The Compact Muon Solenoid Experiment
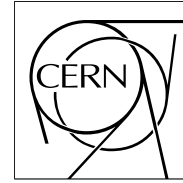
CMS Note

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

Track reconstruction in the CMS tracker

W. Adam,

HEPHY, Vienna

Th. Speer,

Zurich University

B. Mangano,

INFN, Pisa

T. Todorov,

CERN, Geneva, Switzerland and IReS, Strasbourg

Abstract

The track reconstruction based on the combinatorial Kalman filter method is described. The performance of the track reconstruction and the properties of the tracks are shown in detail.

1 Introduction

Track reconstruction in a dense environment needs an efficient search for hits during the pattern recognition stage and a fast propagation of trajectory candidates. In the CMS tracker the first task is simplified by the arrangement of sensitive modules in layers which are practically hermetic for a particle originating from the center of the detector.

The second task uses the fact that the magnetic field is almost constant in a large part of the tracker volume and that also most of the support structure is concentrated on the layers, close to the sensors. During reconstruction the typical step length for propagation of track parameters is in the order of the distance between two layers and a helical track model is adequate. For reconstruction purposes the detailed distribution of passive material as used in the simulation is replaced by an attribution of material to layers. This model simplifies the estimation of energy loss and multiple scattering which can be done at the position of the sensitive elements without requiring additional propagation steps.

Starting from the reconstructed hits, the track reconstruction is decomposed in four logical parts:

- Seed generation

- Pattern recognition, or trajectory building

- Ambiguity resolution

- Final track fit

All these stages are described in this note, with the exception of seed generation which is covered in a separate CMS note [1].

2 Pattern recognition

2.1 Description of the algorithm

The pattern recognition is based on a combinatorial Kalman filter method. The filter proceeds iteratively from the seed layer, starting from a coarse estimate of the track parameters provided by the seed, and including the information of the successive detection layers one by one. On each layer, i.e. with every new measurement, the track parameters are known with a better precision, up to the last point, where they include the full tracker information.

First, a dedicated *navigation* component determines which layers are compatible with the initial seed trajectory. The trajectory is then extrapolated to these layers according to the equations of motion of a charged particle in a constant magnetic field, accounting for multiple scattering and energy loss in the traversed material.

Since several hits on the new layer may be compatible with the predicted trajectory, several new trajectory candidates are created, one per hit. In addition, one further trajectory candidate is created, in which no measured hit is used, to account for the possibility that the track did not leave any hit on that particular layer. This fake hit is called an *invalid hit*.

Each trajectory is then "updated" with the corresponding hit according to the Kalman filter formalism. This update can be seen as a combination of the predicted trajectory state and the hit in a weighted mean, as the weights attributed to the measurement and to the predicted trajectory depend on their respective uncertainties.

All resulting trajectory candidates are then grown in turn to the next compatible layer(s), and the procedure is repeated until either the outermost layer of the tracker is reached or a "stopping condition" is satisfied. In order not to bias the result, all trajectory candidates are grown in parallel. To avoid an exponential increase of the number of trajectory candidates the total number of candidates is truncated at each layer.

2.2 Tunable parameters, regional and partial tracking

This algorithm is configurable through several parameters. Depending on their values, it can provide either a high track finding efficiency, as needed off-line, or very fast CPU performance suitable for its use in the HLT.

These main parameters are (default values in brackets):

- the maximum number of candidates that are propagated at each step (5)

- the inclusion of an invalid hit in the list of compatible hits, when the latter is not empty (always include invalid hit).

- the maximum χ^2 of the hits considered compatible with the predicted track state (30)

- the maximum number of invalid hits, i.e. crossings of sensitive detectors without a measurement (1)

- the maximum number of consecutive invalid hits (1)

- the minimum transverse momentum (0.9)

- the minimum number of hits per track (5)

In addition, it is possible to specify an arbitrary stopping condition, in which case the pattern recognition is interrupted before the end of the tracker is reached (*partial track reconstruction*). Such a condition is typically used in the HLT, since the required accuracy on track parameters is often reached after 5 or 6 hits, and the continuation of the pattern recognition to 12-13 hits would be a waste of CPU time.

If the track reconstruction is constrained to a region, the constraint affects mostly the seed generation phase. The only constraint that can be applied effectively at the pattern recognition stage is the transverse momentum cut.

2.3 Combinatorial behavior of the pattern recognition algorithm

When a trajectory is propagated to a given layer, the uncertainty of the predicted state has a direct effect on pattern recognition, since it determines the compatibility between the trajectory and nearby hits[1]. The number of compatible hits found on a layer determines the increase of the number of trajectories to be propagated, as each initial trajectory is multiplied by the number of hits found. As at most two hits (in the case of detector overlap) are correct[2], a high number of spurious hits[3] can potentially lead to a "combinatorial explosion" if no measure is taken to limit the number of trajectory candidates to be propagated. For each compatible layer, in addition to the measured hits, one invalid hit is added.

Several histograms will thus be shown to illustrate the behaviour of the pattern recognition algorithm. The sample used contains b jets with transverse momenta between 120 and 170 GeV/c, including low luminosity pileup. This sample features dense jets with a high number of tracks, and thus high hit densities. Only seeds of which the two hits are associated to the simulated tracks are followed and analyzed.

The following quantities illustrate the combinatorial complexity of the pattern recognition for a particular layer of the tracker:

- Uncertainty of the predicted state in the transverse ($r\varphi$) and longitudinal (rz) planes. The actual search window is several times bigger than the uncertainty shown in the figures, depending on the actual value of the χ^2 cut.

- Number of compatible hits found for a trajectory candidate, when leaving a particular layer.

- Number of trajectory candidate formed on a particular layer: These histograms show the number of trajectory candidate containing only valid hits which have been formed from the same seed on a particular layer and not been discarded. These are thus the trajectory candidates which will be propagated further. A measure of the contamination by spurious hits is given by the number per seed of trajectory candidates formed on a particular layer containing at least one spurious hit along the trajectory. In these histograms, the solid curves show the trajectory candidates containing only correct hits and the dashed curve shows the trajectory candidates containing at least one spurious hit along the trajectory.

---

[1] With a χ^2-compatibility criterion of 30, the window in which a hit would be accepted on that layer is approximately five times larger than the uncertainty of the predicted state.

[2] A reconstructed hit is deemed *correct* if it could successfully be associated to a simulated hit of that track on that layer.

[3] A reconstructed hit is deemed *spurious* if it could not be associated to a simulated hit of that track on that layer. It was nevertheless close enough to the track to be included during the reconstruction.

2.4 Pattern recognition in the Barrel

As a large fraction of the seeds are composed of hits in the first two pixel layers, the seed trajectories are first propagated from Barrel Pixel layer 2 to Barrel Pixel layer 3. With the limited information available at that stage, the parameters of the trajectory candidates are still poorly defined, and the uncertainties of the predicted states are quite large (Figure 1). Even though the uncertainty is large, due to the very fine granularity and the low occupancy of the Pixel detectors, only two hits are usually compatible (Figure 2). In most cases, these two hits are the correct hit and the invalid hit. As a consequence, the contamination of the trajectory candidates formed on Barrel Pixel layer 3 when the found hits are included is very low (Figure 3).
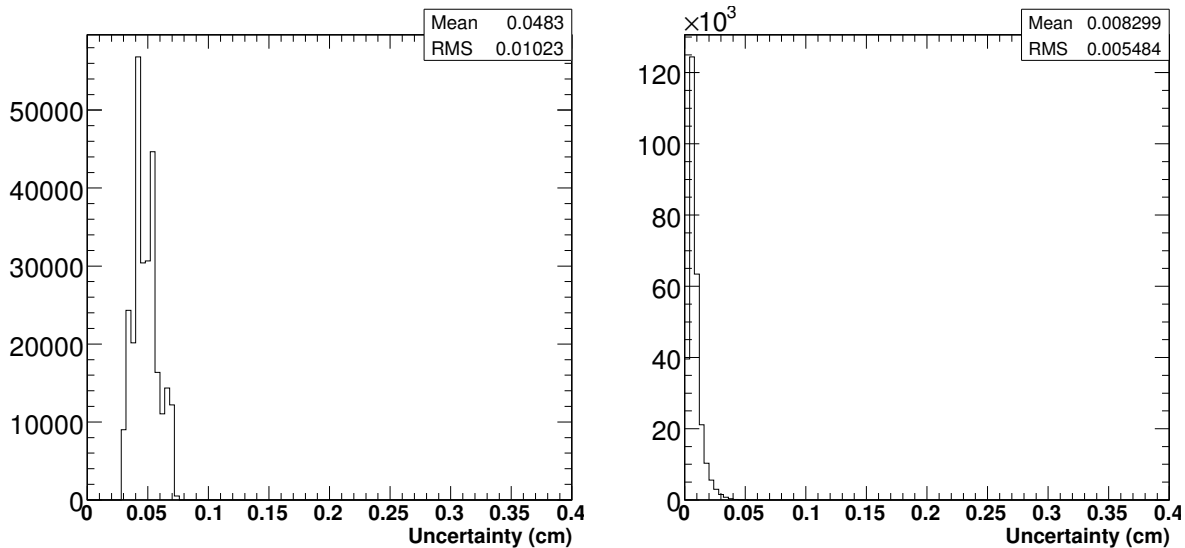
Figure 1: *Uncertainties of the predicted state on Barrel Pixel layer 3 for trajectory candidates leaving Barrel Pixel layer 2 in the $r\varphi$ (left) and rz planes (right).*

When these trajectories are propagated to the first TIB layer, the uncertainty on the predicted state increases and the distribution broader (Figure 4), due to the large gap between the two layers (approximately 13 cm), and the small lever arm of the initial trajectories (approximately 6 cm). As a strip detector with approximately 10 cm-long strips, and thus with a higher occupancy, is now reached, the number of compatible hits on that layer is larger (Figure 5), and the contamination markedly higher (Figure 6). Nevertheless, when the hit is included, the trajectories are now much better defined, with a larger lever arm. When these trajectories are propagated to the next layer, TIB layer 2, the uncertainties of the predicted states are reduced (Figure 7), which in turn reduces the number of spurious hits found on that layer (Figure 8) and the contamination decreases to a negligible level (Figure 9). Indeed, many of the trajectories containing spurious hits have now been discarded, either because they were not retained for a further propagation, or because no compatible hits were found when propagated. From there on, for the subsequent propagations through the rest of the Barrel, the trajectories are well defined and the contamination stays at the same low level.

The mean uncertainties on the predicted state in the different layers in the barrel are summarized in Figure 10 (left) and the average numbers of number of trajectory candidate per seed containing at least one spurious hit along the trajectory is summarized in Figure 10 (right).
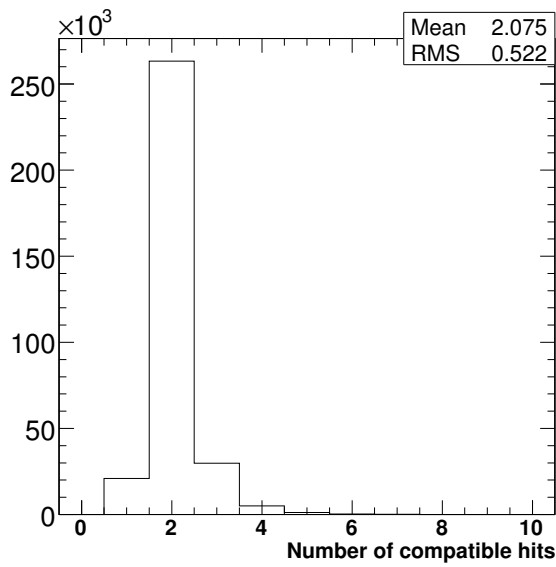
Figure 2: *Number of compatible hits (including the invalid hit(s)) found on Barrel Pixel layer 3 for each trajectory candidate when leaving Barrel Pixel layer 2.*
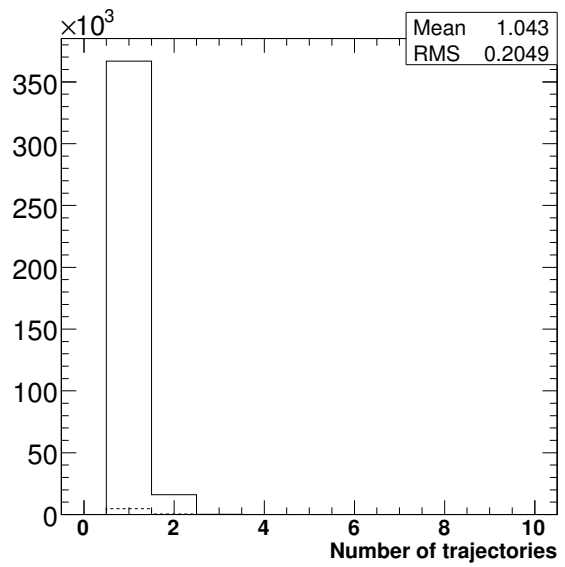

Figure 3: *Total number of trajectory candidates formed (containing only valid hits) for one seed on Barrel Pixel layer 3 during trajectory building. The solid curves show the trajectory candidates containing only correct hits and the dashed curve shows the trajectory candidates containing at least one spurious hit along the trajectory.*
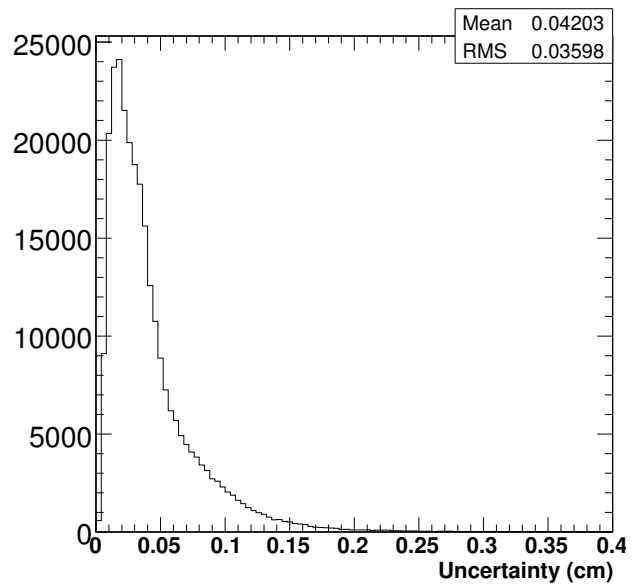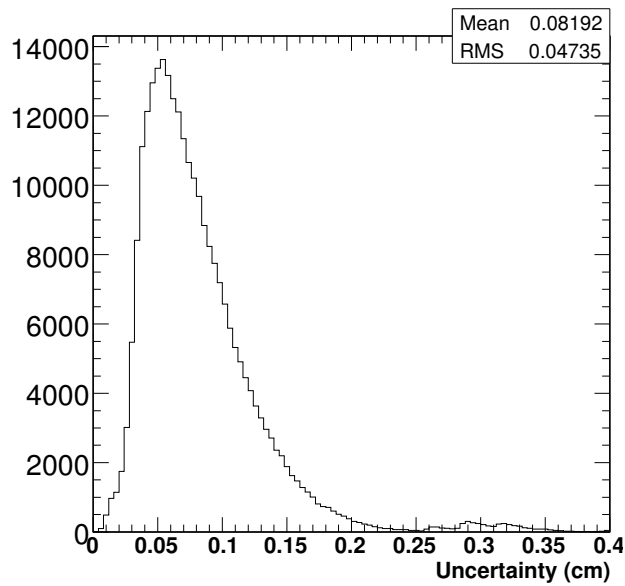

Figure 4: *Uncertainties of the predicted state on TIB layer 1 for all trajectory candidates in the $r\varphi$ (left) and rz planes (right).*
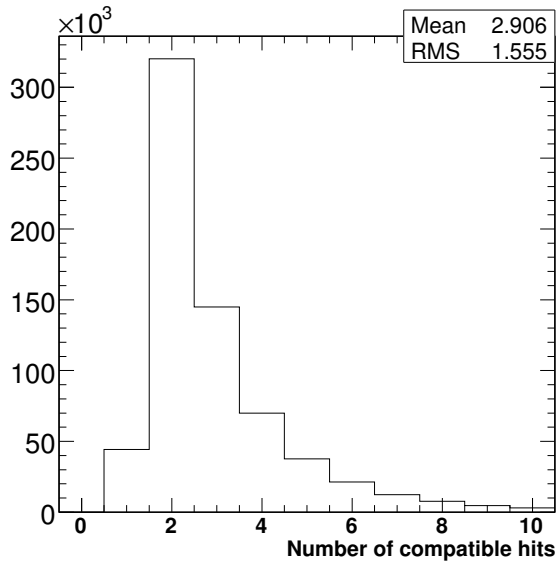
Figure 5: *Number of compatible hits found (including the invalid hit(s)) on TIB layer 1 for each trajectory candidate when leaving Barrel Pixel layer 3.*
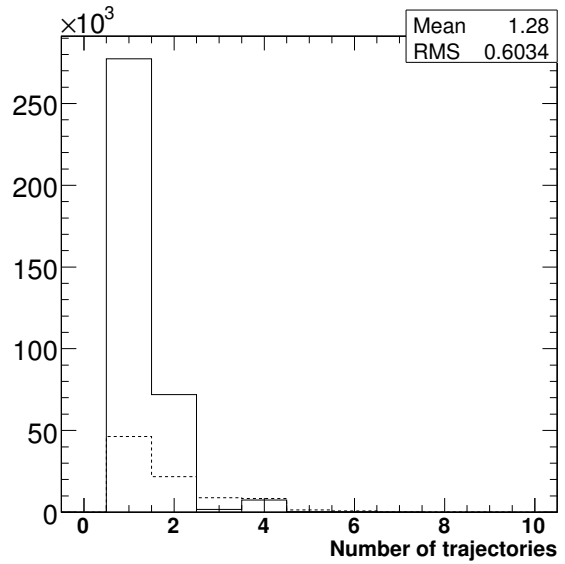

Figure 6: *Total number of trajectory candidates formed (containing only valid hits, refer to caption of Figure 3) for one seed on TIB layer 1 during trajectory building.*
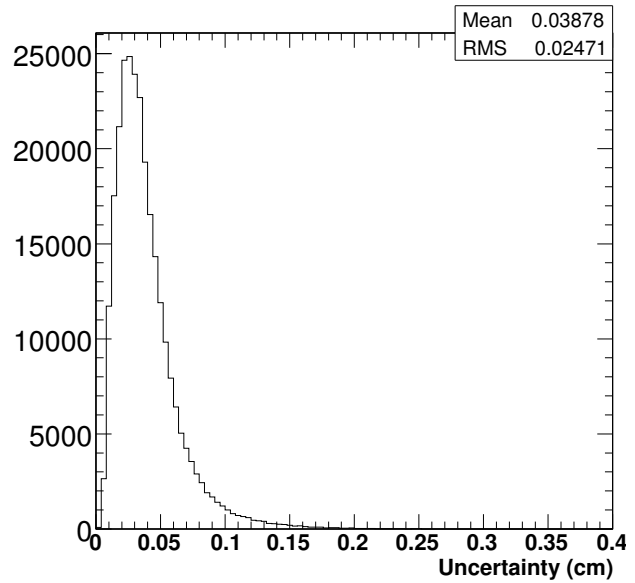

Figure 7: *Uncertainties of the predicted state on TIB layer 2 for trajectory candidates leaving TIB layer 1 in the $r\varphi$ (left) and rz planes (right).*

Figure 8: *Number of compatible hits found (including the invalid hit(s)) on TIB layer 2 for each trajectory candidate when leaving TIB layer 1.*

Figure 9: *Total number of trajectory candidates formed (containing only valid hits, refer to caption of Figure 3) for one seed on TIB layer 2 during trajectory building.*
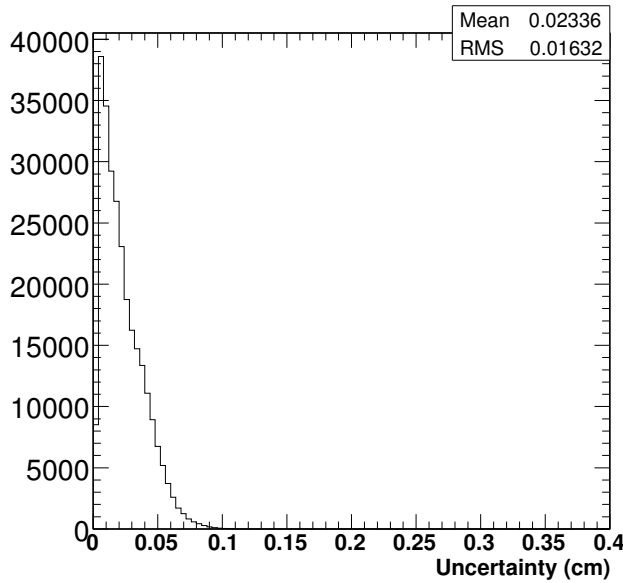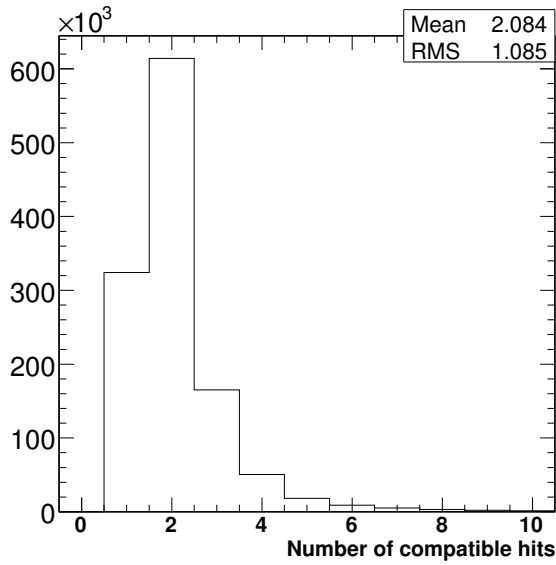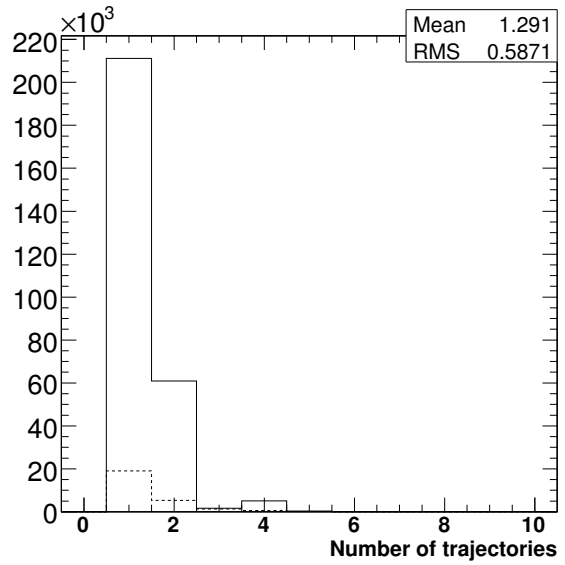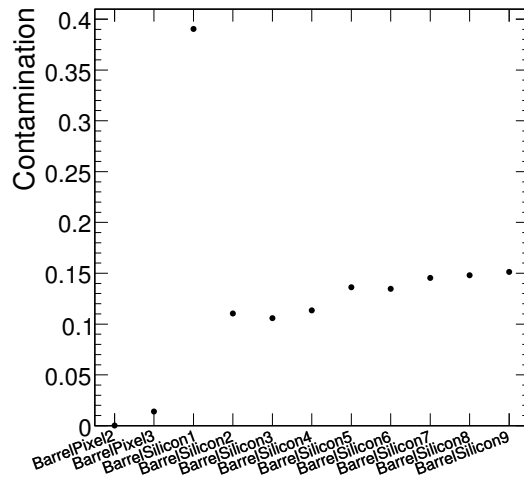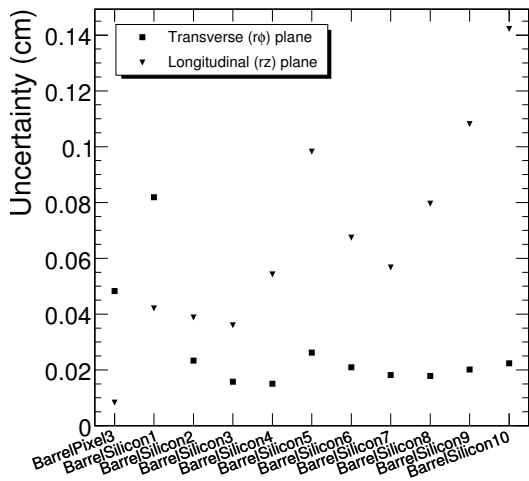

Figure 10: *Mean uncertainties of the predicted state in the Barrel in the transverse ($r\varphi$) and longitudinal (rz) planes (left) and average number of number of trajectory candidate per seed containing at least one spurious hit along the trajectory (right).*

7

2.5 Pattern Recognition in the Forward Tracker

While in the barrel the navigation is quite easy, it is more complex in the transition region. Due to the difficult geometry, the navigation can point to several layers. For example, for high-η tracks leaving Forward Pixel disk 2, all three Inner Disks and the first three Endcap disks could be compatible with the initial trajectory. When these layers are queried for compatible hits, each of these returns at least the invalid hit. As the propagation distance to some of these layers can be quite large, the uncertainties are comparatively large, and the probability of finding a spurious hit increases. There is as such a much higher number of compatible hits (Figure 11) and a higher contamination from spurious hits (Figure 12). Even the number of correct trajectory candidates is higher, since correct hits may be found in several of the compatible layers (as in the example of Forward Pixel disk 2), and each of the thus formed trajectory candidates, even though bypassing an intermediary layer, may continue to be propagated.

Once the trajectory is in the Endcap, the navigation is easy again, the competing trajectories are quickly discarded and the contamination reduced (Figure 13). The mean uncertainties on the predicted state in the different layers in the Forward Tracker are summarized in Figure 14 (left) and the average numbers of number of trajectory candidate per seed containing at least one spurious hit along the trajectory is summarized in Figure 14 (right).
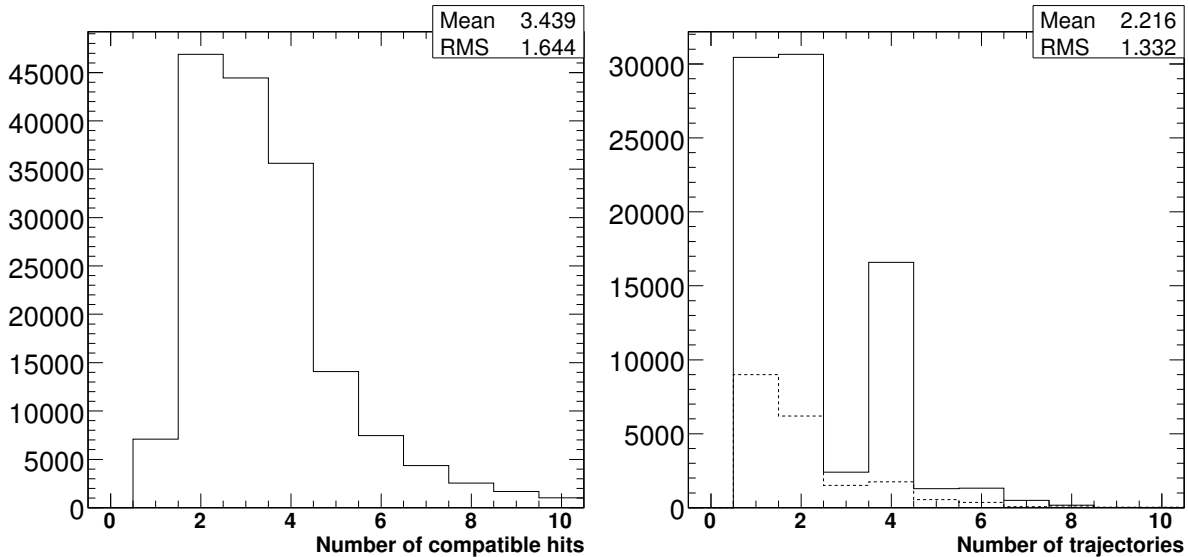

Figure 11: *Number of compatible hits (including the invalid hit(s)) for each trajectory candidate when leaving the Forward Pixel disk 2.*

Figure 12: *Total number of trajectory candidates formed (containing only valid hits, refer to caption of Figure 3) for one seed on TID disk 1.*

3 Ambiguity resolution

Ambiguities in track finding arise because a given track may be reconstructed starting from different seeds, or because a given seed may result in more than one trajectory candidate. These ambiguities, or mutually exclusive track candidates, must be resolved in order to avoid double counting of tracks.

The ambiguity resolution is based on the fraction of hits which are shared between two trajectories. For any pair of track candidates this fraction is defined in the following way:

$$f_{shared} = \frac{N_{shared}^{hits}}{min(N_1^{hits}, N_2^{hits})}$$

where N_1^{hits} (N_2^{hits}) is the number of hits in the first (second) track candidate. If this fraction exceeds a value of 0.5, the track with the least number of hits is discarded, or, if both tracks have the same number of hits, the track with the highest χ^2 value is discarded.
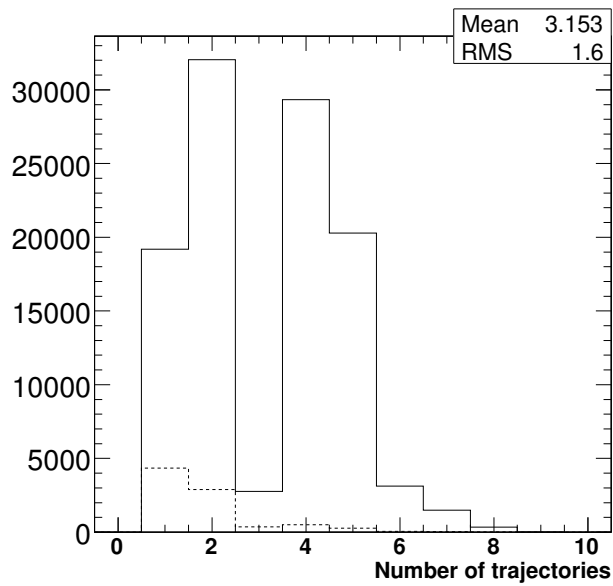
Figure 13: *Total number of trajectory candidates formed (containing only valid hits, refer to caption of Figure 3) for one seed on TEC disk 3 during trajectory building.*
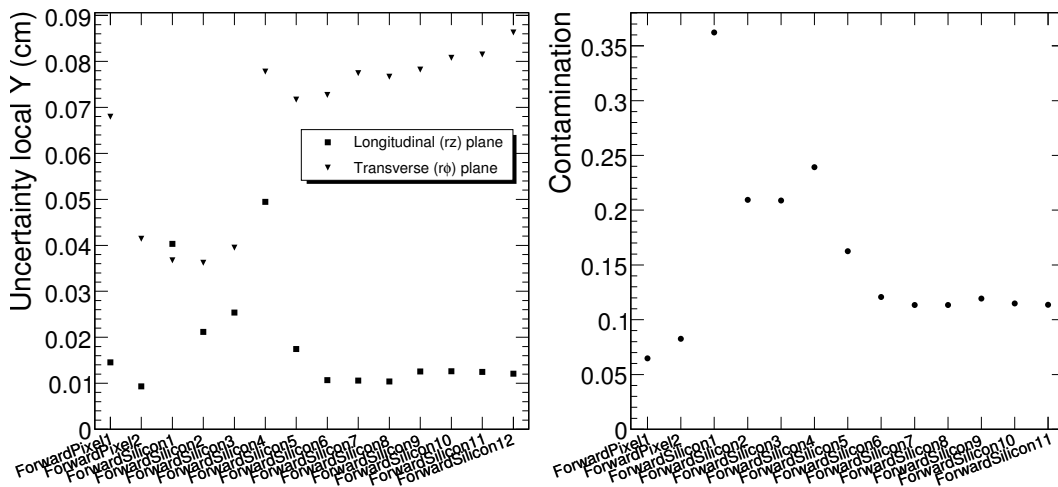

Figure 14: *Mean uncertainties of the predicted state in the Forward tracker in the transverse ($r\varphi$) and longitudinal (rz) planes (left) and average number of number of trajectory candidate per seed containing at least one spurious hit along the trajectory (left).*

The ambiguity resolution is applied twice: the first time on all track candidates resulting from a single seed, and the second time on the complete set of track candidates from all seeds.

4 Track fitting and smoothing

For each trajectory, the building stage results in a collection of hits and in an estimate of the track parameters. However, the full information is only available at the last hit of the trajectory and the estimate can be biased by constraints applied during the seeding stage. Therefore the trajectory is refitted using a least-squares approach, implemented as a combination of a standard Kalman filter and smoother.

The Kalman filter is initialised at the location of the innermost hit with an estimate obtained during seeding. The corresponding covariance matrix is scaled by a large factor in order to avoid any bias. The fit then proceeds in an iterative way through the list of hits. For each valid hit the position estimate is re-evaluated using the current values of the track parameters: information about the angle of incidence increases the precision of the
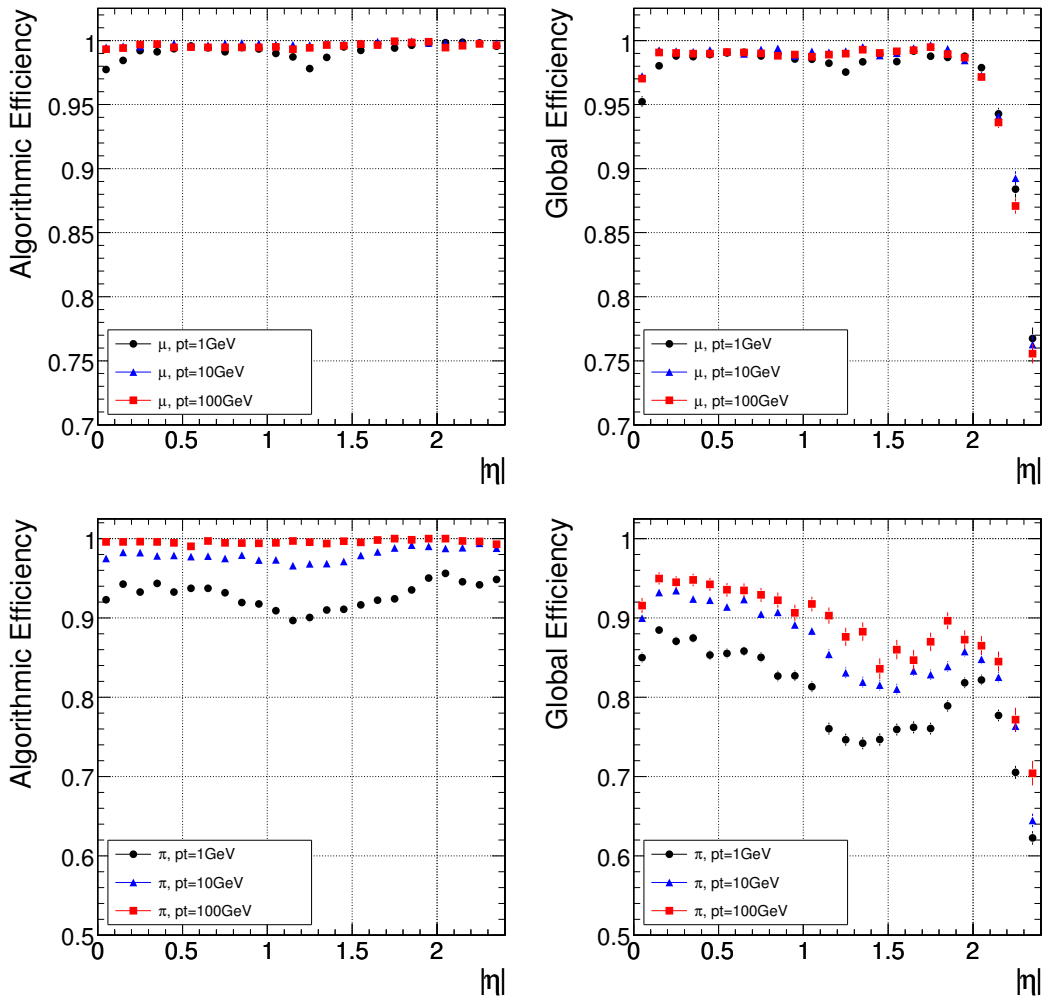
9

Figure 15: Algorithmic (left) and global track reconstruction efficiency (right) for muons (top) and pions (bottom) of transverse momenta of 1, 10 and 100 GeV/c.

measurement especially in the pixel modules. The track parameters and their covariance matrix are updated with the measurement and the trajectory is propagated to the surface associated with the next hit. The track parameters and their covariance matrix are modified according to the estimates for energy loss and multiple scattering at the target surface and the sequence is repeated until the last hit is included.

This first filter is complemented with a smoothing stage: a second filter is initialised with the result of the first one - except for the covariance matrix, which is scaled with a large factor - and run backwards towards the beam line. At each hit the "updated" parameters of this second filter, which contain all information from the outermost hit up to and including the current hit, are combined with the "predicted" parameters of the first filter, i.e. the information from the innermost hit outwards, but excluding the current hit.

This filtering and smoothing procedure yields optimal estimates of the parameters at the surface associated with each hit and, specifically, at the first and the last hit of the trajectory. Estimates on other surfaces, e.g. at the impact point, are then derived by extrapolation from the closest hit.

4.1 Track reconstruction performance

4.1.1 Efficiency and fake rate

The efficiency of reconstructing single tracks with the combinatorial Kalman filter has been estimated using samples of negatively charged muons and pions with transverse momenta of 1, 10 and 100 GeV/c. For this analysis tracks are reconstructed using seeds in the pixel detector and with default settings for the pattern recognition, ambiguity resolution and fitting stages. Reconstructed tracks are required to have a minimum of eight hits, with a hit missing in at most one layer, and $p_{\mathrm{T}} > 0.8$ GeV/c. A reconstructed track is associated to a simulated track if the

latter one contributed to at least 50% of its hits. A simulated track is deemed to be successfully reconstructed if a reconstructed track is associated to it and if the shared hits constitute the majority of all hits it shares with any reconstructed track.

To measure the track reconstruction efficiency, two definitions of efficiency are used, as in the case of the seed generator. The algorithmic efficiency is the efficiency of reconstructing correctly tracks with $p_T > 0.9$ GeV/c which have simulated hits in at least eight layers in the tracker, of which at least two are in the pixel detector, and which are originating in a region compatible with the seeding requirements. It measures directly the performance of the track reconstruction algorithm. It is essentially the efficiency of the pattern recognition stage, the other stages being fully efficient. The global efficiency is the reconstruction efficiency for *all* tracks with $p_T > 0.9$ GeV/c and with a production vertex inside the beam pipe. In addition to the efficiency of the algorithm, it includes the acceptance, hit efficiency and any other factor influencing reconstruction. It mainly differs from the algorithmic efficiency in the forward region, with the loss of coverage by the disks, especially in the pixel system. The fake rate is defined as the fraction of reconstructed tracks which could not be associated with a simulated track of $p_T > 0.7$ GeV/c. Details of these selections can be found in Table 1.

	reconstructed	simulated		
		global efficiency	algorithmic efficiency	fake rate
Nr. of hits	8	0	8	0
Nr. of pixel hits	-	0	2	0
Min. p_T (GeV/c)	0.8	0.9	0.9	0.7
η range	[-2.5,2.5]	[-2.4,2.4]	[-2.4,2.4]	[-2.6,2.6]
TIP / R_{vertex} (cm)	3.5	3.5	0.1	120
LIP / Z_{vertex} (cm)	30	30	15	250
Only signal event	-	yes	yes	no
Particle types	-	e,μ,π,K,p	e,μ,π,K,p	all

Table 1: Selection criteria for reconstructed and simulated tracks as used for the definition of efficiency and fake rate. The first (second) row specifys the minimum number of tracker (pixel) layers with at least one hit. The third and fourth rows list the cuts on transverse momentum and pseudo-rapidity. The following two rows specify the selection on the origin of the track. For reconstructed tracks the transverse (TIP) and longitudinal (LIP) impact parameters are used, for simulated tracks these variables are replaced by the radius and the longitudinal co-ordinate of the production point. The last two rows list further selections for simulated tracks: restriction to tracks from the signal event (rejection of pile up) and selection of stable particles.

The efficiency for single particles is shown in Figure 15. As the algorithmic efficiency indicates, the pattern recognition is fully efficient for pseudo-rapidities up to $|\eta| = 2.4$. For the global efficiency, the drop in the region $|\eta| < 0.1$ is due to the gaps between the sensors in the ladders of the pixel detector at $z = 0$. The alignment of the gaps in the three layers causes some tracks not to have the two required pixel hits. At high $|\eta|$, the drop of efficiency is mainly due to the lack of coverage by the two pairs of forward pixel disks. For hadrons, the global efficiency is lower than for muons because hadrons interact with the material present in the tracker. The lower efficiency of pions at 1 GeV/c is largely due to elastic interactions which are not taken into account in the definition of algorithmic efficiency.

The global and algorithmic efficiencies and the fake rate for tracks in b jets with transverse momenta in the range 120 - 170 GeV/c and including low luminosity pileup are shown in Figure 16. The track selection used for this analysis is a very loose one, and efficiency and fake rate can be tuned by applying additional quality criteria. As an example the changes of global efficiency and fake rate as a function of the cut on the normalised χ^2 or on the number of hits used are also shown in Figure 16.

4.1.2 Resolution

Five parameters are chosen to describe a track: d_0, z_0, ϕ, $\cot \vartheta$ and the transverse momentum p_T. The track parameters are defined at the point of closest approach of the track to the beam axis (this point is called the impact point); d_0 and z_0 hence measure the coordinate of the impact point in the transverse and longitudinal plane ($d_0 = y_0 * \cos \phi - x_0 * \sin \phi$, where x_0 and y_0 are the transverse coordinates of the impact point). ϕ is the azimuthal angle of the momentum vector of the track at the impact point and ϑ the polar angle. The track selection is the one used for the determination of the global efficiency. The resolutions are quantified by the standard deviation of a Gaussian fitted to the distribution of the residuals, i.e. the differences between reconstructed and simulated parameters.
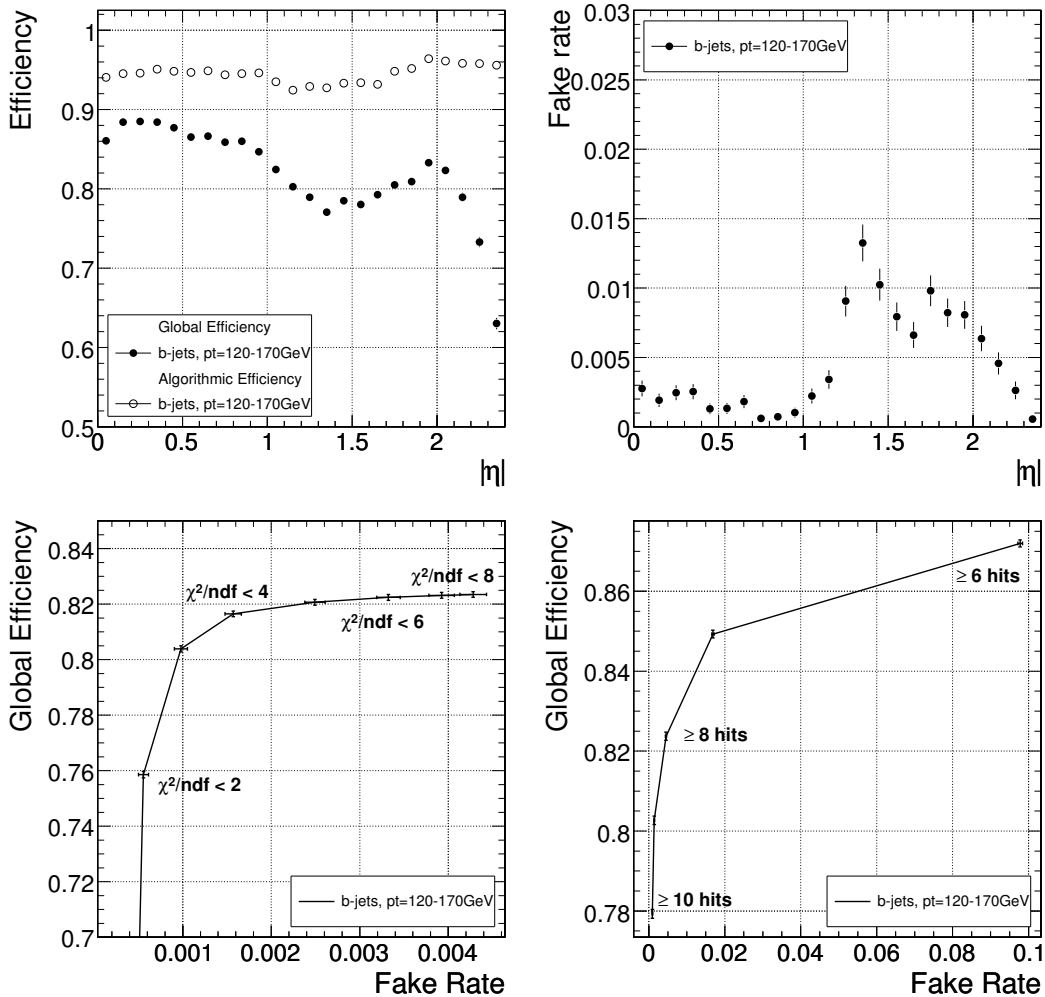
Figure 16: Global track reconstruction efficiency for b jets with transverse momenta between 120 and 170 GeV/c, including low luminosity pileup. Upper row: efficiency (left) and fake rate (right). Lower row: evolution of the average efficiency and fake rate for different cuts on the normalised χ^2 of the track (left) and for different number of hits used (right).

Figure 17 shows the normalised χ^2 and the number of hits used for reconstruction: two quantities which can be used as quality criteria. Figure 18 shows the resolution of the five track parameters for samples of single muons with p_T of 1, 10 and 100 GeV/c.

The resolutions of the transverse and longitudinal impact parameters d_0 and z_0 are shown in Figure 18 a) and b). At high momentum, the d_0 resolution is fairly constant and is dominated by the hit resolution of the first hit in the pixel detector. At lower momenta, the d_0 resolution is progressively degraded by multiple scattering, until the latter becomes dominant. The z_0 resolution of high momentum tracks is also dominated by the hit resolution of the first pixel hit, with multiple scattering dominating at low momentum. The improvement of the z_0 resolution up to a pseudo-rapidity of $|\eta| = 0.5$ can be attributed to the fact that in the barrel, as the angle with which the tracks cross the pixel layers increases the clusters become wider, improving the pixel-hit resolution.

The resolution of the transverse momentum is shown in Figure 18 e). At high momentum, the resolution is around 1-2% up to a pseudo-rapidity of $|\eta| = 1.6$, for higher values of $|\eta|$ the lever arm of the measurement is reduced. The degradation around $|\eta| = 1.0$ is due to the gap between the barrel and the end-cap disks and the degradation beyond $|\eta| = 1.2$ is due to the lower hit resolution of the last hits of the track measured in the TEC ring 7 with respect to the hit resolution in the TOB layers 5 and 6. At a transverse momentum of 100 GeV/c, the material in the tracker accounts for between 20 and 30% of the transverse momentum resolution; at lower momenta, the resolution is dominated by multiple scattering and its distribution reflects the amount of material traversed by the track. The mean value of the residuals for the transverse momentum is shown in Figure 18 f). The bias in the momentum reconstruction at large pseudo-rapidities is due to inhomogeneities of the magnetic field at large radius.

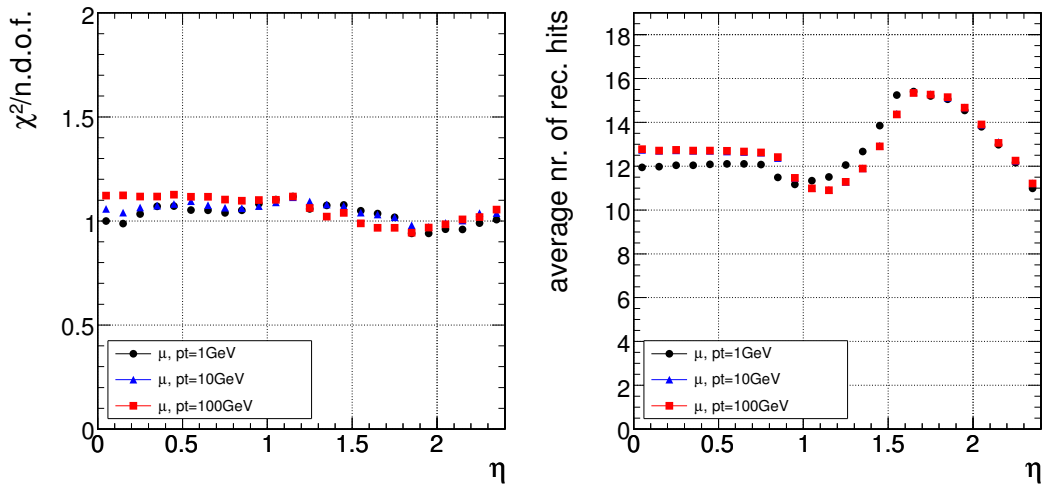This problem can be solved by using a Runge-Kutta propagator, which is in preparation.

Figure 17: Quality criteria for reconstructed single muon tracks with transverse momenta of 1, 10 and 100 GeV/c. Left: normalised χ^2. Right: average number of hits used for reconstruction.

Conclusions

The pattern recognition capabilities of the CMS tracker are analyzed, and the track reconstruction performance is determined in terms of efficiency, fake rate and resolution of the track parameters. The reconstruction task is decomposed into seeding, track finding and fitting stages such that individual components can be tuned to special needs. The system can provide both global and regional reconstruction.

Muons can be reconstructed with full efficiency within the acceptance of the seeding layers, while the efficiency for pions is limited to about 90%, dependent on pseudo-rapidity and on momentum, due to hadronic interactions with the tracker material. For muons with p_T of 100 GeV/c the resolution on the transverse impact parameter is typically 10 μm, almost independently of η, while the relative resolution on p_T is below 2 % for $|\eta| < 1.5$.

References

[1] **CMS NOTE-2006/026**, S. Cucciarelli, M. Konecki, D. Kotlinski, T. Todorov, *"Track reconstruction, primary vertex finding and seed generation with the Pixel Detector"*.
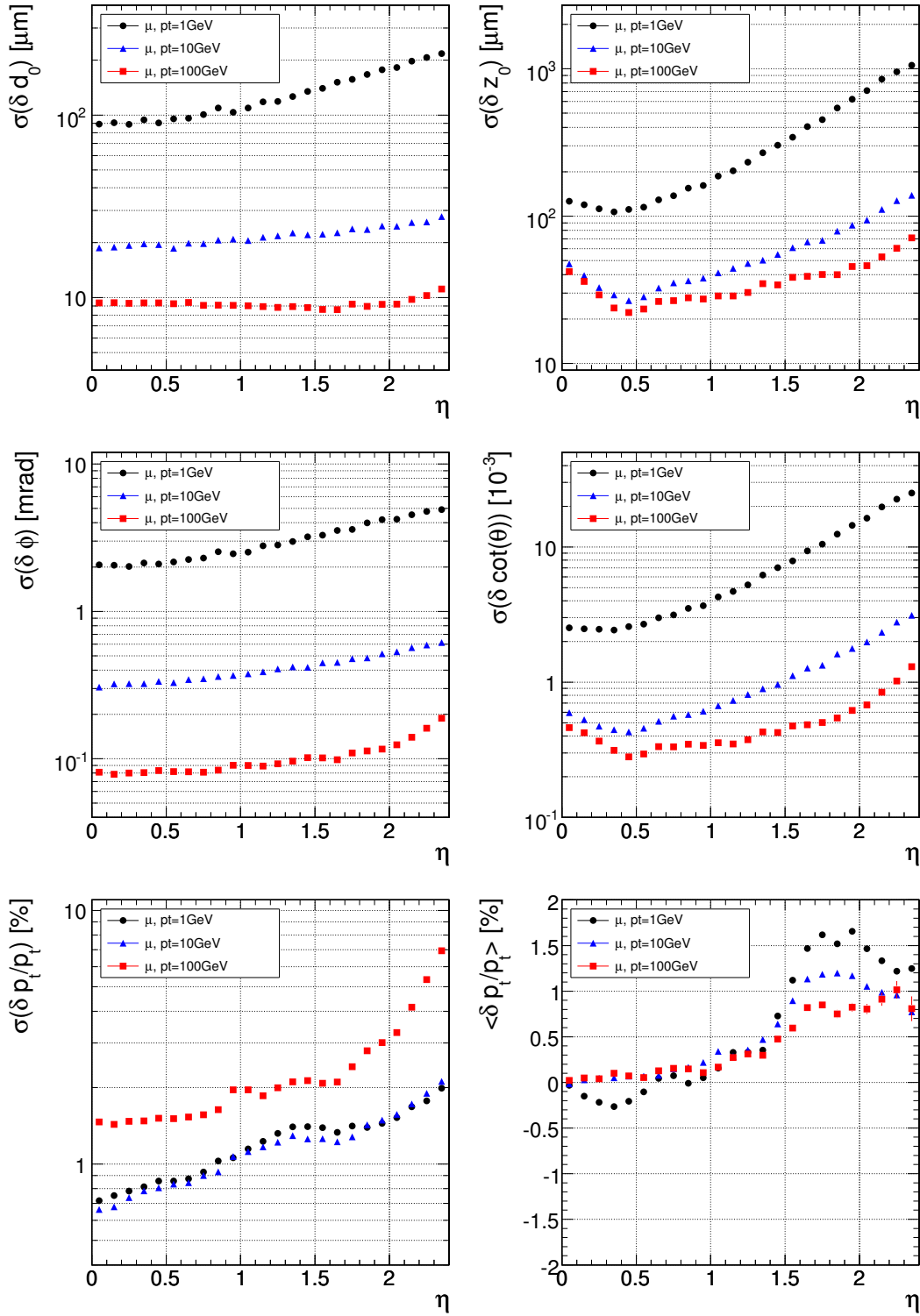
Figure 18: Resolution of the five track parameters for single muons with transverse momenta of 1, 10 and 100 GeV/c. From top to bottom and left to right: a) transverse, b) longitudinal impact parameter, c) ϕ, d) $\cot\vartheta$, and e) transverse momentum. Plot f) shows the mean residual for transverse momentum. The small bias for $|\eta| > 1.5$ is due to the inhomogeneity of the B-field at large radius.

14