Track Reconstruction and Point Source Searches with Antares

Heijboer, A.J.

Citation for published version (APA):
Chapter 4

Muon Track Reconstruction

In this chapter, the algorithms used to estimate the direction, position and energy of the muon from the arrival times and amplitudes of the hits are discussed. The accuracy of the reconstruction of the muon direction determines the pointing accuracy of the detector and is therefore a crucial parameter for searches for point sources of neutrinos.

The reconstruction algorithm consists of four consecutive fitting procedures. The last procedure produces the most accurate result, but requires a priori estimates of the muon track parameters that should be close to the true values. The purpose of the first stages in the chain of fitting procedures is to provide this starting point.

In section 4.1, the variables used to describe the muon track and its relation to the OMs are introduced. The subsequent sections describe the four track fitting methods. In section 4.6, it will be described how these methods are combined to form the full reconstruction algorithm. The performance of the full algorithm is discussed in section 4.7. The energy of the muon is determined by a separate procedure, which will be discussed in section 4.8.

4.1 Track description and relation to the OM

The muon trajectory can be characterised by the direction $\vec{d} = (d_x, d_y, d_z)$ and the position $\vec{p} = (p_x, p_y, p_z)$ of the muon at some fixed time $t_0$. At energies above the detection threshold (10 GeV or so) the muon is relativistic. Hence, the speed of the muon is taken to be equal to the speed of light in vacuum. The direction can be parameterised in terms of the azimuth and zenith angles $\theta$ and $\phi$: $\vec{d} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$. There are thus five independent parameters that are estimated by the reconstruction algorithm. For a given track (i.e. a given $\vec{d}$ and $\vec{p}$) and an OM at position $\vec{q}$, who’s field of view is oriented in the direction $\vec{w}$, the relevant properties of a Cherenkov photon emitted from the muon track are:

- the expected (theoretical) arrival time of the photon ($t^h$).
- the expected photon path length ($b$) and
- the expected cosine of the angle of incidence of the photon on the OM ($a$), i.e. the angle between the direction of the photon and the pointing direction of the PMT.
4.1 Track description and relation to the OM

Muon Track Reconstruction

Figure 4.1: Description of the geometry of the detection of the Cherenkov light. The muon goes through point $\vec{p}$ in the direction $\vec{d}$. The Cherenkov light is emitted at an angle $\theta_C$ with respect to the muon track and is detected by an OM located in point $\vec{q}$. The dashed line indicates the path of the light.

These three quantities completely\(^1\) characterise the position and orientation of the OM relative to the track. They are calculated under the assumption that the light is emitted under the Cherenkov angle w.r.t. the muon and travels in a straight line to the OM. The true arrival time, path length and angle of incidence may differ from these values, since photons are also emitted from secondary electrons and their path is influenced by scattering. The expected time of arrival $t^{th}$ is used in all reconstruction algorithms. The quantities $a$ and $b$ will be used to predict the number of detected photons in sections 4.4 and 4.8.

In order to calculate $t^{th}$, we first define (see figure 4.1)

$$\vec{v} = \vec{q} - \vec{p}. \tag{4.1}$$

The components of $\vec{v}$ parallel and perpendicular to the muon direction are $l = \vec{v} \cdot \vec{d}$ and $k = \sqrt{\vec{v}^2 - l^2}$ respectively. The arrival time of the light in $\vec{q}$ is then given by

$$t^{th} = t_0 + \frac{1}{c} \left( l - \frac{k}{\tan \theta_C} \right) + \frac{1}{v_g} \left( \frac{k}{\sin \theta_C} \right), \tag{4.2}$$

where $v_g$ is the group velocity of light, for which we take the value at 460 nm. The second term is the time it takes for the muon to reach the point where the detected light is

\(^1\)The position of the OM w.r.t. the track has two degrees of freedom, since the situation is symmetric for rotations around $\vec{d}$. The orientation of the OM gives only one degree of freedom, since the response of the OM is (assumed to be) invariant under rotations around $\vec{v}$.
emitted, while the third term is the time it takes the light to travel from that point to \( \vec{q} \). The length of the photon path is given by

\[
b = \frac{k}{\sin \theta_C}.
\]  

(4.3)

The cosine of the angle of incidence of the photon on the OM is given by

\[
a = \left( \vec{v} - d \left( l - \frac{k}{\tan \theta_C} \right) \right) \cdot \vec{w}.
\]  

(4.4)

where \( \vec{w} \) is the pointing direction of the OM. For a head-on collision of a photon with the photocathode \( a = -1 \), whereas \( a = 1 \) means the photon hits the insensitive rear of the OM.

4.2 Linear prefit

The first stage in the track reconstruction procedure is the 'linear prefit'. This is a linear fit through the positions of the hits, with the hit time as independent variable. The position associated with the \( i \)th hit is denoted by \((x_i, y_i, z_i)\). In order to obtain a linear relation between the hit positions and the track parameters, it is assumed that the hits occur on points that are located on the muon track. This is expected to be a reasonable approximation if the length of the muon track in the detector is much larger than the attenuation length of the light. Under this approximation, the following relation holds:

\[
y = H \Theta,
\]

(4.5)

where \( y \) is a vector containing the hit positions, \( y = [x_1, y_1, \ldots, z_n] \), \( \Theta \) is a vector containing the track parameters: \( \Theta = [p_x, d_x, p_y, d_y, p_z, d_z]^T \). These two vectors are related by a matrix containing the hit times:

\[
H = \begin{bmatrix}
1 & c t_1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & c t_1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & c t_1 \\
0 & 0 & 1 & c t_2 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & c t_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 1 & c t_n \\
\end{bmatrix}.
\]

(4.6)

The error estimates on the hit positions are stored in the covariance matrix \( V \), the inverse of which is called \( F \). Uncertainties on the hit times are neglected.

The estimate of the track parameters, \( \hat{\Theta} \), is given by those parameters that minimise the \( \chi^2 \):

\[
\chi^2 = |y - H \hat{\Theta}|^T V^{-1} |y - H \hat{\Theta}|.
\]

(4.7)

This yields

\[
\hat{\Theta} = (H^T V^{-1} H)^{-1} H^T V^{-1} y.
\]

(4.8)
It is now assumed that the covariance matrix is diagonal (i.e. that all errors are uncorrelated) and that the uncertainties on the \(x\), \(y\) and \(z\) components of the position of an individual hit are equal. In this case, \(\mathbf{V}\) and \(\mathbf{V}^{-1}\) are diagonal and equation 4.8 can be expressed as

\[
\hat{\Theta} = \frac{1}{\sum_i f_i\sum_i f_it_i^2 - (\sum_i f_{i\text{ct}})^2} \left[ \sum_i f_i^2t_i^2 \sum_i f_{i\text{ct}x}\sum_i f_i \sum_i f_{i\text{ct}x} - \sum_i f_i^2 \sum_i f_{i\text{ct}x} \sum_i f_i \sum_i f_{i\text{ct}x} \right] 
\]

where \(f_i^{-1}\) is the uncertainty on the position of hit \(i\), and \(t_i\) is the time of that hit.

### 4.2.1 Hit positions

In the previous section it was not specified how the positions associated with the hits are determined. The standard method [67, 68] is simply to use the positions of the OMs on which the hits occur. However, the assumption used in equation 4.5 is that the hit positions are located on the muon track, which is not the case if the OM positions are used. In order to improve the track estimate, a different approach was taken. Here, we try to determine, for each hit, a position which is most likely to be located on the track. If a hit occurs on a given OM, it can be expected that the track passes some distance in front of that OM. This distance is estimated from the amplitude of the hit.

Using simulated events, the spatial distribution of the point of closest approach of the hit to the track is determined as a function of the amplitude of the hit. Figure 4.2a shows this distribution for hits with an amplitude between 0.5 and 1.5 photo-electron (p.e.). The maximum of this distribution is located along the pointing direction of the PMT. The position of the maximum can therefore be characterised by a distance from the OM. The approximate spherical symmetry of the distribution implies that the uncertainties on the \(x\), \(y\) and \(z\)-coordinates of the positions are almost uncorrelated and equally large. They are assumed to be equal to the RMS of the distribution.

Figure 4.2b shows the average distance between the muon track and the PMT, as well as its uncertainty, as a function of the hit amplitude. As expected, hits with high amplitudes are likely to have passed the PMT at close distance, while for low amplitude hits, the position of closest approach is much more uncertain and, on average, further away from the PMT.

### 4.2.2 Constraining the muon velocity

When using equation 4.9 to estimate the direction of the muon, its velocity, \(c \sqrt{d_x^2 + d_y^2 + d_z^2}\), is left free. An attempt was made to improve the estimates of the track parameters by constraining the velocity of the muon to the speed of light in vacuum (\(c\)) using the technique of Lagrangian multipliers. The analytic calculation, which was made under the same assumptions about the covariance matrix that were made in the previous section, shows that the inclusion of this constraint does indeed normalise the velocity to \(c\), but
Figure 4.2: a) The distribution of the points of closest approach of the muon to a PMT in a coordinate system where the positive $z'$-axis points along the orientation $\hat{w}$ of the PMT and where the $x'$-axis is perpendicular to both the $z'$-axis and the vertical. The distribution shown is for hits with an amplitude between 0.5 and 1.5 p.e. b) The average distance of the point of closest approach of the muon track to the OM as a function of hit amplitude. The error bars indicate the spread of this position.

that the estimated direction of the muon is not affected by this constraint. It is therefore easier to use equation 4.9 and to set the velocity afterwards.

4.2.3 Performance

Throughout this chapter, the performance of the reconstruction algorithms will be summarised by the distribution of the reconstruction error, $\alpha_\mu$, which is defined as the angle between the true direction of the muon and the direction of the reconstructed track. Simulated through-going muons are used with an $E_\mu^{-1}$ energy spectrum in the range between 100 GeV and 100 TeV. In order to evaluate the performance of the individual fitting routines, the fits will be performed using hits that occur on OMs within a distance of 150 m from the true muon track and that have time residuals between -150 and 150 ns as calculated using the true muon track.

The distribution of $\alpha_\mu$ is shown in figure 4.3 for the results of the linear prefit with the improved estimation of the hit positions. For comparison, the performance is also shown for the standard method for determination of the hit positions. Compared to the standard version, a small improvement is achieved: the number of tracks which are reconstructed with an error on the direction smaller than $10^\circ$ ($5^\circ$) is increased by about 15% (30%).

4.3 Maximum likelihood fit

Two of the fit procedures used rely on the principle of Maximum Likelihood (ML). For each possible set of track parameters, the probability to obtain the observed events can
be calculated. This probability is called the likelihood of the event. The likelihoods that are used here, only take into account the probability of the time of the hits. In case of uncorrelated hits, the likelihood of the event can be written as the product of the likelihood of the individual hits:

$$P(\text{event}|\text{track}) \equiv P(\text{hits}|\vec{p}, \vec{d}) = \prod_i P(t_i|t_i^{th}, a_i, b_i, A_i), \quad (4.10)$$

where $t_i$ is the time of hit $i$, $A_i$ the hit amplitude and $t_i^{th}$, $a_i$ and $b_i$ were introduced in section 4.1. The ML estimate of the track is defined by the set of track parameters for which the value of the likelihood function is maximal.

Since a full search of the five dimensional parameter space is prohibitively time consuming, standard numerical tools [69] are used to find the maximum of the likelihood function\(^2\). These tools iteratively approximate the maximum by using information on the gradient of the likelihood function. If the likelihood function has multiple maxima, the maximisation algorithm will, in general, not find the global maximum, but converge on a local maximum. The efficiency for finding the global maximum is therefore related to the shape (i.e. the gradient) of the likelihood function and to the quality of the starting point that is used for the fit.

### 4.3.1 Original PDF

In the simplest case, the $a$, $A$ and $b$ dependence in equation 4.10 are neglected and the likelihood is expressed solely in terms of the probability density of the residuals $r_i = t_i - t_i^{th}$. The Probability Density Function (PDF) that is described in this section was developed for the first reconstruction algorithm used in ANTARES [68, 70]. This 'original' PDF is shown in figure 4.4. When comparing this with figure 3.12, it is clear that background hits are not taken into account in this PDF. Instead, a non-physical tail is present for residuals $r < -5$ ns. This tail was included in order to provide a gradient in the likelihood function for negative $r$, which helps the maximisation routine to converge.

\(^2\)In practice, the maximum of L is found by minimising $-\log(L)$
4.3.2 Performance

The performance of the ML estimator has been evaluated in the same way as was done for the linear prefit. Distributions of $\alpha_\mu$ obtained with the original PDF are shown in figure 4.5. To evaluate the dependence on the quality of the track estimate that is used as a starting point for the maximisation procedure, the fit has been performed using the true muon track and random tracks that make an angle $\alpha_0$ of 1°, 10°, and 45° with the true muon track. As expected, the quality of the fit result is degraded when the value of $\alpha_0$ is increased.
4.4 Maximum likelihood fit with improved PDF

As was already mentioned, the original PDF introduced in the previous section does not include background hits. As a result, background hits, especially those with negative residuals, can degrade the performance of the track reconstruction. In this section, a new PDF is presented in which background hits are taken into account.

4.4.1 PDF with background

The background hits are uniformly distributed in time. As a consequence, the PDF of the hit residuals \( P(r) \) can only be normalised if the duration of the event is specified. It is assumed that an event consists of all hits with residuals between \(-T/2\) and \(T/2\). The event duration \( T \) is assumed to be large enough to contain all signal hits.

The PDF has contributions from background hits and from signal hits. The PDF of the signal depends on the amplitude of the hit \( A \): \( P_{\text{sig}}(r|A) \).

The relative contributions from these two types of hits are determined by the expected number of signal and background hits, which are determined from the hit amplitude, and the parameters \( a \) and \( b \):

\[
P(r_i|a_i, b_i, A_i) = \frac{1}{N^T(a_i, b_i, A_i)} [P_{\text{sig}}(r_i|A_i) N_{\text{sig}}(a_i, b_i, A_i) + R^{\text{bg}}(A_i)], \tag{4.11}
\]

where \( R^{\text{bg}}(A_i) \) is the background rate\(^3\) for hits with amplitude \( A_i \) and \( N_{\text{sig}}(a_i, b_i, A_i) \) is the expected number of signal hits as a function of the hit amplitude and \( a \) and \( b \). Finally \( N^T(a_i, b_i, A_i) \) is the total expected number of hits of amplitude \( A_i \) in the event (of duration \( T \)), which is given by

\[
N^T(a_i, b_i, A_i) = N_{\text{sig}}(a_i, b_i, A_i) + R^{\text{bg}}(A_i)T. \tag{4.12}
\]

The factor \( 1/N^T(a_i, b_i, A_i) \) ensures that \( \int_{-T/2}^{T/2} P(r)dr = 1 \) if \( \int_{-T/2}^{T/2} P_{\text{sig}}(r_i)dr = 1 \). The PDF is thus normalised for all values of \( A \), \( a \) and \( b \), provided \( T \) is known. The relative contributions of the signal and background terms in equation 4.11 are independent of the event duration \( T \). However, if the value of \( T \) is incorrect (i.e. if the fit is performed with hits from a larger time window than what is assumed in equation 4.11), the normalisation of the PDF will depend on \( a \) and \( b \) and thus on the track parameters, which is undesirable. The knowledge of \( T \) is thus needed in the reconstruction.

4.4.2 Parameterisation of the PDF

In order to use standard software tools \([69]\) to maximise the likelihood function, it is convenient to parameterise the likelihood function with a continuous, differentiable function. The parameterisation is obtained by fitting a set of histograms obtained from Monte Carlo simulations of muons traversing the detector. For this, the muons have an energy between 100 GeV and 100 TeV and have a spectrum \( \frac{d^2N}{dE}\propto E^{-1} \) (an \( E^{-1} \) spectrum, for short). The\(^3\) Equation (4.11) looks more symmetric for signal and background when \( R^{\text{bg}}(A_i) \) is substituted by \( N^{bg} P^{bg} \), with \( N^{bg} = R^{bg}T \) and \( P^{bg} = 1/T \).
obtained distributions, and the parameterisations obtained from them, correspond to a (weighted) average over this energy range.

The likelihood function was parameterised separately for hits of different amplitudes. In order to obtain enough statistics, the hits were classified into 5 amplitude bins (with boundaries 0, 1.5, 2.5, 5, 10 and $\infty$ p.e.).

**Time dependence**

Figure 4.6 shows the PDF of the time residuals of the signal hits for the different amplitude bins, $P_{\text{sig}}(r|A_i)$, as obtained from the simulation. These distributions have been parameterised using the following functional form:

$$
\frac{dP_{\text{sig}}}{dr} = \begin{cases} 
    A e^{-\frac{(r-r_i)^2}{2\sigma^2}} & \text{if } r < c_1 \\
    B (\alpha r^3 + \beta r^2 + \gamma r + 1) & \text{if } c_1 < r < c_2 \\
    C \frac{e^{-r/c}}{r_+p} & \text{if } r > c_2.
\end{cases}
$$

The peak of the distribution is fitted with a Gaussian function, while the tail has been approximated by the function $\frac{e^{-r/c}}{r_+p}$. These two functions are joined together by a 3rd degree polynomial function. The coefficients of this polynomial are fixed by imposing that the function is continuous and differentiable at the points $r = c_1$ and $r = c_2$. The values of $c_1$ and $c_2$ have been left free in the fit. The parameterisations thus obtained are also shown in figure 4.6. The peaks at $r = 25$ ns that result from the integration of the PMT signal (see section 3.3.2) have not been incorporated in the parameterisation of the PDF. This was done in order to keep the parameterisation as simple as possible and to reduce the possibility of local minima in the likelihood function. The values of the parameters obtained with the fits are shown in table 4.1. It can be seen that, with increasing hit amplitude, the position of the peak shifts to smaller values of the time residual, while its width decreases.

**Number of signal and background hits**

As explained in section 4.4.1, the expected number of signal hits $N_{\text{sig}}$ of a given amplitude, is estimated from the parameters $a$ and $b$. For simplicity, it was assumed that this dependence can be factorised:

$$
N_{\text{sig}}(a, b) = N_{\text{sig}}(b) \times f(a),
$$

Table 4.1: Values used in the parameterisation of the PDF of the hit residuals (see equation 4.13).
Figure 4.6: Distribution of the residuals for different hit amplitudes $A$ and the fitted function that parameterises the distribution. The insets show the same distribution on a linear scale.
where $N_{\text{sig}}(b)$ is the expected number of signal hits as a function of the distance and $f(a)$ describes the dependence on the angle of incidence of the photon on the PMT. The first has been parameterised by the following function:

$$N_{\text{sig}}(b) = \frac{(b + v)e^{\zeta + \xi b}}{b^\psi + \chi},$$

(4.15)

where the values of $v, \zeta, \xi, \psi$ and $\chi$ have been determined by a fit to the simulated events. This function is an extension of the simple function introduced in section 3.4, which does not take into account the readout electronics. The additional factor $(b + v)$ and the additional parameters $\psi$ and $\chi$ allow to incorporate the effects of the simulation of the electronics for all amplitudes. The function $f(a)$ resembles the angular acceptance function of the PMT that was shown in figure 3.10. It is parameterised by a 4th order polynomial. The parameterisations of the functions $N_{\text{sig}}(b)$ and $f(a)$ are obtained from simulations. The results of the simulations are shown in figure 4.7 together with the obtained parameterisations.

Figure 4.7: Left: Number of expected hits as a function of the distance travelled by the photon for different hit amplitudes. The number of background hits detected in a 250 ns time window is also shown. Right: Dependence of the number of hits on the angle of incidence of a photon, emitted under $\theta_C$ for different hit amplitudes. The smooth curves are the parameterisations (see text). For clarity, curves are only shown for three of the five amplitude bins.

The rate of background hits, $R_{\text{bg}}(A)$ is independent of both $a$ and $b$, but it does depend on the amplitude $A$ of the hit. The optical background consists mostly of single photo-electron hits. Due to the resolution of the amplitude measurement of about 30%, the measured amplitude of single p.e. hits will not be precisely 1 p.e. Hits with a measured amplitude below 2 p.e. are assumed to be single photo-electron hits; the background
rate used for these hits in equation 4.11 is $R^{bg}(1) = 60$ kHz. Background hits with amplitudes above 2 p.e. are assumed to be caused by multiple background photons that, by coincidence, are detected within the 25 ns integration time of the ARS. The expected rate $R(A)$ of background hits with amplitude $A > 2$ p.e. is computed by multiplying the single p.e. rate $R(1)$ with the chance probability of detecting $A - 1$ additional background photons within the 25 ns integration time:

$$R^{bg}(A) = R^{bg}(1) \times (25 \text{ ns} \times R^{bg}(1))^{A-1}. \quad (4.16)$$

4.4.3 Performance

The performance of the ML fit with the improved PDF has been evaluated in the same way as the ML fit with the original PDF (section 4.3.1). As a starting point of the fit, tracks were used that were taken at different angles $\alpha_0$ from the true track. The results are shown in figure 4.8. When the fit is started close to the true track ($\alpha_0 = 0^\circ$ or $1^\circ$), the reconstruction error is small (typically $0.3^\circ$). A large improvement is observed compared to the original likelihood fit (see figure 4.5). For larger values of $\alpha_0$, however, the correct solution is less likely to be found. Instead, a local maximum in the likelihood function is found, which is close to the starting point. In some events, the gradient of the likelihood function at the starting point is too small for the maximisation algorithm to start.

![Figure 4.8: Distribution of the reconstruction error $\alpha_0$ obtained from the ML fit with the improved PDF. The fits have been started with track parameters that were taken at different angles $\alpha_0$ from the true muon track.](image)

The strong sensitivity on the quality of the starting points can be understood by the weak dependence of the PDF on large time residuals. If the starting point is bad, the gradient of the likelihood function will be very small or may be dominated by only a small number of (background) hits, which happen to have a small residual with respect to the track used as a starting point. In such cases a local maximum can be found. In conclusion, the ML estimate with the improved PDF presented in this section leads to
accurate track reconstruction, provided a good starting point for the fit is given.

4.5 M-estimator

In the previous section, it was found that the ML estimate is very sensitive to the quality of the track estimate that is used as a starting point. The same is true for the ML estimate using the original PDF discussed in section 4.3. This has been known in ANTARES for a long time and it was the reason to use the linear pre-fit in previous fitting algorithms. In addition, the non-physical tail of the original PDF was added in order to improve the efficiency for finding the correct maximum in the likelihood function.

There seems to be a trade-off between the accuracy of the reconstruction and the efficiency of finding the global maximum of the likelihood function. The latter is expected to be large if the derivative of the fitted function is non-zero for large residuals $r$. In this section, we therefore abandon the notion that the fit function is only meant to describe the data. Instead the behaviour for large $r$ should be such that there is a favourable trade-off between an accurate description of the data and the efficiency of finding the global maximum.

In the literature [71], estimators that work by maximising some function $g$ (in our case $g(r)$), are known as M-estimators. In a sense, the maximum likelihood estimator is a special case of an M-estimator ($g(r) = P(r)$), as is the least squares estimator ($g(r) = -r^2$). Other, more ad hoc, forms of $g$ are often used for so-called robust estimation, i.e. estimation which is insensitive to large fluctuations in a small number of data points. The data points with large fluctuations are usually called 'outliers'. In the case of track fitting in ANTARES, a robust fit is already available, namely the improved ML estimator, which incorporates the outliers (i.e. the background hits) in the PDF. Nevertheless, by choosing an M-estimator that behaves suitably for large residuals, it can be expected that a reasonable track estimate can be obtained without the requirement of an accurate starting point. It has been found that the following function gives good results:

$$g(r) = -2\sqrt{1 + r^2/2} + 2.$$  \hspace{1cm} (4.17)

The resulting M-estimator is called 'L1-L2' [72]. For large values of $r$, this function is linear in $r$, while for small $r$ it is quadratic.

The performance of the M-estimator was found to improve when the hit amplitude of the hits is used as a weighting factor for the time residuals. Furthermore, a term was added, which contains the angular response function $f_{\text{ang}}(a)$ of the optical module (see section 3.3.2). The function that is finally maximised to obtain a track estimate is given by

$$G = \sum_i \kappa (-2\sqrt{1 + Ar_i^2/2}) - (1 - \kappa)f_{\text{ang}}(a_i).$$  \hspace{1cm} (4.18)

The relative weights of the two terms are determined by the parameter $\kappa$, which was optimised using Monte Carlo events. The value used here is $\kappa = 0.05$. The small value of $\kappa$ does not mean that the second term is dominant in the fit, since the influence of the two terms depends on their derivatives w.r.t. the track parameters.
4.5.1 Performance

The performance of the M-estimator fit is shown in figure 4.9. Again, starting values were taken at different angles from the true track. In contrast to the maximum likelihood fits, the result of the M-estimator is hardly influenced by the choice of the starting values. Only when started at 45° from the true track, the performance is somewhat degraded. As expected, the M-estimator is considerably less accurate than the ML methods, but most of the events are reconstructed with an accuracy of a few degrees, which is much more accurate than the result of the linear pre-fit shown in figure 4.3. The performance is adequate as a starting point for the ML fit.

It should be noted, however, that the hits used to obtain the results in figure 4.9 are selected using information about the direction of the true track (see section 4.2.3). If no good starting point is available for the fit, the signal hits cannot be selected in this way and the performance will, in general, be degraded.

4.6 Combining the fitting algorithms

In sections 4.2 to 4.5, four different fitting procedures were discussed. This section describes how they are combined into the final reconstruction program and how the hits are selected that are used for each of the procedures. The full reconstruction algorithm may be summarised as follows:

1. Pre-selection of hits: In the simulated events, background hits are generated in an arbitrary time window around the event. In order to make the algorithm insensitive to the amount of background simulated, a rough, first selection is made. All hits are selected for which $|\Delta t| \leq \frac{d}{v_g} + 100$ ns, where $\Delta t$ is the time difference between a hit
and the hit with the largest amplitude in the sample and $d$ is the distance between the OMs of the two hits. Hits with larger time differences cannot be related to the same muon, unless they have a residual larger than the 'safety factor' of 100 ns. Since the hit with the largest amplitude is virtually always a signal hit, almost no signal hits are rejected.

2. **Linear prefilt**: The first fit is the linear prefilt described in section 4.2. Although not very accurate, it has the advantage that it requires no starting point. It is therefore suited as a first step. The linear prefilt is made with a sub-sample of the hits. Only hits in local coincidences and hits with amplitudes larger than 3.0 p.e. are used. A local coincidence is defined as a combination of 2 or more hits on one floor within 25 ns.

3. **M-estimator fit**: The insensitivity to the quality of the starting point of the M-estimator fit (see section 4.5) makes it a natural choice for the next step. The hits used for this fit are selected on the basis of the result of the prefilt. In order to be selected, a hit must have a time residual w.r.t. the fit calculated from the parameters obtained with the linear prefilt between -150 and 150 ns and a distance from the fitted track of less than 100 m. Hits with an amplitude larger than 2.3 p.e. are always selected.

4. **Maximum likelihood fit with original PDF**: The next step is the ML fit with the original PDF discussed in section 4.3. This fit is performed with hits that are selected based on the result of the M-estimator fit. This time, hits are selected with residuals between $-0.5 \times R$ and $R$, where $R$ is the root mean square of the residuals used for the M-estimator fit. Hits that are part of a coincidence, or that have an amplitude larger than 2.5 p.e. are also selected. The asymmetry in the selection interval reflects the fact that the original PDF is asymmetric.

5. **Repetition of steps 3 and 4 with different starting points**: It was found that the efficiency of the algorithm is improved by repeating steps 3 and 4 with a number of starting points that differ from the prefilt. The result with the best likelihood per degree of freedom, as obtained in step 4, is kept. Four of the additional starting points are obtained by rotating the prefilt track by an angle of 25°. The origin of the rotation is the point on the track that is closest to the centre of gravity of the hits. Four more starting points are obtained by translating the track $\pm 50$ m in the direction $\vec{d} \times \hat{z}$ and $\pm 50$ m in the $\hat{z}$ (i.e. upward) direction. In total, steps 3 and 4 are thus done 9 times. Some additional information about the procedure is kept. The number of starting points that result in track estimates which are compatible with the preferred result (i.e. which give the same track direction to within 1°), will be used in the event selection (section 4.7.1). This number is called $N_{\text{comp}}$. In case the likelihood of one of the results is very good, the remaining starting points are skipped for execution speed.

6. **Maximum likelihood fit with improved PDF**: Finally, the preferred result obtained in step 5 is used as a starting point for the ML fit with the improved PDF. The hit selection is also based on this result: hits are selected with residuals between
-250 and 250 ns and with amplitudes larger than 2.5 p.e. or in local coincidences. Since background is taken into account in the PDF, the presence of background hits in the sample does not jeopardise the reconstruction accuracy. This is reflected in the large time window used for the hit selection. As explained in section 4.4.1, the event duration is an input parameter for the final fit; with this selection $T=500$ ns.

Clearly, the overall algorithm contains many parameters, especially for the hit-selection criteria. While a full optimisation of these parameters was not feasible, the algorithm presented here has been found to perform well over a broad energy range. Nevertheless, further tuning of the algorithm may further improve the performance.

### 4.7 Performance of the full algorithm

In this section the performance of the full reconstruction algorithm is presented. Figure 4.10 shows the result of each of the fitting algorithms when they are used in sequence as described above. The increasing accuracy of the subsequent algorithms is apparent from this figure. Of all reconstructed muon tracks, 1.1% are reconstructed within $1^\circ$ from the true track by the linear prefit (see table 4.2). This number is improved by the repetitive application of the M-estimator and the original ML fit to 38% and 57% respectively. The number is not increased much by the final fit (59%). The accuracy of the events, however, is improved: about 20% more events are reconstructed with an error smaller than $0.1^\circ$. The overall shift of the peak corresponds with a $\sim 20\%$ overall improvement of the angular resolution.

![Figure 4.10: Performance of the different fitting algorithms when used sequentially in the reconstruction program.](image)

While it is of minor importance for neutrino astronomy, the error on the position of the reconstructed muon track is shown in figure 4.11 for completeness. Most tracks are reconstructed within 1 m from the true track. As expected, there is a strong correlation
Table 4.2: Percentage of the events reconstructed within an angle $\alpha_{\mu}$ w.r.t. the true track by the different fitting algorithms that together make up the full reconstruction algorithm.

<table>
<thead>
<tr>
<th>$\alpha_{\mu}$</th>
<th>fit</th>
<th>0.1°</th>
<th>0.3°</th>
<th>1°</th>
<th>3°</th>
<th>10°</th>
</tr>
</thead>
<tbody>
<tr>
<td>final fit</td>
<td></td>
<td>12.0</td>
<td>39.1</td>
<td>59.2</td>
<td>65.6</td>
<td>71.9</td>
</tr>
<tr>
<td>original PDF</td>
<td></td>
<td>7.85</td>
<td>32.1</td>
<td>56.9</td>
<td>65.2</td>
<td>71.8</td>
</tr>
<tr>
<td>M-estimator</td>
<td></td>
<td>2.03</td>
<td>12.5</td>
<td>38.4</td>
<td>56.9</td>
<td>69.6</td>
</tr>
<tr>
<td>linear pref</td>
<td></td>
<td>0.01</td>
<td>0.10</td>
<td>1.10</td>
<td>7.99</td>
<td>39.4</td>
</tr>
</tbody>
</table>

between the angular error and the error on the position. This is demonstrated in figure 4.11 by the distribution of the positional error for events where the muon is reconstructed with a directional error smaller than 1°.

Figure 4.11: Error on the position of the muon track, defined as the minimal distance between the true and the reconstructed muon tracks, for all reconstructed muons and for muons that are reconstructed with an angular error smaller than 1°.

### 4.7.1 Handles on the reconstruction quality

The philosophy adopted in the reconstruction algorithm is to reconstruct as many events as possible without trying to reduce the number of badly reconstructed events by intermediate selection criteria. Rather, selection criteria can be applied afterwards, depending on the demands of the various physics analyses. As a consequence, many events are reconstructed with large errors. In this section, two variables are introduced which can be used to reject badly reconstructed events.

An obvious way to discriminate good and bad events is a cut on the value of the likelihood function at the fitted maximum\(^\text{4}\). As is shown in figure 4.12, there is an overall error

\(^{4}\)For brevity, we treat the likelihood as a dimensionless number; strictly speaking it has units $\text{ns}^{-N}$, with $N$ the number of hits used in the fit.
4.7. Performance of the full algorithm

Figure 4.12: Scatter plot of the value of the likelihood function at the fitted maximum and the number of degrees of freedom of the fit.

linear relationship between \( \log(L) \) and the number of degrees of freedom \( N_{\text{DOF}} \) in the fit. The latter is given by the number of hits used in the fit minus the number of fitted parameters, which is five. Therefore, the variable \( \log(L)/N_{\text{DOF}} \) is used to select the well reconstructed events. Figure 4.13 demonstrates that, as expected, \( \log(L)/N_{\text{DOF}} \) has a high value for well reconstructed events (in this case with directional errors smaller than 1°) and a low value for badly (in this case with errors larger than 45°) reconstructed events.

A second parameter is obtained from information about the multiple starting points that are used for the M-estimator and the subsequent ML estimator. It is the number of tracks compatible with the preferred track \( N_{\text{comp}} \) (see section 4.6). As is shown in figure 4.14, \( N_{\text{comp}} = 1 \) for the majority of the badly reconstructed events, while it can be as large as 9 for well reconstructed events; in that case all of the starting points have resulted in the same track.

We have shown here that these two variables can be used to select well reconstructed muons from badly reconstructed up-going muons. In chapter 5 the same parameters will be used to reject the background of wrongly reconstructed atmospheric muons, which form a more dangerous background than up-going muons.

4.7.2 Error estimates

Besides estimates of the track direction, error estimates are provided by the reconstruction program. In general the \( 1(2)\sigma \) confidence intervals consist of the part of parameter space where \( \log(L) - \log(L_{\text{max}}) < 1/2 \) (2), where \( L_{\text{max}} \) is the maximum value the likelihood function obtained with the fit. Under the assumption that the likelihood function near the fitted maximum follows a multivariate Gaussian distribution, the error covariance matrix \( \mathbf{V} \) can be obtained from the second derivatives of the likelihood function at the
4.7. Performance of the full algorithm

Figure 4.13: Left: Reconstruction error vs. the value of the likelihood per degree of freedom $\log(L)/N_{\text{DOF}}$. Right: The distribution of $\log(L)/N_{\text{DOF}}$ for well reconstructed events (error < 1°) and for badly reconstructed events (error > 45°).

Figure 4.14: Left: Reconstruction error vs. the value of $N_{\text{comp}}$. Right: The distribution of $N_{\text{comp}}$ for well reconstructed events (error < 1°) and for badly reconstructed events (error > 45°).
4.7. Performance of the full algorithm

Muon Track Reconstruction

Figure 4.15: Pull distribution for the zenith angle. The right plot is a close up of the central peak. Distributions are shown for all events and for selected events. The normalisation of the latter distribution reflects the selection efficiency.

fitted maximum:

\[ \left[ V^{-1} \right]_{ij} = -\frac{\partial^2 \log(L)}{\partial x_i \partial x_j} \]  

(4.19)

where \( \mathbf{x} \) is the vector of track parameters: \( \mathbf{x} = (p_x, p_y, p_z, \theta, \phi) \). In particular, we will use the estimated error on the zenith and azimuth angles: \( \hat{\sigma}_\theta \) and \( \hat{\sigma}_\phi \) respectively.

The ratio of the true error on a variable and the error estimate is referred to as the pull. The expectation value of the pull is zero, since the estimate of the parameters are symmetrically distributed around the true value. The properties of ML estimators ensure that, in the limit of an infinite amount of data, the estimates are normally distributed around the true values. This means that the pull distribution should be Gaussian with a width \( \sigma = 1 \).

The distribution of the zenith angle pull (i.e. \( (\hat{\theta} - \theta)/\hat{\sigma}_\theta \)) is shown in figure 4.15. While the distribution is strongly peaked at zero, large tails are present, consisting of events with unreliable error estimates. These can occur when the reconstruction algorithm has converged to a local maximum. Therefore, the pull distribution is also shown for events with \( \log(L)/N_{\text{DOF}} > -5.3 \). This selection improves the pull distribution significantly, which means that a large number of the events with incorrect error estimates are rejected. After this quality cut, the pull distribution is approximately Gaussian (although tails are still present). When fitted in the interval (-2, 2), the \( \sigma \) of the pull distribution is 1.10, which is in reasonable agreement with the expected value of 1. The pull distribution for the azimuth angle is shown in figure 4.16. The same features as the pull distribution of the zenith angle can be distinguished.

The estimate of the error on the direction of the reconstructed muon track, \( \hat{\alpha}_\mu \), is
obtained from $\sigma_\theta$ and $\sigma_\phi$ as follows:

$$\alpha_\mu = \sqrt{\sin^2(\hat{\theta})\sigma_\phi^2 + \sigma_\theta^2}. \quad (4.20)$$

The distribution of the quantity $\alpha_\mu / \hat{\alpha}_\mu$ is shown in figure 4.17. In the case of uncorrelated Gaussian errors, the distribution of this quantity should be proportional to $\alpha_\mu / \hat{\alpha}_\mu e^{-(\alpha_\mu / \hat{\alpha}_\mu)^2}$. The distribution is described reasonably well by this function, although the tail is more pronounced, which is not surprising since tails are also present in the pull distributions for $\theta$ and $\phi$. 

Figure 4.16: Pull distribution for the azimuth angle for all events and for selected events.

Figure 4.17: Distribution of $\alpha_\mu / \hat{\alpha}_\mu$ for all events and for events with a high value of the likelihood. Also shown is the theoretical expectation for this distribution in case the errors are uncorrelated and Gaussian. The normalisation of this function is chosen so that it coincides with the solid histogram.
4.8 Energy reconstruction

The algorithm used for the estimation of the energy is explained in detail in [38] and will be summarised here. As discussed in section 3.2.3, the muon energy loss due to pair production and bremsstrahlung increases strongly with energy. The amount of light emitted by the electro-magnetic cascades resulting from these processes can be used as a measure for the muon energy loss and thus for the muon energy.

The average energy lost by the muon per unit track length is estimated from the sum of the amplitudes of the hits ($A^{\text{tot}}$). Only hits with an amplitude larger than 2.5 p.e. are taken into account, since they originate predominantly from electro-magnetic cascades.

The information on the track direction and position that is provided by the previously described track reconstruction algorithm is used to estimate the acceptance of the detector, which is used to correct the estimate of $\frac{dE}{dx}$:

$$\left\langle \frac{dE}{dx} \right\rangle = A^{\text{tot}} t^{-1}_\mu \left( \frac{1}{N_{\text{PMTs}}} \sum_{i=1}^{\text{PMTs}} \frac{f_{\text{ang}}(a_i)}{b_i} e^{-b_i/\lambda^{\text{abs}}} \right)^{-1},$$

where $b_i$ is the distance traveled by a photon when it travels from the reconstructed track to the PMT, $f_{\text{ang}}(a)$ is the acceptance of the OM, which depends on the angle of incidence of the photon (see figure 3.10). The absorption length, $\lambda^{\text{abs}}$, is taken at a fixed wavelength of 475 nm (see figure 3.7). Finally $l_\mu$ is the 'observable track length', which is defined as the length of the muon track that lies within a cylinder around the instrumented volume. The dimensions of the cylinder are defined by the dimensions of the instrumented volume extended by a distance of $2 \times \lambda^{\text{abs}}$.

Figure 4.18(left) shows the estimated $\frac{dE}{dx}$ as a function of the muon energy. The figure resembles the curves of the average muon energy loss shown in figure 3.4. From the estimate of $\frac{dE}{dx}$, the muon energy is determined using an empirical function, which is also shown in the figure. This function does not extend below $E_\mu = 100$ GeV. For muons below this energy $\frac{dE}{dx}$ is almost independent of the muon energy. If the estimate of $\frac{dE}{dx}$ is smaller than the value corresponding to $E_\mu = 100$ GeV, the muon energy is not reconstructed. A scatter plot of the reconstructed muon energy as a function of the true muon energy is shown in figure 4.18(right). In section 5.2.4 we will say a bit more about the performance of the energy reconstruction.

4.9 Outlook and conclusion

Although the reconstruction algorithm that was presented in this chapter performs well, there is still room for further improvement. A better starting point for the final fit would improve the final results. The method of choosing starting points for the fit could be improved if the positions of the local maxima in the likelihood function are better understood. Furthermore, there are other possibilities for the function $q(r)$ used by the M-estimator, some of which may perform better than the present one. The pointing accuracy can be improved by making the PDF of the time residual of the signal hits dependent on the distance traveled by the photon. The PDF will be peaked if the distance is small, but will be smeared for larger distances due to light scattering and dispersion. Finally,
4.9. Outlook and conclusion

Figure 4.18: Left: Estimate of $\frac{dE}{dx}$ (up to an arbitrary constant) as a function of the simulated muon energy. The empirical function which is used to convert the estimate of $\frac{dE}{dx}$ to the estimate of the muon energy is shown as the grey line. Right: Reconstructed muon energy as a function of the true muon energy. The data in the figures corresponds to events where the muon direction was reconstructed within $1^\circ$ from the true direction.

the energy of the muon could be taken into account, since the PDF of the signal hits broadens with energy (due to the larger contribution of hits from secondary electrons), but also because the relative contribution of background hits depends on the energy.

To conclude, we mention that the track reconstruction algorithm that was described here has been used for various studies of the expected physics potential of ANTARES [38, 73, 74, 75] and of possible future km$^3$-scale neutrino telescopes [76].