Available on CMS information server

CMS AN-NOTE 2010/xxx

CMS Analysis Note

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

February 13, 2010

Electron GSF Tracking Commissioning with first LHC Data

Abstract

The first LHC collisions at center of mass energies of 900 GeV and 2.36 TeV were recorded by the CMS detector in December 2009. The Category Based Electron Identification selection relies on the fBrem variable measurement to categorize the Electrons, hence on the quality of the GSF Tracking. It is then of vital importance its commissioning with the first recorded data. In this analysis we perform a full comparison of GSF track variables as reconstructed in data and in the simulated 900 GeV Monte Carlo events focusing on the relevant observables used in Electron Identification selection.

Contents

1 Introduction

The bremsstrahlung energy loss distribution of electrons propagating in matter is highly non-Gaussian. Since the Kalman filter relies solely on Gaussian probability density functions, it is not necessarily the optimal reconstruction algorithm for electron tracks. To better model the energy loss a Gaussian mixture rather than a single Gaussian can be used and this is implemented in the Gaussian-sum filter (GSF) algorithm [1].

The GSF algorithm has been developed in the CMS reconstruction software and heavily used for electron reconstruction in the CMS detector [2]. The GSF leads to trajectory states for each measurement point. In particular the momentum measurements at the inner ($\mathrm{p_{in}}$) and outer ($\mathrm{p_{out}}$) state are available making possible to give an estimate of the amount of energy loss in the Tracker ($\mathrm{E/fbrem} = (\mathrm{p_{in}} - \mathrm{p_{out}})/\mathrm{p_{in}}$).

The Category Based Electron Identification selection [3] relies on the $\mathrm{E/fbrem}$ measurement to categorize the Electrons and to try to separate them from fakes. The commissioning of the GSF Tracking is therefore crucial to validate the Electron Identification selection.

The first collisions at CMS were recorded in December 2009 at energies of $\sqrt{s} = 900$ GeV and 2.36 TeV. The reconstructed electrons in this data have been studied extensively to commission the electron reconstruction and identification [4]. Unfortunately the number of electron candidates produced in these collisions is low and to better commission pure tracking quantities we have performed a full re-reconstruction of the GSF tracks using standard tracking seeds. This allowed us to greatly increase the available statistic.

In this note we focus on GSF track comparison between data and Monte Carlo.

2 Data and MC Samples

The latest available data reprocessing, labelled as January 29^{th}, has been used in this analysis. The events have been analyzed using the `CMSSW_3_3_6_patch4` version of the CMS reconstruction software. The data have been then compared with the Monte Carlo prediction obtained from the simulated 900 GeV Minimum Bias events: /MinBias/Summer09-STARTUP3X_V8P_900GeV-v1/GEN-SIM-RECO.

In reprocessing the data the standard parameters for electron reconstruction have been used. Also the collection of GSF tracks, that we have analyzed, have been reconstructed starting from the seed collection used for general CTF tracking (`newCombinedSeeds`). Since this study is focused on the commissioning of the GSF algorithm used in electron reconstruction to increase the available statistic it has been applied to all the tracks (mainly produced by pions).

3 Selections

3.1 Event Selection

The available data statistic has been selected by run and lumi-section according to the particular conditions of the detector during the data taking. For the GSF Tracking study we have selected events from collisions at 900 GeV center of mass energy, requiring that the CMS magnetic field was 3.8 T.

Among the selected good runs we have also applied the following selection on each event:

- BPTX
- L1 technical trigger bit 34 (BSC "OR") fired and not 36, 37, 38, 39 (beam halo gas veto),
- hltPhysicsDeclared on
- at least one hit with 2 GeV energy in both HF,
- at least one vertex in the event with $\rho < 2$ cm, |z|<15 cm, with at least 4 degrees of freedom,
- in events with more than 10 reconstructed tracks at least 25% of them must be "high purity".

These selections lead to the run/lumi-section list shown in Table 1.

Run	Lumisec	Events
123815	8-9999	427
123818	2-42	1075
123906	18-28	186
123908	2-12	102
124006	7-9999	65
124008	1-1	46
124009	1-68	10736
124020	12-94	17579
124022	66-179	41341
124023	38-9999	35641
124024	2-83	29343
124025	5-13	2644
124027	24-9999	13301
124030	2-9999	18514
		171000

Table 1: List of the good run/lumi-section used in the analysis with the corresponding number of events passing the selection described in the text.

3.2 Track Selection

The study of the GSF track properties has been carried on with a selected set of good tracks. To separate the badly reconstructed tracks to the good ones with have applied the following criteria:

- the track has to be at least 7 hits.
- the track must be qualified "high purity" and has to be reconstructed during the 0th and 1st iterations of iterative tracking procedure (using only pixel triplet and pair seeds). Since this information is not directly available for GSF track we have first matched the GSF track to one of the reconstructed CTF tracks and then we have used the information stored in the CTF. The matching has been done by hits, requiring the two tracks to have at least 50% of shared hits (fraction computed on the shortest track).

As a cross-check some of the following results have been checked for the full set of reconstructed tracks also.

4 Results

In this section the complete set of results is presented. If not specified the following plots are normalized to the number of data entries and refer to high purity tracks.

Tab. 2 reports the number of events and tracks used in the analysis, together with the number of tracks falling in the different categories. About 67% of the analyzed tracks passed quality cuts, while only few percent (4%) of them could not be associated to any reconstructed CTF track. Looking more closely to the reconstructed electron candidates it turns out that almost 95% of them has an high purity GSF track. The Table 2 also shows how the Monte Carlo simulation foresees about 20% less reconstructed tracks.

	Events	Tracks	Tk/evt	Quality	No High purity	No CTF match
Data	171000	1404937	8.22	946228	398277	60432
MC	494675	33374188	6.82	2393783	868378	116527

Table 2: Number of data and minimum bias events analyzed. The table shows the corresponding number of reconstructed GSF Tracks split in different categories.

Fig. 1 shows the $\mathrm{p_T}$ distribution of the GSF tracks in data and MC. The distributions are normalized to the total number of reconstructed GSF tracks in data and show an excess of data in the Barrel. The transverse momentum ($\mathrm{p_T}$) distributions of GSF tracks are anyway well reproduced.

Fig. 2 (a) and (b) show rapidity (η) distributions for GSF and CTF high purity tracks. There are some discrepancies in the high η regions in both the distributions. The bias is not completely understood but it is introduced by the

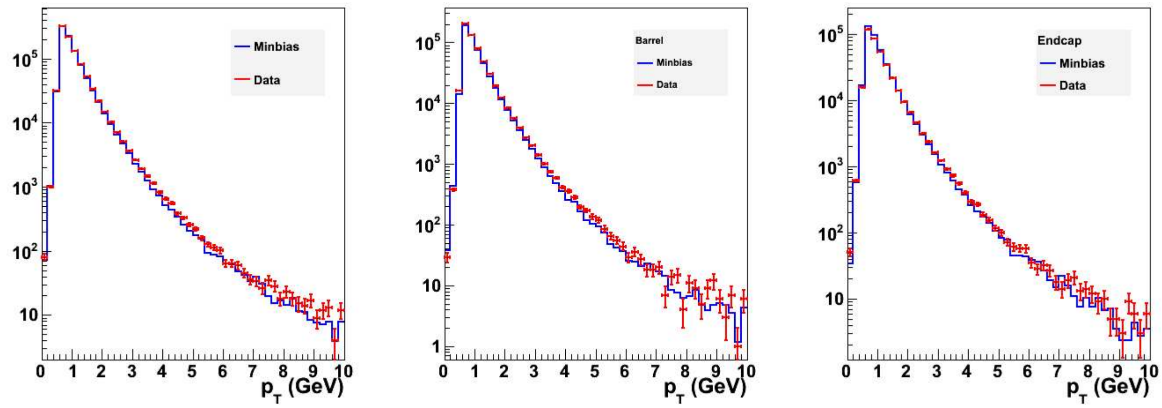


Figure 1: Transverse momentum ($\mathrm{p_T}$) distribution of GSF tracks in data (points) and MC (histogram). The three plots shows, from left to right, the total distribution, tracks in the barrel and tracks in the Endcap. The three plots are normalized to the total number of reconstructed tracks in data.

quality requirements. If the cut is removed and all the tracks are selected, Fig. 2 (c), the agreement between data and MC improves.

As already stated the $\mathrm{E/fbrem}$ variable (defined as $(\mathrm{p_{in}} - \mathrm{p_{out}})/\mathrm{p_{in}}$) gives an estimate of the energy loss of an electron indipendently by the Calorimeter using only the Tracker. This variable is extensively used in the Category Based Electron Identification so it is important to check how well it is reproduced by the MC simulation. Since the large majority of the reconstructed GSF tracks belong to pions, which doesn't brem, we expect a distribution peaked at 0.

The $\mathrm{E/fbrem}$ distribution of the GSF tracks has been studied in $\mathrm{p_T}$ and η bins. Fig. 3 shows the $\mathrm{E/fbrem}$ distributions in bins of the rapidity. The agreement between data and MC is very good in the central part of the detector ($|\eta| < 1.0$). Then, as the quantity of tracker material increases, the peak in data seems to be slightly shifted towards 1 and does not match with the MC expectations. In the Endcap the discrepancy is more pronounced. Fig. 3 shows the $\mathrm{E/fbrem}$ variable in four η bins: two in the barrel and two in the Endcap. The results for the forward and bakcward sides of the detector are shown separately and no particular asymmetry can be revealed.

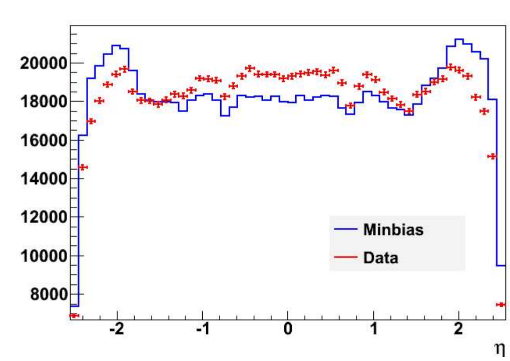
The fbrem binning in $\mathrm{p_T}$ of the GSF track confirms the shift of the peak in the Endcap, showing, at the same time, no $\mathrm{p_T}$ dependence of the effect. These results are in agreement with what has been seen in the analysis of the electron candidates only. Fig. 4 reports the results and Fig. 5 shows the total fbrem distribution in Barrel and Endcap.

Another important variable related to GSF tracks that is used in the Electron Identification selection is the number of missing hits. This is the number of crossed layers without compatible hits in the back-propagation of the track to the beam-line. Fig. 6 shows the missing hits variable comparison in Barrel and Endcap. The distributions show some discrepancies that need to be understood.

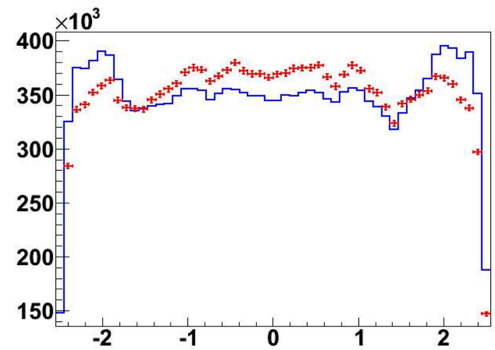
Electron Identification selection uses also an Impact Parameter cut (computed with respect to the reconstructed vertex). Fig. 7 shows the impact parameter distributions which is in good agreement with MC except for the dip around 0.5 cm which is under investigation.

Other GSF tracks distributions have been checked even if not directly correlated with Electron Identification selection. Fig. 8 shows the χ^2 of the GSF Tracks in Barrel and Endcap which is very well reproduced by the MC simulation. Also ratio between the track momentum estimations of the GSF (mode) and of the associated CTF track are well reproduced, the distributions for Barrel and Endcap are shown in Fig. 9.

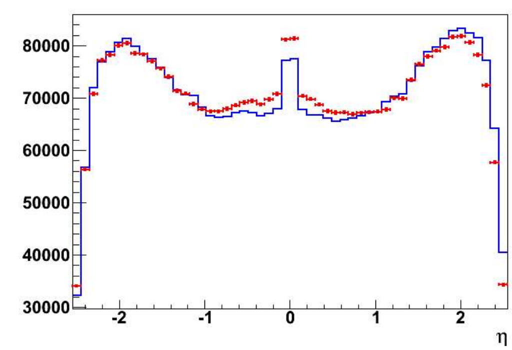
We have also compared the number of hits per track for a GSF track and the corresponding CTF. The variable that has been studied is the ratio of the number of hits of GSF and CTF tracks (nhits(GSF)/nhits(CTF)) and it has been done in four bins of the ratio $\mathrm{p_{out}/p_{in}}$ of the GSF track. It turns out that the four distributions are in good agreement with the simulation. Beside that it is clear how the ratio of number of hits increases as the fraction $\mathrm{p_{out}/p_{in}}$ decreases: the higher is the energy loss in the Tracker the harder is for the CTF pattern recognition to produce a track as long as for the GSF one. Fig. 10 shows the four distributions.



(a)



(b)



(c)

Figure 2: η distribution of the high purity GSF (a) and CTF (b) Tracks. The η distribution of all the tracks is shown in (c). Data are in points, MC in histogram.

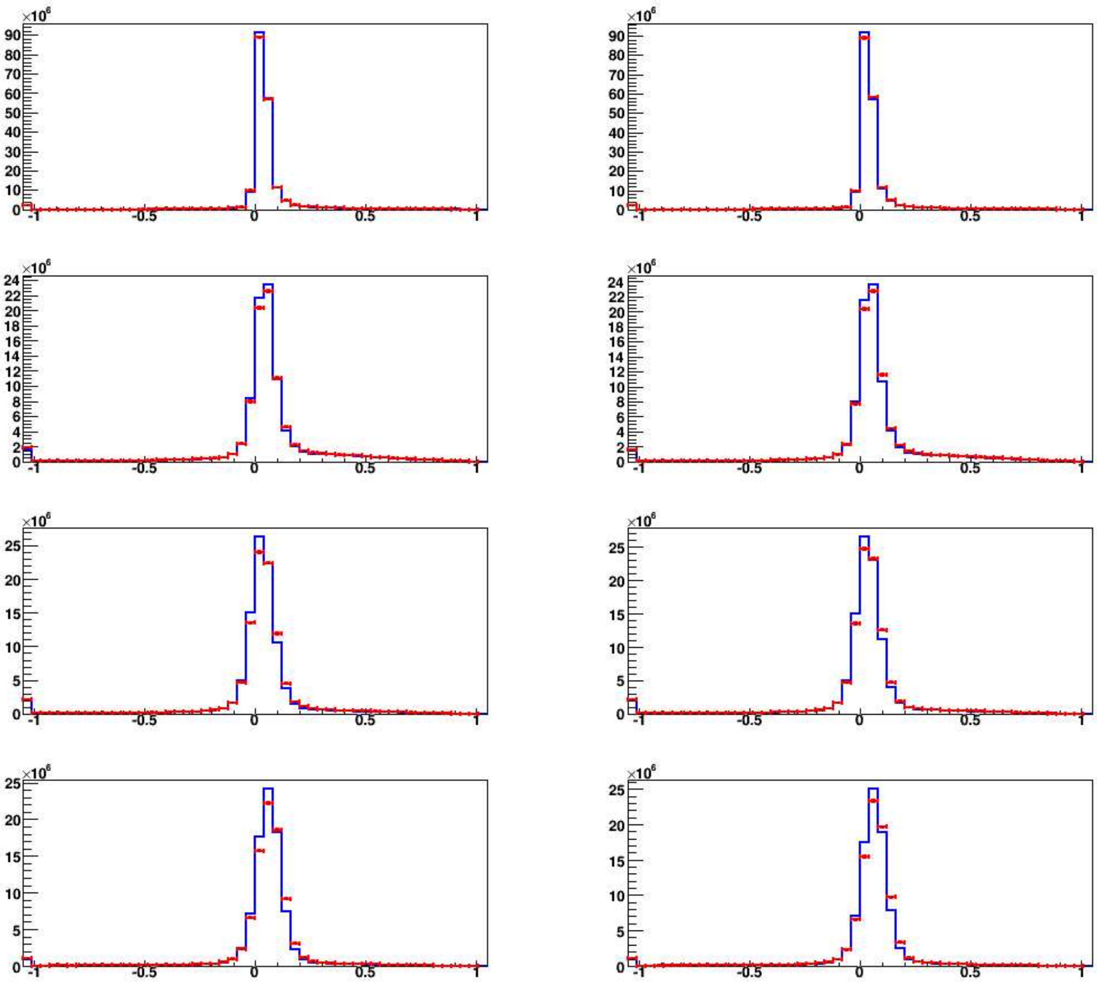


Figure 3: $\mathrm{E}/\mathrm{fbrem}$ distribution in η bins for highpurity GSF Tracks in data (points) and MC (histogram). The four η bins are from top to bottom: $|\eta|$ < 1, $1 < |\eta| < 1.479$,$1.479 < |\eta| < 2$, $|\eta| > 2$. Each row shows the distributions for forward and backward sides of the CMS detector.

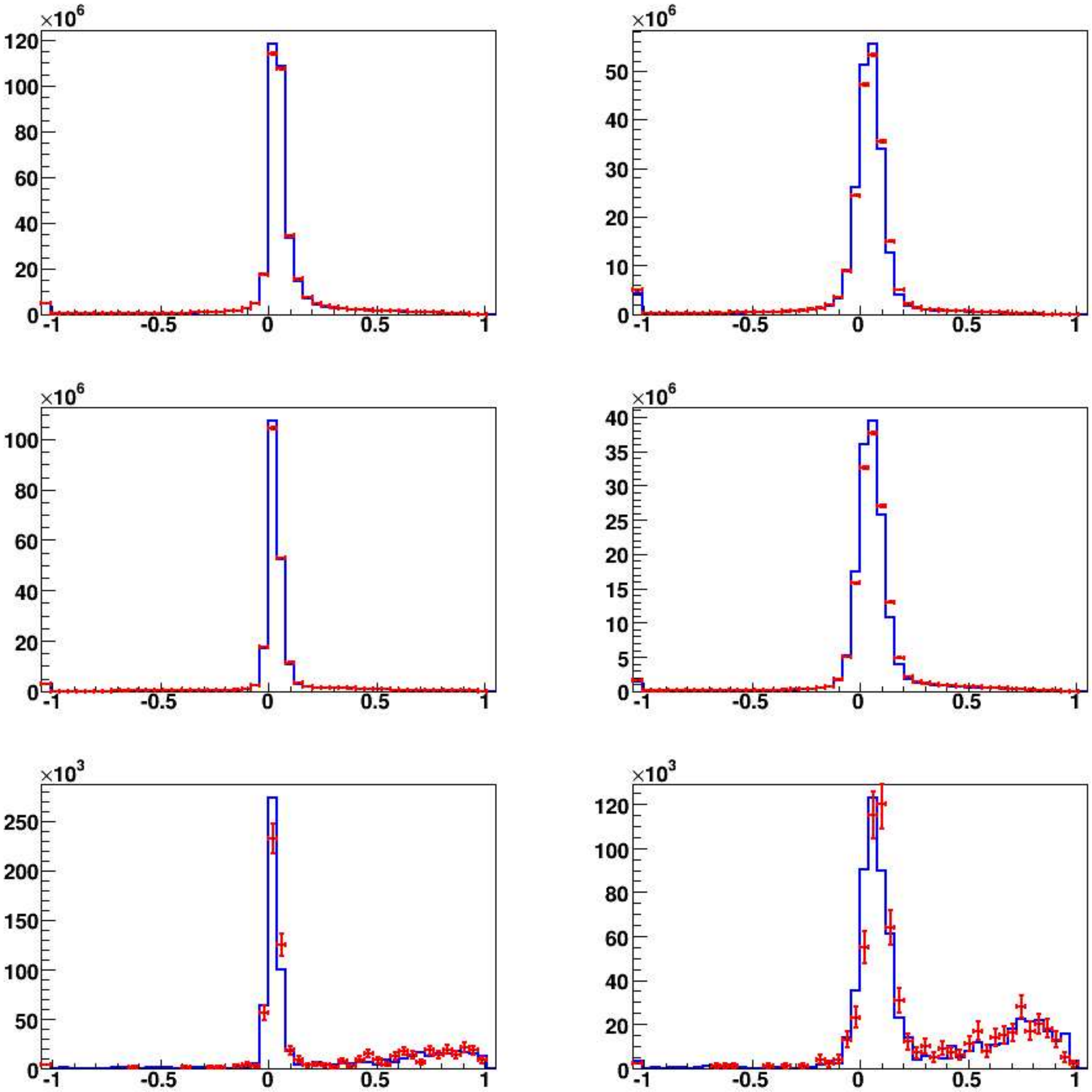


Figure 4: $\mathrm{E/fbrem}$ distribution in $\mathrm{p_T}$ bins for data (points) and MC (histogram). From the top to the bottom the $\mathrm{p_T}$ bins are defined as: $\mathrm{p_T} < 1$ GeV/c, $1 < \mathrm{p_T} < 5$ GeV/c,$\mathrm{p_T} > 5$ GeV/c. Each row shows the distribution for Barrel (left) and Endcap (right).

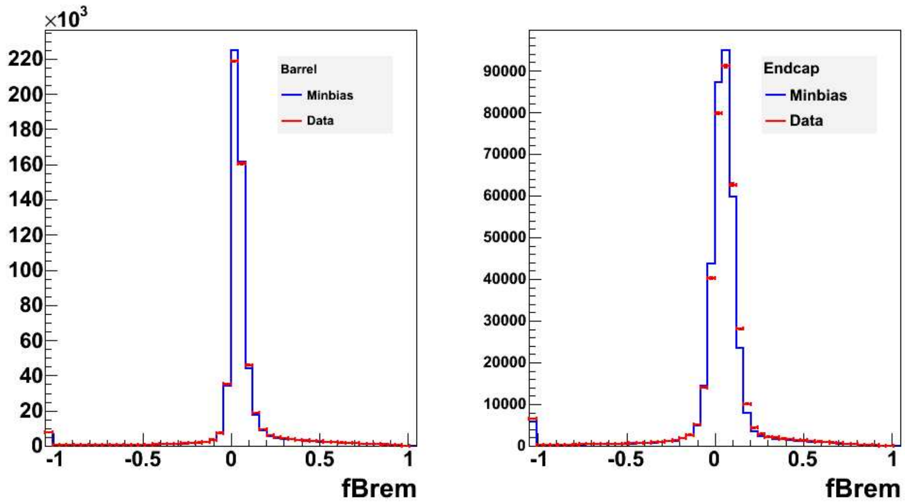


Figure 5: Total $\mathrm{E/fbrem}$ distribution for high purity GSF Tracks in data (points) and MC (histogram). The two plots show Barrel (left) and Endcap (right) distributions.

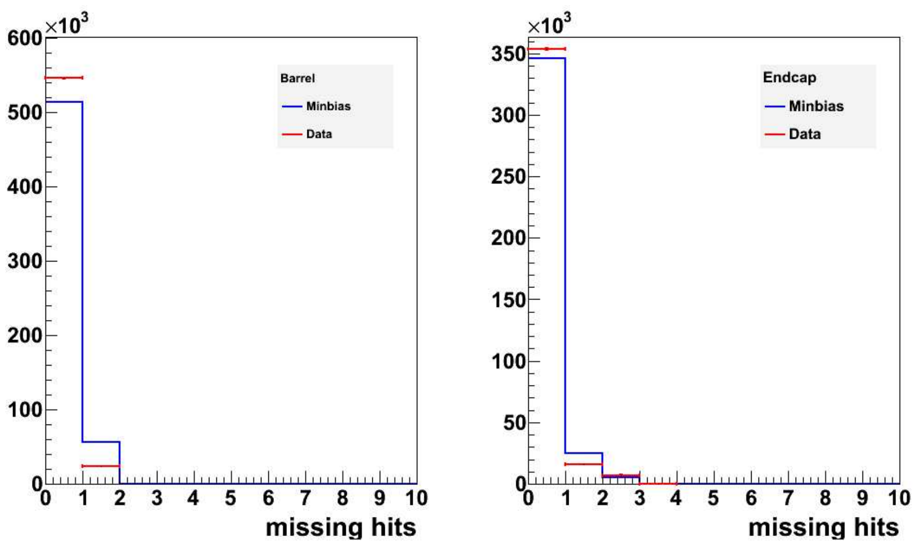


Figure 6: Missing hits distribution for high purity GSF Tracks in data (points) and MC (histogram). The two plots show the distributions in Barrel (left) and Endcap (right).

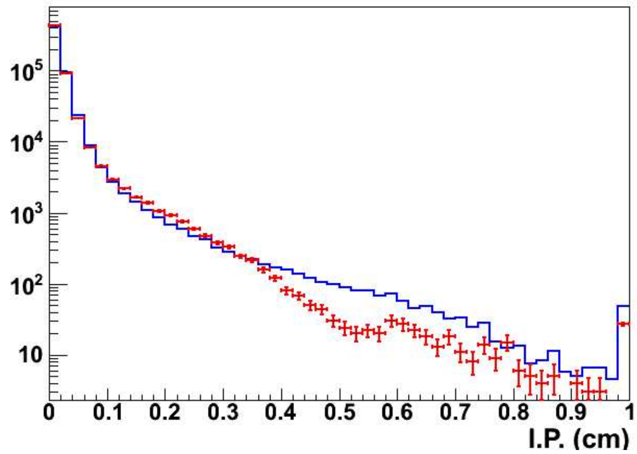


Figure 7: Impact parameter computed with respect to the reconstructed vertex of high purity GSF Tracks in data (points) and MC (histogram).

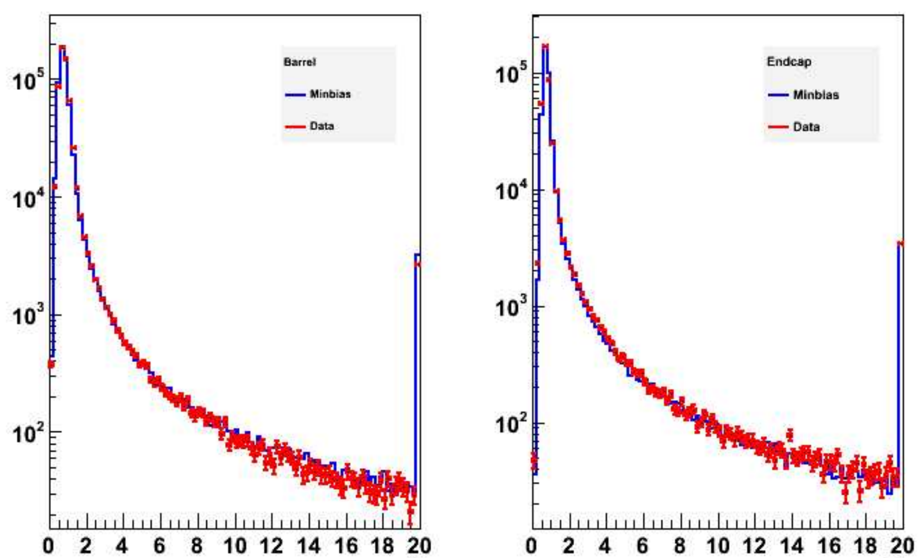


Figure 8: χ^2 distribution for high purity GSF Tracks in data (points) and MC (histogram). The two plots show the distributions in Barrel (left) and Endcap (right).

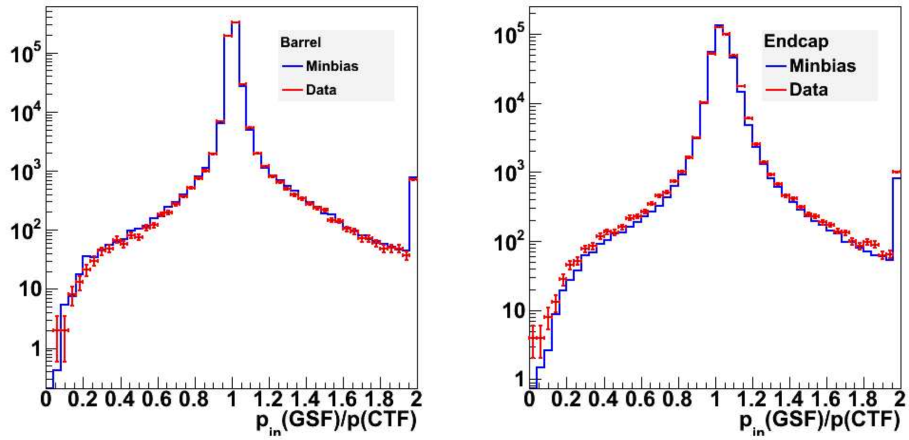


Figure 9: Ratio between the $\mathrm{p_{in}}$ momentum estimated as the mode of the GSF Track and the momentum of the associated CTF track. Data are in points, MC in histogram. The two plots show the distributions in Barrel (left) and Endcap (right).

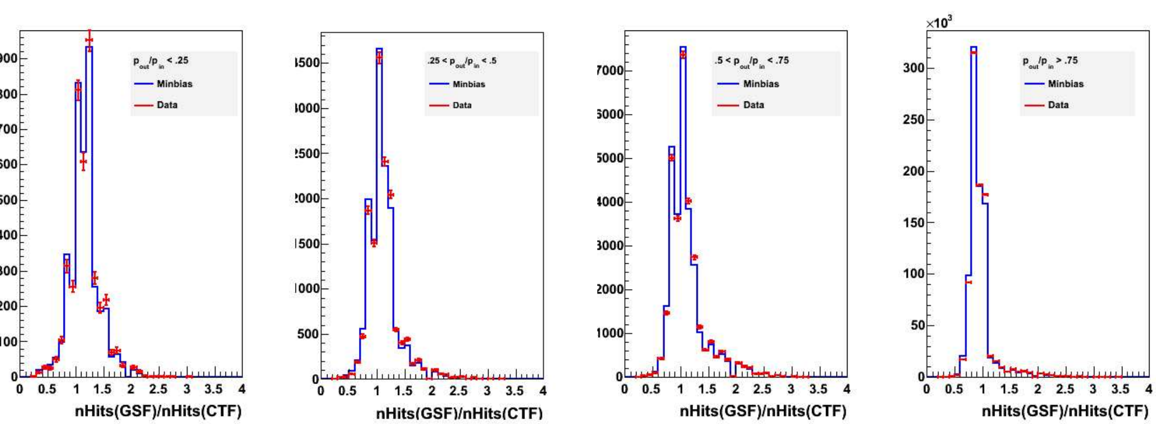


Figure 10: Ratio between number of hits of high purity GSF Tracks and the associated CTF Track in data (points) and MC (histogram). The distributions are shown in bins of the fraction $f = \mathrm{p_{out}}/\mathrm{p_{in}}$, from left to right: $f < .25$, $.25 < f < .5$, $.5 < f < .75$, $f > .75$.

5 Conclusions

In the process of commissioning the Category Based Electron Identification selection we wanted to check the reliability of pure tracking variables, like $\mathrm{E}/\mathrm{fbrem}$ and missing hits, comparing collision data recorded during December 2009 data taking and MC simulated minimum bias events. Unfortunately the number of electron candidates reconstructed during 2009 data taking was not enough. To increase the available statistic we have performed a full re-reconstruction using GSF Tracking algorithm starting from the seed collection used in the general CTF tracking. All the studied variables show a quite good agreement between the collision data and the simulation even if more investigations are needed to understand the remaining discrepancies.

References

[1] W. Adam et al., *'Reconstruction of electrons with the Gasussian-sum filter in the CMS tracker at the LHC', J. Phys. G: Nucl Part. Phys. 31 N9-N20 (2005)*

[2] W. Adam et al., 'Electron Reconstruction in CMS', CMS AN-2009/164, Eur. Phys. J. C 49 (2007) 1099-1116.

[3] S. Baffioni et al., 'Electron Identification in CMS', CMS AN- 2009/173.

[4] ECGelec working group, 'Electron Commissioning from first CMS Data', CMS AN-2010/037 (in preparation).