

Chapter 4

Analysis

4.1 Overview

The procedure to measure the muon decay parameters is summarised in Fig. 4.1. Very generally, the time and position of signals from the detector's wires were used to identify particles, and then characterise events. Wherever possible, the particle trajectories were reconstructed. Information for every event and track was then written into a database; specifically a ROOT tree structure was used[46]. Track selection and cuts were then applied to construct a spectrum in energy and angle. At this point the data spectrum could be fit with a function that allows for non-standard model muon decay parameters. However, a correction would first have to be made for the acceptance and efficiency. Such corrections can introduce their own biases.

Instead of directly fitting the data spectrum, the analysis used the GEANT 3 software to simulate muons entering the detector and decaying into positrons. The detector's response to the particles was simulated in detail; the implementation of GEANT is described in Chapter 6. The simulation was analysed with the same software and cuts as the real data to produce a simulated spectrum, *except* it was generated with hidden values of the muon decay parameters. The spectra from data and simulation are then fit to each other to obtain the *difference* in P_{μ}^{π} , ξ , ρ and δ . All systematic uncertainties are determined on the *difference* in the decay parameters from their hidden values. Only when all the systematic uncertainties are evaluated are the hidden simulation parameters revealed, allowing the result from data to be determined.

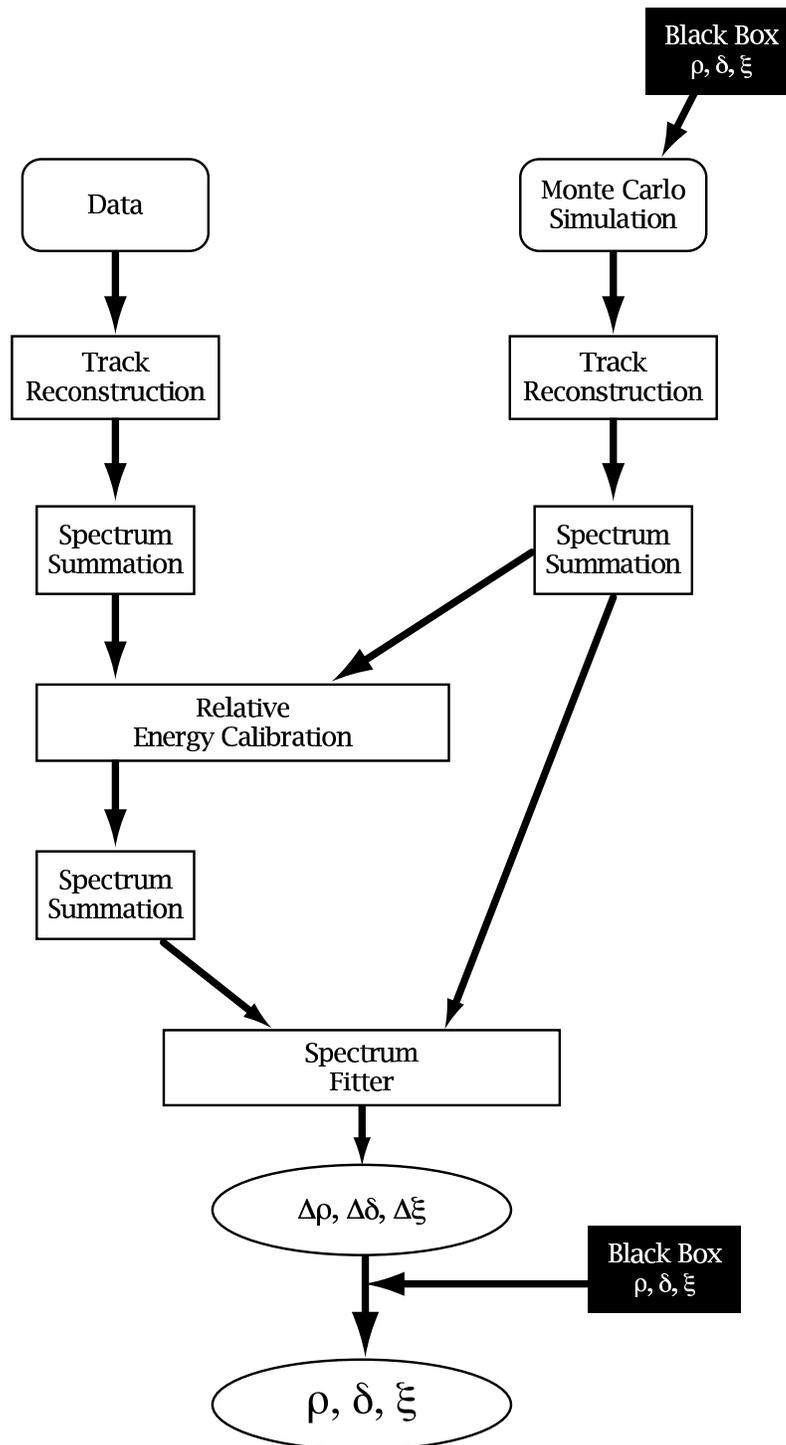


Figure 4.1: Analysis overview. The “Black Box” corresponds to hidden values of the muon decay parameters. This was originally Fig. 5.1 from Ref. [38].

4.2 Track reconstruction

The collaboration used internally written software to identify particles, classify events, and reconstruct trajectories. The code was already operational when the author commenced studies. For the current analysis almost every code module was reviewed and improved where necessary. This section will summarise the important parts of the code, and the changes that were made for this final analysis. The review of the time expansion chamber code appears separately as Chapter 5.

The code described here is processor intensive. Although the data acquisition system could accumulate 2 GB of data in 5 minutes, the subsequent analysis of this data took around 15 hours on a 3 GHz processor. Therefore the track reconstruction was carried out using processors supplied by the WestGrid facility, where the experiment had a continuous allocation of 150 3 GHz CPUs.

4.2.1 Unpacking

The data acquisition system and simulation create data files with the leading and trailing edge times of the signals from the scintillators and wires. These are first “unpacked” according to the following steps:

- Pulses where the time to digital converter reports an error are recovered or discarded as appropriate.
- Wire time offsets (see Section 3.15) are applied, and the leading/trailing edge times are converted to a time and width.
- The trigger scintillator used a “pulse amplitude charge to time converter” (PACT), which converted the integrated charge to a leading/trailing edge. At the unpacking stage, this signal must be converted back to an energy deposit. This is the only place in the electronics where the pulse amplitude/area is used³⁰.

The unpacking code was reviewed for the current analysis, and found to be adequate.

³⁰However this is not the only place where the energy deposit is used. The width of signals from the proportional chambers are used in particle identification, and the two proportional chambers immediately before the target can remove muons that don’t stop in the target, based on their energy deposit. Their operating voltage is lowered to give worse timing resolution, but more sensitive signal widths. In the past a PACT was used on these proportional chambers, but the signal width alone was found to be adequate.

4.2.2 Crosstalk removal

Electronic crosstalk is when a large pulse induces a false smaller pulse on a nearby wire. This can occur inside a drift chamber, or on a pre-amplifier card. Inside the chamber an avalanche at a wire can create photons, and these can undergo a secondary process resulting in ionisation at a second wire. The wires from the detector terminate on pre-amp cards, where they meet in blocks of eight. At this point the wires are close together, and a large pulse can induce smaller pulses on nearby wires. Muons create larger pulses, and therefore the crosstalk is more significant for muons than positrons. In the time expansion chambers, described in Section 3.3, the finer spacing of the sense wires requires guard wires to be placed inbetween, to minimise crosstalk. Their effectiveness is examined more carefully in Chapter 5.

The crosstalk pulses can be identified and removed from software: if a narrow pulse (less than 50 ns wide in a DC, and 60 ns in a PC) occurs within 50 ns of a good pulse, then it is simply removed. In previous analyses, only the 10 (32) wires closest to the good DC (PC) signal were checked for crosstalk. For the current analysis, it was found safe to expand the check to all wires on the same plane. This increased the number of points available to the helix fitting algorithm, which ultimately improved the reconstruction efficiency.

Crosstalk is only present in real electronics; therefore the code to remove crosstalk is disabled when analysing the simulation. This is the only place in the track reconstruction code where the data and simulation are treated differently.

4.2.3 Windowing

A “time window” is started by signals in the proportional chambers. All signals up to $1.05 \mu\text{s}$ after the first PC signal are put into the window. The $1 \mu\text{s}$ is necessary to allow the signals in the drift chamber’s slow drift gas to register at the wires. If additional PC signals occur within $1.05 \mu\text{s}$, a new window is started, and the subsequent signals only appear in this new window. Each window corresponds to a different particle. The procedure is described in more detail in the code author’s (B. Jamieson) thesis[36]. This code was thoroughly reviewed for the current analysis, with only minor changes.

4.2.4 Classification

The identification of the particle in a window uses several pieces of information. The muons are identified from the pulse widths in the upstream proportional chambers. Positrons from the beam line are distinguished from decay positrons since they pass through the entire stack

rather than just one half. Delta electrons are identified, as well as broken trajectories due to a large multiple scatter with detector material, and backscatters from material outside the DC region. Events are then classified according to the particles observed and their time separation. The full list of events can be found in Ref. [36], and was thoroughly reviewed for the current analysis.

4.2.5 Pattern recognition

A helix with centre $(x, y) = (\Delta_x, \Delta_y)$ can be written in the notation

$$\begin{pmatrix} x \\ y \end{pmatrix} = r \begin{pmatrix} \cos(f(z)) \\ \sin(f(z)) \end{pmatrix} + \begin{pmatrix} \Delta_x \\ \Delta_y \end{pmatrix}, \quad (4.1)$$

where (x, y, z) is the position in space, r is the radius, and

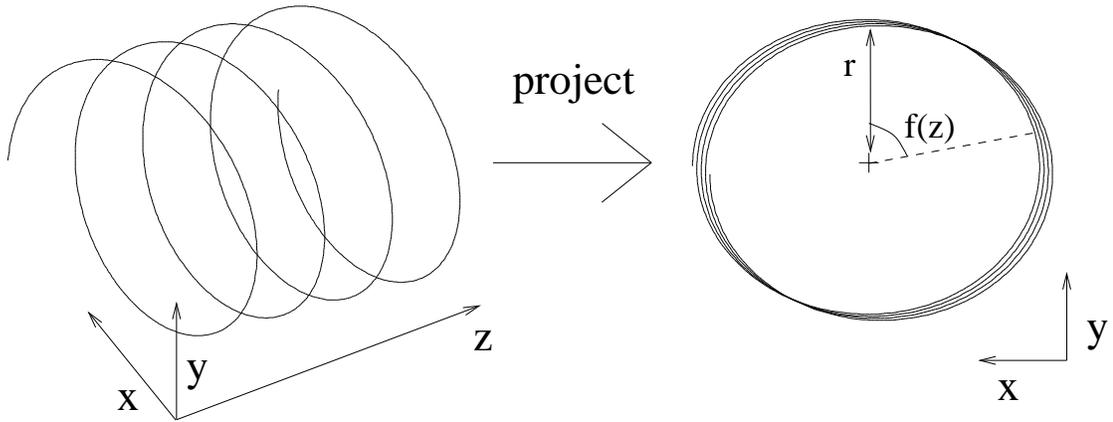
$$f(z) = \frac{2\pi z}{\lambda} + \phi, \quad (4.2)$$

where λ is the wavelength and ϕ is the phase. The pattern recognition algorithm makes a first estimate of the parameters in Eqs. (4.1) and (4.2).

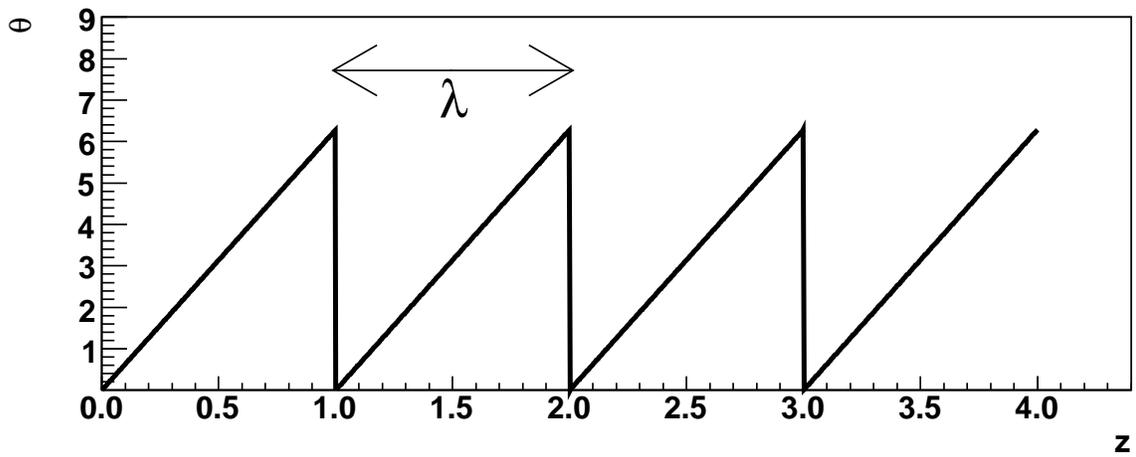
A helix with $\lambda = 1$ and $\phi = 0$ is shown in the left of Fig. 4.2(a). Projecting this helix onto the $x - y$ plane, as shown in the top right of the figure, allows the centre (Δ_x, Δ_y) to be determined, as well as the radius, r . If the angle between each point and the line $x = 0$ is then plotted against z , the wavelength can be determined (see Fig. 4.2(b)). Additionally, ϕ is given by the intercept with the ordinate.

In reality the helix is measured by sampling (x, y) at discrete z locations, which allows the centre and radius to be readily estimated, but there is an ambiguity in determining the wavelength. This is demonstrated in Fig. 4.3(a), where the red circles show the helix phase if it's assumed to be in the range $0 < f(z) < 2\pi$. The other markers show equally possible solutions, displaced by 2π . The figure shows three wavelengths, but mathematically there are an infinite number.

The projection of coordinates from the drift chambers are used in combinations of three to determine the radius. The combination with the smallest χ^2 is kept, and the phase for each combination is determined by the procedure already described. The wavelength degeneracy can be resolved using two additional pieces of information: higher angle tracks will register



(a) Projecting onto the transverse plane.



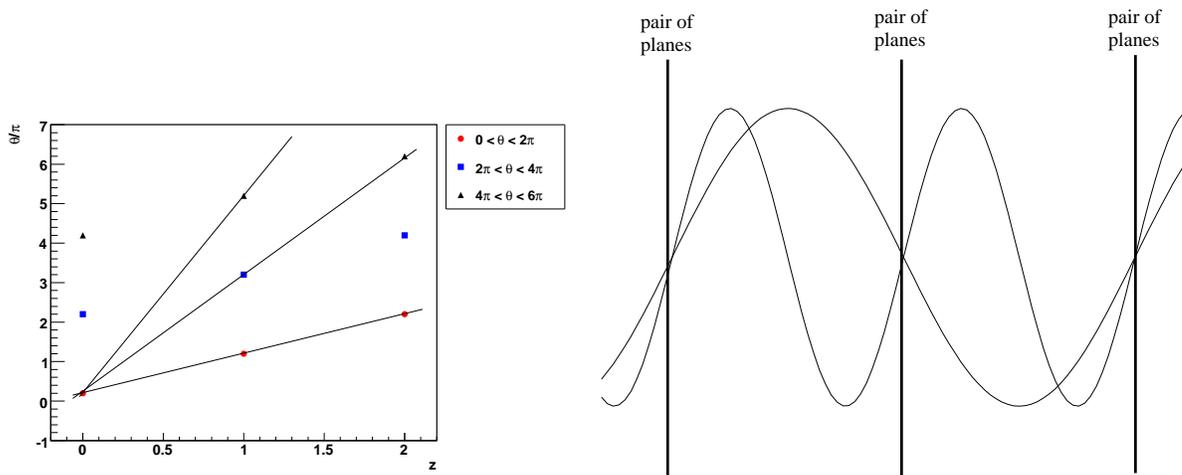
(b) Angle with respect to $x = 0$.

Figure 4.2: Descriptions of a helix.

in more than one drift cell³¹, and tracks cannot be reconstructed if they change by more than 2π between pairs of drift chambers[36]. At this stage there is no estimate made of the positron energy loss, or scatters due to detector material.

The modules in the sparse stacks (see Section 3.8) have been re-arranged since the previous analysis to minimise “magic wavelengths”. This concept is illustrated in Fig. 4.3(b), which shows the worst case scenario where the wavelength is ambiguous. In reality wavelengths close to this situation were poorly reconstructed.

For the current analysis there was an interest in improving reconstruction at larger momentum and smaller angles. Therefore the criteria for keeping candidate points were tuned, and upgraded to depend on track angle, eventually allowing the fiducial to be enlarged. In addition the reconstruction of upstream decays and the analysis software’s ability to separate trajectories were improved.



(a) There is a 2π ambiguity in determining the angle, which leads to an infinite number of possible wavelengths.

(b) The concept of a “magic wavelength” is illustrated. Wavelengths close to the spacing of the chambers in the sparse stack will have poor longitudinal momentum reconstruction.

Figure 4.3: Demonstration of the difficulties in estimating the wavelength of helical trajectories.

³¹Further detail can be found in Section 4.2.1 of Ref. [47].

4.2.6 Fitting the helical trajectory

The pattern recognition in the previous section gives initial estimates to the helix fitting algorithm. The fitter then “swims” helices with parameters close to those determined by the pattern recognition, and iterates using a least squares approach. For the swimming, the trajectory is divided up into segments; for example, a segment could be from one pair of planes to the next. The field is assumed to be uniform over this segment, so that the trajectory is a perfect helix. This approach was validated using the simulation, by comparing the true and reconstructed trajectories.

Initially a fit is made to the positions of the wires, ignoring any drift time information from the chambers. The final iterations then use the drift times³². There is a left-right ambiguity that is resolved iteratively, by choosing the side of the wire that best fits the rest of the trajectory at each iteration. The final fit to drift times includes energy loss and allows for a scatter in the trajectory at pairs of planes.

The helix fitter minimises

$$\chi^2 = \sum_{\text{hits}} \frac{(d_f - d_m)^2}{\sigma_d^2} + \sum_{\text{scatters}} \frac{\theta_s^2}{\sigma_\theta^2}, \quad (4.3)$$

where d_f is the fitted position, d_m is the distance from the space-time relationship, σ_d is the resolution, θ_s is the scatter angle and σ_θ is the theoretical scattering distribution width, which is set to the approximate expression for multiple scattering from the Particle Data Group[48]. The fitting algorithm establishes its own track time, and reports the position and momentum at the first plane that sees the positron, along with a code describing the algorithm’s success. This success code is later used to select good tracks. The helix fitting algorithm is described in more detail in R. MacDonald’s thesis[38]. An example of a clean event is shown in Fig. 4.4. The green hits that stop at the target (most visible on the Z-V figure in the bottom right) are the muon. The blue hits that span the length of the detector correspond to a beam positron. The yellow helix is the reconstructed decay positron trajectory.

Sections of the fitting code were reviewed for the current analysis. A small improvement was made to the algorithm that calculates the scatter angle. The energy loss model used by the fitter was reviewed with the conclusion that bremsstrahlung can be safely neglected, instead assuming all the loss is by ionisation. More detail can be found in R. Bayes’ thesis[49]. There were significant improvements made to the space-time relationship and drift cell resolution, and these will be described later (Sections 4.2.7 and 4.2.8).

³²The proportional chambers are not used in the final helix fit since their resolution is limited by the wire spacing.

Note the algorithm does not reconstruct muons or beam positrons since their transverse momenta (and hence their radii) are very small.

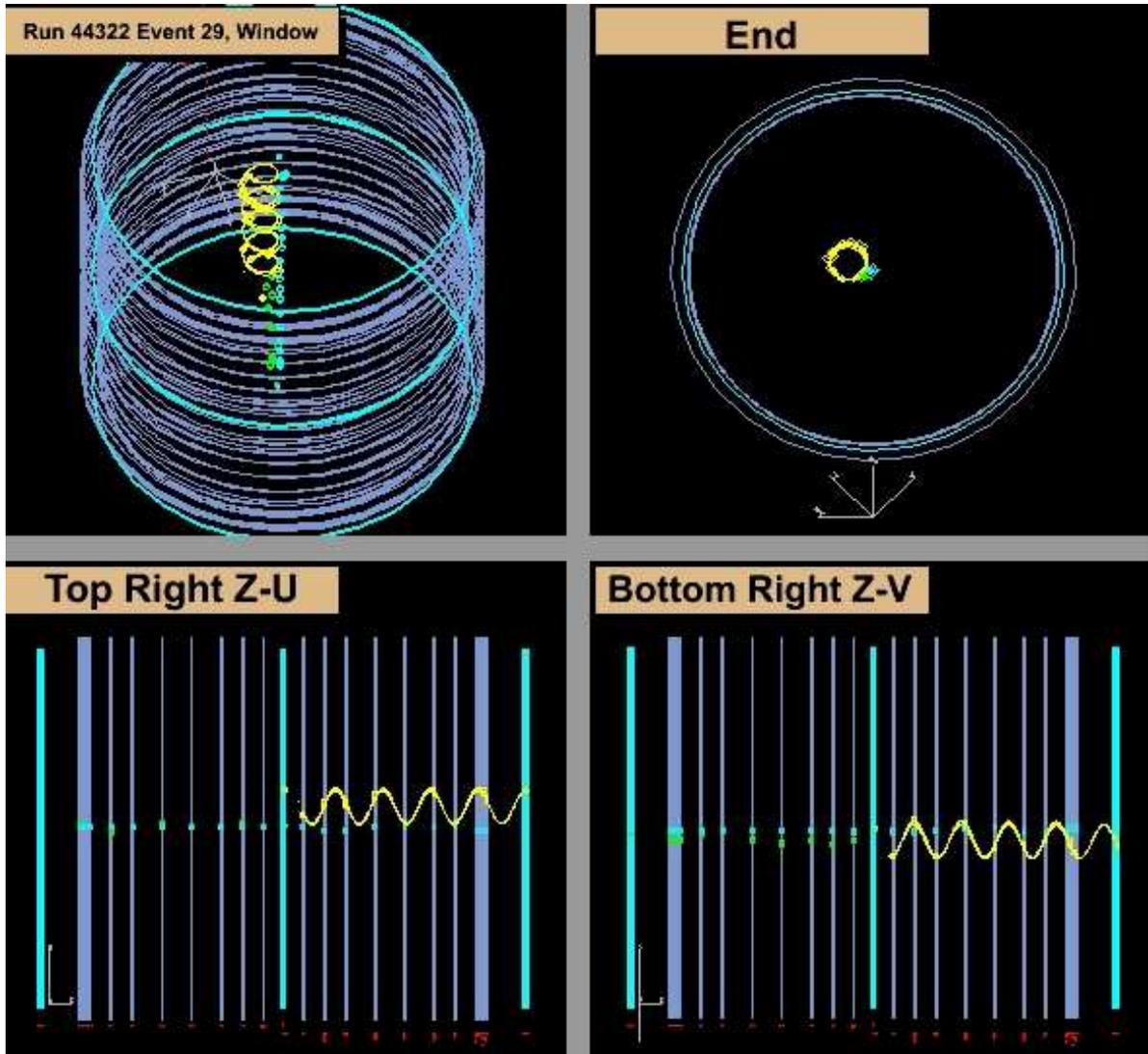


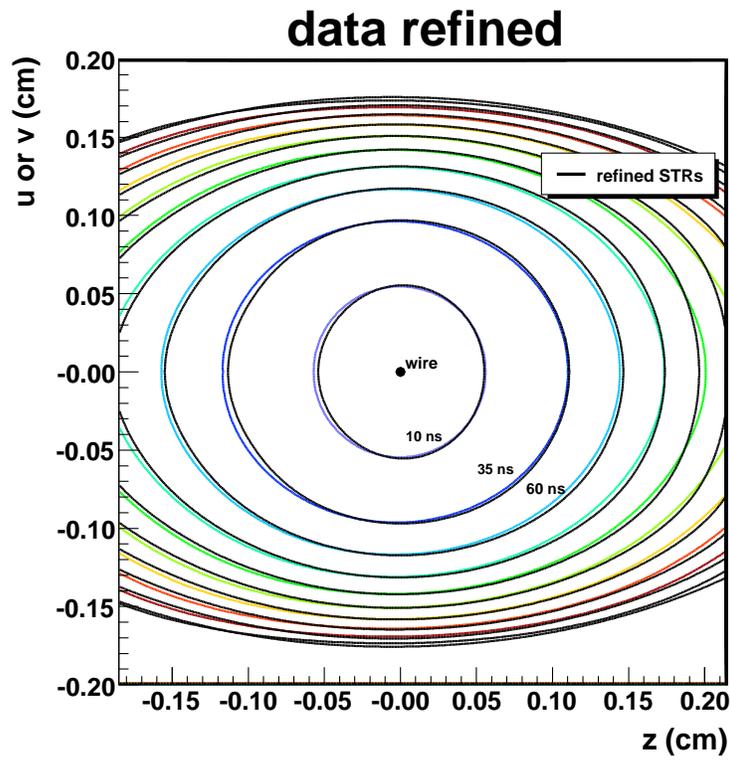
Figure 4.4: A muon (green) stops in the metal foil (centre of the Z-U and Z-V displays) and decays to a positron (yellow). The blue hits that span the detector are a beam positron that passes straight through the detector.

4.2.7 Improved drift chamber space-time-relationship

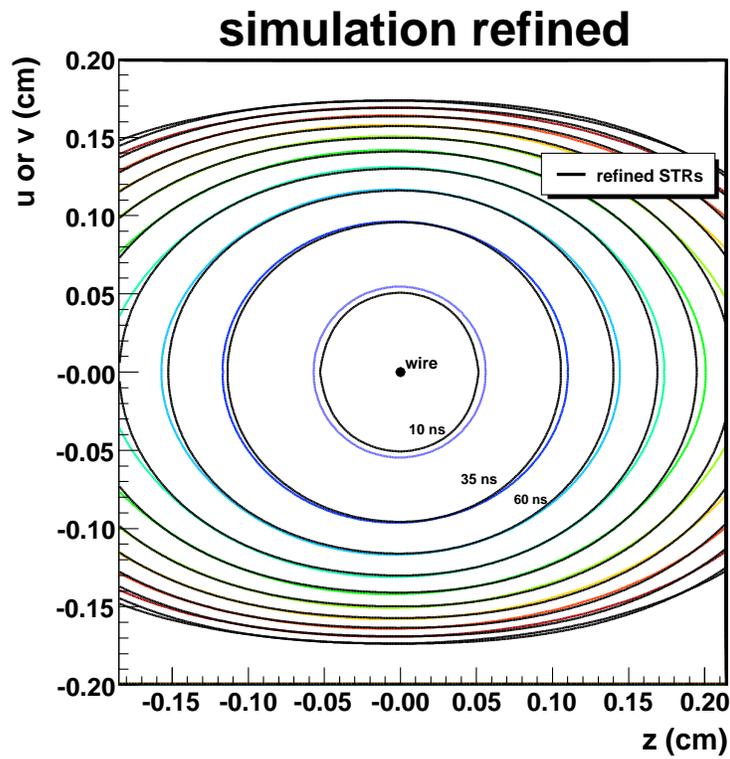
The drift chambers measure the time of arrival of ionisation at a wire, and this is converted into a distance using a drift time map, also known as a space-time-relationship (STR). The STR map for a single drift cell was generated using GARFIELD[45] at intervals of $20\ \mu\text{m}$. For the current analysis the STRs were refined by analysing events using the experiment's helix fitting code, and adjusting the STR at each space point to minimise the average drift time residual. This was performed iteratively, and allowed the STRs to correct for biases in the helix fitting software, and for uncertainties in the GARFIELD inputs such as voltage, gas densities and geometry. Note the previous analysis for $P_{\mu}^{\pi}\xi$ used GARFIELD STRs. The most recent ρ and δ analyses did not use refined STRs during analysis since they were not initially available, but *did* make a correction for them at the end of the analysis.

The refined drift cell isochrones are shown in Fig. 4.5(a), along with the GARFIELD STRs. The significant differences are close to the wire, and in the corners of the cells. The refined STRs reduce the average χ^2 of the positron helices, and improve the momentum bias and resolution. The method was extended to make refined STRs that were dependent on drift plane, correcting for small variations in plane assembly such as wire placement and cathode-to-cathode differences. Although temperature gradients exist in the detector, the STRs for each plane are dominated by mechanical variations, not by temperature. For each drift plane there was just one cell specified; STRs depending on the region of the plane were investigated and found to be unwarranted.

Refined STRs were also produced from the simulation, and these are shown in Fig. 4.5(b). This was done independently from the data, and in this case any differences from GARFIELD must be due to biases in the helix reconstruction software. The simulation was then analysed using these refined STRs, allowing the data and simulation to be treated in the same way; in other words, for both cases the reconstruction biases were absorbed into the STRs. For convenience, the analysis chain is described schematically in Fig. 4.6.



(a) STRs refined from helix fits to data.



(b) STRs refined from helix fits to simulation.

Figure 4.5: Comparison of GARFIELD and refined STRs.

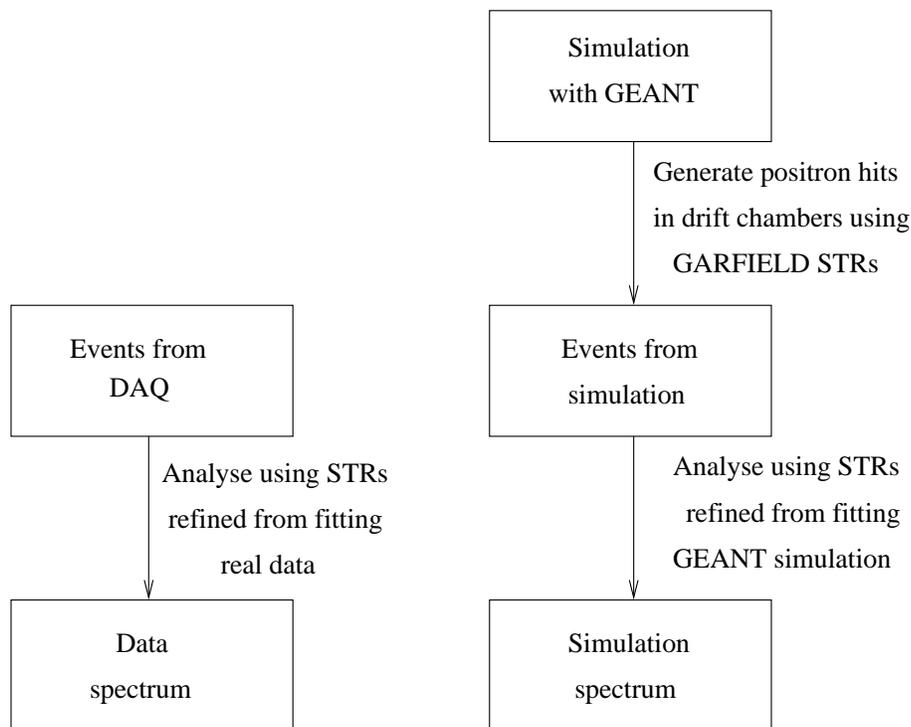


Figure 4.6: Analysis chain for using refined STRs.

4.2.8 Improved resolution function

The helix fitting routine must assign a weight to each point (σ_d in Eq. (4.3)). In the past a constant has been used, $\sigma_d = 100 \mu\text{m}$. A study using the simulation trialed many reasonable functions, and concluded a better choice is the one shown in Fig. 4.7. This is an “effective resolution”, which incorporates the effects of bias in the fitting procedure.

Figure 4.7 shows a degradation of resolution at longer times; this is due to diffusion spreading out ionisation clusters, resulting in several smaller pulses. The baseline in the figure is expected from timing resolution. The resolution from ionisation statistics degrades closer to the wire, but Fig. 4.7 shows that the *effective* resolution is nearly constant. This is because the “left-right” ambiguity dominates the resolution from ionisation statistics; specifically, the wires only record time, and this alone cannot determine which side of the wire the particle went past. The ambiguity is resolved by iterating the fit, keeping the side that is compatible with the rest of the trajectory. In the first iteration, ionisation with distance (x) closer to the wire than 0.1 cm is assigned $\sigma_d = 2x + \sigma(x)$, where σ is taken from Fig. 4.7. In subsequent iterations, as the ambiguity is resolved, this weighting only occurs for signals

closer than 0.05 cm, 0.025 cm etc. The dominance of the left-right ambiguity results in almost no sensitivity to the choice of $\sigma(x)$ below 0.1 cm.

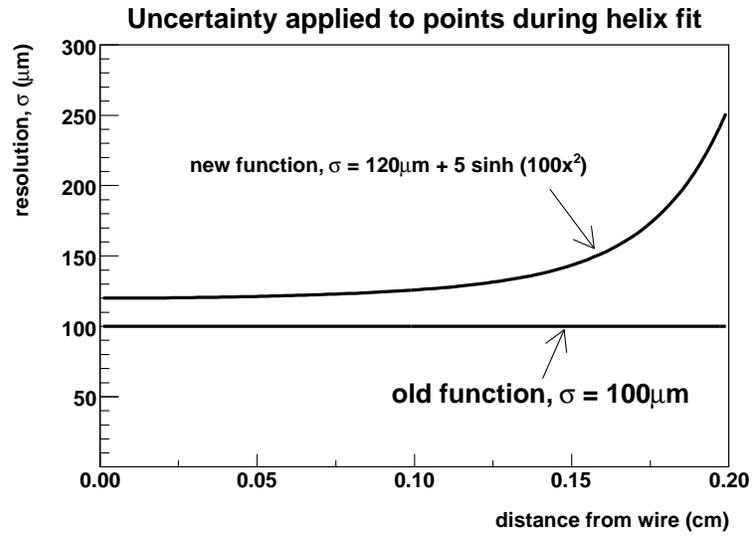


Figure 4.7: Resolution function used during the helix fitting.

4.3 Track selection and cuts

The fit details for each trajectory were placed into a database. Information on the trigger particle, such as its initial energy deposit and final position, was also included in the database. A separate analysis then selected decay positrons from good events, applied cuts, and constructed a decay spectrum in $(p, \cos \theta)$. A separate finely binned spectrum close to the endpoint ($p = 52.8 \text{ MeV}/c$) was also constructed for the purposes of momentum calibration. The selections and cuts have undergone review to ensure they are unbiased. The effect of each selection and cut is shown in Fig. 4.8, where 7% of the data and 13% of the simulation events enter the fiducial. Each of these selections and cuts will now be briefly described.

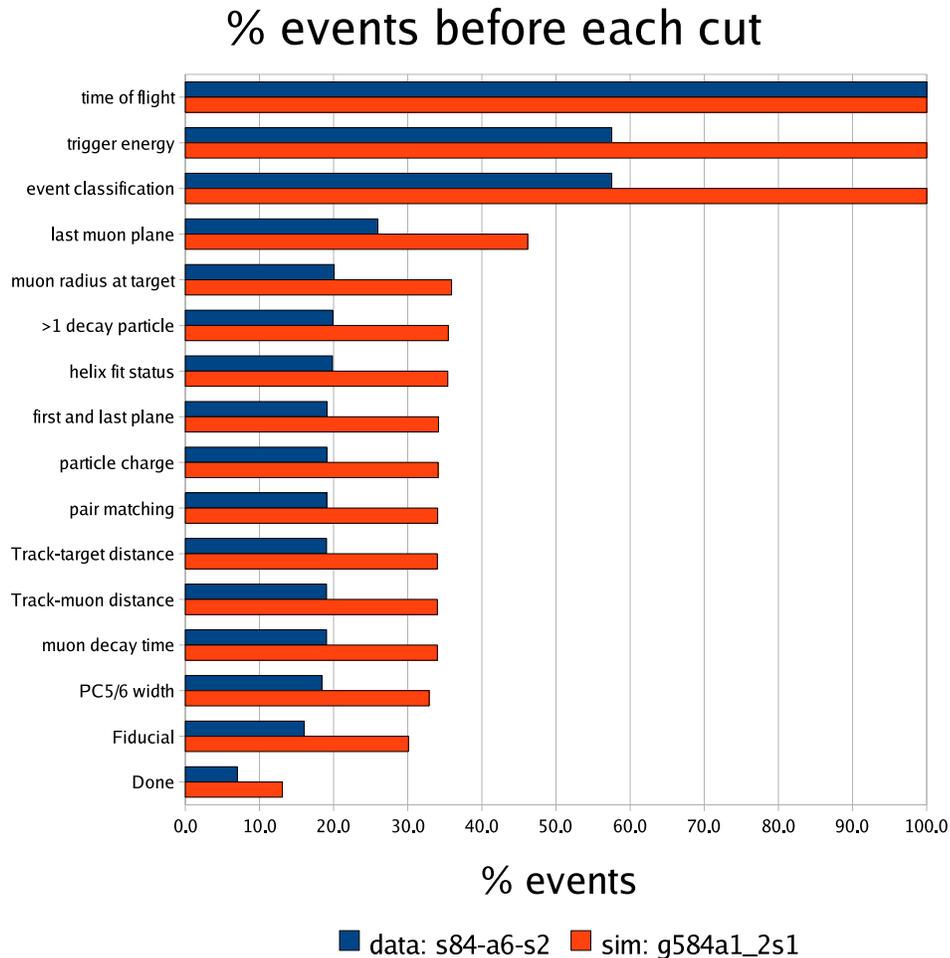


Figure 4.8: Fraction of events before each cut is applied. The yield from the data is lower since a time of flight cut is applied to remove “cloud muons” and pions.

4.3.1 Quality of data

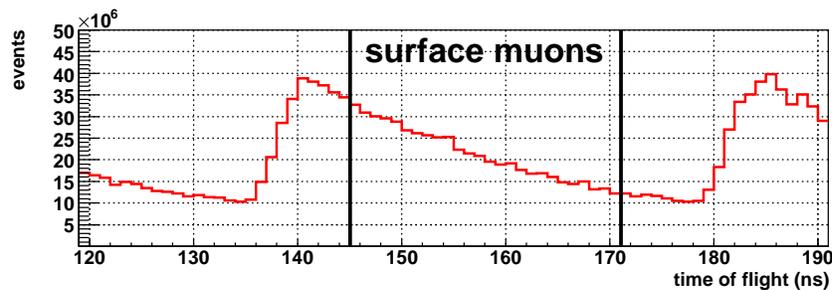
Extensive quality checks eliminated 14% of the data before any track selection or cuts were made; see Section 8.6.

4.3.2 Trigger cuts

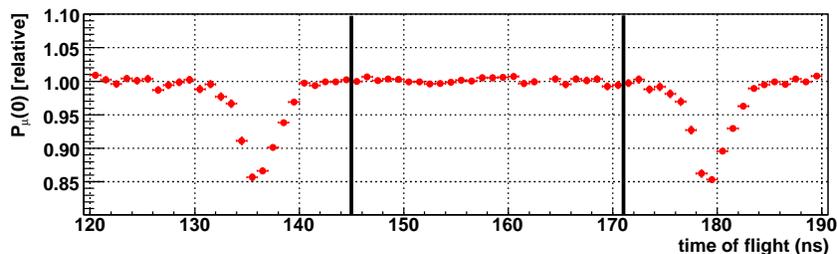
Time of flight

The time of flight through the M13 beamline was recorded. This is the difference in time between a capacitive probe at the proton beam, and the trigger scintillator. As already described in Section 3.2.5, heavier pions took longer to traverse the channel, and “cloud muons” with low polarisation were prompt. Therefore a cut on the time of flight eliminated these particles, leaving only high polarisation muons.

The cut was tuned by observing the asymmetry in the data, as shown in Fig. 4.9, which also shows the conservative setting adopted. This cut was not applied to the simulation since only surface muons and positrons are generated³³.



(a) Number of events for each time of flight.



(b) Relative asymmetry for each time of flight.

Figure 4.9: Tuning the time of flight cut to select surface muons.

³³The capability to simulate proton bunches on the target and the M13 beam line exists, but there is little motivation since the experiment has a high precision muon beam measurement at the end of the channel.

Energy deposit in trigger

Muons were earlier identified using their pulse widths in the four upstream PCs. The energy deposited by the trigger particle in the trigger scintillator is also available, but up to this stage has not been used. A possible cut was investigated to eliminate beam positrons with enough energy to cause a trigger, but was found to be unnecessary since the later event classification cut already removes such triggers. Therefore a cut on energy deposit was not applied in data or simulation.

4.3.3 Event classification cut

The classes of events selected at this stage are shown in Table 4.1. Nearly 90% of events that subsequently pass the remaining cuts are of the simplest type. The remaining $\approx 10\%$ have a beam positron that doesn't interfere with the decay positron's reconstruction. Events with multiple muons have been removed. Note the event classification cut includes an implicit time cut at $1.05 \mu\text{s}$, which removes 38% of events due to the muon's lifetime.

Table 4.1: Fraction of event types in the fiducial.

Event number ^a	Description	% of fiducial events	
		data	simulation
1	μ^+ and decay- e^+ , separated by $> 1.05 \mu\text{s}$	88	87
2	μ^+ , decay- e^+ , beam- e^+ (s), all separated by $1.05 \mu\text{s}$	11	12
6	As (1), with delta- e^- removed	0.6	0.5
7	As (2), with delta- e^- removed	< 0.1	< 0.1
10	As (1), decay- e^+ scattered	< 0.1	< 0.1
11	As (2), decay- e^+ scattered	< 0.1	< 0.1
21	As (1), but beam- e^+ overlapping μ^+ or decay- e^+ in PCs ^b	< 0.1	0.2
22	As (2), but beam- e^+ overlapping any particle in PCs	< 0.1	< 0.1

^a This is an event identifier used internally by the collaboration. See the appendix of Ref. [38] for more detail.

^b Events where the beam- e^+ overlap in the DCs are removed, since DC hits are used to reconstruct the decay- e^+ .

4.3.4 Muon cuts

Last plane

The muon's last hit must be in the PC immediately before the stopping target.

Position at target

Figure 3.21 showed that a cut is needed on the final position of the muon to ensure they stop in the metal rather than surrounding foils. A cut is also needed to keep decay positrons with high transverse momentum inside the tracking region; otherwise they can strike external material (e.g. glass frames) and scatter. A radial cut is made so that

$$r = \sqrt{u_{\text{PC5}}^2 + v_{\text{PC6}}^2} < 2.5 \text{ cm}, \quad (4.4)$$

where $(u_{\text{PC5}}, v_{\text{PC6}})$ are the wire centres in the PCs immediately before the target.

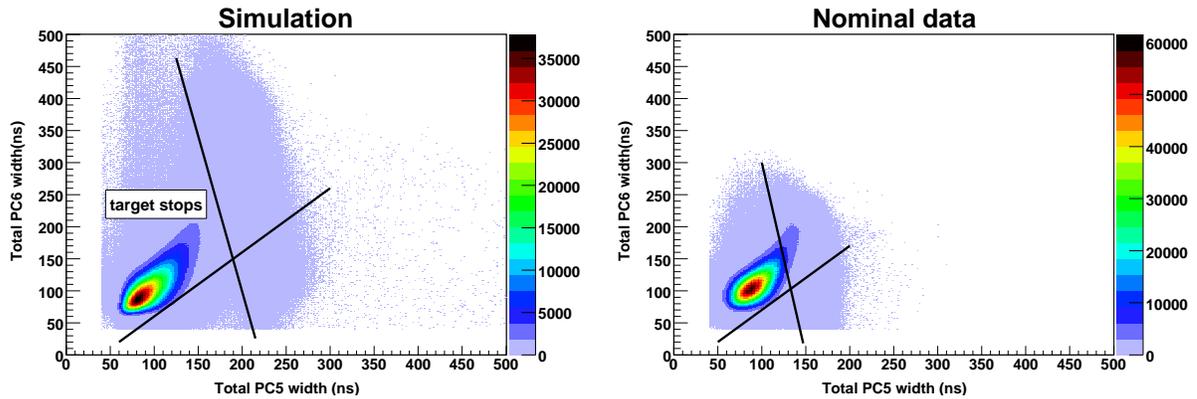
Pulse width at target

The current analysis used a new cut on the pulse widths in the PCs immediately before the target (PC5 and PC6). The voltage on these PCs was deliberately lowered while acquiring data, which increased the sensitivity to muon energy deposit but lowered positron efficiency. The pulse widths indicate the muon energy deposit, which can be used to reject muons stopping in the PC gas rather than the metal target.

Figure 4.10 shows the cut position in simulation and data. The cut must be applied to both data and simulation since it affects the average positron energy loss in the target. The data cut in Fig. 4.10(b) is set very conservatively. Separate cuts were made for multiple hit combinations in each PC.

4.3.5 Decay particle trajectories

The event classification cut allows for events with more than one candidate decay trajectory (see Section 4.3.3). The multiple tracks can originate from a hard scatter (i.e. a large angle multiple scatter which breaks the trajectory to the point where it is reconstructed as two tracks), delta-electrons, decay positrons that scatter off material and cause a second track, or a real beam positron that has survived earlier selections. This section describes the selections and cuts that result in a single decay trajectory.



(a) PC5 and PC6 widths in simulation.

(b) PC5 and PC6 widths in data.

Figure 4.10: Pulse widths in the PCs immediately before the target. Muons that stop in the target can be separated from those stopping in the gas of the final PC. The simulation PC response has not been tuned.

Number of tracks

There must be at least one decay trajectory in the event.

Helix fit status

The helix fitting algorithm reports a code to describe how successful it was in fitting a track. Only completely successful fits are used.

First and last plane

A track must be fully contained in either the upstream or downstream half of the detector. Tracks must not cross the target module. There is no cut on the number of planes crossed.

Particle charge

Only particles with positive charge are accepted, as determined by their direction of winding. There are genuinely negative particles in the detector; principally these are delta-electrons, but there are also small contributions from bremsstrahlung (the γ can undergo pair production), and the process $\mu^+ \rightarrow e^- \bar{\nu}_e \nu_\mu e^+ e^-$. The analysis can incorrectly identify particles as negative; for example a beam positron that only appears in the upstream detector half due to

a scatter in the stopping target can look like an upstream-going decay particle with negative charge.

Pair matching

Pairs of track candidates were extrapolated to the stopping target to determine their closest distance of approach. If they appear to have a common origin, this can indicate a beam positron or delta-electron. The selection has a negligible effect (between 10 and 100 out of 10^6 events are rejected).

Target-positron distance

The track closest to the target is selected.

Muon-positron distance

The experiment has always selected the positron trajectory closest to the muon. For the current analysis, a cut was also made on the distance between the muon and positron, extrapolated to the stopping target. The motivation is that high angle tracks produce more signals in the drift chambers; the reconstruction procedure can choose the wrong signals, and reconstruct the positron as an incorrect low angle track. By comparing the positron trajectory at the target with the last known position of the muon, the incorrect tracks can be reduced. This cut must be applied carefully, since it can remove events with an angle bias.

Decay time

If a muon decays in less than $1.05 \mu\text{s}$, and the decay positron is upstream, then the drift chambers can still receive late ionisation from the muon, reducing the efficiency of upstream positron reconstruction. The event classification selection has already made an implicit time cut on the decay positron at $1.05 \mu\text{s}$. However, this is determined from the PC times, which have a resolution of $\approx 20 \text{ ns}$. An additional cut is now made at $1.05 \mu\text{s}$, using the time determined while fitting the helix, which has improved resolution. There is also a cut made at $9.0 \mu\text{s}$, since this allows $1 \mu\text{s}$ of decay positron information to be analysed up to the data acquisition limit of $10 \mu\text{s}$.

4.3.6 Fiducial region

The helix fitting routine determines the radius and wavelength of each track, from which the transverse momentum is given by

$$p_t [\text{MeV}/c] = 300 \times B [\text{T}] \times r [\text{m}], \quad (4.5)$$

and the longitudinal momentum is given by

$$p_z [\text{MeV}/c] = \frac{300}{2\pi} \times B [\text{T}] \times \lambda [\text{m}]. \quad (4.6)$$

After all the selections and cuts from the previous section, the decay spectrum is reconstructed, which is shown alongside the theoretical spectrum in Fig. 4.11.

The region of $(p, \cos\theta)$ for spectrum fitting (the fiducial) is then chosen to minimise bias and inefficiency, maximise resolution, and ensure any deficiencies are well matched between data and simulation. The fiducial selection is shown in Fig. 4.11(c), and will now be described. Note the boundaries are assigned based on the simulation, and will therefore be described again in Section 6.8. $P_\mu^\pi \xi$ is dominated by systematic uncertainties, so that expanding the fiducial cuts does not help.

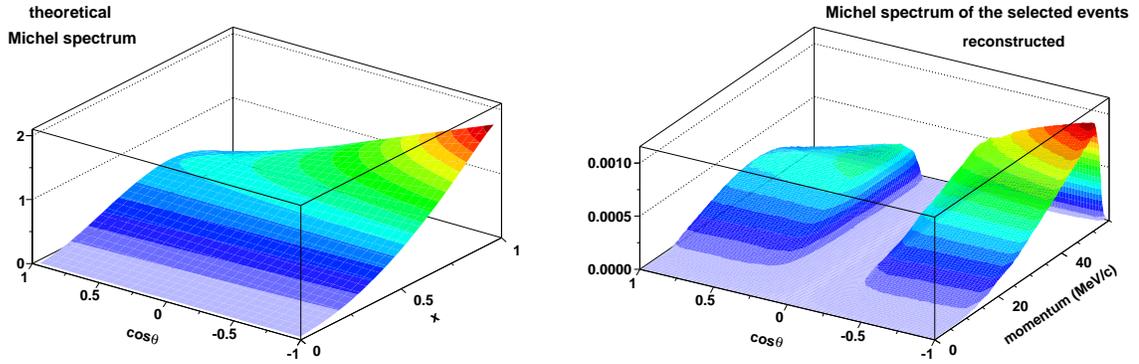
Total momentum, $p < 51.5 \text{ MeV}/c$

There are three reasons for excluding $p > 51.5 \text{ MeV}/c$:

1. At the kinematic endpoint ($p = 52.8 \text{ MeV}/c$) the spectrum is most sensitive to radiative corrections. The exclusion of this region limits the experiment's uncertainty due to radiative corrections.
2. The momentum resolution degrades at higher momenta.
3. The momentum calibration of the spectrum uses the endpoint region ($(52.3 < p < 53.4) \text{ MeV}/c$); in order to be conservative, this region is then excluded from the decay parameter extraction.

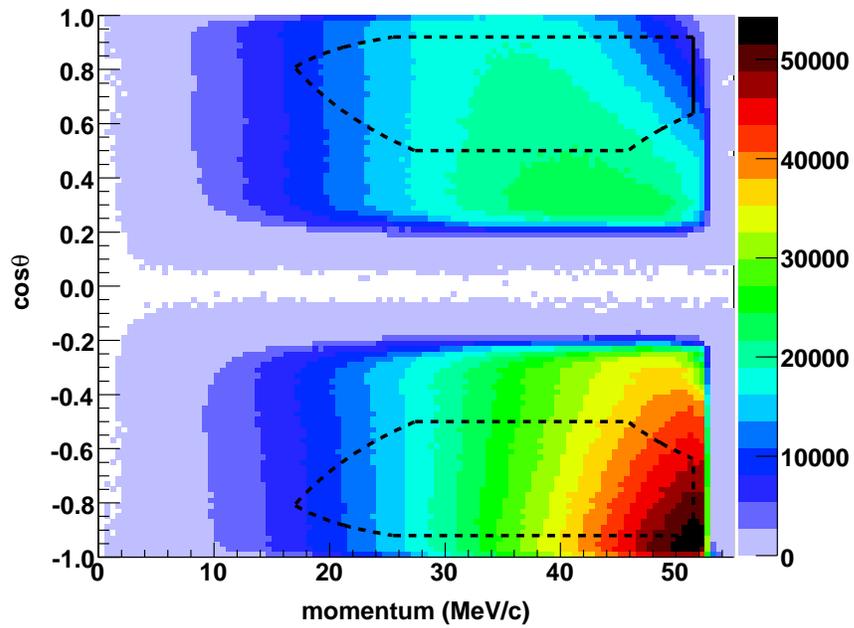
Longitudinal momentum, $|p_z| > 13.7 \text{ MeV}/c$

The sparse stack prior to this $P_\mu^\pi \xi$ measurement had a periodicity of 12.4 cm, leading to a minimum longitudinal wavelength requirement. For the current measurement the periodicity was decreased. However, since statistical uncertainties are not dominant, the longitudinal



(a) Theoretical spectrum.

(b) Reconstructed spectrum.



(c) Reconstructed spectrum with fiducial boundaries.

Figure 4.11: Comparison of theoretical and reconstructed muon decay spectra. The lower plot shows the fiducial region.

momentum cut was unchanged. The systematic uncertainty due to η correlations is also sensitive to this cut.

Angle, $0.50 < |\cos \theta| < 0.92$

At low angles (larger $|\cos \theta|$) the wavelength is poorly resolved, and the helix fitting algorithm fails. At higher angles (smaller $|\cos \theta|$) the reconstruction becomes unreliable due to increased multiple scattering, and there is minimal improvement in $P_\mu^\pi \xi$ sensitivity by including higher angles.

Transverse momentum, $(10.0 < p_t < 39.7) \text{ MeV}/c$

The maximum transverse momentum cut is coupled to the target radius cut of 2.5 cm; together these cuts keep the positron trajectory within the instrumented chambers.

There is a minimum transverse momentum cut to avoid a problem at small angles where the track radius becomes comparable to the wire spacing.

4.4 Extraction of muon decay parameters

Once the data spectrum in $(p, \cos \theta)$ is reconstructed, it could be fit directly by correcting for inefficiencies, biases, and resolution issues. However in a high precision experiment these spectrum corrections are challenging to evaluate. Instead the μ^+ data³⁴ is compared to a simulation that is subject to the same analysis as the data. The advantage is first order cancellation of biases and inefficiencies, since both the data and simulation are analysed with the same reconstruction software and cuts. The degree to which the simulation matches the data therefore dominates the final systematic uncertainty.

The μ^+ spectrum is linear in the muon decay parameters. Recall the expression for the differential decay rate,

$$\frac{d^2\Gamma}{dx d\cos\theta} = k(x) \{F_{IS}(x) + P_\mu \cos\theta F_{AS}(x)\}, \quad (4.7)$$

³⁴For the analysis of the μ^- data acquired in 2007, an “unfolding” approach is being used to correct for reconstruction inefficiencies and biases. The analysis of this data does not require the same precision as the μ^+ results.

where,

$$\begin{aligned}
 k(x) &= \frac{m_\mu}{4\pi^3} W_{e\mu}^4 G_F^2 \sqrt{x^2 - x_0^2} \\
 F_{IS}(x) &= x(1-x) + \frac{2}{9}\rho(4x^2 - 3x - x_0^2) + \eta x_0(1-x) + F_{IS}^{RC}(x), \\
 F_{AS}(x) &= \frac{1}{3}\xi\sqrt{x^2 - x_0^2} \left[1 - x + \frac{2}{3}\delta \left(4x - 3 + \left(\sqrt{1 - x_0^2} - 1 \right) \right) \right] \\
 &\quad + F_{AS}^{RC}(x),
 \end{aligned} \tag{4.8}$$

where the terms were defined in Section 2.1. The expression is linear in³⁵ $P_\mu^\pi \xi$, $P_\mu^\pi \xi \delta$, ρ and η . This allows exact derivatives of the spectrum to be constructed,

$$\begin{aligned}
 S &= \frac{d^2\Gamma}{dx d\cos\theta} \\
 \left. \frac{\partial S}{\partial P_\mu^\pi \xi} \right|_{P_\mu^\pi \xi \delta, \rho, \eta} &= k(x) \cdot \frac{2}{9}(4x^2 - 3x - x_0^2), \\
 \left. \frac{\partial S}{\partial P_\mu^\pi \xi \delta} \right|_{P_\mu^\pi \xi, \rho, \eta} &= k(x) \cdot x_0(1-x), \\
 \left. \frac{\partial S}{\partial \rho} \right|_{P_\mu^\pi \xi, P_\mu^\pi \xi \delta, \eta} &= k(x) \cdot \frac{1}{3} \cos\theta \sqrt{x^2 - x_0^2} (1-x), \\
 \left. \frac{\partial S}{\partial \eta} \right|_{P_\mu^\pi \xi, P_\mu^\pi \xi \delta, \rho} &= k(x) \cdot \frac{2}{9} \sqrt{x^2 - x_0^2} (4x - 3 + \sqrt{1 - x_0^2}),
 \end{aligned} \tag{4.9}$$

These derivatives are independent of the muon decay parameters, and the shape of the functions are shown in Fig. 4.12. The *difference* between the data and simulation spectrum can be used with the derivatives to say how much of each muon decay parameter is needed to make the spectra agree. This procedure is described more thoroughly in Chapter 7 of A. Gaponenko thesis[41], where tests are carried out that show the uncertainties on the decay parameters scale with the square root of available statistics, and the biases for $P_\mu^\pi \xi$ from the procedure are at the 10^{-6} level.

In practice the simulation is *not* generated with the standard model values of the muon decay parameters. Instead hidden values within 1% of the standard model are used for ρ , δ

³⁵Note the spectrum is only linear in $P_\mu^\pi \xi$ if $P_\mu^\pi \xi \delta$ is held constant, and is only linear in $P_\mu^\pi \xi \delta$ if $P_\mu^\pi \xi$ is held constant.

and $P_\mu^\pi \xi$, and η is fixed to the world average³⁶. The hidden values are only revealed after all systematic uncertainties are evaluated on the *difference* between the data and simulation spectra. Since the result from data is unknown to the experimenter, this removes a lot of bias in tuning cuts, rejecting anomalous data, and evaluation of systematic uncertainties.

The technique requires six spectra to be reconstructed with the same analysis software and cuts: data, simulation and one spectra for each of the four derivatives in Eq. (4.9). Since the derivative spectra have negative regions (they are, after all, just a spectrum shape), a positive value is used to generate the $(E, \cos \theta)$ for a decay positron, but the sign of the spectrum is passed to the analysis software, which applies it when reconstructing the derivative spectra.

The $P_\mu^\pi \xi$ result is quoted at the time of muon production, but the data and simulation spectra are constructed with P_μ at the time of decay, after all depolarisation processes have taken place. Algebraically, the difference in the product $P_\mu \xi$ between the data and simulation spectra can be written,

$$\Delta [P_\mu \xi] = P_\mu^{D,data} \xi^{data} - P_\mu^{D,sim} \xi^{sim}, \quad (4.10)$$

where the superscript D in P_μ^D implies that P_μ is at the time of decay, and *sim* refers to the simulation's values. $\Delta [P_\mu \xi]$ is the quantity that uncertainties are assigned to. When the experiment reveals the hidden value of ξ used in the simulation, the result is then stated as,

$$P_\mu^{\pi,data} \xi^{data} = P_\mu^{\pi,sim} \xi^{sim} + \Delta [P_\mu \xi]. \quad (4.11)$$

where the superscript π in P_μ^π implies that P_μ is at the time of production. The experiment claims to accurately simulate the depolarisation between production and decay, which can be written as,

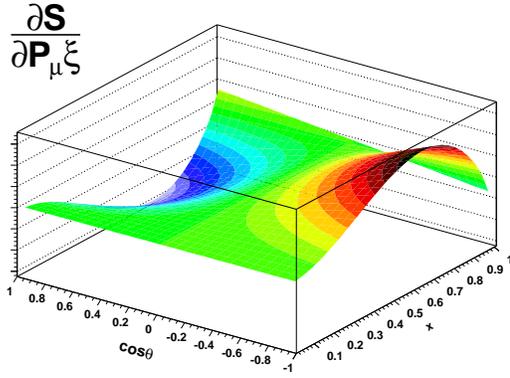
$$P_\mu^{\pi,data} - P_\mu^{D,data} = P_\mu^{\pi,sim} - P_\mu^{D,sim}. \quad (4.12)$$

Equations (4.10) and (4.12) can be re-arranged to yield,

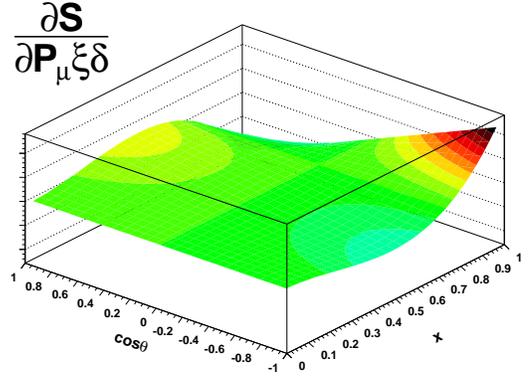
$$\begin{aligned} \Delta [P_\mu \xi] &= P_\mu^{\pi,data} \xi^{data} - P_\mu^{\pi,sim} \xi^{sim} \\ &+ (P_\mu^{D,sim} - P_\mu^{\pi,sim}) (\xi^{data} - \xi^{sim}). \end{aligned} \quad (4.13)$$

³⁶The greatest sensitivity to η comes from measurements of the decay positron transverse polarisation[18]. The energy and angle of the decay positron has low sensitivity. Since the world average value is used, a separate uncertainty on ρ and δ is quoted that depends on η .

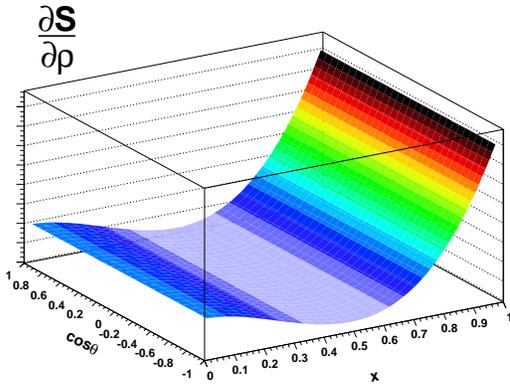
Now $(P_\mu^{D,sim} - P_\mu^{\pi,sim})$ is the amount of depolarisation in the simulation, which is of order 10^{-3} . The term $(\xi^{data} - \xi^{sim})$ depends on the tolerance of the hidden values, which is at most 10^{-2} . Therefore this term is of order 10^{-5} , and can be safely neglected. In other words, even though spectra at the time of decay are being compared, the result for $P_\mu^\pi \xi$ at the time of muon production can be safely extracted.



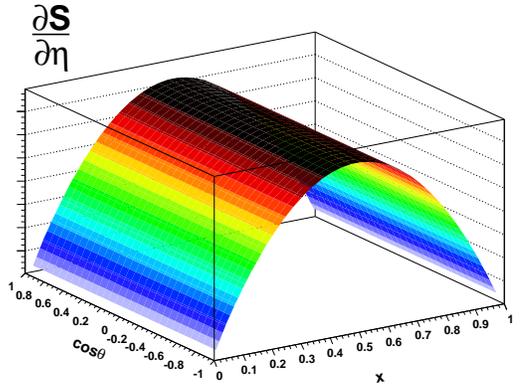
(a) Derivative with respect to $P_\mu^\pi \xi$.



(b) Derivative with respect to $P_\mu^\pi \xi \delta$.



(c) Derivate with respect to ρ .



(d) Derivative with respect to η .

Figure 4.12: The shape of each muon decay parameter's contribution to the total spectrum. The vertical scales are inconsistent between the figures since only the *shape* is important here.

4.5 Momentum calibration

The absolute energy scale can be tested since the maximum positron energy (the “kinematic endpoint”) is given by special relativity as 52.8 MeV, in the limit of massless neutrinos. This region is shown for a $\cos\theta$ slice in Fig. 4.13, where the sharp edge has been smeared by the reconstruction resolution and radiative corrections. The figure shows a fit, where a step function has been convoluted with a Gaussian resolution.

Figure 4.14 shows the step position, as determined from the fits, for a range of angles. The black lines indicate the kinematic prediction. The dependence on reconstructed angle is due to the planar nature of the detector, where path length (and hence energy loss) is proportional to $1/|\cos\theta|$. This is clear when the endpoint positions are re-plotted on this scale, as shown in Fig. 4.15.

The reconstructed endpoints for data and simulation are clearly displaced from the kinematic prediction. The previous $P_\mu^\pi \xi$ analysis had a larger discrepancy since energy loss was not included when fitting the positron helices. The current discrepancy is from reconstruction bias and the energy loss of the positron as it leaves the metal target, before its trajectory can be measured. The bias can be measured directly using the simulation, and indirectly using the data; see Section 6.8.

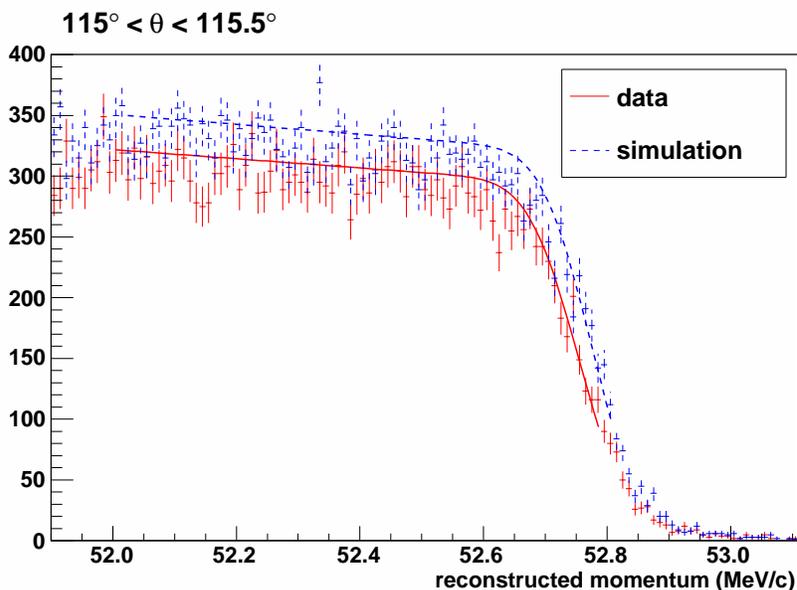


Figure 4.13: Example of fitting the endpoint in data and simulation.

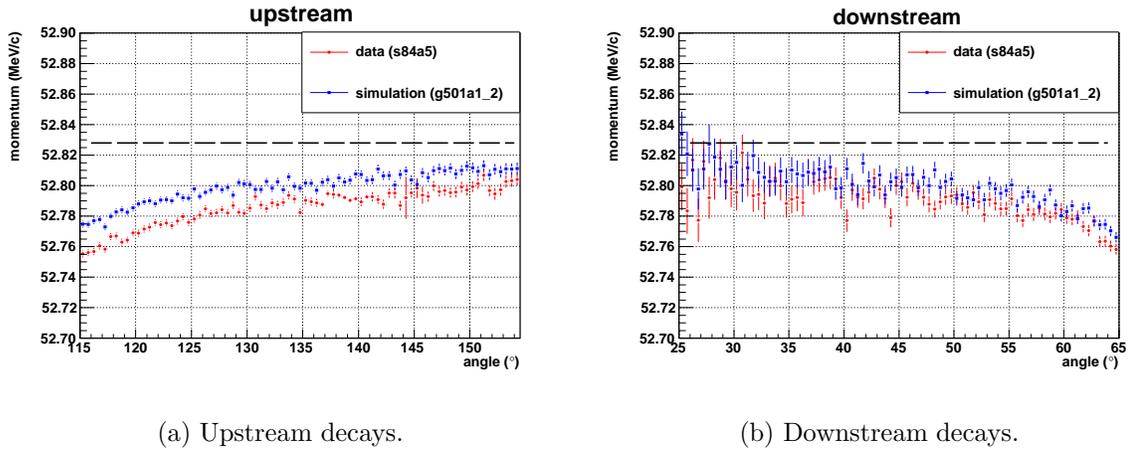


Figure 4.14: Dependence of endpoint position on reconstructed angle. The black dashed line is the expected endpoint from kinematics.

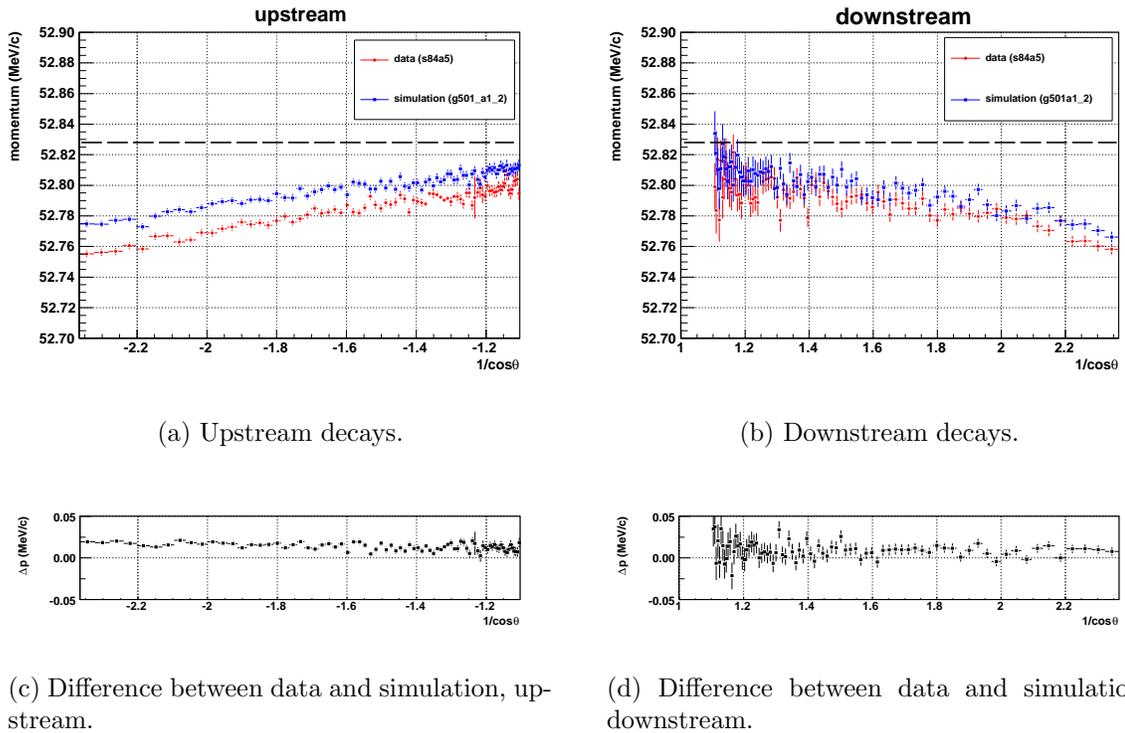


Figure 4.15: Dependence of endpoint position on $|1/\cos\theta|$. The black dashed line is the expected endpoint from kinematics.

The discrepancy from the kinematic endpoint is of interest. However the difference between data and simulation is the relevant measure for the muon decay parameters. The agreement has improved since the previous $P_{\mu}^{\pi} \xi$ analysis, due to the use of STRs that treat both data and simulation the same, and a realistic drift cell resolution function (see Sections 4.2.7 and 4.2.8). The remaining discrepancy has three causes: the mismatch in stopping distribution between data and simulation, the thickness of the stopping target used in the simulation, and the scale of the magnetic field map in data.

The muon stopping distribution in simulation can be tuned with arbitrary precision, since the exact muon stopping location is known. However, in the data, the only available measurement is where the muon was last seen, which was used to regulate the stopping position. This is an imperfect measure due to statistical precision and a background from pion decays at the end of the M13 channel (see Section 9.6). Consequently the positrons in data and simulation see a different amount of target material.

Direct thickness measurements of the target are destructive; instead the thickness for the simulation is an average derived from the mass and density of the foil. The thickness of the 5 cm diameter region where the muons actually stop may not be this average value due to foil non-uniformity. In addition the foil is under tension in data, reducing its thickness. There is an additional indirect measure of the target thickness, from data where the muons were stopped at the entrance to the stack, and the decay positron was reconstructed separately in each half of the detector; again, see Section 6.8.

The stopping distribution and target thickness affect the slope of the relationships in Fig. 4.15. The final discrepancy, the magnetic field scale, affects the vertical offset. In the simulation the same field is used to generate and subsequently analyse. However, in the data, the field map is scaled according to an NMR probe measurement taken at a single location. This probe is positioned at a region where the field is less well known, allowing a possible error in the total field scale used to analyse the data.

Figures 4.15(c) and 4.15(d) show the offset between data and simulation is between 0.010 MeV and 0.020 MeV, with a dependence on $1/|\cos \theta|$. In the previous $P_{\mu}^{\pi} \xi$ analysis, the events in both the data and simulation spectra were shifted according to $\cos \theta$ so that the endpoint agreed with 52.8 MeV. For the current measurement, a shift was applied to either the data or the simulation, so that the endpoints overlapped³⁷. This is a much smaller change than the previous analysis. However, either approach will be seen to produce the same result for $P_{\mu}^{\pi} \xi$.

³⁷For the current analysis there was no model applied to the endpoint. Instead the data and simulation spectra were shifted until the differences in each bin close to the endpoints were minimised.

4.6 Time dependence of depolarisation

In the experiment's previous analysis of $P_\mu^\pi \xi$ the relaxation of the polarisation was determined by fitting

$$G = \frac{N_F - N_B}{N_F + N_B} = K P_\mu(t), \quad (4.14)$$

where F and B are the number of the forward and backward counts within the standard fiducial, and K is a constant that depends on this fiducial. Equation (4.14) can be explicitly written

$$G = \frac{\int_p^q \int_a^b n(x, \cos \theta) dx d \cos \theta - \int_p^q \int_{-b}^{-a} n(x, \cos \theta) dx d \cos \theta}{\int_p^q \int_a^b n(x, \cos \theta) dx d \cos \theta + \int_p^q \int_{-b}^{-a} n(x, \cos \theta) dx d \cos \theta}. \quad (4.15)$$

If the only quantity of interest is the time dependence of G , then a new quantity with the same time dependence can be defined as

$$G' = \frac{\int_p^q \int_a^b w_1 n dx d \cos \theta - \int_p^q \int_{-b}^{-a} w_1 n dx d \cos \theta}{\int_p^q \int_a^b w_2 n dx d \cos \theta + \int_p^q \int_{-b}^{-a} w_2 n dx d \cos \theta}, \quad (4.16)$$

where $w_1(x, \cos \theta)$ and $w_2(x, \cos \theta)$ are weighting terms that can be freely chosen. Since $\cos \theta > 0$ corresponds to upstream decays, and $\cos \theta < 0$ to downstream decays, Eq. (4.16) can be simplified by making $w_1 \propto \cos \theta$ and $w_2 \propto |\cos \theta|$; the sum over the bins, both upstream and downstream, is then

$$G' = \frac{\sum_{\text{US,DS}} w_1 N}{\sum_{\text{US,DS}} w_2 N}, \quad (4.17)$$

where N is the number of integrated counts in the bin.

For the current analysis, it was suggested that w_1 and w_2 are related to the theoretical asymmetry itself[50]. The full expression for the differential decay rate was given in Eq. (4.7). Neglecting radiative corrections and the positron mass, and assuming standard model values for ρ and δ , the differential decay rate is

$$\frac{d^2 \Gamma}{dx d \cos \theta} = x^2 [(3 - 2x) + P_\mu \xi \cos \theta (2x - 1)], \quad (4.18)$$

which is a theoretical asymmetry of

$$A(x, \cos \theta) = P_\mu \xi \cos \theta \frac{x - \frac{1}{2}}{\frac{3}{2} - x}. \quad (4.19)$$

After integration over $\cos \theta$, Eq. (4.19) is shown in Fig. 4.16. The greatest contribution to the integral asymmetry clearly comes from the highest energy positrons; the positrons with $x \approx 0.5$ contribute little to the total asymmetry yet add statistical noise. Positrons with $x < 0.5$ actually detract from the total asymmetry, and add more statistical noise. μ^+ SR experiments have taken advantage of Fig. 4.16 by simply not recording positrons below a threshold energy. However, a better approach is to weight each positron according to how much it contributes to the asymmetry. In other words, w_1 and w_2 need to be closely related to Eq. (4.19).

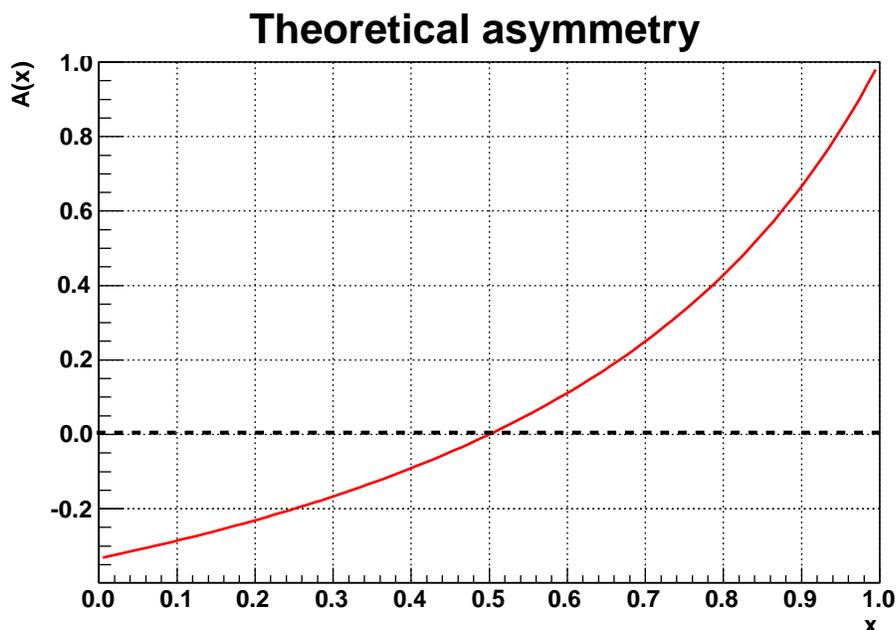


Figure 4.16: Theoretical prediction for the asymmetry, assuming standard model values for the muon decay parameters.

The weighting factors used in the current analysis were

$$w_1 = A(x, \cos \theta) |A(x, \cos \theta)|^n, \quad (4.20)$$

$$w_2 = |A(x, \cos \theta)| |A(x, \cos \theta)|^n, \quad (4.21)$$

$$n = 1. \quad (4.22)$$

Different powers of n were examined, but $n = 1$ combined with rejection of positrons with

$x < 0.59$ was found to be close to optimum³⁸. The combination of weighting the counts, and removing low energy positrons significantly reduces the uncertainty on the depolarisation rate. An example of a fit to a single set is shown in Fig. 4.17, where a fit function $P_\mu(t) = P_\mu(0) \exp(-\lambda t)$ has been used.

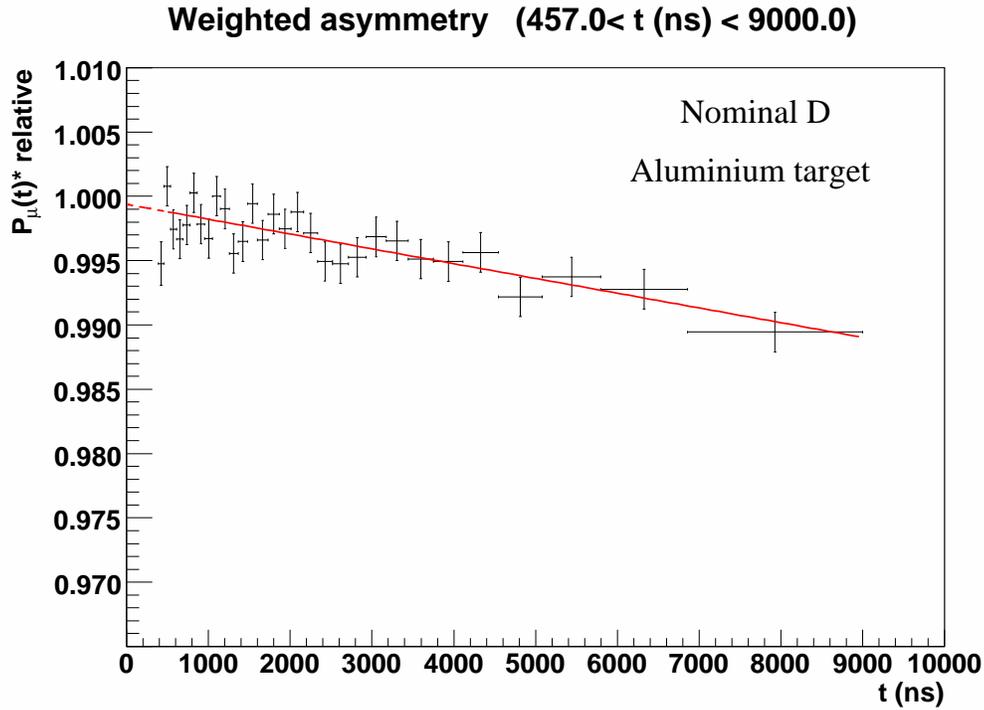


Figure 4.17: Example of an exponential fit to the asymmetry.

³⁸The lower momentum positrons could be included, and their sign could be reversed for $x < 0.5$, but this doesn't improve the statistical precision by a significant amount.