Exploration and extension of an improved Riemann track fitting algorithm

A. Strandlie

NTNU Norwegian University of Science and Technology, P. O. Box 191, N - 2802 Gjøvik, Norway

R. Frühwirth[1]

Institute for High Energy Physics of the Austrian Academy of Sciences, Nikolsdorfer Gasse 18, A - 1050 Wien, Austria

---

Abstract

Recently, a new Riemann track fit which operates on translated and scaled measurements has been proposed. This study shows that the new Riemann fit is virtually as precise as popular approaches such as the Kalman filter or an iterative non-linear track fitting procedure, and significantly more precise than other, non-iterative circular track fitting approaches over a large range of measurement uncertainties. The fit is then extended in two directions: first, the measurements are allowed to lie on plane sensors of arbitrary orientation; second, the full error propagation from the measurements to the estimated circle parameters is computed. The covariance matrix of the estimated track parameters can therefore be computed without recourse to asymptotic properties, and is consequently valid for any number of observation. It does, however, assume normally distributed measurement errors. The calculations are validated on a simulated track sample and show excellent agreement with the theoretical expectations.

Keywords: track fitting, circle fit, Riemann fit, error propagation

PACS: 29.85.-c, 02.50.Tt

Email addresses: are.strandlie@ntnu.no (A. Strandlie), rudolf.fruehwirth@oeaw.ac.at (R. Frühwirth)

[1]Corresponding author

1. Introduction

The trajectory of a charged particle through a high-energy physics detector system is governed by the equations of motion, given by the Lorentz force. In the general case with the presence of an inhomogeneous magnetic field, no analytical solutions to these equations exist, and one has to resort to numerical approaches such as a Runge-Kutta method of some order. In the very simplest case of a vanishing magnetic field, the track model is a straight line. Despite the intrinsic attractiveness of this simple track model, important properties as for instance the momentum and sign of charge of the particle cannot be estimated. The only case enabling such properties to be estimated and at the same time offering an analytical track model is a homogeneous magnetic field with field lines parallel to the beam direction. In this situation, the track model is a helix, or, in the bending plane of the particles, a circle.

Most inner tracking detector systems are therefore embedded in a nearly homogeneous magnetic field. Although general methods such as the Kalman filter [1] or global least-squares estimation [2] can be used in this case, track fitting in the bending plane can also be performed by simple, fast and non-iterative circle fitting methods such as the conformal mapping approach [3], the Karimäki method [4] or the Riemann fit [5]. These non-iterative methods are all based on some kind of simplifying approximation, which in general makes them less precise than more rigorous approaches.

In this paper, we present a thorough study of the precision of a recently proposed, improved Riemann track fit [6]. As suggested by Chernov [7], measurements are transformed in order to achieve invariance under translations and similarity transforms. We show that the improved Riemann fit is significantly more precise than some of the most popular, non-iterative approaches and virtually as precise as the Kalman filter, a global least-squares approach and an iterative, non-linear method.

In addition to estimating the track parameters, a track fitting algorithm should be able to assess the degree of uncertainty of these estimates. These uncertainties and the correlations between them are summarized in the covariance matrix. In [8], the covariance matrix of the track parameters was based on large-sample (asymptotic) properties of the sample covariance matrix of the observations. Here we present the full sequence of error propagation steps from the observations to the final track parameters. It is valid for any number of observations under the assumption of normally distributed measurement errors. The derivation is simpler in the statistically equivalent implementation of the Riemann fit proposed in [9], where the measurements are projected to the paraboloid $z = x^2 + y^2$ rather than to the Riemann sphere.

The paper is organized as follows. After a recollection of the basic concepts of the Riemann track fitting method, the recently introduced improvements to the original algorithm are reviewed. In a simulation study of a generic inner tracking system we show results comparing the precision of the improved Riemann fit with a set of circular track fitting methods. The derivation of the error propagation from the measurements to the estimated circle parameters is then presented and validated with simulated tracks. The paper is concluded by a summary and an outlook to further work.

2. The improved track fit on the Riemann paraboloid

The Riemann paraboloid is positioned on top of the (x, y)-plane with its global minimum at the origin of the plane. We assume that the measured points in the (x, y)-plane are given in Cartesian coordinates, $(u_i, v_i),\ i = 1, \ldots, N$. A covariance matrix $\boldsymbol{V}_i$ is attached to each point. The covariance matrix is arbitrary in principle, but is required to be positive definite in order to avoid problems with rank-deficient matrices during the error propagation.

There are two important special cases. If the radial error of the point (u_i, v_i) can be neglected, its covariance matrix has the form:

$$\boldsymbol{V}_i = \frac{1}{\sqrt{u_i^2 + v_i^2}} \begin{pmatrix} \delta_i^2 u_i^2 + \sigma_i^2 v_i^2 & u_i v_i(\delta_i^2 - \sigma_i^2) \\ u_i v_i(\delta_i^2 - \sigma_i^2) & \delta_i^2 v_i^2 + \sigma_i^2 u_i^2 \end{pmatrix},$$

where σ_i is the standard deviation of the position error in the tangential direction, and δ_i is positive, but much smaller than σ_i, for instance $\delta_i = 0.01 \cdot \sigma_i$.

If the point (u_i, v_i) is a position measurement on a thin plane sensor with normal unit vector $\boldsymbol{a}_i = (a_{i,u}, a_{i,v})^{\mathsf{T}}$, its covariance matrix has the form:

$$\boldsymbol{V}_i = \begin{pmatrix} \delta_i^2 a_{i,u}^2 + \sigma_i^2 a_{i,v}^2 & a_{i,u} a_{i,v} (\delta_i^2 - \sigma_i^2) \\ a_{i,u} a_{i,v} (\delta_i^2 - \sigma_i^2) & \delta_i^2 a_{i,v}^2 + \sigma_i^2 a_{i,u}^2 \end{pmatrix},$$

where σ_i is the standard deviation of the position error of the sensor, and δ_i is again positive, but much smaller than σ_i, for instance $\delta_i = 0.01 \cdot \sigma_i$.

The mapping from the (u, v)-plane to the Riemann paraboloid is given by:

$$\begin{aligned} x_i &= u_i \\ y_i &= v_i \\ z_i &= u_i^2 + v_i^2 \end{aligned}$$

By this mapping, the circle in the plane with the equation

$$(u - u_0)^2 + (v - v_0)^2 = \rho^2$$

is mapped to the plane in 3D space with the equation

$$z - 2xu_0 - 2yv_0 = \rho^2 - u_0^2 - v_0^2$$

A point with position $\boldsymbol{r} = (x, y, z)^{\mathsf{T}}$ satisfying $\boldsymbol{n}^{\mathsf{T}}\boldsymbol{r} + c = 0$ lies in the plane with unit normal vector $\boldsymbol{n}$ and signed distance c from the origin. The plane is fitted to the points $\boldsymbol{r}_i,\ i = 1, \ldots, N$, by minimizing the following objective function:

$$S = \sum_{i=1}^{N} w_i {d_i}^2,$$

where d_i is the distance from the point $\boldsymbol{r}_i = (x_i, y_i, z_i)^{\mathsf{T}}$ to the plane and w_i is its weight. The weights are defined by:

$$w_i \propto 1/\sigma_i^2, \quad \sum_{i=1}^{N} w_i = 1$$

The solution to this minimization problem is a plane with a normal vector $\boldsymbol{n}$ that is the unit eigenvector corresponding to the smallest eigenvalue of the weighted sample covariance matrix $\boldsymbol{A}$, defined as:

$$\boldsymbol{A} = \sum_{i=1}^{N} w_i (\boldsymbol{r}_i - \boldsymbol{r}_0)(\boldsymbol{r}_i - \boldsymbol{r}_0)^{\mathsf{T}},$$

where $\boldsymbol{r}_0$ is the weighted average or center of gravity:

$$\boldsymbol{r}_0 = \sum_{i=1}^{N} w_i \boldsymbol{r}_i$$

Given $\boldsymbol{n}$, c is computed by

$$c = -\boldsymbol{n}^{\mathsf{T}} \boldsymbol{r}_0$$

The parameters $\boldsymbol{n}$ and c of the plane can then be mapped to a set of parameters of the corresponding circle in the (u, v)-plane [9].

We have followed Chernov's [7] suggestion of centering and scaling the measurements before mapping to the paraboloid, in order to achieve invariance of the fit under translations and similarities [6]. Centering is performed by subtracting the average:

$$u_{\mathrm{c},i} = u_i - \overline{u}, \quad v_{\mathrm{c},i} = v_i - \overline{v}, \quad i = 1, \ldots, N$$

with

$$\overline{u} = \frac{1}{N} \sum_{i=1}^{N} u_i, \quad \overline{v} = \frac{1}{N} \sum_{i=1}^{N} v_i$$

The centered measurements $u_{c,i}$ and $v_{c,i}$ are arranged in column vectors $\boldsymbol{u}_c$ and $\boldsymbol{v}_c$. Centered and scaled measurement vectors $\boldsymbol{u}_{cs}$ and $\boldsymbol{v}_{cs}$ are then obtained by:

$$s = b/\sqrt{(\boldsymbol{u}_c^\mathsf{T}\boldsymbol{u}_c + \boldsymbol{v}_c^\mathsf{T}\boldsymbol{v}_c)/N}$$

$$\boldsymbol{u}_{cs} = s \cdot \boldsymbol{u}_c,\ \boldsymbol{v}_{cs} = s \cdot \boldsymbol{v}_c$$

where s is the scaling factor and b an arbitrary, preselected constant [6].

3. Simulation study in a generic cylindrical detector

We have simulated a generic type of a cylindrical detector system embedded in a perfectly homogeneous magnetic field, so that the track model in the bending plane of the particles is a circle. The simulated track sample is the same as the one used in [6]: 10000 tracks coming from the origin with radii of curvature in a range from about 1.5 m to about 750 m. This corresponds to arcs between less than 0.1 degrees and about 20 degrees, following a reasonably flat distribution in this range. There are between 10 and 12 hits per track, and the single hit resolution varies between 0.1 mm and 1.5 mm. The measurement error in the radial direction is assumed to be negligible. We assume no background and thereby implicitly a perfect pattern recognition. The simulation does not include material and detector effects such as multiple scattering, energy loss and sensor misalignment. Measurements in different layers are therefore statistically independent.

We have compared the performance of the modified Riemann fit with a number of other circular track fitting algorithms by considering the mean-square error (MSE) of the residuals $\boldsymbol{\delta}$ of the track parameters, i.e. the estimated track parameters minus the true ones. The MSE is defined by:

$$\mathsf{MSE}[\boldsymbol{\delta}] = \mathsf{det}\big(\Sigma[\boldsymbol{\delta}] + \bar{\boldsymbol{\delta}}\bar{\boldsymbol{\delta}}^\mathsf{T}\big)$$

where $\Sigma[\boldsymbol{\delta}]$ is the sample covariance matrix and $\bar{\boldsymbol{\delta}}$ is the sample mean of the residuals. $\bar{\boldsymbol{\delta}}$ is the least-squares estimate of the bias of the track parameters.

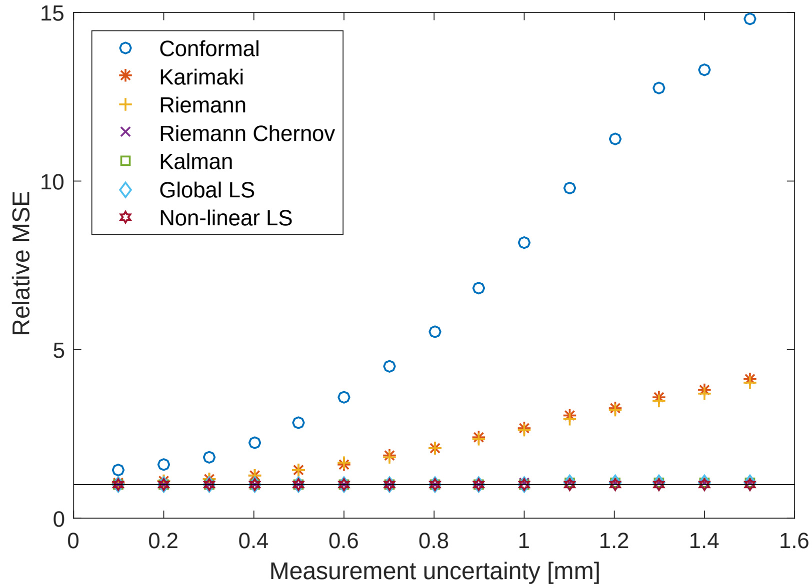

Figure 1: The ratio of the relative generalized mean-square error as a function of the measurement uncertainty.

Figure 1 shows the MSE of the various estimators relative to the baseline, which is an iterative, non-linear least-squares approach using the Levenberg-Marquardt algorithm. Firstly, it can be seen that the modified Riemann fit performs better than the other non-iterative circle fitting algorithms, including the original Riemann track fit. The improvement in general grows with increasing measurement uncertainties. Secondly, the modified Riemann fit is seen to be virtually as precise as the Kalman filter, the global linear least-squares estimator and the non-linear method for the entire range of measurement uncertainties.

A similar plot of the generalized variance, defined as the determinant of the sample covariance matrix, shows no visible difference from Fig. 1. From this we conclude that the bias of all estimators is negligible compared to their spread. For a general error and bias analysis of a wide range of circle fitting algorithms, see [10].

4. Error propagation

Under the assumption of normally distributed position errors, error propagation can be done analytically up to the calculation of the sample covariance matrix $\boldsymbol{A}$, after which point the computation of the required Jacobians is more easily done numerically. We start from the joint error matrix $\boldsymbol{V}$ of the vector of all measurements $\boldsymbol{m} = (\boldsymbol{u}\,;\boldsymbol{v})$.[2] $\boldsymbol{V}$ can also contain correlations between the measurements due to multiple scattering (see [9]).

In what follows, all vectors are column vectors. In addition, we use the following notations. If x, y are random variables, the expectation of x is denoted by $\mathsf{E}[x]$, the covariance of x and y is denoted by $\mathsf{cov}[x\,,y]$, and the variance of x is denoted by $\mathsf{var}[x] = \mathsf{cov}[x\,,x]$. If $\boldsymbol{x}, \boldsymbol{y}$ are random vectors, the expectation vector of $\boldsymbol{x}$ is denoted by $\mathsf{E}[\boldsymbol{x}]$, the cross-covariance matrix $\mathsf{Cov}[\boldsymbol{x}\,,\boldsymbol{y}]$ of $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined by

$$\mathsf{Cov}[\boldsymbol{x}\,,\boldsymbol{y}] = \mathsf{E}\left[\boldsymbol{x}\cdot\boldsymbol{y}^{\mathsf{T}}\right] - \mathsf{E}[\boldsymbol{x}]\cdot\mathsf{E}\left[\boldsymbol{y}^{\mathsf{T}}\right]$$

and the covariance matrix of $\boldsymbol{x}$ is denoted by $\mathsf{Var}[\boldsymbol{x}] = \mathsf{Cov}[\boldsymbol{x}\,,\boldsymbol{x}]$.

If $\boldsymbol{M}$ is a matrix, $\mathcal{S}[\boldsymbol{M}]$ denotes the sum of all elements of $\boldsymbol{M}$. If $\boldsymbol{M}_1, \boldsymbol{M}_2$ are two matrices of the same size, $\boldsymbol{M}_1 \odot \boldsymbol{M}_2$ denotes their element-wise product (Hadamard product). The identity matrix of dimension d is denoted by $\boldsymbol{I}_d$.

We now list the steps that have to be performed to get the covariance matrix of the estimated circle parameters.

A. Centering.

1. Center the measurements:

$$\boldsymbol{u}_{\mathrm{c}} = \boldsymbol{u} - \bar{u}\cdot\boldsymbol{e}, \quad \boldsymbol{v}_{\mathrm{c}} = \boldsymbol{v} - \bar{v}\cdot\boldsymbol{e}, \quad \boldsymbol{m}_{\mathrm{c}} = (\boldsymbol{u}_{\mathrm{c}}\,;\boldsymbol{v}_{\mathrm{c}})$$

   where $\boldsymbol{e} = (1,\ldots,1)^{\mathsf{T}}$ is a column vector of N ones.

[2] The semicolon (comma) denotes vertical (horizontal) concatenation of vectors or matrices.

2. As the centered measurements $\boldsymbol{m}_\mathrm{c}$ are centered again for the computation of $\boldsymbol{A}$ in step E, the joint error matrix $\boldsymbol{V}$ of $\boldsymbol{u}$ and $\boldsymbol{v}$ is not modified, i.e. $\boldsymbol{V}_\mathrm{c} = \boldsymbol{V}$.

B. Scaling.

1. Scale the measurements:

$$q = \boldsymbol{m}_\mathrm{c}^\mathsf{T} \cdot \boldsymbol{m}_\mathrm{c}, \quad Q = \sqrt{q/N}, \quad \boldsymbol{m}_\mathrm{cs} = \boldsymbol{m}_\mathrm{c} \cdot b/Q$$

2. Compute the variance of Q:

$$\boldsymbol{J}_Q = \frac{\partial Q}{\partial \boldsymbol{m}} = \frac{\boldsymbol{m}^\mathsf{T}}{Q \cdot N}, \quad \mathsf{var}[Q] = \boldsymbol{J}_Q \cdot \boldsymbol{V}_\mathrm{c} \cdot \boldsymbol{J}_Q^\mathsf{T}$$

3. Compute the scaled covariance matrix:

$$\boldsymbol{J}_1 = \frac{\partial \boldsymbol{m}_\mathrm{cs}}{\partial \boldsymbol{m}_\mathrm{c}} = b\left(\boldsymbol{I}_{2N}/\sqrt{q/N} - \boldsymbol{m}_\mathrm{c}\boldsymbol{m}_\mathrm{c}^\mathsf{T}/(Nq)^{3/2}\right)$$

$$\boldsymbol{V}_\mathrm{cs} = b^2 \cdot \boldsymbol{J}_1 \cdot \boldsymbol{V}_\mathrm{c} \cdot \boldsymbol{J}_1^\mathsf{T} = \begin{pmatrix} \boldsymbol{V}_{11} & \boldsymbol{V}_{12} \\ \boldsymbol{V}_{21} & \boldsymbol{V}_{22} \end{pmatrix}$$

C. Mapping to paraboloid. Compute $z_i = x_i^2 + y_i^2,\ i = 1, \ldots, N$ and the joint covariance matrix $\boldsymbol{C}$ of $\boldsymbol{r} = (\boldsymbol{x}\,;\boldsymbol{y}\,;\boldsymbol{z}) = (\boldsymbol{r}_1\,;\boldsymbol{r}_2\,;\boldsymbol{r}_3)$:

$$\boldsymbol{C} = \begin{pmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} & \boldsymbol{C}_{13} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} & \boldsymbol{C}_{23} \\ \boldsymbol{C}_{31} & \boldsymbol{C}_{32} & \boldsymbol{C}_{33} \end{pmatrix},$$

with

$$\begin{aligned}
\boldsymbol{C}_{ij} &= \boldsymbol{V}_{ij}, \quad i,j = 1,2 \\
\boldsymbol{C}_{13} &= 2\,\boldsymbol{V}_{11} \odot (\boldsymbol{e} \cdot \boldsymbol{r}_1^\mathsf{T}) + 2\,\boldsymbol{V}_{12} \odot (\boldsymbol{e} \cdot \boldsymbol{r}_2^\mathsf{T}), \quad \boldsymbol{C}_{31} = \boldsymbol{C}_{13}^\mathsf{T} \\
\boldsymbol{C}_{23} &= 2\,\boldsymbol{V}_{21} \odot (\boldsymbol{e} \cdot \boldsymbol{r}_1^\mathsf{T}) + 2\,\boldsymbol{V}_{22} \odot (\boldsymbol{e} \cdot \boldsymbol{r}_2^\mathsf{T}), \quad \boldsymbol{C}_{32} = \boldsymbol{C}_{23}^\mathsf{T} \\
\boldsymbol{C}_{33} &= \sum_{i=1,2} \sum_{j=1,2} 2\,\boldsymbol{V}_{ii} \odot \boldsymbol{V}_{ij} + 4\,\boldsymbol{V}_{ij} \odot (\boldsymbol{r}_i \cdot \boldsymbol{r}_j^\mathsf{T})
\end{aligned}$$

For the proof see Theorem 1 in the appendix. As is usual in error propagation, the unknown expectations of $\boldsymbol{r}_1 = \boldsymbol{x}$ and $\boldsymbol{r}_2 = \boldsymbol{y}$ are replaced by the observed values.

D. Compute center of gravity.

1. Reshape $\boldsymbol{r}$ as a $N \times 3$ matrix of the form $\boldsymbol{r} = (\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3)$ and compute

$$\boldsymbol{r}_0 = \boldsymbol{r}^\mathsf{T} \cdot \boldsymbol{w}$$

2. Compute the elements $C_{0,ij}$ of $\boldsymbol{C}_0 = \mathsf{Var}[\boldsymbol{r}_0]$:

$$C_{0,ij} = \boldsymbol{w}^\mathsf{T} \cdot \boldsymbol{C}_{ij} \cdot \boldsymbol{w},\ i, j = 1, 2, 3$$

E. Subtract center of gravity.

1. Compute the matrix $\boldsymbol{H}$:

$$\boldsymbol{H} = \boldsymbol{I}_N - \boldsymbol{e} \cdot \boldsymbol{w}^\mathsf{T}$$

2. Compute $\boldsymbol{s} = (\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3)$ and $\boldsymbol{D}_{ij} = \mathsf{Cov}[\boldsymbol{s}_i, \boldsymbol{s}_j]$:

$$\boldsymbol{s} = \boldsymbol{H} \cdot \boldsymbol{r}, \quad \boldsymbol{D}_{ij} = \boldsymbol{H} \cdot \boldsymbol{C}_{ij} \cdot \boldsymbol{H}^\mathsf{T},\ i, j = 1, 2, 3$$

F. Computation of $\boldsymbol{A}$*.*

1. As $\boldsymbol{A}$ is symmetric, it has only six independent elements A_α, where $\alpha = 1, \ldots, 6$ enumerates the chosen elements. A possible correspondence $A_\alpha \simeq A_{ij}$ is given by the following table:

α	1	2	3	4	5	6
$(i, j) = \nu(\alpha)$	$(1, 1)$	$(1, 2)$	$(1, 3)$	$(2, 2)$	$(2, 3)$	$(3, 3)$

2. For $\alpha = 1, \ldots, 6$ compute:

$$A_\alpha = \boldsymbol{s}_i^\mathsf{T} \cdot (\boldsymbol{w} \odot \boldsymbol{s}_j),\ \text{with } (i, j) = \nu(\alpha), \quad \alpha = 1, \ldots, 6$$

3. For all pairs (α, β) with $1 \leq \alpha \leq \beta \leq 6$ compute $(i,j) = \nu(\alpha)$ and $(k,l) = \nu(\beta)$. The covariance $\mathsf{cov}[A_\alpha, A_\beta] = E_{\alpha\beta} = E_{\beta\alpha}$ is given by:

$$\begin{aligned} E_{\alpha\beta} = &\mathcal{S}[\boldsymbol{D}_{ik} \odot \boldsymbol{W}_2 \odot \boldsymbol{D}_{jl} + \boldsymbol{D}_{il} \odot \boldsymbol{W}_2 \odot \boldsymbol{D}_{jk}] \\ &+ \boldsymbol{s}_i^\mathsf{T} \cdot (\boldsymbol{D}_{jl} \odot \boldsymbol{W}_2) \cdot \boldsymbol{s}_k + \boldsymbol{s}_i^\mathsf{T} \cdot (\boldsymbol{D}_{jk} \odot \boldsymbol{W}_2) \cdot \boldsymbol{s}_l \\ &+ \boldsymbol{s}_j^\mathsf{T} \cdot (\boldsymbol{D}_{il} \odot \boldsymbol{W}_2) \cdot \boldsymbol{s}_k + \boldsymbol{s}_j^\mathsf{T} \cdot (\boldsymbol{D}_{ik} \odot \boldsymbol{W}_2) \cdot \boldsymbol{s}_l, \end{aligned}$$

For the proof see Theorem 2 in the appendix. As in step C, the unknown expectations of $\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3$ are replaced by the observed values.

G. Computation of $\boldsymbol{n}$ *and* c*.*

1. Complete $\boldsymbol{A}$ to a full symmetric 3×3 matrix and determine the eigenvector $\boldsymbol{n}$ that corresponds to the smallest eigenvalue of $\boldsymbol{A}$.
2. Compute the Jacobian

$$\boldsymbol{J}_2 = \frac{\partial \boldsymbol{n}}{\partial A_\alpha}, \ \alpha = 1, \ldots, 6$$

by numerical differentiation and the covariance matrix of $\boldsymbol{n}$:

$$\boldsymbol{C}_{\boldsymbol{n}} = \mathsf{Var}[\boldsymbol{n}] = \boldsymbol{J}_2 \cdot \boldsymbol{E} \cdot \boldsymbol{J}_2^\mathsf{T}$$

3. Compute the signed distance c:

$$c = -\boldsymbol{n}^\mathsf{T} \cdot \boldsymbol{r}_0$$

4. Compute the joint covariance matrix of $\boldsymbol{n}$ and c (see [8]):

$$\boldsymbol{C}_{\boldsymbol{n},c} = \mathsf{Var}[(\boldsymbol{n}\,; c)] = \begin{pmatrix} \boldsymbol{C}_{\boldsymbol{n}} & -\boldsymbol{C}_{\boldsymbol{n}} \cdot \boldsymbol{r}_0 \\ -\boldsymbol{r}_0^\mathsf{T} \cdot \boldsymbol{C}_{\boldsymbol{n}} & \mathsf{var}[c] \end{pmatrix},$$

with

$$\mathsf{var}[c] = \boldsymbol{n}^\mathsf{T} \cdot \boldsymbol{C}_0 \cdot \boldsymbol{n} + \boldsymbol{r}_0^\mathsf{T} \cdot \boldsymbol{C}_{\boldsymbol{n}} \cdot \boldsymbol{r}_0 + \mathcal{S}[\boldsymbol{C}_{\boldsymbol{n}} \odot \boldsymbol{C}_0]$$

H. Computation of the circle parameters.

1. Compute center and radius of the circle:

$$u_0 = -\frac{n_1}{2n_3}, \quad v_0 = -\frac{n_2}{2n_3}, \quad \rho = \frac{\sqrt{1 - n_3^2 - 4cn_3}}{2n_3}$$

2. Compute the Jacobian $\boldsymbol{J}_3$:

$$\boldsymbol{J}_3 = \frac{\partial(u_0\,;v_0\,;\rho)}{\partial(\boldsymbol{n}\,;c)} = \begin{pmatrix} -\dfrac{1}{2n_3} & 0 & \dfrac{n_1}{2n_3^2} & 0 \\ 0 & -\dfrac{1}{2n_3} & \dfrac{n_2}{2n_3^2} & 0 \\ 0 & 0 & -\dfrac{h}{2n_3^2} - \dfrac{4c + 2n_3}{4hn_3} & -\dfrac{1}{h} \end{pmatrix}$$

with

$$h = \sqrt{1 - n_3^2 - 4cn_3}$$

3. Compute the covariance matrix of $\boldsymbol{p} = (u_0\,;v_0\,;\rho)$:

$$\boldsymbol{C}_{\boldsymbol{p}} = \mathsf{Var}[\boldsymbol{p}] = \boldsymbol{J}_3 \cdot \boldsymbol{C}_{\boldsymbol{n},c} \cdot \boldsymbol{J}_3^{\mathsf{T}}$$

I. Undo scaling and centering.

1. Rescale parameters:

$$\boldsymbol{p}' = \boldsymbol{p} \cdot Q/b$$

2. Error propagation:

$$\boldsymbol{C}_{\boldsymbol{p}'} = \mathsf{Var}[\boldsymbol{p}'] = Q^2/b^2 \cdot \boldsymbol{C}_{\boldsymbol{p}} + \boldsymbol{p} \cdot \boldsymbol{p}^{\mathsf{T}} \cdot \mathsf{var}[Q]\,/b^2$$

3. Undo centering:

$$u_0'' = u_0' + \bar{u}, \quad v_0'' = v_0 + \bar{v}, \quad \rho'' = \rho'$$

J. Transformation to final parameters. Note that u_0'', v_0'', and ρ'' are not at all normally distributed, and that in the limit of a straight line ρ'' tends to infinity. A more reasonable and numerically more stable track representation is for instance given by the parameters $\boldsymbol{q} = (a_0\,; \psi\,; \kappa)$, where a_0 is a signed impact parameter, ψ is the angle of inclination at the point of closest approach to the origin, and κ is a signed curvature of the circle (see also [9]). Their distribution is also much closer to a normal distribution.

The fit and the error propagation up to and including the transformation to the final parameters have been implemented in a MATLAB function that can be obtained from the authors on request.

5. Validation of the error propagation

We have validated the error propagation on the track sample described in Section 3, by analyzing the residuals of the estimated track parameters $\boldsymbol{q} = (a_0\,; \psi\,; \kappa)$ with respect to the true values $\boldsymbol{q}_{\mathrm{t}}$. If the covariance matrix of $\boldsymbol{q}$ is denoted by $\boldsymbol{C}_{\boldsymbol{q}}$, the three standard scores are defined by:

$$t_i = \frac{q_i - q_{\mathrm{t},i}}{\sqrt{\boldsymbol{C}_{\boldsymbol{q},ii}}}, \;\; i = 1, 2, 3$$

The χ^2-statistic c^2 is defined by:

$$c^2 = (\boldsymbol{q} - \boldsymbol{q}_{\mathrm{t}})^{\mathsf{T}} \boldsymbol{C}_{\boldsymbol{q}}^{-1} (\boldsymbol{q} - \boldsymbol{q}_{\mathrm{t}})$$

Its p-value P is obtained by integrating the χ^2-density with three degrees of freedom from zero to c^2. With the correct error propagation, the standard scores follow, at least in good approximation, a standard normal distribution with mean zero and standard deviation one, while the p-value P if approximately

Table 1: Mean and standard deviation of p-values and standard scores versus the measurement uncertainty σ. The bottom row, marked by an asterisk, shows the result with measurement errors from the mixture model (*) and an effective uncertainty $\sigma_{\text{eff}} = 1.07\,\text{mm}$.

σ[mm]	1	2	3	4	5	6	7	8
0.1	0.50	0.29	0.01	−0.02	−0.01	0.99	1.00	1.02
0.2	0.50	0.29	0.02	−0.02	−0.02	0.99	1.01	1.01
0.3	0.50	0.29	0.02	−0.02	−0.02	0.99	1.00	1.01
0.4	0.50	0.29	0.02	−0.02	−0.02	0.99	1.01	1.01
0.5	0.50	0.29	0.02	−0.02	−0.02	0.99	1.00	1.01
0.6	0.51	0.29	0.02	−0.02	−0.02	0.99	1.01	1.04
0.7	0.51	0.29	0.02	−0.02	−0.02	0.99	1.01	1.02
0.8	0.51	0.29	0.01	−0.02	−0.02	0.99	1.01	1.00
0.9	0.51	0.29	0.01	−0.02	−0.01	0.99	1.01	1.03
1.0	0.51	0.29	0.01	−0.02	−0.02	0.99	1.01	1.01
1.1	0.51	0.29	0.02	−0.02	−0.02	0.99	1.01	1.02
1.2	0.51	0.30	0.02	−0.02	−0.02	0.99	1.00	1.01
1.3	0.51	0.30	0.01	−0.02	−0.02	0.99	1.00	1.04
1.4	0.51	0.30	0.01	−0.02	−0.02	0.98	0.99	1.00
1.5	0.51	0.30	0.01	−0.02	−0.02	0.96	0.98	0.98
1.0*	0.50	0.30	0.01	−0.01	−0.01	1.00	1.02	1.02

Columns 1–2: mean and standard deviation of p-values
Columns 3–5: mean of standard scores
Columns 6–8: standard deviation of standard scores

uniformly distributed with mean 0.5 and standard deviation $1/\sqrt{12} \approx 0.289$. Table 1 shows the sample mean and the sample standard deviation of P and of the standard scores $t_i,\ i = 1, 2, 3$. The results are very close to the expected values. It can be observed, however, that the standard scores show a persistent but negligible bias of 1–2% of the standard deviation, which is excellent for a highly non-linear estimator.

The empirical distribution of the standard scores can be compared to a standard normal distribution by means of a quantile-quantile (Q-Q) plot. Figure 2 shows the three Q-Q plots of the standard scores and a histogram of the p-

values. The Q-Q plots show a remarkably good agreement with the normal distribution, and only a few outliers can be observed. On the other hand, in the histogram of the p-values a small excess of about 70 p-values close to 1 can be observed, corresponding to abnormally large values of the χ^2 statistic c^2. These values can be traced back to cases in which the covariance matrix $\boldsymbol{C_q}$ has a very large condition number, which leads to numerical instability in the inversion and to wrong correlations. The fraction of such cases rises from about 0.1% at $\sigma = 0.1\,\text{mm}$ to about 1% at $\sigma = 1.5\,\text{mm}$.

Finally, we have checked the sensitivity of the error propagation to the assumption of normal measurement errors. To this end, we have simulated measurement errors from the following Gaussian mixture:

$$f(x) = 0.95 \cdot \varphi(x\,; 0, \sigma) + 0.05 \cdot \varphi(x\,; 0, 2\sigma) \tag{*}$$

where $\varphi(x\,; 0, \sigma)$ ist the normal density with mean zero and standard deviation σ. The effective standard deviation of the mixture is $\sigma_{\text{eff}} \approx 1.07\,\sigma$. If the Riemann fit is performed with σ_{eff}, the average properties of the standard scores and the p-values are correct, as demonstrated by the bottom row in Table 1. The distribution of the standard scores is very similar to a standard normal, with the exception of the extreme tails. The distribution of the p-value of the χ^2-statistic shows the expected slight U-shape. This is illustrated by Figure 3 which shows the Q-Q plots of the standard scores and a histogram of the p-values with $\sigma = 1\,\text{mm}$ and $\sigma_{\text{eff}} = 1.07\,\text{mm}$, using the mixture model defined above.

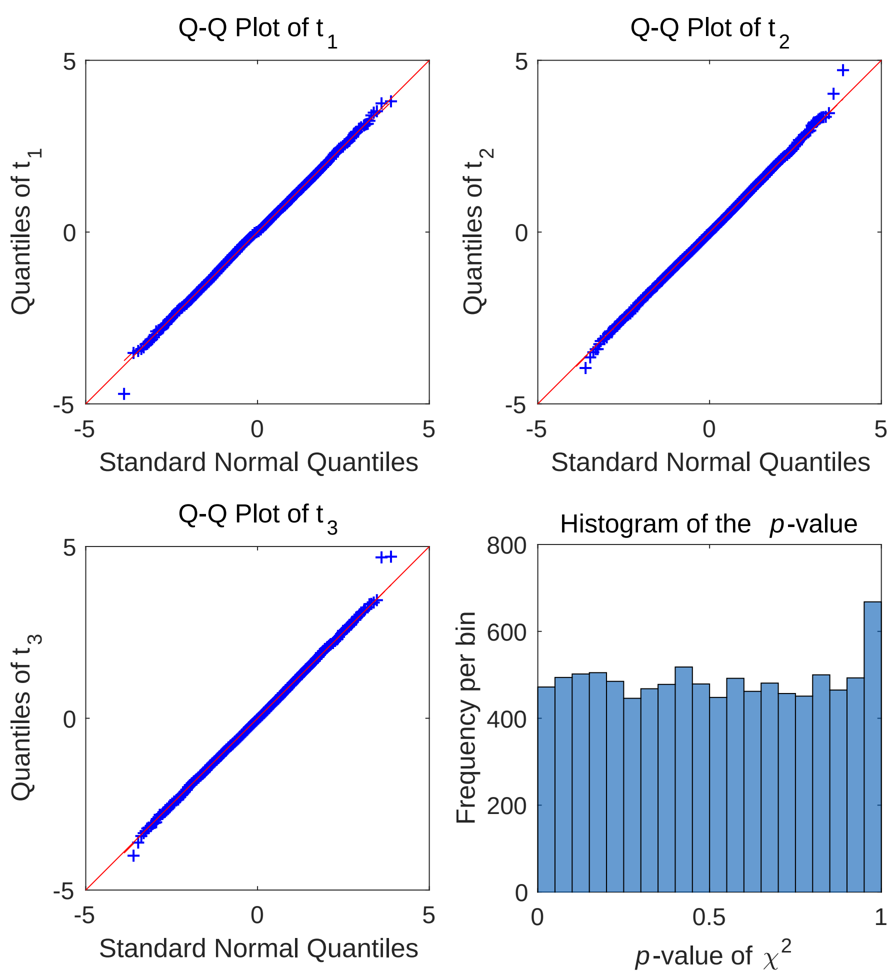

Figure 2: Q-Q plots of the standard scores and histogram of the p-values at the measurement uncertainty $\sigma = 1\,\mathrm{mm}$.

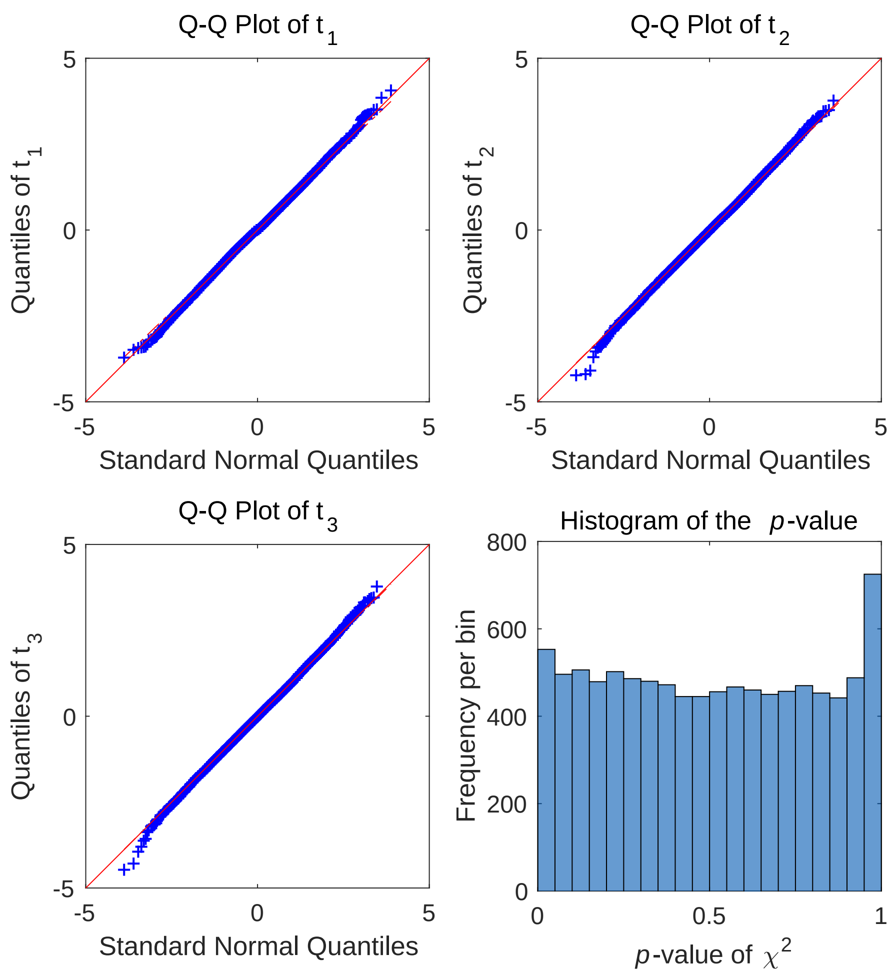

Figure 3: Q-Q plots of the standard scores and histogram of the p-values at the measurement uncertainty $\sigma = 1\,\mathrm{mm}$, with measurement errors generated according to the mixture model defined in the text.

6. Summary and conclusions

We have in this paper further explored the properties of a modified Riemann track fit which operates on translated and scaled measurements [7], making the fit invariant under translations and similarity transforms of the measurements. With these transformations, the fit becomes more precise than other, popular non-iterative track fitting approaches, in particular for large measurement uncertainties. In addition, the modified Riemann fit is demonstrated to be equally precise as more rigorous approaches such as the Kalman filter, a global, linear least-squares method and a non-linear, iterative approach based on the Levenberg-Marquardt algorithm, at least when material and detector effects such as multiple Coulomb scattering, energy loss and misalignment can be neglected. The complete error propagation from the measurements to the final estimated circle parameters has been computed and validated for this case.

Alignment uncertainty can be incorporated into the joint covariance matrix of the position measurements. Multiple Coulomb scattering can be treated in a manner similar to [9]. As the error propagation derived in Section 4 does not depend on the structure of the initial covariance matrix, it will not be affected. Non-negligible energy loss, however, is a different matter as it destroys the circular track model. As a consequence, the Riemann track fit will produce biased estimates. The validation and performance of the Riemann fit under realistic assumptions on the detector material and the sensor misalignment as well as a possible bias correction will be the subject of a subsequent study.

It has also been shown that the fit is robust against small deviations from the assumed normal distribution of the measurement errors. Future developments will concentrate on making the Riemann fit robust against more severe deviations from the normal assumption and additional background observations.

Appendix A Derivation of covariance matrices

Lemma 1. Let $u_i,\ i = 1,\dots,4$ be four correlated standard normal random variables u_i with zero mean, unit variance and correlation coefficients $\rho_{ij},\ i,j = 1,\dots,4$. The products of two, three or four of the variables u_i have the following expectations:

$$
\begin{aligned}
\mathsf{E}[u_i u_j] &= \rho_{ij} \\
\mathsf{E}[u_i u_j u_k] &= 0 \\
\mathsf{E}[u_i u_j u_k u_l] &= \rho_{ij}\rho_{kl} + \rho_{ik}\rho_{jl} + \rho_{il}\rho_{jk}
\end{aligned}
$$

Proof. See [11]. □

Lemma 2. Let $x_i,\ i = 1,\dots,4$ be four normal random variables with means μ_i, variances σ_i^2 and covariances $c_{ij},\ i,j,= 1,\dots,4$. The products of two, three or four of the variables x_i have the following expected values:

$$
\begin{aligned}
\mathsf{E}[x_i x_j] &= c_{ij} + \mu_i\mu_j \\
\mathsf{E}[x_i x_j x_k] &= \mu_i\mu_j\mu_k + \mu_i c_{jk} + \mu_j c_{ik} + \mu_k c_{ij} \\
\mathsf{E}[x_i x_j x_k x_l] &= c_{ij}c_{kl} + c_{ik}c_{jl} + c_{il}c_{jk} + \mu_i\mu_j c_{kl} + \mu_i\mu_k c_{jl} \\
&\quad + \mu_i\mu_l c_{jk} + \mu_j\mu_k c_{il} + \mu_j\mu_l c_{ik} + \mu_k\mu_l c_{ij} + \mu_i\mu_j\mu_k\mu_l
\end{aligned}
$$

Proof. Each of the x_i can be written in the form $x_i = u_i\sigma_i + \mu_i$, where the u_i are standard normal with correlations $\rho_{ij} = c_{ij}/(\sigma_i\sigma_j)$. Expansion of the products and application of Lemma 1 gives the desired expressions. □

Lemma 3. Under the assumptions of Lemma 2, the following statements hold:

$$
\begin{aligned}
\mathsf{cov}\left[x_i, x_j^2\right] &= 2\mu_j c_{ij} \\
\mathsf{cov}\left[x_i^2, x_j^2\right] &= 2c_{ij}^2 + 4\mu_i\mu_j c_{ij} \\
\mathsf{cov}\left[x_i^2, x_i^2 + x_j^2\right] &= 2\sigma_i^4 + 4\mu_i^2\sigma_i^2 + 2c_{ij}^2 + 4\mu_i\mu_j c_{ij} \\
\mathsf{var}\left[x_i^2 + x_j^2\right] &= 2\sigma_i^4 + 4\mu_i^2\sigma_i^2 + 2\sigma_j^4 + 4\mu_j^2\sigma_j^2 + 4c_{ij}^2 + 8\mu_i\mu_j c_{ij}
\end{aligned}
$$

Proof. The proofs follow from special cases of Lemma 2 and are omitted. □

Theorem 1. Let $\boldsymbol{x}_1, \boldsymbol{x}_2$ be two normal random vectors of dimension $N \times 1$ with mean vectors $\boldsymbol{\mu}_i$, covariance matrices $\boldsymbol{C}_{11}$ and $\boldsymbol{C}_{22}$, and cross-covariance matrices $\boldsymbol{C}_{12}$ and $\boldsymbol{C}_{21} = \boldsymbol{C}_{12}^{\mathsf{T}}$. Let $z_k = x_{1,k}^2 + x_{2,k}^2,\ k = 1, \ldots, N$. Then $\boldsymbol{C}_{33} = \mathsf{Var}[\boldsymbol{z}]$ and $\boldsymbol{C}_{i3} = \mathsf{Cov}[\boldsymbol{x}_i\,, \boldsymbol{z}]\,,\ i = 1, 2$ are given by:

$$
\begin{aligned}
\boldsymbol{C}_{33} &= \sum_{i=1}^{2}\sum_{j=1}^{2} 2 \cdot \boldsymbol{C}_{ij} \odot \boldsymbol{C}_{ij} + 4 \cdot \boldsymbol{C}_{ij} \odot (\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j^{\mathsf{T}}) \\
\boldsymbol{C}_{i3} &= 2 \cdot \boldsymbol{C}_{i1} \odot (\boldsymbol{e} \cdot \boldsymbol{\mu}_1^{\mathsf{T}}) + 2 \cdot \boldsymbol{C}_{i2} \odot (\boldsymbol{e} \cdot \boldsymbol{\mu}_2^{\mathsf{T}})
\end{aligned}
$$

Proof. Lemma 3 implies for $m, n = 1, \ldots, N$:

$$
\begin{aligned}
C_{33,mn} &= \mathsf{cov}\left[x_{1,m}^2 + x_{2,m}^2\,, x_{1,n}^2 + x_{2,n}^2\right] \\
&= \mathsf{cov}\left[x_{1,m}^2, x_{1,n}^2\right] + \mathsf{cov}\left[x_{1,m}^2\,, x_{2,n}^2\right] + \mathsf{cov}\left[x_{2,m}^2\,, x_{1,n}^2\right] + \mathsf{cov}\left[x_{2,m}^2\,, x_{2,n}^2\right] \\
&= 2\,C_{11,mn}^2 + 4\,\mu_{1,m}\,\mu_{1,n}\,C_{11,mn} + 2\,C_{12,mn}^2 + 4\,\mu_{1,m}\,\mu_{2,n}\,C_{12,mn} \\
&\quad + 2\,C_{21,mn}^2 + 4\,\mu_{2,m}\,\mu_{1,n}\,C_{21,mn} + 2\,C_{22,mn}^2 + 4\,\mu_{2,m}\,\mu_{2,n}\,C_{22,mn} \\
C_{i3,mn} &= \mathsf{cov}\left[x_{i,m}\,, x_{1,n}^2 + x_{2,n}^2\right] \\
&= \mathsf{cov}\left[x_{i,m}\,, x_{1,n}^2\right] + \mathsf{cov}\left[x_{i,m}\,, x_{2,n}^2\right] \\
&= 2\,\mu_{1,n}\,C_{i1,mn} + 2\,\mu_{2,n}\,C_{i2,mn}
\end{aligned}
$$

Note that $\boldsymbol{C}_{ij} \odot (\boldsymbol{e} \cdot \boldsymbol{\mu}_i^{\mathsf{T}})$ needs fewer multiplications than the equivalent expression $\boldsymbol{C}_{ij} \cdot \mathsf{diag}(\boldsymbol{\mu}_i)$, and that $\boldsymbol{C}_{ij} \odot (\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j^{\mathsf{T}})$ needs fewer multiplications than $\mathsf{diag}(\boldsymbol{\mu}_i) \cdot \boldsymbol{C}_{ij} \cdot \mathsf{diag}(\boldsymbol{\mu}_j)$. □

Theorem 2. Let $\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3$ be three normal random vectors of dimension $N \times 1$ with mean vectors $\boldsymbol{\mu}_i$, covariance matrices $\boldsymbol{D}_{ii}$ and cross-covariance matrices $\boldsymbol{D}_{ij},\ i, j = 1, \ldots, 3$. Let $\boldsymbol{w}$ be a $N \times 1$ column vector of fixed weights and

$\boldsymbol{W}_2 = \boldsymbol{w} \cdot \boldsymbol{w}^\mathsf{T}$. Let A_α and A_β be defined by:

$$A_\alpha = \boldsymbol{s}_i^\mathsf{T} \cdot (\boldsymbol{w} \odot \boldsymbol{s}_j)$$
$$A_\beta = \boldsymbol{s}_k^\mathsf{T} \cdot (\boldsymbol{w} \odot \boldsymbol{s}_l)$$

where $1 \leq i, j, k, l \leq 3$. Then the covariance $\mathsf{cov}[A_\alpha, A_\beta]$ is equal to:

$$\begin{aligned} \mathsf{cov}[A_\alpha, A_\beta] = &\mathcal{S}[\boldsymbol{D}_{ik} \odot \boldsymbol{W}_2 \odot \boldsymbol{D}_{jl} + \boldsymbol{D}_{il} \odot \boldsymbol{W}_2 \odot \boldsymbol{D}_{jk}] \\ &+ \boldsymbol{\mu}_i^\mathsf{T} \cdot (\boldsymbol{D}_{jl} \odot \boldsymbol{W}_2) \cdot \boldsymbol{\mu}_k + \boldsymbol{\mu}_i^\mathsf{T} \cdot (\boldsymbol{D}_{jk} \odot \boldsymbol{W}_2) \cdot \boldsymbol{\mu}_l \\ &+ \boldsymbol{\mu}_j^\mathsf{T} \cdot (\boldsymbol{D}_{il} \odot \boldsymbol{W}_2) \cdot \boldsymbol{\mu}_k + \boldsymbol{\mu}_j^\mathsf{T} \cdot (\boldsymbol{D}_{ik} \odot \boldsymbol{W}_2) \cdot \boldsymbol{\mu}_l \end{aligned}$$

Proof. By definition, $\mathsf{cov}[A_\alpha, A_\beta] = \mathsf{E}[A_\alpha A_\beta] - \mathsf{E}[A_\alpha]\,\mathsf{E}[A_\alpha]$. Then we have:

$$\begin{aligned} \mathsf{E}[A_\alpha] &= \sum_m w_m\,\mathsf{E}[s_{i,m}\, s_{j,m}] = \sum_m w_m (D_{ij,mm} + \mu_{i,m}\,\mu_{j,m}) \\ \mathsf{E}[A_\beta] &= \sum_n w_n\,\mathsf{E}[s_{k,n}\, s_{l,n}] = \sum_n w_n (D_{kl,nn} + \mu_{k,n}\,\mu_{l,n}) \\ \mathsf{E}[A_\alpha\, A_\beta] &= \sum_m \sum_n w_m w_n\,\mathsf{E}[s_{i,m}\, s_{j,m}\, s_{k,n}\, s_{l,n}] \end{aligned}$$

where the sums over m and n run from 1 to N. Using Lemma 2 it is straightforward to show that:

$$\begin{aligned} \mathsf{E}[s_{i,m}\, s_{j,m}\, s_{k,n}\, s_{l,n}] = &D_{ij,mm}\, D_{kl,nn} + D_{ik,mn}\, D_{jl,mn} + D_{il,mn}\, D_{jk,mn} \\ &+ \mu_{i,m}\,\mu_{j,m}\, D_{kl,nn} + \mu_{i,m}\,\mu_{k,n}\, D_{jl,mn} + \mu_{i,m}\,\mu_{l,n}\, D_{jk,mn} \\ &+ \mu_{j,m}\,\mu_{k,n}\, D_{il,mn} + \mu_{j,m}\,\mu_{l,n}\, D_{ik,mn} + \mu_{k,n}\,\mu_{l,n}\, D_{ij,mm} \\ &+ \mu_{i,m}\,\mu_{j,m}\,\mu_{k,n}\,\mu_{l,n} \end{aligned}$$

From this follows:

$$\mathsf{cov}[A_\alpha, A_\beta] = \sum_m \sum_n w_m\, w_n B_{mn}$$

with

$$\begin{aligned}B_{mn} = &D_{ik,mn}\, D_{jl,mn} + D_{il,mn}\, D_{jk,mn}\\ &+ \mu_{i,m}\, \mu_{k,n}\, D_{jl,mn} + \mu_{i,m}\, \mu_{l,n}\, D_{jk,mn}\\ &+ \mu_{j,m}\, \mu_{k,n}\, D_{il,mn} + \mu_{j,m}\, \mu_{l,n}\, D_{ik,mn}\end{aligned}$$

Summing over m and n gives the desired result. □

References

[1] Frühwirth R 1987 *Nucl. Instrum. Methods Phys. Res.* A **262** 444

[2] Frühwirth R et al. 2000 *Data Analysis Techniques for High-Energy Physics* (Cambridge University Press, Cambridge, UK)

[3] Hansroul M, Jeremie H, and Savard D 1988 *Nucl. Instrum. Methods Phys. Res.* A **270** 498

[4] Karimäki V 1991 *Nucl. Instrum. Methods Phys. Res.* A **305** 187

[5] Strandlie A, Wroldsen J, Frühwirth R and Lillekjendlie B 2000 *Comp. Phys. Commun.* **131** 95

[6] Frühwirth R, and Strandlie A 2016 *J. Phys.: Conf. Ser.* **762** 012032

[7] Chernov N 2010 *Circular and linear regression: Fitting circles and lines by least squares* (Boca Raton: CRC Press)

[8] Strandlie A, and Frühwirth R 2002 *Nucl. Instrum. Methods Phys. Res.* A **480** 734

[9] Strandlie A, Wroldsen J, and Frühwirth R 2002 *Nucl. Instrum. Methods Phys. Res.* A **488** 332

[10] Al-Sharadqah A, and Chernov, N 2009 *Electronic Journal of Statistics* **3** 886

[11] Haldane J B S 1942, *Biometrika* **32** No. 3/4 226