UNIVERSITY of CALIFORNIA
Santa Barbara

# Experimental Methods to Measure Standard Model Backgrounds to New Physics Searches Using the CMS Experiment at the Large Hadron Collider

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Physics

by

James Michael Lamb

Committee in charge:

Professor Joseph Incandela, Chair
Professor David Stuart
Professor David Berenstein

May 2009

The dissertation of James Michael Lamb is approved:

---

Professor David Stuart

---

Professor David Berenstein

---

Professor Joseph Incandela, Chair

January 2009

**Experimental Methods to Measure Standard Model Backgrounds to New Physics Searches Using the CMS Experiment at the Large Hadron Collider**



by

James Michael Lamb

*For my father and my mother.*

# Acknowledgements

# Curriculum Vitæ

James Michael Lamb

## Education

| | |
|---|---|
| 2008 | Ph.D. (Physics), University of California, Santa Barbara |
| 2005 | M.A. (Physics), University of California, Santa Barbara |
| 2002 | B.A. (Physics), University of California, Berkeley |

## Publications

Data Driven Estimation of the Invisible Z Background to the SUSY MET Plus Jets Search, CMS Collaboration Physics Analysis Summary, 2008. Available at https://twiki.cern.ch/twiki/bin/view/CMS/PhysicsResults.

The CMS experiment at the CERN LHC, R. Adolphi *et al.* [CMS Collaboration], *JINST 3, S08004* August 14 2008.

CMS Silicon Strip Tracker Integration and Slice Test, J. Lamb, *Proceedings of the Hadron Collider Physics Symposium 2007, Nuclear Physics B (Proc. Supp.) 177+178* March 2008.

A Method of Measurement of the Top Quark Mass using the Mean Decay Length of B Hadrons in tt Events, C.S. Hill, J.R. Incandela, J.M. Lamb, *Phys.Rev.D.71:054029* March 29, 2005.

## Honors and Awards

| | |
|---|---|
| 2005 | Selected by the Department of Energy Office of Science as a delegate to the Lindau Meeting of Nobel Laureates |
| 2002-2003 | Wheelon Fellowship at UC Santa Barbara |
| 1997-2002 | Regents Scholarship at UC Berkeley |
| 1997 | National Merit Scholarship Finalist |

## Abstract

Experimental Methods to Measure Standard Model Backgrounds to New Physics Searches Using the CMS Experiment at the Large Hadron Collider

by

James Michael Lamb

The Large Hadron Collider at CERN in Geneva, Switzerland will produce proton-proton collisions of higher energy than ever before achieved an a laboratory. This thesis describes some of the experimental aspects of hadron collider physics, including the construction of the CMS detector which used to observe the collisions, and the interpretation of measurement results in light of challenging backgrounds. The focus of this this document is the development of a method to use high energy photon events in collider data to estimate the invisible Z background to a search for supersymmetric particles in the MET plus jets topology. A detailed study of the mechanisms by which quark and gluon scattering events can produce the signatures of electrons is presented. The process of the construction of the outer barrel of the silicon charged-particle tracker of CMS is described.

# Contents

# Chapter 1

# Introduction

The Large Hadron Collider (LHC) is a machine designed to collide protons together with energy larger than has ever before been achieved in a laboratory. Inside the protons, quarks and gluons interact. Most frequently they simply scatter, but occassionally their energy is transformed into the heavier particles of matter that can exist in the universe. The Compact Muon Solenoid (CMS) is a machine designed to observe and record these events. Physical theories, the Standard Model and beyond, predict what should happen; what particles should be produced, how often, and with what energy. The experimental technique of hadron collider physics is to observe the collisions; to see what particles are in fact produced, and with what rates, and so to differentiate between the existing theories, and perhaps to drive the creation of new theories by producing something totally unexpected.

## 1.1 The Hadron Collider Detector

CMS is a general-purpose collider detector. That is to say, it is designed to measure as accurately as possible anything that might happen in a collision. Nevertheless its capabilities are limited by the state of the art of radiation detection, and finite financial resources. Thus to understand experimental hadron collider physics one must understand the typical hadron collider detector and its measurement capabilities. The measurement process is illustrated in figure 1.1. The CMS detector is used as an example, but the general principles are in common with other general-purpose collider detectors, such as ATLAS[1] at the LHC and CDF and D0 at the Fermilab Tevatron. The particles produced in a collision travel outward from the interaction point, traversing several different sub-detectors as they go. First they pass through a charged-particle tracker. The tracker is made of several layers of silicon wafers, in a hermetic barrel plus endcap structure saturated in a high magnetic field provided by a solenoidal magnet. Charged particles leave a trace of electron-hole pairs when they traverse the silicon wafers, allowing their position to be recorded at several points in space. The magnetic field bends their trajectories with a radius of curvature proportional to their momentum. In this way, charged particles are identified and their momenta measured. Neutral particles generally pass unmeasured and undisturbed through the tracker.

After traversing the tracker, particles impact the electromagnetic calorimeter (ECAL). The ECAL is designed to measure the energies of electrons and pho-

tons, which normally deposit all of their energy inside that instrument. Other particles, because they are heavier, deposit only a small fraction of their energy in the ECAL. The ECAL functions by initiating electromagnetic showers in an adequately deep layer of material with large atomic Z. The showers result in a large number of low-energy electrons and photons whose total energy can be measured with conventional low energy particle detection techniques such as ionization and scintillation.

Just outside the ECAL in radius is the hadronic calorimeter (HCAL). The HCAL is designed to measure the energies of hadrons (particles composed of quarks). It measures their energies by measuring the energies of particle showers initiated and developed by strong-force interactions between the hadrons and the nuclei of heavy atoms. Just outside the HCAL proper is the solenoidal magnet, which also functions as the last showering layer for the measurement of particularly extended hadronic showers.

Finally, at the outermost radial positions, are the muon detectors. Muon detectors are charged-particle trackers; the CMS detector is designed such that in the vast majority of cases, if a particle has made it to the muon chambers, it is in fact a muon. This works because muons have large mass and also do not interact with the strong force, so that if they are above a certain momentum (a few GeV/c) they are not stopped in the ECAL, HCAL, or any other structure inside the detector. They also have sufficiently long lifetime to travel through the muon

chambers without decaying. The muon detectors are saturated in the return field of the solenoid. This means that the measurement of muons' position in the muon chambers improves the resolution of the momentum measurement, although this improvement is significant only for very high energy muons (hundreds of GeV/c or more), due to the effect of multiple small Coulomb scatterings of the muon in the HCAL, magnet, and muon chambers themselves.

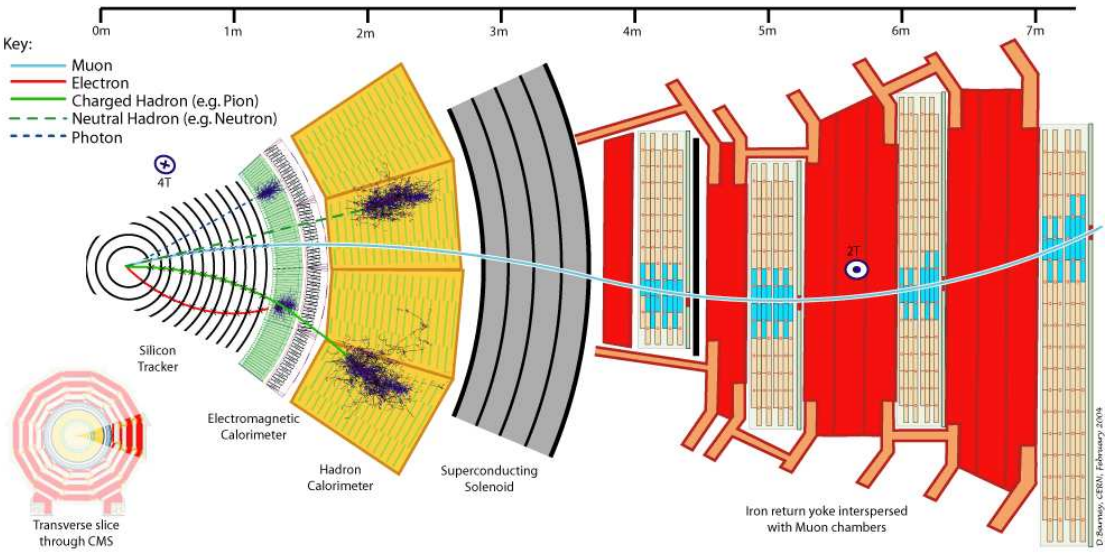


Figure 1.1: Schematic cross-section of the CMS detector.

A large multiplicity, up to hundreds of particles, can be produced in a proton-proton collision. Their energies are measured, their momenta are measured if they are charged, and certain types can be distinguished according to which sub-detector they deposit their energy into. But the the particles that actually interact with the detector are relatively long-lived descendents of fundamental, and typically ephemeral, particles with which the most parsimonious description of the

collision, and indeed all matter, is described. In the following section the spectrum of these fundamental particles is briefly described; then I return to measured quantities and how they can be interpreted to understand the description of the collision in terms of fundamental particles.

## 1.2 The Particles of the Standard Model

A high-energy collision can produce any of the known particles that make up matter in the universe. They can be classified by their properties (and also according to their places in the fundamental theory of matter, the Standard Model of Particle Physics), into the following categories: leptons, quarks, and force-carrying bosons. Leptons include the electron as well as its heavier analogues the muon and the tau, and their partners the neutrinos (one flavor for electron, one for muon, and one for tau). The quarks, the up and down flavors of which make up the proton and neutron and hence most of the matter in the macroscopic word, include also charm, strange, bottom and top flavors. Finally there are the force carriers, including the W and Z bosons, which carry the weak force, the gluon which carries the strong force, and the photon, the particle that carries the electromagnetic force. This categorization is encapsulated in figure 1.2.

After a quark is produced or scattered in a high energy collision, it does not live on its own but rather it rapidly forms a composite particle with one or two

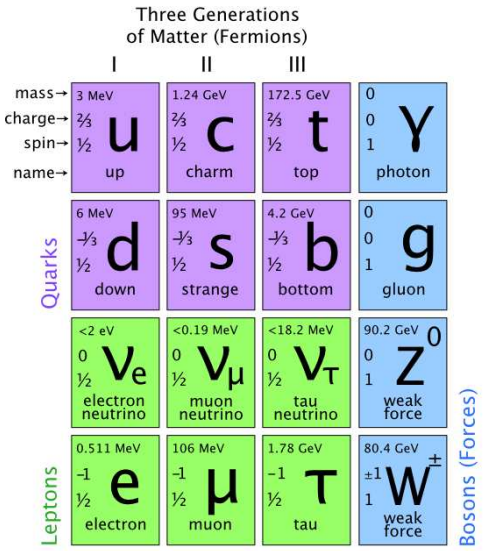


Figure 1.2: Particles in the Standard Model of Particle Physics. Image used with permission (http://commons.wikimedia.org/wiki/User:MissMJ)

other quarks such that the composite's net strong force charge is neutral.[1] As the produced or scattered quark separates from the other strong force-interacting particles involved in the collision, the potential energy due to the strong force becomes so large that it becomes energetically favorable for one or more quark-antiquark pairs to be produced across the separation. This process is called fragmentation or hadronization. Additionally, gluons may be radiated by the bare quarks before hadronization, or produced directly in the collision, and subsequently they split into quark-antiquark pairs which participate in the hadronization. The produced hadrons may be short-lived, and subsequently decay into two or more particles. The end result is that one quark becomes multiple energetic particles in a cone

[1]Such a particle is called a hadron; protons and neutrons are familiar examples.

about the quark's direction of travel[2]. Hadronization happens far faster than any conceivable instrument could measure, so it is never directly observed. Only the resulting collimated group of particles, referred to as a hadronic jet, is directly observable. Figure 1.3 provides a useful look at the spatial distribution of transverse energy in a typical $t\bar{t}$ event. The figure shows a two-dimensional histogram of the locations of stable particles in a single $t\bar{t}$ event, weighted by their transverse energies. The locations of final-state particles directly participating in the collision are indicated. The energy of stable particles produced by quark hadronization is deposited over a relatively wide area.

W and Z bosons can also be produced directly in collisions. They always decay to two fermions (leptons or quarks) but it is generally only possible to identify them when they decay leptonically. A W decays leptonically to an electron, an muon, or tau, and the corresponding neutrino; a Z decays leptonically to two electrons, two muons, two taus, or two neutrinos. W and Z bosons are identified by finding their leptonic decay products. When a W or Z decays to quarks (about 70% of the time) its presence in the event is essentially impossible to determine.

Before trying to understand how the fundamental interaction of the collision is reconstructed, it is useful to see the spectrum of stable[3] particles produced in a collision. Figures 1.4 and 1.5 show the particle production spectrum, and the

[2]The top quark is an exception. It decays to a b quark and a W boson even before hadronization occurs (which is much faster than can ever be measured in any conceivable detector). It is then the b quark which participates in the hadronization.

[3]For these purposes particles are considered stable if there lifetime is sufficiently long that they interact with the detector. An arbitrary lifetime cut-off of $c\tau = 1$ cm is used as a criterion.

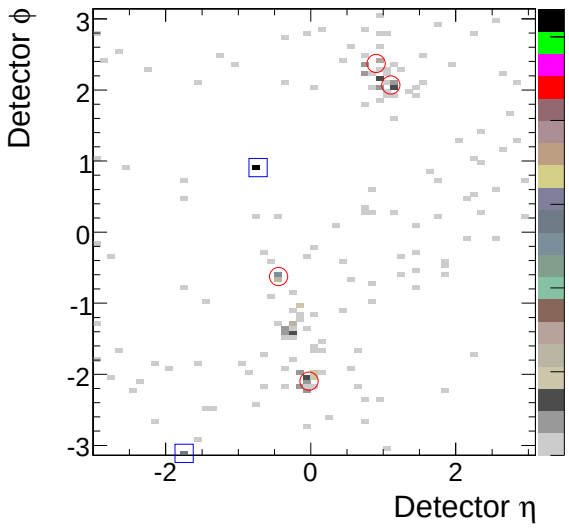


Figure 1.3: Spatial distribution of the transverse energy of stable particles in a typical $t\bar{t}$ event, in the detector coordinate system. Blue squares show the location of an electron and a neutrino from the leptonic decay of a W boson. Red circles show the location of b quarks and quarks from the decay of the other W. The cluster of energy that is unmarked by a circle or square is the the result of a final-state gluon radiated by one of the quarks.

particle production spectrum weighted by transverse energy $E_T$[4], for $t\bar{t}$ events, high energy QCD events, and LM1 Supersymmetry (SUSY) events[5]. The QCD events shown have center-of-mass energy transfer of at least 120 GeV; such events will occur at a rate about 60 times that of $t\bar{t}$ events.

The first thing one should notice in these figures is the high multiplicity of particles; mean multiplicity is over 500 particles per event in $t\bar{t}$. Particle production is dominated by charged pions and photons, both in terms of the numbers of particles and in terms of the fraction of total transverse energy carried by the particles. The vast majority of the photons come from the decay of neutral mesons. Both neutral mesons and charged pions are produced by the fragmentation of quarks. By comparing figure 1.4 with figure 1.5 one can see that the average transverse energy of these pions and photons is quite low, around 1 GeV.

The second important thing to notice in these figures is how similar the distributions look for the three types of events. The main visible difference is that in the three types of events, very different fractions of transverse energy are carried by the relatively small number of electrons, muons and stable weakly-interacting particles[6]. It is apparent that the presence of energetic weakly-interacting particles is a particularly good way to distinguish $t\bar{t}$, SUSY and QCD events for each

[4]Transverse energy, $E_T = sin\theta$ where $\theta$ is the angle to the colliding beam line, is a measure of momentum in the plane transverse to the direction of the colliding protons.

[5]LM1 describes a particular supersymmetric theory (mSUGRA), and choices for a corresponding set of parameters. It is commonly used by the CMS experiment as a benchmark.

[6]Neutrinos from the decay of W in the case of $t\bar{t}$ events; neutrinos and hypothetical *lightest supersymmetric particle* (LSP) in the SUSY events

other.

Particles that interact only through the week force are not directly visible to the detector. The interaction cross-sections are low, and the mean free path for interactions in matter is much larger than the size of the CMS detector (indeed much larger than the size of the Earth). Their presense can only be inferred from a lack of momentum balance in the visible part of the event. Conservation of momentum requires the vector sum of the momenta of all final-state particles in the event, transverse to the colliding proton beam direction, to equal the initial transverse momentum of the pair of colliding protons, which is essentially zero. The negative of the vector sum of transverse momenta of visible particles, called *missing transverse energy* (MET), is a measurement of the summed transverse momenta of the invisible particles in the event. The need to accurately measure MET has two implications for detector design: the importance of precisely measuring the *jet energy*, and the need for detector hermeticity.

The multitude of photons and charged pions seen in the figures anticipates the magnitude of the problem of *fakes*, instances where a particle, or group of adjacent particles, of a given type is mistaken for a particle of another type. *Prompt* photons and electrons, meaning those produced in the primary interaction and not by quark hadronization and subsequent decay, are of central importance to the analysis presented in this thesis. The mixture of photons and charged pions presents a copious background to prompt electrons; as will be discussed in detail

later, electrons can be faked by overlaps of photons and charged pions, and also by charged pions alone when they, with low probability, deposit all of their energy in the ECAL.

Figures 1.6, 1.7 and 1.8 provide another view of the transverse energies of particles in high energy QCD, $t\bar{t}$, and LM1 SUSY events. These figures illustrate the fact that muons and electrons are very useful for selecting $t\bar{t}$ events out of a background of QCD. The figures also highlight how SUSY events might stand out from background events: they have many more energetic leptons, and much more MET, than QCD events, and they have much more hadronic energy than $t\bar{t}$ events.

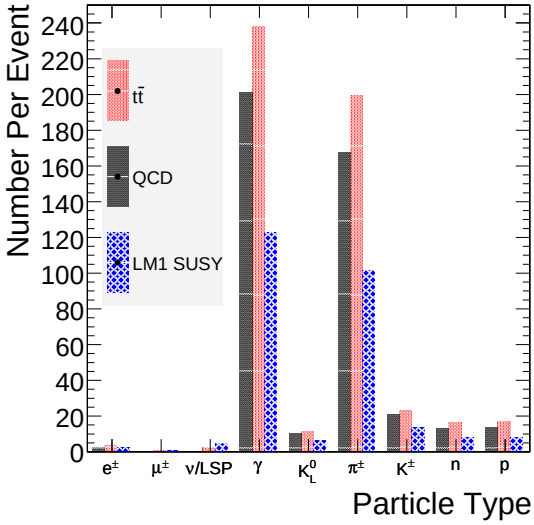


Figure 1.4: Per-event average distribution of the most-frequently produced stable particles in high energy QCD, $t\bar{t}$, and SUSY LM1 events. The QCD events shown have center-of-mass energy transfer of at least 120 GeV; such events will occur at a rate about 60 times that of $t\bar{t}$ events.

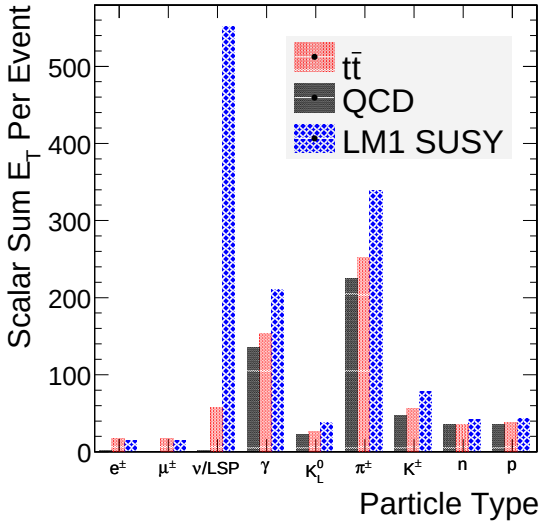


Figure 1.5: Total $E_T$ spectrum of particles in high energy QCD, $t\bar{t}$, and SUSY LM1 events. Each bar in the chart shows the total transverse energy per event of particles of the specified type, averaged over many events.

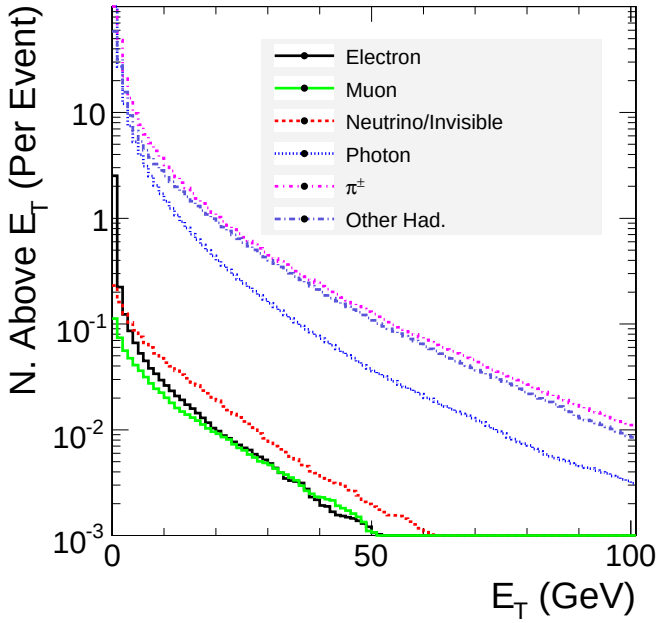


Figure 1.6: $E_T$ of particles in a high energy QCD event sample. The curves show the number of particles above the $E_T$ threshold specified by the X axis.

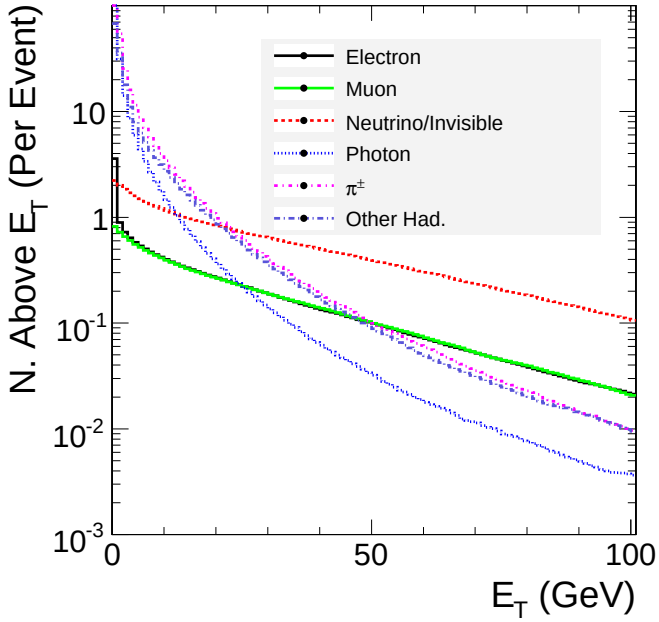


Figure 1.7: $E_T$ of particles in $t\bar{t}$.

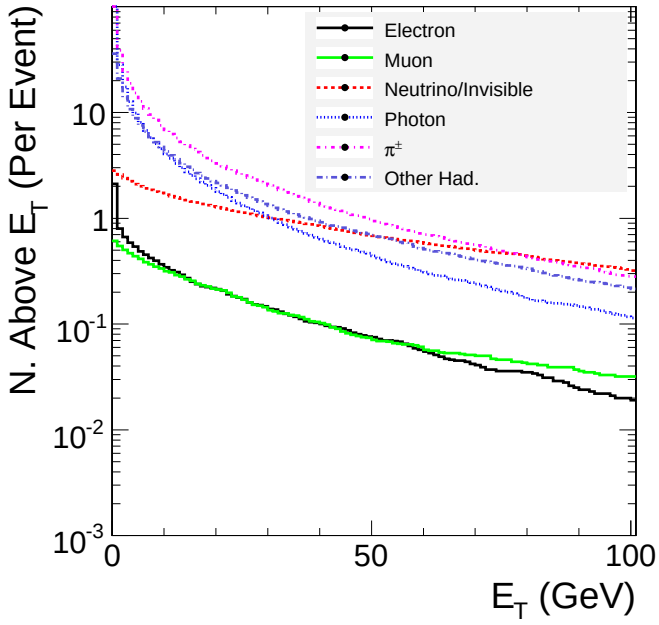


Figure 1.8: $E_T$ of particles in LM1 SUSY.

# Chapter 2

# The CMS Detector

The CMS detector is composed of four principle subdetectors: the inner tracker, the electromagetic calorimeter and pre-shower detector (ECAL), the hadronic calorimeter (HCAL), and the muon spectrometer. In the following sections I present aspects of these subdetectors that are relevent for the work presented in this thesis. Further information can be obtained from [2]. The reader who wishes to gain familiarity with concepts in radiation detection techniques is directed to Ref. [3]. Reference [4] presents an excellent review of calorimetry for collider detectors.

## 2.1 Tracker

CMS uses an all-silicon inner tracker, described in figure 2.1. It is comprised of 13 cylindrical layers in the barrel, and 14 discs in the endcap. The innermost

three barrel layers and innermost two endcap discs are silicon pixels [5]. The remainder is composed of silicon strip detectors. Pixel and strip silicon detectors are made of semiconducting silicon wafers doped to form diode junctions and placed under a reverse bias voltage adequate to fully deplete the silicon of mobile charge. Because the silicon is doped to be a semiconductor, a well-behaved electric field is developed across the wafer cross section. When a charged particle passes through a wafer, it deposits ionization energy which produces electron-hole pairs. The electric field carries the electrons and holes to the outside of the wafer, where they can be recorded by readout electronics. The difference between pixel and strip detectors is in the segmentation of the device. In a pixel detector the silicon wafer is read out in a grid structure, such that the passage of a particle through the silicon can be localized in two coordinate directions. For strip sensors only one spatial coordinate can be reconstructed. Two strip sensors can be placed adjacent to one another in parallel planes, but with an angle between their strips, to allow reconstruction of two spatial coordinates. This is done in several layers of the CMS tracker.

Semiconductor pixel detectors present certain advantages that have made them the de facto requisite choice for the innermost tracking layers of the general purpose detectors at the LHC and future conceivable detectors[1]. Their good resolution in two dimensions is useful for measuring track impact parameters and

[1]Several groups have in fact suggested to replace the entire inner tracker of CMS with pixel detectors for the CMS upgrade.

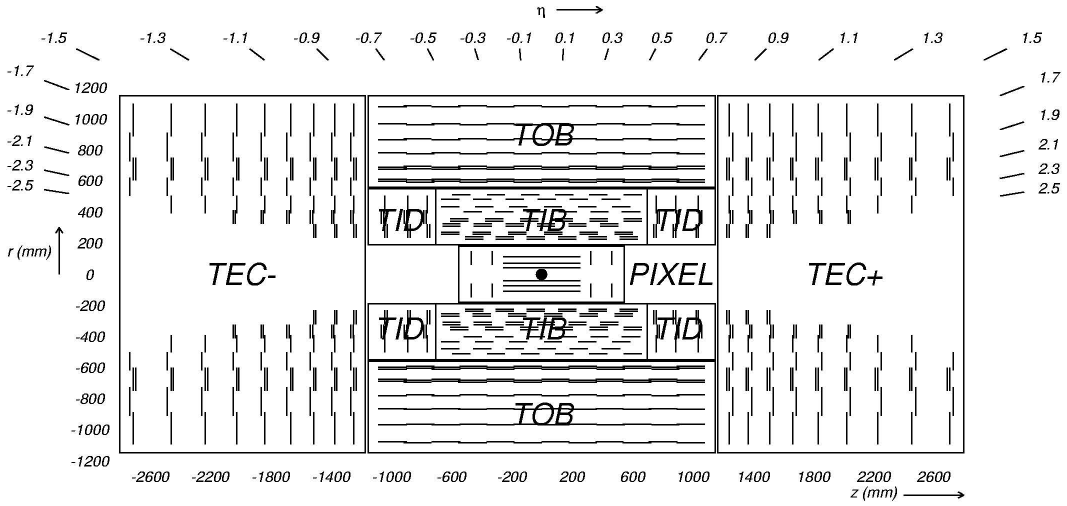


Figure 2.1: Schematic diagram of the CMS inner tracker.

finding secondary vertices. Their fine segmentation allows low occupancy[2] to be maintained even at small radius. Also because of their fine segmentation, pixel sensors present less capacitive load to the readout amplifiers, leading to lower noise. This in turn means that a smaller signal charge deposition is allowable, corresponding to a thinner layer of depleted silicon. The ability of pixel sensors to continue functioning with good signal-to-noise ratio even with a sensor that is only partially depleted means that pixels are less sensitive to radiation damage. This is an important advantage, since being placed close to the interaction point they receive a high radiation dose per unit area.

The performance of the inner tracker is ultimately characterized by its ability to resolve the existence of charged particles and to accurately measure their tra-

[2]Occupancy: average number of particles tracker per readout channel per event

jectories. It should have sufficient granularity to resolve two closely-spaced tracks in order to be able to identify individual tracks in a high-occupancy environment. The tracker should also leave as undisturbed as possible all particles as they pass through. Thus competing requirements are presented. Track finding and trajectory reconstruction are aided by increasing the number of layers of the tracker, and increasing the segmentation of each layer, both of which lead to an increase in the amount of material in the tracker, primarily through the increased number of readout amplifiers and the corresponding increases in low-voltage power supply and cooling needed.

A silicon detector has the advantage that when a particle traverses it, the hit is almost certainly reconstructed. This is due in part to the high signal to noise ratio (S/N) of silicon detectors. The S/N is determined by the amount of charge deposited in silicon when a charged particle traverses it, and by the electronic noise of the readout, which is dominated by the noise of the first stage amplifier. The CMS tracker has been designed to have high signal to noise and a low numbers of layers in an attempt to find a balance between the competing constraints described above. It should be noted that a scheme in which only every other silicon strip channel is read out, as for example was used in the Layer 00 sub-detector [6] of the CDF experiment, could have substantially lowered material budget and cost while maintaining good position resolution and adequate S/N. It remains to be seen whether decreasing the amount of material in the CMS tracker

might have brought it to a more optimized balance (see figure 2.2), but it is clear that it is a highly performing instrument.

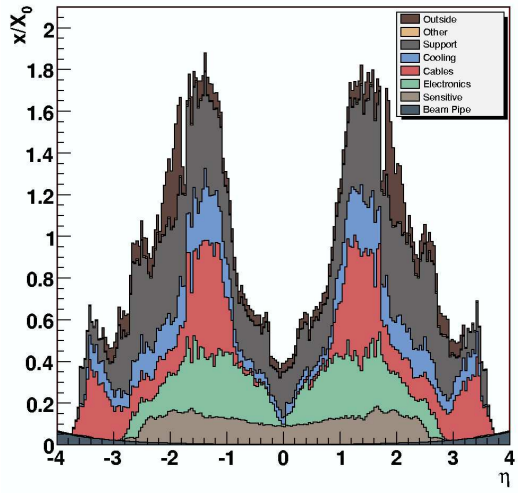


Figure 2.2: Tracker material budget in units of radiation length, from Ref. [2].

## 2.2 Electromagnetic Calorimeter

The CMS electromagnetic calorimeter is a homogenous[3] calorimeter composed of lead tungstate ($PbWO_4$) crystals read out by silicon avalanche photodiodes in the barrel and vacuum phototriodes in the endcaps. Lead tungstate was chosen for the absorbing/scintillating material of the calorimeter because of its short radiation length (0.89 cm), which allows for a compact calorimeter (which is important in order to reduce the volume and hence expense of everything outside the ECAL, including HCAL, magnet, and muon systems), because of its small

[3]Homogenous calorimeter: the same material performs absorbing and scintillating function.

Moliere radius[4] (2.2cm), and especially because of its radiation hardness and its fast scintillation decay time which is comparable to the LHC bunch spacing of 25 ns. Disadvantages of lead tungstate include its low light yield with respect to other scintillators and the temperature dependence of its light output, which varies by about 2% per degree at 18 °C. The readout photodiode response also varies with temperature; the variation in the total response to incident electrons has been measured to be (-3.8±0.4)% per degree C [7].

The ECAL crystals in the barrel are 2.2 cm square on the side facing the interaction point, 2.6 cm square on the side facing away from the interaction point, and 23 cm deep. This corresponds to an angular cross section of 0.0174 x 0.0174 in $\eta - \phi$ space and a depth of about 26 radiation lengths. In the endcap, the crystals are 2.86 cm square at the side facing the interaction point, 3 cm square on the side facing away from the interaction point, and 22 cm deep. This corresponds to about 0.008 radians in the $\theta$ direction, and a depth of about 25 radiation lengths.

Crystals are mounted in such a way as to be a few degrees (3 degrees in the barrel, 2-8 degrees in the endcap) off-pointing from the nominal interaction point, in order to avoid that particle trajectories pass through cracks in between crystals.

The ECAL barrel extends to a pseudorapidity of $|\eta| < 1.49$, and the endcap covers the range $1.479 < |\eta| < 3.0$. The barrel-endcap interface region, approx-

[4]The Moliere radius of a material describes the transverse development of electromagnetic showers and is normally defined such that 90% of the energy deposit of a shower is contained in a cylinder of one Moliere radius.

imately in the range $1.46 < |\eta| < 1.558$, has greatly reduced depth of radiation lengths. The analyses presented in this thesis exclude electrons and photons reconstructed from ECAL energy deposits centered on that portion of the detector.

Between the tracker endcap and the ECAL endcap, there is a *preshower* detector. The primary purpose of the preshower detector is to distinguish single photons from closely-spaced photon pairs, which are produced in the decay of neutral mesons. The preshower detector can also help distinguish electrons from other charged particles, and can improve the resolution of the trajectory of electrons and photons at the point of their impact onto the ECAL.

The preshower is constructed like a two-layer sampling calorimeter. It has two lead absorber layers, each followed by a layer of silicon strip detectors. The first layer is two radiation lengths deep, and the second is one radiation length deep. The principle of operation of the preshower detector is that incident photons will begin showering when they hit the lead absorber, and the shower centroids are measured to high precision by measuring the positions of the electrons of the shower with the silicon strip detectors. The silicon strip detector has a pitch of 1.9 mm. The crystals of the ECAL are 2.2 cm on a side, on the face towards the interaction point. Hence the preshower is able to resolve closely-spaced photon pairs that are indistinguishable from a single photon when reconstructed using the ECAL alone. With three radiation lengths of material comprised in both layers of the preshower detector, approximately 90% of incident photons will begin to

shower (by converting to electrons) before the second silicon strip layer of the preshower. The energies of electrons and photons reconstructed from superclusters in the ECAL endcap are corrected to account for energy deposited in the preshower detector.

The preshower detector is not used in the analyses presented in this thesis. That is because the preshower detector was not expected to be ready for the first data-taking run of CMS[5].

## 2.3 Hadronic Calorimeter

The CMS HCAL consists of four subdetectors: central and endcap calorimeters HB and HE, outer calorimeter HO, and forward calorimter HF. The HB and HE have very similar designs and cover the $\eta$ ranges up to 1.3 and from 1.3 to 3.0, respectively. They are sampling calorimeters that use brass and steel as absorbers and use plastic scintillator. They sit entirely inside the solenoidal magnet. The HO consists of scintillator placed outside the solenoidal magnet; its purpose is to sample showers that extend beyond HB. The forward calorimeter, which extends from $\eta > 3.0$, is not used by analyses presented in this thesis, and will not be described[6].

The depth of HB and HE ranges from 5.8 interaction lengths at $\eta$=0 to about

[5]Since the first run was delayed from fall of 2008 to fall of 2009, the preshower actually should be installed for the first CMS run.

[6]The interested reader should refer to [8]

10 interaction lengths at $\eta$ up to 3. The ECAL adds an additional 1.1 interaction lengths, and HO brings the total minimal calorimeter depth to almost 12 interaction lengths, except at the barrel-endcap interface region.

**HB** The HB consists of 16 absorbing layers. The two outer layers are stainless steel, for structural support. The remainder are brass. Seventeen sampling layers of scintillator are used. The first sampling layer is placed before the first absorbing layer; its purpose is to sample showers developing in inert material between the ECAL and HCAL. The HB depth at normal incidence ($\eta = 0$) is 5.8 interaction lengths. Effective depth increases with polar angle, since HB is nearly uniform in radial depth; at $\eta = 1.3$ it is 10.6 interaction lengths deep. Readout of the HB scintillator is segmented as 0.087 x 0.087 tiles in $\eta - \phi$.

**HE** The design of HE is very similar to that of HB. Notable exceptions are that the segmentation is 0.087 x 0.087 $\eta - \phi$ up to $\eta < 1.6$, but 0.17 x 0.17 for $\eta > 1.6$, and that calorimeter depth is about 10 interaction lengths, uniformly across the entire $\eta$ range of HE.

**HO** The HO consists one scintillator layer to sample hadronic showers developing in the magnet. The magnet adds about $1.4/\sin(\theta)$ of interaction lengths to the calorimeter depth, where $\theta$ is the angle from the beam direction. Because the most-central region of HB has the lowest depth, the most-central segment of HO has an additional 19.5cm-thick layer of iron, and an

additional scintillator layer to read it out. Readout segmentation of HO is similar to that of HB. The HO provides coverage to approximately $\eta < 1.3$.

## 2.4 Muon Spectrometer

The CMS muon spectrometer actually consists of three separate sub-detectors: the drift tube system (DT), the cathode-strip chambers (CSC), and the resistive-plate chambers (RPC). CMS uses these three different gas ionization detector technologies in order to meet the needs of different rate and magnetic field environments in the forward and central rapidity regions and the need for short time resolution to associate muon tracks to the correct LHC bunch crossing.

**DT** The Drift Tube system covers the central rapidity region up to $|\eta| < 1.2$. The central region is a relatively low-rate environment with a low magnetic field.

**CSC** The cathode-strip chamber system gives primary coverage to the rapidity region $1.2 < |\eta| < 2.4$ where the expected rate is high and the magnetic field is large. It gives coverage with a reduced redundancy (reduced number of layers) in the barrel-endcap region of $0.9 < |\eta| < 1.2$.

**RPC** Resistive-plate chambers cover the region $|\eta| < 1.6$. The RPCs are used in CMS in addition to the DT and CSC systems because they provide very fast time resolution, short with respect to the 25 ns bunch spacing of the LHC.

This is advantageous in order to unambiguously determine which bunch crossing a muon came from.

# Chapter 3

# CMS Reconstruction

The CMS experiment follows an analysis paradigm in which raw detector measurements of position and energy deposit are reconstructed into *physics objects* that correspond to the particles initially produced in the collision before any secondary decay or interaction in the detector. Physics objects include electrons, muons, taus, photons, hadronic jets (corresponding to quarks), heavy-flavor hadronic jets (corresponding to b and c quarks), and missing transverse energy (corresponding to neutrinos or other weakly-interacting particles). In this chapter details of the reconstruction that are pertinent to the work presented in this thesis are presented.

## 3.1 Electrons

### 3.1.1 Electron Reconstruction

Electron reconstruction in CMS requires, at the most basic level, a cluster of energy in the ECAL that is matched to a track in the inner tracker. Energy clustering is designed to collect all the energy from an electron incident on the ECAL. Clusters are defined as regions of energy deposit that are centered upon a high-energy seed crystal and which extend from that seed crystal in the $\eta$ and $\phi$ directions until energy deposition either falls below a preset threshold or begins to rise again. The electron energy is taken from the energy of the cluster.[1] The track is used to deduce the charge of the electron, and for background rejection through quality cuts on track momentum and cluster energy matching and track and cluster position matching. A complete description of electron reconstruction in CMS can be found in Ref. [9].

Many features of electron reconstruction are driven by the need to accomodate bremsstrahlung radiation, which is the dominant energy loss mechanism for electrons of energy above about a few tens of MeV.[2]. An electron traversing the CMS tracker from the interaction region to the ECAL passes through an amount of material between about 0.5 and about 1.5 radiation lengths, depending on $\eta$ (see figure 2.2). Hence most electrons lose a significant fraction of their energy to

[1] In certain rare cases when the cluster energy is believed to be badly mismeasured, the electron energy is taken from the measured momentum of the track.

[2] Electrons of interest to analyses at the LHC have energy of at least few GeV/c.

bremsstrahlung. According to Ref. [9], "About 35% of the electrons radiate more than 70% of their initial energy before reaching the ECAL. In about 10% of cases, more than 95% of the initial energy is radiated." Bremsstrahlung energy loss occurs in discrete events that often result in the electron losing a large fraction of its energy to the bremsstralung photon. This drives two important features of the reconstruction: the building of *superclusters*, and changes in the tracking algorithm to allow for drastic momentum changes.

**Supercluster Finding**

If an electron did not emit bremsstrahlung radiation, its energy deposit would be very tightly clustered in the ECAL, with a characteristic size determined by the Moliere radius of the $PbO_4$ crystals that compose the ECAL. About 94% of the energy of an electron incident on the ECAL is deposited in a 3x3 crystal array, and about 97% in a 5x5 crystal array [10]. Bremsstrahlung radiation spreads the energy deposition in $\phi$, because the magnetic field bends the trajectories of electrons, with radii of curvature proportional to their momenta, and it doesn't bend photon trajectories. The result is that a single primary electron can deposit energy in disjoint clusters spread out in the $\phi$ coordinate. Therefore single ECAL clusters are combined along a window that is centered on the seed cluster and is very narrow in $\eta$ and wide in $\phi$.

**GSF Tracking**

Electron track reconstruction uses a different algorithm than the standard algorithm used for generic track finding in CMS. It is called the Gaussian Sum Filter (GSF) algorithm. The GSF algorithm uses a wider parameterization for the effective allowed changes in curvature of tracks between hits.

### 3.1.2 Electron Classification

Electrons are classified according to the amount of bremsstralung radiation they emit. The classifications are used to aid the understanding of the scale and resolution of the electron energy measurement. Classification is based on $f_{brem} \equiv (p_{in} - p_{out})/p_{in}$, where $p_{in}$ and $p_{out}$ are the measurements of the track momentum at its closest approach to the vertex and at its last measurement point (hit nearest the ECAL). It also uses the match between the track momentum and the supercluster energy, the match between the supercluster and track momentum measured at the entrance to the ECAL, and whether the supercluster was formed by just one or by multiple clusters. The classification is as follows:

**Golden:** Supercluster is formed of a single cluster, $f_{brem} < 0.2$, $\Delta\phi(track, supercluster) < 0.15$, $E/p > 0.9$.

**Big brem:** Supercluster is formed of a single cluster, $f_{brem} > 0.5$, $0.9 < E/p < 1.1$.

**Narrow:** Supercluster is formed of a single cluster, $0.9 < E/p < 1.1$, $f_{brem}$ and/or $\Delta\phi(track, supercluster)$ outside the ranges allowed for Golden and Big Brem classes.

**Showering:** Electrons failing the above categories. Multiple clusters, corresponding to separated bremsstrahlung deposits, are allowed. Bad supercluster-track matching in energy and direction is allowed. This category corresponds to poorly-measured electrons involving multiple bremsstrahlung radiations or bremsstrahlung radiation followed by subsequent conversion of bremsstrahlung photons in tracker material. The showering class is the predominant class of electrons.

### 3.1.3 Supercluster Energy Correction

The supercluster energy is corrected in a number of ways to more accurately measure the energy of the electron. The first correction accounts for energy that was deposited in the ECAL but not clustered. This correction is a function of the number of crystals in the seed cluster. Separate correction factors are needed for superclusters in the barrel and in the endcap. After the crystal number-dependent correction is applied, a residual correction factor as a function of $\eta$ is still necessary. One value of the $\eta$-dependent correction factor is used for Golden, Big Brem, and Narrow electrons. A different value is necessary for electrons in the Showering class.

The corrections are ad hoc factors that are currently extracted from Monte Carlo simulated data. They will be extracted from real data when it is available, for example using a tag-and-probe method with $Z \rightarrow ee$ events.[3]

### 3.1.4 Electron Identification

An ECAL supercluster roughly matched to a track doesn't have to come from an electron. Background processes, for example overlaps of charged pions with photons from neutral pion decay, can give rise to the same signature. In order to get a purer sample of electrons, additial selections are applied. The electron identification in use by CMS, detailed in [11] is based on the quantities described below. Distributions of the described quantities are shown in figures 3.1 through 3.5.

$H/E$ The ratio of energy deposited in the HCAL directly behind the ECAL supercluster to the energy of the supercluster. This is useful because particles other than electrons and photons will normally not deposit all of their energy in the ECAL.

$\sigma_{\eta\eta}$ A measure of the width of the supercluster in the $\eta$ direction. The energy deposit of a single electron, while it can be quite elongated in the $\phi$ direction due bremsstrahlung radiation, will be quite narrow in $\eta$. On the other hand,

[3] Tag-and-probe measurements use the Z mass constraint and tight selections for one electron (the tag); the other object (probe) is known to be an electron with minimal explicit selection.

background events resulting from the overlap of hadrons and photons may result in wider distributions of energy in $\eta$ as well as $\phi$.

$\Delta\eta_{in}$ Match in the $\eta$ direction between the electron track, at its point of closest approach to the vertex, with the vector formed between vertex and supercluster. If the track and the supercluster originate from the same particle, the match should be good. If the track and supercluster arise from two or more distinct particles that have momenta close to each other, the match will be worse.

$\Delta\phi_{in}$ Match in the $\phi$ direction between the electron track, at its point of closest approach to the vertex, with the vector formed betwen vertex and supercluster. The motivation for this cut is the same as the cut on $\Delta\eta_{in}$, but larger differences are allowed because the spread in $\phi$ of the electron's energy deposit in the calorimeter means that supercluster $\phi$ is a less precise measurement of the electron's $\phi$ angle.

$E_{seed}/p_{in}$ Ratio of the energy of the supercluster seed to the momentum of the track measured at its innermost point; a minimum value is required.

The CMS reconstruction code provides different levels of electron identification, namely Robust, Loose, and Tight. Robust identification is designed to be simple, easy to understand, and insensitive to expected uncertainties in detector alignment and calibration. It does not use electron classification, nor does it use

a cut on $E_{seed}/p_{in}$. Loose and Tight identification involves tighter cuts, cuts on $E_{seed}/p_{in}$, and electron classification. Electrons passing Tight identification are a subset of those passing Loose identification, which are in turn a subset of those passing Robust identification.

### 3.1.5 Electron Categorization for Selection

The quantites described in the previous section are correlated with the amount of bremsstrahlung emitted by an electron, as measured by $f_{brem} \equiv (p_{in} - p_{out})/p_{in}$, and $E/p_{in}$, the ratio of supercluster energy to inner track momentum. The quantities are also correlated to $f_{brem}$ and $E/p_{in}$ for fake electrons, despite the fact that $f_{brem}$ does not have the physical interpretation of bremsstrahlung radiation for fakes[4]. This fact can be exploited by classifying electron candidates according to $f_{brem}$ and $E/p_{in}$, and using different cut values for different classifications to improve selection efficiency and background rejection. The categorizations used are the following[5]:

1. Low brem electrons: $0.8 < E/p_{in} < 1.2$, $f_{brem} < 0.06$ (barrel), $f_{brem} < 0.1$ (endcap)

2. High brem electrons: $0.8 < E/p_{in} < 1.2$, $f_{brem} > 0.06$ (barrel), $f_{brem} > 0.1$ (endcap)

[4]Amongst all particles produced in a hadronic collision, only electrons experience significant bremsstralung at energies below about 1 TeV; all other charged particles are too heavy.

[5]This categorization is similar but distinct from the categorization described in section 3.1.2.

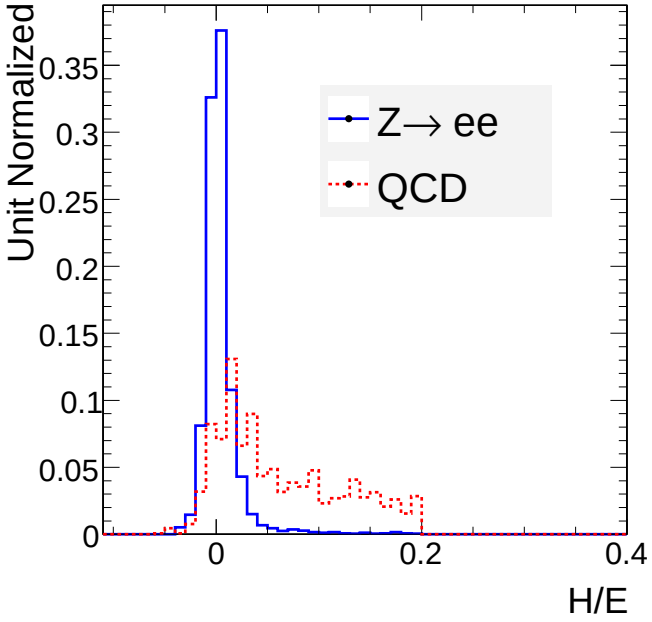


Figure 3.1: $H/E$ for real primary electrons (from $Z \rightarrow ee$) and electrons in QCD events including both instrumental fakes and electrons produced by the decay of hadrons.

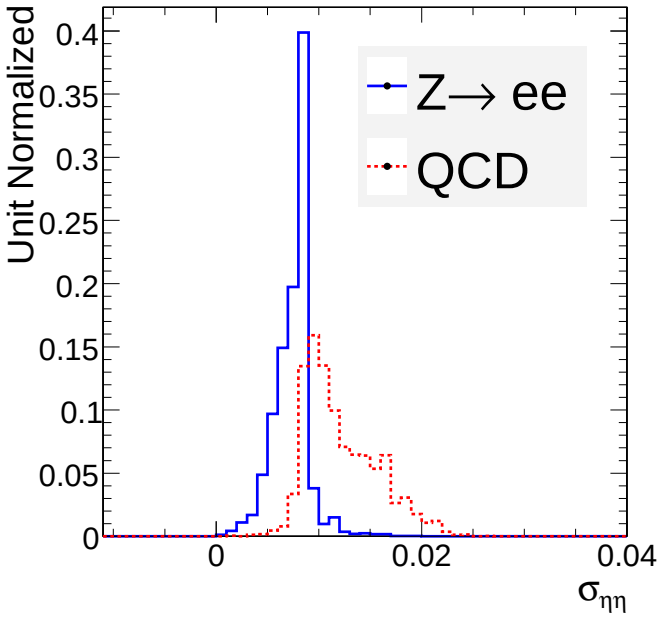


Figure 3.2: $\sigma_{\eta\eta}$ for real primary electrons (from $Z \rightarrow ee$) and electrons in QCD events.

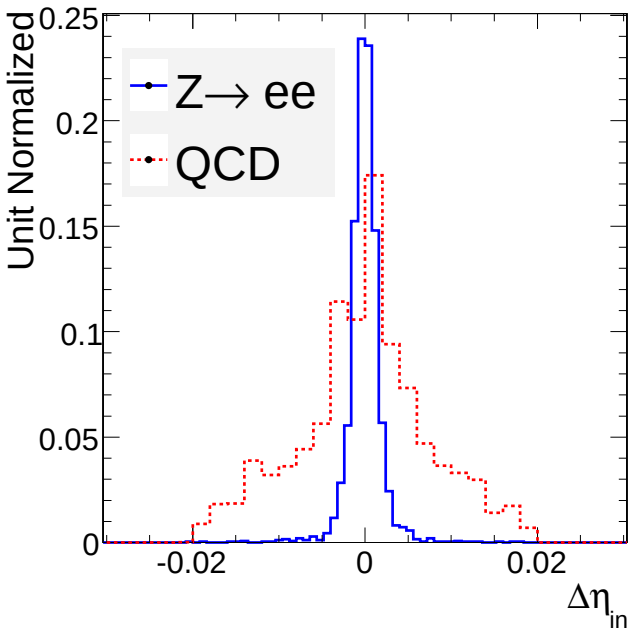


Figure 3.3: $\Delta\eta_{in}$ for real primary electrons (from $Z \rightarrow ee$) and electrons in QCD events.

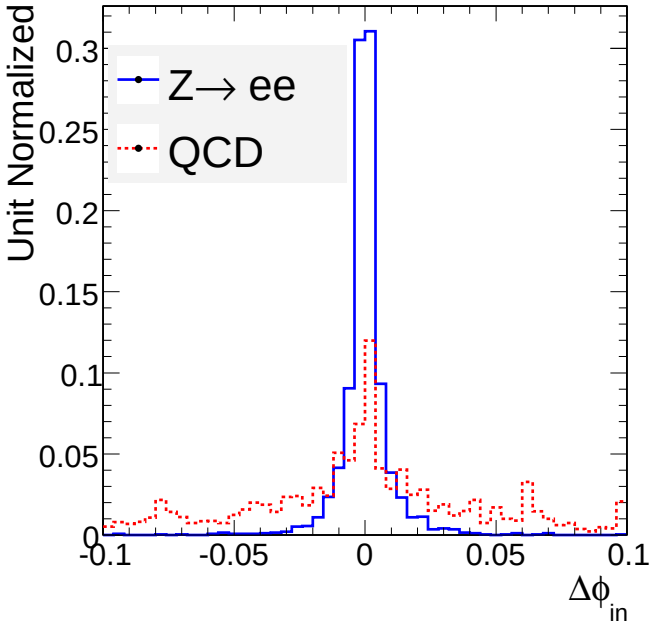


Figure 3.4: $\Delta\phi_{in}$ for real primary electrons (from $Z \rightarrow ee$) and electrons in QCD events.

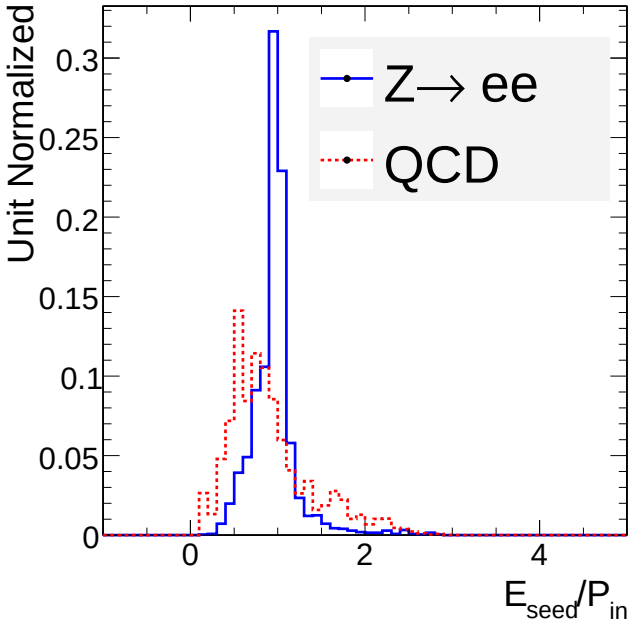


Figure 3.5: $E_{seed}/p_{in}$ for real primary electrons (from $Z \rightarrow ee$) and electrons in QCD events.

3. Badly measured track: $E/p_{in} < 0.8$ or $E/p_{in} > 1.2$

## 3.2 Photons

Photon reconstruction begins, as does electron reconstruction, with the reconstruction of an ECAL supercluster. As with electrons, almost all energy from an incident single photon is deposited in a 3x3 array of crystals, but reconstruction of primary photons is complicated by their material interactions (photon conversions) in the inner tracker. Photon energy is taken from the supercluster energy. The photon momentum direction is taken as the vector between the collision vertex and the supercluster centroid position, unless the photon is determined to be

unconverted, based on a requirement that an unconverted photon should have the ratio of energy in a 3x3 array of clusters containing the supercluster seed, to the total supercluster energy, greater than 0.93. For unconverted photons, the direction is taken as the vector between the vertex and the supercluster seed position, taking into account the *shower max depth* in the ECAL of an unconverted photon (a small correction). The shower max depth is the depth in the ECAL crystals at which, on average, the peak energy deposition occurs.

### 3.2.1 Converted Photons

Primary photons can convert in the material of the inner tracker to a pair of opposite-charge electrons, with a probability of about 50% averaged over the entire $\eta$ range of the tracker. In some cases it is possible to reconstruct the daughter electrons. Conversion finding in CMS is described in Ref. [12]. Tracking of electrons from photon conversions uses an algorithm that first searches for hits in the region of the silicon tracker nearest to the ECAL supercluster, and then seeks to associate additional hits at progressively smaller radii. The smallest radius for which a hit can be associated is assumed to be where the conversion occurred and a second conversion track is then sought, working back toward larger radii.

### 3.2.2 Photon Vertex Correction

An important aspect of photon reconstruction is the correction of the photon momentum direction for the location of the primary vertex. The *luminous region*[6] of the LHC at the CMS collision point will extend several centimeters in the Z coordinate direction (see figure 3.6). Photons aren't measured by the tracking detector, so the point of origin of the photon must be inferred from the other objects in the same event. The small correction of the photon vertex from the nominal interaction point at the center of the luminous region to the true vertex results in a very small correction to the photon's transverse energy measurement. However, the change in the photon direction can be significant during the calculation of tracker isolation (discussed in section 3.3).

The natural correction to the photon vertex is to place it at the *primary vertex*, the measured location of the proton-proton collision. Primary vertex candidates are those points along the beamline from which multiple tracks have a consistent origin. It is possible for more than one primary vertex candidate to be reconstructed, because multiple proton-proton collisions can occur in a single LHC bunch crossing[7], because of errors in track measurement, and because of the presence of long-lived particles that create displaced vertices when they decay some distance from the primary interaction. Therefore the best primary vertex

[6]Luminous region: the region where proton collisions occur.

[7]Multiple proton collisions can occur in a single LHC bunch crossing, but the probability of more than one "interesting" collisions in a single bunch crossing is negligible

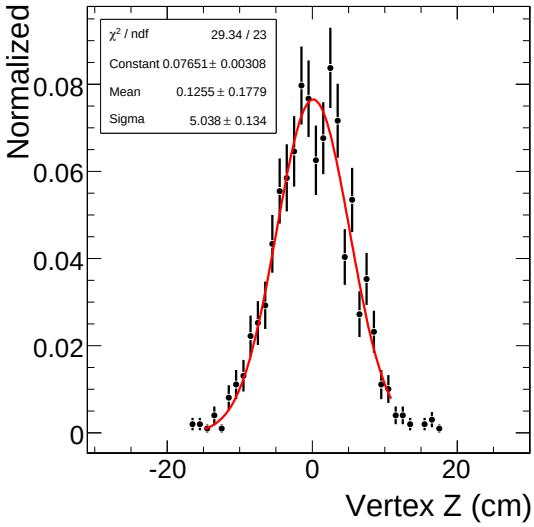


Figure 3.6: Z coordinate of highest-sum-$p_T$ reconstructed primary vertex in photon events generated with photon $E_T$ greater than 300 GeV.

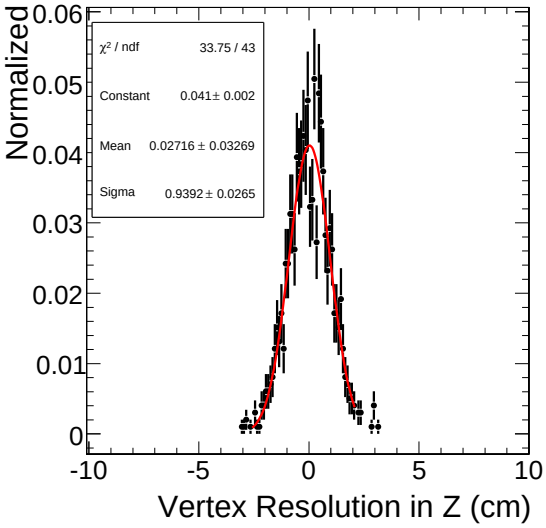


Figure 3.7: Resolution of reconstructed Z coordinate of highest-sum-$p_T$ primary vertex in photon events generated with photon $E_T$ greater than 300 GeV.

is selected according to the criteria that it should have the highest sum $p_T$ of associated tracks.

Figure 3.7 shows the resolution of the reconstructed vertex Z position. The resolution is well-described by a gaussian of mean zero and standard deviation about 1 cm, out to about ±3 cm. The efficiency of finding a primary vertex in this manner is approximately 99% for energetic prompt-photon events, the main event type used in the analysis presented in chapter 5 of this thesis. The efficiency and position resolution of reconstructing the primary vertex depends on how many charged particles there are in the event; events with a smaller number of charged tracks have a slightly lower efficiency and slightly worse resolution of reconstructed primary vertices.

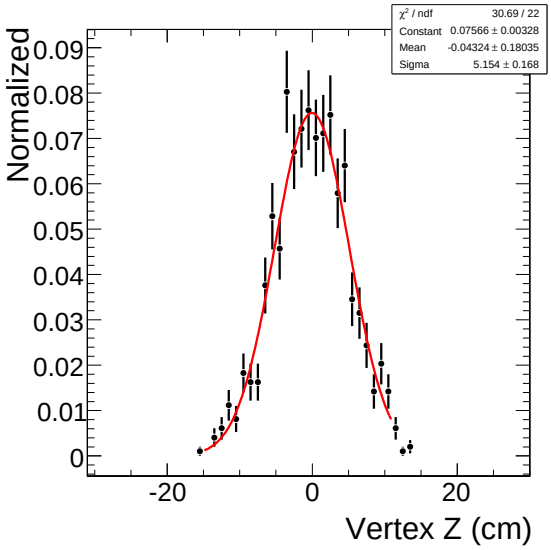


Figure 3.8: Z coordinate of highest-sum-$p_T$ reconstructed primary vertex in photon events generated with photon $E_T$ between 120 and 180 GeV.

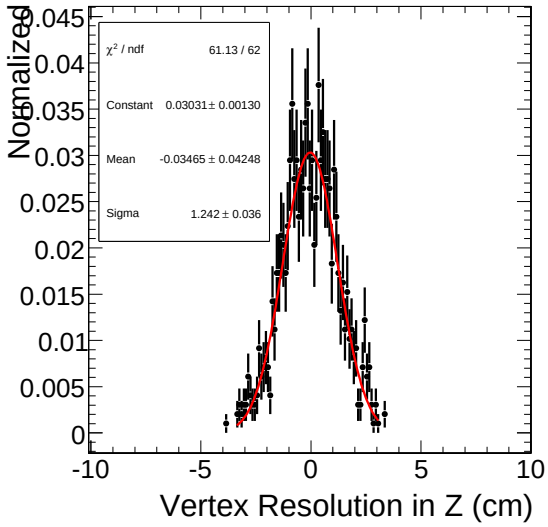


Figure 3.9: Resolution of reconstructed Z coordinate of highest-sum-$p_T$ primary vertex in photon events generated with photon $E_T$ between 120 and 180 GeV.

## 3.3 Isolation

Even after applying the identification cuts decribed in section 3.1.4, samples of reconstructed electron objects are dominated by electron candidates whose origin is in hadronic jets, not prompt electrons from decay of W and Z bosons which are of greater interest for the vast majority of CMS analyses using electrons. Similarly, samples of photon objects are dominated by photons which originate from neutral meson decay in hadronic jets. *Isolation* is a further level of selection that must be applied. Isolation is a measure of the amount of energy that is nearby the electron or photon but is not actually from the electron or photon itself.

### 3.3.1 Tracker Isolation

Tracker isolation is defined as the sum of the $p_T$ of tracks inside a cone about the electron or photon's momentum direction, excluding the track of the electron itself. The most common way to exclude the electron track is to have an *inner veto cone*, inside of which tracks are ignored. The inner veto cone is also useful in photon isolation, to avoid vetoing photons that convert early in the tracker and therefore may have conversion tracks reconstructed in the standard track collection[8]. A typical outer cone size is 0.3 in $\eta - \phi$ space, and a typical inner cone size is 0.02.

For tracks to be considered in the calculation, they are required to originate from the same location in the Z coordinate direction as the electron or photon to be isolated. This requirement is made in order to reject tracks that come from different collisions within the same bunch crossing. In the analysis presented in this thesis, tracks are required to originate within 0.1 cm of the electron vertex Z coordinate, and within 1 cm of the photon vertex Z coordinate; the difference stems from the fact the electron's vertex is known to high precision through its own track, whereas the photon's vertex, being inferred from the event's primary vertex (see section 3.2.2), is known with less precision. The primary vertex resolution can range as high as about 1.25 cm for events with photons of the energy relevant

[8]This is likely to be a small effect since the standard track reconstruction, in contrast to track reconstruction aimed at finding photon conversions, requires hits in the pixel detector, meaning that only very early conversions could be reconstructed in the standard track collection.

to the analysis presented in chapter 5 of this thesis. Thus the 1 cm cut used here, which was taken from the software default, is slightly too small, decreasing the rejection power of photon tracker isolation cuts.

### 3.3.2 Calorimeter Isolation

Calorimeter isolation is the sum of the energies of calorimeter towers in a cone about the electron or photon's direction, excluding those deposits belonging to the electron or photon itself. The exclusion of the electron or photon's own energy is implemented by subtracting the energy of the supercluster. Calorimeter isolation can be calculated and used separately for the ECAL and the HCAL, applying separate cuts for the two quantities and in principle allowing for more effective selection of signal and rejection of background. In most cases the two are combined for simplicity. A typical cone size for calorimeter isolation is 0.3. Electrons or photons with energy above about 100 GeV will have electromagnetic showers that extend past the depth of the ECAL crystals, resulting in some of the electron's energy being deposited in the HCAL. For that reason, *relative* calorimeter isolation is used for electrons and photons above 100 GeV $E_T$ in analyses presented in this thesis; isolation cuts are defined as a percentage of the electron or photon $E_T$, avoiding loss of efficiency due to leakage. A more sophisticated solution that is sometimes used to resolve the leakage problem is to define an inner cone (size typically 0.1 in $\eta - \phi$) inside of which an HCAL isolation sum is calculated

separately. Cuts can then be applied relative to the electron or photon $E_T$ for only this inner HCAL cone, factorizing the effect of leakage and allowing tighter absolute cuts to be used in an outer annulus where leakage is presumed to have no effect. A difficulty with this latter technique is that the spatial distribution of energy deposited in the ECAL from a photon or electron don't typically don't match well to an $\eta\phi$ cone; the energy is often more spread out in $\phi$.

## 3.4 Hadronic Jets

A single quark or gluon scattered in the primary collision becomes multiple hadrons in the fragmentation process described earlier. An accurate measure of the quark or gluon's energy is achieved by clustering the energy deposits in the calorimeters into jets. One basic technique of jet clustering algorithm is to add up all the energy in a cone about some specified direction. The jet reconstruction algorithm used by CMS is based on that technique. It is called the *iterative cone* algorithm, and is described below.

The HCAL readout is segmented in 0.087 x 0.087 rectangles in $\eta - \phi$up to $|\eta| < 1.6$, and 0.17 x 0.17 rectangles thereafter up to $|\eta| < 3.0$. The ECAL readout segmentation is much finer. Correspondingly, calorimeter *towers* are formed by combining HCAL readout cells with the ECAL readout cells that cover approximately the same area in $\eta - \phi$ space. Towers are combined into jets as follows.

1. The calorimeter tower with the largest $E_T$, and with $E_T$ above a certain threshold, defines the direction of a cone of radius 0.5 in $\eta - \phi$ space.

2. Transverse energies of towers within the cone are added to the seed tower $E_T$ to form a *proto-jet*. Only towers having energy above a noise threshold are included. The proto-jet direction is computed as the $E_T$-weighted sum of the directions of its constituent towers.

3. The proto-jet direction is used to define a new cone, and an updated proto-jet is computed. This process is repeated until the proto-jet direction changes by less than 0.001 in $\eta$ and less than 0.001 in $\phi$, or until 100 iterations have been completed, whichever comes first.

4. After iteration is completed, the proto-jet is promoted to jet, and calorimeter towers belonging to the jet are removed from consideration as seeds or constituents of other jets.

5. Steps 1-4 are repeated until there are no more towers above the seed threshold.

Once the jet energy is computed, it is not however equal to the energy of the originating quark or gluon, which is the physical quantity of interest. Instead, it is complicated by several factors:

**Electromagnetic fraction** The response of a calorimeter is not equal with respect to hadronic and electromagnetic interactions. To a large degree that is

because in hadronic interactions some fraction of energy is lost to low-energy nuclear effects which do not result in an ionization or scintillation energy deposit. An electromagnetic energy deposit in a hadronic shower occurs due to the copious production of $\pi_0$ mesons, which decay to photons before they have a chance for further hadronic interaction. The relative fraction of energy that goes to $\pi_0$ production in a hadronic shower is not precisely known, depends on the energy of the hadron that initiates the shower, and fluctuates event by event. Therefore the unequal calorimeter response to hadronic and electromagnetic interaction results in calorimeter nonlinearity of response as a function of incident particle energy, energy scale uncertainty, and poor energy resolution.

**Out-of-cone** Calorimeter jets are formed from fixed-size cones centered around high-energy seeds. Low-momentum charged particles may be bent away from the high-energy particles to which they should be associated.

**Cracks and dead areas** The CMS calorimeter system is designed to be a few degrees off-pointing, so that cracks between calorimeter towers don't project directly to the nominal interaction point of CMS. Nevertheless because of the spread of the true interaction region it is possible for particles to go through cracks in the HCAL, depositing reduced amounts of energy. Furthermore it is inevitable that some detector elements will stop functioning.

In order to compensate for these effects, corrections can be applied to raw mea-

sured jet energies. Jet energy corrections vary as a function of jet $E_T$ and $\eta$. They are derived by balancing energy in di-jet events, in photon plus jet events, and in Z plus jet events where the Z decays to muons or electrons. Corrections extracted from Monte Carlo simulated data are also used. Photon and Z events are especially useful because the energy measurements of photons, electrons and muons aren't subject to as much uncertainty as jet energy measurements. The methodology planned by the CMS experiment for the correction of jet energies taking into acccount these factors is described in Ref. [13].

## 3.5 MET

The summed momemta of all particles involved in a collision must be the same after the collision as before. Since there is no momentum in the transverse plane before the collision, there is also no net momentum in the transverse plane after the collision. Neutrinos and other weakly-interacting particles are not directly detected by CMS, but they do carry momentum that contributes to the balance. Thus weakly-interacting particles are detected by an imbalance of measured momentum. This imbalance of momementum is called missing transverse energy (MET). However MET can also arise from detector effects; any mismeasurement of objects will result in fake MET. Of particular importance for MET is the $E_T$ resolution of hadronic jets.

# Chapter 4

# Electron-Enriched Monte Carlo Samples

The most frequently occurring processes that result from proton-proton collisions are quark and gluon scattering processes, referred to as *QCD production*. Such processes have very low but non-zero probabilities to result in signatures that fake prompt electrons and have production cross-sections several orders of magnitude higher than those of the interesting processes that produce real prompt electrons, such as $t\bar{t}$, W+jets, or hypothetical SUSY signals. One would like to use a Monte Carlo data sample, created using event generators and the CMS detector simulation, to study the rate at which QCD fakes signal events involving electrons, to study electron selection cuts, and to test data-driven methods of estimating backgrounds that involve fakes. Because of the very high cross sec-

tions and low electron fake rates of QCD processes, it is not feasible to simulate a large enough sample of events in order to get a sufficient number of events with fakes; most of the simulation time is wasted on events that do not result in fake electrons.

One way to deal with this problem is to produce a Monte Carlo sample enriched in electrons fakes. To produce an electron-enriched sample means to predict at an early stage whether an fake electron will likely occur, and only run the full detector simulation (which takes up the vast majority of the time in Monte Carlo event production) if an electron fake is likely. The production of Monte Carlo simulated data is conveniently separated into an *event generator* stage and a detector simulation stage. An event generator, such as PYTHIA [14] or MadGraph [15], simulates a proton collision and produces a list of particles and their corresponding momenta to be used as input to a GEANT [16] detector simulation. The event generator stage is typically orders of magnitude faster than the detector simulation stage. Thus a convenient way to produce an enriched sample is to insert a filter that operates on the list of particles produced by the event generator, and only passes events on to the detector simulation if they meet certain criteria.

In order for such an enriched sample to be useful, it must be efficient in selecting events which will have reconstructed fake electrons, it must be efficient in rejecting events that won't have reconstructed fake electrons, and it must not have any important biases. The requirement to avoid important biases means that if the

filter is not 100% efficient in selecting events with fake electrons, the events it misses must not form an important subset of fake-electron events. For example it would be an important bias if the filter selected 90% of electron-fake events, but the 10% it rejected would be the most likely to pass subsequent analysis-level cuts, e.g. isolation, calorimetric MET, or number of jets in the event.

To make an efficient electron-enriching filter, it is necessary to understand the sources of electron fakes. Furthermore understanding the sources of electron fakes also is beneficial for other reasons, such as improving electron selection cuts or creating data-driven background estimation methods. A study of electron fakes using the CMS detector simulation is presented in this chapter.

## 4.1 Sources of Electron Fakes

To find the source of electron fakes the event record of the simulation was examined. The simulation event record stores particles' energy deposits in all *sensitive volumes*[1] of the detector. The simulation also records the complete passage of a particle through the tracker, including any interactions a particle has, and its momenta before and after these interactions. Secondary particles produced inside the tracker are also traced in this way[2].

[1]Sensitive volume: volume designated in the simulation for recording of energy deposit, corresponding to a volume in the physical detector from which some fraction of the energy deposit is read out, e.g. a scintillator tile in the HCAL.

[2]A detailed trace of particles in this fashion is not possible in the calorimeters because of the large number of interactions and secondary particles produced, i.e. the showering.

The set of ECAL crystals belonging to ECAL superclusters of reconstructed electrons from QCD events were examined[3], and the energy deposits were separated according to the categories listed below, based on the *original parent* of the particle making the deposit. The original parent is the stable particle produced by the event generator and passed to the detector simulation, which through subsequent interaction, including possible decay to other particles, deposits all its energy in the detector. For example, if an electron radiated a bremstralung photon that subsequently converted to an electron-positron pair, and the conversion electron contributed energy to the ECAL cluster of a reconstructed electron, the first electron would be the original parent.

All possible original parent particles are assigned to one of the following categories:

1. Electrons
2. Photons (predominantly coming from decay of $\pi_0$ and $\eta$ mesons)
3. Charged Pions
4. Other hadrons stable relative to the time needed to reach the ECAL, eg. protons, neutrons.
5. Hadrons unstable relative to the time needed to reach the ECAL, but with

[3]2869 electrons from QCD $\hat{p_T}$ bins 30-120 GeV were used. These electrons passed Robust electron identification cuts (see section 3.1.4), and loose isolation cuts of 3 GeV/c tracker isolation in a cone of 0.2 and 6 GeV calorimeter isolation in a cone of 0.2, and have $E_T$ of at least 20 GeV.

$c\tau$ longer than 1 cm.

6. other, unidentifiable

Figure 4.1 shows the distribution of the fraction of electrons' energy coming from the various types of parent particle. On average, fake electrons passing Robust identification cuts get over half their energy from photons, with smaller fractions from real electrons, charged pions, and other stable hadrons.[4] The large contribution of photon energy suggests that photon conversions are a source of fake electrons. To check that this is true, the reconstructed tracks of reconstructed electrons were matched to charged, stable generator particles with $E_T$ of at least 5 GeV/c. Figure 4.2 shows the angle between the reconstructed track at the vertex and the nearest stable, charged generator particle. Figure 4.3 shows the efficiency of matching the reconstructed electron's track to the generator particle within a given angle (this is just the integral of figure 4.2). Taking the criterion of an angle less than 0.05 as a "good match", about 85% of reconstructed electrons' tracks are matched to a generator particle. Figure 4.4 shows the ratio of the matching generator particle's $p_T$ to the reconstructed electron's track $p_T$, indicating that the tracks that match in angle also usually match in $p_T$. Recalling that photon conversion electron tracks are not part of the generator record[5], this shows that only up to about 15% of fake electrons could be photon conversions[6]. Some of this

[4]It should be noted that the composition of electron fakes depends strongly on the level of identification cuts used. The focus of this chapter is on electrons passing Robust identification cuts.

[5]Photon conversion electrons are secondary particles produced in the detector simulation

[6]The explicit requirement of matching hits in the pixel detector as part of electron reconstruc-

15% could also include electrons that involve fake tracks from pattern-recognition failures. The figure also shows the ratio of the matching generator particle's $p_T$ to the electron's $E_T$. The electron's $E_T$ is normally substantially higher than that of the generator particle, indicating that there are multiple energy sources contributing to most fakes. The matched generator particles are 60% charged pions, 15% charged kaons, 15% electrons, and about 5% protons, with a few percent taken up by other heavy hadrons.

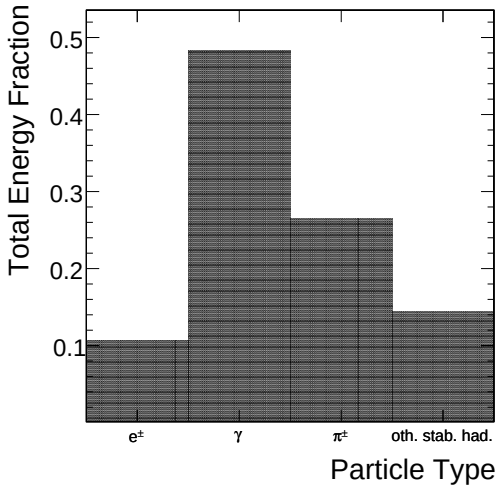


Figure 4.1: Fraction of electron's energy coming from different types of parent particle.

Figure 4.5 shows the histogram of the fraction of energy contributed to fake electrons by the most important particle types. In about 15% of cases, over 90% of the fake's energy comes from photons, providing more evidence for the hypothesis

tion reduces the number of conversion electrons reconstructed by the standard reconstruction sequence. Alternate electron reconstruction sequences that do not require matching pixel hits must explicitly veto photon conversions.

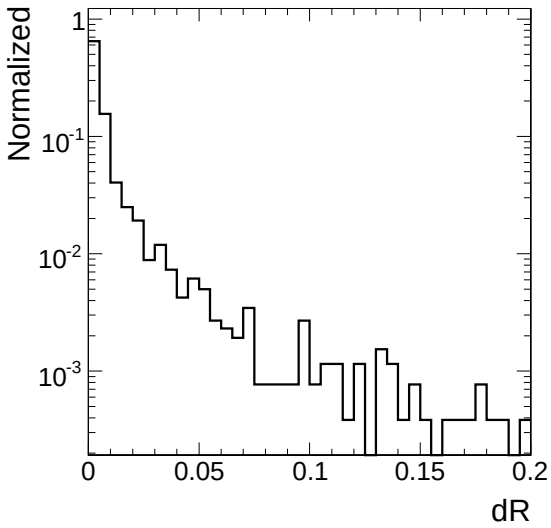


Figure 4.2: Angle $dR \equiv \sqrt{\Delta\phi^2\Delta\eta^2}$ between reconstructed electron track and nearest stable charged generator particle with $p_T$ of at least 5 GeV/c.

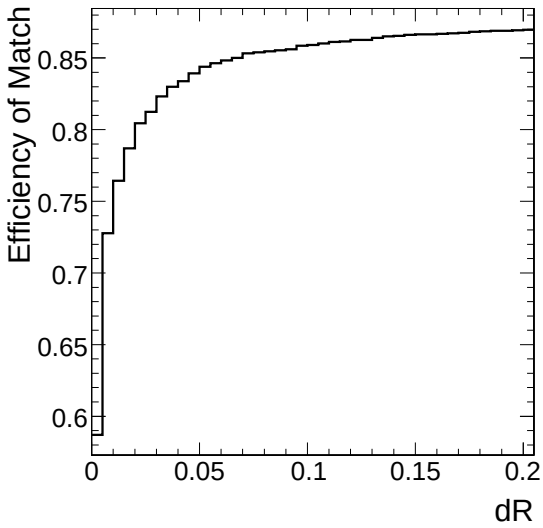


Figure 4.3: Efficiency of matching reconstructed electron track within specified angle to a stable charged generator particle with $p_T$ of at least 5 GeV/c.

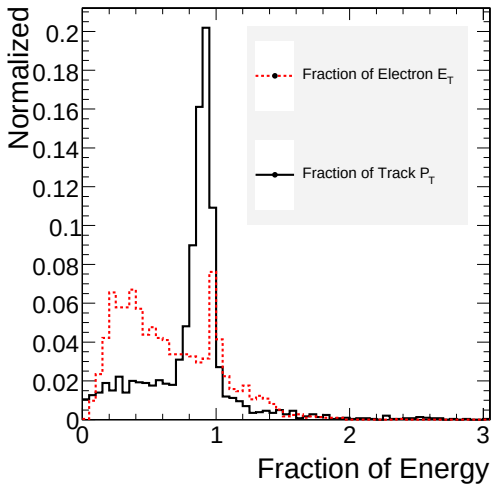


Figure 4.4: Ratio of the matching generator particle's $p_T$ to the reconstructed electron's track $p_T$, and ratio of the matching generator particle's $p_T$ to the electron's $E_T$, for the closest particle matching reconstructed tracks to within dR<0.05.

that about 15% of fake electrons come from photon conversions. Almost 10% of fakes take over 90% of their energy from charged pions. The fraction of fakes that take over 90% of their energy from real electrons is also about 10% (these real electrons come from decay of hadrons in jets).

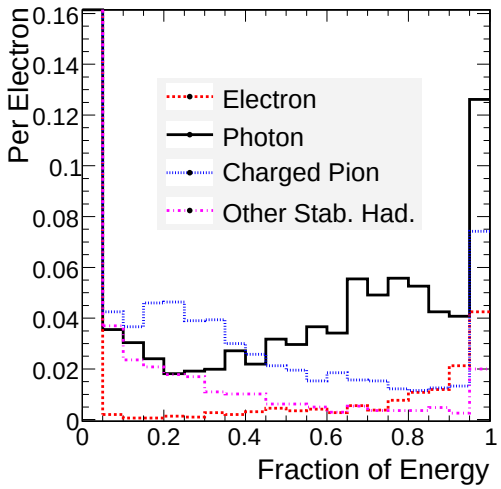


Figure 4.5: Distribution of fraction of reconstructed electron energy coming from different particle types.

Charged hadrons can contribute to fake electrons through two mechanisms. The first is when a charged hadron contributes a track to electron reconstruction, but much of the energy of the fake comes from photons. The second is when an isolated charged hadron deposits nearly all of its energy in the ECAL and so is reconstructed as an electron. Both processes require that the charged hadron deposits most of its energy in the ECAL and very little in the HCAL. This occurs due to rare flucuations in the hadronic showering process, including so-called charge exchange processes. A detailed overview of test beam measurements of

interactions of hadrons in the CMS calorimeter can be found in Ref. [17].

Figure 4.6 shows the distribution of $E_T$ deposited in the ECAL and reconstructed in superclusters by a sample of 20,000 single pions with $p_T$ of 30 GeV/c. The pions were generated with a uniform $\eta$ distribution in the range $|\eta| < 2.5$. The most probable supercluster energy recorded is zero; the ionization energy loss in the ECAL of a minimum-ionizing charged pion is only about 200 MeV, which is well below the 1 GeV seed crystal threshold for the creation of a supercluster. The figure also shows the $E_T$ of electrons which were reconstructed from these superclusters and passed the Robust electron identifications cuts. The electron $E_T$ is slightly higher than supercluster $E_T$ because of energy corrections applied during electron reconstruction. From the 20,000 single pions, 924 electrons with $E_T$ at least 20 GeV were reconstructed, corresponding to a fake rate of about 5% per pion.

## 4.2 Design of an Electron-Enriching Filter

The analysis presented above has implications for electron-enriching filtering. One of the most important is that since fakes frequently involve multiple particles, some clustering of energy is necessary. The clustering must include photons, electrons, and some hadrons. But it is also possible that a single charged particle (typically a pion) will result in a fake. The filter design includes two components to take into account these two cases. It was decided that the filtered sample

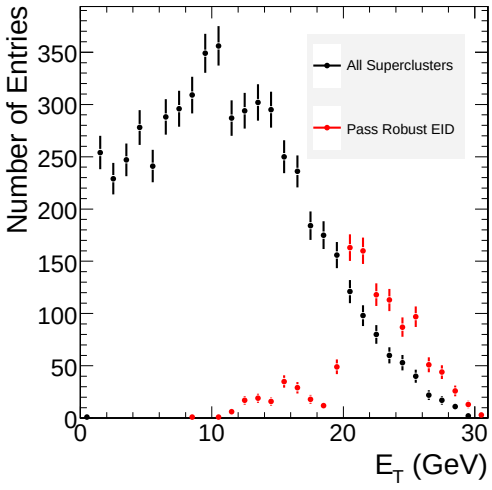


Figure 4.6: Distribution of transverse energy of superclusters from charged pions. All reconstructed superclusters are shown in black. The $E_T$ of electrons that are reconstructed from the pions and pass the Robust electron identification cuts are shown in red. 20,000 single pions are represented in the distribution. Most pions do not result in a reconstructed supercluster.

should focus on isolated electrons, in order to reduce the filter pass rate as much as possible; therefore the filter design uses some isolation criteria that are meant to simulate isolation at the reconstruction level. In both components of the filter, the curvature of charged particle trajectories in a 4T magnetic field is taken into account.

### 4.2.1 Isolated Charged Particle Filter

The first filter component targets events in which a single particle results in an electron fake. It selects events having an electron, charged pion, or charged kaon that is isolated from other particles in the event. The particle is required to

have $E_T$ of at least 20 GeV. Isolation variables are constructed from the generator particles in such a way as to imitate track isolation and calorimeter isolation variables in fully-simulated and reconstructed events. *Tracker-like* isolation is computed as the sum $E_T$ of all other charged particles with $E_T$ of at least 2 GeV in a cone, and *calorimeter-like* isolation is computed as the sum $E_T$ of charged and neutral generator particles other than photons, with $E_T$ of at least 2 GeV. The motivation for excluding photons from the isolation sum is the assumption that they would frequently be included in the supercluster of the charged particle. Also, such events would likely be selected by the clustering filter described in the next section. Moreover it should be noted that by removing photons from the isolation calculation, the pass rate of the filter on spurious events might be increased, but the efficiency on desired events will not be decreased. The tracker-like and calorimeter-like isolation quantities are required to be less than 4 and 7 GeV, respectively, with a cone size of 0.1 in $\eta - \phi$ space.

### 4.2.2 Clustering Filter

The second filter component targets events where multiple particles collectively result in an electron fake. This filter clusters the $E_T$ of generator particles in a rectangular region that is elongated in $\phi$. That shape follows from the behavior of real electrons in the detector. As discussed in detail in section 3.1.1, electrons often emit bremsstralung radiation in the tracker, and because of the

high magnetic field, the corresponding energy deposits are spread out in $\phi$. The ECAL superclustering algorithm is designed to collect all of this energy. Even in fake electron events, where no bremsstralung radiation occurs, superclusters tend to be elongated in $\phi$. Figures 4.7 through 4.10 show the distribution of the widths of superclusters from electron fakes. In the barrel, typically about 90% of the energy is collected within an $\eta$ window of about 0.04, and about 90% of the energy is collected within a $\phi$ window of about 0.45. In the endcap, the cluster shape is more symmetrical: 90% of cluster energy is collected within a region of about 20-25 cm in the X direction, and also 90% of the energy is collected in a region about 20-25 cm wide in the Y direction. Following from these figures, the filter uses a cluster size of 0.4 in $\phi$ by 0.06 in $\eta$ in the ECAL barrel region ($|\eta| < 1.478$), and a circular region of radius 15 centimeters projected onto the location of the ECAL endcap, in the endcap region ($1.478 < |\eta| < 2.4$). The reader may wonder why the filter's cluster width is not chosen to be large enough to correspond to a containment of 95% or more of the typical reconstructed cluster. The reason is that strong tension exists between increasing the filter efficiency for events that will have reconstructed electrons, and keeping the filter pass rate low for events that won't have reconstructed electrons. Filter parameters had to be chosen to balance these two conflicting constraints.

The cluster is required to contain an electron or photon "seed" with $E_T$ of at least 5 GeV, on which the cluster is centered, and to pass cuts requiring tracker-

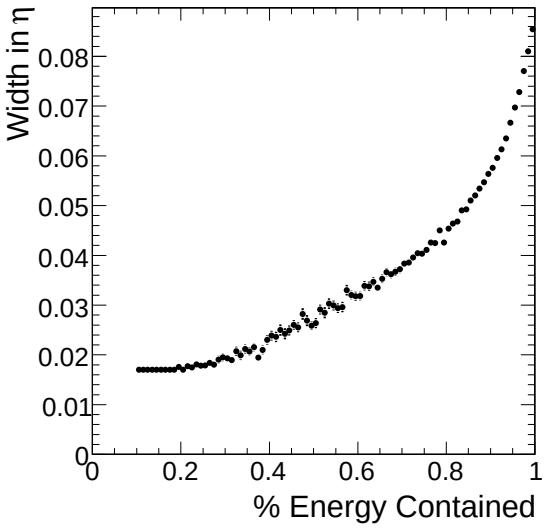


Figure 4.7: Distribution of energy deposition in barrel superclusters in the $\eta$ direction.

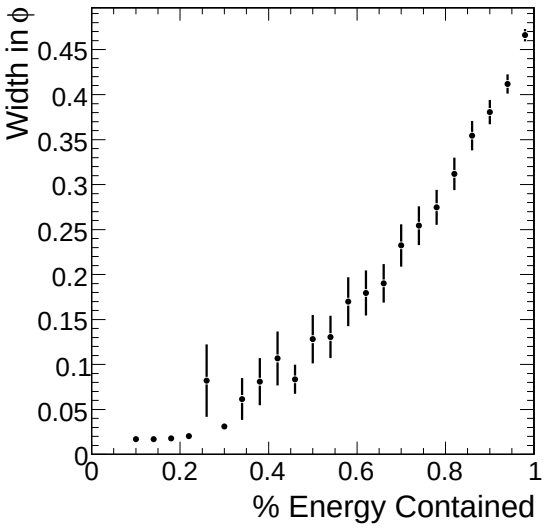


Figure 4.8: Distribution of energy deposition in barrel superclusters in the $\phi$ direction.

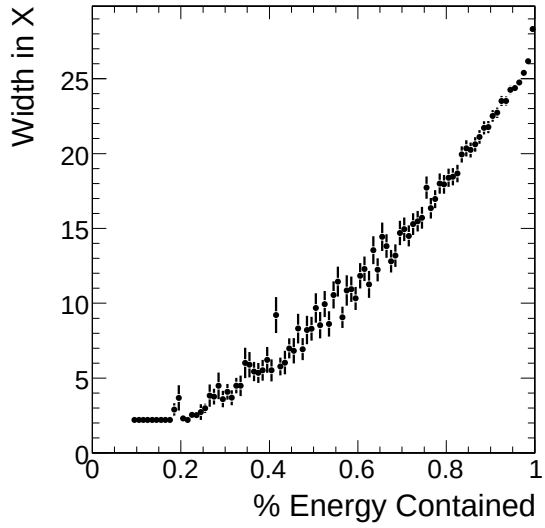


Figure 4.9: Distribution of energy deposition in endcap superclusters in the X direction.

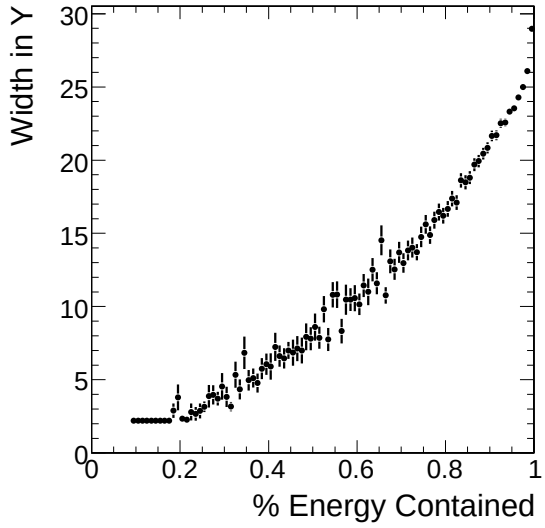


Figure 4.10: Distribution of energy deposition in endcap superclusters in the Y direction.

like isolation less than 5 GeV and calorimeter-like isolation less than 10 GeV in a cone of 0.2. The isolation quantities are computed slightly differently than in the isolated charged particle filter described in the previous section. Tracker-like isolation is computed as the sum $E_T$ of charged stable generator particles that fall outside the cluster window, but inside a cone of size 0.2. Calorimeter-like isolation is computed as the sum of stable charged and uncharged particles other than photons outside the cluster window but inside a cone of 0.2. Additionally they are required to pass a *simulated H/E* cut. Simulated H/E is computed as the ratio of the $E_T$ of hadronic particles other than charged pions and kaons, impinging on the same detector area as the cluster, divided by the $E_T$ of the cluster itself. The cut requires simulated H/E less than 0.5.

### 4.2.3 B/C→e Filter

A separate filter targets electrons that come from decay of b and c hadrons. Decay of b or c hadrons to electrons through multiple steps (e.g. D→K→e) are included. Since in this case the electrons are not fake, and the complete decay chain is recorded in the generator event record, filter design is trivial. Monte Carlo samples produced with this filter and the enriching filters described above are deliberately made to have zero overlap by including a veto of events passing this filter in the production of events using the enriching filters described above.

### 4.2.4 Filter Performance

These filters were used to produce Monte Carlo samples in the CMS Spring08 Production. In order to verify the efficacy of the filters, they were run on unfiltered QCD samples from the same production.

Figures 4.11 and 4.12 show the $E_T$ and $\eta$ distributions of all electron candidates, and electron candidates in events passing the OR-wise combination of the filters described above. The electron candidates were required to have $E_T$ of at least 20 GeV and to pass the PTDR Loose identification cuts described in Ref. [9]. That choice of cuts is similar to but distinct from the Robust identification cuts used in the analysis of fake electrons in the detector simulation described earlier in this chapter, and was chosen for historical reasons. Figures 4.13 and 4.14 show the same thing, but for electrons passing Tight identification cuts, which is currently seen as a more appropriate set of electron identification cuts for single-electron analyses. Both sets of electrons are required to pass isolation cuts with a cone size of 0.2. It is apparent that filter performance is better for the Tight cuts than the Loose cuts. Using Tight cuts decreases the rate of electron candidates passing identification and increases the fraction that come from real electrons and from photon conversions; events containing such objects (high-energy electrons or photons) are very likely to pass the filter. It is also apparant that efficiencies are higher for reconstructed electrons with $E_T$ above 30 GeV/c, compared to $E_T$ between 20 and 30 GeV. That is because particle production in QCD events falls

steeply with increasing $E_T$. Filter $E_T$ thresholds are set to make a compromise between the competing requirements of high filter efficiency on events that will reconstruct an electron candidate, and low filter pass rates on events that won't; this compromise has a bigger effect on lower-$E_T$ electron fakes. Tables 4.1 and 4.2 give a more complete characterization of the filter performance. Efficiencies are shown for different $E_T$ thresholds (20 and 30 GeV) and for different isolation cuts. "Loose" isolation refers to a cut of 3 GeV/c on tracker isolation and 6 GeV on calorimeter isolation. "Tight" isolation refers to a cut of 1 GeV/c on tracker isolation and 3 GeV on calorimeter isolation.

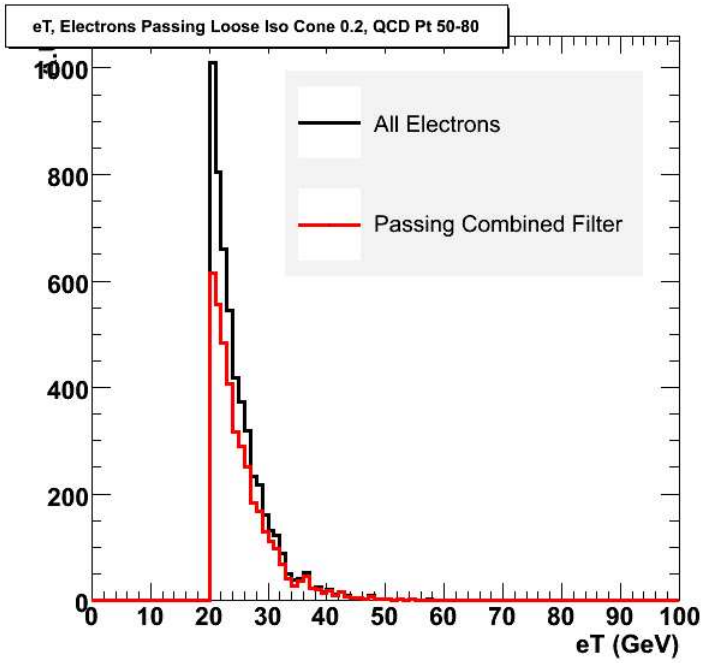


Figure 4.11: $E_T$ distribution of electrons passing PTDR Loose ID cuts.

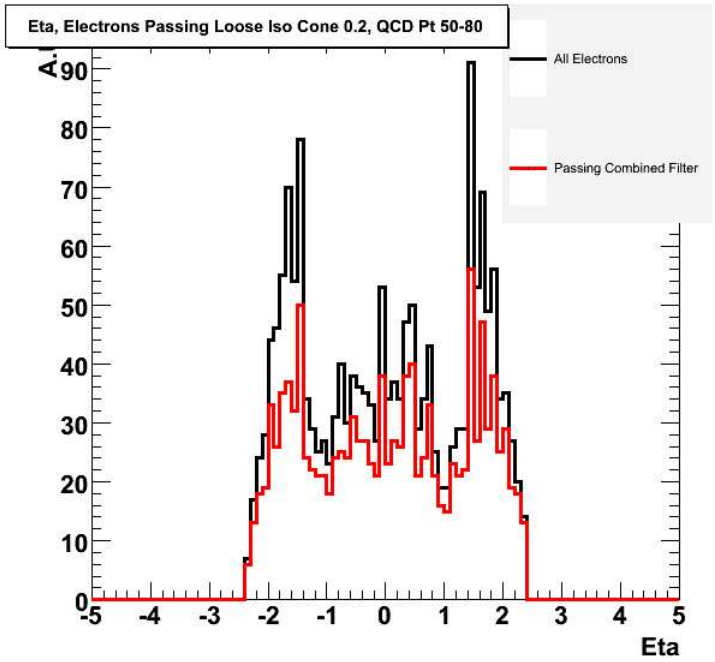


Figure 4.12: The distribution of $\eta$ for electrons passing PTDR Loose ID cuts.

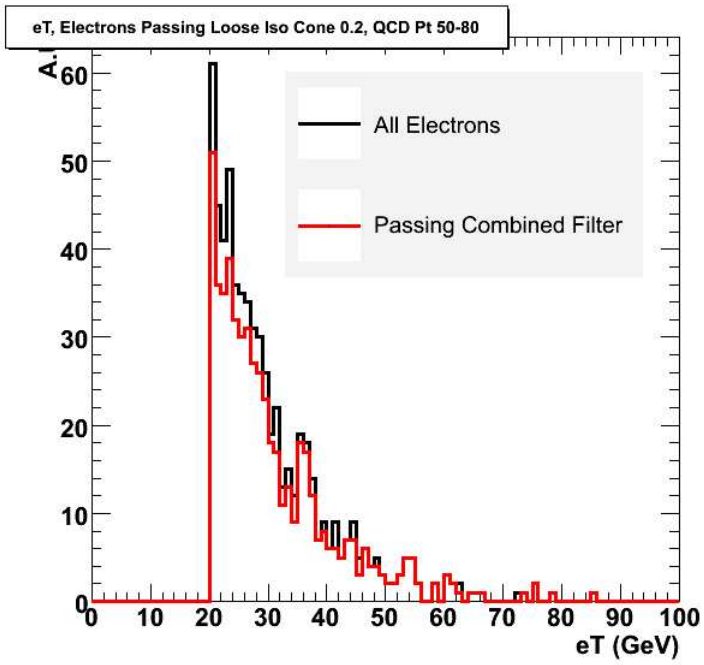


Figure 4.13: $E_T$ distribution of electrons passing Tight cuts.

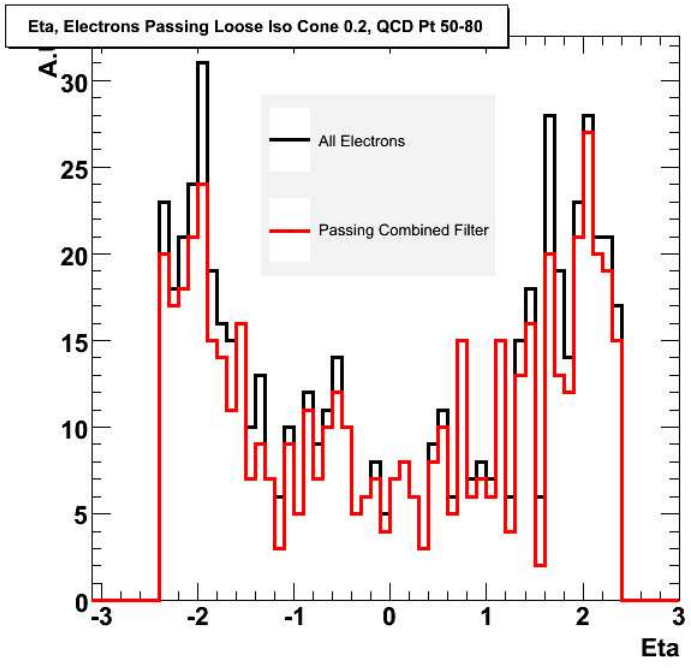


Figure 4.14: The $\eta$ distribution of electrons passing Tight cuts.

| **QCD $\hat{P}_T$** | **1** | **2** | **3** |
|---|---|---|---|
| combined filter, electrons $E_T>$ 20 | | | |
| 30-50 | 0.75 | 0.73 | 0.77 |
| 50-80 | 0.75 | 0.71 | 0.78 |
| 80-120 | 0.75 | 0.70 | 0.73 |
| combined filter, electrons $E_T>$ 30 | | | |
| 30-50 | 0.83 | 0.82 | 0.84 |
| 50-80 | 0.78 | 0.76 | 0.83 |
| 80-120 | 0.76 | 0.74 | 0.76 |
| em-enrich filter, electrons $E_T>$ 20 | | | |
| 30-50 | 0.68 | 0.65 | 0.71 |
| 50-80 | 0.67 | 0.59 | 0.69 |
| 80-120 | 0.63 | 0.56 | 0.64 |
| em-enrich filter, electrons $E_T>$ 30 | | | |
| 30-50 | 0.79 | 0.77 | 0.82 |
| 50-80 | 0.73 | 0.69 | 0.79 |
| 80-120 | 0.67 | 0.64 | 0.69 |
| bcefilter, electrons $E_T>$ 20 | | | |
| 30-50 | 0.069 | 0.086 | 0.058 |
| 50-80 | 0.081 | 0.11 | 0.084 |
| 80-120 | 0.12 | 0.13 | 0.082 |
| bcefilter, electrons $E_T>$ 30 | | | |
| 30-50 | 0.041 | 0.056 | 0.019 |
| 50-80 | 0.051 | 0.075 | 0.043 |
| 80-120 | 0.089 | 0.11 | 0.070 |

Table 4.1: Filter effiencies on electrons passing PTDR Loose electron ID. Columns show efficiencies for electrons passing different levels of isolation selection, as described in the text: 1) loose cuts in a cone of 0.3, 2) loose cuts in a cone of 0.2, 3) tight cuts in a cone of 0.3.

| **QCD** $\hat{P}_T$ | **1** | **2** | **3** |
|---|---|---|---|
| combined filter, electrons $E_T>$ 20 | | | |
| 20-30 | 0.91 | 0.87 | 0.91 |
| 30-50 | 0.90 | 0.88 | 0.92 |
| 50-80 | 0.86 | 0.86 | 0.92 |
| 80-120 | 0.87 | 0.86 | 0.91 |
| combined filter, electrons $E_T>$ 30 | | | |
| 20-30 | 0.82 | 0.81 | 0.91 |
| 30-50 | 0.93 | 0.90 | 0.96 |
| 50-80 | 0.87 | 0.89 | 0.92 |
| 80-120 | 0.86 | 0.82 | 0.86 |
| em-enrich filter, electrons $E_T>$ 20 | | | |
| 20-30 | 0.66 | 0.64 | 0.71 |
| 30-50 | 0.67 | 0.62 | 0.69 |
| 50-80 | 0.58 | 0.53 | 0.69 |
| 80-120 | 0.54 | 0.48 | 0.63 |
| em-enrich filter, electrons $E_T>$ 30 | | | |
| 20-30 | 0.65 | 0.67 | 0.73 |
| 30-50 | 0.82 | 0.78 | 0.88 |
| 50-80 | 0.72 | 0.69 | 0.82 |
| 80-120 | 0.62 | 0.57 | 0.69 |
| bcefilter, electrons $E_T>$ 20 | | | |
| 20-30 | 0.24 | 0.23 | 0.20 |
| 30-50 | 0.23 | 0.26 | 0.23 |
| 50-80 | 0.28 | 0.33 | 0.22 |
| 80-120 | 0.33 | 0.35 | 0.25 |
| bcefilter, electrons $E_T>$30 | | | |
| 20-30 | 0.18 | 0.14 | 0.18 |
| 30-50 | 0.11 | 0.12 | 0.077 |
| 50-80 | 0.15 | 0.19 | 0.10 |
| 80-120 | 0.24 | 0.25 | 0.17 |

Table 4.2: Filter effiencies on electrons passing Tight electron ID. Columns show efficiencies for electrons passing different levels of isolation selection, as described in the text: 1) loose cuts in a cone of 0.3, 2) loose cuts in a cone of 0.2, 3) tight cuts in a cone of 0.3.

# Chapter 5

# Data-Driven Estimation of Vector Boson Backgrounds to New Physics Searches Using Photon Events

## 5.1 Introduction

If supersymmetry (SUSY) exists, CMS has the potential to discover it in early LHC running with an inclusive MET+jets search strategy. Such a strategy, as for example the one outlined in the CMS Physics Technical Design Report [18] and CMS Analysis Note AN-2006/089 [19], looks for events with high MET and

several high-energy jets, along with topological cuts to reject background from QCD events and a lepton veto to reject W, Z, and $t\bar{t}$ events. An irreducible background for this search (and other MET-based searches) is Z boson production in conjunction with energetic jets, where the Z decays to neutrinos giving real missing energy. In Ref. [18] a data-driven method of estimating this background using $Z \to \mu\mu$ production as a "standard candle" is presented. The $Z \to \mu\mu$ events are identified and the calorimetric deposits of the leptons are removed from the event, making the identified events a good proxy for $Z \to \nu\bar{\nu}$. This method however suffers from the small number of $Z \to \mu\mu$ events expected during early LHC running, because of the selective search requirements and the lower branching ratio of Z to charged leptons with respect to Z to neutrinos. One way to deal with this problem of small statistics is to relax some of the search cuts, for example reducing the number of jets required, and extrapolate their effect using Monte Carlo predictions. Using Monte Carlo to extrapolate the effect of cuts on jet variables is disadvantageous because of the difficulty of modeling hadronic jets. This is especially true considering the compilicated nature of the jet cuts in the SUSY MET+jets search, which have different $p_T$ thresholds for the leading, sub-leading, and third jets. Also problematic is the fact that jet activity is highly correlated with MET in $Z \to \nu\bar{\nu}$ events, due to the recoil of the Z boson against the jets.

Another strategy, the one presented here, is to use photon events as a proxy for

$Z \rightarrow \nu\bar{\nu}$ events. At sufficiently high $p_T$, the ratio of the cross sections of photon and Z production is expected to be of order 1, and the events are expected to have similar $\eta$ distributions, similar jet number spectra, and similar distributions of jet $E_T$ and $\eta$. Furthermore much of the residual difference is expected to be theoretically understood and thus correctable. Figure 5.1 shows the ratio of the cross sections of Z and photon production as a function of boson $p_T$, for production in association with 1, 2, and 3 extra hard partons, calculated using MadGraph [15]. Figure 5.2 shows the $\eta$ distributions of Z plus 1 parton and photon plus 1 parton production, also calculated with MadGraph.

In identified photon events the photon $p_T$, which is well-measured by the ECAL, is added to the event calorimetric MET to form a "MET-like quantity" that mimics calorimetric MET in $Z \rightarrow \nu\bar{\nu}$ events. On top of the spectrum of this MET-like quantity correction factors are applied to take into account photon acceptance and identification efficiencies, as well as the Z branching ratio to neutrinos and any other theoretical corrections, giving a prediction for the spectrum of MET in $Z \rightarrow \nu\bar{\nu}$ events. This method was introduced in a CMS public note, Ref. [20], and is expanded upon here.

## 5.2 Hadronic Activity in $\gamma$/W/Z Events

Comparison of photon, Z, and W Monte Carlo simulated data indicates that the hadronic properties of these types of events (number, $E_T$, and $\eta$ distributions

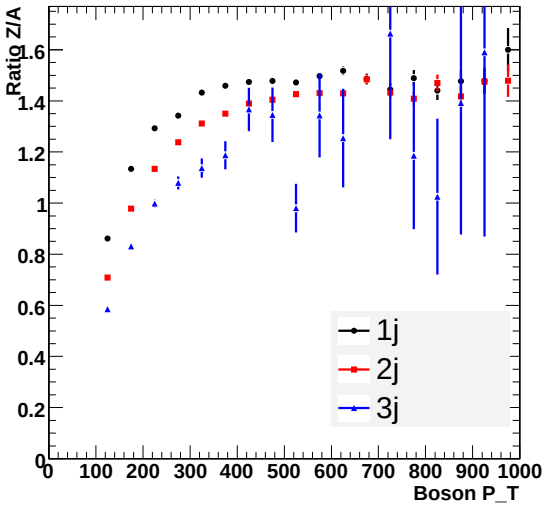


Figure 5.1: Ratio of cross sections of Z and photon production as a function of $p_T$, for production in association with 1, 2 and 3 jets, from Madgraph [15].

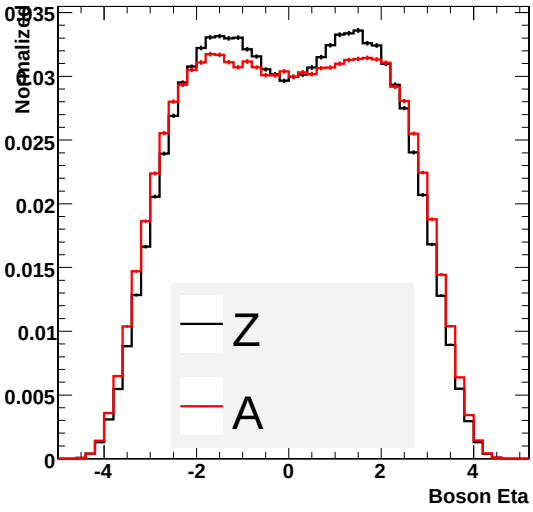


Figure 5.2: Distribution of boson $\eta$ for Z+2j and photon+2j production, from Madgraph. In both cases jets are produced with $p_T > 20$ GeV/c, and with $|\eta| < 5$.

of hadronic jets) are similar. The distributions of number of jets and lead jet $E_T$ were compared for W, Z and photon events generated with ALPGEN [21]. The Monte Carlo samples were produced as part of the CMS Computing Software and Analysis in 2007 (CSA07) excercise. The samples used were generated with a requirement of boson $p_T$ of at least 300 GeV/c. A lower boson $p_T$ cut, for example 200 GeV/c, results in a slighly worse level of agreement. Figure 5.3 shows the distribution of number of jets in W, Z and photon events. Jets are required to have uncorrected $E_T$ of at least 30 GeV. Figure 5.4 shows the ratio of the number events with N jets to the number with N-1 jets, a quantity that is of interest because it is often used as an observable to extrapolate from a

background enriched control region to a signal region. Figure 5.5 shows the ratio of W to photon and Z to photon of the N jets ratio spectrum. Figure 5.6 shows the lead jet $E_T$ distributions. Figure 5.7 shows the ratio of W to photon and Z to photon of the lead jet $E_T$ distributions. This Monte Carlo study supports our assumptions stated in the introduction above, namely that photon events can also be used to make predictions about the hadronic parts of W and Z events.

## 5.3 SUSY MET+Jets Search

The actual search strategy used in CMS early running will almost certainly be somewhat different from what is described in the PTDR. But for the purpose of testing this background subtracting method we adopt that event selection as a model. The selection is as follows:

- MET $> 200$ GeV. Raw calorimeter MET [22] is used.
- Jets are reconstructed with an iterative cone algorithm with a cone radius $\Delta R = 0.5$. At least 3 jets are required with uncorrected $E_T > 30$ GeV and $|\eta| < 3$. Throughout the remainder of this chapter, only jets passing these cuts are used.
- $|\eta| < 1.7$ on the leading jet.
- QCD angular cuts

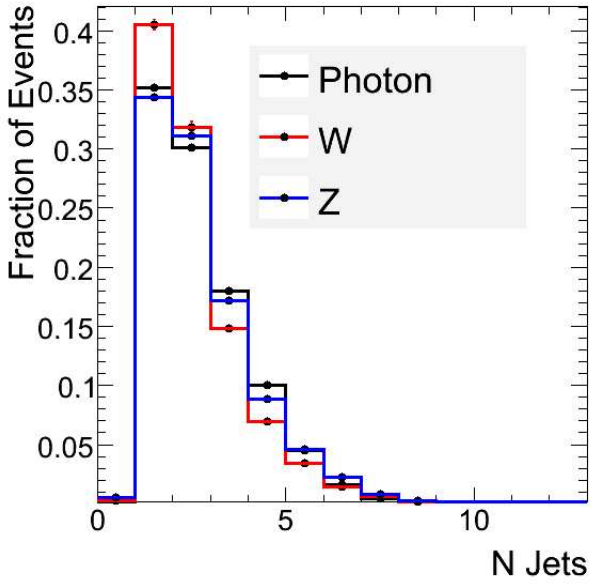


Figure 5.3: Number of jets in W, Z and photon events where the boson has $p_T$ of at least 300 GeV/c. Agreement between the boson types is similar but somewhat worse for a boson $p_T$ cut of 200 GeV/c.

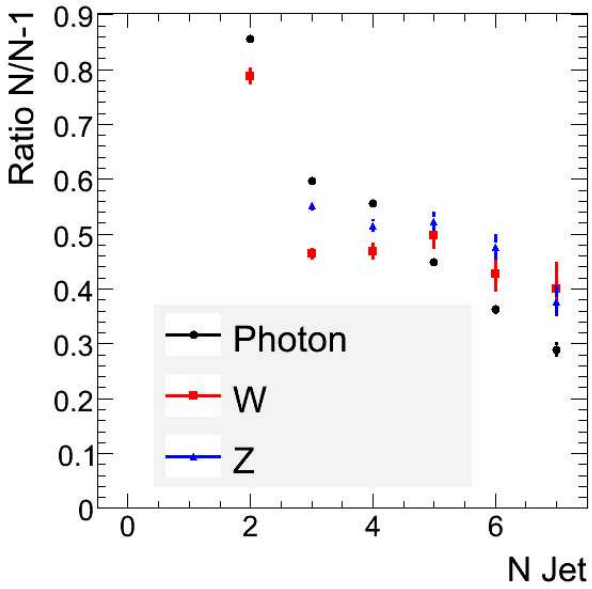


Figure 5.4: Ratio of numbers of events where the boson has N jets to N-1 jets in W, Z and photon events where the boson has $p_T$ of at least 300 GeV/c. A jet is required to have uncorrected $E_T$ greater than 30 GeV, and to lie within $|\eta| < 3$.

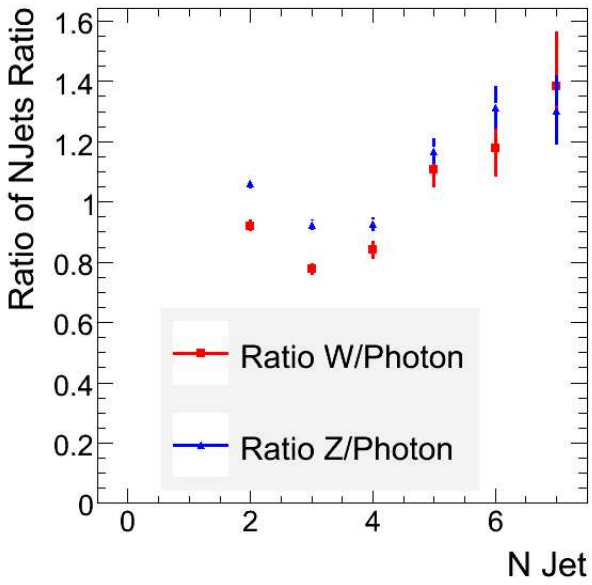


Figure 5.5: The ratio of W to photon and Z to photon N/N-1 jets ratios.

  - $\Delta\phi$ between MET and any jet is required to be $> 0.3$
  - $\Delta\phi$ between MET and the second-leading jet is required to be $> 0.35$
  - $R_1$ and $R_2$ are required to be greater than 0.5, where $R_1$ and $R_2$ are defined as follows: $R_1 \equiv \sqrt{\Delta\phi_1^2 + (\pi - \Delta\phi_2)^2}$ and $R_2 \equiv \sqrt{\Delta\phi_2^2 + (\pi - \Delta\phi_1)^2}$

  These QCD angular cuts are designed to discriminate against QCD events with MET arising from jet mismeasurement.

- Leading jet $E_T > 180$ GeV, second leading jet $E_T > 110$ GeV.
- HT $> 500$ GeV, where HT is defined as the sum of MET and the $E_T$ of the 2nd, 3rd and 4th leading jets. Note that this is a non-standard definition of HT.

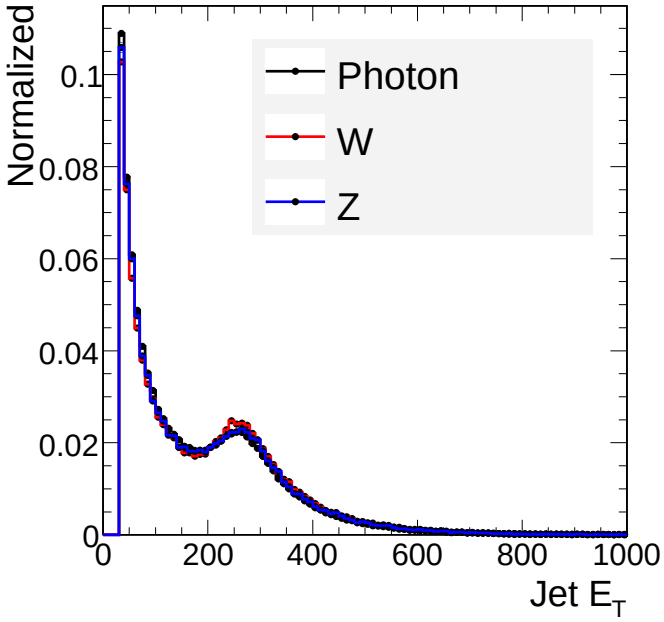


Figure 5.6: Jet $E_T$ for events with boson $p_T$ at least 300 GeV/c. Uncorrected jet $E_T$ is shown.

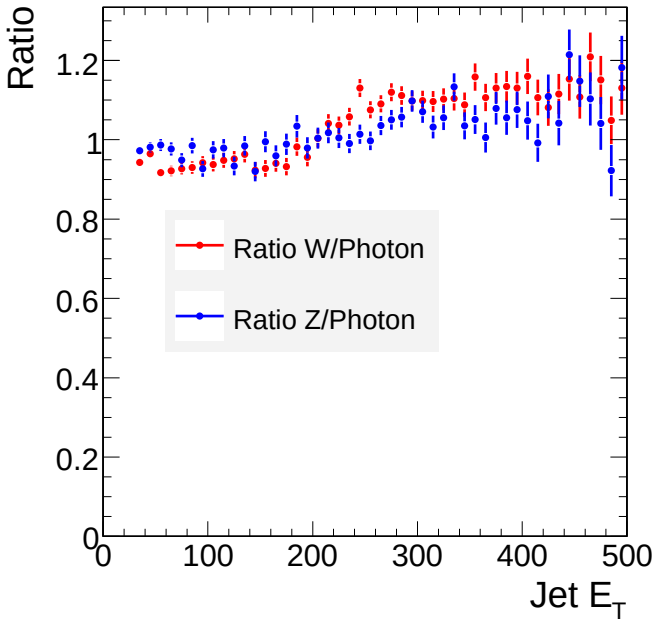


Figure 5.7: Ratio of W to photon and Z to photon jet $E_T$ distributions.

| cut | cumulative selection efficiency |
|---|---|
| MET > 200 GeV | 0.52 |
| >= 3 jets | 0.37 |
| lead jet eta | 0.33 |
| QCD angular | 0.25 |
| lead and 2nd jet et | 0.16 |
| HT >= 500 GeV | 0.14 |

Table 5.1: Selection efficiency on LM1 SUSY signal

The LM1 signal selection efficiency through the selection steps listed above is shown in table 5.1 for comparison to the corresponding results in Ref. [19].

## 5.4 Photon Event Selection

The first step in the photon plus jets channel event selection is to select as pure as possible a sample of prompt photons with $p_T$ of at least 100 GeV. That $p_T$ threshold is low enough to allow the characterization of backgrounds outside of the signal region where the $Z \rightarrow \nu\bar{\nu}$ background will be estimated (roughly photon $E_T > 200$GeV), and also to allow comparison of Z and photon production outside of the signal region, in order to verify the Monte Carlo prediction that photon production is a good proxy for Z production at high $E_T$. We then apply the model SUSY MET + jets search selection, with a few modifications as follows:

- A photon with $p_T$ of at least 100 GeV/c is required. The photon is required to pass isolation cuts and electron veto, both described below. Quality cuts on ECAL cluster shape variables are discussed but are not currently used in

the selection. Fiducial acceptance cuts require that the photon ECAL cluster is within $|\eta| < 2.4$, in order to have tracker coverage, and also require it to be outside of the ECAL barrel-endcap interface region of $1.460 < |\eta| < 1.558$, because the ECAL has a greatly reduced depth of radiation lengths in this region.

- There is no explicit requirement on calorimetric MET, as there is in the standard (MET+jets) selection, since the photon itself is used to mimic the neutrinos of a $Z \rightarrow \nu\bar{\nu}$ event. Instead, the vector sum of calorimetric MET and photon $p_T$ is computed. This quantity, subsequently referred to as the "MET-like" quantity, is directly comparable to calorimetric MET in a $Z \rightarrow \nu\bar{\nu}$ event.

- QCD angular cuts, defined according to angles between MET and jets in the MET+jets search, are in this selection defined according to angles between jets and the MET-like vector built from calorimetric MET and photon $p_T$.

Table 5.3 shows the number of signal and background events selected for different stages of the event selection. Table 5.4 shows the same information, but presented as the relative efficiency of each cut with respect to the number of events selected by the previous cut. Note that the event selection is driven by the SUSY search criteria, which is not *per se* the subject of this thesis.

| **cut level** | **description** |
|---|---|
| 1 | 1 iso. phot eT > 100 GeV |
| 2 | njets >= 3 |
| 3 | lead jet cuts |
| 4 | 2nd jet cuts |
| 5 | QCD angular cuts |
| 6 | HT>500 GeV |
| 7 | MET-like > 200 GeV |

Table 5.2: Defintion of cut levels.

| **cut** | **prompt phot.** | **QCD** | $t\bar{t}$**/W/Z** | **S/B** | $S/\sqrt{(S+B)}$ |
|---|---|---|---|---|---|
| 1 | 50773 | 6382 | 656 | 7.2 | 211 |
| 2 | 4787 | 138 | 73 | 23 | 68 |
| 3 | 1077 | 14 | 19 | 33 | 32 |
| 4 | 698 | 5.7 | 11 | 43 | 26 |
| 5 | 513 | 2.9 | 7.6 | 48 | 22 |
| 6 | 198 | 0.76 | 3.6 | 46 | 14 |
| 7 | 112 | 0.25 | 2.9 | 35 | 10 |

Table 5.3: Signal and QCD background events expected for $100pb^{-1}$. Definition of the cut levels is provided in table 5.2

| **cut level** | **prompt phot.** | **QCD** | $t\bar{t}$**/W/Z** |
|---|---|---|---|
| njets >= 3 | 0.095 | 0.022 | 0.12 |
| lead jet cuts | 0.23 | 0.10 | 0.26 |
| 2nd jet cuts | 0.65 | 0.40 | 0.58 |
| QCD angular cuts | 0.74 | 0.52 | 0.73 |
| HT>500 GeV | 0.39 | 0.26 | 0.48 |
| MET-like > 200 GeV | 0.56 | 0.33 | 0.82 |

Table 5.4: Relative efficiency of each step of the cuts.

### 5.4.1 Photon isolation

Photons are required to pass tracker and calorimeter isolation cuts in order to reject secondary-photon background.

- The tracker isolation variable is computed as the sum $p_T$ of tracks in a hollow cone (inner radius 0.02, outer radius 0.3 in $\eta - \phi$ space) about the vector from the reconstructed primary vertex to the ECAL supercluster centroid, and is required to be less than 1 GeV/c. Tracks from conversion electrons mostly do not contribute to the $p_T$ sum, both because they are generally not reconstructed by the standard track reconstruction sequence, and because of the inner veto cone.

- A relative calorimeter isolation requirement is used in order to achieve a more uniform photon efficiency as a function of the photon transverse energy. The sum of transverse energy in the ECAL and HCAL, computed in a cone of radius 0.3 around the ECAL supercluster centroid, minus the $E_T$ of the photon supercluster, is required to be less than $1\,\mathrm{GeV} + 0.02 \cdot E_T(\gamma)$.

Figures 5.8 and 5.9 show the MC estimation of the efficiency of photon isolation requirements, respectively in the inclusive sample and in events with at least 3 hadronic jets of uncorrected $E_T$ greater than 30 GeV. In Figures 5.10 and 5.11 the efficiency is shown as a function of the photon's pseudorapidity.[1] Figures 5.12

[1]Efficiencies are calculated from Monte Carlo events by looking for reconstructed photons matched to generated prompt photons. A $dR \equiv \sqrt{\Delta\phi^2\Delta\eta^2} < 0.1$ match and $E_T$ match of 10% of the generated photon's $E_T$ is required for a match to be made.

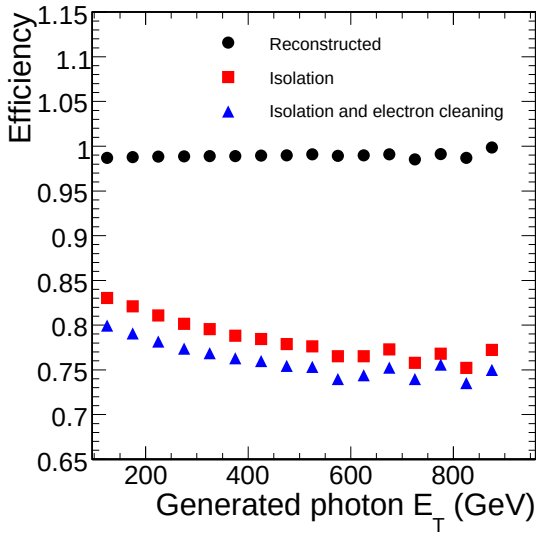


Figure 5.8: Photon selection efficiency as a function of photon $E_T$, evaluated for the inclusive photon sample.

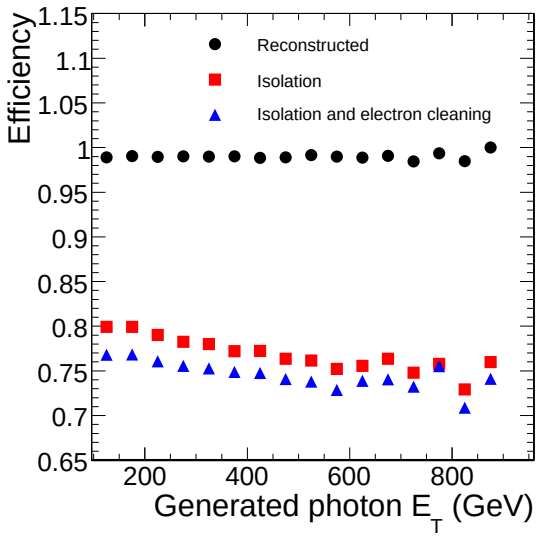


Figure 5.9: Photon selection efficiency in events with at least 3 jets.

and 5.13 show the angular and $E_T$ resolutions of photon reconstruction.

Relative calorimeter isolation cuts are necessary because calorimeter isolation energy increases with increasing photon $E_T$. Both ECAL and HCAL isolation sums increase; the dominant effect is in the HCAL isolation. Potential mechanisms include the following. First, leakage of photon energy from the ECAL to the HCAL, and second, energy deposited in the ECAL by the photon, but not clustered in the ECAL supercluster (a likely effect of photon conversions), could scale linearly with photon $E_T$. Simple arguments suggest that tracker isolation is not expected to scale with increasing photon $E_T$: the possible mechanism would be contributions to the track isolation sum from photon conversions. However, photon conversion tracks rarely contribute to the track isolation sum, and the

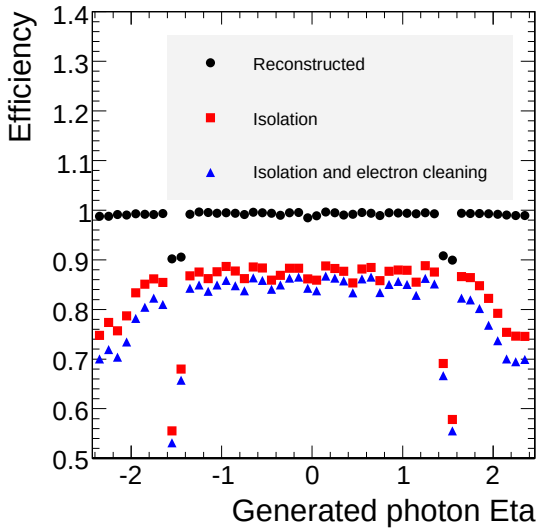


Figure 5.10: Absolute efficiencies as a function of $\eta$ for photon reconstruction and the subsequent isolation and electron-cleaning cuts.

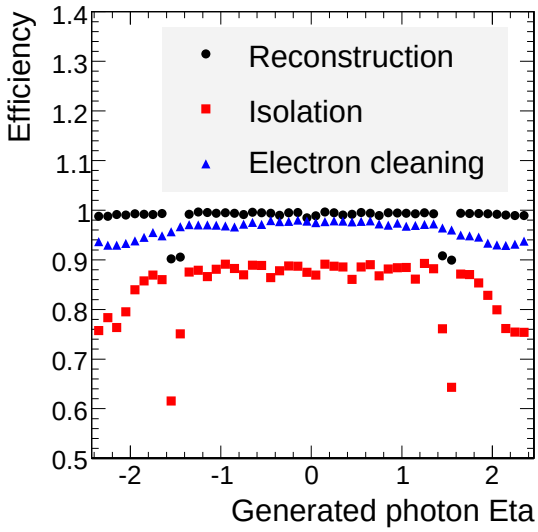


Figure 5.11: Efficiencies of electron reconstruction, isolation cuts, and electron cleaning cuts, evaluated relative to the previous cut step. Reconstruction efficiency is evaluated relative to all generated electrons, isolation efficiency is evaluated relative to all reconstructed electrons, and electron-cleaning efficiency is evaluated relative to all isolated electrons.

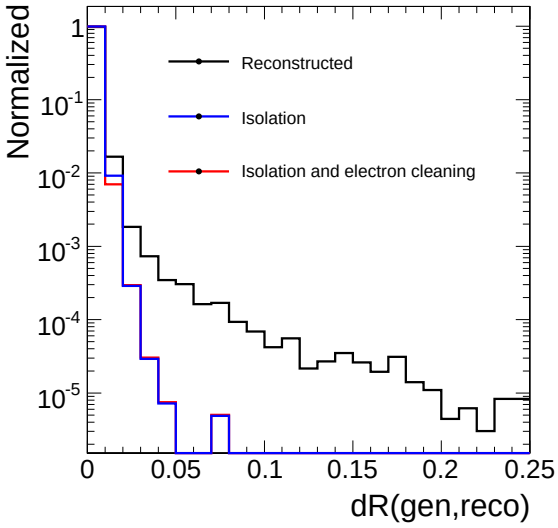


Figure 5.12: Angular resolution of photons, after reconstruction only, after isolation requirements, and after isolation and electron-cleaning requirements.

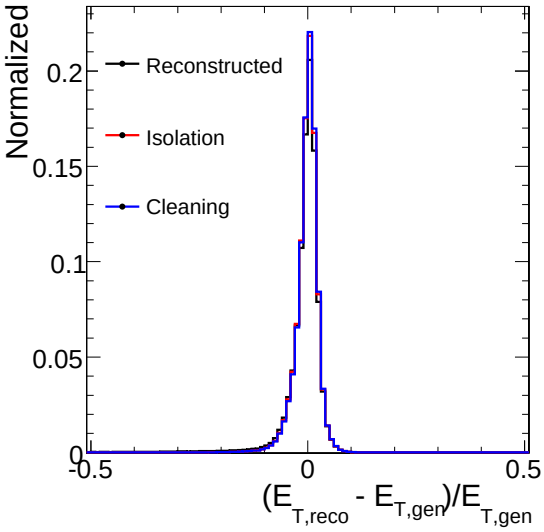


Figure 5.13: $E_T$ resolution of reconstructed photons. A guassian fit to the central region gives a standard deviation of approximately 2%.

photon conversion cross section rises only very weakly with photon energy above 1 GeV. A secondary effect is that events with high photon $E_T$ will have high counter-balancing hadronic energy; an increase in hadronic energy results in increased chance of overlaps of hadronic energy deposits with the photon isolation cone. This secondary effect is not expected to be large. That the effect is not large can be confirmed by looking at the tracker isolation efficiency; the effect of increased hadronic activity in the event should be the same for tracker isolation as for the calorimeter isolations. In fact an *increase* in track isolation efficiency is observed. Figures 5.14 through 5.19 show the isolation efficiency of absolute cuts on track, ECAL, and HCAL isolation sums.

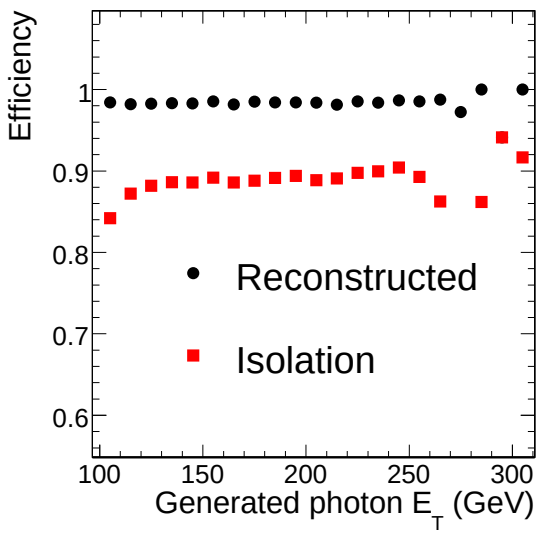


Figure 5.14: Efficiency of the tracker isolation cut of 1 GeV/c as a function of photon $E_T$.

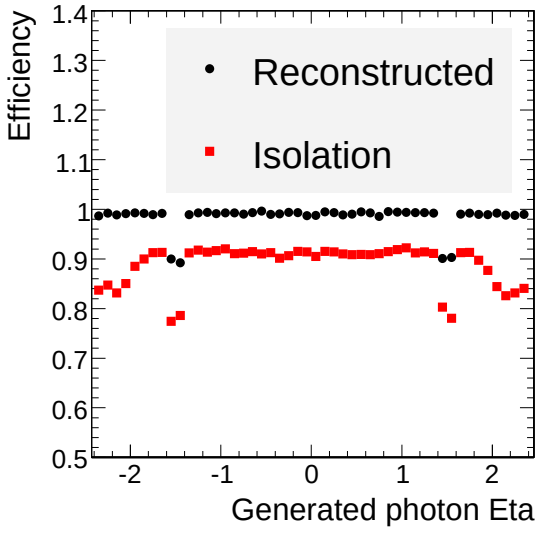


Figure 5.15: Efficiency of the tracker isolation cut of 1 GeV/c as a function of photon $\eta$.

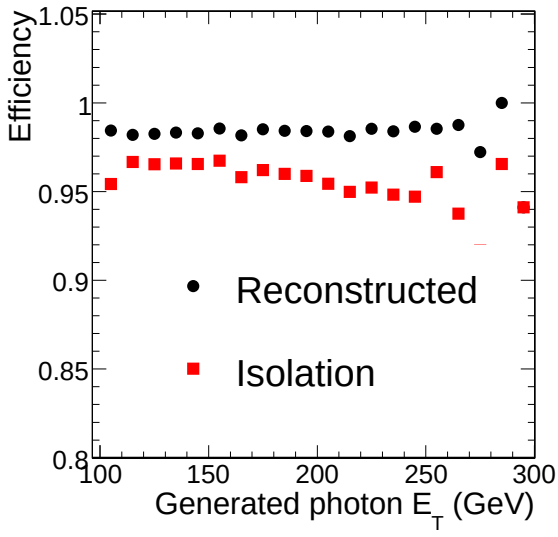


Figure 5.16: Efficiency of an ECAL isolation cut of 3 GeV as a function of photon $E_T$.

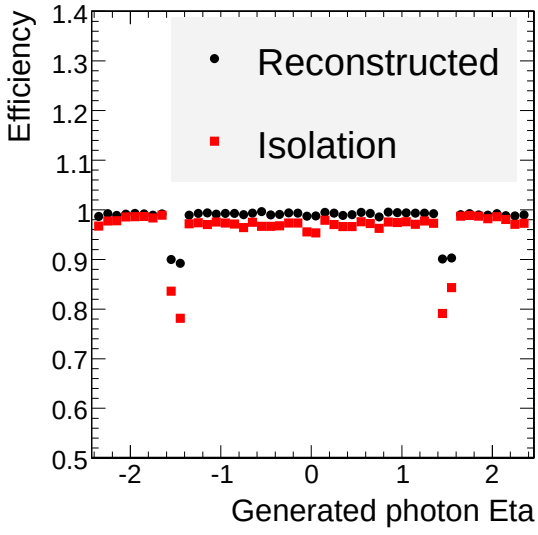


Figure 5.17: Efficiency of an ECAL isolation cut of 3 GeV as a function of photon $\eta$.

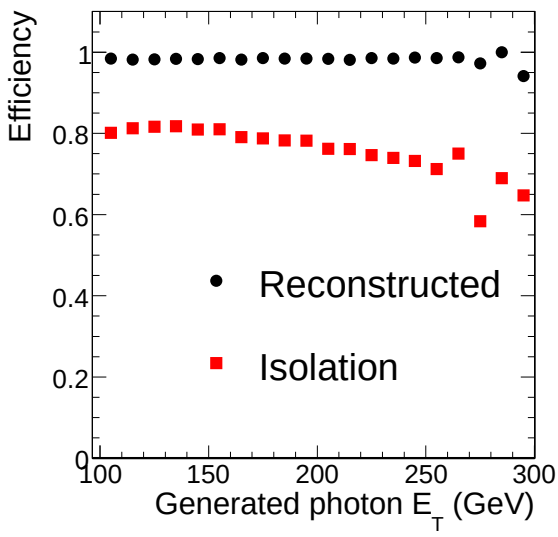


Figure 5.18: Efficiency of an HCAL isolation cut of 3 GeV as a function of photon $E_T$.

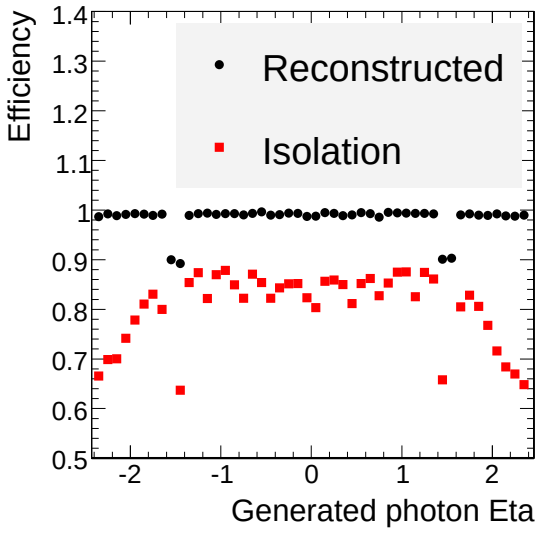


Figure 5.19: Efficiency of an HCAL isolation cut of 3 GeV as a function of photon $\eta$.

### ECAL Shower Leakage

Leakage of photons' energy from the ECAL to the HCAL occurs because, despite the large radiation-length depth of the CMS ECAL, the electromagnetic cascade of incident photons with $E_T$ greater than 100 GeV is not fully contained in the ECAL (the cascade depth depends logarithmically on incident particle energy). The CMS ECAL is nominally about 26 radiation lengths thick, excluding so-called "cracks", regions of decreased thickness (most evident at the barrel-endcap interface). The radiation length, $X_0$, of a material is the mean distance over which a high-energy electron loses all but 1/e of its original energy through bremsstrahlung radiation, and is also equal to 7/9 times the mean free path for

pair production by a high-energy photon [2]. Thus an incident electron or photon has with near-certainty interacted within a few interaction lengths of entering the ECAL. Since the electrons (or photons) produced in this interaction also interact on a similar length scale, and since the number of particles doubles with every interaction, an electromagnetic shower is initiated. The shower continues to develop (producing a doubling of energetic electrons and photons roughly every radiation length, but each with reduced energy), until shower particles become sufficiently reduced in energy that their interactions no longer produce a doubling of the number of high-energy particles. In the case of electrons, this energy is customarily taken as the "critical energy", the energy at which the ionization energy loss per $X_0$ is equal to the electron energy itself; this is also approximately the energy at which the bremsstralung energy loss is equal to the ionization energy loss; the shower stops when the shower particles start to lose most of their energy to ionization rather than bremsstrahlung and pair production.

**EM Shower Depth**

A simple (and oft-quoted) analysis shows that the depth of an electromagnetic (EM) shower depends logarithmically on the energy of the incident particle that triggers the shower. Imagine the ECAL to be divided into thin layers such that the probability $p$ for an interaction in a given layer is small enough that the

[2]The near-equality of the length scales for high-energy electron and photon interaction is why the response of a photon in the ECAL is indistinguishable from that of an electron.

probability for multiple interactions is negligible. Assume that $p$ is the same for electron bremsstrahlung as for photon pair production of electrons, justified by the near-equality of the interaction lengths for these processes. The number of particles produced after layer number L is related, on average, to the number of particles produced after layer (L-1) by $N_L = p \cdot 2N_L + (1-p)N_{L-1}$. By induction, $N_L = (p+1)^L$, if one incident particle started the shower. If each interaction results in an equal energy sharing between daughter particles, after L layers, each particle has on average $E = E_0/(p+1)^L$, where $E_0$ is the incident energy. Taking the stopping condition for an EM shower to be when the average energy is equal to the critical energy $E_C$, at which an electron loses as much energy to ionization as to bremsstrahlung, the shower depth L is proportional to the logarithm of the incident energy divided by the critical energy, $L \propto \ln(E_0/E_C)$.

### 5.4.2 GEANT Simulations

Toy detector simulations were performed using the GEANT4[16] simulation package. The CMS ECAL was simulated as a monolithic block of lead tungstate ($PbWO_4$) 26 radiation lengths thick. Incident photons of various energies were simulated, and energy leakage beyond the end of the calorimeter was recorded. The simulations show that the observed inefficiency of an absolute HCAL isolation cut is consistent with the expected shower leakage of energetic photons incident on the CMS ECAL. Figure 5.20 shows the efficiency of a cut of 3 GeV on energy

leakage beyond 26 radiation lengths in the toy simulation, as a function of incident photon energy. In the actual CMS detector, energetic particles escaping the ECAL crystals must also traverse the ECAL support structure, comprising between zero and two radiation lengths depending on $\eta - \phi$ location, before depositing their energy in the first scintillating layer of the HCAL. Nevertheless figure 5.20 clearly motivates why calorimeter isolation cuts must be defined relative to photon energy for high $E_T$ photons. Figure 5.21 shows the mean energy leakage beyond 26 radiation lengths in the toy simulation. This puts a bound on the significance of energy leakage on photon energy resolution. Up to about 500 GeV the energy leakage is less than 2%, an amount which should be corrected for (and is) but which is not a large effect for the analyses presented in this thesis. Note that ECAL readout for single crystals doesn't saturate until about 1.7 TeV in the barrel and 3.0 TeV in the endcaps[3]; there is a wide energy range where shower leakage may be significant but where ECAL crystals aren't yet saturated.

A data-driven methodology of estimating photon isolation efficiency, using a tag-and-probe method with Z $\rightarrow$ ee events, is discussed in section 5.8.

### 5.4.3 Trigger Efficiency

The trigger efficiency for isolated photons above 100 GeV is shown in Figure 5.22, using the high-$E_T$ EM cluster trigger stream (called HLT1EMHighEt in

[3]At these energies saturation occurs in the analog-to-digital conversion stage of electronic readout.

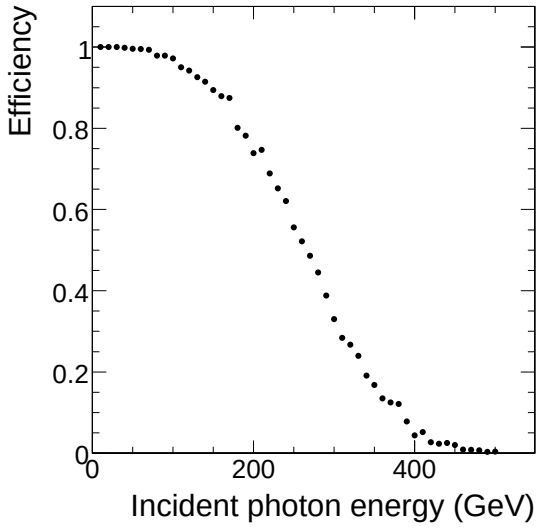


Figure 5.20: Efficiency of a cut of 3 GeV on energy leakage beyond 26 radiation lengths in a toy ECAL simulation described in 5.4.2.

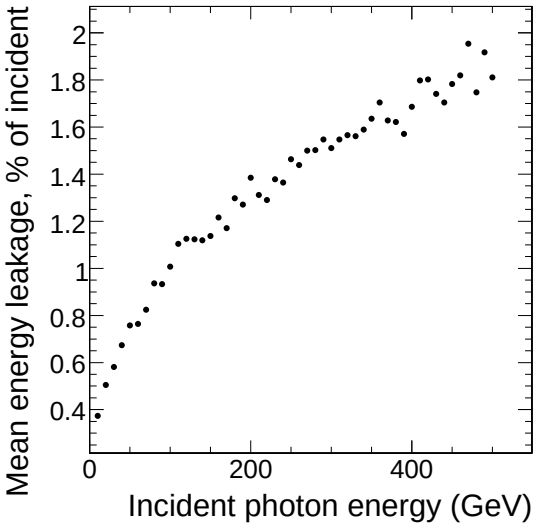


Figure 5.21: Mean energy leakage beyond 26 radiation lengths in the toy ECAL simulation described in 5.4.2.

the CMS trigger naming scheme). The HLT1EMHighEt trigger, which requires an EM cluster with at least 80 GeV $E_T$ and passing loose isolation cuts, is found to have very good efficiency for photon $E_T$ near 100 GeV, but degrades at higher $E_T$[4]. The efficiency of this trigger, integrated over the entire photon $E_T$ spectrum above 100 GeV, is close to 95% relative to events passing analysis cuts at all steps in our selection. Figure 5.23 shows the trigger efficiency as a function of photon $E_T$ for events passing the full event selection up to the HT cut. The photon trigger is not actually applied in the remainder of the analysis.

[4]The degradation of the HLT1EMHighEt trigger is a known problem in CMSSW_1_6_7, which has been fixed in more recent software releases.

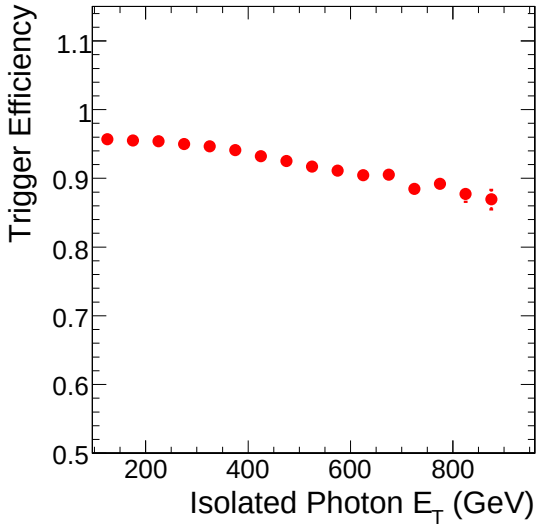


Figure 5.22: Trigger efficiency on events having an isolated reconstructed photon above 100 GeV.

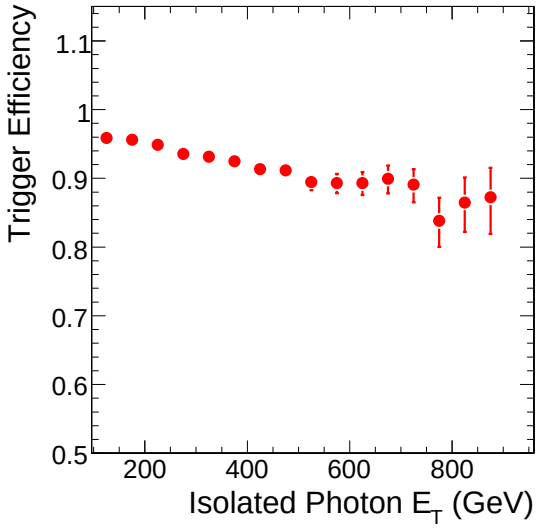


Figure 5.23: Trigger efficiency on events passing the full event selection up to the HT cut.

## 5.5 Backgrounds to photon+jets

The main backgrounds to prompt photons are photons from meson decays produced in QCD jets, and electrons for which no electron track is reconstructed. We estimated the background to our full photon selection using $t\bar{t}$, W, Z, and QCD Monte Carlo simulated data, generated with ALPGEN and fully simulated with the GEANT4-based CMS detector simulation. The QCD Monte Carlo data samples used are listed in table A.1.

### 5.5.1 Background from Electrons Faking Photons

The reconstructed energy deposits of electrons incident on the ECAL are indistinguishable from those of incident photons. Thus in order to reject electrons

from the photon sample, photons are rejected if their superclusters are also reconstructed as electrons. For that to happen, the electron reconstruction code must reconstruct a track that matches the supercluster in $\eta$ and $\phi$, with a loose cut on transverse energy matching. The efficiency of electron cleaning is shown in Figures 5.8, 5.10 and 5.11. The per-electron photon fake rate, after electron veto, is shown in figures 5.39 and 5.40. Because electrons can fake photons, $t\bar{t}$, W, and Z production are possible backgrounds to photons. Given the fake rates, the cross-sections for $t\bar{t}$, W and Z production, and the branching ratio of W and Z to electrons, the expected contribution of these processes to the photon selection is less than 5%. A data-driven methodology to estimate the electron to photon fake rate is presented in 5.9.

### 5.5.2 Sources of Photon Background in QCD Events

Neutral pion and eta mesons, decaying to pairs of photons, are produced copiously in hadronic jets. For $\pi_0$ and $\eta$ with transverse energy above 100 GeV, the energy range relevant for this analysis, the photon pairs are sufficiently collimated to be mostly indistinguishable from prompt photons in their calorimetric response. However, isolation cuts are very effective in rejecting these background photons. Figure 5.24 shows the $p_T$ spectrum of background photons before and after isolation and electron cleaning. Isolation cuts have a rejection power of nearly $10^3$ on QCD photons with $E_T$ above 100 GeV.

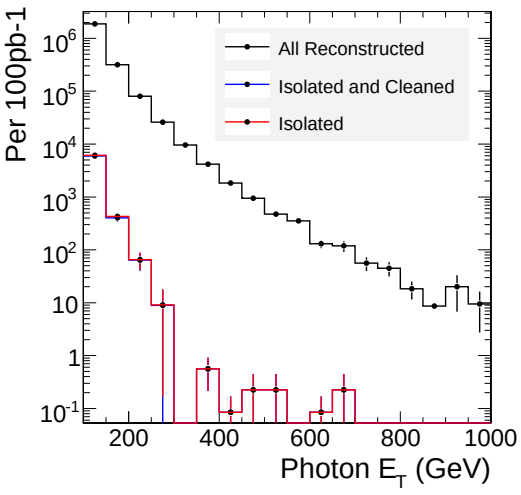


Figure 5.24: Background Photons before and After isolation cuts and electron cleaning. Error bars correspond to Monte Carlo statistics.

### Simulation Study

A simulation-level study was done to determine if any sources other than neutral meson decay are a significant source of photon background from QCD events. All simulated particles depositing energy in ECAL crystals that were clustered into the supercluster belonging to reconstructed isolated photons were identified. The ancestery of these particles was traced to the original parent particle. The original parent is the stable particle produced by the event generator and passed to the detector simulation, which through subsequent interaction, including possible decay to other particles, deposits all or part of its energy in the detector. For example, if an electron radiated a bremstralung photon that subsequently converted to an electron-positron pair, and the conversion electron contributed energy to the ECAL cluster of a reconstructed electron, the first

electron would be the original parent. Figure 5.25 shows the distribution of the fraction of reconstructed photons' energy that comes from photons from $\pi_0$ or $\eta$ meson decay. About 98% of reconstructed photons in QCD events get at least 90% of their energy from such photons. Additional sources of energy deposit in photons superclusters include electrons, $K_L$, $K_S$, charged pions, charged kaons, protons, and neutrons. Also of interest is the fact that reconstructed photons are frequently made up of energy deposits from more than two photons, i.e. the result of more than a single neutral meson decay. Figure 5.26 shows the number of photons needed to account for 90% of the reconstructed photon energy, for all reconstructed photons which get at least 90% of their energy from photons from meson decay. The effect of meson decay products being given asymmetric $E_T$ in the laboratory frame is visible in the first bin.

### 5.5.3 Discrimination by Cluster-Shape Variables

Cluster-shape variables provide some discrimination between prompt and background photons, but for reconstructed photons of $E_T$ greater than 100 GeV the discriminating power is not large enough to justify their use. Cluster shape is correlated with photon-conversion observables used elsewhere (see section 5.6), and furthermore correcting for the efficiency of a cut would induce dependence on Monte Carlo simulation. Photon pairs from the decay of $\pi_0$ and $\eta$ mesons above 100 GeV are highly collimated. Figure 5.27 is a scatter plot showing the angle

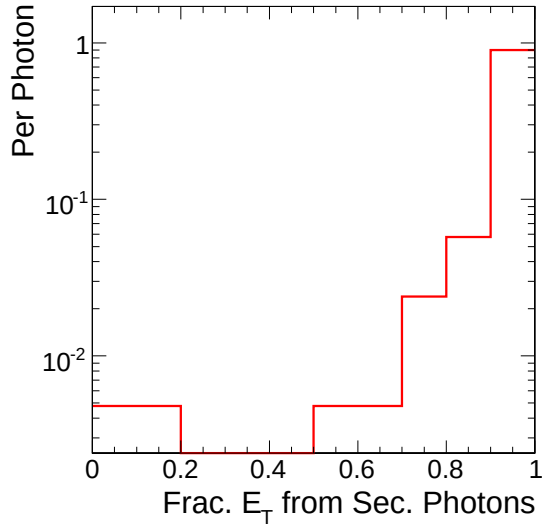


Figure 5.25: Fraction of reconstructed photon's energy coming from photons from meson decay.

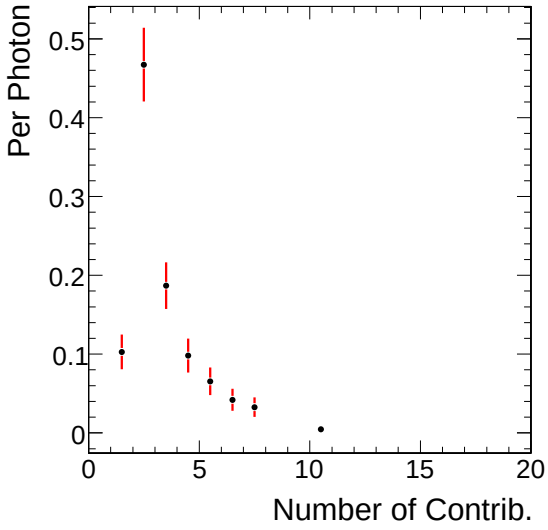


Figure 5.26: Number of photons needed to make up 90% of the energy of reconstructed, isolated photons in QCD events.

between decay photons from $\pi_0$ and $\eta$ mesons with $p_T$ of at least 100 GeV/c versus the relative energy difference of the two photons $|E_T^1 - E_T^2|/(E_T^1 + E_T^2)$. About 99% of the time, the photon decay products of $\pi_0$ and $\eta$ mesons are separated by an angle smaller than the size in $\eta$ and $\phi$ of an ECAL crystal (0.0174 radians). Recall that the 2.2 cm Moliere radius of the ECAL crystals, or radius which contains 90% of the energy of an electromagnetic shower, is approximately equal to the size of the crystals in the direction transverse to the shower axis in the typical case; shower shape variables clearly cannot be very effective in distriminating photon pairs from high-$E_T$ meson decays incident on the ECAL surface.

Figure 5.28 shows the R9 measure in prompt and background isolated photons. The quantity R9, which is the ratio between the energy in the central

3x3 array of crystals of a supercluster relative to the energy of the entire supercluster, is a typical cluster-shape discriminator. The separation of prompt and background photons in R9 is largely a result of the increased probability of a high-$E_T$ reconstructed background photon to have an associated conversion, relative to a reconstructed prompt photon. This increased probability is because reconstructed background photons are usually made up of multiple photons, each one having the same probability to convert as a single prompt photon. The figure also shows the distribution for prompt photons separated into the converted and unconverted cases according to Monte Carlo simulation truth information.

Another cluster shape variable that is often used as a discriminant is $\sigma_{\eta\eta}$, a measure of the cluster width in the $\eta$ direction. Trajectories of conversion electrons do not bend in the $\eta$ direction, thus this quantity is much less sensitive to photon conversion but still could be useful for background discrimination. From figure 5.29 it is clear that $\sigma_{\eta\eta}$ has little discriminating power for high-$p_T$ photons. Because we use conversions to estimate the purity of our selected photon sample, and also because of the difficulty of understanding the efficiency of cluster shape cuts, they are not applied in this analysis.

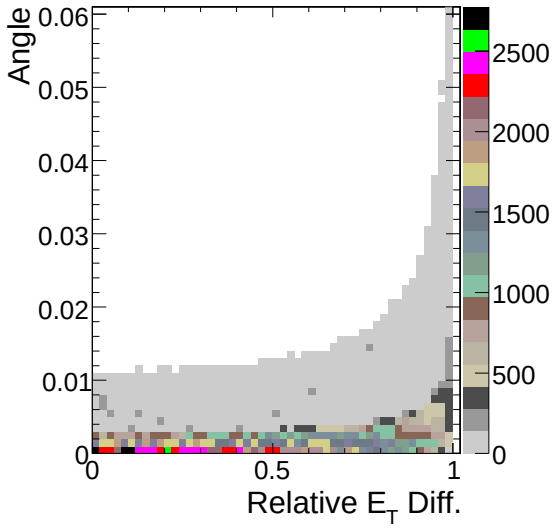


Figure 5.27: Opening angle versus relative energy difference for pairs of secondary photons from $\pi^0$ and $\eta$ mesons with $p_T$ of at least 100 GeV/c.

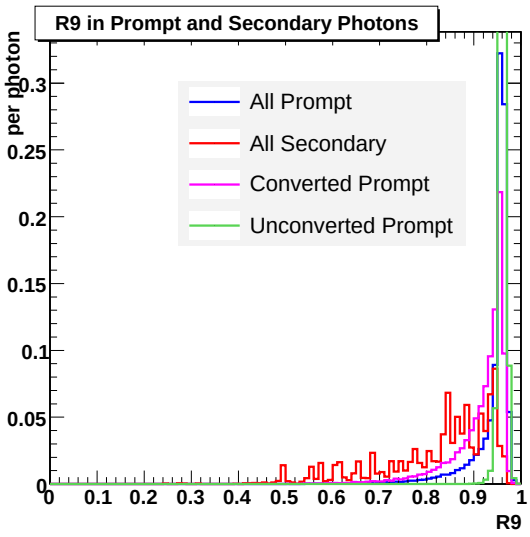


Figure 5.28: R9 in prompt and background photons

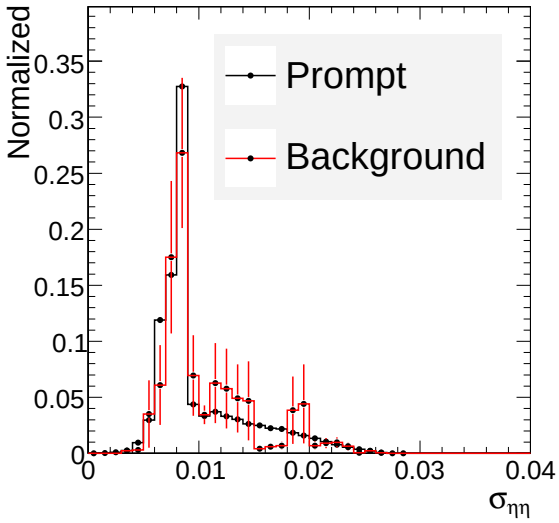


Figure 5.29: $\sigma_{\eta\eta}$, a measure of cluster width in the $\eta$ direction, for prompt and background photons passing photon identification cuts (isolation, electron cleaning).

## 5.6 Data Driven Estimation of the QCD background to Photon+jets

The secondary-photon background to prompt photon production after full selection is expected to be small (see table 5.3). To confirm this expectation in data we plan to exploit the production of conversion electrons in the tracker. Our method is based on the fact that photon conversions associated to prompt and to background reconstructed photons have different properties.

A single prompt photon in the energy range relevant to this analysis has about a 50% probability to convert in the CMS tracker, and if it converts, the sum $p_T$ of the two conversion electrons will equal the photon's reconstructed $E_T$. Background photons are usually made up of two or more highly-collimated secondary photons that are reconstructed as a single photon candidate. Each constituent photon has the same probability to convert as a single prompt photon. Thus for any one background-photon conversion, the $p_T$ of the daughter tracks sums to less, often much less, than the reconstructed photon candidate's $E_T$. With an ideal reconstruction this could be easily exploited to estimate the fraction of secondary photons in a reconstructed photon sample.[5] Such a method is complicated by the fact that, for high $E_T$ photons, conversion tracks are highly collimated. Furthermore conversion electrons may emit bremstralung radiation in

[5]This technique was used by the CDF collaboration in a measurement of the prompt photon production cross section [23].

material interactions, producing tertiary photons, which may themselves convert; a single photon conversion is the beginning of an electromagnetic shower. Thus the accurate reconstruction of high-$E_T$ photon conversions is difficult.

Figure 5.30 shows the number of reconstructed conversions passing quality cuts per isolated reconstructed photon. Figure 5.31 shows a histogram of the sum $p_T$ of the tracks belonging to a conversion, divided by the ECAL supercluster $E_T$ of the photon that seeded the conversion reconstruction. This quantity is subsequently referred to as P/E. In the figure P/E is shown for all events passing the requirement of having an isolated reconstructed photon with $p_T$ of at least 100 GeV/c. Two low-$\hat{p_{\mathrm{T}}}$ QCD samples have been removed from figure for clarity, because they have extremely high weights (weight >100) for a 100 $pb^{-1}$ data sample.[6] These QCD samples do not contribute any events to the final event selection. They do however contribute to the selection requiring only an isolated photon with at least 100 GeV $p_T$, and they have the effect of obscuring the shape of the P/E distribution for secondary photons, due to their high weight. It is assumed that the shape of the P/E distribution for these two $\hat{p_T}$ bins would not be different from the remaining, included, bins, if sufficient statistics were available.

P/E is peaked near 1 for primary photon conversions, and distributed at much lower values for secondary photon conversions, as seen in figure 5.31. Conversion clean-up cuts have been imposed, requiring that two conversion tracks be found,

[6]Because the QCD production cross section is very high, and falls steeply with increasing center-of-mass energy transfer $\hat{p_T}$, QCD Monte Carlo production is binned according to $\hat{p_T}$ and different bins are produced with different equivalent integrated luminosities.

that each track have at least seven hits, and that at least one of the two tracks has a normalized chi-squared less than six. Additionally, only one conversion per reconstructed photon is taken; the one with the highest P/E is kept because this provides the best discrimination between signal and background.

To extract the background photon contamination using conversion information, it's useful to have a control sample of high-$E_T$ background photons in which to study conversions. Semi-isolated photons can be used as such a sample. If isolation requirements are inverted, production of background photons totally dominates production of prompt photons. For the control sample the HCAL isolation sum in a cone of 0.3 is required to be greater than 10 GeV to ensure that the reconstructed photon is part of a hadronic jet and thus a background photon. In order to avoid biasing conversion reconstruction by the presence of extra tracks the secondary photon control sample is required to pass tracker and ECAL isolation cuts in a reduced cone size of 0.2 (the cut values themselves are kept the same). The reduced cone size of 0.2 instead of 0.3 is used in order to maintain the acceptance of background photons. Figure 5.32 shows histograms comparing reconstructed conversions in background photons passing the full selection and in the background-photon control sample defined by the above cuts. Several alternate background control samples, defined by different ranges of HCAL isolation, are displayed in order to demonstrate that the background control sample is not too sensitive to the exact value of the HCAL isolation cut. Table 5.5 shows the

| selection | isolated prompt | isolated secondary | sec. control |
|---|---|---|---|
| N iso. phot | 50773 | 6382 | 13337 |
| N qual. conv. | 7021 | 1380 | 3580 |

Table 5.5: Number of events with photons and with reconstructed conversions passing quality cuts. Shown for isolated prompt photons, isolated secondary photons, and photons in the secondary photon control sample

number of conversions expected in the isolated photon sample, from prompt photons and secondary photons, as well as in the secondary-photon control sample.

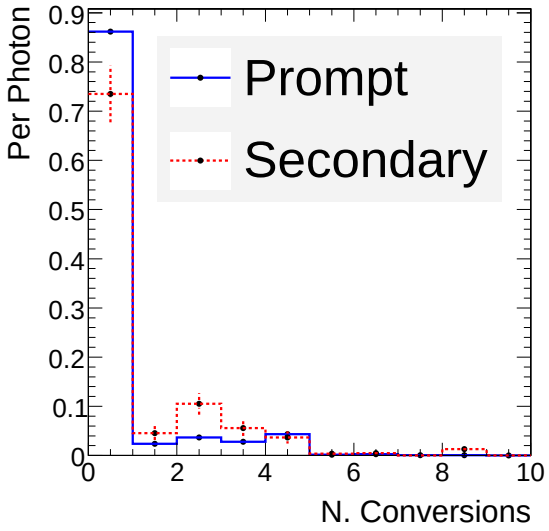


Figure 5.30: Number of reconstructed conversions passing quality cuts, per isolated photon. The figure is extracted from all events having an isolated reconstructed photon with $p_T > 100$ GeV. Error bars correspond to Monte Carlo statistics.

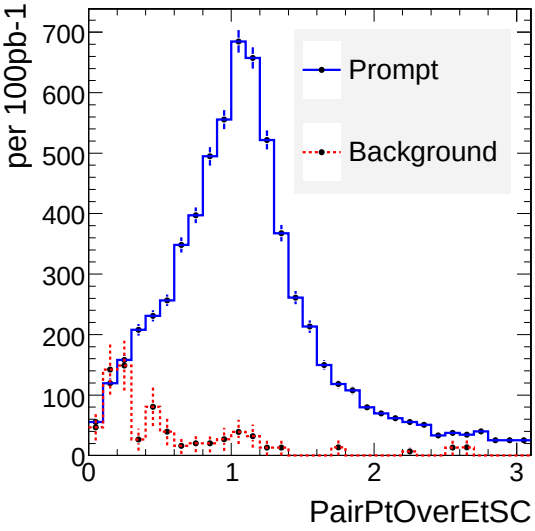


Figure 5.31: P/E, in reconstructed conversions passing quality cuts. The figure is extracted from all events having an isolated reconstructed photon with $p_T > 100$ GeV. Error bars correspond to Monte Carlo statistics.

Ideally templates of P/E for prompt and background photons would be fit the observed P/E distribution in order to extract the fractions of prompt and background photons in the signal sample. But a good technique for extracting a

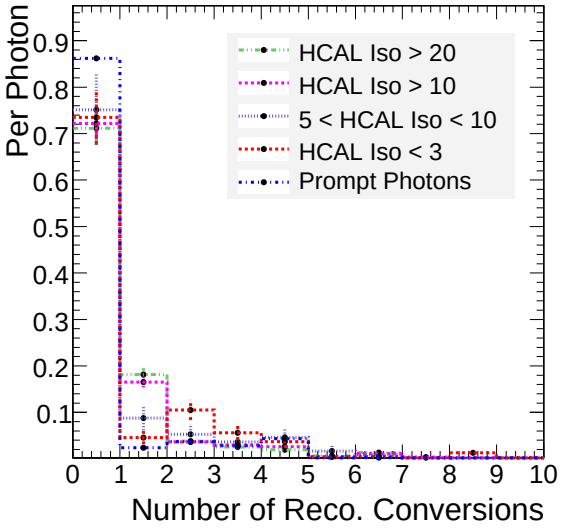


Figure 5.32: Number of reconstructed conversions passing quality cuts, in isolated secondary photons and secondary-photon control sample. Error bars correspond to Monte Carlo statistics.

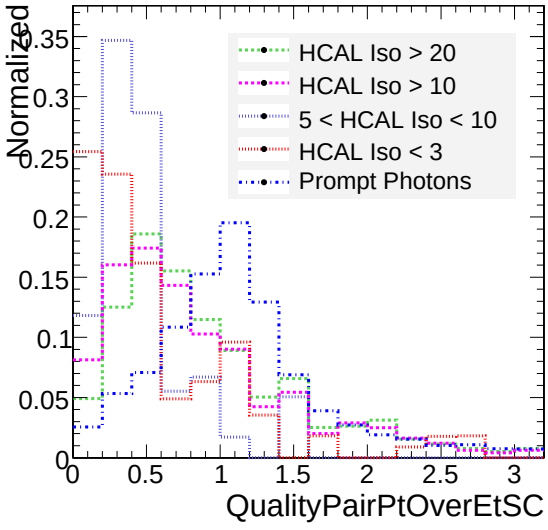


Figure 5.33: P/E, in reconstructed conversions passing quality cuts. Error bars correspond to Monte Carlo statistics.

template P/E distribution for prompt photons from collider data is not at hand. One possibility is to extract the template from Monte Carlo. It remains to be seen whether photon conversions, which are highly dependent on the distribution of material in the tracker, will be well described by the CMS detector simulation without extensive tuning which will not be available for analyses done with early data. However if the background photon P/E template distribution is well-modeled by Monte Carlo simulation, some confidence can be given to the corrected modeling of the prompt photon P/E shape. The P/E distribution is still a very useful cross-check that the sample of isolated photons is dominated by prompt photons, and the variable is promising; with more work a precise background-estimation method should be possible.

### 5.6.1 Number of Hits In a Cone about Photon Direction

Another quantity that might be useful to discriminate between prompt and secondary photons is the number of hits within a cone about the photon direction. Since reconstructed secondary photons have more associated conversions, they have more hits in the silicon tracker. Using hits instead of reconstructed conversions is useful because it avoids systematics associated with uncertainty on conversion reconstruction efficiency, fake rate, and $p_T$ resolution. However, it is not as powerful as using reconstructed conversions because hits not reconstructed into tracks do not provide a direct measure of photon $p_T$.

Since we are concerned with photons passing isolation cuts, the cone about the photon direction should be mostly free of hits from tracks not originating with photon conversions. Residual contamination from low-$p_T$ tracks bending into the isolation cone could be removed provided these tracks are efficiently reconstructed.

Figure 5.34 shows the distributions of number of hits in a cone of 0.2 about the direction of isolated reconstructed photons. Pixel hits are not included. Only *matched* hits from the silicon strip tracker are included. Matched hits occur when a hit in an $r\phi$-oriented strip detector is matched to a hit in the correspoding stereo detector, thus allowing the hit position in the Z coordinate direction to be inferred with some precision. Matched hits are possible only in layers 1, 2, 5 and 6 of the strip tracker barrel, and rings 1, 2 and 4 of the strip tracker endcap. There is a clear difference in the frequency of zero hits for the signal and background distributions. However, using these distributions for a data-driven estimation of background contamination looks difficult because the distributions are very similar in shape.

## 5.7 Data-Driven Estimation Using Relative Isolation

Another possible data-driven estimation of the QCD background is based on the extrapolation of *relative isolation*. Relative isolation is defined as the isolation

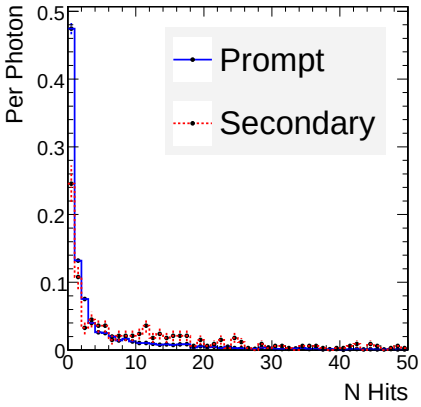


Figure 5.34: Hits in a cone of 0.2 about photon direction. Error bars correspond to Monte Carlo statistics.

sum, as described in section 5.4.1, divided by the $p_T$ of the photon candidate. As with absolute isolation cuts, relative isolation is a strong discriminator between prompt photons and QCD background. The advantage of relative isolation is that it is a more smoothly-varying quantity that can potentially be extrapolated from a region dominated by QCD background (high relative isolation) to the signal region (relative isolation near zero). Figure 5.35 shows relative isolation distributions for signal and background photons with $E_T$ of at least 100 GeV. The number of QCD backround photons with relative isolation near zero is dramatically larger than what would be predicted from an extrapolation from high relative isolation (the background relative isolation displays a peak near zero). Thus the use of relative isolation as a background estimation method is found not to be powerful. It is described here for completeness.

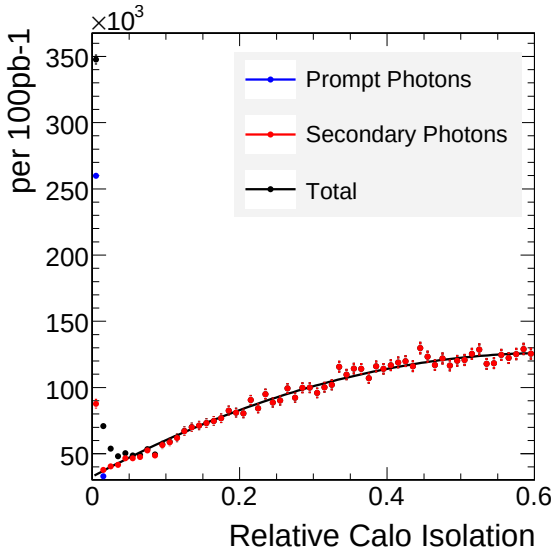


Figure 5.35: Relative calorimeter isolation for photons passing track isolation cuts.

## 5.8 Estimation of Photon Selection Efficiency

In order to validate the Monte Carlo modeling of photon selection efficiency, one could use a sample of prompt photons that is known to be pure by some independent means. As discussed above, such a sample is not at hand. An alternative is to use Z → ee events, with the electron as a proxy for the photon. The CMS calorimeters have nearly the same response to an incident electron as to an incident photon. Thus the ECAL and HCAL responses to *unconverted* prompt photons and to *low-bremsstrahlung* prompt electrons are indistinguishable. With this simplification the effects of energy leakage from the ECAL into the HCAL (and hence into the HCAL isolation sum) and of energy deposited in the ECAL but not clustered with the photon supercluster are measurable directly with electron events.

The comparison of photons and electrons is complicated by material interactions in the tracker; photons convert, and electrons emit bremsstralung radiation. In both cases the effect is to broaden ECAL energy deposition in the $\phi$ direction because electron trajectories bend in a magnetic field proportional to their transverse momenta whereas photons do not. But the degree of broadening is different between electrons and photons.

Inefficiency of tracker isolation cuts on prompt photons arises from two sources: accidental overlaps of unrelated tracks, and tracks arising from photon conversions. Unrelated tracks could come from nearby hadronic jets, from the *underlying event.*[7] The effect of such tracks, which have nothing to do with the photon object *per se*, can be modeled using $Z \rightarrow ee$ or $Z \rightarrow \mu\mu$ events, or even by simulating the photon as an arbitrary vector in a QCD event. To take into account the different amounts of jet activity in Z events and in photon events passing the final event selection which requires much jet activity, efficiency can be parameterized by the number of jets in the event, or angle to nearest jet.

A converted prompt photon may have conversion tracks contributing to the isolation sum. This effect is small because conversion tracks are infrequently reconstructed in the standard track collection. It is further minimized by the use of an inner-radius veto in the isolation cone. An electron may also have conversion tracks contributing to the isolation sum, as a result of bremsstralung

[7]Underlying event: low transverse energy collisions between the remnants of the protons involved in the hard scatter.

and subsequent conversion of bremsstrahlung photons.

The effect of material interactions on the isolation sums of electrons and photons is a priori different because of the reasons discussed above. However, electrons can be separated into low-bremstrahlung and high-bremstrahlung categories, based on the degree to which the reconstructed electron track momentum changes between the point of closest approach to the interaction point and the point of entrance into the ECAL. If necessary, the isolation efficiency of photons can be bounded by that of low-bremsstrahlung and high-bremsstrahlung electrons.

The method for measuring electron isolation efficiency is as follows:

1. Isolate a pure sample of Z $\rightarrow$ ee events using a requirement of one electron passing tight cuts including isolation, and one electron passing looser cuts that don't require isolation. The electron pair is required to have an invariant mass within 9 $GeV/c^2$ of the Z boson mass.

   The following cuts are used:

   - tag electron: $E_T$>20 GeV, isolated
   - probe electron: $E_T$>20 GeV
   - Z-boson mass constraint: $|m_{\mathrm{e^+e^-}} - m_{\mathrm{Z}}| < 9\,\mathrm{GeV}/c^2$

2. Measure the efficiency as the ratio of the number of probe electrons passing the tight selection to the total number of probe electrons. To avoid biasing the measurement, the two leading-$p_T$ electrons in a given event (and only

| **Cut Level** | **Z+jets** | $t\bar{t}$ | **W+jets** | **QCD** | **S/B** |
|---|---|---|---|---|---|
| total events | 618794 | 82680 | 6.7e6 | 2.85e11 | |
| >=1 iso. ele | 116499 | 9881 | 819877 | 1.5e6 | 0.05 |
| >=2 ele, incl. 1 iso. | 40761 | 1305 | 2620 | 1794 | 7 |
| pass Z mass | 36223 | 187 | 432 | 162 | 46 |

Table 5.6: Selection efficiency and signal to background for tag-and-probe measurement of isolation efficiency.

those two) are considered as tag and probe candidates; if both electrons pass the tag criteria, two tag-and-probe pairs are generated for the single pair of electrons.

The backgrounds considered are $t\bar{t}$, W+jets, and QCD multi-jet production. Table 5.6 lists the selection efficiencies on Z+jets signal and on backgrounds, for several stages of the cuts selection.

Figure 5.36 shows the measured and Monte Carlo truth electron isolation efficiency as a function of $E_T$. The red points show the values measured with the tag-and-probe method; the black points show the efficiency calculated using Monte Carlo truth information. Figure 5.37 is the same as figure 5.36, except that it shows only barrel electrons. Figure 5.38 shows the efficiencies versus $\eta$. Backgrounds, which are expected to be small, are not included in these figures. Figure 5.36 also shows the Monte Carlo truth photon isolation efficiency, the quantity which the electron measurement is used to predict.

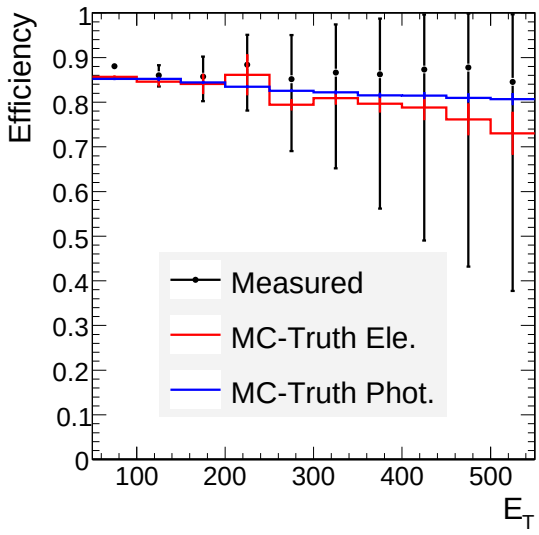


Figure 5.36: Isolation efficiency as a function of reconstructed $E_T$of the electron and photon.

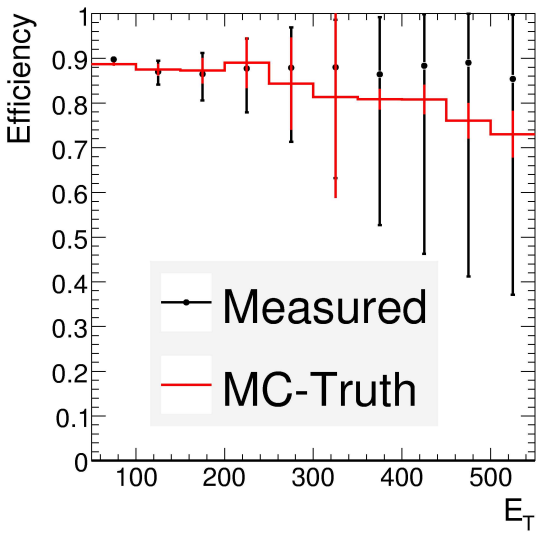


Figure 5.37: Isolation efficiency as a function of reconstructed $E_T$of the electron, barrel ($-1.4 < \eta < 1.4$) only.

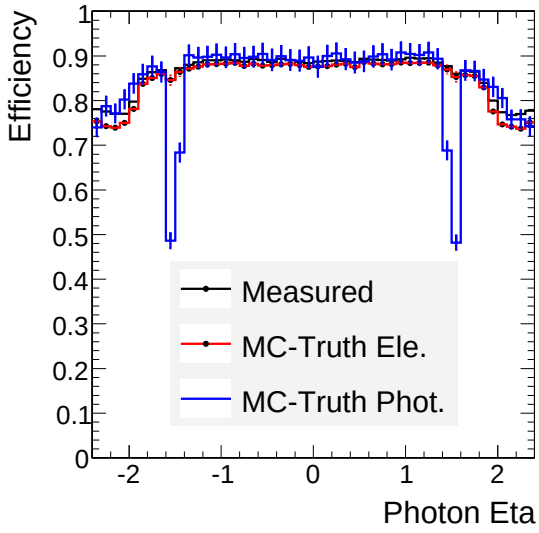


Figure 5.38: Isolation efficiency as a function of reconstructed $\eta$ of the electron and photon.

## 5.9 Estimation of Electron Veto Efficiency and Electron→Photon Fake Rate

As discussed in section 5.5.1, electrons are a background to photons if the electron track is not reconstructed. The electron→photon fake rate can be measured with Z → ee events using a tag and probe methodology similar to what is described in section 5.8, as follows:

1. Isolate a pure sample of Z → ee events using a requirement of isolated superclusters
   - tag supercluster: $p_T$>20 GeV, isolated
   - probe supercluster: $p_T$>20 GeV, isolated
   - Z-boson mass constraint: $|m_{e^+e^-} - m_Z| < 9\,\mathrm{GeV}/c^2$
2. Measure the photon fake rate as the ratio of the number of probe superclusters for which no track is reconstructed to the total number of probe superclusters. To avoid biasing the measurement, the two leading-$E_T$ superclusters in a given event (and only those two) are considered as tag and probe candidates; if both electrons pass the tag criteria, two tag-and-probe pairs are generated for the single pair of superclusters.

Table 5.7 lists the selection efficiencies on Z+jets signal and on backgrounds, for several stages of the cuts selection. Figures 5.39 and 5.40 show the fake rate

| Cut Level | Z+jets | $t\bar{t}$ | W+jets | QCD | S/B |
|---|---|---|---|---|---|
| total events | 618794 | 82680 | 6.7e6 | 2.85e11 | |
| >= 1 iso. SC | 69027 | 7330 | 327433 | 1.78e11 | |
| >= 2 iso. SC | 10591 | 194 | 472 | 9.2e10 | |
| >= 1 iso. SC, inc. 1 ele. | 10386 | 187 | 398 | 117 | |
| pass z mass | 8876 | 20 | 69 | 13 | |

Table 5.7: Selection efficiency and signal to background for tag-and-probe measurement of electron→photon fake rate.

measured from the tag-and-probe, and the Monte Carlo truth fake rate.

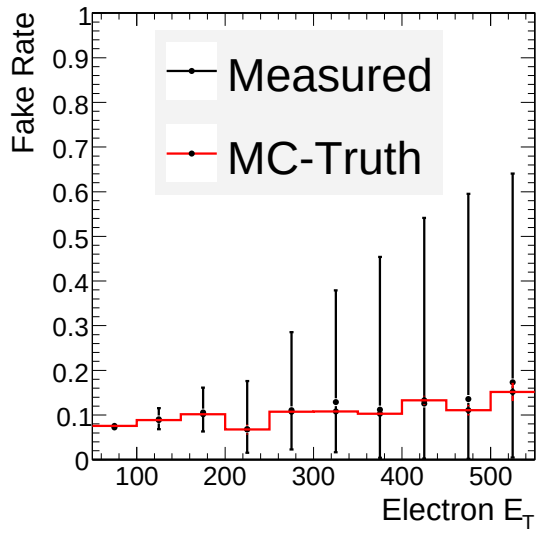


Figure 5.39: Electron→photon fake rate as a function of electron $E_T$.

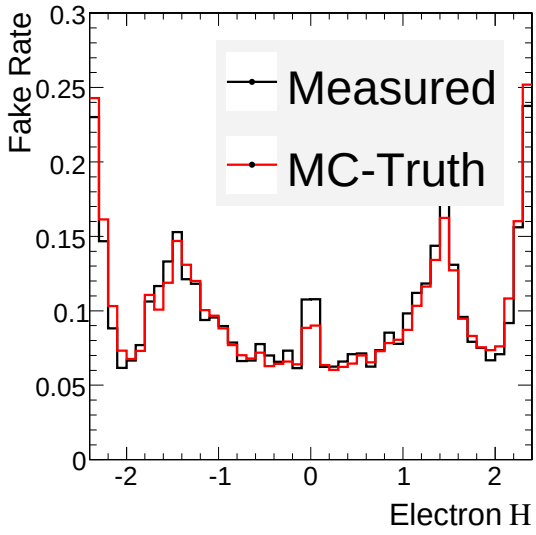


Figure 5.40: Electron→photon fake rate as a function of electron $\eta$.

The overall trend of the fake rate is a gradual rise with electron transverse energy. Bremsstrahulung energy loss by electrons increases linearly with electron energy, and is the dominant energy loss mechanism above about a few hundred MeV (the exact critical energy is dependent on material). Therefore increasingly energetic electrons emit increasing amounts of energy as bremsstrahlung photons. But the fractional energy loss of the electron is approximately con-

stant with increasing energy. The mechanism for electron track loss is not an increasing probability for the electron to lose all or most of it's momentum to a bremsstralung photon, but rather an increasing likelihood of large track curvature changes early in the electron trajectory, leading to pattern-recognition failures. As such, increased rejection of electron→photon fakes could be achieved by changes in pattern recognition (e.g. increasing the allowed track curvature changes in GSF track reconstruction, or using strip-tracker seeded rather than pixel-seeded electrons), with the probable cost of decreased efficiency on converted prompt photons. Given the low rate of electron background, such a change would probably result in a less optimal selection.

### 5.9.1 Determination of Electron Background

Once the electron→photon fake rate is established, the electron background to the photon selection can be estimated by measuring the number of muon events that pass the full photon selection, but with the muons as a proxy for photons. Every process that produces a prompt muon has a corresponding prompt electron production. Hence the electron→photon fake rate could applied to the number of selected muon events (as a function of $p_T$ and $\eta$ of the muon) to estimate the electron background.

## 5.10 Estimation of $Z \rightarrow \nu\bar{\nu}$ MET Spectrum from Photons

Once the photon $p_T$ spectrum is measured for events passing the SUSY MET + jets selection criteria, the *MET-like quantity* that is the vector sum of photon $E_T$ and calorimeter MET is computed. For the purposes of this Monte Carlo simulation study, the combined effects of photon acceptance and reconstruction and isolation efficiency were corrected using an ad hoc conversion factor obtained as the ratio of the transverse energy distributions of generated photons to reconstructed photons. The ad hoc conversion factor was applied directly to the MET-like quantity.

Sources of differences between the $E_T$ spectrum of photon events and $Z \rightarrow \nu\bar{\nu}$ events are listed below. When this measurements is performed on real data, individual corrections that are particular to each source of difference, and are derived from measurements of collider data, must be applied in order to reduce the overall dependence of the measurement on Monte Carlo simulation.

1. The trigger efficiency of the trigger used to select photon events (HLT1EMHighEt) and to select events in the SUSY search (HT+MET).

2. The photon identification and selection efficiency, which is dominated by the efficiency of isolation cuts.

3. The photon acceptance: photons with $|\eta|$ greater than 2.4 are not accepted

by the selection, whereas Z $\rightarrow \nu\bar{\nu}$ events have no such acceptance loss. The efficiency of the photon acceptance will be extracted from Monte-Carlo simulation.

4. The branching ratio of Z to neutrinos.

5. Any differences expected between Z and photons based on theoretical considerations.

### 5.10.1 Acceptance Correction

Photon acceptance is limited to the $\eta$ range for which there is tracker and ECAL coverage. Tracker coverage is required because of the necessity of applying tracker isolation cuts. Invisible Z, however, is accepted out to the full range $\eta$ over which MET is computed, $|\eta| < 3$. After the the photon MET-like distribution is corrected for photon identification efficiency, it must be corrected for this acceptance difference. The acceptance correction will have to depend heavily on Monte Carlo.

There are two components to the acceptance correction: 1) a correction for the fiducial cut removing the barrel-endcap interface region from photon acceptance and 2) a correction for photons outside the tracker acceptance, $|\eta| > 2.4$. The first correction can be made without relying on Monte Carlo simulated data, simply by interpolation of the measured photon $\eta$ distribution. The latter correction must be based on Monte Carlo generated photons, checking that the correction is

consistent with an extrapolation from the measured photon $\eta$ distribution.

### 5.10.2 Comparison of Generated and Reconstructed Photons

The same selection that was run using reconstructed photons was also run using generated photons, in order to demonstrate the magnitude of the required acceptance corrections as well as the level of bias induced by photon acceptance loss and selection efficiency. Figure 5.41 shows the ratio of the number of events passing cuts on reconstructed photons to the number of events passing cuts on generator-level photons. The ratio is shown for events passing on the requirement of a photon with $E_T$ at least 100 GeV, and for events passing the full selection criteria. The ratio is shown as a function of $E_T$, and for the full selection also as a function of MET-like quantity.

In Figure 5.44 the MET-like spectrum is compared to the actual MET spectrum in Monte Carlo simulated $Z \rightarrow \nu\bar{\nu}$ events. Table 5.8 lists the Monte Carlo truth estimate of $Z \rightarrow \nu\bar{\nu}$ events passing the model SUSY selection, the estimate obtained using photon events, and corresponding statistical and systematic errors.

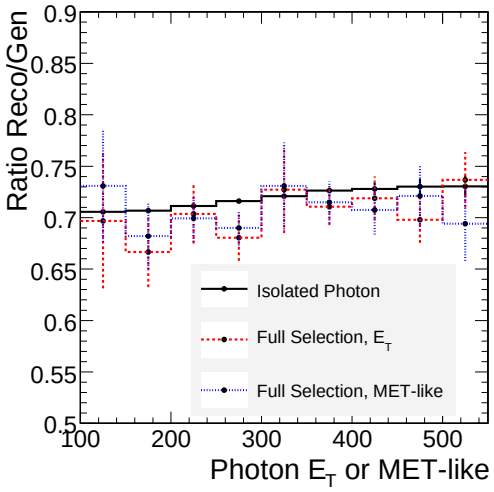


Figure 5.41: Ratio of reconstructed to generator-level photons, after selections.

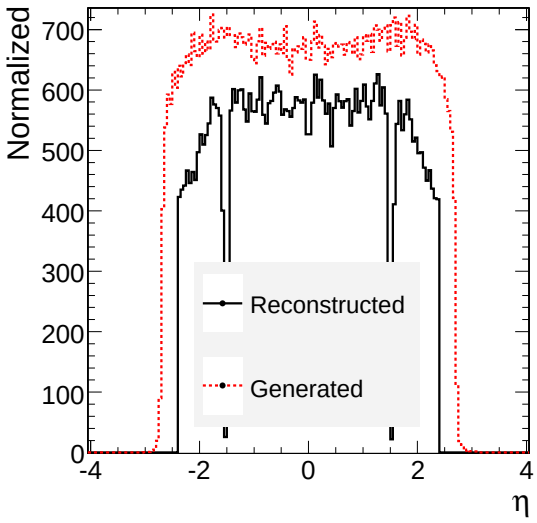


Figure 5.42: Distribution of $\eta$ of reconstructed and generated photons. Photons with $E_T$>100 GeV. Reconstructed photons pass isolation cuts.

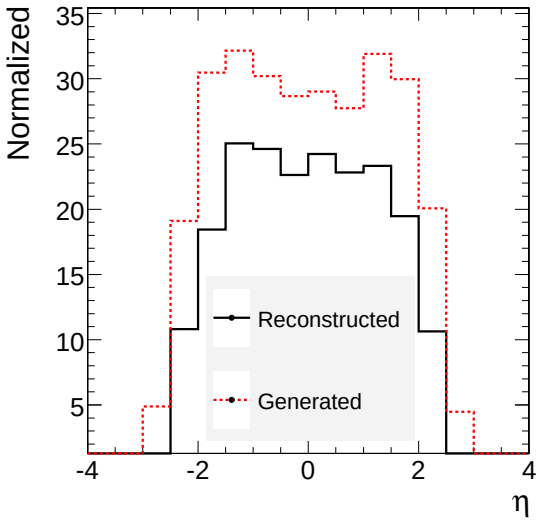


Figure 5.43: Distribution of $\eta$ of reconstructed and generated photons, events passing full selection.

| | |
|---|---|
| **MC-truth $Z \rightarrow \nu\bar{\nu}$ background** | 35 |
| **estimate of $Z \rightarrow \nu\bar{\nu}$ background** | 29 |
| **statistical uncertainty** | 2.7 |
| **systematic photon isolation** | 4.8 |
| **systematic QCD background** | 0.75 |
| **systematic electron background** | 1.5 |

Table 5.8: Monte Carlo truth $Z \rightarrow \nu\bar{\nu}$ background, estimate provided by data-driven methodology, and associated errors

## 5.11 Signal Contamination to Photon Plus Jets

If SUSY or other physics beyond the Standard Model does exist, it could show up in the photon event selection. That would result in an overestimation of the invisible Z background. Depending on the level of contamination a discovery to be missed or reported with a reduced level of significance. To test the sensitivity of the photon selection to signal contamination, the selection was run on Monte Carlo simulations of the CMS SUSY Low Mass (LM) points. The LM points are test points in the mSUGRA parameter space chosen to cover a range of different experimental signatures. The points were chosen to be accessible to early LHC running. Each point also uniquely determines a production cross section. A detailed description of the points can be found in chapter 13 of Ref. [10]. Table 5.9 shows the numbers of events selected for several of the LM points. SUSY contamination of the photon selection is not significant for any of the LM points tested.

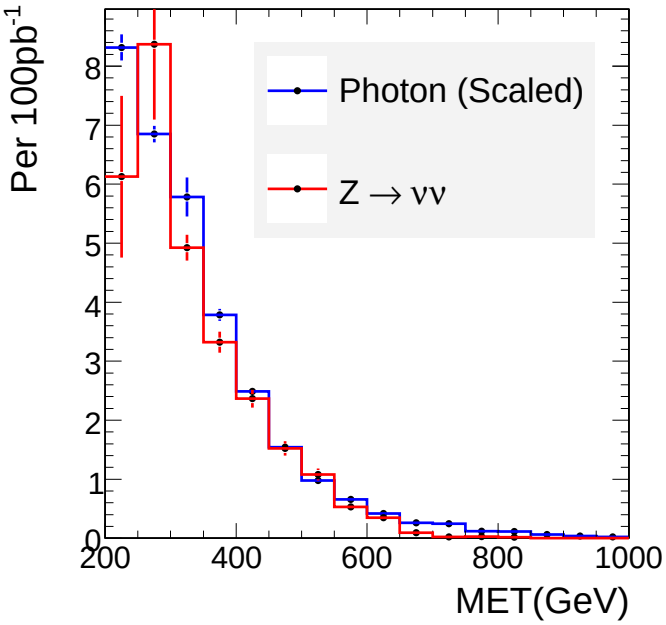


Figure 5.44: The spectrum of the transverse component of MET and photon $E_T$, corrected and scaled according to the Z $\rightarrow \nu\bar{\nu}$ branching ratio, with the quantity it is intended to predict, the MET spectrum in Z $\rightarrow \nu\bar{\nu}$ events.

## 5.12 Comparison of Photon and Z outside the Signal Region

Photon production is similar to Z boson production, but it is not the same. In order to maximize the power of the measurement described in this chapter, the ratio of the two must be known with as much precision as possible. Both the differences in hard production as seen in, for example, figure 5.1, and the potential for a background due to *photon fragmentation*[8] must be considered. However, photon production has not received as much attention as W and Z

[8]Photon fragmentation: the production of a photon that takes nearly all of the quark's momentum in the quark fragmentation process.

| **Cut Level** | **prompt** | **LM1** | **LM2** | **LM3** | **LM4** | **LM5** |
|---|---|---|---|---|---|---|
| 1 iso. phot eT > 100 GeV | 5.1e+04 | 9.6 | 2.4 | 6 | 4 | 2 |
| njets >= 3 | 4.8e+03 | 5.8 | 1.6 | 4.7 | 3.1 | 1.6 |
| lead jet cuts | 1.1e+03 | 4.1 | 1.3 | 3.3 | 2.1 | 1.3 |
| 2nd jet cuts | 7.1e+02 | 2.8 | 1 | 2.5 | 1.8 | 1.2 |
| QCD angular cuts | 5.2e+02 | 2 | 0.73 | 1.6 | 1.3 | 0.82 |
| HT>500 GeV | 2e+02 | 1.7 | 0.66 | 1.2 | 1.1 | 0.7 |
| MET-like > 200 GeV | 1.1e+02 | 1.6 | 0.55 | 1.1 | 0.81 | 0.6 |

Table 5.9: Photon event selection results for SUSY signal types. LM points 6, 7 and 8, not shown, result in smaller contamination. The prompt photon selection is shown in the second column for reference.

production. This is true both on the experimental side and on the theoretical side. Therefore a test of the similarity of photon and Z production using collider data is essential. Fortunately such a test can be made using the CMS experiment by studying photon and Z $\rightarrow \mu\mu$ data while loosening or releasing some of the SUSY search cuts. Figures 5.45-5.49 demonstrate the evolution of the comparison of photon and Z production through the SUSY search cuts, in the following order:

- Boson $p_T$>100 GeV (figure 5.45)
- At least 3 jets with uncorrected $E_T$ > 30 GeV and $|\eta| < 3$ (figure 5.46).
- $|\eta| < 1.7$ and $E_T$ >180 GeV on the leading jet (figure 5.47).
- Second leading jet $E_T$ > 110 GeV (figure 5.48).
- QCD angular cuts (figure 5.49)

Each figure displays the generated Z boson $p_T$ spectrum, the generated photon $p_T$ spectrum, and the reconstructed, isolated photon $p_T$ spectrum, scaled accord-

ing to the reconstruction and isolation efficiency calculated from the Monte Carlo truth photon reconstruction and isolation efficiency. The difference in efficiency of QCD angular cuts on photons and on Z events is not understood.

## 5.13 Conclusions

In this thesis, as was done in [20], it has been shown that high $E_T$ photons are objects that are under good experimental control. Backgrounds, the dominant ones being prompt electrons from W, Z and $t\bar{t}$ production and decay to electrons, in which the electron track is lost, and QCD production of secondary photons through neutral meson decay, are small and their contribution to the photon selection can be bounded with data-driven methods. Photon selection efficiency can be estimated with tag-and-probe methods using $Z \rightarrow ee$ production. It has also been shown to be promising and plausible that the production cross section of high-$p_T$ photons is similar to that of Z bosons and can be used to estimate Z production. Because the use of photon production at high $E_T$ to predict weak boson production is a novel technique at collider experiments, only a limited amount of theoretical and experimental attention has been given to the comparison of photon and Z events. Therefore, a key step in establishing the validity of this technique will be the measurement of the Z to photon ratio in regions of phase space separate from the regions in which photon production will be used to predict the weak boson production backgrounds to new physics. For example

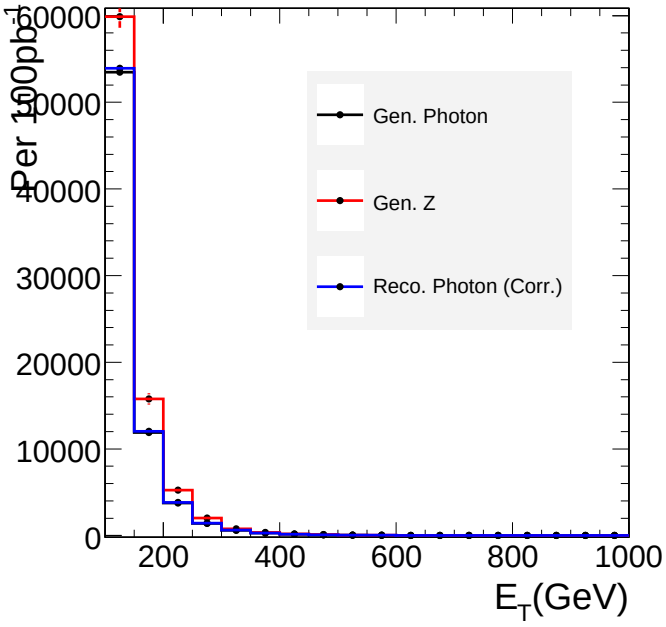


Figure 5.45: $P_T$ distributions of Z bosons, generated photons, and efficiency-corrected reconstructed photons.

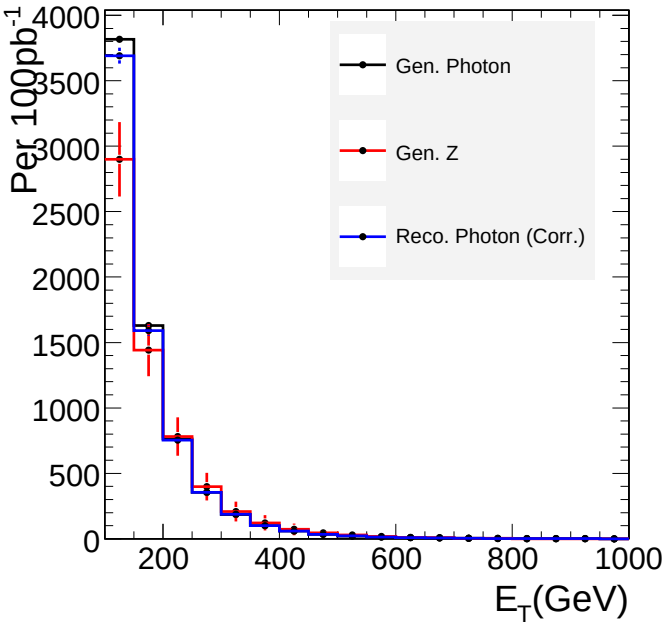


Figure 5.46: $P_T$ distributions of Z bosons, generated photons, and efficiency-corrected reconstructed photons, events passing the cut on number of jets.

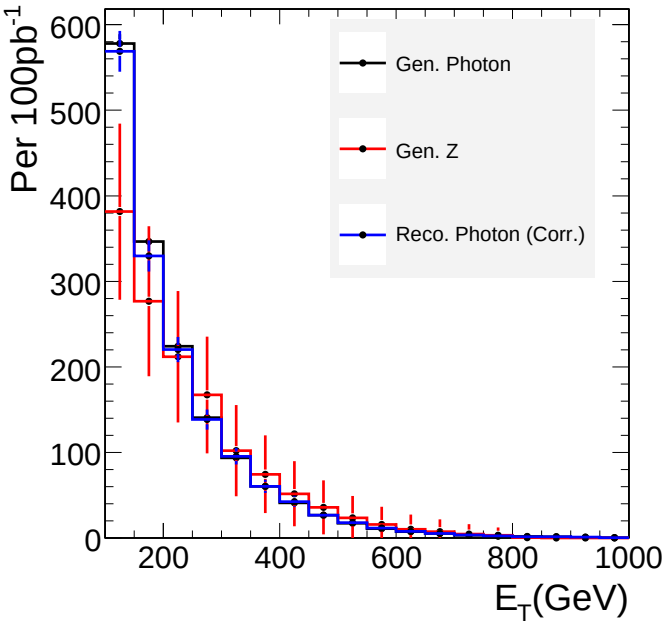


Figure 5.47: $P_T$ distributions of Z bosons, generator-level photons, and efficiency-corrected reconstructed photons, events passing cuts up to lead jet $E_T$ and $\eta$ cuts.

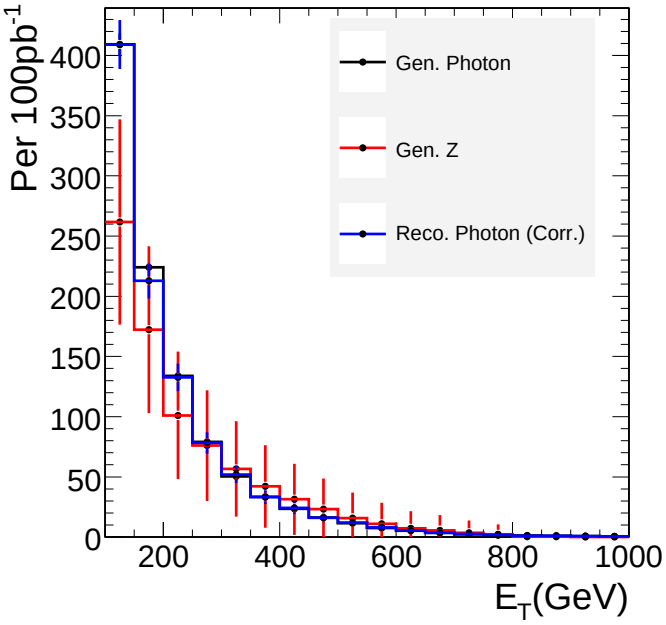


Figure 5.48: $P_T$ distributions of Z bosons, generator-level photons, and efficiency-corrected reconstructed photons, events passing cuts up to second lead jet $E_T$ cut.

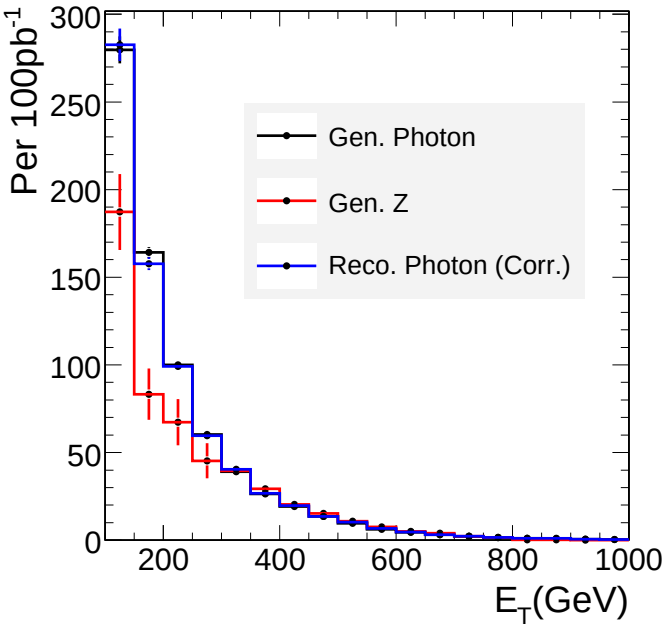


Figure 5.49: $P_T$ distributions of Z bosons, generator-level photons, and efficiency-corrected reconstructed photons, events passing cuts up to QCD angular cuts.

such a comparison could be made in the boson $p_T$ range 100-200 GeV, and with a range of additional selections on number and $E_T$ of hadronic jets. Decreasing the photon $p_T$ range of interest in turn requires a re-evaluation of backgrounds to photon production. Given the independent importance of photon and Z measurements to new physics searches, irrespective of their comparison, and the promising methodology described here, such effort is well warranted.

# Chapter 6

# Construction of the Tracker Outer Barrel

## 6.1 Introduction

The CMS Tracker Outer Barrel (TOB) was integrated at CERN from detector modules assembled and tested at the University of California at Santa Barbara and at Fermi National Accelerator Laboratory. For details of the production and quality control of modules, see respectively Refs. [24] and [25]. Brief descriptions of TOB geometry and the tracker DAQ are given here; for further details refer to Ref. [2]. This chapter describes the quality control and system tests performed during TOB integration, including procedures, performance, and a discussion of problems encountered.

### 6.1.1 TOB Geometry and Terminology

The TOB is made of six layers of silicon strip tracking detectors. Each layer is assembled from *rod* substructures, carbon fiber frames about 1 meter long which hold detector modules in six positions. The TOB detector modules consist of two silicon strip diode wafers, each 500 microns thick and 10 cm by 10 cm in length, joined to form 20 cm long silicon strips. Layers numbered 3 through 6 have *single-sided* modules, single modules with strips oriented parallel to the beamline (z axis), referred to as the r$\phi$ orienttation. The inner two layers, layers 1 and 2, have *double-sided* modules: modules with strips in the r$\phi$ orientation are joined back-to-back with modules whose sensors are rotated with a 100 milliradian stereo angle. Layers 1-4 are made of modules with 512 readout channels with a strip pitch of 180 microns. The outer two layers, layers 5 and 6, contain detector modules with 768 strips and a pitch of 120 microns. Corresponding to the types of detectors in the different layers, there are different types of rods: double-sided (DS) in layers 1 and 2, single-sided four-chip (SS4) in layers 3 and 4, and single-sided six-chip (SS6) in layers 5 and 6.[1] Rods are inserted into the TOB and tested in units of a *cooling-segment*, the smallest set of rods that can be cooled, controlled and read out together, consisting of between 8 and 22 rods each.

[1]The four-chip/six-chip terminology refers to the number of readout chips used per module. Each chip reads out 128 strip channels.

### 6.1.2 DAQ

The data acquisition system (DAQ) of the silicon tracker consists of two systems: the read-out chain to transfer data from the detector to the counting room, and the bidirectional digital link to transfer the clock and trigger to the detector and configure the front-end electronics. During TOB integration final DAQ components were used whenever possible and consistent with having a stable system. Clock and trigger signals were produced by an ad hoc system (TSC+TFB+TTCex) and distributed by the Front End Controller (FEC) together with configuration for front end electronics. The analog readout signals coming from the detector are digitized by the Front End Driver (FED). The FED and FEC are connected to the detector via optical links. The Tracker Xdaq application (version rc0306) was used to access the hardware devices including FEC, FED, and trigger system. We used a prototype of the final tracker power supply system to supply high and low voltage to the rods and low voltage to the control power line. Power supply modules of Service Pack Two were used, along with a mixture of production and pre-production LIC cables [26]. The detector was cooled with the final on-detector cooling manifolds and cooling fluid (C6F14), at a coolant temperature of 15C. The power-supply system was interlocked using a prototype of the final interlock system which used as inputs a subset of the available temperatures from the rods, typically three thermistors measuring the temperature of the rod cooling pipes. A PVSS application was used to control

and monitor the power system and the interlock system.

## 6.2 Silicon Detector Defects

Silicon strip detectors are reverse-biased diode junctions. A charged particle passing through a silicon wafer will deposit energy in the silicon through ionization[2]. Ionizing energy loss ultimately results in the creation of one electron-hole pair for every ≈3.6 eV of energy deposited in the silicon. The reverse-biased p-n junction configuration of the silicon means that the bulk of the silicon is depleted of conduction-band charge (in the absense of ionization by a charged particle), and is under a well-behaved electric field. This allows for efficient collection of the electron-hole pairs by electrodes at the edges of the silicon.

Each diode p-n junction is capacitively coupled to a charge-sensitive front-end amplifier. The coupling is made through a parallel-plate capacitor formed by the semiconductor implant of the diode (the p-type implant in CMS silicon strips), an insulating layer of silicon dioxide, and an aluminum strip connected to the input of the amplifier. Capacitive coupling is used in order to prevent the leakage current of the diode from entering the amplifier input, thus reducing noise and reducing the current-sinking requirements placed on the front-end amplifier. It is especially important for detectors that will be exposed to a high radiation flux, as is the case for the CMS tracker, because radiation damage will drastically increase

[2]Non-ionizing energy loss mechanisms are also present, for example bremsstrahlung radiation (by electrons) and inelestic collisions with silicon nuclei.

the leakage current over the lifetime of the detector. Figure 6.1 shows a schematic representation of a single diode strip of a CMS Tracker Outer Barrel module.

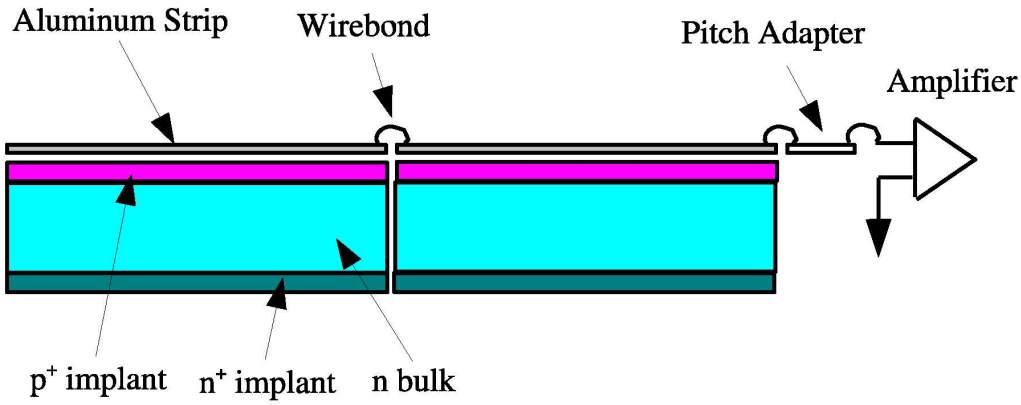


Figure 6.1: Schematic diagram of a TOB silicon strip channel.

During all stages of TOB detector assembly tests were performed to identify the following set of defects:

**Opens** Wirebonds can be broken or missing, or an aluminum strip can be broken. This type of fault, called an *open*, leaves some part of the aluminum strip unconnected to the readout amplifier, and hence signal will not be collected from all or part of the strip. Opens typically occur due to faults in wirebonding; they can also be the result of mechanical damage. A variety of more serious faults, such as shorts or pinholes, are treated by removing the wirebond between the sensor and the pitch adapter, thereby inducing an open. This is done in order to isolate the sensitive front-end amplifier

chip from instabilities caused by these faults.

**Shorts** A short occurs when two adjacent strip channels are electrically coupled. This can occur because of mechanical damage to the aluminum strips (smearing). Shorts typically do not result from contamination of the sensor surface by conducting particles because the sensor, including aluminum strips, is insulated by a thin silicon dioxide layer. Shorts only occassionally result from crossed wirebonds; this is rare because wirebonds are very thin, and covered with aluminum oxide, so conduction between touching wirebonds is actually quite low. Wirebond shorts usually only occur in cases of extreme mechanical damage.

**Pinholes** A pinhole occurs when the p+ implant of the strip diode and the aluminum strip (and hence the amplifier) become ohmically connected. The ohmic coupling of the amplifier input causes the amplifier to behave in unpredictable ways. With the increase in leakage current that will occur with radiation damage during the lifetime of the detector, the amplifier will be overwhelmed with current, causing it to stop functioning. Pinholes can be the result of either electrical or mechanical damage. Transient voltage differences across the capacitor formed by the aluminum strip can cause permanent conductive channels in the silicon dioxide insulating layer between them. Mechanical damage can cause pinholes by pushing part of the aluminum strip across the silicon dioxide into the p+ implant, while main-

taining contact with the rest of the aluminum strip. Such damage can be caused by the wirebonding procedure. Because in the CMS amplifier chip adjacent amplifier channels are coupled by a resistor, a pinhole in one channel can degrade or destroy the functioning of an entire amplifier chip (128 channels). Thus one major goal of the quality control effort of the CMS tracker was to prevent any strip diodes with untreated pinholes from going in to the CMS tracker. Pinholes can be treated by removing the wirebond between the silicon strip and the amplifier (usually done at the sensor to pitch adapter wirebond).

## 6.3 Testing Procedure

The testing procedure used the standard DAQ commissioning runs [27] and some supplemental tests. The aim was to verify the functionality of all components integrated in the TOB.

The procedure is as follows:

1. All rods of one cooling segment are inserted into the TOB
2. Control Ring Test
3. Rods are soldered to cooling manifold
4. Connection Run

5. Timing Run

6. Gain Run

7. VPSP Run

8. Pedestal Run, peak mode without bias

9. Pedestal Run, peak mode with bias

10. Pedestal Run, deconvolution mode

11. Calibration Run

12. DCU Readout Run

### 6.3.1 Control Ring Test

Within a cooling segment, rods are grouped into "control rings". A control ring is a chain of several rods connected to a FEC via a Digital Opto Hybrid Module (DOHM). The DOHM converts the optical control signals coming from the FEC into electrical signals. The Communication and Control Units (CCUs), one of which sits on every rod, distribute the control signal to the modules on the rods. The ring operates under a token-ring communication protocol. A detailed description of the tracker control system can be found in Ref. [28].

The control ring test verifies that the CCUs of all rods on the ring are visible to

the FEC, that the redundancy functionality of the ring[3] is working, and that the primary and secondary digital opto-hybrids of the DOHM are seen by the CCUs of the first and second rods in the ring, respectively. Since this test can be done without powering the readout chips on the rod, it was done before the soldering of the cooling manifolds and attachment of front-end power cables, allowing the rods to be easily removed in case of problems.

### 6.3.2 Connection Run

Every readout laser of a cooling segment is turned on in sequence, with the other lasers off, and its output is searched for in the FED channels. This tells us if any optical connection is broken or very poorly mated, and the mapping between laser and FED channel is necessary for subsequent DAQ running.

### 6.3.3 Timing Run

In the Timing run the front-end chips (APV) are given a synchronization trigger which causes them to emit a clock tick after a configurable delay period. The clock tick signal is read out from the FEDs in scope mode. By changing the relative phase of the trigger and the digitization sampling in the FED using the programmable delay chip that exists on each module (PLL), the entire clock tick can be reconstructed. Using the reconstructed ticks, trigger delays intrinsic to the

[3]The redundancy functionality allows one rod to be skipped in the control ring if its CCU is not functioning. This is achieved by using a parallel circuitry inside the rod that entirely bypasses the CCU.

hardware are measured, and accounted for by setting appropriate delays on the PLLs of every module.

The height of the reconstructed ticks is a more precise measure of the optical link quality than the simple "on or off" measure given by the Connection run. Typically from the Timing run it can be determined if it is necessary to clean the optical connections. Figures 6.2 and 6.3, respectively, show Timing run analysis plots that indicated that on-rod and off-rod optical connections which needed to be cleaned and remated.

### 6.3.4 Gain Run

The data from the TOB is read out via a series of optical fibers comprising a fiber of about one meter length along the rod, then a fiber about 5 meters long to exit the tracker volume, then a fiber a few tens of meters long which is finally optically coupled to the input of the FED. Because of the series of optical couplings, the optical gain of this readout chain varies significantly from channel to channel, and can change every time optical connections are remated. Furthermore the appropriate value of the current bias driving the readout lasers varies from laser to laser and varies strongly with temperature. The Gain run measures the gain of the optical readout chain and the optimal current bias point of the readout lasers by stepping the bias current driving the lasers. Outliers in the distributions of the gains and optimal bias points were searched for in

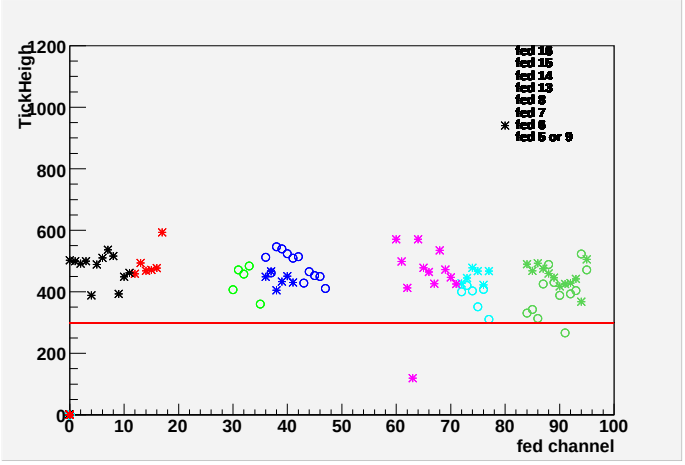


Figure 6.2: Plot of tick heights that indicated an optical connection needed to be reseated at the rod.

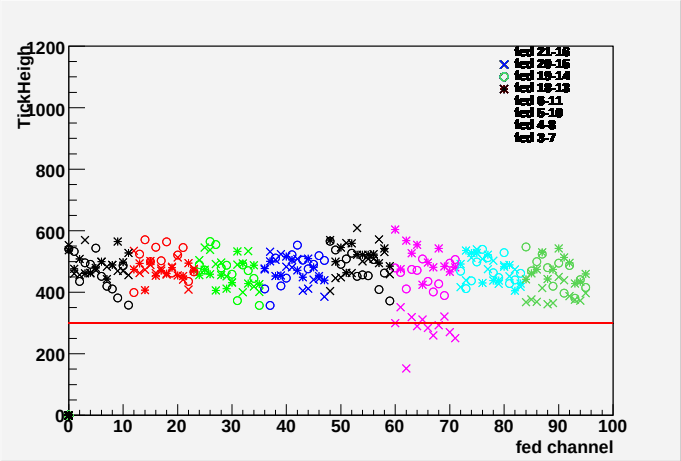


Figure 6.3: Plot of tick heights that indicated optical connections downstream of rods needed to be reseated.

order to identify any malfuntioning optical links, and the results of the run are required for correct operation of the detectors in subsequent runs. However the vast majority of malfunctioning links were not discovered using the Gain run, but by low tickheights measured in the Timing runs.

### 6.3.5 VPSP Run

The analog baseline of the APV chip's output data frame can be shifted by adjusting the chip's VPSP register. In normal tracker operation the baseline should be adjusted to allow for the maximum possible signal range while avoiding biasing the measurement of detector noise by moving the baseline too close to the edge of the linear dynamic range of the readout chain. This adjustment procedure was performed and outliers in the distribution of optimal VPSP values chosen were searched for in order to verify proper functioning of the APVs. No hardware faults were ever discovered with the VPSP run, but it was a useful indicator of DAQ problems.

### 6.3.6 Pedestal Runs

Pedestal runs were taken in deconvolution mode [29], and in peak mode with and without bias. The noise in both modes was verified to be within the expected range, and any unusual features were examined. A comparison of the noise in peak mode with bias voltage on and off was made in order to verify that all

modules could be properly biased; the noise is substantially higher if the module is unbiased, due to an increase in the capacitance seen by the front end amplifiers. Peak mode pedestal runs were used for this check because the intrinsic noise in peak mode is lower, allowing a better seperation of noise levels in biased and unbiased detectors.

To check the overall noise level the noise was normalized to account for the varying of gains of the optical links, and converted the noise into equivalent noise charge (ENC). The tick heights extracted for each laser in the previous timing run were used for the normalization. The conversion is made using the approximate calibration that the APV tick-height is roughly equal to the output expected from an input signal of 200,000 electrons. The expected accuracy of this calibration is better than 10%.

### 6.3.7 Calibration Run

The Calibration run uses the APV's charge injection ciruit and programmable readout latency to step though its response to charge injection. This scan is useful in identifying bad channels, particularly pinholes and damaged APV channels.

### 6.3.8 DCU Readout Run

The Detector Control Unit (DCU) is present on all detector modules, and on every rod's CCU module. It measures the temperatures on the hybrids and

silicon, the bias current, and the values of the low-voltages powering the rods. Particular attention is given to the measured temperatures because they allow us to verify that the rods are properly soldered to the cooling manifolds. A generic calibration of DCU ADC counts to physical value (temperature, voltage, or current) is approximate and masks small effects. Individual calibrations exist in the Tracker Construction Database for most DCUs. Those calibrations were used when they were available.

## 6.4 TOB Performance

### 6.4.1 Noise Performance

Figures 6.4 and 6.5 show histograms of the noise of about 90% of the silicon strips in the TOB, in peak and in deconvolution modes. This sample of strips is representative. The remaining 10% was difficult to include in the plot because of overlaps in data taking and other logistical problems, not because of functional differences. The mean noise is about 2100 electrons when the APV is operating in deconvolution mode and about 1500 electrons when in peak mode. For comparison, the signal of a minimum-ionizing particle passing through a TOB detector at normal incidence is about 40,000 electrons.

The features visible at the low end of the histogram are faulty channels, including dead APV strips and open channels. The high-end tail of the distribution

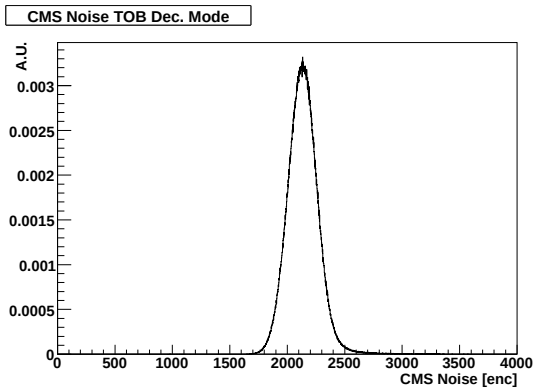


Figure 6.4: Histogram of strip noise in the TOB, deconvolution mode at 400V.

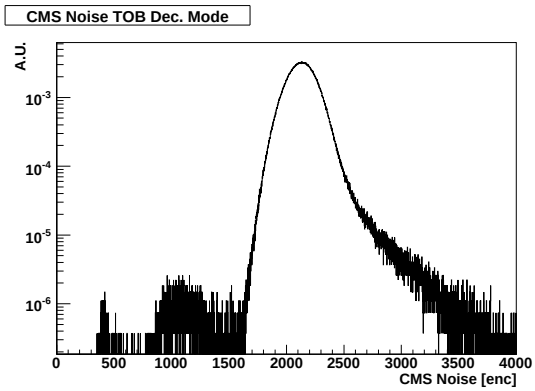


Figure 6.5: Histogram of strip noise in the TOB, deconvolution mode at 400V, logarithmic scale.

is made up of two features: modules with external noise pickup and a small number of modules that are not biased (discussed in section 5). Figure 6.8 shows the TOB noise histogram separately for each layer of the TOB.

### External Noise Pickup

TOB rods are susceptible to external noise pickup. The pickup, which can come from a variety of sources, displays a characteristic profile (often referred to as "wings"), most strongly on the module of the rod that is closest to the CCU and the power supply cable (see figures 6.9 and 6.10). On an event-by-event basis, the noise appears as a "linear common-mode", i.e. it can be removed by subtracting, event by event, a line from the data. Figure 6.11 shows the raw signal[4] for one module for six consecutive events in a run. Figure 6.12 shows the effect on the noise of subtracting a linear fit from every event. Because of the high signal-to-

[4]Raw signal is output of the readout chain, with per-channel pedestal subtracted, but no common-mode noise subtraction.

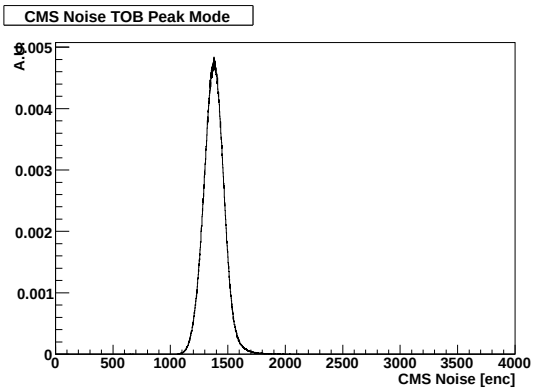


Figure 6.6: Histogram of strip noise in the TOB, peak mode at 400V.

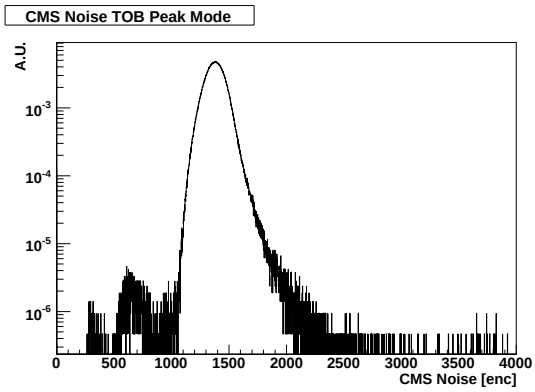


Figure 6.7: Histogram of strip noise in the TOB, peak mode at 400V, logarithmic scale.

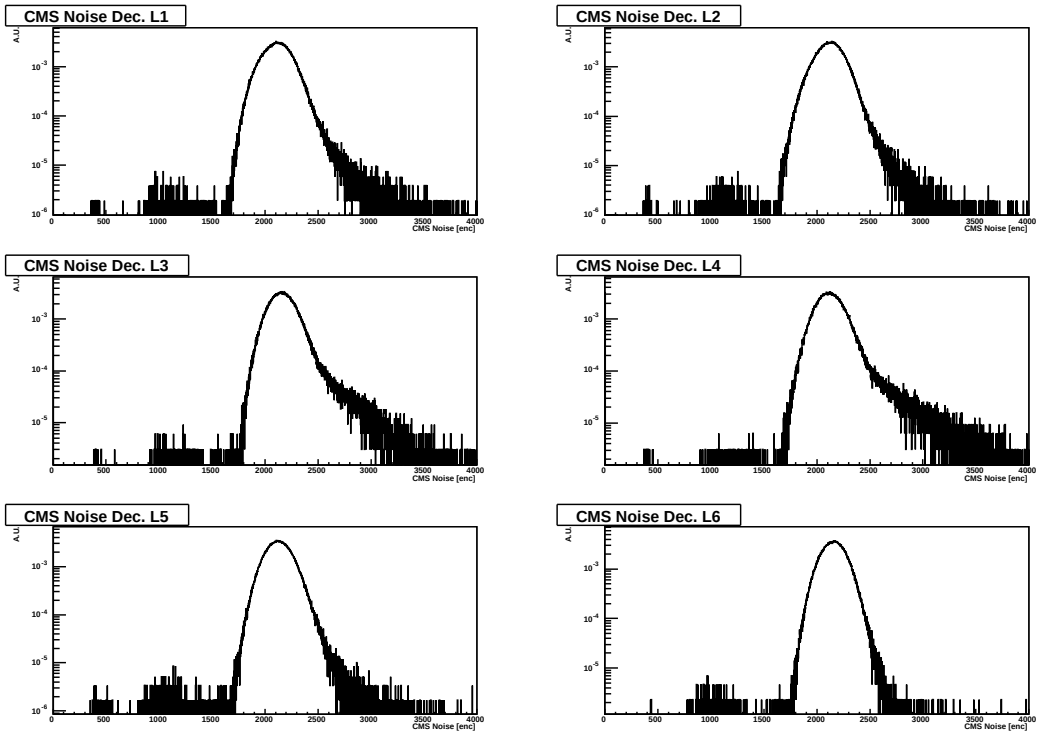


Figure 6.8: Histogram of strip noise in the TOB, separated by layer, deconvolution mode at 400V.

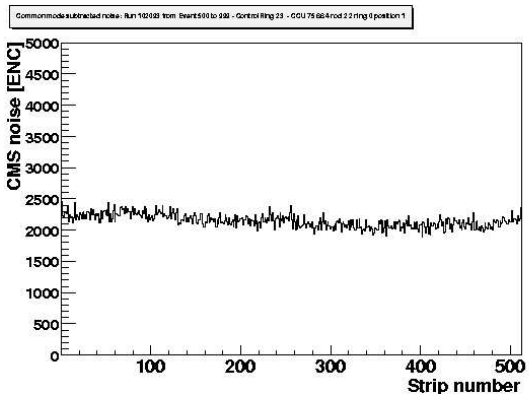


Figure 6.9: Noise profile of module 1 on a layer 4 rod. Module 1 is farthest from the power supply cable and the CCU. It shows no noise pickup.

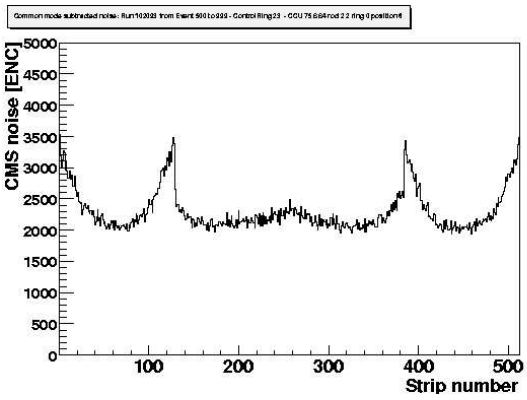


Figure 6.10: Noise profile of module 6 on a layer 4 rod. Module 6 is the module closest to the power supply cable and the CCU. This noise profile is an example of one of the worst cases of noise.

noise ratio of TOB modules, and the relatively low number of strips affected, the noise at current levels should not affect TOB tracking performance. The noise is amenable to standard noise supression techniques, e.g. grounding and shielding changes, proper dressing of cables, and filtering. Preliminary results from tests of the tracker in its final position in the CMS cavern suggest that this noise has been reduced, probably because of improved grounding of power supply cables.

### 6.4.2 Optical Link Performance

As described in section 2, the quality of the optical readout link was verified using the Timing runs and Gain runs. For each of the four possible gain settings of the readout lasers, a measurement of the gain of the link is extracted by varying the bias current to the lasers and measuring the response at the FEDs. The DAQ

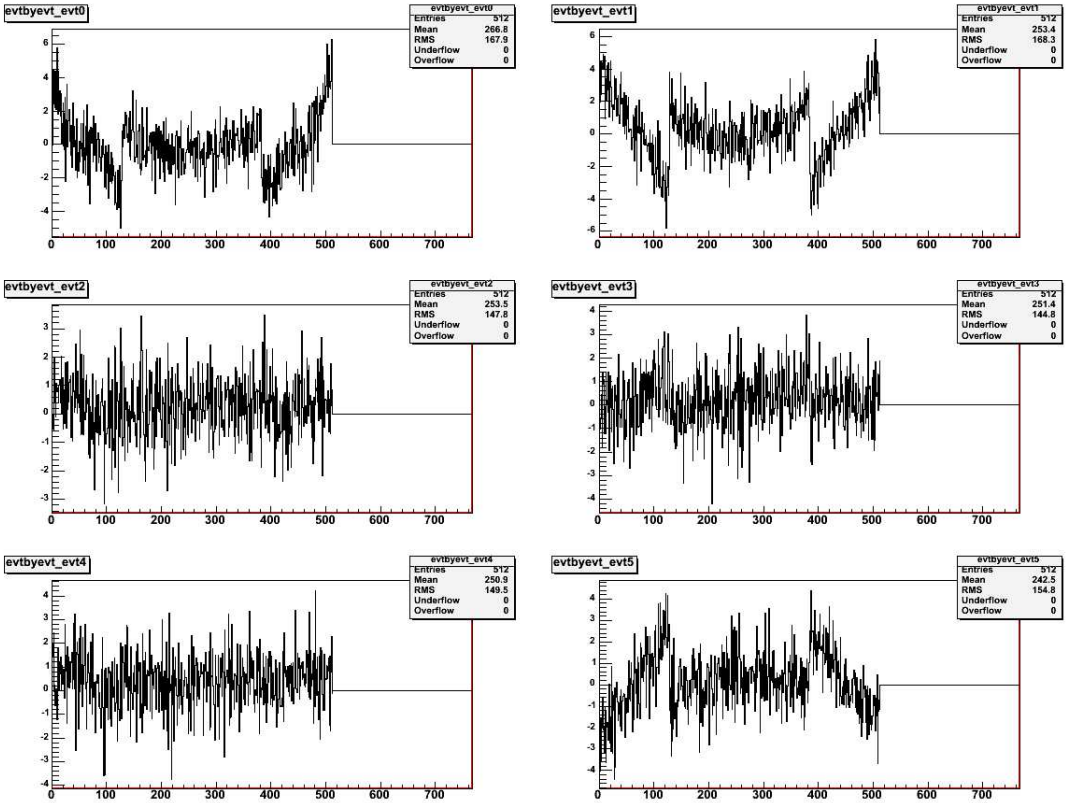


Figure 6.11: Signal on a module, six consecutive events in a run. The scales of the Y axes of the profiles are in ADC counts.

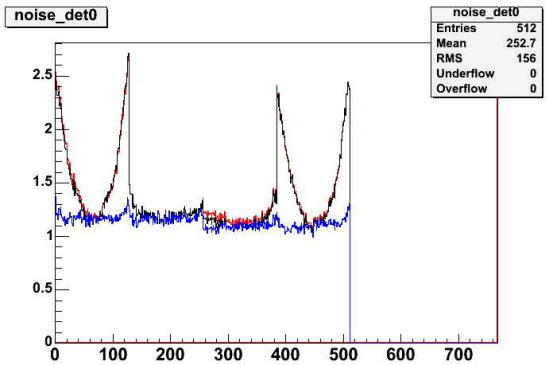


Figure 6.12: The effectiveness of linear common mode subtraction. The red and black profiles are raw noise and median common-mode subtracted noise. The blue profile is linear common-mode subtracted noise. The scale of the Y axis of the profiles is in ADC counts.

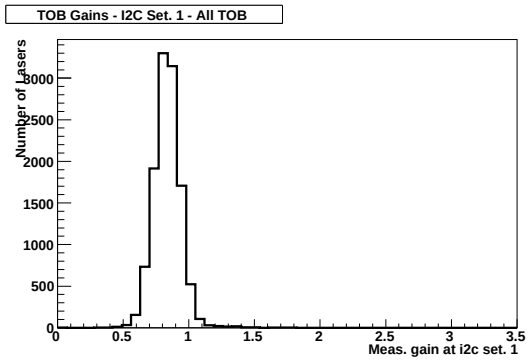


Figure 6.13: Measured optical readout link gains at laser gain setting 1, for all optical readout links of the TOB.

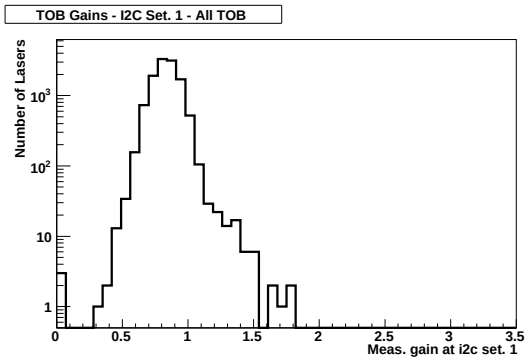


Figure 6.14: Measured link gains at gain setting 1, logarithmic scale.

software used this scan to compute optimal gain and bias settings for the lasers. No numerical cut on the measured gain was chosen to identify an optical link that was not working. Outliers in the gain distribution were searched for as indications that a link was of poor quality. The optical gain setting chosen by the tracker online software was recorded. Figures 6.13 and 6.14 show the measured gains of all lasers in the TOB at gain setting 1. An informal criterion used by the tracker community is that the optimal gain setting should not be the highest possible setting (the possible settings range from 0 to 3). This loosely corresponds to a gain cut of 0.8 at gain setting 2. Figure 6.15 shows the optimal gain settings for all lasers in the TOB.

### 6.4.3 Low-voltage Current and Bias Current Measurements

APV front-end chips on each rod are powered with one 2.5V and one 1.25V channel. The DOHM and the the CCUs of all rods on one control ring are powered

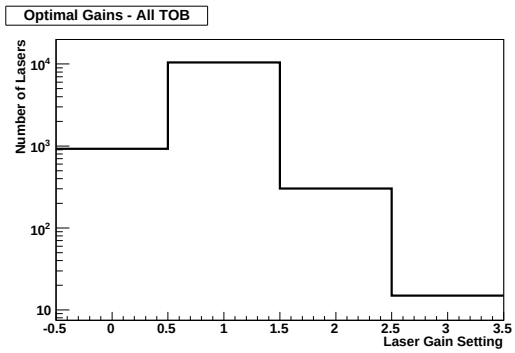


Figure 6.15: Optimal gain settings chosen by the tracker DAQ software for all lasers in the TOB.

by a separate 2.5V channel. Different power consumption is measured for the different rod flavors due to the different number of front-end devices present. The current drawn by control rings varies according to the number of CCUs present in the ring. Results are not presented here for current measurements on the control ring power supply, but verifying the correct current draw of control rings was a critical debugging tool. In particular it should be noted that the DOHM current draw varies according to the quality of its optical connection to the FEC. There are two bias channel per rod, feeding between 3 and 8 detector modules depending on the flavor of rod.

Figures 6.16, 6.17 and 6.18 show the current draws on the 2.5V channel, 1.25V channel, and bias channels, respectively, of all rods in the TOB. The current draws were measured during a pedestal run, with the detectors biased at 400V.

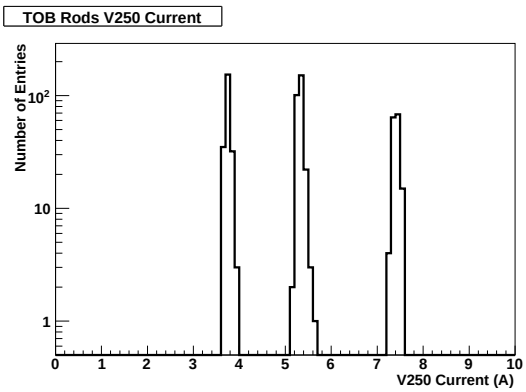


Figure 6.16: Currents on the 2.5V front-end supply channel. The three peaks correspond to the three rod flavors, SS4, SS6 and DS.

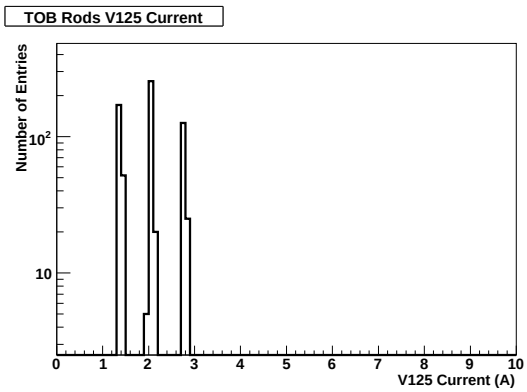


Figure 6.17: Currents on the 1.25V front-end supply channel. The three peaks correspond to the three rod flavors, SS4, SS6 and DS.

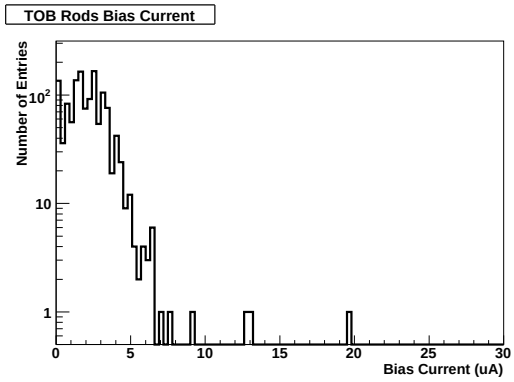


Figure 6.18: Currents on the high-voltage channels biasing TOB silicon detectors. High-current channels were investigated for stability of current with time.

### 6.4.4 Temperature Measurements

The temperatures of the front-end hybrid and the silicon sensors were read out from the DCU after the detectors had been powered and operated long enough to reach thermal equilibrium.

The uniformity of the temperatures across all rods in a given cooling segment was measured in order to verify that the rod cooling pipes were unobstructed and that no detector module showed unusual temperature. On two occasions rod cooling pipes were accidentally obstructed during soldering. These were repaired; all cooling pipes are fully functional. Figure 6.19 shows a temperature measurement of the rods in cooling segment 1.2.1. The elevated temperature of one rod is clearly visible. Figure 6.20 shows the temperature measurement of the same cooling segment after the blocked cooling pipe was repaired[5].

### 6.4.5 Bad Channel Count

Using data from pedestal and calibration charge injection runs bad channels were flagged according to the rules in Table 6.1. More explanation of the different types of channel faults can be found in Ref. [25]. The pedestal and calibration run data was normalized before these flags were calculated. The calibration pulse height was normalized by the multiplicative constant such that the median pulse height of the 128 channels of an APV was 1.0. The calibration pulse rise time was

[5]A typical difference in temperature of 1-2 °C occurs between different module positions reflecting the distribution of power and cooling along the rods.

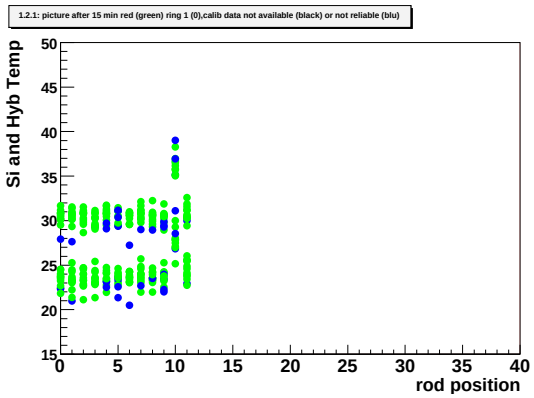


Figure 6.19: Temperatures on segment 1.2.1, showing faulty cooling pipe soldering. The figure shows hybrid and silicon temperatures measured by the DCU.

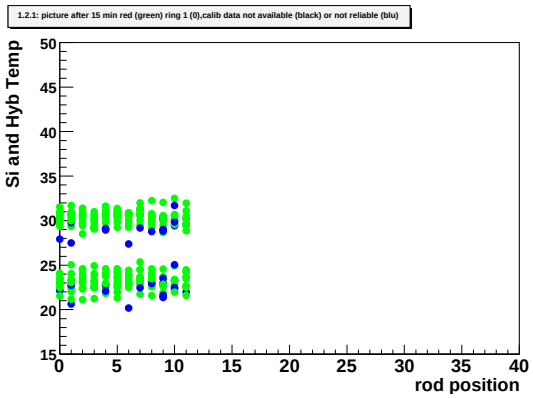


Figure 6.20: Temperatures on segment 1.2.1, after fixing faulty soldering.

normalized by an additive constant such that the median rise time was zero. The noise was normalized by a multiplicative constant such that the mean noise over an APV was 1 ADC count.

After this analysis the bad channel flags for each module were uploaded to the construction database, where all measurement results are stored. This database was used to compare the bad channels before transportation from UCSB and FNAL to CERN, after reception at CERN, and after integration into the TOB. The last measurements before transportation to CERN are performed in the multi-rod test (MRT), in which rods undergo thermal cycling between room temperature and -15°C. The last measurement at room temperature, after the cycling, is used as a reference. Table 6.3 shows the number of bad strips of each fault category. The number of new bad channels with respect to the MRT measurements is shown

| Flag | Corresponding Criteria |
|---|---|
| 2 | Low noise (70-90% normal) AND rise time between -4 and -0.5 |
| 4 | Low noise (70-90% normal) OR rise time between -4 and -0.5 |
| 8 | Low noise (<70% normal) AND rise time between -30 and -4 |
| 16 | Low noise (<70% normal) OR rise time between -30 and -4 |
| 32 | (Low noise (70-90% normal) AND rise time between -30 and -4 )OR<br>( low noise (<70%) AND normalized rise time between -4 and -0.5) |
| 64 | Higher than average noise and no other anomaly |
| 128 | Low calibration pulse peak height (<80% normal) on two neighboring strips |
| 256 | Noisy AND rise time <-30 AND calibration pulse peak height <80% normal |
| 1024 | Anything other anomoly |

Table 6.1: Criteria used in determining bad channel flags.

| DB Flag | Description |
|---|---|
| 2 (unbound1) | Mid-strip open |
| 4 (likeunbound1) | Likely mid-strip open |
| 8 (unbound2) | Strip to pitch adapter open |
| 16 (likeunbound2) | Likely strip to pitch adapter open |
| 32 (likeopen) | Likely open, unknown location |
| 64 (noisy) | Noisy |
| 128 (short) | Short between adjacent strips |
| 256 (likepin) | Likely pinhole |
| 1024 (unknown) | Unknown, usually damaged APV channel |

Table 6.2: Description of bad channel flags.

as well. Note that it should not be taken to mean that the total number of faults necessarily increased. There are also 566 which were flagged bad in MRT but good in TOB. There is some imprecision in the fault flagging, particularly in the flagging of opens and noisy channels.

In total there are only 2695 single strip faults, which is a fraction of 0.087 % of all channels in the TOB. Only a very small number of un-repairable component failures occurred in TOB, including one damaged readout fiber, one module which

| Flag | All in TOB | New in TOB |
|---|---|---|
| 2 (unbound1) | 172 | 16 |
| 4 (likeunbound1) | 652 | 226 |
| 8 (unbound2) | 91 | 3 |
| 16 (likeunbound2) | 243 | 41 |
| 32 (likeopen) | 1097 | 213 |
| 64 (noisy) | 280 | 213 |
| 128 (short) | 69 | 18 |
| 256 (likepin) | 0 | 0 |
| 1024 (unknown) | 91 | 16 |
| **Total** | 2695 | 622 |

Table 6.3: Number of strips flagged as bad.

lost connection to the bias voltage, and one module with a short on the bias line. The total percentage of working channels in the TOB is thus 99.86%.

## 6.5 Problems Encountered

### 6.5.1 Detector Bias Problems

#### Shorts on Modules

A few modules developed shorts on the bias line while in the TOB, after having shown normal bias current in all tests prior to rod insertion. The current on such modules rose erratically after first biasing, eventually exceeding power supply limits, at which time a resistance measurement was used to verify a short circuit. A current plot typical of this situation is shown in figure 6.21. In these cases it was found that the module had silver epoxy that shorted the module backplane bias connection to the module frame. The module frame is made of conductive

carbon fiber and is connected to the global ground. The electrical contacts were of a marginal nature (barely touching) which explains why symptoms weren't observed until well after the silver epoxy was applied. During module production it was necessary to make a retrofit to the module backside bias connections of all modules, in some cases using silver epoxy. In the earliest of these retrofits, the silver epoxy was applied by hand, and sometimes in such a way that shorting to the module's carbon fiber frame was possible (see figure 6.22). Later the application of silver epoxy was automated with a small robot, and the amount applied became more uniform. It is suspected that these marginal connections eventually became shorts only after insertion in the TOB because the ambient humidity was higher in the TOB than at any previous stage in rod or module production, hence increasing the migration of the silver epoxy across the module backplane.

This fault only appeared a few times. It was possible to make a full repair in all cases but one, where it was decided that removing the rod with the faulty module would be too difficult. An obvious concern is that there may be other modules remaining in the TOB which had the same messy silver epoxy application and which will become shorts on a longer timescale. After discovering this problem, a check for this messy silver epoxy joint was instituted on all the rods which at that time were yet to be inserted (which was the majority of the rods of the TOB). Because of this check, and because such modules were produced only in a very limited time, the number of undiscovered cases remaining in the TOB is minimal.

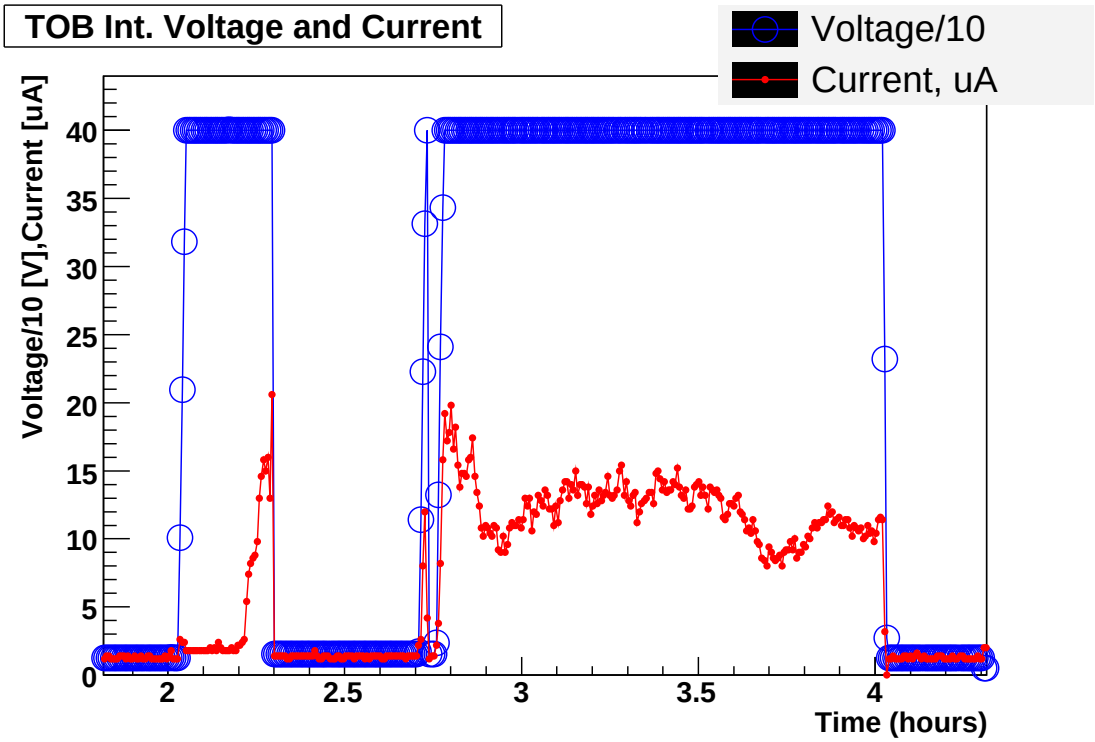


Figure 6.21: Erratic bias current from a module with a silver epoxy short on the backplane.

Figure 6.22: Photograph shows silver epoxy deposited imprecisely, which resulted in an electrical short between the bias line and the module carbon-fiber frame.

#### Shorts on Rod HV Connectors

In three separate incidents shorts between the two high-voltage channels of a rod were observed. In all three cases the affected rod was removed and solder was found on the rod's HV adapter card. The solder was probably deposited accidentally during the soldering of rod cooling pipes to the cooling manifold. The rods were repaired and re-inserted.

#### Open Bias Connections

One case was observed in which a module's bias connection became an electrical open cicuit. Bias was applied to the HV line, but no current was drawn, and the module's noise stayed at the level appropriate for an unbiased sensor. This particular rod was too difficult to remove, because many cables were routed over it, so the cause of the open circuit is not known. It might be a module produced early in the production, which slipped through without having the backside bias retrofit done, and subsequently the backside bias connection failed.

### 6.5.2 Broken Control Wires

Rods that are connected together in a control ring are connected via a short pigtail which comes off the CCU and is bent with a small radius to the adjacent rod. The pigtail uses very narrow gauge wires to minimize material budget. Unfortunately the wires had a tendency to break during the sharp bending necessary

in the final connection. Usually the break occurred in the wires that are used to carry the LVDS communications signal when the redundancy functionality of the ring is used. In a small fraction of cases the lines carrying the primary LVDS signal or the I2C signal between CCU and DOHM were broken. Typically the break occurred where the wires are attached to the connector, or just before the attachment point, underneath the strain-relief gluing. In all cases it was possible to repair the broken connections, either by direct repair while the rod is still inserted in the TOB by adding an external jumper to replace the functionality of the broken wire, or by removing the rod and repairing or replacing the CCU.

This fault was seen rarely, or not at all, in the US rod production sites and CERN rod reception, because it was not necessary to make this sharp bend in the cable in the test stands.

**Control Wire Noise**

It is possible for a wire in a control cable to be partly broken, such that it is still functional for the purposes of passing control signals, but so that it induces noise on the detector modules of that rod itself and its immediate neighbors. The induced noise is similar in shape to the noise pickup discussed in section 3.1, but usually more acute (see figure 6.23), and it cannot be removed by a simple linear fit to the raw signal. In all such cases the control cables wer examined and repaired, either in situ or after removing the effected rod.

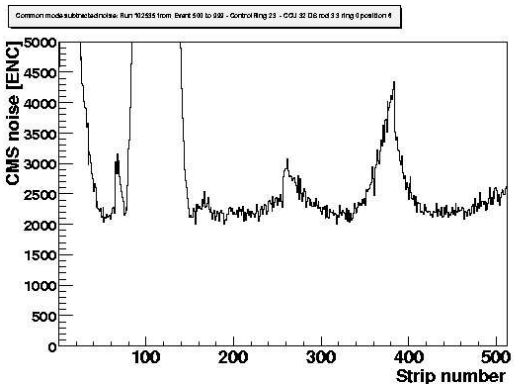


Figure 6.23: Noise caused by broken control wires.

### 6.5.3 Broken Optical Fibers

Tracker optical fibers are sensitive to mishandling. Mating optical fiber connections also requires more care than typically needed for mating electrical connections. During TOB integration a number of fibers and fiber ribbons were broken. Most often the break occurred in the ribbon fan-out where the optical ribbon is split into 12 individual fibers. These breaks cannot be easily repaired since this part of the cable is not easily accessible after the rod has been inserted into the TOB. Initially if an optical ribbon had a broken fiber the rod was removed from the TOB and the ribbon replaced. Later a tool was developed that that allows replacement of the ribbon without removing the rod. One incident occurred in which three optical ribbons were broken. These were repaired in situ: it was possible to splice the broken ribbons back together. Some cases were observed in which the connections of rod optical fibers to ribbons were poorly made. Tooling was developed to re-mate the connection without removing the rod, but this operation

did not have a high success rate, so often the rod was removed. During the later stages of integration a pre-test was implemented to check that the optical ribbon was properly connected to the rod before the rod was inserted. Finally there were cases where the optical fiber on the rod itself was broken, caused by handling problems when the rod was inserted into the TOB. Repairing these cases required that the rod be removed. All broken optical fibers were repaired, except for one in which it was decided that the repair, which required removing the rod, would be too costly in terms of time and risk of damage.

### 6.5.4 APV Synchronization Errors

Two modules were found in the TOB which had synchronization errors from an APV: the pipeline cell addresses did not match for the two APVs on data frames for one laser channel. The rods containing the modules were removed, and we confirmed that the problem was of one APV losing synchronization with respect to the other APVs on the module. After these two occurances an explicit check for this fault was implemented in the rod reception test procedure, and the problem was not seen again in rods inserted in the TOB.

## 6.6 Sector Test

After the +Z end of the TOB was completed, 52 rods were tested together. The number of rods used in the test was limited by the amount of power supplies

and cables that were available. The total number of modules, 408, was about twice the largest number of TOB modules that had been tested together in any previous exercise. Six control rings were chosen, one in each layer. The configuration of rods is shown in figure 6.24. The sector test was a success: full functionality of the tested rods was maintained since the previous test, which was months earlier for some of the rods. All optical fibers were intact, all control ring functionality was intact, and the noise performance was not significantly different (see Figure 6.25).

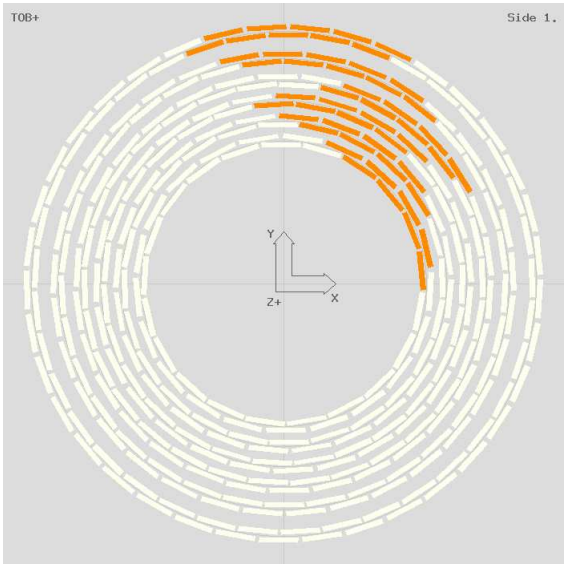


Figure 6.24: Configuration of rods in TOB sector test of October 20th, 2006.

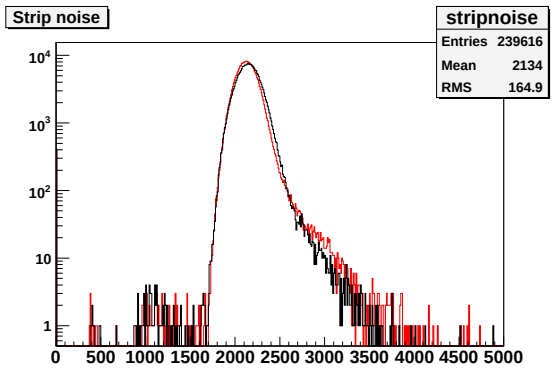


Figure 6.25: Comparision of noise in the sector test of October 20th, 2006 (red) with noise in individual segment validation tests (black).

# Appendix A

| dataset name | events | weight | Phot. ID | full sel. |
|---|---|---|---|---|
| Njet_2j_20_80-alpgen/CMSSW_1_6_7-CSA07 | 4.33e+05 | 5.58e+04 | 0 | 0 |
| Njet_2j_80_140-alpgen/CMSSW_1_6_7-CSA07 | 4.28e+05 | 233 | 9.07e+03 | 0 |
| Njet_2j_140_5600-alpgen/CMSSW_1_6_7-CSA07 | 5.04e+05 | 13.2 | 4.69e+03 | 0 |
| Njet_3j_20_80-alpgen/CMSSW_1_6_7-CSA07 | 6.1e+05 | 2.18e+03 | 0 | 0 |
| Njet_3j_80_140-alpgen/CMSSW_1_6_7-CSA07 | 5.52e+05 | 190 | 4.19e+03 | 0 |
| Njet_3j_140_180-alpgen/CMSSW_1_6_7-CSA07 | | | | |
| Njet_3j_180_5600-alpgen/CMSSW_1_6_7-CSA07 | 2.18e+05 | 13.2 | 1.66e+03 | 0 |
| Njet_4j_20_100-alpgen/CMSSW_1_6_7-CSA07 | 9.74e+05 | 188 | 0 | 0 |
| Njet_4j_100_160-alpgen/CMSSW_1_6_7-CSA07 | 9.97e+05 | 16.3 | 3.18e+03 | 0 |
| Njet_4j_160_200-alpgen/CMSSW_1_6_7-CSA07 | 1.67e+05 | 11.5 | 1.04e+03 | 0 |
| Njet_4j_200_250-alpgen/CMSSW_1_6_7-CSA07 | 7.52e+04 | 9.68 | 445 | 0 |
| Njet_4j_250_400-alpgen/CMSSW_1_6_7-CSA07 | 3.92e+04 | 8.92 | 125 | 0 |
| Njet_4j_400_5600b-alpgen/CMSSW_1_6_7-CSA07 | 2.34e+05 | 0.187 | 19.6 | 0.561 |
| Njet_5j_20_100-alpgen/CMSSW_1_6_7-CSA07 | 1e+06 | 22.2 | 0 | 0 |
| Njet_5j_100_160-alpgen/CMSSW_1_6_7-CSA07 | 3.88e+05 | 11.1 | 701 | 0 |
| Njet_5j_160_200-alpgen/CMSSW_1_6_7-CSA07 | 7.63e+04 | 9.47 | 265 | 0 |
| Njet_5j_200_250-alpgen/CMSSW_1_6_7-CSA07 | 3.49e+04 | 9.61 | 192 | 0 |
| Njet_5j_250_400-alpgen/CMSSW_1_6_7-CSA07 | 2.08e+04 | 9.81 | 88.3 | 0 |
| Njet_5j_400_5600b-alpgen/CMSSW_1_6_7-CSA07 | 2.81e+05 | 0.0844 | 7.77 | 0.675 |
| Njet_6j_20_100-alpgen/CMSSW_1_6_7-CSA07 | 6.77e+05 | 4.72 | 47.2 | 0 |
| Njet_6j_100_180-alpgen/CMSSW_1_6_7-CSA07 | 2.98e+05 | 5.38 | 355 | 0 |
| Njet_6j_180_250-alpgen/CMSSW_1_6_7-CSA07 | 2.36e+05 | 1.27 | 135 | 0 |
| Njet_6j_250_400-alpgen/CMSSW_1_6_7-CSA07 | 9.14e+04 | 1.41 | 67.9 | 0 |
| Njet_6j_400_5600-alpgen/CMSSW_1_6_7-CSA07 | 8.2e+04 | 0.282 | 13 | 1.41 |

Table A.1: The QCD Monte Carlo data samples used, the number of events used per data sample and corresponding weight per $100pb^{-1}$, and the expected number of events per $100pb^{-1}$ after requiring at least one isolated photon with $E_T$above 100 GeV, and after the full event selection.

# Bibliography

[1] The ATLAS Collaboration, G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(S08003), 2008.

[2] The CMS Collaboration, S Chatrchyan et al. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(S08004), 2008.

[3] C. Amsler et al. The Review of Particle Physics, 2008.

[4] Fabjan and Gianotti. Calorimetry for particle physics. *Reviews of Modern Physics*, 75, 2003.

[5] Mendis, Kemeny, and Fossum. CMOS Active Pixel Image Sensor. *IEEE Transactions on Electron Devices*, 41(3), 1994.

[6] Christopher S. Hill. Initial Experience with the CDF layer 00 silicon detector. *Nucl. Instr. and Meth. A*, 511:118–120, 2003.

[7] The CMS Electromagnetic Calorimeter group, P. Adzic et al. Results of the

first performance tests of the CMS electromagnetic calorimeter. *Eur. Phys. J.*, C44(S1.1), 2006.

[8] G. Baiat'ian et al. Design, Performance and Calibration of the CMS Forward Calorimeter Wedges. *CMS Public Note 2006/044*, 2006.

[9] S. Baffioni *et al.* Electron reconstruction in CMS. *Eur. Phys. J. C*, 49(1099), 2007.

[10] CMS Collaboration. CMS Physics Technical Design Report: Volume 1, Detector Performance and Software. *CERN-LHCC-2006-001*, 2006.

[11] M. Sani *et al.* CMS Analysis Note 2008/082: Cut-based electron identification for early search. 2008.

[12] N. Marinelli. CMS Note 2006/006: Track finding and identification of converted photons. 2006.

[13] The CMS Collaboration. CMS Physics Analysis Summary JME-07-002: Plans for Jet Energy Correction at CMS. 2007.

[14] Torbjorn Sjostrand, Stephen Mrenna and Peter Skands. PYTHIA 6.4 physics and manual. *JHEP05*, 026, 2006.

[15] J. Alwall *et al.* MadGraph/MadEvent v4: The New Web Generation. *JHEP*, 0709(028), 2007.

[16] S. Agostinelli et al [GEANT4 Collaboration]. GEANT4 - A Simulation Toolkit. *Nucl. Instr. and Meth. A*, 506:250–303, 2003.

[17] Efe Yazgan for the ECAL/HCAL Collaborations. CMS Conference Report 2008/040: The CMS Barrel Calorimeter Response to Particle Beams from 2 to 350 GeV/c. 2008.

[18] G. L. Bayatian *et al.* [CMS Collaboration]. CMS technical design report, volume II: Physics performance. *J. Phys. G* **34**, 995, 2007.

[19] M. Spiropulu and T. Yetkin. CMS Analysis Note 2006/089:Inclusive missing transverse energy + multijet SUSY search at CMS. 2006.

[20] M. D'Alfonso et al. CMS Physics Analysis Summary SUS-08-002: Data-Driven Estimation of the Invisible Z Background to the SUSY MET Plus Jets Search. 2008.

[21] F. Piccinini R. Pittau M. L. Mangano, M. Moretti and A. D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 0307(001), 2003. [arXiv:hep-ph/0206293].

[22] The CMS Collaboration. CMS Physics Analysis Summary JME-07-001: Missing ET performance in CMS. 2007.

[23] D. Acosta et al., The CDF Collaboration. Direct Photon Cross Section with Conversions at CDF. *Phys. Rev.*, D70(074008), 2004.

[24] A.A. Affolder et al. CMS Note 2004/010: Silicon Tracker Module Assembly at UCSB. 2004.

[25] M. Meschini et al. CMS Internal Note 2004/018: Procedures For CMS Silicon Tracker Module Testing Using ARC Systems. 2004.

[26] E. Focardi, S. Paoletti. CMS Internal Note 2004/055: Technical information on the Low Impedence Cable for the CMS SST tracker. 2004.

[27] L. Mirabito. CMS Internal Note 2004/019: Tracker Data Acquisition user guide Version 1.0. 2004.

[28] F. Drouhin et al. CMS Conference Report 2004/032: The CERN CMS Silicon Strip Tracker Control System. 2004.

[29] N. Bingefors et al. Novel Technique for fast pulse-shaping using a slow amplifier at LHC. *Nucl. Instr. and Meth. A*, 326:112, 1993.