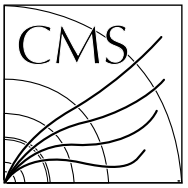
Available on CMS information server

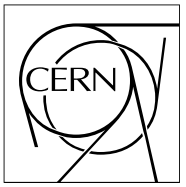
CMS NOTE 2006/40



The Compact Muon Solenoid Experiment

CMS Note

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



February 2006

Electron Reconstruction in CMS

S. Baffioni [a)], C. Charlot [a)], F. Ferri [a) b)], D. Futyan [c)], P. Meridiani [d)], I. Puljak [e)], C. Rovelli [a) b)], R. Salerno [a) b)], Y. Sirois [a)]

a) Laboratoire Leprince-Ringuet, Ecole Polytechnique and IN2P3-CNRS, Palaiseau, France

b) Università di Milano Bicocca and INFN Milano, Milano, Italy

c) University of California, Riverside

d) Università di Roma I and INFN, Rome, Italy

e) Technical University of Split, FESB, Croatia

Abstract

The reconstruction of the energy and momentum of isolated electrons in CMS combining tracking and electromagnetic calorimetry information is described. The emphasis is put on primary electrons with transverse momentum below 50 GeV$/c$. The energy deposited in the electromagnetic calorimeter is measured in clusters of clusters (superclusters) which collect bremsstrahlung photons emitted along the electron trajectory in the tracker volume. The electron tracks are built from seeds in the pixel detector found via a cluster-driven pixel hit matching algorithm, followed by a reconstruction of trajectories in the silicon strip tracker with a Gaussian Sum Filter. Electrons are classified using observables sensitive to the pattern of bremsstrahlung emission and electromagnetic showering in the tracker material. Energy scale corrections depending on the electron class are applied to the supercluster and estimates of associated errors are obtained. The electron energy is deduced from a weighted combination of the corrected supercluster energy and tracker momentum measurements. The electron direction is that of the reconstructed electron track at interaction vertex. The pre-selection of isolated electron candidates for physics analysis is described. Class-dependent observables combining tracking and calorimetry information are discussed for electron identification.

1 Introduction

A strategy for the reconstruction of electrons in CMS is presented in this note. The emphasis is put on the energy measurement, the isolation and identification of primary electrons in the p_T range from 5 to 50 GeV/c. The combination of tracking and calorimetric information allows low p_T electrons to be measured and identified in the challenging kinematics and background conditions relevant for the Standard Model Higgs boson decays $\mathrm{H} \to \mathrm{ZZ}^* \to \mathrm{e}^+\mathrm{e}^-\mathrm{e}^+\mathrm{e}^-$ and $\mathrm{H} \to \mathrm{WW}^* \to \mathrm{e}^+\nu\,\mathrm{e}^-\bar{\nu}$.

The reconstruction of electrons in CMS uses information from the pixel detector, the silicon strip tracker and the electromagnetic calorimetry (ECAL). A brief description of these detectors is given in Section 2. The measurement of the electron energy in the ECAL is hampered by the amount of tracker material which is distributed in front of the ECAL, and by the presence of a strong magnetic field aligned with the collider beam axis (thereafter called z axis). Electrons traversing the silicon layers of the tracker radiate bremsstrahlung photons and the energy reaches the ECAL with a significant spread in the azimuthal direction ϕ. The ECAL clustering, and in particular the building of "superclusters" (clusters of clusters), is designed to take into account the ϕ spread and collect the bremsstrahlung energy. The electron clustering is described in Section 3. Supercluster-driven pixel-seed finding is then used to initiate the building of trajectories in the inner tracker. The electron track reconstruction relies on a dedicated "Gaussian Sum Filter" (GSF) [1, 2] using a specific energy loss modeling. Electron pixel-seed finding and GSF track reconstruction is described in Section 4.

The bremsstrahlung emission introduces, in general, non-Gaussian contributions to the event-by-event fluctuations of the calorimetry and tracking measurements. Additional electron tracks from conversion of secondary photons, actually the first stages of an "electromagnetic showering", contribute to the energy lost in front of the ECAL. More elaborate reconstruction procedures, involving recognition of distinct track-supercluster patterns are in general needed to better disentangle the sources of partial energy containment in the supercluster, adapt the energy scale corrections and estimate associated errors. Different "classes" of electrons are introduced for such purposes in Section 5 where the sources of measurement variations are further discussed. The supercluster energy scale corrections are presented in Section 6. The estimate of the electron energy discussed in Section 7 combines tracking and calorimetry measurements, with track-based information dominating towards low p_T and ECAL-based information dominating towards high energy. The electron direction is obtained from the associated primary track.

The pre-selection of isolated primary electrons is discussed in Section 8. The main background for primary electrons in CMS comes from "fake" electrons from hadron overlaps in jets, but also prompt electrons from semi-leptonic decays of mostly c or b quarks, and possibly electrons from early photon conversions in the tracker material. The bremsstrahlung emission and secondary conversions which accompany real electrons complicates the identification strategy. Whether the electron measurements are compatible with a small amount of bremsstrahlung or, on the contrary, characterized by considerable ϕ spread and secondary conversions, is likely to affect electron identification and background rejection performance. "Good" and "bad" electrons in general have to enter with different weights at physics analysis level due to their different energy-momentum errors and different sample purity. Sophisticated electron reconstruction and identification procedures are essential to recover high efficiencies in particular at low p_T. The definition of variables for electron identification profits from the classification of electron patterns. It also takes benefit of the new observables made available with the GSF track reconstruction method, and in particular of the meaningful track parameters provided at both track ends. Such an electron identification strategy is presented here in a realistic context with filtering, pre-selection, isolation and identification steps. Reconstruction tools have been developed which can be steered for the specific needs of a given analysis in terms of reconstruction efficiency, background rejection and required purity of the physics signal. Electron identification variables adapted to the various class of electron track-supercluster patterns are discussed in Section 9.

The reconstruction studies presented here have been performed in the CMS "ORCA" framework [3] using detailed Monte Carlo simulation of back-to-back $\mathrm{e}^+\mathrm{e}^-$ pairs in the absence of LHC pile-up. Electron samples at either fixed energy (10, 30 and 50 GeV), fixed p_T of (10, 30 and 50 GeV/c), uniformly distributed in energy from 5 to 100 GeV, or uniformly distributed in p_T from 5 to 50 GeV/c have been considered.

2 The CMS Detector

A detailed description of the CMS detector can be found elsewhere [4, 5, 6]. Here only some of the relevant characteristics of the main detectors used for electron reconstruction are presented.

The CMS tracker [5] is a cylindrical detector of 5.5 m in length and 1.1 m in radius. It is equipped with silicon pixel detectors (66M channels) for the innermost part (for radii $R < 15$ cm and for $|z| < 50$ cm) and silicon strip detectors (2.8M channels) for the outer layers ($R < 110$ cm, $|z| < 275$ cm). The pixel detectors provide in general

2 or 3 hits per track, each with a three-dimensional precision of about 10μm in the transverse plane ($R\phi$) and 15μm in z. The strip detectors can provide up to 14 hits per track, with a two-dimensional precision ranging from 10μm to 60μm in $R\phi$. Some of the silicon strip layers are double-sided to provide a longitudinal measurement with a similar accuracy. The tracker acceptance for a minimum of 5 collected hits extends up to pseudorapidities η of about $|\eta| < 2.5$. The efficiency for collecting 2 hits in the pixel detector drops from close to 100% at $|\eta| \simeq 2.1$ to below 70% at $|\eta| \simeq 2.5$ [7].

The material thickness in the tracker volume to be traversed by electrons and photons before reaching the ECAL varies strongly with η. It amounts to about $0.35X_0$ at central pseudorapidities ($\eta = 0$), increases to $\simeq 1.4X_0$ towards the ECAL barrel/endcap transition, and falls back to about $0.8X_0$ at $|\eta| = 2.5$.

The CMS ECAL [6] is made of $PbW0_4$ crystals, a transparent material denser (8.3 g/cm^3) than iron, with a radiation length X_0 of 0.89 cm and a Moliere radius R_M of 2.19 cm. The ECAL is composed of a barrel covering $|\eta| \lesssim 1.48$ and and two endcaps covering $1.48 \lesssim |\eta| \lesssim 3.0$. The barrel is made of 61200 trapezoidal and quasi-projective crystals of approximately $1 \times R_M$ in lateral size and about $25.8X_0$ in depth. The barrel inner radius is of 124 cm. Viewed from the nominal interaction vertex, the individual crystals appear tilted (off-pointing) by about 3° both in polar and azimuthal angles, and the granularity is about $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$ rad. The barrel is divided in two halves, each made of 18 supermodules containing 1700 crystals. Each supermodule is composed of four modules. The endcaps consist of two detectors, a preshower device followed by $PbW0_4$ calorimetry. The preshower is made of silicon strips placed in a 19 cm sandwich of materials including about $2.3X_0$ of Pb absorber. The preshower covers inner radii from 45 cm to 123 cm, corresponding to the range $1.6 < |\eta| < 2.6$. Each endcap calorimeter is made of 7324 rectangular and quasi-projective crystals of approximately $1.3 \times R_M$ in lateral size and about $24.7X_0$ in depth. The crystal front faces are aligned in the (x, y) plane but, as for the barrel, the crystal axes are off-pointing from the nominal vertex in the polar angle by about 3°.

The CMS inner tracking and ECAL detectors are immersed in a 4 T magnetic field parallel to the z axis.

3 Electron Clustering

The electromagnetic showers initiated by electrons (or photons) deposit their energy in several crystals of the ECAL. For a single electron (or photon) reaching the ECAL, most of the energy is collected in a small number of crystals. For a supermodule of the ECAL barrel in the test beam, electrons with an energy E^{e} of 120 GeV impinging at the centre of a crystal for instance deposit about 97% of their incident energy in a 5×5 crystal window [8]. Such simple fixed size arrays of crystals have been shown to allow for best measurement performance for electrons in test beam provided that so-called local containment corrections are applied to account for the variation of the measured energy as function of the shower position with respect to the cluster boundary. Simple clusters made of fixed size arrays have also been considered for the CMS full experiment to measure unconverted photons and as means of simplifying the measurement of a selected sample of "low-radiating" electrons for the intercalibration of crystals.

As mentioned in the introduction, the situation is in general more complicated for the average electron. Electrons traversing the tracker material radiate photons and the energy reaches the ECAL spread in ϕ. Integrated along the electron trajectory the effect can be very large. Figure 1 shows for example the distribution of the fraction of the initial energy radiated by electrons before reaching the ECAL, for electrons of 10, 30 and 50 GeV. Such a distribution is the result of the convolution of the bremsstrahlung spectrum, the finite path to reach the ECAL and the finite initial energy of the electrons. About 35% of the electrons radiate more than 70% of their initial energy before reaching the ECAL. In about 10% of the cases, more than 95% of the initial energy is radiated.

Thus, to obtain a measurement of the electron energy at primary vertex and minimize the cluster containment variations, it is essential to collect bremsstrahlung photons. This is the purpose of the super-clustering algorithms. In the energy range considered in this note the basic "Hybrid" and "Island" clustering algorithms, described in detail in Refs. [9, 10], can be used for electrons in the ECAL barrel and endcaps respectively. The Hybrid algorithm attempts to profit from the simple geometry of the ECAL barrel and exploit the properties of the lateral shower shape in the transverse direction while dynamically searching for separated (bremsstrahlung) energy in the ϕ direction. In the language of the Hybrid super-clustering, what is considered here as a "seed" cluster is a collection over ϕ of contiguous dominoes made of 3 to 5 crystals in η and separated by other such collections by a valley where less than 100 MeV is observed in a domino. The Island algorithm in the endcap builts clusters by connecting rows of crystals containing energies decreasing monotonically when moving away from a seed crystal. Then, superclusters are built by collecting other Island clusters in a ϕ road in both directions around each Island clusters, starting from a list of clusters ordered in E_T, in a procedure called bremsstrahlung recovery. In the language of this algorithm, what is considered here as a "seed" cluster, is a cluster that initiates a bremsstrahlung recovery procedure.

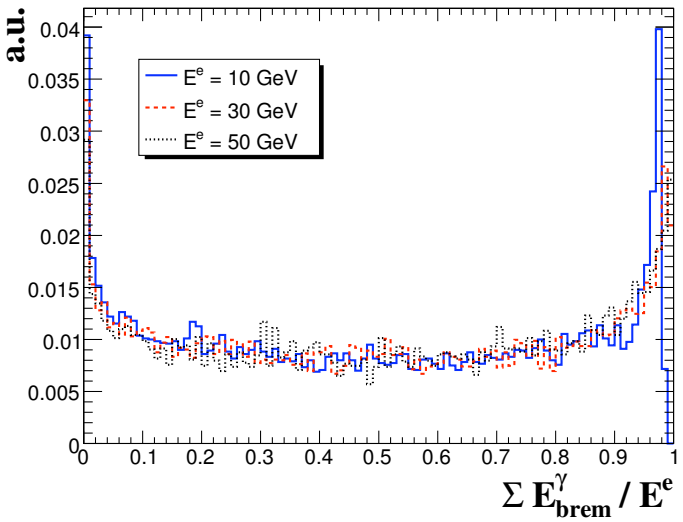


Figure 1: Distribution of the fraction, $\sum E^{\gamma}_{\rm brem}/E^{\rm e}$, of the generated electron energy ($E^{\rm e}$) radiated as bremsstrahlung photons ($\sum E^{\gamma}_{\rm brem}$) for electrons of 10, 30 and 50 GeV. The true emission of bremsstrahlung photons has been integrated up to a radius corresponding to the ECAL inner radius.

A tuning of the supercluster building parameters, with respect to CMS DAQ TDR [9] settings, has been performed for both the Hybrid and the Island algorithms. The minimal E_T threshold for the basic seed cluster of a supercluster has been lowered from the previous default of $E_T^{\rm seed} = 4$ GeV down to $E_T^{\rm seed} = 1$ GeV. This leads to considerable improvement of the efficiency for reconstructing an electron supercluster for low $p_T^{\rm e}$. Integrating over the acceptance in η, this efficiency for back-to-back $\rm e^+e^-$ pairs reaches $\geq 99\%$ for $p_T^{\rm e} \geq 7$ GeV/c and $E_T^{\rm seed} = 1$ GeV, compared to an original efficiency with $E_T^{\rm seed} = 4$ GeV varying from about 65% for $p_T^{\rm e} = 7$ GeV/c to about 93% for $p_T^{\rm e} = 10$ GeV/c. Having lowered the supercluster seed threshold, there is a tendency for extra basic clusters caused by radiated photons with $p_T^{\gamma} > 1$ GeV/c to remain separate and form their own supercluster. With the original ϕ roads for bremsstrahlung recovery extending to ± 0.2 rad in the endcaps and ± 10 crystals (i.e. about ± 0.17 rad) in the barrel, about 8% of back-to-back $\rm e^+e^-$ pairs at $p_T^{\rm e} \geq 7$ GeV/c give three ECAL superclusters. To better collect the bremsstrahlung and reduce (well below 1%) the probability to find a number of superclusters in excess of the number of isolated electrons, the ϕ roads have been increased to ± 0.3 rad in the endcaps and ± 17 crystals (i.e. about ± 0.3 rad) in the barrel.

4 Electron Track Reconstruction

The track reconstruction procedure in CMS [3, 9] is decomposed into four modular components. Firstly, initial tracks called *seeds* are looked for with a *Seed Generator*. Then the *Trajectory Builder* constructs outward all the possible trajectories for a given seed. With the *Trajectory Cleaner* ambiguities among the possible trajectories are solved and a maximum number of track candidates is kept. Finally, the final fit of the track is performed with the *Trajectory Smoother*, which uses all the collected hits to estimate the track parameters at each layer through a backward fit. For electron tracks, in order to better deal with the non-Gaussian fluctuations induced by bremsstrahlung emission, dedicated algorithms have been developped for the seeding and building steps, as well as for the smoothing step where a GSF is used instead of the standard Kalman Filtering (KF) [3, 9] both for forward and the backward fits. These steps are described in the following. The cleaning procedure used for electrons is the same as that used for other types of tracks [3].

4.1 Seed Generation

In order to build a track outward, a seed is created when two hits compatible with a given beam spot are found in the pixel detector.

To tame the many possible hit combinations in the case of electron tracks, the search for seeds better be restricted to a region compatible with a supercluster in the ECAL. In principle, this could be achieved via a simple "regional" restriction (relying on the observation of an ECAL supercluster), at the expense of a more severe fake

track rate which would have to be compensated at later stage when resolving ambiguities to select electron candidates. A somewhat more powerful approach is to start from the basic element which most uniquely characterizes an electron, namely the presence of an electromagnetic supercluster. The supercluster-driven pixel seed finding presents the advantage, for comparable reconstruction efficiencies, of increasing the purity of the sample of candidate electron tracks.

Such a cluster-driven pixel seed finding strategy for the tagging of primary electron-like objects has been successfully developed for robust applications at the High Level Trigger (HLT) [9], where fast and drastic reduction of fake background rates is a key issue. The supercluster-pixel matching takes advantage of the fact that the energy weighted average impact point of the electron and associated external bremsstrahlung photons, as calculated using information from the supercluster in the ECAL, coincides with the impact point that would have been measured for a non-radiating electron of the same initial momentum. This trick works fine provided that the photons from external bremsstrahlung are properly collected. For isolated electrons having p_T's well below HLT trigger thresholds, the tuning of the clustering algorithm parameters described in Section 3 directly contributes to an improvement of the supercluster-driven pixel matching efficiency.

Hits in the pixel layers are predicted by propagation of the energy weighted mean position of the supercluster backward through the magnetic field under both charge hypotheses towards the pixel detector. A first compatible hit is then looked for in the innermost (barrel) pixel layer within a loose $\Delta\phi$ window (adapted to the uncertainty on the $\phi_{\rm sc}$ measurement) and loose Δz interval (adapted to the spread of the interaction vertices). In cases where no hit is found in the innermost layer, the first hit is looked for in the next-to-innermost layer. This accounts for possible pixel finding (reconstruction or algorithmic) inefficiencies. When a first compatible hit is found, a new estimate z_0 for the z coordinate of the primary track vertex is calculated combining the pixel hit found and calorimetry information in the Rz plane. The predicted trajectory is then propagated to look for a second pixel hit in the next pixel layer(s). More details on the supercluster-driven pixel matching algorithm can be found in Ref. [9].

The requirements for the search of the first and second pixel hits have been loosened with respect to the HLT to recover electron detection efficiency at low p_T. The new threshold values are given in Table 1.

Observable	Requirements	
	HLT level	Offline Reconstruction
$\Delta\phi$ - 1$^{\rm st}$ pixel hit	40 mrad	200 mrad
	e^- : $[-25\,\mathrm{mrad}, +15\,\mathrm{mrad}]$	
	e^+ : $[-15\,\mathrm{mrad}, +25\,\mathrm{mrad}]$	
Δz - 1$^{\rm st}$ pixel hit	15 cm	15 cm
$\Delta\phi$ - 2$^{\rm nd}$ pixel hit	2 mrad	10 mrad
Δz - 2$^{\rm nd}$ pixel hit	± 0.05 cm	± 0.07 cm

Table 1: Seeding of electron tracks. Allowed $\Delta\phi$ and Δz search windows for the 1$^{\rm st}$ and 2$^{\rm nd}$ pixel hits of the supercluster-driven seed finding algorithm, for HLT and offline reconstruction.

The effect of the change of the pixel matching criteria is illustrated in Fig. 2 for single electrons at fixed p_T (back-to-back $\mathrm{e^+e^-}$). The two pixel hits found then serve as a seed for the building and fitting of an electron track in the Silicon Tracker Detectors.

4.2 Trajectory Building

Starting from the seed, a trajectory is created. Compatible hits on the next silicon layers are first searched for, then an extrapolation is performed, using a Bethe Heitler modeling of the electron losses and a GSF in the forward fit. This procedure is iterated until the last tracker layer, unless no hit is found in two subsequent layers. The trajectory state at each layer is computed as the weighted mean of the predicted state and of the measured hit. The compatibility among them is defined in term of a χ^2 test. If many hits are found on a compatible layer, many candidate trajectories are grown in parallel. In order not to lose efficiency at this stage, no specific χ^2 cut is applied in the building steps of the GSF tracks, but only the best two candidates (giving the smallest χ^2) are kept. A minimum of five hits is finally required to create a track.

Figure 3 shows the efficiency of electron track reconstruction as function of p_T and η, for isolated electrons with a uniform p_T spectrum between 5 and 50 GeV/c and a uniform η distribution within $|\eta| < 2.5$. The efficiency is defined as the fraction of generated electrons which have a reconstructed electron track with the same charge and which matches in η and ϕ within ± 0.05 units. The reconstructed track parameters are evaluated at the point of closest approach to the generated vertex, with a further backward (i.e., outside-in) fit of the trajectory, so as to

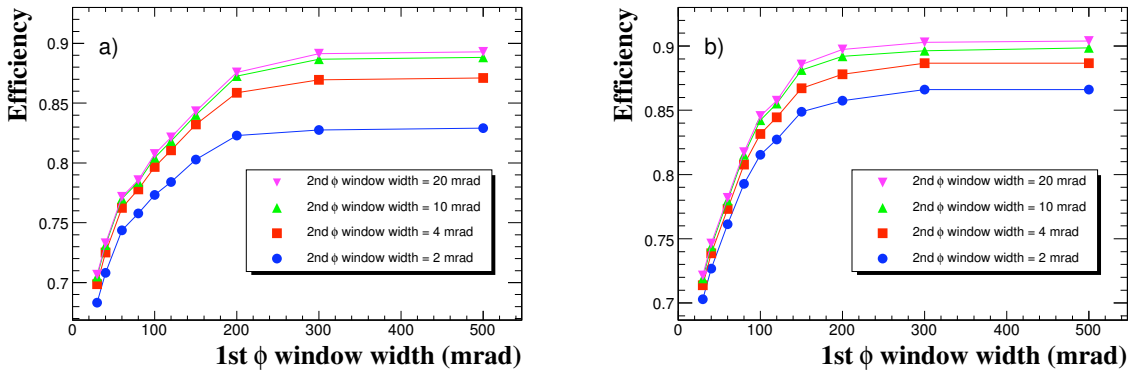


Figure 2: Efficiency of the pixel matching for electron tracks reconstruction as a function of the ϕ window for the finding of the first pixel: a) $p_T^e = 7$ GeV/c electrons and b) $p_T^e = 10$ GeV/c electrons. Plots are shown for varying the setting of the ϕ window, in which a second pixel hit is seeked.

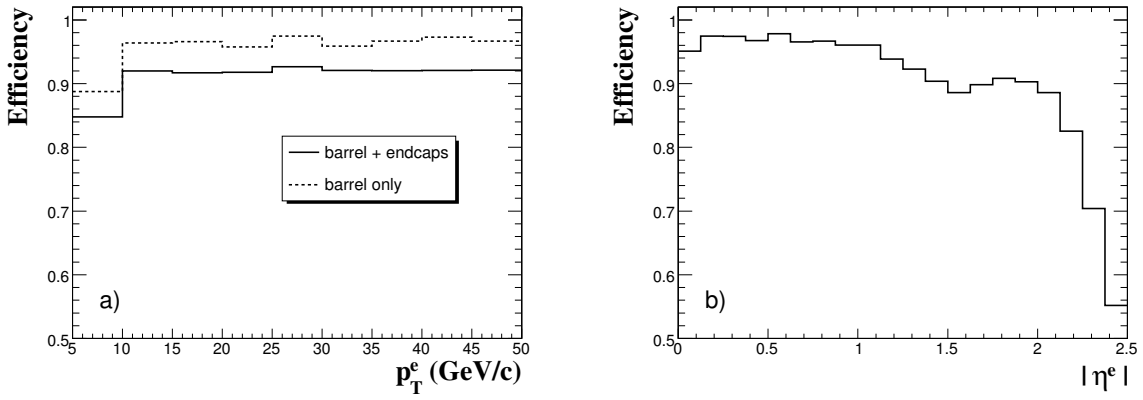


Figure 3: Electron track reconstruction efficiency (a) as a function of p_T and (b) as a function of $|\eta|$, for electrons uniformly distributed in p_T between 5 and 50 GeV/c. In (a), the efficiency is shown averaged over the full ECAL barrel and endcaps η range (full line) and for the barrel only (dotted line).

improve the accuracy of their determination.

A drop in efficiency at low p_T is visible. The algorithm is quite efficient in the full pseudo-rapidity range with a drop at $|\eta^e| = 1.5$ and another one towards $|\eta^e| = 2.4$. The first drop corresponds to the transition region between the ECAL barrel and endcaps and is mainly due to an inefficiency in the reconstruction of superclusters. The second drop is due to the lack of coverage by the pixel endcap disks.

A comparison of the efficiencies obtained with the default KF track finding and with the dedicated KF procedure used in the HLT has been performed in Ref. [11]. The three methods present comparable results at high p_T. At low p_T, an excellent efficiency is obtained for 'e GSF tracks'.

The number of collected hits in the new electron GSF tracks is compared in Figure. 4 with those of the HLT and default KF tracks. The differences arise from the choices of trajectory building parameters and modeling of the energy loss. A very tight χ^2 cut ($\chi^2 < 5$) is used for the hit acceptance in the trajectory building phase of the HLT electron tracks [9]. In this case, the emphasis is put on the initial track parameters and the collection of hits is ended after a significant bremsstrahlung emission. The default (offline) KF requires $\chi^2 < 30$ in the trajectory building and uses a multiple scattering model adapted to reconstruct tracks from charged pions or muons [3, 7]. The new GSF tracks, with the Bethe Heitler energy loss modelling, deals optimally with the effects of Bremsstrahlung losses. The peak of the distribution of the number of collected hits is in the range of 12 to 13, as expected when most of the electron trajectories are followed up to the end of the tracker volume.

It has been checked [12], in the context of photon conversion reconstruction, that electron hits could be also efficiently collected with a Bethe-Heitler energy loss modeling within a KF at the expense of increased χ^2 values in trajectory building steps. This technique, however, would not allow a precision determination of the momentum at vertex, unlike what has been made possible with the use of a GSF (Section 4.3).

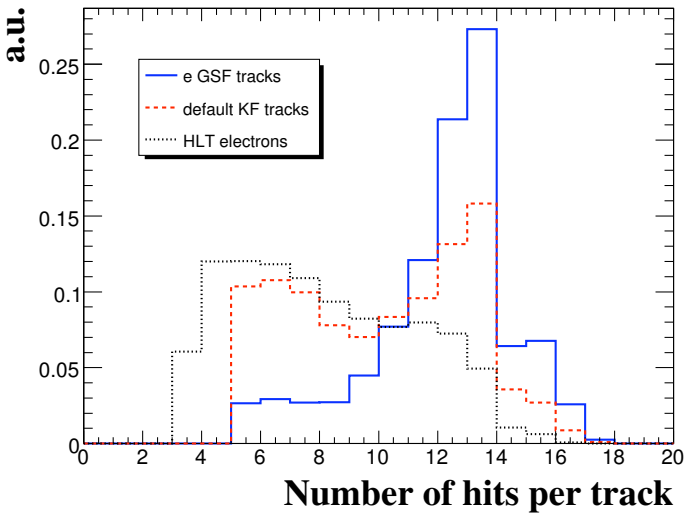


Figure 4: Number of reconstructed hits per track for electrons of $p_T^{\mathrm{e}} = 10$ GeV/c. Distributions are shown for tracks reconstructed with the new method using the Gaussian Sum Filter (full histrogram), Kalman Filter tracks (dashed histrogram) and HLT electron tracks (dash-dotted histogram). The average number of hits collected is related to the cut on the χ^2 increment in the trajectory building steps and modeling of the energy loss, rather than to the choice of the filter. More details are given in the text.

4.3 Track Parameter Measurements

When using the GSF to fit the track, the parameters of all the Gaussian distributions which enter the mixture are available for each hit position. A method to determine the track parameters is, given the track state on each layer, to take the weighted mean of all the components. An alternative way is to rather take only the most probable value (i.e. the mode) of the probability distribution function (PDF), thus giving more importance to the highest weight component. These two methods give quite different results, emphasizing different kind of information. The "weighted mean" method focuses on the average track behaviour and provides best sensitivity to the momentum change along the track due to radiation emission. In contrast, the "mode" method is better suited to obtain an estimation, least affected by bremsstrahlung emission, of the most probable track parameters.

The distributions for the reconstructed transverse momentum at vertex are shown in Fig. 5 for the $p_T^{\mathrm{e}} = 10$ GeV/c case. When taking the mean of the components, a Gaussian distribution with a slightly more pronounced tail towards high p_T values is obtained; the mean residual is compatible with zero. When taking the mode, the distribution is instead well peaked at the correct value but significantly extends at smaller p_T values, a caracteristic of bremsstrahlung losses. Indeed, when a photon is emitted, the track gets more curved than predicted from the most probable value, hence biasing the estimate towards lower p_T values. This behaviour is quite similar to that obtained by tightening the track to follow a non radiating expectation, as done in the HLT electrons procedure. In the following, unless explicitly stated otherwise, the most probable value of the PDF is used to compute the electron track parameters.

4.3.1 Track Parameters at Vertex

In this section, some comparisons between the generated and reconstructed track parameters at vertex are presented.

The difference between the electron pseudorapidity extracted from the momentum at vertex and the generated value is shown in Fig. 6. Also shown in Fig. 6 is the difference between the generated and the reconstructed ϕ. The results are reported for the three track reconstruction algorithms for electrons at $p_T^{\mathrm{e}} = 10$ and 30 GeV/c. The performance of the three algorithms are globally very similar. A slightly better resolution for the ϕ coordinate is observed in the GSF case, in particular at 30 GeV/c.

A comparison of the ratio between reconstructed and generated momentum is shown in Fig. 7 for the three tracking algorithms. The performance of the three algorithms are, here also, globally very similar. However, for both the $p_T^{\mathrm{e}} = 10$ and 30 GeV/c samples, the distributions obtained with the GSF tracks have most probable values

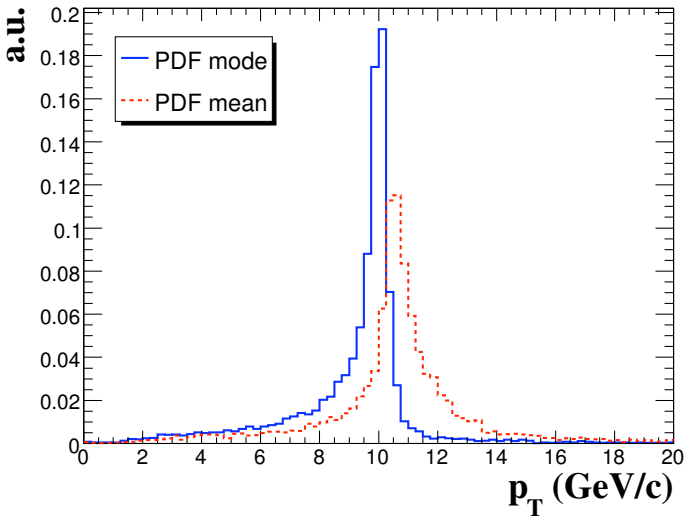


Figure 5: Reconstructed p_T distribution for fixed $p_T^e = 10$ GeV$/c$ electrons, using the most probable value of the Gaussian mixture (full histogram) and using the weighted mean of the Gaussian mixture (dashed histogram) to evaluate the track momentum.

well peaked at 1.0, while the HLT tracks and the standard (default) KF tracks appear slightly biased. The distributions for the HLT and standard KF tracks also show slightly more pronounced tails towards an underestimation of the generated momentum. In contrast, due to the combination of the Gaussians of the mixture, the GSF tracks give slightly larger spread towards an overestimation of the generated momentum. The distributions for the HLT and GSF tracks have very similar full widths at half of their maximum.

In the case of GSF reconstructed tracks, it is expected in addition that the errors provided on the fitted parameters are meaningful. It has been checked that the pull distribution of $1/p$ is indeed unbiased and approximately Gaussian. However, the errors appear to be (uniformly) overestimated and have been rescaled downward to bring the standard width of the pull distribution to 1.0. These errors are used in the combination of tracker momentum and ECAL energy estimates as described in the Section 7.

Finally, a comparison between the new GSF and the default KF reconstructed tracks for the distribution of the transverse impact parameter ($\mathrm{I\!P}_T$) of electrons is shown in Fig. 8. The resolution on $\mathrm{I\!P}_T$ is seen to be sligthly improved with the usage of the new GSF tracks. This parameter is used in the selection of primary electrons (see Section 8).

4.3.2 Track Parameters at the Outermost State

Since the track hits are collected up to the calorimeter, a good estimate of the track parameters at ECAL entrance is possible. This gives the possibility to both improve the matching between the tracker and the calorimeter and estimate the amount of bremsstrahlung radiated by the track using the tracker information only.

Cluster-track matching

Due to the bremsstrahlung emission, the matching between the track and the supercluster is often done using the track parameters at vertex. The track parameters are known with good precision at the initial vertex from the outer-to-inner track fit and, as explained in Section 4.1, the initial track can be matched with the energy weighted average impact point calculated from the supercluster. This, however, does not hold in case of more complicated topologies, as for instance when a radiated photon converts. While the track parameters at the outermost state are known with larger incertainty than those at vertex, they can still be used for the matching.

Figure 9 shows the ratio E/p computed in two ways. In Fig. 9a, the momentum p_{in} evaluated at vertex is used to compare with the supercluster energy E_{sc} ; in Fig. 9b, the momentum p_{out} is taken at the outermost state and the energy E_{seed} is that of the seed of the supercluster. Both distributions exhibit non-Gaussian tails. When using the track parameters at vertex (Fig. 9a), a non-Gaussian tail is seen towards $E/p < 1$. This tail receives contributions both from overestimates of the electron momenta (by the track measurement) and underestimates of the electron energy (by the supercluster). In contrast, when using the track parameters at the outermost state, a non-Gaussian

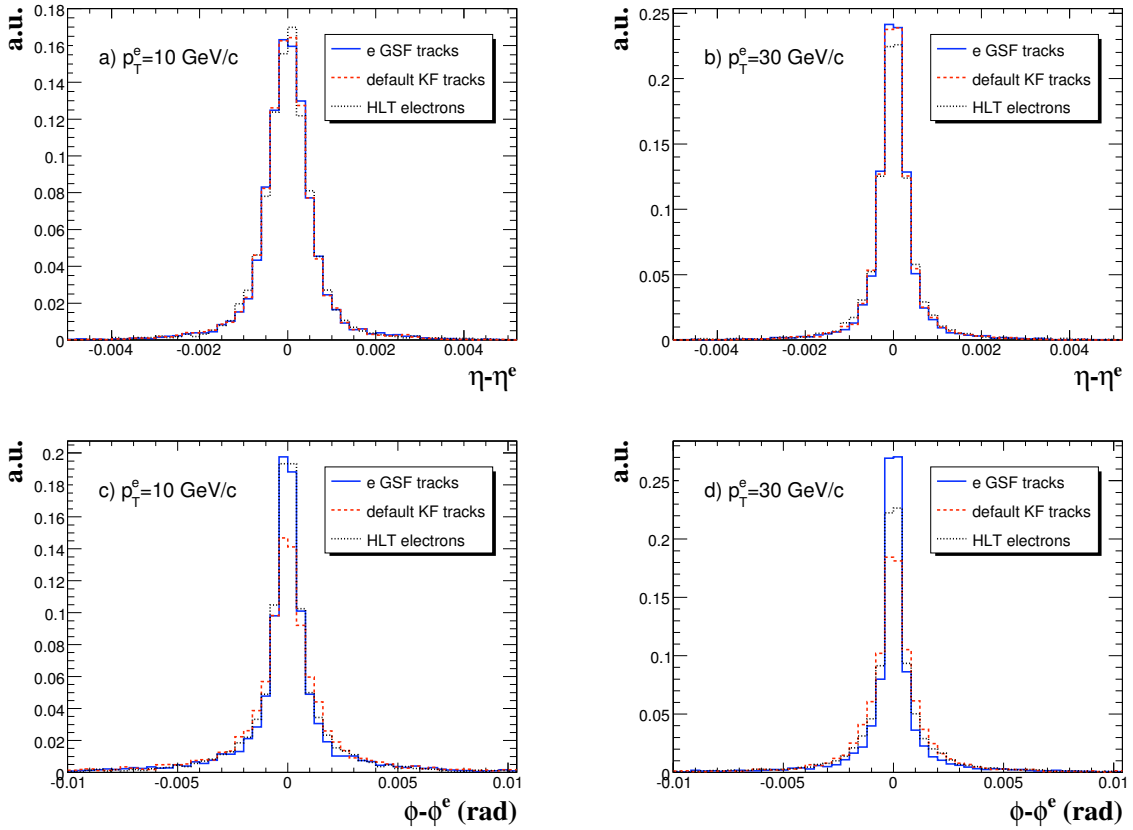


Figure 6: Difference between the reconstructed track direction at vertex and the generated one: a) η direction and c) ϕ direction for $p_T^e = 10$ GeV/c electrons; b) η direction and d) ϕ direction for $p_T^e = 30$ GeV/c electrons. Distributions are shown for Gaussian Sum Filter tracks (full histrogram), Kalman Filter tracks (dashed histrogram) and HLT electron tracks (dash-dotted histogram).

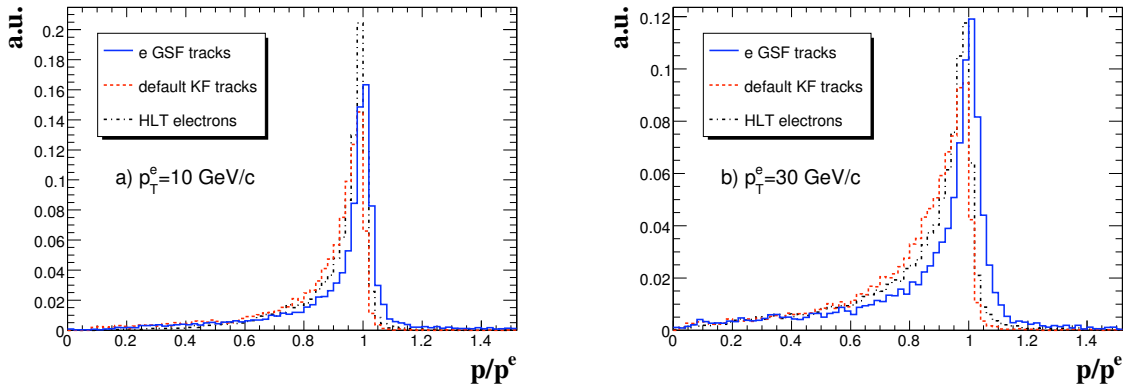


Figure 7: Ratio between the reconstructed momentum at vertex and the generated one for (a) $p_T^e = 10$ GeV/c electrons and (b) $p_T^e = 30$ GeV/c electrons. Distributions are shown for Gaussian Sum Filter tracks (full histrogram), Kalman Filter tracks (dashed histrogram) and HLT electron tracks (dash-dotted histogram).

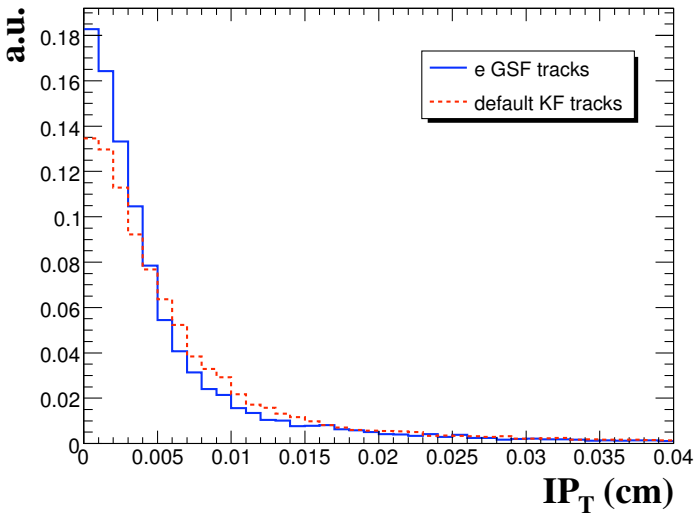


Figure 8: Comparison between the transverse impact parameter distributions for $p_T^{\mathrm{e}} = 10\ \mathrm{GeV}/c$ electrons between GSF tracks (Solid histogram) and KF tracks (dashed histogram).

tail is seen towards $E/p > 1$. This is mainly due to events in which a significant fraction of the initial electron momentum has been radiated, such that $p_{\mathrm{in}}/p_{\mathrm{out}} > 1$, but where the photons nevertheless appear mostly collected in the supercluster seed.

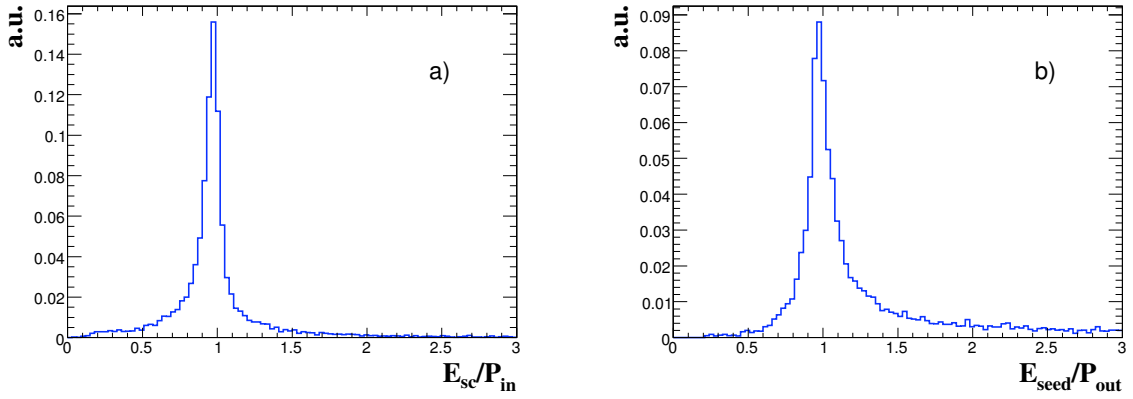


Figure 9: Distributions of E/p for $p_T^{\mathrm{e}} = 10\ \mathrm{GeV}/c$ electrons; a) supercluster energy over track momentum measured at vertex; b) supercluster seed energy over track momentum estimated at the last hit.

The bremsstrahlung estimate

The knowledge of the track momentum at the outermost state gives the possibility of estimating from the track fit the fraction of energy lost by bremsstrahlung. The difference between the magnitude of the momentum at vertex and at the layer of the outermost hit is well visible in Fig. 10a. The track momenta are obtained here via the weighted mean of the GSF which provides a good sensitivity to the momentum evolution along the track. The difference between the momentum magnitude at vertex and at the last hit, $p_{\mathrm{in}} - p_{\mathrm{out}}$, is a measure of the integral amount of bremsstrahlung. Figure 10b shows this difference of momentum measured at both track ends versus the generated energy sum of the emitted photons. A strong correlation is observed. The mean fraction of the energy radiated along the complete trajectory is roughly proportional to the integral amount of material traversed which varies strongly with η. The relative difference between the momenta measured at both track ends, $f_{\mathrm{brem}} = (p_{\mathrm{in}} - p_{\mathrm{out}})/p_{\mathrm{in}}$, is linearly dependent on the radiated energy fraction, whether the momenta are determined with the mode or with the mean of the GSF. This quantity is therefore used in the definition of electron classes (Section 5).

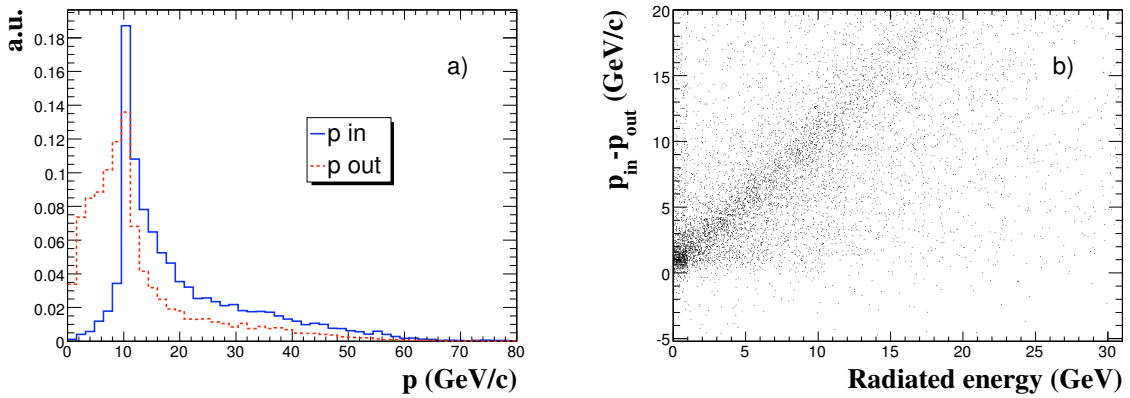


Figure 10: Estimation of the bremsstrahlung radiated energy using the GSF track parameters, for electrons with $p_T^e = 10$ GeV/c; a) track momentum evaluated at the vertex (full histogram) and at the outermost hit (dashed-dotted histogram); b) difference between track momentum at vertex and at outermost state against the energy radiated via bremsstrahlung photons along the electron trajectory. The momentum is obtained from the weighted mean of the Gaussian mixture.

5 Electron Classification

More elaborate reconstruction procedures, involving recognition of distinct track-supercluster patterns, are necessary to distinguish between "well measured" and "badly" measured electrons for later use in physics analyses [13]. This is particularly true towards low p_T^e. Distinct patterns imply, in general, different energy-momentum measurement errors and different electron identification performance.

In addition to the problem of bremsstrahlung collection, the measurement of primary electron energies is affected by energy lost in the tracker material, as illustrated for example in Fig. 11 for electrons of $E^e = 10$, 30 and 50 GeV. As expected the mean energy lost increases with η, that is together with the increasing amount of tracker

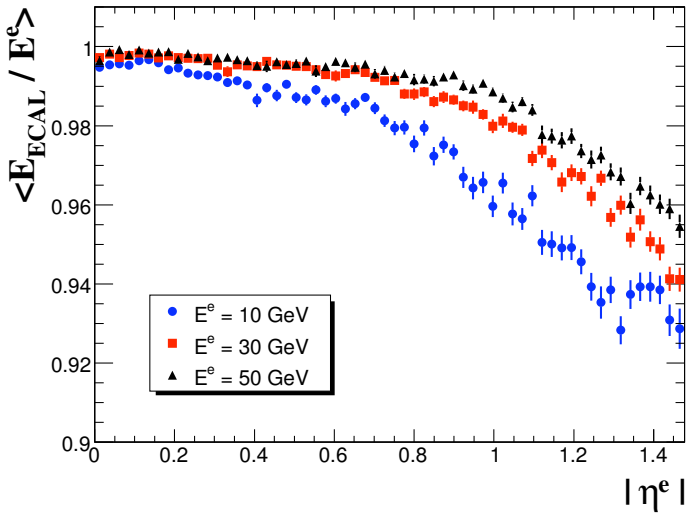


Figure 11: The true energy reaching the ECAL barrel front face normalized to the initial energy as a function of the pseudorapidity η for electrons of 10, 30 and 50 GeV energy. To compute the true energy reaching the ECAL, only photons with energy greater than 10 MeV are considered.

material. On average the energy lost amounts to 7% of the initial energy for 10 GeV electrons heading towards the edge of the barrel. This energy lost comes from e^+e^- pairs from bremsstrahlung photon conversion. In particular, soft secondary electrons get partly, or completely, trapped in the magnetic field and lose most of their energy before reaching the ECAL, or migrate away from the main supercluster. A significant amount of energy lost in this way

is expected for electron measurement patterns corresponding to the first stages of an "electromagnetic showering" in the tracking material.

The effects induced on the energy measurement of single electrons by the bremsstrahlung emission, the partial supercluster energy containment and the energy lost in front of the ECAL are subject to large event-by-event fluctuations.

The energy radiated by a primary electron when traversing each discrete silicon layer fluctuates considerably. The ϕ spread of the energy deposits in the ECAL in general depend on the pattern of bremsstrahlung emission along the electron trajectory and on $p_T^{\rm e}$. The bremsstrahlung photon spectrum has roughly a $1/E^{\gamma}$ dependence leading, for $E^{\gamma} \ll E^{\rm e}$, to an approximately equal number of photons emitted per energy decades. The vast majority of bremsstrahlung photons are low energy photons and populate a continuous band of crystals in ϕ. The fluctuations in the ϕ spread of the energy deposits and partial collection in the supercluster can affect the electron energy resolution.

On average, a handful of hard bremsstrahlung photons are emitted which carry more the 1% of the primary electron energy. Especially for low $p_T^{\rm e}$ electrons given the small curvature radii of their trajectory, a hard bremsstrahlung photon might deposit its energy far enough from the main electron seed cluster to form a separate cluster. This can lead to additional pattern-dependent fluctuations of the supercluster energy containment.

Photons emitted in the innermost silicon layers (or in η regions with more tracker material) are more likely to convert (and thus contribute to energy lost) before reaching the ECAL than photons emitted late (or in η regions with less tracker material). This introduces additional energy measurement fluctuations depending on the bremsstrahlung emission pattern. The fluctuations are expected to be largest in η regions where the mean energy lost is also the largest. Although a correlation of the energy lost with the number of tracker hits in the vicinity of the electron track could be observed [13], an event-by-event estimate is essentially hopeless in presence of low energy secondary electrons curling in the magnetic field. Thus, electrons with showering-like patterns in the tracker material require special η-dependent mean energy scale correction factors. The event-by-event fluctuations of the pattern of bremsstrahlung emission also introduces non-Gaussian fluctuations of the momentum measurement at primary vertex. An early emission of a large amount of radiation is more likely to lead to an underestimation of the initial momentum by a track reconstruction algorithm.

Observables sensitive to the amount of bremsstrahlung radiated along the electron trajectory and to the pattern of photon emission and conversions are therefore used to classify electrons. The main observables used are the "measured bremsstrahlung fraction" $f_{\rm brem}$ introduced in Section 4.3, the ϕ match between the reconstructed track and the supercluster which is is sensitive to the bremsstrahlung collection, the matching between the total energy collected by the supercluster with the momentum measured at the track origin which is sensitive to the energy lost in the tracker material.

These and the pattern of clusters in the superclusters are used to separate electrons in the different classes. The cases where an electron is impacting in the immediate vicinity (less than about half the width in η of a crystal) of ECAL inter-module borders are considered separately. In such cases, a significant fraction of the electron shower can leak behind the ECAL. Whether or not an electron depositing energy near a border can be measured with precision, is to be evaluated in principle by an algorithmic procedure described elsewhere [9]. Here, for simplicity, fiducial regions are defined to group such electrons with those impacting in the transition region $1.4442 < |\eta_{\rm sc}| < 1.5660$ between the ECAL barrel and endcaps [9]. These are kept in a separate *Boundary* class. Such electrons are not further studied in this note and do not enter the following four classes.

- *Golden* electrons. This class represents electrons least affected by radiation emission, with a reconstructed track well matching the supercluster and a well behaved supercluster pattern. It is defined as
  - a supercluster formed by a single cluster (i.e. without observed bremsstrahlung sub-cluster);
  - a measured bremsstrahlung fraction $f_{\rm brem}$ below 0.2;
  - a ϕ matching between the supercluster position and the track extrapolation from last point within $\pm$ 0.15 rad;
  - an $E_{\rm sc}/p_{\rm in}$ value in excess of 0.9.

The Golden electrons are predominantly truly "low radiating" electrons. Electrons which fail to satisfy the Golden electron requirements might be accepted in the following class.

- *Big Brem* electrons. This class contains electrons with a good matching between $E_{\rm sc}$ and $p_{\rm in}$, a well behaved supercluster pattern, and no evidence of energy loss effects from secondary photon conversion despite a very large measured bremsstrahlung fraction. Electrons for which all the bremsstrahlung is radiated in a single step, either very early or very late when crossing tracker silicon layers, can fall in this category. The class is defined as

  - a supercluster formed by a single "seed" cluster;
  - a measured bremsstrahlung fraction above 0.5;
  - an $E_{\rm sc}/p_{\rm in}$ value between 0.9 and 1.1.

A complementary set of electrons, still with a good energy-momentum (at origin) matching, but which fails some criteria for Golden and Big Brem can fall in the following class.

- *Narrow* electrons. In this intermediate class, electrons have a significantly large bremsstrahlung fraction but not has high as for Big Brem, a rather well behaved supercluster (i.e. the bremsstrahlung photons are merged inside the single cluster), but, as for Big Brem, a relaxed track-supercluster geometrical matching. It is defined as
  - a supercluster formed by a single "seed" cluster;
  - an $E_{\rm sc}/p_{\rm in}$ value between 0.9 and 1.1;
  - a measured bremsstrahlung fraction and/or a ϕ matching outside the range of Golden and Big Brem electrons.

Finally, the remaining ("bad") electrons are classified in a fourth class.

- *Showering* electrons. This class contains electrons which failed to enter any of the above classes. It includes electron supercluster patterns involving one or several identified bremsstrahlung sub-cluster(s), or cases where a bad energy-momentum E/p matching is observed. This bad matching is very likely, for instance, in cases of secondary conversion of some early radiated bremsstrahlung for electrons having radiated a large fraction of their initial energy.

Figure 12 presents the raw reconstructed energies, before having applied any corrections, for the different classes in the ECAL barrel and endcaps, and for electron energies between 5 and 100 GeV. Most of the tail in the reconstructed energy spectrum is localized in the showering class of electrons. In particular, the Big Brem and Narrow electron cases appear rather well measured, indicating that the radiated energy is properly collected by the clustering.

The population of the different classes as a function of the η is shown in Fig. 13. The shape of the distribution for the Showering class clearly reflects the η distribution of the material thickness. On the contrary, the observed distribution for the Golden electrons is, as expected, anti-correlated with the material thickness. The small sub-structures observed for instance in the distribution for the population of Narrow electrons at $|\eta| \sim 2.35$ are also a direct consequence of sub-structures in the tracker material distribution [5, 14].

6 Energy Scale Corrections

The lack of containment in cluster reconstructed energy can be corrected for as a function of the measured number of crystals which make up the cluster volume, as used in HLT reconstruction [10]. Figure 14 shows the normalized energy response as function of the number of crystals $N_{\rm cry}$ in the seed cluster of the supercluster which have energy above two standard deviation of the electronic noise, for different ranges of electron initial energy. Electrons from the Golden, Big Brem and Narrow classes as well as from the Showering class, in cases where the supercluster is made of a single seed cluster, are used in Fig. 14. The dependence on $N_{\rm cry}$ can be very well described by a universal $f(N_{\rm cry})$ containment correction function.

This is found to remain true in the case of electrons from the Showering class, when considering $N_{\rm cry}$ from the seed cluster of superclusters which include bremsstrahlung clusters. In contrast, the total number of crystals in the superclusters does not follow a universal function in such cases. Nevertheless,to avoid possible bias, the Showering cases are not considered in the evaluation of the unviserval correction function. The $f(N_{\rm cry})$ correction functions in the endcaps appears to be different from that used in the barrel as a consequence of the different clustering algorithm used.

The reconstructed energy $E_{\rm rec}$, corrected via $f(N_{\rm cry})$ and divided by the generated energy $E^{\rm e}$, is plotted in Fig. 15 as a function of η for different electron classes. The results obtained for the electrons belonging to the Golden, Big Brem and Narrow classes have been found to nicely overlap, such that these classes are grouped together in Fig. 15. The dependence on η of $E_{\rm rec}/E^{\rm e}$ is found to be very much attenuated for these three classes when compared to that for showering electrons. A variation of more than 0.5% remains only in a band of about 0.7 unit in η centred on the transition region between the ECAL barrel and endcaps. The residual η dependence for the different classes in the barrel and endcaps range are parametrized as shown in Fig. 15.

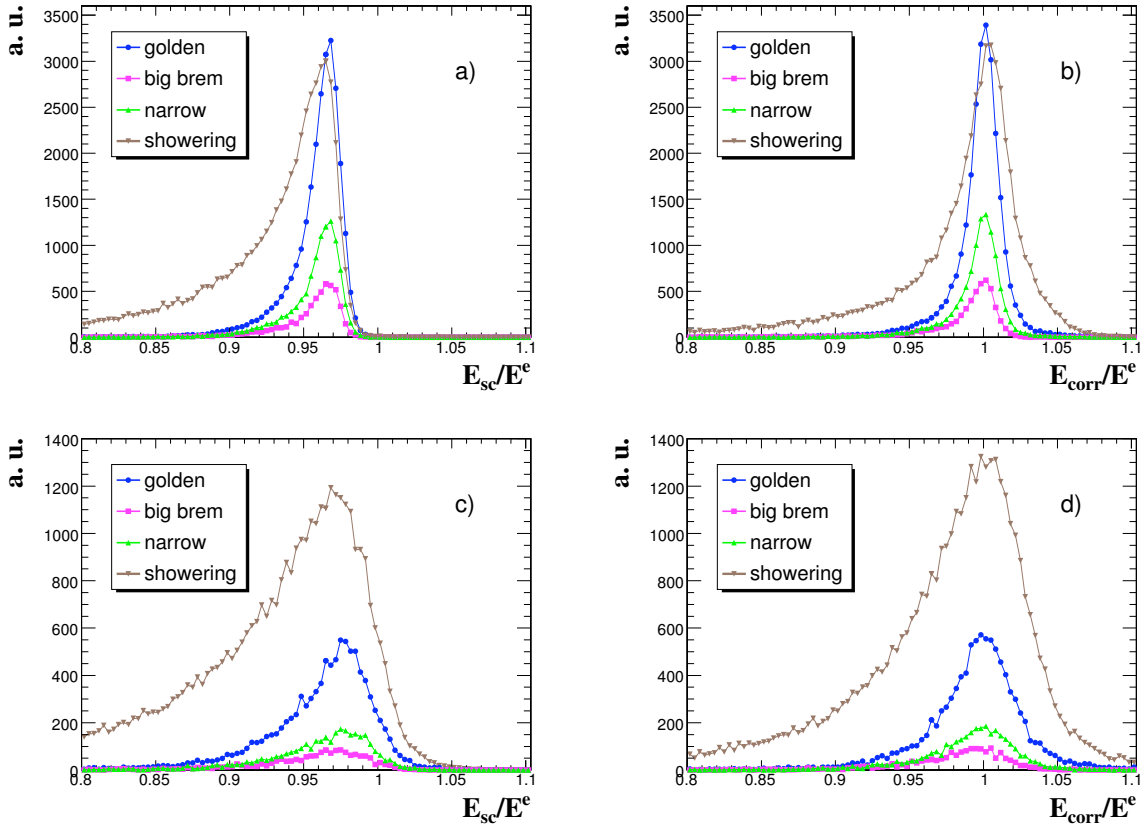


Figure 12: The distribution of the raw and fully corrected reconstructed electron energy normalized to the generated energy, for the different electron classes. Distributions obtained for uniformly distributed electrons of energies between 5 and 100 GeV are shown separately for the η range of the (a,b) barrel and (c,d) endcaps. The distributions are obtained (a,c) before ($E_{\rm sc}$) and (b,d) after ($E_{\rm corr}$) application of all energy scale corrections (as explained in the text). For the endcaps, $E_{\rm sc}$ and $E_{\rm corr}$ includes the energy deposited in the preshower.

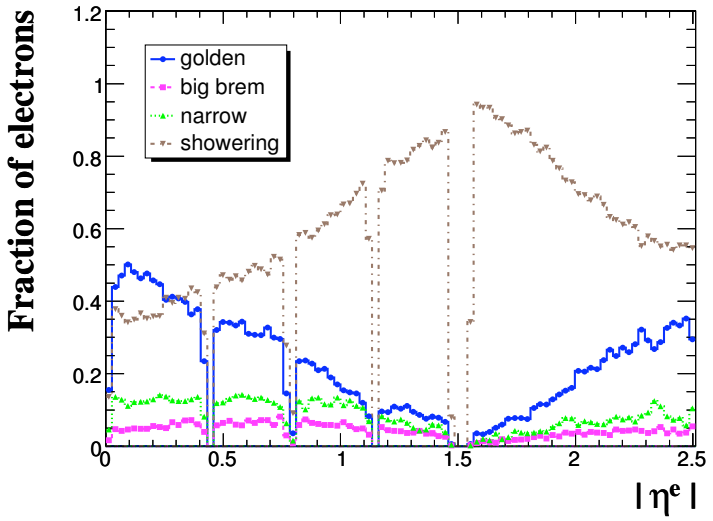


Figure 13: The fraction of electron population in the different classes as a function of the generated pseudorapidity for initial energies between 5 and 100 GeV.

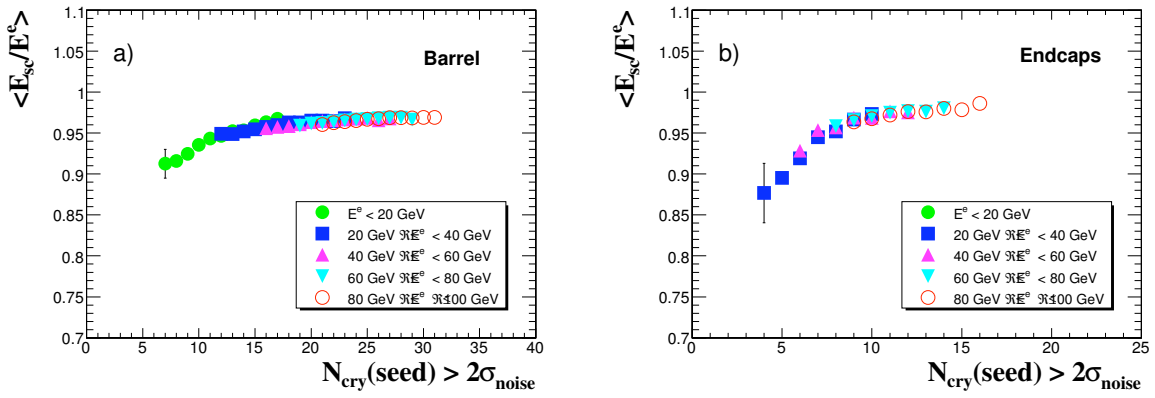


Figure 14: Mean of the raw reconstructed supercluster energy E_{sc} divided by the generator electron energy E^{e} as a function of the number of crystals in the seed cluster that contain an energy greater than $2\sigma_{\mathrm{noise}}$: a) electrons in the ECAL barrel; b) electrons in the ECAL endcaps. In the η range of the ECAL endcaps, E^{e} is taken as the generated initial energy minus the energy measured in the preshower detector. Only electrons with $p_T > 5$ GeV/c and with no identified bremsstrahlung cluster are considered.

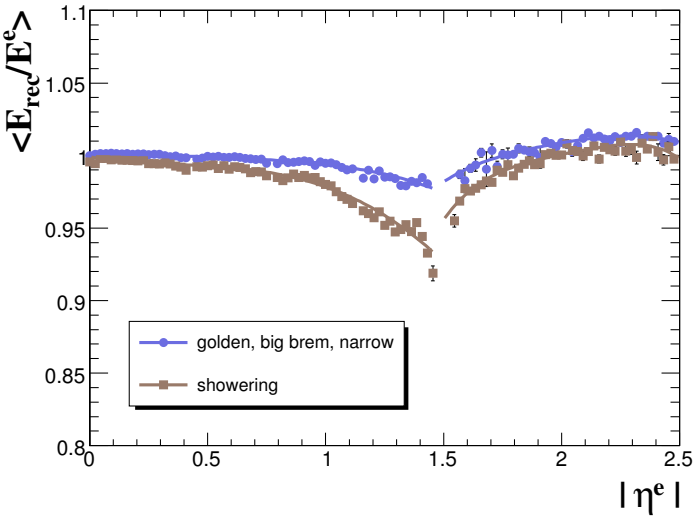


Figure 15: Mean of the reconstructed supercluster energy E_{rec} after $f(N_{\mathrm{cry}})$ correction divided by the generator electron energy E^{e} for different classes and as a function of the pseudorapidity. In the η range of the ECAL endcaps, E^{e} is taken as the generated initial energy minus the energy measured in the preshower detector.

In the case of the endcaps, the energy scale corrections are applied on the clusterized crystal energies. The energy deposited in the preshower is added afterward to the corrected supercluster energy, using a pre-determined intercalibration factor, to obtain the final estimator of the electron energy.

Figure 12 presents the distributions of the reconstructed energies after having applied the corrections described above, for the different classes for the barrel and for the endcaps and for electron energies between 5 and 100 GeV. In addition to the normalization of the response, the overall corrected distribution is narrower, more symmetric and more Gaussian than the uncorrected distribution. A computation of the effective RMS (σ_{eff}), defined as the half width of the smallest window that contains 68.3% of the distribution, shows a gain in the resolution of 5% in the barrel and of 10% in the endcaps over the whole 5 to 100 GeV energy range. A more significant gain in resolution is observed when considering only the low energy part of the spectrum. Finally, such algorithmic corrections are expected to be ultimately tuned on real data e.g. using $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ decays.

7 E-p Combination

The corrected energy measurement E_{corr} provided by electromagnetic calorimeter can be combined with the tracker momentum measurement $p = p_{\text{in}}$ to improve the estimation of the electron momentum at the interaction vertex. In particular at energies of around 15 GeV and below the tracker momentum estimate is more precise than the ECAL energy measurement so that a combination improves the electron energy estimate. Moreover, these two measurements appear to be complementary in the way they are affected by bremsstrahlung radiation in the tracker material, so that the most appropriate measurement can be used depending on the electron class.

In order to build a combined estimate of the electron momentum based on ECAL and tracker measurements, it is useful first to analyse these measurements as a function of a variable affected by the amount of bremsstrahlung radiation.

Figure 16 presents the ratios $E_{\text{corr}}/E^{\text{e}}$ and p/p^{e} as a function of E_{corr}/p for the barrel case. A similar behaviour

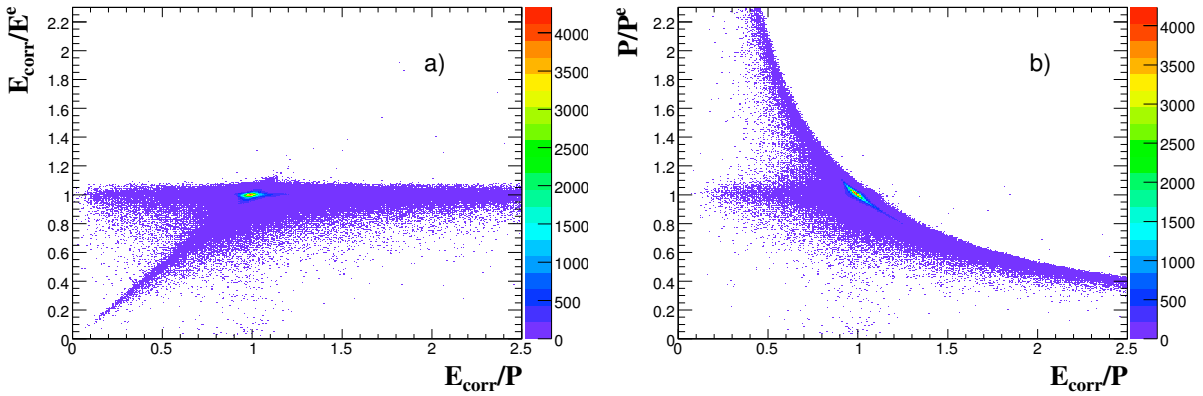


Figure 16: The momentum estimate from the ECAL and the tracker as a function of E/p for electrons in the ECAL barrel: a) corrected supercluster energy E_{corr} normalized to the initial electron energy as a function of E_{corr}/p; b) reconstructed track momentum at origin p normalized to the initial electron energy as a function of E_{corr}/p.

is found for the endcaps. It can be seen that:

- the energy and momentum measurements are very seldom both wrong, under or over estimating the generated energy, since when E_{corr}/p is $\simeq 1$, both are in good agreement with the generated value;
- cases with $E_{\text{corr}}/p > 1$ are almost always due to a momentum underestimate;
- cases with $E_{\text{corr}}/p < 1$ can be due either to an incorrect energy measurement or to a wrong momentum estimate; most of these cases correspond to Showering electrons.

The combination of the two estimates involves the determination of errors. The track errors are discussed in Section 4.3.1. For the error on the supercluster energy, a parametrisation of the resolution for each classes is used.

Figure 17 shows for example the energy resolution obtained as a function of the electron energy for the different classes, in the η range of the ECAL barrel. The resolutions are presented both as obtained from a fit of the Gaussian part of the distributions for each energy bin and as obtained by computing an effective RMS. Golden and Narrow electrons are measured with the best precision.

In the case of the Golden electrons, the energy resolution fit function follows the simple characteristic functional form expected for homegeneous calorimetry, namely $\sigma/E = a/\sqrt{E} \oplus b/E \oplus c$ with a, b and c as free parameters. The fit is shown on Fig. 17 and gives parameters values $a = (2.5 \pm 0.3)\%$, $b = (197 \pm 7)$ MeV and $c = (0.52 \pm 0.01)\%$ compatible with those obtained in Test Beam [8]. The Showering electrons show a deterioration of the asymptotic σ/E towards large energy.

Using these errors on the electron energy E_{corr} from calorimetry and the track momentum p at origin from the tracker, and given the above considerations on their respective sensitivity to bremsstrahlung induced effects, the final electron momentum magnitude is therefore estimated as:

- the weighted mean of E_{corr} and p when $|E_{\text{corr}}/p - 1| < 2\sigma_{E/p}$ with weights defined as the normalized inverse of the variance of each measurement;

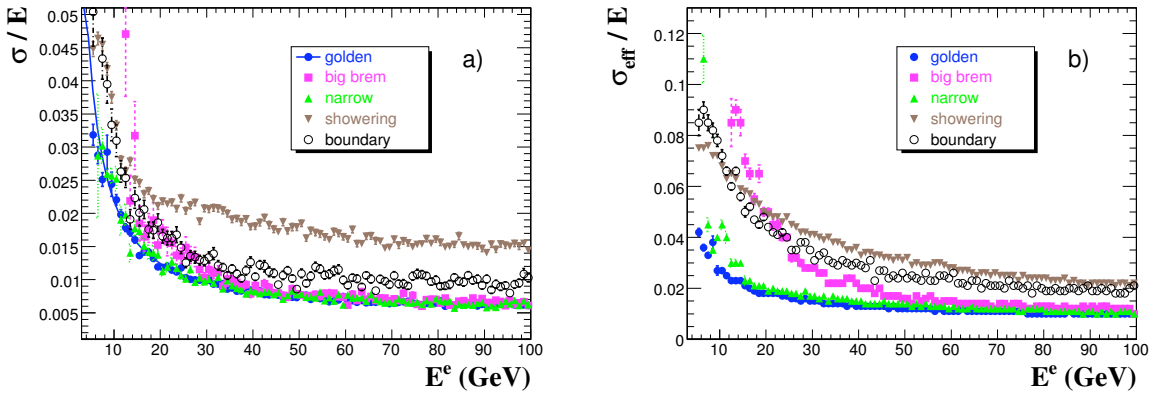


Figure 17: The energy measurement precision as a function of the generated electron energy for the different classes and for electrons in the barrel: a) resolution as obtained from a fit of the Gaussian part of the corrected energy distribution; b) resolution as obtained computing the half width of the smallest window that contains 68.3% of the distribution. The line corresponds to a fit for the golden electrons using a standard calorimeter resolution function, as described in the text.

in cases where $E_{\rm corr}$ and p disagree significantly, the detector offering a priori the best and least biased measurement is used, i.e.

- the energy $E_{\rm corr}$ for $E_{\rm corr} > 15$ GeV, and the momentum p for $E_{\rm corr} < 15$ GeV, when $|E_{\rm corr}/p - 1| > 2\sigma_{E/p}$, where $\sigma_{E/p}$ is computed as the quadratic sum of the $E_{\rm corr}$ and p uncertainties;
- the energy $E_{\rm corr}$ for electrons of the Showering class when $E_{\rm corr}/p < 1 - 2\sigma_{E/p}$.

Figure 18 presents the effective RMS of the combined estimate, together with that of the ECAL and tracker measurements alone. The precision is clearly improved by using the combined estimate with respect to the ECAL only measurement for energies below $\simeq 25$ GeV. At 15 GeV, a factor of about 1.4 in precision is achieved by combining with the tracker measurement.

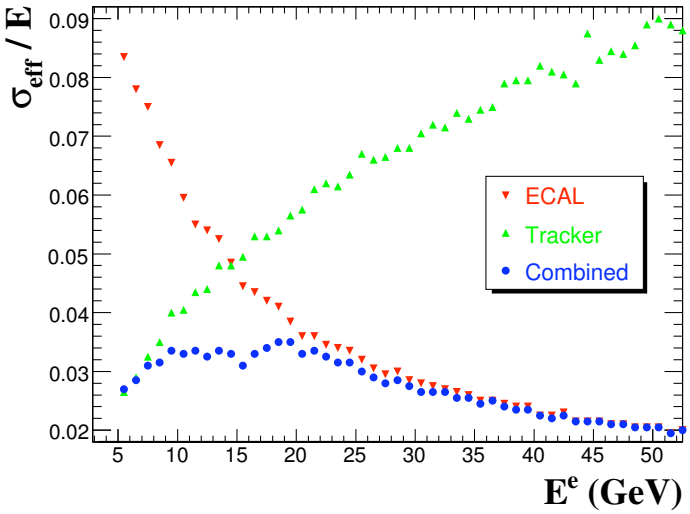


Figure 18: The resolutions as measured by the relative effective RMS of the corrected supercluster energy (downward triangles), the reconstructed track momentum at origin (upward triangles), and of the combined electron momentum estimate (circles) as a function of the electron incident energy for electrons in the ECAL barrel.

8 Electron Isolation and Selection of Primary Electrons

In a "pre-selection" step, a loose geometrical and energy-momentum matching is imposed between the reconstructed electron track and the corresponding supercluster.

Good energy-momentum matching is obtained by matching the corrected supercluster energy $E_{\rm corr}$ with the track momentum $p_{\rm in}$ taken at the closest position from the nominal vertex. The geometrical matching is performed taking the track parameters at interaction vertex, $\eta_{\rm in}$ and $\phi_{\rm in}$, extrapolating to the ECAL assuming a perfect helix, and matching the resulting $\eta_{\rm in}^{\rm extrap.}$ and $\phi_{\rm in}^{\rm extrap.}$ to the energy-weighted position of the supercluster, $\eta_{\rm sc}$ and $\phi_{\rm sc}$. The compatibility of the matched object with the expectations from an electron is reinforced by setting an upper threshold on the fraction of the supercluster energy collected in the hadron calorimeter (HCAL). Electron candidates are therefore defined as:

- a reconstructed electron track initiated by the reconstruction of a supercluster in the ECAL matched with hits in the pixel detector;
- an energy-momentum matching between the supercluster and the track, $E_{\rm corr}/p_{\rm in} < 3$;
- an η geometrical matching $|\Delta\eta_{\rm in}| = |\eta_{\rm sc} - \eta_{\rm in}^{\rm extrap.}| < 0.02$;
- a ϕ geometrical matching $|\Delta\phi_{\rm in}| = |\phi_{\rm sc} - \phi_{\rm in}^{\rm extrap.}| < 0.1$;
- a ratio of the energy deposited in the HCAL tower just behind the electromagnetic seed cluster over the energy of the seed cluster $H/E < 0.2$.

Figure 19 shows the absolute efficiency of such electron candidate definition as a function of p_T and η as measured on electrons from Higgs boson decays $\mathrm{H} \to \mathrm{ZZ}^* \to \mathrm{e}^+\mathrm{e}^-\mathrm{e}^+\mathrm{e}^-$ for $m_{\rm H} = 150$ GeV/c^2. The efficiency

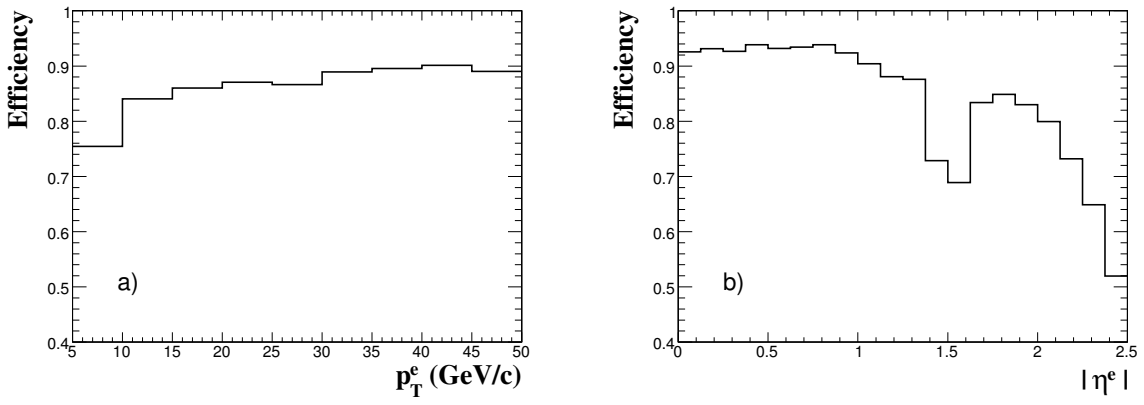


Figure 19: Electron candidate efficiency for electrons from Higgs boson decays $\mathrm{H} \to \mathrm{ZZ}^* \to \mathrm{e}^+\mathrm{e}^-\mathrm{e}^+\mathrm{e}^-$ after preselection and for $m_{\rm H} = 150$ GeV/c^2: a) as a function of $p_T^{\rm e}$; b) as a function of $\eta^{\rm e}$.

is defined as the fraction of generated electrons from the Higgs boson decay which have a matching (Section 4.2) reconstructed track.

The selection of primary electrons is based on a further requirement on the impact parameter with respect to the nominal vertex. As the beam spot is much better defined in the transverse plane, the normalized transverse impact parameter ($\mathrm{I\!P}_T/\sigma_{\mathrm{IP}_T}$) is used to select primary electrons. Figure 20 shows the distributions of the normalized transverse impact parameter for electrons from the SM Higgs boson decaying in $\mathrm{H} \to \mathrm{ZZ}^* \to \mathrm{e}^+\mathrm{e}^-\mathrm{e}^+\mathrm{e}^-$ and for electrons from the corresponding backgrounds, of which $\mathrm{t\bar{t}}$ and $\mathrm{Zb\bar{b}}$ involve electrons from semi leptonic b decays. The LHC pile-up for low luminosity (2×10^{33} cm^{-2}s^{-1}) is included.

Lepton isolation cone requirements can be imposed as a simple and powerful means of suppressing QCD background in multi-lepton physics channels at the LHC. For electrons in CMS, the simplest and most powerful isolation criterion is obtained from tracks originating from a common (primary) vertex. Using track measurements at a primary vertex for the electron isolation avoids the complication due to the severe external bremsstrahlung, photon conversion, and early showering in the tracker material. It moreover allows the question of the identification of possible internal bremsstrahlung photons (associated to final state electrons), which might appear in the ECAL within an isolation cone, to be postponed. Track-based electron isolation must normally be complemented by electron identification requirements.

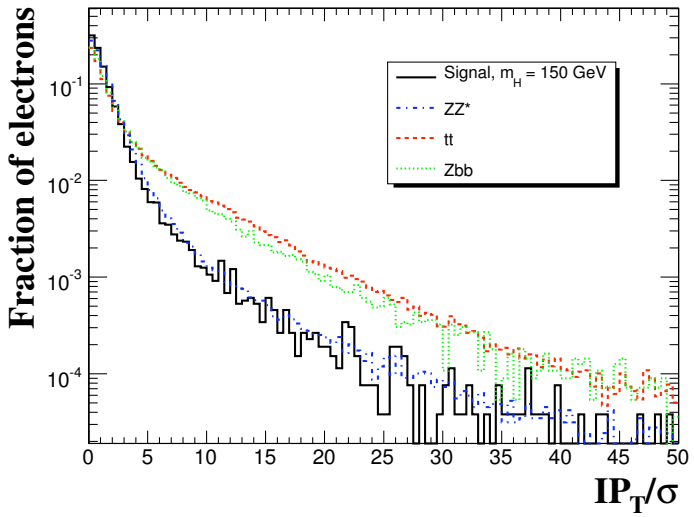


Figure 20: The distributions of the normalized transverse impact parameter $\mathrm{IP}_T/\sigma_{\mathrm{IP}_T}$ for preselected electrons from the SM Higgs boson decaying in $\mathrm{H} \to \mathrm{ZZ}^* \to \mathrm{e}^+\mathrm{e}^-\mathrm{e}^+\mathrm{e}^-$ ($m_{\mathrm{H}} = 150$ GeV/c^2) and for preselected electrons from the ZZ^*, $\mathrm{t\bar{t}}$ and $\mathrm{Zb\bar{b}}$ backgrounds.

The performance of track-based electron isolation are illustrated here by considering the example of the suppression of the $\mathrm{t\bar{t}}$ background to the Higgs boson in the channel $\mathrm{H} \to \mathrm{ZZ}^* \to \mathrm{e}^+\mathrm{e}^-\mathrm{e}^+\mathrm{e}^-$. Reconstructed tracks are considered within an isolation cone in the (η,ϕ) plane of radius $R_{\mathrm{cone}} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ centred on each electron. The tracks are required to have $p_T > 1.5$ GeV/c and $|\Delta \mathrm{IP}_L| < 0.1$ cm, where $\Delta \mathrm{IP}_L$ is the difference between the longitudinal impact parameter and the z position of the primary vertex. The electron isolation variable is then defined as the sum of the p_T of all the tracks satisfying these requirements but the electron one, divided by the electron p_T. The event isolation is finally defined as the request of having all the four electrons from the Higgs boson decay isolated. In order to present purely algorithmic isolation efficiency, the electrons are required to match real Monte Carlo electrons from the Higgs boson decay.

Figure 21 presents the track based isolation efficiency for the Higgs boson signal in the channel $\mathrm{H} \to \mathrm{ZZ}^* \to \mathrm{e}^+\mathrm{e}^-\mathrm{e}^+\mathrm{e}^-$ after preselection as a function of the rejection obtained against the $\mathrm{t\bar{t}}$ background. The efficiency on the Higgs boson signal as a function of the isolation variable and for different cone sizes is also shown. Here also, the LHC pile-up for low luminosity is included.

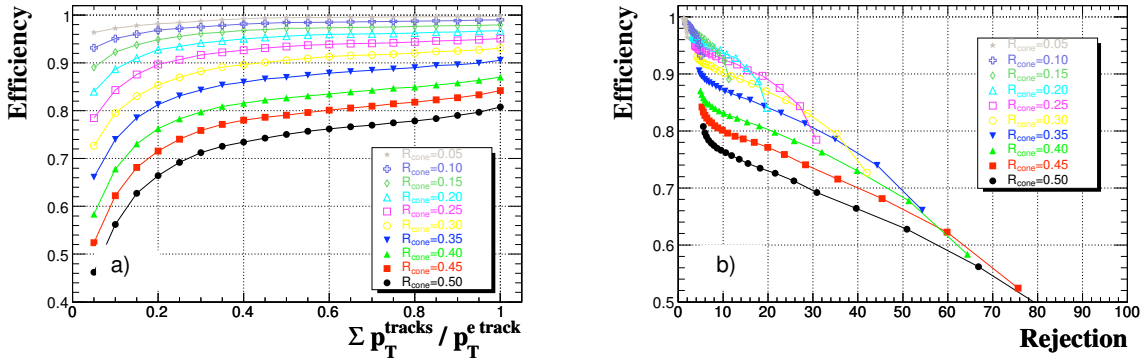


Figure 21: Performance of track isolation in the channel $\mathrm{H} \to \mathrm{ZZ}^* \to \mathrm{e}^+\mathrm{e}^-\mathrm{e}^+\mathrm{e}^-$ after the preselection described in the text: a) efficiency on the Higgs boson signal as a function of the isolation variable and for different cone sizes; b) efficiency for the Higgs boson signal as a function of the rejection obtained against the $\mathrm{t\bar{t}}$ background.

9 Electron Identification

Electron identification makes use of a complete set of estimators. These estimators are combined to establish full compatibility of the observations with expectations from single electrons. The performance (efficiency, rejection power, purity) of this identification depends of course on the degree of isolation imposed on the electron candidates and on the nature of the considered background. It also depends on the quality requirements imposed on the electron objects themselves. In general, the "well measured" and the "badly measured" electrons are likely to be differently affected by possible fake background sources. Finally, the distinction between multi-clusters and single cluster electron patterns is expected to play an important role in the separation of electron from "fake" electrons in QCD jets formed by overlapping particles.

A sample of di-jet QCD events is used here to evaluate the "fake" background and electron identification capabilities using the outer GSF track parameters and the electron classification. The events considered have been generated in two $\hat{p}_T$ bins, $25 < \hat{p}_T < 50\,\mathrm{GeV}/c$ and $50 < \hat{p}_T < 170\,\mathrm{GeV}/c$. A filtering has been applied at generator level to enrich the sample in events which would pass the Level 1 electromagnetic trigger of CMS. Events from the two $\hat{p}_T$ bins are expected to contribute very similar trigger rates. This constitutes what is thereinafter called the QCD jet background.

After pre-selection and the application of a loose track isolation ($\sum_{\mathrm{cone}} p_T^{\mathrm{tracks}}/p_T^{\mathrm{e\,track}} < 0.5$ and $R_{\mathrm{cone}} = 0.35$), the "fake" electron background is significantly reduced. A jet from the QCD jet background is found to have a probability of 18.5 % to give a "fake" reconstructed electron with p_T in the range from 5 to 50 GeV/c. This background is constituted by e.g. isolated π^0's overlapping with $\pi^{\pm}$ (or randomly matching a track at primary vertex) or by $\pi^{\pm}$ interacting early in the ECAL.

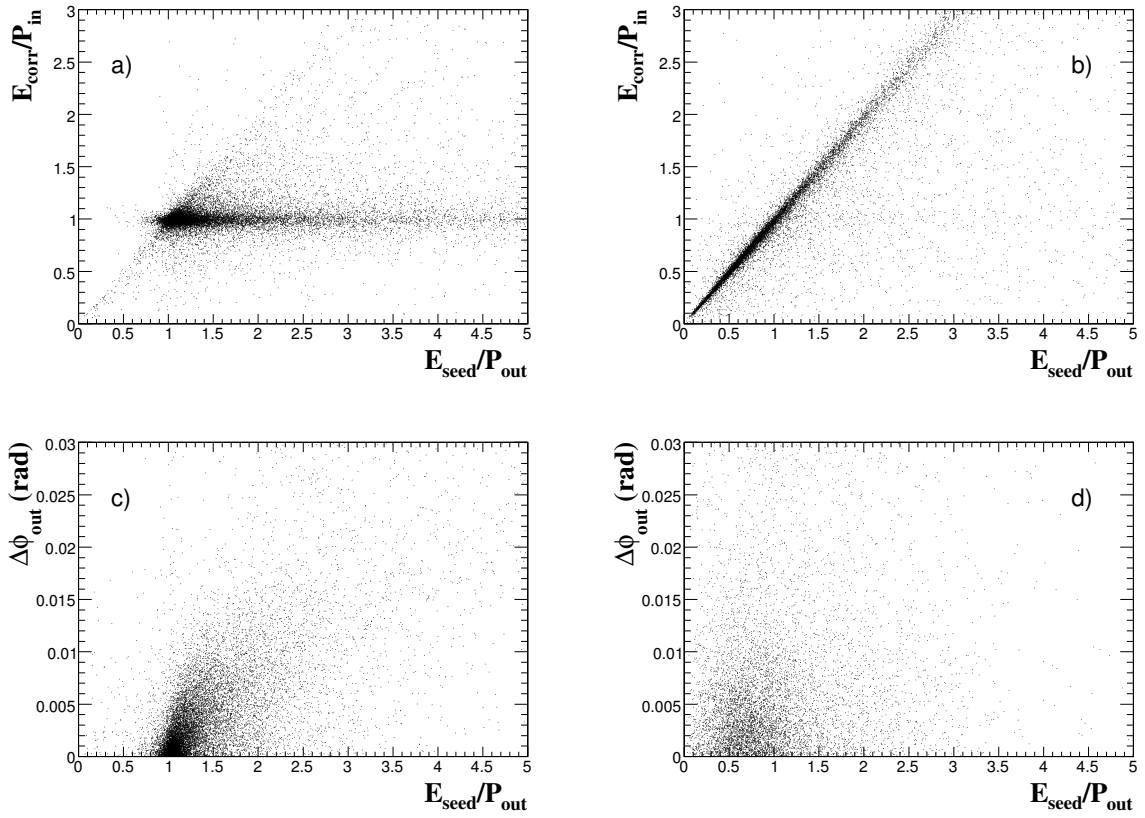


Figure 22: Correlations between different matching variables for electrons and for "fake" electron candidates from QCD jet background: a) $E_{\mathrm{corr}}/p_{\mathrm{in}}$ as a function of $E_{\mathrm{seed}}/p_{\mathrm{out}}$, electrons; b) $E_{\mathrm{corr}}/p_{\mathrm{in}}$ as a function of $E_{\mathrm{seed}}/p_{\mathrm{out}}$, jets; c) $\Delta\phi_{\mathrm{out}}$ as a function of $E_{\mathrm{seed}}/p_{\mathrm{out}}$, electrons; d) $\Delta\phi_{\mathrm{out}}$ as a function of $E_{\mathrm{seed}}/p_{\mathrm{out}}$, jets. Track parameters are either taken at the vertex ($_{\mathrm{in}}$) and matched with the supercluster parameters or at the outermost track position ($_{\mathrm{out}}$) and matched with the seed cluster. Distributions are for the barrel only. Similar distributions with larger spread are obtained for electrons in the endcaps.

The jet background can be further discriminated by a precise matching in energy and position between the

calorimeter cluster and the track and by the use of shower shape variables. Indeed, hadron showers are longer and broader, and subject to larger fluctuations, than electromagnetic showers. The bremsstrahlung, however, affects the electron identification capability. The electron shower shape, in particular in the ϕ projection, appears distorted. On the other hand, the emission of radiation in the tracker volume is a characteristic almost exclusive to electrons.

In addition of using track parameters at interaction vertex, the cluster-track matching can take advantage of GSF track parameters at the outermost layer, as these can better match the electron-induced component of the supercluster. Figure 22 presents correlations between $E_{\text{seed}}/p_{\text{out}}$ and $E_{\text{corr}}/p_{\text{in}}$ and between $\Delta\phi_{\text{out}}$ as a function of $E_{\text{seed}}/p_{\text{out}}$ for electrons and for the jet background. Different patterns resulting from the physics correlation due to bremsstrahlung in the electron case can be seen.

A further improvement in electron identification can be expected from the electron classification. Figures 23 and 24 present the normalized distributions of electron identification variables for electrons of the different classes and for the overall jet background, for electron candidates in the ECAL barrel. The variable H/E was introduced in Section 8. The variable Σ_9/Σ_{25} is the ratio of the energy sums over 3×3 and 5×5 crystal matrices centred on the highest energy (i.e. seed) crystal of the seed cluster. The shape variables $\sigma_{\eta\eta}$ and $\sigma_{\phi\phi}$ are defined from the crystals i and seed crystal s of the seed cluster as

$$\sigma_{\eta\eta} = \sum_{crystals} (\eta_i - \eta_s)^2 \frac{E_i}{E_{\text{seed cluster}}} \tag{1}$$

$$\sigma_{\phi\phi} = \sum_{crystals} (\phi_i - \phi_s)^2 \frac{E_i}{E_{\text{seed cluster}}} \tag{2}$$

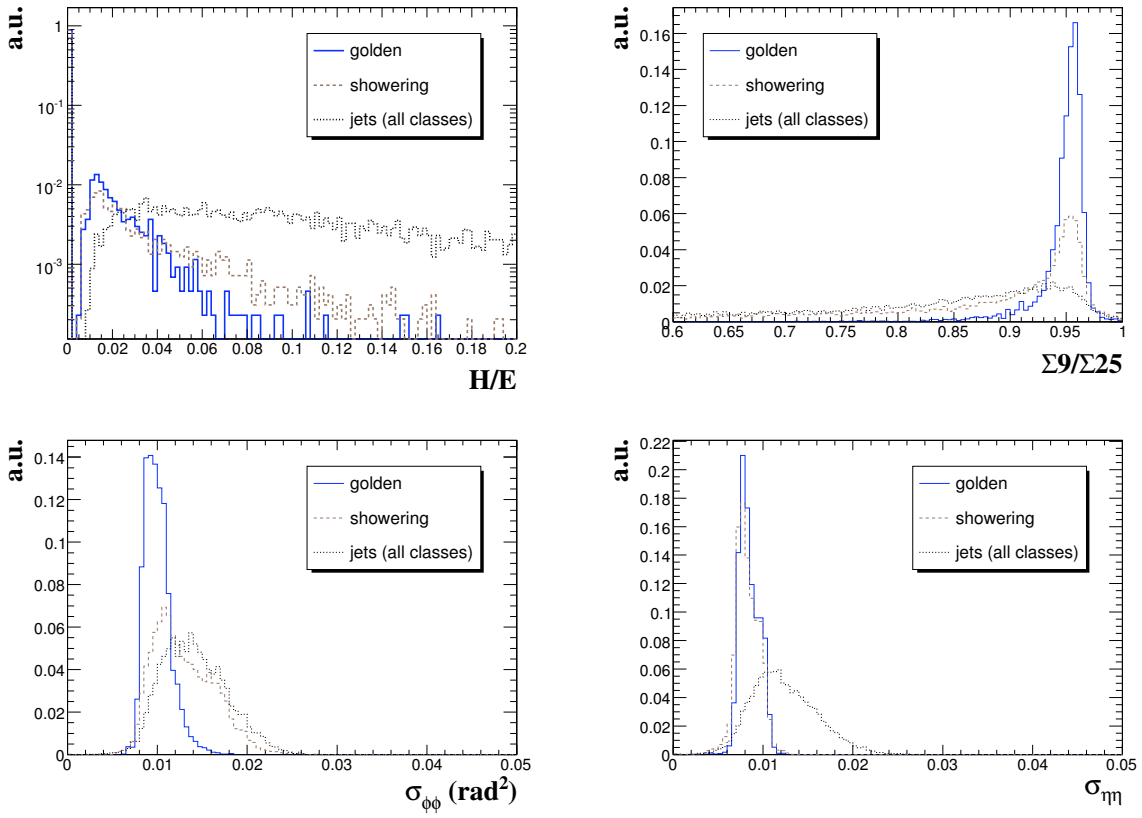


Figure 23: Distributions of electron identification observables in the ECAL barrel. The distributions of shower shape observables are shown for different classes of electrons, and in comparisons with distributions for "fake" electron candidates from QCD jet background. Similar distributions with larger spread are observed for electrons in the endcaps.

The necessity to distinguish electrons from the different classes clearly appears. The effect of bremsstrahlung is visible on the shape variables that involves the ϕ projection, looking at the distribution for Golden electrons

compared to the the distribution for Showering electrons. The distributions of $E_{\text{seed}}/p_{\text{out}}$ and $\Delta\phi_{\text{out}}$ show in addition the behaviour of the Narrow and Big Brem electrons, classes for which the level of bremsstrahlung as measured by the track bremsstrahlung fraction is requested to be high. For instance, the threshold at 2 in $E_{\text{seed}}/p_{\text{out}}$ for the Big Brem class corresponds to a bremsstrahlung fraction $f_{\text{brem}} > 0.5$ as required by construction for this class. Finally, as was already visible on Fig. 22, and expected as a consequence of bremsstrahlung, the shift in $E_{\text{seed}}/p_{\text{out}}$ towards values well above 1 is correlated for electrons with a shift in $\Delta\phi_{\text{out}}$ towards the position of the photon shower.

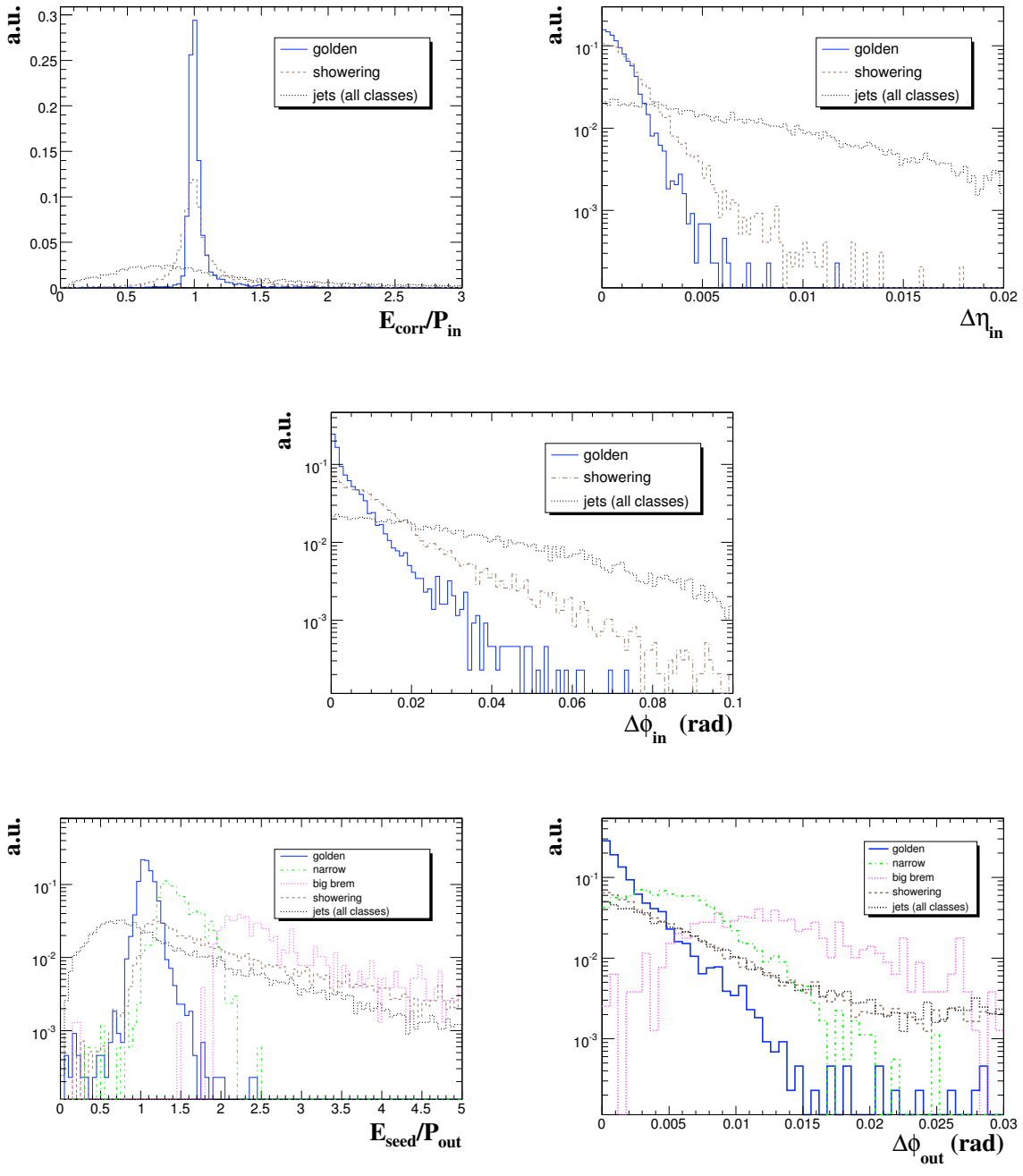


Figure 24: Distributions of electron identification observables in the η range of the ECAL barrel. The distributions of track-cluster matching observables are shown for different classes of electrons, and in comparisons with distributions for "fake" electron candidates from QCD jet background. Similar distributions with larger spread are obtained for electrons in the endcaps.

Considering further the distributions of classified "fake' electrons from the jet background, the following conclusions on the use of identification variables depending on the different electron classes can be drawn:

- the $\sigma_{\eta\eta}$ shape variable is discriminating for all electron classes;
- H/E and $\Delta\eta_{\rm in}$ and $\Delta\phi_{\rm in}$ are discriminating for all electron classes, with a slight loss of discriminating power for Showering electrons;
- $\sigma_{\phi\phi}$ and Σ_9/Σ_{25} shower shape variables involving ϕ projection gives discriminating power for all but Showering electrons;
- $E/p_{\rm in}$ and $E_{\rm seed}/p_{\rm out}$ are the most discriminating for the Showering class;
- $\Delta\phi_{\rm out}$, with track momentum from the outermost track position and energy from the electron seed-cluster, is discriminating only for Golden electrons.

The fraction of electron candidates from the signal and from the jet background which populates the different classes, integrated over the $\eta^{\rm e}$ range below 2.5 and with $p_T^{\rm e}$ in the range from 5 to 50 GeV$/c$, are presented in Table 2. The jet background is found to almost exclusively contribute to the showering class. This is due to the

	ECAL barrel fraction of		ECAL endcaps fraction of	
Class	electron (%)	jets (%)	electron (%)	jets (%)
Golden	24	4	17	2
Narrow	9	1	5	0.5
Big Brem	4	0.5	3	0.2
Showering	53	85	69	96
Boundary	10	9.5	6	1

Table 2: Fraction of electron candidate from the signal and from the jet background in the different classes and in the barrel and in the endcaps part of the ECAL. The signal is constituted an electron sample uniformly distributed in p_T and η electron in the range $5 < p_T < 50$ GeV$/c$, and the jet background by "electromagnetic" jets with p_T in the range $25 < p_T < 50$ GeV$/c$.

fact that this class corresponds to bad E/p match or multi-cluster pattern in the ECAL that are characteristic of the jet background. A large part of the real electron signal also populates the Showering class. As a consequence, the Showering class ultimately limits the overall (i.e. class independent) electron identification performance.

A simple cut approach is finally used to illustrate the electron identification capability based on the electron classification, focusing on the two extreme cases of Golden and Showering electrons. Cuts are defined for each class so as to provide an overall 90% electron efficiency and listed in Table 9, as well as the jet rejection for each class for an overall signal efficiency of 90%. The rejection factors are given as obtained relative to the sample of electron candidates after pre-selection and including the loose track isolation requirement. As can be expected, the rejection power obtained is higher for the Golden than for the Showering electrons. A clear difference is also observed between the barrel and the endcap parts of the ECAL. Using these simple cuts adapted to each pre-selected electron class, an absolute overall jet background fake rate efficiency at the level of 6×10^{-4} is obtained while keeping efficiency on electrons of p_T from 5 to 50 GeV$/c$ at the level of 90%.

The electron identification observables per class can be further combined for an optimal electron identification for a given physics channel using e.g. a likelihood implementation as proposed in Ref. [15].

Cut	ECAL barrel		ECAL endcaps	
	Golden	*Showering*	*Golden*	*Showering*
H/E	0.06	0.14	0.1	0.12
Σ_9/Σ_{25}	0.85	no cut	0.85	no cut
$\sigma_{\eta\eta}$	0.005-0.011	0.005-0.011	0.008-0.022	0.0-0.3
$\sigma_{\phi\phi}$	0.005	no cut	0.01	no cut
E/p	0.9-1.3	0.6-/	0.9-/	0.7-/
$\Delta\eta_{in}$	0.004	0.007	0.007	0.008
$\Delta\phi_{in}$	0.04	0.08	0.06	0.07
$E_{\rm seed}/{\rm p}_{out}$	0.9-1.6	0.75-/	0.6-2	0.8-/
$\Delta\phi_{out}$	0.011	no cut	0.02	no cut
rejection	9.0	7.3	5.9	4.2

Table 3: A typical definition of cuts for electron identification based on classes and for electrons in the barrel and in the endcap parts of the ECAL. The corresponding jet rejection factors are given relative to pre-selected electron candidates. The cuts are chosen so as to provide an overall electron efficiency of 90%.

10 Conclusions

A refined strategy for the reconstruction of electrons in CMS has been presented.

Cluster-driven pixel seed finding has been tuned for applications in physics requiring high efficiencies for low $p_T^{\rm e}$ electrons. Track reconstruction of electrons in CMS in the range $p_T = 5$ to 30 GeV$/c$ has been studied using a Gaussian Sum Filter to build the tracks. An overall strategy for the seed finding, trajectory building and fitting of the track parameters has been put in place in order to build tracks outward with high efficiency at the lowest p_T while keeping good track parameter resolution at the vertex. It is shown that the hits can be collected up to the outermost layers, without introducing a noticeable fake rate. All hits being collected, it is therefore possible to optimize the track parameter estimation. The method proposed here is to use the most probable value of the Gaussian mixture rather than the weighted mean, so as to get a better precision for electrons experiencing only small energy losses from bremsstrahlung emission along their trajectory in the tracker volume. The momentum resolution is comparable to that obtained using the HLT electron track reconstruction procedure. Finally, the proper treatment of the non Gaussian radiation losses allows for a meaningful estimate of the momentum at the outermost hit layer, providing in turn both an estimate of the bremsstrahlung fraction using the tracker information only and the possibility to improve the matching between tracking and calorimetry. Basic electron clustering algorithms for the ECAL barrel and endcaps have been tuned to improve the recovery of bremsstrahlung photons.

Electron classes have been defined according to the different electron topologies. These topologies have been used to define appropriate energy corrections and to serve as basis for electron quality in further analysis involving electrons. The resulting classification allows extreme cases to be separated, e.g. Golden electrons (predominantly truly "low radiating" electrons) having a track well matched in momentum and direction with a supercluster in the ECAL, from Showering electrons having emitted a large fraction of their initial energy as bremsstrahlung photons and therefore having an identified photon sub-cluster or a bad momentum match. Intermediate cases are classified using in particular the estimate of the amount of bremsstrahlung emission provided by the relative difference between the momentum estimate at both track ends. The showering category is shown to contain most of the electrons constituting the tail of the energy measurement from the ECAL. After having disentangled the different effects responsible for bias in the ECAL energy measurement, energy scale corrections are used which depend on a measured supercluster parameter. Only a small η-dependent bias remains for the Golden, Big Brem and Narrow classes and a common parametrized correction is applied. A stronger bias with η remains for the Showering class and this is parametrized separately. Different energy scale corrections are used in the η ranges of the ECAL barrel and endcaps. The resulting corrected distributions are slightly narrower and closer in shape to a Gaussian than the uncorrected distributions. Different energy measurement errors are obtained for the different electron classes. From the corrected ECAL measurement, the track momentum, and their associated errors, a combined estimate is constructed to obtain the final electron momentum at the vertex. As a result, an improvement is seen in the effective RMS of the momentum estimate for electron energies below $\simeq 25$ GeV. At 15 GeV energy, a factor of about 1.4 in precision is obtained by combining the ECAL energy with the tracker measurement.

Possible discriminating variables for the selection, isolation and identification of primary electrons, exploiting different electron measurement patterns, have been presented for electrons in the p_T range from 5 to 50 GeV$/c$. The full sequence of steps for electron reconstruction is described, from supercluster-driven electron track pixel finding and dedicated Gaussian Sum Filter track reconstruction to the track-supercluster matching. New observ-

ables sensitive to the amount of bremsstrahlung photons emitted in the tracker volume and their possible conversion in the tracker material are introduced. It is shown that different electron identification cut strategies and combination of electron identification observables are needed for the various "classes" of electron measurement patterns. The proper treatment of the class-dependent purity of electron samples is likely to have important applications and consequences in physics applications.

11 Acknowledgments

We wish to gratefully thank the CMS PRS ECAL-e/gamma working group for the support. Special thanks to Wolfgang Adam and Teddy Todorov for their precious help on the usage of the GSF algorithm. Thanks to Chris Seez and to Patrick Janot for very useful discussions and suggestions.

References

[1] R. Frühwirth, Comput. Phys. Commun., 100 (1997) 1-16.

[2] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, CMS Note 2005/001 (January 2005) 11pp.

[3] ORCA, "Object-oriented Reconstruction for CMS Analysis", version 8.7.3. (February 2005).

[4] CMS Collaboration, "The Compact Muon Solenoid - Technical Proposal", CERN/LHCC 94-38 (December 1994) 290pp.

[5] CMS Collaboration, "The Tracker Project - Technical Design Report", CERN/LHCC 98-6 (April 1998) 625pp.; *idem*, "Addendum to the Tracker TDR", CERN/LHCC 2000-016 (February 2000) 82pp.

[6] CMS Collaboration, "The Electromagnetic Calorimeter Project - Technical Design Report", CERN/LHCC 97-23 (15 December 1998) 364pp.

[7] S. Cucciarelli et al., CMS Note 2006/026 (January 2006) 22pp.

[8] CMS Collaboration, "CMS Physics Technical Design Report Vol I", CERN/LHCC 2006-001 (February 2006) 547pp.

[9] CMS Collaboration, "The Trigger and Data Acquisition Project", Volume I, CERN/LHCC 2000-038 (December 2000) 599pp.

[10] E. Meschi, T. Monteiro, C. Seez, and P. Vikas, "Electron Reconstruction in the CMS Electromagnetic Calorimeter", CMS Note 2001/034 (June 2001) 16pp.

[11] C. Rovelli, "The CMS Electromagnetic Calorimeter and the Search for the Higgs Boson in the Decay Channel $\mathrm{H} \to \mathrm{WW}^{(*)} \to 2e2\nu$", Doctor degree thesis, University of Milano-Bicocca (Italy) and Ecole Polytechnique (France) (January 2006) 186pp.

[12] N. Marinelli, "Track finding and identification of converted photons", CMS Note 2006/005 (January 2006) 23pp.

[13] F. Ferri, "The CMS Electromagnetic Calorimeter for the Higgs Boson Search $\mathrm{H} \to \mathrm{ZZ}^{(*)} \to 4e^{\pm}$ at the LHC", Doctor degree thesis, University of Milano-Bicocca (Italy) and Ecole Polytechnique (France) (December 2005) 170pp.

[14] D. Abbaneo, "Layout and Performance of the CMS Silicon Strip Tracker", Talk given at the 9th Pisa meeting on Advanced Detector, La Biodola, Isola d'Elba, Italy (May 25-31) 2003.

[15] P. Meridiani, "Optimization of the discovery potential of the Higgs Boson in the decay channel $\mathrm{H} \to \mathrm{ZZ}^{(*)} \to 4e^{\pm}$ with the CMS detector", Doctor degree thesis, University of Roma-1 (September 2004) 199pp.