DRAFT

CMS Physics Analysis Summary

The content of this note is intended for CMS internal use and distribution only

2010/03/21
Head Id: 351
Archive Id: 4127
Archive Date: 2009/12/13
Archive Tag: trunk

Track reconstruction in the CMS Tracker

The CMS Collaboration

Abstract

The main components of the software for the reconstruction of charged particle in the CMS tracker are described. The performance of the reconstruction algorithm are evaluated on simulated samples using the most recent version of the CMSSW software framework.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor: A. Author
PDFTitle: Track reconstruction in the CMS Tracker
PDFSubject: CMS
PDFKeywords: CMS, physics, software, computing, tracking

Please also verify that the abstract does not use any user defined symbols

Contents

1 The CMS Tracker

The CMS tracker [1] has a length of 5.8 m and a diameter of 2.5 m. It is immersed in a coaxial magnetic field of 4 T provided by the CMS solenoid. A schematic drawing of the CMS tracker is shown in Fig. 1. It comprises a silicon pixel detector with 3 barrel layers at radii between 4.4 cm and 10.2 cm and a silicon strip tracker with 10 barrel detection layers extending outwards to a radius of 1.1 m. Each system is completed by endcaps which consist of 2 disks in the pixel detector and 3 plus 9 disks in the strip tracker on each side of the barrel, extending the acceptance of the tracker up to a pseudorapidity of $|\eta| < 2.5$.

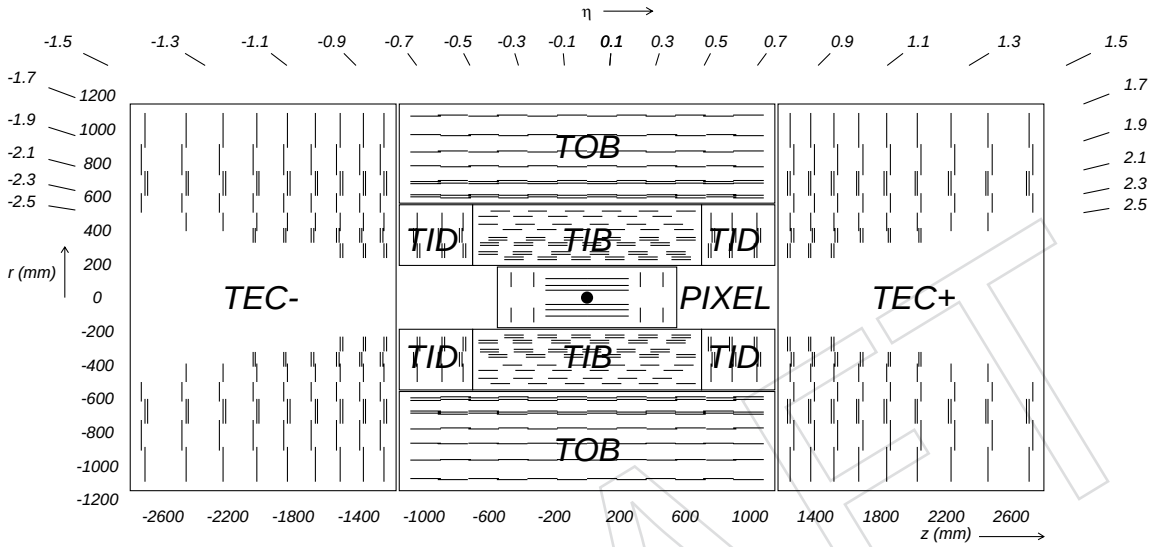

Figure 1: Schematic cross section through the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules which deliver stereo hits.

The pixel detector delivers high precision space-points with resolutions of 15-20 μm. In total it covers an area of about 1 m^2 and has 66 million pixels.

The strip tracker has a total of 9.3 million strips and 198 m^2 of active silicon area. It is composed of three different subsystems. The Inner Barrel and Disks (TIB/TID) cover $r < 55$ cm and $|z| < 118$ cm, and are composed of 4 barrel layers, supplemented by 3 disks at each end. TIB/TID provide r-ϕ measurements with a resolution of 23-35 μm. The Tracker Outer Barrel (TOB) covers $r < 55$ cm and $|z| < 118$ cm and consists of 6 barrel layers providing r-ϕ measurements with a resolution of 35-53 μm. The Tracker EndCaps (TEC) covers the region $124 < |z| < 282$ cm. Each TEC is composed of 9 disks, each containing up to 7 rings of silicon strip modules.

In addition, the modules in the first two layers of the TIB and of the TOB, as well as rings 1 and 2 of the TID and 1, 2 and 5 of the TEC, carry a second micro-strip detector module which is mounted back-to-back with a stereo angle of 100 mrad, in order to provide a measurement of the second co-ordinate (z in the barrel and r on the disks). The achieved single point resolution of this measurement is 230 μm and 530 μm in TIB and TOB, respectively, and varies with pitch in TID and TEC.

Figure 2 shows the material budget of the CMS tracker in units of radiation length.

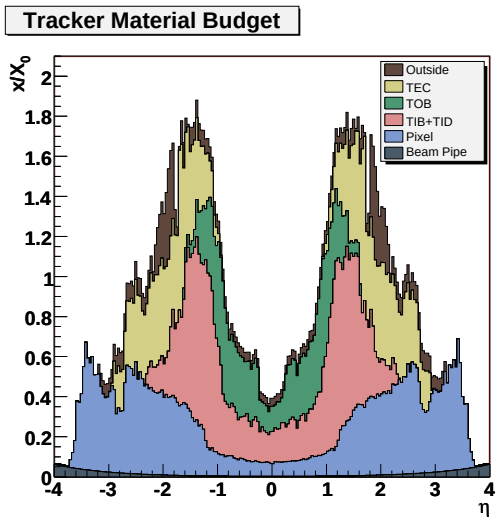

Figure 2: Material budget in units of radiation length as a function of pseudorapidity η for the different sub-detectors.

2 Local hit reconstruction

3 Track reconstruction software modules

The track reconstruction sequence is modularized into 5 logical parts:

- **Local reconstruction** consists in clustering into *hits* the strip and pixel signals produced by charged particles on the silicon detectors of the tracking system. The positions of the hits are estimated along with the corresponding uncertainties.
- **Seed generation** provides initial track-candidates for the full track reconstruction. A seed defines initial trajectory parameters and errors.
- **Pattern recognition** module is based on a global Kalman filter and is responsible of finding the track-candidates that correspond to charged particles of interests. Because trajectories are build in parallel and they are allowed to share position measurements, this modules is also responsible of cleaning the track-candidates collection removing duplicates.
- **Final Track fit** module estimates the parameters of trajectories with ultimate precision by means of Kalman filter and smoother.
- **Track selection** module rejects ghosts requiring that the final tracks match a minimum set of quality cuts.

3.1 Seed generation

The trajectory seeds define the starting trajectory parameters and errors that are used to initiate the iterative track-following procedure employed to identify the track-candidates. The seeds can be obtained externally to the Tracker, using inputs from other subsystems like the muon chambers or the calorimeters. However, the precision of initial trajectory parameters obtained in such a way is generally poor. Also, exclusively those particles that extend to the other subsystems (mostly muons, electrons and converted photons) can be seeded externally to the Tracker. Another strategy is to construct seeds internally. In this case each seed is composed from a small subset of the position measurements in the Tracker itself. Since five parameters (including the trajectory curvature) are needed to start trajectory building, at least 3 hits, or 2

hits and a beam constraint, are necessary to properly define a seed.

Because the seed defines only roughly the initial parameters of the trajectory, the early iterations of the track-finding module employ necessarily pretty large search windows to collect efficiently the rest of the position measurements that correspond to the a charged particle. Although constructing the trajectory seeds on the external layers of the tracking system, where the detector occupancy is lower, can facilitate the pattern recognition, there are three important reasons that encouraged to implement the opposite approach.

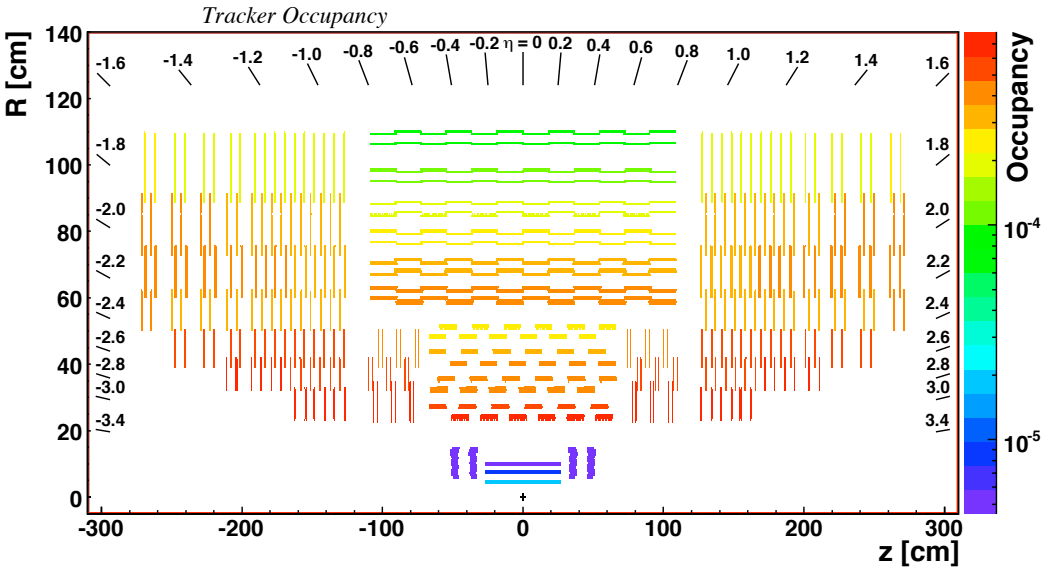

Figure 3: Occupancy of the Silicon Tracker detectors for Minimum Bias events simulated with superimposed pile-up collisions.

Firstly, even if the average occupancy of the *strip* layers of the Tracker system decreases faster than $1/r^2$ for detector placed at bigger and bigger radii, the occupancy of the inner *pixel* subsystem, due to its high density of read-out channels per surface unit, is about one order of magnitude lower than that of the most external strip detectors (Fig. 3). Secondly the pixel sensors provide truly 2-dimensional measurements that, given the same lever arm, allow the construction of trajectory seeds that have better defined parameters than equivalent seeds obtained with strip measurements. Finally, the most important reason to construct seeds in the innermost layers of the tracking system is that, due to the not negligible material budget of the Tracker, many particles produced in LHC collisions suffer destructive interactions before they reach the outermost layers of the tracking system: while most of the muons cross the whole tracker volume without being absorbed (Fig. 4), between 5% and 15% of the pions that are produced in collisions interact inelastically even before they reach the fifth layer of the Tracker (Fig. 5); similarly many electrons lose most of their energy due to bremsstrahlung radiation after they cross few layers. It is therefore apparent that, given the characteristics of CMS detector, a track finder algorithm can be highly efficient only starting the reconstruction of charged particle trajectories on the innermost layers and moving inside-out.

More than 90% of charged particles produced in LHC collisions, which are inside the geometrical acceptance of the Tracker, cross 3 pixel layers and so can be reconstructed starting from trajectory seeds that are obtained from triplets of pixel measurements. Nevertheless, because the barrel layers of the pixel detector are not fully hermetic (Sec. **??**) and the pixel readout is not

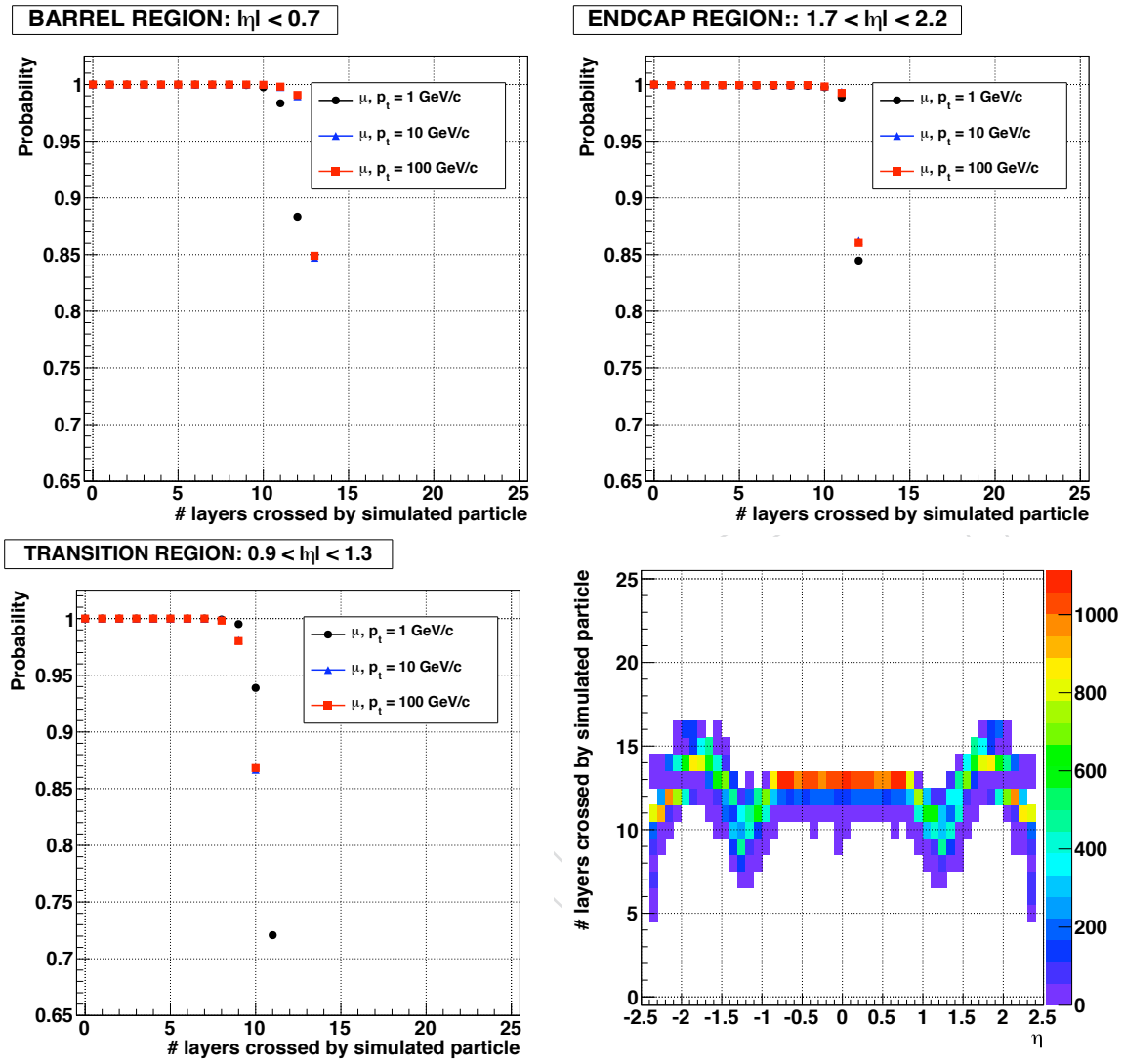

Figure 4: The first three plots show the probability that muons of different momenta are not absorbed by the detector material as a function of the number of Tracker layers that they cross moving outward from the collision point. Probabilities are evaluated for particles moving through the barrel layers (top-left), the endcaps ones (top-right) and the barrel-endcap transition region (bottom-left). The bottom-right plots show the number of crossed layers versus pseudorapidity for muons with transverse momentum equal to 10

fully efficient, seeds are also produced from pairs of pixel measurements plus a vertex or beam-spot constraint. Finally, additional seeds are constructed using a combination of pixel and strip measurements or uniquely strip measurements in order to extend the track reconstruction efficacy beyond the geometrical acceptance of the pixel system (Sec. **??**). In the implementation of the iterative CTF algorithm discussed in Sect. 3.6, four types of seeds are used:

- **Pixel Triplets:** seeds produced from triplets of pixel measurements allow the reconstruction of most of the primary charged particles produced in LHC collisions and facilitate a fast pattern recognition on the remaining layers of the silicon Tracker due to a pretty precise estimate of the trajectory seed parameters. A loose beam-spot constraint is employed to remove from the seeds collection those combination of pixel triplets that are not compatible with particles produced in the luminous region and that correspond presumably to ghosts. However, the beam-spot constraint is not used in the estimation of the trajectory seed parameters and errors.
- **Pixel and Strip pairs with vertex constraint:** using the previous collection of pixel triplets, the vertices corresponding to the primary collision and to the pile-up events are reconstructed. Pairs of tracker measurements are combined with the positions of these vertices to define new trajectory seeds and assure a $\sim$99% seeding efficiency also in the barrel region of the pixel tracker, despite its not complete hermeticity. While the vertex constraint is initially used to define the seed parameters, its position is not used during the final fit of the track in order to remove any bias. In addition to all the pixel measurements, the strip measurements that are used for the construct of this collection of seeds are hits on the two innermost rings of the three innermost TEC layers.
- **Pixel and Strip pairs with beam-spot constraint:** the set of measurements that are used to produce this collection of seeds is the same employed previously, but this time, instead of using a vertex constraint, only a looser beam-spot compatibility is involved in the definition of seed parameters. This seeds collection is designed mainly to recover trajectories corresponding to particles produced from the decay of long-living hadrons, like b-mesons: while these particles are produced within 1-2 mm from the beam line, they are not compatible with the primary vertex position. The bias of the beam-spot constrain is removed during the final fit of the tracks.
- **Strip-only pairs with beam-spot constraint:** these seeds are constructed using only 3-D measurements from the double strip tracker layers/rings. They are formed using the two innermost TIB layers and rings 1-2 of the TID/TEC. They can also be formed using the two innermost TOB layers and ring 5 of the TEC. The goal of this collection is twofold: assuring the reconstruction of trajectories that originate so far from the interaction point (e.g. K_s decay products) that they cross only one, or even zero, pixel layers; allowing the reconstruction of tracks at $\eta \sim 0$ where the pixel detector is intrinsically inefficient. The beam-spot constraint is very loose and its bias is completely removed during the final fit of the tracks.

3.2 Pattern recognition

The pattern recognition module of the CTF algorithm is based on the Kalman filter method. The filter proceeds iteratively from the seed layer, starting from a coarse estimate of the track parameters provided by the trajectory seed, and includes 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. Each iteration of the Kalman filter is implemented in separate steps that

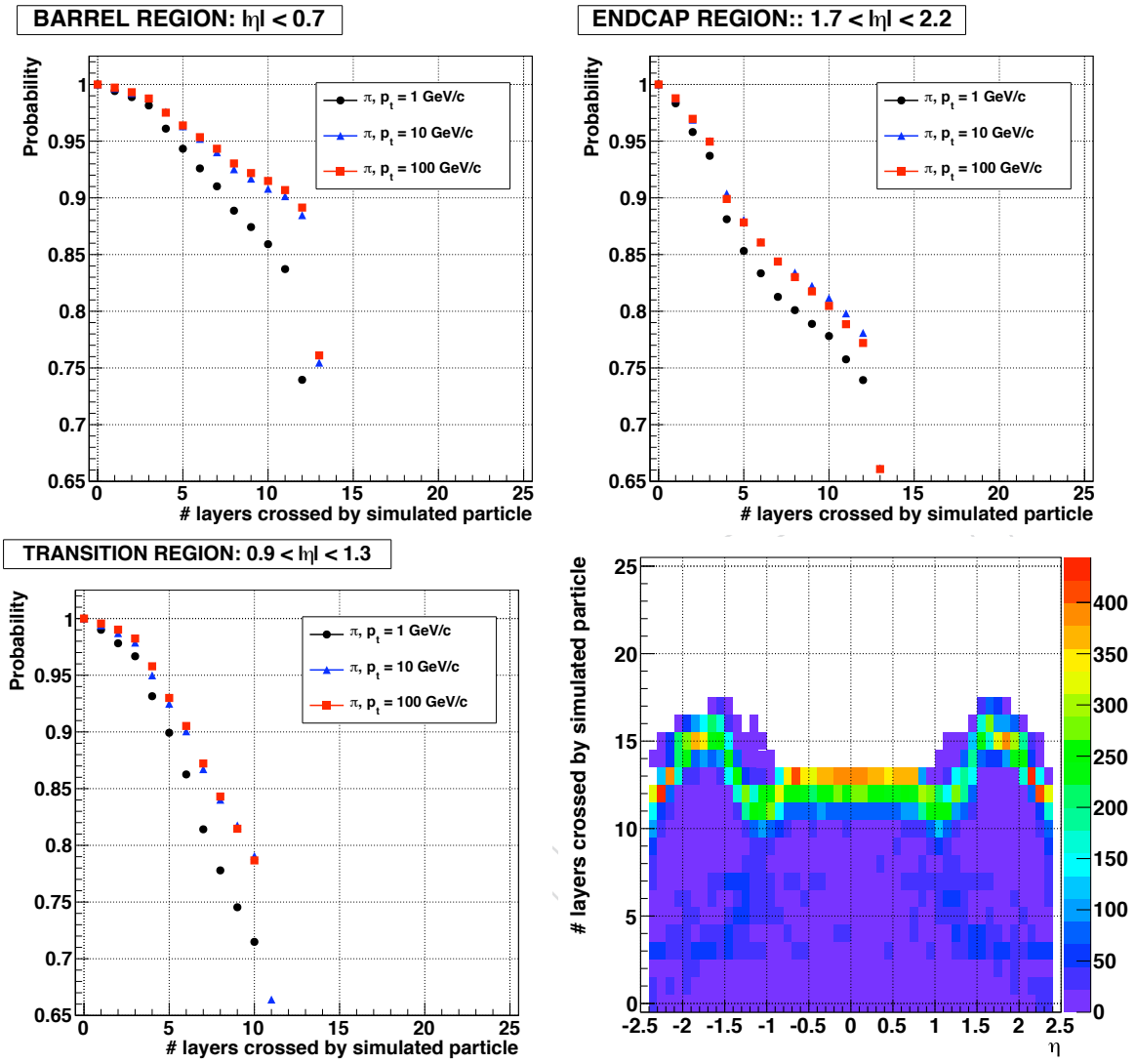

Figure 5: The first three plots show the probability that pions of different momenta are not absorbed by the detector material as a function of the number of Tracker layers that they cross moving outward from the collision point. Probabilities are evaluated for particles moving through the barrel layers (top-left), the endcaps ones (top-right) and the barrel-endcap transition region (bottom-left). The bottom-right plots show the number of crossed layers versus pseudorapidity for pions with transverse momentum equal to 10 GeV/c.

are executed in the following order:

1. **Navigation:** given the track-candidate's updated state on a specific layer, a dedicated navigation service determines which layers of the tracking system, among the adjacent ones, are expected to be intersected by the extrapolated trajectory. In determining the compatibility of a layer surface with the path of the particle, the uncertainty on the trajectory's parameters is considered. The navigation service is configurable and it can be used to build the trajectory both along its momentum and in the opposite direction. During the very first iteration of the trajectory building process, the input of the navigation service is the trajectory state defined by the seed.

2. **Search for compatible detectors:** after the track-candidate state is propagated to the surface of a compatible layer, the list of layer's detectors that are compatible with the trajectory is determined. A detector is considered incompatible with the trajectory when the position of the trajectory's intersection with the detector surface is n sigmas outside the detector's bounds. The default value of n is 3, but this parameter is configurable and it can be changed if track reconstruction efficiency is preferred to the algorithm's speed or vice versa. The propagation of the trajectory parameters (and of the corresponding errors) on the sensor surface involves mathematical operations and routines that are generally time-consuming [**?**]. Hence the code responsible for searching for the compatible detectors has been accurately optimized to limit as much as possible the number of layer's sensors that are queried and to assure, at the same time, an efficiency higher than 99% in finding the relevant sensors.

3. **Search for compatible measurements:** for each compatible detector, the direction of the trajectory on the sensor surface is used to calculate more accurately the drift of the ionization charge carriers inside the silicon bulk and to improve the estimation of the position of the pixel or strip clusters in the detector. A reconstructed clusters is considered compatible with the trajectory when the χ^2 of the predicted residual, as calculated in Eq. **??**, is smaller than a configurable cut (that by default is equal to 30). In order to reduce the processing time, the residuals and the χ^2 are estimated only for those strip clusters that have a chance to be compatible with the trajectory. At the moment, there is not a analogous optimization in the part of code responsible for finding compatible pixel clusters and therefore all the clusters in a pixel detector are analyzed.

4. **State update:** if a measurement is compatible with the trajectory, the track-candidate is extended by this additional point and the trajectory's parameters are updated to the filtered state, which is defined by Eq. **??**. When several measurements on the new layer are compatible with the predicted trajectory, several new track-candidates are created, one per each compatible hit. Finally, one additional track-candidate is created, in which no position measured is added, to account for the possibility that the charged particle did not leave any hit on that particular layer. This fake hit is called an "invalid hit".

All resulting track-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. To avoid an exponential increase of the number of candidates, only a limited number of these are retained at each step, based on their normalized χ^2 and the number of valid and invalid hits. Besides, the building of a trajectory can be stopped as soon as the uncertainty on its parameters is below a given threshold or a minimum number of measurements has been added: this kind of stopping conditions are typically used in the context

of HLT, where the required accuracy on track parameters is often reached after 5 or 6 position measurements are added to the track-candidate, and the continuation of the pattern recognition to 12-13 hits would be a waste of CPU time. 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. The number of compatible hits found on a layer determines the increase of the number of trajectories to be propagated, as the initial trajectories are multiplied by the number of hits found.

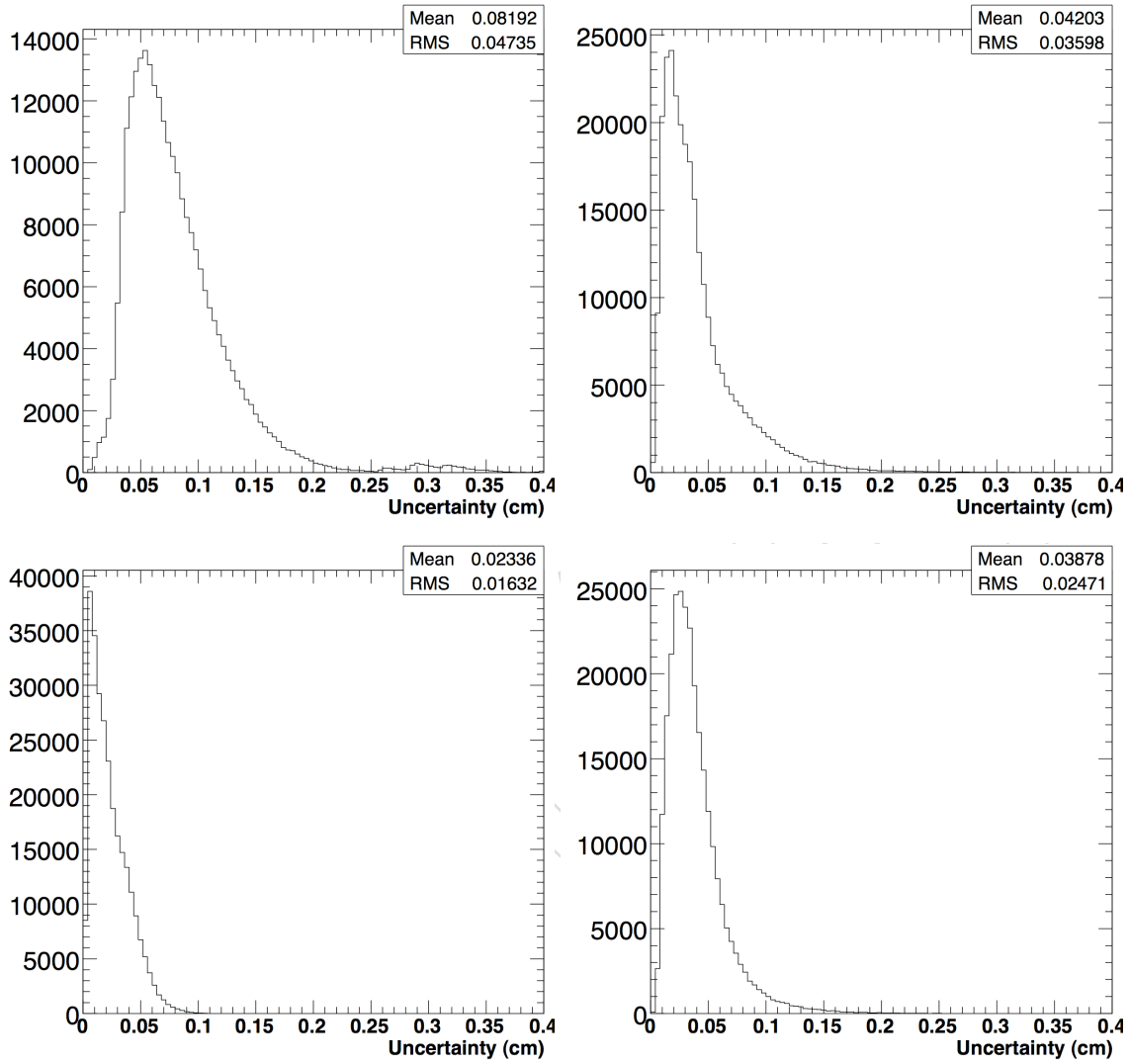

Figure 6: Distributions of the uncertainties on the X (left column) and Y (right column) position coordinates of the trajectory's predicted state during the pattern recognition step of track reconstruction. Coordinates refers the local reference frames of the detectors. Upper row shows histograms for states propagated from the third pixel barrel layer onto the first TIB layer; histograms for states propagated from TIB layer 1 onto TIB layer 2 are on the bottom row.

For particles moving through the barrel region of the Tracker, a large fraction of the seeds are composed of hits in the first three pixel layers (Sec. 3.2), and the parameters of these seeds are propagated from the third pixel layer to the first layer of TIB.

The uncertainty on the position of the predicted state is, at this stage, significant (Fig. 6), due to the large gap between the two layers (approximately 13 cm) and to the short lever arm (about 6 cm) that is used to determine the trajectory's curvature by means of only three pixel measurements. The average uncertainty on the transverse r-ϕ coordinate is about 800 μm, twice larger than that on the longitudinal z coordinate. This is a consequence of the still poorly defined trajectory's momentum and the difficulty in extrapolating the trajectory within the transverse plane.

When a measurement on the first TIB layer is included into the track-candidate, the uncertainty on the trajectory parameters becomes much smaller. In particular, the estimate of the curvature benefits significantly from the larger lever arm that now extends approximately over 20 cm. The position of the extrapolated trajectory on the second layer of TIB is estimated with an average precision of 230 μm in r-ϕ and 380 μm in z (Fig. 6, bottom).

Compared with the first propagation, which was influenced principally by the uncertainty on the trajectory curvature, starting from the second layer of TIB the accuracy of the propagated state is dominated by the precision of the last position measurement added to the track-candidate. For double-sided strip detectors, the z coordinate of a cluster is determined with a precision that is about 10 time worse compared to the r-ϕ coordinate. Besides, on single-sided detectors, the z coordinate is not measured at all and it is constrained only by the length of the sensor's strips (approximately 10 or 20 cm). For this reason, excluded the first extrapolation, the uncertainty on the z coordinate of the predicted state is larger than that on r-ϕ for all the layers of the Strip Tracker (Fig. 7).

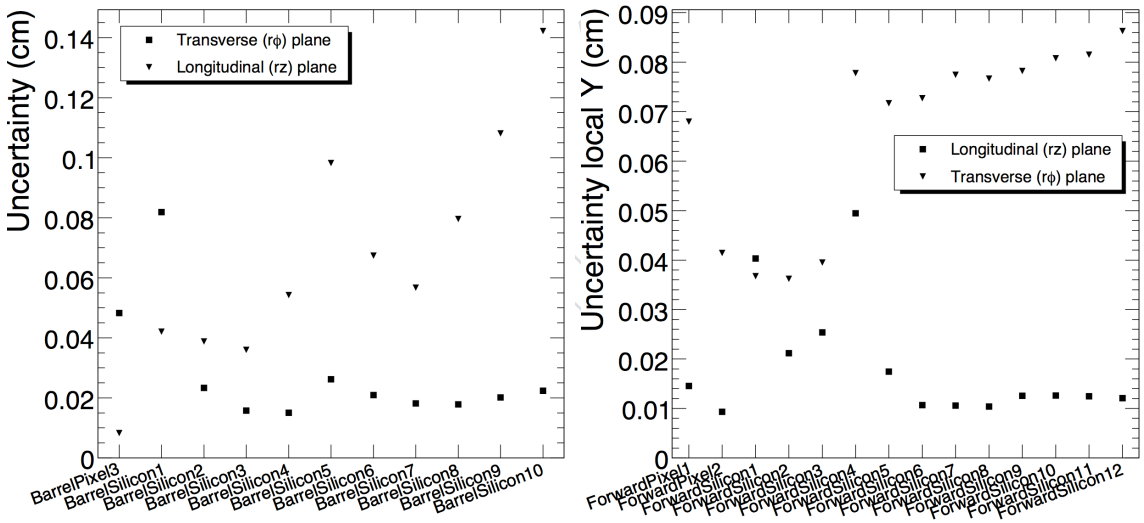

Figure 7: Mean uncertainties of the predicted state in the barrel (left) and forward tracker (right) in the transverse (r-ϕ) and longitudinal (r-z) planes.

The correlation between the uncertainty on the extrapolated trajectory state and the number of compatible measurements affects the distributions shown in Fig. 8: while more than half of the trajectories propagated from the third layer of the pixel barrel prove to be compatible with a spurious hit, this fraction is reduced approximately to one third for trajectories propagated from TIB layer 1 to TIB layer 2.

The specific geometry of the CMS Tracker, which consists of cylindrical barrel layers completed by disks in the two endcaps, facilitates the navigation between barrel layers during the pattern

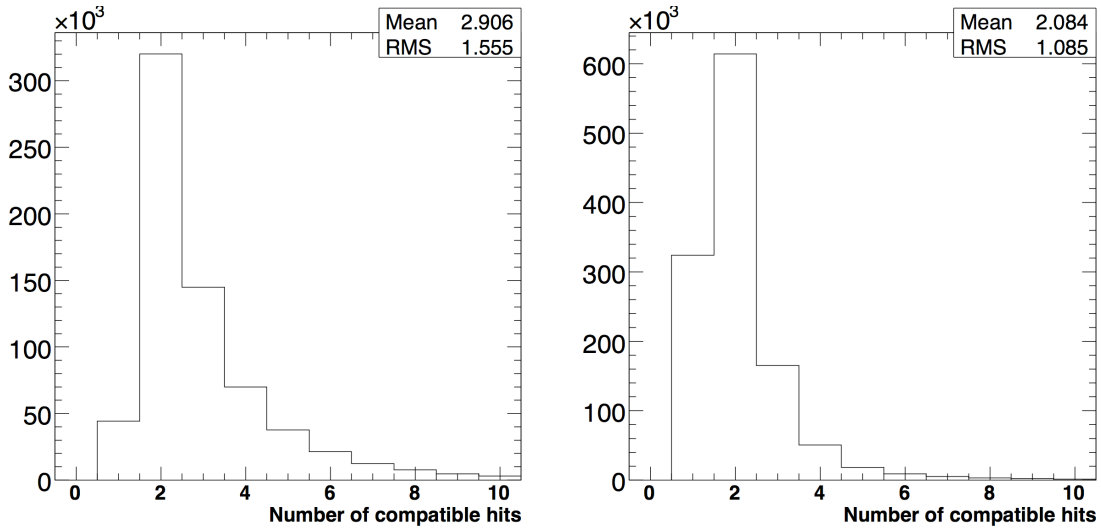

Figure 8: Number of compatible hits (including invalid hit(s)) found per trajectory, when leaving the pixel barrel layer 3 (left) and TIB layer 1 (right). Simulated b-jets events with transverse momenta between 120 and 170 GeV/c are used, with low luminosity pile-up included

recognition stage of track reconstruction. However, the navigation is more complex in the barrel-endcap transition region, where a single layer can be connected to several others. For example, for a trajectory state that is defined on the Forward Pixel disk 2 and is pointing at high η, all three TID disks and the first three TEC ones could be compatible with the 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 also spurious hit increases. Once the trajectory building reaches the endcap, the navigation is easy again and many of the trajectories containing spurious hits are quickly discarded, usually because they fail the condition on the minimum trajectory's χ^2 or that on the number of invalid hits.

3.2.1 Ambiguity resolution

Ambiguities in track-finding arise when the same track is reconstructed twice starting from different seeds or when a given seed develops into more than one final track-candidate. The ambiguity among *mutually exclusive* track-candidates must be resolved in order to avoid that the same charged particle is counted twice, or even multiple times.

The pattern recognition module of the CTF algorithms aims to identify track duplicates through an ancillary service, called trajectory cleaner, that determines the fraction of *shared* hits between two track-candidates:

$$f_{shared} = \frac{N^{hits}_{shared}}{min(N^{hits}_1, N^{hits}_2)} \quad (1)$$

where N^{hits}_1 (N^{hits}_2) is the number of hits in the first (second) track-candidate. If this fraction exceeds the cut value c = 50%, the trajectory cleaner removes the track with the least number of hits from the candidates collection; if both tracks have the same number of hits, that with the highest χ^2 value is discarded. The procedure is repeated iteratively until all the pairs of track-candidates in the collection shares less than half of their hits.

The probability that a track-candidate is a ghost produced by accidentally aligned hits decreases very quickly with the number of measurements on the trajectory and therefore the ambiguity resolution method is biased toward long trajectories. However, such a simple selection criteria, which is successful in most of the cases, fails in some special conditions, like during the reconstruction of high energy hadrons and electrons. As it is shown in section 4.1, sometimes the hits produced by secondary particles (e.g. electrons from conversions or the products of inelastic nuclear interactions) are incorrectly attached by the trajectory builder to the track-candidate corresponding to the primary particle. In this case, it is the shorter track-candidate, containing only the measurements produced by the primary particle, that should be preferred to the longer candidates that include also spurious measurements. According to a more conservative approach, neither of the candidates should be discarded and the ambiguity resolution should be post-poned to a later stage of the reconstruction sequence where additional information becomes available. It has been recently suggested to use a different configuration of the trajectory cleaner for the tracking of electrons and now the cut value c isn't fixed to 50% anymore, but it is a configurable parameter of the tracking algorithm.

3.3 Final Track Fit

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 the standard Kalman filter and smoother approach described in section **??**.

The Kalman filter is initialized 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 full list of hits. For each valid hit the position estimate is re-evaluated again using the current values of the track parameters. This first filter is complemented with the smoothing stage: a second filter is initialized with the result of the first one (except for the covariance matrix, which is scaled with a large factor) and is run backward toward the beam line.

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.

3.4 Rejection of spurious measurements

The Final Track Fit module is also in charge of removing spurious measurements, or *outliers*, that could be included into the track-candidates during pattern recognition. These measurements can be correlated with the otherwise well-defined track, e.g. clusters from δ-rays, or uncorrelated, like hits due to nearby tracks or to electronic noise. Spurious measurements have generally larger residuals than the genuine signal hits produced by the passage of the charged particle through the detector. During the pattern recognition, it is appropriate to discard spurious hits determing the closeness of a measurement to the trajectory by means of the χ^2 of the predicted residuals (eq. **??**). However, the χ^2 of the *smoothed* residual (eq. **??**) is a more powerful decision criterion once the full information from the tracker is available to estimate the trajectory parameters on each sensor surface [**?**].

Once the Kalman smoother is executed, the Final Track Fit module identifies the track's measurement with the largest smoothed residual and removes the hit from the trajectory if its resid-

ual is above a configurable threshold C. The remaining measurements are re-fit with the usual mechanism and the measurement with the next higher residual is found. The procedure is repeated iteratively until all the trajectory measurements have their residuals below the bound C. Every time a measurement is removed from the trajectory, it is replaced by an invalid hit: in this context, the invalid measurement is a sort of placeholder informing the Kalman smoother that, despite the lack of a genuine measurement on that specific sensor, the effects of material (like energy loss and multiple scattering) have still to be accounted during the fit.

The default value (20) for the parameter C has been chosen as the best tradeoff between rejecting the maximum fraction of outliers and removing the minimum number of genuine measurements. About 20% of outliers are pulled out from tracks reconstructed in high-density di-jets events, compared to less than 0.2% of good measurements that are removed [**?**].

While the described method aims to identify spurious measurements using exclusively geometrical information, it is currently under development a more sophisticated approach that exploits also additional attributes of the reconstructed hits, like its cluster charge and cluster width.

3.5 Track Selection

Within an average LHC event with jets the combinatorial track finder described above yields a significant fraction of fake tracks (*ghost* tracks). The fake rate can be effectively reduced by applying quality cuts that remove tracks with a bad fit-χ^2 and a bad compatibility with the event interaction vertices. An optimal way to bring down the overall fake rate retaining the best possible efficiency for true tracks, is to adapt the quality cuts to the track η, p_T and, most of all, to the track number of crossed layers with measurements [2].

In simulated events ghost tracks can be identified by applying an association procedure between reconstructed tracks tracks and Monte Carlo truth particles defined in Section 4. Tracks that are not associated to a simulated true track are called *ghost* tracks.

The relative distribution of matched and ghost tracks decreases exponentially as a function of the track number of crossed layers $dN_{ghost}/dn_{\text{layers}} \sim \exp(-\beta n_{\text{layers}})$ with $\beta = 0.9 - 1.3$ according to the p_T range of the tracks.

Quality cuts to reject fake tracks, have been initially optimized for a single step CTF track reconstruction, adapting the cuts to the track number of layers, p_T and η, and applying cuts on

1. the track χ^2/ν (chi-square per degree of freedom),
2. the track d_0 to the beam spot,
3. the track Δz to the position closest HLT primary vertex,
4. the $d_0/\delta d_0$ transverse compatibility with the beam spot, and
5. the $\Delta z/\delta z_0$ longitudinal compatibility with the closest HLT vertex.

The HLT vertices are taken from the *pixelVertices* collection requiring at least three pixel tracks in the vertex and a vertex Prob$(\chi^2) > 1\%$, if present.

The quality cuts listed above have been optimized as a function of the track p_T and n_{layers}, to reach a maximum quality $Q(\alpha) = \frac{s}{\sqrt{s+\alpha b}}$, where s is the number of selected matched tracks, b is the number of selected fake tracks, and $\alpha \simeq 10$ inflates the ghost track counting to achieve low fake rates (below 1% for QCD events with $\hat{p}_T = 170 - 230$ GeV/c).

In order to adapt the track quality cuts to the expected track resolution for the vertex association, the resolutions on the track measured d_0 and z_0 have been parametrized as

$$\sigma(d_0, z_0 \sin\theta)\,(p_T) = a + \frac{b}{p_T(GeV/c)}$$

where a and b are configurable parameters that can be set to the nominal values $a = 30\mu m$ and $b = 100\mu m$ or tightened to e.g. reduce the fake rate at low p_T.

The dependencies of the optimal cuts with the track n_{layers}, p_T and η have been approximated with the following formulas

- $\chi^2/\nu < \alpha_0 n_{\text{layers}}$
- $|d_0| < \left(\alpha_1 n_{\text{layers}}\right)^{x_1} \sigma_{d_0}(p_T)$
- $|\Delta z| < \left(\alpha_2 n_{\text{layers}}\right)^{x_2} \sigma_{z_0}(p_T, \eta)$
- $|d_0|/\delta d_0 < \left(\alpha_3 n_{\text{layers}}\right)^{x_3}$
- $|\Delta z|/\delta z_0 < \left(\alpha_4 n_{\text{layers}}\right)^{x_4}$.

where α_i and x_j are configurable parameters. For a high purity first reconstruction they have been set to $(\alpha_0 = 0.9, \alpha_1 = 0.3, \alpha_2 = 0.35, \alpha_3 = 0.4, \alpha_4 = 0.4)$ and $(x_1 = x_2 = x_3 = x_4 = 4)$ with resolution parameters $a = 30\mu m$ and $b = 10\mu m$.

In simple words the outcome of the selection optimization is that tracks with many layers ($\geq$10) can be kept with no need of quality cuts, while tracks with fewer layers can be selected with cuts that depend strongly on the track n_{layers} and also on the track p_T and η.

The same kind of track selection is applied at all stages of the iterative tracking described below, with the parameters set to apply looser cuts at each subsequent step.

3.6 Iterative Tracking

An *iterative* tracking procedure [3] has been developed in CMS to preserve high tracking efficiency while minimizing the fake rate. For each iteration, the following steps are applied:

1. A subset of reconstructed hits is selected by removing hits used in previous iterations.

   In the first iteration the complete set of reconstructed hits is available for reconstructing tracks. In the following iterations, hits associated with a track reconstructed in an earlier iteration are removed. The hits are only removed if they are associated with a *highPurity* track and pass a (configurable) χ^2 cut of 30.

2. Seed finding is performed on the available hits. The seeding configuration is the main difference between iterative steps.

3. Track reconstruction (building, filtering, fitting, smoothing) is performed using the available hits. Parameters in each stage can be tuned separately at each iteration to improve performance.

4. The track collection is cleaned following the criteria described in section 3.5.

5. The collection of tracks which pass the cleaning stage is stored.

The iterative tracking comprises six steps, as summarized in Tables 1–4. As mentioned above, the main difference between the iterations is in the initial seeding. However, there are also

significant differences in the parameters used for track reconstruction and track cleaning. The 0^{th} iteration uses pixel-triplet seeds while the 1^{st} iteration uses pixel-pair seeding, allowing it to recover tracks which failed to produce hits in three pixel layers. These two steps suffice to reconstruct the vast majority of high p_T tracks originating from the production vertex. The remaining steps are designed to pick up low p_T tracks and tracks which originate outside the pixel volume. The 2^{nd} and 3^{rd} iterations also use pixel seeding, but since many of the Tracker hits have already been locked by the time they run, they can use a very low p_T cut and a rather loose beam spot constraint, respectively. The final cleaning for these two steps allows tracks to pass either a tight vertex constraint and loose track quality cuts or a loose vertex constraint and tight track quality cuts. Finally, the 4^{th} and 5^{th} iterations use seeds from the Strip Tracker stereo layers, allowing them to find particles produced outside the volume of the Pixel Tracker.

Table 1: The six step iterative tracking configuration, indicating which layers are used to seed the tracks and the p_T and impact parameter cuts applied during seeding. (The symbol '*' indicates that the longitudindal impact parameter is calculated with respect to the reconstructed production vertex, instead of the centre of CMS.)

Iteration	Seeding Layers	p_T cut (GeV/c)	d_0 cut (cm)	z_0 cut (cm)
0	pixel triplets	0.5	0.2	15.9
1	pixel pairs	0.9	0.2	0.2*
2	pixel triplets	0.2	0.2	17.5
3	pixel pairs	0.35	1.2	7.0
4	TIB 1+2 & TID/TEC ring 1+2	0.5	2.0	10.0
5	TOB 1+2 & TEC ring 5	0.8	5.0	10.0

Table 2: The six step iterative tracking configuration gives the cuts applied during track building. In addition to a minimum p_T, N_{hit} is the minimum number of hits, N_{lost} is the maximum number of missing (invalid) hits, and $N_{\text{hit}}^{\text{rebuild}}$ is the minimum number of hits needed in the in-out step to trigger an out-in step.

Iteration	p_T cut (GeV/c)	N_{hit}	N_{lost}	$N_{\text{hit}}^{\text{rebuild}}$
0	0.3	3	1	5
1	0.3	3	1	5
2	0.1	3	1	5
3	0.1	4	0	5
4	0.1	7	0	5
5	0.1	7	0	4

4 Track reconstruction performance

In this section the performance of the CTF algorithm is evaluated on different categories of simulated samples. Firstly, the efficiency and quality of the pattern recognition is analyzed on simple events containing back-to-back pairs of muons, pions or electrons; secondly, the precision of the track-fitting is examined on the same samples through the resolution and pull distributions of the five reconstructed track parameters. Finally, the efficacy of track-finding and track-fitting is re-evaluated on different types of multi-track events that simulate LHC collisions.

A service known as *Track Associator By Hits* has been implemented for evaluating efficiency, resolutions and the other quantities related to the performance of the track reconstruction al-

Table 3: The six step iterative tracking configuration gives the cuts applied during the final cleaning. This cleaning produces one track collection with labels identifying which tracks pass the *Loose* (L), *Tight* (T), and *highPurity* (H) cuts. Steps 2 and 3 have two paths which emphasize track quality (Trk) or production vertex compatibility (Vtx). The cuts shown in this table are the maximum normalized χ^2, the minimum number of layers with a hit, the minimum number of 3D layers with a hit, and the maximum number of missing layers. Note that a layer could have two hits if the layer has a stereo view or if the track passes through an overlap region.

Iteration	χ^2_{max}			min layers			min 3D layers			max lost layers		
	L	T	H	L	T	H	L	T	H	L	T	H
0 & 1	2.0	0.9	0.9	0	3	4	0	3	4	∞	2	2
2 Trk	0.9	0.7	0.5	4	5	5	0	3	3	∞	1	1
2 Vtx	2.0	0.9	0.9	3	3	3	0	3	3	∞	1	1
3 Trk	0.9	0.7	0.5	4	5	5	2	3	4	1	1	1
3 Vtx	2.0	0.9	0.9	3	3	3	2	3	3	1	1	1
4	0.6	0.4	0.3	5	5	6	3	3	3	1	0	0
5	0.6	0.35	0.25	6	6	6	2	2	2	1	0	0

Table 4: The six step iterative tracking configuration gives the cuts applied during the final cleaning. This cleaning produces one track collection with labels identifying which tracks pass the *Loose* (L), *Tight* (T), and *highPurity* (H) cuts. Steps 2 and 3 have two paths which emphasize track quality (Trk) or production vertex compatibility (Vtx). The cuts shown in this table give the parameters associated with the production vertex requirement as described in section 3.5. With the exception of the *Loose* and *Tight* requirements for steps 0 & 1 which use track resolution parameters of $a = 30$ μm and $b = 100$ μm, all of the steps and collections use track resolution parameters of $a = 30$ μm and $b = 10$ μm.

Iteration	Expo-	$\alpha_1 = d_0^{\text{param}}$			$\alpha_3 = d_0^{\text{calc}}$			$\alpha_2 = d_z^{\text{param}}$			$\alpha_4 = d_z^{\text{calc}}$		
	nent	L	T	H	L	T	H	L	T	H	L	T	H
0 & 1	4	0.55	0.30	0.30	0.55	0.40	0.40	0.65	0.35	0.35	0.45	0.40	0.40
2 Trk	4	1.50	1.00	0.90	1.50	1.00	0.90	1.50	1.00	0.90	1.50	1.00	0.90
2 Vtx	3	1.20	0.95	0.85	1.30	1.00	0.90	1.20	0.90	0.80	1.30	1.00	0.90
3 Trk	4	1.80	1.10	1.00	1.80	1.10	1.00	1.80	1.10	1.00	1.80	1.10	1.00
3 Vtx	3	1.20	1.00	0.90	1.30	1.10	1.00	1.20	1.00	0.90	1.30	1.10	1.00
4	3	1.50	1.20	1.00	1.50	1.20	1.00	1.50	1.20	1.00	1.50	1.20	1.00
5	3	1.80	1.30	1.20	1.80	1.30	1.20	1.50	1.20	1.10	1.50	1.20	1.10

gorithm. The associator matches a simulated track to the corresponding reconstructed one if more than 50% of their measurements are connected. The association between the hits of the simulated particles and the clusters in the silicon detectors is possible because the simulation software traces the particles responsible of the signal on each channel of the Tracker. Strips and pixels that are fired by electronic noise are known as well.

4.1 Efficiency and fake rate

Both *global* and *algorithmic* efficiencies are analyzed. The former represents the fraction of charged particles, among all the simulated ones, that have a partner within the collection of reconstructed tracks; the latter considers only that simulated particles that leave at least three measurements on three different layers of the Tracker and, therefore, are track-able by the reconstruction algorithm. While the first kind of efficiency is affected also by properties of the detector, like its geometrical acceptance or its material budget, the algorithmic efficiency is exclusive related to the ability of the tracking algorithm to perform the track-finding task.

The effectiveness of the pattern recognition module is quantified also estimating its hit-finding efficiency and the purity of the measurements on-track: the former is the fraction of hits produced by the simulated particle that are included in the corresponding reconstructed track; the latter is the fraction of not-spurious measurements among all the hits of a track.

The fake rate is defined as the fraction of reconstructed tracks that don't have an associated particle among the simulated ones. This quantity represents the probability that a track provided by the reconstruction algorithm is either a combination of uncorrelated measurements or a genuine trajectory that is reconstructed including a large number of spurious hits. The fake rate is evaluated both requesting that the hit purity is greater than 50% and greater than 75%.

4.1.1 Muons

Muons are the charged particles that are best reconstructed in the Tracker. They mainly interact with the silicon detector through ionization and their energy loss by bremsstrahlung is generally negligible, except when muons are produced with an initial energy higher than about 100 GeV. Therefore these particles usually cross the whole volume of the tracking system, producing detectable hits on *all* the sensitive layers of the apparatus. Finally, muon trajectories are scattered exclusively by multiple-scattering, whose effects are straightforward to include inside the Kalman filter formalism. For isolated muons with a transverse momentum between 1 GeV/c and 100 GeV/c, the global tracking efficiency is higher than 99% in the full η-range of the Tracker acceptance and the efficiency does not depends on the ϕ direction of the particles (Fig. 9, top). The average hit-finding efficiency (Fig. 9, bottom-left) is higher than 99% both in the barrel and in the endcaps; it is still above 97% in the barrel-transition region where the layer navigation is more complicated (Sec. 3.2) and the material budget of the Tracker is more significant. The trajectory contamination due to spurious hits produced by electrical noise or δ-rays is lower than per-mille (Fig. 9, bottom-right).

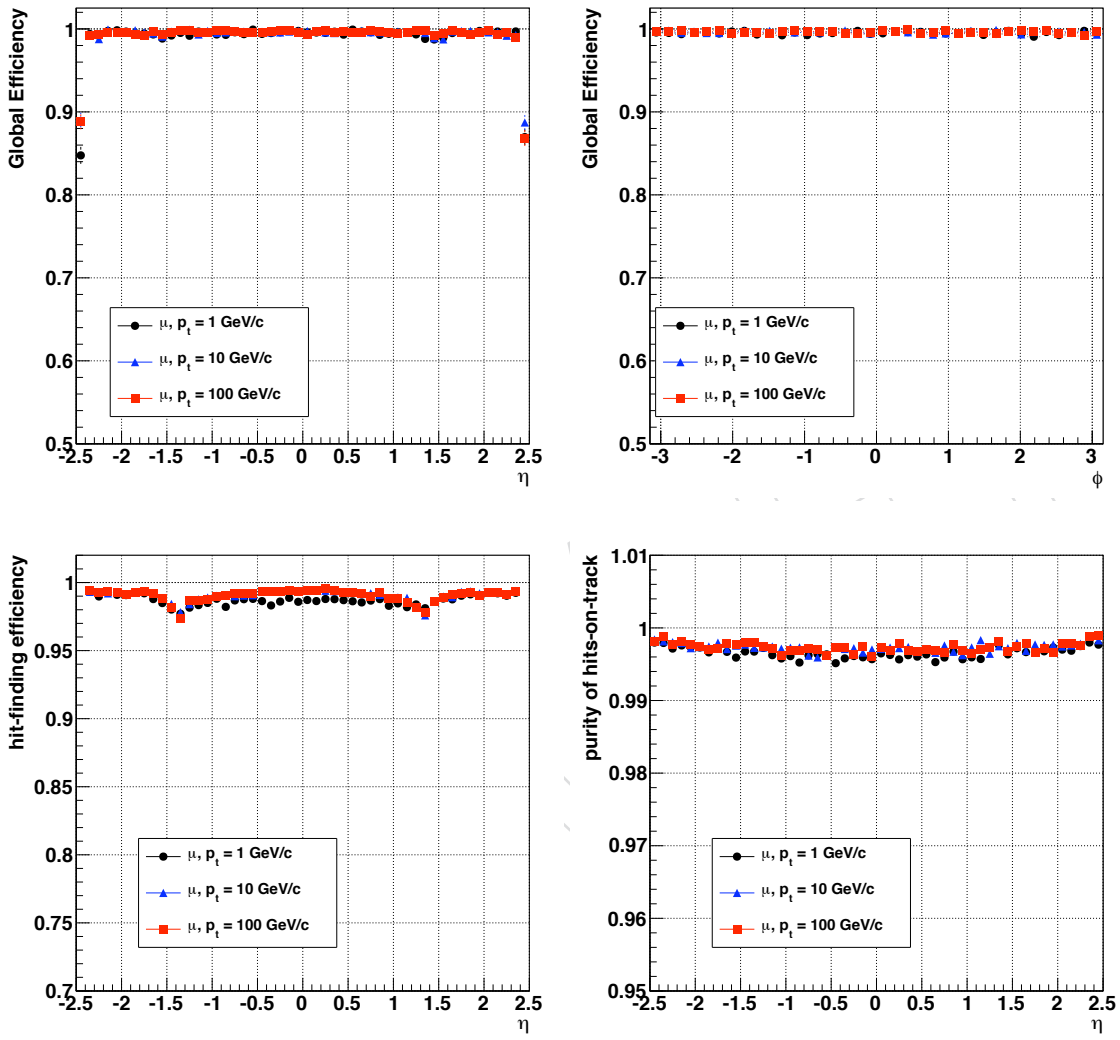

Figure 9: Global track reconstruction efficiency as a function of η (top-left) and ϕ (top-right) for muons of transverse momenta of 1, 10 and 100 GeV/c. Hit-finding efficiency and purity of the measurements on-track (bottom).

4.1.2 Pions

Charged pion particles undergo multiple scattering and energy loss by ionization while they cross the tracker volume, like muons do. However, pion particles, and all hadrons in general, are also subjected to elastic and inelastic *nuclear* interactions. The former effect introduces tails in the distribution of the scattering angle that is expected assuming the multiple scattering as the solely contributor to the stochastic deflection of the particles. In the current implementation of the CTF algorithm, the description of the process noisy (eq. **??**) doesn't account for elastic nuclear interactions and this brings to a lower hit finding efficiency during the reconstruction of hadrons compared to muons (Fig. 10, left).

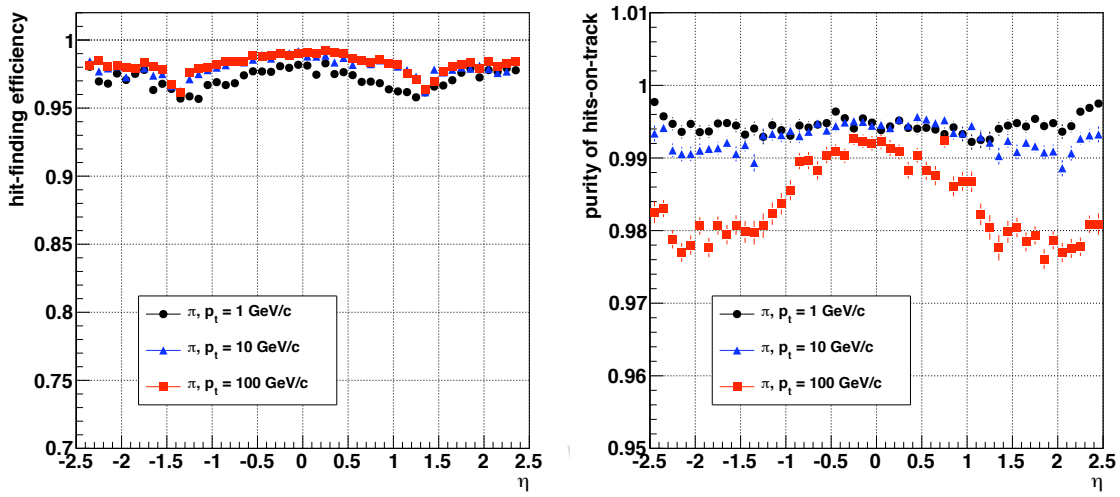

Figure 10: Hit-finding efficiency (left) and hit-purity (right) for pions of transverse momenta of 1, 10 and 100 GeV/c.

The effect of an inelastic nuclear interaction between a hadron and the detector's material is the transformation of the incoming particle into two or more secondary particles. Depending on the location within the Tracker volume where this process occurs, the number of measurements produced by the primary particle on the sensitive layers of the detector can be pretty small, deteriorating the precision with which the trajectory parameters can be estimated (Sec.4.2).

Besides, when the number of crossed layers is less than three, the particle trajectory cannot be reconstructed at all by the CTF algorithm. Inelastic nuclear interactions are the main source of tracking inefficiency for hadrons, in particular in those regions of the Tracker where the material budget is more significant and the probability that particles undergo nuclear interaction is higher (Fig. 11, top-right). However, elastic nuclear interactions, and the consequent hit-finding inefficiency, not only reduce the global efficiency of the pattern recognition software, but they also lower the *algorithmic* efficiency: depending on the η-bin, up to 9% of the simulated particles are not reconstructed (Fig. 11, top-left), despite they leaves measurements on three or more layers of the tracking system. This effect is more significant for low energy hadrons due to their higher cross section for nuclear interactions [**?**].

The secondary particles produced in inelastic processes affect also the hit-purity of the reconstructed trajectories. The products of the nuclear interaction are usually emitted from the nuclear vertex with directions tangent to the trajectory of the incoming particle and therefore it is not uncommon that the trajectory builder combines the hits of the primary particle with those

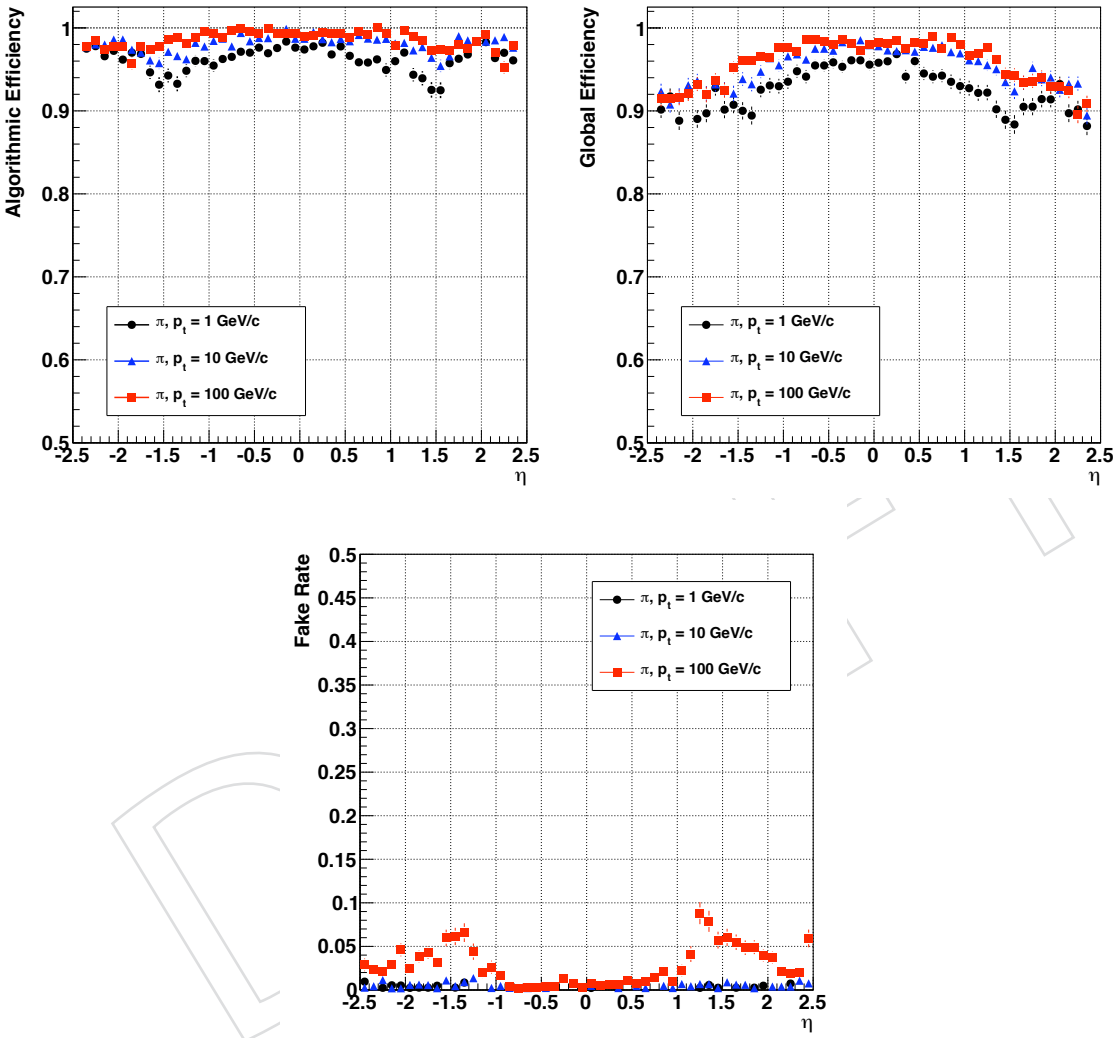

Figure 11: Algorithmic tracking efficiency (top-left), global tracking efficiency (top-right) and fake rate (bottom) as a function of η for pions of transverse momenta of 1, 10 and 100 GeV/c. Results refer to reconstructed tracks that are requested to have a hit-purity greater than 50%.

of a secondary particle. The degradation in the purity of the measurements on track (Fig. 10,

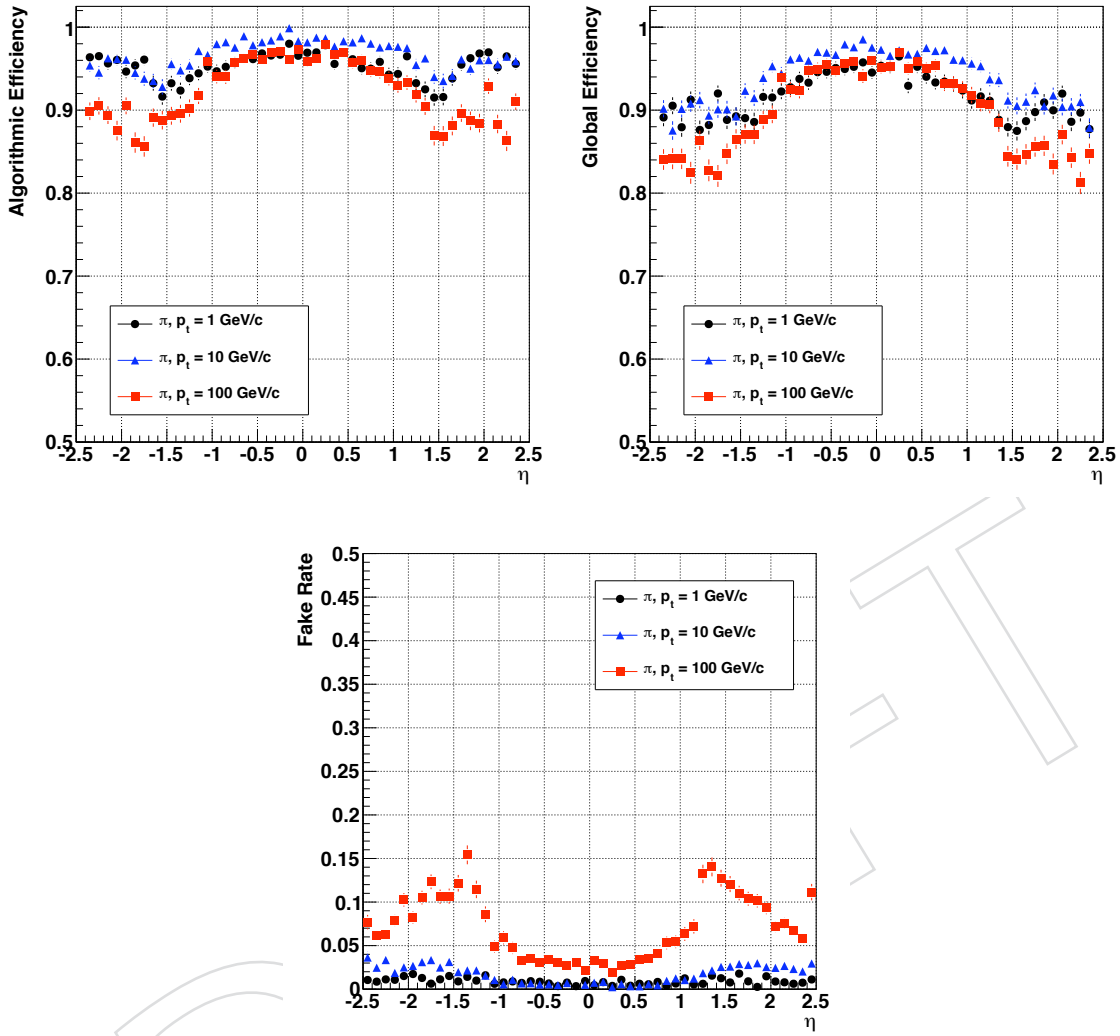

Figure 12: Algorithmic tracking efficiency (top-left), global tracking efficiency (top-right) and fake rate (bottom) as a function of η for pions of transverse momenta of 1, 10 and 100 GeV/c. Results refer to reconstructed tracks that are requested to have a hit-purity greater than 75%.

right) is more accentuated for high energy hadrons than for low energy ones: indeed the probability that the trajectory builder merges together separate tracks increases with the energy of the primary hadrons because of the larger number of secondary particles that are produced in the interaction and because of the smaller variation between the curvature of the primary particle and that of the secondary particles. The merging of hadron trajectories affects significantly the efficiency and fake rate distributions depending on which cut on the hit-purity is used in the definition of correctly-reconstructed tracks: Fig. 11 and Fig. 12 shows efficiency and fake rate distributions that are obtained requesting respectively a hit-purity of 50% and 75% for the reconstructed tracks. The degradation in efficiency and the increase in fake rate are correlated, as expected, and the variation is more significant for the sample with the highest energy particles. In general, the merging of separate trajectories during reconstruction is more common in the transition and endcap regions of the tracker, due to the higher material budget. In the barrel-endcap transition region, the probability the reconstruction algorithm returns

merged trajectories is even higher because of the large extrapolation length during some iterations of the pattern recognition module in which the trajectory builder navigates from barrel to endcap layers. While the fake rate is generally lower than 1% for pions with a transverse momentum of 1 or 10 GeV/c, the probability that a 100 GeV/c pion is incorrectly reconstructed peaks to about 10% for particles produced with $|\eta|$ around 1.5, regardless of a looser cut on the hit-purity (Fit. 11, bottom).

4.1.3 Electrons

Electrons and positrons are subjected to multiple scattering and energy loss by ionization, like the muon leptons, and they are not exposed to the nuclear interactions that affect significantly the reconstruction of charged hadrons. Due to the remarkable material budget of the tracking system, which ranges between 0.4 and 1.8 radiation lengths depending on the particle direction (Sec. **??**), most of the electrons produced in LHC collisions lose a large fraction of their energy by means of bremsstrahlung radiation before they reach the outer layers of the Silicon Tracker. In some cases, the radiated photons have enough energy to convert in electron-positron pairs and initiate authentic electromagnetic showers inside the Tracker volume.

Although the mechanism responsible for the bremsstrahlung radiation is completely different from that of nuclear interaction for hadrons, it influences the performance of the track reconstruction algorithm in a similar way. The radiation of soft photons usually doesn't prevent the reconstruction of electrons, however it introduces tails in the distribution of the electron's energy loss that are not taken into account by the CTF algorithm in its default configuration. The principal consequence of this aspect is a decrease of the hit-finding efficiency, analogous to the reduction caused by elastic nuclear interactions for hadrons. The radiation of high energy photons has an impact on the reconstruction of the electrons which is similar to that of the inelastic nuclear interactions on the reconstruction of charged hadrons: firstly, when an electron loses most of its energy before reaching the outer layer of the tracker, there is a reduction in the number of measurements, recorded by the detector, that are really useful to determine the trajectory's parameters at the collision point; secondly, when a radiated photon converts or creates an electromagnetic shower, the trajectory builder can fail to properly recognize the hits of the secondary particles and it may try to merge the trajectory of the primary electron with other uncorrelated track segments. The latter effect further deteriorates the precision with which the trajectories corresponding to early-radiating electrons are reconstructed and it seems to be the principal source of charge misidentification for electrons.

In its default configuration, the CTF algorithm efficiency for reconstructing electron trajectories that have a hit-purity greater than 50% is above 99% in the barrel region of the Tracker. The efficiency is generally greater than 95% in the full η range of the detector geometrical acceptance (Fig. 13, top-left). If the hit-purity of the reconstructed trajectories is required to be greater than 75%, the global tracking efficiency decreases to about 90% in the transition region, but it remains almost unchanged in the barrel and the endcap regions (Fig. 13, bottom-left). For both cuts on the hit-purity, the fake rate is correlated to the tracking inefficiency and it peaks in the barrel-endcap transition region where the amount of material traversed by the electrons is the largest and the probability of producing electromagnetic showers is maximum (Fig. 13, right-column).

Better performance in term of efficiency and fake rate can be obtained using an ad-hoc configuration of the CTF trajectory builder that is optimized for electron reconstruction. In particular, the service used to extrapolate the trajectory parameters between the layers of the Tracker can activate the inclusion of the process noise term due to the bremsstrahlung radiation. Such a con-

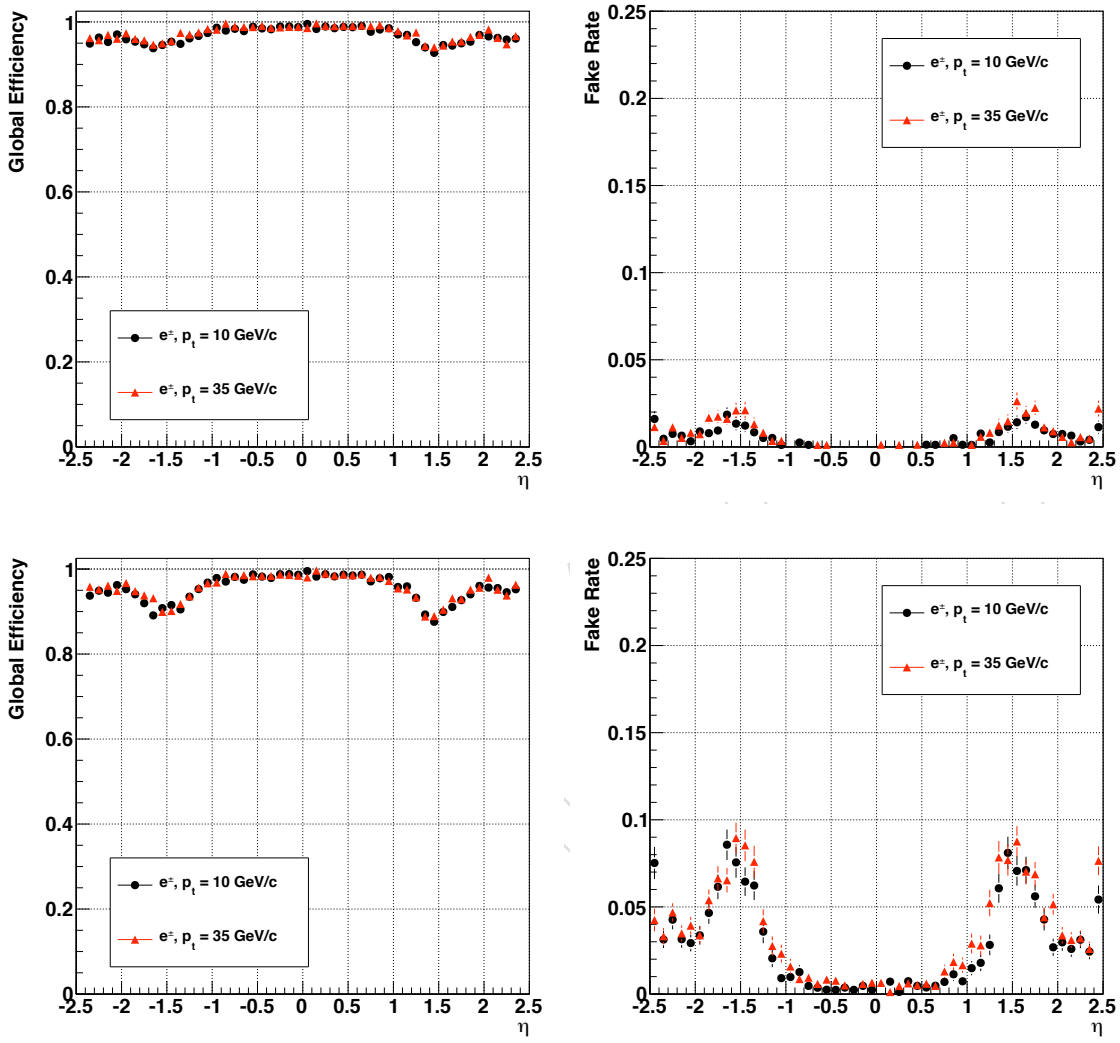

Figure 13: Global tracking efficiency (left) and fake rate (right) for electrons and positrons of transverse momenta of 10 and 35 GeV/c. The histograms are produced with two different requirements on the hit-purity of the reconstructed tracks: 50% purity (top) and 75% purity (bottom).

figuration of the trajectory builder implies a much slower execution of the track reconstruction sequence that is not affordable for the reconstruction all the tracks in a generic LHC collision event. Nevertheless it is used for the reconstruction of specific electron candidates within small regions of interest of the Tracker that are identified by the electromagnetic calorimeter [**?**].

4.2 Track reconstruction resolution

This section analyzes and quantifies the level of precision that is achieved in estimating the parameters of the reconstructed trajectories. In the context of the reconstruction software of CMS, the five parameters used to describe a track are: 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 ($d_0 = y_0 * \cos\varphi - x_0 * \sin\varphi$, 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.

For each of the five track parameters, the resolution is estimated as a function of the pseudorapidity of the simulated charged particle. In every η-bin, the resolution is expressed both as the σ of a gaussian fit of the residuals' distribution[1] and as the RMS of the same distribution. The pulls of the five reconstructed parameters are analyzed to verify whether the uncertainty that is provided for each parameter by the final track fit is a sensible measure of the error of the estimator or not.

Resolutions and pull plots are presented for the same eight samples of isolated muons, pions and electrons that have been used in the previous section to study the efficiency and the fake rate of the reconstruction algorithm.

4.2.1 Muons

Muons are the charged particles whose trajectories are estimated with the best precision in the silicon tracking system of CMS. Since they don't undergo inelastic nuclear interactions and the amount of energy that radiate by bremsstrahlung is negligible (for transverse momenta up to about 100 GeV/c), muon particles that are produced in LHC collisions cross the whole Tracker volume and therefore they can be reconstructed with the largest lever arm available.

Figure 14 shows the resolution of the five track parameters for samples of isolated 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 first two plots of Figure 14. 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 pseudorapidity of $|\eta| = 0.4$ can be attributed to the beneficial effect of charge interpolation in the position estimation of pixel clusters (Sec. **??**): in the barrel, as the angle with which the tracks cross the pixel layers increases the clusters become wider, distributing the signal on more than one pixel and improving the position resolution. Similar considerations about the relation between track momentum and effect of multiple scattering and between track pseudorapidity and pixel hit precision are valid also for the plots of the φ and $\cot\vartheta$ parameters, which are shown in the second and third histogram of the same figure. The resolution of the transverse momentum is shown in the last plot of Figure 14. At high momentum (100 GeV/c), the resolution is around 1-2% up to a pseudo-rapidity of $|\eta| = 1.6$; for higher values of $|\eta|$ the lever arm used for the measurement is reduced. The degradation around $|\eta| = 1.0$ and beyond is due to the gap between the barrel and the end-cap disks (Fig. 1) and 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 (Tab. **??** and **??**). At a transverse momentum

[1] The residual is the difference between the parameter of the reconstructed track and the corresponding value of the simulated particle's parameter.

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. In all the five reported resolution plots, the difference between the sigma of the gaussian fit and the RMS of the residuals distributions is generally very small for the 10 and 100 GeV/c samples, indicating that the distributions of the track parameter errors are properly described by single gaussian functions and have few entries populating the tails. The worse discrepancy between sigma and RMS is about 40% and regards the resolution of d_0 for the 1 GeV/c sample: for tracks with $\eta = 0$, the sigma of the fit is equal to 70 μm compared to a RMS of about 100 μm. **ADD explanation? is it because of the bigger crossing angle and the higher probability of producing δ-rays?**

The precision with which the CMS implementation of the Kalman filter and smoother estimates the *uncertainty* on the reconstructed track parameters is analyzed through the pull distributions of all the five trajectory parameters (Fig. 14). The pull is the distribution of the parameter residual divided by the uncertainty on the reconstructed parameter, as it is provided by the final track fit. If the estimator is unbiased and the parameter uncertainty is evaluate correctly, the pull has a gaussian distribution with mean equal to 0 and sigma equal to 1. All the gaussian functions, obtained from the fit of the pull distributions of the five track parameters, have standard deviations within the range [0.85,1.10], indicating that the parameter errors are generally correctly estimated, but there is still room for improvement. Like for the tails in the resolution distribution, also for the pulls the worse results are those corresponding to the low energy sample. This could be an indication of some limits in the current description of the multiple scattering effects.

Another quantity that can be used to estimate the quality of the track fit is the distribution of the normalized-χ^2 of the reconstructed track . Unlike the pull distribution, the χ^2 includes also contributes from the off-diagonal terms of the error covariance matrix (eq. **??** and **??**). The bottom-right plot of Fig. 14 represents the average normalized-χ^2 as a function of the pseudo-rapidity of the reconstructed tracks in the three muon samples. The values corresponding to 1 GeV/c tracks produced in the barrel region show the worse discrepancy (of about 20%) with the theoretical curve (i.e. a flat distribution equal to 1). The results for the 10 and 100 GeV/c samples are compatible with the expected values within 10%.

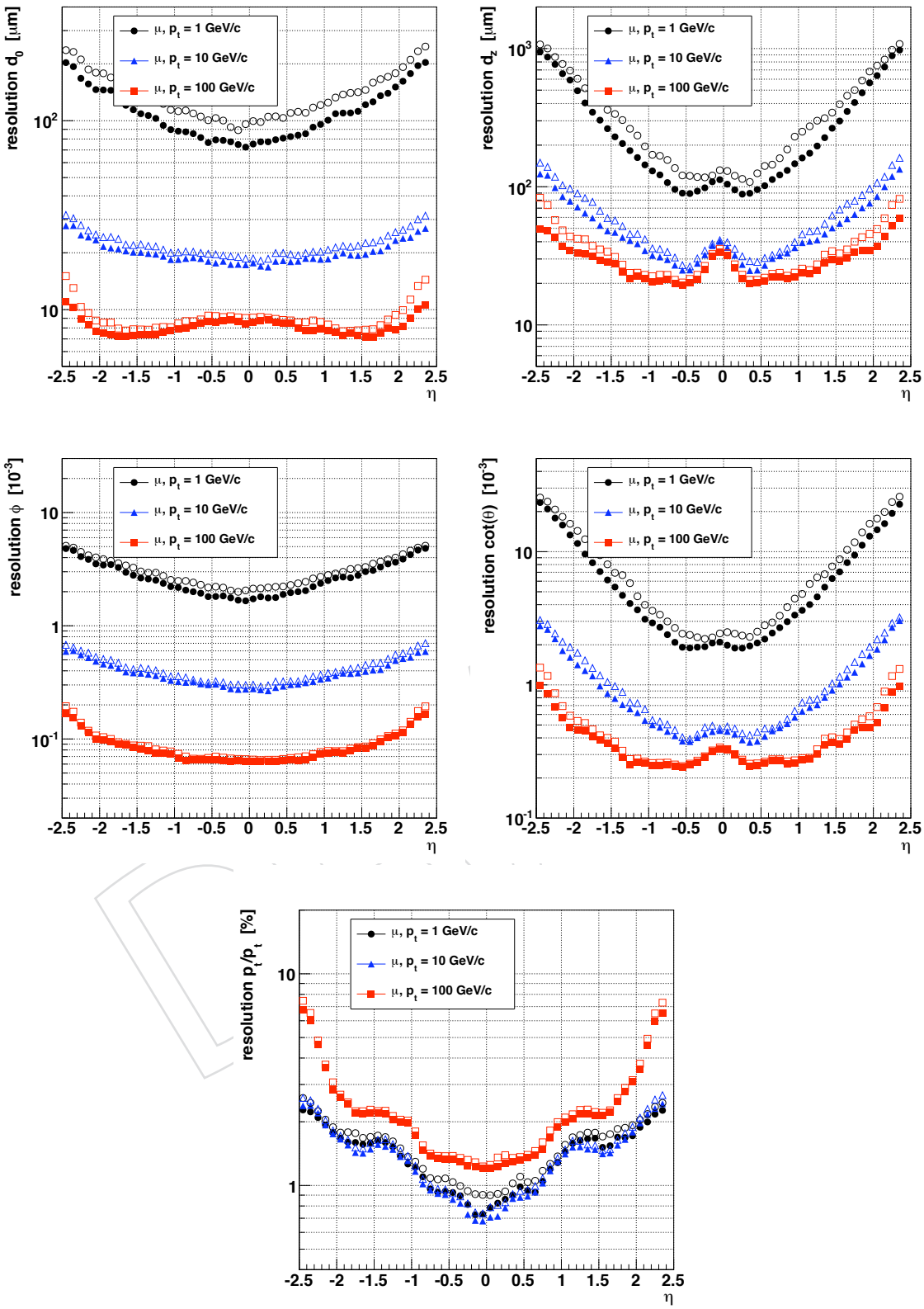

Figure 14: Resolution, as a function of pseudorapidity, of the five track parameters for isolated muons with transverse momenta of 1, 10 and 100 GeV/c. From top to bottom and left to right: transverse momentum, φ, cot ϑ, transverse and longitudinal impact parameter. For each bin in η, the full symbol corresponds to the sigma of a gaussian fit of the residuals distribution and the empty symbol represents the RMS of the same distribution.

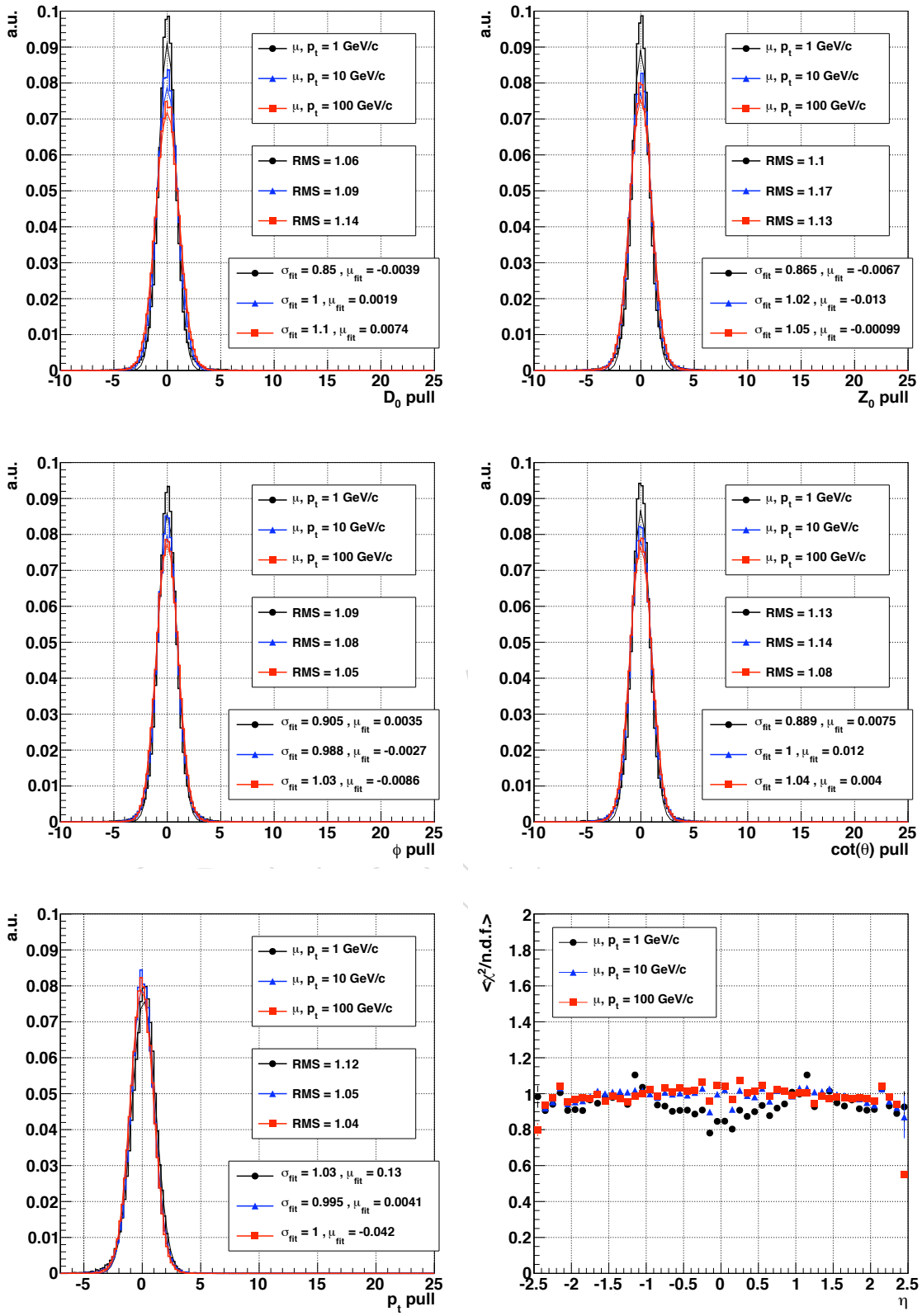

Figure 15: The first five plots correspond to the pull distributions of the five track parameters for isolated muons with transverse momenta of 1, 10 and 100 GeV/c. From top to bottom and left to right: transverse momentum, φ, $\cot\vartheta$, transverse and longitudinal impact parameter. Both the σ and mean of a single-gaussian fit and the RMS of the pull distribution are printed. The sixth plot is the average normalized-χ^2 of reconstructed tracks as a function of pseudorapidity.

4.2.2 Pions

Charged pion particles that don't undergo nuclear interactions behave similarly to muons being subjected to the same multiple scattering effects and to the same mechanism of energy loss by ionization. The trajectories of this subset of pions are reconstructed by the CTF algorithm with a precision which is very close to that achieved for muons. The five plots in Fig. 16 show the resolutions of the five track parameters. As expected, the results obtained with a gaussian fit of the distributions of the residuals are very close to those showed for muon trajectories in Fig. 14. However, the resolutions obtained using the RMS show a different pattern respect to the same kind of resolutions evaluated for the muon samples. The RMS of transverse impact parameter, φ direction and transverse momentum residuals are up to a factor 2 wider than the corresponding results from a gaussian fit. These tracks whose parameters are measured less precisely are mainly pions that interact inelastically; since these particles don't reach the outer layers of the tracking system, their trajectories can only be measured using a shorter lever arm. Considering only the charged pions that succeed to cross the whole Tracker volume without inelastic interactions, the difference between RMS and gaussian fit resolutions is significantly reduced (Fig. 18). While the length of the lever arm affects directly the estimate of the transverse momentum, it influences also the measurement of those quantities, like d_0 and φ, that are obtained propagating the trajectory state within the *transverse* plane from the innermost sensitive surface of the tracking system to the impact point.

The pull distributions of the five trajectory parameters and the average normalized-χ^2 are shown in Fig. 17. Like for the muons samples, the sigmas of the gaussian fit of the pulls are close to the expected value 1, ranging between 0.86 and 1.14. However, the RMS of the same pull distributions are generally wider, extending between 1.18 and 1.65. The uncertainty on the estimates of the reconstructed trajectory parameters is likely underestimated due to the effects of nuclear elastic interactions that are not incorporated in the current version of the trajectory propagator (see also analysis of pion hit-finding inefficiency in Sec. 4.1).

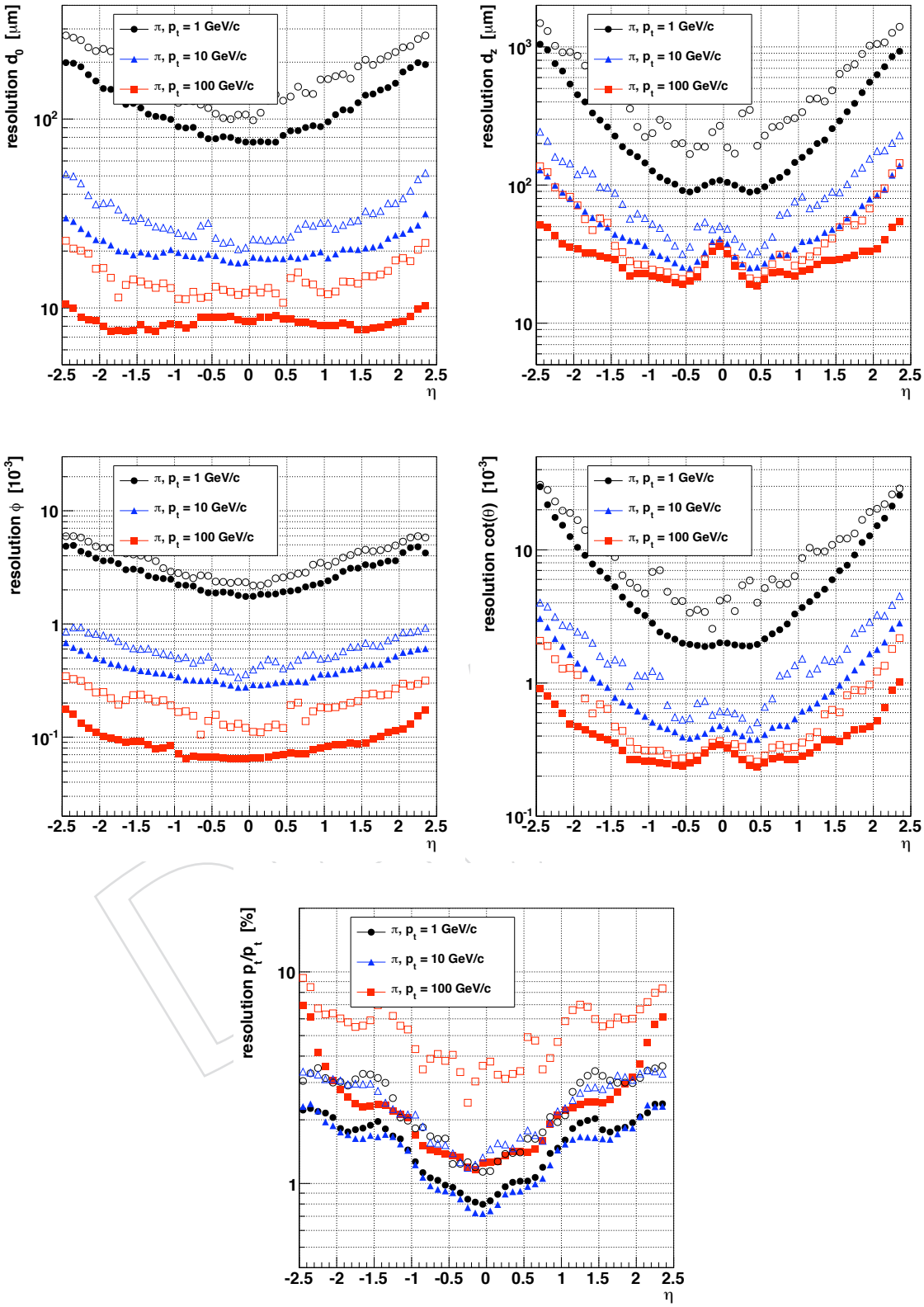

Figure 16: Resolution, as a function of pseudorapidity, of the five track parameters for isolated pions with transverse momenta of 1, 10 and 100 GeV/c. From top to bottom and left to right: transverse momentum, φ, $\cot\vartheta$, transverse and longitudinal impact parameter. For each bin in η, the full symbol corresponds to the sigma of a gaussian fit of the residuals distribution and the empty symbol represents the RMS of the same distribution.

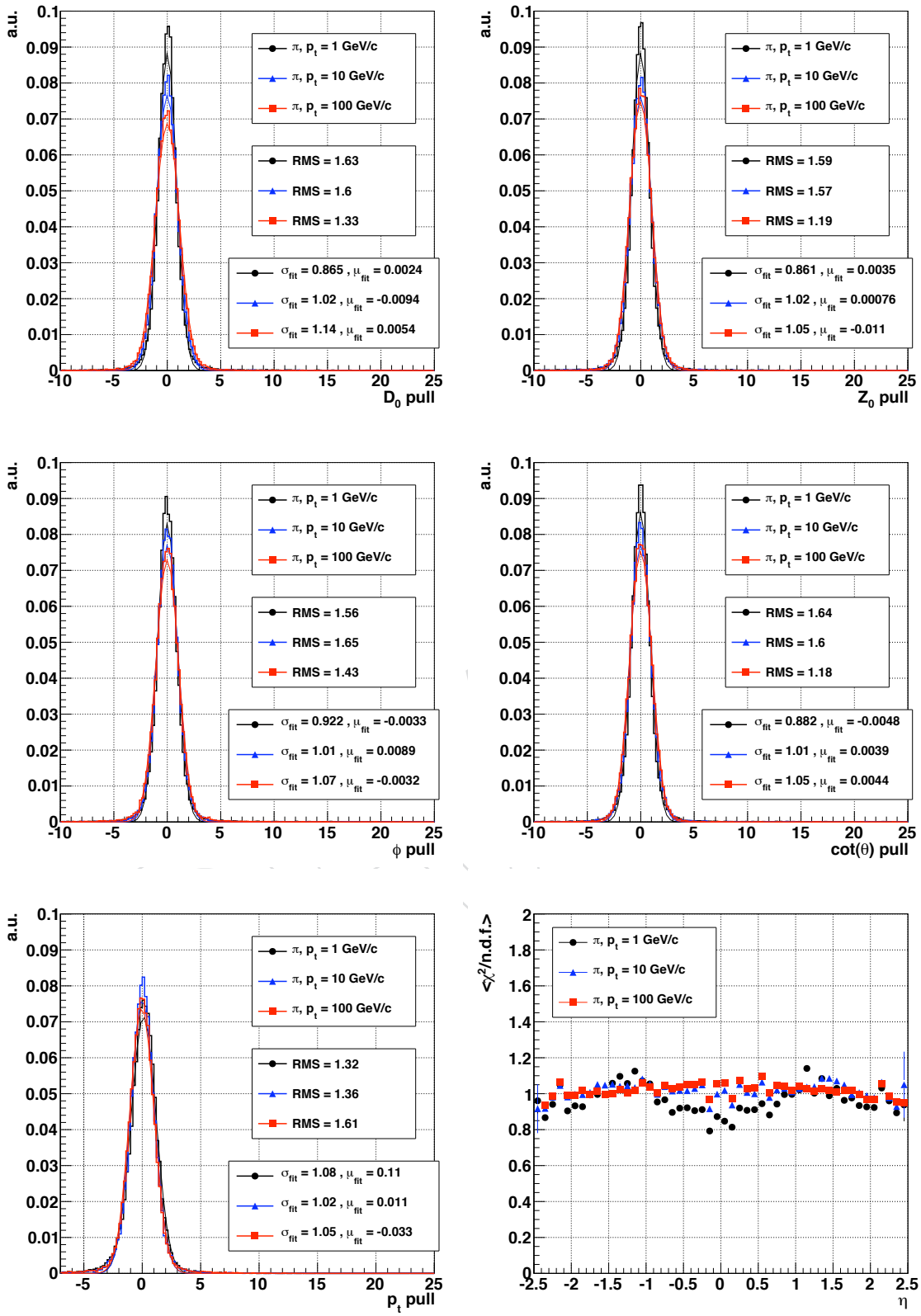

Figure 17: The first five plots correspond to the pull distributions of the five track parameters for isolated pions with transverse momenta of 1, 10 and 100 GeV/c. From top to bottom and left to right: transverse momentum, φ, $\cot\vartheta$, transverse and longitudinal impact parameter. Both the σ and mean of a single-gaussian fit and the RMS of the pull distribution are printed. The sixth plot is the average normalized-χ^2 of reconstructed tracks as a function of pseudorapidity.

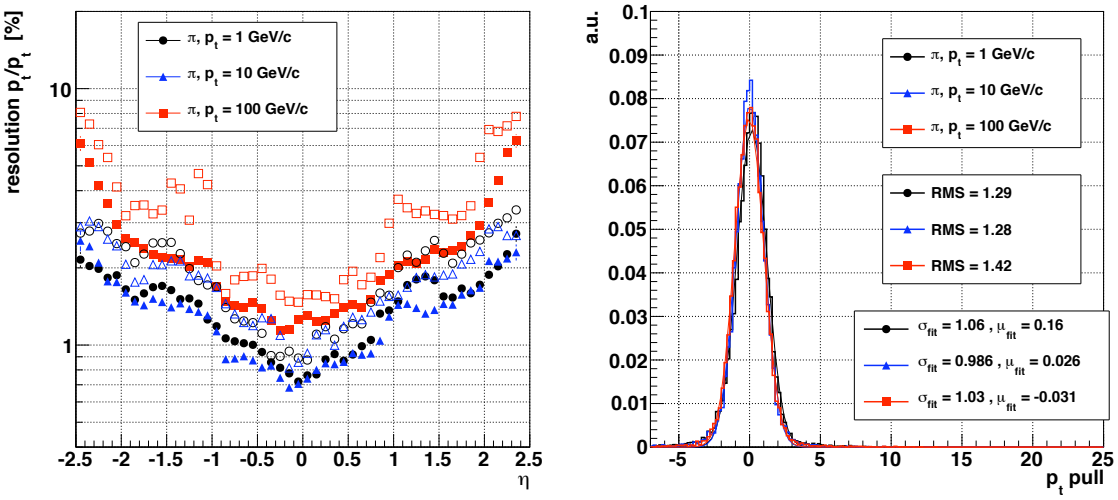

Figure 18: On the left: resolution, as a function of pseudorapidity, of the reconstructed transverse pions for isolated pions with transverse momenta of 1, 10 and 100 GeV/c; for each bin in η, the full symbol corresponds to the sigma of a gaussian fit of the residuals distribution and the empty symbol represents the RMS of the same distribution. On the right: pull distribution of the transverse momentum for the same simulated samples of pions; both the σ and mean of a single-gaussian fit and the RMS of the pull distribution are printed. The two plots have been filled only for those simulated hadrons that reached the outer layers of the tracker without undergoing inelastic nuclear interactions.

4.2.3 Electrons

The resolutions of the five trajectory parameters are shown in Fig. 19 for the two samples of isolated electrons with transverse momenta of 10 and 35 GeV/c. Like for pions, the resolution of the three quantities measured in the transverse plane (d_0, φ and transverse momentum) are significantly different whether they are evaluated with the RMS of the residuals or with a gaussian fit. Due to the effect of bremsstrahlung radiation, some electrons don't reach the outer layers of the tracking system and, like for pions undergoing inelastic nuclear interactions, they can be reconstructed only using a shorter lever arm. Besides, since the process noise due to the bremsstrahlung is not implemented in the default version of Kalman filter and smoother used by the CTF algorithm, the estimates of the trajectory parameters that are reconstructed in the transverse plane have systematically underestimated uncertainties. Pull distributions of the five track parameters are shown in Fig. 20: the trajectory parameters that estimated in the r-z plane have pull distributions centered around zero and with standard deviations differing from 1 by less than 15%; the pull distributions of d_0, φ and transverse momentum measurement have RMS ranging between 3 and beyond 6; finally, the measurement of the transverse momentum is systematically underestimated, as shown by the asymmetric shape of the corresponding pull.

The asymmetric and not-gaussian process noise due to bremsstrahlung can be included in an extended version of the Kalman filter method which is known as Gaussian Sum Filter [**?**]. This different filter has been implemented in CMS [**?**] and it is used for fitting the trajectories of the particles that are reconstructed in the specific electron reconstruction sequence driven by energy deposits in the electromagnetic calorimeter [**?**].

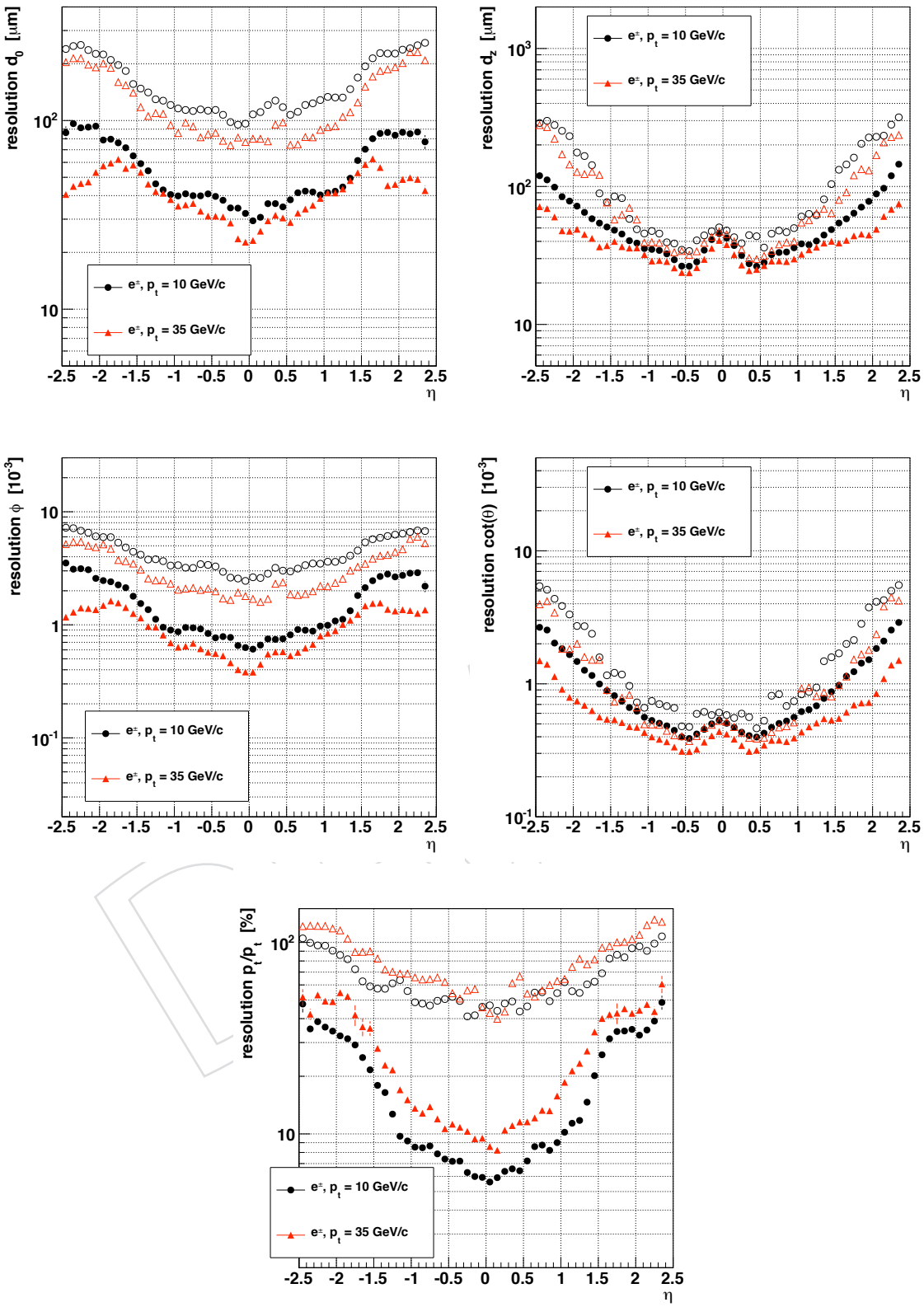

Figure 19: Resolution, as a function of pseudorapidity, of the five track parameters for isolated electrons with transverse momenta of 10 and 35 GeV/c. From top to bottom and left to right: transverse momentum, φ, $\cot\vartheta$, transverse and longitudinal impact parameter. For each bin in η, the full symbol corresponds to the sigma of a gaussian fit of the residuals distribution and the empty symbol represents the RMS of the same distribution.

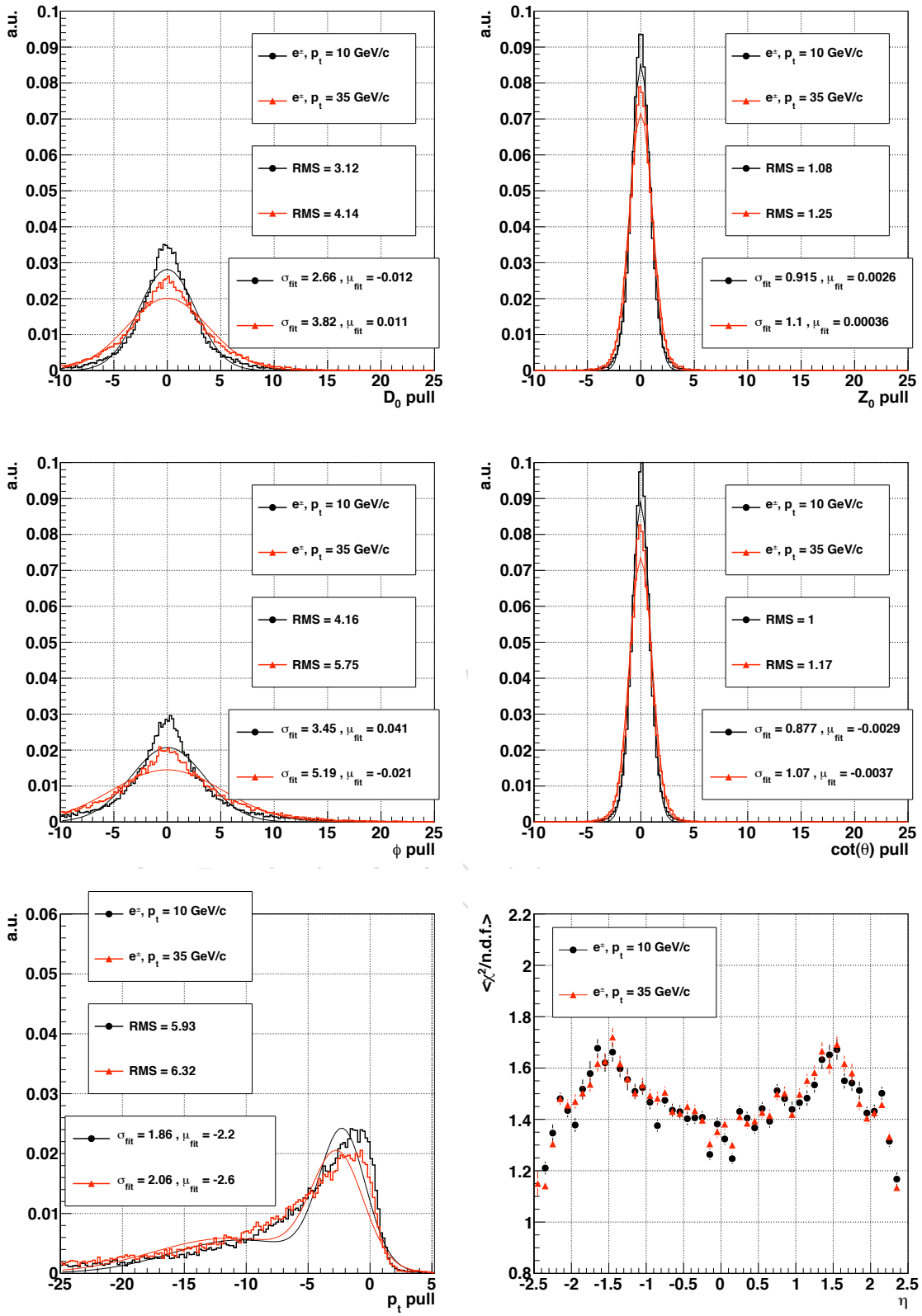

Figure 20: The first five plots correspond to the pull distributions of the five track parameters for isolated electrons with transverse momenta of 10 and 35 GeV/c. From top to bottom and left to right: transverse momentum, φ, $\cot\vartheta$, transverse and longitudinal impact parameter. Both the σ and mean (of the narrower gaussian function) of a double-gaussian fit and the RMS of the pull distribution are printed. The sixth plot is the average normalized-χ^2 of reconstructed tracks as a function of pseudorapidity.

4.3 Tracking performance on simulated LHC collisions

This section analyzes the performance of the CTF tracking software in reconstructing the trajectories of *not-isolated* charged particles that are produced in simulated LHC collisions. Compared to the results shown in Section **??** for isolated particles, the tracking performance that are shown in this section are affected also by a feature that is specific to multi-track LHC collision events: the simultaneous presence, in the Tracker's detectors, of many measurements due to pile-up collisions, low energy spiraling particles (loopers) and hits left from the tens or hundreds of primary particles usually produced in those LHC events that are selected by the Trigger system. In the following some examples of the difficulties encountered by the tracking algorithm during the reconstruction of these events:

- Particles are produced within a very collimated jet and their position hits are closer than the typical uncertainty of the extrapolated trajectory position on the detectors (Fig. 7). In this case, the trajectory builder is unable to assign the measurements to the corresponding trajectories without ambiguity: e.g. hits corresponding to two distinct charged particles may be mixed into two reconstructed tracks that don't describe accurately neither of the actual trajectories of the two particles.
- The trajectories of nearby particles are enough detached on the outer layers of the Tracker and they are correctly identified by the pattern recognition module. Nevertheless, their measurements on the innermost layers are so close to each other that the reconstruction algorithm incorrectly assign the hits to the corresponding trajectories. Distinct particles can be so near that their ionization signals are merged into a single cluster on the innermost silicon detector. In this scenario, even if the individual trajectories are reconstructed and their transverse momentum is generally well-measured, the resolution of the impact parameter is degraded.
- Due to the several hits produced by loopers and particles originating from pile-up collisions, it is possible that completely uncorrelated measurements are arranged as the legitimate trajectory of a charged particle and therefore are reconstructed as a ghost track.

Figure 21 shows, as a function of pseudorapidity, the global tracking efficiency of the CTF algorithm in reconstructing primary charged particles produced in LHC collisions. These simulated events contain di-jets with transverse momenta within intervals of 15-20, 80-120, 120-170 and 3000-3500 GeV/c and multiple-jets produced from $t\bar{t}$ pairs. The efficiency for the three samples with jets of lower transverse momenta (Fig. 21, left) is a sort of average of the reconstruction efficiencies for isolated pions which were shown in Fig. 12. This result is compatible with the fact that the majority of charged particles produced in LHC collisions consists of pions, followed by other hadrons whose trajectories are identified with similar efficacy. The higher Tracker occupancy inside the jets doesn't seem to comprise the pattern recognition in the range of energies of these three di-jets samples. However, the tracking efficiency is significantly degraded within jets of higher energy and that are more collimated: for events with di-jets of more than 3 TeV/c of transverse momentum, the global tracking efficiency decreases from approximately 94% to about 74% for tracks produced in the barrel[2] (Fig. 21, right).

The tracking efficiency as a function of the particle transverse momentum is approximately constant in the p_t range between 5 and 30 GeV/c and then it starts decreasing (Fig. 22, left) because lower energy particles have a larger nuclear cross section. For transverse momenta below 0.8 GeV/c the efficiency decreases quickly from about 80% to zero (Fig. 22, right) due to

[2]Given the 7 TeV beams of the LHC, energy conservation and geometrical arguments requires that di-jets of 3.5 TeV/c transverse momentum can be produced only within $|\eta| < 1.3$.

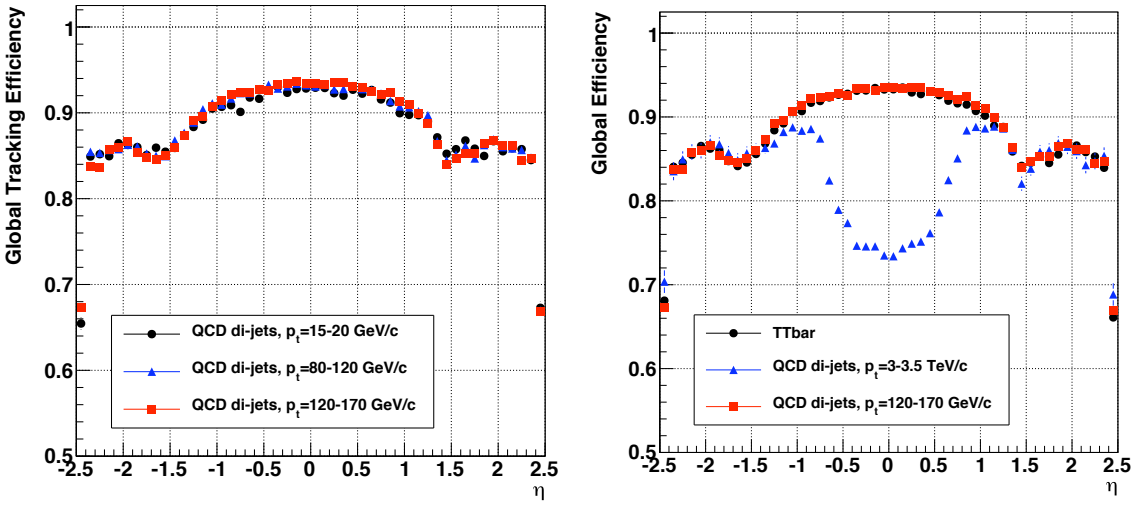

Figure 21: Global tracking efficiency evaluated on simulated LHC collisions containing di-jets with transverse momenta within intervals of 15-20, 80-120 and 120-170 GeV/c (left); the plot on the right contains results also for di-jets of p_t within 3-3.5 TeV/c and multi-jets from $t\bar{t}$ pairs. The efficiency is estimated as a function of pseudorapidity. The events are produced in full simulation including the underlaying event, but without superimposing the pile-up collisions.

several causes:

- The pions-proton nuclear cross section increases rapidly for pions of energies below 0.7 GeV/c.
- The track selection criteria (see Sec. 3.5) are much tighter for low momentum trajectories since they are the main source of fake tracks.
- In order to match the constraints on track reconstruction's timing that are set by the available computing resources, the version of the CTF algorithm used in this study reconstructs only trajectory seeds from hit-triplets (pairs) that have a transverse momentum higher than 0.3 GeV/c (0.6 GeV/c).
- In evaluating the process noise, the trajectory parameters propagator assumes that the mass of the particle is that of a pion, i.e. the most common particle among those originating from LHC collisions. While this assumption is legitimate for relativistic particles, it breaks at low energies.

For the analyzed multi-track samples, the fake rate (Fig. 23, top) has the same functional dependency on η as the fake rate evaluated on isolated pions samples (Sec. 4.1.2). The fake rate generally doesn't vary as a function of the trajectory transverse momentum, except for values below 1 GeV/c: the smaller the energy of an initial trajectory seed, the larger is the search window that is used during the early iterations of the pattern recognition to collect additional trajectory measurements and higher the probability to accept uncorrelated hits and to produce a ghost track. The hit-finding efficiency and hit-purity (Fig. 23, bottom) are similar to those measured for the isolates pion samples and worsen as the energy of the jets increases.

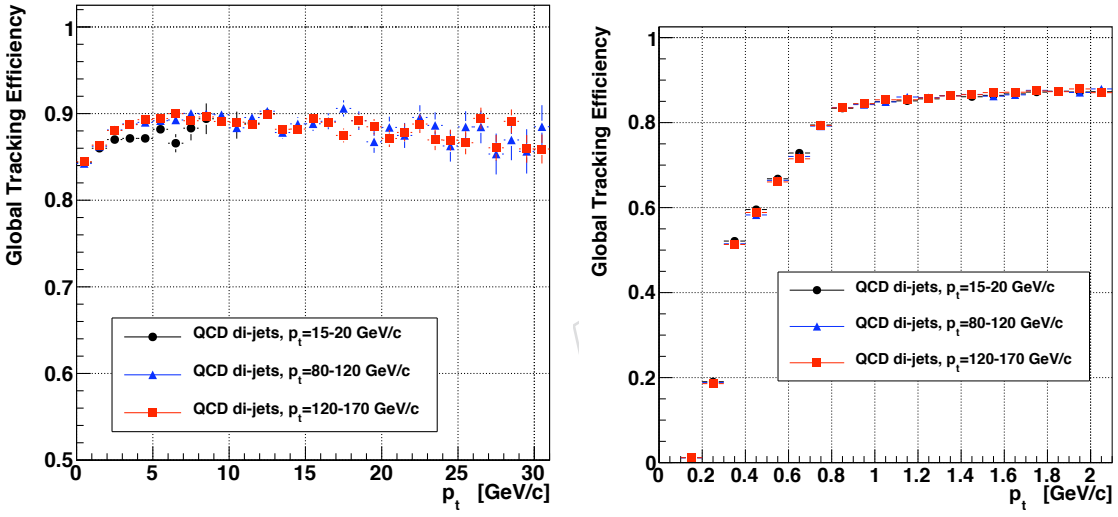

Figure 22: Global tracking efficiency evaluated on simulated LHC collisions containing di-jets with transverse momenta within intervals of 15-20, 80-120 and 120-170 GeV/c; The efficiency is estimated as a function of the transverse momenta of the simulated charged particles (top). The events are produced in full simulation including the underlaying event, but without superimposing the pile-up collisions.

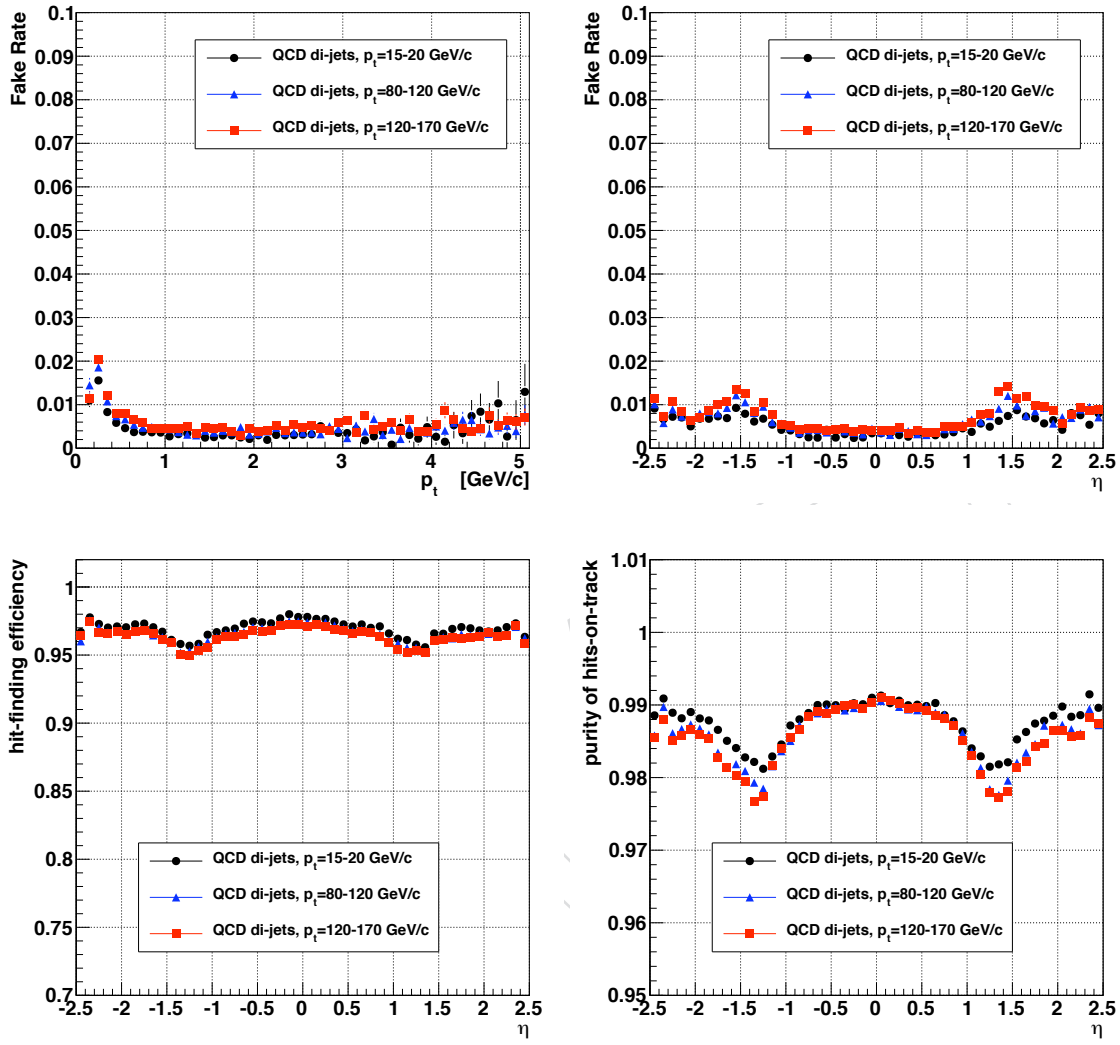

Figure 23: Tracking fake rate as a function of the transverse momentum of the reconstructed trajectory (top-left) and as a function of its pseudorapidity (top-right). Hit-finding efficiency (bottom-left) and hit-purity (bottom-right) as a function of the pseudorapidity of the simulated particle that is reconstructed. All the four sets of histograms are evaluated on simulated LHC collisions containing di-jets with transverse momenta within intervals of 15-20, 80-120 and 120-170 GeV/c. The events are produced in full simulation including the underlaying event, but without superimposing the pile-up collisions.

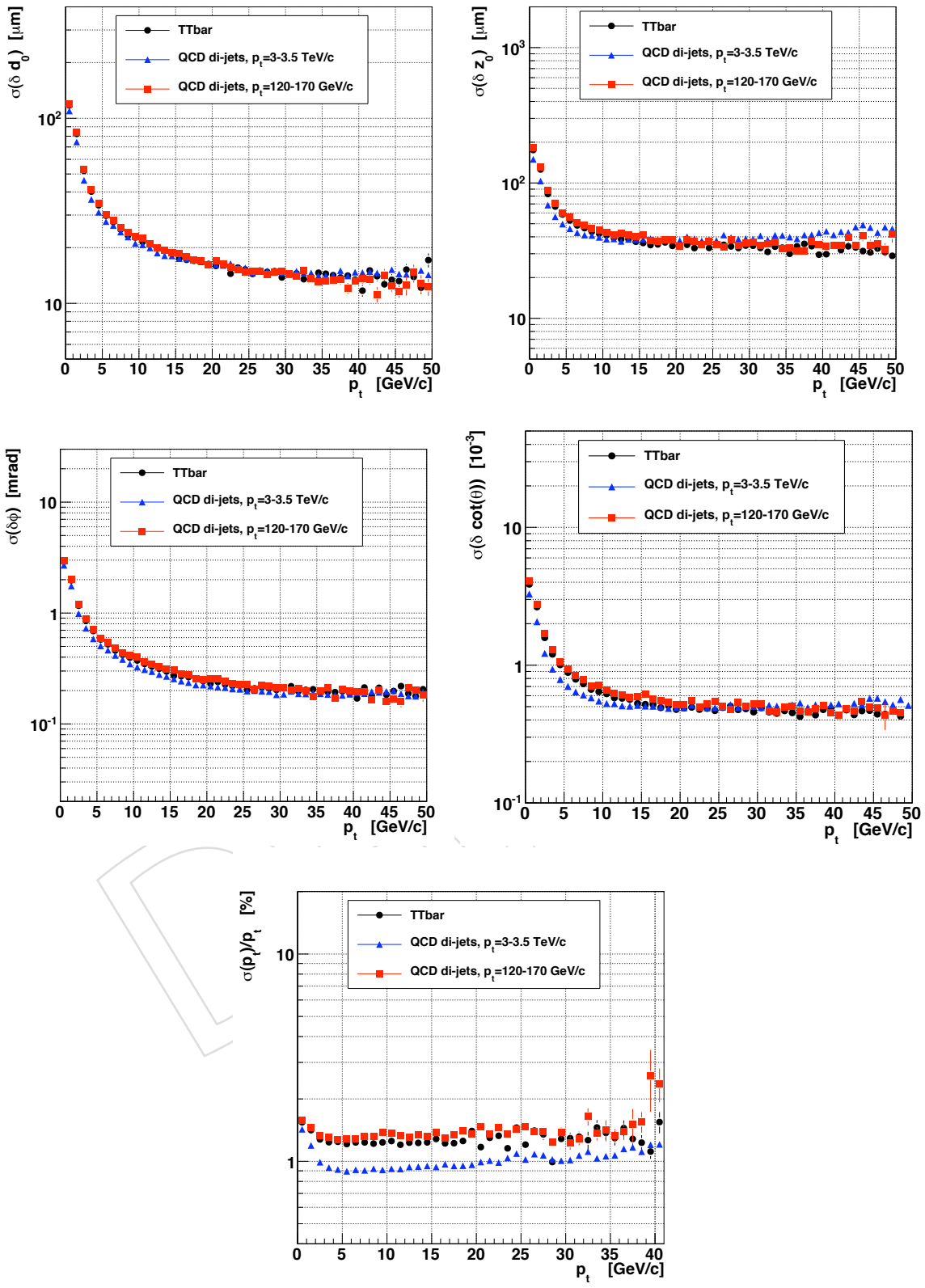

Figure 24: Resolution, as a function of the transverse momentum, of the five track parameters for charged particles produced within 3 cm from the beam-line and $\pm$ 20 cm around the center of CMS in simulated LHC collisions containing di-jets with transverse momenta within 120-170 and 3000-3500 GeV/c and multiple-jets from $t\bar{t}$ pairs. From top to bottom and left to right: transverse momentum, φ, cot ϑ, transverse and longitudinal impact parameter. For each bin in η, the full symbol corresponds to the sigma of a gaussian fit of the residuals distribution. The events are produced in full simulation including the underlaying event, but without superimposing the pile-up collisions.

The di-jets and $t\bar{t}$ events analyzed in this section contain charged particles that are spread over a wide range of energies and, therefore, these simulated samples allow the study of the dependency between the resolution of the track parameters and the particle transverse momentum. In Figure 24 the resolutions of the five trajectory parameters are reported as a function of the particle transverse momentum. Due to the increasing influence of multiple scattering, all the five plots show a significant degradation of the resolution for values of transverse momentum below 10-20 GeV/c. For values *above* the same range, the resolutions of the four parameters measuring d_0, d_z, φ and $\cot\vartheta$ reach a plateau since at that energies the resolutions are dominated by the precision with which the trajectory position is measured on the innermost sensitive layer of the Tracker. The precision in measuring the transverse momentum is maximum for transverse momenta around 5 GeV/c (Fig. 24, bottom) and then it degrades progressively for higher momenta, while the trajectories approximate closer and closer a straight line.

The transverse momentum resolution estimated on the 3-3.5 TeV/c di-jets sample is superior compared to the resolutions measured on the two other kind of events: the explanation is that the jets of the first sample, and the corresponding charged particles, are produced exclusively in the barrel region of the Tracker where the largest lever arm can be used to estimate the trajectory curvature.

4.4 Tracking performance on exotic events with displaced tracks

Iterations 3-5 of the 6-step iterative tracking scheme, described in Sect. 3.6, are dedicated to the reconstruction of displaced tracks. They improve the efficiency for V^0 reconstruction, which is very useful for B physics studies and should also improve particle flow energy resolution. They will also help with γ conversion reconstruction. And finally, they allow searches for exotic physics events, in which long-lived, massive neutral exotic particles decay in the Tracker volume to jets or leptons.

The ability to reconstruction displaced tracks is illustrated using using exotic events simulated with SUSYHIT + PYTHIA + CMSSW. Each event contained a pair of neutral, massive, exotic particles. Each had a mass of 165 GeV/c^2, a mean E_t of 400 GeV and transverse decay length that approximately followed an exponential distribution, with a mean of 24 cm. They each decayed to a pair of jets. The contribution to the tracking efficiency and fake rate from each of the six tracking iterations is shown in Fig. 25. respectively. This figure was made using the *high purity* track selection described in Sect. 3.6.

(If at least 75% of the hits on a reconstructed track are produced by the same particle, then that particle is considered to be successfully reconstructed. Otherwise, it is considered to be a fake track. The tracking efficiency was defined as the fraction of stable, charged particles with $P_t > 0.9$ GeV/c and $|\eta| < 2.4$ that resulted in a successfully reconstructed track. The tracking fake rate was defined as the fraction of reconstructed tracks that were fake.)

References

[1] **CMS** Collaboration, R. *et al.*. Adolphi, "The CMS experiment at the CERN LHC", *JINST.* **3:S08004** (2008).

[2] P. Azzurri and B. Mangano, "Optimal filtering of fake tracks", *CMS Internal Note* **2008/017** (2008).

[3] M. Pioppi, "Iterative Tracking", *CMS Internal Note* **2007/065** (2008).

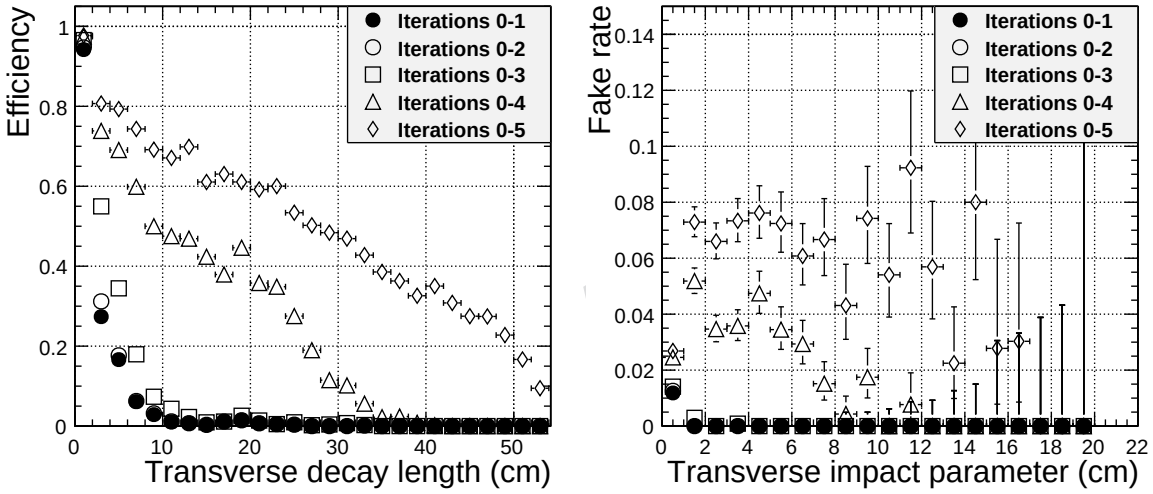

Figure 25: Contributions to the overall tracking performance from the 6 steps of the iterative track-finding. Tracking efficiency (*left plot*) is shown as a function of the transverse distance from the beam-axis of the particle production point. Track fake rate (*right plot*) is shown as function of the transverse impact parameter of the track.