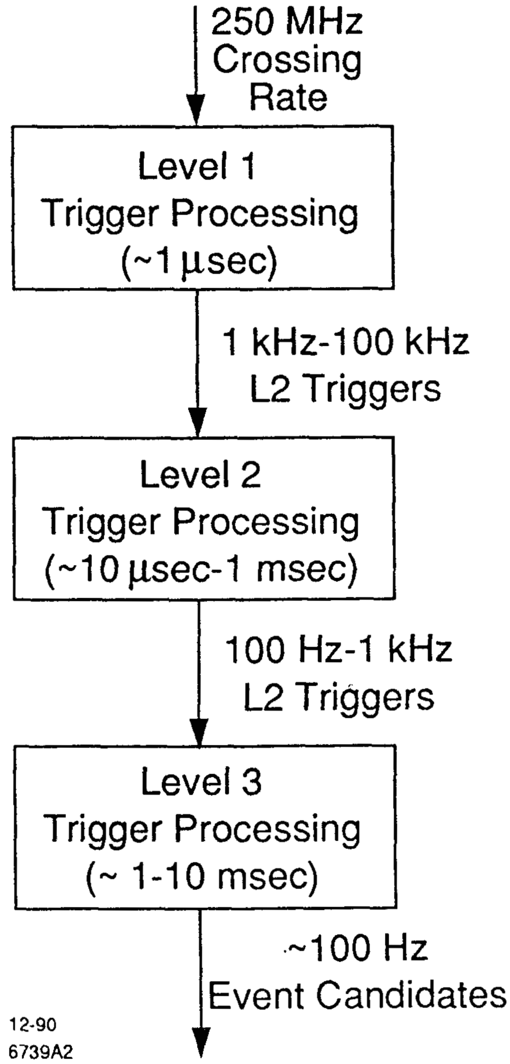


Figure 1. *Trigger rates and processing times at each trigger level.*

Level 2 as much as one millisecond processing time. Such a low rate might eliminate the need for the data acquisition system to buffer multiple events during the *Level 2* decision, or it could possibly eliminate the need for a hardware *Level 2* trigger preceding the *Level 3* trigger. Order 100 Hz, which is near the physics rate, is the lowest possible rate into *Level 3*.

Since the level 3 trigger is usually conceived as a collection of general-purpose microprocessors, its input rate is bounded at approximately 1 kHz, which would allow each processor only one millisecond times the number of processors to select events.

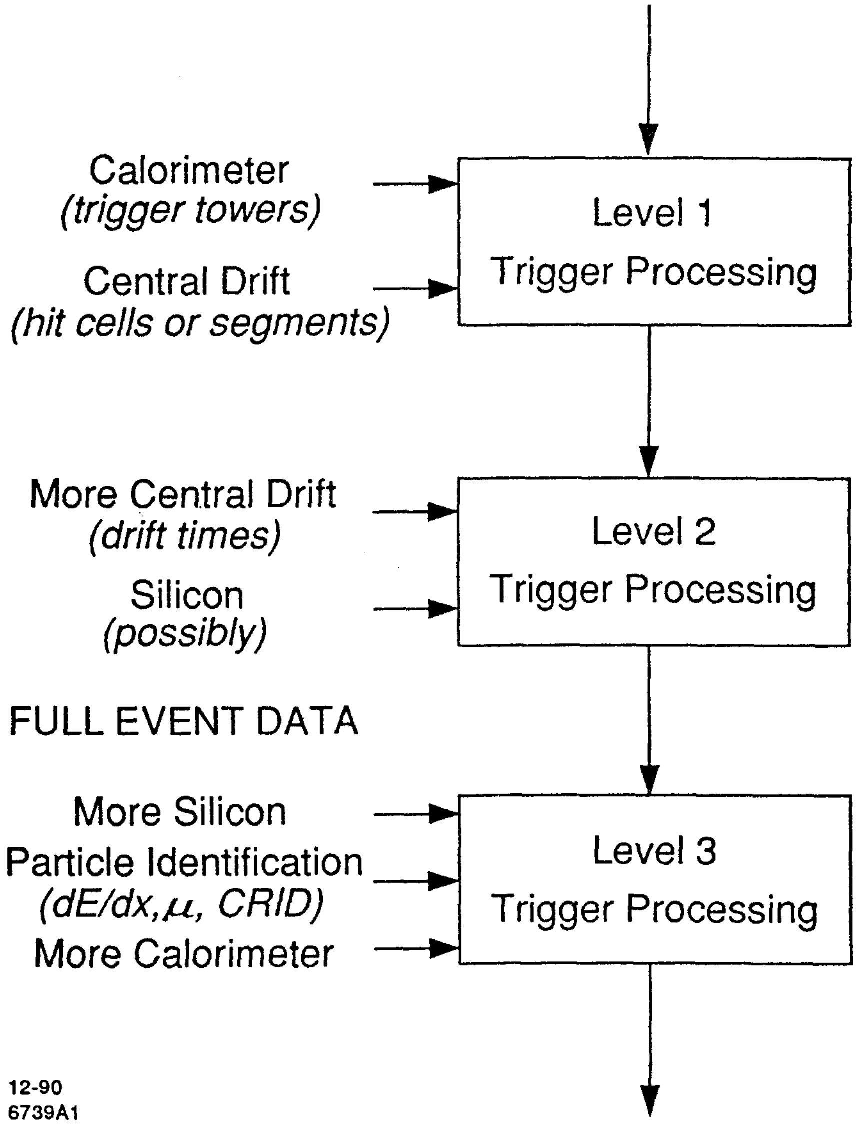


Figure 2. *Detector data input at each trigger level.*

At Level 1, calorimetric triggers will employ total energy measurement or seek patterns of energy deposition, while tracking triggers employ patterns of hit drift chamber cells or track segments

Figure 2 illustrates detector information which is available to and which may be utilized by each trigger level. At *Level 1*, calorimeter triggers will operate on energy deposit in trigger towers and on total energy. Tracking triggers at *Level 1* will operate on either hit cells or track segments in the central drift chamber. At *Level 2*, more detailed drift chamber data, such as drift times, will be used. Vertex information based on OR's of adjacent strips or pixels may also be used. Data from the finely segmented silicon system may be difficult to access on the timescales of the *Level 2* trigger, and are impossible to access in a *Level 1* trigger. At *Level 3*, data from the full digitized event is available. This data includes vertex information from the silicon system, particle identification information such as dE/dx, muon system hits, and ring-imaging or threshold Čerenkov data, and complete calorimetry data. Time-of-flight data, if it exists, would be available as early as *Level 1*. The nature of data used by each level and how the data is used will be further described in sections which follow.

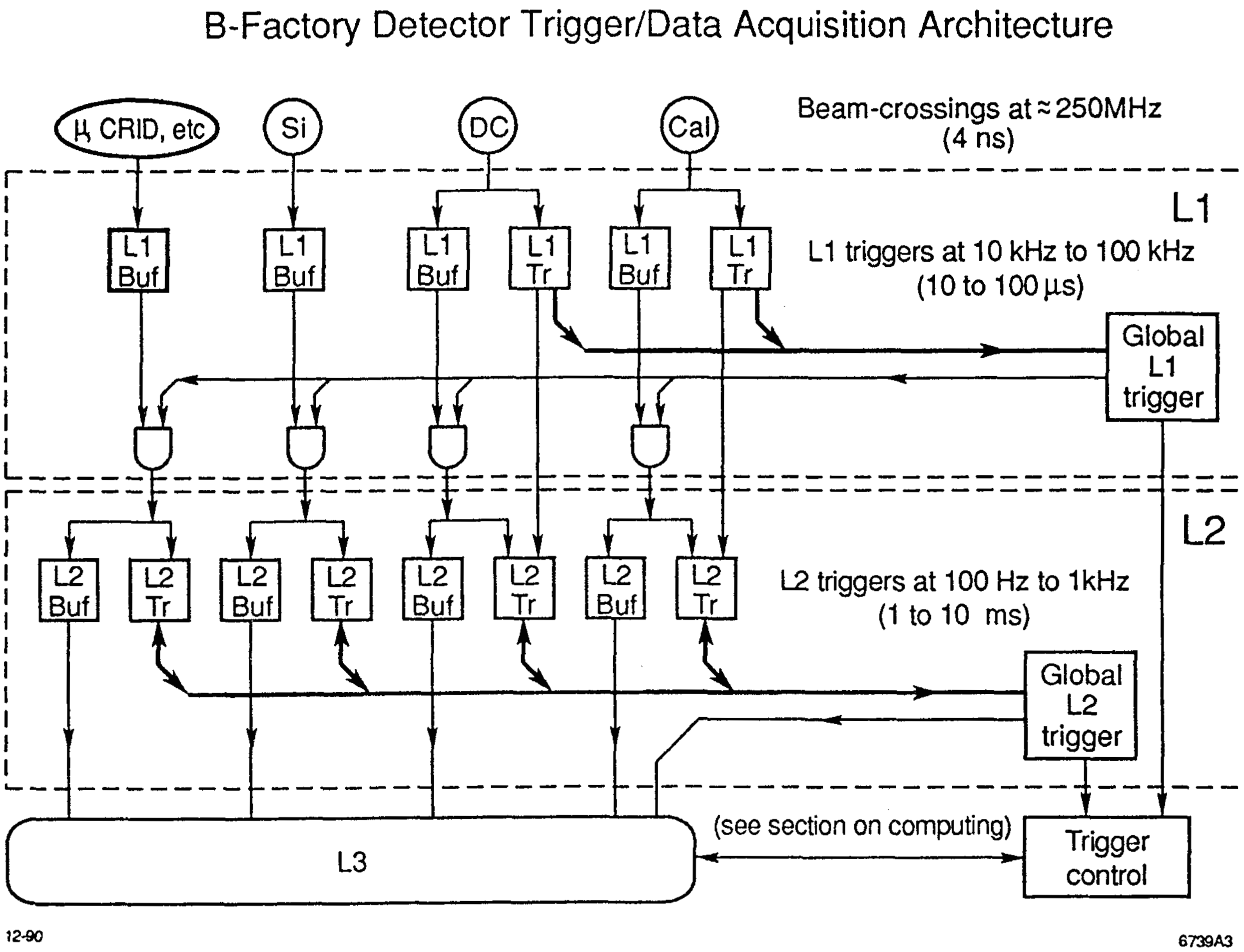


Figure 3. *Data flow through the trigger and data acquisition architecture.*

Data which is used by the *Level 1* trigger will be transported to the trigger processors on paths which are independent of the standard readout (data acquisition) paths. Paths to *Level 2* processors will most likely also be independent. The *Level 3* processors, which are envisaged to be general-purpose microprocessors, will obtain their data through the data acquisition system. This data flow architecture is sketched in Figure 3.

3. Trigger Goals and Efficiencies

The primary goal of the trigger design is to maximize the efficiency for all final states of interest. The annihilation cross-sections and rates for each of the upsilon resonances is shown in Table 1 for luminosities of 3×10^{33} and 10^{34}.

Resonance	Mass (GeV)	R_{had}	R_{tot}	σ_{pt}	σ_{tot}	rate at 10^{34}	at 3×10^{33}
$\Upsilon(1S)$	9.46	22	25	0.97 nb	24 nb	240 Hz	72 Hz
$\Upsilon(2S)$	10.02	10	13	0.87 nb	11 nb	10 Hz	33 Hz
$\Upsilon(3S)$	10.36	7	10	0.81 nb	8 nb	80 Hz	24 Hz
$\Upsilon(4S)$	10.58	4	7	0.78 nb	5.5 nb	55 Hz	16 Hz

Table 1. *Cross-sections and production rates of the* Υ *resonances.*

The dominant physics signatures at the $\Upsilon(4S)$ are listed below :

- $e^+e^- \to B\overline{B}$
- $e^+e^- \to e^+e^-$ (Bhabha)
- $e^+e^- \to \mu\mu$
- $e^+e^- \to \tau\tau$
- $e^+e^- \to e^+e^-X$ (two-photon)

We have considered both tracking-based and calorimetric triggers; the most challenging final states for the trigger are those with both low charged multiplicity and low visible energy, such as the $\tau\tau$ and two-photon processes.

The elemental triggers described here fall into one of two categories : tracking based or calorimetric. The trigger configuration actually used would generally be a logical OR of some subset of these triggers.

The tracking-based triggers are:

2 Track

At least 2 tracks with $|\cos\theta| < 0.8$. These are "central tracks" which penetrate to the outermost drift chamber layer.

1.5 Track

At least 1 track with $|\cos\theta| < 0.8$, and at least 1 track with $|\cos\theta| < 0.9$. Here, the "dipped" track may pass through only the inner drift chamber layers.

The **1.5 Track** trigger is designed to improve acceptance for two-track events, which may be boosted into unfavorable topologies. Single track triggers are expected to result in unacceptably high rates due to backgrounds.

The calorimetric triggers are:

Total Energy

The visible energy in the calorimeter (barrel and endcap) is required to be above a threshold typically of order 1-2 GeV. A low channel-by-channel threshold would be applied prior to summing.

2 Supertower

Contiguous groups of towers are defined to constitute a supertower, and the energy in each supertower must be at least 50% of minimum ionizing, or about 100 MeV.

Topological

It may be necessary to implement a topological requirement in the calorimetric trigger to reduce background due to cosmic rays and beam gas, and other beam related noise. Our Semi-Hemi-Cylinder trigger requires an adjustable degree of azimuthal symmetry in the energy deposition pattern in the barrel calorimeter; this trigger is described in greater detail in Section 5.4. This trigger would be AND'ed with all other triggers, and therefore must be as efficient as possible for desired events.

We have studied the efficiency of the various triggers for physics of interest with the ASLUND Monte Carlo. Our study included both hadronic final states due to $e^+e^- \rightarrow B\overline{B}$ for which efficiencies are expected to be high and the presumably challenging tau pairs. We assume that 3 GeV positrons are colliding with 9 GeV electrons. All Monte Carlo particles with $|\cos\theta| < 0.95$ are included, and the calorimeter response is simply modeled by assuming that for photons and electrons all energy is visible, for muons exactly 200 MeV (1 mip) is visible, and for hadrons the visible energy is evenly distributed between 1 mip and the total energy, or zero and the total energy depending on whether the hadron is charged or neutral, respectively. Supertowers with angular size $2\pi/24(\phi) \times \pi/12(\theta)$ are defined. These are registered as "hit" if the visible energy exceeds 100 MeV. The Semi-Hemi-Cylinders are segmented to match the supertower segmentation; see Section 5 for details.

A plot of the efficiency of the total energy trigger for τ pairs and for the *CP* eigenstate $B^0, \psi K^0_S$ is shown in Figure 4 as a function of the total energy threshold.

A summary of our results for several triggers is given in Table 2. The reader is reminded that the calorimetric triggers are meant to be OR'ed, except for the final topological trigger which would be required for any trigger.

Trigger	$\tau^+\tau^-$	$B^0,\ J/\psi K_S$
Two central tracks **2T**	0.530	0.997
One and "one half" track **1.5T**	0.730	1.000
Two supertowers **2ST**	0.950	1.000
$E_{vis} > 1$ GeV **E**	0.883	1.000
240° semi-hemi-cylinder **SHC**	0.843	1.000

Table 2. *Selected trigger efficiencies.*

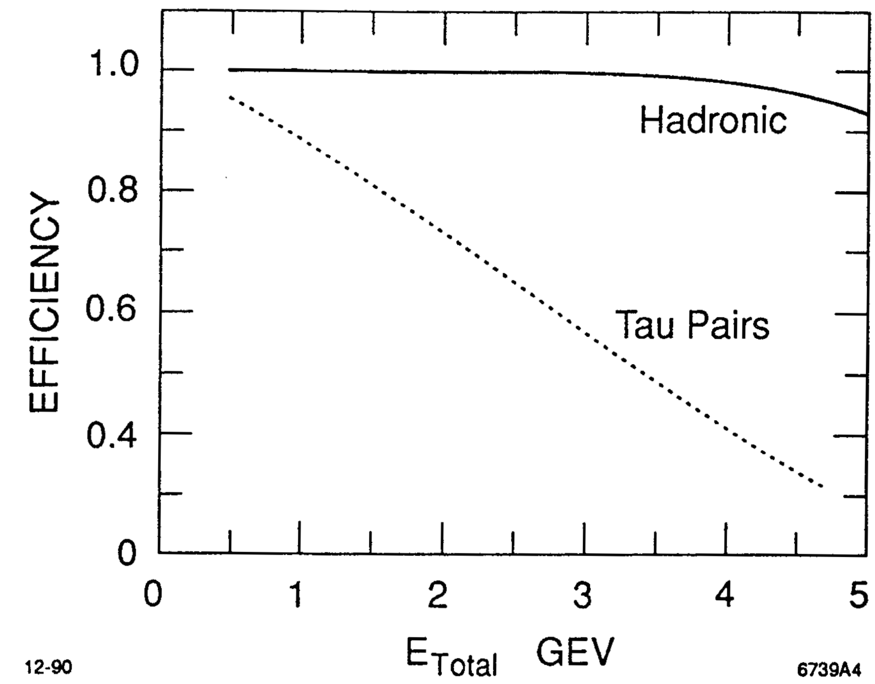


Figure 4. *Efficiency of the total-energy trigger.*

4. Expected Backgrounds

There are four main sources of background

We have considered four sources of background to the trigger: cosmic rays, beam gas scatters within the detector volume, synchrotron radiation, and lost beam particles.

The flux of cosmic rays striking any part of the calorimeter or inner detector is approximately 3 kHz. The experience of the Crystal Ball detector indicates that 90% of the cosmic ray events will be single muons; the remaining 10% will be showering particles, which deposit more energy than the equivalent of two minimum-ionizing particles and illuminate several calorimeter towers. Requiring the presence of charged tracks passing near the interaction point in $r - \phi$ would reduce the cosmic ray rate by nearly two orders of magnitude. Loose geometric cuts, such as requiring energy deposits nearly back-to-back in ϕ, would also reduce the trigger rate.

Based on the experience of the MARK II detector at *PEP*, and assuming the vacuum at the interaction point of the Asymmetric B Factory to be the same as at *PEP*, the rate of two-track triggers from beam-gas interactions is estimated to be 50-100 Hz. Since high vacuum will be difficult to achieve with the high-current beams of the Asymmetric B Factory, beam-gas backgrounds are potentially serious. A vacuum 100 times worse than that at *PEP* could result in a beam-gas trigger rate of 10 kHz, which would be at the limit of the acceptable *Level 1* rate. It appears possible, however, to achieve a vacuum at least as good as that at *PEP*, so this is unlikely to be a major problem.

Ongoing calculations of synchrotron radiation and lost-particle backgrounds indicate that lost-particle backgrounds are dominant, even for low energy photons, and so will be emphasized here. Triggers on high-p_t charged tracks or on high-energy electromagnetic showers in either the barrel or endcap calorimeters have been considered. In each case, requirements of at least one or at least two tracks or showers have been studied. No geometric cuts are applied in this study. The results are summarized in Table 3 for luminosity of 3×10^{33} cm^{-2}s^{-1}. There are two contributions to the trigger rates — random overlap from two different background rays, and two significant tracks or energy deposits from a single ray. The overlap calculation assumes a 1 μs trigger window for both the DC and the calorimeter. The table does not include "fake" charged tracks due to overlap of background hits in the DC. The low-density gas planned for the DC will help to reduce the rate of fake tracks.

Charged particles with $p_t < 100$ MeV (assuming 94 cm calorimeter radius and .7 T magnetic field) will not reach the calorimeter, so the trigger level is likely to be somewhat at or above that value. A minimum-ionizing track will deposit about 200 MeV in a CsI(Tl) calorimeter, or about 350 MeV in a liquid krypton calorimeter.

The background rates calculated for lost beam particles are based on the high energy APIARY 6.0 ring operating at luminosity of 3×10^{33}, and are extremely dependent upon the lattice. The limitations of this calculation are discussed in the chapter on the interaction region. The results are based on limited statistics, the equivalent of 500 μs running time for the calorimeter trigger and 2.9 ms for the charged trigger.

Trigger	Cut (MeV)	Random Overlap Rate (kHz)	Rate from one ray (kHz)	Total Trigger Rate (kHz)
≥ 1 Supertowers	50	–	290	290
	100	–	52	52
	200	–	10	10
≥ 2 Supertowers	50	85	31	116
	100	3	2	5
	200	0.1	< 5	< 5
≥ 1 Track	60	–	20	20
	100	–	4	4
≥ 2 Track	60	0.4	0.3	0.7
	100	0.02	< 0.8	< 0.8

Table 3. *Calculated charged particle and calorimeter trigger rates due to lost particles from the APIARY 6.0 high energy ring during operation at luminosity of 3×10^{33}. The limits are 90% C.L. The cut is on energy per supertower for calorimeter triggers, and track p_t for charged track triggers.*

The background rates from these four sources are manageable. They are, however, sufficiently close to the target of 10 kHz for *Level 1* to be of concern. In particular, the backgrounds from lost beam particles must be reconsidered as the machine design evolves.

5. *Level 1* TRIGGERS

The goal for the trigger rate from *Level 1* is order 10 kHz. Rates as high as 100 kHz may be manageable architecturally for the *Level 2* trigger and for *Level 2* buffers; however, higher rates imply greater system complexity. *Level 1* rates as low as 1 kHz could eliminate the need for multiple event buffering during the *Level 2* trigger, but such low rates will be difficult to achieve.

The goal for the Level 1 trigger is 10 kHz

In order to reduce rates to the 10 kHz Level, the *Level 1* trigger eliminates backgrounds due to low energy (below p_t of 50-100 MeV) tracks and photons from sychrotron radiation and from showers of lost beam particles. In addition, it eliminates most of the background from lost beam particles by requiring that there be two particles (charged or neutral) above the p_t cut. The output of the *Level 1* trigger will then be dominated by cosmic rays and by coincidences of secondaries from showers of one or two lost beam particles. The actual p_t cut implemented will depend upon the resolution of the trigger and upon background rates, but should be sufficiently low to maintain efficiency for two prong final states of annihilation events.

In order to assure and measure efficiency for physics processes, the *Level 1* trigger is comprised of parallel calorimetric and tracking components (hereafter referred to as "nodes") which implement fully independent selection criteria. *Level 1* passes an event candidate on to *Level 2* if either node is satisfied.

5.1. Calorimeter Trigger at Level 1

The electromagnetic calorimeter of the detector at the Asymmetric B Factory must be efficient for low energy photons from π^0 decays. Consequently, the calorimeter will have very good signal-to-noise for individual minimum-ionizing particles, approximately 400:1 signal to electronics noise for a single calorimeter cell of either CsI(Tl) or liquid krypton. We have examined the option of a purely calorimetric trigger at *Level 1* which is efficient for events with two or more minimum-ionizing particles. Such a trigger would accept all cosmic rays which pass through any portion of the calorimeter, yielding a background rate of approximately 3 kHz. Although this rate is within the 10 kHz target of *Level 1*, it could possibly be reduced by requirements on the topology of energy deposit in the calorimeter, and it can be substantially reduced by requiring tracking in some subsequent trigger level.

The Level 1 calorimeter trigger consists of three nodes

The *Level 1* calorimeter trigger consists of three parallel components, or nodes. The principal node provides a trigger on two or more particles, charged or neutral, anywhere in either the barrel or endcap calorimeters. This two-supertower trigger is accomplished by summing nearby calorimeter towers into "supertowers" and then counting the number of supertowers above threshold. The second *Level 1* calorimeter node provides a trigger on total energy deposit in the calorimeters by summing overall all calorimeter towers. The third *Level 1* node is based on particular topologies.

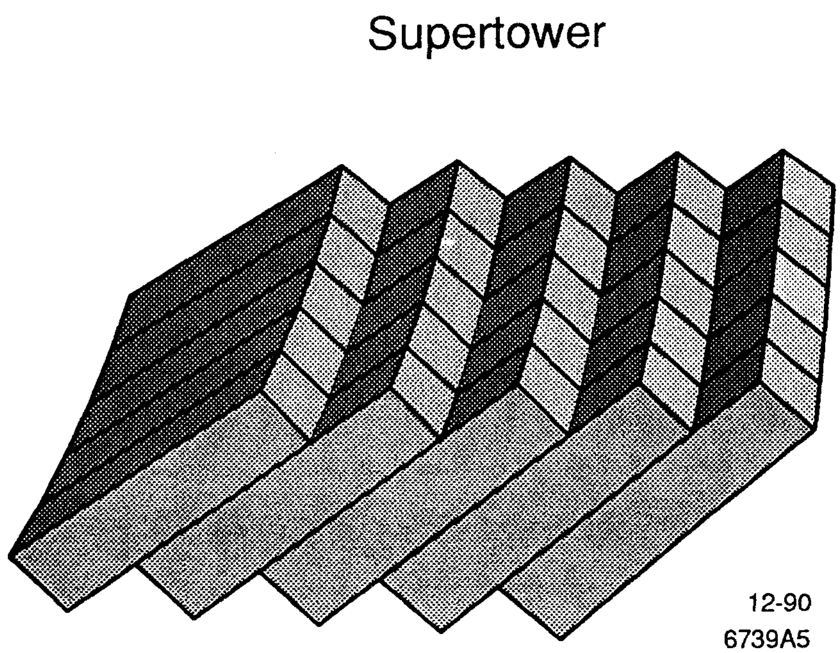
5.2. *Two-Supertower Trigger*

The goal of the two-supertower trigger is to trigger efficiently on any combination of two or more high-p_t charged or neutral particles detected anywhere in the barrel and endcap calorimeters.

A supertower is a 5×5 array of individual towers

The two-supertower trigger requires at least two non-adjacent supertowers

The architecture of this component of the trigger is illustrated in Figure 5. Groups of contiguous calorimeter towers are summed into supertowers of about 5 x 5 individual towers. This reduces the number of calorimeter trigger channels to about 400 from order 10^4 individual towers. It also reduces the probability that the energy of a particle will be split between two trigger channels. The signal-to-noise ratio for a minimum-ionizing particle in a supertower will still be very good, as high as 80:1, enabling an efficient threshold at about one-half the minimum-ionizing signal. To further improve efficiency at boundaries of supertowers, where energy may be split between two supertowers, overlapping sets of supertowers may be formed. We refer to a complete set of supertowers covering the calorimeters as a "layer". The two-supertower trigger requires at least two non-adjacent supertowers above threshold in a single layer as a means of providing a trigger on two particles without double-counting a single track.



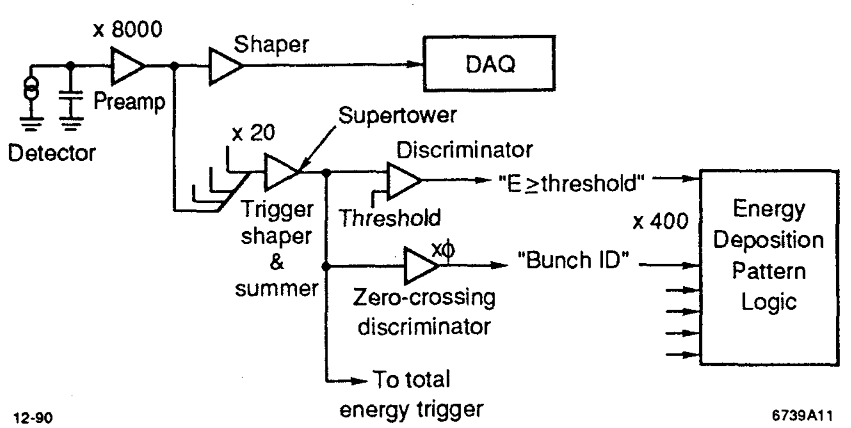


Figure 5. *The two-supertower trigger:*
a. *supertowers formed from sums of calorimeter cells,*
b. *block diagram of the two-supertower trigger.*

The good signal-to-noise ratio allows the threshold on each supertower to be set at about half the energy deposit of a minimum-ionizing particle (E_{mip}). It is unlikely that pileup of low-energy photons will accumulate enough energy in a single supertower (or pair of supertowers) to satisfy this threshold. The single tower rate due to electronic noise alone is:

$$1/\Lambda_{effective} \times erf(-\theta\sqrt(m))$$

where

$\theta =$ threshold in units of σ_E of a tower

$\Lambda_{effective} = 1\ \mu s$,

$m = \#$ of towers in a supertower.

For thresholds of 0.5 E_{mip}, the rate in an individual supertower, or all 1600 supertowers in parallel, is negligible. The rate of two supertowers in coincidence is much less.

For supertowers above threshold, the time of the event can be determined using techniques equivalent to measuring the zero-crossing time of the differentiated signal. The timing resolution of such a trigger at threshold is:

$$\sigma_t < \Lambda_{effective} \times \sqrt{(}m)/\theta\ ;$$

the resolution improves linearly with larger signal. The timing resolution of a single minimum-ionizing particle in a supertower is expected to be a few nanoseconds, and can be improved with signal shaping optimized for timing. The signal time is used to identify the bunch crossing to which the event candidate corresponds. Two supertowers providing a trigger candidate must identify the same crossing. Identification of the bunch crossing also contributes to knowledge of the drift chamber t_0 for use by the *Level 2* tracking node.

5.3. Total-Energy-Deposit Trigger

The total-energy component of the calorimeter *Level 1* node provides redundancy, as well as efficiency for any annihilation events which distribute energy uniformly in neutral particles only. The total-energy signal is formed by summing the energy signals of all supertowers; however, we expect to apply a very low threshold (a few σ) to each supertower before forming the sum over the entire calorimeter. Otherwise, this trigger component will be dominated by electronic backgrounds, including beam pickup and line-synchronous pickup, and correlated beam effects. The tower-by-tower threshold can also be used to make the total-energy-deposit trigger less sensitive to beam-induced backgrounds, by eliminating small energy deposits from low-energy particles and photons before

The total-energy-deposit trigger requires a total energy above threshold of 1-4 GeV forming the total energy sum. With a threshold of between 1 and 4 GeV, the total energy deposit trigger is expected to add little additional rate to the two supertower trigger.

5.4. *A Topological Calorimetric Trigger*

The Semi-Hemi-Cylinder topological calorimeter trigger (SHC) is defined as follows : The event must not contain any "empty" semi-hemi-cylinders in the barrel calorimeter. "Empty" is defined as less than 1/2 mip (about 100 MeV), and a semi-hemicylinder is slightly more than a hemicylinder *i.e.*, the semi-hemi-cylinder covers 180 + ϕ degrees in azimuth. Figure 6 defines a SHC. Note that the noise tolerance of the calorimetric trigger can be improved by ganging channels (to the level of $\sim$ 100) prior to discrimination. That is, a per-channel threshold is set prior to summing. The number of semi-hemi-cylinders defined is taken to match the segmentation used in the supertower definition: 24 azimuthal bins for a total of 24 semi-hemi-cylinders. The value of ϕ would be set by the efficiency for physics events.

The topological calorimetric trigger requires at least 1/2 mip in both semi-hemi-cylinders

The SHC trigger is designed to reduce the primary trigger rate due to cosmic rays, beam-gas scatters and off-energy beam particles, for all of these processes will be characterized by relatively asymmetric energy deposition. For example, off-center cosmic rays will fail the SHC trigger at a rate which depends on the size of the ϕ angle (for $\phi = 60°$, this reduction would be about a factor of 4; see Figure 7.). Analysis of off-energy particles also reveals that the energy deposition due to this process is peaked in azimuth towards the outer edge of the ring, so that even two overlapping events (during the 1 μs primary trigger integration time) would fail the SHC trigger about half the time.

In Figure 7, the efficiency of the SHC trigger is plotted for tau pairs events as a function of the "slop angle" ϕ. It is clear that tau pairs are often rather asymmetric ; for $\phi = 60°$, the efficiency is only 84%. Our conclusion is that a SHC trigger, if necessary, would certainly require sacrificing a sizeable fraction of tau events and probably two-photon events as well.

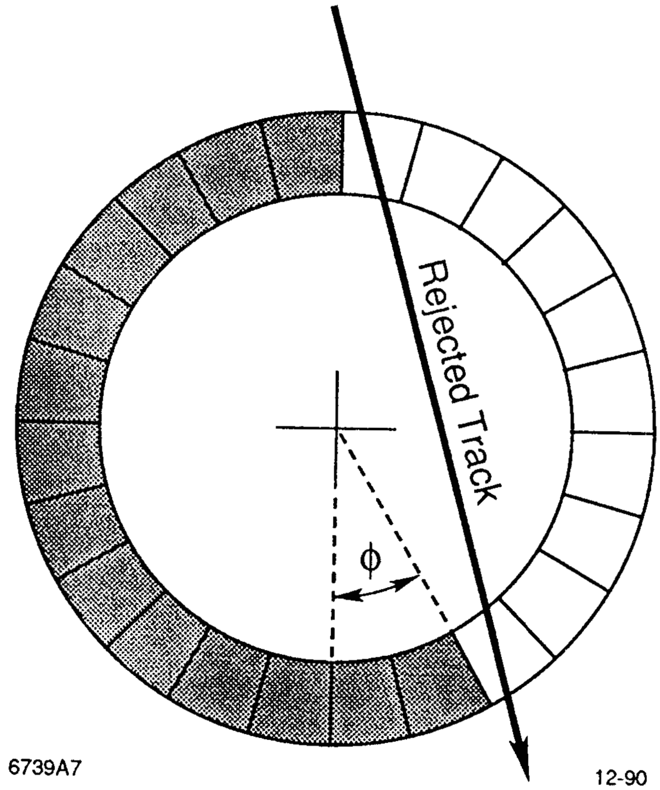


Figure 6. *The geometry of the semi-hemi-cylinder topological calorimeter trigger (SHC). The shaded area indicates the empty SHC in this rejected cosmic ray event. The "slop angle", in this case 30°, is labelled as ϕ.*

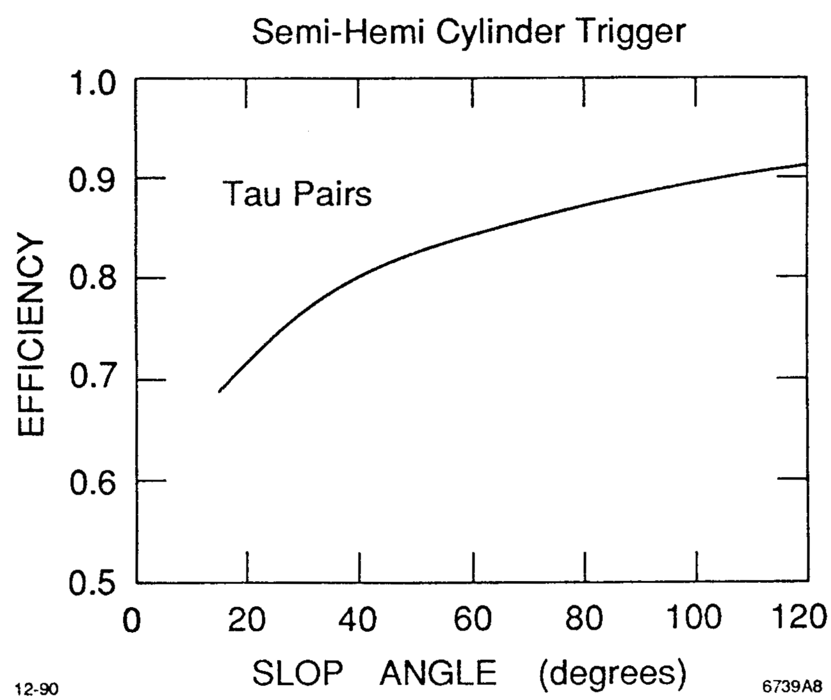


Figure 7. *Efficiency of the SHC trigger.*

5.5. *Tracking Trigger at Level 1*

The tracking node at *Level 1* identifies charged-particle tracks above 50-100 MeV p_t which originate from the region in which the beams interact. It discriminates against low-energy tracks which arise from showers of lost beam particles, and from combinations of random hits. Random hits may arise from synchrotron radiation and other low-energy photons or from out-of-time hits arising from other bunch crossings.

The volume along the beam line from which tracks appear to originate is referred to as the "fiducial" volume for the trigger. The radius of the fiducial volume is determined by the resolution of the track finder and by the radius of the innermost tracking layer used by the trigger. Along the beam line, i.e., in z, the fiducial volume at *Level 1* is determined by the active length of the innermost tracking layer. Consequently, the fiducial volume at *Level 1* will be a few centimeters in radius and a couple meters in length. Less than 100 Hz of cosmic rays, but a few kHz of high-p_t tracks produced by lost beam particles, will traverse this volume.

The tracking node list recognizes track segments in the $r-\phi$ projection of superlayers in the drift chamber and then combines the segments into projected tracks

The tracking node consists of two steps: recognizing segments of tracks in the $r-\phi$ projection of superlayers of the chamber, and linking segments into projected tracks. The segment-finding step will eliminate most out-of-time and other random hits and provide some p_t threshold. The segment linking step will further suppress accidental combinations of random hits and will provide a well-defined p_t threshold. The algorithms by which these steps, particularly segment finding will depend on whether the chamber superlayers are composed of small cells or of jet cells. Figure 8 shows schematically for both small cells and jet cells the logical steps in the tracking trigger.

The tracking node at *Level 1* may require either at least one charged track or at least two charged tracks, depending on its resolution and upon beam-related background rates. If the calorimeter node is efficient for events with two minimum ionizing particles and meets rate requirements, then the tracking node at *Level 1* need only be sufficiently efficient to provide a measure of the efficiency of the calorimeter node.

5.6. Level 1 Tracking Trigger with Small Cells

The small-cell chamber design allows a flexible, efficient trigger that is resistant to random noise. Use of small cells may also allow event time determination to better than ± 1 bunch crossing.

The small-cell chamber design allows a flexible, efficient trigger

The trigger scheme takes advantage of the half-cell-offset chamber geometry. Local track segments are formed in each superlayer. Axial layer segments are linked to form tracks with an associated p_t. Crude stereo information is then added to determine if the track points at the interaction region. Requiring that the innermost superlayer participate in the trigger limits the cosmic-ray rate. Use of stereo information suppresses beam-gas events which do notoriginate in the interaction region, as well as lowering the cosmic-ray rate. Multiplicity and p_t restrictions can be applied to reduce the beam-gas rate further.

Each superlayer of the small-cell drift chamber consists of several layers. The four axial superlayers, moving outward radially, have 3, 4, 5 and 5 layers, respectively. In each case, the layers are rotated clockwise by a half cell with respect to the preceding layer in the superlayer. This geometry is shown in Figure 8a.

A particle passing radially through two adjacent sublayers will deposit ionization on the ith wire of inner layer and on the ith or ith-1 wire of the outer layer. (See Figure 8a.) The sum of the times recorded for the two wires is a constant (the full drift time for the cell). This allows application of a "chronotron" technique. The basic circuit of this technique is shown schematically in Figure 8b. The coincidence unit used can be constructed with shift registers or with tapped delay lines. The tapped delay line coincidence unit is shown in Figure 8c. Shaped, discriminated signals from the ith cell in the inner layer are fed into one end of the line, while signals from the ith cell in the outer layer are fed into the other end of the line. Another delay line is used for the ith inner and ith-1 outer cells. The delay line has the same length as the full drift time, with tap separation corresponding to the inter-bunch spacing. Crossing of the two signals produces a signal on a tap. A beam crossing can then be associated with this signal. In the shift register case, signals are fed to opposite ends of two equal-length registers which are clocked at the crossing rate. The number of register cells corresponds to the drift chamber cell drift time. Having signals in "adjacent" cells of the two registers constitutes an acceptable pair with which a beam crossing can then be identified. The Mark III, which used the chronotron technique with crude cell geometry,[1] achieved better than 20 ns resolution. The spatial resolution expected for the chamber corresponds to about one bunch crossing. Thus one can safely expect temporal resolution, at *Level 1*, of between one and three beam crossings (~5-15 ns).

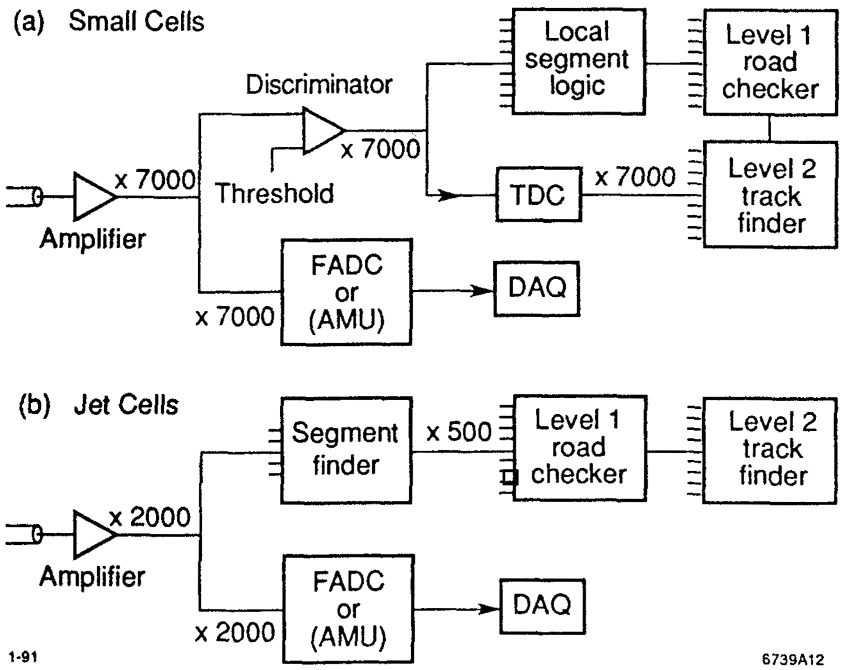


Figure 8. *Block diagram of the logic for tracking triggers for drift chambers with*
a. *small cells,*
b. *jet cells.*

Time-stamped pairs from adjacent layers can now be associated with other pairs in the same superlayer to form track segments. In the first axial superlayer, there are two pairs, which share one cell, which have consistent time stamps. In the second axial superlayer there are three such pairs, with two of them fully independent. For the third and fourth axial superlayers, four pairs are available, two of which are independent. Depending on layer singles rates and layer dead channels, pairs may be AND'ed or OR'ed to produce the time-stamped track segment. The relation of the chronotron coincidence unit and segment-finding logic is shown in Figure 9a.

The track segments in the axial superlayers are formed into tracks by matching the segments found to legitimate track patterns. The block diagram of the axial track finder is shown in Figure 9b. The legitimate track patterns are stored in fast digital logic units (*e.g.* PALs, *etc.*). The pattern matched gives a measure of the p_t of the track. Tracks with acceptable p_t and containing the first and second axial superlayers(A_{12}), the first three axial superlayers(A_{123}), and all axial superlayers(A_{1234}) are counted. The acceptance of the trigger in solid angle can be tuned by imposing requirements upon these track counts. This tuning depends on an acceptable event rate and willingness to cut acceptance for τ pairs. A combination of $A_{12} \geq 2$ and $A_{1234} \geq 1$ gives the desired 1.5 track trigger, while $A_{1234} \geq 2$ corresponds to a two track trigger, which may be required to limit the beam gas rate. This track matching, counting, and selecting process can be completed within 1 μs of the occurrence of the event.

While the axial tracking decision is made, time-stamped pairs and segments are formed in the six stereo superlayers. There are U and V superlayers between each axial superlayer set. Segments are matched with axial tracks that emerge from the axial tracking decision. The block diagram of the 3-D track and vertex finder is shown in Figure 9c. As in the axial case, the number of 3-D tracks with a minimum momentum that point toward the interaction point are counted, and cuts are applied consistent with a 1.5 (or 2) track trigger. It is expected that this decision will not be available until at least 1.5 μs after the interaction. In order to make an earlier z-tracking decision, pads on the inner and outer walls of the drift chamber may be used to form primitive z tracks and to verify that these point to the interaction region. This would be achieved easily by restricting the z range of the inner wall pads. However, this option is sensitive to random hits on the inner wall, requiring a large number of channels with accompanying electronics.

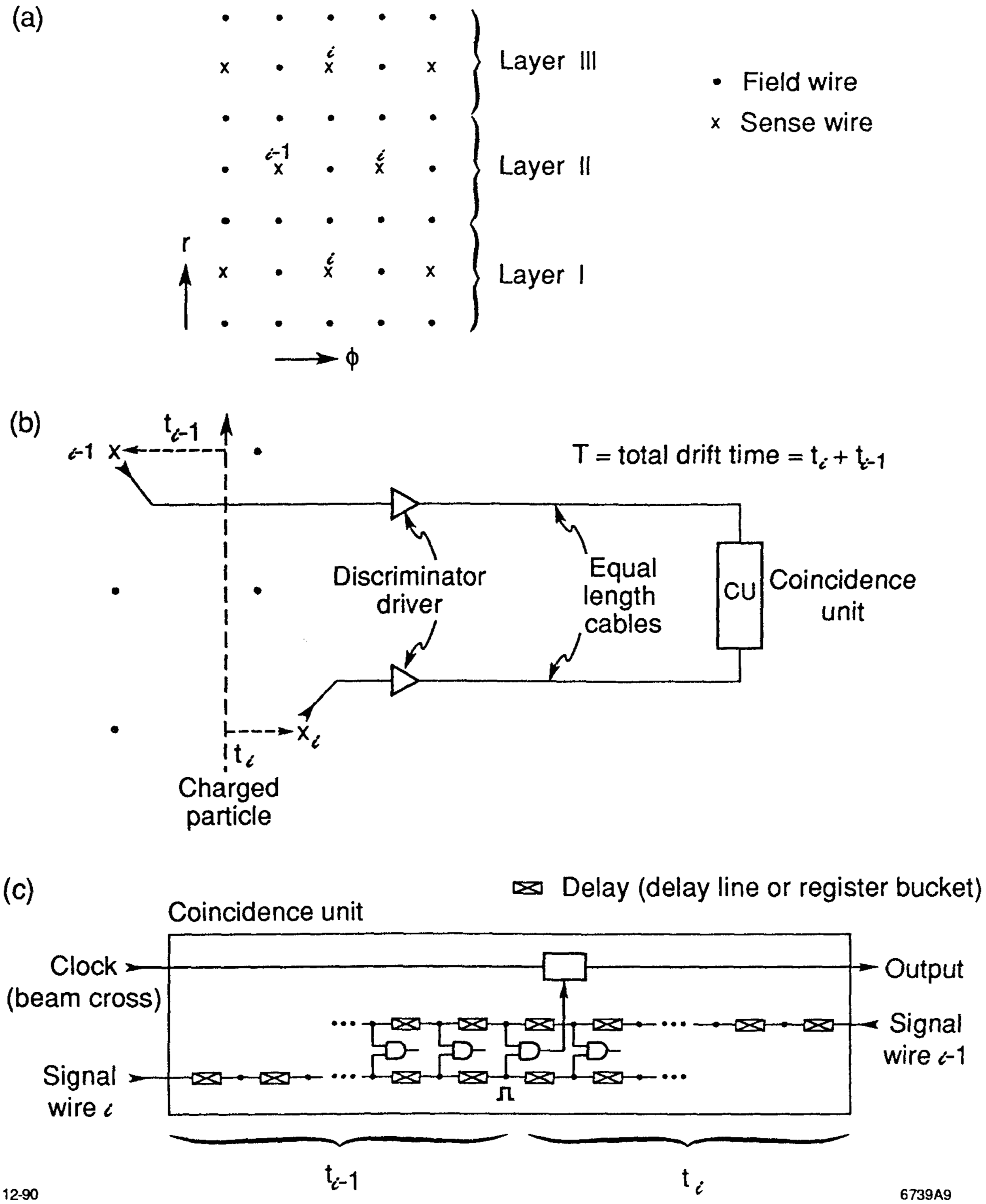


Figure 9. *Chronotron logic for pairs of drift chamber layers.*
a. *Slice in ϕ of axial superlayer 1, with three layers of sense wires.*
b. *Expanded view of sense wires i and $i-1$ with discriminator/drivers, equal-length cables and coincidence unit*
c. *Coincidence unit, showing delay line technique and time-stamping.*

5.7. *Level 1 Tracking Trigger with Jet Cells*

For a drift chamber with jet cells, logic on each cell communicates with the *Level 1* tracking node (DC $L1$ node), as illustrated in Figure 10.

The logic on each jet cell searches for track segment candidates. It examines all hits within the maximum allowable drift time, checks them for consistency with a straight segment, and crudely measures parameters of found segments. Segments are identified with a cell ID, a time stamp, and a coarse measure (approximately 5 bits) of absolute slope in $d\phi/dr$. The cell logic then sends a track-segment message to the DC $L1$ node. A cell which receives only one or two hits within the drift time sends a track-fragment message to the DC $L1$ node.

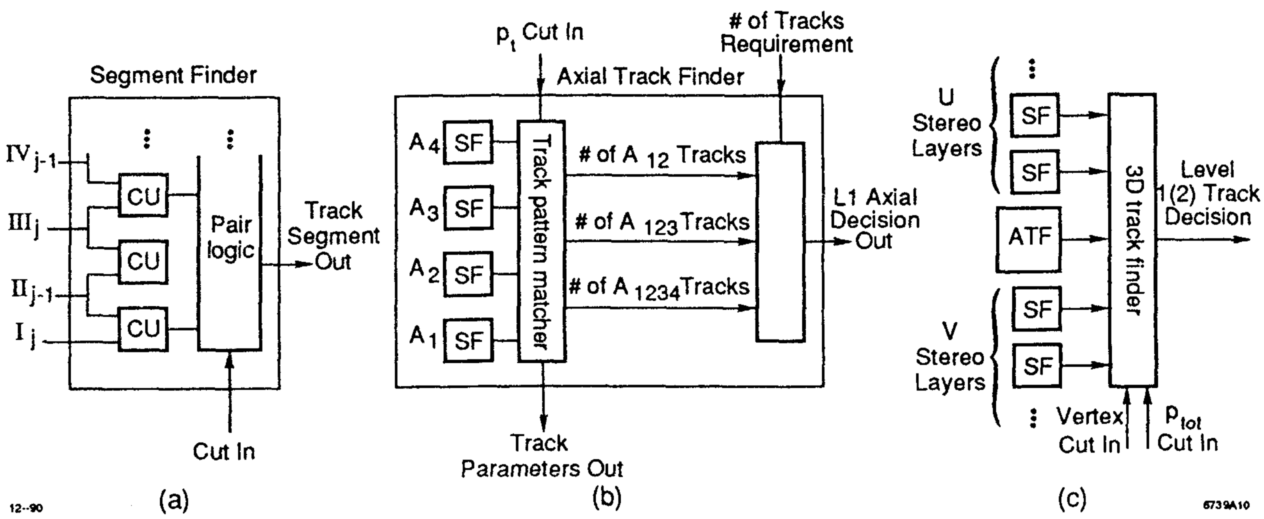


Figure 10. *Track-finding logic for small cell chamber:*
a. *segment finder (SF),*
b. *axial track finder (ATF),* and
c. *3D track and vertex finder.*

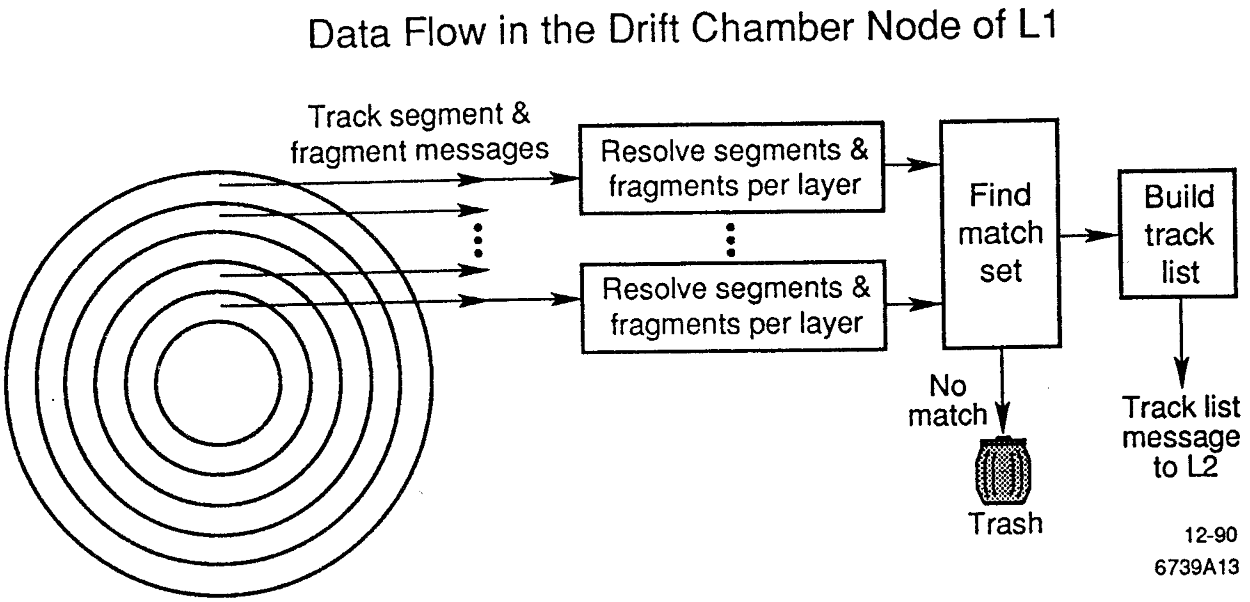
The DC $L1$ node collects messages from segment-finders. The syntax of these messages, for both segments and fragments, is shown in Figure 11. The DC $L1$ node then resolves fragments into segments and looks for a set of track segments (the match set) which:

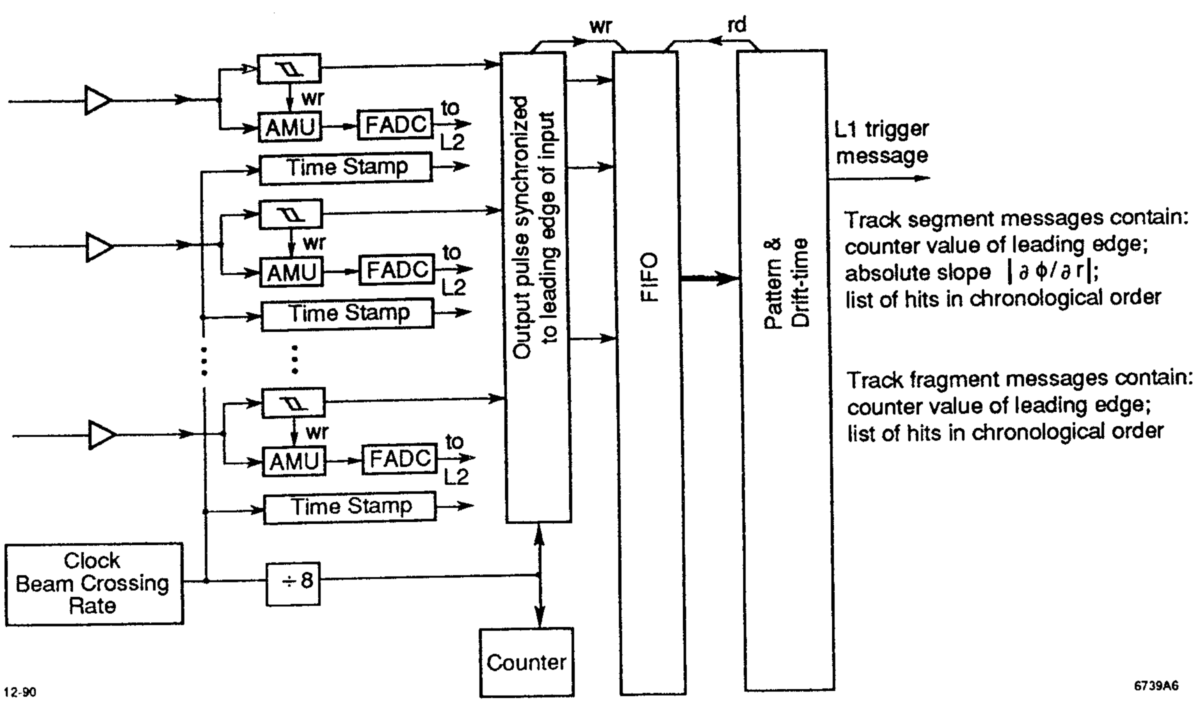
1. involve the first four or five layers,
2. come from cells with similar ϕ in each layer,
3. have similar slope,

4. have consistent time stamps, and
5. project into the L1 fiducial volume in $r - \phi$.

All jet-cell track segment messages within a drift time of the match set are sent along to the *Level 2* trigger for further processing.

Figure 11. *Information flow in the track-finding logic for a central drift chamber with jet cells.*





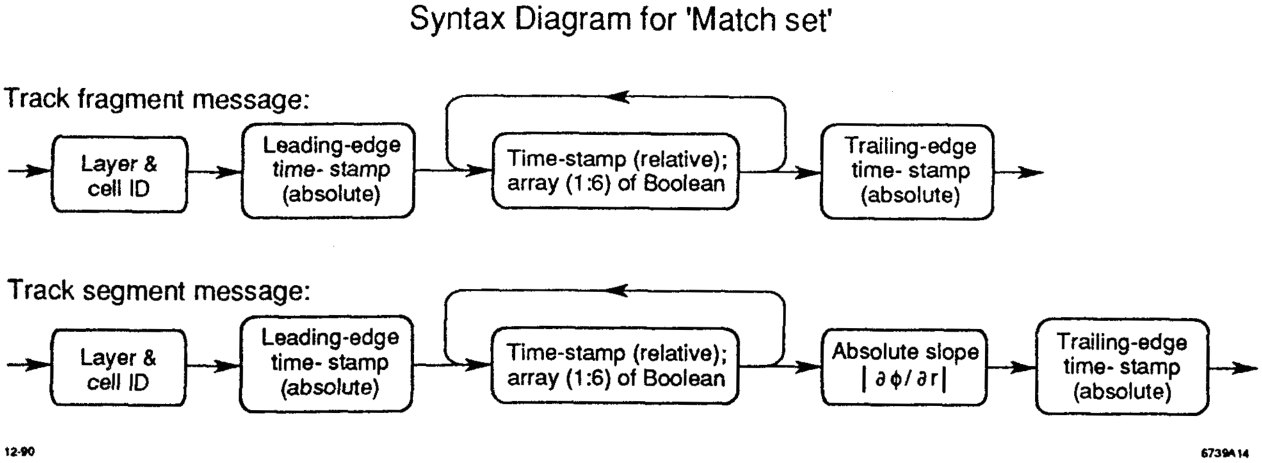


Figure 12. *Syntax of trigger messages for track fragments and track segments.*

6. *Level 2* TRIGGERS

Cosmic rays and lost beam particles dominate the trigger-rate into Level 2

The trigger rate into *Level 2* will be dominated by cosmic rays and by coincidences of secondaries from showers of one or two lost beam particles. These backgrounds have similar two-prong topologies as some interesting physics processes, such as production of τ-pairs; however, background tracks in general will not originate at the beam interaction point. Consequently, the principal manner by which the *Level 2* trigger can reduce the trigger rate is by restricting the fiducial volume for tracks in $r - \phi$ and in z. This is accomplished using tracking information of higher resolution than was used in *Level 1.* With the exception of triggers on totally neutral events, calorimetric information is not important at *Level 2* since it has already been used at *Level 1.* With adequate resolution near the vertex in *Level 2*, the trigger rate from *Level 2* may be dominated by beam-gas backgrounds within the vertex volume.

The rate of cosmic rays into *Level 2*, which may be as high as 3 kHz, is based on the large fiducial volume accepted by the *Level 1* calorimeter node. This rate may be easily reduced by requiring more than one charged-particle track within a smaller fiducial volume. For instance, a fiducial volume comparable to that used in the *Level 1* tracking node, which is a few centimeters in radius and from ten to a few tens of centimeters in length, would result in a cosmic rate of order 50 Hz, or roughly the same as the physics rate with a luminosity of 10^{34} . Further reduction in rate is possible either with further reduction in the *Level 2* fiducial volume or by use of time-of-flight information; however, further reduction at *Level 2* is not critical. The cosmic rate is within the overall trigger rate goal of 100 to 1000 Hz for *Level 2* and is less than other rates at that level. If further reduction in the cosmic rate is necessary, reduction at *Level 3*, where more data and resolution is available, may be more suitable.

The background rate into *Level 2* from charged tracks produced by lost beam particles is expected to be between hundreds of hertz and about 4 kHz (see Table 3). It also can be sharply reduced by restricting the fiducial volume, particularly in z. The rate is expected to drop roughly in proportion to reduction in z of the fiducial volume. Thus, more than one order-of-magnitude reduction is possible by reducing the z acceptance to order five centimeters. Further reduction in the z acceptance will be limited by the length of the beam bunches in z; the fiducial volume cannot be made smaller than the luminous region. Consequently, to further reduce backgrounds from lost beam particles it will be necessary to reduce the fiducial volume in $r - \phi$. Such reduction may be possible using either drift time measurements or by requiring that tracks have associated hits in the silicon vertex detector. (See report of the Vertex Detection Group). Reduction of these backgrounds with reduced acceptance in $r - \phi$ has not yet been studied.

Background from inelastic scattering of beam particles with residual gas within the vacuum at the interaction point (beam-gas) do not dominate the trigger rate into *Level 2*. However, it may be the most difficult to eliminate since the tracks originate along the beam line, rendering acceptance cuts in $r - \phi$ ineffective. As in the case of backgrounds from lost beam particles, cuts on acceptance in z are limited to the length of the bunches. Based on MARK II experience at *PEP*, the beam gas rate may be low, of order 5 - 10 Hz with vacuums comparable to *PEP*. However, if vacuums are two orders of magnitude worse than *PEP* at the interaction point due to the higher beam currents and the smaller beam pipe, beam gas backgrounds may be serious, of order 500 Hz.

If beam-gas backgrounds are severe, it may become necessary to make additional topology cuts at *Level 2*. The asymmetric beam energies make cuts on momentum balance in z unusable; however, cuts on p_t balance, acoplanarity, and if necessary multiplicity may be effective. Any such cut will reduce efficiency for detection of τ pairs, but will probably have little effect on B events. The rate of beam-gas background and techniques to reduce the rate have not yet received ample study.

6.1. *Identifying Bhabhas at Level 2*

Bhabhas will be identified at Level 2 to provide a real-time measure of luminosity and to prescale the Bhabha rate which is written to tape

The *Level 2* trigger identifies (non-radiative) Bhabha events on the basis of their kinematic signature in the calorimeter. The two main goals of recognizing Bhabhas are:

1. to report to the accelerator a real-time measure of luminosity, and
2. to pre-scale Bhabhas should the event rate strain the on-line computing power.

Consequently, Bhabhas are identified with a set of loose cuts for the luminosity monitoring, and a set of stringent cuts for pre-scaling. Other goals include online calibration of the calorimeter towers, as well as of drift chamber gas gain and drift velocity. The *Level 3* trigger might be to provide to the accelerator an online measure of the position of the interaction point as determined by Bhabha events.

Each tower in the calorimeter has associated with it a nominal Bhabha deposition energy, dependent only on its position in θ. If the supertowers are "squarish", the variation of nominal Bhabha deposition energies in constituent towers is small, and the supertower also has a well-defined nominal Bhabha deposition energy.

The classification process is initiated if *Level 2* candidate exhibits:

> exactly two supertowers, whose measured energies are near their nominal Bhabha deposition;

AND the measured energy in no other supertower in any layer (not contiguous to the supertowers above) exceeds a small fraction of the its nominal Bhabha deposition;

AND the entire calorimetric energy sum lies near the nominal Bhabha value.

Then the event is classified as a Bhabha if:

the $\theta - \phi$ positions of the participating supertowers are consistent with the Bhabha topology, as stored in a look-up table;

AND the *Level 2* tracking node reports a sufficiently stiff track pointing at each of the two supertowers;

AND no other tracks are found.

7. *Level 3* TRIGGERS

The trigger rate into *Level 3* will have approximately equal contributions of cosmic rays, backgrounds from lost beam particles, and physics events. In addition, there will be a contribution from beam-gas events which may be comparable to the other contributions, or which may be either larger or smaller depending on the vacuum at the interaction point. If all other backgrounds into *Level 3* are comparable to the physics rate, then it is reasonable to record all events without further selection at *Level 3*. However, a *Level 3* trigger can be implemented to further reduce background or to battle beam-gas backgrounds.

The trigger rate into Level 3 will have equal contributions from cosmic rays, backgrounds from lost beam particles and physics events

The *Level 3* trigger has available the same data and the same processing engines which are available to offline analysis. It differs from offline analysis in that refined detector constants are not available and in that interactive choice of event selection criteria is not possible. However, given enough processing power in *Level 3*, it will be capable of performing online nearly any further event selection done offline. In particular, *Level 3* will be capable of doing refined tracking and some vertexing to further reduce backgrounds which do not come from the interaction point. It will also be capable of applying complex topological cuts which may be necessary to eliminate beam-gas backgrounds which originate within the interaction volume. The residual backgrounds expected from *Level 2* are all low multiplicity backgrounds, which *Level 3* should be able to reconstruct and analyze with modest computing power.

8. Conclusions

This study has focussed on identifying a multilevel trigger architecture that can manage the high-rate environment of the Asymmetric *B* Factory. The challenges posed are all related to the high bunch-crossing frequency in the new machine. Accordingly, greatest consideration has been given to understanding how to reduce machine-related backgrounds to acceptable levels, while maintaining efficiency for all the interesting physics processes available for study.

Reduction at *Level 1* is the most challenging. It confronts the highest rates with the least amount of processing time. Backgrounds from synchrotron radiation and soft lost beam particles must be eliminated at that level. The high resolution of the calorimeter offers the ability to perform sufficient reduction at *Level 1* with calorimeter information only. A *Level 1* tracking trigger would be capable of providing a lower trigger rate, but would require considerable complexity of processing and connection of nearly all central drift chamber wires. Some tracking trigger at *Level 1* will be necessary to determine the efficiency of the *Level 1* calorimeter trigger.

The trigger at *Level 2* will necessarily be based upon tracking to reduce backgrounds of tracks which do not originate from the interaction point. A *Level 3* trigger can be used to perform final, sophisticated event selection, in order to make the amount of recorded data more manageable.

The biggest uncertainty in our understanding of trigger rates is in the rate of beam-gas backgrounds, which depend upon the quality of the vacuum at the interaction point in the difficult environment of high-current beams. In addition, our conclusions on backgrounds from lost beam particles are sensitive to the accelerator lattice. These conclusions should be refined as the accelerator design evolves.

References

1. J.J. Thaler, *et al.*, *IEEE Trans. Nucl. Sci.* **30**, 236 (1983).

An Asymmetric B Factory at KEK

Study Group of the KEK B Factory

Presented by T. Nozaki

Abstract

The status of the Asymmetric B Factory Project at KEK is reported. The design of the detector for the KEK B Factory is mainly discussed.

1. Introduction

In this report we report on the status of the asymmetric B Factory project at KEK.

A Task Force to study feasibility of an asymmetric B Factory at KEK started in June, 1989. The following important dates should be noted:

Jun'89 : A Task Force started.

Oct'89 : The first workshop on asymmetric B Factory was held.

Dec'89 : The second workshop on asymmetric B Factory was held.

Feb'90 : The Task Force report was published[1].

Mar'90 : TPAC (TRISTAN Program Advisory Committee) recommended the Task Force to complete the conceptual design of the machine in 10 months.

Oct'90 : The International Workshop on Accelerators for Asymmetric B Factories was held at KEK.

April'91 The International Workshop on B Factory detectors will be held at KEK

Both the accelerator and the detector were designed based on the following physics requirements.

i) CP violation can be measured within several years operation for $0.1 < \sin\phi < 0.6$.

ii) The B_s^0 mixing parameter x_s can be measured up to 12 within a few years operation.

The design of the accelerator is described briefly in Section 2. Then the design of the detector is discussed in Section 3.

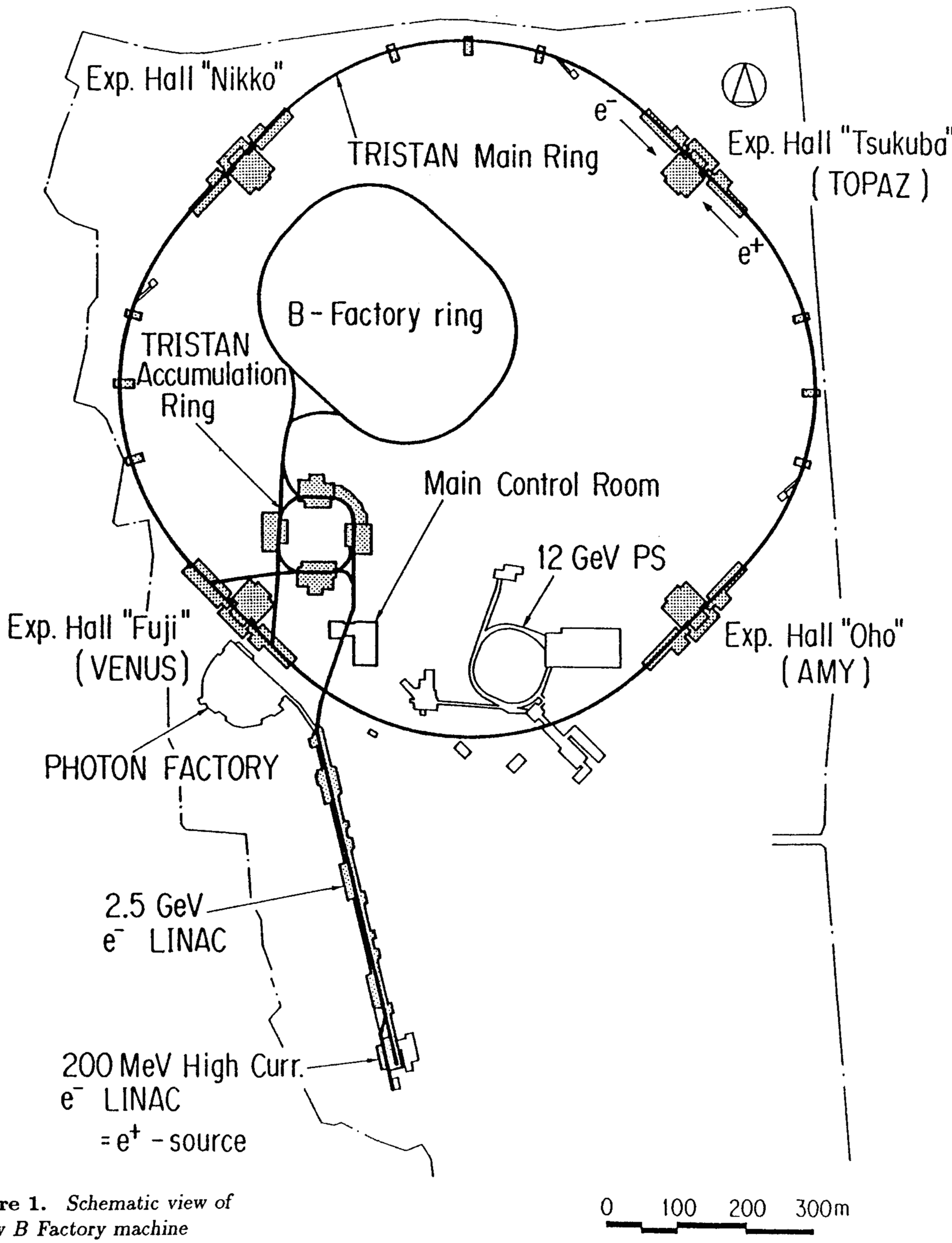


Figure 1. *Schematic view of a new B Factory machine*

2. Accelerator Design

The beam energies and the required luminosities were determined based on the physics requirements and the assumption of the use of the present 8 GeV accumulation ring for the TRISTAN.

i) Measurements of CP violation requires 3.5×8 GeV collision at $\Upsilon(4S)$ with the peak luminosity, $L = 10^{34} cm^{-2}s^{-1}$.

ii) Measurements of B_s^0 mixing requires 2.46×12 GeV collision at $\Upsilon(5S)$ with $L = 10^{33} cm^{-2}s^{-1}$.

The accelerator complex is planned to be constructed in two steps.

Phase 1 : Build a 3.5 and 8 GeV two-ring collider with the same circumference rings in a newly constructed tunnel. An initial luminosity of $2 \times 10^{33} cm^{-2}s^{-1}$ is foreseen and then eventually improved to $10^{34} cm^{-2}s^{-1}$.

Phase 2 : Go to 2.46×12 GeV by adding a 12 GeV booster ring inside the same tunnel if necessary.

The accelerator complex at Phase 1 consists of the upgraded 2.5 GeV linac, the 8 GeV accumulation ring, and 3.5 and 8 GeV collider rings. Their arrangement are shown in Figure 1.

The design parameters of the colliding rings are shown in Table 1. The details of the accelerator is described in ref. 2.

The magnet configuration and beam orbits at the collision point are shown in Figure 2(a). The requirement to make a multi-bunch operation and a head-on collision brings us the present design of the colliding section. Beams are bent by a 0.875 Tesla bending magnet of 40 cm long which is placed 40 cm away from the collision point. A following quadrupole doublet helps to separate beams further by placing quads off the center of beam orbits. The septum magnet bends the low energy beam further. It has to be placed outside the solenoid field of the detector. This requirement strongly restrains the usable solenoid volume. The distance between the surface of the end yoke to the collision point has to be about 2 m. If we have to place a compensation magnet to reduce the effects of the solenoidal magnetic field on the beams, the usable solenoid volume is further reduced in the forward. The current detector is designed for this magnet configuration, but we are also studying another magnet configuration as shown in Figure 2(b) which provides us more detector space.

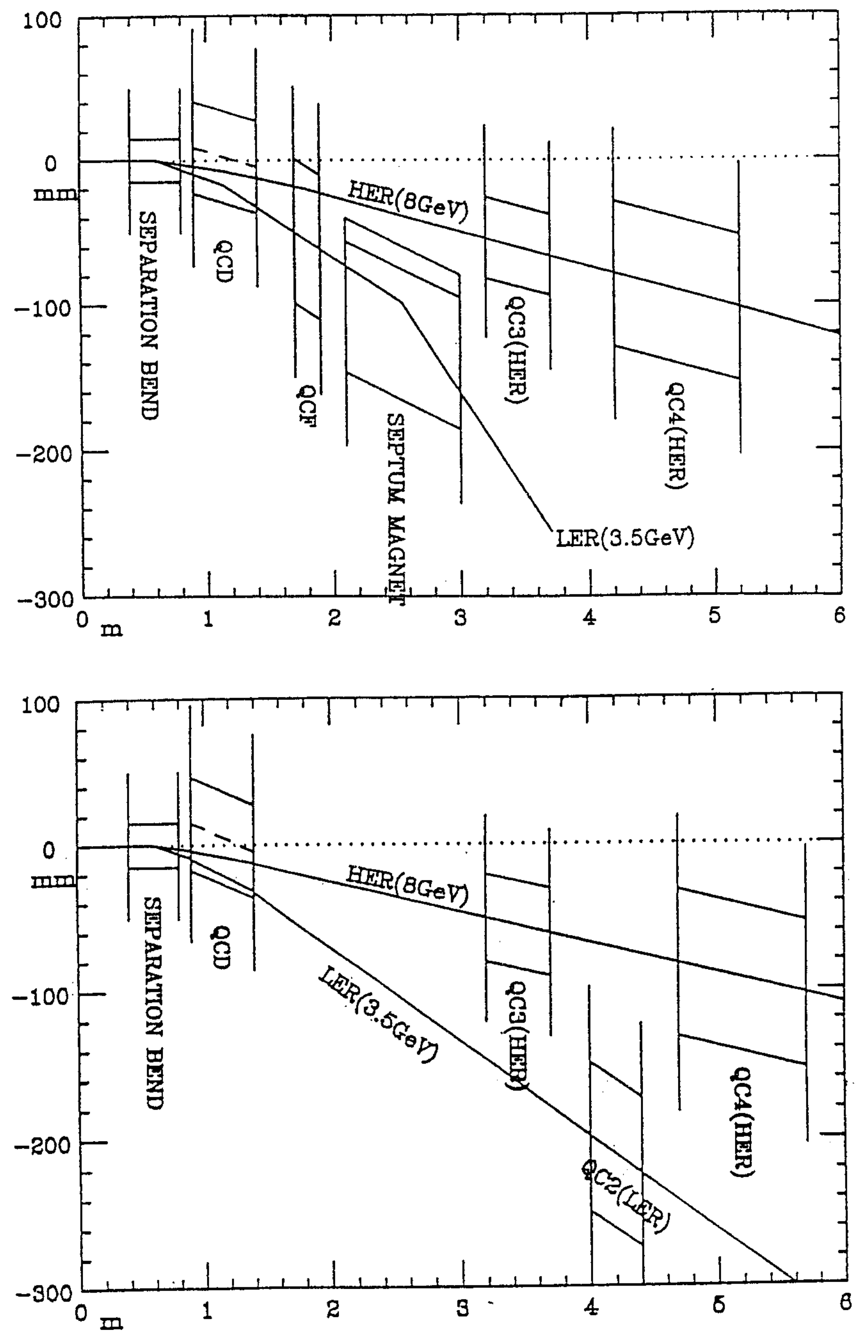


Figure 2.(a) *The magnet configuration and beam orbits at the collision point. a) The first version, b) the new version*

3. Detector Design

We have studied the required performance of the detector elements and the possible configuration of them for the detection of the most important processes such as, $B^0 \to \psi K_S^0$, $B^0 \to \psi K_S^0 \pi^0$, and B_s mixing. Because of the asymmetric collision of electron and positron beams, the detector is also asymmetric in the direction of the beam. Since the general features of B meson decay, namely particle multiplicity, momentum distribution, average momentum in the final state and so on, are known relatively well, the designing of the detector is rather straightforward. Study was made with the assumption of the use of a solenoidal magnet spectrometer. (The dipole magnet option is also interesting because it can be used also as a separation magnet of two beams.) The detector consists of a particle tracker such as a silicon vertex detector (SVD), a precision drift chamber (PDC), and a central drift chamber (CDC), an electromagnetic calorimeter (EMC) such as a CsI crystal array, a particle identifier such as a time-of-flight counter (TOF) and a Ring-Image-Cherenkov counter (RICH) and a muon detector. All of the detector elements except for the SVD and RICH can be constructed using well established techniques. For the SVD and RICH we need R&D. The schematic view of the detector is shown in Figure 3.

All the detector elements cover the angular range between 12^o and 135^o which corresponds to the center-of-mass angular range of $\cos\theta_{cm}$ between -0.95(-0.86) and 0.90(0.95) for 2.3 $\times$ 12 GeV (3.5 $\times$ 8 GeV) collisions. The current angular coverage is optimized for the 2.3 $\times$ 12 GeV collision. If we optimize the angular acceptance for 3.5 $\times$ 8 GeV, the angular range between 15° and 150^o should be covered by the detector.

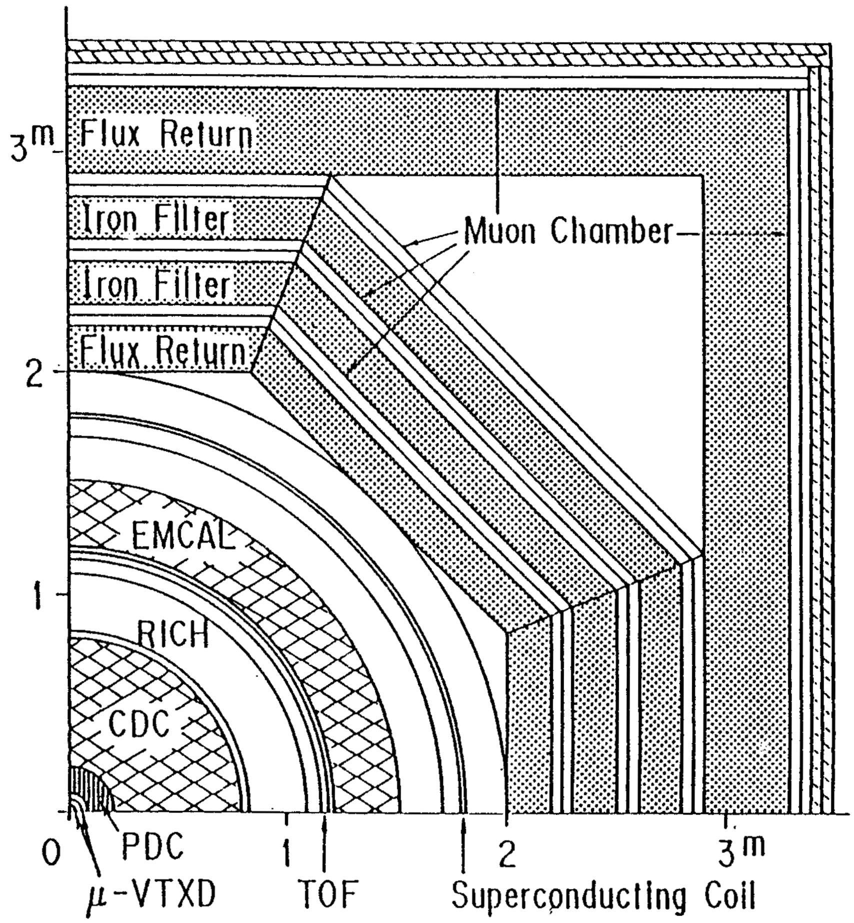
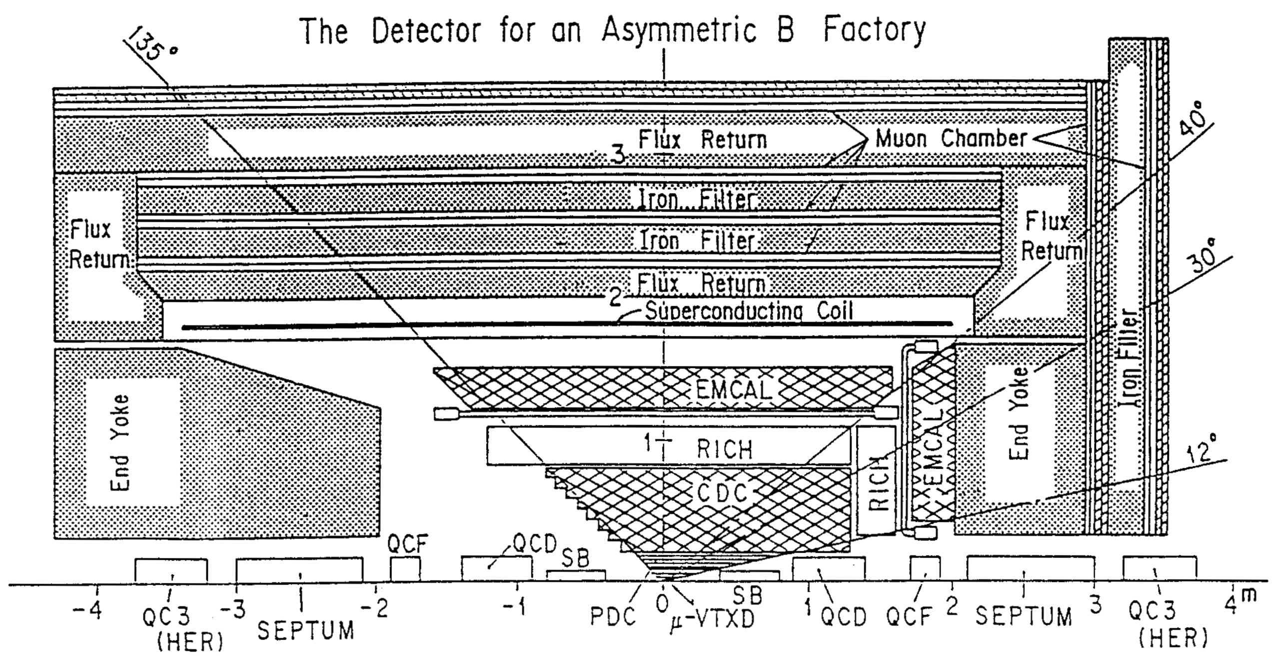


Figure 3. *Schematic view of the detector for the asymmetric B Factory at KEK.*
a. *rφ view,*
b. *rz view*



3.1. Beam Pipe and Vacuum

In order to reduce the effect of multiple scattering in determination of vertex points we have to minimize the radius and the radiation length of the beam pipe at the interaction point. Therefore we use a beryllium beam pipe with 500 μm thickness. The beam background problem may not allow radius of the beam pipe smaller than 2 cm.

The vacuum in the beam pipe has to be better than 10^{-9} torr in order to obtain long life time of the beam. Good vacuum is also very important near the collision point for keeping the beam background in the detector as low as possible. We propose a design of the beam duct whose cross section is shown in Figure 4. The beam duct

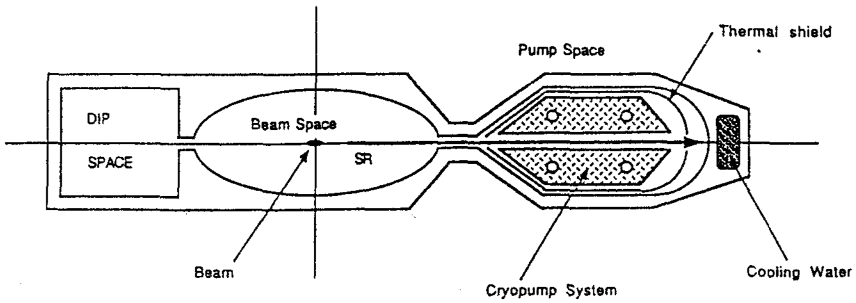


Figure 4. *Proposed beam pipe and pumping system for the arc section of the KEK B Factory*

is divided into two cells, namely, a beam cell and a pump cell. Two cells are connected by a small slit. The electron beam runs in the beam cell and the synchrotron light produced by the electron beam goes through the slit and hits the wall in the pump cell. The gas produced by the synchrotron light is pumped out by a cryogenic pump placed in the pump cell. Therefore outgas due to the synchrotron light should not leak into the beam cell. In order to obtain 10^{-9} torr vacuum we need average distributed pumping speed of 300 $l/s/m$ or gas desorption coefficient of 10^{-6} molecules/photon.

The heating-up of the beam pipe due to the ohmic loss and the higher-order-mode loss has to be also considered. The synchrotron lights generated from the beam separation magnets hit the beam pipe around the collision point and increase the temperature of the beam pipe furthermore. As a result cooling of about 10 kW/m is needed. The cooling system has to be designed not to interfere the particle detection. In order to reduce the heat-up of the beam pipe due to the synchrotron lights we are considering to use a beam pipe of baloon shape which has a larger radius in the region where synchrotron lights hit than the residual part. The feasibility of such a shape of beam pipe is under study.

3.2. Silicon Vertex Detector

Both the measurements of *CP* asymmetry and B_s^0 mixing require the accurate measurement of the distance between two B decay vertices (Δz). For such experiments we need to observe oscillations proportional to $\sin(x_d \Delta z / c\gamma\beta\tau_d)$ or $\cos(x_s \Delta z / c\gamma\beta\tau_s)$, respectively, where $\gamma(\beta)$ is the Lorentz factor (velocity) of the Υ in the laboratory system and $\tau_d(\tau_s)$ is the lifetime of $\Upsilon(4S)(\Upsilon(5S))$. Since the value of x_d is equal to 0.7, accuracy of Δz is only required to be better than half of the mean decay length, $c\gamma\beta\tau_d$ of B_d^0 for the *CP* experiment. As a result the *CP* measurement requires 150μm and 70μm resolutions of Δz for the beam energies of 2.3×12 GeV and 3.5×8 GeV, respectively. Since the value of x_s is considered to be one order larger than the x_d value, we need to have much better accuracy of Δz for the B_s^0 mixing experiment. According to the Monte Carlo simulation we have to have 40 μm resolution of Δz to observe x_s value up to 5 (12) for the beam energies of 3.7×8 GeV (2.5×12 GeV) for the total integrated luminosity[3] of $10^{40}\mathrm{cm}^{-2}$.

In order to achieve such a good Δz resolution we need to measure track coordinates with precision better than 10 μm for normal incidence and 20 μm for inclined incidence at the radial distance of a few cm from the interaction point using a micro vertex detector. Among the presently available devices, the silicon strip detector provides the best position resolution, several μm for normal incidence. If such a high resolution tracking device is used, uncertainty of the measured vertex position is dominated by the error introduced by the extrapolation of the reconstructed track to the collision region. Therefore every effort has to be made to reduce the beam pipe diameter as well as thickness of the beam pipe and silicon detectors. This requires careful protection of the detector against the beam backgrounds and readout of two orthogonal coordinates from one plane, which can be achieved by using double-sided silicon strips.[4]

Inclined Angle Dependence of Position Resolution

The position resolution of a silicon detector deteriorates as a particle enters the detector with large inclined angle because ionization energy deposition per strip becomes smaller in this case. Correctly evaluating this effect is important in the case of a B factory experiment because a large fraction of the charged tracks go forward. We measured the inclined angle dependence of the position resolution and compared the result with an estimate which was obtained using a Monte Carlo simulation.

The measurement was performed in a setup shown in Figure 5(a), where a double-sided detector with 100 μm strip spacing was placed between two one-sided devices with 25 μm strip spacing.

The double-sided type was used only because this measurement was done as a part of evaluating the performance of this type of silicon detector. The two 25 μm devices provide the reference position and the measurement was done at two inclined angles, $\theta = 30°$, and $60°$. Then we compared the measured resolutions with the Monte Carlo prediction.

The Monte Carlo signal in each strip is generated by first assigning a mean signal which is proportional to the pass length. It is then smeared according to the Landau distribution and a noise signal is added. We used two different types of noise signal. One is a Gaussian with the σ= 6 KeV and the other with σ=3.6 KeV, corresponding to the noise signal observed for VLSI (MX3)[6] and for Hybrid (Lecroy HQV810), respectively. The energy deposit of 6 KeV is equal to that of a minimum ionizing particle with 25 μm pass length. Figure 5(b) shows examples of the simulated signals. The strips in an area between two star marks are the ones fired by a particle. The number of those fired strips is counted under the assumption that the incident angle of the particle is known. Such a group of consecutively fired strips is defined as a cluster. If more than one cluster are found, the one with largest total charge is chosen as the correct cluster. The hit position is determined by two different methods: a) Using the pulse heights of the two edge strips the center of the cluster is calculated, b) The average position is calculated by weighting all fired strips with their pulse heights. The distribution of the difference of the calculated position from the true position determines the resolution. The two methods of calculating the hit position in the silicon detector give a similar resolution. Comparison of the measured resolutions with the Monte Carlo prediction is shown in Figure 5(c).

The data are fairly well reproduced by the Monte Carlo prediction. It should be noted that the analog data from every strip have to be recorded to achieve the above resolutions. Also should be noted is that the above resolutions can be achieved only when the incident angle of the particle is known from the other detectors. More measurements are needed before we can make a definitive statement. However, it appears that we can have a space resolution of better than 25 μm at an inclined angle $\theta = 10°$.

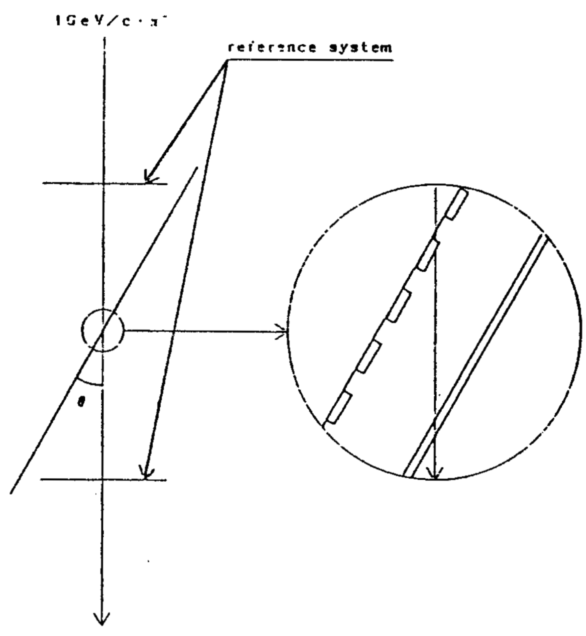


Figure 5. **a.** *A setup that was used to measure the inclined-angle dependence of the position resolution of a double-sided silicon strip detector.*

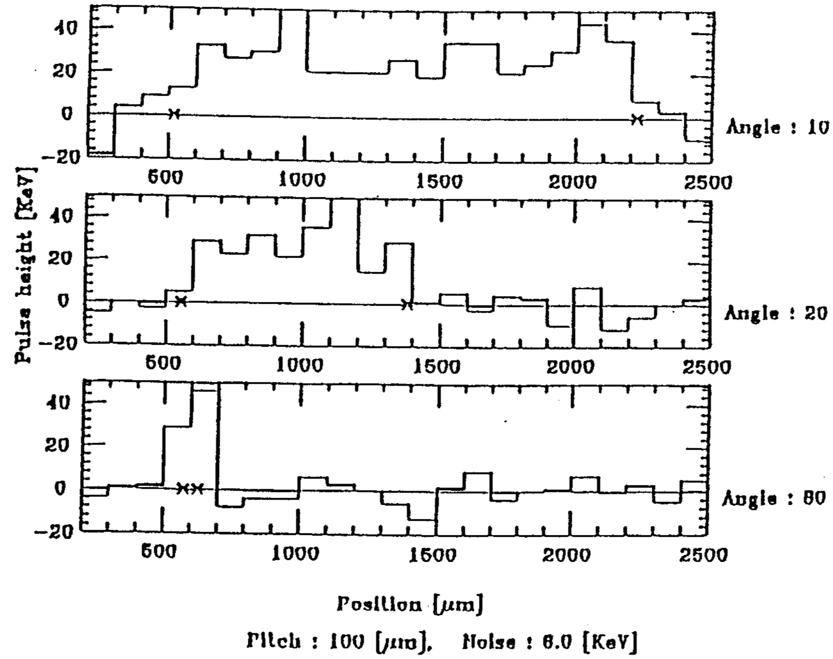


Figure 5. **b.** *Simulated signals for penetrating particles at inclined incidence.*

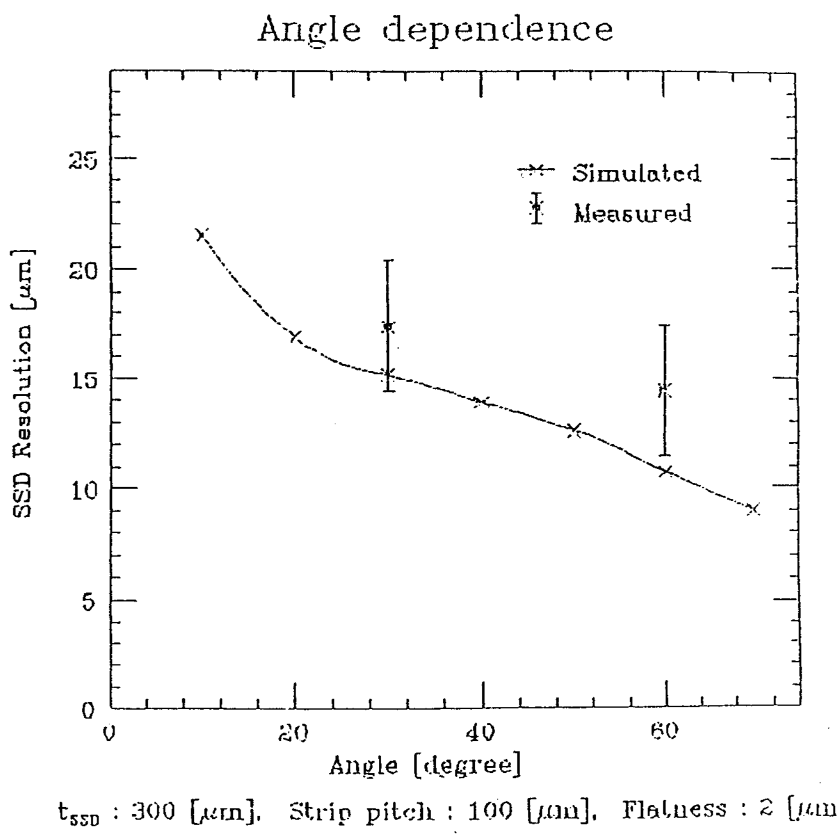


Figure 5. c. *Comparison of the measured position resolutions with the Monte Carlo prediction. In the simulation the flatness of the silicon strip is assumed to be 2 μm.*

Required R&D

As described in the previous section, in order for the silicon strip detector to perform as a high precision vertex detector at a B Factory experiment, it must provide the two-dimensional readout from one plate and it must be equipped with a readout electronics which is fast enough not to cause a significant dead-time to the experiment. Providing a fast output signal is also important because adding the signal from the silicon detector to the fast trigger would be quite powerful to reduce the background trigger rates.

A group of Nagoya university is cooperating with Hamamatsu Photonics Corporation in developing a double-sided silicon device and several prototypes of such detector were manufactured and tested.[5] In these samples, isolation between neighboring n-strips is achieved by placing additional p-strips between adjacent n-strips. A schematic layout of this type is shown in Figure 6. They obtained samples of 6.4 mm $\times$ 6.4 mm $\times$ 300 μm size with the strip spacing of 25 μm, 50 μm, and 100 μm. So far various measurements for evaluating their properties were done mainly on the 100 μm sample. The results indicated high enough resistance between n-strips and between a n-strip and a bias ring. The pulse height correlation between the two sides for 1 GeV π beam is shown in Figure 7.

The bias voltage was 70V and Lecroy HQV810 was used as a readout amplifier. The noise level of this amplifier is about 900 electron equivalent and the peak of the signal distribution corresponds to 22,000 electron equivalent. Clear correlation between p and n-side can be seen. Work will be continued on this type of the two-dimensional readout device. A KEK group is cooperating with P. Weilhammer (CERN) in developing another type of double-sided silicon device in which isolation of n-strips is achieved by a bias applied to n-strips.

Various readout electronics for the silicon vertex detector have been developed using CMOS type VLSI, for example, MX3 at RAL,[6] CAMEX at MPI[7] and SVX at LBL.[8] The first two VLSI's are already used in the LEP and SLC experiments. Three ways of reading out the silicon detector signals which could be pursued for the *B* Factory detector are being developed.

i) Improve the current CMOS type VLSI by reducing number of reset using an amplifier with large dynamic range as well as by using analog switching capacitor technique for sample-and-hold circuits. Another possibility for avoiding dead time is to prepare several sets of amplifiers and sample-and-hold circuits and switch to the other when an amplifier is reset. This scheme seems difficult and need more investigation.

ii) Use a bipolar current preamplifier followed by a shaping amplifier.[9] Only discrete hit information via discriminator is read out. The rise time of this type is faster than 10 ns. In this case, getting fast trigger signal from output hit pattern is straightforward. The spatial resolution is given by (strip spacing)/$\sqrt{12}$. Therefore, in order to have 10 μm resolution, the strip spacing of $\sim$ 35 μm is required and this imposes a tight requirement for size of electronics chip. With present technology, bipolar IC requires larger transistors and resistors than CMOS and it is difficult to achieve required size.

iii) Store output of a current amplifier in CCD, then readout the content when a trigger occurs. In this case, it is quite difficult to get fast output signal for trigger. The current problem is a poor efficiency to store charge from a silicon detector to CCD memory and is under investigation. Application for the *B* Factory detector requires further investigation.

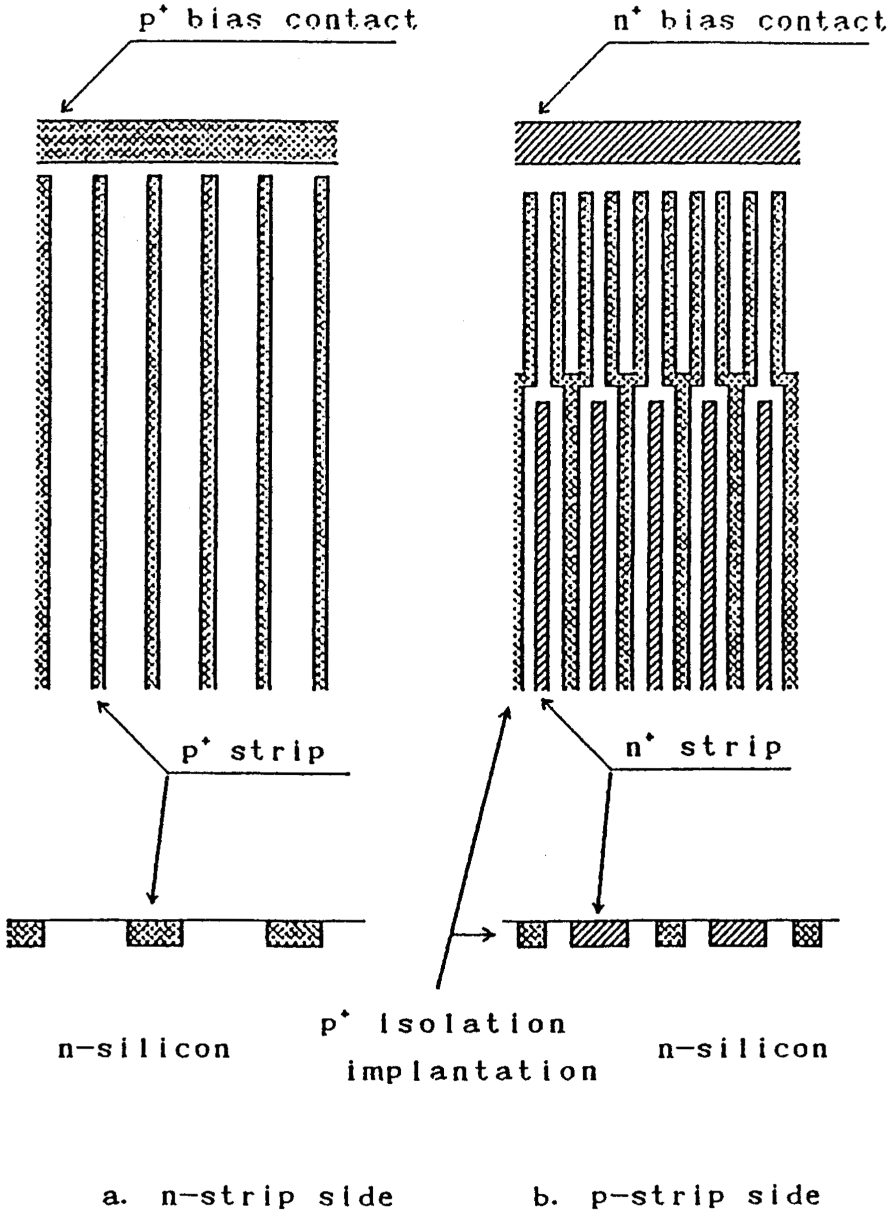


Figure 6. *Schematic view of a double-sided silicon strip detector*

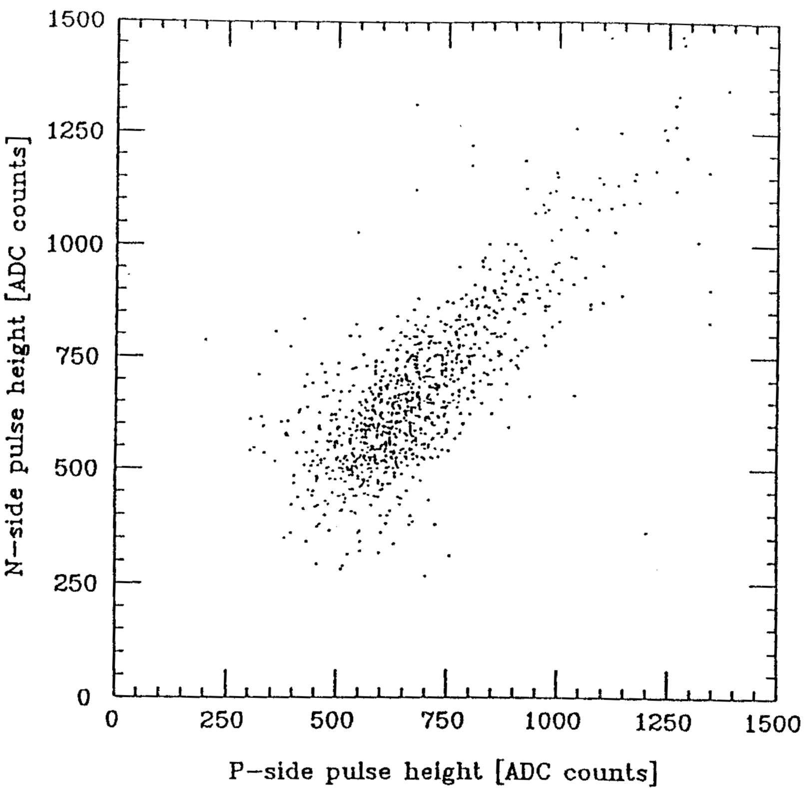


Figure 7. *Pulse height correlation between p- and n-side signals from a double-sided silicon strip detector for a 1 Gev/c π beam*

In order to minimize the material of the silicon detector, signals from z-strips have to be readout through the lines parallel to the ϕ-strip. Double-sided silicon devices with such a readout scheme will be made and their properties and possible problems like cross talk will be studied..

iii) Store output of a current amplifier in CCD, then readout the content when a trigger occurs. In this case, it is quite difficult to get fast output signal for trigger. The current problem is a poor efficiency to store charge from a silicon detector to CCD memory and is under investigation. Application for the B Factory detector requires further investigation.

In order to minimize the material of the silicon detector, signals from z-strips have to be readout through the lines parallel to the ϕ-strip. Double-sided silicon devices with such a readout scheme will be made and their properties and possible problems like cross talk will be studied..

3.3. The Precision Drift Chamber (PDC)

In order to keep good momentum resolution and good dE/dx measurement down to small angle, to detect low p_t tracks efficiently, to extrapolate tracks to the silicon detectors with a good precision, and to provide a fast trigger signal including a z trigger, the inner part of the tracker should have better position resolution and be more finely segmented than the outer part of it. Therefore we locate a precision drift chamber (PDC) just outside the silicon detectors. The PDC has position resolutions better than 100 μm in $r\phi$ and a few 100 μm in z and is located between radius of 8 cm and 20 cm as seen in Figure 8.

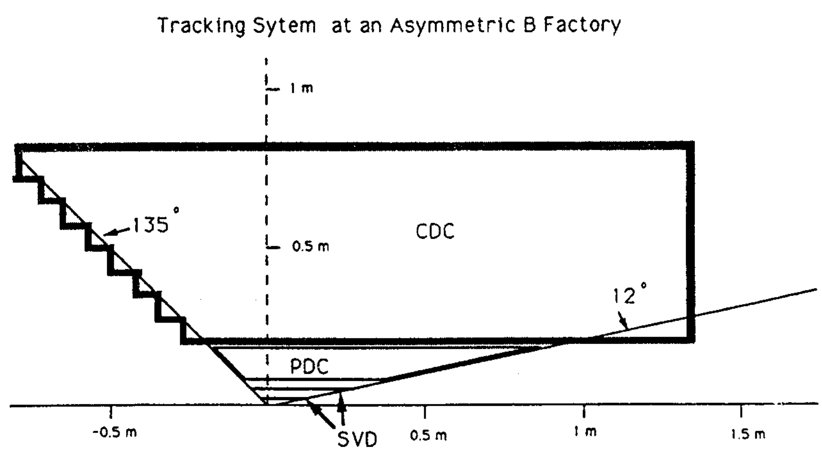


Figure 8. *Schematic view of the tracking detector for the KEK B Factory*

The fast measurements of $r\phi$ and z coordinates allow us to perform $r\phi$ and z trigger. The z trigger is needed to reduce the trigger rate by excluding the background contributions from the beam gas and beam wall events which might be significant at the B Factory because of the small radius of the beam pipe and the high beam current.

We are considering the following detectors as candidates of the PDC.

i) The small cell drift chamber or induction chamber for $r\phi$ measurements and cathode pads for z measurements.

For $r\phi$ measurements a small cell drift chamber or an induction chamber is used. The former is a conventional drift chamber with square shape drift cells which is the same type of chamber as used in the vertex chamber of CLEO II.[10]. The $r\phi$ resolution is about 100 μm and there is no essential problem in building it. The z coordinates are measured using stereo wires as well as cathode pads. The latter is a chamber which consists of alternating anode and potential wires in the symmetric plane with cathode planes on either side.[11] The $r\phi$ coordinates are measured using the ratio of the total charge of the anode and the difference of the induced charges on the potential wires to the left and right of the anode as seen in Figure 9.

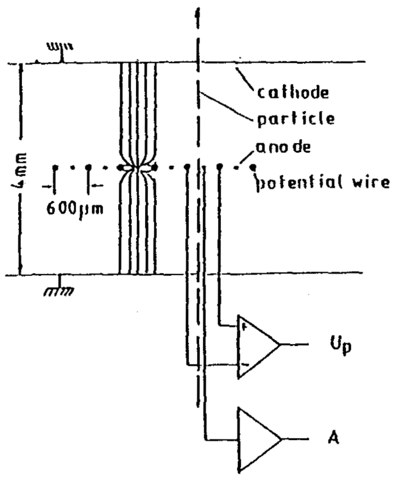


Figure 9. *Schematic view of the induction chamber*

The $r\phi$ resolution is 25 μm for normal incidence, but deteriorates for inclined incidence. This chamber was proposed as a PDC candidate by the PSI group.[12] This chamber has advantages such as high spatial resolution, high rate capability, good double hit resolution, short response time and simplicity of the readout as mentioned in ref.12. The z coordinates are measured using cathode pads. In order to obtain z trigger signals with reasonable rate we need to place at least 4 layers of cathode pads. Although the z resolution for the cathode pads is a few 100 μm for normal incidence, it deteriorates up to about 1 mm when particles enter the detector with large inclined angles. Detailed Monte Carlo study is needed to determine if such a worse z resolution for inclined incidence can still meet the requirements for the PDC. The cell size is chosen to be 8 mm and 10 layers are assembled in the space between r=8 cm and 20 cm.

ii) The ARGUS type micro vertex chamber.

The micro vertex drift chamber of ARGUS (μVDC)[13] provides not only good $r\phi$ resolution of a few tens of μm to 100 μm but also same order of z resolution by using $\pm 45^o$ stereo wires. The schematic view of the chamber is shown in Figure 10.

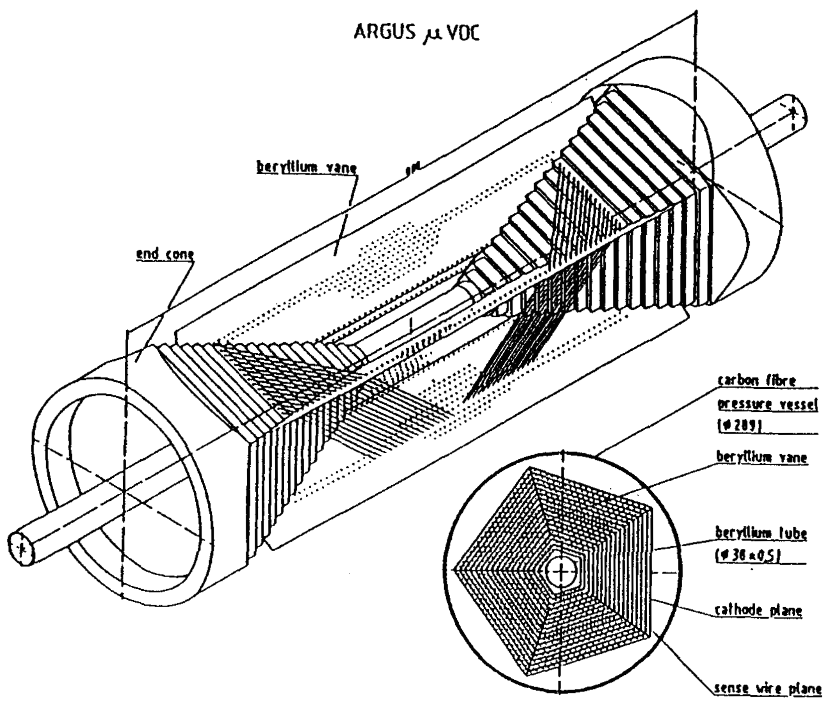


Figure 10. *Schematic view of the ARGUS microvertex drift chamber*

It has to be modified to be used at the B Factory in the following way as proposed by the PSI B Factory group.[12]

i) The gas pressure is changed from 4 atm to 1 atm to minimize the materials.

ii) The slow gas is replaced by the fast one of Ar/Ethane to provide fast signals.

The expected resolution under these conditions is about 100 μm. If we can string stereo wires with 90^o stereo angles we can measure z coordinates directly and obtain fast z signals to be used for fast z trigger as proposed by the PSI group.[12] However, it seems to be practically impossible to support such stereo wires in a reasonable way. We have heard in this workshop that the z trigger is included only in the second level trigger in the ARGUS experiment since there are so many combinations for defining z tracks that decision of trigger takes an order of 1 μs. Therefore we have to study whether it is still possible to obtain fast trigger signals by increasing the number of segmentations. In order to reconstruct tracks using the PDC data alone, the PDC should have at least 5 layers for each of the axial and stereo wires. The detector has a hexagonal shape in the $r\phi$ plane and the stereo wires are supported by the six Be vanes. Wire lengths increase together with the radial lengths as shown in Figure 8. The cell size is chosen to be 8 mm and 10 layers are assembled in the space between r=8 cm and 20 cm.

iii) Drift tubes with cathode pads for z measurements.

If we can produce drift tubes with polygonal shape using conductive plastic with good precision, we can apply cathode pads to z measurements for such drift tubes. Since the gas is not pressurized the expected $r\phi$ resolution is about 100 μm. One of advantages of drift tubes is that broken wires will not destroy a large portion of the detector. Various R&D projects are needed to study the feasibility of such a detector.

3.4. The Central Drift Chamber (CDC)

Just outside the PDC is placed the Central Drift Chamber (CDC). The major tasks of the CDC are shown below. i) Reconstruction of tracks. ii) Measurements of momentum with a resolution better than 0.5-1 % and dE/dx with a resolution better than 5-7 %. iii) Association of the track with the Čerenkov ring image of the RICH detector.

Radial length of 60 cm was chosen to obtain good momentum resolution and good dE/dx resolution, and at the same time to minimize the radius for the following two expensive detectors of RICH and CsI calorimeter. As a result the CDC occupies the radial space between 20 cm and 80 cm as seen in Figure 8. Since the average momentum of particles from B decay is relatively low, for example, 0.7 GeV/c for $\pi^{\pm}$, 1.1 GeV/c for $K^{\pm}$ and 2 GeV/c for leptons, multiple scattering plays an important role in determination of the momentum resolution.

Cell Structure

Jet cell structure is chosen to minimize uncertainty of momentum measurements due to multiple scattering in wires. The CDC consists of several cylinders which are arranged in super-layers of increasing length as shown in Figure 8. Each super-layer consists of several axial or stereo wire layers. It provides a local determination of the track vector, which enables quick estimation of multiplicity and momenta of particles for triggering the data acquisition system and facilitates the recognition of tracks in the off-line analysis. We are considering to choose the wire arrangement used for the central drift chamber of SLD [14]. We arrange eight super-layers (four axial, four stereo), each with six sense wire layers spaced at 5 mm, as shown in Figure 11.

The average drift distance is 2.5 cm. As a result we have 48 measurements by the CDC. Combining it with 10 measurements by the PDC, in total, 58 measurements can be used for determination of space coordinates and dE/dx.

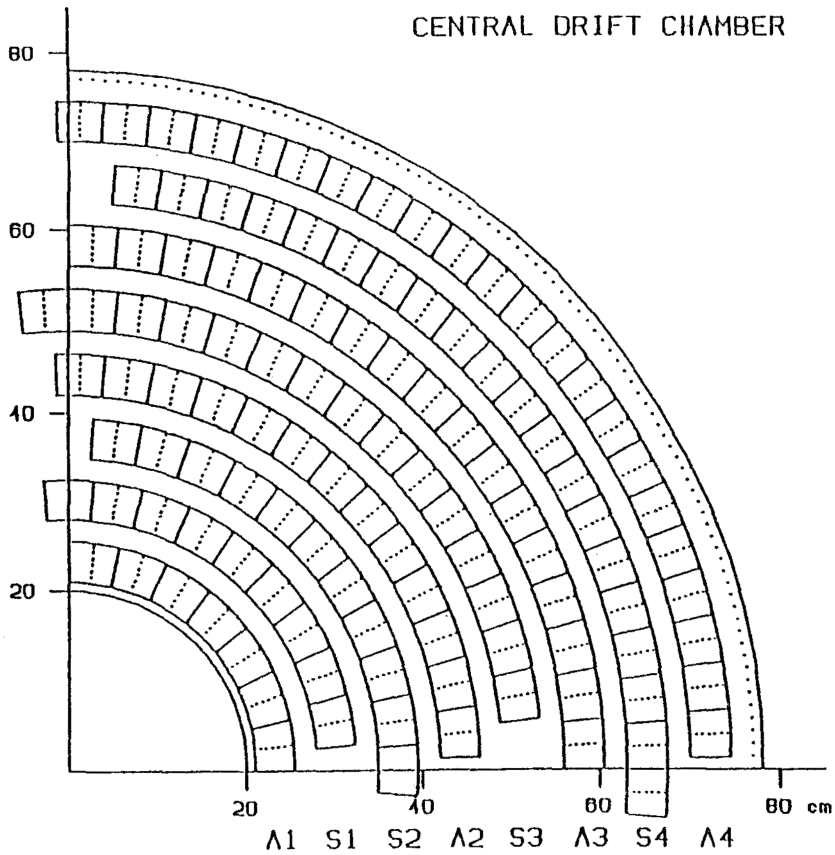


Figure 11. *Cell structure of the central drift chamber*

z measurements

z coordinates are measured using stereo wires with about 2 mm precision. In addition, direct z measurement with a 2 cm precision is also possible if we use the charge division method. But the charge division method will not be adopted since we want to place preamplifiers only in the backward z sides for reducing materials in the forward direction. In order to associate the CDC track with the RICH data we assemble cathode strips just outside the CDC for measuring z coordinates with a few 100 μm precision.

Gas

In order to minimize the uncertainty of momentum measurement due to multiple scattering a low density gas should be used. Presently considered gas is mixture of helium, CO_2 and isobutane, or mixture of helium and DME. The expected momentum resolution for such a design is

$$\frac{\delta p_t}{p_t} = \sqrt{(0.005p_t(GeV/c))^2 + 0.003^2}.$$

The second constant term is the contribution from multiple scattering in gas and wires. If we use Ar and CH_4 gas the term for multiple scattering increases up to 0.8 %.

3.5. The Ring Image Čerenkov Detector (RICH)

The B meson decays into a D meson with a high branching ratio and the D meson decays into a K meson, via following decay chain;

$$\begin{aligned} B^0(\bar{B}^0) \rightarrow & \bar{D}(D) + X, \\ & \bar{D}(D) \rightarrow K^+(K^-) + X. \end{aligned}$$

As a result, the charge of the K meson can be used for identification of the mother B meson. Major source of backgrounds in this method is K mesons originated from light hadrons. However, this contribution is found to be small based on a study using a Monte Carlo simulation.[15,16]

According to the simulation we can identify B flavor with a purity of more than 95% and obtain efficiency of about 40 % for B flavor identification if we can identify the kaon with 100 % efficiency below 3 GeV/c. This value is about 3 times bigger than that obtained from the lepton tagging. Therefore, it is very important for a CP violation measurement to accommodate the best possible detector (or detectors) that can identify K mesons. The conventional methods of particle identification using combination of Time-of-Flight (TOF) and momentum measurements or combination of dE/dx and momentum measurements allow separation of pion and kaon in the momentum range below 1 GeV/c and above 3 GeV/c. The installation of a Ring Image Čerenkov counter (RICH), or equivalently powerful device is needed for the kaon identification between 1 GeV/c and 3 GeV/c. RICH detectors are placed just outside the CDC both in the barrel and endcap regions as seen in Figure 3.

The RICH uses the newly developed technology[17] as described below. The velocity of a particle can be estimated accurately using precise measurements of the radius of the Čerenkov ring image in the photon detector which is located a few tens of cm away from the radiator. We are considering the following different types of RICH detector.

CRID(SLD)/RICH(DELPHI) type

The RICH detectors which were constructed for DELPHI[18] and SLD[19] are using both the liquid and gas radiators along with the high precision mirror to reflect Čerenkov light from the gas into the photon detector to identify particles up to 30 GeV/c. TPC is used for detection of photo-electrons so that the response of the detector is quite slow. This type of detector has the following advantages. i) It has been fully tested. ii) The total number of readout channels is relatively small. iii) It is not so expensive. On the other hand it has the following disadvantages. i) It has a mechanically complicated structure. ii) It demands careful control of temperature of the photoionizing gas of TMAE around 28°C. iii) TPC has a long drift distance so that purity of the gas has to be maintained to be constant. iv) TPC can not be used for the endcap RICH due to the axial magnetic field.

PSI type fast RICH

We can design a simpler and easily maintained RICH for a *B* Factory because particle identification only up to 3 GeV/c is needed for this machine. We follow the design shown in the *B* Factory proposal of PSI [12]. It will be possible to have a required capability of particle separation according to the results of the Monte Carlo simulation by the PSI group. The major modifications from the RICH's of DELPHI and SLD are listed below:

i) A crystal radiator (NaF) is used.

ii) A MWPC with Tri-Ethyl-Amine (TEA) is used for a photon detection.

These modifications lead to the following advantages. i) A mechanical structure is simple because the radiator is solid. A quartz window is not needed to separate a radiator from air so that a light loss is smaller. ii) TEA works in room temperature and no need of careful temperature control. iii) The detector can be made thin (20 - 30 cm). iv) A MWPC works in the endcap region (i.e., not affected by a magnetic field). This type of detector has the following disadvantages. i) The total number of readout channels is very large. ii)It is very expensive due to use of NaF as well as many readout channels and costs more than $ 15 M. iii) Because of new type of detector extensive R&D is required.

Multi-Anode Photomultiplier

It might be possible to use multi-anode photomultipliers (MAPT) for detection of Čerenkov lights. In this case liquid can be used for the radiator. The MAPT is a photomultiplier equipped with many anodes. If the anode segmentation is 1 cm^2 we can determine the radius of a Čerenkov ring with enough precision to be able to identify kaon of 3 GeV/c with 90 % purity.

This detector has the following many advantages. i) Use of photomultipliers is a reliable technique. ii) Efficiency for light collection is high. iii) Fast trigger signal is obtained. iv) Chromatic aberration is small. The disadvantages of this detector are as follows. i) The MAPT is very expensive and costs at least $ 100 per cm^2. The total area where the endcap RICH covers is about 35,000 cm^2. So if we assume that the size of one anode is cm^2 then the total cost only for the endcap RICH amounts at least $ 3.5 M. ii) The materials of MAPT amount about 6% of radiation length.

Necessary R&D

In order to detect Čerenkov rings unambiguously the photon window of a MAPT has to have a polygonal shape and the area of photocathode has to be at least 60 % of the total area of the photon window. The presently available 3 inch ϕ MAPT of square shape just satisfies the above requirements. However, it is desirable to increase the portion of the photocathode by either reducing the thickness of the glass wall or producing a larger size MAPT without increasing the thickness of the glass wall. Also required is to reduce the price of the MAPT. At present we plan to choose type 2 or 3 for the endcap RICH and type 1 for the barrel RICH. Since the expected number of kaons from $\Upsilon(4S)$ decay in the endcap region is about 4 times as many as in the barrel region, major effort will be made for the R&D of the endcap RICH.

3.6. The Time-of-Flight counter (TOF)

Time-of-flight (TOF) is a conventional, however very powerful, method of the particle identification at low energies where most of particles from the Υ decay lie. TOF counters are placed between the RICH and the electromagnetic calorimeter, at r=1.2 m for the barrel part and z=1.7 m for the endcap part as shown in Figure 3. Monte Carlo simulations for $\Upsilon(4S) \rightarrow B\overline{B}$ events were performed for a TOF system with the time resolution of 100 ps and 50 ps. The results indicate that about 89% of all K reaches the TOF. Also indicated is that the purity in the K candidate when they are selected with 90% efficiency is 88% below 2 GeV/c for the 100 ps resolution and 92 % for the 50 ps resolution below 3 GeV/c.

At present, one can build small size (few cm^2) plastic scintillation counters with a better than 100 ps timing resolution. However, general purpose detectors require large area counters, typically 2 to 3 m long, and the achievable resolution is 150 to 200 ps. This worsening is due to variation of light propagation path length along the counter and variation of originating position of the light along the depth of the counter. One possible solution which circumvents the above limitation is to construct a segmented plastic scintillation counter with independent readout at different depth. Plastic scintillating fibers are suitable for construction of such segmented counter. The timing jitter of the read-out photomultipliers can be reduced by using multichannel plates (M*CP*). The total transit time of multichannel based photomultipliers is of the order of 300 ps with a jitter of less than 30 ps. Preliminary results of Monte Carlo simulation of a counter with M*CP* read-out in ref. 20 indicate that 50 ps timing resolution can be achieved for 3m long and 10 cm thick counters.

3.7. The Electromagnetic Calorimeter (EMCAL)

An electromagnetic calorimeter plays an important role in a B Factory detector in the following three aspects:

i) Detect energy and position of photons with good resolution and high efficiency for reconstructing the B meson.

ii) Detect electrons with high efficiency and low background for tagging the B meson.

iii) Contribute to the trigger especially for events in which a large fraction of the energy is carried away by photons.

Among various types of electromagnetic shower counters, the inorganic scintillation crystal, such as NaI or CsI, produces large enough light yield so that the light detection can be done by the photodiode rather than with the high gain photomultiplier. This makes these crystals an ideal electromagnetic shower counter for a B Factory detector because in order to detect low energy photons efficiently, the detector must be located inside the magnet coil. Comparing CsI with NaI, light signal from CsI has better spectral matching to the silicon photodiode. Handling of CsI crystals is much easier because this material is only weakly hygroscopic. One of disadvantages of CsI is its long decay time of scintillation (850 ns, about 10 times longer than the plastic scintillator). This makes it difficult to use the output signal of CsI as a part of the fast trigger. However, it is possible to extract the σ_t=50 ns signal by using an amplifier with a fast shaping time of 250 ns.[12] Another disadvantage of CsI is an inferior radiation resistance as compared to BaF_2 or Lead-Scintillator sandwich counter (Pb-SC).

The barrel part of the calorimeter is placed at r=125-155 cm and the endcap part is placed at z=240-270 cm.

CsI has about 5 times better energy resolution than Pb-Sc. Since the energy spectrum of photons from B decay is peaked at around 70 MeV for 2.3×12 GeV collision a calorimeter is required to have good resolution and detection efficiency down to low energy region. Energy resolution (σ/E) of 6% $\sim$ 9% is expected at 50 MeV for CsI. In case of Pb-Sc, 50MeV corresponds to the minimum detectable energy, while CsI can detect down to 30 MeV. For a CsI calorimeter one can use

$B^0 \to \psi K_S^0(\to \pi^+\pi^-)\pi^0$ [21] and $B^0 \to \psi K_S^0(\to \pi^0\pi^0)$ in addition to the basic decay channel of $B^0 \to \psi K_S^0(\to \pi^+\pi^-)$ for CP measurements. A Monte Carlo simulation shows that the yield of CP events can be increased by 50 to 80 % corresponding to the 50% to 100% probability of helicity zero state of ψK^* as compared to the case where only the basic channel is used. But if we use a Pb-Sc detector such improvement is not obtained due to its worse resolution.

3.8. The Muon Detector

Muons do not interact strongly and this characteristic is used for identification of them. The muon detection system usually consists of muon chambers for position measurement and muon filter (Fe) for hadron reduction and is located outside of the electromagnetic calorimeter. Most of charged π and K cannot penetrate the muon filter and only μ can penetrate. The main background in the muon detection comes from punch-through and decay of charged hadrons. When we increase the thickness of muon filter (Fe), background from charged π can be reduced. But an efficiency for detecting the low momentum muons from the primary B decay is decreased. To solve this problem, one can segment the muon filter (Fe) into several layers and insert muon chambers in between as seen in Figure 3. For this study, we used a simulation program for hadron and electromagnetic shower, GHEISHA.[22] The punch-through probability of hadrons was evaluated when particles incident the following simplified muon detector normally. The simplified muon detector consists of a 57 cm iron shield followed by the six layers of 5 cm gap plus 15 cm iron shield and it is placed 100 cm away from the interaction point. The resulting punch-through probability of hadrons is found to be 0.3 to 0.9 % depending on the muon momentum when detection efficiency of muons is required to be 96 %.

Leptons from the the semileptonic decay of B meson are used for tagging the B meson in CP and B_s mixing experiments. When the mother B meson includes a b quark, it decays into l^- and when the mother B meson includes a $\bar{b}$ quark, it decays into l^+. We studied the tagging efficiency of this method in the process

$$B^0\bar{B}^0 \to (\psi K_S^0) + X \to (l^+l^- + \pi^+\pi^-) + X.$$

We also studied how the vertex information and the muon filter segmentation affect reduction of the background in the lepton from the primary B meson decay. The main physics background in the lepton event in which the lepton originates from $b \to l$ comes from

$$B \to D \to l, \text{ and}$$
$$B \to \tau \to l.$$

The lepton momentum for such processes is much lower than that for the primary B decay. In the tagging method using the lepton, $B \to D, \tau$ and the decay of K and π to μ become main background. In the former case, a primary B decay vertex should be present upstream of the vertex which contains a lepton. If the existence of the upstream B vertex can be found, this type of background can be rejected. It is the same for the latter case because there are many charg ed particles from the secondary decay vertices. z If the difference, Δz, z for the lepton minus z for the charged particles, is positive, the existence of second decay vertex upstream of a vertex containing a lepton candidate can be confirmed.

A Monte Carlo simulation was performed for 12 $\times$ 2.3 GeV beam energies using the punch-through probability calculated above and the vertex information. As a result we obtained an lepton detection efficiency of 64 % and a purity of $B \to l$ of 94 % when we require that p_μ>0.8GeV/c, p_e >0.5 GeV/c, $p^*{}_l$>1.2 GeV/c, and Δz < 150μm for at least two charged particles with momentum larger than 0.25 GeV/c. Herep_μ and p_e are the momenta in the laboratory system, of the muon and electron candidates, respectively and $p^*{}_l$ is the lepton momentum in the Υ(4S) rest system. The electron identification probability is assumed to be 100 % for $p_e > 0.5$ GeV/c.

3.9. Solenoidal Magnet and Compensation Magnet

There are two options for a magnet, a solenoidal magnet and a dipole magnet. We compare features of the two options below.

Solenoidal magnet

A solenoidal magnet has the following advantages. i) Small dead space. Only 2 % dead space for the polar angular acceptance of $\theta > 12^o$. ii) Wires are strung parallel to the beam axis so that they can be supported by endplates. It helps to reduce materials in the barrel region. iii) Symmetry with respect to the beam axis is kept. iv) Many detector analysis programs are already existing.

The disadvantages of the solenoidal magnet are listed below. i) Additional magnets are needed for separating beams at the interaction region. ii) Compensation magnets might be required and they reduce the available volume for the detector.

Dipole magnet

A dipole magnet has the following advantages. i) It can be used as a beam separator. ii) No compensation magnet is needed.

The disadvantages of the dipole magnet are listed below. i) Wires of tracking chambers have to be strung perpendicular to the beam axis so that the wires have to be supported by a barrel outer cylinder. As a result we expect a large amount of materials in front of the barrel electromagnetic calorimeter. ii) Large dead space, about 15 %. iii) Developments of various detector analysis programs are required.

As a result of the above comparison we prefer a solenoidal magnet to a dipole one.

The higher magnetic field provides better momentum resolution while low p_t charged tracks do not reach the detector. The optimum field strength is between 1 and 1.5 Tesla and at present we design the detector for 1 Tesla field. In the detector design given in Figure 3 we assumed to use the VENUS solenoid and modify the return yoke to raise the field strength from the current value of 0.75 Tesla up to 1 Tesla.

Compensation Magnet

Beams are separated by using two sets of magnet configuration each of which consists of a separation bend, a quadrupole and a septum magnet as seen in Figure 2. Since beams pass through these magnets off the axis the affects of the field of the solenoidal magnet on the beams have to be minimized. At the moment we assume that the solenoid field is required to be less than .1 T around the beam separation magnets. In order to achieve such a field we need to place a compensation magnet between the separation magnets and the solenoidal magnet. The above requirement is satisfied if a compensation magnet of 1 M A-turn is placed between the PDC and the return yoke (z = 80 - 200 cm). The solenoid field can be reduced to less than a few hundred Gauss around the beam line at z=80-200 cm. If we have to place such a compensation magnet of about 10 cm radial thickness, the very forward part of the detector configuration shown in Figure 3 has to be removed so that angular acceptance of the detector is reduced. Therefore we are going to study the effects of the solenoid field on the beams in order to check whether the effects are really so strong that a compensation magnet is indispensable or not.

3.10. Trigger and Data Acquisition System

The event rate and the beam repetition rate are very high in the proposed B Factory. Thus, the requirements for the read-out electronics and data-taking system are much different from those at traditional e^+e^- experiments. We consider these requirements for the maximum luminosity of $10^{34}/\text{cm}^2/\text{s}$ since it is the most severe case for them. In this case the beam repetition period will be 1.8 ns and the beam total current will be about 1 A. The expected event rate is about 200 Hz which consists of 75 Hz for physics events, 100 Hz for the cosmic ray background, and 20-30 Hz for backgrounds associated with beams. Here Bhabha events are assumed to be scaled down by factor 100. The data size of one event is assumed to be about 30 KB. As a result the data-acquisition system is required to have overall throughput larger than 15 MB/sec.

The architecture of the high speed electronics and the data acquisition system has been studied at HERA, SSC and other B Factory experiments, and the general scheme has been almost established. Trigger signals generally require about $1\mu s$ to be generated, because they have to wait for signals from the drift chambers. Since the period of the beam crossing at the B Factory is much smaller than this time, the traditional gate-and-clear method can not be applied. The analog signals from the detectors need to be stored until the trigger is generated. For this purpose a fast analog memory (10 ns for read/write) is needed. The analog memory consists of multiple sample-and-hold circuits and switched capacitors, which are formed on a very large scale integrated circuit chip (VLSI).[23] We are also considering another way to store analog signals by using a Surface Acoustic Wave (SAW) delay line. The delay line is made of a single crystal where the surface is manufactured with super accuracy along the lattice of the crystal. A super acoustic wave is induced on the surface by the Piezo effect and it travels along the surface with 3.9 mm/μs speed . When a trigger is generated, only the signals stored in the memory or in the delay pipe line which correspond to the triggered events are digitized and stored again in the digital buffer memory. At this stage it is possible to apply the second level trigger or event selection by delaying those data by digital shift registers. By this method, signal-to-noise ratio will be significantly improved and the load of the data taking system will be reduced. However, the second level trigger requires the data of all the electronics channels to be kept in shift registers of the next step. This makes the system very complicated and the cost will be large. Therefore, we propose to improve the first level analog trigger and to gain the time needed for that by making the capacity of the analog memory larger.

Signals of the TOF counters can be used to enable matching between stored signals and the triggered events as shown in Figure 12.

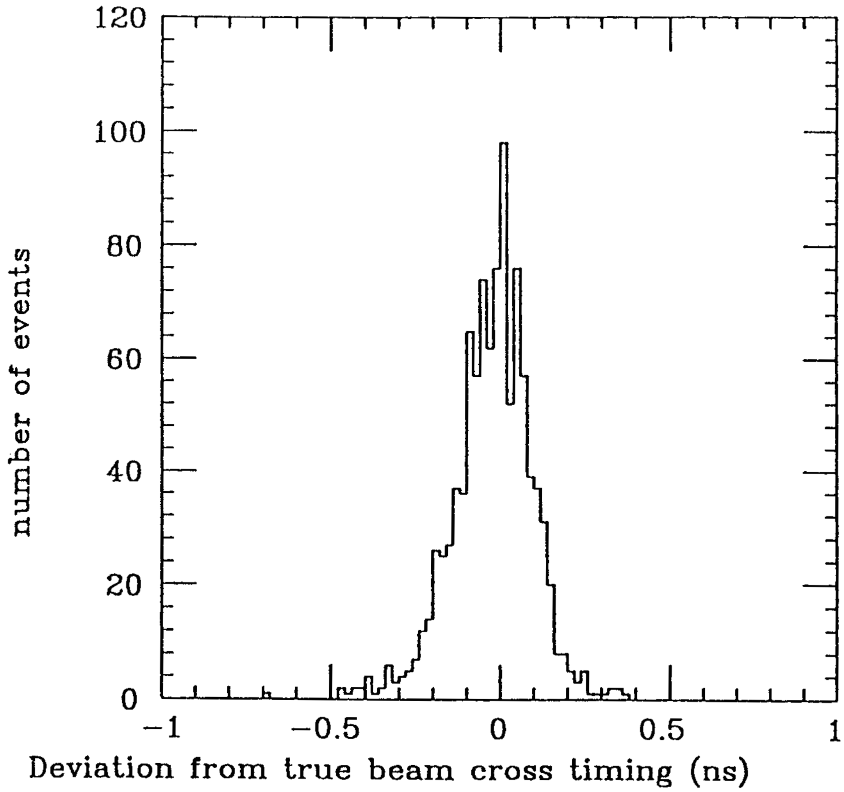


Figure 12. *Monte Carlo simulation of the time difference between the known clock and TOF signals for 100 ps resolution.*

In order to get rid of a pile-up problem the level 1 trigger is needed to be fast (100 ns) and of low rate less than 10 kHz. Such a fast trigger can be obtained by using PDC, CsI and/or SVD signals.

The amount of the data stored in the buffer will be large as mentioned before, and it is not realistic to transfer these data through a single bus. The most natural solution to this is the parallel data transfer on a detector-by-detector basis. The event builder converts these data to parallel data sets on an event-by-event basis. It is possible to write the data built here on a mass storage device through the online computer, but the mass storage device and the data analysis will be big problems. To solve these two problems, an idea of online-computer-farm plays an important role. This analyzes the data to the level of data-summary-tapes on real time, and outputs only background-free summarized data to the mass storage device. By employing this idea, the data can be transferred to the large computer directly, which was proven by the TRISTAN experiments to be a very effective method of data handling. The *CPU* power needed at an experiment of *B* Factory is about 3000 Mips. This *CPU* power can be achieved by using 200 chips of 15Mips microprocessor which are already commercially available. Transfer of digitized data is carried out using the VME bus due to the following advantages. Microprocessors can read data of any other modules without overhead for initialization so that data transfer with high speed of 4MB/sec is feasible. Another advantage is that many kinds of microprocessor boards, memory boards and program development environment are commercially available at low prices.

4. Summary

We have discussed the design of the detector for an asymmetric *B* Factory at KEK after the brief description of the design of the accelerator. All of the detector elements except for the silicon strip detector and RICH can be constructed using well established techniques. We have presented the R&D effforts which have been performed and planned for the double-sided silicon strip detector. All the detector elements cover the angular range between 12° and 135° which is optimized for the 2.3×12 GeV collision. We have shown several new ideas such as use of multi-anode photomultipliers for the RICH detector and use of surface acoustic wave delay lines for storing analog signals in the data acquisition. We have also noticed that it is very important to check if we need to place compensation magnets to reduce the effects of the beam separation magnets on the beams since they reduce the angular acceptance of the detector in the forward direction significantly.

References

1. Task force report on asymmetric *B* Factory at KEK, Feb. 28, 1990.
2. Y. Funakoshi, *et al.*, Contributed paper to the Workshop on Beam Dynamics Issues of High-Luminosity Asymmetric Collider Rings, LBL, Feb. 1990, and KEK Preprint 90-30 (1990).
3. T. Kamitani, reference 1, appendix C1, p. 145.
4. P.Holl, *et al.*, *IEEE Trans. Nucl. Sci.* **36**, 251 (1989); G Batignani, *et al.*, ibid **36**, 40 (1989).
5. H. Tajima, K. Niwa and M. Nakamura, reference 1, appendix C5, p.219.
6. J.C. Stanton, *IEEE Trans. Nucl. Sci.* **36**, 522 (1989).
7. H. Becker, *et al.*, *IEEE Trans. Nucl. Sci.* **36**, 246 (1989).
8. C. Haber,*Proceedings of the Seventh Topical Conference on Proton-Antiproton Collider Physics*, Fermilab (1988).
9. Ikeda, *et al.*, KEK Report A277, 160 (1989).
10. E. Blucher, *et al.*, *Nucl. Instr. Meth.* **A249**, 201 (1986).
11. E. Roderburg, *et al.*, *Nucl. Instr. Meth.* **A252**, 285 (1986).
12. Proposal for an Electron-Positron Collider for Heavy Flavour Physics and Synchrotron Radiation, **PR-88-09** (1988).
13. E. Michel, *et al.*, *Nucl. Inst. Meth.* **A283**, 544 (1989).
14. W. B. Atwood, *et al.*, *Nucl. Instr. and Meth.* **A252**, 295 (1986).
15. R. Aleksan, *et al.*, *Phys. Rev.* **D39**, 1283 (1989).
16. H. Ozaki, *et al.*, reference 1, appendix C3, p. 185.
17. D. Aston, *et al.*, SLAC-PUB-4795; R. Arnold, *et al.*, Nucl. Instr. Meth. **A270** 255 (1987).
18. R. Arnold, *et al.*, CERN/HE 87-01 and CERN/HE 87-08.
19. SLD Design Report SLAC-273, UC-34D (May 1984).

20. R. Stroynowski, *Proceedings of the Workshop on Scintillating Fiber Detector Development for the SSC*, Fermilab (1988), p. 313.

21. B. Kayser, *et al.*, UCLA/89/TEP/60.

22. H. Fesefeldt, *et al.*, PITHA 85/02 (1985).

23. H. Ikeda, Contributed paper to the 8th Hadron Collider Workshop, KEK (1989).

Selected $\mathrm{Ba\overline{B}ar}$ Notes

Note #	Author(s)	Title	Date
24	F. Porter	Dilution Factor in CP Violation Measurement	Jan. 25, 1990
25	A. Snyder, S. Wagner	An Alternate Method of B-Tagging for CP Violation Studies	Feb. 5, 1990
26b	F. Porter	Dilution Factor in CP Violation Measurement	Feb. 9, 1990
31	F. Porter	Luminosity requirement for CP violation measurement in presence of background; Comparison of two CP violation sensitivity estimates.	Mar 22, 1990
32	A. M. Eisner	Bhabhas at an Asymmetric B Factory	Apr 13, 1990
33	F. LeDiberder	Background Study of the $B \to \pi\pi$ Channel	May 1, 1990
34	F. LeDiberder	Precision on CP Violation Measurements and Requirement on the Vertex Resolution	May 1, 1990
35	W. Toki	Parametrization of Decay Length Distributions with Errors	May 1, 1990
38	I. Dunietz, A. Snyder	Additional Bd Decays with Large CP Violation and No Final State Phase Ambiguities	May 8, 1990
39	G. Wormser	The Forward Hybrid Microvertex Detector	May 10, 1990
40	J. Dorfan, S. Komamiya, A. Snyder	Filling in the Detector Variation Matrix	May 9, 1990
41	W. Toki	CP Measurements from Forward-Backward Asymmetry of Decay Lengths	May 17, 1990
42	F. LeDiberder	CP Violation as Seen from the Δz Distribution	May 30, 1990
43	W. Toki	Double Moments, CP Measurements and "CP Moments"	Jun 1, 1990
45	F. LeDiberder, W. Toki, M. Witherell	The Effect of Beam Pipe Radius on the Measurement of CP Asymmetries	Jun 9, 1990
48	V. Luth	Spatial Resolution of Silicon Detectors: a Monte Carlo Study	Jul 19, 1990
49	F. LeDiberder	Precision on the α_S Measurement	Aug 13, 1990
50	F. LeDiberder	On the Optimal High Energy Beam	Aug 13, 1990
52	W. Toki	Estimated CP error in the decay $B \to \psi K^*$	Aug 29, 1990

53	W. Toki	$B \to D^{*}+ D^{*}$ - CP moments	Aug 29, 1990
54	T. Browder, M. Witherell	Limits on Backgrounds for the B Factory Detector	Sep 7, 1990
55	V. Luth	Design of a Vertex Silicon Detector	Sep 7, 1990
56	S. Komamiya	Monte Carlo Study of $B \to \psi + K_S$	Sep 27, 1990
57	S. Shapiro	Silicon pin diode array hybrids as building blocks for a Vertex Detector at an Asymmetric B Factory.	Sep 27, 1990
58	A. Fridman, A. Snyder	Remarks on B Physics with Interactions of pp and $e^+ e^-$	Sep 28, 1990
59	A. Boyarski, T. Glanzman, F. Harris, F.Porter	Computing for a B Factory	Oct 2, 1990
60	J. Va'vra	B Factory forward-backward Vertex Chamber based on the Microstrip gas etc. avalanche	Oct 10, 1990
61	A. Snyder, S. Wagner	Reconstruction, Vertexing and Backgrounds for $B^0 \to D^+D^-$	Dec 10, 1990
63	T. Bolton, et.al.	Triggering in an Asymmetric B Factory	Jan 30, 1991
65	T. Browder, M. Witherell	Limits on Backgrounds for the B Factory Detectors	Mar 1, 1991

☆U.S. GOVERNMENT PRINTING OFFICE: 1991-587-458/59234