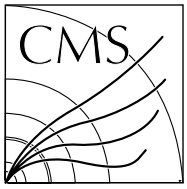
Available on CMS information server

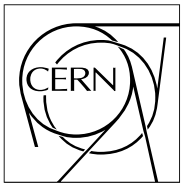
CMS NOTE 2006/022



The Compact Muon Solenoid Experiment

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland



January 30, 2006

A Kalman Filter for Track-based Alignment

E. Widl, R. Frühwirth, W. Adam

Institute of High Energy Physics, Austrian Academy of Sciences, Vienna, Austria

Abstract

An iterative method for track-based global alignment is proposed. It is derived from the Kalman filter and is designed to avoid the inversion of large matrices. The update formulas for the alignment parameters and for the associated covariance matrix are described. The implementation and the computational complexity is discussed, and it is shown how to limit the latter to an acceptable level by restricting the update to detectors that are close in the sense of a certain metrics. The performance of the Kalman filter in terms of precision and speed of convergence is studied in a simplified setup. First results from an implementation in the CMS reconstruction program ORCA are presented, using two sections of the barrel part of the CMS Tracker.

1 Introduction

This note describes an iterative method, derived from the Kalman filter [1, 2, 3], for the global alignment of large tracking systems using charged tracks. The method is iterative, because the alignment parameters are updated after each track. The method is global, because the update is not restricted to the detector units that are crossed by the track. In case the number of detector units to be aligned is very large, it is proposed to limit the update to those detector units that have significant correlations with the ones in the current track trajectory. The update formulas for the alignment parameters and for the associated covariance matrix are derived. It turns out that no large matrices have to be inverted. However, a certain amount of bookkeeping is required in order to restrict the computational load to an acceptable level.

Alignment with tracks requires an already aligned and fixed reference system. All updates are relative to this reference system. The problem of obtaining such a reference system is not discussed here.

In the formalism that is proposed here it is possible to use prior information about the alignment obtained from mechanical and/or laser measurements. It is also possible to fix the position of certain detector units by giving them a large prior weight (small prior uncertainty). A requirement that several detectors move along with each other can be enforced by large prior correlations.

The formulas for the updating of the alignment parameters and their variance-covariance matrix are derived in Section 2. The computation complexity of the method and implementation issues are discussed in Section 3. Section 4 presents results from the alignment of a simplified setup with straight tracks. The precision and speed of convergence is studied under various assumptions on the starting values, the detector resolution, the number of reference detectors and their relative position. In Section 5 results from the barrel part of the CMS Tracker are shown, using tracks from a fast simulation. The conclusions and an outlook to further developments are given in the final Section 6.

2 Sequential Updating of Alignment Parameters

A synopsis of the notation that is used in the following is shown in Table 1. All vectors are supposed to be column vectors.

Assume that there exist alignment parameters $\boldsymbol{d}$ with a covariance matrix $\boldsymbol{D}$. They are to be updated by using information from a track with track parameters $\boldsymbol{x}_{\mathrm{t}}$ and observations $\boldsymbol{m}$. The observations $\boldsymbol{m}$ of a track depend on the track parameters $\boldsymbol{x}_{\mathrm{t}}$ via the track model f:

$$\boldsymbol{m} = \boldsymbol{f}(\boldsymbol{x}_{\mathrm{t}}) + \varepsilon, \quad \mathsf{cov}(\varepsilon) = \boldsymbol{V}.$$

The stochastic vector ε contains the effects of the observation error and of multiple scattering. Its variance-covariance matrix $\boldsymbol{V}$ can be assumed to be known. It is essential that a preliminary track fit is performed in order

Table 1: Synopsis of the notation

$\boldsymbol{d}_{\mathrm{t}}$	vector of true alignment parameters
$\boldsymbol{d}_0$	expansion point of alignment parameters
$\boldsymbol{d}$	current estimate of alignment parameters
$\boldsymbol{d}_i$	subvector of alignment parameters of detector unit i
$\boldsymbol{D}$	covariance matrix of $\boldsymbol{d}$
$\boldsymbol{D}_{ij}$	submatrix of cross-correlations between detector units i and j
$\widehat{\boldsymbol{d}}$	updated estimate of alignment parameters
$\widehat{\boldsymbol{D}}$	covariance matrix of $\widehat{\boldsymbol{d}}$
$\boldsymbol{m}$	observations of the current track
$\boldsymbol{V}$	covariance matrix of $\boldsymbol{m}$
$\boldsymbol{x}_{\mathrm{t}}$	true track parameters of the current track
$\boldsymbol{x}_0$	expansion point of track parameters
$\boldsymbol{x}$	predicted track parameters of the current track
$\boldsymbol{C}$	covariance matrix of $\boldsymbol{x}$
$\widehat{\boldsymbol{x}}$	updated track parameters of the current track

to get a provisional estimate $\boldsymbol{x}$ of the track parameters and of the momentum in particular. As the tracks used for alignment are mainly high-energy minimum-ionizing particles, energy loss can be considered as deterministic and is dealt with in the track model.

If the detector units are not perfectly aligned, the observations also depend on the alignment parameters $\boldsymbol{d}_{\mathrm{t}}$, and the track model has to be modified accordingly:

$$\boldsymbol{m} = \boldsymbol{f}(\boldsymbol{x}_{\mathrm{t}}, \boldsymbol{d}_{\mathrm{t}}) + \varepsilon, \quad \mathsf{cov}(\varepsilon) = \boldsymbol{V}.$$

As is usual, the track model is linearized by a first-order Taylor expansion at expansion points $\boldsymbol{d}_0$ and $\boldsymbol{x}_0$:

$$\boldsymbol{m} = \boldsymbol{c} + \boldsymbol{A}\boldsymbol{d}_{\mathrm{t}} + \boldsymbol{B}\boldsymbol{x}_{\mathrm{t}} + \varepsilon = \boldsymbol{c} + \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \end{pmatrix} \begin{pmatrix} \boldsymbol{d}_{\mathrm{t}} \\ \boldsymbol{x}_{\mathrm{t}} \end{pmatrix} + \varepsilon,$$

where

$$\boldsymbol{A} = \partial \boldsymbol{m}/\partial \boldsymbol{d}_{\mathrm{t}}\big|_{\boldsymbol{d}_0}, \quad \boldsymbol{B} = \partial \boldsymbol{m}/\partial \boldsymbol{x}_{\mathrm{t}}\big|_{\boldsymbol{x}_0}, \quad \boldsymbol{c} = f(\boldsymbol{x}_0, \boldsymbol{d}_0) - \boldsymbol{A}\boldsymbol{d}_0 - \boldsymbol{B}\boldsymbol{x}_0.$$

The expansion point $\boldsymbol{d}_0$ is either the nominal or the currently estimated sensor position, and the expansion point $\boldsymbol{x}_0$ is the result of a preliminary track fit.

In principle, the Kalman filter requires a prediction $\boldsymbol{x}$ of the track parameters, along with its variance-covariance matrix $\boldsymbol{C}$. This prediction has to be independent of the observations $\boldsymbol{m}$. It is conceivable that such a prediction exists, for instance as the result of a vertex and/or kinematic fit constraining some or all of the track parameters. In this case the update equation of the Kalman filter reads:

$$\begin{pmatrix} \widehat{\boldsymbol{d}} \\ \widehat{\boldsymbol{x}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{d} \\ \boldsymbol{x} \end{pmatrix} + \boldsymbol{K}\left(\boldsymbol{m} - \boldsymbol{c} - \boldsymbol{A}\boldsymbol{d} - \boldsymbol{B}\boldsymbol{x}\right),$$

with the gain matrix of the filter:

$$\boldsymbol{K} = \begin{pmatrix} \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C} \end{pmatrix} \begin{pmatrix} \boldsymbol{A}^T \\ \boldsymbol{B}^T \end{pmatrix} \underbrace{\left(\boldsymbol{V} + \boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^T + \boldsymbol{B}\boldsymbol{C}\boldsymbol{B}^T\right)^{-1}}_{\boldsymbol{G}} = \begin{pmatrix} \boldsymbol{D}\boldsymbol{A}^T\boldsymbol{G} \\ \boldsymbol{C}\boldsymbol{B}^T\boldsymbol{G} \end{pmatrix}.$$

In general, however, there is no independent prediction of the track parameters. In this case, the preliminary track parameters $\boldsymbol{x}_0$ are used as the "prediction", but with zero weight in order not to bias the estimation. This is accomplished by multiplying $\boldsymbol{C}$ with a scale factor α and letting α tend to infinity:

$$\boldsymbol{G} = \lim_{\alpha \longrightarrow \infty} \left(\boldsymbol{V} + \boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^T + \alpha \boldsymbol{B}\boldsymbol{C}\boldsymbol{B}^T\right)^{-1} = \boldsymbol{V}_D^{-1} - \boldsymbol{V}_D^{-1}\boldsymbol{B}(\boldsymbol{B}^T\boldsymbol{V}_D^{-1}\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{V}_D^{-1}, \text{ with } \boldsymbol{V}_D = \boldsymbol{V} + \boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^T.$$

Here the Sherman-Morrison inversion formula

$$(\boldsymbol{X} + \boldsymbol{H}\boldsymbol{Y}\boldsymbol{H}^T)^{-1} = \boldsymbol{X}^{-1} - \boldsymbol{X}^{-1}\boldsymbol{H}(\boldsymbol{Y}^{-1} + \boldsymbol{H}^T\boldsymbol{X}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{X}^{-1}$$

has been used (see e.g. [4]). It can easily be verified that $\boldsymbol{G}\boldsymbol{B} = \boldsymbol{0}$. The update equation of the alignment parameters can therefore be simplified to

$$\widehat{\boldsymbol{d}} = \boldsymbol{d} + \boldsymbol{D}\boldsymbol{A}^T\boldsymbol{G}\left(\boldsymbol{m} - \boldsymbol{c} - \boldsymbol{A}\boldsymbol{d}\right).$$

Finally, the update of the covariance matrix can be calculated by linear error propagation:

$$\widehat{\boldsymbol{D}} = \left(\boldsymbol{I} - \boldsymbol{D}\boldsymbol{A}^T\boldsymbol{G}\boldsymbol{A}\right)\boldsymbol{D}\left(\boldsymbol{I} - \boldsymbol{A}^T\boldsymbol{G}\boldsymbol{A}\boldsymbol{D}\right) + \boldsymbol{D}\boldsymbol{A}^T\boldsymbol{G}\boldsymbol{V}\boldsymbol{G}\boldsymbol{A}\boldsymbol{D}.$$

As both terms on the right hand side are positive definite the left hand side is guaranteed to be positive definite as well.

3 Implementation and Computational Complexity

The total number of detector units is denoted by N. The current track crosses a certain number of detector units, denoted by k. If each of them gives a two-dimensional measurement, the dimension $n = 2\,k$ of the observation vector $\boldsymbol{m}$ is small for high-energy tracks, usually not larger than 30. The matrix $\boldsymbol{B}$ is of size $n \times 5$ and is therefore small. The matrix $\boldsymbol{A}$ is a row of N blocks $\boldsymbol{A}_i$ of size $n \times m$, where m is the number of alignment parameters per detector unit (usually 6). For each track, only k out of these N blocks are different from zero. The set of detector units crossed by the current track is denoted by $I = \{i_1, \ldots, i_k\}$. Then the matrix $\boldsymbol{A}$ has the following form:

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{0} \ldots \boldsymbol{0}\ \boldsymbol{A}_{i_1}\ \boldsymbol{0} \ldots \boldsymbol{0}\ \boldsymbol{A}_{i_2}\ \boldsymbol{0} \ldots\ldots \boldsymbol{0}\ \boldsymbol{A}_{i_k}\ \boldsymbol{0} \ldots \boldsymbol{0} \end{pmatrix}.$$

3.1 Update of the Alignment Parameters

The only large matrix in the parameter update is the product $\boldsymbol{D}\boldsymbol{A}^T$. It is a column of N blocks each of which has size $m \times n$. However, only those blocks need to be computed that correspond to the detector units that have significant correlation with the ones in the current track. In order to keep track of the necessary updates, a list L_i is attached to each detector unit i, containing the detector units that have significant correlations with i. This list may contain only i itself in the beginning and grows as more tracks are processed. If there is prior knowledge about correlations, for instance because of mechanical constraints, it can be incorporated in the list and in the initial covariance matrix. The length of the list can hopefully be restricted to a fairly small number, as the correlations between detector units that are far from each other tend to be small. This leads to the following procedure for computing the updated alignment parameters:

1. Update the list L_i for every $i \in I$ (see Subsection 3.3).
2. Form the list L of all detector units that are correlated with the ones crossed by the current track: $L = \bigcup_{i \in I} L_i$. The size of L should be much smaller than N.
3. For all $j \in L$ compute: $(\boldsymbol{D}\boldsymbol{A}^T)_j = \sum_{i \in I} \boldsymbol{D}_{ji}\boldsymbol{A}_i^T$. Each block $\boldsymbol{D}_{ji}$ is of size $m \times m$.
4. Compute: $\boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^T = \sum_{i \in I} \boldsymbol{A}_i(\boldsymbol{D}\boldsymbol{A}^T)_i$.
5. Compute: $\boldsymbol{V}_D = \boldsymbol{V} + \boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^T$ and $\boldsymbol{G}$. All matrices involved are of size $n \times n$.
6. Compute: $\boldsymbol{m'} = \boldsymbol{G}\left(\boldsymbol{m} - \boldsymbol{c} - \sum_{i \in I} \boldsymbol{A}_i\boldsymbol{d}_i - \boldsymbol{B}\boldsymbol{x}\right)$.
7. For all $j \in L$ compute: $\widehat{\boldsymbol{d}}_j = \boldsymbol{d}_j + (\boldsymbol{D}\boldsymbol{A}^T)_j\boldsymbol{m'}$.

The computational complexity of the parameter update is of the order $|L| \cdot |I|$. For the complexity of the list update see Subsection 3.3.

3.2 Update of the Covariance Matrix

In the beginning the covariance matrix $\boldsymbol{D}$ is block-diagonal and contains the prior uncertainty of the alignment parameters, derived from laser alignment and mechanical measurements. If required, it may also contain prior correlations between different detector units. After each track, only the blocks in the list $L = \bigcup_{i \in I} L_i$ need to be updated. This is done in the following way:

For all $j, l \in L$ compute: $\widehat{\boldsymbol{D}}_{jl} = \boldsymbol{D}_{jl} + (\boldsymbol{D}\boldsymbol{A}^T)_j(\boldsymbol{G}\boldsymbol{V}_D\boldsymbol{G} - 2\boldsymbol{G})[(\boldsymbol{D}\boldsymbol{A}^T)_l]^T$.

The computational complexity of the update of the covariance matrix is of the order of $|L|^2$. Restricting the size of the lists L_i is therefore of crucial importance. An algorithm is proposed in the following subsection.

3.3 Update of the Lists L_i

First, a relation "$\sim$" between two different detector units i and j is defined:

$i \sim j \iff i$ and j have been crossed by the same track

The relation is symmetric, but not transitive. On the basis of this relation a distance between different detector units i and j can be defined:

If $i \sim i_1 \sim i_2 \sim \cdots \sim i_n \sim j$ is the shortest chain connecting i to j, the distance is $d(i,j) = n + 1$. In particular, if $i \sim j$, then $d(i,j) = 1$.

With the additional definition $d(i,i) = 0$ it is easy to see that the distance d is a proper metrics, i.e. that

1. $d(i,j) = 0$ if and only of $i = j$,
2. $d(i,j) = d(j,i)$, and
3. $d(i,j) \leq d(i,k) + d(k,j)$ for all k.

Using this distance, the following algorithm for updating the lists $L_i, i \in I$ is proposed:

For all $i \in I$ do:

1. For all $j \in I \setminus \{i\}$ do:

   For all $k \in L_j$ with $d(k,j) < d_{\max}$, add k to L_i and store $d(k,i) = d(k,j) + 1$.
2. If a detector k occurs several time in L_i, keep only the occurrence with the smallest distance $d(k,i)$.

The computational complexity of the list update is of the order $|L| \cdot |I|$. It is assumed that the distance $d(k,i)$ is stored along with k in the list L_i. $d_{\max}$ is the largest distance for which correlations are deemed to be significant.

4 Simulation Experiments in a Simplified Setup

4.1 Description of the Setup

The basic properties of the method have been studied in a simple, small setup. It consists of eight detector layers along the z-axis, with a spacing of 10 cm. In each layer, there is a row of five detector units, each $10 \times 10\,\mathrm{cm}^2$. A schematic view of the setup is shown in Fig. 1.

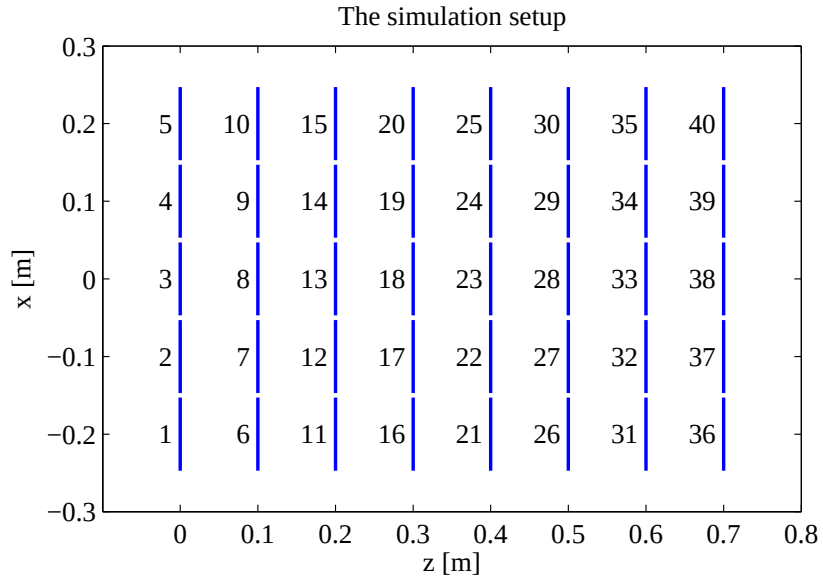


Figure 1: A schematic view of the simulation setup. The direction of the x- and the z-coordinate is shown; the y-coordinate is chosen such that a right-handed coordinate system is formed. The numbers enumerate the detector units.

Straight tracks are simulated in this setup such that each track crosses all detector layers. Neither multiple scattering nor energy loss are simulated. Not all detector units are hit with the same frequency. Figure 2 shows the average number of hits per per track for all detector units. The intersection points of the simulated tracks were smeared by adding an observation error with a standard deviation of 50 μm both in x and in y. The distribution of the observation error was a standard Gaussian with the exception of one simulation run in which a Gaussian mixture was used to study the effect of tails in the observation error.

At least two detector units in different layers have to be fixed to define the reference frame. The effects of taking different pairs of detectors as references will be investigated in the following. All detector units apart from the two reference units were misaligned by shifts in x and y. The shifts were generated randomly by drawing from a Gaussian distribution ten times as wide as the observation error. The positions of the reference units were fixed by giving them a very small prior uncertainty of the order of 0.1 μm. The prior uncertainty of the other units was set to 1 mm. All alignment parameters and the full covariance matrix were updated after each track.

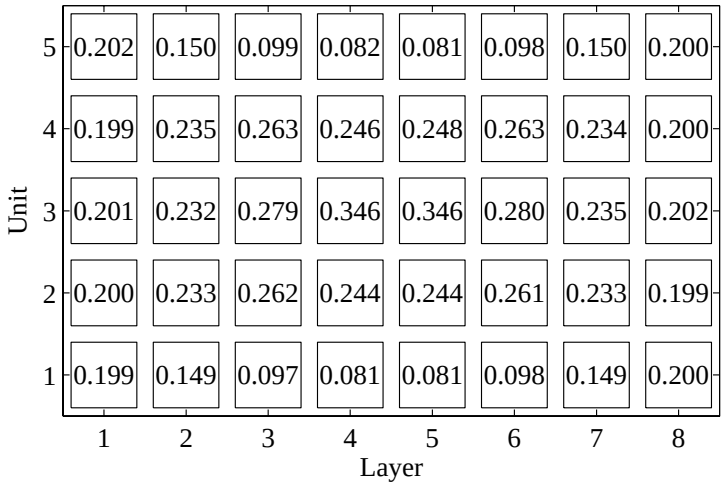


Figure 2: Average number of hits per track versus unit and layer number.

4.2 Precision and Speed of Convergence

The precision of the alignment is measured by the difference $\boldsymbol{\delta}$ between true and estimated shifts:

$$\boldsymbol{\delta} = \widehat{\boldsymbol{d}} - \boldsymbol{d}_{\mathrm{t}}.$$

An example of the evolution of $\boldsymbol{\delta}$ is shown in Fig. 3, separately for x and y shifts. A more quantitative assessment can be made by computing the RMS of $\boldsymbol{\delta}$ after each track. The evolution of the RMS is shown in Fig. 4. In this particular run of 6000 tracks the two reference units were located in the first and last layer. The effects of choosing different locations of the two reference units are investigated in one of the subsections below.

The algorithm provides not only the estimated shifts but also a variance-covariance matrix (error matrix). It has been checked whether the computed errors correspond to the actual deviations by computing the standardized residuals of the estimated shifts with respect to the true shifts. Figure 5 shows their distribution after 500 and 1000 tracks, respectively, for 100 experiments with the same shifts but different tracks. The mean and the standard deviation are very close to their expected value.

The speed of convergence is measured by the number of tracks required to bring the standard errors of all estimated shifts, computed from their variance-covariance matrix, below a certain bound. In the following, we have used a bound of 10 μm. Note that the number of tracks actually seen by a detector unit varies between 10 and 35 percent of the total number of tracks (see Fig. 2).

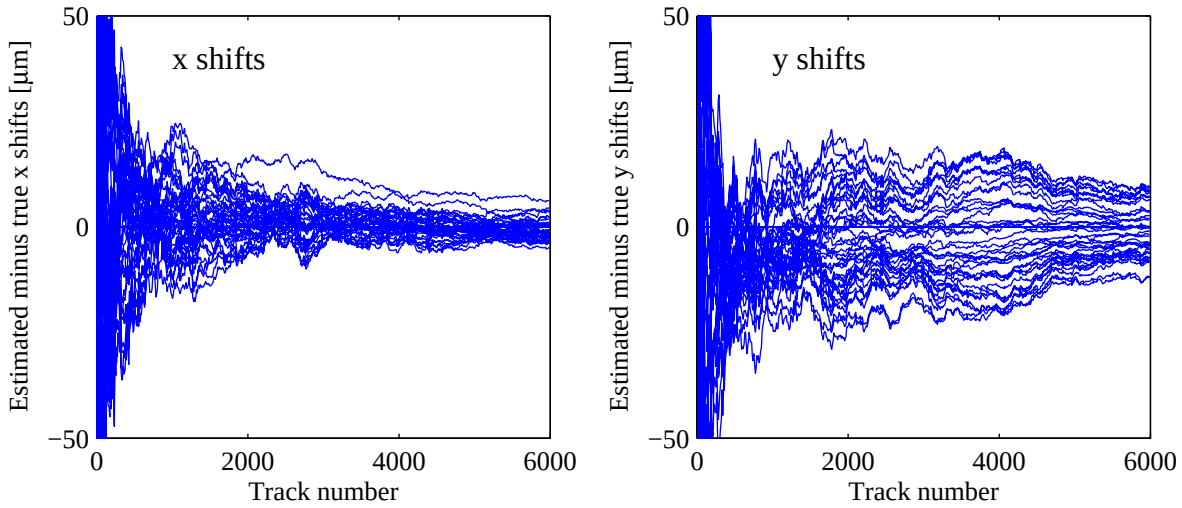


Figure 3: Evolution of the differences between true and estimated shifts.

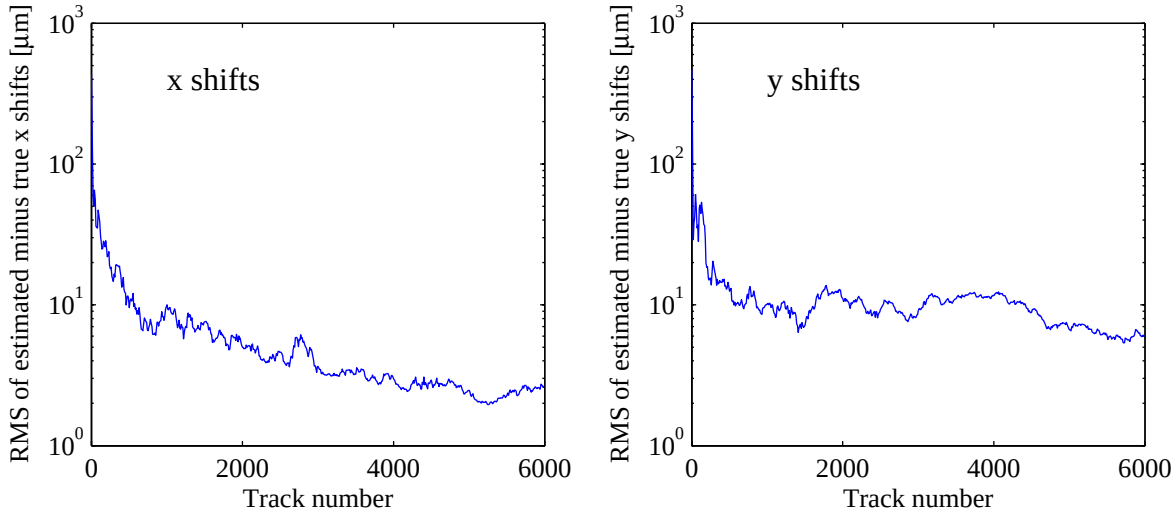


Figure 4: Evolution of the RMS of the differences between true and estimated shifts.

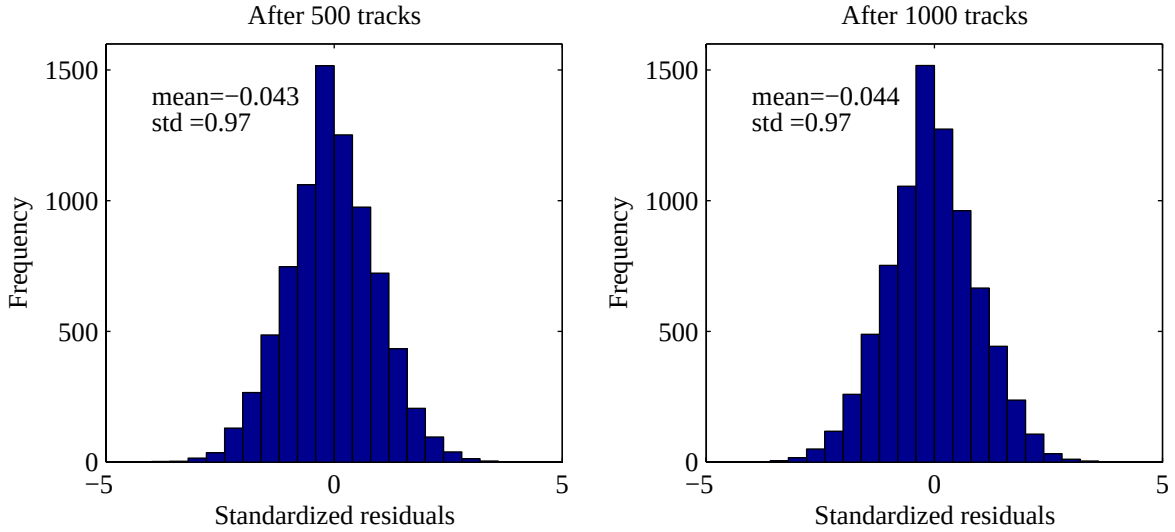


Figure 5: Frequency distribution of the standardized residuals of the estimated shifts w.r.t. the true shifts.

Dependence on the Choice of Reference Detectors

The number and relative position of the reference units has a large influence on the speed of convergence. This influence has been systematically investigated by putting one reference unit in the first layer and a second reference unit in any of the other layers. Figure 6 shows the number of tracks required for convergence as a function of the position of the second reference unit. In the left hand plot, the reference units are at the same edge of the respective layers, whereas in the right hand plot, they are at opposite edges of the respective layers.

As expected, convergence is slowest when the second reference unit is in the layer closest to the first one, and gets faster when the second reference unit is moved to more distant layers. Also, convergence is slower when the reference units are at opposite edges of the respective layers, as in this case they are rarely hit by the same track.

Dependence on the Choice of Starting Values

The estimated shifts should be independent of the initial assumptions about the detector positions. In order to verify this the same sample of tracks has been processed with three different initial values of the shifts:

A: the initial positions are set to the nominal positions ($\boldsymbol{d} = 0$)
B: the initial positions are set to the true positions ($\boldsymbol{d} = \boldsymbol{d}_{\mathrm{t}}$)
C: the initial positions are random, drawn from the same distribution as the actual shifts

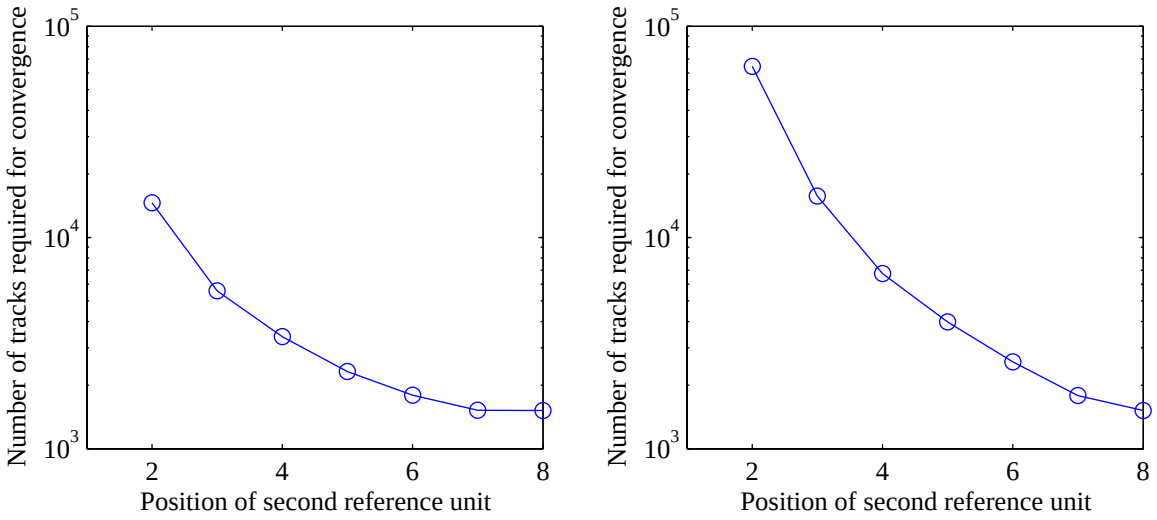


Figure 6: Number of tracks required until convergence. Left hand side: the reference units are at the same edge of the respective layers. Right hand side: the reference units are at opposite edges of the respective layers.

Figure 7 shows the evolution of the RMS of δ under the three assumptions. The reference units are in the first and in the last layer, respectively. It is evident that in this most favourable case the dependence on the starting values has disappeared already after 50 tracks. In the least favourable case, when the reference units are in layers 1 and 2, respectively, it takes several hundred tracks.

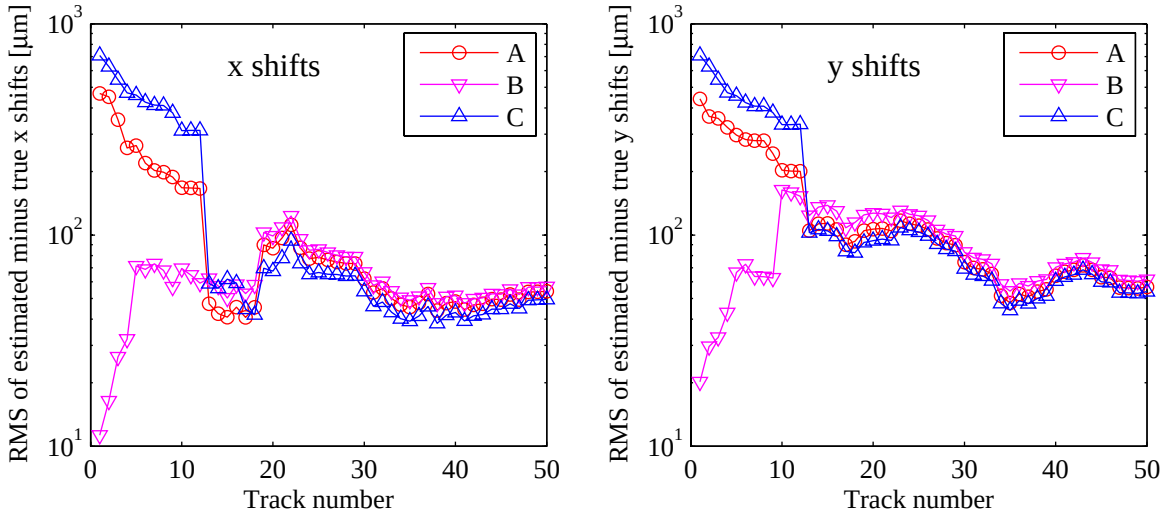


Figure 7: Evolution of the RMS of the differences between true and estimated shifts, for three different starting values of the shifts (see text).

Dependence on the Detector Resolution

It is to be expected that the speed of convergence improves if the precision of the reference units increases. This could be the case, if the reference units are in a pixel detector, while most of the other units are in a strip detector. It is assumed that the reference units are in layer 1 and 2, respectively. Figure 8 shows the evolution of the RMS of δ under three assumptions on the precision of those layers:

A: the precision of the reference layers is equal to the precision of the other layers
B: the precision of the reference layers is twice the precision of the other layers
C: the precision of the reference layers is four times the precision of the other layers

The starting positions are the nominal ones. Again, the behaviour is as expected: better precision of the reference units improves both the speed of convergence and the precision of the estimated shifts.

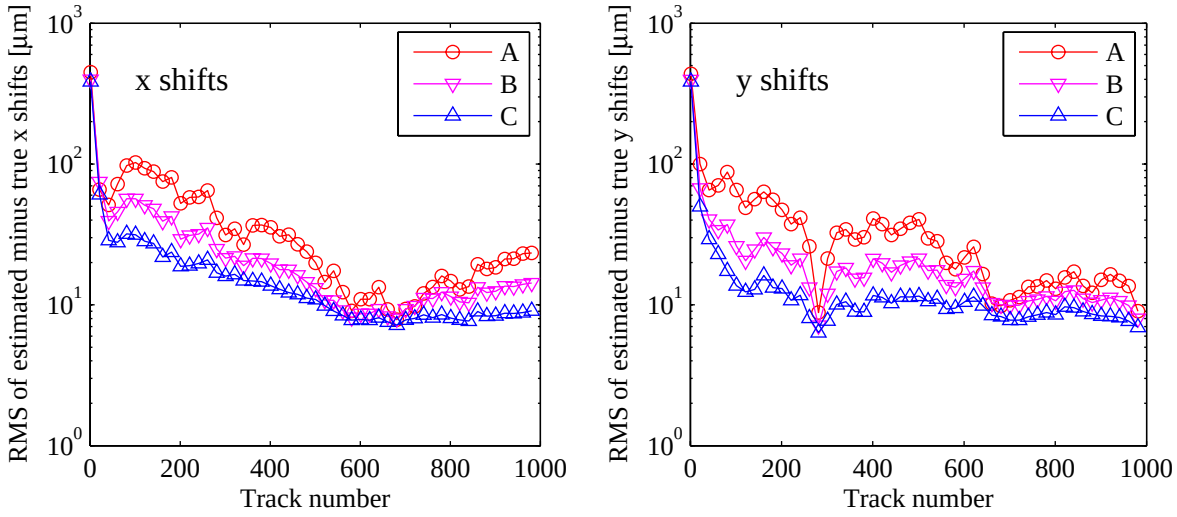


Figure 8: Evolution of the RMS of the differences between true and estimated shifts, for three different precisions of the reference layers (see text).

Effects of Non-Gaussian Smearing

Up to now all the results shown have been obtained with Gaussian smearing of the impact points, using a standard deviation of 50 μm. In the experiment, the distribution of the observation error usually has a Gaussian core with some tails. In order to study the effects of the tails on the convergence of the alignment the Gaussian smearing density was replaced by a two-component Gaussian mixture. The core component has a weight of 90 percent and a standard deviation of 50 μm. The tail component has a weight of 10 percent and a standard deviation of 150 μm, three times as wide as the core. The total standard deviation of the mixture is 67 μm.

Figure 9 shows the evolution of the RMS of δ, the difference between true and estimated shifts, using the same tracks as in Fig. 4. The convergence is somewhat slowed down, because of the larger total width of the observation error. In contrast, the standardized residuals are hardly affected at all. Figure 10 shows their distribution after 500 and 1000 tracks, respectively, for 100 experiments with the same shifts and tracks as in Fig. 5. The mean and the standard deviation are very close to their expected value, and there are no visible tails. If the mixture is scaled such that its total standard deviation is down to 50 μm, the convergence is as fast as with the Gaussian smearing using the same width.

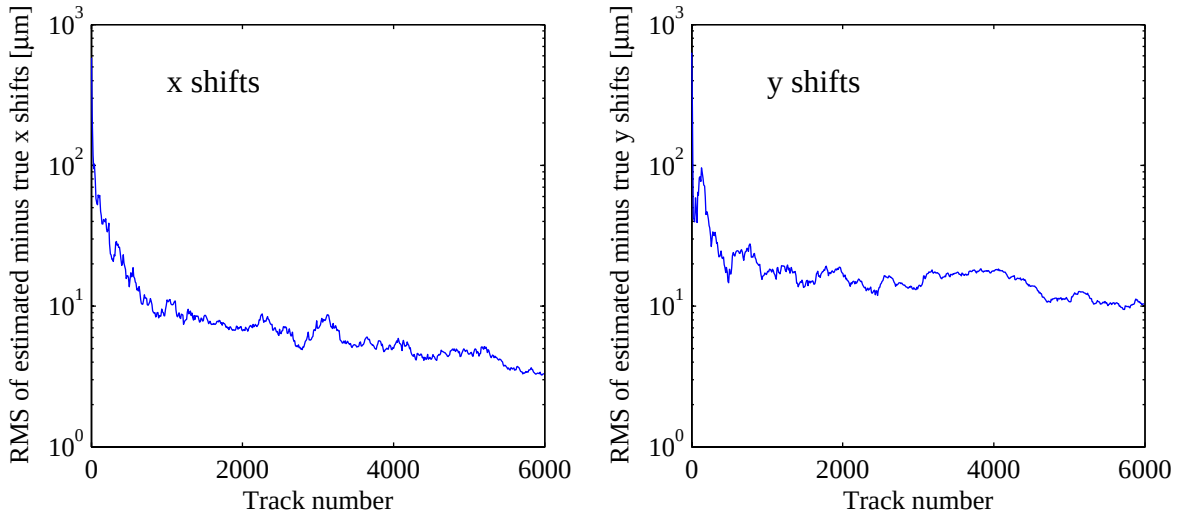


Figure 9: Evolution of the RMS of the differences between true and estimated shifts.

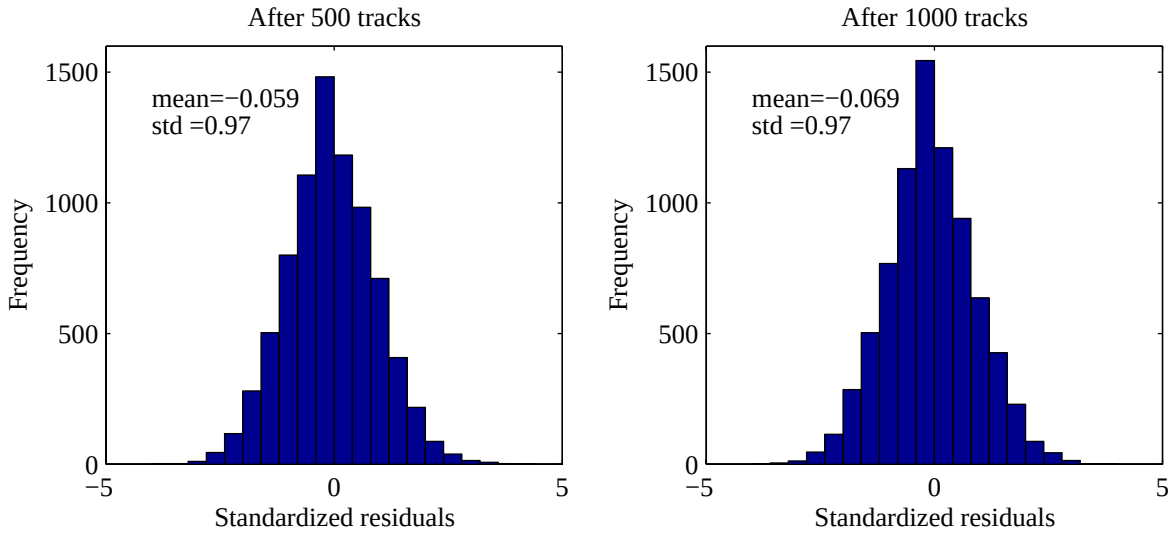


Figure 10: Frequency distribution of the standardized residuals of the estimated shifts w.r.t. the true shifts.

4.3 Correlations between Detector Units

In the beginning, the variance-covariance matrix of the estimated shifts is diagonal. As the tracks are processed, correlations between detector units begin to build up. Depending on the configuration of the reference units, both positive and negative correlations are possible. As an example, Figure 11 shows greyscale-coded plots of the correlation matrix of the x shifts, after 5, 50, 500 and 5000 tracks, respectively. The numbering of the detector units is the one of Fig. 1. In this setup the reference units are at the center of the two central layers (4 and 5). In the plot, the reference units are numbers 18 and 23, easily discernible by their small correlations to the other detectors. The final distinctive pattern emerges already after 50 tracks and changes very little after that.

5 Results from the Barrel Part of the CMS Tracker

5.1 Description of the Setup

The convergence and stability of the method has been investigated in the CMS Tracker [5]. A prototype algorithm was implemented in the CMS reconstruction framework ORCA [6]. Two setups have been used:

- A telescope-like setup (see Fig. 12) containing 14 modules from the Tracker Inner Barrel (TIB) and 30 modules from the Tracker Outer Barrel (TOB). At each event the alignment parameters of all modules and their corresponding covariance matrix are updated.
- A wheel-like setup (see Fig. 13) containing only TIB modules (156 in total). Here the concept of update lists is applied, such that at a given event the update is performed only for the modules with a hit and for the modules that are associated with them via the list (see Subsection 3.3). The update of the alignment parameters is restricted to modules with a distance of at most six from the modules in the current track, according to the definition of Subsection 3.3.

Both setups use the Pixel detector as the reference system, which therefore is not misaligned. The remaining modules are misaligned by shifting them in the local x- and y-directions of the detector plane, where local x is the direction perpendicular to the strips. The shifts are generated randomly by drawing from a Gaussian distribution with $\sigma = 100\,\mu$m. The prior uncertainties of the misaligned modules, i.e. the starting values of the diagonal elements of the variance-covariance matrix $\boldsymbol{D}$, are set to 0.5 mm in x-direction and 0.5 cm in y-direction, respectively. The positions of the reference units are fixed by giving them a very small prior uncertainty of the order of 0.01 μm.

The tracks come from simulated muon and antimuon events (both with $p_T = 100$ GeV), produced with a simplified fast simulation. All simulations are done within a homogeneous 4 Tesla magnetic field and include multiple scattering effects simulated under the same hypotheses as used in the reconstruction. The intersection points are smeared by a Gaussian function using the nominal resolutions for each module type. The tracks are then reconstructed using standard algorithms [7] without any additional information about the misalignment.

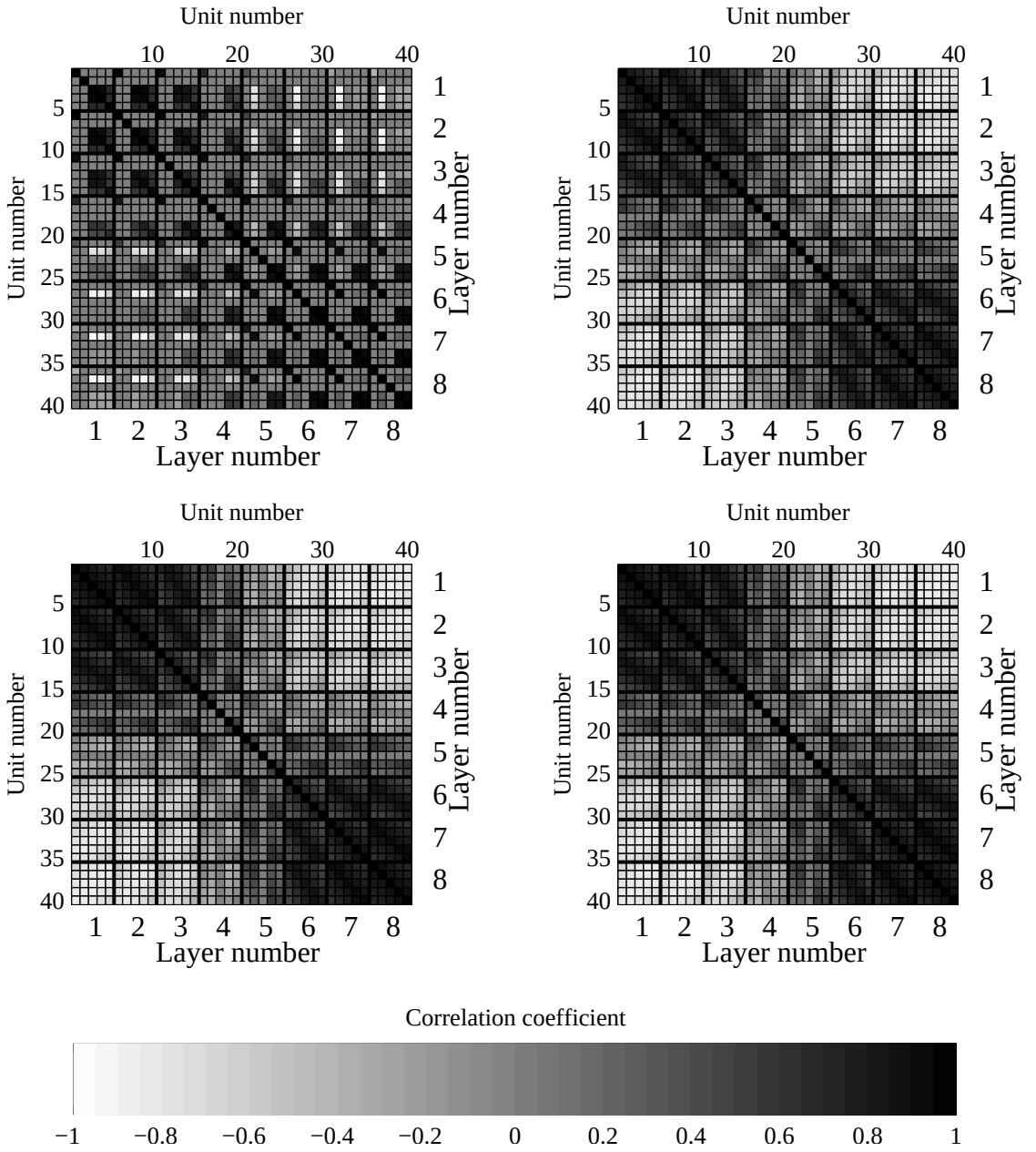


Figure 11: Correlation matrix of the estimated x shifts, after 5 (top left), 50 (top right), 500 (bottom left) and 5000 (bottom right) tracks. The layers are separated by thick black lines. Units 18 and 23 are the reference units.

5.2 Precision and Speed of Convergence

Results from the Wheel-like Setup

Figures 14 and 15 show the evolution of the residuals for the estimates of the x- and y-shifts, respectively. It is obvious that the convergence is slower in the layers that are farther away from the reference system (the Pixel detectors). There are two reasons for this. Firstly, the bigger distance to the reference modules results in bigger uncertainties when propagating information to these layers (compare to section 4.2); secondly, the modules in the outer layers are less often hit by tracks than modules in the inner layers because of the detector geometry, although they are updated with about the same frequency. Nevertheless the convergence is satisfactory, as can be seen from the histograms of the residuals of the x-shifts in Fig. 16.

The choice of the upper distance bound $d_{\max}$ in the the update list is a compromise between computational performance and the attempt to include as many correlations in the update as possible. Figure 17 shows that the choice of $d_{\max} = 6$ does not exclude detector units with significant correlations from the update. A systematic investigation

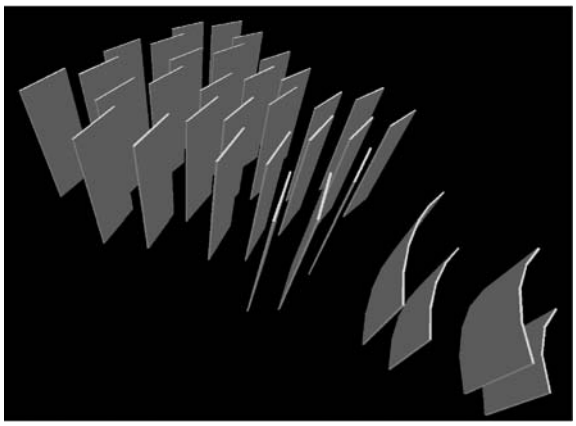


Figure 12: Telescope-like setup.

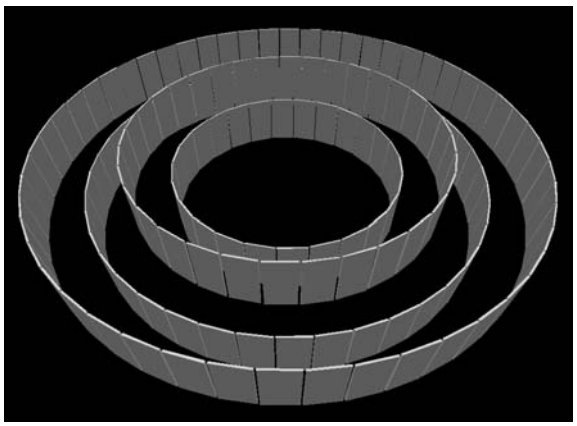


Figure 13: Wheel-like setup.

reveals that in the wheel-like setup a choice of $d_{\max} = 3$ gives in fact the same precision as $d_{\max} = 6$. Table 2 lists the RMS of the final alignment parameters and the time spent in the alignment algorithm for 100,000 tracks, excluding simulation and track reconstruction, as a function of $d_{\max}$. The times have been measured on a 2.2 GHz CPU, an AMD Athlon 64 Processor 3500+.

Table 2: Precision and computing time as a function of $d_{\max}$

$d_{\max}$	1	2	3	4	5	6
σ [μm]	24.75	21.38	20.97	20.95	20.94	20.94
T [s]	472	604	723	936	1152	1319

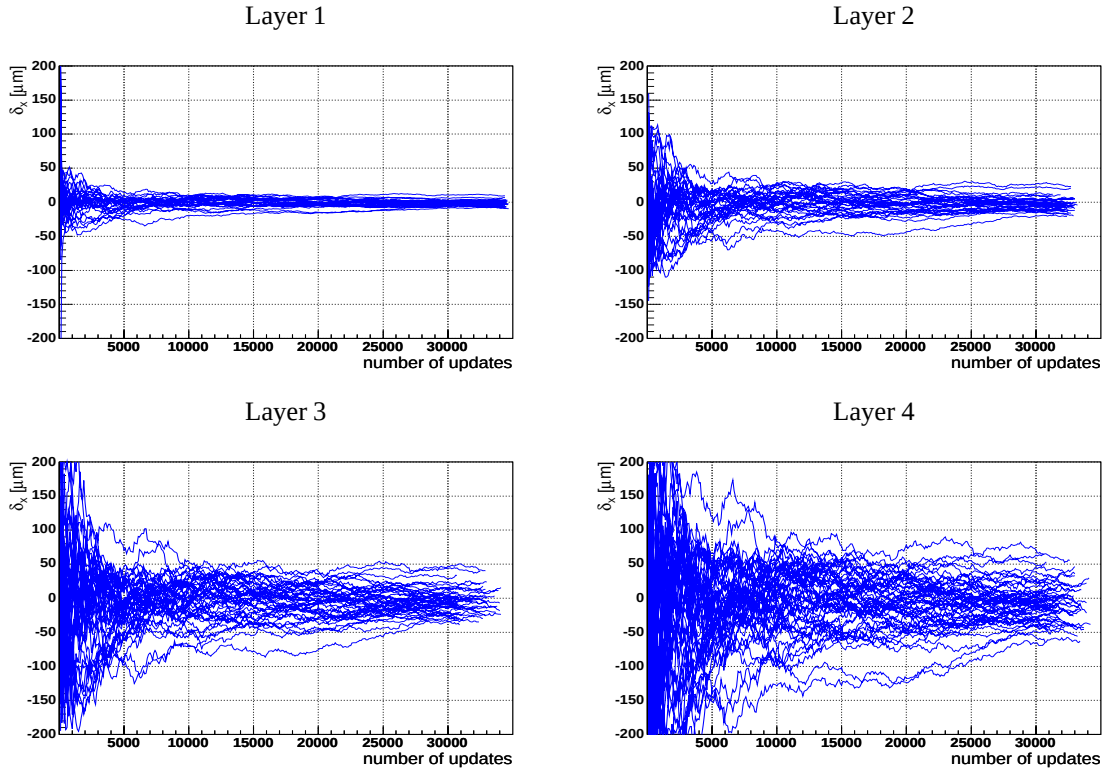


Figure 14: Evolution of the local x-shift residuals for the wheel-like setup. In total 100,000 tracks were processed. Here only correlations between modules with $d(i, j) \leq 6$ were considered.

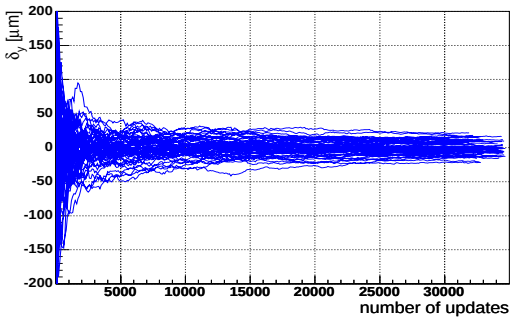


Figure 15: Evolution of the local y-shift residuals for stereo modules from the wheel-like setup. In total 100,000 tracks were processed. Here only correlations between modules with $d(i,j) \leq 6$ were considered.

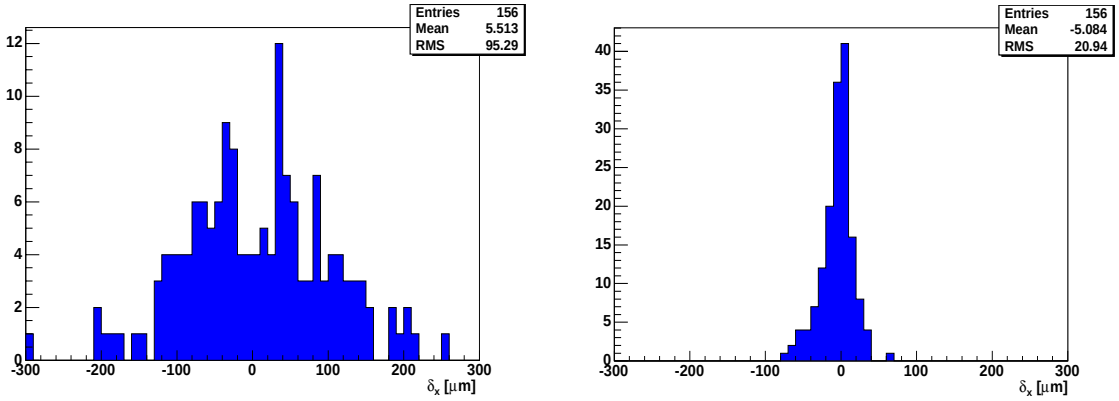


Figure 16: Residuals of the local x-shift estimates before alignment (left) and after 100,000 processed tracks (right). Here only correlations between modules with $d(i,j) \leq 6$ were considered.

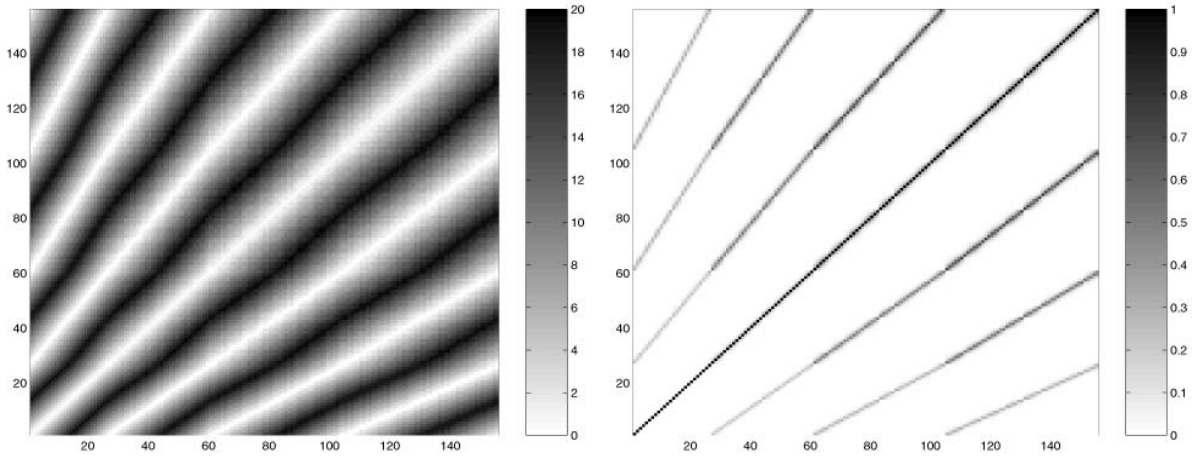


Figure 17: Grayscale-coded visualization of the metrics (left) and the correlation matrix (right) for all modules in the wheel-like setup. Comparing these figures shows that the choice of $d(i,j) \leq 6$ doesn't exclude modules with relevant correlations during update. The modules are ordered by layer and increasing (global) polar angle and indexed from 1 to 156.

Results from the Telescope-like Setup

In this setup modules in the Inner Barrel and in the Outer Barrel of the CMS Tracker were aligned. As expected from the studies in the simplified setup and in the wheel-like setup, convergence is much slower for modules farther away from the reference system. As a consequence, the convergence for TOB-modules does not look as smooth as for TIB-modules, if plotted on the same scale (see Fig. 18). Nevertheless, the standardized residuals are well within the prescribed limits, which shows that the algorithm is working properly (see Fig. 19). The figures also show that the correlation between modules is higher in this setup than in the wheel-like setup, because of their small distance in terms of the metrics d. Figure 20 shows the estimates on the local y-shifts for stereo modules in the Inner and Outer Barrel, respectively. The residuals of the alignment parameters at the start and after 100.000 processed tracks are shown in Fig. 21. The RMS of the final estimates is down to about 2.5 μm, which is well below the required precision.

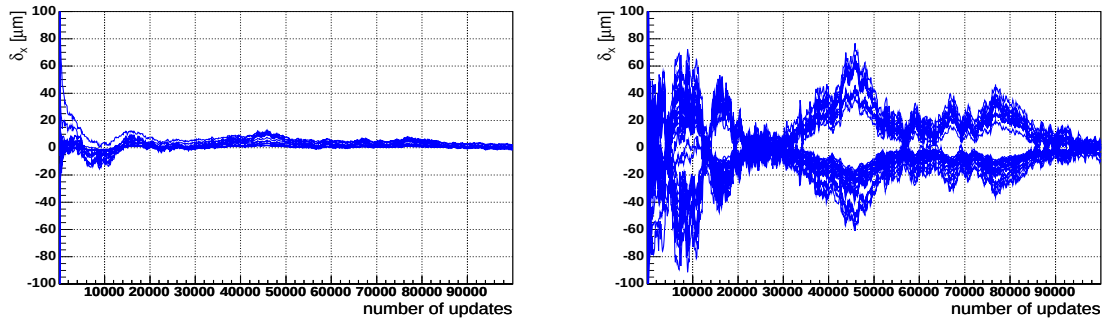


Figure 18: Residuals of the local x-shift estimates for modules in the TIB (left) and the TOB (right). In the TOB the strong correlations of the estimated shifts are visible. The two sets of curves reflect the two different orientations of the local x-axis w.r.t. the global system.

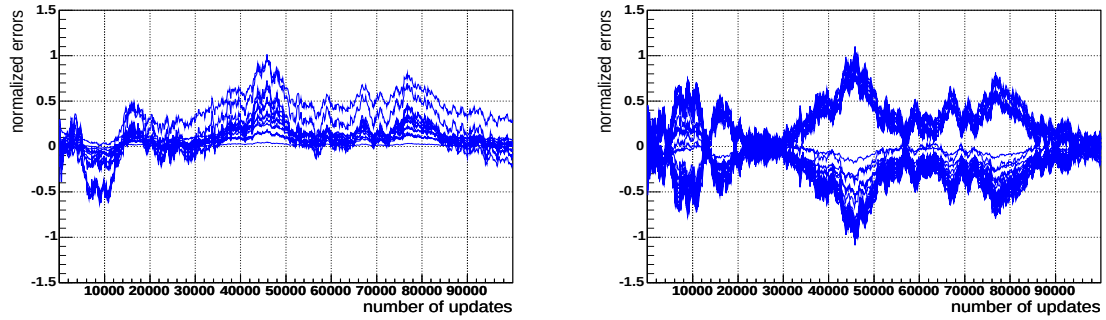


Figure 19: Standardized residuals of the local x-shift estimates for modules in the TIB (left) and the TOB (right).

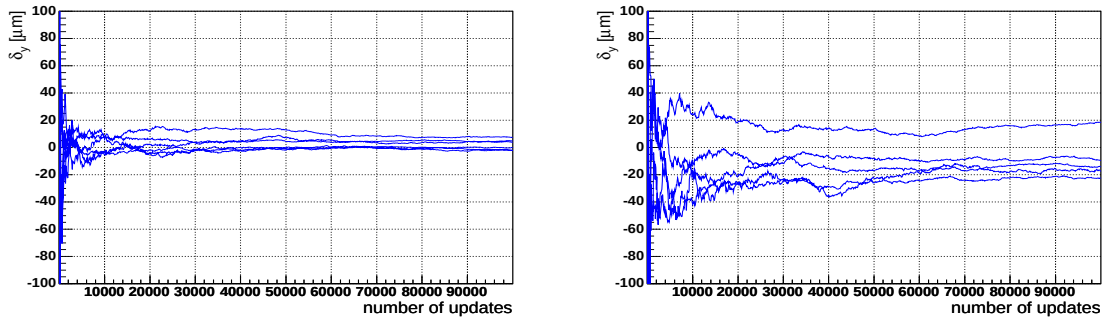


Figure 20: Residuals of the local y-shift estimates for stereo modules in the TIB (left) and the TOB (right).

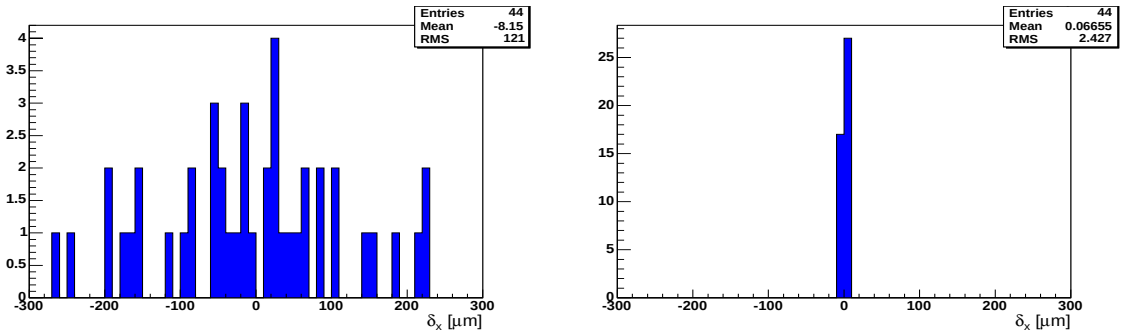


Figure 21: Residuals of the local x-shift estimates before alignment (left) and after 100,000 processed tracks (right).

6 Conclusions and Outlook

A general algorithm for the track-based global alignment of a complex detector system has been presented. The algorithm is derived from the Kalman filter and is designed to avoid the inversion of large matrices. Some basic properties of the method have been explored in a simplified detector. A prototype implementation in ORCA is operational and has been successfully tested on two small sections of the barrel part of the CMS Tracker. Although the method has been shown to work in principle, clearly more development, testing and tuning is required to meet the challenge of a full alignment of the Tracker.

- The viability of the concept of the correlation lists, based on the metrics d, has to be tested on a larger number of modules, and the distance cut $d_{\max}$ has to be optimized. This is particularly important in view of the influence of the maximal distance on the computation time.
- Alternative ways of defining the range of the update will be explored.
- The scaleability of the algorithm and in particular the evolution of the correlations has to be studied on a larger number of modules.
- The simplified fast track simulation has to be replaced by a full simulation.
- In view of the slower convergence of the TOB modules, alternatives to using single tracks are desirable. Using constrained muon pairs from Z- or J/ψ-decays is an auspicious possibility. The extra information from the vertex and kinematics contraints can be fed in via the prior track information. Alternatively, the algorithm can be extended to deal with several tracks instead of a single track. Both possibilities will be implemented and tested.

References

[1] D.E. Catlin, *Estimation, Control, and the Discrete Kalman Filter*. Springer, New York, 1989.

[2] R. Frühwirth, *Nucl. Instrum. and Meth.* **A 262** (1984) 444.

[3] R. Frühwirth, T. Todorov and M. Winkler, *J. Phys. G: Nucl. Part. Phys.* **29** (2003) 561.

[4] R. Zurmühl and S. Falk, *Matrizen und ihre Anwendungen*, 5th edition. Springer, Berlin–Heidelberg, 1984.

[5] The CMS Collaboration, *The Tracker Project Technical Design Report*, CERN/LHCC 98-6.
The CMS Collaboration, *Addendum to the Tracker TDR*, CERN/LHCC 2000-016.

[6] The reference manual and the user guide can be found on [http://cmsdoc.cern.ch/orca].

[7] W. Adam et al., *Track Reconstruction in the CMS Tracker*, CMS Note in preparation.