A Search for Exotic Muon Decays
in the TWIST Muon Decay Spectrum

by

Ryan David Bayes
B.Sc., Simon Fraser University, 2003

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Physics and Astronomy

©Ryan David Bayes, 2005
University of Victoria

Supervisor: Dr. Art Olin

Co-Supervisor: Dr. Robert Kowalewski

ABSTRACT

The search for lepton flavour violation is significant to our understanding of the standard model of particle physics. This measurement uses the muon decay data collected by the TWIST experiment at TRIUMF to search for lepton flavour violation between charged lepton species. Specifically, I searched for the decay $\mu^+ \rightarrow e^+ X^0$, where X^0 is an unknown, undetected, neutral boson, against a background of standard muon decays, $\mu^+ \rightarrow e^+ \nu_e \bar{\nu}_\mu$. I set branching ratios for these decays such that $\mathrm{B}(\mu^+ \rightarrow e^+ X^0) < 2 \times 10^{-5} - 5 \times 10^{-5}$, for X^0 of various masses accessible by the momenta of the muon decay spectrum, to a confidence level of 95%. This improves upon the previous best limits by a factor of three. A correction to the spectrum was required to set a limit on a massless X^0 of $\mathrm{B}(\mu^+ \rightarrow e^+ X^0) < 6.3 \times 10^{-5}$.

Supervisor: Dr. A. Olin (University of Victoria and Science Division, TRIUMF)

Co-Supervisor: Dr. R. V. Kowalewski (Department of Physics and Astronomy)

Contents

List of Tables

List of Figures

Acknowledgements

I would like to extend my gratitude to my TWIST supervisor Art Olin. His input was indispensable for the preparation of this thesis and I benefited greatly from his experience. Similarly I would like to thank the TWIST collaboration for the opportunity to work with them during the exciting time surrounding the publication of their first results. I would specifically like to thank Glen Marshall, for making me feel welcome within the group, and Dick Mishke for the insights on the nature of science. I should also offer my thanks to my fellow graduate students, Blair Jameson, Rob MacDonald, Jim Musser, and Andrei Gaponenko, who developed the framework from which I developed my analysis long before I arrived with the collaboration.

Chapter 1

Introduction

The violation of family symmetries between leptons, also known as lepton flavour violation (LFV), has experienced a revival of interest in recent years due to the discovery of neutrino oscillations. These oscillations amount to a violation of lepton flavour between the neutral lepton species. A standing question within particle physics is whether there is a similar violation of lepton flavour between charged leptons. Here I will describe a search for exotic, lepton flavour violating, decays within the muon decay spectrum measured by the TWIST experiment.

TWIST, the TRIUMF Weak Interaction Symmetry Test, studies the shape of the spectrum from the muon decay $\mu^+ \rightarrow \nu_e \bar{\nu}_\mu e^+$, with the goal of testing the weak interaction to an unprecedented level of accuracy. This goal requires both a significant sample of muon decays, and a high precision apparatus to measure the momentum and angle of the positron produced by the decay. By virtue of this large statistical sample, we can search for LFV in muon decay at a level well below the current experimental limits.

The simplest observable case of this LFV is the case of a muon decaying into an electron (or positron) and a second, unknown, neutral particle. Thus far, such a decay has not been observed. However, the lack of observation of LFV decays in large samples of muon decays allows us to set upper limits on the probability of such decays occurring. TWIST provides an excellent sample of muon decays for such a search.

Many searches for LFV decays have been done to date, and there are many

currently under way. The current best limits to a 90% confidence level as recorded in [1] for the branching ratios of possible muonic LFV decays are,

$$
\frac{\Gamma(\mu^+ \to e^+\gamma)}{\Gamma(\mu^+ \to e^+\nu_e\bar{\nu}_\mu)} \quad < \quad 1.2 \times 10^{-11}
$$

as obtained by the MEGA collaboration

$$
\frac{\Gamma(\mu^+ \to e^+X^0, X^0 \to e^+e^-)}{\Gamma(\mu^+ \to e^+\nu_e\bar{\nu}_\mu)} \quad < \quad 1.0 \times 10^{-10}
$$

as obtained by the SINDRUM collaboration [2]

$$
\frac{\Gamma(\mu^+ \to e^+X^0)}{\Gamma(\mu^+ \to e^+\nu_e\bar{\nu}_\mu)} \quad < \quad 3 \times 10^{-4} \text{ reported by Bryman and Clifford [3]}
$$

$$(1.1)$$

TWIST is not equipped to measure $\mu \to e\gamma$ competitively as it cannot measure the outgoing photon. For this reason I will be focusing on the decay $\mu^+ \to e^+X^0$ where X^0 is an unknown boson which cannot be directly observed. Of the two branching ratios of this type of decay which are shown here, the result given by Bryman and Clifford in [3] is the most comparable to the measurement that I will be describing here. The SINDRUM result looked specifically for the prompt decay of the unknown boson into an electron positron pair. This makes the SINDRUM measurement a complementary measurement to those reviewed by Bryman, and the measurement that I will be describing in this thesis. Bryman's measurement, although the one most commonly quoted to set the limits on the presence of two body decays involving a massive boson in muon decay, was superceded by a later analysis of the his data. Balke [4] reported a branching ratio that improved on the Bryman result by "more than a factor of 2 ... across most of the spectrum". As a consequence I shall be considering both of these measurements in direct comparison to the measurement that I will be describing in this thesis.

Chapter 2

Theory

Within the standard model of particle physics the leptons are grouped into three generations, or families. The charged members of these families, the electron, muon, tau, and their anti-particles, do not change their flavours during interactions due to the vanishingly small masses of their neutrino partners [5]. However, the discovery of neutrino oscillations has provided conclusive evidence that neutrinos do have mass and, subsequently, renewed theoretical interest in the prospect of lepton flavour violation among charged species [6]. A primary target for this interest is in the simplest possible of LFV processes; that is $\mu \to e\gamma$. Alternate two body decay modes involving other flavour mediating particles such as familons, and even particles mediating the quintessence field have also been proposed as the mechanism for LFV processes and all of these require a global analysis of possible decays such as that which I am pursuing here.

2.1   Three Body Muon Decay

Before continuing with a discussion of LFV mechanisms within muon decay, it is important to discuss standard muon decay, via the three body decay $\mu^+ \to e^+ \bar{\nu}_\mu \nu_e$, as it is measured by TWIST. This charge changing interaction can be most generally described using the matrix element

$$M = 4\frac{G_F}{\sqrt{2}} \sum_{\substack{\gamma=S,V,T \\ \epsilon,\mu=R,L}} g^\gamma_{\epsilon,\mu} \langle e_\epsilon|\Gamma^\gamma|(\nu_\epsilon)_n\rangle\langle(\nu_\mu)_m|\Gamma_\gamma|\mu_\mu\rangle. \tag{2.1}$$

3
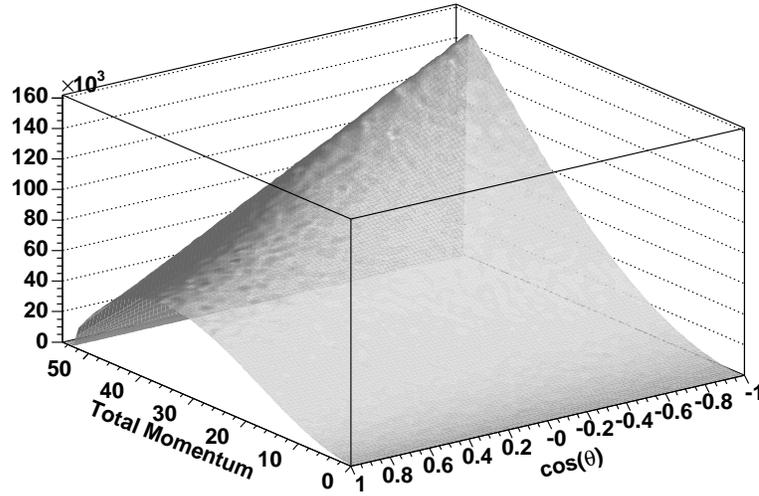
Figure 2.1: A Monte Carlo simulation of the Michel muon decay distribution

Here, the index γ labels the interaction as a 4-vector (V), 4-scalar (S), or 4-tensor (T) interaction, while the indices ϵ and μ denote the chirality of the leptons involved in the interaction. I have also used G_F to represent the Fermi Coupling constant while $g^\gamma_{\epsilon,\mu}$ represents the coupling between a μ handed muon and a ϵ handed positron via the interaction Γ^γ. The chirality of the neutrinos, n and m are uniquely determined from μ, ϵ, and γ. According to the standard model of electroweak interactions, the only interaction allowed is the left handed, vector coupling ($g^\gamma_{\epsilon,\mu} \equiv 1$). All other possible couplings are zero in the standard model [7]. A simulation of the Michel spectrum appears in Fig. 2.1.

The differential decay distribution resulting from this interaction is generally represented in terms of the energy and angle of the outgoing positron.

$$\frac{\partial^2 \Gamma}{\partial x \partial \cos\theta} = \frac{m_\mu}{4\pi^3} W_{e\mu}^4 G_F^2 \sqrt{x^2 - x_0^2} \left(F_{IS}(x) + P_\mu \cos\theta F_{AS}(x) \right) \qquad (2.2)$$

where the positron polarisation has been integrated out of the equation and radiative corrections have been neglected. Here I have used $W_{e\mu} = (m_\mu^2 + m_e^2)/2m_\mu$ to represent the maximum energy attainable by the positron, the reduced energy of the positron $x = E_e/W_{e\mu}$, the reduced mass of the positron $x_0 = m_e/W_{E\mu}$, and the degree of muon polarisation within the measured sample P_μ [8].

If I expand the isotropic and anisotropic parts of the spectrum (F_{IS} and F_{AS} respectively) into functions of the reduced energy,

$$
\begin{aligned}
F_{IS} &= x(1-x) + \frac{2}{9}\rho(4x^2 - 3x - x_0^2) + \eta x_0(1-x) \\
F_{AS} &= \frac{1}{3}\xi\sqrt{x^2 - x_0^2}\left[1 - x + \frac{2}{3}\delta\left(4x - 3 + (\sqrt{1-x_0} - 1)\right)\right].
\end{aligned}
\tag{2.3}
$$

Here the masses of the neutrinos have been neglected as have higher order corrections (although they are included in the simulation). The parameters, ρ, η, δ, and ξ, are referred to as the Michel parameters. Throughout the remainder of this thesis the decay spectrum corresponding to equations 2.2 and 2.3 will be referred to as the Michel spectrum. The measurement of the Michel parameters act as a very sensitive test of the standard model. Any deviations of the Michel parameters from the standard model values of $\rho = \delta = 0.75$, $\xi = 1$, or $\eta = 0$ will imply the existence of physics outside of the standard model [9].

2.2 Two Body Decays

2.2.1 Possible Models

There are a number of proposed mechanisms for the decay of a muon into a positron with an unknown, neutral particle. The common thread within these candidate fields is that their interactions with matter involve derivative couplings. However, the physical motivations for these models are very different.

The first of these candidate particles is the axion. Axions are Nambu-Goldstone bosons resulting from the non-zero vacuum energy of the gauge field that maintains the CP symmetry within strong nuclear interactions [10]. Such a boson would behave as a $U(1)_A$ gauge field (a field mediating axial rotations) with an effective Lagrangian [11]

$$
\mathcal{L}_{eff} = \frac{1}{F_{\mu e}}\bar{\mu}\gamma_\rho e\partial_\rho X_{\mu e}^0,
\tag{2.4}
$$

where $F_{\mu e}$ is the energy scale on which the lepton flavour symmetry is broken. We can use the branching ratio of the $\mu \to e^+ X^0$ to infer this energy scale,

$$
\frac{\Gamma(\mu^+ \to e^+ X^0)}{\Gamma(\mu^+ \to e^+ \bar{\nu}_\mu \nu_e)} = \frac{2.5 \times 10^{14} GeV^2}{F_{\mu e}^2}.
\tag{2.5}
$$

This scale should be observable if $F_{\mu e} \lesssim 10^{12}$ GeV. The value of $F_{\mu e}$ is currently constrained to be greater than 9.9×10^9 GeV due to previously defined upper limits [12]. Improvements on this measurement will drive this lower limit higher.

Another physically interesting example cites the dynamical component of the dark energy field Q, also known as the quintessence, as a possible source of LFV decays [13]. One of the consequences of the quintessence interacting with normal matter is that it will create a time dependant variation in the physical constants. Such a variation can be suppressed by introducing a global shift symmetry into the quintessence field, $Q \rightarrow Q + C$ [14]. This forces the Q field to behave like a pseudo-Goldstone boson and the interaction of the Q field to involve derivative couplings. The resulting effective Lagrangian is again very similar to the axion interaction,

$$\mathcal{L}_{eff} = \frac{1}{\Lambda} \partial_\mu Q (c_{\mu e}^R \bar{f}_{eR} \gamma^\mu f_{\mu R} + c_{\mu e}^L \bar{f}_{eL} \gamma^\mu f_{\mu L}) \tag{2.6}$$

where Λ represents the cutoff energy scale for the interaction. Λ can be constrained using the current limits on the decay $\mathcal{B}(\mu \rightarrow eQ)$. If the coupling constants are unitary, that is $c_{e\mu}^R = c_{e\mu}^L = 1$, then one can produce a branching ratio

$$\frac{\Gamma(\mu^+ \rightarrow e^+ Q)}{\Gamma(\mu^+ \rightarrow e^+ \bar{\nu}_\mu \nu_e)} = \frac{3\pi^2}{\Lambda^2} \frac{1}{(m_\mu G_F)^2}. \tag{2.7}$$

The current limits on the branching ratio give us a lower limit for the cutoff energy scale of $\Lambda \gtrsim 4 \times 10^9$ GeV.

A similar effective Lagrangian occurs in the more general case of other massless spin one gauge bosons other than the photon [15]. If one argues that there is a symmetry for which all standard model fields are singlets a consequence of the interaction of such a field with matter is a derivative coupling. This implies that a singlet gauge field will, in general, produce a coupling that will violate the conservation of Lepton flavour. In this light it is apparent that a general two body search for LFV decays is significant in the search for new physics.

2.2.2   Signal of Two Body Decays

In spite of the wide range of available models which may predict two body decays, the signal indicating their presence is very simple, clean and model independent. The signal resulting from the presence of a two body decay is dictated kinematically

by the mass of the unknown neutral boson emitted from the decay and the spin properties of the boson. Using the conservation of four momentum we know that we will be able to see a resonance peak for a given boson mass, m_X, at a positron momentum

$$p_e = \sqrt{\left(\frac{m_\mu^2 + m_e^2 - m_X^2}{2m_\mu}\right)^2 - m_e^2}. \tag{2.8}$$

An important feature of this equation is that the signal generated by the decay of the muon into a massless (or nearly massless) boson is a peak at the endpoint of the Michel spectrum. This will become important later in the analysis.

If gauge invariance holds for a given field, then the associated gauge bosons should be massless. The masses of the $W^\pm$ and the Z bosons are due to the spontaneous symmetry breaking introduced by the Higgs mechanism. Thus we should expect that the candidate bosons mediating LFV decays should be massless, if their associated symmetries are unbroken. In the case of the axion this is believed to be correct. Similarly, data obtained from the WMAP satellite constrains the mass of Q to be less than $10^{-33}eV$. Jodidio *et. al.* [12] set a 90% upper limit on the branching ratio for a massless boson at

$$\frac{\Gamma(\mu^+ \to e^+ X^0)}{\Gamma(\mu^+ \to e^+ \nu_e \bar{\nu}_\mu)} < 2.6 \times 10^{-6}. \tag{2.9}$$

The data used by Jodidio to produce this result is more sensitive to the endpoint region compared with the data presented here.

The width of this peak is dictated by the life time of the neutral daughter particle. For example, I will consider the width of the resonance if this daughter particle decays while it is inside the detector. If one considers a detector with dimensions on the order of meters then one can reasonably expect such a decay to take place within 10^{-8} seconds after the initial muon decay if it is to take place before the relativistic boson escapes the detector. The width, Γ, produced by such a decay would be

$$\Gamma \sim \frac{\hbar}{\tau} \sim 6.6 \times 10^{-16}\text{eV} \tag{2.10}$$

which is much smaller than any resolution that we can reasonably expect for our detector. On the other hand, given that the resolution (standard deviation) of the response is roughly between 50 keV/c and 70 keV/c, I can set an upper limit on the
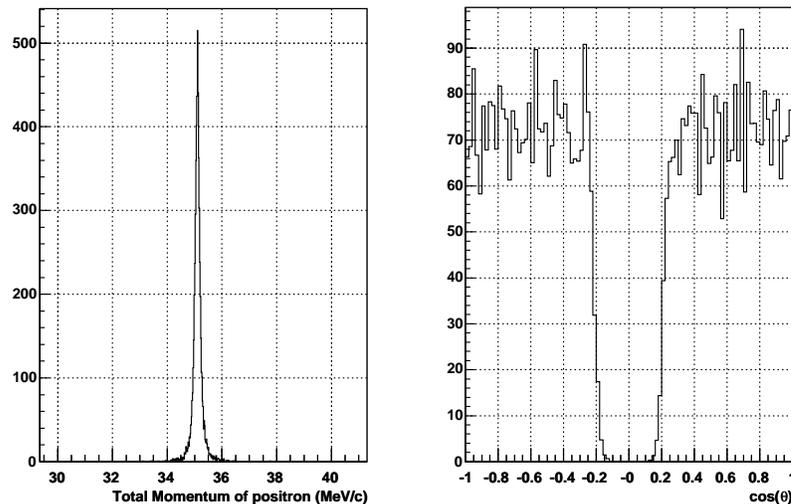
Figure 2.2: A possible signal for a two body decay as it may appear within the TWIST spectrometer. The histogram on the left shows the expected momentum projection of a signal while the histogram on the right shows the projection of the signal in $\cos\theta$. Note that tracks with $|cos\theta| < 0.2$ are not reconstructed by the TWIST spectrometer.

lifetime of a particle resonance dominated by the detector response;

$$\tau \gtrsim \frac{\hbar}{\Gamma} \sim 9 \times 10^{-21} \text{s}. \tag{2.11}$$

This is smaller than the typical timescale for a weak interaction. The resolution scale of the TWIST detector is much larger than the resonance width for a wide range of decay lifetimes. This motivates the use of the detector response to model the shape of the peak. Such a peak may look like Fig. 2.2. This peak is shown superimposed on the Michel spectrum in Fig. 2.3.

Finally I should point out that there is no inherent angular dependence to the signal as it has been described so far. Any angular dependence that may appear will be due to spin interactions. The most general representation of such an anisotropic behaviour is

$$\mathcal{A} = 1 + \mathcal{P} \cos\theta \tag{2.12}$$

where $\mathcal{P}$ represents the degree of anisotropy of the distribution. This anisotropy can
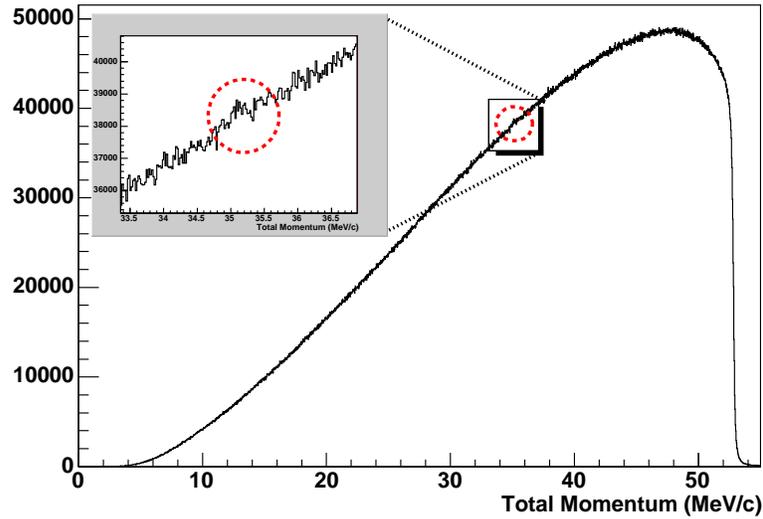
Figure 2.3: The peak shown in Fig. 2.2 added to a standard Monte Carlo simulation of the Michel distribution. This signal has a measurable branching ratio of 1.4×10^{-4}. Note that tracks with $|cos\theta| < 0.2$ are not reconstructed by the TWIST spectrometer.

come from any number of model dependant effects. For example, if some kind of parity violation occurs in the interaction, then one might see some kind of anisotropic signal. In this case we would expect that $\mathcal{P}$ will take on a value related to the total angular momenta of the muon and positron as well as the spin of the boson. Again one would expect the resolution characteristics to dominate the resolution width in momentum producing a signal such as that shown in Fig. 2.4. As TWIST does not measure the polarisation of the positron, there is too little information to make such an observation and draw any conclusions from them. Thus, while it is interesting to set a limit on the branching ratio for an anisotropic two body decay, I can say very little about what such a result means.
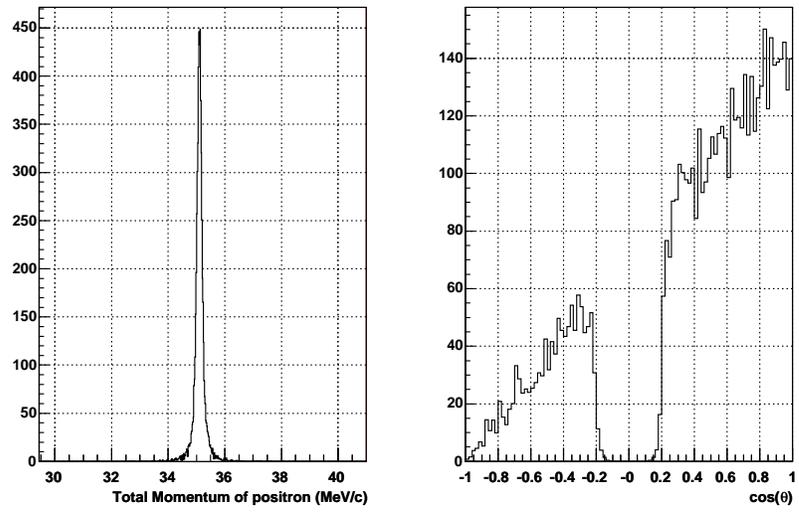
Figure 2.4: Projections in momentum and $\cos\theta$ of a possible signal for a two body muon decay showing positive anisotropy, as the signal will appear in the TWIST detector

Chapter 3

Experimental Setup

Here I give a brief description of the TWIST apparatus. The goal of TWIST is to measure the Michel parameters to an ultimate precision of parts in 10^4. To achieve the precision goals of TWIST, a precision apparatus is required. There are a number of examples where this precision has a direct impact on the experiment. Measuring the anisotropic parameters requires a highly polarised muon beam to achieve the precision goals. Thus the M13 beam-line is designed to transport muons, which have an initial helicity approaching -1, with very little depolarisation. The TWIST spectrometer is similarly designed with these goals of precision in mind. Many of these features highly beneficial to a search for unknown, exotic muon decays. I shall focus on the features of the experimental apparatus that are the most important for this measurement. A more detailed description of the TWIST spectrometer appears in [16].

3.1   Muon Generation and Transport

The M13 beam line transports highly polarised muons from a production target into the TWIST spectrometer. The muons are generated in a graphite or beryllium production target in the TRIUMF proton beam line 1A. As the high energy protons hit the target, they generate pions that quickly decay into muons. The pions that decay on the surface of the target then enter the beam line at 100% negative helicity due to the maximal parity violation of the decay $\pi^+ \rightarrow \mu^+ \nu_\mu$, within the standard
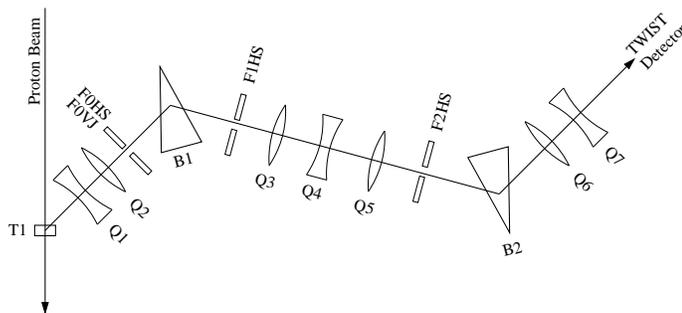
Figure 3.1: A schematic of the M13 secondary beam line

model. This decay generates mono energetic muons at 29.6 MeV/c. The Beam line is tuned specifically to carry muons at this momenta. Pions that do not decay on the surface of the production target tend to scatter, increasing the average helicity of the muons and causing them to lose a significant fraction of their momentum. These muons are generally removed by the momentum selection of the muon channel. Some pions escape the production target to decay within the channel. These muons, also known as cloud muons, provide a source of contamination for a muon beam with a helicity of -1, that can be easily removed from the data set by making a cut on the time between the time when the pions are created and the arrival of the muons in the detector. A source of the beam contamination that is not so easily removed come in the form of positrons resulting from the decay of muons during beam transport. Beam positrons can quite potentially yield a false signal for an exotic decay search. However, steps are taken during the analysis of the data to remove this contamination, making beam positrons an issue to be aware of, but not one of great concern for this analysis.

The beam line is depicted in Fig. 3.1. It consists of two bending magnets interspersed with a number of slits and jaws required for the selection of momenta, and three sets of quadrupoles to focus the beam at a spot within the magnetic field of the detector. This ensures that the initial helicity of the muons is maintained.

Although a highly polarised beam is not as important for an exotic decay search as for the measurement of the Michel parameters, it does provide the benefit of improving one's ability to resolve a decay. If the decay is isotropic, then the decay should be quite more easily resolved in the downstream end of the TWIST detector

than the upstream end due to the high negative polarisation. Similarly, if there is some positive anisotropy to the signal, we should have an even better resolution of the signal. Evidence of this effect will show as a decrease in the error of a signal with positive or zero anisotropy when compared to a signal with negative anisotropy. I will address this point in more detail in Appendix A.

3.2   The TWIST Spectrometer

The primary detector of TWIST is a high precision spectrometer. This spectrometer consists of 44 planar drift chambers (DCs) with 12 multi wire proportional chambers (PCs) contained within a 2 Tesla super-conducting solenoid. A muon scintillator at the far upstream end of the detector provides a low mass trigger for the muon decay events. The chambers are mounted in the detector cradle such that all the chambers are perpendicular to the beam direction (the z direction). Each chamber is separated from its neighbours by a set of precisely machined Sitall spacers. The variation in the thicknesses for these spacers was $\sim 3\mu$m from the design thickness. The Sitalls are arranged in the stack in sets of four. For a given set the variation in the thickness of the Sitalls within that set is less than 0.5 μm. The stack is placed under compression so the z position of a given wire plane is known to high accuracy. Each of the chambers, when inserted into the spectrometer stack, is either oriented +45(U) or -45(V) degrees from vertical when viewed from the upstream end of the detector, and mounted in UV pairs to provide a Cartesian coordinate system. The chambers are then mounted in modules of two or four planar chambers all arranged symmetrically about a central stopping target. As the wire chambers provide the bulk of the information for a given decay event, I will continue by discussing the wire chambers in a little more detail.

3.2.1   Proportional Chambers

The proportional chambers are used to produce the timing information for a given event. For this purpose PCs appear in two distinct sets of modules within the spectrometer. At the far upstream and downstream ends of the detector a module of two UV pairs of proportional chambers to provide a time for the muon entering
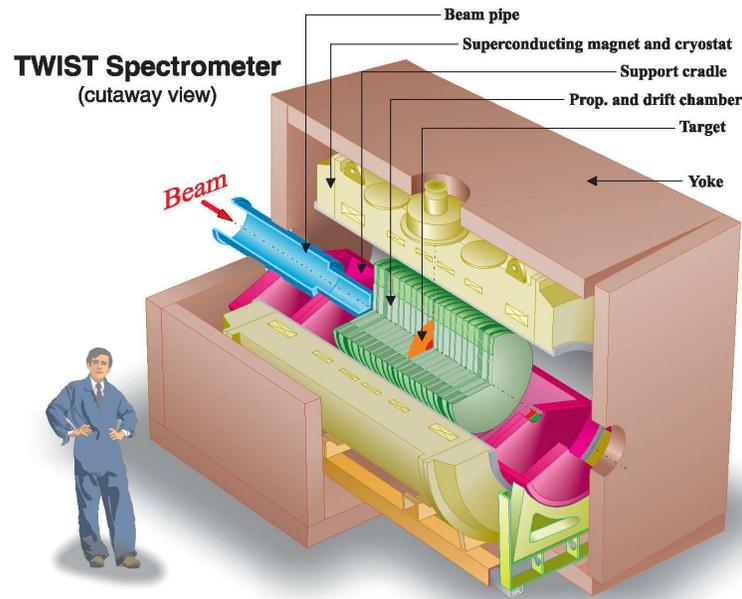
Figure 3.2: A conceptual view of the TWIST spectrometer

the tracking volume (or time of the positron exiting the tracking volume). In the centre of the tracking volume there is another module of four PCs about the target to similarly measure the time when the muon enters the stopping target and when the positron leaves. This information is primarily used the the event classification during the analysis.

Each chamber uses a two 6.35 μm sheets of doubly aluminised mylar foil as a cathode plane and a set of 15 μm tungsten sense wires with a pitch of 2.0 mm as an anode plane. A fast gas (CF_4/isobutane) was used in these chambers to ensure accurate timing. The distance between cathode planes is 4.0 mm. The two chambers at the centre of the target module use the stopping target as the central cathode. During the 2002 running the target was made out of 125 μm mylar sprayed with conducting graphite. This target has since been replaced with a high purity aluminium stopping target due to significant depolarisation within the mylar target. As a result, the average polarisation of the muons at the time of decay, and the best polarisation that we might expect for an anisotropic decay, during the 2002 run period is 0.89.
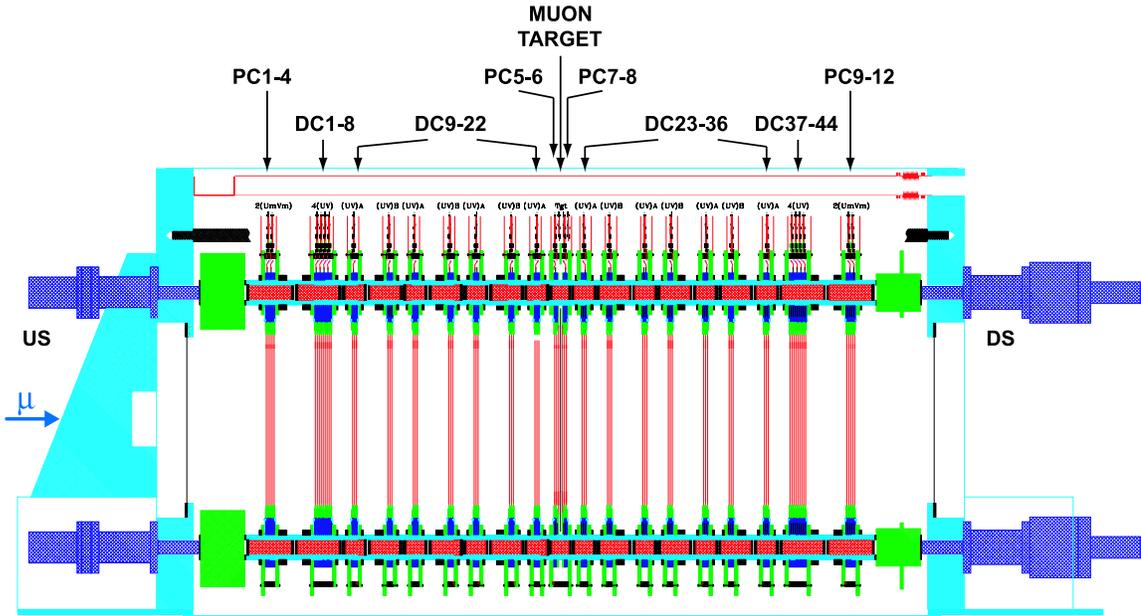
Figure 3.3: A side view of the TWIST detector stack

3.2.2 Drift Chambers

The drift chambers are primarily responsible for the tracking of particles through the detector. These chambers are very much like the proportional chambers in their construction with a few exceptions. The main difference is that the sense wires are strung with a 4 mm pitch to increase the drift distances. A "slower" gas (Dimethyl ether or DME) was chosen to be used in the DCs due to its low Lorentz angle (smaller drift due to the applied magnetic field) and good spatial resolution. The DCs are arranged in seven UV modules to each side of the target module and an eight plane module, called the dense stack, at either end of the tracking volume (where the planes are oriented as VUVUUVUV). The arrangement of the modules in the cradle was meant to eliminate periodicities in the detector. This effort was not completely successful. I will return to this issue in section 4.1.2.

The spatial resolution of the drift chambers is the limiting factor for the resolution of the momentum. Two primary factors limit this resolution; the mechanical precision of the detector and the properties of the drift cell. The mechanical precision issues include the alignment of the wire planes and quality control during the

construction of the planes. The drift cell properties are dictated by the ionisation density and drift velocity of electrons and ions within the gas used in the drift cell. These combine to produce a resolution less than 50 μm for tracks greater than 1mm away from the wire. For track closer than 1mm the resolution increases dramatically as shown in Fig. 3.4 due to the ionisation clustering.
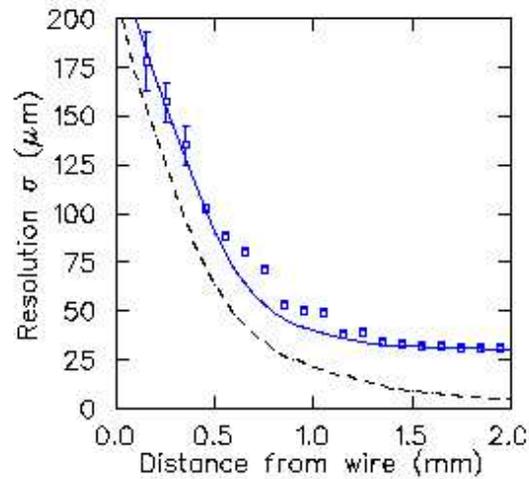
Figure 3.4: Points designate the measured resolution as a function of track distance from the wire. The dashed curve is the resolution from the GARFIELD calculation. The solid line shows includes the quadratic addition of a 30 μm multiple scattering contribution as well as a 1.5 ns time resolution contribution.

Chapter 4

Methods for the Peak Search

4.1 Preliminary Analysis

The analysis that I describe here is ancillary to the primary TWIST analysis. Much of the processing of the raw data required by TWIST for the purpose of measuring the Michel parameters can be used in a search for rare decays with a minimal amount of reprocessing. However, several issues in the processing of the raw data require some attention here as they affect the final analysis of the peak search.

4.1.1 Helix Fitting

After a particle passes through the drift volume, the trigger times (from the upstream scintillator and proportional chambers), and drift times are recorded. It is the job of the helix fitting to first translate these times into positions utilising the GARFIELD software package to set a minimum distance between the wire and the trajectory of a particle. The helix fitter then calculates the trajectories of the positrons given these hit positions to produce the positron's momentum and angle. The equations of motion for a moving particle in a constant uniform magnetic field are

$$
\begin{aligned}
x &= x_0 + \frac{p_t c}{eH} \cos(\frac{eH}{\mathcal{E}} s + \varphi) \\
y &= y_0 + \frac{p_t c}{eH} \sin(\frac{eH}{\mathcal{E}} s + \varphi) \\
z &= z_0 + \frac{p_l}{\mathcal{E}} s.
\end{aligned}
\tag{4.1}
$$

Here I have used p_t, the transverse momentum of a particle and p_l the longitudinal momentum of the particle. I have written this equation is in terms of s, the path length that the particle traverses. As the particle that we are primarily interested in, the positron, is moving very near to the speed of light, s=ct. The values x_0, y_0, z_0, and φ are constants within the fit describing an initial position for the helix. The remaining variables H, e, c, and $\mathcal{E}$, are the magnetic field, positron charge, speed of light and the energy of the positron respectively.

From this we can see that the radius of the helical track gives us the transverse momentum while the "wave number" of the helix gives us the longitudinal momentum of the positron. The value of $\cos\theta$ can be found from the ratio of the longitudinal momentum and the total momentum (p_{tot}). This is a simple proposition theoretically. Complications such as multiple scattering and inhomogeneities of the magnetic field make the problem much less tractable in reality. However, the corrections required by these complications will not be discussed here.

A cut that occurs at this stage of the analysis is a cut on the time between the muon trigger and the appearance of the positron. This time cut is made to ensure the quality of the fit of the track; the tracks made by positrons that result from fast decays can be spoilt by ionisation remaining near the wires due to the passage of the muons. This cut amounts to cutting out all decays that occur less than 700 ns after the muon stops and results in a loss of 36% of the initial statistics.

An important function that is also completed at this stage of the analysis is the event classification. The time information from the PCs is used for this purpose, separating events into one of 31 mutually exclusive event types. These classifications are made considering such things as the time between the muon stop and the subsequent decay into a positron, the number of muons that appear in the event, and whether there exists unwanted background such as beam positrons. Only those events which can be reconstructed with a high efficiency, such as events with a muon and decay positron which are well distinguished in time, are included in the data used for the final analysis. The event types which are kept by the analysis are shown in table 4.1. From the table, one can see that 40% of the decay tracks that occur 700 ns after the muon stop are kept by this cut on the event type. The events cut from the analysis at this stage are typically events where the muon and decay positron

Event Type	Description	Fraction of Events for t > 700 ns
1. Simple, Clean	Muon and decay positron separated by 1050 ns	0.303
2. Time Clean	Muon, decay positron, and beam positron(s), separated by 1050ns	0.0855
6. Simple Delta Cleaned	Muon, decay positron separated by 1050 ns. Delta ray from positron measured	0.0060
7. Time Delta Cleaned	Muon, decay positron, and beam positron(s), separated by 1050ns. Delta ray from positron measured	0.0017
10. Simple Scatter Cleaned	Muon and decay positron separated by 1050 ns. Positron hits on both sides of muon stop	0.00073
11. Time Scatter Cleaned	Muon, decay positron, and beam positron(s), separated by 1050ns. Positron hits on both sides of target.	0.00021
21. Simple Beam Positron Cleaned	Muon and decay positron separated by 1050 ns. Decay positron appears to be within 200 ns of a beam positron.	0.0031
22. Time Beam Positron Cleaned	Muon, decay positron, and beam positron(s) separated by 1050 ns. Decay positron appears to be within 200 ns of a beam positron	0.00089
Total		0.4021

Table 4.1: Fraction of events measured in data for the event types kept by the analysis. A brief description of the event is included. Note that the only particles appearing in the "Simple" events are the particles listed; no other tracks appear in the event.

are not separated by enough time for the respective tracks to be resolved properly (Hence the presence of the implicit cut of decay times less than 1.05 μs) or events where there is no muon and the event was triggered by a beam positron. A far more complete description of the event types and the event classification procedure in [17]. Further cuts will be discussed in Sec. 4.1.2, and summarised in table 4.2.

4.1.2   Important General Results

A few consequences of the reconstruction process are of key significance to the search for two body decays. Although it does not become an issue for the standard analysis of the Michel parameters due to usage of the Monte Carlo simulation in the spectrum fitting (see Sec. 4.2.1), the reconstruction efficiency will become important in the measurement of the two body decay branching ratio. Similarly, the momentum resolution after reconstruction is of central importance to the search for two body decays. The measurement of the resolution as well as some possible reasons for the result of the measurement will be described. Finally the choice of the fiducial region will be discussed in detail.

Acceptance and Reconstruction

During the reconstruction we can expect some events to be lost or misplaced due to inefficiencies in the helix fitter. At the same time it is important to quantify the total number of events lost due to the various cuts on the data made during the analysis. A simple and reasonably easy way to check this is to run a simulation of the experiment and see what happens to the events. Comparing the reconstructed surface muon decay spectrum, after all cuts on the data (which are summarised in Table 4.2), with the surface muon decay spectrum initially created by the Monte Carlo (after a cut on the decay time) gives us a direct measure of the events lost for any given bin in p and $\cos\theta$. The result of the ratio of these spectra is shown in Fig. 4.1. Also shown in this figure is the fiducial volume for this analysis. The effect of the fiducial cut on the statistics is also included in Table 4.2. One of the primary motivations for this fiducial was consideration of region of high reconstruction efficiency in this ratio. The fiducial region will be described in more detail later in this section.

Cut	Ratio of Tracks left by cut in data	Ratio of Tracks left by cut in Monte Carlo
Muon decay time: set at 0.7 μs	0.7264	0.7264
TCAP: cut to remove cloud muons from data set	0.6495	1.0000
Event type: see Table 4.1	0.4021	0.4435
Stopping z: Cut on the muon stopping position	0.7510	0.8838
Muon radius cut at 2.5 cm	0.9672	0.9601
Decay window time: Cut at 10 μs after event start	0.9897	0.9608
Number of tracks in window only one event in window	0.9977	0.9978
Reconstruction error made by the helix fitter	0.9564	0.9548
Cut ensuring events start and stop in the same half of the detector	0.9947	0.9977
Fiducial	0.3492	0.3464
Total fraction of events remaining after cuts	0.0622	0.0863

Table 4.2: A list of cuts made on the data and Monte Carlo simulation. The ratio of candidate decay tracks shown here is the number of tracks remaining after the cut divided by the number of tracks just before the cuts. Note that the candidate tracks that are removed by the TCAP cut are not generated in the simulation; thus the cut does not affect the Monte Carlo tracks

This figure shows the weaknesses and strengths of the TWIST detector reconstruction coupled with the effects of cuts on the data intended to ensure the quality of the data. The losses due to cuts on the event type is the primary cause of the loss of statistics shown in figure 4.1. The cuts on the data leave about 78 percent of the statistics seen in the data generated by the simulation inside the fiducial region. The efficiency of the reconstruction itself is very high. However the reconstruction also generates the fine scale structures seen in figure 4.1. The bands showing a higher reconstruction efficiency compared to the surrounding region are of particular interest. These bands follow lines of constant longitudinal momentum and are typically generated by periodicities within the detector itself. These periodicities produce helices with significant errors. Areas of the phase space are obviously inaccessible at extremely low angles, due to the inability to resolve the transverse component of the momentum at these angles. Furthermore we can see that the number of tracks
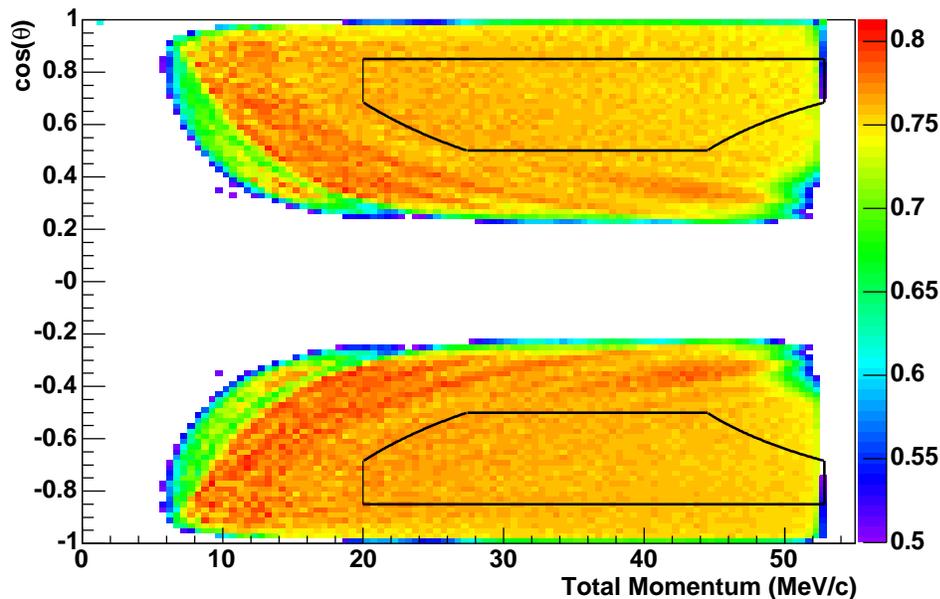
Figure 4.1: Ratio of the number of reconstructed events after the cut on event classification divided by the events accepted by timing cuts before reconstruction for any given cell in the phase space. The black lines describe the fiducial volume that is used for this analysis.

successfully fit drops off significantly at longitudinal momenta below 6 MeV/c. Finally, a sharp drop at $|\cos\theta| < 0.2$ occurs because multiple scattering at these angles impedes the reconstruction of the tracks.

Some of the cuts in Table 4.2 possess a momentum or angle dependant component. In particular high angle events will be removed from the reconstructed spectrum by the event type cut. This makes the summary of cuts appear far more aggressive than Fig. 4.1 would suggest. The existence of multiple candidate decay tracks within a decay window will also increases the number of tracks measured over all, increasing the statistics used in this description of the cuts. Most of these extra candidates are also removed by the event type cut, but subsequent cuts exist to ensure that only one decay track is contained within an event window. Specific examples of such cuts are the cut on reconstruction errors and the cut on the number of decay tracks contained within an event window. A definitive list of the cuts shown in Table 4.2 appears in [18].

A cut that is different between the data and the simulation is the TCAP cut. This is the cut that removes the cloud muons from the spectrum using the time structure of the muon beam. However, the simulated spectrum was generated to match the surface muon spectrum. Therefore no cloud muons were generated. The absence of cloud muons from the simulation should not effect the measurement of exotic decays, however. The change in the polarisation between the surface muons and the cloud muons should not change the proportion of exotic muon decays.

Momentum Resolution

Comparing the generated and measured momenta on an event by event basis, generates a measure of the momentum resolution for all points in the muon decay distribution. Since we know the momentum of each Monte Carlo event before and after reconstruction, we can take the difference of these two values to generate a distribution of deviations of the momentum from the true value. This distribution possesses asymmetric tails. However the tails of this distribution are very small, allowing the distribution to be approximated by a Gaussian. Measuring the Gaussian width for each point in the phase space produces the momentum resolution field shown in Fig. 4.2.

We can see a number of features here that parallel the features seen in the reconstruction efficiency. The sudden increase in the resolution width at high angles parallels the low percentage of reconstructed events appearing at these angles. A particularly interesting feature is that the resolution seems to vary as $|1/\sin\theta|$ for $|\cos\theta|$ between 0.50 and 0.85. We can understand this dependence quite easily if we consider the primary contribution to the resolution to be the finite spatial resolution within the drift cells. If each drift cell measures the positron track to be a distance x away from the wire centre, while the transverse momentum is proportional to the radius of the helix, then the error of this radius is inversely proportional to the square root of the number of hits times the error in the measurement of the track position. Thus,

$$\delta p_t \sim \delta r \propto \frac{\delta x}{\sqrt{n}}. \tag{4.2}$$

Meanwhile we must recall that, the transverse momentum is the sine of the total
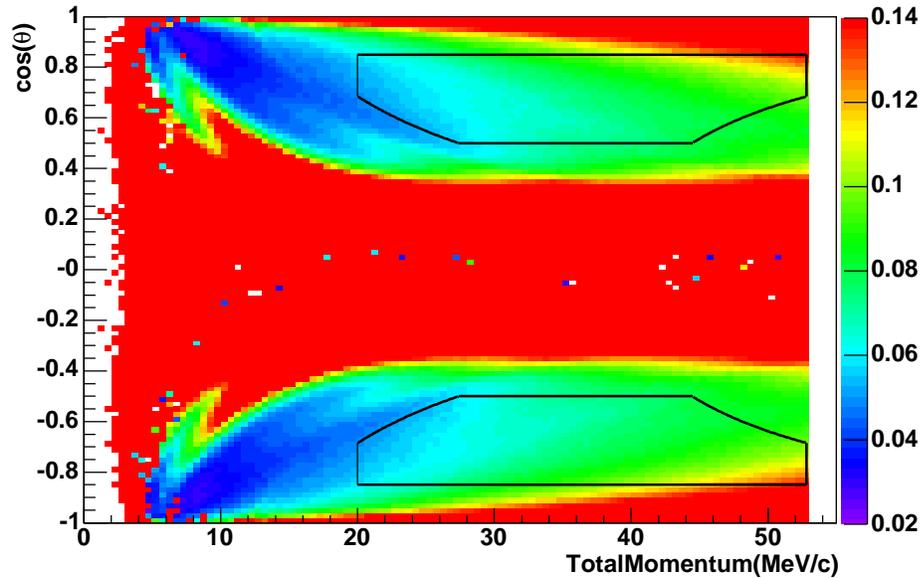
Figure 4.2: Detector resolution as a function of momentum and $\cos\theta$. Colour scale shows the magnitude of the detector resolution, in MeV/c. The black line describes the fiducial volume used in this analysis

momentum. Therefore,

$$
\begin{aligned}
\delta p_{tot} \quad &\sim \quad \frac{\delta p_t}{|\sin\theta|} \\
&\sim \quad \frac{\delta x}{\sqrt{n}|\sin\theta|}.
\end{aligned}
\tag{4.3}
$$

While this model explains the gross behaviour of the detector resolution, it is far from a complete description as there are features that appear in Fig. 4.2 that cannot be entirely understood through this line of reasoning. For example this model does not contain the multiple scattering that is responsible for the increase in resolution at high angles or reflect the possibility of a track passing through multiple drift cells in a plane. For this reason the widths measured directly from the simulation as shown in the figure were used for the detector resolution.

Fiducial Choice

In view of these results, an a priori decision on the fiducial region in which we will conduct our search can be made. The first fiducial cut eliminates positrons that hit the glass frames holding the wire chambers in place. This consists of a cut on the transverse momentum at 38.5 MeV/c. Similarly there are cuts on the angle such that $0.50 < |\cos\theta| < 0.85$ to avoid areas of the fiducial where the reconstruction is poor due to steep tracks or extremely narrow helices.

Of particular note is the cut made on low longitudinal momentum at 13.7 MeV/c. This cut was made with the purpose of avoiding areas where the reconstruction is poor due to detector periodicities mentioned before. When positrons move in the direction of the detector axis at these momenta, an ambiguity in the reconstruction of a track appears, dramatically increasing the error on the reconstruction of the transverse momentum. As the highest longitudinal momentum where this periodicity appears is at 12 MeV/c, this cut at $p_l = 13.7$ MeV/c conservatively removes these poorly reconstructed events from the fiducial.

The final two cuts are necessitated by the goals of the fitting exercise. A fiducial cut on the total momentum was motivated by the number of bins available given the longitudinal momentum cut at low total momenta. The longitudinal momentum cut makes the fiducial more narrow in $\cos\theta$ as the total momentum decreases. As the peak search is looking for mono-energetic peaks in the spectrum, the wider the

fiducial is in $\cos\theta$, the better the fit will be. On the other hand, if the fiducial covers a very small range of $\cos\theta$ at a given momentum, the decreased statistics will produce misleading effects, such as artificial structures in the spectrum or increased error, that confuses the peak search at low momenta. For this reason, a lower bound was placed on the total momentum at 20.0 MeV/c. The upper bound on the total momentum is a little bit more fluid as the only limiting factor is the quality of the simulation and the data. For this reason the upper bound can be allowed to approach the endpoint of the spectrum at 52.827 MeV/c. An upper fiducial limit excluding the endpoint was used for the determination of the Michel parameters in order to reduce the sensitivity of the result to the response of the detector. I can include the endpoint of the spectrum in this search for exotic decays because I am not concerned with producing the best errors on the Michel parameters. I will discuss the sensitivity of the Michel parameters to the inclusion of the endpoint, in section 5.1.1.

4.2 Fitting The Spectrum

To look for the presence of a peak I need to fit the entire muon decay spectrum with a good knowledge of the background. Due to the quality of the TWIST spectrum the only background I need to consider is the Michel decay spectrum. The methods developed for the purpose of the measurement of the Michel parameters are the best tools available for the purpose of determining this background. These methods use the linearity of the Michel spectrum with respect to the Michel parameters to generate a fit based on the Monte Carlo simulation of the spectrum. Generating the fit in this way eliminates the requirement for knowledge of the reconstruction efficiency of the detector, as the data spectrum and the fit to the data spectrum are treated in exactly the same way by the analysis and reconstruction effects will cancel out. The absolute value of the Michel parameters does not need to be known either using this method as one only sees the change in the Michel parameters with respect to some reference value for any given parameter. The reference value can be hidden immediately before the generation of the simulation. This allows TWIST to use a truly blind analysis for the Michel parameters. This ability to fit the Michel parameters in a blind way was maintained within the analysis program described

here, even though blindness to the Michel parameters is not required in the search for exotic decays.

While it is not strictly necessary to refit the Michel parameters every time a fit for an exotic particle is carried out, the analysis procedure described here does exactly that. It turned out to be less work to preserve all of the features of the existing program, making this analysis more sophisticated than it needs to be for the purpose of looking for peaks in the spectrum. On the other hand this measurement of the branching ratio includes the errors on the Michel parameters, which are expected to be small with the uncertainty in the branching ratio, as a direct result of the fitting process.

4.2.1 Fitting the Michel Spectrum

Rather than fitting the shape of the Michel spectrum directly to the data, TWIST exploits the linearity of the equations 2.2 and 2.3 to generate a fit derived from Monte Carlo. Consider the differential muon decay distribution in momentum, p, and $\cos\theta$. If I expand about some preselected values of the Michel parameters, ρ_0, η_0, $P_\mu\xi_0$, and $P_\mu\xi\delta_0$,

$$
\begin{aligned}
\left.\frac{\partial^2\Gamma_{fit}}{\partial x\partial\cos\theta}\right|_{\rho,\eta,P_\mu\xi,\delta} = \ & \left.\frac{\partial^2\Gamma}{\partial x\partial\cos\theta}\right|_{\rho_0,\eta_0,P_\mu\xi_0,\delta_0} \\
& + \ \frac{\partial}{\partial\rho}\left.\frac{\partial^2\Gamma_{MC}}{\partial x\partial\cos\theta}\right|_{\eta_0,P_\mu\xi_0,\delta_0}\Delta\rho \\
& + \ \frac{\partial}{\partial\eta}\left.\frac{\partial^2\Gamma_{MC}}{\partial x\partial\cos\theta}\right|_{\rho_0,P_\mu\xi_0,\delta_0}\Delta\eta \\
& + \ \frac{\partial}{\partial P_\mu\xi}\left.\frac{\partial^2\Gamma_{MC}}{\partial x\partial\cos\theta}\right|_{\rho_0,\eta_0,P_\mu\xi\delta_0}\Delta P_\mu\xi \\
& + \ \frac{\partial}{\partial P_\mu\xi\delta}\left.\frac{\partial^2\Gamma_{MC}}{\partial x\partial\cos\theta}\right|_{\rho_0,\eta_0,P_\mu\xi_0}\Delta P_{mu}\xi\delta \qquad (4.4)
\end{aligned}
$$

For a parameter ϖ, the change between the value of ϖ and ϖ_0 is $\Delta\varpi$. It should be noted that any given derivative of the spectrum in a given parameter is independent of that parameter. These derivatives are simulated and analysed in the same way as the standard Monte Carlo. The data and Monte Carlo spectra are both normalised to one within the fiducial region during the fit. This means that the fit result has an

implicit normalisation equal to the number of Monte Carlo events contained within the fiducial region.

Using this spectrum expansion to generate a fit has the benefit of removing any biases that result from the reconstruction, as all the spectra; data, Monte Carlo and derivatives; are treated in the same way. Thus, for the purpose of a search for peak resulting from exotic decays, effects such as detector periodicity, which could possibly generate a false exotic decay signal already exist in the the simulated Michel spectrum, removing any need to introduce further corrections. However, the availability of Monte Carlo also places a limit on the sensitivity of a peak search. As the sensitivity is based on the quadratic sum of the statistical fluctuations of the data and the Monte Carlo simulation the generation of the Michel spectrum fit by this method increases the error in the branching ratio of any peaks that may be found by the search. To limit the relative significance attached to the statistical fluctuations of the simulation, the Monte Carlo data sets contain roughly 150% more events than the data. Although it may be of interest to measure the branching ratio directly using a smooth Michel Spectrum, such an approach would neglect the effects of the reconstruction that are implicitly accounted for by using this Monte Carlo fitting approach.

4.2.2   Addition of the Peak to the Fit

To search for a two body exotic muon decay one must add a mono-energetic peak to the fit spectrum. As the mono-energetic peak is assumed to have very little intrinsic width, the resonance will have the shape of the response function of the detector. Because the resolution is angle dependent, the response function adds an angular dependence to the peak, even when an isotropic exotic muon decay signal is considered. Various effects, such as hard scattering of the positrons, enhance the tails of the distribution making a Gaussian approximation not quite sufficient for a reasonable estimation of the detector response. Instead a double Gaussian

distribution was adopted to better approximate the response function [19],

$$f_{DG}(p; \bar{p}, \sigma_1(\bar{p}, \cos\theta), \frac{N_2}{N_1}, \frac{\sigma_2}{\sigma_1}) =$$
$$\frac{1}{\sqrt{2\pi\sigma_1^2(\bar{p}, \cos\theta)}(1 + \frac{N_2}{N_1})} \left(\exp\left(-\frac{(p - \bar{p})^2}{2\sigma_1^2(\bar{p}, \cos\theta)} \right) + \frac{N_2}{N_1}\frac{\sigma_1}{\sigma_2} \exp\left(-\frac{(p - \bar{p})^2}{2\sigma_1^2(\bar{p}, \cos\theta)}\frac{\sigma_1^2}{\sigma_2^2} \right) \right).$$

$$(4.5)$$

In this function the ratios N_2/N_1 and σ_1/σ_2 are treated as properties of the detector response and are held constant within the fiducial. The width of the peak $\sigma_1(\bar{p}, \cos\theta)$ was determined directly from the Monte Carlo in the method described in Sec. 4.1.2. The free parameter of this response function is the peak mean, $\bar{p}$. The function is multiplied by a constant to match the amplitude of the fit peak to the amplitude of a peak in the data. As this is a normalised double Gaussian the constant factor will be the area under the peak. The branching ratio, denoted as $\mathcal{B}_{eX^0}$, associated with the peak is the second free parameter associated with the peak and is proportional to the area of the double Gaussian, with the appropriate scaling factors. The function fit to the peak is

$$f_{peak}(p, \cos\theta; \bar{p}, \mathcal{B}_{eX^0}) = \mathcal{B}_{eX^0} A f_{DG}(p, \cos\theta; \bar{p}). \qquad (4.6)$$

where A represents the factor relating the peak amplitude to the branching ratio. I will continue by describing the components of this scaling factor.

The first contribution to the scaling of the peak amplitude explicitly includes the reconstruction efficiency, and the effect of the cuts applied to the data, in the double Gaussian peak. The purpose of this is to make the area measured using the double Gaussian model consistent with the normalisation implicit in the Michel spectrum fitting procedure. As these effects are both dependant on the momentum and angle of the reconstructed decay positron, this scaling must be done for each bin in p and $\cos\theta$. The effective peak in the fit becomes

$$f_{eff}(p, \cos\theta; \mathcal{B}_{eX^0}, \bar{p}) = \mathcal{B}_{eX^0}\frac{N_{rec}(p, \cos\theta)}{N_{true}(p, \cos\theta)}f_{DG}(p, \cos\theta; \bar{p}), \qquad (4.7)$$

where $N_{rec}(p, \cos\theta)$ is the number of events reconstructed in the analysis at a given momentum, p, and $\cos\theta$, after all of the cuts (excepting the fiducial cuts) are applied to the Monte Carlo, while $N_{true}(p, \cos\theta)$ is the number of the surface muon decay

events generated by simulation that pass only the 0.7 μs decay time cut for the same p and $\cos\theta$. This ratio of spectra is exactly the histogram shown in figure 4.1. An example of the results of equation 4.7 was shown in Fig. 2.2.

The remaining contributions to the scaling of the amplitude necessary to generate the branching ratio are constant factors to account for the restricted fiducial. Because the signal for the exotic decay signal is localised in momentum, the factor relating the number of counts contained in the fiducial to the total number of expected counts is number of events expected within the fiducial at $\bar{p}$ integrated over $\cos\theta$, divided by the number of events observed at $\bar{p}$ integrated over all $\cos\theta$. The corresponding factor for the Michel decays is the number of events expected for all momentum and angle divided by the number of events contained by the fiducial (since this factor is meant to scale the denominator of the branching ratio). Including all of these normalisations the spectrum used in the fit becomes,

$$F_{fit}(p, \cos\theta) = \left.\frac{\partial^2\Gamma_{fit}}{\partial p\, \partial\cos\theta}\right|_{\rho,\eta,P_\mu\xi,\delta} +$$
$$\mathcal{B}_{eX^0}\frac{N_{true}}{N_{fid}}\left[\frac{\int_{fid(\bar{p})} N_{true}(\bar{p}, \cos\theta)d\cos\theta}{\int_{-1}^{1} N_{rec}(\bar{p}\cos\theta)d\cos\theta}\right]\frac{N_{rec}(p, \cos\theta)}{N_{true}(p, \cos\theta)}f_{DG}(p, \cos\theta; \bar{p})$$

$$(4.8)$$

Here I have used N_{true} to represent the total number of Monte Carlo events in the full muon decay spectrum, and N_{fid} to represent the number of Monte Carlo events reconstructed within the fiducial. The notation $fid(\bar{p})$ applied to the integral over $\cos\theta$ serves as a reminder that the integral is over the fiducial range of $\cos\theta$ at the momentum, p. After these corrections the value of $\mathcal{B}_{eX^0}$ accurately reflects the required branching ratio. An example of the results of the fit to the peak appears in figure 4.3. For reference, the peak in this figure is overlaid with the residuals of the data and the fit to the Michel spectrum without the peak.

4.2.3 Particulars to this Fitting Method

These functions may now be used to determine the best branching ratio for any two body decays accessible within the muon decay spectrum. To find all possible decays accessible to muon decay, the momentum range is subdivided into a number
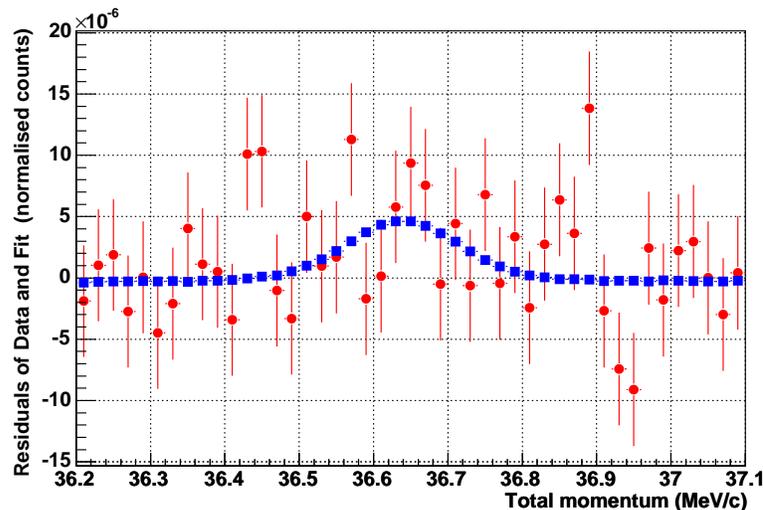
Figure 4.3: Projection of the peak generated using this method (■) overlaid with the difference between the data and the Michel spectrum fit (●). The vertical scale is the number of events per 20 keV/c bin in momentum divided by the total number of events contained in the fiducial range of the histogram.

of ranges, all on the order of the momentum resolution in width (approximately 200 keV/c). For each subrange in momentum the fit is repeated, with the peak mean restricted to be within the subrange and all other parameters left as free parameters.

Fitting a peak to the spectrum was done in a way to find positive peaks in a given region of the spectrum, even in the presence of a more significant negative peak. All fits are done using the ROOT Minuit class [20] to find the best fit in the Michel parameters using the χ^2 statistic. However, one must consider that the best fit in this case does not necessarily produce the most appropriate value for the branching ratio; in this case I want the most positive value within a subrange that still minimises the χ^2. To this end, the exotic decay peak is fit to the residuals of the data and the Michel Spectrum fit a maximum of three times within each momentum subrange. The first fit allows all of the fit parameters to vary freely (excepting the constraint on the mean peak position). If the fit produces a negative branching ratio, the fit is redone using a fixed, positive branching ratio, with the purpose of finding where the residuals of the Michel fit is the most positive. When this minimisation is

complete the χ^2 is again minimised by altering the value of $\mathcal{B}_{eX^0}$, while leaving all the other parameters fixed. This technique was chosen to find the largest positive peak within a given momentum subrange even in the presence of a negative peak which yields better χ^2. A byproduct of this method is that the fits tend toward the edges of the fitting subrange if a maximum does not exist in a given fit range. This tends to result in an unrealistically small error on the position of a peak at the edge of a subrange due to the method used to constrain the momentum. It is for this reason that the branching ratio is not constrained to positive values as I want the error to be as accurate as possible. This method was chosen to find the largest possible peak within a given subrange, while avoiding troughs appearing within the spectrum because of statistical fluctuations, and minimising the number of peaks counted multiple times.

In the early stages of this analysis the fit was done using both the complete spectrum (the distribution differential in energy and $\cos\theta$) and the spectrum integrated over $\cos\theta$. A fit using an angle integrated spectrum is a good way to improve the statistics of a peak and quickly remove any effects that may be due to any small statistical fluctuations which are localised in angle. However, after noting the high angular dependence of the momentum resolution and the subsequent result on the response function, this approach was questioned as being too simplistic. It was finally decided to stop using fits of the angle integrated spectrum when it was discovered that the result of a fit using the complete spectrum generates smaller errors when there is a large amount of background anisotropy. This conclusion is demonstrated in appendix A.

4.3   Definition of Confidence Intervals

Using the branching ratio resulting from the fit we may define a corresponding confidence interval. This is a range of values which contains the true value of an experimental parameter Typically a confidence interval may be defined by locating where the normal distribution of the errors centred at the mean value produces a integral under the curve equal to a prescribed value. However the classical method, also referred to as the Neyman method of defining confidence intervals is not sufficient for

the interpretation of the small signals we expect. If the signal is too small the classi-
cal interval allows for the existence of negative signals, which is an unphysical result
for a branching ratio. A decision must be made after the measurement whether the
Neyman construction is consistent with zero and an upper limit is adopted rather
than a two sided interval. This "flip-flopping" often leads to an over-coverage of the
confidence interval. To avoid having to make this decision, an alternate method of
defining the confidence intervals, advanced by Feldman and Cousins [21], was used
for this analysis.

The Feldman Cousins method for defining the confidence intervals uses an explicit
ordering principle to define the confidence interval. For every value of the measured
parameter x, an ordering function, $R(x; \mu)$ is defined. The purpose of this ordering
function is to determine the relationship between the upper and lower limits of the
confidence interval. With this, an interval $x \in [x_1, x_2]$ can be found which satisfies
$R(x_1; \mu) = R(x_2; \mu)$ and

$$\alpha = \int_{x_1}^{x_2} P(x; \mu) dx \tag{4.9}$$

where α is the required frequency of the true result existing within the interval for
repeated experiments. The confidence interval is therefore defined as the interval
of possible true values for the experimental parameter, $\mu \in [\mu_1, \mu_2]$ where μ_1 is
the smallest μ that allows an x interval that contains the measured value of the
parameter, x_0 (ie. μ such that $x_0 = x_2$) and μ_2 is the largest μ that allows an x
interval containing x_0 (ie. μ such that $x_0 = x_1$).

The ordering function which I use is represented as

$$R(x; \mu) = \frac{P(x; \mu)}{P(x; \mu_{best})}. \tag{4.10}$$

where we define $P(x; \mu_{best})$ to be the maximum likelihood. For physically allowed
values of x, $\mu_{best} = x$ while for values of x which are non-physical μ_{best} is taken to be
zero. This acts as a soft constraint on the summation of the probability distribution
function.

In this particular case the likelihood has a Gaussian distribution of width σ
constrained to positive values of the branching ratio. For this reason a piecewise
definition of $P(x; \mu_{best})$ is used to produce a soft constraint on the lower limit of the
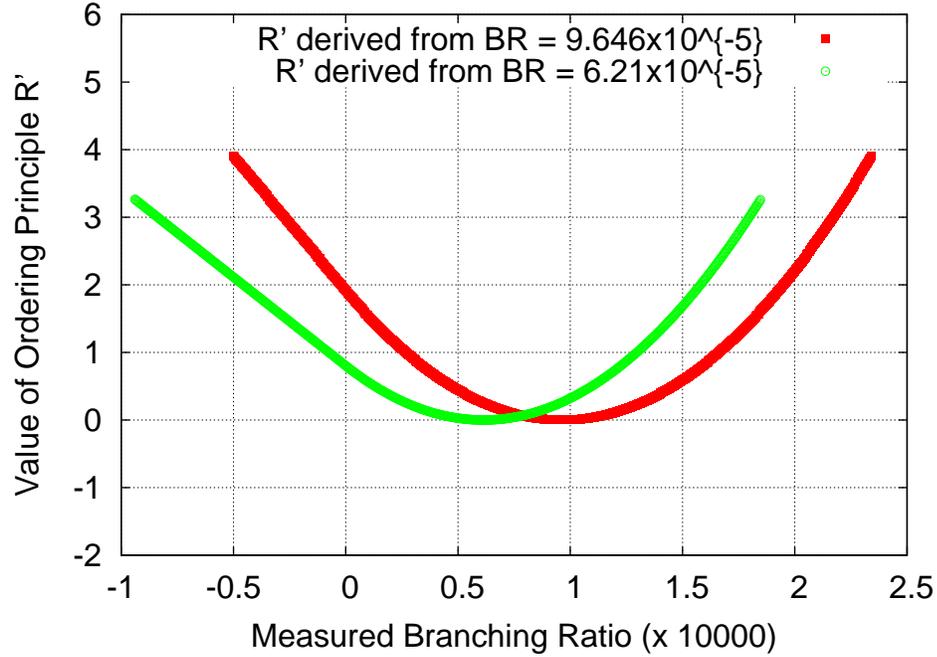
Figure 4.4: The ordering function, R', in the particular case of the addition an isotropic peak to the Michel spectrum, for two different central values of the branching ratio.

confidence limit where the lower limit is less than zero.

$$P(x; \mu_{best}) = \begin{cases} \dfrac{1}{\sqrt{2\pi\sigma^2}} & \text{if } x \geq 0 \\ \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\dfrac{x^2}{2\sigma^2}\right) & \text{if } x < 0 \end{cases} \qquad (4.11)$$

For values of x greater than zero, $P(x; \mu_{best})$ is the maximum likelihood, while for values of x less than zero it is the probability of x appearing should the probability distribution function be centred at zero. This choice impacts on the confidence level by allowing the smooth transition of confidence intervals from central valued to one-sided as the measured branching ratio goes from large branching ratios to branching ratios consistent with zero.

An equivalent expression for the ordering function can be produced by considering the ordering function $R' = -2\ln(R)$, where R is expression 4.10 using $P(x; \mu_{best})$ defined in equation 4.11. For a given measured value of the branching ratio, x, with

a Gaussian distributed error, σ, the ordering principle becomes,

$$R'(x;\mu) = \begin{cases} \dfrac{(x-\mu)^2}{\sigma^2} & \text{if } x \geq 0 \\ \dfrac{\mu^2 - 2x\mu}{\sigma^2} & \text{if } x < 0. \end{cases} \qquad (4.12)$$

Note that the true value of the branching ratio μ is limited to positive values. This requires that the slope of the ordering function when the function is linear (for negative x) is negative with a value greater than zero. This means that it is always possible to find a solution to $R'(x_1;\mu) = R'(x_2;\mu)$ for all allowed μ. The curves described by equation 4.12 for the particular case of adding a peak to the Michel spectrum using two different branching ratios (with different errors) appear in Fig. 4.4.

Using this definition of the ordering principle one may find an algebraic condition where $R'(x_1;\mu) = R'(x_2;\mu)$. The upper and lower limits of the confidence interval can then be found by assuming that $x_1 = x_0$ and $x_2 = x_0$, respectively, and solving equation 4.9 numerically to find μ_2 and μ_1. The 90% confidence band generated using this method appears in Fig. 4.5. This is can be compared directly to the 90% central valued confidence band generated via the Neyman method. The confidence intervals were defined for the branching ratio of a given subrange as a part of the fitting process.
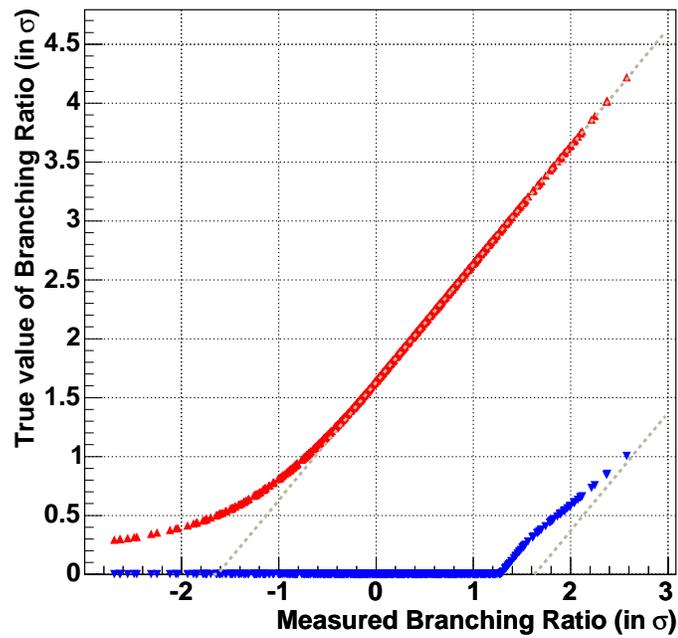
Figure 4.5: Region with a 90% confidence of containing the true value of the branching ratio. For a given branching ratio (horizontal axis) the actual value of the branching ratio lies between the upper limit for the region ($\blacktriangle$) and the lower limit ($\blacktriangledown$). The classical 90% central valued confidence region is contained within the brown dashed lines.

Chapter 5

Analysis Results and Discussion

From the analysis so far described, I produce a series of branching ratios and accompanying confidence intervals for the range of momenta contained within the fiducial. With these plots one can infer the range of values for the branching ratio which are consistent with the measured branching ratio. Should we see a branching ratio consistent with the presence of a previously unknown particle, the mass of that particle can be interpreted directly using equation 2.8.

As the 2002 data was taken in a number of different conditions for the purpose of testing systematics, I will first consider the exclusion plots for each separately. The generation of the exclusion plots for the branching ratios of exotic muon decays have been done for a subset of these differing conditions. I will then show the same result for the sum of the data sets.

5.1   Fit to the Michel Decay Spectrum

The result that I am about to describe use the data collected by TWIST in the fall of 2002. This data has already been well studied by TWIST with the results appearing in [9] and [22]. I will refer to these papers for consistency as the Michel decay is the background for the two body decay signal. However the published results are not the final word in the Michel Parameters used in this study. Changes made to adapt the Michel parameter analysis for the purpose of searching for exotic decays can introduce variations in the Michel parameters extracted from the spectra. Changing

the fiducial region, for example, produces variations in the Michel parameters. In particular the inclusion of regions that are sensitive to the detector response function, such as the region near the decay spectrum endpoint, produces a serious issue that will be given special consideration. Due to the dynamic nature of the fitting method it is also reasonable to expect a small variation in the Michel parameter fit as a result of the presence of the peak. Finally, in order to observe structures in the momentum on the order of the momentum resolution, I chose to bin the distributions with a much finer scale in momentum than that used in the TWIST publications to date. In the standard TWIST Michel parameter analysis the bins in the fitting histograms are 0.5 MeV/c wide in momentum whereas a momentum bin width of 20 keV/c was chosen for this analysis.

5.1.1 Sensitivity of the Michel Parameters

For this particular study ρ has the largest sensitivity to a two body decay signal of all the Michel parameters. The parameter η has only a very minimal impact on the spectrum because it is scaled by the mass of the electron. Thus η was fixed at the best world average and subsequently ignored for the purpose of determining the Michel spectrum. The anisotropic parameters $P_\mu\xi$ and δ have a more significant contribution to the overall shape of the background. However their contribution is orthogonal to that made by the introduction of an isotropic peak. As changes in ρ are the primary source of changes in the energy dependence of the decay spectrum we can expect that ρ will have the largest impact on, and will be impacted the most by, the introduction of a mono-energetic peak. For this reason I will focus this discussion on the value of ρ used in this analysis.

The four data sets and the associated values of ρ for a number of different fiducial volumes are shown in Table 5.1. All of these fits were generated using different fiducial regions. The first column shows the value of ρ obtained from a previous analysis for a fiducial with an upper limit on the total momentum of 50.0 MeV/c and a lower limit on the momentum of 13.7 MeV/c. The second column is the result for ρ generated by this analysis using a histogram bin width of 20 keV/c (as opposed to the 0.5 MeV/c bin width used in the first column) and a lower limit on the total momentum of 20.0 MeV/c (leaving the upper limit unchanged with respect to the first column). The

Data Set	ρ from [18]	ρ for $p_{tot} < 50.0$ and $p_{tot} > 20.0$	ρ for $p_{tot} < 52.827$ and $p_{tot} > 20.0$
Set A	0.75083±0.00083	0.75270±0.00085	0.75230±0.00069
Set B	0.74911±0.00066	0.74914±0.00067	0.74866±0.00055
B = 1.96 T	0.74956±0.00066	0.74993±0.00061	0.74926±0.00054
B = 2.04 T	0.75203±0.00071	0.75267±0.00072	0.75139±0.00059

Table 5.1: Results of fitting the Michel parameter ρ for selected TWIST muon decay spectra. The errors shown are purely statistical and do not reflect the systematic uncertainties.

third column shows the values of ρ measured using the fiducial regions described in full in Sec. 4.1.2 with the upper limit on the total momentum expanded to include the endpoint of the Michel spectrum.

The first two columns show the effect that the choice of binning has on the Michel parameters. This shows a change in the Michel parameters that is reasonable given the errors in all cases but that of set A, where we see a 2σ difference between the two results. This difference only appears in this set and not in the others because data set A was one half the size of the other data sets, with correspondingly lower statistics at low momenta. This allows for much more variation in the central value of *rho* as the statistical errors on a given bin at low momentum become as large as 20% of the counts in the bin. Excluding the areas of the histogram with poor statistics or including areas of higher statistics momenta reduces the difference between the two measurements of of ρ. This argument extends to the other Michel parameters.

Comparing the last two columns affords a view of the effect of changing the fiducial to include the endpoint has on ρ. It appears that the change in the parameters due to the increased fiducial region is not so great as the change due to the binning. The errors shown are statistical only and do not reflect the error due to systematic uncertainty. The systematic uncertainties also depend on the choice of fiducial, but in a non-trivial manner. The inclusion of the endpoint in particular will have a very large effect on the systematics. Since this measurement is only for the determination of the background of a two-body decay signal the deviations in the Michel parameters and the systematic uncertainties generated by choices in the fiducial and the

histogramming of the data can be ignored.

The introduction of the mono-energetic peak itself may generate a change in the Michel parameters. Indeed when one looks at the distribution of the Michel parameters for a given set about the result of the fit without a peak one does see small variations (see Fig. 5.1). However these variations are well below the level of the statistical error in the fit and do not become an issue in the definition of the Michel fit.
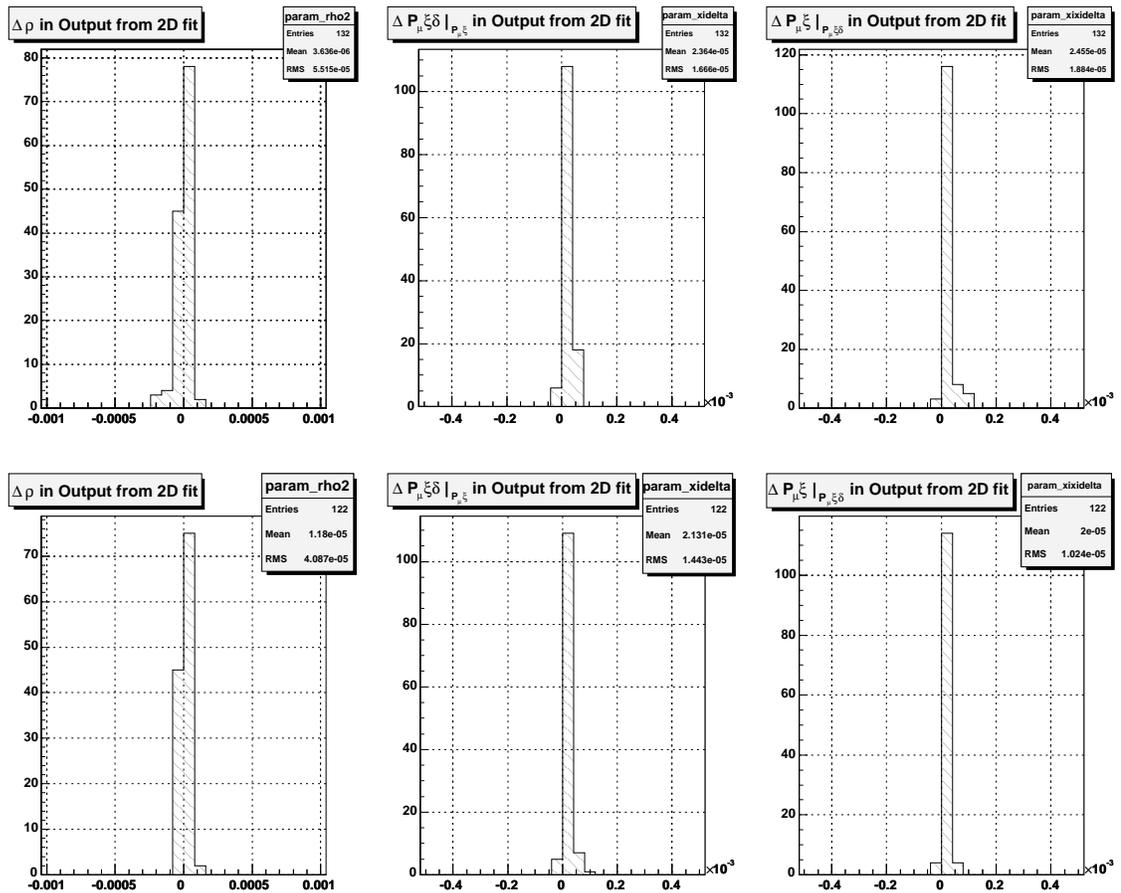

Figure 5.1: The variations in the Michel parameters resulting directly form the inclusion of mono-energetic peaks a various momenta. The top row shows these variations in set B for ρ, and the products $P_\mu \xi \delta$ and $P_\mu \xi$. The second row contain the same variations for set A.

5.1.2   Quantifying the Endpoint Kinematics

A consistent problem in the first TWIST analysis was a mismatch of the momentum resolution, measured at the endpoint of the Michel spectrum, between the data and the Monte Carlo simulation. Typically this difference between the endpoint resolution of the data and the simulation was between 4 keV/c and 7 keV/c, approximately a 5 to 10 percent difference in the resolution. This difference in the momentum resolution does not affect the measurement of the branching ratio. To verify this I added a constant 10 keV/c to the momentum resolution used by the peak fitting algorithm and applied the analysis to a typical data set. No detectable change appeared in the branching ratios measured using this difference altered resolution.

While this difference can be ignored for the purpose of estimating the resolution throughout the rest of the spectrum, the mismatch at the endpoint is of concern because it manifests itself as a peak (or trough) at the endpoint. For this reason this region is excluded from the fiducial in the standard Michel parameter fits. However in this study the endpoint cannot be ignored completely as the endpoint is the primary region of physical interest. For this reason I will provide an estimate of the effect that this resolution mismatch has on the endpoint

The theoretical spectrum has a kinematic cutoff which is described as a step function with an amplitude corresponding to the value of the Michel spectrum at the endpoint. A convolution of the response function with the endpoint this can be approximated as an error function with a width equal to the resolution at the endpoint. Recalling that the resolution varies as $1/|\sin\theta|$ within the fiducial, I can produce a simple model of the endpoint of the Michel spectrum using the standard model values for the Michel parameters.

$$H(x, \cos\theta) = \frac{p^2}{Np_{max}^2}\left(3 - 2\frac{p}{p_{max}} - P_\mu\left(2\frac{p}{p_{max}} - 1\right)\right) \times \mathrm{Erf}\left(\frac{(p_{max} - p)\sin\theta}{\sigma_0}\right)$$

$$(5.1)$$

Here, I have used p_{max} to indicate the momentum of the kinematic cutoff for the spectrum, and σ_0 to indicate the value of the momentum resolution at the spectrum cutoff when $\cos\theta = 0$. The variable N normalises the Michel spectrum to 1 when the spectrum is integrated over the fiducial.

As an example I will use the endpoint resolution values associated with nominal data set B and its Monte Carlo simulation to generate the magnitude of the branching
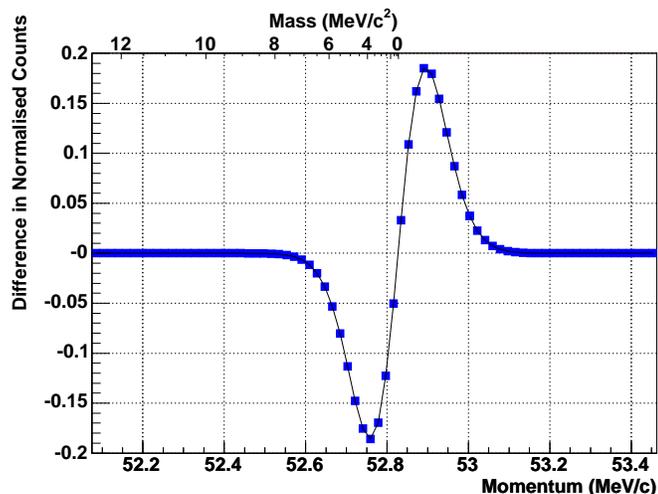
Figure 5.2: The model of the endpoint resolution using the endpoint resolutions from data set B and its associated Monte Carlo. The plot shows the projection of the endpoint difference onto the momentum axis clearly showing a trough for kinematically allowed momenta.

ratio we can expect to measure as a result of the mismatch. For set B, $\sigma_0 = 74.3$ keV/c, while $\sigma_0 = 69$ keV/c for the associated Monte Carlo simulation. The result of taking the difference between the endpoint models using these two values for the resolution is shown in Fig. 5.2. Applying the scaling factors discussed in Sec. 4.2.2 yields a negative branching ratio of 8.7×10^{-5} immediately before the endpoint. As the positive peak from this resolution difference appears at momenta higher than the endpoint, one should not expect to find a positive peak as a result of the mismatch. The magnitude of this peak is not related to statistics. Rather, it is a function of the difference between the resolution of the data and the Monte Carlo simulation.

5.2 Consistency Checks using Single Data Sets

Examination of the analysed data sets independently is good for evaluating anomalies that may have appeared due to the problems unique to single data sets because of running conditions. Exclusion plots were generated for each of the nominal TWIST data sets. This required a minimal amount of preprocessing beyond that which is

required for the measurement of the Michel parameters and allows the meaning-ful consideration of previous studies of the data as a guide for understanding any anomalies that may appear during this study.

Figure 5.3 show the exclusion plots for the four nominal data sets used in this study. These include the mean value of the branching ratio measured along with their upper and lower 95% confidence intervals plotted against the mean momentum of the peak. The boson masses associated with a given peak found in the momentum spectrum is shown on the scale at the top of the plot. We can quite clearly see that the 95% upper confidence limit for all of these sets increases with momentum and lie between 4×10^{-5} and 1×10^{-4}. This is reasonable given the statistics. I can produce an estimate of the expected branching ratio for a given peak in the spectrum based on the statistics for a single data set by assuming that the events are uniformly distributed over a fiducial region uniform in $\cos \theta$ with respect to momentum and that the detector response has a constant momentum resolution of 100 keV/c. Given that there are 1.6×10^7 events in a standard data set, these assumptions generate peaks, due to statistical fluctuations, with branching ratios on the order of 2×10^{-5}. A 95% upper confidence limit on a branching ratio of this amplitude is 5×10^{-5}. Considering that the fiducial and distribution is far more complicated than the situation described, this estimate seems to indicate that the measured results are indeed reasonable.

If one looks at the lower limits of the confidence intervals indicated on these plots one sees several points scattered throughout the sets where the lower limit is non-zero. Put another way, the branching ratio at this mass is not consistent with zero at the 95% confidence level. This is not to say that this is evidence for a new particle, however. In considering a large number of trials, one can always expect to see some outliers. For all of the data sets there are in excess of 500 resolution widths at various momenta. At a 95% level of confidence one can reasonably expect there to be 25 false positive signals. As there are fewer positive signals than this it seems unlikely that there are any systematic effects contained within the data sets that is generating false positive signals. A more detailed discussion of particular peaks with non-zero lower limits will accompany the discussion of the results of the full 2002 data set
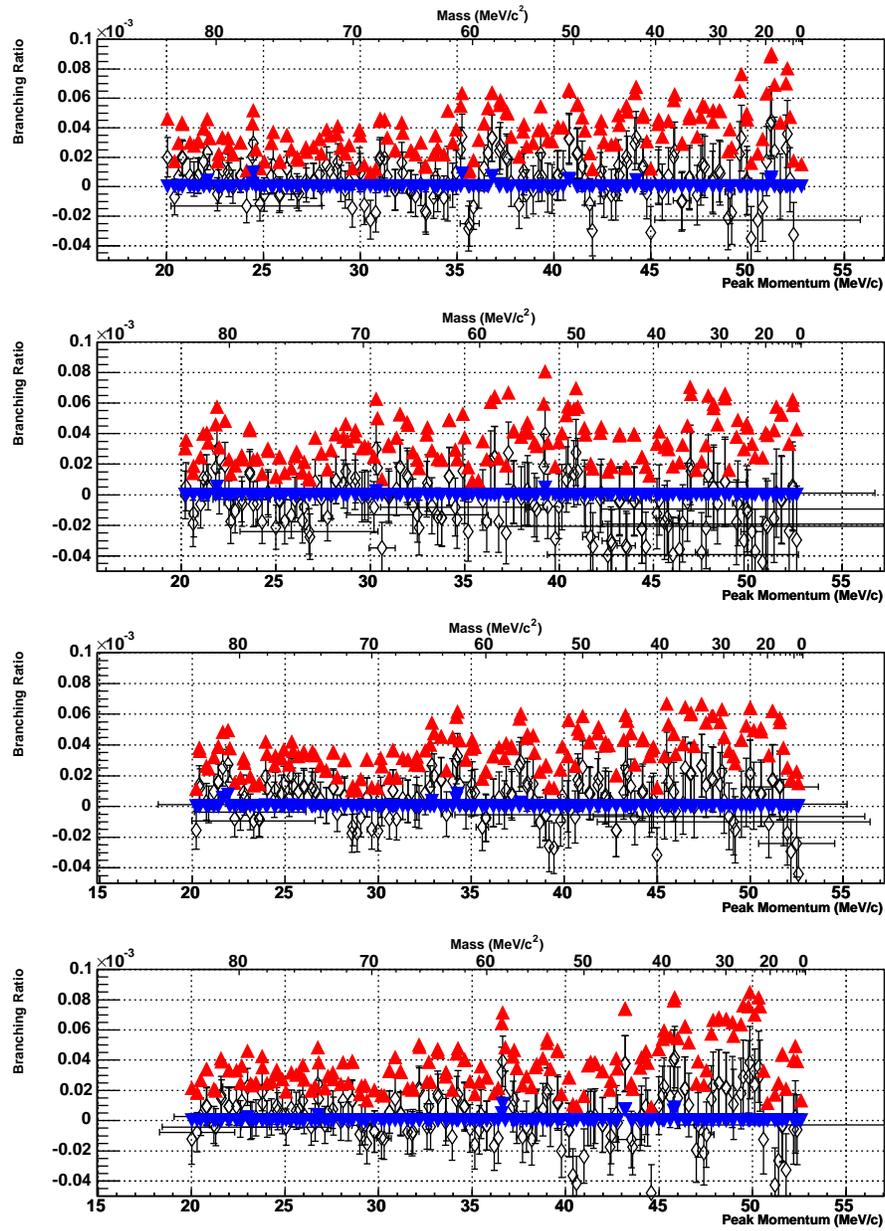
Figure 5.3: Exclusion plot for the presence of isotropic peaks resulting from two body decays in the nominal data sets for momenta approaching the kinematic limit of the spectrum. The figures show the upper limits (▲), mean values (◊), and lower limits (▼) for the branching ratios of, from top to bottom, nominal set B, nominal set A, the set with B=1.96 T, and the set with B=2.04 T.

5.2.1   Treatment of Possible Anisotropy

Anisotropy in the exotic decay signal can introduced by weighting the response function (described in equation 4.5) with $1 \pm \mathcal{P} \cos \theta$ as discussed in equation 2.12. By the arguments outlined in Sec. 2.2.2, the muon polarisations calculated for the nominal data sets in [23] provide an upper limit for the polarisation $\mathcal{P}$ (defined in equation 2.12) expressed by the out going positrons in the decay peak. These polarisations appear in Table 5.2. For the purpose of the measurement of a branching ratio for anisotropic peaks I used values of $\mathcal{P} = \pm 0.89$ to test all of the data sets. Figure 5.4 shows the branching ratios and limits for exotic decays with $\mathcal{P} = \pm 0.89$.

The most prominent feature that appears, when one compares the positive anisotropic result to the negative anisotropic result, is the more restrictive upper bound on the positive anisotropic peak. Indeed, this is more restrictive than the bound set by the isotropic fit for the reasons expressed in Sec. 2.2.2 and the appendix. For similar reasons one can see that the lower 95% branching ratio is not consistent with
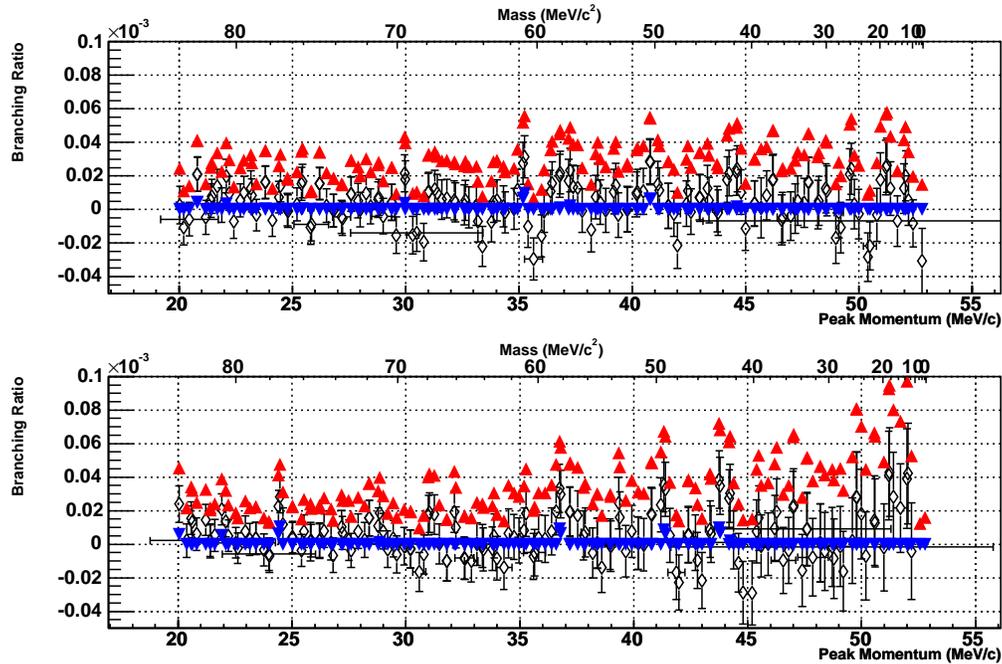

Figure 5.4: The anisotropic branching ratios for Nominal set B with a peak polarisation of +0.89 (positive anisotropy) in the top figure and -0.89 (negative anisotropy) in the bottom figure

Data Set	polarisation
Set A	-0.895
Set B	-0.898
B = 1.96 T	-0.889
B = 2.04 T	-0.893

Table 5.2: Average muon polarisations

zero for more points in the negative anisotropic analysis as opposed to the positive anisotropic analysis.

5.2.2 Monte Carlo Test of the Fitting Method

It became very important during the development of the analysis to check the results generated by the analysis. For this reason I shall discuss a simple check that was used to make sure that the fitting method generated reasonable branching ratios using a simple algorithm to insert a peak into a Monte Carlo data set (where there are no peaks) with a known branching ratio. It is possible to make these artificial peaks with a given anisotropy to test the ability of the algorithm to measure existing anisotropy.

I carried out this test at two different momenta, adding peaks to the Michel spectrum where the spectrum possesses a high anisotropy (at 50.1 MeV/c) and very little anisotropy (at 35.1 MeV/c). At both momenta, three peaks were generated with a branching ratio of 1.4×10^{-4}; one with a positive anisotropy ($\mathcal{P} = 1$), one with a negative anisotropy ($\mathcal{P} = -1$), and one with no anisotropy. Each of these peaks was then fit with $\mathcal{P} = 1$, $\mathcal{P} = -1$, and $\mathcal{P} = 0$. The results for this test appear in table 5.3. Some trends do appear in these numbers. Fitting the peak with the same anisotropy as the generation of the peak always generates a branching ratio consistent with the branching ratio used to generate the peak. Similarly, fitting the peak with an anisotropy opposite to that used in the generation of the peak will suppress the measured branching ratio. Understandably, the existing anisotropy of the Michel spectrum makes peaks with positive anisotropy stand out when they are fit with an isotropic peak. The errors for these results are highly correlated as the

Fit Anisotropy	Generated Anisotropy	$\mathcal{B}_{eX^0}$ at 35.1 MeV/c ($\times 10^{-4}$)	Deviation in σ from 1.4×10^{-4}	$\mathcal{B}_{eX^0}$ at 50.1 MeV/c ($\times 10^{-4}$)	Deviation in σ from 1.4×10^{-4}
+1	+1	1.35±0.25	-0.20	1.22±0.31	-0.58
	0	1.00±0.25	-1.60	0.98±0.31	-1.35
	-1	0.62±0.25	-3.12	0.78±0.31	-2.0
0	+1	1.36±0.31	-0.10	1.82±0.47	0.89
	0	1.31±0.31	-0.29	1.58±0.47	0.38
	-1	1.20±0.31	-0.65	1.34±0.47	-0.13
-1	+1	0.48±0.27	-2.70	1.13±0.56	-0.47
	0	0.89±0.27	-1.76	1.30±0.56	-0.18
	-1	1.09±0.27	-0.91	1.49±0.56	0.16

Table 5.3: Branching ratios produced by fitting peaks with the assumption of positive, negative and zero anisotropy to simulated peaks with positive, negative and zero anisotropy. Statistical noise in the Michel spectrum will strongly affect the results.

same background was used for all of these fits. This obscures the effect that I was looking for in the peak at 50.1 MeV/c as separating the statistical errors in the Michel Spectrum from the error that is due to the peak is a non-trivial issue. Furthermore, statistical fluctuations in the Michel spectrum strongly affect the results. However, there is a trend visible, suggesting that peaks fit with an opposing anisotropy with respect to the anisotropy of the peak itself will produce a suppressed branching ratio.

5.3 Results from the Complete 2002 TWIST Data

To observe the bound for multiple sets I applied this analysis to the sum of the histograms of the data sets. Following this addition of the sets the resulting histogram was analysed as before using summations of the Monte Carlo base sets and the derivative spectra, weighted so that the proportion of a given Monte Carlo set in the final summation matches match the proportion of the data set with the same experimental conditions represented in the in the data set summation, to complete
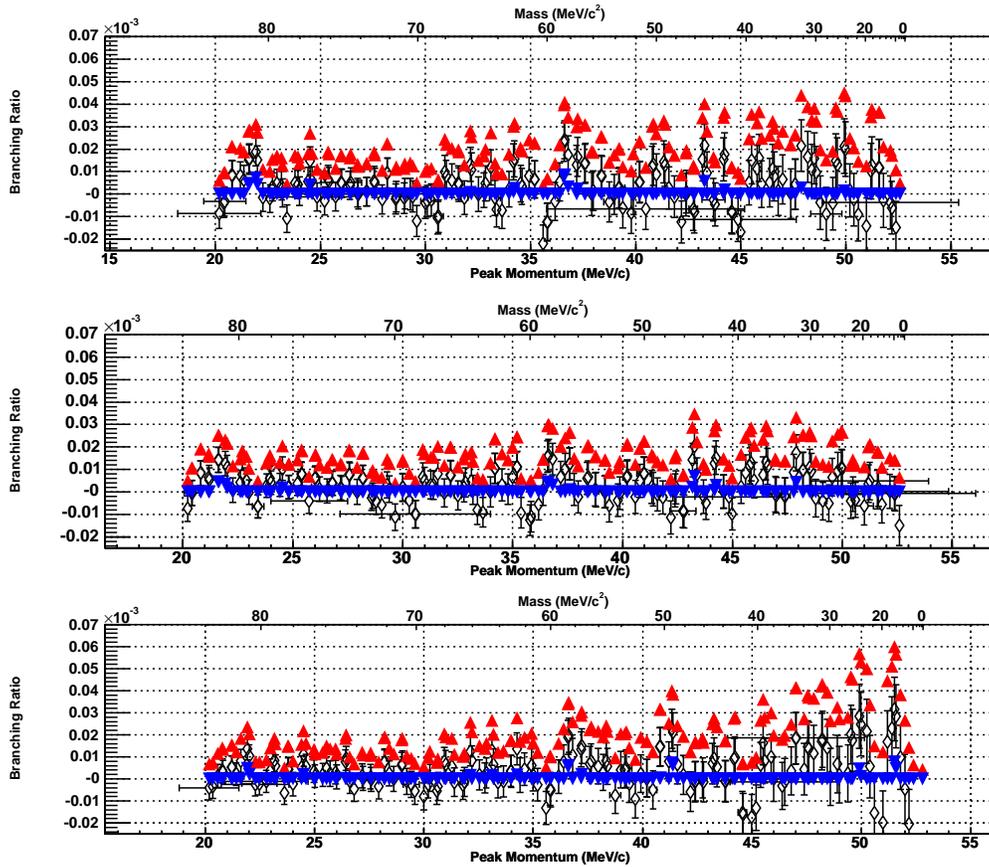
Figure 5.5: Exclusion plots resulting from the analysis of the sum of 2002 data sets. The top figure shows the exclusion plot generated from an isotropic peak. The Middle and bottom plots show the positive and negative anisotropic exclusion plots respectively. This represents an improvement of a factor of two over any single data set

the Michel parameter fit. The summation of the data sets should increase the sensitivity of this measurement by a factor of two. Figure 5.5 shows this improvement over any single data set very clearly. This yields a limit on the branching ratio that is one third of the previous best measurement [4] for the isotropic measurement.

I can also examine this summation of sets for possible decay candidates by looking for peaks with a non-zero 95% lower limits. As in the case of the single data set the number of peaks that we should expect to see is equal to the number of resolution

widths multiplied by the probability of the true branching ratio lying outside of
the 95% confidence limit. In this particular case we should again expect to see a
maximum of 23 instances where the confidence interval associated with the branching
ratio does not contain zero, assuming that the true branching ratio is zero at all
masses and that all measured peaks are entirely due to statistical fluctuations. Cases
where zero is not contained within the confidence interval appears in Table 5.4. In
this table, I have indicated the smallest confidence level that will contain zero for
each of the candidates. Interestingly, some of the cases shown have confidence levels
that exceed 99%. If I consider the largest peak of those shown; the isotropic peak at
36.64±0.05 MeV/c with a 2.87σ branching ratio; and compare the confidence limits
to the (rough) number of resolution widths in the momentum fiducial, I find that I
can expect a maximum of two instances of peaks with this amplitude ($500 \times 0.45\% =$
2.25). By this argument, a candidate peak must exceed 3.1σ before one can claim
that the peak does not entirely appear due to statistical fluctuations. Note that the
peak at 36.64 MeV/c is the peak shown in Fig. 4.3.

The trough, due to the endpoint resolution mismatch, is quite significant with
this summation of sets. To actually measure the significance of this feature, a small
alteration to the peak searching algorithm was made so that it measures the am-
plitude and confidence limits for troughs in the spectrum rather than peaks. The
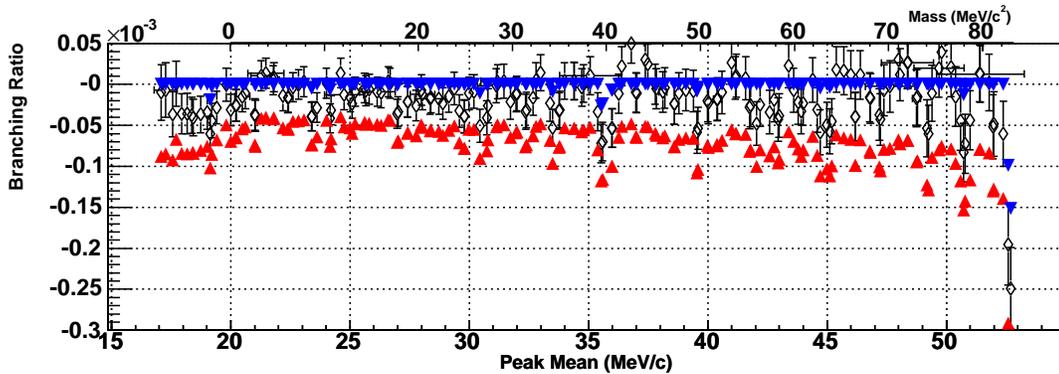

Figure 5.6: Altering a sign in the formulation of the fitting function turns the peak
searching algorithm into a trough searching algorithm. This figure shows the upper
limits (▼), mean values (◊), and lower limits (▲) for the branching ratios.

Peak Mean (in MeV/c)	Branching Ratio of Peak ($\times 10^{-5}$)	Branching Ratio in σ	Minimum Confidence Interval Containing zero
Isotropic Fit			
21.63±0.03	1.59±0.65	2.44	98.54 %
21.93±0.03	1.88±0.62	3.03	99.76 %
24.51±0.03	1.48±0.63	2.35	98.12 %
36.59±0.01	2.31±0.85	2.71	99.32 %
36.79±0.01	1.73±0.85	2.03	95.76 %
36.63±0.05	2.41±0.85	2.84	99.54 %
43.27±0.05	2.15±0.95	2.26	97.62 %
Fit with Positive Anisotropy			
21.63±0.04	1.43±0.57	2.51	98.80 %
21.93±0.04	1.26±0.54	2.33	98.02 %
36.59±0.01	1.57±0.70	2.24	97.50 %
36.64±0.07	1.63±0.70	2.32	97.96 %
36.82±0.07	1.46±0.70	2.08	96.34 %
43.27±0.04	2.02±0.74	2.72	99.34 %
44.24±0.05	1.52±0.74	2.05	95.86 %
47.90±0.04	1.74±0.80	2.18	97.00 %
Fit with Negative Anisotropy			
21.93±0.03	1.37±0.50	2.74	99.36 %
36.59±0.01	1.88±0.77	2.44	98.54 %
36.64±0.05	1.97±0.77	2.56	98.90 %
41.34±0.05	2.27±0.88	2.58	99.02 %
41.39±0.01	2.12±0.88	2.40	98.36 %
51.50±0.07	3.17±1.44	2.20	97.28 %
51.59±0.01	2.84±1.44	1.97	95.12 %

Table 5.4: Peaks that are consistent with a non-zero branching ratio to 95 % confidence in the summed 2002 data set. For each branching ratio the smallest confidence level that contains zero is shown.

result of just this procedure appears in Fig. 5.6. One can see a number of points whose 95% confidence intervals do not contain zero, similar to the number observed in figure 5.5. This can be considered to be a test of how sensitive the peak searching algorithm is to statistical fluctuations; after correcting for detector periodicities and other reconstruction inefficiencies there is no other explanation for troughs in the data. None of the troughs which possess 95% confidence intervals that do not contain zero are particularly significant. The exception is the trough at the endpoint, with a branching ratio of $(-6.7\pm1.4)\times10^{-5}$; a 4.7 σ deviation from zero. This is very close to what was approximated for the branching ratio of the peak resulting from this endpoint resolution mismatch as it was described in Sec. 5.1.2.

To remove the trough produced by endpoint resolution mismatch from the consideration of peaks, I added a trough to the fit that has all the characteristics of the anticipated resolution mismatch. To accomplish this I used the measurement of the resolution at the endpoint to generate the trough via the difference between equation 5.1 using the endpoint resolution of the data (σ_{data}) and the equation 5.1 using the endpoint resolution measured from the Monte Carlo (σ_{MC}). That is to say, the trough was generated using the expression

$$
\begin{aligned}
H_{diff}(p, \cos\theta) \;=\; & \frac{1}{2}\frac{\partial^2\Gamma_{data}}{\partial p\,\partial\cos\theta}\Bigg|_{p=52.4\ \mathrm{MeV/c}} \times \\
& \left(\mathrm{Erf}\left(\frac{(p_{max}-p)\sin\theta}{\sqrt{2}\sigma_{data}}\right) - \mathrm{Erf}\left(\frac{(p_{max}-p)\sin\theta}{\sqrt{2}\sigma_{MC}}\right)\right). \quad (5.2)
\end{aligned}
$$

The resulting distribution was added to equation 4.8. I have used the magnitude of the normalised data spectrum at 52.4 MeV/c to provide and amplitude for the difference of the kinematic endpoints of the data and the Monte Carlo. The exclusion plot for the sum of the data sets at the kinematic endpoint of the muon decay spectrum after including this correction is shown in Fig. 5.7. For the sake of comparison I have also included the exclusion plot for the same range of momenta where the background is not corrected for the resolution difference at the endpoint. From this result I can give a 95 % upper limit to the branching ratio near the kinematic endpoint of 3.4×10^{-5}.

This method assumes that the existence of a peak does not affect the measurement of the characteristics of the endpoint, including the endpoint resolution. These characteristics are used for the energy calibration of the Michel spectrum after re-

construction. Direct investigation conducted by adding an isotropic peak to the simulation of the Michel spectrum shows that the presence of a peak does not significantly affect the momentum resolution measured at the endpoint. For a peak with a branching ratio of $\sim$6$\times$10^{-5} added at or near the endpoint there is a change of $\sim$1.2$\pm$1.3 keV, representing a 1.4% change in the measured resolution at the endpoint.

This correction for the endpoint energy resolution mismatch also generates a change in the average value of the Michel parameters. The largest change was in ρ with an average value that was 0.00035 higher than the average value measured without the correction for the endpoint resolution mismatch. This change is well within the measured statistical error on the value of ρ. The other Michel parameters also have positive correction to the average value resulting from the correction for the endpoint resolution mismatch. However, the correction on the average value of ξ and δ is less than the RMS of the distribution of values generated by the successive measurements and much less than the measured statistical errors on these parameters.
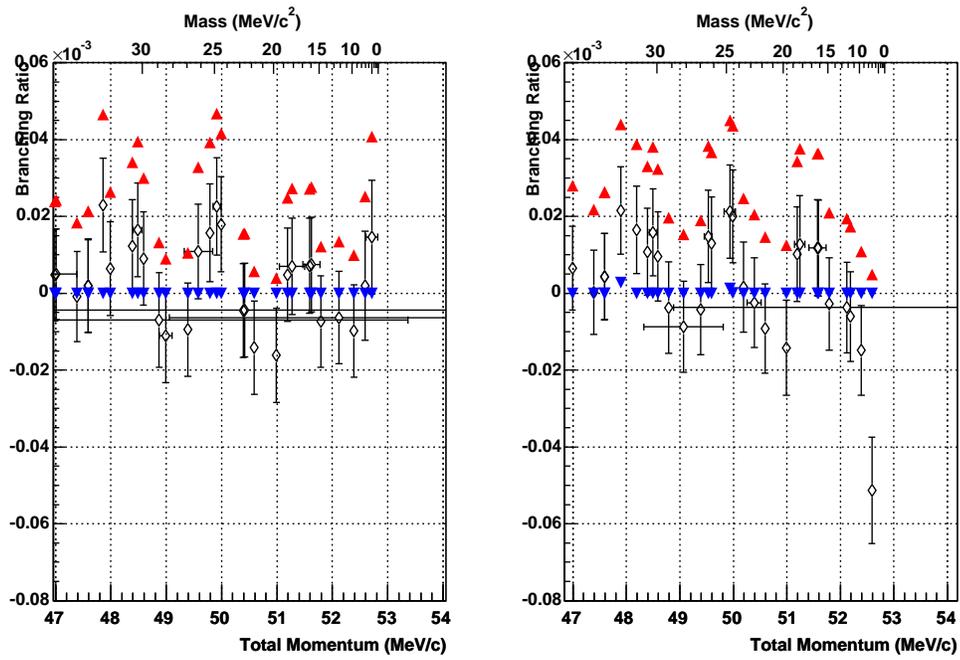
Figure 5.7: Exclusion plot at the kinematic endpoint of the muon decay spectrum. Plot at the left shows the exclusion plot with the correction for the endpoint resolution mismatch. The plot at the right shows the exclusion plot without the correction.

Chapter 6

Conclusions

The data analysed for the TWIST measurement of the Michel parameters can also be used for the measurement of the exotic or forbidden processes. By fitting peaks, using the Michel spectrum as a background, to the data collected by TWIST , I have set model independent 95% upper limits onf the branching ratio of a two body decay. For the purpose of this measurement I focused on the data collected in the 2002 data run period and used for the recently published determination of the Michel parameters [9] [22]. This data consist of four sets, each data set containing $\sim 1.4 \times 10^7$ events after all the data cuts were applied. From each of these data sets I produced a series of branching ratios for the process $\mu \rightarrow e^+ X^0$ given that various masses for the X^0 boson. If all four data sets are included in the estimate the 95% confidence upper limit on the branching ratio is $\mathcal{B}_{eX^0} \leq 2.0 \times 10^{-5} - 5.0 \times 10^{-5}$. The maximum branching ratio appears at higher momenta, near the kinematic endpoint of the Michel spectrum.

I also obtained the branching ratios and accompanying confidence intervals for peaks which have an intrinsic linear anisotropy. This resulted in a more restrictive limit on the branching ratio for the decay for the case using positive anisotropy when compared to the isotropic model. For similar reasons the limit on the branching ratio of a peak showing negative anisotropy is less restrictive than the isotropic case. Using the 2002 data set the 95% upper limit on the positive anisotropic decay between 2×10^{-5} and 3×10^{-5} depending on the momentum, while the 95% upper limit for the branching ratio of the negative anisotropic decay is less than 6.0×10^{-5}.

These measurements compare favourably to previous measurements by Bryman and Clifford and by Balke, of exotic decays involving massive bosons. Although a 95% confidence limit was chosen for this study to limit the number of two body decay candidates, the measured limits are strong enough that they do not suffer by direct comparison to the 90% limits described in these two papers. The 95% upper limit of 1.0×10^{-4} is more than a factor of 6 improvement on the 90% upper limit set by Bryman [3] and a factor of 3 improvement on the 90% upper limit given by Balke [4].

Some caution is required when considering possible zero mass particles because of a mismatch in the energy resolution at the endpoint between the TWIST data and the corresponding Monte Carlo. We can apply a correction to the fit to account for this mismatch. The validity of this correction depends on how the presence of a peak affects the measurement of the energy calibration at the endpoint of the Michel spectrum. However, if the peak is trusted to be small enough that is does not affect the measurement of the energy resolution, then the branching ratio measured for a boson approaching zero mass is 3.4×10^{-5} for the sum of data sets.

It should be pointed out that this is not the best limit now available for the branching ration of a boson with zero mass as the branching ratio set by Jodidio *et.al.* $(\mathcal{B}_{eX^0} \leq 2.6\times10^{-6})$ is a much stronger limit than my limit. This is primarily because the data used in the Jodidio measurement consisted of 1.8×10^7 events collected over a phase space that was much smaller than the fiducial used in this measurement. Specifically, the Jodidio measurement only measures positrons with angles at a very narrow range $(\theta > \arccos(.95))$ downstream of the decay, with an effective momentum range between 37 MeV/c and 68 MeV/c [12]. A particular consequence of this is that the limit set by Jodidio *et.al.* is only sensitive to isotropic peaks and peaks with positive $(\mathcal{P} > 1)$ anisotropy due to their very narrow angular acceptance. The fact that this measurement was only made at the endpoint of the Michel spectrum severely limits the background for an exotic decay, compared to a measurement made away from the endpoint.

The limit on the branching ratio of a particle with zero mass can be used to set limits on the gauge fields described in chapter 2. For the lepton flavour violating interactions involving axions as described by Wilczek [11], this work sets a lower

limit on the symmetry breaking scale; $F_{\mu e} \gtrsim 2.0 \times 10^9$ GeV. Again, a much stronger limit for this quantity was set by Jodidio; $F_{\mu e} \gtrsim 9.9 \times 10^9$ GeV. The upper limit of the cutoff energy scale of the quintessence field can also be calculated. This does constitute an improvement over the previous result as the limit set by Bi *et al* are based on the branching ratio reported by Bryman. Using the numbers in [13], and the limit I set for masses approaching zero, I calculate the cutoff scale $\Lambda \gtrsim 9 \times 10^9$ GeV.

These limits can be improved using available TWIST data. The cuts made on the data for the purpose of this analysis are quite restrictive removing 95% of the data recorded by the experiment to ensure the quality of the remaining data. Some of these cuts may be relaxed to increase the available statistics. This will require a detailed study beyond the scope of this thesis to understand how relaxing any of the current cuts will affect the analysis. A second method of increasing the statistics of this study and improving the limits is to introduce more data sets into the study. The data sets used here represent a small fraction of the data taken in the 2002 running period. The rest of the data was used for the determination of systematic effects. However, if the appropriate Monte Carlo data can be generated, this process can be expanded to include these data sets. There are also data currently available from the 2004 running period that may be used for this determination of the limit on the $\mu \to e^+ X^0$ branching ratio in the near future pending the conclusion of their Michel parameter analyses. Thus, while this thesis does represent a factor three improvement on the branching ratio over previous results, this does not represent the best possible limit on lepton flavour violation available to TWIST.

Bibliography

[1] S. Eidelman, K. Hayes, K. Olive, M. Aguilar-Benitez, C. Amsler, D. Asner, K. Babu, R. Barnett, J. Beringer, P. Burchat, C. Carone, C. Caso, G. Conforto, O. Dahl, G. D'Ambrosio, M. Doser, J. Feng, T. Gherghetta, L. Gibbons, M. Goodman, C. Grab, D. Groom, A. Gurtu, K. Hagiwara, J. Hernández-Rey, K. Hikasa, K. Honscheid, H. Jawahery, C. Kolda, K. Y., M. Mangano, A. Manohar, J. March-Russell, A. Masoni, R. Miquel, K. Mönig, H. Murayama, K. Nakamura, S. Navas, L. Pape, C. Patrignani, A. Piepke, G. Raffelt, M. Roos, M. Tanabashi, J. Terning, N. Törnqvist, T. Trippe, P. Vogel, C. Wohl, R. Workman, W.-M. Yao, P. Zyla, B. Armstrong, P. Gee, G. Harper, K. Lugovsky, S. Lugovsky, V. Lugovsky, and A. Rom, "Review of Particle Physics," *Physics Letters B*, vol. 592, pp. 1+, 2004. [Online]. Available: http://pdg.lbl.gov

[2] R. Eichler *et al.*, "Limits for shortlived neutral particles emitted mu+ or pi+ decay," *Phys. Lett.*, vol. B175, p. 101, 1986.

[3] D. A. Bryman and E. T. H. Clifford, "Exotic muon decay mu → e x," *Phys. Rev. Lett.*, vol. 57, p. 2787, 1986.

[4] B. Balke *et al.*, "Precise measurement of the asymmetry parameter delta in muon decay," *Phys. Rev.*, vol. D37, pp. 587–617, 1988.

[5] A. Masiero, S. K. Vempati, and O. Vives, "Massive neutrinos and flavour violation," *New J. Phys.*, vol. 6, p. 202, 2004.

[6] V. Cirigliano, A. Kurylov, M. J. Ramsey-Musolf, and P. Vogel, "Lepton flavor violation without supersymmetry," *Phys. Rev.*, vol. D70, p. 075007, 2004.

[7] W. Fetscher and H. J. Gerber, "Precision measurements in muon and tau decays," in *Precision Test of the Electroweak Model*, P. Langacker, Ed. World Scientific, 1995, pp. 657–705.

[8] ——, "Muon decay parameters," *Physics Letters B*, vol. 592, pp. 1+, 2004. [Online]. Available: http://pdg.lbl.gov

[9] J. R. Musser, "Measurement of the michel parameter rho in muon decay," *Phys. Rev. Lett.*, vol. 94, p. 101805, 2005.

[10] R. D. Peccei and H. R. Quinn, "Cp conservation in the presence of pseudoparticles," *Phys. Rev. Lett.*, vol. 38, pp. 1440–1443, 1977.

[11] F. Wilczek, "Axions and family symmetry breaking," *Phys. Rev. Lett.*, vol. 49, pp. 1549–1552, 1982.

[12] A. Jodidio *et al.*, "Search for right-handed currents in muon decay," *Phys. Rev.*, vol. D34, p. 1967, 1986.

[13] X.-J. Bi, M.-z. Li, and X.-m. Zhang, "Quintessino as dark matter," *Phys. Rev.*, vol. D69, p. 123521, 2004.

[14] S. M. Carroll, "Quintessence and the rest of the world," *Phys. Rev. Lett.*, vol. 81, pp. 3067–3070, 1998.

[15] B. A. Dobrescu, "Massless gauge bosons other than the photon," *Phys. Rev. Lett.*, vol. 94, p. 151802, 2005.

[16] R. S. Henderson *et al.*, "Precision planar drift chambers and cradle for the twist muon decay spectrometer," *Nucl. Instrum. Meth.*, vol. A548, pp. 306–335, 2005.

[17] B. Jamieson, "Technote 90: Event classification documentation in twist," April 2004, internal TWIST technote.

[18] A. Gaponenko, "A precision measurement of the michel parameter delta," Ph.D. dissertation, University of Alberta, 2005.

[19] R. MacDonald, "Validating the simulation and response function in twist," 2004, presentation at Western Regional Nuclear and Particle Physics Conference.

[20] R. Brun and F. Rademakers, "Root user's guide," 2005, v4.04. [Online]. Available: root.cern.ch/root/doc/RootDoc.html

[21] G. J. Feldman and R. D. Cousins, "A unified approach to the classical statistical analysis of small signals," *Phys. Rev.*, vol. D57, pp. 3873–3889, 1998.

[22] A. Gaponenko *et al.*, "Measurement of the muon decay parameter delta," *Phys. Rev.*, vol. D71, p. 071101, 2005.

[23] J. R. Musser, "Measurement of the michel parameter rho in muon decay," Ph.D. dissertation, Texas A&M University, 2005.

[24] L. Lyons, *Statistics for Nuclear and Particle Physicists*. Cambridge, 1986.

Appendix A

Analytical Estimate of Anisotropic Effects

In this appendix I will quantify the difference in the statistical errors of a mono energetic peak if the spectrum is integrated over angles before the fitting procedure as opposed to maintaining all of the angular information of the spectrum throughout the fitting process. The process of finding peaks within a distribution becomes much easier integrating over angles before implementing a fitting procedure, as the functional peak has no obvious angular behaviour, by construction. However, integrating over angles before the fitting the spectrum fails to discriminate between the sought after isotropic behaviour and say a peak isolated in momentum due to say a very large localised fluctuation in the background.

I consider a simplified model in place of the a mono energetic decay on top of a Michel decay background. For the case where I keep all of the angular information throughout the fit, I consider the χ^2

$$\chi^2_{2D} = \sum_{i=1}^{M} \sum_{j=-N}^{N} \frac{(F(x_i, \cos\theta_j) - G(x_i, \cos\theta_j) - H(x_i, \cos\theta_j))^2}{\sigma^2(x_i, \cos\theta_j)}. \qquad \text{(A.1)}$$

On the other hand, if I integrate over the polar angle before the fit the χ^2 becomes

$$\chi^2_{1D} = \sum_{i=1}^{M} \frac{\left(\sum_{j=-N}^{N} F(x_i, \cos\theta_j) - \sum_{j=-N}^{N} G(x_i, \cos\theta_j) - \sum_{j=-N}^{N} H(x_i, \cos\theta_j) \right)^2}{\sum_{j=-N}^{N} \sigma^2(x_i, \cos\theta_j)}. \quad \text{(A.2)}$$

In both of these expressions for χ^2 the functions $G(x_i, \cos\theta_j)$ and $H(x_i, \cos\theta_j)$ represent the Michel spectrum and the isotropic two body exotic decay spectrum. The function $F(x_i, \cos\theta_j)$, on the other hand represents the experimentally collected data. Note that the sum of bins in $\cos\theta$ are done over 2N bins rather than N bins. This serves as a reminder that $\cos\theta$ is symmetric about zero, and that the summation starts at -1 rather than 0. Thus there are 2N bins in $\cos\theta$ for $\cos\theta$ between -1 and 1, and M bins in x for x between 0 and 1.

Since the actual representations of the decay spectra are complicated, I shall use a simplified model for this exercise. For example a tractable model is,

$$G(x_i, \cos\theta_j) \;\; = \;\; \frac{B}{2NM}(1 - x_i \cos\theta_j) \quad \text{(A.3)}$$

$$H(x_i, \cos\theta_j) \;\; = \;\; \frac{S}{2N}\delta(x_i - \bar{x}). \quad \text{(A.4)}$$

Here I will define the function $\delta(x_i - \bar{x})$ to be equal to one in the momentum bin containing $\bar{x}$, the mean momentum of of the exotic decay, and zero otherwise. The parameters I am changing for the purpose of the fit are the signal amplitude S and the background normalisation B.

One should keep in mind that I am assuming the act of collecting the data $F(x_i, \cos\theta_j)$, to be a Poisson counting process. As such the standard error for a given bin is

$$\sigma(x_i, \cos\theta_j) = \sqrt{F(x_i, \cos\theta_j)}. \quad \text{(A.5)}$$

When I consider the limiting case of $S \rightarrow 0$, this definition will simplify things greatly.

Now I may go through the motions of "fitting" the model. This case is linear in the fit variable, allowing one to make use of a simple least squares fit via the Hessian

[24];

$$H = \begin{pmatrix} \dfrac{\partial^2 \chi^2}{\partial S^2} & \dfrac{\partial^2 \chi^2}{\partial S \partial B} \\ \dfrac{\partial^2 \chi^2}{\partial S \partial B} & \dfrac{\partial^2 \chi^2}{\partial B^2} \end{pmatrix}. \tag{A.6}$$

The inverse of the Hessian generates an error matrix for the fit.

$$H^{-1} = \frac{1}{Det(H)} \begin{pmatrix} \dfrac{\partial^2 \chi^2}{\partial B^2} & -\dfrac{\partial^2 \chi^2}{\partial S \partial B} \\ -\dfrac{\partial^2 \chi^2}{\partial S \partial B} & \dfrac{\partial^2 \chi^2}{\partial S^2} \end{pmatrix}. \tag{A.7}$$

Since I am principally interested in the error on the signal amplitude I can look to the top left matrix element to produce the answer. That is to say,

$$\sigma_S^2 = \frac{\dfrac{\partial^2 \chi^2}{\partial B^2}}{\dfrac{\partial^2 \chi^2}{\partial S^2} \dfrac{\partial^2 \chi^2}{\partial B^2} - \left[\dfrac{\partial^2 \chi^2}{\partial S \partial B} \right]^2} \tag{A.8}$$

If I introduce the notation $f_{ij} = F(x_i, \cos\theta_j)$, $g_{ij} = G(x_i, \cos\theta_j)/B$, and $h_{ij} = H(x_i, \cos\theta_j)/S$, I can succinctly write down general expressions for the elements of the Hessian for the case where all the angular information is maintained.
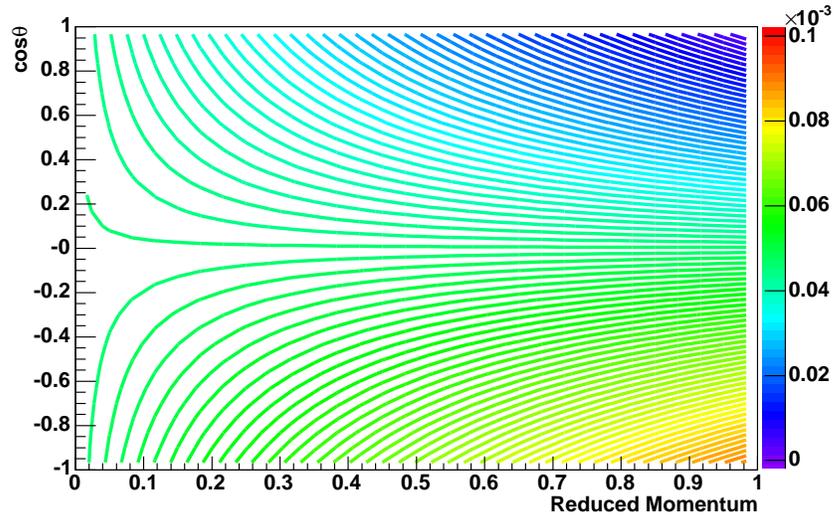

Figure A.1: The toy model discussed for the background function $G(x_i, \cos\theta_j)$. Here I used M = 2750, N = 50, and B = 25

$$\frac{\partial^2 \chi^2}{\partial S^2} = \sum_{ij} \frac{g_{ij}^2}{f_{ij}} \tag{A.9}$$

$$\frac{\partial^2 \chi^2}{\partial B^2} = \sum_{ij} \frac{h_{ij}^2}{f_{ij}} \tag{A.10}$$

$$\frac{\partial^2 \chi^2}{\partial S \partial B} = \sum_{ij} \frac{g_{ij} h_{ij}}{f_{ij}} \tag{A.11}$$

For the simple model described in equations A.3, in the limit of small S, I expect that the background should approximate the "data" when the best fit has been produced. Thus I have a situation that both simulates the expectation of an actual exotic decay search and is analytically tractable. The results of applying such an assumption and approximating the summations with integrations (for large N and M) appear as follows.

$$\frac{\partial^2 \chi^2}{\partial S^2} = \frac{M}{B\bar{x}} \left(\ln(1 + \bar{x}) - \ln(1 - \bar{x}) \right) \tag{A.12}$$

$$\frac{\partial^2 \chi^2}{\partial B^2} = \frac{1}{B} \tag{A.13}$$

$$\frac{\partial^2 \chi^2}{\partial S \partial B} = \frac{1}{B} \tag{A.14}$$

Finally, if I substitute these results into equation A.8, I get the standard error for the signal in terms of the background amplitude.

$$\sigma_S = \sqrt{\frac{B}{\frac{M}{\bar{x}} \ln \frac{1 + \bar{x}}{1 - \bar{x}} - 1}} \tag{A.15}$$

Thus the error of the angle independent case for this model demonstrates a dependence on the mean momentum of the signal which depends strongly on the anisotropic behaviour of the model.

Now, by way of contrast, I consider the angle integrated case. For this case, after integration over $\cos \theta$, the matrix elements become

$$\frac{\partial^2 \chi^2}{\partial S^2} = \frac{M}{B} \tag{A.16}$$

$$\frac{\partial^2 \chi^2}{\partial B^2} = \frac{1}{B} \tag{A.17}$$

$$\frac{\partial^2 \chi^2}{\partial S \partial B} = \frac{1}{B} \tag{A.18}$$

resulting in a standard error of

$$\sigma_S = \sqrt{\frac{B}{M-1}} \tag{A.19}$$

which is of course the error that one should expect to get for a standard deviation of a distribution with uniform weights. This was to be expected since the background distribution is uniform in momentum after integration over $\cos\theta$.

If we compare these two results directly we can see that the method with maintains all of the angular information throughout the fit does indeed produce a smaller expected error, particularly at higher momenta. If I use M=2750 and assume a normalises background the angle integrated case generates an error of 0.019. At zero momentum the case using all of the angular information generates an error of 0.0135.

By changing the model of the signal slightly, I can show the impact of using an anisotropic signal over the anisotropic background. Using the simple anisotropic dependence described in equation 2.12, I will now redefine the model for the idealised signal to be

$$H(x_i, \cos\theta_j) = \frac{S}{2N}\delta(x_i - \bar{x})(1 - \mathcal{P}\cos\theta) \tag{A.20}$$

where $\mathcal{P}$ is between -1 and 1.

Continuing through with the calculation that I have outlined above, I find

$$\frac{\partial^2\chi^2}{\partial S^2} = \frac{M}{B\bar{x}}\left(\ln\left(\frac{1+\bar{x}}{1-\bar{x}}\right)\left(1+\frac{\mathcal{P}}{\bar{x}}\right) - \frac{\mathcal{P}}{\bar{x}}\right) \tag{A.21}$$

which results in a standard deviation of

$$\sigma_S = \sqrt{\frac{B}{\frac{M}{\bar{x}}\left(1+\frac{\mathcal{P}}{\bar{x}}\right)\ln\left(\frac{1+\bar{x}}{1-\bar{x}}\right) - \frac{2M\mathcal{P}}{\bar{x}} - 1}}. \tag{A.22}$$

The standard error on a peak with the described anisotropies are shown as a function of the momentum in figure A.2. This shows that a peak with positive anisotropy will generate a smaller standard deviations at high levels of background anisotropy (with the opposite sign) as compared to the case where the fit peak has no anisotropy. Alternatively, the standard deviation of the negative anisotropic peak on the negative anisotropic background gets worse as the background anisotropy increases. This is much as what one would expect qualitatively.
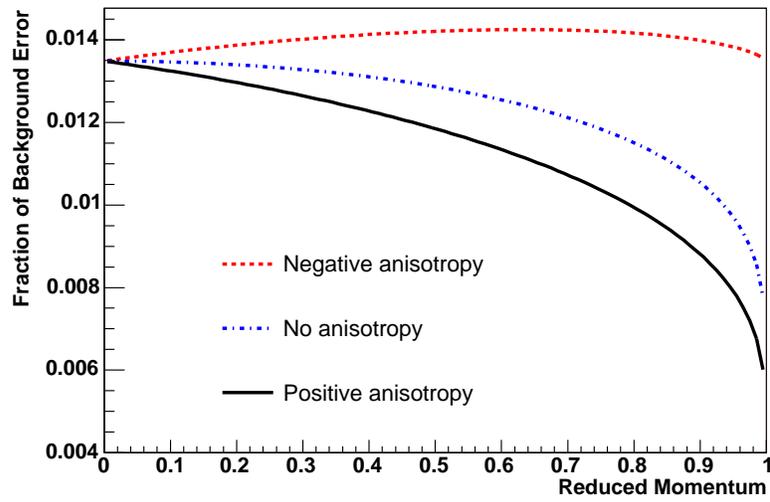
Figure A.2: The standard variance (σ^2) of the signal as a calculated fraction of the background plotted against the reduced momentum. Here we assume that that there are 2750 bins in momentum. Note that the anisotropy of the model is directly proportional to the momentum