A Measurement of the Isolated Photon Cross Section at CMS with square root of $s = 7$ TeV

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Abraham DeBenedetti

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Roger Rusack

September, 2012

Acknowledgements

I want to thank my wife Emily who supports me through my life and has supported me through graduate school and this dissertation. She is kind, understanding and caring. She has unwaveringly been with me and there for me even when I have asked too much of her. My son Robert is an inspiration to me. He is full of life, energy, humor and love.

I want to thank my advisor Roger who has helped me put my ideas into a coherent form understandable by others. He has enabled me to do interesting projects and learn while in graduate school. The CMS group at Minnesota has been supportive and engaging. I have learned many things through my discussions with Jeremy, Yuichi and Giovanni. Through sharing an office and conversations with my fellow graduate students I have come to understand many puzzles and been confronted with new ones. The QCD Photon group at CMS developed with me the techniques used in this dissertation. People from the QCD Photons group made much of the infrastructure behind this analysis exist. Without the CMS collaboration there would be no detector nor would any of the data be understandable.

My Parents and Family have bared with me through this process and understood when to ask about it. Through my fellow graduate students at Minnesota, I have expanded my horizons beyond the problems presented to me in my work.

I would like to thank Dan Cronin-Hennessy, Mikhail Voloshin, Tom Jones and Jim Kakalios who will attempt to understand the ideas I am putting forth in this document.

Dedication

To Emily and Robert

Abstract

We measure the differential cross section for the inclusive production of isolated prompt photons. This measurement uses data recored by the CMS detector at the LHC in 2010 with proton-proton collisions having a center of mass energy of 7 TeV. The measurement of photon production at the LHC adds to our understanding of the Standard Model and is important for future investigations of new physics. The cross section is compared to the Standard Model with next-to-leading-order perturbative QCD calculations.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this dissertation we measure the cross section of isolated photon production at the Compact Muon Solenoid (CMS) on the Large Hadron Collider (LHC) with proton-proton collisions at a center of mass energy of 7 TeV. We measure both how this cross section changes with the photon energy and with the rapidity of the photon. This measurement tests our understanding of the strong interaction and the structure of the proton, and is important to establish an understanding of photon physics at the LHC. We measure the photon production cross-section for photons with energy between 25 GeV and 300 GeV and out to a rapidity of 2.5.

Protons are composed of three valance quarks, two up quarks and one down quark, but these valence quarks account for only about 1% of the proton's mass. The rest of the proton's mass comes from the energy of the strong force holding together the proton. In the Standard Model, the force carrier of the strong force is the gluon. The energy in these gluons that are constantly being exchanged between the constituents of the proton account for a large amount of the mass of the proton.

To probe the internal structure of the proton the LHC collides two protons together. When protons collide at the high energy provided by the LHC most of the internal components, or partons, of the proton miss one another and the collision results in a spray of particles mostly directed down the two beam lines. Occasionally, two of the partons collide in a hard process. By observing the products of these hard collisions, which have resultant particles transverse to the beam line, we observe the component make-up or structure of the proton.

Experiment	Center of Mass (TeV)	Rapidity Covered $> \|y\|$	Energy Range (GeV)	Year	Citation
CDF	1.96	1.0	50-400	2009	[10]
CDF	1.8	0.9	10-65	2004	[11]
CDF	1.8	0.9	11.5-114.7	2002	[12]
CDF	.63	0.9	9.9-33.6	2002	[12]
CDF	1.8	0.9	12.3-114.7	1994	[13]
CDF	1.8	0.9	14-68	1993	[14]
CDF	1.8	0.9	14-68	1992	[15]
D0	1.96	.9	23-300	2006	[16]
D0	1.8	2.5	10-140	1999	[17]
D0	1.8	2.5	10-125	1996	[18]

Table 1.1: Tevatron results for photon cross section measurements

Because the measurement of photon cross sections is a good probe of the Standard Model it has been studied at previous experiments. Results from the Tevatron are summarized in table 1.1. The photon cross section has also been studied on previous colliders to the Tevatron. When comparing our results to previous results it should be noted that the increased energy of the LHC allows us to probe more of the proton structure at equivalent photon energy and rapidity ranges.

We compare the measurement against the prediction of the Standard Model. Photons produced at the LHC provide a probe of Quantum Chromodynamics(QCD). JetPhox is used to calculate the Standard Model prediction of the cross section. JetPhox calculates the cross section to next to leading order (NLO) in the strong coupling (α_s) and performs the calculation of the isolated photon cross section rather than the full cross section. The initial state which describes the structure of a proton is input into the calculation of the strong process in the form of the Parton Distribution Function (PDF). This measurement tests both our understanding of the strong interaction and our understanding of the structure of a proton.

The CMS detector was designed as a general purpose detector to characterize LHC collisions that occur at its center. It is a superior detector for identifying and characterizing collisions containing photons. The primary subdetector used in the analysis is the Electromagnetic Calorimeter (ECAL). It characterizes the energy and location of

electromagnetic showers left by photons emanating from the collision. To further characterize the products of the collisions, we use CMS's Silicon Tracker and the Hadronic Calorimeter. After the detector characterizes the collision products, the Trigger determines if the data for a particular collision should be stored or discarded.

To determine the photon cross section we start with data which was accepted by the Trigger and stored. We then reduce the number of events in this dataset by eliminating events which do not pass a set of criteria that characterizes events containing an isolated photon. In this new enriched dataset, we determine the purity of events with photons. From this we calculate the isolated photon cross section. Once we have found the cross section of photons we compare our measurement to the theoretically predicted value. The largest experimental uncertainty comes from the measurement of the purity of the photon sample and the largest theoretical uncertainty comes from the arbitrary scale used to switch between the perturbative and nonperturbative parts of the calculation.

Chapter 2

Theoretical Motivation

2.1 The Standard Model

The Standard Model is our current understanding of the fundamental particles and forces that makeup the universe. The integer spin particles within the Standard Model are gauge bosons for the three forces, the photon for the electromagnetic force, the W and Z bosons for the weak force and the gluon for the strong force. These forces interact with the spin-$\frac{1}{2}$ fermions. In the standard model, the two types of fermions are those that interact with the strong force and are said to have color charge and those that do not and are said to be color neutral. The color neutral fermions are called leptons and the color charged fermions are called quarks. Both the quarks and leptons have three generations of isospin pairs. The quarks are the up and the down, the charm and the strange, and the top and the bottom. The leptons are the electron, the muon and the tau, each with its corresponding neutrino.

More formally, the Standard Model is constructed as a quantum field theory that is a product of an SU(2) $\otimes$ U(1) group for the electromagnetic and weak interactions, and an SU(3) group for the strong interaction. The gluons are an octet within the SU(3) group. The other gauge bosons come from a triplet and a singlet within the SU(2) $\otimes$ U(1) group. The neutral member of the triplet mixes with the singlet to produce the photon and the Z boson. The other two members of the triplet are the W^+ and W^- bosons. The left handed fermions behave as an SU(2) doublet and in their right handed state they behave as an SU(2) singlet. The quarks also interact with the

Particle	Symbol	Mass	Charge	Spin
Photon	γ	0	0	1
Z boson	Z	91.1876 ± 0.0021 GeV	0	1
W boson	$W^{\pm}$	80.399 ± 0.023 GeV	± 1	1
Gluon	g	0	0	1
Up Quark	u	$2.5^{+0.6}_{-0.8}$ MeV	$\frac{2}{3}$	$\frac{1}{2}$
Down Quark	d	$5.0^{+0.7}_{-0.9}$ MeV	$-\frac{1}{3}$	$\frac{1}{2}$
Charm Quark	c	$1.29^{+0.05}_{-0.11}$ GeV	$\frac{2}{3}$	$\frac{1}{2}$
Strange Quark	s	100^{+30}_{-20} MeV	$-\frac{1}{3}$	$\frac{1}{2}$
Top Quark	t	$172.9 \pm 0.6 \pm 0.9$ GeV	$\frac{2}{3}$	$\frac{1}{2}$
Bottom Quark	b	$4.19^{+0.18}_{-0.06}$ GeV	$-\frac{1}{3}$	$\frac{1}{2}$
Electron	e	510.998910 ± 0.000013 keV	-1	$\frac{1}{2}$
Muon	μ	105.658367 ± 0.000004 MeV	-1	$\frac{1}{2}$
Tau	τ	1.77682 ± 0.00016 GeV	-1	$\frac{1}{2}$
Electron Neutrino	ν_e	> 0	0	$\frac{1}{2}$
Muon Neutrino	ν_μ	> 0	0	$\frac{1}{2}$
Tau Neutrino	ν_τ	> 0	0	$\frac{1}{2}$

Table 2.1: List of particles in the Standard Model taken from the Particle Data Group [1].

strong force. Each named quark represents an SU(3) triplet of different particles with different strong charge. Since the SU(3) symmetry is unbroken and colored particles are confined inside color neutral hadrons it is impossible to put distinct labels on the three different charges. Nevertheless, we know from evidence, like the existence of the Δ^{++} baryon and the branching ratios of the Z boson, that there must be three copies of each of the quarks. In the Standard model there are also three copies of each of the types of fermions having identical quantum numbers but different masses.

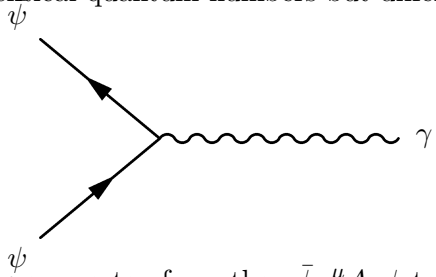

Figure 2.1: Feynman diagram vertex from the $e\bar{\psi}\gamma^{\mu}A_{\mu}\psi$ term of the Lagrangian.

When understanding photon production at a hadron collider we must consider the Quantum Electrodynamic (QED) and the Quantum Chromodynamics (QCD) Lagrangians. Details of the Standard Model can be found in the Particle Data Group review [1] or one can reference texts such as Halzen and Martin [19]. Starting with the QED Lagrangian, we use A_{μ} to represent the photon field, and ψ to represent a fermion field. We define $F_{\mu\nu} \equiv \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$. This gives the Lagrangian:

$$\mathcal{L} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi + e\bar{\psi}\gamma^{\mu}A_{\mu}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.1}$$

where m is the fermion mass, γ^{μ} are the Dirac matrices and e is its electric charge. The first and third terms of the Lagrangian give the free propagators of the fermion and the photon, while the middle term gives the interaction between these two fields. The strength of the coupling is proportional to e the charge of the fermion. This term can be represented by the Feynman diagram shown in Figure 2.1.

In the QCD Lagrangian, we define the interaction between the quark fields q and the strong gauge field or gluons G^{a}_{μ}. We use the T_a to represent the generators of the SU(3) and f_{abc} to define the commutation relations of the group. Similarly to $F_{\mu\nu}$, we define $G^{q}_{\mu\nu} \equiv \partial_{\mu}G^{a}_{\nu} - \partial_{\nu}G^{a}_{\mu} - gf_{abc}G^{b}_{\mu}G^{c}_{\nu}$. The third additional term represents the

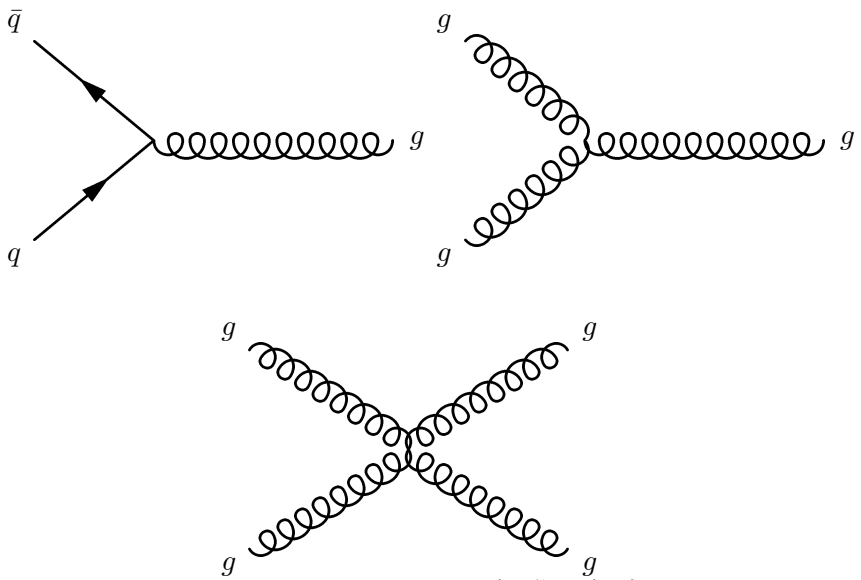

Figure 2.2: Feynman diagram vertex from the $g(\bar{q}\gamma^{\mu}T_a q)G^a_{\mu}$ term and the gluon self interaction from the $G^a_{\mu\nu}G^{\mu\nu}_a$ term of the Lagrangian.

self-interaction of the gluons, since they carry color charge, unlike photons, which are electrically neutral. The QCD Lagrangian is then:

$$\mathcal{L} = \bar{q}(i\gamma^{\mu}\partial_{\mu} - m)q + g(\bar{q}\gamma^{\mu}T_a q)G^a_{\mu} - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a \tag{2.2}$$

where m is the quark mass and g is the strong coupling. As in QED, this Lagrangian has a similar fermion free propagator in the first term, and in the third term, there are terms proportional to "GG" for the gauge boson free propagator. The second term gives an interaction similar to the QED interaction involving the coupling between the fermion the gauge boson. However, within the third term there are terms proportional to "$gGGG$" and "g^2GGGG" which describe the gluon-gluon interactions. The interactions of this Lagrangian are shown in figure 2.2.

In these Lagrangians the strength of the coupling changes with the energy at which one probes the interaction. In QED, there is a polarization of the vacuum which can be understood as pairs of particles being briefly created and destroyed around the primary particle. These particles align such that they screen the charge of the primary particle. This makes the electromagnetic interaction weaker the greater the distance from the primary particle which translates into an increase in the effective coupling with the

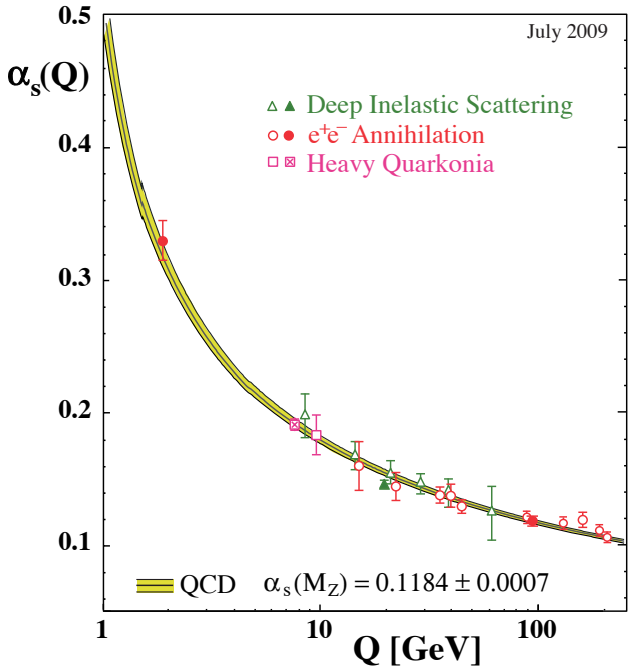

Figure 2.3: This graph from the PDG Review [1] shows the running of the Strong Coupling α_s with the energy it is probed Q.

energy of the probe. On the other hand, in QCD the opposite happens; the coupling becomes weaker as the energy increases. This is known as the running of the strong coupling and is shown in Figure 2.3 through the perturbative domain. The strong force weakens enough at the energy of the hard interaction at the LHC, that it may be treated perturbatively.

We will work within the Standard Model when making theoretical predictions for our observations. Although generally the Standard Model can be used to describe any fundamental process, the variation of the strong coupling constant means that when it is small, precise results can be calculated; but for cases where the strong coupling is large, significant higher order effects prevent perturbative methods from being used to perform a precise calculation. In these circumstances predefined functions are employed. These functions are tuned based on earlier experiments. The calculation of the predicted differential photon cross section is divided into two such parts. First the perturbative with a small strong coupling corresponding to the hard scattering of partons within the protons, and second the nonperturbative part dealing with the structure of the proton and the behavior of the products of the hard interaction.

2.2 Photon Production Process

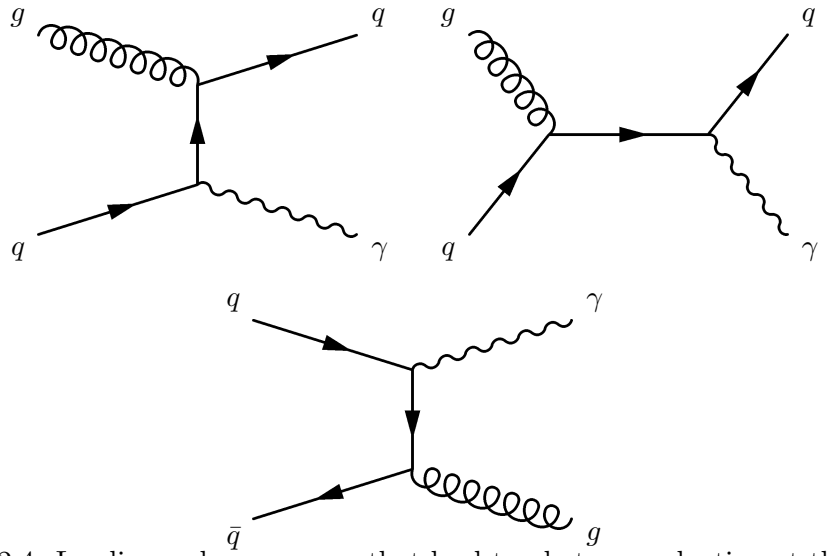

Figure 2.4: Leading order processes that lead to photon production at the LHC

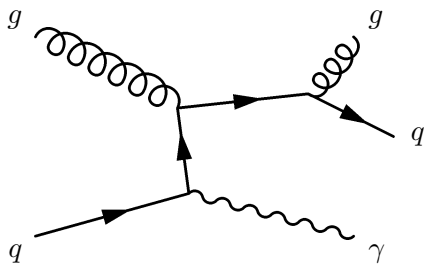

Figure 2.5: Example of next-to-leading order direct photon production.

High transverse energy photon production at the LHC primarily results from a strong process that includes a charged particle, which also couples to a photon. As depicted in figure 2.3 the strong coupling at high energies is small, so the process that results in photons can be expanded convergently in orders of this coupling and thereby calculated perturbatively. The leading order diagrams, representing the approximation matrix elements for direct photon production in proton-proton collisions are shown in Figure 2.4. Nevertheless, this approximation is insufficient to accurately describe data from previous experiments. To bring the prediction into better agreement with the data, one must add the next-to-leading order (NLO) processes. An example of an NLO process is shown in Figure 2.5. Although including the NLO processes increases the predicted rate significantly, the next-to-next-to-leading order (NNLO) terms gives a relatively small change in the rate. The difference between a leading order and an NLO predicted cross section is on the order 100% while the difference between an NLO and an NNLO predicted cross section is closer to 20% [20]. While this is large, when compared to the other theoretical uncertainties in the calculation, it is comparable. That the improvements are less substantial beyond NLO shows that the calculation of the cross section begins to converge at this order. So the research into the theory has centered around improving the NLO predictions and starting to develop techniques for NNLO calculations.

In contrast to the perturbative regime where extensive analytic work is done to make a prediction, in the non-perturbative regime the model is derived from empirical measurements. This non-perturbative part of the theoretical predictions includes the parton distribution functions and the fragmentation functions. The parton distribution functions give the probability of finding a quark or gluon with a particular momentum within a proton, or equivalently, they describe the sharing of the proton's momentum

among the quarks and gluons. The fragmentation functions describe the probability of producing different particles, including photons, after the partons' interaction. The parton distribution functions and the fragmentation functions are discussed in detail in section 2.4.

When calculating the photon production cross section in proton-proton scattering there are two sources of photons, those from the hard interaction and those produced by the fragmentation of outgoing particles. These are both described within the Standard Model, but because of the large coupling in the strong force at low energies when making a prediction one must divide these sources of photons into two seperate categories.

2.3 Details of the Calculation

The calculation of the photon cross section in the Standard Model has been extensively studied theoretically [21–28]. The current state-of-the-art is a computation called JetPhox, which is a calculation to NLO [22, 23]. It has the advantage that it allows the calculation to be matched to the experimental conditions so that direct comparison to the experiment can be performed. In order to make this comparison the authors write a program to perform the calculation.

The basic input to this program is the Standard Model theory and experimental conditions of the measurement, for example the beam conditions, degree of isolation of the photons and angular and transverse energy regions of the cross section to be measured. Then the calculation is performed to NLO using as input external PDFs and fragmentation functions to account for the non-perturbative behavior. In this section, we summarize Cantani et al. in their description of how the calculation is made within JetPhox as described in [22].

To start with we define a coordinate system where the z axis is along the beam line and ϕ around the beam line, and we define an angular variable, pseudorapidity, $\eta \equiv -\ln(\tan\frac{\theta}{2})$, and rapidity, y, a kinematic variable of a particle. For a massless particle, like the photon, rapidity and pseudorapidity are identical.

We then use this coordinate system to define isolation that is a criterion placed on photons to facilitate the comparison with theory. This is important because in proton-proton collisions photons are produced in either directly or indirectly and isolation helps

in the differentiation. Indirect production of a photon can occur by means of a particle produced in collision decaying with a photon as one of the decay products as described by the fragmentation functions. Consequently, a requirement on the photon's isolation preferentially selects direct or prompt photons over photons produced in fragmentation. Additionally, experimentally we are better able to differentiate and measure spatially isolated particles in our detector than photons spatially near other particles. Accordingly, to determine if the theory accurately models the observed phenomenology, we must calculate the isolated photon cross section to compare with our data. We define a spatially isolated photon as one that has less than a maximum value, $E_{T\,max}$, of transverse hadronic energy inside a cone of a size R in (y,ϕ)-space:

$$\sum_a E_T^a \leq E_{T\,max} \quad \text{for particles where} \quad (y_a - y_\gamma)^2 + (\phi_a - \phi_\gamma)^2 \leq R^2 \tag{2.3}$$

Generally, an inclusive cross section calculation may be factorized into a direct and a fragmentation component. Since this cross section excludes non-isolated photons it is not inclusive and may not necessarily be factorized in this way. Instead to determine the cross-section, Cantani et al. calculate the inclusive cross section for a photon with a given value of transverse energy near it. Then they integrate over the required adjacent energies to determine the cross section of non-isolated photons and subtract this cross section from the inclusive cross section. In this way they only ever calculate inclusive cross sections.

In the calculation of the differential cross section $\dfrac{d\sigma}{dp_{\gamma T}dy_\gamma}$, the first step is to define a scale, M_F, that divides the cross section into the two perturbative and non-perturbative components. The perturbative component includes photons that are produced directly from processes above this scale, these are referred to as direct, and the other one, includes photons coming from processes with energy less than this scale. These latter are referred to as fragmentation photons since they are generally associated with jet fragmentation. As there is no physical significance to this scale, the two components cannot be measured separately. So although it is important to the calculation the scale M_F is arbitrary and just an artifact of the calculation. The direct cross section is written as $\hat{\sigma}^\gamma(p_\gamma; \mu, M, M_F)$. Similarly, the direct cross section for the process which leads to

the fragmentation photon is written as $\hat{\sigma}^a(p_a;\mu,M,M_F)$ where a is a quark, an anti-quark, or a gluon which will subsequently fragment into a photon. The fragmentation function, $D_a^\gamma(z;M_F)$, is the probability that the particle a will fragment into a photon with a fraction, z, of its momentum. The momentum of the resulting photon is written p_γ and we can now write the full differential cross section as

$$\frac{d\sigma}{dp_{\gamma T}dy_\gamma} = \left(\sum_a \int_0^1 \frac{dz}{z}\hat{\sigma}^a(p_\gamma/z;\mu,M,M_F)D_a^\gamma(z;M_F)\right) + \hat{\sigma}^\gamma(p_\gamma;\mu,M,M_F). \quad (2.4)$$

Here the direct cross sections, $\hat{\sigma}^\gamma$ and $\hat{\sigma}^a$, are functions of two additional scales μ and M. Although they do not yet appear explicitly, μ is the renormalization scale and defines the value of α_s, and M is similar to M_F except for the initial state parton distribution rather than the final state fragmentation. The choice of the values for the scales is not governed by physical principle, but practically, in order to do the calculation one must fix these to a value. The final uncertainty in the computed values of the cross section is related to variations in the result with the choice of scales.

The calculation is NLO in $\alpha_s(\mu)$, therefore the cross section $\hat{\sigma}$ is divided into $\hat{\sigma}_{Born}$ and $\hat{\sigma}_{HO}$ representing the leading order terms and higher order terms of the cross section in $\alpha_s(\mu)$:

$$\begin{aligned}
\hat{\sigma}^\gamma(p_\gamma;\mu,M,M_F) &= \alpha_s(\mu)\sigma^\gamma_{Born}(p_\gamma;M) &&+(\alpha_s(\mu))^2\sigma^\gamma_{HO}(p_\gamma;\mu,M,M_F),\\
\hat{\sigma}^a(p_a;\mu,M,M_F) &= (\alpha_s(\mu))^2\sigma^a_{Born}(p_a;M) &&+(\alpha_s(\mu))^3\sigma^a_{HO}(p_a;\mu,M,M_F).
\end{aligned} \quad (2.5)$$

The σ_{Born} are the simple $2 \to 2$ tree level cross sections and the σ_{HO} are the inclusive cross sections at higher order. Although the σ^a_{HO} appears to be of third order in $\alpha_s(\mu)$ this is a large enough contribution that it needs to be calculated along with the other three cross sections. The production cross section of particle a is multiplied by a fragmentation function to determine its contribution to the photon cross section.

As an example, we now put the leading order terms into the full cross section.

Explicitly, the leading order terms for the direct and fragmentation contribution are:

$$\begin{aligned} d\sigma^{dir}_{Born}[A+B\to\gamma+jet] &= \frac{1}{8\pi S^2}\sum_{i,j,a}\int \frac{G_{i/A}(x_1,M)}{x_1}\frac{G_{j/B}(x_2,M)}{x_2}\times \\ &\quad |\overline{\mathcal{M}}_B[i+j\to a+\gamma]|^2\times \\ &\quad dy_a\ dy_\gamma\ p_{T\gamma}\ dp_{T\gamma} \\ d\sigma^{frag}_{Born}[A+B\to\gamma+jet+X] &= \frac{1}{8\pi S^2}\sum_{i,j,a,b}\int \frac{G_{i/A}(x_1,M)}{x_1}\frac{G_{j/B}(x_2,M)}{x_2}\times \\ &\quad |\overline{\mathcal{M}}_B[i+j\to a+b]|^2\times \\ &\quad D^\gamma_a(z,M_F)\Theta(z-z_{min})dz\ dy_b\ dy_a\ p_{Ta}\ dp_{Ta} \end{aligned} \tag{2.6}$$

In the direct Born level cross section $G_{i,j/A}(x_{1,2},M)$ are the parton distribution functions, that depend on the momentum fraction, $x_{1,2}$ of the initial parton of species i,j and on the scale M. This process is also proportional to the probability of the matrix element for the Born process given by $|\overline{\mathcal{M}}_B|^2$. The leading order fragmentation photons require a simple Born process in which one of the exiting particles fragments a photon. As before, this fragmentation is added with the fragmentation function $D^\gamma_a(z;M_F)$, where z is the fraction of the transverse momentum of a which goes into the photon and M_F is the scale of the fragmentation function. Here there is also an explicit cutoff Θ for the minimum energy of the fragmented photon.

The higher order tree level cross sections can be built from these using the Altarelli-Parisi splitting functions [29], these describe the modification to the matrix element that an outgoing particle will radiate into an additional particle. This is distinct from fragmenting into that particle, which takes place outside of the perturbative regime.

To incorporate the additional criteria of isolation, the inclusive non-isolated cross section is calculated and subtracted from the total cross section. Similarly to the inclusive cross section, the new inclusive cross section is divided into a part which comes from fragmentation and another one from direct processes

$$\sigma^{is}=\sum_a\int_0^1\frac{dz}{z}\hat{\sigma}^{a,is}\left(\frac{p_\gamma}{z};\frac{z_c}{z},R;\mu,M,M_F\right)D^\gamma_a(z;M_F)+\hat{\sigma}^{\gamma,is}(p_\gamma;z_c,R;\mu,M,M_F) \tag{2.7}$$

The equation for the isolated cross section σ^{is} is similar to the inclusive cross section

in equation 2.4 except this introduces two isolation criteria R and z_c. Where R is the radius in (y,ϕ)-space for the isolation cone, defined earlier, while z_c is the photon's share of the total energy in that isolation cone. In order to perform this calculation, this cross section must be well defined for a isolation criteria, that is it must include no divergences. JetPhox's authors show that if the isolation criteria is unchanged under the following conditions then the corresponding cross section is well behaved and does not diverge [30]. The required criteria are that

1. the momentum of a final state particle is negligible or it is not there at all;
2. two final state particles are emitted in the same direction or they are treated as one;
3. a particle is emitted in the same direction as the photon or its energy is included in the photon's; and
4. a particle is emitted down the beam pipe or its energy was not there in the first place.

These are clearly met with our isolation criteria except for criterion three. By defining an equivalent isolation criteria for a photon with a given energy where the photon's energy is included in the total isolation energy the third criterion can be satisfied as well. Hence the variable z_c and not an isolation energy is used when calculating the inclusive non-isolated cross section.

The calculation is repeated for the next-to-leading order applying similar techniques and criteria. When performing this next-to-leading order calculation in certain parts of the phase space infrared divergences arise. The divergences are normalized then either cancel or are absorbed into the fragmentation function or the parton distribution function where they will cancel because we are measuring a physical process.

2.4 Parton Distribution Function and Fragmentation Functions

The integration performed in JetPhox uses some external functions. These are the Parton Distribution Function and the Fragmentation Functions, which appear as the

$G(x, M)$ and $D_a^\gamma(z, M_F)$ in the integrals. The shape of these functions is determined theoretically with input parameters determined by fitting previous experimental data coming principally from electron-proton and proton-proton colliders. For the calculation of the cross section we use the CT10 [2] parton distribution functions and BFG set II [3] fragmentation functions.

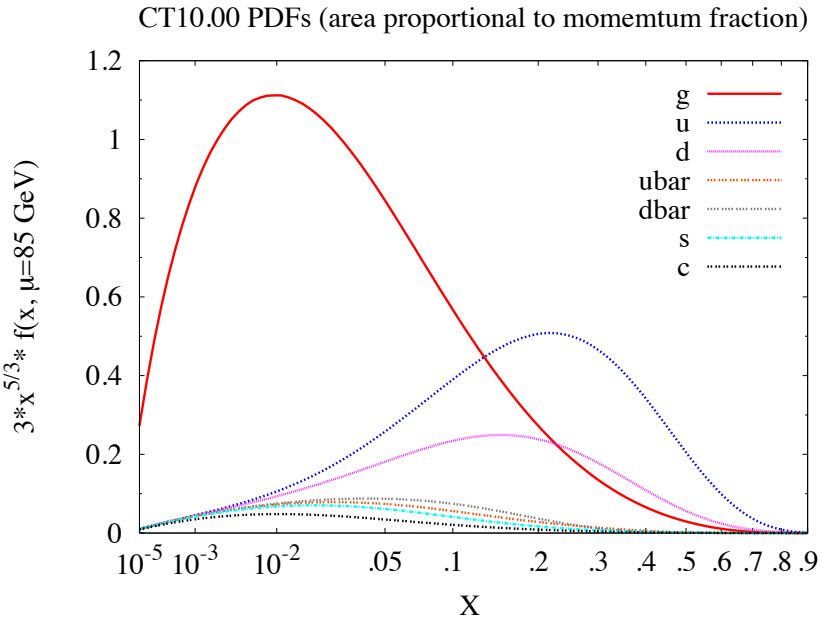

Figure 2.6: The CT10 parton distribution function [2] when $\mu = 85$ GeV. X is the fraction of the proton's momentum in a particular parton and the y axis indicates the probability a parton has share X of the momentum. The graph is scaled such that the share of the proton's momentum in a parton is proportional to its area.

The parton distribution function in figure 2.6 is a description of the initial state of the protons before the interaction. It is a description of the probability that there will be a parton of a particular species with a particular share of the proton's momentum, which is equivalent to being at a particular energy. The high X region is dominated by the three valance quarks, and in figure 2.6 there are twice as many up quarks as down quarks at high x, which is expected for a proton. Carrying a large fraction of the proton's momentum are the gluons binding the proton together. At lower X, they dominate the other species. The first and second diagrams in figure 2.4 require as an

initial state one of each of these partons. The third diagram takes one of these valance quarks and one of the sea antiquarks.

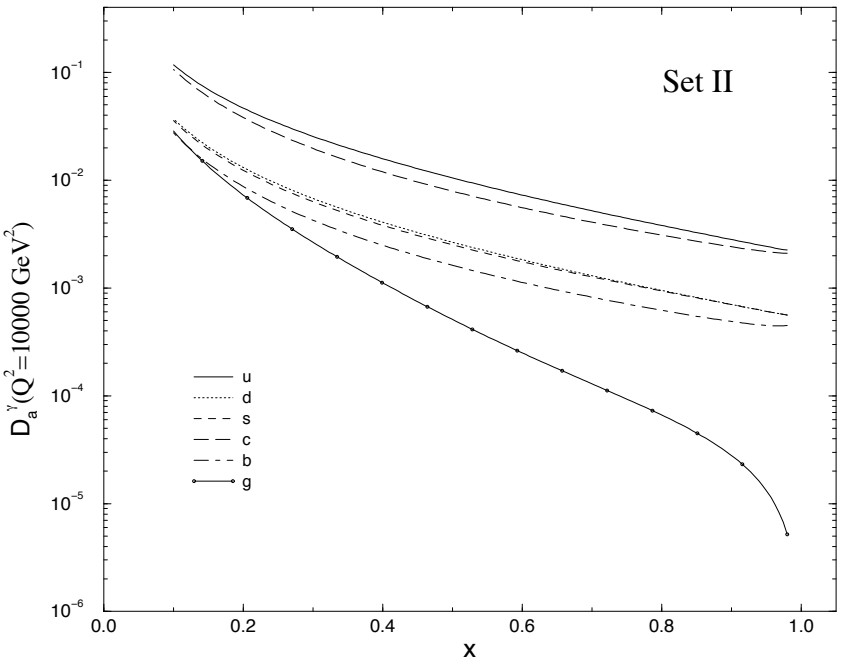

Figure 2.7: D_a^γ is BFG's probability that a particular species of outgoing particle will fragment into a photon with a share X of the initial particle's momentum [3].

Figure 2.7 shows fragmentation functions for particles to fragment into photons. As one would expect, up quarks having twice the absolute charge of down quarks and similar mass are about twice as likely to fragment into photons. Conversely, gluons, being neutral, are the least likely to fragment into photons, which couples to particles in proportion to their charge. Also, a particle is most likely to fragment into a photon with a small fraction of its original momentum. Like the parton distribution functions, the fragmentation functions are not calculated directly from the Standard Model, instead they are derived by fitting previous data.

2.5 Summary

The Standard Model is a well developed and thoroughly tested theory. It describes how isolated photons are produced in proton-proton collisions. At low energies the strong force has a large coupling, one can only directly calculate the direct perturbative stage of the cross section. Although the Standard Model does describe interactions, in this regime, the large coupling means that one cannot use perturbative techniques to perform this calculation. One must rely on data to constrain the parton distribution and fragmentation functions.

With computational techniques described in this chapter, the Standard Model may be used to compute the rate of production of isolated photons. In the next chapter, we describe the apparatus used to measure the photon cross section at the Large Hadron Collider with center of mass energy of 7 TeV.

Chapter 3

The CMS Detector

3.1 Introduction

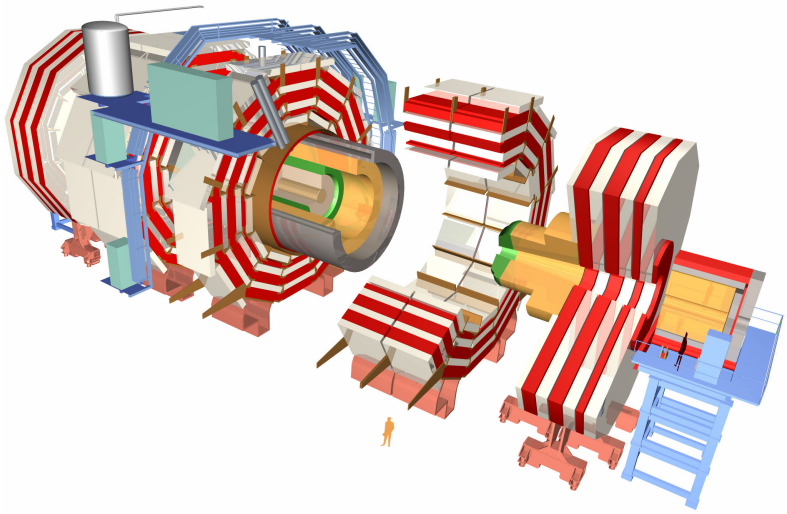

Figure 3.1: A depiction of the CMS detector open and sectioned with a human for scale. It has an onion-like structure. Working outward from the center there is the tracker, Electromagnetic Calorimeter, Hadronic calorimeter, solenoid, and finally muon system and the return yoke for the magnetic solonoid's field.

The Compact Muon Solenoid detector (CMS) [4] is a general purpose detector built at the Large Hadron Collider (LHC) at CERN in Geneva Switzerland. In this chapter,

we explain how high energy particles produced in collisions at the LHC create a signal in the CMS detector and how this signal is selected and recorded. In subsequent chapters we use this data to measure the photon cross section. Performing this measurement requires an understanding of the characterization of photons, quarks, gluons and electrons in the CMS detector.

CMS is designed to measure the products of proton-proton collisions occurring at one of the four interaction points where the two LHC beams collide. The principal goals of the LHC program is to detect the Higgs boson, the final necessary part of the Standard Model, and to find physics beyond the Standard Model, like the source of dark matter. A likely signature for physical phenomena that accounts for dark matter is the production of a massive particle that does not interact with the detector. Accordingly, CMS was designed to measure all the detectable products, thereby allowing by momentum balance us to infer the existence of an unobserved high mass particle in an event. When applied to the measurement of the photon cross section, this means that CMS has a nearly complete acceptance for photons produced at high transverse energy. Previous electroweak measurements indicated we should expect the Higgs boson to be light, but the Large Electron Positron collider (LEP) had excluded a Higgs boson with mass less than 114.4 GeV [31]. In the region just above this exclusion limit, one of the best channels to observe the Higgs boson was by its decay to two photons. For this reason, CMS was designed to measure precisely the energy of a photon so that the invariant mass of the two photon system can be measured precisely. These design requirements also make CMS an ideal detector to measure the photon spectrum.

CMS has a very large solenoid magnet and two types of subdetectors, tracking detectors that record the paths of charged particles in the magnetic field and calorimetric detectors that stop particles and record the energy they deposit. The detector, depicted in Figure 3.1, is structured by its solenoid magnet; inside the magnet moving out from the interaction point is the silicon tracker, then the electromagnetic calorimeter, then the hadronic calorimeter, finally outside the magnet are the muon detector systems. The experiment instruments as much of the solid angle around the interaction point as feasible, to characterize as many of the collision products as possible. Due to the physics emphasis and environmental conditions, like radiation from the LHC, the central region of the detector is more finely segmented and has better resolution. The better

resolution in the central region is an optimization to measure collision products with higher transverse momentum.

The tracking detectors record the curved paths of charge particles as they travel in the magnetic field. The curvature of the track gives the momentum and charge of the particle. The tracker measures the paths of the stable charged particles while minimizing energy loss within it. The silicon tracker has two components: closest to the interaction point is a pixel tracker, and outside that is a strip tracker. The muon detectors are the outermost component of the detector and track the path of charged particles through the return yoke of the magnet.

The calorimeters are designed to stop and record the energy of the outgoing hadrons, photons and electrons. The calorimeters were placed inside the magnet because the walls of the large solenoid have enough material to significantly degrade the calorimeters' ability to measure the electromagnetic and hadronic energy. The calorimetric system has two components the more inner electromagnetic calorimeter (ECAL), just outside the tracker, and the hadronic calorimeter (HCAL), between ECAL and the magnet. ECAL is designed to measure the energy deposited by electrons and photons, which lose their energy by electromagnetic interactions. Hadrons, like charged pions, protons or neutrons, lose their energy by nuclear interaction and consequently require more mass to absorb them.

3.2 The Large Hadron Collider

The Large Hadron Collider(LHC), pictured in Figure 3.2, is a proton-proton collider designed to operate at $\sqrt{s} = 14$ TeV and for this analysis was operating at $\sqrt{s} = 7$ TeV [32]. At each of its four interaction points there is an experiment: two general purpose detectors ATLAS and CMS, and two specialized detectors ALICE for heavy ion collisions and LHCb for bottom quark physics. The energy of the LHC was chosen so that electroweak scale physics, about 100 GeV can be studied in detail.

In its operation protons are injected into the LHC at 450 GeV from the CERN accelerator complex. The LHC accelerates the protons to its operating energy. The beams are focussed such that protons collide in the center of the CMS detector. The design instantaneous luminosity of the LHC will be 10 nb^{-1}/s and can hence deliver

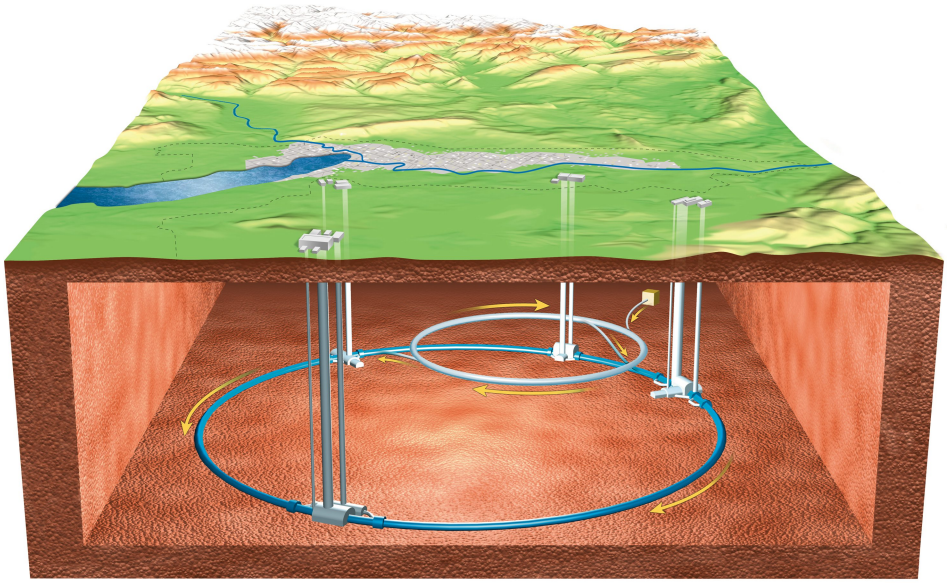

Figure 3.2: The LHC is depicted located 50-100 m under the country side outside of Geneva Switzerland. The large ring represents the tunnel for the LHC. CMS is located in the collision hall on the bottom of the picture. Proceeding clockwise around the LHC are the locations of LHCb, ATLAS, and ALICE. Protons are injected into the LHC by the SPS, the smaller ring. Also show are Lac Leman and Geneva. The Jura mountains can be seen on the bottom right and the Alps are at the top.

100 fb^{-1} in a typical year.

3.3 Particle Signatures in the Detector

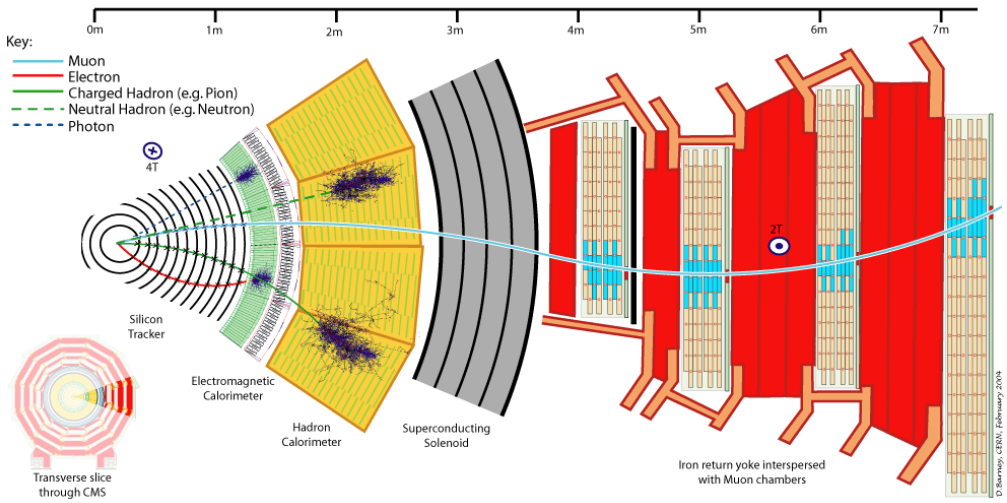

Figure 3.3: Particle interaction inside the CMS detector. Photons deposit energy only in the ECAL, electrons leave a track then deposit their energy in ECAL, and jets, made of many charge and neutral hadrons as well as photons and electrons, to leave track and deposit energy in ECAL and HCAL.

The particles produced in collisions travel out from the interaction point and interact with the surrounding apparatus and these interactions are recorded. Figure 3.4 summarizes the interactions of electrons and photons with matter. The interaction of photons with energies above about 10 MeV is dominated by the pair-production of electrons and positrons. Below this energy they lose energy due to Compton scattering and at even lower energy they are subject to the photoelectric effect. Electrons (and positrons) with energies greater than 7 MeV primarily interact with matter through Bremsstrahlung radiation, where a photon is radiated off reducing the electron's energy. At lower energy the energy loss mechanism is ionization. For the energies of our concern, a photon or an electron cause a chain of pair-production and Bremsstrahlung radiation, called a shower, until all the electrons in the shower are of low enough energy that they ionize the material they pass through. Ideally this shower does not start in the tracker and

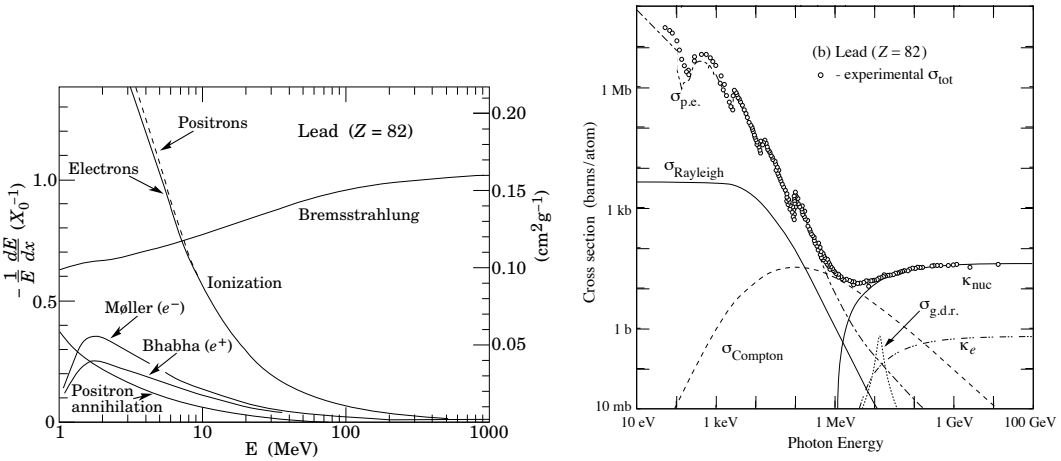

Figure 3.4: These are graphs from the PDG [1] of how electrons and photons interact with matter. The left graph shows how an electron of a given energy loses energy when interacting with Lead. it shows the fraction of the energy lost with each radiation length of material. At hight energy electrons lose most of their energy through Bremsstrahlung radiation and a low energy they lose it through ionization. The right shows the cross section for a photon of a given energy to interact with Lead. At high energy, photons are likely to convert due to interactions with the nucleus and at low energy the photoelectric effect dominates and there is a region around 1 MeV where Compton scattering is dominate.

it is fully contained in ECAL. Nevertheless, due to the material in the tracker, there is a finite probability that the shower will begin in the tracker. The resultant charged particles will be observed in the tracker and the magnetic field will begin to separate the charged particles from the straight path of the neutral particles. So the magnetic field can enlarge the location of the energy deposit in ECAL in the direction perpendicular to the magnetic field. In an additional exception to the ideal case, the shower is not fully contained in ECAL and we observe some leakage from the back into HCAL.

So the typical photon is not observed in the tracker and nearly the full energy is deposited in and measured using ECAL. Similarly, ECAL contains an electron's shower, but the tracker does observe its trajectory as it goes from the interaction point to ECAL. Electrons may be differentiated from photons by the fact that they are registered in the innermost regions of the tracker, before a photon has a significant chance to convert to an electron positron pair.

Unlike electrons and photons, gluons and quarks have color charge and therefore, due to color confinement, do not directly propagate into the detector. They first hadronize into one or more colorless stable particles, which propagate into the detector. This group of particles collectively carry the momentum of the outgoing quark or gluon. The signal from the individual particles may overlap with one another and therefore we think of this collection of particles together as a jet. Particles in a jet may leave several tracks in the tracker and in most cases their energy is not to be entirely contained in ECAL but we reconstruct the full energy of the jet when we combine the energy from ECAL and HCAL.

3.4 The Electromagnetic Calorimeter

The subdetector that is most important to this thesis is the Electromagnetic Calorimeter (ECAL), shown in figure 3.5. It is made with Lead Tungstate ($PbWO_4$) [1] crystals that emit scintillation light with an amplitude proportional to the ionization energy lost in the crystal. It has two components a central barrel region composed of 61200 crystals and two endcap calorimeters with 7324 crystals each; additionally, there is a lead-silicon preshower in front of the endcap calorimeters. The scintillation light in the crystals is collected and measured by avalanche photodiodes (APDs) [5] in the barrel

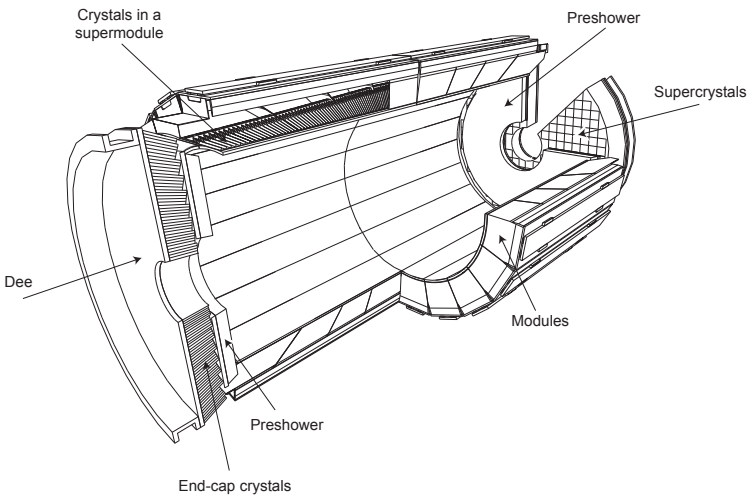

Figure 3.5: A sectioned representation of ECAL from [4]. Depicted are the barrel in the center made of supermodules and the endcap on either end with the preshower in front of it.

and vacuum phototriodes (VPTs) [6] in the endcap calorimeters. The preshower gives additional spatial information about the origin of an electromagnetic shower in the endcap, specifically if there are multiple photons causing the shower.

The barrel of ECAL covers the pseudorapidity of $|\eta| < 1.479$ and a full 360° in ϕ. Each half of the calorimeter is divided in ϕ into 18 identical supermodules each containing 1700 crystals. The endcaps span the pseudorapidity range $1.479 < |\eta| < 3.0$ and are on a plane a distance of 315.4 cm along the beam line from the interaction point. Each endcap is divided in half into two Dee shaped structures which each contain 3662 crystals.

The goal of characterizing electromagnetic showers in CMS makes Lead Tungstate a good choice for ECAL. The high multiplicity of events at the LHC requires that the crystals have a fast signal and contain the shower in a small number of crystals. Due to the radiation environment around the collisions the crystals must be radiation tolerant. Lead Tungstate's high density (8.28 g/cm^3) and therefore low radiation length (.89 cm) means that the calorimeter may be compact while still containing a large fraction of the energy in an electromagnetic shower. Laterally limiting the extent of the shower is

Lead Tungstate's low Molière radius (a parameter characterizing the transverse spread of an electromagnetic shower) of 2.2 cm, thereby confining the shower to a limited number of crystals. These properties mean that in the barrel the crystals need only be 23 cm long in a quasi-projective direction, a depth of 25.8 radiation lengths, and their front faces subtend a cross-section of 0.0174×0.0174 in (η, ϕ)-space or equivalently are 2.2×2.2 mm^2, a lateral dimension comparable to the Molière radius giving optimal spatial resolution. In the endcap calorimeters crystal dimensions are similar(22 cm $\times$ 2.862 cm $\times$ 2.862 cm).

The disadvantages to Lead Tungstate are a relatively low light yield and the variability of its light yield with temperature. The light produced changes by $-2.1\%/^\circ$C at room temperature, so we maintain the subdetector at a temperature of 18°C with precision of about 0.1°C. Because of Lead Tungstate's light yield, the light is collected in the barrel region with APDs, which have a high quantum efficiency and give about 4.5 photoelectrons per MeV of energy deposited. In the endcap the light is collected by VPTs which have lower quantum efficiency than APDs, but this is compensated for by the fact that the VPTs have a larger light collection area and that a particle with the same momentum in a transverse direction to the beam as a particle in the barrel will deposit more energy in the endcap than it would in the barrel. It is necessary to use VPTs because they are more tolerant of the higher radiation environment of the endcap region.

Each crystal in the barrel has two avalanche photodiodes (figure 3.6) whose signals are combined giving 0.5 cm^2 of active light collection area on the crystal end away from the interaction point. APDs are diodes with a high reverse bias voltage, where photons are absorbed at the surface creating electron-hole pairs and the photocurrent is amplified by impact ionization. The resulting quantum efficiency is about 80%. Like the crystals the APD gain has a temperature dependence of $-2.4\%/^\circ$C.

The temperature dependence of the components of ECAL require that the temperature be stabilized and monitored. The temperature is recorded at several points within the electronics of ECAL but most importantly there is a thermister next to the APDs on one in ten of the crystals. This allows the close monitoring of the temperature of both the crystal and the APD. ECAL's temperature is stabilized by a water cooling system which also extracts the heat dissipated by its electronics.

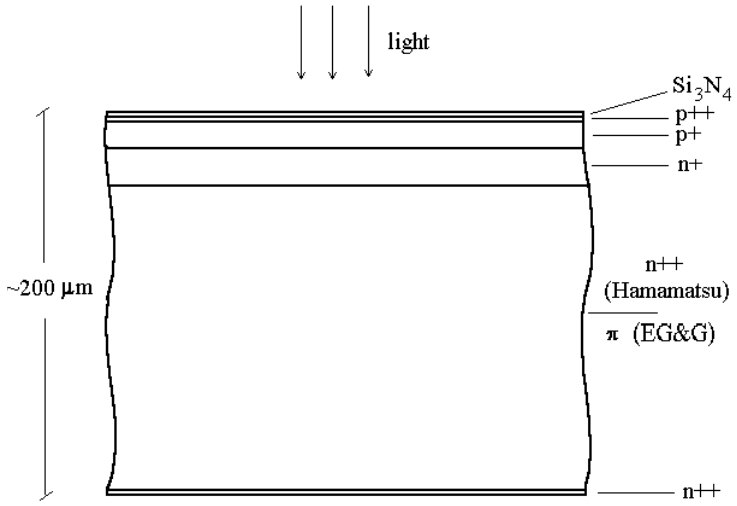

Figure 3.6: Schematic of an APD from the ECAL Technical Design Report [5].

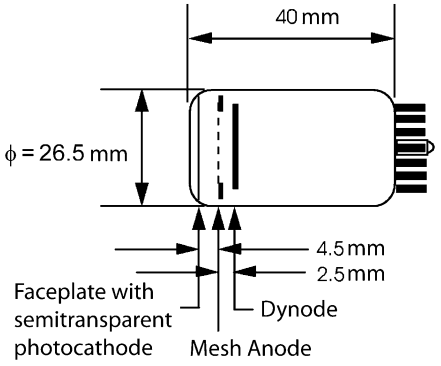

Figure 3.7: Schematic of a VPT from [6]

Due to the higher levels of radiation in the endcap's calorimeters, we use vacuum phototriodes (VPT) [6]. VPTs are photomultiplier tubes (PMT) with a single dynode (figure 3.7), which makes the VPT tolerant to the high magnetic field inside CMS. The quantum efficiency of the VPTs is 22%, much lower than an APD but this is compensated in part by the increased surface area in contact with the crystal.

The electronics used to read out these photodetectors are common throughout ECAL. Starting with the signal from the photodetector, this signal is input into a Multiple Gain Pre-Amplifier (MGPA) [33]. The MGPA outputs three signals with relative gains of 1, 6 and 12 with respect to the original signal. These signals are each sample at 40 MHz and digitized. For each sample the digitization of the signal with the highest unsaturated gain is kept. This gives consistent resolution of the signal between a couple hundred MeV and over 1TeV in a single crystal.

The energy resolution of an electromagnetic calorimeter is usually parameterized as:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2. \tag{3.1}$$

This parameterization disentangles effects coming from different phenomena. The three terms in the parameterization are the stochastic term(S), the noise term(N), and the constant term(C). Each term's contribution to the overall resolution varies with the energy being studied, so the most significant term changes from one enery regime to another. At high energy the constant term dominates while at low energy the noise is of most concern. The constant term comes from contributions that stay constant as a fraction of the overall energy, so their absolute size scales with energy, like uncertainties in the scale factors used to inter-calibrate the individual crystals with one another. The stochastic term comes from sources which involve fluctuations in counting statistics in the signal, like the number of primary photoelectrons collected due to a certain energy deposit. The noise in the electronics appears as a fluctuation in the total energy recorded and therefore the relative effect is negligible when the signal is large.

The performance of the ECAL was measured in a test beam. The test beam resolution of the ECAL was measured to be:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%/\sqrt{\text{GeV}}}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right)^2 + (0.30\%/\text{GeV})^2. \tag{3.2}$$

The test beam environment does not take into account effects present in the LHC environment. Most importantly that in the test beam particles were incident directly on the ECAL and there was not the obstruction of the tracker in front of it. Also in the LHC environment there are contributions to the noise term due to pileup from other collisions and extra energy due to the underlying event. So this should be seen as the performance of the subdetector alone and not as part of CMS. The resolution is measured and monitored *in situ* using standard candles like electrons from the decay of the Z boson and photons from the decay of the π^0 boson. These measurements are consistent with the simulation of the detector [34].

The time of arrival of the signal can be used to differentiate signals from some background sources. Of the initial scintillation signal 80% occurs within the first 25 ns and the rise time of the APD is less than 2 ns and negligible. The current from the photodetector is input into the MGPA where this pulse is stretched to be about a 200 ns [33]. This signal is then sampled every 25 ns with a 12-bit analogue-to-digital converter. From these samples we determine the time of the initial signal to better than 1 ns using the ratios of adjacent samples in the rise and around the peak of the pulse.

To better distinguish photons coming from the decay of neutral pions from photons coming directly from the hard process, we have a preshower detector in front of the ECAL endcap. The preshower is composed of two radiation lengths of lead followed by a sensitive silicon layer followed by a further radiation length of lead with an additional sensitive layer behind it. In this detector incoming photons begin to shower in the lead absorber and the number of showers, and extent of the shower in the silicon layer is determined. The silicon strips in the first layer give position information perpendicular to those in the second layer. Using this information about the first stages of development of the shower, we can determine the structure of a shower more finely than would be possible using only the crystals directly behind it and calculate an energy correction do to the energy lost in the preshower itself.

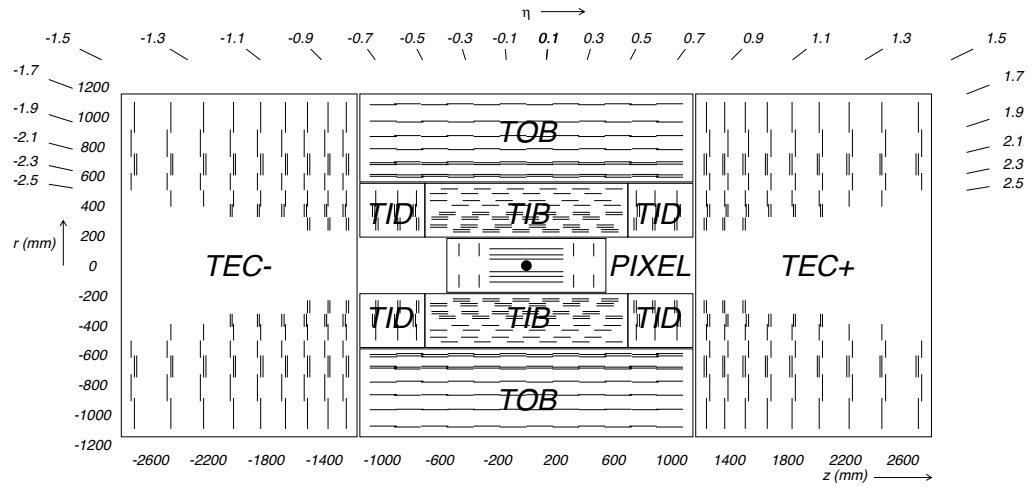

Figure 3.8: Layout of the silicon tracker from [4]. Shown are the inner and outer barrel and disk sections of the silicon strip tracker and at the center the Pixel Tracker.

3.5 The Silicon Tracker

CMS's silicon tracker is situated within ECAL and is used to record the position of charged particles as they pass through it. The tracker is built from two detector technologies an inner silicon pixel detector and an outer silicon strip detector [4, 3.1]. Both use the ionization energy a charged particle deposits as it passes through the silicon to determine the particle's position. A bias voltage is applied and a current is produced proportional to the ionization, which is then recorded. The location of this collected charge indicates the position of the particle as it passes through the silicon. The strip detector and the pixel detector have roughly cylinderical shaped layers around the central region and disk shaped layers covering the ends at high pseudorapidity. As a whole, the tracker detects tracks of charged particles in a pseudorapidity range of $|\eta| < 2.5$.

The pixel detector is the closest detector to the interaction point is highly segmented. This makes it able to resolve interaction vertices, and it is the first detector in which a charged particle leave a signal in CMS. The pixel detector consists of three barrel layers and two disk layers. These are instrumented with 66 million channels each of which give location location information in r, ϕ, and z based on the charge deposited by a particle

going through it.

In this thesis signals from the pixel detector are used to discriminate electrons from photons by determining if a charged particle emerged from the interaction point in the direction of an energy deposit in ECAL. Using the 3 dimensional information given by the pixel tracker, another function of this subdetector is to disentangle the about 20 proton-proton interactions per crossing expected when the LHC reaches its full luminosity. However, an advantage of using early data from the LHC running is that effects from multiple interactions do not influence our results as much as they will later.

The silicon strip tracker is the part of the tracker designed to determine a charged particle's momentum using the magnetic field. As opposed to the pixel detector each layer of the strip detectors provide exact information in 2 dimensions. The momentum of a charged particle moving in a magnetic field is found by measuring the curvature of its track. The strip tracker records the curvature in ϕ as a particle goes out from the interaction point. The strip tracker is also used to reconstruct the momentum of a converted photon from the electron and positron tracks, if they convert early enough that we can observe three or more position measurements.

In this thesis, we use the tracker mainly to reject backgrounds to photons and observe particles near the photons; additionally we must account for the material in the tracker when understanding the observed photon signal in ECAL. We use the most inner pixel tracker to identify tracks pointed directly at our ECAL signal. We also observe tracks reconstructed in the vicinity of the signal and require that our signal is isolated. Because of the structure of the CMS Tracker there is substantial material between the interaction point and ECAL where we detect our signal. This will cause many photons to convert to two electrons, which will cause their signal in ECAL to be spread in ϕ.

A consequence of having a tracking detector in front of the ECAL is that electromagnetic showers often start in the tracker. The tracker has been designed to minimize this problem, but as shown in figure 3.9 in some regions there is a substantial amount of material: at the end of the ECAL barrel $|\eta| > 1$ and in the ECAL endcap there is over a radiation length of material that a particle must traverse before it reaches ECAL. This means that many photons convert before they reach ECAL. This has the effect of spreading their signal in ϕ more than is observed for unconverted photons.

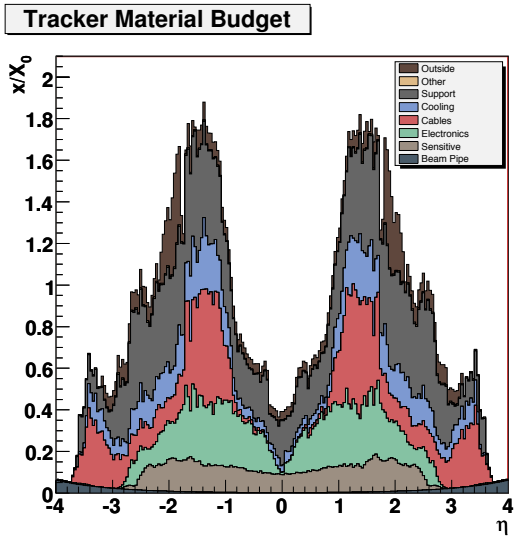

Figure 3.9: The material in the tracker in front of ECAL, in units of radiation length, plotted versus pseudorapidity.

3.6 The Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) is designed to measure energy deposited by hadrons. Often within jets of particles coming from quarks and gluons are hadrons that are stable or stable enough to reach the calorimeteric subdetectors. When they interact they do not shower into small electromagnetic showers contained entirely in ECAL. To measure the energy of these particles one needs much more material for the shower to fully develop than is in ECAL. HCAL is meant to contain and record the shower from hadronic particles in CMS.

HCAL is a sampling calorimeter; going outward from the interaction point HCAL alternates between absorber and active scintillating layers. The result is that the shower develops in the absorber and periodically the scintillator samples the number of charged particles in the shower. The amount of light observed in the scintillator is proportional to the total energy deposited in the subdetector.

Due to the large magnetic field, space constraints from the solenoid, and structural concerns, HCAL uses a brass (70%Cu, 30%Zn) absorber and a stainless steel layer for

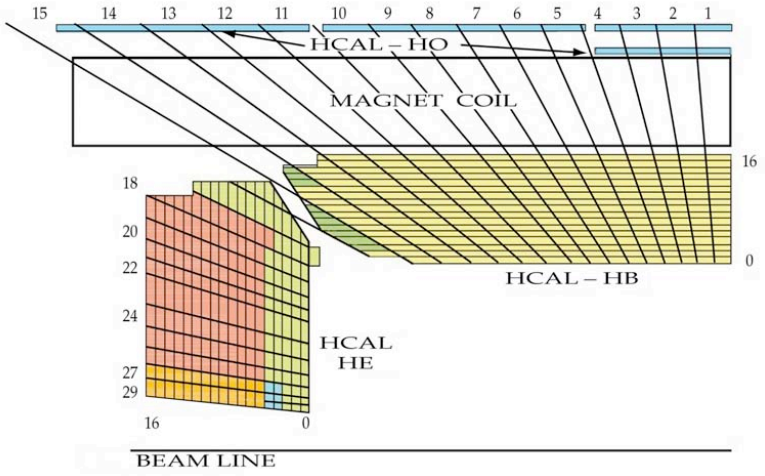

Figure 3.10: Shown is a diagram of a section of the Hadronic Calorimeter. It is divided into a barrel, HB, and endcap, HE, region. Pictured is the tower structure and segmentation of the calorimeter.

structure as the innermost and outermost layers. This absorber gives a total of between 5.92 and 10.6 interaction lengths in the barrel with an additional 1.1 interaction lengths from ECAL and in the endcap about 10 interaction lengths including ECAL.

The sampling of the shower development is done with a plastic scintillator and the signal is collected with wavelength shifting fibers. Except for the thicker innermost and outermost layers, the scintillator is 3.7mm thick. Wavelength shifting fibers run through this scintillator and collect the light and reemit it in the green. Outside of the scintillator tiles the wavelength shifting fibers are coupled by fusing to clear fibers that are used to transport the light to the photodetectors.

These fibers transport the analogue optical signal to hybrid photodiodes (HPD) [35], which are situated on the detector but unlike in ECAL, they do not need to be adjacent to the scintillator. Before the signal is input into the HPD the optical signal from several scintillating regions is combined. This allows an entire tower to be read out by a single HPD channel. Like ECAL, these photodetectors are within the magnetic field; this means that the HPD must be oriented such that the photoelectron trajectory is parallel the magnetic field. Since several optical cables can go to the same photodetector, the HPDs are designed to have several channels per detector.

HCAL is segmented in (η, ϕ)-space to obtain position information about a hadron shower. In the barrel, HCAL is segmented in regions of 0.087×0.087. These regions correspond to, and are projectively behind, the 5×5 sets of crystals in ECAL. In the endcap, the regions are rectangular in (η, ϕ)-space but they gradually take more area in this coordinate system at higher pseudorapidity values. In the barrel the signal from the scintillator at every depth at a given η and ϕ is combined and only the total signal is read out. Similar combining happens in the endcap but there are generally two and at the extreme in η three depth-wise segmentations. Because photons should leave little or no signal in HCAL, in this thesis, we only use quantities corresponding to total energy in HCAL around the position in (η, ϕ)-space of a observed signal in ECAL.

When discriminating photon events from non-photon events in the CMS detector, HCAL is a powerful tool because it is designed to record a signal from jets of particles coming from quarks and gluons. We use this signal behind a deposit of energy in ECAL to determine which deposits are likely to be from photons as opposed to other sources.

3.7 Trigger

In the LHC bunches of protons cross in CMS at a rate of 40 MHz. Because of data storage and computing constraints, CMS is only able to record data at a rate of about 100 Hz. Consequently, CMS uses a two stage trigger to select the events of interest and discards the remainder. It is important that we keep the right information and know which information we are keeping, because if the event is not recorded the information is lost. The competition among multiple physics interests to keep more relevant information for particular physics signals, makes it important that in the trigger we select the minimal number of events to characterize the photon spectrum in CMS.

The first stage of the trigger uses custom-built hardware. This stage, Level-1 (L1), makes a preliminary decision for each LHC crossing whether to keep the event or not. Bandwidth limitations constrain the accept-rate to about 100 kHz of events for further study. In order to make these fast determinations the L1 trigger does not examine the full data in the event. Instead it uses a reduced set of data transmitted from the detector every 25 ns and the remainder of the data is stored on-detector and transmitted only if the event is accepted.

The part of L1 trigger used in this thesis is based on calorimetric data primarily from ECAL. Instead of reading out all the data from ECAL, it is divided into Trigger Towers which then contribute quantities used in the trigger. Since we want to keep (η, ϕ)-space constant, in the barrel, these trigger towers are a 5×5 set of crystals, while in the endcap the number of crystals varies with pseudorapidity. The on-detector electronics sums the energy within each trigger tower. Then it adds this energy to energy in each adjacent trigger tower and determines the maximum of these sums then sends this derived value to the trigger. The highest value in a 4×4 trigger tower region is determined. Then, the four largest of these values are passed to the global trigger where it is determined if this event passes a predetermined criteria, specifically if it is above a threshold and if there less than a relative threshold of energy behind it in HCAL.

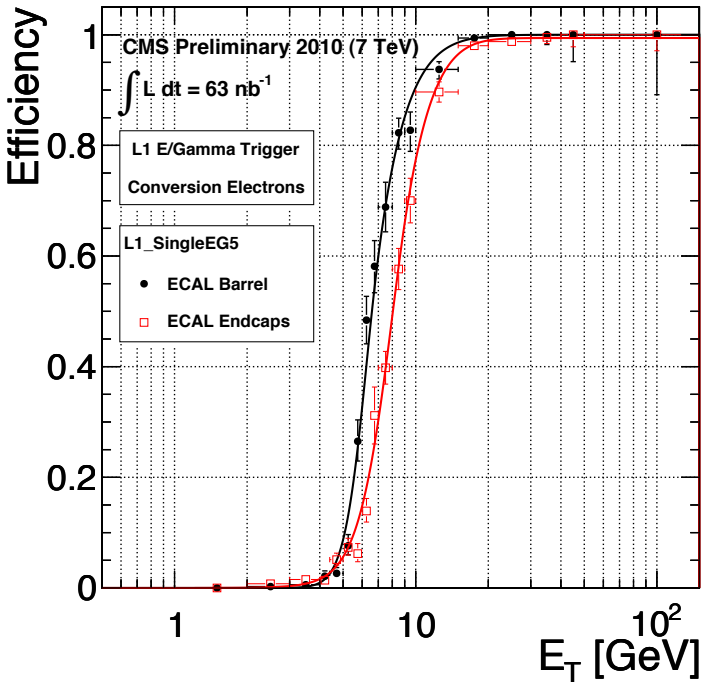

Figure 3.11: This shows the efficiency of the Level-1 trigger for a photon of a given energy, with the trigger having a threshold of 5 GeV. This is taken from minimum bias data and measured using electrons which convert from photons [7].

The turn-on curve for the photon trigger is shown in figure 3.11. This figure shows

the efficiency of applying a 5 GeV threshold to the value passed to the global trigger. This trigger is fully efficient for both the barrel and endcap for a reconstructed energy deposit of about 15 GeV.

If Level-1 accepts the event, all the data from the detector is read out. The data for a single event from different parts of the detector is assembled and passed to the High-Level Trigger (HLT). Unlike the Level-1 trigger, the HLT is a cluster of general purpose computers. An HLT node is passed the complete data for an event and it independently computes if it should accept this event and record it. Having the full data, the node may be programed to analyze quantities that are derived from any part of the detector, but because of processing limitations, in most cases it only does computations on a limited subset of the data. Performing an isolated photon or electron trigger calculation, the node evaluates the data representing the energy in the ECAL crystals, the HCAL and the activity in the region of the tracker in front of the ECAL energy deposit. Using this information, the HLT is programed to work with quantities very similar to those used for offline analysis, and with these quantities, it is programed to accept or reject the event.

The reconstruction of the energy deposit done by the HLT is almost identical to the one done during offline analysis. A prerequisite for an HLT trigger is an "accept" by a specific L1 trigger. The L1 trigger used is one of low enough energy such that it is efficient at the energy of the HLT threshold. Figure 3.12 shows the turn-on curve for the barrel and endcap of an HLT trigger with a threshold of 15 GeV. This trigger requires the L1 to have accepted this event with a threshold of 5 GeV. Because the reconstruction done at the HLT is nearly complete, this trigger becomes efficient very quickly above its threshold. This trigger threshold is used in studies of photons with more then 20 GeV of transverse energy.

3.8 Summary

CMS is designed to record and characterize events resulting from the LHC's proton-proton collision at the center of the detector. As the particles in these events emerge from the interaction point charged particles leave hits in the tracker and bend due to the magnetic field allowing for a determination of its momentum. Outside of the

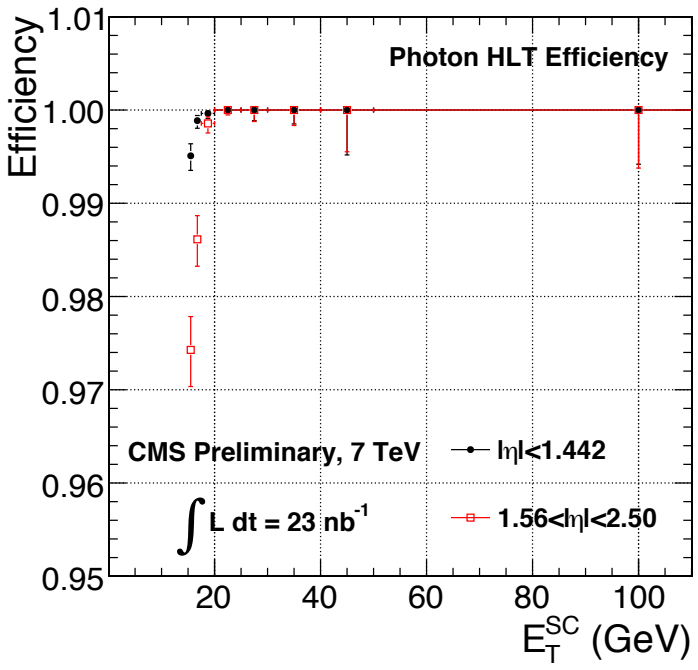

Figure 3.12: This graph shows the efficiency at which an offline super cluster of a given energy will pass the High Level Trigger. HLT The threshold used here is 15 GeV [7].

tracking volume are the calorimeters used to measure hadronic and electromagnetic energy. Using CMS and its Lead Tungstate electromagnetic calorimeter, we are able to excellently characterize photons produced at the LHC and discriminate photons from other signals.

Chapter 4

Analysis

4.1 Introduction

To find the isolated photon cross section we need to measure the number of photons in our detector and the total integrated luminosity for the data set. The comparatively large number of other non-photon energy deposits found in the detector make this measurement challenging. The procedure to measure the number of photons is done in two steps; first, we refine our data set, eliminating most of the events in the data by requiring that the deposits have the attributes of photons; second, we take the remaining events and fit a distribution to the data using the sum of a signal and a background distribution. The result of the fit is an estimate of the number of signal events in the sample. We then estimate how many events were eliminated in the data refining step. We combine this information to measure the total number of signal events in our detector for a known integrated luminosity.

4.2 Data Set

CMS cannot record the data from every bunch crossing, or even from every collision that leaves a signal in the detector. For this analysis, we use data recorded with a trigger designed to record photon signatures in the detector, described in detail in Section 3.7. It is a trigger that primarily depends on signals in the electromagnetic calorimeter. At level-1 it requires a given amount of energy deposited in a set of trigger towers of ECAL.

HLT Threshold	Threshold for Analysis	Effective Integrated Luminosity
$E_T > 20$ GeV	$E_T > 25$ GeV	2.46 pb^{-1}
$E_T > 30$ GeV	$E_T > 35$ GeV	8.23 pb^{-1}
$E_T > 50$ GeV	$E_T > 55$ GeV	17.70 pb^{-1}
$E_T > 70$ GeV	$E_T > 80$ GeV	36.14 pb^{-1}

Table 4.1: We only consider photons above a certain threshold for a particular photon trigger. Depending on when the trigger is running this gives us a different effective luminosity for the different transverse energy bins.

The Higher Level Trigger selection requires that there be more than a certain threshold of energy clustered together in ECAL. When measuring the cross section in a range of transverse energies, we use the lowest physics trigger that is fully efficient for the whole range. The values of the transverse energy thresholds for the analysis and for the trigger are shown in table 4.1. The effective integrated luminosities listed correspond to the luminosity of the collisions when the data were collected. As the instantaneous luminosity of the LHC increased, the threshold of the trigger was increased, so that the data rates did not overload the data acquisition system.

The luminosities in table 4.1 are determined for a given set of triggers using the most forward region of the detector [36]. The average occupancy of the towers in the region $3.5 < |\eta| < 4.2$ for each bunch is recorded without the full detector data being recorded. From the average number of empty towers in this region of the detector we are able to infer the average number of proton-proton collisions per bunch crossing. The constants that go into the calculation are determined with a Van der Meer scan where the beams are incrementally brought together to find the effective overlapping area [37] [38]. Based on this measurement the luminosity is then calculated using the detector to monitor the beam.

4.3 Monte Carlo Simulation

We use Monte Carlo techniques to provide simulated events for this analysis. This simulation is essential to cross check our understanding of particle signatures in CMS, and in some cases simulation is the best way to understand a process in the detector. However whenever possible we base our understanding on real collision data. Although

this simulation represents our understanding of the behavior of particles produced in collisions, we use techniques discussed in chapter 2 and not these simulations to compare finally with the measured cross section.

The simulation is performed in three stages: the collision is simulated and it is determined what the stable products of the collision are; the products are propagated into a model of the detector and it is determined where they deposit energy in the detector; and then the response of the detector to the energy deposited is calculated. For the first stage we use an event generator called PYTHIA 6.4 [39] and for the second we used Geant 4 [40], while the third stage is done with custom software.

The program PYTHIA is used to simulate a set of general QCD events, a set of events that produce photons, and one where Z-bosons are produced via $q\bar{q}$ annihilation. When these events are generated, decay and kinematic restrictions are placed on the particles produced in the event. The cross sections generated using PYTHIA are determined from the leading order processes. These cross sections together with the number of events produced within a sample can be used to generate an effective equivalent integrated luminosity needed to produce these events. After the primary products of the collision are generated, the hadronization and decay of these particles is then simulated until the remaining particles are sufficiently stable to propagate out into the detector. We then used Geant to perform the propagation of these particles through a model of the detector and to calculate the energy losses.

The parameters in PYTHIA have been tuned to reflect the data observed in CMS. The set of parameters that gives the best match to data is designated the "Z2" tune. It is the same as the "Z1" tune described in [41] except it uses the CTEQ 6L [42] parton distribution functions. This tuning was done generally to allow the simulation to match the data and is not specifically tuned for this analysis.

For each simulation set, we use a standard predefined set of processes in PYTHIA. The general QCD events are simulations of $2 \rightarrow 2$ hard processes, where the incoming and outgoing particles are quarks, antiquarks, or gluons. The photon simulation set again simulates $2 \rightarrow 2$ processes, but requires an outgoing photon. The QCD and photon processes are subdivided by the transverse momentum of the outgoing particles before each of these subprocesses is simulated. Finally, in the Z boson sample, the generator is restricted such that Z bosons only decay to electrons with the additional

Process Generated	Transverse Momentum (GeV)	Number of Events	Cross Section (pb)
Photon	0-15	957100	8.42×10^{7}
Photon	15-30	825840	1.72×10^{5}
Photon	30-50	925480	1.67×10^{4}
Photon	50-80	1024608	2.72×10^{3}
Photon	80-120	998215	4.47×10^{2}
Photon	120-170	1023361	8.42×10^{1}
Photon	170-300	1100000	2.26×10^{1}
Photon	300-470	1068904	1.49×10^{0}
Photon	470-800	1083499	1.32×10^{-1}
QCD	0-5	549809	4.84×10^{10}
QCD	5-15	1648096	3.68×10^{10}
QCD	15-30	4154640	8.16×10^{8}
QCD	30-50	2814660	5.31×10^{7}
QCD	50-80	2341546	6.36×10^{6}
QCD	80-120	2408299	7.84×10^{5}
QCD	120-170	3035200	1.15×10^{5}
QCD	170-300	3210080	2.43×10^{4}
QCD	300-470	2149440	1.17×10^{3}
QCD	470-600	1999732	7.02×10^{1}
QCD	600-800	1979055	1.56×10^{1}
Z boson	Invariant Mass > 20	1967607	1.30×10^{3}

Table 4.2: Simulation sets of events

requirement that the intermediate boson has a mass greater than 20 GeV. When events are generated by PYTHIA a cross section is also calculated for the process and a given set of constraints. We use this cross section to compare one generated simulation set to another but not to compare with a measured cross section. The simulation sets generated are listed in table 4.2.

4.4 Differential Cross Section Divisions

Pseudorapidity Divisions
$\|\eta\| < 0.9$
$0.9 < \|\eta\| < 1.4442$
$1.566 < \|\eta\| < 2.1$
$2.1 < \|\eta\| < 2.5$

Transverse Energy Divisions
$25 < E_T < 30$
$30 < E_T < 35$
$35 < E_T < 40$
$40 < E_T < 45$
$45 < E_T < 50$
$50 < E_T < 55$
$55 < E_T < 60$
$60 < E_T < 65$
$65 < E_T < 70$
$70 < E_T < 80$
$80 < E_T < 100$
$100 < E_T < 120$
$120 < E_T < 200$
$200 < E_T < 300$

Table 4.3: We subdivided the photon cross section by pseudorapidity and by transverse energy.

The photon differential cross section is measured in both rapidity and transverse energy and we divide the data that we observe into regions, or bins, in these parameters. Different physical phenomena are probed with photons at different rapidities and energies. However, the number of events we collected in each subdivision, the physical constraints of the sensitive regions of the detector and the resolution of the detector all limit the discrete increments over which we measure the cross section. The binning that we use in this analysis is shown in table 4.3.

4.5 Photon Selection

The first step in the process of determining the isolated photon cross section is to identify events that are likely to have contained an isolated photon by placing requirements on discriminating variables derived from detector measurements. Initially we have a collection of candidate clusters of energy in the ECAL that has passed the HLT requirement. We then reduce the number of candidates due to energy from hadrons by requiring that the candidate cluster is isolated both in the ECAL and in the neighboring detectors. Subsequently, we eliminate candidates caused by electrons, by requiring that the inner layer of the tracker does not detect a charged particle that points towards the candidate. In order that the selection be independent of small fluctuations in the simulation, the threshold for the requirements were made to give a relatively permissive event selection. The permissiveness of the selection we use causes small perturbations in the selection criteria to have little impact on our final sample. Had the requirements tightly constrained the events a small inconsistency in the modeling or reconstruction could lead to a large effect on the final result. This inconsistency would lead to uncertainty, which would limit our ability to probe the cross section distribution.

4.5.1 Discriminating Variables

To differentiate candidates that are photons from those that are background, we construct derived quantities which vary candidate by candidate. These are the discriminating variables. In order for these variables to have some discrimination power, candidates from the signal must on average have different values than the candidates from the background. The variables we use to discriminate signal photons from their QCD background rely on the fact that in general direct photons appear isolated in our detector, and the background candidates are generally associated with other particles in close proximity because they originate from jets. Three discriminating variables are constructed, one from the signal in the ECAL, a second from the HCAL signal and a third from the tracks measured in the tracker.

When constructing a discriminating variable within ECAL we primarily rely on the fact that direct photons are isolated and QCD background is not, but to improve the discrimination power of the variable we consider other attributes of our signal. Photons

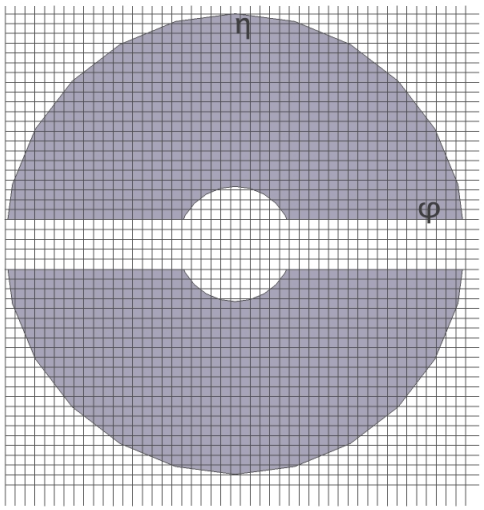

Figure 4.1: The shape in η, ϕ-space of the region in which the energy is summed to construct the ECAL isolation energy overlaid on a grid representing the crystals of ECAL.

that do not convert in the tracker leave a symmetric energy deposition in a region of about ten crystals. However, if the photon has converted in the tracker the magnetic field will cause the energy deposition to broaden in the ϕ direction. We construct the ECAL isolation variable to minimize the contribution of energy from the candidate cluster. Accordingly, we take all the energy found within a cone of radius 0.4 around the center of the candidate cluster and exclude a smaller inner cone with a radius of 0.06, and a strip with a width of 0.087 radians at the same ϕ as the candidate. The resultant shape used to sum this energy is shown if figure 4.1. The discriminating variable ECAL isolation is then defined as the total transverse energy deposited in every crystal within this region.

Figure 4.2 shows the distributions in the ECAL isolation variable for candidates in our data, in the direct photon signal generated by Monte Carlo (squares) and in the background QCD generated by Monte Carlo (triangles). As can be seen, candidates associated with photons are more isolated and have a lower value in the ECAL isolation variable than candidates from background processes. Also plotted is the distribution that is observed in the data (circles).

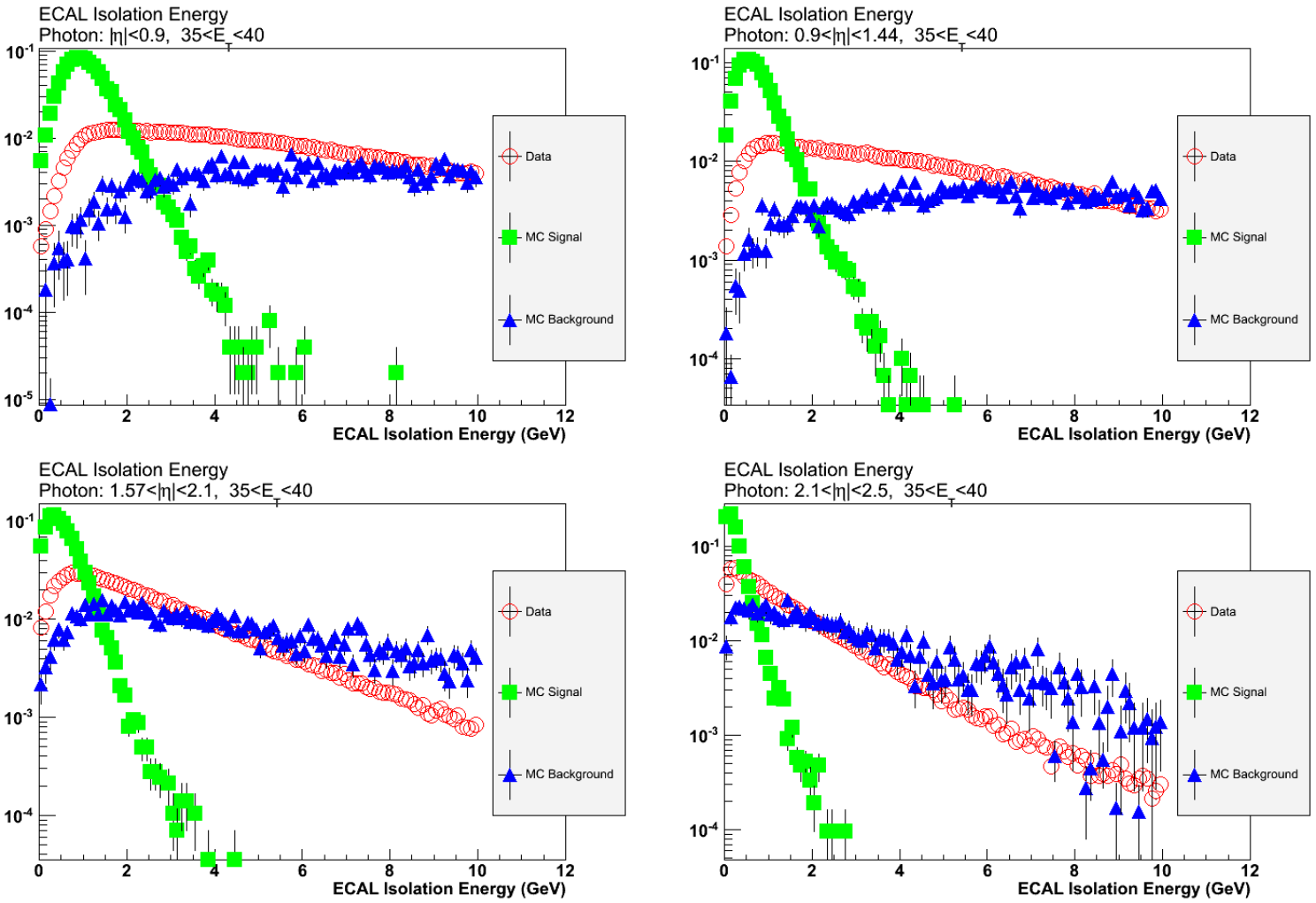

Figure 4.2: Plots of the ECAL isolation for photon candidates with between 35 GeV and 40 GeV of energy, in the pseudorapidity regions $|\eta| < 0.9$(upper left), $0.9 < |\eta| < 1.44$(upper right), $1.57 < |\eta| < 2.1$(lower left), and $2.1 < |\eta| < 2.5$(lower right). The plots show normalized distributions from the previously defined event sets for observed data(circles, 4.2), Monte Carlo photons(squares, 4.3) and Monte Carlo QCD background(triangles, 4.3).

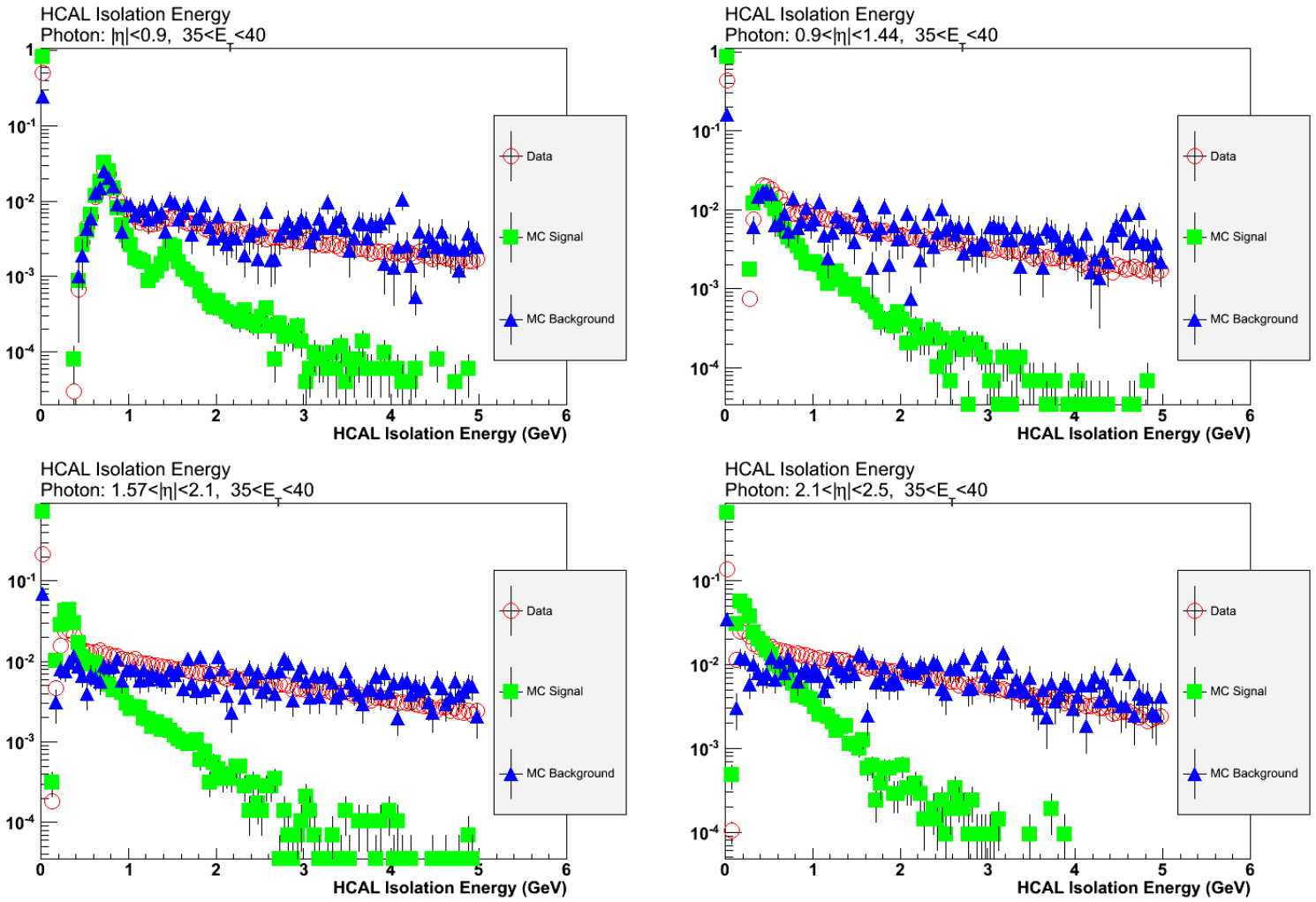

Figure 4.3: Plots of the HCAL isolation for photon candidates with between 35 GeV and 40 GeV of energy, in the pseudorapidity regions $|\eta| < 0.9$(upper left), $0.9 < |\eta| < 1.44$(upper right), $1.57 < |\eta| < 2.1$(lower left), and $2.1 < |\eta| < 2.5$(lower right). The plots show normalized distributions from the previously defined event sets for observed data(circles, 4.2), Monte Carlo photons(squares, 4.3) and Monte Carlo QCD background(triangles, 4.3). These plots are restricted to events with low ECAL isolation.

We define the discriminating variable, HCAL isolation, in a simpler fashion, as photons leave little energy in the HCAL. The value of the HCAL isolation variable is determined by summing the transverse energy in all of the HCAL towers within a cone of 0.4 around the candidate in (η, ϕ)-space. Figure 4.3 shows the distribution of this variable for events from photons (squares) and from all background processes (triangles) each generated by Monte Carlo, and the observed distribution in the data(circles).

In the top left plot of figure 4.3, which is for events from the lowest η division, there is a double peak in the Monte Carlo photon distribution. This is related to the high population seen in the zero bin of the distributions. This happens because, when there is less than 0.6 GeV of energy in a tower, the HCAL is not recorded and the energy is assumed to be zero. In the lowest η region, there are three peaks that correspond each to zero, one, and two towers above threshold. This effect is not present in the higher pseudorapidity regions because the threshold is based on energy and the HCAL isolation is summed in transverse energy, so the threshold effectively becomes smaller.

We also construct a discriminating variable using information from the tracker. We construct the tracker isolation variable to be sensitive to candidate clusters produced in correlation with nearby charged particles, since this is an attribute of QCD background and not of direct photons. For a given candidate, the track isolation variable is consequently defined as the sum of the momentum of the tracks within a cone of 0.4 in (η, ϕ)-space around the candidate cluster, with an exclusion region of 0.04 at the center to prevent photons that convert in the tracker material from contributing to this variable. Also to exclude the tracks from converted photons, tracks that are aligned in η to within a value of 0.015 are excluded. Distributions for this variable are shown in figure 4.4. Again like the HCAL isolation variable, there is a large population of signal events in the zero bin corresponding to the case where there are no tracks in the cone.

Finally, we construct a discriminating variable, associated pixel hits, based on the signal in the pixel detector that is specifically designed to exclude electrons or an atypically showering pion. This particle may appear photon-like in all the other discriminating variables because we specifically ignored tracks pointing directly at the candidate in the tracker isolation variable. If there are hits on the two innermost layers of the pixel detector that are consistent with a track from the vertex of the interaction to the candidate cluster the variable is "true" otherwise it is "false". Distributions of the

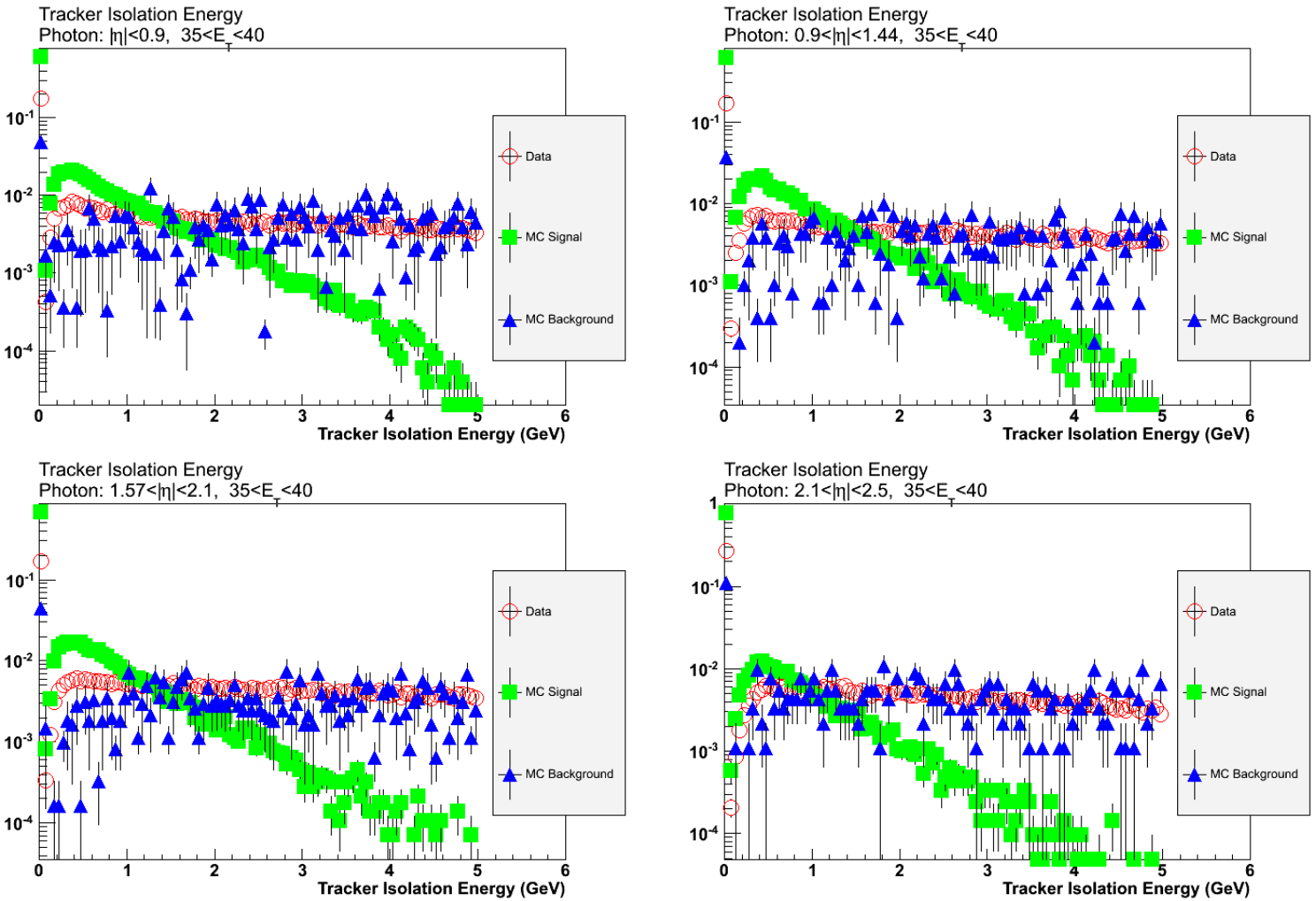

Figure 4.4: Plots of the track isolation for photon canidates with between 35 GeV and 40 GeV of energy, in the pseudorapidity regions $|\eta| < 0.9$(upper left), $0.9 < |\eta| < 1.44$(upper right), $1.57 < |\eta| < 2.1$(lower left), and $2.1 < |\eta| < 2.5$(lower right). The plots show normalized distributions from the previously defined event sets for observed data(circles, 4.2), Monte Carlo photons(squares, 4.3) and Monte Carlo QCD background(triangles, 4.3). These plots are restricted to events with low ECAL isolation and low HCAL isolation.

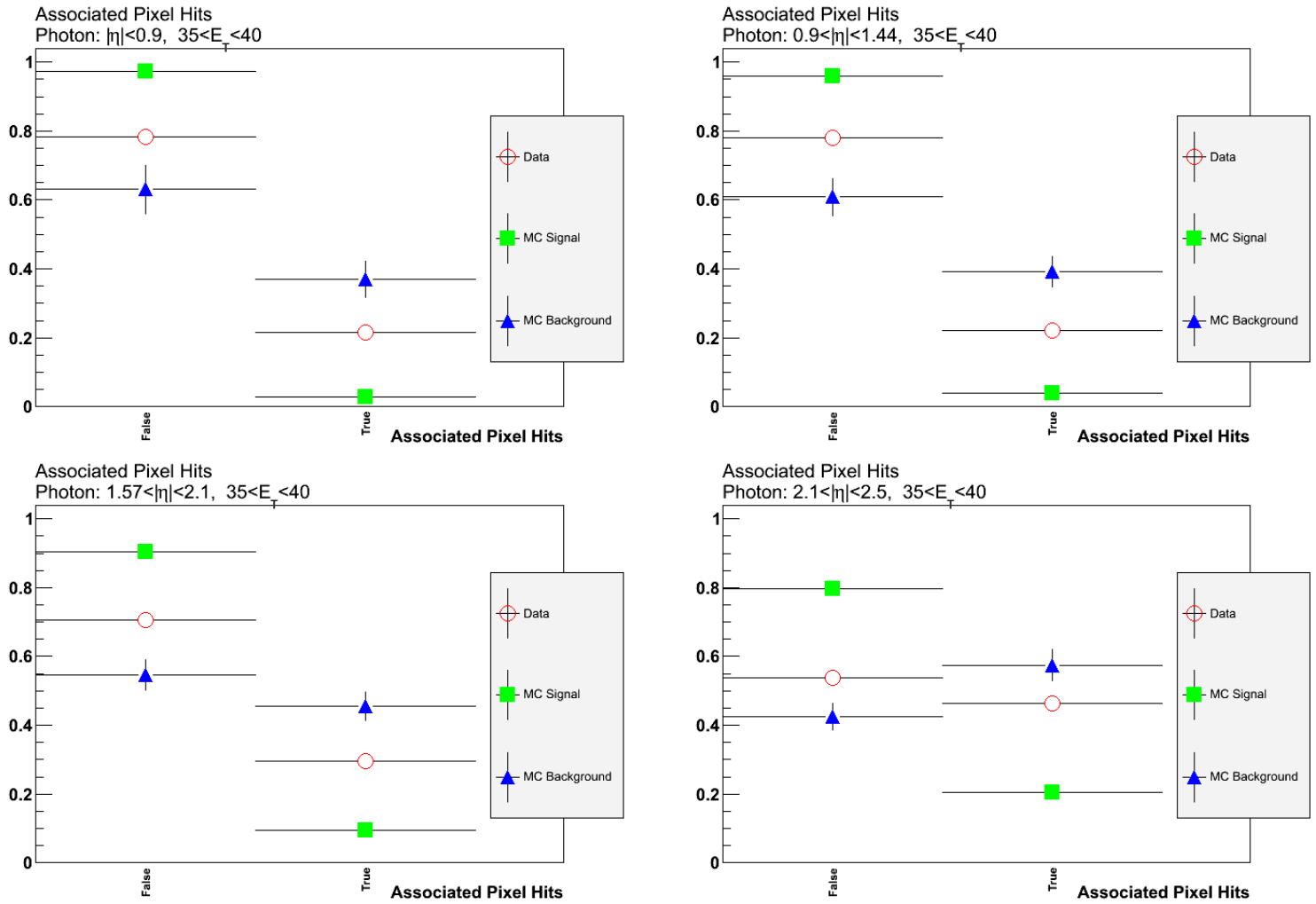

Figure 4.5: Plots of the associated pixel hits for photon candidates with between 35 GeV and 40 GeV of energy, in the pseudorapidity regions $|\eta| < 0.9$(upper left), $0.9 < |\eta| < 1.44$(upper right), $1.57 < |\eta| < 2.1$(lower left), and $2.1 < |\eta| < 2.5$(lower right). These plots are restricted to events with low ECAL, HCAL and Track Isolation.

associated pixel hits discriminating variable are in figure 4.5. Although this variable is intended to discriminate between photons and electrons it also has some discrimination power when applied to the QCD background.

4.5.2 Selection

The selection criteria based on these discriminating variables applied to the candidates are shown in table 4.4. We keep in our sample for further study only candidates that meet these criteria. It is observed that as the energy of the measured photon increases there is a slight increase in the energy of the isolation variables. This effect is largest in the ECAL, but is in general small when compared either to the photon's energy or the energy of the overall threshold. To compensate for this, the value of the thresholds increases with the transverse energy of the candidate in order to maintain a constant efficiency for photons over a range of energies. The values of the thresholds are chosen with the target of making the overall selection 90% efficient.

Variable	Selection Criteria
ECAL Isolation	ECAL Isolation Energy $< 4.2 + 0.003 * E_T$
HCAL Isolation	HCAL Isolation Energy $< 2.2 + 0.001 * E_T$
Track Isolation	Momentum of near tracks $< 2.0 + 0.001 * E_T$
Associated Pixel Hits	"false"

Table 4.4: Selection placed on the data.

4.5.3 Efficiencies

In making a selection to reduce the background contamination in the sample, there is some loss of the photon-signal. To correct for this loss,we need to estimate the efficiency of each of the selections applied to the data. The efficiency is defined as the ratio of the number of photon candidates in our final sample to the number of true photon candidates that were in the original sample. We make this estimation in two steps: first using a set of events generated with the Monte Carlo we estimate the efficiency of the selection, then we correct the Monte Carlo based estimate using a comparison between the efficiency for electrons coming from a decay of a Z boson in events generated by

Monte Carlo and the efficiency for electrons from Z boson decay in data.

4.5.3.1 Monte Carlo Based Efficiencies

To estimate the efficiency for selecting candidate clusters that come from a photon in the data, we use photons generated by the Monte Carlo, apply the selection to these candidates and record the fraction that pass the selection. In figure 4.6, we show the cumulative efficiencies for each successive selection in our pseudorapidity and transverse energy bins. The efficiency of the selection is relatively stable with transverse energy and has a value of about 90% in the barrel. In the endcap, the associated pixel hits selection becomes less efficient as the pseudorapidity increases, due to the geometry, the amount of material in the tracker between the ECAL and the interaction point, and an increase in the background of random hits in the pixel tracker.

4.5.3.2 Tag-and-Probe Efficiencies

Next we correct for differences between the efficiencies found using Monte Carlo events and the true efficiency of applying our selection procedure to the data. To do this we first repeat the measurement of the efficiency performed in section 4.5.3.1 with Monte Carlo events containing electrons from Z boson decay. Then we find the efficiency of the selection on electrons using a "tag-and-probe" technique applied to events in the data. We then use the ratio of the efficiency determined from the simulated electrons and the efficiency found using the "tag-and-probe" technique to calculate a scale factor. Then we correct the efficiencies in 4.5.3.1 by this factor to determine a final estimate of the efficiencies. These corrected efficiencies are then used in determining the cross section.

The principle of the "tag-and-probe" technique is first to obtain an unbiased sample of electrons by using a high purity sample of Z bosons that decay to electrons, and then to use these electrons to measure the efficiencies of the selection. To begin with, we identify electrons that pass tight electron requirements that have been tuned to accept 80% of $W \to e\nu$ events [7]. Then we search for a second cluster of energy in the ECAL in the same event that reconstructs with the first electron to give an invariant mass between 60 GeV and 120 GeV. Clusters found in this way predominately come from the electron-positron decay of a Z boson, and are an unbiased sample of electrons.

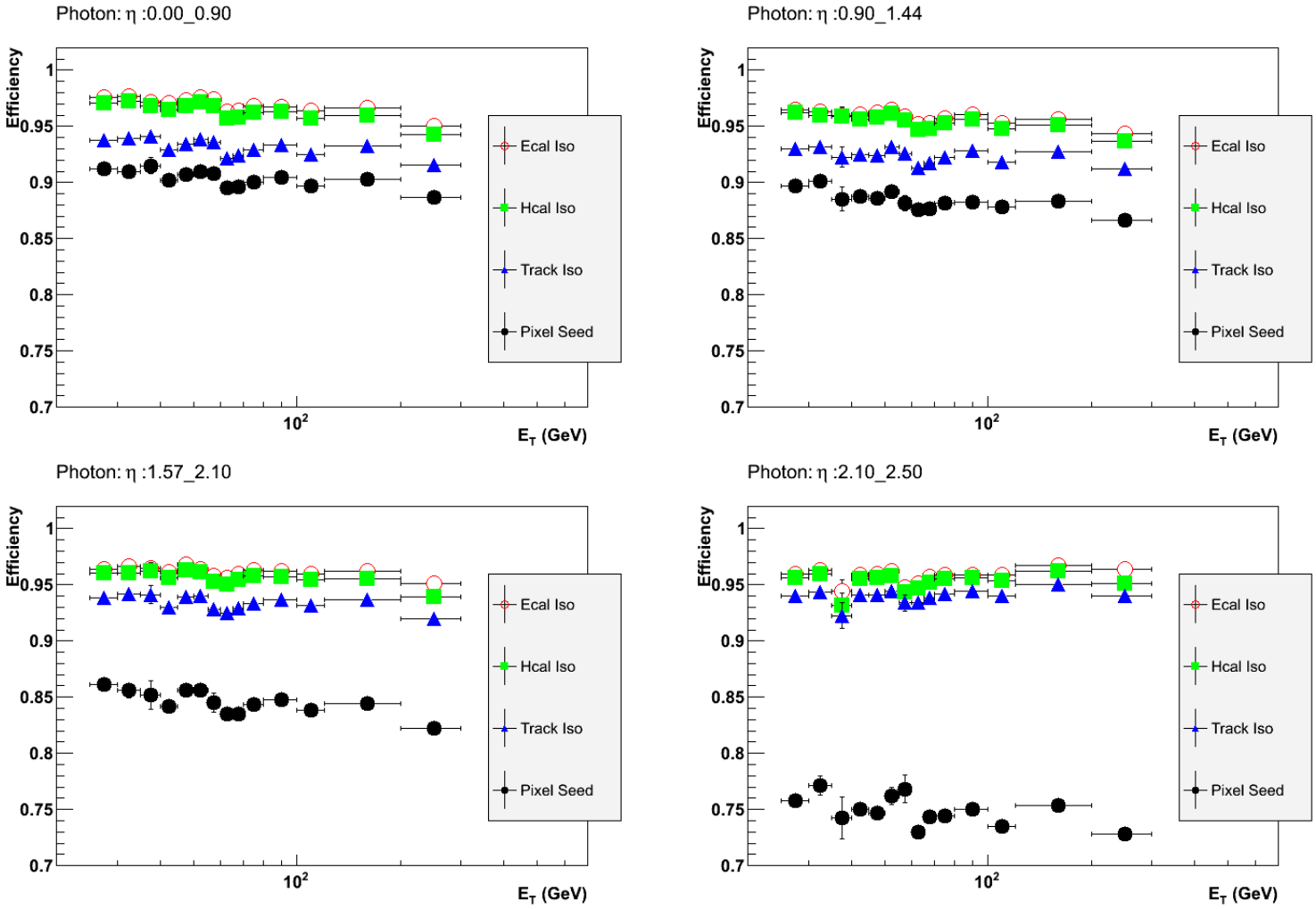

Figure 4.6: Monte Carlo Based efficiency for each successive selection in each of the pseudorapidity regions

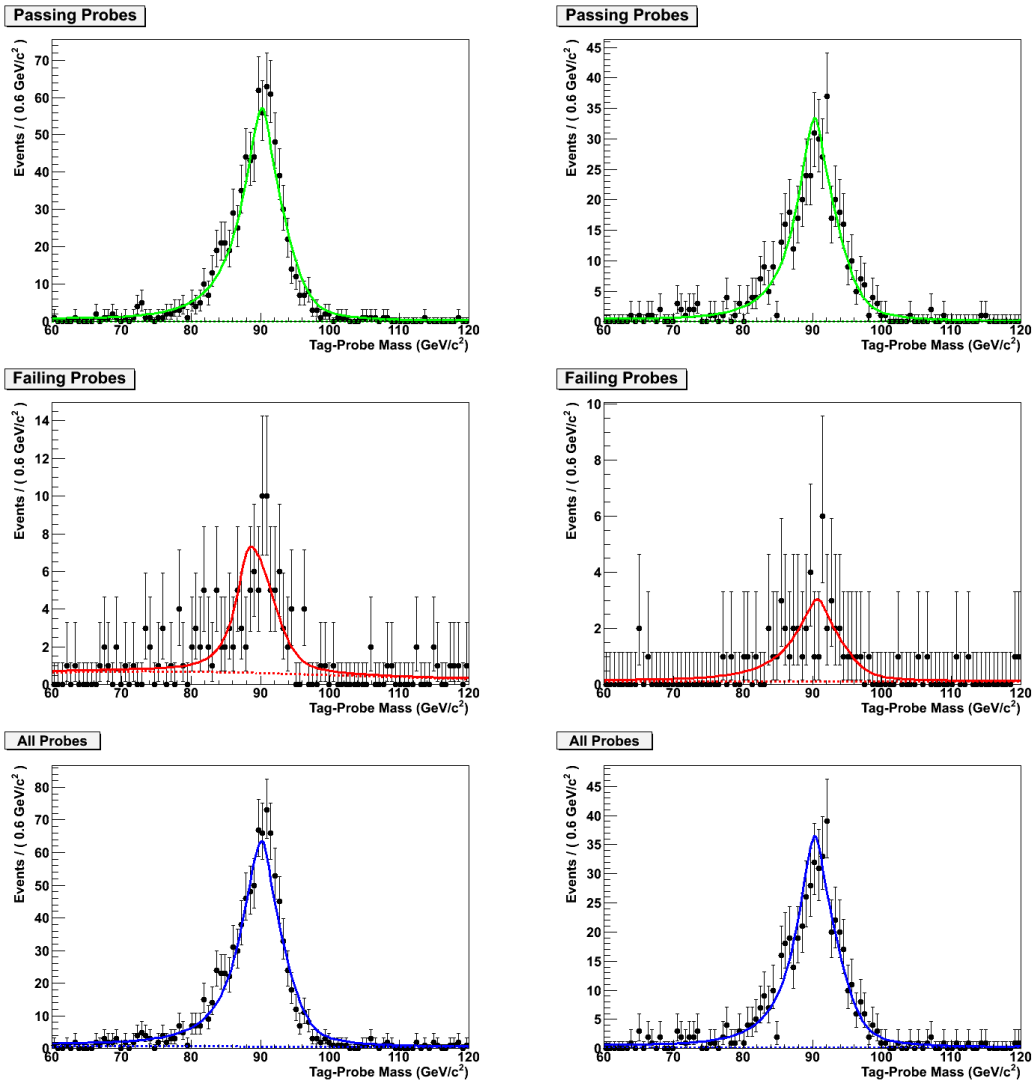

Figure 4.7: Examples of Simultaneous fit of passing and failing the selection of electrons close to the Z boson peak. Both are for the transverse energy bin between 35 GeV and 40 GeV. The Left fit is for the pseudorapidity region between .9 and 1.4442 and the right is for the region between 1.566 and 2.1.

We take these clusters and use them as photon candidates, applying our discriminating variables to them. We exclude the associated pixel hits requirement because this is explicitly included to remove electrons from the sample. Nevertheless, we do include the tracker isolation variable because it does not include momentum from tracks pointing directly at the candidate. We divide these candidates into two categories, those that pass our selection criteria and those that fail. Then we simultaneously fit invariant mass distributions of the two categories of candidates with a function that allows for both true Z boson decay events and background events. Examples of these fits are found in figure 4.7. The efficiency of the electron selection is the ratio of the measured candidates coming from Z boson decay that pass the selection to the total number of candidates coming from Z boson decay. To correct the efficiencies from section 4.5.3.1, we scale them by the ratio of the "tag-and-probe" efficiency to the efficiency found applying the procedure from section 4.5.3.1 to simulated electrons. The corrected photon efficiencies are given in table 4.5.

	$\lvert\eta\rvert < 0.9$	$0.9 < \lvert\eta\rvert < 1.4442$	$1.566 < \lvert\eta\rvert < 2.1$	$2.1 < \lvert\eta\rvert < 2.5$
$25 < E_T < 30$	$0.869\pm^{0.005}_{0.005}$	$0.851\pm^{0.006}_{0.007}$	$0.862\pm^{0.006}_{0.006}$	$0.727\pm^{0.013}_{0.010}$
$30 < E_T < 35$	$0.867\pm^{0.006}_{0.006}$	$0.855\pm^{0.008}_{0.008}$	$0.856\pm^{0.008}_{0.008}$	$0.739\pm^{0.015}_{0.012}$
$35 < E_T < 40$	$0.872\pm^{0.008}_{0.008}$	$0.841\pm^{0.012}_{0.011}$	$0.852\pm^{0.014}_{0.013}$	$0.712\pm^{0.021}_{0.020}$
$40 < E_T < 45$	$0.860\pm^{0.005}_{0.005}$	$0.842\pm^{0.006}_{0.006}$	$0.842\pm^{0.006}_{0.006}$	$0.720\pm^{0.012}_{0.009}$
$45 < E_T < 50$	$0.864\pm^{0.005}_{0.005}$	$0.841\pm^{0.006}_{0.006}$	$0.856\pm^{0.006}_{0.006}$	$0.716\pm^{0.012}_{0.010}$
$50 < E_T < 55$	$0.867\pm^{0.005}_{0.005}$	$0.847\pm^{0.007}_{0.007}$	$0.856\pm^{0.007}_{0.007}$	$0.731\pm^{0.014}_{0.011}$
$55 < E_T < 60$	$0.865\pm^{0.006}_{0.006}$	$0.837\pm^{0.009}_{0.009}$	$0.845\pm^{0.010}_{0.010}$	$0.737\pm^{0.017}_{0.015}$
$60 < E_T < 65$	$0.853\pm^{0.004}_{0.005}$	$0.831\pm^{0.006}_{0.006}$	$0.835\pm^{0.006}_{0.006}$	$0.700\pm^{0.012}_{0.009}$
$65 < E_T < 70$	$0.854\pm^{0.005}_{0.005}$	$0.832\pm^{0.006}_{0.006}$	$0.836\pm^{0.006}_{0.006}$	$0.713\pm^{0.012}_{0.009}$
$70 < E_T < 80$	$0.858\pm^{0.005}_{0.005}$	$0.837\pm^{0.006}_{0.006}$	$0.844\pm^{0.006}_{0.006}$	$0.713\pm^{0.012}_{0.009}$
$80 < E_T < 100$	$0.861\pm^{0.005}_{0.005}$	$0.838\pm^{0.007}_{0.007}$	$0.848\pm^{0.007}_{0.007}$	$0.720\pm^{0.013}_{0.010}$
$100 < E_T < 120$	$0.854\pm^{0.004}_{0.005}$	$0.834\pm^{0.006}_{0.006}$	$0.839\pm^{0.005}_{0.005}$	$0.705\pm^{0.012}_{0.009}$
$120 < E_T < 200$	$0.860\pm^{0.005}_{0.005}$	$0.838\pm^{0.006}_{0.006}$	$0.845\pm^{0.006}_{0.006}$	$0.723\pm^{0.013}_{0.010}$
$200 < E_T < 300$	$0.845\pm^{0.004}_{0.004}$	$0.823\pm^{0.006}_{0.006}$	$0.823\pm^{0.005}_{0.005}$	$0.698\pm^{0.011}_{0.009}$

Table 4.5: Data corrected efficiencies of a true signal candidate passing the selection criteria.

4.6 Background Subtraction Distributions

Section 4.5 described how we remove most of the background using the discriminating variables and determine the efficiency of this selection, what remains is a collection of candidates enriched in photons. The sample nevertheless still contains background events that have passed all of our requirements. To determine the purity of the sample (the fraction of all these candidates that are signal), we use an additional variable, $\sigma_{\eta\eta}$. We determine the expected distributions for photons and for the background in this variable, and we fit a linear sum of the photon and background distributions to the data. The result of this fit gives photon and background components of the overall sample and, therefore, the purity of the sample.

The variable $\sigma_{\eta\eta}$ is the log-weighted second moment of the distribution of the energy in the cluster in the ECAL measured in the pseudorapidity direction. It is defined as:

$$\sigma_{\eta\eta}^2 = \frac{\sum_{i=1}^{25} w_i(\eta_i - \bar{\eta})^2}{\sum_{i=1}^{25} w_i}, \; w_i = \max(0, 4.7 + \ln(\frac{E_i}{E})). \tag{4.1}$$

using the energy deposited in the 25 crystals around and including the crystal that contains the log-energy weighted center of the candidate's ECAL energy deposit. Log weighting is used because as a shower spreads in ECAL the amount of energy deposited decreases exponentially with radius. The background candidates's energy depositions are, on the other hand, expected to be generally broader than for the signal, since energy deposits from background come from multiple photons coming from the decay of hadrons, mostly single or multiple π^0 decays. A massive hadron decaying only to photons kinematically must have a non-zero opening angle and therefore must be separated in the detector. Although we may not be able to discriminate the separate energy depositions, they will be aligned at a random angle so in general the background will leave a broader overall energy deposition. The η direction is chosen for the second moment because due to the magnetic field we expect all showers to be broadened in the ϕ direction. The photon distribution is determined with the Monte Carlo and corrected using data, and the distribution for the background is taken from some data that has been previously eliminated by our selections making it depleted in signal. The procedures used to obtain the distribution used to fit the data are described below.

4.6.1 Distribution of $\sigma_{\eta\eta}$ for Photons

The photon distribution is found using the Monte Carlo as described in section 4.3. To validate this simulation of the electromagnetic showers, we compare the $\sigma_{\eta\eta}$ variable applied to electrons coming from the decay of a Z boson in data and in the simulation, and we observe a slight shift in $\sigma_{\eta\eta}$ between data and Monte Carlo. We measure this shift in each of the pseudorapidity regions and at different energies. The size and variation of this shift is in table 4.6.

To determine the photon distribution we use in the background subtraction fit, we use the distribution from simulated photons and add a constant from table 4.6. Some of the photon distributions are shown in figure 4.8. The corrective constants are small but are needed to faithfully reproduce our observations. In chapter 5 we examine the overall effect of this shift on the results.

Pseudorapidity Bin	Shift	RMS(over different E_t)
$\lvert\eta\rvert < 0.9$	3.2×10^{-5}	9.3×10^{-6}
$0.9 < \lvert\eta\rvert < 1.4442$	8.3×10^{-5}	3.4×10^{-5}
$1.566 < \lvert\eta\rvert < 2.1$	5.9×10^{-4}	1.2×10^{-4}
$2.1 < \lvert\eta\rvert < 2.5$	3.1×10^{-4}	3.5×10^{-4}

Table 4.6: Shift to the Photon $\sigma_{\eta\eta}$ distribution.

4.6.2 Distribution of $\sigma_{\eta\eta}$ for Background Events

We use data to directly determine the distribution of $\sigma_{\eta\eta}$ for background events. For this we require that the set of background events have the same characteristics in the calorimeter as the background that is mixed with the photon signal, but where the photon signal is heavily suppressed. To this end we use a sideband of the track isolation selection to exclude the signal region, but include events that fall near the signal region. The selection criteria for this new disjoint sample are given in table 4.7. Some examples of these distributions are shown in Figure 4.9. In this Figure, The agreement between the Monte Carlo from the signal region and the alternate selection applied to the data appear good, but uncertainties in the shape of this distribution will lead to a large systematic uncertainty; that is discussed in chapter 5.

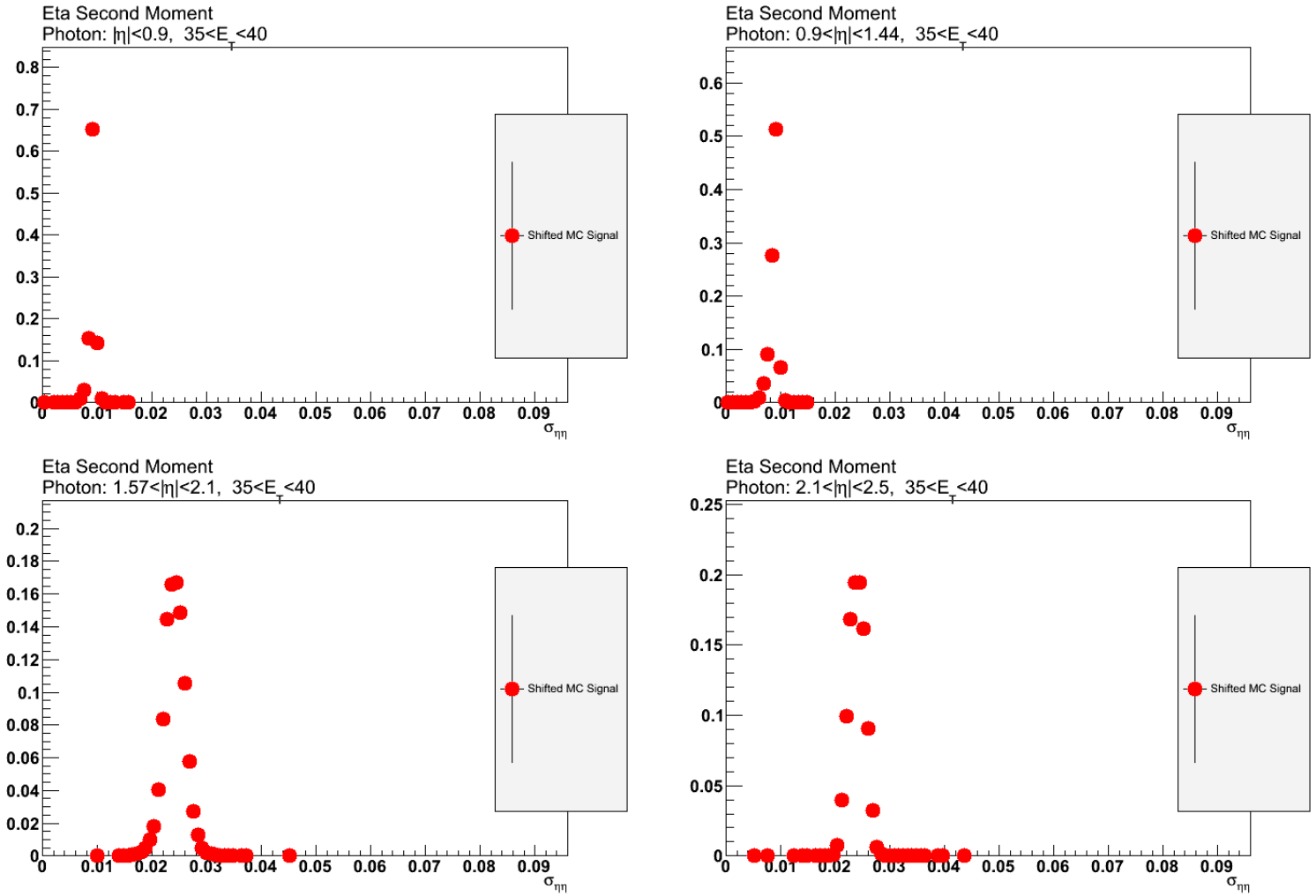

Figure 4.8: Photon distributions for $\sigma_{\eta\eta}$, which will be used in measuring the final signal purity.

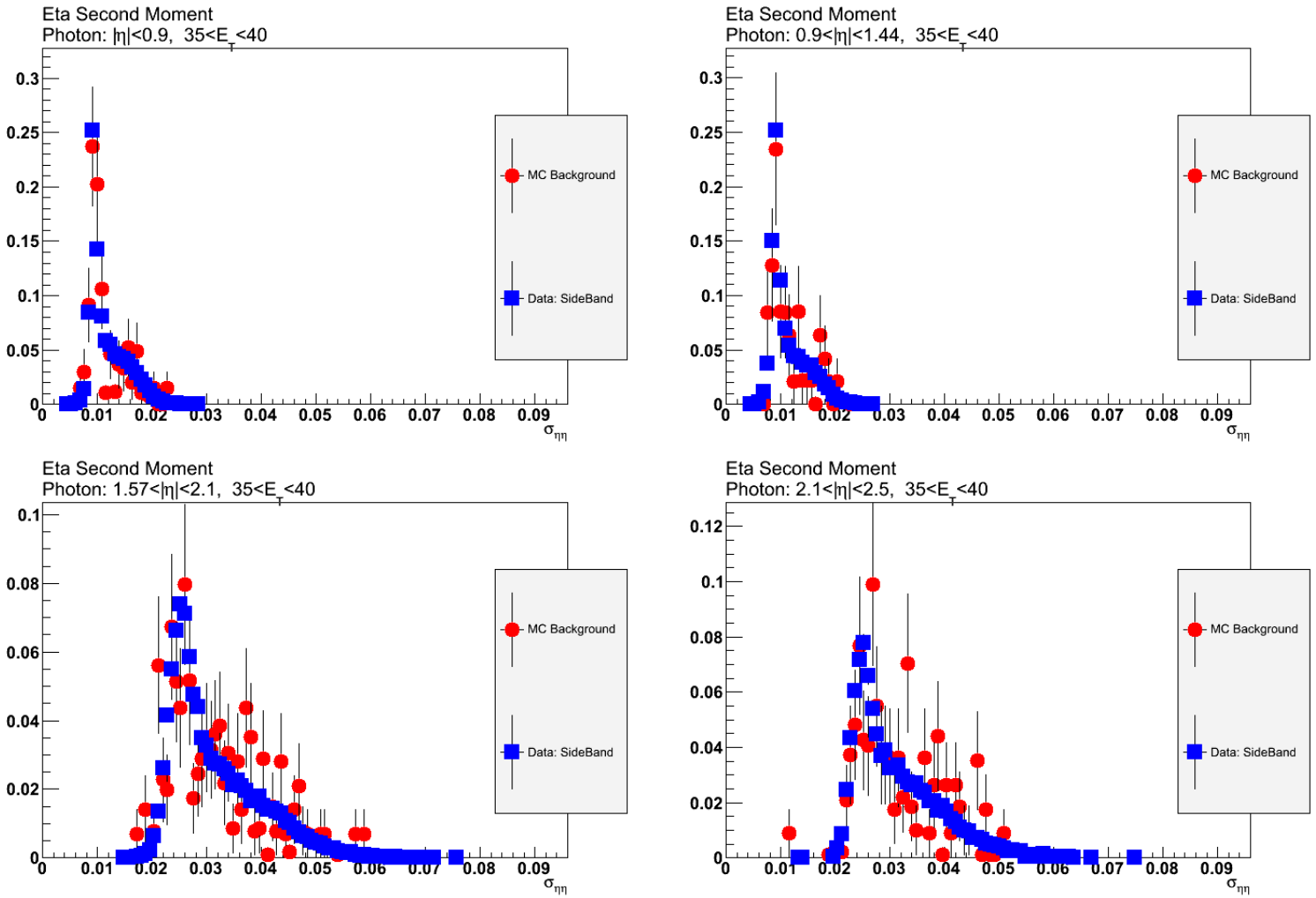

Figure 4.9: Background distributions for $\sigma_{\eta\eta}$, which will be used in measuring the final signal purity. In black are Monte Carlo simulations of the background in the single region. In red is the data taken from a changed selection in the track isolation variable.

Variable	Selection
ECAL Isolation	ECAL Isolation Energy $< 4.2 + 0.003 * E_T$
HCAL Isolation	HCAL Isolation Energy$< 2.2 + 0.001 * E_T$
Track Isolation	$2.0+0.001*E_T$ <Momentum of near tracks$< 5.0+0.001*E_T$
Associated Pixel Hits	"false"

Table 4.7: Selections placed on the data to give background distributions.

4.7 Results

Taking the distributions discussed in sections 4.6.1 and 4.6.2 we now fit the data. To find the optimal linear sum of the $\sigma_{\eta\eta}$ distribution for photons and the $\sigma_{\eta\eta}$ distribution for background for each energy and pseudorapidity bin we use a Extended Maximum Log-likelihood fit [43].

The fit we use assumes Poissonian statistics in each bin and finds the optimal linear sum of the two distributions. For a given energy and pseudorapidity bin, we have distributions with n bins. We then fit the distribution from data, which has N_i candidates in bin i, with a linear sum of the normalized photon and background distributions that have the values in their bins of $D_{\gamma i}$ and D_{Bi} respectively. We scale these distributions by the N_γ for the number of photon candidates and N_B for the number of background candidates. For a given value of N_γ and N_B the Poissonian likelihood of having N_i events in that bin is

$$\mathcal{L}_i = \frac{1}{N_i!} e^{-(N_\gamma D_{\gamma i} + N_b D_{Bi})} (N_\gamma D_{\gamma i} + N_b D_{Bi})^{N_i} \tag{4.2}$$

So, the likelihood distribution for all bins is then $\mathcal{L} = \prod_{i=1}^{n} \mathcal{L}_i$. Letting $C = \sum_{i=1}^{n} (-\ln N_i!)$, the natural logarithm of the likelihood is

$$\begin{aligned} \ln \mathcal{L} &= \sum_{i=1}^{n} \ln \mathcal{L}_i \\ &= C - \sum_{i=1}^{n} (N_\gamma D_{\gamma i} + N_b D_{Bi}) + \sum_{i=1}^{n} N_i \ln(N_\gamma D_{\gamma i} + N_b D_{Bi}) \\ &= C - (N_\gamma + N_b) + \sum_{i=1}^{n} N_i \ln(N_\gamma D_{\gamma i} + N_b D_{Bi}) \end{aligned} \tag{4.3}$$

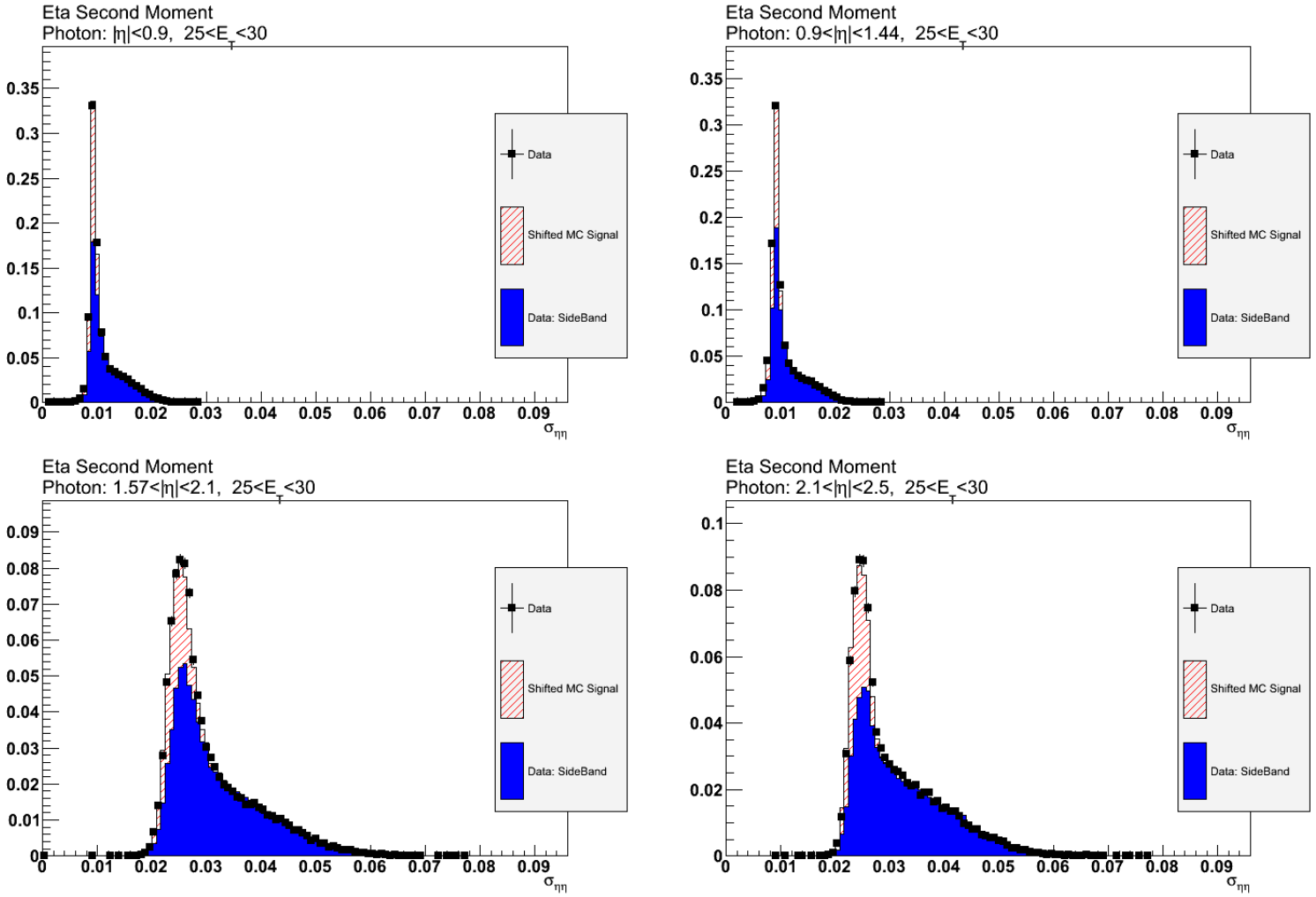

Figure 4.10: Fits of the data with signal and background distributions for the bins with a transverse momentum between 25 GeV and 30 GeV

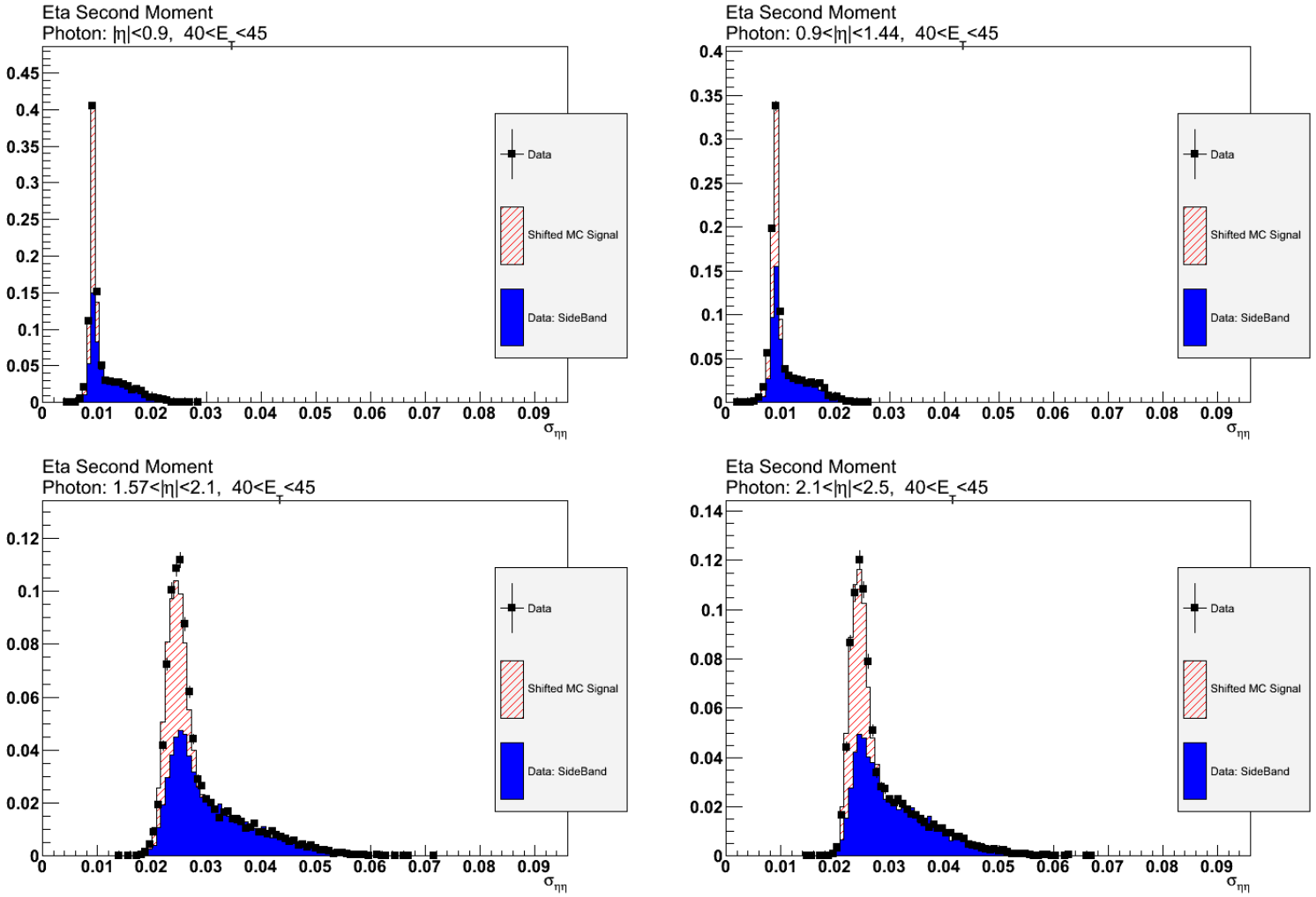

Figure 4.11: Fits of the data with signal and background distributions for the bins with a transverse momentum between 40 GeV and 45 GeV

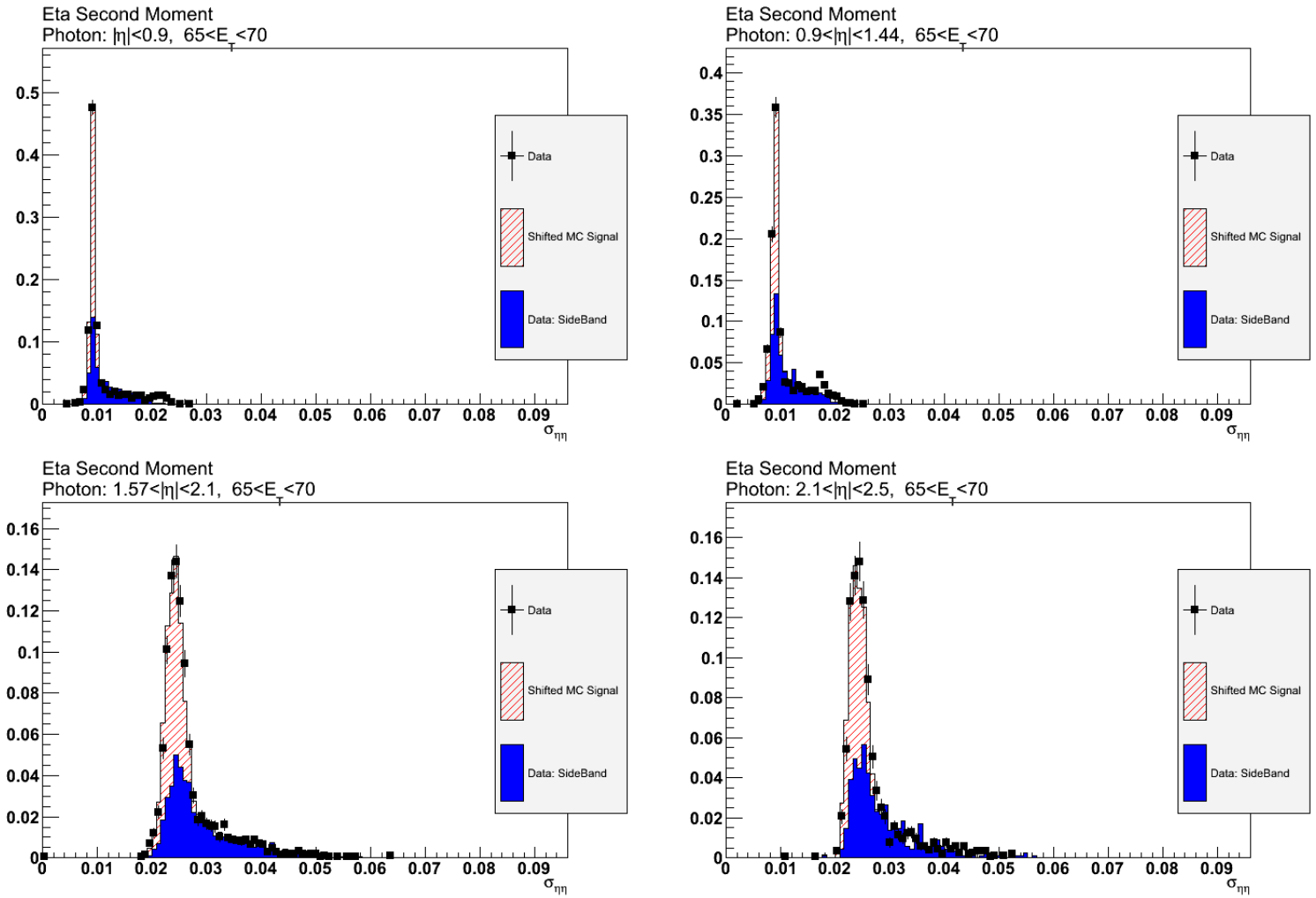

Figure 4.12: Fits of the data with signal and background distributions for the bins with a transverse momentum between 65 GeV and 70 GeV

We then maximize $\ln \mathcal{L}$ for the N_γ and N_B parameters using Minuit [44]. The statistical uncertainty for the parameter N_γ, the number of photons measured, is computed from the covariance matrix of this function. Examples of the fits can be seen in figures 4.10, 4.11 and 4.12.

The yields are in table 4.8. Naively one would expect these yield to monotonically decrease with energy but there are two factors complicating this; firstly, the bin size varies as one goes to high transverse energy; secondly, the integrated luminosity varies due to the changing trigger requirements. Table 4.9 takes into account the total integrated luminosity used for each bin to give a cross section for that region in the differential parameter space.

4.8 Deconvolution

In table 4.9 we report a measured differential isolated photon cross section resulting from energy clusters observed in CMS. The values in table 4.9 include a bin migration due to the resolution of the detector. We must remove the bin migration to compare the cross section to the theoretical predictions. To reconstruct the spectrum absent this effect, we use a package named RooUnfold [45], which conveniently encodes several deconvolution algorithms. Using RooUnfold we apply an iterative Bayesian technique [46] to perform the deconvolution.

As the name would suggest the iterative Bayesian technique is an iterative process based on Bayes theorem. Generally we consider a set of causes C_i, in our case a photon having a specific energy and rapidity, and a set of effects E_i, the fact that we measure a photon as having an energy and rapidity in a certain range. Let $P(E_i)$ be the probability that this photon is observed at a given rapidity and energy, and $P_m(C_j)$ be an estimate of the probability that the photon truly had a specific energy and rapidity. We denote the probability that a photon observed at specific energy and rapidity actually had a specific energy and rapidity as $P(C_i|E_j)$ and the converse, the probability that a photon having a given energy and rapidity be observed at a specific energy and rapidity as $P(E_i|C_i)$. Bayes theorem then states:

$$P_m(C_i|E_j) = \frac{P(E_j|C_i)P_m(C_i)}{\sum_{l=1}^{n_c} P(E_j|C_l)P_m(C_l)} \tag{4.4}$$

	$\|\eta\| < 0.9$	$0.9 < \|\eta\| < 1.4442$	$1.566 < \|\eta\| < 2.1$	$2.1 < \|\eta\| < 2.5$
$25 < E_T < 30$	$1.63 \times 10^4 \pm 247$	$9.73 \times 10^3 \pm 203$	$9.23 \times 10^3 \pm 193$	$6.95 \times 10^3 \pm 145$
$30 < E_T < 35$	$7.93 \times 10^3 \pm 154$	$4.63 \times 10^3 \pm 125$	$4.03 \times 10^3 \pm 119$	$2.97 \times 10^3 \pm 93.4$
$35 < E_T < 40$	$1.38 \times 10^4 \pm 195$	$7.79 \times 10^3 \pm 160$	$7.61 \times 10^3 \pm 151$	$5.38 \times 10^3 \pm 117$
$40 < E_T < 45$	$8.11 \times 10^3 \pm 140$	$4.28 \times 10^3 \pm 112$	$4.61 \times 10^3 \pm 110$	$3.23 \times 10^3 \pm 83.7$
$45 < E_T < 50$	$4.82 \times 10^3 \pm 106$	$2.49 \times 10^3 \pm 86.7$	$2.81 \times 10^3 \pm 82.2$	$1.67 \times 10^3 \pm 64.5$
$50 < E_T < 55$	$2.96 \times 10^3 \pm 80$	$1.76 \times 10^3 \pm 65.2$	$1.76 \times 10^3 \pm 62.3$	$1.18 \times 10^3 \pm 50.2$
$55 < E_T < 60$	$4.46 \times 10^3 \pm 94.2$	$2.43 \times 10^3 \pm 78.6$	$2.65 \times 10^3 \pm 74$	$1.47 \times 10^3 \pm 58.7$
$60 < E_T < 65$	$3.04 \times 10^3 \pm 75.3$	$1.56 \times 10^3 \pm 63.1$	$1.73 \times 10^3 \pm 58.7$	$1.06 \times 10^3 \pm 46.3$
$65 < E_T < 70$	$1.94 \times 10^3 \pm 63$	$1.1 \times 10^3 \pm 51.4$	$1.12 \times 10^3 \pm 48.5$	733 ± 39.4
$70 < E_T < 80$	$2.54 \times 10^3 \pm 70.2$	$1.44 \times 10^3 \pm 57.5$	$1.59 \times 10^3 \pm 53.3$	$1.04 \times 10^3 \pm 43.1$
$80 < E_T < 100$	$4.15 \times 10^3 \pm 92.9$	$2.32 \times 10^3 \pm 76.4$	$2.76 \times 10^3 \pm 70.6$	$1.73 \times 10^3 \pm 56.1$
$100 < E_T < 120$	$1.79 \times 10^3 \pm 57.6$	991 ± 47.1	$1.02 \times 10^3 \pm 42.7$	630 ± 30.6
$120 < E_T < 200$	$1.66 \times 10^3 \pm 53.4$	872 ± 42.7	802 ± 38.6	457 ± 27.1
$200 < E_T < 300$	211 ± 17.4	126 ± 12.2	Fit Not Done	Fit Not Done

Table 4.8: non-Efficiency-corrected yields: events per bin

	$\|\eta\| < 0.9$	$0.9 < \|\eta\| < 1.4442$	$1.566 < \|\eta\| < 2.1$	$2.1 < \|\eta\| < 2.5$
$25 < E_T < 30$	$7867.331 \pm^{125.313}_{125.696}$ pb	$4806.998 \pm^{105.703}_{105.753}$ pb	$4560.099 \pm^{99.634}_{99.631}$ pb	$4048.391 \pm^{107.435}_{98.809}$ pb
$30 < E_T < 35$	$3817.188 \pm^{76.552}_{76.699}$ pb	$2287.415 \pm^{64.039}_{64.058}$ pb	$1993.194 \pm^{60.405}_{60.405}$ pb	$1728.789 \pm^{61.406}_{58.705}$ pb
$35 < E_T < 40$	$1988.664 \pm^{29.820}_{29.923}$ pb	$1150.736 \pm^{24.972}_{24.984}$ pb	$1124.321 \pm^{23.494}_{23.493}$ pb	$937.135 \pm^{25.573}_{23.636}$ pb
$40 < E_T < 45$	$1166.649 \pm^{20.985}_{21.035}$ pb	$631.635 \pm^{17.192}_{17.197}$ pb	$681.392 \pm^{16.763}_{16.762}$ pb	$563.060 \pm^{17.270}_{16.243}$ pb
$45 < E_T < 50$	$692.639 \pm^{15.603}_{15.626}$ pb	$368.251 \pm^{13.066}_{13.069}$ pb	$415.513 \pm^{12.425}_{12.424}$ pb	$291.373 \pm^{12.207}_{11.823}$ pb
$50 < E_T < 55$	$426.043 \pm^{11.713}_{11.725}$ pb	$259.436 \pm^{9.803}_{9.805}$ pb	$259.524 \pm^{9.357}_{9.357}$ pb	$205.920 \pm^{9.374}_{9.125}$ pb
$55 < E_T < 60$	$298.435 \pm^{6.479}_{6.490}$ pb	$167.171 \pm^{5.524}_{5.526}$ pb	$181.887 \pm^{5.211}_{5.211}$ pb	$118.823 \pm^{5.137}_{4.986}$ pb
$60 < E_T < 65$	$203.165 \pm^{5.139}_{5.145}$ pb	$107.204 \pm^{4.395}_{4.396}$ pb	$118.913 \pm^{4.100}_{4.100}$ pb	$85.806 \pm^{4.004}_{3.903}$ pb
$65 < E_T < 70$	$130.038 \pm^{4.265}_{4.268}$ pb	$75.751 \pm^{3.569}_{3.570}$ pb	$77.232 \pm^{3.364}_{3.364}$ pb	$59.313 \pm^{3.338}_{3.280}$ pb
$70 < E_T < 80$	$169.937 \pm^{4.774}_{4.779}$ pb	$98.876 \pm^{4.010}_{4.010}$ pb	$109.290 \pm^{3.725}_{3.725}$ pb	$84.321 \pm^{3.752}_{3.648}$ pb
$80 < E_T < 100$	$135.846 \pm^{3.121}_{3.126}$ pb	$78.113 \pm^{2.628}_{2.628}$ pb	$92.979 \pm^{2.447}_{2.447}$ pb	$68.610 \pm^{2.494}_{2.389}$ pb
$100 < E_T < 120$	$58.792 \pm^{1.911}_{1.912}$ pb	$33.313 \pm^{1.602}_{1.602}$ pb	$34.184 \pm^{1.454}_{1.454}$ pb	$24.988 \pm^{1.281}_{1.254}$ pb
$120 < E_T < 200$	$54.449 \pm^{1.772}_{1.773}$ pb	$29.318 \pm^{1.452}_{1.452}$ pb	$26.989 \pm^{1.310}_{1.310}$ pb	$18.115 \pm^{1.116}_{1.100}$ pb
$200 < E_T < 300$	$6.918 \pm^{0.570}_{0.570}$ pb	$4.239 \pm^{0.411}_{0.411}$ pb	Fit Not Done	Fit Not Done

Table 4.9: Values of cross section found from the fits with statistical errors.

We also define the probability that a photon with a given energy and rapidity be observed at any energy and rapidity as $\epsilon_i \equiv \sum_{j=1}^{n_E} P(E_j|C_i)$. We use Monte Carlo to determine the smearing matrix $P(E_j|C_i)$ and our measured data gives us the number of event in each effect $n(E_j)$. Then we define an iterative process:

$$n_{m+1}(C_i) = \frac{1}{\epsilon_i} \sum_{j=1}^{n_E} n(E_j) P_m(C_i|E_j) \quad , \quad P_{m+1}(C_i) = \frac{n_{m+1}(C_i)}{\sum_{i=1}^{n_C} n_{m+1}(C_i)} \tag{4.5}$$

This gives successively better estimates of $P(C_i)$ on each iteration. Once the results have converge we have a deconvolved distribution which may be compared to theory. The spectrum that is a result of this deconvolution is in table 4.10; this is the final measurement of the cross section.

	$\lvert\eta\rvert < 0.9$	$0.9 < \lvert\eta\rvert < 1.4442$	$1.566 < \lvert\eta\rvert < 2.1$	$2.1 < \lvert\eta\rvert < 2.5$
$25 < E_T < 30$	$899.510 \pm^{14.328}_{14.371} \frac{\text{pb}}{\text{GeV}}$	$945.106 \pm^{20.783}_{20.792} \frac{\text{pb}}{\text{GeV}}$	$892.415 \pm^{19.498}_{19.497} \frac{\text{pb}}{\text{GeV}}$	$1007.988 \pm^{26.750}_{24.602} \frac{\text{pb}}{\text{GeV}}$
$30 < E_T < 35$	$398.421 \pm^{7.990}_{8.005} \frac{\text{pb}}{\text{GeV}}$	$366.782 \pm^{10.269}_{10.272} \frac{\text{pb}}{\text{GeV}}$	$322.193 \pm^{9.764}_{9.764} \frac{\text{pb}}{\text{GeV}}$	$417.689 \pm^{14.837}_{14.183} \frac{\text{pb}}{\text{GeV}}$
$35 < E_T < 40$	$217.735 \pm^{3.265}_{3.276} \frac{\text{pb}}{\text{GeV}}$	$206.364 \pm^{4.479}_{4.480} \frac{\text{pb}}{\text{GeV}}$	$211.414 \pm^{4.418}_{4.417} \frac{\text{pb}}{\text{GeV}}$	$238.762 \pm^{6.515}_{6.022} \frac{\text{pb}}{\text{GeV}}$
$40 < E_T < 45$	$131.557 \pm^{2.367}_{2.372} \frac{\text{pb}}{\text{GeV}}$	$115.074 \pm^{3.132}_{3.133} \frac{\text{pb}}{\text{GeV}}$	$132.444 \pm^{3.258}_{3.258} \frac{\text{pb}}{\text{GeV}}$	$147.728 \pm^{4.532}_{4.262} \frac{\text{pb}}{\text{GeV}}$
$45 < E_T < 50$	$77.169 \pm^{1.739}_{1.741} \frac{\text{pb}}{\text{GeV}}$	$66.608 \pm^{2.363}_{2.363} \frac{\text{pb}}{\text{GeV}}$	$79.165 \pm^{2.367}_{2.367} \frac{\text{pb}}{\text{GeV}}$	$74.021 \pm^{3.101}_{3.004} \frac{\text{pb}}{\text{GeV}}$
$50 < E_T < 55$	$46.283 \pm^{1.272}_{1.274} \frac{\text{pb}}{\text{GeV}}$	$47.242 \pm^{1.785}_{1.786} \frac{\text{pb}}{\text{GeV}}$	$47.818 \pm^{1.724}_{1.724} \frac{\text{pb}}{\text{GeV}}$	$53.364 \pm^{2.430}_{2.364} \frac{\text{pb}}{\text{GeV}}$
$55 < E_T < 60$	$33.421 \pm^{0.726}_{0.727} \frac{\text{pb}}{\text{GeV}}$	$30.384 \pm^{1.004}_{1.004} \frac{\text{pb}}{\text{GeV}}$	$35.774 \pm^{1.025}_{1.025} \frac{\text{pb}}{\text{GeV}}$	$30.176 \pm^{1.304}_{1.266} \frac{\text{pb}}{\text{GeV}}$
$60 < E_T < 65$	$23.014 \pm^{0.582}_{0.583} \frac{\text{pb}}{\text{GeV}}$	$19.454 \pm^{0.798}_{0.798} \frac{\text{pb}}{\text{GeV}}$	$23.048 \pm^{0.795}_{0.795} \frac{\text{pb}}{\text{GeV}}$	$22.223 \pm^{1.037}_{1.011} \frac{\text{pb}}{\text{GeV}}$
$65 < E_T < 70$	$15.455 \pm^{0.507}_{0.507} \frac{\text{pb}}{\text{GeV}}$	$15.174 \pm^{0.715}_{0.715} \frac{\text{pb}}{\text{GeV}}$	$17.163 \pm^{0.748}_{0.748} \frac{\text{pb}}{\text{GeV}}$	$17.021 \pm^{0.957}_{0.941} \frac{\text{pb}}{\text{GeV}}$
$70 < E_T < 80$	$9.461 \pm^{0.266}_{0.266} \frac{\text{pb}}{\text{GeV}}$	$8.888 \pm^{0.360}_{0.360} \frac{\text{pb}}{\text{GeV}}$	$10.896 \pm^{0.372}_{0.372} \frac{\text{pb}}{\text{GeV}}$	$11.347 \pm^{0.505}_{0.491} \frac{\text{pb}}{\text{GeV}}$
$80 < E_T < 100$	$3.587 \pm^{0.083}_{0.083} \frac{\text{pb}}{\text{GeV}}$	$3.279 \pm^{0.111}_{0.111} \frac{\text{pb}}{\text{GeV}}$	$4.329 \pm^{0.114}_{0.114} \frac{\text{pb}}{\text{GeV}}$	$4.311 \pm^{0.157}_{0.150} \frac{\text{pb}}{\text{GeV}}$
$100 < E_T < 120$	$1.718 \pm^{0.056}_{0.056} \frac{\text{pb}}{\text{GeV}}$	$1.611 \pm^{0.078}_{0.078} \frac{\text{pb}}{\text{GeV}}$	$1.782 \pm^{0.076}_{0.076} \frac{\text{pb}}{\text{GeV}}$	$1.716 \pm^{0.088}_{0.086} \frac{\text{pb}}{\text{GeV}}$
$120 < E_T < 200$	$0.358 \pm^{0.011}_{0.011} \frac{\text{pb}}{\text{GeV}}$	$0.301 \pm^{0.015}_{0.015} \frac{\text{pb}}{\text{GeV}}$	$0.314 \pm^{0.015}_{0.015} \frac{\text{pb}}{\text{GeV}}$	$0.284 \pm^{0.017}_{0.017} \frac{\text{pb}}{\text{GeV}}$
$200 < E_T < 300$	$0.039 \pm^{0.003}_{0.003} \frac{\text{pb}}{\text{GeV}}$	$0.039 \pm^{0.004}_{0.004} \frac{\text{pb}}{\text{GeV}}$	Fit Not Done	Fit Not Done

Table 4.10: Cross section measurement with statistical errors. Here the cross section is shown in pb per GeV per unit of rapidity.

Chapter 5

Systematics

5.1 Introduction

When making a measurement it is as important to understand how well one has determined the values as it is to determine the values themselves. We have measured the isolated photon production cross section at the LHC with associated statistical errors. In this chapter we determine the systematic uncertainties on these measured values. While making this measurement we have used the best possible assumptions for the photon and background distributions used to fit the data, and for the efficiency of the selection used to refine the dataset. Here, we determine the variation that we expect because of errors in these assumptions and other aspects of our technique.

5.2 Distribution of Background

After refining the sample by applying the selection in table 4.4, we determine the component of the remaining sample attributable to isolated photon events by fitting the measured distribution of the $\sigma_{\eta\eta}$ variable with a sum of the $\sigma_{\eta\eta}$ distributions expected for photons and for background. In this section, we determine the uncertainty associated with the assumptions we made in determining the shape of the background distribution used in the fit. In section 4.6.2, we used a sideband in the track isolation variable to determine the distribution of $\sigma_{\eta\eta}$ for background events passing the selection. The track isolation variable was used to define this new sample of events because being a

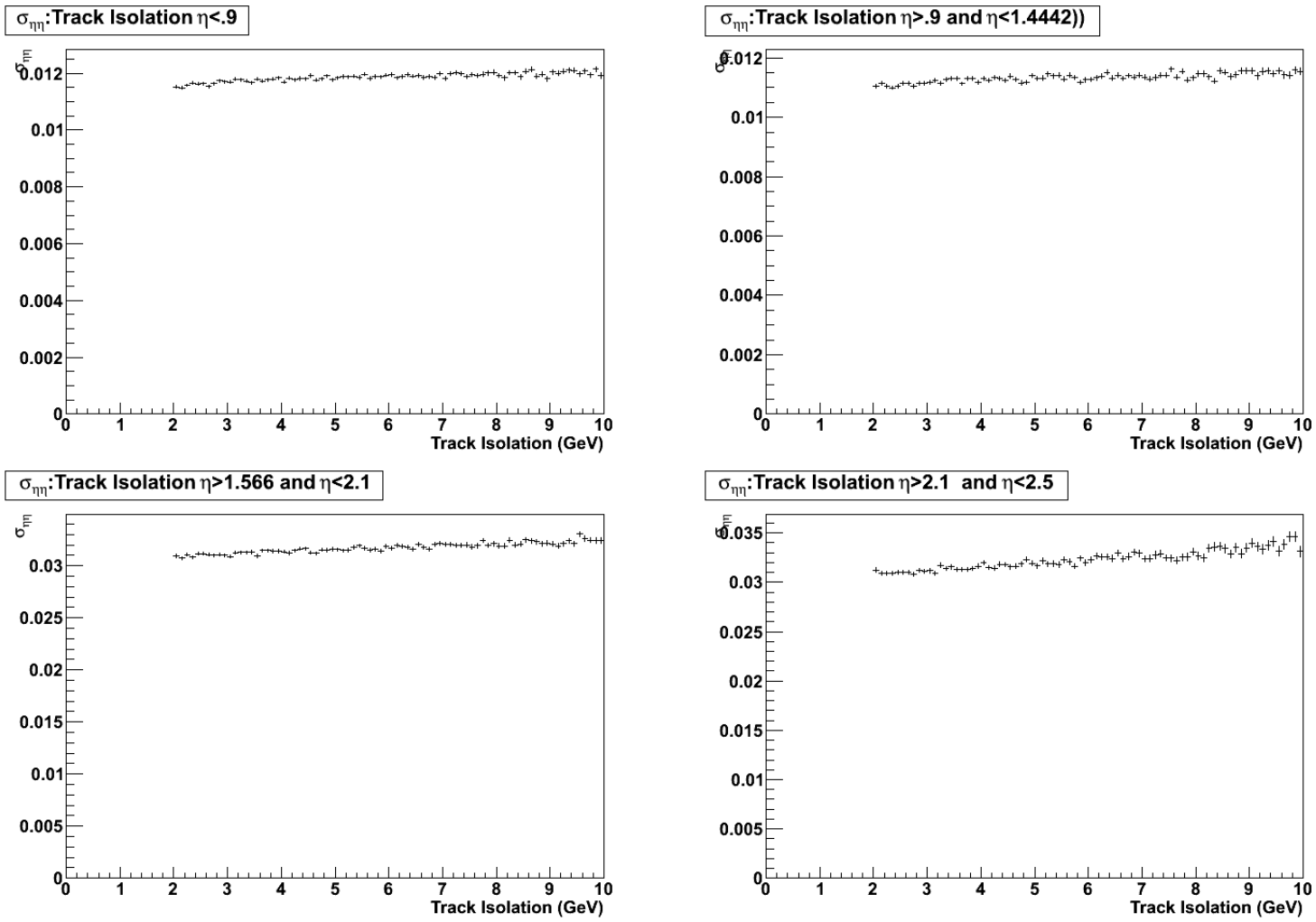

Figure 5.1: The change in the width of the shower in η as a function of the track isolation. The region between 0 GeV and 2 GeV contains the data from which we extract the cross section and are a mix between photons and background. The bins from 2 GeV to 5 GeV are used to approximate $\sigma_{\eta\eta}$ distribution of the background in the region used to determine the cross section.

non-calorimetric variable it is less correlated with $\sigma_{\eta\eta}$, which is derived from values measured in the ECAL. To investigate this correlation, figure 5.1 shows the change in the mean value of $\sigma_{\eta\eta}$ as a function of the track isolation variable in each of the pseudorapidity regions. In this distribution, one can see a small trend toward larger values of $\sigma_{\eta\eta}$ when more momentum is allowed in the track isolation variable. This change is the source of the dominant systematic uncertainty in this analysis.

We estimate this uncertainty by defining 3 new sidebands in the track isolation variable, between 2 GeV and 3.5 GeV, between 3.5 GeV and 5 GeV, and between 5 GeV and 6.5 GeV. Using these new sidebands we refit the data distribution to calculate the number of photons measured. We then compare this measurement to the yield determine in section 4.7. Figure 5.2 shows the relative change that occurs when using each of the three new sidebands instead of the original sideband. The final value used for the uncertainty is the average of each of these values and is also shown in the figure.

5.3 Distribution of Photons

In this section we discuss the investigation of the effect of the uncertainty in the signal $\sigma_{\eta\eta}$ distribution used for photons in determining the photon content. This distribution is taken from the Monte Carlo sample. Then there is a correction applied to it based on a small shift between the data and Monte Carlo distributions for electrons from Z boson decay. To determine the size of this effect we recalculate the cross section fitting the signal and background without applying the shift. The effect of this difference is shown in figure 5.3 and is too small to measure in most cases and in all cases is small enough to be overshadowed by other uncertainties when added in quadrature. The largest deviation this change in the measurement gives is less than 0.1%.

5.4 Statistics of Distribution Used in Fit

Unlike the relatively simple electromagnetic showers from photons and electrons, jets are difficult to model in Monte Carlo. Because $\sigma_{\eta\eta}$ can be sensitive to the particle makeup of the jets we choose to use jets from the data using a sideband of the track isolation variable to determine the shape of the $\sigma_{\eta\eta}$ distribution for the background

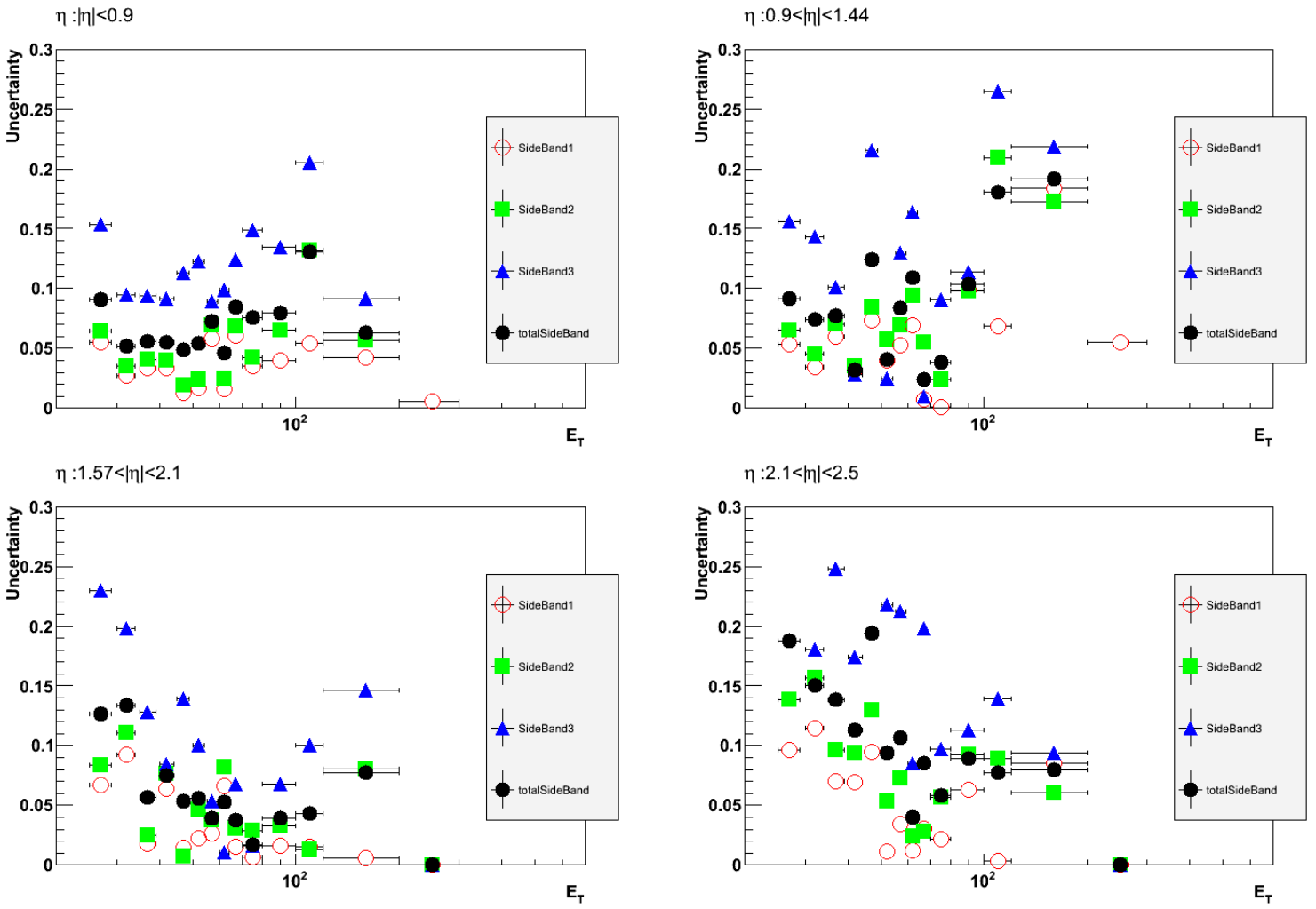

Figure 5.2: The uncertainty determined by looking at the deviation from the determined value using each of three different selections to determine the purity of the signal sample. The black circles show the value used which is the average of these three values.

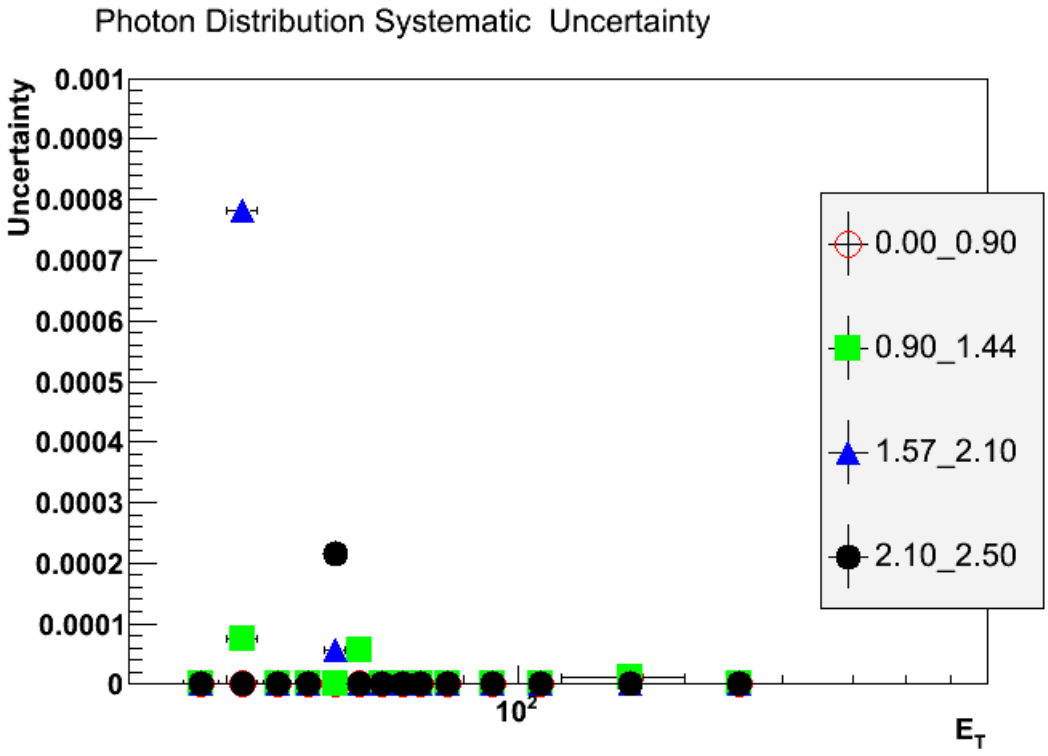

Figure 5.3: The effect of not including a shift in the signal distribution taken from Monte Carlo is shown.

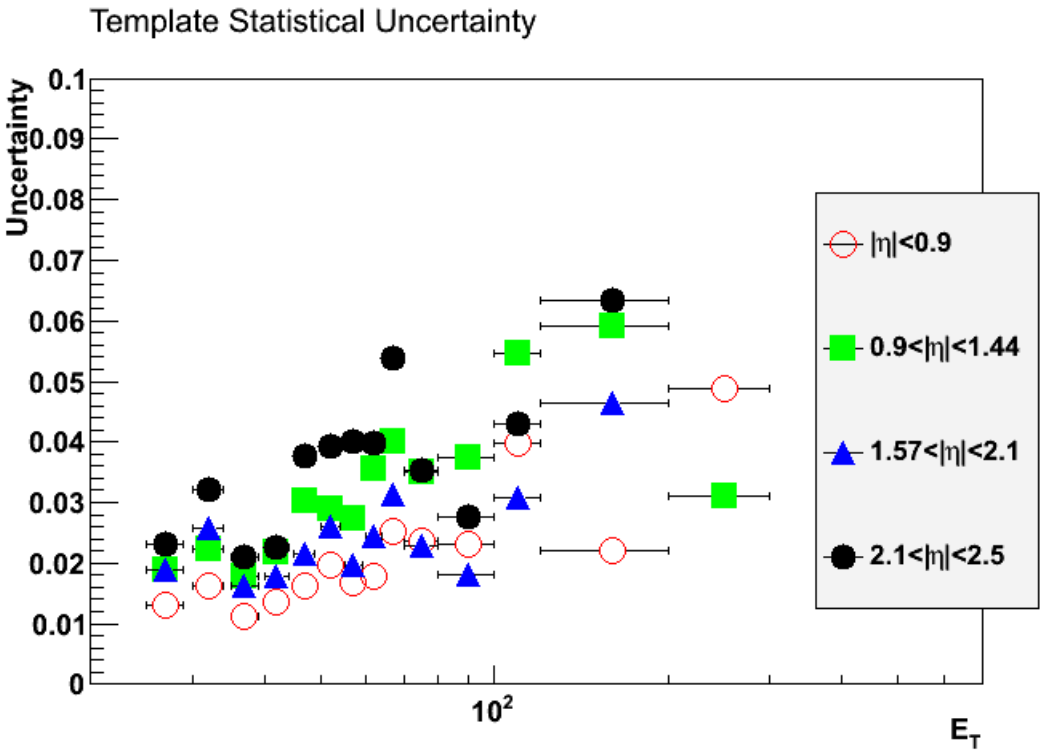

Figure 5.4: The uncertainty due to the number of events in the fitting distribution, in each of the four pseudorapidity regions.

component of the sample. The statistical uncertainty calculated in chapter 4 only takes into account the number of signal events. Because the $\sigma_{\eta\eta}$ distribution of the background used for the fit is taken from data it is not as well populated as it would be if it were taken from Monte Carlo. This leads to an additional statistical uncertainty, which is included here as a systematic because, although it does scale with the amount of data used, one could change the technique of acquiring this distribution to a Monte Carlo based technique or some other and not have an uncertainty, which scales with the data taken. This uncertainty is found by looking at the variation one finds when resampling the distribution thereby measuring the fluctuations due to statistics within this distribution. The systematic uncertainty is shown in figure 5.4.

5.5 Selection Efficiency

The efficiency of the selection we apply to enrich the sample was first determined in Monte Carlo, then there was a correction applied based on the efficiency of electrons from Z boson decays. To determine the uncertainty of the efficiency of this selection we vary the correction by the maximum size of the correction in each of the four pseudorapdidity regions of the detector. This gives the uncertainties in figure 5.5.

5.6 Luminosity

There is an overall 4% uncertainty that CMS assigns to its measurement of the total integrated luminosity [38]. Because this is an overall shift it would act identically in each bin, it is considered separately from the other systematic uncertainties.

5.7 Unfolding

The uncertainty due to the deconvolution procedure is taken from doing this procedure in several ways and combining the variability of the results. The central value is done using a Bayesian deconvolution. The next procedure considered is simply computing the inverse matrix using the same Monte Carlo. This inversion yields results identical to the bayesian deconvolution, so this is not used in the calculation of the uncertainty.

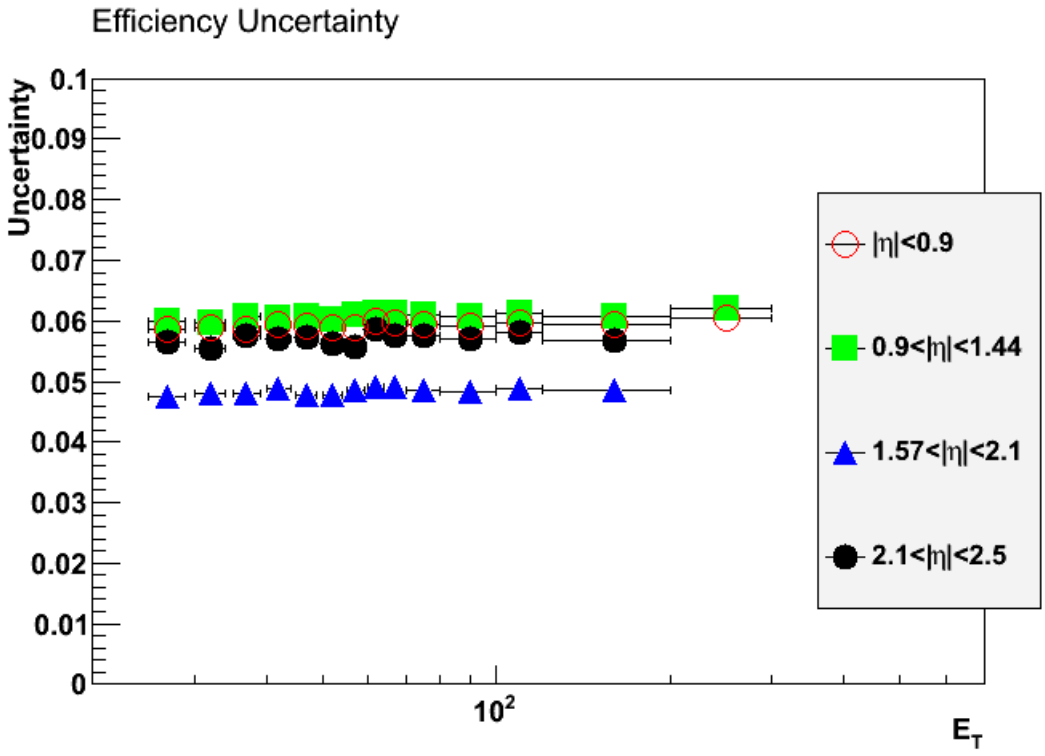

Figure 5.5: The uncertainty of the cross section due to the efficiency of the cuts, in each of the four pseudorapidity regions.

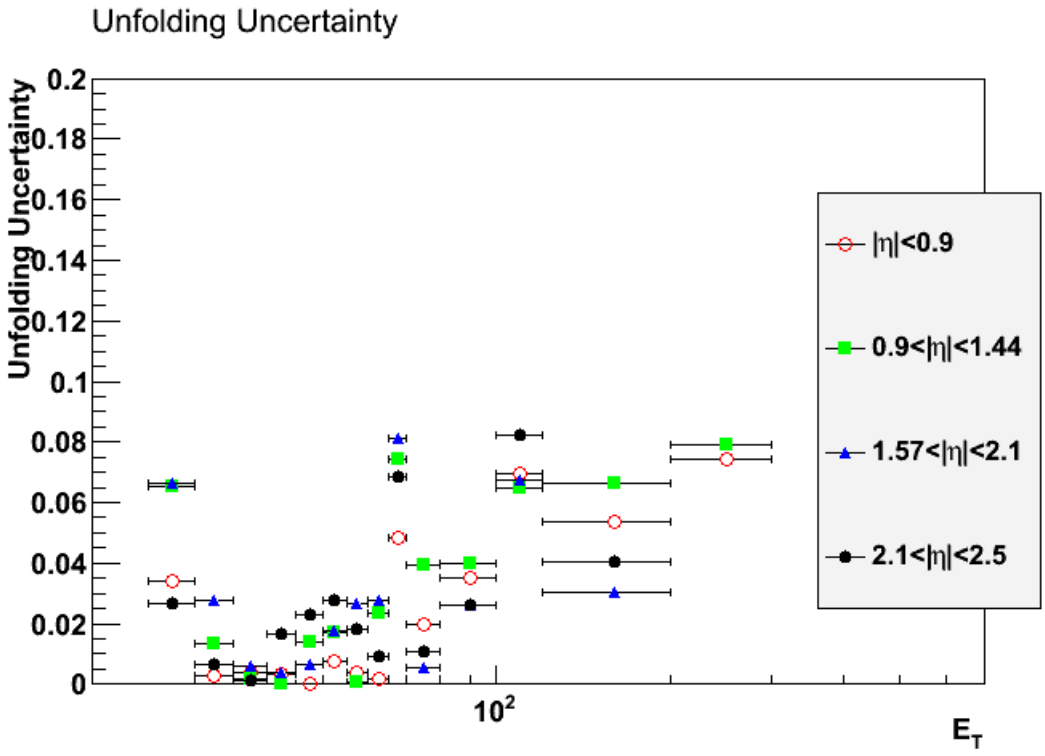

Figure 5.6: The uncertainty calculated by combining different unfolding techniques, in each of the four pseudorapidity regions.

We then apply a bin-by-bin unfolding which is taken from the diagonal of the response matrix defined in section 4.8 and a single value decomposition [47]. The variation of these other deconvolutions from the bayesian technique is used as the uncertainty due to the resolution deconvolution procedure and is shown in figure 5.6.

5.8 Combined Systematics

In order to find the total systematic uncertainty we add each of the uncertainties in quadrature. This gives the final result shown in figure 5.7. For most of the measurement the systematic uncertainty is dominated by the uncertainty in the background distribution. Generally the second most significant source is the systematic uncertainty in the efficiency. Again we note that we do not include the 4% uncertainty in the luminosity in this combined systematic uncertainty.

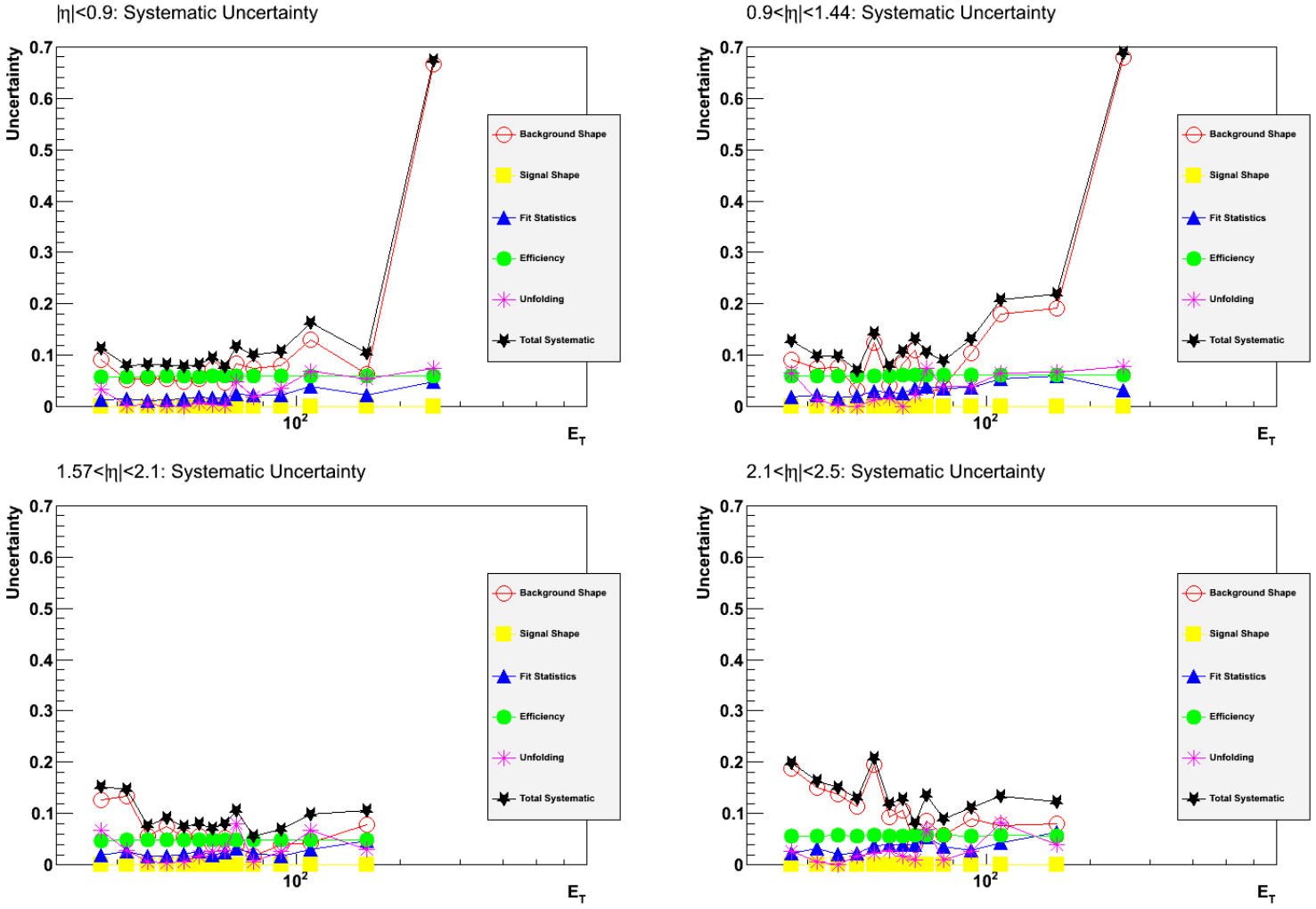

Figure 5.7: Each systematic uncertainty is plotted and total uncertainty is shown in black. The total is the sum in quadrature of all the sources of uncertainty.

Chapter 6

Comparison with theoretical prediction

6.1 Introduction

In chapter 2 the calculation of the prediction of the Standard Model was described. We use JetPhox to calculate the Standard Model isolated photon cross section. The result of this calculation is shown in figure 6.1. This calculation is performed for isolated photons having less then 5 GeV of transverse energy from other particles within a cone of $R = .4$ in (η, ϕ)-space of the photon.

6.2 Theory Uncertainty

To calculate the uncertainty associated with the predicted cross section, we consider the effect of changing the scale of transition between the perturbative calculation and the parton distribution function (PDF). We also consider the uncertainty of the PDF quoted by its authors. Variations to the fragmentation functions have negligible effect on the final cross section measured here. This is because, the imposition of the isolation requirement removes sensitivity to the fragmentation function. When a photon is produced in a fragmentation process it is generally collinear to the original particle. From figure 2.7, we can see that the fragmented particle is unlikely to have a large amount of the energy of the initial particle, therefore any photon produced in a fragmentation

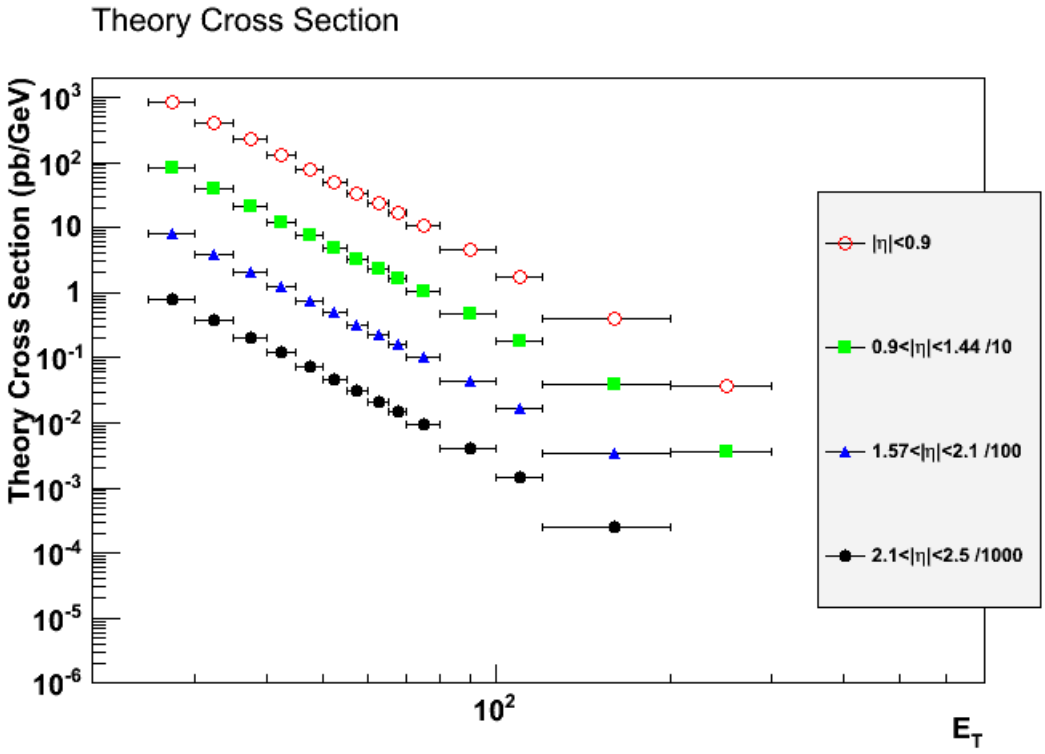

Figure 6.1: The JetPhox prediction for isolated photon cross section is plotted for each pseudorapidity region. For easy of viewing there is a scale factor the separate the different distributions.

process with energy over 25 GeV is very likely correlated with additional particles with over 5 GeV of energy. These photons would not be isolated under the criteria we use, so the contribution of fragmentation processes to this cross section is minimal. Moreover, any uncertainties in the fragmentation function used does not effect the predicted isolated photon cross section.

6.2.1 Scale Uncertainty

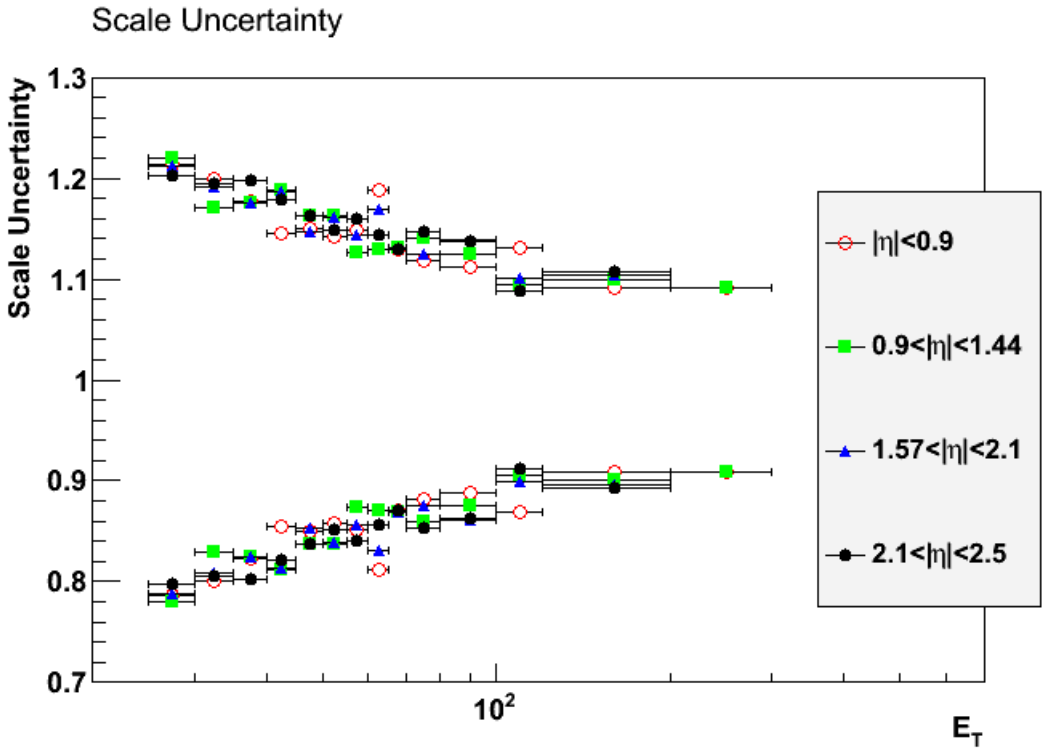

Figure 6.2: The extent of the most extreme variations in calculated values of the cross section when varying the factorization and renormalization scales between twice the transverse energy of the photon and half the transverse energy of the photon are plotted for each pseudorapidity region.

The scale where the calculation switches between the perturbative and nonperturbative is arbitrary, so too is the scale at which we renormalize the strong vertex. Because these are internal variables of the calculation with no physical significance, there is no

way to determine *a priori* the correct value to use. The natural scale in the problem is the transverse energy of the photon. We use this scale for the central value of the cross section. Then we vary the factorization scale and the renormalization scale independently between half the transverse energy of the photon and twice this energy. When varying these scales, the most extreme variation that we observe is shown in figure 6.2.

6.2.2 PDF Uncertainty

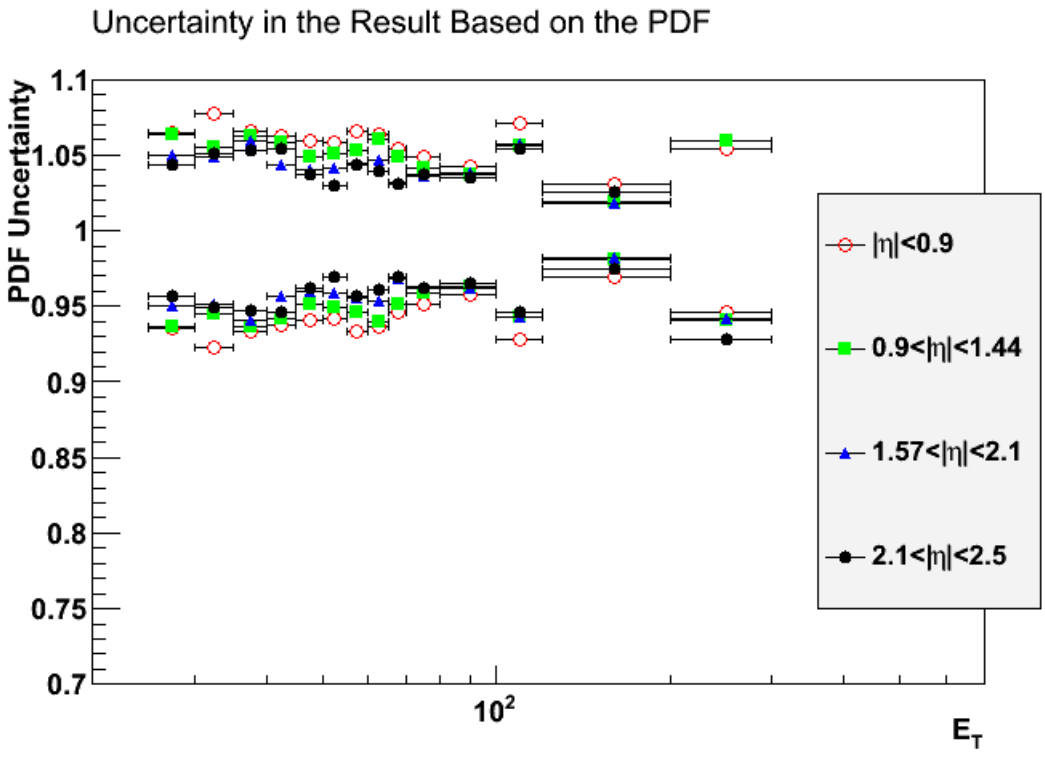

Figure 6.3: The uncertainty due to the variations in the PDF is plotted for each pseudorapidity region.

When understanding the systematic error in the predicted cross section, we also consider the effect of the uncertainty of the PDF used in the calculation. The technique of propagating the PDF uncertainties is described in [48]. The Hessian Method [49] represents the PDF as a set of independent eigenvectors, with corresponding eigenvalues representing the best fit to data. These eigenvalues are then varied up and down within

their uncertainty, this produces a new set of PDFs with two functions for each eigenvector. To calculate the uncertainty on an observable like the photon cross section, the extent of most extreme variation of the observable for the calculation done with the two PDFs resulting from the variation of each eigenvalue are added in quadrature. In this way, the uncertainty shown in figure 6.3 is taken from varying the CT10 PDF within one sigma of its central value. These variations result in changes to the overall cross section and we plot the extent of the changes when compared to the central value.

6.3 Comparison to Measurement

In figure 6.4, we compare the theoretical predictions to measurement. The theory uncertainties are centered around a central value 1.0, meaning agreement with theory, and the statistical and systematic uncertainties on the data are show around the ratio of the measurement to the data. This comparison shows the theoretical calculation to be consistent with the measured values given their respective uncertainties. Over most of the measured bins the uncertainty in the measurement is better than our ability to calculate the standard model predictions.

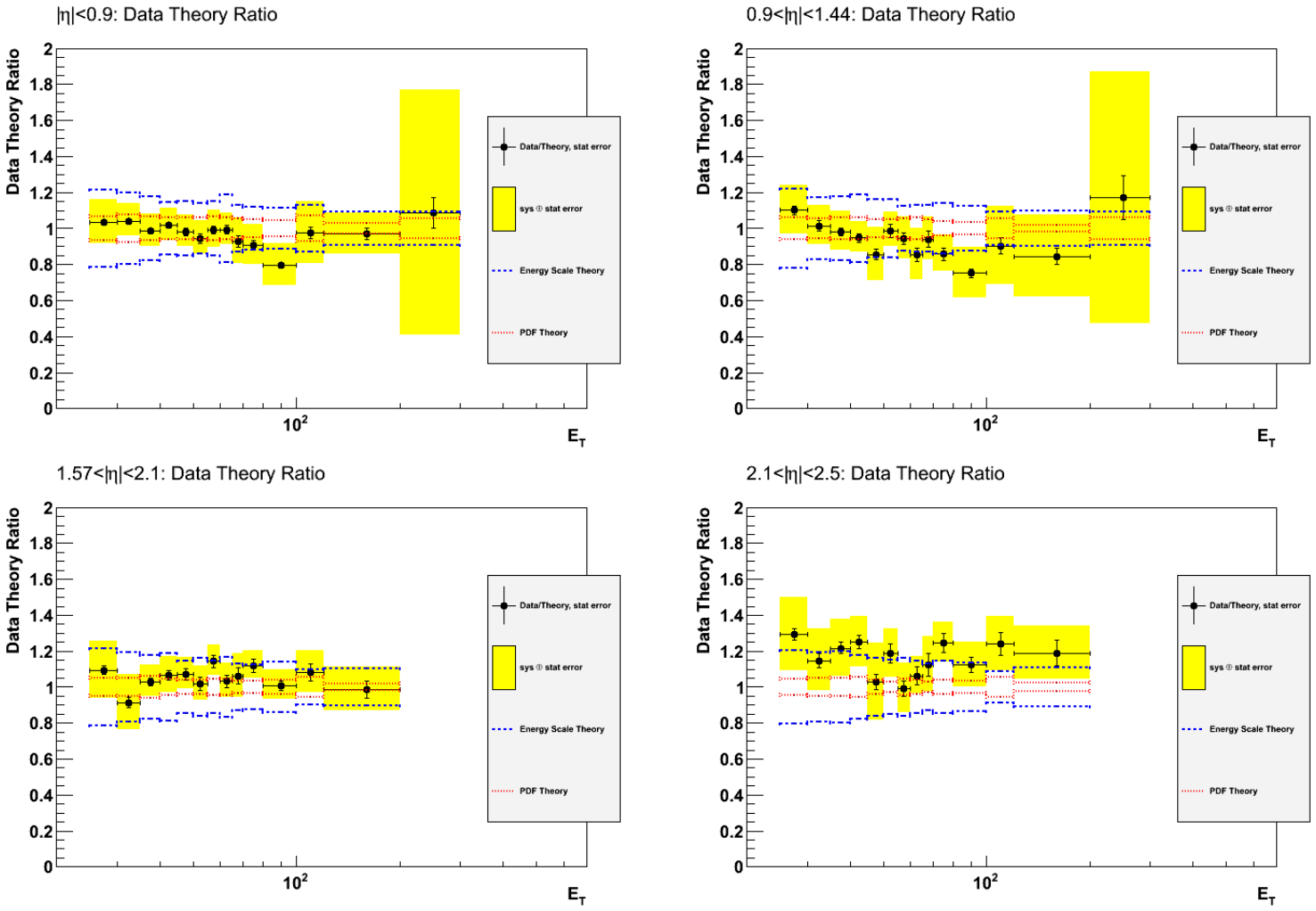

Figure 6.4: Comparison between the measurement and the theory including uncertainty on theory and the systematic and statistical uncertainties on the measurement.

Chapter 7

Conclusion

In this dissertation we have measured the isolated photon cross section at the LHC. It was found to be consistent with the state-of-the-art Standard Model calculation of next-to-leading-order direct photon production. Additionally, this theory calculation relies on input of parton distribution functions (PDFs), describing the structure of the colliding protons. This measurement, enabled by the energy of the LHC, tests the PDF in a region never before explored.

The PDF describes the probability of finding a parton with a certain fraction, x, of the overall proton momentum. This function also changes based on at what energy the proton is probed. The higher the energy of the inspection the smaller the processes that are investigated inside the proton. The energy of this probe is defined to be Q and for the case studied in this thesis is the transverse energy of the out coming photon. Figure 7.1 shows the region of the PDF that the LHC is able to probe. In the figure is also an outline of the region probed in this thesis. This shows that we have made a measurement testing an unexplored region. Because the CDF and D0 regions shown are for jets, they are not directly comparable to the measurement using photons. The D0 and CDF reach for this measurement is close to the region indicated for CDF/D0 Central Jets. Notably, the region shown for this measurement is not centered on the region indicated as the LHC reach. This is because the LHC is not running at its full energy. At its full energy the LHC will be able to probe lower x values of the PDF, and additional LHC data will also have higher Q^2 reach.

CMS has published a measurement of the isolated photon cross section using this

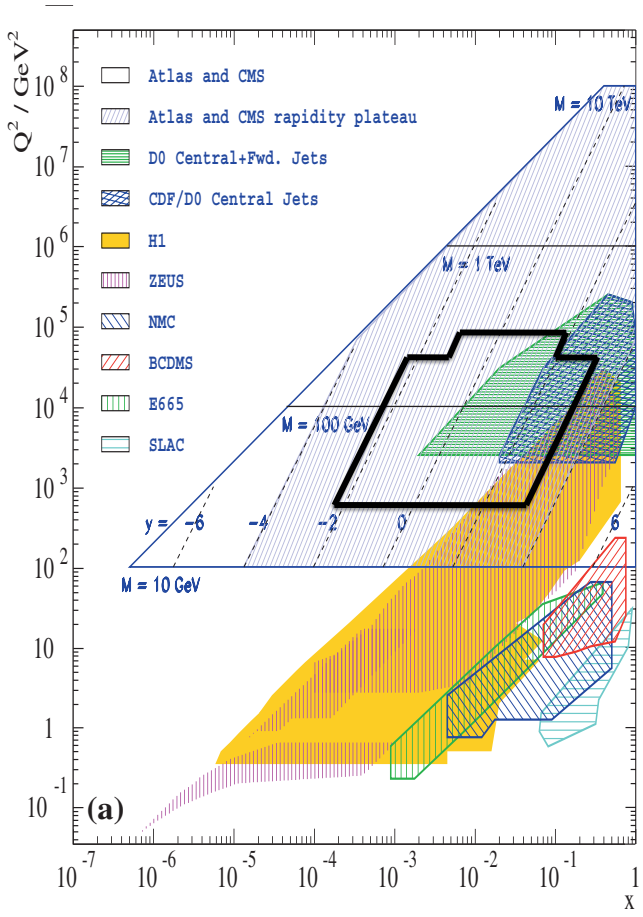

Figure 7.1: Polini plots the region of the PDF probed by past experiments compared to the region the LHC will probe [8, 139]. Overlaid on this plot is the region probed in this thesis.

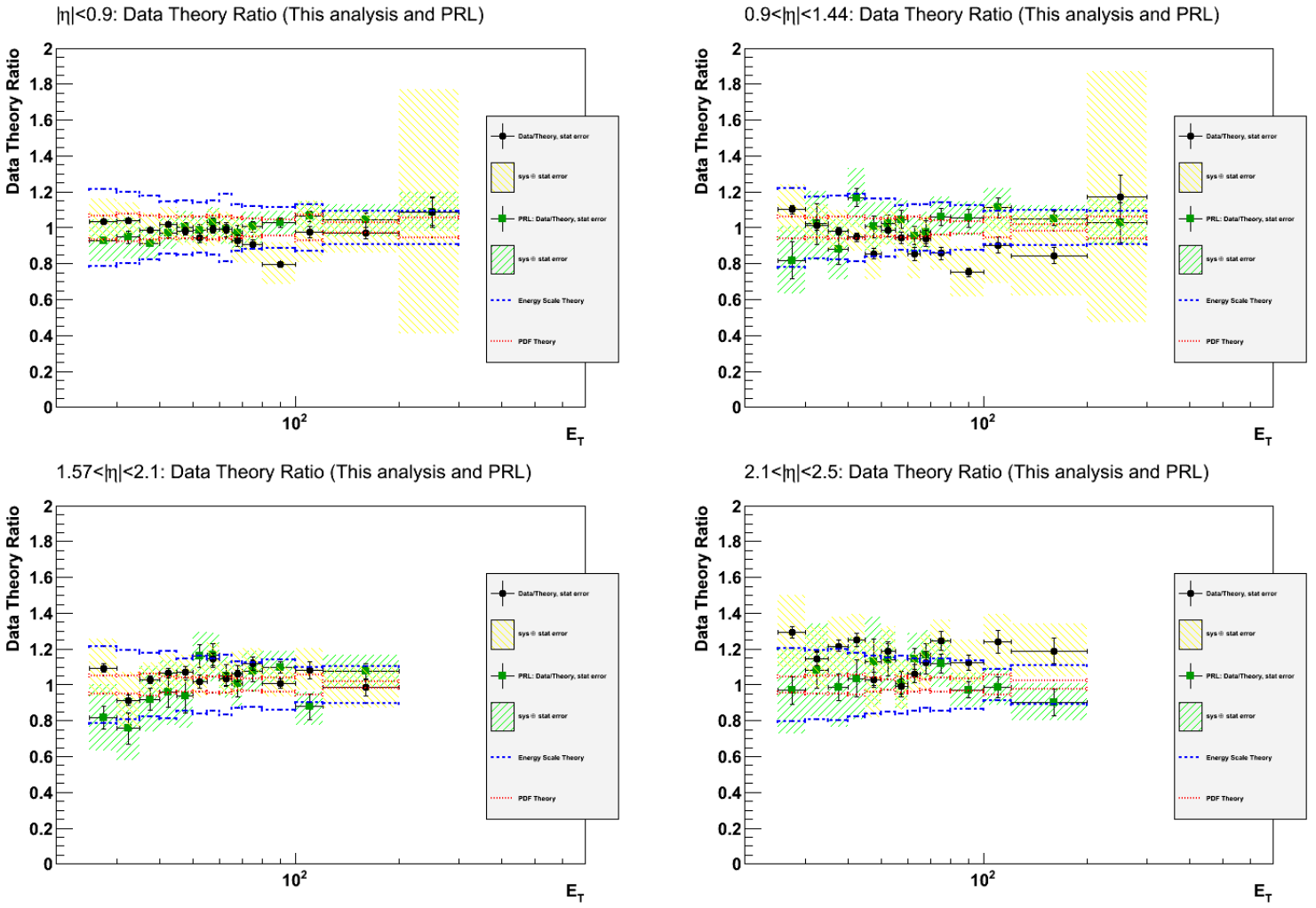

Figure 7.2: Comparison between the measurement made here, the measurement in [9] and the theory including uncertainty on theory and the systematic and statistical uncertainties on the measurements.

data [9]. At low energy, the measurement exclusively used converted photons and at high energy, it used an unbinned fit of the isolation to calculate purity. The comparison between the measurement done in this dissertation and the one in [9] is in figure 7.2. The two measurements are compatible with one another and the uncertainties are of comperable size through most of the measured values.

The measurement done in this dissertation tests our understanding of perturbative QCD and the structure of the proton. It is consistent with similar measurements done by CMS. Additionally the understanding of photons presented here provides a groundwork for further study of LHC data involving photons.

References

[1] K Nakamura and Particle Data Group. Review of particle physics. *Journal of Physics G: Nuclear and Particle Physics*, 37(7A):075021, 2010.

[2] Hung-Liang Lai et al. New parton distributions for collider physics. *Phys. Rev.*, D82:074024, 2010, 1007.2241.

[3] L. Bourhis, M. Fontannaz, and J.P. Guillet. Quarks and gluon fragmentation functions into photons. *Eur.Phys.J.*, C2:529–537, 1998, hep-ph/9704447.

[4] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[5] Bruno Borgi et al. *The CMS electromagnetic calorimeter project: Technical Design Report.* Technical Design Report CMS. CERN, Geneva, 1997.

[6] Kenneth Watson Bell, Robert M. Brown, D.J.A. Cockerill, P.S. Flower, B.W. Kennedy, et al. Vacuum phototriodes for the CMS electromagnetic calorimeter endcap. *IEEE Trans.Nucl.Sci.*, 51:2284–2287, 2004.

[7] CMS Collaboration. Electron reconstruction and identification at sqrt(s) = 7 TeV. *http://cds.cern.ch/record/1299116*, 2010.

[8] T. Čechák, L.L. Jenkovszky, and I. Karpenko. *Nuclear science and safety in Europe.* NATO security through science series: Physics and biophysics. Springer, 2006.

[9] S. Chatrchyan et al. Measurement of the differential cross section for isolated prompt photon production in *pp* collisions at 7 tev. *Phys. Rev. D*, 84:052011, Sep 2011.

[10] T. Aaltonen et al. Measurement of the Inclusive Isolated Prompt Photon Cross Section in p anti-p Collisions at s**(1/2) = 1.96-TeV using the CDF Detector. *Phys.Rev.*, D80:111106, 2009, 0910.3623.

[11] D. Acosta et al. Direct photon cross section with conversions at CDF. *Phys.Rev.*, D70:074008, 2004, hep-ex/0404022.

[12] D. Acosta et al. Comparison of the isolated direct photon cross sections in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.8-TeV and $\sqrt{s}$ = 0.63-TeV. *Phys.Rev.*, D65:112003, 2002, hep-ex/0201004.

[13] F. Abe et al. Precision measurement of the prompt photon cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ tev. *Phys. Rev. Lett.*, 73:2662–2666, Nov 1994.

[14] F. Abe et al. A Prompt photon cross-section measurement in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys.Rev.*, D48:2998–3025, 1993.

[15] F. Abe et al. Measurement of the isolated prompt photon cross section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ tev. *Phys. Rev. Lett.*, 68:2734–2738, May 1992.

[16] V.M. Abazov et al. Measurement of the isolated photon cross section in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96-TeV. *Phys.Lett.*, B639:151–158, 2006, hep-ex/0511054.

[17] B. Abbott et al. The isolated photon cross-section in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys.Rev.Lett.*, 84:2786–2791, 2000, hep-ex/9912017.

[18] S. Abachi et al. Isolated photon cross-section in the central and forward rapidity regions in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.8 TeV. *Phys.Rev.Lett.*, 77:5011–5015, 1996, hep-ex/9603006.

[19] F. Halzen and A.D. Martin. *Quarks and leptons: an introductory course in modern particle physics.* Wiley, 1984.

[20] Robert V. Harlander and William B. Kilgore. Next-to-next-to-leading order Higgs production at hadron colliders. *Phys.Rev.Lett.*, 88:201801, 2002, hep-ph/0201206.

[21] Michael Klasen. Theory of hard photoproduction. *Rev. Mod. Phys.*, 74:1221–1282, Nov 2002.

[22] S. Catani, M. Fontannaz, J.P. Guillet, and E. Pilon. Cross-section of isolated prompt photons in hadron hadron collisions. *JHEP*, 0205:028, 2002, hep-ph/0204023.

[23] Patrick Aurenche, Michel Fontannaz, Jean-Philippe Guillet, Eric Pilon, and Monique Werlen. A New critical study of photon production in hadronic collisions. *Phys.Rev.*, D73:094007, 2006, hep-ph/0602133.

[24] P. Aurenche, R. Baier, and M. Fontannaz. Prompt photon production at colliders. *Phys. Rev. D*, 42:1440–1449, Sep 1990.

[25] H. Baer, J. Ohnemus, and J. F. Owens. Next-to-leading-logarithm calculation of direct photon production. *Phys. Rev. D*, 42:61–71, Jul 1990.

[26] Edmond L. Berger and Jianwei Qiu. Calculations of prompt-photon production in qcd. *Phys. Rev. D*, 44:2002–2024, Oct 1991.

[27] L. E. Gordon and W. Vogelsang. Polarized and unpolarized isolated prompt photon production beyond the leading order. *Phys. Rev. D*, 50:1901–1916, Aug 1994.

[28] W. Vogelsang and A. Vogt. Constraints on the proton ' s gluon distribution from prompt photon production. *Nucl.Phys.*, B453:334–354, 1995, hep-ph/9505404.

[29] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys.*, B126:298, 1977.

[30] S. Catani, M. Fontannaz, and E. Pilon. Factorization and soft gluon divergences in isolated photon cross-sections. *Phys.Rev.*, D58:094025, 1998, hep-ph/9803475.

[31] R. Barate et al. Search for the standard model Higgs boson at LEP. *Phys.Lett.*, B565:61–75, 2003, hep-ex/0306033.

[32] (ed.) Evans, Lyndon and (ed.) Bryant, Philip. LHC Machine. *JINST*, 3:S08001, 2008.

[33] M. Raymond, G. Hall, J. Crooks, and M. French. The MGPA electromagnetic calorimeter readout chip for CMS. *IEEE Trans.Nucl.Sci.*, 52:756–763, 2005.

[34] CMS Collaboration. Electromagnetic calorimeter calibration with 7 TeV data. *http://cdsweb.cern.ch/record/1279350*, 2010.

[35] P. Cushman, A. Heering, and A. Ronzhin. Custom HPD readout for the CMS HCAL. *Nucl.Instrum.Meth.*, A442:289–294, 2000.

[36] CMS Collaboration. Measurement of CMS Luminosity. *http://cdsweb.cern.ch/record/1279145*, 1900.

[37] S van der Meer. Calibration of the effective beam height in the isr. oai:cds.cern.ch:296752. Technical Report CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968.

[38] CMS Collaboration. Absolute luminosity normalization. *http://cdsweb.cern.ch/record/1335668*, Mar 2011.

[39] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006, hep-ph/0603175.

[40] V.N. Ivanchenko. Geant4 toolkit for simulation of HEP experiments. *Nucl.Instrum.Meth.*, A502:666–668, 2003.

[41] Rick Field. Early LHC Underlying Event Data - Findings and Surprises. 2010, 1010.3558.

[42] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, Pavel M. Nadolsky, et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 0207:012, 2002, hep-ph/0201195.

[43] Roger J. Barlow. Extended maximum likelihood. *Nucl.Instrum.Meth.*, A297:496–506, 1990.

[44] F. James and M. Roos. Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations. *Comput. Phys. Commun.*, 10:343–367, 1975.

[45] Tim Adye. Unfolding algorithms and tests using RooUnfold. *http://arxiv.org/abs/1105.1160*, 2011, 1105.1160.

[46] G. D'Agostini. A Multidimensional unfolding method based on Bayes' theorem. *Nucl.Instrum.Meth.*, A362:487–498, 1995.

[47] Andreas Hocker and Vakhtang Kartvelishvili. SVD approach to data unfolding. *Nucl.Instrum.Meth.*, A372:469–481, 1996, hep-ph/9509307.

[48] D Bourilkov, R C Group, and M R Whalley. LHAPDF: PDF use from the Tevatron to the LHC. *http://arxiv.org/abs/hep-ph/0605240*, 2006, hep-ph/0605240.

[49] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, et al. Uncertainties of predictions from parton distribution functions. 2. The Hessian method. *Phys.Rev.*, D65:014013, 2001, hep-ph/0101032.