ELSEVIER

Estimating the angular resolution of tracks in neutrino telescopes based on a likelihood analysis

Till Neunhöffer

Institute of Physics, University of Mainz, Staudinger Weg 7, 55099 Mainz, Germany

Abstract

A semianalytic method to estimate the angular resolution of tracks, that have been reconstructed by a likelihood approach, is presented. The optimal choice of coordinate systems and resolution parameters, as well as tests of the method are discussed based on an application for a neutrino telescope.
© 2006 Elsevier B.V. All rights reserved.

1. Introduction

This paper describes a statistical procedure to extract resolution estimates on a track by track basis. The method was developed for the AMANDA Neutrino Telescope at the South Pole [1], which uses a three-dimensional grid of photosensors imbedded in highly transparent ice to provide spatial and time resolution of Cherenkov photons, that e.g. arise from long muon tracks.

The knowledge of track resolutions is of particular importance in the search for localized sources, such as distant galaxies. The resolution information can in addition be used to suppress mis-reconstructed tracks, that typically are less well defined. The method is not limited to muon reconstruction in neutrino telescopes and can be applied to any experiment in which tracks have been reconstructed with a likelihood approach.

2. Technical aspects of obtaining the confidence ellipse

2.1. Definitions and prerequisites

In all following paragraphs a *pattern of hits* in an event will be modelled by an infinitely long oriented straight line. It will be parameterized by an arbitrary point $\mathbf{r}$ on the track and the zenith $\vartheta \in [0,\pi]$ and azimuth $\varphi \in [0,2\pi]$ angles in a given detector coordinate system.

The procedure presented operates on likelihood functions, thus a likelihood based reconstruction must be available. For any given hypothesis $\{\mathbf{r}_0, \vartheta_0, \varphi_0\}$ and an observed pattern of hits $\mathbf{P}$, a likelihood $\mathscr{L}(\mathbf{r}_0, \vartheta_0, \varphi_0; \mathbf{P})$ must be calculable. The hypothesis $h_{\text{best}} = \{\mathbf{r}_{\text{best}}, \vartheta_{\text{best}}, \varphi_{\text{best}}\}$ with the highest value of $\mathscr{L} = \mathscr{L}_{\text{best}}$ is considered to be the best estimate for the true values. It is assumed that this extremum has already been found.

The uncertainties, which are to be determined, are in essence[1] the errors of ϑ and φ. It is mandatory to use the

E-mail address: Till.Neunhoeffer@uni-mainz.de

---

[1] A coordinate transformation needs to be done first, which will be explained in the following section.

optimal coordinate system for each event and to consider the errors in the determination of **r** as well.

In order to arrive at a confidence ellipse in the two dimensions of track direction, the set in parameter space is being searched for, where the value of the negative logarithmic likelihood function has changed by 1/2 with respect to $-\log \mathscr{L}_{best}$:

$$\Delta(-\log \mathscr{L}) = (-\log \mathscr{L}_{ellipse}) - (-\log \mathscr{L}_{best}) \overset{!}{=} \frac{1}{2}. \quad (1)$$

After discussing the coordinate systems used, the dimension of the problem will be reduced to the two dimensions of direction by eliminating the three dimensions of the location of **r**. This is done using a numerical minimization process. The remaining χ^2-minimization is done analytically to obtain the uncertainty estimators.

The shape of the negative log-likelihood function around its minimum is considered Gaussian. The tests described in Section 4 prove that this assumption is realistic.

2.2. On coordinate systems

There are two sets of coordinate systems used in the method presented here. The first system uses all five coordinates $(x, y, z, \vartheta, \varphi)$, which are connected to the 3-dimensional Cartesian detector coordinate system as illustrated in Fig. 1(a). In the second system the problem is reduced to the 2-dimensional subset of angles ϕ and θ.

The complication with ϑ and φ is that the first has boundaries and the second is periodic. Moreover, at the boundaries of ϑ the variable φ becomes meaningless. In a

Fig. 2. Confidence ellipse in two dimensions with a directional coordinate θ and a spatial coordinate x. The rectangle around the ellipse has the size $2\sigma_\theta$ by $2\sigma_x$. Finding for each θ the corresponding x with extremal likelihood, one arrives at the line shown.

mathematical sense, ϑ and φ are locally orthogonal. However, if one looks at these coordinates in the area surrounding a point, they are distorted—in particular close to $\vartheta = 0$ and $\vartheta = \pi$.

As the behaviour of the likelihood function around the well known best track estimate h_{best} is to be investigated, the coordinates ϑ and φ should be as close to being orthogonal as possible *at that particular direction* ($\vartheta_{best}, \varphi_{best}$). Hence the 3-dimensional detector coordinate system is rotated such that the new x-axis, x', is defined by the track. This results in $\vartheta'_{best} = \frac{\pi}{2}$ and $\varphi'_{best} = \pi$, as illustrated in Fig. 1(b), where the distortion is minimal.

Here and in the following Section 3 only *relative* coordinates $\theta = \vartheta' - \pi$ and $\phi = \varphi' - \pi/2$ with respect to ϑ'_{best} and φ'_{best} will be referred to. This can be thought of as local coordinates in the tangential plane to the unit sphere at $(\vartheta_{best}, \varphi_{best})$ with the origin at the point, where the plane touches the sphere.[2] By choosing a proper rotation, it is assured that the θ-direction is parallel to the ϑ-coordinate, i.e., towards the negative z/z'-axis. The rotation must be such that the track, the z-axis, and the z'-axis are in the same plane.

The local coordinates are displayed in Fig. 1(c).

2.3. Reducing the dimensionality of the problem

The location of the point **r** is of no concern for the discussion of where a track points to. Nevertheless, the correlations of the errors in direction with the errors in determining **r** need to be considered.

The reduction procedure of the dimension shall first be discussed assuming only one spatial dimension x and one directional dimension θ.

After imposing the minimization requirement $\frac{\partial \mathscr{L}}{\partial x}\big|_\theta \overset{!}{=} 0$, the likelihood function $\mathscr{L}(\theta, x(\theta))$ only depends on one
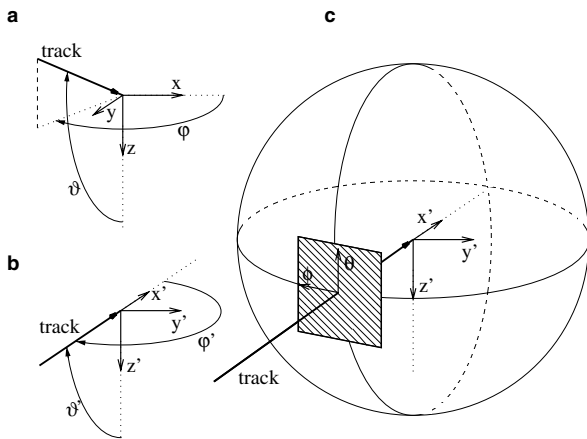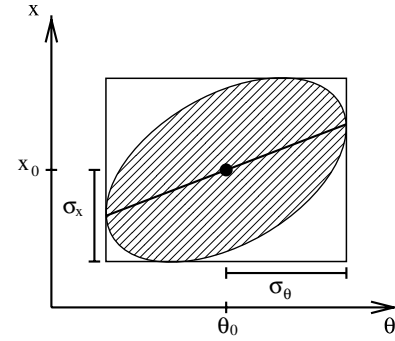

Fig. 1. Figure (a) shows a track in a Cartesian coordinate system (x, y, z). The direction towards the track's origin is described with standard spherical coordinates (ϑ, φ). The downward orientation of the z-axis follows the AMANDA standards [2]. In (b) a rotated coordinate system (x', y', z') is presented, where the x'-axis is defined by the track. To describe directions that are close to the track's direction, relative coordinates $\theta = \vartheta' - \pi/2$ and $\phi = \varphi' - \pi$ are introduced. They are approximately the coordinates in the tangent plane, as shown in (c).

---

[2] This is an approximation that implies that the coordinates on the surface of a sphere can be assumed Cartesian for small distances. The maximum error introduced is of the order of 0.15% for distances less than 5°.

variable θ. The curvature at the minimum of $-\log \mathscr{L}$ reflects the uncertainty in θ *including* its correlation with the uncertainty in x.

Consider the two-dimensional confidence ellipse in (θ, x)-space (Fig. 2). It is tilted versus the θ-axis, which indicates a correlation of θ and x. For each θ one determines the value of x for which the common likelihood function gets maximal. In the strictly Gaussian case, the result is a straight line. The likelihood function along this line is again a Gaussian as a function of θ. The width of this Gaussian is σ_θ as can be seen by straight forward calculation. In this error estimator σ_θ the correlation $\mathrm{cov}(\theta, x)$ is included by construction.

The extension to three spatial and two directional dimensions is straight forward: first one determines a set

$$\Gamma = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12} & \Gamma_{22} \end{pmatrix} \text{ contains the second order parameters.}$$

Using the method of least squares one calculates

$$\chi^2 = \sum_{i=1}^{n} (f(\phi_i, \theta_i; \delta, \beta_1, \beta_2, \Gamma) - l_i)^2$$
$$= \sum_{i=1}^{n} (\delta + \beta_1 \phi_i + \beta_2 \theta_i + \frac{1}{2}(\phi_i, \theta_i)\Gamma(\phi_i, \theta_i)^T - l_i)^2, \quad (3)$$

and requires the minimization conditions

$$\frac{\partial \chi^2}{\partial \delta} = \frac{\partial \chi^2}{\partial \beta_i} = \frac{\partial \chi^2}{\partial \Gamma_{ij}} \overset{!}{=} 0. \quad (4)$$

This leads to a six-dimensional inhomogeneous system of linear[3] equations:

$$\begin{pmatrix} \sum 1 & \sum \phi_i & \sum \theta_i & \frac{1}{2}\sum \phi_i^2 & \sum \phi_i\theta_i & \frac{1}{2}\sum \theta_i^2 \\ \sum \phi_i & \sum \phi_i^2 & \sum \phi_i\theta_i & \frac{1}{2}\sum \phi_i^3 & \sum \phi_i^2\theta_i & \frac{1}{2}\sum \phi_i\theta_i^2 \\ \sum \theta_i & \sum \phi_i\theta_i & \sum \theta_i^2 & \frac{1}{2}\sum \phi_i^2\theta_i & \sum \phi_i\theta_i^2 & \frac{1}{2}\sum \theta_i^3 \\ \sum \phi_i^2 & \sum \phi_i^3 & \sum \phi_i^2\theta_i & \frac{1}{2}\sum \phi_i^4 & \sum \phi_i^3\theta_i & \frac{1}{2}\sum \phi_i^2\theta_i^2 \\ \sum \phi_i\theta_i & \sum \phi_i^2\theta_i & \sum \phi_i\theta_i^2 & \frac{1}{2}\sum \phi_i^3\theta_i & \sum \phi_i^2\theta_i^2 & \frac{1}{2}\sum \phi_i\theta_i^3 \\ \sum \theta_i^2 & \sum \phi_i\theta_i^2 & \sum \theta_i^3 & \frac{1}{2}\sum \phi_i^2\theta_i^2 & \sum \phi_i\theta_i^3 & \frac{1}{2}\sum \theta_i^4 \end{pmatrix} \cdot \begin{pmatrix} \delta \\ \beta_1 \\ \beta_2 \\ (\Gamma)_{11} \\ (\Gamma)_{12} \\ (\Gamma)_{22} \end{pmatrix} = \begin{pmatrix} \sum l_i \\ \sum l_i\phi_i \\ \sum l_i\theta_i \\ \sum l_i\phi_i^2 \\ \sum l_i\phi_i\theta_i \\ \sum l_i\theta_i^2 \end{pmatrix} \quad (5)$$

of points in (ϕ_i, θ_i). For each point, the best guess, i.e., the best likelihood $l_i = -\log \mathscr{L}_i$, is found with respect to **r**.

One can choose the set of points (ϕ_i, θ_i) in a suitable way to simplify Eq. (5). If for each point (ϕ_i, θ_i) also $(\phi_i, -\theta_i)$, $(-\phi_i, \theta_i)$, and $(-\phi_i, -\theta_i)$ are added to the set,[4] one arrives at

$$\begin{pmatrix} \sum 1 & 0 & 0 & \frac{1}{2}\sum \phi_i^2 & 0 & \frac{1}{2}\sum \theta_i^2 \\ 0 & \sum \phi_i^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sum \theta_i^2 & 0 & 0 & 0 \\ \sum \phi_i^2 & 0 & 0 & \frac{1}{2}\sum \phi_i^4 & 0 & \frac{1}{2}\sum \phi_i^2\theta_i^2 \\ 0 & 0 & 0 & 0 & \sum \phi_i^2\theta_i^2 & 0 \\ \sum \theta_i^2 & 0 & 0 & \frac{1}{2}\sum \phi_i^2\theta_i^2 & 0 & \frac{1}{2}\sum \theta_i^4 \end{pmatrix} \cdot \begin{pmatrix} \delta \\ \beta_1 \\ \beta_2 \\ (\Gamma)_{11} \\ (\Gamma)_{12} \\ (\Gamma)_{22} \end{pmatrix} = \begin{pmatrix} \sum l_i \\ \sum l_i\phi_i \\ \sum l_i\theta_i \\ \sum l_i\phi_i^2 \\ \sum l_i\phi_i\theta_i \\ \sum l_i\theta_i^2 \end{pmatrix}. \quad (6)$$

In the Gaussian approximation these points $l_i(\phi_i, \theta_i)$ form a paraboloid. Its parameters are found by analytic minimization, which is discussed next.

2.4. The analytic minimization

In order to find the paraboloid that best fits the points $l_i(\phi_i, \theta_i)$, an analytic χ^2-minimization is performed. The paraboloid $f(\phi, \theta; \delta, \beta_1, \beta_2, \Gamma)$ is parameterized in its polynomial representation:

$$f(\phi, \theta; \delta, \beta_1, \beta_2, \Gamma) = \delta + \beta_1 \phi + \beta_2 \theta + \frac{1}{2}(\phi, \theta)\Gamma(\phi, \theta)^T. \quad (2)$$

The constant parameter is represented by δ and the parameters in first order by β_1 and β_2. The symmetric 2×2-matrix

This leads to three equations:

$$\sum \phi_i^2 \cdot \beta_1 = \sum l_i\phi_i \quad (7)$$
$$\sum \theta_i^2 \cdot \beta_2 = \sum l_i\theta_i \quad (8)$$
$$\sum \phi_i^2\theta_i^2 \cdot (\Gamma)_{12} = \sum l_i\phi_i\theta_i \quad (9)$$

and a three-dimensional system

---

[3] These equations are linear in the parameters δ, β_1, β_2, and Γ, which are to be determined.

[4] This is both a sensible and easily fulfilled demand, as h'_{best} corresponds to $\theta = \phi = 0$.

$$\begin{pmatrix} \sum 1 & \frac{1}{2}\sum \phi_i^2 & \frac{1}{2}\sum \theta_i^2 \\ \sum \phi_i^2 & \frac{1}{2}\sum \phi_i^4 & \frac{1}{2}\sum \phi_i^2\theta_i^2 \\ \sum \theta_i^2 & \frac{1}{2}\sum \phi_i^2\theta_i^2 & \frac{1}{2}\sum \theta_i^4 \end{pmatrix} \cdot \begin{pmatrix} \delta \\ (\Gamma)_{11} \\ (\Gamma)_{22} \end{pmatrix} = \begin{pmatrix} \sum l_i \\ \sum l_i\phi_i^2 \\ \sum l_i\theta_i^2 \end{pmatrix},$$

(10)

which can easily be solved.

The error estimators sought are contained in the parameters of the paraboloid. The three degrees of freedom of δ and $\beta_{1,2}$ correspond to the position of the minimum of the curve. The three parameters in Γ contain the curvature of the paraboloid, which are functions of the resolution, as the covariance matrix C of the Gaussian approximation is the inverse of Γ. From $C = \Gamma^{-1}$ one obtains:

$$\sigma_\phi^2 = (\Gamma^{-1})_{11}$$

(11)

$$\sigma_\theta^2 = (\Gamma^{-1})_{22}$$

(12)

$$\mathrm{cov}(\phi,\theta) = (\Gamma^{-1})_{12}.$$

(13)

The results can also be represented in a more intuitive way using the major and the minor axes σ_1 and σ_2 of the confidence ellipse, and its rotation angle α with respect to the ϕ-axis (Fig. 3). In order to find this alternative representation, the covariance matrix C needs to be diagonalized. That leads to

$$\sigma_{1,2}^2 = \frac{\sigma_\phi^2 + \sigma_\theta^2}{2} \pm \sqrt{\frac{(\sigma_\phi^2 + \sigma_\theta^2)^2}{4} + \det C}.$$

(14)

Choosing $\sigma_1 \geqslant \sigma_2$, the rotation angle α is then given by the expression

$$\tan \alpha = \frac{\sigma_1^2 - \sigma_\phi^2}{\mathrm{cov}(\phi,\theta)}.$$

(15)

In the following discussion a third representation $(\sigma_a, \epsilon, \alpha)$ will be used, where σ_a is connected to the area of the ellipse and ϵ is its excentricity:

$$\sigma_a = \sqrt{\sigma_1 \cdot \sigma_2}$$

(16)

$$\epsilon = \frac{\sigma_1}{\sigma_2}.$$

(17)

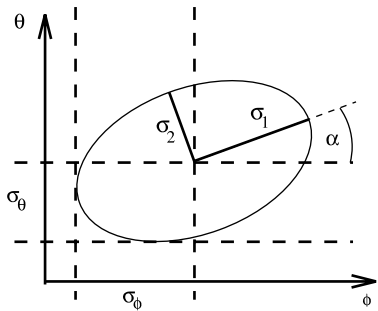The rotation angle α is the same as in the previous parameterization.

Fig. 3. The confidence ellipse in θ and ϕ can be represented either by σ_θ, σ_ϕ, and the covariance $\mathrm{cov}(\theta,\phi)$ or by the major and minor axes σ_1 and σ_2, and the rotation angle α of the ellipse between the major axis and the θ-direction.
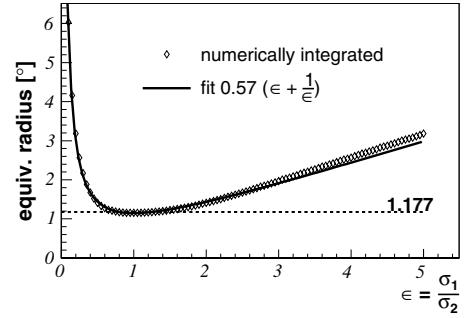

Fig. 4. Obtaining the space angle resolution estimate by applying a correction based on the excentricity. For fixed $\sigma_a = 1°$ and varying ϵ, the *equivalent radius M* is displayed, such that the true direction lies within a circle of radius M with a probability of 50%.

If the track reconstruction arrives at a saddle point rather than at a true minimum in θ and ϕ, unphysical, negative values for the variance estimators occur.

Note, that the resolution estimates are calculated for each pattern of hits separately. Still they are not the *true* values for the discrepancies of true and reconstructed direction, but estimates of the *average* discrepancy for a set of events, which are similar to the one under consideration.

3. Space angle resolution

For a neutrino telescope it is common to quote the median $\overline{\Psi}$ of the spatial angle Ψ between true and reconstructed track direction in order to describe its resolution. This section is devoted to the connection of the error ellipse parameters to the space angle uncertainty ψ. It is convenient to choose the representation of the ellipse using σ_a and ϵ, with α not playing a role in this special context.

The true direction lies within the area given by the error ellipse with a probability of 39.4%. The median of a distribution corresponds to a 50%-probability. Hence as a first step the ellipses need to be enlarged correspondingly.

In order to obtain ψ, one considers the Gaussian likelihood function in its diagonal form, where the coordinates[5] are x_1 and x_2:

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left(\left(\frac{x_1}{\sigma_1}\right)^2 + \left(\frac{x_2}{\sigma_2}\right)^2\right)}.$$

(18)

To obtain the median M, the function $f(x_1, x_2)$ must be integrated from the center $(0,0)$ outwards until a cumulated probability of 0.5 is reached. This integral is done in polar coordinates (radius ρ and angle γ)

$$0.5 \overset{!}{=} \int_0^M \int_0^{2\pi} f(\rho\cos\gamma, \rho\sin\gamma)\rho\,d\gamma\,d\rho.$$

(19)

For the special case of $\sigma_a = \sigma_1 = \sigma_2$ (i.e., $\epsilon = 1$), the calculation can be carried out analytically and yields $M = 1.177 \cdot \sigma_a$.

---

[5] The coordinates x_1 and x_2 are connected to ϕ and θ by a rotation of the angle α.
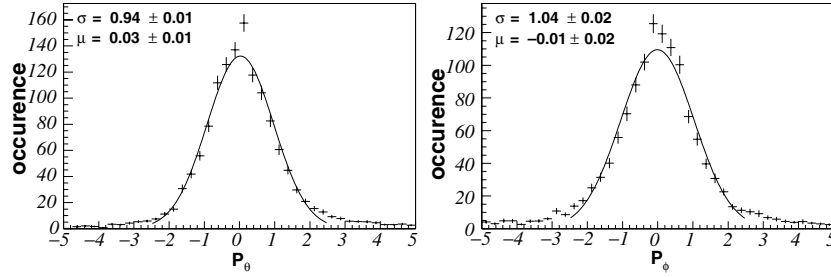
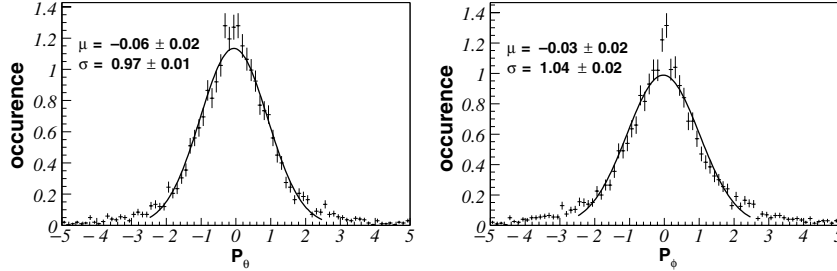Fig. 5. Pull in zenith (left) and azimuth (right) for atmospheric neutrino simulation.

Fig. 6. Distributions of P_ϑ and P_φ as defined in Eqs. (21) and (22) based on split atmospheric muons events.

For the general case (with arbitrary ϵ) the integration must be carried out numerically. This has been done for the case of $\sigma_a = 1°$ and the result is shown in Fig. 4. The median M scales linearly with σ_a.

The curve is symmetric with respect to the exchange of σ_1 and σ_2. To improve the speed of the program an approximation proportional to $(\epsilon + \frac{1}{\epsilon})$ has been applied which results in the *excentricity corrected spatial angle resolution estimation*:

$$\sigma_a^\epsilon = 0.57 \cdot \left(\epsilon + \frac{1}{\epsilon}\right) \cdot \sigma_a. \tag{20}$$

The fit is also shown in Fig. 4. The deviation in the interval $\epsilon \in [1;3]$ is below 5%. For applications where this is too coarse an approximation, the exact results of the numerical integration can be stored in appropriate tables.

4. Discussion

The above algorithms have been tested within the framework of the AMANDA neutrino telescope. The estimates perform in a stable and reliable way and produce sensible results.

To ensure the correctness of the estimations, several testing procedures both in data and in Monte Carlo have been carried out:

(1) For σ_θ and σ_ϕ its *pull* has been studied. It is defined as the ratio of the difference of true and reconstructed direction over the resolution estimator. If the estimation is good, the pull should be Gaussian distributed, centered at zero and with unit width. The corresponding plots are shown in Fig. 5. This investigation can only be done in a Monte Carlo simulation.

(2) For the spatial angle resolution the pull is not a sensible quantity, because angles in space can only be positive. Hence one can study ensembles of events with the same σ_a or σ_a^ϵ, respectively. The median of the true deviation can then be compared to the corresponding σ_a. Again this can only be done in the simulation.

(3) In data the true directions are not known. Thus another approach is being followed. Each event is split into two subevents by assigning every second hit to subevent 1 and the remaining hits to subevent 2. Each subevent undergoes reconstruction and is subjected to the resolution estimation algorithm. That leads to directions[6] ϑ^i, φ^i and errors σ_θ^i and σ_ϕ^i with $i \in \{1,2\}$.

The pull

$$P_\vartheta = \frac{D_\vartheta}{\sigma_\theta^D} = \frac{\vartheta^1 - \vartheta^2}{\sqrt{(\sigma_\theta^1)^2 + (\sigma_\theta^2)^2}} \tag{21}$$

should also be a Gaussian distribution centered at zero with unit width. For P_φ the difference in azimuth angles must be multiplied by $\sin\vartheta$ to make up for the differences in the rotated and not rotated coordinate systems:

$$P_\varphi = \frac{D_\varphi}{\sigma_\phi^D} = \frac{(\varphi^1 - \varphi^2) \cdot \sin\vartheta}{\sqrt{(\sigma_\phi^1)^2 + (\sigma_\phi^2)^2}}. \tag{22}$$

The corresponding distributions for atmospheric muons are displayed in Fig. 6.

---

[6] Note that the reconstructed directions are expressed in standard detector coordinates, whereas the error estimates are obtained in their respective rotated systems. That is due to technical reasons only.

Of course this last check can also be done in the simulation.

In all tests the results were satisfactory. Further details about the tests and their results can be found in [3].

4.1. Concluding remarks

(1) Often one applies quality cuts to the data set, which rely on additional information that is not reflected in the construction of the likelihood function itself. Hence this information is also not used in obtaining the resolution parameters. In this case one expects the selected tracks to be on average closer to the true direction than from the likelihood analysis alone. Consequently one observes a pull with a width smaller than 1.

(2) The resolution parameters—such as σ_a—can efficiently be used as cut variables, as mis-reconstructed events on average have a worse resolution than correctly reconstructed ones.

(3) The resolution estimation on an event-per-event basis can be used as additional input in point source searches. The information is hard to be included in the commonly used binned search algorithms [4]. An appropriate procedure based on maximum likelihood methods that can integrate the new information in a natural way will be presented in a forthcoming publication.

Acknowledgements

References

[1] E. Andres et al., Nature 410 (2001) 441.
[2] J. Ahrens et al., Nucl. Instr. Meth. A 534 (2004) 169.
[3] Till Neunhöffer, Ph.D. Thesis, Mainz 2004, ISBN 3-8322-2474-2 (in German).
[4] J. Ahrens et al., Phys. Rev. Lett. 92 (2004) 71102.