HARP Collaboration

HARP Memo 04-002
10 March 2004
http://cern.ch/dydak/GeneralizedLSF.ps

TPC Track Reconstruction: Generalized Least Squares Fit

F. Dydak and Yu. Nefedov

CERN-HARP-CDP-2004-001
10/03/2004

Abstract

Track fitting in the HARP TPC must account for the fact that, because of the rotational symmetry of the TPC, both x and y cluster coordinates have errors at the same level of importance. Conventional fit algorithms which have only one coordinate with error while all other coordinates are error-free, are not appropriate. A generalized least-squares method is described which is symmetric in all coordinates, and applied in a 3D helix fit of TPC tracks.

1 Introduction

Fitting of clusters in the HARP TPC must cope with the fact that, because of the rotational TPC symmetry, the measurement errors of the transverse cluster coordinates x and y are at the same level of importance. By contrast, conventional fit algorithms have only one coordinate with error (referred to as 'dependent' variable) while all other coordinates (referred to as 'independent' variables) are error-free.

Consider e.g. one straight track in the horizontal direction and another one in the vertical direction. Suppose the error of the y coordinate is used in the fit, while the error in the x coordinate is not taken into account by the algorithm. The fit will do pretty well with the horizontal track and return a reasonable χ^2. However, the vertical track will *de facto* be fitted with a constant error of the x coordinate and will return a meaningless χ^2.

This example suffices to conclude that a fit algorithm with error handling in one coordinate only is inadequate.

Therefore, the concept of 'Generalized Least Squares Fit' (called GLSF below) is introduced and applied in a 3D fit of TPC coordinates. The GLSF is by construction symmetric in all coordinates.

The problem addressed by the GLSF is not new. The earliest trace in the literature which is known to the authors are discussions by Deming [1] in 1931 and 1934. The issue was rediscovered by Cohen [2] in 1953, and by Brown [3] in 1955. The case of a straight-line fit was discussed by York [4] in 1966. His discussion was taken up and expanded by Gerhold [5] in 1969. A comprehensive discussion including the geometrical interpretation of the problem, and the application to a physics case where the procedure proved mandatory for obtaining a bias-free high-precision result, was given by Dydak [6] in 1972.

Nevertheless, the problem and its solution is still not widely known, as demonstrated e.g. by Orear [7] in 1981 and Moreno [8] in 1996, who both discussed the special case of two variables with errors. Texts on Statistics and Data Analysis tend to ignore the issue. A notable exception is Brandt's book of 1992 on data analysis [9].

Section 2 presents the generic algorithm of a GLSF for an arbitrary number of coordinates with correlated errors, and with an arbitrary number of non-linear boundary conditions. Section 3 discusses errors of fit parameters and their correlations, and the χ^2 of the fit. Section 4 presents the geometric interpretation of the algorithm. Section 5 illustrates the simple but frequently occuring case of a straight-line fit in two coordinates with uncorrelated errors. Section 6 presents the algorithm which has been adopted for the 3D-fit of tracks in the HARP TPC.

2 Generic algorithm

Physical quantities can be measured either **directly** (e.g. the measurement of a temperature with a thermometer) or **indirectly** (e.g. the momentum of a track from points along its trajectory). Here, we consider the latter case.

We assume that sufficient measurements are made such that the equations for the wanted physics parameters are over-determined. We ask for that set of physical parameters which has the largest statistical weight or, equivalently, the smallest statistical errors. That optimization is achieved through the requirement of 'maximum likelihood', which coincides for Gaussian errors with the 'least squares' requirement.

The nomenclature used throughout this paper is as follows:

- be i the index of a measurement point in measurement space, with $1 \leq i \leq N$; for example, i is the index of a cluster along a TPC track;
- be s the coordinate index of a measurement point, with $1 \leq s \leq n$; be

$$\vec{X}^i = \begin{pmatrix} X_1^i \\ \cdots \\ X_s^i \\ \cdots \\ X_n^i \end{pmatrix} \ , \ \vec{\bar{x}}^i = \begin{pmatrix} \bar{x}_1^i \\ \cdots \\ \bar{x}_s^i \\ \cdots \\ \bar{x}_n^i \end{pmatrix} \ and \ \vec{x}^i = \begin{pmatrix} x_1^i \\ \cdots \\ x_s^i \\ \cdots \\ x_n^i \end{pmatrix} \tag{1}$$

  the n-component vector of **measured** coordinates of the i.th measurement point, the n-component vector of **approximate** coordinates used in successive iteration steps, and the n-component vector of **best-fit** coordinates of the i.th measurement point; this latter vector is to represent an **unbiased** estimate of the n-component vector of **true** coordinates; generally, capital letters denote measured coordinates while small letters denote approximate or fitted coordinates; for example, x_s^i with $1 \leq s \leq 3$ denote the x, y^i and z best-fit coordinates of the i.th TPC cluster.

- be r the index of a fit parameter, with $1 \leq r \leq k$; be

$$\vec{\alpha} = \begin{pmatrix} \alpha_1 \\ \cdots \\ \alpha_r \\ \cdots \\ \alpha_k \end{pmatrix} \tag{2}$$

  the k-component vector of fit parameters; for example, α_r with $1 \leq r \leq 5$ is one of five parameters of a helical TPC track;

- be l the index of a boundary condition, with $1 \leq l \leq m$, which the best-fit coordinates and the best-fit parameters must satisfy:

$$F_l(\vec{x}^i, \vec{\alpha}) = 0 \ , \tag{3}$$

  or in vectorial form

$$\vec{F}^i = 0 \ , \tag{4}$$

  where we introduced the symbolic m-component vector

$$\vec{F}^i = \begin{pmatrix} F_1(\vec{x}^i, \vec{\alpha}) \\ \cdots \\ F_l(\vec{x}^i, \vec{\alpha}) \\ \cdots \\ F_m(\vec{x}^i, \vec{\alpha}) \end{pmatrix} \ ; \tag{5}$$

for example, a TPC track must satisfy three boundary conditions which together describe a helix in 3D space.

A boundary condition or partial derivative thereof which is not evaluated at the best-fit coordinate positions but at approximate coordinate positions, is labelled by a bar (e.g. $\bar{F}$). An evaluation at the measured coordinate positions is labelled by a hat (e.g. $\hat{F}$).

With no measurement errors, k equations or, equivalently, k vectors $\vec{X}^i$, are sufficient to determine the k parameters $\alpha_1, \alpha_2, ...\alpha_k$. In practice, one has $N \gg k$, and the optimum set of parameters must be obtained from an over-determined system of equations.

We start with the *ansatz* that the probability density of each coordinate measurement X_s^i is a Gaussian around the true location:

$$p(X_s^i) = \frac{1}{\sigma_{X_s^i}\sqrt{2\pi}} \exp\left[-\frac{(X_s^i - x_s^i)^2}{2\sigma_{X_s^i}^2}\right] . \tag{6}$$

In many cases, the coordinate measurement errors will be correlated. Accordingly, we make the *ansatz* that the probability density of the vector $\vec{X}^i$ of coordinate measurements is a multivariate Gaussian distribution around the true locations. With the $n \times n$ covariance matrix of the measured coordinates X_s^i,

$$\mathbb{V}^i = \begin{pmatrix} \mathrm{var}(X_1^i) & cov(X_1^i, X_2^i) & \cdot & \cdot & \cdot & cov(X_1^i, X_n^i) \\ cov(X_2^i, X_1^i) & \mathrm{var}(X_2^i) & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ cov(X_n^i, X_1^i) & cov(X_n^i, X_2^i) & \cdot & \cdot & \cdot & \mathrm{var}(X_n^i) \end{pmatrix} , \tag{7}$$

the probability density of the vector $\vec{X}^i$ reads as

$$p(\vec{X}^i) = \frac{1}{(2\pi)^{n/2}\sqrt{|\mathbb{V}^i|}} \exp\left[-\frac{1}{2}(\vec{X}^i - \vec{x}^i)^{\mathrm{T}}(\mathbb{V}^i)^{-1}(\vec{X}^i - \vec{x}^i)\right] . \tag{8}$$

The inclusion of the covariance matrix ensures that not only measurement errors of individual coordinates but also their correlations are taken into account.

Taking error correlation into account is particularly important for TPC coordinates: while the primary transverse cluster coordinates r and $r \cdot \phi$ are not correlated, the derived coordinates x and y are. At the same time, this example highlights that the correlation matrix will in general depend on the index i of the measurement point.

In the case of uncorrelated measurements, where the covariance matrix takes the diagonal form

$$\mathbb{V}^i = \begin{pmatrix} (\sigma_{x_1}^i)^2 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & (\sigma_{x_2}^i)^2 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & (\sigma_{x_n}^i)^2 \end{pmatrix} , \tag{9}$$

Eq. [8] reduces to

$$p(\vec{X^i}) = \frac{1}{(2\pi)^{n/2}\sigma^i_{x_1}\sigma^i_{x_2}\cdots\sigma^i_{x_n}} \exp\left[-\sum_{s=1}^{n}\frac{(X^i_s - x^i_s)^2}{2(\sigma^i_{x_s})^2}\right] . \tag{10}$$

It is stressed that in all considerations of error correlations, only correlations between coordinate errors within the **same** measurement point are considered. By contrast, coordinate errors of **different** measurement points are considered uncorrelated throughout. We will make use of this feature in the discussions of Section 3.

The maximum likelihood requirement demands that the 'best-fit' coordinate vector $\vec{x^i}$ is to be determined such that the probability density Eq. [8] is maximal. This requires that

$$\chi^2 = \sum_{i=1}^{N}(\vec{X}^i - \vec{x}^i)^{\mathrm{T}}(\mathbb{V}^i)^{-1}(\vec{X}^i - \vec{x}^i) \tag{11}$$

takes its minimum value. In case of uncorrelated errors, this reduces to the familiar 'least squares' requirement, where

$$\chi^2 = \sum_{i=1}^{N}\sum_{s=1}^{n} p^i_s(X^i_s - x^i_s)^2 \tag{12}$$

takes a minimum; here, the 'weight' p^i_s of the i.th measurement of the x_s coordinate is defined as

$$p^i_s = \frac{1}{(\sigma^i_{x_s})^2} . \tag{13}$$

At this point the difference of the GLSF to the Standard Least Squares Fit becomes apparent: the GLSF has the 'best-fit' coordinates x^i_s explicitly as **additional** unknowns to be determined, whereas in the Standard Least Squares Fit the parameters α_r are the only unknowns. For the 'best-fit' coordinate vector $\vec{x^i}$ to satisfy the $N \times m$ boundary conditions Eqs. [3], the traditional method of solving an optimization problem with boundary conditions is employed: introducing Lagrange multipliers as further unknowns.

The variation of χ^2 Eq. [11] reads with the n-component vector

$$\delta\vec{x^i} = \begin{pmatrix} \delta x^i_1 \\ \cdots \\ \delta x^i_s \\ \cdots \\ \delta x^i_n \end{pmatrix} \tag{14}$$

as follows:

$$\delta\chi^2 = -2\sum_{i=1}^{N}(\delta\vec{x^i})^{\mathrm{T}}(\mathbb{V}^i)^{-1}(\vec{X}^i - \vec{x}^i) = 0 . \tag{15}$$

Without loss of generality, we replace the constant factor -2 by 1.

The variation of the boundary conditions Eqs. [3] is

$$\delta F_l(\vec{x^i}, \vec{\alpha}) = \sum_{s=1}^{n}\frac{\partial F_l(\vec{x^i}, \vec{\alpha})}{\partial x_s}\delta x^i_s + \sum_{r=1}^{k}\frac{\partial F_l(\vec{x^i}, \vec{\alpha})}{\partial \alpha_r}\delta\alpha_r = 0 , \tag{16}$$

or, in the vector notation of Eqs. [4],

$$\delta\vec{F}^i = (\partial_x \mathbb{F}^i)\delta\vec{x}^i + (\partial_\alpha \mathbb{F}^i)\delta\vec{\alpha} = 0 \; , \tag{17}$$

where we introduced the m-component vector

$$\delta\vec{F}^i = \begin{pmatrix} \delta F_1(\vec{x}^i, \vec{\alpha}) \\ \cdots \\ \delta F_l(\vec{x}^i, \vec{\alpha}) \\ \cdots \\ \delta F_m(\vec{x}^i, \vec{\alpha}) \end{pmatrix} \; , \tag{18}$$

the k-component vector

$$\delta\vec{\alpha} = \begin{pmatrix} \delta\alpha_1 \\ \cdots \\ \delta\alpha_r \\ \cdots \\ \delta\alpha_k \end{pmatrix} \; , \tag{19}$$

the $m \times n$ matrix

$$\partial_x \mathbb{F}^i = \begin{pmatrix} \frac{\partial F_1^i}{\partial x_1} & \cdot & \cdot & \cdot & \frac{\partial F_1^i}{\partial x_n} \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \frac{\partial F_m^i}{\partial x_1} & \cdot & \cdot & \cdot & \frac{\partial F_m^i}{\partial x_n} \end{pmatrix} \tag{20}$$

and the $m \times k$ matrix

$$\partial_\alpha \mathbb{F}^i = \begin{pmatrix} \frac{\partial F_1^i}{\partial \alpha_1} & \cdot & \cdot & \cdot & \frac{\partial F_1^i}{\partial \alpha_k} \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \frac{\partial F_m^i}{\partial \alpha_1} & \cdot & \cdot & \cdot & \frac{\partial F_m^i}{\partial \alpha_k} \end{pmatrix} \; . \tag{21}$$

We multiply each of Eqs. [17] with a Lagrange multiplier λ_l^i and sum all $N \times m$ equations with Eq. [15]. We order according to the variables' variations and set the respective coefficients to zero. This leads to the following $N \times n$ equations

$$(\mathbb{V}^i)^{-1}(\vec{X}^i - \vec{x}^i) + (\partial_x \mathbb{F}^i)^{\mathrm{T}} \vec{\lambda}^i = 0 \; , \tag{22}$$

where we introduced the m-component vector

$$\vec{\lambda}^i = \begin{pmatrix} \lambda_1^i \\ \cdots \\ \lambda_l^i \\ \cdots \\ \lambda_m^i \end{pmatrix} \; ; \tag{23}$$

and the k equations

$$\sum_{i=1}^{N} (\partial_\alpha \mathbb{F}^i)^{\mathrm{T}} \vec{\lambda}^i = 0 \; . \tag{24}$$

Together with the $N \times m$ boundary conditions $\vec{F}^i = 0$ (Eqs. [4]), we have for the $N \times n$ unknowns x_s^i, plus $N \times m$ unknowns λ_l^i, plus k unknowns α_r, together $N(m+n)+k$ unknowns, the same number of equations. Therefore, a solution of the above system of equations is in principle possible, and in case of linear boundary conditions the solution will be unique.

The solution of the above system of equations proceeds in four steps:

1. we express the x_s^i in terms of λ_l^i and α_r in the $N \times n$ Eqs. [22];
2. we insert the x_s^i into the $N \times m$ boundary conditions Eqs. [4]; this operation yields $N \times m$ equations with the only variables λ_l^i and α_r;
3. we then express the λ_l^i in these $N \times m$ euqations in terms of the k parameters α_r;
4. we finally insert the λ_l^i into Eqs. [24]; this operation yields a system of k equations for the wanted k parameters α_r.

The first step gives the result

$$\vec{x}^i = \vec{X}^i + \mathbb{V}^i (\partial_x \mathbb{F}^i)^{\mathrm{T}} \vec{\lambda}^i \; , \tag{25}$$

which reduces for uncorrelated errors to

$$x_s^i = X_s^i + \sum_{l=1}^{m} \lambda_l^i \frac{1}{p_s^i} \frac{\partial F_l(\vec{x}^i, \vec{\alpha})}{\partial x_s} \; . \tag{26}$$

The second step meets an inherent difficulty: it can only be made in explicit analytic form after linearizing the boundary conditions Eqs. [3, 4] in the form of a Taylor expansion around good guesses of the unknown best-fit coordinates and parameters. This means of course that the best-fit results cannot be obtained by just solving the system of equations discussed above, but through convergence after several iterations.

Fortunately, good initial guesses for the best-fit coordinates $\vec{x}^i$ exist: the measured coordinates $\vec{X}^i$.

The linearization of the boundary conditions Eqs. [3, 4]] by Taylor expansion gives

$$\begin{aligned} F_l(\vec{x}^i, \vec{\alpha}) \;\; &\cong \;\; F_l(\vec{\bar{x}}^i, \vec{\bar{\alpha}}) \\ &+ \;\; \sum_{s=1}^{n} \frac{\partial F_l(\vec{\bar{x}}^i, \vec{\bar{\alpha}})}{\partial x_s} (x_s^i - \bar{x}_s^i) + \sum_{r=1}^{k} \frac{\partial F_l(\vec{\bar{x}}^i, \vec{\bar{\alpha}})}{\partial \alpha_r} (\alpha_r - \bar{\alpha}_r) = 0 \; , \end{aligned} \tag{27}$$

or, equivalently, in vectorial form

$$\vec{F}^i \cong \vec{\bar{F}}^i + (\partial_x \bar{\mathbb{F}}^i)(\vec{x}^i - \vec{\bar{x}}^i) + (\partial_\alpha \bar{\mathbb{F}}^i)(\vec{\alpha} - \vec{\bar{\alpha}}) = 0 \; . \tag{28}$$

The variation of Eqs. [27, 28] gives

$$\delta F_l(\vec{x}^i, \vec{\alpha}) \cong \sum_{s=1}^{n} \frac{\partial F_l(\vec{\bar{x}}^i, \vec{\bar{\alpha}})}{\partial x_s} \delta x_s^i + \sum_{r=1}^{k} \frac{\partial F_l(\vec{\bar{x}}^i, \vec{\bar{\alpha}})}{\partial \alpha_r} \delta \alpha_r = 0 \; , \tag{29}$$

or, equivalently,

$$\delta\vec{\bar{F}}^i \cong (\partial_x\bar{\mathbb{F}}^i)\delta\vec{x}^i + (\partial_\alpha\bar{\mathbb{F}}^i)\delta\vec{\alpha} = 0\ , \tag{30}$$

where $\mathbb{F}$ is replaced everywhere by $\bar{\mathbb{F}}$, i.e. the respective terms are not evaluated at the **best-fit** coordinate positions $\vec{x}^i$, but at the **approximate** values $\vec{\bar{x}}^i$ at each iteration step. Of course, Eqs. [30] look formally the same as Eqs. [17] above.

Now we can proceed with the second step. Inserting the $\vec{x}^i$ from Eqs. [25] – again replacing all terms $\partial\mathbb{F}^i$ by $\partial\bar{\mathbb{F}}^i$ – into the linearized boundary conditions Eqs. [28] gives the $N \times m$ equations

$$\vec{\bar{F}}^i + (\partial_x\bar{\mathbb{F}}^i)(\vec{X}^i + \mathbb{V}^i(\partial_x\bar{\mathbb{F}}^i)^{\mathrm{T}}\vec{\lambda}^i - \vec{\bar{x}}^i) + (\partial_\alpha\bar{\mathbb{F}}^i)(\vec{\alpha} - \vec{\bar{\alpha}}) = 0\ , \tag{31}$$

from which we extract, in the third step, the $N \times m$ Lagrange multipliers $\vec{\lambda}^i$ in terms of the parameters $\vec{\alpha}$:

$$\vec{\lambda}^i = [(\partial_x\bar{\mathbb{F}}^i)\mathbb{V}^i(\partial_x\bar{\mathbb{F}}^i)^{\mathrm{T}}]^{-1}[-\vec{\bar{F}}^i - (\partial_x\bar{\mathbb{F}}^i)(\vec{X}^i - \vec{\bar{x}}^i) - (\partial_\alpha\bar{\mathbb{F}}^i)(\vec{\alpha} - \vec{\bar{\alpha}})] \tag{32}$$

For uncorrelated errors, this expression reduces for the l.th boundary condition to

$$\lambda_l^i = \bar{P}_l^i\ [-F_l(\vec{\bar{x}}^i, \vec{\bar{\alpha}}) - \sum_{s=1}^{n} \frac{\partial F_l(\vec{\bar{x}}^i, \vec{\bar{\alpha}})}{\partial x_s}(X_s^i - \bar{x}_s^i) - \sum_{r=1}^{k} \frac{\partial F_l(\vec{\bar{x}}^i, \vec{\bar{\alpha}})}{\partial \alpha_r}(\alpha_r - \bar{\alpha}_r)]$$

with the overall 'weight'

$$\bar{P}_l^i = \frac{1}{\sum_{s=1}^{n} \frac{1}{p_s^i}(\frac{\partial F_l(\vec{\bar{x}}^i, \vec{\bar{\alpha}})}{\partial x_s})^2}\ .$$

In the fourth and final step, we insert the $\vec{\lambda}^i$ into Eqs. [24] – once more replacing all terms $\partial\mathbb{F}^i$ by $\partial\bar{\mathbb{F}}^i$ –, which results in the following system of k linear equations for the k parameters $\vec{\alpha}$:

$$\sum_{i=1}^{N}(\partial_\alpha\bar{\mathbb{F}}^i)^{\mathrm{T}}[(\partial_x\bar{\mathbb{F}}^i)\mathbb{V}^i(\partial_x\bar{\mathbb{F}}^i)^{\mathrm{T}}]^{-1}[-\vec{\bar{F}}^i - (\partial_x\bar{\mathbb{F}}^i)(\vec{X}^i - \vec{\bar{x}}^i) - (\partial_\alpha\bar{\mathbb{F}}^i)(\vec{\alpha} - \vec{\bar{\alpha}})] = 0\ . \tag{33}$$

or, after rearrangement,

$$\sum_{i=1}^{N}(\partial_\alpha\bar{\mathbb{F}}^i)^{\mathrm{T}}[(\partial_x\bar{\mathbb{F}}^i)\mathbb{V}^i(\partial_x\bar{\mathbb{F}}^i)^{\mathrm{T}}]^{-1}(\partial_\alpha\bar{\mathbb{F}}^i)\vec{\alpha} = \tag{34}$$

$$\sum_{i=1}^{N}(\partial_\alpha\bar{\mathbb{F}}^i)^{\mathrm{T}}[(\partial_x\bar{\mathbb{F}}^i)\mathbb{V}^i(\partial_x\bar{\mathbb{F}}^i)^{\mathrm{T}}]^{-1}[-\vec{\bar{F}}^i - (\partial_x\bar{\mathbb{F}}^i)(\vec{X}^i - \vec{\bar{x}}^i) + (\partial_\alpha\bar{\mathbb{F}}^i)(\vec{\bar{\alpha}})]\ .$$

At this point, we introduce the $N \times k$ matrix $\bar{\mathbb{A}}$ (the rationale behind this step will become clear only in the discussion of errors, see Section 3):

$$\bar{\mathbb{A}} = \begin{pmatrix} \sqrt{\bar{\mathbb{R}}^1}(\partial_{\alpha_1}\bar{\mathbb{F}}^1) & \cdot & \cdot & \cdot & \sqrt{\bar{\mathbb{R}}^1}(\partial_{\alpha_k}\bar{\mathbb{F}}^1) \\ & & \cdot & & \\ & & \cdot & & \\ & & \cdot & & \\ \sqrt{\bar{\mathbb{R}}^N}(\partial_{\alpha_1}\bar{\mathbb{F}}^N) & \cdot & \cdot & \cdot & \sqrt{\bar{\mathbb{R}}^N}(\partial_{\alpha_k}\bar{\mathbb{F}}^N) \end{pmatrix}\ , \tag{35}$$

where we introduced the short-hand notation

$$\bar{\mathbb{R}}^i = [(\partial_x \bar{\mathbb{F}}^i)\mathbb{V}^i(\partial_x \bar{\mathbb{F}}^i)^{\mathrm{T}}]^{-1} \; . \tag{36}$$

for this $m \times m$ matrix. The notation $\sqrt{\bar{\mathbb{R}}^i}$ refers to the matrix which, when multiplied with its transposed matrix, reproduces the matrix $\bar{\mathbb{R}}^i$ (according to the Cholesky decomposition [10], this is possible for a symmetric and positive-definite matrix, as is the case here). Further, through the use of the m-component vector

$$(\partial_{\alpha_r} \bar{\mathbb{F}}^i) = \begin{pmatrix} \frac{\partial F_1(\vec{\bar{x}}^i, \vec{\bar{\alpha}})}{\partial \alpha_r} \\ \cdots \\ \frac{\partial F_l(\vec{\bar{x}}^i, \vec{\bar{\alpha}})}{\partial \alpha_r} \\ \cdots \\ \frac{\partial F_m(\vec{\bar{x}}^i, \vec{\bar{\alpha}})}{\partial \alpha_r} \end{pmatrix} \; , \tag{37}$$

each element of the above matrix $\bar{\mathbb{A}}$ represents in fact a vector with as many components as the number of boundary conditions.

We further introduce the N-component vector

$$\vec{a} = \begin{pmatrix} \sqrt{\bar{\mathbb{R}}^1}[-\vec{\bar{F}}^1 - (\partial_x \bar{\mathbb{F}}^1)(\vec{X}^1 - \vec{\bar{x}}^1) + (\partial_\alpha \bar{\mathbb{F}}^1)\vec{\bar{\alpha}}] \\ \cdots \\ \sqrt{\bar{\mathbb{R}}^i}[-\vec{\bar{F}}^i - (\partial_x \bar{\mathbb{F}}^i)(\vec{X}^i - \vec{\bar{x}}^i) + (\partial_\alpha \bar{\mathbb{F}}^i)\vec{\bar{\alpha}}] \\ \cdots \\ \sqrt{\bar{\mathbb{R}}^N}[-\vec{\bar{F}}^N - (\partial_x \bar{\mathbb{F}}^N)(\vec{X}^N - \vec{\bar{x}}^N) + (\partial_\alpha \bar{\mathbb{F}}^N)\vec{\bar{\alpha}}] \end{pmatrix} \; , \tag{38}$$

where the same calculational conventions apply.

It is understood that in multiplications involving the matrix $\bar{\mathbb{A}}$ or the vector $\vec{a}$, the first step is always the contraction of the m components of each element to a scalar number, by multiplying with the pertinent transposed vector.

The above conventions ensure that the general element $A_{\alpha_r \alpha_{r'}}$ of the

$$(k \times m) \cdot (m \times n) \cdot (n \times n) \cdot (n \times m) \cdot (m \times k) = k \times k$$

matrix $\bar{\mathbb{A}}^{\mathrm{T}} \bar{\mathbb{A}}$ is indeed the scalar number

$$A_{\alpha_r \alpha_{r'}} = \sum_{i=1}^{N} (\partial_{\alpha_r} \bar{\mathbb{F}}^i)^{\mathrm{T}} [(\partial_x \bar{\mathbb{F}}^i)\mathbb{V}^i(\partial_x \bar{\mathbb{F}}^i)^{\mathrm{T}}]^{-1} (\partial_{\alpha_{r'}} \bar{\mathbb{F}}^i) \tag{39}$$

Analogously, the

$$(k \times m) \cdot (m \times n) \cdot (n \times n) \cdot (n \times m) \cdot (m \times 1) = k \times 1$$

matrix or, equivalently, k-component vector $\bar{\mathbb{A}}^{\mathrm{T}} \vec{a}$ is indeed

$$\bar{\mathbb{A}}^{\mathrm{T}} \vec{a} = \sum_{i=1}^{N} (\partial_\alpha \bar{\mathbb{F}}^i)^{\mathrm{T}} [(\partial_x \bar{\mathbb{F}}^i)\mathbb{V}^i(\partial_x \bar{\mathbb{F}}^i)^{\mathrm{T}}]^{-1} [-\vec{\bar{F}}^i - (\partial_x \bar{\mathbb{F}}^i)(\vec{X}^i - \vec{\bar{x}}^i) + (\partial_\alpha \bar{\mathbb{F}}^i)\vec{\bar{\alpha}}] \; . \tag{40}$$

Equations [34] can then be written as

$$\bar{\mathbb{A}}^{\mathrm{T}}\bar{\mathbb{A}}\,\vec{\alpha} = \bar{\mathbb{A}}^{\mathrm{T}}\vec{a}\;, \tag{41}$$

with the solution for the k wanted parameters

$$\vec{\alpha} = (\bar{\mathbb{A}}^{\mathrm{T}}\bar{\mathbb{A}})^{-1}\,\bar{\mathbb{A}}^{\mathrm{T}}\vec{a}\;. \tag{42}$$

With $\vec{\alpha}$ known, the Lagrange multipliers $\vec{\lambda}^i$ can be determined from Eqs. [32]. They are used to calculate the best-fit values of coordinate positions $\vec{\bar{x}}^i$ from Eqs. [25]. The thus determined best-fit values of coordinate positions and parameters serve as improved guesses for the next iteration step. After satisfactory convergence of the results, the iteration can be terminated.

3 Errors and covariance matrix

The calculation of the errors, and their correlations, of the k wanted parameters $\vec{\alpha}$ proceeds again in the approximation of a linear Taylor expansion of the boundary conditions. For the **parameters** $\vec{\alpha}$, the expansion is made around their approximate values in each iteration step, alike the procedure discussed earlier in Section 2 Unlike the earlier procedure, the expansion of the **coordinate positions** is made around the set of **measured** values, for **all** iteration steps. To highlight the evaluation of the boundary conditions at the position of the measured coordinates, we use the notation $\hat{F}$ and $\hat{\mathbb{F}}$ *in lieu* of $\bar{F}$ and $\bar{\mathbb{F}}$, respectively.

Accordingly, we define the following variations around the approximate values: the n-component vector of coordinate variations

$$\vec{v}_x^{\,i} = \vec{x}^i - \vec{X}^i \tag{43}$$

and the k-component vector of parameter variations

$$\vec{v}_\alpha = \vec{\alpha} - \vec{\bar{\alpha}}\;. \tag{44}$$

The boundary conditions then read as

$$\vec{F}^i \cong \vec{\hat{F}}^i + (\partial_x\hat{\mathbb{F}}^i)\vec{v}_x^{\,i} + (\partial_\alpha\hat{\mathbb{F}}^i)\vec{v}_\alpha = 0\;. \tag{45}$$

We note the equalities

$$\delta\vec{v}_x^{\,i} = \delta\vec{x}^i$$

and

$$\delta\vec{v}_\alpha = \delta\vec{\alpha}\;,$$

which permit to repeat all calculational steps outlined in Section 2, i.e. setting up the same system of equations and applying the same method for their solution. Unless explicitly stated, we use also the same nomenclature.

Accordingly, using the matrix (see Eq. [35])

$$\hat{\mathbb{A}} = \begin{pmatrix} \sqrt{\hat{\mathbb{R}}^1}(\partial_{\alpha_1}\hat{\mathbb{F}}^1) & \cdot & \cdot & \cdot & \sqrt{\hat{\mathbb{R}}^1}(\partial_{\alpha_k}\hat{\mathbb{F}}^1) \\ & \cdot & & & \\ & \cdot & & & \\ & \cdot & & & \\ \sqrt{\hat{\mathbb{R}}^N}(\partial_{\alpha_1}\hat{\mathbb{F}}^N) & \cdot & \cdot & \cdot & \sqrt{\hat{\mathbb{R}}^N}(\partial_{\alpha_k}\hat{\mathbb{F}}^N) \end{pmatrix} , \tag{46}$$

we can directly give the solution for the variation of the k parameters as

$$\vec{v}_\alpha = (\hat{\mathbb{A}}^{\mathrm{T}}\hat{\mathbb{A}})^{-1}\,\hat{\mathbb{A}}^{\mathrm{T}}\vec{b}\,. \tag{47}$$

However, the vector $\vec{b}$, which is constructed analogously to the vector $\vec{a}$ in Section 2 (see Eq. [38]), is – apart from the evaluation of the boundary conditions and their derivatives not at approximate coordinate positions, but at their measured positions – different from the vector $\vec{a}$ in two aspects: the term proportional to $(\vec{X}^i - \vec{\bar{x}}^i)$ is now identical to zero, and the term $(\partial_\alpha\hat{\mathbb{F}}^i)\vec{\alpha}$ has been absorbed into the parameters' variation. Accordingly, the vector $\vec{b}$ reads as

$$\vec{b} = \begin{pmatrix} \sqrt{\hat{\mathbb{R}}^1}[-\vec{\hat{F}}^1] \\ \cdots \\ \sqrt{\hat{\mathbb{R}}^i}[-\vec{\hat{F}}^i] \\ \cdots \\ \sqrt{\hat{\mathbb{R}}^N}[-\vec{\hat{F}}^N] \end{pmatrix} , \tag{48}$$

Here, not the expectation value of $\vec{v}_\alpha$ is of interest but its variance. According to the rule of variance calculation of the product of a matrix with a vector, we obtain from Eqs. [47] the variance

$$\mathrm{v}ar\;\vec{v}_\alpha = (\hat{\mathbb{A}}^{\mathrm{T}}\hat{\mathbb{A}})^{-1}\,\hat{\mathbb{A}}^{\mathrm{T}}\,(\mathrm{v}ar\;\vec{b})\,\hat{\mathbb{A}}(\hat{\mathbb{A}}^{\mathrm{T}}\hat{\mathbb{A}})^{-1}\,. \tag{49}$$

The covariance matrix $\mathrm{v}ar\;\vec{b}$ has the form

$$\mathrm{v}ar\;\vec{b} = \sigma^2 \begin{pmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & 1 \end{pmatrix} , \tag{50}$$

since the N measurements, other than the coordinate measurements within one measurement point, are independent of each other. Furthermore, because of the normalization by appropriate weights, all N measurements have the same variance σ^2, the expectation value of which is unity.

In retrospect, we recognize that the matrix $\hat{\mathbb{A}}$ and the vector $\vec{b}$ were defined with the goal in mind to make possible this generic assertion on $\mathrm{v}ar\,\vec{b}$. What matters in practical calculations, is the $m \times m$ matrix $\hat{\mathbb{A}}^{\mathrm{T}}\hat{\mathbb{A}}$ only, for the calculation of both the best-fit parameters and their correlation matrix.

After inserting this covariance matrix into Eq. [49] we obtain the result

$$\mathrm{v}ar\;\vec{v}_\alpha = \sigma^2(\hat{\mathbb{A}}^{\mathrm{T}}\hat{\mathbb{A}})^{-1}\,. \tag{51}$$

4 Geometric interpretation

The GLSF has an elegant geometrical interpretion which we give here for the special case of uncorrelated errors of the coordinates of a measurement point, and for one boundary condition (it is conjectured that the interpretation holds also for the case with correlated errors and for several boundary conditions).

We suppose that the GLSF has successfully converged after a few iterations, and that the best-fit coordinates and the best-fit parameters are all numerically determined.

We start from Eqs. [26] which for the i.th measurement point ($1 \leq i \leq N$), the s.th coordinate ($1 \leq s \leq n$) and for one boundary condition reads as

$$x_s^i = X_s^i + \lambda^i \frac{1}{p_s^i} \frac{\partial F^i}{\partial x_s} \; . \tag{52}$$

This is the equation of the straight line in measurement space which connects the measured point $\vec{X}^i$ with the associated best-fit point $\vec{x}^i$. The direction of the straight line is given by the n-component vector

$$\begin{pmatrix} \frac{1}{p_1^i} \frac{\partial F^i}{\partial x_1} \\ \cdots \\ \frac{1}{p_s^i} \frac{\partial F^i}{\partial x_s} \\ \cdots \\ \frac{1}{p_n^i} \frac{\partial F^i}{\partial x_n} \end{pmatrix} \; . \tag{53}$$

The best-fit point is located on the hyper-sheet which represents the boundary condition. The vector which is perpendicular to the tangential plane in the best-fit point, is given by

$$\vec{n}_{hs}^i = \begin{pmatrix} \frac{\partial F^i}{\partial x_1} \\ \cdots \\ \frac{\partial F^i}{\partial x_s} \\ \cdots \\ \frac{\partial F^i}{\partial x_n} \end{pmatrix} \; . \tag{54}$$

From this follows that the line connecting the points $\vec{X}^i$ and $\vec{x}^i$ will only in case of equal weights p_s^i, i.e. of equal measurement errors in all coordinates, be perpendicular to the hyper-sheet.

We now construct an error-ellipsoid around the measured point $\vec{X}^i$, with half-axes σ_s^i. This error-ellipsoid is given by

$$p_1^i (u_1 - X_1^i)^2 + \ldots + p_s^i (u_s - X_s^i)^2 + \ldots + p_n^i (u_n - X_n^i)^2 = 1 \tag{55}$$

The tangential plane in a given point $\vec{v}^i$ on the i.th error-ellipsoid is given by

$$p_1^i (u_1 - X_1^i)(v_1^i - X_1^i) + \ldots + p_s^i (u_s - X_s^i)(v_s^i - X_s^i) + \ldots + p_n^i (u_n - X_n^i)(v_n^i - X_n^i) = 1 \tag{56}$$

The vector perpendicular to the tangential plane to the error-ellipsoid in the point $\vec{v}^i$ is

$$\vec{n}^i_{ee} = \begin{pmatrix} p^i_1(v^i_1 - X^i_1) \\ \cdots \\ p^i_s(v^i_s - X^i_s) \\ \cdots \\ p^i_n(v^i_n - X^i_n) \end{pmatrix} . \tag{57}$$

We take for the point $\vec{v}^i$ specifically the point where the straight line from $\vec{X}^i$ to $\vec{x}^i$ crosses the error-ellipsoid. For this, it follows from Eqs. [52] and [55] that the pertinent λ^i_0 must take the numerical value

$$\lambda^i_0 = \pm \frac{1}{\sqrt{\sum_{s=1}^n \frac{1}{p^i_s} \left[\frac{\partial F^i}{\partial x_s}\right]^2}} . \tag{58}$$

The vector perpendicular to the tangential plane to the error-ellipsoid in the point $\vec{v}^i$ is then

$$\vec{n}^i_{ee} = \begin{pmatrix} \lambda^i_0 \frac{\partial F^i}{\partial x_1} \\ \cdots \\ \lambda^i_0 \frac{\partial F^i}{\partial x_s} \\ \cdots \\ \lambda^i_0 \frac{\partial F^i}{\partial x_n} \end{pmatrix} . \tag{59}$$

A comparison with the vector $\vec{n}^i_{hs}$ shows that both vectors are parallel.

Therefore, the position of the best-fit point $\vec{x}^i$ on the hyper-sheet is such that the tangential plane to the hyper-sheet in this point is parallel to the tangential plane to the error-ellipsoid around the measured point $\vec{X}^i$, in the point where the straight line connecting $\vec{X}^i$ with $\vec{x}^i$ crosses the error-ellipsoid.

Figure 1 illustrates this geometric interpretation in 3D measurement space.

5 Application to straight-line fit

As a simple example of the GLSF we discuss its application to the problem of fitting a straight line through N measured points $(X^i \pm \sigma^i_x, Y^i \pm \sigma^i_y)$. We assume that the measurements of the two coordinates x and y are uncorrelated.

The boundary condition (see Eq. [4]) for the 'best-fit' points (x^i, y^i) is

$$F^i = y^i - kx^i - d = 0 , \tag{60}$$

with the slope k and the intercept d as parameters.

This special form of the boundary condition permits to get away with less approximate values of the coordinates and parameters than needed in the algorithm for the general case.

The fit aims at minimizing (see Eq. [12])

$$\chi^2 = \sum_{i=1}^N \left[p^i_x (X^i - x^i)^2 + p^i_y (Y^i - y^i)^2 \right] , \tag{61}$$

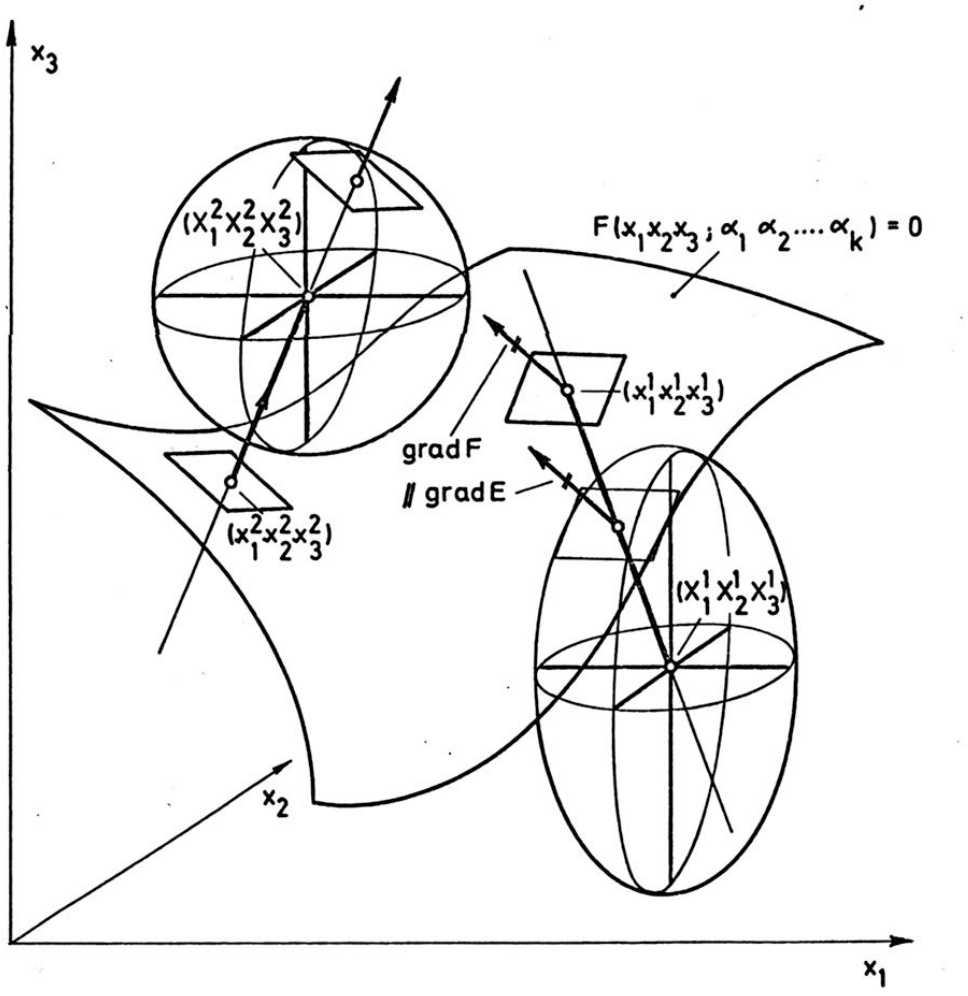

Figure 1: Measured points, error-ellipsoids and best-fit hyper-sheet in 3D measurement space.

where $p_x^i = 1/(\sigma_x^i)^2$ and $p_y^i = 1/(\sigma_y^i)^2$.

From the equation $\delta\chi^2 = 0$ (see Eq. [15]) and the N equations $\delta F^i = 0$ (see Eqs. [17]) follow after the multiplication with N Lagrange multipliers λ^i the $2N + 2$ equations (see Eqs. [22] and [24]):

$$-k\lambda^i + p_x^i(X^i - x^i) = 0 \tag{62}$$

$$\lambda^i + p_y^i(Y^i - y^i) = 0 \tag{63}$$

$$\sum_{i=1}^{N} \lambda^i \bar{x}^i = 0 \tag{64}$$

$$\sum_{i=1}^{N} \lambda^i = 0\,, \tag{65}$$

in addition to the N boundary conditions $F^i = 0$, for the $3N + 2$ unknowns $\vec{x}$, $\vec{y}$, λ^i, k and d. The linear dependence of the boundary conditions on the coordinates necessitates only

in Eq. [64] the use of approximate coordinates $\bar{x}^i$ *in lieu* of the best-fit coordinates x^i.

According to Eqs. [26] we calculate the best-fit coordinates

$$\begin{aligned} x^i &= X^i - \frac{k}{p_x^i}\lambda^i \\ y^i &= Y^i + \frac{1}{p_y^i}\lambda^i \ , \end{aligned} \tag{66}$$

and insert them into the N boundary conditions Eqs. [60]. This gives the Lagrange multipliers

$$\lambda^i = \bar{P}^i(kX^i - Y^i + d) \ , \tag{67}$$

where $\bar{P}^i$ is the approximate 'weight' of the i.th measurement point,

$$\bar{P}^i = \frac{1}{1/p_y^i + \bar{k}^2/p_x^i} \ , \tag{68}$$

for the calculation of which the use of an approximate slope $\bar{k}$ is necessary.

Now we insert these Lagrange multipliers into Eqs. [64] and [65]. This yields two linear equations for the two parameters k and d:

$$k\sum \bar{P}^i X^i \bar{x}^i + d\sum \bar{P}^i \bar{x}^i = \sum \bar{P}^i Y^i \bar{x}^i \tag{69}$$

$$k\sum \bar{P}^i X^i + d\sum \bar{P}^i = \sum \bar{P}^i Y^i \ , \tag{70}$$

where all sums run from 1 to N. The best-fit parameters are then

$$k = \frac{(\sum \bar{P}^i)(\sum \bar{P}^i Y^i \bar{x}^i) - (\sum \bar{P}^i Y^i)(\sum \bar{P}^i \bar{x}^i)}{(\sum \bar{P}^i)(\sum \bar{P}^i X^i \bar{x}^i) - (\sum \bar{P}^i X^i)(\sum \bar{P}^i \bar{x}^i)} \tag{71}$$

$$d = \frac{(\sum \bar{P}^i Y^i)(\sum \bar{P}^i X^i \bar{x}^i) - (\sum \bar{P}^i X^i)(\sum \bar{P}^i Y^i \bar{x}^i)}{(\sum \bar{P}^i)(\sum \bar{P}^i X^i \bar{x}^i) - (\sum \bar{P}^i X^i)(\sum \bar{P}^i \bar{x}^i)} \ . \tag{72}$$

We note that these expressions for the straight-line parameters k and d look formally like those obtained in a Standard Least Squares Fit where only the y coordinate has an error, by contrast to the x coordinate which is error-free. The difference is the weight of the measurement point which is for the Standard Least Squares Fit

$$P^i = p_y^i = 1/(\sigma_y^i)^2 \ ,$$

by contrast to the weight given by Eq. [68] for the GLSF.

After the measured coordinates X^i and Y^i have initially been used as approximate cooordinates, better approximate coordinates $\bar{x}^i$ can now be calculated from Eqs. [66] with the help of the Lagrange multipliers from Eqs. [67]. Also, better approximate weights $\bar{P}^i$ can now be calculated with the result for k. With these better approximations the next iteration is made, and so on, until convergence of the results at the wanted level of precision is achieved.

Figure 2 illustrates the main issue of the GLSF: complete symmetry between both coordinates with respect to the treatment of their errors. Besides the measured points with their

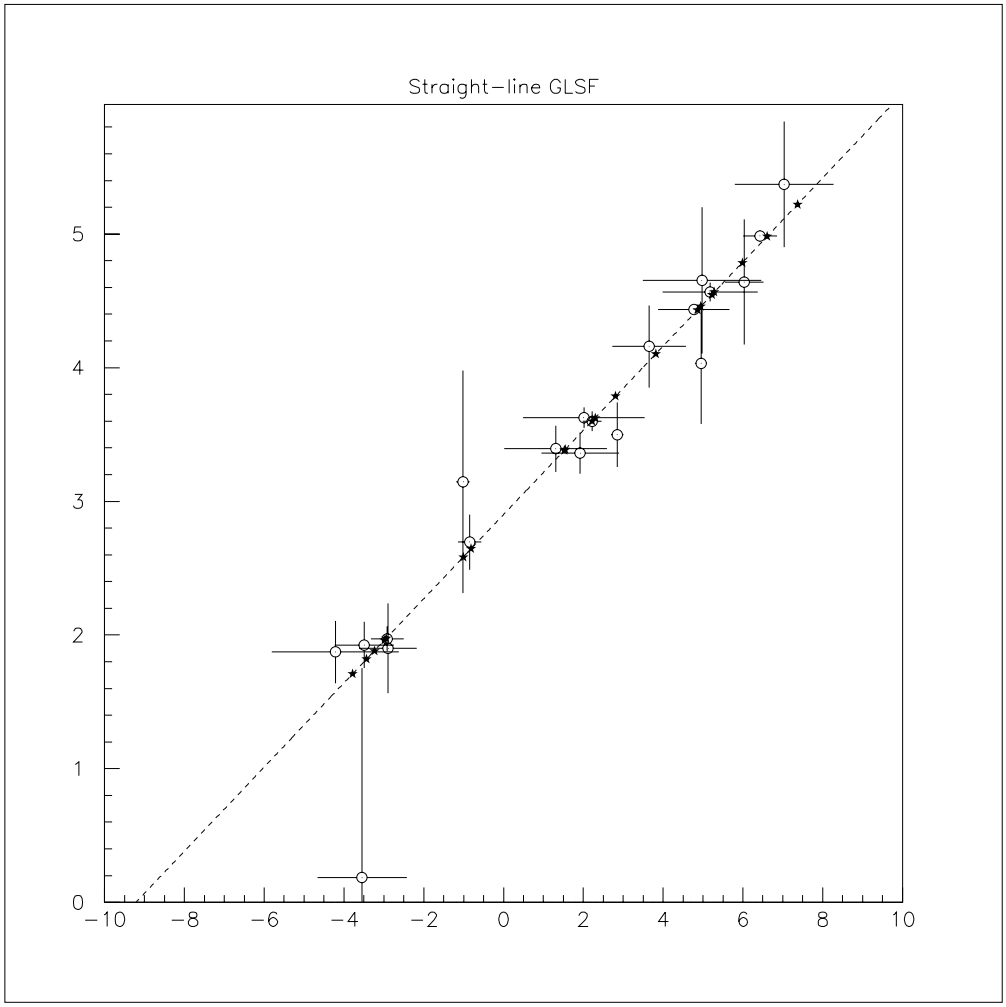

Figure 2: Measured points with errors, GLSF best-fit straight line, and best-fit points.

errors, also the best-fit points are shown. Table 1 gives the results for the slope k and the intercept d of straight-line fits to the measured points shown in Fig. 2, for the Standard Least Squares Fit as well as for iterative steps of the GLSF. The difference between the results of the Standard Least Squares Fit and the GSLF is noticeable. For the GSLF, in many cases already the first pass with the measured coordinates as approximate coordinates in Eqs. [71] and [72] will give numerical results for the parameters k and d with sufficient precision.

According to the considerations outlined in Section 3, the errors of the parameters are given by

$$\sigma_k^2 = \frac{\sum \bar{P}^i}{(\sum \bar{P}^i)(\sum \bar{P}^i (X^i)^2) - (\sum \bar{P}^i X^i)^2} \tag{73}$$

$$\sigma_d^2 = \frac{\sum \bar{P}^i (X^i)^2}{(\sum \bar{P}^i)(\sum \bar{P}^i (X^i)^2) - (\sum \bar{P}^i X^i)^2} \, , \tag{74}$$

where again the sums run from 1 to N, and the same comment as above applies regarding the comparison with the variance of the parameters in the Standard Least Squares Fit.

Table 1: Best-fit values of slope and intercept of straight-line fits of the measured points shown in Fig. 2.

	Slope	Intercept
True values	0.3	3.
Standard Fit	0.3161	2.9276
GLSF Pass No.1	0.3138	2.9034
GLSF Pass No.2	0.3149	2.9015
GLSF Pass No.3	0.3149 ± 0.0177	2.9014 ± 0.0679

6 Application to TPC track fit

6.1 Covariance matrix of TPC coordinate measurements

The HARP TPC has the pads organized in concentric rings. Therefore, the directly measured coordinates are the two transverse coordinates r and $r{\cdot}\phi$, and the longitudinal coordinate z. The errors of these three coordinates are considered mutually uncorrelated.

In the transformation from (r, $r{\cdot}\phi$, z) to the righthanded Cartesian system (x, y, z), the z coordinate error remains uncorrelated while the x and y coordinate errors become correlated among themselves, depending on the azimuthal ϕ angle. This correlation is taken into account in the GLSF.

In the coordinate system of uncorrelated transverse coordinates r and $r{\cdot}\phi$, the covariance matrix is

$$\mathbb{V}^* = \begin{pmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_{r\phi}^2 \end{pmatrix} . \tag{75}$$

We transform this matrix into the one of a coordinate system rotated by the azimuthal angle ϕ. The result is

$$\mathbb{V} = \begin{pmatrix} \sigma_r^2 \cos^2\phi + \sigma_{r\phi}^2 \sin^2\phi & \sin\phi\cos\phi(\sigma_{r\phi}^2 - \sigma_r^2) \\ \sin\phi\cos\phi(\sigma_{r\phi}^2 - \sigma_r^2) & \sigma_r^2 \sin^2\phi + \sigma_{r\phi}^2 \cos^2\phi \end{pmatrix} . \tag{76}$$

6.2 Algorithm for 3D TPC track fit

The standard HARP coordinate system (x, y, z) is employed: looking downstream in the $+z$ direction, the $+x$ coordinate points to the left, and the $+y$ coordinate points up.

A charged particle follows in the TPC's uniform solenoidal magnetic field a helical trajectory with the helix axis parallel to the z axis. We parameterize the helix by the following set of three equations with five parameters:

$$\begin{cases} x = a + r\cos\psi \\ y = b + r\sin\psi \\ z = c + k\psi \ . \end{cases} \tag{77}$$

We utilize the symbol ψ for the azimuthal angle parameter of the helix, by contrast to the symbol ϕ which denotes the azimuthal angle of the rotational TPC geometry.

The TPC measures the coordinate vector of the i.th space point along a track,

$$\vec{X}^i = \begin{pmatrix} X^i \\ Y^i \\ Z^i \end{pmatrix} , \tag{78}$$

with the covariance matrix

$$\begin{aligned} \mathbb{V}^i &= \begin{pmatrix} \mathrm{var}(X^i) & \mathrm{cov}(X^i, Y^i) & \mathrm{cov}(X^i, Z^i) \\ \mathrm{cov}(Y^i, X^i) & \mathrm{var}(Y^i) & \mathrm{cov}(Y^i, Z^i) \\ \mathrm{cov}(Z^i, X^i) & \mathrm{cov}(Z^i, Y^i) & \mathrm{var}(Z^i) \end{pmatrix} \\ &= \begin{pmatrix} (\sigma_r^i)^2 \cos^2\Phi^i + (\sigma_{r\phi}^i)^2 \sin^2\Phi^i & \sin\Phi^i \cos\Phi^i((\sigma_{r\phi}^i)^2 - (\sigma_r^i)^2) & 0 \\ \sin\Phi^i \cos\Phi^i((\sigma_{r\phi}^i)^2 - (\sigma_r^i)^2) & (\sigma_r^i)^2 \sin^2\Phi^i + (\sigma_{r\phi}^i)^2 \cos^2\Phi^i & 0 \\ 0 & 0 & \sigma_z^2 \end{pmatrix} , \end{aligned} \tag{79}$$

where $R^i = \sqrt{(X^i)^2 + (Y^i)^2}$ and $\Phi^i = \arctan(Y^i/X^i)$ are the polar coordinates of the transverse projection of the space point $\vec{X}^i$. The transverse coordinate correlation is modelled according to Eq. [76].

The above helix parameterization gives, in vector notation, the set of boundary conditions

$$\vec{F}^i = \begin{pmatrix} x^i - a - r\cos\psi^i = 0 \\ y^i - b - r\sin\psi^i = 0 \\ z^i - c - k\psi^i = 0 \end{pmatrix} \tag{80}$$

with 5 parameters

$$\vec{\alpha} = \begin{pmatrix} a \\ b \\ r \\ c \\ k \end{pmatrix} . \tag{81}$$

The GLSF determines best-fit coordinates $\vec{x}^i$ and best-fit parameters $\vec{\alpha}$. For reasons of 'easy' mathematics in the special case of a helix fit (with a view to exploiting the linearity of the boundary conditions in the parameters $\vec{\alpha}$), we introduce **additional** best-fit parameters: the azimuthal angles $\vec{\psi}$, to be determined by the GLSF (these new parameters are fully correlated with the coordinates $\vec{x}^i$, and therefore do not impair the precision of the parameters $\vec{\alpha}$):

$$\vec{\psi} = \begin{pmatrix} \psi^1 \\ \cdots \\ \psi^i \\ \cdots \\ \psi^N \end{pmatrix} . \tag{82}$$

First, we calculate according to Eq. [20] the 3×3 matrix of derivatives of the boundary conditions w.r.t. the coordinates x, y, and z, with the particularly appealing result

$$\partial_x \mathbb{F}^i = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{83}$$

Analogously, we calculate the $3 \times N$ matrix of derivatives of the boundary conditions w.r.t. the coordinates $\vec{\psi}$:

$$\begin{aligned} \partial_\psi \mathbb{F}^i &= \begin{pmatrix} \frac{\partial F_1}{\partial \psi_1} & \cdots & \frac{\partial F_1}{\partial \psi_i} & \cdots & \frac{\partial F_1}{\partial \psi_N} \\ \frac{\partial F_2}{\partial \psi_1} & \cdots & \frac{\partial F_2}{\partial \psi_i} & \cdots & \frac{\partial F_2}{\partial \psi_N} \\ \frac{\partial F_3}{\partial \psi_1} & \cdots & \frac{\partial F_3}{\partial \psi_i} & \cdots & \frac{\partial F_3}{\partial \psi_N} \end{pmatrix} \\ &= \begin{pmatrix} 0 & \dots & 0 & r\sin\psi^i & 0 & \dots & 0 \\ 0 & \dots & 0 & -r\cos\psi^i & 0 & \dots & 0 \\ 0 & \dots & 0 & -k & 0 & \dots & 0 \end{pmatrix} . \end{aligned} \tag{84}$$

Finally, we calculate according to Eq. [21] the 3×5 matrix of derivatives of the boundary conditions w.r.t. the parameters $\vec{\alpha}$:

$$\partial_\alpha \mathbb{F}^i = \begin{pmatrix} -1 & 0 & -\cos\psi^i & 0 & 0 \\ 0 & -1 & -\sin\psi^i & 0 & 0 \\ 0 & 0 & 0 & -1 & -\psi^i \end{pmatrix} . \tag{85}$$

Because the boundary conditions are linear in the coordinates $\vec{x}^i$ and the parameters $\vec{\alpha}$, they can be written as

$$\vec{F}^i = (\partial_x \mathbb{F}^i)\vec{x}^i + (\partial_\alpha \mathbb{F}^i)\vec{\alpha} = \vec{x}^i + (\partial_\alpha \mathbb{F}^i)\vec{\alpha} = 0 . \tag{86}$$

We now insert according to Eq. [25] the $\vec{x}^i$ into the boundary conditions Eqs. [86] and obtain

$$\vec{F}^i = \vec{X}^i + \mathbb{V}^i \vec{\lambda}^i + (\partial_\alpha \mathbb{F}^i)\vec{\alpha} = 0 , \tag{87}$$

from where we extract the Lagrange multipliers $\vec{\lambda}^i$ in terms of the parameters $\vec{\alpha}$:

$$\vec{\lambda}^i = -(\mathbb{V}^i)^{-1} \left[\vec{X}^i + (\partial_\alpha \mathbb{F}^i)\vec{\alpha} \right] . \tag{88}$$

We insert these Lagrange multipliers first into the N equations

$$(\partial_\psi \mathbb{F}^i)^{\mathrm{T}} \vec{\lambda}^i = 0 \tag{89}$$

which arise from the new parameters $\vec{\psi}$ in analogy to Eq. [24]. The simple form of the matrix $\partial_\psi \mathbb{F}^i$ (see Eq. [84]) permits to re-write these equations in the form

$$(\vec{T}^i)^{\mathrm{T}} \vec{\lambda}^i = 0 , \tag{90}$$

where we introduced the 3-component vector

$$\vec{T}^i = \begin{pmatrix} r\sin\psi^i \\ -r\cos\psi^i \\ -k \end{pmatrix} . \tag{91}$$

This gives the N equations

$$\left(\vec{T}^i\right)^T (\mathbb{V}^i)^{-1} \left[\vec{X}^i + (\partial_\alpha \mathbb{F}^i)\vec{\alpha} \right] = 0 . \tag{92}$$

Then, we insert the Lagrange multipliers $\vec{\lambda}^i$ into the five equations

$$\sum_{i=1}^{N} (\partial_\alpha \mathbb{F}^i)^{\mathrm{T}} \vec{\lambda}^i = 0 . \tag{93}$$

which arise from the parameters $\vec{\alpha}$ according to Eq. [24]. This gives the five equations

$$\sum_{i=1}^{N}(\partial_\alpha \mathbb{F}^i)^T \left(V^i\right)^{-1}\left[\vec{X}^i + (\partial_\alpha \mathbb{F}^i)\vec{\alpha}\right] = 0\ . \tag{94}$$

Altogether, we have now $N+5$ equations for the $N+5$ unknowns $\vec{\psi}^i$ and $\vec{\alpha}$.

First we solve Eqs. [94] for the five parameters $\vec{\alpha}$. Approximating $(\partial_\alpha \mathbb{F}^i)$ by $(\partial_\alpha \bar{\mathbb{F}}^i)$, we obtain

$$\vec{\alpha} = \mathbb{D}^{-1}\vec{d}\,, \tag{95}$$

where we use the 5×5 matrix

$$\mathbb{D} = \sum_{i=1}^{N}(\partial_\alpha \bar{\mathbb{F}}^i)^{\mathrm{T}}(\mathbb{V}^i)^{-1}(\partial_\alpha \bar{\mathbb{F}}^i) \tag{96}$$

and the 5-component vector

$$\vec{d} = -\sum_{i=1}^{N}(\partial_\alpha \bar{\mathbb{F}}^i)^T(\mathbb{V}^i)^{-1}\vec{X}^i\ . \tag{97}$$

Then, we solve Eqs. [92] for the angles $\vec{\psi}^i$. Because of the nonlinear sin and cos functions, we must resort to a linearization of the elements of the $\partial_\alpha \mathbb{F}^i$ matrix w.r.t. ψ^i in order to obtain a system of linear equations. With approximate angles $\bar{\psi}^i$, we can write

$$(\partial_\alpha \mathbb{F}^i)\vec{\alpha} \cong (\partial_\alpha \bar{\mathbb{F}}^i)\vec{\alpha} + \vec{\bar{T}}^i(\psi^i - \bar{\psi}^i)\ ,$$

where $\partial_\alpha \bar{\mathbb{F}}^i$ and $\vec{\bar{T}}^i$ are evaluated at the approximate angles $\bar{\psi}^i$.

We insert this approximation for $(\partial_\alpha \mathbb{F}^i)\vec{\alpha}$ into Eqs. [92] and obtain the N angles ψ^i:

$$\psi^i = \bar{\psi}^i - \frac{(\vec{\bar{T}}^i)^{\mathrm{T}}(\mathbb{V}^i)^{-1}\left[\vec{X}^i + (\partial_\alpha \bar{\mathbb{F}}^i)\vec{\alpha}\right]}{(\vec{\bar{T}}^i)^{\mathrm{T}}(\mathbb{V}^i)^{-1}\vec{\bar{T}}^i}\ . \tag{98}$$

The final set of best-fit parameters $\vec{\alpha}$ and best-fit angles $\vec{\psi}$ is obtained by iterating the solutions of Eqs. [95] and [98]. We start from an initial set of angles $\bar{\psi}^i$ and parameters $\vec{\bar{\alpha}}$, and calculate from Eq. [95] an improved set of parameters. With these, we calculate from Eqs. [98] improved angles. With a good initial estimate of parameters, convergence is achieved after few iterations. After convergence, the covariance matrix for the five parameters $\vec{\alpha}$ is calculated according to Eq. [51] as

$$\mathrm{var}\,\vec{\alpha} = \mathbb{D}^{-1}\ . \tag{99}$$

Figure 3 shows an example of a 2D GLSF of the transverse projection of 'measured' points along a helical track in the TPC. Besides the measured points, also the pertinent best-fit points are shown.

Table 2 shows the fit results for a standard fit, and for the first 3 passes of the GLSF.

The GLSF algorithm of the 3D helix fit of TPC tracks has been coded accordingly, and is now routinely used for TPC track fitting.

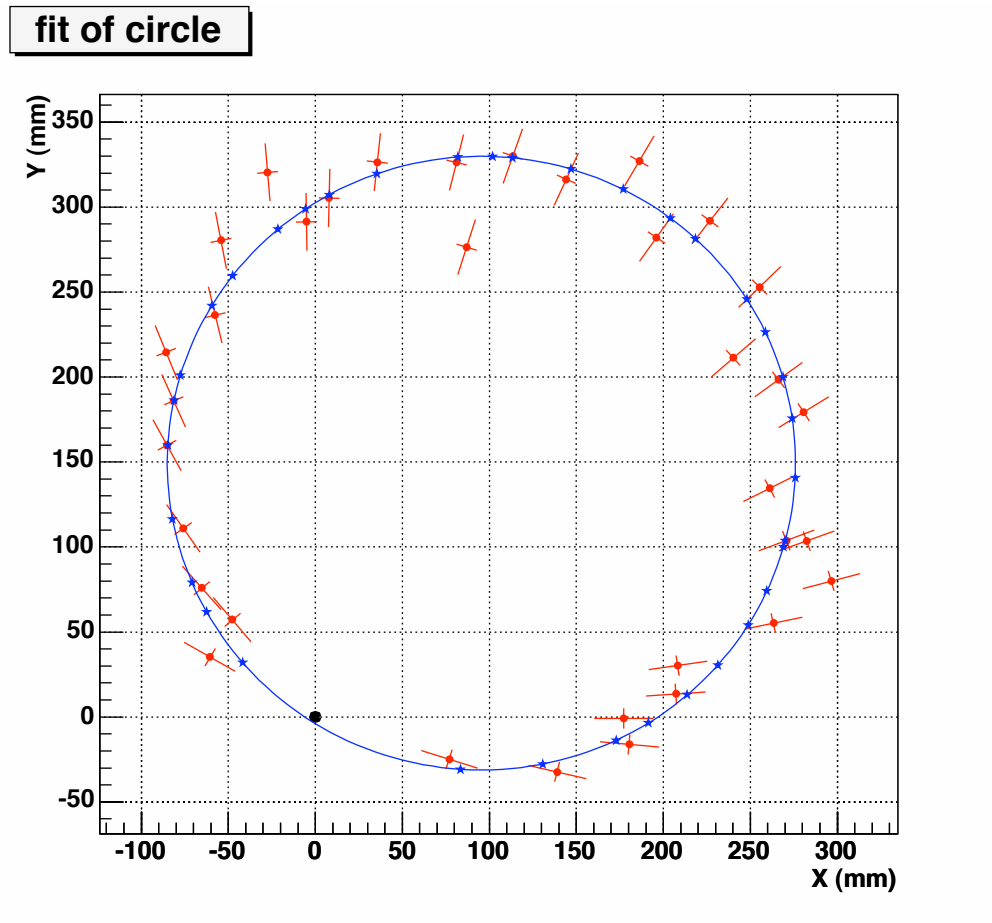

Figure 3: Measured points with errors, GLSF best-fit circle, and best-fit points.

References

[1] W.E. Deming, Phil. Mag. **11** (1931) 146; *ibid.* **17** (1934) 804.

[2] E.R. Cohen, Rev. Mod. Phys. **25** (1953) 709.

[3] D. Brown, Report No. 937, Ballistic Research Laboratories, Aberdeen Proving Ground, Md (USA), 1955 (unpublished).

[4] D. York, Can. J. Phys. **44** (1966) 1079.

[5] G.A. Gerhold, Am. J. Phys. **37** (1969) 156.

[6] F. Dydak, PhD Thesis at the Philosophical Faculty of the University of Vienna, 1972 (unpublished).

[7] J. Orear, Note CLNS-81/496, Cornell University, NY (USA), 1981; published in Am. J. Phys. **50** (1982) 912.

[8] C. Moreno, Meas. Sci. Technol. **7** (1996) 137.

Table 2: Parameters of the GLSF circle to the points shown in Fig. 3; a and b denote the x and y coordinates of the centre, and r the radius of the circle.

	a	b	r
True values	100.	150.	180.
Standard fit	95.6	150.0	182.7
GLSF Pass No.1	95.1	149.0	180.6
GLSF Pass No.2	95.3	149.2	180.6
GLSF Pass No.3	95.4 ± 1.5	149.3 ± 1.8	180.5 ± 2.1

[9] S. Brandt, Datenanalyse, BI Wissenschaftsverlag, Mannheim (Germany), 1992 (in german language).

[10] See e.g. *http://mathworld.wolfram.com/Cholesky/CholeskyDecomposition.html*.