# The global covariance matrix of tracks fitted with a Kalman filter and an application in detector alignment

W. D. Hulsbergen

*National Institute for Subatomic Physics (Nikhef), Amsterdam, The Netherlands.*

**Abstract**

We present an expression for the covariance matrix of the set of state vectors describing a track fitted with a Kalman filter. We demonstrate that this expression facilitates the use of a Kalman filter track model in a minimum $\chi^2$ algorithm for the alignment of tracking detectors. We also show that it allows to incorporate vertex constraints in such a procedure without refitting the tracks.

## 1. Introduction

Minimum $\chi^2$ algorithms for the alignment of tracking detectors generally come in two flavours, namely those that ignore and those that do not ignore the correlations between hit residuals. The former are sometimes called *local* or *iterative* methods while the latter are called *global* or *closed-form* methods [1]. The advantage of the closed-form methods is that for an alignment problem in which the measurement model is a linear function of both track and alignment parameters the solution that minimizes the total $\chi^2$ can be obtained with a single pass over the data.

The covariance matrix for the track parameters is an essential ingredient to the closed-form alignment approach [2]. If the track fit is performed using the standard expression for the least-squares estimator (sometimes called the *standard* or *global* fit method), the computation of the covariance matrix is a natural part of the track fit. This is why previously reported implementations of the closed-form alignment procedure (e.g. [3, 4, 5, 6, 7, 8]) make use of the standard fit.

In contrast most modern particle physics experiments rely on a Kalman filter track fit [9, 10] for default track reconstruction. The Kalman filter is less computa-

*Email addresses:* `wouter.hulsberg@nikhef.nl` (W. D. Hulsbergen)

tionally expensive than the standard fit and facilitates an easy treatment of multiple scattering in the form of process noise. However, the computation of the covariance matrix in the common Kalman track fit is not complete: The correlations between track parameters at different position along the track are not calculated. In the presence of process noise these correlations are non-trivial. Consequently, the result of the common Kalman track fit cannot be used directly in a closed-form alignment procedure.

In this paper we present the expressions for the computation of the global covariance matrix — the covariance matrix for all parameters in the track model — in a Kalman filter track fit. We show how this result can be used in an alignment procedure. Furthermore, using similar expressions we demonstrate how vertex constraints can be applied in the alignment without refitting the tracks in the vertex. To illustrate that our approach leads to a functional closed-form alignment algorithm, we present some results obtained for the alignment of the LHCb vertex detector with Monte Carlo simulated data.

An important motivation for extending the Kalman track fit for use in a closed-form alignment approach is that the estimation of alignment parameters is not independent of the track model. Typically, in closed-form alignment procedures the track model used in the alignment is different from that used in the track reconstruction for physics analysis, which in practise is always a Kalman filter. Sometimes the track model in the alignment is simplified, ignoring multiple scattering corrections or the magnetic field. The imperfections in the track model used for alignment will partially be absorbed in calibration parameters. Consequently, in order the guarantee consistency between track model and detector alignment, it is desirable to use the default track fit in the alignment procedure.

The Kalman filter has also been proposed for the estimation of the alignment parameters themselves [11]. This method for alignment is an alternative formulation of the closed-form alignment approach that is particularly attractive if the number of alignment parameter is large. Our results for the global covariance matrix of the Kalman filter track model and for vertex constraints can eventually be applied in such a Kalman filter alignment procedure.

## 2. Minimum $\chi^2$ formalism for alignment

To show that the global covariance matrix of the track parameters is an essential ingredient to the closed-form alignment approach, we briefly revisit the minimum $\chi^2$ formalism for alignment. Consider a track $\chi^2$ defined as

$$\chi^2 \;=\; [m - h(x)]^T \; V^{-1} \; [m - h(x)]\,, \tag{1}$$

m是track在测量平面上测到的坐标，那么h就是拟合track在测量平面上的坐标，两者相减就是测量坐标的residual, m矢量的长度是探测器的平面数

where $m$ is a vector of measured coordinates, $V$ is a (usually diagonal) covariance matrix, $h(x)$ is the measurement model and $x$ is the vector of track parameters. Note that Eq.1 is a matrix expression: $m$ and $h$ are vectors and $V$ is a symmetric matrix, all with dimension equal to the number of measurements.

For a linear expansion of the measurement model around an initial estimate $x_0$ of the track parameters,

$$h(x) = h(x_0) + H(x - x_0),$$ 泰勒展开

where

$$H = \left. \frac{\partial h(x)}{\partial x} \right|_{x_0}$$

is sometimes called the derivative or projection matrix, the condition that the $\chi^2$ be minimal with respect to $x$ can be written as

$$0 \equiv \frac{\mathrm{d}\chi^2}{\mathrm{d}x} = -2H^T V^{-1} \left[ m - h(x_0) - H(x - x_0) \right].$$

The solution to this system of equations is given by the well known expression for the least squares estimator

$$x = x_0 - C H^T V^{-1} \left[ m - h(x_0) \right], \tag{2}$$

where the matrix $C$ is the covariance matrix for $x$

$$C = \left( H^T V^{-1} H \right)^{-1}. \tag{3}$$

If the measurement model is not linear, *i.e.* if $H$ depends on $x$, expression Eq.2 can be applied iteratively, until a certain convergence criterion is met, for example defined by a minimum change in the $\chi^2$. In that case it makes sense to write Eq.2 in terms of the first and second derivative of the $\chi^2$ at the current estimate $x_0$

$$x - x_0 = - \left( \left. \frac{\mathrm{d}^2\chi^2}{\mathrm{d}x^2} \right|_{x_0} \right)^{-1} \left. \frac{\mathrm{d}\chi^2}{\mathrm{d}x} \right|_{x_0}$$

and regard the iterative minimization procedure as an application of the Newton-Raphson method.

We now consider an extension of the measurement model with a set of calibration parameters $\alpha$,

$$h(x) \rightarrow h(x, \alpha).$$

The parameters $\alpha$ are considered common to all tracks in a particular calibration sample. We estimate $\alpha$ by minimizing the sum of the $\chi^2$ values of the tracks simultaneously with respect to $\alpha$ and the track parameters $x_i$ of each track $i$,

$$\overset{\text{Global Parameter}}{\frac{\partial \sum_i \chi_i^2}{\partial \alpha}} \;=\; 0 \qquad \text{and} \qquad \forall_i \; \overset{\text{Local Parameter}}{\frac{\partial \chi_i^2}{\partial x_i}} \;=\; 0. \tag{4}$$

Please, note that the index $i$ refers to the track and not to a component of the vector $x$. We will omit the index from now on and consider only the $\chi^2$ contribution from a single track.

The number of parameters in the minimization problem above scales with the number of tracks. If the number of tracks is large enough, a computation that uses an expression for the least squares estimator analogous to Eq.2 is computationally too expensive. A more practical method relies on a computation in two steps. First, track parameters are estimated for an initial set of calibration parameters $\alpha_0$. Subsequently, the total $\chi^2$ is minimized with respect to $\alpha$ taking into account the dependence of $x_i$ on $\alpha$, *e.g.* through the total derivative

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} \;=\; \frac{\partial}{\partial \alpha} \;+\; \frac{\mathrm{d}x}{\mathrm{d}\alpha}\frac{\partial}{\partial x}. \tag{5}$$

The derivative matrix $\mathrm{d}x/\mathrm{d}\alpha$ in Eq.5 follows from the condition that the $\chi^2$ of the track remains minimal with respect to $x$, which can be expressed as

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\frac{\partial \chi^2}{\partial x} \;=\; 0$$

and results in

$$\frac{\mathrm{d}x}{\mathrm{d}\alpha} \;=\; -\frac{\partial^2 \chi^2}{\partial \alpha \partial x}\left(\frac{\partial^2 \chi^2}{\partial x^2}\right)^{-1}. \tag{6}$$

Note that if the problem is linear this derivative is independent of the actual value of $x$ or $\alpha$. Consequently, in this limit this expression remains valid even if the track $\chi^2$ was not yet minimized with respect to $x$.

The condition that the total $\chi^2$ of a sample of tracks be minimal with respect to both track and alignment parameters can now be expressed as

$$0 \;\equiv\; \frac{\mathrm{d}\chi^2}{\mathrm{d}\alpha} \tag{7}$$

For $M$ alignment parameter this defines a system of $M$ coupled non-linear equations. In analogy with the procedure introduced for the track $\chi^2$ minimization above we search for a solution by linearizing the minimum $\chi^2$ condition around an initial value $\alpha_0$ and solving the linear system of $M$ equations

只要 h(x(alpha),alpha) 函数是 alpha 的线性函数，就可以用 P3 的做法，做出 P3 的倒数第二式，对于 alpha 就是(8)式

$$\left.\frac{\mathrm{d}^2\chi^2}{\mathrm{d}\alpha^2}\right|_{\alpha_0} \Delta\alpha = -\left.\frac{\mathrm{d}\chi^2}{\mathrm{d}\alpha}\right|_{\alpha_0} \tag{8}$$

for $\Delta\alpha$. In the remainder of this section we derive the expressions for these derivatives.

To simplify the notation we define the residual vector of the track

$$r = m - h(x, \alpha)$$

and its derivative to $\alpha$

$$A_{k\ell} \equiv \frac{\partial r_k}{\partial \alpha_\ell}.$$

We linearize $r$ around the expansion point $(x(\alpha_0), \alpha_0)$, and using Eq.6 obtain for any total derivative to $\alpha$

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} = \frac{\partial}{\partial\alpha} - A^T V^{-1} H C \frac{\partial}{\partial x}.$$

(The minus sign appears because $H$ is the derivative of $h$ and not of $r$.) In this expression we have substituted the covariance matrix for $C$ for $x$. The first and second derivatives of the $\chi^2$ contribution of a single track are now given by

$$\frac{\mathrm{d}\chi^2}{\mathrm{d}\alpha} = 2A^T V^{-1} \left(V - HCH^T\right) V^{-1} r, \tag{9}$$

$$\frac{\mathrm{d}^2\chi^2}{\mathrm{d}\alpha^2} = 2A^T V^{-1} \left(V - HCH^T\right) V^{-1} A. \tag{10}$$

The matrix

$$R \equiv V - HCH^T \tag{11}$$

that appears in these expressions is the covariance matrix for the residuals $r$. This matrix is in general singular and its rank is the number of degrees of freedom of the fit.

If the track parameters $x$ for which the residuals $r$ and $H$ are calculated, are actually those that minimize the track's $\chi^2$ for the current set of alignment constants $\alpha_0$, the residuals satisfy the least squares condition $H^T V^{-1} r = 0$ and the first derivative to $\alpha$ reduces to

$$\frac{\mathrm{d}\chi^2}{\mathrm{d}\alpha} = 2A^T V^{-1} r. \tag{12}$$

Consequently, if $V$ is diagonal, the derivative to a particular parameter $\alpha_j$ only receives contributions from residuals for which $\partial r_i/\partial\alpha_j$ does not vanish.[1] An important consequence of this is that if there are additional contributions to the tracks $\chi^2$, in particular hits in subdetectors that we do not align for, constraints from a vertex fit or multiple scattering terms, then these terms only enter the derivative calculation through the track covariance matrix $C$. We will exploit this property in the next section when we discuss the use of a Kalman filter track model for alignment.

The expressions Eq.12 and Eq.10 can now be used to evaluate the first and second derivative for an initial calibration $\alpha_0$ over a given track sample and inserted in Eq.8 to obtain an improved calibration. If the residuals are non-linear in either track parameters or alignment parameters, several iterations may be necessary to minimize the $\chi^2$.

If the alignment is sufficiently constrained, the second derivative matrix can be inverted and the covariance matrix for the alignment parameters is given by

$$\mathrm{Cov}(\alpha) \;=\; 2\left(\frac{\mathrm{d}^2\chi^2}{\mathrm{d}\alpha^2}\right)^{-1}.$$

Ignoring higher order derivatives in $\alpha$, the change in the total $\chi^2$ as the result of a change $\Delta\alpha$ in the alignment parameters can be written as

$$\Delta\chi^2 \;=\; \frac{1}{2}\,\frac{\mathrm{d}\chi^2}{\mathrm{d}\alpha}^{T}\,\Delta\alpha \;=\; -\Delta\alpha^T\,\mathrm{Cov}(\alpha)^{-1}\,\Delta\alpha.$$

Consequently, the change in the total $\chi^2$ is equivalent to the significance of the alignment correction. The quantity $\Delta\chi^2$ is a useful measure for following the convergence of an alignment.

## 3. The global covariance matrix in the Kalman filter track fit

In the global method for track fitting a track is modelled by a single $N$ parameter vector (usually $N = 5$) at a fixed position along the track. Multiple scattering can be incorporated in this model by introducing explicit parameters for the kinks at scattering planes. The parameters that minimize the $\chi^2$ and the corresponding covariance matrix follow from the application of the least squares estimator Eq.2.

[1] In other words, if $\alpha_i$ is an alignment parameter of module $X$, only hits in module $X$ contribute to the first derivative of the $\chi^2$ to $\alpha_i$. 这么说的前提是module X上的alpha_j变化只会影响位于module X上测量的resiudal r_i
但我感觉这个前提不大对，只要一个 alpha 发生变化，所有module上的residual应该都会有所变化

In the Kalman filter method for track fitting [10] a track is modelled by a separate $N$ dimensional track parameter vector (or *state vector*) at each measurement (or *node*). The state vectors are related by a *transport* function, which follows from the equation of motion of the charged particle. In the absence of multiple scattering the state vectors are one-to-one functions of one-another and hence fully correlated. In the presence of multiple scattering the correlation is reduced by introducing so-called process noise in the propagation of the state vector between neighbouring nodes.

As we have seen in the previous section the closed-form method for alignment uses the vector of residuals $r$ and a corresponding covariance matrix $R$. The covariance matrix for the residuals can be computed from the global covariance matrix of the track parameters. However, the correlations between the state vectors at different nodes are normally not calculated in the Kalman filter: They are either not computed at all (if the smoothing is done as a weighted average of a forward and backward filter) or (if the Rauch-Tung-Striebel smoother formalism is applied) only implicitly and only between neighbouring nodes.

To derive an expression for the covariance matrix of all parameters in the Kalman filter track model we use the notation of reference [10] for the linear Kalman filter, in particular

- $x_k$ is the state vector at node $k$ after accumulating the information from measurements $\{1, \ldots, k\}$;
- $C_k$ is the covariance of $x_k$;
- $x_k^n$ is the state vector at node $k$ after processing all $n$ measurements.

In the following we first calculate the correlation between $x_{k-1}$ and $x_k$, which we denote by $C_{k-1,k}$. From this we proceed with the correlation matrix between $x_{k-1}^n$ and $x_k^n$. The correlation between any two states $k$ and $l$ then follows from the observation that the correlation between these states occurs via intermediate states.

In the notation of [10] we have for the prediction of state $k$ from state $k-1$,

$$x_k^{k-1} \;=\; F_{k-1} x_{k-1},$$

where $F$ is the Jacobian or transport matrix. The covariance of the prediction is given by

$$C_k^{k-1} \;=\; F_{k-1} C_{k-1} F_{k-1}^T \;+\; Q_{k-1}$$

where $Q_{k-1}$ is the process noise in the transition from state $k-1$ to $k$. The full covariance matrix for the pair of states $(x_{k-1}, x_k^{k-1})$ is then given by

$$\mathrm{Cov}(x_{k-1}, x_k^{k-1}) = \begin{pmatrix} C_{k-1} & C_{k-1}F_{k-1}^T \\ F_{k-1}C_{k-1} & F_{k-1}C_{k-1}F_{k-1}^T + Q_{k-1} \end{pmatrix}$$

In the Kalman filter track fit we now proceed by adding the information of measurement $k$ to obtain a new estimate for the state in procedure that is called *filtering* and leads to state vector $x_k$. The remaining measurements $k+1, \ldots, n$ are processed with prediction and filter steps in the same fashion. Afterwards a procedure called *smoothing* can be applied to recursively propagate the information obtained through measurements $k+1, \ldots, n$ back to node $k$. The smoothed state vector at node $k$ is labelled by $x_k^n$ and its covariance by $C_k^n$. To derive the expression for the covariance matrix of the smoothed states $x_k^n$ and $x_{k-1}^n$ we first present the following lemma.

Suppose we have two observables $a$ and $b$ with covariance matrix

$$\begin{pmatrix} V_{aa} & V_{ab} \\ V_{ba} & V_{bb} \end{pmatrix}$$

Now suppose we have obtained a new estimate of $\tilde{a}$ with variance $\tilde{V}_{aa}$ by adding information. We can propagate the new information to $b$ with a least squares estimator, which gives

$$\begin{aligned} \tilde{b} &= b + V_{ba}V_{aa}^{-1}(\tilde{a} - a) \\ \tilde{V}_{bb} &= V_{bb} + V_{ba}V_{aa}^{-1}(\tilde{V}_{aa} - V_{aa})V_{aa}^{-1}V_{ab} \\ \tilde{V}_{ab} &= V_{ab} + (\tilde{V}_{aa} - V_{aa})V_{aa}^{-1}V_{ab} = \tilde{V}_{aa}V_{aa}^{-1}V_{ab} \end{aligned} \tag{13}$$

This expression also holds if $a$ and $b$ are vectors. It can be derived by minimizing the following $\chi^2$

$$\chi^2 = \begin{pmatrix} \tilde{a} - a \\ \tilde{b} - b \end{pmatrix}^T \begin{pmatrix} V_{aa} & V_{ab} \\ V_{ba} & V_{bb} \end{pmatrix}^{-1} \begin{pmatrix} \tilde{a} - a \\ \tilde{b} - b \end{pmatrix} + (\tilde{a} - m)^T \, V_m{}^{-1} \, (\tilde{a} - m)$$

(where $m$ with variance $V_m$ is the additional information for $a$) with respect to $\tilde{a}$ and $\tilde{b}$.

Substituting $x_{k-1}$ for $b$, $x_k$ for $a$ and $x_k^n$ for $\tilde{a}$ in Eq.13 we obtain for the correlation between the smoothed states

$$C_{k-1,k}^n = C_{k-1}F_{k-1}^T \left(C_k^{k-1}\right)^{-1} C_k^n = A_{k-1}C_k^n \tag{14}$$

where we have used the definition of the smoother gain matrix [10]

$$A_{k-1} \;=\; C_{k-1} F_{k-1}^T \left( C_k^{k-1} \right)^{-1} . \tag{15}$$

For the smoothed state $x_{k-1}^n$ and its covariance we find

$$\begin{aligned} x_{k-1}^n &= x_{k-1} + A_{k-1}(x_k^n - x_k^{k-1}), \\ C_{k-1}^n &= C_{k-1} \;+\; A_{k-1} \left( C_k^n - C_k^{k-1} \right) A_{k-1}^T . \end{aligned}$$

These are the Rauch-Tung-Striebel smoothing expressions as found in [10]. The gain matrix in Eq.15 can be written in different forms, *e.g.*

$$A_{k-1} \;=\; (F_{k-1})^{-1} \left( C_k^{k-1} - Q_{k-1} \right) \left( C_k^{k-1} \right)^{-1} .$$

This expression shows explicitly that $A_{k-1} = (F_{k-1})^{-1}$ if there is no process noise ($Q = 0$). Therefore, as one expects, without process noise the smoothed states in the Kalman filter are just related by the transport equation.

Once we have the calculated the off-diagonal element $(k-1, k)$, we proceed to the next diagonal $(k-2, k)$. The correlation between states $k-2$ and $k$ can be calculated by performing a simultaneous smoothing of states $k-2$ and $k-1$: In the argument above we substitute a new vector $(x_{k-2}, x_{k-1})$ for $b$, rather than just the state $x_{k-1}$. The result for the correlation $(k-2, k)$ is

$$C_{k-2,k}^n \;=\; C_{k-2,k-1}^n \, (C_{k-1}^n)^{-1} \, C_{k-1,k}^n .$$

This expression can also be derived with a simpler argument: The origin of the correlation between state vectors is the transport. Therefore, the correlation $(k-2, k)$ occurs only through the correlations $(k-2, k-1)$ and $(k-1, k)$. Following the same reasoning the next diagonal becomes

$$\begin{aligned} C_{k-3,k}^n &= C_{k-3,k-2}^n \left( C_{k-2}^n \right)^{-1} C_{k-2,k-1}^n \left( C_{k-1}^n \right)^{-1} C_{k-1,k}^n \\ &= C_{k-3,k-2}^n \left( C_{k-2}^n \right)^{-1} C_{k-2,k}^n , \end{aligned}$$

which shows that the calculation can be performed recursively. By substituting the smoother gain matrix we can write this in the following compact form

$$C_{k-1,\ell}^n \;=\; C_{k-1,k}^n \, (C_k^n)^{-1} \, C_{k,\ell}^n \;=\; A_{k-1} C_{k,\ell}^n \qquad k \leq \ell . \tag{16}$$

If the gain matrices are temporarily stored, then if $N$ is the dimension of the state vector, the calculation of each off-diagonal element in the full covariance matrix of state vectors requires about $N$ multiplications and $N$ additions. The total number of operations by far exceeds the numerical complication of the standard Kalman filter. However, we have found that for tracks traversing the entire LHCb tracking system, with a total of about 30 measurement coordinates, the computational cost for the global covariance matrix with the procedure above was smaller than that of the Kalman filter track fit itself. This is because the LHCb track fit is largely dominated by integration of the inhomogeneous magnetic field and the location of intersections with detector material.

Now that we have calculated the full covariance matrix of all states $x_k^{(n)}$, the elements of the covariance matrix $R$ of the residuals are simply given by

$$R_{k\ell} \;=\; V_k \delta_{kl} - H_k C^n_{k,\ell} H^T_\ell \,. \tag{17}$$

This completes the recipe for using a Kalman filter track model in the alignment of tracking detectors.

We have argued below Eq.12 that the cancellation that takes place between Eq.9 and Eq.12 is important when considering the Kalman track fit for alignment. This can be explained as follows. If we were to use the Kalman filter track model in a global $\chi^2$ fit, the $\chi^2$ would contain explicit contributions for the difference in the state vectors at neighbouring nodes,

$$\chi^2_{k-1,k} \;=\; (x_k - F_{k-1} x_{k-1})^T \, (Q_{k-1})^{-1} \, (x_k - F_{k-1} x_{k-1}) \,.$$

These contributions are equivalent to the terms that constrain scattering angles in the conventional track model for a global track fit. As they represent additional constraints to the $\chi^2$, they must also appear in the matrix $V$ and the residual vector $r$ in Eq.9. It is only because of the minimum $\chi^2$ condition for the track parameters that their contribution in the derivatives to the alignment parameters vanishes.

## 4. Vertex constraints

The expressions in Eq.13 can also be used to include vertex or mass constraints in an alignment procedure. First, we propagate the track parameters to the estimated position of the vertex. We label the track parameters at that position with $x_0^n$ and its covariance by $C_0^n$. The correlations between these track parameters and those at the position of each measurement can be computed with the procedure outlined in the previous section.

For clarity we now drop the superscript $n$ and replace it with a superscript $(i)$ that labels the track in the vertex: The state of track $i$ at the vertex is $x_0^{(i)}$

with covariance $C_0^{(i)}$. As a result of the vertex fit (which we can implement as the Billoir-Frühwirth-Regler algorithm [12, 10]) we obtain the new ‘constrained’ track parameters $\tilde{x}_0^{(i)}$ with covariance $\tilde{C}_0^{(i)}$. The change in the track parameters can be propagated to the track states at each measurement using Eq.13, which gives for the state vector at node $k$

$$\tilde{x}_k^{(i)} \;=\; x_k^{(i)} \;+\; C_{k,0}^{(i)} \left(C_0^{(i)}\right)^{-1} (\tilde{x}_0^{(i)} - x_0^{(i)}) \tag{18}$$

and for the covariance

$$\tilde{C}_{k,\ell}^{(i)} \;=\; C_{k,\ell}^{(i)} + C_{k,0}^{(i)} \left(C_0^{(i)}\right)^{-1} (\tilde{C}_0^{(i)} - C_0^{(i)}) \left(C_0^{(i)}\right)^{-1} C_{0,\ell}^{(i)}. \tag{19}$$

The constrained residuals for track $j$ then become

$$\tilde{r}_k^{(i)} \;=\; m_k^{(i)} - h(\tilde{x}_k^{(i)}) \tag{20}$$

and the covariance matrix $R$ in Eq.17 can be computed using the new track state covariance $\tilde{C}_{k,\ell}^{(i)}$.

The vertex fit also gives us the covariance $\tilde{C}^{(i,j)}$ between any two tracks $i$ and $j$ in the vertex. This allows to compute the correlation between any two states in any two tracks as follows

$$\tilde{C}_{k,\ell}^{(i,j)} \;=\; C_{k,0}^{(i)} \left(C_0^{(i)}\right)^{-1} \tilde{C}_0^{(i,j)} \left(C_0^{(j)}\right)^{-1} C_{0,\ell}^{(j)} \tag{21}$$

Inserting this into the multi-track equivalent of Eq.17 gives the full correlation matrix for the residuals on *all* tracks. If the number of tracks in the vertex is large, the computation of the global covariance matrix for all states on all tracks is rather CPU time consuming. Therefore, in practical applications it makes sense to compute the correlation only for a subset of hits close to the vertex.

This completes the ingredients for including vertex constraints in the calculation of the alignment derivatives. Eventual mass constraints or other kinematic constraints are included implicitly if they are applied during the vertex fit.

## 5. Application to the alignment of the LHCb tracking system

The LHCb tracking system consists of a silicon vertex detector (VELO) and a spectrometer [13]. For the track based alignment of this system a closed-form alignment algorithm has been implemented in the LHCb software framework. This algorithm, which uses the standard LHCb Kalman filter track fit, will be described in detail in a future publication [14]. Here we briefly illustrate the effect of correlations between residuals and the applications of vertex constraints, using the alignment of the VELO system as an example.[2]

The VELO system consists of 21 layers of double sided silicon detectors with radial strips on one side and concentric circular strips on the other. Each layer consist of two half circular disks called *modules*. The modules are mounted onto two separate support structures, the *left* and *right* VELO halves. The two halves can be moved independently in the direction perpendicular to the beam ($z$) axis in order to ensure the safety of the detectors during beam injection. The alignment of the VELO system is of crucial importance to the physics performance of the LHCb experiment.

For the analysis described here we have simulated deformations in the VELO detector in such a way that it bows along the $z$-axis: we introduced a bias in the $x$ and $y$ position of each module that was approximately proportional to $(z - z_0)^2$, where $z - z_0$ was the $z$ position of the module relative to the middle of the VELO. The reason to choose this particular misalignment is that it corresponds to a correlated movement of detector elements: Such deformations — sometimes called 'weak modes' — are inherently difficult to correct for with an alignment method that ignores the correlations between residuals in the track fit.

Tracks from a sample of simulated minimum bias interactions were reconstructed using a 'cheated' pattern recognition, assigning VELO hits to a track based on the Monte Carlo truth. We required at least 8 hits per track. The tracks were fitted with the standard LHCb track fit, taking scattering corrections into account as process noise. Tracks were accepted for alignment if their $\chi^2$ per degree of freedom was less than 20. In a perfectly aligned detector this cut only excludes a tiny fraction of reconstructed tracks, namely those with a kink due to a hadronic interaction. Primary vertices were reconstructed using the standard LHCb primary vertex finder.

To validate the implementation of the algorithm we have performed two tests. First, we have checked the calculation of the residual covariance matrix $R$ by comparing it to a numerical computation. A single track was refitted after changing one measurement coordinate $m_i$ by a numerically small value. The $i$-th row of $R$

[2]An alternative algorithm for the VELO alignment is described in [8]. It uses a standalone global track fit with a straight line track model, suitable only in the field-free region of the detector.

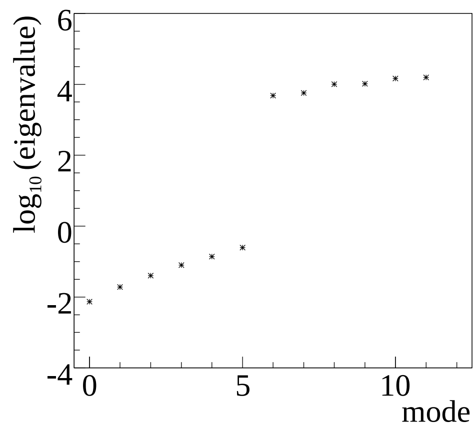


Figure 1: Logarithm of the eigenvalue versus eigenvalue index ('mode') in the alignment of the position and rotation of two halves of the LHCb VELO.

follows from (for $j \neq i$)

$$R_{ij} = -\frac{\delta r_j^{(i)}}{\delta r_i^{(i)}} R_{ii},$$

where $\delta r_j^{(i)}$ is the change in residual $r_j$. (The computation of the diagonal element $R_{ii}$ is part of the standard Kalman fit procedure.) This test has shown that the numerical uncertainty in the correlations coefficients of $R$ is typically of order $10^{-4}$, which is good enough for the purpose of detector alignment.

Second, we have analyzed the eigenvalue spectrum of the second derivative matrix Eq.10. Without an external reference system the global translations and rotations of a tracking system are unconstrained in the alignment procedure. Such unconstrained degrees of freedom lead to vanishing eigenvalues in the derivative matrix and, if left untreated, result in a poorly converging alignment. (See *e.g.* [5].) Unconstrained degrees of freedom can be removed with Lagrange constraints or by omitting the corresponding eigenvector from the solution to the linear system in Eq.8. However, to test the implementation of the calculations in the global alignment algorithm, the identification of the vanishing eigenvalues is a powerful tool: If the zero eigenvalues corresponding to the global movements are observed, we can be confident that the computation of both the matrix $R$ and the alignment derivatives $\partial r/\partial\alpha$ is correct (or at least consistently wrong).

Figure 1 shows the eigenvalues for the alignment of the position and rotation of the two VELO halves. (The eigenvalues are plotted versus an arbitrary index

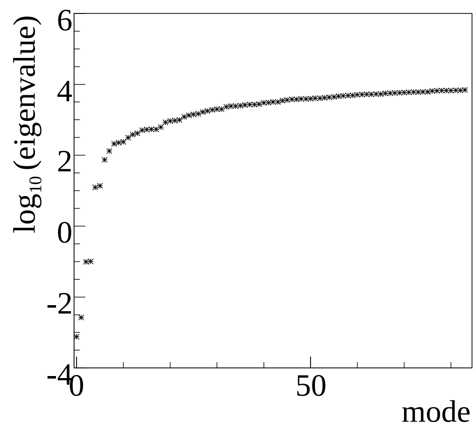


Figure 2: Logarithm of the eigenvalue versus eigenvalue index (‘mode’) in the alignment of the position of modules in the LHCb VELO.

that increases with the size of the eigenvalue.) The total number of alignment parameters is 12. To define the scale of the eigenvalues the derivative matrix was rescaled following the recipe in [6]: the numerical value of the eigenvalue is roughly equal to the number of hits contributing to the corresponding linear combination of alignment parameters. As can be seen in the figure the eigenvalue distribution splits in two: The six smaller eigenvalues correspond to the global rotation and translation, whereas the six larger eigenvalues correspond to the relative alignment of the two detector halves. Note that if correlations between residuals are ignored, the linear equations Eq.8 split in independent parts for the two aligned objects and all eigenvalues are of about the same size.

One may wonder why the eigenvalues corresponding to the global movements are not ‘numerically’ zero, in contrast with the analysis reported in [5]. The reason for this is a feature of the Kalman filter: In the Kalman filter the state vector is seeded with a finite variance even before a single measurement is processed. The variance must be large enough to have negligible weight in the variance of the state vector after all measurements are processed, but it must be small enough to make the computation of the filter gain matrix numerically stable. The finite value of the seed variance essentially fixes the track in space. We have observed that the value of the small eigenvalues is indeed sensitive to the variance of the seed. For practical purposes the bias from the Kalman filter seed is not important.

To test the alignment procedure for the misalignment scenario presented above

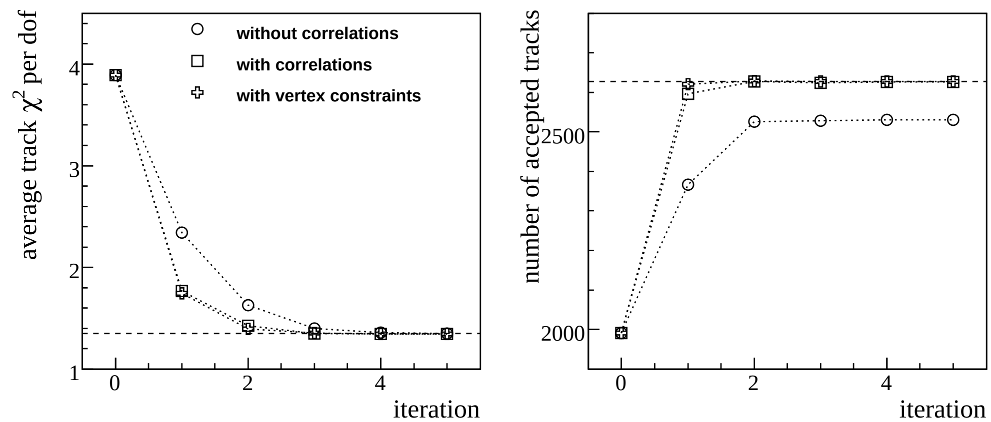


Figure 3: Number of selected tracks (right) and average $\chi^2$ per track (left) as a function of the number of alignment iterations for 3 alignment scenarios, namely ignoring correlations between residuals, not ignoring those correlations and including vertex constraints. The dashed line represents the result for a perfectly aligned detector.

we aligned the position of each module in $x$ and $y$, corresponding to a total of 84 alignment parameters. We omitted the $z$ translation and rotations to simplify the analysis. The eigenvalue distribution, shown in figure 2, reveals 4 unconstrained degrees of freedom. These correspond to the global translation in $x$ and $y$ and — originating from the planar geometry of the detector — shearings in the $xz$ and $yz$ plane. We constrain these degrees of freedom with Lagrange constraints.

We report here two figures of merit that we use to judge the convergence of the alignment procedure, namely the number of selected tracks and the average $\chi^2$ of selected tracks, both as a function of the alignment iteration. The results are shown in figure 3 for 3 different scenarios: First, we entirely ignore correlations between residuals, which means that off-diagonal elements in the matrix $R$ in equation Eq.11 are assumed zero. Second, we compute these correlations with the recipe outlined in section 3. Finally, we also include vertex constraints with the expressions given in section 4. As can be seen in the figure the scenario with correlations converges faster than the scenario without. Furthermore, in the scenario without correlations less tracks survive the $\chi^2$ cut even after 5 iterations.

The difference in convergence behaviour is mostly because there are two kinds of tracks. Though most tracks pass only through a single VELO half, there is a small fraction that passes through small regions in which detectors from both halves overlap. When correlations between hit residuals are taken into account, tracks that pass through a single half do not carry any weight in determining the relative positions of the two halves, because the contribution to the $\chi^2$ is invariant to the global position of the detector half. Therefore, the relative position of the two halves is fully sensitive to the tracks that pass through both halves. On the contrary, if correlations are ignored, every track fixes the position of any detector element in space. As a result the overlap tracks get a much smaller weight in determining the relative position and convergence becomes poor.

This problem can be partially overcome by explicitly enhancing the fraction of overlap tracks in the sample, *e.g.* by down-sampling the tracks that do not pass through the overlap regions. Such a strategy is applied in the alignment of the Babar vertex detector [15]. An important advantage of the closed-form algorithm is that it is not necessary to remove tracks with a small weight in the alignment as the algorithm inherently weights the information contained in the residuals correctly.

## 6. Conclusions

In this paper we have presented how the most popular track fitting method, the Kalman filter, can be used in a closed-form alignment procedure for tracking detectors. Our contribution is summarized in expression Eq.16 which shows how the correlations between state vectors can be computed recursively by using the smoother gain matrix. We have also shown how vertex constraints can be included without refitting the tracks. Using an implementation of this formalism in the LHCb software framework we have illustrated for a simple misalignment scenario of the LHCb vertex detector the importance of correlations between residuals in the track fit. A more detailed analysis of the performance of the alignment algorithm to the LHCb tracking system will be reported in due course [14].

## Acknowledgments

The author would like to thank G. Raven for posing the question concerning correlations between residuals in the Kalman filter and for his patient proof reading of this manuscript. The alignment software that was used for the analysis in section 5 was developed in close collaboration with J. Amoraal, A. Hicheur, M. Needham and L. Nicolas and G. Raven,

## References


[1] For an overview, see V. Blobel, “Software alignment for tracking detectors”, Nucl. Instrum. Meth. A **566** (2006) 5.

[2] V. Blobel and C. Kleinwort, PHYSTAT02 proceedings, “A new method for the high-precision alignment of track detectors,”, arXiv:hep-ex/0208021.

[3] I. Belotelov, A. Lanyov and G. Ososkov, “Data-driven alignment of the Hera-B Outer Tracker”, Phys. Part. Nucl. Lett. **3** (2006) 335.

[4] C. Kleinwort, “H1 alignment experience”, Proceedings of the first LHC Detector Alignment Workshop, Geneva (2006)

[5] P. Brückman, A. Hicheur and S. J. Haywood, “Global $\chi^2$ approach to the alignment of the ATLAS silicon tracking detectors”, ATL-INDET-PUB-2005-002; P. Brückman de Renstrom, S. Haywood, in: L. Lyons, M.K. nel (Eds.), PHYSTAT05 proceedings, IC Press, 2006.

[6] A. Bocci and W. Hulsbergen, “TRT alignment for SR1 cosmics and beyond”, ATL-COM-INDET-2007-011.

[7] G. Flucke, P. Schleper, G. Steinbruck and M. Stoye, “A study of full scale CMS tracker alignment using high momentum muons and cosmics,”, CERN-CMS-NOTE-2008-008.

[8] S. Viret, C. Parkes and M. Gersabeck, “Alignment procedure of the LHCb Vertex Detector,” Nucl. Instrum. Meth. A **596** (2008) 157.

[9] R. Kalman, “A new approach to linear filtering and prediction problems”, Journal of Basic Engineering, 35 (1960).

[10] R. Fruhwirth, “Application of Kalman Filtering to track and vertex fitting,” Nucl. Instrum. Meth. A **262** (1987) 444.

[11] R. Fruhwirth, T. Todorov, M. Winkler 2003 J. Phys. G: Nucl. Part. Phys. 29 561; E. Widl and R. Fruhwirth, “A large-scale application of the Kalman alignment algorithm to the CMS tracker,” J. Phys. Conf. Ser. **119** (2008) 032038.

[12] P. Billoir, R. Frühwirth and M. Regler, "Track element merging strategy and vertex fitting in complex modular detectors", Nucl. Instrum. Meth. A **241**, 115 (1985)

[13] A. Augusto Alves *et al.* [LHCb Collaboration], JINST **3** (2008) S08005.

[14] J. Amoraal *et al.*, "Alignment of the LHCb tracking system", *in preparation*.

[15] D. N. Brown, A. V. Gritsan, Z. J. Guo and D. Roberts, "Local Alignment of the BABAR Silicon Vertex Tracking Detector", arXiv:0809.3823 [physics.ins-det].