TRACK FITTING WITH MULTIPLE SCATTERING: A NEW METHOD

Pierre BILLOIR

Laboratoire de Physique Corpusculaire, Collège de France, Paris, France

An analytical calculation of the variance is performed, in some simple cases, for standard least-squares estimators of track parameters (accounting for independent measurement errors only); comparison is made with optimal estimators (accounting also for scattering errors, correlated between one point and the following ones). A new method is proposed for optimal estimation: the points measured on the track are included backwards, one by one, in the fitting algorithm, and the scattering is handled locally at each step. The feasibility of the method is shown on real events, for which the geometrical resolution is improved. The algorithm is very flexible and allows fast programmation; moreover the computation time is merely proportional to the number of measured points, contrary to the other optimal estimators.

1. Introduction

Charged particles going through matter are affected by random deviations due to multiple scattering. The uncertainty on their initial fitted parameters (position and momentum components) arises from two contributions: on the one hand the contribution of the measurement errors, which is a decreasing function of the number of measured points; on the other hand the contribution of the multiple scattering errors, which cannot be reduced below some minimum values, because the detectors contain at least a gas at atmospheric pressure, and some denser parts. These scattering uncertainties become predominant for low momentum particles, or for very accurate measurements.

There are essentially two ways to perform a geometrical track fit when the scattering errors are not negligible:

1) *A standard fit* takes only the measurement errors into account, and provides estimators for the track parameters and their covariance matrix (as resulting from the measurement errors only). Such estimators are not optimal (some information is lost). Their actual covariance matrix can be obtained by adding afterwards the multiple scattering contribution.

2) *An optimal fit* makes use of the full ($n \times n$) covariance matrix of the n measurements (including multiple scattering, i.e. correlation terms). This is usually realized by using the Gauss–Markow theorem, and needs then the inversion of this matrix [1]; other methods introduce extra parameters to describe the scattering angles [1].

The precision of the estimators corresponding to both methods has already been compared by Drijard [2] by numerical evaluation, for the curvature and angle parametes only. In this paper we give analytical expressions of the covariance matrices, and we consider also position parameters, because the kinematical resolution depends on the precision of these parameters when the tracks are extrapolated to a vertex. Moreover, it is necessary to calculate correctly the covariance matrix of the parameters (accounting for multiple scattering) in order to use the χ^2 as a goodness-of-fit criterion in a subsequent vertex fit or kinematical analysis, or to detect another source of errors.

In sect. 2 we calculate the matrices resulting from the standard method in some simple cases, and we show that an increase of the number of measured points may lead to an increasing error in the parameters; in sect. 3 we describe a recursive solution of the optimal method, hopefully less expensive than the brute-force matrix inversion involved in the Gauss–Markov method; in sect. 4 we determine recursively the optimal covariance matrix of the parameters in the same cases as in sect. 2 in order to evaluate the gain in

precision. In sect. 5 we present results obtained with real events from the OMEGA spectrometer at CERN.

Let us recall the expression of the variance of the projected scattering angle for a particle of charge Z, velocity v, momentum p, after passing through ξ radiation lengths:

$$\Delta\alpha^2 = \left(\frac{KZ}{pv}\right)^2 \xi,$$

where K is about 15 MeV (with a correcting factor at low velocities).

The main point for our purpose is that for a given particle with a given momentum, $\Delta\alpha^2$ is proportional to ξ.

2. The standard fit: evaluation of the error matrix

2.1. General conventions

We suppose that all measured points on the track are equidistant (interval l), with the same precision, and that their number n is large enough to replace the summations over the points by integrations (or, equivalently, to hold only the terms of highest degree in n).

We consider slightly curved tracks in an uniform vertical magnetic field, so that they have a straight vertical projection, and their horizontal circular projection can be handled as a parabola in order to perform easier analytical calculations. These conditions are often realistic, and our conclusions will hold qualitatively in more general cases. The calculation could be extended to various track/detector configurations.

2.2. Scattering at one point of the track

We suppose that the particle encounters matter concentrated at a distance L_0 after the first measured point. Let $n_0 = L_0/l$ be the number of points before L_0, and $\Delta\alpha^2 = (KZ/pv)^2\xi$ the variance of the scattering angle α in both vertical and horizontal directions. α is a quasi-Gaussian variable, independent of the measurement errors.

At n fixed abscissas $x_i = il$, the coordinate y is measured in the horizontal plane with precision σ_h and z in the vertical plane with precision σ_v. Their deviations from the theoretical ones (which would be obtained without errors and scattering) are illustrated in fig. 1.
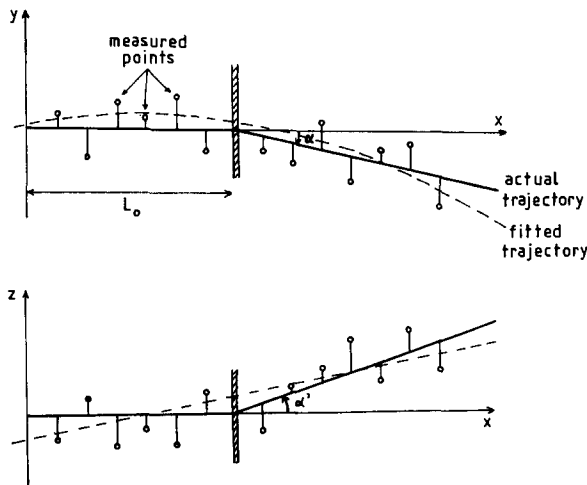

Fig. 1. Deviations of the measured coordinates from the theoretical trajectory.

The vertical projection is then parametrized by $z = Z + bx$ and the horizontal one by $y = Y + ax + cx^2/2$ (with $c = 1/R$ when the track has a direction close to x axis). The parameters to be fitted are, in the first case, Z and b, and their variances can be calculated [4] as functions of σ, $\eta = n/n_0$ and $\rho^2 = n_0^3 l^2 \Delta\alpha^2/\sigma^2$

$$\sigma_Z^2 = \frac{4}{n}\sigma_v^2 \left[1 + \frac{(\eta - 1)^4}{4\eta^3} \rho_v^2 \right]$$

$$\sigma_b^2 = \frac{12}{n} \left(\frac{\sigma_v}{L} \right)^2 \left[1 + \frac{(\eta - 1)^4 (2 + \eta)^2}{12\eta^3} \rho_v^2 \right].$$

In the second case, the variances of the parameters Y, a, c are:

$$\sigma_Y^2 = \frac{9}{n}\sigma_h^2 \left[1 + \frac{(2\eta - 5)^2 (\eta - 1)^4}{36\eta^5} \rho_h^2 \right]$$

$$\sigma_a^2 = \frac{192}{n} \left(\frac{\sigma_h}{L} \right)^2 \left[1 + \frac{(\eta - 3)^3 (\eta + 5)^2 (\eta - 1)^4}{192\eta^5} \rho_h^2 \right]$$

$$\sigma_c^2 = \frac{180}{n} \left(\frac{\sigma_h}{L^2} \right)^2 \left[1 + \frac{5(\eta - 1)^4}{4\eta^5} \rho_h^2 \right].$$

In these expressions the first term represents the well known contributions of the measurement errors.

In order to show how the errors on the parameters depends on the measured length (i.e. the number of points), figs. 2 and 3 (solid curves) represent the square roots of the variances (in normalized units) as function of n/n_0 for various values of ρ: this quantity expresses the relative importance of the scattering errors for a given track/detector configuration; under usual experimental conditions, it can be much greater than 1 for particles with momentum around 1 GeV/c.

These curves often have an absolute minimum (and sometimes also a secondary minimum): so we are led to the notion of optimal measurement length: using points beyond this length worsens the precision of the estimator. Unfortunately the optimal length is generally not the same for all parameters; in many cases the best precision for the curvature is obtained by using all measurements.

It is important to note that the mean squared residual r^2 is not affected in the same way by the scattering errors: one finds, for large n:

$$r_v^2 \simeq \sigma_v^2 \left[1 + \frac{(\eta - 1)^3}{3\eta^4} \frac{\rho_v^2}{n_0} \right] \quad \text{in the vertical plane}$$

$$r_h^2 \simeq \sigma_h^2 \left[1 + \frac{(\eta - 1)^3 (4\eta^2 - 15\eta + 15)}{12\eta^6} \frac{\rho_h^2}{n_0} \right] \quad \text{in the horizontal plane.}$$

So the relative increase of the residuals due to the multiple scattering is generally much smaller that the increase of the variances of the fitted parameters. This implies that the χ^2 criterion applied to the residuals fails to indicate whether the standard estimator is reliable in spite of the scattering.

When the particles encounter several slices of matter inside the measurement range, the successive scattering angles are independent: so the variances and the mean squared residuals can be calculated by adding the contribution of each slice.

2.3. Scattering uniformly distributed along the track

This approximation is exact for bubble chambers and realistic for many detectors. The variance of the parameters and the mean squared residuals are deduced from the formulae found in 2.2, with an integration over L_0 (as above, the number n of measurement points is assumed to be large enough to allow this approximation).

We call $\delta\alpha^2$ the variance of the scattering angle between two successive points (i.e. over a length l); the dimensionless quantity $q = l\delta\alpha/\sigma$ plays a role analogous to that of ρ in 2.2. One finds:

$$\left. \begin{array}{l} \sigma_Z^2 = \dfrac{4\sigma_v^2}{n}\left(1 + \frac{1}{420}n^4 q_v^2\right) \\[2mm] \sigma_b^2 = \dfrac{12}{n}\left(\dfrac{\sigma_v}{L}\right)^2\left(1 + \frac{13}{420}n^4 q_v^2\right) \\[2mm] r_v^2 = \sigma_v^2\left(1 + \frac{1}{420}n^3 q_v^2\right) \end{array} \right\}$$ in the vertical plane.
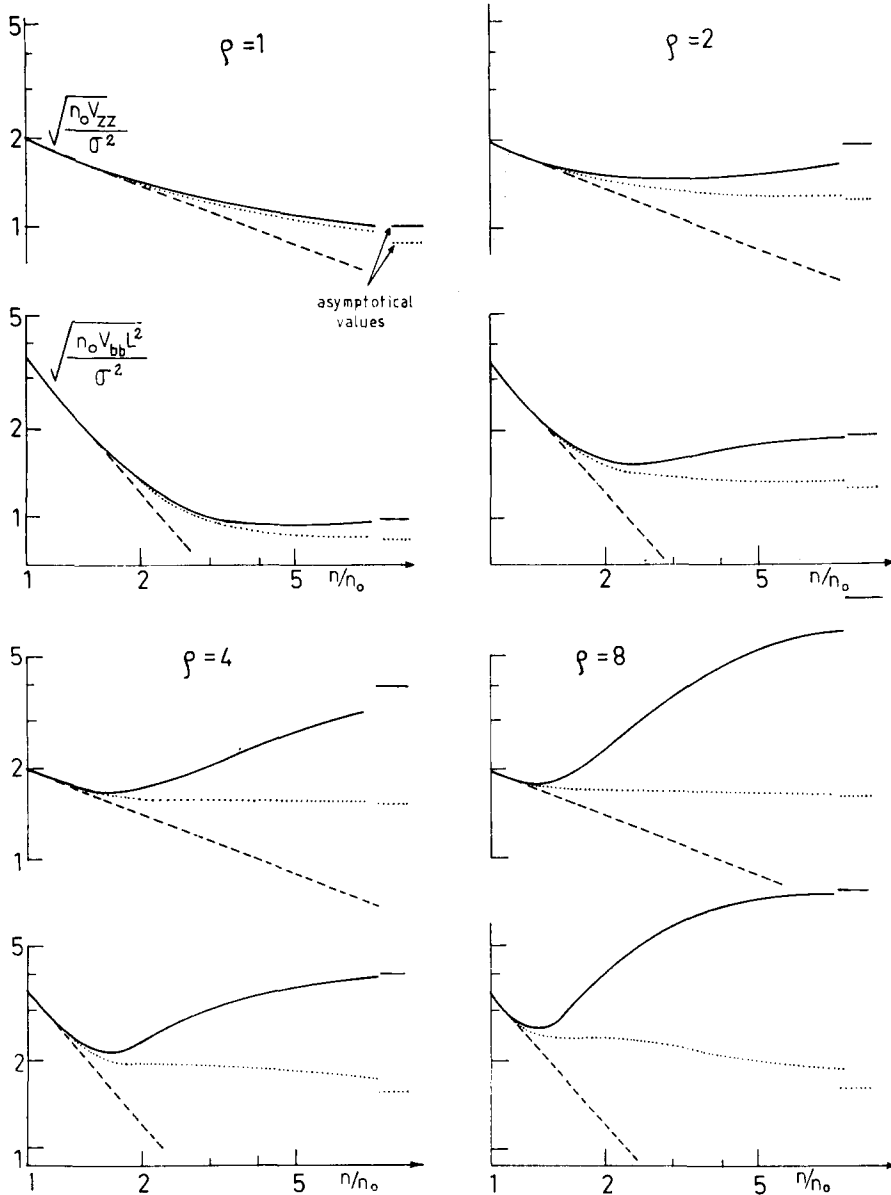

Fig. 2. Errors on the track parameters (position and slope at first point, and eventually curvature) as functions of the measured length for both standard and optimal estimators: *straight track, scattered at one point.*
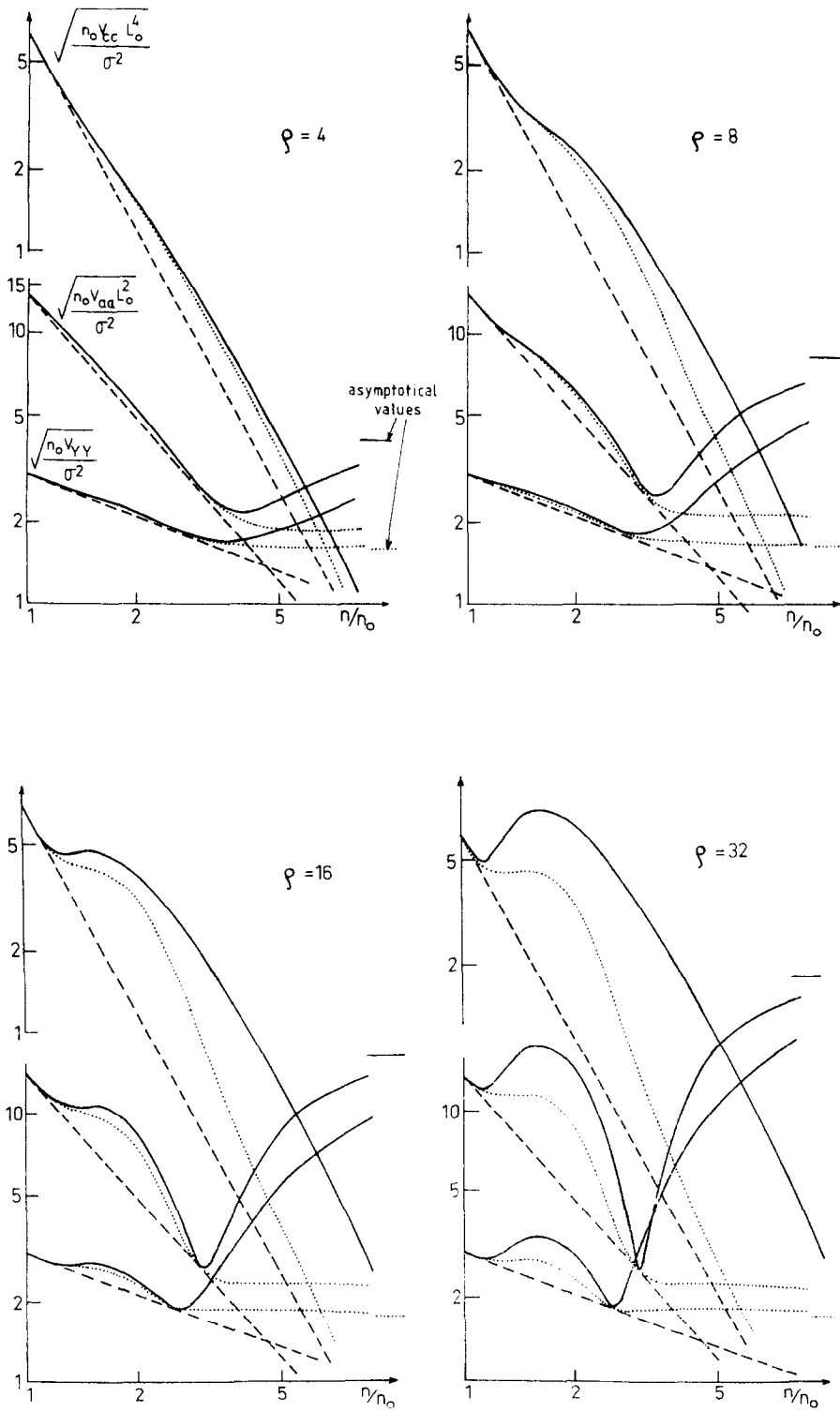
Fig. 3. Errors on the track parameters, as in fig. 2 but for: *curved track, scattered at one point.*
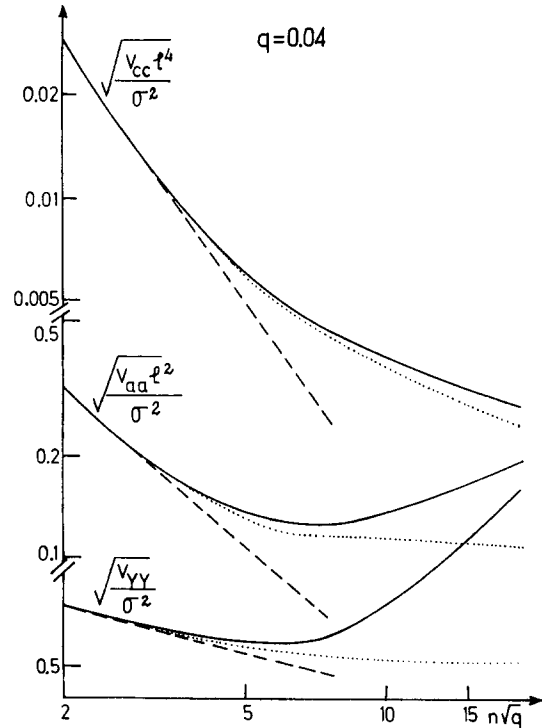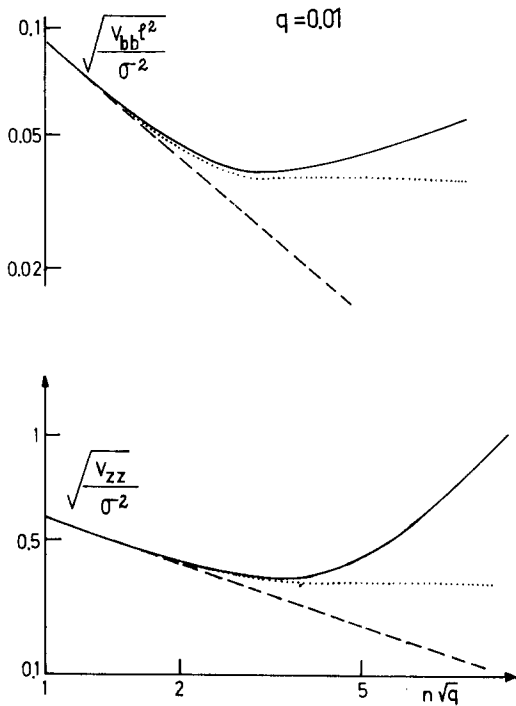
Fig. 4. Errors on the track parameters, as in fig. 2 but for: *straight track with uniformly distributed scattering.*

Fig. 5. Errors on the track parameters, as in fig. 2 but for: *curved track with uniformly distributed scattering.*

$$\sigma_Y^2 = \frac{9}{n}\sigma_h^2\left(1 + \tfrac{1}{5670}n^4 q_h^2\right)$$

$$\sigma_a^2 = \frac{192}{n}\left(\frac{\sigma_h}{L}\right)^2\left(1 + \tfrac{1}{840}n^4 q_h^2\right)$$

$$\sigma_c^2 = \frac{180}{n}\left(\frac{\sigma_h}{L^2}\right)^2\left(1 + \tfrac{1}{504}n^4 q_h^2\right)$$

$$r_h^2 = \sigma_h^2\left(1 + \tfrac{1}{2520}n^3 q_h^2\right)$$

in the horizontal plane.

The dependence of σ_z, σ_b, σ_Y, ... on n (with l fixed) is shown in figs. 4 and 5: the errors have a minimum (at $n \simeq 3.2/\sqrt{q_v}$ for σ_z and σ_b, at $n \simeq 6.8/\sqrt{q_h}$ for σ_Y and σ_a) excepted σ_c, which decreases to zero. As a conclusion no optimal length can be found to minimize the errors on all parameters simultaneously.

Here again the residuals are generally much less affected by the multiple scattering than the errors on the parameters (as a matter of fact, r^2 increases as n^3 instead of n^4 for the variances); in most cases they reflect the measurement errors only.

3. The optimal fit: recursive method

3.1. Description of the method

3.1.1. General features

The track is measured at abscissas x_1, x_2 ... x_N. Knowing the best estimators of the track parameters *at*

x_n, starting from the measurements at x_n, x_{n+1} ... x_N, we want to define the best estimators of the parameters *at* x_{n-1} starting from the measurements at x_{n-1}, x_n, ... x_N. Here we have to do with the true parameters, i.e. including the effect of the previous scatterings.

This involves:

- scattering between x_{n-1} and x_n, that is, loss of information with respect to the simple extrapolation of the parameters and of their covariance matrix.
- extra information thanks to the x_{n-1} measurement.

We suppose the distance δx_n between x_{n-1} and x_n to be short, so that the scattering affects the angles but not the position: in other words we can consider that the matter is concentrated at the measurement points, and that the track can be assimilated to a broken line; anyway, one can introduce intermediate scattering points where no measurement is made. This assumption simplifies the calculations, but it is not essential to the method.

We will first describe this (backward) recursive estimator for tracks in a plane, parametrizing each elementary segment as a portion of parabola (with the same notations as in sect. 2).

3.1.2. Accounting for the scattering

Let y_n^t, a_n^t and c_n^t be the true values of the parameters at x_n, *after* scattering; the true values at this point *before* scattering are y_n^t, $a_n^t + \alpha_n$ and c_n^t, where α_n is a quasi-Gaussian variable of variance $\delta\alpha_n^2$. Let y_n^{opt}, a_n^{opt} and c_n^{opt} be the best estimators of y_n^t, a_n^t and c_n^t in terms of y_n^m, y_{n+1}^m ... y_N^m (measured coordinates at x_n, x_{n+1} ... x_N) and V_n their covariance matrix. These estimators, as random variables, are clearly independent of α_n. Then they are also the best estimators of y_n^t, $a_n^t + \alpha$ and c_n^t, with a covariance matrix V_n^* equal to V_n where $\delta\alpha_n^2$ has been added to the V_{aa} term.

In the framework of information theory: if a_n^{opt} is an efficient estimator of a_n^t, it is also efficient for $a_n^t + \alpha_n$.

3.1.3. Propagating to the previous point

The transformation from the parameters at x_n (before scattering) to those at x_{n-1} (after scattering) is one-to-one, without information loss. Then y_{n-1}', a_{n-1}' and c_{n-1}' defined by:

$$\begin{bmatrix} y_{n-1}' \\ a_{n-1}' \\ c_{n-1}' \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & -\delta x_n & \dfrac{\delta x_n^2}{2} \\ 0 & 1 & -\delta x_n \\ 0 & 0 & 1 \end{bmatrix}}_{D_n} \begin{bmatrix} y_n^{opt} \\ a_n^{opt} \\ c_n^{opt} \end{bmatrix} \tag{1}$$

are the best estimators of y_{n-1}^t, a_{n-1}^t and c_{n-1}^t in terms of y_n^m, y_{n+1}^m, ... y_N^m. Their covariance matrix is:

$$V_{n-1}' = D_n \left(V_n + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \delta\alpha_n^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) D_n^t = D_n V_n^* D_n^t. \tag{2}$$

3.1.4. Adding the previous measurement

Now we add the information brought by y_{n-1}^m. This information concerns y_{n-1}^t, a_{n-1}^t and c_{n-1}^t, and also α_n, which is not useful for our purpose (it could perhaps be used for particle identification at low momentum).

The likelihood of the whole set of measurements (y_n^m ... y_N^m), as a function of y, a, c (parameters to be fitted at x_{n-1}) is given, up to a constant factor, by $\exp[-\frac{1}{2}\Delta p^t (V_{n-1}')^{-1}\Delta p]$ where the components of Δp are $y-y_{n-1}'$, $a - a_{n-1}'$ and $c - c_{n-1}'$. The likelihood of the y_{n-1}^m measurement is simply $\exp[-(y-y_{n-1}^m)^2/2\sigma^2]$.

Since the errors are independent, the likelihood functions can be multiplied; the maximum likelihood is

obtained by minimizing the quadratic function:

$$\Delta p^{t}(V'_{n-1})^{-1}\Delta p + \frac{(y - y^{m}_{n-1})^{2}}{\sigma^{2}} \tag{3}$$

with respect to y, a, c.

The best estimators y^{opt}_{n-1}, a^{opt}_{n-1} and c^{opt}_{n-1} are then solutions of the linear system:

$$\left((V'_{n-1})^{-1} + \begin{bmatrix} \dfrac{1}{\sigma^{2}} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right)\begin{bmatrix} y - y'_{n-1} \\ a - a'_{n-1} \\ c - c'_{n-1} \end{bmatrix} = \begin{bmatrix} \dfrac{y^{m}_{n-1} - y'_{n-1}}{\sigma^{2}} \\ 0 \\ 0 \end{bmatrix} \tag{4}$$

and their covariance matrix:

$$V_{n-1} = \left((V'_{n-1})^{-1} + \begin{bmatrix} \dfrac{1}{\sigma^{2}} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right)^{-1} \tag{5}$$

Eqs. (2) to (5) can be written with the information matrices (i.e. the inverted covariance matrices, in the Gaussian hypothesis).

Eq. (2) gives:

$$I'_{n-1} = (D^{t}_{n})^{-1}\underbrace{\left((I_{n})^{-1} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \delta\alpha^{2}_{n} & 0 \\ 0 & 0 & 0 \end{bmatrix}\right)^{-1}}_{I^{*}_{n}}(D_{n})^{-1}$$

with

$$(D_{n})^{-1} = \begin{bmatrix} 1 & \delta x_{n} & \dfrac{\delta x^{2}_{n}}{2} \\ 0 & 1 & \delta x_{n} \\ 0 & 0 & 1 \end{bmatrix}$$

Moreover:

$$I^{*}_{n} = \left((I_{n})^{-1} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \delta\alpha^{2}_{n} & 0 \\ 0 & 0 & 0 \end{bmatrix}\right)^{-1} = I_{n} - \frac{\delta\alpha^{2}_{n}}{1 + \delta\alpha^{2}_{n}(I_{n})_{aa}}I_{n}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}I_{n}$$

$$= I_{n} - \underbrace{\frac{\delta\alpha^{2}_{n}}{1 + \delta\alpha^{2}_{n}(I_{n})_{aa}}\begin{bmatrix} (I_{n})^{2}_{Ya} & (I_{n})_{Ya}(I_{n})_{aa} & (I_{n})_{Ya}(I_{n})_{Yc} \\ \cdots & (I_{n})^{2}_{aa} & (I_{n})_{aa}(I_{n})_{ac} \\ \cdots & \cdots & (I_{n})^{2}_{ac} \end{bmatrix}}_{\text{information loss}} \tag{6}$$

Eq. (4) becomes:

$$\left(\underbrace{(D^{t}_{n})^{-1}I_{n}(D_{n})^{-1}}_{I'_{n-1}} + \begin{bmatrix} \dfrac{1}{\sigma^{2}} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right)\begin{bmatrix} y - y'_{n-1} \\ a - a'_{n-1} \\ c - c'_{n-1} \end{bmatrix} = \begin{bmatrix} \dfrac{y^{m}_{n-1} - y'_{n-1}}{\sigma^{2}} \\ 0 \\ 0 \end{bmatrix} \tag{7}$$

and eq. (5) gives:

$$
I_{n-1} = I'_{n-1} + \underbrace{\begin{bmatrix} \dfrac{1}{\sigma^2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\text{information brought by the measurement}}.
\tag{8}
$$

This formulation avoids the explicit matrix inversions involved in eqs. (4) and (5).
Table 1 summarizes the recursive algorithm.

3.1.5. Starting the recursion
This can be done by two practically equivalent means:
– Using the last three points to calculate y, a, c at x_{N-2}, and their covariance matrix (including the scattering contribution at x_{N-2} and x_{N-1} if necessary).
– Starting the recursion at the last point, which defines y_N with variance σ^2, but leaves a and c fully undetermined: one can assign to them approximate initial values and very large arbitrary variances: then the recursive algorithm can be applied without risk of matrix singularity: the information brought by y_{N-2}^m fits a and c at x_{N-2}, and the fitted values are nearly independent of the initial ones.

3.2. Generalizations

We consider now a track in 3-dimensional space with an inhomogeneous magnetic field. This is defined by 5 parameters (e.g. y, z, a, b and $1/p$ at fixed x). Moreover the measured coordinates are not necessarily y or z at fixed x.

Table 1
One step of the recursive estimator

	Estimator	Information matrix
	at x_n: $y_n^{\text{opt}} a_n^{\text{opt}} c_n^{\text{opt}}$	I_n
scattering at x_n	↓	↓
	at x_n: $y_n^{\text{opt}} a_n^{\text{opt}} c_n^{\text{opt}}$	$I_n^* = (I_n^{-1} + A_n)^{-1}$
propagation from x_n to x_{n-1}	↓	↓
	at x_{n-1}: $y_{n-1}^* a_{n-1}^* c_{n-1}^*$ given by: $\begin{bmatrix} y_{n-1}^* \\ a_{n-1}^* \\ c_{n-1}^* \end{bmatrix} = D_n \begin{bmatrix} y_n^{opt} \\ a_n^{opt} \\ c_n^{opt} \end{bmatrix}$	$I_{n-1}^* = (D_n^t)^{-1} I_n^* (D_n)^{-1}$
measurement at x_{n-1}	↓	↓
	at x_{n-1}: $y_{n-1}^{\text{opt}} a_{n-1}^{\text{opt}} c_{n-1}^{\text{opt}}$ given by the equation: $(I_{n-1}^* + M_n)\begin{bmatrix} y_{n-1}^{opt} - y_{n-1}^* \\ a_{n-1}^{opt} - a_{n-1}^* \\ c_{n-1}^{opt} - c_{n-1}^* \end{bmatrix} = M_n \begin{bmatrix} y_{n-1}^m - y_{n-1}^* \\ 0 \\ 0 \end{bmatrix}$	$I_{n-1} = I_{n-1}^* + M_n$

In the parabolical schematization:

$$
D_n = \begin{bmatrix} 1 & -\delta x_n & \delta x_n^2/2 \\ 0 & 1 & -\delta x_n \\ 0 & 0 & 1 \end{bmatrix} \quad A_n = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \delta\alpha^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad M_n = \begin{bmatrix} 1/\sigma^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.
$$

We describe hereafter the 3 stages of each step of the recursion. Instead of the azimuth ϕ and the dip λ, we continue to us the slope parameters $a = p_y/p_x = \tan \phi$ and $b = p_z/p_x = \tan \lambda/\cos \phi$: this avoids trigonometrical calculations. We use also: $d = Ze/p$ ($B \cdot d$ is then the curvature for a track perpendicular to the magnetic field).

3.2.1. scattering at x_n

This affects only λ and ϕ (or a and b). Let ξ_n be the number of radiation lengths X_0 crossed from x_{n-1} to x_n. For a homogeneous medium:

$$\xi_n = \frac{x_n - x_{n-1}}{X_0 \cos \lambda \cos \phi} = \frac{\delta x_n}{X_0} \sqrt{1 + a^2 + b^2} \,. \tag{9}$$

The scattering angles α and β in two perpendicular planes containing the track direction are independent random variables of variance $\delta \alpha_n^2 = (KZ/pv)^2 \xi_n$. If one of these planes is vertical, the scattering angle in this plane represents exactly the variation of λ and the other one of the variation of ϕ multiplied by cos λ; with the parameters ϕ and λ, to account for the scattering consists in adding $\delta \alpha_n^2/\cos^2\lambda$ to $V_{\phi\phi}$ and $\delta \alpha_n^2$ to $V_{\lambda\lambda}$.

With the parameters a and b, one finds easily, that one has to add:

$$
\begin{aligned}
(1 + a^2 + b^2)(1 + a^2)\delta\alpha_n^2 \quad &\text{to} \quad V_{aa} \\
(1 + a^2 + b^2)(1 + b^2)\,\delta\alpha_n^2 \quad &\text{to} \quad V_{bb} \\
(1 + a^2 + b^2)ab \qquad \delta\alpha_n^2 \quad &\text{to} \quad V_{ab}.
\end{aligned}
\tag{10}
$$

In terms of information matrices, we get:

$$I_n^* = \left(V_n^*\right)^{-1} = \left((I_n)^{-1} + \delta\alpha_n^2 \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{\mathbf{A}} & 0 \\ 0 & 0 & & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right)^{-1}, \tag{11}$$

where

$$\mathbf{A} = (1 + a^2 + b^2) \begin{bmatrix} 1 + a^2 & ab \\ ab & 1 + b^2 \end{bmatrix}.$$

The matrix inversions can be simplified in the same way as in sect. 3.1.4.:

$$I_n^* = I_n - \delta\alpha_n^2 I_n \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{\mathbf{a}} & 0 \\ 0 & 0 & & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} I_n, \tag{12}$$

with

$$\mathbf{a} = \left(A^{-1} + \delta\alpha_n^2 \begin{bmatrix} I_{aa} & I_{ab} \\ I_{ab} & I_{bb} \end{bmatrix} \right)^{-1}.$$

3.2.2. Propagation from x_n to x_{n-1}:

The track extrapolation from a point to the previous one must be accurate: the cumulated error over the

whole length of the track should remain small with respect to the position uncertainty resulting from the fit. Any suitable method for this extrapolation allows to calculate also the (5×5) matrix of derivatives analogous to D_n in eq. (1), either by analytical differentiation, or by finite differences.

As a matter of fact, on the left-hand side of the equation which should generalize eq. (4), the matrix V'_{n-1} can be calculated with an approximation for D_n, assuming the field B to be constant along the track between x_{n-1} and x_n. If the curvature is not too strong, we obtain for the parameters y, z, a, b and d:

$$D_n \simeq \begin{bmatrix} 1 & 0 & -\delta x_n + F_1 \delta x_n^2/2 & F_2 \delta x_n^2/2 & F_3 \delta x_n^2/2 \\ 0 & 1 & G_1 \delta x_n^2/2 & -\delta x_n + G_2 \delta x_n^2/2 & G_3 \delta x_n^2/2 \\ 0 & 0 & 1 - F_1 \delta x_n & -F_2 \delta x_n & -F_3 \delta x_n \\ 0 & 0 & -G_1 \delta x_n & 1 - G_2 \delta x_n & -G_3 \delta x_n \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{13}$$

with

$$F_3 = e\left[bB_x + abB_y - (1 + a^2)B_z\right]; \qquad F_1 = d\left[aF_3/e^2 + e(bB_y - 2aB_z)\right]$$

$$F_2 = d\left[bF_3/e^2 + e(B_x + aB_y)\right]$$

$$G_3 = e\left[-aB_x + (1 + b^2)B_y - abB_z\right]; \qquad G_1 = d\left[aG_3/e^2 - e(B_x - bB_z)\right]$$

$$G_2 = d\left[bG_3/e^2 + e(2bB_y - aB_z)\right],$$

where $e^2 = 1 + a^2 + b^2$.

3.2.3. Addition of a measurement

In many cases each measurement consists in determining y or z, or a combination $t = \lambda y + \mu z$, at a given fixed x. Here it is useless to build space points from several raw measurements: such a procedure implies sometimes a loss of information.

The likehood of measurement t_{n-1}^m is now, in the Gaussian approximation:

$$\exp\left(-\frac{1}{2}\frac{(\lambda y + \mu z - t_{n-1}^m)^2}{\sigma_n^2}\right)$$

and the linear system (4) becomes:

$$\left((V'_{n-1})^{-1} + \frac{1}{\sigma^2}\begin{bmatrix} \lambda^2 & \lambda\mu & 0 & 0 & 0 \\ \lambda\mu & \mu^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}\right)\begin{bmatrix} y - y'_{n-1} \\ z - z'_{n-1} \\ a - a'_{n-1} \\ b - b'_{n-1} \\ d - d'_{n-1} \end{bmatrix} = \frac{1}{\sigma^2}\begin{bmatrix} \lambda(t_{n-1}^m - \lambda y'_{n-1} - \mu z'_{n-1}) \\ \mu(t_{n-1}^m - \lambda y'_{n-1} - \mu z'_{n-1}) \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{14}$$

Of course, if y and z (or two independent combinations of y and z) are available at the same x, both measurements can be included together in one equation like (14).

If the information refers to a non-linear combination of y and z, it can generally be linearized around the measured values.

When the measured quantity depends only on x, especially in the case of optical measurements, one can transform the raw information into information at fixed x; this transformation is a projection along the direction of the track. Several examples are described in detail in ref. [4].

3.3. Extension to energy loss

The energy loss can be taken into account in natural way: it is enough, in the progagation, in the propagation stage 3.2.2., to add a suitable quantity to the *d* parameter. Moreover, it is possible to introduce a (Gaussian) uncertainty on this energy loss, as in 3.2.1., by adding a contribution to V_{dd} in the same time as V_{aa}, V_{bb} and V_{ab} at the scattering stage.

4. Comparison between the optimal fit and the standard fit

4.1. Introduction

Our aim is to determine the gain in precision on the parameters, brought by the optimal estimators, with the same model as in 2.2. and 2.3., and the same notations.

In the optimal fit, the variances of the parameters cannot increase as measurement points are added (in other terms, addition of measurement cannot cause a loss of information). Thus, for $n \to \infty$ they must tend toward some limits (possibly zero). Again we use the Gaussian approximation: so the covariance matrix V of the fitted parameters is the inverse of the information matrix I.

4.2. Scattering at one point

The expressions of the measured variables show that the measurements after L_0 are independent of those before L_0, so that the total information matrix on the parameters (Z and b) at the origin is the sum of the contributions brought by the two portions of track delimited by the scatterer

For the second portion (after L_0), we calculate first the information on the parameters at $x = L_0$ (including the effect of the scattering uncertainty), and then we propagate it back to $x = 0$. At $x = 0$ we add the information of the first portion (before L_0), where no scattering uncertainty occurs.

This calculation gives the covariance matrix of the parameters, hence their errors as represented in fig. 2a and b (dotted curves). These curves show that the gain in precision of the optimal estimator is large when ρ is large and/or when $n \geq 3n_0$ (however at very large values of n the gain vanishes for the curvature).

Let us note that in this case (scattering at only one point) the optimal fit can be reduced to two standard fits (one for each track portion, before and after L_0) connected together by the three elementary operations defined in sect. 3: accounting for the scattering at $x = L_0$ for the second portion, propagating the information matrix of this portion back to the first point, and adding the information of both portions. An analogous strategy could be adopted with few scattering points, provided that each portion of track between two scatterers has a sufficient number of points to perform a standard fit: the information brought by each portion could then be merged recursively as described in sect. 3.

4.3. Uniformly distributed scattering

The formalism developed in sect. 3 gives a recursive expression of the optimal information matrix: if the track has $n + 1$ points, the last n points give an information matrix I_n for the parameters at the second point; then the first measurement can be added within one step of the recursion: so we obtain I_{n+1} for the parameters at the first point.

For Z and b parameters (straight line fit in the vertical plane) this gives:

$$I_{n+1} = \begin{bmatrix} 1 & 0 \\ l & 1 \end{bmatrix} \left(I_n^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \delta\alpha^2 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & l \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} \sigma_0^{-2} & 0 \\ 0 & 0 \end{bmatrix}.$$

For Y, a and c (parabolic fit in the horizontal plane):

$$I_{n+1} = \begin{bmatrix} 1 & 0 & 0 \\ l & 1 & 0 \\ l^2/2 & l & 1 \end{bmatrix} \left(I_n^{-1} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \delta\alpha^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & l & l^2/2 \\ 0 & 1 & l \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} \sigma_h^{-2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

From these formulae we can deduce an expansion of I_n in powers of $\delta\alpha^2$, valid for small q (let us recall $q = l\delta\alpha/\sigma$) up to $n = 1/\sqrt{q}$. As a result, the optimal covariance matrix, at first order in $\delta\alpha^2$, is identical to the standard one, found in sect. 2. thus the *standard* fit is *quasi-optimal* for $n \leq 1/\sqrt{q}$ (as can be seen in figs. 4 and 5) and it is useless to search for others estimators (e.g. giving to the measurements weights decreasing with n).

The recursion relations enable us to calculate also the limits of the variances for $n \to \infty$. they can be expanded in powers of $\sqrt{q}$:

$$(\sigma_Z^\infty)^2 = \sqrt{2\sigma_v^3/\delta\alpha} \left(1 - \sqrt{\frac{q_v}{2}} - \tfrac{5}{8} q_v + \ldots \right).$$

$$(\sigma_b^\infty)^2 = \sqrt{\frac{2\sigma_v\delta\alpha^2}{l}} \left(1 - \sqrt{\frac{q_v}{2}} + \tfrac{1}{8} q_v + \ldots \right).$$

$(\sigma_Y^\infty)^2$ and $(\sigma_a^\infty)^2$ have the same expressions as $(\sigma_Z^\infty)^2$ and $(\sigma_b^\infty)^2$ where σ_v is replaced by σ_h, and q_v by q_h.

Figs. 4 and 5 show the comparison between the standard uncertainties and the optimal ones.

5. Implementation of the recursive method

5.1. Description of the experiment

We used real events from the CERN WA13 experiment, which was realized in the Omega spectrometer equipped with optical spark chambers; its aim was to study hadronic large-angle reactions giving in the final state two charged particles or two neutrals decaying in charged mode; detailed description of the experiment can be found in refs. [3] and [4]. Fig. 6 summarizes the measurement apparatus: in their measured section, the tracks crossed mainly the spark chambers (0.06 L_{rad} for the whole set of chambers) and a 1 cm thick scintillator S6.

We used events produced at 3 GeV/c, and we did not apply kinematical cuts: the momenta of the reconstructed tracks ranged from a few hundred MeV/c; the outgoing particles were not identified.

5.2. Programmation of the algorithm

The recursive method was programmed in the following way:

− All matrix operations were written explicitly, without subscripted variables or loops, in order to save computation time (this is possible because all matrices involved have the same dimension).

− The propagation between two points (generally separated by 6 to 10 cm) was realized in two steps by a fourth-order Taylor expansion of the trajectory; the propagation of the information matrix used the second order approximation (cf. (3.13)). The magnetic field was assumed to be uniform between these two points. In most cases this calculation was much too accurate; it was needed for low-momentum particles emitted at large angle.

Every time the distance between two measured points exceeded 10 cm, dummy points were inserted.

− The variances of the scattering angles were calculated in two velocity hypothesis: ultrarelativistic (valid for π) and proton.

− The measured points were extracted with their weights defined at fixed x, from ROMEO (standard

reconstruction program for Omega events), just before the final fit, so that both fitting procedures worked on exactly the same input.

– The starting value of the curvature parameter was estimated with three points (at the beginning, in the middle and at the end of the track); for the slope parameters, the last two points were used, with a curvature correction. As a matter of fact, the fitted values do not appear to depend strongly on the initial ones. After 3 or 4 recursion steps, the gaps between the extrapolated trajectory and the measured points are compatible with the measurements errors, and moderate deviations on the starting values are automatically corrected; for example, estimating the slope without curvature correction between the last two points has no appreciable effect on the fitted values.

The computation time for one recursion step (propagation + measurement + scattering) amounted to about 360 μs on CDC CYBER 750, equivalent to 160 μs on CDC 7600 (the computation of the magnetic field is not included). The time for a whole track is obviously proportional to the number of points on this track.

5.3. Test of the precision

From the parameters fitted at the first measured point we extrapolated the tracks backwards to find intersections, and we determined the points of closest approach between each pair of tracks with opposite sign.

Fig. 7 shows the distributions of the minimal distances obtained with the standard program ROMEO, and with our algorithm (for both mass hypotheses).
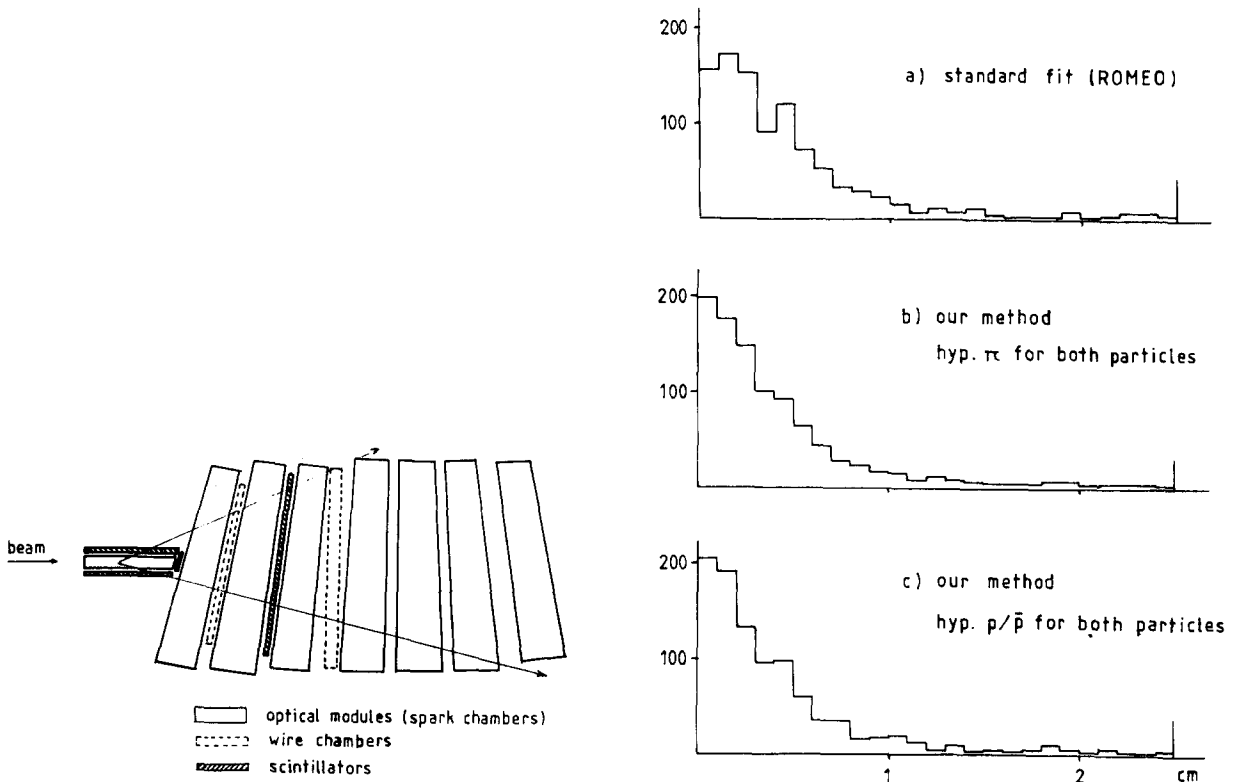

Fig. 6. Layout of the WA13 experiment in the Omega Spectrometer.

Fig. 7. Comparison between the standard fitting method of ROMEO and the recursive implementation of the optimal estimator: distance, at the point of closest approach, between two tracks extrapolated backwards to the vertex.

The precision is clearly improved by our method. The improvement is not very strong, but anyway the theoretical gain that we could deduce from sect. 4 for this experiment, is not enormous. Moreover there was matter accumulated around the vertex (the hydrogen target was surrounded by scintillators) which worsened its geometrical precision, and for which optimization cannot help. For the same reason, no appreciable improvement can be expected in the kinematical resolution.

6. Conclusion

The standard fitting procedures, applied to slow tracks measured over a long length, and/or with high accuracy, lead to important losses of precision on the geometrical parameters. They can be improved by taking account of the measurements up to a certain length only; this optimal length can be evaluated from analytical approximations of the variances. In many cases, especially in a homogeneous medium, this "truncated" fit is not far from being optimal. Its main drawback is that the optimal length depends on the choice of parameter for which one wants the highest precision; moreover it was calculated assuming Gaussian independent measurement errors, and we do not know how it would be modified by small uncorrected geometrical distorsions, or by non-Gaussian errors (e.g. in a MWPC).

The optimal fit (without information loss) is usually realized by calculating the whole covariance matrix of the measurements, or by adding extra parameters to describe the scattering [1]; these methods involve handling big matrices. In this paper we described a new implementation of the optimal estimator: the parameters are fitted backwards by introducing the measured points one at a time from the end to the beginning of the track. The elementary steps are thus: including one or several raw measurements, and/or one elementary scattering: from one point to the next one the parameters and their information (or weight) matrix are propagated by a local polynomial parametrization.

Such a procedure is very flexible: it can be applied to various types of detectors, and especially to composite detectors. Moreover it does not require more computation time than the standard ones (perhaps less); contrary to the "big matrix" optimal methods, the time used is merely proportional to the number of points.

The implementation of this algorithm on real events shows an improvement of the geometrical accuracy; its order of magnitude corresponds to what could be expected. The fitted values of the parameters appear to be stable with respect to the starting value chosen at the end of the track: thus no iteration is needed to obtain the best estimate; the algorithm stabilizes after 3 or 4 steps along the track.

Finally we point out that this method, in some configurations, could be able to perform at the same time the pattern recognition and the geometrical reconstruction of the tracks, since it gives in a short time the best extrapolation of a track candidate built up with the most external points.

References

[1] H. Eichinger and M. Regler, CERN 81-06, Review of track-fitting methods in counter experiments.
[2] D. Drijard, CERN/D.Ph. II/Prog. 69-4, Handling of multiple scattering in track reconstruction.
[3] R. Schwarz, Thesis, Neuchâtel (1982).
[4] P. Billoir, Thesis, Paris VI (1983).