MEASUREMENT OF THE DIFFERENTIAL CROSS SECTION OF PHOTON AND JET EVENTS IN pp COLLISIONS AT $\sqrt{s} = 7$ TeV

Chuanzhe Lin
Zhouning, Fujian, P.R. China

B.S., University of Science and Technology of China, 2007

A Thesis presented to the Graduate Faculty of the University of Virginia in Candidacy for the Degree of Doctor of Philosophy

Department of Physics

University of Virginia
Jan, 2014

1. Reviewer:

2. Reviewer:

3. Reviewer:

4. Reviewer:

Day of the defense:

MEASUREMENT OF THE DIFFERENTIAL CROSS SECTION OF PHOTON AND JET EVENTS IN pp COLLISIONS AT $\sqrt{s} = 7$ TeV

Chuanzhe Lin

Department of Physics
University of Virginia

A thesis submitted for the degree of
Doctor of Philosophy

Jan, 2014

We present a measurement of the triple-differential cross section, $d^3\sigma/(dp_T^\gamma d\eta^\gamma d\eta^{jet})$, in photon+jets final states using a data sample from proton-proton collisions at $\sqrt{s} =$ 7 TeV. This sample corresponds to an integrated luminosity of 2.14 fb^{-1} collected by the Compact Muon Solenoid (CMS) detector at the Large Hadron Collider (LHC). Photons and jets are reconstructed within a pseudorapidity range of $|\eta| < 2.5$, and are required to have transverse momentum in the range $40 < p_T^\gamma < 300$ GeV and $p_T^{jet} > 30$ GeV, respectively. The measurements are compared to theoretical predictions from the SHERPA leading-order Monte Carlo event generator and the next-to-leading-order perturbative QCD calculation from JETPHOX. The predictions are found to be consistent with the data over much of the examined kinematic region.

To my father Wenjie for making me who I am.

To my wife Xiaomei for being the love of my life, my strength and support.

To my son Aaron for bringing the most joy to my life.

Acknowledgements

My first and sincere appreciation goes to my Ph.D. advisor, Prof. Bob Hirosky, for all I have learned from him and for his continuous help, organizational support, and professional advice throughout my graduate study. I would like to thank Vasundhara Chetluru for her patient guidance in the QCD photon analysis and for always finding time to answer my questions. I have many thanks to my colleagues who also engaged in photon+jets analysis for their collaboration and effective teamwork: Darko Mekterovic, Ocalan Kadir, and Sudha Ahuja.

The University of Virginia high energy physics group provides me with a great platform to be engaged in a range of interesting and useful projects. I would like to thank Prof. Chris Neu and Prof. Brad Cox for their mentorship. I owe my thanks to Sasha Ledovskoy for the stimulating discussion in LED and endcap projects. Many thanks also go to all other members in UVA CMS group. I would like to give a special thanks to my officemate, Huong Nguyen, for her help and encouragement.

I am also grateful to other members of my dissertation committee, Pro. Hank Thacker and Pro. Jeff Holt for their precious time.

I owe my loving thanks to my wife Xiaomei Mu. She is the one who always supports, understands, and motivates me. She took good care of our lives when I was fully occupied by my research. She gave me the best gift in my life, our precious baby Aaron Lin. Moreover, I would like to express my deep gratitude to my father Wenjie Lin and my grandparents Xiurong Zhang and Ruoyuan Lin for making me who I am. And I am grateful to my parents-in-law Qijuan Hui and Hong Mu for their unconditional support and beliefs in me. I also want to thank Qun Zhang and Yongzhe Lin for their moral support.

Contents

List of Figures

LIST OF FIGURES

List of Tables

Chapter 1

INTRODUCTION

The field of Elementary Particle Physics focuses on the study of the fundamental objects of the universe and their corresponding interactions. From the development of quantum mechanics through the experimental discoveries of a "zoo" of new particles since the 1930s, the parallel development of theories and discoveries by experiment have resulted in a remarkable insight into the fundamental structure of the universe. The Standard Model of particle physics summarizes our best understanding of the properties of the fundamental particles and the interactions between them via three types of forces. This theory matured in the early 1970s as a collaborative effort of scientists around the world and has become a well-tested physics theory, explaining most of our experimental results and accurately predicting many phenomena.

However, the Standard Model still does not explain the complete picture. Gravity is not incorporated in this theory. It does not describe dark matter, dark energy or the mass hierarchy of the elementary particles. How much do we know about dark matter? Do the quarks themselves have a deeper composite structure? How have we come to exist in a matter-dominated universe? Are there undiscovered fundamental forces or symmetries in nature? Do we live in a fine-tuned or more natural universe? Many high energy physicists are pursuing goals to complete the Standard Model and to explore physics beyond the Standard Model. The Large Hadron Collider (LHC), which has surpassed any predecessor, becoming the largest and highest-energy particle accelerator ever constructed, is a primary tool for finding more of these missing pieces. The Compact Muon Solenoid (CMS) is constructed as a general-purpose detector to capture signals from particles produced by collisions in the LHC. It enables us to

perform comprehensive tests of the Standard Model and to search for new physics at the TeV energy scale.

This thesis presents the first measurement of the triple-differential cross section for photon+jets production using data collected within the CMS detector. The data corresponds to an integrated luminosity of 2.14 fb^{-1} from proton-proton collision at $\sqrt{s} = 7$ TeV. Studies of events with a photon and one or more jets in the final state provide a direct probe of hard quantum chromodynamics (QCD) interactions (7, 8, 9, 10, 11). Because the production of photon+jets events is dominated by the Compton-like process ($qg \to q\gamma$) at LHC energies, the differential cross sections of various angular configurations are sensitive to the gluon parton distribution functions (PDFs) of the proton over large ranges of parton momentum fraction x and hard interaction scales Q^2 (12, 13). Therefore, measurements of the cross sections can help to constrain future PDF models and provide information for improving phenomenological Monte Carlo models, as well as testing the applicability of fixed-order perturbative calculations over a wide range of kinematic regions. Photon+jets events are a major source of background to standard model measurements, most notably that for a light neutral Higgs boson decaying via $\text{H} \to \gamma\gamma$ (14), as well as beyond-the-standard-model searches for signatures of extra dimensions (15) and excited quarks (16), among others. Photon+jets events can also be used to calibrate jet energies (17), and for the modeling of missing transverse energy distributions attributed to the presence of non-interacting particles (18).

This thesis is organized as follows. Chapter 2 presents a brief review of the Standard Model of particle physics (especially Quantum Chromodynamics) and Monte Carlo methods used for theoretical predictions. Chapter 3 describes the experimental apparatus of the accelerator Large Hadron Collider (LHC) and Compact Muon Solenoid (CMS), which provided the data that was used to perform this measurement. Chapters 4 and 5 describe the event reconstruction and selection for the measurement of the triple-differential cross section for the photon+jets events. Chapter 6 describes the details of the cross section measurement, including the photon efficiency calculation, photon purity evaluation and unfolding of the detector effects. Chapter 7 summarizes the result of the cross section measurement and gives the conclusion.

Chapter 2

THEORY

2.1 Standard Model

The Standard Model (SM) of particle physics is the most successful theory describing the fundamental particles and their interactions. In the SM, the fundamental particles are grouped into categories of fermions and bosons. Fermions are the building blocks of matter, while bosons are often carriers of forces. The SM describes three types of interactions including electromagnetic, weak, and strong forces. Technically, the SM is expressed in the mathematical framework of quantum field theory, where the Lagrangian summarizes the dynamics of the theory. Each particle is described in the form of a field. The local $U(1)_Y \times SU(2) \times SU(3)$ gauge symmetry gives rise to the fundamental interactions of electromagnetic, weak, and strong forces. From the middle of 20th century, when Sheldon Glashow proposed the unification of electromagnetic and weak interactions (19), through the recent detection of the Higgs boson at the LHC (20), the development of the SM has been constantly driven by theoretical advances and experimental discoveries.

There are twelve fermions in the SM, each having half integer spin and obeying the Pauli Exclusion Principle according to the spin-statistics theorem. They can be further divided according to different criteria. Based on if they participate in the strong interaction, fermions can be categorized into two subgroups as quarks and leptons, which are shown in the purple and green sections in Fig. 2.1. Quarks participate in strong interactions and are subject to the phenomenon called color confinement. Unlike leptons, they cannot exist independently. Instead, they are bound into triplets or

doublets, where triplets are called baryons and doublets are called mesons. Additionally, fermions can be divided into three generations distinguished by flavor and mass, as shown in the first three columns in Fig. 2.1. Between generations, fermions of the same charge share identical interactions. The higher the generation, the greater the mass of the corresponding particles (neutrinos might be an exception, but this question remains unanswered). This mass hierarchy allows particles of higher generations to decay to lower generations.

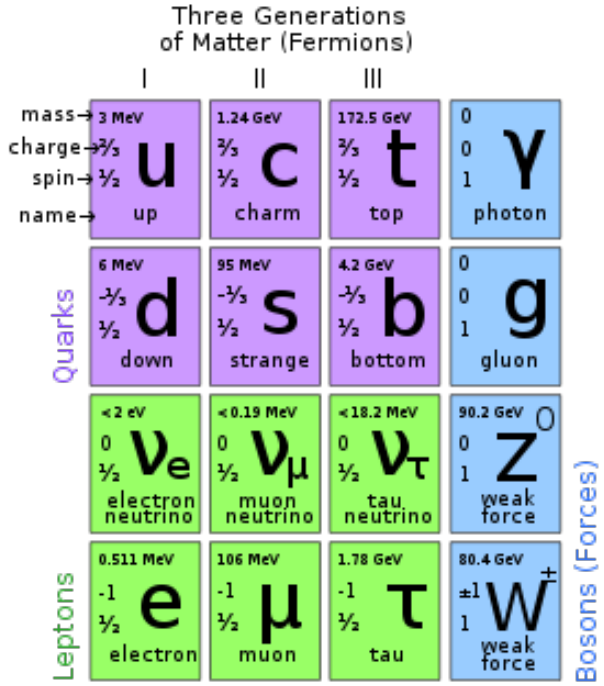


Figure 2.1: Fermions and gauge bosons in the Standard Model

Another category of fundamental particles in the SM is bosons. In contrast to fermions, bosons have integer spin and obey Bose-Einstein statistics, allowing multiple particles to exist in the same state. In the SM, bosons include the Higgs boson and four force-carrying gauge bosons. The Higgs boson is the consequence of the mechanism that causes some fundamental particles to have mass, which are otherwise required to be massless due to symmetry constrains. In the SM, the Higgs boson carries zero spin, electric charge, or color charge. It is very unstable and decays into other particles with lower mass. The four force-carrying gauge bosons are listed in the last column in

Fig. 2.1. They realize the three basic interactions of the fundamental particles. Interactions between fermions happen through the exchange of force-carriers. The photon is the force-carrier mediating the electromagnetic interaction. Quantum electrodynamics (QED) is a quantum field theory of the electromagnetic force. It describes all electromagnetic phenomena associated with charged fundamental particles, like the Compton scattering process shown in Fig. 2.2. Three massive gauge bosons, $W^{\pm}$ and neutral Z^0, mediate the weak interaction. Their large masses account for the short range of weak interaction. Fermions interact weakly through the exchange of $W^{\pm}$ or Z^0 bosons. The weak interaction mediated by the charged W boson is called the charged current interaction and is illustrated in the beta decay process shown in Fig. 2.3. The weak interaction involving the Z boson, neutral current interaction, is illustrated by the quark-antiquark annihilation process shown in Fig. 2.4. Unification of electromagnetic and weak forces is accomplished under the $SU(2)\times U(1)_Y$ gauge group. The corresponding gauge boson of $U(1)_Y$ is the B^0 boson of weak hypercharge and those of SU(2) are W bosons (W^+, W^0, and W^-) of weak isospin. From the Higgs mechanism in the SM, the photon, and the physical $W^{\pm}$ and Z^0 bosons are produced by spontaneous symmetry breaking of $SU(2) \times U(1)_Y$. The spontaneous symmetry mixes the B^0 and W^0 resulting in the photon and the Z^0 boson as shown in the following equation:

$$\begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B^0 \\ W^0 \end{pmatrix}; \tag{2.1}$$

where θ_W is the weak mixing angle. The Higgs Mechanism also describes the generation of masses for the $W^{\pm}$ and Z^0 bosons. The remaining interaction in the SM is the strong interaction, mediated by the gluon bosons. Quantum chromodynamics (QCD) is a non-abelian gauge theory with the group SU(3). Quarks interact with each other via the exchange of gluons. The confinement property of QCD prevents the existence of free quarks. On the other hand, the interaction between quarks and gluons becomes very weak in high-energy interactions, an effect known as asymptotic freedom. Perturbative QCD (pQCD) is based on the effect of asymptotic freedom. In this thesis, the analysis of the cross section for producing photon+jets events in proton-proton collisions at center-of-mass energy $\sqrt(s) = 7$ TeV provides a testing ground for pQCD. More discussion of QCD is provided in the next section.

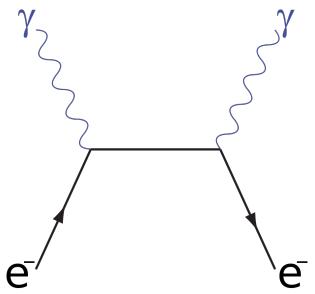


Figure 2.2: The Feynman diagram of the Compton scattering process.

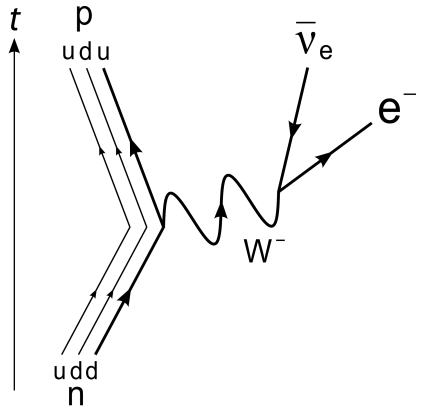


Figure 2.3: The Feynman diagram of beta decay through the weak interaction.

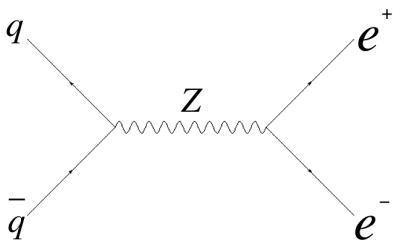


Figure 2.4: The Feynman diagram of quark-antiquark annihilation with an intermediate Z boson.

In summary, the SM of particle physics provides a framework that demonstrates our best understanding of how the fundamental particles interact with each other. Theoretically, the SM is expressed using gauge fields based on the symmetry group $U(1)_Y \times SU(2) \times SU(3)$. The corresponding Lagrangian depends on 19 parameters, whose numerical values are determined by experiments. Even though the SM is the best description of the subatomic world so far, it is not the complete picture. For example, it does not incorporate another fundamental force, gravity. Furthermore, there are still important questions that the SM does not answer, like neutrino oscillation, the matter-antimatter symmetry, the nature of dark matter and dark energy, etc. The new information from the LHC and other high energy experiments will facilitate the search for clues to more of these missing pieces.

2.2 Quantum Chromodynamics

Quantum chromodynamics is based on the non-abelian gauge theory of $SU(3)$. It provides the description of strong interactions between quarks and gluons, which make up hadrons. Color charge is the generator of $SU(3)$, which includes three kinds of color typically designated as red, green, and blue. To preserve the local symmetry, eight gauge bosons (gluons) are introduced to function as the carriers mediating the strong interaction. Two phenomena, confinement and asymptotic freedom, are peculiar properties of QCD. All features of QCD dynamics are encoded in the corresponding Lagrangian:

$$\mathcal{L} = -\frac{1}{4} F^a_{\mu\nu} F^{\mu\nu}_a + \bar{\psi}^i_q (i(\gamma^\mu D_\mu))_{ij} \psi^j_q - m_q \bar{\psi}^i_q \psi_{qi}, \tag{2.2}$$

where $F^a_{\mu\nu}$ is the field strength tensor for the gluon field A^a_μ defined as:

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_s f^{abc} A^b_\mu A^c_\nu \tag{2.3}$$

and D_μ is the covariant derivative in QCD defined as:

$$(D_\mu)_{ij} = \delta_{ij} \partial_\mu - i g_s t^a_{ij} A^a_\mu. \tag{2.4}$$

The notations in above equations is explained as following:

- A^a_μ is the gluon field with color index a and Lorentz vector index μ;
- g_s is the strong coupling and determines the strength of strong interaction and α_s defined by $\frac{g_s^2}{4\pi}$ is often used as the effective QCD coupling;
- f^{abc} are the structure constants of $SU(3)$;
- ψ^i_q are the quark field with color index i and flavor q;
- γ^μ are the Dirac matricies expressing the vector nature of strong interaction;
- t^a_{ij} are half of the Gell-Mann matrices of $SU(3)$, which are hermitean and traceless;
- m_q is the quark mass.

In Equation 2.2, when the first term is expanded by substituting Equation 2.3, this part not only contains quadratic terms in the derivatives of the gauge boson, but also includes the following terms:

$$g_s f^{abc}(\partial_\mu A^a_\nu) A^b_\mu A^c_\nu - \frac{1}{4} g_s^2 f^{abc} f^{ade} A^b_\mu A^c_\nu A^d_\mu A^e_\nu \tag{2.5}$$

The non-Abelian terms of the form $g_s f^{abc} A^b_\mu A^c_\nu$ gives rise to the self-interaction of gauge bosons via both three-point and four-point interactions. These self-interactions are the key difference between abelian and non-abelian gauge theories. The second term in Equation 2.2 demonstrates the kinematics of quarks and their interaction via the gluon field with color charge changing interactions. The third term describes the free quarks with mass m_q.

When calculating physical quantities in QCD, as in other quantum field theories, renormalization is applied in order to avoid ultraviolet divergences. In the renormalization procedure, the coupling is redefined by the renormalized coupling α_s and a finite scale μ as $\alpha = \alpha(\alpha_s, M/\mu)$, where M is the ultraviolet cutoff. All the divergences are absorbed in the renormalized coupling. Then the physical quantity does not depend on the ultraviolet cutoff M. In QCD, the coupling runs logarithmically with the energy scale by following the beta function:

$$\mu^2 \frac{d\alpha_s}{d\mu^2} = \frac{d\alpha_s}{d \ln \mu^2} = \beta(\alpha_s) = -\alpha_s^2(b_0 + b_1\alpha_s + b_2\alpha_s^2 + ...), \tag{2.6}$$

where $b_0 = \frac{11C_A - 4n_f T_R}{12\pi} = \frac{33-2n_f}{12\pi}$ defines the 1-loop beta-function coefficient. The first term is due to the contribution from gluon loops, the second is due to quark loops, and n_f is the number of contributing flavors. Beyond b_0, b_1 and b_2 refer to 2-loop and 3-loop coefficients. By setting a reference scale $\mu^2 = M_Z^2$, Equation 2.6 can be solved as:

$$\alpha_s(\mu^2) = \alpha_s(M_Z^2)\frac{1}{1 + b_0\alpha_s(M_Z^2)\ln\frac{\mu^2}{M_Z^2} + \mathcal{O}(\alpha_s^2)}, \tag{2.7}$$

where $\alpha_s(M_Z^2) = 0.1184 \pm 0.0007$. A consequence of the fact that $b_0 > 0$ is that the QCD coupling decreases with higher energy scale. The slope of the running coupling becomes flatter with energy scale crossing more massive quark flavor thresholds, because the small coupling strength at high energy scales is proportional to n_f. This explains the phenomenon of asymptotic freedom, which allows accurate predictions via perturbation theory. On the other hand, running the coupling towards the direction of lower energy scales, Equation 2.7 can approximately be rewritten as:

$$\alpha_s(\mu^2) = \frac{1}{b_0 \ln\frac{\mu^2}{\Lambda^2}}, \tag{2.8}$$

where $\Lambda \sim 200MeV$. This shows that the value of coupling diverges rapidly below 1 GeV, where the perturbative approach is no longer valid. The running of the coupling constant is shown in Fig. 2.5.

In the hard scattering process, the partons (quarks and gluons) are asymptotically free in short time scales equivalent to high energy scales, and perturbative QCD (pQCD) can be used to calculate the parton-level cross sections. The corresponding Matrix Elements (MEs) can be calculated systematically at fixed orders of the strong coupling α_s. Beyond the lowest order calculation of pQCD, two kinds of corrections need to be taken into consideration. Due to the radiation of a real parton by the initial or final state particle, additional jets change the topology of the final state. This effect refers to real-emission corrections. There also exist configurations of collinear and soft additional patrons causing infrared (IR) divergences in the QCD calculation. Secondly, the loop factors coming from virtual partons change the number of available paths through phase space. This effect is called as virtual corrections. There is also an IR

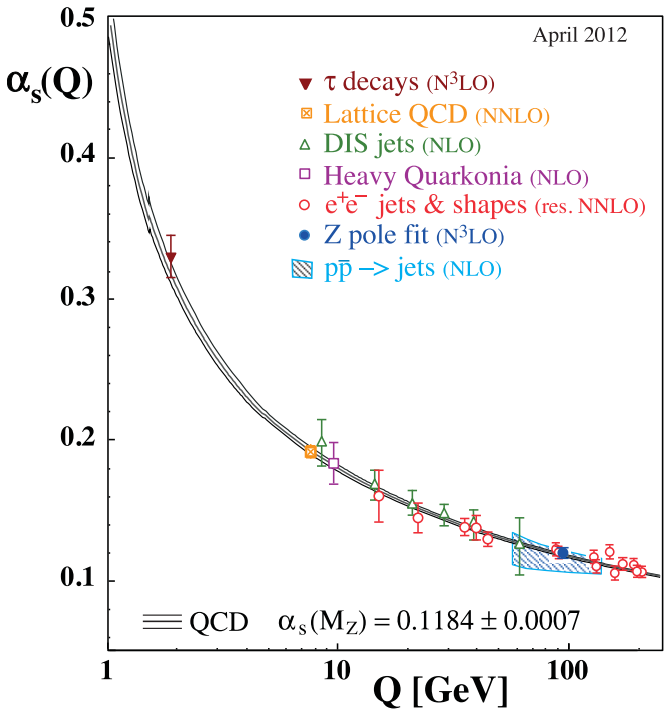


Figure 2.5: The evolution of strong coupling α_s as a function of energy scale Q. The measurements from experiments are shown in the legend and the value of strong coupling at energy scale of M_Z is shown at the bottom of the figure. (1)

divergence in the extra phase space from loop integrals. Notably, when the real-emission and virtual corrections are combined together, stated by the Kinoshita-Lee-Nauenberg (KLN) theorem, the IR divergence from the two factors cancel exactly against each other and make the whole pQCD calculation IR finite.

When the process involves hadrons in the initial state, the partons are confined in the hadrons. Perturbative QCD is not able to predict the parton content in the hadron. The Parton Distribution Functions (PDFs) $f_i(x, \mu_F^2)$ are introduced to describe the probability of finding a parton with specific flavor that carries a fraction x of the hadron momentum in the hadron. The factorization scale μ_F separates the perturbative region of the PDF model from the divergent collinear region, which may be absorbed in the parton functions as a normalization. The cross sections in processes with initial-state hadrons are calculated by convoluting parton-level cross sections with the corresponding hadron PDFs. Since PDFs are non-perturbative and not feasible to calculate in lattice QCD, they are parameterized and fitted based on experimental data, mainly from Deep Inelastic Scattering (HERA), fixed target and Hadron Collider experiments (TEVATRON and LHC), as shown in Figure 2.6. The Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equation can be used to extrapolate the PDFs from one perturbative resolution scale to another. An example is given in Fig. 2.7, which shows the MSTW NLO PDFs at scales $\mu_F^2 = 10\ GeV^2$ and $\mu_F^2 = 10^4\ GeV^2$ (2). At low $Q^2 = 10\ GeV^2$, the proton structure is dominated by valence quarks specially at $x = 0.2$, while at high $\mu = 100GeV$ large gluon and sea-quark distributions yields increase with small x value and valence quarks make a progressively smaller contribution.

2.3 Prompt Photon Production

The prompt photon production mechanism at the LHC is discussed in this section. At Leading Order (LO) in pQCD, prompt photons are produced by the direct photon process and the fragmentation photon process. The direct (or prompt) photon process includes the Compton-like scattering process $qg \to q\gamma$ and the quark anti-quark annihilation process $q\bar{q} \to g\gamma$. The Feynman diagrams of these two processes are shown in Fig. 2.8a. From the topological point of view, the recoiling photon and associated hadronic object in the final state are most likely back-to-back in the plane transverse

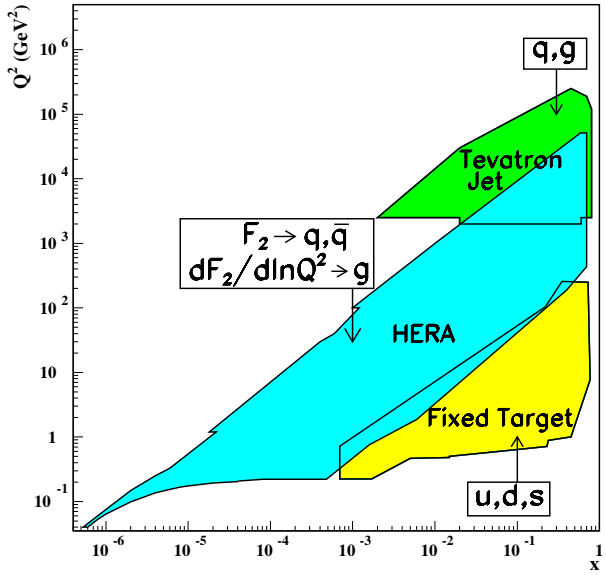


Figure 2.6: Kinematic regions of Q^2 versus x probed by fixed target and collider experiment, shown together with the parton distribution regions that are most strongly constrained by various data (1)

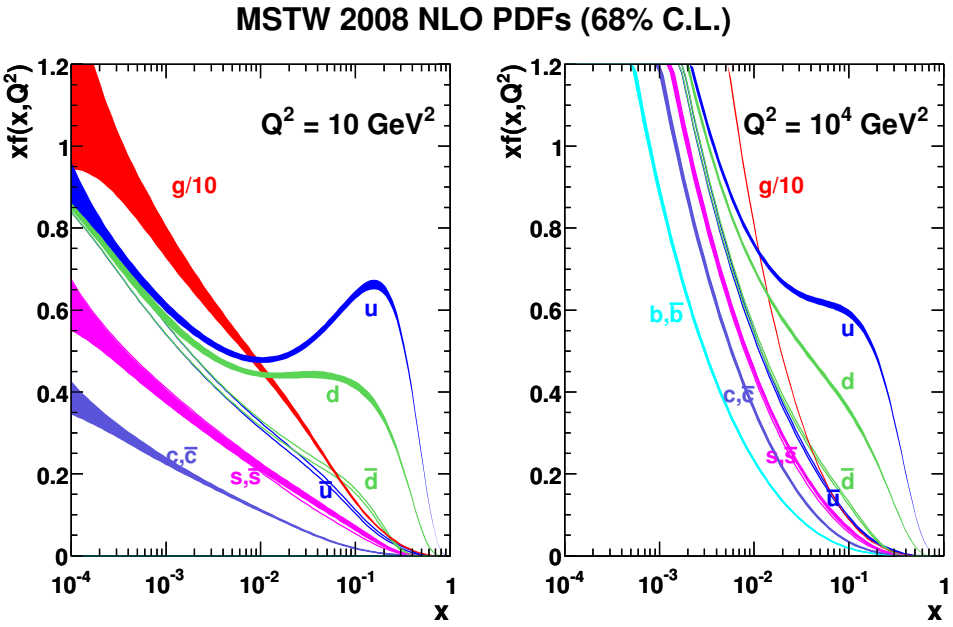


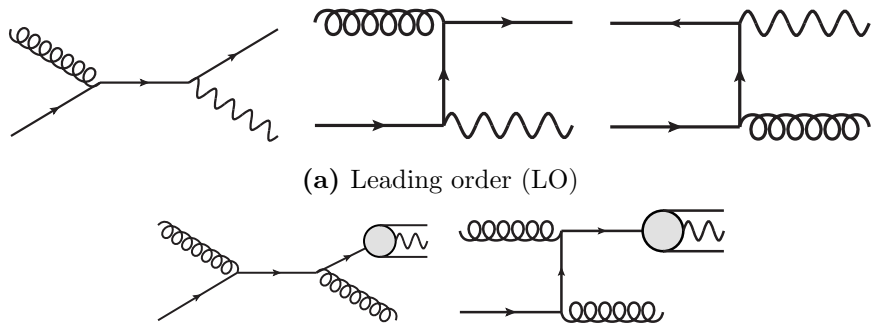
Figure 2.7: MSTW 2008 NLO PDFs at $Q^2 = 10$ GeV^2 and $Q^2 = 10^4$ GeV^2. (2)

to the beam direction and equal in transverse momentum. For the fragmentation photon process, a photon is produced by the fragmentation of a high momentum parton, examples of which are shown in Fig. 2.8b . In this process, the photon is most probably accompanied by a hadron shower. Most photons from fragmentation are suppressed by applying a photon isolation requirement. Only the case where the photon carries away most of the momentum of the fragmenting parton contributes to the measurement of prompt photon production. At LO, the differential cross section for prompt photon production can be written as:

$$\frac{d^2\sigma}{dp_T^\gamma d\eta^\gamma} = \frac{d^2\sigma^D}{dp_T^\gamma d\eta^\gamma} + \sum_a \frac{d^2\sigma_k^F(\mu_R, \mu_F, \mu_f)}{dp_T^a d\eta^a} \otimes D_{\gamma/a}(\mu_F), \tag{2.9}$$

where σ^D is the cross section of the direct part and σ^F describes the fragmentation part. For the fragmentation part, the cross section for photon production depends on the parton type and corresponding fragmentation function $D_{\gamma/k}$.

Beyond LO, there are higher order contributions to direct photon production. Next-to-leading order (NLO) calculations include gluon radiation in initial, intermediate or



(a) Leading order (LO)

(b) Fragmentation

Figure 2.8: Leading order (LO) and fragmentation Feynman diagrams for photon production.

final states, and virtual loops in the Feynman diagrams. Some of these examples are shown in Fig. 2.9.

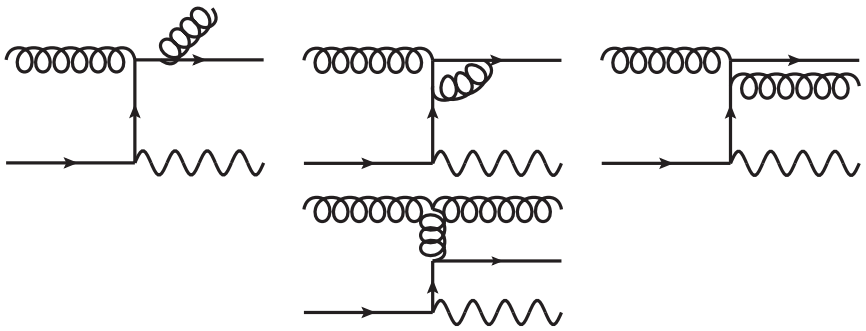


Figure 2.9: Next-to-Leading order (NLO) direct photon production.

2.4 Motivation of Prompt Photon Measurement

Prompt photons generated from hard interaction processes have a clean signal signature without local hadronization signatures, which is different from jets coming from quarks and gluons. Prompt photon production in association with jets in proton-proton collisions provides an important testing ground for pQCD predictions at large hard-scattering scale (Q^2) and over a wide range of the parton momentum fraction (x). In particular, prompt photon production in the LHC environment is dominated by the

Compton scattering process, the rate of which is sensitive to the gluon content in the proton. The measurements of this process may be used to improve constraints to the parton distribution functions of the proton (13). The higher collision energy of the LHC and the wider rapidity range of the CMS detector allow us to access lower ranges of parton momentum fraction, compared to previous collider experiments. In the leading order approximation, the $x_{1,2}$ regions of parton momentum fraction space of two initial interacting partons can be determined by $x_{1,2} = \frac{p_T^\gamma}{\sqrt{s}}(\exp(\pm y^\gamma) + \exp(\pm y^{jet}))$ (21). For example, at $p_T^\gamma = 40\ GeV$ and $-2.5 < y^{\gamma,jet} < 2.5$, the $y^\gamma \times y^{jet} > 0$ region covers adjacent x_1 and x_2 intervals ($0.00093 < x_1 < 0.01142$ and $0.01142 < x_2 < 0.13922$), while for events with $y^\gamma \times y^{jet} < 0$, the x_1 and x_2 have the same intervals ($0.00618 < x_1, x_2 < 0.07532$).

Another benefit of the photon plus jet analysis is for measurements of a low mass Higgs boson. At the LHC, a Higgs boson decaying to di-photons provides a clean signature and serves as one of the most powerful channels for the detection of a Higgs boson around the low mass of 125 GeV (14), some examples of which are shown in Fig. 2.10. Photon plus jets events are one of the main contributions to the background of the di-photon signature. The calculation of the K-factors for the photon plus jets background for $H \to \gamma\gamma$ analysis requires the understanding of the cross section measurement of photon plus jets directly (14).

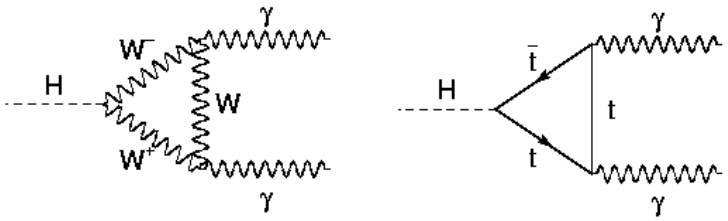


Figure 2.10: The decay of the Higgs boson into di-photon. In the SM, the Higgs boson does not directly couple to photons due to the masslessness of photons. So the decay of the Higgs boson to di-photon is mediated by a loop mediated process.

Photon plus jets events are also a major source of background to beyond-the-standard-model (BSM) searches. In the LHC, searches for dark matter (DM) rely on detection through the reaction of $q\bar{q} \to \gamma\chi\bar{\chi}$, where the photon is radiated by one of

the incoming quarks (15). In addition, the final state of $\gamma + \not{E}_T$ is sensitive to models of extra large spatial dimensions (15, 22). Additionally, in scenarios where quarks have excited states or themselves are composites, the presence of such excited quarks can potentially be accessed in photon plus jets final states (16). Another example comes from Supersymmetry (SUSY) searches. In the model of gauge mediated supersymmetry breaking (GMSB), the lightest neutralino decays to a graviton, which is the lightest supersymmetric particle (LSP), and a photon (23). A better understanding of photon plus jets events will be important for all these BSM searches.

In physics object reconstruction at the detector level, jet objects suffer poor energy resolution due to the parton hadronization process and the limited precision of the Hadronic Calorimeter. By comparison, especially in CMS, the high energy resolution of the Electromagnetic Calorimeter enables the accurate measurement of photon energy. So the production of prompt photons and jets from QCD provides a mechanism to calibrate the jet energy by taking advantage of the precision of photon energy measurement.

2.5 Monte Carlo (MC) Prediction Tools

In the high-energy collision environment, the structure of events is complex and difficult to predict from first principles. The use of MC event generators provides a powerful tool to subdivide the problem into more manageable pieces, some of which can be calculated from first principles, while others are described by appropriate models with parameters tuned to data. Compared with real life of experiments, where only the initial and final states are known, MC event generators provide a detailed description of the whole physical process in each event. Then the experimental observables can be predicted and compared with data. Therefore, MC event generators can be used to study various aspects of physics analysis, like trigger design, evaluating acceptances, planning analysis strategies, understanding of separating signal and background processes, estimating smearing effects of the detector, and so on.

For a typical MC event generator, an event includes the following building blocks (24, 25, 26):

- Hard Subprocesses: At the LHC, many interesting physical processes involve large momentum transfers. The simulation of this hard subprocess with large invariant

momentum transfer is the core component. In the high energy environment, where QCD has the property of asymptotic freedom, the hard subprocesses can be described by perturbation theory using Feynman diagrams. The cross section σ for a scattering subprocess $ab \to n$ at hadron collider can be described through

$$\begin{aligned} \sigma &= \sum_{a,b} \int_0^1 dx_a dx_b \int f_a^{h_1}(x_a, \mu_F) f_b^{h_2}(x_b, \mu_F) d\hat{\sigma}_{ab\to n}(\mu_F, \mu_R) \\ &= \sum_{a,b} \int_0^1 dx_a dx_b \int d\Phi_n f_a^{h_1}(x_a, \mu_F) f_b^{h_2}(x_b, \mu_F) \times \frac{1}{2\hat{s}} |M_{ab\to n}|^2(\Phi_n; \mu_F, \mu_R), \end{aligned} \tag{2.10}$$

where $f_a^h(x, \mu)$ represents the parton distribution functions; $\hat{\sigma}_{ab\to n}(\mu_F, \mu_R)$ is defined as the parton-level cross section for initial patrons a and b to produce final state n. This cross section depends on the final phase space $d\Phi_n$, the corresponding matrix element squared $|M_{ab\to n}|^2(\Phi_n; \mu_F, \mu_R)$ and the parton flux $\frac{1}{2\hat{s}} = \frac{1}{2x_a x_b s}$, where s is the hadronic center-of-mass energy squared. The matrix element $|M_{ab\to n}(\Phi_n; \mu_F, \mu_R)|$ can be calculated from the Feyman rules derived from the Lagrangian and depends on the final phase state $d\Phi_n$, the factorization scale μ_F, and the renormalization scale μ_R.

- Parton Showers: Hard subprocesses describe the process of large momentum transfer and the momenta of out-going jets. However, this is not sufficient to represent the exclusive physical process, which includes the internal structure of multi-jets with accompanying particles. Similar to QED photon radiation from accelerated charged particles, the accelerated colored partons will emit QCD radiation in the form of gluons, which will emit further radiation. These QCD radiations can be modeled by parton showering, which is an approximation scheme to provide higher-order corrections to the hard subprocesses. The parton shower simulation algorithm is formulated as an evolution, where momentum transfers are happening down from the high energy scale associated with the hard subprocess to the low energy scale, where patrons are bound into hadrons. The corresponding dominant contributions in the parton showers are collinear parton

splitting and soft gluon emission. These processes are associated with both incoming and out-going partons related to the hard subprocess, which are called final-state showers and initial-state showers (24, 25, 26).

- Hadronization: The cascading patrons generated by parton showers evolve to the low energy scale (∼1 GeV), where the QCD running coupling increases steeply. At this scale, perturbation theory becomes invalid. Hadronization in MC simulates the formation of observed final-state hadrons from the non-perpurbative phase space. Currently, there is no available non-perturbative technique for hadronization calculation. So phenomenological models are used to calculate hadronization process, like the *string model* (27, 28) and *cluster model* (29).

- Underlying Event (UE): Hadron collider events including a hard subprocess also have extra activity coming from UE. The main contribution of UE is multiple-parton-interactions (MPI) and pile-up. Other contributions include initial and final-state radiation and beam-beam remnants (BBR). Most of the MPI are soft interactions that cannot be reconstructed as jets directly. However, they will still affect the final-state activity by increasing the summed E_T distributions. The modeling of MPI is important for MC to match real experiment. The detailed MC model for perturbative MPI can be found in (24, 30). For the varying number of semi hard parton-parton interaction, the average interaction rate is calculated by perturbative QCD and the variation for different events are generated by Poissonian statistics (30). Pile-up is additional proton-proton interactions occuring in the same time (in-time pile-up) and previous bunch crossings (out-of-time pile-up), the possibility of which is proportional to the instantaneous luminosity. The particles produced by pile-up may overlap with those of the event of interest and influence the physics object reconstruction. The simulation of pile-up is studied by using minimum-bias data (31);

- Detector Simulation: In hadron collider experiments, event signals are collected through different types of particle detectors. The accurate and comprehensive simulations of particle detectors are of particular importance. The object-oriented simulation toolkit, *GEANT4*, has been developed to satisfy this need. *GEANT4*'s functionality includes tracking, geometry, physics models and hits, a large set of

long-lived particles, materials and elements, over a wide energy range (32). It describes the geometry and material of the detector elements. The kinematics of the particles are tracked through detectors by simulating their interactions in the detector matter.

Different programs on the market provide general-purpose event generators. They combine different approaches and models to simulate the components of the physics processes described above. In this analysis, three MC programs are used for varying purposes. They are PYTHIA, SHERPA, and JETPHOX separately.

- PYTHIA: The PYTHIA (33) event generator is a Leading-Order Monte Carlo generator. Leading-order matrix elements are used to simulate hard sub-processes, which involve 2 incoming particles and 1 or 2 outgoing particles. In this analysis, PYTHIA 6.424 is used for MC samples of the prompt photon (signal) and QCD (background). The corresponding sub-processes are summarized in Table 2.1.

Table 2.1: Prompt photon and QCD sub-processes generated by PYTHIA 6.424.

prompt photon sub-processes	QCD sub-processes
$q_i\bar{q}_i \to g\gamma$	$q_iq_j \to q_iq_j$
$q_ig \to q_i\gamma$	$q_i\bar{q}_i \to q_k\bar{q}_k$
$gg \to \gamma\gamma$	$q_i\bar{q}_i \to gg$
$gg \to g\gamma$	$gg \to q_k\bar{q}_k$
$g_i\bar{g}_i \to \gamma\gamma$	$q_ig \to q_ig$
	$gg \to gg$

- SHERPA: The SHERPA (34) event generator is a general-purpose LO Monte Carlo event generator. It merges a flexible tree-level matrix-element generator (AMEGIC++ and Comix) with a parton-shower scheme (CSSHOWER++) by Catani, Krauss, Kuhn and Webber (CKKW). This method enables SHERPA to deliver a good description of multi-jet final states and has been extensively validated with previous collider data.

- JETPHOX: JETPHOX (35, 36, 37) is a Next-to-Leading Order Monte Carlo event generator used to calculate the reaction p p $\to \gamma$/hadron + jet + X. A

combination of phase-space slicing and subtraction methods are applied in the calculation to model the soft and collinear singular parts of the perturbative matrix elements. JETPHOX is able to be interfaced with standard sets of PDFs, which enables the user to evaluate the dependence of the computed cross section on the PDF set. In this analysis, it is interfaced with LHAPDF version 8.5.4 in order to apply a recent set of the proton PDFs (CT10). JETPHOX also models the isolation requirements for photons at the parton level, which can be used to match the corresponding cut at the detector level.

Chapter 3

EXPERIMENTAL APPARATUS

3.1 Large Hadron Collider

The Large Hadron Collider (LHC) is a circular particle accelerator in Geneva, Switzerland, built by the European Organization for Nuclear Research (CERN). It resides in a tunnel of 27 kilometers in circumference and is built as deep as 175 meters underground, as shown in Fig 3.1. The LHC has surpassed all preceding colliders to become the world's largest and highest-energy particle accelerator. Its synchrotron is designed to collide two opposing proton beams at 7 TeV per beam. Chronologically, LHC achieved 1.18 TeV per beam and surpassed Tevatron's record of 0.98 TeV per beam in 2009. Energies were ramped to run at 3.5 TeV per beam in 2010 and kept running throughout 2011, providing the data used in this thesis. In 2012, the energy was successfully increased to 4 TeV per beam and continued operating at this energy till the end of 2012. With such high collision energy, the LHC provides an unprecedented opportunity to explore the fundamental laws of the universe and address various Big Questions of physics, including "what is mass"; "What is 96 % of the universe made of?"; "Why is there is so little antimatter?"; "What was matter like within the first moments of the Universe?"; "Do extra dimensions of space really exist?", and so on. Moreover, proton-proton collisions at such high energies also allows us to understand more details of the existing Standard Model. Especially, investigation of QCD interactions not only provides insight into parton distribution, but also allows tests for new phenomena and

scales of interactions, such quark compositiness.

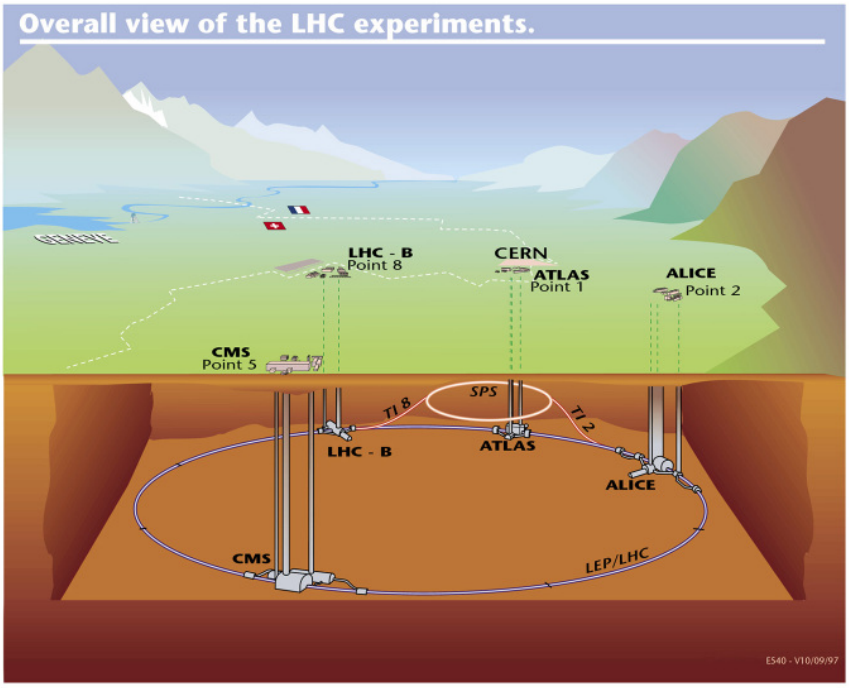


Figure 3.1: Overall view of the LHC accelerator.

In order to explore the different types of physics questions mentioned above, there are six experiments conducted at the LHC, as shown in Fig 3.2. Each of them is characterized by its special particle detectors. CMS and ATLAS are two general-purpose detectors designed to capture particle signals produced by collisions in the accelerator. The goal of these is to investigate the largest range of physics produced in these high energy data. The two independently designed detectors can be used to cross-confirm any new discoveries. More detail about CMS will be provided in the next section. By comparison to the above two large-size experiments, there are two medium-size experiments, A Large Ion Collider Experiment (ALICE) and the Large Hadron Collider Beauty (LHCb), which serve more specific functions. ALICE studies the physics of strongly interacting matter at extreme energy densities, while LHCb specializes in studying the "beauty quark" in order to investigate the matter-antimatter asymmetry of the universe. In addition, the TOTal Elastic and diffractive cross section Measurement (TOTEM) and Large Hadron Collider Forward (LHCf) detectors are

much smaller in size and focus on the production of "forward particles" in the high energy collisions.

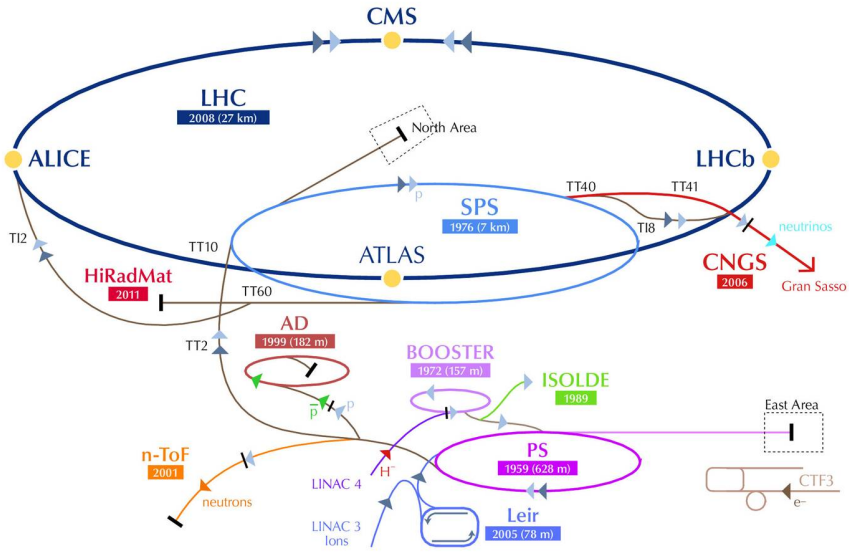


Figure 3.2: The LHC accelerator complex at CERN.

In addition to the collision energy, luminosity is another important feature of an accelerator, determining the production rate of final states from the physics collisions. The higher the luminosity, the more events will be generated for study. The luminosity of colliding bunches is described using following equation:

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

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 normalized transverse beam emittance, 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 peak luminosity designed for the LHC is $\mathcal{L} = 10^{34} cm^{-2} s^{-1}$ and data from the 2011 run spanned running conditions with a maximum instantaneous luminosity of 3.55×10^{33} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Given the instantaneous luminosity profile, the integrated luminosity can be calculated by $L_{int} = \int \mathcal{L} \mathrm{d}t$, which corresponds to the number of collisions during a time interval. The cumulative integrated luminosity within 2010, 2011 and 2012 is shown

in Fig. 3.3. In run 2011A completed in August 2011, the LHC delivered an integrated luminosity of 2.63 fb^{-1} and CMS recorded 2.38 fb^{-1} providing the dataset used in this thesis.

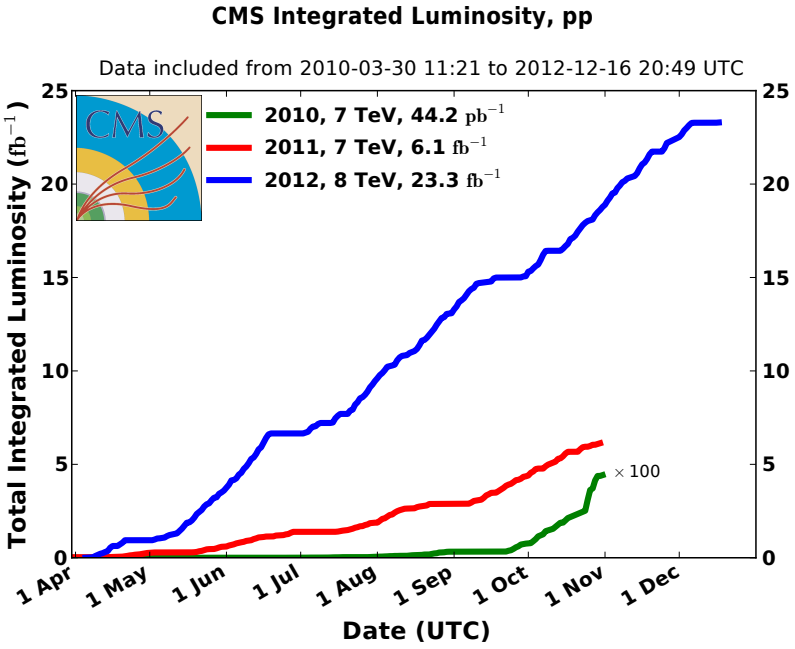


Figure 3.3: Cumulative integrated luminosity versus time in 2010, 2011 and 2012.

3.2 Compact Muon Solenoid

The Compact Muon Solenoid (CMS) experiment is a general-purpose particle physics apparatus for detecting particles produced in LHC collisions. The CMS detector enables us to perform comprehensive tests of the Standard Model (SM) and to search for Higgs bosons and new physics at the TeV energy scale. It is located at Point 5 of the LHC ring near the town of Gex, France, while ATLAS symmetrically sits at Point 1 of the ring adjacent to the main CERN site. CMS is 21 meters long, 15 meters height and weighs about 12,500 tons.

Being a multi-purpose particle physics experiment, it is designed to have the following features.

- Good charged particle tracking and large magnetic field to bend charged particles and measure their momentum.
- High granularity of electromagnetic and hadronic calorimeters to determine particle position.
- Good time resolution of detectors to function in the high frequency particle collision environment.
- Good energy resolution of calorimeters to precisely measure the energy of photons and electrons.
- Good muon identification system.
- Radiation-hard detectors and front-end electrons to withstand the large flux of particles coming from high intensity particle collision.

CMS is a state-of-art detector satisfying the above requirements and its distinguishing features includes a high-field solenoid, a full-silicon-based inner tracking system and homogeneous scintillating-crystals-based electromagnetic calorimeter (Fig. 3.4). The CMS tracking system, including pixel detector and silicon tracker, forms the inner-most part of CMS. Electromagnetic and hadronic calorimeters follow the tracking system. The above three sub-detector systems are located within a solenoid magnet producing a field of 3.8 T. The outer-most layer of CMS is the muon detection system, which is embedded in the steel return yoke of the solenoid and resides in the field of about 2 T.

Before introducing the details of each sub-detector, it is useful to have a coordinate system used to describe the CMS geometry, as shown in Fig 3.5. The origin is centered at the nominal collision point at the center of the detector. The x-axis points radially towards the center of LHC ring; the y-axis points vertically upward; and the z-axis points along the beam direction and satisfies the right-hand rule. Besides this Cartesian coordinate system, cylindrical and spherical coordinate systems are often used in high energy experiments. The azimuthal angle ϕ is measured from the x-axis in the x-y plane; the radial coordinate in this plane is denoted by r; and the polar angle θ is measure from the z-axis. A number of derived variables are also commonly used. Pseudorapidity is defined as $\eta = -\ln\tan(\theta/2)$; the momentum and energy transverse to z-axis is denoted

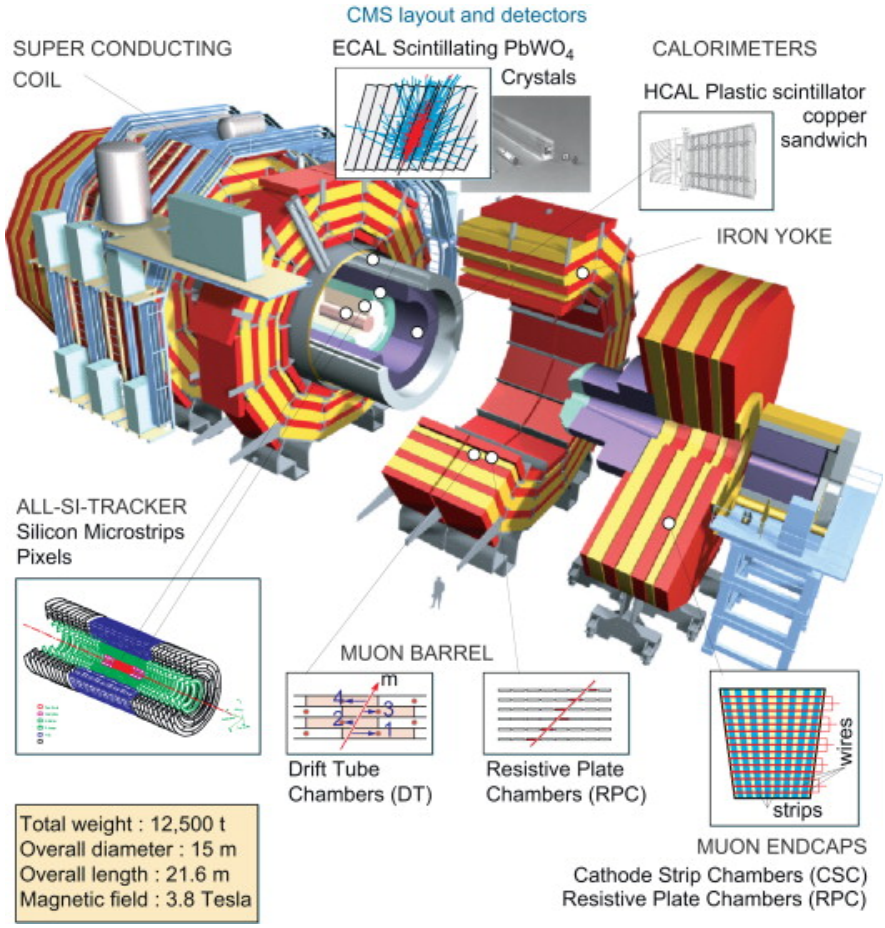


Figure 3.4: The components of CMS (3)

by p_T ($p_T = \sqrt{(p_x^2 + p_y^2)}$) and E_T ($E_T = E \sin\theta$); and the imbalance of energy measured in the transfer plane is denoted by E_T^{miss} ($E_T^{miss} = \not{E}_T = -\sqrt{((\sum_{i=1}^{n} E_x^i)^2 + (\sum_{i=1}^{n} E_y^i)^2)}$).

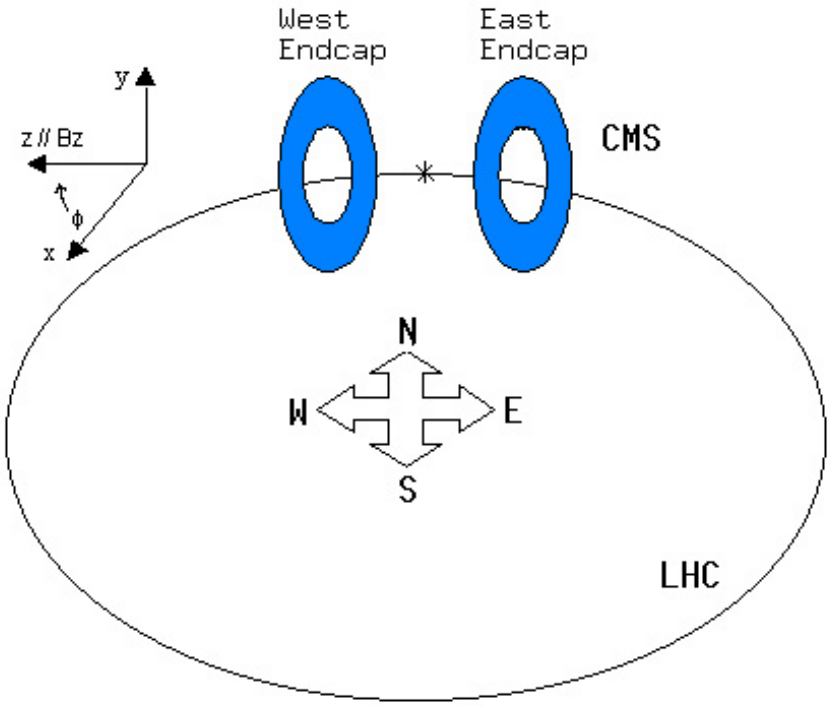


Figure 3.5: The coordinate system of CMS.

3.2.1 Tracking Detector

The inner most sub-detector of CMS is the tracking system, which is designed to precisely measure charged particle trajectories and the vertices of hard interactions (38). Operating at the designed LHC instantaneous luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$, the collisions will generate thousands of particles from an average of more than 20 proton-proton interactions occurring within the tracking system every 25 ns. In this environment, the technology of the tracking system is required to feature high granularity, fast response time, and radiation hardness. In addition, the detector material should be kept at a minimum so that multiple scattering, photon conversion and nuclear interactions can be minimized within the tracking volume. Based on these requirements, silicon detector

technology is applied to construct the tracking system in CMS. The system includes two types of trackers, a silicon pixel tracker and silicon strip tracker. When a charged particle traverses the trackers, it causes ionization currents and the pixel/strip trackers will generate electric signals. Then the corresponding signals will be amplified and detected as hits within the detector.

The pixel tracker is closest to the beam collision region. It is essential for the reconstruction of secondary vertices, which may come from b and τ decay. It also provides precise measurement of tracking points in r, ϕ and z, which form seed tracks for outer track reconstruction and high level triggering. The pixel cells on this tracker have size of $100 \times 150\ \mu m^2$. The entire pixel tracker covers the pseudorapidity range from $-2.5 < \eta < 2.5$. It is composed of three barrel layers (BPix) and two endcap disks (FPix), show in Fig 3.6. The BPix layers have length of 53-cm and are located at mean radii of 4.4, 7.3 and 10.2 cm. The FPix disks are placed on each side of the beam collision region at $z = \pm 34.5$ and $z = \pm 46.5$ cm and cover the radius range from 6 to 15 cm. The semi-conductor technique used here is an n+ pixel on n-substrate detector design that is able to operate while being partial depleted (39). For the barrel layers, the electrons will have a Lorentz drift to the collecting pixel implant due to the magnetic field of CMS. This drift will lead to charge spreading over more than one pixel. The signal pulses from these pixels are used for charge interpolation, which improves the spacial resolution in the position on individual hits to 15-20 μm. For the FPix layers, the detectors are tilted at 20° to induce charge-sharing. This charge-sharing between neighboring pixels enables a position resolution of about 15 μm. With radiation damage, a reduction of the depletion depth or the increase in bias voltage will reduce the effect of charge-sharing and influence the future spacial resolution.

At increasing radius, the silicon strip tracker follows the pixel detector. It is comprised of single sided p-on-n type silicon micro-strip sensors. Like the pixels, the strip sensor generates signal currents when ionized by a charged particle. The whole construction of the silicon strip tracker includes four components, Tracker Inner Barrel (TIB), Tracker Inner Disks (TID), Tracker Outer Barrel (TOB) and Tracker Outer Endcap (TEC). The TIB consists of four concentric cylinders with length of 1400 mm along the z axis and located at radii of 255.0, 339.0, 418.5 and 498.0 mm. The TID is composed of three disks at both sides of the beam collision region between $z = \pm 800$ mm and ± 900 mm and spans a radius from about 200 mm to 500 mm. The TOB

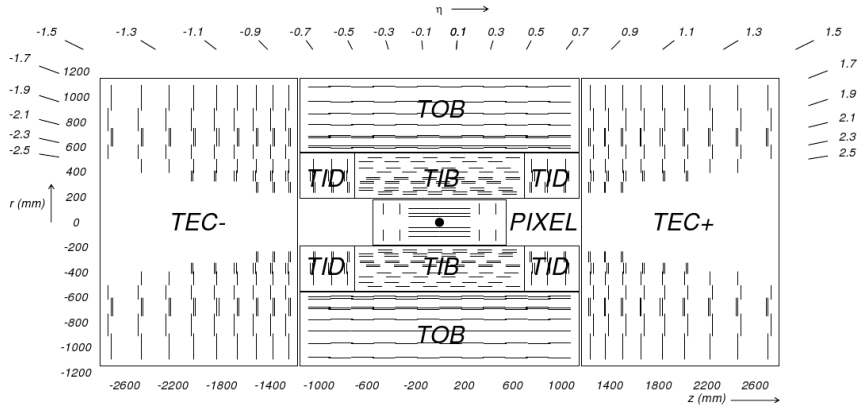


Figure 3.6: Cross-section view of the CMS tracker in the $y(r) - z$ plane. The lines in the pixel detector represent ladders and paddles where the detectors are mounted. And the lines in the strip tracker represent the silicon detectors.(40)

has three outer and three inner cylinders with a length of 2180 mm along the z axis and covers a radius from 555 mm to 1160 mm. The TEC includes nine disks between $z = \pm 1240$ and ± 2800 at each side of the collision region and extends radially from 220 mm to 1135 mm. The whole silicon strip tracker covers the pseudorapidity region $-2.5 < \eta < 2.5$.

To measure the positions of tracks with high resolution, a precise knowledge of the orientation of tracker modules is also very important, because deviations from nominal positions can be caused by the assembly precision, deformation due to temperature effects and magnetic field, etc. A precise alignment is required to determine the true module positions and orientation. The alignment process takes advantage of data from the tracker assembly and the Laser Alignment System (LAS). The LAS uses infrared laser beams with a wavelength of 1075 nm to calibrate the position of some tracker modules. And the system generates alignment information for tracker substructures (TIB, TOB and TEC discs) on a continuous basis at the level of 100 μm. Meanwhile, two well-known track-based alignment algorithms are used to calculate true positions, the Millepede algorithm (41) and Kalman filter (40). The whole alignment system ensures that good measurement resolution is maintained for tracking parameters.

3.2.2 Electromagnetic Calorimeter

The ECAL is the most important detector for the photon analysis in this thesis. One of the motivations of its design is the search for a low-mass Higgs Boson that decays to a final state with two photons. The ECAL is required to have fast response, fine granularity, radiation resistance, good energy resolution and to operate in 4 T magnetic field (38). The ECAL is a hermetic and homogeneous calorimeter composed of 61,200 lead tungstate $PbWO_4$ crystals mounted in the central barrel part and 7,324 crystals in each of two endcaps. The barrel part of the ECAL covers the pseudorapidity range $|\eta| < 1.479$. The front faces of crystals are positioned at a radius of 1.29 m and have a cross-section of approximately 0.0174×0.0174 in $\eta - \phi$ or 22×22 mm^2. The end caps are oriented vertically and cover the rapidity range of $1.479 < \eta < 3.0$. The longitudinal distance between the front plane of the end cap plane and the beam collision center is 3.154 m.

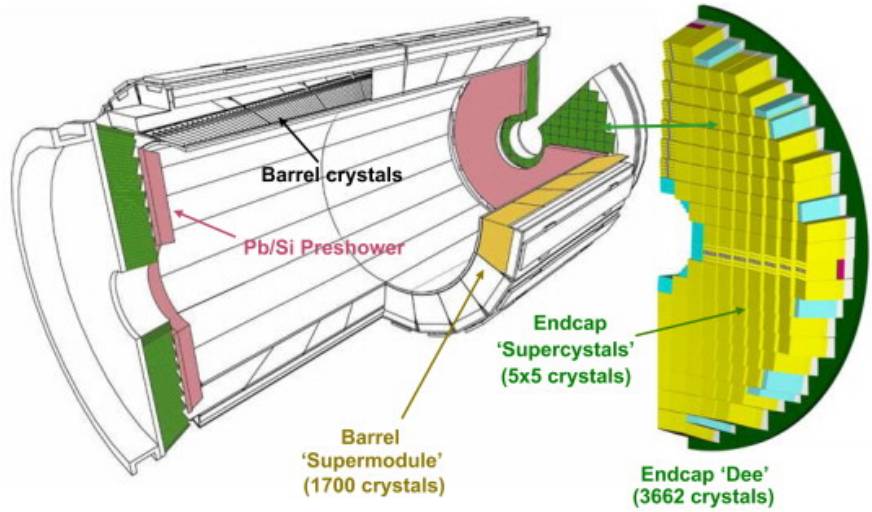


Figure 3.7: Layout of the CMS ECAL, showing the barrel supermodules with coverage up to $\eta = 1.48$, the two endcaps with coverage up to $\eta = 3.0$ and the preshower detectors with coverage of $1.65 < \eta < 2.6$. (42)

The $PbWO_4$ crystals have a high density (8.28 g/cm^3), short radiation length (0.89 cm) and small Moliere radius (2.2 cm). Exploiting these characteristics the ECAL is constructed as a compact calorimeter with fine granularity. Additionally, about 80% of the light from scintillation decay in the crystals is emitted in 25 ns, which is of the same

scale as the LHC bunch crossing time. The emitted photons are measured using two types of photodetectors. Avalanche photodiodes (APDs) are used as photodetectors in the barrel and vacuum photodiodes (VPTs) in endcaps. An APD is a electronic device that use semiconductor materials to convert light to electricity according to photoelectric effect with a high gain through avalanche multiplication. In the barrel, a pair of APDs is mounted on each crystal and each with an active area of 5x5 mm^2. The APD use a bulk n-type silicon structure behind the p-n junction and are operated at a gain of 50. VPTs are vacuum phototriodes, which have a single gain stage of about 10.2 at zero field. The corresponding mean quantum efficiency is about 22% at 430 nm. In the endcaps, VPTs are glued to the back of each crystal. Each of these is 25 mm in diameter and has an active area of approximately 280 mm^2. Compared to the APDs, VPTs have lower quantum efficiency and internal gain. But this is offset by their larger active area. There are preshower detectors of 20 cm thickness installed in front of the endcap crystals to assist in identifying neutral pions within a fiducial region of $1.653 < \eta < 2.6$. They also improve the granularity of the position determination for photons and electrons. There are two layers in the preshower. A lead radiator initiates electromagnetic showers from incoming photons and electrons, after which silicon strip sensors measure the deposited energy and the transverse shower profiles.

The energy resolution of the ECAL system can be parametrized as in the following equation (43)

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

where S is the stochastic term, N the noise term, and C the constant term. Contributions to the stochastic term include event-to-event fluctuations in the shower development, photon statistics contribution and measurement uncertainty of energy deposited in the preshower absorber. The noise term models effects of electronics noise, digitization and pileup. The constant term arises from non-uniformity of longitudinal light collection, inter-calibration errors and leakage of energy from the back of the crystal. The parameters in the above equation also depend on the energy reconstruction algorithm. For energy constructed by summing 3x3 crystal arrays, the energy resolution was found to be (43):

$$(\frac{\sigma}{E})^2 = (\frac{2.8\%}{\sqrt(E)})^2 + (\frac{12\%}{E})^2 + (0.3\%)^2, \tag{3.3}$$

where E is in GeV. In this case, the energy resolution for 120 GeV electrons or photons is about 0.5 %. The precise measurements of the ECAL system facilitate the study of photons and electrons in collider events.

3.2.3 Hadronic Calorimeter

The hadronic calorimeter (HCAL) resides between the outer extent of the ECAL (r=1.77m) and the inner surface of the solenoid (r=2.95m) (38). The HCAL measures hadron jets and contributes to indirectly detect neutrinos through apparent missing transverse energy. The HCAL is designed to be a non-magnetic compact absorber and have a high radiation tolerance. The HCAL is composed of four parts: HCAL Barrel (HB), HCAL endcap (HE), HCAL outer (HO), and HCAL forward (HF). The HB, HE and HO sections are composed of a brass absorber/plastic scintillator structure, while a *Čerenkov* detector made of quartz fibers is used in the HF to withstand the higher radiation environment.

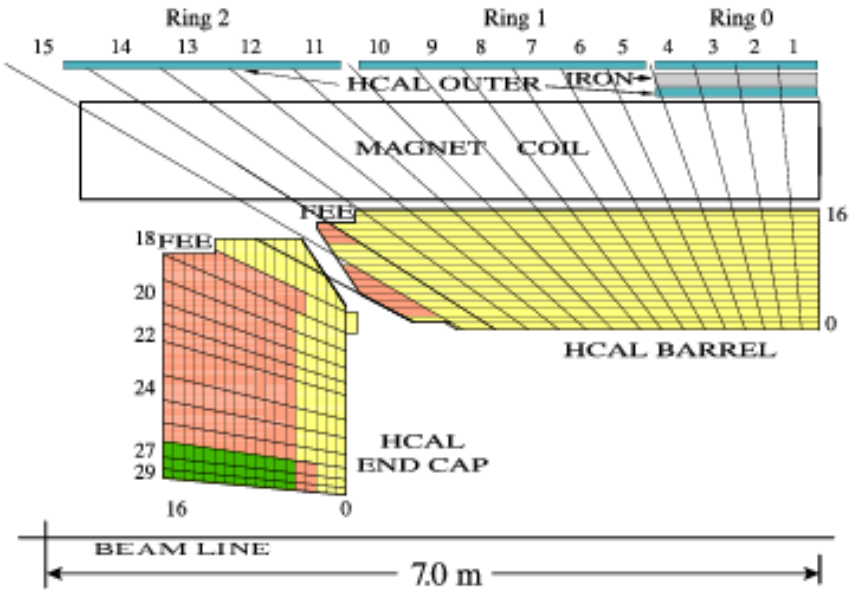


Figure 3.8: A quarter cross-section view of the CMS HCAL detector in $y-z$ plane.(4)

The HB covers the pseudorapidity range $|\eta| < 1.305$ with towers of $\delta\eta \times \delta\phi = 0.087 \times 0.087$ in a single longitudinal sampling layer. HO is located outside the solenoid

to provide sufficient containment for hadron showers and functions as tail catcher. The HE spans the pseudorapidity range of $1.3 < |\eta| < 3$, and the correpsonding granularity varies from $\delta\eta \times \delta\phi = 0.087 \times 0.087$ for $\eta < 1.6$ to $\delta\eta \times \delta\phi = 0.17 \times 0.17$ for $\eta > 1.6$. The HB, HE and HO are all tile-fibre sampling calorimeters with alternating layers of brass absorber and plastic scintillator. When hadron jets traverse the HCAL, the hadronic shower is generated by the absorber layers. Then the shower particles initiate scintillation light when they travel through the scintillator layers. Finally, wavelength-shifting (WLS) fibers connected to the scintillator tiles read out the light signals, which are amplified by hybrid photodiodes (HPDs).

In the forward pseudorapidity region of $3.0 < |\eta| < 5.0$, the radiation environment becomes extremely harsh. The HF is built using different techniques from other HCAL components. Quartz fibers and steel are chosen as the active and passive media. The particle showers, initiated in the steel, generate Cherenkov light in the quartz fiber, when they are above the Cherenkov threshold ($E \geq 190$ keV for electrons). The generated Cherenkov light is transferred to photo-multiplier tubes (PMT) through light-guide fibers. To exploit the different longitudinal profiles of showers based on the gamma factor of incident particles, long quartz fibers (165 cm) and short quarts fibers (143 cm) are used together in the HF to distinguish the contributions from EM and hadronic showers.

The energy resolution of HCAL can be expressed by the following equations:

$$(\frac{\sigma}{E})^2 = (\frac{90\%}{\sqrt{(E)}})^2 + (4.5\%)^2, \text{for the HB and HE sections} \tag{3.4}$$

and

$$(\frac{\sigma}{E})^2 = (\frac{198\%}{\sqrt{(E)}})^2 + (9.0\%)^2, \text{for the HF section.} \tag{3.5}$$

The first term on the right hand side is the stochastic term describing statistical fluctuations and showers fluctuations. The second term is the constant term accounting for detector non-uniformity and calibration uncertainty. By comparison with ECAL, absorber and scintillator are separated in HCAL, which increases the sampling fluctuations and reduces energy resolution. And the layer structure provides longitudinal information for the particle shower, which is not accessible in the ECAL system.

3.2.4 Muon Detectors

Muons are important signatures for decays of heavy particles such as a standard model Higgs boson, which can decay into ZZ with a Z further decaying into two muons. Unlike other electromagnetic and hadronic particles, most of which can be captured by the ECAL and HCAL systems, muons can penetrate through these calorimeters with minimum energy loss for γ factors typical at LHC energies. So the muon system is designed to identify muons by tracking their trajectories beyond the calorimeters and to provide momentum measurement and triggering. Like the inner tracker system, the high magnetic field in the CMS return yoke enables good muon momentum resolution and trigger capability based on transverse momentum measurements. Inside the muon system, three different types of gaseous particle detectors are implemented for detection of muons (38).

In the barrel region, drift tube (DT) chambers span the pseudorapidty range $|\eta| < 1.2$ and include four stations, three of which measure muon position in the $r - \phi$ plane and remaining one identifies the position in z direction. The tubes used here have cross section of 42×13 mm^2 and are filled with Ar and CO_2. An anode wire in the center of each tube is surround by aluminum cathodes. When a muon traverses through the gas volume, ionization electrons are collected by the central anode wires. The DT chambers have sufficient time resolution for associating muon hits with a particular beam collision.

In the higher pseudorapidty region of $0.9 < |\eta| < 2.4$, cathode strip chambers (CSC) are used to detect muons. These provide fast response time, fine segmentation, and radiation resistance. The CSCs have four stations in each endcap and each CSC contains six layers. Each layer has a plane of cathode strips running radially outward from the beam axis, providing a measurement of the $r - \phi$ plane coordinate. Perpendicular to the strips, anode wires enable the measurement of muon η.

To differentiate between multiple hits in a chamber (likely to occur at high luminosity), a complementary muon detector with fast response is required to provide accurate trigger information. Resistive plane chambers (RPC) fulfill this function and cover the whole barrel and partial endcap region for $|\eta| < 1.6$. The RPCs are double-gap chambers with two parallel plates of anode and cathode. The incoming muon will ionize electrons from a gas medium, which will cause a further avalanche of electrons across a

high voltage gap. This signal provides a fast measure of muon momentum and is used to make trigger decisions. The corresponding time resolution is about one nanosecond.

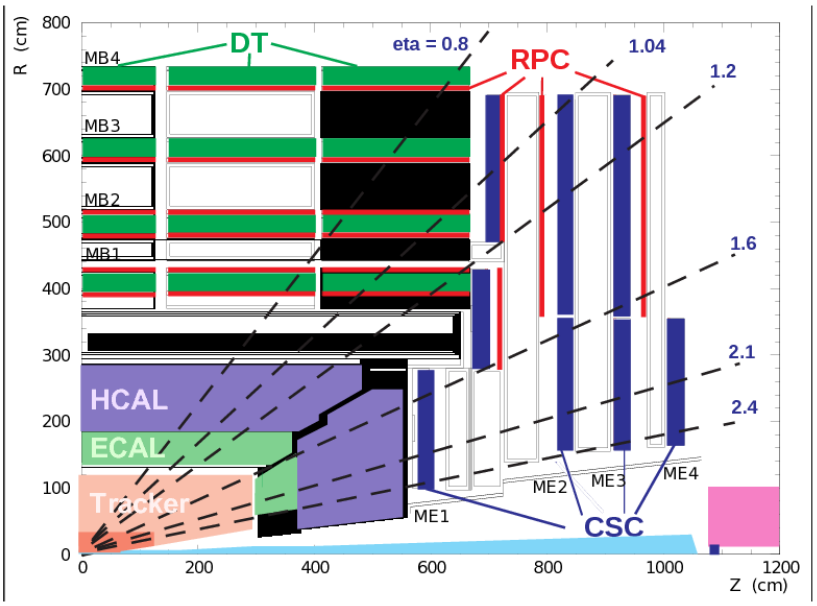


Figure 3.9: A quarter cross-section view of the CMS muon detector in the $y - z$ plane, showing four DT stations in the barrel, the four CSC stations in the endcap, and the RPC station (5).

3.2.5 Trigger System

The LHC is designed to provide a proton-proton beam crossing every 25 ns, corresponding to a frequency of 40 MHz. Combining the approximately 10^8 channels in its electronic readout with a high collision frequency, the CMS detector can generate several Terabyte of data each second. This large amount of data must be reduced for storage and processing with current computing resources. This can be accomplished while preserving the acceptance for interesting and new physics phenomena by selecting events according to well designed criteria. The trigger system is designed to select events and reduce the data acquisition rate. The system includes two stages: Level-1 (L1) Trigger and High-Level Trigger (HLT) (38). The L1 trigger aims to reduce the input rate of 40 MHz to an output rate of 100 kHz, and is largely implemented using

custom programmable electronics. The HLT utilizes more complex software algorithms to further reduce the rate to around 300 Hz.

The L1 Trigger consists of local, regional and global components for both calorimeter and muon systems. In this thesis, we used data passing the L1_SingleEG15 trigger selection, which is based on the calorimeter trigger system (44). At the front-end of the system, the Local Trigger, also called Trigger Primitive Generator (TPG), generates signals of energy deposits in the calorimeter trigger towers. The sum of the transverse energies measured in ECAL crystals or HCAL read-out towers is first transferred to the Regional Calorimeter Trigger (RCT). Then the RCT identifies photon and electron candidates, transverse energy sums, and muon-related information. Afterwards, the Global Calorimeter Trigger receives all information from the RCT and further identifies jets, the total transverse energy, the missing transverse energy, jet multiplicity and the scalar transverse energy sum of all jets (H_T) above a pre-defined threshold. It also provides the highest-p_T-ranked isolated and non-isolated photon and electron candidates. These results are used to make accept (reject) decisions at the rate of the beam crossings.

After an event is accepted by the L1 trigger, the whole event is read out and transmitted to a computing farm, where the HLT software algorithms are applied. There are different algorithms designed to satisfy various physics priorities. In order to study the photon+jet final state, the HLT for single photons is used in this thesis. There are three steps applied in the selection of electrons and photons (44). First, the calorimeter information is used to form energy clusters. Second, by identifying if energy deposits in the ECAL can be matched with hits in the pixel detector, samples will be separated into categories of electron and photon candidates above various thresholds in transverse energy. Finally, the full track reconstruction, seeded from the pixel detector hits, is used to select electrons.

Due to the limited trigger bandwidth, the single photon samples are prescaled when they pass the corresponding HLT. Especially for low transverse momentum photons, where prescale values can be above 1000. This means that the size of the data sample will be reduced by 1000 times limiting our ability to perform studies of special low-$p_T$subsamples within the photon+jet events, for example photons accompanied by heavy flavor jets.

3.3 Computing System

Events passing trigger system selection are stored for offline analysis. Given the large amount of data produced by the CMS experiment, there are several requirements for the offline computing system, including hardware and software.

- A large scale system is required to efficiently perform data reduction and pattern recognition.
- It must be highly flexible to support a variety of data analysis tasks.
- It must operate stably and facilitate uniform maintenance and upgrade tasks.

The CMS software provides an application framework to satisfy a wide range of physics analyses. The architecture consists of a common framework, physics modules, and interface between the two (38). The event is the central component of the CMS data model. It may contain raw digitized data, reconstructed and calibrated detector-level data, or high-level analysis objects. When events pass through the application framework, physics modules can work on them sequentially to apply reconstruction, selection or analysis algorithms. Each physics module can be insulated and independent from others. The framework configures modules parameters, schedules module sequencing and provides access to global services and utilities.

The CMS computing model uses a hierarchical architecture of tiered centers to take advantage of computing resources from collaborating institutes around the world (45). A single Tier-0 center, hosted at CERN, stores RAW data from the online system, produces first-pass reconstruction (RECO) data, and transfers RAW and RECO data to Tier-1 centers. Seven large Tier-1 centers are hosted at collaborating national labs and computing centers. These store RAW (RECO) data from the Tier-0 along with simulated data, perform second-pass reconstruction, and carry out data skimming and analysis. About 45 Tier-2 centers are hosted at CMS institutes and perform analysis activities, simulate data events using Monte Carlo methods, and support specialized activities including detector and algorithm studies. Outside the management scope of LHC projects, there are also many Tier-3 centers hosted by remaining CMS institutes. These facilitate assessment of data in Tier-1s and Tier-2s and enable local data analysis. The high energy group at the University of Virginia functions as one of these local

Tier-3s. All CMS computing centers are integrated into a coherent system based on grid computing middleware, defined by the Worldwide LHC Computing Grid (WLCG) project. The Grid enables user to submit remote jobs to a world-wide network of computing and data storage resources through a standardized interface.

Chapter 4

EVENT RECONSTRUCTION

The final state of the event studied in this analysis includes a photon object and at least one jet object within the required kinematic region. Effectively reconstructing these objects and accurately measuring the corresponding kinematic properties, such as energy and position, are critical to the result of the analysis. The following sections will describe details relating to the reconstruction of the observable objects for photons and jets.

4.1 Photon Reconstruction

Photon showers deposit most of their energy in crystals in the ECAL, where approximately 97% of the incident energy of a single unconverted photon is contained in a 5×5 array of crystals. Due to the material in the tracking system in front of the ECAL, a photon may convert into an e^+e^- pair before reaching the calorimeter. In the strong magnetic field, conversion electrons separate due to Lorentz forces and incident EM particles spread their energy in the ϕ direction. In order to correlate energy deposits in the crystals in the ECAL to incident EM particles, a clustering algorithm is implemented to measure the photon energy. This is referred to as super-clustering (38). This idea is applied to photon reconstruction in both the barrel and endcap regions. There are some differences in the super-clustering algorithms for photons in the barrel and endcap regions to account for differences in detector structures.

The first step of super-clustering is to build basic clusters. The list of reconstructed hits (RecHits) in each ECAL crystal is sorted in order of decreasing transverse energy.

A threshold of 0.5 GeV is applied to each hit to be considered as a seed of a basic cluster. In the ECAL barrel, dynamic clustering is used. Beginning from the seed of a basic cluster, the algorithm first traverses RecHits along the ϕ direction and adds the energy of each RecHit until a rise in energy is encountered or there are no more continuous RecHits. Then the algorithm repeats this first step starting from the nearby RecHits along the η direction. In the ECAL endcap, island clustering is applied. The basic cluster is constructed as a fixed 5×5 array centered on the seed crystal. The area of the basic cluster is projected to the two layers of the preshower in front of the ECAL endcap and the corresponding energy deposited in the preshower detectors also is added to the basic cluster.

The second step of super-clustering is to cluster the basic clusters. The algorithm starts from the basic cluster with highest energy and defines a window of 5 crystal-widths in $\eta \times 17$ crystal-widths in ϕ. Any basic clusters located within the window are jointed together to form a super-cluster.

The energy deposited within a super-cluster represents most of the energy deposited by an electron or photon in the ECAL. An energy correction is applied to compensate for remaining unclustered energy due to the following sources. First, the fraction of energy collected by a super-cluster depends on the shower position with respect to the cluster boundary, which is also described as the "local containment". Second, there is considerable amount of energy loss due to rear leakage for showers close to the barrel inter-module and inter-supermodule regions, where there are cracks aligned at 3° with respect to a line from the beam collision center. Third, photon conversion due to the tracking system spreads energy over a larger area compared with unconverted photons. The fraction of clustered energy systematically varies as a function of number of crystals in the cluster due to zero suppression effects. Fourth, the fraction of clustered energy varies as function of η in the barrel region as a result of changing detector material thicknesses with incident angle. Corrections to these observations are determined on the basis of Monte Carlo simulation.

In addition to energy measurement, photon position is another important measurement. The position of a super-cluster is calculated by the weighted mean of crystal

position using the logarithm of the crystal energy:

$$x = \frac{\Sigma x_i \times W_i}{\Sigma W_i}, \tag{4.1}$$

where x_i is the position of crystal i and $W_i = W_0 + log\frac{E_i}{\sum\limits_j E_j}$. The weight W_i is constrained to be positive. Otherwise, it is set to zero. W_0 is optimized to have the value of 4.2 (46).

4.2 Jet Reconstruction

Unlike a photon, which is a well-defined physical object, a jet is defined by the choice of algorithm that is used to describe the signature of an outgoing quark or gluon. Because of QCD color confinement, particles with a color charge cannot exist in isolation. So quarks or gluons carrying a color charge combine with quarks and anti-quarks spontaneously created from the vacuum and form colorless hadrons, a process called hadronization. The produced hadrons are collected by particle detectors. A jet can be reconstructed by the combination of hadrons using various algorithms. Since hadrons deposit most of their energy in the calorimeter detectors (ECAL and HCAL), the traditional jet reconstruction algorithm mainly uses information from calorimeter detectors and the jet energy resolution depends on the energy resolution of the corresponding detectors. In CMS, the HCAL energy resolution varies from $\frac{90\%}{\sqrt{E}}$ to $\frac{120\%}{\sqrt{E}}$ for different detector modules. The energy resolution uncertainty from the HCAL dominates for the reconstructed jets. In order to improve energy resolution, the concept of particle flow is introduced, which takes advantage of information from tracking systems. Particle flow (PF) jets are used to reconstruct jets based on PF candidates (47).

The reconstruction of PF candidates is performed with a combination of the information from each CMS sub-detector. The input information includes charged-particle tracks, calorimeter clusters, and/or muon tracks. The tracker has a momentum resolution vastly superior to that of the calorimeters and provides a precise measurement of the charged-particle direction from the primary vertex. So it is used to measure the momentum of charged hadrons. An iterative-tracking strategy (48) is implemented to reconstruct charged-particle tracks with high efficiency and low fake rate. Calorimeter

clustering is implemented to reconstruct energy deposits in the ECAL and HCAL, separating neutral particles from charged hadrons. The link algorithm is used to create blocks of fundamental objects, including linking tracks to calorimeter clusters, clusters in different calorimeter layers, and control tracks to muon tracks. Muons, electrons, neutral hadrons, and charged hadrons are identified based on the block structure.

The particle-flow candidates are further clustered into jets by using the anti-k_T algorithm with radius parameter $R = 0.5$ (49). The anti-k_T algorithm is an idealized cone algorithm and has the properties of being infrared and collinear safe. In this algorithm, the clustering process loops through all possible paris of PF candidates and combines the pair with minimum distance of

$$d_{ij} = \frac{\delta_{ij}^2}{R^2 max(k_{Ti}^2, k_{Tj}^2)}, \tag{4.2}$$

where $\delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ and k_{Ti}, η_i and ϕ_i are respectively the transverse momentum, rapidity and azimuth of particle i. This pairwise-clustered PF candidate becomes a new candidate for the next round of clustering. This process is repeated until $d_{ij} > \frac{1}{k_{Ti}^2}$. This algorithm allows soft (low transverse momentum) particles preferentially to cluster with hard (high transverse momentum) ones. As a result of this assignment procedure, when the two showers from two hard particles overlap with each other, $R < \delta_{ij} < 2R$, the one with higher transverse momentum will be conical in shape and the other will be partly conical.

After jets are reconstructed from the above steps, their energy measurement must be corrected for further usage. A bias of jet energy measurement comes from various resources, including pile-up and underlying event, electromagnetic energy fraction, detector geometry and cracks, effects of clustering algorithms and so on (17). Jet energy calibration is performed to relate the energy measurement of the detector jet to the energy of the corresponding particle jet on average using the in-situ measurement of the absolute (vs p_T) jet energy scale. The correction is implemented as a multiplier $C(p_T^{raw}, \eta)$ on the raw transverse momentum p_T^{raw}:

$$p_T^{cor} = C \times p_T^{raw}, \tag{4.3}$$

where p_T^{raw} is the raw jet transverse momentum, C is the correction factor, and p_T^{cor} is the corrected jet transverse momentum. The correction factor C is further divided

into four categories:

$$C = C_{offset}(p_T^{raw}) \times C_{MC}(p'_T, \eta) \times C_{rel}(\eta) \times C_{abs}(p''_T), \tag{4.4}$$

$$p'_T = C_{offset} \times p_T^{raw}, \tag{4.5}$$

$$p''_T = C_{rel} \times C_{MC} \times p'_T, \tag{4.6}$$

where C_{offset} is the correction factor for extra energy due to noise, pileup, and the underlying event; C_{MC} is the MC calibration factor to remove non-uniformities in η and the non-linearity in p_T; C_{rel} and C_{abs} account for the residual difference between data and simulation.

Chapter 5

EVENT SELECTION

5.1 Dataset

This analysis uses around 2.14 fb^{-1} of integrated luminosity recorded with the CMS detector at the LHC during 2011. The four data-sets used for this thesis are summarized in Table 5.1 along with the run periods and recorded luminosity. This thesis only uses the runs certified as "Good" by the CMS data validation team. The data sets used are summarized in Table 5.2.

Table 5.1: Dataset names, run ranges and recorded luminosity of the data used in this thesis.

Dataset Name	Run Range	Recorded Luminosity (pb^{-1})
/Photon/Run2011A-May10ReReco-v1/AOD	160431 - 163869	216.204
/Photon/Run2011A-PromptReco-v4/AOD	165088 - 167913	934.059
/Photon/Run2011A-05Aug2011-v1/AOD	170722 - 172619	373.349
/Photon/Run2011A-PromptReco-v6/AOD	172620 - 173692	667.180

Table 5.2: Dataset names and corresponding official CMS JSON catalog files used.

Dataset Name	JSON file
May10ReReco-v1	Cert_160404-163869_7TeV_May10ReReco_Collisions11_JSON_v3.txt
PromptReco-v4 & v6	Cert_160404-177053_7TeV_PromptReco_Collisions11_JSON.txt
05Aug2011-v1	Cert_170249-172619_7TeV_ReReco5Aug_Collisions11_JSON_v2.txt

The Monte Carlo samples used in this thesis are listed in Tables 5.4∼5.6. The MC samples are generated primarily using the leading-order event generator PYTHIA version 6.4 (50) for a center-of-mass energy of 7 TeV and using CTEQ6L (51) as PDF. The hard scattering sub-processes simulated are listed in Table 5.3. The underlying event (UE) tune Z2 is used in all samples (31). A MC-generated Z signal is also used as a control sample for photons and is listed in Table 5.6.

Table 5.3: Pythia processes for signal and background samples.

Signal Processes	Background Processes
$q_i\bar{q}_i \to g\gamma$	$q_iq_j \to q_iq_j$
$f_i\bar{f}_i \to \gamma\gamma$	$q_i\bar{q}_i \to q_k\bar{q}_k$
$q_ig \to q_i\gamma$	$q_i\bar{q}_i \to gg$
$gg \to \gamma\gamma$	$q_ig \to q_ig$
$gg \to g\gamma$	$gg \to q_k\bar{q}_k$

Table 5.4: Monte Carlo datasets of PYTHIA γ+jet samples used in this thesis.

MC dataset name	$p_{\hat{T}}$ range
PYTHIA γ+jet Z2 tune	
/G_Pt-0to15_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	0 - 15
/G_Pt-15to30_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	15 - 30
/G_Pt-30to50_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	30 - 50
/G_Pt-50to80_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	50 - 80
/G_Pt-80to120_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v2/AODSIM	80 - 120
/G_Pt-120to170_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	120 - 170
/G_Pt-170to300_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v2/AODSIM	170 - 300
/G_Pt-300to470_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	300 - 470
/G_Pt-470to800_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	470 - 800
/G_Pt-800to1400_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	800 - 1400
/G_Pt-1400to1800_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	1400 - 1800
/G_Pt-1800_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	1800 - ∞

Table 5.5: Monte Carlo datasets of PYTHIA QCD dijet samples used in this thesis.

MC dataset name	$p_{\hat{T}}$ range
PYTHIA QCD dijet Z2 tune	
/QCD_Pt-5to15_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	5 - 15
/QCD_Pt-15to30_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	15 - 30
/QCD_Pt-30to50_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	30 - 50
/QCD_Pt-50to80_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	50 - 80
/QCD_Pt-80to120_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	80 - 120
/QCD_Pt-120to170_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	120 - 170
/QCD_Pt-170to300_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	170 - 300
/QCD_Pt-300to470_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	300 - 470
/QCD_Pt-470to600_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	470 - 600
/QCD_Pt-600to800_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	600 - 800
/QCD_Pt-800to1000_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	800 - 1000
/QCD_Pt-1000to1400_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	1000 - 1400
/QCD_Pt-1400to1800_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	1400 - 1800
/QCD_Pt-1800_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	1800 - ∞

Table 5.6: MC datasets of PYTHIA Z→ee used in this thesis.

MC dataset name	$p_{\hat{T}}$ range
PYTHIA Z→ee Z2 tune	
/ZJetToEE_Pt-0to15_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	0 - 15
/ZJetToEE_Pt-15to20_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	15 - 20
/ZJetToEE_Pt-20to30_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	20 - 30
/ZJetToEE_Pt-30to50_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	30 - 50
/ZJetToEE_Pt-50to80_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	50 - 80
/ZJetToEE_Pt-80to120_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	80 - 120
/ZJetToEE_Pt-120to170_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	120 - 170
/ZJetToEE_Pt-170to230_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	170 - 230
/ZJetToEE_Pt-230to300_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	230 - 300
/ZJetToEE_Pt-300_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	300 - ∞

5.2 Trigger Selection

Events used in this thesis were selected by the two-level trigger system of the CMS detector. Both "Level-1" (L1) and the "High-level'" triggers are based on energy deposited in the ECAL. The clustering algorithms used in the L1 and HLT triggers were described in chapter 3. A threshold of 20 GeV is required for 5×5 crystal energy deposits at L1.

Five different HLT paths, corresponding to different p_T thresholds for super-clusters, were used to maximize the sample size within each p_T range. They are given in Table 5.7 together with their run-ranges, recorded and effective luminosities, number of events passing offline selection and p_T range for which they are used. For each p_T range only one HLT path is used. The choice of p_T ranges is guided by the requirement to have maximally efficient triggers, as explained in Section 6.2.

The effective luminosity, which is used to normalize the cross-section, is smaller than the recorded luminosity due to the pre-scales applied on the respective trigger paths that vary as a function of time.

For offline selection we require an event to have a high-quality primary vertex within $\pm$ 24 cm of the nominal interaction point along the proton beam axis (z-direction).

5.3 Photon Identification

In order to select a higher purity photon sample, a number of variables are calculated to improve of photon identification. The variables used for photon identification are listed below:

- $\frac{H}{E}$ is the ratio of energy deposited in the HCAL in a cone with radius size of 0.15 ($\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.15$) to the energy deposited in the ECAL super-cluster. Compared with jet objects, a photon candidate tends to deposit most of its energy in the ECAL system. The value of $\frac{H}{E}$ for photon is typically below 0.05.

- $\sigma_{i\eta i\eta}$ is a shower shape variable describing the width of the energy deposit in the ECAL system along the η direction. The $\sigma_{i\eta i\eta}$ is defined as:

$$\sigma^2_{i\eta i\eta} = \frac{\sum_i^{5\times5} (\eta_i - \bar{\eta})^2 w_i}{\sum_i^{5\times5} w_i}, \; w_i = \max\left(0, 4.7 + \log(E_i/E_{5\times5})\right), \tag{5.1}$$

Table 5.7: HLT trigger paths used in the thesis. In the trigger names, given in the first column, the numbers 30, 50, 75, 90 and 135 represent super-cluster p_T thresholds in GeV. The second column gives the run range during which the trigger was active. Recorded luminosities, and effective luminosities in parentheses, are in the third column. Number of events passing offline selection requirement is given in the fourth column. The last column gives p_T ranges for which the triggers are used.

HLT trigger path	Run range	Rec.(Eff.) (pb^{-1})	Events	p_T range (GeV)
HLT_Photon30_CaloIdVL_v1-7	160431-173692	2190.8(3.53)	1208139	40-60
HLT_Photon30_CaloIdVL_v1-3	160431-165087	2190.8(11.67)	1231920	60-85
HLT_Photon50_CaloIdVL_v1-4	165088-173692			
HLT_Photon75_CaloIdVL_v1-7	160431-173692	2190.8(276.93)	1756132	85-100
HLT_Photon75_CaloIdVL_v1-3	160431-165087	2190.8(361.80)	1844488	100-145
HLT_Photon90_CaloIdVL_v1-4	165088-173692			
HLT_Photon75_CaloIdVL_v1-3	160431-165087	2190.8(2190.6)	1220065	145-300
HLT_Photon125_v1-2	165088-166967			
HLT_Photon135_v1-2	167039-173692			

where E_i and η_i are the energy and η of the i^{th} crystal within the 5×5 electromagnetic cluster, $E_{5\times5}$ is the energy sum of the 5×5 cystals, and $\bar{\eta}$ is the energy weighted average η of 5×5 crystals. For jets, this tends to have a broader shower shape compared with photon candidates.

- Iso_{ECAL} is the sum of the transverse energy from crystals in an annulus with $0.06 < \Delta R < 0.4$ around the photon momentum direction. The included crystals are required to have energy larger than 80 MeV. The contribution within a rectangular strip $\Delta\eta \times \Delta\phi = 0.06 \times 0.4$ is excluded to prevent counting energy associated with photon conversions.

- Iso_{HCAL} is the sum of the transverse energy deposited in the HCAL in an annulus with $0.15 < \Delta R < 0.4$. The inner cone size of 0.15 corresponds to the region used to measure $\frac{H}{E}$.

- Iso_{TRK} is the sum of the transverse energy of collected tracks in an annulus with $0.04 < \Delta R < 0.4$. The tracks are required to be compatible with originating from

the primary vertex. The contribution within a rectangular strip of $\Delta\eta \times \Delta\phi = 0.015 \times 0.4$ is excluded as in the case of Iso_{ECAL}.

- Iso is sum of the isolation variables described above defined as

$$Iso = Iso_{ECAL} + Iso_{HCAL} + Iso_{TRK} \tag{5.2}$$

The Iso variable combines the distinguishing power from Iso_{ECAL}, Iso_{HCAL} and Iso_{TRK}. Photon candidates typically have Iso less than 10 GeV, while jets tend to have a larger Iso value due to hadronization and fragmentation in association with any EM clusters.

- The pixel veto is a boolean variable. It is true when there is no hit in the first two inner layers of the silicon pixel tracking detector that is consistent with tracks matching the location and energy of the photon candidate in the ECAL. This variable is used to exclude electron candidates.

The photon identification criteria used in this thesis are listed in Table 6.9

Table 5.8: Photon identification criteria.

Variable	Selection
pixel seed	require none
$\sigma_{i\eta i\eta}$(Barrel)	< 0.01
$\sigma_{i\eta i\eta}$(Endcap)	< 0.028
H/E	< 0.05

5.4 Jet Identification

Jets of cone size $\Delta R = 0.5$ are reconstructed using the anti-k_t clustering algorithm (52) employing CMS particle-flow objects as the input. Jets with a $p_T^{jet} > 30$ GeV are considered and are required to pass selection criteria (53) to remove noise. The following selection variables, also known as *Jet Id*, are used for the selection (54):

- Neutral Hadron Fraction <0.99
- Neutral EM Fraction <0.99

- Number of Constituents >1
- And for $-2.4 < \eta < 2.4$ in addition apply
  - Charged Hadron Fraction >0
  - Charged Multiplicity >0
  - Charged EM Fraction <0.99

Because particle-flow reconstruction involves summing of ECAL and HCAL towers, photons are also reconstructed as jets. Cleaning of the jet collection from these fake jets is performed by requiring that the jet does not overlap with the leading photon within a radius of $\Delta R < 0.5$. After jet cleaning, only those jets which satisfy the *Jet Id* requirements are considered.

Chapter 6

ANALYSIS

6.1 Definition of Cross Section Measurement

In this analysis, the triple differential cross section for photon+jet production is defined in the following form:

$$\frac{d^3\sigma}{dE_T^\gamma d\eta^\gamma d\eta^{jet}} = \frac{1}{\Delta E_T^\gamma \cdot \Delta\eta^\gamma \cdot \Delta\eta^{jet}} \frac{N^\gamma \cdot U}{L \cdot \epsilon \cdot p^\gamma}, \tag{6.1}$$

where ΔE_T^γ is the bin width for the prompt photon transverse energy[1], and $\Delta\eta^\gamma$ ($\Delta\eta^{jet}$) is the bin width for the pseudorapidity of the prompt photon (leading jet), N^γ is the number of photons measured, p^γ is the photon purity, U is the unfolding factor used to correct photon reconstruction effects due to detector smearing, L is the corresponding integrated luminosity, and ϵ is the total measurement efficiency. This measurement spans the transverse momentum range of $40 < E_T^\gamma < 300\,\text{GeV}$ and $p_T^{jet} > 30\,\text{GeV}$ for photons and jets, respectively, and is performed in four regions of pseudorapidity for the photon ($|\eta^\gamma| < 0.9$, $0.9 \leq |\eta^\gamma| < 1.4442$, $1.56 \leq |\eta^\gamma| < 2.1$ and $2.1 \leq |\eta^\gamma| < 2.5$) and two regions of pseudorapidity for the leading transverse momentum jet ($|\eta^{jet}| < 1.5$ and $1.5 \leq |\eta^{jet}| < 2.5$). The integrated luminosity and its corresponding uncertainty are measured in Ref. (55). Measurements of remaining components used in determining the cross section are discussed in the following sections.

[1]For massless particles $E_T \equiv p_T$. By convention, E_T is often used to designate calorimeter-based measurement.

6.2 Photon Detection Efficiency

The photon detection efficiency measurement can be decomposed into four factors summarized in the following equation:

$$\epsilon_{total} = \epsilon_{trigger} \cdot \epsilon_{RECO} \cdot \epsilon_{ID} \cdot \epsilon_{PMV}. \tag{6.2}$$

Each efficiency component is defined as follows:

- $\epsilon_{trigger}$ represents the probability for a reconstructed signal photon to be selected by the trigger system,
- ϵ_{RECO} stands for the probability for a signal photon produced inside the detector geometrical acceptance to be reconstructed by the clustering algorithms,
- ϵ_{ID} is the probability for a reconstructed signal photon to pass the photon identification criteria,
- ϵ_{PMV} is the probability for a reconstructed signal photon not to match any reconstructed track in the pixel detector that is consistent with the primary vertex.

6.2.1 Photon Trigger Efficiency

There are three categories of single photon HLT paths applied at the HLT level to acquire data in 2011. Isolated triggers, `HLT_Photon*_CaloIdVL_IsoL`[1], employ both photon isolation and shower shape selections to define the most strict selection criteria among the three types of triggers. Non-isolated triggers, `HLT_Photon*_CaloIdVL`, only employ a shower shape selection. Lastly, for photons with high transverse momentum, the triggers, `HLT_Photon*`, employ neither isolation nor shower shape selections. This is acceptable where there are relatively low fake photon rates. In order to increase the statistics of photon candidates to facilitate use of the template fitting method of extracting signal purity (Section 6.3), the second and third types of single photon triggers are used. With increasing instantaneous luminosity of the LHC, some triggers are pre-scaled during data acquisition to satisfy the online trigger system bandwidth limit.

[1] "*" represents the online p_T threshold

The HLT efficiency "turn-on curves" are derived versus photon p_T, in order to find the plateau defining where the HLT path becomes maximally efficient. The trigger efficiency is estimated by the Tag and Probe technique in $Z \rightarrow e^-e^+$ events (57). The tag is a well-reconstructed electron identified using tight electron selection criteria, while the second electron is used to test (probe) the probability of satisfying the photon selection criteria. Using this technique, the turn-on curves of `HLT_Photon75_CaloIdVL`, `HLT_Photon125`, and `HLT_Photon135` are demonstrated in Figs 6.1–6.3. From these results, it is observed that the single photon HLT paths reach maximal efficiency about 10 GeV above their online p_T thresholds. The corresponding maximal efficiencies in both the Barrel ($|\eta| < 1.4442$) and Endcap ($1.566 < |\eta| < 2.5$) regions are shown in Tables 6.1–6.3. In general, non-isolated trigger paths with a shower shape requirement (`*HLT_Photon*_CaloIdVL`) are observed to be fully efficient in both Barrel and Endcap regions. The `HLT_Photon125` trigger is observed to be fully efficient in our fiducial acceptance, whereas `HLT_Photon135` is found to be fully efficient only in the Barrel. The inefficiency of `HLT_Photon135` in the Endcap is consistent with statistical uncertainty and is counted as a systematic error. Non-isolated HLT paths are almost fully efficient in the fiducial region of this analysis, assuring that their systematic uncertainties are negligible.

To examine the effect of pile-up, the single photon HLT efficiency measurement versus number of reconstructed vertices is studied. For example, in Fig. 6.4 the dependence of the HLT efficiency on N_{PV} is shown for the non-isolated HLT path. It is observed that HLT efficiencies have negligible dependence on pile-up effects in 2011A data.

Table 6.1: Efficiency of single photon HLT path for HLT_Photon75_CaloIdVL as a function of photon p_T in the Barrel ($|\eta| < 1.4442$) and Endcap ($1.566 < |\eta| < 2.5$) regions.

Method	Photon p_T (GeV)	Barrel	Endcap
`Fitting method`	85 – Inf.	$99.8 \pm 0.2\%$	$100 \pm 1.5\%$
`Counting method`	85 – Inf.	$99.8 \pm 0.3\%$	$100 \pm 1.5\%$

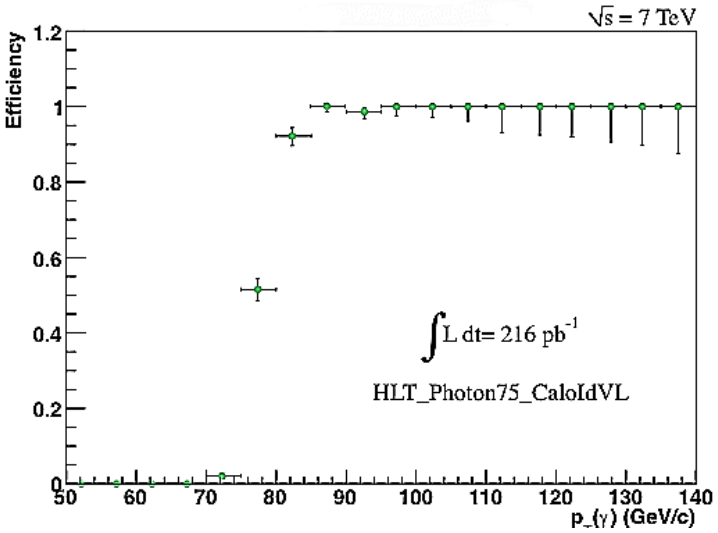


Figure 6.1: Dependence of single photon HLT efficiency for HLT_Photon75_CaloIdVL on reconstructed photon p_T (turn-on curve).

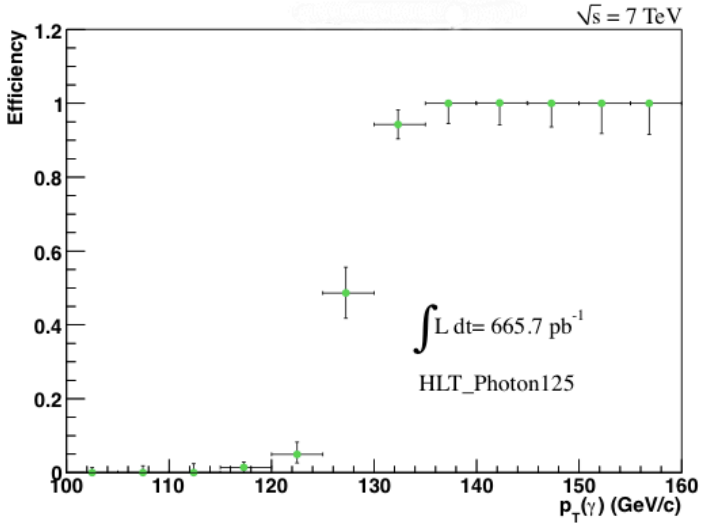


Figure 6.2: Dependence of single photon HLT efficiency for HLT_Photon125 on reconstructed photon p_T (turn-on curve).

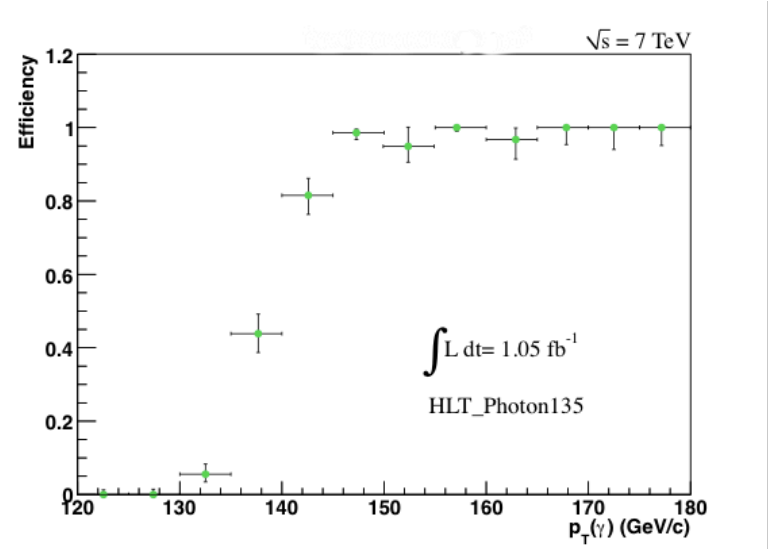


Figure 6.3: Dependence of single photon HLT efficiency for HLT_Photon135 on reconstructed photon p_T (turn-on curve).

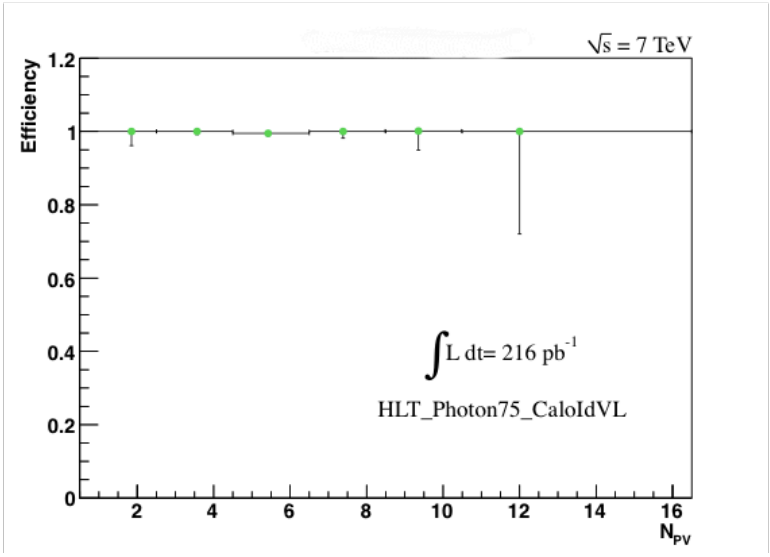


Figure 6.4: Dependence of single photon HLT efficiency for HLT_Photon75_CaloIdVL on number of primary vertices N_{PV}.

Table 6.2: Efficiency of single photon HLT path for HLT_Photon125 as a function of photon p_T in the Barrel ($|\eta| < 1.4442$) and Endcap ($1.566 < |\eta| < 2.5$) regions.

Method	Photon p_{T} (GeV)	Barrel	Endcap
`Fitting method`	135 - Inf.	$100 \pm 0.4\%$	$100 \pm 4.2\%$
`Counting method`	135 - Inf.	$100 \pm 0.4\%$	$100 \pm 4.0\%$

Table 6.3: Efficiency of single photon HLT path for HLT_Photon135 as a function of photon p_T in Barrel ($|\eta| < 1.4442$) and Endcap ($1.566 < |\eta| < 2.5$) regions.

Method	Photon p_{T} (GeV)	Barrel	Endcap
`Fitting method`	145 - Inf.	$100 \pm 0.6\%$	$93.8 \pm 4.7\%$
`Counting method`	145 - Inf.	$99.8 \pm 0.3\%$	$93.2 \pm 4.8\%$

6.2.2 Photon Reconstruction Efficiency

The photon reconstruction algorithm described in Section 4.1 identifies prompt photon objects with high efficiency. To precisely measure the corresponding efficiency, MC samples of PYTHIA photon+jet events listed in Table 5.4 are studied to determine the ratio of reconstructed photons that match photons at generator level to all photons at the generator level. Photons at generator level are required to be isolated by satisfying $Iso < 5$ GeV to match the photon shower shape criteria for reconstructed photons as determined by simulation studies. Photon reconstruction efficiencies are displayed in Fig. 6.5 as a function of photon p_T in different photon η regions. Numerical values of photon reconstruction efficiencies are also summarized in Table 6.4. The results show that the efficiencies are in general above 98%.

Table 6.4: Efficiency of photon reconstruction in bins of generated photon p_T and η.

Photon p_T	$\lvert\eta\rvert < 0.9$	$0.9 < \lvert\eta\rvert < 1.4442$	$1.566 < \lvert\eta\rvert < 2.1$	$2.1 < \lvert\eta\rvert < 2.5$
40–60	$98.78 \pm 0.02\%$	$98.20 \pm 0.04\%$	$98.71 \pm 0.03\%$	$98.12 \pm 0.05\%$
60–85	$98.90 \pm 0.02\%$	$98.36 \pm 0.03\%$	$98.84 \pm 0.03\%$	$98.32 \pm 0.04\%$
85–100	$98.93 \pm 0.02\%$	$98.46 \pm 0.03\%$	$98.76 \pm 0.03\%$	$98.22 \pm 0.04\%$
100–145	$98.98 \pm 0.02\%$	$98.56 \pm 0.03\%$	$98.91 \pm 0.02\%$	$98.46 \pm 0.04\%$
145–300	$99.00 \pm 0.01\%$	$98.67 \pm 0.02\%$	$98.81 \pm 0.02\%$	$98.43 \pm 0.03\%$

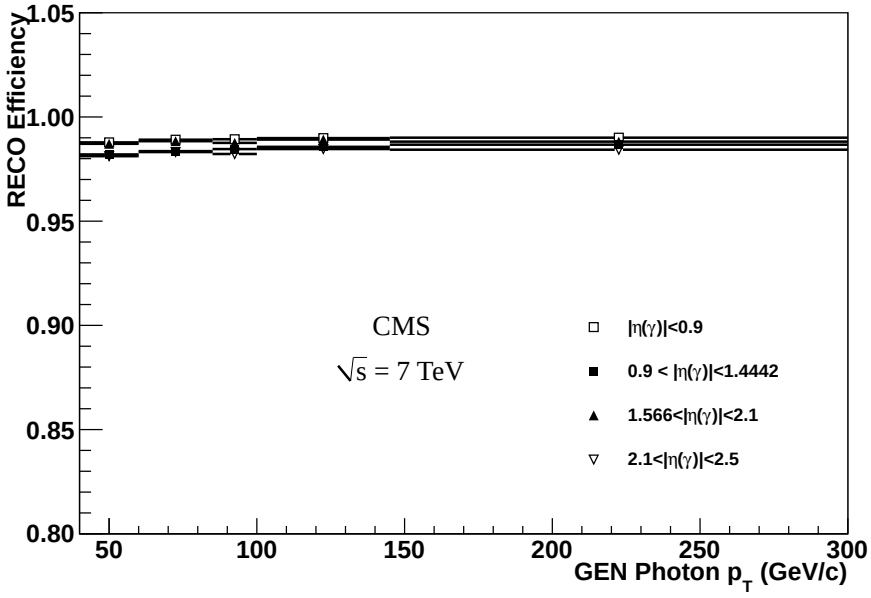


Figure 6.5: Monte Carlo photon reconstruction efficiencies as a function of generated photon p_T in four photon η regions.

6.2.3 Photon Identification Efficiency

The following photon selection criteria are applied to remove background events and improve the purity of the selected photon candidates.

- H/E < 0.05
- Shower shape ($\sigma_{i\eta i\eta}$) < 0.010 for ECAL Barrel ($|\eta| < 1.4442$)
- Shower shape ($\sigma_{i\eta i\eta}$) < 0.028 for ECAL Endcap ($1.566 < |\eta| < 2.5$).

These criteria may also remove some of the real photon objects. The efficiency for photon identification quantifies this effect. Two methods are employed to measure this efficiency. The first method is based on the MC samples of PYTHIA photon+jet events listed in Table 5.4. The efficiency is calculated by the ratio of reconstructed photons passing the above selection criteria divided by all reconstructed photons. The corresponding result is displayed in Fig. 6.6 as a function of reconstructed photon p_T in different photon η regions. The numerical results are listed in Table 6.5. The

second method uses the Tag and Probe technique. This method provides the photon identification efficiencies in Table 6.6. Due to statistical limitations of the dataset, the second method is not able to provide efficiencies for all photon p_T and η bins. The corresponding result is used as a cross-check for the efficiencies calculated from MC samples that are used for the differential cross-section calculation. The difference between the data and MC-driven results is taken as a systematic uncertainty and is listed in Table 6.7. Based on the above result, photon identification efficiencies are in general larger than 93%. The efficiencies are relatively smaller for photons located in the most central η region.

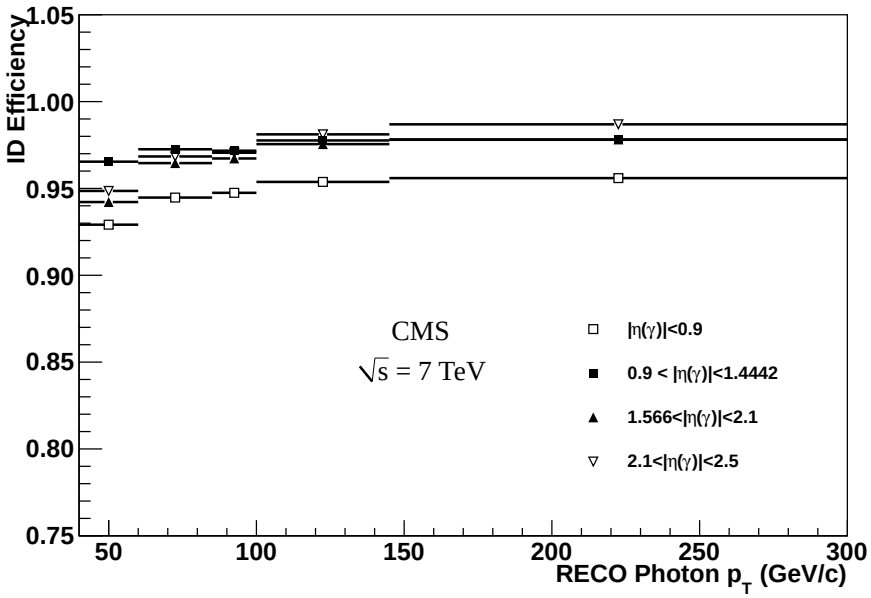


Figure 6.6: Monte Carlo photon ID efficiencies as a function of reconstructed photon p_T in four photon η regions.

6.2.4 Photon Pixel Veto Efficiency

A pixel veto is required in the photon selection criteria to exclude electron candidates passing other photon identification criteria. The corresponding efficiencies for different photon η regions are measured by $Z \to \mu\mu\gamma$ events. This measurement is performed in a dedicated analysis (58). Both data and MC samples of $Z \to \mu\mu\gamma$ events are used

Table 6.5: Efficiency of the isolation template photon ID selecton in bins of reconstructed photon p_T and η from Monte Carlo.

Photon p_T	$\lvert\eta\rvert < 0.9$	$0.9 < \lvert\eta\rvert < 1.4442$	$1.566 < \lvert\eta\rvert < 2.1$	$2.1 < \lvert\eta\rvert < 2.5$
40-60	$92.93 \pm 0.06\%$	$96.54 \pm 0.05\%$	$94.22 \pm 0.07\%$	$94.86 \pm 0.07\%$
60-85	$94.48 \pm 0.05\%$	$97.25 \pm 0.04\%$	$96.46 \pm 0.05\%$	$96.84 \pm 0.05\%$
85-100	$94.75 \pm 0.05\%$	$97.17 \pm 0.04\%$	$96.73 \pm 0.05\%$	$97.06 \pm 0.05\%$
100-145	$95.37 \pm 0.04\%$	$97.76 \pm 0.03\%$	$97.55 \pm 0.04\%$	$98.11 \pm 0.04\%$
145-300	$95.60 \pm 0.03\%$	$97.80 \pm 0.02\%$	$97.82 \pm 0.03\%$	$98.70 \pm 0.03\%$

Table 6.6: Efficiency of the isolation template photon ID selection in bins of reconstructed photon p_T and η from Tag and Probe data.

Photon p_{T} (GeV)	Barrel	Endcap
40-45	$95.1 \pm 1.3\%$	$92.2 \pm 3.2\%$
45-50	$94.4 \pm 1.2\%$	$97.1 \pm 2.5\%$
50-Inf.	$92.5 \pm 1.6\%$	$100 \pm 2.3\%$

Table 6.7: Systematic uncertainty of isolation template photon ID selection efficiency in bins of reconstructed photon p_T in Barrel ($\lvert\eta\rvert < 1.4442$) and Endcap ($1.566 < \lvert\eta\rvert < 2.5$) regions.

Photon p_{T} (GeV)	Barrel	Endcap
40-45	2.5%	3.0%
45-50	1.2%	3.5%
50-Inf.	4.5%	5.0%

to calculate the photon pixel veto efficiency. The results are listed in Table 6.8. The ratio between results from data and MC are consistent with each other. The efficiency decreases with higher $|\eta_\gamma|$. The data results are used in the differential cross-section measurement.

Table 6.8: Photon pixel veto efficiency for different photon η regions.

$\|\eta_\gamma\| region$	ϵ_{PMV}^{data}	ϵ_{PMV}^{MC}	$\frac{\epsilon_{PMV}^{data}}{\epsilon_{PMV}^{MC}}$
[0,0.9)	$97.12^{+0.37}_{-0.38}\%$	$97.80^{+0.08}_{-0.09}\%$	$0.9931^{+0.0038}_{-0.0039}\%$
[0.9,0.14442]	$95.77^{+0.51}_{-0.55}\%$	$96.57^{+0.15}_{-0.15}\%$	$0.9918^{+0.0053}_{-0.0057}\%$
[1.556,2.1)	$91.30^{+0.91}_{-0.98}\%$	$90.93^{+0.32}_{-0.34}\%$	$1.0040^{+0.0097}_{-0.0103}\%$
[2.1,2.5]	$77.47^{+2.13}_{-2.25}\%$	$79.18^{+0.73}_{-0.75}\%$	$0.9784^{+0.0225}_{-0.0237}\%$

6.3 Purity

In a high-energy proton-proton collision, prompt photons (signal photons) are produced via processes such as quark-gluon Compton-like scattering, quark-antiquark annihilation, and to lesser extent initial and final state radiation. Photons that are produced in the decay of neutral hadrons inside of jets may fake prompt photons and constitute background objects, which contaminate signal samples. The photon selection criteria in Table 6.9 are applied to diminish background events. However, a considerable contamination from the neutral hadrons still remains even after the selection procedure. In the data sample, the events passing the photon selection criteria, called photon candidates, are a mixture of both the signal and background events. Taking advantage of the shape differences between signal and background, the template fitting method is used to determine the component of signal events in the photon candidates.

6.3.1 Isolation Template Description

The isolation variable (Iso) is used as the template to distinguish signal and background events. It describes the energy deposited in the detectors around the photon shower and is defined as described in Section 5.3. The underlying event and pileup contribute to the isolation energy of the signal photons. In contrast, neutral hadrons that fake

prompt photons usually are embedded in jets. Hence they tend to have more energy around them in addition to the contributions from the underlying event and pileup.

6.3.2 Template Fitting Procedure

The signal template is modeled using photon+jets MC events generated with PYTHIA listed in Table 5.4 and parameterized by the convolution of a lifetime function with a Gaussian,

$$Iso_S^{\gamma}(\vec{p},\alpha) = \exp(\alpha x) \otimes \text{Gaussian}(x,\mu,\sigma), \tag{6.3}$$

where $\vec{p} = (\mu, \sigma)$ describes the peak of the signal template and α describes the tail of the signal template.

The background template is obtained from data using a background enriched sample collected from a sideband region, defined by inverting the shower shape selection requirement and requiring $\sigma_{i\eta i\eta} > 0.011(0.030)$ in the barrel (endcap) regions (Table 6.10). The background distribution is parameterized using an inverse ARGUS function (59),

$$Iso_B^{\gamma}(z,\vec{q}) = \begin{cases} \left[1 - e^{z(x-q_1)}\right] \cdot [1 - q_2(x-q_1)]^{q_3} & ;\ x \geq q_1 \\ 0 & ;\ x < q_1, \end{cases} \tag{6.4}$$

where z describes the shape of the background template in the signal-dominated region and $\vec{q} = (q_1, q_2, q_3)$, where q_1 describes the starting point of the background template, and q_2 and q_3 describe its shape in the background-dominated region.

The signal purity is determined by fitting the signal and background template functional forms to data, $N_S * Iso_S^{\gamma} + N_B * Iso_B^{\gamma}$, and minimizing an extended χ^2 defined as:

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{N_i - (N_S S_i(\vec{p},\alpha) + N_B B_i(z,\vec{q}))}{\sigma_{N_i}} \right)^2 + \left(\frac{(z - z_{\text{central}})}{\sigma_z} \right)^2, \tag{6.5}$$

where N_S and N_B are the number of signal and background events, n is the number of bins in the templates, N_i the observed number of events for the i-th Iso^{γ} bin with uncertainty σ_{N_i}, S_i and B_i are the per bin normalizations of the corresponding signal

and background templates. The parameters can be categorized by those that most directly model the signal- ($\vec{p}$ and z) and background-dominated (α and $\vec{q}$) regions. The parameter that describes the peak in the signal template is allowed to vary in the fit to correct for differences between data and MC in the region of low isolation energy. This technique was verified using photons from $z \to \mu\mu\gamma$ events in Section 6.3.4. In the low Iso^{γ} region, the background distribution is constrained by the distribution of sideband data, allowing the parameter z to vary based on the value z_{central} and constrained by the prior uncertainty σ_z. An example of the resulting templates is shown in Fig. 6.7.

Table 6.9: Photon identification criteria.

Variable	Selection
pixel seed	require none
$\sigma_{i\eta i\eta}$(Barrel)	< 0.01
$\sigma_{i\eta i\eta}$(Endcap)	< 0.028
H/E	< 0.05

Table 6.10: Photon sideband selection.

Variable	Selection
pixel seed	require none
$\sigma_{i\eta i\eta}$(Barrel)	> 0.011 && < 0.015
$\sigma_{i\eta i\eta}$(Endcap)	> 0.035 && > 0.040
H/E	< 0.05

The template fitting results for different photon η orientations are shown in Figs 6.8–6.11.

The numerical results of the purity determination for each kinematic bin are shown in Table 6.11, and the results of the corresponding statistical error in Table 6.12.

6.3.3 Pileup Reweighting for Monte Carlo sample

As mentioned above, photon+jets MC samples are used to simulate the signal template. These Monte Carlo samples have been generated including a number of minimum bias events distributed with a uniform plus Poisson tail distribution to model the number of pileup interactions. However, it is evident from distributions of the number of pileup

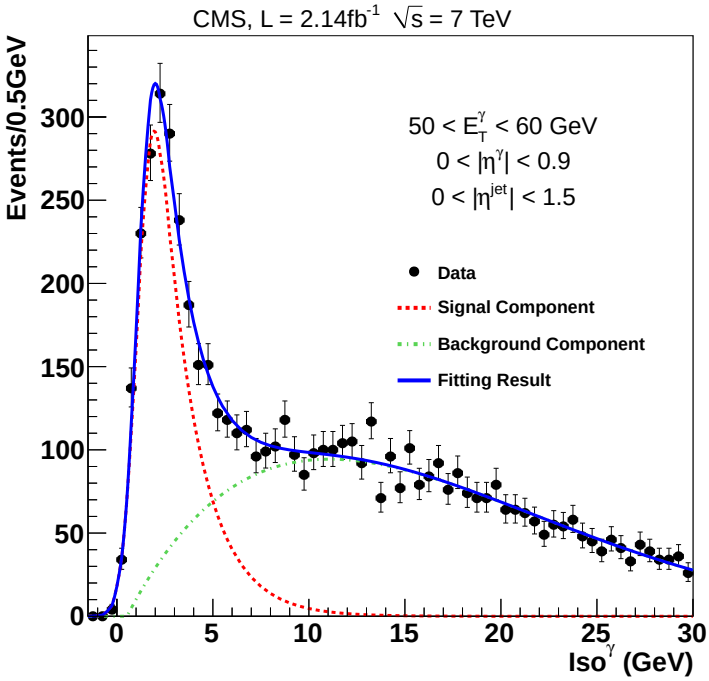


Figure 6.7: Example of fit to Iso^{γ} distribution using signal and background templates.

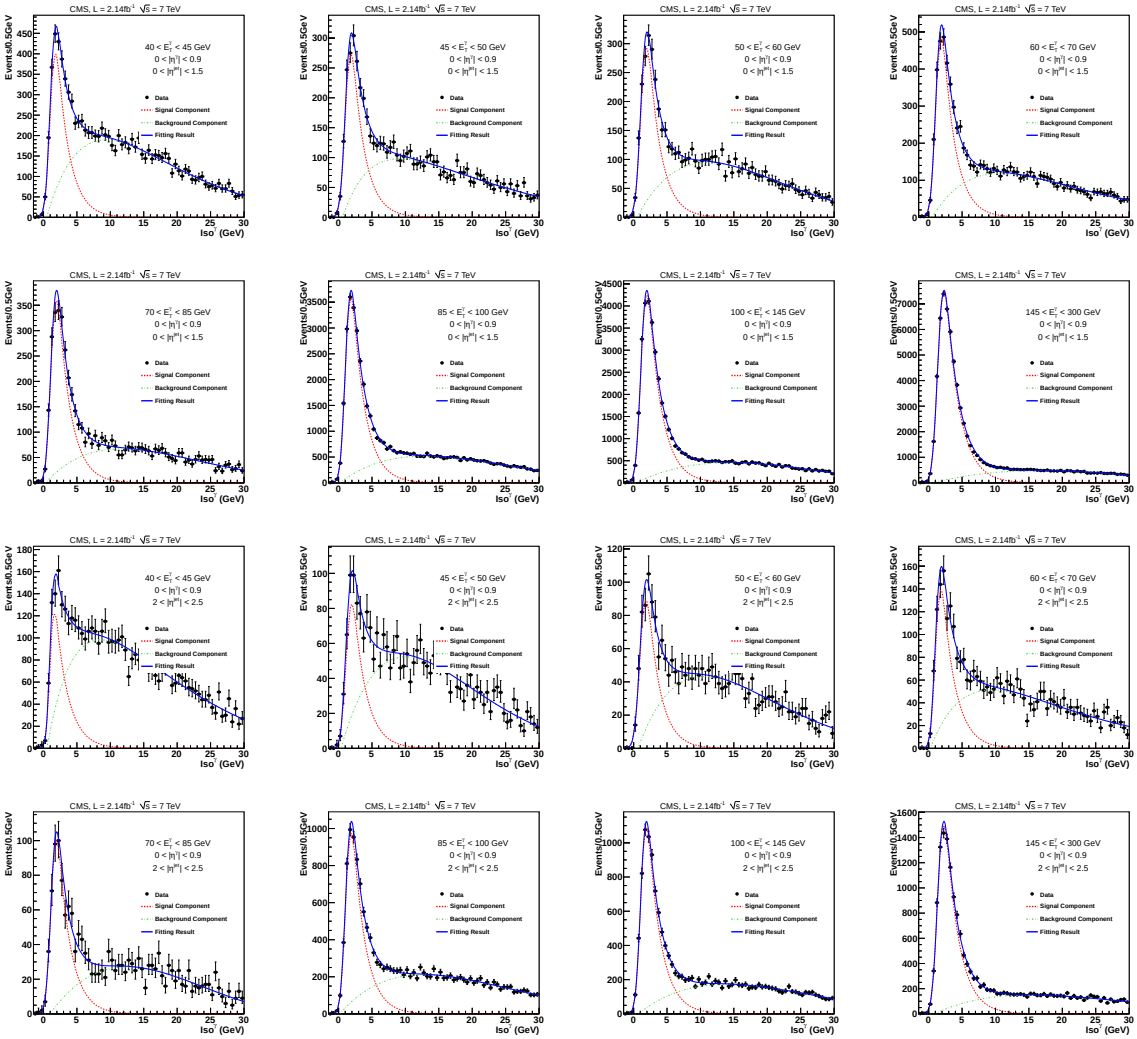


Figure 6.8: Measured isolation distributions (points with error bars) for photons with $0 < |\eta^\gamma| < 0.9$. The extended-χ^2 fit result (blue) is overlaid in each plot with the component for background (green) and the component for signal (red) shown separately.

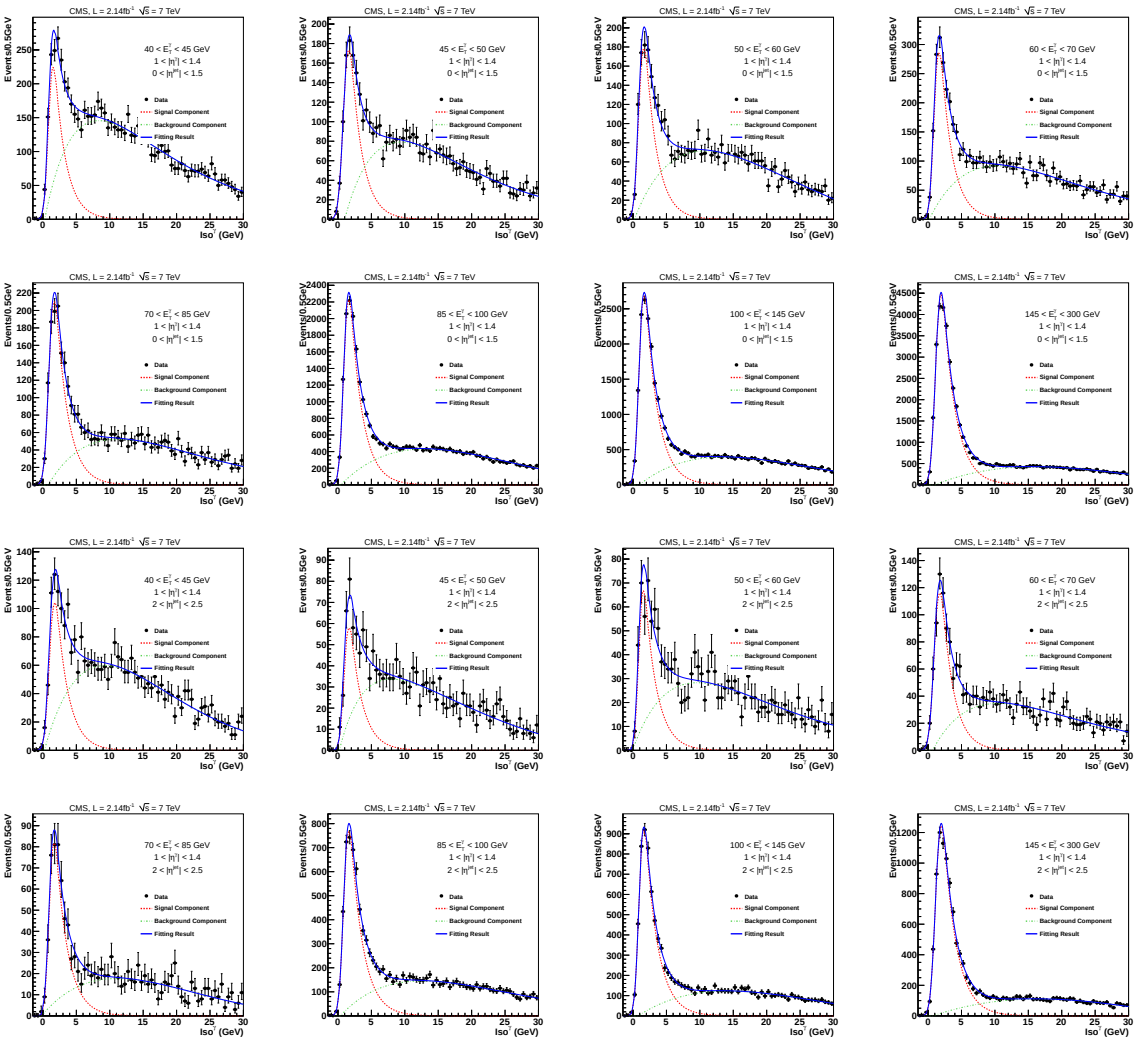


Figure 6.9: Measured isolation distributions (points with error bars) for photons with $0.9 < |\eta^\gamma| < 1.4442$. The extended-$\chi^2$ fit result (blue) is overlaid in each plot with the component for background (green) and the component for signal (red) shown separately.

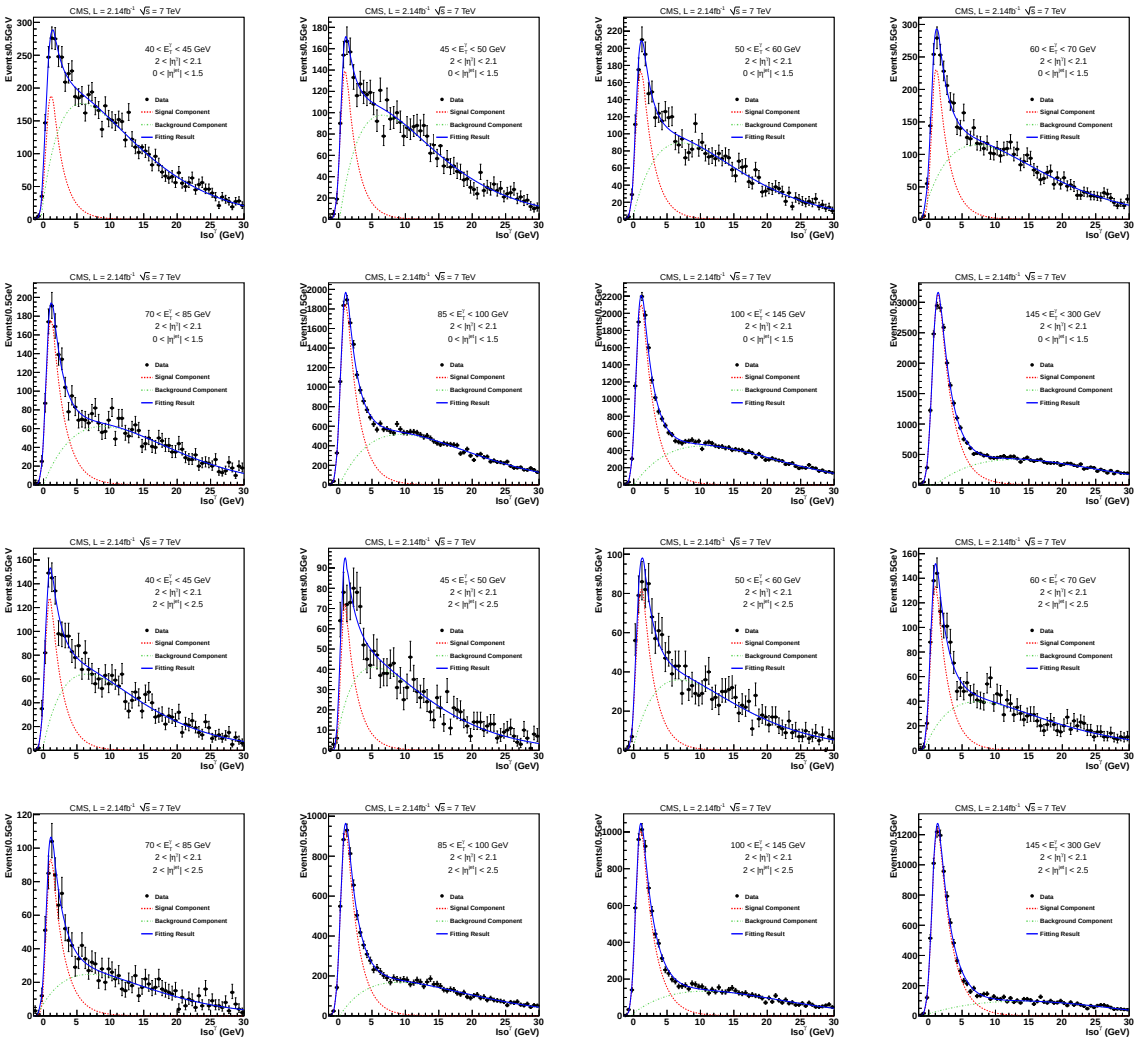


Figure 6.10: Measured isolation distributions (points with error bars) for photons with $1.566 < |\eta^\gamma| < 2.1$. The extended-$\chi^2$ fit result (blue) is overlaid in each plot with the component for background (green) and the component for signal (red) shown separately.

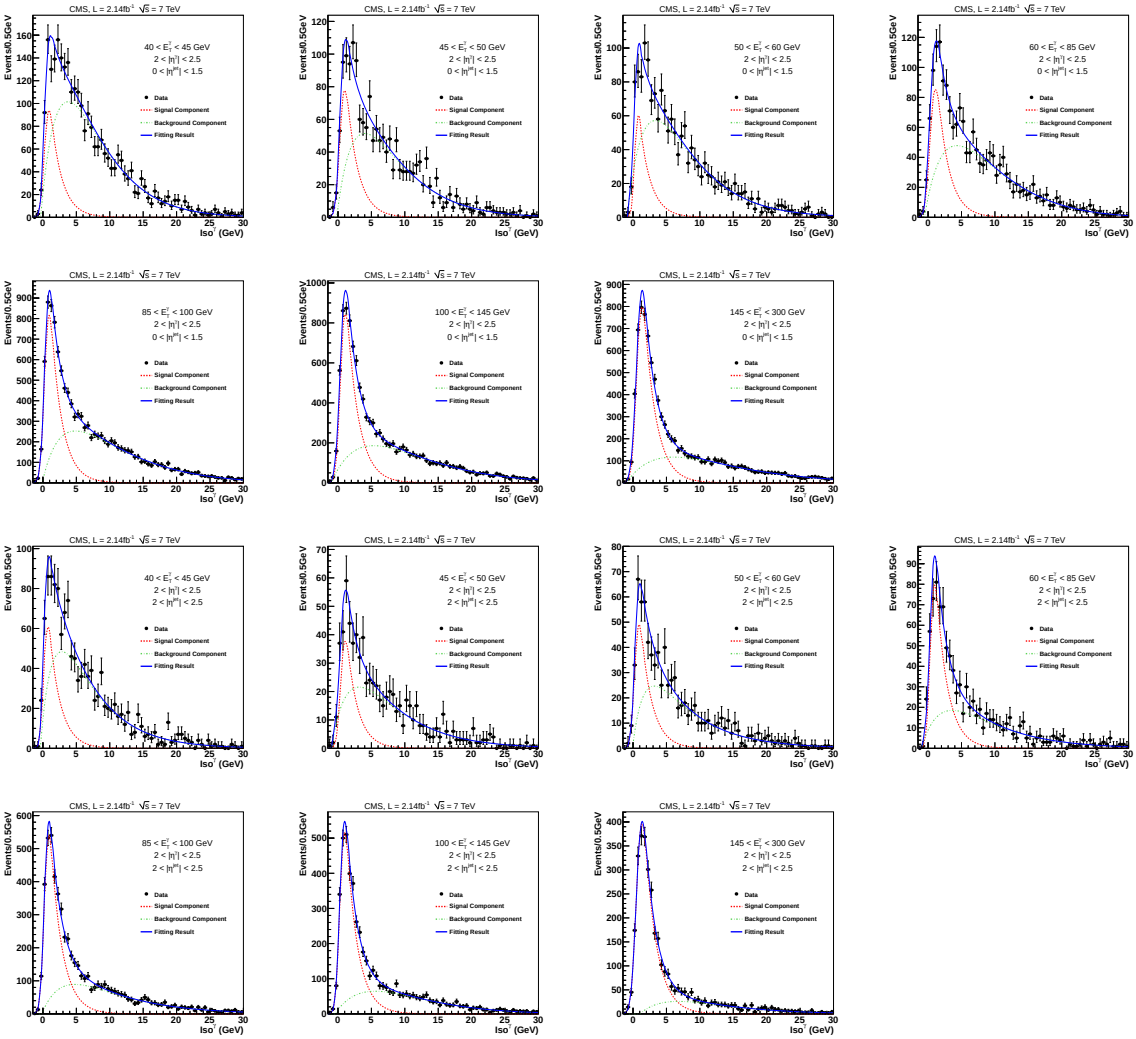


Figure 6.11: Measured isolation distributions (points with error bars) for photons with $2.1 < |\eta^\gamma| < 2.5$. The extended-$\chi^2$ fit result (blue) is overlaid in each plot with the component for background (green) and the component for signal (red) shown separately.

Table 6.11: Purity results from isolation template fits

p_T^γ	$\lvert\eta^\gamma\rvert < 0.9$		$0.9 < \lvert\eta^\gamma\rvert < 1.4442$		$1.566 < \lvert\eta^\gamma\rvert < 2.1$		$2.1 < \lvert\eta^\gamma\rvert < 2.5$	
(GeV)	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$
40–45	25.06%	15.71%	18.10%	21.45%	15.70%	25.64%	18.89%	25.22%
45–50	30.33%	20.82%	25.55%	20.65%	19.27%	22.58%	27.62%	31.13%
50–60	33.18%	25.50%	26.24%	24.16%	24.69%	28.34%	21.47%	33.79%
60–70	38.23%	29.35%	30.03%	33.13%	26.35%	33.36%	31.27%	44.48%
70–85	44.81%	37.12%	36.80%	39.38%	31.02%	40.62%		
85–100	50.20%	39.79%	41.64%	42.16%	35.18%	44.90%	38.58%	54.74%
100–145	58.41%	48.32%	48.54%	50.57%	40.57%	52.18%	46.35%	60.47%
145–300	70.75%	60.41%	60.52%	62.84%	53.83%	64.71%	54.56%	74.22%

Table 6.12: Statistical Uncertainty of Purity result from Isolation template

p_T^γ	$\lvert\eta^\gamma\rvert < 0.9$		$0.9 < \lvert\eta^\gamma\rvert < 1.4442$		$1.566 < \lvert\eta^\gamma\rvert < 2.1$		$2.1 < \lvert\eta^\gamma\rvert < 2.5$	
(GeV)	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$
40–45	3.27%	7.15%	5.23%	6.72%	7.73%	5.56%	18.96%	25.16%
45–50	4.23%	7.49%	4.09%	12.28%	6.54%	10.34%	10.93%	29.36%
50–60	3.58%	6.34%	4.45%	10.30%	5.56%	7.65%	11.92%	14.56%
60–70	2.55%	5.32%	3.58%	5.11%	4.52%	5.05%	7.15%	8.49%
70–85	3.10%	5.03%	3.68%	6.30%	4.47%	5.38%		
85–100	0.84%	1.72%	1.06%	1.94%	1.29%	1.77%	3.34%	2.98%
100–145	0.70%	1.46%	0.93%	1.56%	1.33%	1.60%	3.45%	2.67%
145–300	0.49%	1.19%	0.67%	1.24%	0.82%	1.26%	3.36%	2.84%

interactions from data and MC in Fig. 6.12 that the MC setting does not match the experimental conditions. In order to correct for this difference, we reweight MC events based on the number of pileup interactions from the simulation truth to match the pileup distribution for data, which is derived by using the instantaneous luminosity together with the total proton-proton inelastic cross-section to generate a pileup distribution, correctly weighted by the per-bunch-crossing luminosity distribution over the entire data-taking period. The calculation follows the standard CMSSW recipe (60). The right plot in Fig. 6.12 shows the reweighting factors for the photon+jets MC samples.

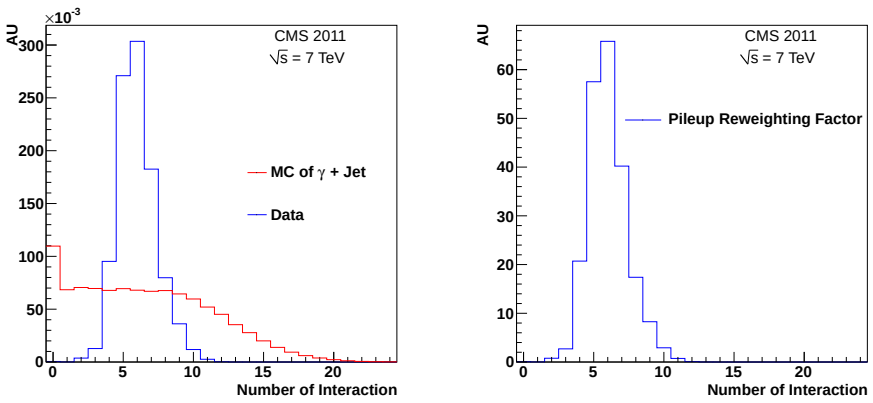


Figure 6.12: Left plot shows the distributions of the number of pileup interactions with blue line refering to the distribution of data and red line refering to the distribution of MC. Right plot shows the reweighing factors versus the number of pileup interactions for MC.

6.3.4 Signal Template Uncertainty

The uncertainty in the photon purity relating to the signal template depends on the difference of template shapes between photon+jets MC samples and the real photon+jets data. Due to the irreducible background contamination, data with pure prompt photons is not available. Alternatively, $Z \to \mu\mu\gamma$ events have clean signatures to identify and very small background contamination. They provide a clean data sample to study the properties of pure photon objects that can further be used to quantify the difference between MC and data. The data and MC samples used for this study are listed in Table 6.13.

Table 6.13: $Z \to \mu\mu\gamma$ MC and data samples list

Samples	Dataset Name
Data	DoubleMuRun2011A-PromptReco-v4
MC	DYToMuMu_M-20_CT10_TuneZ2_7TeV-powheg-pythia

To obtain pure $Z \to \mu\mu\gamma$ events, we apply following selection criteria similar to a dedicated $Z \to \mu\mu\gamma$ study (58).

The muon selection criteria are as follows:

- the muon must be reconstructed as a global muon,
- the global muon track fit should have a χ^2/ndof < 10,
- the global muon must have at least one valid muon chamber hit matched to the global fit,
- the muon must be reconstructed also as a tracker muon
- the tracker muon must match to at least two muon stations,
- the muon inner track must have at least one hit in the silicon pixel detector,
- the muon inner track must have more than 10 hits (pixels + strips),
- $|d_{xy}(\mathbf{o})| < \mathbf{1}$ cm where $d_{xy}(\mathbf{o})$ is the (approximate) inner track impact parameter in the transverse plane calculated with respect to the origin of the CMS coordinate system,
- $|d_{xy}| < 2$ mm where d_{xy} is the (approximate) global track impact parameter in the transverse plane calculated with respect to the beam spot position, and
- $p_T^{\mu} > 10$ GeV where p_T^{μ} is the muon transverse momentum.

The di-muon selection criteria are as follows:

- the two muons must have opposite charge,
- $m_{\mu^+\mu^-} \in [40, 80]$ GeV, where $m_{\mu^+\mu^-}$ is the di-muon system invariant mass.

The photon selecion criteria are as follows:

- the selection criteria listed in Table 6.9 are applied,
- $Iso < 13GeV$
- $p_T^{\gamma} > 20GeV$

The $\mu\mu\gamma$ candidate selection criteria are as follows:

Based on the $\Delta R(\mu^{\pm}, \gamma)$ distance of the muons and the photon in the η-ϕ plane we define the near and far muons μ_{near} and μ_{far}:

$$\Delta R(\mu_{\text{near}}, \gamma) = \min \Delta R(\mu^{\pm}, \gamma),$$
$$\Delta R(\mu_{\text{far}}, \gamma) = \max \Delta R(\mu^{\pm}, \gamma).$$

- $Iso_{HCAL}^{\mu_{\text{near}}} < 1$ GeV where $Iso_{HCAL}^{\mu_{\text{near}}}$, the near muon HCAL isolation, is the sum of the transverse energy of HCAL towers within a cone of $\Delta R < 0.3$ around the near muon direction,
- $Iso_{TRK}^{\mu_{\text{far}}} < 3$ GeV where $Iso_{TRK}^{\mu_{\text{far}}}$, the far muon tracker isolation, is the sum of the transverse momentum of tracks with $p_T > 1.5$ GeV within a cone of $\Delta R < 0.3$ around the far muon direction, vetoing a cone of 0.015 around that direction,
- $Iso_{ECAL}^{\mu_{\text{far}}} < 1$ GeV where $Iso_{ECAL}^{\mu_{\text{far}}}$, the far muon ECAL isolation, is the sum of transverse energy of ECAL rec hits within a cone of $\Delta R < 0.3$ around the far muon direction,
- $Iso_{TRK}^{\text{corr}}\gamma < 2\text{GeV} + 0.001 p_T$ where $Iso_{TRK}^{\text{corr}}\gamma$ is the hollow-cone photon tracker isolation corrected for the presence of the near moun the p_T of which is removed from the sum,
- $\Delta R(\mu_{\text{near}}, \gamma) < 1$ where $\Delta R(\mu_{\text{near}}, \gamma)$ is distance between the photon and the near muon the $\eta - \phi$ plane,
- $p_T(\mu_{\text{far}}) > 15$ GeV where $p_T(\mu_{\text{far}})$ is the far muon transverse momentum,
- $m_{\mu^+\mu^-\gamma} \in [75, 105]$ GeV where $m_{\mu^+\mu^-\gamma}$ is the $\mu^+\mu^-\gamma$ system invariant mass,
- $\min \Delta\eta(\mu^{\pm}, \gamma) \begin{cases} > 0.04 \text{ for photon in the barrel} \\ > 0.08 \text{ for photon in the endcaps} \end{cases}$, or $\min \Delta\phi(\mu^{\pm}, \gamma) > 0.3$.

After applying the above event selection criteria, the Iso variable distributions of barrel photons from $Z \to \mu\mu\gamma$ MC and data samples are shown in Fig. 6.13 and fitted by a convolution of a lifetime function with a Gaussian distribution (Eq. 6.3). The overlay of these two distributions is shown in Fig. 6.13. The fit results from MC and data samples are shown in Table 6.14. The last column in Table 6.14 is calculated by dividing the absolute value of parameter differences in data and MC by the average of that parameter. The percentage difference for parameters ($\vec{p}$) that describe the peak in the template shows are less than 1%. The parameter α that characterizes the right-side tail of template has a difference of about 2.6%. As mentioned above, the parameter α is fixed during the template fitting procedure. To estimate the corresponding uncertainty, we shift the α by $\pm 5\%$ and compare the new purity fitting results with nominal one. Then the difference in purity is quoted as an uncertainty of the signal template in Table 6.15.

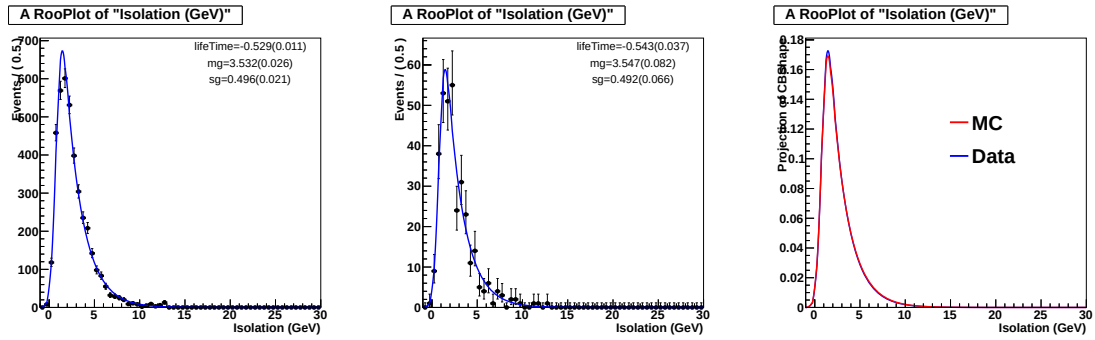


Figure 6.13: The Iso variable distribution (left) for barrel photon candidates in the $Z \to \mu\mu\gamma$ MC samples. The middle figure is the same distribution for $Z \to \mu\mu\gamma$ data samples. The right figure is the comparison between the fitting results of the left two distributions.

Table 6.14: The parameters resulting from fits for the Iso distribution of photon candidates in $Z \to \mu\mu\gamma$ MC and data samples

Parameters	MC	Error (MC)	Data	Error (Data)	Difference
α	-0.529	0.011	-0.543	0.037	2.61%
p_1	3.532	0.026	3.547	0.082	0.42%
p_2	0.496	0.021	0.492	0.066	0.81%

Table 6.15: Uncertainty corresponding to signal template

p_T^γ	$\lvert\eta^\gamma\rvert < 0.9$		$0.9 < \lvert\eta^\gamma\rvert < 1.4442$		$1.566 < \lvert\eta^\gamma\rvert < 2.1$		$2.1 < \lvert\eta^\gamma\rvert < 2.5$	
(GeV)	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$
40–45	2.97%	3.19%	2.88%	4.51%	3.45%	3.24%	3.70%	5.15%
45–50	2.73%	2.80%	1.94%	6.43%	3.77%	6.78%	3.82%	3.69%
50–60	2.01%	2.04%	2.16%	7.56%	2.99%	3.52%	4.91%	4.13%
60–70	2.27%	2.79%	8.80%	2.26%	3.52%	7.06%	4.47%	3.04%
70–85	4.28%	1.95%	2.08%	1.77%	7.89%	2.69%		
85–100	1.94%	2.09%	1.69%	2.08%	2.23%	2.10%	3.84%	3.15%
100–145	1.53%	1.90%	1.64%	1.52%	1.97%	1.79%	4.12%	2.50%
145–300	1.29%	2.21%	1.35%	1.15%	1.59%	1.36%	4.11%	2.23%

6.3.5 Background Template Uncertainty

To model background events in the photon candidates, we use a template constructed from data events passing the sideband selection in Table 6.10. Differences in the shape between this template and the true background template affect the accuracy of the purity measurement. The background events within the photon candidates cannot be accessed directly. Therefore the QCD di-jets MC samples in Table 5.5 are used to study effects due to differences between sideband events and background events. Events are required to be located in the kinematic region: 40 GeV $< |p_T^\gamma| <$ 145 GeV, $0 < |\eta^\gamma| < 1.4442$ and $0 < |\eta^{jet}| < 2.5$. The method is implemented in three steps as follows:

First we apply the signal selection in Table 6.9 to both photon+jets and QCD PYTHIA MC samples to model the signal and background templates. And we apply the sideband selection in Table 6.10 to the QCD MC samples to get a QCD sideband template. We then use Eq. 6.3 to fit the signal template and Eq. 6.4 to fit the background template and sideband template. The three templates and corresponding fits are shown in Fig. 6.14.

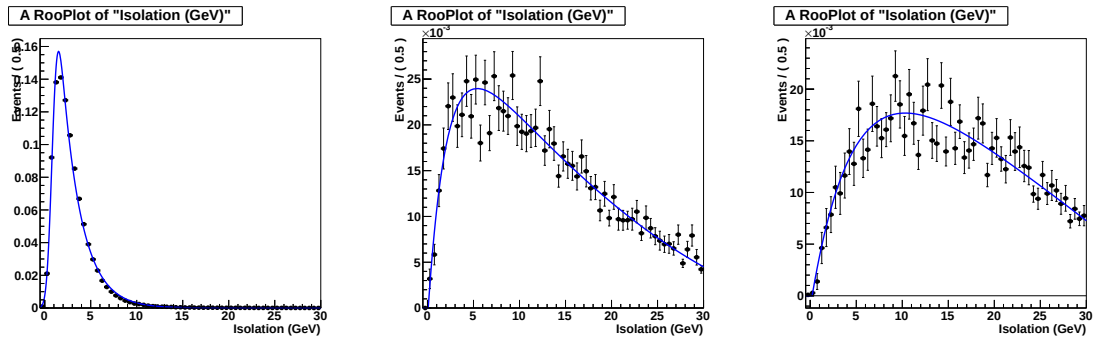


Figure 6.14: The signal template (left) from γ+jet PYTHIA MC samples. The background template (middle) from QCD PYTHIA MC samples. The sideband template (right) from QCD PYTHIA MC samples.

Next we use the functions representing the signal (Iso_{signal_MC}) and background ($Iso_{background\_MC}$) templates to construct pseudodata distributions. For the construction of pseudodata distributions, we also require two more input parameters, total number of events (N_{events}) and signal purity (θ), which can be transformed to signal

($N_{events} \times \theta$) and background ($N_{events} \times (1-\theta)$) event numbers. We can write a smooth template distribution for pseudodata using the following equation:

$$Iso_{pseudodata} = N_{events} \times \theta \times Iso_{signal\_MC} + N_{events} \times (1-\theta) \times Iso_{background\_MC}. \quad (6.6)$$

To account for the effect of statistical uncertainties in the data affecting the purity estimation, we fluctuate each bin content of the histogram of the pseudodata template distribution randomly according to a Poisson distribution. For example beginning with a total event number of 10,000 and signal purity of 0.5, the corresponding template distributions of pseudodata before and after fluctuations are shown in Fig. 6.15.

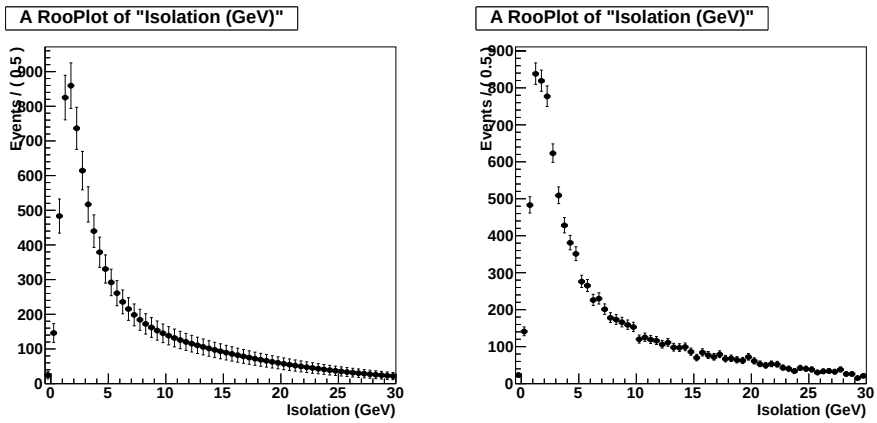


Figure 6.15: The template distribution (left) of pseudodata before statistic fluctuations. The template distribution (right) of pseudodata after fluctuations. The pseudodata distributions shown are generated using a total number of 1000 events and signal purity of 0.5.

Finally we use the signal and sideband template functions to perform the purity fitting following the procedure described in Section 6.3.2. The result from one pseudoexperiment is shown in Fig. 6.16. We define the purity error due to the difference between sideband template and background template in a given trial by comparing the signal purity value from the fit result to the input value:

$$error = \frac{Purity_{\text{default}} - Purity_{\text{pseudoexperiment}}}{Purity_{\text{pseudoexperiment}}}, \quad (6.7)$$

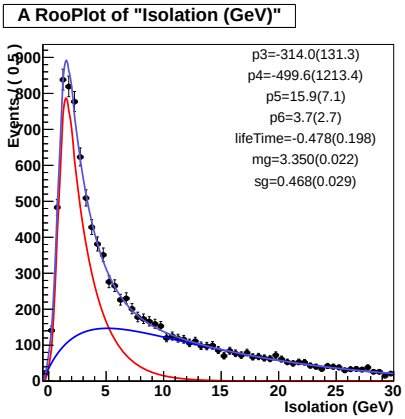


Figure 6.16: Pseudodata fitted by signal template and sideband template. The pseudodatas shown are generated using a total number of 1000 events and signal purity of 0.5.

where $Purity_{\text{default}}$ is the priori purity used to generate the pseudodata.

The above example determines the variation in the signal purity due to the mismatch of sideband template and background template for a single pseudo-experiment. To estimate the statistical uncertainty on the purity due to the variation in the fit, we repeat the process and perform the purity fit 500 times. We generate a distribution of the estimated purity error for the ensemble of pseudo-experiments and fit this distribution using Gaussian function. The mean of Gaussian function is used as the purity correction factor. And the standard deviation is used to estimate the uncertainty due to the limited statistics in the templates. An example of the fitted distribution for all pseudo-experiments is shown in Fig. 6.17. This procedure is repeated for each measurement using the corresponding total event counts and signal purity. The corresponding results are shown in Table 6.16 and 6.17.

Based on above study, we see that the mismatch between the background template and sideband template is likely to generate a bias in the purity determination. We use Table 6.16 as reference to calculate the correction factor for the central purity values. The correction factors are shown in Table 6.18. The statistical uncertainties of the correction factors are estimated from the results corresponding in Table 6.17. The results are shown in Table 6.19.

Another uncertainty relating to the background template is from the parameter q_1,

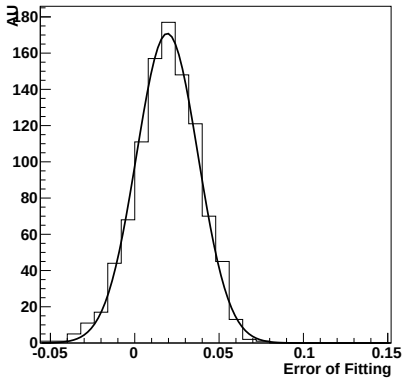


Figure 6.17: The distribution of the variation of the purity determination due to the limited statistics is fitted by Gaussian function.

Table 6.16: Mean value of Gaussian function used to fit the purity error distribution for each measurement using the corresponding total event counts and signal purity

Statistics \ Purity	0.2	0.3	0.4	0.5	0.6	0.7
2000	9.92%	8.84%	6.69%	5.73%	4.46%	3.97%
4000	7.38%	4.49%	4.90%	4.27%	2.80%	2.86%
6000	6.06%	4.06%	2.97%	2.89%	2.63%	2.30%
10000	5.89%	3.83%	2.51%	1.95%	1.43%	1.47%
15000	5.82%	3.58%	2.49%	1.68%	1.22%	0.92%
20000	5.76%	3.40%	2.31%	1.65%	1.16%	0.78%

Table 6.17: σ value of Gaussian function used to fit the purity error distribution for each measurement using the corresponding total event counts and signal purity

Statistics \ Purity	0.2	0.3	0.4	0.5	0.6	0.7
2000	6.57%	4.81%	4.27%	3.18%	2.98%	2.26%
4000	5.56%	4.38%	2.93%	2.29%	2.09%	1.55%
6000	5.19%	3.64%	2.70%	2.25%	1.75%	1.33%
10000	3.76%	2.88%	2.28%	1.76%	1.52%	1.15%
15000	3.15%	2.54%	1.90%	1.47%	1.31%	1.09%
20000	2.79%	2.10%	1.64%	1.31%	1.11%	0.99%

Table 6.18: Correct factor of purity value

p_T^γ	$\lvert\eta^\gamma\rvert < 0.9$		$0.9 < \lvert\eta^\gamma\rvert < 1.4442$		$1.566 < \lvert\eta^\gamma\rvert < 2.1$		$2.1 < \lvert\eta^\gamma\rvert < 2.5$	
(GeV)	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$
40–45	4.21%	9.49%	7.13%	7.17%	8.46%	6.63%	8.57%	11.90%
45–50	4.05%	8.29%	5.42%	11.36%	7.73%	10.79%	9.70%	10.96%
50–60	3.66%	8.13%	5.31%	10.61%	5.67%	9.34%	11.30%	10.49%
60–70	3.66%	6.03%	4.14%	7.82%	5.06%	7.56%	5.90%	7.20%
70–85	3.45%	8.03%	5.20%	8.96%	5.33%	8.27%	2.71%	3.56%
85–100	1.51%	2.42%	2.17%	2.35%	2.80%	2.06%	2.03%	2.71%
100–145	1.08%	1.78%	1.68%	1.82%	2.24%	1.68%	1.70%	2.89%
145–300	0.64%	1.24%	1.00%	1.14%	1.36%	1.00%		

Table 6.19: Uncertainty of background template

p_T^γ	$\lvert\eta^\gamma\rvert < 0.9$		$0.9 < \lvert\eta^\gamma\rvert < 1.4442$		$1.566 < \lvert\eta^\gamma\rvert < 2.1$		$2.1 < \lvert\eta^\gamma\rvert < 2.5$	
(GeV)	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$
40–45	3.48%	6.42%	5.05%	6.53%	5.60%	5.54%	7.14%	6.75%
45–50	3.65%	6.62%	4.44%	6.90%	5.71%	6.62%	5.81%	6.61%
50–60	3.37%	5.74%	4.45%	6.47%	4.54%	5.50%	6.83%	6.00%
60–70	2.60%	5.17%	3.73%	4.36%	4.10%	4.29%	4.73%	3.90%
70–85	2.85%	4.44%	3.58%	4.90%	4.63%	4.31%		
85–100	0.93%	1.88%	1.25%	2.06%	1.45%	1.96%	2.17%	1.85%
100–145	0.76%	1.59%	1.04%	1.70%	1.30%	1.64%	1.93%	1.82%
145–300	0.58%	1.22%	0.74%	1.30%	0.94%	1.23%	1.77%	1.59%

which affects the description of the overlap of the background and signal distributions. As described earlier, we fix the parameter q_1 in Eq. 6.4. To estimate the corresponding uncertainty from q_1, we shift its value by its $\pm$ 1 S.D. from the background template fitting result. The corresponding shifts in the purity values are used to estimate the uncertainty. The results are shown in Table 6.20.

The full systematic uncertainty on the purity calculation comes from the modeling of the signal and background templates and is taken as the quadratic sum of these uncertainties. The systematic uncertainty is evaluated independently for each bin and ranges from $1 \rightarrow 30\%$ with decreasing photon transverse momentum.

Using the template fitting method and uncertainty estimation, the final result of the purity estimation, shown as a function of p_T^γ in Fig. 6.18 for four different photon η orientations, increases with transverse momentum of the photons. The main contribution to the systematic uncertainty on the photon signal purity is due to uncertainty in modeling the shape of the background template.

6.4 Unfolding

After correcting for the detection efficiency and purity, photon yields are calculated for bins in ranges of p_T^γ, η^γ and η^{jet}. Due to the effect of detector resolution, calibration, etc., these measured values may slightly differ from the true values as a result of events produced in one bin migrating to another. The last step of our measurement is to correct for this effect to obtain the differential cross-sections in bins defined by the true values of p_T^γ, η^γ and η^{jet}. This procedure, called unfolding, builds a map of photon yields between true and measured values. Since only MC samples provide both true and measured values, the unfolding is derived based on the photon+jets PYTHIA MC samples listed in Table 5.4.

The unfolding is realized with a dedicated software package RooUnfold (61). An iterative (Bayesian) method (6) provides 3D unfolding and is implemented in our analysis to account for the influence of event migration between different bins. The result is cross-checked by a less precise but robust bin-by-bin correction method, which only takes the ratio of true and measured photon yields in each bin as the correction factor.

The same binning is used for both true and measured variables within the acceptance defined by the experiment. Outside of the acceptance region, additional bins were used:

Table 6.20: Uncertainty corresponding to parameter q_1

p_T^γ	$\lvert\eta^\gamma\rvert < 0.9$		$0.9 < \lvert\eta^\gamma\rvert < 1.4442$		$1.566 < \lvert\eta^\gamma\rvert < 2.1$		$2.1 < \lvert\eta^\gamma\rvert < 2.5$	
(GeV)	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$
40–45	1.00%	3.07%	3.64%	3.50%	2.68%	0.12%	4.12%	0.05%
45–50	0.17%	7.81%	1.58%	11.71%	9.58%	2.72%	3.50%	32.50%
50–60	0.99%	2.20%	2.51%	6.63%	4.05%	6.96%	20.02%	13.44%
60–70	0.75%	3.78%	5.32%	1.55%	3.92%	7.43%	9.70%	8.15%
70–85	2.01%	0.60%	1.42%	0.03%	2.68%	11.12%		
85–100	0.09%	0.28%	0.35%	0.61%	0.60%	0.45%	0.90%	0.05%
100–145	0.26%	0.26%	0.37%	0.78%	1.33%	1.75%	4.16%	1.03%
145–300	0.08%	0.01%	0.03%	0.62%	3.94%	2.19%	2.37%	2.88%

Table 6.21: Total uncertainty of purity result

p_T^γ	$\lvert\eta^\gamma\rvert < 0.9$		$0.9 < \lvert\eta^\gamma\rvert < 1.4442$		$1.566 < \lvert\eta^\gamma\rvert < 2.1$		$2.1 < \lvert\eta^\gamma\rvert < 2.5$	
(GeV)	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$	$\lvert\eta^{jet}\rvert < 1.5$	$1.5 < \lvert\eta^{jet}\rvert < 2.5$
40–45	11.18%	22.22%	17.66%	22.67%	19.70%	19.72%	78.65%	26.65%
45–50	13.73%	24.44%	17.25%	25.86%	22.69%	36.08%	30.82%	48.78%
50–60	12.80%	19.51%	17.09%	24.20%	18.78%	118.66%	33.30%	31.86%
60–70	7.41%	16.91%	14.10%	15.64%	19.43%	16.85%	27.49%	20.55%
70–85	9.52%	14.53%	12.78%	16.72%	15.39%	17.46%	22.86%	18.32%
85–100	4.92%	6.63%	5.84%	7.26%	7.73%	8.31%	8.63%	13.05%
100–145	3.62%	6.10%	4.74%	5.04%	6.16%	6.20%	36.29%	8.04%
145–300	2.53%	3.95%	3.36%	3.58%	5.76%	3.84%	15.43%	6.18%

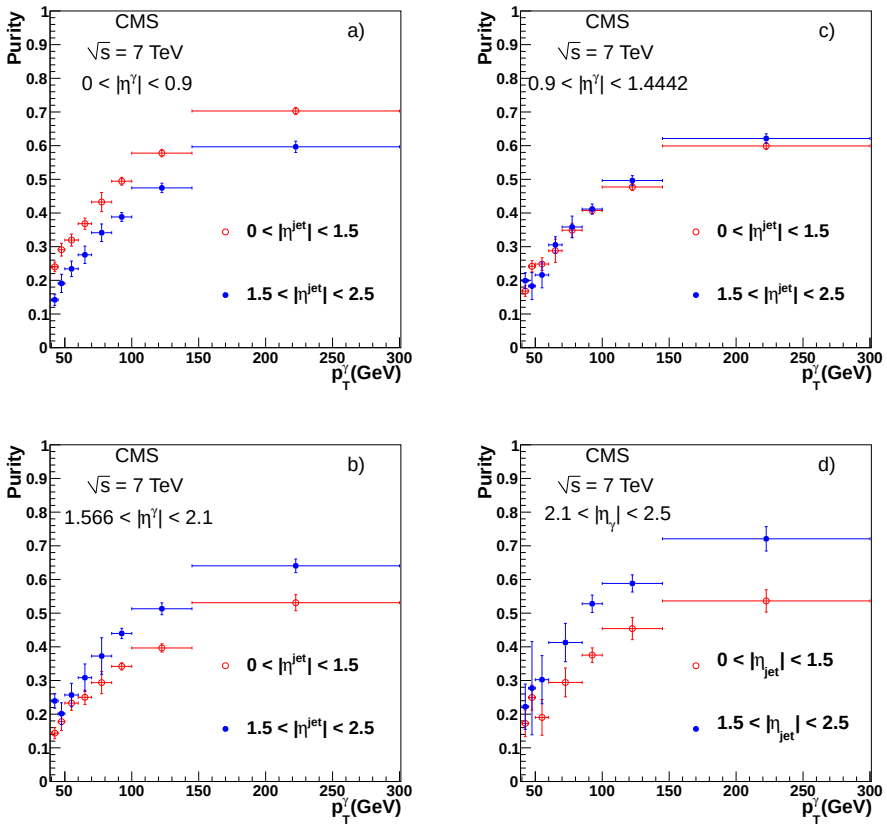


Figure 6.18: Signal purity versus photon transverse momentum corresponding to different photon η orientations. In each figure, the red line refers to events with the jet located in the ECAL barrel and the blue line refers to events with the jet located in the ECAL endcap

- 30-40 GeV and 300-500 GeV for p_T^γ
- 1.4442-1.566 and 2.5-2.6 for $|\eta^\gamma|$
- 2.5-2.7 for $|\eta^{jet}|$

For the unfolding procedure, MC samples are divided into two statistically independent sets of equal size. The first set is used to obtain the unfolding correction factors, while the second set is used for testing. To perform the testing, the measured distribution is unfolded by the correction factors from first set and compared with the true distribution. This closure test shows the precision of the iterative method with limited MC statistics. An example result of the test is shown in Fig. 6.19, where the ratio of corrected-measured and true distributions is given as a function of p_T^γ for the most central photon and jet pseudorapidity bins. The unfolding correction factors are shown in Fig. 6.20.

There are two factors contributing to the uncertainty of the unfolding procedure. The first comes from the limited size of the MC samples resulting an approximate 15% increase in the statistical uncertainty of the cross section after unfolding. The second comes from the differences between MC and data, where the correction factor from the unfolding is applied. Two methods are applied to check the sensitivity of the unfolding procedure due to modeling differences. First, the weights of individual MC samples are artificially modified by factors from 0.5 to 2 as a function of $\hat{p_T}$ (event energy scale) for training, while the weights remain the same for testing. The corresponding result shows that the effect is negligible. The second method to estimate the effect is to calculate the ratio of the measured and MC photon yields for each bin and correct MC training distributions with these ratios. The unfolding result shows that, for most of bins, this will introduce an approximately $1 \rightarrow 2\%$ variation. For some bins the uncertainty is larger, namely those where the ratio factor differed significantly from the neighboring bins. It would be an overestimate to take that change as a systematic uncertainty for a particular bin because part of that change comes from statistical and systematic uncertainties associated with cross-section calculation. Since the uncertainty is found to be mainly dependent on p_T^γ, the relative systematic uncertainty is quoted by using a p_T^γ-dependent average of relative systematic uncertainties for different η^γ and η^{jet} bins. The only exceptions are the bins $p_T^\gamma > 70$ GeV for $|\eta^{jet}| > 1.5$ and $|\eta^\gamma| > 1.566$ which had much higher uncertainty and were calculated separately. The results are shown in

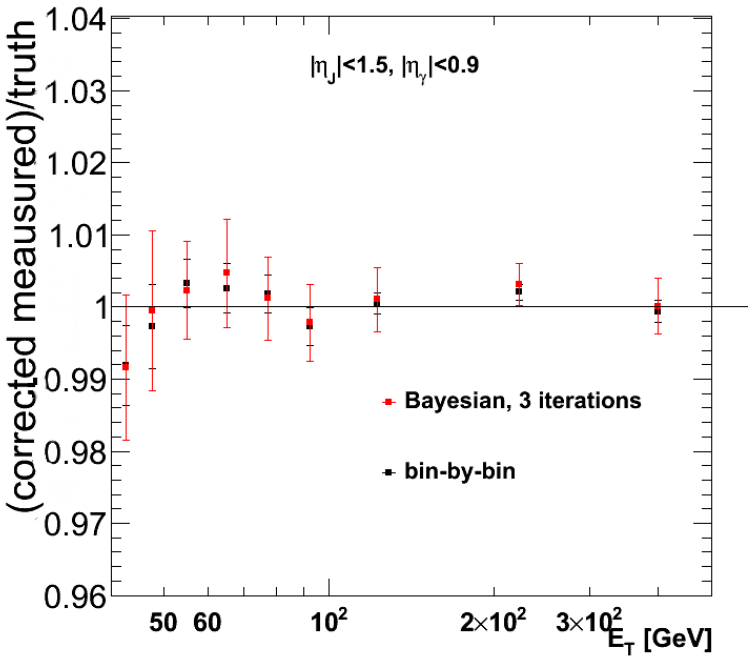


Figure 6.19: Test of unfolding correction. The red squares show the ratio of corrected measurement and truth with statistical uncertainty calculated by an iterative (Bayesian) method (6), while the black squares show the ratio determined by using the bin-by-bin correction method.

Table 6.22. Another difference examined between data and MC is the photon energy resolution. By comparing $Z \to ee$ data and MC, it is found that energy resolution in data is worse than in the simulation (62). The difference also depends on whether the photon is in the barrel or endcap region of the ECAL, what selections are applied and what data sets are being compared. This effect is studied by comparing the default unfolding with one where the MC reconstructed photon energy is additionally smeared in the following way: $(p_T^{reco} smeared) = p_T^{gen} + (p_T^{reco} - p_T^{gen}) \times 1.3$, where factor 1.3 was chosen as the typical ratio between data and MC energy resolutions. The result, in Table 6.22, shows that the difference in energy resolution between data and MC has only a marginal effect on unfolding.

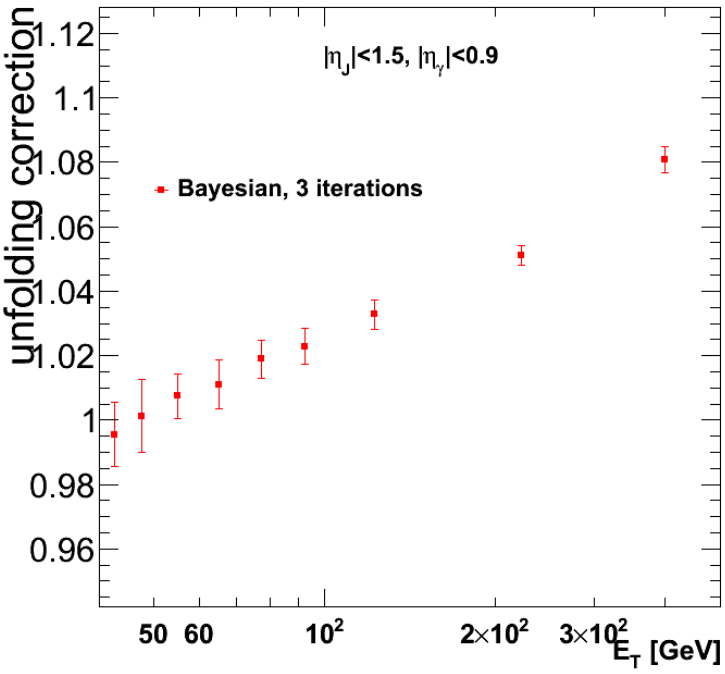


Figure 6.20: The example of unfolding correction factors calculated by an iterative (Bayesian) method (6) for kinematic range $\eta_{jet} < 1.5$ and $\eta_\gamma < 0.9$. Error bars correspond to the statistical uncertainty.

Table 6.22: First two rows: systematical uncertainty on unfolding due to the shape difference between MC and data photon yields. The difference was obtained by correcting the MC training distributions with the ratio of data and MC yields. The numbers are averaged over different η^{γ} and η^{jet} bins. The only exceptions are the bins $p_T^{\gamma} > 70GeV$ for $|\eta^{jet}| > 1.5$ and $|\eta^{\gamma}| > 1.566$ which were calculated separately. Last row: systematical uncertainty on unfolding due to difference in photon energy resolution between MC and data. The numbers are averaged over different η^{γ} and η^{jet} bins.

p_T^{γ} (GeV)	40-45	45-50	50-60	60-70	70-85	85-100	100-145	145-300
$\|\eta^{jet}\| < 1.5$ $\|\eta^{\gamma}\| < 1.4442$	1.8%	2.4%	2.5%	2.2%	1.0%	1.3%	1.3%	1.0%
$\|\eta^{jet}\| > 1.5$ $\|\eta^{\gamma}\| > 1.566$	1.8%	2.4%	2.5%	2.2%	5.0%	5.0%	4.0%	2.0%
Energy resol.	1.0%	0.8%	0.6%	1.0%	0.7%	0.6%	0.4%	0.6%

6.5 Systematic Uncertainties

The contributions to the systematic uncertainty of the differential cross section include those from the efficiency measurement, purity measurement, unfolding factor, and integrated luminosity. The corresponding systematic uncertainties for the efficiency, purity and unfolding factor are listed in Table 6.23. This table also shows the total systematic uncertainty obtained by adding all contributions in quadrature. At low p_T^{γ} the systematic uncertainty is dominated by the purity determination. This is also the region where the uncertainty is the highest. At high p_T^{γ} usually the most significant contribution comes from the determination of the reconstruction efficiency. The systematic uncertainty for the determination of the integrated luminosity is 2.2% (55).

Table 6.23: Contributions to the relative systematic uncertainty (in percent) in the cross section measurement from efficiency, unfolding, and purity calculations. The total systematic uncertainty is obtained by adding all the contributions in quadrature. The numbers in the last column represent the ranges of uncertainties obtained in different η^{γ} and η^{jet} bins. An additional 2.2% luminosity uncertainty is not included in the totals.

$\lvert\eta^{\gamma}\rvert < 1.4442$				
p_T^{γ} (GeV)	efficiency (%)	unfolding (%)	purity (%)	total (%)
40-45	2.5	2.1	4.9 - 9.3	5.9 - 9.9
45-50	1.2	2.5	4.9 - 17	5.5 - 17
50-60	4.5	2.6	4.2 - 13	6.7 - 14
60-70	4.5	2.4	3.7 - 11	6.3 - 13
70-85	4.5	1.2	4.6 - 5.7	6.6 - 7.4
85-100	4.5	1.4	2.2 - 3.1	5.2 - 5.6
100-145	4.5	1.4	1.8 - 2.5	5.0 - 5.4
145-300	4.5	1.2	1.4 - 2.6	4.9 - 5.3
$1.556 < \lvert\eta^{\gamma}\rvert < 2.5$				
40-45	3.0	2.1	6.9 - 9.9	7.8 - 11
45-50	3.5	2.5	8.6 - 38	9.6 - 38
50-60	5.0	2.6	7.2 - 25	9.1 - 25
60-70	5.0	2.4	7.0 - 12	9.0 - 14
70-85	5.0	1.2 - 5.0	10 - 13	11 - 15
85-100	5.0	1.4 - 5.0	2.8 - 4.6	5.9 - 8.0
100-145	5.0	1.4 - 4.0	2.8 - 6.3	5.9 - 8.2
145-300	5.0	1.2 - 2.1	2.9 - 5.1	6.1 - 7.3

Chapter 7

RESULTS AND CONCLUSION

7.1 Results of the triple-differential cross section measurement

In the previous chapter, all the required elements are calculated separately for the measurement of the triple-differential cross section $d^3\sigma/(dp_T^\gamma d\eta^\gamma d\eta^{jet})$. The final results of the measurement of the triple-differential cross section for $|\eta^{jet}| < 1.5$ and $1.5 < |\eta^{jet}| < 2.5$ are shown in Figs. 7.1 and 7.2. The measured cross sections are compared to theoretical predictions based on perturbative QCD using the leading order MC event generator SHERPA (v1.3.1) (34) and the full next-to-leading order calculation of JETPHOX (v1.2.2) (35). The SHERPA generator implements the calculation utilizing higher-order tree-level matrix elements and parton shower modeling (63). For generating events involving prompt photons (64), the SHERPA generator combines the photon and QCD parton tree-level matrix elements with a QCD+QED parton shower using the formalism given in Ref. (63). In order to directly compare with experimental measurements, the fragmentation photon component is also included in the SHERPA generator. The SHERPA generator predictions agree well with the photon+jets measurements from the Tevatron (65). In this analysis, the photon+jets events generated by SHERPA include final states with up to three additional jets and uses the CTEQ6 (66) parton distribution functions. The corresponding parameter settings use the default choices for renormalization (μ_R) and factorization (μ_F) scales equal to p_T^γ. For JETPHOX, the CT10 (67) NLO PDFs are used with $\mu_R = \mu_F = \mu_f = p_T^\gamma/2$, where μ_f defines the fragmentation scale. To model the experimental

selection requirements, the hadronic energy around the photon within the $R < 0.4$ cone is required to be less than 5 GeV. An estimate of the theoretical predictions due to the choice of theory scales is obtained by independently varying μ_R, μ_F, μ_f by factors 0.5 and 2.0. The uncertainty on the predictions due to the choice of PDF is determined from the 40 (52) component error sets of CTEQ6M (CT10) and evaluated using the master equations as given by the 'modified tolerance method' recommended in Ref. (68). Figure 7.3 shows the ratios of the measured triple-differential cross section to theoretical predictions. The determination of the photon signal purity contributes the main systematic uncertainty relating to this measurement. The central values of the cross section, the statistical uncertainty, and the total systematic uncertainty are summarized in Tables 7.1, 7.2, 7.3, and 7.4.

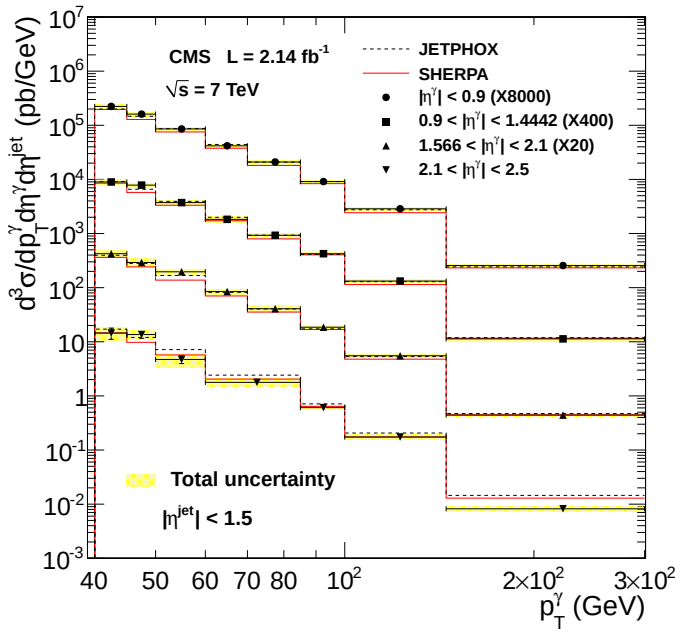


Figure 7.1: Differential cross-sections for $|\eta^{jet}| < 1.5$. The measured cross sections (markers) in four different ranges of η^{γ} are compared with the SHERPA tree-level MC (solid line) and the NLO perturbative QCD calculation from JETPHOX (dashed line). Error bars show statistical uncertainties and the shaded bands correspond to the total experimental uncertainties.

In addition to the triple-differential cross section measurement, the ratios of cross sections with different angular orientations between the photon and the leading jet are studied. The advantage of this measurement is that the uncertainties from the

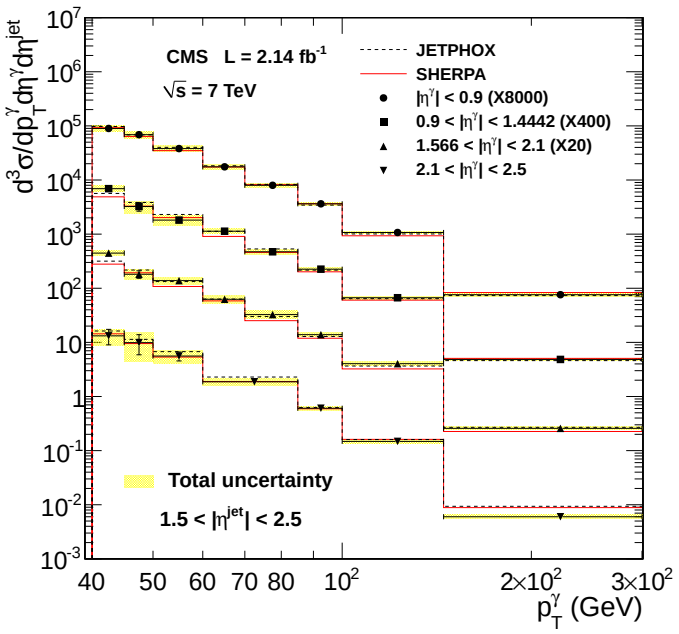


Figure 7.2: Differential cross-sections for $1.5 < |\eta^{jet}| < 2.5$. The measured cross sections (markers) in four different ranges of η^{γ} are compared with the SHERPA tree-level MC (solid line) and the NLO perturbative QCD calculation from JETPHOX (dashed line). Error bars show statistical uncertainties and the shaded bands are the total experimental uncertainties.

integrated luminosity and the reconstruction efficiencies are largely canceled. In the earlier study performed by the D0 experiment at the Tevatron (65), the photons are located in the region $|\eta^{\gamma}| < 1.0$ and the jet is required to be either in the central ($|\eta^{jet}| < 0.8$) or in the forward region ($1.5 < |\eta^{jet}| < 2.5$). The analysis performed by the ATLAS experiment at the LHC (69) studied photons in the pseudorapidity range of $|\eta^{\gamma}| < 1.37$, using data corresponding to an integrated luminosity of 37 pb^{-1}. In this analysis, the kinematic regions of $|\eta^{\gamma}| < 0.9$ and $|\eta^{jet}| < 1.5$ or $1.5 < |\eta^{jet}| < 2.5$ are considered. The ratios of cross sections with various angular orientations between the photon and the leading jet are displayed in Fig. 7.4. The corresponding predictions from SHERPA and JETPHOX are consistent with data, except for cases of photons measured in the largest η and p_T regions.

Table 7.1: The triple-differential cross sections $d^3\sigma/(dp_T^\gamma d\eta^\gamma d\eta^{jet})$ for photons and jets located in the central region with statistical and systematic uncertainties, compared to predictions from JETPHOX and SHERPA. A 2.2% luminosity uncertainty is included in the systematic uncertainty. The final two columns show the ratio of data to JETPHOX (D/J) and SHERPA (D/S), respectively.

$\lvert\eta^\gamma\rvert < 0.9$ and $\lvert\eta^{jet}\rvert < 1.5$					
p_T^γ	Cross section (pb/GeV)			Ratio	
(GeV)	DATA	JETPHOX	SHERPA	D/J	D/S
40-45	27.9±1.0±1.8	24.9	24.5	1.12±0.08	1.14±0.08
45-50	20.1±1.0±1.2	18.3	16.0	1.10±0.08	1.26±0.10
50-60	10.70±0.40±0.75	10.8	9.41	0.99±0.08	1.14±0.09
60-70	5.22±0.16±0.35	5.53	4.71	0.94±0.07	1.11±0.08
70-85	2.62±0.09±0.20	2.61	2.26	1.00±0.08	1.16±0.10
85-100	1.14±0.01±0.06	1.14	1.04	1.00±0.06	1.09±0.06
100-145	0.358±0.003±0.020	0.344	0.303	1.04±0.06	1.18±0.07
145-300	0.0320±0.0002±0.0017	0.0302	0.0290	1.06±0.06	1.10±0.06
$0.9 < \lvert\eta^\gamma\rvert < 1.4442$ and $\lvert\eta^{jet}\rvert < 1.5$					
p_T^γ	Cross section (pb/GeV)			Ratio	
(GeV)	DATA	JETPHOX	SHERPA	D/J	D/S
40-45	22.4±1.4±1.9	22.8	21.3	0.98±0.10	1.05±0.11
45-50	19.6±1.0±1.3	16.4	14.4	1.19±0.10	1.36±0.11
50-60	9.32±0.50±0.76	9.82	8.32	0.95±0.09	1.12±0.11
60-70	4.57±0.20±0.58	4.99	4.32	0.92±0.12	1.06±0.14
70-85	2.32±0.10±0.16	2.33	1.99	1.00±0.08	1.17±0.10
85-100	1.06±0.01±0.06	1.03	1.01	1.03±0.06	1.05±0.06
100-145	0.331±0.004±0.018	0.322	0.285	1.03±0.06	1.16±0.07
145-300	0.0283±0.0003±0.0015	0.0298	0.0291	0.95±0.05	0.97±0.05

7.2 Conclusion

In conclusion, this thesis describes the measurement of the triple-differential cross section for at least one photon and one jet using a data sample with an integrated luminosity of 2.14 fb^{-1} collected by the CMS experiment in proton-proton collisions at $\sqrt{s} = 7$ TeV. This is the first measurement of the cross sections for production of photon+jet final states from the CMS experiment. The triple-differential cross section ($d^3\sigma/(dp_T^\gamma d\eta^\gamma d\eta^{jet})$) is measured as a function of the transverse momentum of the photon for various orientations in pseudorapidity between the leading photon and jet. The individual cross sections determined for eight different photon and jet pseu-

Table 7.2: The triple-differential cross sections $d^3\sigma/(dp_T^\gamma d\eta^\gamma d\eta^{jet})$ for photons located in the central region and jets located in the forward region with statistical and systematic uncertainties, compared to predictions from JETPHOX and SHERPA. A 2.2% luminosity uncertainty is included in the systematic uncertainty. The final two columns show the ratio of data to JETPHOX (D/J) and SHERPA (D/S), respectively.

$\lvert\eta^\gamma\rvert < 0.9$ and $1.5 < \lvert\eta^{jet}\rvert < 2.5$					
p_T^γ	Cross section (pb/GeV)			Ratio	
(GeV)	DATA	JETPHOX	SHERPA	D/J	D/S
40-45	11.2±1.0±1.1	12.2	11.6	0.92±0.12	0.97±0.13
45-50	8.59±0.82±1.04	8.52	7.94	1.01±0.16	1.08±0.17
50-60	4.76±0.36±0.43	5.02	4.36	0.95±0.11	1.09±0.13
60-70	2.19±0.14±0.20	2.29	2.17	0.96±0.11	1.01±0.11
70-85	0.998±0.061±0.074	1.04	1.02	0.96±0.09	0.97±0.09
85-100	0.454±0.009±0.027	0.429	0.455	1.06±0.07	1.00±0.06
100-145	0.134±0.002±0.008	0.126	0.116	1.06±0.06	1.15±0.07
145-300	0.0095±0.0001±0.0005	0.0091	0.0104	1.04±0.06	0.91±0.05
$0.9 < \lvert\eta^\gamma\rvert < 1.4442$ and $1.5 < \lvert\eta^{jet}\rvert < 2.5$					
p_T^γ	Cross section (pb/GeV)			Ratio	
(GeV)	DATA	JETPHOX	SHERPA	D/J	D/S
40-45	17.3±1.3±1.8	14.1	12.2	1.22±0.15	1.42±0.18
45-50	8.1±1.5±1.4	9.62	8.23	0.84±0.21	0.98±0.25
50-60	4.54±0.61±0.66	5.77	5.05	0.79±0.16	0.90±0.18
60-70	2.83±0.18±0.22	2.82	2.27	1.00±0.10	1.25±0.13
70-85	1.18±0.09±0.09	1.33	1.15	0.89±0.10	1.03±0.11
85-100	0.563±0.013±0.034	0.541	0.503	1.04±0.07	1.12±0.07
100-145	0.167±0.003±0.010	0.161	0.151	1.04±0.06	1.11±0.07
145-300	0.0121±0.0002±0.0007	0.0115	0.0127	1.05±0.06	0.96±0.05

dorapidity regions enable access to a wide range of parton momentum fraction x and parton interaction momentum scale Q^2. The ratio of cross sections for photon and jets in several orientations are also presented, in which the uncertainties from efficiency and luminosity measurement are reduced. The detailed evaluations of the photon efficiency calculation, photon purity determination and unfolding of detector effects are presented. The results are also compared with the theoretical calculations from the QCD tree level predicition of SHERPA and the NLO prediction of JETPHOX. The predictions from SHERPA are found to be lower than those from JETPHOX. The measured cross sections in general agree with both MC predictions within systematic

Table 7.3: The triple-differential cross sections $d^3\sigma/(dp_T^\gamma d\eta^\gamma d\eta^{jet})$ for photons located in forward region and jets located in the central region with statistical and systematic uncertainties, compared to predictions from JETPHOX and SHERPA. A 2.2% luminosity uncertainty is included in the systematic uncertainty. The final two columns show the ratio of data to JETPHOX (D/J and SHERPA (D/S), respectively.

$1.556 < \lvert\eta^\gamma\rvert < 2.1$ and $\lvert\eta^{jet}\rvert < 1.5$					
p_T^γ	Cross section (pb/GeV)			Ratio	
(GeV)	DATA	JETPHOX	SHERPA	D/J	D/S
40-45	21.2±2.0±1.9	19.8	18.1	1.07±0.14	1.17±0.15
45-50	14.6±1.4±2.0	14.0	12.1	1.04±0.17	1.21±0.20
50-60	9.82±0.67±0.92	8.38	6.89	1.17±0.14	1.43±0.17
60-70	4.23±0.26±0.39	4.10	3.51	1.03±0.11	1.20±0.13
70-85	2.04±0.11±0.24	2.02	1.77	1.01±0.13	1.15±0.15
85-100	0.928±0.019±0.058	0.868	0.842	1.07±0.07	1.10±0.07
100-145	0.276±0.005±0.017	0.267	0.239	1.04±0.07	1.16±0.08
145-300	0.0221±0.0003±0.0016	0.0236	0.0223	0.94±0.07	0.99±0.07
$2.1 < \lvert\eta^\gamma\rvert < 2.5$ and $\lvert\eta^{jet}\rvert < 1.5$					
p_T^γ	Cross section (pb/GeV)			Ratio	
(GeV)	DATA	JETPHOX	SHERPA	D/J	D/S
40-45	14.5±3.4±1.6	17.1	14.5	0.85±0.22	1.00±0.26
45-50	13.6±2.0±1.3	12.0	9.77	1.13±0.20	1.39±0.25
50-60	4.72±0.76±1.19	7.17	5.71	0.66±0.20	0.83±0.25
60-85	1.78±0.16±0.24	2.42	2.05	0.74±0.12	0.87±0.14
85-100	0.607±0.031±0.044	0.713	0.641	0.85±0.08	0.95±0.08
100-145	0.174±0.008±0.015	0.206	0.174	0.84±0.08	1.00±0.10
145-300	0.0082±0.0004±0.0006	0.0145	0.0129	0.56±0.05	0.63±0.06

uncertainties for most of the measured kinematic regions. The disagreement between the measured result and the theoretical predictions found in the kinematic region of highest η^γ and p_T^γ may imply the need for an improved theoretical description of the differential distributions for the photon in the forward regions. These regions can be explored in more detail using larger data sets collected by the CMS detector.

Table 7.4: The triple-differential cross sections $d^3\sigma/(dp_T^\gamma d\eta^\gamma d\eta^{jet})$ for photons and jets located in forward region with statistical and systematic uncertainties, compared to predictions from JETPHOX and SHERPA. A 2.2% luminosity uncertainty is included in the systematic uncertainty. The final two columns show the ratio of data to JETPHOX (D/J and SHERPA (D/S), respectively.

$1.556 < \lvert\eta^\gamma\rvert < 2.1$ and $1.5 < \lvert\eta^{jet}\rvert < 2.5$					
p_T^γ	Cross section (pb/GeV)			Ratio	
(GeV)	DATA	JETPHOX	SHERPA	D/J	D/S
40-45	22.3±1.4±1.8	15.8	14.0	1.41±0.14	1.60±0.16
45-50	9.1±1.4±1.1	10.9	9.66	0.83±0.17	0.94±0.19
50-60	6.92±0.68±0.84	6.65	5.39	1.04±0.16	1.28±0.20
60-70	3.13±0.21±0.42	3.15	2.92	0.99±0.15	1.07±0.16
70-85	1.63±0.11±0.25	1.50	1.26	1.09±0.18	1.29±0.22
85-100	0.694±0.017±0.055	0.643	0.596	1.08±0.09	1.16±0.10
100-145	0.202±0.004±0.015	0.183	0.162	1.10±0.08	1.25±0.10
145-300	0.0129±0.0002±0.0008	0.0135	0.0113	0.96±0.06	1.14±0.08
$2.1 < \lvert\eta^\gamma\rvert < 2.5$ and $1.5 < \lvert\eta^{jet}\rvert < 2.5$					
p_T^γ	Cross section (pb/GeV)			Ratio	
(GeV)	DATA	JETPHOX	SHERPA	D/J	D/S
40-45	13.2±4.2±1.4	16.2	14.4	0.81±0.27	0.92±0.31
45-50	9.9±4.0±3.7	11.4	9.51	0.87±0.48	1.04±0.57
50-60	5.6±1.0±1.0	6.75	5.36	0.83±0.22	1.04±0.27
60-85	1.87±0.18±0.23	2.29	1.88	0.82±0.13	0.99±0.15
85-100	0.607±0.029±0.051	0.628	0.593	0.97±0.09	1.02±0.10
100-145	0.148±0.006±0.011	0.160	0.161	0.92±0.08	0.92±0.08
145-300	0.0060±0.0003±0.0004	0.0094	0.0088	0.64±0.06	0.68±0.06

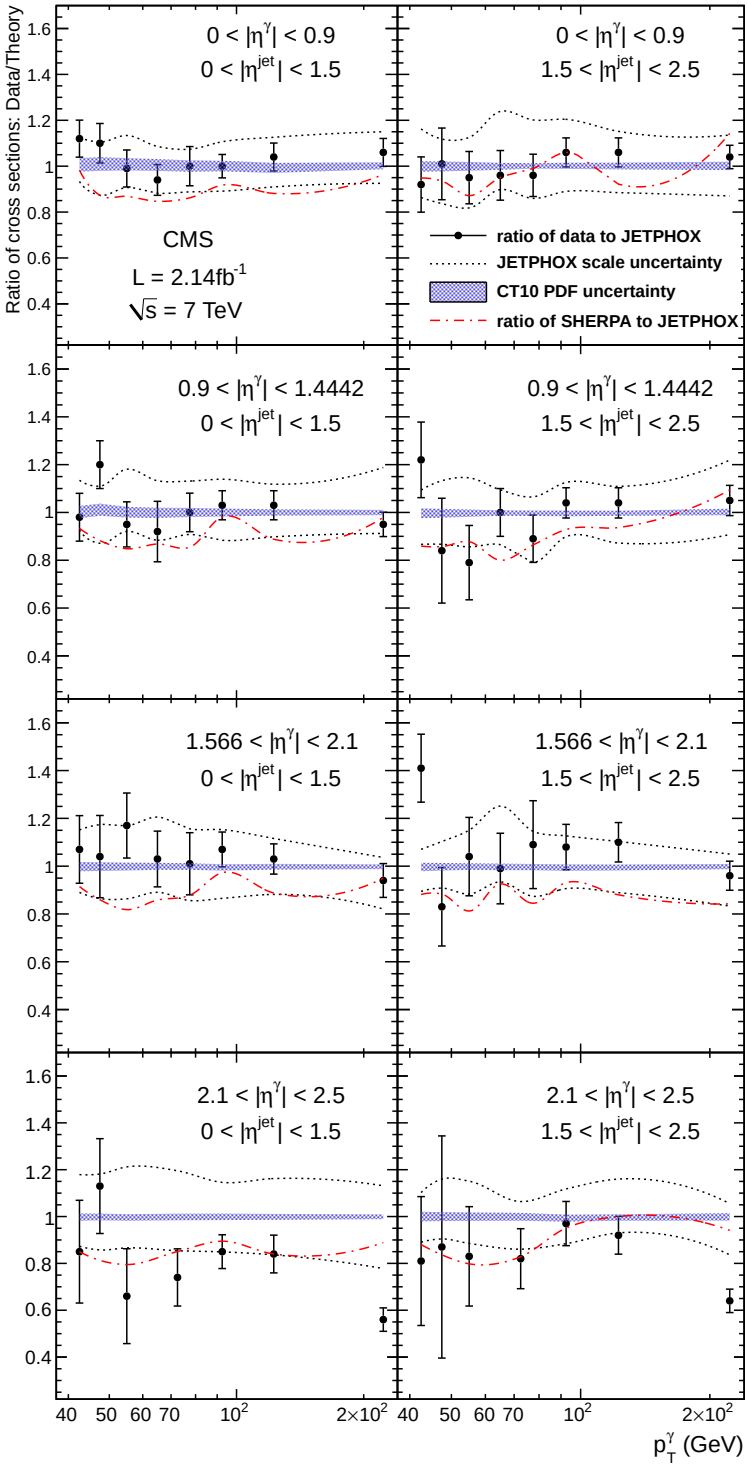


Figure 7.3: The ratios of the measured triple-differential cross sections to the NLO QCD prediction using JETPHOX with the CT10 PDF set and scales $\mu_{R,F,f} = \frac{1}{2}p_T^{\gamma}$. The error bars show the statistical and systematic uncertainties added in quadrature. The two dotted lines represent the effect of varying the theoretical scales as described in the text. The shaded bands correspond to the CT10 PDF uncertainty. The dash-dotted lines show the ratios of the SHERPA predictions to JETPHOX.

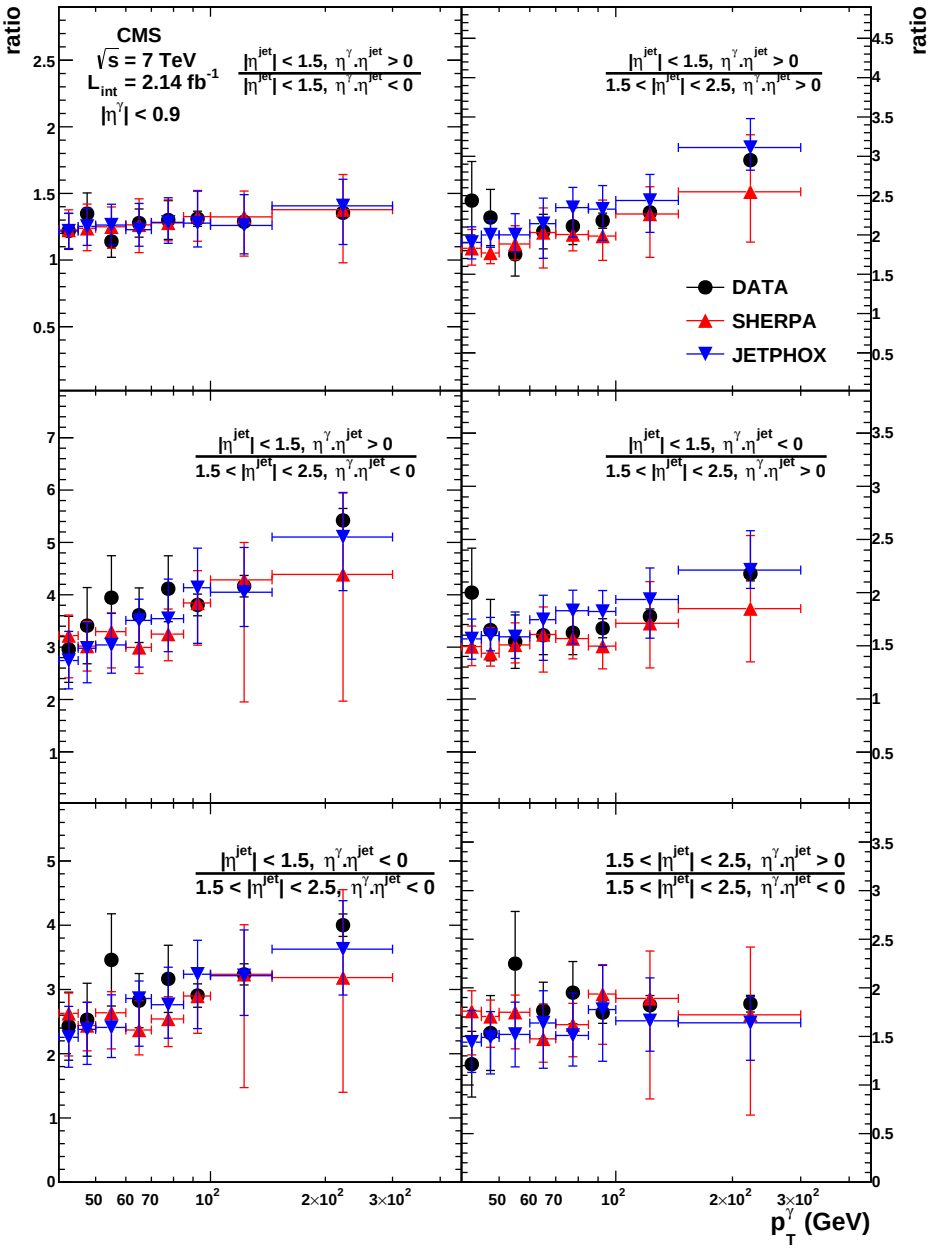


Figure 7.4: Ratios of the triple-differential cross sections for the various jet orientations with respect to the photon. The error bars on the theoretical predictions correspond to statistical and systematic uncertainties.

Bibliography

[1] J. Beringer *et al.*, "Review of Particle Physics (RPP)," *Phys.Rev.*, vol. D86, p. 010001, 2012. v, 10, 12

[2] A. Martin, W. Stirling, R. Thorne, and G. Watt, "Parton distributions for the LHC," *Eur.Phys.J.*, vol. C63, pp. 189–285, 2009. v, 11, 13

[3] F. Hartmann and A. Sharma, "Multipurpose detectors for high energy physics, an introduction," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 666, no. 0, pp. 1 – 9, 2012. Advanced Instrumentation. v, 26

[4] S. Chatrchyan *et al.*, "Performance of CMS Hadron Calorimeter Timing and Synchronization using Test Beam, Cosmic Ray, and LHC Beam Data," *JINST*, vol. 5, p. T03013, 2010. v, 32

[5] S. Chatrchyan *et al.*, "Performance of CMS muon reconstruction in *pp* collision events at $\sqrt{s} = 7$ TeV," *JINST*, vol. 7, p. P10002, 2012. vi, 35

[6] G. D'Agostini, "A Multidimensional unfolding method based on Bayes' theorem," *Nucl. Instrum. Meth.*, vol. A362, pp. 487–498, 1995. vii, 82, 87, 88

[7] P. Aurenche and J. Lindfors, "Direct photon production beyond leading order in qcd," *Nuclear Physics B*, vol. 168, no. 2, pp. 296 – 314, 1980. 2

[8] P. Aurenche, A. Douiri, R. Baier, M. Fontannaz, and D. Schiff, "Prompt photon production at large pt in qcd beyond the leading order," *Physics Letters B*, vol. 140, no. 12, pp. 87 – 92, 1984. 2

[9] J. F. Owens, "Large-momentum-transfer production of direct photons, jets, and particles," *Rev. Mod. Phys.*, vol. 59, pp. 465–503, Apr 1987. 2

[10] P. Aurenche, R. Baier, M. Fontannaz, J. F. Owens, and M. Werlen, "Gluon content of the nucleon probed with real and virtual photons," *Phys. Rev. D*, vol. 39, pp. 3275–3286, Jun 1989. 2

[11] W. Vogelsang and A. Vogt, "Constraints on the proton ' *s* gluon distribution from prompt photon production," *Nucl.Phys.*, vol. B453, pp. 334–354, 1995. 2

[12] D. d'Enterria and J. Rojo, "Quantitative constraints on the gluon distribution function in the proton from collider isolated-photon data," *Nucl.Phys.*, vol. B860, pp. 311–338, 2012. 2

BIBLIOGRAPHY

[13] L. Carminati, G. Costa, D. D'Enterria, I. Koletsou, G. Marchiori, *et al.*, "Sensitivity of the LHC isolated-gamma+jet data to the parton distribution functions of the proton," *EPL*, vol. 101, p. 61002, 2013. 2, 15

[14] S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys.Lett.*, vol. B716, pp. 30–61, 2012. 2, 15

[15] S. Chatrchyan *et al.*, "Search for Dark Matter and Large Extra Dimensions in pp Collisions Yielding a Photon and Missing Transverse Energy," *Phys.Rev.Lett.*, vol. 108, p. 261803, 2012. 2, 16

[16] S. Bhattacharya, S. S. Chauhan, B. C. Choudhary, and D. Choudhury, "Quark Excitations Through the Prism of Direct Photon Plus Jet at the LHC," *Phys.Rev.*, vol. D80, p. 015014, 2009. 2, 16

[17] T. C. collaboration, "Determination of jet energy calibration and transverse momentum resolution in cms," *Journal of Instrumentation*, vol. 6, no. 11, p. P11002, 2011. 2, 42

[18] CMS Collaboration, "Missing transverse energy performance of the CMS detector," *Journal of Instrumentation*, vol. 6, p. 9001, Sept. 2011. 2

[19] S. Glashow, "Partial Symmetries of Weak Interactions," *Nucl.Phys.*, vol. 22, pp. 579–588, 1961. 3

[20] S. Chatrchyan *et al.*, "A new boson with a mass of 125-GeV observed with the CMS experiment at the Large Hadron Collider," *Science*, vol. 338, pp. 1569–1575, 2012. 3

[21] P. Aurenche, A. Douiri, R. Baier, M. Fontannaz, and D. Schiff, "Prompt Photon Production at Large p(T) in QCD Beyond the Leading Order," *Phys.Lett.*, vol. B140, p. 87, 1984. 15

[22] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, "The Hierarchy problem and new dimensions at a millimeter," *Phys.Lett.*, vol. B429, pp. 263–272, 1998. 16

[23] H. Baer, P. G. Mercadante, F. Paige, X. Tata, and Y. Wang, "Lhc reach for gauge mediated supersymmetry breaking models via prompt photon channels," *Physics Letters B*, vol. 435, no. 12, pp. 109 – 117, 1998. 16

[24] A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. Hoche, *et al.*, "General-purpose event generators for LHC physics," *Phys.Rept.*, vol. 504, pp. 145–233, 2011. 16, 18

[25] B. Webber, "Parton shower monte carlo event generators," *Scholarpedia*, vol. 6, no. 12, p. 10662, 2011. 16, 18

[26] T. Sjostrand, "Monte Carlo Generators," pp. 51–74, 2006. 16, 18

[27] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, "Parton fragmentation and string dynamics," *Physics Reports*, vol. 97, no. 2-3, pp. 31 – 145, 1983. 18

[28] B. Andersson. Cambridge University Press, 1998. 18

[29] D. Amati and G. Veneziano, "Preconfinement as a property of perturbative qcd," *Physics Letters B*, vol. 83, no. 1, pp. 87 – 92, 1979. 18

[30] T. Sjöstrand and M. van Zijl, "A multiple-interaction model for the event structure in hadron collisions," *Phys. Rev. D*, vol. 36, pp. 2019–2041, Oct 1987. 18

[31] R. Field, "Min-Bias and the Underlying Event at the LHC," 2012. 18, 46

[32] e. a. S. Agostinelli, "Geant4a simulation toolkit," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250 – 303, 2003. 19

[33] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP*, vol. 0605, p. 026, 2006. 19

[34] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, *et al.*, "Event generation with SHERPA 1.1," *JHEP*, vol. 0902, p. 007, 2009. 19, 91

[35] S. Catani, M. Fontannaz, J. Guillet, and E. Pilon, "Cross-section of isolated prompt photons in hadron hadron collisions," *JHEP*, vol. 0205, p. 028, 2002. 19, 91

[36] P. Aurenche, M. Fontannaz, J.-P. Guillet, E. Pilon, and M. Werlen, "A New critical study of photon production in hadronic collisions," *Phys.Rev.*, vol. D73, p. 094007, 2006. 19

[37] Z. Belghobsi, M. Fontannaz, J.-P. Guillet, G. Heinrich, E. Pilon, *et al.*, "Photon - Jet Correlations and Constraints on Fragmentation Functions," *Phys.Rev.*, vol. D79, p. 114024, 2009. 19

[38] C. collaboration, "Cms physics technical design report," 27, 30, 32, 34, 35, 37, 39

[39] T. C. Collaboration, "The cms experiment at the cern lhc," *Journal of Instrumentation*, vol. 3, no. 08, p. S08004, 2008. 28

[40] D. Sprenger, M. Weber, R. Adolphi, R. Brauer, L. Feld, *et al.*, "Validation of Kalman Filter alignment algorithm with cosmic-ray data using a CMS silicon strip tracker endcap," *JINST*, vol. 5, p. P06007, 2010. 29

[41] K. Kaschube, "First alignment of the complete CMS silicon tracker," *IEEE Nucl.Sci.Symp.Conf.Rec.*, vol. 2009, pp. 1189–1193, 2009. 29

[42] R. Brown and D. Cockerill, "Electromagnetic calorimetry," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 666, no. 0, pp. 47 – 79, 2012. Advanced Instrumentation. 30

[43] S. Chatrchyan *et al.*, "Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7$ TeV," 2013. 31

[44] A. Zabi, "Triggering on electrons and photons with the CMS experiment at the LHC," *J.Phys.Conf.Ser.*, vol. 404, p. 012056, 2012. 36

[45] D. Bonacorsi, "The cms computing model," *Nuclear Physics B - Proceedings Supplements*, vol. 172, no. 0, pp. 53 – 56, 2007. 37

BIBLIOGRAPHY

[46] T. Awes, F. Obenshain, F. Plasil, S. Saini, S. Sorensen, and G. Young, "A simple method of shower localization and identification in laterally segmented calorimeters," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 311, no. 12, pp. 130 – 138, 1992. 41

[47] "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET," 2009. 41

[48] M. Pioppi, "A pre-identification for electron reconstruction in the cms particle-flow algorithm," *Journal of Physics: Conference Series*, vol. 119, no. 3, p. 032039, 2008. 41

[49] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm," *JHEP*, vol. 0804, p. 063, 2008. 42

[50] Sjostrand, T. and Mrenna, S. and Skands, P., "Pythia 6.4 physics and manual." 2006. 46

[51] J. Pumplin, D. Stump, J. Huston, H. Lai, P. Nadolsky, and W. Tung, "New generation of parton distributions with uncertainties from global qcd analysis," *JHEP*, vol. 0207, p. 012, 2002. 46

[52] M. Cacciari, G. Salam, and S. G., "The anti-kt jet clustering algorithm," *JHEP*, vol. 0804:063, 2008. 50

[53] "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET," Tech. Rep. CMS-PAS-PFT-09-001, CERN, 2009. Geneva, Apr 2009. 50

[54] CMS Collaboration, "Calorimeter jet quality criteria for the first cms collision data," *CMS Physics Analysis Summary*, vol. CMS-PAS-JME-09-008, 2009. 50

[55] "Absolute calibration of the luminosity measurement at cms: Winter 2012 update," Tech. Rep. CMS-PAS-SMP-12-008, CERN, Geneva, 2012. 53, 89

[56] V. Chetluru and K. Ocalan, "Photon plus jet: Single photon hlt path efficiency measurements for the 2011 dataset," *CMS Note*, vol. CMS AN-2011-314, 2011.

[57] "Measuring Electron Efficiencies at CMS with Early Data," 2008. 55

[58] J. Veverka, "Photon performance measurements with $z \to \mu\mu\gamma$ events selected at 7 tev," *CMS Note*, vol. CMS AN-2012-043, 2012. 60, 72

[59] H. Albrecht *et al.*, "Search for hadronic $b \to u$ decays," *Physics Letters B*, vol. 241, no. 2, pp. 278 – 282, 1990. 63

[60] "Pileup mc reweighting utilities," 71

[61] T. Adye, "Unfolding algorithms and tests using RooUnfold," *ArXiv e-prints*, May 2011. 82

[62] A. Bornheim, J. Veverka, V. Timciuc, and Y. Ma, "Measuremant of cms ecal scale and resolution performance from $z \to ee$ decays and determination of the signal line shape $h \to \gamma\gamma$," *CMS Analysis Note*, vol. CMS AN-2011-204, 2011. 88

[63] S. Hoeche, F. Krauss, S. Schumann, and F. Siegert, "QCD matrix elements and truncated showers," *JHEP*, vol. 0905, p. 053, 2009. 91

[64] S. Hoeche, S. Schumann, and F. Siegert, "Hard photon production and matrix-element parton-shower merging," *Phys.Rev.*, vol. D81, p. 034026, 2010. 91

[65] V. Abazov *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.*, vol. B666, pp. 435–445, 2008. 91, 93

[66] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, *et al.*, "New generation of parton distributions with uncertainties from global QCD analysis," *JHEP*, vol. 0207, p. 012, 2002. 91

[67] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, *et al.*, "New parton distributions for collider physics," *Phys.Rev.*, vol. D82, p. 074024, 2010. 91

[68] D. Bourilkov, R. C. Group, and M. R. Whalley, "LHAPDF: PDF use from the Tevatron to the LHC," 2006. 92

[69] G. Aad *et al.*, "Measurement of the production cross section of an isolated photon associated with jets in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," *Phys.Rev.*, vol. D85, p. 092014, 2012. 93

Declaration