First studies of T-station alignment with simulated data

A. Hicheur, M. Needham, L. Nicolas
Laboratoire de Physique des Hautes Energies, École Polytechnique Fédérale de Lausanne

J. Amoraal, W. Hulsbergen, G. Raven
NIKHEF, Amsterdam

March 19, 2009

Abstract

The alignment of the tracking stations using Kalman-fitted tracks from the standard LHCb track fit has been studied. The alignment procedure is briefly presented, followed by two practical use-cases using Monte Carlo simulated data. The first scenario presents the alignment of the IT and OT layers using beam-gas events at 450 GeV with no magnetic field. The second scenario is the alignment of the OT layers and IT boxes, layers and ladders with magnet-on minimum bias events at nominal energy. The results are validated by comparing the $J/\psi$ and $K_s$ mass resolution and bias obtained with the misaligned and aligned geometries. It is shown that the $J/\psi$ mass resolution is not degraded after alignment.
1 Introduction

The LHCb tracking system (see Fig. 1) consists of a Vertex Locator (VELO) and a large area silicon detector (Trigger Tracker) located upstream of the dipole magnet of the experiment and three stations (the T-stations) located downstream of the magnet. The latter are divided into the Inner and Outer Tracker (IT and OT). The three IT+OT stations are composed of 3 stations × 4 boxes × 4 layers × 7 = 336 silicon micro-strips IT ladders and 3 stations × 4 layers × 22 modules = 264 straw-tube OT modules, respectively.

Performance studies are concentrated on so called ‘Long tracks’ that traverse the entire spectrometer [1]. For these tracks a momentum resolution of between 4% is achieved with simulated data. To achieve this precision with real data the position of the sensors has to be determined with an accuracy well below the hit resolution, \( \sim 60 \, \mu m \) and \( \sim 200 \, \mu m \), for IT and OT, respectively.

2 Alignment method

The method used to fit for the alignment parameters is the minimisation of a \( \chi^2 \), based on residuals \( r \) of reconstructed tracks, with respect to both alignment and track parameters (see [2, 3]). The minimisation leads to an
expression for the alignment parameters $a$ as follows:

$$
a = -\left( \sum_t \frac{\partial r_t}{\partial a} V^{-1} (V - HCH^T) V^{-1} \frac{\partial r_t}{\partial a} \right)^{-1} \times \left( \sum_t \frac{\partial r_t}{\partial a} V^{-1} r_t \right)
$$

The sum is made over the tracks in the sample. $\frac{\partial r_t}{\partial a}$ represents the derivatives of residuals with respect to the alignment parameters. $V$ is the measurements (diagonal) error matrix. $V - HCH^T$ is the covariance matrix of the residuals, $H$ is the derivative of the residuals with respect to the track parameters, and $C$ is the track parameter covariance matrix. The novelty in this study is that we use a Kalman filter track fit in the alignment procedure, following the recipe in [4] to compute the global covariance matrix $C$. This allows us to use the standard LHCb Kalman-based track-fitting tools and data model in the alignment procedure. This has the advantage that energy loss and multiple scattering in the detector are correctly accounted for.

3 Software and procedure

The software was developed in the LHCb Gaudi framework [5]. The core alignment algorithm, which processes the track sample and fills the algebra objects, is used in conjunction with a set of tools performing tasks such as track selection, spectral analysis, regularisation of the problem and solving.

The general sequence (see Fig. 2) starts with pattern recognition, followed by track fit, track quality selection, computation of the alignment derivatives, solving and update of the alignment constants. To obtain an analytic solution, the method assumes a linear dependence of the residuals with respect to alignment parameters. To resolve non-linearities, we iterate over the sequence until convergence. Convergence is obtained when the total $\chi^2$ shows no significant change in subsequent iterations.

The framework is flexible enough to allow several options for alignment: one can align for groups of sensors or decide to turn off degrees of freedom which we are not sensitive to.
Figure 2: Flow diagram of the software procedure used for the tracking-stations alignment. The pattern recognition can optionally be re-run at each iteration.

3.1 Weak mode suppression

The $\chi^2$ minimum reached by the alignment method may be invariant under linear combinations of alignment parameters, known as weak modes. These are global translations or rotations of the full block of T-stations, or more complicated distortions. To avoid the influence of weak modes on the solution, constraints are applied. Global translations, rotations and shearings cannot be aligned for and should therefore be removed in the solving procedure. This could be done with Lagrange multipliers or by performing a spectral analysis (removing the eigenvectors corresponding to the weak modes). We can also include in the procedure the fit to the origin vertex of the tracks, which helps constraining further the alignment parameters. However, since the origin vertex is located in the Velo region, vertex fitting has little influence on the T-stations alignment. These constraints are described in detail in [6].

Another way of constraining one sub-detector is to align for it with respect to another fixed sub-detector. For example, aligning for the T-stations using long tracks fixes both the Inner and Outer tracker to the nominal position of the Velo and Trigger Tracker. This corresponds to implementing simultaneously the T-stations internal alignment and the Velo–T-stations relative alignment.
Finally, the alignment of a sub-detector can also be constrained by fixing one or several of its components.

4 Event and track selection

Since the alignment $\chi^2$ uses primarily track information, it is sensitive to the effect of ghost tracks and particles with kinks in their trajectory due to hadronic interactions in the detector material. If these tracks are used in the alignment procedure they can cause the algorithm to converge to a false minimum. Therefore, it is important to obtain a sample of tracks with a minimum contamination of ghost and other bad tracks.

4.1 Cuts on track quality

One way of dealing with bad tracks is to cut on the track fit $\chi^2$. For the long tracks it is also possible to cut on the $\chi^2$ of the track fit match: $\chi^2_m$. This is defined as

$$\chi^2_m = \chi^2_{tot} - \chi^2_T - \chi^2_{Velo},$$

where $\chi^2_T$ is the $\chi^2$ of the segment in the tracking stations and $\chi^2_{Velo}$ is the $\chi^2$ of the segment in the Velo. This is true in our case because we do not use the TT hits in the fit. These variables need to be used with care as selection criteria since their quality is degraded in a misaligned detector. Therefore, an evolving strategy has been developed. In the first iteration of the procedure a loose cut on the track fit $\chi^2$ is made which is then tightened in subsequent iterations. To develop this strategy the distributions of selected, not selected (ie tracks which are known to have interactions) and ghosts have been studied with the ideal geometry and the misalignment scenario described in Section 6.1. Fig 4 shows the distribution of the $\chi^2$/dof in the misaligned case. It can be seen that cutting on the track fit $\chi^2$/dof at 100 will not bias the sample of selected tracks. Therefore, this is chosen as the starting value in the iterative procedure. Fig 3 shows the same distribution in the case of the ideal detector. From this plot it can be seen that a reasonable cut to apply at the end of the scheme is $\chi^2$/dof < 10.

Figs. 5 and 6 show the same plots for the case of $\chi^2_m$. In this case an initial cut value of 100 seems reasonable. This is reduced to 30 during the iteration procedure.
Figure 3: Distribution of track fit $\chi^2$/dof for (a) selected, (b) not-selected and (c) ghost tracks with the ideal geometry.

Tables 1 and 2 summarise the performance of the cuts described above in removing ghost\(^1\) for long tracks using the misaligned geometry (initial case)

\(^1\)Ghost tracks are tracks which cannot be associated to any Monte Carlo simulated track. They can be for example fake matches between Velo- and T-segments, results of hadronic interactions, results of fake combinations of hits or decays in flight. See [8] for a
Figure 4: Distribution of track fit $\chi^2$/dof for (a) selected, (b) not-selected and (c) ghost tracks with the misalignment scenario presented in 6.1.

and the ideal geometry (final case) respectively. Tables 3 and 4 show the same results in the case of T-tracks.

These tables and plots don’t show a clear possibility for reducing the ghost detailed study.
rate and the number of interaction tracks. This is due to the fact that we consider tracks going both through the Inner and the Outer Trackers. Tracks going only through the IT are much more sensitive to the cuts on the track $\chi^2$/dof and on the track fit match $\chi^2$. This has been studied in Appendix A. Due to the better resolution of the IT, this fact is very important as a high
Figure 6: Distribution of track $\chi^2_m$ for (a) selected, (b) not-selected and (c) ghost tracks with the misalignment scenario presented in Section 6.1.

purity in this sub-detector fixes the Outer Tracker to better results.
Table 1: Efficiencies of selection of good tracks and interactions and remaining ghost rate of the cuts on the long track $\chi^2$ and $\chi^2_m$ in the misaligned case.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Cut</th>
<th>Efficiency</th>
<th>Ghost rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Selected</td>
<td>Not selected</td>
</tr>
<tr>
<td>Minimum bias</td>
<td>No cut</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td></td>
<td>$\chi^2$/dof &lt; 100</td>
<td>100 %</td>
<td>99.99 %</td>
</tr>
<tr>
<td></td>
<td>$\chi^2_m &lt; 100</td>
<td>99.07 %</td>
<td>93.68 %</td>
</tr>
<tr>
<td>L0 stripped</td>
<td>No cut</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td></td>
<td>$\chi^2$/dof &lt; 100</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td></td>
<td>$\chi^2_m &lt; 100</td>
<td>98.92 %</td>
<td>95.06 %</td>
</tr>
</tbody>
</table>

Table 2: Efficiencies of selection of good tracks and interactions and remaining ghost rate of the cuts on the long track $\chi^2$ and $\chi^2_m$ in the aligned case.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Cut</th>
<th>Efficiency</th>
<th>Ghost rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Selected</td>
<td>Not selected</td>
</tr>
<tr>
<td>Minimum bias</td>
<td>No cut</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td></td>
<td>$\chi^2$/dof &lt; 10</td>
<td>99.99 %</td>
<td>99.55 %</td>
</tr>
<tr>
<td></td>
<td>$\chi^2_m &lt; 30</td>
<td>98.23 %</td>
<td>82.83 %</td>
</tr>
<tr>
<td>L0 stripped</td>
<td>No cut</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td></td>
<td>$\chi^2$/dof &lt; 10</td>
<td>100 %</td>
<td>99.82 %</td>
</tr>
<tr>
<td></td>
<td>$\chi^2_m &lt; 30</td>
<td>98.52 %</td>
<td>85.14 %</td>
</tr>
</tbody>
</table>

Table 3: Efficiencies of selection of good tracks and interactions and remaining ghost rate of the cuts on the T-track $\chi^2$ in the misaligned case using beam–gas data.

<table>
<thead>
<tr>
<th>Cut</th>
<th>Efficiency</th>
<th>Ghost rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Selected</td>
<td>Not selected</td>
</tr>
<tr>
<td>No cut</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td>$\chi^2$/dof &lt; 100</td>
<td>100 %</td>
<td>99.98 %</td>
</tr>
</tbody>
</table>

4.2 Other cuts

Several other cuts were investigated to reduce the ghost rate further and obtain better convergence of the algorithm. The results shown below have
Table 4: Efficiencies of selection of good tracks and interactions and remaining ghost rate of the cuts on the T-track $\chi^2$ in the aligned case using beam–gas data.

<table>
<thead>
<tr>
<th>Cut</th>
<th>Efficiency Selected</th>
<th>Efficiency Not selected</th>
<th>Ghost rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cut</td>
<td>100 %</td>
<td>100 %</td>
<td>6.71 %</td>
</tr>
<tr>
<td>$\chi^2$/dof &lt; 10</td>
<td>99.96 %</td>
<td>99.11 %</td>
<td>6.55 %</td>
</tr>
</tbody>
</table>

been obtained on a sample of 20'000 minimum bias long tracks reconstructed with the ideal geometry. The tracks are required to cross at least one element of the tracking stations.

It has been shown for the Long Tracking [7] that most ghosts originate in the matching between the Velo- and T-seeds. These ghosts appear when there are many reconstructed segments in either of the sub-detectors. In order to reduce this source of ghosts, events with many Velo or Inner Tracker clusters are rejected. High Outer Tracker occupancy is not as big a problem as the IT case because of the lower track density. The cuts are chosen to reduce the ghost rate to an acceptable level. For the time being, 400 IT clusters and 900 Velo clusters are chosen. These cuts are summarised in Table 5.

Poorly reconstructed tracks also have a negative impact on the alignment results, as discussed in section 4.1. For example, tracks with large multiple scattering in the detector will tend to be badly fitted. As this process most strongly effects low energy tracks, it can be reduced by cutting on the track momentum at 10 GeV (see Table 5).

Finally, the ghost rate can be further reduced by cutting on the track pseudo-rapidity. The LHCb acceptance in this variable is between 1.9 and 4.9. Ghosts tend to have a large or even unphysical pseudo-rapidity in the Velo [8]. A reasonable cut is at $\eta < 5.2$.

Combining these four cuts leads to an overall selection efficiency for real tracks of 11.57 % with a ghost rate of 7.75 %.
Table 5: Summary of the cuts applied to reduce the ghost rate.

<table>
<thead>
<tr>
<th>Cut</th>
<th>Efficiency</th>
<th>Ghost rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cut</td>
<td>100 %</td>
<td>17.48 %</td>
</tr>
<tr>
<td># IT Clusters &lt; 400</td>
<td>55.99 %</td>
<td>11.90 %</td>
</tr>
<tr>
<td># Velo Clusters &lt; 900</td>
<td>38.45 %</td>
<td>10.60 %</td>
</tr>
<tr>
<td>P &gt; 10 GeV</td>
<td>35.32 %</td>
<td>15.16 %</td>
</tr>
<tr>
<td>η &lt; 5.2</td>
<td>99.98 %</td>
<td>16.73 %</td>
</tr>
</tbody>
</table>

5 Alignment with beam–gas interactions

During the machine startup only one beam will circulate in the beam pipe. In this period LHCb will acquire events from beam–gas interactions occurring in front of the detector, giving the opportunity to calibrate, commission and align the detector. In the following sections, the scenario for which the alignment has been studied is reported along with the achieved results.

5.1 Scenario

The Monte Carlo data used for this test is a sample of 50,000 beam–gas minimum bias events with a simulated centre-of-mass energy of 450 GeV. This data sample was reconstructed using the ideal geometry. However, the alignment procedure is performed using a misaligned condition database. This database was obtained by misaligning both the Inner Tracker and Outer Tracker layers with translations along the measurement axis and rotations around both the beam and vertical axes. The amplitude of the misalignments for each of the alignable detector elements has been chosen to follow Gaussian distributions with widths of 0.3 mm (IT Tx), 2.5 mm (OT Tx), 2.5 mrad (IT rotations) and 1 mrad (OT rotations). The details of this misalignment scenario are given in Table 6.

No misalignment has been added for the translations along the strips and beam axes nor for the rotation around the measurement direction and vertical axis as these degrees of freedom are those to which the alignment procedure is the least sensitive. The amplitude chosen for the misalignments reflects a reasonable day–1 scenario.

To simulate the situation of first day alignment, a standalone T-stations
<table>
<thead>
<tr>
<th>Detector</th>
<th>dof</th>
<th>Amplitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT layers</td>
<td>Tx [mm]</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>Rz [mrad]</td>
<td>2.5</td>
</tr>
<tr>
<td>OT layers</td>
<td>Tx [mm]</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>Rz [mrad]</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 6: Summary of the misalignments applied in the beam–gas scenario.

alignment has been performed. The T stations were aligned using T-tracks, removing the constraint to the Velo that would appear if using long tracks. In addition, the magnet is assumed to be switched off, removing the knowledge of the momentum and hence the possibility to cut on this variable. This also means that we cannot properly account for multiple scattering leading to larger track parameter errors and therefore a loss in statistical precision.

Both the Inner and Outer Tracker have been aligned at the level of layers, for the two degrees of freedom that have been misaligned\(^2\). In order to remove weakly constrained degrees of freedom, four IT layers in each stack of boxes and four OT layers are frozen to 0. These four layers are the first X and stereo layers as well as the last stereo and X layers.

In order to gain in precision, the drift time information should be used for the OT. Without this, the resolution decreases from 200 \(\mu\text{m}\) to 2.5 mm/\(\sqrt{12}\) \(\sim\) 1 mm. However, if the drift-time is used, the left-right ambiguity of the measurement must be resolved. Unfortunately, the convergence of the algorithm is less stable in this case. The strategy chosen to deal with this is to first get close to the minimum by running 4 iterations with drift times off and to then use the power of drift time information for the remaining iterations.

5.2 Results

There are several possibilities to monitor the performance of the alignment jobs. As described in section 3, the alignment procedure minimises the total track \(\chi^2\) with respect to both the track and alignment parameters. The first way of looking at alignment results is to check that the normalised total sum of track \(\chi^2/\text{dof}\) has indeed converged. Fig. 7 shows the convergence of this parameter as a function of the iteration number.

\(^2\)The four other degrees of freedom are frozen.
Two different regions can be seen on this plot. The first one covers the first four iterations where we see a clear drop in the total $\chi^2$. The second region, from iteration 4 to 8, shows a big jump and then a plateau. The plateau indicates that the procedure has converged. As explained in section 3, this does not mean that the detector elements have reached a stable position, due to the weak modes. We will see below that it is the case in our example. The jump is explained by the OT drift times being turned on after four iterations. Turning on the drift times indeed decreases the hit error and thus increases the $\chi^2$.

It is clear that the alignment algorithm has found a minimum of the total track $\chi^2$. However, we still need to check if we have found a stable position of our detector elements or if we are suffering from weak modes. This is done by looking at the flow of the alignment parameters as a function of the iteration number. In this note, we concentrate on the translation in the 'X' direction, which is the measurement direction and the parameter we are the most sensible to. Fig. 8 (a) and (b) show this flow for all the IT and OT layers (each line representing one layer). The displacement shown in the iteration 0 corresponds to the input misalignment set in the conditions. It appears as a first correction to the default geometry and the goal of the alignment is to have all the curves converging to 0. The curves then show the value of the Tx alignment parameter for each detector element, including the correction
given by the current iteration. Fig. 8 (b) is a zoom on fig. 8 (a).

![Figure 8: Convergence of the Tx alignment parameter for each of the IT and OT layers (beam gas scenario). (b) is a zoom of (a).](image)

On the left-hand-side plot of fig. 8, we see that the OT layers, although misaligned by a large amount in the beginning converge well towards 0 in 5 to 6 iterations. All layers end up within 10 μm of the ideal geometry, to which it should converge. We will return to the discussion of the results in Section 5.3.

In detail, we can see that the IT layers converge in 3 to 4 iterations to less than 2 μm of the ideal geometry. This corresponds to less than 5% of the IT resolution (which is \( \frac{200}{\sqrt{12}} \approx 57.7 \) μm). Although this looks like a very good result, we still need to validate it. This will be the subject of section 5.4.

Another way of looking at the results is to calculate the residual misalignment after the alignment procedure has converged. This is done by subtracting the output misalignment (as given by the alignment algorithm) to the input misalignment set in the conditions. Fig. 9 shows this value for all the IT and OT layers.

These plots correspond to a projection on the y axis of the last point of each graph in fig. 8. The points with no error bars sitting on the line at 0
residual misalignment are the layers that were kept fixed during the alignment procedure as constraints.

Another constraint that was not addressed before is the constraint on IT coming from the fact that we simultaneously align IT and OT layers. This constrains the IT layers to the position of the OT layers through the overlap regions between the two sub-detectors. However, when aligning only the IT layers, we see a similar precision, meaning that this constraint to the OT layers doesn’t improve much the alignment results. This is not a surprise as the constraint is mainly driven by the intrinsic resolution, which is about 4 times better for IT than for OT.

5.3 Further studies

There are two issues that need to be addressed at this point. The first one is that the convergence is achieved in 5 iterations. With the method in use, we expect to need only 1 to 2 iterations. The second issue is the fact that some OT layers have a final position at about 10 standard deviations from the true position (ideal geometry). In other words, the error on the alignment parameters seem to be under-estimated. This is quite surprising and has been studied in detail in the following sub-sections.
5.3.1 Effect of drift time

The first effect that was studied is the influence of the use of OT drift times in the alignment procedure. Using drift times gives a better resolution, but slows the convergence of the algorithm. The results shown in Figs. 8 (with drift times) and 9 can be compared to those shown in Fig. 10 (no drift times). We see here that the convergence is not slower. This means that the 4 iterations were sufficient to retrieve all the information necessary for the alignment procedure. However, the second consequence of using drift times can indeed be seen. The precision of the final position after alignment for the OT layers is increased by a factor 10 when using the power of the drift times information. This is a straight-forward proof that this information is needed for the alignment of the Outer Tracker.

![Figure 10: (a) Alignment parameters evolution and (b) (Input - output) Tx misalignment for each of the IT and OT layers and (beam gas scenario, no drift times).](image)

5.3.2 Effect of the cut on track quality

The cut on the track quality has been discussed in detail in Section 4.1. However, we don’t know yet the effect of such a cut on the results of the alignment procedure. The same alignment scenario as shown in Section 5.2
has been run without the evolving cut on the track $\chi^2$/dof (see Fig. 11) and with a fixed cut at $\chi^2$/dof $< 10$ (shown in Fig. 12).

![Graphs showing alignment parameters evolution and input-output Tx misalignment](image)

Figure 11: (a) Alignment parameters evolution and (b) (Input - output) Tx misalignment for each of the IT and OT layers and (beam gas scenario, no cut on the track $\chi^2$/dof).

These two results show two very important things. First, the alignment procedure suffers from poor quality tracks, which are the consequence of a very misaligned detector. The alignment procedure never recovers from the effect of these tracks. We can also see that the IT suffers more from these bad tracks. This is not a surprise as the IT resolution is about 4 times better than the OT resolution.

The second thing we learn from these results is that cutting too hard on the track quality in the first iteration also worsens the results. This is due to the fact that with a very misaligned detector, cutting too hard on the track $\chi^2$/dof rejects a large number of tracks, which biases the track sample. In this case the alignment procedure seems then not to be able to recover from this loss.
5.3.3 Effect of low momentum tracks

As explained in Section 5.1, aligning with beam-gas data prevents from using the momentum information in the process. In this Section, we show one effect of this. A momentum estimate (and error estimate) is still not performed during the track fit, and so the multiple scattering is not taken into account. But we use the Monte Carlo truth information to reject low momentum tracks. A cut is set at 10 GeV/c. The results from Fig. 13 show that the convergence is faster when the low momentum tracks are removed. The IT layers alignment parameters converge in 1 to 2 iterations as opposed to 3 to 4 without the momentum cut. However, as one can see in the right-hand-side plot, the final position of the layers is slightly worse with the cut than without.

One more thing that can be noticed here is related to the issue of the underestimated errors. Compared to the other results shown in previous subsections, the errors here seem much better estimated with layers sitting only 3 to 5 standard deviations away from the true position. This is to be compared with the 10 standard deviations seen before. This error estimate problem can hence be related to the fact that, with no magnetic field, we have no
knowledge on the track momentum and cannot take the multiple scattering into account.

In this study, the momentum cut was performed by using the Monte Carlo truth information. When running the alignment algorithm on real data, this is obviously not possible. There are however other possibilities, which could be the aim of further studies. The tracks going through to IT and OT could be extrapolated to either the calorimeters or to the muon system in order to link the track to an information on its energy and hence be able to cut out the low-momentum tracks.

5.3.4 Reference alignment job (ideal geometry)

The goal of the alignment procedure is to get a result that is equal to the result starting from an ideal alignment (the default geometry with no misalignment added). In Fig. 14, we show the result of the same scenario as shown before, only starting from a null misalignment. These plots, and especially the plot on the right, could serve as a reference. What we see here is that the results are the same as those shown in Figs. 8 and 9, to 1 μm.
This fact shows that the alignment algorithm is very robust against misalignments, that it does not depend from the initial misalignment of our detectors. Said another way, it means that the results of our alignment procedure only depends on the track selection, use of drift time and other such parameters.

![Figure 14](image)

Figure 14: (a) Alignment parameters evolution and (b) (Input - output) Tx misalignment for each of the IT and OT layers and (beam gas scenario, ideal geometry).

### 5.3.5 Adding more degrees of freedom

In all the previous results, we were aligning the IT and OT layers with 4 IT layers in each stack and 4 OT layers frozen to the ideal position. This is a large number of constraints and it doesn’t allow us to get the IT position with respect to the OT, for example. When we will align our tracking stations with real data, we will be very interested in such information. But, of course, the more degrees of freedom we have, the more we will suffer from weak modes and the precision will worsen.

This is what has been tested below. In the plots of Figs. 15 and 16, we show the results of the alignment procedure with no constraint on the IT layers (all layers are free to move in the measurement direction and to rotate around the beam axis). In addition to that, in the latter plots, we show the effect of the cut on the momentum (taken from the Monte Carlo truth) at
10 GeV. We see that the convergence for the IT layers is 3 iterations slower in these cases. Also, the final position is now within 40 $\mu$m instead of 2 $\mu$m with the IT constraints. This is not a surprise but is a good indication of what result we can expect when we will align our detectors with real data. Also, it can be seen that the layers closer to the Outer Tracker are less well aligned than the others. The OT layers are also less well aligned than in the original scenario, but to a smaller extent. This is a combined effect of the layers being pulled away from the true position, while still being constrained by the fixed layers.

![Figure 15: (a) Alignment parameters evolution and (b) (Input - output) Tx misalignment for each of the IT and OT layers and (beam gas scenario, no IT layers constrained).](image)

5.3.6 Summary

There are a few conclusions we can draw from these studies. The first one is that, as expected, we gain a lot from the power of the drift times information. Running 4 iterations without drift times is enough to reach a minimum, which is then a good starting point for the use of the drift times information.

The second conclusion is that the best way to reject poor-quality tracks is to use the evolving $\chi^2$ cut described before. Not cutting, or cutting at a fixed value gives worse results in terms of final precision.
We’ve also seen that we suffer from the lack of knowledge of the track momentum. This can be seen in the bad estimation of the alignment parameters errors. Cutting on the true momentum (taken from the Monte Carlo truth) in order to reject low momentum tracks reduces the problem. When running the alignment algorithm on real magnet-off data, this momentum information could be taken from the calorimeters by linking the track to clusters in this sub-detector. The tracks could also be extrapolated to the muon system, which would allow for a selection of high momentum tracks (only these reach the muon system). The multiple scattering, leading to larger track parameter errors becomes less of a problem as soon as we have a momentum estimate. Indeed, the track fitting method we use accounts for the multiple scattering, which is a novelty in detector alignment.

A demonstration has been shown that the alignment procedure is robust against initial misalignment. Starting from a very misaligned detector (OT layers moved by up to 2.5 mm) or from the ideal geometry leads to the same final position.

Finally, we have seen that introducing more degrees of freedom (releasing the constraints on the IT layers) leads to a slower convergence as well as to
a worse alignment precision. Part of the loss in precision is due to correlated movements: this is also reflected in the larger estimated statistical errors.

5.4 Validation

In section 5.2 three methods of checking the alignment procedure convergence were discussed. However, this is not enough to ensure the results of the job are good enough for physics studies. Indeed, if the procedure converges but the detector elements still show some residual systematics, the physics performance of the detector could be degraded. Therefore, the quality of the alignment obtained has been studied using $K_S$ and $J/\psi$ decays. The results of these studies are described in the next two sections.

5.4.1 $J/\psi \rightarrow \mu^+\mu^-$ studies.

A sample of 65'500 inclusive $J/\psi$ events from DC06 DST production is re-fitted in DaVinci using a standard loose $J/\psi$ selection on three different geometry databases:

1. Unaligned (before alignment job).
2. Aligned (after alignment job).
3. Default (ideal geometry).

The database 2 is the output of the alignment job.

To check the validity of the alignment results, the distribution of the dimuon mass resolution $(M(\mu\mu)_{\text{rec}} - M(\mu\mu)_{\text{true}}))$ (see fig. 17) and of the track $\chi^2$/dof (fig. 19) are then compared for the three different cases. The mass bias with respect to the true dimuon mass and the mass resolution are plotted as a function of the particle’s momentum (see respectively fig. 18 (a) and (b)). If the alignment job has been done correctly, then the results for 2 and 3 should be very similar.

From fig. 17, we see that the shape of the distribution of the dimuon mass is fully recovered after alignment. This plot can be broken up into the dimuon mass bias (difference in the mean of the distribution) and dimuon mass resolution (width of the distribution) as a function of the particle momentum.
Figure 17: Dimuon mass resolution using the unaligned, aligned and default geometry (beam gas scenario).

Figure 18: (a) Dimuon mass bias and (b) dimuon mass resolution as a function of the J/ψ momentum (beam gas scenario).

These two variables are shown respectively in fig. 18 (a) and (b). At both low and high momenta, the difference in the bias between aligned and default geometries increases up to 0.5 MeV. However, the mean of the momentum distribution is at 80 GeV where the bias is equal for the aligned and default databases. Fig. 18 (a) also shows that most of the effect of the misalignment is removed after the alignment job, especially the high rise at large momenta.

This recovery is even more obvious in the mass resolution. At low momentum (between 10 and 60 GeV), the difference between the aligned and default database doesn’t exceed 1 %. This difference increases to 2 % for high mo-
mentum tracks.

Figure 19: Track $\chi^2$/dof for tracks in true $J/\psi$ candidates (beam gas scenario).

Finally, the distribution of the tracks $\chi^2$/dof also shows the same behaviour with the shape being recovered when aligning the T-stations (see fig. 19).

5.4.2 $K_S \rightarrow \pi^+\pi^-$ studies

The $K_S$ mass distribution has been studied with a sample of $\sim 100,000$ L0 selected minimum bias events from the DC 06 production [8]. Candidates were selected using the standard loose selection for $K_S$ decays that occur in the Velo. To simplify the analysis events with only one reconstructed primary vertex were considered. Fig. 20 (a) shows the mass distribution obtained. The S/B is around 0.8. This is increased to 5.9 for a 17 % loss in efficiency by making the additional requirements that the $\chi^2$ of the vertex be less than 20 and that the flight distance between the primary and the $K_S$ decay vertex be greater than 5 cm (Fig. 20 (b)).

Although the shape of the profiles of the mass bias and resolution are well recovered (with respect to the default geometry database) when aligning the T-layers, the difference between the default and the unaligned cases is very small. This is especially the case for the mass resolution at low momentum (10 % difference at the mean momentum -- $\sim 17$ GeV -- and decreasing for lower momentum values). If we compare these plots to the results shown in section 5.4.1, we see that the validation process with $K_S$ is less powerful than the $J/\psi$ case.
Figure 20: Mass distribution for candidates selected by (a) the loose K_S selection and (b) the K_S selection described in the text. The result of a fit to a Gaussian plus a flat background component is superimposed.

Figure 21: (a) Bias on the K_S mass and (b) K_S mass resolution versus p/GeV (beam gas scenario).

6 Studies with Magnet On

As soon as the LHC startup phase is finished, proton–proton collisions at $\sqrt{s} = 14$ TeV will be the main focus in LHCb. These high energy collisions will allow the detector to be used to its full extent. The magnet will be switched on at this point giving a momentum measurement. In addition, the Velo will be usable in its closed position, which may not be the case in the
first phase. This in turn will allow to use long tracks (ie tracks going through the Velo, the trigger tracker and the tracking stations) for the alignment. The Velo having a very precise standalone alignment (partly due to a high number of measurement planes), it can be used as a reference point for the alignment of the inner and outer tracker.

The sections below describe the scenario that has been studied using long tracks from minimum bias events produced and reconstructed with the magnet on.

6.1 Scenario

A sample of 20’000 minimum bias events reconstructed with the default geometry has been used for this analysis. The data was simulated with a centre-of-mass energy of 14 TeV and the magnetic field turned on to its nominal value. During the alignment procedure, the tracks are refitted using updated geometrical information from a misaligned database. In the scenario studied here, only translations along the local x axis (measurement direction) have been applied at the level of boxes, layers and ladders individually. The amplitude of these misalignments follows a flat distribution of 1 mm in width for the boxes, 100 µm for the layers and 50 µm for the ladders except for the innermost ones (the most illuminated during a run) which are misaligned with a value of up to 100 µm. The summary of these values can be found in table 7.

To solve for alignment, it is preferable to disentangle coherent movements
Table 7: Summary of the misalignments applied in the default (magnet on) scenario. (*) Except for the innermost ladder of each layer where this value can reach 100 µm.

<table>
<thead>
<tr>
<th>Detector</th>
<th>dof</th>
<th>Amplitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT</td>
<td>Tx [mm]</td>
<td>1.0</td>
</tr>
<tr>
<td>layers</td>
<td>Tx [mm]</td>
<td>0.1</td>
</tr>
<tr>
<td>ladders</td>
<td>Tx [mm]</td>
<td>0.05 (*)</td>
</tr>
<tr>
<td>OT layers</td>
<td>Tx [mm]</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Rz [mrad]</td>
<td>0.15</td>
</tr>
</tbody>
</table>

at the box and layer level before determining the positions of the individual smallest units (ladders). Therefore a multi-step approach was used for IT, aligning for detector elements deeper in the geometry tree at each step (see Fig. 23). At first, only the IT boxes are aligned, removing the constraint to the outer tracker by requiring tracks with no OT hits. At the end of this first job, the xml conditions are written and used in the second job where the IT layers are aligned, requiring again no OT hits on the tracks. In parallel to that, the OT layers are also aligned, with no constraint to the IT (requiring no IT hits on the tracks). Once the IT and OT layers have been aligned separately, they are again aligned together. This step is used to align the two detectors with respect to each other after their standalone internal alignment. Finally, the IT ladders and OT layers are again aligned together.

![Diagram of the multi-step alignment procedure used in the magnet-on scenario.](image)

During the whole chain, the detector elements are only aligned along the measurement direction (local Tx), while fixing Ty, Tz and Rx (the movements
to which the track $\chi^2$ is the least sensible), except during the first step where the boxes are still aligned along Ty. The movements that are neither aligned for, nor frozen are left free. As described in section 6, the Velo is taken as a reference point during the whole procedure. This exempt us from having to use other external constraints such as those described in section 3.

6.2 Results

In this section, we present only the results of the last step of the whole alignment procedure, namely the simultaneous alignment of the IT ladders and OT layers. As already explained in section 5.2, there are several ways of looking at the results of the alignment job. The flow of the 348 alignment parameters is not presented here because this plot would not be clear.

Fig. 24 shows the convergence of the normalised total sum of track $\chi^2$ (ie. the normalised track $\chi^2$ averaged over the track sample). During the iterative process, the number of tracks used for the alignment reduces, due to the cut on the track fit match $\chi^2$), from 293'000 down to 284'000 tracks. We see that in this last step of the alignment procedure, 3 iterations are needed to converge. This is the same number as we have already seen in section 5.2. However, the drift times are off in this present scenario.

Fig. 25 (a) shows the difference between the input (as set in the conditions) and the output (the result of the alignment job) misalignments for each of the IT ladders and OT layers. Each point on this plot corresponds to one detector element, starting from the first OT layer (station 1, layer X1), then moving on to the IT ladders (station 1, top box, layer X1, ladder 1) until the last ladder (station 3, C-side box, layer X2, ladder 7). We see on this plot that all the detector elements have converged within 100 $\mu m$ of the input misalignment. Moreover, we see in this distribution that only a few outliers are found outside 40 $\mu m$ on either side of 0.

It can also be seen in this plot that the alignment errors are correlated. Further displaced ladders have larger errors, which indicates that there are some correlated movements to which the algorithm is not sensitive (for example the X-scale).

In order to have a better feeling of the alignment precision, we can project all of these results on the y axis of fig. 25 (a), leading to the distribution in the right-hand-side plot of the same fig. The FWHM of the distribution is 30 $\mu m$, which corresponds to a Gaussian $\sigma$ of 13 $\mu m$, only 20 % of the IT.
Figure 24: Convergence of the normalised total sum of track $\chi^2$ (magnet-on scenario). The number of tracks is reduced during the iterative process (due to the cut on the $\chi^2$) from 293'000 to 284'000.

Figure 25: (a) (output - input) Tx misalignment for each of the IT ladders and OT layers and (b) alignment resolution in the Tx direction (magnet-on scenario).
resolution. The large number of tracks (about 290’000) used for this analysis induces that the statistical error on this result cannot be reduced more.

In addition to this narrow distribution, we can also see that the mean of the distribution is shifted by about 10 $\mu$m. This shift is actually a weak mode related to the fact that we don’t constraint the magnet bending. This means that the whole tracking stations (actually every sub-detectors situated after the magnet) can be moved sideways without the total sum of track $\chi^2$ being changed.

6.3 Validation

The same method and same data samples described in section 5.4 are used to validate the results of this magnet-on scenario.

6.3.1 $J/\psi \rightarrow \mu^+\mu^-$ studies.

The shape of the distribution of the dimuon mass resolution is again very well recovered after alignment (see Fig. 26). It can be seen that the mass resolution and bias as a function of the $J/\psi$ momentum (see Fig. 27), with the aligned database match very well the default results. The difference in resolution is less than 1%, even at high momentum, which is even better than in the first scenario studied. On the other hand, the results with the misaligned database is very different, which shows that the alignment procedure is correctly doing its job.

The same conclusion can be drawn for the distribution of the track $\chi^2$/dof (Fig. 28) which is fully recovered after the alignment, validating to a high confidence the procedure in use.

6.3.2 $K_S \rightarrow \pi^+\pi^-$ studies

As already seen in section 5.4.2, the results of the validation using the $K_S$ sample are much less convincing than those using a $J/\psi$ sample. The shape of the profiles and distributions seem to be recovered here as well. But as we can see in Fig. 29, the difference between the misaligned and aligned geometries is very small (near the mean of the momentum spectrum, around 17 GeV, this difference is only $\sim$2%). Therefore $K_S$ seems not to be a candidate as good as $J/\psi$ to validate the alignment results.
Figure 26: Dimuon mass resolution using the unaligned, aligned and default geometry (magnet-on scenario).

Figure 27: (a) Bias on the J/ψ mass and (b) J/ψ mass resolution versus p/GeV (magnet-on scenario).

7 Summary

In this note, we have presented a new method to align the LHCb detector using Kalman-fitted tracks coming from the standard track fit. The new feature in the procedure lies in the derivation of the global track covariance matrix after the Kalman-filter track-fit (see [4]).

We then presented two different realistic scenarios on which this alignment procedure has been tested. The first one used simulated beam-gas events with 450 GeV protons and no magnetic field. Starting from a realistic day-1
misalignment scenario for both IT and OT layers, we were able to align the detector to a good precision. Refitting tracks from J/ψ dimuon decays, we showed that the alignment procedure is able to fully recover the shape of the track $\chi^2$/dof distribution as well as resonance mass bias and resolution. The largest difference in mass resolution between the ideal case (using default geometry) and the geometry after alignment is of the order of 2-3 % for high momenta (above 80 GeV).

A second scenario has been presented where we jointly aligned OT layers and IT boxes, layers and ladders starting from a realistic day-1 misalignment. For
this second exercise, we used minimum bias events produced at a proton-proton centre-of-mass energy of 14 TeV, with the LHCb magnet on. The procedure has been adapted here to the high misalignment complexity and big number of degrees of freedom by aligning step by step, starting from a coarse granularity (IT boxes and OT layers separately) and moving to the finest granularity (IT ladders and OT layers together). We showed in particular that we are able to align the detector to a precision good enough to not affect the J/ψ mass resolution by more than 1-2 % (for momenta above 80 GeV).

These results are very promising a few months before the first beam-gas or collisions data will be acquired at LHCb. More work is needed in order to analyse the first TED and cosmics data and some new ways for rejecting weak modes will also be investigated.
References


A Cut on the (fit match) $\chi^2$ for tracks going through IT only

In Section 4.1, the $\chi^2$ and fit match $\chi^2$ of selected, not selected and ghost tracks were presented for tracks going through the tracking stations. In this appendix, we present the same distributions for tracks going through IT only (no OT hits on the tracks). We see (by comparing the plots in each of the figures 31 to 34 or looking at Tables 8 and 9) that the cut on the track quality (especially the $\chi^2_m$) is much more powerful with tracks going through IT only. This implies that when these cuts are applied, the sample of tracks selected to align the Inner Tracker will be much less polluted than the sample selected to align the Outer Tracker.

Table 8: Selection efficiencies and remaining ghost rate after cuts on the long track (going through IT only) $\chi^2$ and $\chi^2_m$ with misaligned geometry.

<table>
<thead>
<tr>
<th>Cut</th>
<th>Efficiency</th>
<th>Ghost rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Selected</td>
<td>Not selected</td>
</tr>
<tr>
<td>No cut</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td>$\chi^2$/dof $&lt; 100$</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td>$\chi^2_m &lt; 100$</td>
<td>95.28 %</td>
<td>74.58 %</td>
</tr>
</tbody>
</table>

Table 9: Selection efficiencies and remaining ghost rate after cuts on the long track (going through IT only) $\chi^2$ and $\chi^2_m$ with ideal geometry.

<table>
<thead>
<tr>
<th>Cut</th>
<th>Efficiency</th>
<th>Ghost rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Selected</td>
<td>Not selected</td>
</tr>
<tr>
<td>No cut</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td>$\chi^2$/dof $&lt; 10$</td>
<td>99.99 %</td>
<td>98.92 %</td>
</tr>
<tr>
<td>$\chi^2_m &lt; 30$</td>
<td>96.67 %</td>
<td>69.50 %</td>
</tr>
</tbody>
</table>
Figure 31: Distribution of track fit $\chi^2$/dof for (a) selected, (b) not-selected and (c) ghost tracks (going through IT only) with the ideal geometry.
Figure 32: Distribution of track fit $\chi^2$/dof for (a) selected, (b) not-selected and (c) ghost tracks (going through IT only) with the misalignment scenario presented in 6.1.
Figure 33: Distribution of track $\chi^2_m$ for (a) selected, (b) not-selected and (c) ghost tracks (going through IT only) with the ideal geometry.
Figure 34: Distribution of track $\chi^2_m$ for (a) selected, (b) not-selected and (c) ghost tracks (going through IT only) with the misalignment scenario presented in 6.1.