Proceedings of CTD 2020
PROC-CTD2020-02
June 29, 2020

Global least squares alignment with Kalman Filter fitted tracks

Pawel Brückman de Renstrom

Institute of Nuclear Physics
Polish Academy of Sciences
ul. Radzikowskiego 152
31-342 Krakow, Poland

ABSTRACT

The Kalman Filter approach to fitting charged particle trajectories is widespread in modern complex tracking systems. At the same time, the global fit of the detector geometry using Newton-Raphson fitted tracks remains the baseline method to achieve efficient and reliable track-based alignment which is free from weak-mode biases affecting physics measurements. A brief reminder of the global least squares formalism for track-based alignment and how Kalman Filter fitted tracks can be equivalently used for the global fit as well as potential computational benefits and use of additional constraints are reviewed.

PRESENTED AT

Connecting the Dots Workshop (CTD 2020)
April 20-30, 2020

1 Introduction

Contemporary high energy physics experiments are equipped with high precision semiconductor tracking systems of high complexity. In order to exploit their full potential an accurate determination of their geometry is indispensable. This is usually achieved using so-called track-based alignment, whereby reconstructed tracks are used as survey (gauge) tools for the detector elements.

The Global χ^2 fit of the detector geometry using *least squares*-fitted (LS) tracks remains the baseline method to achieve efficient and reliable track-based alignment which is free from weak-mode biases affecting physics measurements. However, the *Kalman Filter* (KF) approach to fitting charged particle trajectories is widespread in modern complex tracking systems. The natural question arises: Can KF-fitted tracks be used for the the Global χ^2 alignment, and at what cost? In the following we recall the basic formalism and illustrate it with some very simplistic examples in order to gain better insight into the mechanism, in hope of adding a pedagogical value.

2 Track models

There is a number of different fitting methods and corresponding track models. They generally split into global, usually iterative, fit approaches and sequential update ones. From physics point of view, the only relevant information returned by the fit are track parameters and their covariance at its origin, usually the point of closest approach to the production vertex. In case of tracks used for alignment, however, these are rather track parameters at subsequent measurement planes together with their full covariance, notably the off-diagonal elements representing correlations between measurements in different tracking elements. An alignment fit additionally uses derivatives of track residuals* w.r.t. the track parameters which may take very different form depending on the track model.

2.1 The Least Squares track model

In the LS global approach track fit is based on the minimization of the χ^2 defined as:

$$\chi^2 = \boldsymbol{\rho}^T V^{-1} \boldsymbol{\rho}, \quad \text{with} \quad \boldsymbol{\rho} \equiv \begin{pmatrix} \boldsymbol{r} \\ \boldsymbol{\theta} \end{pmatrix}, \quad V \equiv \begin{pmatrix} \Omega & 0 \\ 0 & \Theta \end{pmatrix}, \tag{1}$$

where $\boldsymbol{r}$ represents a vector of residuals of measurements associated along the trajectory and $\boldsymbol{\theta}$ is a vector of trajectory deflection angles due to multiple Coulomb scattering (MCS) on the subsequent measurement planes. $\boldsymbol{V}$ is an explicitly diagonal matrix representing measurement uncertainties and MCS standard deviations, hence is of size $n_{\mathrm{r}} + 2n$, with n being the number of measurement planes and n_{r} the number of measurements† Tracks are parameterized by a single set of global parameters and two deflection angles per material surface (≈ measurement plane):

$$\boldsymbol{\pi} \equiv \begin{pmatrix} \boldsymbol{\tau} \\ \boldsymbol{\theta} \end{pmatrix}, \quad \boldsymbol{\tau} = (d_0, z_0, \phi, \theta, Q/p) \tag{2}$$

corresponding to a vector of $5 + 2n$ parameters.

Track parameter covariance matrix is given by:

$$C = \mathrm{Cov}(\boldsymbol{\pi}) = \left(H^T V^{-1} H\right) \quad \text{with} \quad H \equiv \frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\pi}}. \tag{3}$$

The crucial matrix H of size $(n_{\mathrm{r}} + 2n) \times (5 + 2n)$ is dense and generally highly nontrivial.

*A residual is defined as a distance between the position of the signal cluster and the track intersection with the sensitive detector element, usually measured in its local reference frame.

†It may, in general, be larger than n if a single plane provides more than one measurement, e,g, pixel detectors.

2.2 The Kalman Filter track model

The Kalman Filter track fitting is based on sequential update of the state vector during the forward *filtering* and backward *smooting* processes [1]. A track is represented by n state vectors and their corresponding covariance matrices:

$$\boldsymbol{x}_k \equiv (l_1, l_2, \phi, \theta, Q/p)_k, \quad C_k \tag{4}$$

resulting in $5n$ track parameters:

$$\boldsymbol{\pi} \equiv (\boldsymbol{x}_1, ..., \boldsymbol{x}_n) \, . \tag{5}$$

Residuals and their covariance are reduced to the measurement ones only:

$$\boldsymbol{\rho} \equiv \boldsymbol{r} = \boldsymbol{m} - \boldsymbol{h}(\boldsymbol{x}), \quad \boldsymbol{V} \equiv \Omega, \tag{6}$$

while the scattering Θ contributes as the process random noise $\boldsymbol{w}$ in the update of the state vector:

$$\boldsymbol{x}_{k+1} = F_k \boldsymbol{x}_k + \boldsymbol{w}_k, \qquad C^k_{k+1} = F_k C^k_k F^{\mathrm{T}}_k + Q_k, \tag{7}$$

where F_k is the state propagator and $Q_k = \mathrm{Cov}(\boldsymbol{w}_k)$. The derivatives $H \equiv \frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\pi}}$, although given by a large $(n_{\mathrm{r}} + 2n) \times 5n$ matrix are usually trivial and sparse. The full covariance matrix of the track parameters appears large $(5n \times 5n)$:

$$C = \mathrm{Cov}(\boldsymbol{\pi}) = \begin{pmatrix} C_1 & C_{1,2} & \dots & C_{1,n} \\ C_{2,1} & C_2 & \dots & C_{2,n} \\ \dots & \dots & \dots & \dots \\ C_{n,1} & C_{n,2} & \dots & C_n \end{pmatrix} \tag{8}$$

and is never fully constructed in the normal KF process. Only the diagonal sub-matrices C_k (5×5) are obtained in:

$$\begin{array}{lll} \text{filtering :} & C^k_k = (\mathbf{1} - K_k H_k)\, C^{k-1}_k, & K_k = C^{k-1}_k H^T_k \left(V_k + H_k C^{k-1}_k H^T_k\right)^{-1} \\ \text{smoothing :} & C_k = C^k_k + S_k (C_{k+1} - C^k_{k+1}) S^T_k, & S_k = (F_k)^{-1} (C^k_{k+1} - Q_k)(C^k_{k+1})^{-1}, \end{array} \tag{9}$$

where, K_k is the Kalman *gain matrix* while S_k is the commonly called the *smoother gain matrix*. The smoothing step is necessary to propagate the full information about all measurements on the track back to its origin.

3 The Global χ^2 alignment

The track-based alignment of the tracking system relies on the minimization of the χ^2 summed over all considered tracks [2]:

$$\chi^2_{\mathrm{global}} = \sum_i \chi^2_i, \qquad \frac{\mathrm{d}\chi^2_{\mathrm{global}}}{\mathrm{d}\boldsymbol{\alpha}} = 0, \qquad \boldsymbol{\rho} \equiv \boldsymbol{\rho}(\boldsymbol{\pi}(\boldsymbol{\alpha}), \boldsymbol{\alpha}), \tag{10}$$

where both the the fitted tracks and the position detector measurements explicitly depend on alignment parameters $\boldsymbol{\alpha}$. The full derivative with respect to the alignment parameters is then given by:

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\alpha}} = \frac{\partial}{\partial \boldsymbol{\alpha}} + \frac{\mathrm{d}\boldsymbol{\pi}}{\mathrm{d}\boldsymbol{\alpha}} \frac{\partial}{\partial \boldsymbol{\pi}} = \frac{\partial}{\partial \boldsymbol{\alpha}} - A^T V^{-1} H C \frac{\partial}{\partial \boldsymbol{\pi}}, \quad \text{where} \quad A \equiv \frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\alpha}}. \tag{11}$$

In the limit of small corrections, the χ^2_{global} minimization problem can be solved using a linear expansion[‡]:

$$\Delta\alpha = -\underbrace{\left(\mathcal{M}^{-1}\right)}_{\mathrm{Cov}(\alpha)} \mathcal{R} \tag{12}$$

[‡]In practice this condition is sufficiently well satisfied. Nonetheless, alignment usually allows for a few iterations in order to mitigate residual nonlinearities.

with

$$\begin{aligned} \mathcal{R} = \frac{1}{2}\frac{\mathrm{d}\chi^2}{\mathrm{d}\boldsymbol{\alpha}} &= \sum_{\text{tracks}} A^T V^{-1} R V^{-1} \boldsymbol{\rho}, \\ \mathcal{M} = \frac{1}{2}\frac{\mathrm{d}^2\chi^2}{\mathrm{d}\boldsymbol{\alpha}^2} &= \sum_{\text{tracks}} A^T V^{-1} R V^{-1} A. \end{aligned} \tag{13}$$

$R = V - HCH^T$ is the full covariance of the residuals of the track fit, the cornerstone of the alignment principle. Any track model that is capable of returning this information is suitable to provide input for alignment.

4 The Kalman Filter tracks for alignment

In order to use KF-fitted tracks for alignment one needs to recover the off-diagonal elements of the C matrix. It has been shown that these can be obtained recursively using the already known smoother gain matrix [3]:

$$C_{k,l} = S_k C_{k+1,l}. \tag{14}$$

The procedure requires $125n(n-1)$ multiplications, provided S_k are known from the track fit. Calculation of $R = V - HCH^T$ appears faster for the KF than LS tracks thanks to the sparse nature of H.

Constraints on KF track parameters can be introduced just as in the LS fit [2] by adding pseudo-measurements in the filtering. It can be thought of as providing a meaningful prior.

If several tracks have been fitted to a common vertex, the covariance of each track at the vertex position $C_0^{(i)}$ gets modified by the additional measurement [4]. This can be propagated forward to the states at subsequent measurement planes to yield their full covariance:

$$\tilde{C}_{k,l}^{(i)} = C_{k,l}^{(i)} + C_{k,0}^{(i)} \left(C_0^{(i)}\right)^{-1} (\tilde{C}_0^{(i)} - C_0^{(i)}) \left(C_0^{(i)}\right)^{-1} C_{0,l}^{(i)} \tag{15}$$

Even more importantly, the vertex introduces precious correlations between any two states in any two tracks:

$$\tilde{C}_{k,l}^{(i,j)} = \mathbf{0} + C_{k,0}^{(i)} \left(C_0^{(i)}\right)^{-1} (\tilde{C}_0^{(i,j)} - \mathbf{0}) \left(C_0^{(j)}\right)^{-1} C_{0,l}^{(j)} \tag{16}$$

All above additional constraints allow to mitigate weak modes of alignment which lead to systematic deformations and can bias reconstructed track parameters.

5 Super-simple examples

In the following we illustrate solutions to alignment problem using the simplest possible setups and ignoring process noise (MCS) which in real case always has to be included.

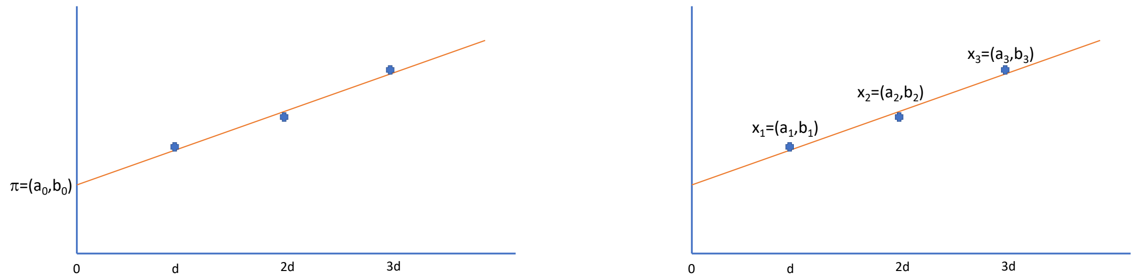
5.1 LS track fit

Let us consider a LS fit of a straight line through three equidistant vertical planes providing measurements with a common uncertainty σ, as schematically illustrated in Fig.1a. The measurement covariance, derivative matrix and the resulting covariance matrix of the fitted parameters $\boldsymbol{\pi} = (a_0, b_0)$ are given by:

$$V = \begin{pmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix}, \quad H = \frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\pi}} = \begin{pmatrix} d & 1 \\ 2d & 1 \\ 3d & 1 \end{pmatrix}, \quad C = (H^T V^{-1} H)^{-1} = \begin{pmatrix} \frac{\sigma^2}{2d^2} & \frac{-\sigma^2}{d} \\ \frac{-\sigma^2}{d} & \frac{7\sigma^2}{3} \end{pmatrix}$$

Hence, we get the full covariance of the fitted residuals:

$$R = V - HCH^T = \begin{pmatrix} \frac{\sigma^2}{6} & \frac{-2\sigma^2}{6} & \frac{\sigma^2}{6} \\ \frac{-2\sigma^2}{6} & \frac{4/\sigma^2}{6} & \frac{-2\sigma^2}{6} \\ \frac{\sigma^2}{6} & \frac{-2\sigma^2}{6} & \frac{\sigma^2}{6} \end{pmatrix}.$$



(a) (b)

Figure 1: A fit of a straight line through three equidistant measurements (in the vertical direction) with common uncertainties. (a) the LS fit, (b) the KF fit.

5.2 KF track fit

Now, let us perform a KF fit through the same three equidistant measurements with a common uncertainty σ, as illustrated in Fig.1b. The derivative matrix takes a particularly simple, sparse form as track parameters are defined in the local reference frame of the measurement plane. The track parameter[§] covariance matrix is obtained by the filtering and smoothing to get the block-diagonal elements, as shown black in Eq. 9. The remaining off-diagonal ones (in blue) are recovered recursively using Eq. 14:

$$H = \frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\pi}} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} \frac{\sigma^2}{2d^2} & \frac{-\sigma^2}{2d} & \frac{\sigma^2}{2d^2} & \frac{-\sigma^2}{2d} & \frac{\sigma^2}{2d^2} & \frac{-\sigma^2}{2d} \\ \frac{-\sigma^2}{2d} & \frac{5\sigma^2}{6} & 0 & \frac{\sigma^2}{3} & \frac{\sigma^2}{2d} & \frac{\sigma^2}{6} \\ \frac{\sigma^2}{2d^2} & 0 & \frac{\sigma^2}{2d^2} & 0 & \frac{\sigma^2}{2d^2} & 0 \\ \frac{-\sigma^2}{2d} & \frac{\sigma^2}{3} & 0 & \frac{\sigma^2}{3} & \frac{\sigma^2}{2d} & \frac{\sigma^2}{3} \\ \frac{\sigma^2}{2d^2} & \frac{\sigma^2}{2d} & \frac{\sigma^2}{2d^2} & \frac{\sigma^2}{2d} & \frac{\sigma^2}{2d^2} & \frac{\sigma^2}{2d} \\ \frac{-\sigma^2}{2d} & \frac{\sigma^2}{6} & 0 & \frac{\sigma^2}{3} & \frac{\sigma^2}{2d} & \frac{5\sigma^2}{6} \end{pmatrix}$$

Despite very different form of the derivatives and track covariance, the resulting full covariance of the fitted residuals is exactly identical to the one obtained with the LS fit:

$$R = V - HCH^T = \begin{pmatrix} \frac{\sigma^2}{6} & \frac{-2\sigma^2}{6} & \frac{\sigma^2}{6} \\ \frac{-2\sigma^2}{6} & \frac{4/\sigma^2}{6} & \frac{-2\sigma^2}{6} \\ \frac{\sigma^2}{6} & \frac{-2\sigma^2}{6} & \frac{\sigma^2}{6} \end{pmatrix},$$

which is the sought after result.

5.3 The Global χ^2 alignment solution

As a next step we find solution to the alignment, or rather the alignment parameter weight matrix $\mathcal{M}$, which after inversion gives the covariance of the alignment parameters. For sake of simplicity, let us assume that each measurement belongs to a vertical measurement plane with just a single degree of freedom (DoF), i.e. can move in the measurement direction, as shown in Fig. 2.

The derivatives of the residuals w.r.t. alignment parameters are trivially obtained and give us the system of linear equations for the alignment problem:

$$A \equiv \frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\alpha}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{M} = A^T V^{-1} R V^{-1} A = \begin{pmatrix} \frac{1}{6\sigma^2} & \frac{-2}{6\sigma^2} & \frac{1}{6\sigma^2} \\ \frac{-2}{6\sigma^2} & \frac{4}{6\sigma^2} & \frac{-2}{6\sigma^2} \\ \frac{1}{6\sigma^2} & \frac{-2}{6\sigma^2} & \frac{1}{6\sigma^2} \end{pmatrix}, \quad \mathcal{R} = \begin{pmatrix} \frac{r_1 - 2r_2 + r_3}{6\sigma^2} \\ \frac{-2r_1 + 4r_2 - 2r_3}{6\sigma^2} \\ \frac{r_1 - 2r_2 + r_3}{6\sigma^2} \end{pmatrix}.$$

[§]Recall, that in the KF model tracks are parameterized at each measurement plane, in our case by three pairs of local track parameter (a_i, b_i); $\boldsymbol{\pi} = (a_1, b_1, a_2, b_2, a_3, b_3)$.

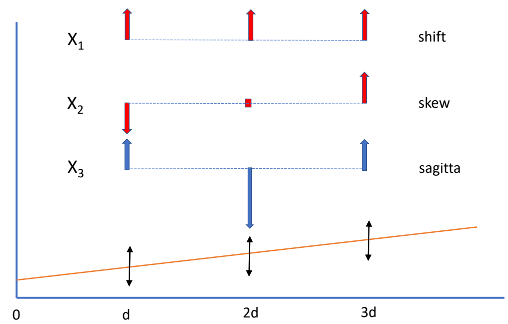


Figure 2: Alignment of three measurement planes with one DoF (vertical translation) each. X_1, X_2, X_3 show the three eigenmodes of the solution. The modes marked in red are singular in the unconstrained solution.

The matrix $\mathcal{M}$ is singular leading to two *weak modes*¶ of the alignment solution:

$$\lambda_1 = 0, \quad X_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \lambda_2 = 0, \quad X_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \lambda_3 = \frac{1}{\sigma^2}, \quad X_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

5.4 Adding constraints on the Global χ^2 alignment solution

The free-flying solution obtained in the previous section does not represent a meaningful solution in practical terms. In order for it to be so, one needs to impose additional, well motivated constraints in order to mitigate the weak (singular) modes. In alignment problems there are three types of such constraints: on alignment parameters (or their linear combinations), on individual track parameters and on common vertices of several tracks‖. The first one is straightforward and not discussed here. The latter two enter the alignment solution via the covariance of the fitted track parameters.

In our simple example let us assume a perfect knowledge of the a_0 parameter. In the KF process this is equivalent to using a prior for a_0 with an infinitesimally small uncertainty. The constraint results in removing one of the two weak modes, the skew:

$$\lambda_1 = 0, \quad X_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \lambda_2 = \frac{1}{\sigma^2}, \quad X_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \lambda_3 = \frac{1}{\sigma^2}, \quad X_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

The resulting $\mathcal{M}$ matrix is still singular due to the remaining global translation, though. This one could be removed by either introducing an arbitrary constraint on the motion of any of the measurement planes or imposing a constraint on the b_0 track parameter.

Another and very powerful constraint comes with the requirement of a common vertex of two or more tracks. In order to illustrate this, let us consider the setup shown in Fig. 3. Two tracks (A and B) are fitted through three detector elements. For the sake of maximal simplicity only the two elements at position d are free to move in the vertical direction, while the measurement plane at $2d$ remains frozen. The two eigenmodes of the solution, X_1 and X_2 would both be singular without any further constraints.

Imposing a requirement of the common vertex ($y^A(0) = y^B(0) = b_0$) correlates parameters of the two tracks. At the vertex position ($x = 0$) the joint covariance matrix of the two track parameters ($a_0^{\mathrm{A}}, b_0^{\mathrm{A}}, a_0^{\mathrm{B}}, b_0^{\mathrm{B}}$)

¶In alignment problems, it is customary to use the term *weak mode* for both strictly singular eigenmodes of the solution as well as the eigenmodes with unacceptably large uncertainties. Here, we are dealing with explicitly singular solutions, of course.

‖Other methods of mitigating weak modes are known, e.g. using tracks with different topology, notably cosmic muons. However, this does not require any special treatment from the formal point of view and as such is not discussed here.

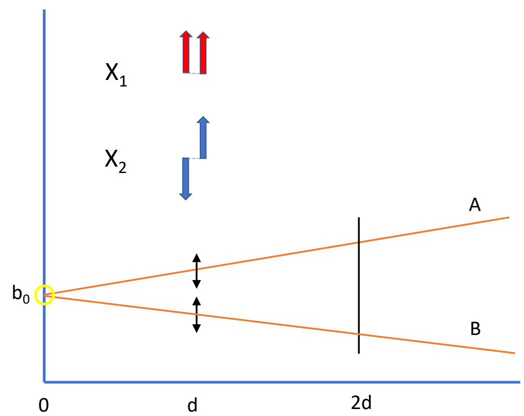


Figure 3: Alignment of two measurement planes with one DoF (vertical translation) each and using two tracks. X_1, X_2 show the eigenmodes of the solution. The mode marked in blue looses its singularity after imposing the common vertex constraint.

reads:

$$C_0^{\mathrm{AB}} = \begin{pmatrix} \frac{11\sigma^2}{10d^2} & \frac{-15\sigma^2}{10d} & \frac{9\sigma^2}{10d^2} & \frac{-15\sigma^2}{10d} \\ \frac{-15\sigma^2}{10d} & \frac{25\sigma^2}{10} & \frac{-15\sigma^2}{10d} & \frac{25\sigma^2}{10} \\ \frac{9\sigma^2}{10d^2} & \frac{-15\sigma^2}{10d} & \frac{11\sigma^2}{10d^2} & \frac{-15\sigma^2}{10d} \\ \frac{-15\sigma^2}{10d} & \frac{25\sigma^2}{10} & \frac{-15\sigma^2}{10d} & \frac{25\sigma^2}{10} \end{pmatrix}$$

The matrix $\mathcal{M}$ is singular due to the unconstrained vertex position. However, the other singularity is removed via correlation between A and B:

$$\lambda_1 = 0, \quad X_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \lambda_2 = \frac{4}{5\sigma^2}, \quad X_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

As a result, such geometry is free from deformation which would otherwise compromise vertexing capability of the detector.

5.5 Conclusions

In track-based alignment problems tracks are merely tools providing correlations between measurements recorded by detector elements. Kalman Filter provides a track fit which is equivalent to the least squares one but represented in a different model. The off-diagonal elements of the covariance matrix of the track parameters, i.e. correlations between different track Kalman states, can be recovered recursively using the smoother gain matrix. The full calculation does not appear computationally more intense than in the LS case. It may, however, be more numerically stable thanks to very simple structure of the H matrix, hence democratic treatment of all measurements. In turn, H matrix of the LS track fit is dense and highly nontrivial. Constraints on track parameters can be introduced as in the LS by adding pseudo-measurements in the filtering process. Other constraints, such as common vertex or invariant mass of several tracks, etc. can be equally introduced with full propagation to Kalman states which does not involve large matrix inversions.

ACKNOWLEDGEMENTS

I am grateful to ATLAS tracking and alignment experts for many fruitful discussions. To name just a few, I owe special thanks to Matthias Danninger, Stephen Haywood, Salvador Marti I Garcia, Anthony Morley and Andreas Salzburger.

The work has been partly supported by the Polish Ministry of Science and Higher Education grant DIR/WK/2016/13 and the National Science Centre grant 2017/25/B/ST2/01234.

References

[1] R. Frühwirth, "Application of Kalman filtering to track and vertex fitting", Nucl. Instr. Meth. **A262**, 444 (1987).

[2] P. Brückman de Renstrom et al. (ATLAS Collab.), "Common Framework Implementation for the Track-Based Alignment of the ATLAS Detector", ATL-SOFT-PUB-2014-003.

[3] W. Hulsbergen, "The Global covariance matrix of tracks fitted with a Kalman filter and an application in detector alignment", Nucl. Instr. Meth. **A600**, 471 (2009).

[4] P. Billoir, R. Frühwirth, M. Regler, "Track element merging strategy and vertex fitting in complex modular detectors", DELPHI 85-26 PROG-21.