High-p_T Photon Production in Association with Jets at $\sqrt{s} = 7$ TeV using the Compact Muon Solenoid at the Large Hadron Collider

Michael B. Anderson

Under the supervision of Professor Sridhara Dasu

At the University of Wisconsin–Madison

We present a study of jet production in association with high p_T photons in proton-proton collisions at $\sqrt{s} = 7$ TeV from the 2010 run of the Large Hadron Collider using the Compact Muon Solenoid detecter. We show the number of events versus jet multiplicity along with ratios of $\sigma(\gamma + (\geq n)-\mathrm{jets})/\sigma(\gamma + (\geq n-1)-\mathrm{jets})$, where n is the number of jets. The reasonable agreement between data and predictions from MadGraph Monte Carlo provides confidence in using MadGraph for new physics searches that involve photons and jets.

HIGH-p_T PHOTON PRODUCTION IN ASSOCIATION WITH JETS AT $\sqrt{s} = 7$ TeV USING THE COMPACT MUON SOLENOID AT THE LARGE HADRON COLLIDER

by

Michael B. Anderson

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Physics)

at the

UNIVERSITY OF WISCONSIN – MADISON

2011

Acknowledgements

I would like to thank professor Dan McCammon for taking a chance on me and giving me a job in his X-ray astrophysics lab. He gave me, and many undergraduates, our first experience working in a real experimental physics lab. Professor Sridhara Dasu gave me a job in the Wisconsin CMS group, which allowed me the chance to contribute to the computing side of the collaboration while developing my programming skills. Kevin Flood gave me more technical guidance on using physics software than any individual has in my graduate experience, even though it was for a short duration.

My work would not have been possible without my officemate Will Maier. Thanks to Will, I regained my appreciation and passion for programming, and learned programming tips to make physics analysis much easier. In addition, the other graduate students and I are fairly knowledgable and adept with computers, but without Will Maier, Dan Bradley and Ajit Mohapatra here, doing particle physics research would have been an endless stream of frustrations with the software and computing facilities.

With Jeff Klukas I helped work on a package in the Python programming language called `rootplot`, which is available at `packages.python.org`. It has been used by me and several other graduate students to make our analysis work much easier. I am glad Jeff had the ambition to start that project, and that he let me contribute to it.

I'm happy to have been part of a large group where I haven't had to go through the tougher parts of graduate school alone. So thank you Christos Lazaridis, Kira Grogg, and Marc Weinberg for your input and support.

And finally of course, I am grateful to both my parents, Bruce and Yolanda Anderson, for their support over the years. Without their encouragement, and their appreciation for academics, I would not have come this far.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Physics have shown that atoms are made of smaller particles consisting of electrons, protons, and neutrons in the early 20th century. Later, protons and neutrons themselves were shown to be also be composite particles - that is, they had substructure. They have built a sub-atomic equivalent of a "periodic table."

The current theory of elementary particles and their interactions at high energies is an accurate, successful theory known as the Standard Model of particle physics. It is the culmination of decades of research beginning in the 1950's using particle accelerators to reach greater energies and to probe smaller length scales. The predictions of the Standard Model have been verified extensively by many experiments making high precision measurements often to 0.1% precision [1].

1.1 Standard Model

Particles interact via three forces described in the Standard Model: electromagnetism, weak force, and the strong force. Electromagnetism is the familiar force between charged particles, and in everyday life it is the cause of things such as light, radio

Electric Charge	1st Generation	2nd Generation	3rd Generation
$0e$	e-neutrino	μ-neutrino	τ-neutrino
$-1e$	electron	muon	tau

Table 1.1: Standard model table of the leptons. These particles do not interact via the strong force. Electric charge is in units of the electron charge, e.

waves, and the behavior of elements. The weak force is involved with the decay of certain radioactive elements and is involved with the process of fusion within stars. The strong force is more practically known for binding protons and neutrons together within atoms. The gravitational force is not incorporated within the framework of the Standard Model. Gravity is by far the weakest of the four fundamental forces, and though it must be taken into account over large-scales, it is negligible in individual particle interactions.

Leptons and Quarks

Within the framework of the Standard Model there are two types of particles that make up matter: leptons and quarks. Perhaps the best known lepton is the electron, found in atoms. Leptons do not interact via the strong force. In contrast, quarks do feel the strong force, and they may combine to form other particles such as protons. Both quark and lepton particles exist in three generations and are shown in tables 1.2 and 1.1. Particles within the same row of these tables have equal electric charge, but each generation is generally heavier than the last. As one moves from left to right on the table the mass of these particles increases while stability decreases, leading to shorter lifetimes. Both quarks and leptons are spin-1/2 fermions.

All the matter that naturally occurs on this planet is entirely made up of particles found only within the first generation of matter. Protons, for example, are composed of two up quarks and a down quark while neutrons contain two down quarks and an

Electric Charge	1st Generation	2nd Generation	3rd Generation
$+\frac{2}{3}e$	up	charm	top
$-\frac{1}{3}e$	down	strange	bottom

Table 1.2: Standard model table of the quarks, constituents of hadrons. Electric charge is in units of the electron charge, e.

up quark. Add in the electron, and one has all the ingredients required to make all the atoms on the periodic table of elements.

With each type of force is a an associated charge. The electromagnetic force is associated with electric charge, the weak force with weak charge (sometimes called flavor charge), while the strong force is associated with color charge. The gravitational force also has a charge referred to more commonly as mass.

The 2nd and 3rd generation of particles are only made at high energies. These particles were much more abundant when the universe was very young and hot. These particles, after creation, rapidly decay into lighter particles and thus eventually leave only particles from the first generation of matter. The table of quarks was completed with the discovery of the top quark in 1995 at Fermilab.

Bosons

The forces of nature are described by the exchange of particles known as bosons. These particles have integer spin, obey Bose-Einstein statistics, and are each associated as the carrier of a specific force. They are listed in table 1.3. The strong force is mediated by gluons, the weak force by the W and Z bosons, and the electromagnetic force by photons. The only known force not included within the Standard Model is gravity - which is difficult to study in particle physics because it is incredibly weak compared to the other three forces.

Photons, the carrier of electric charge, do not themselves carry electric charge,

and similarly the carriers of the weak force, the W and Z bosons, do not carry weak charge. This is in contrast with the carriers of the strong force, gluons, which do in fact carry color charge.

The carriers of the weak force, the W and Z bosons, were predicted by the Standard Model to exist years before they were discovered at the UA1 and UA2 experiments at CERN. Since then, studies of the lifetime of the Z boson have led to the conclusion that there is indeed likely only 3 types of neutrinos [2]. At least, if there is a fourth generation then it is unlike the other three in that it would not have a low-mass neutrino.

Force	Name	Symbol
Electromagnetic	Photon	γ
Weak	W and Z Bosons	W, Z
Strong	Gluon	g

Table 1.3: Standard model table of the bosons, the force-carriers.

1.2 Higgs Boson

A final particle predicted to exist by the Standard Model of particle physics is the Higgs Boson. It has yet to be discovered in high-energy particle experiments perhaps due to it having a relatively high mass. This boson was predicted to exist as a result of the Higgs mechanism [3, 4, 5], which is a necessary in order for particle theory to give the W and Z bosons mass while still retaining predictive power (by not containing infinities in the theoretical equations). Higgs Bosons, if they exist, would interact with particles that have mass, and should be detectable via their decay to other particles.

1.3   Quantum Electrodynamics

Electromagnetism was the first of the four known forces to be described by a quantum mechanical field theory known as Quantum Electrodynamics (QED) [6]. Quarks and leptons that interact via the electromagnetic force do so through the exchange of a photon. In QED photons cannot interact with other photons directly.

The coupling strength of the electromagnetic force, α, is a function of momentum transfer, Q^2, between interacting, charged particles as follows:

$$\alpha(Q^2) = \frac{\alpha(m^2)}{1 - \frac{\alpha(m^2)}{3\pi} \log \frac{Q^2}{m^2}} \tag{1.1}$$

where $\alpha(m^2)$ is more commonly called the *fine structure constant*. The value of $\alpha(m^2)$ has been experimentally measured and theoretically calculated to be roughly $1/137$ [7]. From this equation it can be seen that the coupling strength, α, increases with increasing momentum transfer of a particle collision.

1.4   Quantum Chromodynamics

The modern understanding of the strong force began with the discovery of quarks as the building blocks of protons and other particles known as hadrons. Interactions involving the strong force are described by a realativistic quantum mechanical field theory called quantum chromodynamics [8, 9]. Unlike QED, which has a single, electrically neutral carrier (the photon), QCD has 8 carriers of the strong force, all known as gluons, but each carrying different color charge.

Gluons mediate the strong force, and themselves also carry color charge, which causes QCD to be different from QED in a significant way: gluons can interact with other gluons directly. This self-coupling causes the potential to increase linearly with distance resulting in a strong binding of quarks within hadrons such as the proton.

The coupling strength of the strong force, α_s, is often expressed as follows:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \frac{\alpha_s(\mu^2)}{12\pi}(33 - 2n_f)\log\frac{Q^2}{\mu^2}} \tag{1.2}$$

where μ^2 is a chosen momentum scale, and n_f is the number of quarks. This shows that the strong coupling constant decreases with increasing momentum transfer of a collision. At small distance scales, the strong coupling constant decreases, and this is called asymptotic freedom.

The fact that the strong potential field energy increases with separation between colored particles allows quark/anti-quark pairs to be created out of the vacuum to form colorless, bound states. For the strong force, the only stable bound states involve 3 quarks or a quark-anti-quark pair [10]. 3-quark bound states are known as hadrons (such as protons and neutrons), whereas states of quark-anti-quark pairs are known as mesons (such as π^0 or ρ particles).

1.5 Composite Particles

In nature, only color-neutral combinations of quarks exist together. One way to achieve this is with a quark-anti-quark combination where the quark and anti-quark carry opposite color so that the color combination "sums" to zero. Another way to achieve a color-neutral combination is with three quarks where each quark carries a different color charge so that the sum is neutral. Two-quark combinations are known as mesons (such as π^0 or ρ mesons), while three-quark combinations are known as baryons (such as protons and neutrons). Figure 1.1 shows two simplified examples of composite particles and their simplified quark content. Also contained within these particles (but not shown in the figure) are gluons which mediate interactions between the quarks.
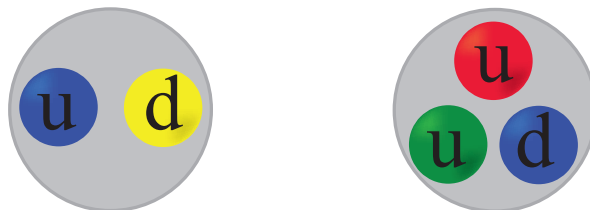
Figure 1.1: Examples of color-neutral combinations of quarks. A quark-anti-quark combination, the pion, π^0, is shown (left). A three-quark combination, the proton, is also shown (right).

When a composite particle collides with another composite particle - such as in a proton-proton collision - it results in the collision of a quark (or gluon) from one of the particles with a quark (or gluon) from the other. At high enough collision energy, this can result in the disintegration of the composite particle and is then called an "inelastic collision".

1.6 Cross Sections

The measure of the probability for certain processes to occur during a collision of particles is the cross section, σ, and is usually measured in units of barns [b] ($1\text{b} = 10^{-28}\text{m}^2$). Beginning with an incident beams of particles, and measuring outgoing, scattered particles, we can measure the probability of certain processes to have occurred. This probability, the cross-section, is related to the rate of the process and the beam flux as follows:

$$\frac{d\sigma}{d\Omega} = \frac{\text{Scattered flux/solid angle}}{\text{Incident flux/solid angle}} \tag{1.3}$$

Integrating this over the solid angle, $d\Omega$, we obtain the total cross section. The rate of a certain interaction depends on the interaction probability derived from the

dynamics of underlying theory, QED and QCD. The reaction rate, W, is defined as:

$$W = \frac{2\pi}{\hbar}|M_{if}|^2 \rho_f \tag{1.4}$$

Where M_{if} is the matrix element relating the initial and final-states, and ρ_f is the phase space available to the final state. Matrix elements can be derived using Feynman rules from the underlying theory (see the following section 1.8).

When the colliding beams of particles is made of up hadrons, which are composite particles, there is the added complication of needing to define parton density functions (PDFs). Hadrons are made up of particles such as quarks and gluons, and it is these particles which actually collide. PDFs cannot be calculated from theory, but are instead determined by fits to measurements of structure functions. Various groups have released experimentally determined parton distribution functions such as the Coordinated Theoretical-Experimental Project on QCD (CTEQ) [11].

The cross section for processes where two incoming protons producing an outgoing photon plus anything else ($P_1 + P_2 \rightarrow \gamma + $anything) is given by the following expression from QCD:

$$\sigma(P_1 + P_2 \rightarrow \gamma + \text{anything}) = \sum_{i,j,k} \int dx_1 dx_2 dx_3 F_{i/1}(x_1) F_{j/2}(x_2) D_{k/3}(x_3) \hat{\sigma}_{ij}(x_1, x_2)$$

$$\tag{1.5}$$

Where P_1 and P_2 are the incoming protons, from which a parton from each actually contributes to the hard scattering. The momentum fraction of the interacting partons are $x_{1,2} = p_{1,2}/P_{1,2}$. The sum is over all the parton types in the protons, and the integral is over all the momentum fractions. The parton distribution functions give the probability of getting a certain kind of parton with momentum fraction x. These PDFs are denoted by $F_{1,2}$, and they depends on the factorization scale μ_F chosen. The fragmentation function, D, describes the probability of getting a specific particle
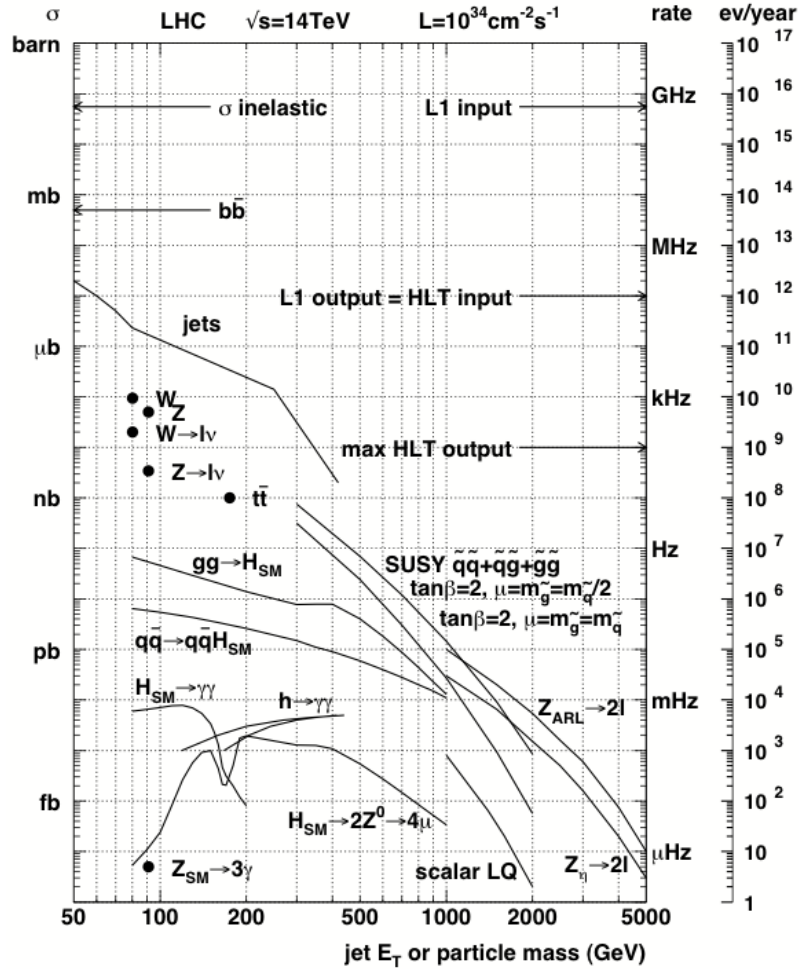
Figure 1.2: Production cross section for various physics processes from pp collisions [12]. The rate of events per year is shown on the far right axis for design luminosity of the Large Hadron Collider.

with some fraction x of the original momentum. The final term, σ_{ij} is the cross section for the parton interaction.

1.7 Parton Distribution Functions

The parton distribution functions (PDFs) used in 1.5 describe content of protons by the momentum fraction carried by quarks and gluons. Figure 1.3, shows example

parton distribution functions from the CTEQ collaboration.
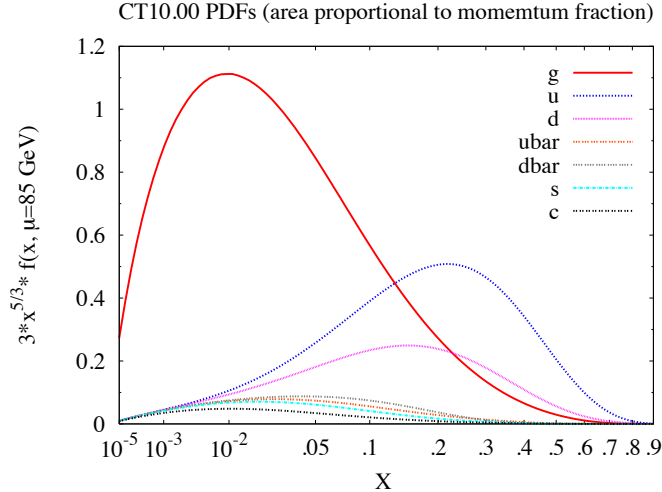
Figure 1.3: Parton distribution functions in the central fit from the CTEQ collaboration [11] for $\mu = 85\,\text{GeV}$. The variable x is the ratio of the parton's momentum to the total momentum of the hadron it is in.

1.8   Parton-Level Interaction Cross Sections

Equation 1.5 contains a term which which depends on parton-level interactions. These can be determined by creating what are referred to as Feynman diagrams which show incoming particles coming in, interacting, and creating outgoing particles. Examples of Feynman diagrams are shown in figure 1.5. Using Feynman rules, mathematical terms can be derived and used in the calculation of the parton-level cross section.

For the quark-gluon scattering processes shown in figure 1.5 these parton-interaction cross sections are:

$$\frac{d\sigma}{d\hat{t}} = \pi\alpha\alpha_s e_q^2 \left(-\frac{1}{3s^2}\right)\left(\frac{\hat{u}}{\hat{s}} + \frac{\hat{s}}{\hat{u}}\right) \tag{1.6}$$

and for $q\bar{q}$ annihilation:

$$\frac{d\sigma}{d\hat{t}} = \pi\alpha\alpha_s e_q^2 \left(-\frac{8}{9s^2}\right)\left(\frac{\hat{u}}{\hat{t}} + \frac{\hat{t}}{\hat{u}}\right) \tag{1.7}$$

where e_q is the charge of the quark, and $\hat{s}$, $\hat{t}$, $\hat{u}$ are Mandelstam variables [13].

The simplest Feynman diagrams that one can draw connecting incoming and outgoing particles is referred to as the lowest-order (LO) terms. For more accuracy, more complicated Feynman diagrams can be drawn, which allows calculation of next-to-leading-order (NLO) terms. Such Feynman diagrams will include loops, or more vertices, such as in figure 1.6. The number of parton-level diagrams increases rapidly with the number of vertices drawn, however, the contribution to the total cross section also decreases with each additional vertex.

1.9   Direct Photons

Collisions of protons that produce high-energy, direct photons (as opposed to those created in the decay of mesons) are useful for studying QCD over several orders of magnitude in hard-scattering energy. The energy and position of photons can be measured accurately with detectors known as electromagnetic calorimeters. A schematic of a proton-proton collision giving rise to a direct photon figure 1.4. Such events also typically produce a relatively narrow collection of hadrons in the opposite direction. These collections of hadrons are referred to as *jets*.

Shown in figure 1.5 are Feynman diagrams showing leading-order processes that contribute to production of direct photons. The coupling of photons to quarks is well-understood, and measurements of direct photons can provide insight to the gluon distribution inside protons. As an be seen in 1.5, gluons appear in all the leading-order processes producing direct photons.

Experiments which collide hadrons such as at the Tevatron and now at the LHC do not generally measure inclusive photon production. This is because the production rate of background photons from meson decay is several orders of magnitude higher.

This background arises from the decay of mesons such as the π^0 and ρ. Background photons are contained within jets, and so direct photons are identified by requiring that they are isolated from other hadrons. To achieve this, it is typically required that the sum of energy within a certain radius around the photon be below some value. This reduces the background, but has the side-effect of reducing the contribution to direct photons from fragmentation.
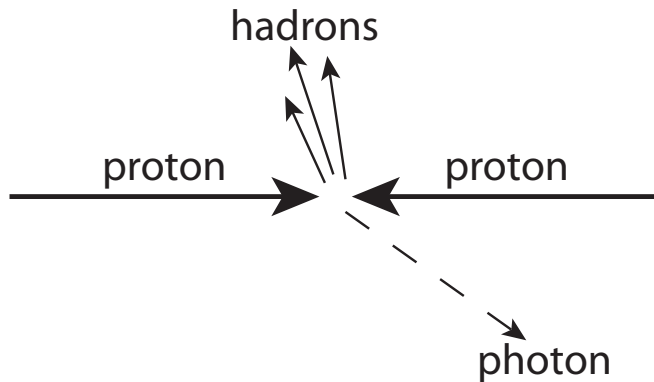
Figure 1.4: Schematic of a collision between two high-energy protons traveling in opposite directions. This collision occasionally creates a high-energy photon with hadrons produced in roughly the opposite direction so as to conserve transverse momentum.

Several types of hard-scattering interactions contribute to create direct photons. Feynman diagrams of leading order interactions are shown in 1.5. These include gluon compton scattering and quark anti-quark annihilation.

The inclusive cross section for the production of a non-isolated photon with momentum p^γ is given by [14]:

$$\frac{d\sigma}{dp_\mathrm{T}d\eta} = \sum_i \int_0^1 \left(\frac{dz}{z}\hat{\sigma}^a(p^\gamma/z; \mu, M, M_F)D_a^\gamma(z; M_F) + \hat{\sigma}^\gamma(p^\gamma; \mu, M, M_F) \right) \quad (1.8)$$

The first term corresponds to 'fragmentation' photons, while the second term corresponds to 'direct' photons. The term $\hat{\sigma}^a$ describes the production of a parton (quark or gluon) in the hard interaction, and D_a^γ is the fragmentation function for that parton

into a photon. The term $\hat{\sigma}^\gamma$ is the direct photon production by the hard interaction, and so does not contain a fragmentation function.

Both terms depend on the renormalization scale μ, the factorization scale M of the initial-state parton distributions and on the factorization scale M_F of the photon fragmentation function.

The cross sections for the terms $\hat{\sigma}^a$ and $\hat{\sigma}^\gamma$ are given to next-to-leading-order in α_s:

$$\hat{\sigma}^a(p; \mu, M, M_F) = \left(\frac{\alpha_s(\mu)}{\phi}\right)^2 \sigma^a_{\text{Born}}(p; M) + \left(\frac{\alpha_s(\mu)}{\phi}\right)^3 \sigma^a_{\text{HO}}(p; \mu, M, M_F) \qquad (1.9)$$

$$\hat{\sigma}^\gamma(p; \mu, M, M_F) = \left(\frac{\alpha_s(\mu)}{\phi}\right) \sigma^\gamma_{\text{Born}}(p; M) + \left(\frac{\alpha_s(\mu)}{\phi}\right)^2 \sigma^\gamma_{\text{HO}}(p; \mu, M, M_F) \qquad (1.10)$$

The expressions for σ_{Born} and σ_{HO} (higher-order corrections), as well as further details can be found in [15, 16, 17], where NLO calculations are done and extended to isolated photons.
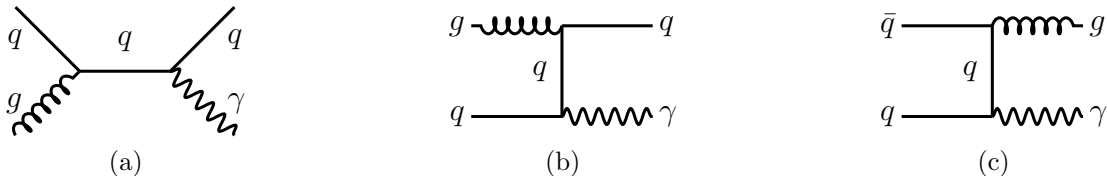
Figure 1.5: Leading order Feynman diagrams contributing to the γ+jet cross section. Figures (a) and (b) are quark-gluon compton scattering, (c) is quark anti-quark annihilation.

1.10 Jets

In contrast with the photon from the hard scattering, the outgoing parton from the hard scattering cannot be detected directly. Due to color-confinement, these particles
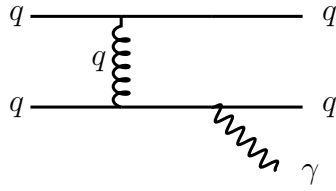
Figure 1.6: Leading order Feynman diagram contributing to the bremstrahlung photons.

Figure 1.7: Leading order Feynman diagrams for quark-gluon scattering (a) and quark-antiquark annihilation (b). These processes are backgrounds for γ+jet events.

cannot exist by themselves, and so they combine with quarks and antiquarks from the vacuum to form large collections of hadrons commonly known as *jets*. This formation of a collection of hadrons from outgoing quarks or gluons from a hard scattering interaction is a process known as hadronization. Essentially, QCD perturbation theory is valid at short distances, but at long distances becomes a stronger and stronger interaction causing the breakdown of pertubation theory. Different phenomenological models exist for hadronization, and are described later on in section 4.2.

Essentially, quarks and gluons from a hard scattering event produce showers of hadrons by branching such as $q \rightarrow q + g$, $g \rightarrow g + g$, and $g \rightarrow q + \bar{q}$. According to QCD, the branching of gluons is favored over the branching of quarks, and so gluon jets are expected to contain more particles (higher multiplicity) than quark jets [18].

Particle detectors see jets as cone-like formation of hadrons. Unlike photons, which are detected using one type of particle detector (electromagnetic calorimeters), jets are found using information from at least three detectors including the tracking

system, electromagnetic calorimeter, and hadronic calorimeter. This is discussed in further detail in 5.3.

1.11 Photon production in association with jets

For several searches of new physics at hadron colliders, signals are an excess of events in broader distributions where backgrounds are events with photons and jets. Processes where jets are produced with a high-p_T photon are a background for physics searches beyond the Standard Model especially when produced in association with missing transverse energy, E_T^{miss}. Thus searches for new physics require a good understanding of the QCD backgrounds. In addition, the study of jet production in association with high-p_T photons provides a test of perturbative quantum chromodynamics (pQCD) predictions and allows for the possibility to constrain parton distribution functions of the proton [11]. Theoretical predictions for the cross section of isolated photon production exist [14].

Production of vector bosons in association with jets has been studied for the W and Z. Next-to-leading-order (NLO) predictions for W+jets events at the Large Hadron Collider are available for production of up to W+4jets [19]. For Z+jets, NLO predictions at the Tevatron were made up to Z+3jets [20]. Production of those vector bosons with jets are a large background for new physics as well as top quark studies. The ATLAS collaboration measured W+jets production [21], as well as Z+jets production [22]. The CMS collaboration made measurements of W+jets and Z+jets as well [23].

This thesis shows the first analysis of photon production in association with jets with the CMS detector.

1.12 Summary

Photon production in hadronic collisions has been studied for decades. High transverse momentum photons are produced primarily by point-like particles within hadrons. Measurements of direct photons is important for investigating the structure of hadrons. Specifically, these measurements can can provide insight on the the quark/gluon content of protons.

The primary obstacle to overcome when measuring the photon+jet cross section is finding and removing the events where a jet leaves an energy deposit in the detector similar to a direct photon. Although most jets hadronize into many charged and uncharged particles, a few (on the order of 1/1000 jets) hadronize in such a way that a single π^0 or similar meson carries away most of the energy and decays into two photons. These two photons may then deposit their energy in such a way to appear similar to a single photon.

This thesis is the first measurement of high-$p_{\rm T}$ photon production in association with jets with the CMS detector at the Large Hadron Collider. The final results are ratios of cross-sections of events with a photon and jets in the final state as a function of jet multiplicity.

Chapter 2

Previous Studies of Direct Photons

Several experiments have measured inclusive and isolated photon cross sections from hadronic collisions at various center-of-mass energies. The dominant production method of high energy photons at hadron colliders is the reaction $gq \rightarrow \gamma q$ [24]. Photon production in high energy collisions still needs better predictions as data cannot be simultaneously fitted with a single set of scales and structure functions [25]. Whereas good agreement between data and theoretical predictions has been shown by recent measurements at the Tevatron, RHIC, HERA, and the LHC. These experiments measured photon production for various ranges of pseudorapidity (defined as $\eta = -\ln\left(\tan\left(\theta/2\right)\right)$ where θ is the angle between the photon's momentum and the beam axis), and for different ranges of photon transverse energy, E_{T}.

2.1   CDF

The CDF Collaboration at the Tevatron has measured the isolated, direct photon cross section $(d^2\sigma/dp_{\mathrm{T}}^{\gamma}d\eta^{\gamma})$ from $p\bar{p}$ collisions for two center-of-mass energies: 0.63 and 1.8 TeV [26]. This includes photons with $|\eta| < 0.9$ and $10\,\mathrm{GeV} < p_{\mathrm{T}} < 120\,\mathrm{GeV}$.

Data is compared to a next-to-leading order QCD calculation [24] derived using the CTEQ5M parton distributions [27] and using p_T for the renormalization, factorization, and fragmentation scales. A difference between data and NLO QCD theory was observed when plotted as a function of photon p_T, and is particularly pronounced at low p_T. Data and prediction are shown in figure 2.1. The purity of the photon sample is lowest at low photon p_T and increases with photon p_T.

A possible reason put forth by CDF for the observed discrepancy is the lack of a complete description of the initial-state parton shower in the NLO QCD calculation. Such a shower can give a recoil effect to the photon+jet system.
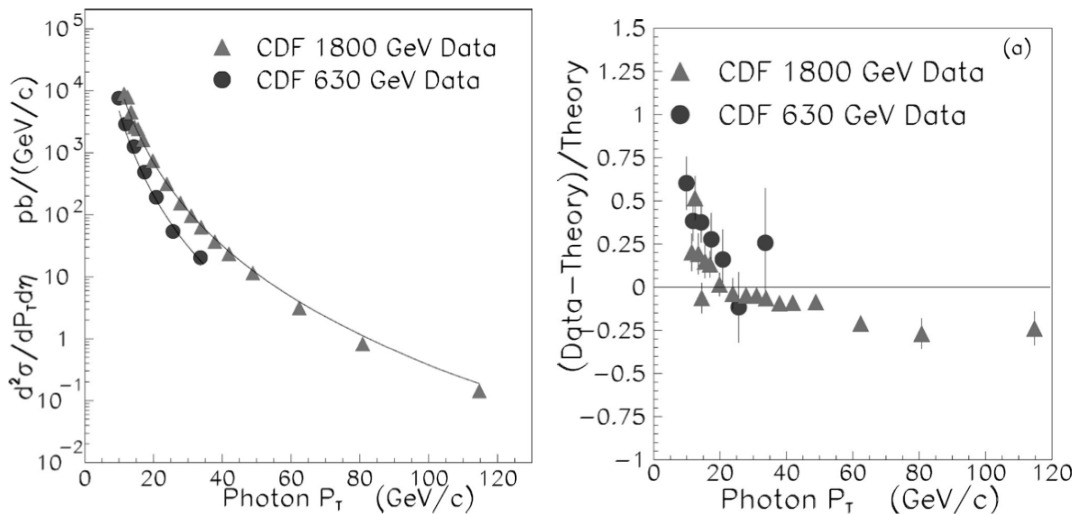
Figure 2.1: Inclusive photon cross section at two center-of-mass energies compared to NLO QCD theory (left) with data from the CDF detector [26]. Comparison with NLO QCD is shown as a ratio (right).

2.2   DZERO

The DØ Collaboration at the Tevatron has studied the process $p\bar{p} \rightarrow \gamma + \text{jet} + X$ using $1.0\,\text{fb}^{-1}$ of data [28]. Photons in the range $30 < p_T^\gamma < 400\,\text{GeV}$ and with central

rapidity $|y^\gamma| < 1.0$ were used. The jet was required to be either central $(|y^{\text{jet}}| < 0.8)$ or forward $(1.5 < |y^{\text{jet}}| < 2.5)$.

Four different rapidity configurations between the photon and leading jet were used for measuring the cross section: central jets with $y^\gamma \cdot y^{\text{jet}} > 0$ or $y^\gamma \cdot y^{\text{jet}} < 0$, and forward jets with $y^\gamma \cdot y^{\text{jet}} > 0$ or $y^\gamma \cdot y^{\text{jet}} < 0$. By taking ratios, factors such as integrated luminosity cancel out leading to lower experimental systematic uncertainty. The total x and $Q^2 (\equiv (p_{\text{T}}^\gamma)^2)$ covered by these measurements are $0.007 < x < 0.8$ and $900 < Q^2 < 1.6 \times 10^5 \, \text{GeV}^2$.

Data is compared to next-to-leading order QCD predictions obtained using JET-PHOX, with CTEQ6.5M parton distribution functions. Next-to-leading order QCD predictions failed to describe the shape of the p_{T} dependence of the cross section. Different values for fragmentation, factorization and renormalization scales were tried, $\mu_{R,F,f} = \mu_0, 0.5\mu_0$, and $2\mu_0$ where $\mu_0 = p_{\text{T}}^\gamma f(y^*)$. In addition, different PDFs were also used: MRST04 [29], Alekhin02 [30], and Zeus05 [31]. Despite trying different PDFs and variations of scale, the next-to-leading order QCD predictions failed to describe the shape of the p_{T} dependence of the cross section. This showed a need for improved theoretical predictions for γ+jet production in hadron collisions.

2.3 RHIC

The PHENIX collaboration has collected data to measure both the inclusive [32] and isolated [33] rate of photon production at center-of-mass 200 GeV at the RHIC collider. It was the first collaboration to measure both, and has helped increase the understanding of the role of the fragmentation component in prompt photon production. The isolation method enhanced signal purity. These data cover an energy range of $4 < p_{\text{T}}^\gamma < 17 \, \text{GeV}$. Photons were detected in the pseudorapidity range of
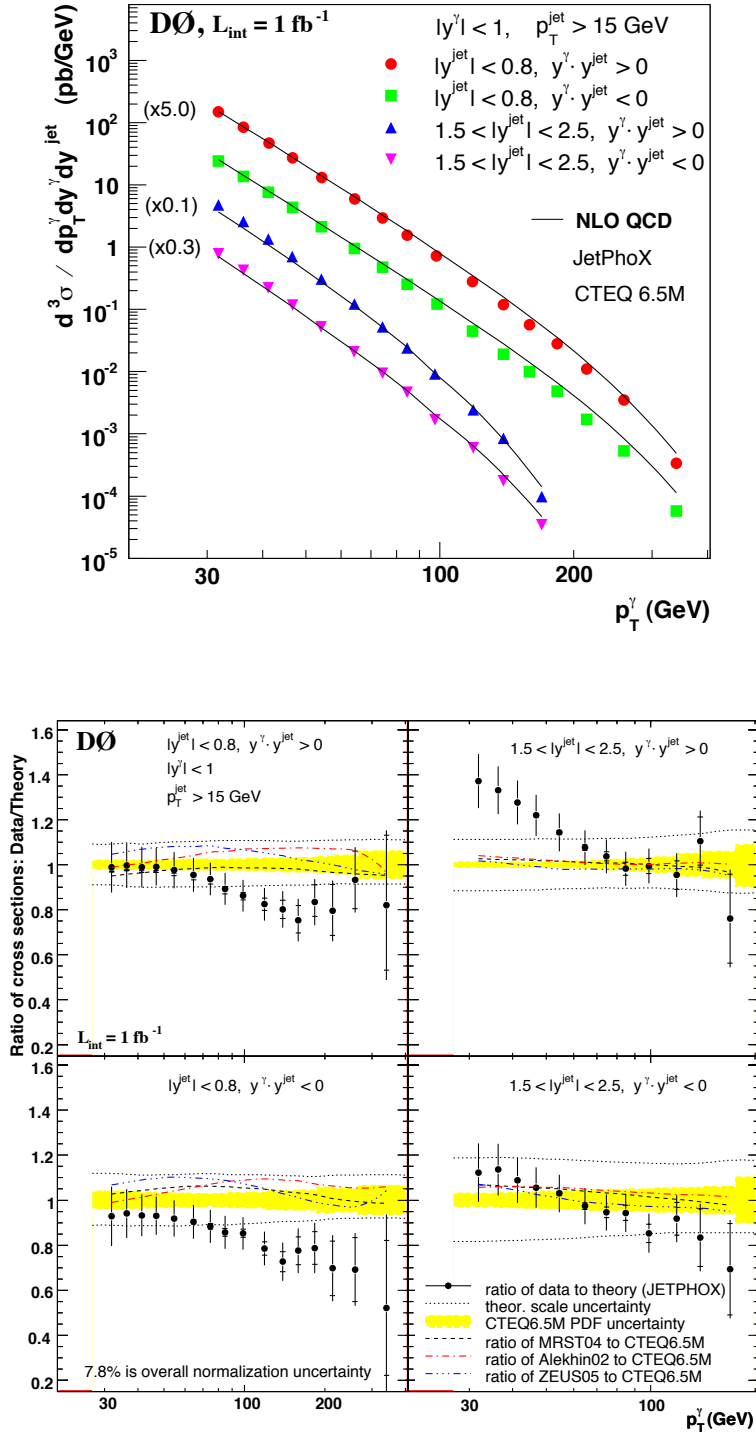
Figure 2.2: Measured differential cross section for $p\bar{p} \to \gamma + \text{jet} + X$ for for rapidity ranges (left) by the DØ collaboration [28]. Ratios with NLO QCD predictions also shown (right).

$-0.3 \leq \eta^\gamma \leq 0.3$ and phi range of $-0.73 \leq \phi \leq 0.73$. Hadronic energy in a cone of radius 0.5 was included only in a narrow region $-0.35 \leq \eta^\gamma \leq 0.35$, $-\pi/4 \leq \phi \leq \pi/4$. There, the hadronic energy in the cone around the photon candidate was required to be less than 10% of the photon momentum. The experimental result was compared to NLO pQCD calculations [16]. These NLO pQCD calculations used the CTEQ6 parton distribution functions and the GRV parton-to-photon fragmentation function [34]. The NLO pQCD calculation agrees with measurement well.
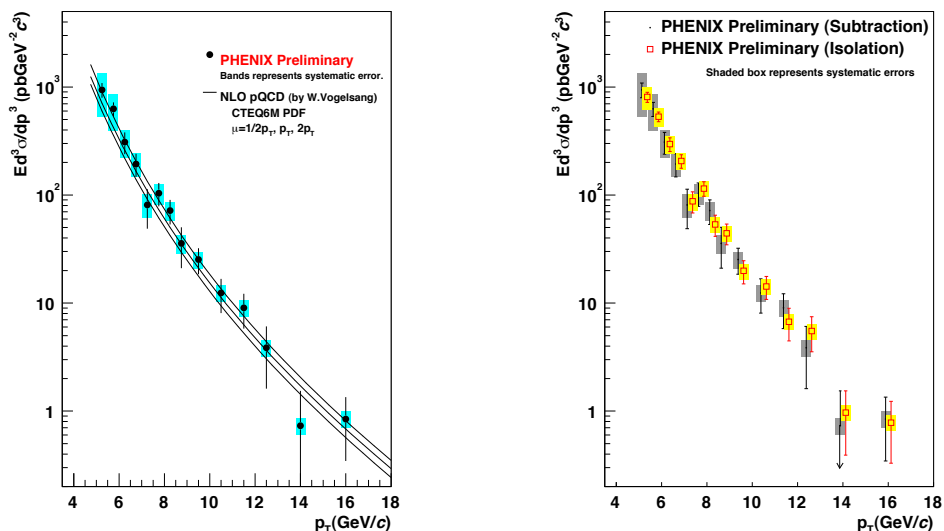
Figure 2.3: Prompt photon cross section as a function of p_T as measured by the PHENIX experiment [32, 33]. Curves are NLO pQCD prediction with 3 different scales.

2.4   H1

The H1 Collaboration at the HERA accelerator has studied prompt photon production from electron-proton scattering [35]. Low energy photons in the range $6 < E_T^\gamma < 15$ GeV were measured in pseudorapidity range of $-1.0 < \eta^\gamma < 2.4$. Cross sections were also measured as a function of transverse energy and pseudorapidity of the jet.

The data sample corresponds to an integrated luminosity of $340\,\text{pb}^{-1}$. The theory predictions used for comparison were a QCD calculation based on the collinear factorization in NLO (FGH) [36], and a QCD calculation based on k_T factorization approach (LZ) [37]. The measured total cross section was higher than both theory predictions by over 15%. Differential cross section comparisons between data and theory are shown in figure 2.4.
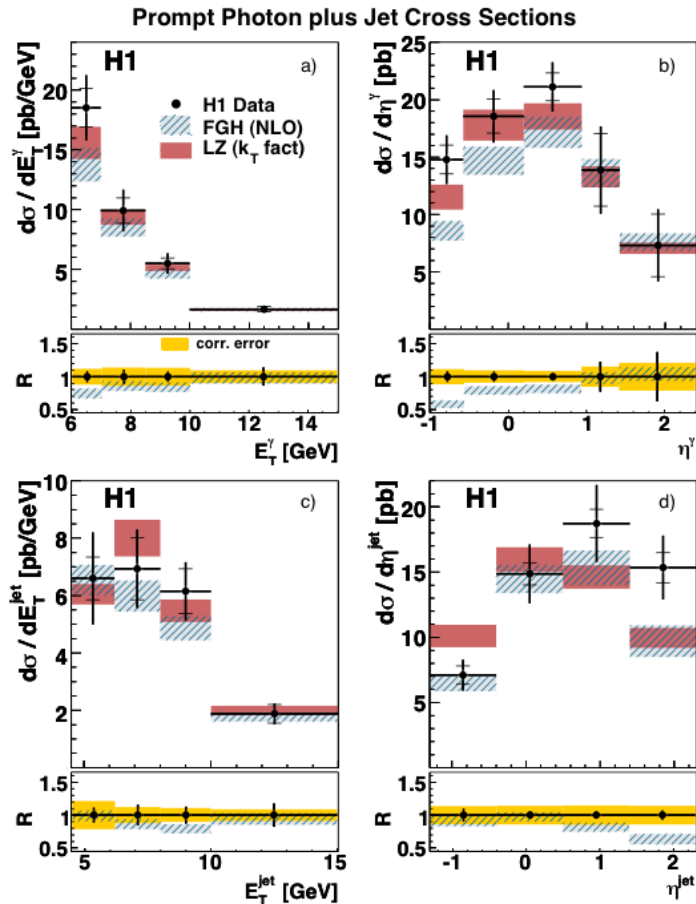
Figure 2.4: Differential cross sections for prompt photon plus jet events as measured by the H1 experiment [35]. Data is compared to NLO QCD predictions from two methods, FGH [36] and LZ [37].

2.5 ZEUS

The production of prompt photons with an accompanying jet, was been studied in ep collisions at a centre-of-mass energy of 318 GeV with the ZEUS detector at HERA using an integrated luminosity of 77 pb^{-1} [38]. Cross sections were measured when the E_T of the photon was larger than 5 GeV and the jet E_T was larger than 6 GeV. The differential γ+jet cross sections were reconstructed as functions of the transverse energy, pseudorapidity and x_γ, the fraction of the incoming photon momentum taken by the photon-jet system.

The theoretical predictions for the γ+jet differential cross sections were performed at NLO with additional higher-order terms [39]. The renormalisation and factorisation scales were set to $\mu_R = \mu_F = E_T^\gamma$.

The prompt-photon data disagree with the available MC predictions which predict a less steep rise of the cross sections with decreasing E_T^γ. The discrepancy was diminished when compared to the NLO calculations.

The ZEUS detector also studied the production of prompt photons with an accompanying jet. The effective transverse momentum of partons, k_T, in the proton were measured in the γp center-of-mass energy range of 134 to 251 GeV. It used a total integrated luminosity of 38.6 pb^{-1} [40]. Their results were in agreement with an observed trend in hadron-hadron scattering for k_T to rise with increasing energy.

2.6 ATLAS

The ATLAS Collaboration at the LHC measured the production of isolated prompt photons from pp collisions at a center-of-mass energy of $\sqrt{s} = 7\,\mathrm{TeV}$ [41]. Photons were measured within the transverse energy range of $15 \leq E_T^\gamma \leq 100\,\mathrm{GeV}$, and the

24
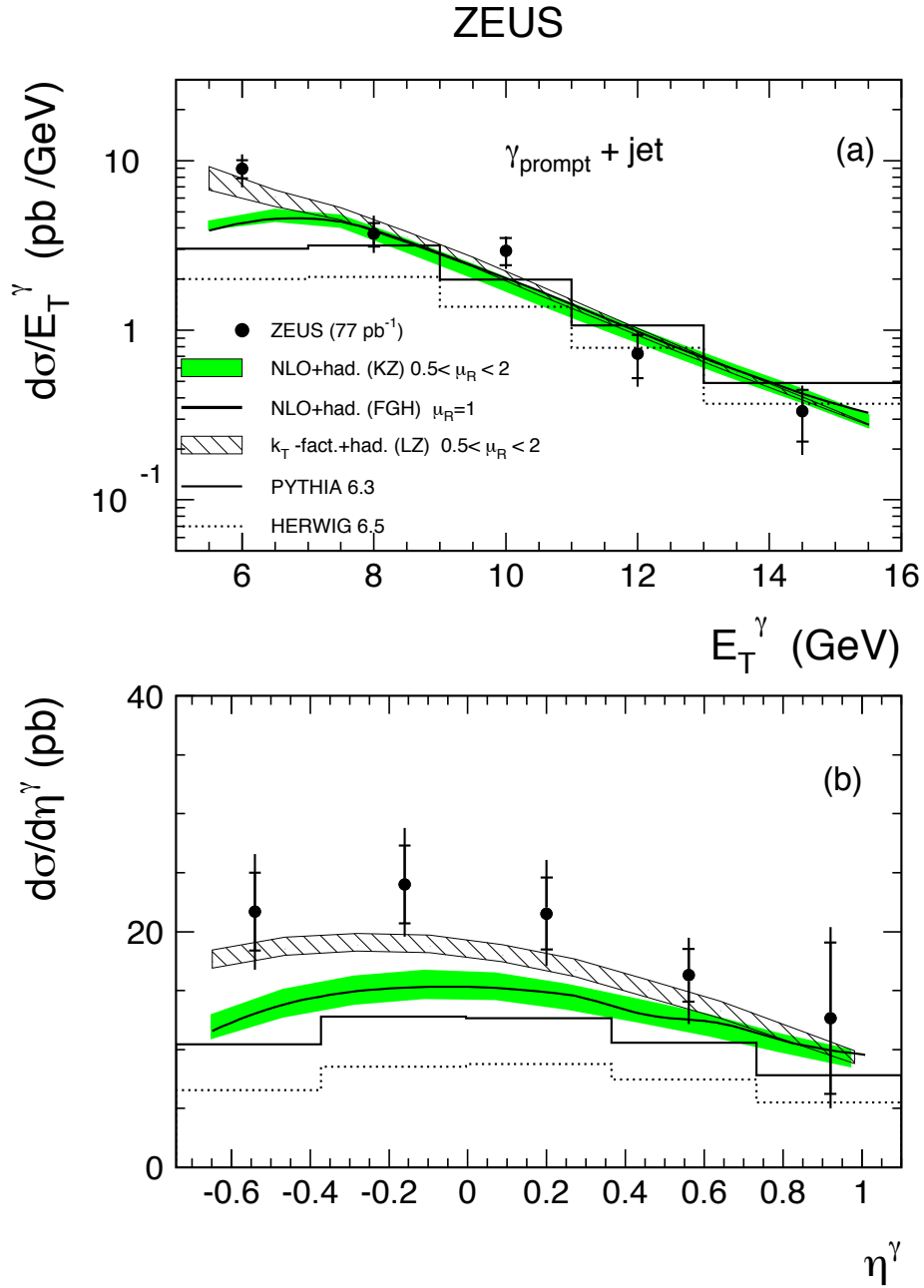
Figure 2.5: Differential cross sections for prompt photon plus jet events as measured by the Zeus experiment [38]. The theoretical QCD predictions are also shown from the methods KZ, FGH [36] and LZ [37].

psedorapidity range $|\eta^\gamma| < 1.37$ and $1.52 \leq |\eta^\gamma| \leq 1.81$. The total amount of data used for their analysis is 880 nb^{-1} collected in 2010.

Their measurement of prompt photons includes both "direct" photons (which come from the hard scattering) as well as "fragmentation" photons, which are the result of the fragmentation from a high-p_T parton from the hard scattering.
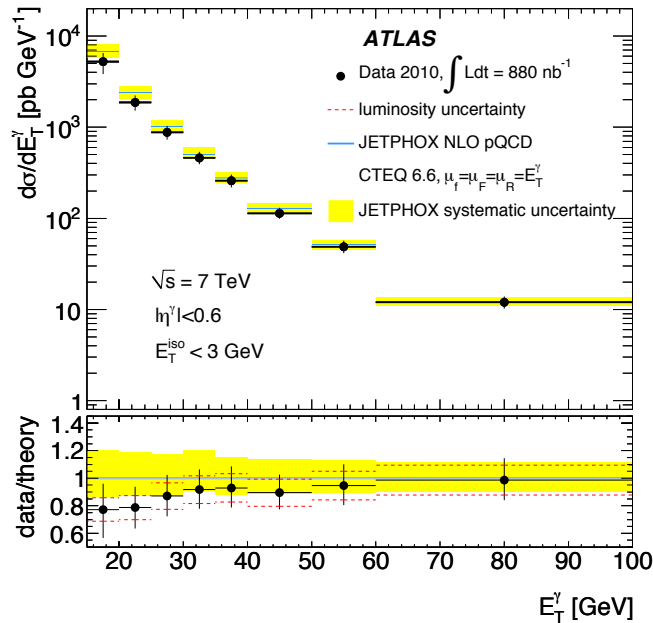
Figure 2.6: Measured by ATLAS (dots) [41] and expected (full line) differential cross section for inclusive prompt photon production. Ratio between measurement and theory is shown in bottom plot.

The predicted isolated prompt photon production cross section as a function of photon transverse energy is calculated with the Monte Carlo program JETPHOX. This program calculates contributions from both direct and fragmentation photons using full next-to-leading order QCD calculations. The renormalization (μ_r), factorization (μ_F), and fragmentation (μ_f) scales are set to the photon transverse energy. They found in general that the measured cross section agrees with theoretical predictions for photons with transverse energy $E_\mathrm{T} > 25\,\mathrm{GeV}$. At lower E_T, where values

of $x_\mathrm{T} = 2E_\mathrm{T}^\gamma/\sqrt{s}$, the NLO QCD predictions are less accurate. There, the measured cross section is found to be lower than the predictions.

2.7 CMS

Inclusive production of isolated, prompt photons has been studied in pp collisions at $\sqrt{7} =$ TeV with the CMS detector at the LHC [42]. The data sample corresponds to an integrated luminosity of 2.9 pb^{-1}. The differential cross section as a function of the photon transverse energy $(d^2\sigma/dE_\mathrm{T}^\gamma d\eta^\gamma)$ was measured. Photons were required to have a pseudorapidity $|\eta| < 1.45$ and $E_\mathrm{T} > 21 GeV$, covering the kinematic region $0.006 < x_\mathrm{T} < 0.086$. Next-to-leading-order pQCD predictions were made using JETPHOX 1.1 [14] with CT10 PDFs [43] and BFG set II of fragmentation functions [44]. The renormalization, factorization and fragmentation scales were set to E_T^γ. The measured differential cross section was found to be in agreement with the next-to-leading-order perturbative QCD calculations and is shown in figure 2.7.
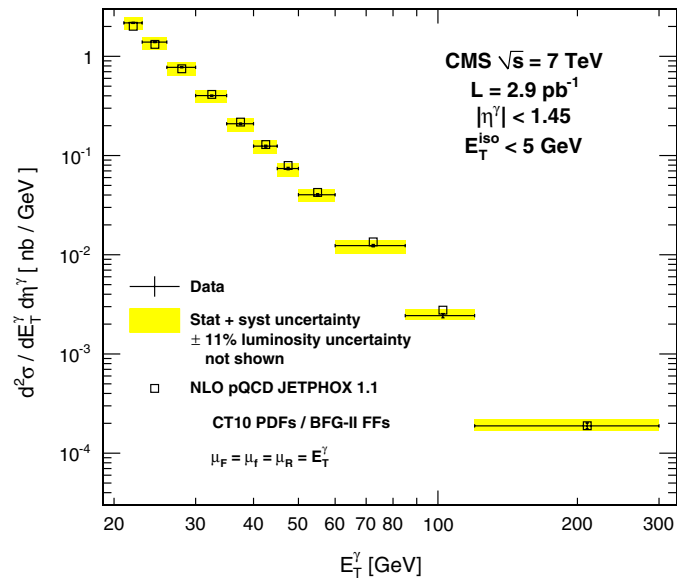
Figure 2.7: Measured isolated prompt photon differential cross section as a function of E_{T}^{γ} measured by the CMS detector [42]. Also shown are NLO pQCD prediction from JETPHOX.

Chapter 3

Experimental Setup

3.1   Large Hadron Collider

Located underneath the border between Switzerland and France, near Geneva, is an underground ring with a circumference of 27 km. The ring lies between 45 m and 170 m below the surface of the earth (Fig. 3.1) and it contains over a thousand dipole magnets that direct two oppositely-rotating beams of protons. Each beam is accelerated to energies of 3.5 TeV so that the total center-of-mass energy upon collision is 7 TeV. More details on the accelerator complex at CERN can be found in [45], but a brief description follows.

To be more cost efficient, the LHC was built in the existing tunnel that was constructed for the LEP accelerator. LEP was built between 1984 and 1989 to accelerate and collide electrons and positrons. The tunnel itself has an internal diameter of 3.7 m, which, due to this small size, forced a magnet design that would allow for the two beams to operate very close together.

The design center-of-mass energy for proton-proton collisions is $\sqrt{s} = 14$ TeV, but for the 2010 run from May to November it was 7 TeV. To control protons with
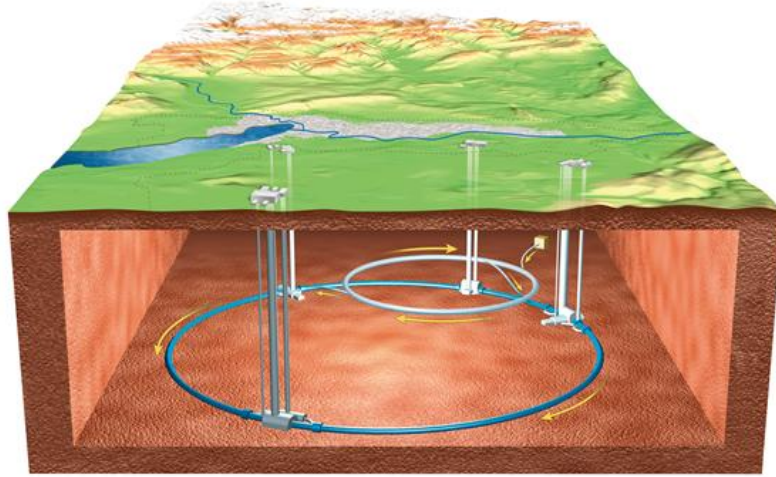
Figure 3.1: The accelerator complex at CERN outside of Geneva, Switzerland is about 300 feet below the surface. The main, large ring has a circumference of 27 km (image not to scale).

energy in the TeV range the super-conducting NbTi magnets require, in total, over 11,000 A of current to produce magnetic fields of a maximum of 8.4 T. The beams of protons are "bunched" such that there is 25 ns between crossings of the proton beams at design luminosity. At design luminosity, 2808 bunches of protons will be collided.

Along the LHC beampipe are several detectors, including two general, all-purpose particle detectors known as ATLAS and CMS. In addition, there is the ALICE detector built specifically for when the LHC collides heavy ions (Pb-Pb), and also the LHCb detector built for studying b-physics.

The production rate of collision events is proportional to the luminosity of the LHC machine:

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F \tag{3.1}$$

where N_b is the number of particles per bunch, n_b is the number of bunches per beam, f_{rev} is the revolution frequency, γ_r is the relativistic gamma factor, ϵ_n is the

31
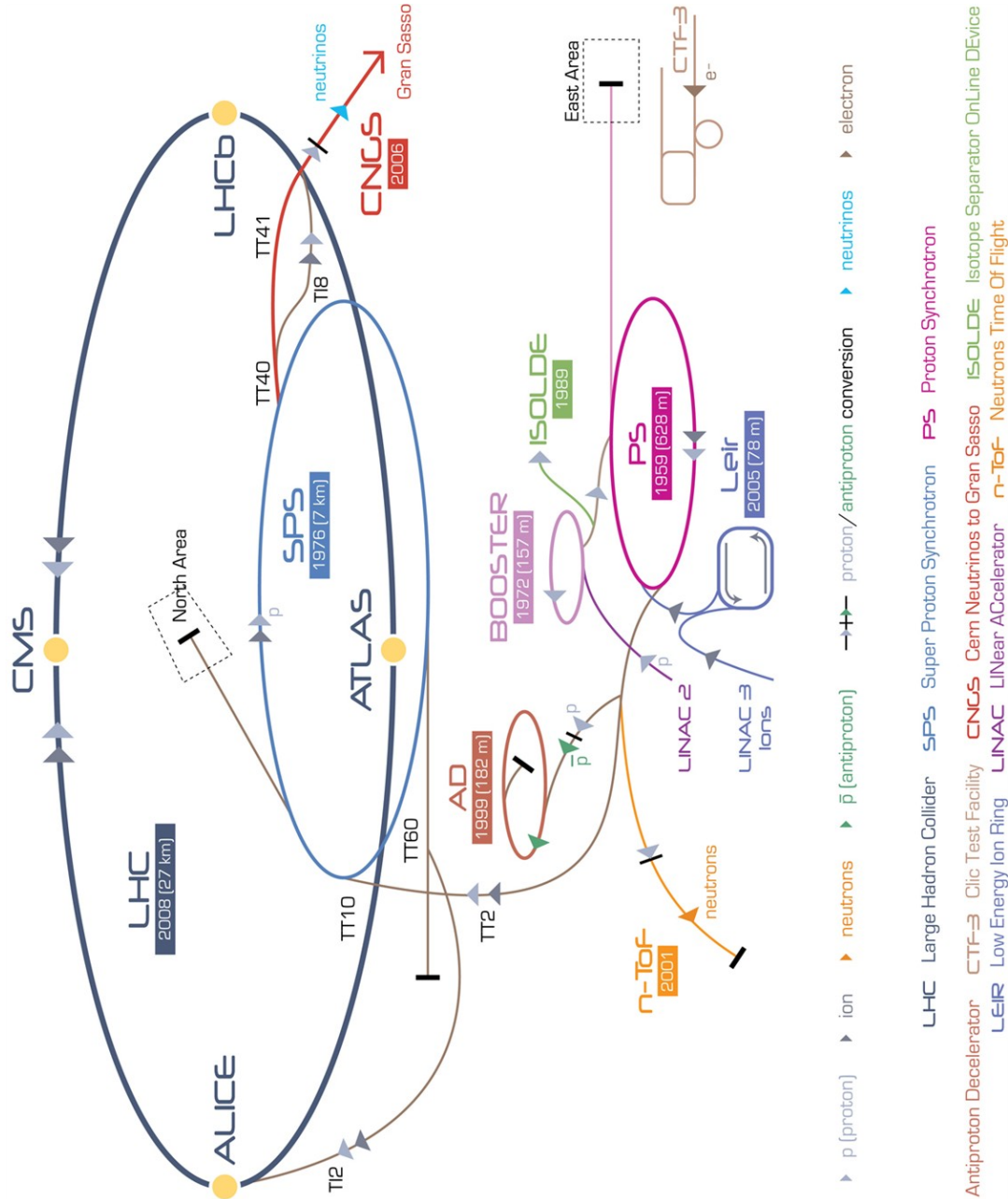
Figure 3.2: The Large Hadron Collider is the last ring in a complex chain of proton accelerators. The smaller accelerators are used to help boost the proton to their final energies and provide beams to a whole set of experiments.

normalized transverse beam emmittance, β^* is the beta function at the collision point, and F is the geometric luminosity reduction factor due to the crossing angle at the interaction point. The design peak luminosity of the LHC is $\mathcal{L} = 10^{34}\,\mathrm{cm^{-2}\,s^{-1}}$, and during the 2010 run a peak luminosity of $\mathcal{L} = 2 \times 10^{32}\,\mathrm{cm^{-2}\,s^{-1}}$ was reached.

The LHC accelerator team and the CMS luminosity group collaborated to use Van der Meer scans to measure the beam profiles (This process involves one beam being swept stepwise across the other while measuring the collision rate as a function of beam displacement). They also measured currents using beam pickup in coils. These measurements resulted in 11% uncertainty in the instantaneous luminosity at the start of the experiment [46], but were later refined to 4% for operation during 2010. Relative luminosity monitoring is done using the hadronic forward calorimeter and pixel hits. This allows for integration to obtain total luminosity per lumi-section.

The amount of data delivered to the detectors is measured as the integral of the luminosity, $L_{\mathrm{int}} = \int \mathcal{L}dt$, and is shown for the 2010 run in figure 3.3. The CMS experiment recorded a total integrated luminosity of $43 \pm 5\,\mathrm{pb^{-1}}$ of pp collisions during this time. In order to use the highest quality data recorded by the detector, the amount of data used in this thesis is $36\,\mathrm{pb^{-1}}$.

3.2 Compact Muon Solenoid

On the French side of the LHC is the 12,500-ton Compact Muon Solenoid (CMS) detector (fig. 3.4). The CMS detector is $21\,\mathrm{m}$ in length and $16\,\mathrm{m}$ in height. It contains several subdetectors detectors (trackers, calorimeters, and muon chambers) that track and measure particles over a wide-range of energies to help measure or search for many physics processes. Protons from the LHC beam-pipe are led in from the left and right sides of the detector to the very center where they collide right at
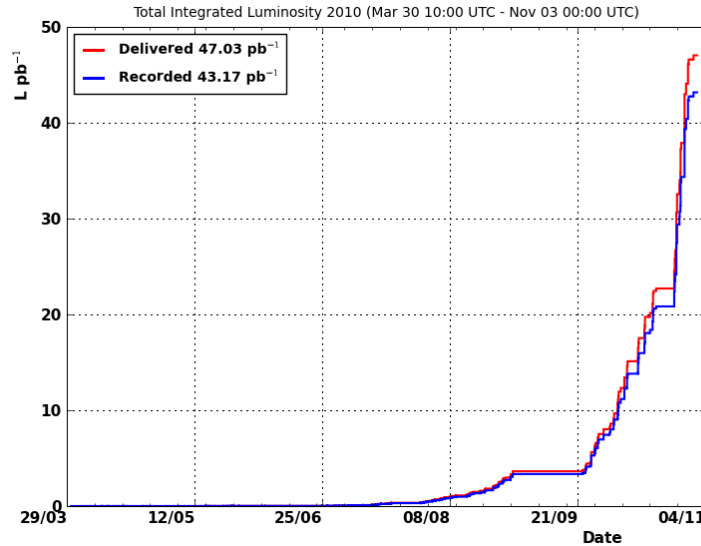
Figure 3.3: The total integrated luminosity delivered by the LHC machine (red) and recorded by the CMS detector (blue) in 2010 as a function of time[47].

the heart of CMS. Details on the CMS detector can be found in [48], while a brief description of the systems relevant to this analysis follows.

When designing the CMS detector, it was important to take into account the crossing time of the proton beams at the LHC. Between each bunch of protons is the relatively short time of 25 ns. When the LHC reaches design luminosity ($\mathcal{L} = 10^{34}\,\mathrm{cm^{-2}\,s^{-1}}$), interesting events will occur with an average of 20 inelastic collisions superimposed. That is, roughly 1,000 charged particles will emerge from the interaction region every 25 ns. Thus, it is important that detector elements have a response time of less than 25 ns. To reduce the effects of these "pile-ups", detectors with high granularity and good time resolution were used.

The Tracker system is located just outside the beampipe beginning at a radius of 4.4 cm from the interaction point. This silicon tracker measures the paths of particles such as muons, electrons and hadrons, which are charged. Closest to the interaction
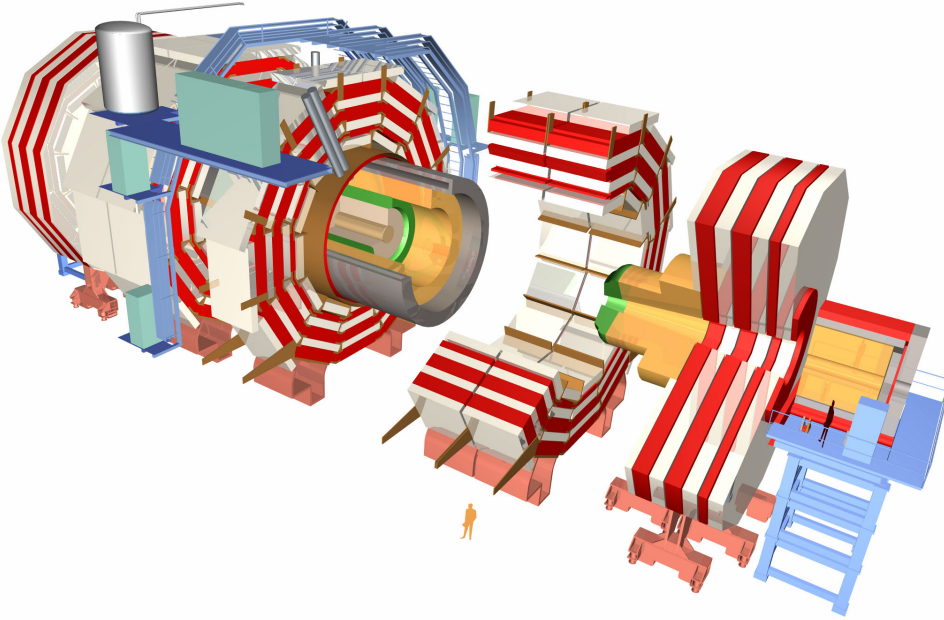
Figure 3.4: 3D Drawing of the CMS detector with person shown for scale.

point are silicon pixel detectors that measure the primary and secondary vertices of proton interactions in both x-y and z dimensions. Outside this are 10 layers of silicon microstrip detectors. Both the pixel and microstrip detectors provide the granularity and precision required to reduce effects from pile-up events.

Outside the Tracker system is the Electromagnetic Calorimeter (ECAL). This subdetector consists of over 76,000 lead tungstate ($PbWO_4$) crystals. These provide a precise measurement of photon energy and positions. They have short radiation lengths and Moliere lengths – 0.89 cm and 2.2 cm respectively. They are also fast – within 25 ns 80% of the light is emitted. The front of each crystal has a cross section of 22x22 mm^2 and a depth corresponding to 25.8 radiation lengths. In its entirety, the ECAL covers up to an $|\eta|$ of 3.0.

Outside the ECAL is a Hadron Calorimeter (HCAL) made up of brass plates and scintillator tiles. The HCAL covers up to an $|\eta|$ of 3.0. Beyond that, from $3 < |\eta| < 5$,

is the forward hadronic calorimeter (HF).

These three subdetectors (Tracker, ECAL and HCAL) are contained within the 6 m diameter superconducting solenoid that provides a field of 3.8 T. The magnetic field it creates runs parallel to the beampipe so that the products of particle interactions and decays can have their momentum determined by measuring their radius of curvature through the Tracker.

Coordinate System

CMS uses a right-handed coordinate system, whose origin is defined to be the the nominal interaction point. The x-axis points to the center of the LHC ring, the y-axis points vertically upward, and the z-axis points west, along the beam direction towards the Jura Mountains. The polar angle θ is measured from the z-axis, while the azimulthal angle ϕ is measured from the x-axis in the x-y plane. The radial coordinate, r, is defined in the x-y plane. Another variable, pseudorapidity is defined as $\eta = -\ln(\tan\theta/2)$, and is useful at hadron colliders because the particle production rate is roughly flat as a function of pseudorapidity.

Tracking System

The tracking system was designed to obtain precise measurements of charged particle trajectories, as well as the vertices of hard interactions. It is contained in a homogeneous magnetic field of 3.8 T, and thus it obtains the transverse momentum of particles by measuring the radius of curvature as they travel through the detector. It has a length of 5.8 m and a diameter of 2.5 m.

To keep occupancy low, the inner part of the tracker is composed of a pixel detector with three layers in the barrel of CMS at radii between 4.4 cm and 10.2 cm.
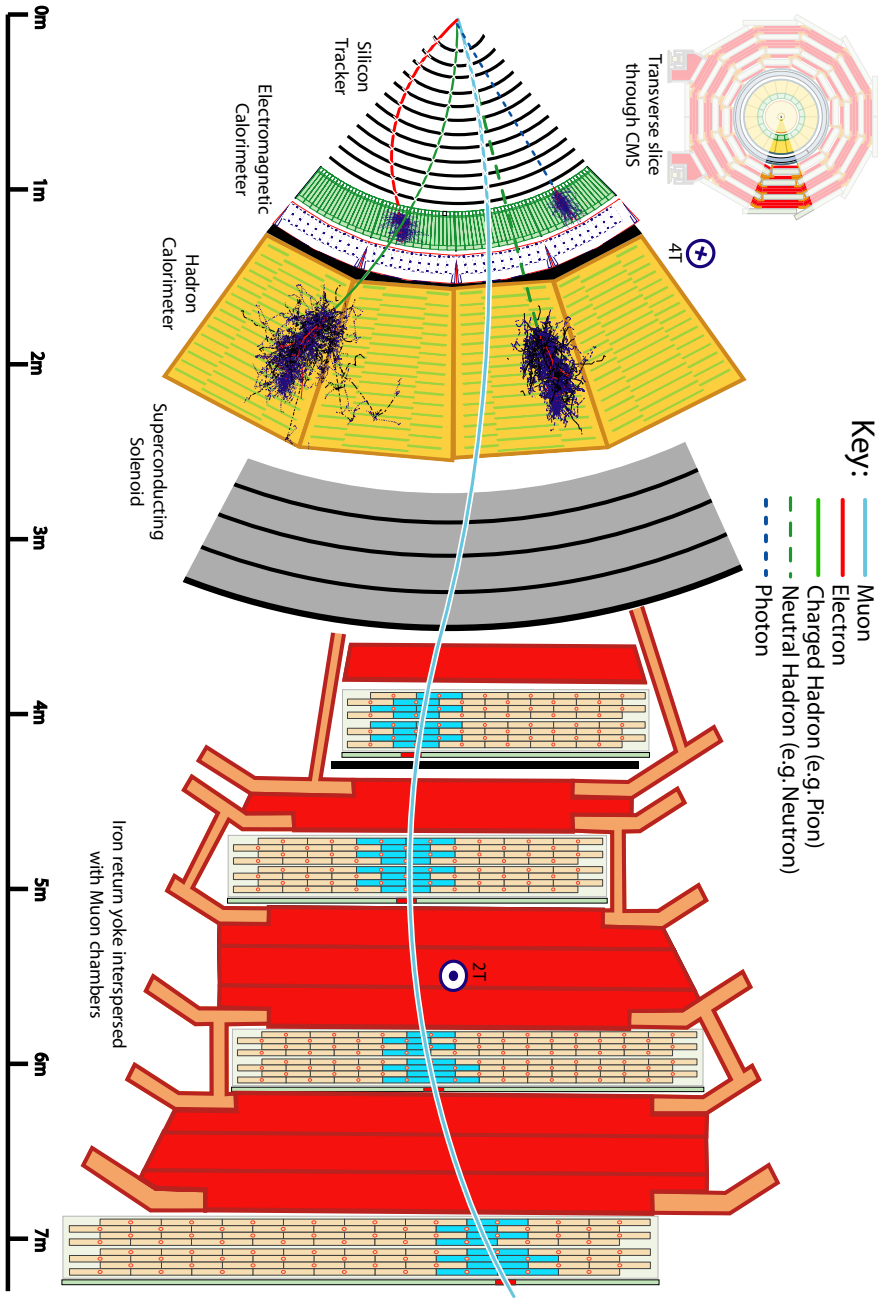
Figure 3.5: A slice of the CMS detector and how various particles would travel through it. Photons (blue dashed lines) deposit all of their energy in the Electromagnetic Calorimeter.
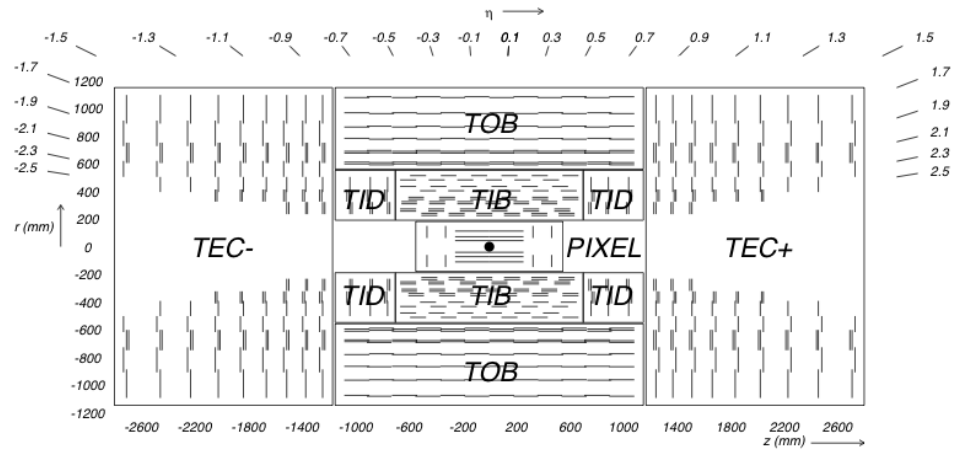
Figure 3.6: Schematic of a slice the CMS tracker parallel to the beam line [48]. Each line represents a module of the detector.
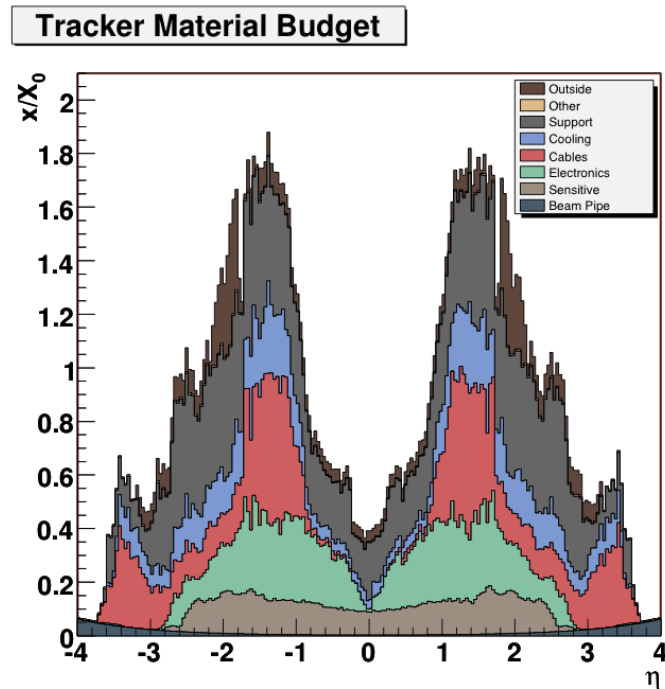
Figure 3.7: Material budget of the Tracker System in CMS in units of radiation length [48].

The resolution in the $r - \phi$ plane is $\sim 10\,\mu$m, and in the $r - z$ plane is $\sim 20\mu$m. Beyond that is a silicon strip tracker with 10 layers in the barrel region out to a radius of 1.1 m. It is divided into the Tracker Inner Barrel for $|r| < 65$ cm and the Tracker Outer Barrel (TOB) for $65 < |r| < 110$ cm. The TIB and the TOB have a resolution of $230\,\mu$m and $530\,\mu$m respectively. In total, the tracking system has fine granularity and covers the region $|\eta| < 2.5$ with an area of about $200\,\text{m}^2$ of active silicon sensors.

The Pixel Tracker detector is comprised of two disks in the forward region, and in the central region it is three cylindrical layers as shown in figure 3.6.

A major design constraint to the tracking system was that the amount of material must be kept low in order to limit the amount of interactions and photon conversions that particles undergo before reaching the ECAL subdetector. The material in the tracker has 0.4 to 1.8 radiation lengths depending on η shown in figure 3.7.

Electromagnetic Calorimeter

One of the major motivations for the design of the ECAL was the need for the potential discovery of a low-mass neutral Higgs Boson which can decay to two photons. Electrons and photons interact with the electrons within atoms inside the electromagnetic calorimeter through Compton scattering and the photoelectric effect to create showers of photons and e^+/e^- pairs. The material chosen for the ECAL was PbWO_4, which has a high density (8.28g/cm^3), low radiation length (0.89 cm) and narrow Moliere radius (2.2 cm). These qualities allow the system to be compact and have high granularity.

The ECAL system is composed of 61,200 crystals in the barrel, and 13,648 crystals in the endcap of CMS. When a high energy photon or electron strikes one of these
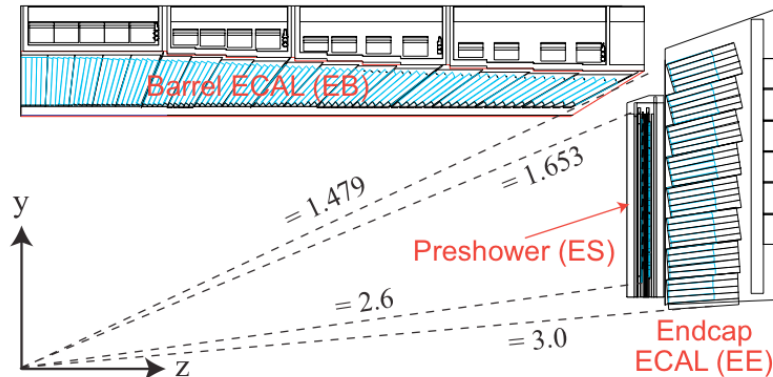
Figure 3.8: Schematic of a transverse section through the ECAL, showing the barrel ($|\eta| < 1.479$) and the endcap ($1.479 < |\eta|$) regions.

crystals, about 80% of the light is emitted in the LHC design bunch crossing of 25 ns.

In the barrel of CMS, the crystals have a front face of $2.2 \times 2.2 \text{cm}^2$ and are 23cm long (25.8 radiation lengths). The crystals are read out by avalanche photo diodes (APD). The barrel ECAL system covers the region $|\eta| < 1.479$.

In the endcap, the ECAL crystals have a front face cross-section of $2.86 \times 2.86 \text{cm}^2$ and a length of 22.0cm (24.7 radiation lengths). They are read out by vacuum phototriodes (VPT).

In the region $1.65 < |\eta| < 2.5$ there is a Pb-Si preshower (ES) detector between the ECAL and the nominal interaction point. This subdetector was designed to improve the discrimination between isolated photons and neutral mesons (such as π^0). It is a sampling calorimeter with silicon sensors between lead radiators.

The energy resolution of the ECAL system can be parametrized into three terms:

$$\left(\frac{\sigma}{E}\right) = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C \tag{3.2}$$

where S is the intrinsic stochastic term, N is the noise term, and C is the constant

term. The stochastic term arises from fluctuations in the lateral shower containment. The noise term comes from noise from electronics, digitization and pileup. The constant term comes from intercalibration errors, non-uniformity of light collection, and leakage of energy from the back of the crystals. The values of the three terms are determined by a fit to the resolution from an electron test beam which is shown in figure 3.9. Above $20\,\mathrm{GeV}$, the resolution of the ECAL ($\sigma(E)/E$) is better than 1%.
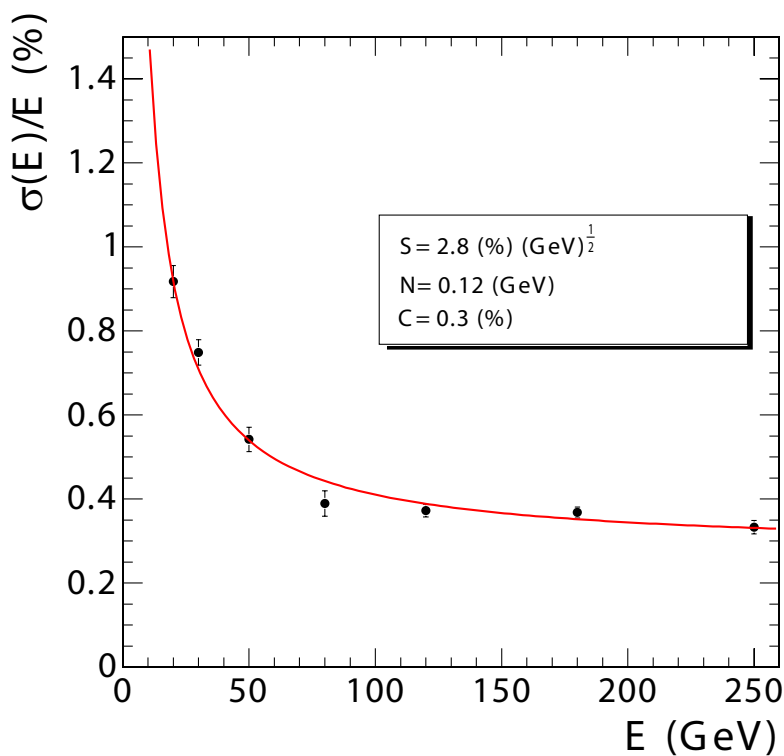
Figure 3.9: Energy resolution of the ECAL, $\sigma(E)/E$, as a function of energy from an electron-beam test [48]. The fit returns values for the stochastic (S), noise (N), and constant (C) term.

Hadronic Calorimeter

Beyond the ECAL lies the hadronic calorimeter where hadronic particles deposit most of their energy. The HCAL is made of a dense material, brass, which causes hadrons

to interact with their nuclei and creates a cascade of particles (such as K mesons, protons, and neutrons). These particles pass through scintilator tiles interweaved between absorber plates in the HCAL.

The hadronic calorimeter in the barrel of CMS (called HB) covers the range $|\eta| <$ 1.3. It lies between the ECAL and the solenoid, with a total width of 1.18 m. Outside the solenoid there is an additional scinitilator tile which uses the solenoid as an absorber, this is the hadronic outer calorimeter (HO). The hadronic calorimeter in the endcap (called HE) covers the range $1.3 < |\eta| < 3.0$.

In the far forward region is the hadronic forward calorimeter (HF), which covers the range $3.0 < |\eta| < 5.0$. The front of the HF is 11.2 m from the interaction region. Due to high multiplicity in this far forward region, quartz fibers are used as the active medium.

The energy resolution of the hadronic calorimeter to hadronic jets was measured and found to be better than 10% of jets with $E_\mathrm{T} > 40\,\mathrm{GeV}$[49].

Solenoid and Magnetic Field

The magnetic field in the CMS detector is created by a superconducting magnetic solenoid with an inner diameter of 5.9 m, and length of 12.5 m. It was designed to create an inner magnetic field up to 4 T. At full current, the energy stored in the magnetic field is 2.6 GJ. Outside the solenoid, the return field is directed by the iron yoke which consists of 5 wheels and 2 endcaps (each made of 3 disks).

Trigger System

The design collision rate at the LHC is 40 MHz, and the amount of data needed to record all information for an event is on the order of 250 kB. To store data for each
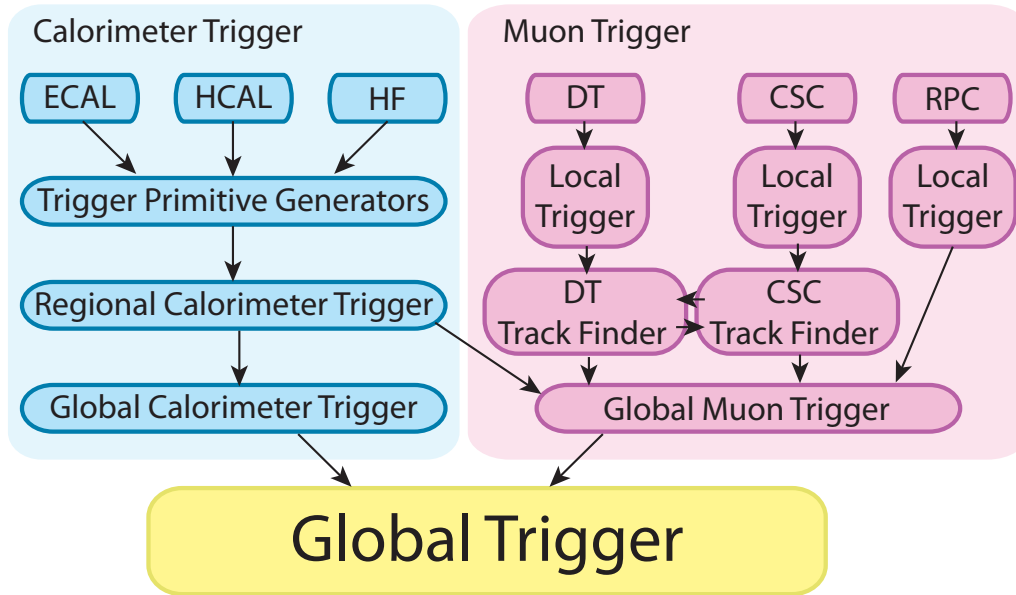
Figure 3.10: Schematic of how information passes through the various systems within the Level-1 Trigger.

pp collision would then require writing several TB/s, which is far too much data for any storage system currently available. The main requirement of the trigger system is to select the more interesting events to reduce the recorded event rate to a more manageable 100 Hz for data storage.

Another reason for the rate reduction of stored events is shown in figure 1.2. The production cross section for physics processes such as Supersymmetry or Higgs Bosons are several orders of magnitude lower than the inelastic pp cross section. The trigger is the first step in selecting interesting physics processes.

To achieve rate reduction by a factor of 10^6, CMS employs a 2-level trigger system. These are referred to within CMS as the Level-1 Trigger (L1) and the High Level Trigger (HLT).

The Level-1 trigger (L1) has been constructed using programmable electronics. These are Field Programmable Gate Arrays, as well as Application Specific Inte-

gration Circuits and programmable memory lookup tables. The L1 trigger reduces the input rate of 40 MHz down to below 100 kHz. The L1 trigger does not use full-resolution information from the various subdetectors, but instead uses coarsely segmented data to make a trigger decision. Information from the tracker subdetector is not used at L1 because it is too slow. A schematic of the flow of information through the level 1 system is shown in figure 3.10.

The Level-1 trigger has two main systems: the calorimeter trigger and the muon trigger. The calorimeter trigger provides information on objects such as the 8 highest energy e/γ candidates, the 8 highest energy jets, 4 leading τ jets, among other quantities. Using the Drift Tubes (DT) and Resistive-Plate-Chambers (RPC) in the barrel, and Cathode-Strip-Chambers (CSC) in the endcap, the muon trigger system finds the top 4 highest energy muon candidates. Both the calorimeter and muon trigger systems provide this information to the global trigger to make the final decision on whether to keep an event.

If an event passes the decision by the Level-1 global trigger, it moves on to the High Level Trigger system [50]. The HLT is software-based and runs on commercially available processors. It uses the full resolution information and so uses the same (or similar) algorithms as used offline for computations such as object isolation or momentum.

3.3   Computing

The amount of data needed to record all information for each pp collision event with the CMS detector is on the order of 250 kB. The purpose of the computing system of CMS is to store, transfer, and process the data from recorded events. Like the members of the CMS collaboration, the computing systems are also distributed

around the world. Thanks to modern broadband internet connections (10 to 100 Gbps), data collected by the detector can quickly make it to the users who need it for analysis.

Computing centers of the CMS collaboration are organized into "tiers". There is a single Tier-0 center located at CERN, there are 8 Tier-1 centers around the world, and finally, over 50 Tier-2 centers located at universities and institutes.

The Tier-0 center has several functions. It must accept data directly from the online detector system and copy it to permanent mass storage (tape). It performs the first pass of reconstructing the RAW data to provide RECO datasets. Both the RAW and RECO dataseets are transferred then to Tier-1 centers. A dataset is not removed from the Tier-0 until it is safely stored at two Tier-1 centers.

The Tier-1 centers provide long-term storage of RAW data. When reconstruction software improves or updates, the RAW data is reconstructed again to produce new RECO datasets. It transfers these datasets in a reduced format, Analysis Object Data (AOD) to Tier-2 centers.

The Tier-2 centers vary widely in computing power and storage. Some Tier-2 centers are larger than Tier-1 centers in other countries. The production of Monte Carlo samples largely takes place at Tier-2s. In addition, user analysis is completed with the resources at Tier-2s.

Some of my own main support work of the CMS experiment was user-support at the Tier-2 center located at the University of Wisconsin-Madison. In addition, I helped with the production of Monte Carlo simulated data. This involved using software to create and submit hundreds to thousands of compute "jobs" for each sample of simulated data. I helped create software tools to monitor and control the flow of these jobs across several Tier-2 sites.
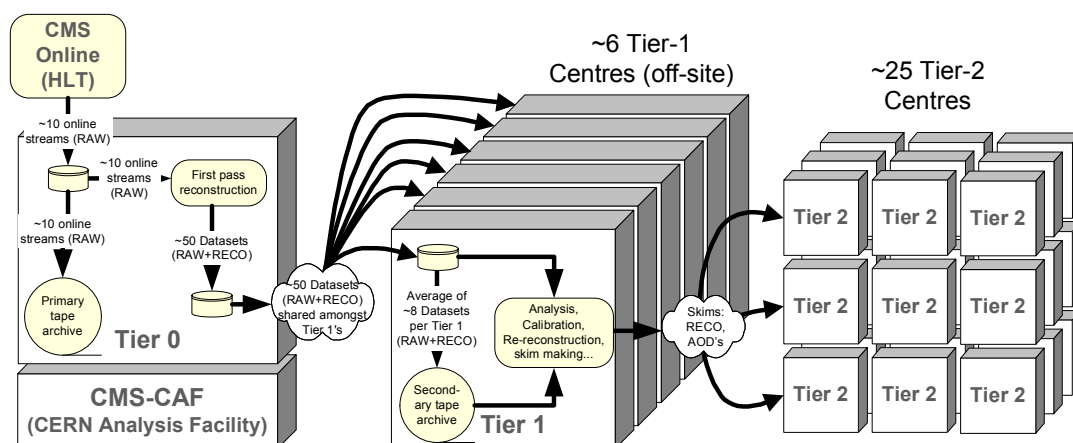
Figure 3.11: Schematic showing the flow of data from the CMS detector through the various "tiers" of computing centers down to Tier-2 centers where user analysis takes place [48].
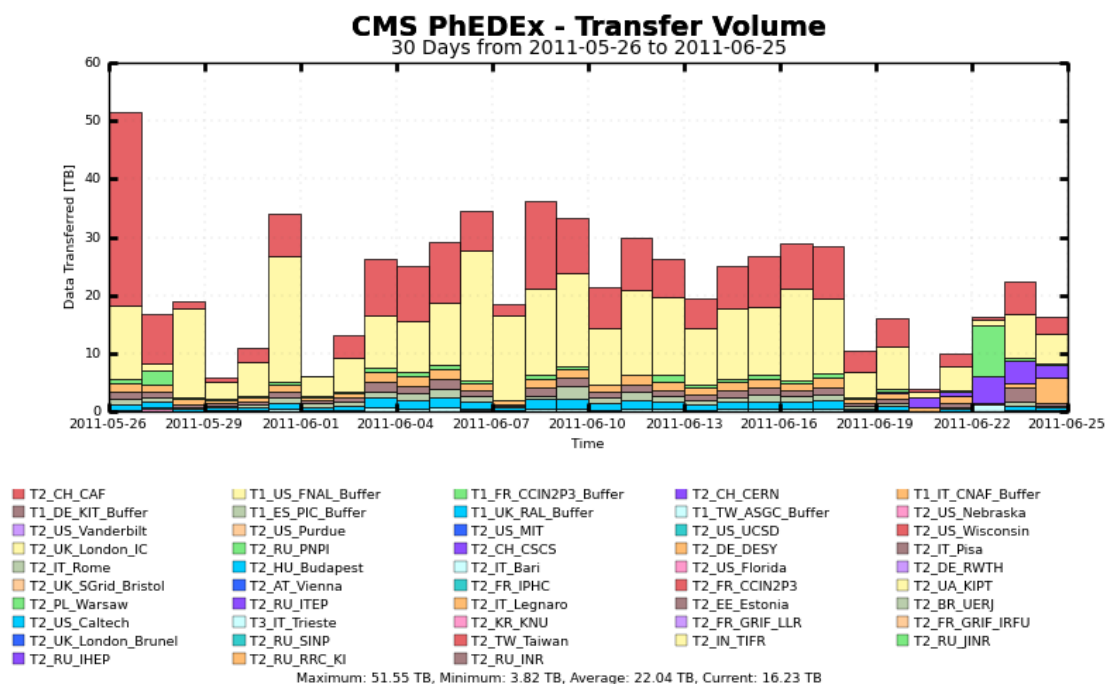
Figure 3.12: Volume of data transferred from the CERN Tier-0 center during a typical month with CMS recording *pp* collisions. In total, it is several hundred TB of data transferred from CERN each month.

Chapter 4

Monte-Carlo Event Simulations

Predicting the results of colliding protons involves modeling of several things including modeling the subatomic makeup of a proton, the calculation for scattering amplitudes, the decay of unstable particles, and the hadronization of quarks and gluons into jets. This is generally labeled as the "event generation". After that, the response of the detector to these final state particles must be modeled. Knowledge of detector materials and positions of these materials is necessary for accurate modeling of the detector response.

Monte Carlo simulations are used for event generation and modeling of detector response. This helps when designing particle detectors and helps define strategies for finding signatures experimentally. Monte Carlo simulations are designed to reproduce as closely as possible the collision processes taking place in detectors. These simulations produce hypothetical events with distributions and rates predicted by theory. Pseudo-random numbers are used and weighted by probability of underlying processes occurring in particle collisions. Monte Carlo event generators create a list of final state particles from a hard scattering process which is then passed to another simulator to determine the response of the detector to these particles.

The software for many Monte Carlo simulators allow the computation of matrix elements with a fixed number of partons in the final state. Most often, they describe a given final state to the lowest order in perturbation theory and do not include calculation of Feynman diagrams with loops in them. Calculating the value of Feynman diagrams with loops is complicated and typically involves procedures to eliminate divergences at high energy and high momentum (this procedure is referred to as 'regularization'). Often, approximate methods are used to obtain high-order contributions perturbation series.

4.1 Hard Scattering and QCD Radiation

The calculation of any production cross section requires knowledge of the distribution of the momentum fraction x of the partons (quarks and gluons) in the incoming hadrons in the relevant kinematic range. Thus, the first step in simulating proton-proton collisions requires probabilities of finding partons of specific flavor and carrying certain fraction, x of the proton momentum. Such probabilities are giving by parton distribution functions (PDF). PDFs cannot be calculated perturbatively, instead they are determined by global fits to data from deep inelastic scattering, Drell-Yan, and jet production at high energy. These distribution functions are also dependent on the square of the momentum exchanged between colliding partons, which determines the energy scale of the collision, Q^2.

The PDFs used in the analysis described in this thesis is from the Coordinated Theoretical-Experimental Project on QCD (CTEQ) group[51]. The most recent PDFs were used, CTEQ6.1 [52], and provide the most accurate description available to collision data. These PDFs are fed into parton *evolution* equations, which are models of perturbative QCD (pQCD) [53]. These extrapolate the Q^2 and x dependence of
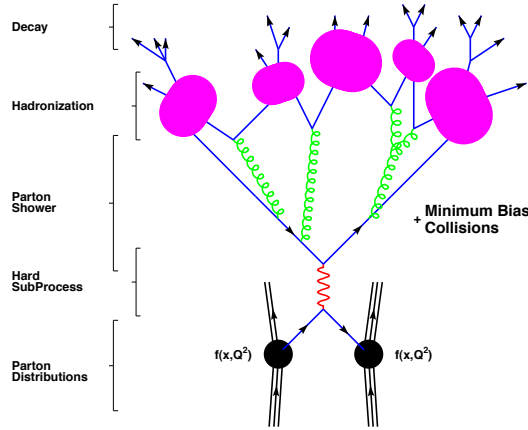
Figure 4.1: A diagram of a generated event which shows incoming hadrons where a parton from each interact in a hard scattering process, and the outgoing partons undergo showering and hadronization [54].

given PDFs.

4.2 Showering and Hadronization

After simulating a hard scattering event, the final state contains quarks and gluons. Some Monte Carlo generators begin with leading order hard subprocesses and higher order effects are added by evolving the event with parton showering, which splits partons into pairs of other partons. Whereas others include some next-to-leading-order perturbative QCD diagrams leading to 3 or more partons in the final state before beginning the parton showering process. Resulting partons are grouped together into colorless hadrons (hadronization) in accordance with the quark confinement phenomenon, described by approximate models based on QCD theory. A schematic of this whole process is shown in figure 4.1.

Hadronization is the process that begins with a single colored parton into a collection of colorless particles. This process cannot be modeled by perturbative QCD, as

hadronization takes place over relatively larger distances compared to the hard inter-action. In addition, hadronization involves scattering with low momentum transfer. A common model for this is the Lund String Model [55], described below.
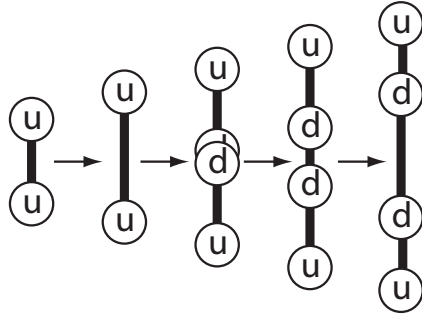
Figure 4.2: An example of hadronization of an initial state of two quarks where another quark pair is created from the energy stored in the color field.

To start, the Lund String Model represents a color field as a one-dimensional tube (or 'string') between two pairs of colored quarks. Due to the three-gluon coupling, color flux lines do not spread out over all space (unlike electromagnetic fields), but are constrained to a thin region. The amount of energy per unit length of the tube is phenomenologically found to be $\kappa \approx 1 \, \text{GeV}/fm$. New pairs of quarks $(q\bar{q})$ can be created from available energy in the field. The original system of quarks breaks into smaller pieces iteratively, until only colorless hadrons remain - which may themselves decay into leptons or stable hadrons. The final system of particles emerge aligned along the original $q\bar{q}$-axis. A depiction of an example of this process is shown in figure 4.2.

Beam remnants are the colored partons left over in the proton after a parton is pulled out to participate in the hard scattering event. They have small transverse momentum (of order 1 GeV), but are connected through color fields to the hard scattering. Within Monte Carlo simulators such as PYTHIA, these beam remnants

are allowed to interact (a multi-parton interaction), shower and decay themselves. What results from these beam remnants is generally referred to as the "underlying event".

4.3   Detector Simulation

After the simulation of the hard scattering event to final-state particles, a simulation is performed on their interaction with the CMS detector. Descriptions of detector material and locations are stored and used to determine how the particles will interact as they traverse through the detector. This stage of the simulation is performed within the CMSSW framework by the GEANT4 software package [56]. All aspects of the detector simulation process are included in this package including: geometry of the detector, its materials, the tracking of particles through those materials, and through electromagnetic fields. The core of GEANT4 is a set of physics models to handle interactions of particles with materials over a large energy range (from a few hundred eV up to a few TeV).

After the reaction of the detector to these particles as they pass through the detector is simulated, reconstruction of the event is performed by the same software used to reconstruct true data.

4.4   MC Simulator Programs

PYTHIA

Some of the simulated events generated for this thesis use the software program PYTHIA [57]. This uses leading-order matrix elements to simulate hard scattering

between partons. It was primarily developed for hard-scattering processes that involve 2 incoming particles and 1 or 2 outgoing particles. The subprocesses simulated for the signal samples used in this analysis are dominated by $q\bar{q} \to g\gamma$, and $qg \to q\gamma$ [58, 59]. Samples for backgrounds (QCD) were not used from PYTHIA in this analysis.

MADGRAPH

The MadGraph matrix element generator [60] software allows one to generate amplitudes for any process up to 9 final state particles in any process. MADGRAPH automatically generates amplitudes for relevant subprocesses and produces mappings for the integration over phase space. Events are passed as parton level files in the standard format called Les Houches format (LHE files). These parton level files are then passed to PYTHIA, which takes care of the parton showering and the fragmentation part before passing the event to detector simulation. The matching procedure of matrix elements to parton shower (ME-PS) also happens in at this stage. It is important to perform the merging that avoids double counting of emissions in overlapping phase space regions. There are two methods available to perform the merging, the CKKW algorithm [61, 62], and the MLM scheme [63].

The MADGRAPH samples used in this thesis contain tree level calculations for parton-level processes of up to $2 \to 4$ (so for example, up to $pp \to \gamma + 3$ jets). They were produced with cuts on the scalar sum of the outgoing parton transverse momentums, $H_T = \sum p_T$.

4.5   Monte Carlo Samples

The MC samples used in this thesis are listed in table 4.1. The PYTHIA samples were only used for an additional final generator-level comparison. The MADGRAPH

samples were used to obtain template distributions for true photons, for finding the response matrix, and also determining efficiency of photon selection.

The PYTHIA samples have the default underlying event tune Z2, while the MAD-GRAPH samples use the D6T tune. These tunes simply control the amount of activity in the underlying event and are described in more detail in [64]. These samples were reconstructed with CMS software version `3_8_4_patch2` in the Fall of 2010.

Process	Kinematic cuts (in GeV, $c = 1$)	σ (pb)
MADGRAPH γ+jets	$20 < H_\mathrm{T}$	5.95×10^4
MADGRAPH QCD	$50 < H_\mathrm{T}$	3.71×10^7
PYTHIA γ+jet	$15 < \hat{p}_\mathrm{T} < 300$	1.91×10^5

Table 4.1: Summary of Monte Carlo simulated datasets used in this analysis. Note that H_T is the scalar sum of transverse momentums of outgoing partons, and $\hat{p}_\mathrm{T}$ is the transverse momentum of the hard interaction.

The PYTHIA γ+jet samples set the parameter, `MSEL = 10`. This specifically turns on simulation of five subprocesses: $q\bar{q} \to g\gamma$, $ff \to \gamma\gamma$, $qg \to q\gamma$, $gg \to \gamma\gamma$, and $gg \to g\gamma$ in the hard interaction of the proton collision. The dominate processes out of those is the $q\bar{q}$ and qg interaction.

The MADGRAPH samples were generated by MadGraph/MadEvent version `4.4.30`. The processes simulated in the γ+jets samples were: $pp \to \gamma + 1j; \gamma + 2j; \gamma + 3j$, where j are jets (an outgoing parton, which was a sum of from all quark flavors). The processes simulated in the MADGRAPH QCD samples were, specifically: $pp \to 2j; 3j; 4j$. The QCD scale in the MADGRAPH samples was set to $91.1880\,\mathrm{GeV}$.

Chapter 5

Event Reconstruction

5.1   Track Reconstruction

The strong magnetic field of 3.8 T causes sufficient curvature in the trajectories of charged particles that it allows the measurement of their momentum. Charged particles leave 'hits' in the Tracker subdetector, and they are reconstructed as follows [65]. First, the seeding step: triplets or pairs of hits in the pixel part of the tracker or the inner layers of the strip tracker are used as track candidates. Next, track candidates are propagated using a Kalman filter technique to find new compatible hits. If not enough hits are found, the track candidate is rejected. Then, a final track fit is performed and all the track parameters are estimated combining all the associated hits. A quality flag is then assigned to the track based on several quantities including: track normalized χ^2, transverse impact parameter d_{xy} with respect to the beamspot, longitudinal impact parameter d_z with respect to the closest HLT primary vertex, and others described in more detail in [65]. Many fake tracks are removed by the reconstruction software with requirements placed on the fit χ^2, and compatibility with the beam line or primary vertices of the event.

5.2 Vertex Reconstruction

The location of a proton-proton interaction (a "vertex") is reconstructed from tracks selected by a transverse impact parameter with respect to the beamspot. The tracks are clustered along the z-axis by requiring a separation of at least 1 cm to the next cluster. These clusters are then fit by an adaptive vertex fitter. This adaptive vertex fitter does not reject outlying tracks. Instead, it it down-weights outliers with a weight, given by:

$$w_i(\chi_i^2) = \frac{\exp(-\chi_i^2/2T)}{\exp(\exp(-\chi_i^2/2T) + \exp(-\chi_c^2/2T))}$$

where w_i is the probability that a track, i, belongs to a vertex. The constant χ_c^2 defines the threshold where the weight is equal to $1/2$. Below this value, a track is considered an outlier. The variable T is a parameter that controls the shape of the functional dependence. More details of vertex reconstruction can be found in [66] and [67].

A data-driven method was applied to measure the efficiency of reconstruction of the primary vertex and the resolution of the position. This method splits tracks randomly into two sets. Both sets are fit independently with the same procedure. The resolution is determined by a gaussian fit to the distribution of the difference between the two vertex positions. This resolution depends on the number of tracks used in the fit to the vertex and on their momentum.

The efficiency of reconstructing a vertex was calculated by considering one of the vertices a "tag" and the other a "probe". The efficiency is the fraction of reconstructed probe vertices that are within a z distance smaller than 5σ of the tag vertex. This efficiency was found to be almost 100% when there are more than two tracks in the event with momenta greater than 0.5 GeV.

5.3 Jet Reconstruction

Some outgoing particles from a hard scattering event, such as gluons and quarks, hadronize into collimated collections of baryons, mesons and other particles, which are collectively called "jets". It is these final state particles which are ultimately detected by a particle detector. Clustering the deposits of energy or tracker-measured momentum vectors from a large collection of particles into one or more jets is done by jet clustering algorithms.

Jets used in this analysis are clustered using particles from the list of particles reconstructed using the "particle-flow" algorithm [68], which combines the information from all sub-detectors to identify and measure all particles produced in the collision, namely charged hadrons, photons, neutral hadrons, muons, and electrons. The resulting list of particles are then used by the jet clustering algorithms to construct a higher-level objects and observables such as jets.

Anti-k_t Algorithm

An extension of the k_t algorithm [70], and the Cambridge/Aachen inclusive jet algorithm[71, 72] is the anti-k_t algorithm[69]. Consider typical event with a few hard particles, and many soft particles. Often, the hard particles are well separated in space. Labeling the transverse momentum of the hard particles as k_{t1}, k_{t2}, ... and the separation between hard and soft particles as Δ_{ij}, a variable is defined:

$$d_{ij} = \min\left(1/k_{ti}^2, 1/k_{tj}^2\right) \frac{\Delta_{ij}^2}{R^2}$$

The d_{ij} between separated soft particles will be large, while this value will be small between a hard particle and a nearby soft particle. This causes soft particles to cluster with hard particles long before they are clustered among themselves. If a
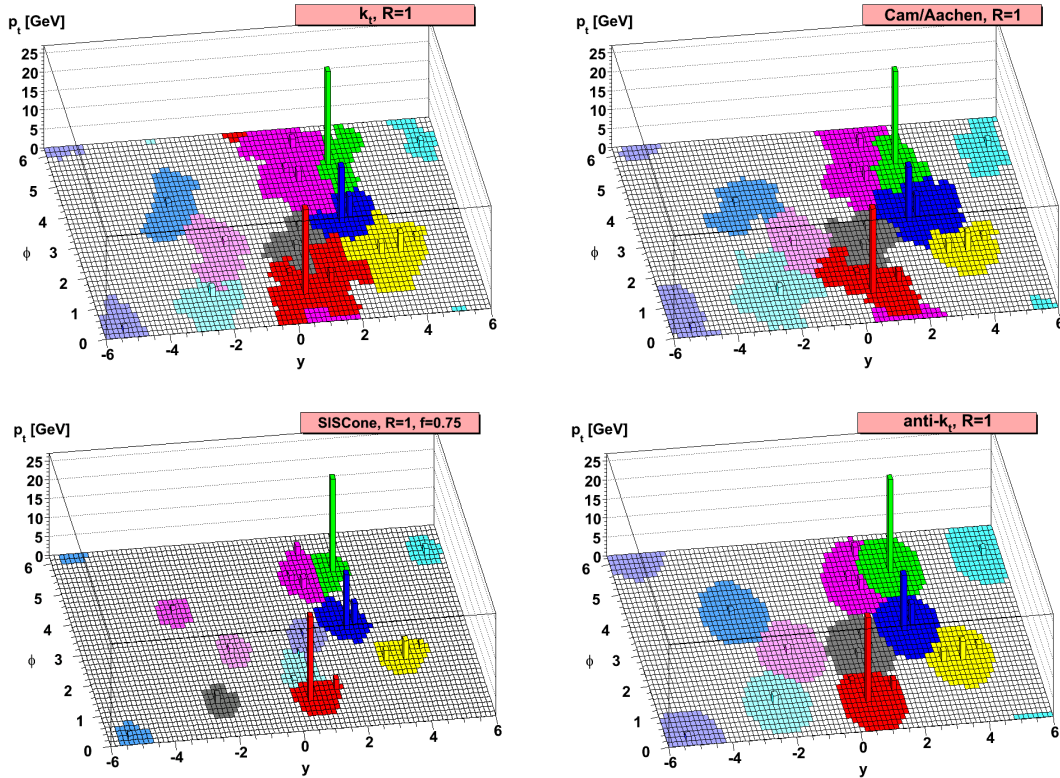
Figure 5.1: Taken from [69], an example of a parton-level event with soft particles clustered with four different jet clustering algorithms. The anti-k_t algorithm was used to cluster jets used in this analysis.

hard particle has no neighboring hard particles within a distance of $2R$, it will cluster with all the soft particles with a circle of radius R - a perfectly conical jet.

However, if another hard particle is present and $R < \Delta_{12} < 2R$, then there will be two hard jets and neither can be perfectly conical. The jet with the higher momentum particle will be reconstructed as conical, while the neighboring jet will not and will miss the part that overlaps with the conical jet. In the case where $k_{t1} \simeq k_{t2}$, both jets will only be partially conical.

The most important feature of the anti-k_t clustering algorithm is that only hard particles can change the shape of jets, soft particles do not. This causes the boundary of jets to be resilient to soft radiation, thus making it collinear safe, unlike the iterative

cone algorithm.

Energy Corrections

Generally, energy measured by the CMS detector is different from the true particle-level jet energy. Primarily, this is caused by the non-uniform and non-linear response of the calorimeters in CMS. Further complicating accurate measurements of jet energy is the presence of noise from electronics and additional proton collisions in the same bunch crossing. To make measurements of jet energy as accurate possible, corrections are made to the energy scale of reconstructed jets.

Two primary corrections are made. The first is a relative correction that removes variations versus jet η relative to a central, uniform, "control" region of the detector. The second correction is an absolute correction which removes variations versus jet p_{T}. Corrected jet energies are calculated as follows:

$$E_{\mathrm{corrected}} = (E_{\mathrm{uncorrected}} - E_{\mathrm{offset}}) \times C_{\mathrm{rel}}(\eta, p_{\mathrm{T}}'') \times C_{\mathrm{abs}}(p_{\mathrm{T}}') \tag{5.1}$$

where p_{T}'' is the transverse momentum of the jet corrected for offset, and p_{T}' is the transverse momentum of the jet corrected for offset and η dependence, or more explicitly $p_{\mathrm{T}}' = p_{\mathrm{T}}'' \times C_{\mathrm{rel}}(\eta, p_{\mathrm{T}}'')$.

The total jet energy correction factors fall in the range of ~ 1 to ~ 1.3 for particle-flow jets with $|\eta| < 2.5$. Further details about jet energy corrections and calibration are in [49].

5.4   Photon Reconstruction

Photon reconstruction begins with energy deposited in the ECAL of CMS. With the ECAL crystals in CMS having a Moliere radius of 2.2 cm (the same as the physical
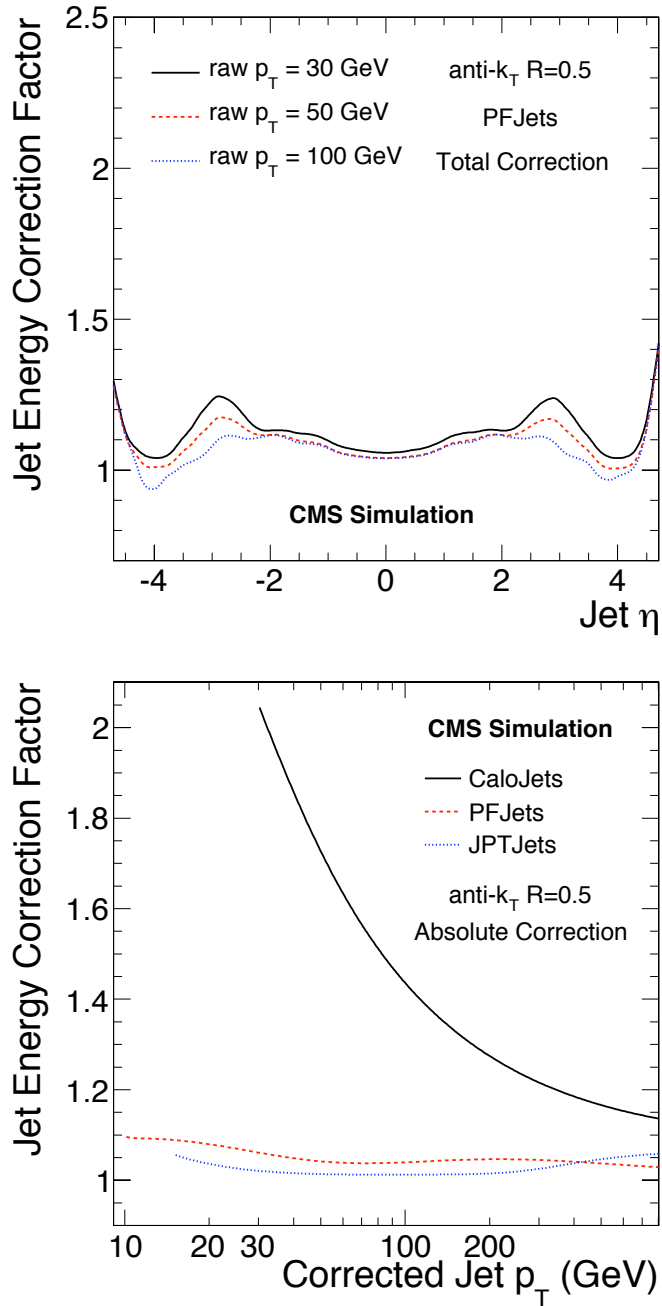
Figure 5.2: Total jet energy correction factors derived from simulation for particle flow jets at $\sqrt{s} = 7\,\text{TeV}$ from [49]. Total jet energy correction factors, $C(p_\text{T}, \eta)$, as function of jet η for different values of uncorrected jet p_T (top). Absolute jet energy correction factors, C_abs, as a function of corrected jet p_T for different jet reconstruction algorithms (bottom). CaloJets and JPTJets are two other jet clustering algorithms, but only PFJets are used in this analysis.

width of their front face), a photon with GeV of energy deposits 95% of it's energy into an array of 5×5 crystals. Material between the interaction region and the ECAL cause roughly half of direct photons to convert into e^+e^- pairs, resulting in a deposits of energy more spread along ϕ (due to the presence of the magnetic field from the solenoid).
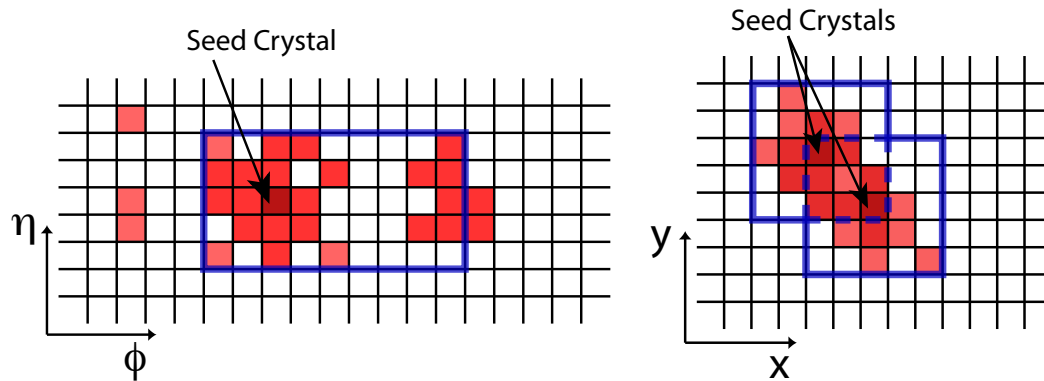
Figure 5.3: Example diagram of deposits of energy into crystals of the ECAL CMS clustered together with the Hybrid Supercluster Algorithm (left) and Multi5x5 Algorithm (right). The Hybrid Supercluster Algorithm is used in the Barrel, the Multi5x5 is used in the endcap.

The ECAL crystal arrays have different geometry in the barrel and endcap, and in addition, the magnetic field is different, so energy deposits is grouped together by different algorithms: a "Hybrid Clustering Algorithm" in the barrel, and a "Multi5 $\times$ 5" algorithm in the endcap [73].

The hybrid clustering algorithm results in clusters of energy roughly 5 crystals wide in η and up to 35 crystals wide in ϕ as depicted in figure 5.3 (left). The hybrid clustering algorithm works as follows:

1. The crystal with the highest E_T which is not a part of any other cluster becomes the seed to a new cluster. The E_T must be above a minimum threshold of 1 GeV.

2. Two crystals on each side of the seed crystal along η join the cluster (5 crystals wide in η, 1 crystal wide in ϕ), this is the original "domino" of this cluster.

3. Staying at the same η, begin stepping along ϕ and add these 5×1 dominos as long as energy within the domino is above a minimum threshold of 350 MeV.

4. Stop adding dominos once you have reached ± 17 crystals from original seed crystal.

5. Repeat starting from step 1.

In the endcap of CMS, due to the different geometry as described in section 3.2, the "Multi5 $\times$ 5" algorithm is used, and example of which is shown in figure 5.3 (right). The steps in the Multi5 $\times$ 5 algorithm are listed here:

1. The crystal with the highest E_T which is not a part of any other cluster becomes the seed to a new cluster. The E_T must be above a minimum threshold of 180 MeV.

2. The array of crystals 5×5 centered on the seed crystal join the cluster.

3. Crystals along the edge of this 5×5 array are allowed to be the seeds of new 5×5 clusters if they are a local maximum when compared to neighboring crystals.

4. Overlapping arrays of 5×5 crystals are joined together into one supercluster.

5. Repeat starting from step 1.

For both algorithms, the location of a supercluster is determined by a weighted average of all the positions in the supercluster. Each crystal (i) in the supercluster (SC) is given a weight: $w_i = \max\left(0, 4.7 + \log(E_i/E_{SC})\right)$. If the energy in the tower of the hadronic calorimeter directly behind the supercluster is low enough to give a
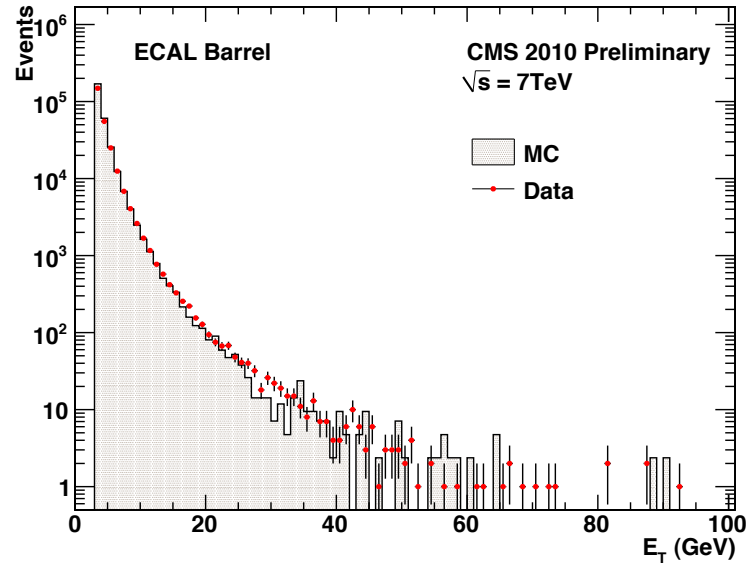
(a)
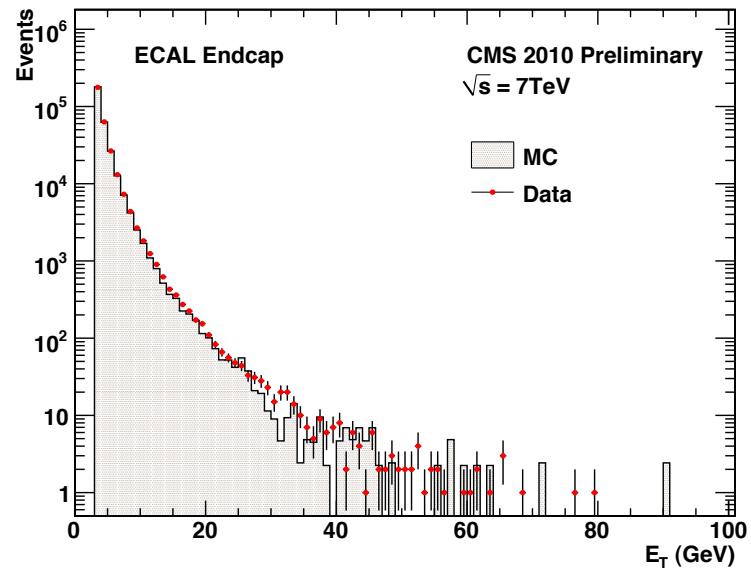
(b)

Figure 5.4: Photon objects are found after first building superclusters. Shown here is the spectra of transverse energy, E_{T}, for superclusters in the barrel (a), and endcap (b) of CMS from minimum bias events [74].

ratio of hadronic to electromagnetic energy of below 0.15, the supercluster is placed into the collection of photon candidates. The direction of the momentum of a photon candidate is determined by connecting a line from the primary vertex to the position of the supercluster.

Figure 5.4 shows the distribution of uncorrected energy for superclusters in the barrel and endcaps of CMS. Events used in figure 5.4 had to pass a trigger that required at least one bunch crossing within the interaction point, and other selection used to suppress beam-related backgrounds described in [74].

Figure 5.5 shows an event display for a typical γ+jet event. The photon is isolated from other energy in the tracker, ECAL, and HCAL, and can be seen opposite in ϕ to the jet. The jet appears as a collimated collection tracks (green), and energy in the ECAL (red) and HCAL (blue).
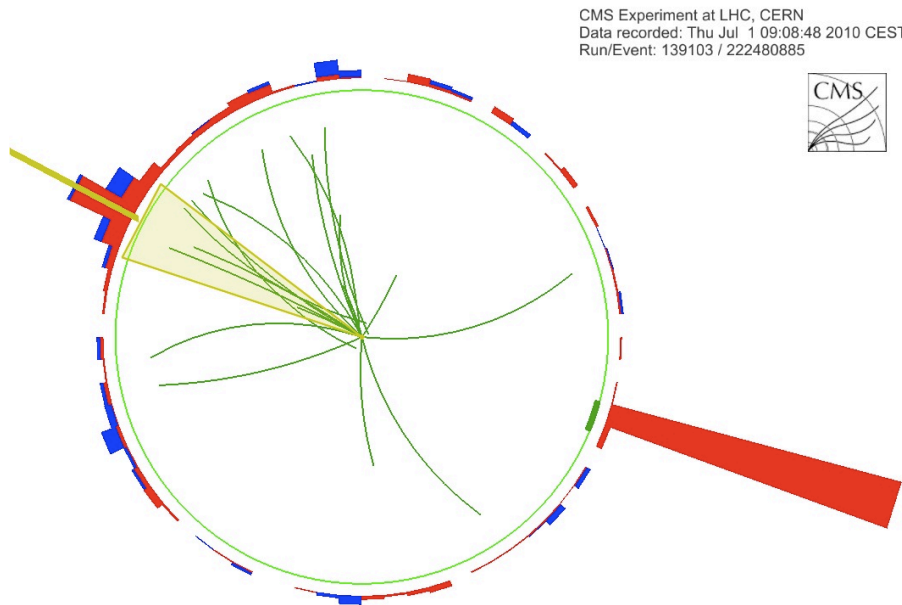
Figure 5.5: Event display of an example γ+jet event in r-ϕ view. The photon candidate is the isolated deposit of energy in the ECAL (red). The jet, opposite in ϕ is a mixture of deposits in the tracker (green), ECAL (red), and HCAL (blue).

5.5 Anomalous ECAL Energy Deposits

Ionization of Avalanche Photodiodes

Anomalous signals in the ECAL barrel are a source of background for detecting and measuring particles from a pp collision event. These energy deposits a believed to be caused by direct ionization of an avalanche photodiode by highly ionizing particles (such as protons) [74]. These deposits of energy are observed in single crystals and are referred to as "spikes". The rate of these spikes scale with $\sqrt{s}$ at a level consistent with the increase of charged particle multiplicity from collisions [75].

The pulse shape of a signal in the ECAL allows discrimination between energy deposits from true electromagnetic showers and those from direct ionization of the avalanche photodiode. For examples, signals produced by electromagnetic showers are peaked at a time of zero (relative to pp collision time), while the timing found for "spikes" is peaked at -10 ns. The time difference is due to the fact that electromagnetic shows take time to develop and reach a maximum, while direct ionization of the avalanche photodiode results in a much shorter rise-time. Figure 5.6 shows the distribution for a topological variable, $(1 - E4/E1)$, where $E4$ is the sum of energy in 4 adjacent crystals to the seed crystal, and $E1$ is the energy in the seed crystal. A value for $(1 - E4/E1)$ of near 1.0 indicates a deposit of energy localized almost entirely within a single crystal. Also shown in figure 5.6 is the reconstructed time of the seed crystal, t_{seed}. Even after a cut on the topological variable $(1 - E4/E1)$, a peak at -10 ns is still visible, and this is a result of spikes that are non-isolated and thus have lower values of $(1 - E4/E1)$.
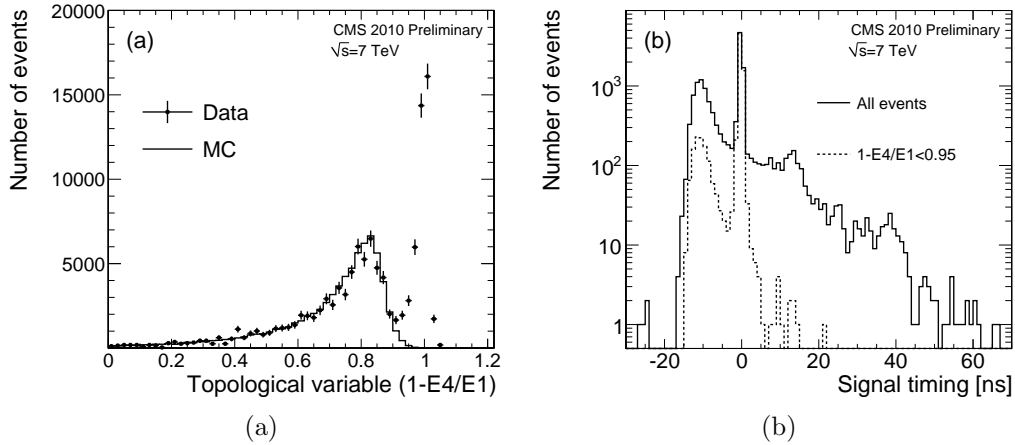
Figure 5.6: From [74], distribution of the the variable $(1 - E4/E1)$ for the highest energy deposit in each event for data and simulation (a). Values near 1.0 are energy deposits with almost all the energy in a single crystal. In (b) is shown the reconstructed time corresponding to the maximum of the signal pulse for the highest energy crystal.

Beam Halo

Outside of the CMS detector, protons from the LHC proton beams may collide with atomic nuclei of gas atoms or other material along the beam line [76]. These interactions produce a "halo" of particles which tend to be nearly parallel to the proton beam. These "halo" particles are made of muons, mesons, and baryons. Muons from beam halo can penetrate through large amounts of material including the endcaps of CMS. These muons may also interact with material in CMS and create photons from bremsstrahlung. The photons from these interactions will be detected by the ECAL and tend to be relatively isolated.

Figure 5.7 shows the timing of the seed crystal to superclusters vs their pseudo-rapidity. Photons which are the result of pp collisions at the interaction point are reconstructed with a $t_{\text{seed}} \approx 0$, while photons which are the result of interactions from beam halo tend to have a negative t_{seed}.
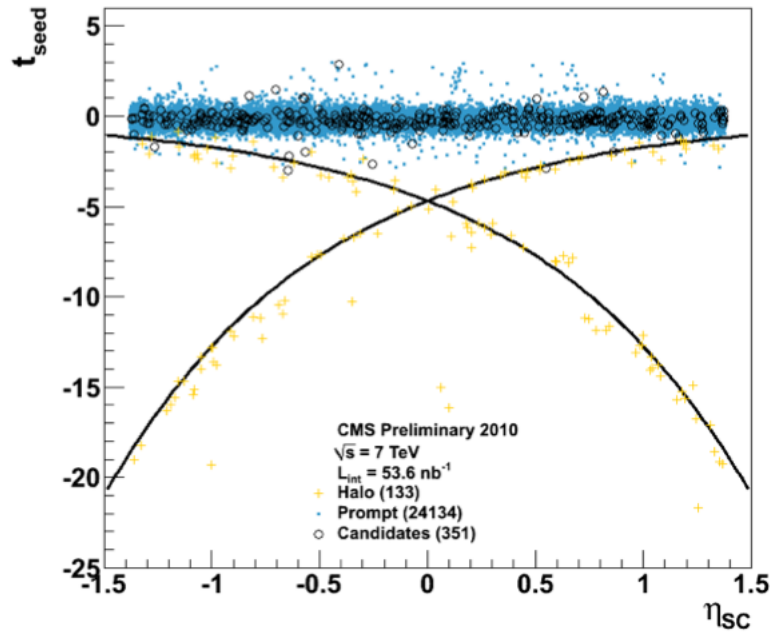
Figure 5.7: From [76], reconstructed time ($t_{\rm seed}$) of the for superclusters vs pseudorapidity. Photons from a *pp* collision are reconstructed with $t_{\rm seed}$ near 0, while superclusters arising from beam halo have negative $t_{\rm seed}$.

Chapter 6

Event Selection

The peak instantaneous luminosity of the LHC at the end of the run during 2010 was $\mathcal{L} = 2.07 \times 10^{32} \, \mathrm{cm}^{-2} \, \mathrm{s}^{-1}$. A collision event requires on the order of $200 \, \mathrm{kB}$ to record relevant all data. The rate of proton collisions at the interaction point in CMS is on the order of MHz. This is far too much data to record and process offline, and so this rate is first reduced by the CMS Trigger, which essentially acts as a filter to record the events most relevant for physics analysis. The Trigger system reduces the rate of recorded events from MHz down to $\sim$300 Hz. After the trigger, selections are made offline to select high quality events with jets and an energy-isolated photon.

6.1   Trigger

The first level of event filtering at CMS is done by custom hardware and is designed to filter an event rate of 40 MHz down to the rate of 300-400 kHz. It is described in general in 3.2, while the selection of photon trigger candidates is described here.

Photon and electron candidates as well as transverse energy sums per calorimeter region are determined by the Regional Calorimeter Trigger (RCT) [77]. Candidates

are found with an algorithm that uses a 3×3 trigger tower sliding window centered on trigger towers out to an $|\eta| = 2.5$. Each trigger tower has an (η, ϕ)-coverage of 0.087×0.087 up to an $|\eta| = 1.74$, beyond which trigger tower size increases with increasing η. Level-1 photon candidates are created when the E_T sum in neighboring pairs of trigger towers is above a threshold [78]. The threshold was $5\,\mathrm{GeV}$ in the earliest stages of data taking, but increased to incrementally to $12\,\mathrm{GeV}$ as luminosity increased. The 4 most energetic isolated and nonisolated candidates with the highest E_T above the minimum threshold from the Level-1 trigger are passed to the global trigger which decides whether to keep or throw away an event. Events that pass move next to the High level Trigger.

Selection for photons by the High Level Trigger relies on information from only the ECAL and HCAL subdetectors. Photon candidates are constructed using the same algorithm used offline (described in 5.4) to cluster energy deposits in the ECAL. Loose requirements are placed on these photon candidates to reduce jets faking photons, and reduce anomalous signals caused by interaction of heavily ionizing particles in the APD of the CMS barrel [74]. The requirements of the single-photon HLT paths are on the hadronic to electromagnetic ratio ($H/E < 0.15$), and and on a ratio of energy in the crystals that make up the photon's supercluster ($E_{3\times3}/E_{\mathrm{SC}} < 0.98$). See section 5.4 for a description superclusters. In addition, the photon candidate is required to have an E_T greater than a threshold. This threshold changed due to increasing luminosity of the LHC, from 20 GeV in the early runs up to 70 GeV in the later runs:

Run Range	HLT Path Name	Int. Lumi (pb^{-1})
138564 - 143962	HLT_Photon20_Cleaned_L1R	2.46
144010 - 147116	HLT_Photon30_Cleaned_L1R	5.81
147196 - 148058	HLT_Photon50_Cleaned_L1R_v1	9.47
148822 - 149294	HLT_Photon70_Cleaned_L1R_v1	18.4

Table 6.1: Summary of the single photon triggers used and corresponding run ranges. The number in the path names refer to the minimum E_T threshold, and "Cleaned" refers to the fact that there is show-shape cut.

6.2   Fiducial Cuts

Selection requirements were placed on the location of photon and jet candidates to be within the acceptance of certain subdetectors. Photon reconstruction is problematic in the border region between the ECAL barrel and endcap. For photons, it is important to be able to accurately determine energy and isolation, so to be within the acceptance of the ECAL and Tracker subdetectors, photons were required to have $\eta < 1.442$ or $1.566 < |\eta| < 2.5$. For jets, which were reconstructed with the particle flow algorithm using information from all subdetectors, it is important to be able to determine its constituent particles to be able to more accurately determine its transverse momentum. Jets were required to be have $|\eta| < 2.4$ to be well within acceptance of the Tracker.

6.3   Vertex

Knowledge of the location of the primary event vertex is important for determining the outgoing angle of the hard scattering photon, which affects the accuracy of the E_T measurement of the photon. In addition, events with a vertex far from the nominal interaction may be from anomalous interactions (see section 5.5). Figure 6.1 shows an example distribution of the z-location of the primary vertex in typical pp collisions.

To select high-quality events, requirements are placed on the primary interaction vertex of the event. The location of the vertex along the z-axis (parallel to the beam line) is required to be within ±24 cm of the nominal interaction point. A good hard-collision event has a high number of degrees of freedom due to the emergence of many tracks from the collision point, hence a minimum number of degrees of freedom on the primary vertex was placed at 4 on the reconstructed vertices. To filter out beam-scraping events, at least 10 tracks with $p_\mathrm{T} > 0.25\,\mathrm{GeV}$ were required to be present.
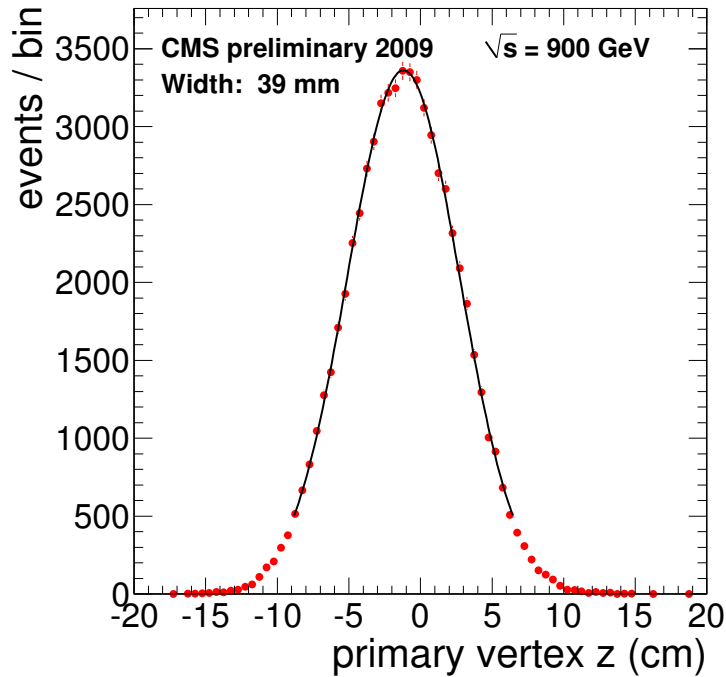
Figure 6.1: Distribution of the z-location of the primary vertex from a single run [65].

6.4   Jet Identification

Jets used in this analysis are reconstructed using the particle flow method [79]. This method identifies particles such as charged hadrons, neutral hadrons, electrons, pho-

tons, using information from all sub-detectors within CMS. These particles are labeled as "particle-flow candidates" and then are clustered into jets with the anti-k_t cluster algorithm [69] with a cone size of $\Delta R = 0.5$. Jet energy corrections [80] are applied to correct p_T and make the jet response flat vs η. Jets are only counted if they have $p_T > 30\,\text{GeV}$, are within the acceptance of the CMS tracker ($|\eta| < 2.4$), and are further than $\Delta R = 0.5$ from the leading high-p_T photon. Very loose identification is applied to jets, described in [81] and listed here:

- Charged Hadron Fraction > 0.0

- Neutral Hadron Fraction < 1.0

- Charged Multiplicity > 0.0

- Charged Electromagnetic Fraction < 1.0

- Neutral Electromagnetic Fraction < 1.0

The above requirements are placed on the collection of particle-flow candidates that contained within the jet. These requirements are very loose and primarly remove jets arising from anomalous signals (see section 5.5) or noise in one sub-detector.

6.5   Photon Identification

The largest background to direct photons is from jets which fragment in such a way that one or a few, light mesons such as the π^0 carry most of the energy. These mesons primarily decay into two photons (the branching fraction of $\pi^0 \to \gamma\gamma$ is 98.8%). If the meson has enough energy, the photons may be boosted enough to deposit a narrow deposit of energy in the ECAL, and thus resemble the deposit of a single direct photon.

To discriminate between direct photons and jets, we employ several variables. Some involve the shape of the energy deposit in the ECAL, others involve the sum of energy in a cone centered around the main energy deposit. A "photon" which arises from a jet has a higher probability of containing other tracks or energy deposits nearby. It tends to not be as isolated as direct photons.

Variables used to discriminate between direct photons are photons which arise from jets are described in the following subsections.

Energy Cluster Transverse Profile

A shower-shape variable of the energy deposit in the ECAL which is useful in discriminating true prompt photons from background is $\sigma_{i\eta i\eta}$. This is a modified second moment of the energy deposit in the ECAL. It is computed about the mean η position using logarithmic weights and is defined as:

$$\sigma_{i\eta i\eta}^2 = \sum_{i=1}^{25} w_i(\eta_i - \bar{\eta}_{5\times5})/\sum_{i=1}^{25} w_i$$

where the sums are over the 5×5 crystals centered on the seed crystal of the photon, w_i are the weights of each crystal, and are equal to:

$$w_i = \max\left(0, 4.7 + \ln\frac{E_i}{E_{5\times5}}\right).$$

The energy of each crystal is E_i, and the η position of each crystal is η_i. The variable $E_{5\times5}$ is the sum of energy in the 25 crystals, and $\bar{\eta}$ is the average η position of the same set of crystals. This variable, $\sigma_{i\eta i\eta}$, measures the extent of the energy deposit in the η direction. Photons produced directly in hard scattering leave deposits of energy in the ECAL, and the distribution of $\sigma_{i\eta i\eta}$ is at relatively small values and narrow. Deposits of energy in the ECAL which are isolated, but which arise from hadronic

decays have a distribution of $\sigma_{i\eta i\eta}$ which has a much wider distribution, with a tail towards relatively larger values.

The magnetic field of CMS causes bending for charged particles along the ϕ direction, so even if a photon converts ($\gamma \rightarrow e^+e^-$), its shower profile along the η direction is not disturbed. In addition, $\sigma_{i\eta i\eta}$ distributions were found to be unaffected by the number of reconstructed vertices in the event (and thus unaffected by event pileup).

Cluster Energy Profile: ECAL vs HCAL

Photons and electrons deposit most of their energy in the ECAL, whereas hadrons deposit most of their energy in the HCAL. A variable that is useful for identification of photons is the hadronic to electromagnetic fraction (H/E): the ratio between energy in the HCAL and the energy in the ECAL in a cone of radius $\Delta R = 0.15$ centered on the photon. The ECAL has ~ 25 radiation lengths of material, so direct photons deposit almost all of their energy in the ECAL itself. A significant amount of energy in the HCAL directly behind the ECAL is evidence of the presence of other particles in association with the candidate photon.

Track Isolation

Labelled as Iso_{TRK}, this quantity is the sum of transverse momentum, p_{T}, of all tracks in a cone of radius $\Delta R = 0.4$. This isolation cone has its tip at the event's primary vertex and the other end centered on the ECAL energy deposit. Iso_{TRK} has values close to 0 for photons from a hard scattering event, while photons which are embedded in jets have higher values. To keep this quantity low for converted photons as well, tracks within $\Delta R = 0.04$ or $\Delta \eta = 0.015$ are not included in the sum.
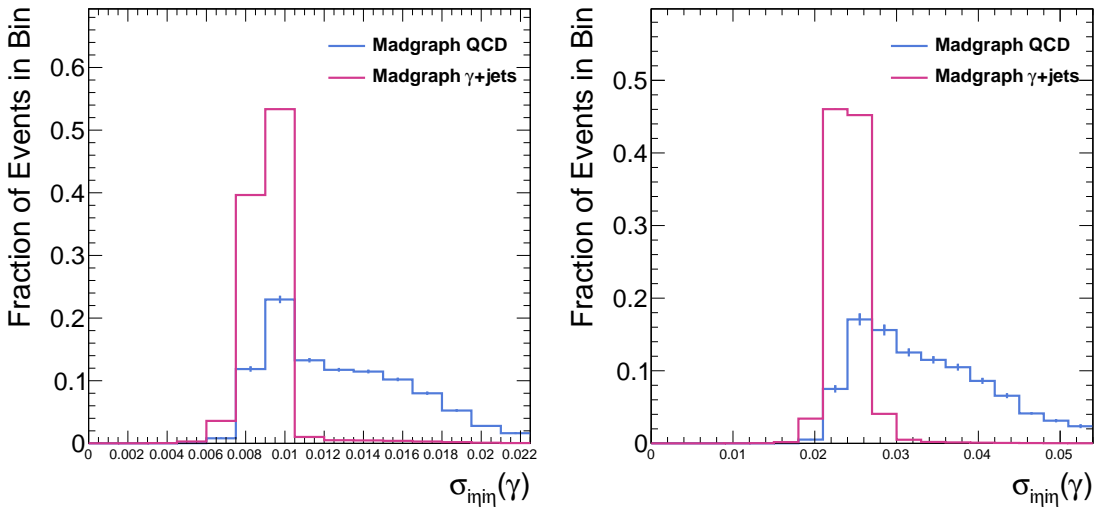
Figure 6.2: Area-normalized distribution of $\sigma_{i\eta i\eta}$ for leading photon in signal and background monte carlo. Shown, are photons in the barrel (left), and photons in the endcap (right)
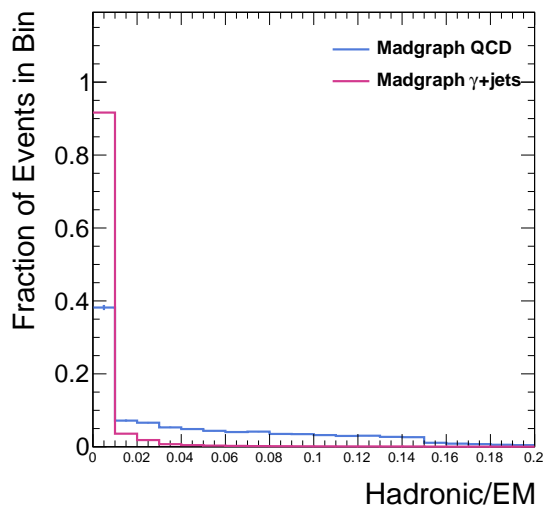
Figure 6.3: Area-normalized distribution of the ratio of hadronic to electromagentic energy for leading photon in signal and background monte carlo.

Figure 6.4: Area-normalized distribution track isolation sum for leading photon in signal and background monte carlo.

ECAL Isolation

Labeled as Iso_{ECAL}, this quantity is the sum of transverse energy E_{T} in crystals located within a radius of $\Delta R < 0.4$ centered on the photon's super cluster. Iso_{ECAL} has values close to 0 for photons from a hard scattering event, while photons which are embedded in jets have higher values. To keep this sum low for converted photons, crystals within a radius of $\Delta R = 3.5$ or within $\Delta\eta = 2.5$ crystals of the seed crystal are not included in the sum.

HCAL Isolation

Labeled as Iso_{HCAL}, this quantity is the sum of transverse energy in the HCAL towers in a cone of radius $\Delta R = 0.4$around the photon's supercluster. True photons will deposit their energy in the ECAL, and even if they convert into an e^+e^- pair, those electrons will deposit their energy in the ECAL. Thus, photons from a hard scattering have values of Iso_{HCAL} near 0, while photons embedded within jets have higher values.

Energy within $\Delta R = 0.15$ is not included within the sum as it is already used in the H/E variable.

Loose Photon Isolation Selection

The photon isolation sums defined above are used together for a set of requirements for photon isolation identification. The exact values of the cuts used are the same used in the CMS measurement of the inclusive isolated photon production [42] and are listed in table 6.2.

The isolation requirements remove the majority of fake photons. However, a substantial background remains primarily from fragmentation of partons into neutral mesons which may be isolated and decay into two photons and leave a narrow deposit of energy - much like a true photon from the hard scattering event. The cluster shape variable $\sigma_{i\eta i\eta}$, described above, is used to measure the amount of true photons in the selected sample. This process, finding the purity of the photon sample, is described in the next chapter, in section 7.2.

To remove events where the photon is the result of interactions beam halo or "spikes" (described in 5.5) a cut was placed on a shower shape variable, $\sigma_{i\phi i\phi}$ (defined similarly to $\sigma_{i\eta i\eta}$, but in the ϕ plane) and the reconstructed time of the most energetic crystal (t_{seed}). These requirements are described in more detail in [82] and are 2 dimensional: $\sigma_{i\phi i\phi} > 0.009$ or $t_{\text{seed}} > -1.5$. This cut is 100% for monte carlo signal photons, but removes 99% of photons arising from beam halo.

The number of events remaining after the stages of selection are shown in table 6.3. The isolation requirements remove a significant amount of events where the photon candidate is actually jet itself or a photon within a jet.
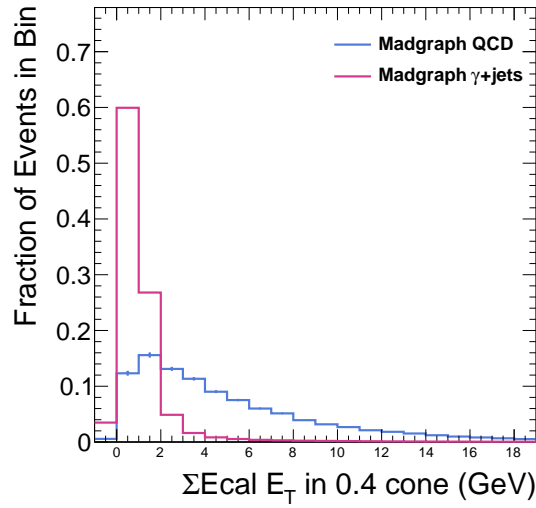
Figure 6.5: Area-normalized distribution ECAL isolation sum for leading photon in signal and background monte carlo.
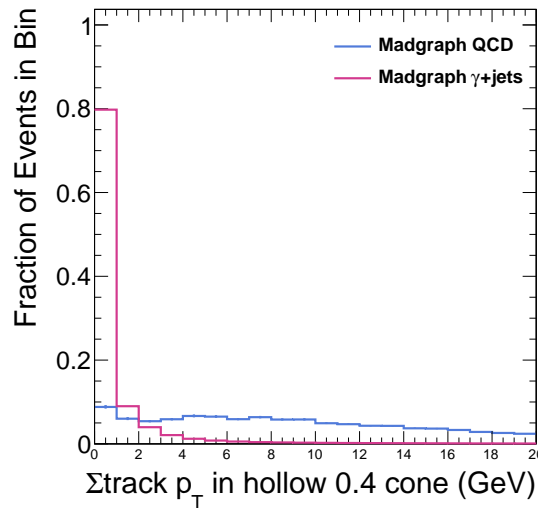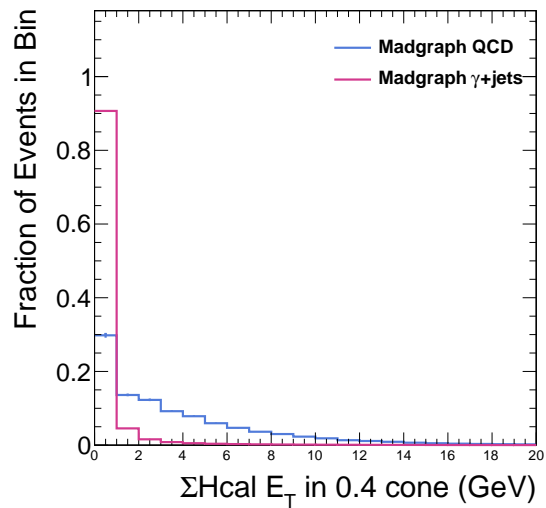

Figure 6.6: Area-normalized distribution HCAL isolation sum for leading photon in signal and background monte carlo.

Variable	Cut
$Iso_{\mathrm{TRK}} <$	$2.0\,\mathrm{GeV}$
$Iso_{\mathrm{ECAL}} <$	$4.2\,\mathrm{GeV}$
$Iso_{\mathrm{HCAL}} <$	$2.2\,\mathrm{GeV}$
$H/E <$	0.05

Table 6.2: Selection cuts for loose photon isolation identification.

Selection	Number of remaining data events
Events in original "Photon 2010" dataset	25 M
Trigger, Vertex, Photon p_{T} and η selection	645,867
Photon passing Isolation	51,905
Jet passing p_{T} and loose selection	42,248

Table 6.3: Remaining number of events in data after selection stages. The name of the original 2010 dataset containing events that had passed any of the active photon triggers is "/Photon/Run2010B-Dec22ReReco_v1/AOD".

Chapter 7

Analysis

As mentioned in section 1.11, production of vector bosons in association with jets has been studied for the W and Z. This thesis presents results of high-p_{T} photon production in association with jets. The methods for identification of photons and performing of signal extraction are based on the methods used in the measurement of the differential isolated prompt photon production cross section with CMS [83, 84]. The data in this analysis uses an integrated luminosity of $36\,\mathrm{pb}^{-1}$, which corresponds to all the the 2010 data sample where consistent single photon triggers were available. Systematic uncertainties (such as integrated luminosity) are reduced or removed completely by measuring ratios of cross sections of γ+jet, such as $\sigma(\gamma+ \geq n\mathrm{jets})/\sigma(\gamma+ \geq (n-1)\mathrm{jets})$. Jet-production ratios such as these depend strongly on the p_{T} requirements of the vector boson [20]. For this analysis we require the photon have to $p_{\mathrm{T}} > 75\,\mathrm{GeV}$.

7.1 Comparison of Data to Monte Carlo

Distributions of basic variables for the highest p_{T} photon are shown in figure 7.1. MADGRAPH is scaled up by an arbitrary factor 1.7 to better compare the shape of distributions between data and monte Carlo. After arbitrary normalization, the data is reasonably modeled by the sum of MADGRAPH simulations of γ+jet and QCD.

7.2 Signal Extraction

To extract the yield of true photons in the data, we use a shower shape variable, $\sigma_{i\eta i\eta}$. This transverse shape of the cluster of energy in the ECAL is defined in section 6.5. This shower shape variable, $\sigma_{i\eta i\eta}$, tends to be small for true photons, but larger for other types of energy deposits (like $\pi^0 \to \gamma\gamma$).

The photon yield is obtained by fitting this distribution as a two-component sum of signal (γ+jets) and background (QCD). The fit was performed using an Extended Maximum Likelihood fit. The distribution is fit over the range $0.006 < \sigma_{i\eta i\eta} < 0.02$ in the barrel, and the range $0.02 < \sigma_{i\eta i\eta} < 0.05$ in the endcap. The signal shape is taken from Monte Carlo, while the background shape is taken from a side-band selection on data in a similar manner as done in [82]. This side-band selection is same as the signal selection except for a different requirement on track isolation: $2.0 < Iso_{\mathrm{TRK}} < 5.0\,\mathrm{GeV}$. This track isolation variable, Iso_{TRK}, is chosen because it has the smallest correlation with the other photon selection variables. Different bounds on the track isolation for the sideband selection were used, and the effect on the final results is smaller than the effect of uncertainty on the jet energy scale, see section 8.1.

Data is binned by location of the lead photon (barrel or endcap) and by the
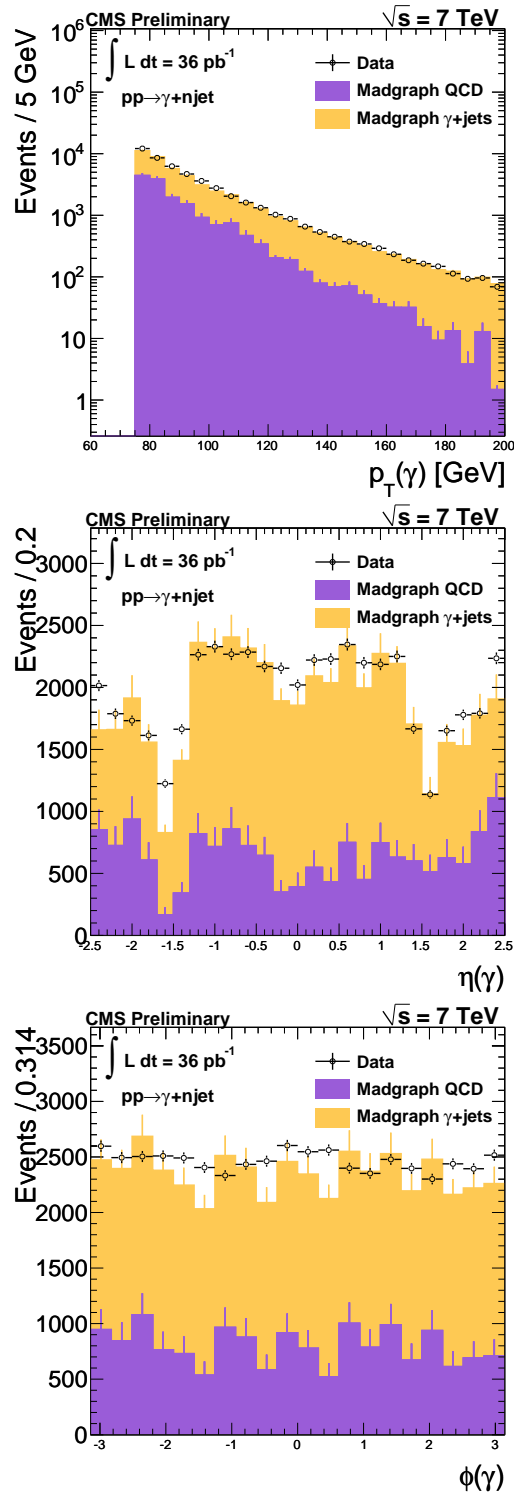
Figure 7.1: Distributions of the leading photon's p_{T}, η, and ϕ in data compared to MadGraph Monte Carlo. The dip near $|\eta| = 1.5$ is due to fiducial cuts near the ECAL gap between the barrel and endcap.

exclusive number of jets found. Then the lead photon's $\sigma_{i\eta i\eta}$ distribution in each of those bins from data is fit using a binned likelihood function of signal and background distributions. Figures 7.2 and 7.3 show distributions for $\sigma_{i\eta i\eta}$ for barrel and endcap photons. The jet multiplicity distributions are then scaled down to the amount of signal determined by the result of the fit. For the two highest jet multiplicities when the photon is in the endcap, too little data is available to do reliable fits and so we use the average of the signal fraction found for the lower jet multiplicity bins, 56%. Uncertainty in signal fraction has a smaller effect on systematics of the final results than uncertainty on jet energy scale.

N Jets	Barrel	Endcap
0	52 ± 2	55 ± 2
1	68 ± 1	56 ± 1
2	73 ± 2	51 ± 2
3	73 ± 4	60 ± 6
4	84 ± 8	-
5	80 ± 3	-

Table 7.1: Signal fraction in data of photons as determined by fits to the lead photon's $\sigma_{i\eta i\eta}$. Data was first binned by exclusive number of jets in the event that have $p_T > 30\,\mathrm{GeV}$, and then the fits were performed. Photons have $p_T > 75\,\mathrm{GeV}$ and pass loose isolation ID.

7.3   Acceptance

Measurements presented in this thesis are done within the photon and jet phase space regions (in $|\eta|$ and p_T) described previously. To provide model independent results, we do not correct for acceptance of the detector. Correcting cross sections for acceptance necessarily relies on the use of Monte Carlo simulations, and we wish to avoid the theoretical uncertainty in the model used to calculate the acceptance.

Figure 7.2: Measured $\sigma_{i\eta i\eta}$ distributions for photons in the barrel for different numbers of jets plotted for data, sideband (data), and MADGRAPH MC. Figure (a) is $\gamma + 1$ jet, (b) $\gamma + 2$ jets, etc.
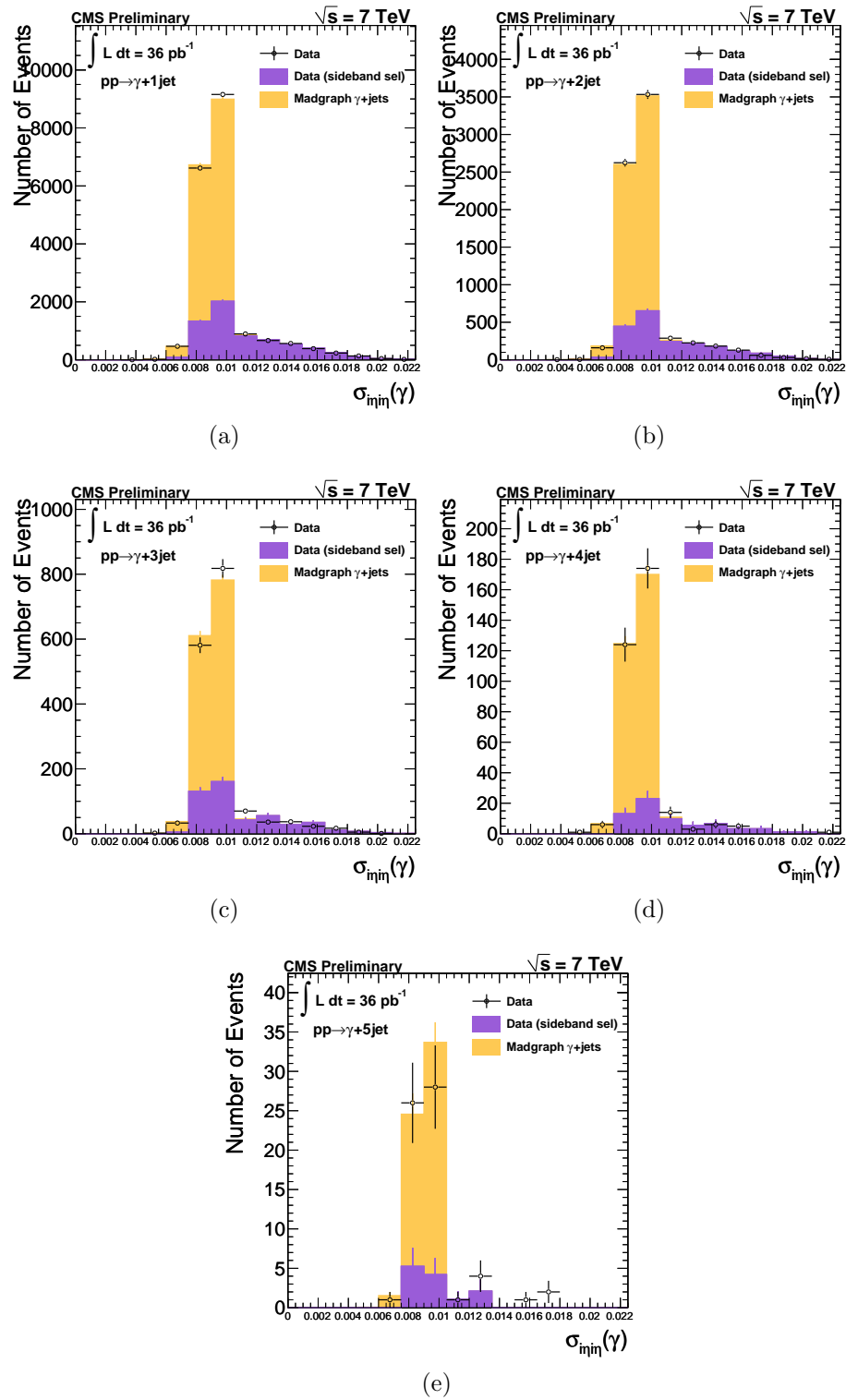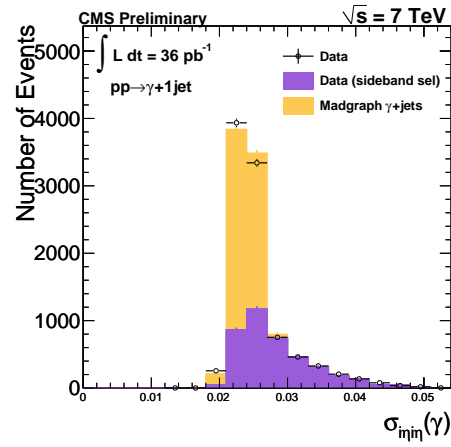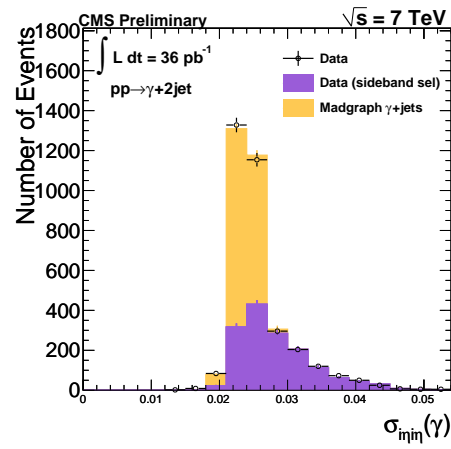
(a)

(b)
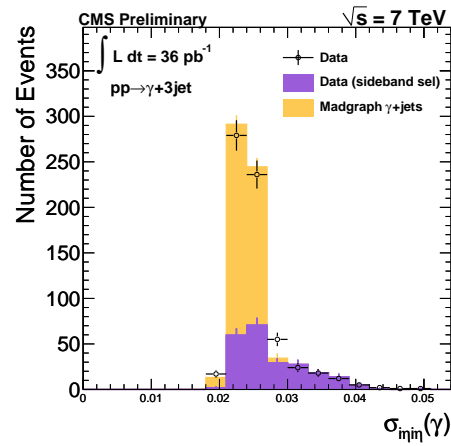
(c)

Figure 7.3: Measured $\sigma_{i\eta i\eta}$ distributions for photons in the endcap for different number of jets plotted for data, sideband (data), and MADGRAPH MC. Figure (a) is $\gamma + 1$ jet, (b) $\gamma + 2$ jets, etc.

7.4 Efficiency

Efficiencies are determined in events with a lower p_T lead photon ($p_\mathrm{T} > 30\,\mathrm{GeV}$) in both data and Monte Carlo. Due to the high p_T requirement on the photon ($p_\mathrm{T} > 75\,\mathrm{GeV}$), statistics are too limited to use the Tag and Probe method on $Z \to e^+e^- + \mathrm{jets}$ events directly to obtain efficiency corrections. MadGraph Monte Carlo gives results similar to data, which gives confidence in using Monte Carlo corrections for events with higher p_T photons. The systematics on the final results due to using Monte Carlo corrections from pythia vs MadGraph is smaller than the systematics due to uncertainty on the jet energy scale.

First, we begin with events that contain at least two lepton candidates which have an invariant mass in the range 60-120 GeV. One lepton, the "tag", is required to pass standard tight electron identification requirements (discussed below and listed in table 7.2). The other lepton, the "probe", is required to pass selection to at least be a photon candidate. The yield of signal events is obtained before and after requiring the probe to pass cuts to determine efficiency of the selection. To determine signal yield, the invariant mass distribution of the tag-probe pair is fitted assuming a two component contribution: signal and background. The efficiency is simply the ratio of the signal yield after selection to signal yield before selection. Photon efficiency is the product of three parts: reconstruction efficiency, isolation efficiency, and trigger efficiency.

Tag electrons are required to pass isolation requirements, and to pass cuts that remove electrons arising from photon conversions ($\gamma \to e^+e^-$). These requirements are described in the following paragraphs.

Three types of isolation are applied and are described in detail in section 6.5. Track isolation, Iso_TRK is the sum of p_T of tracks with $p_\mathrm{T} > 0.7\,\mathrm{GeV}$, each within an

annulus of $0.04 < \Delta R < 0.3$ centered on teh track direction at the vertex. ECAL isolation, Iso_{ECAL} is the sum of E_{T} of recontructed hits within a cone of $\Delta R < 0.3$ but with $\Delta \eta > 3$crystals centered around the ECAL supercluster. Only hits with $E > 0.08\,\mathrm{GeV}$ or $E_{\mathrm{T}} > 0.1\,\mathrm{GeV}$ in the endcap are included in the sum. HCAL Isolation is the sum of E_{T} of HCAL towers within $0.15 < \Delta R < 0.3$ centered around the supercluster.

To remove electrons from conversions, it is required that there are no missing hits in the tracker between the event vertex and the first measured hit of the reconstructed electron track. True electrons from the hard scattering at the vertex will have hits starting with the first layer in the tracker. Photons which later convert to electrons will not have hits in one or more layers of the tracker. Additionally, electrons from photon conversions will have a pair electron of opposite sign. The variable $\Delta \cot(\phi)$ is the difference in the angle between the two electron tracks. Another variable $dist$, is the distance of closest approach between the two tracks. If both $\Delta \cot(\phi)$ and $dist$ are less than 0.02 or there is any missing hits, the electron is rejected as arising from a converted photon.

Variable	Barrel	Endcap
$Iso_{\mathrm{TRK}}/E_{\mathrm{T}} <$	0.09	0.04
$Iso_{\mathrm{ECAL}}/E_{\mathrm{T}} <$	0.07	0.05
$Iso_{\mathrm{HCAL}}/E_{\mathrm{T}} <$	0.10	0.025
Missing hits $=$	0	0
$\Delta \cot(\phi) >$	0.02	0.02
$dist >$	0.02	0.02
$\sigma_{i\eta i\eta} <$	0.01	0.03
$\Delta\phi_{\mathrm{in}} <$	0.06	0.03
$\Delta\eta_{\mathrm{in}} <$	0.004	0.007
$H/E <$	0.04	0.025

Table 7.2: Maximum values for variables of the "tag" electrons used in the Tag-and-Probe method. Referred to within the CMS collaboration as working-point 80 selection (WP80) for elections because 80% of true electrons pass this selection.
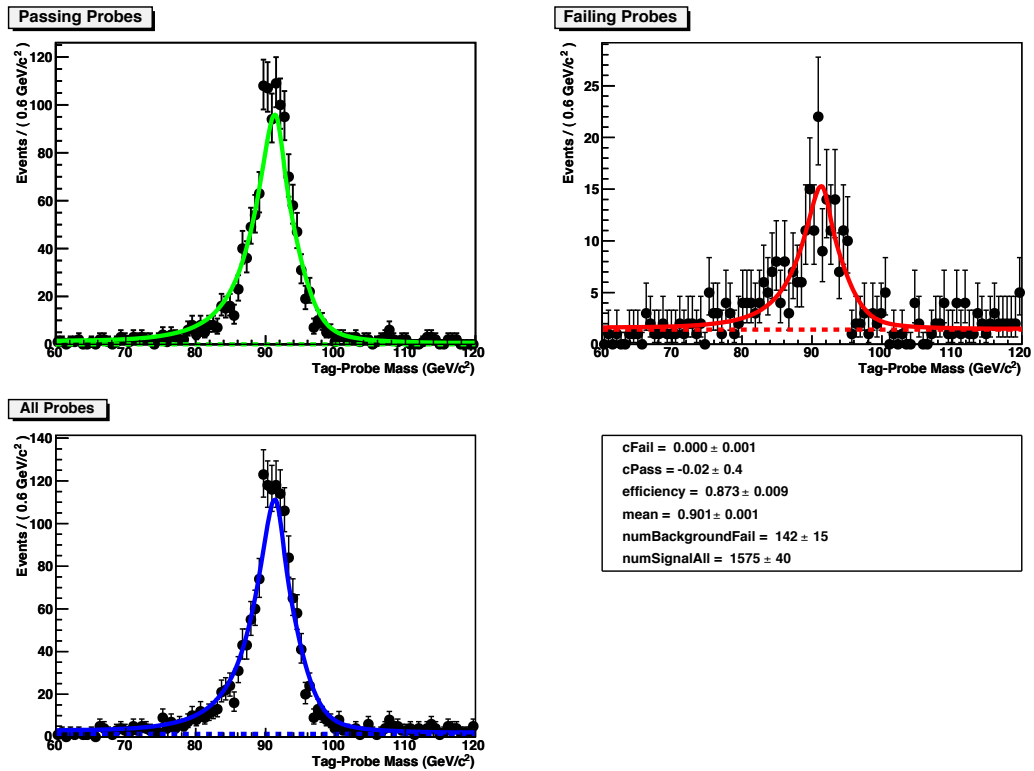
Figure 7.4: Example Tag-and-Probe results. The "tag" electron was required to pass WP80 selection, while the "probe" electron was treated as a photon and tested for passing loose photon selection.

The Trigger efficiency is $\sim100\%$ in our events, as a photon with $p_\mathrm{T} > 75\,\mathrm{GeV}$ is required to be present - well above the p_T turn-on for all triggers used. Reconstruction efficiency is also very high, because photon reconstruction only requires a clustered deposit of energy in the ECAL with a total $E_\mathrm{T} > 10\,\mathrm{GeV}$, far below our offline cut of 75 GeV. Thus the only efficiency that must be corrected for is photon isolation efficiency.

Photon isolation efficiencies are obtained as a function of the jet multiplicity in the event and are shown in table 7.3 for when the photon is in the barrel, and table 7.4 for when the photon is in the endcap. Efficiency of photon isolation selection is lower in the endcap compared to the barrel due primarily to the presence of more particles

in the endcap. Photon isolation efficiencies for higher p_T photons are obtained from Monte Carlo and shown in tabel 7.5. Efficiencies from MadGraph Monte Carlo are used to correct the data. There is not enough data at higher multiplicities to determine the efficiency. The largest difference between data and Monte Carlo is roughly 4%, and this is used as a source of systematic uncertainty in the final result.

N Jets	Data T&P	MadGraph T&P (Z+jets)	MadGraph (γ+jets)
0	89 ± 0.3	91 ± 0.04	86 ± 0.3
1	87 ± 1	88 ± 0.1	86 ± 0.1
2	81 ± 3	85 ± 0.3	79 ± 0.2
3	-	84 ± 0.8	74 ± 0.4
4	-	-	68 ± 0.8

Table 7.3: Efficiency of loose photon isolation cuts in the barrel (for photons with $p_T > 30\,\text{GeV}$) as a function of of the number of jets in the event that have $p_T > 30\,\text{GeV}$ in data and Monte Carlo samples.

N Jets	Data T&P	MadGraph T&P (Z+jets)	MadGraph (γ+jets)
0	94 ± 0.9	93 ± 0.1	92 ± 0.3
1	90 ± 3	91 ± 0.3	89 ± 0.2
2	89 ± 7	89 ± 0.8	85 ± 0.4
3	-	90 ± 2	83 ± 0.8
4	-	-	80 ± 2

Table 7.4: Efficiency of loose photon isolation cuts in the endcap (for photons with $p_T > 30\,\text{GeV}$) as a function of of the number of jets in the event that have $p_T > 30\,\text{GeV}$ in data and Monte Carlo samples.

The efficiencies for photon isolation selection in the MadGraph Monte Carlo reasonably agrees with data. This is true for both the Z+jets and γ+jets sample. Thus, the efficiencies obtained from MadGraph were chosen to use to correct data.

N Jets	Barrel		Endcap	
	PYTHIA	MADGRAPH	PYTHIA	MADGRAPH
0	86 ± 0.8	77 ± 2	94 ± 0.9	90 ± 2
1	92 ± 0.3	81 ± 0.6	94 ± 0.5	87 ± 1
2	90 ± 0.5	74 ± 0.6	93 ± 1	82 ± 1
3	88 ± 1	69 ± 0.8	91 ± 3	80 ± 2
4	85 ± 3	64 ± 1	90 ± 8	79 ± 5
5	85 ± 7	59 ± 3	82 ± 14	72 ± 11

Table 7.5: Efficiency of loose photon isolation cuts (for photons with $p_\mathrm{T} > 75\,\mathrm{GeV}$) as a function of of the number of jets in the event that have $p_\mathrm{T} > 30\,\mathrm{GeV}$ in γ+jet Monte Carlo samples.

7.5 Unfolding

Due to finite detector resolution, there are migration effects which must be corrected for to find the true number of γ+jet events as a function of jet multiplicity. This migration is a variable smearing effect introduced by the detector. The Baysian unfolding method [85] is used in the RooUnfold package [86]. This algorithm uses a regularization parameter to prevent statistical fluctuations in the data from appearing as structure in the unfolded distribution. For this iterative algorithm, the regularization parameter is set to the recommendation of the RooUnfold authors at $k_\mathrm{Bayes} = 4$ is used.

The first step to unfolding involves building a response matrix that maps the true distribution to the measured one (both must be binned). This matrix is constructed from Monte Carlo samples. For a given number of events in a bin of the true distribution, T_j, the response matrix element R_{ij} gives the fraction of events that end up in the measured bin M_i. Unfolding applies this response matrix to measured data to find the "true" distribution.

To create the response matrix, the generated photon was required to have $p_\mathrm{T} > 75\,\mathrm{GeV}$, $|\eta| < 2.5$ and not in $1.4442 < |\eta| < 1.566$. Jets were only counted if they

were above the p_T threshold of 30 GeV and had $|\eta| < 2.4$. The difference between the η requirements for the photon and jets is because accurate measurement of jet p_T is dependent on being well within acceptance of the Tracker. Both generated and reconstructed PF jets used the same anti-k_T algorithm with cone size of $\Delta R = 0.5$.

Response matrices were obtained from both MadGraph and pythia Monte Carlo signal samples, and are shown in figure 7.5. The response matrices shown are row-normalized to sum to 1 to make comparison easier. The overall normalization of the response matrix does not effect the final results of the unfolding process. Before unfolding there were two jet multiplicity distributions: one for when the photon was in the barrel, and one for when the photon was in the endcap. The two jet multiplicity distributions were summed before unfolding.

The data, before and after the unfolding process is shown in figure 7.6. The effect of the unfolding on the distribution of jet multiplicities ranges from a few % for the 1 and 2-jet bins, and 50% or more for the 5-jet bins or higher. The effect of using a different response matrix (from pythia or MadGraph) is shown in figure 7.7.
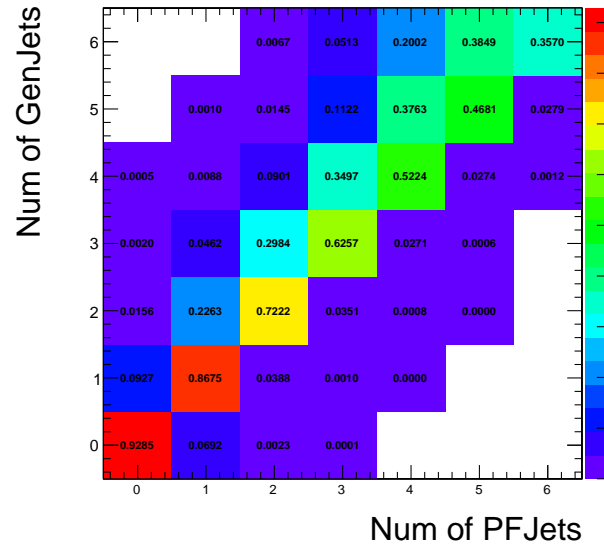
(a)
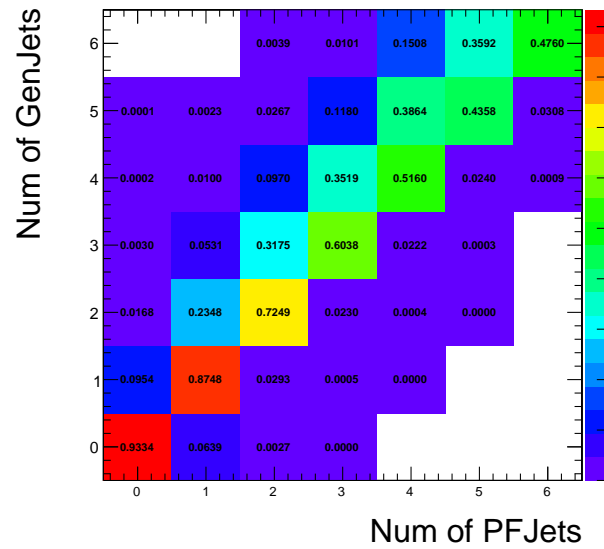
(b)

Figure 7.5: Response matrix from simulated γ+jet events in bins of exclusive jet multiplicity for jets from MADGRAPH (a) and PYTHIA (b). Generated jets and reconstructed particle flow jets were only counted if they had $p_{\mathrm{T}} > 30\,\mathrm{GeV}$. Rows are normalized to sum to 1 to make comparison easier.

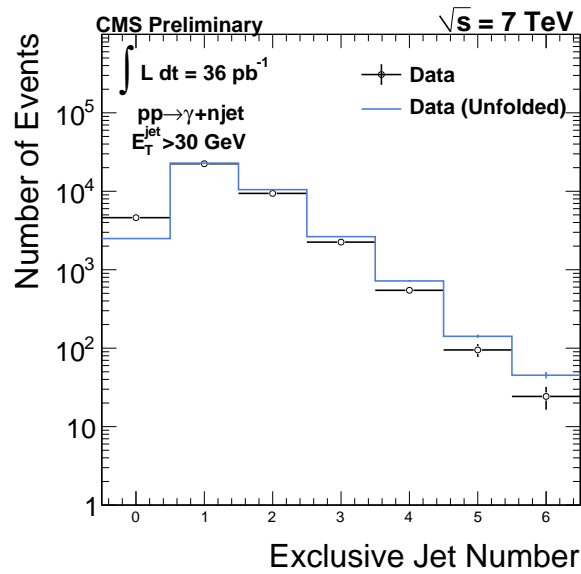Figure 7.6: Jet multiplicity distribution before and after unfolding. The unfolding process has the most effect at the larger jet multiplicities (i.e. 50% for multiplicity of 5 jets).
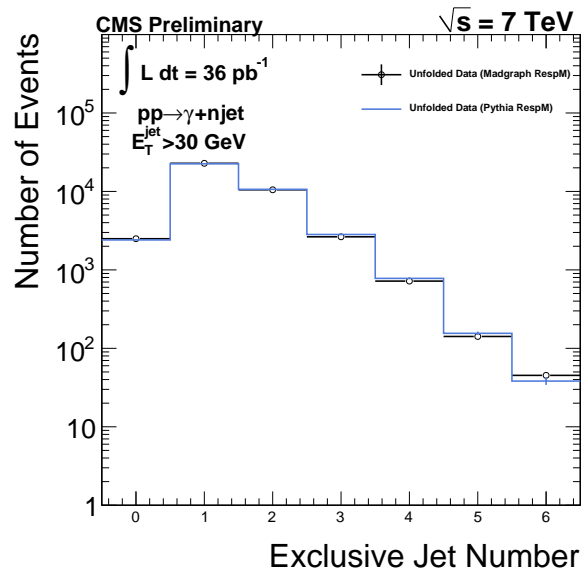
Figure 7.7: Jet multiplicity distribution after unfolding using response matrices from PYTHIA compared to MADGRAPH. The difference is largest at the highest jet multiplicities at about 10%.

Chapter 8

Results

The measurement made is ratios of cross sections of events with a photon and n or more jets. The exclusive cross section is defined as:

$$\sigma(\gamma + n\,\text{jets}) = \frac{N_s(n) \cdot U(n)}{\epsilon(n) \cdot \int \mathcal{L}} \tag{8.1}$$

where N_s is the number signal events, U is an unfolding factor, ϵ is the efficiency, and $\int \mathcal{L}$ is the integrated luminosity. The factors N_s, ϵ were found for each number of jets, n. The integrated luminosity is a constant, and is just the total amount of data used (in our case it is $36\,\text{pb}^{-1}$). The unfolding factors are determined by the unfolding process described in 7.5.

Once the exclusive cross sections are determined, the inclusive cross sections are found by summing the cross sections for n or more jets:

$$\sigma(\gamma + \geq n\,\text{jets}) = \sum_{i=n}^{\infty} \sigma(\gamma + i\,\text{jets}) \tag{8.2}$$

The integrated luminosity, $\int \mathcal{L}$, is the same in all the terms, and so cancels when the ratios are taken. The has the advantage of removing any uncertainty associated with the amount of data taken.

Table 8.1: $\sigma(\gamma + \geq n \text{ jets})/\sigma(\gamma + \geq 1 \text{ jets})$, the jet multiplicities normalized to the "1 or more jets" inclusive cross section.

Jet multiplicity	σ ratio	uncertainty	
		statistical	systematic
≥ 2 / ≥ 1 jets	0.36	± 0.003	$+0.05/-0.02$
≥ 3 / ≥ 1 jets	0.09	± 0.001	$+0.02/-0.005$
≥ 4 / ≥ 1 jets	0.02	± 0.0005	$+0.003/-0.004$
≥ 5 / ≥ 1 jets	0.005	± 0.0003	$+0.001/-0.001$

Table 8.2: $\sigma(\gamma + \geq n \text{ jets})/\sigma(\gamma + \geq (n-1) \text{ jets})$, the ratio of jet multiplicities.

Jet multiplicity	σ ratio	uncertainty	
		statistical	systematic
≥ 2 / ≥ 1 jets	0.36	± 0.003	$+0.05/-0.02$
≥ 3 / ≥ 2 jets	0.25	± 0.004	$+0.03/-0.003$
≥ 4 / ≥ 3 jets	0.26	± 0.007	$+0.02/-0.07$
≥ 5 / ≥ 4 jets	0.23	± 0.01	$+0.03/-0.05$

After background subtraction, efficiency correction, and unfolding the final distributions binned by exclusive jet multiplicity are translated to binned by *inclusive* jet multiplicities. The results for $\sigma(\gamma + \geq n\text{jets})/\sigma(\gamma + \geq 1\text{jet})$ are shown in figure 8.1. The results for $\sigma(\gamma + \geq n\text{jets})/\sigma(\gamma + \geq (n-1)\text{jets})$ are shown in figure 8.2.

The MADGRAPH monte carlo simulation agrees well with data, while the PYTHIA simulations fall well below data. This is expected, because the PYTHIA simulations only contain matrix elements for hard scatterings with 1 photon and 1 jet in the final state, any more jets found in the event are simply the result of fragmentation from the original jet. The MADGRAPH simulations contained matrix elements for final states with higher jet multiplicity.

Figure 8.1: Ratio of cross sections $\sigma(\gamma+ \geq n\text{jets})/\sigma(\gamma+ \geq 1\text{jet})$ for data, MAD-GRAPH, and PYTHIA MC. MADGRAPH samples used simulated up to $\gamma + 3$ jets, while PYTHIA simulated up to $\gamma + 1$ jet. Asymmetry in the systematics is likely due to the unfolding process.
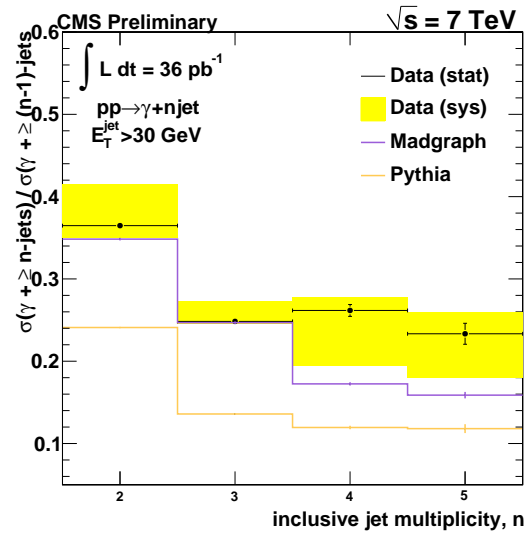

Figure 8.2: Ratio of cross sections $\sigma(\gamma+ \geq n\text{jets})/\sigma(\gamma+ \geq (n-1)\text{jets})$ for data, MADGRAPH, and PYTHIA MC. MADGRAPH samples used simulated up to $\gamma+3$ jets, while PYTHIA simulated up to $\gamma + 1$ jet. Asymmetry in the systematics is likely due to the unfolding process.
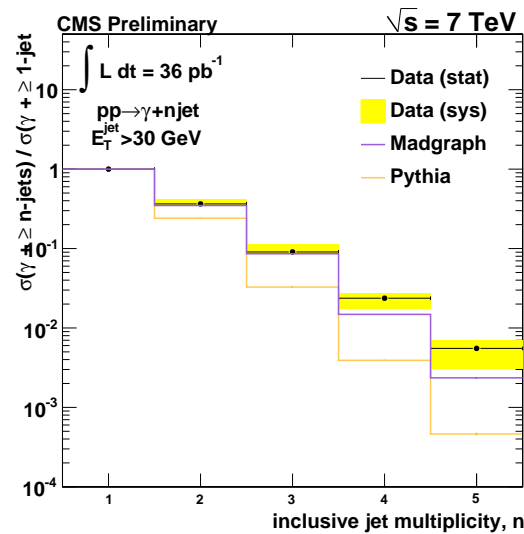
8.1 Systematic Uncertainty

The largest source of systematic uncertainty for the γ+jets measurements is the uncertainty in jet energy. Over or underestimation of jet energy affects the counting of the number of jets in events. Uncertainty in jet energy comes from several sources [80], including:

- Uncertainty in flavor composition between jets used to extract energy corrections, and jets in γ+jets events. This will add roughly 2% uncertainty on jet energy.

- Subtraction of energy due to pileup corrections removes roughly 500 MeV of energy from jets in events without pile-up present.

- Uncertainty in the jet energy corrections that are applied as a function of jet η and p_T.

The uncertainty in jet energy was calculated for each jet in every event using the sources listed above. Due to the jet threshold $p_\mathrm{T} > 30\,\mathrm{GeV}$, the sensitivity to the underlying event is small. The 1σ uncertainty (on order of a few GeV) on energy was added (subtracted) to each jet's energy to have a collection of jets with over(under)-estimated jet energies. Then, the number of jets passing the p_T threshold were then counted from these collections. Overestimation of jet energy decreases the number of events with 0 or 1 jet, and increases the number of events with higher jet multiplicity. Underestimation of jet energy has the opposite effect. As the number of jets present in the event increases, the greater the effect on total uncertainty from jet energy uncertainties. The effect of uncertainty in jet energy on the final results are listed in tables 8.4 and 8.3.

Another contribution to the total systematic uncertainty comes from the knowledge of the shape of the background component. The sideband selection has more activity in the isolation cone than is present in the true background, leading to an overestimation of the signal yield. In addition, the sideband selection contains some signal, leading to an underestimation of the signal yield. The bounds on the sideband selection used on the data to obtain the background template distribution was varied by $\pm 1\,\mathrm{GeV}$ and the effect on the final results are listed in tables 8.4 and 8.3.

Uncertainty arising from the photon isolation efficiency corrections are also considered. The largest difference between data and monte carlo for photon efficiency is roughly 4% in one of the jet multiplicity bins. This value of 4% is used as the systematic uncertainty from efficiency in the final result.

Uncertainty arising from the unfolding process are estimated by changing the MC generator used to make the response matrix. Figure 7.7 shows the effect of using different Monte Carlo signal samples to derive the response matrix. A relatively small dependence on the origin of the response matrix is visible.

The effect of all these systematics on the final results are shown in tables 8.1 and 8.2, and the details of each source of uncertainty are shown in tables 8.3 and 8.4. The measurement of the cross section ratios is limited by systematic uncertainty, and is dominated by uncertainty on the jet energy scale.

Table 8.3: Relative systematic uncertainties on the final jet multiplicity ratios of $(\geq n \text{ jets})/(\geq 1 \text{ jet})$ in $\gamma + n$ jets events.

	$\geq 2/ \geq 1$ jets	$\geq 3/ \geq 1$ jets	$\geq 4/ \geq 1$ jets	$\geq 5/ \geq 1$ jets
JES $(+/- 1\sigma)$	+13/-4 %	+23/-5 %	+15/-18 %	+27/-36 %
Signal Fraction	+4 %	+8 %	-8 %	-10 %
Efficiency	± 4 %	± 4 %	± 4 %	± 4 %
Unfolding	-1 %	+7 %	+8 %	+9 %
Total	+14/-6 %	+26/-6 %	+17/-20 %	+29/-38 %

Table 8.4: Relative systematic uncertainties on the final jet multiplicity ratios of $(\geq n$ jets$)/(\geq (n-1)$ jets$)$ in $\gamma + n$ jets events.

	$\geq 2/ \geq 1$ jets	$\geq 3/ \geq 2$ jets	$\geq 4/ \geq 3$ jets	$\geq 5/ \geq 4$ jets
JES $(+/- 1\sigma)$	+13/-4 %	+9/-1 %	+7/-13 %	+11/-22 %
Signal Fraction	+4 %	+4 %	-15 %	-3 %
Efficiency	±4 %	±4 %	±4 %	±4 %
Unfolding	+2 %	+4 %	+1 %	-4 %
Total	+14/-6 %	+11/-1 %	+8/-26 %	+12/-23 %

Chapter 9

Conclusion

The production of jets in association with high-pt photons has been studied with the CMS detector at the Large Hadron Collider with an integrated luminosity of $36 \, \text{pb}^{-1}$. Ratios of the production of $pp \rightarrow \gamma + \text{jets}$ were measured. Photons were required to have $E_\text{T} > 75 \, \text{GeV}$ and be within $|\eta| < 2.5$, and pass loose isolation requirements. Jets were counted if they had $p_\text{T} > 30 \, \text{GeV}$, and were within $|\eta| < 2.4$. Jets were reconstructed with the anti-k_t algorithm with a cone size of $\Delta R = 0.5$, and were required not to overlap with the photon.

The cross section ratios measured were $\sigma(\gamma+ \geq n\text{jets})/\sigma(\gamma+ \geq (n-1)\text{jets})$ and $\sigma(\gamma+ \geq n\text{jets})/\sigma(\gamma+ \geq 1\text{jet})$. The ratios were compared to leading-order predictions from PYTHIA and MADGRAPH monte carlos. The PYTHIA samples simulate only one photon and one jet in the final state, and rely on fragmentation of the jet to create more jets. The MADGRAPH samples used contain leading-order matrix elements for up to $\gamma + 3 \, \text{jets}$. The data and MADGRAPH monte carlo prediction agreed within errors for ratios when 2-jet or 3-jet events were in the numerator. The data deviated from MADGRAPH predictions for higher jet multiplicities, as expected. Possibilities for future analysis include comparing data to MADGRAPH simulations with higher

jet multiplicities, or comparing to NLO predictions if they become available.

The reasonable agreement of MADGRAPH predictions to these measurements provides confidence in using MADGRAPH for new physics searches in which photons and jets are in the final state. It also provides confidence in our understanding of the detector and reconstruction of photons and jets.

9.1 Outlook

This thesis used all the high-quality data recorded in year 2010 with the CMS detector. As of this writing, in the summer of 2011, the LHC accelerator is delivering in a single 12-hour fill the same amount of data that was delivered in the whole 2010 run. The CMS detector has now recorded a total of over $1\ \mathrm{fb}^{-1}$, and will continue to record proton collision events until the end of 2012. At that point the LHC will begin a 2-year shutdown for upgrades. In that amount of data ($\sim 20\,\mathrm{fb}^{-1}$), we expect a few hundred million events of γ+jets (with $p_\mathrm{T}^\gamma > 15\,\mathrm{GeV}$). Certainly, the amount recorded may be be significantly less than that due to the increasingly high thresholds on p_T of the single photon triggers used to select these events. Nevertheless, we expect to measure jet multiplicity distributions to higher values. Additional measurements characterizing the jet kinematics can also be made. such measurements will strengthen our confidence in monte carlo programs enabling a thorough exploration of new physics using the increasing amount of data from the LHC.

With the increasing statistics, it will be possible to improve uncertainties such as those due to photon selection efficiencies with data and will remove the need to use simulations to find such efficiencies.

Bibliography

[1] Precision Electroweak Measurements and Constraints on the Standard Model (2008). `0811.4682`.

[2] Hagiwara, K. *et al.* Review of particle properties. *Phys. Rev. D* **66**, 010001 (2002).

[3] Englert, F. & Brout, R. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.* **13**, 321–323 (1964).

[4] Higgs, P. W. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.* **13**, 508–509 (1964).

[5] Higgs, P. W. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.* **145**, 1156–1163 (1966).

[6] Barger, V. & Phillips, R. *Collider Physics* (Westview Press, Boulder, CO, 1991).

[7] Hanneke, D., Fogwell, S. & Gabrielse, G. New Measurement of the Electron Magnetic Moment and the Fine Structure Constant. *Physical Review Letters* **100**, 120801–+ (2008). `0801.1134`.

[8] Gross, D. J. & Wilczek, F. Asymptotically free gauge theories. i. *Phys. Rev. D* **8**, 3633–3652 (1973).

[9] Gross, D. J. & Wilczek, F. Asymptotically free gauge theories. ii. *Phys. Rev. D* **9**, 980–993 (1974).

[10] Perkins, D. H. *Introduction to High Energy Physics* (Cambridge University Press, 2000).

104

[11] Lai, H.-L. *et al.* New parton distributions for collider physics. *Phys. Rev.* **D82**, 074024 (2010). `1007.2241`.

[12] CMS Collaboration. *CMS TriDAS project: Technical Design Report; 1, the trigger systems.* Technical Design Report CMS (CERN, 2000).

[13] Mandelstam, S. Determination of the pion-nucleon scattering amplitude from dispersion relations and unitarity. general theory. *Phys. Rev.* **112**, 1344–1360 (1958).

[14] Catani, S., Fontannaz, M., Guillet, J. P. & Pilon, E. Cross-section of isolated prompt photons in hadron hadron collisions. *JHEP* **05**, 028 (2002). `hep-ph/0204023`.

[15] Aurenche, P., Baier, R., Fontannaz, M. & Schiff, D. Prompt photon production at large pt scheme invariant qcd predictions and comparison with experiment. *Nuclear Physics B* **297**, 661 – 696 (1988). URL `http://www.sciencedirect.com/science/article/pii/0550321388905536`.

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

[17] Aversa, F., Chiappetta, P., Greco, M. & Guillet, J. P. Higher-order corrections to qcd jets. *Physics Letters B* **210**, 225 – 232 (1988). URL `http://www.sciencedirect.com/science/article/pii/0370269388903772`.

[18] Pumplin, J. Quark-gluon jet differences in z decay. *Phys. Rev. D* **48**, 1112–1116 (1993).

[19] Berger, C. F. *et al.* Precise Predictions for $W + 4$-Jet Production at the Large Hadron Collider. *Phys. Rev. Lett.* **106**, 092001 (2011).

[20] Berger, C. F. *et al.* Next-to-leading order QCD predictions for $Z, \gamma^* + 3$-jet distributions at the Tevatron. *Phys. Rev. D* **82**, 074002 (2010).

[21] Aad, G. *et al.* Measurement of the production cross section for W-bosons in association with jets in pp collisions at $\sqrt{s} = 7 \, \text{TeV}$ with the ATLAS detector (2010). `1012.5382`.

[22] ATLAS Collaboration. Measurement of the production cross section for Z/γ in association with jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS Detector. Tech. Rep. ATLAS-CONF-2011-001, CERN, Geneva (2011).

[23] CMS Collaboration. Rates of Jets Produced in Association with W and Z Bosons in pp Collisions at $\sqrt{7}$ TeV. *CMS PAS EWK* **10-012** (2011).

[24] Gluck, M., Gordon, L. E., Reya, E. & Vogelsang, W. High-p_{T} photon production at $p\bar{p}$ colliders. *Phys. Rev. Lett.* **73**, 388–391 (1994).

[25] Aurenche, P. *et al.* A critical phenomenological study of inclusive photon production in hadronic collisions. *Eur. Phys. J.* **C9**, 107–119 (1999). `hep-ph/9811382`.

[26] CDF Collaboration. 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.D* **65** (2002). URL `http://arxiv.org/abs/hep-ex/0201004`.

[27] Lai, H. *et al.* Global qcd analysis of parton structure of the nucleon: Cteq5 parton distributions. *The European Physical Journal C - Particles and Fields* **12**, 375–392 (2000). URL `http://dx.doi.org/10.1007/s100529900196`. 10.1007/s100529900196.

[28] Abazov, V. M. *et al.* Measurement of the differential cross-section for the production of an isolated photon with associated jet in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$-TeV. *Phys. Lett.* **B666**, 435–445 (2008). `0804.1107`.

[29] Martin, A., Roberts, R., Stirling, W. & Thorne, R. Physical gluons and high-et jets. *Physics Letters B* **604**, 61 – 68 (2004). URL `http://www.sciencedirect.com/science/article/pii/S0370269304015059`.

[30] Alekhin, S. I. Parton distributions from deep-inelastic-scattering data. *Phys. Rev. D* **68**, 014002 (2003).

[31] ZEUS Collaboration. An NLO QCD analysis of inclusive cross-section and jet-production data from the ZEUS experiment. *The European Physical Journal C - Particles and Fields* **42**, 1–16 (2005). URL `http://dx.doi.org/10.1140/epjc/s2005-02293-x`. 10.1140/epjc/s2005-02293-x.

[32] Adler, S. S. *et al.* Mid-rapidity direct-photon production in p+p collisions at sqrt(s) = 200-GeV. *Phys. Rev.* **D71**, 071102 (2005). `hep-ex/0502006`.

[33] Okada, K. Measurement of prompt photons in $\sqrt{s} = 200\,\text{GeV}$ pp collisions (2005). `hep-ex/0501066`.

[34] Glück, M., Reya, E. & Vogt, A. Erratum: Parton fragmentation into photons beyond the leading order. *Phys. Rev. D* **51**, 1427 (1995).

[35] Aaron, F. D. *et al.* Prompt Photons in Photoproduction at HERA. *Eur. Phys. J.* **C66**, 17–33 (2010). `0910.5631`.

[36] Fontannaz, M., Guillet, J. P. & Heinrich, G. Isolated prompt photon photoproduction at NLO. *Eur. Phys. J.* **C21**, 303–312 (2001). `hep-ph/0105121`.

[37] Lipatov, A. V. & Zotov, N. P. Prompt photon photoproduction at HERA in the k(T)- factorization approach. *Phys. Rev.* **D72**, 054002 (2005). `hep-ph/0506044`.

[38] Chekanov, S. *et al.* Measurement of prompt photons with associated jets in photoproduction at HERA. *Eur. Phys. J.* **C49**, 511–522 (2007). `hep-ex/0608028`.

[39] Zembrzuski, A. & Krawczyk, M. Photoproduction of isolated photon and jet at the DESY HERA (2003). `hep-ph/0309308`.

[40] Chekanov, S., Derrick, M., Krakauer, D. & Magill, S. Study of the effective transverse momentum of partons in the proton using prompt photons in photoproduction at hera. *Physics Letters B* **511**, 19 – 32 (2001). URL `http://www.sciencedirect.com/science/article/pii/S0370269301006153`.

[41] ATLAS Collaboration. Measurement of the inclusive isolated prompt photon cross section in pp collisions at sqrt(s) = 7 TeV with the ATLAS detector (2010). `1012.4389`.

[42] CMS Collaboration. Measurement of the isolated prompt photon production cross section in *pp* collisions at $\sqrt{s} = 7\,\text{TeV}$. *Phys. Rev. Lett.* **106**, 082001 (2011).

[43] Lai, H.-L. *et al.* New parton distributions for collider physics. *Phys. Rev. D* **82**, 074024 (2010). `1007.2241`.

[44] Bourhis, L., Fontannaz, M. & Guillet, J. Quark and gluon fragmentation functions into photons. *The European Physical Journal C* **2**, 529–537 (1998). URL `http://www.springerlink.com/content/hq3kqjwng4ebku0d/`.

[45] Evans, L. & Bryant, P. LHC Machine. *Journal of Instrumentation* **3**, S08001 (2008). URL `http://stacks.iop.org/1748-0221/3/i=08/a=S08001`.

[46] CMS Collaboration. Measurement of CMS Luminosity. *PAS-EWK-10-004* (2010). URL `http://cdsweb.cern.ch/record/1279145`.

[47] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults2010

[48] CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation* **3**, S08004 (2008).

[49] CMS Collaboration. Jet Performance in pp Collisions at 7 TeV. *CMS-PAS-JME-10-003* (2010).

[50] Adam, W. *et al.* The CMS high level trigger. *Eur. Phys. J.* **C46**, 605–667 (2006). `hep-ex/0512077`.

[51] Nadolsky, P. M. *et al.* Implications of CTEQ global analysis for collider observables. *Phys. Rev.* **D78**, 013004 (2008). `0802.0007`.

[52] Stump, D. *et al.* Inclusive jet production, parton distributions, and the search for new physics. *JHEP* **10**, 046 (2003). `hep-ph/0303013`.

[53] Altarelli, G. & Parisi, G. Asymptotic freedom in parton language. *Nuclear Physics B* **126**, 298 – 318 (1977). URL `http://www.sciencedirect.com/science/article/pii/0550321377903844`.

[54] Dobbs, M. & Hansen, J. B. The HepMC C++ Monte Carlo event record for High Energy Physics. *Computer Physics Communications* **134**, 41 – 46 (2001). URL `http://www.sciencedirect.com/science/article/pii/S0010465500001892`.

[55] Andersson, B., Gustafson, G., Ingelman, G. & Sjstrand, T. Parton fragmentation and string dynamics. *Physics Reports* **97**, 31 – 145 (1983).

[56] Agostinelli, S. *et al.* Geant4–a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506**, 250 – 303 (2003).

[57] Sjöstrand, T., Mrenna, S. & Skands, P. Z. PYTHIA 6.4 Physics and Manual. *JHEP* **05**, 026 (2006). `hep-ph/0603175`.

[58] Halzen, F. & Scott, D. M. Hadroproduction of photons and leptons. *Phys. Rev. D* **18**, 3378–3389 (1978).

[59] Bengtsson, H. U. The lund monte carlo for high-pt physics. *Computer Physics Communications* **31**, 323 – 355 (1984). URL `http://www.sciencedirect.com/science/article/pii/0010465584900183`.

[60] Alwall, J. *et al.* MadGraph/MadEvent v4: The New Web Generation. *JHEP* **09**, 028 (2007). `0706.2334`.

[61] Catani, S., Krauss, F., Kuhn, R. & Webber, B. R. QCD Matrix Elements + Parton Showers. *JHEP* **11**, 063 (2001). `hep-ph/0109231`.

[62] Krauss, F. Matrix elements and parton showers in hadronic interactions. *JHEP* **08**, 015 (2002). `hep-ph/0205283`.

[63] Mangano, M. L., Moretti, M. & Pittau, R. Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b} + n$-jets as a case study. *Nuclear Physics B* **632**, 343 – 362 (2002). URL `http://www.sciencedirect.com/science/article/pii/S0550321302002493`.

[64] CMS Collaboration. Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 7\,\text{TeV}$ and Comparison with $\sqrt{s} = 0.9\,\text{TeV}$. *PAS-QCD-10-010* (2010).

[65] CMS Collaboration. Tracking and Vertexing Results from First Collisions. *CMS Note* (2010).

[66] Schmidt, A. Performance of Track and Vertex Reconstruction and B-Tagging Studies with CMS in pp Collisions at sqrt(s)=7 TeV. *CMS CR-2011-047* (2011).

[67] Fruehwirth, R., Waltenberger, W. & Vanlaer, P. Adaptive Vertex Fitting. *CMS IN-2007/008* (2007).

[68] CMS Collaboration. ParticleFlow Event Reconstruction in CMS and Performance for Jets, Taus, and E_T^{miss}. *CMS PFT-2009/001* (2009).

[69] Cacciari, M., Salam, G. P. & Soyez, G. The anti-k_t jet clustering algorithm. *Journal of High Energy Physics* **2008**, 063 (2008). URL `http://stacks.iop.org/1126-6708/2008/i=04/a=063`.

[70] Catani, S., Dokshitzer, Y. L., Seymour, M. H. & Webber, B. R. Longitudinally invariant K(t) clustering algorithms for hadron hadron collisions. *Nucl. Phys.* **B406**, 187–224 (1993).

[71] Dokshitzer, Y. L., Leder, G. D., Moretti, S. & Webber, B. R. Better Jet Clustering Algorithms. *JHEP* **08**, 001 (1997). `hep-ph/9707323`.

[72] Wobisch, M. & Wengler, T. Hadronization corrections to jet cross sections in deep-inelastic scattering (1998). `hep-ph/9907280`.

[73] Anderson, M. *et al.* Review of clustering algorithms and energy corrections in ECAL. *CMS IN-2010/008* (2008).

[74] CMS Collaboration. Electromagnetic calorimeter commissioning and performance with 7 TeV data. *CMS PAS EGM-10-002* (2010).

[75] CMS Collaboration. Transverse-Momentum and Pseudorapidity Distributions of Charged Hadrons in *pp* Collisions at $\sqrt{s} = 7\,\mathrm{TeV}$. *Phys. Rev. Lett.* **105**, 022002 (2010).

[76] Miceli, T., Tripathi, M., Askew, A., Gershtein, Y. & Gaultney, V. Data-Driven Determination of Beam Halo Contamination in Photon Candidates. *CMS AN-2010/115* (2010).

[77] Chumney, P. *et al.* Level-1 regional calorimeter trigger system for cms. *Proceedings of Computing in High Energy Physics and Nuclear Physics* (2003). URL `http://arxiv.org/abs/hep-ex/0305047`.

[78] CMS Collaboration. Electron reconstruction and identification at $\sqrt{s} = 7\,\mathrm{TeV}$. *CMS PAS EGM-10-004* (2010).

[79] CMS Collaboration. Commissioning of the Particle-Flow Event Reconstruction using high-energy LHC Collisions in the CMS Detector. *CMS PAS PFT* **1-002** (2010).

[80] CMS Collaboration. Jet Energy Corrections determination at $\sqrt{s} = 7$ TeV. *CMS PAS JME* **10-010** (2010).

[81] Saoulidou, N. Particle flow jet identification criteria. *CMS Note* **CMS AN AN-10-003** (2010).

[82] CMS Collaboration. Isolated Photon Reconstruction and Identification at $\sqrt{s} = 7$ TeV. *CMS PAS EGM* **10-006** (2010).

[83] Measurement of the isolated prompt photon production cross section in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev. Lett.* **106** (2011). `1012.0799`.

[84] CMS Collaboration. Measurement of the Isolated Prompt Photon Production Cross Section in pp Collisions at sqrt(s) = 7 TeV (2010). `1012.0799`.

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

[86] Adye, T. et. al. RooUnfold: ROOT unfolding framework (2011).