Evidence for $B \to K\eta'\gamma$ decays at the Belle detector

Robin Wedd

THE UNIVERSITY OF
MELBOURNE

Submitted in fulfilment of the requirements
of the degree of Doctor of Philosophy

The School of Physics
The University of Melbourne

September 2008

Abstract

This thesis presents an investigation of B meson decay via the process $b \to s\gamma$. A search for the highly suppressed decay $B \to K\eta'\gamma$ is performed using data recorded at the $\Upsilon(4S)$ resonance by the Belle detector at the KEKB accelerator facility in Tsukuba, Japan. The data sample was obtained from an integrated luminosity of 605.44 fb^{-1} and contains 657 million $B\overline{B}$ pairs.

A measurement of the branching fraction of the flavour-changing neutral current process $B \to K\eta'\gamma$ will provide a highly sensitive probe of unknown beyond Standard Model physics. The virtual-loop induced interaction can gain contributions from any undiscovered heavy particles that couple to Standard Model quarks. Any discrepency between the measured value and the prediction of the Standard Model could be an indication of new physical phenomena.

We report the first evidence of the decay $B^+ \to K^+\eta'\gamma$ with a significance of 3.29 standard deviations and a branching fraction of

$$\mathcal{BF}(B^+ \to K^+\eta'\gamma) = (3.42^{+1.23+0.34}_{-1.13-0.40}) \times 10^{-6}$$

where the first error is statistical and the second is systematic. No significant evidence of the decay $B^0 \to K^0\eta'\gamma$ is found. An upper limit at the 90% confidence level for the decay is calculated to be

$$\mathcal{BF}(B^0 \to K^0\eta'\gamma) < 6.0 \times 10^{-6}$$

This is to certify that

i. this thesis comprises only my original work towards my PhD,

ii. due acknowledgement has been made of all other material used,

iii. this thesis is less than 100,000 words in length, exclusive of tables, bibliographies and appendices.

Robin Wedd

Acknowledgements

This theis is first and foremost a product of the love and support I have received from my parents Sue and Tony Wedd. Without their unfailing assistance and encouragement, even when it seemed I was losing the way, I would never have reached this point in my research career. Thanks guys, this thesis is dedicated to you.

Contents

List of Figures

List of Tables

Chapter 1

Theoretical Framework and Motivation

1.1 Introduction

This dissertation describes a measurement of B meson decays to the final state $K\eta'\gamma$ undertaken at the Belle detector in Tsukuba, Japan. The measurement utilises data recorded by the Belle detector over the period 1999 to 2006. Chapter 2 describes the Belle detector itself, the KEKB accelerator which provides the B mesons, and the process of data collection and simulation. Chapter 3 details the selection process by which B mesons decaying to $K\eta'\gamma$ are separated from other B meson decays and all backgrounds. Measuring the number of $B \to K\eta'\gamma$ decays that remain in the data after selection is the subject of Chapter 4. The estimation of the selection efficiency, the uncertainty due to detector and modelling inaccuracies, and the final calculation of the frequency of the $B \to K\eta'\gamma$ process is shown in Chapter 5.

The following chapter provides the theoretical grounding and motivation for the measurement. The Standard Model of particle physics is presented, focusing on flavour physics in the weak sector. The phenomenology of flavour-changing neutral currents (FCNC) and the $B \to K\eta'\gamma$ process is described.

1.2 The Standard Model

The current knowledge of fundamental particle physics is contained in a single theory called the Standard Model (SM). According to the SM the universe consists of spin-$\frac{1}{2}$ fermions that communicate via the exchange of spin-1 bosons. Three types of fermion-fermion interactions are described in the SM by the use of a gauge field theory based on the symmetry group

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{1.1}$$

where $SU(3)_C$ is the special unitary group of 3×3 matrices that describe interaction via the strong force, $SU(2)_L$ is the special unitary group of 2×2 matrices that describe weak interaction, and $U(1)_Y$ is the

Leptons				Quarks			
Flavour		Mass (GeV/c^2)	Charge	Flavour		Mass (GeV/c^2)	Charge
electron neutrino	ν_e	$(0-0.13)\times10^{-9}$	0	down	d	0.0035-0.006	-1/3
electron	e	0.000511	-1	up	u	0.0015-0.0033	2/3
muon neutrino	ν_μ	$(0.009-0.13)\times10^{-9}$	0	strange	s	0.104	-1/3
muon	μ	0.1057	-1	charm	c	1.27	2/3
tau neutrino	ν_τ	$(0.04-0.14)\times10^{-9}$	0	bottom	b	4.3	-1/3
tau	τ	1.7768	-1	top	t	171	2/3

Table 1.1: Standard Model fermions. 'Charge' denotes the electromagnetic charge.

group of all complex numbers of absolute value 1 and describes the electro-magnetic (EM) interaction [1]. Gravity is not included in the SM; a quantum theory of gravity is yet to be discovered.

Fermions can be divided into two groups; those that feel the strong force (quarks) and those that do not (leptons). Quarks fractional EM charge while leptons have either integer or zero EM charge. There are three generations of leptons and quarks, with each generation comprising two quarks and two leptons. The six quark and six lepton 'flavours' are shown in Table 1.1. The different generations of fermions couple to the spin-1 bosons in identical fashion; the generations are differentiated only by mass and their individual flavour quantum numbers (strangeness etc.). As can be seen in Table 1.1 the fermion mass range covers many orders of magnitude.

The SM gauge bosons and the interactions in which they are exchanged are listed in Table 1.2. All fermions have anti-particle partners which have opposite EM charge. The neutrino anti-partners

The strong force is mediated by eight spin-1 gluons. The strong force charge is called 'colour' (denoted by the subscript 'C' in equation 1.1) and the theory describing the strong force is thus referred to as 'quantum chromo-dynamics', or QCD. Colour charge comes in three versions- red, green and blue- in analogy to the primary colours of the visual spectrum. Quarks can have any one of these charges; anti-quarks can have any one of anti-red, anti-green or anti-blue. Gluons carry a super-position of colour/anti-colour states. Quarks never appear in isolation due to the colour confinement mechanism, which causes the force between quarks to increase with distance until the quark-antiquark production threshold is reached. This due to the self-interaction of the gluons. Quarks must form colour-neutral states to be stable, either quark-antiquark pairs with opposite colour (mesons) or triplets with one of each of the three colours or anti-colours (baryons).

The weak and EM forces have been successfully unified at high energies and are described by the combined $SU(2)_L \otimes U(1)_Y$ symmetry group, where the subscripts 'L' and 'Y' refer to the fact that the weak interaction couples only to left-handed particles (and right-handed anti-particles) and particle hyper-charge, respectively. The Higgs mechanism provides a means through which the symmetry of the electro-weak sector is spontaneously broken by a non-zero vacuum expectation value. This reduces the symmetry of the electro-weak sector to that of $U(1)_{EM}$. In the process the fermions as well as the W and Z bosons are given mass, while the photon is left massless. The massless nature of the photon enables the EM force

to act over longer ranges than the weak force, which is short-ranged due the large masses of the W and Z bosons. All fermions with electric charge interact electro-magnetically via photon exchange. All fermions (anti-fermions) with left-handed (right-handed) chirality interact via the weak force. The Higgs mechanism results in an additional boson also named Higgs; it is the only particle in the SM that is yet to be discovered. Table 1.2 gives the lower mass bound provided by the search at the LEP collider at CERN [2].

The Standard Model is one of the most successful endeavours of modern physics. Its elegant symmetry-based construction has been shown to be highly accurate in explaining and predicting experimental results. Several missing elements and unexplained phenomena indicate that it is not a complete theory of the physical universe however. The particle masses are among 18 free parameters which have no explanation within the SM and must be measured experimentally. A complete theory will need to give some reason for the observed mass hierarchy. The Higgs mechanism provides this theoretically but has yet to be confirmed.

If the Higgs boson is discovered then a problem in the formulation of the Higgs mechanism will also need to be solved. The Higgs boson mass receives corrections from loop interactions mediated by every particle with which it couples. Any theory of quantum gravity requires very heavy particles with mass of the order of the Planck energy scale: the coupling of the Higgs to these particles increases the Higgs mass by many orders of magnitude, making it incompatible with the SM. Solving this divergence requires either enormously precise fine-tuning to cancel the contributing corrections or a new theory of particle physics. Two theories that address this problem are super-symmetry and models with extra spacial dimensions.

Another phenomenon requiring explanation is the present matter-dominated state of the universe. At its inception the universe is believed to have been matter-antimatter symmetric. Sakharov defined three necessary conditions to explain the evolution from the symmetric to the current asymmetric state, now known as the Sakharov Conditions [3]:

- **Baryon number violation.** Baryon number (B) is defined as $B = \frac{N_q - N_{\overline{q}}}{3}$, where N_q and $N_{\overline{q}}$ are the number of quarks and anti-quarks, respectively.

- **A lack of thermal equilibrium.** In which the rate of expansion of the universe is greater than the rate of the baryon number violating interactions.

Type	Mass GeV/c^2	EM Charge	Force Mediated
Photon (γ)	$< 10^{-18}$	0	Electro-magnetic
$W^{\pm}$	80.40	± 1	Weak
Z	91.19	0	Weak
Gluon (g)	$0^{\dagger}$	0	Strong
Higgs (H)	> 114.4	0	-

†: Theoretical mass value. Masses of up to a few MeV/c^2 may not ruled out.

Table 1.2: The gauge bosons of the Standard Model. The mass limits for the photon and the Higgs are at the 95% confi dence level.

- ***CP* symmetry violation.** In which the production of matter is preferred over antimatter in the decay of elemental particles.

The SM is known to violate baryon number in the non-perturbative regime. The expansion period of the early universe provides the necessary lack of thermal equilibrium. The last condition- that of the violation of the combined charge conjugation (C) and parity (P) symmetries- was discovered in kaon decays in 1964 [4]. Since then *CP* violation has been the focus of intense experimental investigation. All current measurements of *CP* violation have an understood source within the SM, however the magnitude of the total measured violation is not enough to explain the matter-antimatter asymmetry. New sources of violation must be found.

Explanations for the astrophysical phenomena of dark matter and dark energy must also be found in a comprehensive theory. Dark matter, a massive form of matter that interacts only weakly and gravitationally, was hypothesised to explain the observed apparant non-Newtonian rotational distributions of galaxies [5]. While a modification of Newtonian principles might provide a solution, further evidence for the existence of dark matter is provided by the mass distribution in the collision of two galaxies in the Bullet cluster [6]. Dark energy is the theoretical force driving the acceleration of the expansion of the universe observed in surveys of Type 1a supernovae [7]. Dark energy is calculated to make up over 70% of the energy density of the universe.

Another sign that the SM is not a complete theory is the recent observation of flavour mixing in atmospheric and solar neutrinos [8] [9]. The SM describes neutrinos as massless but mixing implies that they must have some mass. The SM can be modified to include massive neutrinos, with consequences of lepton number violation and perhaps a new source of *CP* violation within the lepton sector.

The weak interaction is fertile ground for the investigation of inconsistencies within the SM. Flavour physics in the weak sector is described in detail in the following section.

1.3 Flavour Physics in The Weak Sector

The weak interaction has the unique property within the SM of being able to change fermion flavour. The most common example of this is the decay of a u quark to a d quark in the beta decay of a neutron to a proton via W boson emission (Figure 1.1). Much experimental effort in the past few decades has been devoted to investigating the exact nature of flavour physics in the weak interaction. The two bosons that mediate the weak force give rise to two types of weak interaction: the charged and neutral interactions. This section describes the current theoretical understanding and the motivation for investigation in this area.

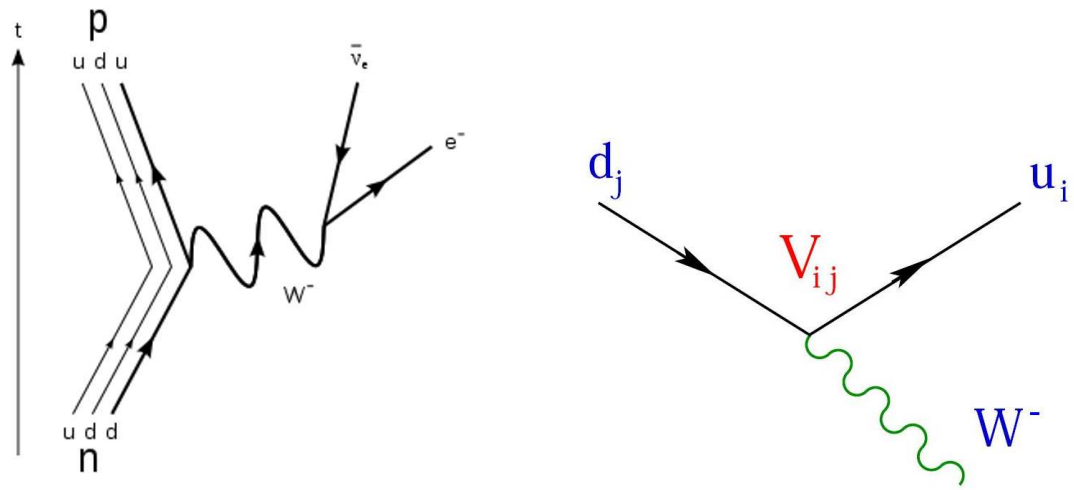


Figure 1.1: Feynman diagrams of neutron beta decay (left) and the generalised flavour-changing charged-current weak transition (right), showing the responsible element of the CKM matrix.

1.3.1 The Charged Current

The charged current involves the exchange of the charged W bosons. The emission of a W^+ or W^- boson by a fermion changes its charge and its flavour. The quark weak charged current element of the SM Lagrangian is given by the expression

$$\mathcal{L}_{CC} = -\frac{g}{2\sqrt{2}} \left(W^+_\mu \overline{u} \gamma^\mu (1-\gamma^5) V_{\mathrm{CKM}} d + h.c. \right) \tag{1.2}$$

where d and $\overline{u}$ are vectors of down-type quarks and up-type anti-quarks, repectively, and V_{CKM} is the Cabbibo-Kobayashi-Maskawa mixing matrix [1]. V_{CKM} is a unitary 3×3 matrix which couples the quark mass eigenstates to the weak eigenstates. By convention the matrix transforms up-type quarks into down-type quarks, and is written as

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.3}$$

Each element of the CKM matrix describes the magnitude of the coupling between an up-type quark and a down-type quark. This is represented by the right-hand Feynman diagram in Figure 1.1.

The CKM matrix is parameterised by three Euler angles and six phases. Five of the six phases can be eliminated by suitable choices of phase differences. The remaining phase is complex and irreducible and is the source of all *CP* violation within the SM.

The *CP* symmetry is the combination of the charge conjugation symmetry, which transforms a particle into its anti-partner, and the parity symmetry, which reflects spacial co-ordinates. Both *C* and *P* have been

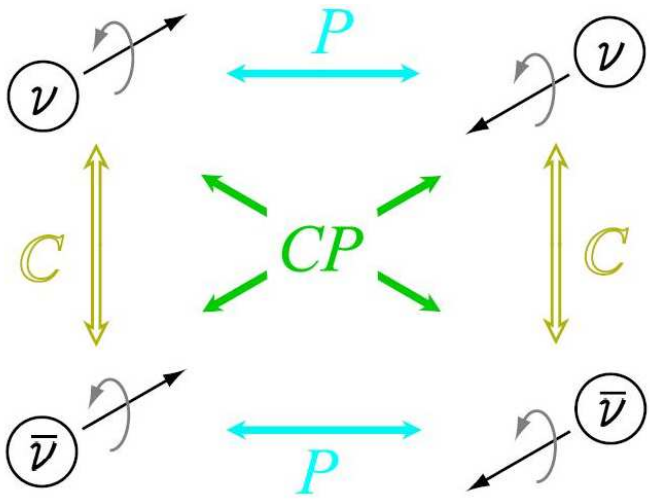


Figure 1.2: The action of the *C*, *P* and combined *CP* transformations on neutrino states. *P* relates left-handed helicity states (left) to right-handed states (right), while *C* relates neutrino (top) to anti-neutrino (bottom). Only the left-handed neutrino (top left) and the right-handed anti-neutrino (bottom right) have been found to exist.

found to be violated maximally in the SM, which is why $SU(2)_L$ is the form of the weak interaction Lagrangian. This is demonstrated by the fact that no right-handed neutrinos or left-handed anti-neutrinos have been discovered in any interactions. *P* transformation will change a left-handed neutrino into a right-handed neutrino, which does not exist. *C* transformation will change a left-handed neutrino into a left-handed anti-neutrino, which also does not exist. The combined operation is needed to relate the existent neutrino states, as shown in Figure 1.2.

Though the CKM matrix can be parameterised in several useful ways, its properties are demonstrated well by the Wolfenstein parameterisation [10]:

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2+\frac{1}{2} & 1 \end{pmatrix} \tag{1.4}$$

which is correct up to terms of order λ^4. As $\lambda = \sin\theta_{\mathrm{Cabbibo}} \approx 0.22$, the magnitude of the couplings between quarks in the same generation- the diagonal terms- are close to unity, the couplings between the first and second generations are of order λ, the couplings between the second and third are of order λ^2, and the couplings between the first and third generations are the smallest at order λ^3. *CP* violation enters the matrix via the complex element in these extreme off-diagonal λ^3 terms. It is parameterised by η: *CP* violation will be present in the SM for non-zero values of η.

The unitarity of V_{CKM} allows the imposition of the following constraints:

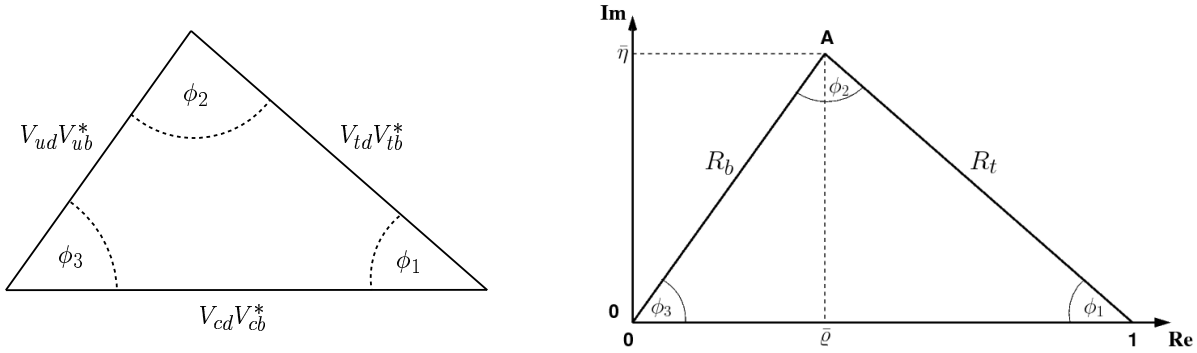


Figure 1.3: The Unitarity Triangle before (left) and after (right) normalisation.

$$\sum_{i=u,c,t} V_{ij}^* V_{ik} = \delta_{jk} \quad \text{and} \quad \sum_{i=d,s,b} V_{ij}^* V_{kj} = \delta_{ik} \tag{1.5}$$

which leads to a number of equations each requiring the sum of three complex quantities to equate to zero. These can be represented as triangles in the complex plane, called the unitary triangles. One such triangle is defined by the equation

$$V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0 \tag{1.6}$$

When the phase convention such that $V_{cb}^* V_{cd}$ is real is taken, and the above relation is normalised by $|V_{cb}^* V_{cd}|$, the resultant triangle is called the Unitary Triangle. This triangle is of special interest as the length of its sides are of approximately the same order. In the absence of *CP* violation the unitarity triangles would collapse onto the real axis, with no imaginary componant. Convention fixes the length of the base to unity and the vertices to (0,0), (1,0) and $(\overline{\rho}, \overline{\eta})$, where

$$\overline{\rho} \equiv \left(1 - \frac{\lambda^2}{2}\right) \rho, \quad \overline{\eta} \equiv \left(1 - \frac{\lambda^2}{2}\right) \eta \tag{1.7}$$

Figure 1.3 shows the Unitarity Triangle in the Argand plane both before and after normalisation. The side lengths of the triangle are

$$R_b \equiv \left| \frac{V_{ub}^* V_{ud}}{V_{cb}^* V_{cd}} \right| = \sqrt{\overline{\rho}^2 + \overline{\eta}^2} \tag{1.8}$$

$$R_t \equiv \left| \frac{V_{td}^* V_{tb}}{V_{cb}^* V_{cd}} \right| = \sqrt{(1 - \overline{\rho}^2) + \overline{\eta}^2} \tag{1.9}$$

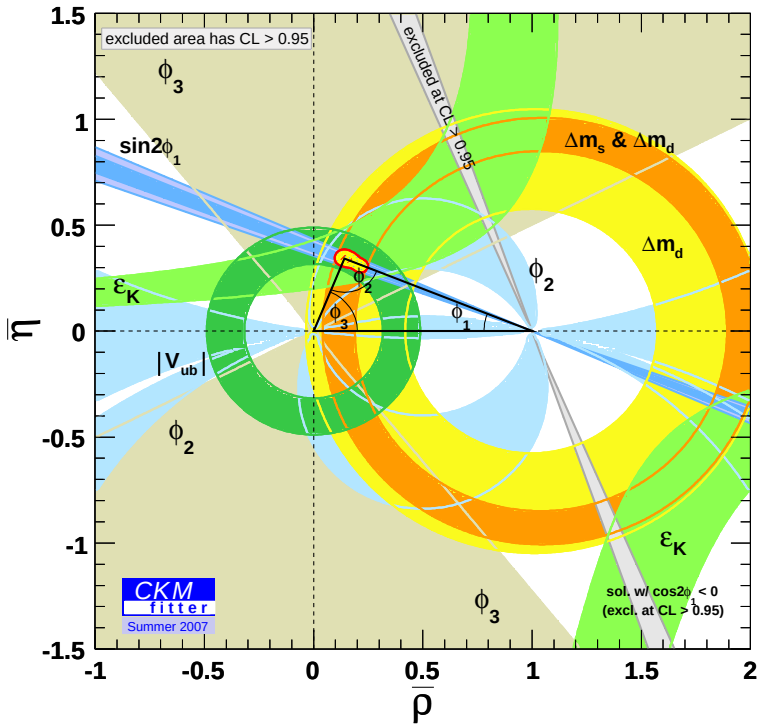


Figure 1.4: Experimental constraints on the Unitarity Triangle provided by a fit by the CKMfitter group in the summer conferences of 2007 [11].

Table 1.3: Experimental values for the Unitary Triangle parameters [11].

$\phi_1 = (21.7 \pm 1.0)^\circ$	$R_b = 0.371 \pm 0.015$
$\phi_2 = (90.7^{+4.5}_{-2.9})^\circ$	$R_t = 0.925^{+0.018}_{-0.030}$
$\phi_3 = (67.6^{+2.8}_{-4.5})^\circ$	$R_c = 1.0$

Table 1.4: Experimental values for the CKM matrix components [11].

$V_{ud} = 0.9740 \pm 0.00018$	$V_{us} = 0.2265 \pm 0.00077$	$V_{ub} = 0.0036 \pm 0.00017$
$V_{cd} = 0.2264 \pm 0.00076$	$V_{cs} = 0.9732 \pm 0.00018$	$V_{cb} = 0.0405 \pm 0.00320$
$V_{td} = 0.0087 \pm 0.00033$	$V_{ts} = 0.0407 \pm 0.00090$	$V_{tb} = 0.9992 \pm 0.00004$

and the angles are

$$\phi_1 \equiv arg\left(-\frac{V_{cb}^* V_{cd}}{V_{td}^* V_{tb}}\right) \tag{1.10}$$

$$\phi_2 \equiv arg\left(-\frac{V_{td}^* V_{tb}}{V_{ud}^* V_{ub}}\right) \tag{1.11}$$

$$\phi_3 \equiv arg\left(-\frac{V_{ud}^* V_{ub}}{V_{cb}^* V_{cd}}\right) \tag{1.12}$$

The current constraints on the Unitarity Triangle are shown in Figure 1.4. R_t is constrained using measurements of neutral B meson mixing to obtain $\triangle m_d$ and $\triangle m_s$ which give access to $|V_{tb}|$ and $\frac{|V_{td}|}{|V_{ts}|}$. R_b is constrained by measuring the rates of $b \to u$ and $b \to c$ transitions to gain access to $\frac{|V_{ub}|}{|V_{cb}|}$. The angles ϕ_1, ϕ_2 and ϕ_3 can be determined through the measurement of B meson decays to *CP* eigenstates: states in which the final state particles are invariant under *CP* transformation. ϕ_1 is the most strongly constrained angle from time-dependant *CP* asymmetry measurements of the $b \to c\overline{c}s$ transition $B^0 \to J/\psi K_S^0$. ϕ_2 is accessed using $b \to u\overline{u}d$ decays such as $B^0 \to \rho^+\rho^-$ and $B^0 \to \pi^+\pi^-$. ϕ_3 is the most difficult angle to determine accurately. The interference of $b \to c$ and $b \to u$ transitions in $B \to DK$ decays are used to get the best constraint. The current world averages for the side lengths and angles are listed in Table 1.3. The experimentally determined values of the CKM matrix parameters are listed in Table 1.4.

1.3.2 The Neutral Current

The element of the SM Lagrangian that describes the quark weak neutral current can be expressed as [1]

$$\mathcal{L}_{NC} = -\frac{e}{2\sin\theta_W \cos\theta_W} Z_\mu \sum_f \left[\overline{f}\gamma^\mu (T_3^f - 4T_3^f |Q_f| \sin^2\theta_W - T_3^f \gamma_5) f\right] \tag{1.13}$$

where f and $\overline{f}$ denote the quarks and anti-quarks. Unlike the charged current, the quark neutral current Lagrangian is identical whether it is expressed in terms of mass eigenstates, as it is here, or in terms of weak eigenstates. This means that no analogue of the CKM matrix arises in the neutral current and mixing of flavours via the Z boson does not occur within the SM.

The mechanism behind this can be seen in the transformation of the quark weak eigenstates (to which the weak force couples) to their mass eigenstates (which we measure in the lab). The quark weak eigenstates can be expressed in terms of their corresponding mass eigenstates as

$$d'_L \equiv \quad S_d d_L \,, \qquad u'_L \equiv \quad S_u u_L \tag{1.14}$$

$$d'_R \equiv \quad S_d U_d d_R \,, \quad u'_R \equiv \quad S_u U_u u_R \tag{1.15}$$

where d_L, d_R and u_L, u_R are vectors of down-type and up-type quark mass eigenstates with left- and right-

handed helicity, respectively, the primes denote the vectors of weak eigenstates, and S_d, S_u, U_d and U_d are unitary (ie, $S^\dagger S, U^\dagger U = 1$) matrices that derive from the diagonalisation of the mass matrices that translate the eigenstates. Thus in the coupling of the weak eigenstates of two down- or up-type quarks to the Z boson;

$$\overline{d}'_L d'_L = \overline{d}_L S_d^\dagger S_d d_L = \overline{d}_L d_L \tag{1.16}$$

$$\overline{u}'_L u'_L = \overline{u}_L S_u^\dagger S_u u_L = \overline{u}_L u_L \tag{1.17}$$

as contrasted with the coupling of an up-type quark to a down-type quark via the W boson;

$$\overline{u}'_L d'_L = \overline{u}_L S_u^\dagger S_d d_L \equiv \overline{u}_L V_{CKM} d_L \tag{1.18}$$

This is the Glashow-Iliopoulos-Maiani (GIM) mechanism [12], which derives from the fact that the unitary matrices S and U are identical for the different generations but different for up- and down-type quarks. Thus each component in the mass eigenstate vectors in equation 1.18 are combinations of the components of the weak eigenstates vectors with mixing defined by V_{CKM}, whereas the components of the mass eigenstate vectors in equations 1.16 and 1.17 are each pure weak eigenstates.

Processes that change quark flavour without changing EM charge- called flavour-changing neutral current (FCNC) interactions- are therefore forbidden at tree-level in the SM and must proceed via second-order loop interactions. This involves the emission and subsequent re-absorption of a W boson, resulting in a final state equivalent to that of a tree-level Z boson emission. These highly suppressed interactions are the focus of the next Section.

1.4 Penguin Decays and $b \to s\gamma$

Loop-induced interactions are the only possible FCNC processes within the SM. Loop phenomenology was introduced to the SM by Vainshtain, Zakharov and Shifman in 1975 to explain some experimental results in kaon physics [13]. Radiative loop processes in which one of the loop-mediating particles emits a gauge boson are commonly referred to as "penguin" decays. Figure 1.5 is one of the diagrams included in the paper by John Ellis that coined the name "penguin" [14]. The W coupling the two quarks at each vertex is not shown in Figure 1.5; at low-energy the Fermi coupling constant (G_F) can be used to describe the four-fermion coupling. The Feyman diagram of the same process in electro-weak theory with W^+ boson exchange is shown in Figure 1.6(b).

Penguin transitions of the b quark have two possible decay paths: $b \to s$ or $b \to d$. The loop is mediated by a W boson and an up-type quark, one of which radiates a boson. The choice of radiated boson leads to the classification of penguin diagrams as: electro-magnetic (radiated photon); electro-weak (virtual photon

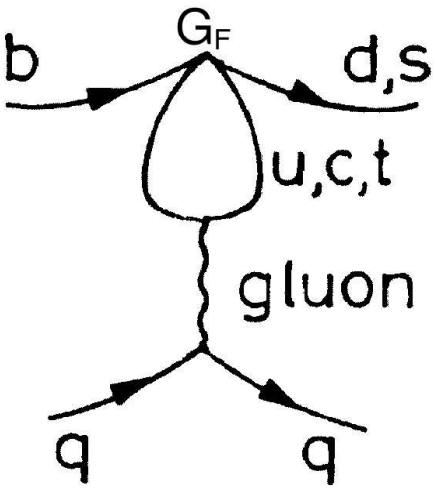
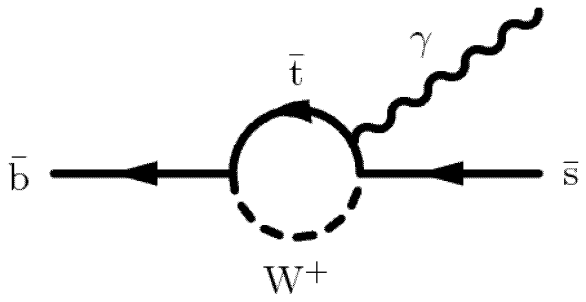


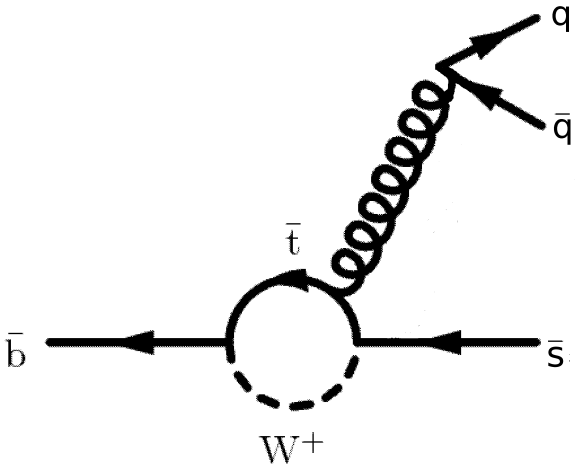
Figure 1.5: One of the original penguin diagrams from the paper by J. Ellis *et al.* [14] showing a b quark decaying to a d or s quark via a u,c or t quark loop with a radiated gluon. The gluon decays into a $q\overline{q}$ pair.

or Z decaying to a lepton/anti-lepton pair); and gluonic (radiated gluon decaying to a $q\overline{q}$ pair). Feynman diagrams for examples of these processes can be seen in Figure 1.6.

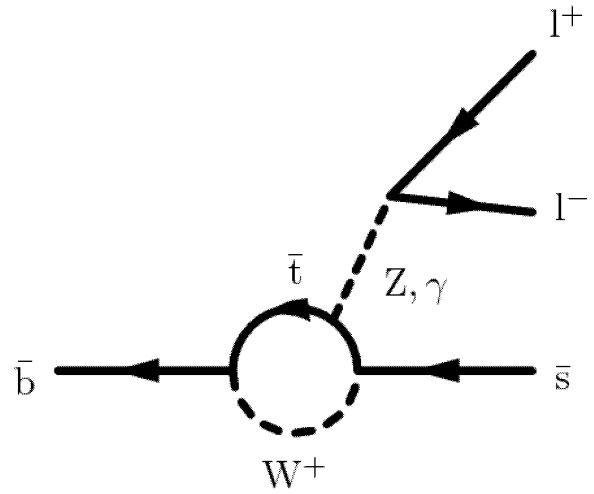
FCNC decays are powerful probes of unknown physics beyond the SM. Any undiscovered heavy particles that couple to SM quarks can mediate the loop instead of the W and t. The measured properties of the decays will then be shifted away from SM predictions, and a comparison of theoretical and experimental results could lead in an indirect detection of new physics. Penguin decays are highly suppressed due to the nature of second-order weak processes; this makes them difficult to study, but also provides the sensitivity needed to detect small contributions from beyond SM physics to the measured rates of decay. Theoretical particles that could mediate the loop include a charged Higgs replacing the W, or the super-symmetric partners of a quark and a gauge boson from super-symmetry theory. The electro-magnetic penguin decay of a b quark to a s quark via photon emission ($b \to s\gamma$) provides the most stringent test of physics beyond the SM, and is the focus of this analysis. Feynman diagrams for two possible beyond-SM contributions to $b \to s\gamma$ are shown in Figure 1.7.



(a) Electro-magnetic penguin decay.

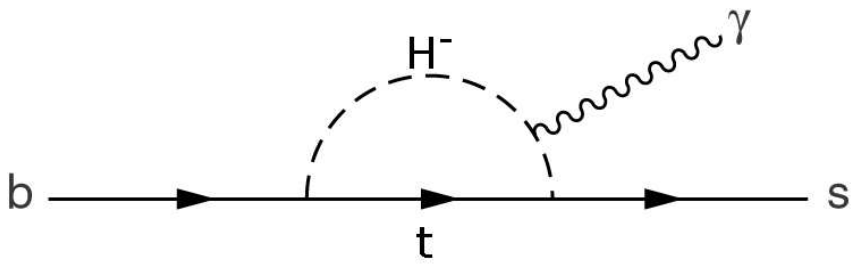


(b) Gluonic penguin decay.

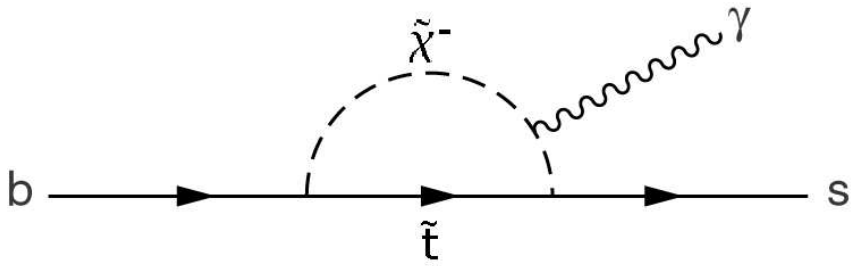


(c) Electro-weak penguin decay.

Figure 1.6: Feynman diagrams showing a $\overline{b}$ quark decaying to a $\overline{s}$ quark via a $\overline{t}/W^+$ loop with a radiated boson.



(a) Charged Higgs boson/top mediated loop.



(b) Super-symmetric chargino/stop mediated loop.

Figure 1.7: Feynman diagrams showing $b \to s\gamma$ transitions mediated by possible beyond-SM particles.

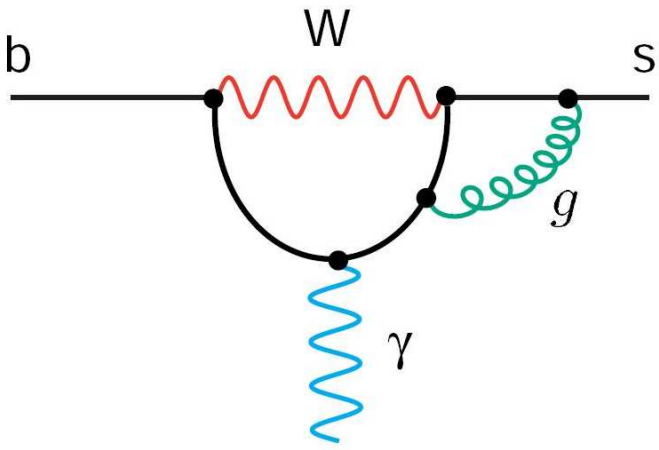


Figure 1.8: An example of a QCD correction to $b \to s\gamma$.

1.4.1 $b \to s\gamma$

The SM theoretical branching fraction ($\mathcal{BF}$) of $b \to s\gamma$ is calculated using the effective Hamiltonian [16]

$$H_{eff}(b \to s\gamma) = \frac{4G_F}{\sqrt{2}}|V_{ts}^* V_{tb}| \sum_{i=1}^{8} C_i(\mu) O_i(\mu) \tag{1.19}$$

where V_{ts}^* and V_{tb} are elements of the CKM matrix, C_i are the Wilson co-efficients, O_i are local operators, and μ is the renormalisation scale which is chosen to be of order m_b. The assumption that local operators of dimension greater than 6 are suppressed by powers of $\frac{1}{m_W}$ or $\frac{1}{m_t}$ reduces the contributing operators to

$$O_2 = (\overline{s}_L \gamma_\mu c_L)(\overline{c}_L \gamma^\mu b_L), \quad O_7 = \frac{e}{16\pi^2} m_b \overline{s}_L \sigma_{\mu\nu} F^{\mu\nu} b_R, \quad O_8 = \frac{g_s}{16\pi^2} m_b \overline{s}_L \sigma_{\mu\nu} G_a^{\mu\nu} t_a b_R \tag{1.20}$$

where e and g_s are the EM and strong force coupling constants, respectively, t_a is a SU(3) generator and $F^{\mu\nu}$ and $G_a^{\mu\nu}$ are the EM and QCD field strength tensors, respectively.

The effective Hamiltonian is formed from the full SM theory by utilising the operator product expansion (OPE) method to integrate out the heaviest SM particles- the electroweak bosons and the t quark- whose dependence is then contained in the Wilson co-efficients. These co-efficients are required to match the full theory at an energy scale of order $m_{W,t}$, and are then evolved down to the scale of m_b by use of a renormalisation group derived from the scale independence of H_{eff}. This leaves a low-energy five-quark effective theory, which matches up to the full theory when scaled up in energy.

Within this framework the QCD corrections to the decay rate can be calculated [16]. The decay rate is given by

$$\Gamma(b \to s\gamma) = \frac{G_F^2 \alpha m_b^5}{32\pi^4} |V_{ts}^* V_{tb}|^2 \left[|C_{7\gamma} m_b|^2 + O\left(\alpha_s, \frac{\Lambda_{QCD}^2}{m_b^2}\right) \right] \tag{1.21}$$

where the second term describes the QCD corrections at order α_s. These corrections give very large contributions to the decay rate- the dominant ones enhance the total rate by a factor of two- and derive from hard-gluon exchange, bremsstrahlung and virtual interactions. An example of such a correction is shown in Figure 1.8. They bring in large logarithms of the form

$$\alpha_s^n \log^n\left(\frac{m_b}{M_{W,t}}\right) \tag{1.22}$$

which have to be summed up to the order of α_s desired. This is a highly non-trivial enterprise which has involved years of work from many theorists.

As quarks are always found in a bound state, corrections must also include effects from the second 'spectator' quark in the meson. Experimental measurements access $b \to s\gamma$ when the b is bound with a light u or d quark within a B meson: $B \to X_s\gamma$; where X_s denotes the hadronic recoil component which

consists of the s quark decay product, the spectator u or d quark, and any other quarks that form in the hadronisation process. It is the b quark which decays in the $B \to X_s\gamma$ process. The calculated theoretical $b \to s\gamma$ rate is approximated to the $B \to X_s\gamma$ rate using the ratio [15]

$$\frac{\Gamma(B \to X_s\gamma)}{\Gamma(B \to X_c e\overline{\nu}_e)} \simeq \frac{\Gamma(b \to s\gamma)}{\Gamma(b \to ce\overline{\nu}_e)} \tag{1.23}$$

from which it follows that

$$\mathcal{BF}(B \to X_s\gamma) \simeq \frac{\Gamma(b \to s\gamma)}{\Gamma(b \to ce\overline{\nu}_e)}\mathcal{BF}(B \to X_c e\overline{\nu}_e) \tag{1.24}$$

This is based on the assumption that the quark-quark interactions within the meson are the same for both processes. $\mathcal{BF}(B \to X_c e\overline{\nu}_e)$ has been measured to better than 4% accuracy. Normalising to the semi-leptonic rate in this manner also removes the strong dependence of the calculation on m_b.

The $B \to X_s\gamma$ decay rate at leading order (LO) or α_s^0 was calculated in 1994 to be $\mathcal{BF}$ $(B \to X_s\gamma)$= $(2.8 \pm 0.8) \times 10^{-4}$ when normalised to the total B decay rate [15]. This was updated in 2002 to next-to-leading order (NLO) or α_s^1 as $\mathcal{BF}$ $(B \to X_s\gamma)$= $(3.57 \pm 0.30) \times 10^{-4}$ [17]. The current state of knowledge is of the order of α_s^2, involving diagrams with up to three gluonic loops. This NNLO calculation was published in January 2007 and gave $\mathcal{BF}$ $(B \to X_s\gamma)$= $(3.15 \pm 0.23) \times 10^{-4}$ [18]. The alteration of the central value and the reduction of the error on the predicted rates is testimony to the importance of QCD corrections in this calculation.

Experimental measurements of $\mathcal{BF}$ $(B \to X_s\gamma)$ are performed in one of two ways: the inclusive or the sum of exclusive modes methods. The inclusive method involves recording the high energy photon without reconstruction of the hadronic X_s component. Photon backgrounds are then subtracted from the spectrum and the total number of recorded $B \to X_s\gamma$ decays measured by integrating over the spectrum. The method itself is quite model-independentbut is subject to inaccuracy in the background subtraction. The spectrum has only been measured to a lower E_γ energy limit of 1.7 GeV and must therefore be extrapolated to energies below this by use of a theoretical model. This process results in analysis model dependencies.

The sum of exclusive modes method proceeds by recording the high energy photon and reconstructing X_s as a kaon (K^+, K^-, K^0 or $\bar{K}^0$, which contain the strange quark) plus charged and neutral pions. Some analyses also include other light unflavoured mesons such as the η. The backgrounds are suppressed by the kinematic selection criteria used in the X_s reconstruction. The need to estimate the total X_s decay phase space the analysis is sensitive to requires a model of the hadronisation process. This process is poorly understood and the estimation introduces model dependencies into the analysis.

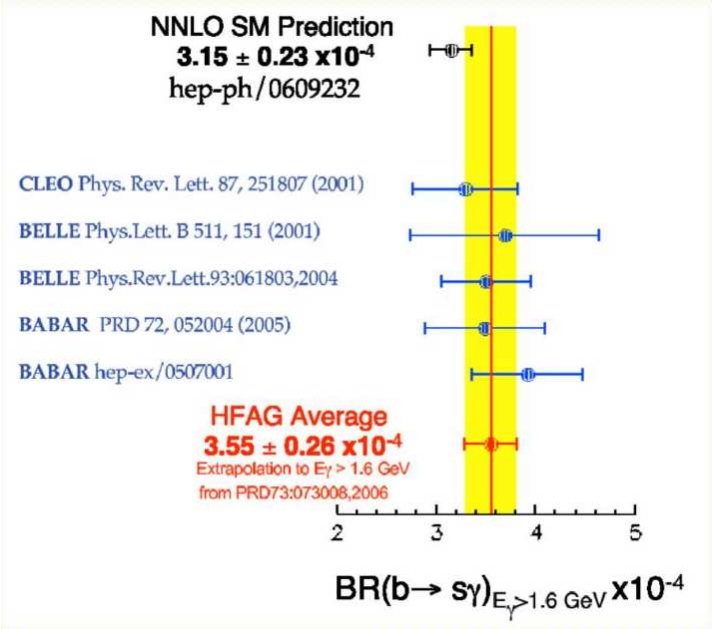


Figure 1.9: The world experimental average and the theoretical prediction of $\mathcal{BF}$ $(B \to X_s\gamma)$.

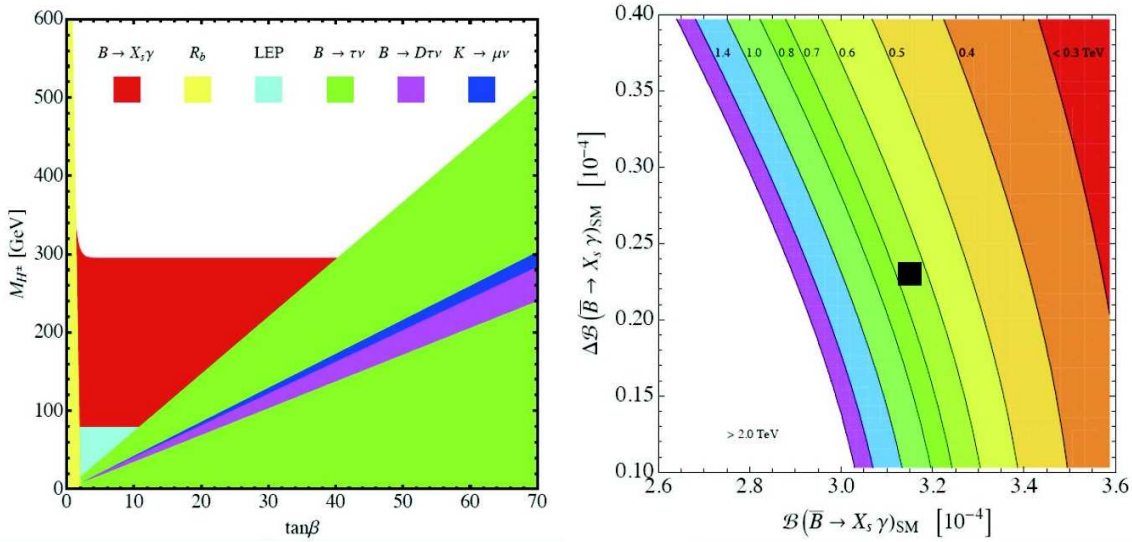


Figure 1.10: The constraints imposed by the experimental bounds on $\mathcal{BF}$ $(B \to X_s\gamma)$ on beyond SM theories. The left plot shows the restrictions on the Higgs mass as a function of $\tan\beta$ for the type-II two-Higgs-doublet model, with the region dis-allowed by $\mathcal{BF}$ $(B \to X_s\gamma)$ in red. The right plot shows the limits on the compactifi cation scale of the universal extra-dimension model type 6 as a function of the $\mathcal{BF}$ $(B \to X_s\gamma)$ central value and error. The black square is the present SM theoretical value.

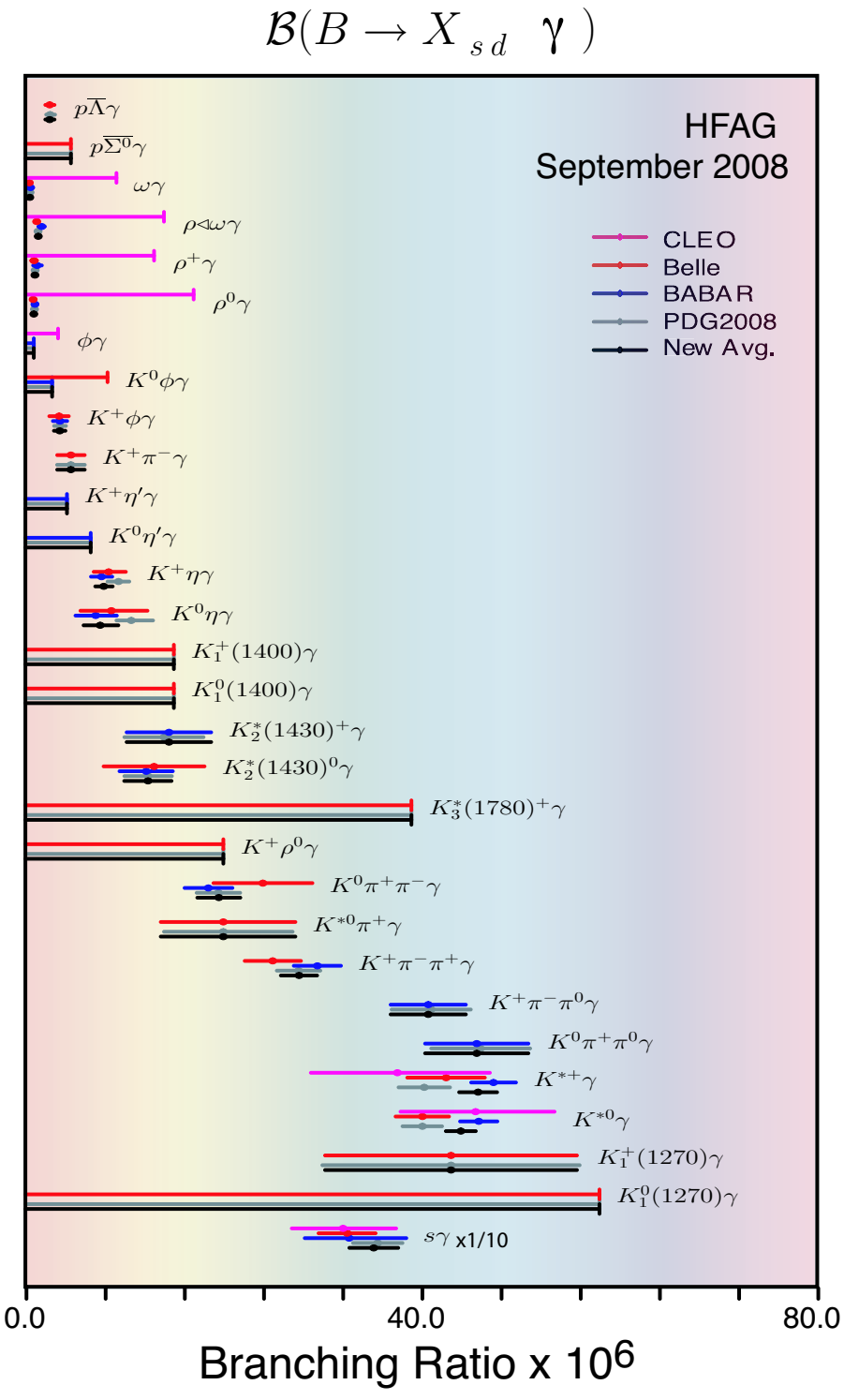


Figure 1.11: A graphical summary of the measurements of $B \to X_s\gamma$ and $B \to X_d\gamma$ exclusive state branching fractions. Also shown is world average inclusive branching fraction at $1/10^{th}$ of the true value. States in which the error bars overlap zero are upper limits at 90% confi dence.

The current world experimental average is $\mathcal{BF}\ (B \to X_s\gamma) = (3.55 \pm 0.24^{+9}_{-10} \pm 0.3) \times 10^{-4}$ [22] where the errors are the combined statistical and systematic, the systematic due to the shape function used to extrapolate the photon spectrum below the minimum photon energy measured, and the fraction of the $\mathcal{BF}$ due to the $b \to d\gamma$ process. This average and the results that produced it are compared to the NNLO theoretical prediction in Figure 1.9. Experiment agrees with the SM prediction, though there is a slight tension of 1.3 standard deviations between the two. Any influence of new physics on the $\mathcal{BF}$ is not evident within the levels of theoretical and experimental uncertainty. This agreement is a powerful tool in restricting the parameter space of beyond SM theories of physics. All theories which include new heavy particles that couple to SM quarks must not introduce extra elements to the $b \to s\gamma$ decay rate which would alter the measured $\mathcal{BF}\ (B \to X_s\gamma)$ beyond experimental bounds. For example, the restrictions for the type-II two-Higgs-doublet model and the universal extra-dimension model are shown in Figure 1.10 [21].

It is also useful to study exclusive modes of the $B \to X_s\gamma$ process. In this case the hadronic X_s is reconstructed in certain specific form. These specific reconstructions are also sensitive probes of beyond SM physics. They additionally provide information on the hadronisation process of X_s which reduces the modelling errors involved in the sum of exclusive $B \to X_s\gamma$ measurement described above. Only approximately 35% of the inclusive decay rate can be attributed to measured exclusive modes. The current state of experimental data on exclusive $B \to X_s\gamma$ modes is summarised in Figure 1.11.

1.5 $B \to K\eta'\gamma$

$B \to K\eta'\gamma$ is an un-measured exclusive $b \to s\gamma$ mode. A Feynman diagram of the process $B^- \to K^-\eta'\gamma$ is shown in Figure 1.12. In 2006 B. Aubert *et al.* published upper limits of $\mathcal{BF}\ (B^+ \to K^+\eta'\gamma) < 4.2 \times 10^{-6}$ and $\mathcal{BF}\ (B^0 \to K^0\eta'\gamma) < 6.6 \times 10^{-6}$ from the analysis of $211 fb^{-1}$ of integrated luminosity collected by the BaBar detector at the Stanford Linear Accelerator Center [53]. The measured $\mathcal{BF}$ were

$$\mathcal{BF}(B^+ \to K^+\eta'\gamma) = (1.9^{1.5}_{1.2} \pm 0.1) \times 10^{-6} \tag{1.25}$$

$$\mathcal{BF}(B^0 \to K^0_S\eta'\gamma) = (1.1^{2.8}_{2.0} \pm 0.1) \times 10^{-6} \tag{1.26}$$

As described above, measuring exclusive $b \to s\gamma$ modes provides a further probe of physics beyond the SM and improves our knowledge of the X_s hadronisation process. There is additional interest in $B \to K\eta'\gamma$ due to the history of $B \to K\eta/\eta'$ modes. The η and η' mesons are electrically neutral pseudo-scalar particles with quark composition defined by the superpositions of states $(u\overline{u} + d\overline{d} - 2s\overline{s})/\sqrt{6}$ and $(u\overline{u} + d\overline{d} + s\overline{s})/\sqrt{3}$, respectively. There is some mixing between the two states, meaning that both the η and the η' have some probability of existing in either of the quark state superpositions. They differ in

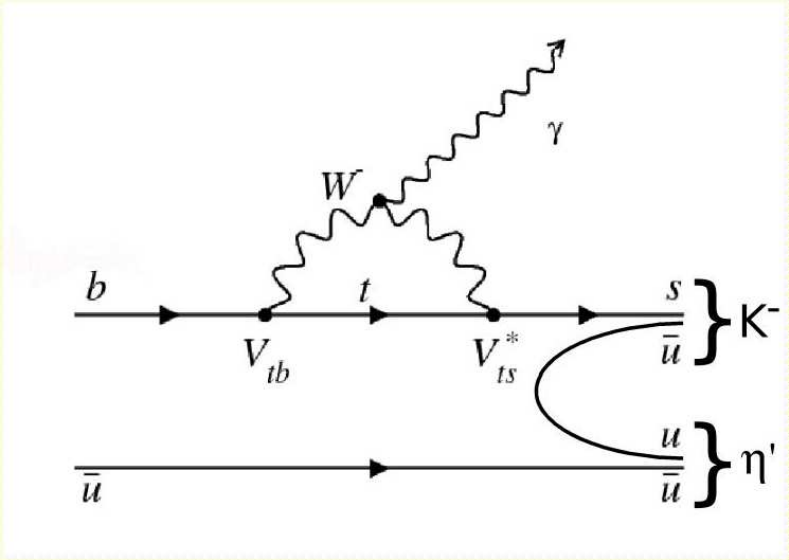


Figure 1.12: Feynman diagram of the $B \to K\eta'\gamma$ process.

quantum numbers from the π^0 only in isospin: $I = 0$ for π^0, $I = 1$ for η and η'. Their place in the spin=0 meson nonet is shown in Figure 1.13.

When the branching fractions were measured there was no known reason for the enhancement of $B \to K\eta'$ and suppression of $B \to K\eta$ with respect to $B \to K\pi$: the world average branching fractions are $\mathcal{BF}\ (B^+ \to K^+\eta') = (7.02 \pm 0.25) \times 10^{-5}$, $\mathcal{BF}\ (B^+ \to K^+\eta) = (2.7 \pm 0.9) \times 10^{-6}$, and $\mathcal{BF}\ (B^+ \to K^0\pi^+) = (2.31 \pm 0.10) \times 10^{-5}$ [19]. Some theorists believed it was the first sign of new physics. It was demonstrated in 2002 that the relative decay rates can be accounted for as the constructive or destructive interference of non-singlet penguin amplitudes, and the correct $\mathcal{BF}$s can be calculated using QCD factorisation, albeit with large errors [20]. It would further these investigations if a similar calculation for the decays $B \to K\eta'\gamma$, $B \to K\eta\gamma$ and $B \to K\pi\gamma$ could be compared to the experimentally measured values.

$B^+ \to K^0\pi^+\gamma$ is very well measured coming through the $K^*(892)^+ \to K^0\pi^+$ resonance, with a world average of BR$(B^+ \to K^*(892)^+\gamma) = (4.03 \pm 0.26) \times 10^{-5}$. $B \to K\eta\gamma$ has also been measured, with a world average of BR$(B^+ \to K^+\eta\gamma) = (9.4 \pm 1.1) \times 10^{-6}$. The measurement of the branching fraction of the exclusive mode $B \to K\eta'\gamma$ is the goal of this analysis.

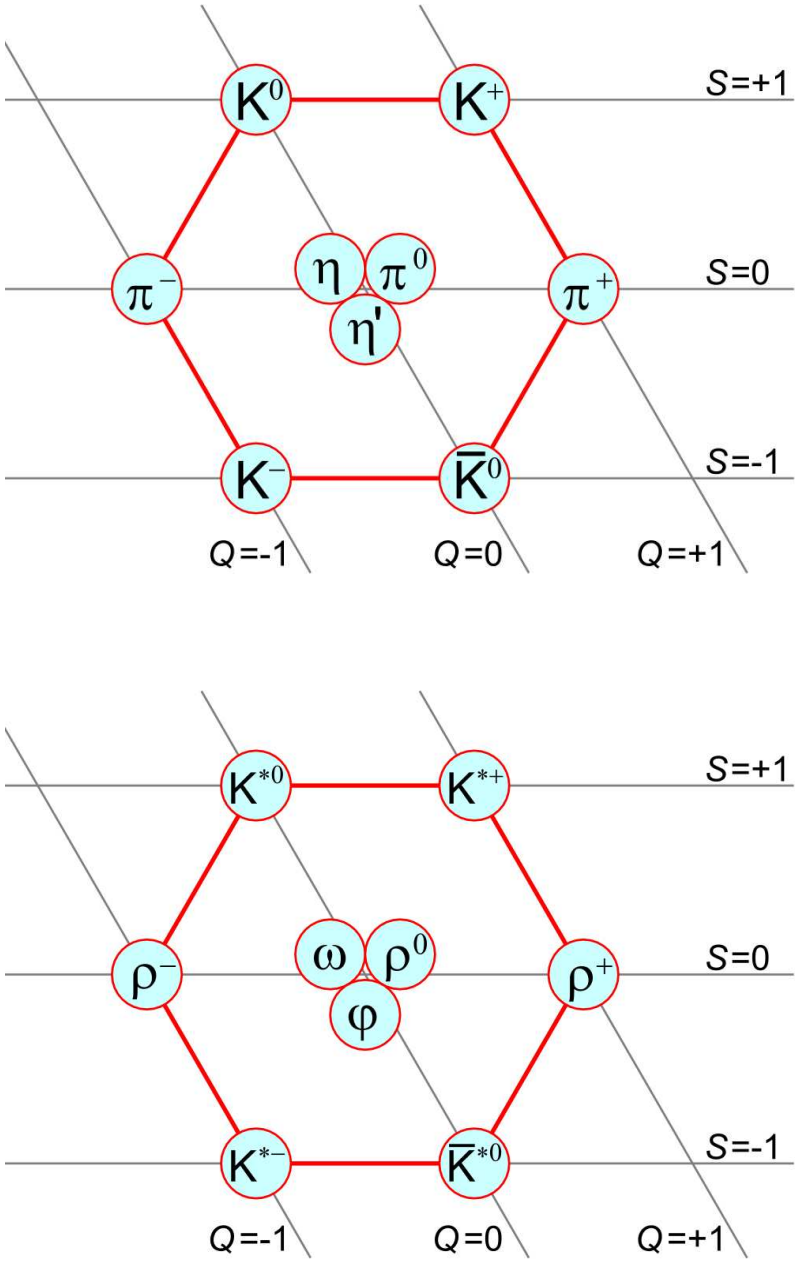


Figure 1.13: The nonets of spin=0 mesons (top), and spin=1 mesons (bottom).

Chapter 2

The Belle Experiment

The Belle experiment is based at the KEK high-energy research facility in Tsukuba, Japan, and consists of the KEKB asymmetric electron-positron collider and the Belle concentric particle detector. The experiment was commissioned in early 1999 and began taking data in June of the same year. The experiment is operated by the Belle collaboration, a partnership of more than 350 physicists from over 50 institutions in 14 countries around the globe.

2.1 B-Factories and the $\Upsilon(4S)$

CP violation was first observed in 1964 at the Brookhaven Laboratory, USA, in the decay of K^0 mesons produced by proton bombardment of a fixed beryllium target [4]. At the time only the u, d and s quarks were known to exist and the combined charge and parity symmetries were believed to be unbroken. The discovery of decays violating the combined symmetry sparked an intense theoretical effort to understand the mechanism by which they proceed, culminating in the proposal in 1973 by Kobayashi and Maskawa of a model in which *CP* is violated by an irreducible complex phase in the matrix describing quark mixing via the weak interaction [29]. The c quark had been predicted and confirmed experimentally in the intervening years, but this matrix- now known as the Cabbibo-Kobayashi-Maskawa (CKM) matrix- required the existence of at least three generations of quarks to provide the mechanism for *CPV*. The KM model thus predicted the existence of a further family of quarks beyond what had been discovered- the b and t quarks. Kobayashi and Maskawa were awarded the Nobel Prize in 2008 for their work. The model and it's consequences are described in more detail in the preceding Chapter.

The KM model was given support in 1977 by the discovery of the b quark at the E288 experiment at Fermilab [32] and by its compatibility with observed *CP* violation in the kaon sector. Carter, Bigi and Sanda realised in 1981 that it predicted large *CPV* in certain B meson decays, and the path was open for the model to be tested thoroughly. Experiments were initiated to produce samples of B mesons for the

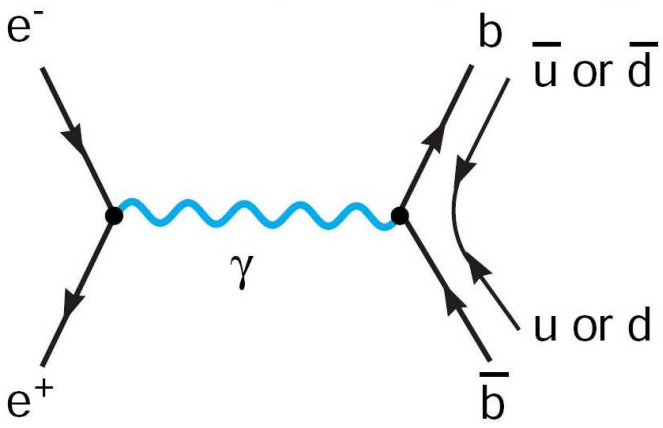


Figure 2.1: $e^+\,e^-$ annihilation to form the $\Upsilon(4S)$ which decays almost immediately to a $B\overline{B}$ pair.

study of *CPV* phenomenology in b quark decay. These so-called 'B-factories' were designed to produce very large samples of Bs in a clean environment in order to accurately measure the *CP*-violating phase and other small elements of the CKM matrix. Table 2.1 lists the major high energy physics experiments from around the world that have produced B physics results. The Belle detector on the KEKB accelerator and the BABARdetector on the PEP-II accelerator are the latest generation of high-luminosity B factories. The next generation of experiments are in development; LHCb will begin operation in the middle of 2009 and an upgrade of the KEKB accelerator and the Belle detector will start in 2010.

The Υ vector mesons are bound $b\bar{b}$ states with quantum numbers $J^{PC} = 1^{--}$ and masses in the range 9 - 11 GeV/c^2. Figure 2.2 shows the hadronic cross-section in $e^+\,e^-$ annihilation as a function of CM energy in the region of the Υ resonances. The mass of the $\Upsilon(4S)$ resonance is just above threshold to produce a $B\overline{B}$ pair, and it therefore decays in this manner almost immediately via the strong interaction more than 96% of the time [33], as shown in Figure 2.1. This makes the $\Upsilon(4S)$ the ideal resonance for the production of many B mesons in a clean environment where all subsequent decay particles can be attributed to either one or the other B meson. The rate at which the $\Upsilon(4S)$ decays to B^+B^- has been found to be very consistent with the rate to $B^0\overline{B}^0$, with the ratio of the branching fractions measured as

$$\frac{\mathcal{B}(\Upsilon(4S) \to B^+B^-)}{\mathcal{B}(\Upsilon(4S) \to B^0\overline{B}^0)} = 1.04 \pm 0.07 \pm 0.04 \tag{2.1}$$

where the first error is statistical and the second systematic [34]. In this analysis the branching fractions are assumed to be equal. The $b\bar{b}$ production cross-section at the $\Upsilon(4S)$ resonance is about 1.1 nano-barn.

B-factories accelerate and collide particles at a CM energy ($\sqrt{s}$) of 10.58 GeV to produce $\Upsilon(4S)$ mesons. The B mesons are almost at rest in the $\Upsilon(4S)$ CM frame. To enable time-dependant B meson decay analysis, the particles collide at asymmetric energies. This gives the $\Upsilon(4S)$ a non-zero momentum in the lab frame and the B mesons a small Lorentz boost in the experiment. The decay times of the B mesons can then be determined by high precision measurements of their flight lengths. One such factory is

the KEKB collider.

Table 2.1: The B-factories that have been constructed around the world.

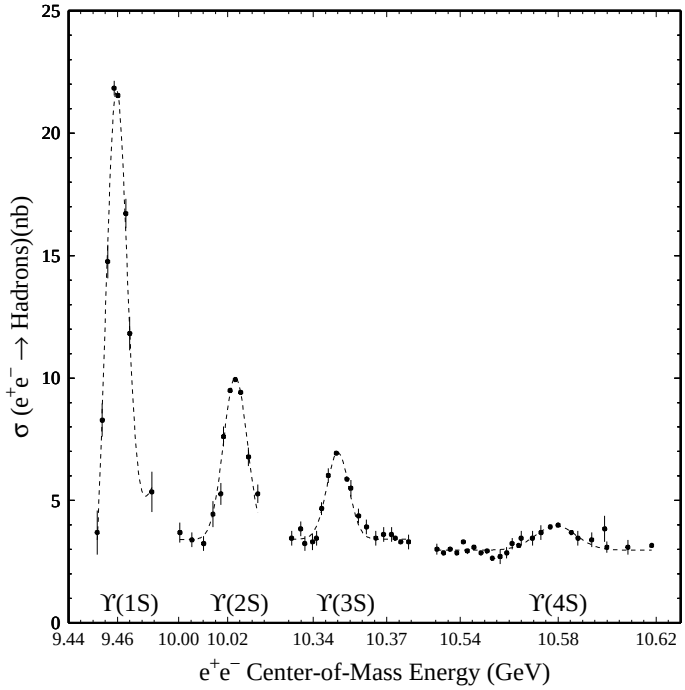
Institute	Accelerator	Type	Detector(s)
DESY	HERA	p-Fixed Target	HERA-B
Cornell	CESR	Symmetric e^+ e^-	CLEO II, III
CERN	LEP	Symmetric e^+ e^-	Aleph, Delphi, Opal, and L3
KEK	KEKB	Asymmetric e^+ e^-	Belle
SLAC	PEP-II	Asymmetric e^+ e^-	BABAR
FNAL	Tevatron	$p\bar{p}$	BTeV
CERN	LHC	pp	LHCb

2.2 The KEKB Storage Ring

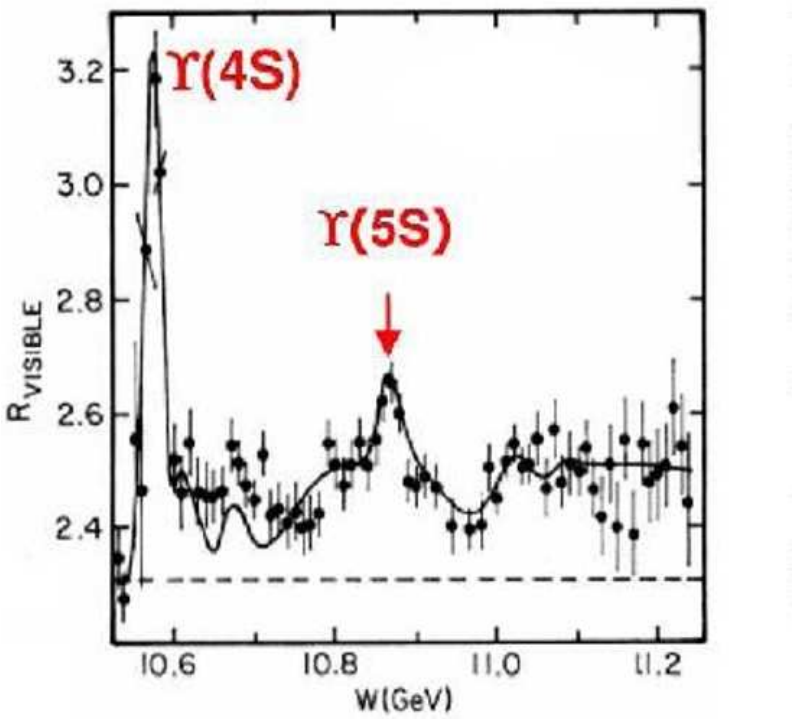
The KEKB asymmetric electron-positron collider is a circular particle accelerator 3 km in diameter. A schematic of the ring is shown in Figure 2.3. The linear accelerator (linac) situated at the 'Fuji' area of the ring injects 8.0 GeV electrons into the high energy ring (HER) and 3.5 GeV positrons into the low energy ring (LER). The different beam energies- and thus different beam bending and tuning requirements- necessitate two separate rings within the single tunnel. RF cavities in the 'Fuji', 'Nikko' and 'Oho' areas compensate for energy lost due to synchrotron radiation. The single interaction point (IP), around which the Belle detector sits, is located in the 'Tsukuba' section opposite the linac. At the IP the beams intersect at an angle of 22 milli-radians; this finite crossing-angle reduces beam-beam interactions away from the IP and removes the need for separation magnets within the detector volume. This also means that the electron and positron bunches do not collide head-on as would happen at zero crossing angle. This raises the effective beam cross-sectional area and causes a reduction in the specific luminosity of the collisions. To compensate for this, specialised RF cavities called Crab Cavities' were installed on the beam-line in January 2007. Crab cavities give each bunch a kick to effectively rotate it to face the colliding bunch directly; this is demonstrated in the left-hand plot of Figure 2.4. The installation occured after the data used in this experiment (see Section 3.1); they are mentioned here for completeness.

The main design parameters of KEKB are listed in Table 2.2. The accelerator was designed to achieve a luminosity ($\mathcal{L}$) of $1 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Luminosity is a measure of the rate of $B\overline{B}$ production, and is most intuitively expressed as

$$\mathcal{L} = \frac{N_- N_+ f}{4\pi\sigma_x^*\sigma_y^*} \tag{2.2}$$



(a) Cross-section in nb and the resonances up to $\Upsilon(4S)$.



(b) The $\Upsilon(4S)$ and $\Upsilon(5S)$ resonances normalised to the theoretical muon-pair cross-section.

Figure 2.2: Hadronic cross-section in e^+ e^- annihilation as a function of CM energy.

where N_- and N_+ are the numbers of electrons and positrons in each bunch, respectively, f is the frequency of collision, and σ_x^* and σ_y^* are the horizontal and vertical beam sizes at the IP, respectively. Thus to increase $\mathcal{L}$ the number of particles in each bunch and the number of bunches in the ring must be maximised and the beams must be focused as much as possible. Luminosity is more commonly expressed as

$$\mathcal{L} = 2.17 \times 10^{34} \xi_y (1 + r)(\frac{I \cdot E}{\beta_y^*})_{\pm} \tag{2.3}$$

where ξ_y is the beam-beam tune shift parameter, E is the beam energy, I is the beam current, r is the aspect ratio of the beam shape, and β_y^* is the vertical beta function at the IP. The $\pm$ subscript refers to the HER or LER; the beams may be tuned independently. Tuning the accelerator parameters to increase luminosity has been on-going since KEKB was commissioned. The record peak luminosity of $1.712 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ was achieved in November 2006. This is equivalent to about 10^8 $B\overline{B}$ pairs per year. Table 2.2 shows the parameters that were in place as of 19^{th} of May 2008. Computer modelling of the effect of the Crab cavities predicted an increase in luminosity to above $2 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, however since the Crab cavity installation beam instabilities have restricted the maximum beam currents: specific $\mathcal{L}$ ($\mathcal{L}$ per bunch crossing) has increased as expected but the lower beam currents have prevented an increase in total luminosity.

Figure 2.5 shows the performance of KEKB over its lifetime until early 2008, including daily peak luminosity, daily peak beam currents, the daily recorded integrated luminosity and recording efficiency, and the total accumulated integrated luminosity recorded. Over $850\,\mathrm{fb}^{-1}$ has been recorded thus far, giving KEKB and Belle the greatest recorded integrated luminosity of any high energy physics experiment ever conducted. The data used in this analysis comprises $605\,\mathrm{fb}^{-1}$ of the full data set, as some of the data is in the processsing phase and some was taken at the $\Upsilon(5S)$ and $\Upsilon(3S)$ resonances (see Sections 2.4.6 and 3.1 for more details on the data set analysed.)

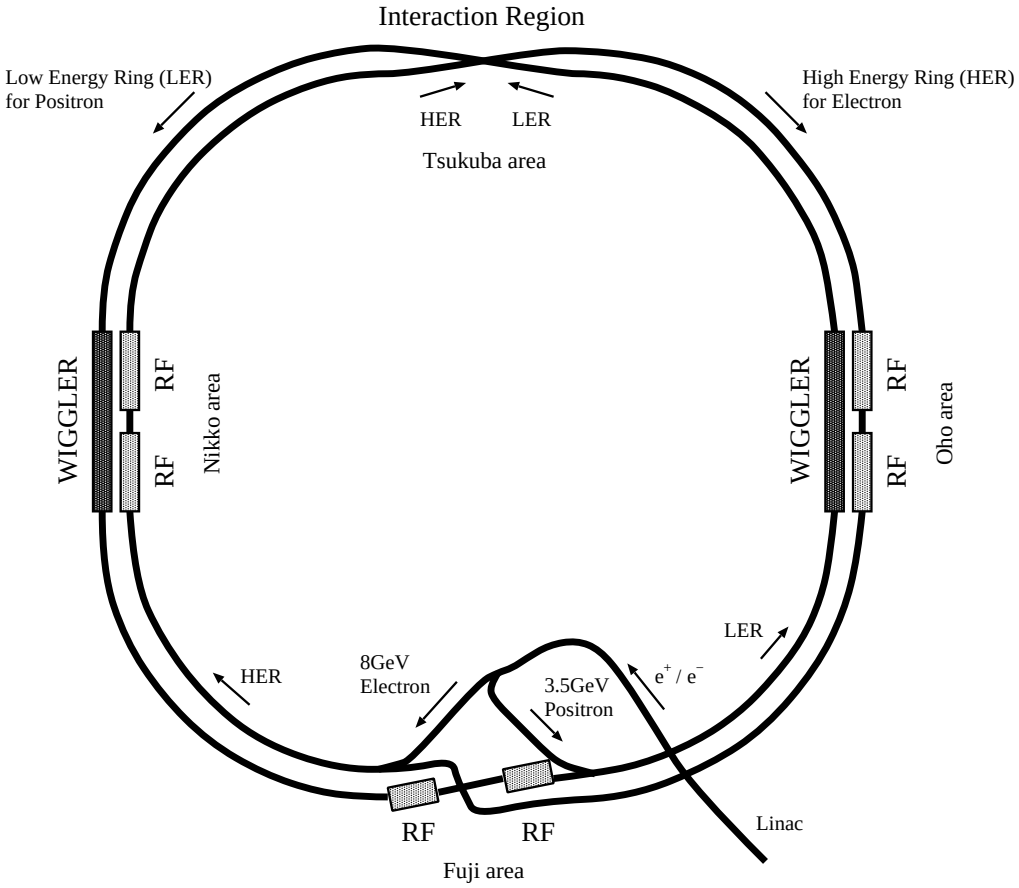


Figure 2.3: A schematic representation of the KEKB storage rings and linear accelerator. The Belle detector is situated at the interaction point.

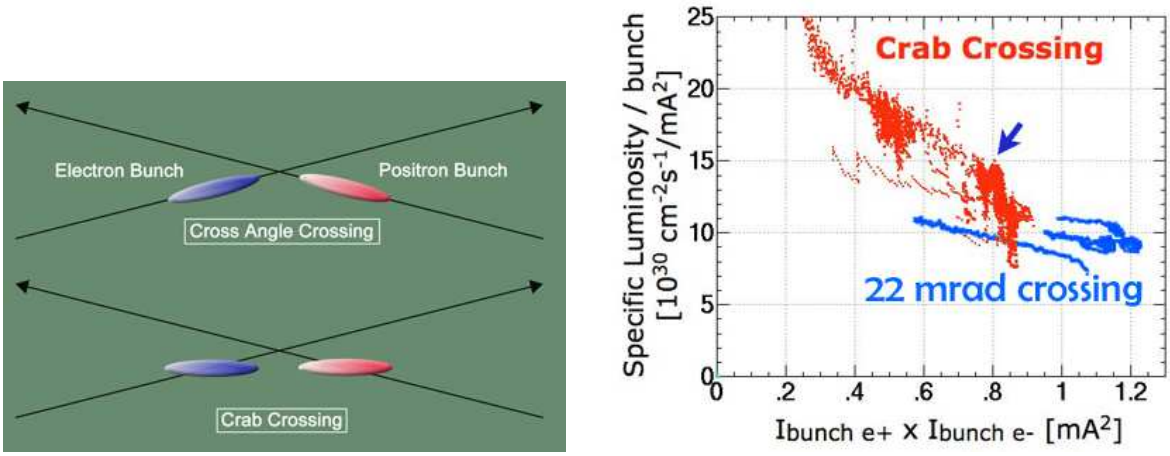


Figure 2.4: The left-hand diagram depicts the beam-crossing without (top) and with (bottom) crab cavities in operation. The right-hand figure shows the expected luminosity per bunch divided by the product of the beam currents, plotted as a function of the beam current product.

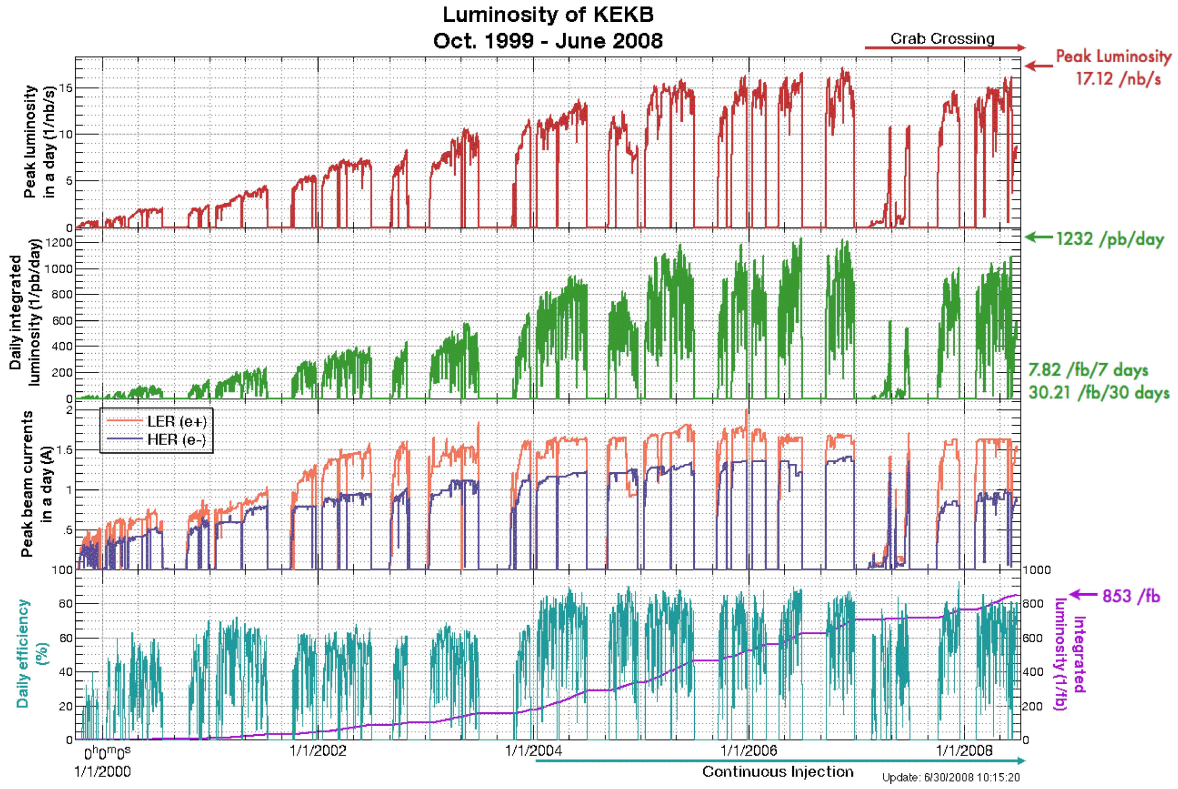


Figure 2.5: KEKB delivered luminosity. From top to bottom: Peak daily instantaneous luminosity, daily integrated luminosity, peak daily beam currents, and daily Belle efficiency with the total integrated recorded luminosity overlaid.

Table 2.2: KEKB accelerator design parameters [38].

Parameter	LER	HER	Unit
Energy (E)	3.5	8.0	GeV
Circumference (C)	3016.26		m
Crossing angle (θ_x)	± 11		mrad
Tune shifts (ξ_x/ξ_y)	$0.039/0.052$		
Beta function at IP (β_x^*/β_y^*)	$0.33/0.01$		m
Vert. beam size at IP (σ_y^*)	1.9		μm
Horiz. beam size at IP (σ_x^*)	80.0		μm
Beam current (I)	2.6	1.1	A
Number of bunches	5000		
Natural bunch length (σ_z)	0.4		cm
Energy spread (σ_E/E)	7.1×10^{-4}	6.7×10^{-4}	
Bunch spacing (s_B)	0.59		m
Particles per bunch (N)	3.3×10^{10}	1.4×10^{10}	
Emittance ($\varepsilon_x/\varepsilon_y$)	$1.8 \times 10^{-8}/3.6 \times 10^{-10}$		m
Synchrotron tune (ν_s)	$0.01 \sim 0.02$		
Betatron tune (ν_x/ν_y)	$45.52/45.08$	$47.52/46.08$	
Momentum compaction factor (α_p)	$1 \times 10^{-4} \sim 2 \times 10^{-4}$		
Energy loss per turn (U_0)	$0.81^{\dagger}/1.5^{\ddagger}$	3.5	MeV
RF voltage (V_c)	$5 \sim 10$	$10 \sim 20$	MV
RF frequency ($f_{\rm RF}$)	508.887		MHz
Harmonic number (h)	5120		
Longitudinal damping time (τ_ε)	$43^{\dagger}/23^{\ddagger}$	23	ms
Total beam power (P_b)	$2.7^{\dagger}/4.5^{\ddagger}$	4.0	MW
Radiation power ($P_{\rm SR}$)	$2.1^{\dagger}/4.0^{\ddagger}$	3.8	MW
HOM power ($P_{\rm HOM}$)	0.57	0.15	MW
Bending radius (ρ)	16.3	104.5	m
Length of bending magnet (l_B)	0.915	5.86	m

†: without wigglers, ‡: with wigglers

Table 2.3: KEKB accelerator parameters as of 19/5/2008.

Parameter	LER	HER	Unit
Tune shifts (ξ_x/ξ_y)	$0.039/0.052$		
Beta function at IP (β_x^*/β_y^*)	$0.90/0.0059$		m
Beam current (I)	1.605	0.934	A
Number of bunches	1584+1		
Vert. beam size at IP (σ_y^*)	1.1		μm
Bunch spacing (s_B)	2.1		m
Horizontal Emittance (ε_x)	1.5×10^{-8}	2.4×10^{-10}	m
Synchrotron tune (ν_s)	-0.024		
Betatron tune (ν_x/ν_y)	$45.51/43.56$	$44.51/41.60$	
Momentum compaction factor (α_p)	3.17×10^{-4}	3.38×10^{-4}	
Beam lifetime	94@1605	158@934	min.@mA

2.3 The Belle Detector

2.3.1 Overview

The Belle detector is a large-solid-angle asymmetric magnetic spectrometer. It is constructed of concentric layers of sub-detectors designed to provide momentum and position information via magnetic spectroscopy, energy measurements via electromagnetic calorimetry, and particle identification (PID) discrimination through energy loss and penetration depth data. Figure 2.6 is a cut-away diagram of the detector showing all 7 sub-detectors, the solenoid which provides a 1.5 T magnetic field, and the electron and positron beam-lines. The seven sub-detectors from the inner-most outwards are:

- Silicon Vertex Detector (*SVD*)
- Central Drift Chamber (*CDC*)
- Aerogel Cerenkov Counter (*ACC*)
- Time of Flight scintillator (*TOF*)
- Electromagnetic Calorimeter (*ECL*)
- K_L^0 and μ Detector (*KLM*)
- Extreme Forward Calorimeter (*EFC*)

Figure 2.6 also shows the standard Belle co-ordinate system. The x-axis is defined as outwards to the KEKB ring, the y-axis as vertical, and the z-axis as the opposite direction to the positron beam-line. Cylindrical co-ordinates are defined as follows: the polar angle (θ) runs from 0° to 180° along the z-axis; the azimuthal angle (ϕ) from 0° to 360° in the x-axis; and the radial distance $r = \sqrt{x^2 + y^2}$. The overall structure of the Belle detector is asymmetric in the positive z direction due to the energy asymmetry of the collisions. This is visible in the side-on view of the detector in Figure 2.7. The detector is divided into three sections: the barrel, which lies parallel to the positron beam axis, and the two end-caps which extend radially outwards from the beam line at either end of the barrel. The barrel lies from 34° to 127°, the forward end-cap from 17° to 34° and the backward end-cap from 127° to 150°.

The components of the Belle detector are described in detail in the following sections.

2.3.2 Beam Pipe

The two beam pipes of the HER and LER which maintain the vacuum through which the beams must travel merge at the IP. This section of the beam pipe is the inner-most piece of the detector and all particles must traverse its walls to reach the Belle detector proper. The material in the pipe must be kept to a minimum to avoid Coulomb scattering which degrades the resolution of the *SVD* and must also be able to withstand

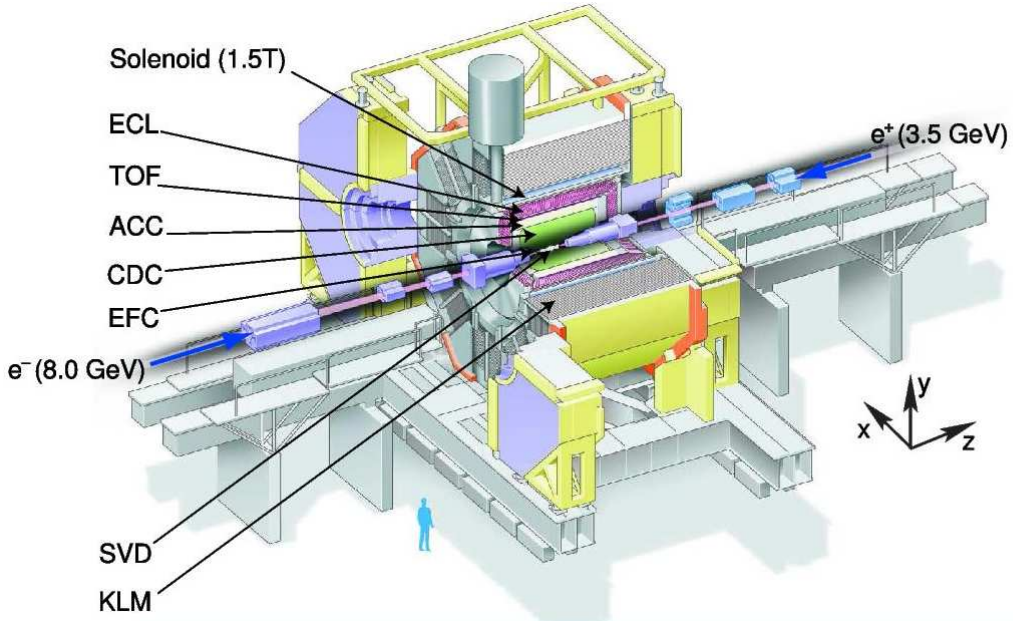


Figure 2.6: A cut-away diagram of the Belle detector.

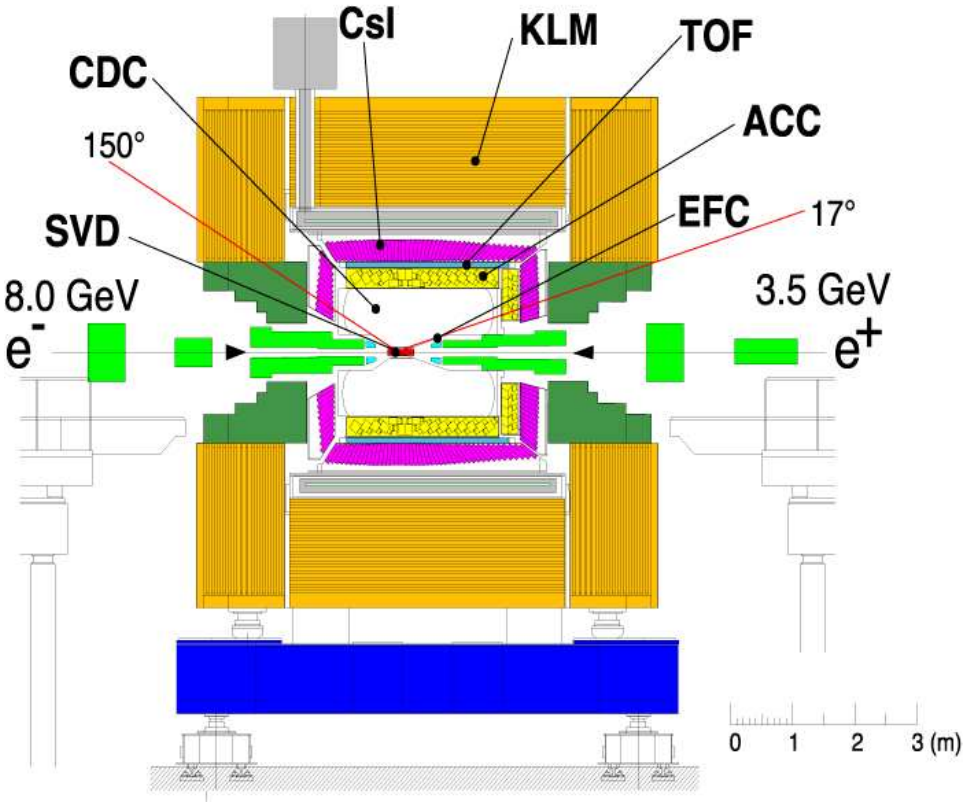


Figure 2.7: A side-on schematic of the Belle detector.

the beam-induced heating which can be up to several hundred watts. These attributes are supplied by the double-walled beryllium cylinder shown in Figure 2.8. The two 0.5 mm thick walls are separated by a 2.5 mm gap through which helium gas is circulated as a coolant. The beryllium is covered in a $20\,\mu$m thick layer of gold foil, which absorbs low energy X-rays from synchrotron radiation.

2.3.3 Silicon Vertex Detector

The closest detector to the IP is the *SVD*, a silicon strip based charged particle detector. The *SVD* is designed to provide high resolution position information for the reconstruction of the B meson decay vertices, essential for the study of time-dependant *CP* violation.

The original *SVD*1 was upgraded in October 2003 to the *SVD*2. Both are constructed of double-sided silicon strip detectors (DSSDs). *SVD*1 had 3 layers at 30 mm, 45.5 mm and 60.5 mm with a total of 102 DSSDs. *SVD*2 has four layers at 20 mm, 43.5 mm, 70 mm and 88 mm with a total of 138 DSSDs. *SVD*1 has an active area with range $23^\circ < \theta < 139^\circ$, while *SVD*2 extends to $17^\circ < \theta < 150^\circ$. The upgrade to *SVD*2 also provided higher vertex resolution, better low-momentum particle tracking and greater radiation hardness. Figure 2.9 shows two diagrams of *SVD*1 and a close-up of one of the DSSD ladders. Figure 2.10 shows the configuration of *SVD*2.

A DSSD is a single $300\,\mu$m thick silicon wafer. Both surfaces of the silicon are etched with a total of 1280 sense strips with 640 read-out pads. One side is p-doped silicon with p^+ read-out strips that are parallel to the beam line. These strips have a pitch of $24\,\mu$m and measure the particle position in ϕ. The other side has n^+ read-out strips perpendicular to the beam at a pitch of $42\,\mu$m to measure the particle z-position. When the silicon is placed under reverse bias, a charged particle traversing the depleted region will liberate electron-hole pairs which drift to the appropriately biased side of the DSSD where their charge is read by the read-out pads. This is demonstrated graphically in Figure 2.11. Using several layers of DSSDs mean θ can also be measured.

2.3.4 Central Drift Chamber

The next sub-detector from the IP is the *CDC*. The *CDC* is the backbone of the charged particle detection system. It has three purposes; to reconstruct the trajectory of charged particles travelling through the detector system in three dimensions (called 'tracks'), to precisely measure such particle's momenta, and to register the energy loss per distance travelled ($\mathrm{d}E/\mathrm{d}x$). The *CDC* is important to the analysis of $B \to K\eta'\gamma$ as the reconstructed decay modes (detailed further in Chapter 3) include at least three charged particles, and up to six. The momentum and trajectory resolution of these particles will have a significant impact on the ability to accurately reconstruct the decaying B meson.

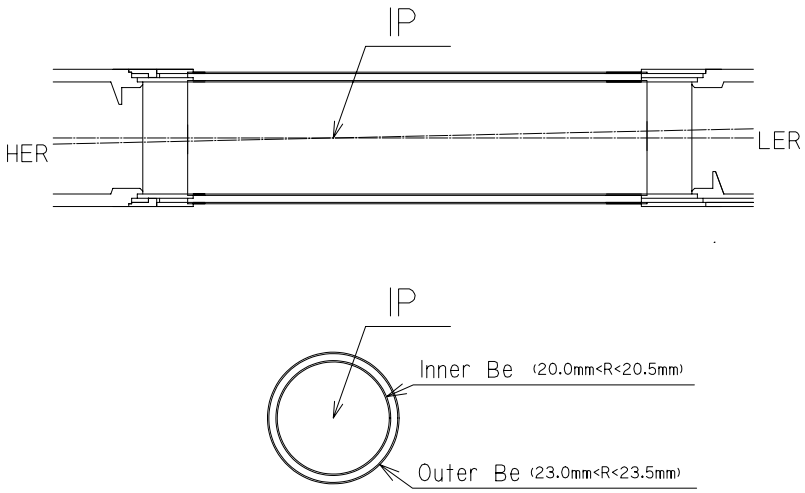


Figure 2.8: Cross section and side-view of the Belle beam pipe.

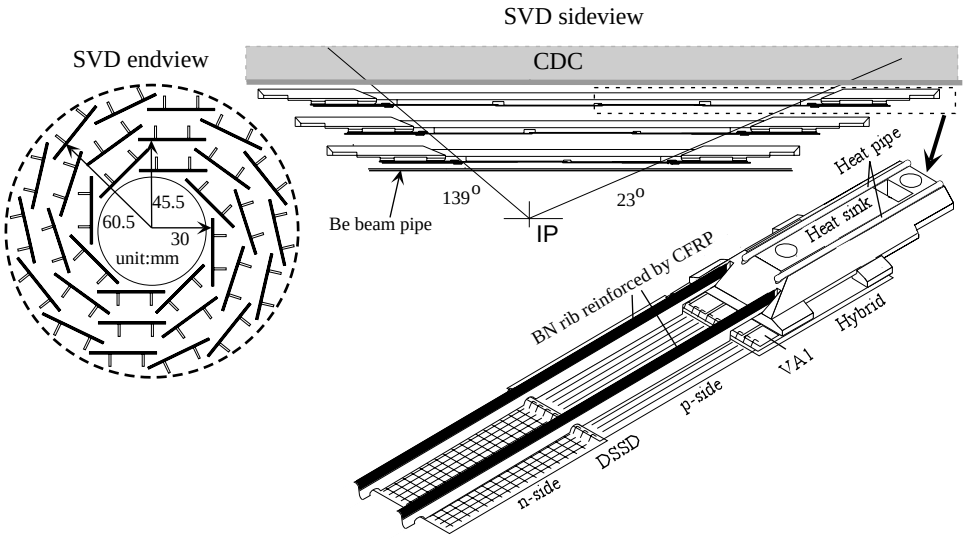


Figure 2.9: End and side views of the *SVD*1 confi guration and a close-up of one of the DSSD ladders.

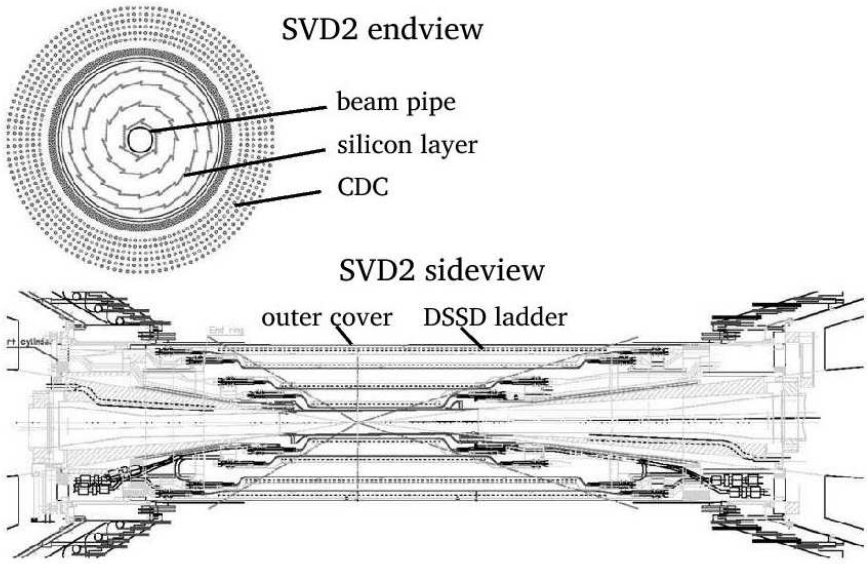


Figure 2.10: End and side views of the *SVD*2 confi guration.

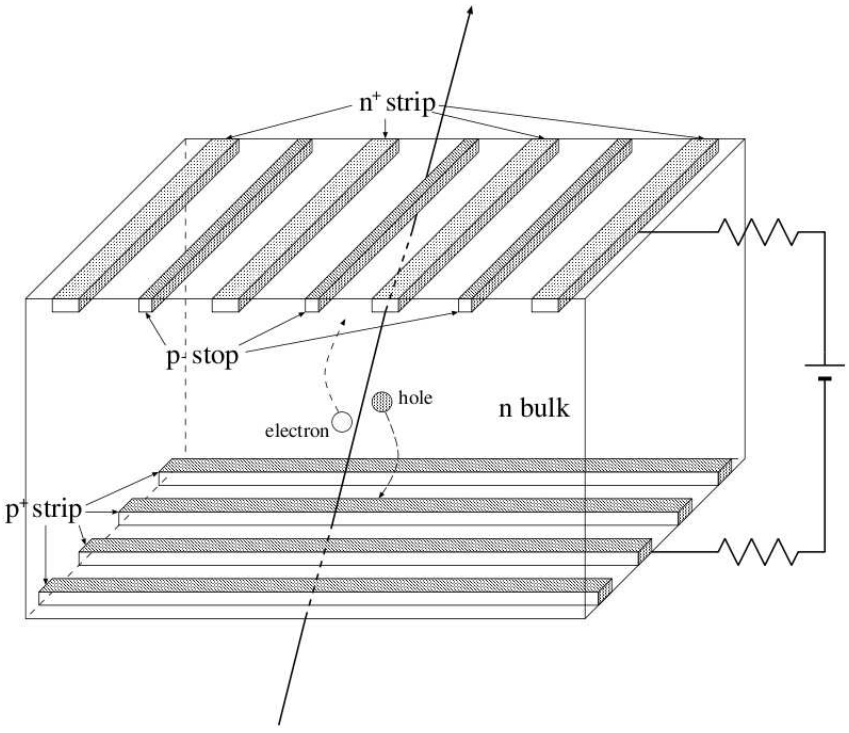


Figure 2.11: Diagram of a DSSD.

The *CDC* is a gas ionisation drift chamber. A one-to-one helium/ethane mixture is ionised by the passage of charged particles and the liberated electrons are accelerated by a 2.4 kV electric field towards the anode or 'sense' wires. These wires then record the deposited charge, or 'hit'. The *CDC* contains 8400 sense wires, each surrounded by 6 negatively biased 'field' wires to form a *CDC* 'cell'. There are 50 cylindrical layers of cells and three cathode-strip layers. About half of the cells are parallel to the z-axis (axial cells) and half placed at a small angle (small-angle-stereo cells) to provide 3-dimensional track reconstruction; the cathode strip layers further improve z resolution. The structure of the *CDC* chamber and of the individual cells is shown in Figure 2.12. The helium/ethane mix is chosen as a low-Z medium to suppress Coulomb scattering while still providing good $\mathrm{d}E/\mathrm{d}x$ resolution via the ethane component.

The *CDC* is within the 1.5 T magnetic field supplied by the Belle solenoid (described further in section 2.3.9). Charged particles traversing the detector feel a force transverse to the field and proportional to their momentum and describe helical paths. The *CDC* measures the five independent parameters that define a helix: the positive or negative curvature; the slope; and the closest approach of the helix to the IP in three dimensions, known as the impact parameter. The curvature is proportional to the particle's transverse momentum, the slope is proportional to its momentum in the z-direction, and the sign of the curvature gives the particle charge. The impact parameter can be extrapolated from the curvature of the helix, however it is provided with greater accuracy by the *SVD*; the *SVD* tracking information is combined with the *CDC* to improve particle trajectory resolution. The combined detector output gives the following performance in momentum and trajectory resolution:

$$\frac{\sigma_{p_t}}{p_t} = (0.19 p_t \oplus 0.34)\% \tag{2.4}$$

$$\sigma_{xy} = \left(19 \oplus \frac{50}{p\beta \sin^{3/2}\theta}\right)\ \mu\mathrm{m} \tag{2.5}$$

$$\sigma_z = \left(36 \oplus \frac{42}{p\beta \sin^{5/2}\theta}\right)\ \mu\mathrm{m} \tag{2.6}$$

where p is in $\mathrm{GeV/c}$ and $\oplus$ denotes a sum in quadrature.

The mean of the pulse heights recorded by the sense wires gives a measure of the $\mathrm{d}E/\mathrm{d}x$ of a track. The ionising energy loss is described by the Bethe-Bloch formula. For a relativistic particle in a known material, the $\mathrm{d}E/\mathrm{d}x$ depends mainly on the particle's velocity; $\beta\gamma = \frac{p}{Mc}$. Thus for a set momentum, the $\mathrm{d}E/\mathrm{d}x$ of a track depends on the mass of the particle. In this way particle species can be differentiated. Figure 2.13 shows $\mathrm{d}E/\mathrm{d}x$ as a function of momentum for charged pions and kaons, protons and electrons. The *CDC* provides a three standard deviation (σ) separation of pions and kaons with momenta up to 0.8 $\mathrm{GeV/c}$ and 2σ for momenta above 2.0 $\mathrm{GeV/c}$. Between 0.8 $\mathrm{GeV/c}$ and 2.0 $\mathrm{GeV/c}$ the $\mathrm{d}E/\mathrm{d}x$ profiles of pions and kaons overlap too much for useful differentiation.

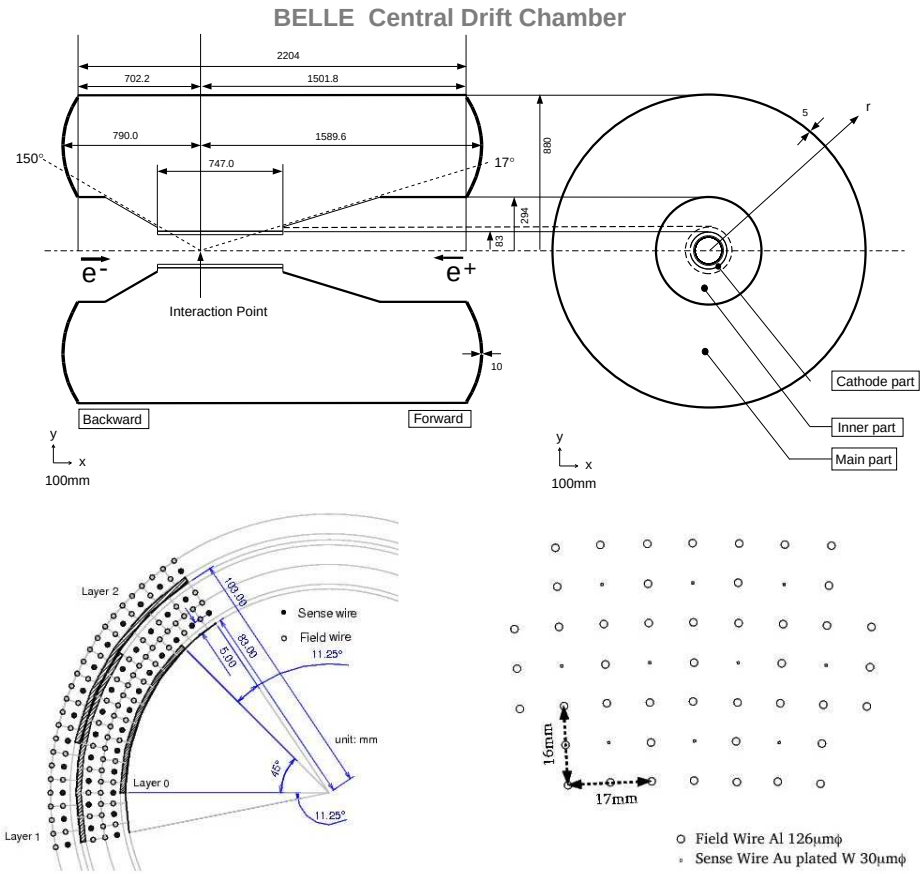


Figure 2.12: The top images show the general structure of the *CDC* from the side and end-on perspective, and the lower show the cell structure and wiring of the chamber.

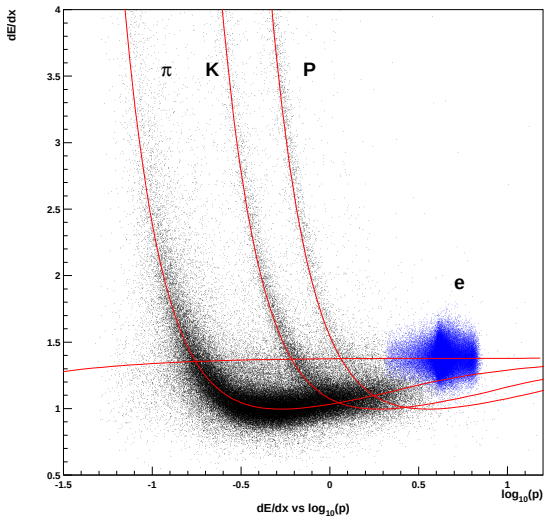


Figure 2.13: dE/dx vs. momentum for different particle species as measured by the *CDC* in collision data.

2.3.5 Aerogel Čerenkov Counter

The identification of particle species, or PID, is very important in the Belle experiment. B meson decays produce copious numbers of charged kaons and pions and obtaining good differentiation between the two is essential for accurate reconstruction. The *ACC* is solely a PID detector, designed to separate kaons and pions with momentum in the region at which $\mathrm{d}E/\mathrm{d}x$ and the time-of-flight measurements of the *TOF* (described in the next section) have little power.

The *ACC* takes advantage of the fact that a particle traversing a material at velocity greater than the speed of light in that medium produces an electromagnetic shock-wave in the form of Čerenkov radiation:

$$v > v_t (= c/n) \tag{2.7}$$

where v_t is the threshold velocity equal to the speed of light divided by the refractive index of the medium (n) and v is the particle velocity. This light is emitted in a conical wave-front with a velocity-dependant angle of emission.

The *ACC* is constructed of silica aerogel with refractive indices chosen so that pions with momentum in the range 1-4 $\mathrm{GeV/c}$ will be above v_t, while kaons in the same momentum range, being around three times heavier than pions, will be below v_t. Aerogel is a colloidal form of $Si0_2$ with unique properties. It has greater than 95% porosity, making it very light, and is of extremely high optical clarity. 960 blocks of aerogel in 0.2 mm thick aluminium boxes are arrayed in the barrel region of the detector, and 228 aerogel blocks housed in carbon-fibre reinforced plastic (CFRP) are in the end-caps. Figure 2.15 shows the design of the individual *ACC* modules for the barrel and end-caps. The specific refractive index of the aerogel block depends on its position in the array, and range from $n = 1.01$ to $n = 1.03$, as shown in Figure 2.14. The Čerenkov radiation is read by fine-mesh photo-multiplier tubes (FM-PMT), one for each block in the end-caps and two for each in the barrel, which convert it to an electronic signal. Figure 2.16 shows the pulse height response of the *ACC* to kaon candidates selected using *CDC* and *TOF* information compared to that of electrons and positrons from Bhabha ($e^+e^- \to e^+e^-$) scattering events.

2.3.6 Time of Flight Detector

The *TOF* detector provides further PID information via an independent measurement of a particle's velocity. For a given momentum a higher mass particle will take a longer time to reach the *TOF* after the interaction at the IP. The particle mass can then be determined from the *TOF* timing information using the relation

$$M = p \cdot \sqrt{(\frac{cT}{D})^2 - 1} \tag{2.8}$$

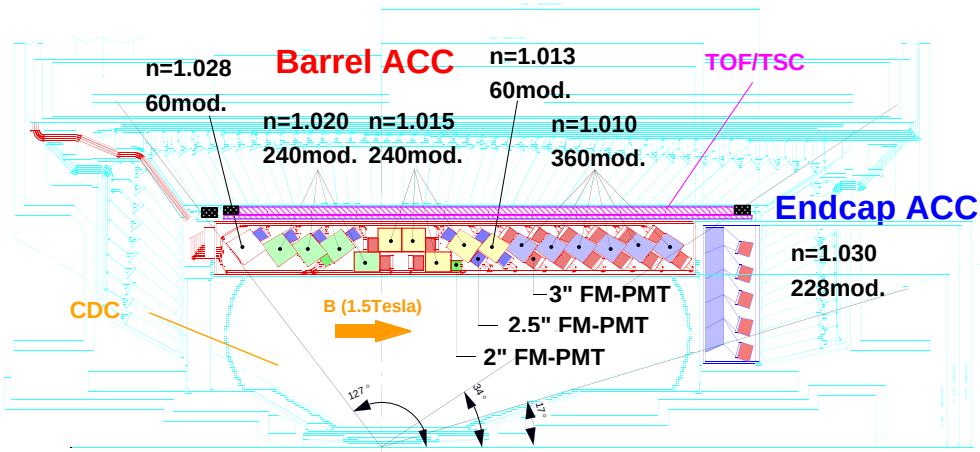
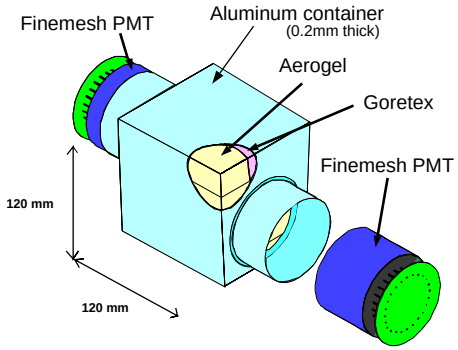


Figure 2.14: Schematic of the *ACC* design.

a) Barrel ACC Module



b) Endcap ACC Module

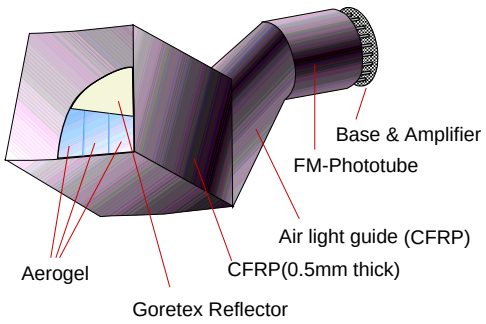


Figure 2.15: *ACC* modules for the barrel region (top) and the end-caps (bottom).

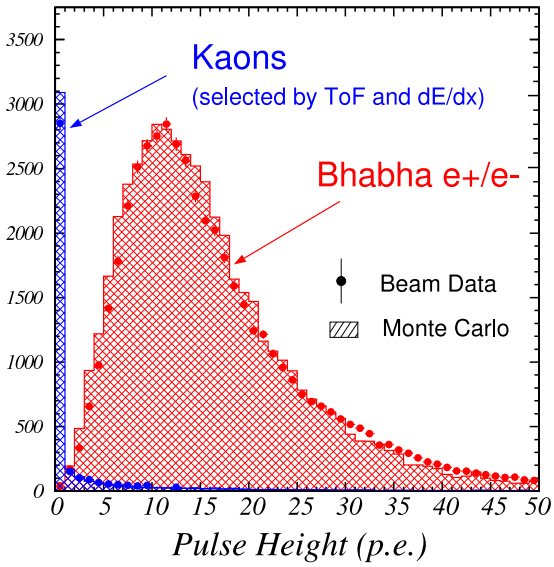


Figure 2.16: *ACC* pulse height spectra for kaon and e^+/e^- candidates. The points are beam data, the shaded histograms Monte Carlo simulation.

where M is the mass and p the momentum of the particle, T is the time from the collision and D the distance travelled.

The *TOF* is situated beyond the *ACC* at a radius of 1.2 meters and is constructed of 4 cm thick plastic scintillators in conjunction with two FM-PMTs mounted on both ends. The *TOF* also has 0.5 cm thick trigger scintillator counters (TSCs) coupled to single FM-PMTs that provide fast information to the trigger system (described in detail in section 2.4). Two *TOF* scintillators and one TSC make up one module of the *TOF* detector, as shown in Figure 2.18. 64 modules placed around the IP measure the time from the collision to the particle hitting one of the scintillators in the region $33° < \theta < 121°$.

The *TOF* has a timing resolution of 100 ps. The timing information is most useful for particles with momentum lower than 1.2 GeV/c. Figure 2.19 shows a histogram of particle mass as calculated using equation 2.8 with clear peaks for kaons, pions and protons. Also shown is the K^+/π^+ separation as a function of momentum in units of σ. The particle momentum regions in which each of the PID detectors are compared in Figure 2.17.

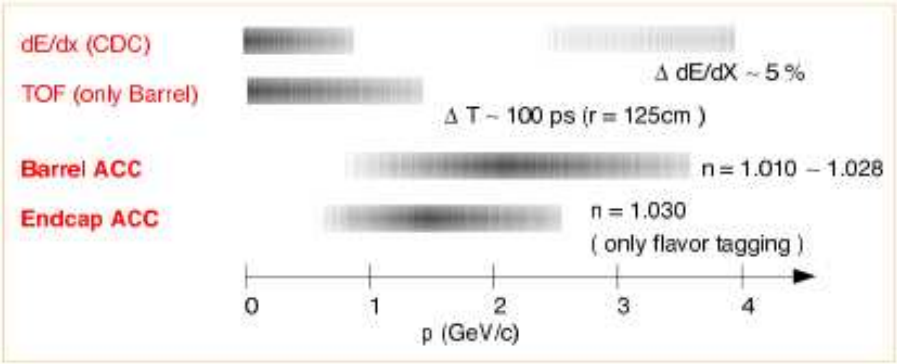


Figure 2.17: The momentum regions in which each PID detector is most useful.

2.3.7 Electromagnetic Calorimeter

The *ECL* is designed to detect photons, electrons and positrons and measure their energies and trajectories. This is very important for radiative decay analysis, since a large portion of the B meson energy will be carried by the primary decay photon, or 'signal' photon. The detector must be sensitive to a large range of photon energies, from the very high energy photons produced in $e^+e^- \to \gamma\gamma$ anihillation and radiative Bhabha ($e^+e^- \to e^+e^-\gamma$) events as well as radiative B decay (up to $\sim$ 4 GeV), down to the decay photons from low-momentum $\pi^0 \to \gamma\gamma$ decays (down to 20 MeV), which are important to a lot of B physics at Belle. Efficient tagging and reconstruction of electrons and positrons is also essential for the flavour-tagging of neutral Bs (discussed further in section 2.3.6) and the analysis of semi-leptonic B decay.

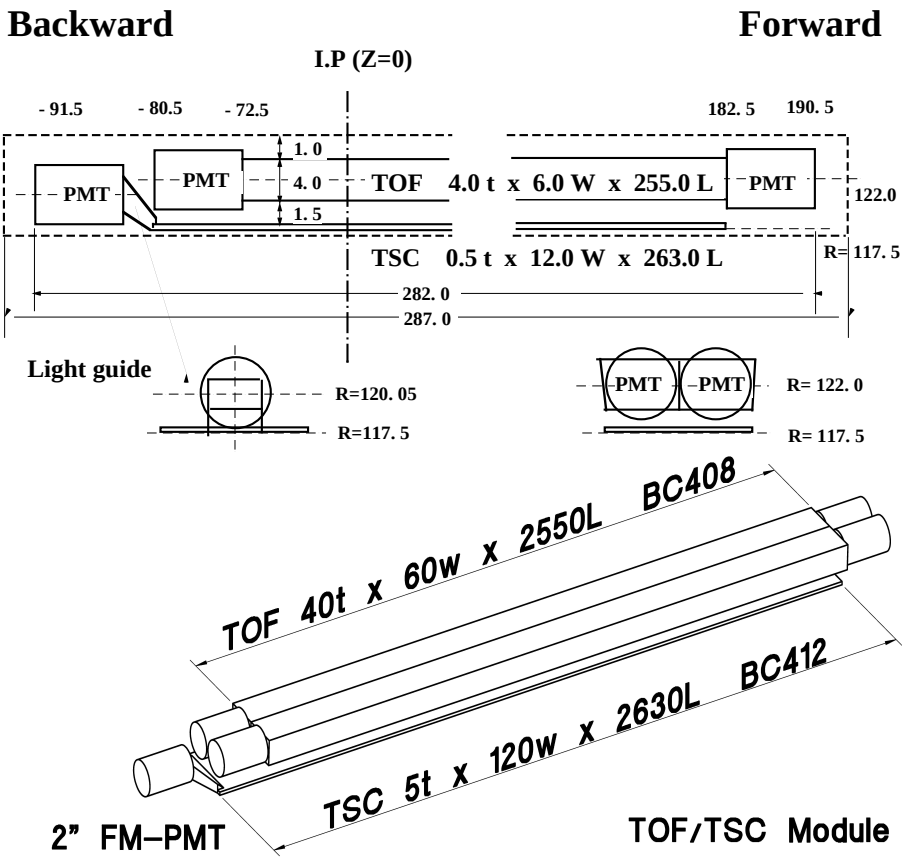


Figure 2.18: Schematics of the *TOF* design.

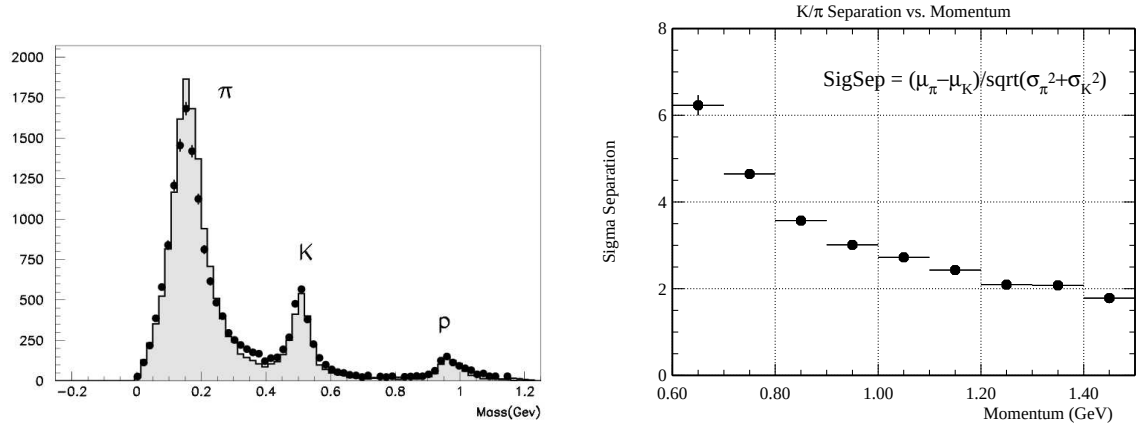


Figure 2.19: Particle mass as calculated from the *TOF* timing and *CDC* momentum information (left), and the kaon/pion separation provided by the detectors as a function of momentum (right).

The *ECL* is constructed of 8736 Thallium-doped Cesium Iodide scintillating crystal towers. The crystals are typically 30 cm in length, equivalent to 16.2 radiation lengths for photons and electrons. They have inner faces $5.5\ \mathrm{cm}^2 \times 5.5\ \mathrm{cm}^2$ in area; this is a compromise between a highly segmented array for precision spacial information and the energy resolution of a single crystal detector. Incident photons pair-produce via interaction with crystal nuclei. The subsequent electron and positron radiate bremsstrahlung photons which then also pair-produce, inducing an electro-magnetic particle shower with the crystal. Coulomb scattering creates a lateral shower spread. The shower proceeds to create more particles until eventually all the energy is in the form of ionisation or excitation photons, which are read out by a pair of silicon PIN photo-diodes coupled to the rear of every crystal. This is a destructive detection process; the 16.2 radiation lengths mean that the initiating particle deposits almost all of its energy within the *ECL*. The towers all point to face the IP to minimise the possibility of photons being lost in the dead material between the crystals. The *ECL* barrel contains 6624 crystals in the region $32.3^\circ < \theta < 128.7^\circ$, the forward end-cap has 1152 crystals in $12.4^\circ < \theta < 31.4^\circ$, and the backward end-cap 960 and covers $130.7^\circ < \theta < 155.1^\circ$. The *ECL* structure is shown in Figure 2.21.

Other charged particles will also deposit some energy within the *ECL* via ionisation. Only photons, electron and positrons will interact strongly with the detector material and initiate large showers. The energy deposited in the *ECL* by $1\,\mathrm{GeV}$ electrons and charged pions is compared in Figure 2.22. To differentiate photons and electrons, whose shower shapes will be very similar, the showers are matched with tracks in the *CDC*. The *ECL* energy resolution has been measured to be

$$\frac{\sigma_E}{E} = \frac{0.066\%}{E} \oplus \frac{0.81\%}{E^{1/4}} \oplus 1.34\% \tag{2.9}$$

where the first term describes electronic noise in the photo-diodes, the second term the shower leakage fluctuation, and the third the systematic uncertainties. The position resolution is measured to be

$$\sigma_{pos} = \left(0.27 + \frac{3.4}{E^{1/2}} + \frac{1.8}{E^{1/4}}\right) \mathrm{mm} \tag{2.10}$$

Both the energy and position resolutions increase with higher particle energy, which is important for the hard gamma rays studied in this analysis.

2.3.8 Extreme Forward Calorimeter

Belle also has a calorimeter in the extreme forward and backward sections of the detector; the *EFC*. The two segments of the *EFC* cover the regions $6.4^\circ < \theta < 11.5^\circ$ and $163.3^\circ < \theta < 171.2^\circ$ and are constructed from 320 crystal scintillators of radiation-hard Bismuth Germinate (Bi_4Ge_3O-12) coupled to photo-diodes. The *EFC* requires greater radiation hardness than the *ECL* in order to withstand the high flux of energetic particles at small angles near the beam pipe. The *EFC* has lower energy resolution than the *ECL* (7-10% at particle energies of 1 to $3\,\mathrm{GeV}$) and cannot distinguish between electrons and photons

as the *CDC* does not cover the same regions. As such it is not generally used in physics analysis and its main roles are as a beam monitor and an on-line luminosity monitor. The configuration of the crystals in the *EFC* is shown in Figure 2.20.

2.3.9 Solenoid

The Belle NbTi/Cu super-conducting solenoid magnet provides a 1.5 T field to all sub-detectors except the *KLM*. The field is not required for calorimetry, but the solenoid is placed beyond the *ECL* to reduce the material between it and the IP. The iron structure of the *KLM* acts as a return path for the magnetic flux. The structure of the solenoid is shown in Figure 2.23, and the results of a mapping of the magnetic field within the solenoid is shown in Figure 2.24.

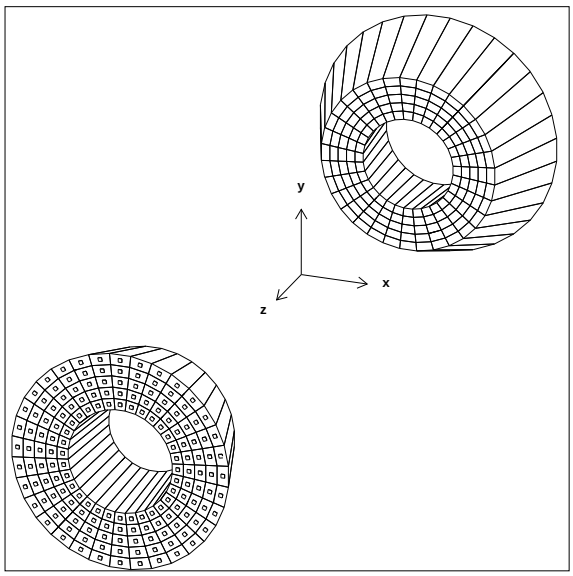


Figure 2.20: Schematic of the *EFC* structure.

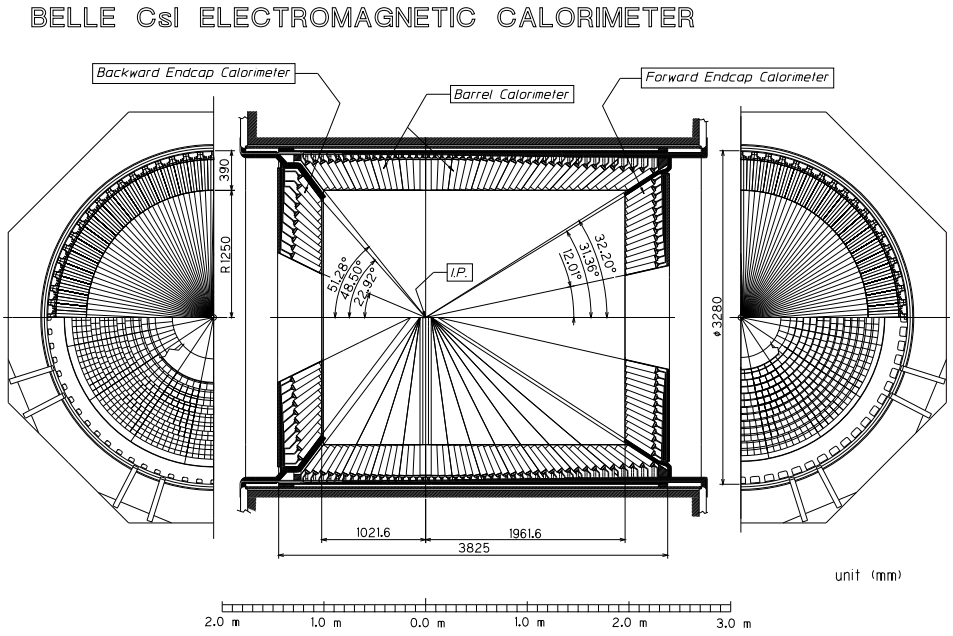


Figure 2.21: *ECL* structure in the barrel and end-caps.

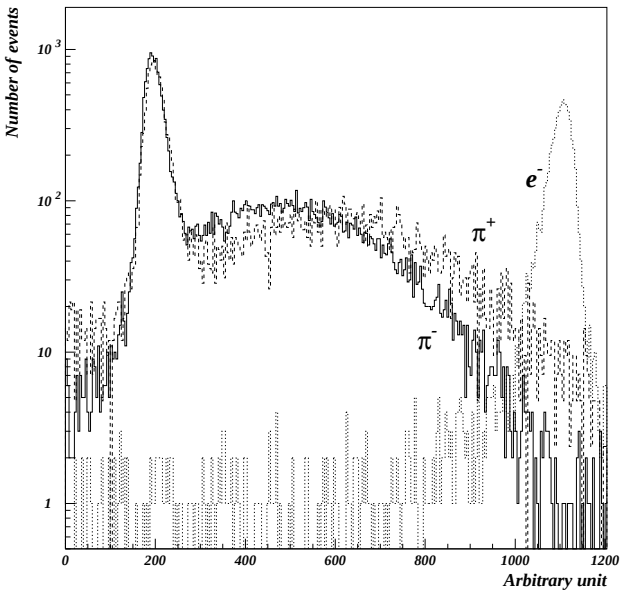


Figure 2.22: The energy deposited in the *ECL* by 1 GeV electrons and charged pions.

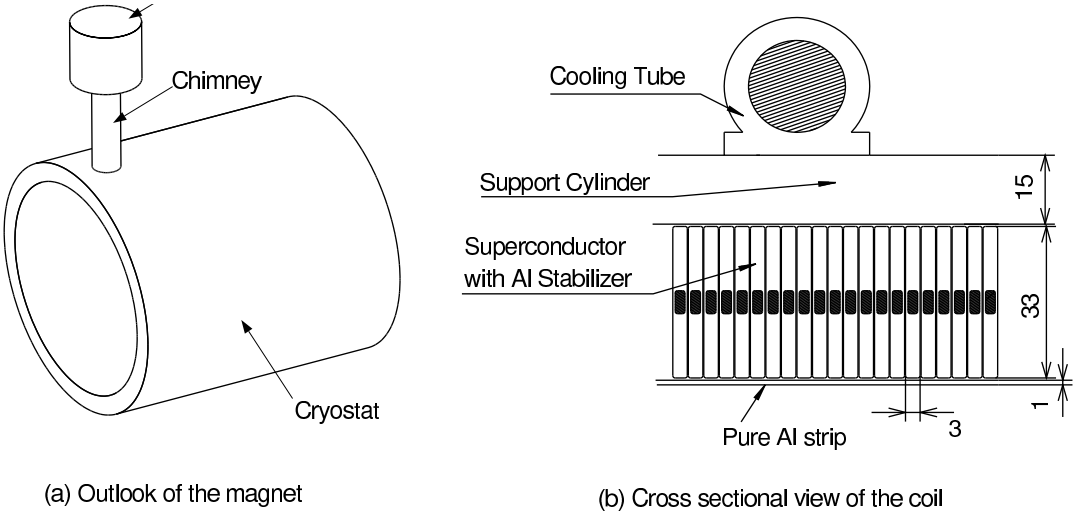


Figure 2.23: Schematic of the main Belle solenoid structure (a), and the coil cross-section (b).

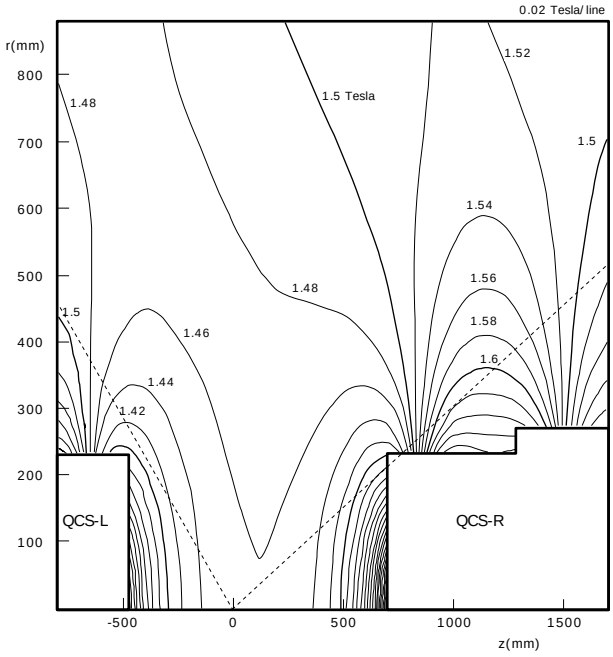


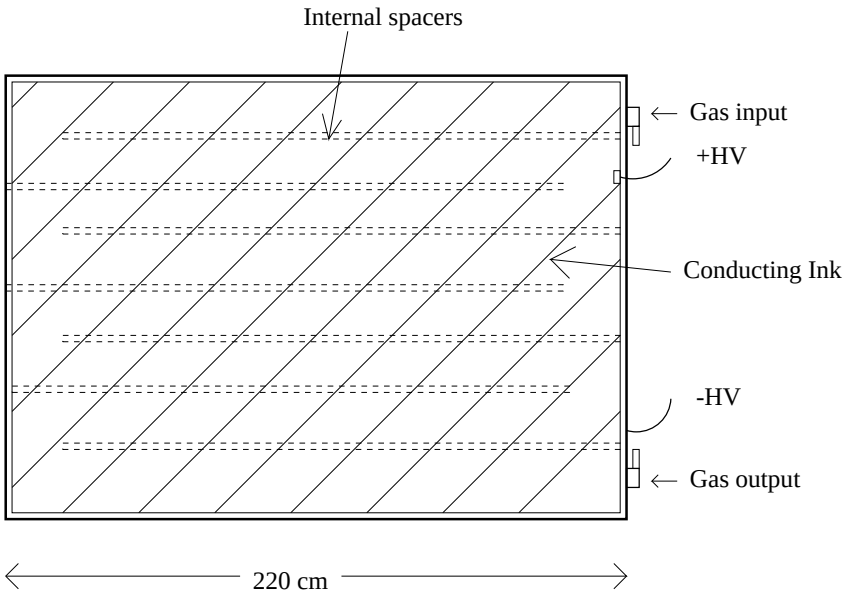
Figure 2.24: A contour plot of the magnetic field within the Belle solenoid. The mapping was performed using the QCS-L and QCS-R magnets shown in the plot. They are the final focusing quadrople magnets and are located within the solinoid.

2.3.10 K_L^0/μ Detector

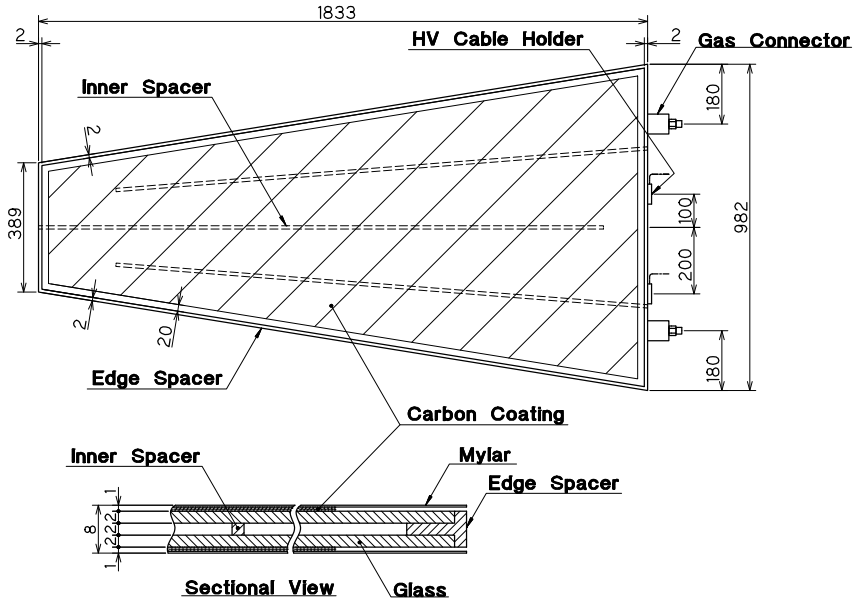
The K_L^0/μ detector is designed to detect K_L^0 mesons and identify muons. It is constructed from alternating 4.7 cm thick iron plates and 3.7 cm thick active *KLM* detector plates. The iron provides most of the 3.9 radiation lengths seen by K_L^0 mesons, while the detector plates register the passage of ionising particles. The detector plates consist of two glass-electrode resistive-plate counters (RPCs) sandwiched between layers of read-out strips in the θ and ϕ directions. An RPC has an active gaseous region between two highly resistive glass parallel plate electrodes. Charged particles ionise a streamer in the gas which results in a local discharge of the resistive plates, inducing a signal in the read-out strips. This structure is shown in Figure 2.26.

K_L^0 particles live long enough to travel beyond the *ECL* and interact primarily via the strong force. They are detected by the hadronic showers of ionising particles they induce. The detector up until the *KLM* is about a radiation length for K_L^0 mesons, mostly from the *ECL*. Showers initiated in the *ECL* will continue into the *KLM*; K_L^0s will deposit most of their energy within the iron of the *KLM* proper. The detector provides position information for the K_L^0 but no useful energy information is gained as a significant proportion of the shower will generally not be within the *KLM*.

Muons on the other hand will not interact via the strong force but do have an electro-magnetic cross-section; they will lose energy mostly through ionisation processes. They penetrate the *ECL* easily and will continue through most of all of the *KLM*. Deeply penetrating *KLM* tracks that are able to be matched with a track in the *CDC* are identified as muons.



(a) Barrel RPC.



(b) End-cap RPC.

Figure 2.25: The *KLM* detector plate schematics and cross-sectional diagram.

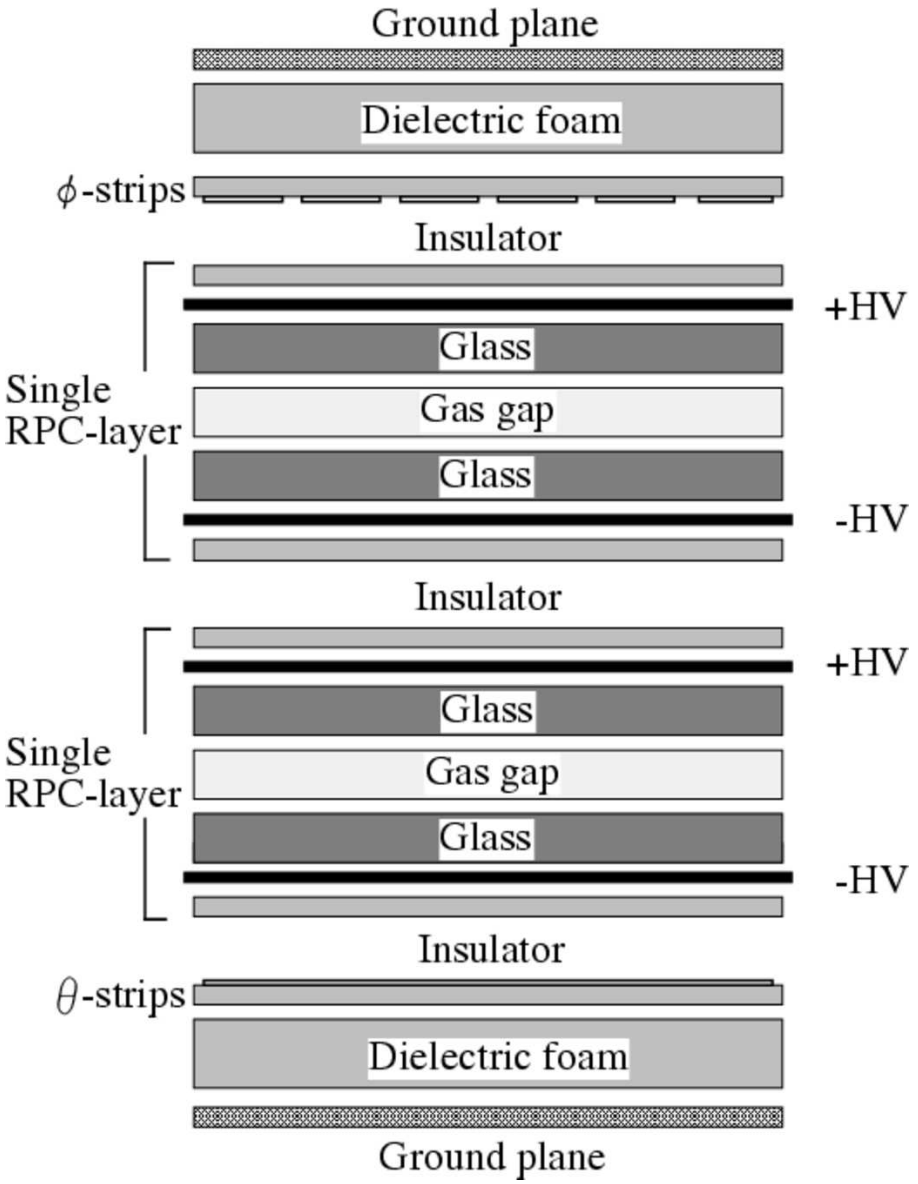


Figure 2.26: *KLM* detector cross-section.

2.4 Trigger, Data Acquisition and Event Processing

The trigger system is a multi-tiered hard- and soft-ware based selection process by which the decision whether or not to record the output signals of the entire detector system is made. This is performed on an event-by-event bases, where 'event' is used interchangeably to mean an e^+e^- collision or the stored detector response recording that collision. Physically interesting events include hadronic interactions ($e^+e^- \rightarrow q\overline{q}$ where $q\overline{q}$ =$u\overline{u}$,$d\overline{d}$,$s\overline{s}$,$c\overline{c}$ or $b\overline{b}$), leptonic events ($e^+e^- \rightarrow l^+l^-$, where l^+l^-=e^+e^-, μ^+ μ^- or τ^+ τ^-), and photon-pair events ($e^+e^- \rightarrow \gamma\gamma$). The total rate of desirable physical events is around 100 Hz at a luminosity of $1 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The major backgrounds to these processes are from beam related backgrounds, including beam interaction with residual gas in the beam pipe and with the beam pipe wall itself, as well as other sources such as synchrotron radiation and cosmic rays. The background rate is expected to be around 120 Hz, taking the total rate to around 220 Hz. The beam background rate is highly dependant on the beam current and other accelerator conditions however, and accurate estimation is difficult. The processes and rates are listed in Table 2.4. The trigger's role is to suppress the backgrounds while preserving as many interesting events as possible. With the increases in luminosity over the Belle operation, the trigger system now runs at around 500 Hz and is designed to be capable of dealing with rates of up to 1300 Hz with occupancies of 5%. Low occupancy levels are essential for distinguishing the incoming event data.

The trigger is divided into four levels: the hardware level 0 (L0) and level 1 (L1) triggers, and the software level 3 (L3) and level 4 (L4) triggers. Following the hardware triggers the Belle data acquisition system (DAQ) translates the sub-detector outputs to computer readable form for off-line interpretation by the software triggers. Events that pass all levels of the trigger are then recorded in data summary tape (DST) form. Figure 2.27 shows the trigger/DAQ/data-recording system graphically.

Table 2.4: The expected rates of physical processes at a luminosity of $1 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ at the $\Upsilon(4S)$ resonance.

Process	Rate (Hz)
$\Upsilon(4S) \rightarrow B\overline{B}$	12
Hadronic Continuum	28
$e^+e^- \rightarrow \mu^+\mu^-$,$\tau^+$ τ^-	16
$e^+e^- \rightarrow e^+e^-$	4.4
$e^+e^- \rightarrow \gamma\gamma$ (prescaled by 1/100)	0.4
$\gamma\gamma \rightarrow$anything ($p_t > 0.3\,\mathrm{GeV}/c$)	35
Beam Background	$\mathcal{O}(100)$
Cosmic Rays	20
Total	220

2.4.1 Level 0 Trigger

The first hardware trigger consists of a signal sent from the *TOF* TSC scintillators to the *SVD* causing the *SVD* to enter a 'hold' state in which the recorded hits are stored briefly. This is required because the circuit shaping time of the *SVD* (1 μs) is shorter than the level 1 trigger latency (2.2 μs). The short *SVD* shaping time reduces high occupancies caused by beam backgrounds.

2.4.2 Level 1 Trigger

The level 1 hardware-based trigger is the main triggering system of the Belle detector. Information from each of the sub-detectors except the *SVD* and the *ACC* is fed into the the global decision logic (GDL) system, as shown in Figure 2.28. The *CDC* provides charged particle tracking and multiplicity information. The *ECL* provides total deposited energy and cluster multiplicity information. The *KLM* provides a highly efficient muon trigger. The *EFC* triggers on Bhabha and two-photon events. The TSCs in the *TOF* provide the event timing signal and further charged particle multiplicity information. When timing information from the *TOF* is unavailable the *ECL* is used. The total triggering information is passed to the GDL within 1.85 μs of the collision and it issues a decision within 2.2 μs. If the GDL decides to keep the event, the information is passed to the level 3 trigger via the data acquisition system (DAQ) described below. The efficiency of this trigger is 99.5% for hadronic events.

A further level 1.5 hardware trigger was introduced with the upgrade to *SVD*2. It uses *SVD* vertexing information available 25.6 μs after the event to remove beam-gas interactions in which the event vertex is displaced from the IP. If an event is vetoed by this trigger it aborts the passing of the event information from the level 1 trigger to the software triggers, which takes around 50 μs. This removes background events earlier in the trigger chain, reducing dead time.

2.4.3 Level 3 Trigger

The level 3 software trigger runs on a Linux PC farm. It reconstructs the charged particle tracks using a fast track finding algorithm and rejects events having no track with z impact parameter less than 5 cm. The total energy deposited in the *ECL* is required to be greater than 3 GeV. This reduces the overall data rate by 50-60% while retaining 99% of interesting physics events.

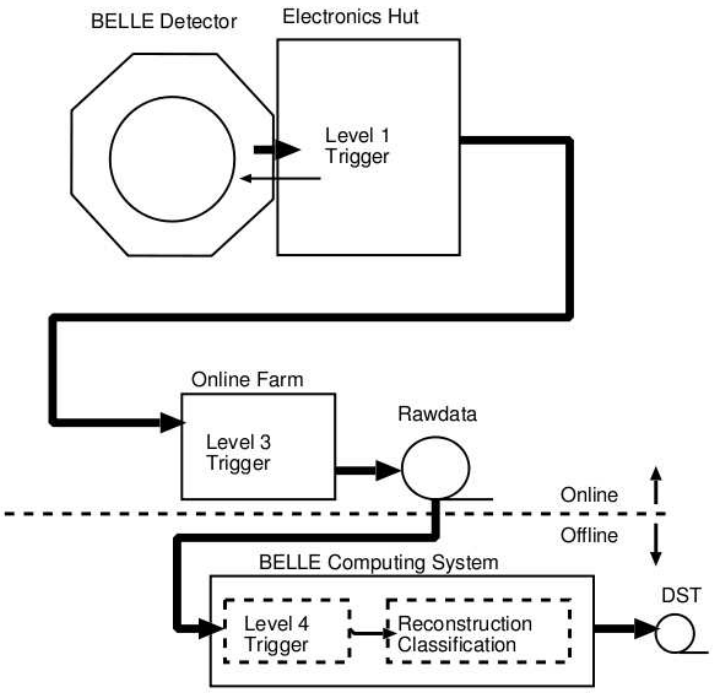


Figure 2.27: Overview of the Belle trigger system. The production of the DSTs has now been shifted to online.

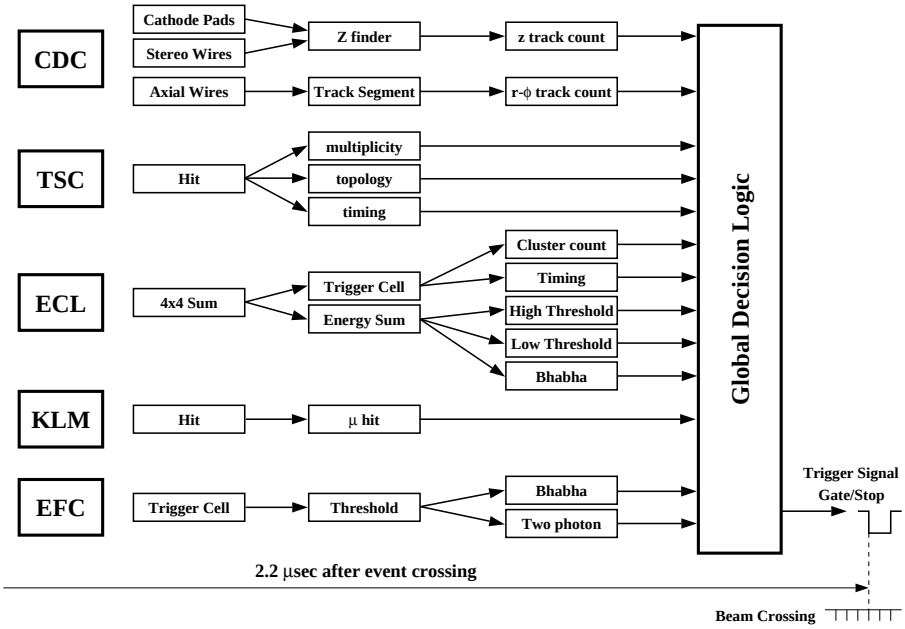


Figure 2.28: Level 1 hardware trigger system.

2.4.4 Level 4 Trigger

The level 4 software trigger removes events just before full reconstruction. This trigger is similar to the level 3 trigger, with tighter tracking requirements of at least one track of transverse momentum greater than 300 $\mathrm{MeV/c}$ and impact parameters less than 1 cm and less than 4 cm in the r and z dimensions, respectively. The total *ECL* energy must also be greater than 4 GeV. This selection retains 99.8% of hadronic events while removing 73% of the total trigger rate. Unlike the level 3 trigger, events rejected by the level 4 trigger are retained in their raw data form. This trigger reduces the time and CPU requirements of full event reconstruction rather than the rate and size of data storage.

2.4.5 Data Acquisition System

The DAQ is designed to process data passing the level 1 trigger into a form usable by the off-line PC farm while keeping a dead-time fraction of less than 10%. As shown in Figure 2.29, seven parts of the DAQ dedicated to each of the sub-detectors run in parallel. In all sub-detectors except the *KLM* and *SVD* the integrated charge pulse of a signal is proportional to the energy deposited in the sub-detector. The pulses are converted to timing information by Q-to-T modules and digitised by time digital converters (TDCs). No Q-to-T conversion is needed for the *KLM* as the pulse heights provide no useful energy information. The DSSDs on the *SVD* are read out by on-board chips and processed by flash-analogue to digital converters (Flash ADCs).

When the event passes the level 1 trigger and the sequence control receives a signal from the GDL, the seven parallel data processes are fed into the event builder, which combines the information into a single event record. This is passed to the PC farm where it is filtered by the level 3 trigger and written onto tapes as raw data. An average hadronic event occupies 30 kilobytes of storage. At a trigger rate of 500 Hz this translates to 15 megabytes of data storage per second.

2.4.6 Data Processing

Events that pass the level 3 trigger are stored as raw data for the lifetime of the experiment. Events that pass the level 4 trigger undergo full reconstruction, which converts it into higher-level data structures for ease of use in analysis. For example, the four-vectors of all reconstructed charged tracks, photon and electron candidates in the *ECL*, and muon candidates from the *KLM* are stored as such to be used in further reconstruction.

The fully reconstructed events are stored as DST files. These are then analysed further using individual sets of selection criteria called ‘skims’ to classify events into hadronic, Bhabha, τ-pair, μ-pair and two photon events. The hadronic sample is skimmed into many sub-categories of B decays. The resulting sub-data sets are saved into mini-DST (MDST) files which are used in high-end analysis. The MDST files

are accessed through the Belle analysis framework (BASF) [39], software written by the collaboration for the purpose.

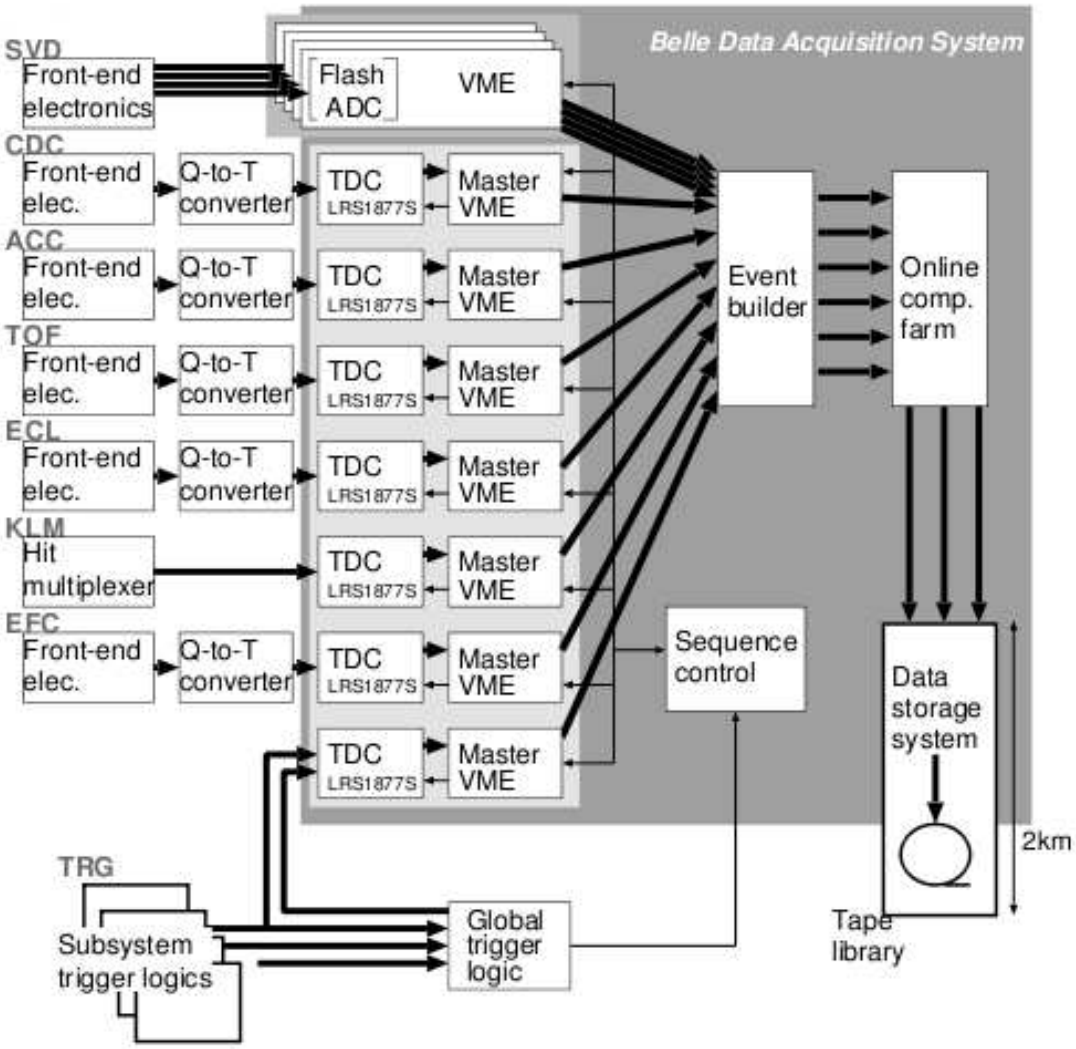


Figure 2.29: Overview of the Belle DAQ system.

2.5 Event Simulation

The response of the Belle detector to certain physical events needs to be well understood before the said events can be analysed. Interesting physical events are often very rare and submerged in overwhelming backgrounds. Simulating the physical properties and the detector response to both the desirable events and the backgrounds using Monte Carlo (MC) techniques enables the efficient separation of the two and the estimation of any uncertainties involved in the process.

MC simulated data is generated at Belle in a two-step process. The first simulates the under-lying particle physics involved in e^+e^- collisions and $\Upsilon(4S)$ decays. This is achieved using the EvtGen simulation program [40] which is a module of BASF. It contains the properties of all the relevant particles as measured by particle physics experiments around the world, such as branching fractions, charges, lifetimes and masses. Information on the collision environment such as the electron and positron energies and the crossing angle of the beams is also included. The particle decays are modelled, and the four-vectors of each particle calculated at the point of stability ($c\tau > 1\,\mathrm{cm}$).

In the second step the generated events are passed to the GSIM module which simulates the Belle detector. GSIM is based on the GEANT3 [41] package created at CERN. The particles generated by EvtGen are propagated and each sub-detector's response modelled. The simulation includes the particle's interactions with both the active and dead material in the detector. GSIM is also responsible for modelling the decay of longer lived particles such as K^0_S. Environmental backgrounds such as beam-gas interactions and cosmic rays are included by embedding randomly triggered real data events into each simulated event. GSIM is constantly updated with new detector properties as they evolve.

MC for continuum background events ($e^+e^- \rightarrow q\overline{q}$ where $q\overline{q}$ =$u\overline{u}$,$d\overline{d}$,$s\overline{s}$,$c\overline{c}$) are generated using JETSET [42] which is based on the Lund string fragmentation model [43]. Some of the MC generated in the large $B\overline{B}$ background sets (see section 3.1.2) were generated using the older program QQ98 rather than EvtGen. QQ98 was developed for use by the CLEO [44] collaboration and modified for use at Belle before EvtGen was adopted.

Chapter 3

Event Reconstruction

The next three chapters describe the experimental techniques used in the measurement of $\mathcal{BF}(B \to K\eta'\gamma)$. This chapter focuses on the reconstruction and selection of $B \to K\eta'\gamma$ events and the supression of backgrounds to the decay. Chapter 4 describes the modelling of $B \to K\eta'\gamma$ and background events, the optimisation of the modelling process and the extraction of the number of $B \to K\eta'\gamma$ events in the data sample after selection. Chapter 5 shows calculations of the efficiency of the selection criteria, the errors inherent in the reconstruction and modelling processes, and the final measured branching fraction of $B \to K\eta'\gamma$.

3.1 Data Set

3.1.1 Experimental data

The data set analysed in the search for $B \to K\eta'\gamma$ decays was recorded by the Belle detector over the period January 2000 to December 2006. It totals $604.55\ \mathrm{fb}^{-1}$ of integrated luminosity at the $\Upsilon(4S)$ resonance (on-resonance), and $68.27\ \mathrm{fb}^{-1}$ at a center of mass (CoM) energy $\sim 50\,\mathrm{MeV}$ below the $\Upsilon(4S)$ resonance (off-resonance). Off-resonance data is collected as a background estimation tool. The data was accumulated over 24 active periods, or "experiments", with detector and accelerator configurations varying from experiment to experiment. The convention at Belle is to use even numbered experiments for calibration and detector/accelerator studies, and odd numbered experiments as physics runs. Part of most Belle physics experiments is allocated to off-resonance data collection. The total integrated on- and off-resonance luminosity recorded for each experiment is detailed in Figure 3.1. To date over $850\,\mathrm{fb}^{-1}$ of integrated luminosity has been recorded by the detector however not all of this is available for analysis.

The number of $B\overline{B}$ events ($N_{B\overline{B}}$) within this data sample is calculated using the equation:

$$N_{B\overline{B}} = N_{\mathrm{on}} - \frac{\epsilon_{\mathrm{on}}}{\epsilon_{\mathrm{off}}}\frac{\mathcal{L}_{\mathrm{on}}}{\mathcal{L}_{\mathrm{off}}}N_{\mathrm{off}} \tag{3.1}$$

where N is the number of events recorded, $\mathcal{L}$ is the luminosity, ϵ is the HadronB skim efficiency, (see Section 3.2) and 'on' and 'off' refer to on-resonance and off-resonance respectively. $N_{B\overline{B}}$ for each experiment and the total for the data set analysed is shown in Figure 3.1.

Table 3.1: The on-resonance and off-resonance accumulated luminosity, and the number of $B\overline{B}$ pairs in each Belle experiment.

Experiment	On-res(fb^{-1})	Off-res(fb^{-1})	$\#B\overline{B}$ ($\times 10^6$)
7	5.928	0.594	$6.4587^{+0.1615}_{-0.0976}$
9	4.440	0	$4.7597^{+0.0286}_{-0.0473}$
11	8.132	1.211	$8.8509^{+0.0517}_{-0.0518}$
13	10.739	1.203	$11.6998^{+0.2393}_{-0.2392}$
15	12.682	1.402	$13.5679^{+0.0963}_{-0.1055}$
17	11.181	0.853	$12.4588^{+0.3301}_{-0.3301}$
19	24.953	3.562	$27.1705^{+0.1676}_{-0.1676}$
21	4.375	0	$4.3371^{+0.0540}_{-0.0676}$
23	6.266	1.416	$6.4755^{+0.0675}_{-0.0989}$
25	25.741	1.671	$28.0008^{+0.3329}_{-0.1605}$
27	25.427	3.745	$28.1814^{+0.2110}_{-0.1516}$
31	17.827	2.393	$19.6587^{+0.3045}_{-0.3031}$
33	17.619	2.722	$19.3022^{+0.3000}_{-0.2987}$
35	16.733	1.944	$18.5262^{+0.2861}_{-0.2855}$
37	61.658	6.078	$67.1819^{+1.0326}_{-1.0319}$
39	43.639	6.315	$47.0818^{+0.7265}_{-0.7246}$
41	59.937	5.657	$64.0134^{+0.9863}_{-0.9857}$
43	56.989	6.524	$61.5614^{+0.9493}_{-0.9474}$
45	13.048	2.315	$14.3538^{+0.2218}_{-0.2215}$
47	37.577	3.438	$41.2186^{+0.6406}_{-0.6393}$
49	27.293	2.586	$29.7271^{+0.4648}_{-0.4634}$
51	38.935	4.825	$41.8919^{+0.6605}_{-0.6590}$
55	73.514	7.821	$80.2472^{+1.2462}_{-1.2439}$
Total	604.55	68.27	656.725 ± 8.940

3.1.2 Monte Carlo data

Large simulated Monte Carlo (MC) data samples are needed to model the decay kinematics, detection efficiencies and expected variable distributions of both $K\eta'\gamma$ events and all types of possible backgrounds. The optimisation of the event selection criteria and background suppression techniques depend heavily on the amount and accuracy of available MC. Samples of $B \to K\eta'\gamma$ MC and three types of background MC ("continuum", "generic $B\overline{B}$" and "rare $B\overline{B}$") form the basis of the set of discriminating selection criteria described in the following Sections. The samples are described below.

Signal Monte Carlo

B decays to the following final states (modes) are modelled in the $B \to K\eta'\gamma$ ("signal") MC:

$$B^+ \to K^+\eta'\gamma + c.c. \tag{3.2}$$

$$B^0 \to K^0_S\eta'\gamma + c.c. \tag{3.3}$$

The decay of the neutral B is only generated when it includes a K^0_S meson, though half the decays in data will include a K^0_L meson instead. K^0_L mesons are detected by the *KLM* detector, as described in the previous chapter. No energy information is given by the *KLM* and the inclusion of K^0_Ls in the reconstruction without their measured energies would reduce the accuracy of the analysis. K^0_Ss are only generated as decaying in the $\pi^+\,\pi^-$ mode ($\mathcal{BF}(K^0_S \to \pi^+\pi^-) = 69.2\%$), as this is the optimal mode for reconstruction within the detector. The next-most common decay, $\pi^0\,\pi^0$ ($\mathcal{BF}(K^0_S \to \pi^0\pi^0) = 30.7\%$), cannot be reconstructed efficiently due to the poor energy resolution of low energy γs in the *ECL* and would also give no improvement to the analysis. These restrictions of phase space sensitivity result in a $\sim 65\%$ reduction in the reconstruction efficiency of $B^0 \to K^0_S\eta'\gamma$.

The η' meson decay is generated in the two modes listed below, with their individual measured $\mathcal{BF}$s [19];

$$\eta' \to \rho^0\gamma \qquad 29.4\% \tag{3.4}$$

$$\eta' \to \eta\pi^+\pi^- \qquad 44.6\% \tag{3.5}$$

and the η meson decay in two:

$$\eta \to \gamma\gamma \qquad 39.3\% \tag{3.6}$$

$$\eta \to \pi^+\pi^-\pi^0 \qquad 22.7\% \tag{3.7}$$

The ρ^0 meson decays to $\pi^+\pi^-$ $\sim 100\%$ of the time and the π^0 meson to $\gamma\gamma$ $\sim 98\%$ of the time. They are generated in these modes only. This gives six final states and their complex conjugates:

$$B^+ \to K^+\eta'\gamma \to K^+(\rho^0\gamma)\gamma + c.c \qquad 29.4\% \qquad (3.8)$$

$$B^0 \to K^0_S\eta'\gamma \to K^0_S(\rho^0\gamma)\gamma + c.c \qquad 10.1\% \qquad (3.9)$$

$$B^+ \to K^+\eta'\gamma \to K^+(\eta\pi^+\pi^-)\gamma \to K^+((\gamma\gamma)\pi^+\pi^-)\gamma + c.c \qquad 17.6\% \qquad (3.10)$$

$$B^0 \to K^0_S\eta'\gamma \to K^0_S(\eta\pi^+\pi^-)\gamma \to K^0_S((\gamma\gamma)\pi^+\pi^-)\gamma + c.c \qquad 6.1\% \qquad (3.11)$$

$$B^+ \to K^+\eta'\gamma \to K^+(\eta\pi^+\pi^-)\gamma \to K^+((\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma + c.c \qquad 10.1\% \qquad (3.12)$$

$$B^0 \to K^0_S\eta'\gamma \to K^0_S(\eta\pi^+\pi^-)\gamma \to K^0_S((\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma + c.c \qquad 3.5\% \qquad (3.13)$$

The listed $\mathcal{BF}$s are the proportion of either the charged or neutral B meson decay to $K\eta'\gamma$ that the final state represents [19]. This gives the analysis sensitivity to 57.1% of the total $B^+ \to K^+\eta'\gamma$ decay final states and 19.7% of the total $B^0 \to K^0\eta'\gamma$ decay final states. These are called the 'daughter branching fractions' and are included in the calculation of the measured $\mathcal{BF}$ ($B \to K\eta'\gamma$) shown in Section 5.3.

The complex conjugate form of the above processes are implied throughout the rest of this thesis.

EvtGen was used to generate fifty thousand signal MC events for each of the six final states, for a total of 600,000 $K\eta'\gamma$ MC events. GEANT3 was used to model the detector response to the generated decays as described in Section 2.5. Half was modelled on experiment 27 parameters and backgrounds and half on experiment 37. This means half the sample is generated using SVD1 geometry, efficiency and accuracy and half SVD2.

In the EvtGen process, the $\Upsilon(4S)$ vector (spin=1, odd parity) meson decays into the two pseudoscalar (spin=0, odd parity) mesons of the $B\overline{B}$ pair with a decay amplitude of 1. To generate the three body process $B \to K\eta'\gamma$, one B meson then decays into the prompt photon (the "signal" photon) and a pseudo-resonance labelled X_s, the s indicating the presence of a strange quark. The pseudo-resonance is modelled as a vector particle (spin=1, odd parity) with a mass placed randomly between $1.52\,\mathrm{GeV}/c^2$ and $2.7\,\mathrm{GeV}/c^2$. This mass range is chosen to provide the correct distribution of the signal photon energy. The distribution is randomised to remove the need for a model-dependant distribution structure. The X_s then decays into the two pseudoscalar mesons K and η'.

The second B meson decays 'generically'- that is, in all possible channels as recorded in the Particle Data Group (PDG) review of 2004 [55]. These are dominated by the $b \to c$ transition.

Continuum Monte Carlo

The background for this analysis is divided into three components. The componant expected to dominate is from interactions in which the electron-positron collision does not result in the production of a $b\overline{b}$ pair but in a pair of lighter quarks; $u\overline{u}$, $d\overline{d}$, $s\overline{s}$ or $c\overline{c}$. These interactions are collectively called $q\overline{q}$ or 'continuum' background and occur in about 75% of hadronic interactions at the $\Upsilon(4S)$ resonance. Continuum back-

ground is modelled using $1703.98\,\mathrm{fb}^{-1}$ of QQ/EvtGen and GSIM generated MC produced by the Belle collaboration for use in all analyses.

Generic $B\overline{B}$ Monte Carlo

The $b \to c$ transition dominates the decay modes at the $\Upsilon(4S)$ resonance and are a background for $b \to s$ analyses like this. This type of background is also called 'generic' $B\overline{B}$ background and is modelled using $2846.11\,\mathrm{fb}^{-1}$ of QQ/EvtGen and GSIM generated MC.

Rare $B\overline{B}$ Monte Carlo

The $b \to u, d, s$ transitions, or charmless B decays, are much less common than $b \to c$. They are labeled 'rare' $B\overline{B}$ background. This background is expected to include decay modes with similar or identical final state particles as $B \to K\eta'\gamma$ and could therefore be mistaken as signal events. The rare $B\overline{B}$ background is modelled with a very large MC sample produced by a Belle collaborator equivalent to $24{,}850\,\mathrm{fb}^{-1}$ of integrated luminocity.

3.2 Hadronic event selection

The triggering system described in Section 2.4 prevents many non-hadronic events being recorded by the Belle DAQ. Some of these types of events do enter the data set however, including those from $e^+e^- \to e^+e^-, \mu^+\mu^-, \tau^+\tau^-, e^+e^-\gamma, \gamma\gamma$, and beam backgrounds interactions. A series of general hadronic selection criteria has been developed at Belle to augment the triggering system. These criteria are compiled in a software skim called HadronB [56], which is run over the recorded data set and removes more of theses non-hadronic events. The result is used by most analyses conducted at Belle. It has an efficiency of 99.1% for $B\overline{B}$ events and 79.5% for hadronic continuum ($q\overline{q}$) events, while removing 99.998% of Bhabha (e^+e^-) and radiative Bhabha ($e^+e^-\gamma$) events, 99.6% of $\gamma\gamma$ events, 95.1% of $\tau^+\,\tau^-$ events, and 91.0% of beam-gas background events. This removes the need for large MC samples to model these backgrounds.

The HadronB selection criteria are:

- **Track Multiplicity:** The number of *good* charged tracks must be greater than 3. A *good* track satisfies: $p_t > 0.1\,\mathrm{GeV}/c$, where p_t is the momentum transverse to the positron beam axis; $dr < 2.0\,\mathrm{cm}$, where dr is the distance of closest approach to the z-axis; and $|dz| < 4.0\ \mathrm{cm}$, where $|dz|$ is the distance of closest approach to the IP in the z direction.
- **Visible Energy:** The total visible energy, defined as the sum of track and photon energies, must satisfy;

$$E_{vis} \geq 0.20\sqrt{s} \tag{3.14}$$

where $\sqrt{s}$ is the total CoM energy.

- **Momentum Balance:** The z-component of the momentum sum of *good* charged tracks and *good* photons must satisfy;

$$\left|\sum p_z c\right| \leq 0.5\sqrt{s} \tag{3.15}$$

Good photons are defined as ECL clusters of $E > 100\,\mathrm{MeV}$ with no associated CDC track.

- **Primary Vertex:** A vertex fit of all *good* charged tracks in the event must satisfy;

$$|dz| < 3.5\,\mathrm{cm} \tag{3.16}$$

$$dr < 1.5\,\mathrm{cm} \tag{3.17}$$

This removes beam-gas interactions that do not originate at the IP.

- **Cluster Energy Sum:** The energy sum of ECL clusters with $E > 100\,\mathrm{MeV}$ within ($17^\circ \leq \theta \leq 150^\circ$) must satisfy;

$$0.1 \leq \frac{\sum E_{ECL}}{\sqrt{s}} \leq 0.80 \tag{3.18}$$

- **Cluster Multiplicity:** There must be two or more clusters of $E > 100\,\mathrm{MeV}$ in the barrel region of the ECL ($-0.7 \leq \cos\theta \leq 0.9$)

- **Average Cluster Energy:** The average cluster energy within the ECL must satisfy;

$$\frac{\sum E_{ECL}}{n_{cluster}} < 1.0\,\mathrm{GeV} \tag{3.19}$$

- **Heavy Jet Mass:** The heavy jet mass (M_{jet}) must satisfy the conditional requirement of;

$$M_{jet} > 0.25 E_{vis} \quad \textbf{if} \quad M_{jet} \leq 1.8\,\mathrm{GeV}/c^2 \tag{3.20}$$

where M_{jet} is calculated by dividing all event tracks and photons into two hemispheres defined by the plane perpendicular to the event thrust axis. The invariant mass of all the particles in each hemisphere is calculated with a pion mass hypothesis for all charged tracks, and the largest taken as M_{jet}. The event thrust axis ($\vec{n}$) is defined as the direction which maximises the momentum vector sum of all event particles;

$$Max(\frac{\sum_i |\vec{n} \cdot \vec{p}_i|}{\sum_i |\vec{p}_i|}) \tag{3.21}$$

where $\vec{p}_i$ is the 3-momentum of the i^{th} particle.

- **Conditional Cluster Energy Sum:** A more stringent requirement on the ECL cluster energy sum is enforced if it meets the conditional requirement of;

$$\frac{\sum E_{ECL}}{\sqrt{s}} > 0.18 \quad \textbf{if} \quad M_{jet} < 1.8\,\mathrm{GeV}/c^2 \tag{3.22}$$

3.3 Event selection

The following section describes the discriminating variables that are used to separate $B \to K\eta'\gamma$ events from other interactions. The efficiency of the selection process is examined in Chapter .

3.3.1 Signal photon selection

The primary identifier of a radiative B meson decay is an isolated high energy photon. This photon is monochromatic at the quark level but gains a finite width in the lab frame due to several factors:

- The Fermi momentum of the b quark within the B meson.
- The B meson lab momentum.
- Detector resolution effects.

ECL clusters with no matching CDC track and energy in the range $1.8\,\mathrm{GeV} < E_{CoM} < 3.4\,\mathrm{GeV}$ are taken as candidate signal photons.

The B mesons decay isotropically as they are spin-0 particles. While the boost of the $\Upsilon(4S)$ gives the signal photon some directionality in the lab frame, the photons created still generally have large transverse momentum. One of the backgrounds to radiative photons in this energy range are photons from initial state radiation (ISR) interactions, where either the e^- or the e^+ emitts a hard photon before collision. This shifts the interaction CoM energy below the $b\overline{b}$ production threshold and results in a final state of $e^+e^- \to q\overline{q}\gamma$ ($q = u, d, s, c$). ISR photons peak in the forward and backward directions of the beam pipe and produce high levels of background in the EFCs. To reduce this background, only clusters from within the barrel ($32° < \theta < 129°$) of the ECL are considered.

The largest backgrounds to the signal photon are from the decay of high momentum π^0 and η particles into $\gamma\gamma$. The momentum required of these two light mesons in order to produce an ECL cluster within the signal energy range boosts the photons, back-to-back in the π^0/η frame, to be highly co-linear. Two

isolation conditions must be met by the signal photon candidates: there must be no other cluster of energy greater than 20 MeV detected within 30cm of it in the ECL, and they must satisfy $0.95 \leq E_9/E_{25} \leq 1.0$. The variable E_9/E_{25} is the ratio of energy deposited within the central 9 CsI(Tl) scintillating calorimeter crystals of the shower divided by the energy deposited within the central 25 crystals. A high ratio is an indicator of a collimated shower, and a low ratio indicates a shower with large lateral spread, as would be the case if two separate showers overlapped.

Photons from π^0 and η decays are further suppressed using a probability density function (PDF) based algorithm created by P. Koppenburg [35] for his inclusive $b \to s\gamma$ analysis at Belle. The algorithim first combines candidate signal photons with every other photon in the event in turn and calculates each pair's invariant mass ($m_{\gamma_1\gamma_2}$). Probabilities for the pair being from either a π^0 or η decay are then assigned based on $m_{\gamma_1\gamma_2}$ and the energy of the secondary photon (E_{γ_2}). The 2D PDFs used in assigning the likelihoods were trained on samples of $q\overline{q}$, $b \to c$ and $b \to u, d, s$ MC. They are shown in Figure 3.1. Signal photon candidates are required to have $Pr(\pi^0) < 0.25$ and $Pr(\eta) < 0.25$ for all possible combinations.

If more than one photon from an event passes selection, the highest energy candidate is taken.

3.3.2 Charged tracks

The detection of charged particles is handled primarily by the CDC. The information from the CDC axial and stereo sense wires is matched to hits in the SVD and the combined detector information is fitted assuming a helical path. The parameters which describe this helix: the radius of curvature, the pitch and the pivot point co-ordinates; provide the transverse momentum, longitudinal momentum and dr of the track. The inclusion of the SVD information improves the pivot point measurement and the fitted momentum resolution.

Any charged tracks recorded in the detector must first pass the standard *good* track selection described earlier in this chapter: $p_t > 0.1\,\mathrm{GeV}/c$, $dr < 0.20$ cm, and $|dz| < 4.0$ cm. Any event with less than three charged tracks passing these cuts is discarded, as this is the minimum required to reconstruct $K\eta'\gamma$. Further cuts of $dr < 0.15$ cm and $p_{CoM} > 100$ MeV/c are applied. These requirements reduce the number of poorly reconstructed tracks and those from beam-wall and beam-gas interactions. K^0_S daughter pions are exempt from the these final cuts.

The only charged particles required to reconstruct the modes of $B \to K\eta'\gamma$ analysed are π^+/π^- and K^+/K^-. Electron and positron contamination is reduced by imposing a cut of $\mathcal{L}_e < 0.6$, where $\mathcal{L}_e$ is the likelihood output of the electron identification software algoithim used at Belle [36]. Electrons are identified by large showers in the ECL and their energy:momentum (E/p) ratio. E/p peaks close to 1.0 for electrons with momentum in the GeV range, whereas the heavier π^+ and K^+ particles will have distributions peaking at a lower ratio. $\mathcal{L}_e$ combines the outputs of **atc_pid** (described in detail below) with shower shape and energy deposition information from the ECL cluster.

The remaining tracks are differentiated into kaon and pion candidates using the identification likelihood

ratio $\mathcal{R}_{K,\pi} = \mathcal{L}_K/(\mathcal{L}_K + \mathcal{L}_\pi)$. $\mathcal{L}_\pi$ and $\mathcal{L}_K$ formed from the number of photoelectrons ($N_{p.e.}$) detected in the ACC, the energy loss (dE/dx) measured in the CDC, and their time-of-flight measured in the TOF. Tracks with $\mathcal{R}_{K,\pi} > 0.6$ are taken as kaon candidates and tracks with $\mathcal{R}_{K,\pi} < 0.9$ as pion candidates. The overlap means no tracks are lost from reconstruction.

3.3.3 π^0 reconstruction

Clusters in the ECL of energy greater than 50 MeV are combined in pairs to form candidate π^0 mesons. Constructed π^0 mesons with invariant mass in the range $0.119\ \mathrm{GeV}/c^2 < M_{\gamma\gamma} < 0.152\ \mathrm{GeV}/c^2$ and laboratory system momenta $p_{\pi^0} > 100\ \mathrm{MeV}/c$ are taken as candidates.

3.3.4 K^0_S meson reconstruction

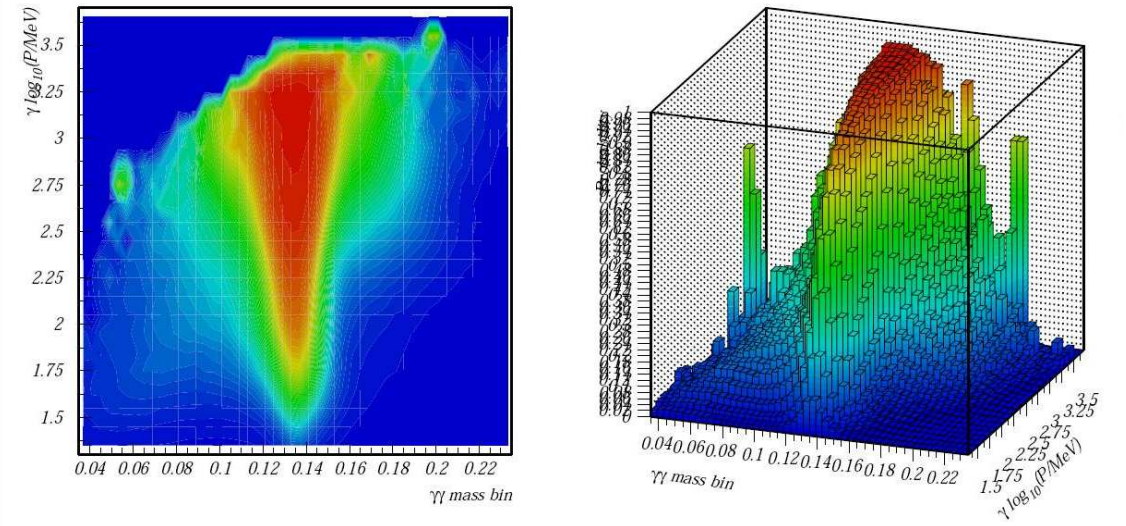
K^0_S candidates are formed from $\pi^+\ \pi^-$ pairs. The tracking algorithim selects pairs of oppositely charged tracks that are likely to have originated from a common source. These pairs must pass a set of momentum dependant cuts on:

- **Proximity to the IP:** dr. The smallest of the two tracks minimum distance from the IP in the x-y plane.
- **Distance between tracks:** z_dist. The distance between the two tracks at their point of closest proximity.
- **Flight length:** fl. The distance between the IP and the reconstructed vertex point in the x-y plane.
- **Angle between the momentum vector and reconstructed vertex vector:** $d\phi$. The azimuthal angle between the momentum vector and decay vertex vector of the candidate.

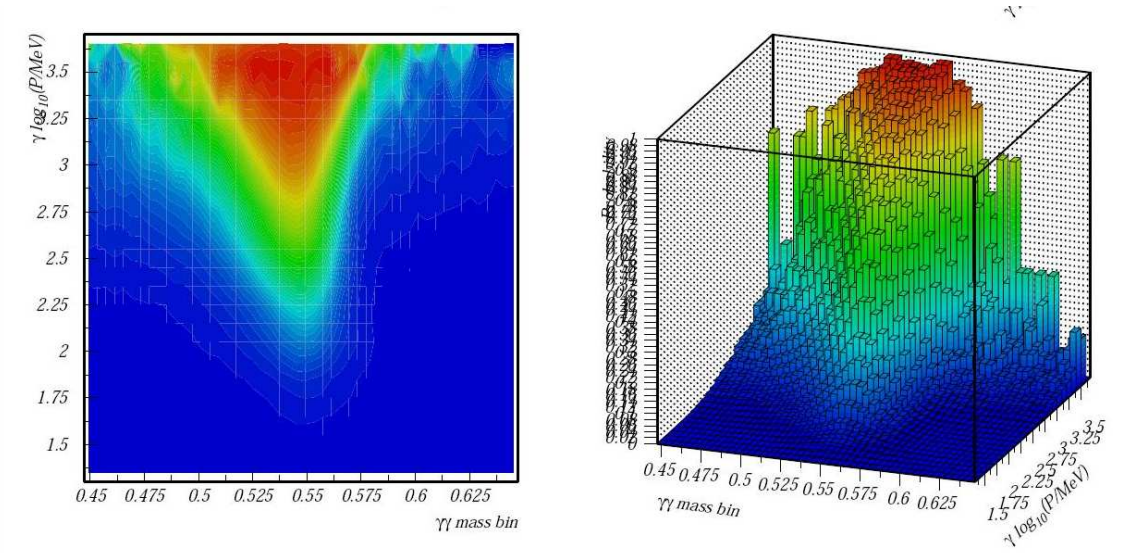
These are called the **goodKs** cuts [31] and are shown in Table 3.2. Candidates who pass and whose invariant mass lies within $10\,\mathrm{MeV}/c^2$ of the nominal K^0_S mass [19] are included in the reconstruction.

Momentum (GeV/c)	dr(cm)	$d\phi$(radians)	z_dist(cm)	fl(cm)
< 0.5	> 0.05	< 0.3	< 0.8	-
$0.5 - 1.5$	> 0.03	< 0.1	< 1.8	> 0.08
> 1.5	> 0.02	< 0.03	< 2.4	> 0.22

Table 3.2: The **goodKs** cuts.



(a) The π^0 PDF in surface and lego plots.



(b) The η PDF in surface and lego plots.

Figure 3.1: The 2D PDFs used to assign π^0 and η likelihoods to signal γ candidates, shown in the $log_{10}(E_{\gamma_2})$ vs. $m_{\gamma_1\gamma_2}$ plane [35].

3.3.5 η meson reconstruction

η mesons are reconstructed via two channels: $\eta \to \gamma\gamma$ and $\eta \to \pi^+\pi^-\pi^0$. The $\eta \to \gamma\gamma$ candidates are chosen by combining two ECL clusters of energy greater than $100\,\mathrm{MeV}$ and selecting those with $490\,\mathrm{MeV}/c^2 < M_{\gamma\gamma} < 590\,\mathrm{MeV}/c^2$. They are also required to satisfy $|\cos\theta_{hel}| < 0.9$ where $\cos\theta_{hel}$is the helicity angle of the η decay, defined as

$$\cos\theta_{hel} = \frac{\vec{p}^{\,\prime}_{\eta'} \cdot \vec{p}^{\,\prime}_{\gamma}}{|\vec{p}^{\,\prime}_{\eta'}||\vec{p}^{\,\prime}_{\gamma}|} \tag{3.23}$$

where $\vec{p}^{\,\prime}_{\eta'}$ is the momentum of the η' in the η rest frame, and $\vec{p}^{\,\prime}_{\gamma}$ is the momentum of the candidate photon in the η rest frame. Truly reconstructed η mesons will have a uniform distribution, whereas random combinations of photons, which are typically more energetically asymmetric, have a distribution that peaks at ± 1.

The $\eta \to \pi^+\pi^-\pi^0$ channel is reconstructed by combining two oppositely charged tracks that have passed the pion selection criteria and a pair of photons that have passed the criteria for inclusion as a π^0. The combined four vector is required to have invariant mass $537\,\mathrm{MeV}/c^2 < M_{\pi^+\pi^-\pi^0} < 560\,\mathrm{MeV}/c^2$.

A mass-constrained fit is then applied to all reconstructed η mesons and their momentum adjusted according to the results.

3.3.6 ρ^0 meson reconstruction

The ρ^0 meson decays to a π^+ π^- pair with a likelihood of $\approx$ 100%. Pairs of oppositely charged tracks passing pion selection are combined and those with $550\,\mathrm{MeV}/c^2 < M_{\pi^+\pi^-} < 850\,\mathrm{MeV}/c^2$ are taken. A vertex fit of the pair which constrains the two track to originate from a common point must succeed with a χ^2 of less than 100 using the kvertex fitter package [30]. Candidates must also pass an helicity cut similar to the η cut: $|\cos\theta_{hel}| < 0.9$ where

$$\cos\theta_{hel} = \frac{\vec{p}^{\,\prime}_{\eta'} \cdot \vec{p}^{\,\prime}_{\pi}}{|\vec{p}^{\,\prime}_{\eta'}||\vec{p}^{\,\prime}_{\pi}|} \tag{3.24}$$

As the ρ^0 is a vector meson and it decays to two pseudoscalars, its helicity follows a $\cos^2\theta_{hel}$ distribution. The distribution of combinatorial background will be approximately flat. The helicity distributions of $K\eta'\gamma$ MC and $q\overline{q}$ MC are shown in Figure 3.2.

3.3.7 η' meson reconstruction

The η' meson candidates are reconstructed in two channels: $\eta' \to \rho^0\gamma$ and $\eta' \to \eta\pi^+\pi^-$. The ρ^0 candidates are combined with photons of energy greater than $200\,\mathrm{MeV}$. η candidates are combined with pairs of oppositely charged pion candidates.

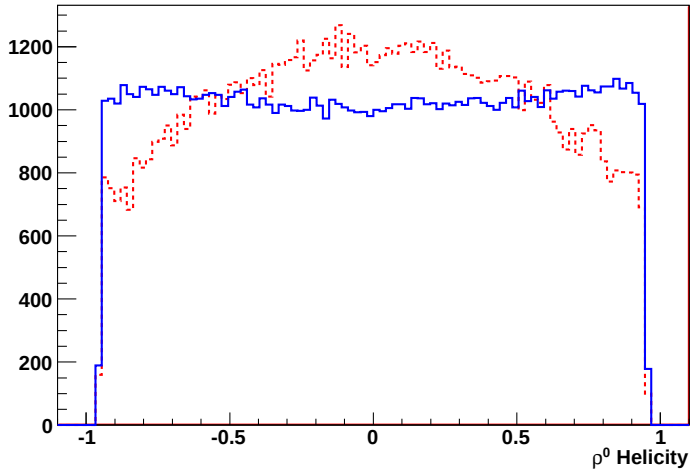


Figure 3.2: The helicity distributions of ρ^0 candidates. $q\overline{q}$ MC is in solid blue, truly reconstructed ρ^0 candidates from $K\eta'\gamma$ MC in dashed red. The $q\overline{q}$ MC distribution has been normalised to the same area as the $K\eta'\gamma$ MC.

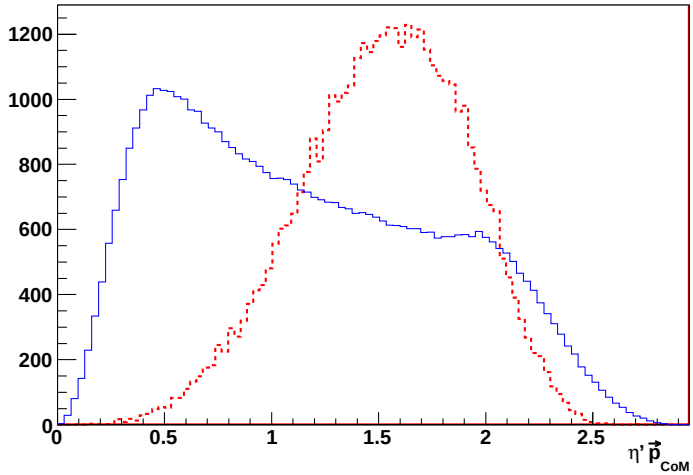


Figure 3.3: η' momentum in the center of mass frame. The solid blue histogram is $q\overline{q}$ MC, the dashed red is $K\eta'\gamma$ MC. The $q\overline{q}$ MC histogram has been normalised to the same area as the $K\eta'\gamma$ MC histogram.

Section 4.3 describes three different techniques for modelling the data distributions and extracting the number of $B \to K\eta'\gamma$ events in the data sample; a three dimensional (3D) mode-by-mode fitter, a 3D combined fitter, and a 2D fitter. The three methods have different selection criteria for $M_{\eta'}$ depending on whether it is included as a fit variable and how the fit is performed. Invariant masses in the range $930\ \mathrm{MeV}/c^2 < M_{\rho^0\gamma} < 985\ \mathrm{MeV}/c^2$ and $940\ \mathrm{MeV}/c^2 < M_{\eta\pi^+\pi^-} < 975\ \mathrm{MeV}/c^2$ are defined as the fitting regions for $M_{\eta'}$ in the 3D mode-by-mode fitting procedure. The 3D combined fitting method treats all final states in the same way: both $M_{\rho^0\gamma}$ and $M_{\eta\pi^+\pi^-}$ must satisfy $930\ \mathrm{MeV}/c^2 < M_{\eta'} < 985\ \mathrm{MeV}/c^2$. The 2D fitting procedure has tighter requirements as $M_{\eta'}$ is not included it the fit and therefore does not need to encompass sideband regions for stable function modelling: the $M_{\eta'}$ selection cuts for 2D fits are defined as $945\ \mathrm{MeV}/c^2 < M_{\rho^0\gamma} < 970\ \mathrm{MeV}/c^2$ and $950\ \mathrm{MeV}/c^2 < M_{\eta\pi^+\pi^-} < 965\ \mathrm{MeV}/c^2$. These ranges also define the signal region for the two 3D fitters.

Reconstructed η' candidates must have momentum $p_{CoM} > 1.0\,\mathrm{GeV}/c$ where p_{CoM} is the magnitude of the momentum of the η' calculated in the CoM frame. This frame was chosen to maximise the separation between signal and continuum MC. The lorentz boost into the CoM frame improves the momentum resolution of truly reconstructed η''s and smears the distribution of random combinations. The distribution of p_{CoM} for $K\eta'\gamma$ MC and $q\overline{q}$ MC is shown in Figure 3.3.

3.3.8 B-meson reconstruction

Charged and neutral kaons are combined with η' candidates to form an X_s pseudo-resonance. The high energy of photons from $b \to s\gamma$ decays provides an upper limit for the mass of the X_s system: in this analysis the invariant mass of the kaon-η' combination is required to be below $3.4\,\mathrm{GeV}/c^2$. This is to a certain extent redundant after the signal photon energy selection is in place, however it does have some power to remove combinatorial and continuum backgrounds and will not reduce the measured phase space.

The candidate signal photon is then combined with the X_s system to form a B meson candidate. Two kinematic variables are formed from the momentum and energy of the B candidate and the CoM energy provided by the colliding electron-positron pair. As $\Upsilon(4S) \to B\overline{B}$ is a two-body process each B meson will carry half the total CoM energy of the collision: $\sqrt{s}/2 = E_{\mathrm{beam}} = 5.29\,\mathrm{GeV}$, where E_{beam} is called the 'beam energy'. E_{beam} is measured at Belle every ~ 100 runs and has a spread of about $3\,\mathrm{MeV}$. By replacing the measured B candidate energy with E_{beam} and calculating the invariant mass we form the kinematic variable 'beam constrained mass' (M_{BC}):

$$M_{\mathrm{BC}} \equiv \sqrt{E_{\mathrm{beam}}^2 - (\vec{p}^{\,*}_{X_s} + \vec{p}^{\,*}_{\gamma})^2 c^2} \tag{3.25}$$

where $\vec{p}^{\,*}_{X_s}$ and $\vec{p}^{\,*}_{\gamma}$ are the CoM momenta of the X_s combination and the signal photon respectively. As the single measurement of E_{beam} carries a higher resolution than the multiple energy measurements in the reconstructed B, replacing the measured B energy with E_{beam} improves the mass resolution of correctly

reconstructed candidates. To further increase the mass resolution the momentum of the signal photon is rescaled such that $E_{\rm beam} = p^*_\gamma + E^*_{X_s}$, as detector momentum resolution is lower for photons than for charged particles. Correctly reconstructed B mesons will peak at the mass of the B meson: $5.28\,{\rm GeV}/c^2$. Randomly reconstructed Bs will be distributed in the shape of a kinematic endpoint with a maximum of $M_{\rm BC} = 5.29\,{\rm GeV}/c^2$. These can be seen in Figure 3.4.

The second kinematic variable is called the 'energy difference' (ΔE) and is defined as

$$\Delta E \equiv E_B - E_{\rm beam} \tag{3.26}$$

where E_B is the energy of the candidate B meson. Correctly reconstructed B mesons will peak at $\Delta E = 0$, while random combinations will be distributed essentially randomly. The width of ΔE is a function of the momentum resolution of the final state particles. In radiative decay analyses the signal photon has the lowest momentum resolution and gives a distinctive low-energy tail to the ΔE distribution, as shown in Figure 3.4.

Candidate B mesons are retained if they satisfy the conditions

$$|\Delta E| < 0.3\ {\rm GeV} \tag{3.27}$$

$$5.20\ {\rm GeV}/c^2 < M_{\rm BC} < 5.29\ {\rm GeV}/c^2 \tag{3.28}$$

The signal region is defined as

$$-0.1\ {\rm GeV} < \Delta E < 0.07\ {\rm GeV} \tag{3.29}$$

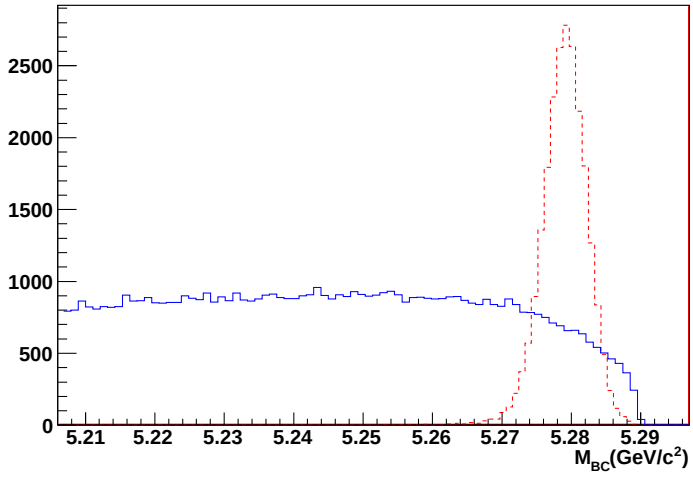
$$5.27\ {\rm GeV}/c^2 < M_{\rm BC} < 5.29\ {\rm GeV}/c^2 \tag{3.30}$$

The sideband region is defined as

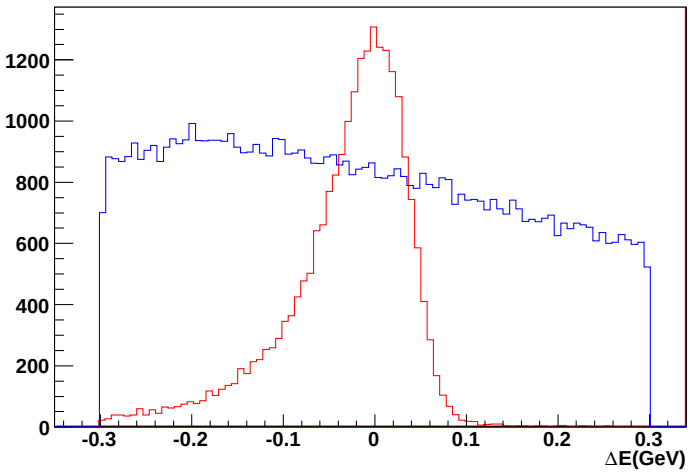
$$-0.3\ {\rm GeV} < \Delta E < -0.2\ {\rm GeV}\ \ {\rm or}\ \ 0.1\ {\rm GeV} < \Delta E < 0.3\ {\rm GeV} \tag{3.31}$$

$$5.20\ {\rm GeV}/c^2 < M_{\rm BC} < 5.26\ {\rm GeV}/c^2 \tag{3.32}$$

The data events in these sideband regions are used for background estimation in Section 3.5.



(a) $M_{\rm BC}$ distribution in MC.



(b) ΔE distribution in MC.

Figure 3.4: The distributions of $M_{\rm BC}$ and ΔE. The solid blue histogram is $q\overline{q}$ MC, the dashed red is correctly reconstructed $K\eta'\gamma$ MC. The $q\overline{q}$ MC histogram has been normalised to there times the area of the $K\eta'\gamma$ MC histogram.

3.4 B Backgrounds

The majority of the events from $b \rightarrow c$ decays that pass selection cuts consist of D mesons (D^+, D^0, $D^*(2010)^+ \rightarrow (D\pi)$, $D^*(2007)^0 \rightarrow (D^0\pi^0)$) combined with one or more light unflavoured mesons (π^+, π^0, ρ^0, ρ^+, η, ω), where the D decays into a kaon plus one or more pions. The background mode making the largest contribution to the signal region (Equations 3.29 and 3.30) was found to be $B^+ \rightarrow D^0\rho^+$ for $B^+ \rightarrow K^+\eta'\gamma$ and $B^0 \rightarrow D^+\rho^-$ for $B^0 \rightarrow K^0_S\eta'\gamma$. Invariant mass scans were performed for the D decay modes shown in Table 3.3. B decays including ϕ and K^* also contribute and scans were performed for their decays including kaons as shown in Table 3.3. The candidate signal kaon was combined with all other particles in the event in turn to form the combinations listed and the resulting invariant mass plots examined.

Table 3.3: Mass scan decay modes.

Particle	Decay Mode
D^0	$K^+\pi^-$
D^0	$K^+\pi^-\pi^0$
D^0	$K^+\pi^-\pi^-\pi^+$
D^+	$K^+\pi^+\pi^-$
D^+	$K^0_S\pi^+$
D^+	$K^0_S\pi^+\pi^0$
ϕ	K^+K^-
$K^*(892)^0$	$K^+\pi^-$
$K^*(892)^+$	$K^+\pi^0$

Only $D^0 \rightarrow K^-\pi^+ + c.c.$ shows strong peaking behaviour in the background MC samples. B candidates that have a combination of the charged signal kaon and any other charged pion with $1.84GeV < M_{K^-\pi^+} < 1.89GeV$ are discarded. Figure 3.5 shows the $M_{K^-\pi^+}$ distribution for signal MC and generic BB MC.

Few events from the $b \rightarrow u, d, s$ MC sample pass the selection cuts. Three peaking background modes were identified: $B \rightarrow J/\psi K$, $B \rightarrow K\eta'\eta$ and $B \rightarrow K\eta'\pi^0$. When the J/ψ decays to η' γ, $B \rightarrow J/\psi K$ has identical final state particles to the signal modes and the γ can fall within the lower bounds of the signal photon energy requirement. To suppress this background a veto is placed on events that have a combined η' γ invariant mass within $\pm 25\,\mathrm{MeV}/c^2$ of the nominal J/ψ mass [19].

The other peaking backgrounds, $B \rightarrow K\eta'\eta$ and $B \rightarrow K\eta'\pi^0$, pass the selection cuts when the η or π^0 decays with asymmetric γ energies in the lab frame. The second lower energy γ is not included in the B reconstruction. When the second γ is recorded in the detector, these events are suppressed by the π^0 and η vetoes detailed above. If the second γ is missing- when it doesn't reach the ECL due to intervening material or it passes through a crack in the detector hermeticity- only the kinematic variables M_{BC} and ΔE provide discrimination. No further suppression is applied.

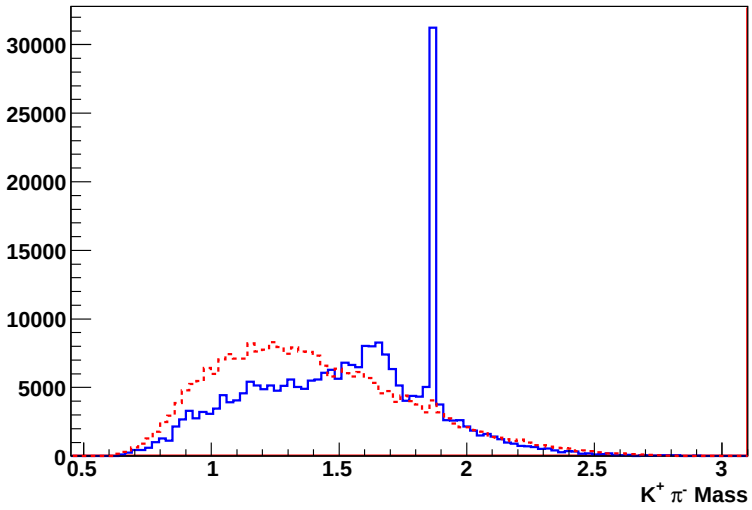


Figure 3.5: $M_{K^-\pi^+}$ for signal K^- candidates combined with all π^+s in the event. The solid blue histogram is $b \to c$ MC, the dashed red is $K\eta'\gamma$ MC. The $b \to c$ MC histogram has been normalised to the same area as the $K\eta'\gamma$ MC histogram.

3.5 Continuum suppression

For this analysis the most important background is that of $e^+e^- \to q\overline{q}$ events ($q = u, d, s, c$). The high energy γ in these events generally originates from a π^0 or η decay to $\gamma\gamma$. The isolation cuts described in the previous chapter remove many of these events, but more discrimination is required in the form of targeted continuum background suppression.

As previouly mentioned, the $\Upsilon(4S)$ meson is just above the energy threshold required to produce a $B\overline{B}$ pair. When an $\Upsilon(4S)$ decays to a $B\overline{B}$ pair the B mesons are therefore almost at rest in the CoM frame. When a light quark pair is created instead of an $\Upsilon(4S)$ the mass difference between the produced quark pair and that of an $\Upsilon(4S)$ becomes momentum. This results in a pair of high momentum hadronic jets being recorded in the detector with collimated event topology in the CoM frame. The B mesons, with little momentum and spin 0, decay almost isotropically. This topological difference is exploited via event shape variables. In this analysis 12 modified Fox Wolfram moments [45] and the transverse event energy are combined into a Fisher discriminant [46]. Additional discrimination is provided by the use of $\cos(\theta_B)$ and Δz. The definition of these variables follows.

3.5.1 SFW and the Fisher discriminant

Fox-Wolfram moments provide a quantification of the sphericity of a group of vectors, in this case the momentum 3-vectors of the particles in an event, and are defined as:

$$\mathcal{H}_l = \sum_{i,j} \frac{p_i p_j}{E_{vis}^2} P_l(\cos\theta_{ij}) \tag{3.33}$$

where *i* and *j* run over the all charged tracks, photons and K_S^0 recorded in an event, p_i and p_j are the particle momenta, θ_{ij} is the angle between the particles, E_{vis} is the total visible energy in the event, and P_l is the l^{th} order Legendre polynomial. It has been found that by dividing the event particles into those that are incorporated into the candidate B meson and all other particles in the event, the moments provide greater discrimination. Thus the Fox-Wolfram moments become

$$\mathcal{R}_l^{so} = \sum_{i,j} \frac{p_i^s p_j^o}{E_{vis}^2} P_l(\cos\theta_{i,j}) \tag{3.34}$$

$$\mathcal{R}_l^{oo} = \sum_{i,j} \frac{p_i^o p_j^o}{E_{vis}^2} P_l(\cos\theta_{i,j}) \tag{3.35}$$

$$\mathcal{R}_l^{ss} = \sum_{i,j} \frac{p_i^s p_j^s}{E_{vis}^2} P_l(\cos\theta_{i,j}) \tag{3.36}$$

where p_i^s is the momentum of the i^{th} particle in the B candidate and p_i^o is the momentum of the i^{th} particle of the non-B candidate particles. These are the so-called Super Fox-Wolfram (SFW) moments. The $\mathcal{R}_l^{ss}$ moments are found to be correlated with $M_{\rm BC}$ and are not used. The $\mathcal{R}_l^{so}$ moments are further divided into those in which the non-B candidate particle is charged or neutral. Another moment which combines the B candidate particle momenta and the total missing momenta from the event is also included, taking the total moments used to:

$$\mathcal{R}_l^{oo} = \sum_{i,j} \frac{p_i^o p_j^o}{E_{vis}^2} P_l(\cos\theta_{i,j}) \tag{3.37}$$

$$\mathcal{R}_l^{cso} = \sum_{i,j} \frac{p_i^s p_j^{co}}{E_{vis}^2} P_l(\cos\theta_{i,j}) \tag{3.38}$$

$$\mathcal{R}_l^{nso} = \sum_{i,j} \frac{p_i^s p_j^{no}}{E_{vis}^2} P_l(\cos\theta_{i,j}) \tag{3.39}$$

$$\mathcal{R}_l^{ms} = \sum_{i} \frac{p_i^s p^m}{E_{vis}^2} P_l(cos\theta_i) \tag{3.40}$$

where p_i^{co} is the i^{th} particle momentum of the charged non-B candidate particles, p_i^{no} is the i^{th} particle momentum of the neutral non-B candidate particles, p^m is the total missing momentum of the event, and $\cos\theta_i$ is the angle between the p_i^s vector and the missing momentum vector. The moments from 0^{th} to 4^{th} order are included in the analysis, however the 1^{st} and 3^{rd} order $\mathcal{R}_l^{nso}$ and $\mathcal{R}_l^{ms}$ moments are also found to be correlated with $M_{\rm BC}$ and neglected. The sixteen modified Fox-Wolfram moments are combined into a Fisher discriminant ($\mathcal{F}$). The total event energy in the plane transverse to the beam direction (E_t) is also

found to have discriminating power and is included as well:

$$\mathcal{F} = \sum_{n=1}^{4} \alpha_n \mathcal{R}_l^{oo} + \sum_{n=1}^{4} \beta_n \mathcal{R}_l^{cso} + \sum_{n=2,4} \gamma_n \mathcal{R}_l^{nso} + \sum_{n=2,4} \delta_n \mathcal{R}_l^{ms} + \epsilon E_t \tag{3.41}$$

where α_n, β_n, γ_n, δ_n and ϵ are coefficients trained on signal and continuum MC to maximise the separation between signal and background, and E_t is the transverse event energy.

The Fisher discriminant gets its name from the technique of determining α_n, β_n, γ_n, δ_n and ϵ [46]. The method proceeds in this manner: if the variables in $\mathcal{F}$ are the set of x_i, then the values of the coefficients that maximise the separation between signal and background are found using by minimising

$$D = \frac{\sum_i \lambda_i (\mu_i^S - \mu_i^B)}{\sqrt{\sum_{i,j} \lambda_i \lambda_j (U_{ij}^S + U_{ij}^B)}} \tag{3.42}$$

where μ_i^S and μ_i^B are the means of the x_i distributions for signal and background, respectively, λ_i denote the coefficients, and U_{ij}^S and U_{ij}^B are elements of the signal and background covariant matrices for the series of measurements, respectively. The coefficients that minimise D can be expressed as

$$\lambda_i = \sum_j \frac{(\mu_j^B - \mu_j^S)}{(U_{ij}^S + U_{ij}^B)} \tag{3.43}$$

This method utilises any correlations between the input variables to maximise the separation between signal and background. The samples of $K\eta'\gamma$ MC and either $q\overline{q}$ MC or sideband data are analysed and U_{ij}^S, U_{ij}^B, μ_i^S, and μ_i^B extracted. Then α_n, β_n, γ_n, δ_n and ϵ are determined using Equation 3.43. The $q\overline{q}$ MC was originally used in the Fisher training but was replaced with sideband data to increase accuracy. The continuum suppression used in the 3D mode-by-mode method is trained on $q\overline{q}$ MC; the other two fitters use sideband data (see Section 4.3).

Any particles that fail to be reconstructed in the event (particles that pass through gaps in the detector hermeticity and badly reconstructed tracks, for example) will alter the overall shape of the event. The distribution of $\mathcal{F}$ therefore varies with the total reconstructed event mass. It been found that $\mathcal{F}$ provides the best suppression when the missing mass is small: when the two B mesons have been reconstructed accurately [47]. The output of $\mathcal{F}$ for each individual reconstructed final state is therefore divided into 7 bins of missing mass squared (mm^2). Each of the 42 resulting distributions is fitted with a asymmetric Gaussian in an unbinned maximum likelihood (ML) fit (see Section 4.1 for more details on ML fitting). The fitted functions are normalised and used as PDFs to calculate $\mathcal{L}_S^{SFW}$ and $\mathcal{L}_{q\overline{q}}^{SFW}$; the likelihood of any individual event being a $B \to K\eta'\gamma$ decay or from the continuum, respectively.

Very few $q\overline{q}$ MC events are tagged as $\eta \to \pi^+\pi^-\pi^0$ decays. This leaves inadequate statistics in each mm^2 bin for an accurate fit. The charged and neutral $\eta \to \pi^+\pi^-\pi^0$ modes are therefore combined and treated as one with respect to the training and fitting of $\mathcal{F}$.

The output of $\mathcal{F}$ for all modes of signal and continuum MC is shown in Figure 3.6.

3.5.2 Cosθ_B

$\cos\theta_B$ is the cosine of the polar angle between the reconstructed B meson 3-vector and the negative direction of the positron beam axis measured in the CoM frame. B mesons are produced in a p-wave state, as a pair of pseudoscalar mesons produced from the decay of a vector meson. Correct reconstructions will have a $(1-\cos^2\theta_B)$ dependence in $\cos\theta_B$ while random combinations will have an essentially uniform shape. These distributions are fitted with 2^{nd} order Chebychev polynomials, the resulting functions normalised to form PDFs, and $\mathcal{L}_S^{\cos\theta_B}$ and $\mathcal{L}_{q\overline{q}}^{\cos\theta_B}$ calculated in the same manner as $\mathcal{L}_S^{SFW}$ and $\mathcal{L}_{q\overline{q}}^{SFW}$. Figure 3.6 shows the distributions for $K\eta'\gamma$ and $q\overline{q}$ MC.

3.5.3 Δz

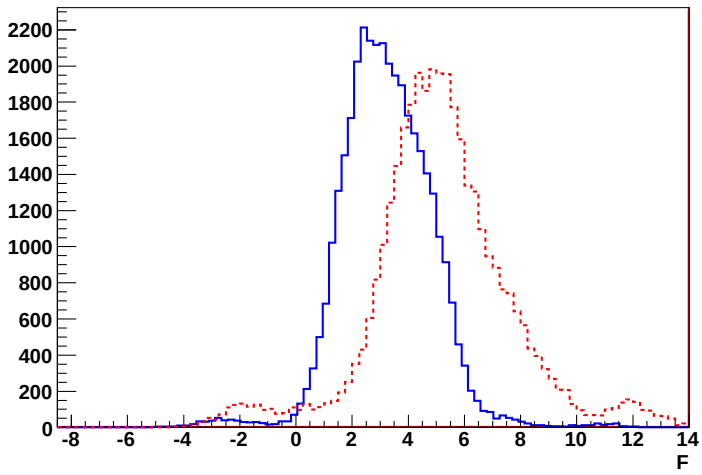
As the B mesons are created with a significant momentum in the lab frame and have a lifetime of the order of pico-seconds, they travel a measurable distance before decaying. Particles from continuum events, on the other hand, will generally have a common vertex close to the IP. The distance along the beam axis between the signal B vertex and the vertex of the other B in the CoM frame (ΔZ) is used in addition to $\mathcal{F}$ and $\cos\theta_B$ if $\Delta Z < 0.1$cm. When $\Delta Z > 0.1$cm, the discriminating value of $|dz|$ is lost and it is removed from consideration. The distribution for $K\eta'\gamma$ and $q\overline{q}$ MC is shown in Figure 3.6. The $K\eta'\gamma$ MC is fitted with a Gaussian centered at $\Delta Z = 0$ convolved with an exponential, modelling the detector resolution and physical decay probability of the particle respectively. The $q\overline{q}$ MC is fitted with a triple Gaussian. PDFs are created from the fits and $\mathcal{L}_S^{\Delta Z}$ and $\mathcal{L}_{q\overline{q}}^{\Delta Z}$ calculated.

3.5.4 Likelihood ratio

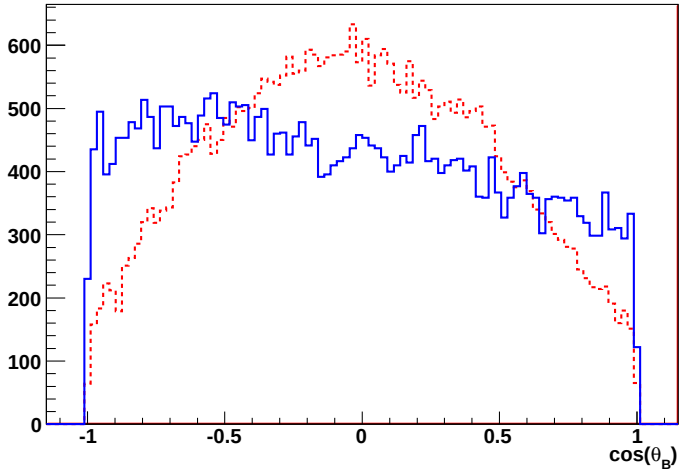
The three calculated likelihoods are combined into a likelihood ratio:

$$\mathcal{LR} = \frac{\mathcal{L}_S}{(\mathcal{L}_S + \mathcal{L}_{q\overline{q}})} \tag{3.44}$$

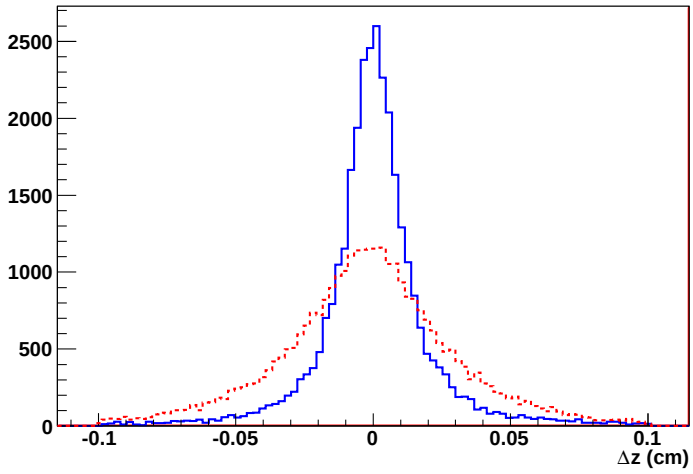
where $\mathcal{L}_S$ and $\mathcal{L}_{q\overline{q}}$ are the sum of the likelihoods extracted from the PDFs for signal and continuum respectively. If $|\Delta Z| < 0.1$cm then all three likelihoods are summed, if $|\Delta Z| > 0.1$cm or if one of the B vertex fits does not converge, then only $\cos\theta_B$ and $\mathcal{F}$ are used. The distribution of the $\mathcal{LR}$ is shown in Figure 3.8.



(a) The distribution of $\mathcal{F}$ in MC.



(b) The distribution of $\cos\theta_B$ in MC.



(c) The distribution of Δz in MC.

Figure 3.6: The distributions of the variables included in $\mathcal{LR}$. Signal MC is in dashed red and sideband data is in solid blue. The sideband data is normalised to the same number of events as the signal MC.

3.5.5 Figure of Merit

Where to place the cut on $\mathcal{LR}$ is determined by normalising the MC to the expected number of events in $604.55\ \mathrm{fb}^{-1}$, assuming the central values of the branching fractions of Babar's analysis of $B \to K\eta'\gamma$ (Equations 1.25 and 1.26), and optimising a Figure of Merit ($\mathcal{FOM}$) of the form

$$\mathcal{FOM} = \frac{\mathcal{N}_S}{\sqrt{\mathcal{N}_S + \mathcal{N}_{q\overline{q}}}} \tag{3.45}$$

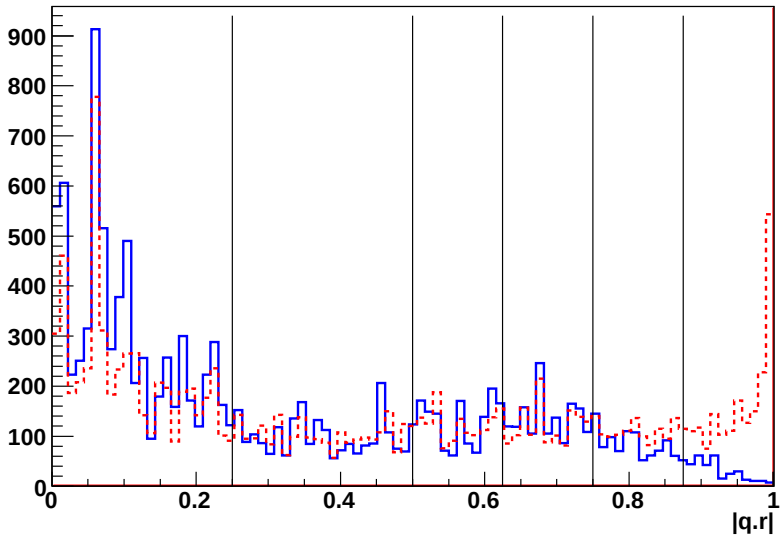
where $\mathcal{N}_S$ is the number of expected signal events that pass a particular $\mathcal{LR}$ cut and $\mathcal{N}_{q\overline{q}}$ is the number of expected continuum events passing the same cut.

The $\mathcal{FOM}$ is optimised for each final state individually in six bins of the flavour-tagging variables $|q.r|$ for neutral B decays and $(q.r.K_{charge})$ for charged. The flavour tagging module *Hamlet* [54] was developed for use in $\mathcal{CP}$ asymmetry analysis. The flavour of a neutral B meson decaying to a $\mathcal{CP}$ eigenstate ($\mathcal{B}_{\mathcal{CP}}$) is determined by reconstructing the final state of the second B ($\mathcal{B}_{tag}$). *Hamlet* takes all particles not used to construct the signal B candidate and tries to reconstruct one of several B decays which have highly recognisable flavour characteristics. The most probable flavour of $\mathcal{B}_{tag}$ is returned as the variable q and a tag quality factor r is assigned. The flavour is +1 for B^0 or B^+ and -1 for $\overline{B}^0$ or B^-, and r ranges from 0 for an event in which no flavour tagging was possible to 1 for an event in which $\mathcal{B}_{tag}$ was unambiguously tagged as a certain flavour.

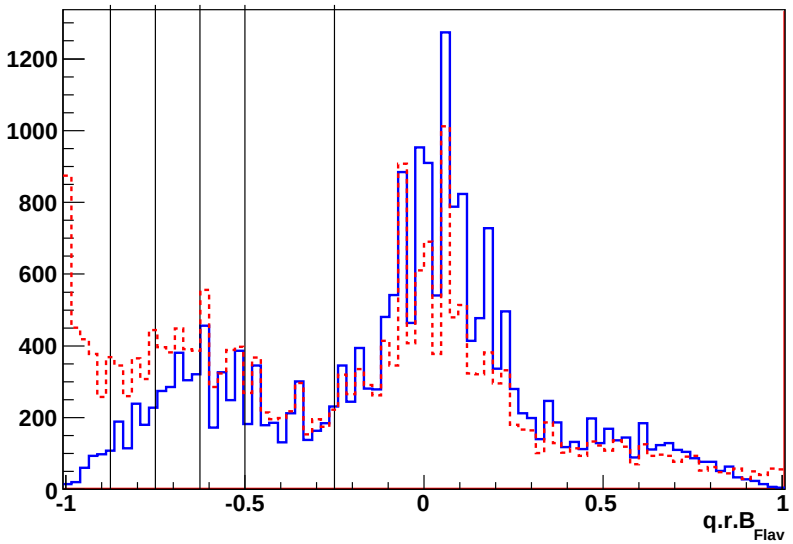
The tag-side B is not fully reconstructed. The estimation of flavour comes from a single charged particle: a lepton,kaon, Λ or slow pion. The charge of the high momentum lepton in semi-leptonic B decays provides the highest confidence tag, however the charge of the kaon in $B \to K^+X$ decays, the slow pion in $B \to D^{*-}X \to (\overline{D}^0\pi^-)X$ decays, and the $\overline{\Lambda}$ in $\overline{b} \to \overline{c} \to \overline{s}$ decays also provide useful information about the flavour of $\mathcal{B}_{tag}$.

Tagging *per se* is not relevant to this analysis. However, as continuum events are less likely to be successfully tagged by *Hamlet*, the distribution of r for $q\overline{q}$ events is markedly different from that of $B\overline{B}$ events. In addition, when the signal candidate flavour is also known- when the final state includes a charged kaon- the flavour of a successfully tagged $\mathcal{B}_{tag}$ is much more likely to be the opposite of the signal B in correctly reconstructed events. The variables q and r are based on the particle identification described above and have little dependence on event shape. The signal/continuum differences in q and r are utilised by dividing charged candidates into six bins of $(q.r.B_{Flav})$ and neutral candidates into six bins of $|q.r|$ and optimising the $\mathcal{FOM}$ for each bin separately. Figure 3.7 shows the distributions of $(q.r.B_{Flav})$ and $|q.r|$ for $K\eta'\gamma$ MC and sideband data and the optimisation bins for each variable.

The $\mathcal{FOM}$ is calculated in $\mathcal{LR}$ cut increments of 0.02, giving fifty points for possible $\mathcal{LR}$ cut placement. The point with the highest $\mathcal{FOM}$ is taken as the cut. Figure 3.8(b) shows the $\mathcal{FOM}$ points of the most negative $(q.r.B_{Flav})$ bin of $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$ over-laid on the $\mathcal{LR}$ distribution from which the $\mathcal{FOM}$ was calculated.

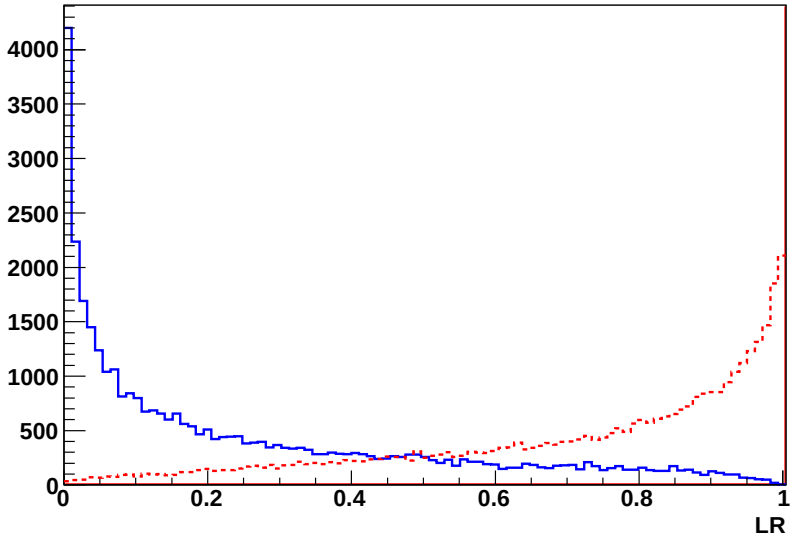


(a) The distributions of $|q.r|$.

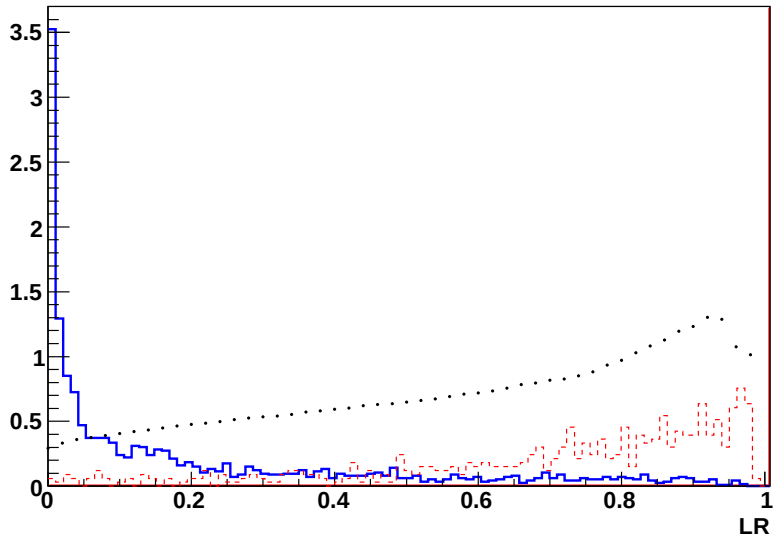


(b) The distributions of $(q.r.B_{Flav})$.

Figure 3.7: The distributions of $|q.r|$ and $(q.r.B_{Flav})$. $K\eta'\gamma$ MC is in red and sideband data is in blue. The optimisation bins are shown as black vertical lines.



(a) The distribution of $\mathcal{LR}$ for the 2D fitter.



(b) $\mathcal{FOM}$ and $\mathcal{LR}$ for $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$ events with $-1.0 < (q.r.B_{Flav}) < -0.875$. The y-axis scale is the magnitude of the $\mathcal{FOM}$.

Figure 3.8: The distribution of $\mathcal{LR}$ and $\mathcal{FOM}$ for the 2D fitter. $K\eta'\gamma$ MC is in red and sideband data is in blue.

Again, due to a lack of $q\overline{q}$ MC and sideband data statistics in the $\eta \to \pi^+\pi^-\pi^0$ final states, these modes are not divided into flavour tagging bins for $\mathcal{LR}$ cut optimisation. The optimisation for the charged and neutral $\eta \to \pi^+\pi^-\pi^0$ final states is performed separately.

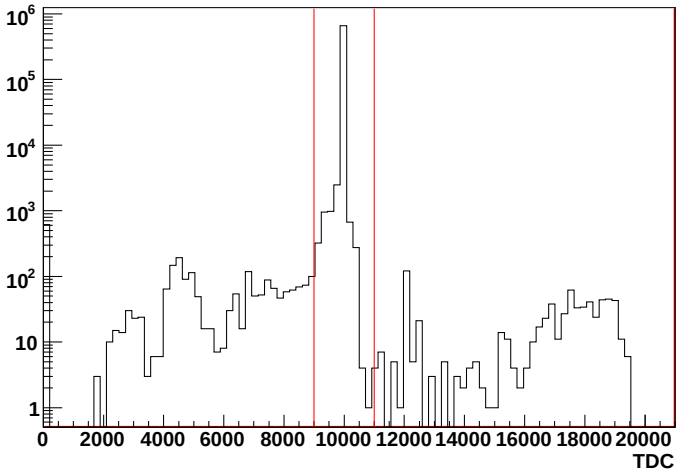
As previously mentioned, the three fitting methods described in the next Chapter have different selection criteria for $M_{\eta'}$ depending on whether it is included as a fit variable and how the fit is performed. The continuum suppression optimisation was performed separately for each of these sets of criteria. The process removes 98.15% of sideband data and passes 38.10% of the $K\eta'\gamma$ MC for the 2D fitter, and removes 96.38% of sideband data and passes 40.95% of $K\eta'\gamma$ MC for the 3D combined fitter. For the 3D mode-by-mode fitter, 97.23% of $q\overline{q}$ MC is removed, while 38.28% of $K\eta'\gamma$ MC is passed. Tables of the maximum $\mathcal{FOM}$ and the chosen $\mathcal{LR}$ cut for each optimisation region for all three fitting methods are shown in Appendix A.

3.6 Off-Time Backgrounds

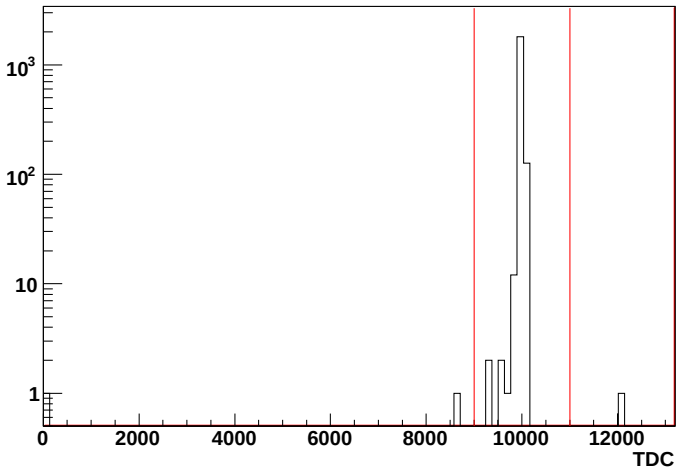
Previous analyses at Belle [48] [49] have found a non-negligible background arising from events in which some of the ECL crystals remain hot after the previous bunch-crossing. When a $e^+e^- \to e^+e^-$ (Bhabha) or $e^+e^- \to \gamma\gamma$ event is recorded in the detector, the trigger system discards it and no data is recorded. The high energy clusters these backgrounds deposit in the ECL have decay times of the order of micro-seconds (μs) however, making it possible for some of the energy in these clusters to be added to the ECL energy in the next beam crossing 2 nano-seconds (ns) later. This extra energy can act to shift a lower energy cluster from the next interaction into the signal region, or, as $e^+e^- \to e^+e^-$ and $e^+e^- \to \gamma\gamma$ events will produce very high energy ECL clusters, the decaying cluster energy can fall within the signal region by itself.

There are two ways of quantifying the off-time contributions at Belle: using the timing information of each Trigger Cell (a grouping of ECL cells that is used by the trigger), or the time-stamp of the Bhabha veto-trigger which is tripped by the detection of a Bhabha-like event in the ECL. The Trigger Cell timing (TDC) information includes all types of events removed by the trigger and is the most general test of the off-time contribution. The TDC information has only been recorded for data from experiment 39 onwards however. The off-time contribution will be assessed using the TDC information; if there is a non-negligible contribution a Bhabha veto-based cut will be required for data experiment 39 and earlier.

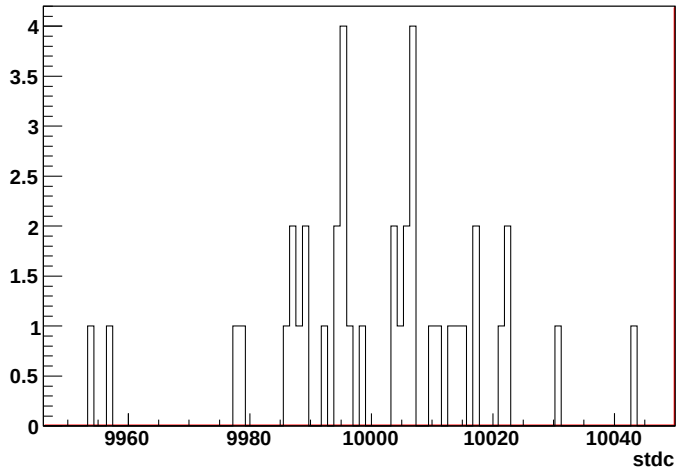
The event selection criteria were applied to the $24.52\ \mathrm{fb}^{-1}$ of off-resonance data from experiments 39, 41, 43, 47 and 49 and the timing information studied. To make the check as broad as possible, the timing spectrum under a much looser set of cuts was also investigated.



(a) TDC after loose selection.



(b) TDC after full selection.



(c) TDC after full selection and continuum supression.

Figure 3.9: TDC distributions for off-resonance data.

Figure 3.9 shows the TDC information for the signal γ cluster of the ECL, in arbitrary units. The red lines show the defined on-time region: $9000 \rightarrow 11000$. An event lying outside of this region is one in which the off-time energy component dominates the energy of the signal γ. The left plot is the TDC distribution under minimal cuts of $M_{\rm BC} > 5.0$ GeV/c^2, $|\Delta E| < 0.5$ GeV, 0.9 GeV/$c^2 < M_{\eta'} < 1.0$ GeV/c^2, 1.7 GeV $< E_\gamma < 3.4$ GeV, $E_9/E_{25} > 0.9$, loosened PID cuts and loosened ρ^0 and η mass cuts. The middle plot is the distribution after all the selection criteria described in Section 3.3, and the right-most plot of shows the distribution after all selection criteria and continuum suppression. Under the loose cuts 3110 events, or 0.45%, were outside the on-time region. After the full set of selection cuts 3 events, or 0.15% lay outside. Of these, one event lies within the signal box: this event has TDC=0, which means no timing information was available. None of the three events survive the continuum suppression cuts.

From this result the off-time background is assumed to be negligible.

3.7 Best candidate selection

After the selection cuts and continuum suppression are in place, multiple B candidates still remain in some events. Table 3.4 shows the average number of candidates for each event in $K\eta'\gamma$ MC after all selection criteria. The tighter selection criteria of the 2D analysis results in fewer multiple candidates than the either of the 3D analyses. The best B meson reconstruction is chosen based on the following:

- Vertex fits are performed using the charged particles in each of the candidates. K^0_S daughter pions are excluded from the fit. The candidate with the lowest B vertex χ^2 is taken as the preferred B.
- If two candidates have the same charged particles but different neutral particles, or the same charged particles in a different combination, they will have the same B vertex χ^2. In that case the lowest vertex χ^2 from the $\rho^0 \rightarrow \pi^+\ \pi^-$ or the $\eta' \rightarrow \eta\pi^+\pi^-$ is taken, depending on the reconstruction.
- Multiple candidates still remain in some events, but due to the previous condition they will either all be reconstructed with a ρ^0, or all with an η. In the first case the candidate with the highest energy photon from the $\eta' \rightarrow \rho^0\gamma$ decay is taken, and in the second case the lowest $\chi^2_{M_\eta}$ is taken, defined as

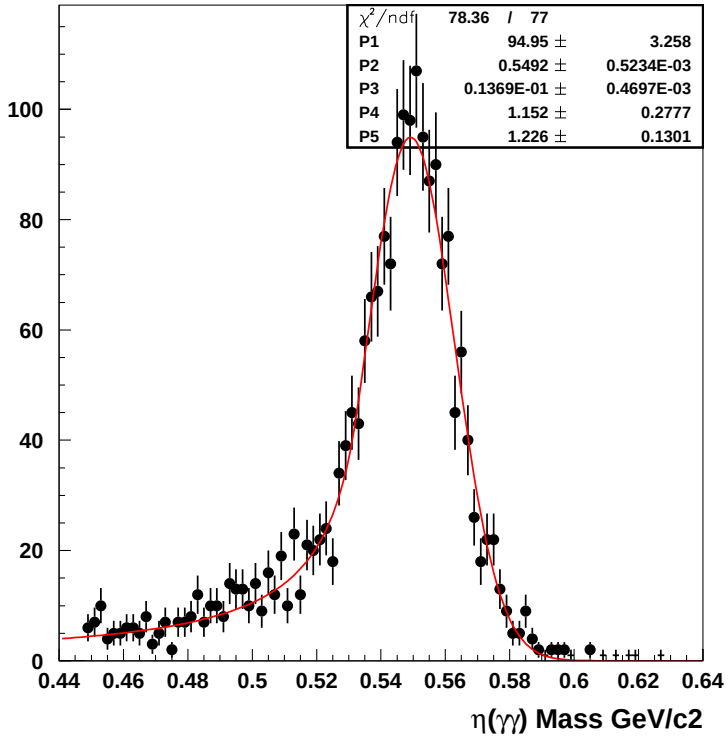
Final State	3D Mode-by-Mode Analysis	3D Combined Analysis	2D Analysis
$K^+\eta'(\rho^0\gamma)\gamma$	1.122	1.110	1.064
$K^0_S\eta'(\rho^0\gamma)\gamma$	1.125	1.113	1.089
$K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	1.217	1.280	1.137
$K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	1.216	1.305	1.143
$K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	1.633	1.805	1.482
$K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	1.697	1.849	1.499

Table 3.4: Mean number of candidates per event in signal MC in the fitting region and the signal box for the different sets of cuts.

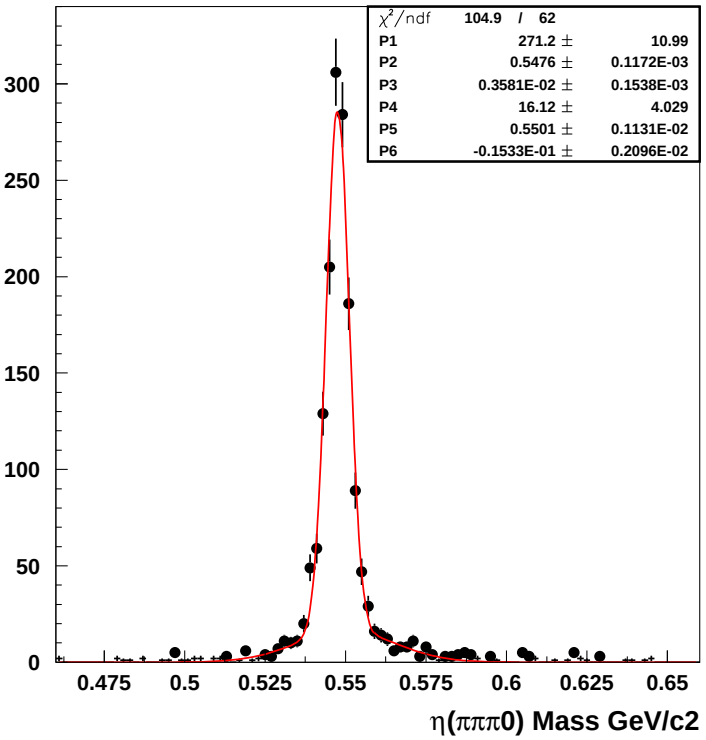
$$\chi^2_{M_\eta} = \frac{(M(\eta) - M_\eta)}{\sigma_\eta} \tag{3.46}$$

where $M(\eta)$ is the mass of the reconstructed η candidate, M_η is the nominal η mass [19], and σ_η is the width of the M_η peak from fitting correctly reconstructed η candidates in $K\eta'\gamma$ MC. Figure 3.10 shows the fits to $K\eta'\gamma$ MC to extract σ_η.

This technique selects the correct candidate in 76.2% of cases.



(a) Fit to $M_{\eta\to\gamma\gamma}$ distribution.



(b) Fit to $M_{\eta\to\pi^+\pi^-\pi^0}$ distribution.

Figure 3.10: Fits to M_η from correctly reconstructed $K\eta'\gamma$ MC.

Chapter 4

Data Analysis

4.1 Maximum Likelihood Formalism

The proportion of $B \rightarrow K\eta'\gamma$ decays in the events that pass selection is extracted by comparing the data distribution in several discriminating variables to an analytic function. The analytic function is a combination of probability density functions (PDFs) created from the expected distributions of $B \rightarrow K\eta'\gamma$ and all major backgrounds. PDFs are by definition normalised to one:

$$\int_a^b f(x)dx = 1 \tag{4.1}$$

where $f(x)$ is the PDF, and a and b define the fitting region. The PDFs are determined by unbinned maximum likelihood (ML) fits to Monte Carlo simulated data distributions. In the fitting process the parameters which define the functions are varied to determine the values which best describe the MC distribution. The vector of parameters ($\vec{\alpha}$) of the fitting function that best describe the distribution of a set of measurements (x_i) are chosen by maximising the likelihood function defined as

$$\mathcal{L}(\vec{\alpha}) = \Pi_i f(x_i;\vec{\alpha}) \tag{4.2}$$

where $f(x_i;\vec{\alpha})$ is the function describing the distribution. The $\mathcal{L}$ function quantifies the agreement between the data and the function. As $\mathcal{L}$ is often very small, its logarithm is more commonly used. The parameters are tuned in a series of steps to find the values that give the largest $\ln\mathcal{L}$, or equivalently the values that give the smallest $-\ln\mathcal{L}$. The minimisation is performed in this analysis by the numerical minimisation package Minuit [52].

The method of maximum likelihood was pioneered in 1912 by R. A. Fisher [51]. In this analysis it is chosen over the χ^2 method as the means of tuning parameters. Both methods compare function and data in defined bins of the chosen discriminating variables. When some bins in the distribution have very few

entries the Gaussian approximation to the Poisson distribution is no longer accurate, and events must be described by Poissonian statistics. The χ^2 method assumes Gaussian errors on the number of entries in each bin, and this can lead to inaccuracies and bias in the fitting process. The maximum likelihood method treats the bin errors as Poisson distributed. This makes it possible to take the limit $N_{bins} \to \infty$, in which all bins contain either one or zero events. No structural information is lost to the binning process in this "unbinned" maximum likelihood method, making it the technique of choice for low-statistics analyses.

Functions are fitted to the MC distributions of $B \to K\eta'\gamma$ and the three largest backgrounds: $q\overline{q}$, $b \to c$ and $b \to u, d, s$. These are combined into a single PDF which is used in an 'extended' unbinned ML fit to the data distribution. In an extended fit the function normalisations rather than the function parameters are optimised to values that best suit the distribution. The PDF is no longer normalised to 1 but to the total number of events in the fitted sample:

$$\int_a^b f(x)dx = \mathcal{N} \tag{4.3}$$

where $f(x)$ is the combined PDF and $\mathcal{N}$ is the total number of events. The fitter determines the most likely proportions of $\mathcal{N}$ attributed to signal or background events based on the shape of the PDFs and shape of the data distribution. In this way the signal yield- the number of $B \to K\eta'\gamma$ decays within the data sample- is extracted.

Once Minuit returns a stable minimum in the $-\ln\mathcal{L}$ function, the uncertainties in the parameter values are described by the width of $-\ln\mathcal{L}$ at the minimum. The error on each parameter is found by varying that parameter in a scan of $-\ln\mathcal{L}$ near the minimum. The change in the parameter value that gives an increase in $-\ln\mathcal{L}$ by 0.5 is the error at one standard deviation (σ). This method is designed to find the correct errors in all cases, including those in which non-linear fitting functions and/or low statistics cause the $-\ln\mathcal{L}$ function to be asymmetric around the minimum. This method is important for low statistics rare decay searches like this, as other error estimation tools which assume a symmetric $-\ln\mathcal{L}$ function can give inaccurate results. The scan and the calculation of these asymmetric errors is performed by MINOS, a sub-routine of Minuit. All the fit results and parameter errors given in this thesis are performed using this Minuit-MINOS technique.

Once the fit to data has been performed and the signal yield obtained, it is important to quantify the probability of obtaining this result relative to any other result- in particular, relative to the null hypothesis. This is called the 'significance' of the result, and the likelihood function gives a means of estimating it. The value of $\mathcal{L}$ at any two points in parameter space can be compared in terms of the standard deviation of a Gaussian by creating a variable that obeys a χ^2 distribution:

$$-2\ln(\mathcal{L}_0/\mathcal{L}_{\max}) \tag{4.4}$$

where $\mathcal{L}_{\max}$ is the maximum value of the likelihood function and $\mathcal{L}_0$ is the value of likelihood function

when the signal yield of the fit is fixed zero. The square root of this variable is the significance in units of σ:

$$\sqrt{-2\ln(\mathcal{L}_0/\mathcal{L}_{\max})} \tag{4.5}$$

The significance is a measure of how unlikely the result is to actually be the null hypothesis: the greater the difference between $\mathcal{L}$ and $\mathcal{L}_0$, the less likely the result is to be merely statistical fluctuation. If the significance of a hypothesis- eg, for a particular B decay having occurred in the sample- is 3σ (99.73% of the likelihood density) or more there is said to be 'evidence' for the hypothesis. For a decay to be considered to have been unequivocally observed it is usually required to have a significance of 5σ (99.9999%) or greater.

The discriminating variables used in the fit are chosen based on their mutual independence and the expected differences between the signal and background distributions. The kinematic variables defined in Equations 3.25 and 3.26, $M_{\rm BC}$ and ΔE, are chosen. ΔE is a variable based on the reconstructed energy of the B meson candidate and has width dictated by the momentum resolution of the final state particles. $M_{\rm BC}$ on the other hand is based on the B momentum and $E_{\rm beam}$. As previously mentioned the B meson has low momentum in the CoM frame; this momentum contributes little to the $M_{\rm BC}$ width, which is actually dominated by the spread of $E_{\rm beam}$. This makes $M_{\rm BC}$ nearly independent of final state particle momentum resolution and thus nearly independent of ΔE. This is less accurate when the final state includes photons, as the ECL momentum resolution is much lower than that of the CDC: ΔE-$M_{\rm BC}$ independence needs to be tested in radiative decay analyses such as this. As shown in Figure 3.4, the distributions of $B \to K\eta'\gamma$ and $q\overline{q}$ events are expected to be very different in these two variables.

The invariant mass of the reconstructed η' is also tested as a fitting variable. While all types of backgrounds include some real η' mesons, $B \to K\eta'\gamma$ events are expected to peak more strongly around the measured value of $M_{\eta'}$ and this could provide some additional discriminating power to the fit. The calculation of $M_{\eta'}$ includes some of the same momentum measurements as ΔE and there could be dependence between the two. The independence of all three fitting variables and the discriminating power of the $M_{\eta'}$ distribution are tested in the following Sections.

4.2 Correlation Study

The simplest way to fit a multi-dimensional distribution is with the product of factorised PDFs. Each PDF describes the distribution of one variable with no interdependence between the dimensions. Before attempting to use such functions to model $K\eta'\gamma$ and the backgrounds, the three prospective fit variables are checked for correlations. Any significant correlations between the fitting variables would compromise the ability of factorised PDFs to model the multi-dimensional distributions.

Correlations between variables in MC are assessed by calculating the mean of one variable in fine bins of the other. 2D histograms of the three variables $M_{\rm BC}$, ΔE and $M_{\eta'}$ in $K\eta'\gamma$ MC are shown in Figure 4.1. The means and their errors are shown as red points with error bars. Only correctly reconstructed events are shown in the plots: no correlations were found in randomly reconstructed events.

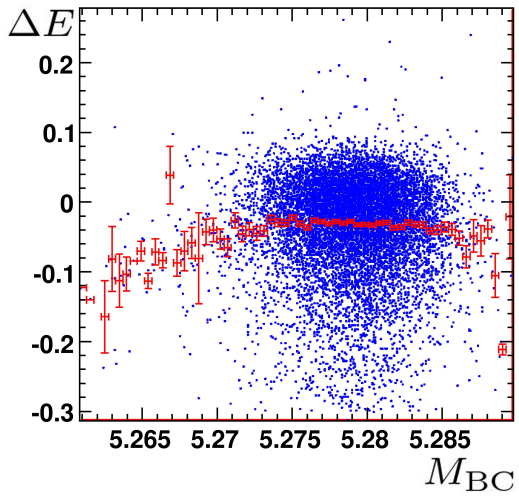
Correlations are seen as a shift in the mean of one variable in the scan across the other variable. In the $M_{\rm BC}$ vs. $M_{\eta'}$ plot (middle) no significant correlation is visible between the variables within the statistical scope of the study. There is some correlation visible the ΔE vs. $M_{\rm BC}$ plot (left)- the mean of ΔE can be seen to drop as $M_{\rm BC}$ increases. As mentioned before, this is due to the poor momentum resolution of the final state photons in both variable calculations. In the ΔE vs. $M_{\eta'}$ plot (right), definite correlation is visible as the mean value of $M_{\eta'}$ shifts higher with increasing ΔE. This correlation is not unexpected as the calculation of ΔE (Equation 3.26) includes the reconstructed η' four-vector. It is dealt with by implementing a mass-constrained fit on the candidate η' meson. The fit assumes that a real η' has been reconstructed, assigns it the nominal η' mass of $0.958\,{\rm GeV}/c^2$ [19], and re-calculates its momentum. The re-fitted momentum of the η' is used to calculate the modified variables $M_{\rm BC}^{mcf}$ and ΔE^{mcf}. The original unmodified η' momentum is used to calculate $M_{\eta'}$, removing the common factor. Throughout the remainder of this thesis references to $M_{\rm BC}$ and ΔE will refer to the modified variables $M_{\rm BC}^{mcf}$ and ΔE^{mcf} unless otherwise stated.

The effect of the mass-constrained fit on the correlations can be seen in Figure 4.2. The upwards trend of the $M_{\eta'}$ mean as ΔE^{mcf} increases has been greatly reduced, and now lies within the statistical errors of the mean. The $M_{\rm BC}$ vs. ΔE correlation also shows reduction as the effect of the common photon momentum measurements is reduced by the momentum recalculation.

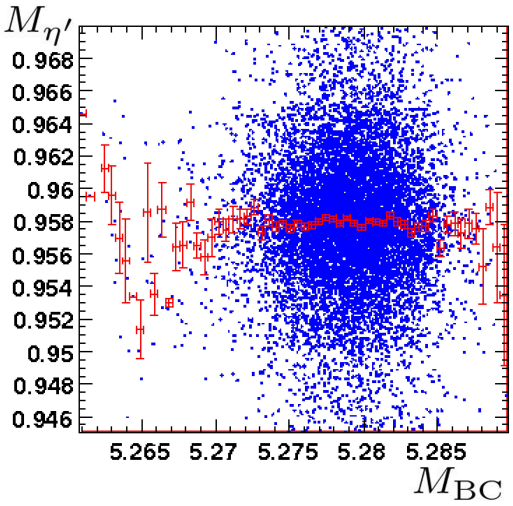
The correlation plots for the background MC samples are shown in Appendix B. All plots use the modified variables $M_{\rm BC}^{mcf}$ and ΔE^{mcf}. The $q\overline{q}$ MC shows no signs of correlation between $M_{\eta'}$ and either ΔE^{mcf} or $M_{\rm BC}^{mcf}$ before or after the continuum suppression and best candidate selection is applied. There is some suggestion of correlation in the $M_{\rm BC}^{mcf}$ vs. ΔE^{mcf} plots; it is within the statistical error of the study and is not adjusted for. The $b \to u, d, s$ MC is also uncorrelated within the statistical errors of the study, however previous analyses at Belle have found correlations between $M_{\rm BC}$ and ΔE, and these will be taken into account in the fitting procedure.

The $b \to c$ MC (Figure 4.3) shows correlations between $M_{\rm BC}$ and ΔE before the continuum suppres-

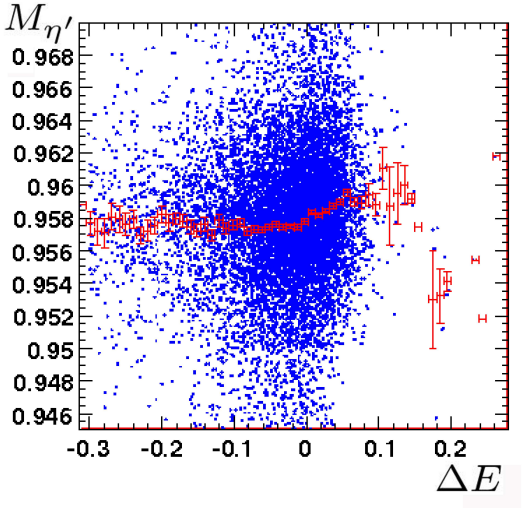
sion and best candidate selection is applied. After the full set of cuts the correlation is within the statistical error of the study (see Appendix B), but this correlation will also be adjusted for during the fitting procedure.



(a) $M_{\rm BC}$ vs. ΔE.

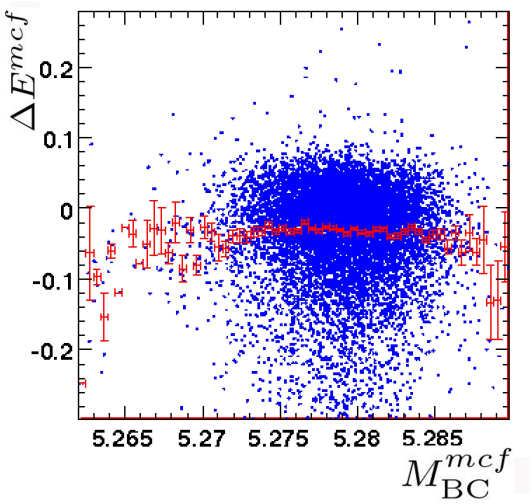


(b) $M_{\rm BC}$ vs. $M_{\eta'}$.

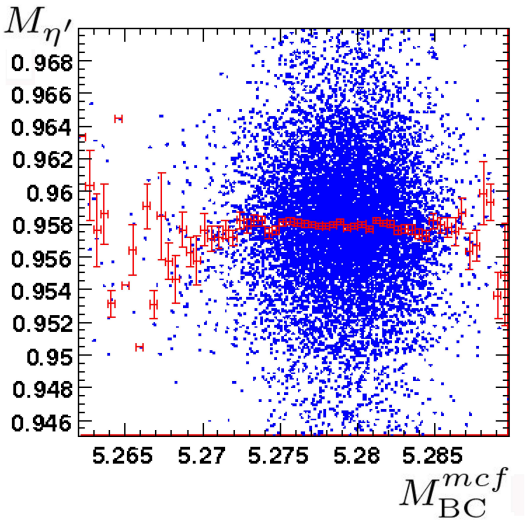


(c) ΔE vs. $M_{\eta'}$.

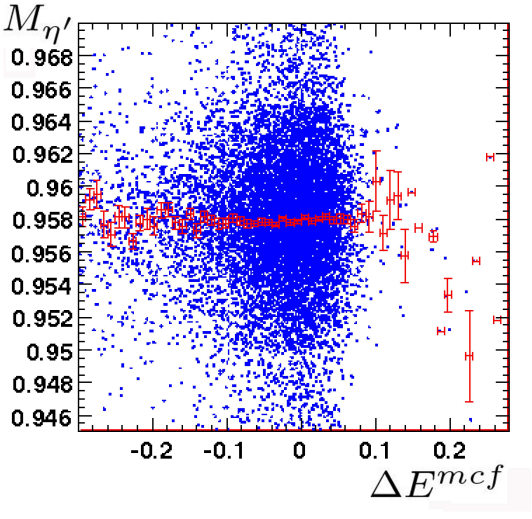
Figure 4.1: Scatter plots of the fi tting variables. The red points and error bars are the means of the y-axis variable in a range of the x-axis variable- 0.46 MeV bins for $M_{\rm BC}$, and 10 MeV bins for ΔE.



(a) $M_{\rm BC}^{mcf}$ vs. ΔE^{mcf}.

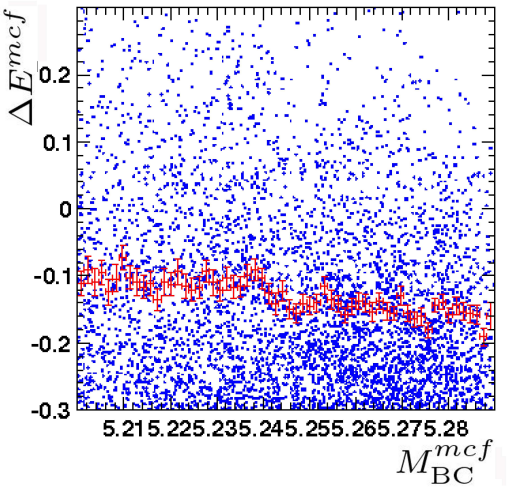


(b) $M_{\rm BC}^{mcf}$ vs. $M_{\eta'}$.

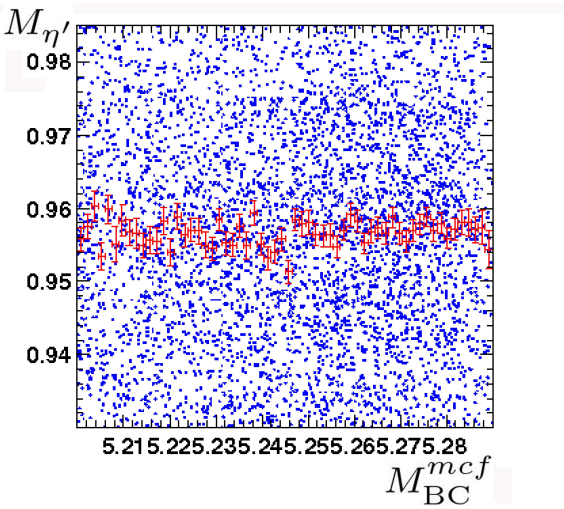


(c) ΔE^{mcf} vs. $M_{\eta'}$.

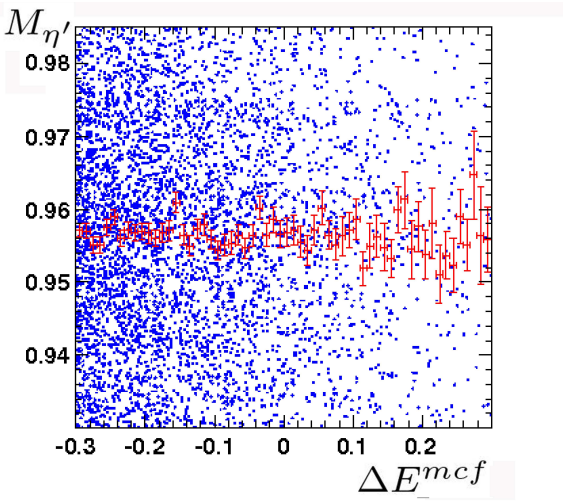
Figure 4.2: Scatter plots of the fi tting variables after the mass constrained fi t to the η' candidate has been applied.



(a) $M_{\rm BC}^{mcf}$ vs. ΔE^{mcf}.



(b) $M_{\rm BC}^{mcf}$ vs. $M_{\eta'}$.



(c) ΔE^{mcf} vs. $M_{\eta'}$.

Figure 4.3: Scatter plots of the fitting variables for $b \to c$ MC. The red points and error bars are the means of the y-axis variable in a range of the x-axis variable- 0.46 MeV bins for $M_{\rm BC}^{mcf}$, and 10 MeV bins for ΔE^{mcf}.

4.3 PDF modelling

Several methods of fitting the data were studied before a technique was decided upon. Two types of 3-dimensional fitter and a 2-dimensional fitter are compared for accuracy, stability and significance. The 3D fitters use an extended unbinned ML fit to $M_{\rm BC}$, ΔE and $M_{\eta'}$. The 2D fitter uses $M_{\rm BC}$ and ΔE with tight cuts on $M_{\eta'}$.

This section describes the process of modelling the MC using PDFs for each of the three fitting methods. MC samples are created with the signal-to-background ratios expected in the data set and fitted as a test of each fitting methods ability to accurately extract the correct numbers of $K\eta'\gamma$ and background MC events.

In Section 4.4 the three methods are tested for accuracy and stability more stringintly using larger numbers of scaled MC samples. In order for the fit results to be trusted the chosen fitting method should be as accurate as possible in reproducing the MC signal-to-background ratios. The stability of the fitting method is of even greater importance in chosing the correct method. If any of the methods failed to find a stable minimum in the likelihood function for a significant number of the trail fits they were deemed too unstable to be trusted to find a stable minimum in the fit to data.

4.3.1 3D Mode-by-mode Fitting Method

This technique involves fitting each of the six reconstructed final states (modes) individually in a 3D unbinned ML fit: three charged final states and three neutral giving six fits in total. The benefits of this method are the ability to tune the fitting ranges of $M_{\eta'}$ for each type of η' reconstruction, and individually calculated yields and statistical errors for each mode.

The global signal yields, $\mathcal{BF}$s and significances are calculated via a simultaneous fit to the charged final states and another to the neutral final states. A simultaneous fit sums the likelihood functions of the included individual fits and minimises the combination. In this way the information from each fit is included.

The fitting ranges for each variable are shown in Table 4.1, and the functions used to model each type of distribution are shown in Table 4.2.

The ΔE distribution of $K\eta'\gamma$ was modelled with a Crystal Ball line shape function (CBLS) [57] combined with a Gaussian function. A standard Gaussian function is defined as

Fit Variable	Final State	Fitting Range	Signal Region
$M_{\rm BC}$ (GeV/c^2)	All	$5.20 < M_{\rm BC} < 5.29$	$5.27 < M_{\rm BC} < 5.29$
ΔE (GeV)	All	$-0.3 < \Delta E < 0.3$	$-0.1 < \Delta E < 0.07$
$M_{\eta'}$ (GeV/c^2)	$K\eta'(\rho^0\gamma)\gamma$	$0.93 < M_{\eta'} < 0.985$	$0.945 < M_{\eta'} < 0.97$
	$K\eta'(\eta\pi^+\pi^-)\gamma$	$0.94 < M_{\eta'} < 0.975$	$0.95 < M_{\eta'} < 0.965$

Table 4.1: The fitting ranges and signal regions of each fitted variable for the 3D mode-by-mode method.

MC Type	Fit Variable	Function
$K\eta'\gamma$	$M_{\rm BC}$	CBLS
	ΔE	CBLS + Gaussian
	$M_{\eta'}$	1^{st} Order Chebyshev + BW
$q\overline{q}$	$M_{\rm BC}$	ARGUS
	ΔE	1^{st} Order Chebyshev
	$M_{\eta'}$	1^{st} Order Chebyshev $\pm$ BW, or Keys PDF
$b \rightarrow c$	$M_{\rm BC}$	ARGUS
	ΔE	2^{nd} Order Chebyshev
	$M_{\eta'}$	1^{st} Order Chebyshev $\pm$ BW
$b \rightarrow u, d, s$	$M_{\rm BC}$	} 2D Histogram PDF
	ΔE	
	$M_{\eta'}$	BW $\pm$ 1^{st} Order Chebyshev

Table 4.2: The functions used to model MC samples for the 3D mode-by-mode method. CBLS denotes a Crystal Ball line shape function, BW a Breit-Wigner function, and 1^{st} or 2^{nd} order Chebeyshev a 1^{st} or 2^{nd} order Chebyshev polynomial respectively. The 2D Histogram PDF runs over $M_{\rm BC}$ and ΔE.

$$f(x) = Ne^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \tag{4.6}$$

where N is the normalisation factor, μ is the mean and σ the width of the Gaussian. A CBLS function is a Gaussian with a power-law tail on one side and is used to describe radiative energy loss in an invariant mass distribution. It is defined as

$$f(x) = N \cdot \begin{cases} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} & , \frac{x-\mu}{\sigma} > -|\alpha| \\ (\frac{n}{|\alpha|})^n e^{-\frac{1}{2}|\alpha|^2} (\frac{n}{|\alpha|} - |\alpha| - \frac{x-\mu}{\sigma})^{-n} & , \frac{x-\mu}{\sigma} \leq -|\alpha| \end{cases} \tag{4.7}$$

where n defines the shape of the power-law function and α defines the point at which the Gaussian and power-law functions swap.

The combined CBLS and Gaussian functions have common means and relative widths: $\sigma_{gauss} = A \times \sigma_{CBLS}$, where A is a floated value. Separately floated means for the two functions would be highly correlated and might introduce bias into the fitting process. The dependency between the widths allows the use of a single calibration factor for A rather than one for each of σ when assessing the difference between the shapes of the ΔE distributions of MC and data (see Section 4.5).

The $M_{\rm BC}$ distribution of $K\eta'\gamma$ is also modelled using a CBLS.

The combinatorial background of the $M_{\eta'}$ distribution is modelled with a Chebychev polynomial of the 1^{st} order, which is combined with a Breit-Wigner function to model the correctly reconstructed η' mesons. Chebyshev polynomials describe the same shapes as standard polynomials but a different organisation of the power terms in the Chebyshev form reduces correlations between the coefficients and provides a more stable fit. The Breit-Wigner function describes the production probability distribution of a particle

resonance in terms of the center-of-mass energy or invariant mass:

$$f(x) = \frac{1}{(x-\mu)^2 + \frac{1}{4}g^2} \tag{4.8}$$

where x is the center-of-mass energy, μ is the mean invariant mass of the resonance and g is the width.

The product of the PDFs describing the three variables $M_{\rm BC}$, ΔE and $M_{\eta'}$ is used to fit the $K\eta'\gamma$ MC distributions in a 3D unbinned ML fit. The $K\eta'\gamma$ MC distributions and the shape of the functions after fitting the distributions is shown in Figure 4.4 for the final state $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$. The other fits and distributions are in Appendix C. All parameters in the functions as well the relative fractions of the combined functions are floated in the fits.

The $q\overline{q}$ ΔE distribution is modelled using a 1^{st} order Chebyshev polynomial. Another 1^{st} order Chebyshev polynomial is used to model the $M_{\eta'}$ distribution, combined with a Breit-Wigner when there is evidence of a true η' peak (the final states $K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$ and $K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$). These functions did not provide a good fit of the $K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$ $M_{\eta'}$ distribution and a Keys PDF [59] was used instead. A Keys PDF is a combination of Gaussians centred at each MC point, with width dependant on the surrounding MC density; areas with higher density have lower σ, those with lower density have higher σ.

The $M_{\rm BC}$ distribution is modelled using an ARGUS function [58]. The ARGUS function describes the probability density of randomly constructed invariant mass candidates in the region of a kinematic end-point. In this case the end-point is where the B candidate has zero momentum and $M_{\rm BC} = E_{\rm beam} = 5.29\,{\rm GeV}/c^2$. The ARGUS function is defined as

$$f(x) = Nx\sqrt{1-(\frac{x}{x_0})^2}\, e^{\xi(1-(\frac{x}{x_0})^2)} \tag{4.9}$$

where N is the function normalisation, ξ is the shape parameter and x_0 is the end-point of the function.

The product of these PDFs is used to fit the $q\overline{q}$ MC distributions in a 3D unbinned ML fit. The $q\overline{q}$ MC distributions and the shape of the fitted functions is shown in Figure 4.5 for the final state $B^+ \to K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$. The other fits and distributions are in Appendix C. All parameters in the functions are floated in the fits except the end-point of the ARGUS function which is fixed to $5.29\,{\rm GeV}/c^2$.

The ΔE distribution of $b \to c$ events is modelled with a 2^{nd} order Chebyshev polynomial. The $M_{\rm BC}$ distribution is modelled with an ARGUS function, and the $M_{\eta'}$ distribution with either a 1^{st} order Chebyshev polynomial or a Breit-Wigner function or a combination of the two depending on the shape of the distribution. The product of these PDFs is used in a 3D unbinned ML fit to the $b \to c$ MC. The MC distributions and the shape of the fitted functions is shown in Figure 4.6 for the final state $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$. The other fits and distributions are in Appendix C. All parameters in the functions are floated in the fits except the end-point of the ARGUS function which is fixed to $5.29\,{\rm GeV}/c^2$.

The correlations found between $M_{\rm BC}$ and ΔE in the previous section are not modelled in the $b \to c$

functions. The number of expected $b \to c$ events in each fit is low enough that the correlations do not have a significant impact on the results. The $M_{\rm BC}$ vs. ΔE plot shown in Figure 4.3 includes all final states: each fit will have approximately one sixth of the events shown.

The $M_{\rm BC}$ and ΔE distributions of rare B decays ($b \to u, d, s$) are modelled with a 2D histogram PDF to compensate for known correlations between the two variables for this background. The PDF takes its shape directly from a 2D histogram of the $M_{\rm BC}$:ΔE distribution. The $M_{\eta'}$ distribution is fitted with either a Chebyshev polynomial or a Breit-Wigner function or a combination of the two. The product of these PDFs is used in a 3D unbinned ML fit to the $b \to u, d, s$ MC. The MC distributions and the shape of the fitted functions is shown in Figure 4.7 for the final state $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$. The 2D histogram PDF is also shown in two dimensions. The other fits and distributions are in Appendix C.

For each final state, the 3D PDFs for $K\eta'\gamma$ and the three backgrounds were summed to form the final 3D PDFs that are tested for suitablity for fitting the data.

Scaled MC Trial for the 3D Mode-by-Mode Fitting Method.

Scaled MC samples that resemble the expected data distributions were assembled and fitted as a test of the accuracy of the fitter in reproducing the number of $K\eta'\gamma$ and background events in each trail sample. This is not a statistically significant test of the fiting method, however a stable accurate result shows the method is suitible to be tested further.

The number of events expected in $700\,\mathrm{fb}^{-1}$ of data for each of the MC types was randomly selected from the full MC samples. The level of integrated luminosity was chosen to be between the $605\,\mathrm{fb}^{-1}$ available for analysis and the total amount of data recorded at the time these tests were conducted. As the research progressed it became clear that $605\,\mathrm{fb}^{-1}$ of data would be all that would be available. Later tests of the fitting methods are normalised to this lower level (see Section **??**).

Signal MC was scaled using the formula

$$\text{Expected Events} = (\#B\text{s in } 700\,\mathrm{fb}^{-1}) \times \mathcal{BF}(B \to K\eta'\gamma) \times \prod \times \varepsilon \tag{4.10}$$

where the daughter $\mathcal{BF}$ ($\prod$) is the proportion of the total decay chain that is reconstructed (as detailed in Section 3.1.2), the efficiency (ε) is the proportion of signal MC events of that final state that pass the selection criteria, and central values of BABAR's results from their $B \to K\eta'\gamma$ analysis [53] are used for $\mathcal{BF}$ ($B \to K\eta'\gamma$):

$$\mathcal{B}(B^+ \to K^+\eta'\gamma) = (1.9^{+1.5}_{-1.2} \pm 0.1) \times 10^{-6} \tag{4.11}$$

$$\mathcal{B}(B^0 \to K^0_S\eta'\gamma) = (1.1^{+2.8}_{-2.0} \pm 0.1) \times 10^{-6} \tag{4.12}$$

The daughter $\mathcal{BF}$s and the counted and fitted $K\eta'\gamma$ MC εs are shown in Table 4.3. The counted signal ε is the proportion of $K\eta'\gamma$ MC events that pass the selection criteria. The expected number of events in 700 fb^{-1} of data is also shown for each $K\eta'\gamma$ mode. The number of $K\eta'\gamma$ MC events included in each scaled sample is the expected events rounded to the nearest whole number, except for $B^0 \to K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$ which is rounded to 1.

Final State	Daughter $\mathcal{BF}$	ε	Expected Events
$K^+\eta'(\rho^0\gamma)\gamma$	0.295	0.03116	12.27
$K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	0.175	0.05074	11.80
$K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	0.101	0.02223	2.99
$K^0_S\eta'(\rho^0\gamma)\gamma$	0.102	0.01611	1.27
$K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	0.060	0.02262	1.05
$K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	0.035	0.00932	0.25

Table 4.3: The daughter $\mathcal{BF}$ s of the reconstructed final states, and the counted $K\eta'\gamma$ MC ε, the fitted $K\eta'\gamma$ MC ε and expected number of $K\eta'\gamma$ events in 700 fb^{-1} for the 3D mode-by-mode method.

The background MC samples were scaled to 700 fb^{-1} according to their equivalent integrated luminosities: 1703.98 fb^{-1} of $q\overline{q}$ MC, 2846.11 fb^{-1} of $b \to c$ MC, and 24,850 fb^{-1} of $b \to u, d, s$ MC. The background and $K\eta'\gamma$ scaled MC samples were combined and fitted in an extended unbinned ML fit using the 3D PDFs described above. All parameters were fixed except the $q\overline{q}$ M_{BC} ARGUS shape parameter, the $q\overline{q}$ ΔE Chebyshev parameter, and the $q\overline{q}$ $M_{\eta'}$ Chebyshev parameter. Floating these parameters more accurately models the $q\overline{q}$ distribution, increasing the accuracy of the fit and lowering the systematic errors involved in the fitting process.

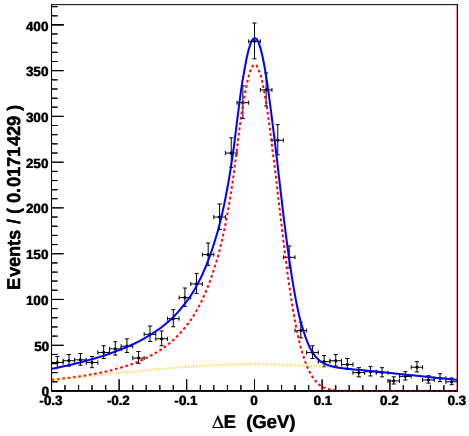
The normalisations of each component of the total PDF can be floated in the extended ML fit and the most likely numbers of $K\eta'\gamma$, $q\overline{q}$, $b \to c$ and $b \to u, d, s$ events extracted. The fitter has trouble distinguishing between the $b \to c$ and $q\overline{q}$ background components however, as they are modelled using similar functions. If both yields are allowed to float this leads to a result with unacceptably large errors on the yields of the two components. To solve this problem the $b \to c$ yield was fixed to the number of events expected from the MC scaling. The $b \to u, d, s$ is a relatively small component and its yield was also fixed to the number expected from the MC scaling calculation. The effect of fixing these components on the measured signal yield is examined in Section 5.2.

The results of the trial fits to scaled MC for each final state are shown in Figures 4.8 to 4.13. Also shown are the distributions of each variable and the fitted PDFs after projection into the signal region of the other two variables, as defined in Table 4.1. The projections help to visualise the signal component. The number of MC events in the distributions and the yields returned by the fits for each component are shown in Tables 4.4 to 4.9. Also shown is the calculated significance of the signal component, as defined in Equation 4.5.

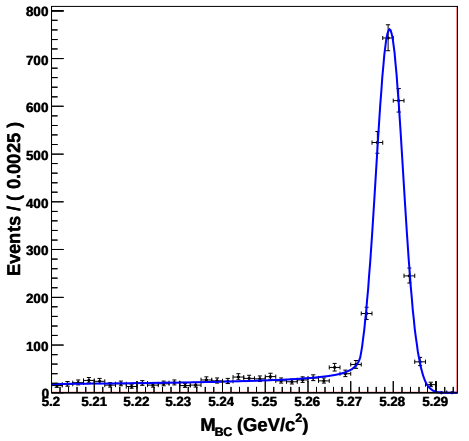
The fits show a tendency to over-estimate both the signal yield and the $q\overline{q}$ yield, though the results are well within statistical error of the true numbers of MC events. The final state expected to have the best signal to background ratio, $B^+ \to K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$, is the most accurate result with the highest fit significance.

Simultaneous fits to the scaled MC were performed for the charged modes and the neutral modes. The two simultaneous fits combine the likelihood functions of the three included final states and fit the distributions of each simultaneously by minimising the combined $-\ln\mathcal{L}$. The returned yields are shown in Tables 4.10 and 4.11. The fitted distributions are shown in Appendix C. The fits found stable minima in the likelihood function and reproduced the input MC yields well. The significance of the fit to the charged modes is 4.19σ, which would be the first discovery of the $B \to K\eta'\gamma$ decay if it were to be reproduced in the fit to data.

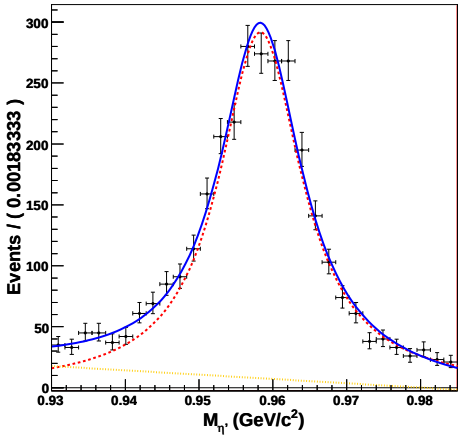
Further tests of the 3D mode-by-mode fitting method are detailed in Section 4.4.



(a) ΔE distribution for $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$ events in signal MC. The CBLS function is in dashed red, the Gaussian in dotted orange, and the combined function in solid blue.

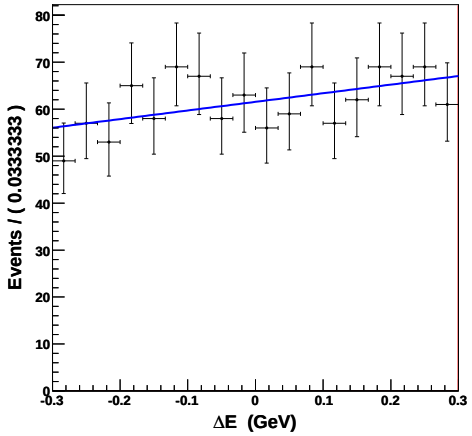


(b) $M_{\rm BC}$ distribution for $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$ events in signal MC.

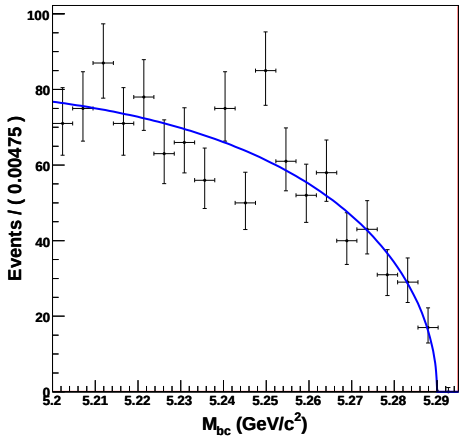


(c) $M_{\eta'}$ distribution for $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$ events in signal MC. The Breit-Wigner function is in dashed red, the Chebyshev polynomial in dotted yellow and the combined function in solid blue.

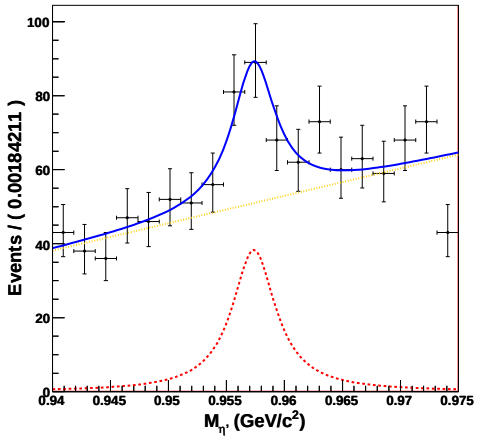
Figure 4.4: Signal MC modelling for the 3D mode-by-mode fit of the final state $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$.



(a) ΔE distribution for $B^+ \rightarrow K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$ events in $q\overline{q}$ MC.

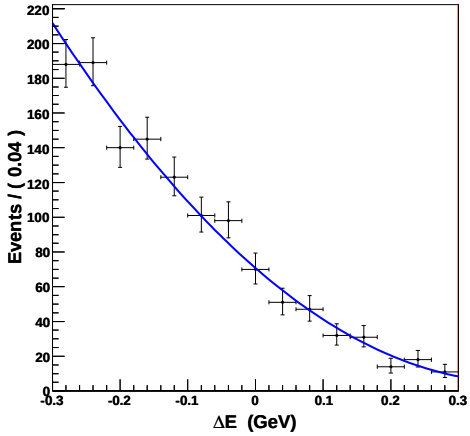


(b) $M_{\rm BC}$ distribution for $B^+ \rightarrow K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$ events in $q\overline{q}$ MC.

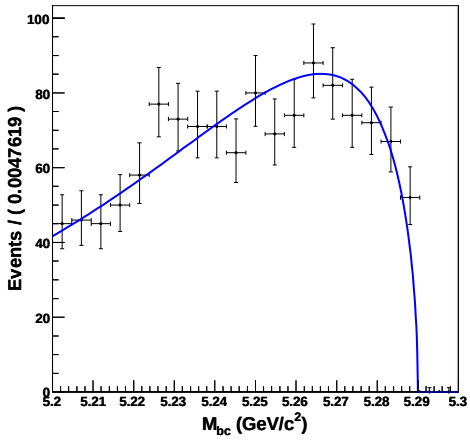


(c) $M_{\eta'}$ distribution for $B^+ \rightarrow K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$ events in $q\overline{q}$ MC. The Breit-Wigner function is in dashed red, the Chebyshev polynomial in dotted yellow and the combined function in solid blue.

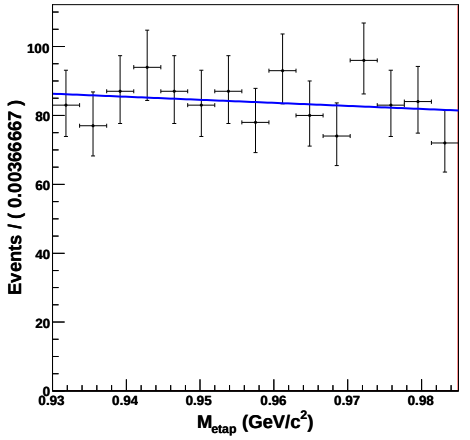
Figure 4.5: $q\overline{q}$ MC modelling for the 3D mode-by-mode fit of the final state $B^+ \rightarrow K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$.



(a) ΔE distribution for $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$ events in $b \to c$ MC.

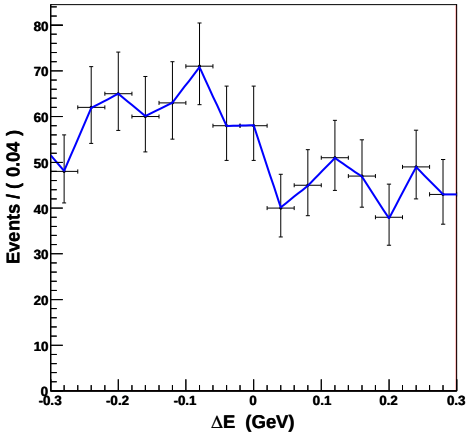


(b) $M_{\rm BC}$ distribution for $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$ events in $b \to c$ MC.

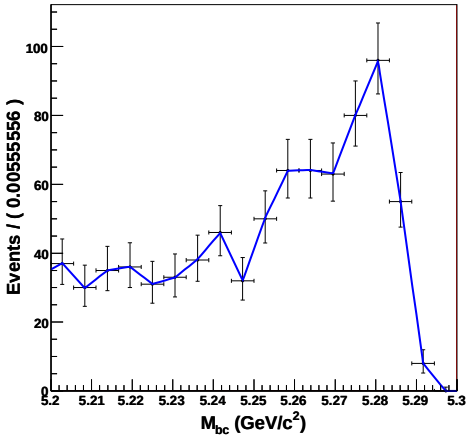


(c) $M_{\eta'}$ distribution for $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$ events in $b \to c$ MC.

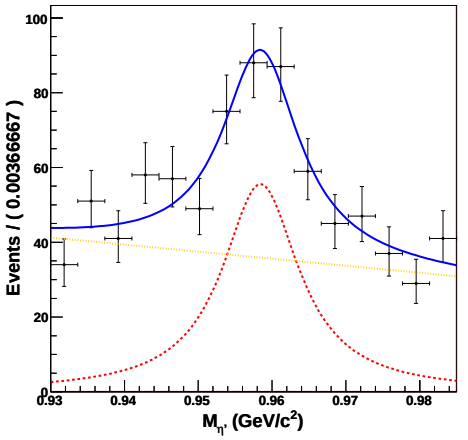
Figure 4.6: $b \to c$ MC modelling for the 3D mode-by-mode fit of the final state $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$.



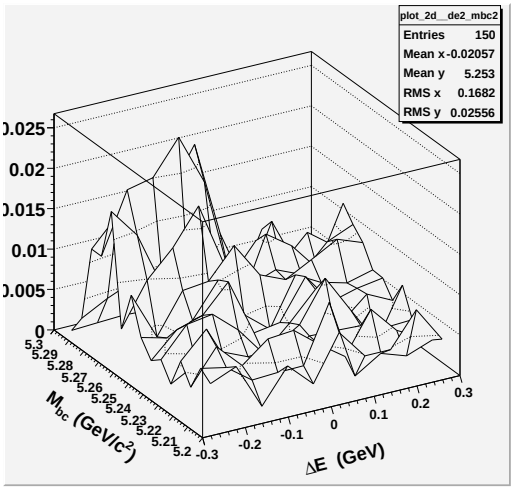
(a) ΔE distribution for $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$ events in $b \to u, d, s$ MC.



(b) $M_{\rm BC}$ distribution for $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$ events in $b \to u, d, s$ MC.



(c) $M_{\eta'}$ distribution for $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$ events in $b \to u, d, s$ MC. The Breit-Wigner function is in dashed red, the Chebyshev polynomial in dotted yellow and the combined function in solid blue.



(d) 2D $M_{\rm BC}$ vs. ΔE distribution for $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$ events in $b \to u, d, s$ MC.

Figure 4.7: $b \to u, d, s$ MC modelling for the 3D mode-by-mode fit of the final state $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$.

Scaled MC Fitting Trial to $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$.

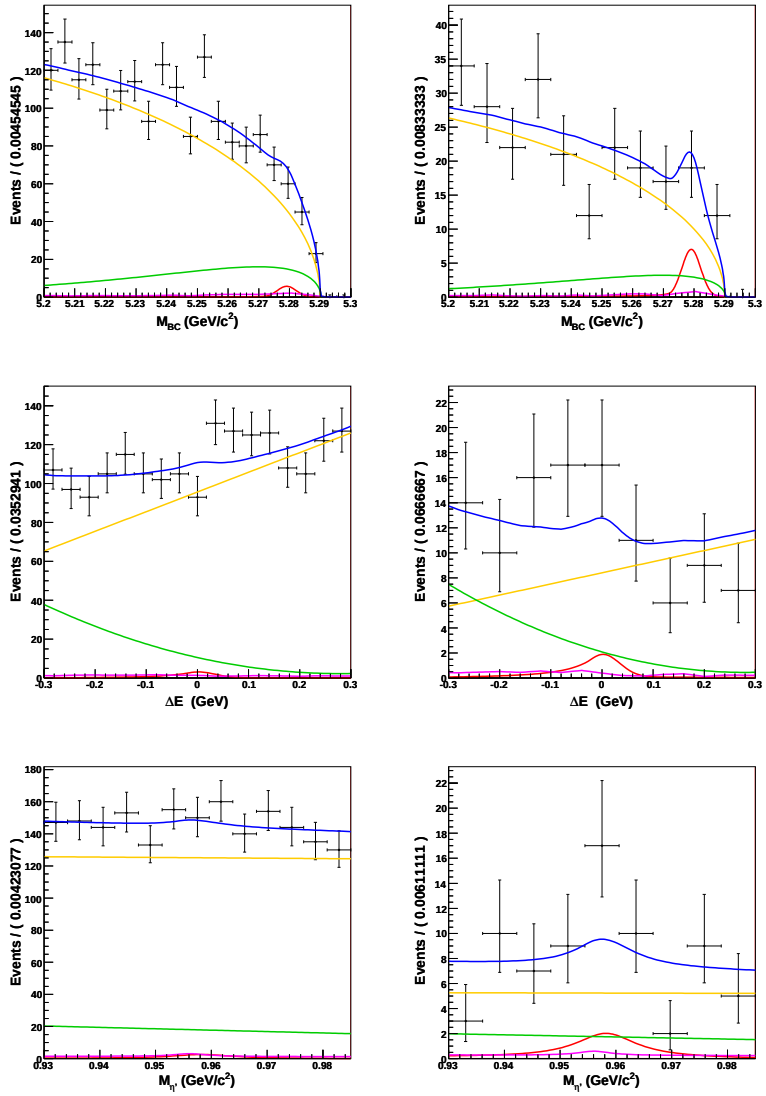


Figure 4.8: The scaled MC fitting trial to $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$. From top to bottom, the rows show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. The left column is the full fitting region, the right is the projections to the signal region defined in Table 4.1. The $K\eta'\gamma$ function is shown in red, $q\overline{q}$ in orange, $b \to c$ in green, $b \to u, d, s$ in magenta, and the combined function in blue.

MC Type	Input Events	Fit Yield	Significance
$K\eta'\gamma$	12	$15.180^{+11.501}_{-9.953}$	1.612σ
$q\overline{q}$	2501	$2576.7^{+54.442}_{-53.604}$	
$b \to c$	309	Fixed	
$b \to u, d, s$	22	Fixed	

Table 4.4: The input MC events, fit yields and fit significance for the scaled MC fitting trial to $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$.

Scaled MC Fitting Trial to $B^0 \to K^0_S\eta'(\rho^0\gamma)\gamma$.

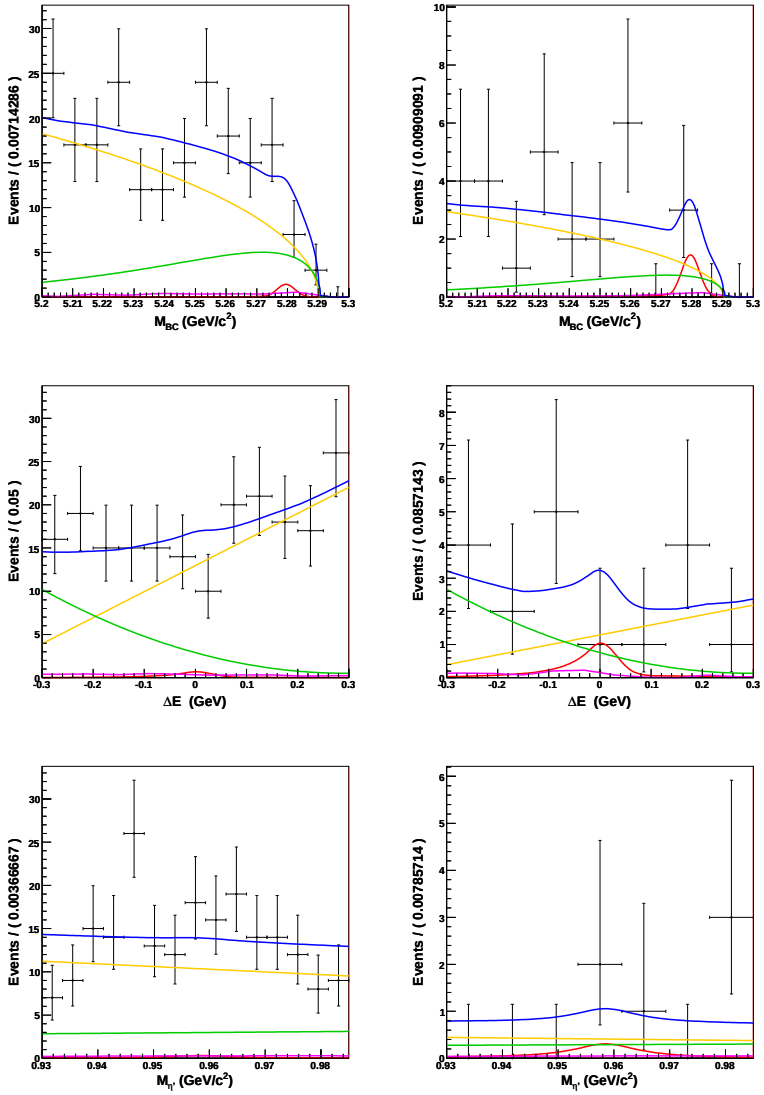


Figure 4.9: The scaled MC fitting trial for $B^0 \to K^0_S\eta'(\rho^0\gamma)\gamma$. From top to bottom, the rows show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. The left column is the full fitting region, the right is the projections to the signal region defined in Table 4.1. The $K\eta'\gamma$ function is shown in red, $q\overline{q}$ in orange, $b \to c$ in green, $b \to u, d, s$ in magenta, and the combined function in blue.

MC Type	Input Events	Fit Yield	Significance
$K\eta'\gamma$	1	$1.6262^{+3.4182}_{-2.2871}$	0.6633σ
$q\overline{q}$	239	$256.35^{+17.724}_{-17.056}$	
$b \to c$	59	Fixed	
$b \to u, d, s$	4	Fixed	

Table 4.5: The input MC events, fit yields and fit significance for the scaled MC fitting trial to $B^0 \to K^0_S\eta'(\rho^0\gamma)\gamma$.

Scaled MC Fitting Trial to $B^+ \to K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$.

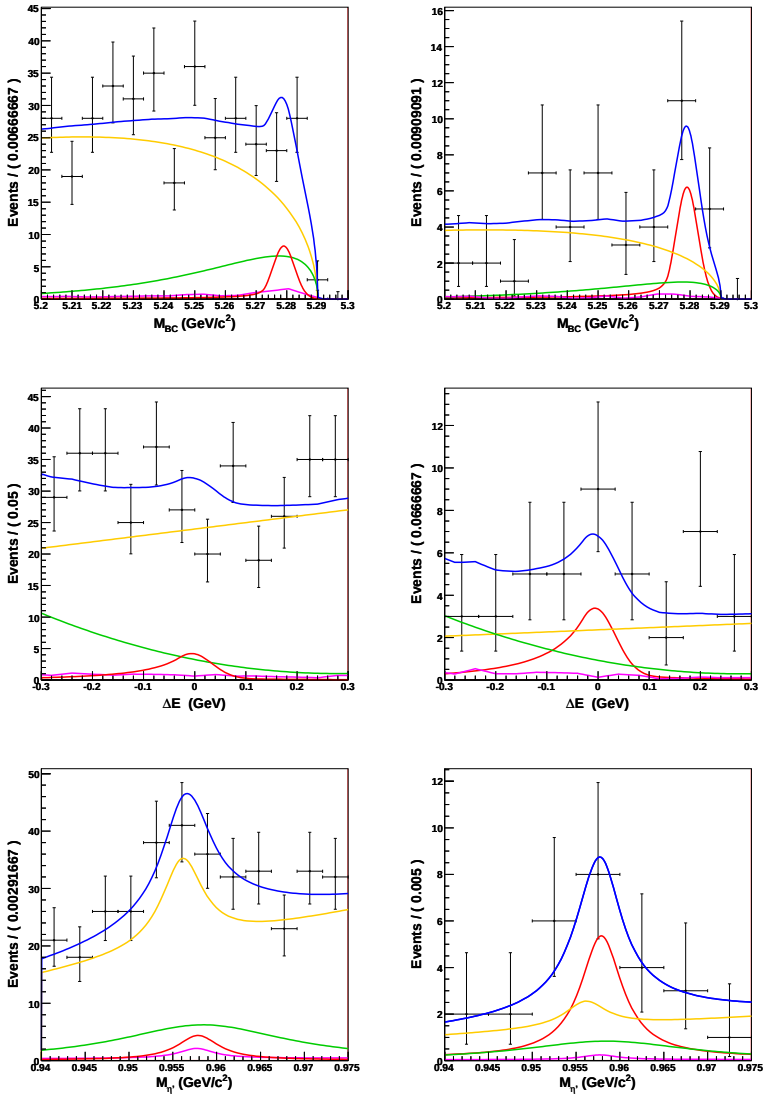


Figure 4.10: The scaled MC fitting trial for $B^+ \to K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$. From top to bottom, the rows show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. The left column is the full fitting region, the right is the projections to the signal region defined in Table 4.1. The $K\eta'\gamma$ function is shown in red, $q\overline{q}$ in orange, $b \to c$ in green, $b \to u, d, s$ in magenta, and the combined function in blue.

MC Type	Input Events	Fit Yield	Significance
$K\eta'\gamma$	12	$12.887^{+6.7126}_{-5.6951}$	2.709σ
$q\overline{q}$	455	$453.57^{+23.250}_{-22.461}$	
$b \to c$	67	Fixed	
$b \to u, d, s$	9	Fixed	

Table 4.6: The input MC events, fit yields and fit significance for the scaled MC fitting trial to $B^+ \to K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$.

Scaled MC Fitting Trial to $B^0 \to K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$.

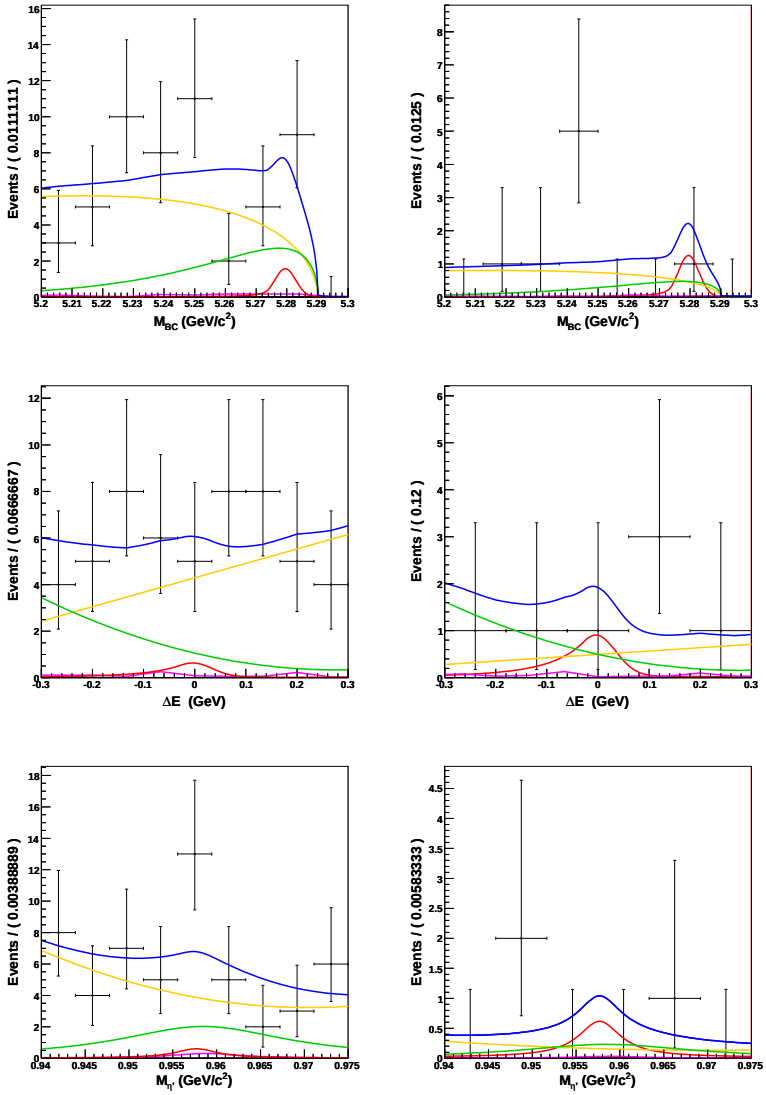


Figure 4.11: The scaled MC fitting trial for $B^0 \to K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$. From top to bottom, the rows show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. The left column is the full fitting region, the right is the projections to the signal region defined in Table 4.1. The $K\eta'\gamma$ function is shown in red, $q\overline{q}$ in orange, $b \to c$ in green, $b \to u, d, s$ in magenta, and the combined function in blue.

MC Type	Input Events	Fit Yield	Significance
$K\eta'\gamma$	1	$1.2362^{+3.4664}_{-2.4199}$	0.3561σ
$q\overline{q}$	39	$40.738^{+7.9134}_{-7.1741}$	
$b \to c$	12	Fixed	
$b \to u, d, s$	1	Fixed	

Table 4.7: The input MC events, fit yields and fit significance for the scaled MC fitting trial to $B^0 \to K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$.

Scaled MC Fitting Trial to $B^+ \to K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$.

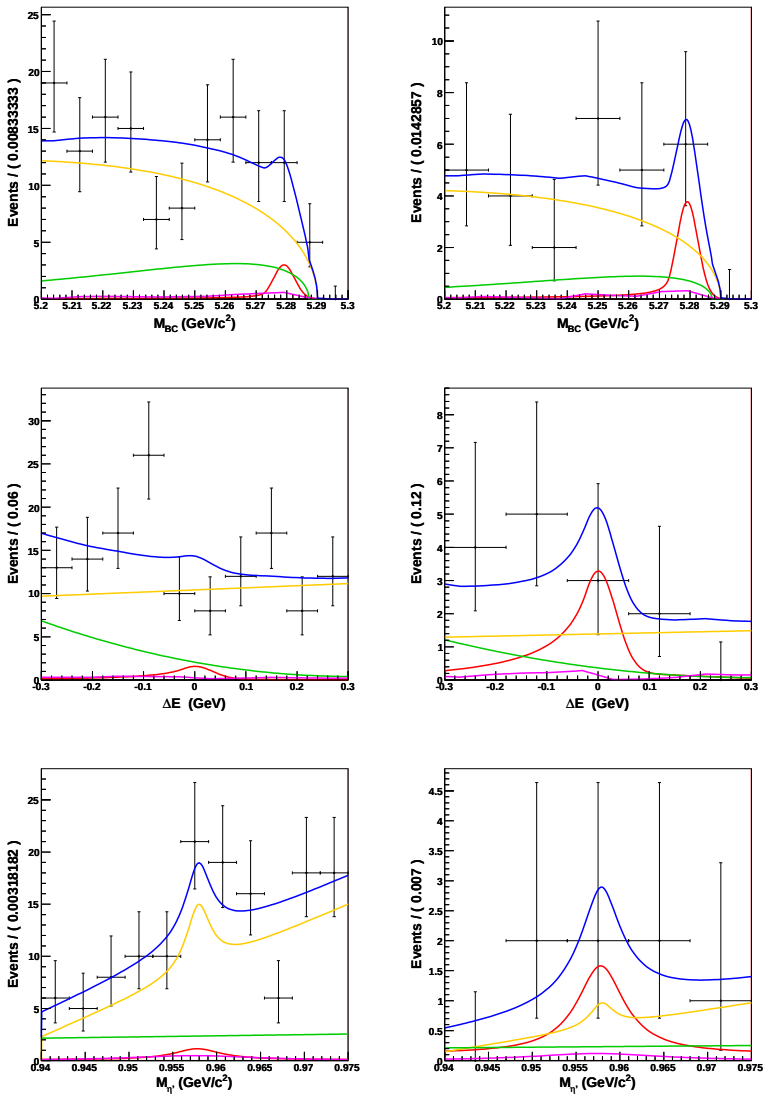


Figure 4.12: The scaled MC fitting trial for $B^+ \to K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$. From top to bottom, the rows show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. The left column is the full fitting region, the right is the projections to the signal region defined in Table 4.1. The $K\eta'\gamma$ function is shown in red, $q\overline{q}$ in orange, $b \to c$ in green, $b \to u, d, s$ in magenta, and the combined function in blue.

MC Type	Input Events	Fit Yield	Significance
$K\eta'\gamma$	3	$3.9045^{+4.5989}_{-3.4202}$	1.183σ
$q\overline{q}$	106	$104.76^{+12.337}_{-11.674}$	
$b \to c$	26	Fixed	
$b \to u, d, s$	3	Fixed	

Table 4.8: The input MC events, fit yields and fit significance for the scaled MC fitting trial to $B^+ \to K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$.

Scaled MC Fitting Trial to $B^0 \to K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$.

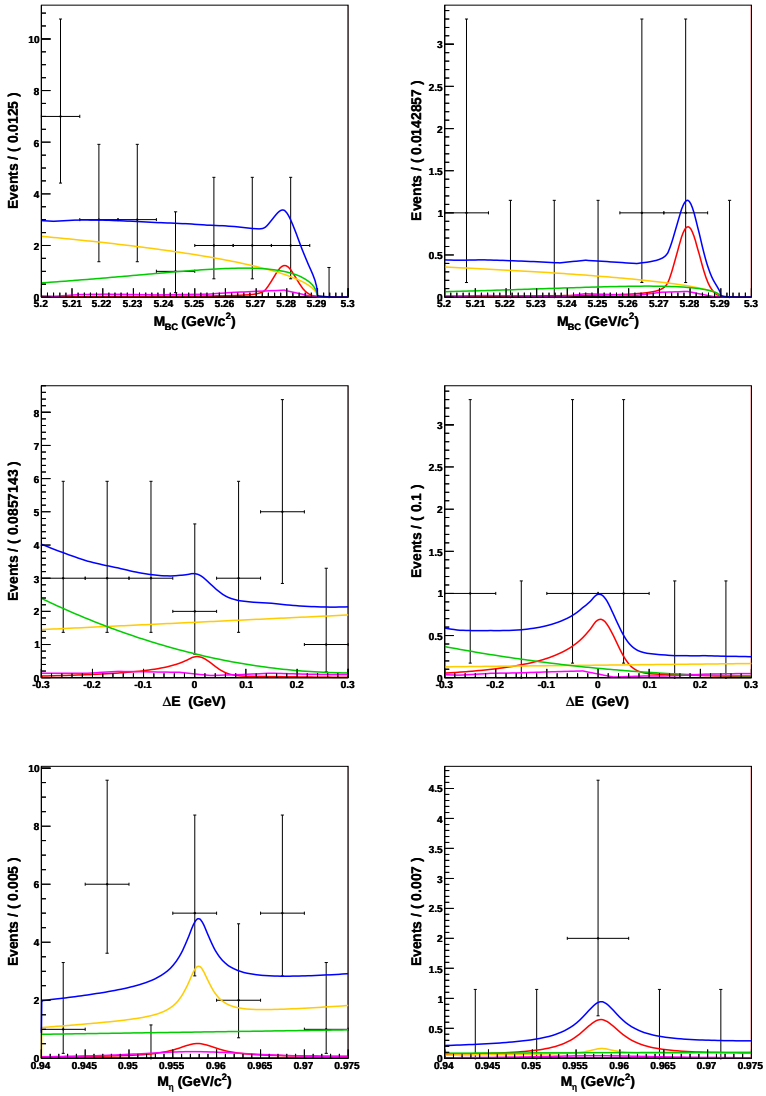


Figure 4.13: The scaled MC fitting trial for $B^0 \to K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$. From top to bottom, the rows show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. The left column is the full fitting region, the right is the projections to the signal region defined in Table 4.1. The $K\eta'\gamma$ function is shown in red, $q\overline{q}$ in orange, $b \to c$ in green, $b \to u, d, s$ in magenta, and the combined function in blue.

MC Type	Input Events	Fit Yield	Significance
$K\eta'\gamma$	1	$1.3239^{+1.9695}_{-1.1089}$	1.361σ
$q\overline{q}$	11	$12.942^{+4.6199}_{-3.9829}$	
$b \to c$	7	Fixed	
$b \to u, d, s$	1	Fixed	

Table 4.9: The input MC events, fit yields and fit significance for the scaled MC fitting trial to $B^0 \to K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$.

Mode	MC Type	Input Events	Fit Yield	Significance
$K^+\eta'(\rho^0\gamma)\gamma$	$K\eta'\gamma$	12	$15.001^{+7.596}_{-12.320}$	4.19σ
	$q\overline{q}$	2501	$2515.1^{+59.520}_{-53.864}$	
	$b \to c$	309	Fixed	
	$b \to u, d, s$	22	Fixed	
$K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	$K\eta'\gamma$	12	$18.835^{+7.182}_{-7.496}$	
	$q\overline{q}$	455	$456.66^{+25.386}_{-23.886}$	
	$b \to c$	67	Fixed	
	$b \to u, d, s$	9	Fixed	
$K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	$K\eta'\gamma$	3	$3.5545^{+4.507}_{-3.355}$	
	$q\overline{q}$	106	$104.59^{+12.784}_{-12.047}$	
	$b \to c$	26	Fixed	
	$b \to u, d, s$	3	Fixed	

Table 4.10: The input MC events, fit yields and fit significance for the simultaneous scaled MC fitting trial to all three charged final states.

Mode	MC Type	Input Events	Fit Yield	Significance
$K^0_S\eta'(\rho^0\gamma)\gamma$	$K\eta'\gamma$	1	$1.988^{+4.440}_{-3.320}$	1.20σ
	$q\overline{q}$	239	$234.82^{+17.903}_{-17.056}$	
	$b \to c$	59	Fixed	
	$b \to u, d, s$	4	Fixed	
$K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	$K\eta'\gamma$	1	$3.443^{+3.506}_{-2.566}$	
	$q\overline{q}$	39	$36.233^{+7.482}_{-6.839}$	
	$b \to c$	12	Fixed	
	$b \to u, d, s$	1	Fixed	
$K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	$K\eta'\gamma$	1	$0.869^{+1.625}_{-0.848}$	
	$q\overline{q}$	11	$13.963^{+4.610}_{-3.861}$	
	$b \to c$	7	Fixed	
	$b \to u, d, s$	1	Fixed	

Table 4.11: The input MC events, fit yields and fit significance for the simultaneous scaled MC fitting trial to all three neutral final states.

4.3.2 Combined 3D Fitting Method

This technique combines the distributions of the three charged modes into one 3D unbinned ML fit and the three neutral modes into another, reducing six 3D fits to two. This increases the number of events in each fit which may improve stability and accuracy. The fitting range of $M_{\eta'}$ can no longer be individually defined for each final state as is given in Table 4.1. Instead it is set to $0.93\,\mathrm{GeV}/c^2 < M_{\eta'} < 0.985\,\mathrm{GeV}/c^2$, and the $M_{\eta'}$ signal region is defined as $0.945\,\mathrm{GeV}/c^2 < M_{\eta'} < 0.97\,\mathrm{GeV}/c^2$. The fitting ranges and signal regions for M_{BC} and ΔE remain the same. As described in Section 3.5.5, the Fisher discriminant was re-trained, the $\mathcal{LR}$ re-calculated and the $\mathcal{FOM}$ re-optimised for new $M_{\eta'}$ fitting ranges, this time using sideband data rather than $q\overline{q}$ MC.

The same functions are used to describe the $K\eta'\gamma$ distributions for the combined method as for the mode-by-mode method: each final state is still fitted separately to model the shapes. The PDFs of the three modes included in each fit are then weighted according to the $\prod$ and ε of each final state and combined to form a single $K\eta'\gamma$ PDF. The $q\overline{q}$ distributions are also modelled using the same functions as the 3D mode-by-mode method, with a Breit-Wigner plus a 1^{st} order Chebyshev polynomial used for the $M_{\eta'}$ distribution for both charged and neutral final states.

The M_{BC} and ΔE distributions of the $b \to c$ component are modelled using a 2D Keys PDF to include the correlations between the two variables. As previously mentioned, a Keys PDF is a combination of Gaussians centred at each MC point with width dependant on the surrounding MC density. The 2D Keys PDF includes correlations in a similar fashion as the 2D histogram PDF used to describe the $b \to u, d, s$ component.

Keys PDFs depend heavily on the accuracy of the MC used to determine the parameters. When the number of MC events in the fits is low this can result in a fluctuating, highly MC-dependant Keys PDF. The $\mathcal{LR}$ cut is therefore relaxed for the determination of the 2D Keys PDF shape to maximise the available statistics. This relaxation could introduce bias in the fit if the shape of M_{BC} and/or ΔE is dependant on the $\mathcal{LR}$ cut. Correlation plots for $\mathcal{LR}$, M_{BC} and ΔE in $b \to c$ MC were generated and are shown in Appendix B. There are no large correlations between the variables, however the mean of M_{BC} increases slightly as the $\mathcal{LR}$ cut rises above 0.9. This is taken to be acceptable. The fitters will be tested in the following sections to investigate biases inherent in the process: any biases introduced by the M_{BC}:$\mathcal{LR}$ correlation will be visible in the tests. The $b \to c$ $M_{\eta'}$ distribution was also modelled using a Keys PDF with the $\mathcal{LR}$ cuts relaxed as there were no correlations found between $M_{\eta'}$ and $\mathcal{LR}$ (Appendix B).

All functions in the 3D combined fit are listed in Table 4.12. The results for the fits to $b \to c$ MC charged final states are shown in Figure 4.14.

The M_{BC} and ΔE distributions of $b \to u, d, s$ MC are also fit with a 2D Keys PDF: no correlations were found between $\mathcal{LR}$ and M_{BC} or ΔE for $b \to u, d, s$ MC (Appendix B), so the $\mathcal{LR}$ cut was relaxed for the training of this function as well. The $M_{\eta'}$ distributions are modelled with a Breit-Wigner plus a 1^{st}

order Chebyshev polynomial function. The distributions and fits are shown in Appendix C.

Table 4.13 shows the changes to the efficiencies and the number of expected events for each final state due to the change in $M_{\eta'}$ fitting ranges and use of sideband data instead of $q\overline{q}$ MC in the continuum suppression. The total efficiency for the combined samples is the average efficiency of the included final states weighted by each state's proportion of the total daughter branching fraction.

Scaled MC Trial for the 3D Combined Fitting Method.

Scaled MC samples were prepared using the method described above. One event of each neutral final state was included in the scaled neutral $K\eta'\gamma$ sample. The scaled MC samples were fit in extended unbinned ML fits, the results of which are shown in Figures 4.15 and 4.16. The returned yields of the floated components and the included numbers of MC in the samples are listed in Tables 4.14 and 4.15.

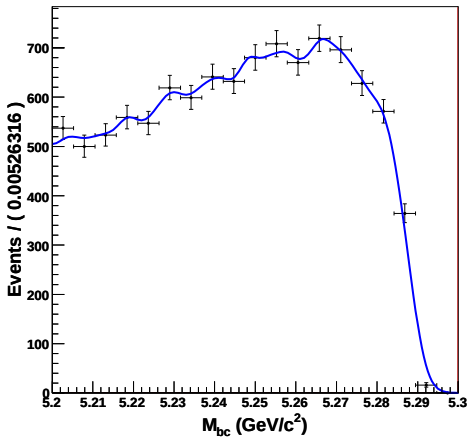
Both trial fits find stable minima in their likelihood functions and return accurate yields for both the $K\eta'\gamma$ and the $q\overline{q}$ componants of the samples. The wider $M_{\eta'}$ windows for the final states including an η pass larger levels of background MC, resulting in a lower signal significance for the fit to charged final states. The significance drops below the level required for evidence of $B \to K\eta'\gamma$ decays. The significance for the fit to neutral final states is similar to that for the 3D mode-by-mode method. Section 4.4 shows the further tests applied to the 3D combined fitting method.

MC Type	Fit Variable	Function
$K\eta'\gamma$	$M_{\rm BC}$	CBLS
	ΔE	CBLS + Gaussian
	$M_{\eta'}$	BW + 1^{st} Order Chebyshev
$q\overline{q}$	$M_{\rm BC}$	ARGUS
	ΔE	1^{st} Order Chebyshev
	$M_{\eta'}$	1^{st} Order Chebyshev + BW
$b \to c$	$M_{\rm BC}$	} 2D Keys PDF
	ΔE	
	$M_{\eta'}$	Keys PDF
$b \to u, d, s$	$M_{\rm BC}$	} 2D Keys PDF
	ΔE	
	$M_{\eta'}$	1^{st} Order Chebyshev + BW

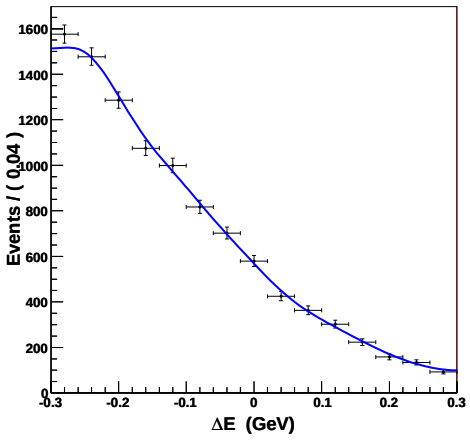
Table 4.12: The functions used to model MC samples for the combined 3D method. CBLS denotes a Crystal Ball line shape function, BW a Breit-Wigner function, and 1^{st} order Chebeyshev a 1^{st} order Chebyshev polynomial respectively. The 2D Keys PDFs run over $M_{\rm BC}$ and ΔE.

Final State	ε	Expected Events
$K^+\eta'(\rho^0\gamma)\gamma$	0.02710	10.67
$K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	0.05139	12.01
$K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	0.02974	4.01
Total Charged	0.03507	26.69
$K^0_S\eta'(\rho^0\gamma)\gamma$	0.01437	1.13
$K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	0.03335	1.55
$K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	0.01453	0.39
Total Neutral	0.02075	3.07

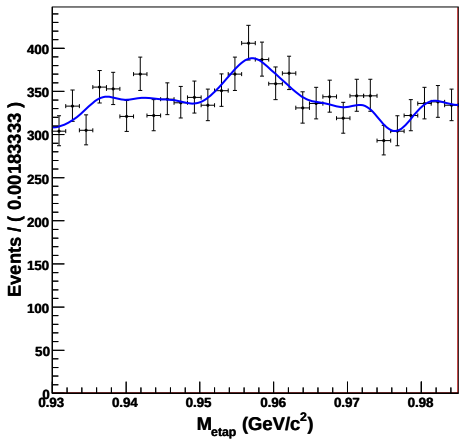
Table 4.13: The counted MC effi ciencies (ε) and expected events in 700 fb^{-1} for the combined 3D fi tting method.



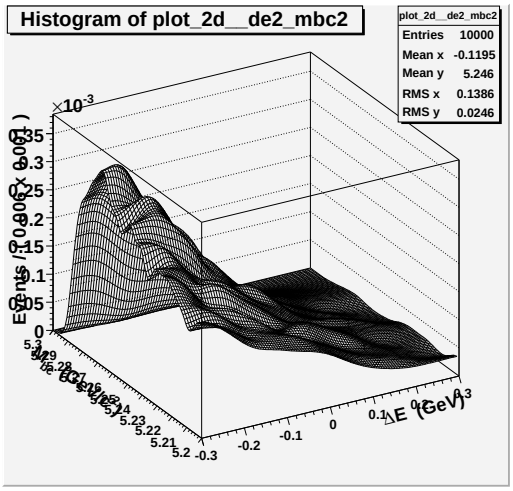
(a) Combined 3D fit to the $M_{\rm BC}$ distribution of $b \to c$ MC.



(b) Combined 3D fit to the ΔE distribution of $b \to c$ MC.



(c) Combined 3D fit to the $M_{\eta'}$ distribution of $b \to c$ MC.



(d) The 2D Keys PDF used to model the $M_{\rm BC}$ and ΔE distributions of $b \to c$ MC.

Figure 4.14: Combined 3D fits to $b \to c$ MC.

Scaled MC Fitting Trial for 3D Combined Charged Modes.

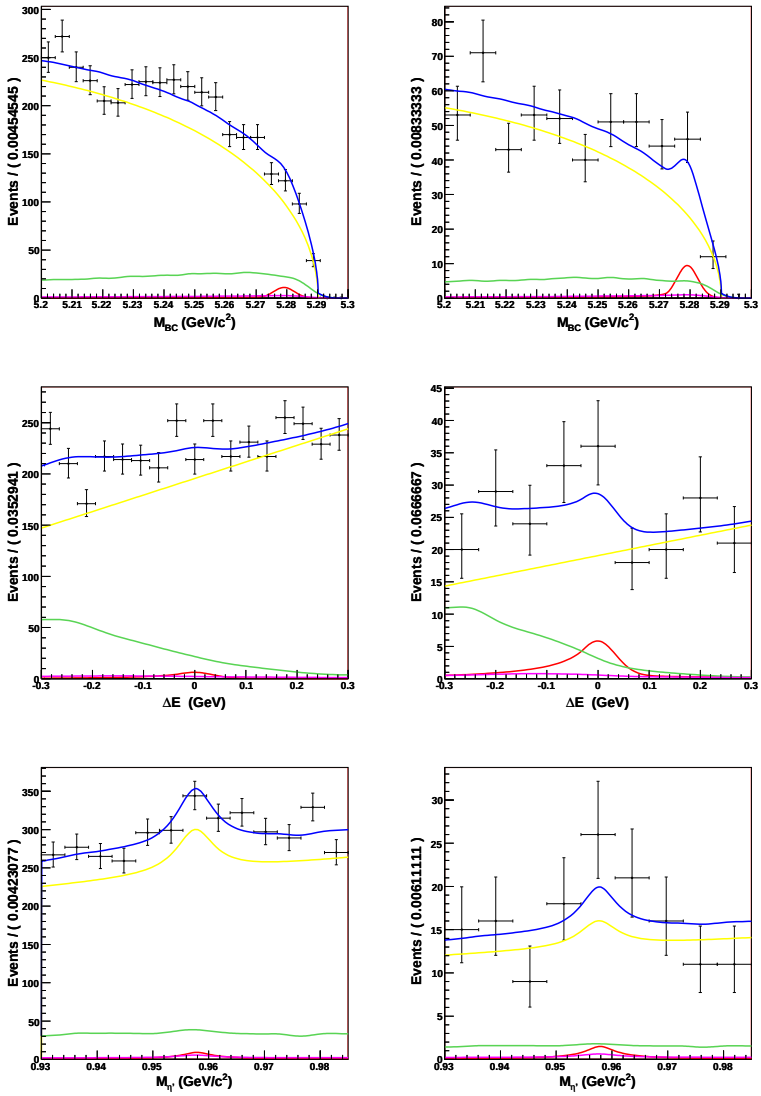


Figure 4.15: The scaled MC fitting trial for 3D combined charged modes. From top to bottom, the rows show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. The left column is the full fitting region, the right is the projections to the signal region defined in Table 4.1. The $K\eta'\gamma$ function is shown in red, $q\overline{q}$ in orange, $b \to c$ in green, $b \to u, d, s$ in magenta, and the combined function in blue.

MC Type	Input Events	Fit Yield	Significance
$K\eta'\gamma$	27	$27.574^{+15.066}_{-13.745}$	2.145σ
$q\overline{q}$	3322	$3336.2^{+62.96}_{-62.93}$	
$b \to c$	443	Fixed	
$b \to u, d, s$	36	Fixed	

Table 4.14: The input MC events, fit yields and fit significance for the 3D scaled MC fitting trial to combined charged modes.

Scaled MC Fitting Trial for 3D combined neutral modes.

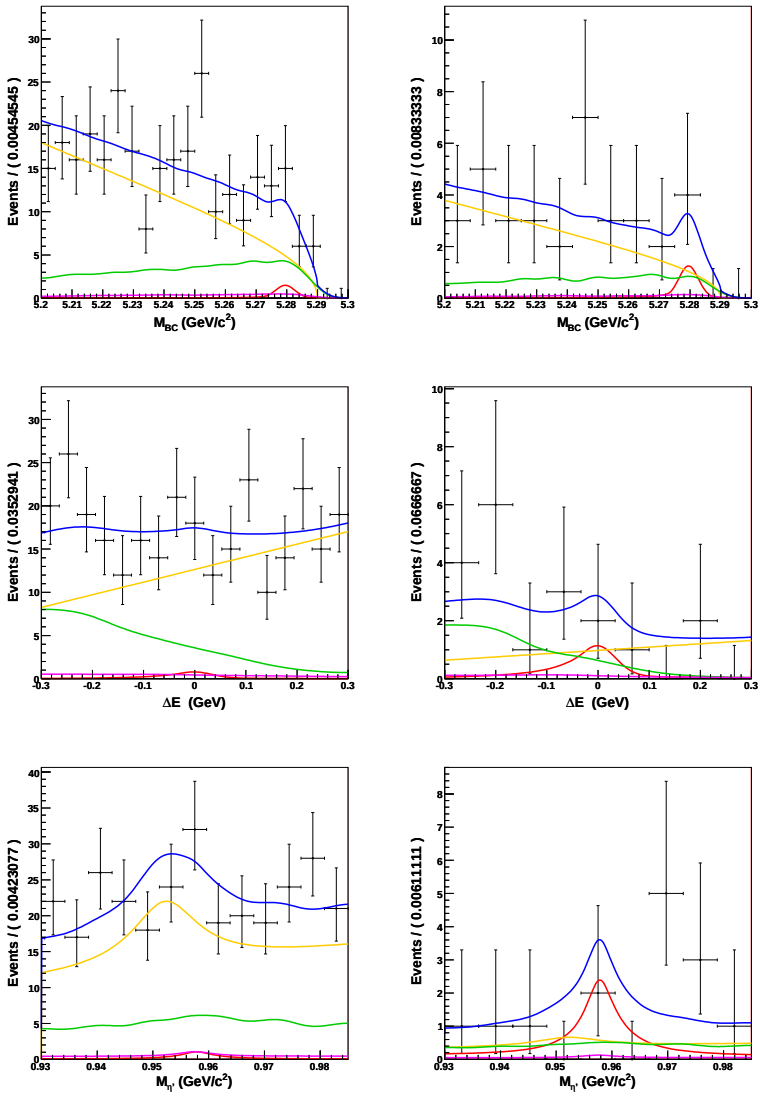


Figure 4.16: The scaled MC fitting trial for 3D combined neutral modes. From top to bottom, the rows show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. The left column is the full fitting region, the right is the projections to the signal region defined in Table 4.1. The $K\eta'\gamma$ function is shown in red, $q\overline{q}$ in orange, $b \to c$ in green, $b \to u, d, s$ in magenta, and the combined function in blue.

MC Type	Input Events	Fit Yield	Significance
$K\eta'\gamma$	3	$3.7330^{+5.0863}_{-3.4932}$	1.243σ
$q\overline{q}$	332	$326.48^{+21.110}_{-20.3521}$	
$b \to c$	87	Fixed	
$b \to u, d, s$	7	Fixed	

Table 4.15: The input MC events, fit yields and fit significance for the 3D scaled MC fitting trial to combined neutral modes.

4.3.3 2D Fitting Method

To assess whether or not the inclusion of $M_{\eta'}$ as a fit variable improves the data modelling process, the fit variables are reduced to $M_{\rm BC}$ and ΔE and the cuts on $M_{\eta'}$ changed to $945\ \mathrm{MeV}/c^2 < M_{\rho\gamma} < 970\ \mathrm{MeV}/c^2$ and $950\ \mathrm{MeV}/c^2 < M_{\eta\pi\pi} < 965\ \mathrm{MeV}/c^2$. The fitting ranges and signal regions for $M_{\rm BC}$ and ΔE remain the same. The Fisher discriminant was again re-trained, the $\mathcal{LR}$ re-calculated and the $\mathcal{FOM}$ re-optimised for these new windows.

The same functions are used for the $K\eta'\gamma$ component in the 2D fit as for the 3D fitting techniques, however, the individual final states are no longer modelled separately. The different $K\eta'\gamma$ final states are combined into one distribution and modelled with a single PDF. The results for these fits to MC are shown in Appendix E.

Instead of using a 2D Keys PDF, the correlations found between $M_{\rm BC}$ and ΔE for $b \to c$ MC are taken into account in the 2D fitting process by the introduction of a ΔE dependency in the shape parameter of the ARGUS function describing the $M_{\rm BC}$ distribution and an alteration in the normalisation condition of the PDF. The formula for the standard ARGUS function (Equation 4.9 with $x = M_{\rm BC}$) becomes

$$ARGUS(M_{\rm BC}) = NM_{\rm BC}\sqrt{1-(\frac{M_{\rm BC}}{x_0}^2)}e^{(\xi+C\Delta E)(1-(\frac{M_{\rm BC}}{x_0})^2)} \tag{4.13}$$

where C is a floated parameter. The ΔE factor allows the shape of the ARGUS to change across the range of ΔE.

The ARGUS is defined as a function of $M_{\rm BC}$ conditional on ΔE. Conditional PDFs use a different normalisation condition from standard PDFs. A standard 1D PDF is normalised as

$$\int F(x)dx \equiv 1 \tag{4.14}$$

A 1D PDF in which x is conditional on y however is normalised to 1 for all values of y, ie,

$$\int F(x|y)dx \equiv 1 \quad \forall\, y \tag{4.15}$$

A conditional PDF describes the x distribution *given* the values of y, and thus also describes any correlations between x and y. Taking the product of the conditional PDF describing the $M_{\rm BC}$ distribution and the regular PDF describing the ΔE distribution gives a PDF that is normalised over both $M_{\rm BC}$ and ΔE in identical manner as the product of two regular PDFs, ie,

$$\int F(x|y)\cdot F(y)dxdy \equiv \int F(x,y)dxdy \equiv 1 \tag{4.16}$$

No $M_{\rm BC}$ dependence or conditional probability was needed for the Chebyshev polynomial used to describe the ΔE distribution. The distributions of $M_{\rm BC}$ in three statistically equivalent bins of ΔE, and

vice versa, with the fitted function overlaid, are shown in Figure 4.17 for the charged final states. The shape of the $M_{\rm BC}$ distribution changes with the projected range of ΔE, and the conditional PDF can be seen to describe this change. The ΔE distribution is well modelled by the standard 2^{nd} order Chebyshev polynomial.

The fit does not converge to a stable minimum when using a conditional PDF to model the $b \to c$ component of the neutral final states, and a 2D Keys PDF was used instead. The $\mathcal{LR}$ cuts are relaxed for this fit and for the 2D Keys PDF modelling the $b \to u, d, s$ contribution. The functions used in the 2D fitting method are listed in Table 4.16.

Scaled MC Trial for the 2D Fitting Method.

Scaled MC samples were prepared using the same method as described above, with changes to ε due to the change in $M_{\eta'}$ selection cuts, the Fisher function training and the $\mathcal{FOM}$ optimisation as shown in Table 4.17. One event of each neutral final state was included in the scaled neutral $K\eta'\gamma$ sample. The scaled MC samples were fit in an extended unbinned ML fit, the results of which are shown in Figures 4.18 and 4.19. The returned yields of the floated components and the included numbers of MC in the samples are listed in Tables 4.18 and 4.19.

Both the 2D trial fits also find stable minima in their likelihood functions and return accurate yields for both the $K\eta'\gamma$ and the $q\overline{q}$ componants. The signal significance of the 2D trial fit to the charged final states is inbetween that of the 3D mode-by-mode method and the 3D combined method. The significance of the trial fit to neutral modes is approximately equal to the 3D combined trial fit. In the next section more stringent tests give the means of determining the optimum fitting method to use.

MC Type	Fit Variable	$K^+\eta'\gamma$ Function	$K^0_S\eta'\gamma$ Function
$K\eta'\gamma$	$M_{\rm BC}$	CBLS	CBLS
	ΔE	CBLS + Gaussian	CBLS + Gaussian
$q\overline{q}$	$M_{\rm BC}$	ARGUS	ARGUS
	ΔE	1^{st} Order Chebyshev	1^{st} Order Chebyshev
$b \to c$	$M_{\rm BC}$	ARGUS(Conditional)	} 2D Keys PDF
	ΔE	2^{nd} Order Chebyshev	
$b \to u, d, s$	$M_{\rm BC}$	} 2D Keys PDF	} 2D Keys PDF
	ΔE		

Table 4.16: The functions used to model MC samples for the 2D method. CBLS denotes a Crystal Ball line shape function, BW a Breit-Wigner function, and 1^{st} or 2^{nd} order Chebeyshev a 1^{st} or 2^{nd} order Chebyshev polynomial respectively. The 2D Histogram PDF runs over $M_{\rm BC}$ and ΔE.

Final State	ε	Expected Events
$K^+\eta'(\rho^0\gamma)\gamma$	0.01444	5.78
$K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	0.03544	8.41
$K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	0.02122	2.91
Total Charged	0.02210	17.10
$K^0_S\eta'(\rho^0\gamma)\gamma$	0.01380	1.09
$K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	0.02393	1.11
$K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	0.00805	0.22
Total Neutral	0.01596	2.42

Table 4.17: The counted MC effi ciencies (ε) and the expected number of events in 700 fb^{-1} for the 2D fi tting method.

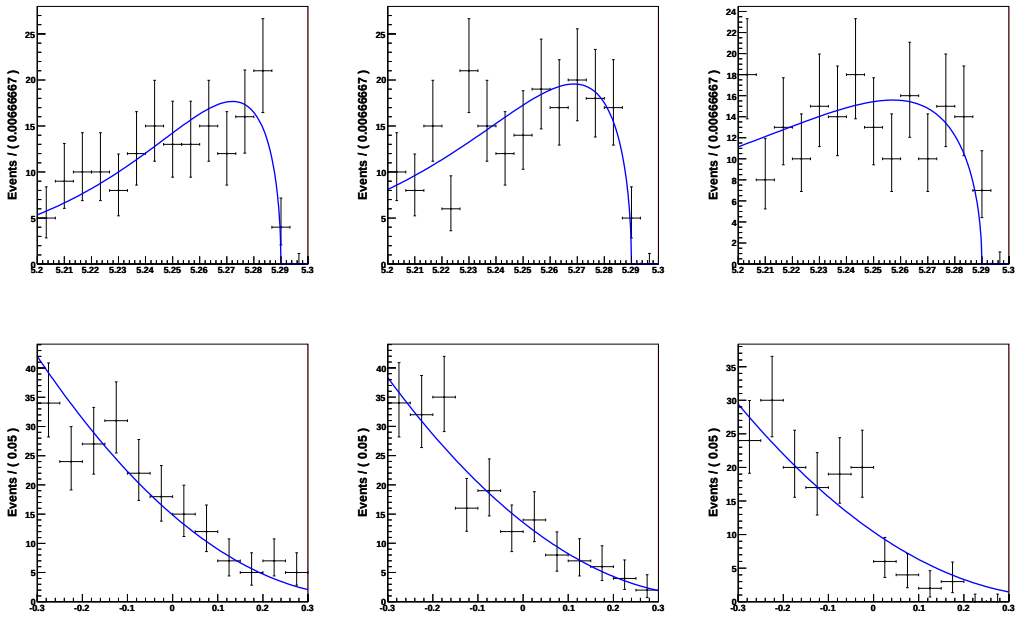


Figure 4.17: Projections of one fi tting variable in statistically equivalent bins of the other for the charged modes of the 2D fi tting method over $b \to c$ MC. The top row are the M_{BC} distributions in bins of, from left to right, $-0.3 < \Delta E/\,\mathrm{GeV} < -0.21$, $-0.21 < \Delta E/\,\mathrm{GeV} < -0.07$, and $-0.07 < \Delta E/\,\mathrm{GeV} < 0.3$. The bottom ΔE distributions in bins of $5.20 < M_{\mathrm{BC}}/\,\mathrm{GeV}/c^2 < 5.24$, $5.24 < M_{\mathrm{BC}}/\,\mathrm{GeV}/c^2 < 5.27$, and $5.27 < M_{\mathrm{BC}}/\,\mathrm{GeV}/c^2 < 5.295$.

Scaled MC Fitting Trial for the 2D Fit to Charged Modes.

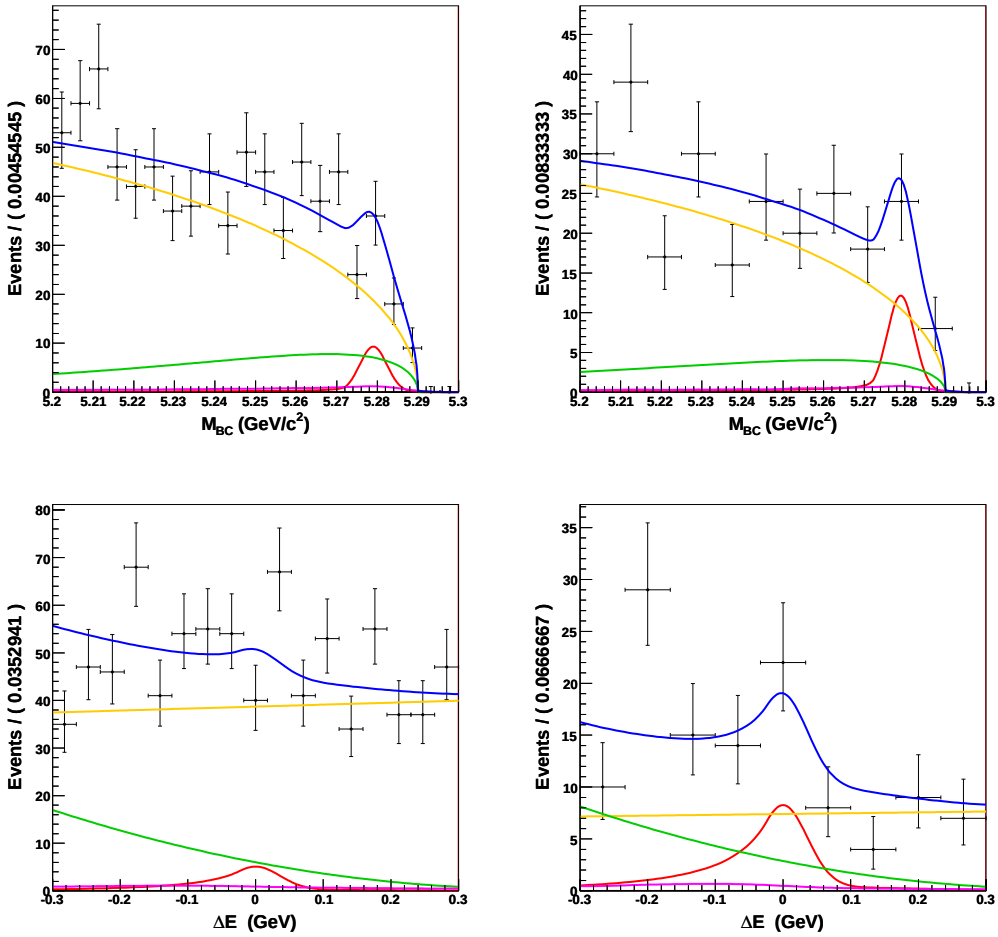


Figure 4.18: The scaled MC trial for the 2D fit to charged modes. The top plots show $M_{\rm BC}$, the left show ΔE. The left column is the full fitting region, the right the projections to the signal region. The $K\eta'\gamma$ function is shown in red, $q\overline{q}$ in orange, $b \to c$ in green, $b \to u, d, s$ in magenta, and the combined function in blue.

MC Type	Input Events	Fit Yield	Significance
$K\eta'\gamma$	17	$21.898^{+9.6401}_{-8.6368}$	2.824σ
$q\overline{q}$	651	$704.50^{+29.495}_{-28.675}$	
$b \to c$	127	Fixed	
$b \to u, d, s$	15	Fixed	

Table 4.18: The input MC events, fit yields and fit significance for the 2D scaled MC fitting trial to combined charged modes.

Scaled MC Fitting Trial for the 2D Fit to Neutral Modes.

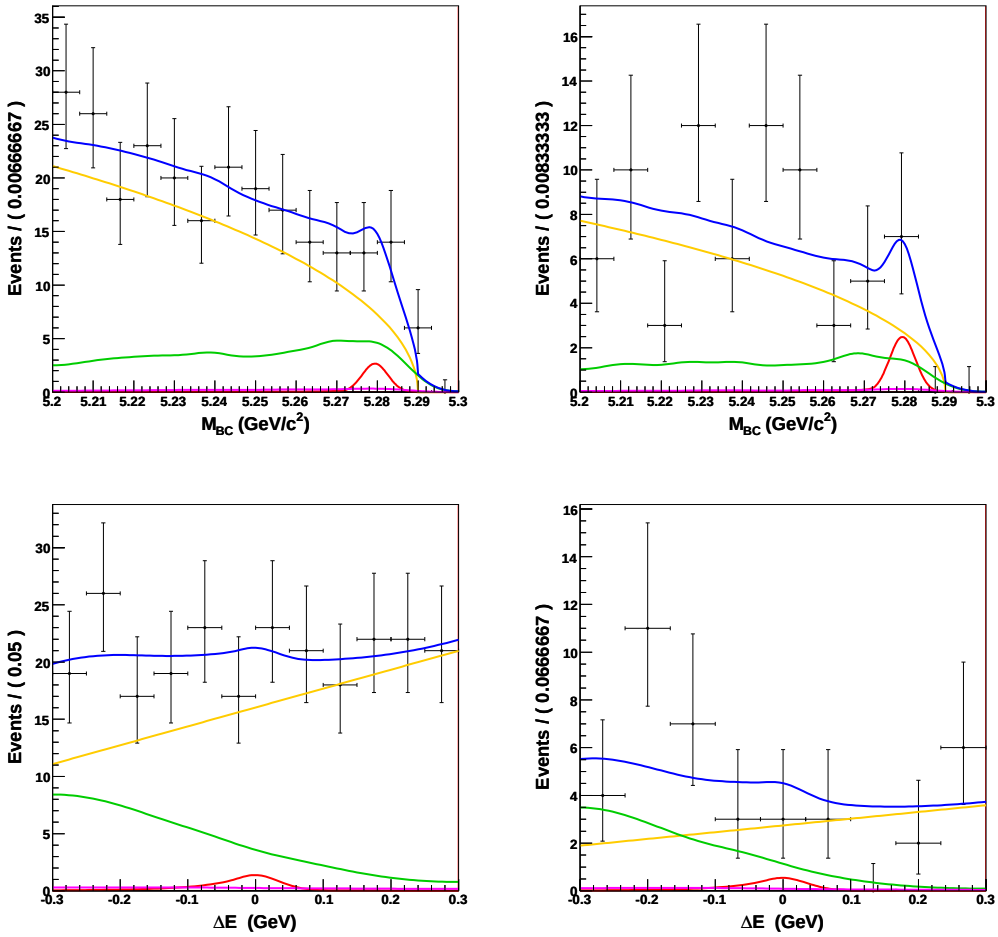


Figure 4.19: The scaled MC trial for the 2D fit to neutral modes. The top plots show $M_{\rm BC}$, the left show ΔE. The left column is the full fitting region, the right the projections to the signal region. The left plots show $M_{\rm BC}$, the right show ΔE. The $K\eta'\gamma$ function is shown in red, $q\overline{q}$ in orange, $b \to c$ in green, $b \to u, d, s$ in magenta, and the combined function in blue.

MC Type	Input Events	Fit Yield	Significance
$K\eta'\gamma$	3	$3.5252^{+4.2094}_{-3.2908}$	1.086σ
$q\overline{q}$	193	$192.02^{+15.893}_{-15.379}$	
$b \to c$	49	Fixed	
$b \to u, d, s$	3	Fixed	

Table 4.19: The input MC events, fit yields and fit significance for the 2D scaled MC fitting trial to combined neutral modes.

4.4 Fitting Method Tests

In order to determine the stability and accuracy of the three different fitting methods, many samples of MC were generated and fitted using each technique. Two MC generating methods were utilised; the "toy Monte Carlo" method, and full simulation using EvtGen and GSIM.

Toy MC uses the initialised shape of the fitting function to accept or reject randomly generated points within the parameter space of the function, creating a randomised MC sample that is heavily based on the input function. Toy MC has the advantage of the speed and ease with which very large samples can be generated. The use of toy MC in the statistical examination of fit accuracy is a good test of biases introduced by the choice of function parameterisation and parameter boundaries, and those due to the fitting of low statistical samples. The last factor in particlular is very relevant to rare decay mode analyses like this one.

The use of toy MC does not test the assumptions that are included in the creation of the initial fitting function. An example of such assumptions is the neglect of any correlations between fitting variables below a certain level. Small correlations can introduce biases that will not be tested by toy MC, as points randomly created within the multi-dimensional parameter space will not include such correlations. In order to test these assumptions, fully simulated MC is required. To acheive this the full samples of signal and background MC that were used in creating the fitters were randomly divided into as many fully independent samples as possible, using the technique described in Section 4.3. These are then used in a statistical study of the fitter attributes.

Tests based on both of these MC generation techniques are described in detail below.

4.4.1 Toy Monte Carlo test

2500 samples of toy MC were generated for each of the three fitting techniques described in the Section 4.3. The relative numbers of $q\overline{q}$, $b \to c$, $b \to u, d, s$ and $K\eta'\gamma$ events included in each sample was scaled to $604.55\,\mathrm{fb}^{-1}$ of integrated luminosity.

Each of the samples was fit with the associated fitter and the distributions of the $K\eta'\gamma$ and $q\overline{q}$ yields, the floated function parameters and the errors associated with each were recorded. The accuracy of each fitter was determined by calculating the "pull" of the yields, as defined by

$$\mathrm{Pull}(N_i) = \frac{N_i^{fit} - N_i^{true}}{\sigma_i^{fit}} \tag{4.17}$$

where N_i^{fit} and N_i^{true} are the value returned by the fit and true input value of the yield i, respectively, and σ_i is the absolute value of the returned statistical error on N_i^{fit}. The asymmetric MINOS errors are used, as described in Section 4.1. σ_i is the negative MINOS error when the numerator is positive, and the positive MINOS error when the numerator is negative.

The pull distribution for a statistically significant sample of fits will be normally distributed around zero

for an unbiased fitter. If the mean of the distribution (μ_{pull}) is within statistical error of 0 and the width (σ_{pull}) is within statistical error of 1, the fitter is shown to be unbiased for the factors tested by toy MC. A μ_{pull} significantly away from 0 indicates a bias in the fitted yield, while a σ_{pull} significantly away from 1 indicates a problem with the calculated errors on the yield. The pull distributions are fitted with a Gaussian function to determine μ_{pull} and σ_{pull}.

Figure 4.20 shows the distribution of the pull, the returned yield and the signal significance from the 2500 fits for the 2D fitter; Figure 4.21 shows the same distributions for the 3D combined fitter. Similar plots for the individual and simultaneous fits of the 3D mode-by-mode fitter are shown in Appendix F. Table 4.20 gives the results of the tests for all three fitters.

The simultaneous fits of the 3D mode-by-mode method perform poorly in the test; the fits to both the charged and neutral final states significantly underestimate the number of signal events in the samples and only 60.8% of the fits to neutral final states converge to a stable minimum. The instability of this method renders it unusable for the fit to data.

The 3D combined and the 2D fitters perform well in stability and accuracy. They both show fractional negative biases in the returned signal yield at approximately two standard deviations. This is an acceptable level of inaccuracy which can be adjusted for in the fits to data. Most importantly, both fitters display the stability required for reliable use.

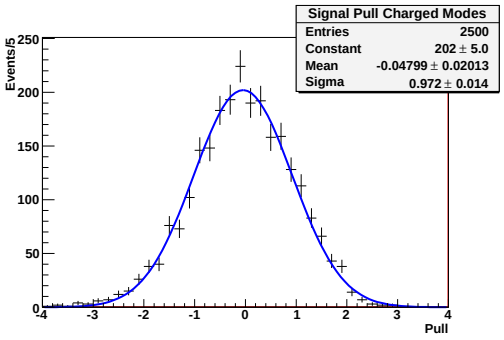
As above, the significance is defined as $\sqrt{-2ln(L_0/L_{max})}$, where L_{max} and L_0 are the minimised values of the likelihood functions when the signal yield is floated or fixed to zero, respectively. The significance is assigned the sign of the returned signal yield: a negative significance indicates a negative returned signal yield. This tagging of the un-physical results provides a continuous significance distribution across $\sigma = 0$, allowing a Gaussian fit estimation of the mean expected significance.

The 2D method returns a $\approx 10\%$ higher signal significance than the 3D combined method for the fits to charged final states, while for the neutral final states the 3D combined method returns a $\approx 20\%$ higher significance. This is important as a result of above 3σ is required to be considered evidence of $B \to K\eta'\gamma$ decay. The charged channel has the highest chance of a positive outcome as the reconstruction efficiency

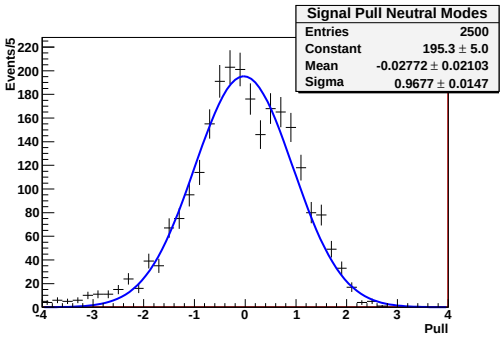
Fitter	Final States	Fits (%)	Input	Yield	μ_{pull}	σ_{pull}
3D	Charged	98.4	24	10.12 ± 0.24	-1.447 ± 0.023	1.056 ± 0.017
Mode-by-Mode	Neutral	60.8	3	-2.313 ± 0.202	-1.298 ± 0.052	1.364 ± 0.038
3D	Charged	99.8	23	22.65 ± 0.30	-0.046 ± 0.020	0.957 ± 0.014
Combined	Neutral	100.0	3	2.824 ± 0.096	-0.046 ± 0.022	0.952 ± 0.015
2D	Charged	100.0	15	14.76 ± 0.19	-0.048 ± 0.020	0.972 ± 0.014
	Neutral	100.0	3	3.034 ± 0.082	-0.028 ± 0.021	0.968 ± 0.015

Table 4.20: The results of the toy MC test for the three fitters. The 'Fits' column gives the percentage of the 2500 fits that converged successfully. The 'Input' column gives the number of $K\eta'\gamma$ events in each fitted sample.

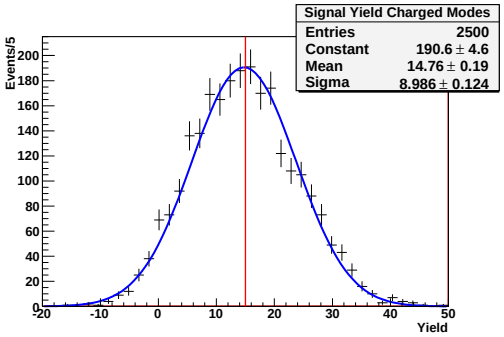
is much greater than that of the neutral modes; optimising for discovery will favour the 2D method on the basis of these results.



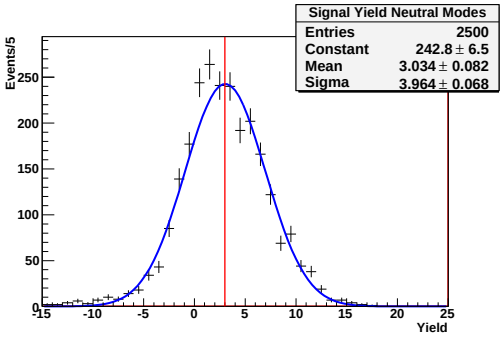
(a) Pull distrubution for the charged modes.



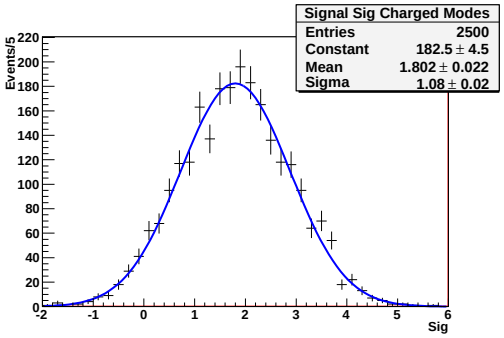
(b) Pull distrubution for the neutral modes.



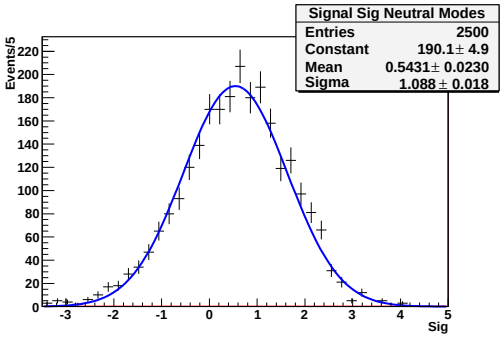
(c) Yield distrubution for the charged modes. The red line is the true signal value for the fit.



(d) Yield distrubution for the neutral modes. The red line is the true signal value for the fit.

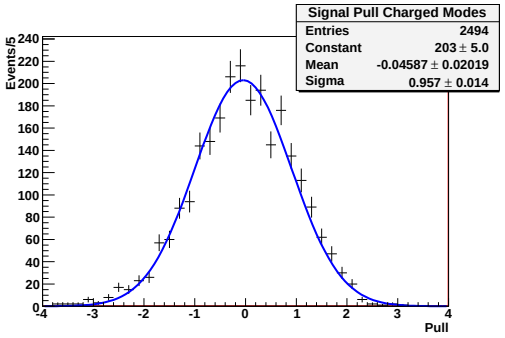


(e) Significance distrubution for the charged modes.

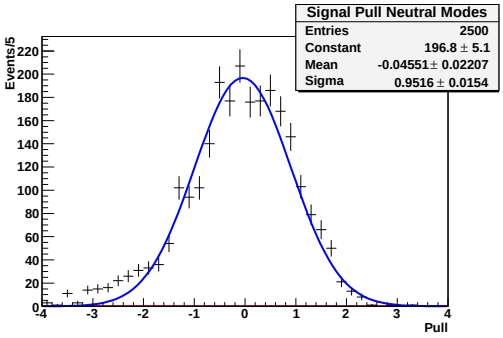


(f) Significance distrubution for the charged modes.

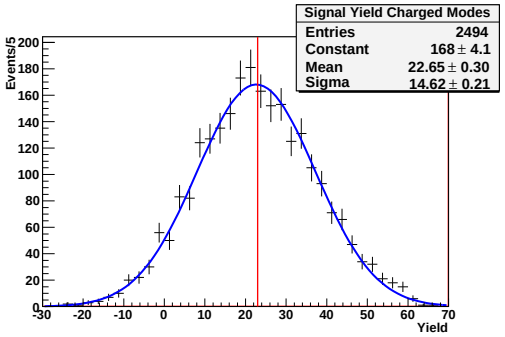
Figure 4.20: The results of the toy MC test for the 2D fitter.



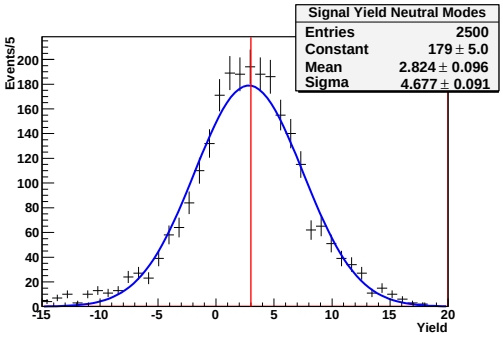
(a) Pull distrubution for the charged modes.



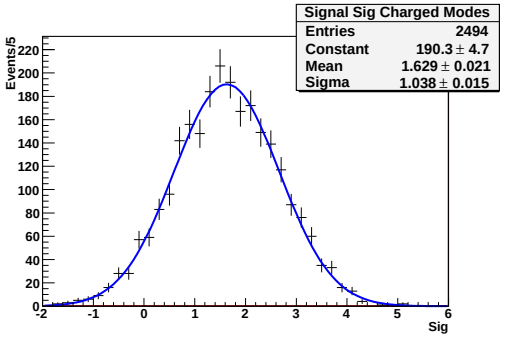
(b) Pull distrubution for the neutral modes.



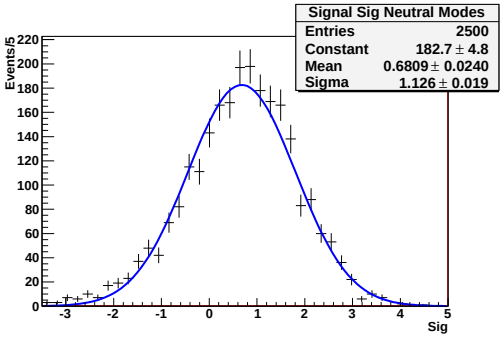
(c) Yield distrubution for the charged modes. The red line is the true signal value for the fit.



(d) Yield distrubution for the neutral modes. The red line is the true signal value for the fit.



(e) Significance distrubution for the charged modes.



(f) Significance distrubution for the neutral modes.

Figure 4.21: The results of the toy MC test for the 3D combined method fitter.

4.4.2 GSIM ensemble test

The GSIM ensemble test used MC generated using the full simulation method of EvtGen and GSIM rather than the toy MC process. The number of MC events passing all selection criteria is divided by the number of events expected in $700\,\mathrm{fb}^{-1}$ to find the total number of independent pseudo-experiments possible. As very few $B \to K\eta'\gamma$ events are expected in $700\,\mathrm{fb}^{-1}$ of data, the signal component could be generated quickly for many pseudo-experiments. The expected number of $b \to u, d, s$ events was also small and the available $b \to u, d, s$ MC was enough for many pseudo-experiments. The production of the $q\overline{q}$ and $b \to c$ MC samples has been an on-going project of the entire collaboration since its inception and any significant increase in their size was not feasible within the scope of this analysis. The samples referred to in Section 3.1.2 are all that are available and this will place constraints on the number of pseudo-experiments possible.

The limiting factor was found to be the number of $b \to c$ MC events. As the $\mathcal{LR}$ cut was found to be uncorrelated with M_{BC} and ΔE for $b \to c$ events it was completely relaxed to maximise statistics. Twenty fully independent pseudo-experiments of $b \to c$ MC can be formed for each of the three fitting methods.

Although the equivalent integrated luminosity of the total $q\overline{q}$ MC sample was less than that of the $b \to c$ MC sample, the continuum suppression was far more effective at removing $q\overline{q}$ events than $b \to c$ events. This means that the $\mathcal{LR}$ cut for $q\overline{q}$ MC does not need to be relaxed as far to create the same number of pseudo-experiments. The correlation plots for M_{BC}, ΔE and $M_{\eta'}$ vs. $\mathcal{LR}$ for $q\overline{q}$ MC are shown in Appendix B. Correlations can be seen between $\mathcal{LR}$ and both M_{BC} and ΔE. Changing the $\mathcal{LR}$ cuts will change the shape of the M_{BC} and ΔE spectra so it was desirable to leave the $\mathcal{LR}$ cuts as close to the optimum as possible. The $\mathcal{LR}$ cut was relaxed only as much as was required to create twenty samples.

The pseudo-experiments were constructed in the same manner as the trial MC fits shown in Section 4.3. For each final state the number of events expected of each type of background were selected randomly from the MC that pass selection and combined into a pseudo-experiment. Each is equivalent to the background component of the previous trial fits shown in Section 4.3. Twenty pseudo-experiments were created for each of the three different fitting methods.

$K\eta'\gamma$ MC was then embedded in these background MC pseudo-experiments. Twenty $K\eta'\gamma$ MC samples for each final state were randomly selected from the full $K\eta'\gamma$ MC sample and combined with the corresponding background pseudo-experiments. The expected sizes of the background components could be estimated with reasonable accuracy as they have been the focus of prolonged study, but the number of $K\eta'\gamma$ events in the data was a complete unknown. To investigate the behaviour of each fitting method when different amounts of $K\eta'\gamma$ MC was embedded in the background, the ramdom selection of $K\eta'\gamma$ events was repeated ten times for each of the twenty background pseudo-experiments based on each of the final states. The size of the samples ranged from four events below the expected number of $K\eta'\gamma$ events to five events above. For those final states with less than five expected events, samples of sizes from 0 to 10 were created. Each of the groups of $K\eta'\gamma$ samples were then combined with the background pseudo-

experiments, creating two hundred full pseudo-experiments for each final state: 10 groups of twenty, with each group differing in the number of of $K\eta'\gamma$ MC samples embedded. For the 3D combined and the 2D methods, the charged final state pseudo-experiments and neutral final state pseudo-experiments were then merged.

The pseudo-experiments were fitted using the relevant fitting method. For each of the ten groups of twenty fits the mean of the returned signal yields was calculated and plotted. The results are shown in Figures 4.22 to 4.24. The black points are the means of the fits to the twenty pseudo-experiments and the blue points are the true numbers of embedded $K\eta'\gamma$ MC. For an unbiased fitter the black points will be within statistical error of the blue points. The means are fitted with a 1^{st} order Chebyshev polynomial to estimate the linearity of the fitter response across the different amounts of embedded $K\eta'\gamma$ MC. The results of the linear fits are shown on the plots. Deviation from $p_0 = 0.0$ indicates a constant bias and deviation from $p_1 = 1.0$ indicates a bias that is a function of the number of true $K\eta'\gamma$ events.

All of the GSIM test fits for the 3D mode-by-mode method simultaneous fits to charged final states succeeded in finding a stable minimum. The means of the signal yields (Figure 4.22) had a large increasing positive bias, equivalent to approximately 10 events. Separation of the final states showed the greatest over-estimation to be in the $K^+\eta'(\rho^0\gamma)\gamma$ mode. An investigation of a similar bias in the 3D combined method (detailed below) found it to originate in $q\overline{q}$ component the $K^+\eta'(\rho^0\gamma)\gamma$ final state. As the selection criteria for $K^+\eta'(\rho^0\gamma)\gamma$ are identical for the 3D mode-by-mode method and the 3D combined method, the bias here most likely stems from the same source. The conflicting nature of the biases in the toy MC and GSIM tests are a further indication of the unsuitability of this fitting technique for the low statistics involved in this analysis.

The simultaneous fits to neutral final states fail to converge for 40% of the pseudo-experiments, confirming the instability shown in the toy MC test. Most of the means are within error of the true values, though the linear fit is skewed to the negative by a $\approx$ 5 event yield under-estimation when there are no $K\eta'\gamma$ MC events in the sample. The fitter here shows greater accuracy than in the toy MC test, however the fail rate is again too high for the fitter to be considered stable enough for use.

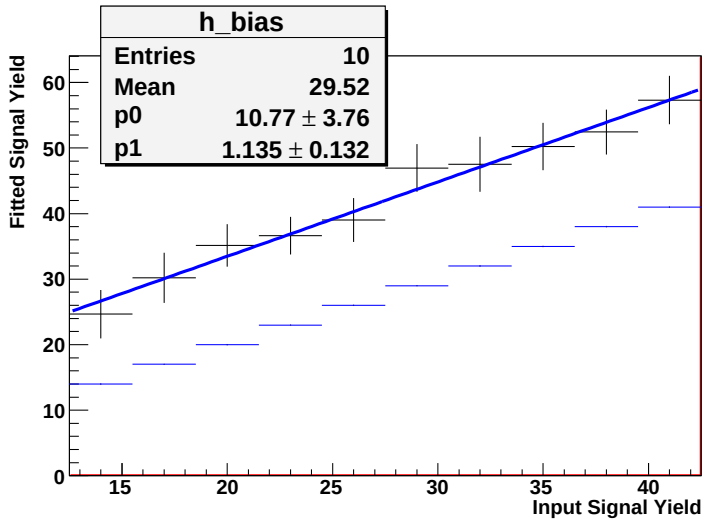
The results for the 3D combined fits to charged final states (Figure 4.23) also show a significant increasing positive bias of approximately 5 to 7 events. In an effort to identify the source of this bias, GSIM tests were run with and without the relaxation of the $b \to c$ $\mathcal{LR}$ cuts when creating the 2D Keys PDF; similar results were obtained. Further tests were run over samples that include only a single type of background combined with the signal, samples that include only two of the three reconstructed final states, and samples in which $M_{\eta'}$ was excluded from the fit. The investigation indicates the bias as originating in the $q\overline{q}$ MC component tagged as $K^+\eta'(\rho^0\gamma)\gamma$ events. When this final state is not included in the pseudo-experiments all the signal yield means are within error of the true values and the linear fit describes a constant negative bias of $\approx$1.5 events. When the $q\overline{q}$ component is not included in the pseudo-experiments the bias is negligible at low numbers of true $K\eta'\gamma$ events, rising to $\approx$2 events when the number of true $K\eta'\gamma$ events is

at maximum. The constructed PDF is unable to model this component of the GSIM MC accurately. The $K^+\eta'(\rho^0\gamma)\gamma$ final state has by far the highest $q\overline{q}$ background of any of those reconstructed: the signal to noise ratio for this mode is low. While no correlations were found between the fitting variables for $q\overline{q}$ MC, the fact that this bias was not evident in the toy MC test suggests that this is the root cause. As the selection criteria for the 3D mode-by-mode fitter are almost identical for $K^+\eta'(\rho^0\gamma)\gamma$ this source will also be part of that bias.

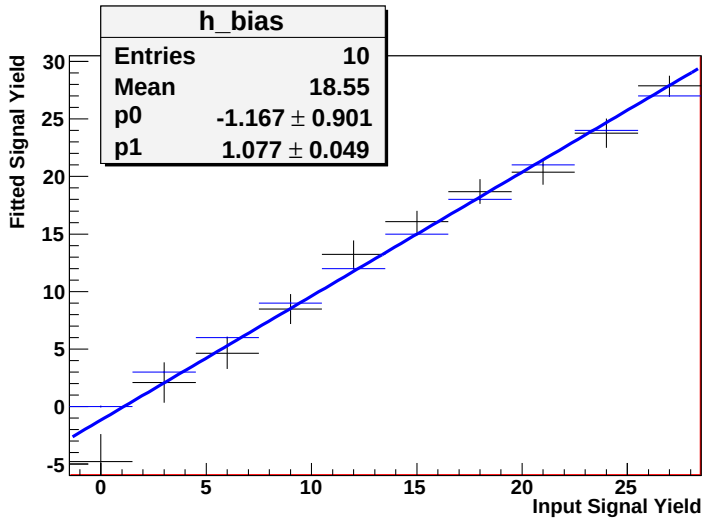
The results for the 3D combined fits to neutral final states are more accurate than the charged final states, with no evidence of bias seen within statistical error. The 3D combined fitter is more stable than the 3D mode-by-mode method: all of the fits to the pseudo-experiments converge to a stable minimum for both the charged and neutral final states.

The results for the 2D fits to the pseudo-experiments are shown in Figure 4.24. None of the 2D fits fail to converge and no large bias is seen in either result. All of the means for the charged final states are within error of the true values, though the linear fit gives a negative bias that is shifting towards the positive. This bias is zero at 17.7 events, close to the expected number of 17 events. The 2D fit, with tighter $M_{\eta'}$ and $\mathcal{LR}$ cuts, has less than 20% of the $q\overline{q}$ background of the 3D combined fit. The affect of any small correlations in the $q\overline{q}$ distribution will be reduced accordingly.

The means of the returned yields of the 2D fits to neutral final states are within error at low $K\eta'\gamma$ event numbers, but show an increasing positive bias as the number of events gets higher. This dependency would require a correction to the neutral signal yield if a high yield is measured; the results of the previous $K\eta'\gamma$ analysis (Equations 4.10) predict it to be low. Tests based on the returned yields of the fit to data will be give a more accurate bias assessment.

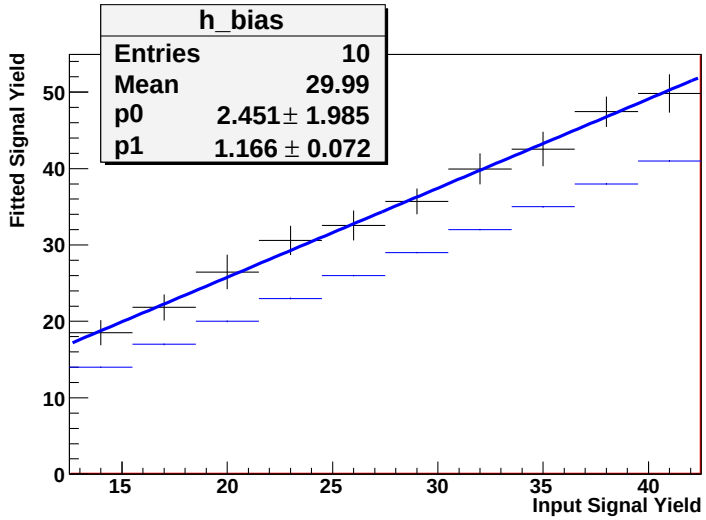


(a) GSIM test results for the 3D mode-by-mode method fit to charged modes.

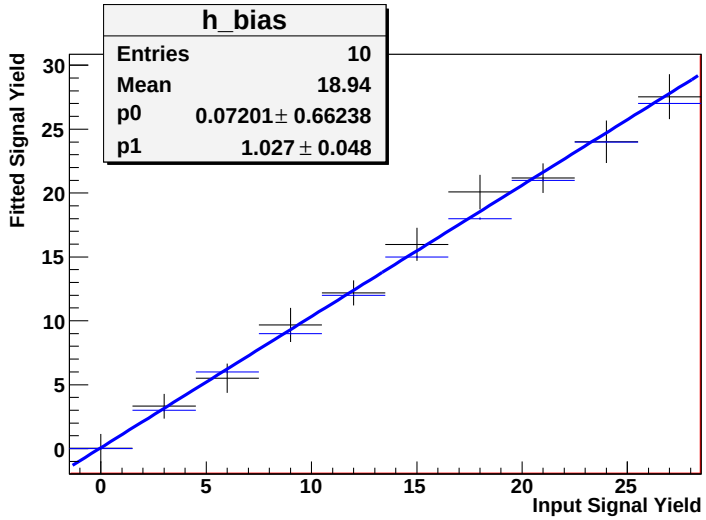


(b) GSIM test results for the 3D mode-by-mode method fit to neutral modes.

Figure 4.22: The results of the GSIM test for the simultaneous fits of the 3D mode-by-mode method. The y-axis is the returned signal yield, the x-axis the true number of $K\eta'\gamma$ events. The black points with vertical error bars are the means of the 20 fits to GSIM MC for each number of $K\eta'\gamma$ events. The blue points are the true number of $K\eta'\gamma$ events. The blue line is the fit to the returned yield means.

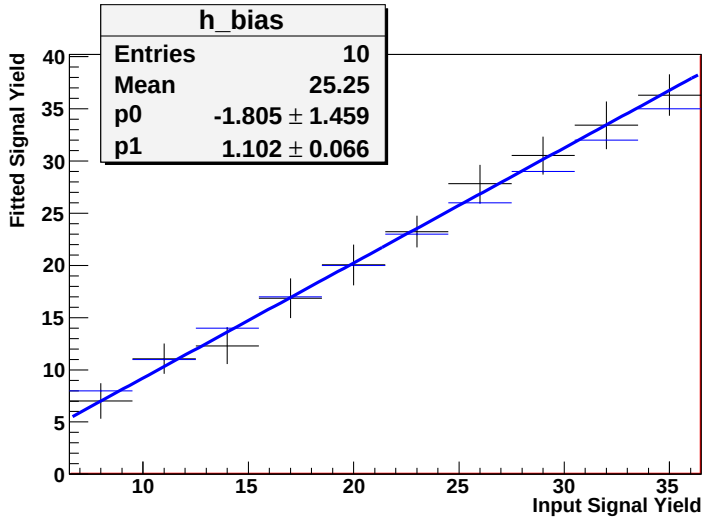


(a) GSIM test results for the 3D combined method fit to charged modes.

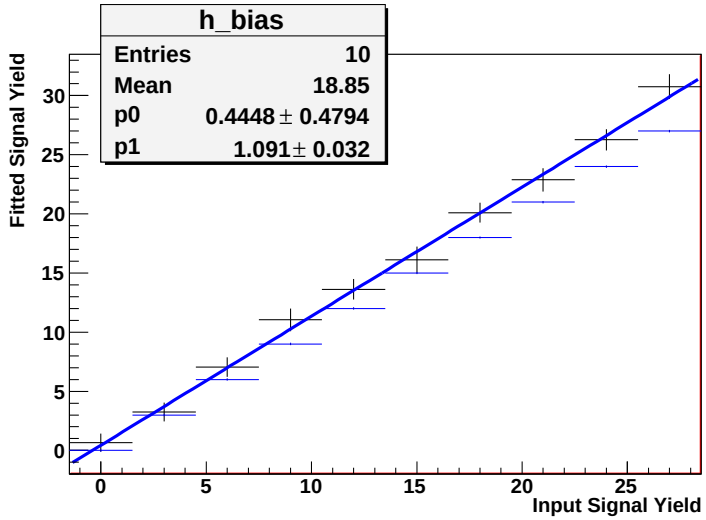


(b) GSIM test results for the 3D combined method fit to neutral modes.

Figure 4.23: The results of the GSIM test for the 3D combined method. The y-axis is the returned signal yield, the x-axis the true number of $K\eta'\gamma$ events. The black points with vertical error bars are the means of the 20 fits to GSIM MC for each number of $K\eta'\gamma$ events. The blue points are the true number of $K\eta'\gamma$ events. The blue line is the fit to the returned yield means.



(a) GSIM test results for the 2D method fit to charged modes.



(b) GSIM test results for the 2D method fit to neutral modes.

Figure 4.24: The results of the GSIM test for the 2D method. The y-axis is the returned signal yield, the x-axis the true number of $K\eta'\gamma$ events. The black points with vertical error bars are the means of the fits to GSIM MC for each number of $K\eta'\gamma$ events. The blue points are the true number of $K\eta'\gamma$ events. The blue line is the fit to the returned yield means.

4.4.3 Fitter Comparison

The results the tests of the 3D mode-by-mode method showed it to be inaccurate and unstable. The toy MC fits did not have any large biases when the final states are fitted individually, but when the simultaneous fits were performed there were strong negative biases for both charged and neutral final states. Minuit was unable to find a single stable minimum in the combined likelihood function of the PDFs for the three final states. The GSIM test confirmed this instability, showing a large positive bias for the charged final states. The simultaneous fitter was finding different minima for each test, both of them far away from the desired result. The stability of the fits to the neutral final states was also very poor in both tests. While the GSIM test showed greater accuracy than the toy MC test, over 40% of the fits in both tests either failed to find a stable minimum and/or had likelihood functions for which MINOS could not calculate asymmetric errors.

The toy MC results for the 3D combined method had the least bias of the different methods of fitting the charged final states, however the GSIM test showed a ≈ 5 event positive bias. This bias stemmed from $q\overline{q}$ MC reconstructed as $K^+\eta'(\rho^0\gamma)\gamma$, most likely due to small correlations between the fitting variables. The neutral final states GSIM test showed no bias within statistical error, while the toy MC test had a negative bias of ≈ 0.22 events.

The 2D fitting method performed in a similar fashion as the 3D combined method in the toy MC test of fits to the charged final states, and slightly better for the neutral final states. This method was by far the most accurate in the GSIM test results for charged final states and had minimal bias in the neutral final states.

While the technique of fitting each reconstructed final state separately in the 3D mode-by-mode method returned the highest expected signal significance, it was too unstable to be trusted. Of the two stable fitting methods the 2D fitter was the most accurate in the GSIM test and had the highest expected signal significance for the charged final states.

The 2D fitter overall was found to have the least bias and greatest stability of the different methods. The introduction of $M_{\eta'}$ as a third fitting dimension increases the amount of information available for the fit, however this did not correspond with greater accuracy or significance. The variable $M_{\eta'}$ was utilised more effectively as a criteria for selection rather than a fitting variable. The 2D fitter is used throughout the rest of this analysis. Further statistical tests of fitter accuracy were performed based on the results of the fit to data and are presented in Section 4.7.

4.5 Signal PDF calibration

The Monte Carlo generation technique is known to be imperfect and differences in the shapes of the kinematic variables $M_{\rm BC}$ and ΔE between MC and data are expected. Control samples are utilised to calibrate MC without compromising the objectivity of the analysis. They are required to be as kinematically similar to the studied decay as possible and to have a large branching fraction to minimise statistical error in the calibration. The decay $B \to K^*(892)\gamma \to K\pi\gamma$ was identified as having similar kinematic properties to $B \to K\eta'\gamma$. The differences in final state particle multiplicity is off-set by $B \to K^*(892)\gamma$ also being a radiative decay and thus including a similar low energy tail in the ΔE distribution. $B \to K^*(892)\gamma$ has a $\mathcal{BF}$ approximately 300 times higher than $B \to K\eta'\gamma$, meaning a significantly lower statistical error on the fit.

The full data set of $604.55\,\mathrm{fb}^{-1}$ along with 1.481 million $K^*(892)\gamma$ MC events and $149.4\,\mathrm{fb}^{-1}$ of $q\overline{q}$ MC were analysed in as similar a fashion to the $B \to K\eta'\gamma$ analysis as possible. Identical selection criteria were used for all particles the two analyses have in common. $K^*(892)$ candidates of invariant mass within $\pm 50\,\mathrm{MeV}$ of the nominal $K^*(892)$ mass [19] were accepted. The Fisher discriminant was re-trained and the $\mathcal{FOM}$ re-optimised on the $K^*(892)\gamma$ and $q\overline{q}$ MC. A best candidate selection based on minimum B vertex χ^2 and the closest $K^*(892)$ mass to the nominal value was used when multiple candidates were reconstructed from the same event.

The ΔE distribution of the $B \to K^*(892)\gamma$ MC in the region of $M_{\rm BC} > 5.27\,\mathrm{GeV}/c^2$ was fit in an unbinned ML fit using the same CBLS plus Gaussian PDF as was used for $B \to K\eta'\gamma$. The data was fit in the same region with the same function but with α_{CB}, N_{CB}, the relative fractions of the CBLS and Gaussian and the function width scaling factor fixed to the results of the MC fit. A 1^{st} order Chebyshev polynomial was added to model the background, fixed to the shape obtained from a fit to the $q\overline{q}$ MC. The relative fractions of the signal and background functions, and the σ and the μ of the CBLS PDF were floated. The fit results are shown in Figure 4.25.

The $M_{\rm BC}$ distribution of $B \to K^*(892)\gamma$ MC in the region $-0.1\,\mathrm{GeV} < \Delta E < 0.07\,\mathrm{GeV}$ was fit using a CBLS PDF. The data was fit with the same function with α_{CB} and N_{CB} fixed to the results of the MC fit. An ARGUS function with end-point fixed to $5.29\,\mathrm{GeV}/c^2$ and floating shape parameter was added to model the background. The fit results are shown in Figure 4.26.

The optimised parameters of the ΔE and $M_{\rm BC}$ CBLS PDFs from the data and MC fits were compared, with results shown in Table 4.21. The ratio of the σs of the ΔE distributions is within statistical error of

Calibration	$M_{\rm BC}$	ΔE
$\sigma_{mc}/\sigma_{data}$	$1.105^{+0.045}_{-0.044}$	$1.002^{+0.097}_{-0.088}$
$\mu_{mc} - \mu_{data}$	$-0.0017^{+0.0001}_{-0.0001}$	$-0.0072^{+0.0034}_{-0.0035}$

Table 4.21: The control sample calibration results.

1 and no calibration is required. The means of the ΔE distributions differ significantly. The central value shown in Table 4.21 was used to calibrate the mean of the CBLS PDFs used to model the $B \to K\eta'\gamma$ ΔE distribution for both charged and neutral final states.

Both the σ and μ of the $K^*(892)\gamma$ MC $M_{\rm BC}$ distribution differ significantly from the values in the data distribution. The $B \to K\eta'\gamma$ CBLS functions for $M_{\rm BC}$ are calibrated using these results.

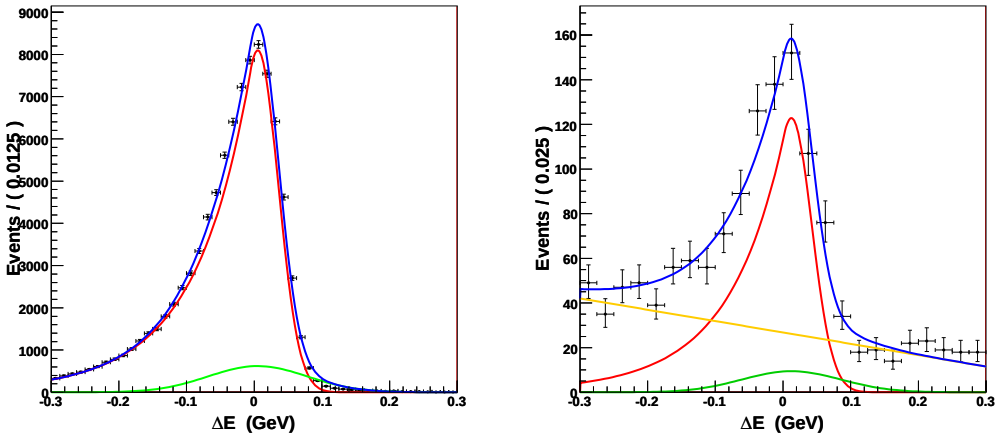


Figure 4.25: The fits to $B \to K^*(892)\gamma$ ΔE distributions. The left plot is signal MC, the right the data fit.

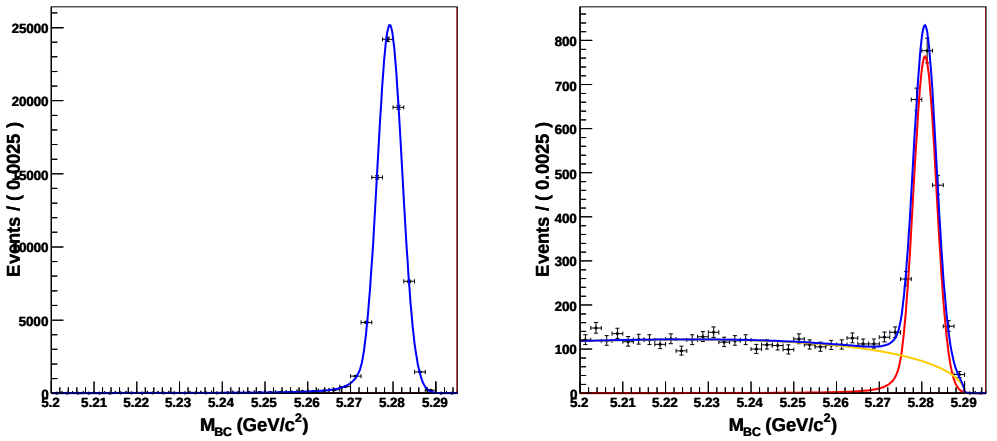


Figure 4.26: The fits to $B \to K^*(892)\gamma$ $M_{\rm BC}$ distributions. The left plot is signal MC, the right the data fit.

4.6 2D Fit to Data

The full available data set with an integrated luminosity of $604.55\,\mathrm{fb}^{-1}$ was processed using the selection cuts, continuum suppression and best candidate selection for the 2D fitting method as detailed in the preceding chapters and sections. The resulting M_{BC} and ΔE distributions were fitted in a 2D unbinned extended ML fit using the combined 2D PDFs shown in Section 4.3.3, calibrated as in Section 4.5.

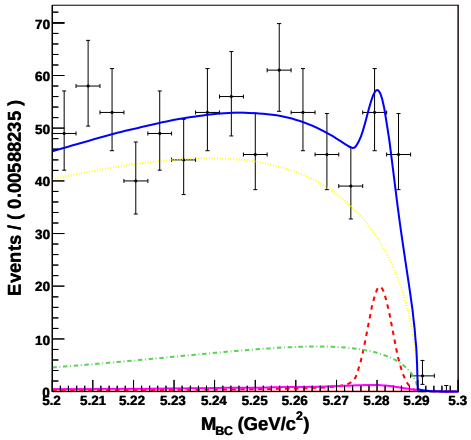
The results for the fit to charged modes is shown in Figure 4.27, the fit to neutral modes in Figure 4.28. Tables 4.22 and 4.23 give the initialised values and returned yields of the charged and neutral fits, respectively.

There is evidence of the decay $B^+ \to K^+\eta'\gamma$ with a signal significance of 3.43σ. The fit returns a signal yield of $32.61^{+11.75}_{-10.75}$ events and a $q\overline{q}$ yield of $630.0^{+28.7}_{-28.0}$ events. This is the first evidence discovered of this exclusive mode of the FCNC decay $b \to s\gamma$.

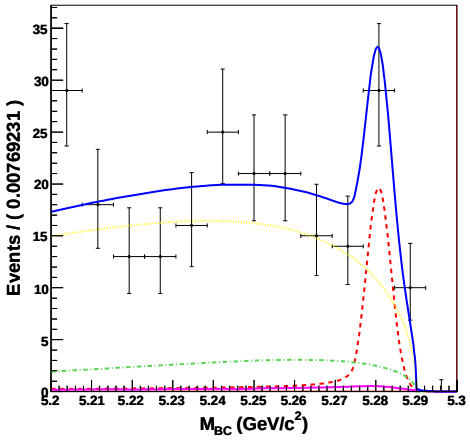
The fit to neutral modes returns a signal yield of $5.08^{+4.99}_{-3.96}$ events and a $q\overline{q}$ yield of $190.9^{+16.2}_{-15.5}$ events. The fit has a signal significance of 1.33σ.

The $M_{\eta'}$ distributions of events within the signal box for the two fits are in Figure 4.29. Both histograms have means close to the nominal $M_{\eta'}$ value of $0.95778\,\mathrm{GeV}/c^2$ [19]. Figure 4.30 shows the E_γ distributions of events within the signal box for the two fits. Figure 4.31 shows the M_{Xs} distributions of events within the signal box for the two fits. The statistics are too low for any structure to be attributed to these plots.

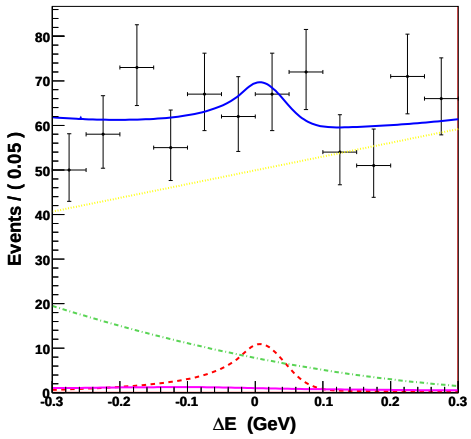
2D Fit to Data for Charged Modes.



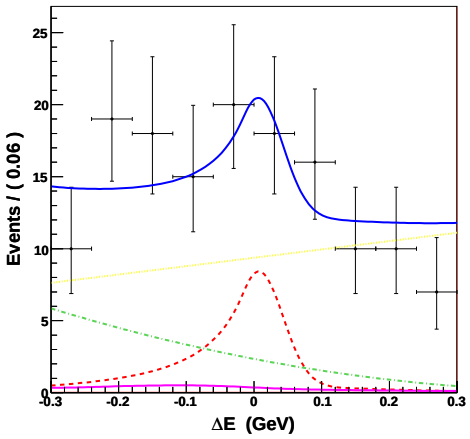
(a) Fit to the $M_{\rm BC}$ distribution of data in the full fitting region.



(b) Fit to the $M_{\rm BC}$ distribution of data in the ΔE signal region.



(c) Fit to the ΔE distribution of data in the full fitting region.



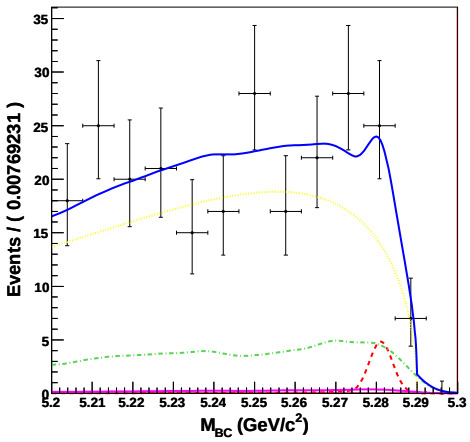
(d) Fit to the ΔE distribution of data in the $M_{\rm BC}$ signal region.

Figure 4.27: The 2D fit to data for charged modes. The $K\eta'\gamma$ function is shown in dashed red, $q\overline{q}$ in dotted orange, $b \to c$ in dash-dotted green, $b \to u, d, s$ in solid magenta, and the combined function in solid blue.

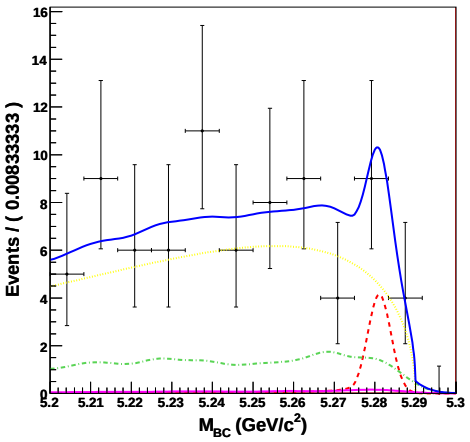
Event Type	Initialisation	Fit Yield	Significance
$K\eta'\gamma$	15	$32.61^{+11.75}_{-10.75}$	3.43σ
$q\overline{q}$	756	$630.0^{+28.7}_{-28.0}$	
$b \to c$	110	Fixed	
$b \to u, d, s$	12	Fixed	

Table 4.22: The initialised and returned fit yields, and fit significance for the 2D fit to data for the charged modes.

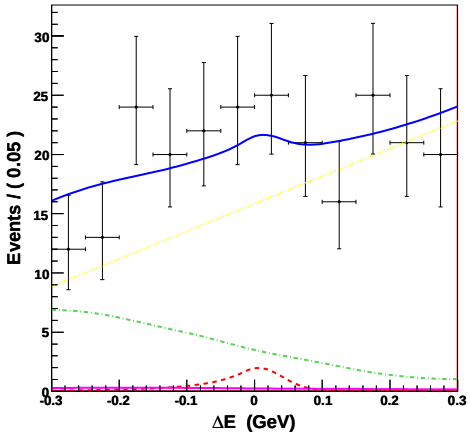
2D Fit to Data for Neutral Modes.



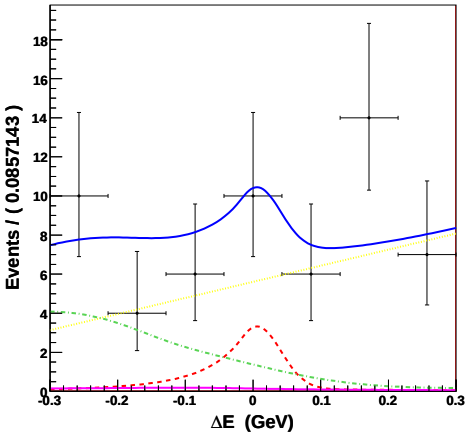
(a) Fit to the $M_{\rm BC}$ distribution of data in the full fitting region.



(b) Fit to the $M_{\rm BC}$ distribution of data in the ΔE signal region.



(c) Fit to the ΔE distribution of data in the full fitting region.

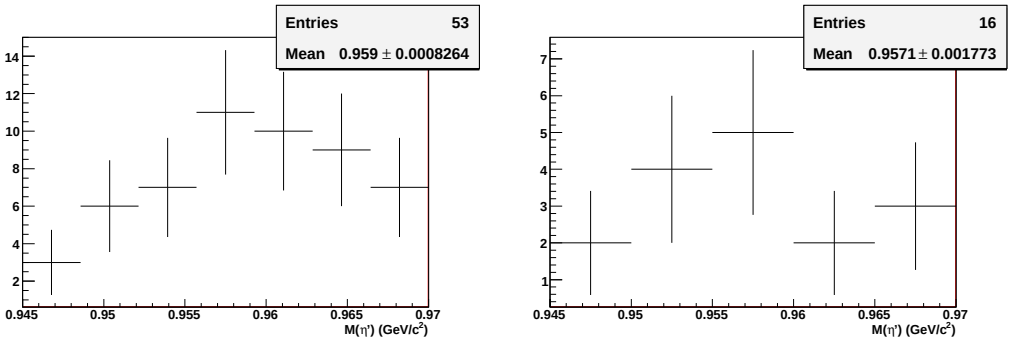


(d) Fit to the ΔE distribution of data in the $M_{\rm BC}$ signal region.

Figure 4.28: The 2D fit to data for neutral modes. The $K\eta'\gamma$ function is shown in dashed red, $q\overline{q}$ in dotted orange, $b \to c$ in dash-dotted green, $b \to u, d, s$ in solid magenta, and the combined function in solid blue. The left plots show $M_{\rm BC}$, the right show ΔE. The upper row is the full fitting region, the lower the projections to the signal region.

Event Type	Initialisation	Fit Yield	Significance
$K\eta'\gamma$	2	$5.08^{+4.99}_{-3.96}$	1.33σ
$q\overline{q}$	188	$190.9^{+16.2}_{-15.5}$	
$b \to c$	45	Fixed	
$b \to u, d, s$	3	Fixed	

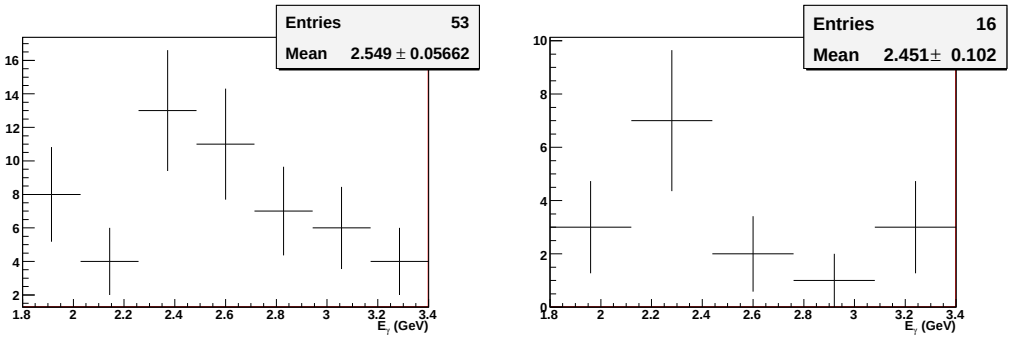
Table 4.23: The initialised and returned fit yields, and fit significance for the 2D fit to data for the neutral modes.



(a) The $M_{\eta'}$ distributions within the signal box for charged modes.

(b) The $M_{\eta'}$ distributions within the signal box for neutral modes.

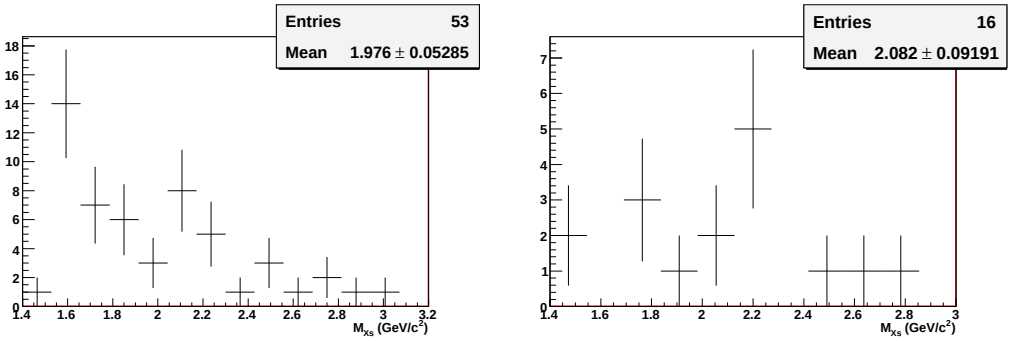
Figure 4.29: The $M_{\eta'}$ distributions within the signal box for data.



(a) The E_γ distributions within the signal box for charged modes.

(b) The E_γ distributions within the signal box for neutral modes.

Figure 4.30: The E_γ distributions within the signal box for data.



(a) The $M_{X\,s}$ distributions within the signal box for charged modes.

(b) The $M_{X\,s}$ distributions within the signal box for neutral modes.

Figure 4.31: The $M_{X\,s}$ distributions within the signal box for data.

4.7 Bias study

The toy MC and GSIM ensemble tests performed in Section 4.4 were based on the amounts of $K\eta'\gamma$ and background expected from MC efficiencies and the result from the previous $K\eta'\gamma$ analysis. Tests based on the returned yields from data provide a more accurate assessment of any bias inherent in the fitting method.

The same techniques for the generation and fitting of toy MC samples were used. 2500 charged and 2500 neutral toy MC samples were generated from the PDF shapes and yields found in the fits to data. The yields were rounded to the nearest whole number. The signal pull and yield distributions from the fits to these samples are shown in Figure 4.32.

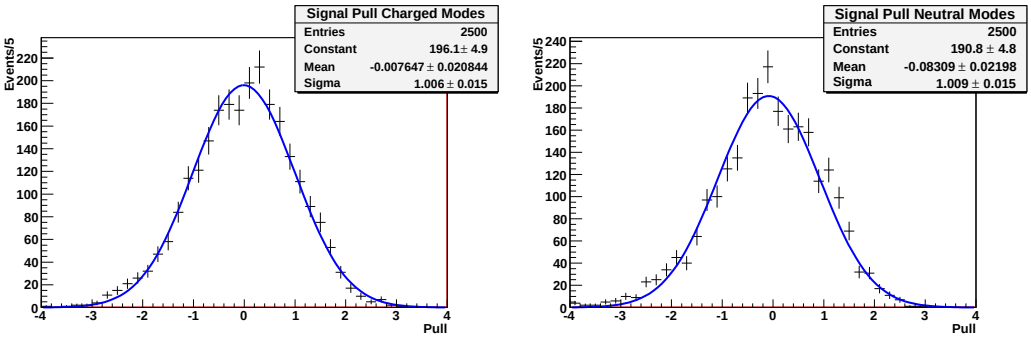
The test of the charged modes fit showed no bias in the pull distribution. The yield distribution is 0.24 events or 0.72% from the true yield of 33 events, at a significance of 1.14σ. No correction of the signal yield or any additional error was required based on these results.

The test of the neutral modes fit shows a negative bias in both the pull and the yield distributions. The bias in the yield was equivalent to 0.171 events, or 3.42%, at a significance of 1.88σ. The mean of the pull distribution was 3.78σ away from zero. This is not large enough to require correction of the signal yield but this bias and the similar one evident in the GSIM test described below were included in the systematic errors calculated in Section 5.2.

A GSIM ensemble test was also performed using a different technique to the one used in Section 4.4.2. To increase the statistical accuracy of the test, toy MC samples were generated for the $q\overline{q}$ and $b \to c$ components instead of using GSIM MC. Using toy MC for theses two backgrounds enabled the generation of many more pseudo-experiments. This combined toy/GSIM MC does not test the $q\overline{q}$ or $b \to c$ backgrounds for any source of bias that toy MC is insensitive to. It does test the response of the 2D fitter across different levels of embedded $K\eta'\gamma$ events to much higher accuracy than the full GSIM MC test. Any bias stemming from the $q\overline{q}$ and $b \to c$ components would have been visible as a constant across all points of the previous fully GSIM ensemble test (Figure 4.24). The small fluctuations found in that test vary across the points, and can thus be deduced to originate in the only varying component: the $K\eta'\gamma$ MC. The use of toy MC instead of GSIM MC for some background components will not affect this.

This assumption was tested before the combined toy/GSIM MC test was trusted. The agreement between the fully GSIM method and the toy/GSIM combination was assessed by comparing the results over a limited number of pseudo-experiments. 23 fully independent $q\overline{q}$ and $b \to c$ GSIM MC samples were available for both charged and neutral fits, based on the yields obtained from the fit to data. Toy MC samples were generated of the same size and number. The $b \to u, d, s$ components were randomly selected from the large GSIM MC sample available. Samples of $K\eta'\gamma$ MC were randomly selected with size ranging from 18 to 45 for the charged modes, and 0 to 27 for the neutral. The final states included in the $K\eta'\gamma$ MC were weighted according to their expected ratios based on their efficiencies and daughter branching fractions. Samples of each increment were then combined with the different background MC samples and

fitted.



(a) Pull distribution for the charged modes.

(b) Pull distribution for the neutral modes.

(c) Yield distribution for the charged modes.

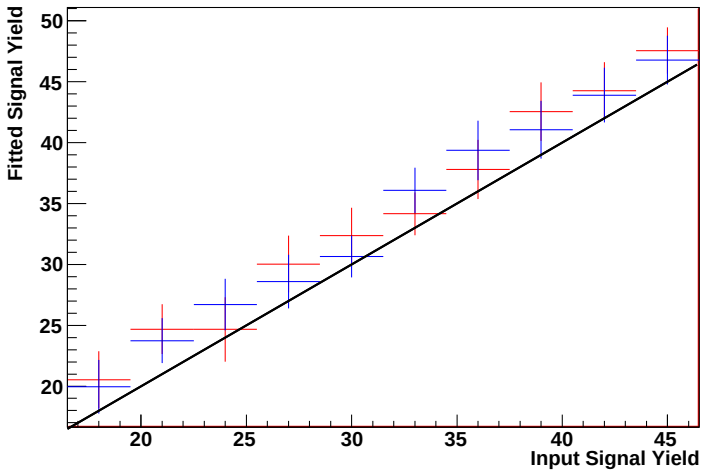
(d) Yield distribution for the neutral modes.

Figure 4.32: The pull and yield distributions from the data based toy MC. The top row is the pull distributions, the bottom the yield distributions. The left plots show the charged modes, the right plots the neutral modes.

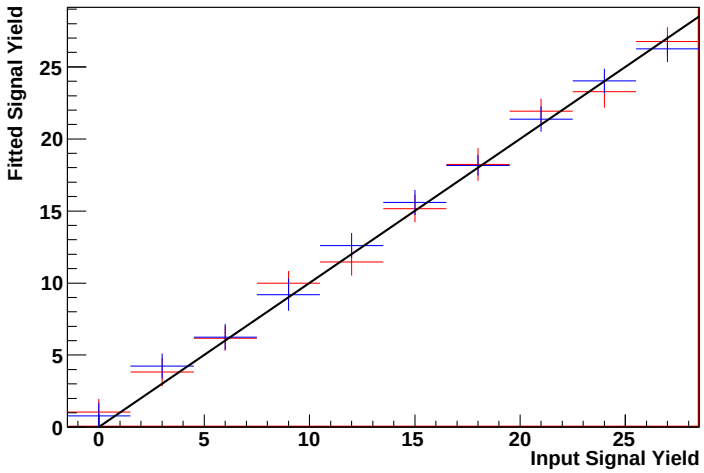
Figure 4.33 compares the results for the two methods. The blue points are the mean returned yields for each increment of included $K\eta'\gamma$ MC using entirely GSIM MC for background, the red points are from the combined GSIM/toy MC samples. The black line indicates the unbiased line $y = x$. All points from the two methods agree within the statistical error of the study with no sign of any consistent shift in the mean. The charged final states signal yield was over-estimated by both of the methods, though not to a highly significant degree. No difference in the bias estimation was visible between the two forms of samples in the fits to the charged or neutral final states. From these results the combined GSIM/toy MC method was taken to be as reliable as the full GSIM method.

The number of pseudo-experiments was increased to 500 and the combined GSIM/toy MC tests run again. The results are shown in Figure 4.34. The mean returned yields for the charged final states generally showed a small positive bias, similar to the low statistics tests. The parameter p_0 from the linear fit was 1.03σ from zero and the slope was within error of one: no correction was applied to the signal yield. Instead the p_0 central value was taken as a systematic error. This was equivalent to 1.5%.

The results for the neutral modes showed no bias in the region of interest around 5 signal events, however there was an increasing tendency to underestimate the true number of signal events as event number rose. Correcting the data yield by the slope of the linear fit gave a modified yield of 4.91. This is in good agreement with the toy MC pull results shown above. No correction was applied to the signal yield; instead, the larger of the measured deviations in the two tests was taken as a systematic error. This was the slope of the linear fit in the toy/GSIM ensemble test, equal to an uncertainty of 3.5%.

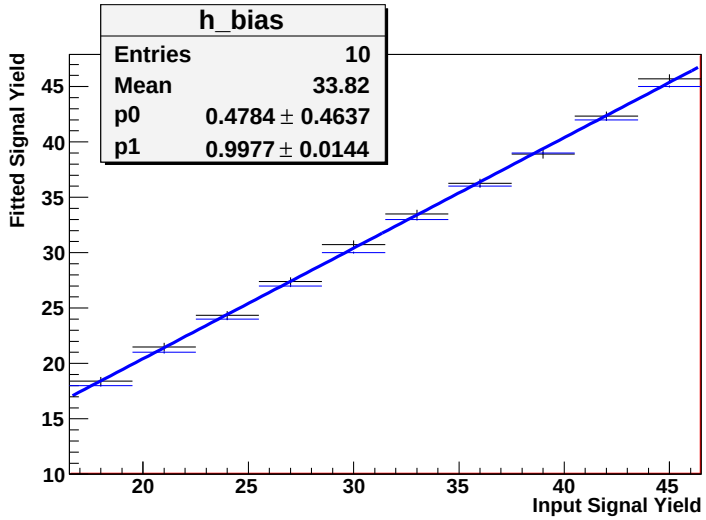


(a) Charged modes.

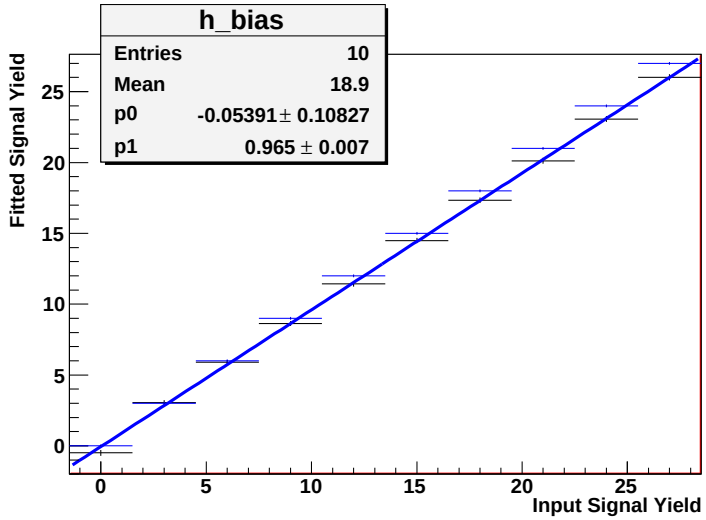


(b) Neutral modes.

Figure 4.33: Comparison of the ensemble test methods. The red points are the means of the returned yields from the combined GSIM/toy MC method, the blue from the full GSIM method. The black line shows $y = x$.



(a) Charged modes.



(b) Neutral modes.

Figure 4.34: The results of the data based GSIM/toy MC ensemble test. The y-axis is the returned signal yield, the x-axis the true number of $K\eta'\gamma$ events. The black points with vertical error bars are the means of the 20 fits to GSIM MC for each number of $K\eta'\gamma$ events. The blue points are the true number of $K\eta'\gamma$ events. The blue line is the fit to the returned yield means.

Chapter 5

Efficiencies, Systematic Errors and Branching Fractions

5.1 Efficiency calibration

The MC production process is not a perfect one and there is expected to be some disparity between simulated events and real data in decay dynamics, detector response and particle reconstruction. These differences can bias the signal efficiency (ε) which will in turn change the measured branching fraction of $B \to K\eta'\gamma$. This section describes the calibration of ε.

5.1.1 M_{Xs} efficiency calibration

In the $B \to K\eta'\gamma$ MC generation process described in Section 3.1.2 the X_s pseudo-resonance is modelled with a flat mass distribution between $1.52\,\mathrm{GeV}/c^2$ and $2.7\,\mathrm{GeV}/c^2$. The estimation of the signal reconstruction efficiency from $K\eta'\gamma$ MC (Table 4.17) is based on the assumption that this modelling is an accurate representation of the true M_{Xs} distribution in data. Figure 4.31 in the previous Chapter shows the data distribution of M_{Xs} within the signal region does not follow this assumption. If the efficiency has any dependence on M_{Xs} then the MC efficiency estimation will be incorrect.

Figure 5.1 shows the generated versus reconstructed M_{Xs} distributions in $K\eta'\gamma$ MC. There is a strong correlation between M_{Xs} and the number of reconstructed events. Figure 5.2 shows ε calculated in 10 bins of M_{Xs} from $1.52\,\mathrm{GeV}/c^2$ to $2.7\,\mathrm{GeV}/c^2$ in $K\eta'\gamma$ MC. The efficiency is calculated as the number of events reconstructed in a certain M_{Xs} bin divided by the number of events generated in that bin, with each final state weighted according to its proportional daughter branching fraction. The generated and reconstructed events in each bin are correlated variables; the Poissonian errors on each are summed as such and are shown as error bars in Figure 5.1.

The efficiencies in Table 4.17 are the means of the efficiency across the M_{Xs} distribution. To obtain more realistic figures for ε, background-subtracted M_{Xs} distributions are weighted by the efficiency distributions in Figure 5.2. The M_{Xs} distributions in data from the sideband and full fitting regions for the charged modes are shown in Figure 5.3 and for the neutral modes in Figure 5.4. The sideband distributions are weighted according to the ratio:

$$\frac{(\#q\overline{q}\text{ MC events in fitting region})}{(\#q\overline{q}\text{ MC events in sideband region})} \tag{5.1}$$

and subtracted from the full fitting region distributions to create the background-subtracted distributions shown in Figure 5.5. Similar structure is observed in the background-subtracted distribution and the distribution from the signal region for the charged final states (Figure 4.31).

Each event in the background-subtracted distributions is assigned an efficiency based on the M_{Xs} bin in which the event lies. If an event is outside the range $1.52\,\mathrm{GeV}/c^2$ to $2.70\,\mathrm{GeV}/c^2$ then the efficiency of the closest bin is assigned. The weighted average of all the events is taken as the global efficiency:

$$\varepsilon = \frac{1}{N}\sum_{N}\varepsilon_{bin} \tag{5.2}$$

where N runs over all events in the background-subtracted distribution, and ε_{bin} is the efficiency of the bin in which the event lies.

The error on ε is estimated by varying the efficiency assigned to each bin by the errors in Figure 5.2 and calculating the change in ε. All errors are then added in quadrature. The efficiencies calculated are 0.0259 ± 0.0004 for the charged final states and 0.0180 ± 0.0003 for the neutral.

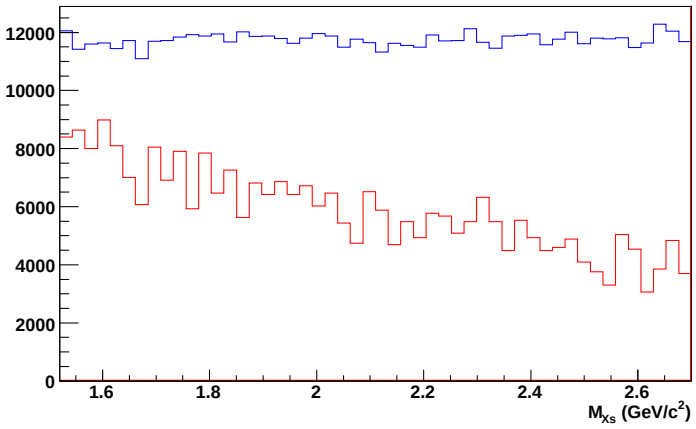
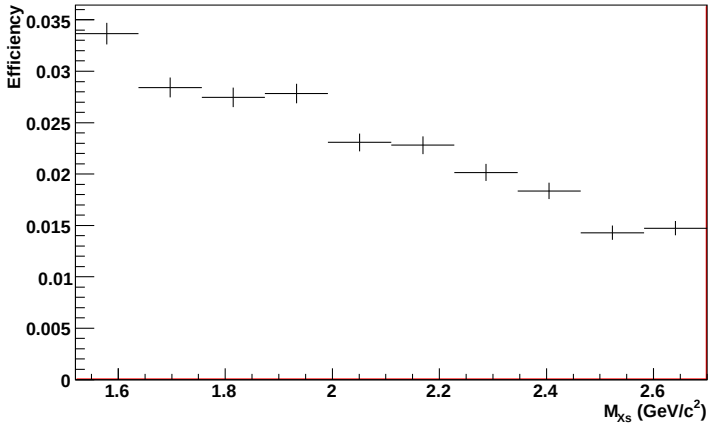
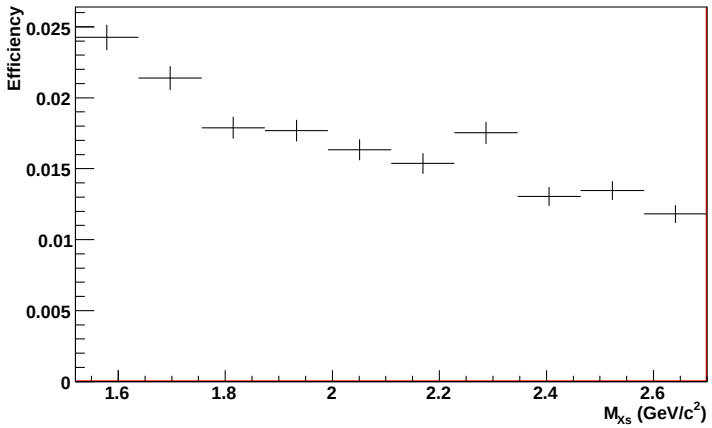


Figure 5.1: The generated M_{Xs} (blue) and the M_{Xs} of events passing all selection cuts (red) in $K\eta'\gamma$ MC. The red histogram has been normalised to half the number in the blue histogram. The M_{Xs} effi ciency dependence is clearly visible.

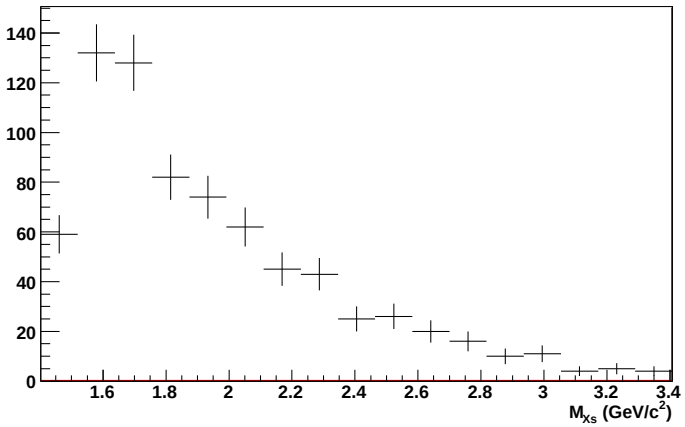


(a) Charged modes.

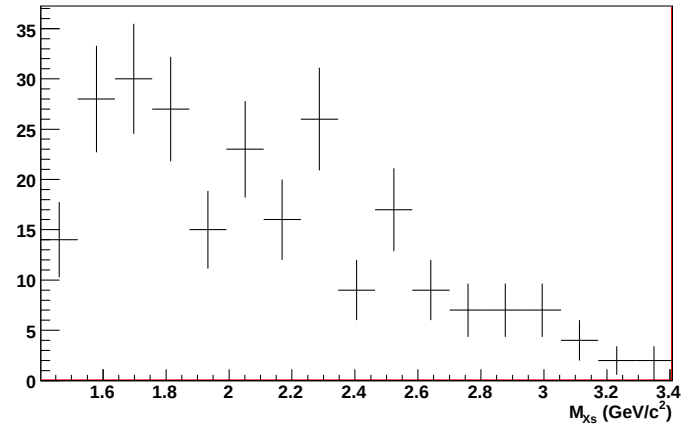


(b) Neutral modes.

Figure 5.2: The signal reconstruction effi ciency as a function of the reconstructed M_{Xs}.

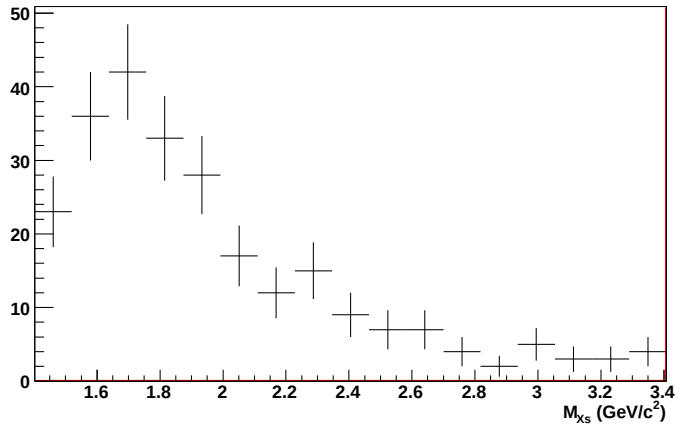


(a) M_{Xs} distribution in the full fitting region.

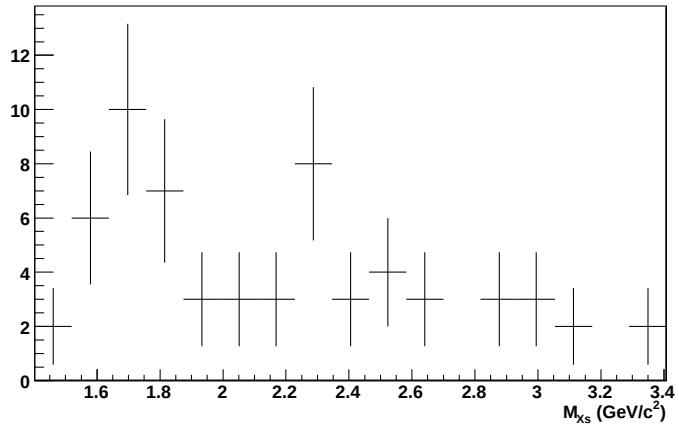


(b) M_{Xs} distribution in the sideband region.

Figure 5.3: The M_{Xs} distributions for charged modes in data. The upper plots are the full fitting region, the lower from the sideband region.

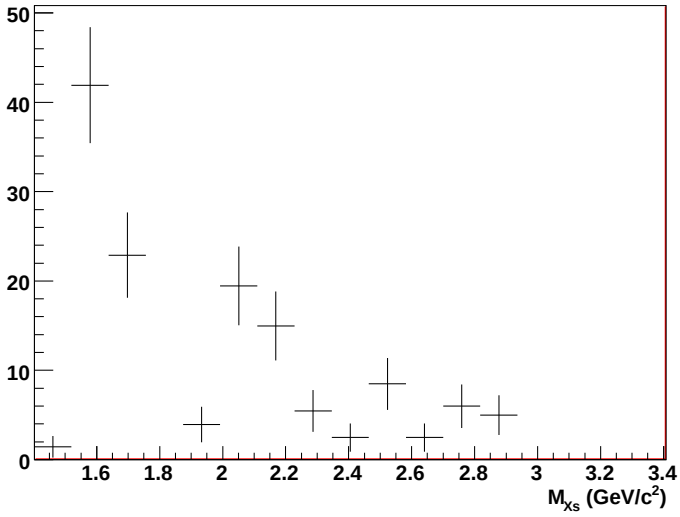


(a) M_{Xs} distribution in the full fitting region.

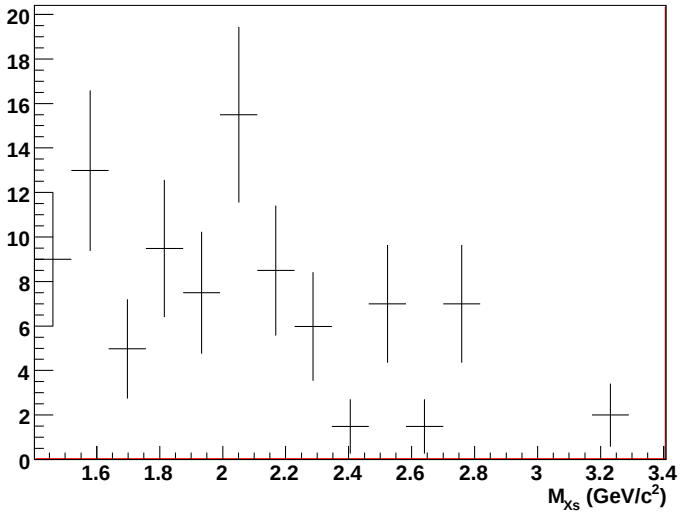


(b) M_{Xs} distribution in the sideband region.

Figure 5.4: The M_{Xs} distributions for neutral modes in data. The upper plots are the full fitting region, the lower from the sideband region.



(a) Charged modes.



(b) Neutral modes.

Figure 5.5: The background-subtracted M_{Xs} distributions for data in the full fitting region.

5.1.2 K/π ID efficiency calibration

The difference in the K/π ID efficiency between MC and data is expected to be significant. It is dependant on the specific $\mathcal{R}_{K,\pi}$ selection cuts used and must be assessed for each analysis independently. The K/π ID ε ratio is calculated using a program provided by the Belle particle ID (PID) group [65] based on a 142 fb^{-1} sample of inclusive D^{*+} decays to $D^0\pi^+ \rightarrow (K^-\pi^+)\pi^+$. This mode provides numerous charged kaons and pions in a clean environment. After selection criteria are placed on the D^{*+} and D^0, the ratio of K^+ (π^+) before and after each $\mathcal{R}_{K,\pi} < / > 0.1, 0.2, ...0.9$ cut are taken as the K/π ID efficiencies. This is done for both the 142 fb^{-1} data sample and a large MC sample and the data/MC efficiency ratios calculated.

Efficiency ratios are provided from the study in fine bins of particle momentum and direction for each $\mathcal{R}_{K,\pi}$ cut. Each kaon and pion in $K\eta'\gamma$ MC that passes the selection criteria is assigned an ε ratio based on it's momentum, direction and the $\mathcal{R}_{K,\pi}$ cut applied. The corrections are then calculated as a weighted average of the ε ratios:

$$R_\epsilon = \frac{1}{N}\sum_l n_l R_l \tag{5.3}$$

where l denotes the $p : \cos\theta : \mathcal{R}_{K,\pi}$ bin, R_l is the ε ratio in bin l, n_l is number of candidates falling within bin l, and N is the total number of candidates.

Three corrections were thus calculated; a K^+ ID ε correction ($R_\epsilon^{K^+}$) and π^+ ID ε corrections ($R_\epsilon^{\pi^+}$) for both the charged and neutral final states. This was done for SVD1 and SVD2 separately as the charged particle tracking changes substantially between the detectors; the correction is then the average value of the two weighted according to the total SVD1 and SVD2 integrated luminosity.

$R_\epsilon^{\pi^+}$ for both the charged and neutral final states is a function of the number of π^+ in the $K\eta'\gamma$ MC sample that passes selection. The correction for each reconstructed final state is the product of the calculated R_ϵ of each π^+ in that final state. This product is then weighted by the proportion of that final state in the $K\eta'\gamma$ MC sample that passes selection, and the sum of the three included final states taken as the total $R_\epsilon^{\pi^+}$ for each efficiency:

$$R_\epsilon^{\pi^+} = \sum_m (\prod_n R_\epsilon^n) Frac_m \tag{5.4}$$

	$R_\epsilon^{K^+}$	$R_\epsilon^{\pi^+}$	Total correction
$K^+\eta'\gamma$	1.0037 ± 0.0086	0.9784 ± 0.0111	0.9820 ± 0.0140
$K_S^0\eta'\gamma$	NA	0.9510 ± 0.0138	0.9510 ± 0.0138

Table 5.1: The results for the K/π ID data/MC ε ratio study.

where m runs over the included final states, n runs over each π^+ in final state m, R_ϵ^n is the efficiency ratio for the n^{th} π^+, and $Frac_m$ is the fraction of mode m in the $K\eta'\gamma$ MC sample passing selection. There is never more than one K^+ in any reconstructed mode, so no more is needed for $R_\epsilon^{K^+}$.

The total correction for each signal ε is then the product of $R_\epsilon^{\pi^+}$ and $R_\epsilon^{K^+}$. The results are shown in Table 5.1. The central values are used to correct the signal reconstruction efficiencies calculated in the last Section, and the errors are included in the systematic error calculations in Section 5.2.

5.1.3 $\mathcal{LR}$ efficiency calibration

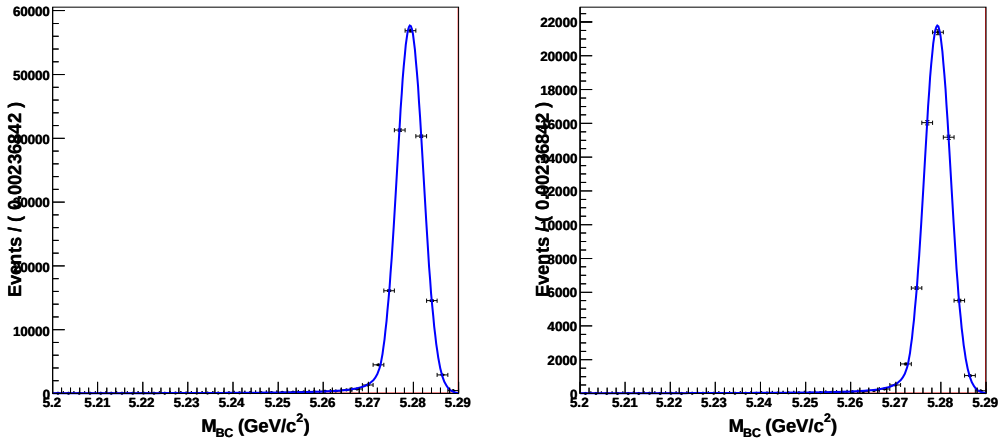
The efficiencies of the $\mathcal{LR}$ cuts are also expected to be significantly different for data and MC. They are calibrated using the same $B \to K^*(892)\gamma$ control sample as was used for the $K\eta'\gamma$ PDF calibration in Section 4.5. The efficiencies of both the charged and neutral final states are calibrated using the same sample.

The Fisher co-efficients and $\mathcal{LR}$ cut values were trained and optimised using $B \to K^*(892)\gamma$ MC and sideband data. Extended unbinned ML fits to $M_{\rm BC}$ with a requirement of $\Delta E > -0.1\,{\rm GeV}$ and $\Delta E < 0.07$ are used to extract the yields before and after the $\mathcal{LR}$ cuts for data and MC.

A CBLS PDF was used to fit the $M_{\rm BC}$ distribution of the $B \to K^*(892)\gamma$ MC. A CBLS with α_{CB} and N_{CB} fixed to the results of the MC fit was used to fit the data. An ARGUS function with a fixed endpoint and floating shape parameter was added to model the background in the fit to data. The fit results are shown in Figure 5.6.

The data/MC ε ratio for the $\mathcal{LR}$ cut is calculated to be $0.9811^{+0.0375}_{-0.0372}$. The ratio is within statistical error of 1.0, so the central value is not used to calibrate the efficiency of $K\eta'\gamma$. The error is included as a systematic in Section 5.2.

The final ε for the 2D fit to the charged final states is 0.0254 ± 0.0004. For the neutral states it is 0.0171 ± 0.0003.



(a) Fit to MC before $\mathcal{LR}$ cuts.

(b) Fit to MC after $\mathcal{LR}$ cuts.

(c) Fit to data before $\mathcal{LR}$ cuts.

(d) Fit to data after $\mathcal{LR}$ cuts.

Figure 5.6: Fits to $B \to K^*(892)\gamma$ data and MC $M_{\rm BC}$ distributions.

5.2 Systematic error study

The systematic errors are intended to encompass all the differences between data and MC that are too small to require calibration as well as the uncertainty involved in the fitting process. The total systematic errors are shown in Table 5.2 below. The total errors are the combination of the systematics in Table 5.3 which effect the efficiency calculation, and the systematics in Table 5.5 which concern the uncertainty in the signal yield. The latter type of errors are calculated by varying the data fit parameters and observing the effect on the signal yield. All errors are assumed to be uncorrelated; the total error is the quadratic sum of the individual errors. The source of each error and the methods of estimation are discussed in the following Section.

Error Type	$B^+ \to K^+\eta'\gamma$		$B^0 \to K^0_S\eta'\gamma$	
	$+\sigma(\%)$	$-\sigma(\%)$	$+\sigma(\%)$	$-\sigma(\%)$
Total error	10.09	-11.76	16.51	-19.36

Table 5.2: Total calculated systematic errors for $B \to K\eta'\gamma$.

5.2.1 MC Efficiency

The error on the signal ε after calibration using the M_{Xs} distribution and the K/π ID data/MC efficiency ratio is taken as a systematic error.

Error Type	$B^+ \to K^+\eta'\gamma$		$B^0 \to K^0_S\eta'\gamma$	
	$+\sigma(\%)$	$-\sigma(\%)$	$+\sigma(\%)$	$-\sigma(\%)$
MC Efficiency	1.47	-1.47	1.85	-1.85
K/π ID efficiency	1.41	-1.41	1.45	-1.45
Signal γ reconstruction efficiency	2.80	-2.80	2.80	-2.80
π^0 reconstruction efficiency	0.80	-0.80	0.47	-0.47
K^0_S reconstruction efficiency	-	-	4.50	-4.50
η reconstruction efficiency	3.43	-3.43	3.61	-3.61
Track reconstruction efficiency	3.84	-3.84	4.97	-4.97
$\mathcal{LR}$ cut efficiency	3.75	-3.72	3.75	-3.72
$N_{B\overline{B}}$ error	1.36	-1.36	1.36	-1.36
J/ψ Veto	0.22	-0.22	0.39	-0.39
D^0 Veto	0.54	-0.54	-	-
Cross-Feed	-	-6.00	6.00	-
Total error	7.44	-9.56	11.12	-9.36

Table 5.3: Systematic errors for $B \to K\eta'\gamma$, independent of the data fit.

5.2.2 K/π ID efficiency

The error is taken from the K/π ID efficiency calibration described in the last Section.

5.2.3 Signal γ reconstruction efficiency

A data/MC difference of $\pm 2.8\%$ on the reconstruction efficiency of the signal γ is found from study of radiative Bhabha ($e^+e^- \to e^+e^-\gamma$) events [62]. $29.24\,\mathrm{fb}^{-1}$ of data and a large MC sample are analysed. Radiative Bhabha events are selected with no photon requirements. The selection criteria are;

- There are only two charged tracks in the event, with opposite charge. Both tracks need to satisfy electron ID requirements.
- The total ECL cluster energy is between $11\,\mathrm{GeV}$ and $12\,\mathrm{GeV}$, indicating that all the event energy is carried by electrons, positrons and photons.
- The vector of missing momentum from the two charged tracks is within $33^\circ < \theta < 128^\circ$, has energy between $2\,\mathrm{GeV}$ and $3\,\mathrm{GeV}$, and is separated from both the tracks by more than 20°.
- Only single clusters are associated with each of the charged tracks with $E_9/E_{25} < 0.95$ to remove final state radiation and bremsstrahlung events.
- The mm^2 of the missing vector is within $-0.2\,\mathrm{GeV}/c^2 < mm^2 < 0.3\,\mathrm{GeV}/c^2$. This is to remove radiative Bhabha events in which there is more than one photon. The mm^2 of these events is shifted to the positive.

The most energetic ECL cluster within 20° of the missing momentum vector is taken as the candidate photon. The energy and direction of the photon should be close to the missing energy and missing momentum vector calculated from the two charged tracks. The efficiency is defined as the number of events that have a candidate photon of energy within -0.1 GeV to +0.08 GeV of the missing energy divided by the total number of events that pass the radiative Bhabha selection criteria. The data/MC efficiency ratio is found to be 1.01 ± 0.02. A study of the systematics of this method estimates a 2% error; the quadratic sum of this error and the error on the efficiency ratio is taken as the total uncertainty on the signal photon reconstruction efficiency.

5.2.4 π^0 reconstruction efficiency

The π^0 reconstruction efficiency data/MC difference is assessed from a study of inclusive η decays [63]. Two methods are used to measure the efficiency ratio:

Method 1:

$$\frac{\varepsilon_{Data}(2\pi^0)}{\varepsilon_{MC}(2\pi^0)} = \frac{N_{Data}(\eta \to 3\pi^0)/N_{MC}(\eta \to 3\pi^0)}{N_{Data}(\eta \to \gamma\gamma)/N_{MC}(\eta \to \gamma\gamma)} \tag{5.5}$$

which assumes $\varepsilon_{Data}(\pi^0 \to \gamma\gamma)/\varepsilon_{MC}(\pi^0 \to \gamma\gamma) = \varepsilon_{Data}(\eta \to \gamma\gamma)/\varepsilon_{MC}(\eta \to \gamma\gamma)$.

Method 2:

$$\frac{\varepsilon_{Data}(2\pi^0)}{\varepsilon_{MC}(2\pi^0)} = \frac{N_{Data}(\eta \to 3\pi^0)/N_{MC}(\eta \to 3\pi^0)}{N_{Data}(\eta \to \pi^+\pi^-\pi^0)/N_{MC}(\eta \to \pi^+\pi^-\pi^0)} \tag{5.6}$$

which assumes $\varepsilon_{Data}(\pi^0 \to \gamma\gamma)/\varepsilon_{MC}(\pi^0 \to \gamma\gamma)_{\eta\to\pi^+\pi^-\pi^0} = \varepsilon_{Data}(\pi^0 \to \gamma\gamma)/\varepsilon_{MC}(\pi^0 \to \gamma\gamma)_{\eta\to 3\pi^0}$. This second method also assumes that the charged pion tracking efficiency is the same for MC and data. Both methods assume that the systematics involved in photon detection are cancelled out by taking the ratios. The single π^0 efficiency ratio is then

$$\frac{\varepsilon_{Data}(\pi^0)}{\varepsilon_{MC}(\pi^0)} = \sqrt{\frac{\varepsilon_{Data}(2\pi^0)}{\varepsilon_{MC}(2\pi^0)}} \tag{5.7}$$

The selection criteria for the decays are:

- $\eta \to \gamma\gamma$ photon energies must be greater than $50\,\mathrm{MeV}$ and the η momentum must be greater than $2\,\mathrm{GeV/c}$ in the lab frame.
- The π^0 decay angle (θ) of $\eta \to 3\pi^0$ must satisfy $|\cos\theta| < 0.5$.
- Pairs of photons with invariant mass in the range $0.06\,\mathrm{GeV}/c^2 < M_{\gamma\gamma} < 0.26\,\mathrm{GeV}/c^2$ are taken as π^0 candidates.

The invariant mass distributions of the three decay modes $\eta \to \gamma\gamma$, $\eta \to \pi^+\pi^-\pi^0$ and $\eta \to 3\pi^0$ in data and MC are then fitted and the ratios shown above calculated for different γ energy and π^0 momentum requirements.

The selection cuts of $E_\gamma > 50\,\mathrm{MeV}$ and $p_{\pi^0} > 100\,\mathrm{MeV}$ in the $B \to K\eta'\gamma$ analysis give a π^0 reconstruction efficiency ratio of $0.979 \pm 0.018 \pm 0.02$ where the first error is statistical and the second systematic. This is the result of Method 1; Method 2 agrees within error. The $K\eta'\gamma$ signal efficiency is not corrected from this result; instead the quadratic sum of the two errors on the efficiency ratio is taken as a systematic error. This translates to a $\pm 2.69\%$ error on $B \to K\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$. This is then scaled according to the proportion of these decays in the MC sample used to estimate the efficiencies, giving $\pm 0.80\%$ for the charged final states and $\pm 0.47\%$ for the neutral.

5.2.5 Track reconstruction efficiency

The reconstruction efficiency of charged tracks in the detector is calculated from a study of partially versus fully reconstructed D^* decays for tracks of momentum greater than $250\,\mathrm{MeV}/c$ [60], and a GSIM MC based study for those with momentum $< 250\,\mathrm{MeV}/c$ [61]. The former technique reconstructs the decay chain

$$\begin{aligned} D^* \rightarrow & D^0 \pi_{slow} \\ & D^0 \rightarrow \pi^+ \pi^- K^0_S \\ & K^0_S \rightarrow \pi^+ \pi^- \end{aligned} \tag{5.8}$$

where one of the pions from the K^0_S decay is not included in the reconstruction. In similar fashion to the signal photon reconstruction efficiency study, the masses of the K^0_S and the D^0 and the momentum vector of the K^0_S are used to constrain the momentum vector of the missing pion (π_{miss}), and the efficiency is defined as the ratio of the number of found π_{miss} to the total number of D^* decays.

The MC-based study involves embedding a single charged pion MC track in full events of data and MC and searching for the track within the noisy environment. The success rate of the searches in data and MC are compared. This technique is very accurate but the reconstruction of the embedded MC track can introduce other data/MC discrepancies and it is only used in the momentum region for which the D^* analysis has the poorest resolution.

The results of these studies are contained in a program provided by the tracking group that calculates the tracking efficiency error based on particle type and momentum. The results for other charged particles are inferred from the results for charged pions. The efficiency ratios are experiment-dependant, so $B \rightarrow K\eta'\gamma$ MC was generated based on the parameters of each Belle experiment individually and weighted according to each experiments integrated luminosity. The tracking program was run over this MC and the mean error calculated for each track in each reconstructed final state. The mean errors of all tracks in a certain final state were linearly summed, giving a total error for each final state. A weighted average based on the proportion of each final state in the $K\eta'\gamma$ MC sample used to calculate the signal efficiency is taken as the total tracking efficiency error for each measured branching fraction: $\pm 3.84\%$ for the charged final states, and 4.97% for the neutral.

5.2.6 K^0_S reconstruction efficiency

The error on the reconstruction of the K^0_S mesons is assessed from a study of $D^+ \rightarrow K^0_S \pi^+$ and $\rightarrow K^- \pi^+ \pi^+$ decays [64]. The data/MC efficiency ratio is calculated as

$$\begin{aligned} \frac{\varepsilon_{Data}(K^0_S)}{\varepsilon_{MC}(K^0_S)} = & \frac{\varepsilon_{Data}(K^+ ID)}{\varepsilon_{MC}(K^+ ID)} \times \left(\frac{\varepsilon_{Data}(track)}{\varepsilon_{MC}(track)} \right)^2 \\ & \times \frac{N_{Data}(D^+ \rightarrow K^0_S \pi^+)}{N_{MC}(D^+ \rightarrow K^0_S \pi^+)} \times \frac{N_{MC}(D^+ \rightarrow K^- \pi^+ \pi^+)}{N_{Data}(D^+ \rightarrow K^- \pi^+ \pi^+)} \end{aligned} \tag{5.9}$$

where $K^+ ID$ signifies the charged kaon ID selection and $track$ signifies the charged track reconstruction. The selection criteria for the D^+ decays are

- K^0_S candidates must pass the **goodKS** cuts.
- $p_t > 0.2\,\mathrm{GeV/c}$ for all charged tracks except the K^0_S daughter pions.
- π^0 candidates must be within the range $130\,\mathrm{MeV}/c^2 < M_{\gamma\gamma} < 141\,\mathrm{MeV}/c^2$.
- $|M_{K^-\pi^+\pi^+} - M_{D^+}| < 24\,\mathrm{MeV}/c^2$.

The signal yields N_{MC} and N_{Data} are extracted from fits to the $M_{K^0_S\pi^+\pi^0} - M_{K^0_S\pi^+}$ and $M_{K^+\pi^-\pi^-\pi^0} - M_{K^+\pi^-\pi^-}$ mass difference distributions. The data/MC efficiency ratio is found to be 0.979 ± 0.045. The error is taken as the systematic error on K^0_S reconstruction efficiency.

5.2.7 η reconstruction efficiency

The $\eta \to \gamma\gamma$ reconstruction efficiency is assumed to suffer from the same systematics as π^0 and is estimated in the same fashion. From the $E_\gamma > 100\,\mathrm{MeV}$ selection cut the $\eta \to \gamma\gamma$ reconstruction efficiency ratio is taken as 0.954 ± 0.023 [63]. The signal ε is not adjusted as not all reconstructed final states include an $\eta \to \gamma\gamma$ decay. To be conservative the central value minus the error is taken as the systematic error on $B \to K\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$. The central value translates to a $\pm 6.90\%$ error which is then scaled according to the proportion of these decays in the MC sample used to calculated the signal efficiency. This gives $\pm 1.27\%$ for the charged final states and $\pm 3.61\%$ for the neutral.

The uncertainty on the $\eta \to \pi^+\pi^-\pi^0$ reconstruction efficiency is assumed to be encompassed by the π^0 reconstruction uncertainty and the charged particle tracking uncertainty of the π^+ π^- pair.

5.2.8 $\mathcal{LR}$ cut efficiency

This error is taken from the $\mathcal{LR}$ efficiency calibration as described in the Section 5.1.

5.2.9 $N_{B\overline{B}}$ error

This is an official Belle statistic taken from the Belle website [66]. For experiments 7-55 the number of $B\overline{B}$ pairs is $(656.725 \pm 8.940) \times 10^6$ which equates to an error of $\pm 1.36\%$.

5.2.10 J/ψ Veto

The number of $K\eta'\gamma$ MC events removed by the $M_{\eta'\gamma}$ veto around the J/ψ mass was varied by the Poissonian error at 3σ and the effect on the efficiency was taken as the systematic error.

5.2.11 D^0 Veto

The number of $K\eta'\gamma$ MC events removed by the $M_{K^-\pi^+}$ veto around the D^0 mass was varied by the Poissonian error at 3σ and the effect on the efficiency was taken as the systematic error.

5.2.12 Cross-feeds

The possibility of a $B \to K\eta'\gamma$ decay being reconstructed as an incorrect final state is known as 'cross-feed' and was investigated in the $K\eta'\gamma$ MC sample passing selection. Table 5.4 lists the true versus reconstructed decay mode for the six $K\eta'\gamma$ final states. In this analysis the effect of cross-feed is entirely due to the mis-identification of K^+ as K^0_S or *vice versa*, as any decay with a wrongly reconstructed η' but correct K will still enter the correct fit distribution. It was found that of the $K\eta'\gamma$ events passing selection to enter the fit to charged final states 7.2% were actually from B^0 meson decay and 1.5% of those entering the fit to neutral final states were from B^+ meson decay. This results in a +5.7% shift in the signal efficiency for the charged final states and a -5.7% shift for the neutral. This effect also contributes to the signal yield from the fit to data, as the $K\eta'\gamma$ PDF shapes were fitted to the MC samples in which cross-feed was included. The efficiencies were thus not calibrated, as the cross-feed effect is present in both the numerator and denominator of the calculation of the branching fractions (See Equation 5.10) and the branching fraction central value is unchanged.

The important factor is the difference in the magnitude of the cross-feeds between data and MC. There is no way to measure the cross-feed in data with any accuracy; instead, a systematic error of -6% is taken for both the charged final states and a $+6\%$ for the neutral.

	Reconstruction					
True Decay	1	2	3	4	5	6
$K^+\eta'(\rho^0\gamma)\gamma$ (1)	98.4%	1.1%	0.5%	0.0%	0.5%	0.0%
$K^0_S\eta'(\rho^0\gamma)\gamma$ (2)	6.3%	93.3%	0.0%	0.2%	0.2%	0.0%
$K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$ (3)	0.3%	0.0%	98.3%	1.1%	0.0%	0.3%
$K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$ (4)	0.0%	0.5%	9.5%	89.7%	0.0%	0.3%
$K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$ (5)	0.0%	0.0%	0.9%	0.0%	98.2%	0.8%
$K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$ (6)	0.0%	0.6%	0.0%	1.8%	17.6%	79.8%

Table 5.4: True versus reconstructed decay modes for $K\eta'\gamma$ MC passing selection criteria. The number (1-6) after the true decay mode indicates that modes' reconstruction column.

5.2.13 $K\eta'\gamma$ PDF Shape

The fixed parameters in the $K\eta'\gamma$ PDF shapes were individually varied by 1σ of their errors from the fits to MC and the data distribution re-fitted with the altered parameters. The parameters varied and the resulting percentage change in the returned signal yield are shown in Table 5.6. All positive shifts in yield were summed in quadrature as were all negative shifts, to give the total systematic errors.

Error Type	$B^+ \to K^+\eta'\gamma$		$B^0 \to K^0_S\eta'\gamma$	
	$+\sigma(\%)$	$-\sigma(\%)$	$+\sigma(\%)$	$-\sigma(\%)$
$K\eta'\gamma$ PDF shape	5.34	−5.33	8.01	−12.29
$K\eta'\gamma$ PDF Calibration	2.81	−2.83	5.33	−6.94
$b \to c$ PDF shape	1.22	−1.38	–	–
$b \to c$ Yield	0.52	−0.43	6.21	−8.40
$b \to u, d, s$ Yield	2.45	−2.49	2.38	−2.33
Bias Study	1.47	−1.47	3.50	−3.50
Total error	6.81	−6.85	12.21	−16.95

Table 5.5: Systematic errors for $B \to K\eta'\gamma$ extracted by varying the fit parameters and those due to vetoes and cross-feed.

5.2.14 $K\eta'\gamma$ PDF Calibration

The corrections to the $K\eta'\gamma$ PDF shown in Section 4.5 were individually varied by 1σ of their errors and the data distribution re-fitted with the altered parameters. The parameters varied and the resulting percentage change in the returned signal yield are shown in Table 5.7. Again, All positive shifts in yield were summed in quadrature as were all negative shifts, to give the total systematic errors.

5.2.15 $b \to c$ PDF Shape

The fixed parameters in the $b \to c$ PDF shapes were individually varied by 1σ of the error on the fits to MC and the data distribution re-fitted with the altered parameters. The parameters varied and the resulting percentage change in the returned signal yield are shown in Table 5.8.

5.2.16 $b \to c$ PDF Yield

The error on the expected number of $b \to c$ events in the data sample was taken to be the Poissonian error. The fixed yields of the $b \to c$ component of the fit to data were varied by 3σ of the error on the yields and the data distribution re-fitted with the altered parameters.

5.2.17 $b \to u, d, s$ PDF Yield

The fixed yields of the $b \to u, d, s$ component of the fit to data were varied by 3σ of the Poissonian error on the yields and the data distribution re-fitted with the altered parameters. In the fit to neutral final states, 3σ is larger than 100%; the yield is varied by 100%.

5.2.18 Bias Study

The biases found on the 2D fitter using the Toy MC pull test and the GSIM/Toy MC ensemble test (Section 4.7) are included as systematic errors. The neutral mode error is from the slope (p_1) of the linear fit to the

GSIM/Toy MC ensemble test results. For the charged modes, the value of p_0 in the linear fit is taken as the systematic.

Error Type	$B^+ \to K^+\eta'\gamma$		$B^0 \to K^0_S\eta'\gamma$	
	$+\sigma(\%)$	$-\sigma(\%)$	$+\sigma(\%)$	$-\sigma(\%)$
$\Delta E_{\sigma rel}$	-0.16	0.18	-0.88	0.83
ΔE_{μ}	-0.54	0.49	-1.99	2.02
ΔE_{σ}	-2.81	2.8	-6.27	3.56
ΔE_{α}	3.20	-3.16	3.94	-3.67
ΔE_{n}	2.30	-2.27	2.70	-2.68
ΔE_{frac}	0.82	-0.87	1.66	-1.78
$M_{\mathrm{BC}\mu}$	-0.41	0.42	-1.71	1.78
$M_{\mathrm{BC}\sigma}$	-0.57	0.58	0.18	-0.23
$M_{\mathrm{BC}\alpha}$	1.22	-1.32	3.07	-6.32
M_{BCn}	1.32	-1.44	2.92	-6.34
Total Error	5.34	-5.33	8.01	-12.29

Table 5.6: $K\eta'\gamma$ PDF parameters varied and the resulting percentage shift in the signal yield.

Error Type	$B^+ \to K^+\eta'\gamma$		$B^0 \to K^0_S\eta'\gamma$	
	$+\sigma(\%)$	$-\sigma(\%)$	$+\sigma(\%)$	$-\sigma(\%)$
$M_{\mathrm{BC}\sigma}$	2.31	-2.26	-0.69	0.36
$M_{\mathrm{BC}\mu}$	-0.98	1.04	-2.86	3.05
ΔE_{μ}	-1.39	1.21	-6.28	4.35
Total Error	2.81	-2.83	-6.94	5.33

Table 5.7: The uncertainties related to the signal PDF calibration.

Error Type	$B^+ \to K^+\eta'\gamma$		$B^0 \to K^0_S\eta'\gamma$	
	$+\sigma(\%)$	$-\sigma(\%)$	$+\sigma(\%)$	$-\sigma(\%)$
ΔE_{C0}	-0.09	0.077	-	-
ΔE_{C1}	-0.70	0.69	-	-
$M_{\rm BCsp}$	-0.58	0.71	-	-
$M_{\rm BCA}$	0.72	-1.03	-	-
Total Error	1.22	-1.37	-	-

Table 5.8: Generic BB PDF parameters varied and the resulting percentage shift in the signal yield.

5.3 Branching fractions, signifi cance and upper limits

The branching fraction of $B \to K\eta'\gamma$ is defined as

$$\mathcal{BF} = \frac{Y_{K\eta'\gamma}}{\varepsilon \times N_{B\overline{B}} \times \prod} \tag{5.10}$$

where $Y_{K\eta'\gamma}$ is the measured yield, ε is the signal efficiency after corrections, $N_{B\overline{B}}$ is the number of $B\overline{B}$ pairs in the data sample, and $\prod$ is the total daughter branching fraction the analysis is sensitive to.

The measured yields, daughter branching fractions, efficiencies, and the calculated branching fractions for the fits to charged and neutral $B \to K\eta'\gamma$ decays are shown in Table 5.9. The statistical and systematic errors on the branching fractions are also shown. The significance of the measurements shown in Section 4.6 includes only the statistical error. To include the systematic errors in the significance the likelihood functions were smeared by convolution with a Gaussian of varying width defined by

$$\sigma = \sigma_1 \oplus \sigma_2 \times (\frac{\mathcal{BF}}{\mathcal{BF}_{\rm cv}}) \tag{5.11}$$

where σ_1 is the total additive systematic error shown in Table 5.5, σ_2 is the multiplicative systematic error shown in Table 5.3, and $\mathcal{BF}_{\rm cv}$ is the $\mathcal{BF}$ central value shown in Table 5.9.

Figure 5.7 shows the normalised likelihood of the two fits as a function of $\mathcal{BF}$. The black curves are the likelihood including statistical errors only. The red curves are the smeared likelihood functions. The significances (σ') in Table 5.9 are calculated from the smeared likelihood function.

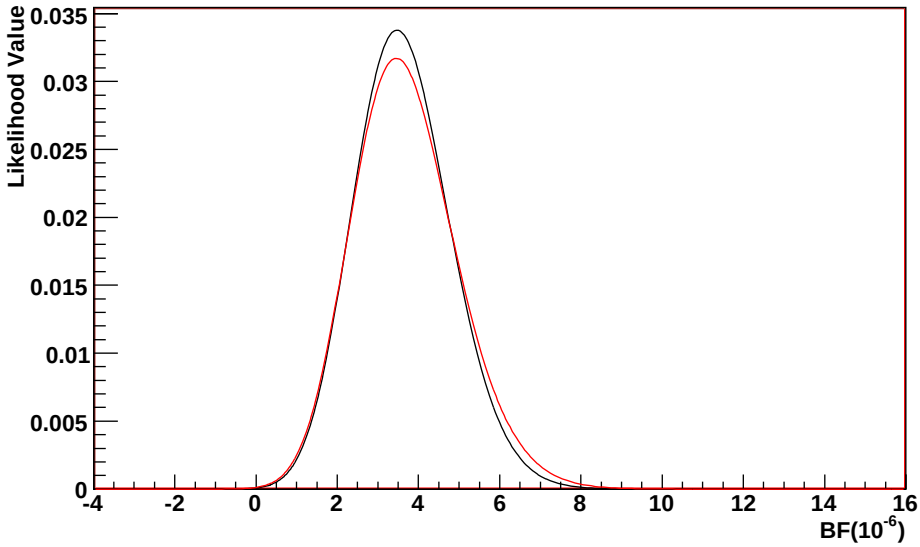
An upper limit ($\mathcal{UL}$) at 90% confidence level is calculated by integrating the smeared likelihood function in the positive $\mathcal{BF}$ region. The $\mathcal{UL}$ is the $\mathcal{BF}$ at which 90% of the total integrated likelihood is included.

The first evidence found for the decay $B^+ \to K^+\eta'\gamma$ is measured at 3.29σ significance. No evidence for the decay $B^0 \to K^0_S\eta'\gamma$ is found, and an upper limit ($\mathcal{UL}$) of 6.0×10^{-6} is calculated at a 90% confidence level.[1]

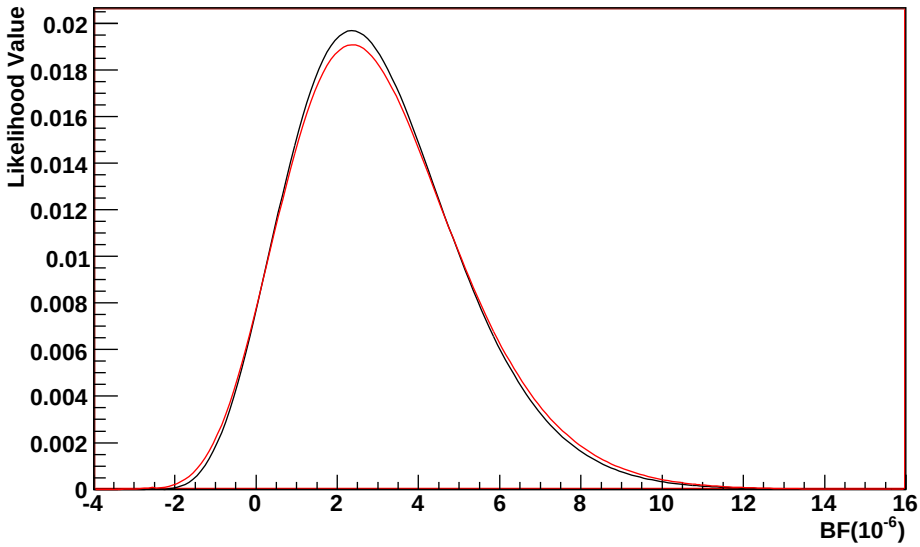
Mode	$Y_{K\eta'\gamma}$ (events)	ε	$\prod$	$\mathcal{BF}$ ($\times 10^{-6}$)	σ'	$\mathcal{UL}$ ($\times 10^{-6}$)
$B^+ \to K^+\eta'\gamma$	$32.61^{+11.75}_{-10.75}$	0.0254	0.571	$3.42^{+1.23+0.34}_{-1.13-0.40}$	3.29	5.3
$B^0 \to K^0\eta'\gamma$	$5.08^{+4.99}_{-3.96}$	0.0171	0.197	$2.30^{+2.26+0.38}_{-1.79-0.44}$	1.31	6.0

Table 5.9: The yields ($Y_{K\eta'\gamma}$), effi ciencies (ε), daughter $\mathcal{BF}$ s ($\prod$), measured $\mathcal{BF}$ s, fi t signifi cances including systematics (S') and $\mathcal{UL}$s for the measured decays. The fi rst error on the $\mathcal{BF}$ is the statistical, the second the systematic.

[1] These results were presented at the International Conference on High Energy Physics (ICHEP) in July of 2008. A conference paper entitled *'Evidence for B to K eta' gamma Decays at Belle'* (arXiv:0810.0804v1 [hep-ex]) is available online at http://arxiv.org/abs/0810.0804



(a) Likelihood function for the fit to charged modes.



(b) Likelihood function for the fit to neutral modes.

Figure 5.7: The likelihood functions of the fits to data as a function of branching fraction. The black curve includes only statistical errors, the red includes both statistical and systematic errors.

Chapter 6

Conclusion

An investigation of B meson decays to the final state $K\eta'\gamma$ has been performed using 605.44 fb^{-1} of data collected at the $\Upsilon(4S)$ resonance by the Belle detector in the years 1999 to 2006.

Backgrounds to the signal decay from $e^+e^- \rightarrow q\overline{q}$ transitions were suppressed using event shape variables that utilise topological differences between the jet-like $q\overline{q}$ events and isotropic $\Upsilon(4S) \rightarrow B\overline{B}$ events. Their discriminating power was optimised using the method of the Fisher Discriminant.

A significant background was found from B decays to final states including a D meson and one or more light unflavoured mesons. To deal with this, charged kaons that gave an invariant mass within 25 MeV/c^2 of the nominal D mass when combined with any other charged pion in the event were removed from consideration. Background events from $B \rightarrow J/\psi K \rightarrow (\eta'\gamma)K$ decays were also suppressed using a veto on reconstructions in which $M_{\eta'\gamma}$ is within 25 MeV/c^2 of the nominal J/ψ mass.

Detailed studies of several multi-dimensional fitting methods were conducted due to the need for high accuracy and confidence in the stability of the method chosen to analyse the data. A 2D method fitting the distributions of the variables M_{BC} and ΔE in unbinned maximum likelihood fits proved to be the most reliable and was used to extract the $K\eta'\gamma$ yield from the selected data events. No significant inherent biases were found in the fitting process using two methods of statistical analysis. An in-depth systematic error analysis was conducted to estimate the uncertainties involved in the detection, reconstuction and measurement of the decay.

No significant signal was found for the decay $B^0 \rightarrow K^0\eta'\gamma$. The measured branching fraction was

$$\mathcal{BF}(B^0 \rightarrow K^0\eta'\gamma) = (2.30^{+2.26+0.38}_{-1.79-0.44}) \times 10^{-6}$$

with a signal significance of 1.31σ. An upper limit at the 90% confidence level for this decay was calculated to be

$$\mathcal{BF}(B^0 \rightarrow K^0\eta'\gamma) < 6.0 \times 10^{-6}$$

This is an improvement of 10% on the previous upper limit.

The branching fraction of the decay $B^+ \to K^+\eta'\gamma$ was measured to be

$$\mathcal{BF}(B^+ \to K^+\eta'\gamma) = (3.42^{+1.23+0.34}_{-1.13-0.40}) \times 10^{-6}$$

where the first error is statistical and the second is systematic. The signal significance was 3.29 standard deviations away from the null hypothesis. This is the first evidence discovered of this flavour-changing neutral current decay. This discovery increases our knowledge of the hadronisation process in the $B \to X_s\gamma$ interaction. The measurement also provides information with which to test the current knowledge of theoretical particle physics. A comparison with a theoretical prediction of $\mathcal{BF}(B^+ \to K^+\eta'\gamma)$ will test QCD factorisation through the ratio $\mathcal{BF}(B^+ \to K^+\eta'\gamma)$:$\mathcal{BF}()$:$\mathcal{BF}(B \to K\pi\gamma)$. Such a comparison will also test theories of beyond SM physics through the effect such theories have on the $b \to s\gamma$ loop interaction.

Appendix A

$\mathcal{LR}$ Cuts

Final State	Flavour Tagging Region	Max. $\mathcal{FOM}$	LR cut
$K^+\eta'(\rho^0\gamma)\gamma$	$-0.100 \leq (q.r.B_{Flav}) < -0.875$	0.844188	0.94
	$-0.875 \leq (q.r.B_{Flav}) < -0.750$	0.365850	0.94
	$-0.750 \leq (q.r.B_{Flav}) < -0.625$	0.255866	0.86
	$-0.625 \leq (q.r.B_{Flav}) < -0.500$	0.237540	0.84
	$-0.500 \leq (q.r.B_{Flav}) < -0.250$	0.264840	0.96
	$-0.250 \leq (q.r.B_{Flav}) < 1.000$	0.449271	0.82
$K^0_S\eta'(\rho^0\gamma)\gamma$	$0.000 \leq \|q.r\| < 0.250$	0.134553	0.86
	$0.250 \leq \|q.r\| < 0.500$	0.108277	0.98
	$0.500 \leq \|q.r\| < 0.625$	0.080338	0.96
	$0.625 \leq \|q.r\| < 0.750$	0.080367	0.96
	$0.750 \leq \|q.r\| < 0.875$	0.097710	0.88
	$0.875 \leq \|q.r\| < 1.000$	0.352057	0.96
$K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	$-0.100 \leq (q.r.B_{Flav}) < -0.875$	1.692620	0.84
	$-0.875 \leq (q.r.B_{Flav}) < -0.750$	0.644906	0.92
	$-0.750 \leq (q.r.B_{Flav}) < -0.625$	0.648410	0.88
	$-0.625 \leq (q.r.B_{Flav}) < -0.500$	0.503548	0.86
	$-0.500 \leq (q.r.B_{Flav}) < -0.250$	0.503516	0.82
	$-0.250 \leq (q.r.B_{Flav}) < 1.000$	0.903294	0.88
$K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	$0.000 \leq \|q.r\| < 0.250$	0.225114	0.88
	$0.250 \leq \|q.r\| < 0.500$	0.226669	0.96
	$0.500 \leq \|q.r\| < 0.625$	0.192886	0.92
	$0.625 \leq \|q.r\| < 0.750$	0.176225	0.90
	$0.750 \leq \|q.r\| < 0.875$	0.268177	0.96
	$0.875 \leq \|q.r\| < 1.000$	0.473678	0.94
$K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	No Flavour Tagging	0.929328	0.80
$K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	No Flavour Tagging	0.257173	0.86

Table A.1: The optimised $\mathcal{FOM}$ value and corresponding $\mathcal{LR}$ cut for each flavour tagging bin and final state in the 3D mode-by-mode method.

Final State	Flavour Tagging Region	Max. $\mathcal{FOM}$	LR cut
$K^+\eta'(\rho^0\gamma)\gamma$	$-0.1 \leq (q.r.B_{Flav}) < -0.875$	0.715514	0.96
	$-0.875 \leq (q.r.B_{Flav}) < -0.75$	0.299452	0.94
	$-0.75 \leq (q.r.B_{Flav}) < -0.625$	0.207126	0.88
	$-0.625 \leq (q.r.B_{Flav}) < -0.5$	0.192197	0.88
	$-0.5 \leq (q.r.B_{Flav}) < -0.25$	0.216263	0.96
	$-0.25 \leq (q.r.B_{Flav}) < 1.0$	0.362846	0.88
$K^0_S\eta'(\rho^0\gamma)\gamma$	$0.0 \leq \|q.r\| < 0.25$	0.108662	0.90
	$0.25 \leq \|q.r\| < 0.50$	0.0894537	0.98
	$0.50 \leq \|q.r\| < 0.625$	0.0653321	0.92
	$0.625 \leq \|q.r\| < 0.75$	0.0652263	0.96
	$0.75 \leq \|q.r\| < 0.875$	0.0790728	0.90
	$0.875 \leq \|q.r\| < 1.0$	0.304741	0.94
$K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	$-0.1 \leq (q.r.B_{Flav}) < -0.875$	1.57928	0.90
	$-0.875 \leq (q.r.B_{Flav}) < -0.75$	0.568788	0.94
	$-0.75 \leq (q.r.B_{Flav}) < -0.625$	0.520242	0.92
	$-0.625 \leq (q.r.B_{Flav}) < -0.5$	0.386136	0.88
	$-0.5 \leq (q.r.B_{Flav}) < -0.25$	0.383142	0.88
	$-0.25 \leq (q.r.B_{Flav}) < 1.0$	0.697317	0.82
$K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	$0.0 \leq \|q.r\| < 0.25$	0.178944	0.84
	$0.25 \leq \|q.r\| < 0.50$	0.176379	0.96
	$0.50 \leq \|q.r\| < 0.625$	0.116759	0.88
	$0.625 \leq \|q.r\| < 0.75$	0.112932	0.88
	$0.75 \leq \|q.r\| < 0.875$	0.148814	0.94
	$0.875 \leq \|q.r\| < 1.0$	0.437932	0.94
$K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	No Flavour Tagging	0.790696	0.78
$K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	No Flavour Tagging	0.227526	0.88

Table A.2: The optimised $\mathcal{FOM}$ value and corresponding $\mathcal{LR}$ cut for each flavour tagging bin and final state in the 3D combined method.

Final State	Flavour Tagging Region	Max. $\mathcal{FOM}$	LR cut
$K^+\eta'(\rho^0\gamma)\gamma$	$-0.100 \leq (q.r.B_{Flav}) < -0.875$	1.30173	0.92
	$-0.875 \leq (q.r.B_{Flav}) < -0.750$	0.657573	0.98
	$-0.750 \leq (q.r.B_{Flav}) < -0.625$	0.533148	0.92
	$-0.625 \leq (q.r.B_{Flav}) < -0.500$	0.433832	0.86
	$-0.500 \leq (q.r.B_{Flav}) < -0.250$	0.531591	0.92
	$-0.250 \leq (q.r.B_{Flav}) < 1.000$	0.964162	0.92
$K^0_S\eta'(\rho^0\gamma)\gamma$	$0.000 \leq \|q.r\| < 0.250$	0.319025	0.78
	$0.250 \leq \|q.r\| < 0.500$	0.300454	0.98
	$0.500 \leq \|q.r\| < 0.625$	0.226999	0.96
	$0.625 \leq \|q.r\| < 0.750$	0.183196	0.98
	$0.75 \leq \|q.r\| < 0.875$	0.342161	0.84
	$0.875 \leq \|q.r\| < 1.000$	0.472024	0.96
$K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	$-0.100 \leq (q.r.B_{Flav}) < -0.875$	1.89646	0.92
	$-0.875 \leq (q.r.B_{Flav}) < -0.750$	0.918172	0.94
	$-0.750 \leq (q.r.B_{Flav}) < -0.625$	0.997831	0.92
	$-0.625 \leq (q.r.B_{Flav}) < -0.500$	0.958511	0.90
	$-0.500 \leq (q.r.B_{Flav}) < -0.250$	1.09227	0.84
	$-0.250 \leq (q.r.B_{Flav}) < 1.000$	1.82653	0.84
$K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$	$0.000 \leq \|q.r\| < 0.250$	0.700408	0.88
	$0.250 \leq \|q.r\| < 0.500$	0.509081	0.94
	$0.500 \leq \|q.r\| < 0.625$	0.301839	0.88
	$0.625 \leq \|q.r\| < 0.750$	0.423725	0.96
	$0.750 \leq \|q.r\| < 0.875$	0.397003	0.94
	$0.875 \leq \|q.r\| < 1.000$	0.49606	0.94
$K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	No Flavour Tagging	1.95637	0.76
$K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$	No Flavour Tagging	0.494206	0.88

Table A.3: The optimised $\mathcal{FOM}$ value and corresponding $\mathcal{LR}$ cut for each flavour tagging bin and final state in the 2D method.

Appendix B

Background Correlation Plots

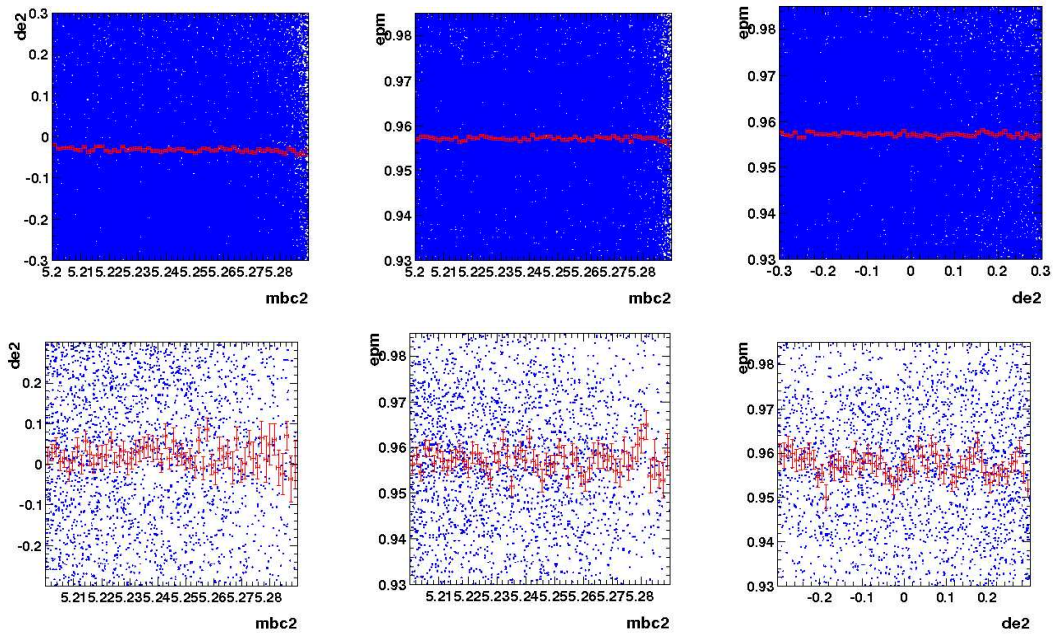


Figure B.1: Scatter plots of the fitting variables for $q\overline{q}$ MC. The red points and error bars are the means of the y-axis variable in a range of the x-axis variable- 0.46 MeV bins for $M_{\rm BC}$, and 10 MeV bins for ΔE. From left to right the plots are: $M_{\rm BC}$:ΔE, $M_{\rm BC}$:$M_{\eta'}$, and ΔE:$M_{\eta'}$. The top row are MC with only reconstruction cuts, the bottom are after $q\overline{q}$ suppression and best candidate selection.

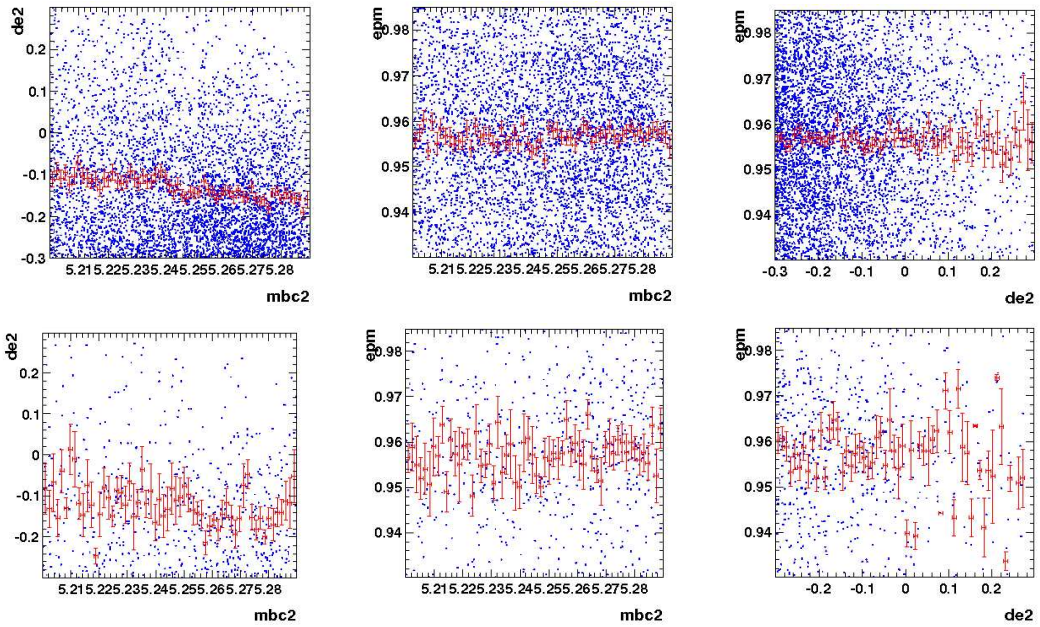


Figure B.2: Scatter plots of the fitting variables for $b \to c$ MC. The red points and error bars are the means of the y-axis variable in a range of the x-axis variable- 0.46 MeV bins for $M_{\rm BC}$, and 10 MeV bins for ΔE. From left to right the plots are: $M_{\rm BC}$:ΔE, $M_{\rm BC}$:$M_{\eta'}$, and ΔE:$M_{\eta'}$. The top row are MC with only reconstruction cuts, the bottom are after $q\overline{q}$ suppression and best candidate selection.

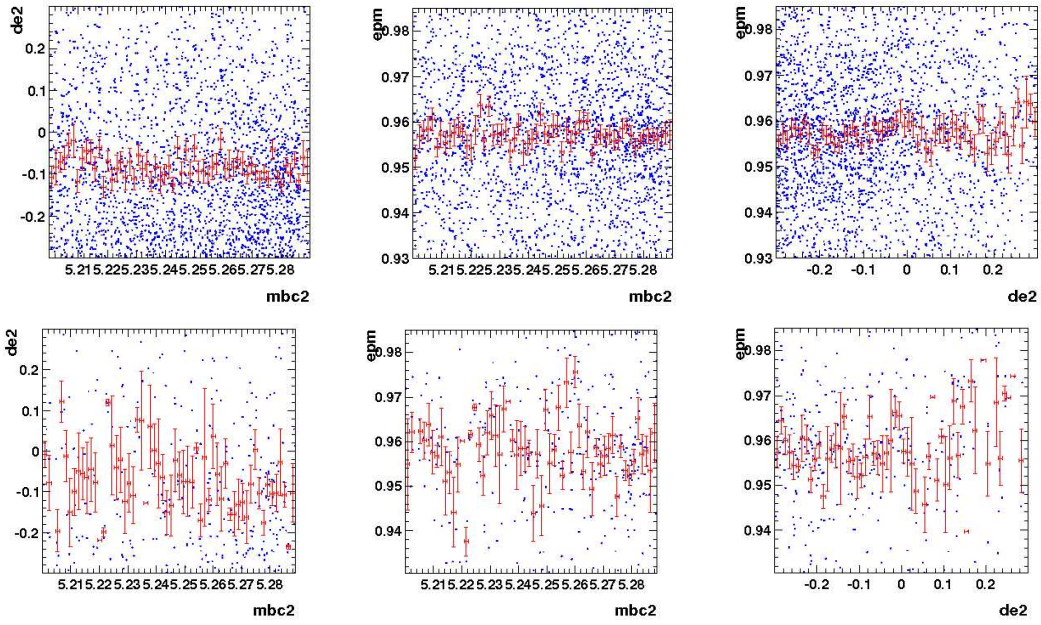


Figure B.3: Scatter plots of the fitting variables for $b \to u, d, s$ MC. The red points and error bars are the means of the y-axis variable in a range of the x-axis variable- 0.46 MeV bins for $M_{\rm BC}$, and 10 MeV bins for ΔE. From left to right the plots are: $M_{\rm BC}$:ΔE, $M_{\rm BC}$:$M_{\eta'}$, and ΔE:$M_{\eta'}$. The top row are MC with only reconstruction cuts, the bottom are after $q\overline{q}$ suppression and best candidate selection.

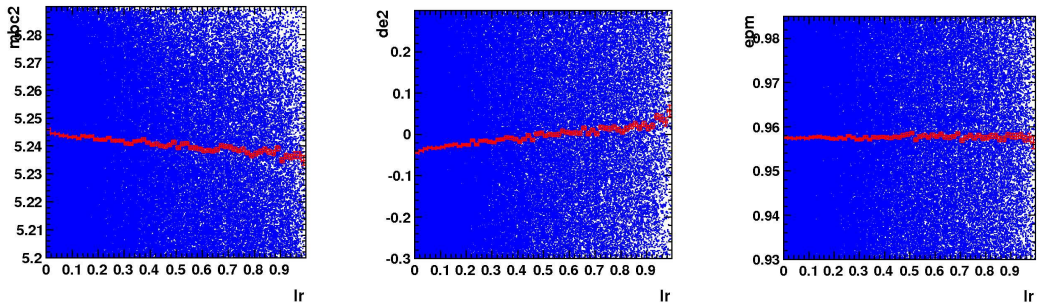


Figure B.4: Scatter plots of the fitting variables $M_{\rm BC}$ and ΔE versus $\mathcal{LR}$ in $q\overline{q}$ MC. The red points and error bars are the means of the y-axis variable in a range of the x-axis variable; 100 bins of $\mathcal{LR} = 0.01$. From left to right the plots are: $M_{\rm BC}$:$\mathcal{LR}$, ΔE:$\mathcal{LR}$ and $M_{\eta'}$:$\mathcal{LR}$.

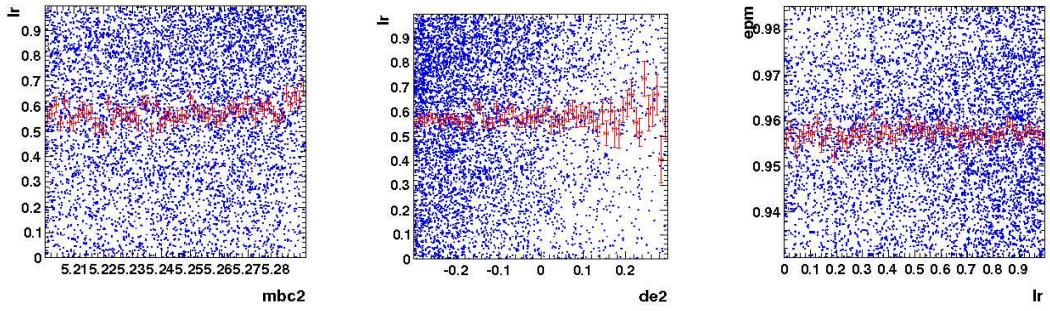


Figure B.5: Scatter plots of the fitting variables $M_{\rm BC}$ and ΔE versus $\mathcal{LR}$ in $b \to c$ MC. The red points and error bars are the means of the y-axis variable in a range of the x-axis variable; 100 bins of $\mathcal{LR} = 0.01$. From left to right the plots are: $M_{\rm BC}$:$\mathcal{LR}$, ΔE:$\mathcal{LR}$ and $M_{\eta'}$:$\mathcal{LR}$.

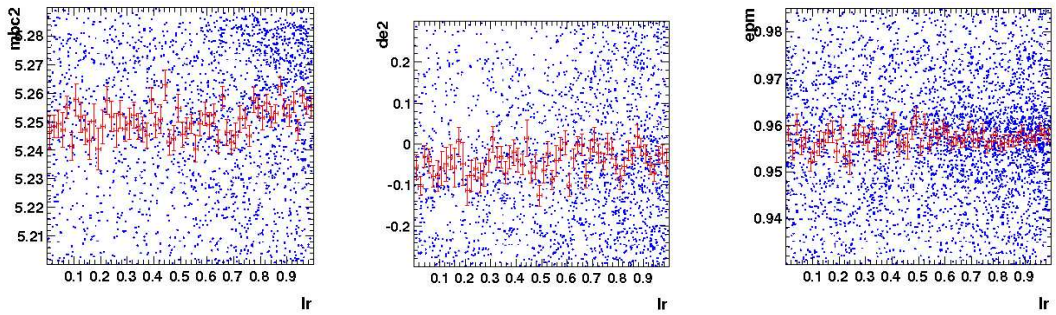


Figure B.6: Scatter plots of the fitting variables $M_{\rm BC}$ and ΔE versus $\mathcal{LR}$ in $b \to u, d, s$ MC. The red points and error bars are the means of the y-axis variable in a range of the x-axis variable; 100 bins of $\mathcal{LR} = 0.01$. From left to right the plots are: $M_{\rm BC}$:$\mathcal{LR}$, ΔE:$\mathcal{LR}$ and $M_{\eta'}$:$\mathcal{LR}$.

Appendix C

3D mode-by-mode MC fi ts

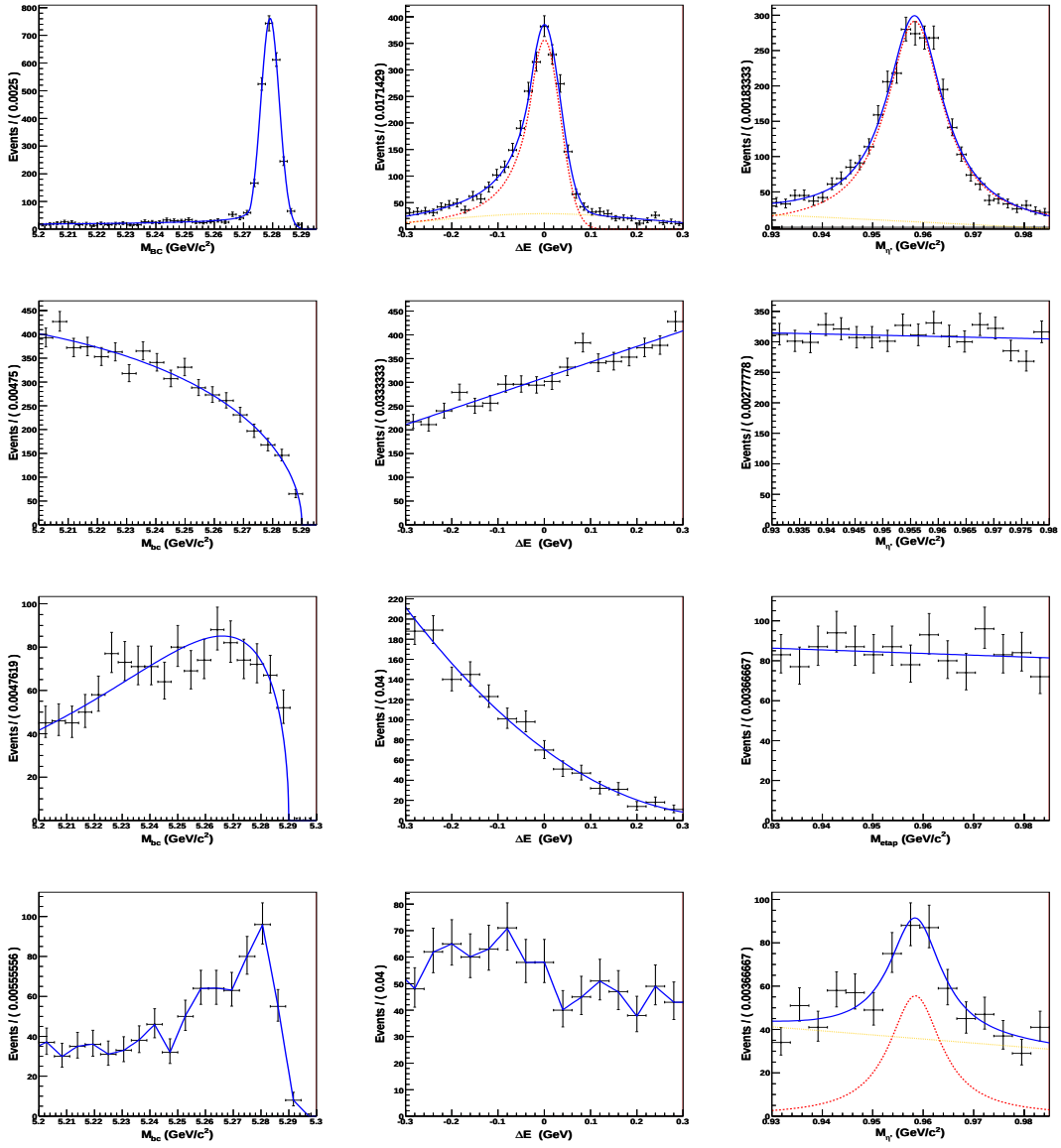


Figure C.1: 3D fits to MC for final state $B^+ \to K^+\eta'(\rho^0\gamma)\gamma$. Left to right, the columns show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. Top to bottom, the rows show $K\eta'\gamma$ MC, $q\overline{q}$ MC, $b \to c$ MC and $b \to u, d, s$ MC.

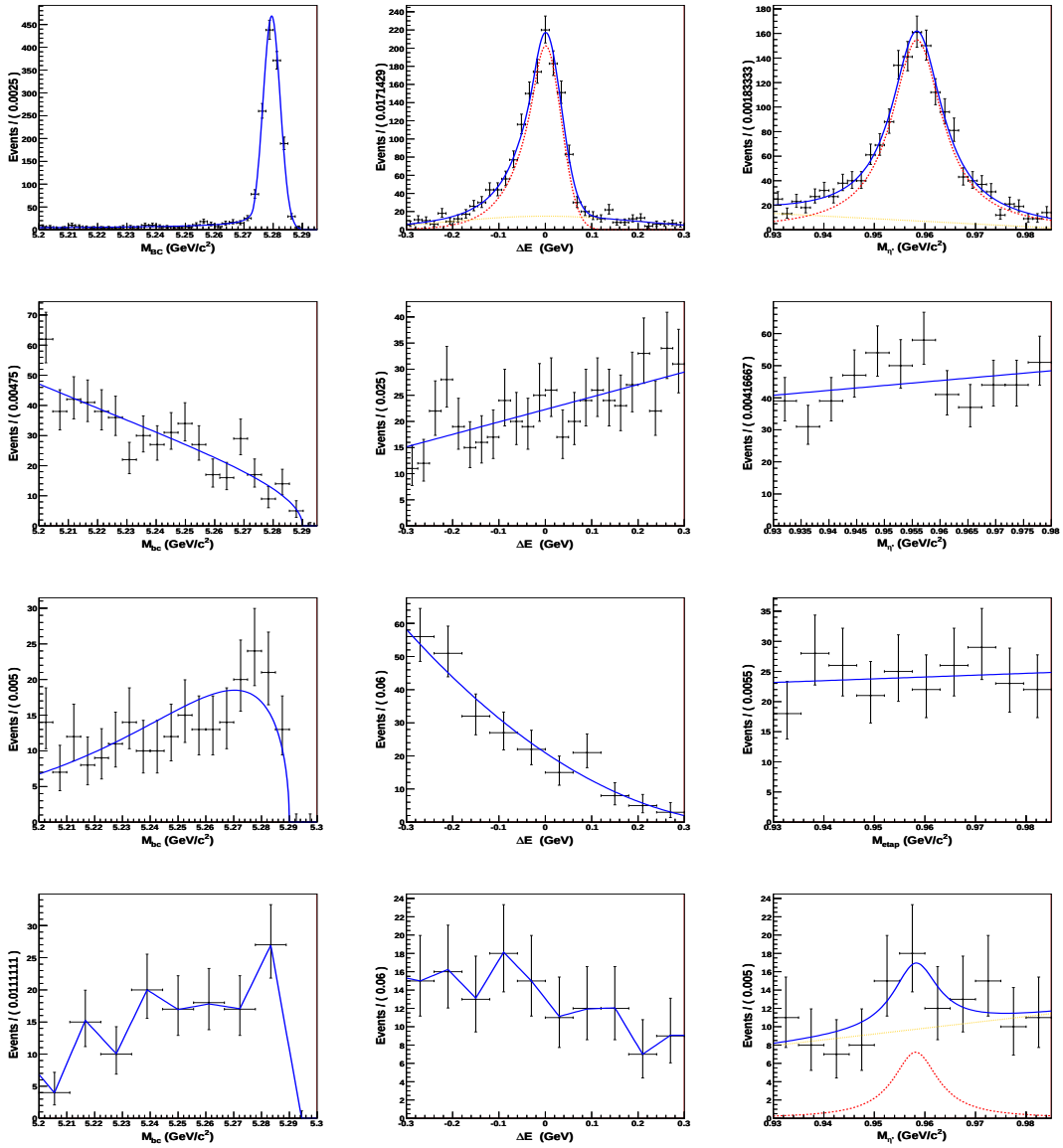


Figure C.2: 3D fits to MC for final state $B^0 \to K^0_S\eta'(\rho^0\gamma)\gamma$. Left to right, the columns show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. Top to bottom, the rows show $K\eta'\gamma$ MC, $q\overline{q}$ MC, $b \to c$ MC and $b \to u, d, s$ MC.

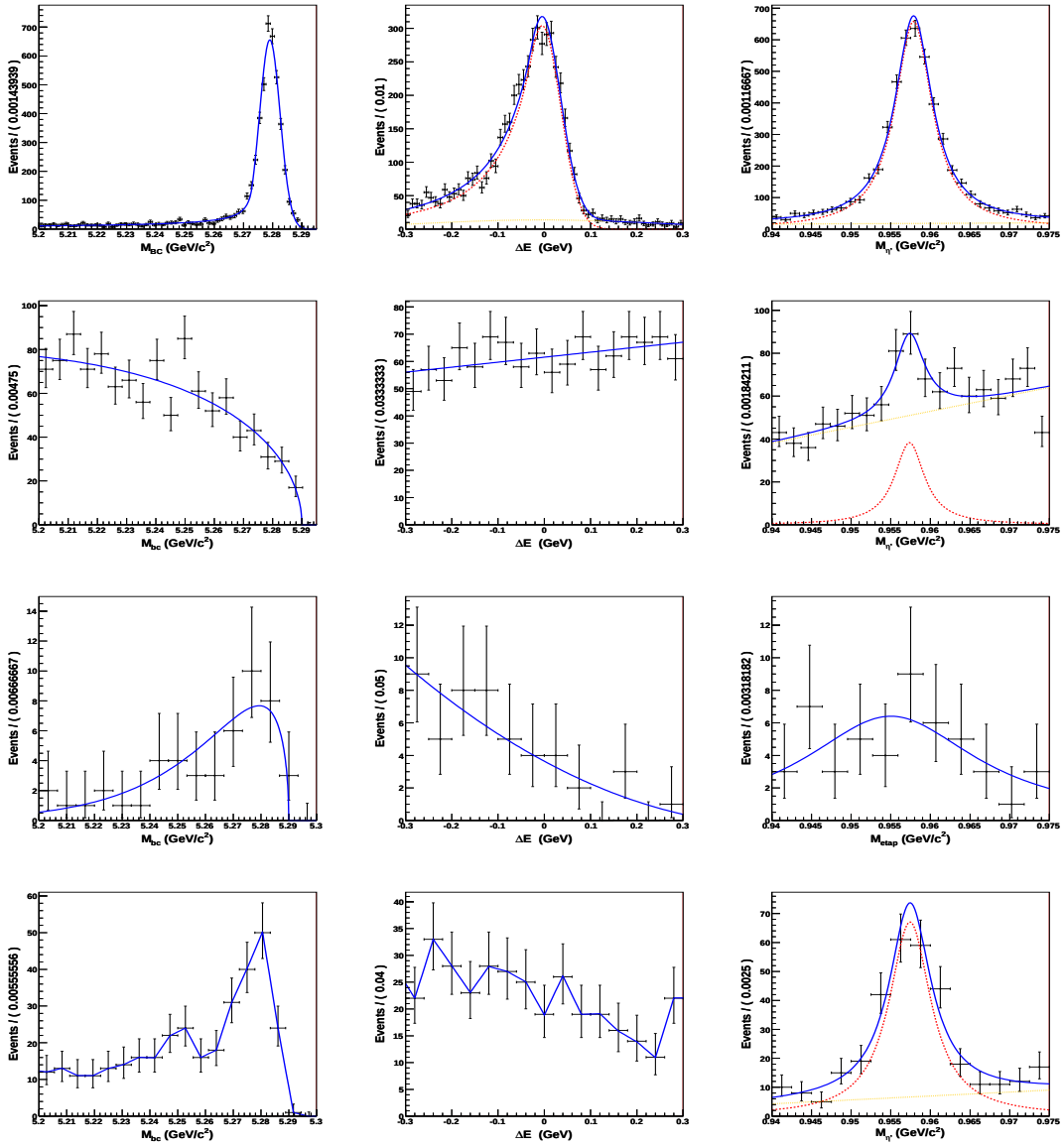


Figure C.3: 3D fits to MC for final state $B^+ \to K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$. Left to right, the columns show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. Top to bottom, the rows show $K\eta'\gamma$ MC, $q\overline{q}$ MC, $b \to c$ MC and $b \to u, d, s$ MC.

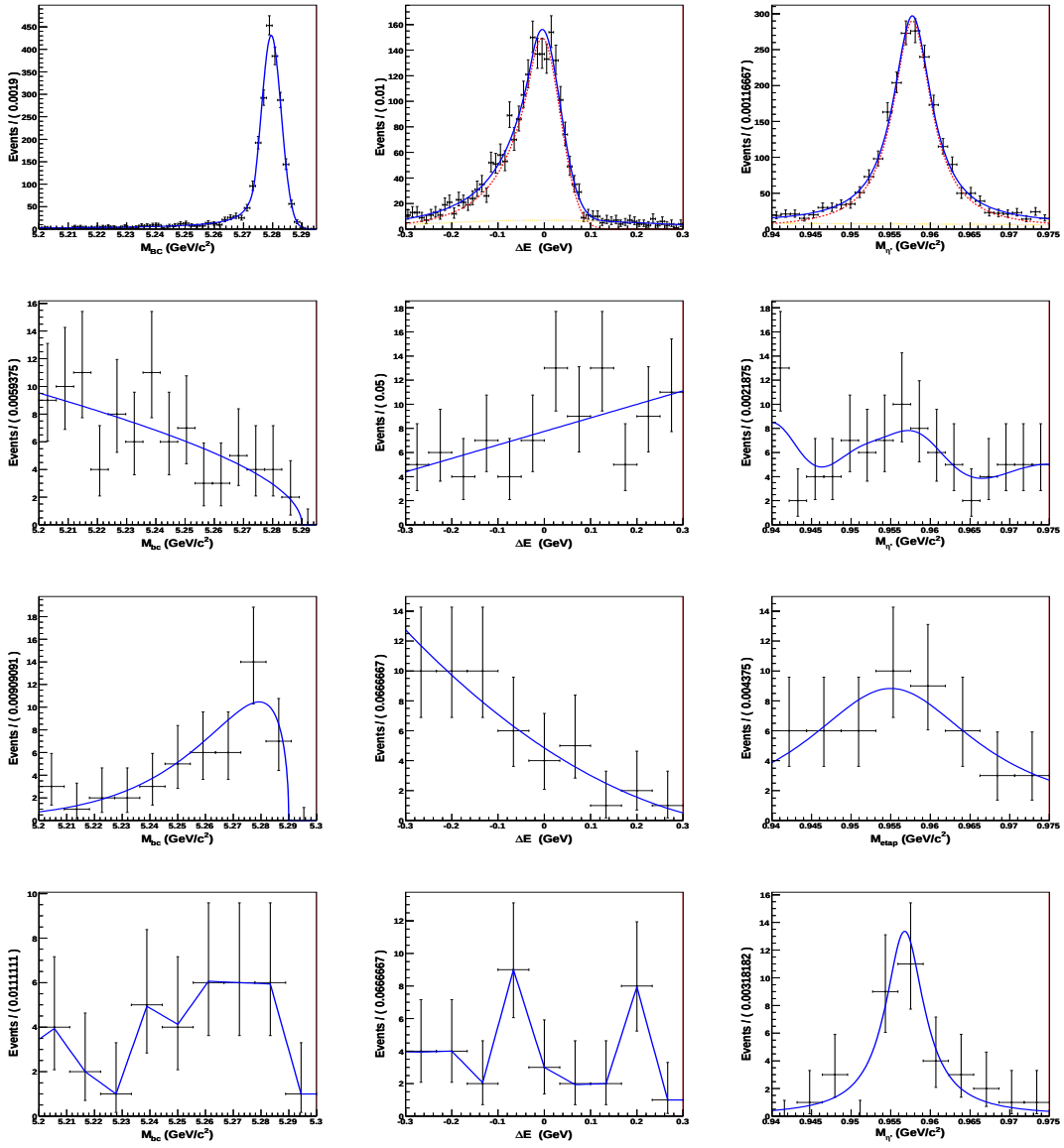


Figure C.4: 3D fits to MC for final state $B^0 \to K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$. Left to right, the columns show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. Top to bottom, the rows show signal MC, $q\overline{q}$ MC, $b \to c$ MC and $b \to u, d, s$ MC.

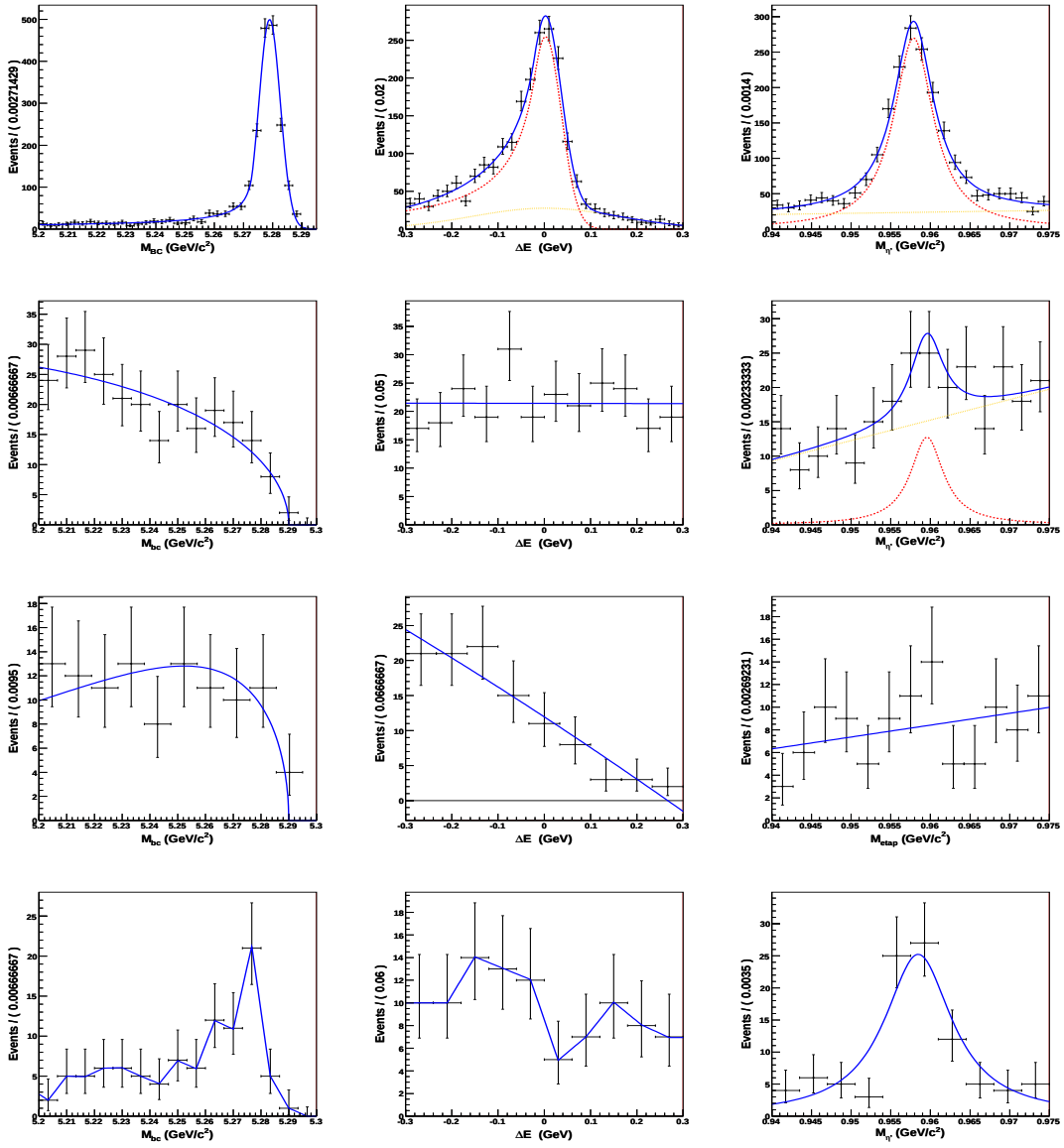


Figure C.5: 3D fits to MC for final state $B^+ \to K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$. Left to right, the columns show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. Top to bottom, the rows show signal MC, $q\overline{q}$ MC, $b \to c$ MC and $b \to u, d, s$ MC.

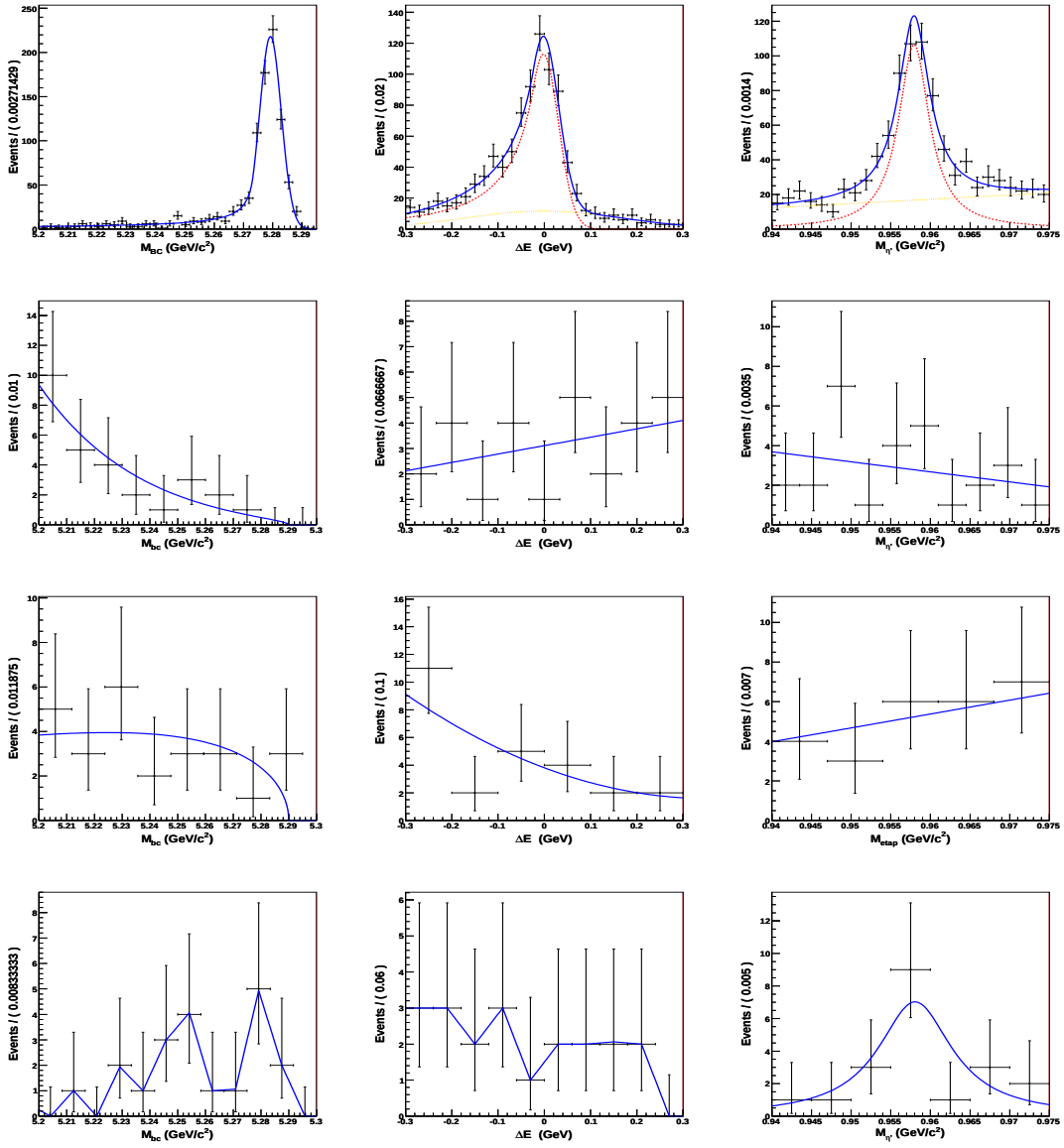


Figure C.6: 3D fits to MC for final state $B^0 \to K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$. Left to right, the columns show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. Top to bottom, the rows show signal MC, $q\overline{q}$ MC, $b \to c$ MC and $b \to u, d, s$ MC.

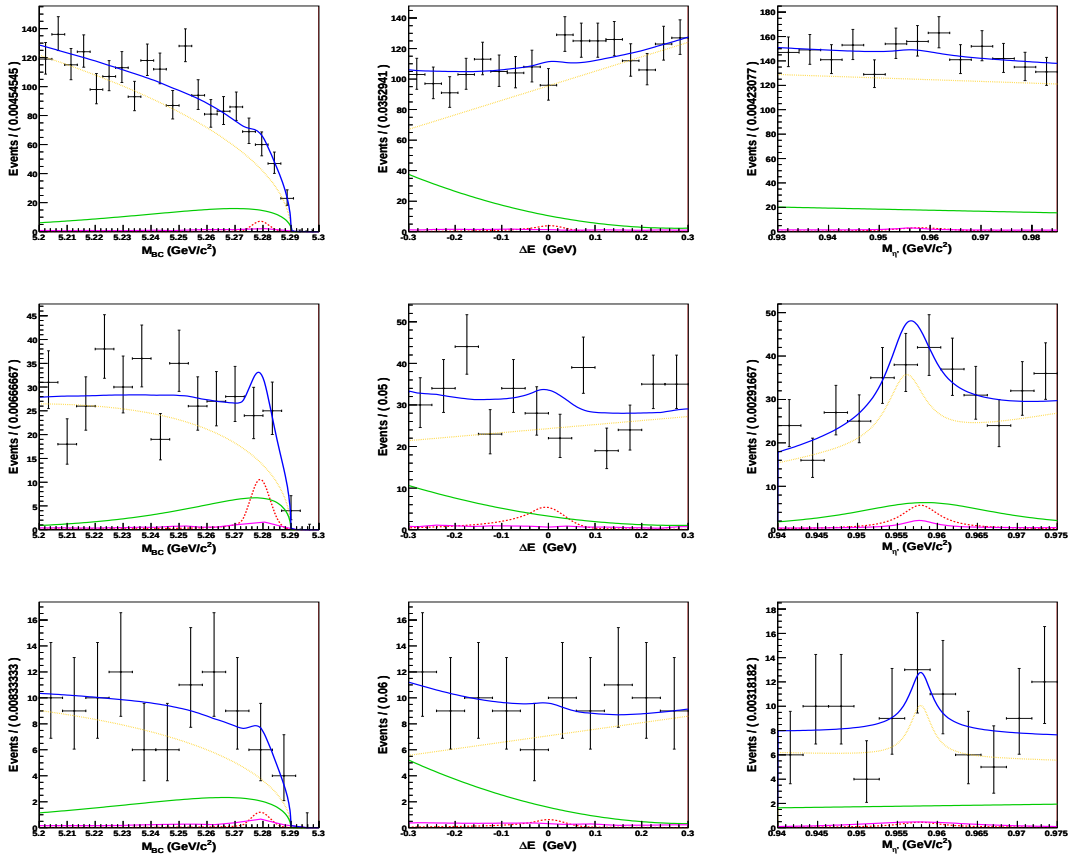


Figure C.7: Simultaneous 3D fit to MC for charged final states. Left to right, the columns show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. The signal function is shown in red, $q\overline{q}$ in orange, $b \to c$ in green, $b \to u, d, s$ in magenta, and the combined function in blue. The top row is final state $K^+\eta'(\rho^0\gamma)\gamma$, the middle $K^+\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$ and the bottom $K^+\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$.

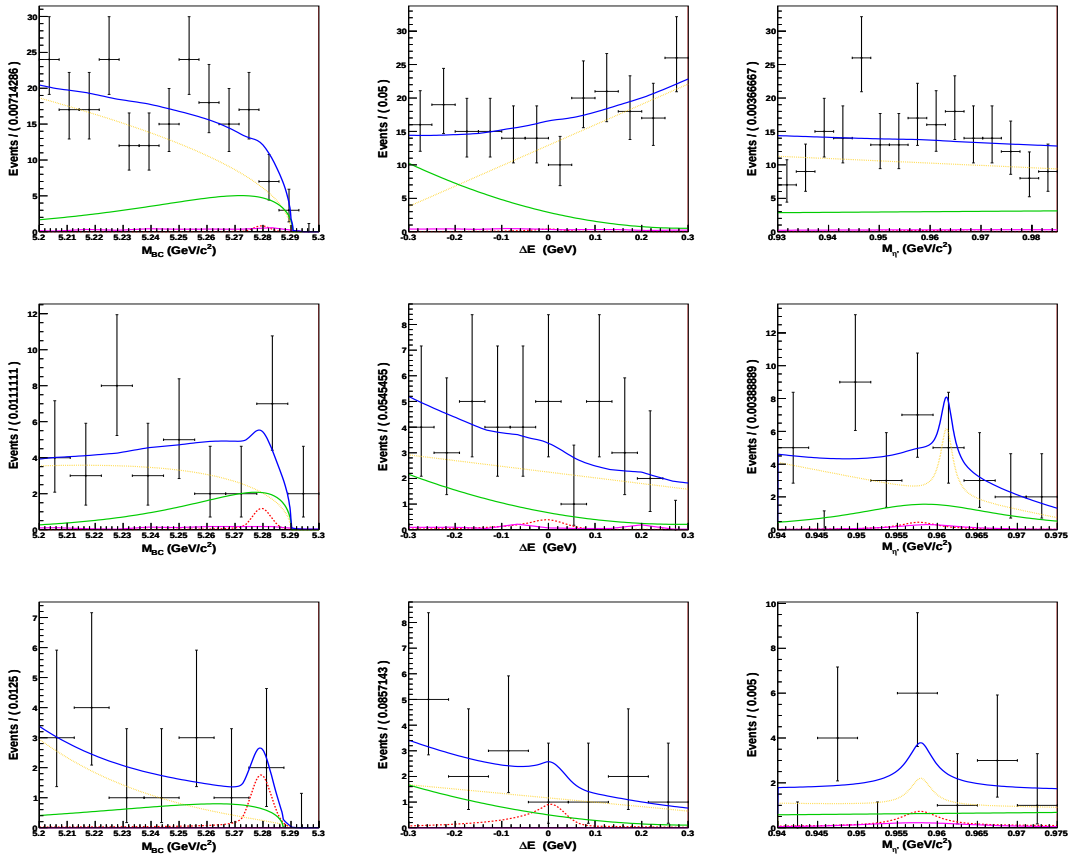


Figure C.8: Simultaneous 3D fit to MC for neutral final states. Left to right, the columns show fits to $M_{\rm BC}$, ΔE and $M_{\eta'}$. The signal function is shown in red, $q\overline{q}$ in orange, $b \to c$ in green, $b \to u, d, s$ in magenta, and the combined function in blue. The top row is final state $K^0_S\eta'(\rho^0\gamma)\gamma$, the middle $K^0_S\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$ and the bottom $K^0_S\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$.

Appendix D

Combined 3D MC fi ts

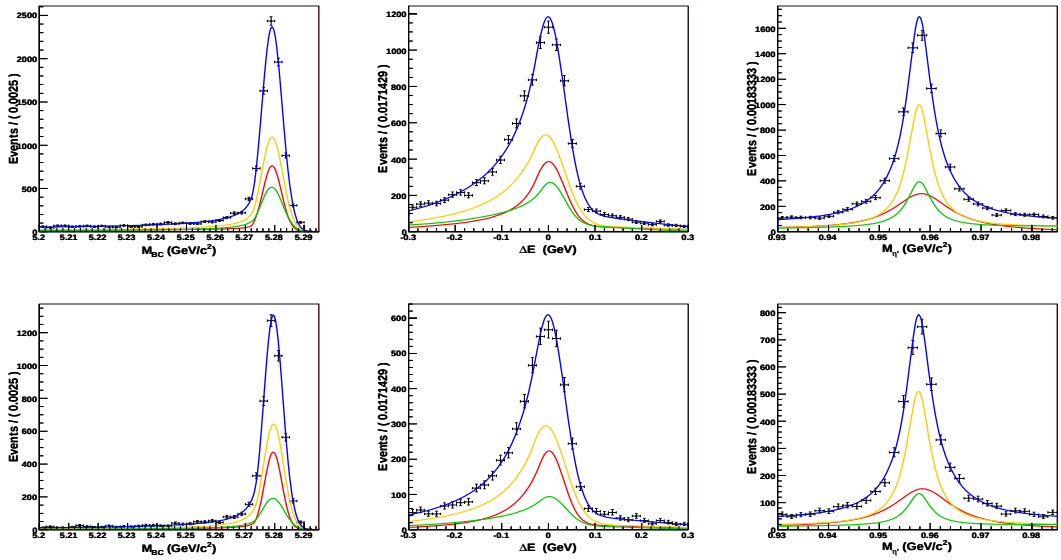


Figure D.1: Combined scaled 3D fi ts to signal MC. Left to right, the columns show fi ts to M_{BC}, ΔE and $M_{\eta'}$. $K\eta'(\rho^0\gamma)\gamma$ is in red, $K\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$ is in orange, and $K\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$ is in green. The top row is the charged modes, the bottom row the neutral.

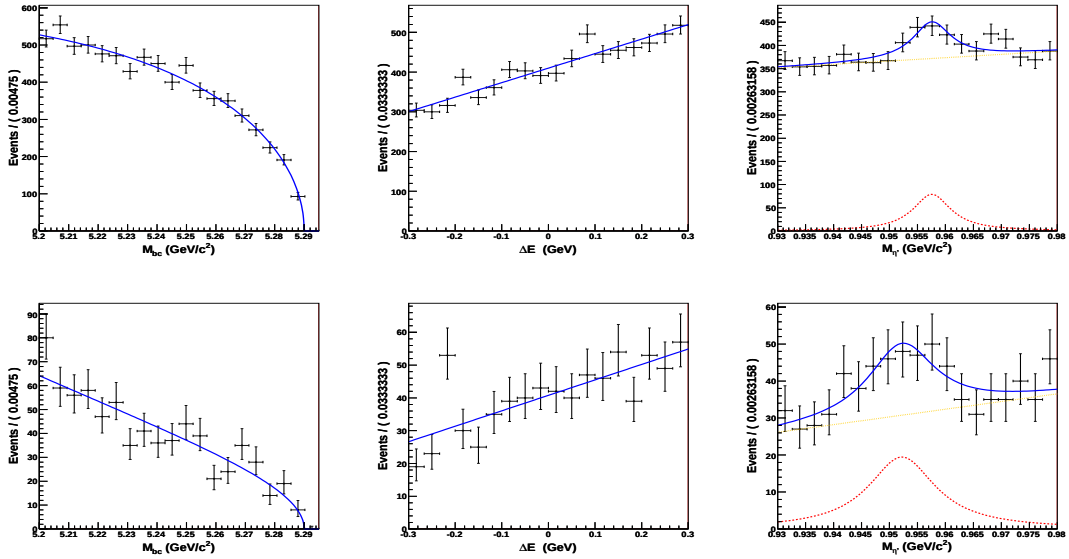


Figure D.2: Combined 3D fits to $q\bar{q}$ MC. Left to right, the columns show fits to M_{BC}, ΔE and $M_{\eta'}$. The top row is the charged modes, the bottom row the neutral.

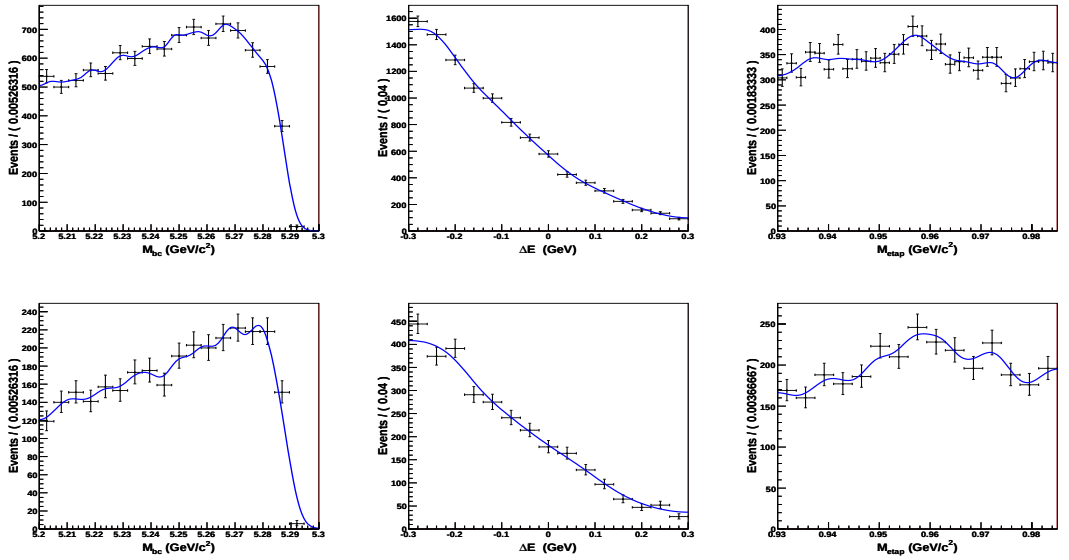


Figure D.3: Combined 3D fits to $b \to c$ MC. Left to right, the columns show fits to M_{BC}, ΔE and $M_{\eta'}$. The top row is the charged mode, the bottom row the neutral.

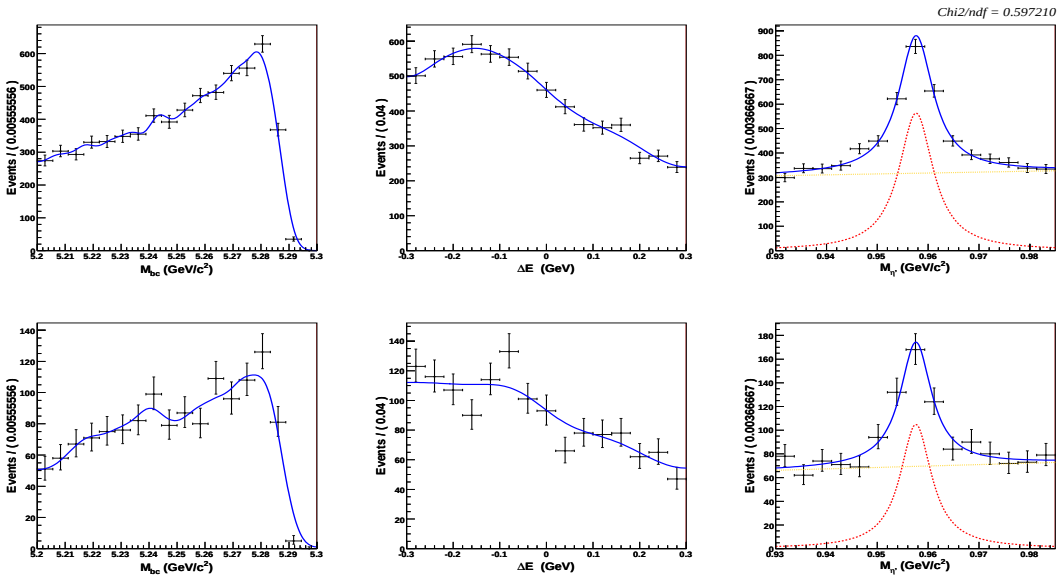


Figure D.4: Combined 3D fits to $b \to u, d, s$ MC. Left to right, the columns show fits to M_{BC}, ΔE and $M_{\eta'}$. The top row is the charged modes, the bottom row the neutral.

Appendix E

2D MC fits

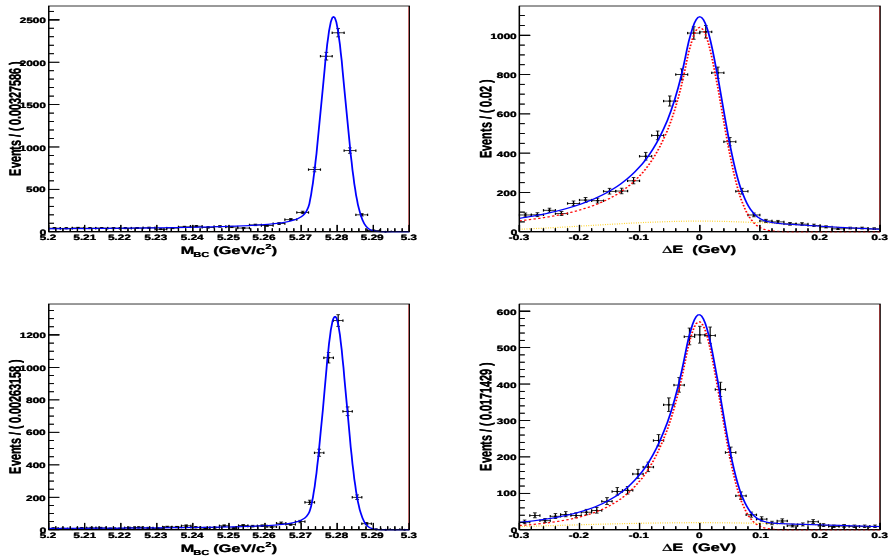


Figure E.1: 2D fits to signal MC. The left plots show M_{BC}, the right show ΔE. The top row are the charged final states, the bottom row the neutral.

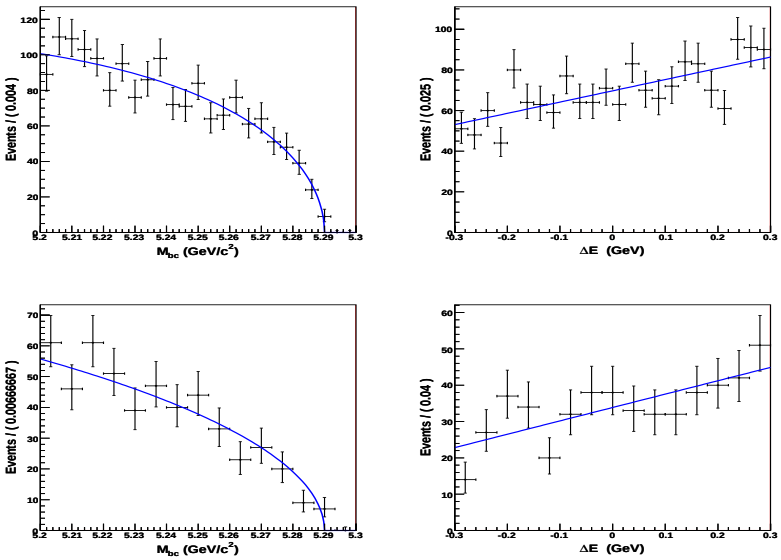


Figure E.2: 2D fits to $q\overline{q}$ MC. The left plots show M_{BC}, the right show ΔE. The top row are the charged final states, the bottom row the neutral.

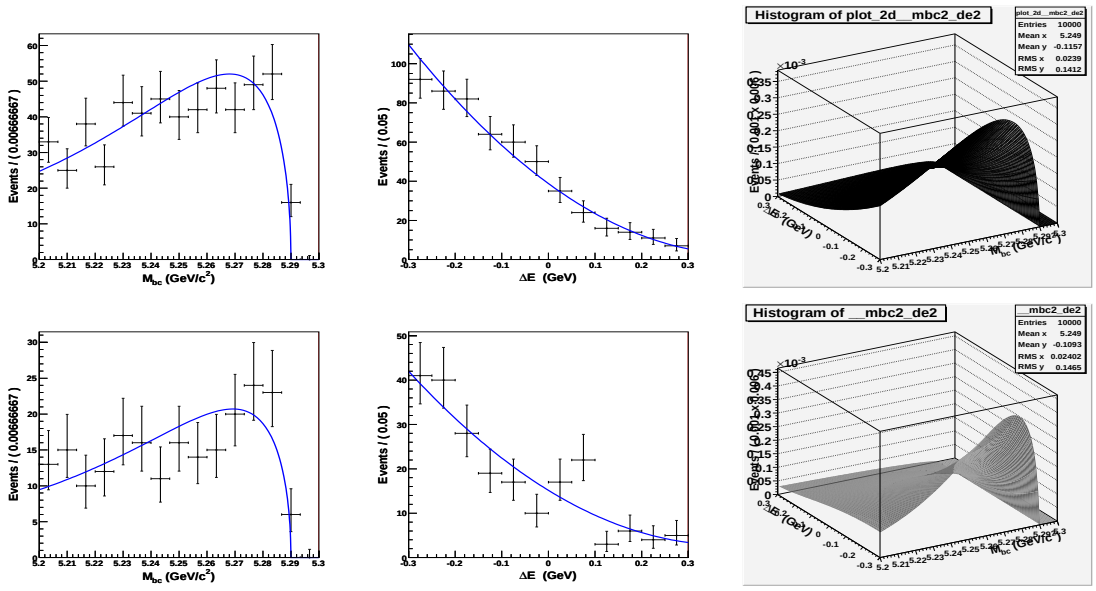


Figure E.3: 2D fits to $b \to c$ MC. The left plots show M_{BC}, the middle show ΔE and the right show the 2D function. The top row are the charged final states, the bottom row the neutral.

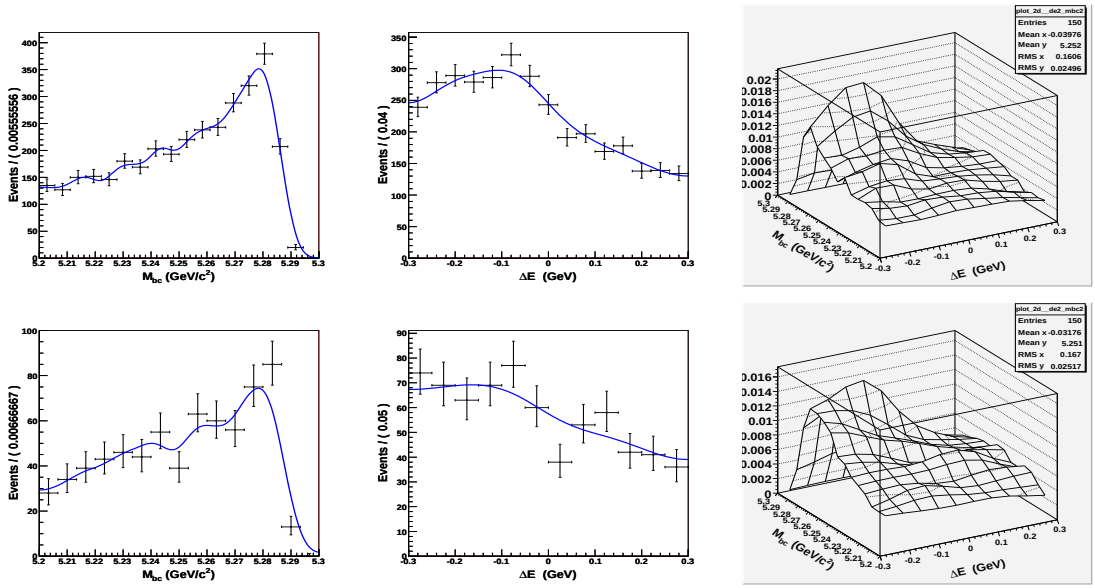


Figure E.4: 2D fits to $b \to u, d, s$ MC. The left plots show $M_{\rm BC}$, the middle show ΔE and the right show the 2D function. The top row are the charged final states, the bottom row the neutral.

Appendix F

Toy MC Results

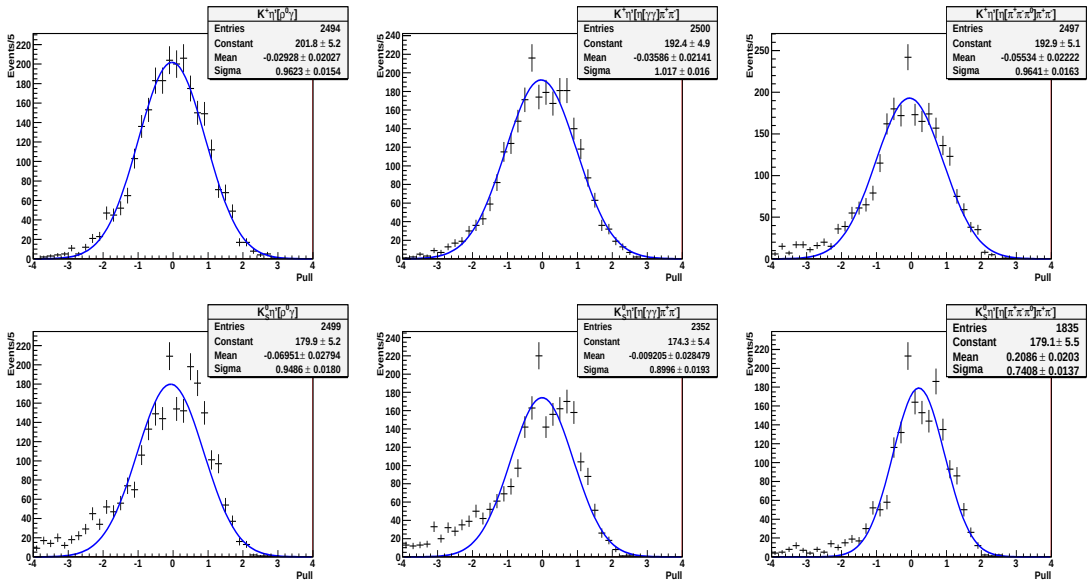


Figure F.1: The pull distributions and fits for the 3D mode-by-mode method fitter. The top row shows the charged final states, the bottom the neutral final states. From left to right, $K\eta(\rho^0\gamma)\gamma$, $K\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$ and $K\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$.

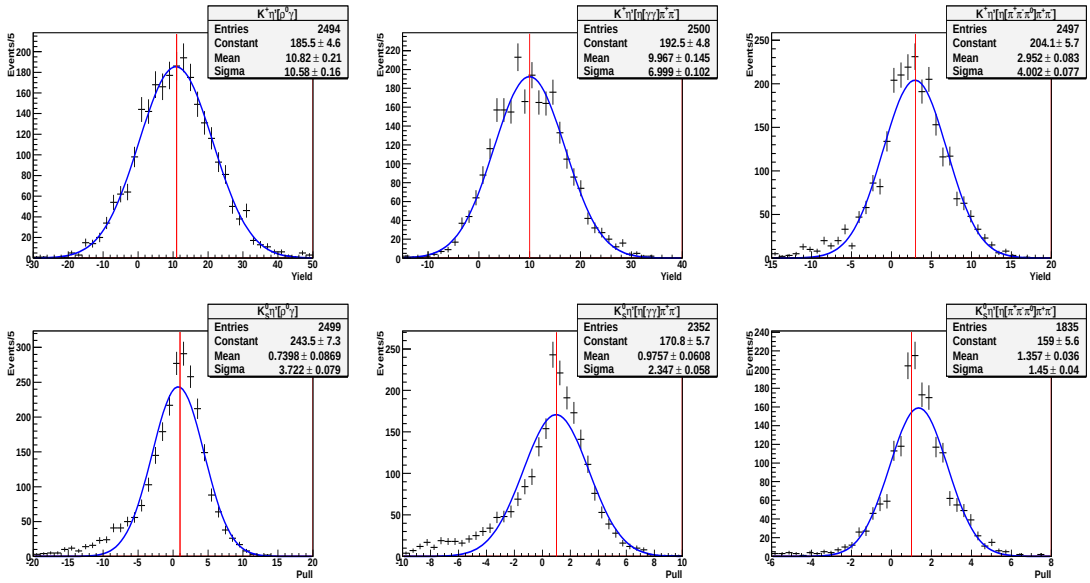


Figure F.2: The yield distributions and fits for the 3D mode-by-mode method fitter. The top row shows the charged final states, the bottom the neutral final states. From left to right, $K\eta(\rho^0\gamma)\gamma$, $K\eta'(\eta(\gamma\gamma)\pi^+\pi^-)\gamma$ and $K\eta'(\eta(\pi^+\pi^-\pi^0)\pi^+\pi^-)\gamma$. The red lines are the true signal value for the fits.

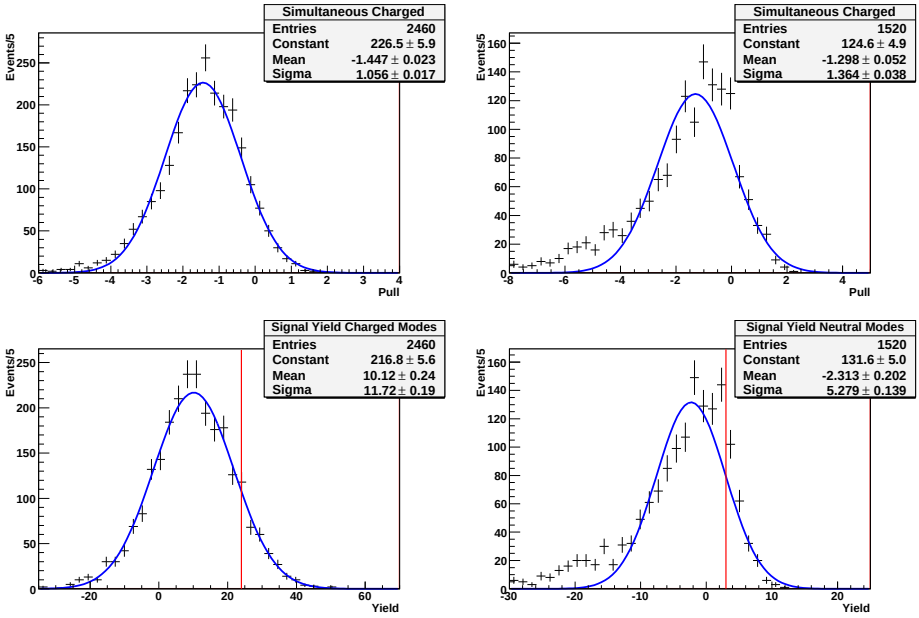


Figure F.3: The pull distributions (top row) and signal yield distributions (bottom row) for the simultaneous fits. Left is the charged final states, right the neutral final states. The red lines are the true signal value for the fits.

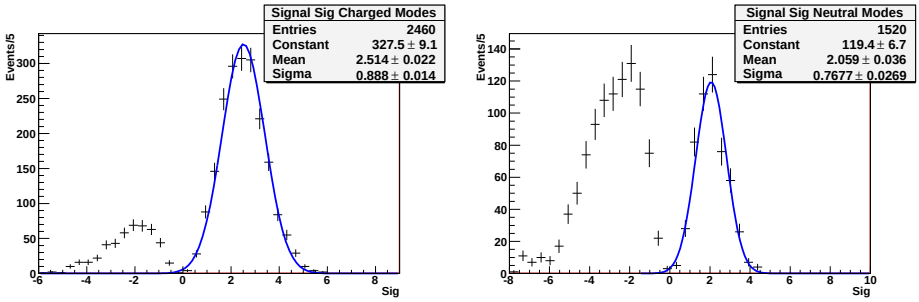


Figure F.4: The significance distributions for the simultaneous fits. Left is the charged final states, right the neutral final states.

Bibliography

[1] A. Pich, *arXiv:hep-ph/0705.4264*, 2007.

[2] The LEP Higgs working group, Phys. Lett. B **565**, 61 (2003).

[3] A. Sakharov, Pisma Zh.Eksp.Teor.Fiz. **5**, 32 (1967).

[4] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, Phys. Rev. Lett. **13**, 138 (1964).

[5] J.R. Browstein and J.W. Moffat, Astronomical J **636**, 721 (2006).

[6] D. Clowe *et al.* , Astronomical J **648**, L109 (2006).

[7] A.G. Riess *et al.* , Astronomical J **166**, 1009 (2006).

[8] Y. Fukuda *et al.* , Phys. Rev. Lett. **81**, 1562 (1998).

[9] Q.R. Ahmad *et al.* , Phys. Rev. Lett. **89**, 011301 (2002).

[10] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983)

[11] CKMfitter Group (J. Charles et al.), Eur. Phys. J. **C41**, 1-131 (2005).

[12] S.L. Glashow, J. Iliopoulos and L. Maiani, Phys. Rev. D2, 1285 (1970).

[13] A.I. Vainshtein, V.I Zakharov and M.A. Shifman, J. Exp. Theor. Phys. Lett. **22**, 55 (1975).

[14] J. Ellis, M.K Gaillard, D. Nanopoulous, and S. Rudaz, Nucl. Phys. B **131**, 285 (1977).

[15] A.J. Buras, M. Misiak, M. Munz, and S. Pokorski, Nucl. Phys. B **424**, 374 (1994).

[16] B. Grinstein, R. Springer, and M.B. Wise , Phys. Lett. **202**, 138 (1988).

[17] A.J. Buras, A. Czarnecki, M. Misiak, J. Urban, Nucl. Phys. B **631**, 219 (2002).

[18] M. Misiak *et al.* , Phys. Rev. Lett. **98**, 022002 (2007).

[19] Particle Data Group, W.-M. Yao *et al.* , J. Phys., G **33**, 1-1232 (2006).

[20] M. Beneke and M. Neubert, hep-ph/0210085v1, (2002).

[21] U. Haisch, hep-ph/0805.2141v1, (2008).

[22] Heavy Flavour Averaging Group, E. Barberio *et al.* , hep-ex/0704.3575 (2007).

[23] M. Misiak *et al.* , Phys. Rev. Lett. **98**, 022002, (2007)

[24] M. Nakao *et al.* , Phys. Rev. **D69**, 112001 (2004).

[25] H. Yang *et al.* , Phys. Rev. Lett. **94**, 111802 (2005).

[26] S. Nishida *et al.* , Phys. Rev. Lett. **89**, 231801 (2002).

[27] A. Drutskoy *et al.* , Phys. Rev. Lett. **92**, 051801 (2004).

[28] S. Nishida *et al.* , Phys. Lett. B **610**, 23 (2005).

[29] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49** 652 (1973).

[30] J.Tanaka, Belle Internal Note 193, (1998).

[31] F. Fang, Belle Internal Note 323, (2000).

[32] S. W. Herb *et al.* , Phys. Rev. Lett. **39**, 252 (1977).

[33] B. Barish *et al.* , Phys. Rev. Lett. **76**, 1570 (1996).

[34] J. P. Alexander *et al.* , Phys. Rev. Lett. **86**, 2737 (2001).

[35] P. Koppenburg, Belle Internal Note 665, (2004).

[36] K. Hanagaki *et al.* , Nucl. Instrum. Meth. A **485**, 490 (2002)

[37] A. Abashian *et al.*, Nucl. Instrum. Methods. A **479**, 117 (2002).

[38] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods. A **499**, 1 (2003), and other papers in this volume.

[39] R. Itoh, Belle Internal Note 161, (1996).

[40] D. J. Lange, Nucl. Instrum. Methods. A **462**, 152 (2001)

[41] R. Brun *et al.* , CERN Report DD/EE/84-1, 1984.

[42] T. Sjostrand, Comput. Phys. Commun. **82**, 74 (1994).

[43] B. Andersson *et al.* , Phys. Rept. **97**, 31 (1983).

[44] Cleo Collaboration. See http://www.cornell.edu/

[45] The Fox-Wolfram moments were introduced in G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978). The modified moments used in this Letter are described in S. H. Lee *et al.* , Phys. Rev. Lett. **91**, 261801 (2003).

[46] R. A. Fisher, *The use of multiple measurements in Taxonomic problems*, Annals of Eugenics, **7** 179 (1936).

[47] H. Kakuno *et al.* , Belle Internal Note 786, (2005).

[48] S. Villa, Belle Internal Note 811, (2005).

[49] A. Limosani, Belle Internal Note 1035, (2008).

[50] I. Nakamura, Belle Internal Note 774, (2005).

[51] R. A. Fisher, *On an Absolute Critereon for Fitting Frequency Curves*, Messenger of Mathematics, **41** 155 (1912).

[52] F. James, *Minuit User Guide: CERN Computer Center Program Library*, CERN-D506.

[53] B. Aubert *et al.* , Phys. Rev. D **74**, 031102(R) (2006).

[54] H. Kakuno *et al.*, Nucl. Instrum. Methods. A **533**, 516 (2004).

[55] Particle Data Group, S. Eidelman *et al.* , Phys. Lett. B **592**, 1-1109 (2004).

[56] B. Casey *et al.* , Belle Internal Note 390, (2001).

[57] T. Skwarnicki, Ph.D. Thesis, Institute for Nuclear Physics, Krakow 1986; DESY Internal Report, DESY F31-86-02 (1986).

[58] ARGUS Collaboration, H. Albrecht *et al.* , Phys. Lett. B **241**, 278 (1990).

[59] Kernel Estimation in High-Energy Physics, *K. Cranmer*, Comput. Phys. Commun. 136 (2001) 198207, hep-ex/0011057.

[60] P. Koppenburg, Belle Internal Note 621, (2003).

[61] G. Majumder, Belle Internal Note 641, (2003).

[62] H.W. Kim, M. Nakao, S. Nishida, H. Ikeda, Y. Ushiroda, R.S. Lu, Belle Internal Note 499, (2001).

[63] S.W. Lin, Belle Internal Note 645, (2003).

[64] M. Nakao, Belle Internal Note 513, (2003).

[65] http://belle.kek.jp/group/pid_joint/kid/kid.html

[66] http:/belle.kek.jp/secured/nbb/nbb.html