

6. Neutrino Flavour and Phase Estimators

The data analysis undertaken in this thesis will distance itself from the traditional Soudan 2 scanning methods [7][47]. The data are not studied on an event-by-event basis but a statistical analysis is favoured instead. In this chapter, the experimentally measured quantities that are characteristic of the flavour, energy and direction of a neutrino interaction will be introduced. They will be referred to as *estimators*. Their performance will be evaluated by studying the MCNO event sample. The complete event samples will be considered, where all scattering processes are taken into account, including neutral current interactions. The analysis of the real data sample and its comparison to the MCNO and MCAF expectation follows in the subsequent chapters.

6.1 Flavour estimators

Two approaches have been employed in order to extract the flavour content of the events on a statistical basis between ν_e and ν_μ interactions in the Soudan 2 detector. The first, Γ , looks at the shape of the event as a whole by considering its moment of inertia eigenvalues, while the second, Λ_{46} , looks at the microstructure of the event by examining the relative position of neighbouring hits.

6.1.1 Cumulative Flavour Distributions

Before proceeding with the discussion on flavour estimators, let us make a point on the cumulative (integrated) distributions that will be used in this chapter in order to extract the separation efficiency and misidentification matrices of the flavour estimators.

The flavour estimators described in the following sections divide the events into ν_e -like and ν_μ -like. For both Γ and Λ_{46} , the ν_e -like events are found at low values of the estimator, while the ν_μ -like events populate the high values of the estimator. In order to illustrate the point the reader is invited to glance at the top two Γ distributions in page 126. From the ν_e -like and ν_μ -like distributions of either Γ or Λ_{46} , we wish to extract the separation efficiency at any value of the estimator. For this purpose cumulative distributions of each flavour are compared. However, the cumulative distributions for ν_e and ν_μ differ: for a certain value of the flavour estimator, the ν_e -like distribution gives the fraction of ν_e events *above* that value, while the ν_μ -like distribution gives the fraction of ν_μ events *below* that value (see Γ cumulative distributions in page 128, third plot from the top). The misidentification of ν_e and ν_μ interactions will be equal at the *crossing point* of the ν_e and ν_μ curves (when these are plotted on top of each other) and will be used as a measure of the flavour separation efficiency of the estimators. This procedure simplifies the definition of a flavour cut and the construction of truth matrices.

6.1.2 Moments of Inertia: Γ

Intuitively, a track event is “long and thin” whereas a shower event is “short and broad”. The first flavour estimator is defined by the ratio of the longitudinal to the transverse mean-squared components of the event:

$$\Gamma = \log_{10} \left(\frac{\overline{z'^2}}{\overline{r'^2}} \right) \quad (6.1)$$

where the mean is over the number of hits in the event box, N , and the radial component is given by $r'^2 = x'^2 + y'^2$. The rotated orthogonal co-ordinate system x' - y' - z' has its origin at the centre of gravity of the event and is defined by the Moment of Inertia (MoI) Tensor eigenvectors (see §3.3 and fig. 3.8). The moment of inertia of the event about the z' -axis is minimum, i.e. the z' -axis coincides with the event axis. By re-writing the expression as

$$\Gamma = \log_{10} \left(\frac{\sum_{i=1}^N z_i'^2 / N}{\sum_{i=1}^N (x_i'^2 + y_i'^2) / N} \right) \quad (6.2)$$

and then by adding and subtracting the same terms in the nominator,

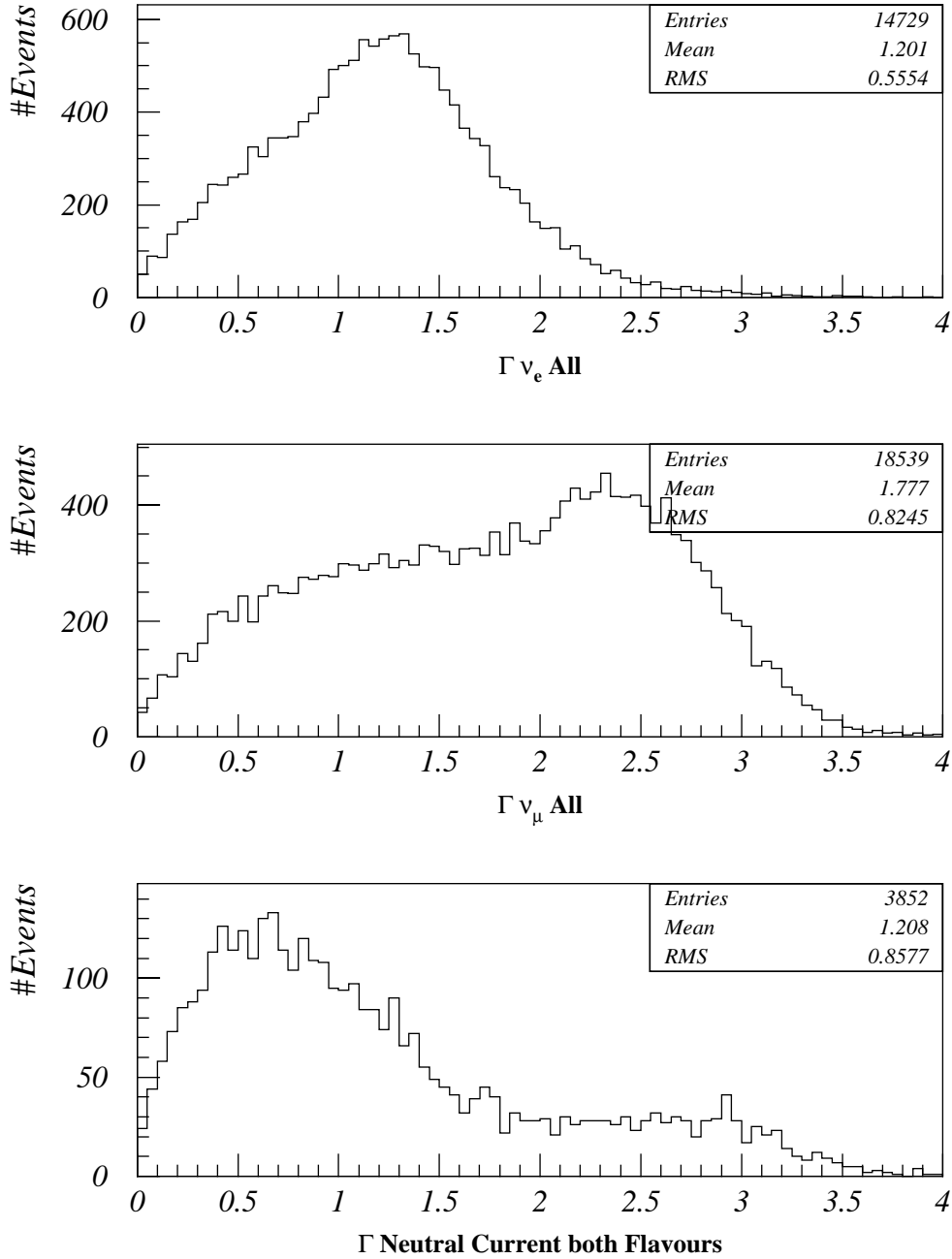
$$\Gamma = \log_{10} \left(\frac{\sum_{i=1}^N (x_i'^2 + z_i'^2) + \sum_{i=1}^N (y_i'^2 + z_i'^2) - \sum_{i=1}^N (x_i'^2 + y_i'^2)}{2 \sum_{i=1}^N (x_i'^2 + y_i'^2)} \right) \quad (6.3)$$

it is observed that Γ can be constructed by the moment of inertia (MoI) eigenvalues of the event about its centre of gravity (as defined in §3.3):

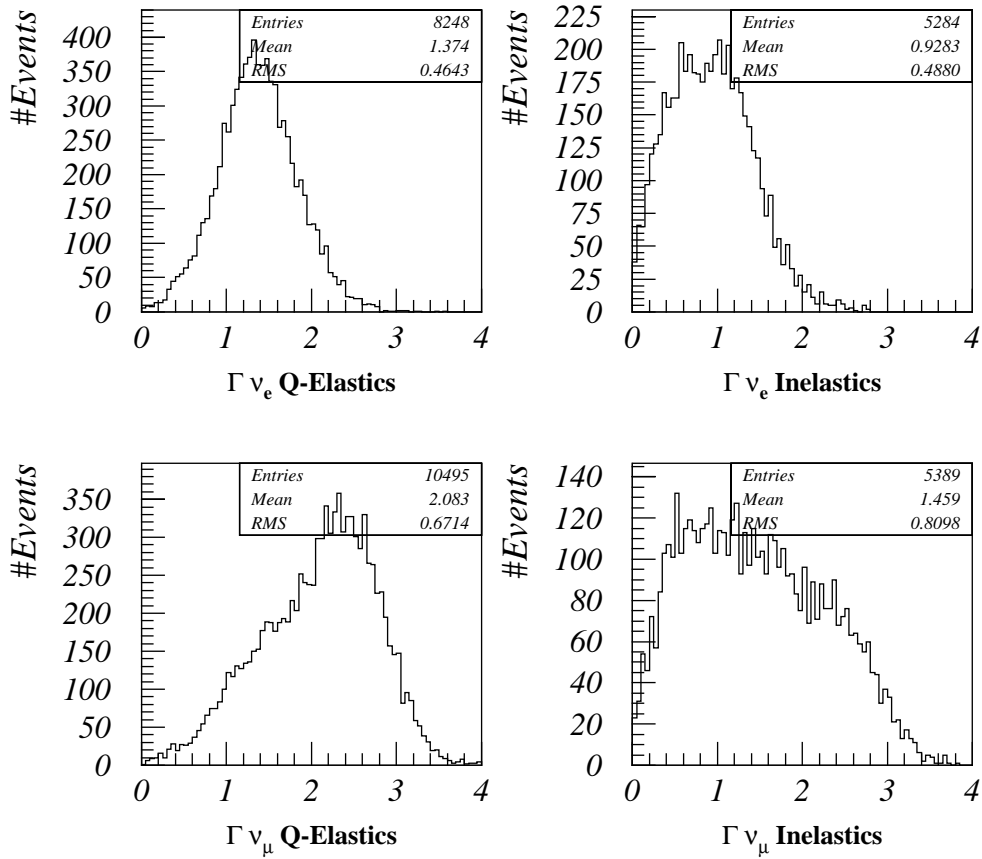
$$\Gamma = \log_{10} \left(\frac{\text{MoI}_B + \text{MoI}_C - \text{MoI}_A}{2 \text{MoI}_A} \right) \quad (6.4)$$

By definition, long and thin muon events will be described by large values of Γ , say above 2, where the ratio of longitudinal to transverse components exceeds 100, while lower values of Γ will be characteristic for rounder electron shower events. Indeed, less than 10% of ν_e events, compared to 50% of ν_μ events, have $\Gamma > 2$ (figs. **6.1**, **6.3**). The neutral current interactions, which are common to both flavours, will produce very round events ($\Gamma < 1.5$) but also exhibit a long tail up to $\Gamma = 4$, due to short, straight proton tracks knocked out of nuclei (fig. **6.1**).

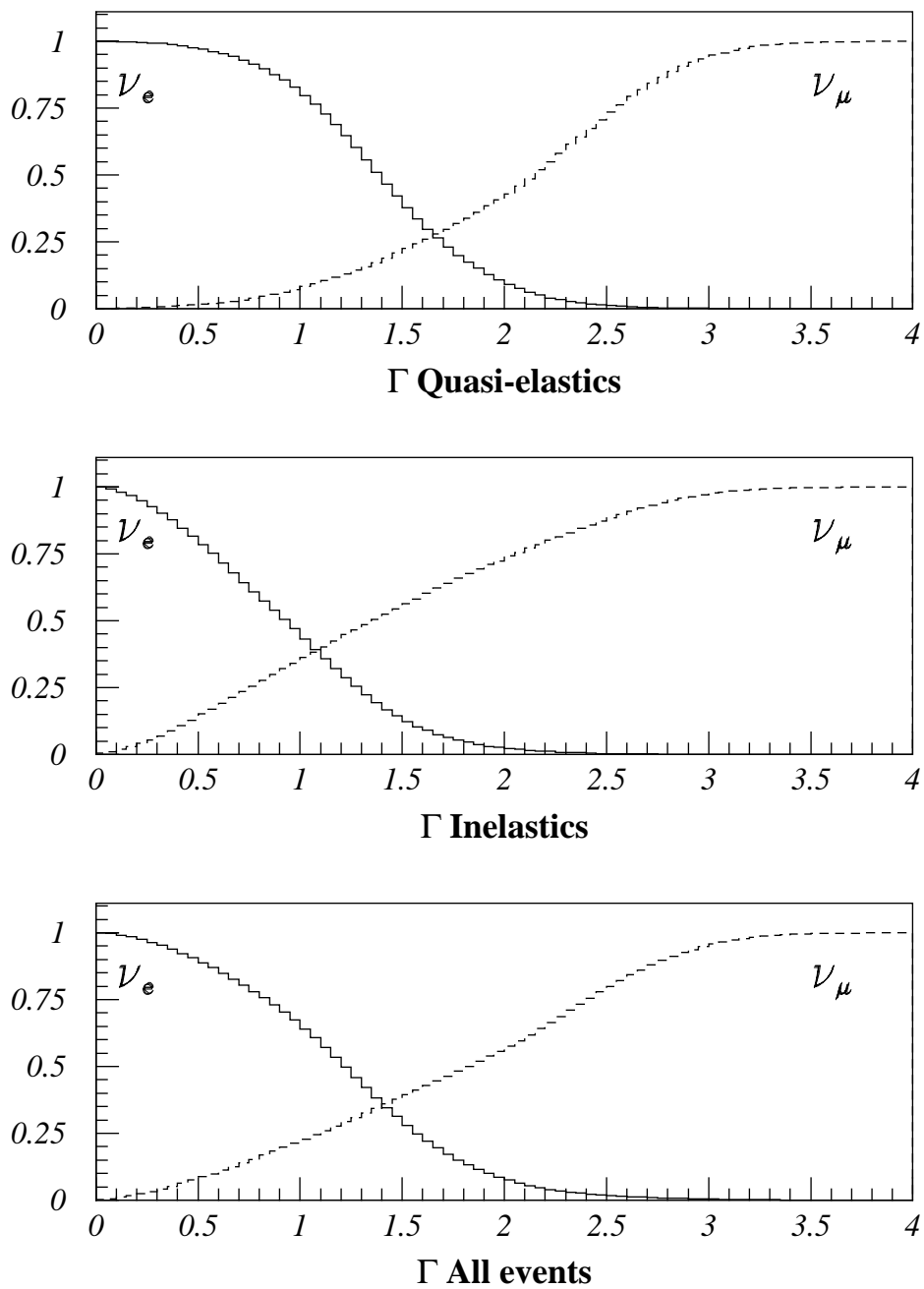
The misidentification between the complete ν_e and ν_μ samples at the crossing point of $\Gamma = 1.4$ is 35% (fig. **6.3**). The misidentification performance improves to 30% when a low energy cut of 40 hits (≈ 850 MeV) is applied (fig. **6.4**). Such a cut divides the MCNO sample into two equally populated event sets. The MC charged current data may be divided into quasi-elastic and inelastic interactions, a distinction that cannot be made for real data and is only of academic interest. As expected, the misidentification is lower for quasi-elastic events where the lepton dominates and flavour tagging is easier (18% and 25% with and without an energy cut respectively, figs. **6.2**, **6.3**, **6.4**).



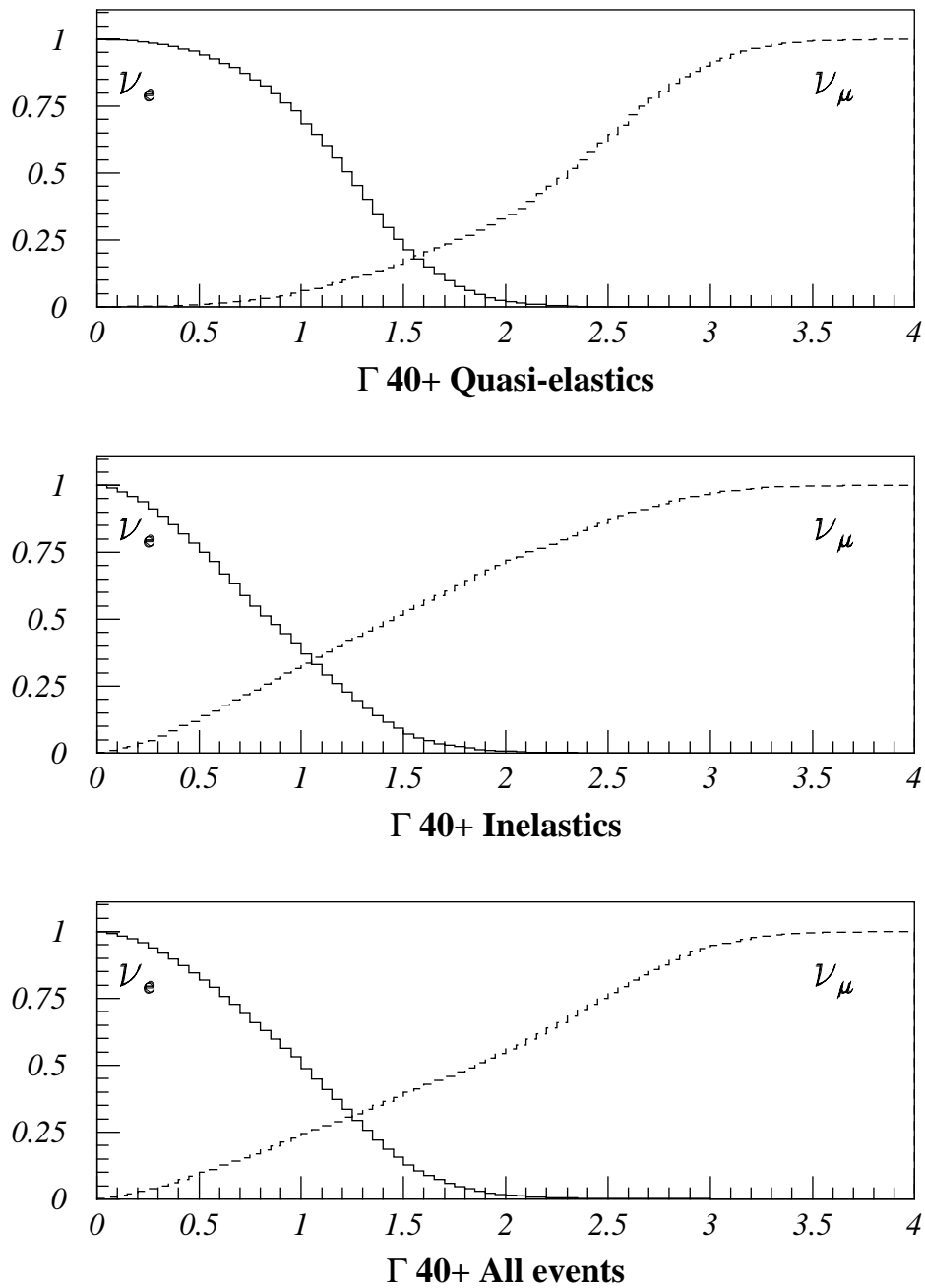
• **Figure 6.1:** Distributions of Γ for ν_e and ν_μ and Neutral Current MCNO events that have passed the RINSE cuts. The top two plots are for all scattering processes, including neutral current interactions. Tau neutrinos, ν_τ , which would be produced in $\nu_\mu \leftrightarrow \nu_\tau$ oscillation modes (among others) would manifest their presence as Neutral Current events because the mean energy of the atmospheric neutrino spectrum is well below the tau production threshold.



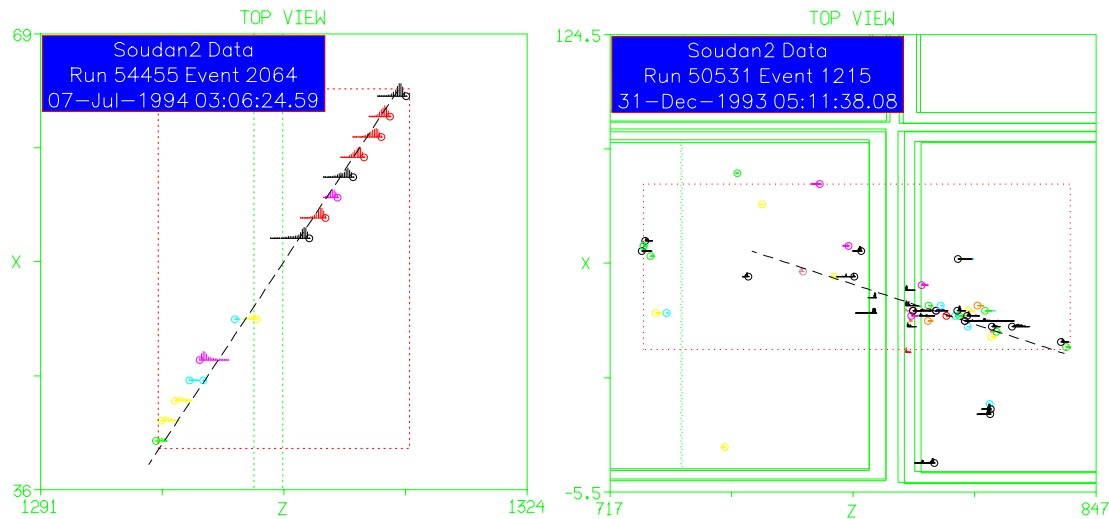
• **Figure 6.2:** Distributions of Γ for ν_e and ν_μ quasi-elastic and inelastic MCNO events that have passed the RINSE cuts.



• **Figure 6.3:** Cumulative distributions of Γ for ν_e and ν_μ quasi-elastic, inelastic and all interactions (including neutral current) of MCNO events that have passed the RINSE cuts.



• **Figure 6.4:** Cumulative distributions of Γ for ν_e and ν_μ quasi-elastic, inelastic and all interactions (including neutral current) of MCNO events of 40 or more 3D hits that have passed the RINSE cuts.



• **Figure 6.5:** Pictures (x-z projection) of a short (13 hits) straight track (left) and a 36-hit shower (right). The scale in the plots is in cm, while the dashed line corresponds to the outline of the event box.

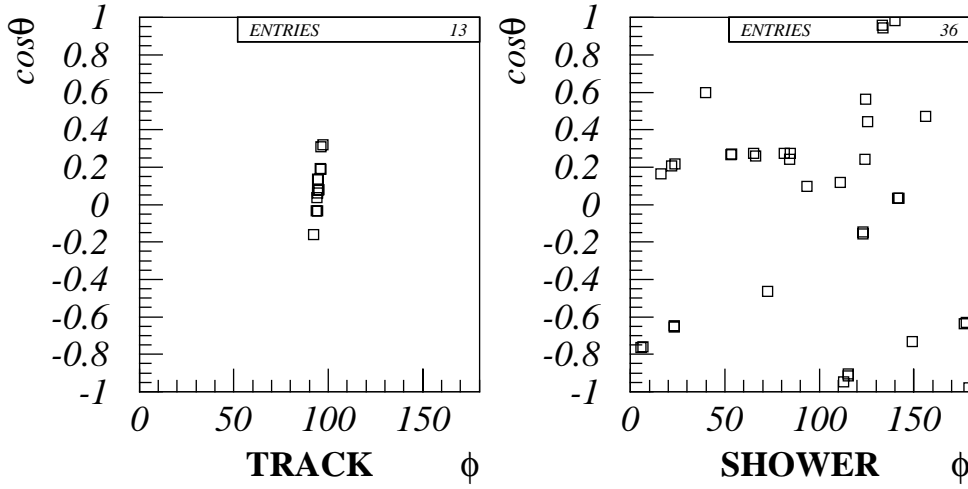
6.1.3 Vectors joining neighbouring hits: Λ_{46}

The second flavour algorithm looks at the microstructure and “trackiness” of the event. For a track event, where the hits lie along a line, the vectors joining neighbouring hits are expected to point, more or less, in the same direction. In a shower event, the hits are distributed almost randomly relative to each other and the vectors joining neighbours may point in all directions (fig. 6.5).

The following discussion is valid for any definition of neighbouring hits, which will be dealt with in due time, so as to ease the readability of the text. However, it is deemed necessary to introduce a couple of examples which require a definition of neighbouring hits. For the sake of argument, let *only nearest neighbours* be used. In this simple case the number of vectors joining neighbouring hits is equal to the number of hits in the event.

- **Power Spectrum of Distribution on a Sphere**

Consider the set of vectors between neighbouring hits in the event box. When normalised to unity they will form a distribution on the surface of a unit sphere, $f(\cos \vartheta, \phi)$, which will be



• **Figure 6.6:** Scatter plots of $\cos\vartheta$ versus ϕ for $f(\cos\vartheta, \phi)$ of the track and shower of fig. 6.5, where only vectors between neighbouring hits are used. The coordinate system for each event has been chosen such that the event axis comes out of the page through the middle of the plot. The track distribution is narrow about the event axis while the shower distribution is very broad.

quite localised for a track event but will be much broader for a shower event (fig. 6.6). A “localisation” estimator looking for density hotspots on the sphere surface should therefore resolve the flavour of the event. This problem is remarkably similar to the cosmic background radiation explored with data from the Cobe satellite [77], where the analysis involved taking the power spectrum of $f(\cos\vartheta, \phi)$. The latter can be expressed in term of spherical harmonics $Y_{LM}^{\pm}(\cos\vartheta, \phi)$:¹⁵

$$f(\cos\vartheta, \phi) = \sum_{L=0}^{\infty} \sum_{M=0}^L [a_{LM}^+ Y_{LM}^+(\cos\vartheta, \phi) + a_{LM}^- Y_{LM}^-(\cos\vartheta, \phi)] \quad (6.5)$$

where, according the classical formalism,

$$Y_{LM}^+(\cos\vartheta, \phi) = \sqrt{\frac{2L+1}{2\pi} \frac{(L-M)!}{(L+M)!}} P_L^M(\cos\vartheta) \cos(M\phi) \quad (6.6)$$

$$Y_{LM}^-(\cos\vartheta, \phi) = \sqrt{\frac{2L+1}{2\pi} \frac{(L-M)!}{(L+M)!}} P_L^M(\cos\vartheta) \sin(M\phi)$$

for $M \neq 0$, while

$$\begin{aligned} Y_{L0}^+(\cos \vartheta, \phi) &= \sqrt{\frac{2L+1}{4\pi}} P_L^0(\cos \vartheta) \\ Y_{L0}^-(\cos \vartheta, \phi) &= 0 \end{aligned} \quad (6.7)$$

for $M = 0$ [28]. The coefficients a_{LM}^\pm will be obtained by multiplying $f(\cos \vartheta, \phi)$ by the spherical harmonics $Y_{LM}^\pm(\cos \vartheta, \phi)$, which are orthogonal functions, and by integrating over $\cos \vartheta$ and ϕ . Because $f(\cos \vartheta, \phi)$ is a sum of δ -functions, the expression for a_{LM}^\pm is simplified to

$$a_{LM}^\pm = \sum_{\substack{\text{vectors} \\ (\cos \vartheta, \phi)}} Y_{LM}^\pm(\cos \vartheta, \phi). \quad (6.8)$$

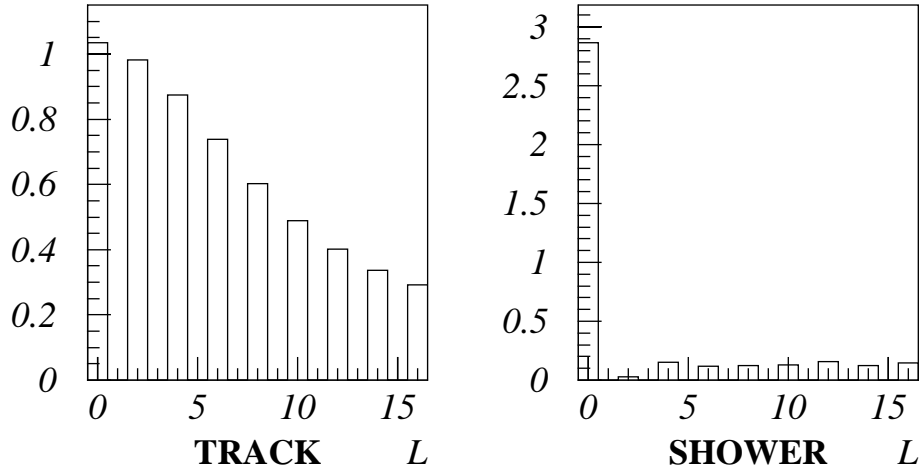
Finally, the power spectrum of the event as a function of ‘‘angular momentum’’ L is given by

$$W(L) = \frac{1}{N} \frac{1}{2L+1} \sum_{M=0}^L \left((a_{LM}^+)^2 + (a_{LM}^-)^2 \right) \quad (6.9)$$

where N is the number of vectors between neighbouring hits considered in the calculation. The summation over M removes the significance of the chosen co-ordinate system used in defining $\cos \vartheta$ and ϕ , which, for each event, is arbitrarily chosen so that the event axis coincides with $\cos \vartheta = 0$ and $\phi = 90^\circ$.

In order to illustrate the power of this calculation, consider the track event and shower event discussed above, still using only vectors between the nearest neighbours. The power spectra of these two events exhibit very different behaviour (fig. 6.7): the $L = 0$ term, which is proportional to the number of vectors, dominates the spectrum for the shower event, while it is not very different from the higher L terms for the almost straight track. We may choose to define a flavour estimator, Λ_{46} , by

¹⁵ The Spherical Harmonics functions were obtained from the FORTRAN routine RASLGF of the CERN library. I would like to thank Prof. G. Efstathiou for providing me with his equivalent routine which goes up to very high values of L (the CERN routine only reaches up to $L=16$). The results from both routines were identical.



• **Figure 6.7:** Power spectra of distributions $f(\cos \vartheta, \phi)$ (fig. 6.6) for the track and shower of fig. 6.5. The $L=0$ term is proportional to the number of vectors populating $f(\cos \vartheta, \phi)$. The odd L terms are zero because the vectors between neighbouring hits have no sense of direction and $f(\cos \vartheta, \phi)$ lies on a hemisphere. The difference in shape is typical and will be exploited to extract a flavour estimator between track-like and shower-like events.

$$\Lambda_{46} = \frac{W(L=4) + W(L=6)}{W(L=0)}, \quad (6.10)$$

which for this particular track has a value close to 2 while for the shower it is close to zero. Similarly other combinations, such as Λ_{24} , Λ_{68} or even Λ_{2468} may be defined.

• Definition of Neighbouring Hits

The above discussion is valid under any definition of neighbouring hits. Apart from nearest neighbours, the following options have also been explored:

- 1) Only neighbours not further apart than 5 cm are used.
- 2) Only neighbours not further apart than 10 cm are used.
- 3) Only neighbours not closer than 2 cm and not further apart than 10 cm are used.
- 4) Only neighbours not further apart than 15 cm are used; this figure corresponds to approximately one radiation length in the Soudan 2 detector.

The conclusions drawn from nearest neighbours remain unchanged when using any of the other four options. The flavour misidentification is of the order of 18% *irrespective* of the choice of definition of neighbouring hits, even when the possible definitions of Λ , such as Λ_{24} , Λ_{68} etc, are taken into account.

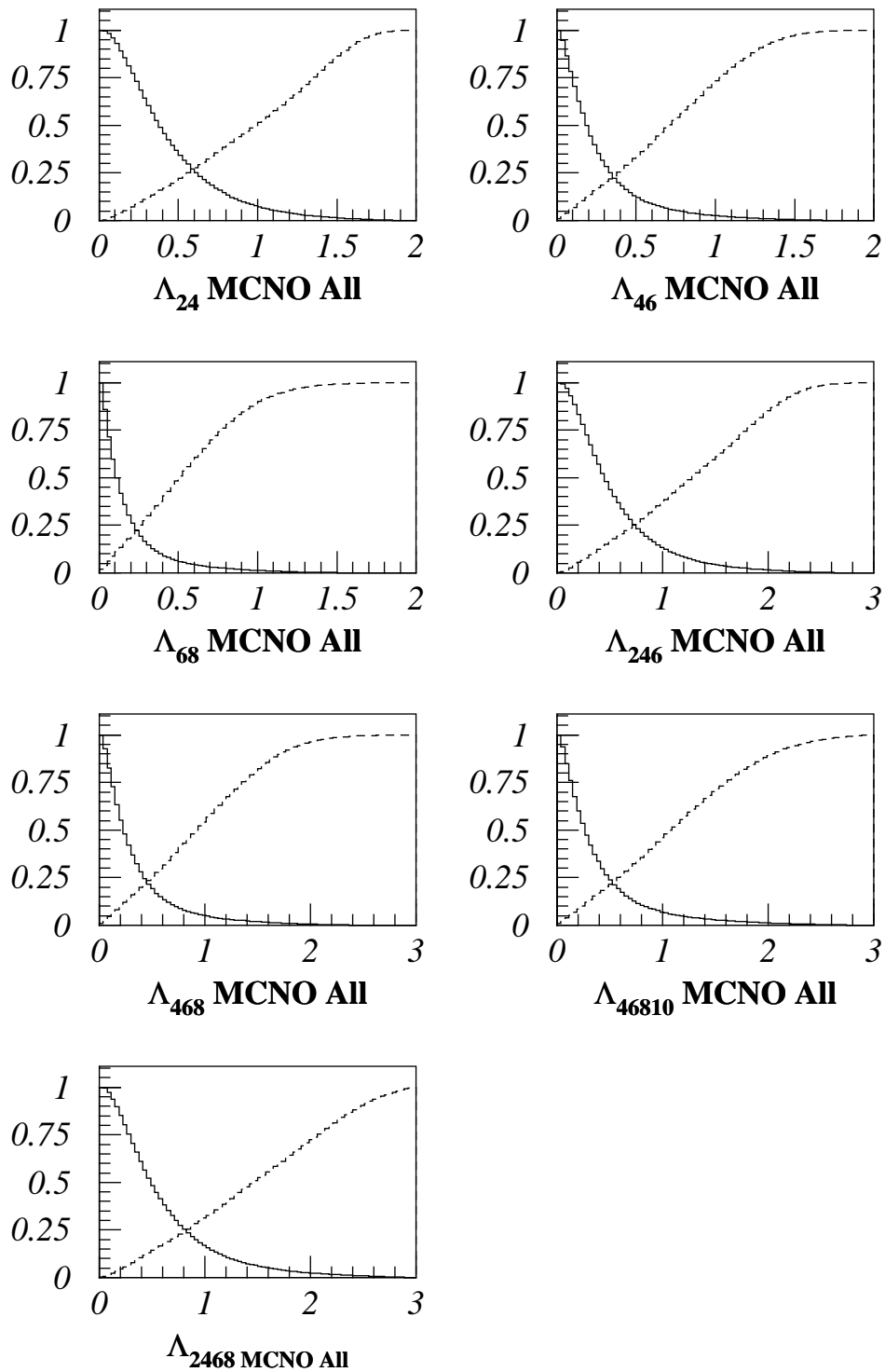
The choice was eventually taken on grounds of systematic effects between real and MC data. In particular Λ may be affected by irregularities of the main detector operation. Small differences in the drift speed of adjoining drift tubes are *not* included in the MC and may distort the event in the small scale: if a pulse is delayed by $1\mu\text{s}$ (0.5 cm) relative to its closest neighbour (1-2cm apart), the vector joining them may be rotated by as much as 30° . Such an effect may produce a bias between MC and real data. For such effects to become negligible, it was decided to define nearest neighbours according to option (4): all vectors between hits not further apart than 15 cm will be used.

- **Choice of Λ Flavour Estimator**

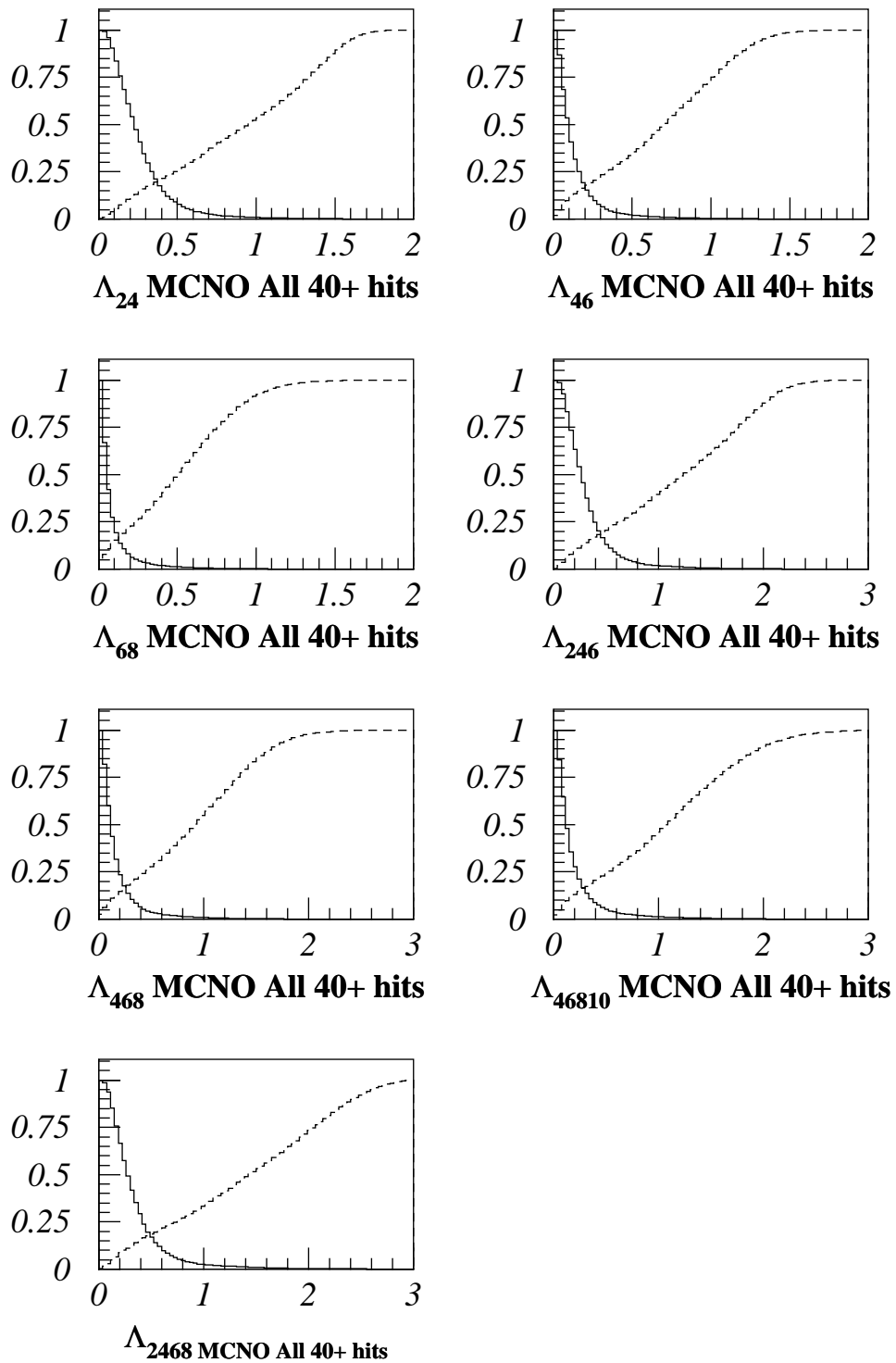
The quantities Λ_{24} , Λ_{46} , Λ_{68} , Λ_{246} , Λ_{468} , Λ_{2468} and Λ_{46810} are extracted from the power spectrum of each event in a fashion similar to Eq. 6.10. Their cumulative ν_e and ν_μ (all events) MCNO distributions are used to extract the misidentification at the point where this is equal for ν_e and ν_μ (fig. 6.8). The results are summarised in Table 6.1. It is observed that Λ_{46} offers the lowest misidentification figure of 22%, which occurs at a crossing value of 0.35. It is also interesting to observe that the dipole $L = 2$ term of an event's power spectrum reduces the flavour separation efficiency of the candidate estimator Λ . Higher L terms are probing into the event with greater detail and carry more flavour separation power. The law of diminishing returns comes into action for values of L above 8.

The exercise is repeated with a low energy cut of 40 hits (≈ 850 MeV); as a result, the misidentification is lowered for all flavour estimator candidates (fig. 6.9 and Table 6.1).

From the above, Λ_{46} is chosen as the “trackiness” flavour estimator.



• **Figure 6.8:** Cumulative distributions of the Λ -flavour estimator candidates for MCNO events that have passed the RINSE cuts.



• **Figure 6.9:** Cumulative distributions of the Λ -flavour estimator candidates for MCNO events of 40 or more 3D hits that have passed the RINSE cuts.

Flavour Estimator	All event types MCNO		All event types MCNO 40 or more hits	
	Crossing point	Misidentification	Crossing Point	Misidentification
Λ_{24}	0.58	0.27	0.37	0.19
Λ_{46}	0.35	0.22	0.20	0.16
Λ_{68}	0.24	0.22	0.13	0.16
Λ_{246}	0.73	0.24	0.44	0.18
Λ_{468}	0.45	0.23	0.25	0.17
Λ_{46810}	0.52	0.23	0.28	0.17
Λ_{2468}	0.82	0.24	0.47	0.17

• **Table 6.1:** Summary of cumulative distributions of Λ -flavour estimators candidates. The flavour misidentification is equal between ν_e and ν_μ flavours at the crossing point of the ν_e and ν_μ cumulative distributions. The performance of all estimators is improved by applying a low energy cut of 40 hits, as shown in the two columns on the right.

6.1.4 Further Discussion on Λ_{46}

In the data analysis that follows in the next chapter, the data sample will be analysed as a whole and no separation between single-prong and multi-prong events will be performed, as in the case of the scanned analysis.

It is intuitively simple to understand the mechanism of flavour separation for single-prong events using Λ_{46} . In order to get a quantitative feel of the flavour separation achieved by Λ_{46} for multi-prong events, the MCNO data is divided into quasi-elastic and inelastic interactions according to the MC truth. This distinction cannot be made for real data so this exercise is only of academic interest. However, the flavour misidentification for each process will be not calculated at the optimum crossing value of Λ_{46} for each of the two classes of events but at the crossing point determined by the *whole* data sample.

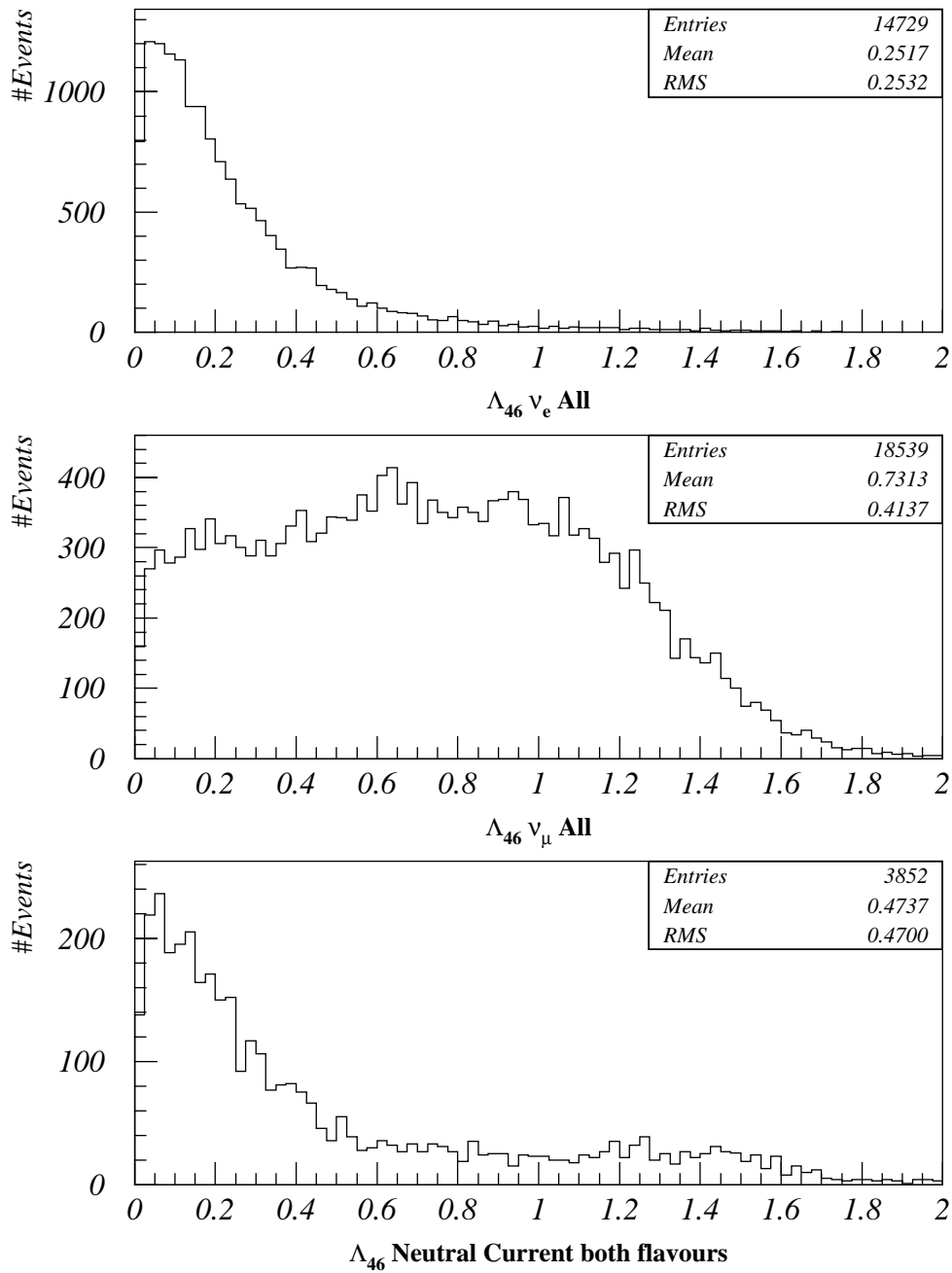
The distributions of Λ_{46} for each flavour and for neutral current interactions separately are shown in fig. 6.10. For each flavour, the Λ_{46} distributions for quasi-elastic and inelastic interactions separately are shown in fig. 6.11.

In terms of Λ_{46} , inelastic charged current events present a more rounded shape relative to their quasi-elastic counterparts because of the hadronic activity at their vertex. Hence, the misidentification crossing value of Λ_{46} is lower for inelastic processes than it is for elastic ones, as is attested by the cumulative distributions (fig. 6.12). Table 6.2 summarises the misidentification between ν_e and ν_μ interactions for quasi-elastic and inelastic processes, where an event is assigned ν_e flavour for $\Lambda_{46} < 0.35$ and ν_μ flavour for $\Lambda_{46} > 0.35$. The chosen value of Λ_{46} corresponds to the crossing point between ν_e and ν_μ interactions when *all* scattering modes (including neutral currents) are considered. Flavour separation is possible in the inelastic sample where less than 10% of ν_e events and 40% of ν_μ are erroneously identified. The unevenness between the two flavours is due to the fact that the crossing point for inelastic events alone is lower than that for all events (fig. 6.12). As expected, the picture is even better in the quasi-elastic sector.

The exercise is repeated by applying a low energy cut. Intuitively, one expects low energy inelastic events to carry little flavour information content, simply because the lepton cannot be identified amidst the hadronic activity. On the other hand, the lepton may be easily distinguished in high-energy inelastic processes. In particular, the presence or, equally significantly, the absence of a long track or shower may give away the flavour of the neutrino interaction.

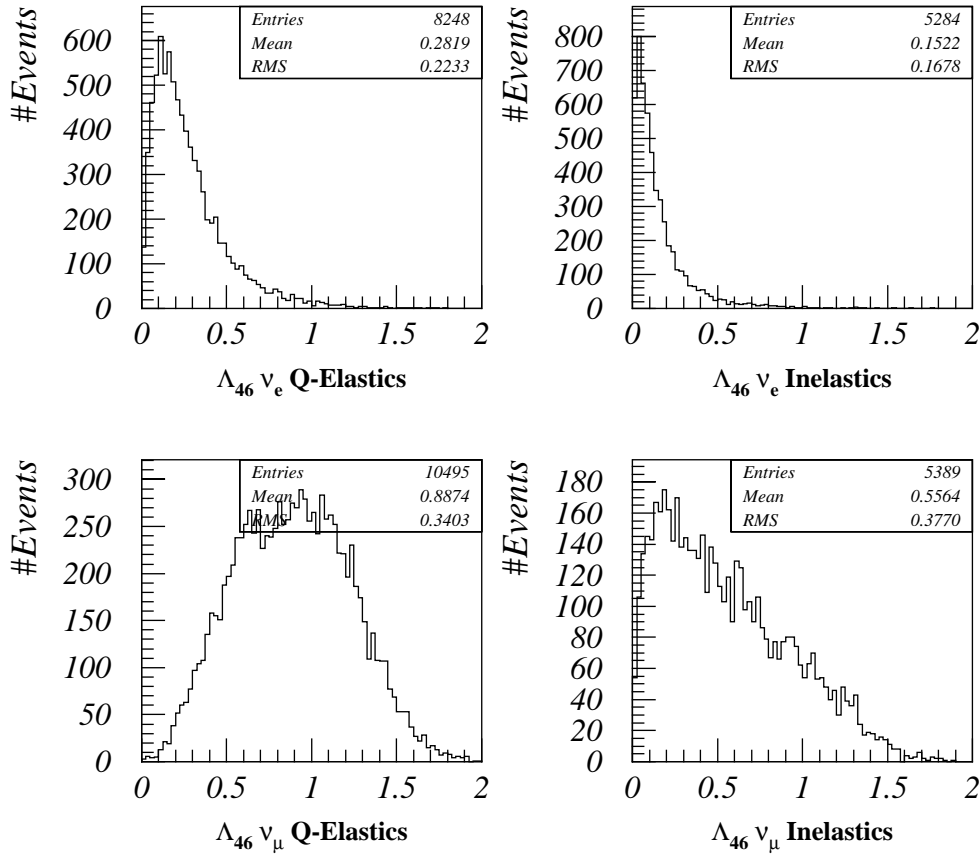
	MCNO Interactions			MCNO Interactions 40 or more hits		
	Q-Elastic	Inelastic	All	Q-Elastic	Inelastic	All
Crossing point	0.35			0.20		
ν_e Mis-ID	0.27	0.09	0.22	0.22	0.10	0.16
ν_μ Mis-ID	0.05	0.38		0.02	0.20	

• **Table 6.2:** Flavour Misidentification using Λ_{46} for event samples with and without a 40 3D Hit cut. The crossing point has been chosen such that the mis-id between the two flavours is the same for the complete event sample. The crossing value for Λ_{46} is energy dependent and is therefore different for the samples of large events.

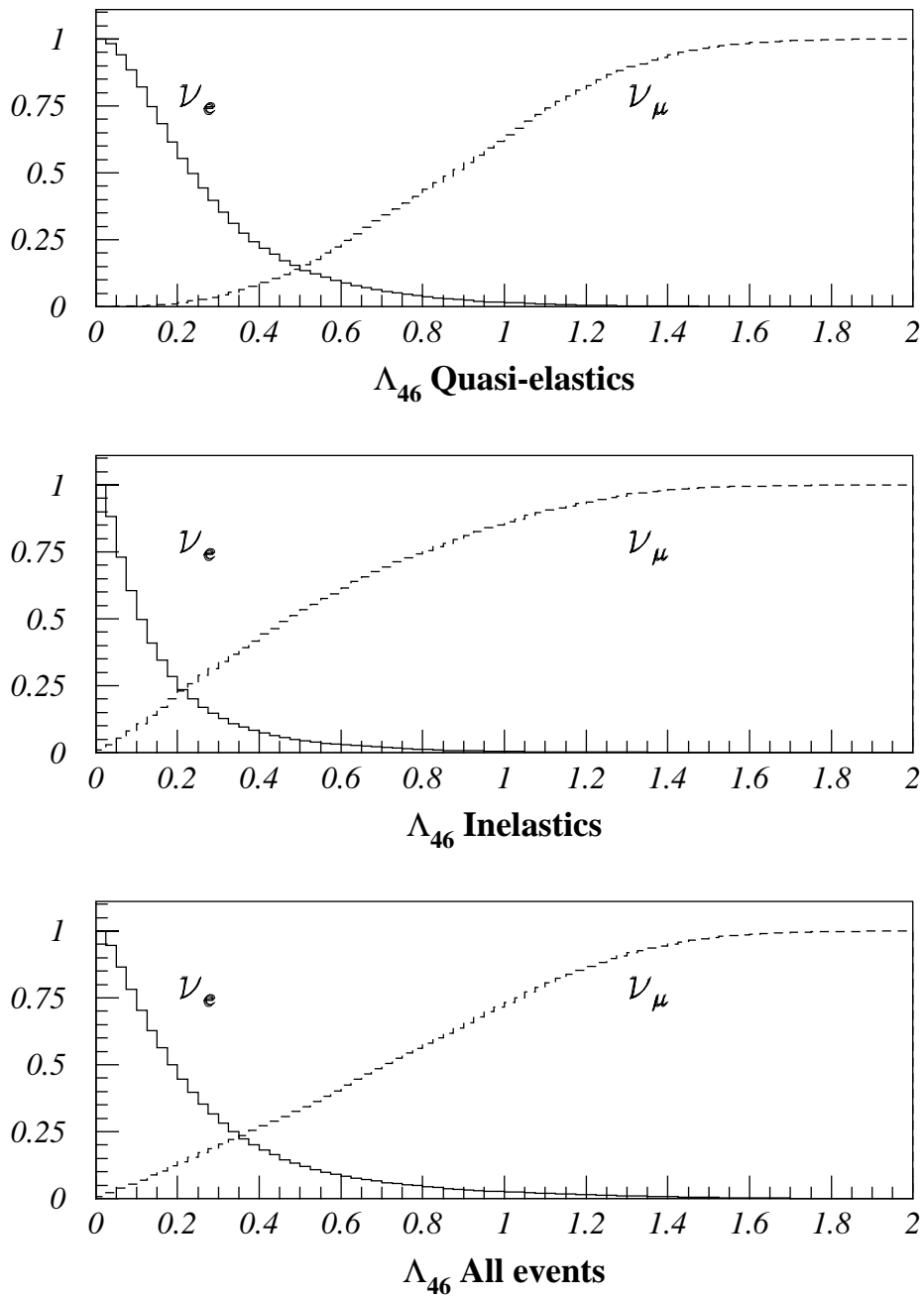


• **Figure 6.10:** Distributions of Λ_{46} for ν_e and ν_μ and Neutral Current MCNO events that have passed the RINSE cuts. The top two plots are for all scattering processes, including neutral current interactions. Tau neutrinos, ν_τ , which would be produced in $\nu_\mu \leftrightarrow \nu_\tau$ oscillation modes (among others) would manifest their presence as Neutral Current events because the mean energy of the atmospheric neutrino spectrum is well below the tau production threshold.

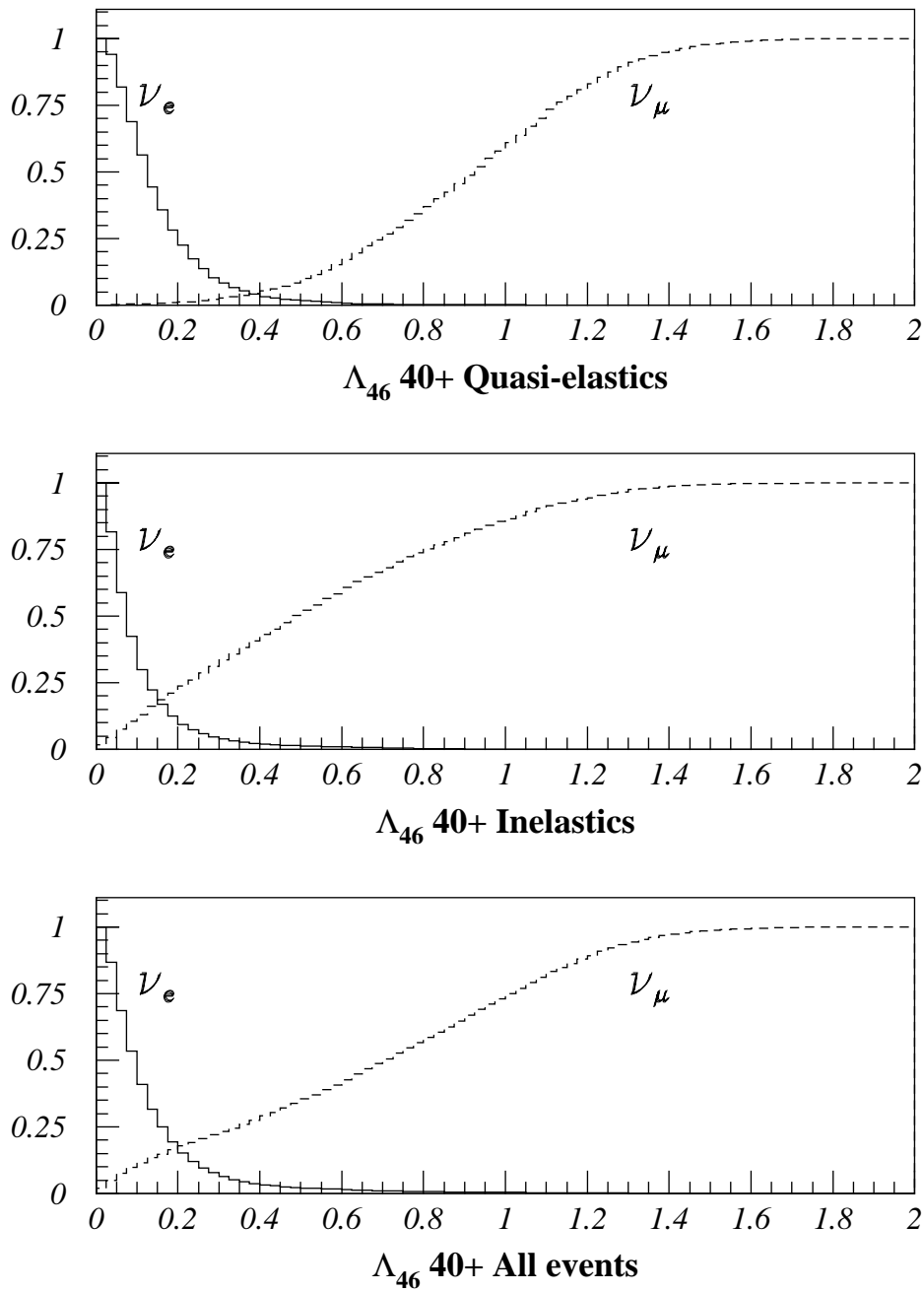
In this context, a large event is defined by having 40 or more 3D hits in the event box. The statistics of the MCNO sample decrease by a factor of 2, but better flavour separation is indeed observed (fig. 6.13 and Table 6.2). At the crossing value of $\Lambda_{46} = 0.2$ the misidentification for the whole data sample reduces to 16%, while 10% of ν_e events and 20% of ν_μ are now assigned flavour erroneously.



• **Figure 6.11:** Distributions of Λ_{46} for ν_e and ν_μ quasi-elastic and inelastic MCNO events that have passed the RINSE cuts.



• **Figure 6.12:** Cumulative distributions Λ_{46} for ν_e and ν_μ quasi-elastic, inelastic and all interactions (including neutral current) of MCNO events that have passed the RINSE cuts.



• **Figure 6.13:** Cumulative distributions of Λ_{46} for ν_e and ν_μ quasi-elastic, inelastic and all interactions (including neutral current) of MCNO events with 40 or more 3D hits that have passed the RINSE cuts.

6.2 Definition of Energy Dependent Flavour Cut

In this section a flavour cut separating ν_e and ν_μ events will be defined. This could simply be a cut at some value of Λ_{46} , such as the crossing point of 0.35. However, the Λ_{46} crossing point is energy dependent, as it has already been demonstrated by applying a 40 hit cut, and a simple cut on Λ_{46} would be a compromise. For this reason the event sample is divided into 7 energy bands according to the number of 3D hits in the event box. The crossing Λ_{46} value for each of these bands has been recorded into Table 6.3.

The flavour cut is simply defined by the number of hits and Λ_{46} . An event will be flagged as ν_e -like if its value of Λ_{46} is below the crossing value of its energy band. Conversely, ν_μ -like events are those with Λ_{46} above the crossing value of their energy band.

The flavour misidentification matrix is given in Table 6.4 for (i) all events and (ii) events with 40 or more hits in the event box. This method is an improvement on a universal cut on Λ_{46} of 0.35, where the mis-ID was 22% (Table 6.2).

Number of Hits	Λ_{46} Crossing Value	ν_e - ν_μ Mis-ID
11-20	0.5	0.32
21-30	0.4	0.25
31-40	0.33	0.18
41-60	0.27	0.15
61-80	0.2	0.14
81-100	0.15	0.14
Above 100	0.08	0.21

• **Table 6.3:** Flavour cut. Λ_{46} crossing value as a function of the number of hits.

All	Λ_{46} -Hits measured		
MC truth		ν_e	ν_μ
	ν_e	0.807	0.193
	ν_μ	0.217	0.783

40+	Λ_{46} -Hits measured		
MC truth		ν_e	ν_μ
	ν_e	0.878	0.122
	ν_μ	0.156	0.844

• **Table 6.4:** Misidentification matrix after applying the Λ_{46} -Hits flavour cut for all events (left) and after rejecting events with less than 40 Hits (right).

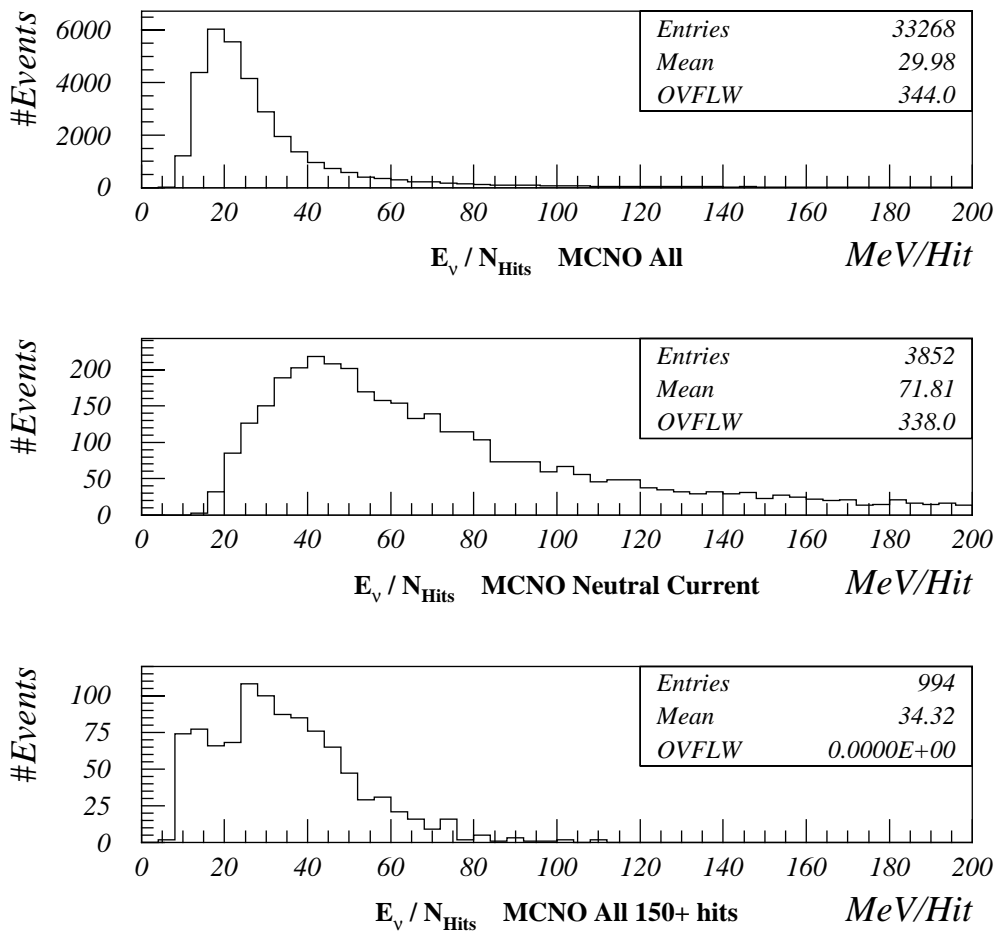
6.3 Neutrino Energy (E_ν) Estimator

Knowledge of the neutrino energy is important in the estimation of the neutrino oscillation phase, L/E_ν , which would give a handle in the determination of Δm^2 , if of course neutrinos do oscillate.

The most straightforward neutrino energy estimator is the number of hits in the event box. The Soudan 2 tracking calorimeter also offers the pulse-height of the reconstructed hits, but this proved not to offer additional information towards the estimation of the neutrino energy. Either way, hits or pulse-height measure the visible energy only, i.e. the energy that has been deposited through charged particles, and this may be far off the actual neutrino energy. In particular, we have no handle on E_ν for neutral current events. Also, the visible momentum of a lepton in a charged current interaction will be affected by the Fermi momentum of the struck nucleon or parton. Moreover, the energy resolution of the calorimeter is expected to differ for ionisation deposited by hadrons and ionisation deposited by leptons.

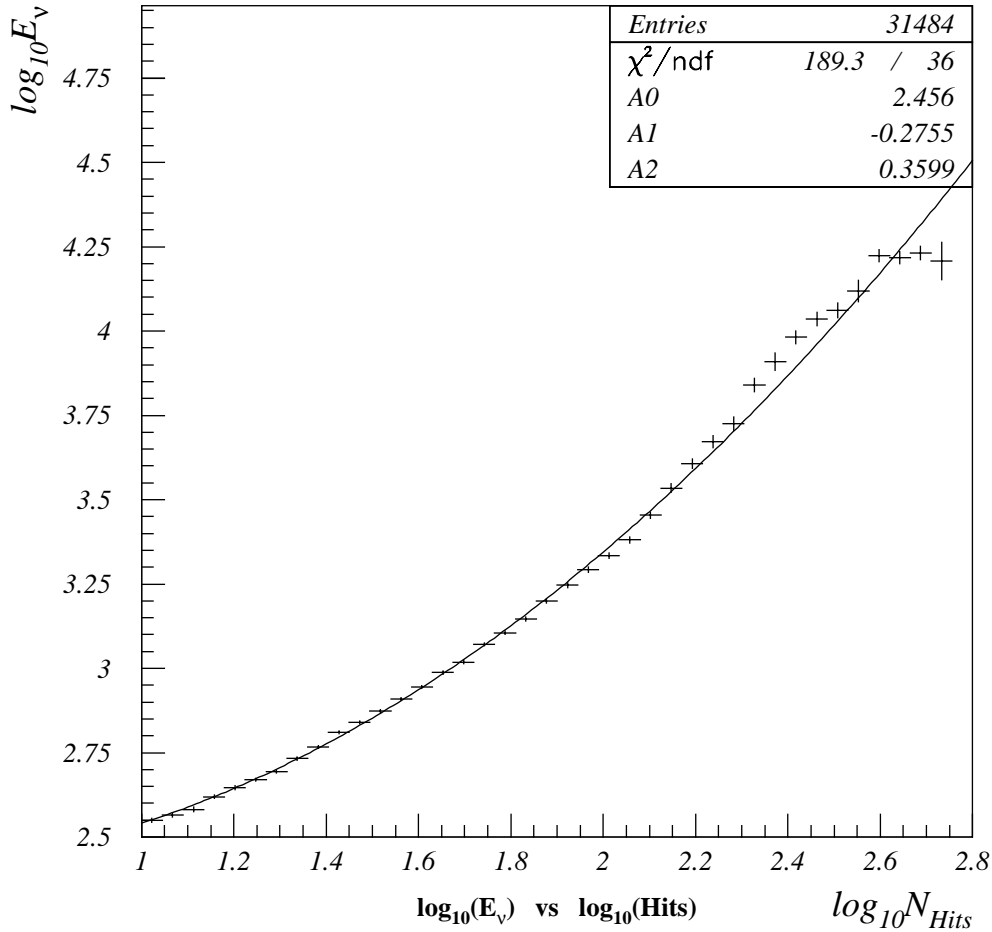
The above points illustrate the reasons why the estimation of neutrino energy is not a simple counting exercise. However, no bias is introduced in the analysis if the MC correctly models the Physics and MC data are treated in an identical fashion to the real data.

The number of 3D hits, N_{Hits} , in the event box is the starting point for the energy estimator. The MCNO histogram of the ratio of E_ν/N_{Hits} has mean value of ≈ 30 MeV/Hit and a very long tail, which is attributed to neutral current interactions (fig. 6.14). One could simply estimate the neutrino energy by multiplying the number of 3D hits in the box by the average value of ≈ 25 MeV/Hit, when disregarding the effect of the tail to the mean of the distribution. However, the ratio of E_ν/N_{Hits} may not be constant as a function of the number of 3D Hits.



• **Figure 6.14:** Ratio of the true neutrino energy to the number of 3D hits in the event box for all MCNO events that have passed the RINSE cuts (top) and for their subset which is due to neutral current interactions (middle) and for the subset of events with 150 or more hits in the event box.

One may go a step further and fit the profile distribution of E_ν versus N_{Hits} , as shown in fig. 6.15. The statistics of the MCNO data sample are scarce above 150 hits and above 2 GeV of true neutrino energy, which is why the logarithms (base 10) of E_ν and N_{Hits} are plotted, instead of the quantities themselves. A 2nd degree polynomial fit was conducted. Events of 150 or less hits and E_ν/N_{Hits} above 75 MeV/Hit were rejected from the calibration because they carry a very large weight, offsetting the mean of E_ν . Large events



• **Figure 6.15:** Profile histogram of the mean logarithm (base 10) of the neutrino energy versus the logarithm (base 10) of the number of 3D hits in the event box for MCNO events that have passed all RINSE cuts. A 2nd degree polynomial function has been fitted. The error bars correspond to the error of the mean value of neutrino energy and have been taken into account in the fits. Events in the tails of the E_ν/N_{Hits} distribution are neglected, as discussed in the text.

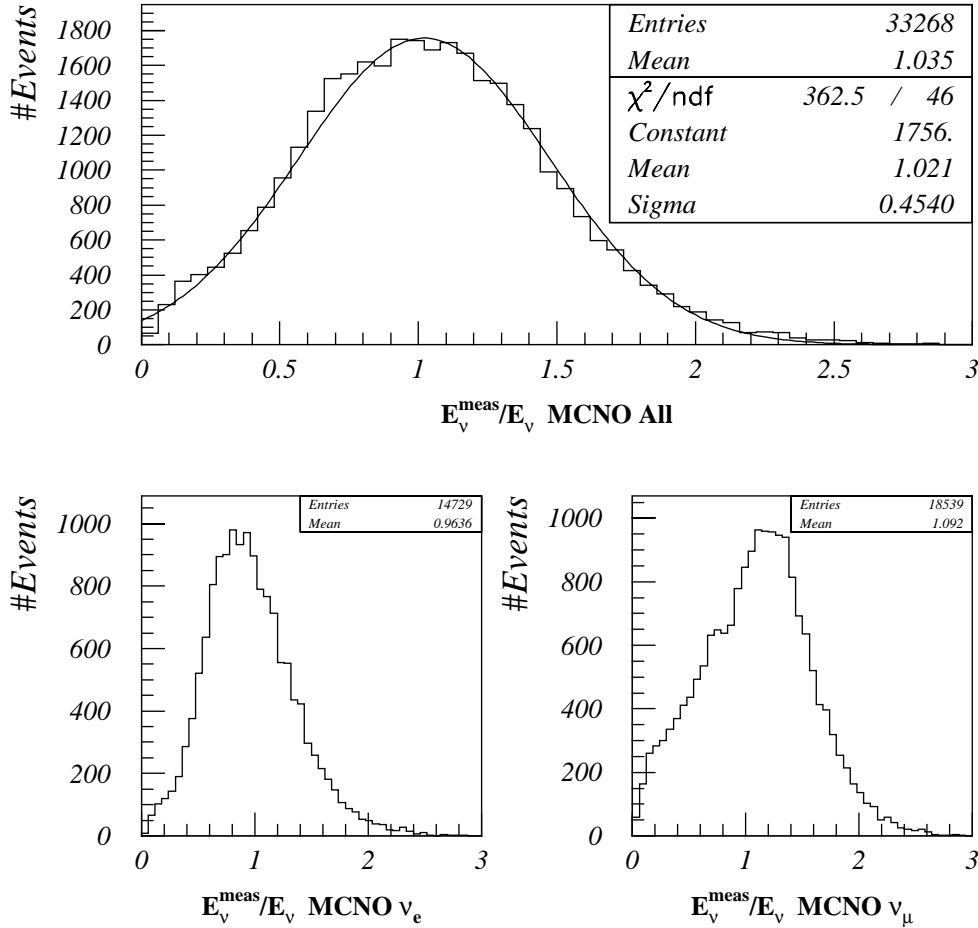
of more than 150 hits with E_ν/N_{Hits} above 150 MeV/Hit were similarly rejected. These cuts are justified by the distributions of fig. 6.14. The results of the fits are shown in Table 6.5 and will be used in estimating the neutrino energy in MeV:

$$E_\nu^{\text{meas}} = 10^f,$$

$$\text{where } f = A_0 + A_1 \times l + A_2 \times l^2, \quad (6.11)$$

$$\text{where } l = \log_{10}(N_{\text{Hits}})$$

The MCNO distribution of $E_\nu^{\text{meas}}/E_\nu$ can be fitted with a Gaussian of mean value 1.02 and



• **Figure 6.16:** Histogram of the ratio of measured neutrino energy (as given by eq. 6.11) over true neutrino energy for MCNO events that have passed all RINSE cuts (top) and for the ν_e and ν_μ MCNO subsets (bottom). The complete MCNO sample can be fitted with a Gaussian of mean value 1.02 and sigma 45% of the true energy.

	A_0	A_1	A_2
Value	2.456	-0.275	0.360
Error	0.797×10^{-2}	0.936×10^{-2}	0.292×10^{-2}

• **Table 6.5:** Results of neutrino energy fit as a function of N_{Hits} , as given by Eq. 6.11.

standard deviation equal to 45% of E_ν (fig. 6.16). This means that the neutrino energy is known to about half an order of magnitude. It is worth pointing out that the above method underestimates E_ν for ν_e interactions while it has the opposite effect for ν_μ interactions (fig. 6.16). This can be understood by the fact that the pulse-matching efficiency of PMT is better for tracks than for showers, because in the latter case there is greater ambiguity as to the pairing of anode and cathode multiplexed hits. Hence, showers have fewer 3D hits than tracks of the same energy. The effect is of the order of less than 10% and is not important in the final analysis: the largest uncertainty in the neutrino L/E_ν phase comes from the estimation of L , as it will be discussed next.

6.4 Angular Resolution of Neutrino Direction and Distance (L) Estimator

The distance, L , the neutrino has travelled to reach the detector is determined by its zenith angle, α , through the relationship $L(\alpha) = \sqrt{R^2 \cos^2 \alpha + 2Rd + d^2} - R \cos \alpha$, where R is the radius of the earth (6400km) and d is the distance vertically above the detector that neutrinos are produced, typically 20km (fig. 2.3 and Eq. 2.26). The direction and sense of travel of the neutrino, which determine its zenith angle, are therefore necessary in estimating L .

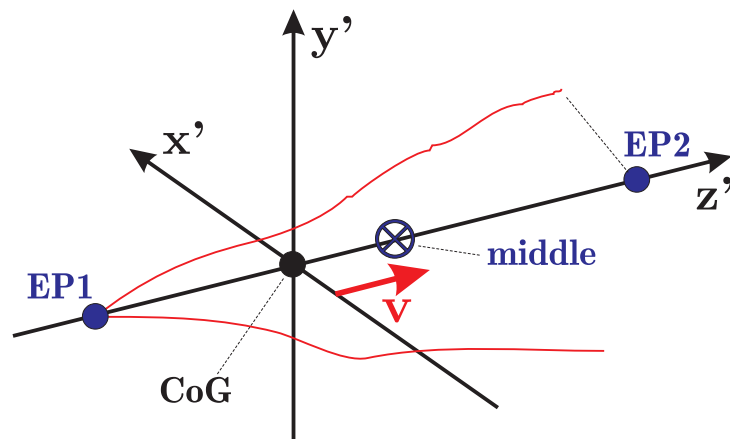
The event axis (see §3.3) is chosen to describe the direction of travel, but is ambiguous as far as the *sense* of travel is concerned. The latter is to be retrieved from other features of the event. We know that electromagnetic showers deposit more ionisation near their origin and less as they develop. Therefore, along the event axis the centre of gravity

(CoG) of the event will lie between the middle and the vertex of the event. This will also be the case in inelastic interactions where hadrons are produced at the vertex. For example, recoil protons that typically range out within 20 cm in the Soudan 2 detector [45] will result in a heavier event profile near the vertex.

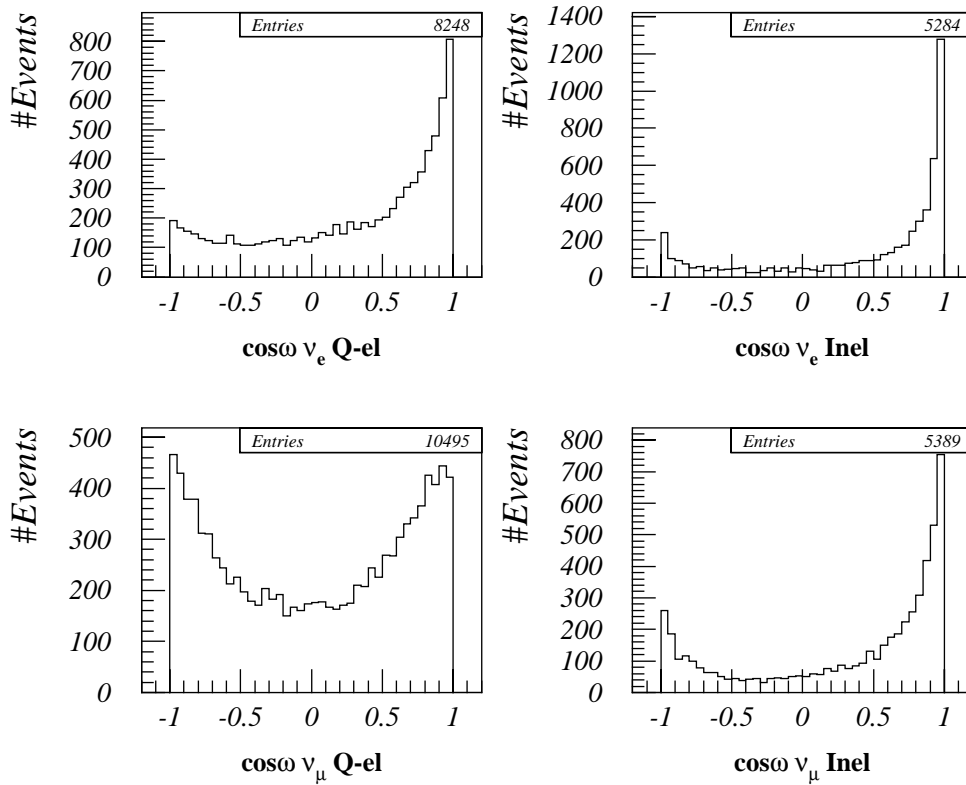
The above observation is exploited in defining the sense of direction of the event by the vector joining the centre of gravity of the event (see §3.3) to the *middle of the event*. The latter is the midpoint between the *two end points* (TEP), defined as the points on the event axis marked by projecting onto the event axis the first and last hits in the event box along the event axis itself (fig. 6.17). By definition the middle of the event lies on the event axis.

The angular resolution of the reconstructed neutrino direction is studied on the MCNO sample with and without a low hit (energy) cut of 40 3D hits (figs 6.18 and 6.19). A few points are worthy of attention:

- The distributions of the cosine of the angle between reconstructed and true



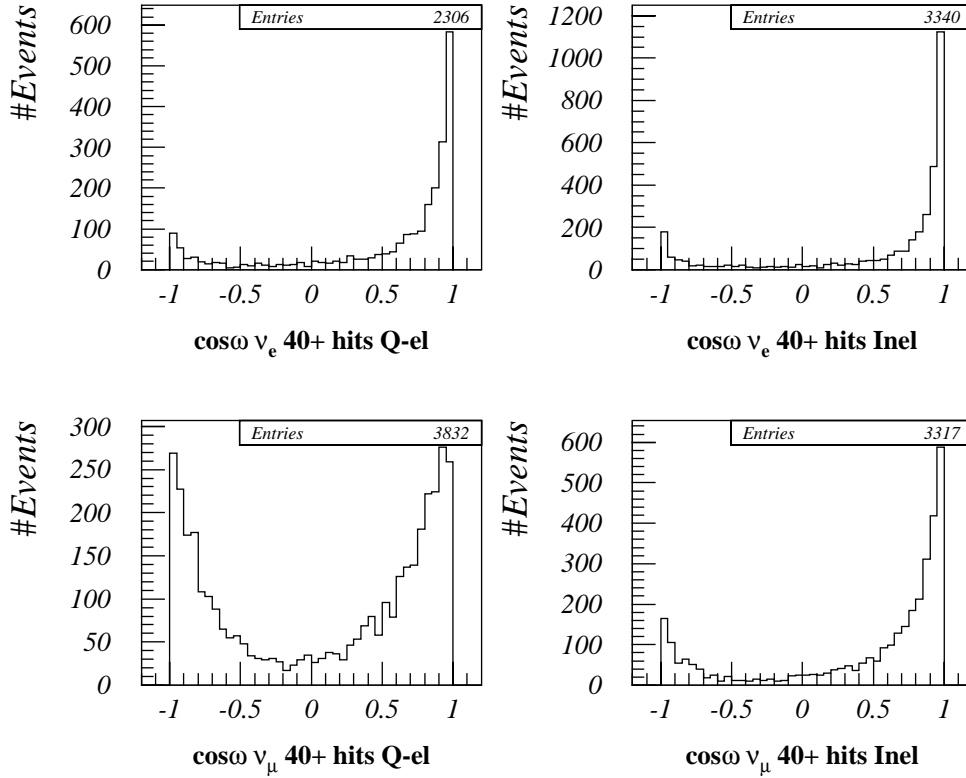
• **Figure 6.17:** Schematic illustrating the Two End Points algorithm. End Points EP1 and EP2 are marked by the projection on the event axis of the two extreme 3D hits in the event box along the event axis. The middle of the event is midway EP1 and EP2. The vector joining the Centre of Gravity (CoG) and the middle of the event defines the reconstructed neutrino direction of travel.



• **Figure 6.18:** Distributions of the cosine of the angle, ω , between the event axis given sense by the TEP algorithm and the true neutrino direction for MC quasi-elastic and inelastic ν_e and ν_μ interactions.

neutrino direction are narrower for ν_e than for ν_μ interactions. This is due to multiple scattering, a process which changes the path of muon tracks in one direction, but which does not affect the direction of travel of a shower. Hence, the initial direction of the lepton is preserved in ν_e interactions but not in ν_μ ones.

- The direction and sense of the incident neutrino vector can be well resolved for both quasi-elastic and inelastic ν_e interactions.
- Quasi-elastic ν_μ interactions offer no sense of neutrino direction because the TEP algorithm described above is not optimised for a straight track where the middle of the event may fall on either side of the centre of gravity of the event with equal probability. The same is not true for ν_μ inelastic processes however, where the



• **Figure 6.19:** Distributions of the cosine of the angle, ω , between the event axis given sense by the TEP algorithm and the true neutrino direction for MCNO quasi-elastic and inelastic ν_e and ν_μ events that have 40 or more hits and have passed the RINSE cuts.

extra hadronic activity at the event vertex will force the centre of gravity in that direction and away from the middle of the event. This results in a better resolution of the neutrino sense of direction for this class of events.

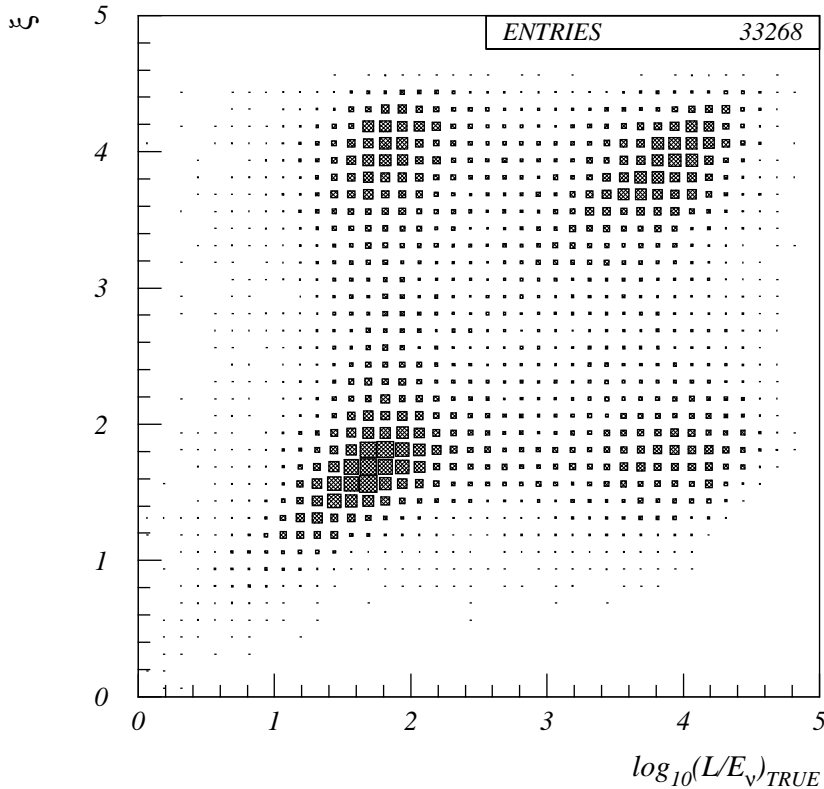
- The overall performance of the algorithm is improved for large events of 40 or more hits. The picture for ν_e interactions is very good, but for the ν_μ quasi-elastic it is still hopeless. However, the angular resolution for large inelastic ν_μ processes is good.
- On the other hand, events with less than 40 hits carry little directional information. This is mainly due to the effect of the Fermi momentum of the struck nucleon on the visible momentum of the event.

6.5 L/E_ν estimator

The L/E_ν estimator, ξ , is simply defined as the logarithm of the ratio of the corresponding estimators for L , given by eq. 2.26 and described in the previous section, and E_ν , as fitted in eq. 6.11 by the number of 3D hits in the event box:

$$\xi = \log_{10} \left(\frac{L(\alpha)}{E_\nu(\text{hits})} \right) \quad (6.12)$$

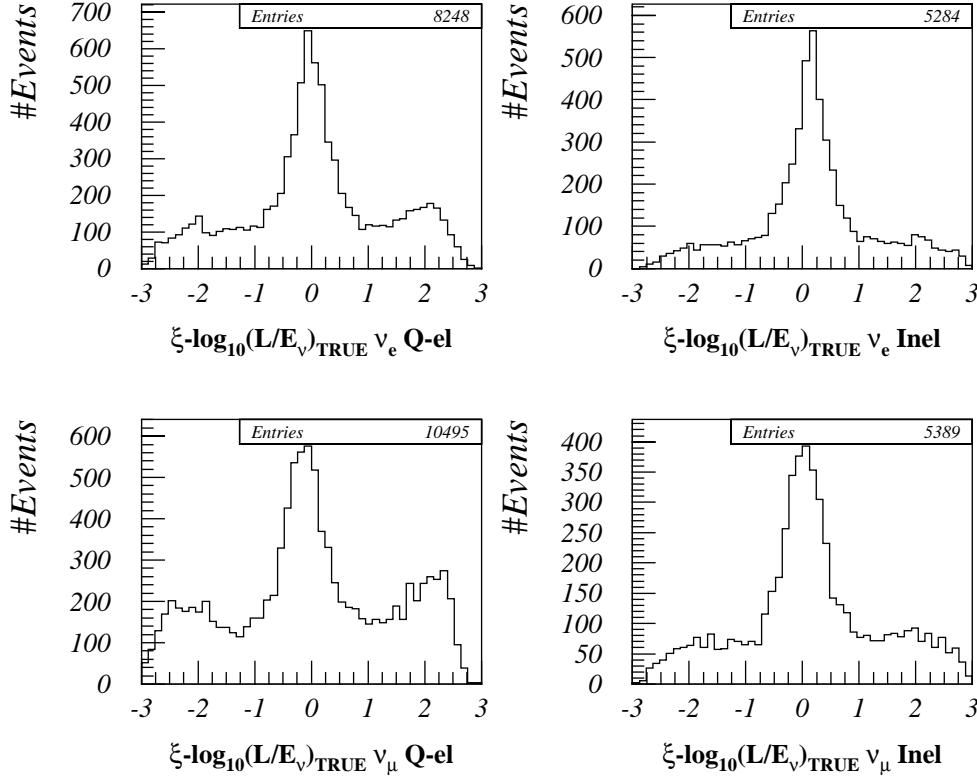
The ξ vs $\log_{10}(L/E_\nu)_{\text{truth}}$ scatter plot for all MCNO data (fig. 6.20) demonstrates four bumps, two along the one-to-one diagonal, which correspond to the measurement and the truth agreeing within one order of magnitude, and the other two along the anti-correlation



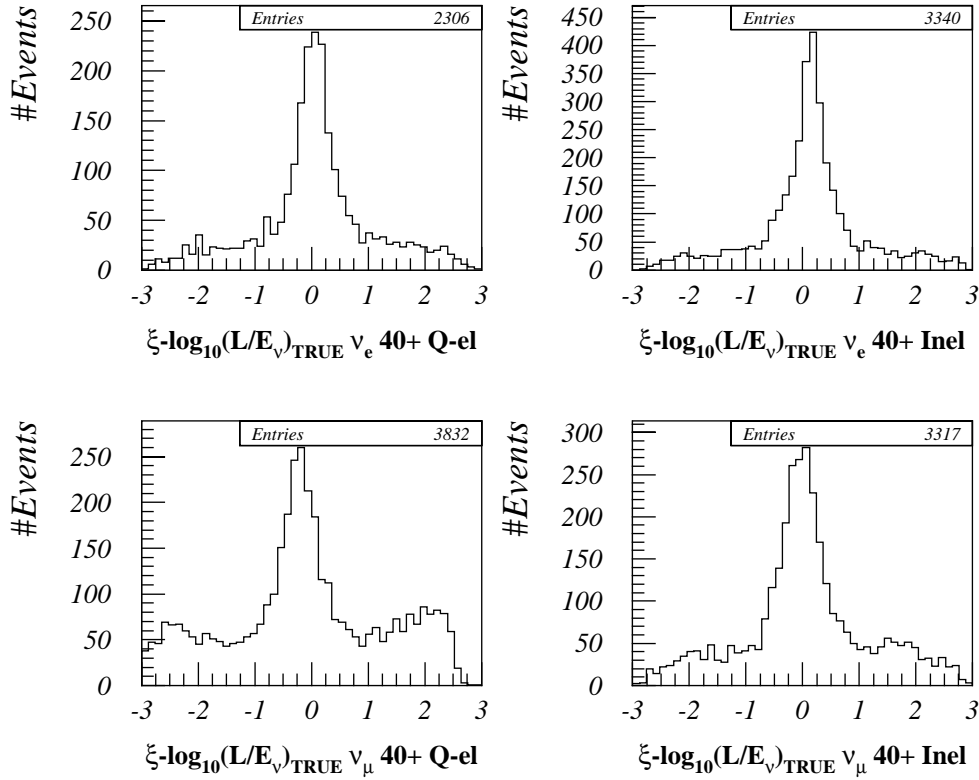
• **Figure 6.20:** Scatter plot of ξ versus $\log_{10}(L/E_\nu)_{\text{truth}}$ for MCNO atmospheric neutrino events that have passed the RINSE cuts. The plot is presented in ‘box’ form because of the very high statistics of this event sample. The population of each bin is proportional to the linear dimension of the box and not its area. L/E_ν is measured in km/GeV.

diagonal, where the neutrino was assigned a sense of travel opposite to the truth. However, it is not simple to extract quantitative conclusions from the scatter plot and the projection $\xi - (L/E_\nu)_{\text{truth}}$ is used instead (fig. 6.21).

As expected from the individual behaviour of L and E_ν estimators examined in previous sections, ξ correlates well with $\log_{10}(L/E_\nu)_{\text{truth}}$ for ν_e events, where the two small side-peaks, corresponding to the “ghost” bumps in the scatter plot, are low for both quasi-elastic and inelastic interactions. Inelastic ν_μ interactions are also resolved, but not as well, whereas the sum of the side-peaks is equal to the central peak in the quasi-elastic ν_μ processes which offer no L/E_ν information whatsoever. The central peak in all cases has width at half maximum of about one decade, which is the L/E_ν resolution reached with these simple algorithms. The resolution of the peak does not improve by applying a low



• **Figure 6.21** Distributions of $\log_{10}(L/E_\nu)_{\text{truth}}$ for MCNO quasi-elastic and inelastic ν_e and ν_μ events that have passed the RINSE cuts. L/E_ν is measured in km/GeV.



• **Figure 6.22** Distributions of $\log_{10}\left(\frac{L}{E_{\nu}}\right)_{\text{truth}}$ for MCNO quasi-elastic and inelastic ν_e and ν_{μ} events of 40 or more hits that have passed the RINSE cuts. L/E_{ν} is measured in km/GeV.

energy cut of less than 40 hits but the size of the side bumps decreases (fig. 6.22).

Finally, the underestimation and overestimation of the neutrino energy for ν_e and ν_{μ} interactions respectively, as observed in fig. 6.16, is also evident here, producing a less than 10% shift of the $\xi - \left(\frac{L}{E_{\nu}}\right)_{\text{truth}}$ distributions about the central value of zero, which is small compared to the L/E_{ν} resolution of one decade. The shift in the energy estimation between the two neutrino flavours is therefore not important in the subsequent L/E_{ν} analysis.

6.6 Conclusions

It has been demonstrated that the flavour of interactions of atmospheric neutrino data in the Soudan 2 detector can be statistically separated by quantities such as Γ and Λ_{46} . This is particularly true for high energy events.

Γ was historically the first attempt in defining a flavour estimator. Its separation efficiency is lower than that of Λ_{46} . However, Γ is still of value, because its operation is orthogonal to that of Λ_{46} . The latter is based on the microstructure of the event and the relative positions of neighbouring hits, while Γ looks at the shape of the event as a whole. The data analysis that follows will predominantly use Λ_{46} and the energy dependent flavour cut that is based on it. On the other hand, Γ will be used to qualitatively support the conclusions drawn by Λ_{46} .

A cut separating neutrino flavours has been defined by applying an energy dependent cut on Λ_{46} . Its identification efficiency is 81% for ν_e and 78% for ν_μ interactions when the complete data sample is analysed, i.e. including single-prong and multi-prong and neutral current events, unlike in the Soudan 2 scanning analysis.

The neutrino direction can be determined within 40° for ν_e interactions but its sense is typically ambiguous for ν_μ interactions, unless associated hadronic activity allows for the determination of the vertex .

The performance of both flavour and L/E_ν estimators improves for large events, typically above 40 hits (≈ 850 MeV). Moreover, the background contamination of the **GOLD** data sample is significantly reduced at high energies, as it has been pointed out in the previous chapter. Therefore, the large event sample may be a good place to look for new Physics effects. Incidentally, this is where the Kamiokande group has reported on a zenith angle distribution that favoured the neutrino oscillation hypothesis. However, this has not been the case for the later Super-Kamiokande data, which exhibits the same effect at lower energies.