



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Nuclear Instruments and Methods in Physics Research A 496 (2003) 172–182

**NUCLEAR  
INSTRUMENTS  
& METHODS  
IN PHYSICS  
RESEARCH**  
Section A[www.elsevier.com/locate/nima](http://www.elsevier.com/locate/nima)

# A maximum likelihood method for particle momentum determination

Tingjun Liu\*, William Molzon

*Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA*

Received 1 July 2002; accepted 9 July 2002

---

## Abstract

We discuss a maximum likelihood method for determining a charged particle's momentum as it moves in a magnetic field. The formalism is presented in both rigorous and approximate forms. The rigorous form is valid when random processes include multiple scattering, energy loss and detector spatial resolution. When the measurement error is dominated by multiple scattering, it takes a particularly simple approximate form. The validity of both formalisms extends to include non-Gaussian multiple scattering distribution.

© 2003 Elsevier Science B.V. All rights reserved.

PACS: 02.50.Cw; 02.50.Ng; 02.50.Ph

Keywords: Likelihood; Fitting

---

## 1. Introduction

Track fitting (determining the best estimate of the kinematic properties of a particle from a set of measured positions) is a standard problem in experimental particle physics. This paper discusses a method for determining the momentum of a charged particle moving in a magnetic field, using a set of position measurements and including the effects of multiple scattering, energy loss and detector spatial resolution.

This work is motivated by considering the measurement of 105 MeV electrons in an experiment [1] to search for the coherent conversion of

muons to electrons in the field of a nucleus. To reduce backgrounds, it is required that the momentum be measured with a precision of order  $\delta P/P = 0.001$ . Fig. 1 shows three-dimensional and cut views of a typical electron trajectory in the detector; the particle crosses the detector 7 times and each time one measurement of  $x$ ,  $y$  and  $z$  is recorded. The detector is in a uniform 1 T magnetic field, which corresponds to  $r = 25$  cm for a typical electron that has a total momentum of 105 MeV/ $c$  and a momentum transverse to the magnetic field of 75 MeV/ $c$ . The detailed construction of the detector is not important for this discussion, but for completeness we briefly describe it. It consists of three layers of 5 mm diameter straw tubes arranged in a cylinder and eight *vanes* extending radially outward from the cylinder; all individual detector elements are

---

\*Corresponding author. Tel.: +1-949-824-6444.

E-mail addresses: [tjliu@uci.edu](mailto:tjliu@uci.edu) (T. Liu),  
[wmolzon@uci.edu](mailto:wmolzon@uci.edu) (W. Molzon).

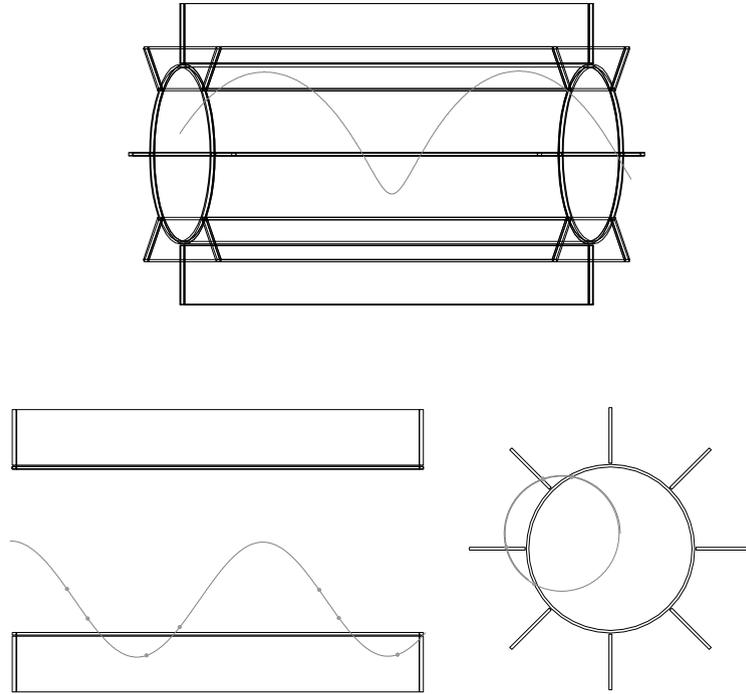


Fig. 1. The plot shows three-dimensional, side and end views of a typical electron trajectory in the detector [1]. The 105 MeV electron crosses the detector 7 times in a uniform 1 T region. The crossings are indicated by dots.

parallel to the field direction. Each detector plane corresponds to  $\sim 0.1 \text{ g/cm}^2$  of path length traversed by a typical particle's trajectory. The axial coordinate will be measured by capacitive coupling to foils on the outside of the straw tubes. The radius of the cylinder is 40 cm, the radial extent of each vane is 30 cm and the detector length is 200 cm.

Throughout this discussion, we only consider the case of uniform  $B$  field and note that the method can be readily generalized to any well measured  $B$  field. The particle's momentum is determined by the shape of the helical trajectory. Denoting the particle's perpendicular, axial and total momenta as  $P_{\perp}$ ,  $P_{\parallel}$  and  $P$ , respectively, one has

$$P_{\perp} = Cr \quad (1)$$

$$P_{\parallel} = C \frac{dz}{d\phi} \quad (2)$$

$$P = \sqrt{P_{\perp}^2 + P_{\parallel}^2} \quad (3)$$

where  $r$  is the radius of the helix,  $dz/d\phi$  is the axial displacement per radian and  $z$  is along the magnetic field direction.  $C$  is a constant, and  $C = |q|B/c$  in Gaussian units, where  $q$  is the charge,  $B$  is the value of the magnetic field and  $c$  is the speed of light.

Assuming  $n$  sequential position measurements on the trajectory,  $(x_1, y_1, z_1), \dots, (x_n, y_n, z_n)$ , one can reconstruct the helix path to calculate the particle momentum. However, the particle's actual trajectory deviates from a helix due to multiple scattering and energy loss in the detector. Finite position resolution of the detector also contributes to deviations of measured positions from those of a perfect helix. When the particle momentum is low, multiple scattering is often the dominant source of error, as is the case in this experiment [1]. This introduces point-to-point correlations between the deviations of the measured positions from those of an ideal trajectory.

One way of dealing with such correlations is to use an error matrix that accounts for the effect of

scattering. In general, the elements of this matrix strongly depend on the measured positions, complicating the analysis. Furthermore, use of an error matrix explicitly assumes a Gaussian form for the error probability distribution; this is not true, since the Moliere scattering angle distribution is not Gaussian at large scattering angles.

Various other techniques already exist to deal with the effects of multiple scattering, including use of the Kalman filter [2–4] and fitting of a kinked trajectory [5]. However, these methods suffer from the same drawbacks mentioned above, namely, they apply to Gaussian errors and linear systems. Consequently, they are only good for small perturbations from an ideal trajectory. Furthermore, these techniques are most effective for determining the particle's coordinates and do not directly provide the best estimate of the particle momentum.

Here we present a maximum likelihood algorithm to estimate the particle momentum, which removes any assumption about Gaussian error distributions. We start with a general discussion of the maximum likelihood method in Section 2, then we apply these ideas to derive the best momentum estimate in Section 3. We will first derive the rigorous form of the likelihood function, taking into account the effects of multiple scattering, energy loss in detector and detector spatial resolution. We then simplify the expression when multiple scattering is the dominant source of error and show that the approximate form retains good accuracy.

## 2. General discussions of the likelihood method

Our goal is to determine a set of  $k$  parameters  $\alpha_1, \dots, \alpha_k$ , from  $N$  measurements  $(\mu_1, \dots, \mu_N)$ . For example, the parameters can be a particle's momentum and its initial position, and the measurement can be the position coordinates as it goes through the detector system. The likelihood method is used to determine the most likely values of those parameters given the measurements. Without loss of generality, we assume all  $\alpha$ 's and  $\mu$ 's are continuous variables. In that case,  $L$  should be understood as a likelihood density. In order to

calculate it, we need to consider finite intervals in the parameter and measurement space.

In general, the  $N$  measurements are constrained by the detector geometry so that they are not all independent of each other. Consider another measurement of  $(\mu_1 + \Delta\mu_1, \mu_2 + \Delta\mu_2, \dots, \mu_N + \Delta\mu_N)$ , also consistent with the detector geometry, where all  $\Delta$ 's are infinitesimally small. We can always assume that only the first  $m$  ( $m \leq N$ )  $\Delta$ 's are independent, and the other ones are linear combinations of the first  $m$   $\Delta$ 's. As a result, one has the following  $(N - m)$  equations:

$$\Delta\mu_i = M_{ij} \Delta\mu_j, \quad j = 1, \dots, m \quad i = (m + 1), \dots, N. \quad (4)$$

An example of non-independent measurements are the coordinates  $x$ ,  $y$  and  $z$  in a planar detector, for which only two of them are independent.

We are to find the probability  $L$  that the  $k$  parameters, with values  $\alpha_1, \dots, \alpha_k$  within infinitesimal variations  $\Delta\alpha_1, \dots, \Delta\alpha_k$ , produce the  $N$  measurements  $(\mu_1, \dots, \mu_N)$  within  $\Delta\mu_1, \dots, \Delta\mu_m$  of the first  $m$   $\mu$ 's. We note that all  $\Delta$ 's are only mathematical symbols having nothing to do with the actual finite detector resolution.

As a trivial generalization of Bayes' theorem [6],  $L$  can be written as

$$L = L_\alpha L_\beta \quad (5)$$

where  $L_\alpha$  is the a priori probability that the  $k$  parameters take values  $\alpha_1, \dots, \alpha_k$  within infinitesimal variations  $\Delta\alpha_1, \dots, \Delta\alpha_k$ . It is usually set to be a constant when performing the likelihood method, assuming all hypotheses are a priori equally likely.

$L_\beta$  is the probability of the  $k$  parameters producing the  $N$  measurements  $(\mu_1, \dots, \mu_N)$  within  $\Delta\mu_1, \dots, \Delta\mu_m$  of the first  $m$   $\mu$ 's. It is a function of both the  $\mu$ 's and the  $\alpha$ 's. In general, it also depends on the values of  $M_{ij}$ . Determining the functional form of  $L_\beta$  is the central task of the likelihood method.

## 3. Determining particle momentum using the maximum likelihood method

Next we utilize the general principles outlined above to derive the best estimate of the particle

momentum. Throughout the discussion, we assume that detected signals are correctly associated with the particles that we are measuring. In addition, we assume the detector thickness is negligible compared with the helical radius such that scattering can be treated as happening at one space point. In this section, we will first derive the rigorous form of the likelihood function, taking into account the effects of multiple scattering, energy loss in the detector and detector spatial resolution. Then we derive its approximate form when multiple scattering is the dominant source of error. Lastly we discuss the accuracy and performance.

### 3.1. The rigorous form of the likelihood function

We start with the simplest case where the only deviation from an ideal trajectory comes from multiple scattering at positions where the trajectory is measured. In that case,  $P$  is conserved;  $P_{\perp}$  and  $P_{\parallel}$  are not. We have  $3n$  successive measured coordinates:  $(x_1, y_1, z_1), \dots, (x_n, y_n, z_n)$  at detector elements  $1-n$ . Assuming all detectors are two-dimensional, one has  $N = 3n$  and  $m = 2n$ . If we choose  $\Delta x_1, \dots, \Delta x_n, \Delta y_1, \dots, \Delta y_n$  to be the independent variables, Eq. (4) becomes

$$\Delta z_i = s_{x_i} \Delta x_i + s_{y_i} \Delta y_i, \quad i = 1, \dots, n \quad (6)$$

where  $s_{x_i}$  and  $s_{y_i}$  depend on the orientation of detector element  $i$  if it is planar. If the detector is not planar, one needs to consider the tangent plane to the detector surface at the measurement point  $i$ .

We define five parameters  $P, n_1, n_2, \alpha_1$  and  $\alpha_2$ . The parameter  $P$  refers to the magnitude of the particle's initial momentum into the detector, and its direction is specified by  $n_1$  and  $n_2$ . The parameters  $\alpha_1$  and  $\alpha_2$  refer to  $x$  and  $y$  coordinates where the particle intercepts the first detector element. Taking  $L_{\alpha}$  to be constant, we have from Eq. (5)

$$L(P, \alpha_1, \alpha_2, n_1, n_2) \propto \delta(\alpha_1 - x_1) \delta(\alpha_2 - y_1) \Delta x_1 \Delta y_1 \times \left\{ \prod_{i=1}^{n-1} f_i(\theta_i) \Delta \Omega_i \right\} \quad (7)$$

where  $\delta$  represents the Dirac delta function, and  $f_i(\theta_i) d\Omega$  is the probability that a particle scatters

with angle  $\theta_i$  into  $d\Omega$  at position  $i$ .  $\Delta \Omega_i$  is the solid angle for a particle of momentum  $P$  following a helical trajectory exiting detector element  $i$  at point  $(x_i, y_i, z_i)$  to enter detector element  $(i+1)$  at point  $(x_{i+1}, y_{i+1}, z_{i+1})$  within  $\Delta x_{i+1}$  and  $\Delta y_{i+1}$ . Again, we are working under the assumption that the only deviation from an ideal trajectory comes from multiple scattering.

One can write down the following relation for each  $i$  ( $1 \leq i \leq (n-1)$ ):

$$\Delta \Omega_i \equiv Q_i \Delta x_{i+1} \Delta y_{i+1} \quad (8)$$

where  $Q_i$  is a non-negative Jacobian term arising from coordinate transformation. Its value depends on the geometry of detector element  $(i+1)$  as well as the helical trajectory between the two points  $(i)$  and  $(i+1)$ . It depends only on  $P$  and not on  $\alpha_1, \alpha_2, n_1$  or  $n_2$ . With no magnetic field or when  $P$  is very large such that the  $B$  field can be neglected, the particle follows a straight line trajectory and  $Q_i$  is independent of  $P$  and hence can be dropped. In general, when the  $B$  field cannot be neglected, the calculation of  $Q_i$  is straightforward but tedious. More details on  $Q_i$  will be discussed in the appendix.

In order to determine the most probable value of  $P$ , we define a likelihood as a function of  $P$  by integrating Eq. (7):

$$\begin{aligned} L(P) &\equiv \int \int \int \int L(P, \alpha_1, \alpha_2, n_1, n_2) d\alpha_1 d\alpha_2 dn_1 dn_2 \\ &\propto \int \int \int \int \delta(\alpha_1 - x_1) \delta(\alpha_2 - y_1) \\ &\quad \times \Delta x_1 \Delta y_1 \left\{ \prod_{i=1}^{n-1} f_i(\theta_i) \Delta \Omega_i \right\} d\alpha_1 d\alpha_2 dn_1 dn_2 \\ &= \int \int \Delta x_1 \Delta y_1 \left\{ \prod_{i=1}^{n-1} f_i(\theta_i) \Delta \Omega_i \right\} dn_1 dn_2 \quad (9) \end{aligned}$$

$$\begin{aligned} &= \Delta x_1 \Delta y_1 \Delta \Omega_1 \left\{ \prod_{i=2}^{n-1} f_i(\theta_i) \Delta \Omega_i \right\} \\ &\quad \times \int \int f_1(\theta_1) dn_1 dn_2 \quad (10) \end{aligned}$$

$$= \Delta x_1 \Delta y_1 \Delta \Omega_1 \left\{ \prod_{i=2}^{n-1} f_i(\theta_i) \Delta \Omega_i \right\} \quad (11)$$

$$= \prod_{i=2}^{n-1} f_i(\theta_i) \prod_{i=1}^{n-1} Q_i \left\{ \prod_{i=1}^n \Delta x_i \Delta y_i \right\} \\ \propto \prod_{i=2}^{n-1} f_i(\theta_i) \prod_{i=1}^{n-1} Q_i. \quad (12)$$

In going from Eq. (10) to (11), we used the relation that  $\int \int f_1(\theta_1) dn_1 dn_2 = \int \int f_1(\theta_1) d\Omega_1 = 1$ , if  $f_1$  is properly normalized. Note that Eq. (12) has a straightforward and intuitive interpretation: the overall likelihood is simply the product of the individual scattering probability at each detector plane because the scatterings are independent of each other. This way, we have effectively diagonalized the “error matrix”. Notice that  $f_1(\theta_1)$  and  $f_n(\theta_n)$  are missing because the incident angle entering the detector and the final exit angle are not related to the measured quantities.

The form of  $f_i(\theta_i)$  is approximately Gaussian [7]:

$$f_i(\theta) \propto \exp\left(-\frac{\theta^2}{\langle \theta^2 \rangle}\right)$$

with  $\langle \theta_i^2 \rangle^{1/2} = 21 \text{ MeV}(z/Pv)\sqrt{t_i/X_{0i}}$  where  $z$  is the particle’s charge,  $v$  is its velocity,  $X_{0i}$  is the radiation length of the  $i$ th material, and  $t_i$  is the path length traversed in the material, which depends on the detector thickness and the incident angle. This angle can be determined iteratively and does not need to be known to high precision. We note that the above Gaussian approximation breaks down for large values of  $\theta_i$ , due to the non-Gaussian Moliere scattering distribution.

To calculate  $\theta_i$ , we notice that the path between each pair of hits is helical, although the entire trajectory is not. For any value of  $P$ , one can reconstruct the path segment between points  $(i-1)$  and  $(i)$ . The incident direction at detector element  $(i)$  is simply the tangent of this path at point  $(i)$ . Similarly, the exit direction can be deduced by reconstructing the helical trajectory between points  $(i)$  and  $(i+1)$  and taking the tangent at point  $(i)$ . The angle between these two directions gives  $\theta_i$ . Therefore  $L(P)$  is fully calculable, and hence can be maximized. Taking the

helical trajectory between points  $(x_i, y_i, z_i)$  and  $(x_{i+1}, y_{i+1}, z_{i+1})$  for example, we denote the radius of the helix as  $r$  and the azimuthal angle subtended as  $\Delta\phi$ . If  $0 < \Delta\phi < 2\pi$ , one then has

$$\sin\left(\frac{1}{2}\Delta\phi\right) = \frac{d}{2r} \quad (13)$$

where

$$d = \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2}$$

and the  $B$  field is assumed to be along the  $z$  direction. Eq. (13) has two solutions, one when  $0 < \Delta\phi \leq \pi$  and other when  $\pi < \Delta\phi < 2\pi$ :

$$\Delta\phi_1 = 2\sin^{-1}\left(\frac{d}{2r}\right) \quad (14)$$

and

$$\Delta\phi_2 = 2\pi - 2\sin^{-1}\left(\frac{d}{2r}\right). \quad (15)$$

If we allow the possibility that  $\Delta\phi > 2\pi$ , we then have, in the general case,

$$\Delta\phi = 2\sin^{-1}\left(\frac{d}{2r}\right) + 2n\pi, \quad n = 0, 1, 2, \dots$$

or

$$\Delta\phi = 2\pi - 2\sin^{-1}\left(\frac{d}{2r}\right) + 2n\pi, \quad n = 0, 1, 2, \dots$$

The non-negative integer  $n$  represents the number of complete helical turns in this segment. The allowed values of  $n$  depend on the detector geometry. In many cases only  $n = 0$  is allowed, which we assume is the case here to simplify the following discussions. Therefore we have two equations relating  $\Delta\phi$  and  $r$ , Eqs. (14) and (15).

Inserting Eqs. (1) and (2) into Eq. (3), one has

$$r^2 + \frac{(\Delta z)^2}{(\Delta\phi)^2} = \frac{P^2}{C^2} \quad (16)$$

where  $\Delta z = z_{i+1} - z_i$ . Inserting the two choices for  $\Delta\phi$  into Eq. (16), one gets

$$\left(\frac{P}{Cd}\right)^2 = \left(\frac{r}{d}\right)^2 + \frac{(\Delta z/d)^2}{(2\sin^{-1}(d/2r))^2} \quad (17)$$

or

$$\left(\frac{P}{Cd}\right)^2 = \left(\frac{r}{d}\right)^2 + \frac{(\Delta z/d)^2}{(2\pi - 2\sin^{-1}(d/2r))^2} \quad (18)$$

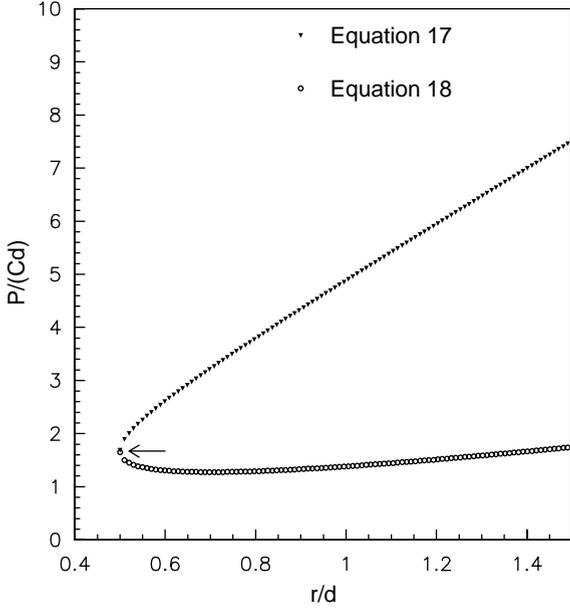


Fig. 2. An example of  $P/Cd$  versus  $r/d$  from Eqs. (17) and (18), for a specific case when  $\Delta Z/d = 5$ , plotted for  $r/d < 1.5$ , a value arbitrarily chosen. The two curves monotonically increase to infinity beyond  $r/d = 1.5$ . These two equations correspond to  $\Delta\phi \leq \pi$  and  $> \pi$ , respectively. The arrow points to where the two curves connect at  $\Delta\phi = \pi$  when  $r/d = 0.5$  and  $(P/Cd)^2 = \frac{1}{4} + (\Delta z/d)^2/\pi^2$ .

where  $C$  is a constant defined earlier. Fig. 2 shows an example of  $P/Cd$  versus  $r/d$  from Eqs. (17) and (18) for a specific case when  $\Delta Z/d = 5$ . The arrow points to where the curves for the two equations connect at  $\Delta\phi = \pi$  when  $r/d = 0.5$  and  $(P/Cd)^2 = \frac{1}{4} + (\Delta z/d)^2/\pi^2$ .

The value of  $r$  (and hence  $\Delta\phi$ ) can be calculated in terms of  $P$  from Eq. (17) or (18). It is instructive to look at Fig. 2 for the general behavior of these two equations. We first note that  $r/d$  always exceeds 0.5. The RHS of Eq. (17) is an increasing function of  $r/d$  with a minimum value,  $m_1 = \frac{1}{4} + (\Delta z/d)^2/\pi^2$  at  $r/d = 0.5$ ; Eq. (17) has a solution only if  $(P/Cd)^2 \geq m_1$ . The RHS of Eq. (18) is a little more complicated. Starting from  $r/d = 0.5$ , it is a decreasing function of  $r/d$  until it reaches its minimum value of  $m_2$ , then it increases monotonically. As a result, for  $(P/Cd)^2 < m_2$  there is no solution for  $r$  in either case, therefore there is a lower cut-off on  $P$ , at  $P = Cd\sqrt{m_2}$ , and Eq. (18)

has one solution while Eq. (17) has no solution at this value of  $P$ . For  $m_1 > (P/Cd)^2 > m_2$ , Eq. (18) has two solutions while Eq. (17) has no solution. For  $(P/Cd)^2 \geq m_1$ , each of Eqs. (17) and (18) has one solution. To summarize, there are a total of two solutions for  $(P/Cd)^2 > m_2$ , and the two solutions degenerate into one at  $(P/Cd)^2 = m_2$ . There is no solution for  $(P/Cd)^2 < m_2$ . Each solution of  $r$  and  $\Delta\phi$  corresponds to two possibilities of the helical portion between the two points, one going clockwise and one going counter-clockwise. For a known  $B$  field and particle charge, only one possibility remains. Hence we have shown that  $\theta_i$  can be derived in terms of  $P$ ,  $x_{i-1}$ ,  $y_{i-1}$ ,  $z_{i-1}$ ,  $x_i$ ,  $y_i$ ,  $z_i$ ,  $x_{i+1}$ ,  $y_{i+1}$  and  $z_{i+1}$ .

As we have seen, there are two solutions for  $r$  and  $\Delta\phi$  when  $(P/Cd)^2 > m_2$ . Consequently, one needs to examine both solutions and exclude the incompatible ones, for which the trajectory will have additional crossings at the detector between points  $i$  and  $(i+1)$ . When both segments are compatible, they both have to be considered. As a result, Eq. (12) needs to be summed over all allowed combinations of segments:

$$L(P) = \sum_{j=1}^{N_c} \left\{ \prod_{i=2}^{n-1} f_i(\theta_i^j) \prod_{i=1}^{n-1} Q_i^j \right\} \quad (19)$$

where the summation  $j$  goes over all allowed segment combinations. In practice, this sum is usually dominated by one particular combination and Eq. (19) reduces to Eq. (12).

The feature of a cut-off momentum reflects the fact that  $P$  cannot be arbitrarily low given the measured points. Since each pair of adjacent points gives one such cut-off, when considering all the  $n$  points, the maximum of these  $(n-1)$  cut-off values is the overall cut-off in  $P$ , which we denote as  $P_c$ . As a result, Eq. (19) is defined only for  $P \geq P_c$ , and  $L$  is taken to be zero for  $P < P_c$ .

When including the effect of energy loss in the detector,  $L$  takes the following form:

$$L(P) = \int_0^P dP_1 \int_0^{P_1} dP_2 \dots \int_0^{P_{n-2}} dP_{n-1} h(P, P_1) \times \sum_{j=1}^{N_c} \left\{ \prod_{i=2}^{n-1} g_i(\theta_i^j, P_{i-1}, P_i) \prod_{i=1}^{n-1} Q_i^j \right\} \quad (20)$$

where  $P_i$  ( $i = 1, \dots, (n-1)$ ) is the particle's momentum between hits ( $i$ ) and ( $i+1$ ). We define  $h(P, P_1) dP_1$  to be the probability that the particle with initial momentum  $P$  has final momentum  $P_1$  within  $dP_1$  after going through the first detector element. For  $i = 2, \dots, (n-1)$ ,  $g_i(\theta_i^j, P_a, P_b) d\Omega dP_b$  is the probability of the particle with initial momentum  $P_a$  being scattered into the final solid angle space and momentum intervals with final momentum  $P_b$  in the  $i$ th detector element. One can calculate  $\theta_i^j$  the same way as before, except that each segment now has a different momentum value.

If one also takes into account the detector spatial resolution, Eq. (20) should be replaced by

$$L(P) = \int \dots \int dx'_1 dy'_1 \dots dx'_n dy'_n \times L'(P, x'_1, y'_1 \dots x'_n, y'_n) \prod_{i=1}^n r_i(x_i, y_i, x'_i, y'_i) \quad (21)$$

where  $(x_i, y_i)$  denote the measured coordinates and  $(x'_i, y'_i)$  denote the actual coordinates. The difference between them represent the measurement error and is characterized by the resolution function  $r_i(x_i, y_i, x'_i, y'_i)$ . The function  $r_i$  is the probability density function that measurement  $(x_i, y_i)$  corresponds to actual coordinates  $(x'_i, y'_i)$ . The expression  $L'(P, x'_1, y'_1 \dots x'_n, y'_n)$  is the same as  $L(P)$  as defined in Eq. (20), except that it is evaluated for the actual coordinates.

To summarize, Eq. (21) is the rigorous form of the likelihood of  $P$ , when random processes include multiple scattering, energy loss and detector spatial resolution. It may be too complicated for practical use, and we next develop its approximate forms to be used in actual calculations.

### 3.2. The approximate forms of the likelihood function

Both multiple scattering and spatial resolution introduce deviations in the positions with respect to those on a perfect helix. The contribution from multiple scattering is on the order of  $l \langle \theta_i^2 \rangle^{1/2}$ , where  $l$  is the typical helix length between two hits.

In many cases this contribution to the deviation is much larger than that due to the spatial resolution and the latter can be neglected. In this case, Eq. (21) reduces to Eq. (20). For example, in this experiment [1], the deviation from multiple scattering is  $\sim 0.3$  cm, and the spatial resolution is 0.02 cm for the transverse coordinates and 0.15 cm for the longitudinal coordinate, and we find that neglecting the spatial resolution in the likelihood formulae (while keeping the effect of spatial resolution in the Monte Carlo simulation) gives good results in determining the momentum, as will be shown in Section 3.3.

Eq. (20) can be further simplified by assuming an average energy loss at each detector. One then has

$$L(P) \approx \sum_{j=1}^{N_c} \left\{ \prod_{i=2}^{n-1} f_i(\theta_i^j(P_{i-1}, P_i)) \prod_{i=1}^{n-1} Q_i^j \right\} \quad (22)$$

in which we explicitly write down  $\theta_i$  as a function of  $P_{i-1}$  and  $P_i$ . Instead of being integration variables,  $P_1 \dots P_{n-1}$  are now determined from  $P$ . They are given by

$$P_i = P - \sum_{j=1}^i t_j C_{ej} \quad (23)$$

where  $C_{ej}$  is the average energy loss per unit length for the  $j$ th detector element, and  $t_j$  is the path length traversed through the  $j$ th detector element.

The approximation in Eq. (23) reduces the likelihood with respect to the maximum value determined for a fit that includes the effects of energy loss. It also introduces an uncertainty in the energy on the order of the dispersion in the energy loss in the detector. This approximation compensates for the average energy loss, whereas the actual energy loss distribution falls off rapidly on the low-energy side and has longer high-energy tails. As a result, it introduces low-energy tails in the resolution function. No significant high-energy tails are induced which is crucial for this experiment [1].

We stated earlier that calculations of the  $Q$ 's are, in general, tedious. Fortunately, in many cases they vary much less with  $P$  compared with other terms such that they can be taken as constants and

dropped. Eq. (22) then becomes

$$L(P) \approx \sum_{j=1}^{N_c} \left\{ \prod_{i=2}^{n-1} f_i(\theta_i^j(P_{i-1}, P_i)) \right\}. \quad (24)$$

Eq. (24) is much simpler to use compared with Eq. (21). In the next section we will show that it also has good accuracy.

### 3.3. Accuracy and performance

The uncertainty in the determination of  $P$  can be estimated by how fast  $L$  falls off when  $P$  deviates from  $P_m$ , the value of  $P$  when  $L$  is maximal. If the likelihood curve is near Gaussian, the uncertainty can be defined by fitting  $L(P)$  to a Gaussian function around  $P = P_m$ . In general, a measure of the quality of the fit is the shape of  $L(P)$ , and a selection criteria based on this shape can be used to eliminate backgrounds. This selection will depend on the details of the backgrounds one is trying to eliminate.<sup>1</sup> For example, in the case of experiment [1] where most of the unwanted backgrounds come from high-energy tails in the resolution function exceeding  $\sim 1$  MeV/ $c$ , one can minimize backgrounds by requiring that the likelihood value fall by more than some arbitrary factor when  $\Delta P \sim 1$  MeV/ $c$ , where  $\Delta P = P_m - P$ .

In principle, the shape of  $L(P)$  (as opposed to its absolute value) alone dictates the uncertainty in the determination of  $P$ . However, in practice, we also require a minimum value for  $L(P_m)$  to ensure a reliable measurement. This requirement will reduce the probability of finding the wrong trajectory when the detector signals are contaminated with noise signals. In addition to deriving  $P_m$  and the shape of  $L(P)$ , the proposed fitting method evaluates the individual scattering angle,  $\theta_i$ , at each detector crossing for a given momentum. One can additionally require that the values of  $\theta_i$  evaluated at  $P = P_m$  not exceed some

maximum, although this is highly correlated with requiring that  $L(P_m)$  exceed some value. This requirement on  $\theta_i$  is necessary when the form of  $f_i(\theta_i)$  is not well known for large values of  $\theta_i$ .

Eq. (24) is an approximate form for  $L$ , but it gives good results. It has been used in a Monte Carlo simulation of this experiment [1] for momentum determination by maximizing the value of  $L$  as calculated by the above equation. The simulation includes the full effects of multiple scattering and energy straggling. Spatial resolution of 0.02 cm for the transverse coordinates and 0.15 cm for the longitudinal coordinate is also included in the simulation. When the momentum determined by Eq. (24) is compared with the actual momentum entering the detector (around 105 MeV/ $c$ ), a resolution of  $\sigma \sim 150$  keV/ $c$  is achieved.

Fig. 3 shows an example of the shape of  $L$  as a function of  $P$  for a typical simulation event. Again,  $L$  is calculated using Eq. (24). The curve peaks at about 104.73 MeV/ $c$ , while the particle's actual momentum into the detector is 104.85 MeV/ $c$ . One sees that  $L(P)$  is narrow, with a FWHM of  $\sim 200$  keV/ $c$ . The overall cut-off momentum is  $P_c = 100.3$  MeV/ $c$ .

## 4. Conclusions

In conclusion, we have devised a likelihood algorithm to determine charged particle momenta and assess their uncertainties by using position measurements in a magnetic field. Both the rigorous and approximate forms are presented. Eq. (21) is the rigorous form when random processes include non-Gaussian multiple scattering, energy loss and detector spatial resolution.

This method is particularly simple and useful when multiple scattering is the dominant source of uncertainties. Because the scatterings are independent of each other, the likelihood is simply the product of the individual scattering probability at each detector plane, modified by some Jacobian terms. The calculated scattering angles at each scattering point have very sensitive dependence on  $P$ , which makes the likelihood method effective in the momentum determination.

<sup>1</sup>In general, if the objective is to minimize a cost function arising from the incorrect determination of  $P$ , the cost function will depend on the shape of  $L(P)$  [8], and the most probable value of  $P$  may not be the optimal estimate depending on the form of the cost function. However, we ignore this subtlety in this paper and use  $P_m$  as the estimate of  $P$ .

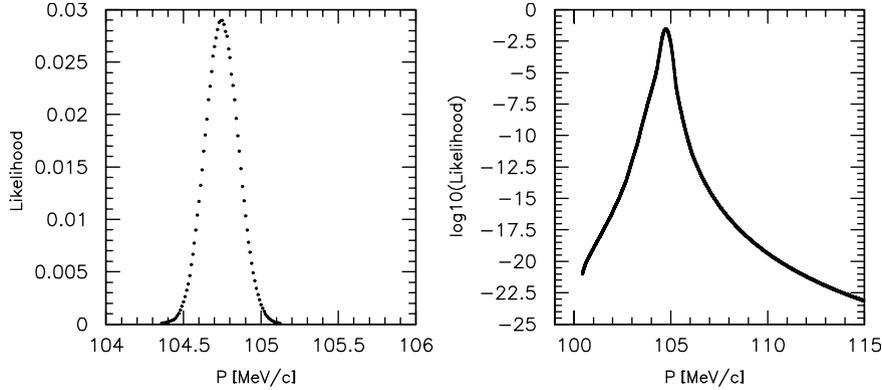


Fig. 3. An example of the likelihood  $L$  (left) and  $\log_{10}(L)$  (right) versus momentum  $P$  from a Monte Carlo generated event based on Eq. (24). Here  $L$  peaks at 104.73 MeV/ $c$ , while the particle's actual momentum entering the detector is 104.85 MeV/ $c$ .

The likelihood formulae can be simplified under various conditions. If the detector's spatial resolution contributes little to the deviations from a perfect helical trajectory, the spatial resolution can be safely ignored; if the dispersion in the energy loss in the detector is much smaller than the required energy resolution, the particle's energy loss in each detector can then be approximated by the calculated mean energy loss; if the dependence of Jacobian terms in the scattering probability do not depend strongly on  $P$ , they can be dropped. When all of the above conditions apply,  $L$  can be approximated by Eq. (24), which is relatively simple to implement without sacrificing accuracy in the momentum measurement.

### Acknowledgements

This work is supported by DOE Grant DE-FG03-91ER40679. We also thank an anonymous referee for helpful suggestions in making this presentation more coherent and clear.

### Appendix A. More discussions on the $Q$ 's

We here examine the  $Q$ 's as defined in Eq. (8) in more detail, mainly for mathematical curiosity. We calculate its value for the segment between points ( $i$ ) and ( $i + 1$ ) by differentiating Eqs. (13)

and (16). The calculations are straightforward although tedious, and one of the authors has derived the following expression:

$$Q = \left| \frac{\tan \theta \sin \theta}{2r^3(\Delta\phi)^3 \cos(\Delta\phi/2) + 2(\Delta z)^2 d} \times \left( \Delta z \Delta\phi s_y + \frac{2(\Delta z)^2}{d} - \frac{2\Delta z}{d} r \Delta\phi \cos\left(\frac{\Delta\phi}{2}\right) s_x \right) \right| \quad (\text{A.1})$$

where the symbols used in Eqs. (13)–(18) are carried over, and  $s_x$  and  $s_y$  are the same as defined in Eq. (6). They refer to detector element ( $i + 1$ ) and the coordinates are conveniently chosen such that  $z$  is along the  $B$  field, and the line joining points ( $i$ ) and ( $i + 1$ ) is in the  $xz$  plane. The angle between this line and the  $z$ -axis is denoted by  $\theta$ , and  $\tan \theta = d/\Delta z$ . Note that  $Q$  has different values for the two solutions of the trajectory. In general,  $Q$  has a complicated dependence on  $P$  through its dependence on  $\Delta\phi$  and  $r$ . We examine two special cases below.

In the limit when  $\Delta\phi \rightarrow 0$ , which happens when  $d/r = |q|Bd/cP_{\perp} \ll 1$  such that the helical portion is approximately a straight line,  $r \Delta\phi$  approaches  $d$ , and Eq. (A.1) reduces to

$$Q = \frac{\sin^2 \theta}{d^2} (\cos \theta - s_x \sin \theta) \quad (\text{A.2})$$

as can be readily verified when taking all trajectories to be straight lines. In this case,  $Q$

has no dependence on  $P$  and hence can be safely dropped.

Next we examine a striking feature of  $Q$ , that is it diverges when  $\gamma \equiv 2r^3(\Delta\phi)^3 \cos(\Delta\phi/2) + 2(\Delta z)^2 d = 0$ . Obviously  $\Delta\phi > \pi$  when that happens. Furthermore, we show that this happens at the cut-off momentum when  $r = r_m$  and the RHS of Eq. (18) reaches its minimum value of  $m_2$ . We remind ourselves that  $P$  in  $L(P)$  is bounded from below by  $P_c$ , the overall cut-off momentum, which is the maximum of the  $n - 1$  cut-off momenta derived for each pair of adjacent hits. Without loss of generality, we assume that  $P_c$  is defined by the segment joining points ( $i$ ) and ( $i + 1$ ). Denoting the RHS of Eq. (18) as

$$f(r) \equiv \left(\frac{r}{d}\right)^2 + \frac{(\Delta z/d)^2}{(2\pi - 2\sin^{-1}(d/2r))^2},$$

one can calculate its derivative as

$$f'(r) = \frac{\gamma}{r^2 d^2 (\Delta\phi)^3 \cos(\Delta\phi/2)}. \quad (\text{A.3})$$

Since  $f'(r_m) = 0$ , one has  $\gamma = 0$  at  $r = r_m$  or  $P = P_c$  such that  $Q$  diverges.

Probability conservation requires  $\int L(P) dP$  to converge. Because of the divergence of  $Q$ , this may present a problem in Eq. (22). To check this, we examine the segment that defines  $P_c$  and require the convergence of  $\int_{P_c}^{P_c+\varepsilon} Q dP$  for the two solutions for the trajectory, where  $\varepsilon$  is small and positive.<sup>2</sup> To check this, we Taylor-expand  $f(r)$  around  $r_m$ . It is easy to show that its second derivative is positive, i.e.  $f''(r_m) > 0$ , so one has  $f(r) \approx (P_c/Cd)^2 + \frac{1}{2}f''(r_m)(r - r_m)^2$  and

$$f'(r) \approx f''(r_m)(r - r_m). \quad (\text{A.4})$$

Eq. (18) becomes  $P - P_c \approx (C^2 d^2 / 4P_c) f''(r_m) (r - r_m)^2$  and the two solutions of  $r$  become

$$r - r_m \approx \pm \frac{2}{Cd} \sqrt{\frac{P_c}{f''(r_m)}} (P - P_c)^{1/2} \quad (\text{A.5})$$

<sup>2</sup>The two solutions for the trajectory approach each other when  $P$  approaches  $P_c$  and they degenerate into one at  $P = P_c$ . The two segments reinforce each other in the sum of contributions to  $L$  in Eq. (19).

From Eqs. (A.3)–(A.5), one has

$$\gamma \approx \pm \frac{2}{Cd} r^2 d^2 (\Delta\phi)^3 \cos\left(\frac{\Delta\phi}{2}\right) \times \sqrt{P_c f''(r_m)} (P - P_c)^{1/2}. \quad (\text{A.6})$$

Eqs. (A.1) and (A.6) give the leading term in  $Q$  when  $P$  approaches  $P_c$ :

$$Q \approx c_P (P - P_c)^{-1/2} \quad (\text{A.7})$$

where the constant  $c_P$  is the absolute value of the expression

$$\frac{\tan \theta \sin \theta Cd}{2r^2 d^2 (\Delta\phi)^3 \cos(\Delta\phi/2) \sqrt{P_c f''(r_m)}} \times \left( \Delta z \Delta\phi s_y + \frac{2(\Delta z)^2}{d} - \frac{2\Delta z}{d} r \Delta\phi \cos\left(\frac{\Delta\phi}{2}\right) s_x \right)$$

evaluated at  $P = P_c$ . This  $c_P$  is the same for the two solutions of  $r < r_m$  and  $r > r_m$ . For either solution, we have  $\int_{P_c}^{P_c+\varepsilon} Q dP = 2c_P \sqrt{\varepsilon}$  and it does converge.

In general, the inclusion of the  $Q$ 's will change the shape of the likelihood function. The change is usually small except at  $P = P_c$  where  $Q$  approaches infinity. This surprising feature would make  $P_c$  the most probable momentum, regardless of the quality of the helix fit. In practice, this problem is resolved because the objective is to estimate  $P$  within some tolerance  $\varepsilon$ , such that we need to compare the integrals of  $\int_{P_c}^{P_c+\varepsilon} L dp$  and  $\int_{P_m-(1/2)\varepsilon}^{P_m+(1/2)\varepsilon} L dp$ . Usually the former is negligible compared with the latter, such that  $P = P_c$  is not a viable choice. Furthermore, when  $P$  is away from  $P_c$ , the  $Q$ 's usually vary much less with  $P$  compared with the dependence of scattering angles on  $P$ . As a result, the  $Q$ 's can be dropped, and the approximation is valid.

Lastly, we point out that the divergence of  $L$  caused by that of  $Q$  at  $P = P_c$  will disappear if the detector spatial resolution or energy loss dispersion in the detector is taken into account in formulating  $L$ , such as in Eqs. (20) and (21).

## References

- [1] M. Bachman, et al., A search for  $\mu^- N \rightarrow e^- N$  with sensitivity below  $10^{-16}$ , AGS Proposal E940, 1997, This

document and the updates of the experiment can be most easily found at <http://mecop.ps.uci.edu/>.

- [2] E. Wolin, L. Ho, Nucl. Instr. and Meth. A 329 (1993) 493.
- [3] R. Fruhwirth, Nucl. Instr. and Meth. A 262 (1987) 444.
- [4] P. Billoir, Nucl. Instr. and Meth. 225 (1984) 352.
- [5] G. Lutz, Nucl. Instr. and Meth. A 273 (1988) 349.
- [6] J. Mathews, R.L. Walker, Mathematical Methods of Physics, 2nd Edition, Addison Wesley Publishing company, Inc., Menlo Park, CA, 1970.
- [7] D.H. Perkins, Introduction to High Energy Physics, 3rd Edition, Addison Wesley Publishing company, Inc., Menlo Park, CA, 1986.
- [8] The analytic Sciences Corporation, Applied Optimal Estimation, 1974.