A STUDY OF EXCLUSIVE CHARMLESS SEMILEPTONIC B DECAYS WITH THE CLEO DETECTOR

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Véronique Boisvert

May 2002

A STUDY OF EXCLUSIVE CHARMLESS SEMILEPTONIC B DECAYS WITH THE CLEO DETECTOR

Véronique Boisvert, Ph.D.

Cornell University 2002

We report on a study of exclusive semileptonic $b \to u$ decays in 9.7 million $B\bar{B}$ events accumulated with the CLEO detector produced in the Cornell Electron Storage Ring (CESR). We reconstruct candidates in the exclusive decay modes $B \to [\pi^{\pm}, \pi^0, \rho^{\pm}, \rho^0, \omega, \eta]\ell\nu$, where the charged lepton is an electron or a muon. We use the hermeticity of the CLEO detector to infer the neutrino four momentum. The ISGW2 theoretical calculation of the $B \to X_u$ form factors is used to determine total branching fractions, the partial widths as a function of the momentum transfer q^2, and an estimate of the CKM matrix element $|V_{ub}|$. We find a value for the $B \to \pi^{\pm}\ell\nu$ branching fraction of $(1.470 \pm 0.179^{+0.129}_{-0.150}) \times 10^{-4}$ and a value of $(1.657 \pm 0.228^{+0.421}_{-0.481}) \times 10^{-4}$ for the $B \to \rho^{\pm}\ell\nu$ branching fraction. We also report a value of $(0.837 \pm 0.306^{+0.082}_{-0.092}) \times 10^{-4}$ for the $B \to \eta\ell\nu$ branching fraction. Finally, we find a value of $|V_{ub}|$ of $(2.913 \pm 0.128^{+0.165}_{-0.220}) \times 10^{-3}$.

BIOGRAPHICAL SKETCH

Véronique Boisvert was born to Line Fleurent and Michel Boisvert on August 24^{th} 1973 in Montréal, Canada. Her years as an only child happily ended in 1984 when her sister Marie-Michèle was born. Véronique attended Collège Regina Assumpta for her five years of high school. This is where she decided that she would pursue a career in science, and physics was her first choice. She then continued her education by spending the following two years at Collège-de-Bois-de-Boulogne, where she majored in natural sciences. While at this college she was part of the science club and participated in science fairs, leading her to be chosen to be part of the Québec delegation at the International Science fair of 1992 held in Amarillo, Texas. She joined the Université de Montréal to get her B.Sc. in physics from 1992-1995. While working as a summer researcher with Prof. David London, she decided that her research area in graduate school would be the phenomenology of particle physics.

Those plans were changed when she entered the dynamic physics department at Cornell University in the fall of 1995. She got her M.S. in physics in 1999 having chosen the field of experimental particle physics. While at Cornell she also got involved in various activities pertaining to the improvement of the academic life of graduate students. Twice, in 1997 and 2001, she gave a talk to the High-Energy Physics Advisory Panel on recommendations on how to improve the field for young physicists. She is now heading to CERN where she will be a research Fellow working on one of the LHC experiment.

To my family and Line, Michel et Marie-Michèle,

ACKNOWLEDGEMENTS

I would foremost thank the wonderful crew of CESR and CLEO personnel that have built and maintained the wonderful facilities at the Laboratory of Nuclear Studies, allowing me to do such an exciting analysis.

I would like to thank my advisor, Lawrence, who made all this work possible in the first place. I am truly honored to be his first of what I am sure will be a long list of accomplished students. Lawrence's ability to grasp the theoretical issues pertaining to experimental analyses is something that I admire greatly. I feel I have become an experimentalist from learning so much from Lawrence. I know for a fact that my experience as a graduate student at Cornell would have been completely different if it had not been for my interaction with Persis. Her constant encouragement was a blessing. Her joining SLAC is a loss for Cornell and the laboratory, but it is a great gain for the field of High-Energy physics. Another faculty which truly makes the CLEO atmosphere as special as it is is Ed Thorndike. I am reminded daily of Ed's ability to interact with the various members of the collaboration. Ed is the special friend of all the graduate students: his much appreciated sense of humor makes us keep faith that one day we might stop making stupid mistakes! Finally I must thank David Cassel for his insightful advice through the years.

Performing an analysis can sometimes be quite an adventure. I feel I was extremely lucky to have interacted not only with quite amazing faculty but also with remarkable students and post-docs. I am specially grateful to my officemates and $|V_{ub}|$ partners, Tom and Andreas. Andreas's organized way of approaching any project is something I hope I can emulate when I am a post-doc myself. Tom's ability to understand any experimental concept (although he will deny this) has made physics discussion quite deep and rewarding. But more than my interaction with them in the lab, I feel I have a special connection with Tom and Andreas. They sometimes acted as surrogate boyfriends by patiently listening to me and offering me comfort and advice. I hope we will be in touch for a long time to come.

The Cornell physics department is known for being very convivial. All the students help each other during the first few years of classes. I want to thank my classmates for their support, specially Chad, Alex, Erik A., Basu, Anja, Iya and Jeff N. I doubt I will again be in a collaboration as friendly and cooperative as CLEO. I have made many friends on CLEO over the years and I will miss them. At this point I doubt I will be able to list everybody, but here is a try: Alain, Adrian, Abi, Dave C., Bruce, Ye, Ken, Dan B., Ed P., Georg, Dave A., Jesse, Adam, Stefan, Chris S., Pete Z., Naresh, Charles, Roy, Elliott, Werner, Anton, Toni H., Mark P., Chris J., Toni R.. I also particularly want to thank Craig P. for all his incredible work for CLEO and the special relationship we have. I would also like to thank Gregg for being such a good and funny friend. I also want to thank the women of the lab which have made such a big impact on my well-being at the lab: Silvia, Jana, Lauren, Bonnie, Anto, Hannah and Nadia, you girls are

the best, Silvia, I'm so excited to be close to you again, and Jana you have been a great support for me and I can't wait to hear of all your accomplishments. A special thanks goes to Andrew for making me learn so much about myself.

I also have to thank my second family in Ithaca, my various housemates. I would particularly like to thank Vern, you are a truly exceptional woman. Tammie and McKenna, you hold a very dear and special place in my heart, good luck to you and we'll keep in touch. Mary, back in Chile, I hope to visit you at some point. Mike, Jana and Chris L., I'm really glad I met you. To my "special" friends: Omer, I hope you will remain safe and sound, Mark you are an important friend to me, please stay in touch.

Finally, my family and I have had a rocky time for the last few years, but I feel that we have grown closer and I'm very grateful for this. I feel now more than ever that my place is in Montreal, and I hope that the dice will be favorable.

Peace on Earth.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Overview

At about the 15 billion-year mark in the great race of Nature, one interesting milestone was achieved: the human being. Even more astonishing than all of the technological, social, political and cultural achievements of this amazing species is their desire to understand the world around themselves which led them to study what is now known as physics, chemistry and biology. It's to answer this intrinsic need of human nature that I decided to become a physicist. Physicists push back the limit of our knowledge of Nature, but their work is only meaningful if they can transmit that knowledge to the rest of the human community. This is not only true for physicists: any scholar of any field has the responsibility to share his or her discoveries. The *Philosophia Doctoris* thesis used to serve as the traditional method of propagating such information. In modern times, scientific research is so complex that *Ph.D* theses fail to render to the general public, or to fellow university scholars, even the gist of the study at hand. It is important to me to somewhat restore this tradition. In this first section I take the reader on a journey of particle physics. I first set the stage for what modern particle physics is about, then I introduce the main actors relevant to my research, finally, I describe what my research tries to accomplish. In the second part of this chapter, I accompany the readers familiar with the field through a grand tour of the Standard Model.

1.1.1 The Stage

In the expression "particle physics" resides the word particle. The classic image that comes to mind is that of a billiard ball, of a specific mass, and other characteristics. It can hit other balls, and if we use conservation laws like those of momentum (the product of mass and velocity) and energy, we can predict exactly the trajectory of such a ball. This was also the picture that early physicists had of elementary particles. The atom was thought of in terms of a nucleus of protons and neutrons, with one or more electrons orbiting around it. This classical picture quickly fell apart when confronted with experiments. When we enter a regime where systems are very small, we need to use quantum mechanics as opposed to the classical mechanics of billiard balls. Elementary particles certainly fall into the small system category. The revolutionary idea behind quantum mechanics is the realization that the natural world is described by means of observation and measurements, and such actions by definition perturb the system under study. This leads to the introduction of probabilities, since until we observe particular characteristics of a system, the system contains all the different possibilities within it. If any observation impacts the system under study, can we think of the smallest perturbation possible? This corresponds to one quantum of action, which is the product of the amount of energy involved in the perturbation by the amount of time the perturbation took. The numerical value of this quantum of action is about 1×10^{-34} Joules second.

Elementary particles are not only small, they usually move very fast, so fast as to approach the speed limit: the speed of light in vacuum. Again classical physics at low speed is not adequate for systems at high speed. We need the theory of special relativity. For the case of elementary particles we need to put those two extremes together, quantum mechanics with special relativity, to give us quantum field theory. One consequence of picturing particles in this context is that when we probe particles at high energies we see that they are in fact surrounded by a cloud of virtual particles. They are virtual because they violate conservation of energy, but since they exist for a very short time, the quantum of action is not violated. An experimental consequence of the presence of such particles is to partially screen the charge of the particle they are surrounding. Hence intrinsic characteristics like the charge of a particle is actually dependent on the scale of the experimental problem.

We have just seen that the concept of "what is a particle" is greatly modified as we go into the realm of elementary particles that are small and go fast. What about the concept of force? We can think of a force as the source of why oppositely charged particles attract each other while particles of the same charge repel each other; this is the electric force. An apple falling from a tree involves the gravitational force. The classic representation of such forces is that of an action taking place across some distance. This gets modified in the quantum field context: the force is carried by a messenger, a virtual particle that carries the force back and forth between the particles feeling the force. The carriers are different depending on which force is at play.

We are now ready to look in detail at what are the elementary particles and basic forces according to the model that has successfully succeeded under numerous experimental tests, the so-called Standard Model. We start out with the atom, since most people are familiar with it: some number of electrons surround a nucleus made up of protons and neutrons. It was found experimentally that the electrons are elementary particles, while the protons and neutrons are composite particles: they are made out of quarks. To give an idea of the relative size involved in an atom, the Particle Data Group (PDG) has a good analogy: if the protons and neutrons were 10cm across, then the quarks and electrons would be less than 0.1mm in size and the entire atom would be about 10km across! The electrons and quarks are two branches of the same family called fermions; so called because these particles have an intrinsic angular momentum (we call it spin) in units of a half-fraction of the unit of action seen previously. Although the word spin evokes the classic image of an object rotating, spin is a purely quantum quantity and does not mean that the electrons or quarks are rotating. But the spin has influence on experimental characteristics: for example, it determines the behavior of the particle in a magnetic field. Another consequence of having half-integer of spin is that the fermions obey Fermi statistics, which lead to the Pauli exclusion principle: no two fermions can be in the same quantum state. One consequence of this principle is the structure of the periodic table of elements.

We now introduce some more fermions. In many particle reactions involving the electron, there is an elusive partner that accompanies it: the neutrino. It is elusive because it does not carry any electric charge, and it is believed to have very tiny mass. Neutrinos are nonetheless crucial since they are produced copiously; for example, our Sun produces millions of those neutrinos all the time, and they reach the Earth and go through our bodies without us noticing. The exact value of their mass is the subject of ongoing research, since their mass combined with the fact that there are so many of them could have drastic consequences on the structure of the Universe. The electron and neutrino have two sisters each: the muon and its partner the muon neutrino, and the tau, and its partner the tau neutrino. The muons and taus are heavier "copies" of the electrons. The electron, muon, tau and their associated neutrinos are referred to as the leptons. Aside from producing them in particle accelerators, we find muons and taus only in cosmic showers coming from space and reaching our atmosphere. Another type of particle that is also produced in particle accelerators and not commonly found around us is the anti-particle. The anti-electron has the same mass and spin as the electron, except it has opposite charge, hence it was called the positron. When a particle encounters its anti-particle partner they annihilate, releasing their combined energy.

We continue our tour of the atom: the proton is made out of two up (u) quarks and one down (d) quark, while the neutron is made out of two down quarks and one up quark. Everything in our everyday life is made out of those two kind of quarks, but just like the electron, the u and d quarks have siblings: the charm (c) and strange (s) quarks for the second generation, and the top (t) and bottom

(b) quarks for the third generation. The quarks are also fermions, so they carry half-integer units of spin. Particles that are made out of quarks are called hadrons, and they come in two classes: those that have three quarks (baryons), and those that contain a quark and an anti-quark (mesons). An example of a meson is the pion (π), which contains an up quark and an anti-down quark when it has positive charge. Since quarks carry half-integer spin, then the baryons also carry half-integer spins, while the mesons carry integer spin. Instead of being fermions, the mesons obey Bose-Einstein statistics, and so are referred to as bosons. One consequence of bosonic statistics is that the Pauli exclusion principle does not apply for these particles, hence we can have peculiar states of matter where all the bosons particles are in the same state.

Mesons are composite particles, but there are also elementary particles that have integer units of spin: the force carriers. The force carrier of the electromagnetic interaction (electricity and magnetism) is the photon, the bosons associated with the weak force are the two charged W's and the neutral Z, and the boson associated with the strong force is called the gluon. The strong interaction is responsible for keeping quarks together inside hadrons. For example, inside a proton the three quarks keep exchanging gluons back and forth. Just as the electromagnetic interaction acts on the electric charge of the particles feeling the force, the strong interaction acts on a charge that we call “color”. Hence quarks not only carry electric charge, they also carry color charge. There are two possible states for the electric charge (+ or -) but there are three possible states for the color charge (red, blue, green, or any other combination of names that refer to three colors). Unlike the electric charge, hadrons cannot have a net color charge associated with themselves, the colors associated with the quarks must add up once inside the hadrons, so that the hadron is colorless (or "white"). Also unlike the electromagnetic interaction, the force carriers of the strong interaction, the gluons, carry color charge themselves (remember, the photon is electrically neutral). If the protons and neutrons are colorless, then what binds the nucleus together? There is a residual strong force coming from the colored quarks and gluons inside the protons and neutrons; this is analogous to the residual electrical interaction inside a molecule which binds together various neutral atoms.

Another fundamental interaction is the weak interaction. It is the only known interaction that applies to neutrinos. For example, the weak interaction is responsible for the decay of the neutron (see figure 1.1). One of the down quarks inside the neutron turns into an up quark by emitting a virtual W^- boson, which then decays into an electron and an anti-electron neutrino. The up quark combines with the two spectator quarks that were inside the neutrons, to form a proton. This interaction is called weak because if we were to compare the strength of that interaction with that of the strong force for two up quarks at some distance, the strength of the weak interaction relative to the electromagnetic interaction would be down by a factor of 10^{-4}, while the strength of the strong interaction relative to the electromagnetic one would be 60. The charge associated with the weak interaction is called "flavor". Leptons and quarks both carry flavors, but the flavor is different for an electron than a neutrino and different for an up quark than a down quark. Another fundamental interaction, and we are reminded everyday of its presence, is the gravitational interaction. It is not described by the Standard Model, because it has little influence on elementary particles since its strength is proportional to mass. Gravity is the main actor involved in the global structure of galaxies and the creation of stars. The common belief among physicists is that at some point early in the creation of the Universe, all interactions were unified into one single interaction. As the Universe cooled down, that interaction broke down into four separate interactions. So far physicists have managed to unify the weak interaction with the electromagnetic one. The daunting challenge facing theorists nowadays is to combine the quantum field theory relevant to the Standard Model with the General Relativity involved in describing the gravitational interaction, a unification of the infinitely small with the infinitely big. Table 1.1 summarizes the various elementary particles and interactions.

1.1.2 The Synopsis

Now that we have set the stage of modern particle physics, it is useful to peek at the script of my analysis to help put things in perspective, in view of the next few sections. I intend to make clear how I go about to use the CLEO detector to detect a B meson that decayed into a hadron that contains an up quark (pion, rho, omega or eta), a charged lepton (electron or muon) and a neutrino. The goal of looking at this particular decay is two fold. First, this reaction involves a transition from a b quark to a u quark and the rate of this reaction is proportional to a fundamental parameter of the Standard Model, V_{ub}, for which I get a measurement. Second, the rate distribution as a function of kinematic variables sheds light on the inner

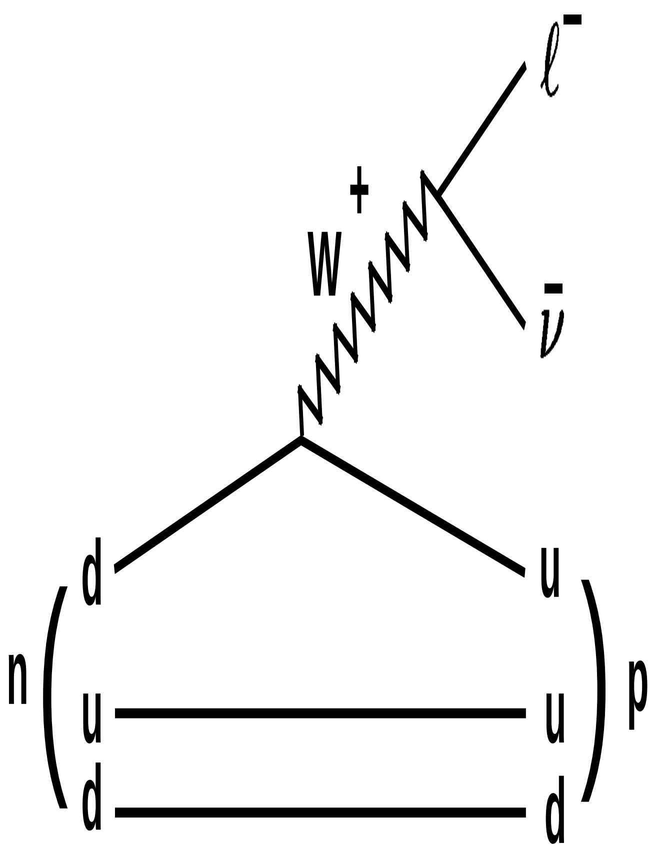

Figure 1.1: A neutron decays into a proton using the weak interaction.

Table 1.1: Elementary particles and interactions

Interactions:	Electromagnetic	Strong	Weak		Gravitational
Acts on:	electric charge	color	flavor		mass-energy
particles experiencing:	electrically charged	quarks, gluons	quarks	leptons	all
			u c t	e μ τ	
			d s b	ν_e ν_μ ν_τ	
carriers:	photon	gluons	W^+, W^-, Z^0		graviton

working of the hadronic part of this decay. At present, very little experimental information is available about this particular transition, getting a measurement of this distribution will help the community understand better this murky area of QCD physics.

1.1.3 The Actors

With the synopsis of the analysis in mind we now consider the various actors involved in my research. Taking another look at the neutron decay, we see that the weak interaction is at play in the decay of the down quark. The strong interaction is also involved since there are spectator quarks that somehow go from being part of a neutron to being part of a proton. Gluons are being exchanged throughout this process, and specifically how many of them and of what energy has consequences on the final products of this reaction. Looking now at one of the reactions that is the subject of this thesis, we see a similar picture (see Figure 1.2): there is an anti-B meson composed of a b quark and a anti-d quark, the b quark decays into a u quark, via the emission of a virtual W^- weak boson. The W^- boson then decays into an electron and an anti-electron neutrino. The anti-d quark then makes a pion with the u quark.

We have just determined that in order to perform my analysis I need B mesons, and I also need to be able to tell if the B meson decayed to a lepton (electron or muon), a neutrino, and a pion, for example. The first order of things is to be able to produce B mesons. To do so, the accelerator collides electron and positron against each other. The energy of this collision is just right so that there

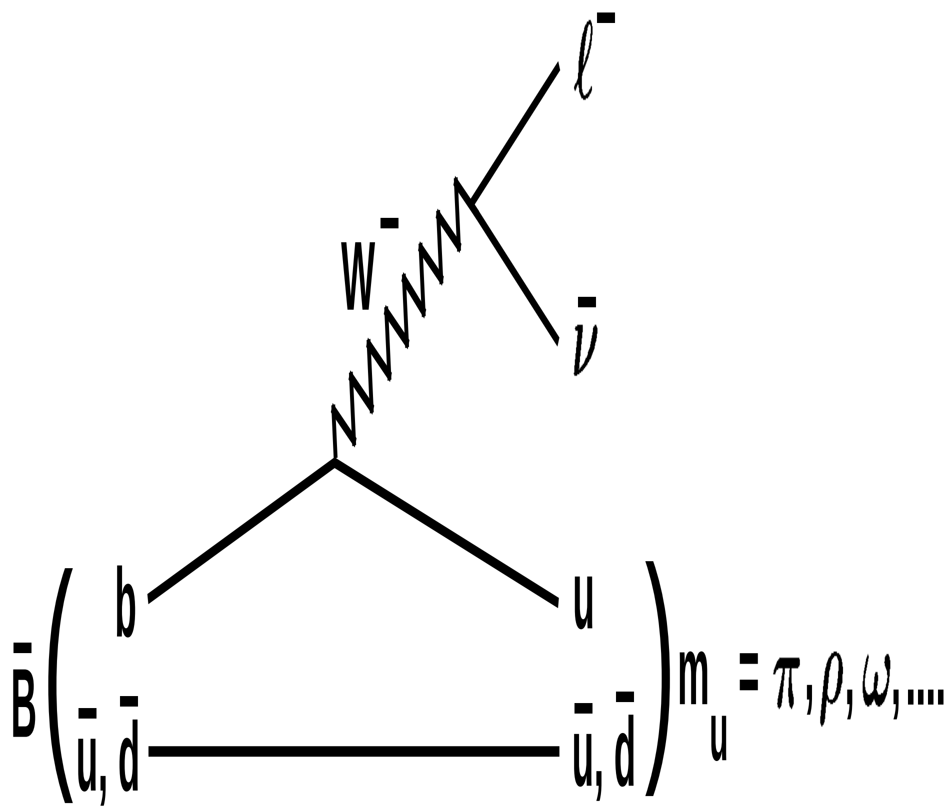

Figure 1.2: A B meson decays into a hadron containing an up quark, a lepton and a neutrino.

is enough to produce the heavy b quarks, which hadronize into a pair of B and anti-B mesons, but not too much energy, so that the B mesons are almost at rest. This makes the reconstruction of the event easier. For example, no other particles are created along with the B meson pair. In order to get the colliding electrons we heat up a metal filament. Electrons in the atoms get enough energy so that they can escape the atom. This happens in the Linear Accelerator (Linac), which accelerates the electrons to 150 million electron volts (your TV tube accelerates electrons to 20,000 electron volts). To get the positrons, there is a metal plate in the middle of the Linac. It gets bombarded by electrons and X-rays as well as electrons and positrons come out of the interactions. The positrons are then selected using the fact that they have charge opposite to the electrons. After the Linac, the electrons and the positrons get transfered into the synchrotron, which further accelerates the particles. First it's the positron turn, and when they have reached 5 billion electron volts, they get transfered into the Storage Ring (CESR). It takes the particles about 2000 revolutions around the synchrotron for them to reach the required energy, this takes place in about one hundredth of a second. Once the electrons have been transfered to CESR as well, collisions can happen. More information on the accelerator can be found in section 3.1.

Once the collision happens at the center of CLEO, the equivalent of 10 billion electron volts of energy is released in new particles that go flying away. That amount of energy is equivalent to the amount of energy produced by a 100 W light-bulb turned on for only 16 picoseconds! This is not very impressive, why are we not producing new particles by just turning lights on? The key feature is the density of energy: the energy produced by the synchrotron is given to each colliding electron, while the light-bulb is sharing its energy among billions of particles.

For our case of interest, the energy generated go into producing a pair of B mesons. In order to tell if a B meson decayed into a lepton, a neutrino, and a pion, detectors are needed that register exactly what happened: how many particles were created, where they went, what they were, how much energy they carried. Different types of detectors answer the different questions. For example, at a radius of about 8cm from the interaction point, there is a one meter radius cylinder filled with gas. In the cylinder are about 40,000 wires strung along the cylinder axis. The wires are held at a certain voltage. When a charged particle comes into this cylinder it ionizes the gas atoms, separating the electrons from the atom. The newly released electrons drift toward the sense wires and collect there, sending an electrical impulse along the wire. The cylinder is bathed in a magnetic field, which bends the trajectory of the charged particle. Just how much the trajectory is bent depends on the value of the magnetic field and on the momentum (remember, product of mass and velocity) of the charged particle. By aligning all the wires that reported a signal, we can reconstruct the curved trajectory of the charged particle, giving us two pieces of information: where it went and how fast it was going. We have another detector that gives us the velocity of the particle. It is called the time of flight detector since it tells us how long the particles took to go a fixed distance from the interaction point, where there is a plastic scintillator which registers the passage of the particle. Putting the momentum and velocity information together we can get the mass of the particle, and so its identity. Overall, the CLEO detector is a multipurpose detector that gives excellent charge and neutral particle detection and measurement. It is 6m on a side and consists of about 900,000KG of iron. It has about 50 000 individual detector elements. The CLEO collaboration consists of about 150 physicists from about 25 institutions in the US and Canada. More information on the CLEO detector can be found in section 3.2.

1.1.4 The Play

We now turn our attention toward the fundamental interactions involved in my decay of interest. The strong interaction is involved in making the final bound state of quarks, the pion. We can examine two extreme kinematic cases (see Figure 1.3). In one case, most of the initial energy is used up to create the virtual W boson, so that not much energy is left for the final pion. In that case, the electron and neutrino go flying back to back (situation b) in figure). From the quark's point of view, not much disturbance happens for the spectator anti-d quark, and so presumably not many gluons have to be exchanged to make the final pion. This is in contrast to the opposite case where most of the initial energy goes into the final pion, so that now the electron and neutrino are going in one direction, while the pion is going in the opposite direction (situation c) in figure). From the quark's perspective a lot happens to the anti-d quark since it goes from being part of a relatively slow B meson to a fast π. A lot of gluons have to be exchanged in this case. We can see that from the strong interaction's point of view, the first case is more favored than the second case. Concretely this means that the probability for a particular interaction to follow the first case is higher than the probability to follow the second case. A way to see experimentally if this is true is to count how many reactions occurred each way. Of course, there are all the cases in between those two extremes. By labeling the events according to, say, how much energy was given to the lepton-neutrino pair (called q^2), we get an experimental distribution that has a direct connection to the strong interaction. We call this distribution, the form factor relevant to the reaction. This is one of the main goals of my research: to get the form factors as a function of q^2 and to compare the distributions of various theories, each of which use different calculations. Seeing which calculations are favored by the experimental results will give some understanding of how the strong interaction behave in this specific circumstance, and hopefully, future calculations for a myriad of other processes will be able to make use of the new information.

Can we gain new information from the weak part of the reaction? The b quark decays into a u quark: this is a jump from the third to the first generation, and is not very likely. Exactly how likely is represented by a number called V_{ub}. Each kinematically possible quark transition is quantitavely labeled by how probable it is. Figure 1.4 shows the various quark transitions allowed and the thickness of the arrows is correlated with how likely each transition is. The Standard Model does not predict what the numerical value is for V_{ub}, we have to measure it. Whether the value that I measure is in agreement with other quark transition probabilities sheds light on the validity of the Standard Model.

In a given parameterization of the different quark transition probabilities, the size of the value of V_{ub} is also related to a phenomenon called CP violation. CP violation is related to the notion that physical reactions must obey some gen-

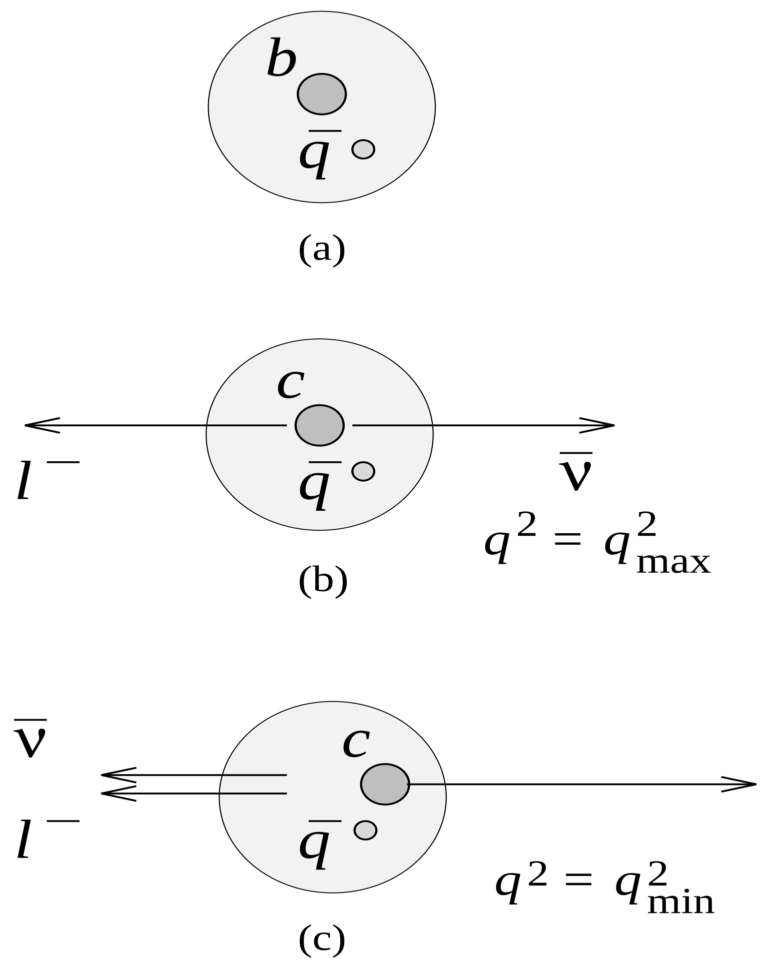

Figure 1.3: Different kinematic limits of a heavy to heavy quark transition.

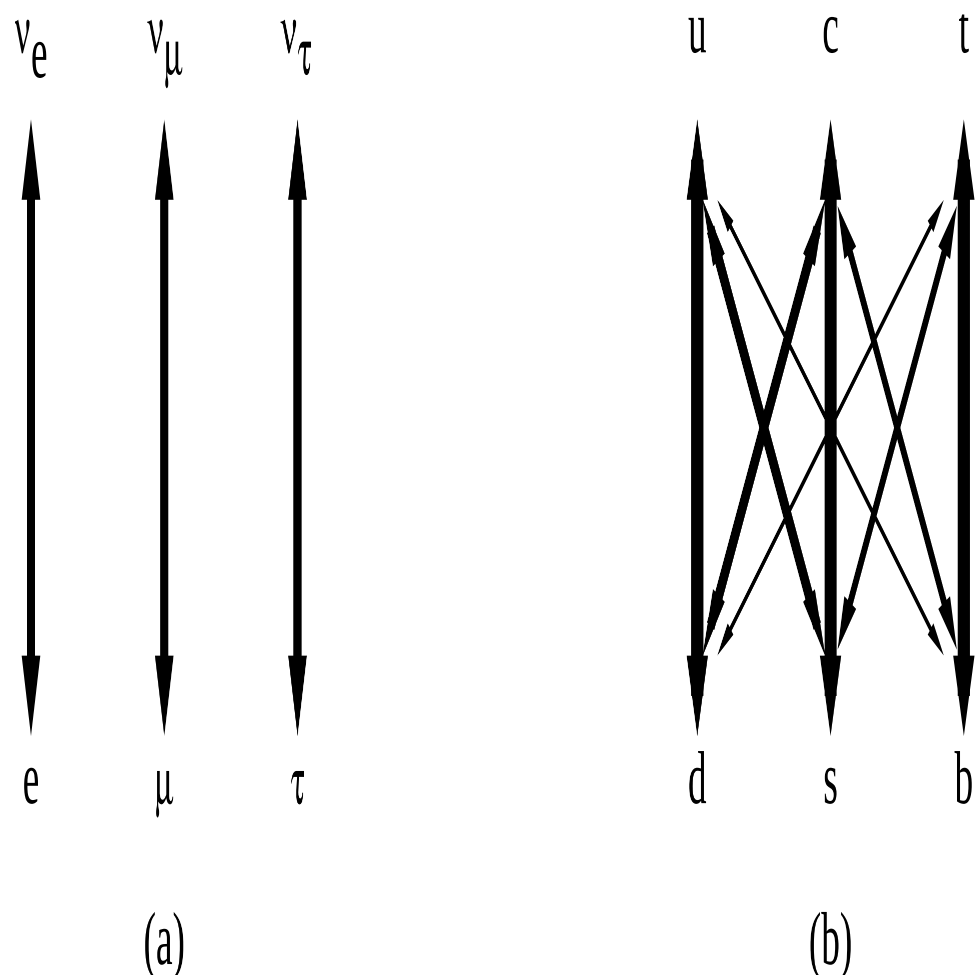

Figure 1.4: Quark and lepton transitions.

eral symmetry principles. For example, we would expect that a physical reaction involving particles would have the same probability of happening as the same reaction involving the corresponding anti-particles. Changing particles for their anti-particles is applying the charge conjugation (C) operator. We would also expect that the mirror image of a reaction would have the same probability of happening as the original reaction. Flipping the space coordinates and the spin projection of the particles is applying the parity (P) operator. Finally, we would expect that there is no preferred time direction to a reaction. Flipping the space coordinates while conserving the same spin projection to a reaction is applying the time (T) operator. It turns out that the weak interaction does not conserve charge conjugation and does not conserve parity either. We also have evidence that even the product of both C and P operators is not conserved in weak interactions, although the amount by which it is violated is small. Besides having some philosophical implications, like the fact that the Universe seems to distinguish between "right" and "left", CP violation also has some connection as to why we exist at all. At the beginning of the Universe, particles and anti-particles kept annihilating each other. The fact that our current Universe is made out of particles means that the particles won out against anti-particles at some point. The weak interaction and its CP violation could be one mechanism for this to happen. It is important to note that the amount of CP violation that the Standard Model predicts, and which seems to be confirmed by experimental measurements, including my V_{ub} measurement, is not enough to satisfy the amount of CP violation needed by models of the early Universe. This is one of the signs that the Standard Model is not the end of the story.

1.1.5 Structure of this thesis

The rest of this chapter reviews more in detail the foundations of the Standard Model and how it connects to $B \to X_u \ell\nu$. Chapter 2 reviews what we know theoretically about the $B \to X_u \ell\nu$ decays. Chapter 3 goes into more details on the experimental apparatus. Chapter 4, 5 and 6 go over the different steps of measuring the branching fractions, while chapter 7 describes the V_{ub} measurement and interprets the results in light of the theoretical models.

1.2 The Standard Model

In order to put the measurements described in this research into context, I take a tour of the Standard Model and beyond. I have chosen a particular lens for this tour: I want to describe how symmetries lead to the structure of the different interactions. Gauge symmetries are the central themes around which revolve the next sections. None of the following work is new or original to me; I summarize references [1] to [11].

1.2.1 Symmetries and conservation laws

Emma Noether's theorem is very powerful for it relates mathematically, in the Lagrangian formulation, the connection between non-observable absolute quantities (symmetries) and the corresponding conservation law. We can classify symmetries in four categories: space-time symmetries, symmetries of identical particles, discrete symmetries, and internal symmetries. Table 1.2 [1] lists the various symmetries and their corresponding conservation laws.

The first seven symmetries are global symmetries since they involve transformations that have the same value at each space-time coordinate. The Standard Model is obtained by requiring the last three symmetries to be local symmetries, or local gauge invariances. This means that the corresponding Lagrangian remains invariant under a transformation which has different values at each space-time coordinate. This more restrictive requirement has the drastic consequence of requiring the existence of forces: force carriers must be emitted to satisfy these constraints. For example, the torque from the gauge invariance of the fields's phase creates the photon, the torque on the color field creates the gluons, the torque on the weak isospin doublets creates the vector boson, and the torque on space-time creates the gravitons. We now examine how this happens for each symmetry.

1.2.2 Quantum Electrodynamics (QED)

In this section we see how a local gauge invariance gives rise to QED. The simplest local gauge invariance that one could think of is a simple phase transformation of the fields [1]. We first take the Lagrangian of a free fermion:

[1] It is interesting to note, as described in [2], that the expression gauge invariance comes from Weyl, who wanted to make interactions invariant under a space-time dependent change of scale. His geometric idea did not work, but we kept the word scale, or gauge.

Table 1.2: Noether's theorem in relativistic quantum theory [1]

Absolute quantity	Symmetry	Conservation law
abs. spatial position abs. time abs. spatial direction abs. velocity	space translation time translation rotation Lorentz transf.	momentum energy angular momentum gen. of Lorentz group
distinguishing identical particles	permutation of identical particles	Fermi Dirac or Bose Einstein statistic
absolute right or left absolute sign of charge	change $\vec{x} \rightarrow -\vec{x}$ change particles to their anti-particles	Parity Charge conjugation
Absolute phase of a charged field absolute difference between mix of colored quarks absolute difference between mix of leptons or neutrinos	change of phase change of color change lepton into its neutrino	electric charge color generators weak isospin generators

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi, \tag{1.1}$$

where the γ^μ are the Dirac matrices. We can impose a phase transformation to the field $\psi' = e^{i\alpha(x)}\psi$, then we get:

$$\mathcal{L}' = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - \bar{\psi}\gamma^\mu(\partial_\mu \alpha)\psi \tag{1.2}$$

We see that we have an extra term in $\mathcal{L}'$ compared to $\mathcal{L}$. The extra term can be taken care of if we introduce a vector potential, A_μ, transforming the derivative into a covariant derivative, $\mathcal{D}_\mu \equiv \partial_\mu + iqA_\mu$, where q is the charge. Then the Lagrangian is

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - q\bar{\psi}\gamma_\mu A_\mu \psi. \tag{1.3}$$

Applying the phase transformation to the fermion field, we now get

$$\mathcal{L}' = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - \bar{\psi}\gamma^\mu(\partial_\mu \alpha)\psi - q\bar{\psi}\gamma^\mu A'_\mu \psi \tag{1.4}$$

and we see that to get $\mathcal{L}' = \mathcal{L}$, the vector potential must change by a gradient, $A'_\mu = A_\mu - \frac{1}{q}\partial_\mu \alpha$. At this point, we could identify the vector potential with the electromagnetic potential. This redefinition of the potential will not change the electromagnetic strength defined as

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{1.5}$$

Now the electromagnetic Lagrangian looks like

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - J^\mu A_\mu, \tag{1.6}$$

where we can identify $J^\mu = q\bar{\psi}\gamma^\mu\psi$ with the electromagnetic current. Indeed we get conservation of charge, $\partial_\mu J^\mu = 0$, when we use the Dirac equations for the fermion fields.

To reproduce the complete Lagrangian of electrodynamics we need a kinetic term associated with the vector potential, representing the propagation of free photons,

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - J^\mu A_\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{1.7}$$

This Lagrangian is still invariant under the local phase transformation, although if we were to add a mass term for the photon,

$$\mathcal{L}_\gamma = \frac{1}{2}m^2 A_\mu A^\mu, \tag{1.8}$$

then we would not get invariance. The transformation involved with $A_\mu^{'}$ generates new terms that can not be canceled. From this local gauge invariance nicely emerges all of electrodynamics.

As mentioned in [3] a phase transformation can be generalized to a 1×1 unitary matrix, and the group of all such matrices is called $U(1)$. So the symmetry shown here is a $U(1)$ gauge invariance. We will now see that the other interactions in the Standard Model also have a similar form of gauge invariance. However in those cases the groups that will be connected to the symmetry are represented by non-

Abelian matrices (matrices that do not commute), which will make the description more complex but richer.

1.2.3 Quantum Chromodynamics (QCD)

Continuing our tour of the Standard Model through the lens of symmetries, we can extend the idea of a $U(1)$ symmetry to the $SU(3)$ group, which is relevant to the case of QCD, describing a quark of some flavor that comes into three colors. If we write the Lagrangian for a particular flavor we get

$$\mathcal{L}_| = \bar{q_j}(i\gamma^\mu\partial_\mu - m)q_j, \tag{1.9}$$

where $j = 1, 2, 3$ for the three colors. The idea is to make the Lagrangian invariant under a local phase transformation of the quark fields, like $q(x) \to Uq(x)$. The matrix U is a 3×3 unitary matrix which can be written as $U = e^{iH}$, where H is a Hermitian 3×3 matrix. It can be decomposed into 9 real numbers as $H = \theta(x)1 + \alpha_a(x)T_a$, where the T_a are the Gell-Mann matrices and there is a sum over $a = 1 \ldots 8$. We now have $U = e^{i\theta(x)}e^{i\alpha_a(x)T_a}$.

In the development of QED we already experimented with a phase transformation like $\theta(x)$, so we can concentrate on the $\alpha_a(x)$ part. It turns out that the matrix $e^{i\alpha_a(x)T_a}$ has determinant 1, so it belongs to the $SU(3)$ group, (the S is for special) and we see here that $U(3) = U(1) \times SU(3)$. As previously stated, the idea is to see that by making the Lagrangian invariant under $SU(3)$, the phenomenological features of QCD nicely emerge. We are here in the presence of a non-Abelian group: the Gell-Mann matrices do not commute, but they obey the following commutation relation,

$$[T_a, T_b] = if_{abc}T_c, \tag{1.10}$$

where the f_{abc} are real constants.

The trick that we learned from QED is to replace the derivative by a covariant derivative and introduce new gauge fields

$$\mathcal{D}_\mu = \partial_\mu + igT_aG^a_\mu. \tag{1.11}$$

The eight new gauge fields, G_μ are the gluons, they are the bosons carrying the force of QCD. The Lagrangian now reads

$$\mathcal{L} = \bar{q}(i\gamma^\mu\partial_\mu - m)q - g(\bar{q}\gamma^\mu T_a q)G^a_\mu. \tag{1.12}$$

We want $\mathcal{L}$ to remain invariant under the $U(3)$ transformation

$$\begin{aligned} q' &= Uq \approx [1 + i\alpha_a(x)T_a]q(x) \\ \overline{q'} &\approx [1 - i\alpha_a(x)T_a]\overline{q(x)}. \end{aligned}$$

The transformed Lagrangian reads,

$$\begin{aligned} \mathcal{L}' &= (1 - i\alpha_aT_a)\bar{q}(i\gamma^\mu\partial_\mu - m)(1 + i\alpha_aT_a)q \\ &\quad +g((1 - i\alpha_bT_b)\bar{q}\gamma^\mu T_a(1 + i\alpha_bT_b)q)G'^a_\mu \\ &= \bar{q}(i\gamma^\mu\partial_\mu - m)q + \bar{q}(i\gamma^\mu\partial_\mu - m)q(i\alpha_aT_a) - (i\alpha_aT_a)\bar{q}(i\gamma^\mu\partial_\mu - m)q \\ &\quad -T_a\bar{q}\gamma^\mu\partial\alpha_aq + g(\bar{q}\gamma^\mu T_aq - i\alpha_bT_b\bar{q}\gamma^\mu T_aq + \bar{q}\gamma^\mu T_ai\alpha_bT_bq \\ &\quad +i\alpha_b\bar{q}\gamma^\mu(T_aT_b - T_bT_a)q)G'^a_\mu \\ &= \bar{q}(i\gamma^\mu\partial_\mu - m)q - T_a\bar{q}\gamma^\mu\partial\alpha_aq - g\bar{q}\gamma^\mu T_aqG'^a_\mu + gf_{abc}\alpha_b\bar{q}\gamma^\mu T_cqG^c_\mu. \end{aligned}$$

Looking at that last equality, we see that to have $\mathcal{L}' = \mathcal{L}$ we must require the transformation law,

$$G'^{a}_{\mu} = G^{a}_{\mu} - \frac{1}{g}\partial\alpha_{a} - f_{abc}\alpha_{b}G^{c}_{\mu}. \tag{1.13}$$

We can now also add the kinetic term for the gluon fields,

$$\mathcal{L} = \bar{q}(i\gamma^{\mu}\partial_{\mu} - m)q - g(\bar{q}\gamma^{\mu}T_{a}q)G^{a}_{\mu} - \frac{1}{4}G^{a}_{\mu\nu}G^{\mu\nu}_{a}. \tag{1.14}$$

To preserve invariance of the new kinetic term $G^{a}_{\mu\nu}$ must be defined as,

$$G^{a}_{\mu\nu} \equiv \partial_{\mu}G^{a}_{\nu} - \partial_{\nu}G^{a}_{\mu} - gf_{abc}G^{b}_{\mu}G^{c}_{\nu}. \tag{1.15}$$

The same argument that required the photon to be massless in QED applies here, hence the gluons must also be massless. The extra term in equation 1.15 is different than in QED and has the profound consequence of adding self-interaction of the gluons to the Lagrangian, reflecting the fact that gluons carry colors. It is interesting to realize that this comes about because the $U(3)$ group is non-Abelian while the $U(1)$ group is an Abelian group.

We look at some more differences between QED and QCD. In quantum field theory, an elementary particle is not thought of as a point particle, but more like a point particle surrounded by a cloud of virtual particles. For example, an electron is surrounded by virtual positron-electron pairs. Hence, there is a vacuum polarization: the positrons tend to align themselves with the central electron, and the measured electric charge is dependent on the distance of the test charge, or said differently, the charge of the electron increases as we get closer to it. We say that the coupling constant runs. The relationship between the running coupling constant and the electric charge is given by $\alpha(Q^2) \equiv e^2(Q^2)/4\pi$, where $Q^2 \equiv -q^2$ is the momentum transfer characteristic to the reaction at hand. In QED the expression for the running coupling constant is ([4])

$$\alpha(Q^2) = \frac{\alpha(\mu^2)}{1 - \frac{\alpha(\mu^2)}{3\pi} log \frac{Q^2}{\mu^2}} \tag{1.16}$$

at large Q^2, where μ is the scale characteristic of the renormalization scheme chosen; a different μ leads to a different expansion of the amplitude. The physical observable is the amplitude squared,$|\mathcal{M}|^2$, and should not depend on μ. This is ensured by using the renormalization group equation,

$$\mu \frac{d\mathcal{M}}{d\mu} = (\mu \frac{\partial}{\partial \mu}|_e + \mu \frac{\partial e}{\partial \mu} \frac{\partial}{\partial e})\mathcal{M} = 0, \tag{1.17}$$

where e is the coupling constant. We see from equation 1.16 that as Q^2 increases, the running coupling constant increases as well. It is useful to rewrite the coefficient of the *log* as

$$\frac{\alpha(\mu^2)}{4\pi}(-\frac{4}{3}). \tag{1.18}$$

The analogous term in the running coupling constant of QCD, α_s, is

$$\frac{\alpha_s(\mu^2)}{4\pi}(-\frac{2}{3}n_f - 5 + 16), \tag{1.19}$$

where n_f is the number of flavors. The first term is the simple gluon loop going into a quark anti-quark pair, and we find an equivalent term in QED when the photon annihilates into a pair of an electron and a positron. In QED, the number of "flavors" is one, and there is a factor of two in the definition of α compared to α_s. The other two terms in the QCD expression result from gluon self-interactions. The consequence is that these contributions have an anti-screening effect: a red gluon attracts other red gluons around it instead of anti-red gluons. If we enter the "cloud of red gluons" we therefore see less and less "redness". Mathematically, since the coefficient is positive in the denominator, as Q^2 increases, α_s decreases. This is referred to as asymptotic freedom. It is interesting to note that at a distance of one fermi, the coupling constant is about one, so that perturbation theory is not possible as a tool to compute QCD reactions. At a momentum transfer of about $Q^2 = (30GeV)^2$, then $\alpha_s \approx 0.1$, which is more amenable to perturbative calculations.

A final difference between QED and QCD is the fact that although particles can carry charge, they can not carry color. This feature of QCD is related to the fact that quarks are confined into hadrons. Quark confinement makes QCD processes rather complicated since we never deal with individual quarks. This will be crucial for our decay of interest.

1.2.4 Electroweak Interaction

In the previous section, we saw how the fact that quarks come in three colors led to a gauge theory involving the $SU(3)$ symmetry. It would be tempting to apply the same trick to quarks and leptons since they come in weak doublets containing two members. Maybe we could get a gauge theory of weak interactions involving the $SU(2)$ symmetry. Yang and Mills developed the mechanism of such a gauge theory, but they were unsuccessful in their attempt to apply such a symmetry to the proton-neutron isospin doublet. One problem associated with making the weak interaction an $SU(2)$ gauge theory is the fact that the bosons carrying the interaction are heavy, and so the range of the force is short (hence the name "weak"). We will now see how an elegant mechanism such as the Higgs mechanism provides the essential tool to cure this problem. First, we will review the effective phenomenology of the weak interactions.

Weak Interaction

Since we are interested in an effective theory we first review all of the experimental evidences that lead to the "guess" for the form of the weak Lagrangian.

As shown in table 1.2 it seems natural to think that all physical processes have to obey some general symmetries: for example the mirror image of any reaction should also happen (Parity, $(p_x, p_y, p_z) \rightarrow (-p_x, -p_y, -p_z)$), or there should not be any difference between a reaction involving particles, from a reaction involving their associated anti-particles (Charge Conjugation).

In 1956, Lee and Yang made a survey of all the weak interaction data. They were particularly puzzled about the decay of a kaon, since this particle seemed to be decaying sometimes to two pions, sometimes to three. This was a concern since these two final states have opposite Parity. For Parity to be a good symmetry, a state of a definite Parity can not change to a state of opposite Parity. Therefore, Parity appeared to be violated by the weak interaction.

To settle the issue of Parity violation Lee and Yang proposed an experiment that was carried out by C.S. Wu. She aligned nuclei of Cobalt 60 so that they would all have a particular direction of nuclear spin. Then she recorded the direction of the electrons emitted when the nucleus underwent β decay. Her results showed that most of the electrons were emitted in the direction of the nuclear spin. For Parity to be conserved in this reaction, the electrons should have been emitted equally in both directions.

Other experimental evidence suggested that Parity was maximally violated in weak decays. To understand this, we need to introduce the definition of helicity, which is the dot product of the spin with the direction of the momentum of the particle. For example, for a particle of spin 1/2, the helicity (λ) can be right handed (when the spin points in the direction of the momentum) or left handed (opposite). Now, let's suppose there is an observer going at a velocity greater than the velocity of the particle, then the direction of the momentum appears to be reversed, so the helicity changes sign (or handedness).

For a massless particle, the situation is special. A massless particle moves at the speed of light and, therefore all observers agree on its velocity. As a result its helicity is fixed. For the case of the photon, this translates into the fact that there is no longitudinal polarization.

The helicity of the neutrino has been determined experimentally. Observations of the $\pi^+ \to \mu^+\nu_\mu$ decay, revealed that the anti-muons emitted were always left handed. It was then inferred that the neutrinos had to be left handed also (the anti-neutrinos are then right handed). This observation would imply that the neutrinos are massless [2].

Depending on how an object reacts to the Parity operator we get the following naming scheme:

$$\begin{aligned} P(s) &= s \\ P(pseudo\, s) &= -pseudo\, s \\ P(v) &= -v \\ P(pseudo\, v) &= v \end{aligned}$$

where s is a scalar and v is a vector. From Quantum Field Theory, we can determine that the parity of a fermion is opposite to the parity of an anti-fermion, while the parity of a boson is the same as the parity of an anti-boson. We have assigned the quark to have an intrinsic parity of +1, so the parity of an anti-quark is -1. For a composite system, the parity is the product of the intrinsic parities. In general for a meson, the parity is given by $P = (-1)^{l+1}$, where l is the orbital angular momentum between the two quarks in the meson.

Continuing our review of experimental evidences that will lead to an effective Lagrangian, we now turn to the Charge Conjugation symmetry. It is interesting

[2] Recent results from neutrino physics involving measurements of solar neutrino oscillations and atmospheric oscillations strongly suggest that neutrinos have mass, although probably so tiny that such an effect would have gone undetected in the π^+ decay experiment. Another possibility is that the mass term is carried by a "sterile" neutrino, which would not have interacted in such an experiment.

to note that only particles that are their own anti-particles are eigenstates of the Charge Conjugation operator (C). Just as for Parity, Charge Conjugation is a multiplicative number. For a meson, the Charge Conjugation number is given by $C = (-1)^{l+s}$, where s is the spin of the meson.

Let's reconsider the $\pi^+ \rightarrow \mu^+\nu_L$ reaction. If we apply Parity then we get the reaction $\pi^+ \rightarrow \mu^+\nu_R$, which does not exist, so weak decays violate parity, as mentioned earlier. If we apply C then we get the reaction $\pi^- \rightarrow \mu^-\bar{\nu_L}$ which does not exist either, so weak decays violate Charge Conjugation as well. Now if we take the product of C and P (CP) we get the reaction $\pi^- \rightarrow \mu^-\bar{\nu_R}$, which does exist, so CP is conserved in this reaction. We will return later to the phenomenon of CP violation.

The above discussion points us at a possible expression for the weak current. For example, if we have a lepton neutrino interaction, involving the exchange of a W^- the current is given by

$$\bar{u}_{\nu_e}\gamma^\mu \frac{1}{2}(1-\gamma^5)u_e, \tag{1.20}$$

where u is the spinor associated with the incoming particle, while $\bar{u}$ is the adjoint spinor associated with the outgoing particle. The spinor u satisfies the Dirac equation (in momentum space),

$$(\gamma^\mu p_\mu - m)u = 0,$$

while the adjoint is given by $\bar{u} \equiv u^\dagger\gamma^0$. The γ^μ are the Dirac matrices

$$\gamma^\mu = (\beta, \beta\alpha)$$
$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$$

,

where the matrices α^i's, and β are given by

$$\alpha^i = \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix} \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix},$$

and where the σ^i are the three usual Pauli matrices. For anti-particles we have a spinor denoted by v and which satisfy the other Dirac equation,

$$(\gamma^\mu p_\mu + m)v = 0.$$

The complete wave function solution is $\psi = u(p)e^{-ip\cdot x}$ and solves the fundamental Dirac equation, involving the hamiltonian H of the interaction in question,

$$H\psi = (\alpha \cdot P + \beta m)\psi. \tag{1.21}$$

The current of equation 1.20 is sometimes referred to as the V-A current, which comes from the different combinations of the wave function with the gamma matrices:

$$\bar{\psi}\psi = scalar$$
$$\bar{\psi}\gamma^\mu\psi = vector(V)$$

$$\begin{aligned}\bar{\psi}\gamma^5\gamma^\mu\psi &= axialvector(pseudovector)(A)\\ \bar{\psi}\gamma^5\psi &= pseudoscalar\end{aligned}$$

.

The next step toward writing an amplitude of a reaction is to find how the the mediating particle interacts with the external particles, this is the propagator. For massive, spin 1 particles, the propagator can be written as,

$$\frac{-i(g^{\mu\nu} - q^\mu q^\nu/M^2)}{q^2 - M^2}, \tag{1.22}$$

where q is the momentum of the boson, of mass M, mediating the interaction. In our decays of interest we have $q^2 \ll M^2$ since the mass of the W boson is around 80 GeV. We can then approximate the propagator as being $ig^{\mu\nu}/M^2$, where $g^{\mu\nu}$ is associated with the weak coupling g_w.

Weak interactions allow for different generations of quarks to interact. If we look at the case of three quarks, u,d, and s, the two possible vertices are: a vertex involving a W^+ going into a u and d quark with strength $cos\theta_c$ and a vertex involving a W^+ going into a u and s quark with strength $sin\theta_c$. In 1963 Cabibbo introduced the angle θ_c, which was measured to be small as expected (13.1°), so that transitions across generations were less likely than within the same generation.

This angle allowed several decay rates to be computed successfully. There was a puzzle though with the case of the decay $K^0_L \to \mu^+\mu^-$. The amplitude for this decay was predicted to be proportional to $sin\theta_c cos\theta_c$; but the measured branching fraction is far less than this value. In 1970, Glashow, Iliopoulos and

Maiani proposed to introduce a fourth quark, c, and this was before the J/ψ (a $c\bar{c} resonance$) was even discovered!. This quark is involved in the loop of the K_L^0 diagram and cancels almost exactly the contribution given by the diagram with a u quark in the loop. This is called the GIM mechanism.

The bigger picture of these $sin\theta_c$ and $cos\theta_c$ factors can be interpreted, by convention, as the fact that the weak bosons couple to the *rotated* states of the lower member of the quark doublets (we could have chosen to rotate the upper member of the doublets),

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}$$

where the weak eigenstates are related to the mass eigenstates by,

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} cos\theta_c & sin\theta_c \\ -sin\theta_c & cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}.$$

Kobayashi and Maskawa generalized the Cabibbo matrix to include the 3^{rd} generation of quark. To get the right strength of a particular amplitude one needs to take the corresponding element of the CKM matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud}V_{us}V_{ub} \\ V_{cd}V_{cs}V_{cb} \\ V_{td}V_{ts}V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}.$$

We will discuss the CKM matrix in detail in the next section.

Electroweak Unification

The reader will notice that in the expression for the Yukawa coupling we have used a different notation for the spinor as compared to equation 1.20. The connection between those two is the projection operator

$$u_L = \frac{1}{2}(1 - \gamma_5)u, \tag{1.23}$$

which projects out the left handed helicity component. This operator finds its usefulness in the process of unifying the weak interaction with the electromagnetic one. At first it seems like those two interactions share more differences than common features. Their relative strength is very different, which comes from the photon being massless and the weak interaction mediators being massive. Finally, the structures of their current also differ: the electromagnetic current is purely vector-like while the weak current contains a vector-like part and an axial vector-like part. Nonetheless, the photon must be involved in the weak interaction, since the W bosons are themselves charged. If the electromagnetic current is to be conserved, there must be a $\gamma W^+ W^-$ coupling term somewhere ([6]).

We hinted that the solution for making the structure of the weak current look more vector-like was to use the projection operator,

$$j_\mu^- = \bar{u}^{\nu_e}\gamma^\mu \frac{1}{2}(1 - \gamma^5)u^e = \bar{u}_L^{\nu_e}\gamma^\mu u_L^e. \tag{1.24}$$

The u_L spinors are known as chiral spinors. The electromagnetic current can also be written in terms of those chiral spinors:

$$j_\mu^{em} = -\bar{u}_L^e \gamma^\mu u_L^e - \bar{u}_R^e \gamma^\mu u_R^e. \tag{1.25}$$

The weak current can be viewed as involving a left-handed doublet,

$$\chi_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L ,$$

so that the charged weak currents become

$$j_\mu^\pm = \bar{\chi}_L \gamma_\mu \tau^\pm \chi_L, \tag{1.26}$$

where the τ matrices are related to the first two Pauli matrices,

$$\tau^\pm = \frac{1}{2}(\tau^1 \pm i\tau^2). \tag{1.27}$$

This structure is similar to the isospin structure used to describe the proton and neutron as two states of the nucleon, which implies that the current just defined is invariant under the $SU(2)$ symmetry. To complete the $SU(2)$ structure we need a third invariant weak current,

$$\begin{aligned} j_\mu^3 &= \bar{\chi}_L \gamma_\mu \frac{1}{2} \tau^3 \chi_L \\ &= \frac{1}{2} \bar{u}_L^{\nu_e} \gamma^\mu u_L^{\nu_e} - \frac{1}{2} \bar{u}_L^e \gamma^\mu u_L^e. \end{aligned}$$

It would be tempting to associate this current with the weak neutral current, but the Z boson also couples to right-handed spinors. We can push the analogy with the isospin system and make use of the hypercharge,

$$Y = 2(Q - I_3), \tag{1.28}$$

where Q is the electric charge and I_3 is the third component of isospin. The from for Y inspires a definition for a weak hypercharge current that involves both right-handed and left-handed spinors:

$$\begin{aligned} j_\mu^Y &= 2j_\mu^{em} - 2j_\mu^3 \\ &= -2\bar{u}_R^e \gamma^\mu u_R^e - \bar{u}_L^e \gamma^\mu u_L^e - \bar{u}_L^{\nu_e} \gamma^\mu u_L^{\nu_e}. \end{aligned}$$

This current is invariant under the combined symmetry $SU(2)_L \times U(1)_Y$. In summary, we have combined QED and the weak interaction, so that the electroweak currents are given by,

$$\vec{j}_\mu = \frac{1}{2}\bar{\chi}_L \gamma_\mu \vec{\tau} \chi_L \tag{1.29}$$

$$j_\mu^Y = 2j_\mu^{em} - 2j_\mu^3. \tag{1.30}$$

The unification is not perfect since we still have two groups each with an independent coupling, so the idea is to relate those two couplings.

The last piece missing in this electroweak theory is the identification of a propagator of the interaction, responsible for the currents. The structure of the $SU(2)_L \times U(1)_Y$ symmetry leads to a basic electroweak interaction of the form,

$$-igj_\mu \dot{W}^\mu + \frac{g'}{2} j_\mu^Y B^\mu, \tag{1.31}$$

which involves a massless triplet of vector bosons W_μ and a massless singlet boson B_μ. The question is then how to relate those vector bosons to the physical states that are the massless photon and the massive $W^\pm$ and Z^0? The answer is that the symmetry is broken via the Higgs mechanism. The next sections describes how this process results in the physical massive bosons.

The Higgs Mechanism

The thread we have been following so far has been how symmetries of Nature shape the various fundamental interactions. From a field theory point of view, there are two conditions that determine whether we have an exact symmetry ([2]):

- The Lagrangian is invariant under the symmetry (we have seen several examples of this in previous sections)
- The unique physical vacuum is also invariant under the symmetry

There are two types of situations that can spoil the exact symmetry case. First, if we do not have an exact symmetry, but the symmetry breaking effect is small, then all is not lost and we can write the Lagrangian as:

$$\mathcal{L} = \mathcal{L}_{symm} + \epsilon\mathcal{L}_{symm.break.} \tag{1.32}$$

A good example of this situation is

$$\mathcal{L} = \mathcal{L}_{strong} + \mathcal{L}_{EM}. \tag{1.33}$$

The strong interaction is invariant under the isospin symmetry, and the effect of isospin violation is due to the eletromagnetic interaction. The second situation arises when we do have an exact symmetry of the Lagrangian but the dynamics are such that the vacuum states are not invariant under the symmetry. In this case, we talk about spontaneous symmetry breaking.

For a very illustrative example of this last case, we look at a continuous symmetry breaking ([3]). We take the following Lagrangian, dependent on two real fields ϕ_1 and ϕ_2:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi_1)(\partial^\mu \phi_1) + \frac{1}{2}(\partial_\mu \phi_2)(\partial^\mu \phi_2) + \frac{1}{2}\mu^2(\phi_1^2 + \phi_2^2) - \frac{1}{4}\lambda^2(\phi_1^2 + \phi_2^2)^2. \quad (1.34)$$

This Lagrangian is invariant under rotation in $\phi_1\phi_2$ space ($SO(2)$). We can write the potential of this Lagrangian as,

$$V = -\frac{1}{2}\mu^2(\phi_1^2 + \phi_2^2) + \frac{1}{4}\lambda^2(\phi_1^2 + \phi_2^2)^2, \quad (1.35)$$

so that we can see that the minimum of this potential is a circle of radius μ/λ (this is the wine bottle potential) given by,

$$\phi_1^2|_0 + \phi_2^2|_0 = \frac{\mu^2}{\lambda^2}. \quad (1.36)$$

In this interesting situation, the vacuum state does not lie at the zero of the potential, as is usual. We pick an arbitrary vacuum state around which the Feynman rules can be used,

$$\phi_1|_0 = \frac{\mu}{\lambda}, \phi_2|_0 = 0. \quad (1.37)$$

A judicious change of basis allow the new fields to have their ground state lying on the 0 of the potential,

$$\eta \equiv \phi_1 - \frac{\mu}{\lambda}, \xi \equiv \phi_2. \tag{1.38}$$

The Lagrangian in terms of the new fields now reads,

$$\mathcal{L} = [\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2 \eta^2] + [\frac{1}{2}(\partial_\mu \xi)(\partial^\mu \xi)] + \dots . \tag{1.39}$$

The Lagrangian in the new basis has lost its $SO(2)$ symmetry (it's hidden), but now the spectrum of particles is apparent. The mass term of the η field can be read off as $\approx \sqrt(2)\mu$, the ξ field is massless, and the other terms are interaction terms between the fields. According to the Goldstone theorem, for every continuous symmetry breaking there is the appearance of a massless scalar field, called a Goldstone boson. At this point it is not clear how this mechanism can help the interaction fields acquire mass, since it introduces even more massless particles. We will see that there is a delightful interplay between our good friend the local gauge invariance and the Higgs mechanism that does all the magic.

The Higgs mechanism in the Abelian case

In order to see how the magic happens in a simple case, we look at the case where the local gauge symmetry is Abelian. Taking our last example of the two real fields, we combine them into one complex field, $\phi = \phi_1 + i\phi_2$,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^*(\partial_\mu \phi) + \frac{1}{2}\mu^2(\phi^* \phi) - \frac{1}{4}\lambda^2(\phi^* \phi)^2. \tag{1.40}$$

The $SO(2)$ is replaced with the now familiar $U(1)$ symmetry,

$$\begin{aligned}
\mathcal{D} &\equiv \partial_\mu + iqA_\mu \\
F_{\mu\nu} &\equiv \partial_\nu A_\mu - \partial_\mu A_\nu \\
\phi \rightarrow \phi' &= e^{iq\alpha(x)}\phi(x) \\
A_\mu \rightarrow A'_\mu &= A_\mu - \partial_\mu \alpha(x).
\end{aligned}$$

The invariant Lagrangian is then,

$$\mathcal{L} = \frac{1}{2}\mathcal{D}_\mu \phi^* \mathcal{D}_\mu \phi + \frac{1}{2}\mu^2(\phi^*\phi) - \frac{1}{4}\lambda^2(\phi^*\phi)^2 - \frac{1}{16\pi}F_{\mu\nu}F^{\mu\nu}. \tag{1.41}$$

We make the same change of basis around the same particular ground state that we chose in the previous example, so that the Lagrangian in the new basis looks like,

$$\begin{aligned}
\mathcal{L} &= [\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2\eta^2] + [\frac{1}{2}(\partial_\mu \xi)(\partial^\mu \xi)] \\
&\quad +[-\frac{1}{16\pi}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}(q\frac{\mu}{\lambda})^2 A_\mu A^\mu - 2iq\frac{\mu}{\lambda}\partial_\mu \xi A^\mu] + \ldots
\end{aligned}$$

As before, the η field has acquired mass, and we have a massless ξ field, but now the field A_μ has also acquired some mass! Unfortunately there is also a suspicious looking term that involves the A_μ and the ξ fields. This is where the beauty of having an invariant Lagrangian comes into play, since we only have to use a judicious choice of gauge and we can still use the form of equation 1.41. To make more apparent the choice of gauge, we can rewrite the offending terms as followed,

$$\frac{1}{2}(q\frac{\mu}{\lambda})^2(A_\mu + \frac{1}{2iq\frac{\mu}{\lambda}}\partial_\mu\xi)(A^\mu + \frac{1}{2iq\frac{\mu}{\lambda}}\partial^\mu\xi). \tag{1.42}$$

We choose the following gauge,

$$A_\mu \to A'_\mu = A_\mu + \frac{1}{2iq\frac{\mu}{\lambda}}\partial_\mu\xi, \tag{1.43}$$

which corresponds to a particular phase of the field,

$$\phi \to \phi' = e^{i\frac{\lambda}{2\mu}\xi}\phi. \tag{1.44}$$

This gauge forces the imaginary part of ϕ to 0. This particular gauge, which involves having only physical states in the Lagrangian, is called the unitary gauge (U-gauge). We can use the same equation 1.41 in the new gauge,

$$\mathcal{L} = [\frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \mu^2\eta^2] - \frac{1}{16\pi}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}(q\frac{\mu}{\lambda})^2 A_\mu A^\mu. \tag{1.45}$$

We see explicitly that the η field still has mass, this is the Higgs boson, the vector field has acquired mass, and we do not have any ξ field anymore. It is interesting to note that before the symmetry breaking happened, we had four particle degrees of freedom: two scalar fields and two polarizations of the massless gauge vector A_μ. After the symmetry breaking we are left with one scalar field and three polarizations of the massive vector fields, so still four particle degrees of freedom. This is where the expression "the vector boson (the A_μ) eats the Goldstone boson (the ξ) to acquire mass" comes from.

The Higgs mechanism applied to $SU(2)_L \times U(1)_Y$

Following [2], we start by just looking at the leptonic part of applying the Higgs mechanism to $SU(2)_L \times U(1)_Y$. The $SU(2)_L$ symmetry has three massless gauge bosons associated with it $b^1_\mu, b^2_\mu, b^3_\mu$, while the $U(1)_Y$ has only one, $\mathcal{A}_\mu$. The Lagrangian has a part associated with the gauge fields and a part associated with the leptons. The gauge part has a kinetic term written in terms of the field-strength tensors, which we've defined in previous sections for both the non-Abelian case of $SU(2)$ and the Abelian case of $U(1)$. We can write the lepton part of the Lagrangian as,

$$\mathcal{L}_{leptons} = \bar{u}_R i\gamma^\mu(\partial_\mu + \frac{ig'}{2}\mathcal{A}_\mu Y)u_R + \bar{u}_L i\gamma^\mu(\partial_\mu + \frac{ig'}{2}\mathcal{A}_\mu Y + \frac{ig}{2}\tau \dot{b}_\mu)u_L, \tag{1.46}$$

where g and $\frac{g'}{2}$ are the couplings for $SU(2)_L$ and $U(1)_Y$ respectively. We now use the Higgs mechanism to render mass to the gauge bosons. We introduce a doublet of complex scalar fields,

$$\phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}.$$

The new part of the Lagrangian corresponding to this doublet of fields is

$$\mathcal{L}_{scalar} = (\mathcal{D}^\mu\phi)^\dagger(\mathcal{D}_\mu\phi) - V(\phi^\dagger\phi). \tag{1.47}$$

The covariant derivative is defined as usual involving the gauge bosons. We take the potential to be the same as we've used in the previous section for the continuous symmetry. This is a Standard Model assumption: until we discover the precise mechanism of symmetry breaking of the electroweak system, we will not know what the potential is,

$$V(\phi^\dagger\phi) = \mu^2(\phi^\dagger\phi) + |\lambda|(\phi^\dagger\phi)^2. \tag{1.48}$$

We can also add to the Lagrangian an interaction term that involves the scalar doublet and the leptons. The simplest coupling symmetric under $SU(2)_L \times U(1)_Y$ is given by

$$\mathcal{L}_{Yukawa} = -G_e(\bar{u}_R\phi^\dagger u_L + \bar{u}_L\phi u_R). \tag{1.49}$$

We can now apply the spontaneous symmetry breaking process by noting that the minimum of the potential is given by,

$$\frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda}. \tag{1.50}$$

Thus we can choose the following ground states,

$$\phi_1 = \phi_2 = \phi_4 = 0, \phi_3^2 = -\frac{\mu^2}{\lambda} \equiv v^2, \tag{1.51}$$

which can be rewriten as:

$$< \phi >_0 = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}.$$

We have said in the previous section that the Goldstone theorem states that for every symmetry breaking there will be the creation of a Goldstone boson (which is going to be massless and be eaten up to give mass to the gauge boson). In a more mathematical sense this is equivalent to saying that there will be a Goldstone boson created if a particular generator of the symmetry group does not leave the vacuum invariant, that is, if $G < \phi >_0 \neq 0$. The generators of $SU(2)$ are the τ matrices and applying each of them to $< \phi >_0$ does not give 0. The generator of $U(1)_Y$ is a number Y, so applying it to $< \phi >_0$ does not give 0 either. It is interesting to note that if we apply the electric charge as the generator, which is a mix of τ_3 and Y according to the Gell-Mann-Nishijima relation, then we get,

$$Q < \phi >_0 = \frac{1}{2}(\tau_3 + Y) < \phi >_0 = \frac{1}{2}\begin{pmatrix} 0 \\ -\frac{v}{\sqrt{2}} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} = 0.$$

Hence the charge operator leaves the vacuum invariant. As a result, the photon will remain massless while the three other gauge bosons will eat up their respective Goldstone bosons and acquire mass.

The next step would be to rewrite the Lagrangian with the new fields defined around the minimum of the potential. We know that by choosing the U-gauge we ensure that the physical spectrum of particles is apparent:

$$\phi \to \phi' = \begin{pmatrix} 0 \\ \frac{v+\eta}{\sqrt{2}} \end{pmatrix}.$$

The Yukawa term now reads

$$\mathcal{L}_{Yukawa} = -\frac{G_e v}{\sqrt{2}}\bar{u}u - \frac{G_e \eta}{\sqrt{2}}\bar{u}u. \tag{1.52}$$

The first term can be interpreted as the mass of the leptons, while the second term is an interaction term between the leptons and the η field. The scalar term becomes

$$\mathcal{L}_{scalar} = \frac{1}{2}(\partial^\mu \eta)(\partial_\mu \eta) - \mu^2\eta^2$$
$$+\frac{v^2}{8}[g^2|b^1_\mu - ib^2_\mu|^2 + (g'\mathcal{A})\mu - gb^3_\mu)^2] + \ldots.$$

We can read off the mass of the η field to be $M^2 = -2\mu^2$. The η field is the Higgs boson of the Standard Model.

The charged gauge fields can be defined as

$$W^\pm_\mu \equiv \frac{b^1_\mu \mp ib^2_\mu}{\sqrt{2}}. \tag{1.53}$$

With this definition the mass term of the charged gauge bosons become manifest, since now we have

$$\frac{g^2v^2}{8}(|W^+_\mu|^2 + |W^-_\mu|^2), \tag{1.54}$$

and hence $M^\pm_W = \frac{gv}{\sqrt{2}}$. We can also write down the part of the Lagrangian describing the interactions between the vector bosons and the leptons and make a comparison with the phenomenological Lagrangian found in previous sections for the Electroweak interaction. The connection between the two gives the relationship between g and the coupling found experimentally,

$$\frac{g^2}{8} = G_F\frac{M^2_W}{\sqrt{2}}. \tag{1.55}$$

This gives us a value for v, the vacuum expectation value of the Higgs field,

$$v = (G_F\sqrt{2})^{-1/2} = 174\,\mathrm{GeV}. \tag{1.56}$$

We can also define the neutral boson fields as

$$Z_\mu = \frac{-g'\mathcal{A}_\mu + g b^3_\mu}{\sqrt{g^2+g'^2}}, \tag{1.57}$$

$$A_\mu = \frac{g\mathcal{A}_\mu + g' b^3_\mu}{\sqrt{g^2+g'^2}}. \tag{1.58}$$

The masses of the Z^0 particle, $M_Z^0 = \sqrt{g^2+g'^2}\,\frac{v}{2}$ and the photon, $m_\gamma = 0$, become apparent. We see that the neutral bosons are mixed. It is useful to introduce a weak mixing angle such that $g' = g\tan\theta_W$ so that now we have

$$Z_\mu = -\mathcal{A}_\mu sin\theta_W + b^3_\mu cos\theta_W \tag{1.59}$$

$$A_\mu = \mathcal{A}_\mu cos\theta_W + b^3_\mu sin\theta_W \tag{1.60}$$

The unification of the weak and electromagnetic interactions is now complete since we can relate each coupling constant to each other, with a relationship that is found by making the connection between the Lagrangian of the neutral bosons and their phenomenological counter part:

$$g = \frac{e}{sin\theta_W}, g' = \frac{e}{cos\theta_W}. \tag{1.61}$$

1.2.5 The CKM matrix

We come back to the CKM matrix introduced in section 1.2.4. The Standard Model does not predict the values of the CKM elements; we need to measure them. The matrix is unitary by construction. This constraint along with the fact that phase differences are unphysical leave four independent real parameters, for three generations of quarks. These correspond to three angles and one phase factor, so the CKM matrix contains complex elements. The PDG uses the set of angles $\theta_{12}, \theta_{13}, \theta_{23}$ and the phase factor δ_{13}.

Although the Standard Model does not predict the individual elements, some information can be extracted from the CKM matrix when written in a useful form. To do this we can use the approximation that $cos\theta_{13}$ is very close to unity ($V_{ub} = sin\theta_{13}e^{-i\delta_{13}}$ is very small). Also, we can expand all the cos and sin terms. Wolfenstein chose the expansion parameter to be $\lambda = sin\theta_C \approx 0.22$. The CKM matrix can be written, using the four independent parameters, A, λ, ρ and η as

$$V = \begin{pmatrix} 1 - 1/2\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - 1/2\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4).$$

If we examine the magnitudes of the different elements we notice that the bigger the step in generation, the weaker the element: V_{ud}, V_{cs}, V_{tb} are of order 1, while the elements involving the 1^{st} and 2^{nd} generation, V_{us}, V_{cd}, are of order λ, the elements involving the 2^{nd} and 3^{rd} generation, V_{cb}, V_{ts}, are of order λ^2 and finally the last level (and the hardest to measure) involves the elements from the 1^{st} and 3^{rd} generation, V_{ub}, V_{td}, which are of order λ^3. Table 1.3 summarizes the latest results and signatures for each element: ([7],[8],[9]).

Using the unitarity constraint ($V^\dagger V = 1$) we can write down six equations relating the different CKM elements. The most interesting one involves combina-

Table 1.3: Measuring the CKM elements

CKM	Experimental Signature	Measurement
$\|V_{ud}\|$	nuclear beta decays	0.9735±0.0008
	neutron decays	
$\|V_{cs}\|$	$K \to \pi e \nu_e$ with $D \to K e \nu_e$ decays	1.04±0.16
	Charmed-tagged W decays	
$\|V_{tb}\|$	$t \to b\ell\nu$ decays	$\frac{\|V_{tb}\|^2}{V_{td}\|^2+V_{ts}\|^2+V_{tb}\|^2} = 0.99 \pm 0.29$
$\|V_{us}\|$	$K \to \pi e \nu_e$ decays	0.2196±0.0023
	hyperon decays	
$\|V_{cd}\|$	$\nu_\mu d \to \mu c$ decays	0.224±0.016
$\|V_{cb}\|$	$B \to D^* \ell\nu$ decays	CLEO: 0.0462±0.0032
	Inclusive $b \to c\ell\nu$	
$\|V_{ts}\|$	$b \to s\gamma$ decays	$\frac{\|V_{ts}^* V_{tb}\|^2}{\|V_{cb}\|^2} = 0.95(1+0.01\ \rho)$
$\|V_{td}\|$	B_d, B_s mixing	$\|V_{tb}^* V_{td}\|$=0.0083±0.0016
$\|V_{ub}\|$	$B \to \pi, \rho, \ell\nu$ **decays**	**this thesis**
	Inclusive $b \to u\ell\nu$	CLEO: 0.00408 ± 0.00063

tions of similar order and is the product of the 3^{rd} row of $V^\dagger$ and the 1^{st} column of V:

$$V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0. \tag{1.62}$$

If we set the different elements which are of order unity to one, and if we divide each members by $V_{cb}^* V_{cd}$ then we get the triangle of base length one, shown in figure 1.5.

From the previous discussion about the experimental signatures of the different elements, it is clear that only the processes $b \to u\ell\nu$, $b \to c\ell\nu$ and $B^0 - \bar{B}^0$ mixing are necessary to completely determine the unitary triangle. The problem is that those reactions involve large hadronic contributions which results in large uncertainties on the different elements.

The hope is that the angles α, β, and γ can be extracted with reduced hadronic uncertainties. The angle β has been determined experimentally by the new B factories, BaBar and BELLE. They essentially looked at the process $B \to J/\psi K_s$, which is very clean both theoretically and experimentally. The current average value is $sin 2\beta = 0.77 \pm 0.08$ [5]. To determine the angle α the B factories will have to perform what is called an isospin analysis. Finally, the angle γ is even harder to measure experimentally. There are various methods proposed to measure this angle ([9]). One of these method is to build triangles relating various decay amplitudes to each other.

The point of measuring all the various pieces of the unitary triangle independently is of course to see if the triangle closes, and see if there is some new physics

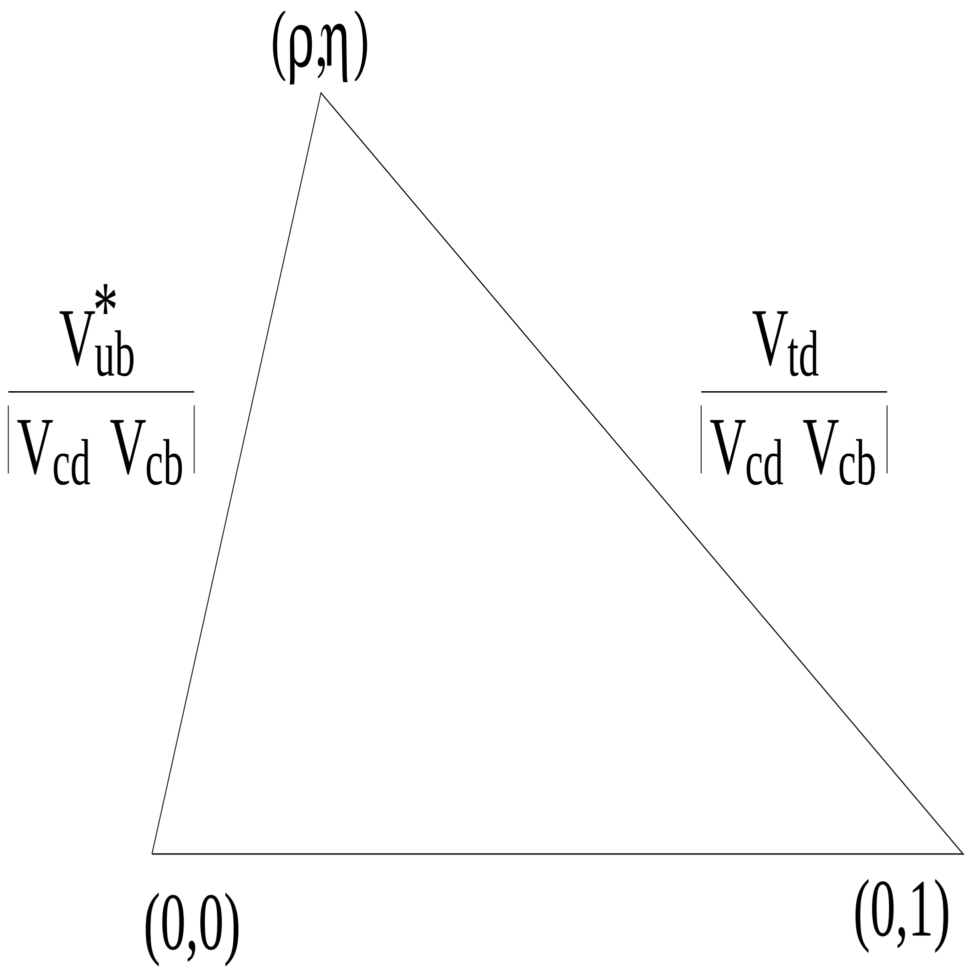

Figure 1.5: Unitarity triangle built from the CKM matrix.

beyond the Standard Model that shows up in one of those measurements. Since this type of new physics hunt is indirect (we are not looking for bumps indicating new particles), there is usually no single clean measurement that can clearly show the new effect. Different processes that would be controlled by the same combinations of CKM elements can have different non-Standard Model contributions. Extracting CKM quantities assuming the Standard Model would then result in inconsistent parameters.

1.2.6 CP violation

We have seen in section 1.2.4 that the V-A structure of the weak interaction produces Parity violation and Charge Conjugation violation, but leaves CP invariant. Here we study how CP might be violated.

If we look at the the following quark scattering process: ab $\rightarrow$ cd, then the amplitude is proportional to,

$$\mathcal{M} \propto V_{ca} V_{db}^{*} (\bar{u_c} \gamma^{\mu} (1 - \gamma^5) u_a)(\bar{u_d} \gamma^{\mu} (1 - \gamma^5) u_b).$$

If we apply the CP operator on the reaction we get the process $\bar{a}\bar{b} \rightarrow \bar{c}\bar{d}$, for which the amplitude can be written as,

$$\mathcal{M}' \propto V_{ca}^{*} V_{db} (\bar{u_a} \gamma^{\mu} (1 - \gamma^5) u_c)(\bar{u_b} \gamma^{\mu} (1 - \gamma^5) u_d). \tag{1.63}$$

Since an Hamiltonian is always hermitian we have $\mathcal{M}' = \mathcal{M}^{\dagger}$. For CP invariance to hold, application of the operator CP on the $\mathcal{M}$ amplitude must result in $\mathcal{M}'$.

We need to apply the operator CP to the different currents. The end result is that we find the amplitude to be ([4])

$$CP\mathcal{M} \propto V_{ca}V_{db}^{*}(\bar{u_a}\gamma^{\mu}(1-\gamma^5)u_c)(\bar{u_b}\gamma^{\mu}(1-\gamma^5)u_d).$$

This amplitude is the same as the one in equation 1.63, except for the position of the "*" on the different CKM elements. If the CKM matrix were real, then we would have CP invariance. Because there is a complex number in the CKM matrix, the SM predicts some CP violation.

The unitarity triangle shown in figure is related to CP violation through the fact that its area J (for Jarlskog), quantifies the amount of CP violation because it is proportional to the imaginary part of the CKM matrix: $J \approx \lambda^6 A^2 \eta$. The other two triangles that we could have chosen also have their area equal to J. One triangle involves K decays, the other B_s decays. We mostly hear about the B_d triangle because it has all sides of roughly the same size, which makes it easier experimentally to measure all the sides and angles.

The study of CP violation is a very hot subject of research within HEP these days. The reasons for this are that it's the least tested aspect of the Standard Model. Also most New Physics models predict additional sources of CP violation. Finally, our own existence implies that there was some amount of CP violation shortly after the Big Bang, but baryogenesis requires a larger level of CP violation than is predicted by the Standard Model.

There are three types of CP violation in meson decays: CP violation in the mixing of the neutral mesons, CP violation in the decay of either the neutral or charged mesons and finally, CP violation in the interference of decays of the neutral mesons with and without mixing.

One way to measure CP violation in mixing is to look at semileptonic asymmetries of the neutral B meson,

$$a_{SL} = \frac{\Gamma(\bar{B}^0(t) \to X\ell^+\nu) - \Gamma(B^0(t) \to X\ell^-\nu)}{\Gamma(\bar{B}^0(t) \to X\ell^+\nu) + \Gamma(B^0(t) \to X\ell^-\nu)}. \tag{1.64}$$

CLEO has a measurement of such an asymmetry, and the world average value is ([5]) $a_{SL} = (0.2 \pm 1.4) \times 10^{-2}$. Unfortunately, the theoretical predictions for this quantity involve large QCD effects, but we do know that the asymmetry can not be larger than $\mathcal{O}(\frac{\Delta\Gamma_B}{\Delta M_B})$, where $\Delta\Gamma_B$ is the difference in width between the two neutral B mesons, while ΔM_B is their mass difference. The current precision of the measurements is still not high enough to either validate or refute the Standard Model. This type of CP violation is sometimes also referred to as indirect CP violation.

To measure CP violation in decay, one can look at the asymmetry involving charged final states,

$$a_{f^\pm} = \frac{\Gamma(B^+ \to f^+) - \Gamma(B^- \to f^-)}{\Gamma(B^+ \to f^+) + \Gamma(B^- \to f^-)}. \tag{1.65}$$

If we choose two amplitudes such that $A_2 << A_1$ then the theoretical expression for the asymmetry can be written as

$$a_{f^\pm} = -2(A_2/A_1) sin(\delta_2 - \delta_1) sin(\phi_2 - \phi_1), \tag{1.66}$$

where δ_2, δ_1 are strong phases, and ϕ_2, ϕ_1 are weak phases. Unfortunately, this asymmetry also suffers from large theoretical uncertainties since it involves the strong phases, which usually depend on QCD effects that are hard to calculate, like final states interactions and hadronization. This type of CP violation is sometimes referred to as direct CP violation.

There is hope, though, concerning the third type of CP violation. For this kind, one can measure the time dependence of the asymmetry involving CP eigenstate final modes,

$$a_{f_{CP}} = \frac{\Gamma(\bar{B}^0(t) \to f_{CP}) - \Gamma(B^0(t) \to f_{CP})}{\Gamma(\bar{B}^0(t) \to f_{CP}) + \Gamma(B^0(t) \to f_{CP})}. \tag{1.67}$$

Since this measurement requires the time dependence of the B decay, one gains much leverage by building an asymmetric machine, leading to a displaced vertex of one B relative to the other. As we have mentioned in the previous section, the B factories have measured the CP eigenstate ψK_S decay mode, for which the asymmetry can be written as

$$a_{f_{CP}} = sin(2\beta) sin(\Delta m_B t). \tag{1.68}$$

The magnitude of the asymmetry is given by $sin(2\beta)$ and it is significantly different than zero. This is the first direct observation of CP violation in the B system. It is also the first precision test of the Standard Model concerning CP violation, and we're happy to report that it passed.

1.3 Beyond the Standard Model

We have just gone through a quick overview of how the Standard Model is a possible ground on which to grow fundamental laws governing elementary particles and how they interact. Although we know that it can't be the end of the story. For example, as mentioned before, our own existence implies that there is more CP violation than what the Standard Model predicts. We now briefly discuss the main reasons why the Standard Model is unsatisfactory and also look into possible extensions.

Modern neutrino physics has already provided experimental hints that the massless neutrinos predicted by the Standard Model need to be remodeled. We now review how the massless neutrino is predicted by the Standard Model. Coming back to the lepton current of equation 1.20, the leptons have a Yukawa-type coupling with the Higgs field,

$$\mathcal{L}_{Yukawa} = Y^e_{ij}\bar{u}^i_L \phi u^j_R, \tag{1.69}$$

where the i, j refer to the flavors. This coupling eventually leads to a mass term for the charged lepton and predicts that neutrinos are massless. The simplest way ([5]) to account for some neutrino mass is to add a dimension-five term that makes the neutrino mass explicit,

$$\mathcal{L}^{dim-5}_{Yukawa} = \frac{Y^\nu_{ij}}{M}\bar{u}^i_L u^j_L \phi\phi, \tag{1.70}$$

where the parameter M is a mass. This term is nonrenormalizable and implies that this extended version of the lepton sector of the Standard Model is valid only until the scale M. Because of the structure of this new dimension-five term, the couplings Y^ν are symmetric. The mass terms for the leptons are now given by ([5]),

$$M_e = \frac{v}{\sqrt{2}} Y^e, M_\nu = \frac{v^2}{2M} Y^\nu. \tag{1.71}$$

In complete analogy with the quark sector, we can always find unitary matrices $V_{eL} V_\nu$ such that

$$\begin{aligned} V_{eL} M_e M_e^\dagger V_{eL}^\dagger &= diag(m_e^2, m_\mu^2, m_\tau^2) \\ V_\nu M_\nu M_\nu^\dagger V_\nu^\dagger &= diag(m_1^2, m_2^2, m_3^2). \end{aligned}$$

In the mass basis, the interaction between the $W^\pm$ and the lepton is given by

$$\mathcal{L}^\ell_{W^\pm} = \frac{g}{\sqrt{2}} \bar{u}_{\ell L} \gamma^\mu (V_{eL} V_\nu^\dagger) u_{\nu L} W_\mu^\pm. \tag{1.72}$$

The unitary 3×3 matrix $V_{MNS} = V_{eL} V_\nu^\dagger$ (the Maki-Nakagawa-Sakata matrix) is the mixing matrix for the leptons just like the CKM matrix is the mixing matrix for the quarks. Because of the Majorana nature of the neutrinos in equation 1.70, we are not free to make a phase transformation to the neutrino fields, in contrast to the situation with the charged leptons and the quarks. As a result, the MNS matrix has three CP violating phases, as opposed to the single one in the CKM matrix. Because of this new MNS matrix, we now have flavor changing interaction in the lepton sector, making it similar to the quark sector.

None of the Yukawa couplings are predicted by the SM; we have to measure all the masses of the fundamental particles. This touches on another shortcoming of the SM: the fact that there is such a difference in scale between the lightest and heaviest fundamental particle. Just among the quarks we go from ~ 0.004 GeV for the up quark to $\sim$170 GeV for the top quark. Actually, the top quark is the only one that has the right mass scale: we have seen that the Higgs mechanism provides the mass of the quarks, leptons and gauge bosons of the Electroweak interactions via its vacuum expectation value (vev). This value is about 100 GeV, so we would expect all the masses to be around that value, which is only the case for the top quark and the W and Z bosons. Another issue connected with this is the mass of the Higgs particle itself: since it's a scalar field we would expect that it acquires mass through all of its quantum effects. This means that its mass should be around 10^{16} GeV, clearly not the 100 GeV that seems to emerge from all the Electroweak precision data. This is known as the Hierarchy Problem.

Another unnatural feature of the SM is related to what is called the strong CP problem. In the SM Lagrangian, there are nonperturbative QCD terms that violate CP and induce an electric dipole to the neutron. We expect a relationship like ([5]):

$$d_N \approx 5 \times 10^{-16} \theta_{QCD}\, e\, cm \tag{1.73}$$

The experimental bound on d_N is less than $6.3 \times 10^{-26} e\,cm$, so θ_{QCD} should be less than 10^{-10}. The theoretical prediction for this quantity is that it should be of order one, so the small measured value seems to indicate possible new physics.

The main issue with the Standard Model seems to be the hierarchy problem. Fundamentally, there are two ways to solve this problem: 1) there can be some careful fine tuning and cancelations that take place to prevent the quantum corrections that drive up the Higgs mass, or 2) the Planck scale is the same as the Electroweak scale.

If you were to take a poll among theorists about their favorite contender for possible physics beyond the Electroweak scale, the vast majority would answer Supersymmetry (SUSY). This is particularly interesting in view of the fact that aside from the lack of any experimental observation of SUSY parameters, there needs to be several versions of SUSY that fix this or that theoretical shortcoming. Also there is a multiplication of unpredicted parameters that one has to measure with SUSY models, even though we already feel uncomfortable with the 18 parameters from the SM. Nevertheless, SUSY elegantly solves the fine tuning of the SM and it looks like it's a necessary part of any Grand Unified Theory, where the electroweak and strong interactions come together.

In 1882 Darbaux factored second order differential operators into the product of two first order operators. This is the most basic implementation of the idea of supersymmetry ([6]). The same idea was used in Dirac's formulation of the harmonic oscillator writing it as $H = h\omega(a^+a = \frac{1}{2})$. Since the Hamiltonian is the generator of time translations, we can say that we've decomposed time translations into more fundamental operations. In SUSY, we have both spatial and time translations embodied in a super algebra. For example, under the group $SU(2)\times SU(2)$, operators transform under (1/2,1/2). To apply the super algebra is to use more fundamental objects which transform under (1/2,0) and (0,1/2). Physically speaking this translates into postulating that there are fermion and boson multiplets: for every fermion there is a bosonic partner (sleptons, squarks) and for every boson there is a fermionic partner (photinos, gluinos). The fine tuning of the hierarchy problem happens because for each offensive fermionic loop in the quantum corrections of the Higgs field, there is an equal and opposite bosonic loop that cancels it ([11]).

It is interesting to note that the spectrum of sparticles is decoupled from the particle spectrum, so that the physics at the EW scale is consistent with the SM predictions (and hence with observations). Even the supersymmetry Higgs sector can be accommodated with a SM-like Higgs around 100 GeV.

Another possible solution to the hierarchy problem is to postulate the existence of extra spatial dimensions. Those extra dimensions would be compactified so that there would be no direct evidence of them in our everyday world. The relation between how the size of these extra dimensions influences the constants at the Planck scale is given by ([11])

$$M_{Planck}^{4d} = (M_{Planck}^{(4+n)d})^{\frac{n+2}{2}} R^{\frac{n}{2}}, \tag{1.74}$$

where R^n is the volume of the compactified n-dimensional space. This line of thought is an appealing candidate for developing quantum gravity, since in this model only the graviton would be able to radiate into the "bulk" (the extra dimensional volume), making its strength as feeble as we experience it. It's interesting to see that such a model has exotic experimental predictions for the next generation of colliders such as the production of black holes.

Finally, a less popular view is to consider the Higgs particle as a composite particle, so that it would be protected from quantum corrections. Technicolor models have problems in that their predictions have difficulties standing up to the EW precision measurements.

In summary, we have seen that the CKM matrix is a cornerstone in the complete understanding of fundamental particles and their interactions. Furthermore, precise kinematic observations of decays such as the transition of a b quark to a u quark semileptonically can be used as a probe of the QCD physics involved, leading to much improved tools to help interpret more accurately a whole range of crucial measurements. In the next chapter, we look more in detail at the dynamic of such a decay and the different tools that are used in the strong interaction part of the decay.

CHAPTER 2

WHAT WE KNOW ABOUT THE $B \to X_U \ell\nu$ DECAYS

2.1 Dynamics of the decays

Now that we have layed out the backgroud picture for the weak interations and specially the role of the CKM matrix, we can turn to looking more specifically to the exclusive decays $B \to \pi\rho\omega\eta\ell\nu$. We will first look at their amplitudes and decay rates and then we will look at what sort of information we can gain from analyzing their Dalitz plot.

2.1.1 Amplitudes and decay rates

We are interested in deriving an expression for the differential decay rate and the width of the exclusive decays involved in the transition $b \to u\ell\nu$. These expressions lead to the determination of V_{ub}, once we have the experimental results. We show the complete derivation for the case of $B \to \pi\ell\nu$. It follows very closely the formalism used in [18]. Figure 1.2 shows the decay in the rest frame of the B meson. We choose the z axis in the direction of the W. The angle θ_ℓ is the angle between the lepton momentum, in the W rest frame, and the direction of the W, in the B rest frame.

The four-momenta of the B meson and of the pion are as followed, in the B rest frame: $p^\mu = (M_B, 0, 0, 0), k^\mu = (E_\pi, 0, 0, -p_\pi)$. The W meson is a spin one particle and we can write down it's helicity polarization vectors as,

$$\begin{aligned}
\epsilon^\mu_{W_+} &= \frac{1}{\sqrt{2}}(0, -1, -i, 0) \\
\epsilon^\mu_{W_-} &= \frac{1}{\sqrt{2}}(0, 1, -i, 0) \\
\epsilon^\mu_{W_0} &= \frac{1}{\sqrt{q^2}}(p_\pi, 0, 0, q_0), \\
\epsilon^\mu_{W_t} &= \frac{1}{\sqrt{q^2}}(q_0, 0, 0, -p_\pi)
\end{aligned}$$

where $q^2 = (p + k)^2$, and so, in the rest frame of the B, $q_0 = M_B - E_\pi$. We can also write these polarization vectors in the W rest frame,

$$\begin{aligned}
\epsilon^\mu_{W_+} &= \frac{1}{\sqrt{2}}(0, -1, -i, 0) \\
\epsilon^\mu_{W_-} &= \frac{1}{\sqrt{2}}(0, 1, -i, 0), \\
\epsilon^\mu_{W_0} &= (0, 0, 0, -1) \\
\epsilon^\mu_{W_t} &= (1, 0, 0, 0)
\end{aligned}$$

The differential decay rate is given by,

$$d\Gamma = \frac{(2\pi)^4}{2M_B}|\mathcal{M}|^2 d\Phi, \tag{2.1}$$

where $\mathcal{M}$ is the decay amplitude and $d\Phi$ is the differential amount of phase space. The amplitude is given by,

$$\mathcal{M} = -i\frac{G_F}{2}V_{ub}L^{\mu}H_{\mu}. \tag{2.2}$$

We see the expressions for the leptonic current and the hadronic matrix element,

$$\begin{aligned} L^{\mu} &= \bar{u_\ell}\gamma^{\mu}(1-\gamma_5)v_{\nu} \\ H^{\mu} &= \langle\pi|\bar{q'}\gamma^{\mu}(1-\gamma_5)Q|M\rangle \\ &= f_+(q^2)(p+k)^{\mu} + f_-(q^2)(p-k)^{\mu}, \end{aligned}$$

where f_+, f_- are form factors. The term proportional to f_- vanishes when contracted to the lepton current, in the limit of massless leptons. With the above expressions we can find the helicity amplitude of the W. Since both the pion and the B meson are spinless, the only helicity available for the W is zero,

$$\begin{aligned} H_0 &= \epsilon^{*\mu}_{W0}H_{\mu} \\ &= \epsilon^{*0}_{W0}H_0 + \epsilon^{*3}_{W0}H_3 \\ &= \frac{2}{\sqrt{q^2}}p_{\pi}M_B f_+(q^2). \end{aligned}$$

In the differential decay rate we need the magnitude squared of the amplitude,

$$|\mathcal{M}|^2 = \frac{G_F^2|V_{ub}|^2}{2}L^{\mu\nu}H_{\mu\nu}. \tag{2.3}$$

We start with the leptonic current,

$$
\begin{aligned}
L^{\mu\nu} = L^{\mu}L^{\nu *} &= (\bar{u_\ell}\gamma^\mu(1-\gamma_5)v_\nu)\,(\bar{v_\nu}(1+\gamma_5)\gamma^\nu u_\ell) \\
&= \frac{1}{2}Tr\left[p_\ell\gamma^\mu(1-\gamma_5)p_\nu\gamma^\nu(1-\gamma_5)\right] \\
&= 8\left[p_\ell^\mu p_\nu^\nu + p_\ell^\nu p_\nu^\mu - (p_\ell \cdot p_\nu)g^{\mu\nu} + i\epsilon^{\mu\alpha\nu\beta}p_{\nu\alpha}p_{\ell\beta}\right].
\end{aligned} \tag{2.4}
$$

We can rewrite equation 2.3 in the following way,

$$
|\mathcal{M}|^2 = \frac{G_F^2|V_{ub}|^2}{2}L^{\mu'\nu'}q_{\mu'\mu}q_{\nu'\nu}H^{\mu\nu}, \tag{2.5}
$$

this form allows us to insert the completeness relation of the W polarization vectors,

$$
\begin{aligned}
|\mathcal{M}|^2 &= \frac{G_F^2|V_{ub}|^2}{2}\sum_{mm'}\sum_{\tilde{m}\tilde{m}'}L^{\mu'\nu'}\epsilon_{\mu'}^{m}\epsilon_{\nu'}^{*\tilde{m}}g_{mm'}g_{\tilde{m}\tilde{m}'}\epsilon_{\mu}^{*m'}\epsilon_{\nu}^{\tilde{m}'}H^{\mu\nu} \\
&= \frac{G_F^2|V_{ub}|^2}{2}\sum_{m}\sum_{\tilde{m}}L^{\mu'\nu'}\epsilon_{\mu'}^{m}\epsilon_{\nu'}^{*\tilde{m}}\eta_m\eta_{\tilde{m}}\epsilon_{\mu}^{*m}\epsilon_{\nu}^{\tilde{m}}H_{\mu\nu}.
\end{aligned} \tag{2.6}
$$

We can now evaluate the amplitude piece by piece starting with the lepton sector,

$$
L^{m\tilde{m}} = L^{\mu'\nu'}\epsilon_{\mu'}^{m}\epsilon_{\nu'}^{*\tilde{m}}\eta_m\eta_{\tilde{m}}, \tag{2.7}
$$

We work in the W rest frame in which we can write the four-momenta of the lepton and the neutrino as,

$$
\begin{aligned}
p_\ell^\mu &= (p_\ell, p_\ell sin\theta_\ell, 0, -p_\ell cos\theta_\ell) \\
p_\nu^\mu &= (p_\ell, -p_\ell sin\theta_\ell, 0, p_\ell cos\theta_\ell).
\end{aligned}
$$

Because we are in the W rest frame, we have $p_\ell = \sqrt{q^2}/2$ and $(p_\ell \cdot p_\nu) = q^2/2$. We could now evaluate all the different possible helicity combinations (+,-,0,t) for m and $\tilde{m}$ using equation 2.7. Looking ahead, evaluation of the hadronic piece of equation 2.5 tells us that only the helicity zero is possible, as we mentioned before, so m and m' are zero and we only need to evaluate L^{00},

$$\begin{aligned}
L^{00} &= L_{33}\epsilon^3_{W0}\epsilon^{*3}_{W0} \\
&= L_{33} \\
&= 8[2(-p_\ell^2 cos^2\theta_\ell) + 2p_\ell^2)] \\
L^{00} &= 16p_\ell^2 sin^2\theta_\ell.
\end{aligned}$$

We can write down the hadronic part as,

$$\begin{aligned}
H_{m\tilde{m}} &= \epsilon^{*\mu}_{Wm}\epsilon^{\nu}_{W\tilde{m}}H_{\mu\nu} \\
H_{00} &= \epsilon^{*\mu}_{W0}H_\mu\epsilon^{\nu}_{W0}H^*_\nu \\
&= H_0^2.
\end{aligned}$$

We can now write down the amplitude as,

$$|\mathcal{M}|^2 = \frac{G_F^2|V_{ub}|^2}{2}16p_\ell^2 sin^2\theta_\ell H_0^2. \tag{2.8}$$

The three body decay phase space is given by,

$$d\Phi = \frac{1}{16(2\pi)^7}\frac{p_\pi}{M_B}dcos\theta_\ell dq^2, \tag{2.9}$$

where θ_ℓ is the angle between the direction of the lepton in the W rest frame and the dirextion of the W in the B rest frame. The differential decay rate is then

$$\begin{aligned} d\Gamma &= \frac{(2\pi)^4}{2M_B}\frac{G_F^2|V_{ub}|^2}{2}16p_\ell^2 sin^2\theta_\ell \frac{4}{q^2}p_\pi^2 M_B^2|f_+(q^2)|^2\frac{1}{16(2\pi)^7}\frac{p_\pi}{M_B}dcos\theta_\ell dq^2 \\ \frac{d\Gamma}{dq^2 dcos\theta_\ell} &= \frac{G_F^2|V_{ub}|^2}{(2\pi)^3 4}sin^2\theta_\ell p_\pi^3|f_+(q^2)|^2. \end{aligned}$$

Integrating over the angle we get,

$$\frac{d\Gamma}{dq^2} = \frac{G_F^2|V_{ub}|^2}{24\pi^3}p_\pi^3|f_+(q^2)|^2. \tag{2.10}$$

One important feature of this last expression is the momentum dependence: there is a power of three for the pion decay and we would find a power of one for the vector case. In general the dependence goes like p^{2L+1}, where L is the lowest allowed orbital angular momentum. The power of the momentum influences the shape of the lepton spectrum. The other important feature is that we need to know the dependence of the form factor as a function of q^2. This dependence is the subject of much theoretical activity, as we shall see in the following sections.

For the case of the ρ and ω, the fact that these particles have one unit of spin, allows all possible polarization of the W. The complete expression for the differential decay rate can be found in many places, in [19] for example. We now turn toward the dynamics of the decays.

2.1.2 The Dalitz plot

A Dalitz plot is essentially a map of the probability of the different decay configurations. We show the Dalitz plot for the $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ decays using the ISGW2 model in fig 2.1 and fig 2.2 respectively. The y axis of the Dalitz plots is q^2, which is the momentum transfer to the virtual W boson. The x axis is the energy of the lepton.

We can derive several expressions for q^2:

$$
\begin{aligned}
q^2 &= (p_\ell + p_\nu)^2, \\
q^2 &= (p_B - p_\pi)^2, \\
q^2 &= M_B^2 + m_\pi - 2M_B E_\pi,
\end{aligned}
$$

where the different p's are four momenta. The last equation has been obtained using the 2^{nd} expression, assuming the B is at rest.

To analyze the Dalitz plot we can look at the various regions of q^2. First let us look at the region around the maximum q^2 (also called zero recoil region). This situation happens when the daughter pion receives no momentum and then $q^2 = (M_B - m_\pi)^2$. In this situation the mass of the W is taking most of the available energy, and so it is produced nearly at rest. This has the consequence of producing the lepton and the neutrino nearly back to back.

At parton level, the u quark does not get much momentum from the transfer so if the u quark mass was close to the b quark mass nothing much would have

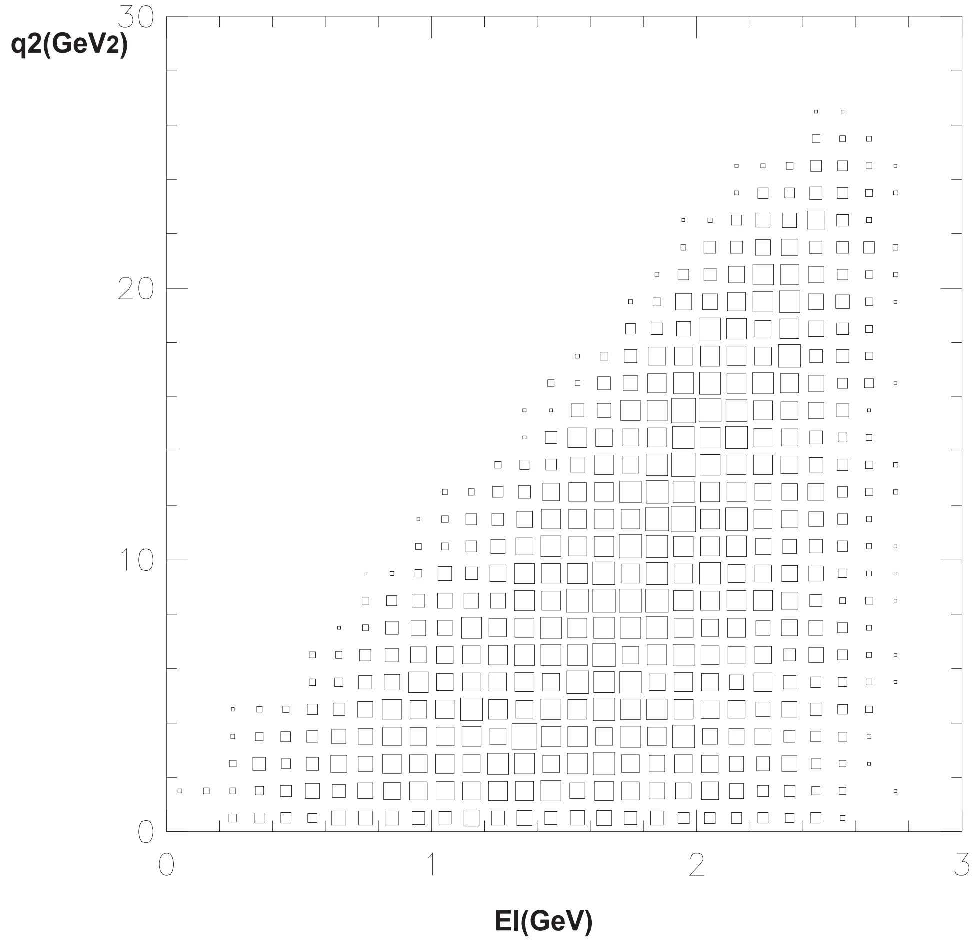

Figure 2.1: Dalitz plot for $B \to \pi \ell \nu$ using ISGW2.

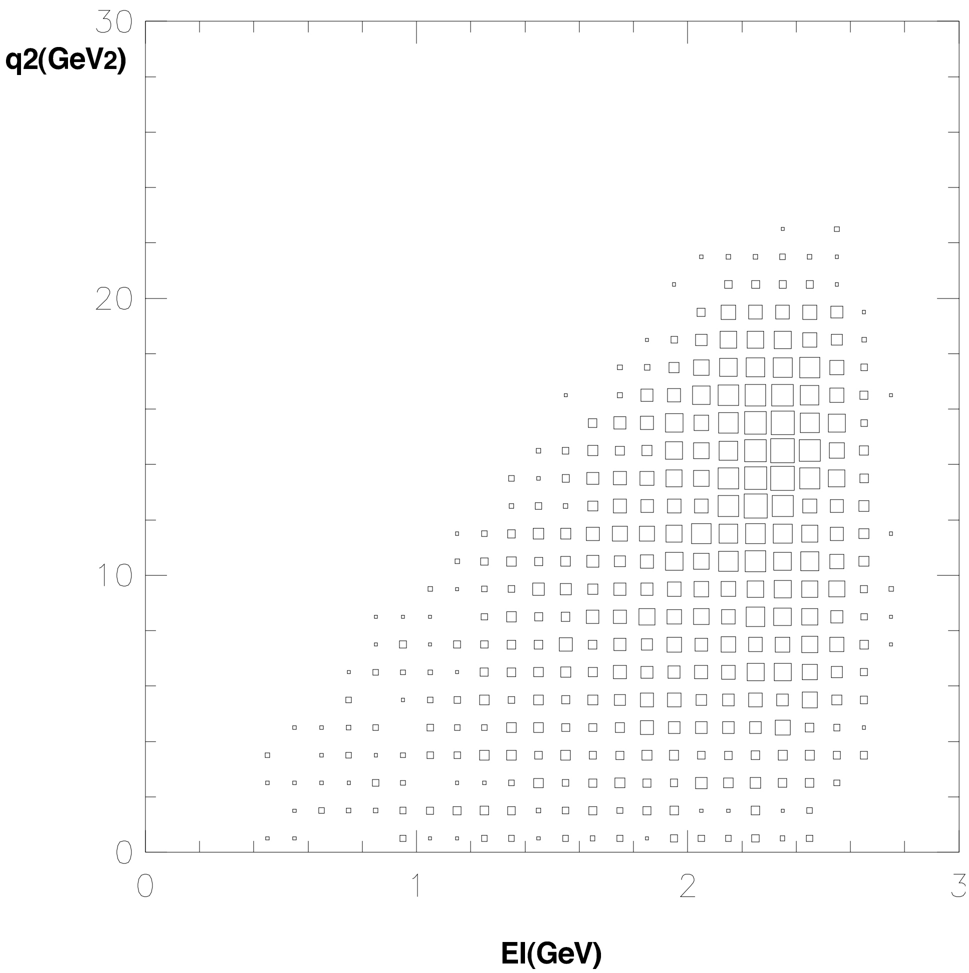

Figure 2.2: Dalitz plot for $B \to \rho \ell \nu$ using ISGW2.

changed in the reaction. This is roughly the underlying idea of the Heavy Quark Symmetry (HQS). The form factors of the hadronic part of the decay favor this situation. In this region of q^2 the relativistic corrections that depend on the mass of the heavy quark is negligible. For the case of $b \to u$ this last remark is not as valid as for the case of $b \to c$. Still, the form factors favors this situation. It is for this reason that the density of points is greater in the top of both Dalitz plots.

The other extreme situation is for q^2 minimum. In that case, the value of q^2 is the mass of the lepton squared, so essentially zero. At q^2 minimum the lepton and the neutrino go flying parallel to each other and opposite to the daughter hadron. In this situation the hadron gets the largest recoil momentum. At the quark level, the u quark got momentum in the reaction so it is moving quickly with respect to the spectator quark. Many gluons must be exchanged between the two quarks to form a bound state. This situation is clearly messier than the previous situation, so the form factors do not favor it. This the reason why the density of points in that region for the Dalitz plot is not as high. It is interesting to note that the theoretical calculations are much harder at lower q^2 because of all the gluons dynamic and the relativistic nature of the particles involved.

There is another factor that influences the q^2 distribution, and hence the Dalitz plot: the spin of the daughter meson. If we take the example of the pion being the daughter meson: we've seen in the previous section that the decay rate is proportional to p_π^3 so this suppresses the rate at high q^2. If we compare with the case of the ρ being the daughter meson, then in this case the decay rate was found to be proportional to the first power in momentum; so this does not suppress the rate as much at high q^2.

We can look at it in a different way: at q^2 max the lepton and the neutrino are nearly back to back, which implies that the third component of spin for the lepton-neutrino system is +1 or -1 but not 0. The B meson has total spin of 0. In the case of the pion, which has also spin 0, there is nothing available to cancel the +1 or -1, so the orbital angular momentum is of value 1. This is referred to as a P wave decay. On the other hand, the ρ has spin 1, so the orbital angular momentum can take any value from 0 to 2. The decay can be an S, P or D wave. This also has the consequence of not suppressing the rate of the ρ at high q^2.

The different situations of q^2 are summarized in figure 1.3 and we show the respective q^2 distribution in figures 2.3 and 2.4.

We now turn our attention to the lepton energy distribution. The important factors for this distribution are the V-A structure of the weak current, and the spin of the daughter meson. The u quark has mostly an helicity of -1/2, while the lepton has almost purely an helicity of -1/2. In the case of the ρ meson, since its spin is 1, then the spectator quark can have either helicity +1/2 or -1/2. This leads to a helicity of the ρ to be either 0 or -1. This in turns leads to the helicity of the W to be either 0 or -1. This has consequences on the energy of the lepton. When the W has helicity -1, then the helicity of the lepton is mostly -1/2 and so because of that the lepton is likely to be emitted in the W direction. Because leptons that go in the direction of the W are given higher energy (in the B rest frame) because of the boost, we get what is known as the lepton forward-backward

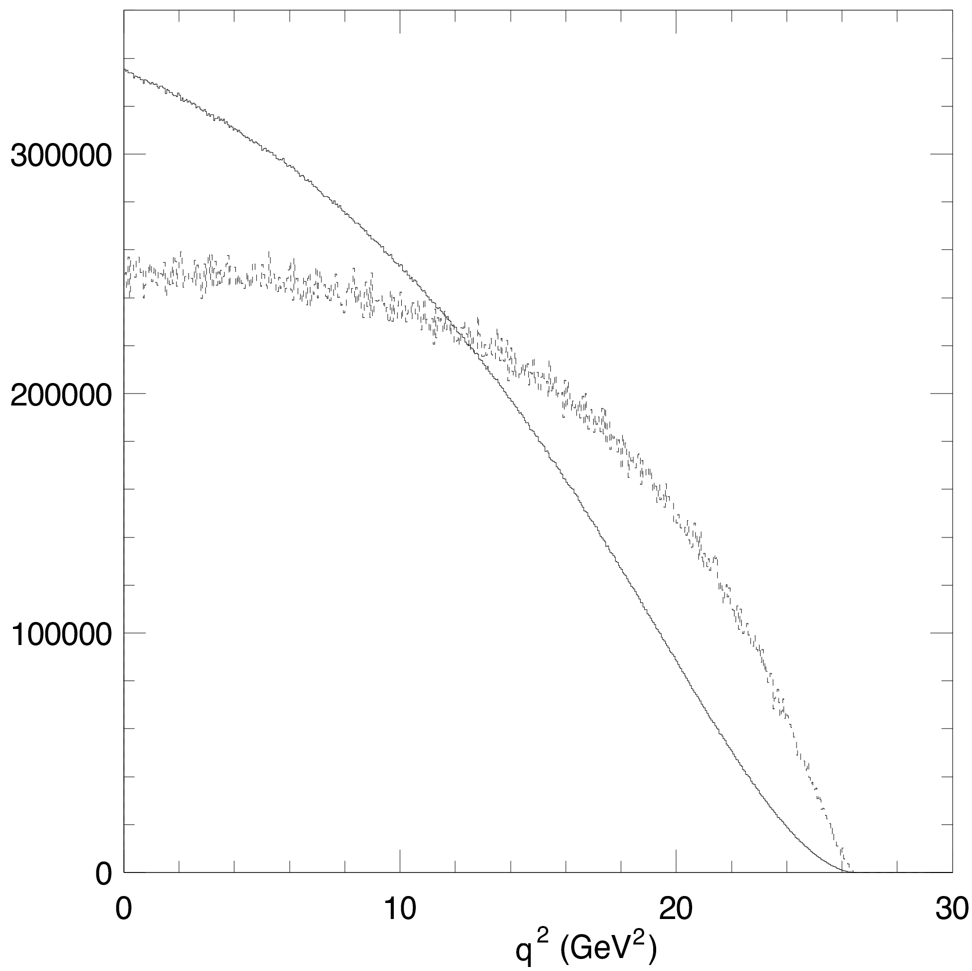

Figure 2.3: q^2 distribution for $B \to \pi \ell \nu$. The solid curve is for ISGW2, dashed curve is for a licht-cone sum rule calculation.

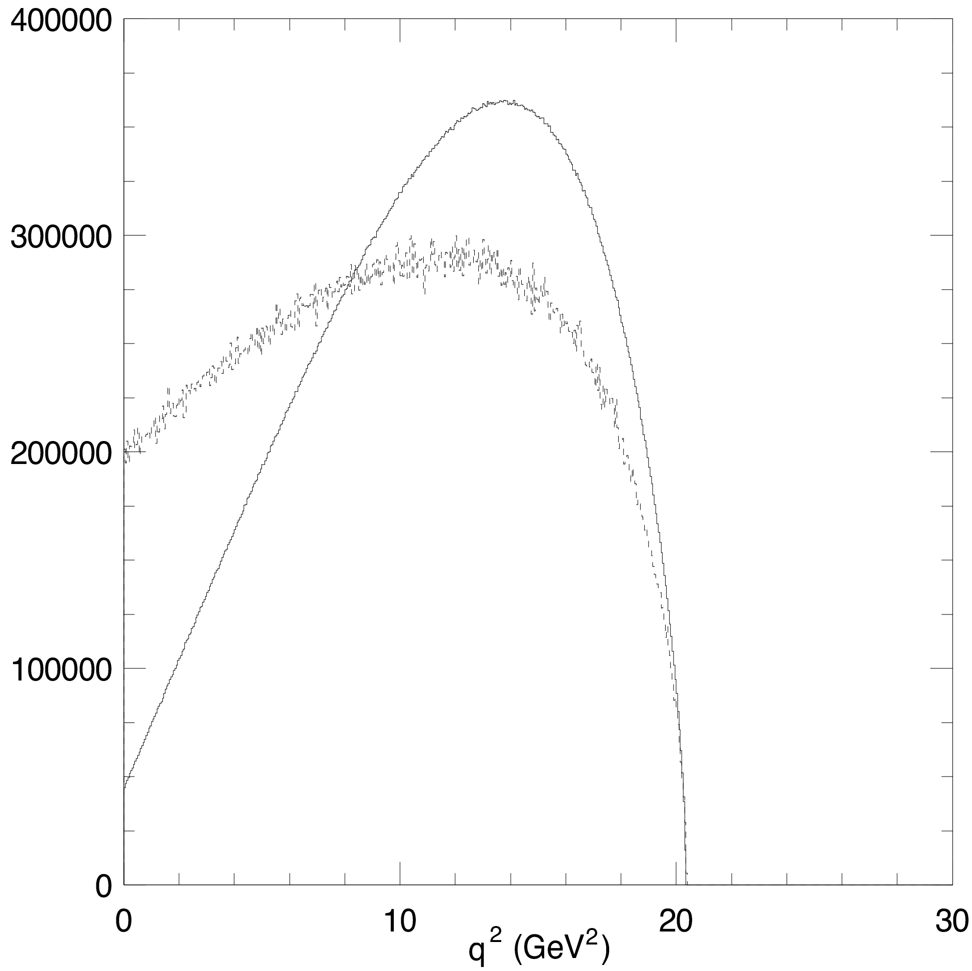

Figure 2.4: q^2 distribution for $B \to \rho \ell \nu$. The solid curve is for ISGW2, dashed curve is for a licht-cone sum rule calculation.

asymmetry. The way to see it in the Dalitz plot is that there is a higher density of points at higher lepton energy for a constant q^2. Another way of seeing it is that the distribution of the rho is roughly proportional to $(1+cos\theta_\ell)^2$ (θ_ℓ has been defined in the previous section) as shown in figure 2.5.

If we look at the case of the pion, the d quark has no choice but to have helicity $+1/2$, so the helicity of the W is 0 and we lose the helicity information.The distribution associated with the pion is proportional to $sin^2\theta_\ell$ as shown in figure 2.6. The absence of lepton forward-backward asymmetry also leads to a softer lepton spectrum for the pion mode compared to the rho mode. The lepton energy spectra are shown in fig 2.7 and 2.8 for pion and rho respectively. Table 2.1 summarizes what we have learned about the Dalitz plot.

2.2 The Form Factors

The theoretical challenge in weak semileptonic decays is to describe the form factors present in the hadronic matrix elements. Those form factors are important in order to predict the rate of the different decays. The prediction can then be used experimentally to obtain a value for the different CKM elements. A powerful handle used to determine experimentally the behavior of the form factor is the q^2 distribution of the decay.

We have seen in section 1.2.3 the basic ideas behind QCD. Trying to describe the form factor behavior's in terms of QCD is very attractive. Unfortunately, for the present we need to build models to fill the gaps where QCD is non-perturbative and

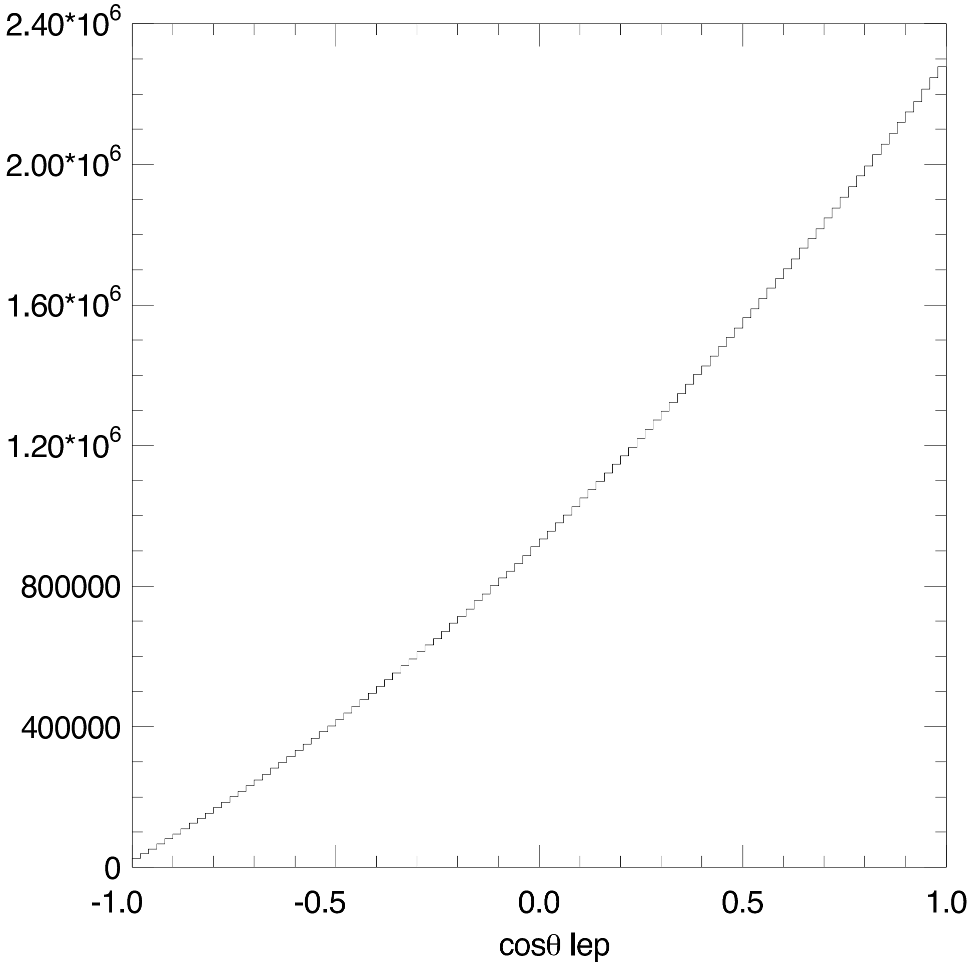

Figure 2.5: $cos\theta_\ell$ spectrum for $B \rightarrow \rho\ell\nu$ for ISGW2.

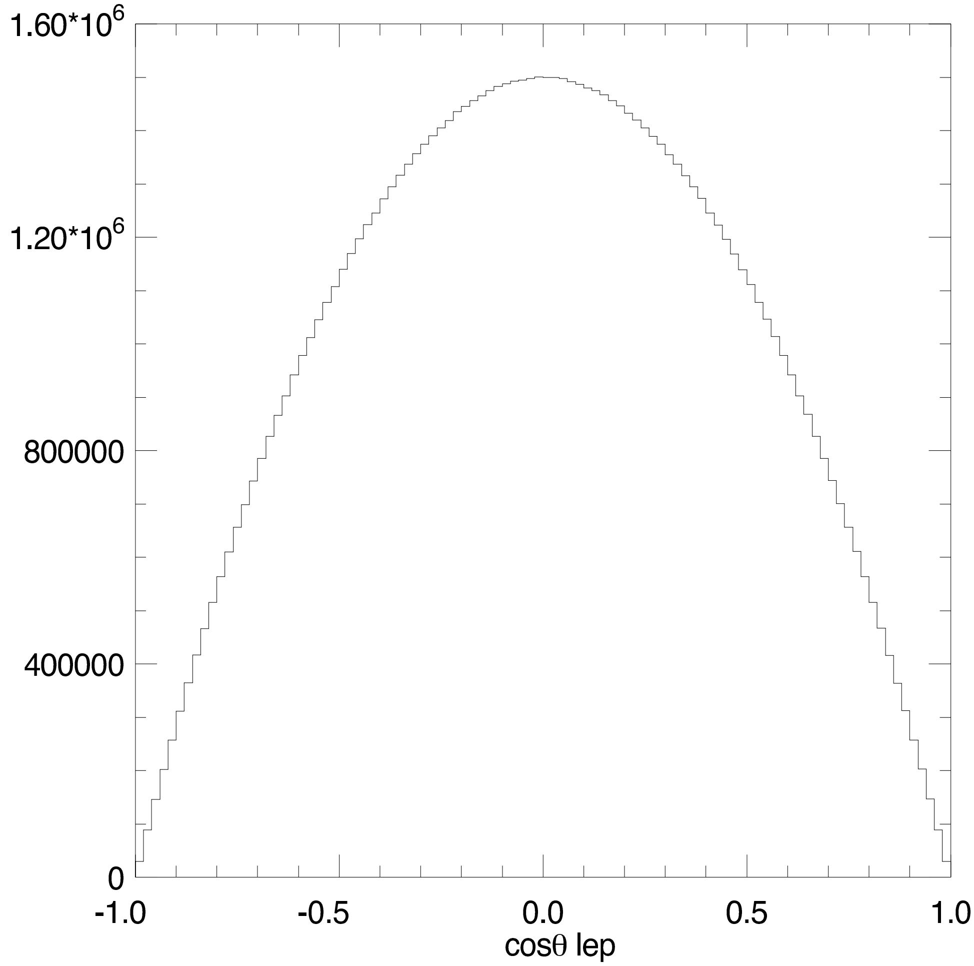

Figure 2.6: $cos\theta_\ell$ spectrum for $B \to \pi\ell\nu$ for ISGW2.

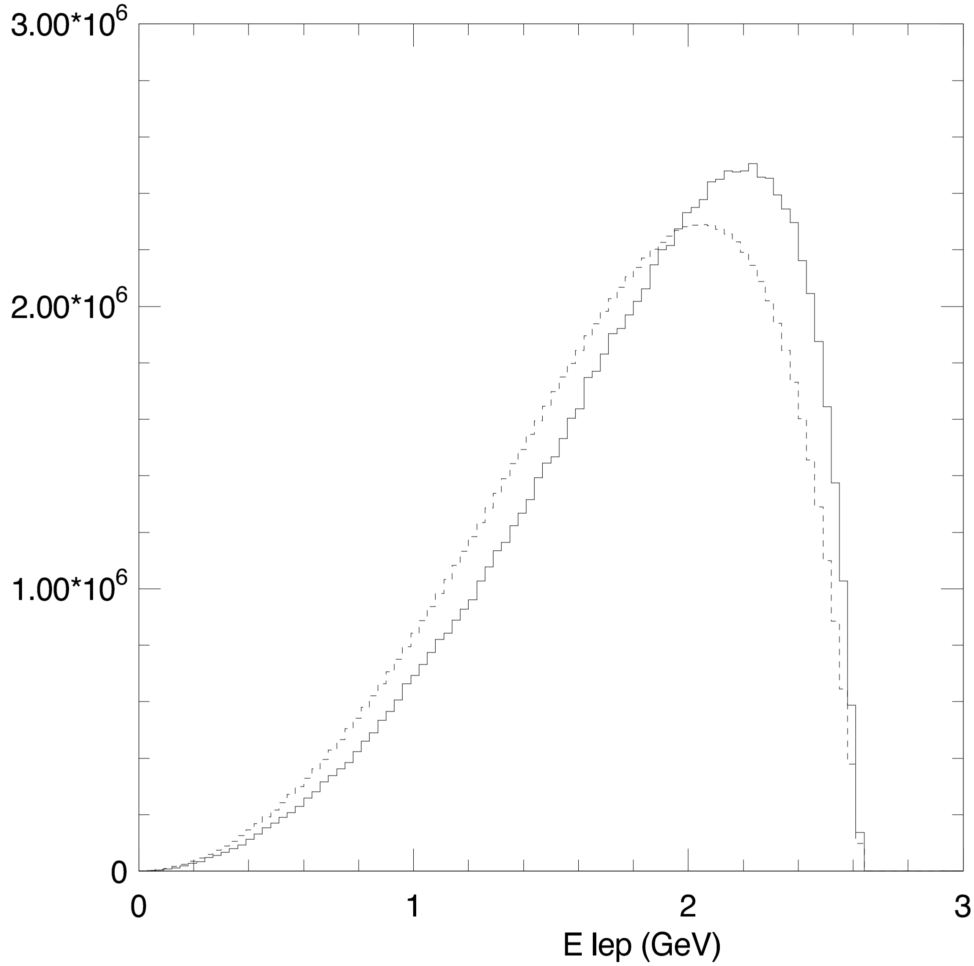

Figure 2.7: Lepton energy spectrum for $B \to \pi \ell \nu$. Solid curve if for a light cone sum rule calculation, dashed curve is for ISGW2.

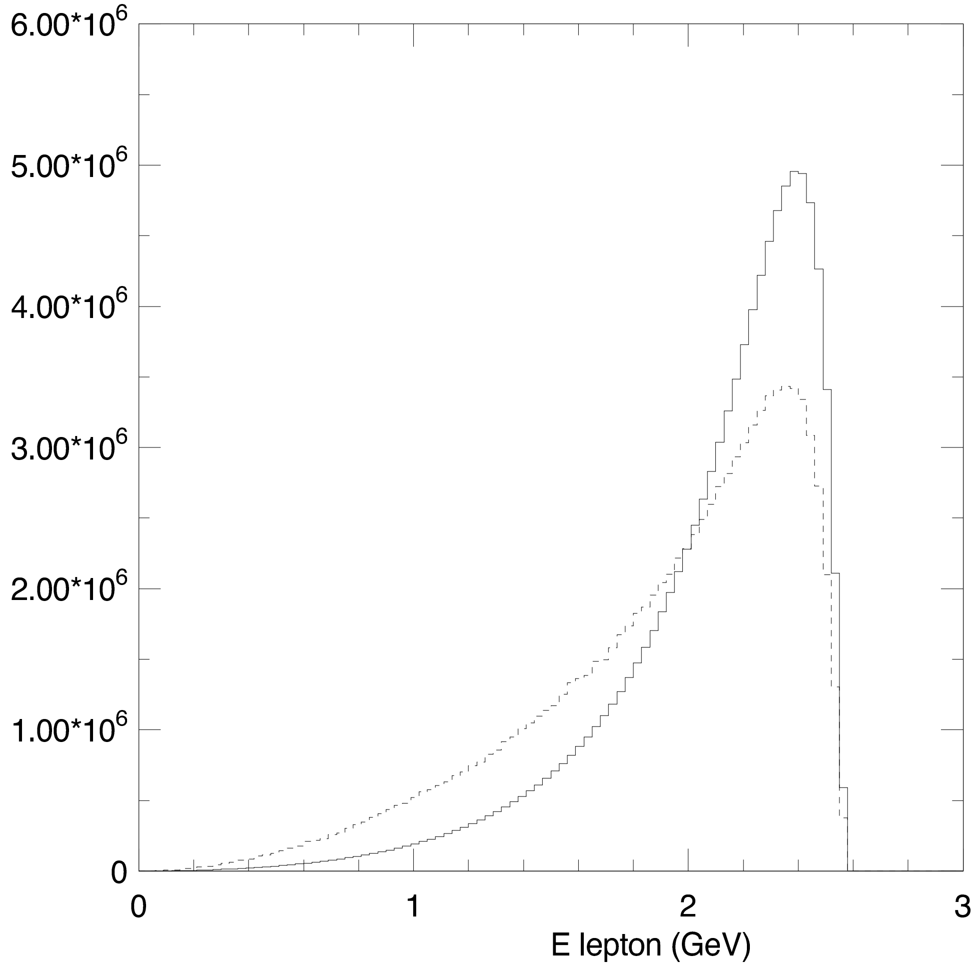

Figure 2.8: Lepton energy spectrum for $B \to \rho \ell \nu$. Solid curve if for ISGW2, dashed curve is for a light cone sum rule calculation.

Table 2.1: Summary of the information for the Dalitz plot

Dalitz region	q^2	Comment
Top	maximum (O recoil)	the momentum of the daughter mesons is small the lepton and neutrino are back to back the form factors favor this region the density of points is high particles are unpolarized distribution of cos is uniform
Middle	middle	the W helicity of -1 starts dominating over +1 (case of ρ) excess of points at high lepton energy
Bottom	minimum (max recoil)	maximum recoil for the π and ρ relativistic situation the π, ρ and W have helicity of 0 no asymmetry in the distribution of $cos\theta_\ell$ depletion of points at low and high lepton energy

numerical solutions are not quite there yet to take over. We review the fundamental ideas between some QCD-based tools, and some models used in particular for the $b \to u\ell\nu$ case.

2.2.1 Heavy Quark Symmetry and HQET

Historically, one of the problem physicists had with the Quark Model was the fact that it was not possible to "see" directly the individual quarks. The explanation for this, which comes naturally from the field theory of QCD, is the asymptotic freedom of quarks: the strong coupling constant, α_s, becomes weaker and weaker in processes involving large momentum transfer. Large momentum transfer can probe short distance reactions. On the other hand, for large distance processes, α_s, is strong, and the system is usually non-perturbative; we cannot make expansions in order α_s. Another consequence is the confinement of quarks and gluons inside hadrons. We can actually use the size of the hadron to get an estimate of the energy scale that distinguishes the weak and strong regime of α_s,

$$R_{had} \approx 1/\Lambda_{QCD} \approx 1fm.$$

We find Λ_{QCD}, the energy scale below which a non-perturbative approach is necessary, to be about 0.2GeV. Is is then natural to classify quarks based on that energy scale: heavy quarks have $m_Q \gg \Lambda_{QCD}$, they are the c, b and t quarks. The other three quarks are considered light quarks in this particular limit.

For heavy quarks, the coupling constant is small, and QCD behaves very much like QED, for which the electromagnetic coupling constant is also small. Actually the hydrogen atom formalism offers a very nice treatment of quarkonium states made of heavy quarks.

The particular system of interest to us involves the B meson, which is made out of one heavy quark, but also one light quark. The heavy quark is then surrounded by a complicated system composed of interacting light quark pairs and gluons; a system sometimes referred to as the “brown muck”. The interesting observation is that in order to have information about the heavy quark's quantum numbers, information about small distance scales is required. The soft gluons, part of the brown muck, can only resolve larger distance scales, so the light degrees of freedom of the heavy quark meson are blind to the mass and spin of the heavy quark, they only feel its color field. This leads to the important conclusion that in the infinite quark-mass limit, if the only things different are the mass and the spin of the heavy quark, the configuration of the light degrees of freedom is the same.

Isgur and Wise were the first to see the implication of this principle: if a heavy quark of velocity v and spin s is replaced by a different heavy quark with different spin, then, as long as the velocities are the same, the light degrees of freedom do not change. This can be directly applied to some semileptonic decays involving transition between heavy quarks, like $b \to c$ for example. This flavor symmetry has an interesting analogy: different isotopes have about the same chemistry, since the electron clouds do not care about the mass of the nucleus. Of course in real life, quark masses are not infinite, and so there will be corrections of order Λ_{QCD}/m_Q.

Since it is possible to factor out the dependence of the mass of the heavy quark, the hadronic matrix element can be written in terms of a function, $\xi(v \cdot v')$, involving the four-velocity of the initial heavy quark (v) and the four-velocity of the final quark (v'). If we now use the heavy quark symmetry, we can change one of the heavy meson into a different heavy meson, and the matrix element is still a function of the same function ξ. This function is called the Isgur-Wise function and it is a universal form factor, valid for any heavy to heavy meson decay. A common variable used for the dot product of the two velocities is w [1]. w is also the boost between the frame of the initial meson and the final meson.

The variable w can also be related with q^2 in the following way,

$$w = \frac{m_Q^2 + m_{Q'}^2 - q^2}{2 m_Q m_{Q'}},$$

where m_Q and m'_Q are the masses of the initial and final mesons respectively. We see that in the particular case of maximum q^2, which corresponds to the final meson being at rest in the rest frame of the initial meson, then the two four-velocities are the same and we have $(v \cdot v') = 1$. This is sometimes referred to as the zero recoil configuration. We also explained in the previous section how this particular configuration is favorable from the form factor point of view. This also means that the Isgur-Wise form factor is maximum and in fact it determines its normalization,

$$\xi(1) = 1.$$

Based on these fundamental principles, it is possible to write an effective theory that deals especially well with systems for which the heavy quark symmetry applies [20]. It is an effective theory because in a sense we do not need all the high energy

[1] It comes from the French pronunciation: double-v

details of the full theory if we are working in a domain of low energies. The prescription to build such an effective theory goes as follows: first, we integrate out all the heavy fields since they do not matter in the low energy regime. This leads to a non local theory since in the full theory the heavy fields had an influence over a short distance, $\Delta x \approx 1/M$. The next step is to rewrite the action as an infinite series of local terms. This series takes care of reproducing the long distance physics of the full theory but not the short distance one. The short distance physics is present because of the hard gluons inside the meson. To take this into account, we need to add up the small distance effects by using perturbative method, this is sometimes referred to as matching.

In the particular case of building a Heavy Quark effective theory (HQET), the mass of the heavy quark is the high energy scale and Λ_{QCD} is the low energy scale at which we are working. Obviously it is not possible to completely remove the heavy quark, but it is possible to integrate out some negligible terms in the full heavy quark spinor. Finally, to write the effective theory it is useful to make use of Luke's theorem, which states that in the limit of zero recoil, there are no corrections of order $1/m_Q$ in the hadronic matrix elements.

HQET has been proven to be very useful for decays involving heavy quarks like, for example, semileptonic decays $B \to D^{(*)}\ell\nu$. In these cases HQS provides a normalization for the form factors at zero recoil, and one can extrapolate over the kinematical range of q^2. For the case of a B meson going into a meson containing a light u quark, the situation is rather different, and HQET can not be applied.

There is no normalization possible in that case, and the principles of HQS can only be applied in the region of q^2 max.

HQS, predicts scaling relations, in the case of zero recoil, for the $\pi \ell \nu$ form factors,

$$f_+(t_m) + f_-(t_m) \approx m_b^{-1/2}$$
$$f_+(t_m) - f_-(t_m) \approx m_b^{+1/2},$$

where, m_b is the b quark mass. We have mentioned before how in the heavy quark limit, the b quarks in the B and B^* mesons act as a static color source and the spin decouples. Hence, in the HQ limit the B and B^* are degenerate. Since the mass difference between the B and B^* is less than the pion mass, we can then expect a contribution to the form factors, from the B^* for the $B \to \pi \ell \nu$ decay. We expect a scaling relation of the type,

$$f_+^{B^*} \approx m_b^{3/2}.$$

We can then conclude that this contribution dominates the rate in the region near q^2_{max}, but it is not expected to dominate the entire range of q^2.

In order to perform comparisons with experiments, we need more than useful relations that hold around symmetry limits. Since any hadronic decay is truly described by non-perturbative physics, only non-perturbative methods will lead to accurate predictions. One such method is lattice QCD.

2.2.2 Form factors on the lattice

Using lattice QCD is in a sense comforting, since in this case we do not have to worry about the justification of the physical approximations: there are none. Also there are some number of methods available to estimate the errors associated with lattice QCD predictions, which is also very useful when the experimentalists are trying to extract the theoretical dependence of their results.

Brief introduction to lattice QCD

[2]

Given a particular physics problem, field theorists usually solve it by evaluating the corresponding vacuum expectation value. This expression can be evaluated from the path integral of an action, $S[\Phi]$,

$$\langle 0|G[\Phi]|0\rangle = \frac{\int \mathcal{D}\Phi e^{-S[\Phi]}G[\Phi]}{\int \mathcal{D}\Phi e^{-S[\Phi]}}. \tag{2.11}$$

As the name invokes, the idea behind lattice QCD is to replace the continuum of space and time by a grid of points. The spacing between each site is a and the total length of the grid is L. Field values are evaluated at each site, but gauge fields are evaluated at each link in between the sites. This discretization turn path integrals into normal integrals, functionals into functions and derivatives into differences. Putting the continuum into a grid is a reasonable approximation when the following condition is met:

[2]Based on [21] and [22]

$$a \ll \text{length scale of particular problem} \ll L.$$

Unfortunately, even though equation 2.11 has been reduced to a more "computable" form, it is not enough: there are too many integration variables in the system and also usually $e^{-S[\Phi]}$ is sharply peaked. The solution is to generate N random field configurations and arrange them so that they are distributed according to the probability distribution $P[\Phi] \approx e^{-S[\Phi]}$. Then the vacuum expectation value is given by the average over all the configurations,

$$\langle 0|G[\Phi]|0\rangle \approx \bar{G} = \frac{1}{N}\sum_{i=1}^{N} G[\Phi^i].$$

The question is how good this expression is for a finite N. If N is sufficiently large, the distribution approaches a Gaussian distribution, and $\bar{G}$ is the mean of that distribution. In that case we also know how to get the width of the distribution,

$$\sigma^2 \approx \frac{\frac{1}{N}\sum_i G^2[\Phi^i] - \bar{G}^2}{N}.$$

For a particular problem that we can solve analytically, we can compute,

$$\sigma^2 \approx \frac{\langle 0|G^2[\Phi]|0\rangle - \langle 0|G^2[\Phi]|0\rangle^2}{N},$$

and with this prescription we have a way of estimating the errors of the simulation.

How do we generate such configurations in practice? Basically, the configurations are sets of random numbers, which are drawn from a particular distribution. One type of method used relies on the generalized Langevin equation,

$$m\ddot{x} = -\frac{\partial V(x)}{\partial x} - \beta\dot{x} + \eta(t),$$

which describes the motion of a particle of mass m under the influence of a randomly acting force from the potential V(x) and the influence of a dissipative force characterized by β. The vector η is a random vector from an arbitrary distribution. From this second order differential equation it is possible to write a 1^{st} order equation for x and the momentum. It is then possible to write by how much x and the momentum change over one time step.

In order to apply this formalism to lattice field theory, we make the following association:

$$\begin{aligned} x(n) &\rightarrow \Phi^n(x) \\ p(n) &\rightarrow \Pi^n(x) \text{fake field} \\ V(x) &\rightarrow S[\Phi] \\ \beta &\rightarrow 1 \\ m &\rightarrow M. \end{aligned}$$

The starting point is to choose completely arbitrary values for each fields Φ and Π. To build the next iteration we can use

$$\Delta\Phi(x) = \frac{\Delta t \Pi(x)}{M}$$

$$\Delta\Pi(x) \;\; = \;\; -\Delta t \left(\frac{\Pi(x)}{M} + \frac{\partial S[\Phi]}{\partial \Phi(x)} \right) + \sqrt{2\Delta t}\,\eta(x).$$

This has the end result of distributing all the Φ^n according to $e^{-S[\Phi]}$, as desired. Note that if the interval Δt between each updating step is small, then two neighbouring configurations do not look much different; they are strongly correlated. We then need to throw away most of the configurations generated and keep only those that are separated by enough updating steps, Δn. It turns out that the size of Δn necessary to have a reasonable distribution, is proportional to L. When we want to approach the continuum limit, we increase L, which has terrible consequences in terms of how many configurations one must generate. This is the phenomenon called "critical slowing down".

We mentioned above that the gauge fields are specified on links joining sites instead of on each site, which we would naively expect for a field. The gauge fields are specified by the line integral, over a link, of the exponential of the gauge field. The reason we have to do this is to preserve gauge invariance, which is the only symmetry we really care about. The reason is: because we are dealing with quantum theories, the fields have structure at all scales, which is what leads to quantum fluctuations. The fact that we need to discretize the theory when we project it onto the lattice has the effect of an ultra-violet cutoff. We have the right behavior at low energy scale, but the high energy effects are not correctly reproduced. The way out of this situation is to renormalize all the running constants found on the lattice so that they give the same physics as the continuum physics. If one wants to reduce the disretization errors, we can add local terms to the original

Lagrangian. These new terms also have running constants associated with them, that much be renormalized. This renormalization is accomplished by relating the new running constants with the ones that were in the original Lagrangian. If the spacing of the lattice is small enough, then asymptotic freedom acts in our favor and perturbation theory can be used to compute all the renormalization factors. What about gauge invariance? If it is not preserved then renormalization theory cannot be used to relate all the constants, and each constant we would need to be tuned separately. That would be extremely time consuming.

The choice of the expansion parameter is important. Since the lattice operators involve an exponential, it is possible to expand it in powers to evaluate it. The terms in the expansion are proportional to the lattice spacing so it is reasonable to think that the terms get smaller and smaller with increasing power of the lattice spacing, but there is a factor a in front of the expansion. This factor has the consequence that the different terms in the expansion are not negligible. An improvement was found that involved rescaling all the fields by a common factor, so that the expansion would go always decreasing in size. This is called the "tadpole improvement". The name comes from the fact that the higher order terms in fields in the expansion involve high-energy loop diagrams, called tadpole diagrams.

Lattice calculations for $B \to \pi, \rho\ell\nu$

We have mentioned before that in our decay of interest the available recoil is large because the final mesons are so light . Large momentum means being able to probe short distances, hence we need smaller spacings when we want to simulate particles with higher momenta. Nowadays, the available range of lattice calculations for the π and ρ modes is a momentum from 0 to about 1GeV. This region corresponds to being very near q^2 max. Extrapolation over the whole range is problematic and introduces model dependency for which the uncertainties are uncontrolled theoretically. For the case of $B \to \pi\ell\nu$, this means that only 20% of the total decay rate is available. There are currently two main approaches to deal with the heavy b quark: one can perform the calculation at the charm mass and then use HQET to do the appropriate extrapolation. This approach is taken by the UKQCD group, for example. The other approach is to use a relativistic action to perform the calculations, but interpret the results in a non-relativistic fashion, allowing one to stay at the b quark mass. This approach is taken by the Fermilab group.

Another characteristic of lattice calculations is that the mass of the u and d quark are too small to be directly simulated on the lattice. Usually calculations are done at roughly the strange quark mass and the results are extrapolated to the physical values. This is called the chiral extrapolation.

Recently the Fermilab lattice group ([23])has worked on lattice calculations in a specific region of q^2, that could lead to a robust value of V_{ub} since no model dependence would be introduced by extrapolating to lower q^2 values. They have also studied the effect of lattice spacing and found only a mild dependence, though their systematic uncertainty receives a contribution from this effect. The expression that they propose is:

$$|V_{ub}^2| = \frac{12\pi^3}{G_F^2 M_B} \frac{1}{T_B(0.4, 1.0)} \int_{0.4}^{1.0} dp \frac{d\Gamma_{B\to\pi\ell\nu}}{dp}, \tag{2.12}$$

where T_B contains the calculations of the form factors and they find: $T_B(0.4, 1.0) = 0.55^{+0.15+0.09+0.09}_{-0.05-0.12-0.02} \pm 0.06 \pm 0.09 GeV^4 + 10 - 20\%$. The first error is the statistical error, the second is the effect of the chiral extrapolation, the third is the effect of the lattice spacing, the fourth is the effect of matching the discrete calculations to the continuum, the fifth is an error coming from linking the lattice units to the physical units, and parts of it reflects the quenched approximation effect. The last 10-20% is the estimate of what the quenched approximation effect could be [3]. The effect of this approximation varies from situation to situation. For example, for the case of the ρ mode this can have a big effect, since there must be a quark pair created from the vacuum. On the other hand, systems like the Υ systems are well approximated because the valence quarks are really the most important players. Recently there have been unquenched calculations performed to determine the decay constants of the B and D mesons: those results drive the uncertainty estimate shown here.

The UKQCD lattice group ([24], [25]) uses a different procedure for the chiral extrapolation. They also report form factors over the whole range of q^2 and for both the π and ρ modes. For their π mode extrapolation over the whole range of q^2 they choose the parameterization developed by Becirevic and Kaidalov (BK) ([26]), which proposes to write the form factor to depend explicitly on the B^* pole, and the other higher states can be written relative to the first pole,

[3] The quenched approximation means that quark polarization loops are ignored.

$$f^+(q^2) = \frac{c_B(1-\alpha)}{(1-\frac{q^2}{M_{B^*}^2})(1-\frac{\alpha q^2}{M_{B^*}^2})}, \tag{2.13}$$

where c_B, α are parameters related to the influence of the B^* pole and other higher states and can be fit for using lattice results. It is important to note that the FNAL lattice results can also be extrapolated over the whole q^2 range using this parameterization. There are other possible parameterizations for extrapolating at lower q^2 that make use of dispersion relations ([27],[28]). The concept of dispersive bounds is introduced in the next section, where we review the other QCD tool for getting the form factors: the QCD Sum Rules.

2.2.3 QCD Sum Rules

QCD Sum Rules were developed more than 20 years ago by Shifman, Vainshtein and Zacharov (SVZ). They are powerful tools used in hadron phenomenology. The idea is to consider a hadron in terms of interpolating quark currents taken at large virtualities ([29]). The fact that the quarks are highly virtual means that we are interested in a short distance process for which pertubation theory can be used. We can write such an amplitude as a correlation function of quark currents,

$$\Pi_{\mu\nu}(q) = (q_\mu q_\nu - q^2 g_{\mu\nu})\Pi(q^2), \tag{2.14}$$

where q is the four momentum transfer of the reaction and has $q^2 < 0$. Our reactions of interest involve both short-distance and long-distance regimes. The long-distance part can be thought of as a sum over all possible hadron states and their polarizations. For $q^2 > 0$ we can make use of the unitarity relation and insert of complete set of intermediate states,

$$2Im\Pi_{\mu\nu}(q) = \sum_n < 0|j_\mu|n >< n|j_\nu|0 > d\tau_n(2\pi)^4\delta^{(4)}(q - p_n), \tag{2.15}$$

where $d\tau_n$ is the phase space volume of all the states, p_n is the four momentum of a state n. We need to be able to relate the $\Pi(q^2)$ of equation 2.14 to its imaginary part involved in the sum over hadronic states. A dispersion relation can do just that. We can represent all the hadronic states as lying on the real axis of a complex plane and take a complex integral on a contour in that plane. The Cauchy formula gives us,

$$\Pi(q^2) = \frac{1}{\pi}\int_{q_0}^{\infty} ds \frac{Im\Pi(s)}{s - q^2 - i\epsilon}. \tag{2.16}$$

The sum rule equates this last expression to a sum of local operators. Some of the local operators represent the short-distance physics and can be computed perturbatively, others represent the long-distance physics and can be calculated in the quark-hadron duality approximation. This approximation can be viewed as saying that

$$Im\Pi(s) \rightarrow Im\Pi^{pert}(s)|_{s\rightarrow+\infty}. \tag{2.17}$$

We can then see that the limitations of the QCD sum rules are the use of the quark-hadron duality approximation and also the fact that the sum of local operators is truncated.

Nowadays, semileptonic phenomenology makes use of Light-Cone Sum Rules (LCSR). This variation of the SVZ sum rules uses an expansion of the quark currents near the light-cone, which means that the quarks are traveling near the light-cone $x^2 \sim 0$, where x^2 is the space-time interval $x^2 = x_0^2 - \vec{x}^2$. Typically we have this condition as $q^2 \to -\infty$, since

$$x^2 \sim \frac{1}{-q^2}. \tag{2.18}$$

This corresponds to highly virtual quarks. Using the LCSR approach avoids the irregularities of the truncated sum of local operators. Another difference of the LCSR calculations is that the sum is performed on nonlocal operators, classified according to their twist instead of their dimension ([30]). The twist of an operator is the difference between its dimension and its spin. The lowest twist possible is two since an operator without any derivative in it has dimension three and the Lorentz spin one. Since sum rules in general involve quarks at high virtualities, the LCSR calculations are most adequate in the region of high recoil of the daughter meson, that is at low values of q^2. Hence, LCSR calculations are complementary to those obtained with the lattice simulations. Calculations for $B \to \rho \ell \nu$ mode get form factors to an accuracy of 15%. These calculations take into account radiative corrections and higher twists calculations, and the authors believe that they cannot be improved. Added to this error is the irreducible error of the quark hadron duality approximation mentioned before. This additional uncertainty is guessed to be around 10%.

The authors in [31] perform LCSR calculations for the $B \to \pi\ell\nu$ mode. They take into account the B^* pole influence referred to previously. They find about an 18% uncertainty on the form factor at $q^2 = 0$. They think that the accuracy of the form factor can be reduced if higher twist calculations are performed for this mode. The authors do not refer explicitly to the quark hadron duality approximation used, but presumably an additional 10% uncertainty must also be considered in this mode.

2.2.4 Quark models

It is clear that the most desirable way to get a precise calculation of the dynamics of heavy quark decay is to go with non-perturbative QCD. However such results are hard to obtain and sometimes, useful approximations and analogies can be made which helps getting theoretical predictions that can be compared to the experimental results. We can then learn whether the approximations were justified.

ISGW2

One of the most popular phenomenological models on the market nowadays is a model developed in the eighties by Isgur, Scora, Grinstein and Wise (ISGW). At that time much emphasis was put on inclusive semileptonic measurements. The hope was to measure the CKM matrix elements V_{cb} and V_{ub} from the inclusive lepton spectra. Then, the theoretical model was ACCMM, a "QCD corrected parton model". Unfortunately, the ratio of the mentioned CKM elements extracted with that model turned out to be much smaller than anyone imagined. The ISGW model was introduced to offer an alternative to this situation.

In the ISGW notation, the letter t is used for q^2 and t_m represents the q^2 max situation, or zero recoil. The ISGW model is a non-relativistic estimate of the intercept $f_i(0)$ and of the charge radii (eq 2.19) of the Lorentz invariant form factors $f_i(t_m - t)$.

$$r_i \equiv \left[6 \frac{df_i(o)}{d(t_m - t)} \right]^1 /2 \tag{2.19}$$

The term charge radius will be clearer later. The expressions found, rely on the observation that for every form factor there is an associated partial wave amplitude. For example, in the case of the transition of a pseudoscalar to a pseudoscalar, the expansion can be written as,

$$\langle P'(p')|V^{\nu}(0)|P(p)\rangle = f_+(p+p')^{\nu} + f_-(p-p')^{\nu},$$

where p and p' relates to the initial and final pseudoscalars respectively. In this case, there are two partial wave amplitudes.

The description of the ISGW model as being non-relativistic applies more than just to the assumption that the external momenta of the initial and final mesons are themselves non-relativistic. It also applies to the internal motion of the light degrees of freedom of the meson. In the second version of the model (ISGW2, which is what is now commonly used), the parameters like the intercepts and the charge radii, involved in the form factors parameterization, includes some relativistic corrections.

Actually, ISGW makes the assumption that the form factors found in the approximation that the light quarks $m_u, m_d, m_s \gg \Lambda_{QCD}$ can be extrapolated into the domain of the physical quark values.

Also, the model has been designed for matrix elements involving resonances only; it does not include non resonant decays. The contribution of the latter was believed to be "small".

The fundamental idea behind the ISGW model is to treat the heavy-light mesons like quarkonia, a bound state of two heavy quarks. This means that all the extra quark pairs and gluons are ignored. We have previously mentioned that the quarkonium system is relatively easy to treat, and is in analogy with the hydrogen atom. In the ISGW model they make the supplementary simplification of using the harmonic oscillator wave function instead of the full quarkonium wave function.

Having described the various approximations used by this model, we now show the parameterization of the various form factors,

$$ff(t) = ff(t_m)\left[1 + \frac{1}{6N} r^2 (t_m - t)\right]^{-N},$$

where r, as mentioned above, is the charge radius, as in the expression,

$$ff(t) = ff(t_m)\left[1 - \frac{1}{6} r^2 (t_m - t) + \cdots\right]$$

and where $N = 2 + n + n'$. The n and n' are the oscillator quantum numbers of the initial and final wave functions. For example, if the decay involves an S wave amplitude to an S wave amplitude then $n = n' = 0, N = 2$, if the decay involves a transition between an S wave and a P wave then we have $n = 0, n' = 1, N = 3$.

As $N \to \infty$, the form factors acquire the Gaussian formalism associated with the harmonic oscillator.

As mentioned before, large hadronic recoils are possible in $b \to u\ell\nu$ decays. This means that there is a large contribution from the $1P$ and $2S$ states in the inclusive lepton spectrum. Since ISGW only takes into account the lowest resonances, this model can only be used in the very end of the inclusive lepton spectrum.

Going from the ISGW to the ISGW2 model, the exclusive spectra show a harder end point. Actually, there was a change of a factor of five between the two versions, for the pion decay rate.

In summary, ISGW is adequate for the great majority of semileptonic decays, and predict a small model dependence. This is not the case for decays of the type $b \to u\ell\nu$. The reasons being that, the final mesons have a large available recoil, and they have a rather relativistic nature. Also, these decays are far from any symmetry limit. ISGW2 gives a factor of two of uncertainty in the $B \to \pi\ell\nu$, and a factor of about 50% uncertainty for $B \to \rho(\omega)\ell\nu$.

ISGW2 also predict that much improvements can be made if one makes full use of the analogous D semileptonic decays, since HQS symmetry predicts nice parallels between those two types of decays. CLEO has started working on a $D \to \pi\ell\nu$ analysis using the data collected with the new particle identification detector.

Relativistic spectral representations

The authors in [34] perform a fully relativistic treatment for their quark model. They get parameterizations of the form factors for both the $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ over the full range of q^2 values. In their model, "the transition form factors are given by relativistic double spectral representations through the wave functions of the initial and final mesons both in the scattering and the decay regions". In order to obtain those spectral representations they write down the amplitude in terms of vector, axial-vector and tensor currents. Then they go to the region of space-like values ($q^2 < 0$) and choose a proper integral in the complex plane. The form factors in the physical q^2 region are obtained by analytical continuation. The obtained spectral functions are ensured to obey QCD constraints in the heavy quark limit. One disadvantage of this approach is that the model depends on parameters like the effective constituent quark masses. In order to reduce this dependence, the authors used lattice results at high values of q^2 as experimental inputs to their model.

They compare the value of their form factors with lattice results at high q^2 and with LCSR results at low q^2. The agreement is good in general. There is one form factors in the $B \to \rho\ell\nu$ case for which their value is slightly greater than the LCSR value by more than the 15% error quoted by the LCSR group.

Skewed parton distributions

Another quark model used for the $B \to \pi\ell\nu$ mode is to model the $B \to \pi$ transition with skewed parton distributions ([35]). Those distributions are non-forward matrix elements of non-local currents. The fact that they are skewed means that there is a change of variable performed and a new variable, the skewedness parameter, is introduced. Performing the integral of the parton distributions over different domains of skewedness, allows one to clearly separate the various dynamical conditions of the reaction. At $q^2 \sim 0$, large recoil, there is the overlap region, while at q^2_{max} the annihilation and the resonance contributions dominate. The term overlap means that for this region of q^2 we have an overlap of the B meson wave function with the π meson wave functions. Annihilation means that the B meson emits the $b\ u$ quark pair and the final pion is made out of the remaining partons. Finally, the resonance contribution is the effect of the nearby B^* pole that must be considered when close to q^2_{max}. The authors assure that the skewedness method ensures that no double counting is possible over the different regions.

They find that their total uncertainty for the form factors is about 20-25%. They compare their distribution of form factors with other theoretical determinations and find good agreement. In particular, they find good agreement with the LCSR approach of [31], which is very close to their approach, since it makes use of a light-cone wave function.

2.2.5 Summary

This chapter presented the relevant theoretical ideas and motivations for an experimental analysis of the charmless semileptonic exclusive decays: $B \to \pi, \rho, \omega\eta\ell\nu$. We have learned that since these decays involve a heavy to light quark transition, useful symmetries like Heavy Quark Symmetry are not directly applicable.

Nowadays there are several theoretical ways to approach how to calculate the non perturbative hadronic element present in these decays. Some of the methods involve brute force calculations of QCD; some of the methods make phenomenological assumptions. One thing is clear, some experimental direction coming from a q^2 distribution measurement for these decays will greatly help the theory community. We now look at the apparatus that enables us to make such measurements.

CHAPTER 3

OBSERVING THE $B \to X_U \ell\nu$ DECAYS

To be able to probe every corner of the Standard Model and its possible extensions we need to go back in time, get ever closer to the creation time of the Universe. As we go back in time, the Universe was hotter and particles of very high mass were created. To be able to study those particles we need a high density of energy. We now turn to the machine that provides enough energy to create a pair of B mesons. We also look into the various detectors necessary to make our measurement.

3.1 CESR

Particle accelerators for high-energy physics are big machines that are hard to build and operate and cost a lot of money. There are less than ten such machines around the world. Some of them accelerate electrons and positrons, some collide protons with anti-protons, or even with other protons, others collide electrons and protons. The technical issues and the physics are quite different for the different machines. The Cornell Electron positron Storage Ring does not operate at the highest energy now possible for an accelerator. Instead its center of mass energy is

10.58GeV, enough to create a $B\bar{B}$ meson pair. Before we go in the details of how the pair is created, we briefly review the different parts of CESR.

The first stage of the accelerator complex, as shown in figure 3.1, is the linear accelerator (linac). A cathode filament is heated and electrons are boiled off. They get accelerated down a 30m beam pipe through the action of eight accelerating cavities. Those cavities generate an electric field from a radio wave, and the electric field increases the energy of the electrons. The electrons reach 300MeV at the end of the linac. To create positrons, a tungsten plate is put about half-way down the linac. When the electrons hit the plate, out of the interactions come a spray of electrons, positrons and X-rays. The positrons get selected and accelerated down the linac to about 140MeV.

The next step is the accelerating part. The synchrotron accelerates either electrons or positrons with four radio frequency cavities. There are 192 bending magnets that give a circular orbit to the particle. The radius of curvature of the particles is given by,

$$R = \frac{p}{qB}, \tag{3.1}$$

where p is the momentum of the particle, q its charge and B the magnetic field strength. We see that as the particles gain energy and momentum, to keep a constant radius, a varying magnetic field is needed. The name synchrotron comes from this particular synchronization. It takes about 4000 revolutions around the synchrotron for the particles to reach 5.29GeV. This takes place in about one hundredth of a second.

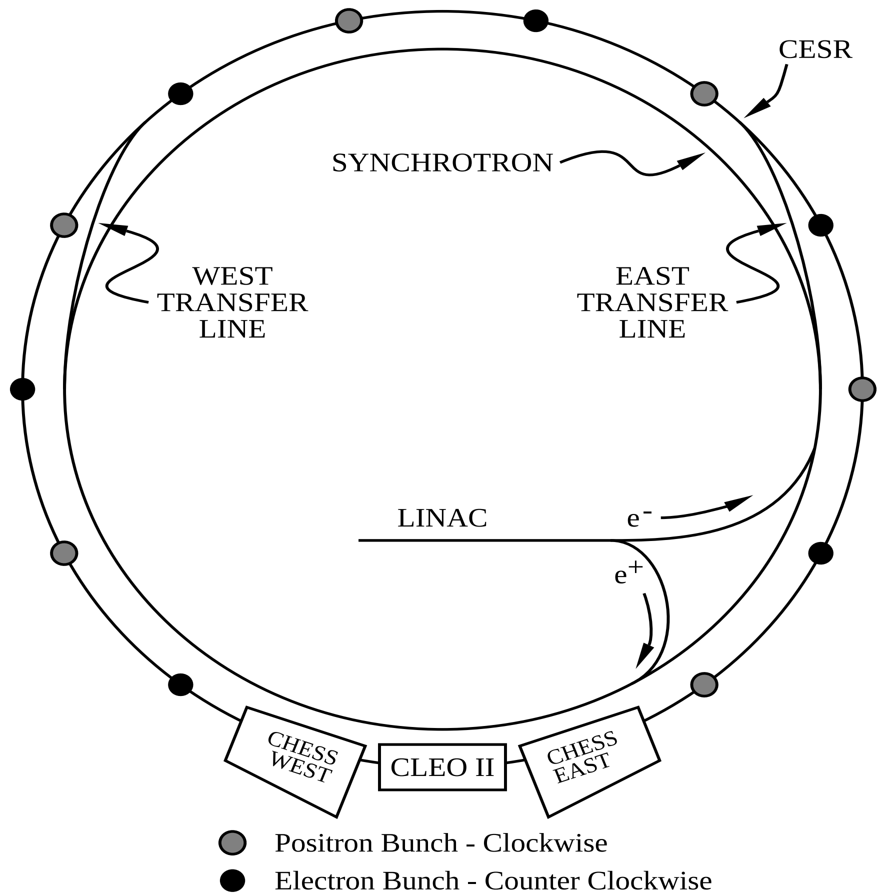

Figure 3.1: Diagram of the various pieces of the accelerator complex.

The final step is to inject first the positrons, then the electrons into the storage ring. Although the storage ring does not provide any acceleration to the particles, its technical requirements are formidable due to the long period of time that the particles stay in it. To allow the particles to stay in the storage ring for about an hour, we need a very good vacuum, of the order of 10^-8 torr, which is several billions times less dense than normal atmosphere. To reach such a vacuum, the metal of the beam pipe can be heated up to 150 C, so that any trapped gas can be driven out with vacuum pumps. Those pumps use a discharge to ionize the air molecules, which then get collected on electrodes. We now look at the various elements around CESR.

Around CESR there are 86 dipole magnets which bend the trajectory of the particles according to

$$F = q\vec{v} \times \vec{B}, \tag{3.2}$$

where F is the inward force, q is the charge of the particle while v is its velocity, and B is the magnetic field provided by the dipole. Those dipole magnets are electromagnets, and there are two types of iron used, one for the soft bend magnets and one for the hard bend magnets. There are 98 quadrupoles around CESR, providing focusing either in the horizontal plane or in the vertical one, by the use of:

$$F = -kx, \tag{3.3}$$

where x is the distance from the center of the axis. Finally, there are 84 sextupoles which correct for the chromatic effect of the quadrupoles, that is they make the momentum distribution of the particles more uniform.

Due to the centripetal acceleration felt by the particles, they radiate off some energy, at the rate of about 1.2MeV per revolution. Those X-rays are a problem for the particle physicists, but they are useful tools for other scientists: the X-ray beams produced by synchrotron radiation are many thousand times brighter and more collimated than the X-rays produced at conventional X-ray laboratories. The Cornell High-Energy Synchrotron Source provides experimental stations to users that wish to make use of these X-rays. As for the particle physicists, they have to restore the lost energy. Radio-Frequency cavities are used around the ring. In the last few years, CESR has replaced its copper cavities for super-conducting ones, providing a better performance efficiency of the cavities. The CESR RF frequency is 500 MHz.

The effect of the linac RF frequency is to group particles together, instead of having a continuous stream of particles. The configuration is to have trains of cars of particles, electrons in one direction and positrons in the other direction, in the same beam pipe. We are currently running in a 9x5 configuration, which means that there are nine trains each separated by 284ns and each train contains five cars, separated by 42ns each. The current record (as of 2002) for the amount of current per beam is 365mA, which corresponds to 2.3×10^{18} particles per car. The trains do not go in a circle, they follow a pretzel-like orbit and collide at a slight horizontal crossing angle of about 2mrad. To prevent collisions to happen in other places than at the center of the detector, there are electrostatic separators around CESR as well. At the interaction point, the cars are about 2cm long, 0.3mm wide and 8μm high. At the position of the collision there is a luminous region, the beam spot, which is smaller than the size of each car: about 13mm long, 300μm wide, and 6μm high. Getting the beam size to be as small as possible is a criteria for a successful accelerator.

The main criteria for a competitive accelerator is defined with its luminosity ([13]),

$$\mathcal{L} = fn\frac{N_{e+}N_{e-}}{A}, \tag{3.4}$$

where f is the frequency of revolution of the particles, n is the number of cars per beam, $N_{e+}N_{e-}$ are the number of particles per beam and A is the cross-section of the beam. The record luminosity for CESR at this point is 1.25 $\times 10^3 3 cm^{-2} s^{-1}$. The reason that the luminosity is important is that when integrated over time and multiplied with the cross-section of a particular reaction, it gives the number of events produced. As we probe rarer reactions, we need more integrated luminosity to get a statistically meaningful sample. Integrated luminosity has units of inverse area, a useful unit is the barn, equaled to $10^{-28} m^2$. A record day for CESR is to collect 73 pb^{-1} (April 2001), this is to be compared to a typical day of about 15 pb^{-1} back around 1996. For information about how the energy of the beams is determined, see [12].

$B\bar{B}$ production

We have just seen that at the interaction point we get a collision of an electron and a positron. Sometimes the electron and positron scatter off each other; Bhabha scattering is the most prevalent reaction happening at CESR. We keep some fraction of these events for calibration of the detectors. Another common process is to get a pair of muons. The cross-section for this reaction is 0.775nb at the CESR energy. The muon pairs have a $1 + cos^2\theta$ distribution, where θ is the polar angle.

Another possibility is that out of the energy of this collision, a bound state of a b and an anti-b quark is produced. This is the Υ resonance, first discovered at FNAL in 1977. There can be different resonances of this bound-state based on the amount of energy available. The next figure 3.2 shows the rate of production of particles vs energy. For the first three resonances, there is annihilation of the b and anti-b quark, while for the fourth resonance, we have enough energy that a pair of light quarks is produced, making up a pair of $B\bar{B}$ mesons (discovered with CLEO in 1983). The fact that the $\Upsilon(4S)$ resonance is very broad is an indication that there are several new decay channels at that energy. The initial state, made out of an electron positron pair, enforces the virtual photon that gives rise to the b quark pair to have a third component of spin to be ± 1. The B mesons are spinless so they have a $sin^2\theta$ distribution, where θ is the polar angle.

Finally, there is another reaction that can happen from the initial electron positron pair: they do not interact and, in place, the photons that radiate from the beam particles interact, while the electrons and positrons go down the beam pipe. This is referred to as two photon physics. Because of the difference in initial

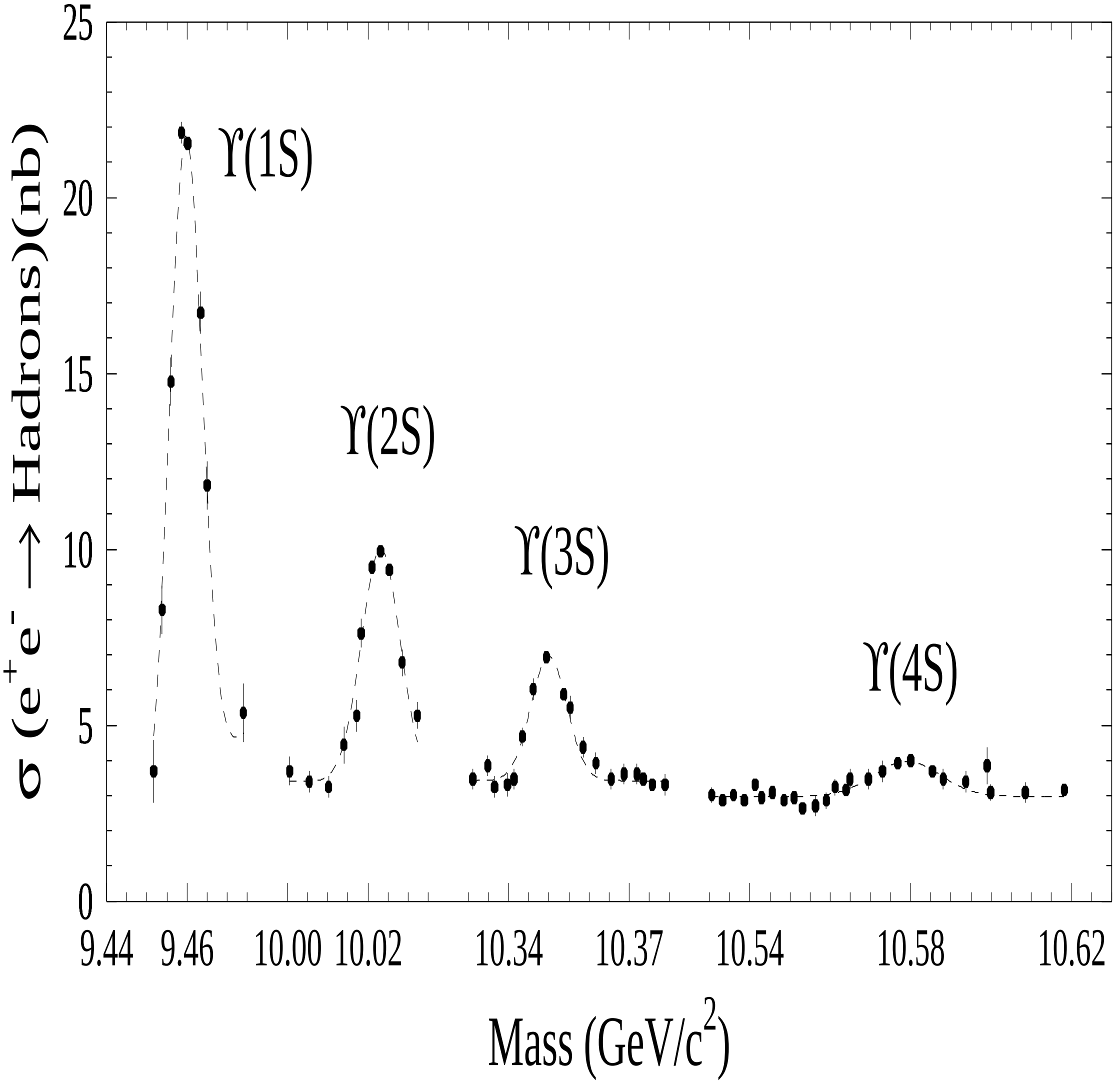

Figure 3.2: Hadronic cross section as a function of the resonance mass.

states, the quantum numbers possible in those reactions are different than the ones in conventional B meson decays, allowing for a very different kind of physics to be performed.

3.2 CLEO

We have just seen how we get a B meson pair at the interaction point. The B meson has a short lifetime (1.6ps) and it decays into various stable and unstable particles. The particle physicists can be compared to a detective: a crime was committed (some initial decay involving the B meson), and we are only left with clues (detection information from the final particles); our job is to use all the evidences to reconstruct what happened. We will first review what happens when particles go through matter, and then we will briefly look into how each subdetector of CLEO makes use of those principles.

3.2.1 Interaction of particles with matter

Most interactions that happen between some incident particle and some material is electromagnetic in nature, hence we can distinguish cases that involve charged particles and cases that involve neutral particles. As for the charged particles, we also have to distinguish between the particular situation of the electrons and positrons compare to the other heavier charged particles. In general, we can say that there are two main things that can happen to a charged particle as it goes through matter ([14]): 1) it loses energy 2) it is deflected from its original direction.

We can identify two main reasons for this: 1) inelastic collision with the atomic electrons of the material, 2) elastic scattering from the nuclei. There can also be other less frequent processes like emission of Cherenkov radiation, nuclear reactions and bremsstrahlung radiation.

The energy loss of a particle is mainly due to the inelastic collisions with the atomic electrons, there is little energy loss from the nuclei interaction since the mass of the nucleus is typically much larger than the mass of the incident particle. Although the energy transfered from the incident particle to the atomic electrons is small, there are so many such collisions, that the energy loss is significant even after short distances in a given material. There can be either soft collisions, where only excitations happen, or hard collisions, for which ionization of the atom happen. Sometimes the ionization electrons have enough energy to produce secondary ionization, we call these primary electrons δ-rays. The energy loss due to ionization is given by the Bethe-Bloch equation (see [14] for example), which relates the energy loss for a given distance, dE/dx, to an expression which does not depend on the mass of the incident particle, but does depend on its velocity, and is somewhat independent of the material. The Bethe-Bloch expression gives the average energy loss, but there are many fluctuations. Those are dominated by the close primary interactions, so that we actually get a Landau distribution of energy loss as a function of distance, which has a long tail. If we look at the energy dependence of the Bethe-Bloch expression, we first see that it is dominated by its dependence on velocity: $1/\beta^2$, then at about a velocity of 0.96c, there is a minimum, and we call the particles at this point minimum ionizing. This point is the same for all particles of the same charge. Finally, past this point, there is a relativistic rise.

We mentioned that the case of electrons and positrons is special, and that is due to the small mass of these particles. There are two ways an electron or positron can lose its energy: by ionization and by bremsstrahlung. For the ionization part, the Bethe-Bloch expression is modified in two ways: 1) since the mass is small, the electron or positron gets deflected from the collision with the atomic electrons 2) since the incident particle and the target particle are identical particles, there are some corrections to be applied.

Bremmstrahlung is the german word for "braking radiation", in fact it represents the emission of electromagnetic radiation from the scattering of the electron or positron with the electric field of the nucleus. At energies below the hundred of GeV, electrons and positrons are the only particles for which radiation loss is significant. Radiation loss by muons is 40000 times smaller than for electrons. Since the energy loss from ionization is roughly constant for electrons and positrons and the energy loss from radiation is a function of the particle's energy, for high-energy electrons, radiation loss is the dominant process of energy loss for those particles. The amount of energy loss in that case is given by ([13]),

$$(\frac{dE}{dx})_{rad} = -\frac{E}{X_0}, \tag{3.5}$$

where X_0 is the radiation length of the material, and we can see that it corresponds to that thickness of material that reduces the mean energy of a beam of electrons by a factor e. The energy at which the energy loss from collision is the same as the energy loss from radiation is called the critical energy.

The second electromagnetic process that we referred to earlier is elastic collision with the nucleus. Those are Coulomb scatterings, and as far as most detectors are concerned, there are many such scatterings, which we name, multiple scatterings. The incident particle zigzag across the detector and the net effect is a deflection from its initial position and momentum, which limits the precision of the detecting elements. The cross section for such a process is given by the Rutherford expression.

We now turn our attention to neutral particles. Of main concern to CLEO is the detection of photons. Since they have no electric charge there cannot be any inelastic collisions with the atomic electrons. Instead there are three main processes involving the photon and the material: 1) photoelectric effect 2) Coulomb scattering 3) pair production.

The photoelectric effect is the absorption of the photon by the atomic electron, which gets ejected from the atom. To have conservation of momentum we need the presence of the nucleus. Coulomb scattering refers to the scattering of an incident photon on a free electron. In a material, the atomic electrons are not free, but since the energy of the incident photon is much higher than the binding energy of the electron, then Coulomb scattering can still apply. Since the cross section of the photoelectric effect goes like $1/E^3$, the one for Coulomb scattering goes like $1/E$ and the one for pair production is constant with energy, any photon of more than about 10MeV loses energy primarily due to the pair production process.

Pair production is the creation of an electron positron pair from the energy of the initial photon. To have conservation of momentum, the nuclei are used for the recoil momentum. In theory, the pair production is very similar to the bremsstrahlung process, so that the conversion length is connected to the radiation length, $\lambda_{pair} \approx \frac{9}{7}X_0$. This relationship between pair production and bremsstrahlung leads to electron-photon showers: an incident photon in a material loses energy through pair production, then the electron and positron emit photons through bremsstrahlung, and so on, until the electrons and positrons reach the critical energy, at which point they lose their energy through collisions. We would get the same pattern if the initial particle was an electron or a positron. The maximum depth of the cascade is a function of the radiation length and the critical energy and is an important factor in choices of detector material. The transverse dimension of the shower is given by the Molière radius and is also different depending on the type of material.

Finally, neutral hadrons, like neutrons or K_L, are going to mostly interact via nuclear interactions. Since the rate of this process is much less than the rate of electromagnetic interactions, much more material is needed to contain those showers. We now look at the specific case of CLEO.

3.2.2 The CLEO subdetectors

During tours to visitors I like to say that High Energy physics is the intersection of two types of physics: the very precise type, we are dealing after all with resolution of the order of the micron sometimes, but also with the very bulky type, a drift chamber that contains thousands of wires and is about 2m long is not an easy piece of equipment to manipulate. The huge size of high energy physics collaboration is a reflection of this fact: we need a lot of people to take care of the multiple challenges and details of each square inch of instrumented area. We now take a tour of the various kinds of detectors present in CLEO. There is no exhaustive description of the construction aspects, which can be found in [15], we rather concentrate on the physics processes going on and their associated challenges. Summary tables of the relevant parameters of each detector can be found at the end of this section. Figure 3.3 shows a view of the overall CLEO detector.

Drift Chambers

We saw in the previous section, how ionization is a big part of the interaction between incoming particles and detector material. To detect the trajectory of the incoming charged particle, drift chambers rely on the charged particle to ionize the gas, so that released electrons drift toward a wire held at some voltage. The CLEO II detector contains three subdetectors that make use of this detection process: the precision tracker (PT), the vertex detector (VD), and the main drift chamber (DR). We show the layout of those drift chambers in figures 3.4 and 3.5. Typically, a sense wire is held at a high voltage and is surrounded by some field wires held at ground. This configuration creates a field cell around the sense wire. As the ionized electron gets closer to the sense wire, it reaches a high field region that creates an avalanche of ions, the charge of which is collected. Although we know which wire collected the charge, we do not know where on the wire the charge

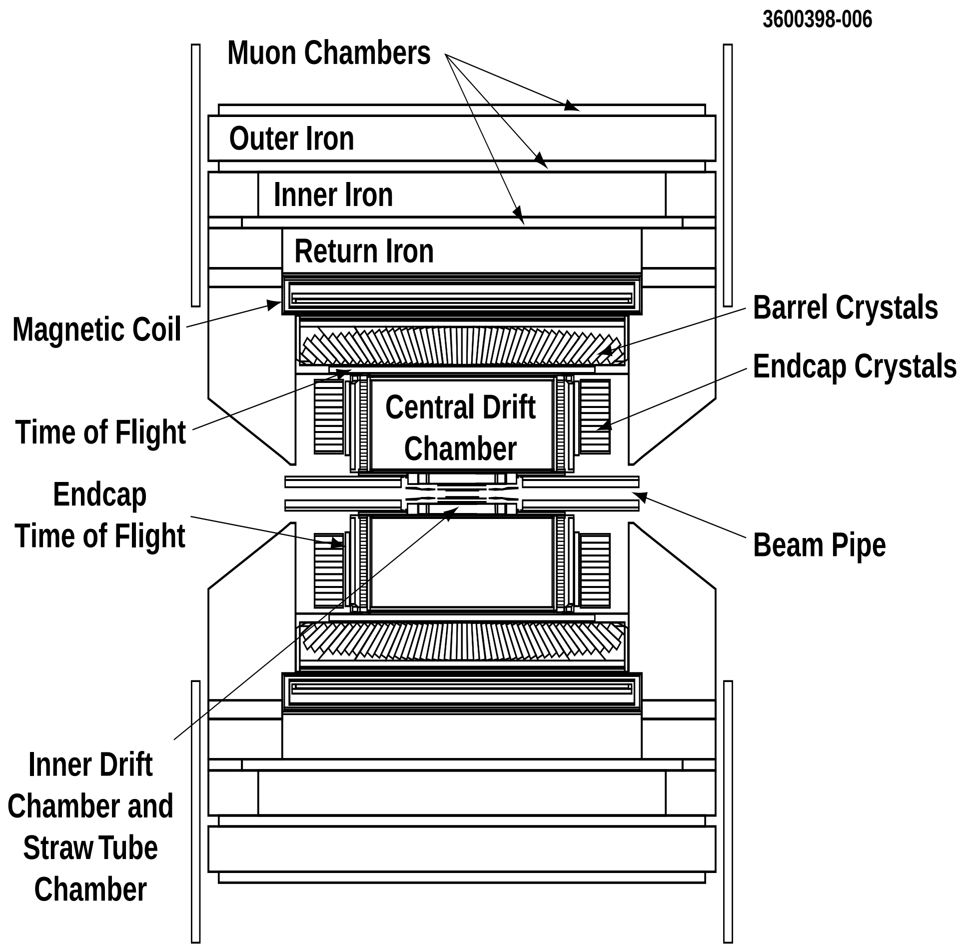

Figure 3.3: R-z view of the CLEO II detector.

was collected. To get some z position information, three methods are used: 1) the vertex detector is read out at both ends, so that based on the charge difference some estimate of the z position can be inferred, 2) the main drift chamber has some of its layers not perfectly aligned in z, they have a slight stereo angle, which allow for some z information, 3) finally, there are four cathode planes segmented along z at the inner and outer edges of the VD and DR. The cathodes rely on the image charge produced by the avalanche near the sense wire to induce an equal charge on the corresponding cathode pad.

The drift chambers are in a 1.5T magnetic field, provided by the super-conducting magnet. We can see from the following expression that we can get the momentum of the particle, if we know the radius of the trajectory's curvature,

$$pc = Be\rho, \tag{3.6}$$

where ρ is the radius of curvature. We also realize that since the drift chamber has a radius of about 1 meter, particles with a momentum less than about 225 MeV/c cannot exit the chambers. They spiral inside the chamber and leave tracks that we call curlers.

The presence of curlers in the drift chamber raises the issue of track finding and fitting. Tracks are essentially built by putting together track segments in superlayers, which are groups of layers, and then seeing if those segments can be combined to make a track. The goal of track fitting is to get as accurately as possible the position and momentum of the particle at the place of its creation, along with the right error matrices of these quantities. In the era that we refer to

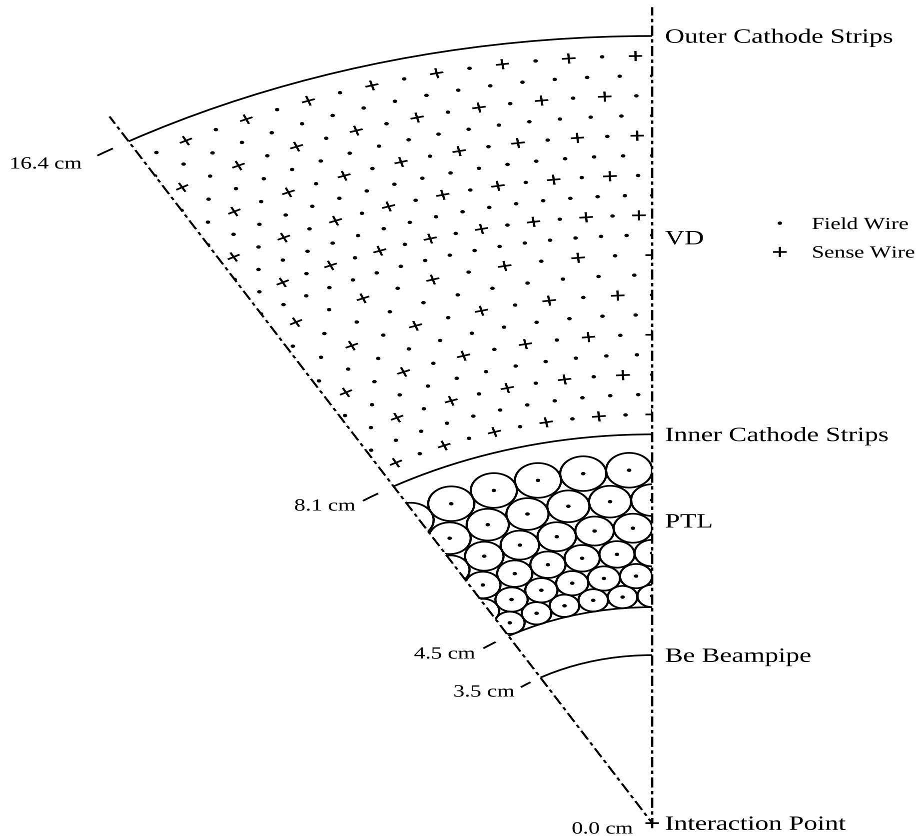

Figure 3.4: Cells layout of the inner drift chambers.

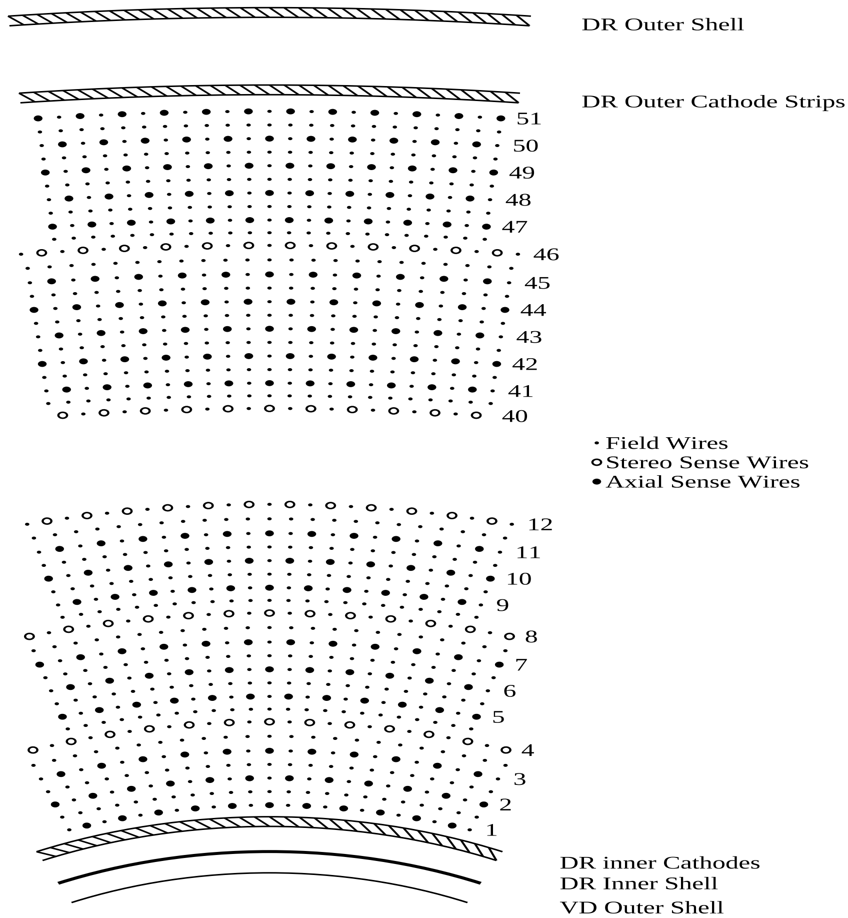

Figure 3.5: Cells layout of the main drift chamber.

pre-compress, a simple chi square fit was applied on the string of hits found with pattern recognition, to make a helix trajectory. In the re-compress era, we started using what is called the kalman filter, which in our case is only applied for track fitting. It takes into account energy loss and multiple scattering, so that each hit is taken into account independently in the track fit. Using this method we accurately get a track fit for each particle hypothesis: electron, muon, pion, kaon and proton.

In the previous section we described how the specific ionization of a particle is a function of its velocity. The drift chamber also provides a measurement of dE/dx, which combined with the momentum information determines the mass of the particle, and is used for particle identification. In figure 3.6 we show the various bands as a function of momentum, for the different particle species.

For the CLEO II.V version of the detector, the Argon Ethane gas of the drift chamber was replaced with a mixture of Helium and Propane. The new gases reduced the multiple scattering effect, improving the charge collection efficiency. Improved tracking and dE/dx performance resulted from this change.

Finally, aside from the drift chambers, the muon chambers also rely on ionization for charged particle detection. Muons are penetrating particles and they are the only particles remaining after the iron of the magnet flux return. Three superlayers of three layers each are inserted at an absorption length of 3, 5 and 7. A cross section of a superlayer is shown in figure 3.7.

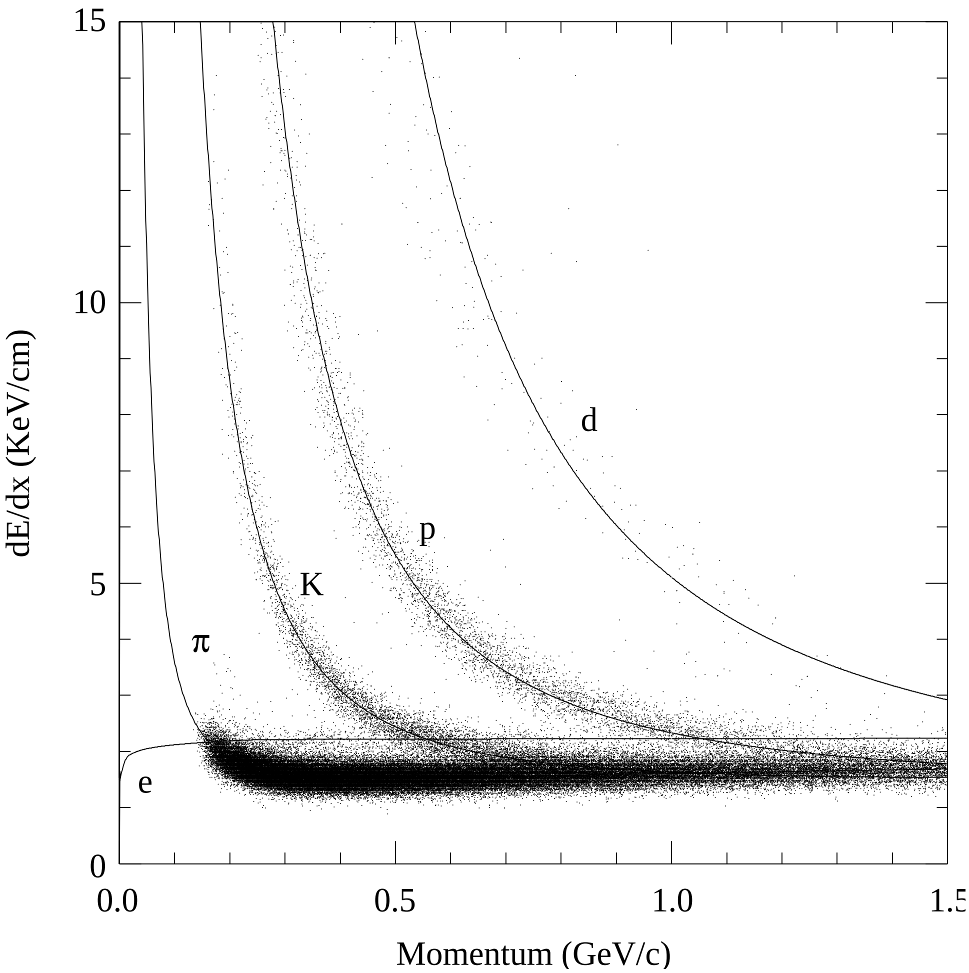

Figure 3.6: The dE/dx bands vs momentum for the main drift chamber.

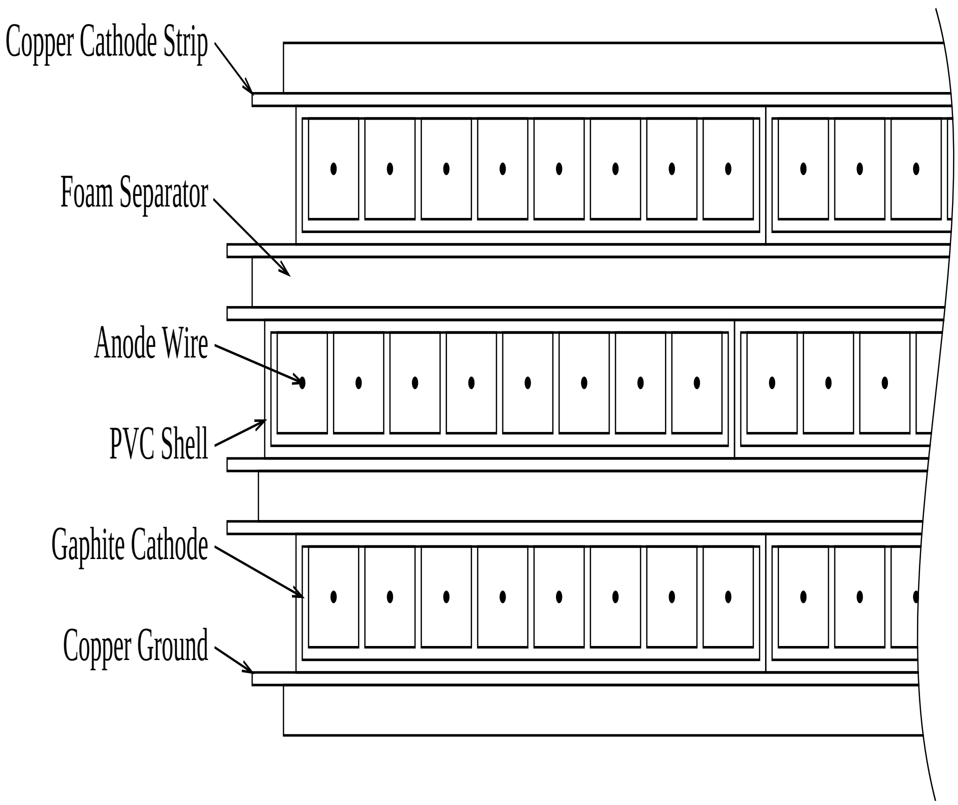

Figure 3.7: A superlayer of the muon detector.

Semiconductor Detectors

Semiconductor detectors have been in use in nuclear physics for a long time, but is is relatively recently that its general use has propagated to high energy physics experiments ([14]). Similarly to the ionization happening in the gas of drift chambers, an incoming charged particle creates electron-hole pairs which are collected by an applied electric field. A big difference with gas ionization is that the minimum energy required to produce such a pair is much less than the energy necessary to ionize a gas. A disadvantage of these detectors is the amount of material they put in the path of the particles and their susceptibility to radiation damage.

For CLEO II.V the semiconductor detector is a three layer double-sided silicon microstrip detector (SVX). Both $r\phi$ and rz information is read out. It replaced the PT of CLEO II and improved the impact parameter in $r\phi$ by a factor of two while the rz resolution improved by an order of magnitude ([16]). Figure 3.8 shows the various layers of the SVX detector.

Scintillators and Calorimeter detectors

Other detectors that make use of crystals's conducting properties are scintillators. In those detectors, some material is doped, so that when charged particles liberate electrons and holes, they get captured by the doping element. This activation center gets excited and decays with some light emission. This light is then captured by photodiodes.

The CLEO time of flight (TOF) detector makes use of this process by recording the time it took for a particle to go from the interaction point to the time of flight

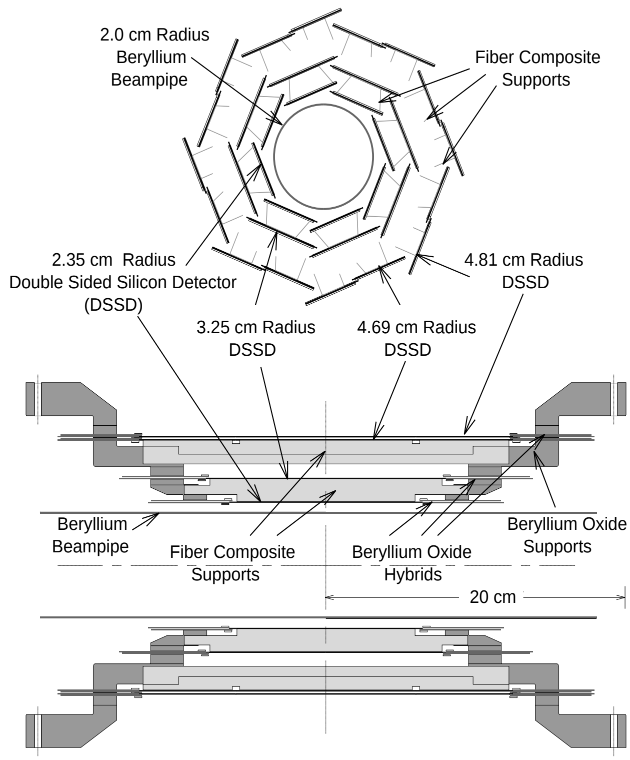

Figure 3.8: The three layers of the Silicon VerteX detector.

crystals. This gives information on the velocity of the particle and can be used to get the identity of the particle when combined with its momentum measurement. Figure 3.9 shows the different bands of $1/\beta$ as a function of momentum, for the different species.

If one puts a dense enough crystal in the path of the particles, this becomes an absorption detector. The CLEO electromagnetic calorimeter (CC) is made of the inorganic crystal Cesium Iodide and it is doped with Thallium atoms. We have seen in the previous section how an electromagnetic shower is created from the incoming photon, electron or positron. The various electrons and positrons produced in the shower ionize electrons and holes in the crystal, which then activate the Thallium atoms.

Trigger and DAQ

If CLEO were to record data for every CESR bunch crossing, the rate would be of the order of 3.6MHz. Fortunately, the amount of interesting events is only about a few Hz. CLEO has a three hardware level trigger system before a software filter is applied to make a final decision about which events are written to a permanent storage device. The first trigger level, L0, tries to determine if there were any charged particles or neutral particles in the events, it takes information from the TOF, the VD and the CC. It reduces the rate to about 10kHz. The next trigger level, L1, makes use of the TOF, VD, DR and the CC and combined with L0 takes a few microseconds to process. After L1, the rate is down to about 50Hz. The next trigger level, L2, takes more detailed tracking information and reduces the

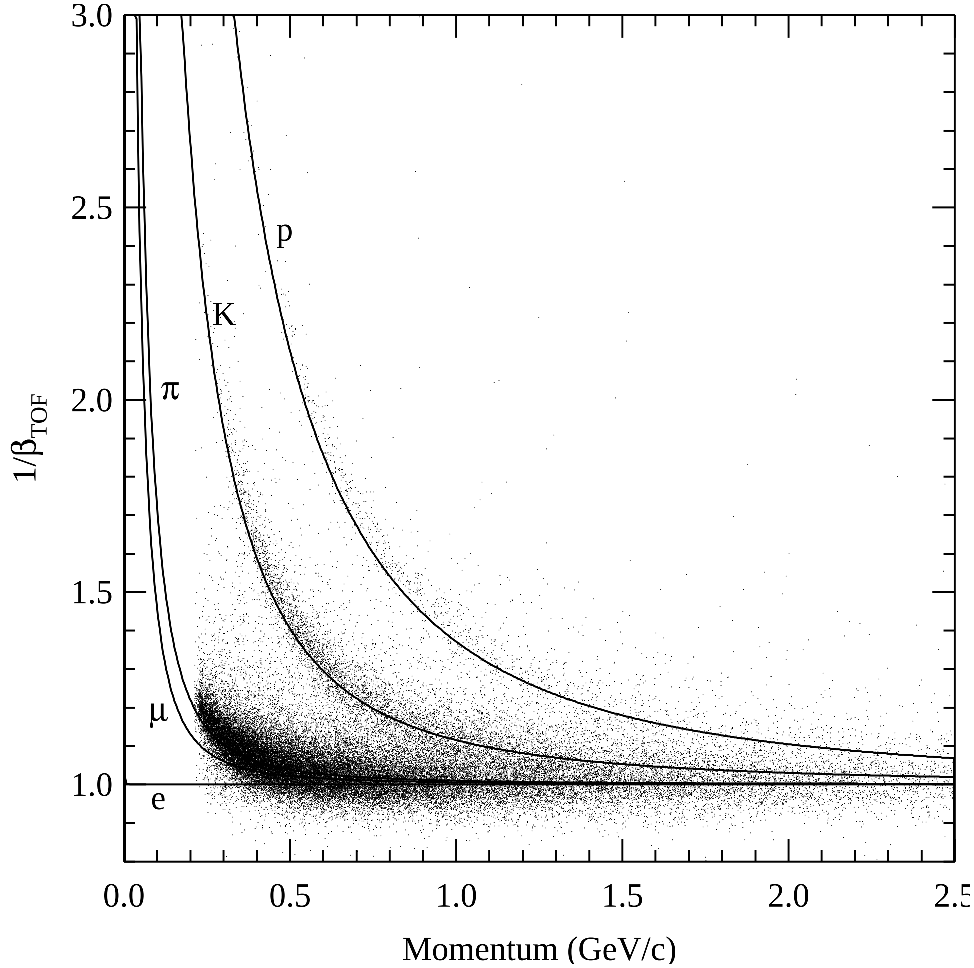

Figure 3.9: $1/\beta$ from Time of Flight vs momentum.

readout by a factor of two. The software trigger L3, gives an additional 30,40% reduction. The trigger efficiency for $B\bar{B}$ events is 99.8% and 100% for $B \to X_u \ell \nu$ events.

The read out of the information is done differently for the different detectors. For example, for the drift chambers, the trigger sends a stop signal, which then makes the drift chamber read out send its timing information contained in its buffer from before the stop. This takes about 2 μs. On top of digitizing the information from the electronics, some sparsification is also necessary to reduce the event size: for example, only channels that recorded a non zero value are kept. In CLEO II the digitization and sparsification take about 13.5 μs. A typical event size for CLEO II is 8 kbytes. For CLEO II.V, it is much higher because of the SVX.

Summary of the CLEO detector

We summarize the geometrical aspects as well as the resolutions of each detector in tables 3.1 through 3.4. Ar is for Argone, C_2H_6 is ethane, He is helium, C_3H_8 is propane, and DME is dimethyl ether.

3.3 Monte-Carlo simulation

This analysis depends crucially on Monte-Carlo events: as we will see in the next chapters, there are background events mixed in with our signal events. In order to extract the fraction of events representative of our signal modes, we need some Monte-Carlo simulation to represent both the signal events and the background

Table 3.1: Information about the central subdetectors

Detector Name	Radius	Comments
beam pipe	3.5 cm	II: beryllium, silver coating
	2.0cm	II.V: 2 concentric cylinders, water in 0.5mm gap
		10μm gold coating
PT	4.7-7.2cm	6 layers of aluminized mylar tubes
		384 gold plated tungsten wires
		gas: $\leq$ 1992: 50:50 Ar-C_2H_6
		$\geq$ 1992: DME
		Non existent in II.V ($\geq$1995)
SVX	2.35-4.81cm	3 double-sided layers of Silicon strips
		96 wafers, 300μm thick
		pitch of $r\phi$ side: 28 μm, rz:100 μm
		resolution for tracks perpendicular to beam:
		$r\phi$:$\sim$ 20 μm, rz:$\sim$ 27 μm
		Non existent in II ($\leq$1995)

Table 3.2: Information about the VD and DR subdetectors

Detector Name	Radius	Comments
VD	7.5-16cm	70 cm long, 10 layers
		800 nickel-chromium sense wires
		2272 aluminium field wires
		Inner and outer cathode planes
		gas: 50:50 Ar-C_2H_6
		read out at both ends
DR	17.5-95cm	2m long, 51 layers, 92% coverage
		40 axial, 11 stereo (up to 6°)
		12240 gold plated tungsten sense wires
		36240 gold plated (Al/copper beryllium)
		field wires
		gas: II: 50:50 Ar-C_2H_6
		II.V: 60-40 He-C_3H_8
		resolution: $\frac{\delta p_\perp}{p_\perp}^2 = (0.0015 p_\perp)^2 + (0.0050)^2$
		$\sigma\phi = 1mrad$, $\sigma\theta = 4mrad$

Table 3.3: Information about the TOF and CC subdetectors

Detector Name	Radius	Comments
TOF	0.95-1.02cm	$\sim$ 5cm thick Bicron BC-408 plastic scintillators Barrel: 64 tubes, long UV transp. lucite light guides phototubes outside magnet, 86% coverage resolution: $\sim$ 150ps Endcap: 28 tubes, phototubes on scintillators inside magnet, extend coverage to 96% resolution: $\sim$ 300ps
CC	1.02-1.44 cm	7800 Thallium doped Cesium Iodide crystals $\sim$ 5cm x $\sim$ 5cm x $\sim$ 30 cm 4 photodiodes on each crystal, inside magnet Barrel: 6144 crystals, 48 z rows x 128 azimuthal near vertex pointing geometry E resolution for 1GeV photon: $\sim$2.2% Endcap: 828x2, E resolution for 1GeV photon:$\sim$ 2.8%

Table 3.4: Information about the outer subdetecors

Detector Name	Radius	Comments
Coil	~ 1.5m	superconducting, axial 1.5T, uniform to 0.2%
		cooled by liquid helium
		over 95% of tracking volume
MU	≥ 1.74m	plastic streamer counters
		Barrel: embedded in iron flux return
		8 octants, 3superlayers at 36,72,108cm (3,5,7 X_0)
		Endcap: forward and backward region
		each superlayer: 3 staggered layers
		each layer: 8 rectangular plastic tubes
		anode wire in center of rectangle
		1 side of rectangle: perpendicular copper strips
		gas: 50:50 Ar-C_2H_6
		resolution: 2.4cm, z: 2.8-5.5cm

events. Fortunately, CLEO has a lot of experience in generating simulated events, and also, systematic uncertainties take into account possible mismodeling of the simulated events.

There are three main stages to generating simulated events: the physics of the generated events, the simulation of the detector and the processing of the simulated detector responses. For the first step, we rely as much as possible on published measurements, a lot of them from CLEO, of B decays, or of continuum events[1]. For the remaining decays, theoretical predictions or educated guesses are used for the different branching fractions. It it interesting to note that only about 50% of the branching fractions of the B meson have been measured. In the special case of our analysis, the B decays that most influence the results come from the $b \to c\ell\nu$ and $b \to u\ell\nu$ reactions. CLEO has a long experience in measuring charmed semileptonic decays. The decays $B \to D\ell\nu$, $B \to D^*\ell\nu$, $B \to D_1\ell\nu$, $B \to D_2^*\ell\nu$, and even non-resonant decays like $B \to D\pi\ell\nu$ have all been studied at CLEO. Furthermore, in some cases, the form factors were measured and since Heavy Quark Symmetry holds relatively well for the charm system, this information can be used for other charmed modes. Nonetheless, part of the systematic uncertainties come from varying wildly the relative rate of these decays. Also treated in the systematic uncertainties is the influence of various theoretical assumptions about how to generate $b \to u\ell\nu$ events.

[1]Continuum events are made of jets of light quarks: u,d,c,s, as well as τ pairs, and two photon interactions

The next step is to simulate the detector response to the passage of the generated particles. It is crucial that a thorough description of each piece of material be included in the GEANT based simulation. As we have seen earlier, the amount of material a particle encounter, whether it is instrumented or not, has crucial consequences on its energy loss or direction. Some of the detectors are simulated according to a first principle approach: for the drift chambers, ions are generated according to a Poisson distribution; while other detectors simulate some macroscopic features: electromagnetic showers are not created by simulating the various chain reactions and atom excitations but rather by producing a signal of a certain intensity based on data studies. In both cases, there are "knobs" (called constants!) associated with key parameters. Those knobs are turned to reproduce the data and also take into account the aging of the various detectors. The end result is a very good agreement of low level features between the MC and the data. For example, in simulating the drift chambers, we get 1% agreement between data and Monte-Carlo for the tracking efficiency, and we get up to 5% agreement between data and Monte-Carlo for the resolution ([17]).

Finally, the same PASS2 program used to take all the raw detector signals and turn them into reconstructed information is used to treat the Monte-Carlo. This ensures that no bias is introduced between the data and the Monte-Carlo.

CHAPTER 4

MEASURING THE BRANCHING FRACTIONS

4.1 Current status of the measurements

In 1996 CLEO reported on the exclusive branching ratios of $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$; the decay $B \to \omega\ell\nu$ was also studied [36] [37] [38]. To get to those measurements isospin relations were used and also the fact that the quark model predicts similar decay rates for $B^+ \to \omega\ell^+\nu$ and $B^+ \to \rho^0\ell^+\nu$.

$$
\begin{aligned}
\Gamma(B^0 \to \pi^-\ell^+\nu) &= 2\Gamma(B^+ \to \pi^0\ell^+\nu) \\
\Gamma(B^0 \to \rho^-\ell^+\nu) &= 2\Gamma(B^+ \to \rho^0\ell^+\nu) \\
&\sim 2\Gamma(B^+ \to \omega\ell^+\nu)
\end{aligned}
$$

In 1998 CLEO released an updated measurement of the $B \to \rho\ell\nu$ branching ratio and $|V_{ub}|$ [39]. That analysis also looked at the q^2 distribution of the decay and extracted the form factor slope. The method used then was slightly different than the one used in the first measurement. Also, because of the very strong lepton momentum cut used, the q^2 distribution could not discriminate very much between different theoretical models.

Another approach to extract $|V_{ub}|$ is to look at inclusive measurements $B \to X_u \ell \nu$, where the X_u are all the possible final states. Unfortunately, because of the very large backgrounds, one has to put a very strong cut on the lepton momentum, which induces much model dependence uncertainty on the measurement. There are interesting ideas about using other kinematic variables like q^2 and the mass of the recoiling hadronic system, which could reduce the model dependence ([40]).

Our analysis updates the original $B \to \pi, \rho, \omega \ell \nu$ analysis with the addition of $B \to \eta \ell \nu$ decay. We get a precise value for $|V_{ub}|$. We also extract the q^2 distribution for the various decays. The method uses the neutrino reconstruction technique, that was used in the previous exclusive $b \to u \ell \nu$ analyses as well as in a measurement of the branching fraction of $B \to D \ell \nu$ [41].

The following section describes the event selection cuts required for our particular analysis. The following two chapters will describe the fitting technique and the results, as well as the systematic uncertainties, pertaining to the branching ratio measurements. Appendix A details the various tools needed to perform a neutrino reconstruction analysis.

4.2 Selecting $B \to X_u \ell \nu$ events

4.2.1 Overview of the analysis

Before going into the details of how to select events representing $B \to X_u \ell \nu$ candidates, we lay down the road map of the analysis. As mentioned earlier, we assign a neutrino 4-momentum based on the missing energy and momentum of the entire event. The track fitting of Trackman approved tracks provides the best representation of the energy and momentum for the charged particles, while the Splitoff approved showers from the calorimeter provide the best representation for the energy and momentum of the neutral particles. We define "good" tracks, from which we compute the missing momentum, to be Trackman-approved, and also have a total momentum less than 5.0GeV/c. Similarly, "good" clusters are Splitoff-approved clusters. For vertices, like $K_s \rightarrow \pi^+\pi^-$ or $\Lambda \rightarrow p\pi^-$ or photon conversions, the vertex information from the kinematic fit provides the best representation of the momentum and energy for those particles since the track parameters assume the particle came from the beam point.

We combine the approved tracks and showers with the beam information to obtain the neutrino 4-momentum as follows,

$$
\begin{aligned}
p_{sum} &= \sum_{tracks} p_i + \sum_{clusters} p_i + \sum_{vertices} p_i \\
E_{miss} &= 2E_{beam} - p_{sum}(4) \\
\vec{p}_{miss} &= p_c r\vec{os}sing - \vec{p}_{sum} \\
p_{miss}(4) &= E_{miss} \\
p_\nu &= (|\vec{p}_{miss}|, \vec{p}_{miss}).
\end{aligned}
$$

where $p_c r\vec{os}sing$ accounts for the small crossing angle. We assign the neutrino energy to be the magnitude of the missing momentum, because the missing momentum resolution is significantly better than the missing energy resolution. This comes from the fact that any mistakes can only be additive in the scalar energy case, while they can potentially cancel out in the vector momentum case. We show in figure 4.1 the resolution for the neutrino momentum and energy. We show both the cases where there are no extra particles in the event, from which we extract the intrinsic resolution (FWHM/2.36), and the case where such particles are present. Extra particles refer to K_L, neutrons, and extra neutrinos. For K_L and neutrons, those particles sometimes interact partially in the calorimeter, which then reduce the error in the missing energy and momentum. The intrinsic resolution is similar for the different hadronic signal modes, and also similar between CLEO II and II.V. The missing momentum resolution is significantly better than for non kalman-fitted data.

$$
\begin{aligned}
\sigma_{E_{miss}} &\approx 250 MeV \\
\sigma_{\vec{p}_{miss}} &\approx 85 MeV
\end{aligned}
$$

Now that we have a neutrino 4-momentum, we can combine it with an identified lepton, and a reconstructed hadron to make a B candidate. We consider the beam constrained mass (refered to as M_B), and the energy difference ΔE defined as,

$$
\begin{aligned}
\Delta E &= (E_\nu + E_\ell + E_{had}) - E_{beam} \\
M_B &= \sqrt{(E_{beam}^2 - |\alpha\vec{p}_\nu + \vec{p}_\ell + \vec{p}_{had}|^2)} \\
\alpha &= 1 - \frac{\Delta E}{E_\nu}.
\end{aligned}
$$

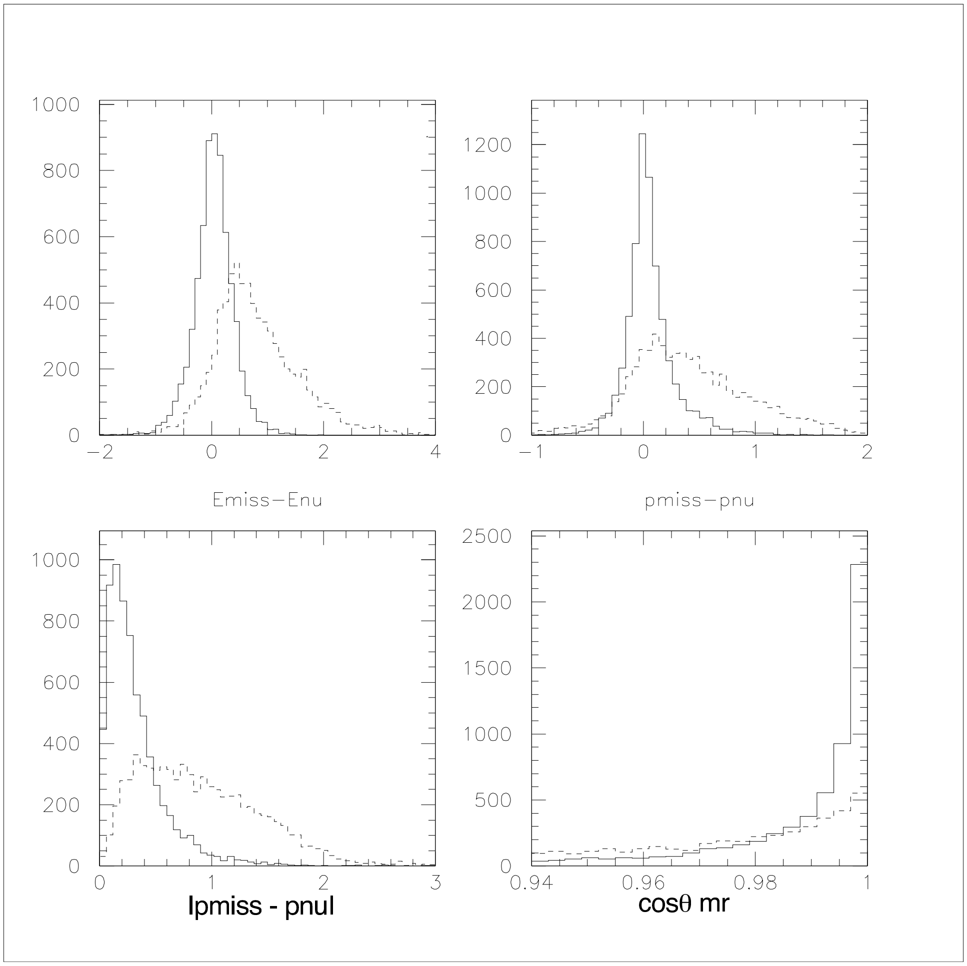

Figure 4.1: Neutrino resolution for signal π MC. Top left is $E_{miss} - E_{\nu}$, Top right is $|\vec{p_{miss}}| - |\vec{p_{\nu}}|$, bottom left is $|\vec{p_{miss}} - \vec{p_{\nu}}|$ and bottom right is the angle between the reconstructed and generated neutrino. The solid curve shows events for which there are no extra particles; the dashed curve shows events that had at least one extra particle. The normalization is absolute.

Table 4.1: GSR and SR boundaries

GSR	$5.175 \leq M_B \leq 5.3025$	$-0.75 \leq \Delta E \leq 0.25$
SR	$5.265 \leq M_B \leq 5.2875$	$-0.15 \leq \Delta E \leq 0.25$

For perfect energy and momentum conservation we would have $\Delta E = 0$ and the beam constrained mass equal the B hadron mass. The factor α enforces energy conservation ($\Delta E = 0$), by scaling the neutrino momentum to close the vector sum made by the lepton, the hadron and the B momentum magnitude (that we get from the beam constraint).

To extract the different branching fraction of the various hadron modes, we fit the ΔE vs M_B distribution over a so called Grand Signal Region (GSR), while extracting the yields in the Signal Region (SR), defined in table 4.1. Figure 4.2 shows the Grand Signal Region and the delimited Signal Region for the π MC.

The ΔE window is asymmetric because the signal tends to be asymmetric and the $b \to c\ell\nu$ (referred from now on as btoc) background tends to reconstruct at negative ΔE values, as we can see in figures 4.3 and 4.4. The signal is asymmetric because hadronic splitoffs tend to increase the energy associated with hadrons, and missing particles tend to increase the reconstructed neutrino energy. The btoc background piles up at negative ΔE values because the lepton momentum is typically softer for btoc events. When we have extra particles they increase the neutrino momentum, which increases ΔE which in turn can smear a btoc event into the signal region.The ρ cross-feed into π piles up at negative ΔE values because

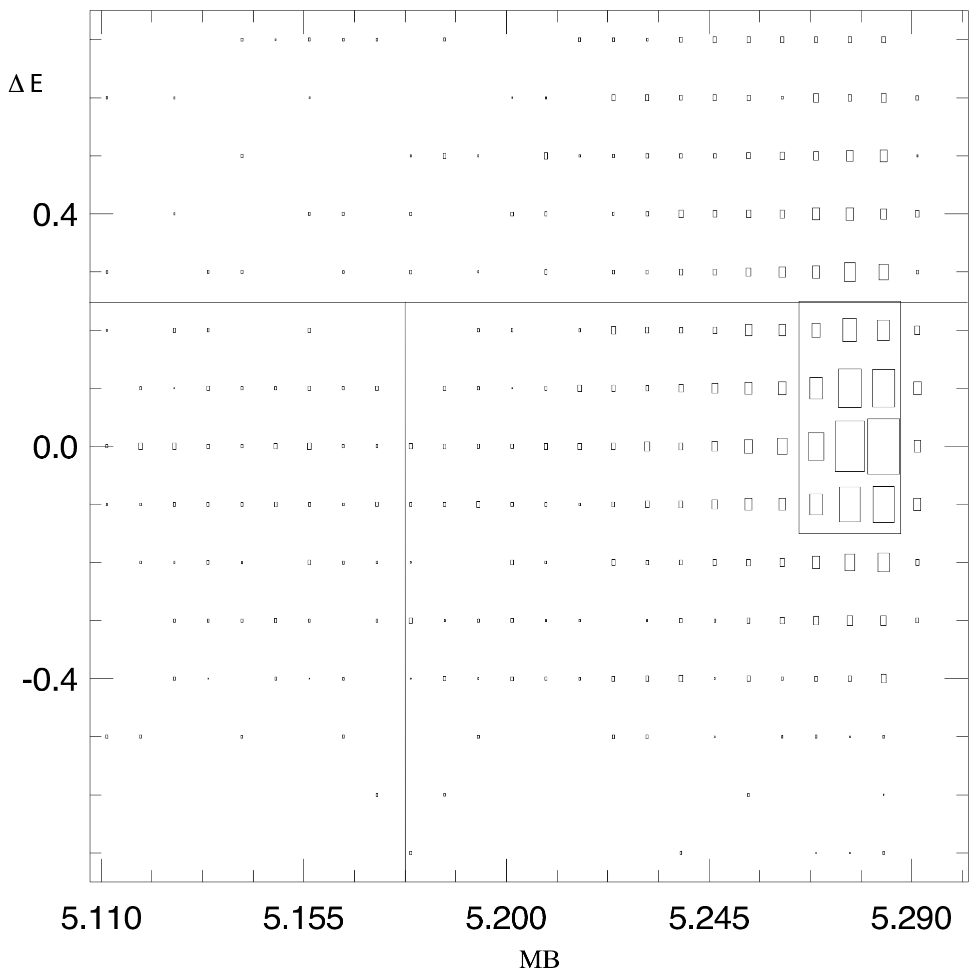

Figure 4.2: ΔE vs M_B for signal π ISGW2 MC.

the signal π is likely to be one of the π coming from the real ρ, and so the event is missing the other π energy to get $\Delta E = 0$. Conversely, in the ρ reconstruction case, the π cross-feed reconstructs at positive ΔE values because the true π, neutrino and lepton energy already add up to E_{beam}, so adding in a random π to build a ρ makes $\Delta E > 0$. Also, the π momentum is in general harder than the ρ one, as can be seen in figure 4.5

Aside from measuring the various branching fraction of the different signal modes, we also measure the q^2 distribution, which will allow discrimination among various theoretical form factors models and so reduce the model dependence uncertainty on $|V_{ub}|$. Having the neutrino momentum helps the resolution of q^2, particularly if we scale the neutrino momentum by the factor α defined earlier,

$$q^2 = (\alpha p_\ell + p_\nu)^2. \tag{4.1}$$

For π signal MC we get a q^2 resolution of about 310 MeV^2 when we use the factor α, 540 MeV^2 otherwise. Given the statistics for the different modes, we choose to fit in three q^2 bins. The width of each q^2 bin is the same for all modes, 8 MeV^2. The plots shown in figures 4.6 and 4.7 show that the area in each q^2 bin according to ISGW2 is given in table 4.2. The ω and the η modes are not statistically significant enough to be fitted in three q^2 bins, so for those modes, the entire q^2 region is fitted.

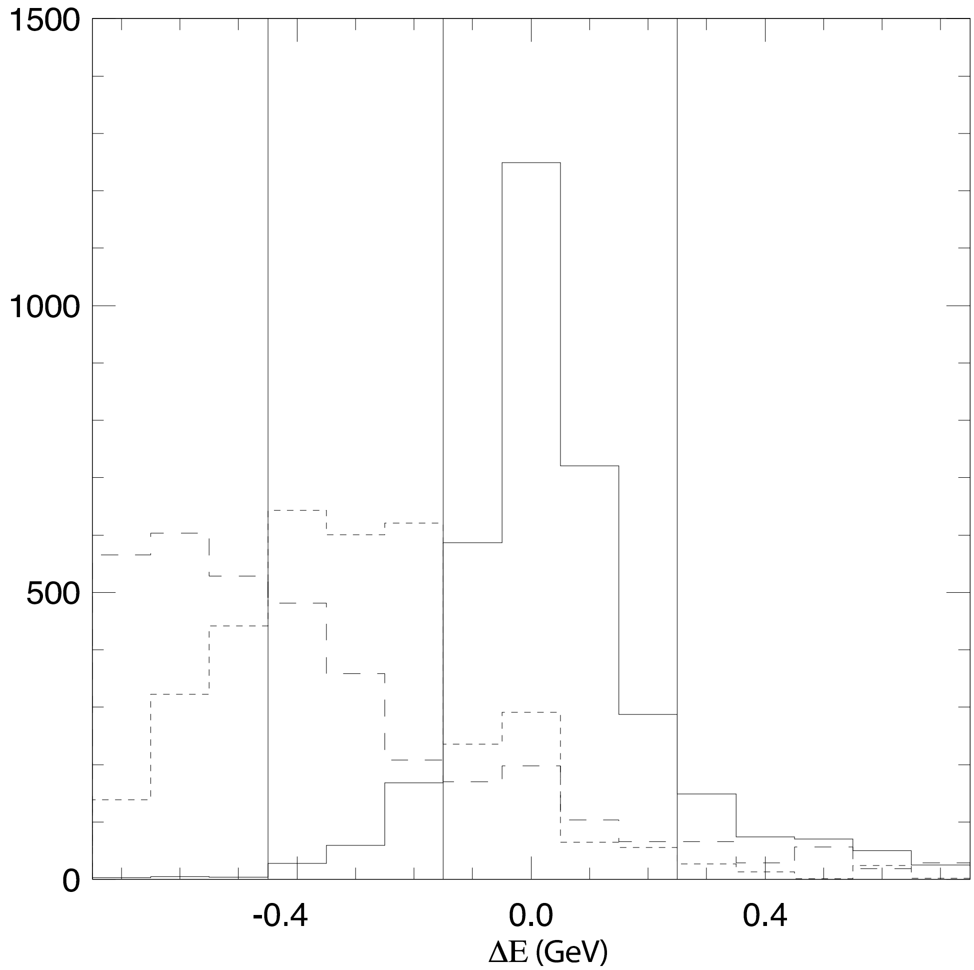

Figure 4.3: ΔE in the M_B signal region for π. Solid is π signal MC, short dashed is ρ cross-feed and long dashed is btoc MC. The vertical lines represent the various bins as described in the next section. The normalizations are arbitrary.

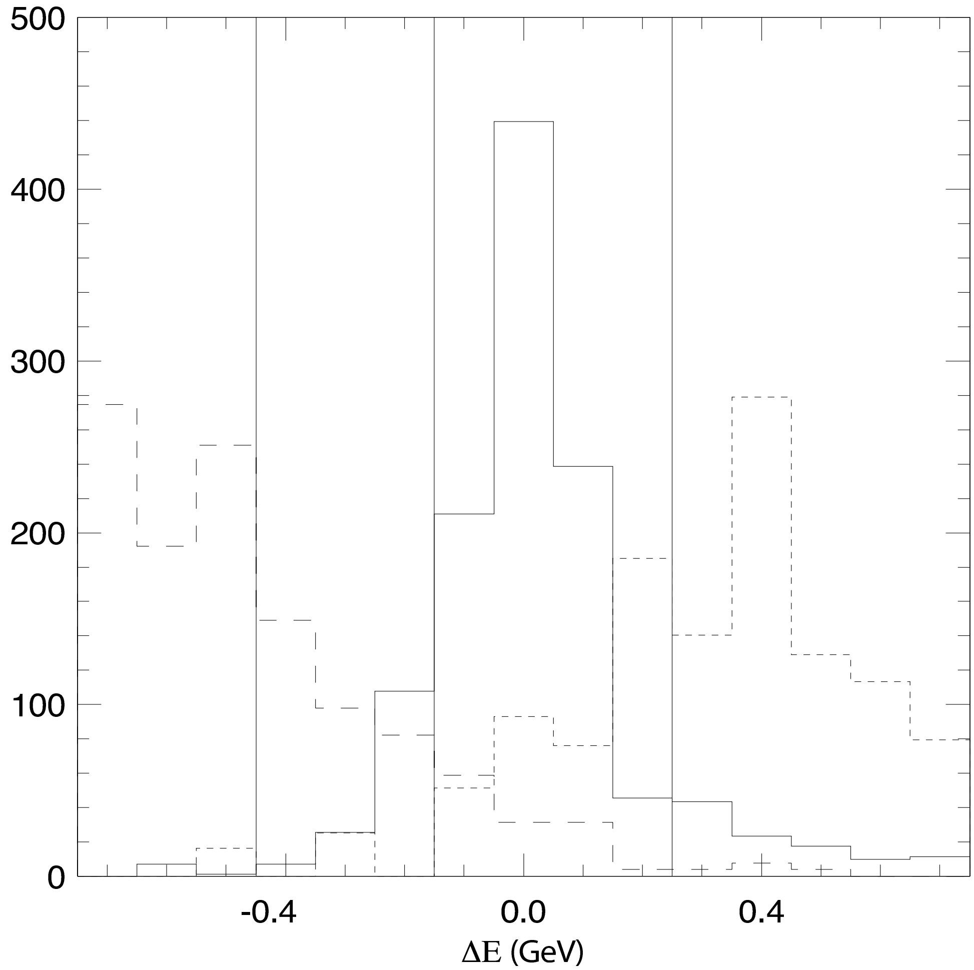

Figure 4.4: ΔE in the M_B signal region for ρ. Solid is ρ signal MC, short dashed is π cross-feed and long dashed is btoc MC. The vertical lines represent the various bins as described in the next section. The normalizations are arbitrary.

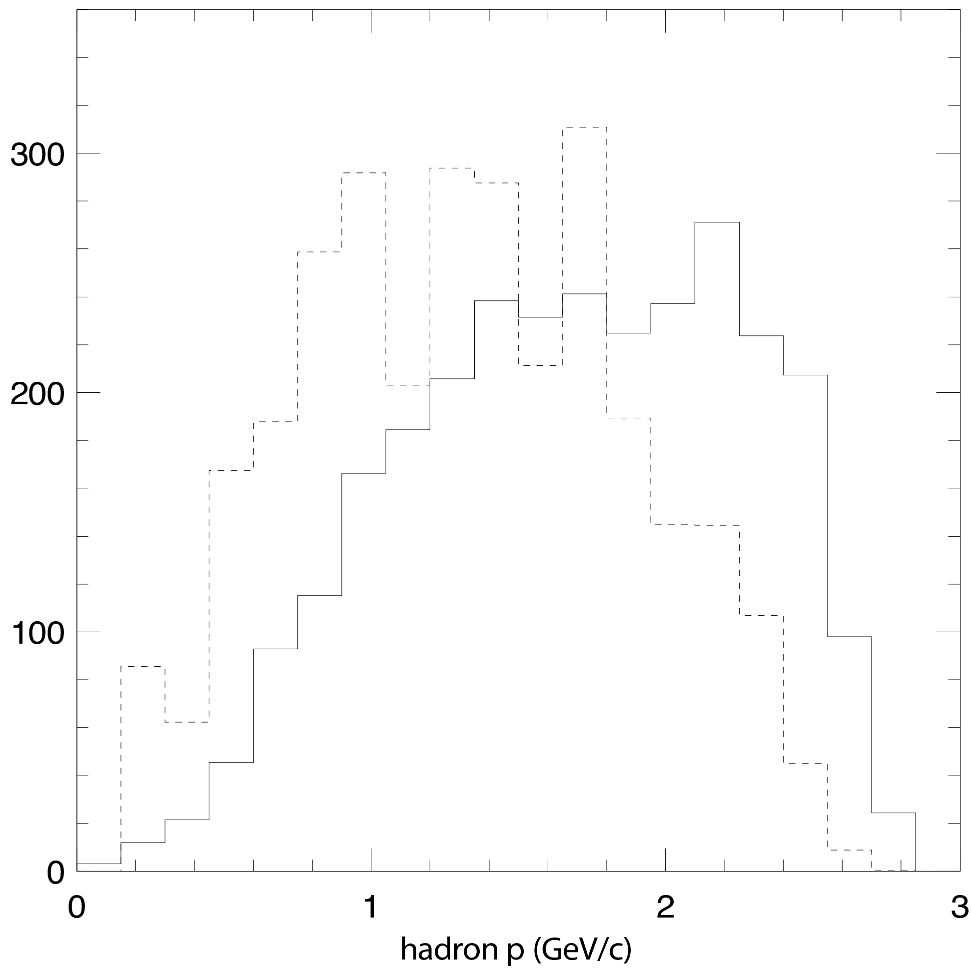

Figure 4.5: Solid is π momentum and dashed is ρ momentum. The normalization is arbitrary.

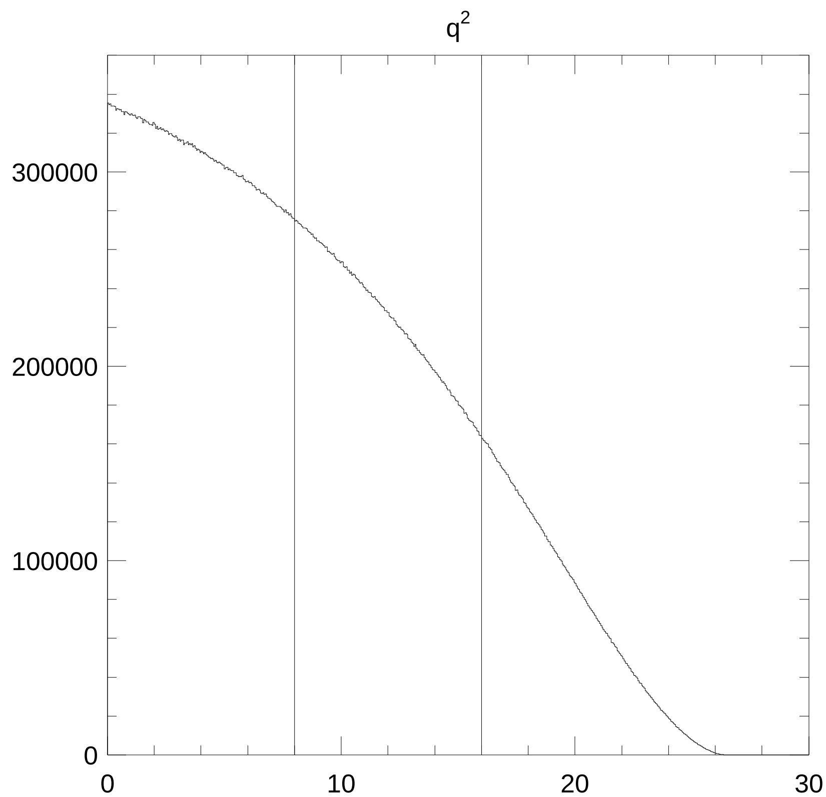

Figure 4.6: $\frac{d\Gamma}{dq^2}$ for the π mode according to ISGW2. The vertical lines show the different q^2 bins.

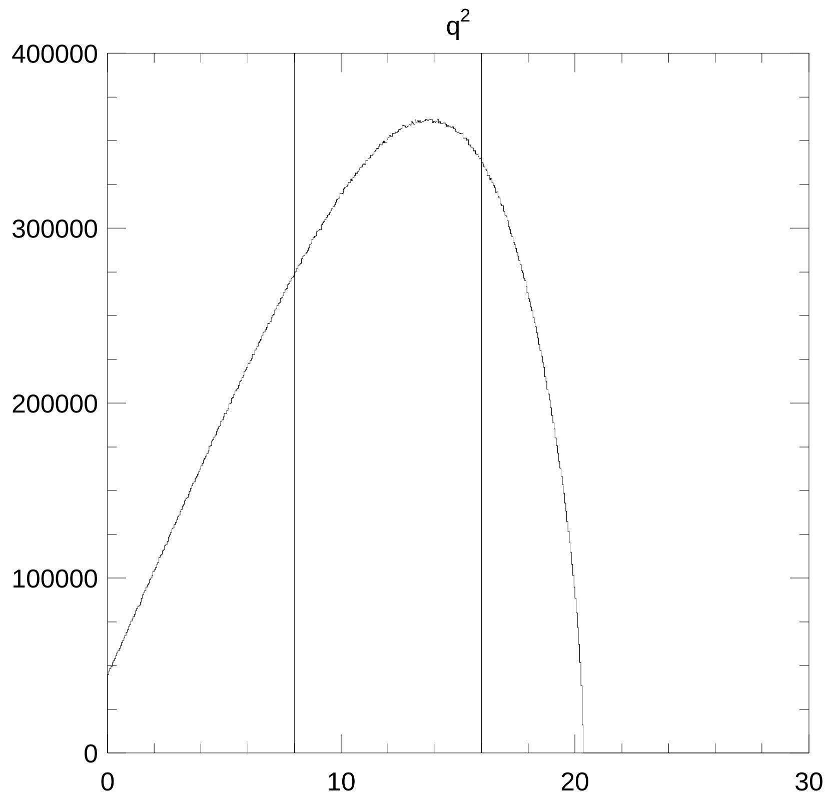

Figure 4.7: $\frac{d\Gamma}{dq^2}$ for the ρ mode according to ISGW2. The vertical lines show the different q^2 bins.

Table 4.2: Fraction of events in each q^2 bin according to ISGW2

mode	$q_1^2(0-8GeV^2)$	$q_2^2(8-16GeV^2)$	$q_3^2(16-q_{max}^2GeV^2)$
π	49%	36 %	15 %
ρ	26%	54%	20 %

4.2.2 Cuts and backgrounds

We now describe the cuts that eliminate the various backgrounds and select our signal events. The cuts, summarized at the end of the section, were optimized using signal MC for all the signal modes, as well as some generic $B\bar{B}$ MC (not containing any $b \to u\ell\nu$ events), some continuum MC, and some $b \to u\ell\nu$ inclusive MC (not containing our signal events), all independent from the fitting samples. We looked at the figure of merit for the charged π and the charged ρ modes. The figure of merit (fom), defined as

$$fom = \frac{S^2}{S + B_{xfeed} + 1.2B_{btoc} + 3B_{cont} + B_{btou}}, \tag{4.2}$$

is the ratio of the amount of signal events (S) over the total statistical uncertainty. B_{xfeed} refers to the amount of the other signal modes that are feeding through the particular signal mode that we are studying, B_{btoc} refers to the generic $B\bar{B}$ background, B_{cont} refers to the continuum background, and B_{btou} refers to the other $b \to u\ell\nu$ modes that are feeding down into the signal mode. The factors of 1.2 and 3 accounts for the statistics of those samples used to model those backgrounds, since we have five times the data sample for the generic $B\bar{B}$ MC, and half the data sample for off resonance running.

Global Event cuts

In order to require hadronic events, we ask for $Klas = 10$ events. The Klas code determines what kind of physics the event contains based on information like the number of tracks and their momenta, other possibilies are beam gas events, radiative bhabhas, etc.

We have already mentioned that the criteria for selecting good tracks is that they be Trackman-approved. In the case that a track is Trackman-approved but had a bad z-fit, the event is poorly reconstructed and the neutrino momentum is probably not reliable so we get rid of the event.

In a perfect event, Trackman selects all the good tracks representing the particles and we should get a net charge (ΔQ) of the event equal to zero. There is some justification for accepting events that have $\Delta Q = -Q_{lep}$, a net charge of the opposite polarity of the signal lepton in the event. The reason for this is that the tracking efficiency falls off for soft pions. If the other B in the event decayed into a D^* meson, which decayed into a D meson, the accompanying pion is soft and this signal event might very well have net charge different than zero. Those events represent 20% of all the mistakes that happen in the events for which net charge is +1, or -1. We looked a the fom for the cases of $\Delta Q = 0, -Q_{lep}, +Q_{lep}$ and $|\Delta Q| = 1$, and it turns out that accepting $|\Delta Q| = 1$ in addition to $\Delta Q = 0$ gives the best figure of merit.

Once we have constructed a neutrino 4-momentum, as described above, we require the polar angle of the momentum vector to satisfy $|cos\theta_\nu| < 0.96$. If particles excaped down the beam pipe, their momenta would be included in the missing momentum.

Charged lepton requirements

A powerful part of selecting the exclusive signal events is the choice of the signal charged lepton. Since the u quark is lighter than the c quark, the hadrons in a $b \to u$ transition are lighter than the hadrons produced in a $b \to c$ transition, which in turn, produces leptons of higher momentum. In fact, the first evidence for the $b \to u$ transition was achieved by looking at the end-point of the lepton spectrum ([45]). This means that a hard lepton cut eliminates much of the btoc background, which is one of the most important backgrounds of the analysis. This cut comes with a price: in addition to having a reduced signal efficiency, we must rely on models to correct for our acceptance over only a limited lepton momentum spectrum. The shape predicted by the different form factor models affect the branching fractions and the $d\Gamma/dq^2$ distribution, although the associated uncertainty is small in comparison to the other uncertainties on those measurements. The dominant uncertainty on $|V_{ub}|$ comes from the normalization of the form factor shape. On the other hand, getting a precise $d\Gamma/dq^2$ distribution will help determine the various QCD calculations which will in turn put more constraints on the possible normalizations. Since it is one of the goals of this analysis to discriminate among models, it is worth trying to keep the lepton cut as low as possible.

For muons, we have different selection criteria depending on whether the muon is the only lepton in the event or not. At first, we call a muon any track that satisfies the following criteria:

$$\begin{aligned} mudepth(track) &\geq 5 \\ |\vec{p}| &> 1.2 \\ |cos\theta| &< 0.85 \\ mod(muquality(track), 10000) &= 0 \end{aligned}$$

or

$$\begin{aligned} 3 \leq mudepth(track) &< 5 \\ 1.0 \leq |\vec{p}| &< 1.75 \\ |cos\theta| &< 0.71 \\ muquality(track) &= 0 \end{aligned}$$

where mudepth refers to the number of interactions lenghts of material before the muons would range out. Muquality refers to whether there were missing hits within a subregion of the muon detector. After we have made sure that only one lepton is in the event, we ask if the muon satisfies the first set of requirements, these are the signal muons.

In this analysis we use a package [1] which makes an overall likelihood based on the ratio of energy of the shower over the momentum of the track, the amount of separation between the track and the shower and the dE/dx information from the track. The efficiencies were determined using embedded radiative bhabha events and 2 photon events. It was found that the radiative bhabha sample suffered from some background contamination, but requiring the tracks to have 40% of possible hits improved the purity of the sample. We make that requirement both in the data and in the MC. The loss of efficiency from the embedding goes from 6% at low momentum to 3% for high momentum tracks, in the good barrel portion of the calorimeter. The fake rates from hadrons were determined using a K_S data skim for pions and a D^* data skim for kaons. One has to take into account both the efficiency and the fake rate for each momentum bin in order to decide on the adequate electron identification cut. The efficiency of identifying an electron as a function of momentum is shown in figure 4.8. The fake rates are discussed in the next section.

Now that we have identified leptons, we can count them. We require that there is only one lepton in the event. If there was an additional lepton, there would likely be an additional neutrino, and that extra particle momentum would distort the signal neutrino momentum. We also require that the lepton track be a “good” track, as defined previously. The kinematics of the decay require the lepton momentum to be lower than 2.84 GeV/c. As for the lower limit, it will be a mode dependent value since the lepton momentum spectrum depends strongly

[1] The Rochester Electron ID package (REID)

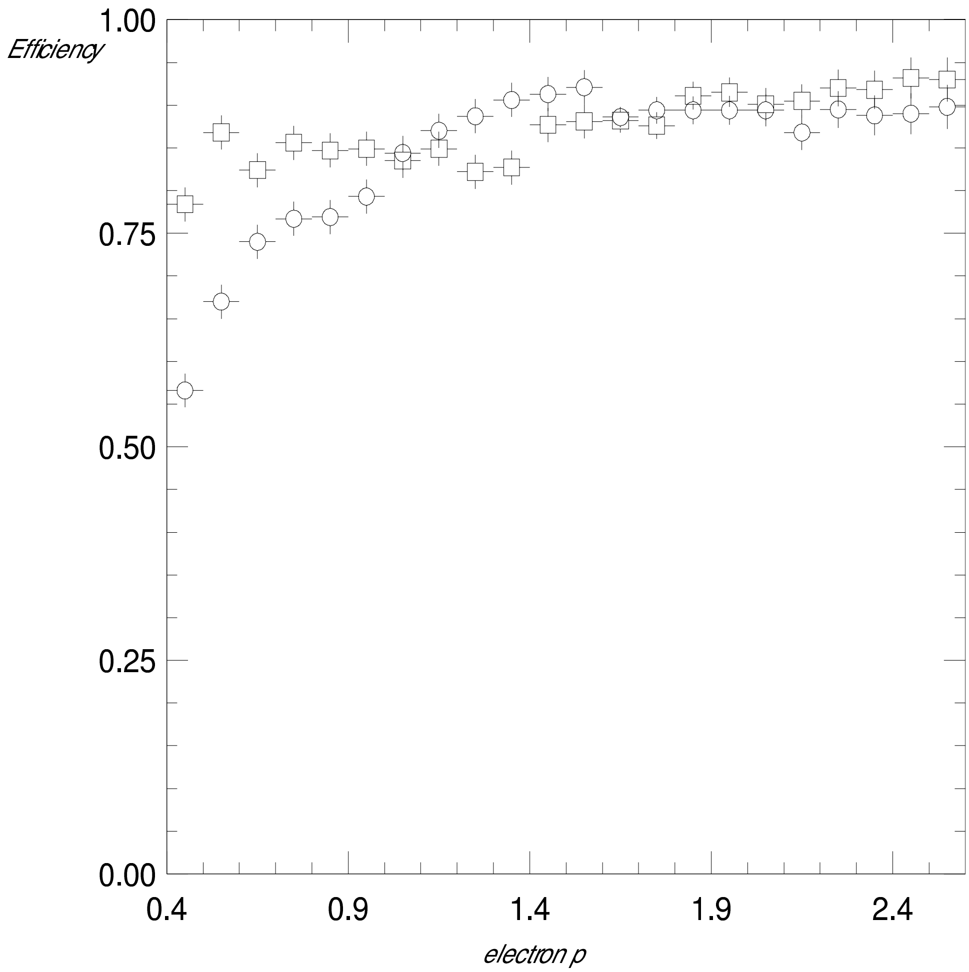

Figure 4.8: Efficiency vs electron momentum. Square is for the Good Barrel, circle is for the Non Good Barrel.

Table 4.3: Lepton momentum requirements

Modes	Lepton cut
π,π^0,η	1.0GeV/c
ρ,ρ^0,ω	1.5 GeV/c

Table 4.4: ψ and $\psi(2S)$ mass windows

mode	ψ (m=3.09688)	$\psi(2S)$ (m=3.686)
ee	3.02-3.13	3.675-3.705
$\mu\mu$	3.06-3.13	3.675-3.705

in the hadron spin. The lepton momentum cuts for the various modes are shown next in table 4.3.

Figure 4.9 shows the lepton momentum distribution for π signal MC, ρ signal MC and btoc MC.

A potential background for $B \to \pi\ell\nu$ comes from $B \to c\bar{c}K_L$, in which the $c\bar{c}$ meson decays to a pair of lepton and one of the lepton is misidentifies. The K_L plays the role of the neutrino. To eliminate such events, we combine the signal lepton with every other track of opposite charge in the event and see if it falls within the following mass windows for the ψ and $\psi(2S)$:

Finally, we consider $cos\theta_{lep}$, where θ_{lep} is the angle between the lepton in the W rest frame and the direction of the W in the B rest frame. This angle is strongly

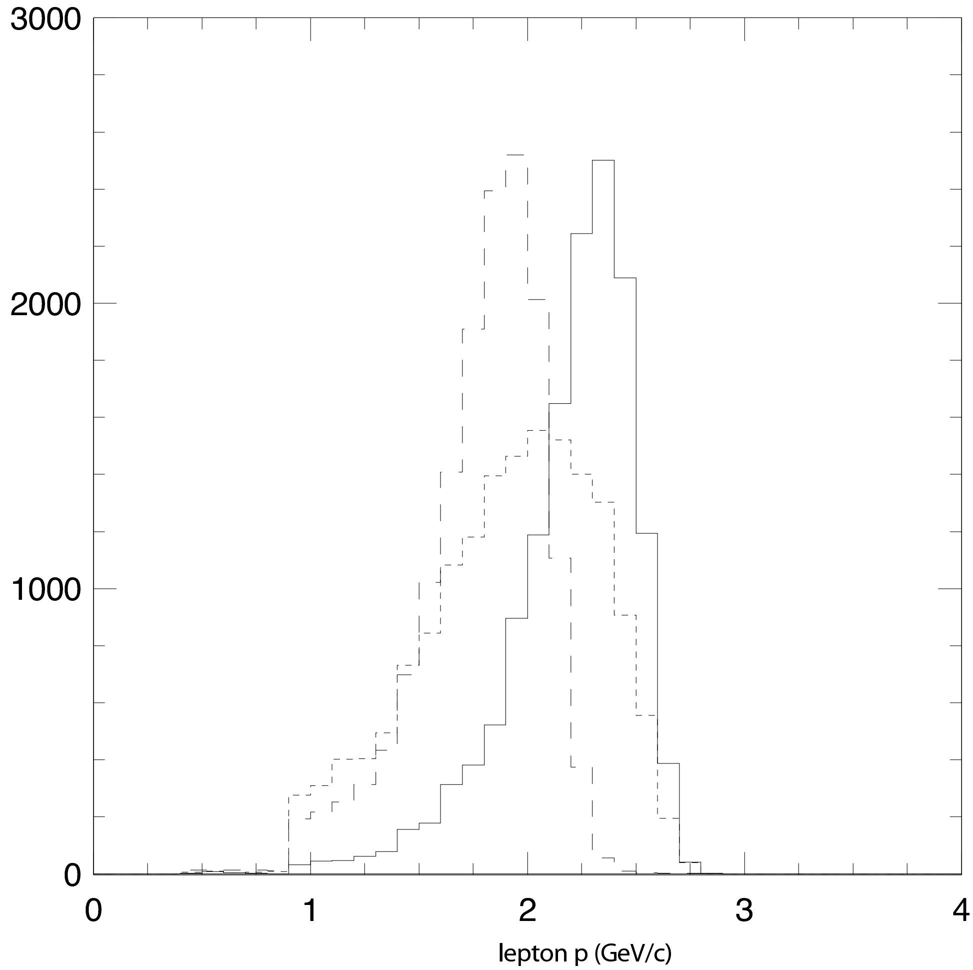

Figure 4.9: Lepton momentum distribution. Solid is ρ signal MC, short dashed is π signal MC, and long dashed is btoc MC. The normalizations are arbitrary

correlated to the dynamics of the decay. Background is expected to be roughly flat, while signal will have a distinctive shape that depends on the meson spin. In the pseudoscalar modes, the amplitude consists only of a longitudinal piece, the angle in this case follows a $sin^2\theta_{lep}$ distribution. In this case the signal and background distribution are not clearly separated. The situation is better in the vector modes since then the amplitude is dominated by the transverse polarization and the V-A structure of the weak current, and the resulting distribution is forward peaked.(see section 2.1.2 for a more detailed description of the dynamics of the decays and a prediction of the $cos\theta_{lep}$ distributions). We apply a $cos\theta_{lep} > 0$ in the vector modes, based on the figure of merit. We verified with various models that such a cut would not significantly bias the q^2 distribution. Figure 4.10 shows the $cos\theta_{lep}$ distribution for signal MC and btoc MC.

Fakes

Particle identification is not perfect and misidentification can lead to two possibilities: leptons can get identified as hadrons and hadrons can get identified as leptons. The first case comes from the inefficiency of the lepton cuts. It can only remove events from the signal region. The second case comes from the fact that a pion, kaon or proton can mimic a lepton response in the detectors. For example, hadrons interact in the calorimeter, and the resulting showers occasionally look like an electromagnetic shower. Another phenomenon is the fact that around 900MeV/c of momentum kaons stop in the calorimeter, leaving a cluster of energy that results in E/p $\sim$ 1. This combined with the fact that kaons have an ioniza-

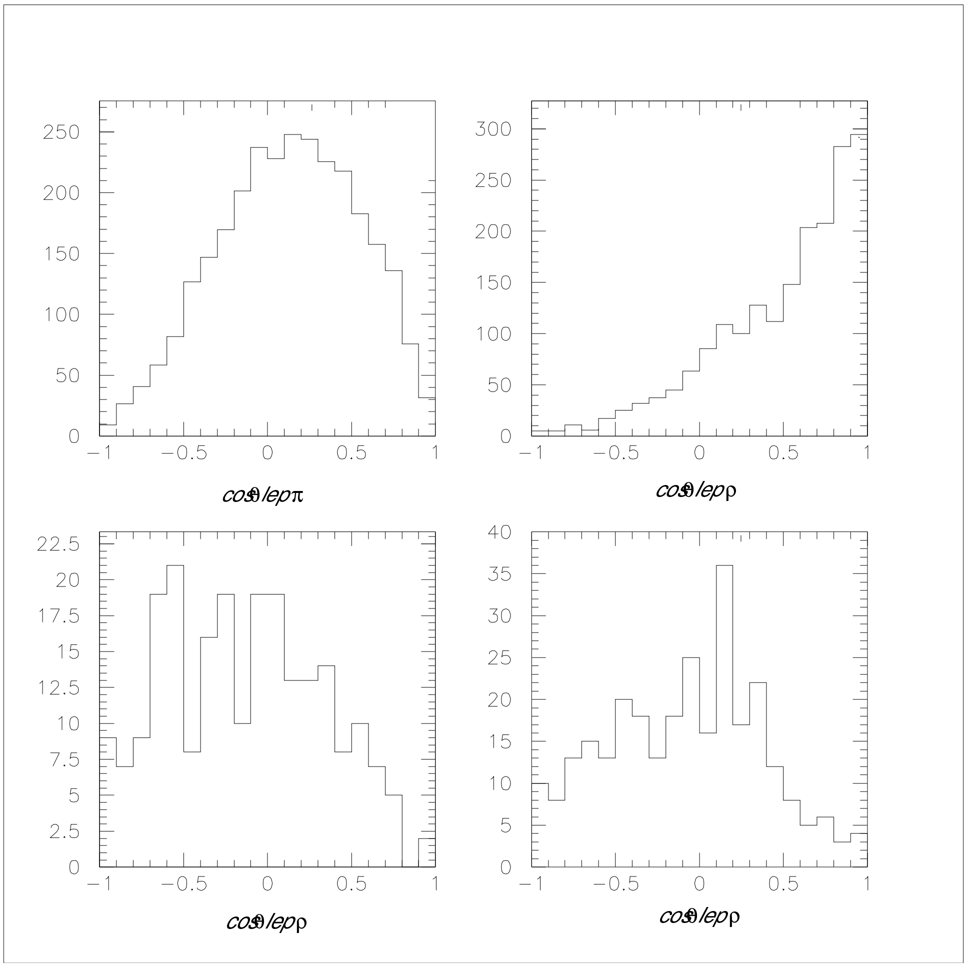

Figure 4.10: $cos\theta_{lep}$ distribution. Top right is for π signal MC, top left is forρ signal MC, bottom is btoc MC, left is CLEO II and right is II.V

tion de/dx band (see previous chapter) that merges with the electron band at near that momentum results in a significant probability for kaons to fake electrons. A veto based on time of flight, which discriminates well between kaons and electrons around 1.0GeV/c, significantly suppresses this fake rates. For that reason, we require some time of flight information in addition to the likelihood requirement for the data. Finally, anti-protons annihilate in the calorimeter, increasing their probability to fake electrons. Consequently we only use the anti-proton fake probability [2]. The fake probability for pions, kaons and protons are shown in figures 4.11 and 4.12, as a function of momentum for the electron case. Figures 4.13 and 4.14 show the case of the muon, both for the conditions of faking a muon that we use for signal, and for faking a muon used in the multiple lepton veto. We distinguish between two regions of the calorimeter: the Good Barrel and the Non-Good Barrel. The former corresponds to the region of the calorimeter with $|cos\theta| < 0.7$ while the latter corresponds to $|cos\theta| > 0.7$.

Having the possibility of a hadron faking a lepton might make a nonleptonic data event fall into our signal region. We determine the backgroundfrom hadrons faking the signal lepton by using data, not MC. For that reason, we run on non-leptonic data and treat successively each track as the signal lepton track and give this combination a weight equal to the fake rate probability of this track. Since in data we do not know for sure if the hadron was a pion, kaon or proton, we look in MC what is the relative population of each species as a function of momentum.

[2] We use the results from the Minnesota study ([46]) to estimate our fake probability.

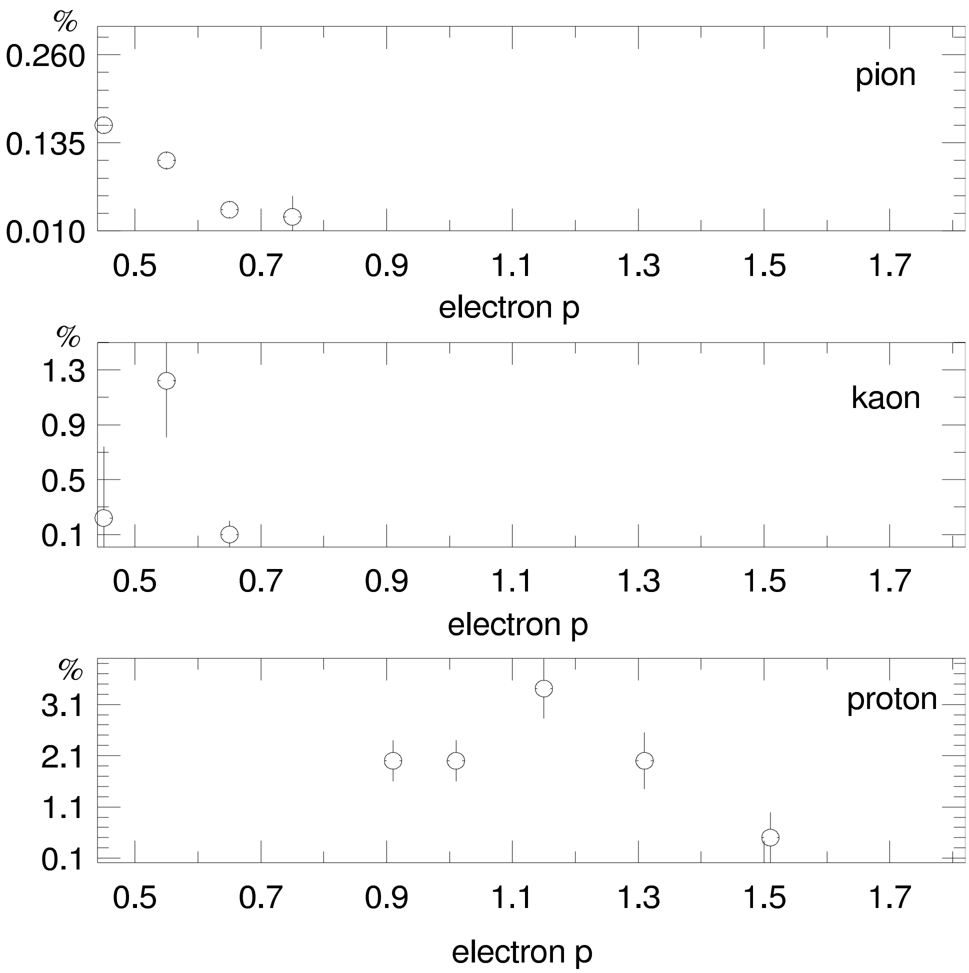

Figure 4.11: Fake rates for electrons vs momentum, in the Good Barrel.

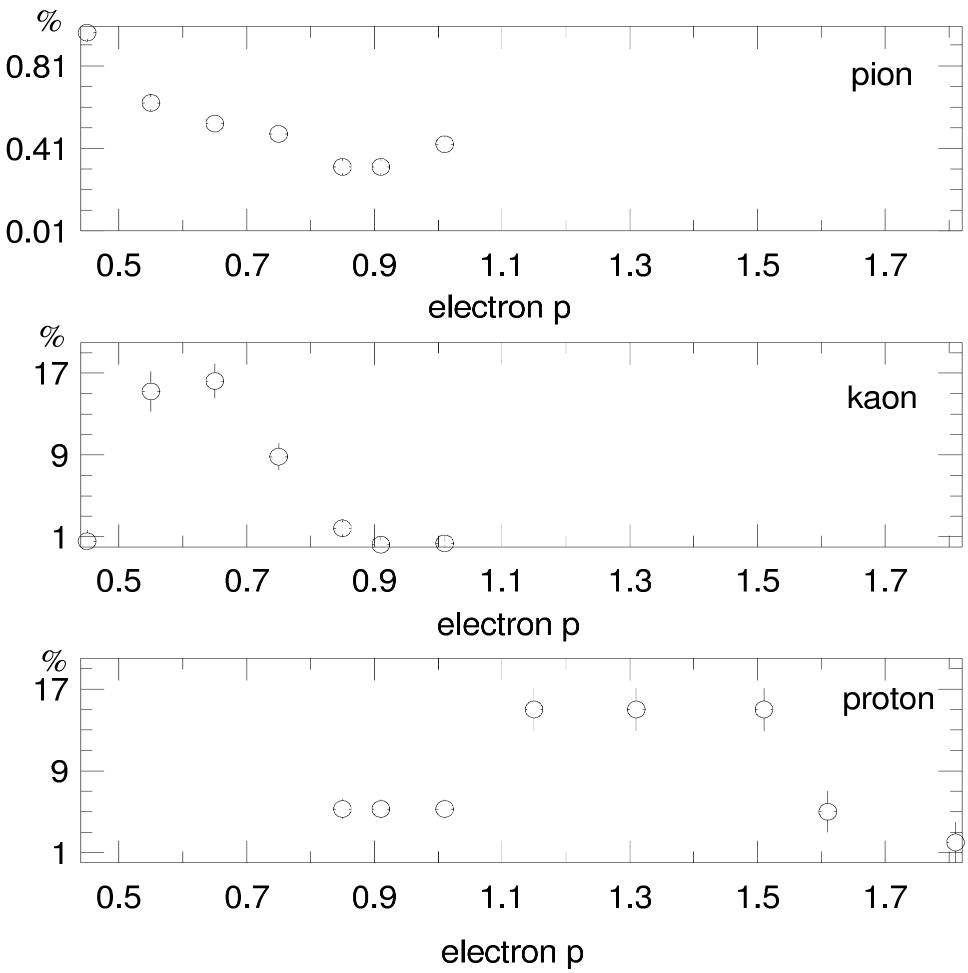

Figure 4.12: Fake rates for electrons vs momentum, in the Non Good Barrel.

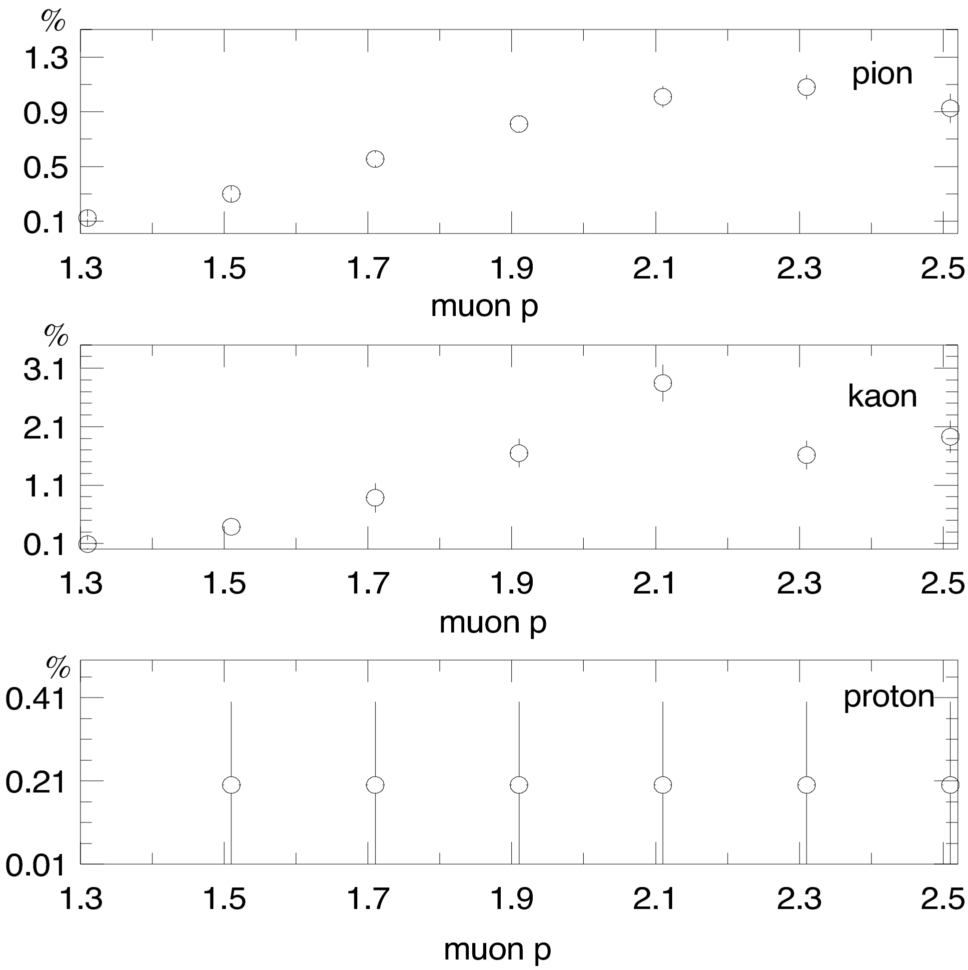

Figure 4.13: Fake rates for signal muons vs momentum.

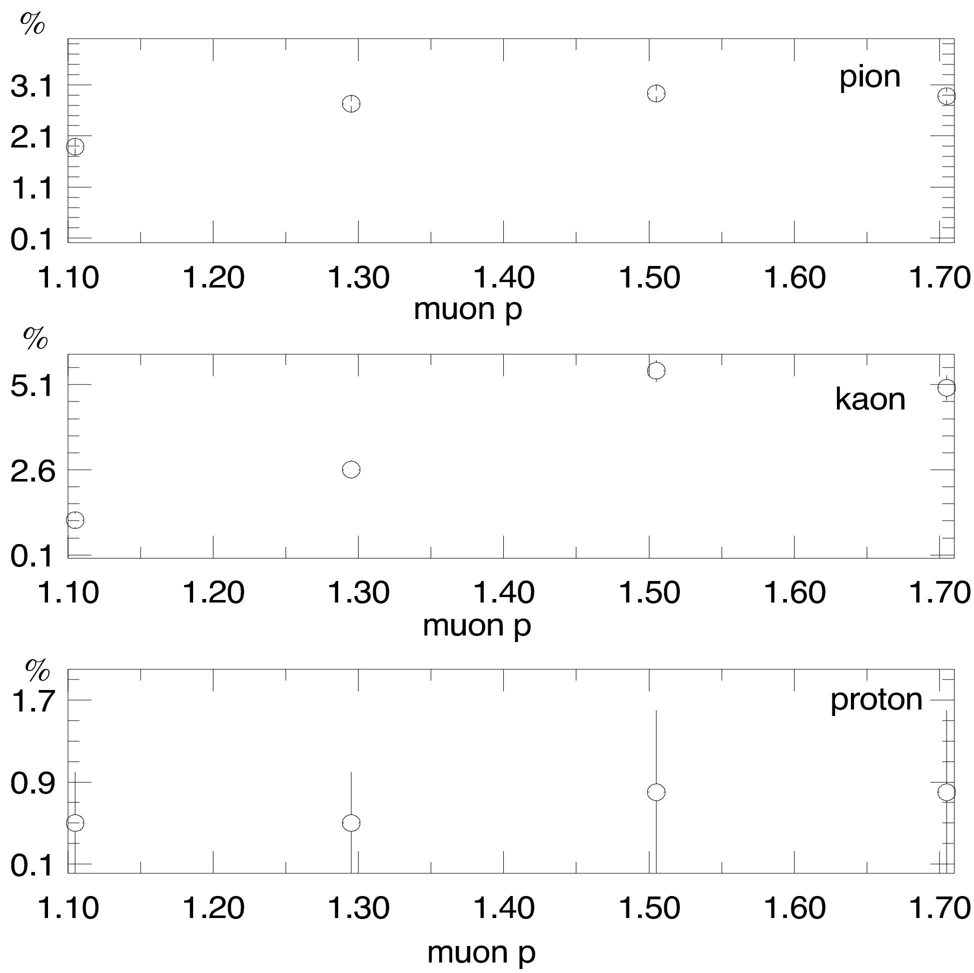

Figure 4.14: Fake rates for veto muons vs momentum.

Combining this information with the fake rate probability in momentum bin we get an average fake probability as a function of momentum.

A hadron can also fake a secondary lepton in semileptonic decay and result in the loss of events. To account for this in the semileptonic MC events containing an identified lepton, we compute the total fake probability of all the hadronic tracks in the event and compare this to a random number. If the random number is greater than the total fake probability, we reject the event. In MC we know with certainty the identity of the hadrons, so we do not need the average fake probability that we used in the nonleptonic data.

Hadron candidates requirements

At this point in the analysis, we have a signal neutrino and a signal lepton. We need a signal hadron in order to reconstruct the B meson. To identify hadrons, we use the combination of the de/dx information and the time of flight information, when they are available [3]. We reweight the probability of being a particular hadron based on the relative population of the hadrons as seen in the semileptonic $B\bar{B}$ MC. If the identification failed, the default identity of the track is a pion.

We reconstruct six $B \to X_u \ell\nu$ modes: $\pi^{\pm}$, π^0, $\rho^{\pm} \to \pi^{\pm}\pi^0$, $\rho^0 \to \pi^+\pi^-$, $\omega \to \pi^+\pi^-\pi^0$ and η. For the η mode we look at two decay channels: $\eta \to \pi^+\pi^-\pi^0$ and $\eta \to \gamma\gamma$. Table 4.5 shows the multiplicty of tracks required for the different modes. These allow for some separation with the btoc background. η_p and η_g refer to the η decaying into three pions and two photons respectively. When the

[3]We use the Rochester's list of bad time of flight runs ([47]).

Table 4.5: Track multiplicity for each signal mode

Mode	Min. tracks	Max. tracks
$\pi^{\pm}$	4	10
π^0	4	8
$\rho^{\pm}$	4	10
ρ^0	6	10
ω	4	10
η_p	4	10
η_g	4	8

hadronic decay mode involves charged pion tracks, those tracks should be good tracks, and of opposite charge to the lepton in the case of the $\pi^{\pm}$ and $\rho^{\pm}$ mode, or of opposite charge to each other in the ρ^0, ω and η_p modes. The signal neutral pion mode requires that the signal π^0 momentum be smaller than 5 GeV/c. The ω mode has the additional requirement of a Dalitz plot cut of 0.4 on the amplitude squared relative to the maximum amplitude of the plot. The neutral pions requirements for the different modes are summarized in table 4.6

Each mode, except the charged pion mode, is labeled according to its mass. Table 4.7 summarizes the different mass windows for the different signal modes.

Figures 4.15 and 4.16 show the ρ and ω mass distribution respectively.

Table 4.6: Neutral pions requirements.

π^0 requirements	π^0 and η_g modes	$\rho^\pm$ mode	ω and η_p modes
Minimum π^0 Energy	0	325 MeV	225 MeV
Max. mass dev. (σ)	6	1.8	2
Max. χ^2	7	8	7
Splitof approved showers:	yes	yes	yes
Minimum shower Energy	30 MeV	30 MeV	30 MeV

Table 4.7: Mass information for the various signal modes.

requirements	π^0 and η_g	$\rho^\pm$ andρ^0	ω and η_p
mass of hadron (MeV):	134.976, 547.45	769.9	781.94, 547.45
resolution (MeV):	8, 13	8.5, 4	8.5, 6
natural width (MeV):	n/a	151	8.4, n/a
number of mass bands:	5	9	9, 5
mass width:	2 σ	190 MeV	20 MeV
central bands in fit:	1	3	3, 1

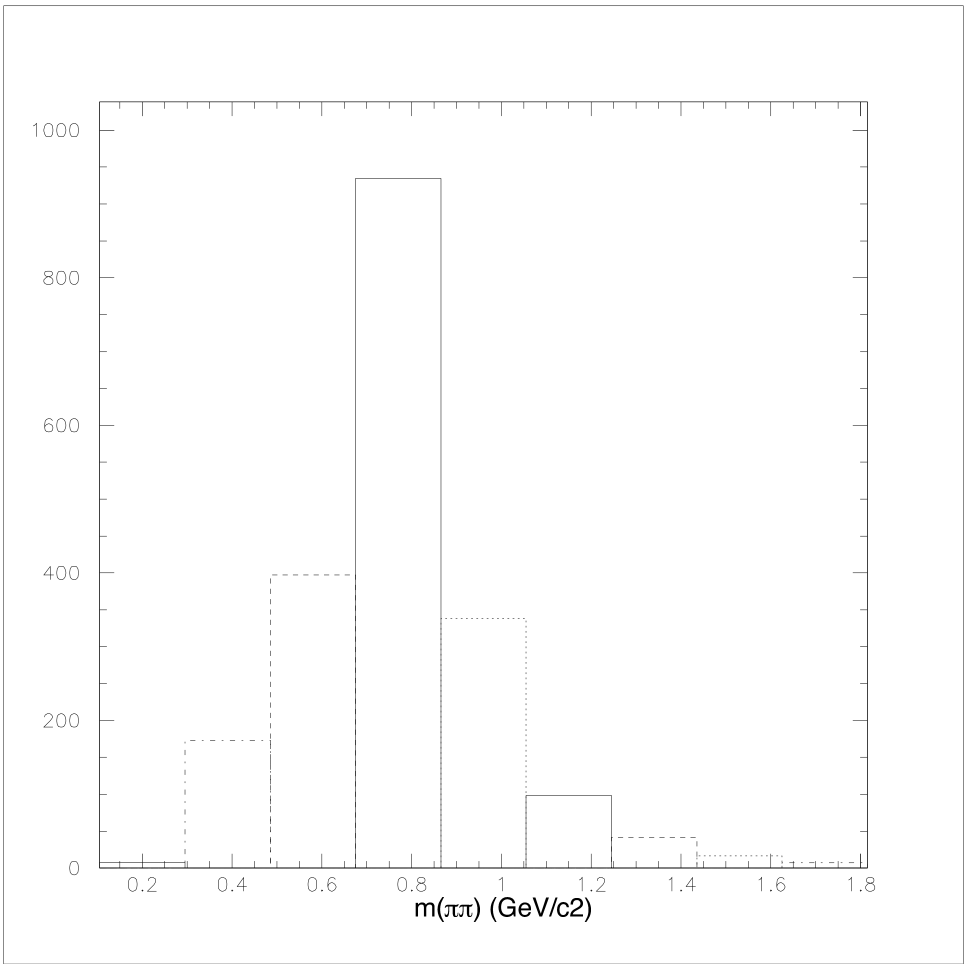

Figure 4.15: Reconstructed $\pi\pi$ mass distribution from $\rho^{\pm}$ signal MC. Each bin is a mass band.

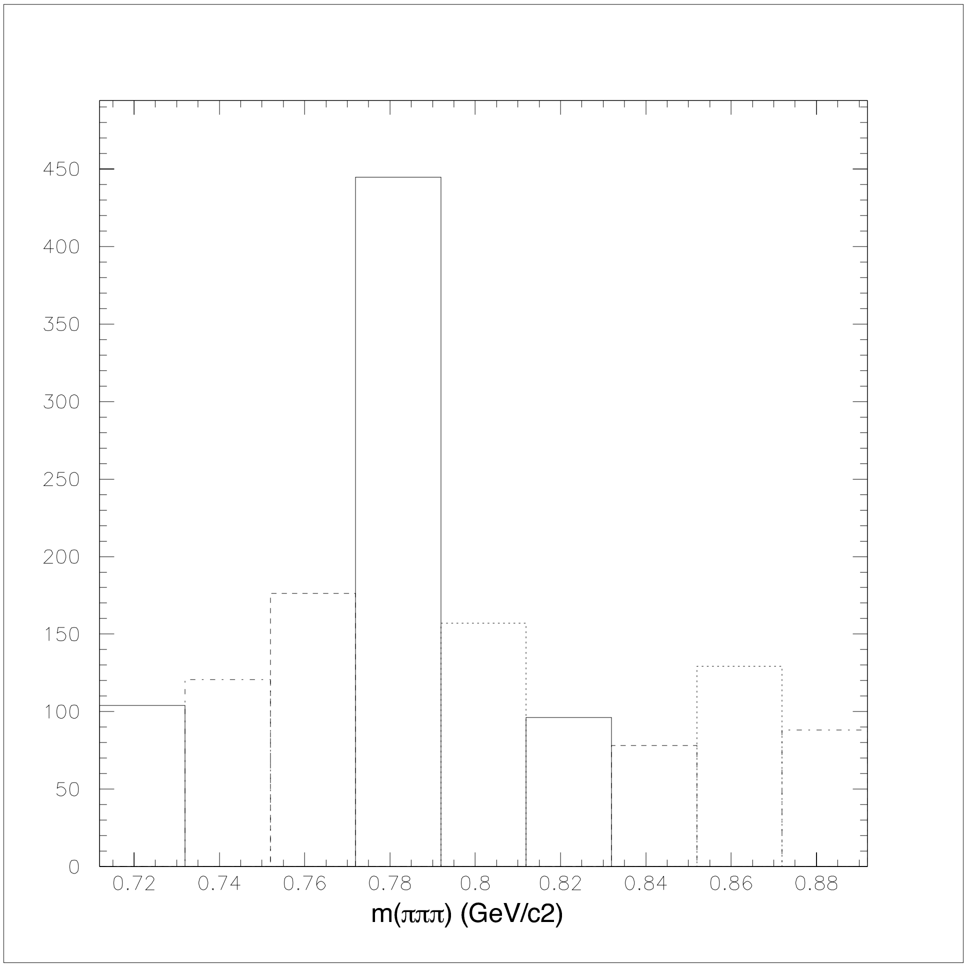

Figure 4.16: Reconstructed $\pi\pi\pi$ mass distribution from ω signal MC. Each bin is a mass band.

For the II.V sample, there is an additional requirement for each track entering the hadron candidate selection: if the track has no r-phi hits or no r-z hits, the track is rejected.

Once the hadron candidate is built, we evaluate if this combination passes the continuum suppression cut (describe in the next section). If it does then we compute the beam constrained mass using the signal lepton and neutrino. If the beam constrained mass is greater than 5.175 GeV, the candidate hadron is kept. We typically get more than one candidate per event for any given hadron mode, even in signal MC. We choose one signal candidate (per mode) based on $|\Delta E|$.

Continuum suppression

A signigicant background comes from the continuum ($e^+e^- \to q\bar{q}$) under the $\Upsilon(4S)$ resonance. One example of how continuum events might mimic signal events is for one of the quark to hadronize into a charmed meson and for this charmed meson to decay semileptonically. That lepton-neutrino pair can then combine with a random pion in the event. We have two ways of dealing with this background: we use a continuum suppression cut to reject as many of those events as possible. For the remaining events, we make use of the off-resonance data, assuming that the continuum at a lower energy than the $\Upsilon(4S)$ looks the same as the continuum under the resonance after accounting for the beam energy difference. The continuum suppression cut is designed to discriminate between the jetty topology of continuum events versus the uniform topology of B events. In general, one has to be careful since some cuts introduce a bias as a function of q^2, while others do not. In this analysis, we only have to worry about the q^2 dependence within one of our q^2 bin. We optimized a geometrical cut which is a function of R2 and $cos\theta_{thrust}$. R2 is the second moment of the Fox-Wolfram ratio and approaches one as an event approaches a perfectly uniform distribution of partiles in 4π. $cos\theta_{thrust}$ is the angle between the thrust of the signal, the lepton and hadron candidates, and the thrust of the rest of the event. The cut is separately optimized for every q^2 bin, since the low region of q^2 contains the most continuum background (the high recoil dynamic produces jetty events). It is also optimized independently for every mode, and for $\Delta Q = 0$ and $|\Delta Q = 1|$.

We show the $cos\theta_{thrust}$ in figure 4.17. The maximum lepton cut of 2.84 Gev/c previously referred to also rejects continuum events containing leptons with momenta greater than the $b \to u\ell\nu$ endpoint.

The "V" cut

The most important background, after continuum suppression, comes from the $b \to c\ell\nu$ transition. Normally those events should not fall into the signal region: the charmed mesons are heavier than the charmless mesons, so their leptons are softer, but more importantly, the various hadrons present in those events should not make an acceptable beam constrained mass. It's only in events for which something wrong happened that such an event fall into our signal region. Unfortunately this happens rather often. For example, 90% of btoc events in the signal region have extra particles in them, half are K_L and the other half are neutrinos from $b \to c \to s\ell\nu$ transition. When there is an extra particle it increases the magnitude

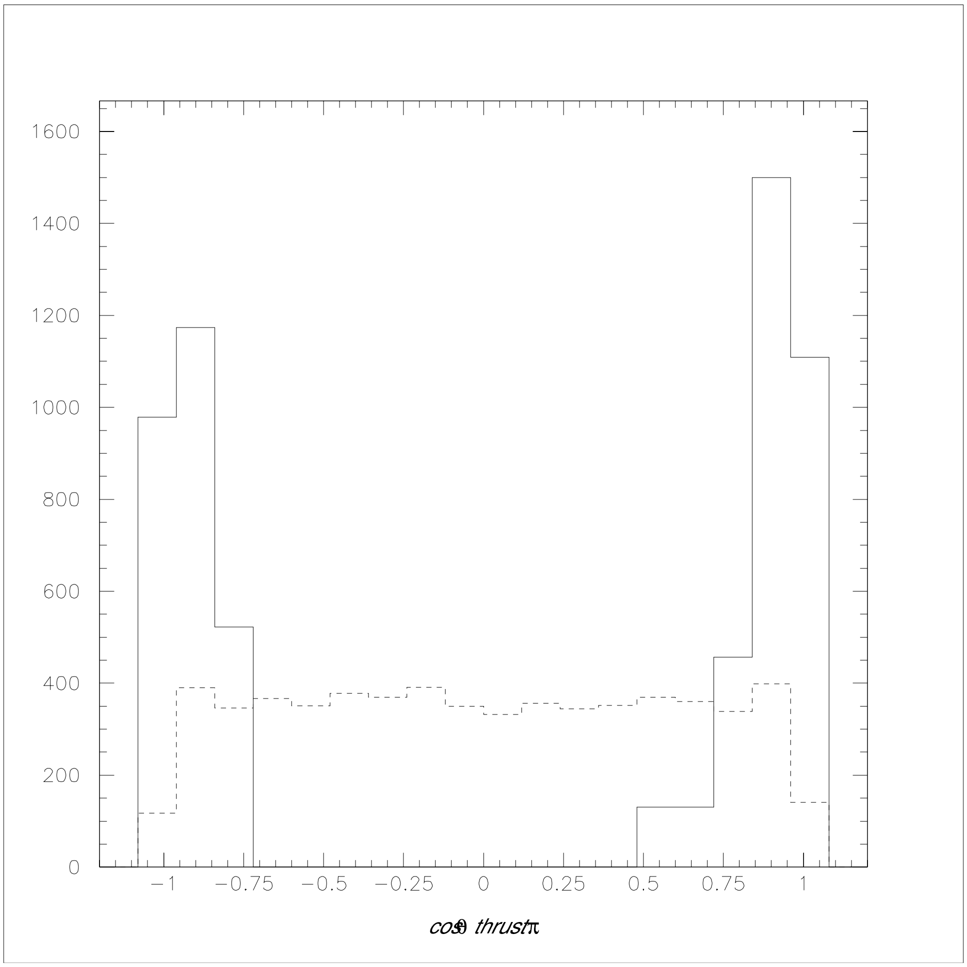

Figure 4.17: $cos\theta_{thrust}$ distribution. Solid is off resonance data, while dashed is π signal MC. The normalizations are arbitrary.

of the neutrino momentum; it is then easier to find a random soft track or set of tracks which, combined with the rather soft lepton, makes an acceptable beam constrained mass.

As mentioned previously the possible missed particles can be K_L, neutrons and neutrinos. K_L and neutrons can leave part of their energy in the calorimeter, if they do, that fraction of energy is taken into account and the mismeasurement is reduced, unless those showers are labeled as splitoffs. It is for the case that they do not interact at all that their entire energy is included in the neutrino energy and momentum [4].

All is not lost though, there are observables that can distinguish reasonably well if an extra particles was present. If we only had neutrinos as undetected particles, then, since their mass is essentially zero, the missing mass squared of the event should also be close to zero. If the missing mass squared is far away from zero there is a chance that there was an extra particle. Also, the missing energy of the event shows a different distribution for events with and without extra particles. Other observables, somewhat related to the previous ones, are the missing momentum of the event and the total energy of all good tracks in the event[5].

[4]The most recent $B \to \tau\nu$ analysis [48] made use of the fact that the number of kaons in an event is related with the probability of having a K_L in the event or not. For the present analysis, the figure of merit of such a cut indicated that it was not a useful cut to make.

[5]One thing that was tried was to build a Fisher discriminant using those variables, to distinguish between events with extra particles and events without. We can not train the Fisher by saying that signal is signal MC and background is btoc MC since then the Fisher will use its most powerful separating feature: the lepton momentum, so the resulting Fisher cut would be biased in q^2. Unfortunately,

We use a geometrical cut in the 2D plane of some combination of MM^2_{ev}, E_{miss} and p_{miss}. There are two observations that one can deduct from

$$\begin{aligned} MM^2_{ev} &\equiv E^2_{miss} - |\vec{p}_{miss}|^2, \\ \sigma MM^2_{ev} &\approx 2E_{miss}\sigma E_{miss}. \end{aligned}$$

The first one is that any geometrical contour in two variables has an equivalent contour for the other combination of two variables, the second observation is that the resolution in MM^2_{ev} is proportional to the true E_{miss}, the true neutrino energy, so actually p_{miss} is a better representation of that quantity. Indeed we see a clearer V shape when the y axis is p_{miss}, than when it is E_{miss}. Ultimately, we looked at what the figure of merit had to say on the various shapes and variables: a slightly asymmetric V cut on E_{miss} vs MM^2_{ev} for $\Delta Q = 0$ gave the best result. For the $|\Delta Q| = 1$ events, the nominal, tighter symmetric cut was kept. Table 4.8 gives the values used for the cut. Figure 4.18 and 4.19 show the V cut for π signal MC events and btoc MC events respectively, in the π signal region for $\Delta Q = 0$ and $|\Delta Q| = 1$.

Summary of all cuts and backgrounds

The following table summarizes the cuts, and to which background they are destined.

Tables 4.9 and 4.10 list the Signal Regions efficiencies for the different modes.

---

the Fisher trained on events with extra particles or without did not give enough separation between signal MC and btoc MC.

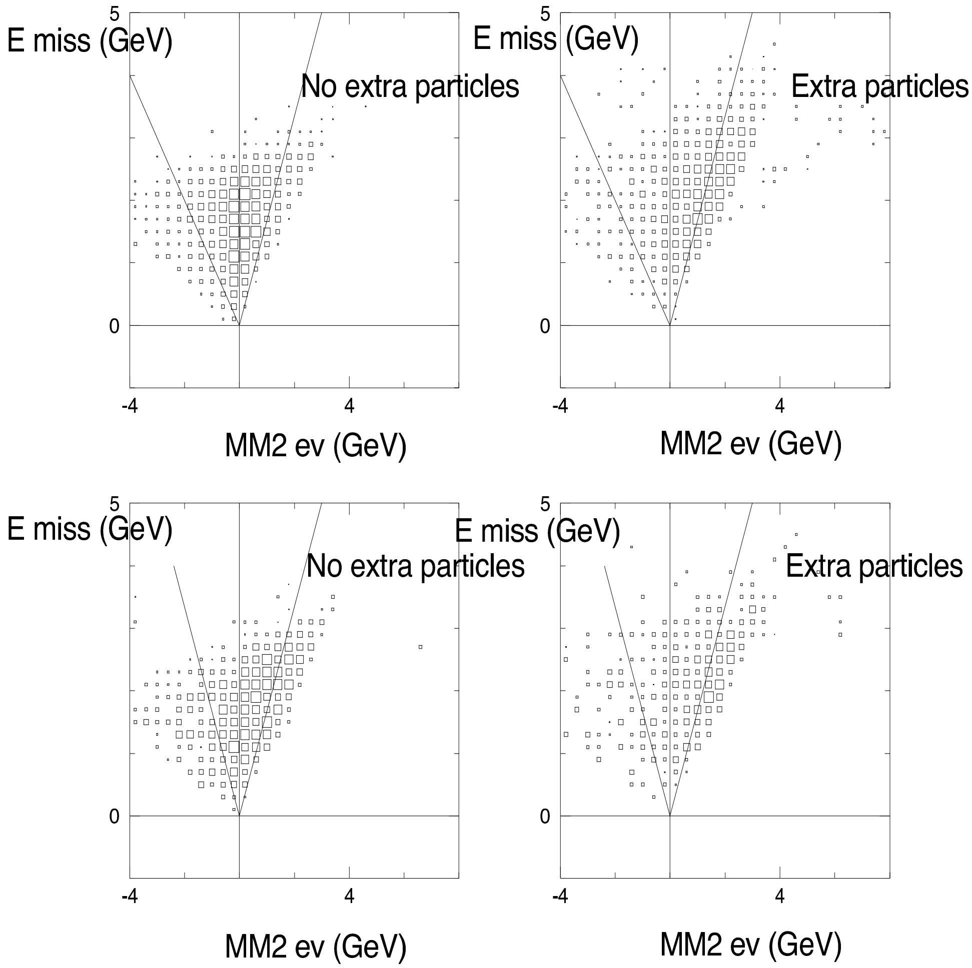

Figure 4.18: E_{miss} vs MM_{ev}^2 for π signal MC. Top is $\Delta Q = 0$, bottom is $|\Delta Q| = 1$, while left is with no extra particles and right has at least one extra particle.

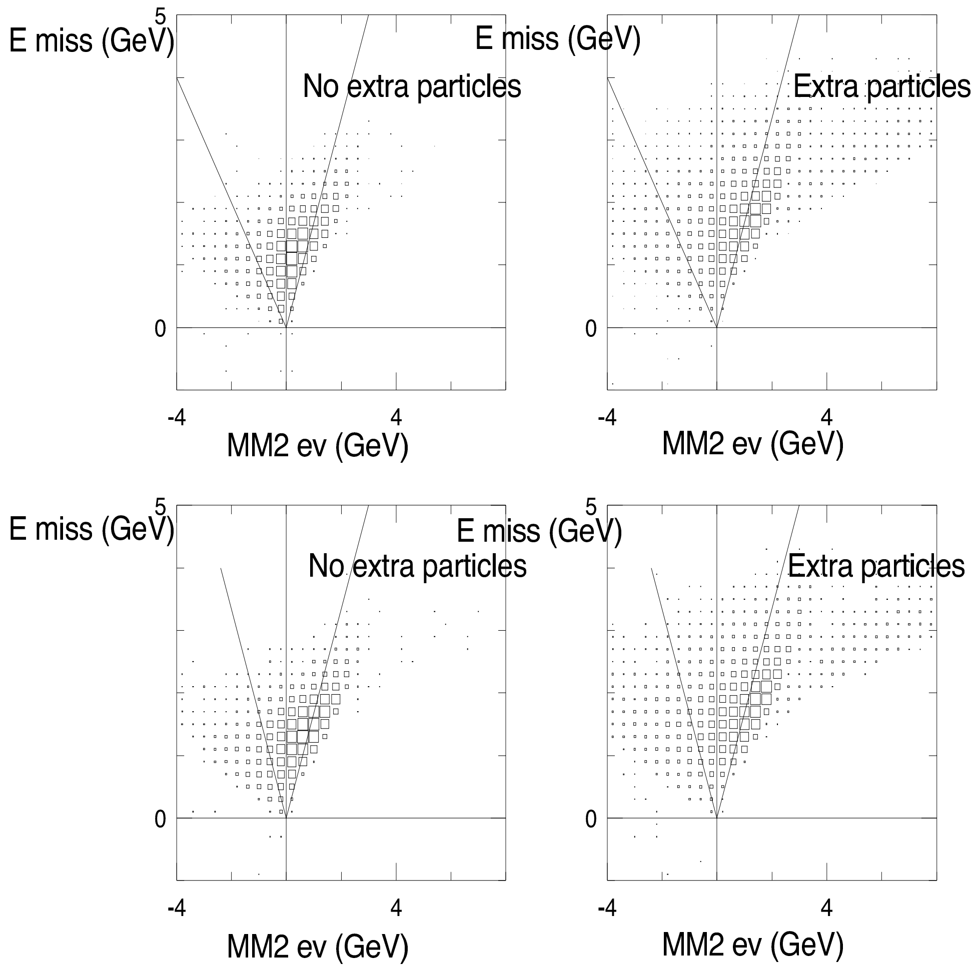

Figure 4.19: E_{miss} vs MM^2_{ev} for btoc MC. Top is $\Delta Q = 0$, bottom is $|\Delta Q| = 1$, while left is with no extra particles and right has at least one extra particle.

Table 4.8: Summary of all cuts

Cuts	type of background
Vcut: $\Delta Q = 0$: -600MeV-1GeV $\|\Delta Q\| = 1$: -600MeV-600MeV	K_L^0 and extra neutrinos $b \to c(u)\ell\nu$
$\Delta Q = 0 + \|\Delta Q\| = 1$	good ν reconstruction
$\vec{p}_\ell \geq 1.0GeV, 1.5GeV$	$b \to c\ell\nu$
Nlep=1	good ν reconstruction extra neutrinos
$cos\theta_{lep} > 0$ (vector modes)	$b \to c\ell\nu$
kincd,tng	good ν reconstruction
Ntracks: mode dep.	$b \to c\ell\nu$
$R2$ vs $\|c_{thrust}\|$ (per mode, q^2,ΔQ) tng(lep)	continuum suppression good ν reconstruction
$\|cos\theta_\nu\| < 0.96$	good ν reconstruction
hadronic event class	continuum suppression
psicut	$B \to \psi K_L$

Table 4.9: SR efficiencies , separately for $\Delta Q = 0$ and $|\Delta Q = 1|$ and separately for CLEO II and II.V, in percent

mode		signal	btoc	btou	fakes	cont
pip	dq0 II	3.16	0.00045	0.010	0.000066	0.000080
	dq1 II	0.80	0.00030	0.0063	0.000045	0.000020
	dq0 II.V	2.41	0.00032	0.010	0.00023	0.000022
	dq1 II.V	0.94	0.00021	0.0044	0.00021	0.000033
pi0	dq0 II	2.07	0.00013	0.0053	0.000028	0
	dq1 II	0.49	0.000076	0.0021	0.000013	0
	dq0 II.V	1.77	0.00012	0.0031	0.00013	0.000011
	dq1 II.V	0.56	0.000064	0.0045	0.000078	0.000011
rhp	dq0 II	1.17	0.00025	0.023	0.000063	0.000060
	dq1 II	0.27	0.00018	0.013	0.000017	0
	dq0 II.V	0.88	0.00024	0.019	0.00017	0.000022
	dq1 II.V	0.35	0.00018	0.012	0.000083	0
rho0	dq0 II	1.99	0.00058	0.045	0.000095	0.000040
	dq1 II	0.45	0.00021	0.014	0.000036	0.000020
	dq0 II.V	1.40	0.00035	0.039	0.00029	0.000022
	dq1 II.V	0.42	0.00022	0.014	0.00011	0

Table 4.10: SR efficiencies , separately for $\Delta Q = 0$ and $|\Delta Q = 1|$ and separately for CLEO II and II.V, in percent

mode		signal	btoc	btou	fakes	cont
ome	dq0 II	0.52	0.000046	0.010	0.0000071	0
	dq1 II	0.096	0.000027	0.0027	0.0000036	0
	dq0 II.V	0.40	0.000024	0.0081	0.000027	0
	dq1 II.V	0.11	0.000037	0.0027	0.000015	0
etp	dq0 II	1.08	0.000034	0.0031	0.0000020	0
	dq1 II	0.20	0.000047	0.0015	0.0000023	0
	dq0 II.V	0.76	0.000029	0.0026	0.0000079	0
	dq1 II.V	0.20	0.000015	0.00044	0	0
etg	dq0 II	1.44	0.000045	0.00076	0.000011	0
	dq1 II	0.25	0.000010	0.0014	0.000002	0
	dq0 II.V	1.19	0.000014	0.00083	0.000040	0
	dq1 II.V	0.27	0.000014	0.0017	0.0000083	0

Tables 4.11, 4.12 and 4.13 list the SR efficiency for each mode and each q^2 bin. They show both the cross feed among modes but also among q^2 bins. Those numbers are for both $\Delta Q = 0$ and $|\Delta Q| = 1$, and both CLEO II and II.V mixed.

Table 4.11: Efficiencies in q^2 bins for reconstructing $\pi^\pm$, in percent

True Mode		q_1^2	q_2^2	q_3^2	Sum
$\pi^\pm$	Sum	1.226	1.669	0.664	3.559
	q_1^2	2.421	0.054	0.001	
	q_2^2	0.082	4.513	0.056	
	q_3^2	0.000	0.129	4.400	
π^0	Sum	0.000	0.002	0.045	0.046
	q_1^2	0.000	0.000	0.000	
	q_2^2	0.000	0.004	0.009	
	q_3^2	0.000	0.003	0.283	
$\rho^\pm$	Sum	0.007	0.021	0.023	0.051
	q_1^2	0.011	0.000	0.000	
	q_2^2	0.004	0.033	0.011	
	q_3^2	0.000	0.060	0.133	
ρ^0	Sum	0.041	0.075	0.054	0.169
	q_1^2	0.067	0.001	0.000	
	q_2^2	0.020	0.137	0.013	
	q_3^2	0.000	0.170	0.335	
ω	sum				0.005
η_p	sum				0.001
η_g	sum				0.003

Table 4.12: Efficiencies in q^2 bins for reconstructing ρ^0, in percent

True Mode		q_1^2	q_2^2	q_3^2	Sum
$\pi^{\pm}$	sum	0.013	0.083	0.185	0.281
	q_1^2	0.051	0.031	0.003	
	q_2^2	0.000	0.138	0.107	
	q_3^2	0.000	0.000	0.626	
π^0	sum	0.000	0.006	0.035	0.041
	q_1^2	0.000	0.000	0.000	
	q_2^2	0.000	0.006	0.008	
	q_3^2	0.000	0.012	0.154	
$\rho^{\pm}$	sum	0.013	0.154	0.162	0.329
	q_1^2	0.033	0.014	0.004	
	q_2^2	0.008	0.244	0.064	
	q_3^2	0.001	0.093	0.627	
ρ^0	sum	0.553	1.934	0.837	3.324
	q_1^2	2.029	0.182	0.006	
	q_2^2	0.049	3.418	0.171	
	q_3^2	0.000	0.228	3.678	
ω	sum				0.040
η_p	sum				0.002
η_g	sum				0.007

Table 4.13: Efficiencies for the remaining modes, summed over q^2, in percent

Rec. Mode	q_1^2	q_2^2	q_3^2	Sum
π^0	0.886	1.115	0.405	2.406
$\rho^\pm$	0.350	1.344	0.583	2.276
ω				0.927
η_p				1.071
η_g				1.539

CHAPTER 5

FITTING PROCEDURE AND RESULTS

In the previous chapter, we described the selection of candidate events. In this chapter we describe how, from the selected events in data, we can extract the subset of which are most likely to be signal events and not background events. To do this we use a Maximum Likelihood fit that takes the MC distributions for the backgrounds and the signal and adjusts them to fit the data distributions.

5.1 Fitting procedure

A binned maximum log Likelihood fit

In order to compute the π and ρ branching fractions, we need to extract the yield of each mode in the data. As mentioned in the previous chapter, we fit over the $\Delta E - M_B$ region called the Grand Signal Region, more precisely, we subdivide this plane into 11 subregions, as shown in figure 5.1 and described in table 5.1.

We let the fit determine the relative proportion of each source to obtain the best representation of the data distributed over those subregions. One could minimize a chi-square built from the comparison of the data and the sum of all the sources in each bin. In our case, since we are dividing up our data into q^2 bins, as well

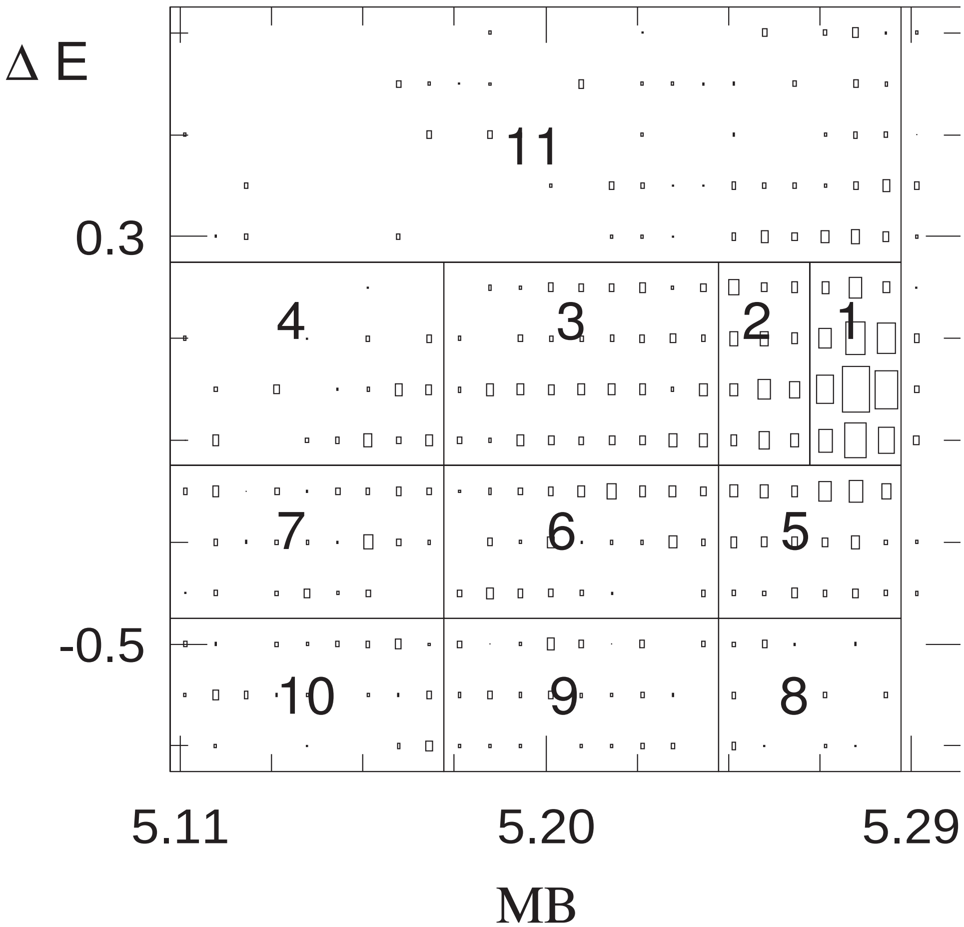

Figure 5.1: Grand Signal region with fit bins for reconstructed ρ mode. Top left is ρ signal MC, top right is π signal MC, bottom left is btoc MC, bottom right is btou MC

Table 5.1: Subregions of the GSR fit plane

Subregion	ΔE	M_B	Subregion	ΔE	M_B
1	-0.15-0.25	5.265-5.2875	7	-0.45- -0.15	5.1075-5.175
2	-0.15-0.25	5.2425-5.265	8	-0.75- -0.45	5.2425-5.2875
3	-0.15-0.25	5.175-5.2425	9	-0.75- -0.45	5.175-2.2425
4	-0.15-0.25	5.1075-5.175	10	-0.75- -0.45	5.1075-5.175
5	-0.45- -0.15	5.2425-5.2875	11	0.25-0.75	5.1075-5.2875
6	-0.45- -0.15	5.175-5.2425			

as in $\Delta Q = 0$ or $\Delta Q = 1$, we do not have enough data in certain bins to ensure that a Gaussian approximation of the data is adequate. We have to use a Poisson distribution for the data. We can still find the right proportion of each source by maximizing the likelihood. We are then performing a binned maximum likelihood fit. Another complication arises from the fact that we have to take into account the statistical nature of the different sources. We only have five times the amount of data for our generic btoc MC, and the off resonance data is clearly statistically limited. The prescription for a chi-square fit would be to add the errors coming from the sources, along with the data errors. In the likelihood approach, we realize that the events coming from the sources and that are being used in the fit, actually come from some unknown expected number of events that are also Poisson distributed, so we can take all this into account in the likelihood. This straight forward prescription is described in detail in [49], we describe here our particular situation.

The different sources that represent the data are the off resonance data representing the continuum background, the fake lepton data (on and off resonance), the btoc background MC, the btou other background MC, and the signal MC. The data yield used in the likelihood expression is not continuum and fake background subtracted. Those backgrounds, determined with data, are considered sources in the fit. The different bins that we need to consider in the fits are: 7 $\Delta E - M_B$ bins (Subregions 1, 2, 3, 5, 6, 8, 9), 7 modes, 2 ΔQ, 3 mass bands in the vector modes and 3 q^2 bins.

We eliminated the high ΔE region from our fit to reduce potential biasing of the $B \to \pi\ell\ nu$ branching fraction. This mode feeds into the high ΔE region of the ρ mode (see figure 4.4) and we found the $B \to \rho\ell\ nu$ branching fraction to be sensitive to mismodeling of the btoc background in this region.

We also decided that the three lowest M_B regions were constraining the btoc background too much given the high level of combinatorics in those regions, and so we do not include those regions in the fit either.

The btoc background is allowed to float independently according to the mode, the ΔQ and the reconstructed q^2 bin. The off resonance and the fakes are normalized based on luminosity (corrected for energy dependence and cross section). The btou other is represented by only one scale factor.

We now describe the formulation of the likelihood fit. The Poisson distribution is the limit of the binomial distribution but with a probability of success (p) which tends toward zero and a number of trials (N) which tends toward infinity, so that the mean $\mu = Np$ is still finite. The probability of observing r events is given by

$$P(r) = \frac{\mu^r e^{-\mu}}{r!}. \tag{5.1}$$

This probability distribution is not symmetric, so that the peak of the distribution does not correspond to the mean. If the mean is greater than about 20, then we recover the Gaussian distribution, for which the peak is the mean. If we have n independent observations $x_1, x_2, \ldots x_n$, from a theoretical distribution, $f(x, \theta)$, where θ is the parameter to be estimated, then the probability of observing the sequence of values is given by the Likelihood Function

$$\mathcal{L}(x, \theta) = f(x_1, \theta) f(x_2, \theta) \ldots f(x_n, \theta). \tag{5.2}$$

This probability is maximum for observed values, so to find θ, we take the derivative of $\mathcal{L}$ with respect to θ and set it to 0. For the case that the function f is the Poisson distribution, we get

$$\mathcal{L}(x, \mu_i) = \prod_{i=1}^{n} \frac{\mu_i^{x_i}}{x_i!} e^{-\mu_i}, \tag{5.3}$$

where μ is the mean (which we are trying to estimate). Here we can interpret μ as the estimated number of events in the i^{th} bin. n is the total number of bins and x_i is the number of data events in the i^{th} bin. We get,

$$\sum_{i=1}^{n} x_i \quad = \quad N_D$$

$$
\begin{aligned}
N_j &= \sum_{i=1}^{n} a_{ji} \\
\mu_i &= N_D \sum_{j=1}^{m} \frac{P_j a_{ji}}{N_j},
\end{aligned}
$$

where N_D is the total number of data events, j refers to a source, N_j is the total number of events from that source, a_{ji} is the observed number of events from source j in bin i. P_j is the normalization of the source j and is what we are trying to determine from the fit. It is simpler to evaluate the natural logarithm of the expression

$$
ln\mathcal{L} = \sum_{i=1}^{n} x_i ln\mu_i - \sum_{i=1}^{n} \mu_i - \sum_{i=1}^{n} ln(x_i!). \tag{5.4}
$$

So far we have not taken into account the effect of the MC statistics. As we alluded previously, we recognize that the a_{ji} events come from some true A_{ji}. Those events are also Poisson distributed since we have $A_{ji} \ll N_j$. The Log Likelihood can now be written as

$$
ln\mathcal{L} = \sum_{i=1}^{n} x_i ln\mu_i - \mu_i - ln(x_i!) + \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ji} lnA_{ji} - A_{ji} - ln(a_{ji}!). \tag{5.5}
$$

We can also rewrite,

$$
\begin{aligned}
\mu_i &= \sum_{j=1}^{m} p_j A_{ji} \\
p_j &= \frac{N_D P_j}{N_j}.
\end{aligned}
$$

We use MINUIT to minimize $-2ln\mathcal{L}$ by variation of the different p_j for the different sources.

It would be a CPU intensive effort to let the fit find the various A_{ji}. Fortunately we can express the various A_{ji}'s in terms of the a_{ji} at the cost of having to solve some polynomial equations, which we obtain by differentiating with respect to the p_j's and to the A_{ji}'s. We find

$$A_{ji} = \frac{a_{ji}}{p_j t_i + 1}, \tag{5.6}$$

where the t's satisfy

$$\frac{x_i}{1 - t_i} = \sum_{j=1}^{m} \frac{p_j a_{ji}}{p_j t_i + 1}, \tag{5.7}$$

and are polynomials of up to order 27. We use Newton's method to solve for the t's. The procedure of the fit is then to solve for each t_i, using the above equations, and use those in the appropriate equations for the A's, so that the log likelihood expression can be computed for each individual bin.

There is a special case when one of the sources has zero observed events. We have to be able to allow the dominant source to fluctuate to a non-zero value. Dominant here means that not only does this source contribute somewhat significantly to that bin, but also, that source is definitely statistically limited. This is sometimes the case for the btoc background and for the off resonance. We must be careful about the off resonance of the fakes sample since in that case its contribution is negative to the total yield for that bin. Hence it is never picked to be fluctuated up.

Normalization conditions

There are nice normalization conditions that come out of the previous formalism. For example, we get

$$\sum_i A_{ji} = \sum_i a_{ji} \; \forall j. \tag{5.8}$$

What this normalization condition means is that the predicted amount of the source should sum up to the observed amount, only the distribution among the different bins differs. We get another similar condition

$$N_D = \sum_i \sum_j p_j a_{ji}. \tag{5.9}$$

Since N_D represents the total number of data events, we see that it is built from the total number of events from the different sources, scaled by their right amount, once we have summed over all bins and sources.

Those normalization conditions should hold perfectly at the minimum of the fit and we verify that they do in our nominal fit.

Continuum Smoothing

Before describing the nominal fit and the results, there is an additional feature in performing our maximum likelihood fit. For the case of the off resonance data, the statistics is so sparse that the Poisson distribution can not be expected to fluctuate up to give a proper representation of this background over all the individual bins. To address this feature we perform a continuum smoothing procedure. We take the $\Delta E - M_B$ distributions with and without the continuum suppression cut and look in MC for a possible shape change as a result of this cut. The MC is a cocktail of $q\bar{q}$ MC, τ pairs and data containing fakes, all properly normalized to the data luminosity. In the fit we use the off-resonance data distributions with no continuum suppression cuts as our observed number of off resonance events for each individual bins. We then scale this amount to represent the effect of the continuum suppression cut. Finally,we renormalize the scaled yields so that the sum over the whole $\Delta E - M_B$ plane adds up to the off resonance yield with the cut applied. So, in short, we are not creating new off resonance events, we are merely redistributing them in a more uniform way over the $\Delta E - M_B$ plane. Figure 5.2 shows the continuum smoothed M_B distribution on top of the distribution which does not have any continuum smoothing applied.

5.2 The Results

Nominal Fit

As mentioned above, the btoc background is fitted separately for each mode and independently for the two ΔQ conditions. It is possible to also fit the btoc background separately for each q^2 bin. We expect the statistical error to go up, as the fit is given more freedom, but we are in fact trading some amount of systematic error for this increase of statistical error. We did verify that the errors were consistent when we fit with this particular situation. We also see the fitted scale factors

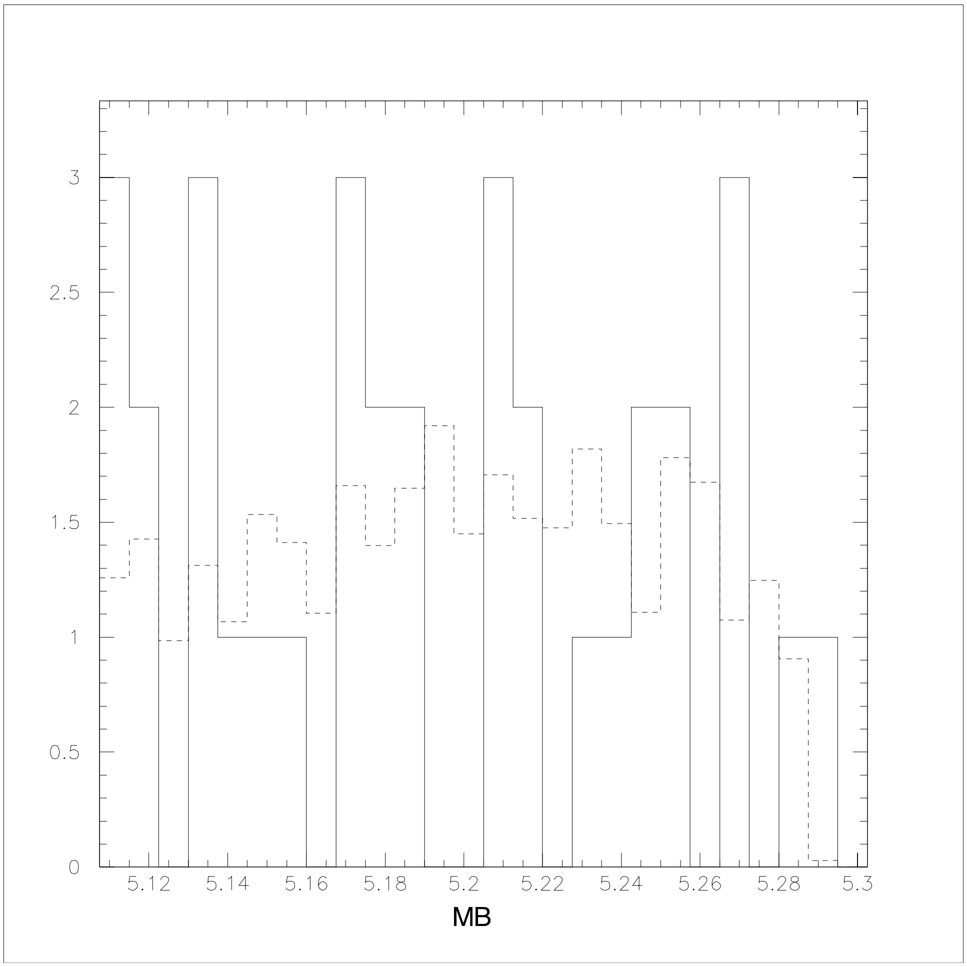

Figure 5.2: First q^2 bin of $\pi^{\pm}$. Solid is off resonance data, dashed is continuum smoothed off resonance data.

to be consistent across the three q^2 bins, hence the nominal fit only allows for one scale factor for all three q^2 bins.

There is a similar situation regarding the btou other background. This background is fitted using only one total scale factor. We could however let it float and let the fit decide on the best scale factor, but we do not since it is likely to be affected too much by btoc mismodeling. We currently have two different modeling of the btou other background: the first is the ISGW2 model (described in section 2.2.4), which only has resonant modes. The second model, used in the nominal fit, is a mix of ISGW2 for the resonant modes and of an inclusive HQET calculation of $d\Gamma/dq^2dM_X^2dE_\ell$ [50]. The appropriate amount of resonant vs non-resonant modes is made by assigning the difference between the total resonant exclusive rate and the total inclusive rate to be due to the non-resonant pieces.

Our dataset consists of the entire samples of CLEO II (3.14 fb^{-1}) and II.V (6.03 fb^{-1}). The off resonance data sample has 1.61 fb^{-1} and 2.94 fb^{-1} of luminosity, respectively. We get the branching fraction from the efficiency corrected yield, N, using

$$\mathcal{B} = \frac{N}{4N_{B\bar{B}}f_{00}}, \tag{5.10}$$

where the factor of four accounts for the fact that we have two B's per event and we use both electrons and muons as the charged lepton and f_{00} is the fraction of neutral B's vs charged B's, so around two. Tables 5.2, 5.3, 5.4, 5.5 and 5.6 summarize the results of the nominal fit.

We see that the η signal is only about 2.7σ significant.

Table 5.2: Yields for q_1^2

$\Delta Q0$	$\pi^{\pm}$	π^{0}	$\rho^{\pm}$	ρ^{0}
Signal	17.6	6.4	3.8	3.2
xfeed	1.1	0.4	0.6	0.6
cont.	4.0	1.7	1.2	3.9
btoc scale	1.01 $\pm$0.03	0.86 $\pm$0.05	0.91 $\pm$0.02	0.90 $\pm$0.01
btoc	4.9	0.5	5.8	10.3
btou	0.2	0.3	1.8	2.9
data	20.0	12.0	16.0	20.0
$\Delta Q1$				
Signal	5.5	1.8	1.3	0.7
xfeed	0.5	0.1	0.2	0.2
cont.	2.0	0.7	1.3	0.4
btoc scale	1.01 $\pm$ 0.04	1.09$\pm$0.07	0.90 $\pm$ 0.02	0.90 $\pm$ 0.02
btoc	3.0	0.0	2.9	3.5
btou	0.1	0.1	1.1	0.6
data	8.0	7.0	5.0	10.0

Table 5.3: Yields for q_2^2

$\Delta Q0$	$\pi^{\pm}$	π^0	$\rho^{\pm}$	ρ^0
Signal	48.6	16.8	18.1	15.3
xfeed	3.8	1.9	3.1	5.1
cont.	4.9	2.1	5.4	5.7
btoc scale	1.01 $\pm$0.03	0.86 $\pm$0.05	0.91 $\pm$0.02	0.90 $\pm$0.01
btoc	17.4	6.0	11.0	23.2
btou	1.7	0.6	4.9	10.8
data	70	31	50	70
$\Delta Q1$				
Signal	16.0	4.7	5.9	3.8
xfeed	1.9	0.6	1.4	1.5
cont.	2.5	1.1	0.4	3.0
btoc scale	1.01 $\pm$ 0.04	1.09$\pm$0.07	0.90 $\pm$ 0.02	0.90 $\pm$ 0.02
btoc	11.9	2.7	9.2	12.1
btou	0.7	0.4	2.8	3.7
data	28	15	17	20

Table 5.4: Yields for q_3^2

$\Delta Q0$	$\pi^{\pm}$	π^0	$\rho^{\pm}$	ρ^0
Signal	17.5	5.6	10.1	8.0
xfeed	10.7	4.3	4.1	6.5
cont.	4.7	0.9	3.1	2.4
btoc scale	1.01 $\pm$0.03	0.86 $\pm$0.05	0.91 $\pm$0.02	0.90 $\pm$0.01
btoc	13.1	4.1	5.4	6.3
btou	2.3	0.9	2.2	4.6
data	40	16	15	30
$\Delta Q1$				
Signal	6.9	1.8	3.4	2.7
xfeed	6.6	3.3	1.6	3.8
cont.	3.5	0.5	0.2	1.8
btoc scale	1.01 $\pm$ 0.04	1.09$\pm$0.07	0.90 $\pm$ 0.02	0.90 $\pm$ 0.02
btoc	9.2	4.8	3.1	3.2
btou	1.5	1.0	1.4	1.6
data	31	10	10	11

Table 5.5: Yields for all q^2

$\Delta Q0$	ω	η_p	η_g
Signal	4.8	3.3	8.1
xfeed	0.5	0.3	0.8
cont.	2.2	0.7	1.3
btoc scale	0.73 $\pm$0.03	0.85 $\pm$0.09	0.98 $\pm$0.12
btoc	2.3	2.4	2.7
btou	2.2	0.0	0.6
data	16	8	20
$\Delta Q1$			
Signal	1.1	0.8	1.6
xfeed	0.1	0.3	0.3
cont.	1.1	0.8	0.2
btoc scale	0.73 $\pm$ 0.05	1.33$\pm$0.18	1.13 $\pm$ 0.19
btoc	2.3	2.9	1.4
btou	1.2	0.4	0.7
data	8	1	7

Table 5.6: Efficiency corrected yields and branching fractions.

$N_{\pi^\pm}$	907 ± 213	1407 ± 210	537 ± 176	
$N_{\rho^\pm}$	576 ± 240	1680 ± 336	959 ± 161	
N_{η_g}				630 ± 230
Branching Fractions ($\times 10^{-4}$)				
	q_1^2	q_2^2	q_3^2	total
π	0.468 ± 0.110	0.725 ± 0.108	0.277 ± 0.091	**1.470 ± 0.179**
ρ	0.297 ± 0.124	0.866 ± 0.173	0.494 ± 0.083	**1.657 ± 0.228**
η				**0.837 ± 0.306**
-2 ln $\mathcal{L}$	394 for 406-21 d.o.f			

The following figures show some ΔE and M_B plots resulting from the fits. We can see that the ω mode plots in 5.12 show very little signal.

It is also interesting to look at how the fit agrees in the other regions of the $\Delta E - M_B$ plane. We show the π mode in subregion 2 (figures 5.14,5.15,5.16) which is just below the Signal Region in M_B. In those figures the M_B plot is for $-0.15 < |\Delta E| < 0.25$, while the ΔE plots is for $5.2425 < M_B < 5.265$. We also show the ρ mode in subregion 6 (figures 5.17,5.18,5.19), for which the M_B plot is for $-0.45 < |\Delta E| < -0.15$, while the ΔE plots is for $5.175 < M_B < 5.2425$.

Finally, we examine how the data yield is well represented in terms of the various sources, for each bin in the fit. In figure 5.20 we show a plot where the x axis is the bin number and the y axis is the different yields. We show these yields for two different $\Delta - EM_B$ regions: the Signal Region and subregion number 5. Also on that figure, we show a plot where the x axis are those same individual bins while the y axis is the log likelihood value. We see the correspondance between a small value of log likelihood (the fit tries to maximize the log likelihood) and a bin where the different sources do not match the data yield. In general, the agreement is good.

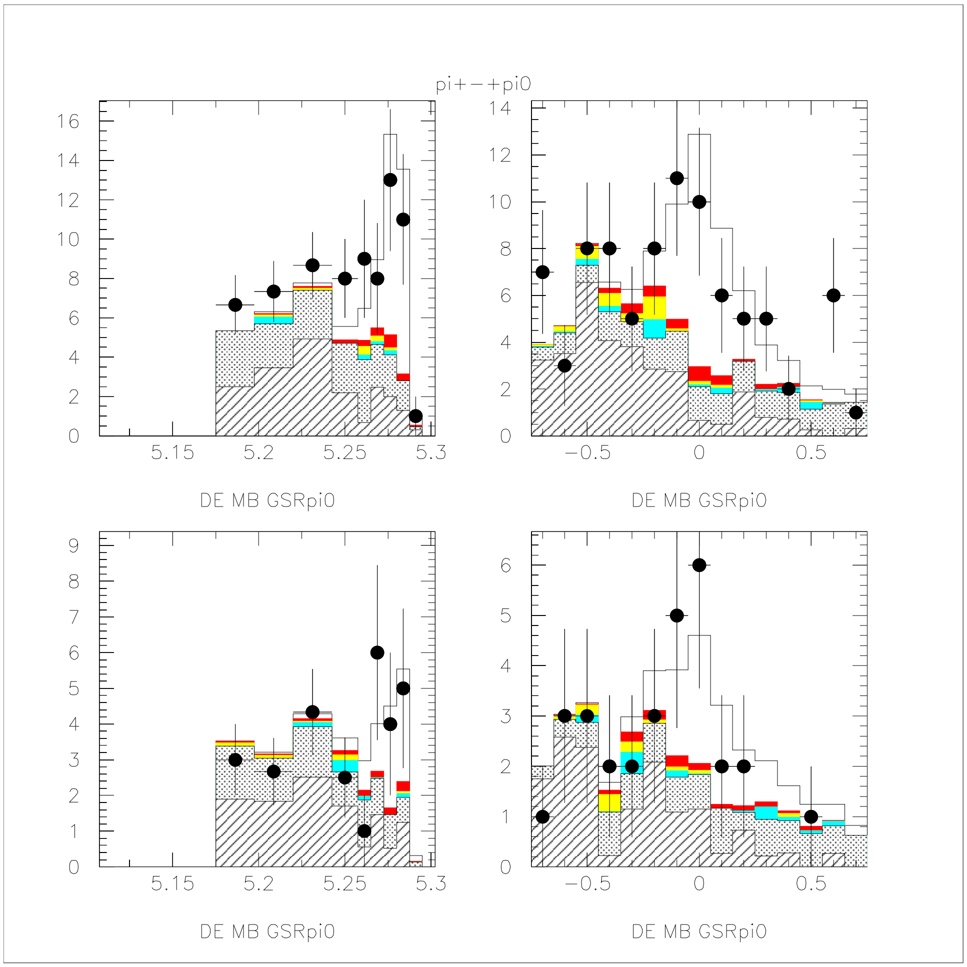

Figure 5.3: π mode plots for q_1^2. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

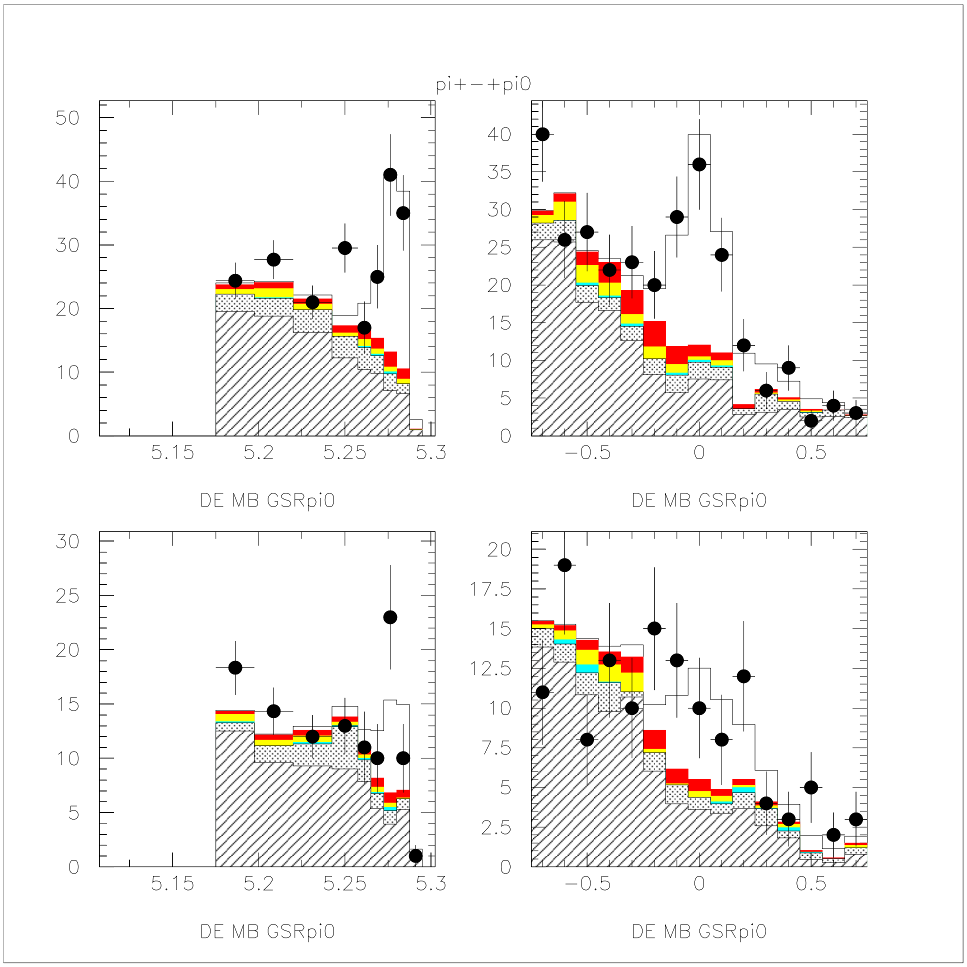

Figure 5.4: π mode plots for q_2^2. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

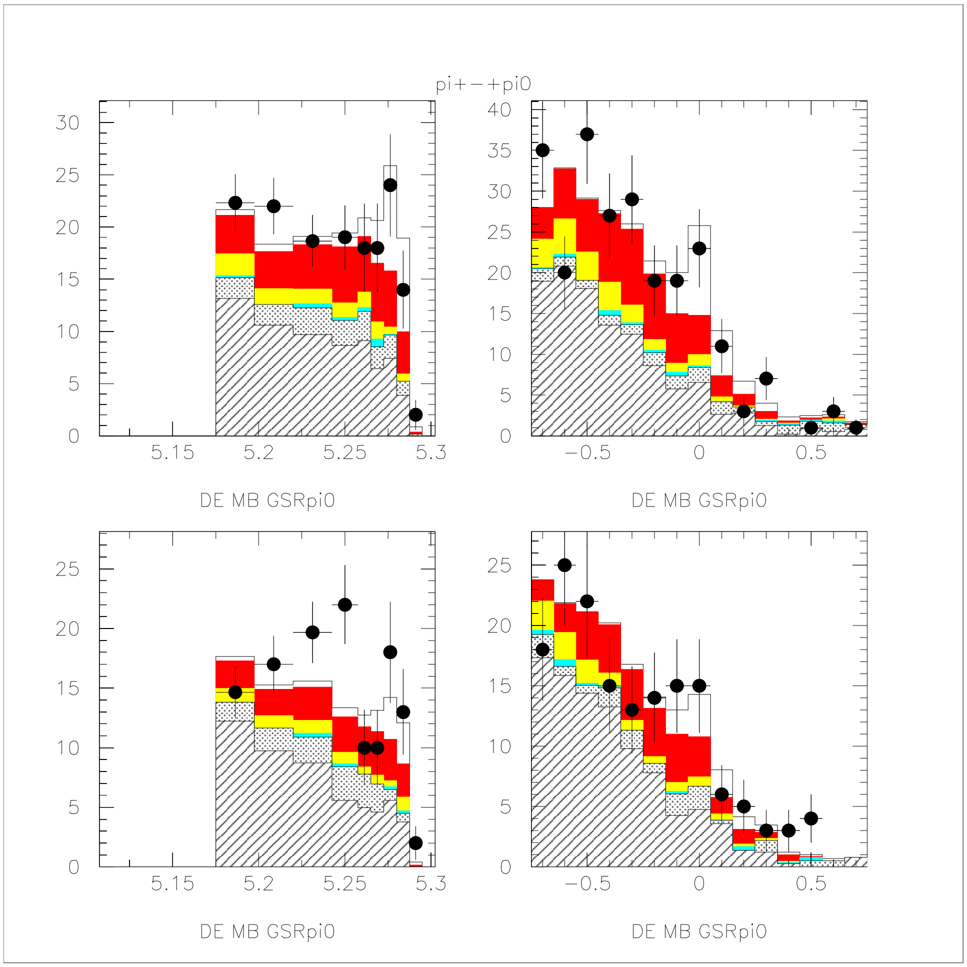

Figure 5.5: π mode plots for q_3^2. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

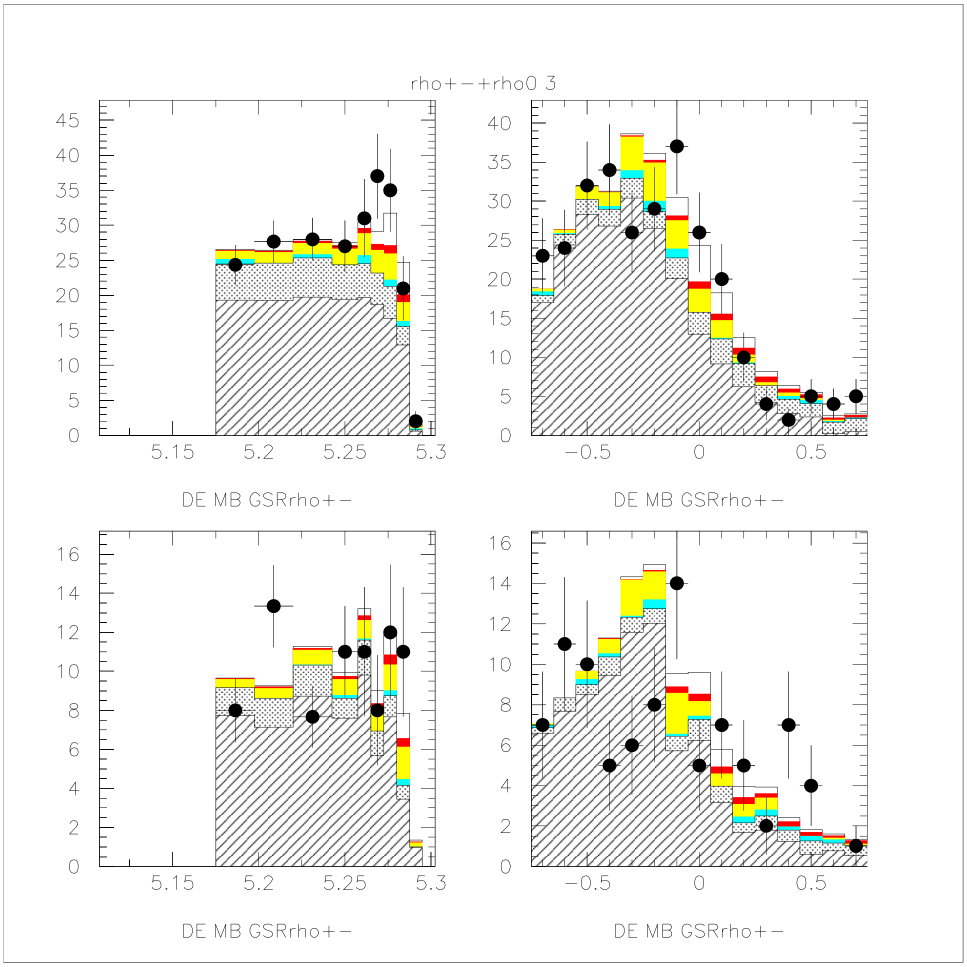

Figure 5.6: ρ mode plots for q_1^2, central 3 mass bands. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

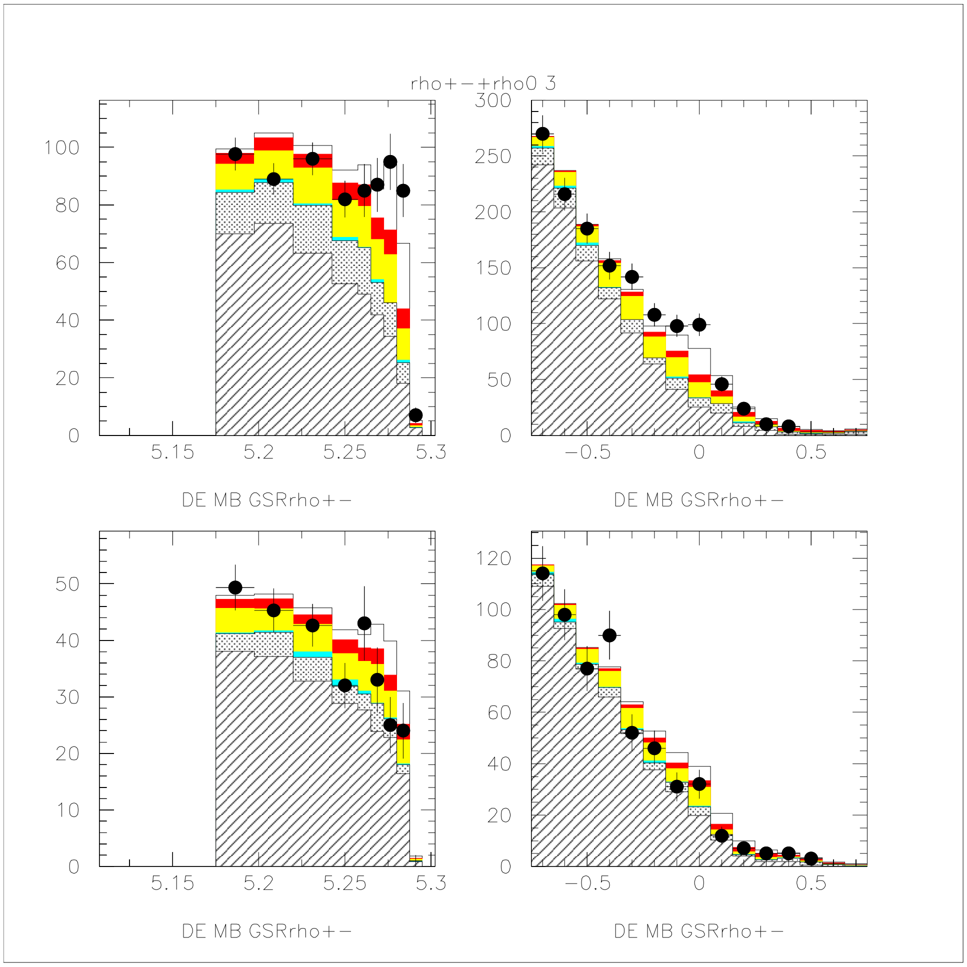

Figure 5.7: ρ mode plots for q_2^2, central 3 mass bands. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

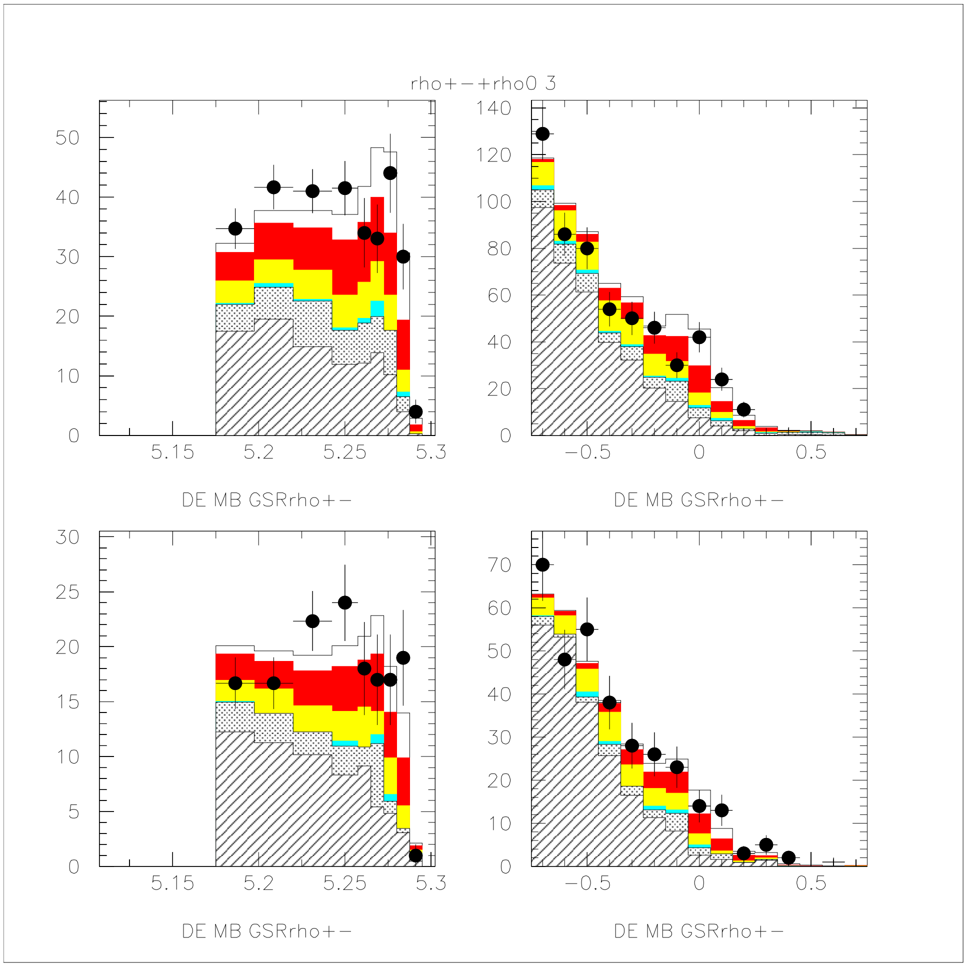

Figure 5.8: ρ mode plots for q^2_3, central 3 mass bands. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

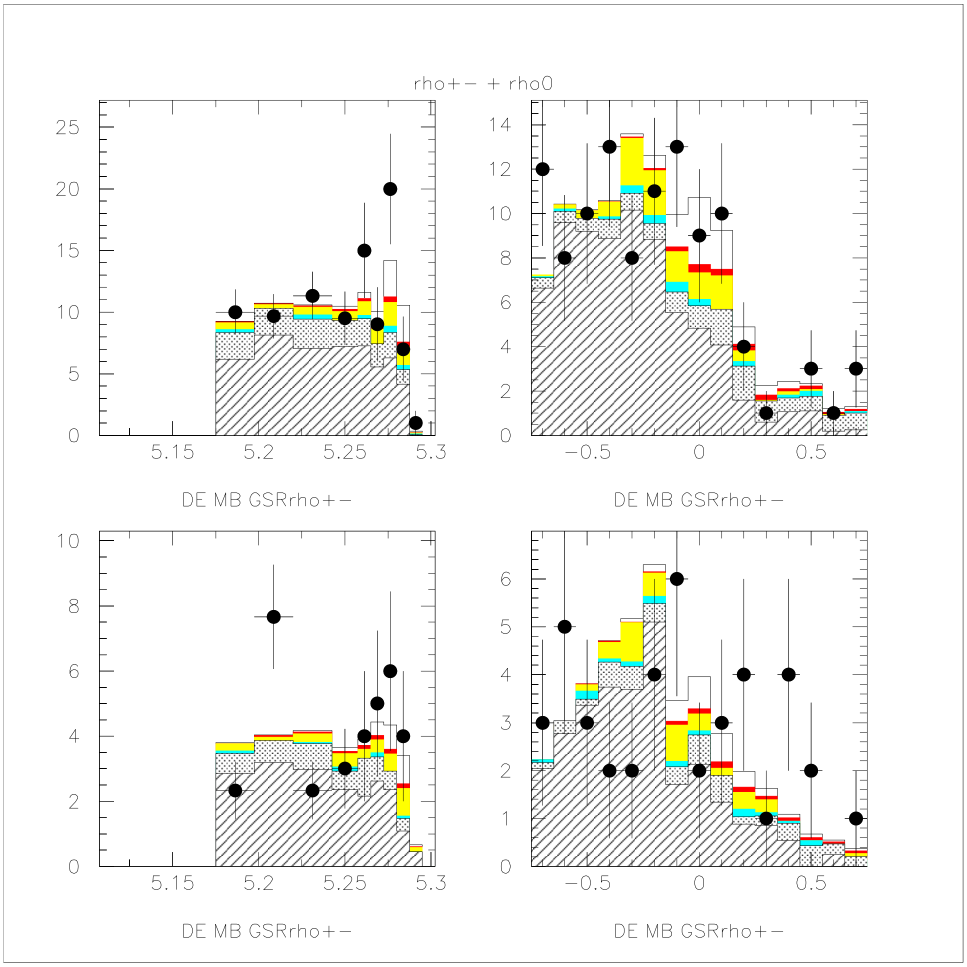

Figure 5.9: ρ mode plots for q_1^2, central mass band. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

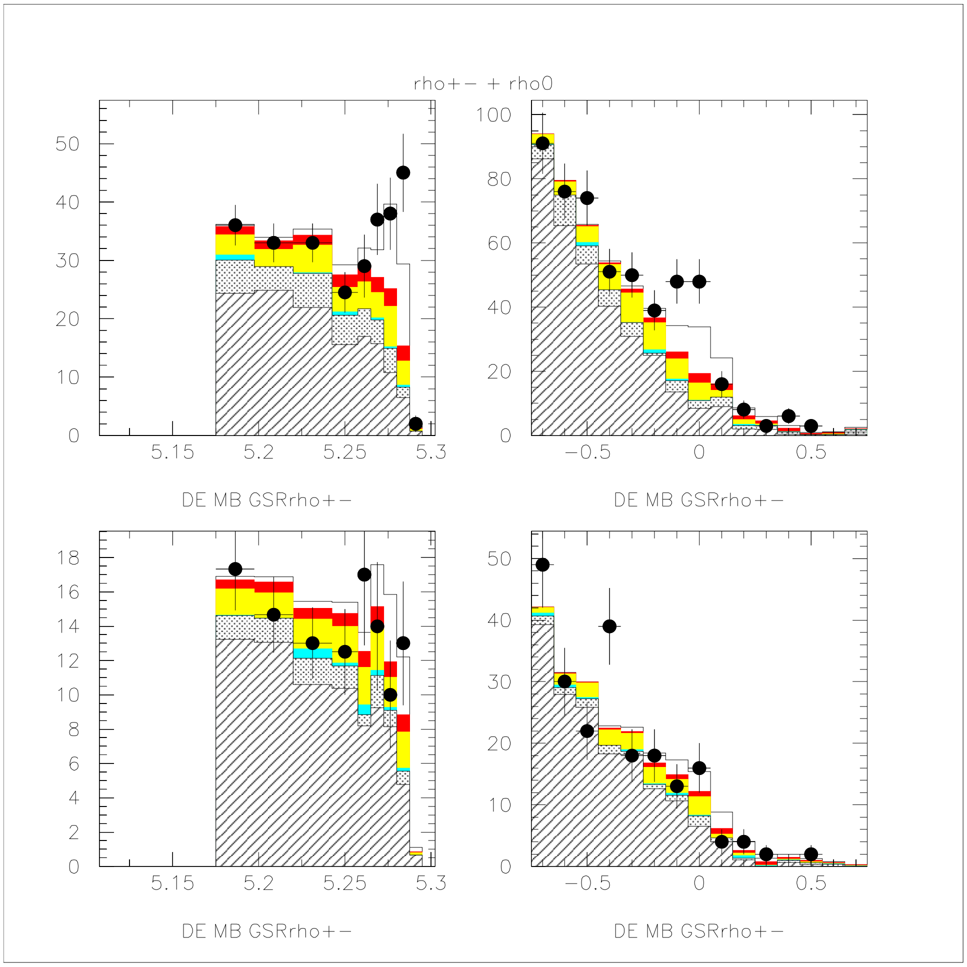

Figure 5.10: ρ mode plots for q_2^2, central mass band. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

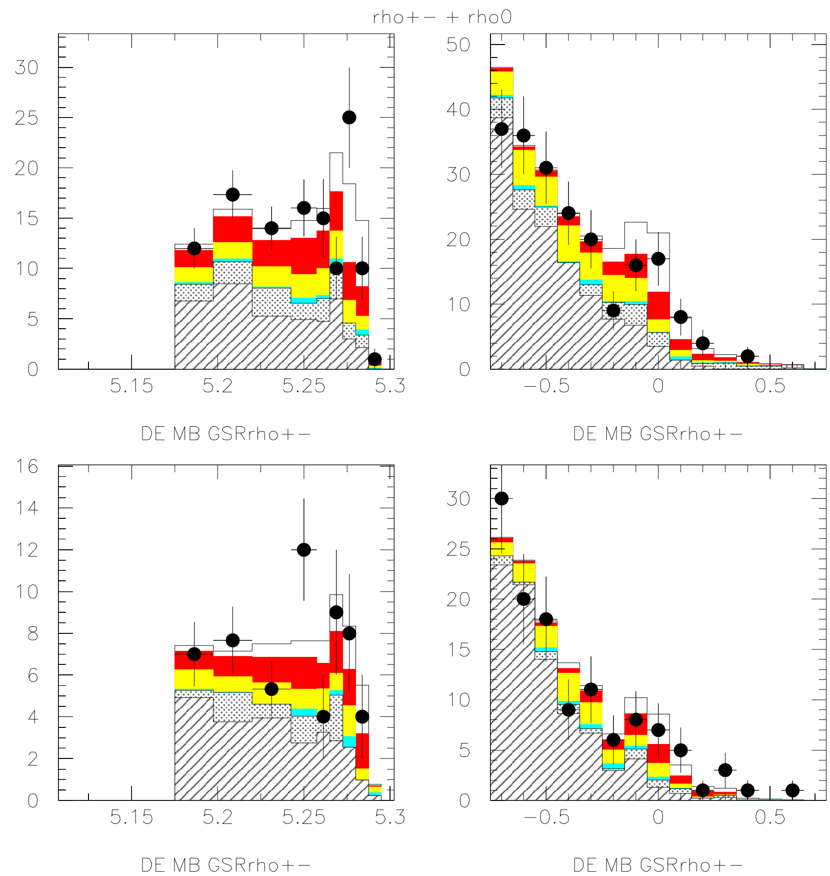

Figure 5.11: ρ mode plots for q_3^2, central mass band. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

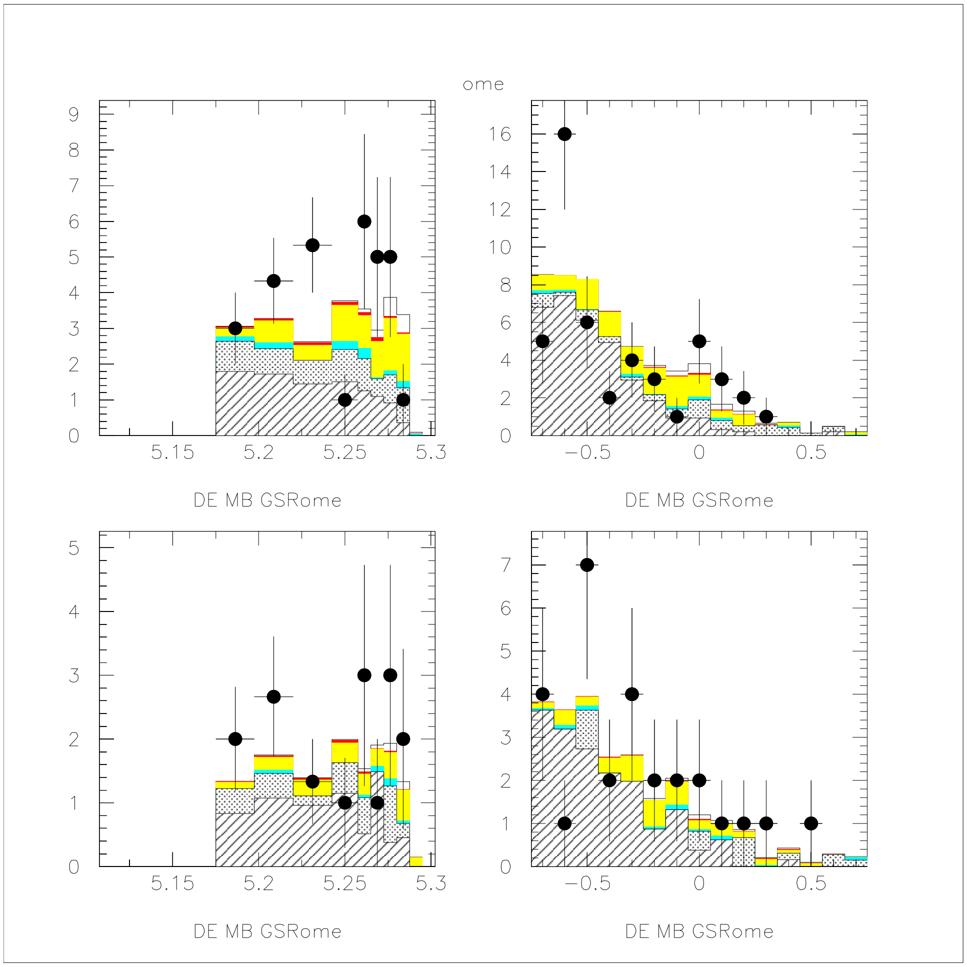

Figure 5.12: ω mode plots for all q^2. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

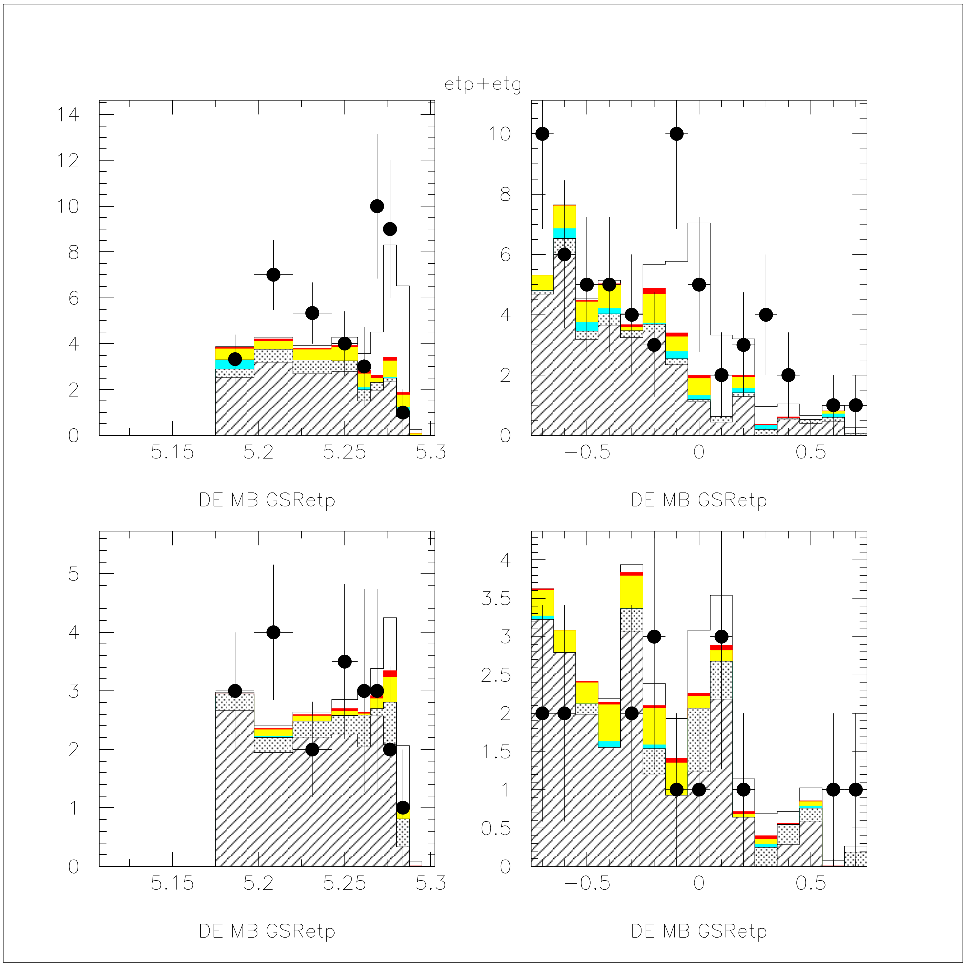

Figure 5.13: η mode plots for all q^2. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

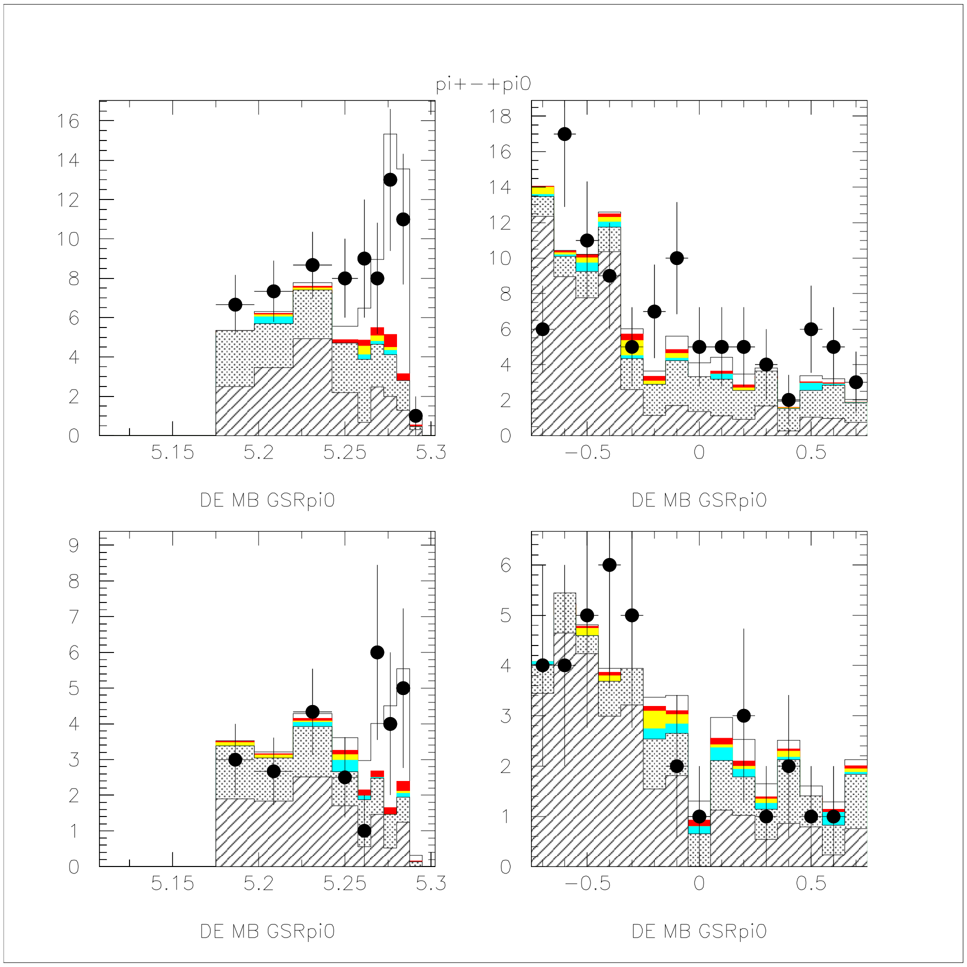

Figure 5.14: π mode plots for q_1^2 in subregion 2. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

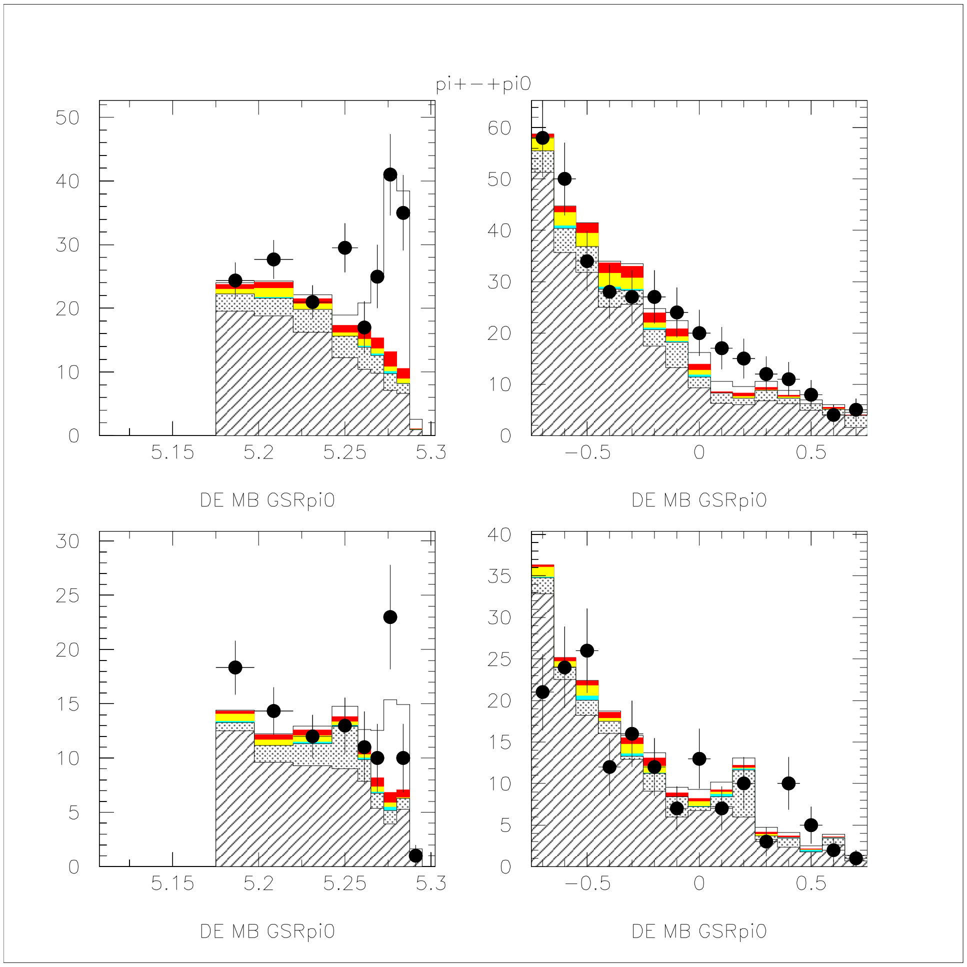

Figure 5.15: π mode plots for q_2^2 in subregion 2. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

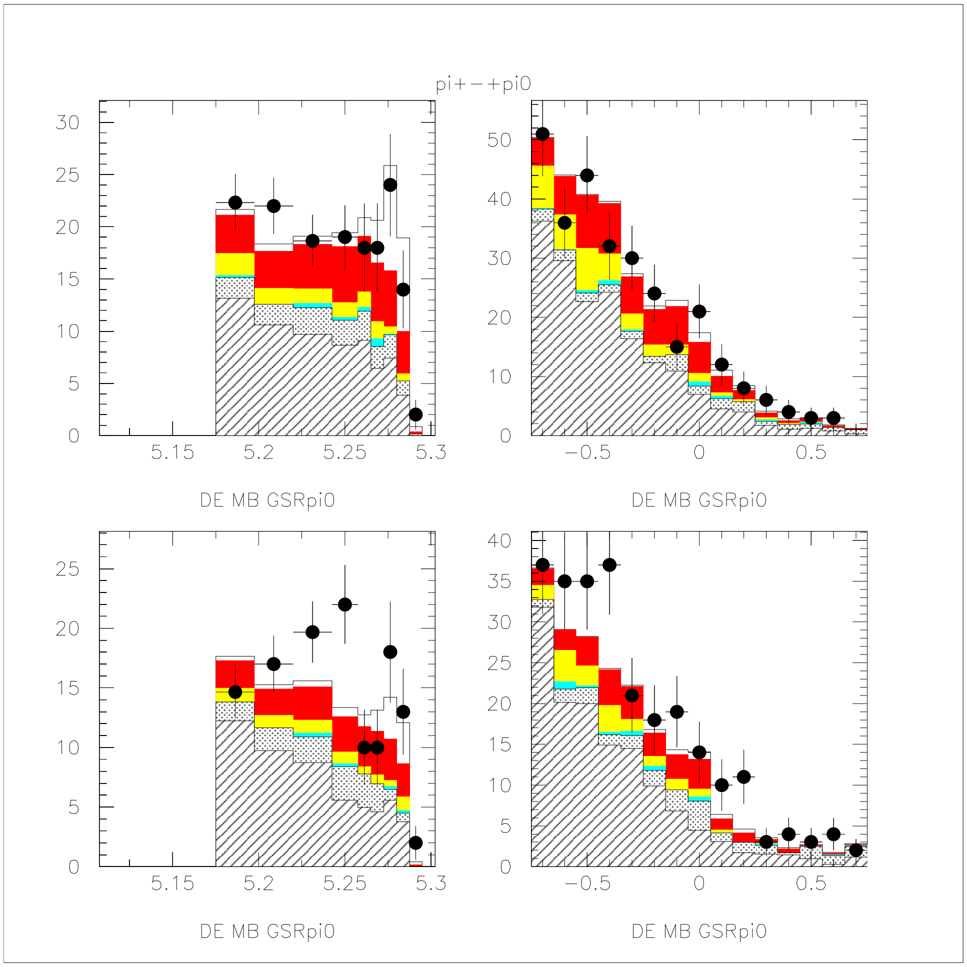

Figure 5.16: π mode plots for q_3^2 in subregion 2. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

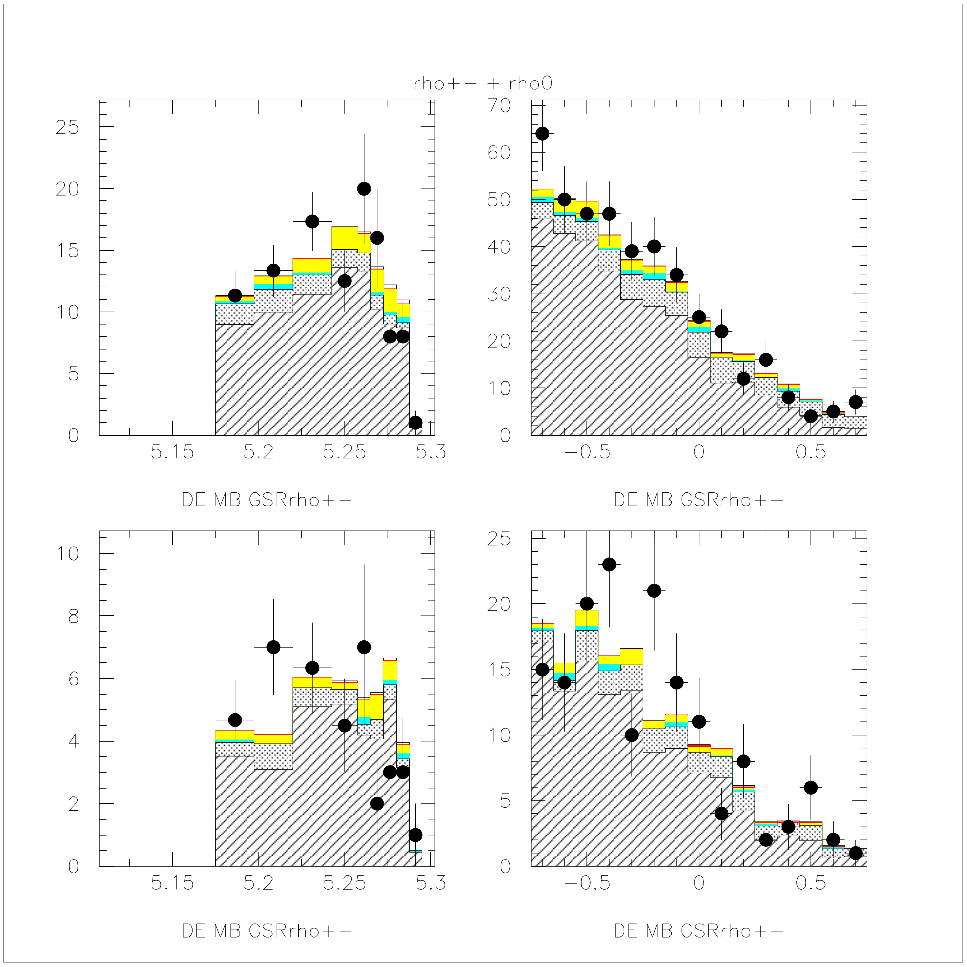

Figure 5.17: ρ mode plots for q_1^2, central mass band in subregion 6. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

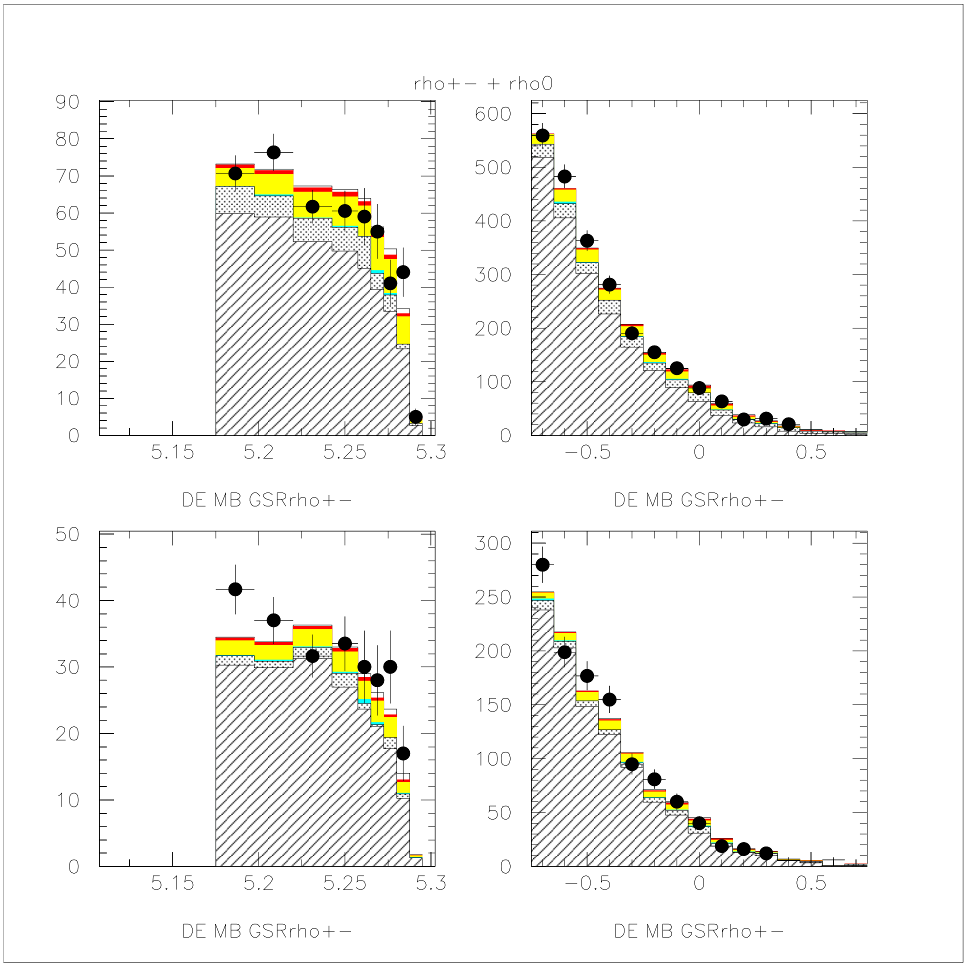

Figure 5.18: ρ mode plots for q_2^2, central mass band in subregion 6. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

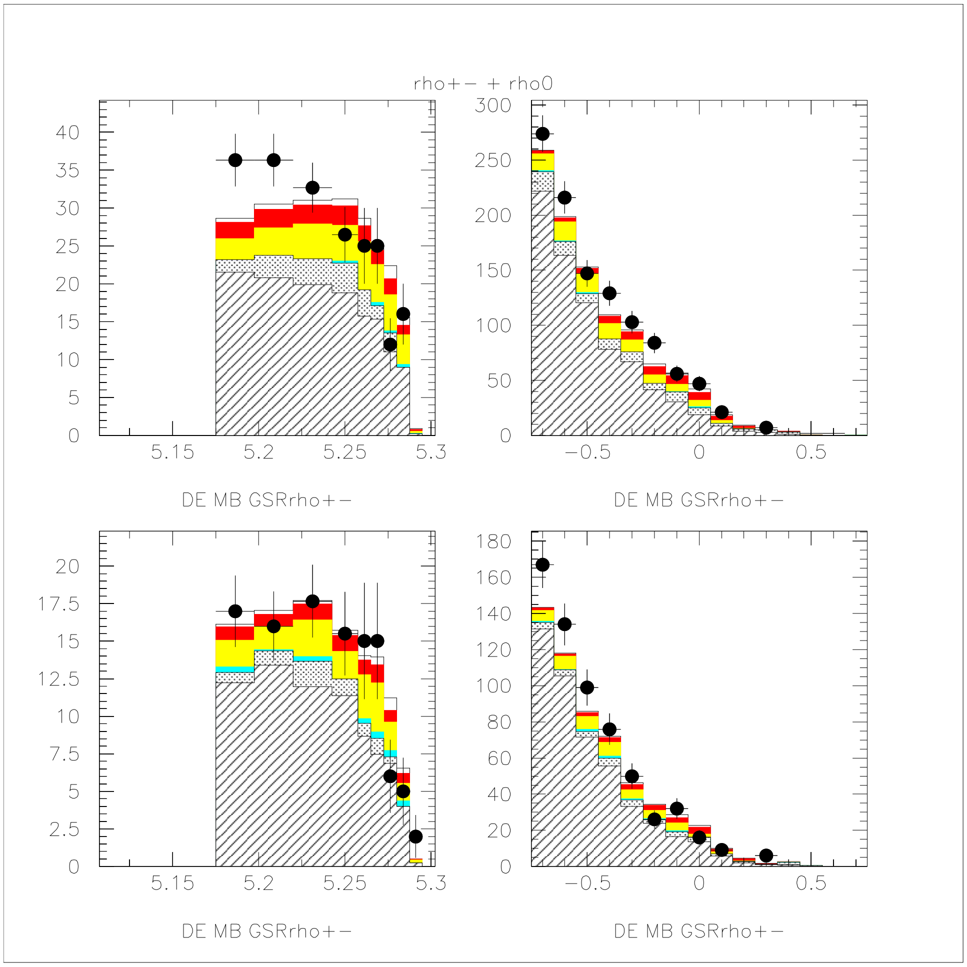

Figure 5.19: ρ mode plots for q_3^2, central mass band in subregion 6. Top is for $\Delta Q = 0$, bottom is for $\Delta Q = \pm 1$. Open histogram is signal, black is cross-feed, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatched histogram is btoc.

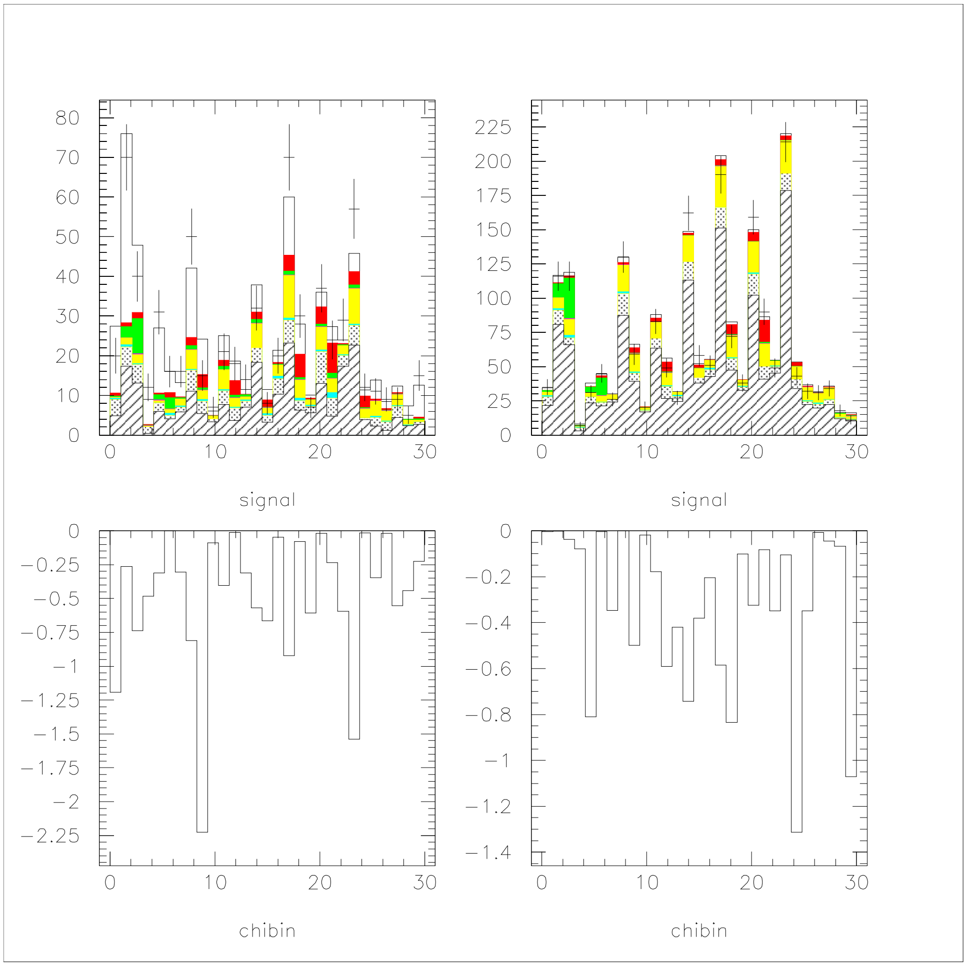

Figure 5.20: Top: Yield vs bin number: the order is $\pi_1^\pm, \pi_2^\pm, \pi_3^\pm, \pi_1^0, \pi_2^0, \pi_3^0, 5(\rho_1^\pm, \rho_2^\pm, \rho_3^\pm), 5(\rho_1^0, \rho_2^0, \rho_3^0), 4(\omega), \eta_p, \eta_g$. Open histogram is signal, black is cross-feed from same mode ($\pi^\pm \leftrightarrow \pi^0$), dark grey is cross-feed from other mode ($\pi \leftrightarrow \rho$), light grey is btou, darker grey is fakes, dotted histogram is continuum, hatched histogram is btoc. Bottom: $ln\mathcal{L}$ vs individual bins, for $\Delta Q = 0$. Left is Signal Region, right is region 5.

CHAPTER 6

SYSTEMATIC UNCERTAINTIES AND CHECKS

We have just seen in the previous two chapters how we select our events of interest, and how we fit them to get the different branching fractions. The final errors on the branching fractions or V_{ub} include a statistical part, a systematic part and a model dependence part. We now turn to determining the systematic uncertainties associated with the branching fractions. First, we look at some basic checks of the analysis.

6.1 Checks of the analysis

Kinematic Distributions

Several checks are performed to test the stability of the analysis. First we show some lepton momentum plots (figures 6.1 through 6.6), we see how the signal tend to populate the end point of the lepton spectrum, while the btoc background is at lower momentum values. It is interesting to note the correlation between a particular q^2 bin and the domain of lepton momentum values for the π and ρ modes.

Figures 6.7 through 6.12 show the hadronic mass distribution. The last bin of the plots contains the plot overflows. Figures 6.13, 6.14 and 6.15 show the $cos\theta_{lep}$ distribution, we see how the signal tend to populate the forward part of the spectrum. It is also interesting to note how the various backgrounds populate differently the various q^2 bins when projecting on the $cos\theta_{lep}$ axis. All the normalizations are taken from the fit results. In general we see good agreement between the data and MC, even though those variables were not part of the fit.

Stability of the fit

We can also test the stability of the fit by relaxing some of the cuts used in the analysis and comparing those fits with the nominal fit. We varied the lepton momentum cut, the "V" cut, and the $cos\theta_{lep}$ cut. We also performed the fit by using only the electrons and only the muons as the signal lepton candidate. We show in table 6.1 by how many sigmas the branching fractions differ [1]. We see no significant bias.

Also, we have run the fit on the CLEO II only part of the data set, using the same cuts as the previous analysis ([36]) and with the q^2 bins summed, so that we can compare our result with the previous analysis ones. For both analyses, the signal and the btou other models are ISGW2. The differences between the analyses are that the 1995 analysis used about 2/3 of the CLEO II data and used

[1] The sigma being defined as the difference of the central values over the quadrature sum of the uncertainties.

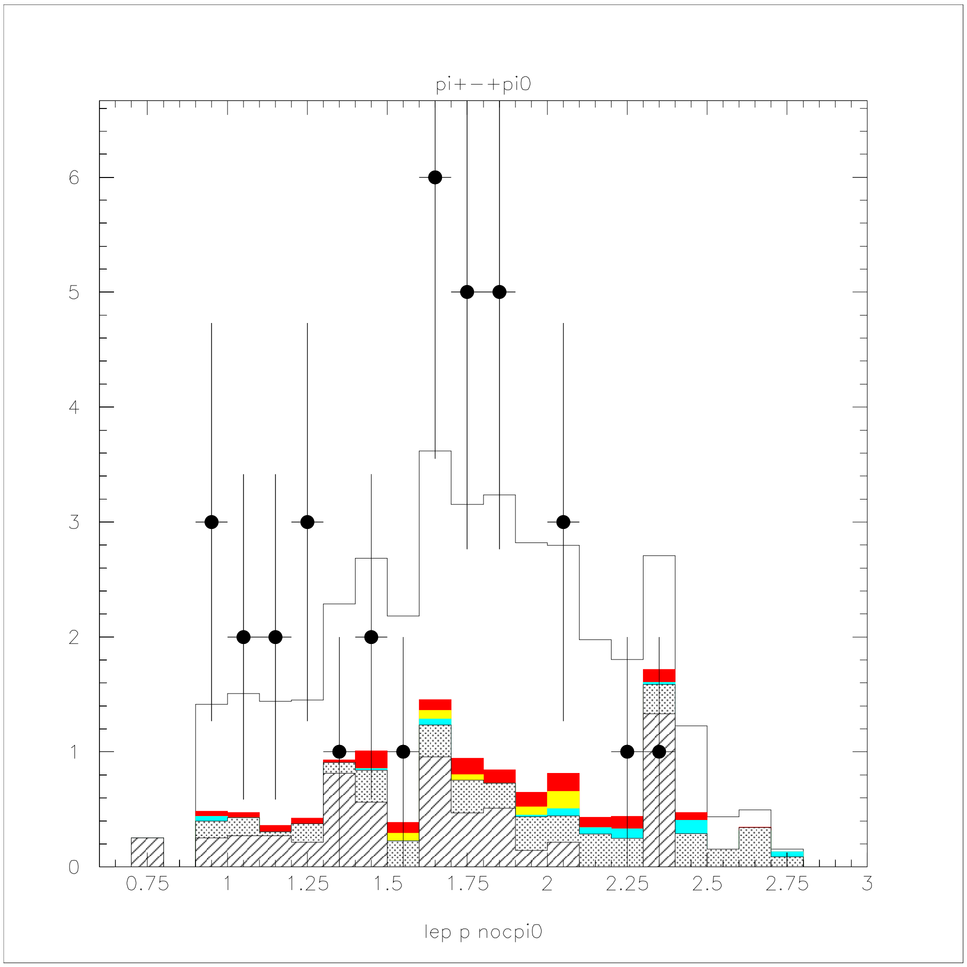

Figure 6.1: π mode lepton momentum for q_1^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

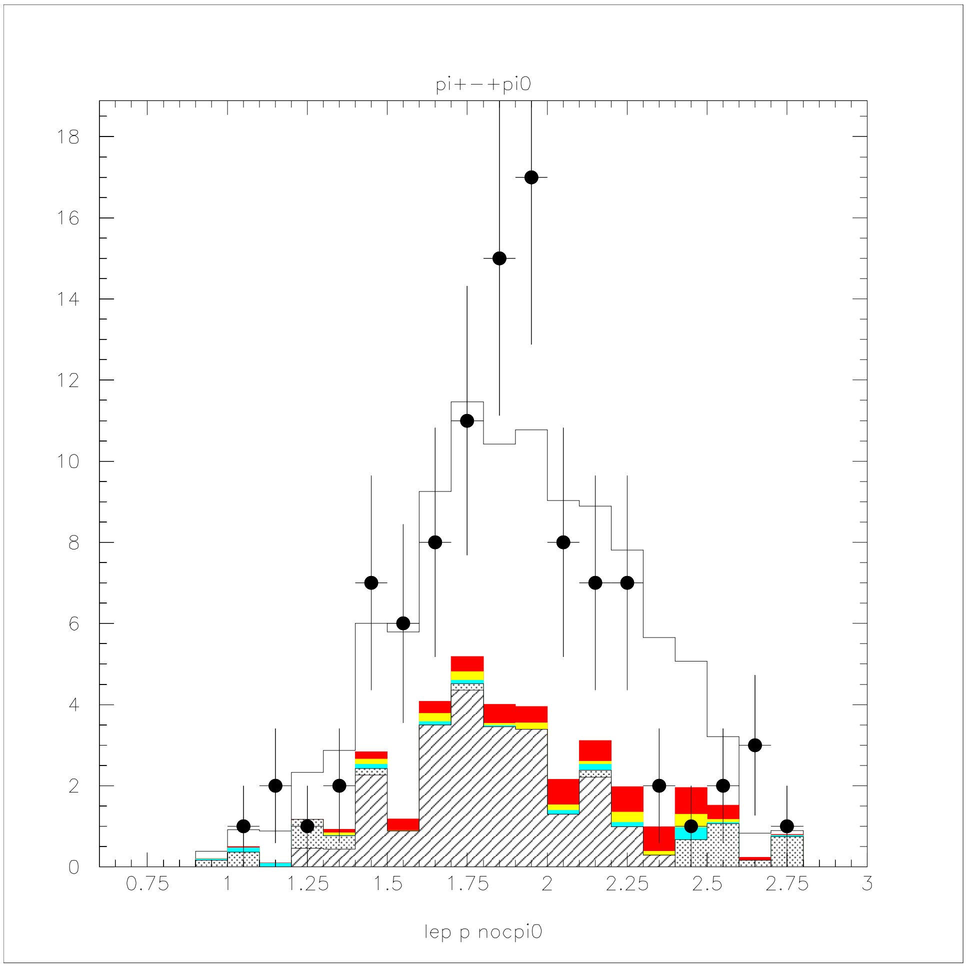

Figure 6.2: π mode lepton momentum for q_2^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

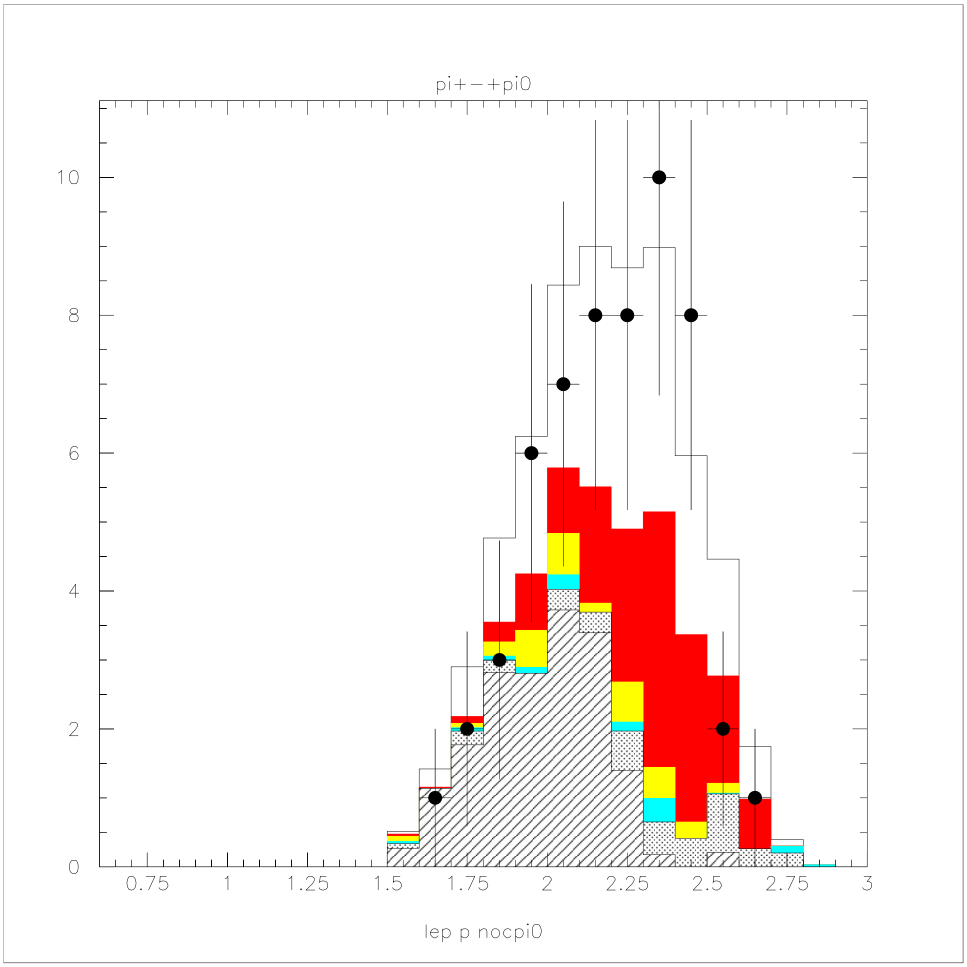

Figure 6.3: π mode lepton momentum for q_3^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

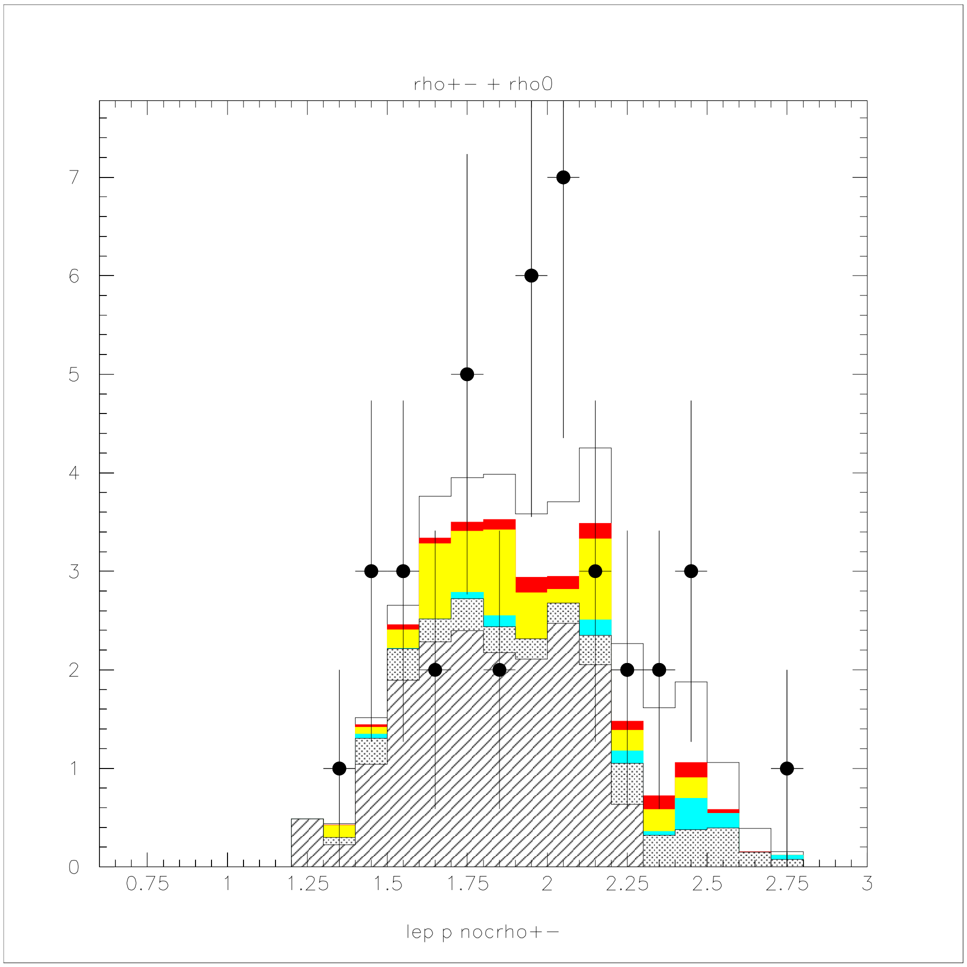

Figure 6.4: ρ mode lepton momentum for q_1^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

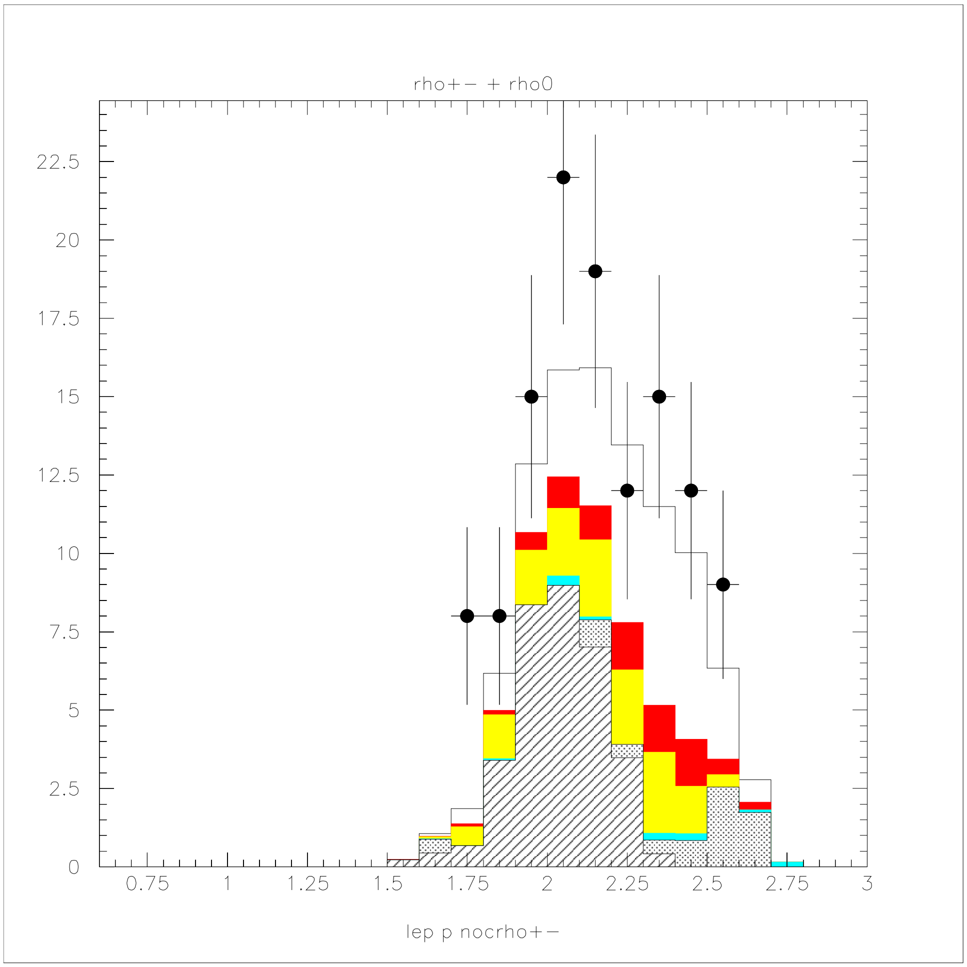

Figure 6.5: ρ mode lepton momentum for q_2^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

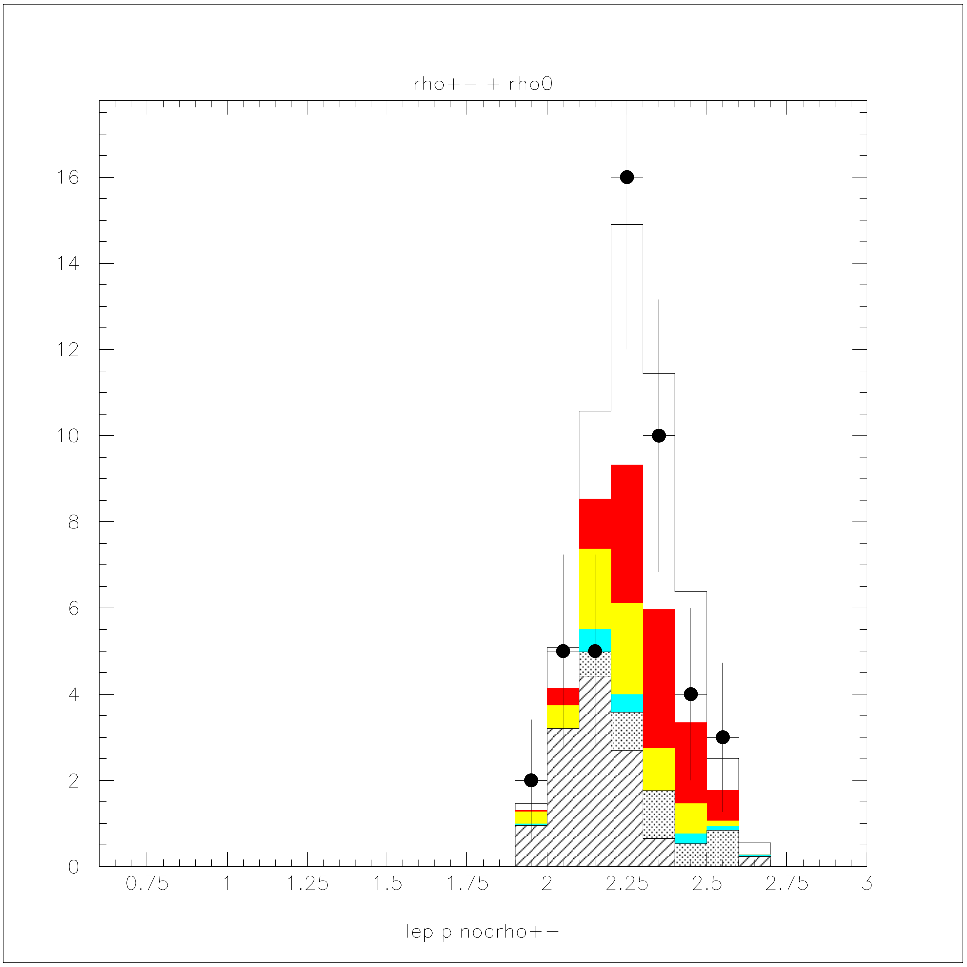

Figure 6.6: ρ mode lepton momentum for q_3^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

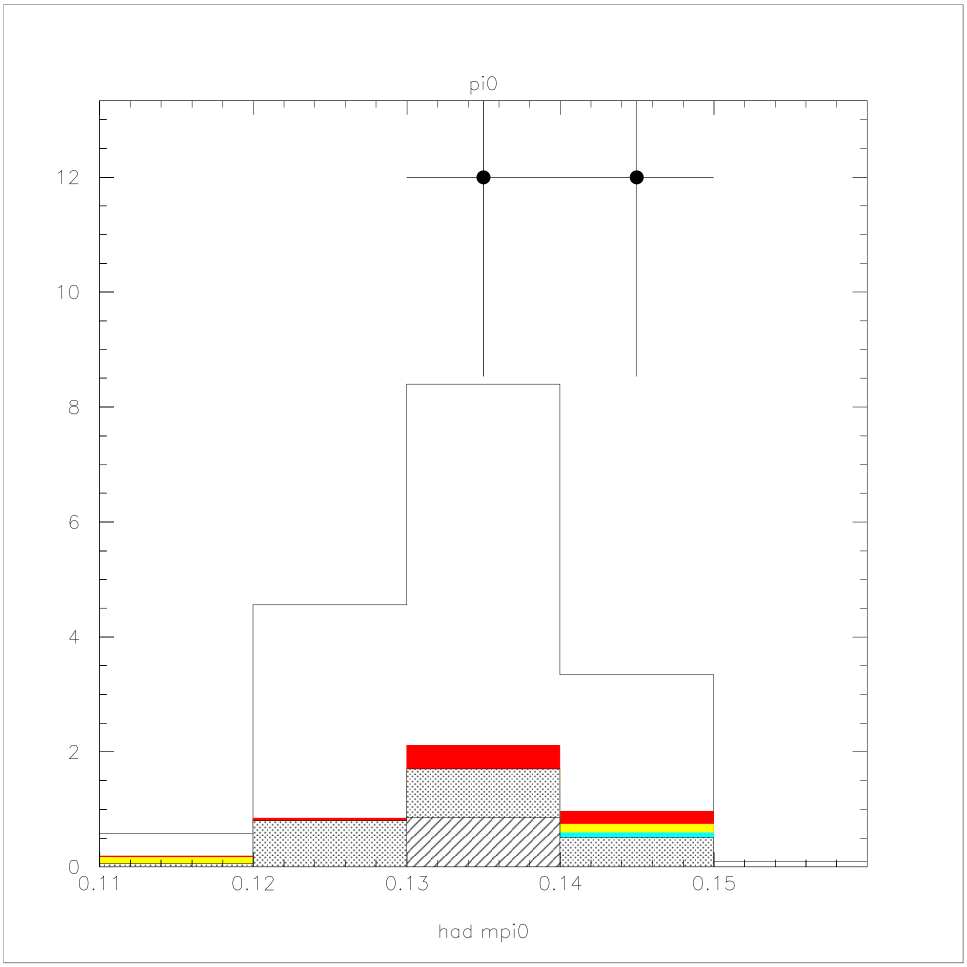

Figure 6.7: $m_{\gamma\gamma}$ for the π^0 mode for q_1^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

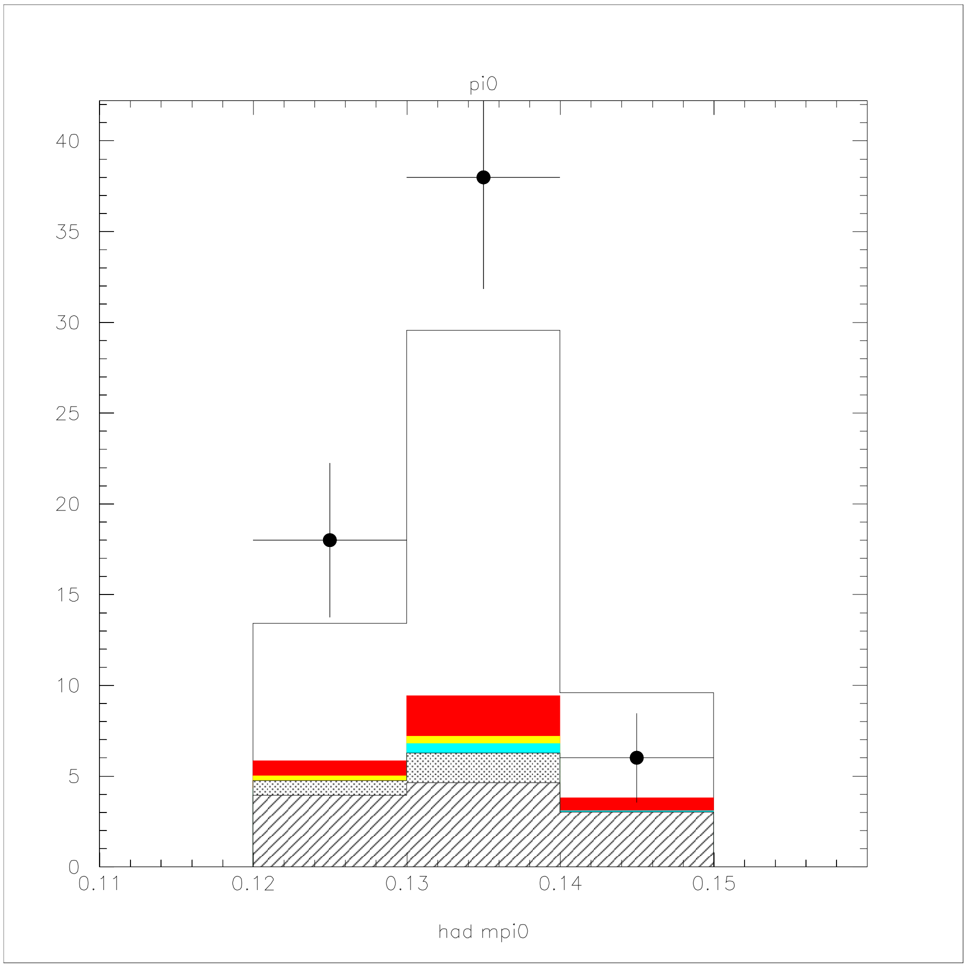

Figure 6.8: $m_{\gamma\gamma}$ for the π^0 mode for q_2^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

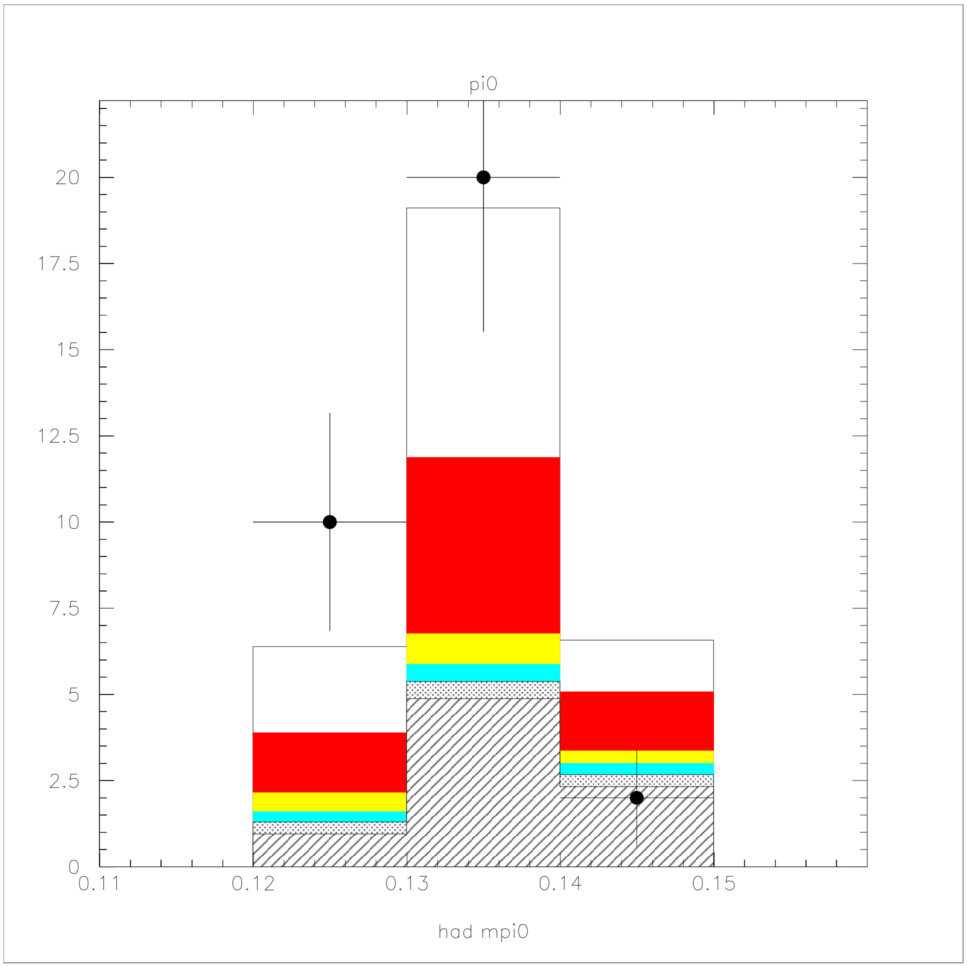

Figure 6.9: $m_{\gamma\gamma}$ for the π^0 mode for q_3^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

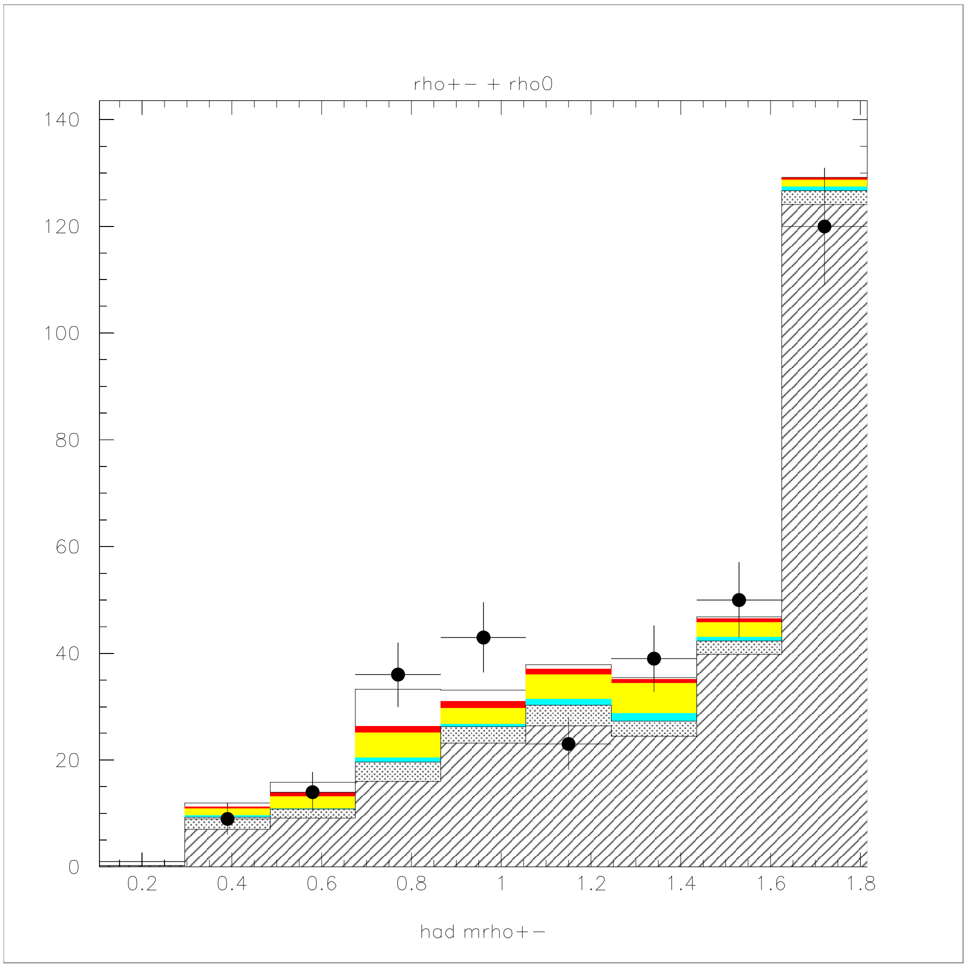

Figure 6.10: Combined $m_{\pi^{\pm}\pi^{0}}$ and $m_{\pi}^{+}\pi^{-}$ for the $\rho^{\pm}$ and ρ^{0} mode for q_1^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

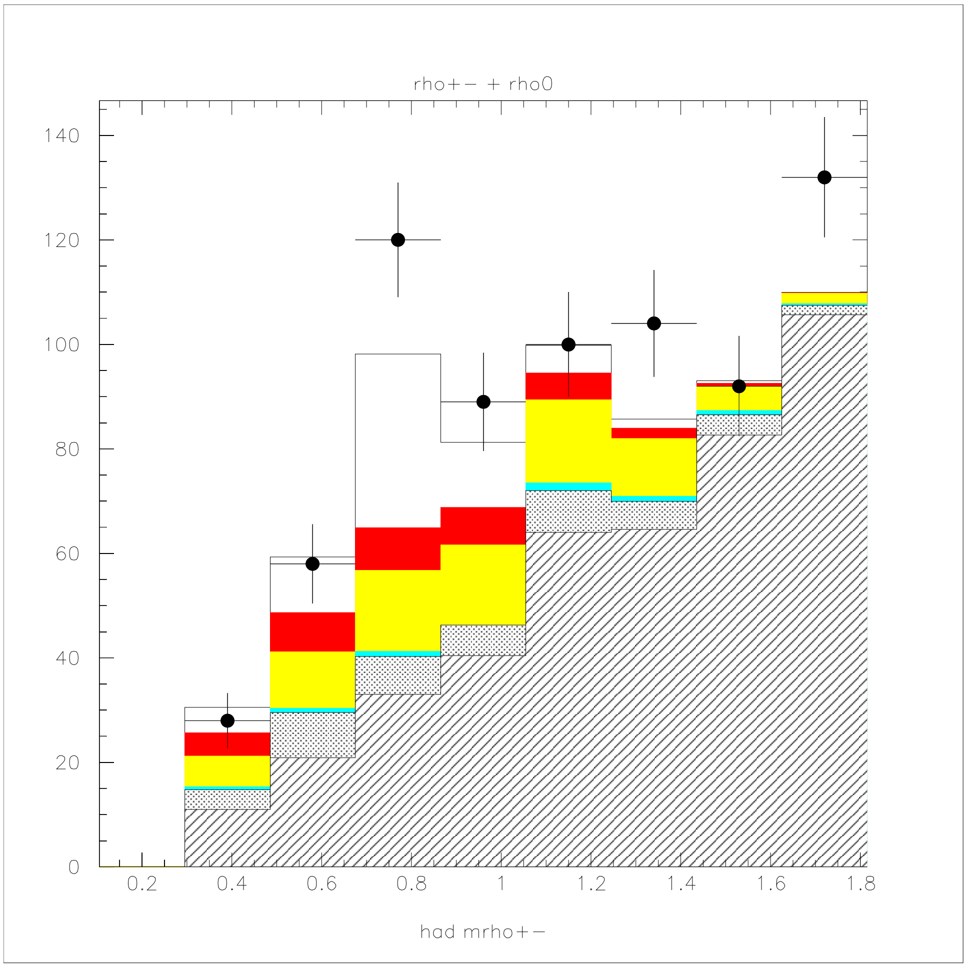

Figure 6.11: Combined $m_{\pi^{\pm}\pi^{0}}$ and $m_{\pi}^{+}\pi^{-}$ for the $\rho^{\pm}$ and ρ^{0} mode for q_2^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

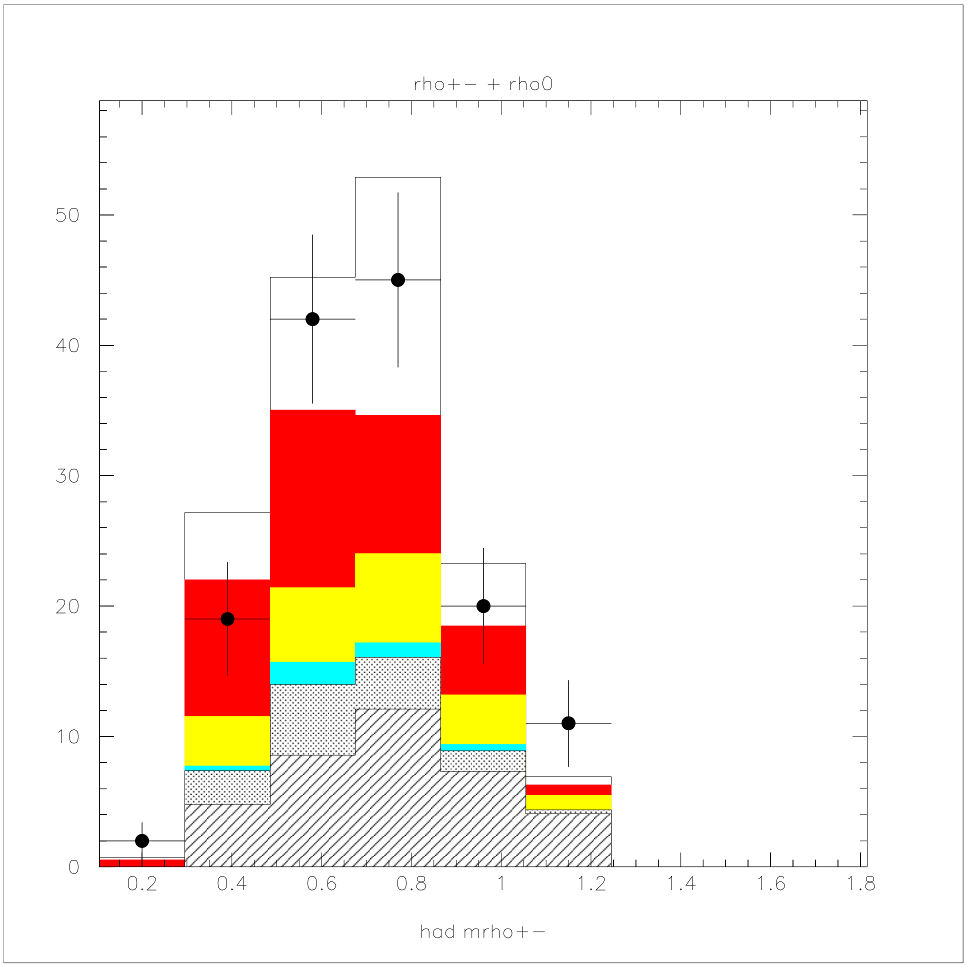

Figure 6.12: Combined $m_{\pi^\pm\pi^0}$ and $m_\pi^+\pi^-$ for the $\rho^\pm$ and ρ^0 mode for q_3^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

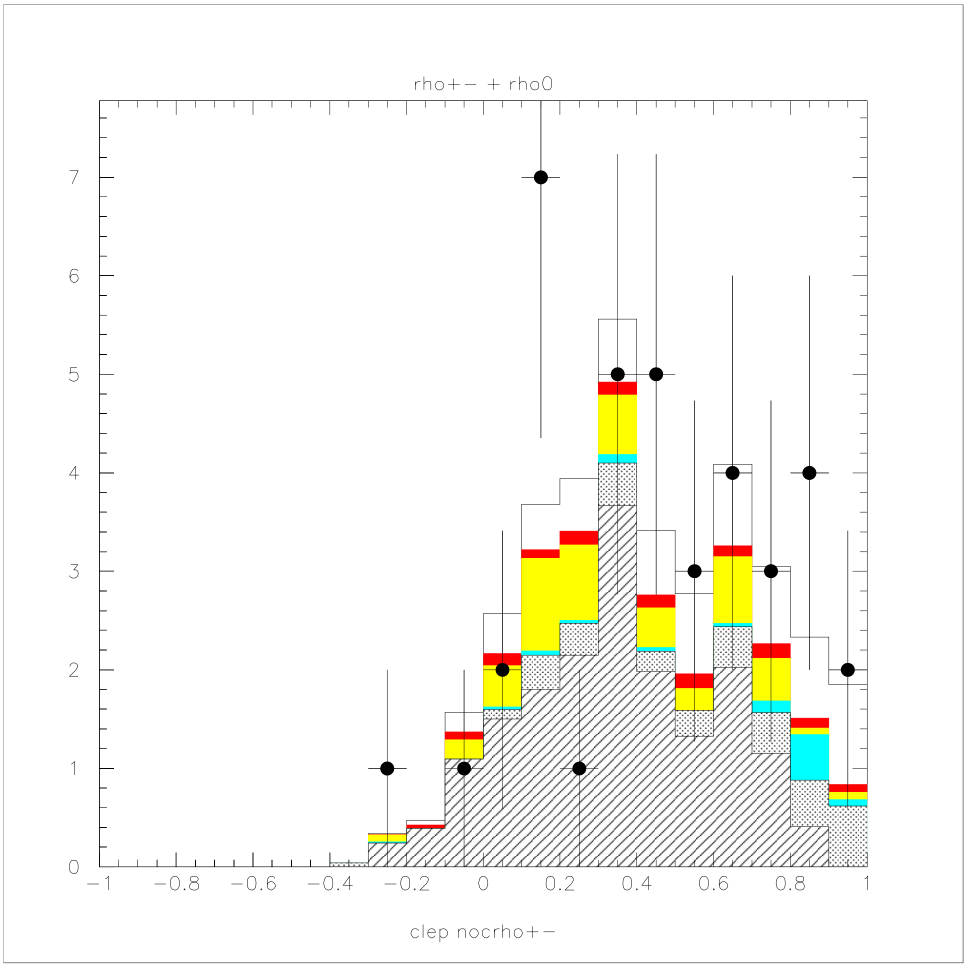

Figure 6.13: ρ mode $cos\theta_{lep}$ for q_1^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

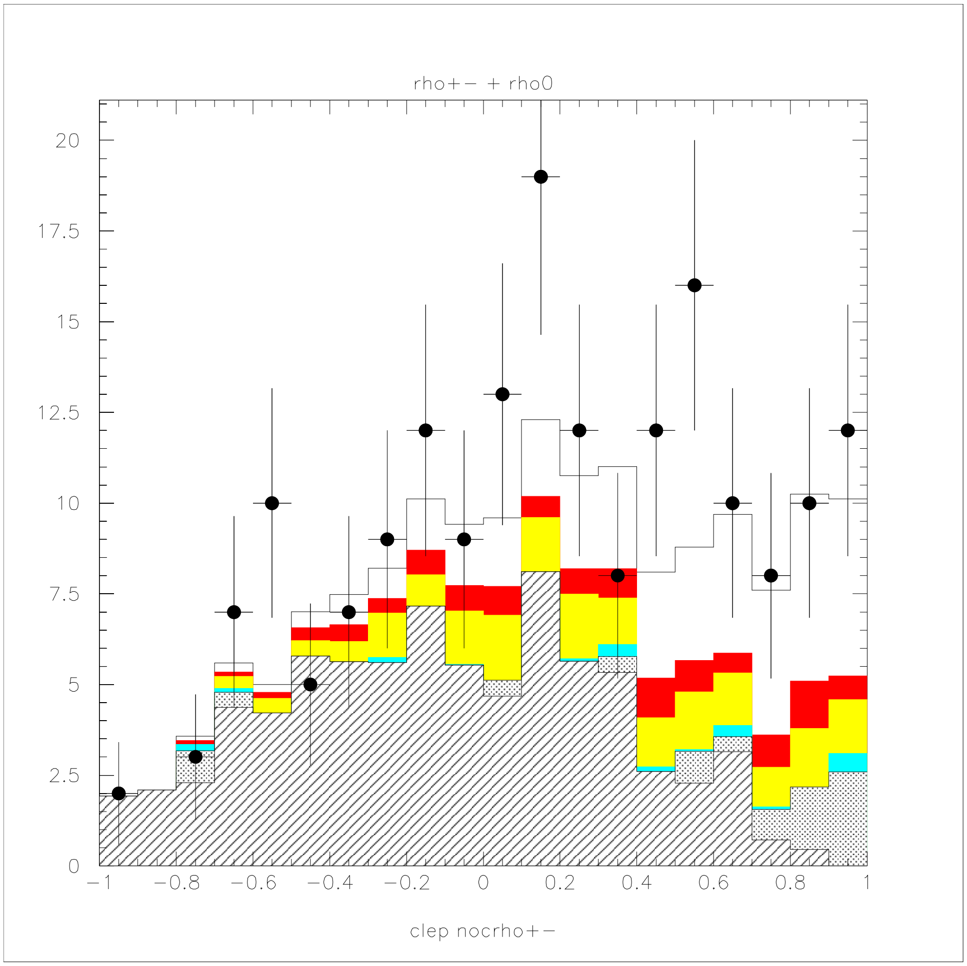

Figure 6.14: ρ mode $cos\theta_{lep}$ for q_2^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

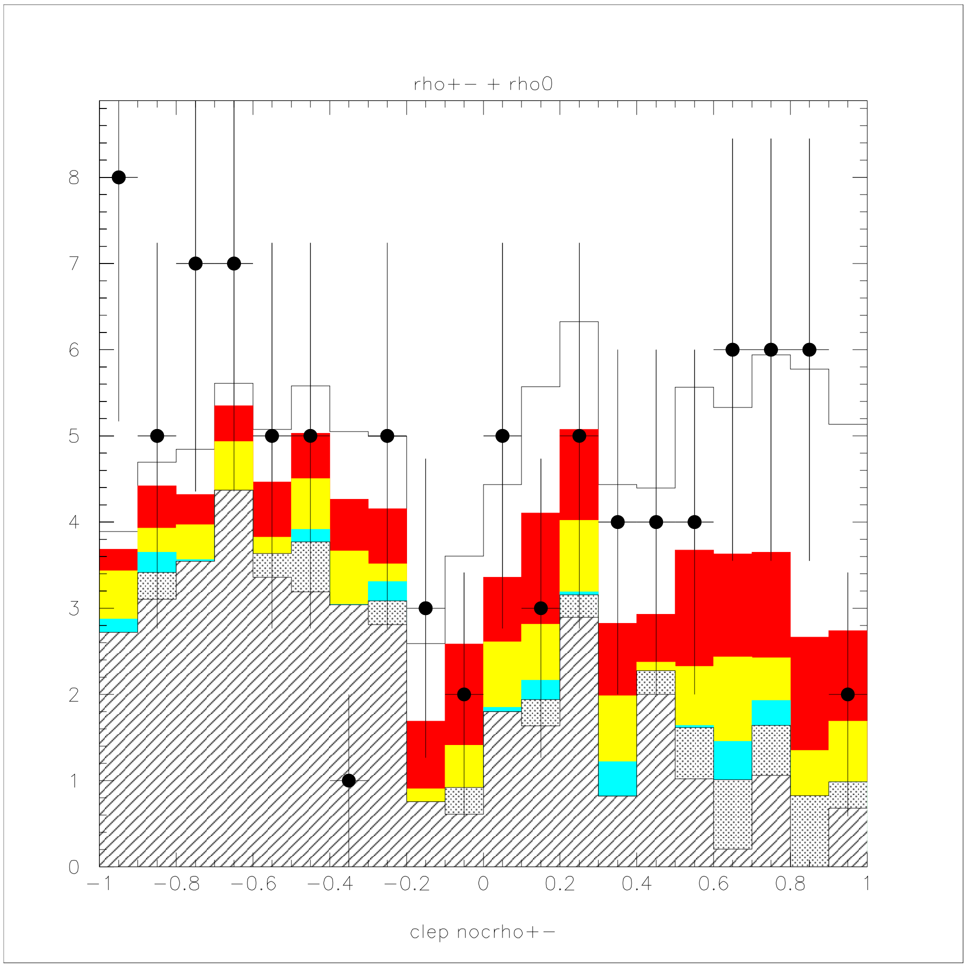

Figure 6.15: ρ mode $cos\theta_{lep}$ for q_3^2 and $\Delta Q = 0$. Open histogram is signal, black is cross-fees, light grey is btou, dark grey is fakes, dotted histogram is continuum, hatch histogram is btoc.

Table 6.1: Testing the stability of the fit

Cut	new value	mode	q_1^2	q_2^2	q_3^2	total
"V" cut	$\Delta Q0 : -1.2 - 0.8$	π	0.29	0.46	-0.06	0.25
	$\Delta Q1 : -0.8 - 0.8$	ρ	-0.37	-0.62	-2.22	-0.75
		η				0.22
lep mom	$p_{\pi,\eta} \geq 1.5GeV/c$	π	-3.72	1.35	-1.34	-1.22
	$p_\rho \geq 2.0GeV/c$	ρ	1.36	-3.43	5.35	-0.84
		η				0.66
$cos\theta_{lep}$ cut	No cut	π	0.00	0.58	-0.59	0.06
		ρ	-0.70	0.30	6.80	0.89
		η				0.37
electrons	muons difference:	π	0.53	-0.94	1.17	0.42
		ρ	0.32	-2.73	1.02	-1.51
		η				-0.85

Table 6.2: Comparisons with previous analysis

Analysis	$\mathcal{B}(\pi \times 10^{-4})$	$\mathcal{B}(\rho \times 10^{-4})$
current	1.77 ± 0.39	2.55 ± 0.33
1995 (ISGW2)	2.04 ± 0.47	2.24 ± 0.37

the pre-compress version of the tracking software. The comparisons are shown in table 6.2 and we conclude that the two analyses give the same results.

Testing the fitter

Finally, we have done some tests using the MC to see if the fitter is performing adequately. There are several things to test: given some known branching fraction of the MC, does the fitter return the right value, and is the treatment of the Poisson statistics adequate given the amount of luminosity of the data. To address the first point, we build some mock data by summing all of the sources in the amount that the fit is expecting. We fit this mock data with the same MC that was used in the mock data and we do reproduce the branching fraction used in the MC.

To address the second issue, we randomly select a number of events out of our MC sample that is representative of the data luminosity. For example, if we generated ten times more signal MC than the data luminosity, we need to randomly select one tenth of the MC sample. With those events, we build a mock data distribution and we fit it with the remainder of the MC for each source. If we repeat this procedure several times (N), we can make a distribution of the N branching fractions obtained, and their errors. We should expect that the mean of the branching fraction distribution be close to the branching fraction that was put in the MC. We also expect that the width of the branching fraction distribution should be close to the mean of the error distribution.

To perform this test we build the mock data to contain signal MC, btoc MC and btou other MC. Ideally, we would include $q\bar{q}$ MC to represent the continuum background. The problem is that based on the available amount of $q\bar{q}$ MC generated by the CLEO MC farms, we need to take 60% of the sample in each N try. This means that all the tries are highly correlated to each other and correlated to the parent sample. The statistical fluctuations present in the parent distribution get reproduced in each try, ending up into a coherent sum of these statistical fluctuations, hence the branching fractions are biased due to these fluctuations. In figures 6.16 and 6.17 we show a Gaussian fit to the 100 tries of the π branching fraction and of the error on that branching fraction reported by the fit, where the mock data and the fit contain btoc, btou other and signal MC. Figures 6.18 and 6.19 show the same thing for the ρ mode. In figure 6.20 we show the $-2ln\mathcal{L}$ quantity reported for each fit.

In table 6.3 we summarize the results from using btoc btou other and signal MC in the mock data. The description of each column is as followed: mode, expected branching fraction, mean of branching fraction distribution, width of branching fraction distribution, mean of the error distribution of the branching fractions, sigma associated with how the mean and the expected branching fraction agree, error on error is the percentage difference between the width and the mean of the

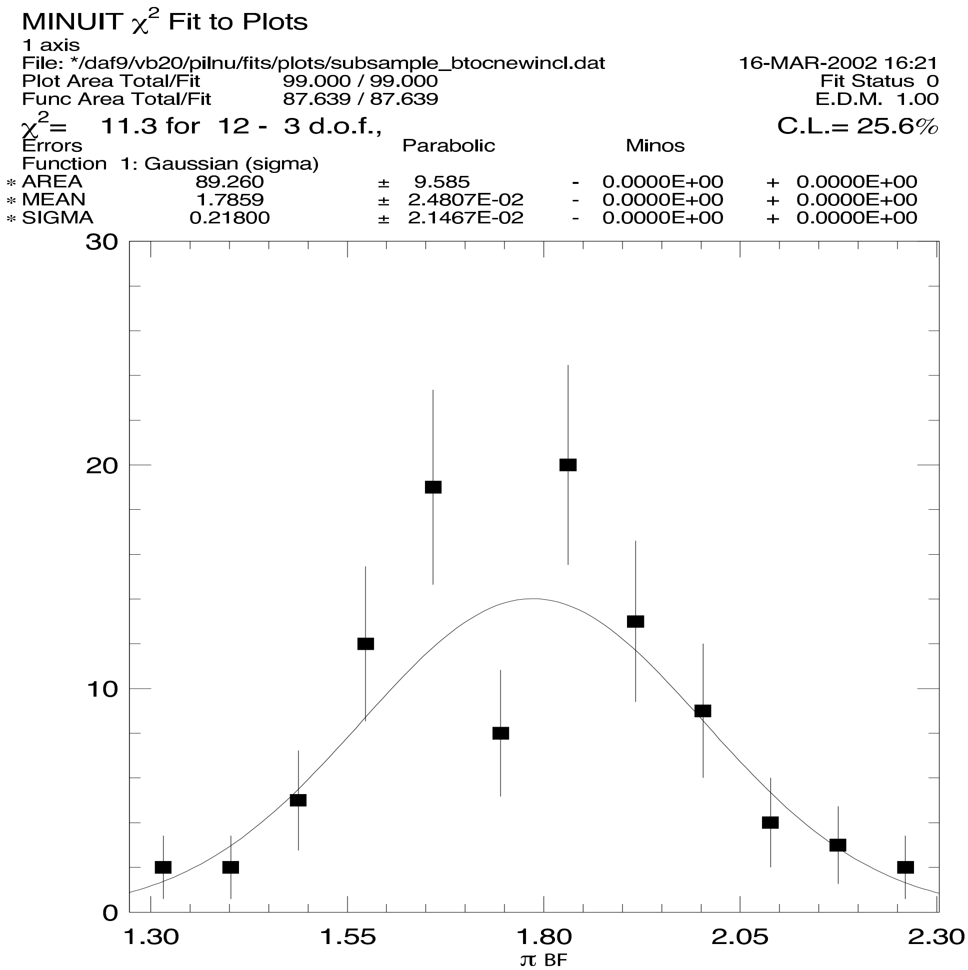

Figure 6.16: $B \to \pi\ell\nu$ branching fractions for 100 mock data fits.

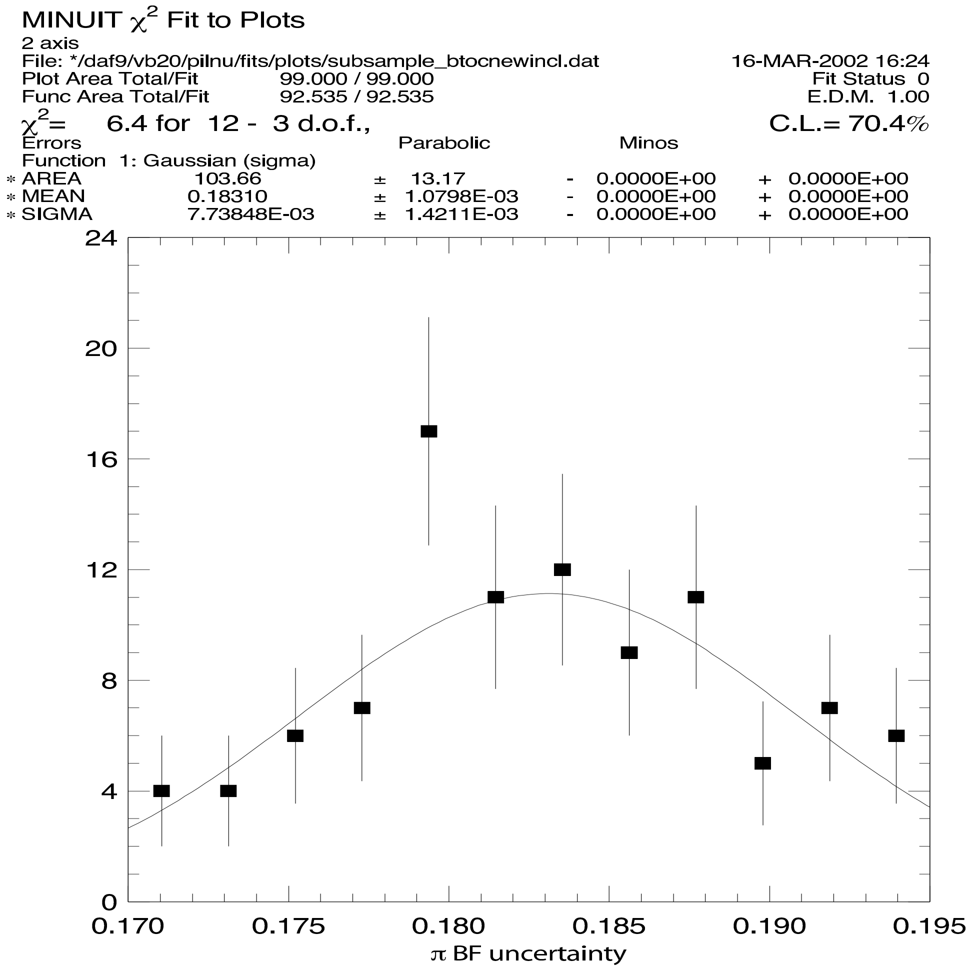

Figure 6.17: $B \to \pi\ell\nu$ branching fraction errors for 100 mock data fits.

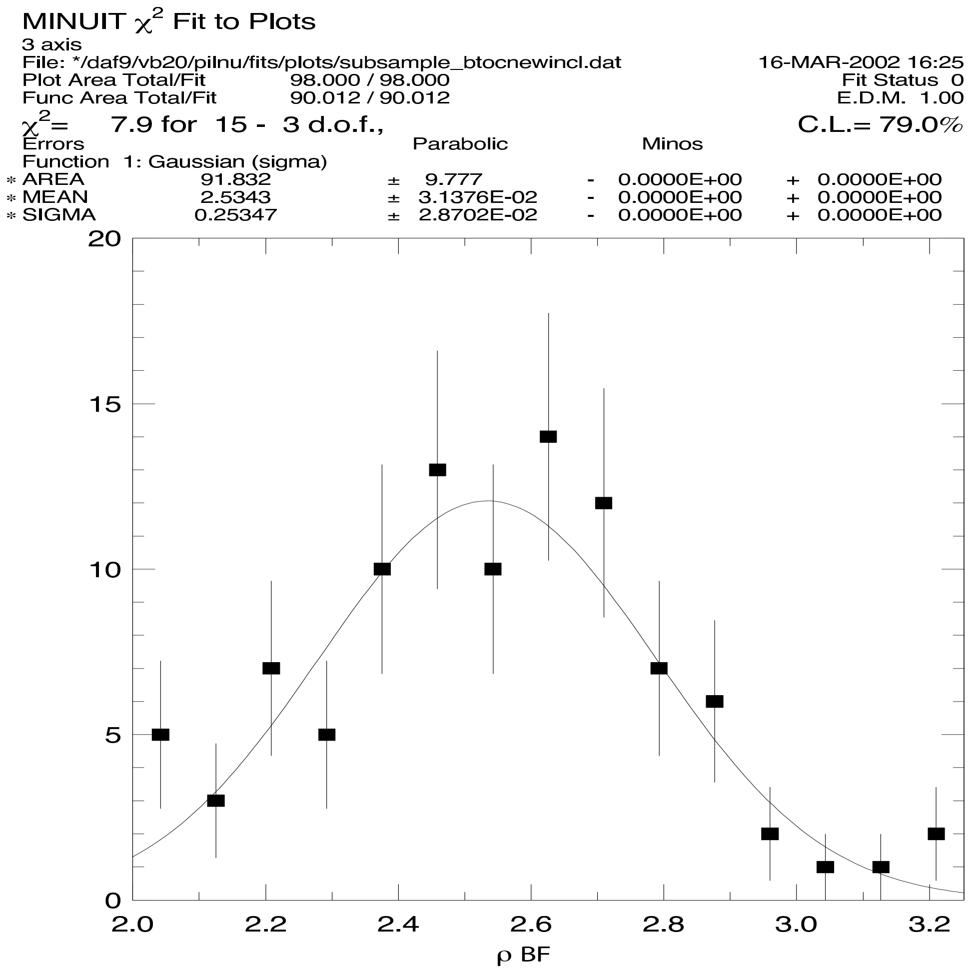

Figure 6.18: $B \to \rho\ell\nu$ branching fractions for 100 mock data fits.

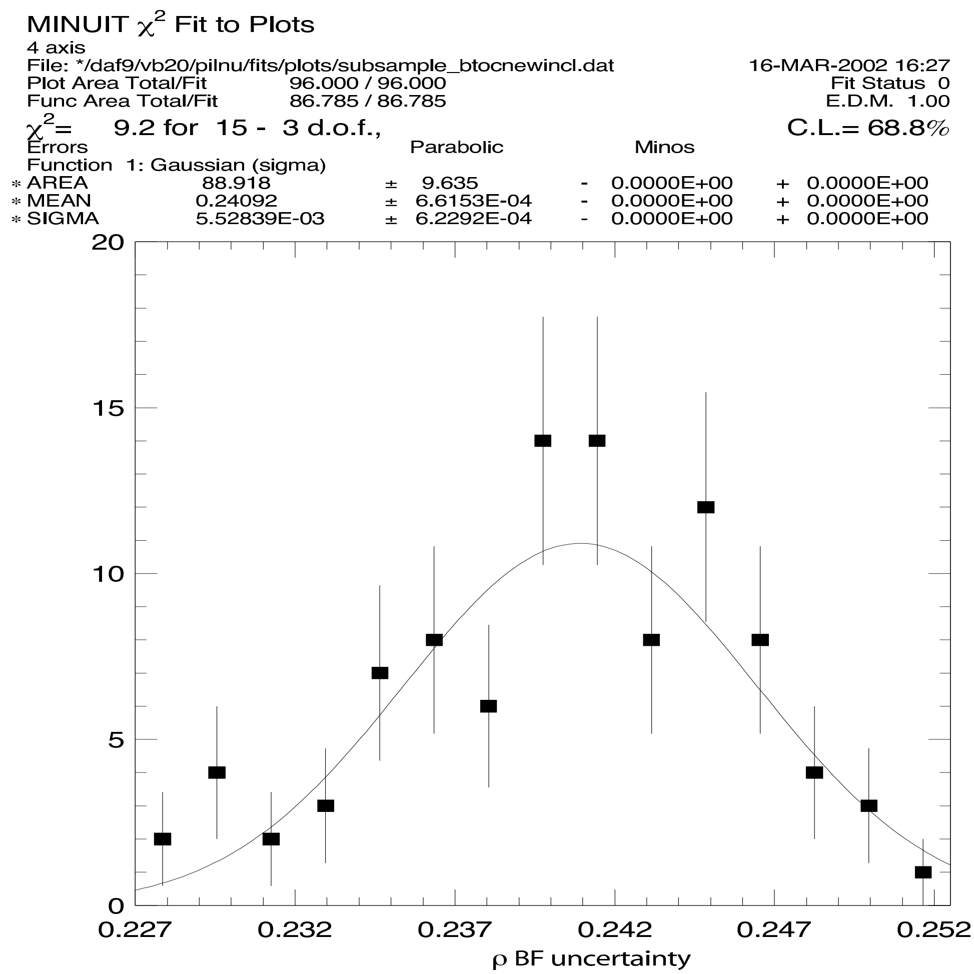

Figure 6.19: $B \to \rho\ell\nu$ branching fraction errors for 100 mock data fits.

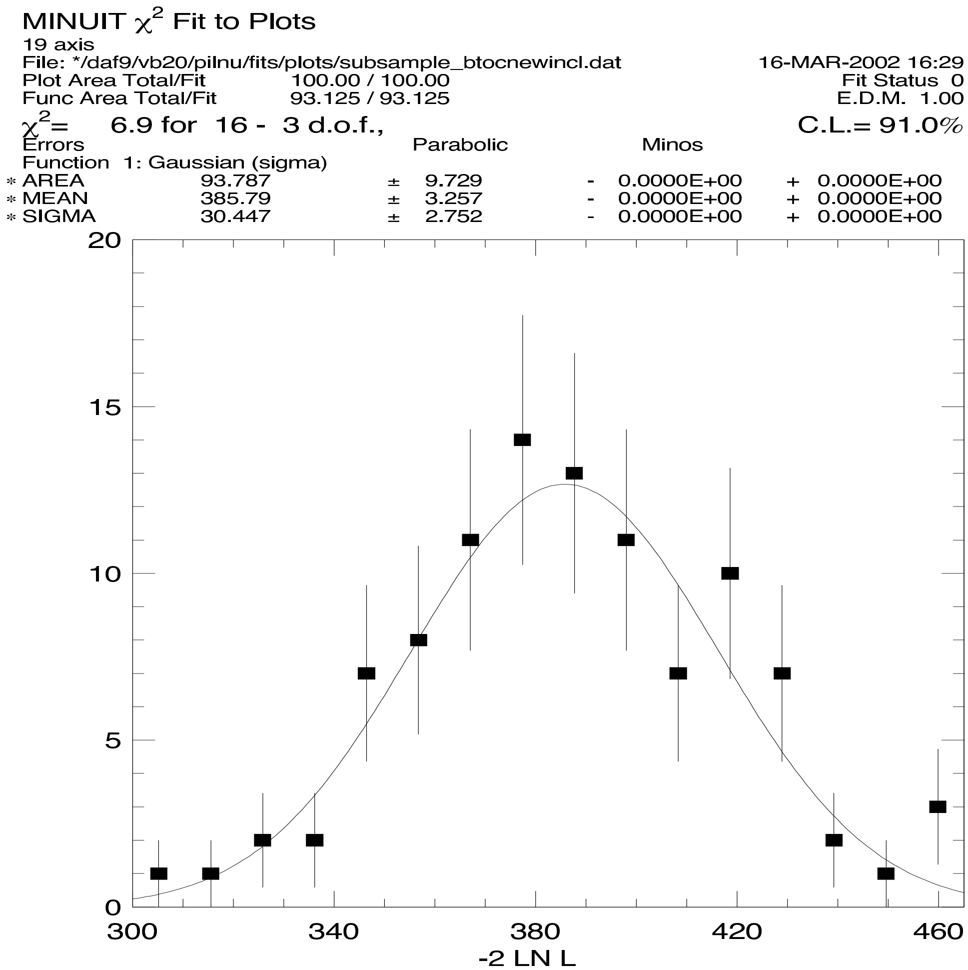

Figure 6.20: $-2ln\mathcal{L}$ for 100 mock data fits.

Table 6.3: Summary of mock data fitter test

mode	exp BF	mean	width	error mean	sigma	error on error	data error
πq_1^2	0.9	0.88	0.132	0.129	1.52	2.3%	23.5%
πq_2^2	0.65	0.63	0.107	0.101	1.87	5.6%	14.9%
πq_3^2	0.26	0.26	0.072	0.083	0.00	-15.3%	32.9%
π	1.8	1.784	0.216	0.183	0.74	15.3%	12.2%
ρq_1^2	0.64	0.646	0.125	0.132	-0.48	-5.6%	41.8%
ρq_2^2	1.33	1.386	0.213	0.183	-2.63	14.1%	20.0%
ρq_3^2	0.5	0.52	0.095	0.082	-2.11	13.7%	16.8%
ρ	2.5	2.532	0.255	0.242	-1.25	5.1%	13.8%
η	0.9	0.908	0.384	0.327	-0.21	14.8%	36.6%
mean of $-2ln\mathcal{L}$		$-2ln\mathcal{L}$ from data fit				σ	
386		394				0.3	

error distribution, data error comes from the data fit and so is not related to the mock data tries.

We see that the fitter does not introduce a bias on the branching fractions. We also see that it does a pretty good job at estimating the error. We know the 12% statistical uncertainty for the π mode to 15% of itself and the same with the η mode. As for the ρ, we know its 14% statistical uncertainty to 5% of itself.

6.2 The systematic uncertainties

This analysis depends heavily on MC. since we use such distributions in the fit. Our systematic uncertainties essentially come from imperfections of the MC simulation. The main ingredient affected by this is the neutrino reconstruction. We also study how a different composition of exclusive modes in the btoc background affects the analysis. We do a similar exercise with regard to the btou other background. Part of the btou other background is the non-resonant contribution ($B \to \pi\pi\ell\nu$) which, understandably, mainly affects the ρ mode. We study this uncertainty in detail. Finally, we discuss the remaining uncertainties related with the fake lepton sample, the continuum smoothing procedure, etc. A summary of all the systematic uncertainties is shown at the end of this section.

Detector simulation

How well the MC simulates the response of the detector is crucial in our analysis since the neutrino reconstruction depends on every track and shower in the detector. To quantify the possible effects on the results from discrepancies between the data and the MC, we vary knobs related to detector simulation and tabulate how much each knob affects the signal efficiencies, the yields and the branching fractions. The final uncertainty is a combination of the effect on those three quantities. We expect that the change in the branching fraction from the change in efficiency and the change in yield to cancel to some extent because if signal got out of the signal region and background got in, the efficiency will be lower in about the same way that the yield will be lower. However, we choose not to only take the change in branching fraction as the total uncertainty, since our simulation of the cancelation for a given knob can be disturbed by other effects. We choose to take the following combination,

$$\delta_{knob} = \Delta_{BF} \oplus \frac{1}{3}(\Delta_{yield} \oplus \Delta_{\epsilon}). \tag{6.1}$$

The choice of 1/3 is somewhat arbitrary but seems to be of the correct order.We describe each knob in turn, and summarize their effect at the end. Note that each knob is applied on the MC events only.

- **Photon-finding efficiency:** (PHOTON) we throw away 3% of clusters from photons. The effect of this knob is to lower the efficiency as well as the yield. Some amount of signal left the signal region, while some background went in. This is an example of a cancelation between the change in efficiency and the change in yield, so that the change in branching fraction is small. The standard for CLEO analyses is to take a loss of 2% in efficiency, so our uncertainty is: $\sigma = \frac{2}{3}\delta_{PHOTON}$.

- **Track-finding efficiency:** (TRFND) we throw away 0.75% (2.6%) of tracks with momentum greater (less) than 0.25 GeV/c. Those numbers come from the embedding study done for the $B \to D^* \ell \nu$ analysis ([51]). We take the full amount of this uncertainty, so $\sigma = \delta_{TRFND}$.

- **Fiducial cuts:** (FIDU) we apply cuts of $|cos\theta| < 0.93$ for tracks and $|cos\theta| < 0.91$ for photons, since around the beam line is a hard region to simulate. Those are extreme cuts, we decide to take as the uncertainty: $\sigma = \frac{1}{6}\delta_{FIDU}$.

- **Splitoff simulation:** (SPLSIMU) we randomly increase the number of hadronic splitoffs by a mean of 0.029 splitoffs per hadron. The energy of the added shower is taken randomly from the MC splitoff energy spectrum. To study the effect of the splitoff package we have compared a sample of $\gamma\gamma \to K_S K_S$ in data and MC, since those events only contain hadronic showers, so a lot of possible splitoff showers and no photon showers. We compared the multiplicities of splitoff showers in data and MC to obtain the value of the knob, see figure 6.21 for an example. We take the full amount of this uncertainty, so $\sigma = \delta_{SPLSIMU}$.

- **Splitoff Neural Net:** (SPLNN) We smear the value of the neural net output variable of the splitoff package, we make real photons look more like splitoffs and real splitoffs look more like photons. We obtained the value of the smearing by comparing the neural net output variable in data and MC for different energy bins, for the sample described in the previous knob. Figure 6.22 shows an example of such a comparison. We take the full amount of this uncertainty, so $\sigma = \delta_{SPLNN}$.

- **Shower Energy resolution:** (SHWRES) we degrade the shower energy resolution by 10%, so $E_{shower} \to E_{true} + 1.1(E_{shower} - E_{true})$. This is probably

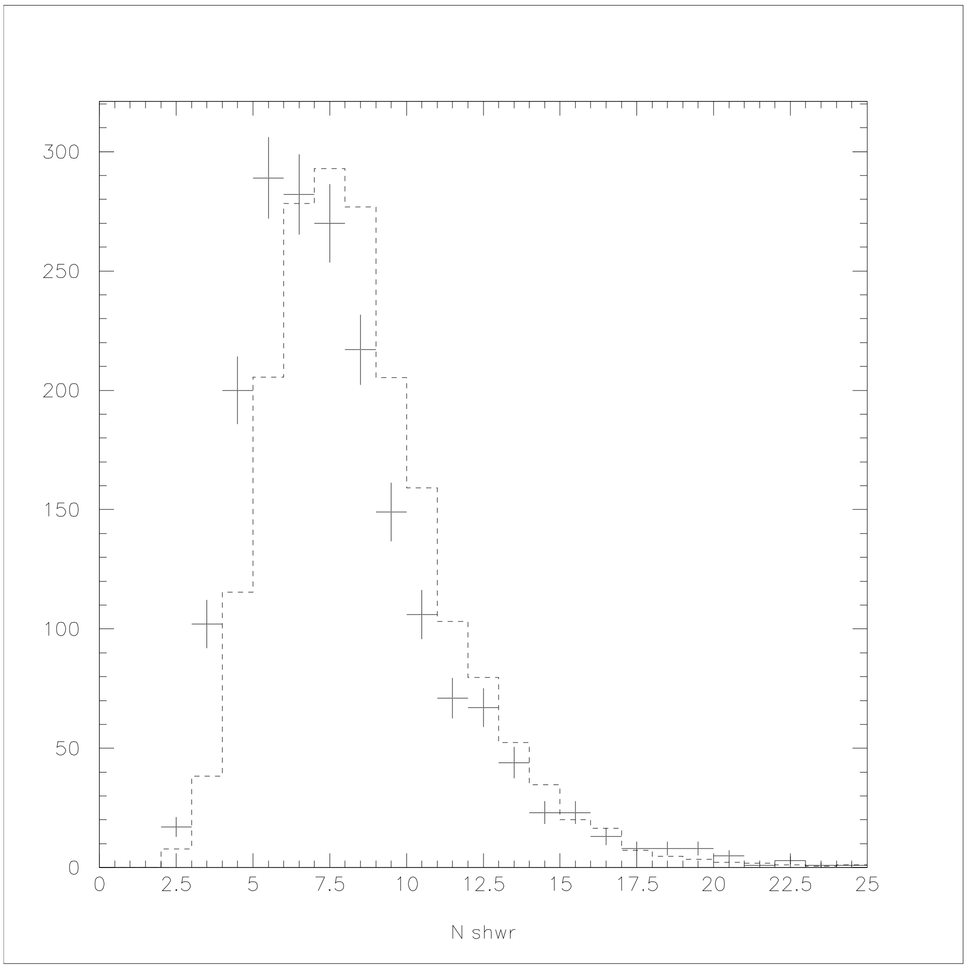

Figure 6.21: Number of showers in $\gamma\gamma \to K_S K_S$ event. Dashed histogram is MC, points are data.

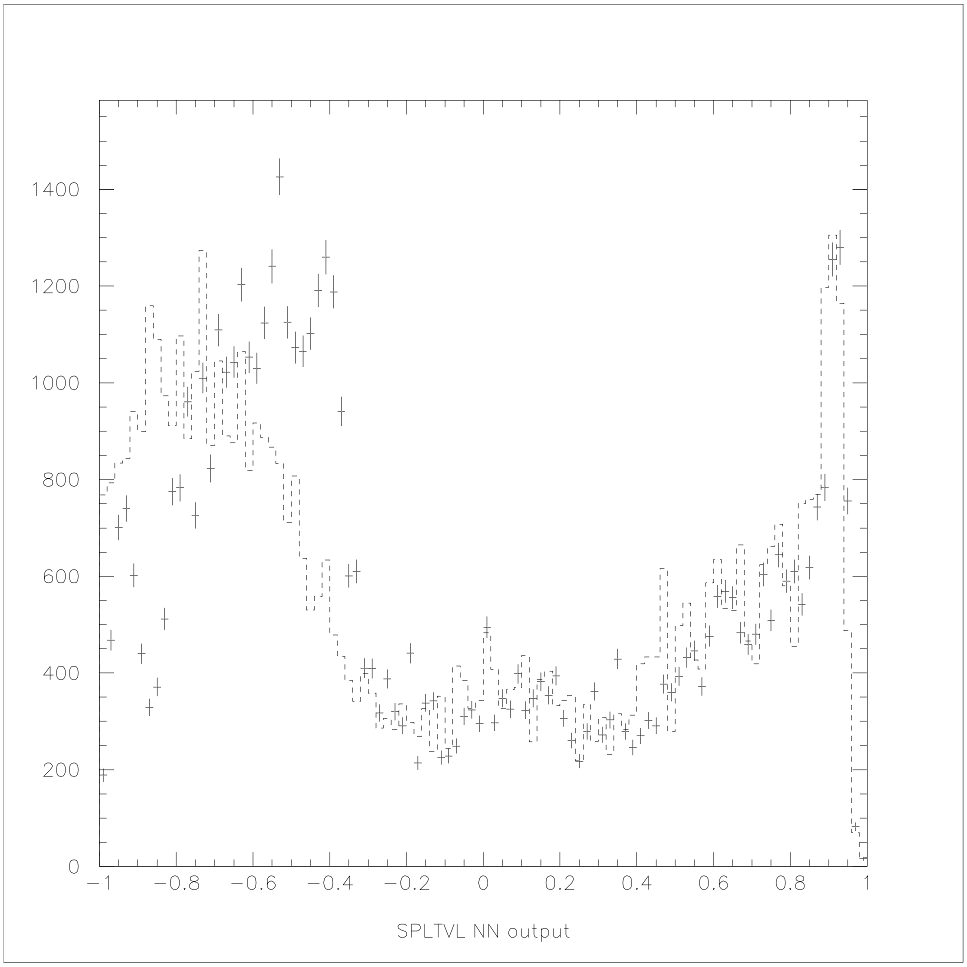

Figure 6.22: Neural Net output variable, -1 is photon like and +1 is splitoff like. Points are MC before smearing, dashed histogram is MC after smearing.

a conservative estimate of the resolution, but we still take the full amount of this uncertainty, so $\sigma = \delta_{SHWRES}$.

- **Track Momentum resolution:** (TRKRES) we degrade the track momentum resolution by 10%, so $P_{track} \rightarrow P_{true} + 1.1(P_{track} - P_{true})$. This is also probably a conservative estimate of the resolution, but we still take the full amount of this uncertainty, so $\sigma = \delta_{TRKRES}$.

- **Charged Particle ID:** (PID) we shift the sigmas of de/dx by 0.25 and of time of flight by 0.5, away from the true value. We take the full amount of this uncertainty, so $\sigma = \delta_{PID}$.

- K_L **Energy deposition:** (KLEDEP) we increase the amount of energy the K_L leave in the calorimeter by 20%. We take the full amount of this uncertainty, so $\sigma = \delta_{KLEDEP}$.

- K_L **Yield:** (KLYLD) for this knob we compared the yield of K_S in data and MC, since K_S's are produced at the same rate as K_L's. A discrepancy indicates a problem with the physics model, not a detector effect. We found that there were more K_S in data than in MC, as shown in figure 6.23. We apply a *correction* to our MC samples, where events containing a K_L are reweighted by 1.072^N where N is the number of K_L in the event. We made sure that the momentum dependence of the yield was comparable between data and MC. There is an uncertainty on the reweight number and that is the source of the uncertainty for this knob. The uncertainty is, and we take the full amount of this uncertainty, so $\sigma = \delta_{KLYLD}$.

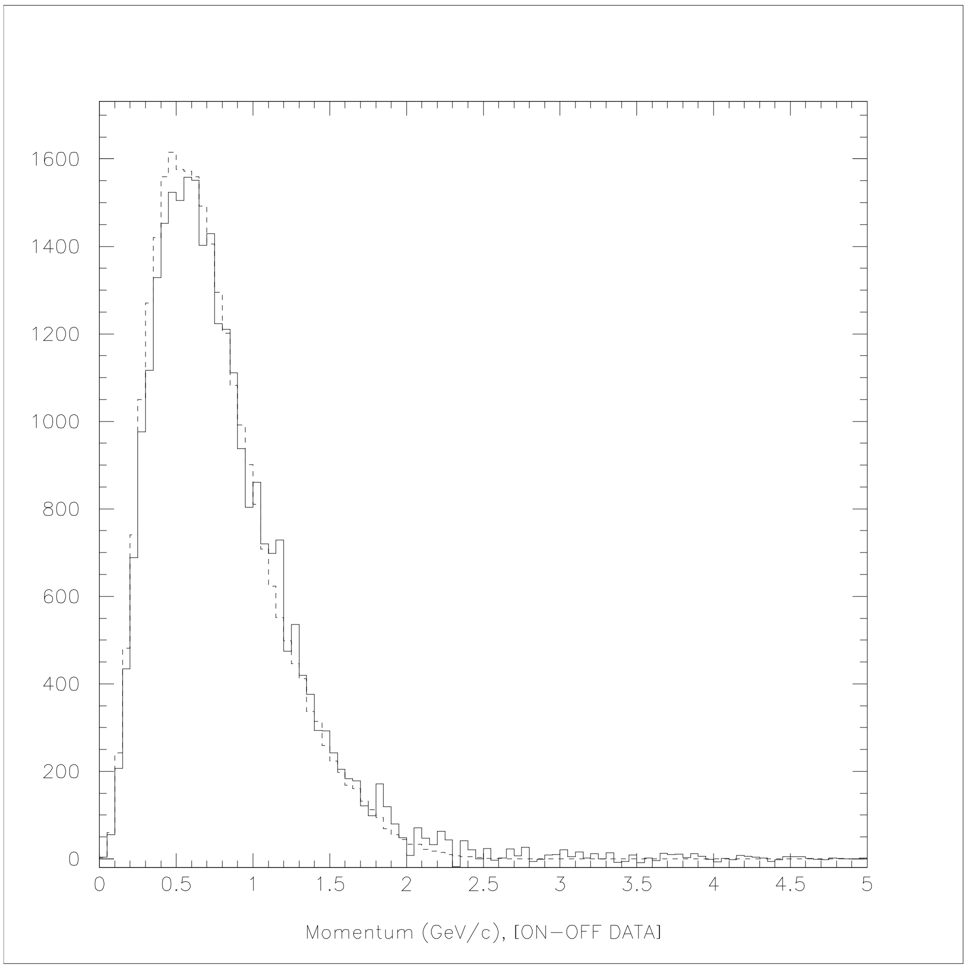

Figure 6.23: K_S momentum distribution. Solid histogram is data, dashed histogram is MC.

- **Secondary Lepton spectrum:** (SECLEP) Secondary leptons come from the process $b \to c \to s\ell\nu$, and have a soft momentum. The rate of this process affects our cut on the number of leptons in the event, for example. Because of the soft momentum, we often cannot identify those leptons as leptons, resulting in accepted events with a poorly-reconstructed neutrino. We find that the MC contains a branching fraction of 10.41% for this process, while the world average branching fraction is $8.3 \pm 0.4\%$ (PDG). So we have an excess of 20% in the MC for this process. We apply a *correction* to the MC samples by correcting the shape of the generated secondary lepton spectrum to match the measured one. The measured spectrum is taken as a convolution of the latest CLEO D meson momentum spectrum ([52]) with the semileptonic D meson momentum spectrum from the DELCO collaboration ([53]). We show the comparison of the two spectra in figure 6.24. We vary the reweighting procedure by $\pm 1\sigma$ to get an uncertainty associated with this knob. We take the full amount of this uncertainty, so $\sigma = \delta_{SECLEP}$.

Table 6.4 shows the effective uncertainty coming from each knob.

btoc modeling

In order to get a systematic uncertainty related with the btoc background modeling, we vary the relative branching fractions of $B \to D/D^{*}/D^{**}/D(n\pi)\ell\nu$ processes while keeping the total semileptonic rate constant. We vary each branching fraction by about 30% of its measured value. It might be argued that this is a somewhat extreme amount of vary the measured branching fraction, but this is to

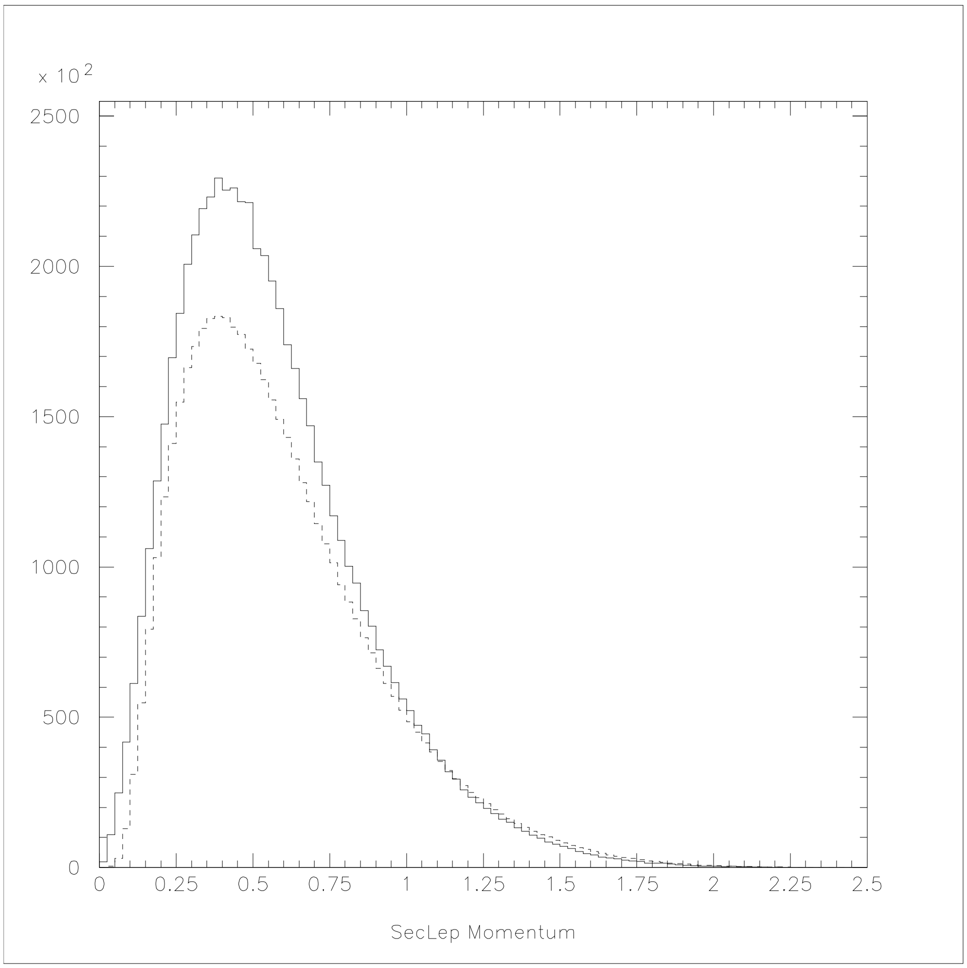

Figure 6.24: Secondary lepton momentum. Solid is MC, dashed is CLEO+DELCO data. The normalization is absolute.

Table 6.4: Knob turning systematic uncertainties, in percent

Knob	scale	π	πq_1^2	πq_2^2	πq_3^2	ρ	ρq_1^2	ρq_2^2	ρq_3^2	η
smear track params.	1.00	1.12	0.83	1.85	8.12	1.20	7.75	4.67	1.79	0.90
reduce photon-finding eff.	0.67	2.68	3.95	3.91	7.03	3.99	3.83	6.19	4.21	9.47
reduce track-finding eff.	1.00	2.90	4.65	4.80	8.97	2.88	12.13	5.72	5.44	3.68
smear photon shwr E's	1.00	2.33	1.64	4.79	2.15	7.25	5.44	15.34	0.90	2.96
Incr. K0L shower E's	1.00	1.32	1.03	1.70	2.43	1.32	0.42	1.33	2.14	3.88
incr. n of splitoff clusters/hadron	1.00	3.25	3.84	4.48	0.85	3.37	9.70	6.90	3.96	0.80
shift dE/dx and TOF sigmas high	1.00	2.18	3.08	3.25	8.55	1.10	20.52	2.41	6.67	1.55
fiduc. polar-angle cuts to MC trks shwrs	0.17	1.93	2.30	2.83	3.58	2.42	2.85	1.87	3.76	1.74
smear SPLITF NN output variable	1.00	1.67	0.98	2.27	4.45	2.19	1.91	4.31	1.25	3.38
secondary-lepton-spectrum correction	1.00	0.90	3.78	2.89	2.88	2.40	38.01	6.86	0.62	3.64
K_L yield correction	1.00	0.15	0.23	0.14	0.40	0.98	0.39	1.28	0.79	0.39
TOTAL		6.83	9.30	10.95	17.92	10.49	47.15	21.31	11.52	12.60

compensate for not varying the form factors of the various btoc transitions. It is important to note that the other B in signal MC as well as in the btou other MC, also has this variation applied to it, in addition to both B mesons in the btoc MC. Table 6.5 lists the effect of each variation.

Putting the $B \to D\ell\nu$ and $B \to D^*\ell\nu$ branching fractions to zero is an extreme change since they dominate the $b \to c$ rate, so we only include the variations high and low in the systematic uncertainty. We see that even such extreme variations result in a relatively small effect.

btou modeling

We already mentioned that for the btou other we have two models: one using ISGW2 to generate the $B \to X_u\ell\nu$ modes that are not our signal modes, and one that makes use of HQET and contains some prescription for non-resonant decays. Our nominal fit uses the second model (InclGen), and has the normalization fixed. The normalization was determined using the latest measurement of the $b \to u\ell\nu$ inclusive end point analysis ([54]), and also using the measured branching fractions for the signal modes from this analysis, in an iterative process. The uncertainty on the normalization comes mainly from the uncertainties on the current branching fractions. For a systematic uncertainty associated with the btou modeling, we vary the normalization factor by $\pm 1\sigma$ and also take into account the difference with using the ISGW2 model for the btou other background. Table 6.6 summarizes the systematic uncertainty coming from the btou other modeling.

Table 6.5: Btoc modeling systematic uncertainties, in percent.

Knob	scale	π	πq_1^2	πq_2^2	πq_3^2	ρ	ρq_1^2	ρq_2^2	ρq_3^2	η
$B \to D\ell\nu$ High	1.00	0.38	1.24	0.00	0.66	0.32	6.06	1.12	0.99	0.00
$B \to D\ell\nu$ Low	1.00	0.13	0.62	0.00	0.00	0.00	9.09	1.12	0.99	0.94
$B \to D\ell\nu$ Zero	0.00	[3.18]	[13.04]	[0.21]	[3.29]	[3.53]	[136.36]	[12.36]	[11.88]	[13.21]
$B \to D^*\ell\nu$ High	1.00	0.38	1.24	0.21	0.00	0.64	12.12	2.25	1.98	2.83
$B \to D^*\ell\nu$ Low	1.00	0.51	1.24	0.21	0.66	0.96	9.09	2.25	1.98	0.94
$B \to D^*\ell\nu$ Zero	0.00	[13.72]	[37.27]	[4.43]	[17.76]	[100.32]	[154.55]	[162.92]	[73.27]	[69.81]
$B \to D^{**}\ell\nu$ High	1.00	0.13	0.62	0.21	0.66	0.32	3.03	1.12	0.00	0.94
$B \to D^{**}\ell\nu$ Low	1.00	0.00	0.00	0.21	0.66	0.64	3.03	1.12	0.99	1.89
$B \to D^{**}\ell\nu$ Zero	1.00	0.25	0.62	0.84	1.97	0.64	9.09	2.25	0.99	3.77
$B \to Dnr\ell\nu$ High	1.00	0.38	0.00	0.42	0.66	0.64	0.00	0.56	0.99	0.00
$B \to Dnr\ell\nu$ Low	1.00	0.38	0.00	0.42	0.66	0.64	0.00	0.56	0.99	0.94
$B \to Dnr\ell\nu$ Zero	1.00	1.40	0.62	1.48	1.97	0.96	3.03	1.12	1.98	0.94
TOTAL		1.70	2.48	1.85	3.22	2.03	21.43	4.70	4.20	5.50

Table 6.6: Btou other modeling systematic uncertainty.

mode	% unc. btou modeling
πq_1^2	0.2
πq_2^2	1.1
πq_3^2	5.6
π	1.5
ρq_1^2	30.5
ρq_2^2	25.2
ρq_3^2	13.5
ρ	22.7
η	2.8

We see that for the π and η mode this systematic uncertainty is well under control, but it is clearly not the case for the ρ mode: this is by far the dominant uncertainty for this mode. For comparison, in the previous analysis, this uncertainty was 7.5%. We have tracked down the different ingredients that increased the uncertainty :

- Adding the II.V data increased the uncertainty to 10%
- Using the new cuts increased the uncertainty to 15%, which we expect since the cuts are much looser now
- Using the new way of calculating the scale factor for the btou other increased the uncertainty to 21%
- Using InclGen as the btou other model increased the uncertainty to 25%
- Not fitting over subregions 4,7,10 and having a higher M_B requirement for the hadron candidate reduced the uncertainty to 20%
- Finally, iterating over the btou other scale factor and using the new higher scale factor increased the uncertainty to 22.7%

It is important to note that when we optimized cuts, we only took into account the statistical effect and not the systematic one. It is clear that some of cuts should be tighter in light of this uncertainty. We have increased the lepton momentum cut to 2.0GeV/c and we would gain only a few percent reduction of the uncertainty. This is good news, since a higher lepton momentum cut would distort the various

q^2 bins. The other two cuts under consideration are the continuum suppresion and the "V" cut.

Non-resonant contribution

The dominant systematic error in the previous analysis was the effect of $B \to \pi\pi\ell\nu$ decays on the ρ branching fraction result. In this analysis, we are using a btou other MC sample that contains some non-resonant contribution. This means that the amount of non-resonant decays predicted by this model is taken into account in the level of btou other. Although this new btou model is based on HQET and on reasonnable assumptions about how to assign a decay to be exclusive or inclusive, it is still probably far from perfect, and one has to worry about the mismodeling of the non-resonant piece affecting the ρ branching fraction result.

To accomplish this, we generate some $B \to \rho\ell\nu$ MC where the ρ decays to two π^0's. That way we select out the portion of the non-resonant that will maximally affect the analysis. We reconstruct this mode and allow for it as a new source in the maximum likelihood fitter. The amount that the fitter seems to favor give us an idea of the mismodeling of that part in the btou other background. Dipion non-resonant decays are more than just $\pi^0\pi^0$, what about $\pi^\pm\pi^0$ and $\pi^+\pi^-$?

Since pions are spinless, they are bosons, and so their total wave function should be symmetric. Combining two pions together can give an isospin sum of 0,1 or 2, but since the two pions come from the combination of up and down quarks, only isospin 0 or 1 are allowed. If one looks at the decomposition of the I=1 and I=0 pieces in terms of the various $\pi^+\pi^-$ combinations, one finds that the I=1 part is antisymmetric and the I=0 is symmetric. This implies that in the I=1 part, there should be an odd orbital angular momentum between the two pions to make the wave function completely symmetric (since the spin part of the wave function is symmetric). Having a dipion pair with I=1 and L=1, is the ρ particle, given that the invariant mass of the pair is close to the mass of the ρ. Writing down the decomposition of $\pi^{\pm}\pi^{0}$, $\pi^{+}\pi^{-}$ and $\pi^{0}\pi^{0}$ using the Clebsch-Gordan coefficients, we find a ratio of 2:1:0. In the case of I=0, then we need an even orbital angular momentum, so it cannot be a ρ particle. In that case, we find a ratio of 0:2:1 (taking into account the fact that we can also have $\pi^{-}\pi^{+}$, in addition to $\pi^{+}\pi^{-}$). Once we know how much $\pi^{0}\pi^{0}$ non-resonant the fit is comfortable with, we see that twice that amount represents the level of $\pi^{+}\pi^{-}$ non-resonant.

Remaining systematic uncertainties

There are some systematic uncertainties left that do not enter in the previous categories. There is an uncertainty from the fake lepton sample. We run the analysis on the fake sample and change the fake rates by $\pm 1\sigma$. We assign a 1% systematic uncertainty for all modes and q^2 bins.

In our procedure for the continuum smoothing in section 5.1, we described how we fit MC M_B distributions to obtain the shape bias that the continuum suppresion cut could introduce. To get a systematic uncertainty we generate a random number between -1 and 1 and use it to scale the uncertainty on the shape parameters. We repeat this procedure five times with different seeds for the random numbers. Table 6.7 shows the assigned systematic uncertainties.

Table 6.7: Continuum smoothing systematic uncertainties, in percent

mode	% unc. cont. smooth.
πq_1^2	2
πq_2^2	0.2
πq_3^2	2
π	1
ρq_1^2	10
ρq_2^2	1
ρq_3^2	2
ρ	3
η	2

In our nominal fit we have assumed the relative production of charged and neutral B meson to be the same, $f^{+-}/f^{00} = 1$. We vary this fraction by $\pm 1\sigma$ using the latest measurement from CLEO: $f^{+-}/f^{00} = 1.04 \pm 0.08$ ([55]). This ratio enters in the denominator of the branching fraction expression, and also determines the normalization of the neutral modes vs the charged modes, since we are using the isospin relation. We have also assumed the ratio of B meson lifetimes to be unity. We vary this fraction by the measured value of $\tau_{B^+}/\tau_{B^0} = 1.062 \pm 0.029$ by $\pm 1\sigma$. The ratio comes into the normalization of the neutral modes vs the charged modes. We have also varied the isospin assumption, in the nominal fit we use a ratio of 2, so we vary this ratio to be 1.43 or 1.7, as suggested by Diaz-Cruz. Table 6.8 summarizes the uncertainties.

Finally, we assign a 2% uncertainty on the lepton identification efficiency, and also a 2% uncertainty on the number of B mesons. Those translate directly into a 2% uncertainty on the branching fractions.

Summary

Table 6.9 shows a summary of all the systematic uncertainties. Table 6.10 summarizes the final results for the branching fractions.

Notice that the π branching fraction has a factor of two better uncertainty than the previous analysis, with the statistical uncertainty being of the same size as the systematic uncertainty. The ρ branching fraction has the same combined statistical and systematic uncertainty as the previous analysis, but we have the bonus of the

Table 6.8: Various systematic uncertainties, in percent

mode	f_{+-}/f_{00}	τ_{B^+}/τ_{B^0}	isospin
πq_1^2	-2.6	-0.1	0.0
πq_2^2	-2.3	-0.3	0.0
πq_3^2	-2.2	-0.5	-0.2
π	-2.4	-0.2	0.0
ρq_1^2	2.5	-4.2	1.9
ρq_2^2	-1.0	-1.4	2.7
ρq_3^2	-0.1	-2.1	2.3
ρ	0.0	-2.1	2.4
η	4.1	-1.4	-0.1

Table 6.9: Summary of all systematic uncertainties, in percent

Systematic	π	πq_1^2	πq_2^2	πq_3^2	ρ	ρq_1^2	ρq_2^2	ρq_3^2	η
ν reconstr.	7	9	11	18	10	47	21	12	13
btoc mod.	1.7	2.5	1.9	3.2	2	21.4	4.7	4.2	5.5
btou other	1.5	0.2	1.1	5.6	22.7	30.5	25.2	13.5	2.8
cont. sm.	1	2	0.2	2	3	10	1	2	2
fakes	1	1	1	1	1	1	1	1	1
f_{+-}/f_{00}	2.4	2.6	2.3	2.2	0	2.5	1	0.1	4.1
τ_{B^+}/τ_{B^0}	0.2	0.1	0.3	0.5	2.1	4.2	1.4	2.1	1.4
isospin	0	0	0	0.2	2.4	1.9	2.7	2.3	0.1
lep id	2	2	2	2	2	2	2	2	2
luminosity	2	2	2	2	2	2	2	2	2
Upper Unc.	**8.4**	10.3	11.8	19.6	**25.4**	61.1	33.4	19.1	**15.5**
Non Resonant	-5	-5	-5	-5	-14	-14	-14	-14	-5
Lower Unc.	**9.7**	11.5	12.9	20.2	**29.0**	62.7	36.3	23.7	**16.2**

Table 6.10: Branching Fractions

Mode	Branching Fraction ($\times 10^{-4}$)	Total Unc.
πq_1^2	$0.468 \pm 0.110^{+0.048}_{-0.054}$	$\pm$ 0.12 (26%)
πq_2^2	$0.725 \pm 0.108^{+0.086}_{-0.093}$	$\pm$ 0.14 (20%)
πq_3^2	$0.277 \pm 0.091^{+0.054}_{-0.056}$	$\pm$ 0.11(39%)
π	$\mathbf{1.47} \pm 0.179^{+0.123}_{-0.143}$	$\pm$ **0.229 (16%)**
ρq_1^2	$0.297 \pm 0.124^{+0.181}_{-0.186}$	$\pm$ 0.22 (75%)
ρq_2^2	$0.866 \pm 0.173^{+0.290}_{-0.314}$	$\pm$ 0.36 (41%)
ρq_3^2	$0.494 \pm 0.083^{+0.095}_{-0.117}$	$\pm$ 0.14 (29%)
ρ	$\mathbf{1.657} \pm 0.228^{+0.422}_{-0.481}$	$\pm$ **0.53 (32%)**
η	$\mathbf{0.837} \pm 0.306^{+0.129}_{-0.136}$	$\pm$ **0.34 (40%)**

q^2 distribution, which will lead to a lower model dependence uncertainty on $|V_{ub}|$. The η branching fraction is dominated by the statistical uncertainty.

CHAPTER 7

MEASURING V_{ub} AND CONCLUSION

The last three chapters described the branching fraction measurement of the decays $B \to (\pi, \rho, \omega, \eta)\ell\nu$, in three q^2 bins. They are important measurements in their own right since they are the dominant exclusive modes making up the inclusive $b \to u$ rate, and may be some information can be extracted from them related to factorization (by comparing those to the $B \to \pi\pi$ branching fraction for example). The branching fractions are also the necessary ingredients for getting a $|V_{ub}|$ measurement from exclusive decays. There are two main steps in getting a $|V_{ub}|$ measurement with complete uncertainties: the first step is to get the statistical and systematic uncertainty, the second step is to get the model dependence uncertainty. We show here the first step of this process, since we have used a single model throughout our discussion, ISGW2. We first describe the fitting procedure, and then we repeat the final results of this study, along with some concluding remarks.

Table 7.1: γ_u (in ps x 100) from ISGW2.

Mode	q_1^2	q_2^2	q_3^2
π	0.04728	0.03441	0.01340
ρ	0.03682	0.07638	0.02868

7.1 The $|V_{ub}|$ fit

To fit for $|V_{ub}|$ we perform a simple chi-square fit of the six different partial branching fractions (three for the π mode and three for the ρ mode). The chi-square can be written as,

$$\chi^2 = \sum_{jk} \frac{[\mathcal{B}_j - \mathcal{B}_{th_j}][\mathcal{B}_k - \mathcal{B}_{th_k}]}{c_{jk}\sigma_j\sigma_k}, \tag{7.1}$$

where c_{jk} are the correlation coefficients from the branching fraction fit, σ is the statistical uncertainty on the measured branching fraction and $\mathcal{B}_{th}$ are the theoretical branching fractions, given by

$$\mathcal{B}_{th} = \tau_B \gamma_u V_{ub}^2, \tag{7.2}$$

where τ_B is the B lifetime, γ_u is the integral over the form factors for the particular q^2 bin and for the model in question. Table 7.1 shows the values used for γ_u (in ps x 100) coming from ISGW2.

Table 7.2 shows the correlation coefficients for the nominal fit results.

Table 7.2: Correlation coefficients for the nominal fit.

	π_1	π_2	π_3	ρ_1	ρ_2	ρ_3
π_1	1.0	-0.039	0.004	-0.088	0.017	0.002
π_2		1.0	-0.033	-0.022	-0.126	0.059
π_3			1.0	0.004	0.009	-0.465
ρ_1				1.0	-0.039	0.043
ρ_2					1.0	-0.342

To get the central value and statistical uncertainty on $|V_{ub}|$, we use the central values of the branching fractions and their statistical uncertainties. Figures 7.1 and 7.2 show the result of the nominal $|V_{ub}|$ fit for $d\Gamma/dq^2$ for the π and ρ mode respectively, including systematic uncertainties. The fit finds a value of $|V_{ub}| = (2.913 \pm 0.128) \times 10^{-3}$, with a $\chi^2 = 14.1$, which corresponds to a probability of 1.5%.

To get the systematic uncertainties, we run the $|V_{ub}|$ fit using as the branching fractions the central values coming from the systematic change, as well as the associated statistical uncertainty. We choose to investigate the dominant branching fraction systematics: the btou other scale factor, the shower Energy smearing and the increased number of splitoff showers per hadron. We guesstimate the contribution from the $\pi\pi$ non-resonant uncertainty based on the effect of this uncertainty on the previous analysis. Table 7.3 summarizes the systematic uncertainty.

The final result is:

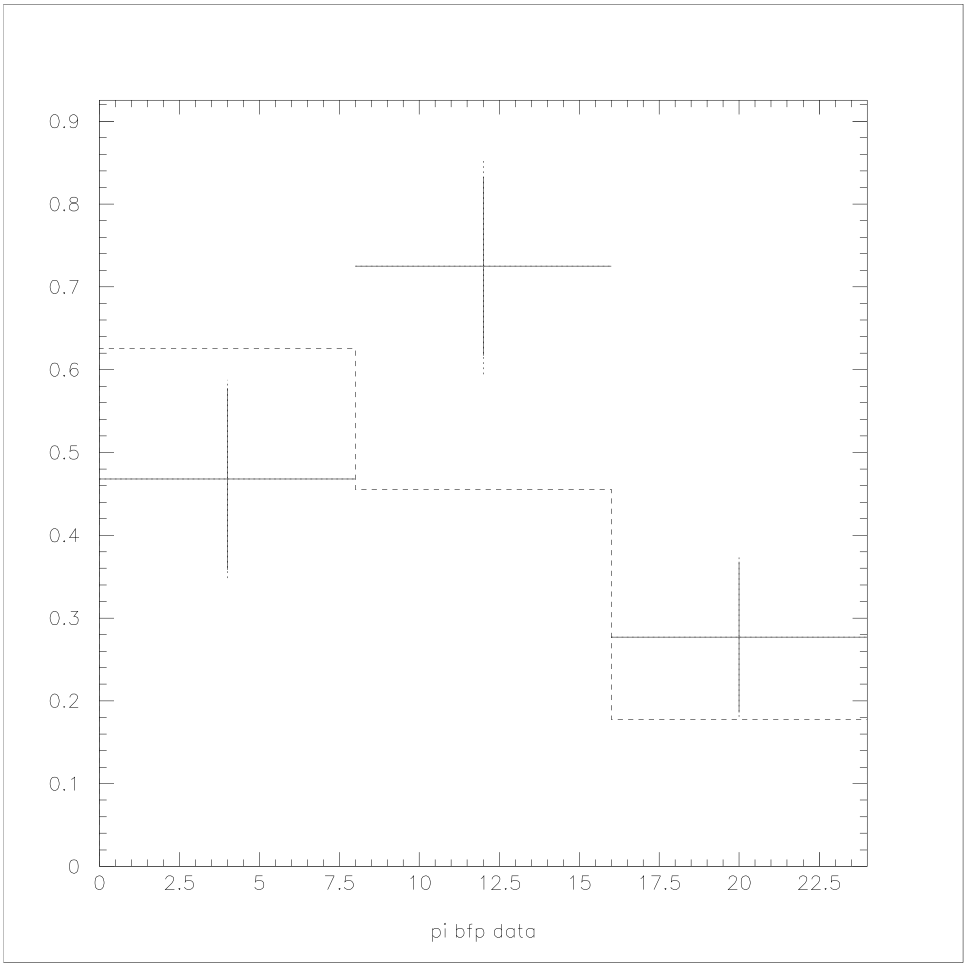

Figure 7.1: $d\Gamma/dq^2$ vs q^2 for π mode. Solid are the measured branching fractions with statistical uncertainties, dashed are the theoretical branching fraction (where $|V_{ub}|$ is from fit), dotted are the systematic uncertainties on the measured branching fractions.

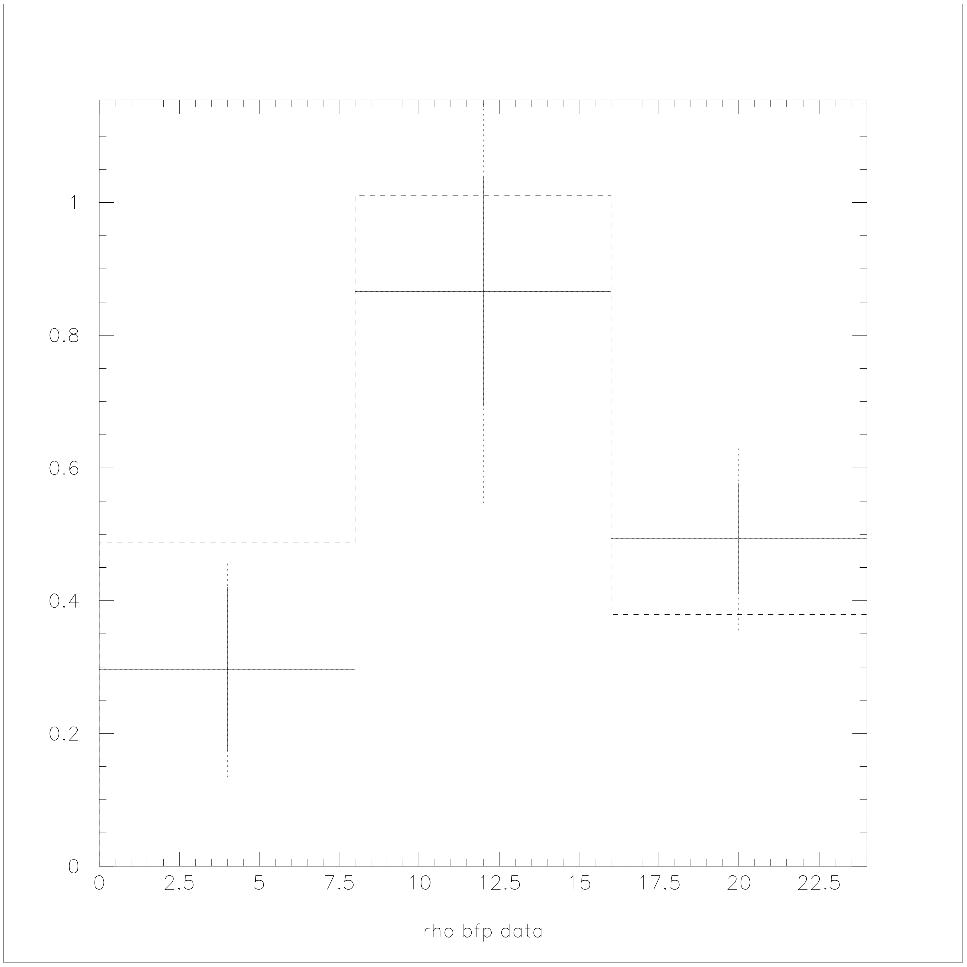

Figure 7.2: $d\Gamma/dq^2$ vs q^2 for ρ mode. Solid are the measured branching fractions with statistical uncertainties, dashed are the theoretical branching fraction (where $|V_{ub}|$ is from fit), dotted are the systematic uncertainties on the measured branching fractions.

Table 7.3: Systematic uncertainties on $|V_{ub}|$.

systematic	$\|V_{ub}\| \times 10^{-3}$	stat. unc.	Prob	χ^2	%
Nominal	2.913	0.128	1.48%	14.1	
btou scale+1σ	2.752	0.135	0.243%	18.5	
btou scale-1σ	3.063	0.123	2.89%	12.5	5.34%
shower E smear	2.924	0.133	0.939%	15.2	0.41%
spltf shw/had	2.859	0.130	0.851%	15.5	1.84%
Upper total:					5.7%
$\pi\pi$ NR					-5%
Lower total:					-7.6%

$$\boxed{\text{Current Measurement:}|V_{ub}| = (2.913 \pm 0.128^{+0.165}_{-0.220}) \times 10^{-3}} \tag{7.3}$$

We compare this result to the average of the previous exclusive CLEO measurements ([39]):

$$\text{CLEO exclusive 1998:}|V_{ub}| = (3.25 \pm 0.14^{+0.21}_{-0.29} \pm 0.55) \times 10^{-3} \tag{7.4}$$

where the last uncertainty comes from model dependence, and was the dominant uncertainty. We see that combining statistical and systematic uncertainties the current measurement is marginally better than the previous measurement (8.7% vs 9.9%), but we do expect some improvement regarding the model dependence in the future because of the information provided by the various q^2 bins.

We also compare this result to the latest inclusive CLEO measurement ([54]):

$$\text{CLEO inclusive 2002:}|V_{ub}| = (4.08 \pm 0.34 \pm 0.44 \pm 0.29) \times 10{-}3 \tag{7.5}$$

To be competitive with this measurement (15% total uncertainty), the model dependence for our measurement would have to be equal or better than 12.5%. We do not think this is a reasonable expectation, since the uncertainty on the normalization from the models are of the order 15-20%. It is important to note that the inclusive measurement has an additional unknown systematic uncertainty coming from the quark hadron duality assumption. If one assumes a 12.5% model dependence uncertainty, then our measurement is about 2σ away from the inclusive measurement. If this discrepancy is real, it could indicate that the quark hadron duality assumption is not as unsignificant as currently assumed.

7.2 Conclusion

We repeat here the main results obtained with this measurement:

Mode	Branching Fraction ($\times 10^{-4}$)	Total Unc.
πq_1^2	$0.468 \pm 0.110^{+0.048}_{-0.054}$	$\pm$ 0.12 (26%)
πq_2^2	$0.725 \pm 0.108^{+0.086}_{-0.093}$	$\pm$ 0.14 (20%)
πq_3^2	$0.277 \pm 0.091^{+0.054}_{-0.056}$	$\pm$ 0.11(39%)
π	$\mathbf{1.47} \pm 0.179^{+0.123}_{-0.143}$	$\pm$ **0.229 (16%)**
ρq_1^2	$0.297 \pm 0.124^{+0.181}_{-0.186}$	$\pm$ 0.22 (75%)
ρq_2^2	$0.866 \pm 0.173^{+0.290}_{-0.314}$	$\pm$ 0.36 (41%)
ρq_3^2	$0.494 \pm 0.083^{+0.095}_{-0.117}$	$\pm$ 0.14 (29%)
ρ	$\mathbf{1.657} \pm 0.228^{+0.422}_{-0.481}$	$\pm$ **0.53 (32%)**
η	$\mathbf{0.837} \pm 0.306^{+0.129}_{-0.136}$	$\pm$ **0.34 (40%)**

$$\boxed{|V_{ub}| = 2.913 \pm 0.128^{+0.165}_{-0.220}} \tag{7.6}$$

As previously stated, to get a $|V_{ub}|$ measurement close to 15%, we need a model dependence close to 12.5%, which we think is unlikely with the current status of QCD calculations. Although both the branching fraction measurements as well as the $|V_{ub}|$ measurement are consistent with both the previous exclusive analysis and the current inclusive analysis, it is interesting to note that the ρ branching fraction is about 30% smaller than the previous measurement.

Most of the difference in the ρ mode arises from the use of a different scale factor in scaling the btou other background: since the new scale factor is bigger (by about a factor of two), there is more btou other background and so less $\rho\ell\nu$ signal. We have also seen how the btou other is, by far, the source of the dominant branching fraction systematic uncertainty for this mode. It is clear that the focus should now be on this effect. Although the large systematic uncertainty from the btou other background is not very important as a systematic uncertainty on $|V_{ub}|$, the decreased branching fraction leads to a smaller central value for $|V_{ub}|$. The new central value is about 10% smaller than the previous exclusive measurement and about 30% smaller than the current inclusive measurement.

In figure 7.3 we show a CKM fit[1] to current values that determine the ρ η plane (see chapter 1 for more information on the CKM triangle) ([56]). We see that the current measurement of $|V_{ub}|$ is only somewhat consistent with the 1σ band of $sin2\beta$. We stress that those are impressionistic arguments and that with the current size of uncertainties no scientific judgment can be inferred from the measurements.

What is clear from the previous $|V_{ub}|$ measurement is that the dominant uncertainty is from the model dependence. The model dependence comes through on a

[1] Values used in the fit: $|V_{ub}|$ = 2.913$\pm$0.444, $|V_{ud}|$=0.97394 $\pm$ 0.00089, $|V_{us}|$=0.2200 $\pm$0.0025,$|V_{cd}|$=0.224$\pm$0.014, $|V_{cb}|$=40.4D-03 $\pm$ 1.3D-03 $\pm$ 0.9D-03, Δm_d= 0.489$\pm$ 0.008,Δm_s= 0.0 $\pm$ -1.0, $|\epsilon_K|$= 2.271D-03$\pm$ 0.017D-03, $\sin 2\beta$=0.793 $\pm$ 0.102

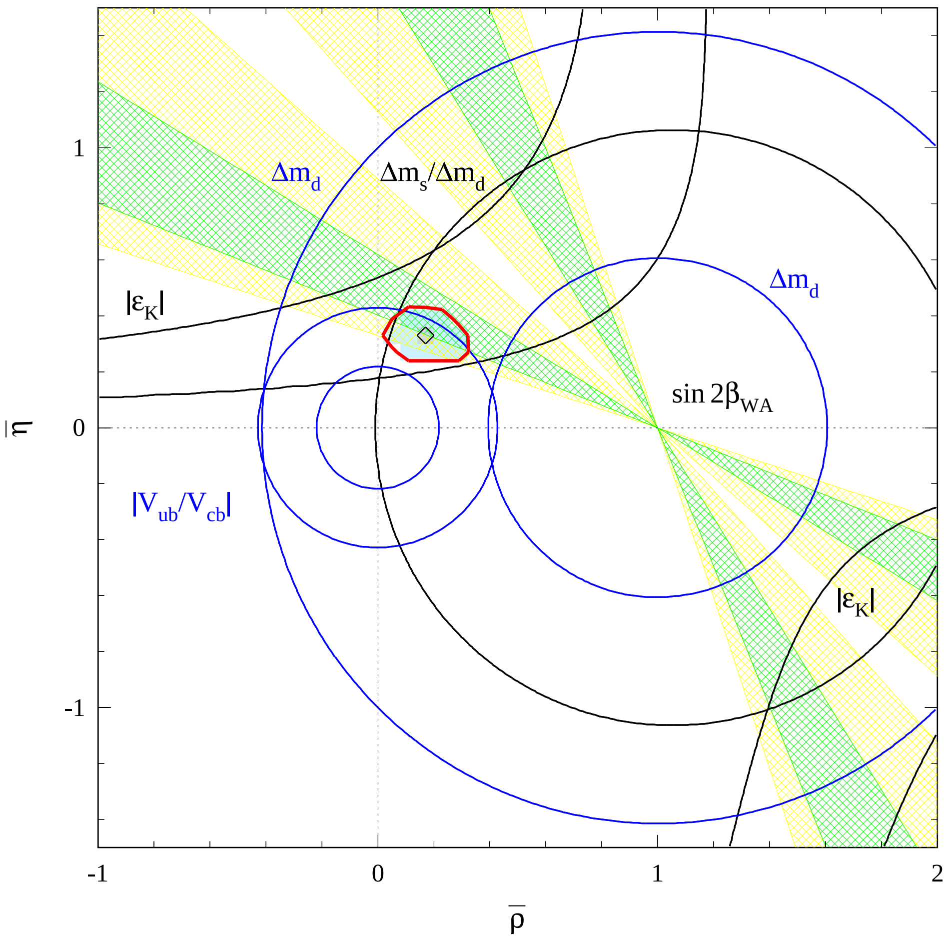

Figure 7.3: CKM triangle from global fit, using current values including this thesis's $|V_{ub}|$ value. The small diamond shape is the 1σ contour which is surrounded by the larger 2σ contour.

small level for the branching fraction since the efficiency of reconstructing signal events is dependent on the model, but mostly it is an important factor for the $|V_{ub}|$ extraction through the various γ_u's that are the integral of the form factors. In this study, we have only used the ISGW2 model, although we have described different models and QCD calculations in Chapter 2. We can readily say that the agreement between the partial branching fraction measurements and the ISGW2 model is not stellar, in view of the small probability of chi-square obtained for the $|V_{ub}|$ fit (1.5%).

In conclusion, we can say that the branching fraction measurements shown in this study are the necessary tools to get a precise value of $|V_{ub}|$. Furthermore, they will shed much needed light on the status of form factor calculations for heavy to light quark transitions.

APPENDIX A

TOOLS FOR NEUTRINO RECONSTRUCTION

The neutrino reconstruction technique assumes that the missing energy and momentum (see section 4.2) of an event arises solely from the elusive neutrino in the signal decay. This technique relies heavily on the hermeticity of the detector. For CLEO II, the coverage is 95% and 98% of 4π for the tracking and calorimetry volumes respectively. For CLEO II.V, the tracking coverage goes down to 93%, because of the SVX detector.

To obtain a respectable neutrino 4-momentum resolution, one must have good tracking performance. This means that for every particle we want one and only one representation, either a track or a calorimeter shower, and not both. These requirements are rather different than those for standard exclusive reconstruction analyses, for which one is interested in the best representation of a particular particle and does not necessarily care if there is more than one.

In this section we review various tools and studies that were made over the years to improve the neutrino 4-momentum resolution.

“Trackman” and “Splitoff”

As just mentioned, the tracking software of the drift chamber is not optimized for a neutrino reconstruction analysis, for example, it can make two distinct tracks out of hits coming from a single particle (ghost tracks). When the particle curls in the drift chamber, each semi-circle is found as a distinct track. In both cases we are interested in the best representation of the particle, and we want to reject the spurious tracks. Trackman was designed with that goal in mind. The main categories of problematic tracks that trackman tries to address are: ghosts, when a string of hits has more than one track associated with them, curlers, these are subdivided according to how many semi-circles are found, from two to 4, and finally decays in flight or scatters. Another criteria used to distinguish among various curler categories is whether the various tracks have a z component fit. Detailed documentation can be found in reference [43]. It is important to note that it was tuned using an earlier version of the tracking software which did not include KALMAN fitting (see section 3.2) for example. We found that the performance was not significantly different for the new fitter and for II.V, although one could imagine possible improvements from using the SVX detector.

The situation is similar in the case of electromagnetic showers from the calorimeter. Charged particles leave both track and shower signals in the detector. We want to use the more precise track information, so we need to discard the showers coming from the interaction of charged particles with the calorimeter. Hadrons can also interact via nuclear interactions with the Cesium Iodide crystal calorimeter. These can produce energetic particles that travel through the crystals and interact, making distant “splitoff” showers nearby. We need to eliminate all the showers associated with the charged particles, so that ideally only those showers associated with neutral particles would remain.

Neutral particles are primarily photons, but they can also be neutrons, K_L and neutrinos. The energy left in the calorimeter by neutrons and K_L is only a fraction of the actual energy of the particle, since the calorimeter does not have enough material for such hadrons to leave all of their energy. From a neutrino reconstruction analysis point of view the partial energy measurement from the neutrons and K_L is still valuable information. A software package, splitoff, was written to get rid of charged particles showers and keep neutral particle showers. It is based on a neural net which is trained to distinguish between splitoffs shower and photon shower geometries. Details about the splitoff package can be found in [37].

Good particles

We can evaluate our hermeticity performance through MC studies. To asses how well we do with tracking performance, we must know what particles we would ideally want. A utility was created to identify which of the charged particles we would want to reconstruct for each MC event. We call these good particles. There are two classes of mistakes: we either miss a particle, or we have an extra track.

The criteria for deciding on the good particles are that they need to be charged, stable and produced before the calorimeter. Within those criteria there remain particles that we are not really interested in. For example, the particle could be the daughter of a pion, of a muon or of a kaon, that decayed in flight. It could be a particle that resulted from a nuclear interaction between an initial particle and a piece of the detector. Finally, it could be an electron from a delta-ray. From a neutrino reconstruction point of view, the perfect situation for these processes is to keep the parent track for the missing momentum and energy calculation, but to keep the daughter track for track shower matching. The worse case scenario would be to keep both tracks, and then end up with an extra track, or to miss both particles. An intermediate situation would be to identify either one, so that at least the missing energy and momentum calculation is not too far off. A test of whether we have an inclusive list of all the good particles is to sum up the charge from each good particle and see if it adds up to zero.

Mistakes

Once we have decided on the set of good particles, we can determine for a particular event, if we indeed found those good particles, and only those. As mentioned earlier, we can have two classes of mistakes: 1) a missed particle, or 2) an extra track. Before discussing those two types of mistakes, it is interesting to note that the total number of mistakes per event for the CLEO II dataset is 1.42 ± 0.02 compared to 1.83 ± 0.04 for the II.V dataset, when trackman is used. Even after correcting for the acceptance of the detector ($|cos\theta| \leq 0.7$), and the effect of low momentum tracks ($|\vec{p}| \geq 0.25$ GeV/c), we find the number of mistakes per event to be 0.83 ± 0.01 (II) and 1.02 ± 0.02 (II.V). All the numbers shown in these sections come from studies that were performed on a sample of signal $B^0 \to \pi^- \ell^+ \nu$ MC.

We can also consider the net charge, the sum of all approved tracks, of the reconstructed events. We have found the following general correlations of mistakes with net charge. When the net charge of the event is the same sign as the signal lepton, then on average the mistake is a missed particle, (for example the signal π^+ going down the beam pipe). When the net charge is the opposite charge of the signal lepton, then on average the mistake is a missed particle, such as the soft pion from a D^{*+} from the other B decay. When extra tracks remain, they arise mainly from electrons from photon conversion, and also from tracks from a nuclear interaction in the detector.

We compare the types of mistakes between the II and the II.V datasets. We first notice an increase in the number of mistakes that are extra particles coming from photon conversions of $\sim 35\%$ in II.V; from nuclear interaction the increase in number of mistakes is $\sim 40\%$. The increase in number of mistakes from missed soft pions is $\sim 25\%$ in the II.V dataset. These are all due to the increased amount of material, and the reduced tracking coverage of the II.V configuration.

Nuclear Interactions

Extra tracks from the nuclear interactions of particles with the detector material is a prevalent mistake. We can attempt to improve this situation by implementing an algorithm that identifies daughters of a nuclear interactions in an event [1]. We could then do several things: 1) throw out the event on the basis that the missing

[1] It is at least easier than to try to recuperate missing soft pions. Information can be found on a software package designed to do stand-alone tracking in order to reconstruct low momentum particles in [44]

momentum is inherently unreliable, 2) flag the daughters as bad tracks that should not be included in the missing momentum computation 3) and attempt to identify the incoming track, or even incoming hits, left by the particle that caused the nuclear interaction.

Several situations might arise when we have a nuclear interaction related with whether we detect the parent particle, and/or the daughter particles of the nuclear interaction. Here is the break down of the various possibilities, as seen in signal MC:

• 17% of the time, we find no tracks for either the parent or the daughters. That means that we have a missed particle, but we can not do anything about it.

• 61% of the time, we find a track for the parent and none for the daughters, this is the best situation as far as neutrino reconstruction is concerned.

• 13% of the time, we find one or more daughters and no track for the parent. If we actually only detected one daughter, it will carry a fraction of the parent particle momentum, so that the missing momentum might not be too disturbed. But to be on the safe side, we would want to find that nuclear interaction to possibly fix up the missing momentum.

• 7% of the time, we find one track for the parent and one or more tracks for the daughters. In this case, we should discard the daughter tracks.

The cuts that give good results for identifying tracks that come from nuclear interactions are: (based on a similar study in [42])

• choose a track identified as a proton, but nos as an anti-proton. Anti-protons come primarily from Λ decays not nuclear interactions.

• the distance of closest approach to the beam should be less than 1 mm

• the number of drift chamber hits should be greater than 10

• the track should be a good track

• try to make a good vertex (using a kinematic fit) with another track (which fills the same track requirements as for the proton track, except for the particle identification and the charge conditions). The probability of χ^2 of the vertex should be greater than 0.1%

• make a global vertex with all the selected partners. The probability of χ^2 of that vertex should be greater than 0.5% and the vertex should be at least the radius of the beam pipe.

• if no partner was found for the proton, flag this track as a daughter of a nuclear interaction

• if the global vertex is not good, drop tracks until a good vertex is found

• if two nuclear interaction vertices contain a common track, take the one with the best probability.

Those cuts were optimized using Monte-Carlo. To define the efficiency and purity of the algorithm we used the tool described above, which identifies good tracks in a reconstructed MC event. Those cuts were also selected after various attempts were made to raise the efficiency of finding cases of nuclear interactions. For example, we tried allowing pions as seed tracks. To raise the purity, we tried vetoing the track if it was labeled as part of a good vertex. We also tried to take only vertices that were within a certain range of known material. Finally we also tried looking at the track's trackman code to see if that would help the decision, since Trackman attempts to identify tracks from nuclear interaction, but its efficiency and purity is very low.

The above cuts identify nuclear interactions with an efficiency of 23.5% and a purity of 71.2%. We can look at the resulting improvements in the different cases: when we found both the daughters and the parent tracks, just using trackman gives a 57% chance of doing the right thing for neutrino reconstruction, while, if we add the algorithm to the trackman decision, we get a 63% chance. Even though a 6 % improvement would not seem like much, if it does translate to a 6 % improvement in neutrino momentum resolution, it would be quite interesting. Probably more important would be the change in the fraction of events with no mistakes.

One disadvantage of the previous algorithm is that it is actually quite good at identifying Λ's, so that although this would probably be a small effect for this particular analysis, it could have a big effect for other analyses. The next step would be to look at the situation of the parent and decide on a global scheme to utilize this algorithm and study the improvement it made to the analysis. We leave this task to the next generation of neutrino reconstruction adepts.

REFERENCES

[1] G.Cohen-Tannoudji, M. Spiro, *La matière espace-temps*, Librairie Arthème Fayard, 1986

[2] Chris Quigg, *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, Frontiers in Physics v. 56, Addison Wesley

[3] David Griffiths, *Introduction to elementary particles*, John Wiley and Sons

[4] Francis Halzen, Alan D. Martin,*Quarks and Leptons: An Introductory Course in Modern Particle Physics*, John Wiley and Sons

[5] Yosef Nir, *CP Violation A New Era*, hep-ph/0109090

[6] Jonathan L. Rosner, *The Standard Model in 2001*, hep-ph/0108195

[7] *Review of Particle Physics*, Volume 15, number 1-4, 2000

[8] Brian K. Heltsley, *CKM status and Prospects*, CLNS 01/1765

[9] P.F. Harrison, H. R. Quinn, Editors *The BaBaR Physics Book*, SLAC-R-504

[10] Wei-Shu Hou, James G. Smith, Frank Wurthwein,*Determination of the Phase of Vub from Charmless Hadronic B Decay Rates*, hep-ex/9910014

[11] Nir Polonsky, *Supersymmetry: Structure and Phenomena*, Springer-Verlag Heidelberg, Vol. m68, (hep-ph/0108236)

[12] V. Boisvert, *The New Energy Program*, CBX 00-70

[13] Donald H. Perkins, *Introduction to High Energy physics*, Addison-Wesley Publishing.

[14] W.R. Leo, *Techniques for Nuclear and Particle Physics Experiments*, Springer Verlag.

[15] Y. Kubota, (CLEO collaboration) *The CLEO II Detector*, CLNS 91/1122

[16] Alexey Ershov, *Beauty Meson Decays to Charmonium*, Harvard thesis, May 2001

[17] V. Boisvert, *The Central Detector Recompress and II.V Monte Carlo tuning*, CBX 98-70

[18] A. Ryd, *Measurement of B meson semileptonic decay form factors*, UCSB thesis

[19] J. Richman, P.Burchat, *Leptonic and Semileptonic Decays of Charm and Bottom Hadrons*, UCSB-HEP-95-08

[20] M. Neubert, *Heavy Quark Symmetry*, SLAC-PUB-6263

[21] G.P.Lepage, *Lattice QCD for small computers*, TASI 93 lectures

[22] G.P. Lepage, *The analysis of algorithms for lattice field theory*, CLNS 89/971

[23] A. El-Khadra, A. Kronfeld, P. Mackenzie, S. Ryan, J. Simone, *Semileptonic decays $B \to \pi\ell\nu$ and $D \to \pi\ell\nu$ from lattice QCD*, Physical Review D Volume 64 014502

[24] UKQCD Collaboration, *Improved $B \to \pi\ell\nu$ form factors from the lattice*, Physics Letters B 486 (2000) 111-117

[25] UKQCD Collaboration, *Lattice-Constrained Parametrizations of Form Factors for Semileptonic and Rare Radiative B decats*, Physics Letters B 416 392 (1998)

[26] D. Becirevic, A. B. Kaidalov, *Comment on the heavy $\to$ light form factors*, Physics Letters B 478(2000) 417-423

[27] Laurent Lellouch, *Exclusive semileptonic B decays: lattice results and dispersive bounds*, hep-ph/9912353

[28] L. Lellouch, *Lattice-constrained unitarity bounds for $B \to \pi\ell\nu$ decays*, Nuclear Physics B 479 (1996) 353-391

[29] P. Colangelo, A. Khodjamirian, *QCD Sum Rules, A modern perspective*, hep-ph/001010175

[30] Patricia Ball, V. M. Braun, *Exclusive Semileptonic and Rare B meson decays in QCD*, Physical Review D, 58 094016 (1998)

[31] A. Khodjamirian, R. Ruckl, S. Weinzierl, C. W. Winhart, and O. Yakovlev, *Predictions on $B \to \pi\ell\nu$, $D \to \pi\ell\nu$ and $D \to K\ell\nu$ from QCD light-cone sum rules*, Physical Review D, 62 114002 (2000)

[32] N. Isgur, M. Wise,*Influence of the B^* resonance on the $B \to \pi\ell\nu$*,Phys. Rev. D 41, 151 (90)

[33] D. Scora, N. Isgur,*Semileptonic meson decays in the quark model: An update*, Phys. Rev. D, 52, 2783 (95)

[34] D. Melikhov, B. Stech, *Weak form factors for heavy meson decays: An update*, Phys. Rev. D, 62 014006 (2000)

[35] Th. Feldmann, P. Kroll, *Skewed parton distributions for $B \to \pi$ transitions*, Eur. Phys. J. C, 12, 99-108 (2000)

[36] CLEO collab., *First Measurement of the $B \to \pi\ell\nu$ and $B \to \rho(\omega)\ell\nu$ branching fractions*, CLNS 96-1419

[37] L. Gibbons et. al., *Exclusive $B \to \pi/\rho/\omega\ell\nu$ via neutrino reconstruction*, CBX 95-6.

[38] L. Gibbons et. al., *Exclusive $B \to \pi/\rho/\omega\ell\nu$ Update*, CBX 96-26

[39] CLEO coll., *Measurement of the $\mathcal{B}(B \to \rho\ell\nu), |V_{ub}|$, and the form factor slope in $B \to \rho\ell\nu$ decay*, CLEO CONF 98-18

[40] C. W. Bauer, Z. Ligeti and M. Luke, *A model independent determination of* $|V_{ub}|$, Phys. Lett. **B479**, 395 (2000) [hep-ph/0002161].

[41] K.A. Bloom,*Analysis of semileptonic decays of B mesons to D mesons*, Cornell thesis, 1997

[42] A. D. Foland, *Lenght scale in the CLEO II.V detector*, CBX 98-59

[43] S. Roberts, L. Gibbons, E. Thorndike, *Trkman the Next Generation*, CBX 96-103

[44] D. Chin, A. Magerkurth, M. Marsh, M. Palmer, J. Thaler, *BINGO: A set of Object Oriented Tools for Hierarchical Pattern Recognition and Track Fitting*, submitted to CHEP 2000

[45] Fulton R, et al (CLEO collaboration). *Observation of B-Meson Semi-leptonic Decays to Non-Charmed CLEO Final States* Phys. Rev. Lett. 64:16-20 (1990)

[46] T. Riehle, *Recompress Fake Rates*, CBX Draft, June 11 2000

[47] A. Lyon and E. Thorndike, *Study of ToF Performance*, CBX 00-37

[48] P. Hopman, J. Alexander, J.G. Thayer, *A Search for B+ -¿ tau+ nutau Using Full Event Reconstruction* , CBX 00-26

[49] Roger Barlow and Christine Beeston, *Fittin using finite Monte Carlo samples*, Computer Physics Communications 77(1993) 219-228

[50] F. De Fazio and M. Neubert, *$B \to X_u l \overline{\nu}_l$ decay distributions to order α_s*, JHEP **9906**, 017 (1999) [hep-ph/9905351].

[51] B. Berger, *Slow Track Embedding Studies I: Prelib Efficiency Results.* CBX 98-36

[52] L. Gibbons et al. (CLEO collaboration), *The Inclusive Decays $B \to DX$ and $B \to D^* X$*, Phys. Rev.D56:3783-3802 (1997)

[53] R.M. Baltrusaitis et al. (ARGUS III collaboration), *A Direct Measurement of charmed D^+ and D^0 Semileptonic Branching Rations.*,Phys.Rev.Lett.54:1976 (1985)

[54] A. Bornheim et al. (CLEO collaboration). *Improved Measurement of $|V_{ub}|$ with Inclusive Semileptonic B Decays*, CLNS 01-1767

[55] J. P. Alexander, et al (CLEO collaboration). *Measurement of the Relative Branching Fraction of $\Upsilon(4S)$ to charged and neutral B meson pairs*, CLNS 00/1670

[56] Plot graciously provided by Stefan Anderson.