CMS Physics Analysis Summary

Contact: cms-pog-conveners-tracking@cern.ch 2010/07/22

Tracking and Primary Vertex Results in First 7 TeV Collisions

The CMS Collaboration

Abstract

After a previous run at lower energies, the first 7 TeV proton-proton collisions produced by the LHC have been recorded by the CMS experiment in 2010. We use the silicon tracking detector and the 3.8 T superconducting solenoid to reconstruct the trajectories, momenta and vertexes of charged particles. We report on the performance of this track, vertex, and beam line reconstruction for the first time at these higher beam energies.

1 Introduction

Charged particle reconstruction is one of the central pieces of information in understanding the processes that take place during the 7 TeV collisions of protons in the LHC collider [1]. The CMS tracking detector [2] consists of an inner silicon pixel detector composed of three barrel layers and two forward/backward disks, and up to ten barrel silicon strip layers and three inner and nine outer forward/backward strip disks. The first measurement comes from the innermost barrel pixel detector located at an average radius from the center of the beam pipe of 4.4 cm, and the barrel strip tracker extends out to a radius of 1.1 m. The detector provides tracking coverage out to pseudorapidities of 2.5. The silicon tracker sits inside the 3.8 T superconducting solenoid in the heart of the CMS experiment [2] and is the primary detector used to measure the momentum and spatial origin of charged particles before they pass into the calorimeters and muon chambers. This information is then used as a main component in the reconstruction of hadrons, electrons, muons, taus and b-quarks. A fast, pixel-only version of the tracking and vertexing information is available at the high level trigger ("online") and will be described in detail in an upcoming note.

We analyze the data collected by CMS at a center of mass energy of 7 TeV in the early 2010 run of the LHC. In this note we report on the performance of the CMS experiment at reconstructing tracks, primary vertexes, the LHC's luminous region (beam line), multiple interactions of the proton bunches, and finally background events which come from beam-gas interactions.

1.1 Data and Simulated Samples and Event Selection

The results presented in this analysis summary have been obtained using the data corresponding to 1.1 nb^{-1} of integrated luminosity in the runs with colliding beams at the center-of-mass energy of 7 TeV, the solenoidal magnetic field at the nominal value of 3.8 T and the silicon pixel and the silicon strip tracker detectors enabled with the nominal high voltage bias applied to the sensors. Due to the relatively low LHC luminosity the CMS readout was triggered by the beam scintillator counter (BSC) trigger to collect minimum-bias collision events and by the beam pick-up timing detector to detect the passage of the beam bunches [3]. The analysis of the impact parameter resolution in Section 3 uses slightly more data (10.9 nb^{-1}) in order to enhance the sample of higher momentum tracks.

Most of the analyses presented in this summary reduce the background from non-collision events further and select useful events for tracking studies by requiring events to have: one primary vertex reconstructed with at least four degrees of freedom (see Section 2); the reconstructed position along the beam line within $\pm$15 cm of the nominal CMS detector center; the distance in the transverse plane from the nominal beam line within 2 cm; and a fraction of "high purity" tracks [3] larger than 25% if the number of reconstructed tracks is larger than 10.

The simulated events used are minimum-bias events (about 10 million) produced with the PYTHIA 6.4 [4] event generator, tune D6T [5] at a center-of-mass energy of 7 TeV. Due to the observed discrepancy in the track multiplicity between data and PYTHIA 6.4 tune D6T simulation, we use the minimum-bias events produced with the PYTHIA 8.1 event generator, tune 1 [6], at center-of-mass energy of 7 TeV for the analyses that are sensitive to the overall track multiplicity. Pythia 6.4 is used when not otherwise specified. The events produced with the event generators are then processed with a simulation of the CMS detector response based on GEANT 4 [7].

The alignment and calibration uncertainties and dead channel maps of the detector have been included in the simulation. The average beam line position and width have been tuned in the

simulation to match the real data.

2 Primary Vertex Reconstruction

In the primary vertex reconstruction [8], the measurements of the location and uncertainty of an interaction vertex are computed from a given set of reconstructed tracks. The prompt tracks originating from the primary interaction region are selected based on the transverse impact parameter significance with respect to the beam line (see section 4), number of strip and pixel hits, and the normalized track χ^2. To ensure high reconstruction efficiency in the minimum bias events, there is no requirement on the track transverse momentum. The selected tracks are then clustered based on their z coordinates at the point of closest approach to the beam line. Vertex candidates are formed by grouping tracks that are separated in z by less than a distance $z_{\text{sep}} = 1$ cm from their nearest neighbor. Candidates containing at least two tracks are then fit with an adaptive vertex fit [9] to compute the best estimate of vertex parameters such as position and covariance matrix, as well as the indicators of the success of the fit, such as the number of degrees freedom of the vertex and track weights of the tracks in the vertex.

In the adaptive vertex fit, each track in the vertex is assigned a track weight between 0 and 1 based on its compatibility with the common vertex. For a track consistent with the common vertex, its weight is close to 1. The number of degrees of freedom is defined as $n_{\text{dof}} = 2\sum_{i=1}^{nTracks} w_i - 3$, where w_i is the weight of the i^{th} track. It is thus strongly correlated to the number of tracks compatible with the primary interaction region. For this reason, the number of degrees of freedom of the vertex can be used to select real proton-proton interactions.

2.1 Primary Vertex Resolution

The primary vertex resolution depends strongly on the number of tracks used in fitting the vertex and the p_{T} of those tracks. In this section, we introduce a data-driven method to measure the resolution as a function of the number of tracks in the vertex, which is referred to as the "split method." The tracks used in the vertex in an event are split evenly into two different sets. During the splitting procedure, the tracks are ordered in descending order of p_{T} first and then grouped in pairs starting from the track with largest p_{T}. In each pair, tracks are randomly assigned to one or the other track set. This procedure ensures that the two split track sets have the same kinematic distributions on average. These two different track sets are then fit independently with the adaptive vertex fitter. The distributions of the difference in the fitted vertex positions for a given number of tracks are fit with a single Gaussian distribution to extract the resolution. The fit range is constrained to be within 1.5 times the root-mean-square (RMS) for the histograms with less tails and 1 times RMS for the histograms with large tails.

Figure 1 shows the measured primary vertex resolutions in x (a), y (b), and z (c) as a function of the number of tracks. Results are shown for both the data and the simulation (Pythia8 Tune 1), and a good agreement in the curves is seen. The resolutions in x and y are observed to be consistent. For the minimum bias events at 7 TeV, the resolutions in $x(y)$ and z are found to be close to 25 μm and 20 μm for the primary vertexes using more than 30 tracks.

The difference between the measured vertex positions, divided by the sum in quadrature of the uncertainties reported by the fit, is referred to as the "pull." The pulls are roughly flat and close to 0.9, indicating that the vertex fit error is slightly overestimated. This is because the input track fit errors are slightly overestimated.

To examine the effect of the sum of the transverse momenta of the tracks in the vertex, we can study the resolution versus the number of tracks in the vertex for different average transverse

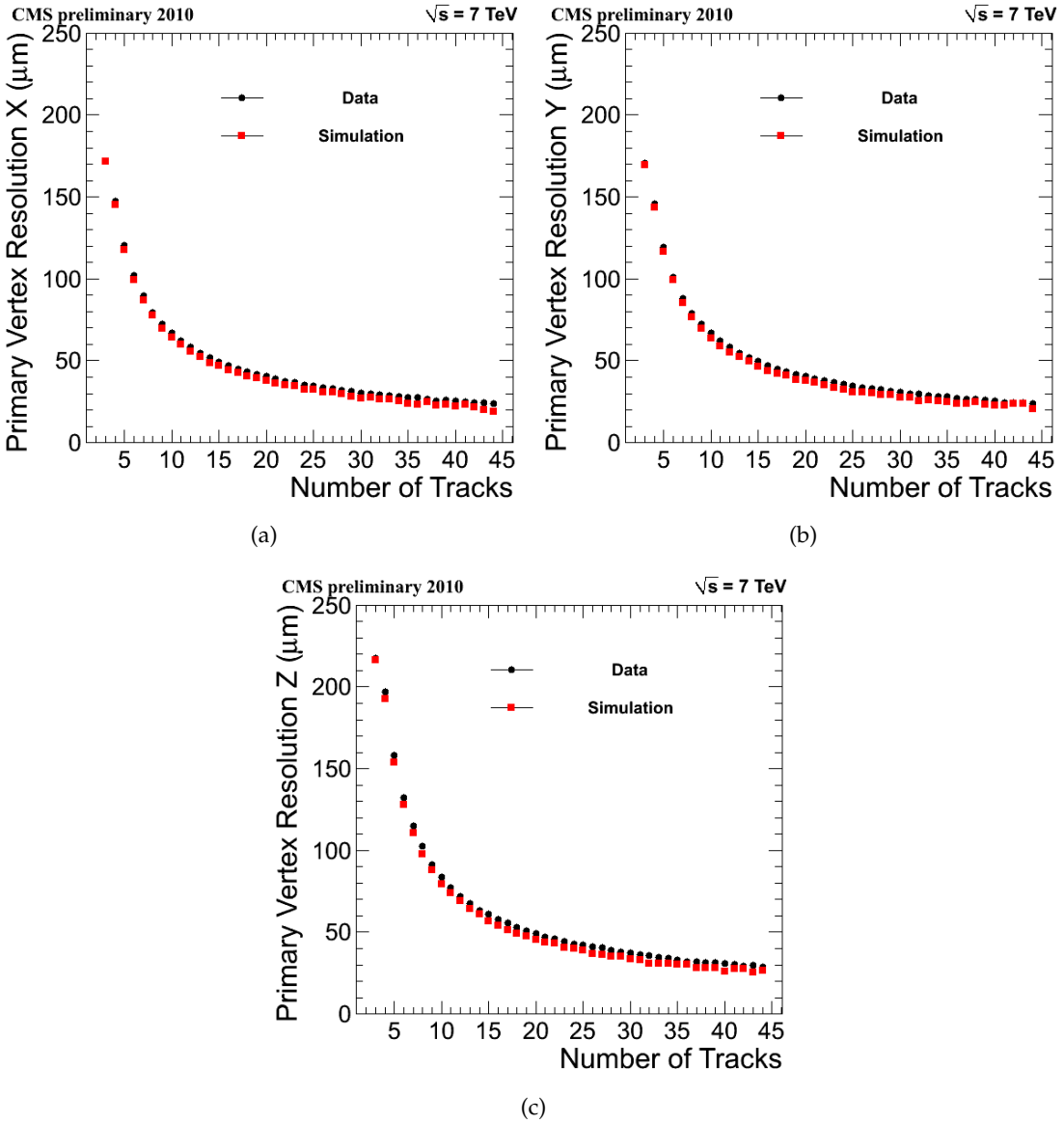


Figure 1: Primary vertex resolution in x (a), y (b), and z (c) as a function of the number of tracks used in the fitted vertex. The Pythia8 Tune 1 is used in the simulation.

momentum $\overline{p_{\mathrm{T}}}$ of tracks in the vertex. Figure 2 shows the vertex resolution in x (a), y (b), and z (c) as a function of the number of tracks for different average transverse momentum $\overline{p_{\mathrm{T}}}$ ranges. The resolution differs for the different average transverse momentum $\overline{p_{\mathrm{T}}}$ regions shown, the simulation is able to describe the data fairly well.

2.2 Primary Vertex Efficiency

Given a set of tracks clustered in z which forms a reconstructed vertex, we estimate the vertex reconstruction efficiency by asking how often its position is consistent with the true value. In this estimation, the tracking efficiency is not considered. Similar to the primary vertex resolution, the primary vertex efficiency also strongly depends on the number of tracks in the cluster. The split method described in previous section can be used to measure the primary vertex reconstruction efficiency as a function of the number of tracks in the cluster. To decou-

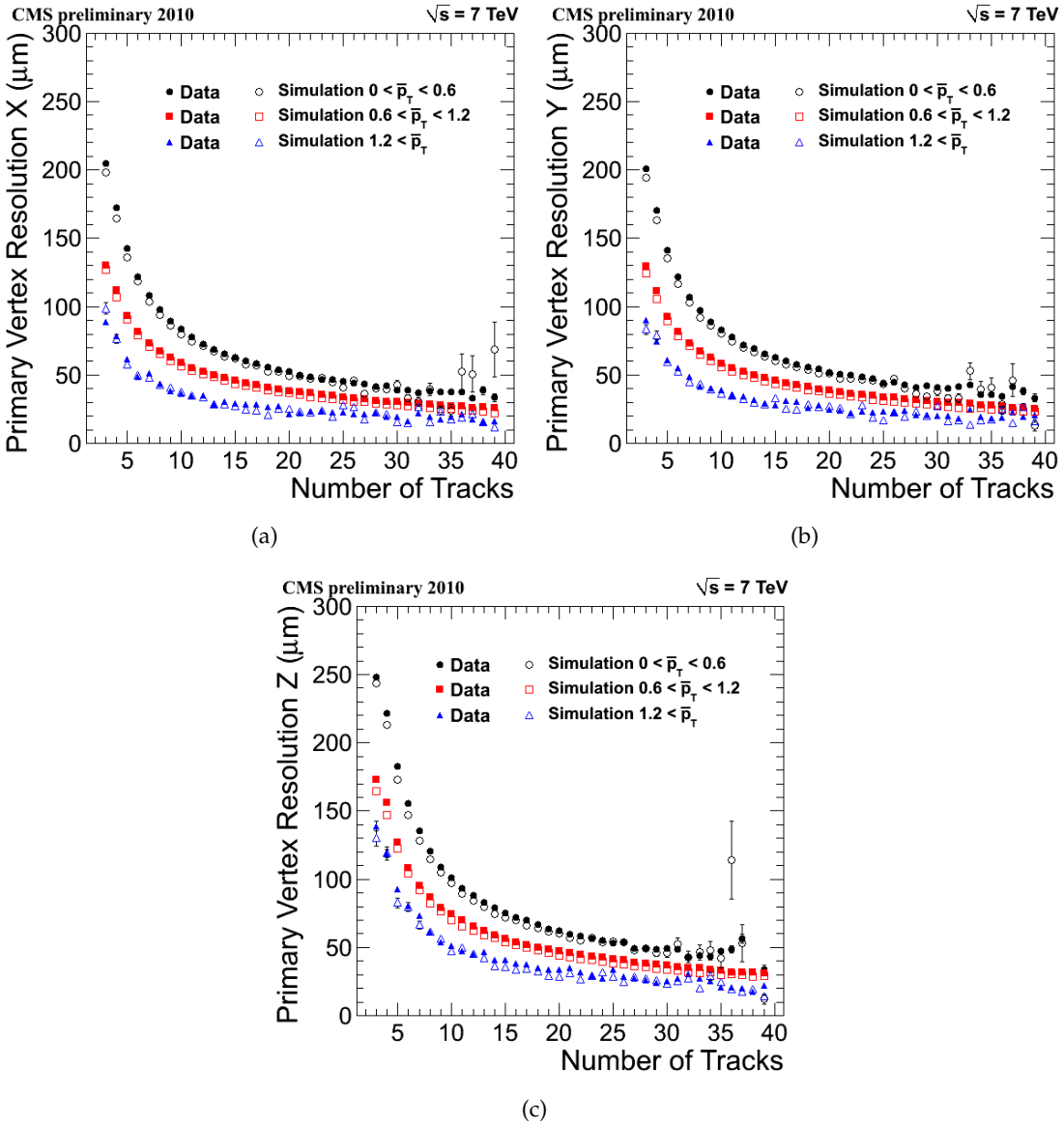


Figure 2: Primary vertex resolution in x (a), y (b), and z (c) as a function of the number of tracks for different average transverse momentum $\overline{p_{\mathrm{T}}}$. The Pythia8 Tune 1 is used in the simulation.

ple the primary vertex efficiency from the fake rate of reconstructed tracks, we suppress fakes by requiring all tracks to have a transverse momenta of 0.5 GeV.

In the split method, the tracks used in the primary vertex in an event are ordered first in descending order of p_{T} and then split into two different sets, with 2/3 (1/3) of the tracks assigned to the tag (probe) track sets. The asymmetric splitting is used to increase the number of vertexes with low numbers of tracks. The tag and probe track sets are then fit independently with the adaptive vertex fitter to extract the primary vertex reconstruction efficiency.

The efficiency is calculated by how often the probe vertex is matched to the original vertex given that the tag vertex is reconstructed and matched to the original vertex. A tag or probe vertex is considered to be matched to the original vertex if the tag or probe vertex position in z is within 5σ from the original vertex. The σ is chosen to be the larger value of the vertex fit

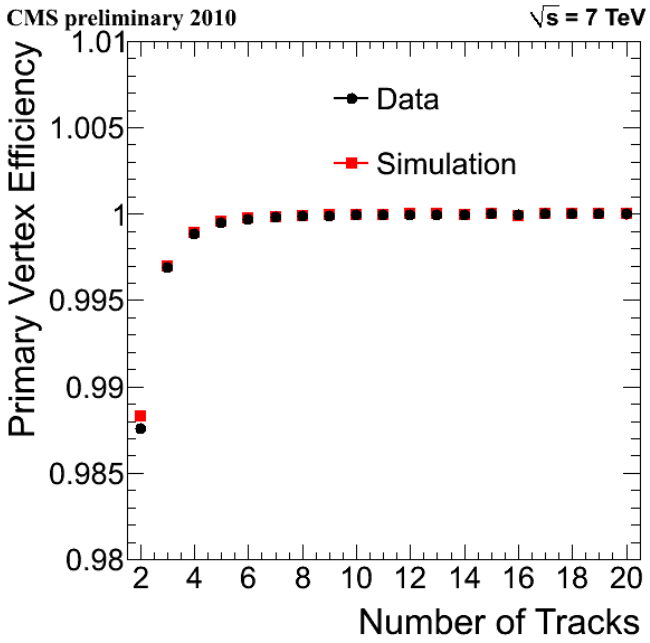


Figure 3: Primary vertex efficiency as a function of the number of tracks in a cluster.

error between the tag or probe and the original vertex.

Figure 3 shows the measured primary vertex efficiency as a function of the number of tracks that are clustered in z. The results obtained using the split method described above are applied to both data and simulation and good agreement between the two is observed. The primary vertex efficiency is estimated to be close to 100% if there are more than two tracks with transverse momenta greater than 0.5 GeV in the vertex.

2.3 Multiple Interactions

Although the instantaneous luminosity in the early collision data is far below the design luminosity of the LHC, the luminosity per bunch crossing was already high enough to produce multiple collisions in a few percent of the events. The possibility of multiple primary interactions in the same bunch crossing is taken into account by a simple clustering step as described in the previous section.

Vertexes separated by $\mathcal{O}\left(z_{\text{sep}}\right)$ or less are not separated by this procedure and are merged into a single reconstructed vertex. Depending on the track content and separation of the vertexes, the result of the subsequent adaptive vertex fit will often be very close to the vertex with higher multiplicity.

Tracks separated by more than z_{sep} from the true collision point are likely to be split off from the vertex by this procedure. For z_{sep} larger than the typical z–resolution this has little impact on the reconstructed vertex position because such tracks would either be down-weighted by the adaptive vertex fit or have very poor resolution. For very soft interactions with a small number of mostly low resolution tracks, splitting can lead to the complete loss of the real vertex. Multiple split-off tracks on the other hand may lead to an additional reconstructed vertex near the main vertex.

The choice of the clustering distance represents a trade-off between merging of nearby vertexes for large z_{sep} and false vertexes from vertex splitting for small z_{sep}. Given the low probability

of multiple collisions in the early data, a relatively large clustering distance of $z_{\text{sep}} = 1$ cm was chosen in order to safely group all tracks of the same interaction within the same cluster. Results for a smaller clustering distance $z_{\text{sep}} = 2$ mm are shown for comparison.

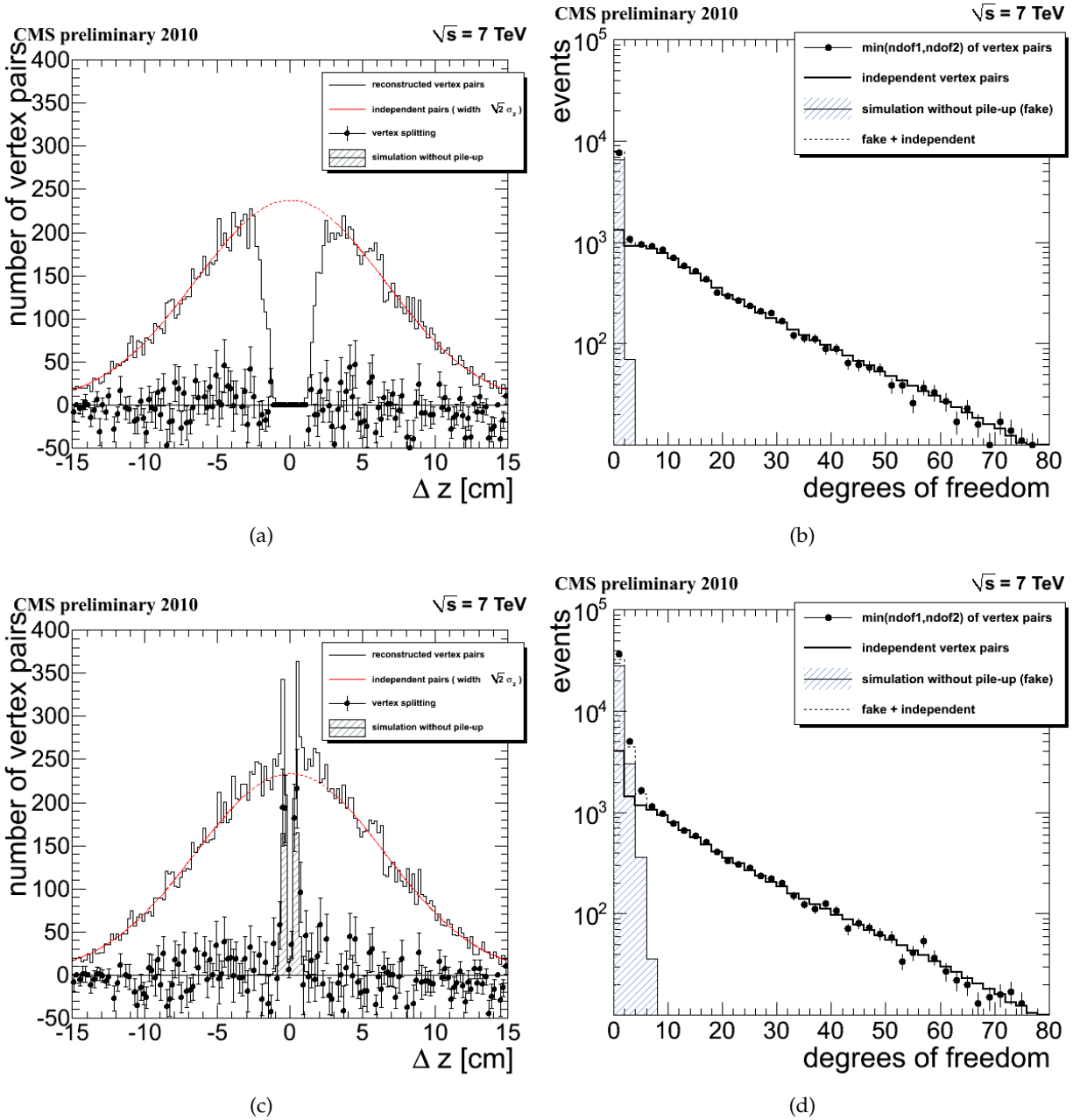
On an event-by-event basis, it is in general not possible to determine whether a given vertex corresponds to a single vertex, or is the result of merging two closely spaced interactions. Likewise, the result of splitting cannot be distinguished with certainty from two genuine collisions event-by-event. However, the rate of merging and splitting in an event sample can be determined from the distribution of the vertex pair position.

The z coordinate of collisions has a distribution that is Gaussian over several orders of magnitude. The width of the luminous region, $\sigma_z \sim 3$–5 cm, can vary from fill to fill and also slowly during fills. When two or more interactions occur in the same bunch crossing, their true positions are independent of each other. The distance $\Delta z = z_1 - z_2$ between any pair of two independent vertexes is distributed with a width of $\sqrt{2}\sigma_z$ while their average position $\bar{z} = \frac{1}{2}(z_1 + z_2)$ has a width of $\sigma_z/\sqrt{2}$. For well separated pairs this is also expected for the reconstructed z positions while near $|z_1 - z_2| \lesssim z_{\text{sep}}$ vertexes are merged and the number of pairs falls below this expectation.

Splitting, on the other hand, leads predominantly to vertex pairs with $z_1 - z_2$ close to z_{sep} and hence $|z_1 - z_2| \ll \sigma_z$. The pair average position is close to the true collision point and therefore distributed with width σ_z. Neglecting other types of false vertexes, the contribution of vertex splitting to a set of vertex pairs is found by fitting the $\bar{z}$ distribution with a sum of two Gaussian with widths of σ_z and $\sigma_z/\sqrt{2}$. Figures 4(a) and 4(c) show the number of events corresponding to the amplitude of the Gaussian with width σ_z in bins of Δz for a sample of events collected with a minimum bias trigger in single CMS run (135149). A significant vertex splitting component is only visible for $\Delta z < 1$ cm for $z_{\text{sep}} = 0.2$ cm. In total it amounts to a probability of 0.3% of splitting off a vertex with $n_{\text{dof}} > 4$. No significant vertex splitting is found for $z_{\text{sep}} = 1$ cm. The distribution of vertex pairs reconstructed in a sample of simulated events without multiple primary interactions describes well the shape of the observed vertex splitting. The simulation underestimates the rate of vertex splitting and a common scale factor 1.5 is applied to all simulated distributions shown in Figs. 4(a) and 4(c).

For large $|\Delta z|$ the distribution of vertex pairs is a given by a Gaussian with width $\sqrt{2}\sigma_z$ as expected for independent pairs. The deficit in the central region – after subtracting split vertexes – is approximately the number of multiple vertexes that were not reconstructed but merged with another vertex. This merging is found to be 23% for the coarse clustering with 1 cm and 7% for the finer clustering with 2 mm. The width of the luminous region in this case was $\sigma_z = 4.6$ cm.

The distribution of the smaller n_{dof} of any two vertexes in the same event from an unbiased event sample is shown in Fig. 4. It is compared to the expected distribution, $p(n)$, for collisions derived from the observed n_{dof} distribution of single, well isolated vertexes in the same unbiased sample, $p_0(n)$, with $p(n) = 2p_0(n) \times \sum_{n'>n} p_0(n')$. The excess from split vertexes at small n_{dof} values is well described by the simulation.



(a) (b)

(c) (d)

Figure 4: Properties of vertex pairs reconstructed with a clustering parameter $z_{\text{sep}} = 1$ cm (a), (b), and 2 mm (c), (d). (a) and (c) show the $\Delta z = z_1 - z_2$ distribution pairs of vertexes fulfilling $n_{\text{dof}} > 4$. The data points with error bars show the fitted number of correlated pairs from vertex splitting. The distribution of $\min(n_{\text{dof1}}, n_{\text{dof2}})$ of the vertex pairs is shown in (b) and (d). For comparison, vertex pairs from a simulation of events without multiple primary interactions is superimposed (hatched histograms).

3 Track Impact Parameters

The precision in measuring the track Impact Parameter (IP) depends strongly on the track direction and momentum. Using a data-driven technique, the resolutions of both the transverse and the longitudinal impact parameters are estimated for several values of the track η, ϕ and p_T. These resolutions are measured from data and are compared with predictions from simulations.

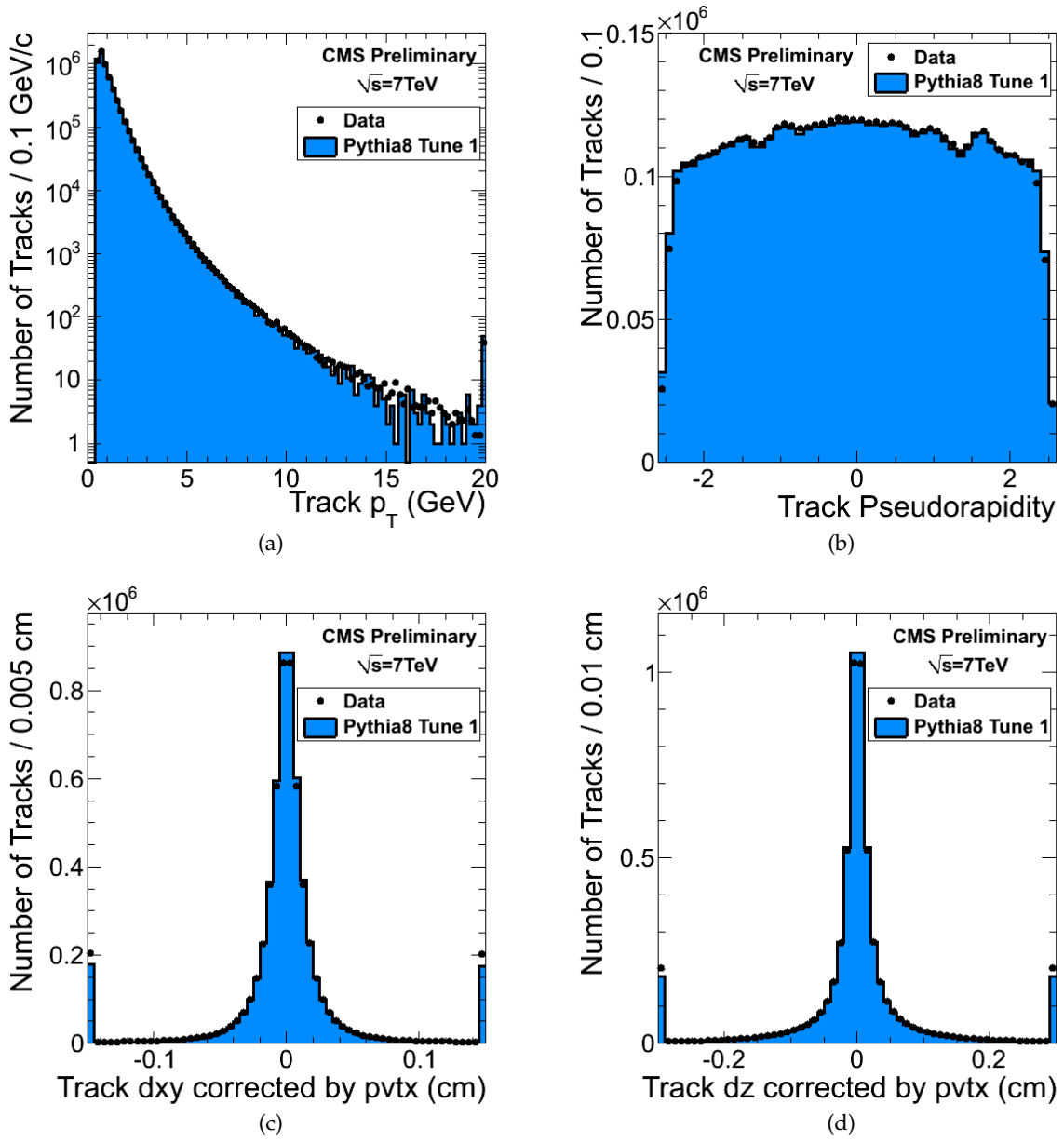


Figure 5: Comparison of data (points) and simulation (blue histogram) for the following measured track parameters: (a) p_T; (b) η; (c) transverse impact parameter; (d) longitudinal impact parameter.

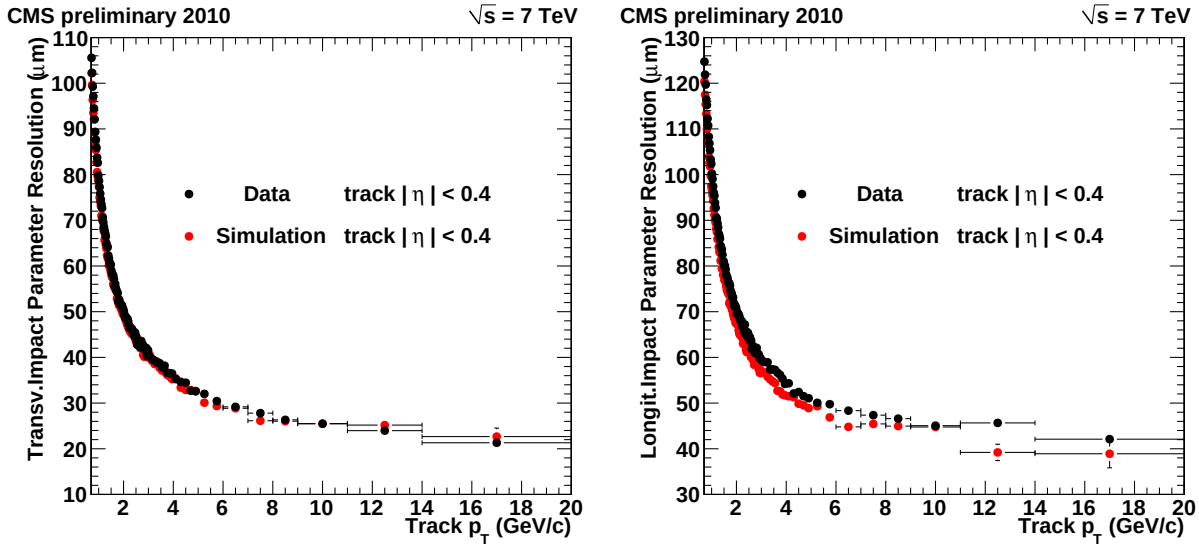
3.1 Basic Track Parameter Distributions

Before describing the algorithm used to estimate impact parameter resolutions, basic track parameter distributions (detailed definitions of the CMS track parameter conventions and de-

scriptions of track selection can be found in Reference [8]) are illustrated. Figure 5 shows a comparison of data and simulation (Pythia8 Tune 1) for the following distributions: (a) transverse momentum, p_T; (b) pseudorapidity, η; (c) transverse impact parameter, d_{xy}, with respect to the primary vertex; and (d) longitudinal impact parameter, d_z, with respect to the primary vertex.

3.2 Track Impact Parameter Resolution

The analysis described in this section is based on the 7 TeV data collected by CMS up to the 27th of May 2010 and corresponding to 10.9 nb^{-1}. In addition to the general selection detailed in Section 1.1, the events used for the measurement of the IP resolutions are required also to pass the uncorrected 6 GeV jet trigger. The usage of a common trigger ensures that the tracks used in both data and simulation are comparable in terms of track multiplicity and distribution of particle kinematic variables. The measurement of the impact parameter resolution starts from the selection of high quality tracks that have a high probability of having been produced promptly in the pp collision: a track must have its p_T greater than 0.3 GeV/c and valid measurements on at least 7 consecutive layers of the tracker, including a measurement on the innermost pixel layers (either the barrel or one of the endcap disks). Simulation studies predict that this simple selection is expected to reduce the fraction of fake tracks to the per mil level. For transverse momenta smaller than 4 GeV/c (20 GeV/c), the fraction of non-prompt tracks that are selected is less than 2% (10%) of the total.



(a) (b)

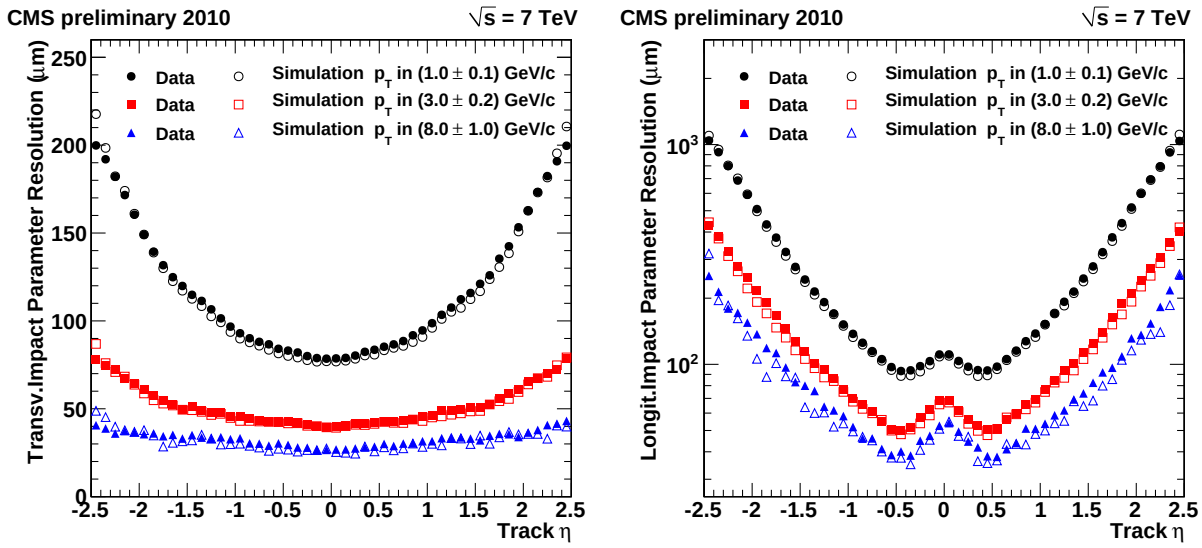
Figure 6: Measured resolution of the track transverse (a) and longitudinal (b) impact parameter as a function of the track p_T. Only central tracks with $|\eta| < 0.4$ are considered. Black and red symbols correspond to results from data and simulation, respectively.

For each track passing these criteria, the unbiased position of the collision point is determined using all and only the other tracks in the event with the vertex fitter described in Section 2. The uncertainty on the position is estimated from the vertex fit and it is used to filter the newly reconstructed vertexes. If the errors on the x and y (z) coordinates of the vertex position are within 15–37 μm (20–36 μm), a vertex-track pair is created and used in the next step of the analysis. These cuts on the position error have been chosen as a trade-off between selecting vertexes that are very precisely reconstructed and having enough vertexes passing the selec-

tion. Given the trigger, track and vertex selections, 124M and 39M pairs have been chosen for the analysis in the data and simulation, respectively. The transverse and longitudinal impact parameters of the track are evaluated with respect to the position of the matched vertex. Each calculated value is used to fill histograms corresponding to different bins of the track η, ϕ and p_{T}. In each of these, the distribution of the measured IP is centered around zero and has a spread that depends on three terms: the uncertainty on the track impact parameter itself due to the error on the track innermost measurement and the extrapolation of the track parameters to the primary vertex through the wall of the beam-pipe; the uncertainty on the primary vertex position; the fraction of selected tracks corresponding to genuinely displaced particles from decay of heavy-flavor and K_s or from conversions.

The final IP distribution is the convolution of three probability density functions corresponding to the previous three terms. Since the fraction of selected tracks which correspond to non-prompt particles is relatively small, this term is ignored and the function used to fit the data is simplified to the convolution function of only two Gaussians. The width of the first Gaussian, related to the vertex position uncertainty, is estimated from simulation. The width of the second is what we define as the impact parameter *resolution* and it is the only remaining unknown that is extracted from the fit.

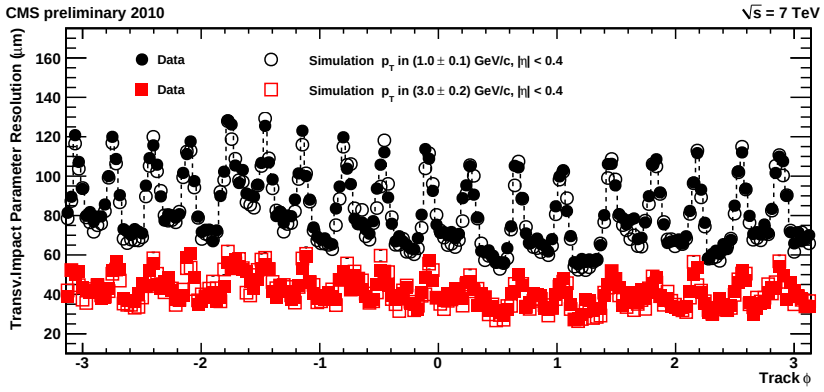
Figure 6 shows the transverse and longitudinal impact parameter resolution as a function of the track p_{T}. For higher track momenta, charged particles are less deflected by multiple scattering while traversing the material of the beam pipe and, as expected, this translates into a more precise impact parameter measurement. Neglecting the contribution of non-prompt particles does not compromise the measurement of the IP resolution for transverse momenta below 4 GeV/c, but it introduces a small bias in the measurement for higher p_{T} values. According to simulation, the resolutions estimated with the described data-driven technique are overestimated by about 5-10% (20%) for transverse momenta around 8 GeV/c (20 GeV/c). This bias is expected to affect the results shown in Fig. 6 for both the data and simulation.



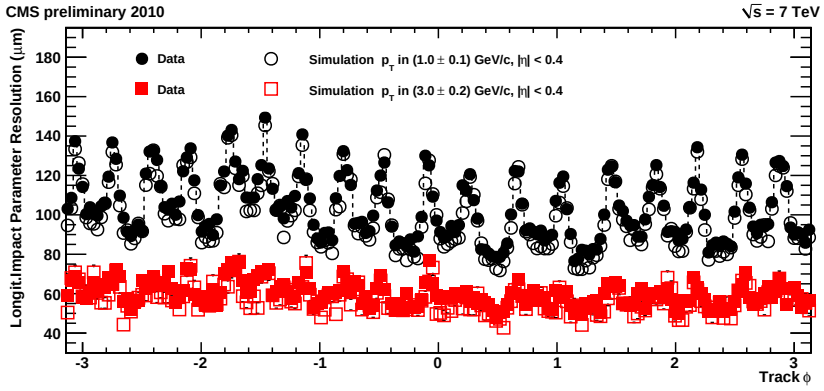
(a) (b)

Figure 7: Measured resolution of the track transverse (a) and longitudinal (b) impact parameter as a function of the track η for transverse momenta in 1.0 ± 0.1 GeV/c (circles), in 3.0 ± 0.2 GeV/c (squares) and in 8.0 ± 1.0 GeV/c (triangles). Filled and open symbols correspond to results from data and simulation, respectively.

Figures 7(a) and 8(a) show the transverse impact parameter resolution as a function of the track η, ϕ for transverse momenta in the ranges $1.0 \pm 0.1\,\text{GeV}/c$, $3.0 \pm 0.2\,\text{GeV}/c$ and $8.0 \pm 1.0\,\text{GeV}/c$. Both the longitudinal and transverse impact parameter resolution gets worse for increasing values of $|\eta|$ primarily because of the larger amount of material traversed by such particles. The plot of the resolution as a function of ϕ shows 18 peaks which correspond to the 18 cooling pipes of the innermost layer of pixel sensors. The plot has also a $\sin(\phi)$ modulation which is due to the displacement of the beams in the transverse plane with respect to the center of the innermost cylindrical layer of the tracking system.



(a)



(b)

Figure 8: Measured resolution of the track transverse (a) and longitudinal (b) impact parameter as a function of the track ϕ for transverse momenta in $1.0 \pm 0.1\,\text{GeV}/c$ (circles) and in $3.0 \pm 0.2\,\text{GeV}/c$ (squares). Fillled and open symbols correspond to results from data and simulation, respectively.

Similarly, Fig. 7(b) and 8(b) show the longitudinal impact parameter resolution as a function of the track η and ϕ for the same selection on the track transverse momentum. Compared to the transverse impact parameter, the plot for the resolution of the longitudinal impact parameter increases much faster as a function of $|\eta|$ because of the geometrical relation between the track's polar angle and the impact parameter in the longitudinal plane. Also, the longitudinal impact parameter resolution has a special feature at $|\eta| \approx 0.5$. Particles traversing a pixel module at such an angle deposit their ionization charge on more than one single pixel cell, allowing the improvement of the hit position determination by using the charge barycenter.

Overall, except for $\sim 10\%$ discrepancies at values of high track momenta and pseudorapidity, the agreement between the measured resolutions and the predictions from simulation is within

2-3%.

4 Reconstruction of the LHC Beam Line

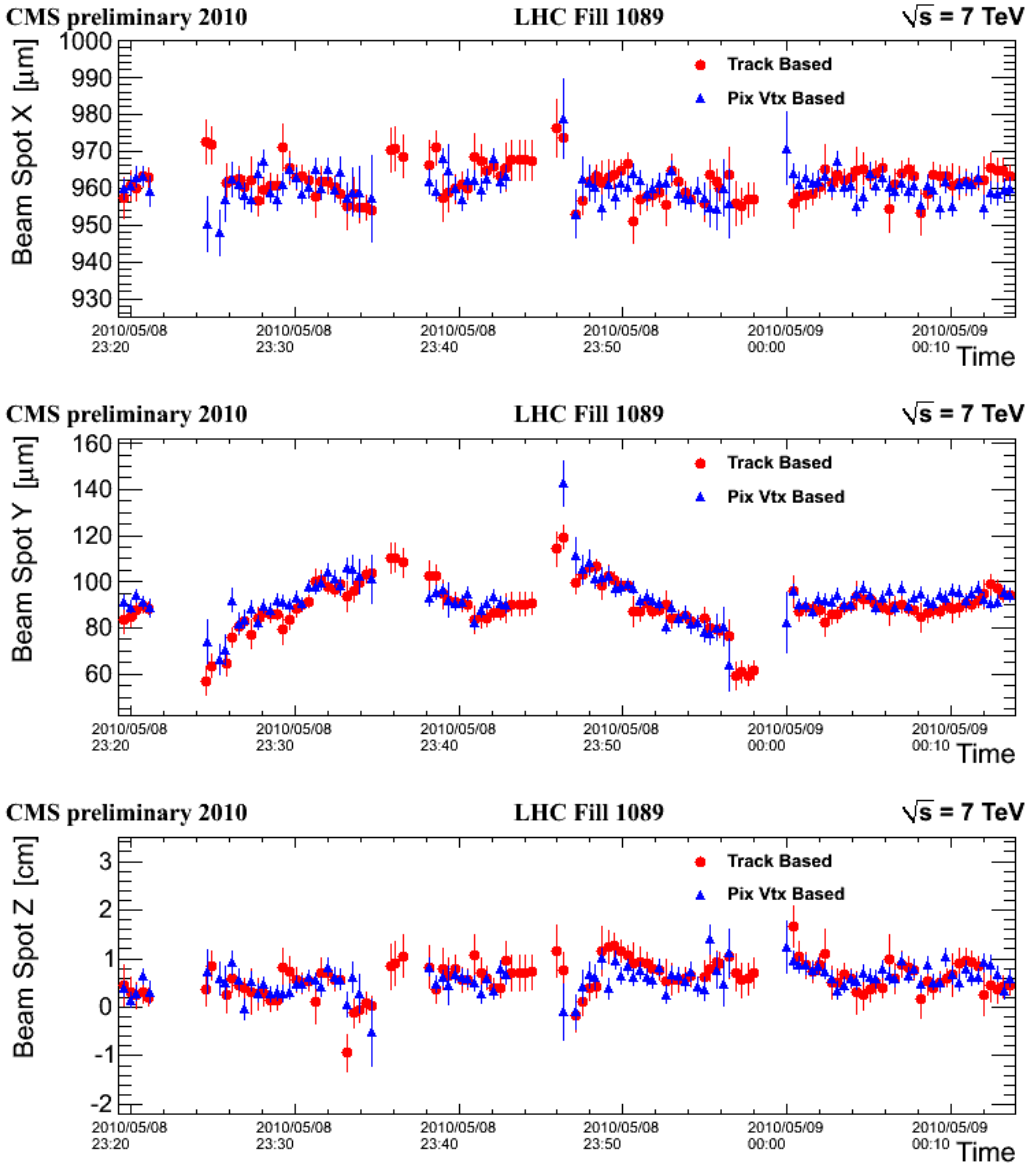


Figure 9: Fitted x_0 (top), y_0 (middle) and z_0 (bottom) positions of the beam line as a function of time during an LHC fill where a luminosity scan was performed.

The beam line represents the three-dimensional profile of the luminous region where the LHC beams collide at CMS. The beam line is determined in an average over many events, in contrast to the event-by-event primary vertex which gives the precise position of a single collision. A good measurement of the position and slope of the beam line is an important component of the event reconstruction. The beam line position can be used, especially in the high level trigger, as an estimate of the primary interaction point prior to the reconstruction of the primary vertex and even as the primary interaction point in low multiplicity data.

4.1 Determination of the Beam Line Position

The beam line position refers to the center of the luminous region. This position can be determined in two ways. The first is through the reconstruction of primary vertexes (see Section 2). The reconstructed vertexes map out the collisions and thus the shape of the beam line. The mean position in x, y, and z can be determined from a likelihood fit to the 3D distribution of vertexes.

The second method exploits a correlation between the transverse impact parameter (d_{xy}) and the azimuthal angle of tracks (ϕ_0) that exists when the beam line is displaced from the expected position. To first order the d_{xy} for tracks coming from the primary vertex can be parameterized by

$$d_{xy}(\phi_0, z) = x_0 \cdot \sin\phi_0 + \frac{dx}{dz} \cdot \sin\phi_0 \cdot z - y_0 \cdot \cos\phi_0 - \frac{dy}{dz} \cdot \cos\phi_0 \cdot z, \quad (1)$$

where x_0 and y_0 are the position of the beam at $z = 0$, and $\frac{dx}{dz}$ and $\frac{dy}{dz}$ are the x and y slopes of the beam. The beam line fit [10] uses an iterative χ^2 fit to exploit this correlation between d_{xy} and ϕ_0. With a sample of 1000 tracks, the position can be determined with a statistical precision of ~ 5 μm.

Two independent programs are used to determine the beam line position in quasi-real-time during collisions, exploiting the two different algorithms. In one program tracks from the standard reconstruction are used in the $d - \phi$ fit. In the other program a fast reconstruction is done using a special reconstruction that uses only hits in the pixel detector. Pixel hits from three different layers are selected and fit using a fast pattern recognition and helix fitter [11]. These pixel-only tracks are then passed to the standard vertexing algorithm to determine the position of the primary vertex. The likelihood fit is applied to the distribution of these primary vertexes to extract the beam line position.

In both cases a fit is attempted after every luminosity section (LS) [1]. A minimum of 150 tracks (35 pixel vertexes) is required, while typically more than 500 tracks ($300 - 400$ pixel vertexes) are reconstructed. Tracks and vertexes are accumulated for up to five LS before the fit is reset. Results of the fit are fed back to the LHC to monitor beam parameters. Figure 9 shows the fitted positions as a function of time during one part of one fill in which LHC performed a luminosity scan. While the beam position is normally stable, during this scan the position in y was adjusted by LHC operators during the fill. No points are plotted for LS where an insufficient number of tracks or vertexes were reconstructed to fit the beam line. The plots show that the two methods give consistent results and are able to track well the movements of the beam.

4.2 Determination of the Beam Line Width

The size of the luminous region, described by the beam width in the transverse plane (σ_x and σ_y) and the length in z (σ_z), is also determined with two methods. The first is with the primary vertex distribution, where the widths are returned by the likelihood fit described above. Figure 10 shows the beam widths as a function of time for primary vertexes reconstructed with pixel-only tracks and primary vertexes reconstructed from tracks with both pixel and strip hits. Note that the size of the beam is growing during the fill, due to the growth of the emittance. A second method for measuring the beam width is to use the event-by-event correlations between the transverse impact parameters of two tracks from the primary event vertex.

The displacement of the interaction point within the interaction region profile introduces a common displacement for the trajectories of all particles from the primary interaction. This

[1] Events recorded by CMS are grouped in blocks of time, called luminosity sections, with a length of 23 seconds.

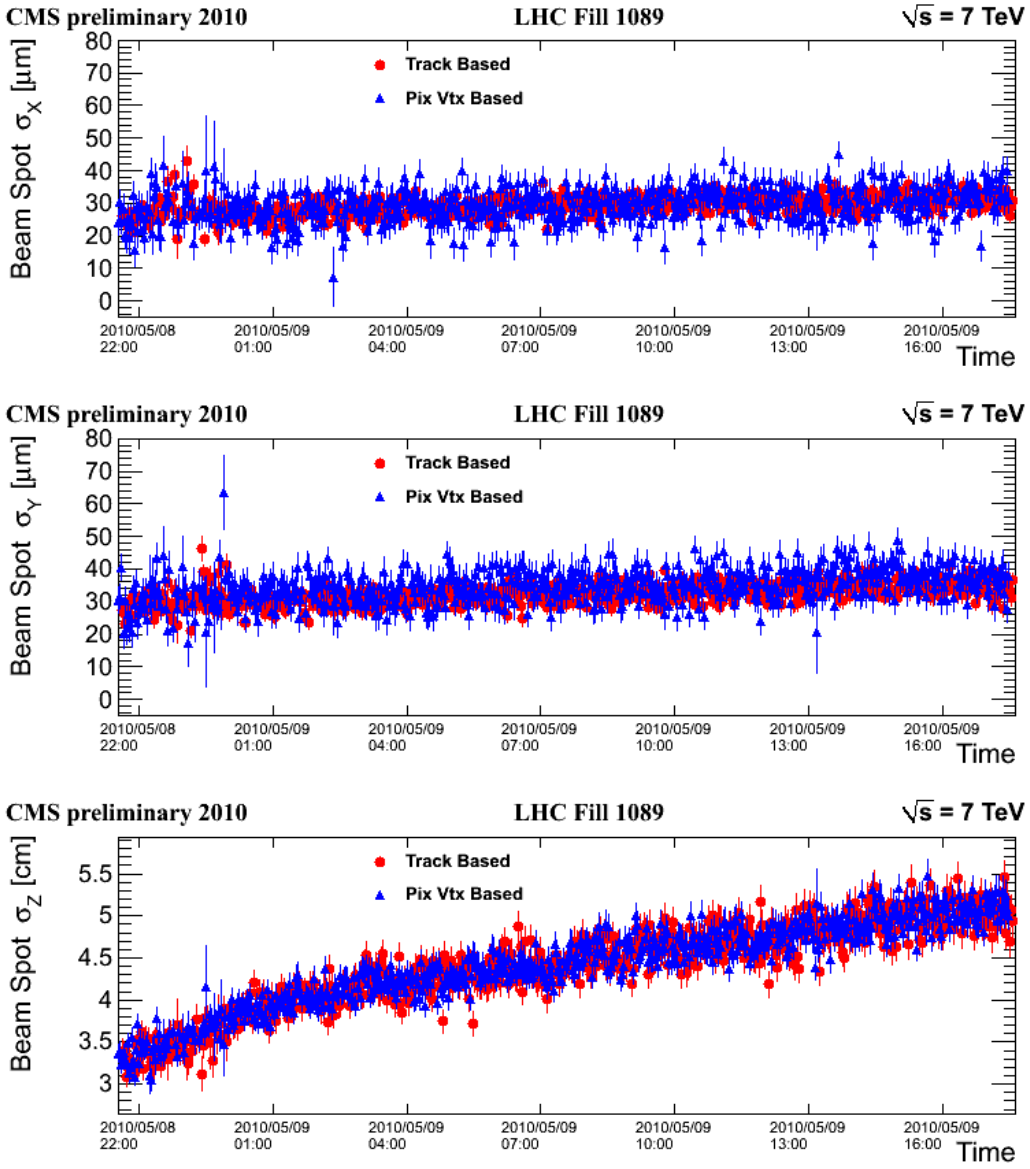
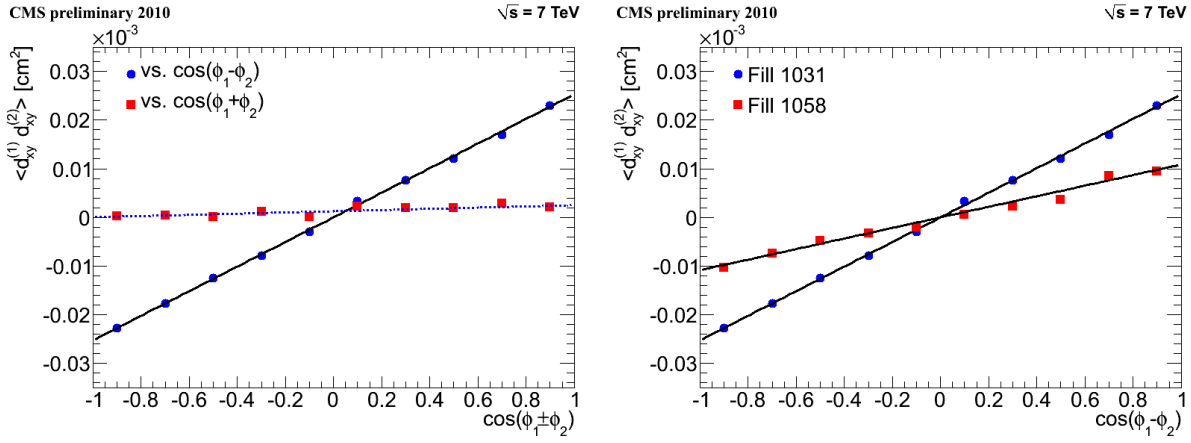


Figure 10: Fitted beam widths σ_x (top), σ_y (middle) and length σ_z (bottom) of the beam line as a function of time during one LHC fill.

shift of the trajectories results in a correlation between the transverse impact parameters of tracks with respect to the nominal beam position. The strength of the correlation reflects the transverse size of the beam. For a non-tilted beam ellipse, the correlation between the transverse impact parameters of two tracks from the primary vertex of the same interaction (labeled (1) and (2)) can be expressed by the expectation value

$$\left\langle d_{xy}^{(1)}\, d_{xy}^{(2)} \right\rangle = \frac{\sigma_x^2 + \sigma_y^2}{2} \cos(\phi_1 - \phi_2) + \frac{\sigma_y^2 - \sigma_x^2}{2} \cos(\phi_1 + \phi_2) \tag{2}$$

where ϕ_1 and ϕ_2 are the azimuthal angles of the tracks measured at the point of closest approach to the beam. A particular feature of this correlation is that its size is independent of primary vertex and impact parameter resolutions, and thus does not require corrections to remove the contributions of the resolution. Assuming non-correlation of $\phi_1 - \phi_2$ and $\phi_1 + \phi_2$, the



(a) (b)

Figure 11: (a) $\left\langle d_{xy}^{(1)}\, d_{xy}^{(2)} \right\rangle$ vs. $\cos(\phi_1 - \phi_2)$ in filled circles, with fit results shown by the solid line, and vs. $\cos(\phi_1 + \phi_2)$ in filled squares, with fit results shown by the dashed line, for a single LHC fill. (b) $\left\langle d_{xy}^{(1)}\, d_{xy}^{(2)} \right\rangle$ vs. $\cos(\phi_1 - \phi_2)$ for two LHC fills with different beam widths.

coefficients in Equation 2 can be obtained as the slopes of straight lines fitted to the respective dependence of $\left\langle d_{xy}^{(1)}\, d_{xy}^{(2)} \right\rangle$.

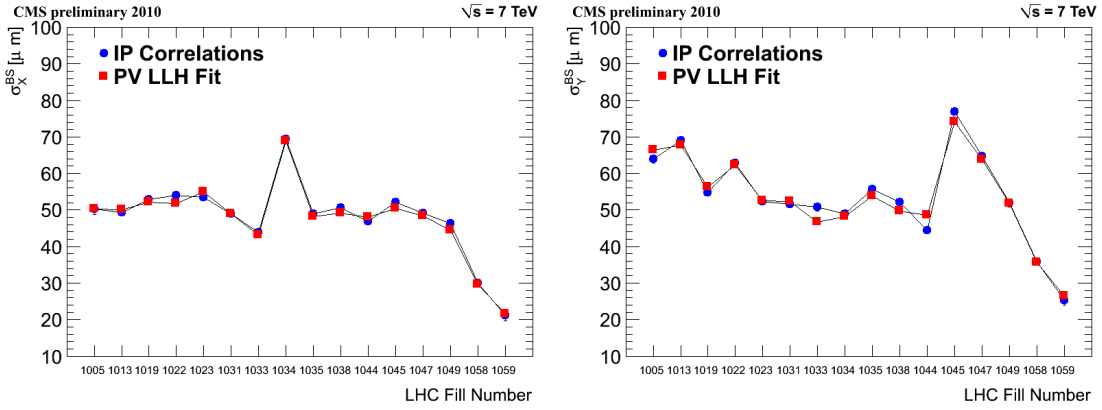
Figure 11(a) demonstrates the extraction of the beam line widths σ_x and σ_y using this method. The x-axis shows $\cos(\phi_1 - \phi_2)$ for the filled circles, fitted by the solid line, and $\cos(\phi_1 + \phi_2)$ for the filled squares fitted by the dashed line, respectively. In both cases the expected linear dependence is very clearly established. The slope of the solid line is indicative of the squared mean of σ_x and σ_y. The positive slope of the dashed line indicates that σ_y is slightly larger than σ_x.

Figure 11(b) shows a comparison of the impact parameter correlation as function of $\cos(\phi_1 - \phi_2)$ for two fills with significantly different beam widths. The squared average beam width is directly visible as the fitted function value at $\cos(\phi_1 - \phi_2) = 1$.

Figures 12(a) and 12(b) show a comparison of the transverse width from the likelihood fit to the primary vertex distributions, along with the results obtained using the two-track impact parameter correlations. In this case both methods are using the full offline track reconstruction. The fit to the distribution of primary vertexes includes the errors on the reconstructed primary vertexes. These errors are scaled based on the observed pulls of the primary vertexes. The uncertainty on the pull introduces a systematic uncertainty on the width of a few μm. The systematic uncertainty from the two-track impact parameters correlation method is estimated to be $\sim 2\mu$m based on studies of simulated samples comparing the fitted widths to the input values.

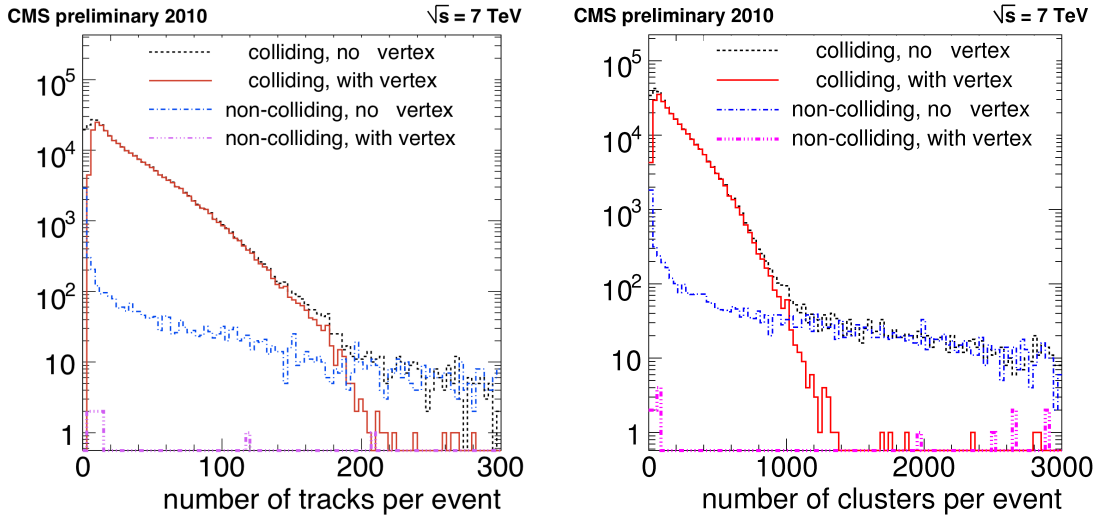
5 Beam Background Rejection

During all collision runs, irrespective of the energy of the beams, events with an unexpectedly large number of tracks and pixel occupancy were observed. Such events are characterized by a large number of clusters in the pixel detector. The pixel clusters recorded in the barrel detector tend to be very long and parallel the beam line direction. The cluster shapes are, in general,



(a) (b)

Figure 12: Fitted transverse beam line widths σ_x (a) and σ_y (b) by LHC fill, comparing results from the two-track impact parameter correlations (blue) and the primary vertex likelihood fit (red).



(a) (b)

Figure 13: Distributions of number of tracks per event (a) and number of pixel clusters per event (b). Collision events before primary vertex selection are shown as dashed (black) line while collision events after primary vertex selection are shown as solid (red) line. Non-colliding beam events before primary vertex selection are shown as dot-dashed (blue) line while non-colliding beam events after primary vertex selection are shown as triple-dot-dashed (magenta) line.

incompatible with the direction of particles coming from the nominal interaction region. The density of pixel clusters in these events decreases with the distance from the detector center. However, anomalously large activity is also observed in the first layers of the silicon strip detector as well as in the forward calorimeters.

These events are consistent with simulated beam-gas interactions, which produce sprays of hadrons traveling at small angles with respect to the beam line. Asymmetric activity is observed in the forward calorimeters, as expected for beam-gas interactions produced away from

the nominal interaction point. The rate of beam-gas interactions is estimated to be approximately 0.5 Hz / 10^{10} protons / bunch.

To identify and reject the beam background events, various methods have been investigated. Requiring that the event have a high quality primary vertex very close to the nominal interaction point is a powerful rejection technique. The primary vertex is required to have at least four degrees of freedom, to be no more than 2 cm in the radial direction and 15 cm in the beam direction away from the nominal interaction point. This method provides a very clean sample of real proton-proton collision events. The purity of the sample can be evaluated by investigating the distributions in Fig. 13 which shows the number of tracks and number of pixel clusters in 7 TeV colliding and non-colliding beams, before and after applying the primary vertex selection. The distributions of tracks and pixel clusters from colliding beams exhibit long tails which are not compatible with expectations from simulation. We identify the source of these tails to be the beam background events. To prove this hypothesis we study the same track and pixel cluster distributions produced by non-colliding beams. Since the beams are not colliding, the corresponding distributions are dominated by beam background events from beam-gas interactions. The track and cluster multiplicities from non-colliding beams have shapes which are qualitatively different than the corresponding shapes from colliding beams, but they indeed have very long tails with shapes that match the tails from colliding beam distributions. Figure 13 also shows the non-colliding distributions after the vertex selection. As expected, a very small fraction of beam background events from non-colliding beams survives the vertex selection. This fraction is at the sub-percent level. Finally, the track and cluster distributions are shown for colliding beams after applying the vertex selection. Indeed, the long beam background tails are highly suppressed after applying the vertex selection.

Another method used to identify beam background events is based on the number of tracks in the event which pass the quality criteria described in Reference [3] as "high purity." The very large number of hits produced by beam background events lead to a high rate of fake tracks. These fake tracks fail the tight "high purity" criteria. Requiring a large fraction of high quality tracks per event is also an effective way to suppress beam backgrounds.

Alternatively, to identify beam backgrounds, one can use the fraction of pixel clusters with shapes which are incompatible with the track direction. The compatibility between the expected cluster shape, based on the track direction, and the observed cluster shape is quantified by a χ^2 probability [12]. In beam background events, particles from a beam-gas interactions far from the nominal interaction point travel at small angles with respect to the beam line. The pixel hits produced by these particles have long shapes along the beam directions. Track reconstruction assumes that particles originate from near the interaction region. Due to the large number of pixel hits, fake tracks originating from the interaction region are reconstructed as a result of random combinations of hits. These fake tracks are expected to produce pixel clusters much shorter than the long clusters from beam backgrounds. Beam background events are suppressed by requiring small fraction of pixel clusters incompatible with their track direction.

Event selections which require both a high quality primary vertex and a low fraction of hits incompatible with track directions or a high fraction of good quality tracks are used to provide samples of collision events free of beam background contamination.

6 Conclusion

We find very good performance and understanding of our detector in this first period of 7 TeV running of the LHC. The reconstruction of charged particles underlies the large majority of all

physics analyses at CMS. It is crucial to be able to estimate these performance parameters in a data-driven method which does not rely heavily upon simulation. At the same time, this then allows us to not only better understand our tracking detector and reconstruction algorithms, but to correctly describe them in our simulations. This lays a solid foundation for exploring the exciting physics taking place in the high energy proton-proton collisions of the LHC.

References

[1] "LHC Design Report". CERN-2004-003.

[2] CMS Collaboration, "The CMS Experiment at the CERN LHC", *JINST* **3** (2008) S08004.

[3] CMS Collaboration, "Tracking and vertexing results from first collisions", *CMS Physics Analysis Summary* **TRK-10-001** (2010).

[4] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 Physics and Manual; v6.420, tune ATLAS", *JHEP* **05** (2006) 026.

[5] P. Bartalini and L. Fanò, eds., "First International Workshop on Multiple Partonic Interactions at the LHC (MPI08)", volume DESY-PROC-2009-006. Perugia, Italy, October, 2008. `http://www-library.desy.de/preparch/desy/proc/proc09-06.pdf`.

[6] T. Sjöstrand, S. Mrenna, and P. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput.Phys.Commun.* **178** (2008) 852–867.

[7] GEANT4 Collaboration, "GEANT4: A Simulation Toolkit", *NIM* **A506** (2003) 250.

[8] T. Speer, K. Prokofiev, R. Fruehwirth et al., "Vertex Fitting in the CMS Tracker", *CMS Note* **2006/032** (2006).

[9] R. Fruewirth, W. Waltenberger, and P. Vanlaer, "Adaptive Vertex Fitting", *CMS Note* **2007/008** (2007).

[10] T. Miao, H. Wenzel, and F. Yumiceva, "Beam Position Determination Using Tracks", *CMS Note* **2007/021** (2007).

[11] S. Cucciarelli, M. Konecki, D. Kotlinksi et al., "Track Reconstruction, Primary Vertex Finding and Seed Generation with the Pixel Detector", *CMS Note* **2006/026** (2006).

[12] M. Swartz, D. Fehling, G. Giurgiu et al., "A new technique for the reconstruction, validation and simulation of hits in the CMS pixel detector", *PoS* (Vertex 2007) 035.