MEASUREMENT OF DIFFERENTIAL PHOTON-JET CROSS SECTION FROM 7 TEV PP COLLISIONS IN THE CMS EXPERIMENT AT THE LHC

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

KADİR ÖCALAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF PHILOSOPHY OF DOCTORATE
IN
PHYSICS

MARCH 2012

Approval of the thesis:

MEASUREMENT OF DIFFERENTIAL PHOTON-JET CROSS SECTION FROM 7 TEV PP COLLISIONS IN THE CMS EXPERIMENT AT THE LHC

submitted by **KADİR ÖCALAN** in partial fulfillment of the requirements for the degree of **Philosophy of Doctorate in Physics Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Mehmet T. Zeyrek _____
Head of Department, **Physics**

Prof. Dr. Ali Murat Güler _____
Supervisor, **Physics Dept., METU**

Examining Committee Members:

Prof. Dr. Ramazan Sever
Physics Dept., METU _____

Prof. Dr. Ali Murat Güler
Physics Dept., METU _____

Prof. Dr. Erhan Onur İltan
Physics Dept., METU _____

Prof. Dr. Altuğ Özpineci
Physics Dept., METU _____

Assoc. Prof. Dr. Kerem Cankoçak
Physics Engineering Dept., Istanbul Technical University _____

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: KADİR ÖCALAN

Signature :

ABSTRACT

MEASUREMENT OF DIFFERENTIAL PHOTON-JET CROSS SECTION FROM 7 TEV PP COLLISIONS IN THE CMS EXPERIMENT AT THE LHC

Öcalan, Kadir
Ph.D., Department of Physics
Supervisor : Prof. Dr. Ali Murat Güler

March 2012, 125 pages

Prompt photons are produced primarily by quark-gluon compton scattering and quark-anti-quark annihilation mechanisms that their measurements are driven by several motivations at hadron colliders. Measurement of prompt photons can be used for probing perturbative Quantum Chromodynamics (pQCD). Prompt photons are produced in the final states of important decays providing evidence for low mass Higgs boson and new physics searches. Compact Muon Solenoid (CMS) is one of the multi-purpose experiments conducted at the Large Hadron Collider (LHC) to study the Standard Model (SM) physics including prompt photons and new theories in proton-proton collisions. In this thesis, photon reconstruction is discussed along with the photon identification methods in the CMS experiment. Photon efficiency measurements are presented as an important ingredient for photon+jet cross section measurement. Photon High Level Trigger (HLT) efficiencies, reconstruction (RECO) efficiencies, and identification (ID) efficiencies are presented that are measured from collision data recorded by the CMS detector and Monte Carlo simulation data. Efficiency corrected differential photon+jet cross section measurement results are presented in this study. The collision data used in this thesis corresponds to 2.2 /fb integrated luminosity collected by the CMS detector at a center-of-mass energy of 7 TeV from 2011 LHC proton-proton collision running.

Keywords: LHC, CMS Experiment, Collider Photon Physics, Photon Efficiency Measurement, Photon-Jet Cross Section

ÖZ

LHC CMS DENEYİNDE 7 TEV PP ÇARPIŞMALARINDA DİFERANSİYEL FOTON-JET TESİR KESİTİ ÖLÇÜMÜ

Öcalan, Kadir
Doktora, Fizik Bölümü
Tez Yöneticisi : Prof. Dr. Ali Murat Güler

Mart 2012, 125 sayfa

Öncül fotonlar öncelikli olarak quark-gluon compton saçılması ve quark-anti-quark annihilasyon mekanizmalarında üretilir ve bu fotonların hadron çarpıştırıcılarında ölçülmesi için birçok motivasyon vardır. Öncül fotonların ölçümü pertürbatif Kuantum Renk Dinamiği teorisinin (pQCD) araştırılması için kullanılabilir. Öncül fotonlar önemli bozunumların son durumlarında üretilmekte olup, düşük kütle Higgs bozonu ve yeni fizik aranması için bulgu teşkil edebilir. Compact Muon Solenoid (CMS) deneyi, Büyük Hadron Çarpıştırıcısı'nda (LHC) gerçekleştirilen çok amaçlı deneylerden biri olup, öncül fotonlar dahil Standard Model (SM) fiziği ve yeni fizik teorilerini proton-proton çarpışmalarında çalışmaktadır. Bu tezde, CMS deneyi foton yeniden yapılandırma ve foton belirlenmesi yöntemleri tartışılmaktadır. Foton+jet tesir kesiti ölçümü için önem arzeden foton verimlilik ölçümleri sunulmaktadır. CMS dedektörü tarafından kaydedilen çarpışma verisinden ve Monte Carlo benzeşim verisinden ölçülen, foton Yüksek Seviye Tetikleyici (HLT) verimlilikleri, yeniden yapılandırma verimlilikleri (RECO) ve belirlenmesi (ID) verimlilikleri sunulmaktadır. Verimlilik bakımından düzeltilmiş diferansiyel foton+jet tesir kesiti ölçümü sonuçları bu çalışmada sunulmaktadır. Bu tezde kullanılan çarpışma verisi CMS dedektörüyle toplanan 2.2 /fb entegre lüminosite değerine tekabül etmekte olup, kütle merkezi enerjisi 7 TeV'lik LHC'nin 2011 yılındaki

proton-proton çarpışması çalışmasından elde edilmiştir.

Anahtar Kelimeler: LHC, CMS Deneyi, Çarpıştırıcı Foton Fiziği, Foton Verimlilik Ölçümü, Foton-Jet Tesir Kesiti

To my wife Hümeyra

ACKNOWLEDGMENTS

I would like to express my deep gratitude to my supervisor Prof. Dr. Ali Murat Güler (METU, Ankara, Turkey) for supporting my graduate study with doctorate research at CERN. I am grateful to Prof. Dr. Ramazan Sever (METU, Ankara, Turkey) and Assoc. Prof. Dr. Kerem Cankoçak (ITU, Istanbul, Turkey) for their supports and thesis follow-ups.

I indebted to the members of the CMS Collaboration QCD Photons group with whom I have interacted and worked together during the course of my Ph.D. research studies. Particularly, I would like to express my gratitude to the CMS Egamma HLT group convener and also our analyis coordinator Dr. Vasundhara Chetluru (Fermilab, Illinois, USA) for her help and motivations through my involvement in the QCD Photons physics programs and High Level Trigger (HLT) studies. This thesis would not have been completed without her guidance, support, and expertise on photon physics. It has always been a great pleasure for me to work with her. Next, I am grateful to the CMS former QCD Photons group convener, Dr. Shin-Shan Eiko Yu (National Central University, Chung-Li, Taiwan) for the kind helps and directions which speeded up my orientation to the group.

I have many thanks to my colleagues that I have been working with for their cooperation: Sudha Ahuja (University of Delhi, Delhi, India), Chuanzhe Lin (University of Virginia, Charlottesville, USA), and Dr. Darko Mekterovic (National Central University, Chung-Li, Taiwan). The physics results presented in this work are as mine as theirs. I have special thanks to Dr. Michael Anderson (University of Wisconsin, Madison, Wisconsin, USA) for the helps to run the software package for photon efficiency measurement.

My sincere thanks go to the CMS former Egamma HLT and current Standard Model Physics (SMP) groups convener Dr. Jeffrey Berryhill (Fermilab, Illinois, USA) for providing us valuable feedbacks for trigger efficiency measurements. We are grateful to the convener of the CMS former QCD Physics and current Standard Model Physics (SMP) groups convener Dr. Konstantinos Kousouris (CERN, Geneva, Switzerland) for supporting our physics analysis efforts. I also thank former QCD Photons group convener and current Standard Model Physics

(SMP) Photons object group coordinator Dr. Susan Gascon-Shotkin (IPN Lyon/Université Claude Bernard Lyon 1, Villeurbanne, France) for her very useful discussions on the physics analysis presented in this thesis.

I owe my loving thanks to my wife Hümeyra to whom this thesis dedicated. During the research and the preperation stages of this work, she was the one to support, understand, and motivate me. She took care of our lives when I was fully occupied by my doctorate research. Moreover, I would like to express my heart-felt gratitude to my father and my mother for everything they gave me. I am thankful to my parents and sisters for their understandings and beliefs in me. The last but not least, I would like to thank Dr. Ayhan Ekti and Hülya Ekti for their endless supports.

This thesis work is financially supported by the Turkish Atomic Energy Agency (TAEK) with Project No:05K120010 and Scientific and Technological Research Council of Turkey (TUBITAK) with Project No:108T324.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

Quantum Chromodynamics (QCD) part of the Standard Model provides a description for hard processes in hadron-hadron interactions [13], [14]. Prompt photon with or without associated jet production is one of the most attractive QCD related field at the hadron colliders [18], [19]. The experimental studies of prompt photon production are driven by several motivations. Prompt photon production cross section measurement provides information on the parton distribution functions (PDFs) for testing perturbative QCD (pQCD). The production of prompt photons at large transverse momentum are sensitive to the presence of new physics.

The Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) (Geneva, Switzerland) is mainly a hadron collider that operates at the highest energies in the world [32]. It collides proton beams head-on with a center-of-mass energy of 7 TeV. Created particles in the collisions are recorded by the special detectors in a number of experiments dedicated to the LHC. The Compact Muon Solenoid (CMS) experiment is one of two large general-purpose particle physics detectors constructed at the LHC [36], [37]. The LHC experiments including the CMS target to test various predictions of the Standard Model and to search for new physics models.

Prompt photon reconstruction in the CMS experiment is based on clustering energy deposits in the electromagnetic calorimeter (ECAL) [51]. Reconstructed photon candidates are discriminated from background processes which are primarily electrons and jets with the help of sets of isolation variables in photon identification stage [52]. Signal photons are estimated by using the so-called signal extraction template isolation requirements. High purity photon events are further counted and processed for photon with or without associated jet cross section measurements.

Photon efficiency measurement is one of the most important ingredient of cross section measurement in high energy physics experiments. Efficiencies of photon triggering, reconstruction, and identification need to be measured from both the collision data and simulation data in order to provide appropriate corrections for final photon spectra. The Tag and Probe is the most widely used technique to estimate photon efficiencies in the CMS experiment [70], [71].

The study of prompt photon production with associated jet in the final state of high energy collisions provides means for testing perturbative QCD (pQCD) predictions such as for probing the gluon distribution in the proton [20]. In addition, prompt photons (and associated jet production) provide an estimate of the background for important searches within and beyond the Standard Model. They constitute a background for the searches of low mass Higgs boson ($H\rightarrow \gamma\gamma$), the Randall-Sundrum gravitons, Gauge Mediated Supersymmetry, and Universal Extra Dimensions with two energetic photons in the final state.

The thesis organization can be summarized here. In Chapter 2 of this thesis, the Standard Model of High Energy Physics including brief description of QCD theory is reviewed. The production mechanisms and physics motivations of prompt and isolated photons are discussed along with their production in association with jets. Theory prediction generator tools are also discussed for prompt photon production cross section measurements in this chapter. Chapter 3 of this thesis is reserved for describing general design and operation of the LHC. The CMS experiment is extensively described in this chapter by providing detailed information on its coordinate system and sub-dedectors including Tracker, Calorimeters, Muon system, and Triggers with their physics objectives. The Computing model of the CMS data processing at the LHC is also summarized in this part. Chapter 4 provides descriptions of important steps in cross section measurement in the CMS experiment: photon reconstruction, photon identification with isolation variables, and photon signal extraction templates. The most recent measurements of isolated prompt photon with(out) associated jet cross section in experiments are also reviewed in this chapter. Chapter 5 focuses on presenting photon trigger, reconstruction, and identification efficiencies for prompt photon in association with jet (photon+jet) cross section measurement with the collision data recorded by the CMS experiment during the year of 2011 at $\sqrt{s} = 7$ TeV and Monte Carlo simulation samples. Efficiency results are presented as a function of photon transverse energy and pseudorapidity in this dedicated chapter. In Chapter 6, measurement of the triple differential cross section for the γ+jet final

state production is presented using data collected with the CMS detector from proton-proton collisions at the LHC with $\sqrt{s} = 7$ TeV. The photons and jets are required to be reconstructed with a maximum pseudorapidity of $|\eta^{\gamma}, \eta^{jet}| < 2.5$ and to have transverse momenta in the ranges of $40 < p_T^{\gamma} < 300$ GeV and $p_T^{jet} > 30$ GeV, respectively. The differential cross section $d^3\sigma/dp_T^{\gamma}d\eta^{\gamma}d\eta^{jet}$ measurement results are given in various possible kinematic regions of photon and jet objects. The final results are reported after correcting photon spectra by photon efficiency in detection, photon purity in signal extraction, and unfolding factors in detector reconstruction effects. Theoretical comparisions to data-driven results are also presented from the Pythia and the Sherpa event generations. Finally in Chapter 7, the concluding remarks for photon efficiency and differential γ+jet cross section measurement results are discussed. Future outlook and work plan for photon (in association with jets) production measurements in the CMS experiment are summarized.

CHAPTER 2

THEORY

The Standard Model (SM) of High Energy Physics is the most well-established particle physics theory as its predictions were confirmed precisely by several experimental measurements. Quantum Chromodynamics (QCD) sector of the SM describes the hard scattering processes in the final states of hadron-hadron interactions. Energetic prompt photon production at hadron colliders is one of the most motivated QCD related fields, since its measurement is highly significant for probing perturbative QCD (pQCD), searches for the Higgs boson which is the missing piece of the SM, and searches for the new physics theories beyond.

2.1 The Standard Model

The Standard Model (SM) of Particle Physics is the most satisfying model to explain particles and their interactions with high degree of precision. It is a quantum field theory which describes the unification of electromagnetic interaction and weak interaction into an electroweak (EWK) sector, and includes a Quantum Chromodynamics (QCD) sector for the description of the strong interactions. In the SM framework, particles are grouped as fermions (matter particles) and bosons (interaction particles).

The matter we observe in the universe is composed of elementary particles, having either half-integer or integer spins. The particles with half-ineteger spin, so-called fermions can be splitted into leptons and quarks with three generations. Leptons are electron, muon, tau, and their corresponding neutrinos (e.g., electron neutrino, ν_e). Quarks come in six flavors (up, down, charm, strange, top, and bottom), two lightest of them make up protons and neutrons. Free quarks have not yet been observed; they instead come in compositions of two ($q\bar{q}$) or

three (qqq). Collections of quarks and anti-quarks form other, composite particles known as hadrons. Hadrons are divided into mesons and baryons: mesons (including pions, kaons, and so on) comprise a quark and an anti-quark while baryons (including proton and neutron) are three quark states. For each fermion, a corresponding anti-particle exists. A fermion (e.g., electron, up quark) and an anti-fermion (positron, anti-up quark) have the same mass, but opposite electric charge. Table 2.1 lists leptons, quarks, and their properties for completeness [1].

Table 2.1: Properties of leptons and quarks ($spin-\frac{1}{2}$). Masses without uncertainties are shown based on the evaluations of the Particle Data Group (PDG). The elementary charge, e, equals to 1.602×10^{-19} C.

Fermions	1^{st} *generation*	2^{nd} *generation*	3^{rd} *generation*
Leptons	electron(e) m_e= 0.511 MeV $q_e = -e$ e neutrino (ν_e) $m_{\nu_e} < 2$ eV $q_{\nu_e} = 0$	muon (μ) m_μ=105.66 MeV $q_\mu = -e$ μ neutrino (ν_μ) $m_{\nu_\mu} < 0.17$ MeV $q_{\nu_\mu} = 0$	tau (τ) m_τ=1776.84 MeV $q_\tau = -e$ τ neutrino (ν_τ) $m_{\nu_\tau} < 15.5$ MeV $q_{\nu_\tau} = 0$
Quarks	up(u) m_u= 2.55 MeV $q_u = \frac{2e}{3}$ down(d) m_d= 5.04 MeV $q_d = -\frac{e}{3}$	charm (c) m_c=1.27 GeV $q_c = \frac{2e}{3}$ strange (s) m_s=105 MeV $q_s = -\frac{e}{3}$	top (t) m_t=171.3 GeV $q_t = \frac{2e}{3}$ bottom (b) m_b=4.20 GeV $q_b = -\frac{e}{3}$

Interactions between fermions occur through the exchange of force carriers, which are known as bosons (interaction particles). The list of bosons which are known with their properties is given compactly in Table 2.2. The photon is massless, has no charge, and mediates electromagnetic interactions known as Quantum Electrodynamics (QED) [2]. Particles interacting through their charge, such as an electron orbiting the nucleus of an atom, exchange photons. Two massive bosons, the charged $W^{\pm}$ and the neutral Z, mediate the weak interactions. The weak force has a short range and allows fermions to decay and change flavor. The W and Z bosons can interact with all leptons and quarks. The electromagnetic interactions mediated by the exchange of photons (QED) and the weak interactions described by the W and Z bosons are unified in the SM into a single electroweak (EWK) theory [3], [4], [5].

The unification of electromagnetic and weak forces requires additional massive particle, the

Table 2.2: Gauge bosons (*spin*-1) of the SM with their corresponding masses and interactions they appear in.

Gauge Bosons	Mass	Interaction
photon (γ)	0.0	electromagnetic
neutral weak boson (Z)	91.2 GeV	weak neutral
charged weak bosons ($W^{\pm}$)	80.4 GeV	weak charged
gluon (g)	0.0	strong (color)

Higgs boson [6], [7]. Masses of the SM particles are explained by means of their interactions with the Higgs boson, including its self-interaction, which gives special status to the Higgs boson within the SM framework. However, the Higgs boson has not yet been observed experimentally. Discovery of the Higgs boson would be a major achievement in particle physics, though the complete description of the Higgs mechanism is not much relevant for this thesis.

The gluon which is massless and chargeless like the photon, carries the strong force. This strong force binds quarks into mesons ($q\overline{q}$) and baryons (qqq) through gluon exchange and holds the nucleus of an atom together. Quarks and gluons are collectively known as partons. The strong interactions of the partons are described by the Quantum Chromodynamics (QCD) theory. Although, all fermions interact weakly, only quarks interact by the strong force. Only the electrically charged particles, quarks, electrons, muons, and taus, interact electromagnetically. Neutrinos have neither color (dislike quarks) nor electric charge, only interact by the weak force. The gravity is the force not included in the SM framework which indicates that the SM theory is not complete and encourages for searches of theories beyond.

The SM is based on the gauge symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. Every group has a coupling constant associated with: g_s (related with α_s) for the strong interactions, g (related with the Fermi constant G_F) for the weak interactions, and g' (related with electric charge e) for the electromagnetic interactions. Further detailed and pedagogical descriptions of the SM can be found elsewhere ([8], [9], [10], [11], and [12]).

2.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) was introduced by Gell-Mann and Fritzsch in 1972 to describe the behavior of quarks being hold together by strong force carried by gluons [13],

[14]. QCD is the renormalizable non-abelian gauge field theory based on $SU(3)_C$ group, where C refers to color quantum number which is introduced to refer to three possible states of the quarks. The corresponding Lagrangian, that contains quark q_k and gluon G^A_α fields, is given by [12], [15]:

$$\mathcal{L}_{QCD} = -\frac{1}{4}F^A_{\alpha\beta}F^{\alpha\beta}_A + \sum_j^{flavors} \overline{q}_j(i\gamma^\mu D_\mu - m_j)q_j \tag{2.1}$$

where $F^A_{\alpha\beta}$ is the field strength tensor for the spin-1 gluon field G^A_α,

$$F^A_{\alpha\beta} = \partial_\alpha G^A_\beta - \partial_\beta G^A_\alpha - g_s f^{ABC} G^B_\alpha G^C_\beta, \tag{2.2}$$

$D_\mu = \partial_\mu + ig_s G^\alpha_\mu t^\alpha$ is the covariant derivative, g_s is the strong coupling constant, t^α are the Gell-Mann matrices, and f^{ABC} are the structure constants of $SU(3)$. Capital indexes A, B, C run over the 8 degrees of freedom of the gluon fields and index j runs over the quark flavors.

The charge associated to the strong interaction is the color. The color charge is introduced by the fact that quarks, as fermions, must satisfy the Pauli exclusion principle. The validity of the color quantum numbers and the choice of $SU(n_C = 3)_C$ is confirmed by many measurements that directly probe the value of n_C [16].

The non-abelian nature (which allows interaction between particles) of the QCD leads to two important characteristics of the strong interaction: *color confinement* and *asymptotic freedom*. The strong running coupling constant, $\alpha_s = g_s^2/4\pi$ changes with the scale of the interaction. It is assumed to be asymptotically diverging at low energies (large distances) making therefore impossible to produce isolated quarks. When a pair of quarks ($q\overline{q}$) begin to separate from each other, the exchanged gluons interact with each other and the strong coupling constant, α_s, increases. This increasing force either binds the quarks together at low energy scale or it breaks when the energy density of the color field between the quarks is great enough to create an additional $q\overline{q}$ pair. At the end, there are only colorless bound states (hadrons) favored. This situation is explained as color confinement and believed to be related to the process of jet formation. The top quark constitutes an exception in the sense that due to its huge mass, it decays before it *hadronise*.

At high energies (short distances) α_s decreases (Figure 2.1). Then, strong interactions proceed

via color fields of reduced strength, quarks and gluons behave as essentially free particles. This situation is called asymptotic freedom and is totally supported by the results from deep inelastic scattering (DIS) experiments. The asymptotic freedom is the basis of the perturbative approach to calculate QCD observables (pQCD).

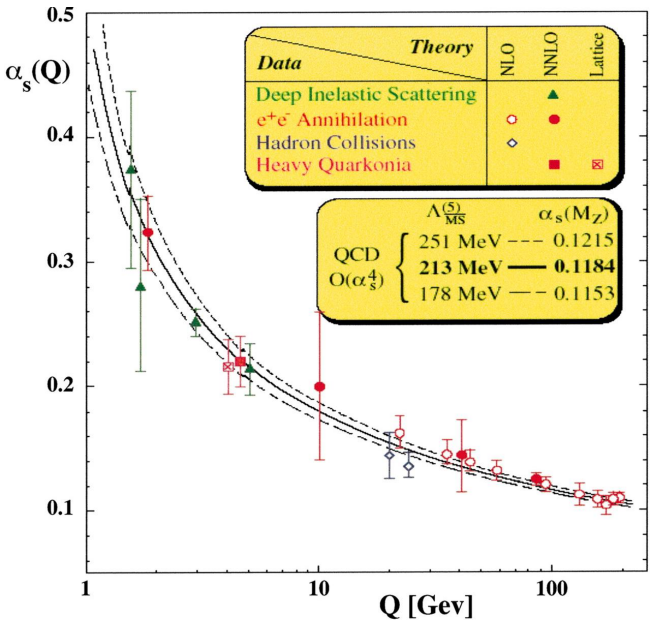

Figure 2.1: The running strong coupling α_s as a function of the energy scale Q.

The amplitude of a strong interaction process at a given momentum transfer scale, q^2, can be parameterized in terms of the coupling strength of the strong force $\alpha_s(q^2)$. A conventional definition of α_s, at leading order can be given as in the following form:

$$\alpha_s(q^2) = \frac{12\pi}{(11n_c - 2n_f)ln\frac{q^2}{\Lambda^2_{QCD}}}, \tag{2.3}$$

where n_c is the number of colors (which is 3 in the SM) and n_f is the number of flavors (which is 6 for large q^2 and smaller for small q^2)) of the quarks with mass less than the energy scale Q and Λ_{QCD} is the QCD scale, which is the only adjustable parameter of QCD and depends on the momentum scale of the interaction. It marks the energy scale at which α_s becomes large and the perturbative approach is no longer valid. For most processes, the measured value of Λ_{QCD} is consistent with 200 MeV. The running strong coupling presented in Eq. 2.3 shows that for large q^2 (short distances), the coupling becomes small (asymptotic freedom)

but at low q^2, the coupling approaches to unity. Hence, high-q^2 processes can be described by perturbative calculations but low-q^2 interactions need to rely on phenomenological model.

2.3 Hadron-Hadron Interactions

Interactions between hadrons do not involve the entire hadron at high momentum transfer, since they are composite particles. Instead of this picture, one of the partons from a hadron interacts with one from the other hadron. Protons as hadrons are made up of three quarks, *uud*, these quarks exchange and radiate gluons that produce virtual quarks within the proton after splitting into $q\overline{q}$ pairs. It is likely that any of these quarks may interact, making $q\overline{q}$ interactions possible. The quark-gluon interactions are more likely to occur at hard scattering processes in proton-proton collisions. Likelihood of a particular interaction occurring is formulated in terms of experimentally observable quantity, cross section. Cross section (σ) is a measure of probability for certain processes occuring during an interaction and given usually in units of barns (b) (1b = $10^{-28}m^2$). The rate of events produced for a particular process depends on the luminosity, $\mathcal{L}$ which is the total number of p-p interactions per unit time and the cross section as in the following expression:

$$Rate = \frac{dN}{dt} = \sigma\mathcal{L}. \tag{2.4}$$

The cross sections for interactions involving hadrons in the initial or the final state are calculated by using QCD formalism. The QCD factorization theorem is used to calculate cross section of any process involved in hard scattering. According to this theorem, a cross section can be factorized into the high-energy (perturbative) processes and the low-energy (non-perturbative) physics effects by introducing a factorization scale. The production cross section of a final state X in a proton-proton collision ($p + p \rightarrow X$) can be calculated as:

$$\sigma(pp \rightarrow X) = \sum_{ij} \int dx_1 dx_2 f_i^p(x_1, \mu_F) f_j^p(x_2, \mu_F) \hat{\sigma_{ij}}(x_1, x_2, \hat{x}, \mu_F^2, \mu_f^2) D(\hat{x}, X, \mu_f^2) \tag{2.5}$$

where:

- p_1 and p_2 are momenta of the colliding partons, x_1 and x_2 are the fractions of the proton momenta (P_1 and P_2) carried by the colliding partons, ($x_{1,2} = p_{1,2}/P_{1,2}$),

- the sum over indices (i and j) extends over the partons of the incoming protons,

- the scales μ_F (factorization scale) and μ_f (fragmentation scale) represent the scales at which the hard scattering process, calculated with pQCD, is separated from the initial and final states respectively,

- the functions $f_i^p(x, \mu_F)$ are known as Parton Distribution Functions (PDFs) and represent the probability that a parton i carries a fraction x of the proton momentum at a scale μ_F,

- $\hat{\sigma_{ij}}(x_1, x_2, \hat{x}, \mu_F^2, \mu_f^2)$ represents the partonic cross section for the production of a state $\hat{x}$ from a pair of partons (i, j),

- the function $D(\hat{x}, X, \mu_f^2)$ represents the probability that the partonic state $\hat{x}$ evolves into the final state X. These type of functions are called fragmentation functions.

The fragmentation functions only appear in the calculation when the final state particle is the result of the fragmentation process of the parton produced in the hard scattering process. For example, prompt photons ($X = \gamma$ in Eq. 2.5) can be produced directly in the hard interaction collision (direct photons) or as a result of the fragmentation of a quark or a gluon into a photon (fragmentation photons). In the first case no fragmentation function is needed for describing the process but in the second case the contribution to the cross section will depend on the related fragmentation functions.

2.4 Prompt Photons

Prompt photons are referred to both direct photons and fragmentation photons in this study. Direct photons are produced primarily in the hadron-hadron interaction as opposed to fragmentation photons created as a result of the collinear fragmentation of a parton [18], [19]. The study of these photons plays an important role at hadron colliders, from probing perturbative QCD (pQCD) to providing evidence for new physics within and beyond the Standard Model.

2.4.1 Production Mechanisms

Direct photons are produced through the quark-gluon compton scattering and the quark-anti-quark annihilation mechanisms at leading order (LO) as shown in Figure 2.2 and Figure 2.3. They are produced in association with collection of partons, so-called jet(s), in the final state. The recoiling photon and the associated jet are mostly back-to-back in the ϕ plane and in equal balance in terms of their energies.

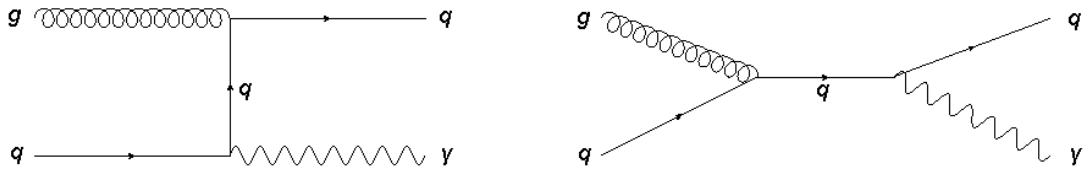

Figure 2.2: The quark-gloun compton scattering process yielding a direct photon at the leading order (LO).

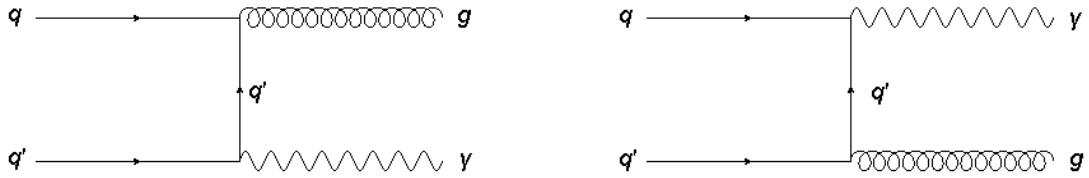

Figure 2.3: The quark-anti-quark annihilation process yielding a direct photon at the leading order (LO), where q′ refers to $\bar{q}$.

There are also higher order contributions to direct photon production mechanism. In these effects, initial or final state radiation is emitted in addition to associated multiple jets in the final state. Figure 2.4 shows some examples of these higher order effects to the direct processes. Such Feynman diagrams consist of the emission of a quark or gluon from one of the initial or final state legs, or through virtual loops where a gluon is emitted and then becomes reabsorbed.

In addition to the direct production mechanisms, prompt photons can also be produced through the bremsstrahlung diagrams as shown in Figure 2.5. These photons are no longer created in the hard scattering but instead through the fragmentation of a final state parton. Fragmentation photons can provide a significant source of isolated photons.

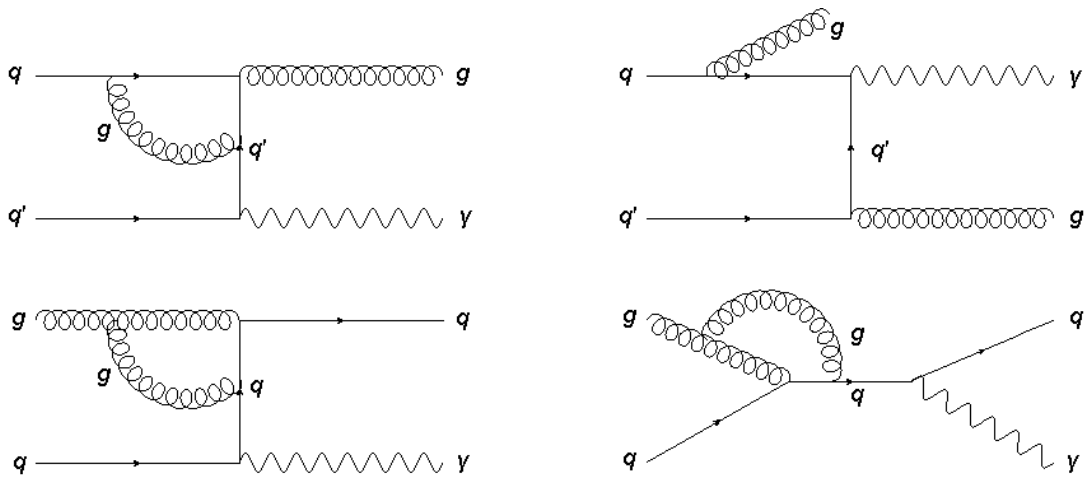

Figure 2.4: Some examples of higher order (NLO) Feynman diagrams contributing to direct photon production, where q′ refers to $\bar{q}$.

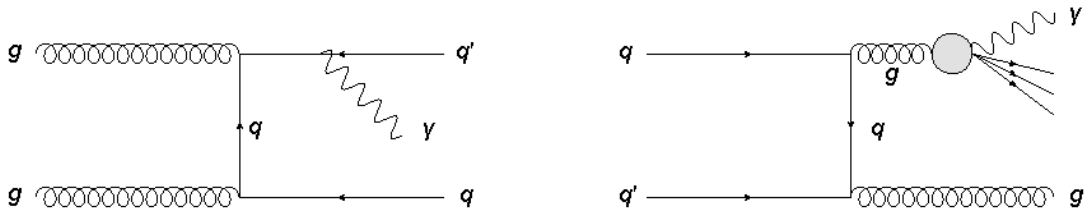

Figure 2.5: Bremsstrahlung Feynman diagrams contributing to fragmentation photon production, where q′ refers to $\bar{q}$.

2.4.2 Isolated Prompt Photons

Calculations of the prompt photon production has a major challenge in identification of the photon signal from an experimental point of view. The task of identification of the prompt photons is complicated by the decays of light mesons such as π^0 and η into two photons. Light mesons are produced in great abundance in hadron-hadron interactions as they decay hadronically, hence all photon measurements must have highly efficient techniques to remove these decays from signal photon production. In fixed target experiments, the cross section of photon is measured generally energies less than 10 GeV. At these low energies, the decays of light mesons into two photons can be resolved by the corresponding calorimeter detectors. Therefore, fixed target experiments can measure the prompt photon cross section inclusively. However, measurements at hadron collider experiments are performed at higher energies, making it difficult to disentangle between the decays and single photon showers in an event-

by-event basis. To get rid of this background, hadron colliders usually measure only the isolated part of the prompt photon cross section.

For the prompt photon production, background is contained within jets as discussed above at hadron colliders. In order to account for background from the decays of light mesons, prompt photons are identified by requiring that they are isolated from other hadrons. For this purpose, it is typically required that the sum of transverse energy deposited within a certain radius around the photon to be below some value. In addition, shower profile of the photon in the electromagnetic calorimeters is used to remove background contribution. These requirements reduce the background, but has the side-effect of reducing the contribution to prompt photons from fragmentation. Prompt photon isolation methods (hence identification) in the CMS experiment will be discussed in great details in the analysis chapters of this thesis.

2.4.3 Associated Jets

The prompt photons are produced in association with partons as already been demonstrated in Figure 2.2 and Figure 2.3 in hadron-hadron interactions. Here, prompt photons from the hard scattering can be detected directly as opposed to the outgoing energetic parton. As mentioned before, partons cannot exit by themselves due to color confinement, they instead combine with quarks and anti-quarks from vacuum to form large collections of hadrons widely known as jets. This formation of jets from outgoing quarks or gluons from a hard scattering interaction is named as *hadronization*. Basically, quarks and gluons from a hard interaction produce showers of hadrons such as $q \rightarrow q + g$, $g \rightarrow g + g$, and $g \rightarrow q + \overline{q}$. According to QCD theory, the branching of gluons is favored over the branching of quarks, and so gluon jets are expected to contain more particles than quark jets in hadron-hadron interactions.

From an experimental point of view, particle detectors measure jets as hadrons formed in a cone. Unlike photons, which are detected using one type of particle detector such as electromagnetic calorimeter, jets are needed to be measured by at least three detectors including the tracking system, electromagnetic calorimeter, and hadronic calorimeter [20]. Furthermore, definition of jet is algorithmic by jet cone algorithms. Conversely, definition of photons are absolute that is cone algorithms are not used to define a photon. Associated jet calorimetric energy resolution in the detector is not as good as photon energy resolution measured only by the electromagnetic calorimeter.

2.4.4 Physics Motivation

The primary means of testing pQCD at hadron colliders is based on measurements on jet observables such as multiplicity, energy, momentum, azimuthal correlations, and so on. Jets are by far the most abundant events produced at hadron colliders and hence they provide great statistical power out to very high transverse energies in geometrical acceptance of detector. However, definition of jet is algorithmic and parametric, not absolute. Jet algorithms are still in development and the current jet algorithm favored by the LHC experiments (called the $anti - k_t$ algorithm [21]) is roughly four years old.

Prompt photons provide another way of directly probing the hard scattering process, and thus pQCD. The direct production of photons, through either an annihilation or compton process at the leading order provides an extremely clean probe of hadron-hadron interactions. The photon emerges from the hard interaction without any effects from formation of jets, making the photon well defined at both the parton and particle levels. Moreover, the existence of the compton diagram means that their production rate is sensitive to the gluon content of the proton [22]. This value is constrained at low $x_T = 2E_T/\sqrt{s}$ by measurements of deep inelastic scattering and photo-production experiments. The LHC will probe a much broader range of x_T values, with the capability of providing valuable input to global fits of parton distribution functions.

Another motivation for the study of prompt photon production at the hadron colliders is searches for low mass Higgs boson. The Higgs boson does not couple directly to photons, or to any massless particles. However, it can produce two photons (diphoton) in the final state via an intermediate fermion or vector-boson loop. These loops contain other particles to which the Higgs does couples strongly. In this Higgs production mechanism, there is a colored loop of bottom or top quarks that connects to the colliding gluons. In a Higgs decay ($H \rightarrow \gamma\gamma$), there is an electrically charged loop of W bosons or top quarks that connects to the photons. Such a diphoton signature is extremely clean, and represents one of the most powerful channels for observing the Higgs around the low mass of 120 GeV.

Looking beyond the Standard Model, measurement of prompt photons is very crucial as well since they are present in the final states of many signatures of new physics. The first example is the searches for The Randall-Sundrum (RS) model [23]. This model was proposed to solve

the hierarchy problem, by claiming that the observable universe is a four-dimensional brane contained within a five-dimensional bulk space. In this model (and in other models including large extra dimensions) gravity operates in five dimensions while the other known forces only operate within the brane. The warped geometry of the bulk space leads to a tower of excited resonances of the gravitational field, the first of which corresponds to the normally massless graviton, with charge=0 and spin=2. The massive graviton can then decay into Standard Model fermions or bosons, including into pairs of photons.

In parallel to the RS model, other theories with large extra dimensions predict the production of diphoton pairs. In these models, photon pairs are in combination with missing transverse energy due to the escape of massless gravitons in a cascade decay from the first Kaluza-Klein (KK) excitation [24]. In this case, the signal would appear as an excess in the transverse energy distribution for events containing two energetic photons.

Measurement of photon production is also important for Supersymmetry (SUSY) searches. One model of symmetry breaking within SUSY can be given as an example which is known as gauge mediated supersymmetry breaking (GMSB) (see, for example [25]). In this sub-model, the lightest supersymmetric particle (LSP) is the gravitino, the supersymmetric partner to the graviton. The lightest neutralino then decays to a gravitino and a photon. In production of a pair of heavy supersymmetric particles, they can decay in cascades to neutralinos (which then decay), producing jets, two photons, and two gravitinos.

This is not the complete list of new physics that can be probed with photons, but the ability of photon physics to play a role in a broad range of theoretical models that extend beyond the Standard Model is discussed with examples. Prompt photon production at high energies constitutes not only signal but also an irreducible background for important searches, such as $H \rightarrow \gamma\gamma$, RS Graviton, extra-dimensions or SUSY (as discussed before) with energetic photons.

Finally, measurement of prompt photons has one significant advantage over jets such as they are well measured with the electromagnetic calorimeters rather than hadronic calorimeters, resulting in an improved photon energy resolution over jet energy resolution.

2.5 Predictive Tools

There are several so-called generator tools available in the market for making theoretical predictions of prompt photon production rates (and hence cross sections) from hadron-hadron interactions. The most significant and widely used generator programs in the LHC experiments (the ones used in the analysis parts of this thesis are among those) are outlined in this section.

Mainly, Monte Carlo programs can be divided into two categories. First type (such as PYTHIA) includes all the steps in event generation from the parton distribution functions to the final state particles. The hard interaction portion of the cross section calculation for $2 \rightarrow 2$ processes is based on pQCD to leading order in this type of generators. The second type of generator programs such as MadGraph, concentrate on more complex matrix elements and they can produce multiple parton final states. These generators are more reliable for producing events with multiple jets with the correct kinematics and cross sections. They usually do not include the hadronization process and therefore are often interfaced with the other, more complete generators to get full event information.

2.5.1 Pythia

The *PYTHIA* event generator is a leading-order (LO) parton showering Monte Carlo program [27], and is widely used in collider physics to generate a broad range of QCD and EWK events in addition to events for new physics searches. It relies on QED radiation emitted off quarks in the initial state (ISR) and final state (FSR). The hadronization of partons is modeled with the Lund string model in PYTHIA [28]. The fragmentation component is modeled as FSR, without any knowledge of the fragmentation functions.

Generated events which are from PYTHIA version 6.4 simulations are used in this thesis for the signal photon production. The sub-processes simulated for the signal samples used in this analysis are dominated by $q\bar{q} \rightarrow g\gamma$ and $qg \rightarrow q\gamma$. Both the signal and background hard scattering sub-processes simulated are summarized in Table 2.3.

Table 2.3: Signal and background sub-processes (decay channels) simulated by the PYTHIA generator program of version 6.4.

Signal Processes	Background Processes
$q_i\bar{q}_i \rightarrow g\gamma$	$q_i q_j \rightarrow q_i q_j$
$f_i\bar{f}_i \rightarrow \gamma\gamma$	$q_i\bar{q}_i \rightarrow q_k\bar{q}_k$
$q_i g \rightarrow q_i\gamma$	$q_i\bar{q}_i \rightarrow gg$
$gg \rightarrow \gamma\gamma$	$q_i g \rightarrow q_i g$
$gg \rightarrow g\gamma$	$gg \rightarrow q_k\bar{q}_k$

2.5.2 Madgraph

The *MadGraph* matrix element software has capability of generating amplitudes for any process up to 9 final state particles in any process [29]. MadGraph automatically calculates amplitudes for corresponding subprocesses and produces mappings for the integration over phase space. Events are passed as parton level files in the standard format called Les Houches format (LHE files). These parton level files are then passed to PYTHIA, which takes care of the parton showering and the fragmentation part before passing the event to detector simulation.

The MadGraph simulation samples used in this thesis contain tree level calculations for parton level processes of up to $2 \rightarrow 4$ (such as, up to $pp \rightarrow \gamma + 3$ jets). The samples were produced with cuts on the scalar sum of the transverse momenta of the outgoing jets, $H_T = \sum p_T$.

2.5.3 Sherpa

The *SHERPA* is a relatively new multi-purpose leading order event generator [30]. Treatment of prompt photon production in SHERPA differs from that of PYTHIA by modeling the fragmentation functions as a part of the parton shower. This makes predictions with next-to-leading order accuracy while still providing parton-shower merging that can be used to study the detector response.

SHERPA has in-built matrix-element generators, AMEGIC++ & Comix and a phase-space generator Phasic. Thus, SHERPA can be used both as a cross section integrator as well as a parton-level event generator. For better description of multijet processes Comix is used within SHERPA program.

2.5.4 Jetphox

The *JETPHOX* generator [26] is a next-to-leading order (NLO) Monte Carlo program. It calculates the double differential cross sections $d^2\sigma/(dE_T^\gamma d\eta^\gamma)$ for A+B $\rightarrow$ C+D+X processes, where C and/or D can be photons, and A and B are partons with properties given by parton distribution functions. This generator program takes the definitions of the fragmentation functions, the parton distribution functions, and some common experimental constraints as inputs. Modelling of the leading order direct photon contributions in JETPHOX is almost the same with PYTHIA.

The default capability of JETPHOX is to calculate cross section of inclusive prompt photon production, regardless of the isolation of the photon from additional hadronic activity (background). However in JETPHOX, the isolation energy of prompt photons at the parton level can be estimated, and therefore isolated prompt photon cross section can be obtained.

Measurements of prompt photon, photon+jet, and diphoton production at the LHC include comparisons to SHERPA as well as JETPHOX. Prompt photon simulation with SHERPA has been recently validated by its authors. There is one other leading order parton shower Monte Carlo program which is HERWIG [31] in addition to those given above. HERWIG is different from PYTHIA in terms of its parton showering model (which is angle ordered instead of p_T ordered), its treatment of hadronization (the cluster model instead of the Lund string model), and its treatment of the underlying event.

CHAPTER 3

EXPERIMENTAL SETUP

The European Organization for Nuclear Research (CERN) is an international organization whose primary purpose is to operate the world's largest and highest-energy particle accelerator, the Large Hadron Collider (LHC). The LHC collides two beams of protons or lead ions head-on at very high energies. Created particles in the collisions are recorded by the special detectors in a number of experiments dedicated to the LHC. The Compact Muon Solenoid (CMS) experiment is one of two large general-purpose particle physics detectors constructed at the LHC. The LHC experiments including the CMS target to test various predictions of the High Energy Physics Standard Model and to search for new physics models.

3.1 Large Hadron Collider

The Large Hadron Collider (LHC) uses an underground ring which is between 45 m and 170 m below the surface of the earth and 27 km in circumference. The LHC ring housed an electron-positron collider (LEP) originally, constructed at the French-Swiss border near Geneva, Switzerland. The LHC project is carried out by the European Organisation for Nuclear Research (CERN) and designed to collide opposing proton beams at a centre-of-mass energy ($\sqrt{s}$) of 14 TeV with an unprecedented luminosity of $10^{34} cm^{-2} s^{-1}$. It can also collide heavy lead (Pb) ions at an energy of 2.8 TeV per nucleon with a peak luminosity of $10^{27} cm^{-2} s^{-1}$.

There are six experiments conducted at the LHC, each experiment is distinct by means of its unique particle detector. The two large experiments, ATLAS (A Toroidal Lhc ApparatuS) and CMS (Compact Muon Solenoid), are based on general-purpose detectors to analyze

abundance of particles produced by the collisions in the accelerator. They are designed to investigate a wide range of physics possible including precision measurements of the Standard Model and searches for the Higgs boson and Supersymmetry. Having two independently designed detectors is essential for cross-confirmation of any new discoveries. The two medium-size experiments, ALICE (A Large Ion Collider Experiment) and LHCb (Large Hadron Collider Beauty), have specialized detectors for analysing the LHC collisions in relation to heavy ion and beauty quark physics, respectively. The last two experiments, TOTEM (TOTal Elastic and diffractive cross section Measurement) and LHCf (Large Hadron Collider Forward), are much smaller in size. They are designed to focus on measurements of cross section and forward particles (protons or heavy ions). Those particles that brush past each other as the beams collide, rather than meeting head-on are called forward particles. The ATLAS, CMS, ALICE, and LHCb detectors are installed in four huge underground caverns located around the ring of the LHC. The four experimental sites along the LHC ring are shown in Figure 3.1. The CMS detector is situated approximately 100 m underground at P5 (Interaction Point in 5th LHC octant). The detectors used by the TOTEM experiment are located near the CMS detector, whereas those used by LHCf are near the ATLAS detector.

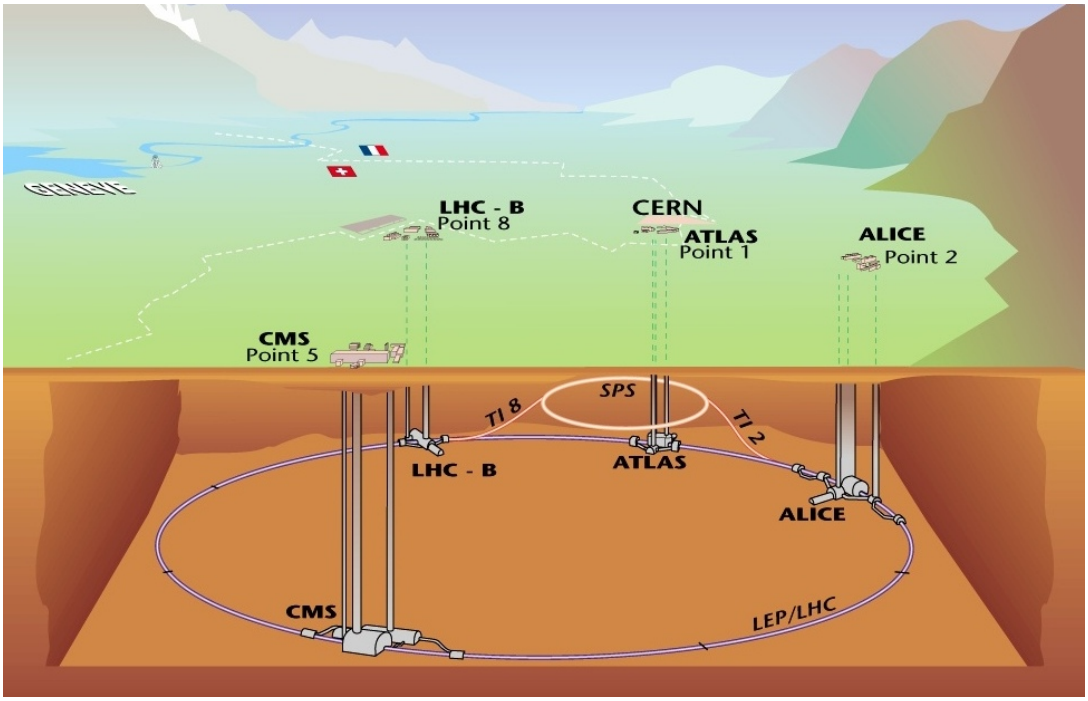

Figure 3.1: Overall view of the LHC experiments above and underground seen from north side of the ring.

The LHC complex is composed of a number of particle accelerator units as shown in Figure 3.2 ([33]). Each accelerator unit injects the beam into the next one in the sequence, which takes over to bring the beam to an even higher energy. Protons are obtained by removing orbiting electrons from hydrogen atoms and boosted to 50 MeV by one of the two linear accelerators, LINAC2 (LINAC3 for lead ions) which uses radio frequency (RF) cavities. They are then injected into the Proton Synchrotron Booster (PSB) which ramps up the beam energy to 1.4 GeV, also by making use of RF cavities. The beam is then fed to the Proton Synchrotron (PS) where it is accelerated to 25 GeV. Protons are then sent to the Super Proton Synchrotron (SPS) and accelerated to 450 GeV. They are finally transferred to the LHC ring both in a clockwise and an anti-clockwise directions where they are accelerated for 20 minutes to their nominal energy of 3.5 TeV per beam (7 TeV is the design energy which will be achieved in later years) [32].

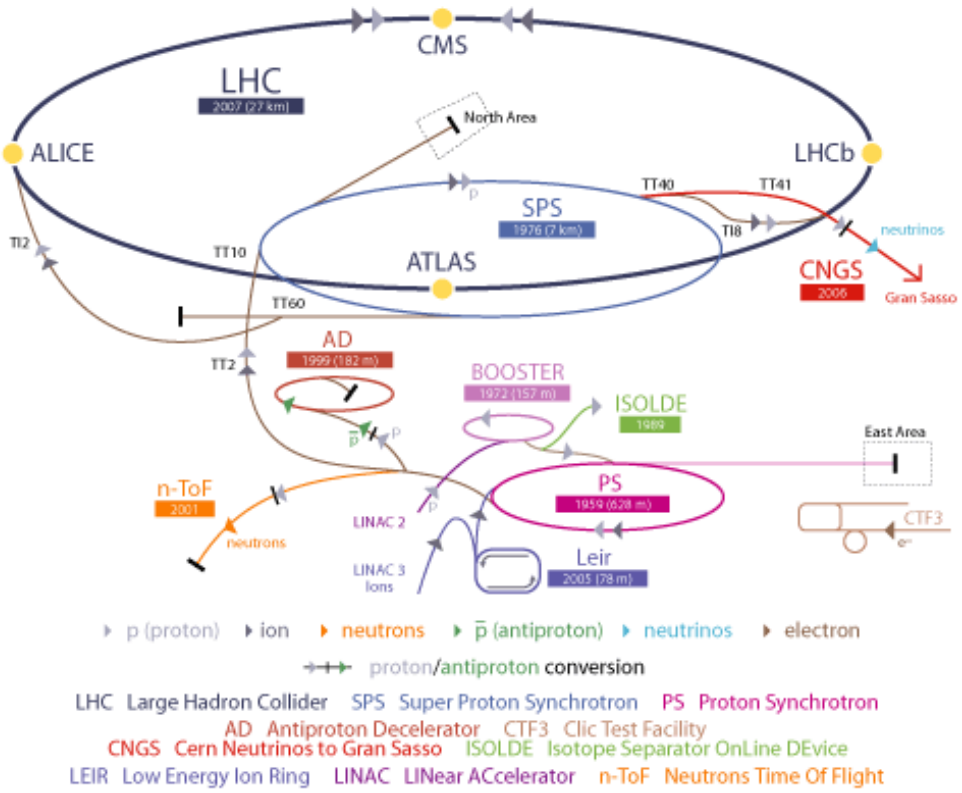

Figure 3.2: The LHC accelerator complex at CERN.

Inside the LHC ring, the beams circulate in opposite directions in two separate beam pipes. Special dipole magnets are utilized to direct the beams in opposing directions. The LHC

dipole magnets have a compact, twin-bore design to fit into the existing 3.7 m tunnel width. The magnets are made of superconducting niobium titanium (NbTi) and rest inside a vessel filled with liquid Helium to enable operation at temperature below 2 K. They have two beam channels in the same structure and cooling system which are surrounded by shells of superconducting coils creating the magnetic field which keeps the beams to follow a circular path. The cross section of the LHC dipole magnet is demonstrated in Figure 3.3 ([34]). In short, 1232 dipole magnets, each with a field of 8.3 Tesla, are used to bend the proton beams around the LHC ring. In addition, 392 quadrupole magnets are also used to keep the beams focused, in order to maximize the probability of interaction between the particles at the intersection points, where the two beams will cross.

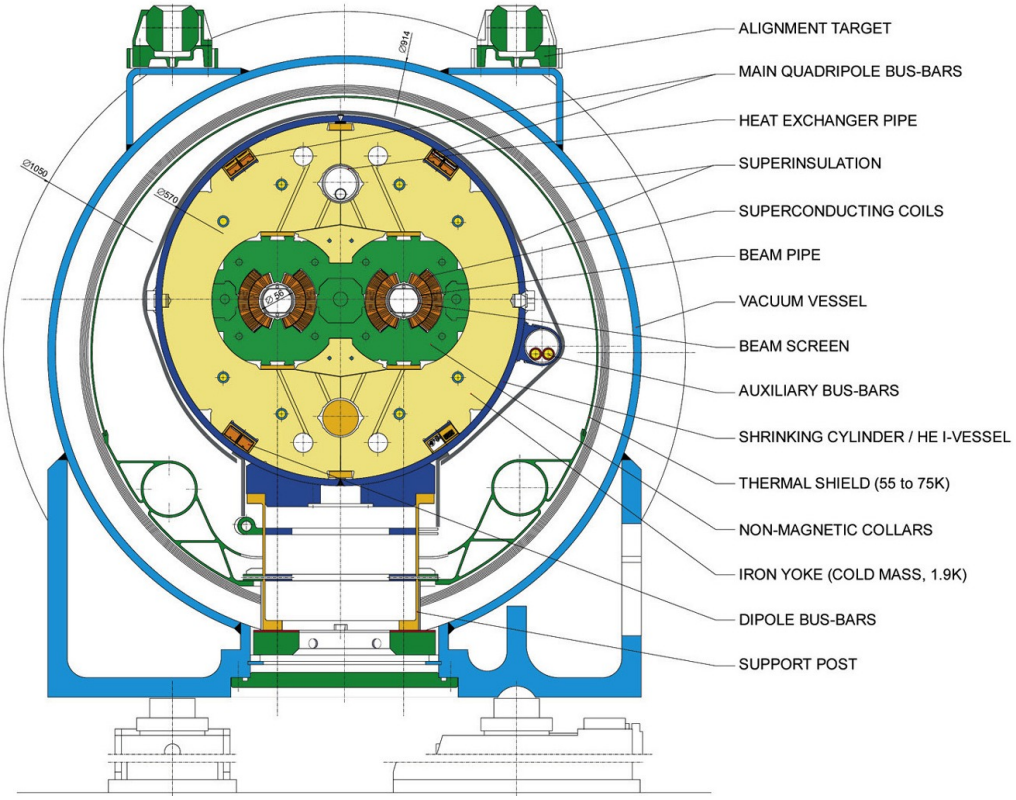

Figure 3.3: Schematic view of the LHC dipole magnet.

Particles inside the beam are grouped into cylindrical bunches with a diameter and a length of approximately 16 μm and 8 cm at the interaction point, respectively. The LHC filling scheme for the proton beam assumes a bunch separation of 25 ns. Trains of 72 occupied and 12 empty bunches, each made in one PS cycle, are used to fill the SPS with an injection kicker

gap of 220 ns. A longer gap of 950 ns is allocated after 3-4 consecutive trains for injection of beam into the LHC. Almost 4 cycles of SPS fill the entire LHC, requiring nearly 3 minutes. Longer gaps between bunch trains and one missing train at the end of the LHC cycle allow for the synchronisation, resetting of electronics, and obtaining calibration data in the absence of collision events. There are 3564 spaces available during each LHC cycle, 2808 are filled with actual particle bunches and each bunch contains 10^{11} protons [35] nominally. Some of the LHC proton beam parameters are further summarised in Table 3.1.

Table 3.1: Some of the LHC beam parameters for nominal proton runs as of late 2011.

Beam Parameters	Values
Luminosity	$10^{33}\ cm^{-2}s^{-1}$
Design Luminosity	$10^{34}\ cm^{-2}s^{-1}$
Energy at collision	7 TeV
Design energy at collision	14 TeV
Magnetic field	4 T
Number of bunches	2808
Number of particles per bunch	1.15×10^{11}
Bunch spacing	25 ns
Number of collision events per beam crossing	~ 20

The production rate of collision events is proportional to luminosity delivered by the LHC machine. Luminosity is given roughly as a function of the number of protons in each bunch (N_b), the number of bunches in the beams (n_b), the revolution frequency of the beams (f), and the size of the beams (effective collision area, A_{eff}) as

$$\mathcal{L} = \frac{N_b^2 \cdot n_b \cdot f}{A_{eff}} \tag{3.1}$$

for two beams colliding with the same parameters. Given the luminosity, *Integrated luminosity* is calculated as $L = \int \mathcal{L} dt$, which corresponds directly to the number of collisions collected in a given time interval. As discussed earlier, quadrupole magnets are used to focus the beams to 16.7 μm. The smaller beam size implies a smaller effective collision area, A_{eff}, and a higher luminosity. The revolution frequency is known to be 11.246 kHz. The number of particles, number of bunches, and the frequency of revolution quantities which are well known. The effective cross-sectional area of the beam at a collision point depends on the longitudinal and lateral spread of the proton bunches and can be expressed as:

$$A_{eff} = \frac{4\pi \cdot \epsilon_n \cdot \beta^\star}{\gamma_r \cdot F} \tag{3.2}$$

where ϵ_n is the normalized transverse beam emittance (which is nominally 3.5 μm), $\beta^\star$ is the beta function (a measure of the beam width which is nominally 0.55 m) at the collision point, γ_r is the Lorentz gamma factor, and F is a factor that accounts for the fact that the beams do not collide head-on, but rather cross with some angle.

Furthermore, there is one another important relation which illustrates the connection between luminosity and the physics process in terms of the observable quantity, cross section σ:

$$\sigma = \frac{N_{obs}}{A \cdot \epsilon \cdot L}, \tag{3.3}$$

where σ is the cross section, N_{obs} is the number of observed events, A is the acceptance (fraction of the events that can be found by the detector), ϵ is the efficiency of the selection, and L is the integrated luminosity.

3.1.1 Running Conditions in 2011

In 2011, the LHC delivered 5.72 fb^{-1} total integrated luminosity of data for the experiments and the CMS detector recorded the amount of 5.20 fb^{-1} data with high efficiency as shown in Figure 3.4. This amount of data collected is far beyond the expectations that were before the start of the run. The run taking started in the middle of March and lasted up until to the end of October, 2011. Some important records in the proton-proton running are summarized below:

- Peak instantaneous luminosity: 3.55 x 10^{33} $cm^{-2}s^{-1}$ in fill 2256,
- Delivered luminosity in one fill: 123 pb^{-1} in fill 2219,
- Maximum luminosity in one day: 136 pb^{-1},
- Maximum luminosity delivered in one week: 538 pb^{-1},
- Maximum luminosity delivered in one month: 1614 pb^{-1}.

The LHC was switched in stable beams for 1364 hours total in this year. After the end of proton running, the CMS entered into heavy ion running phase which was followed by the machine development stage. In 2012, the LHC is aimed to provide 8 TeV collisions data for CMS with a bunch spacing of 50 ns.

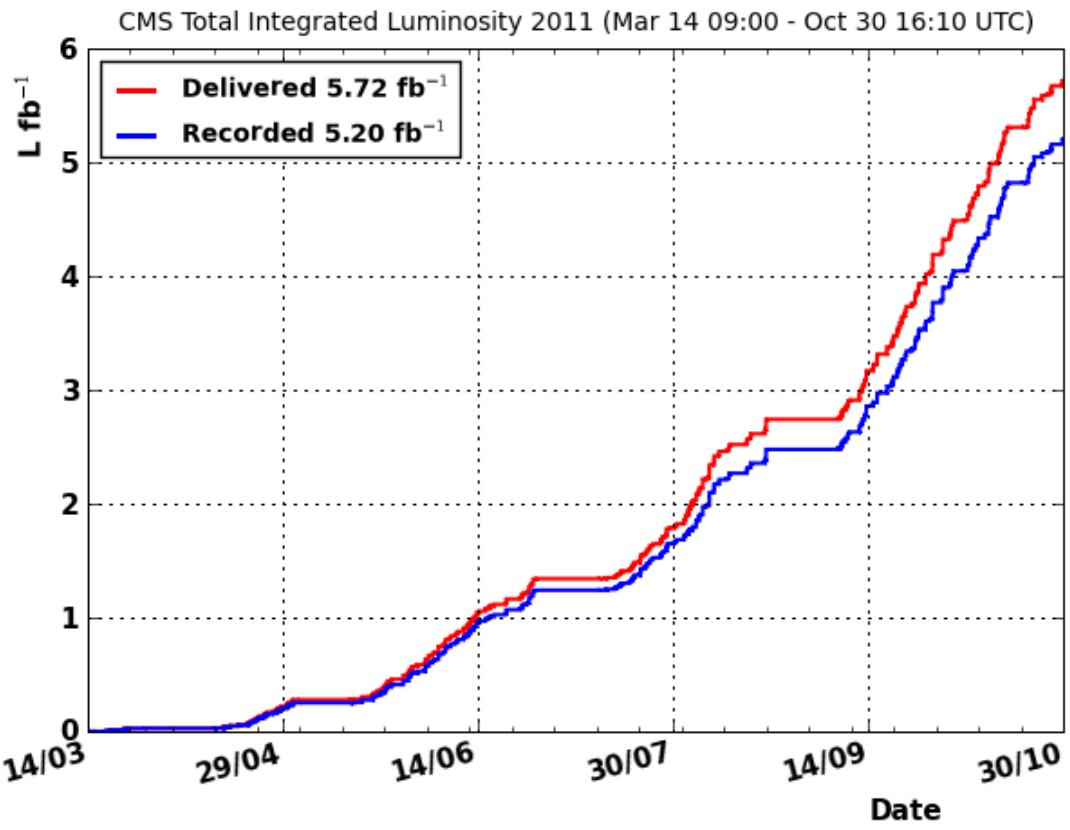

Figure 3.4: Total amount of integrated luminosity delivered by the LHC and recorded by the CMS during the 2011 proton-proton collisions running.

3.2 Compact Muon Solenoid

Located at one of the four interaction points where two proton beams are diverted into collisions, the Compact Muon Solenoid (CMS) is one of two general purpose detectors to search for the Higgs boson and new physics signatures at the LHC (Figure 3.5) [36], [37]. CMS is composed of several sub-detectors, each with specific purpose. The inner-most part of the detector is the tracker, followed by the electromagnetic calorimeter and the hadronic calorimeter. These three sub-detectors are within a solenoid magnet, which is surrounded by muon cham-

bers. The location of the calorimeters inside the solenoid creates a much more compact design (hence the name), and better detection of particles in the calorimeters since they do not first traverse the material of the magnet. The CMS magnet is a NbTi solenoid that cretaes a large magnetic field of 3.8 T. The field bends charged particles in the tracker to enable measurement of their momentum. An iron yoke interspersed with the muon chambers around the solenoid carries a return field of 2 T, making further momentum measurements of muons possible. The main features of the CMS sub-detectors can be summarised as good charged particle tracking, good muon identification, good electromagnetic and hadronic energy resolutions, and good missing transverse energy and jet resolutions.

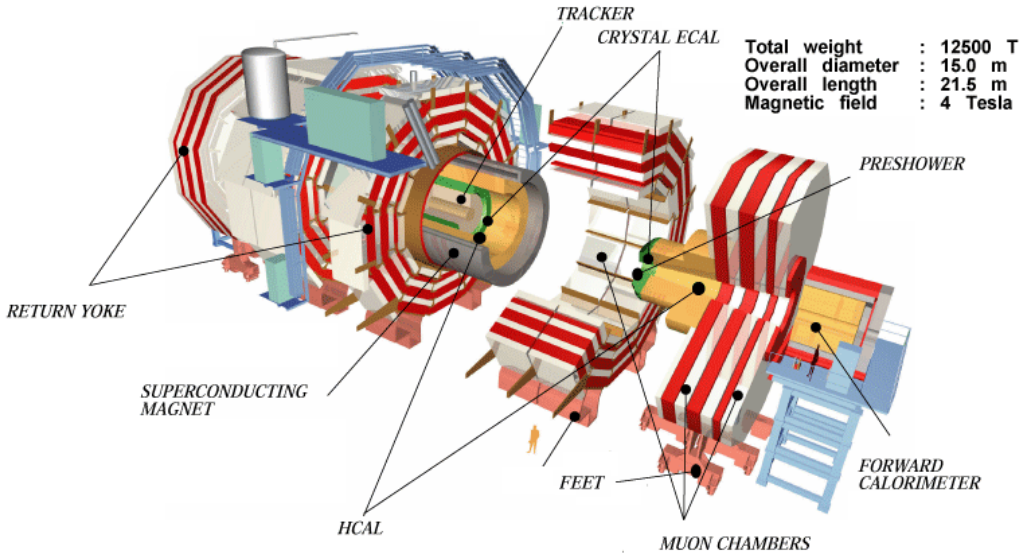

Figure 3.5: An illustration of the major sub-components of the CMS detector.

In CMS, all electrically charged particles leave a trail of ionization in the central tracker, photons and electrons deposit all of their energies in the electromagnetic calorimeter while hadrons leave their energies in the hadronic calorimeter. Muons have low interaction rate in the calorimeters, which are identified by the muon tracking system at high radius; the magnetic field acts in opposite direction outside the solenoid hence the muon track points back to its vertex. Thus, the particles are detected by means of their interactions with different sub-detectors. Figure 3.6 shows a slice of the detector with the main sub-components labeled along with particles detected.

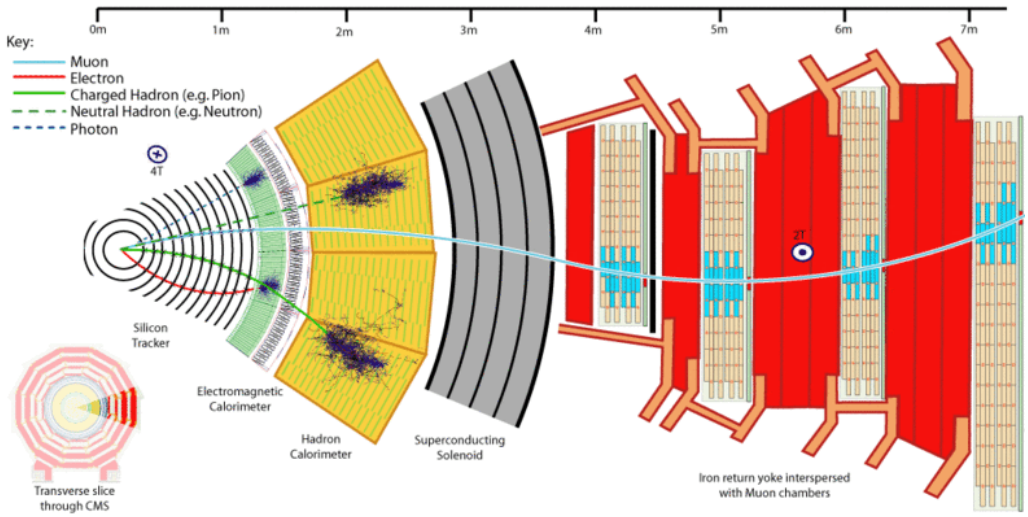

Figure 3.6: An illustration of transverse slice of CMS detector with passages of key particles through the sub-components.

3.2.1 Coordinate System

CMS is cylindrical in shape with a diameter of 14.6 m and length of 21.6 m. It is divided into three major sections: the barrel, two endcaps, and two forward regions. CMS uses a right-handed Cartesian coordinate system with the origin at the nominal interaction point. The x-direction points radially inward towards the center of the LHC ring, the y-axis points vertically up, and the z-axis points west along the beam direction. The radial distance, r is defined in the x-y (transverse) plane. The azimuthal angle, ϕ is measured from the x-axis in the x-y plane. The polar angle, θ is measured between the line connecting the coordinate to the interaction point and the z-axis, but ”pseudorapidity“, η is used almost always, since the distribution of particles in η is almost flat. Pseudorapidity is defined as

$$\eta = -ln[tan(\frac{\theta}{2})]. \tag{3.4}$$

For the calorimeters, ”barrel” refers to the region of $|\eta| < 1.4442$, “endcap“ to the region $|\eta| > 1.566$ and $|\eta| < 3.0$. The gap between the barrel and endcap is used for cables and has fewer detecting components. The hadron forward region covers $3.0 < |\eta| < 5.0$. The tracker system extends to $|\eta| < 2.5$. Additionally, distance in the $\eta - \phi$ plane is another useful coordinate quantity which is calculated as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{3.5}$$

As a general overview, momentum and energy are measured transverse to the beam, p_T and E_T respectively, which are computed from their x and y components (i.e. $E_T^2 = E_x^2 + E_y^2$, for massless particles). The energy imbalance is measured in the transverse plane that is denoted by E_T^{miss} (missing transverse energy). The transverse component of momentum is calculated as $p_T = psin\theta$.

3.2.2 Tracker

The closest sub-detector to the beam interaction point is the tracker system [38], [39]. The tracker uses millions of silicon sensors to measure the trajectory (paths and hence momentum) of charged particles such as high energy muons, electrons, and hadrons up to $|\eta| < 2.5$ through the magnetic field. In the tracker, two different silicon technologies are used: pixels and strips. Nearest to the beam pipe, the high particle flux requires pixel sensors whereas further out the occupancy drops sufficiently to allow silicon strip sensors to be used. When charged particles travel through the tracker, they cause ionization currents, then the pixels and strips generate electric signals which are amplified and detected. The idea of using silicon sensors is to make it possible to have thin layers, and thus short response times and excellent position resolution. The location and layout of the entire tracker is shown in Figure 3.7 ([40]).

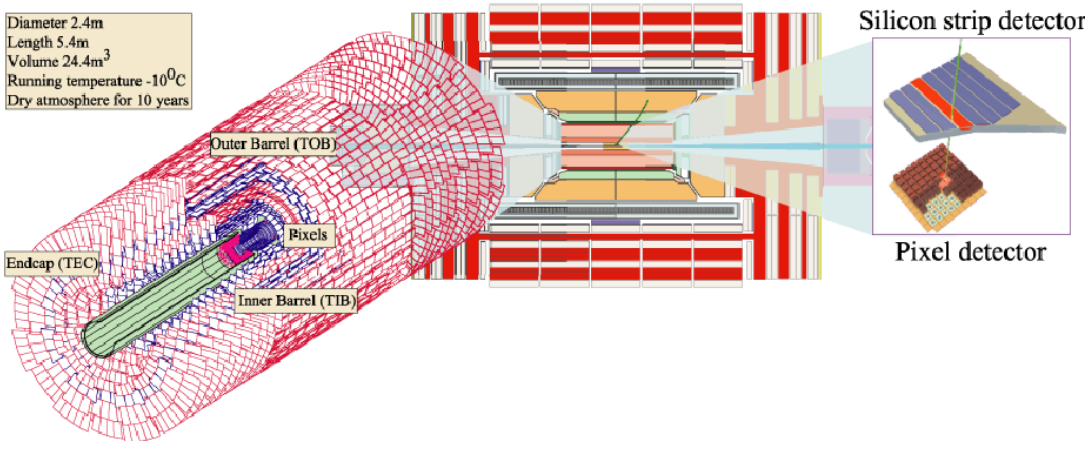

Figure 3.7: Layout of the CMS tracker with sub-components, pixels, and strips.

The pixel detector contains 65 million pixels, allowing it to track the paths of short-lived particles emerging from the collision with high accuracy. It consists of 3 barrel layers with 2 pairs of endcap disks and provides two hit coverage up to $|\eta| = 2.5$. The barrel layers are located at radii of 4 cm, 7 cm and 11 cm and are 53 cm long. The two endcap disks are located at $|z| = 34.5$ cm and 46.5 cm with radii of 6 cm to 15 cm. Each layer is spilt into tiny segments, each a little silicon sensor, 100 μm by 150 μm, about two hairs widths. When a charged particle passes through it distributes enough energy for electrons to be ejected from the silicon atoms, and thus creating electron-hole pairs. Each pixel uses an electric current to collect these charges on the surface as a small electric signal.

In order to maximize the vertex resolution, each pixel has dimensions of about 100 μm $\times$ 150 μm $\times$ 250 μm. The pixel spatial resolution is increased by using analogue signal interpolation of the charge sharing induced by the large Lorentz drift in the magnetic field. Thus the barrel pixel layers are collinear to the beam and the endcaps are arranged in a turbine-like geometry with blades rotated by 20^o. The resolution is 10 μm in $r - \phi$ and 20 μm in $r - z$.

Surrounding the pixels are 11.4 million silicon strips covering the range of $|\eta| < 2.5$. The strip tracker is split into 2 systems, inner and outer, each of which has barrel and endcap sections. There are ten strips layers in the barrel covering a radius of 20 cm to 110 cm from the beam axis. The barrel is divided into the Tracker Inner Barrel (TIB) for $|r| < 65$ cm and Tracker Outer Barrel (TOB) for $65 < |r| < 110$ cm. The TIB and TOB have a resolution of 230 μm and 530 μm respectively and are oriented along the beam axis. The Tracker Inner Disks (TID) is next to the TIB in filling the gap between the barrel and endcap. The Tracker EndCap (TEC) has 9 layers from 124 cm $< |z| <$ 282 cm. Both the TID and TEC are oriented radially to the beam direction.

The tracker transverse momentum (p_T) resolution (σ_{p_T}/p_T) is given by the following useful expressions:

$$\frac{\sigma_{p_T}}{p_T} \approx (15 p_T \oplus 0.5)\%(TeV), |\eta| < 1.6, \tag{3.6}$$

$$\frac{\sigma_{p_T}}{p_T} \approx (60 p_T \oplus 0.5)\%(TeV), |\eta| < 2.5. \tag{3.7}$$

3.2.3 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) surrounds the tracker system which is in place to measure energies of particles [41], [42]. As of particular interest, energies of electrons and photons (they also appear in decay channels of Higgs boson and new physics searches) are measured. CMS has chosen 76,832 scintillating lead tungstate ($PbWO_4$) crystals which produce electromagnetic showers and transmit their energy (light). As electrons pass through the crystal they release bremsstrahlung photons that convert to electron pairs, resulting in a shower of light collected by vacuum photodiodes (barrel) or vacuum phototriodes (endcap). Lead tungstate crystals are radiation hard, have short radiation lengths ($X_0 = 0.89$ cm) and are fast that is 80 percent of light is emitted within 25 ns. The ECAL is sectioned into the barrel (EB), the endcap (EE), and a preshower (PS) detector in the endcaps. A drawing of the ECAL and sub-components is shown in Figure 3.8 ([40]).

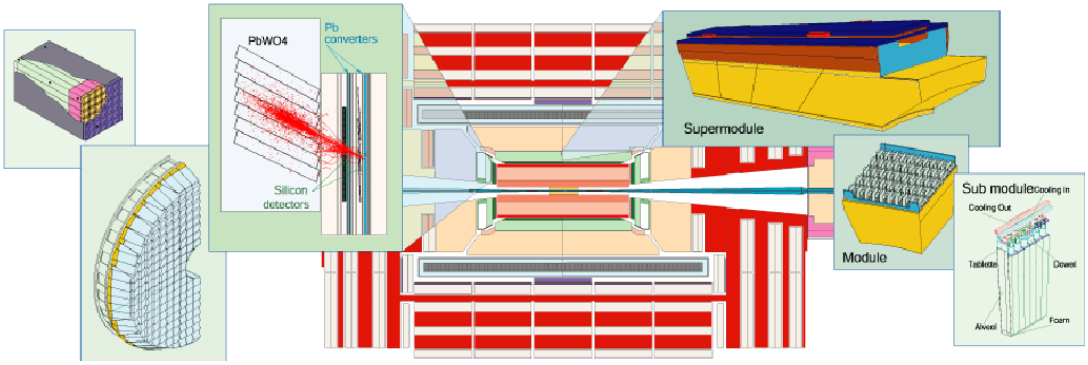

Figure 3.8: Drawing of the CMS ECAL sub-components, demonstrating the organization of the crystals.

The ECAL barrel (EB) begins at a radius of 129 cm and covers the range $|\eta| < 1.4442$. 61,200 crystals are grouped together to form one of 36 identical *supermodules*. Each supermodule covers half of the barrel length. The crystals are mounted 3^o off axis from the nominal vertex position to avoid energy leakage between crystals. Each crystal measures $22 \times 22 \times 230\ mm^3$ and covers an area of $\Delta\eta \times \Delta\phi = 0.0174 \times 0.0174$. These crystals are read out by silicon avalanche photodiodes (APDs).

The ECAL endcaps (EE) are located 314 cm from the vertex and cover the range $1.56 < |\eta| < 3.0$. Each endcap is constructed from two *Dees* consisting of semi-circular aluminium plates

mounting crystals in groups of 5×5 crystals, known as *supercrystals*. The endcaps crystals are tilted off axis in an x-y grid. There are 7,234 crystals of dimensions $28.6 \times 28.6 \times 220$ mm^3 in each endcap. Vacuum *phototriodes* are used for the endcap readout as they are more radiation hard than APDs.

A preshower detector (PS) is comprised of two planes of lead absorber (Pb-Si), followed by silicon strip detectors, which is placed in the endcaps, covering $1.65 < |\eta| < 2.5$. These are used to help identification of neutral pions (such as rejecting photons from π^0 decays) in the endcap region where the average pion energy is high enough to make resolving individual photons difficult resulted by the calorimeter granularity.

The energy (E) resolution (σ/E) of the ECAL is excellent which is reported as

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{41.5MeV}{E}\right)^2 + (0.3\%)^2. \tag{3.8}$$

3.2.4 Hadronic Calorimeter

Surrounding the ECAL is the hadronic calorimeter (HCAL) which measures energies, positions, and arrival times of hadrons such as proton, neutron, kaon, and pion in the form of jets [43]. It provides indirect measurements of non-interacting uncharged particles such as neutrinos hence the missing energy. The HCAL was placed inside the magnet coil in order to minimize non-Gaussian energy resolution tails and to provide good containment. This required a compact absorber and left little room for the active medium. Brass was chosen for the absorber because it is non-magnetic and has a short interaction length (λ). Tile/fiber technology was chosen for the active medium together with wavelength-shifting fibre readouts. A drawing of the HCAL and sub-components is shown in Figure 3.9 ([40]).

The barrel section (HB) spans the region $|\eta| < 1.305$ and is read out in towers of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ in a single longitudinal sampling. The HB has 15 brass plates comprising 6.5 λ thus additional layers, known as the hadronic outer (HO), are placed behind the coil to act as a tail catcher. HO uses the same tower geometry as the barrel and extends the barrel HCAL depth to over 10 λ.

Hadron endcap (HE) disks are located either side of the solenoid coil and span the region

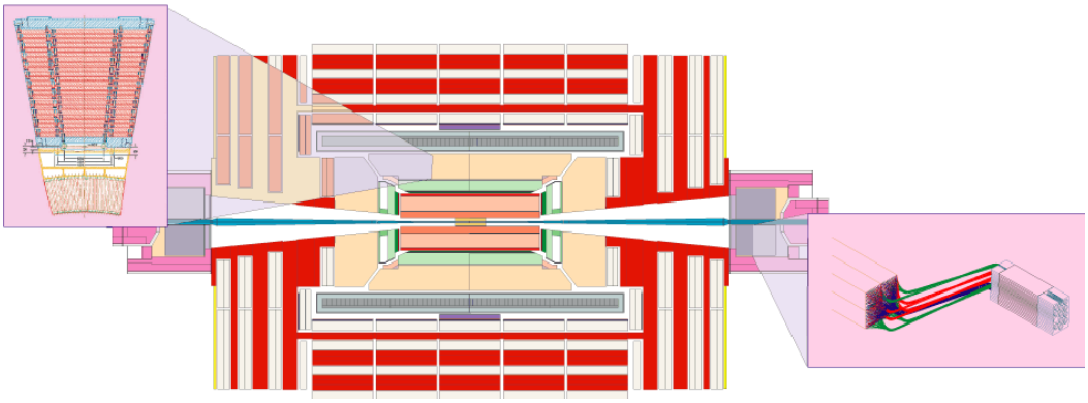

Figure 3.9: Drawing of the CMS HCAL sub-components. On the left, flowerthe brass and scintillator of the HB/HE are given. To the right the steel and quartz of the HF secor are demonstrated.

$1.305 < |\eta| < 3.0$ with towers varying in size from $\Delta\eta \times \Delta\phi = 0.087 \times 0.8 \times 0.35 \times 0.8$. Similar to HB, HE is made up of brass and scintillator layers.

To extend the η coverage, a forward calorimeter, hadronic forward (HF), is located at the edges of the CMS covering a region of $3.0 < |\eta| < 5.0$. This detector is constructed from steel and quartz fibers. These materials result to shorter and narrower jets which is useful in the high flux forward region. HF is based on Cherenkov radiation, signal generation is produced in the quartz fibers when the charged particles in the the shower electromagnetic component exceed the corresponding Cherenkov thresholds. Further, charged hadrons are detected mainly through π^0 production. The produced Cherenkov light (signal) is guided to "Photo Multiplier" tubes (PMT) by the fibers. To account for the different responses for electrons and pions, HF uses two different sets of fibers with different lengths, which are read-out separately. Long (165 cm) and short (143 cm) quartz fibers run through HF with a 5 mm separation. The long fiber extend to the front face of the detector, while the short fibers end 22 cm before the front face. HF lies outside the strong magnetic field of the CMS, thus PMTs are utilized. Long and short fibers, which are readout separately, distinguish electromagnetic and hadronic showers in HF [44].

The energy (E) resolution (σ/E) of the HCAL is given as

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{90\%}{\sqrt{E}}\right)^2 + (4.5\%)^2, \tag{3.9}$$

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{198\%}{\sqrt{E}}\right)^2 + (9.0\%)^2. \quad (3.10)$$

where Eq. 3.9 is for the HB and HE sectors, and Eq. 3.10 is for the HF sector. The first term is the âstochasticâ term covering statistical fluctuations and shower fluctuations. The second term is a constant term to cover detector non-uniformity and calibration uncertainty. Since the HCAL is a sampling calorimeter, the resolution of the HCAL is inferior to the ECAL partly, also because hadronic showers have fewer particles which increase the statistical fluctuations.

3.2.5 Muon System

The outer-most layer of the CMS is the muon chambers positioned in $|\eta| < 2.4$ as shown in Figure 3.10 [45]. Since muons can penetrate several meters of iron without interacting, unlike most particles they are not stopped by any of the CMS calorimeters. The muon system detects muons (like electrons and positrons but 200 times heavier) which is constructed at the very edge of the experiment. Muon detectors are placed at four layers or stations in the barrel and endcap sections of the iron flux return.

Three different types of gaseous detector are used due to the varying radiation and magnetic environments. The barrel section, covering $|\eta| < 1.2$, uses drift tubes (DTs). The high muon and neutron background environment of the endcaps, $0.9 < |\eta| < 2.4$, requires cathode strip chambers (CSCs) to be used instead of DTs. Resistive Plate Chambers (RPCs) are used in both the barrel and part of the endcaps, $|\eta| < 2.1$. The magnetic field outside the solenoid is increased by an iron yoke, interspersed with the muon detectors. The field strength is about 2 T and bends the muons enough so that they leave a curved track from which the transverse momentum can be calculated.

The barrel region contains 250 chambers of up to 12 planes of drift tubes. The individual DTs have a cross section of 42×13 mm^2 and are filled with Ar and CO_2. Each drift tube consists of a central anode wire surrounded by aluminum cathodes. The induced charge has a maximum drift length of 2 cm or 380 ns. The spatial resolution is 100 μm in the $r\phi$ plane, and 150 μm in the z direction.

The two endcaps use 540 CSCs each of which is trapezoidal and contains 6 gas gaps. Each gap has a plane of radial cathode strips with perpendicular anode wires. The muon position is

measured from the charge sharing of the radial cathode strips. Each station provides a muon vector with a resolution of $\sim 200\ \mu m$ in ϕ and 10 mrad in direction.

There are 610 RPCs which consist of a gas gap enclosed by two graphite-coated bakelite plates. The graphite forms a cathode with an aluminum strip used to read out the generated signal. The RPCs have a time resolution of ~1 ns which makes them useful for identifying the bunch crossing time and useful for triggering.

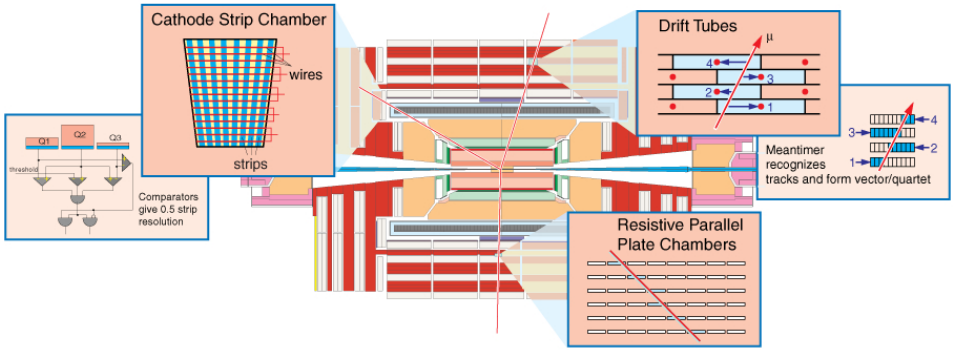

Figure 3.10: Drawing of the CMS muon detectors including cathode strip chambers (CSCs), drift tubes (DTs), and resistive place chambers (RPCs).

3.2.6 Trigger System

The design collision rate at the LHC is 40 MHz (bunch crossing interval is 25 ns), and the amount of data needed to record all information for an event is on the order of 250 kB. To store data for each proton-proton collision would then require writing several TB per second, which is far too much data for any storage system currently available and storage systems place a limit of ~200 MB per second. Moreover, there is no way that data from all these events can be read out, and even if they could, most of the data would be less likely to indicate new phenomena; they might be low energy regime collisions for instance, instead of energetic head-on interactions.

For the purpose of reducing the rate of recorded collision events for data storage and selecting only the most interesting events among the hard collision activities, the trigger system plays an important role within CMS dedector workflow. In particular, the trigger system reduces the recorded event rate to a more manageable 300 Hz in the data processing and storage [46]

and allows only the most promising and important events for full reconstruction.

The CMS trigger system achieves a rate reduction of collision events by a factor of 10^6. The trigger system is mainly divided into two parts; a fast hardware Level-1 (L1) trigger and a software High Level Trigger (HLT). The L1 trigger uses custom-built electronics hardware and reduces the event rate from 40 MHz to 100 kHz. The HLT runs on a commodity computing farm and reduces the rate further to $\sim$ 300 Hz for offline data storage.

3.2.6.1 Level-1 Trigger

The first step in reducing the number of events by selecting only 50-100 kHz of the most interesting events is the L1 trigger. It uses hardware algorithms that can make decisions in less than 3.2 μs. The L1 trigger utilizes information from calorimeters and muon systems with the rest of the event stored in a memory until the decision is reached. Events are selected for electrons, photons, jets, E_T, taus, muons, and combinations of these objects by the L1 system.

The L1 trigger has separate the Regional Calorimeter Trigger (RCT), the Global Calorimeter Trigger (GCT), and muon triggers, that are all fed into a global trigger, the Global Trigger (GT). The RCT readouts both the ECAL and HCAL calorimetry in coarse grained samples of one HCAL tower (for the HCAL, a tower is summed layers of tiles in depth) or 5×5 ECAL barrel crystals, which corresponds to an area of $\Delta\eta \times \Delta\phi = 0.087{\times}0.087$. The RCT identifies jet and photon/electron candidates (primitives). The results of the RCT algorithm are sent to the GCT where the candidates are counted and sorted. The GCT creates a transverse energy (E_T) ordered list of each primitive type. The calorimeter trigger results are then combined with the muon trigger results in the GT which uses threshold cuts on the primitives to make the accept/reject decision. At this stage, only 100 kHz of events pass the selection algorithms and accepted for the High Level Trigger. The L1 trigger flow is illustrated in Figure 3.11 along with the various L1 trigger systems.

3.2.6.2 High Level Trigger

Once accepted by the L1 trigger, the event is readout and combined by an event builder via a switched network capable of a data transmission rate of 1 TB per second. Aftermath, the

event is sent to a processor in the computing farm which runs the High Level Trigger (HLT) [47], [48] software algorithms. The HLT further filters out non-interesting events (mostly from QCD processes) for archiving rate reduction to almost 300 Hz. If an event passes the HLT it is passed to the offline computing system together with a list of all primitives passing the thresholds, which are called the trigger bits.

The HLT selection for electrons and photons are concerned in this thesis which will be discussed specifically. This selection follows in three stages in the CMS experiment. In the first stage, the calorimeter information is taken into account. In the next step, hits in the pixel detectors consistent with an electron candidate is concerned. Passing or failing the energy matching in the ECAL with hits of the pixel detector splits the electromagnetic triggers into two categories: (single and double) electron candidates and photon candidates at significantly higher energy thresholds. In the last stage, the selection of electrons uses full track reconstruction, conditioned from the pixel hits obtained at the matching stage.

In the HLT calorimeter reconstruction step, energy and position measurements follow clustering, but the reconstruction of electrons has special importance since the transverse momentum thresholds for triggering on electrons are much lower than those for photons. This comes from the fact that the amount of material traversed in the tracker volume is almost one radiation length for certain polar angles, thus the primary challenge for ECAL clustering is to include all energy radiated by electrons.

Photons that have converted in the tracker material are adequately reconstructed by the electron algorithms. For the final higher precision reconstruction of photons it is foreseen that unconverted photons will be reconstructed using energy sums of fixed arrays of crystals and that tracker information will be used to assist the clustering for converted photons. Such algorithms that improve the photon energy resolution, are not necessary for the HLT selection.

For the photon HLT paths, an energy deposit is identified in the ECAL first, and only if it passes the desired energy threshold requirement is a track reconstructed from the tracker information. The ECAL deposit and track are then matched, and if they pass a loose or tight selection, the event is saved. More detailed information on reconstruction of the HLT objects is given in Chapter 5. Since the luminosity raised quickly in the preiod of the 2011 data taking, the HLT requirements also changed quickly with respect to the 2010 HLT paths: the more luminosity implies that the more events need to be filtered out, the tighter the selection

needs to be in order to keep the rate reasonable lower.

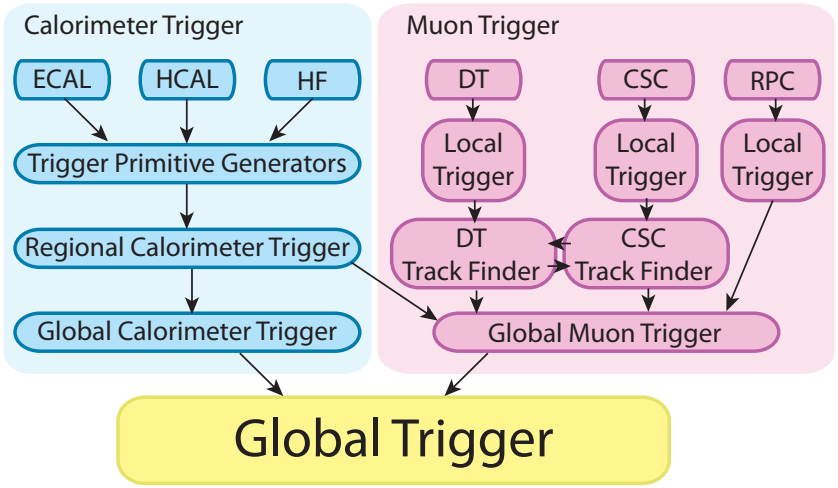

Figure 3.11: Schematic view of the various trigger systems within the L1 trigger.

3.3 Computing Model

The CMS Computing Model copes with computing requirements for storage, processing, and analysis of proton-proton collision data recorded. Given that the total rate of data produced by the CMS online trigger system is ~230 MB/s, these large amounts of data need large computing power and storage space to be able to efficiently extract physics results. In addition, large size of the CMS collaboration, users are from around the world, requires data to be accessible and distributed world wide.

3.3.1 Tiered Architecture

The CMS Computing Model meets computing needs by making use of the hierarchy of computing *Tiers* (tiered architecture) (Figure 3.12) as described in the CMS Computing Project [49]. There is a Tier-0 center and a CMS-CAF center located at CERN, there are 8 Tier-1 centers around the world, and finally, over 50 Tier-2 centers located at institutions and universities.

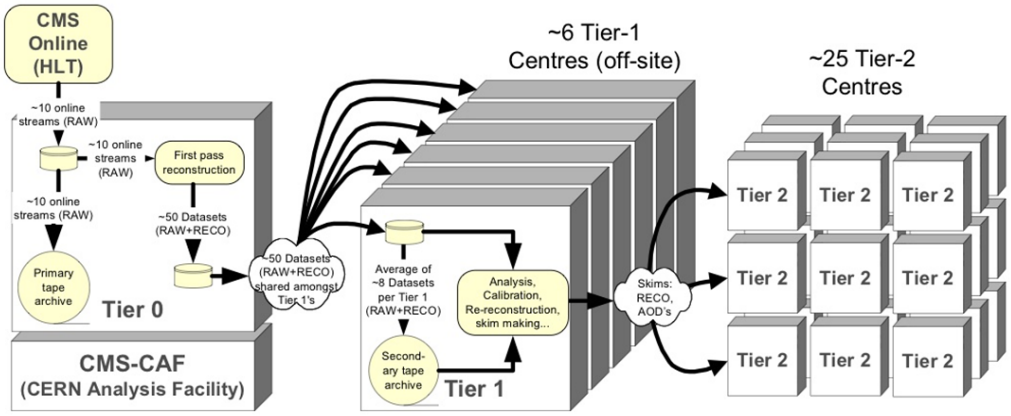

Figure 3.12: Tiered architecture and data flow of the CMS Computing system.

A single Tier-0 center at CERN accepts data from the CMS Online Data Acquisition system (DAQ), archives the data and performs prompt first pass reconstruction. Online streams of data arrive in a 20 day which is called input buffer. They are split into *Primary Datasets* that are then linked together to form reasonable file sizes. Primary Dataset RAW data contain the full recorded information from the detector and a record of the trigger decision. The RAW data are archived to tape at Tier-0 and sent to reconstruction nodes in the Tier-0. Resultant RECO (reconstructed) data are produced by applying detector-specific filtering and correction of the digitised data, cluster and track-finding and primary and secondary vertex reconstruction, as well as particle identification algorithms. Reconstruction is the most CPU-intensive activity in the CMS data processing chain. The resulting RECO events contain high-level physics objects, plus a full record of the reconstructed hits and clusters used to produce them. Both the RAW and RECO datasets are transferred then to Tier-1 centers. A dataset is not deleted from the Tier-0 until it is safely copied at two Tier-1 centers.

A CMS CERN Analysis Facility (CAF) provides access to the full RAW dataset and focusses on the latency-critical detector, trigger, calibration and analysis activities. Activities are directly coupled to the operations and performance of the CMS detector. In particular, CMS-CAF is responsible of diagnosing detector problems, providing trigger performance services such as reconfiguration, optimization and the testing of new algorithms, and calibrations required by the high level trigger or the initial reconstruction pass. These activities will have the highest priority at the CAF and will take priority over all other activities. The other tasks are supporting the analysis of CMS wide and CERN based users, recording and book keeping ser-

vices, provide the central information repository for data management, storing conditions and calibration data, providing the main software repositories and the documentation repositories, and other services associated with all Tier-1 centers.

A set of large Tier-1 centers in CMS collaborating countries provide services for data archiving, reconstruction, calibration, skimming and other data-intensive analysis tasks. Tier-1s receive data (RAW and RECO) from Tier-0 for long-term storage. When reconstruction software improves or updates, the RAW data is reconstructed again to produce new RECO datasets. The main task of Tier-1s to serve data to analysis groups running selections, skims, and test reprocessing. Most analysis products are sent to Tier-2s for iterative analysis works. It transfers these datasets in a reduced format, Analysis Object Data (AOD) to Tier-2 centers.

A more numerous set of Tier-2 centers (including Turkish Grid services provided by TUBITAK ULAKBIM), smaller but with substantial CPU resources, provide capacity for analysis and Monte Carlo simulation. Tier-2 centers rely upon Tier-1s for accessing to large datasets and securing storage of the new data products they produce. Each Tier-2 is associated with a particular Tier-1 center. Three types of use of Tier-2 resources are worth of discussing: local community use that is some fraction of the Tier-2 center resources will be fully under the control of their own local users, CMS controlled use which is Tier-2 resources will also be used for organized activities allocated top-down by CMS, and opportunistic use by any CMS member.

3.3.2 Worldwide LHC Computing Grid

The purpose of the Worldwide LHC Computing Grid (WLCG) project is to build and maintain a data storage and analysis infrastructure for the entire high energy physics community that will use the LHC [50]. The data from the LHC experiments including CMS will be distributed over the world, according to a tiered architecture discussed before. A primary backup will be recorded on tape at CERN, the Tier-0 center of LCG. After initial processing, this data will be distributed to a series of Tier-1 centers, large computer centers with sufficient storage capacity. The Tier-1 centers will make data available to Tier-2 centers, each consisting of one or several collaborating computing facilities, which can store sufficient data and provide adequate computing power for specific analysis tasks.

The Grid computing connects computers that are scattered over a wide geographic area, allowing their computing power to be shared. Just as the World Wide Web enables access to information, computer grids enable access to computing resources. These resources include data storage capacity, processing power, sensors, visualization tools and more. Thus, grids can combine the resources of thousands of different computers to create a massively powerful computing resource, accessible from the comfort of a personal computer.

The grid was designed to provide an advanced computing infrastructure suitable for collaborative problem solving within science and engineering. Resources, both computational and storage, were shared among collaborating institutions within a dynamic *Virtual Organisation* (VO). The VO comprises individuals based at different institutions around the world all working towards a common goal. The ultimate aim of grid computing was to provide ever present access to resources such that the user did not need to know where their work was carried out. They simply interacted with the grid and resources were provided. The grid was named by analogy with the power grid: users should consume computing power much as they consume electricity, without knowing the details of how or where it was generated.

The Worldwide LCG Project [50] depends upon several other projects for the supply of much of the specialized software used to manage data distribution and access as well as job submission and user authentication and authorization, known as the Grid *middleware*. These projects include *Globus*, *Condor*, *the Virtual Data Toolkit* and *the gLite* toolkit. The majority of the computing resources made available by the members of the collaboration are operated as part of the EGEE (Enabling Grids for E-sciencE) Grid, a consortium of national Grid infrastructures and computing centres from 34 countries including Turkey. Other resources are operated as part of other grids, such as the Open Science Grid (OSG) in the United States and the Nordic Data Grid Facility (NDGF). Achieving interoperability between different grids without compromising on the functionality constitutes a major challenge.

The interaction that the Grid provides between users and components of a complex software, middleware architecture can be seen in Figure 3.13. The User Interface (UI) was the gateway to the grid for users. The installed software allowed the user to manage data and submit computational jobs.

Data was stored on a site *Storage Element* (SE) and recorded in a global file catalogue, known

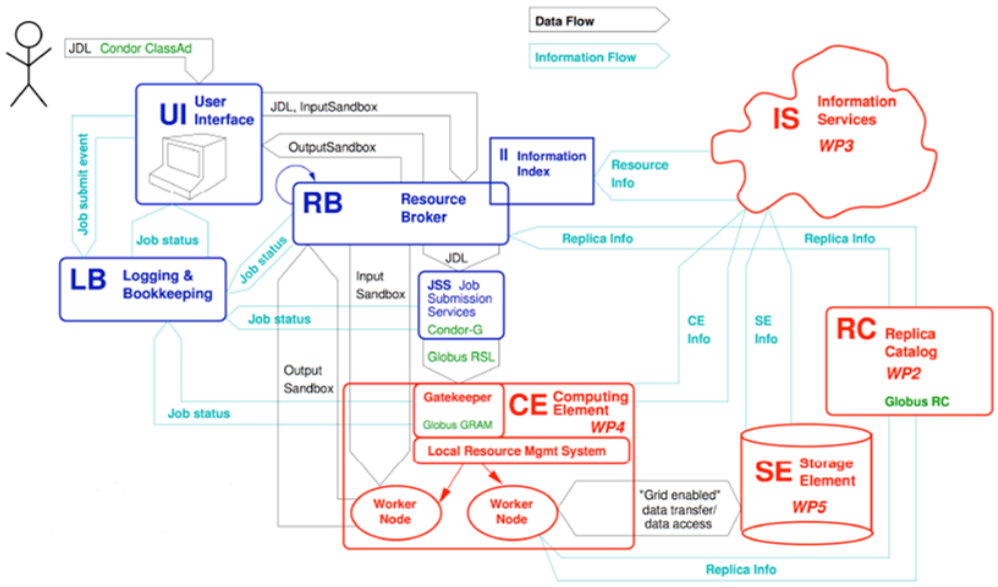

Figure 3.13: An overview of the major LCG components.

as the Replica Location Service (RLS) or *Replica Catalogue*. By using the client tools a user could copy data to/from and between SEs. Queries on the RLS were used for data discovery and location. The Workload Management System (WMS) was provided by *the Resource Broker* (RB). The role of RBs was to accept users jobs and take responsibility for assigning them to a site for processing. *The Computing Element* (CE) controlled the processing at a site, receiving jobs from the RB and scheduling them for execution on the site *Worker Nodes* (WNs). *The Logging and Bookkeeping* (LB) system recorded changes in job state and was queried by users to determine the state of their jobs. The Information Services (IS) listed all the RBs, CEs and SEs and was used by each system, and end users, to discover and obtain information about the other components.

CHAPTER 4

PHOTON MEASUREMENTS IN THE CMS EXPERIMENT

Photon measurement begins with reconstruction of energy deposit in the electromagnetic calorimeter (ECAL) of the CMS detector by employing dedicated clustering algorithms. Photon supercluster is corrected for better energy resolution at this stage. Reconstructed photon objects are discriminated from backgrounds which are primarily electrons and jets with the help of sets of isolation variables. Isolation requirements ensure that signal photons identified are no longer contaminated by backgrounds. Purity of signal photons are further verified by signal extraction templates that are based on variable fittings. Isolated real photons are counted and processed by offline analyses for the cross section measurement. Photon reconstruction and identification along with the ingredients and the most recent studies of isolated photon cross section are reviewed in this scope.

4.1 Photon Reconstruction

In CMS, photon candidates are primarily reconstructed from the energy deposits in the ECAL called superclusters (SC). In the ECAL barrel region ($|\eta| < 1.4442$), superclusters are formed from the energy sum clustered in a rectangle of crystals 35 wide in ϕ and 5 wide in η (Figure 4.1). This allows to recover energy emitted from photon conversion into electron-positron pairs via photon conversion, $\gamma \rightarrow e^+e^-$, since electrons and positrons are bent by the 3.8 T magnetic field to separate in the ϕ direction. In the ECAL endcap region ($1.566 < |\eta| < 2.5$), superclusters are formed by one or more arrays of 5×5 crystals (Figure 4.1). The superclusters are corrected for energy to cope with interactions with the material in front of the ECAL. Photon candidates are then built with the corrected superclusters, assigning the candidate momentum to the location of the primary vertex ([51], [52], and references therein).

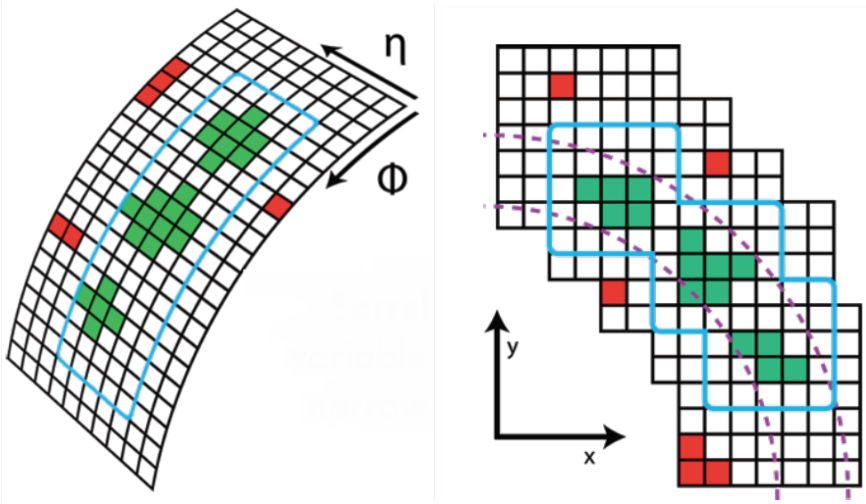

Figure 4.1: Illustrations of energy deposit superclustering in the electromagnetic calorimeter. The lines in light blue show the grouping of crystals in the ECAL Barrel (left) and Endcap (right) regions.

4.1.1 Clustering Algorithms

Dedicated algorithms are used to efficiently reconstruct the full photon energy which are called *clustering algorithms*. These algorithms differ as a result of different arrangement of the crystals in the barrel and endcap regions of the ECAL. Both the barrel and endcap clustering algorithms group the crystals which are associated with individual energy deposits (hence electromagnetic showers) and collect showers which are close in η but spread in ϕ direction.

In the ECAL barrel region, *Hybrid algorithm* is used to exploit the $\eta - \phi$ geometry of this part of the detector. The Hybrid algorithm collects both energies in individual showers and set of showers compatible with bremsstrahlung from the crystals. The result of the procedure are superclusters made up of several showers at constant η but spread in ϕ.

In the ECAL endcap region, superclustering proceeds similarly, but a slightly different algorithm is used since the crystals in the endcap are not arranged in an $\eta - \phi$ geometry as in the barrel. Thus the same idea of collecting energy deposits within a window in η and ϕ is implemented by the *Multi5x5* algorithm. This algorithm adds fixed 5×5 crystal clusters together for the process of superclustering.

4.1.2 Energy Corrections

A high resolution of the ECAL is vital in every new physics search with electrons and/or photons in the final state. A precise measurement of energy deposit in the calorimeter improves the resolution of missing transverse energy, which is a signature for beyond the known physics. There are different reasons that are influencing the ECAL energy resolution negatively. One of them is the interaction of particles with matter (the tracker material), which results in bremsstrahlung and photon conversions and causes the energy of electromagnetic particles measured in ECAL to be underestimated. The function of supercluster energy corrections is necessary to compensate this energy loss, or in other words to control energy scales in the ECAL.

Several correction factors are applied to the total supercluster energy to obtain high energy resolution after the clustering step. Correction factors are applied in a way to achieve a homogeneous response in the full calorimeter volume. The corrected photon energy can then be given as:

$$E_\gamma = F_\gamma \times G \times \sum_{i}^{N} c_i A_i, \tag{4.1}$$

where A_i are the digital amplitudes measured by the ECAL readout electronics and c_i are a set of coefficients that equalise the response of the different crystals, G is named as a global energy scale and defined such that the sum of the amplitudes of a 5×5 crystal matrix multiplied by G amounts to the total energy of an incident unconverted photon. The factor F_γ represents a correction to the supercluster energy, which takes into account the differences in shower containment in the calorimeter as a function of the supercluster η. This factor is also in place to correct for the different response of the algorithm to different topologies and the effect of the non-linear distribution of the material in front of the calorimeter. After applying energy corrections, the superclusters are used to reconstruct photons (and also electrons), and to seed electron track reconstruction.

4.2 Photon Identification

Decays of neutral mesons are produced in association with jets. The jet background to prompt photon production can be reduced by requiring the reconstructed superclusters to be *isolated*, i.e. limiting the amount of energy carried by other particles surrounding them. In order to suppress the jet background contribution, the following variables are used in CMS to isolate (identify) signal prompt photons [52]:

- Hadronic over electromagnetic ratio (H/E): fraction of hadronic energy to total electromagnetic energy inside a cone of $\Delta R = 0.15$ around photon direction. For photons this ratio should be low, while for jets, which carry both hadronic and electromagnetic energy, it is generally higher.

- Shower shape ($\sigma_{i\eta i\eta}$): transverse shape of the electromagnetic cluster calculated with logarithmic weights as

$$\sigma^2_{i\eta i\eta} = \frac{\sum_i^{5\times5} w_i(\eta_i - \overline{\eta}_{i5\times5})^2}{\sum_i^{5\times5} w_i}, w_i = max(0, 4.7 + ln\frac{E_i}{E_{5\times5}}) \tag{4.2}$$

  where E_i and η_i are the energy and pseudorapidity of the i^{th} crystal within 5×5 electromagnetic cluster and $E_{5\times5}$ and $\eta_{5\times5}$ are the energy and the pseudorapidity of the entire 5×5 cluster. Since the trajectory in η of a photon is not affected by the magnetic field, magnitude of shower shape in η should be small, whereas for a π^0 it will tend to be larger.

- Tracker Isolation (Iso_{TRACK}): computed as the sum of the transverse momenta of all the tracks in a full cone ($\Delta R_o = 0.4$) centered around a line joining the primary vertex to the cluster. Tracks in an inner cone ($\Delta R_i = 0.04$) and a rectangular strip of $\Delta\eta \times \Delta\phi = 0.015 \times 0.4$ are not included to avoid misidentification of converted photons. Since prompt photons are generally isolated, they should have a lower value of it, while photons from fragmentation and decay processes will be generally accompanied by other tracks, and hence a higher value of tracker isolation.

- ECAL Isolation (Iso_{ECAL}): computed as the E_T sum of the energy deposits in the approximately 1250 individual crystals located in a cone ($\Delta R_o = 0.4$), centered around the supercluster. Inner cone of radius $\Delta R_i = 0.06$ and a strip of dimensions $\Delta\eta \times \Delta\phi =$

0.04×0.4 are excluded. Like tracker isolation, it tends to have a lower value for isolated photons.

- HCAL Isolation (Iso_{HCAL}): computed as the sum of the transverse energy in the HCAL towers in a hollow cone with inner radius of ΔR_i = 0.15 and an outer radius of ΔR_o = 0.4 which is centered around the photon supercluster. As with the other isolation variables, its value is lower for prompt photons.

In Figure 4.2, photon isolation variables of H/E, Iso_{ECAL}, Iso_{HCAL}, and Iso_{TRACK} are given without any pre-selection on the event. The distributions are compared between Pythia simulation samples in which signal photon comes from the hard processes (not from fragmentation of decays) and real collision data recorded by the CMS detector in 2011. The Monte Carlo and real data shapes are normalized to unit area on histograms to be able make comparisons regadless of statistics of the data samples. In Figure 4.3, photon cluster shape variable, $\sigma_{i\eta i\eta}$ is compared between simulation and data samples in the ECAL Barrel and Endcap regions separately. The distributions are obtained without any pre-selection on the event and the shapes are scaled to unit area for ease of comparision.

The above photon isolation requirements are together known as *photon signal region preselection*. In addition to this photon ID selections, photon superclusters are required not to match pixel hits consistent with a track from the interaction region. This *pixel seed veto* criteria ensures that background from electrons is suppressed by requiring the absence of a short track segment that is built from either two or three hits in the silicon pixel detector and consistent with an electron track matching the observed location and energy of the photon candidate.

4.3 Signal Extraction

Background photons that are produced in the decay of neutral hadrons (inside jets) may fake and contaminate signal prompt photon candidates. Photon identification (isolation) selections can remove such bakground to some extent as discussed in the previous section. However, a considerable contamination from the neutral hadrons may remain even after applying the offline photon ID selections. For that reason, two *template techniques* are developed to estimate the purity of signal events within the candidates that pass photon identification selection in

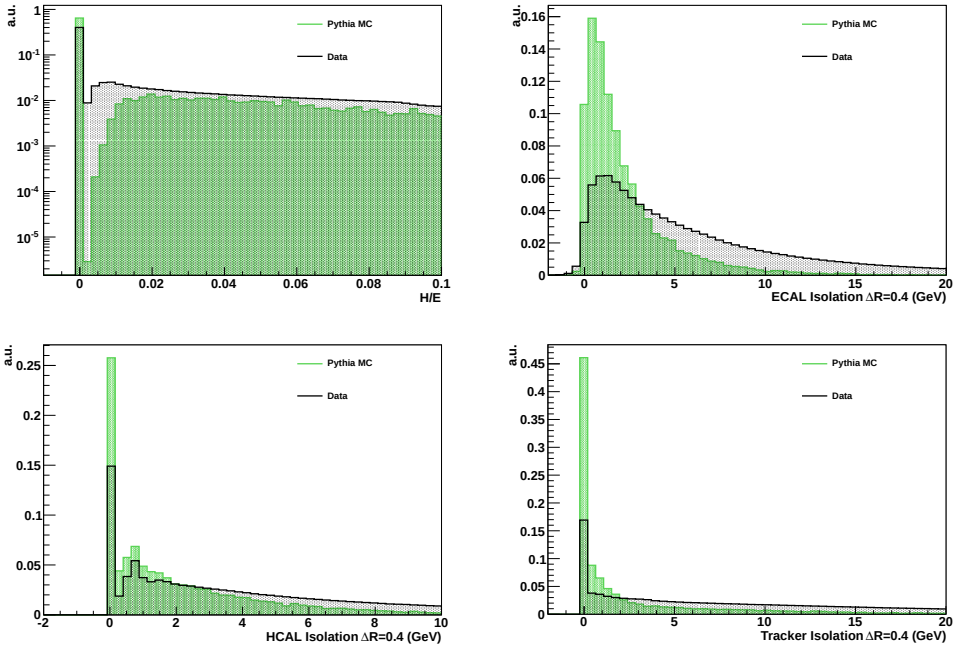

Figure 4.2: Photon isolation variables are shown from Pythia Monte Carlo samples and real data recorded by the CMS detector in 2011. The shapes are scaled to unit area for comparisions for both Barrel and Endcap photons.

the CMS experiment. These methods take advantage of differences in variable distributions between signal events and background events. In order to extract the fraction of signal events, the combinition of two distributions from signal and background are used to fit the distribution of all candidates [53].

4.3.1 Shower Shape Template

When a photon candidate measured in the ECAL, it will give its energy through some number of crystals. The shower shape variable, $\sigma_{i\eta i\eta}$, is used to measure the width of the photon supercluster in the η direction. For signal events, $\sigma_{i\eta i\eta}$ distribution tends to have a narrow peak while for background it tends to have a wider peak with a heavy tail extending to a higher value on $\sigma_{i\eta i\eta}$.

In the shower shape template, the photon candidates are collected by requiring events to satisfy the photon identification criteria (Table 4.1). The signal template is taken from the γ+jet Monte Carlo samples and the background template is taken from real data satisfying the pho-

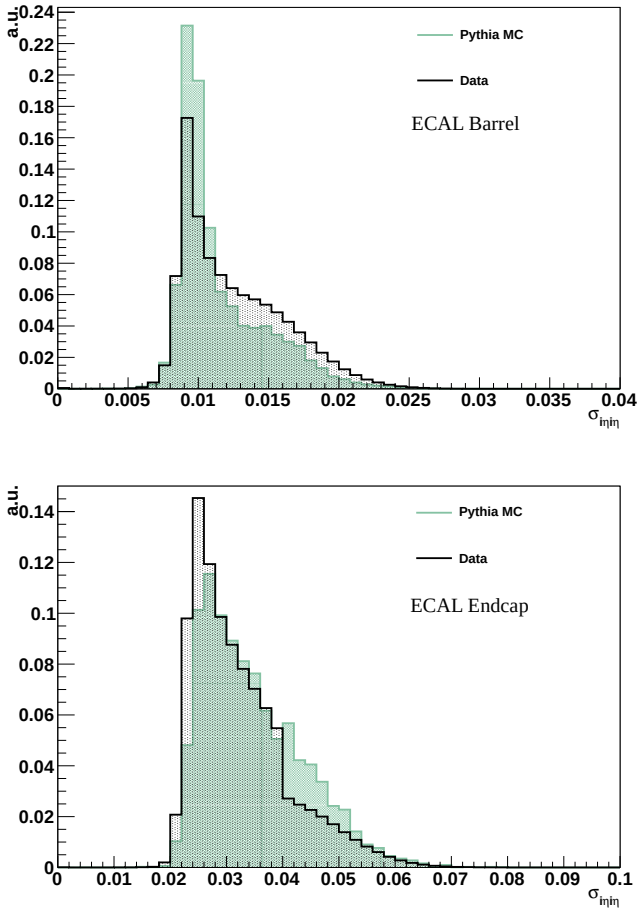

Figure 4.3: Photon shower shape variable is shown from Pythia Monte Carlo samples and real data recorded by the CMS detector in 2011. The shapes are scaled to unit area for ease of comparisions in the Endcap Barrel (left) and the Endcap (right) regions.

ton sideband selection given in the Table 4.2. After the signal and background templates are determined, the combinition of them are *fitted* to the $\sigma_{i\eta i\eta}$ distribution in data by an extended maximum likelihood method.

For each bin in photon p_T, the data is fitted with the function, $f(\sigma_{i\eta i\eta}) = N_S S(\sigma_{i\eta i\eta}) + N_B B(\sigma_{i\eta i\eta})$, where N_S and N_B refer to the estimated number of signal and background events. The fit is usually performed by a *binned extended maximum likelihood method* where negative log likelihood function is minimized as in the following way:

Table 4.1: Photon identification criteria for signal candidate selection in the shower shape template.

Variable	Selection
pixel seed	require none
Tracker Isolation	$< 2.0 + 0.002 \cdot E_T^{\gamma}$ GeV
ECAL Isolation	$< 4.2 + 0.012 \cdot E_T^{\gamma}$ GeV
HCAL Isolation	$< 2.2 + 0.005 \cdot E_T^{\gamma}$ GeV
H/E	< 0.05

Table 4.2: Photon sideband selection for background candidate selection in the shower shape template.

Variable	Selection
pixel seed	require none
Tracker Isolation	$> 2.0 + 0.002 \cdot E_T^{\gamma}$ GeV
Tracker Isolation	$< 5.0 + 0.002 \cdot E_T^{\gamma}$ GeV
ECAL Isolation	$< 4.2 + 0.012 \cdot E_T^{\gamma}$ GeV
HCAL Isolation	$< 2.2 + 0.005 \cdot E_T^{\gamma}$ GeV
H/E	< 0.05

$$\mathcal{L} = -\ln L = -(N_S + N_B) + \sum_{i=1}^{n} N_i \ln(N_S S_i + N_B B_i) \tag{4.3}$$

where n is the number of bins in $\sigma_{i\eta i\eta}$, N_i is the observed number of events for the i-th $\sigma_{i\eta i\eta}$ bin and S_i and B_i are the values of the corresponding components in that bin; N is the total number of data events in the given E_T bin.

4.3.2 Isolation Template

Photon candidate can deposit its energy in the Tracker, ECAL and HCAL sub-systems. In the isolation template method, deposited energy for signal shape and background shape should differ. For this purpose the variable, Iso, is used to measure the energy deposited around the center of photon candidate shower. The variable Iso is the summation of the isolation energy variables: $Iso = Iso_{TRK} + Iso_{ECAL} + Iso_{HCAL}$, where the isolation variables are defined previously. In this template, the signal photon candidates are required to satisfy the photon idendification criteria listed in Table 4.3 and taken from the γ+jet Monte Carlo samples. The background template is taken from real data satisfying the photon sideband selection given in the Table 4.4.

Table 4.3: Photon identification criteria for signal candidate selection in the isolation template.

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 4.4: Photon sideband selection for background candidate selection in the isolation template.

Variable	Selection
$\sigma_{i\eta i\eta}$ (Barrel)	> 0.011
$\sigma_{i\eta i\eta}$ (Endcap)	> 0.035
H/E	< 0.05

The signal shape in this template is fitted by the convolution of a lifetime function with *Gaussian distribution*:

$$S(x) = \frac{1}{p_0} \times e^{(p_2^2/2p_0^2)-((x-p_1)/p_0)} \times \left[1 - Freq\left(p_2/p_0 - (x-p_1)/p_2\right)\right], \tag{4.4}$$

where the parameters p_1 and p_2 describe the peak of template and are freely floated in the template fitting, p_0 is the parameter effecting tail of the signal template, and x is the Iso variable. For the background template, the inverted *ARGUS* function is used for the fitting:

$$B(x) = \left[1 - e^{p_3(x-p_4)}\right] \times \left[1 - p_5(x - p_4)\right]^{p_6} . \tag{4.5}$$

The parameters p_3 and p_4 control left-side peak of the template which determines the background in the signal dominated region. During the template fitting, p_3 is contsrained and p_4 is fixed. The other two parameters, p_5 and p_6 describe the right-side tail of the template, where the data are dominated by background events. These two parameters will be freely floated in template fitting.

After determing the signal and background templates, data are fit with $f(Iso) = N_S S(Iso) + N_B B(Iso)$, where N_S and N_B refer to the estimated number of signal and background events. The fit is performed by a binned extended minimum χ^2 method where the following χ^2 function is minimised:

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{N_i - (N_S S_i + N_B B_i)}{error_{N_i}}\right)^2 + \left(\frac{(p_0 - p_0(central))}{p_0(uncertainty)}\right)^2 + \left(\frac{(p_3 - p_3(central))}{p_3(uncertainty)}\right)^2, \tag{4.6}$$

where n is the number of bins in template, N_i is the observed number of events for the i-th Isolation bin, and S_i and B_i are the values of the corresponding components in that bin; N is the total number of events from data; p_0(central) and p_0(uncertainty) is the background fitting results for p_0; p_3(central) and p_3(uncertainty) is the background fitting result for p_3.

The parameters p_1, p_2, p_5 and p_6 are fixed in the fitting procedure. The background starting point (p_4) is fixed. The parameter of the power function describing signal tail (p_0) and background turn-on power (p_3) are constrained using the last two terms of the χ^2 function.

The results of this fitting procedure for the signal photon and background photon are shown as examples in Figure 4.4 to Figure 4.7 for a selection in which a photon p_T is between 40-45 GeV in different pseudorapidity regions of photon and jet. As in the shower shape template, these data fitted results are used to exract the purity of the photon samples measured from data and modelled from Monte Carlo simulation samples.

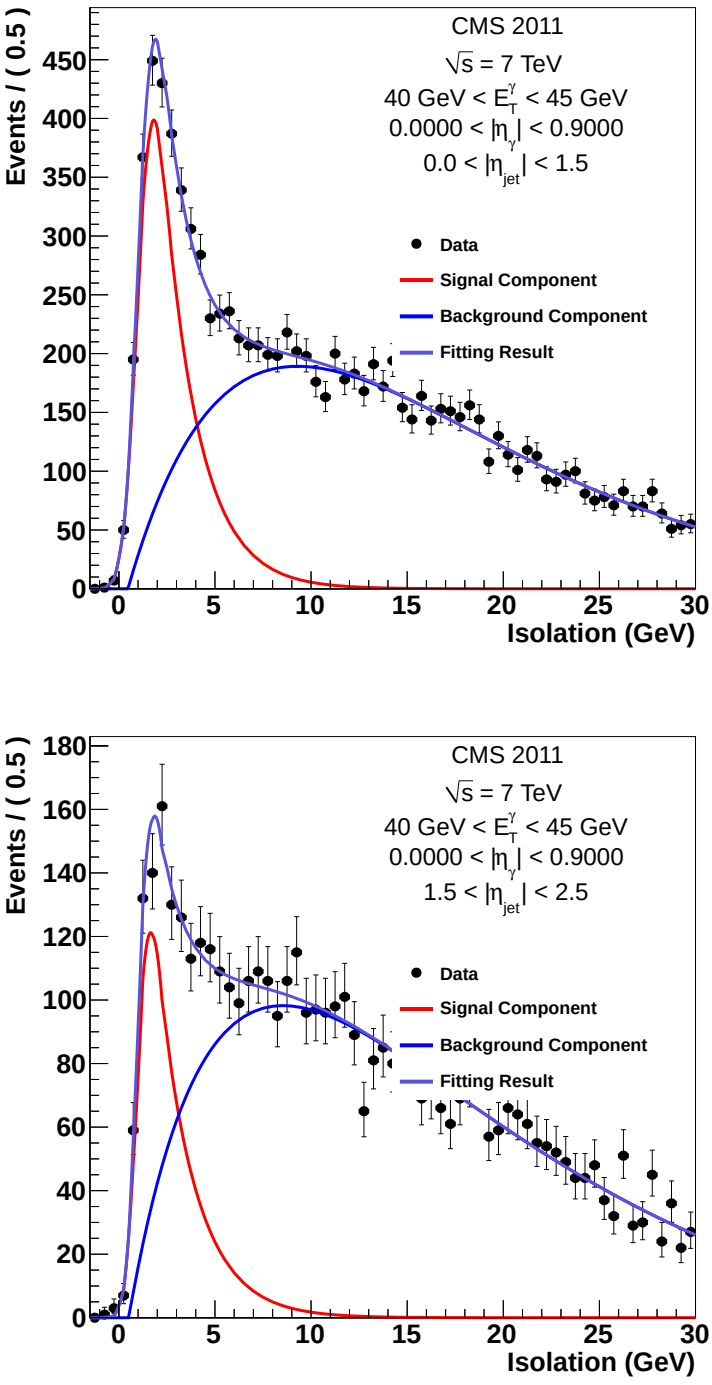

Figure 4.4: Examples to the isolation template fit results for $\eta^{\gamma} < 0.9$: Measured Iso distributions (points with error bars). The extended χ^2 fit result (light blue line) is overlaid in each plot and the component for background (dark blue line) and the component for signal (red line) are shown separately. The top plot is for $0 < \eta^{jet} < 1.5$ and the bottom one is for $1.5 < \eta^{jet} < 2.5$.

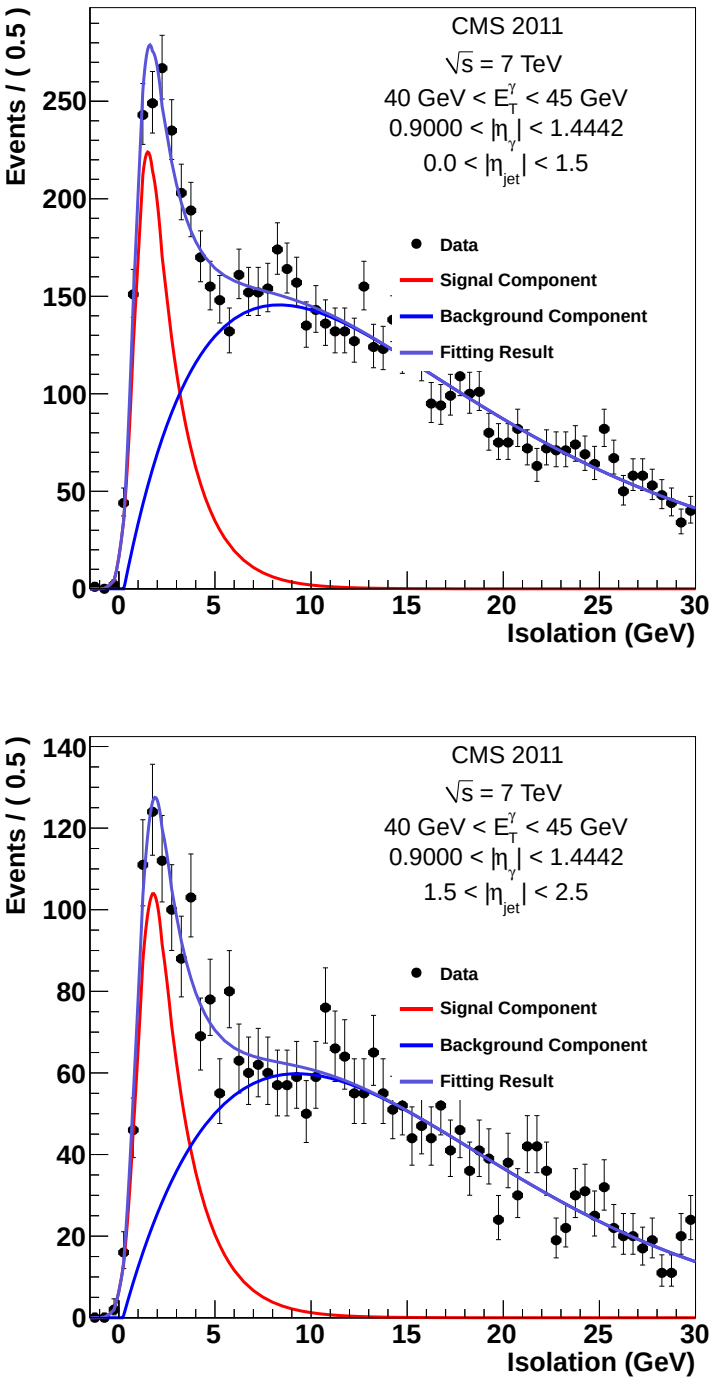

Figure 4.5: Examples to the isolation template fit results for $0.9 < \eta^{\gamma} < 1.442$: Measured Iso distributions (points with error bars). The extended χ^2 fit result (light blue line) is overlaid in each plot and the component for background (dark blue line) and the component for signal (red line) are shown separately. The top plot is for $0 < \eta^{jet} < 1.5$ and the bottom one is for $1.5 < \eta^{jet} < 2.5$.

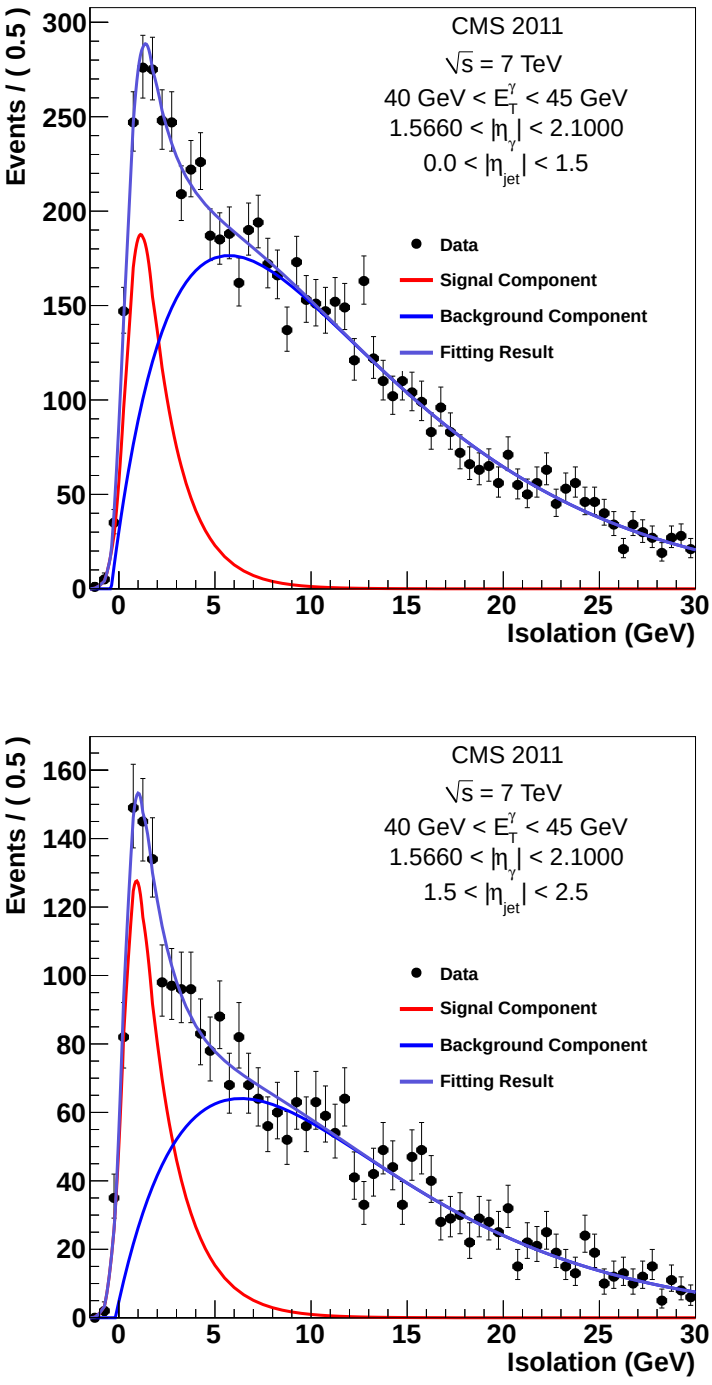

Figure 4.6: Examples to the isolation template fit results for $1.566 < \eta^{\gamma} < 2.1$: Measured Iso distributions (points with error bars). The extended χ^2 fit result (light blue line) is overlaid in each plot and the component for background (dark blue line) and the component for signal (red line) are shown separately. The top plot is for $0 < \eta^{jet} < 1.5$ and the bottom one is for $1.5 < \eta^{jet} < 2.5$.

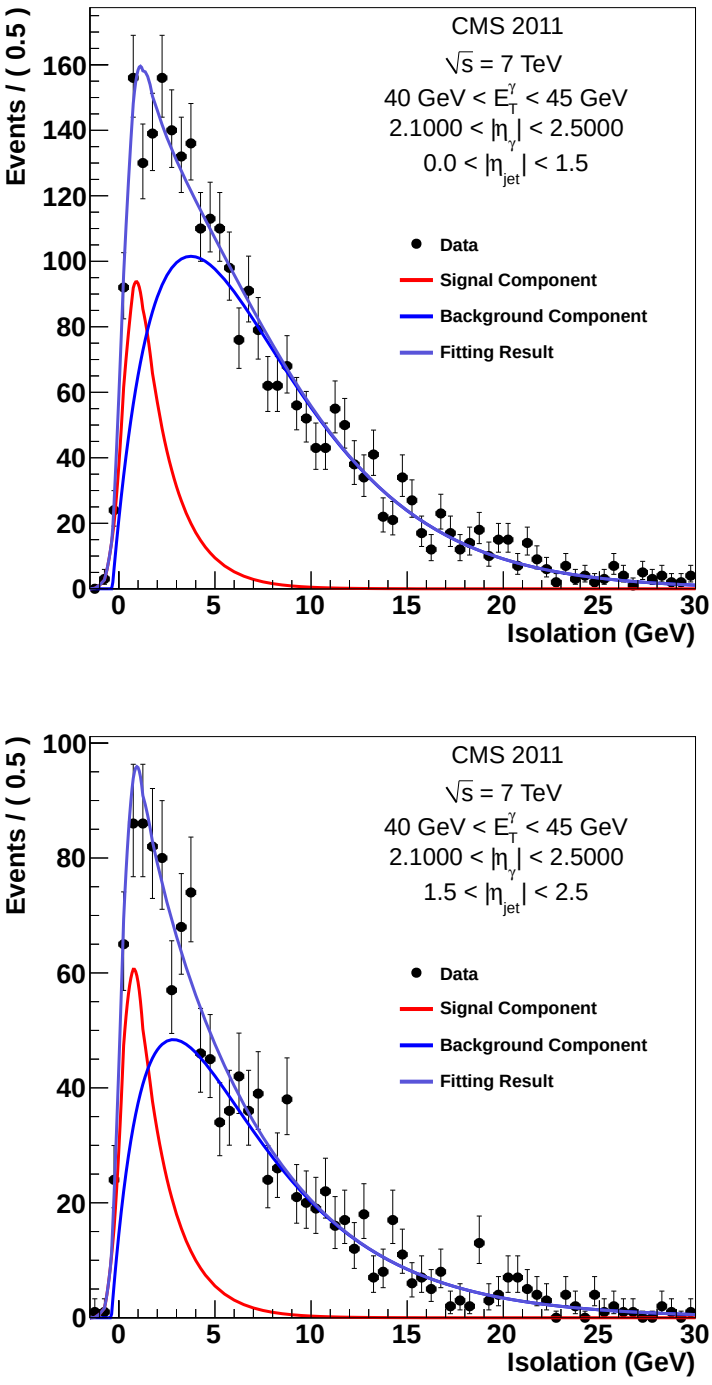

Figure 4.7: Examples to the isolation template fit results for $2.1 < \eta^{\gamma} < 2.5$: Measured Iso distributions (points with error bars). The extended χ^2 fit result (light blue line) is overlaid in each plot and the component for background (dark blue line) and the component for signal (red line) are shown separately. The top plot is for $0 < \eta^{jet} < 1.5$ and the bottom one is for $1.5 < \eta^{jet} < 2.5$.

4.4 Cross Section Measurements

In high energy physics experiments including the CMS, isolated prompt photon double differential cross section is defined as in the following form of

$$\frac{d^2\sigma^{\gamma}_{isolated}}{dE^{\gamma}_{T}d\eta^{\gamma}} = \frac{1}{\Delta E^{\gamma}_{T}\cdot\Delta\eta^{\gamma}}\frac{N^{\gamma}_{signal}\cdot U}{L\cdot\epsilon}, \tag{4.7}$$

where ΔE^{γ}_{T} and $\Delta\eta^{\gamma}$ are isolated photon transverse energy and pseudorapidity variations respectively, N^{γ}_{signal} is measured number of signal photons (photon yield), L is corresponding integrated luminosity, U is unfolding factor which includes corrections for photon reconstruction effects, and ϵ is the product of trigger, reconstruction, and identification efficiencies ($\epsilon = \epsilon_{trigger}\times\epsilon_{RECO}\times\epsilon_{ID}$).

If the final state involves prompt photons produced in association with jet(s), then double differential cross section given above is slightly modified to take into account geometrical acceptance of jet. Hence, photon+jet triple differential cross section can be given as in the following modified form

$$\frac{d^3\sigma^{\gamma}_{isolated}}{dE^{\gamma}_{T}d\eta^{\gamma}d\eta^{jet}} = \frac{1}{\Delta E^{\gamma}_{T}\cdot\Delta\eta^{\gamma}\cdot\Delta\eta^{jet}}\frac{N^{\gamma}_{signal}\cdot U}{L\cdot\epsilon}, \tag{4.8}$$

where $\Delta\eta^{jet}$ is the pseudorapidity size of the associated jet and the other elements remain the same as they are defined in cross section formulation of Eq. 4.7.

4.5 Previous Results

The earliest measurements of prompt photon production were carried out at the ISR (Intersecting Storage Rings) collider at CERN [54], [55]. Further, the series of other studies established prompt photons as a useful probe of hadron interactions [56], [57], [58]. Here, the most recent prompt photon production measurements at RHIC, Tevatron, HERA, and the LHC by various collaborations are summarized in terms of cross section measurements and theoretical comparisons. These collaborations measured photon production for various ranges of photon pseudorapidity and for different ranges of photon transverse energy.

The PHENIX collaboration at the Relativistic Heavy Ion Collider (RHIC) measured both the inclusive and isolated photon production rate at center-of-mass energy of 200 GeV ([59]). In these measurements, photons are detected in different pseudorapidity and phi angle regions within the photon energy range of 4-17 GeV. Their measurements provided a deep understanding into the fragmentation contribution to prompt photon production.

More recent measurements of prompt photon production cross section at hadron colliders were performed at the Tevatron, in $p\overline{p}$ collisions at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV. The measurement was performed by the DØ Collaboration [60] is based on 380 pb^{-1} data and covers a pseudorapidity range $|\eta(\gamma)| < 0.9$, with a transverse energy range $23 < E_T < 300$ GeV. Later, the measurement by the CDF Collaboration [61] is based on 2.5 fb^{-1} data and covers a pseudorapidity range $|\eta(\gamma)| < 1.0$, with a transverse energy range of $30 < E_T < 400$ GeV. Both DØ and CDF studies present measurements of an isolated prompt photon cross section which are in good agreement with next-to-leading order pQCD calculations, with a slight excess observed in the CDF data between 30 and 50 GeV transverse energy band. Their cross section measurement results, compared with JETPHOX predictions, are shown in Figure 4.8.

Prompt photon production with or without an accompanying jet was studied extensively in electron-positron collisions at a centre-of-mass energy of 318 GeV at HERA. The ZEUS colaboration reported γ+jet cross section for the transverse energy of the photon and the jet larger than 5 and 6 GeV, respectively from an integrated luminosity of 77 pb^{-1} [62]. The differential γ+jet cross sections were reconstructed as functions of the transverse energy, pseudorapidity, and x_γ^{obs} scale which is the fraction of the incoming photon momentum taken by the photon-jet system. Their predictions are based on leading logarithm parton shower Monte Carlo models and next-to-leading-order (NLO) QCD generally which underestimate the cross sections for the transverse energies of prompt photons below 7 GeV, while the k_T-factorisation ([63]) QCD calculation agrees with the data better. When the minimum transverse energy of prompt photons is increased to 7 GeV, both NLO QCD and the k_T-factorisation calculations were shown to be in good agreement with the data (Figure 4.9).

Later in 2010, the ZEUS collaboration reported differential cross sections of inclusive photon production using an integrated luminosity of 320 pb^{-1} [64]. In this measurement, the

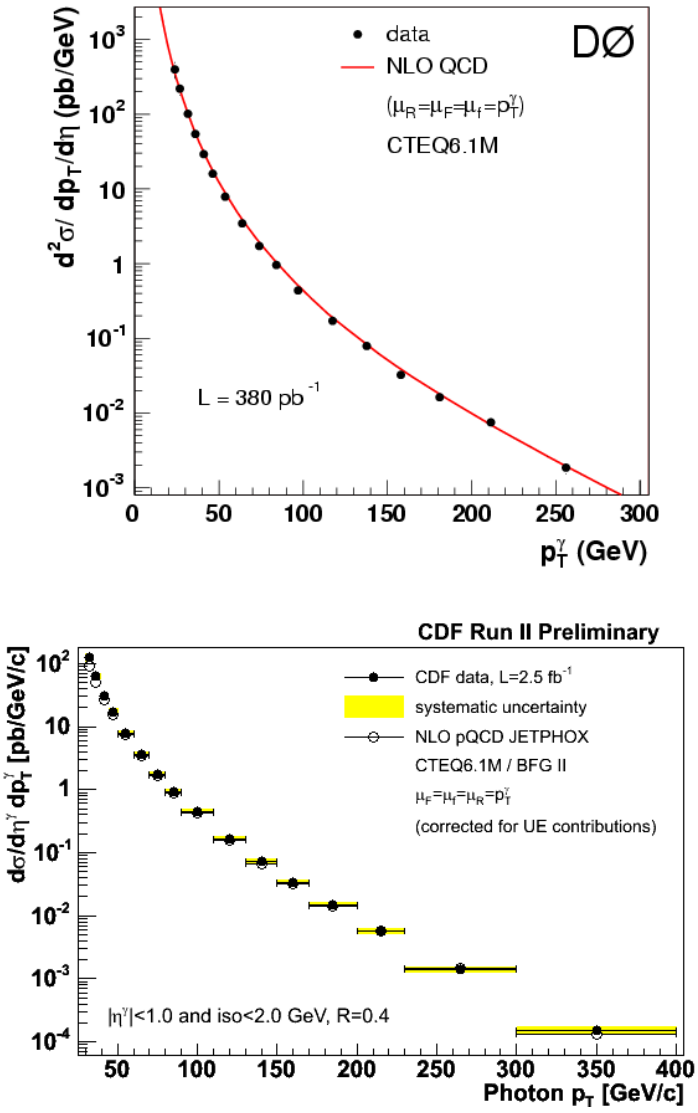

Figure 4.8: Measured inclusive isolated prompt photon cross section as a function of photon p_T (E_T) compared to NLO pQCD predictions with JETPHOX in DØ experiment (top) and in CDF experiment (bottom) at Tevatron.

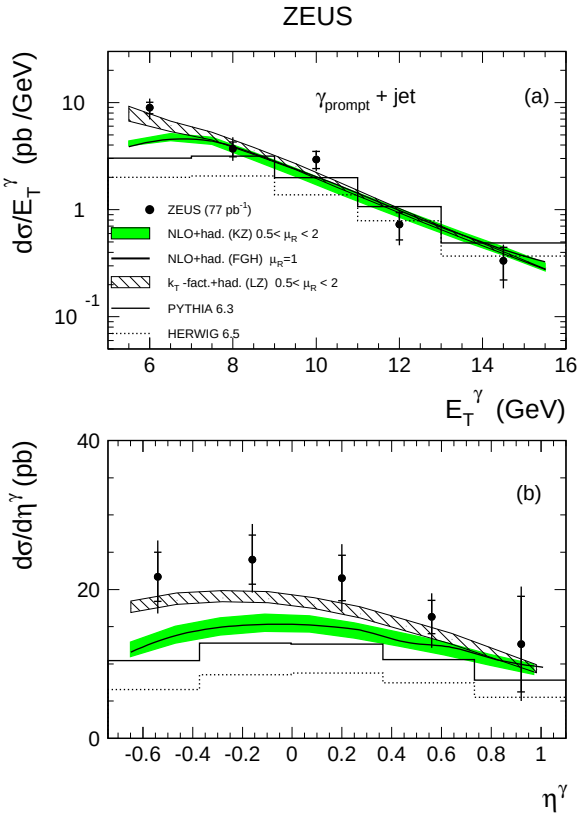

Figure 4.9: Measured differential cross sections for prompt photon plus jet events by the ZEUS experiment at HERA as a function of photon E_T and η^γ. The theoretical pQCD predictions are shown from different NLO methods.

isolated photon production was reconstructed in transverse energy and pseudorapidity ranges and pQCD predictions were shown to have reasonable description of the data over most of the kinematic range studied. In addition, the H1 collaboration measured both the inclusive prompt photon and γ+jet productions differential cross sections at HERA based on data corresponding to a total integrated luminosity of 340 pb^{-1} [65].

In the most recent measurement by the ATLAS experiment at the LHC [66], the differential cross section for the inclusive production of isolated prompt photons in pp collisions at a center-of-mass energy $\sqrt{s}$ = 7 TeV is presented. The ATLAS measurement covers the pseudorapidity ranges $|\eta| < 1.37$ and $1.52 < |\eta| < 2.37$ in the transverse energy window of $45 < E_T < 400$ GeV. The results are reported from an integrated luminosity of 35 pb^{-1} and are compared with next-to-leading order perturbative QCD calculations. This study also reports that their measurements are in good agreement with theoretical predictions over four

orders of magnitude in cross-section as shown in Figure 4.10.

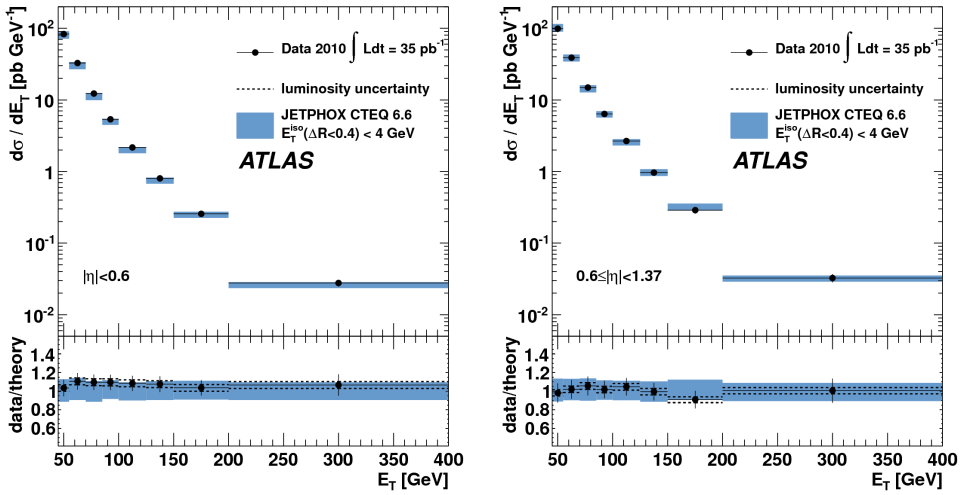

Figure 4.10: Measured and expected inclusive prompt photon production cross-sections, and their ratio, as a function of the photon E_T in two selected different pseudorapidity ranges by the ATLAS experiment at LHC.

In the CMS experiment, the first measurement of the inclusive photon production differential cross section is performed with a relatively limited amount of collected data, 2.9 pb^{-1} at $\sqrt{s}$ = 7 TeV in 2010 [67]. The double differential cross section for the isolated prompt photons is measured as a function of the photon transverse energy E_T in pp collisions. Photons are required to have a pseudorapidity $|\eta| < 1.45$ which corresponds to reconstructed photons in the ECAL barrel region. In addition, photon transverse energy is required to be $21 < E_T < 300$ GeV. The measured cross section is found to be in agreement with next-to-leading-order perturbative QCD calculations provided by the JETPHOX as given in Figure 4.11.

The most recent measurement of the inclusive photon production cross section by CMS came in 2011 with a higher amount of data and extended pseudorapidity regions [67]. For this measurement, an integrated luminosity of 36 pb^{-1} at $\sqrt{s}$ = 7 TeV is used providing much statistics for the final results compared to the previous measurement. This measurement covers the pseudorapidity range of both the ECAL barrel and endcap $|\eta| < 2.5$ and the transverse energy range $25 < E_T < 400$ GeV.

In this inclusive photon cross section measurement, photon candidates are identified with two complementary methods, one based on photon conversions in the silicon tracker and the other

one is based on the isolated energy deposits in the electromagnetic calorimeter. The measured cross section is given as a function of photon E_T in four pseudorapidity regions (two for the barrel and two for the endcap). The next-to-leading-order perturbative QCD calculations are also found to be consistent with the measured cross section in this study (Figure 4.11).

In addition to isolated single photon cross section measurements at the colliders, triple cross section of an isolated photon in association with jet(s) is also measured by the DØ Collaboration [69]. In this measurement, 1.0 fb^{-1} of data collected by the DØ detector is used at a center-of-mass energy of 1.96 TeV. Photons are reported to be reconstructed in the central rapidity region $|\eta(\gamma)| < 1.0$ with transverse momenta in the range $30 < p_T(\gamma) < 400$ GeV while jets are reconstructed in either the central $|\eta(jet)| < 1.0$ or forward $1.5 < |\eta(jet)| < 2.5$ rapidity intervals with $p_T(jet) > 15$ GeV. The triple photon+jet differential cross section is measured as a function of photon transverse momentum in four regions, differing by the relative orientations of the photon and the jet in rapidity. Further in this study, ratios between the differential cross sections in each region are also reported. Next-to-leading order QCD predictions using different parameterizations of parton distribution functions and theoretical scale choices are compared to the data (Figure 4.12).

This study is the first measurement that has been performed for the γ+jet differential triple cross section. The same measurement is initiated at the CMS experiment with higher luminosity and energy. The CMS results of this measurement is to be published in a journal as of this thesis writing with an improved measurements of detector photon efficiency and photon purity at higher energies with larger statistics [53].

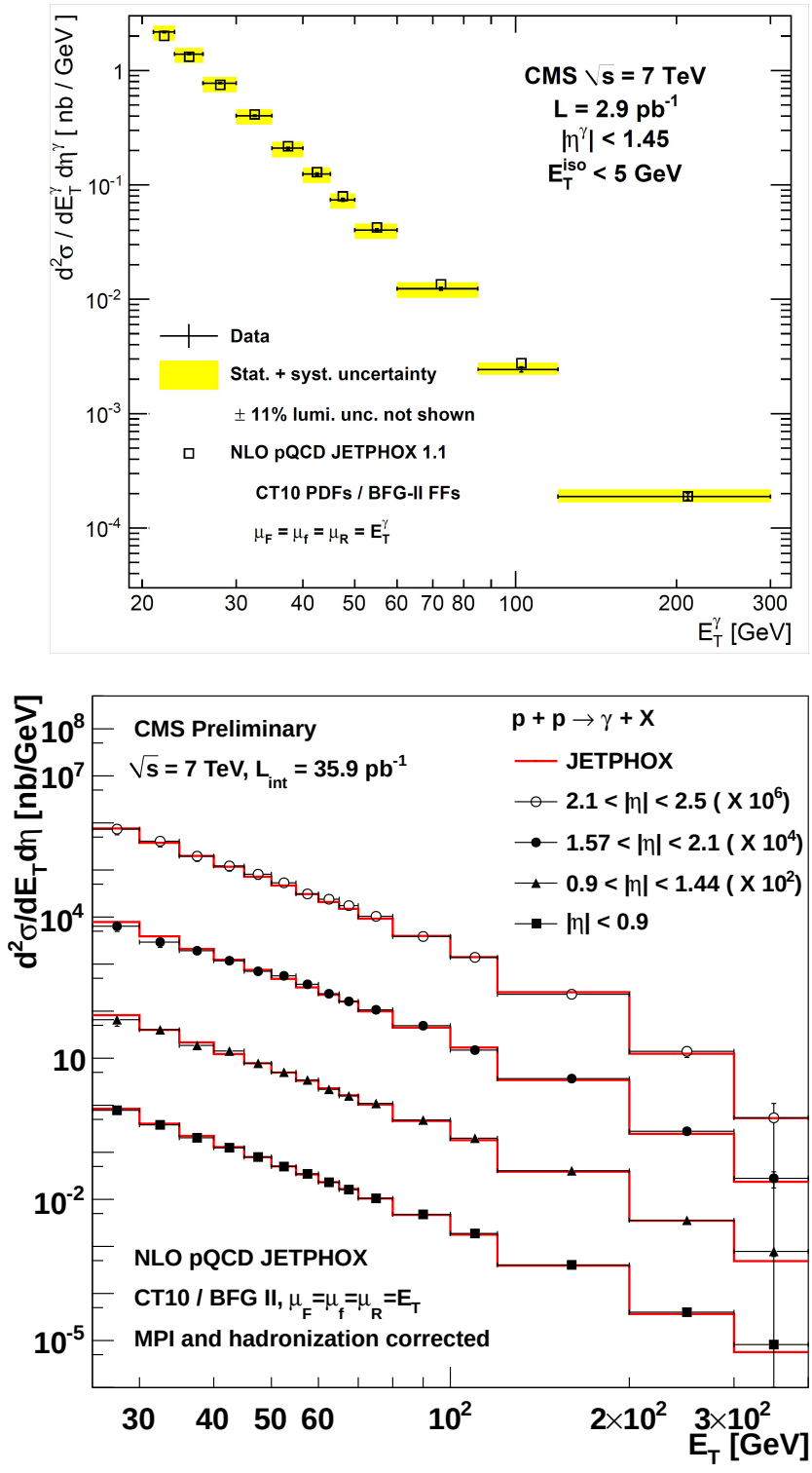

Figure 4.11: Top plot is the first measurement of the isolated prompt photon differential cross section and NLO pQCD predictions, as a function of photon E_T for photon $|\eta|$ <1.45 by the CMS experiment. Bottom plot is the last measurement of isolated inclusive prompt photon differential cross sections as a function of photon transverse energy E_T in the four pseudorapidity regions of the CMS electromagnetic calorimeter and the predictions from JETPHOX by the CMS experiment.

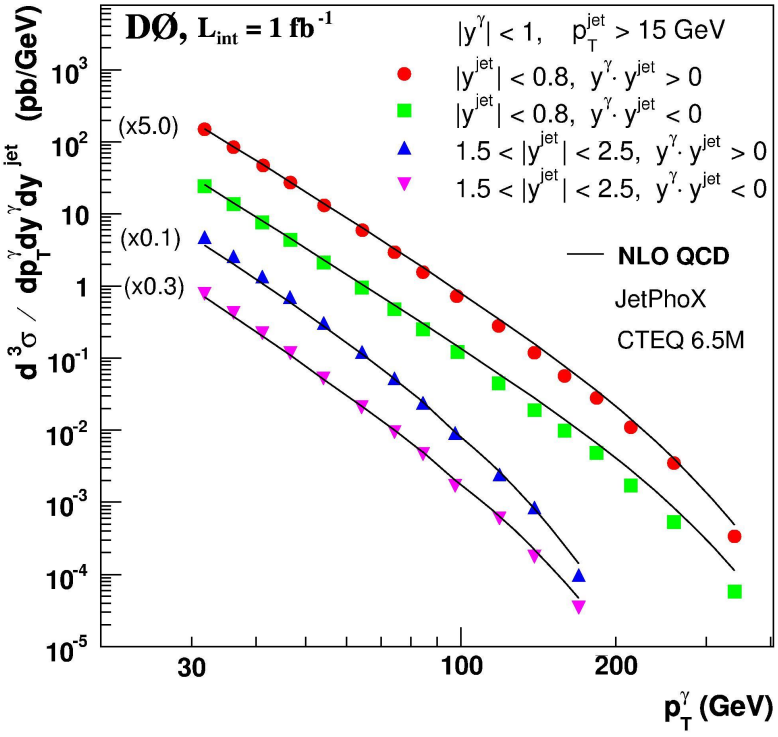

Figure 4.12: First measurement of the triple differential $p\overline{p} \rightarrow \gamma + jet + X$ cross section as a function of photon transverse momentum for the four measured rapidity intervals by the DØ collaboration. The data are compared to the theoretical NLO QCD predictions using the JETPHOX with the CTEQ6.5M PDF set and renormalization, factorization, and fragmentation scales are set to be $\mu_R = \mu_F = \mu_f = p_T(\gamma)f(y)$.

CHAPTER 5

PHOTON EFFICIENCY MEASUREMENTS WITH 2011 DATA

Efficiency measurement is one of the most important ingredient of particle detection in experimental high energy physics. Almost all the physics analyses rely on the detected particles such as photons, electrons, muons, and jets in an efficiently measured way. In a real detector environment, detection of the particles would not give the exact number of produced particles in the presence of background. This comes from the issue that particles could not be triggered, reconstructed, and identified as they are in reality. Therefore, efficiency measurement needs to be performed to determine the ratio of the detected particles to the actual particles produced. This chapter focuses on photon efficiency measurements with 2011 CMS data recorded at $\sqrt{s}$ = 7 TeV.

5.1 Tag and Probe Technique

The Tag and Probe (T&P) is a powerful data-driven technique developed to measure any user defined object efficiency from data at CMS by exploiting di-object resonances [70], [71]. Such resonances decay into two leptons, e.g., $Z \rightarrow l^+l^-$, $J/\psi \rightarrow l^+l^-$, or $Y \rightarrow l^+l^-$ (where $l = e, \mu$). In this technique, resonances are reconstructed as pairs with one leg passing a tight selection (tag) and the second leg passing a loose selection (probe). Passing probes are defined according to type of the efficiency to be estimated. Lineshapes of the passing probe and the failing probe (with tag object) are fit separately with a signal + background model. The efficiency is then computed from the ratio of the signal yields in these two lineshapes. This procedure is repeated in bins of the probe variables (e.g. p_T, η, ϕ,...) to compute efficiency histograms as a function of those variables.

5.1.1 T&P Software

The workflow for an efficiency measurement using the tag and probe software tool is splitted into two steps [72]. The first step is performed by using the *TagProbeFitTreeProducer* tool which defines the tags and the probes that are collections of *reco::Candidate* objects. The tags and probes are combined into pairs with the help of *CandViewShallowCloneCombiner* class. Probe variables and passing probe objects are defined in this step and stored as *ROOT* trees for the use in the later step. This step is achieved by running the configuration of *Photon_TagProbeTreeProducer_cfg.py*.

The second step is performed by using the *TagProbeFitTreeAnalyzer* tool. The tool takes the information obtained in the first step and calculates efficiencies by using different methods as a function of variable of interest in the form of histogram canvases. The results are stored as unbinned *RooDataSet* format in *ROOT* framework and also 1-D efficiency plots saved as *RooPlot* objects. In this step, highly configurable file *testTagProbeFitTreeAnalyzer_Zee.py* is used. Overall, tag and probe workflow is given as a drawing in Figure 5.1.

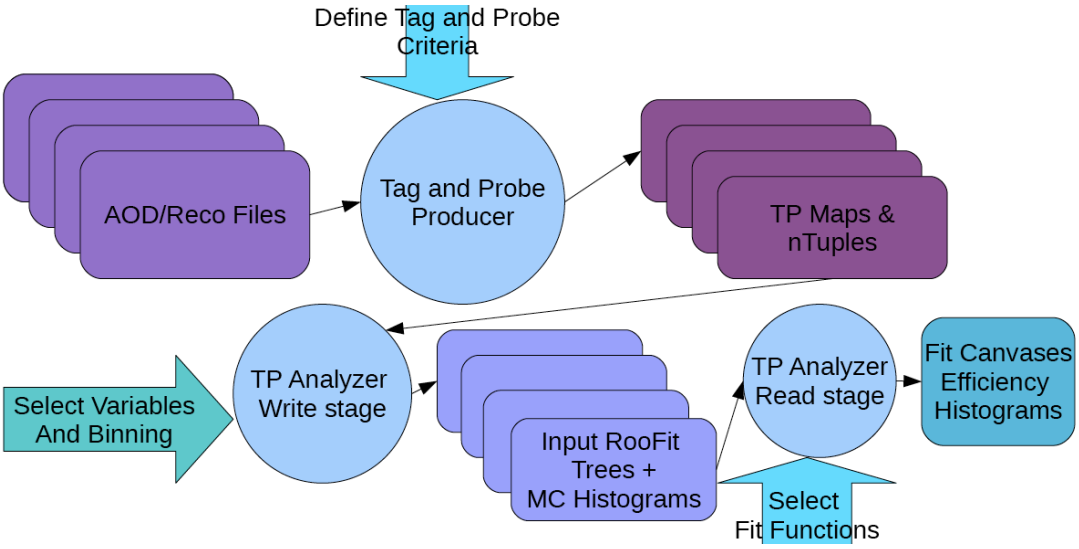

Figure 5.1: Drawing of the steps in efficiency measurement with the Tag and Probe software package.

5.1.2 Efficiency Calculation

In T&P technique, the Z boson mass constraint from the dielectron pair is taken into account to determine the efficiency. The tag is the electron passing a set of tight selection criteria and the probe is an electron supercluster (SC) passing photon selection criteria. Invariant mass of the tag and probe pair ($M_{e\gamma}$) is required to fit into the region of (M_Z) mass window. Aftermath, efficiency is calculated as taking the ratio of the probes passing the selection criteria to all probes,

$$\epsilon_{trigger} = \frac{N_{TP_P}}{N_{TP_P} + N_{TP_F}} \tag{5.1}$$

where N_{TP_P} and N_{TP_F} are the number of tag and probe events in which probe passes or fails certain selection criteria respectively. Selection criteria for probe photons will be discussed in more details for this analysis chapter where become relevant.

In efficiency calculation, if a probe also satisfies the tag selection, then there will be two tags and hence two probes. For this reason, there will be double counting of tag and probe pairs in such events. If we include these events where probes are also tags, then the efficiency formula in Eq. 5.1 should be modified in the following way:

$$\epsilon_{trigger} = \frac{2N_{TT} + N_{TP_P}}{2N_{TT} + N_{TP_P} + N_{TP_F}} \tag{5.2}$$

where the term N_{TT} is the number of tags and probes in which the probes pass also the tag selection. Two other terms N_{TP_P} and N_{TP_F} are the tag and probe pairs in which probes are also tags but pass and fail the selection criteria respectively. If there is more than one probe in some events, these events will be rejected since efficiency calculation can be affected in a biased way.

5.2 Trigger Efficiency

In this section, efficiencies of single photon HLT paths in 2011 data taking are given by using T&P technique on Z$\rightarrow e^-e^+$ events. Naming of 2011 single photon HLT paths has the form of HLT_PhotonE$_T$_WPs_version. Here WPs stands for working points decided by the CMS

Egamma HLT subgroup for electron and photon paths. WPs refer to sets of isolation cuts applied at the HLT level to keep the trigger rate under control as instaneous luminosity increases. WPs for photon triggers include either *CaloIdX* (X=VL,L,T or VT) or both *CaloIdX* and *IsoY* (Y=VL,L, or T) referring to photon object selection (isolation) at the HLT level. Photon object HLT WPs are summarized in Table 5.1. Introduced later in the 2011A data taking, some photon trigger paths with high E_T threshold do not contain WPs which meant that there is no need for further rate reduction for that paths.

Table 5.1: Egamma WPs (working points) for photon HLT paths.

WPs	Iso	CaloId Barrel	CaloId Endcap
VeryLoose 'VL'	Iso_{ECAL} <6.0+0.012×E_T Iso_{ECAL} <4.0+0.005×E_T Iso_{TRACK} <4.0+0.002×E_T	H/E<0.15 $\sigma_{i\eta i\eta}$ <0.024	H/E<0.10 $\sigma_{i\eta i\eta}$ <0.040
Loose 'L'	Iso_{ECAL} <5.5+0.012×E_T Iso_{ECAL} <3.5+0.005×E_T Iso_{TRACK} <3.5+0.002×E_T	H/E<0.15 $\sigma_{i\eta i\eta}$ <0.014	H/E<0.10 $\sigma_{i\eta i\eta}$ <0.035
Tight 'T'	Iso_{ECAL} <5.0+0.012×E_T Iso_{ECAL} <3.0+0.005×E_T Iso_{TRACK} <3.0+0.002×E_T	H/E<0.10 $\sigma_{i\eta i\eta}$ <0.011	H/E<0.075 $\sigma_{i\eta i\eta}$ <0.031
VeryTight 'VT'	NA	H/E<0.05 $\sigma_{i\eta i\eta}$ <0.011	H/E<0.05 $\sigma_{i\eta i\eta}$ <0.031

For photon and jet cross section measurement analysis with 2011 data, there are three sets of single photon HLT paths in use. The first set is signal trigger paths having the structure of HLT_PhotonE_T_CaloIdVL_IsoL (where E_T = 30 GeV, 50 GeV, 75 GeV, and 90 GeV). The second trigger set is background trigger paths with the structure of HLT_PhotonE_T_CaloIdVL (where again E_T = 30 GeV, 50 GeV, 75 GeV, and 90 GeV). The last set of trigger paths that are used in measurement of γ+jet triple cross section has the form of HLT_PhotonE_T (where E_T = 125 GeV and 135 GeV). Trigger paths in the third set are neither isolated nor cut on shower shape variable at the HLT level and their transverse energy thresholds are higher than the ones in the first and the second sets of trigger paths.

Single photon HLT path efficiency can directly be measured from collision data by using the Tag and Probe technique on Z→ e^-e^+ events [73]. In this technique, the tag is an electron which is well-reconstructed by a stringent electron selection criteria and the probe is a photon required that the electron supercluster (SC) to pass photon selection criteria.

5.2.1 Data-sets

This analysis is performed by using *4_2_8* version of *CMSSW* and *V04-00-09* version of Tag and Probe package. The 2011A data used in this analysis splits into four data-sets given in Table 5.2 with run ranges and accessible recorded luminosities. Data consists of both *PromptReco* and *ReReco* photon primary data-sets.

Table 5.2: Data-sets used, run ranges, and recorded luminosities in 2011A era.

Data-sets	Run range	Recorded lum. (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

To avoid complications in the efficiency measurement, only the run range for when the trigger path was unprescaled is used. Such as for the path *HLT_Photon50_CaloIdVL_IsoL*, it is unprescaled in the run range 161217-163261 and this run range corresponds to an effective luminosity of 40.95 pb^{-1}. The unprescaled run ranges and the effective luminosities for the single photon HLT paths in 2011A era are given in Table 5.3. We used the following most current official JSON files to have good luminosity sections from the runs:

- Cert_160404-163869_7TeV_May10ReReco_Collisions11_JSON_v3.txt,
- Cert_160404-177053_7TeV_PromptReco_Collisions11_JSON.txt,
- Cert_170249-172619_7TeV_ReReco5Aug_Collisions11_JSON_v2.txt.

Table 5.3: Unprescaled run ranges and corresponding effective luminosities for the photon HLT paths.

HLT trigger path	Run range	Eff. (pb^{-1})
HLT_Photon50_CaloIdVL_IsoL	161217-163261	40.95
HLT_Photon75_CaloIdVL	160431-163869	216.06
HLT_Photon75_CaloIdVL_IsoL	160431-165633	355.5
HLT_Photon90_CaloIdVL_IsoL	165088-167913	934.0
HLT_Photon125	165088-166967	665.70
HLT_Photon135	167039-173198	1053.98

5.2.2 Tag and Probe Selection

Tags are GsfElectrons that pass $E_T > 20$ GeV within the detector acceptance excluding the transition region (i.e. having $|\eta| < 1.4442$ and $1.566 < |\eta| < 2.5$). The tag electrons are required to pass the selection WP80 from the Vector Boson Task Force (VBTF) (Table 5.4). Such a selection has 80% signal efficiency [74].

Further, tag electrons are required to match to the single electron and the first leg of double electron trigger paths where they are unprescaled. The single and double electron HLT paths which are used in this analysis are listed in Table 5.5 and in Table 5.6 respectively.

Table 5.4: List of Working Point 80% electron selection criteria.

Variables	Selection in Barrel	Selection in Endcaps
Missing hits in inner pixel layers	0	0
Distance of the partner track	>0.02	>0.02
$\Delta cot\theta$ of the partner track	>0.02	>0.02
Track isolation in dR=0.3 / electron E_T	<0.09	<0.04
ECAL isolation in dR=0.3 / electron E_T	<0.07	<0.05
HCAL isolation in dR=0.3 / electron E_T	<0.10	<0.025
Shower shape: $\sigma_{i\eta i\eta}$	<0.01	<0.03
Track-super cluster matching: $\|\Delta\phi\|$	<0.06	<0.03
Track-super cluster matching: $\|\Delta\eta\|$	<0.004	-
H/E	<0.04	<0.025

Table 5.5: Single electron HLT paths and their unprescaled run ranges.

HLT trigger path	Run range
HLT_Ele27_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_v1	160431-161176
HLT_Ele27_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_v2	161217-163261
HLT_Ele27_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_v3	163270-163869

Probes are both barrel ($|\eta| < 1.4442$) and endcap ($1.566 < |\eta| < 2.5$) photons passing $E_T > 20$ GeV. Probe photons are further selected with the two set of photon identification requirements. For the single photon HLT paths on which both isolation and shower shape cuts are applied at HLT level (i.e. **_CaloIdVL_IsoL*), the following selection is employed at the analysis level:

Table 5.6: Double electron HLT paths and their unprescaled run ranges.

HLT trigger path	Run range
HLT_Ele32_CaloIdT_CaloIsoT_TrkIdT_TrkIsoT_SC17_v1	165088-165633
HLT_Ele32_CaloIdT_CaloIsoT_TrkIdT_TrkIsoT_SC17_v2	165970-166967
HLT_Ele32_CaloIdT_CaloIsoT_TrkIdT_TrkIsoT_SC17_v3	167039-167043
HLT_Ele32_CaloIdT_CaloIsoT_TrkIdT_TrkIsoT_SC17_v4	167078-167913
HLT_Ele32_CaloIdT_CaloIsoT_TrkIdT_TrkIsoT_SC17_v5	170722-173198

- H/E < 0.05
- Tracker Isolation < 2.0 + 0.002 $\times p_T$ GeV
- ECAL Isolation < 4.2 + 0.012$\times p_T$ GeV
- HCAL Isolation < 2.2 + 0.005$\times p_T$ GeV
- Shower shape ($\sigma_{i\eta i\eta}$) < 0.020 for ECAL Barrel ($|\eta| < 1.4442$)
- Shower shape ($\sigma_{i\eta i\eta}$) < 0.039 for ECAL Endcap ($1.566 < |\eta| < 2.5$).

For the single photon HLT paths on which only shower shape cut is applied at HLT level (i.e. **_CaloIdVL*), probe photons are cut on only H/E and $\sigma_{i\eta i\eta}$ variables in this analysis (this selection is employed also for the HLT paths without isolation and shower shape cuts at HLT level which are *HLT_Photon125* and *HLT_Photon135*):

- 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 ($1.566 < |\eta| < 2.5$).

Above photon selection cuts are applied at the analysis level and tighter than photon isolation and shower shape cuts which are implemented at HLT level. Finally, we impose an invariant mass cut on the tag-probe pair which fits into the mass window of the Z boson, 60 GeV < $M_{e\gamma}$ < 120 GeV in this study. This requirement ensures a high-purity sample of tag-probe pairs.

5.2.3 Method Description

The number of tag and probe events are determined by two methods in the Tag and Probe technique. The first one is the *counting method* in which the number of tag-probe pairs are counted as the name suggests, within an invariant mass near the Z peak (i.e. 60 - 120 GeV in this analysis). The second method is the *fitting method* that performs a fit to the tag-probe invariant mass to determine the amount of signal and background. In this study, both counting and fitting methods are used to calculate efficiencies of single photon HLT paths.

In the Tag and Probe HLT efficiency measurement, efficiency of an HLT path ($\epsilon_{trigger}$) is calculated as the ratio of the number of probes passing an HLT path (having an HLT object match) to the total number of probes (including both failing and passing an HLT path) in accord with Eq. 5.1 given previously. As in this measurement probe photons are reconstructed objects, thus $\epsilon_{trigger}$ represents the probability for a reconstructed signal photon to be selected by the HLT system.

5.2.4 Results

CMS trigger menus have been changed to reduce trigger rates regarding substantial increase of instaneous luminosity in 2011 data taking period. For the single photon trigger paths, this operation is accomplished by raising E_T thresholds and implementing isolation requirements (i.e. *CaloIdVL* and *IsoL*). Given high E_T thresholds and isolation cuts at HLT level, statistics of probe photons passing and failing the single photon HLT paths are studied. Numbers of probe photons passing and failing the corresponding HLT requirements are shown in Table 5.7.

Table 5.7: Number of probe photons passing and failing the HLT paths.

HLT path	E_T (GeV)	Passing	Failing	Total
HLT_Photon50_CaloIdVL_IsoL	> 50	1686	417	2103
HLT_Photon75_CaloIdVL	> 75	889	154	1043
HLT_Photon75_CaloIdVL_IsoL	> 75	1209	218	1427
HLT_Photon90_CaloIdVL_IsoL	> 90	1728	237	1965
HLT_Photon125	> 125	351	36	387
HLT_Photon135	> 135	624	80	704

Number of passing and failing photons are also compared in Barrel ($|\eta| < 1.4442$) and Endcap ($1.566 < |\eta| < 2.5$) regions seperately as a function of tag-probe invariant mass variable in Figure 5.2 and Figure 5.3.

HLT efficiency turn-on curves are further studied to understand where the single photon trigger paths become maximally efficient. For this purpose, single photon HLT effcencies as a function of photon E_T are given in Figure 5.4. The single photon trigger paths are observed to have a plateau which is 10 GeV above HLT E_T thresholds. Thus, single photons HLT paths should only be used 10 GeV above their online thresholds in cross section measurements or in any other analyses involving photon object.

In the following subsections, unprescaled photon HLT path efficiencies in 2011 data are reported by making use of two complementary methods which are *counting* and *fitting* (see estimated efficiencies of prescaled triggers in Appendix A). Efficiencies are measured 10 GeV above E_T thresholds at HLT level where we found that trigger paths are maximally efficient.

5.2.4.1 Counting Method

The efficiencies measured by the counting method for single photon HLT paths are given in Table 5.8 in bins of the probe photon's E_T and η. The uncertainity of the efficiency results is the binomial statistical error as reported from the Tag and Probe package.

In the counting method, isolated trigger paths are found to be less efficent than non-isolated trigger path *HLT_Photon75_CaloIdVL*. Non-isolated trigger path is found to be 100% efficent. While *HLT_Photon125* is fully efficient in both endcap and barrel, *HLT_Photon135* is fully efficient only in the barrel. Efficiency numbers are observed mostly to be larger in barrel than in endcap which can also mimic presence of more sufficient statistics in barrel than in endcap.

5.2.4.2 Fitting Method

The invariant mass of the tag-probe pair is fitted using the convolution of a Bereit Wigner and a Crystal-Ball, the default fitting method in the Tag and Probe package. The fits return the amount of signal and background, which are then used to compute efficiencies of the HLT paths.

Table 5.8: Efficiency of single photon HLT paths as a function of the probe photon E_T in barrel and endcaps as determined by the counting method.

HLT path	Probe E_{T} (GeV)	$-\eta$ Endcap	Barrel	$+\eta$ Endcap
HLT_Photon50_CaloIdVL_IsoL	60 - Inf.	93.5 ± 3.5	97.8 ± 0.8	96.4 ± 3.4
HLT_Photon75_CaloIdVL	85 - Inf.	100 ± 1.4	99.8 ± 0.3	100 ± 1.6
HLT_Photon75_CaloIdVL_IsoL	85 - Inf.	97.7 ± 2.2	97.3 ± 0.7	95.6 ± 2.7
HLT_Photon90_CaloIdVL_IsoL	100 - Inf.	95.6 ± 2.1	97.3 ± 0.5	95.1 ± 2.6
HLT_Photon125	135 - Inf.	100 ± 3.7	100 ± 0.4	100 ± 4.4
HLT_Photon135	145 - Inf.	92.6 ± 5.4	99.8 ± 0.3	94.8 ± 4.1

In Table 5.9, single photon HLT efficiencies are given from the fitting method in bins of probe photon's E_T and η. Non-isolated path *HLT_Photon75_CaloIdVL* is observed to be fully efficient in both barrel and endcap while isolated trigger paths are 96-100% efficient in the acceptance except *HLT_Photon50_CaloIdVL_IsoL*. HLT_Photon50 is unprescaled only for its first version and its later versions are quickly and highly prescaled so that there is no enough statistics for this path to have much reliable efficieny calculation. Isolated single photon trigger paths are more efficient in Barrel than in Endcap which is also related to number of events triggered. *HLT_Photon125* is fully efficient whereas *HLT_Photon135* is fully efficient in only barrel.

Table 5.9: Efficiency of single photon HLT paths as a function of the probe photon E_T in barrel and endcaps as determined by the fitting method.

HLT path	Probe E_{T} (GeV)	$-\eta$ Endcap	Barrel	$+\eta$ Endcap
HLT_Photon50_CaloIdVL_IsoL	60 - Inf.	93.5 ± 2.8	97.8 ± 1.4	96.4 ± 2.5
HLT_Photon75_CaloIdVL	85 - Inf.	100 ± 1.4	99.8 ± 0.2	100 ± 1.6
HLT_Photon75_CaloIdVL_IsoL	85 - Inf.	97.7 ± 1.6	97.5 ± 0.6	98.3 ± 1.8
HLT_Photon90_CaloIdVL_IsoL	100 - Inf.	95.6 ± 1.8	97.3 ± 0.5	100 ± 0.8
HLT_Photon125	135 - Inf.	100 ± 3.7	100 ± 0.4	100 ± 4.5
HLT_Photon135	145 - Inf.	92.5 ± 4.2	100 ± 0.6	95.0 ± 5.3

Dependancy of HLT efficiencies on photon η, photon ϕ angle, and event number of reconstructed primary vertices N_{PV} are also shown in Figure 5.5, Figure 5.6, and Figure 5.7 respectively which are obtained from the fitting method. In HLT efficiency distributions as a function of N_{PV}, it could be observed that efficiencies are not strongly affected by pile-up in collision data as these distributions are almost flat in number of reconstructed vertices. Therefore, no pile-up subtraction from hard collision data is needed to apply at this stage.

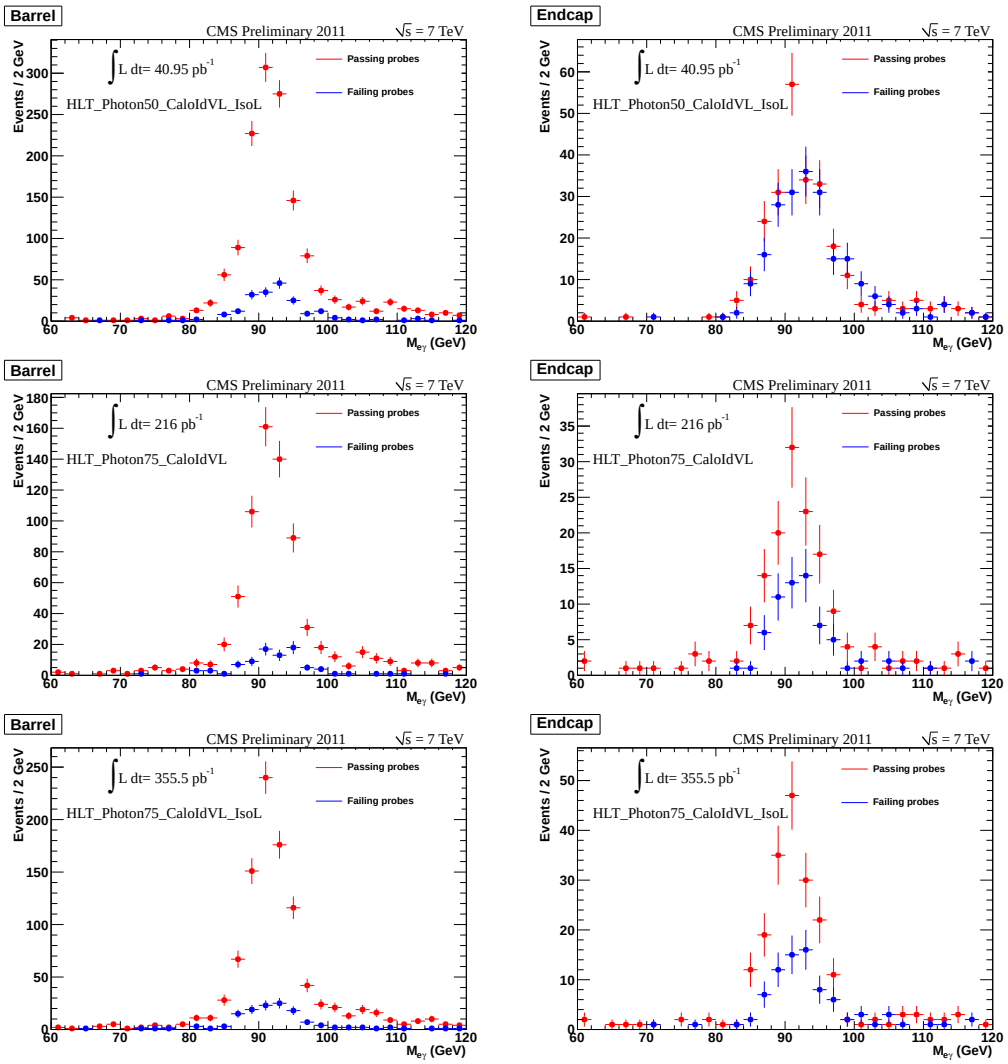

Figure 5.2: Invariant mass distributions for events when probe photons pass and fail to have an HLT object match to HLT paths (for 50 GeV and 75 GeV threshold triggers) in Barrel (left) and in Endcap (right).

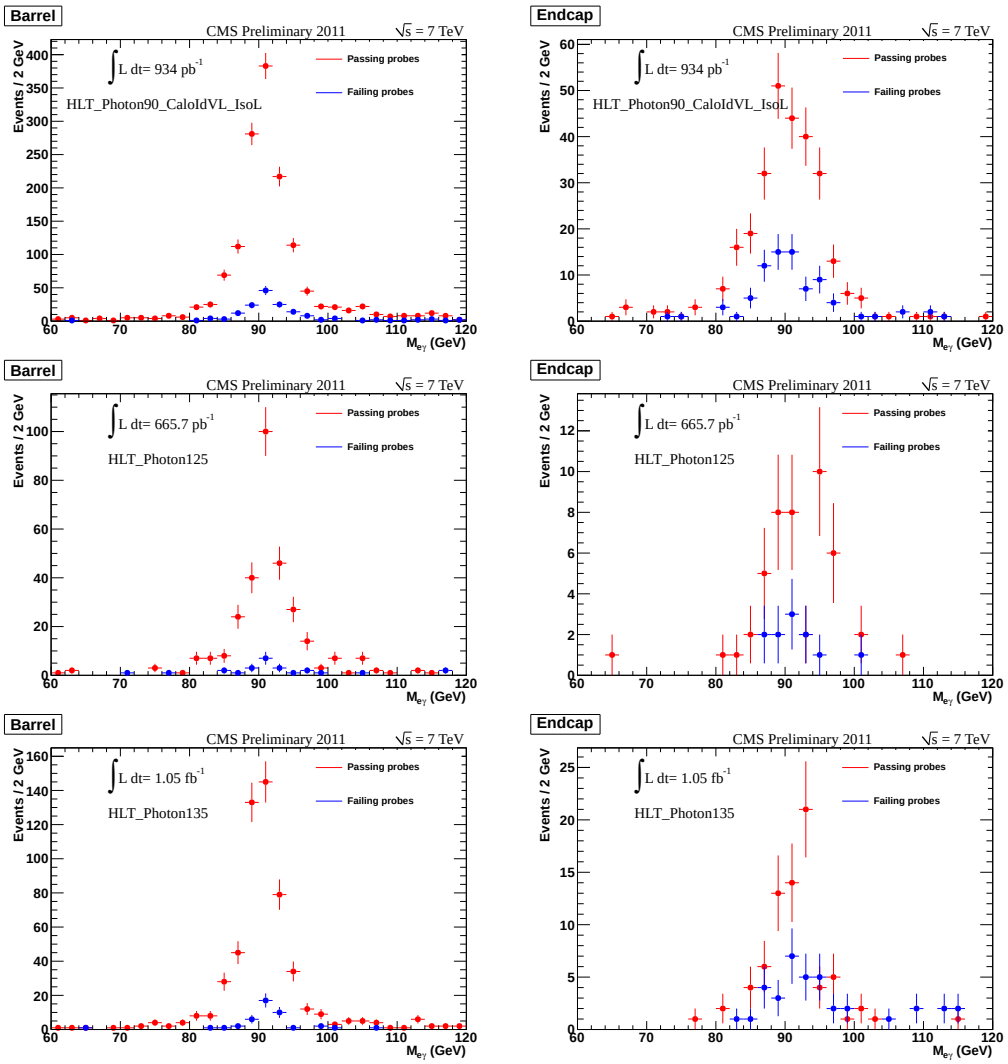

Figure 5.3: Invariant mass distributions for events when probe photons pass and fail to have an HLT object match to HLT paths (for 90 GeV, 125 GeV, and 135 GeV thresholds triggers) in Barrel (left) and in Endcap (right).

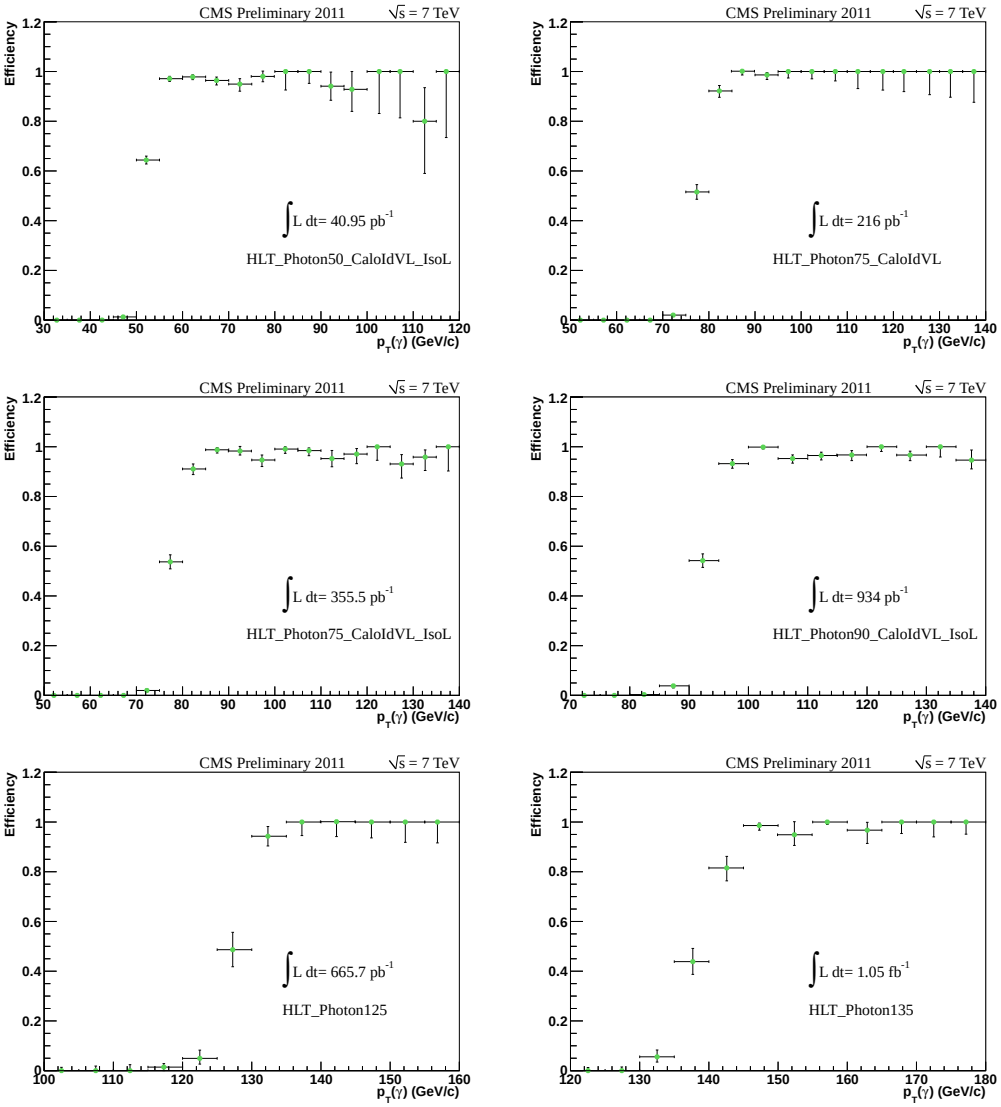

Figure 5.4: Dependancy of single photon HLT efficiencies on probe photon's E_T (so-called turn-on curves).

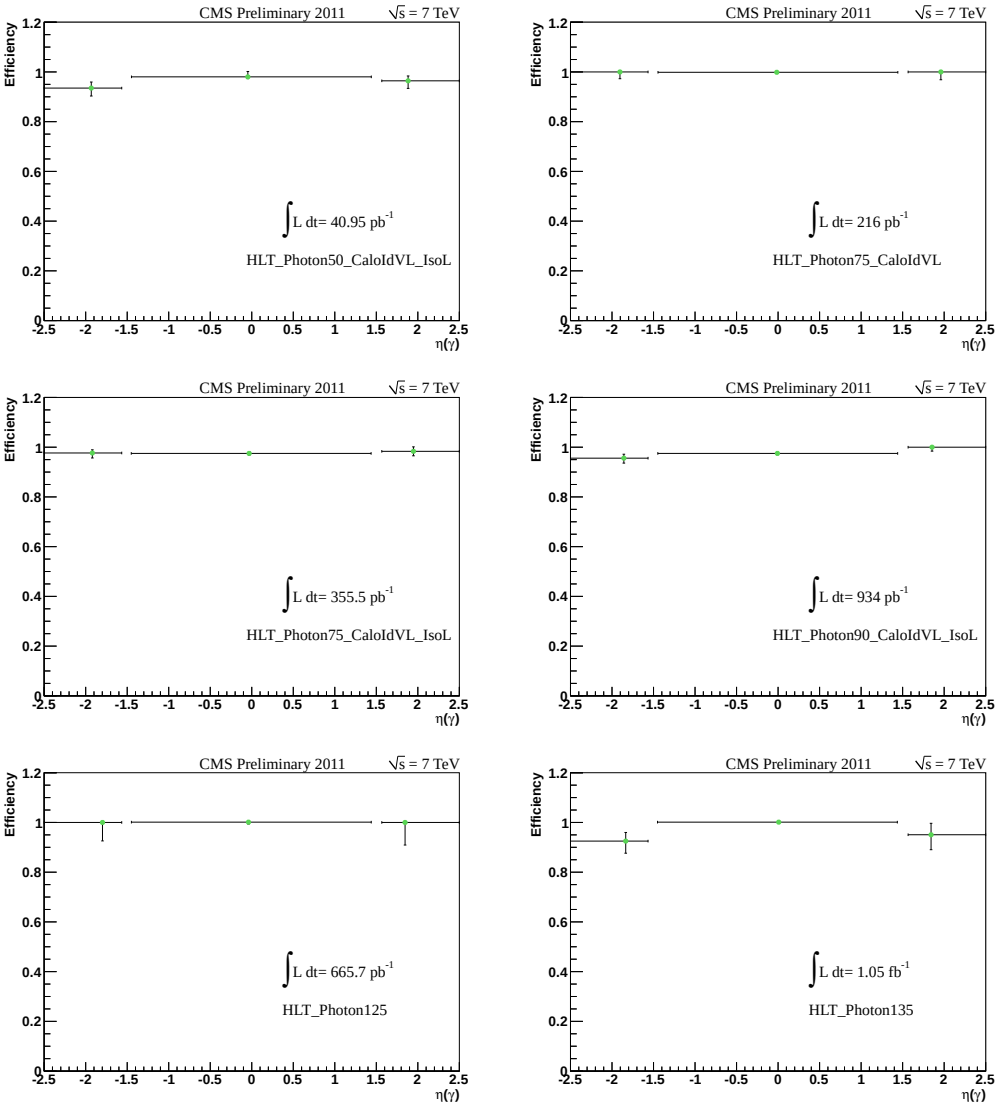

Figure 5.5: Dependancy of single photon HLT efficiencies on probe photon's η by the fitting method.

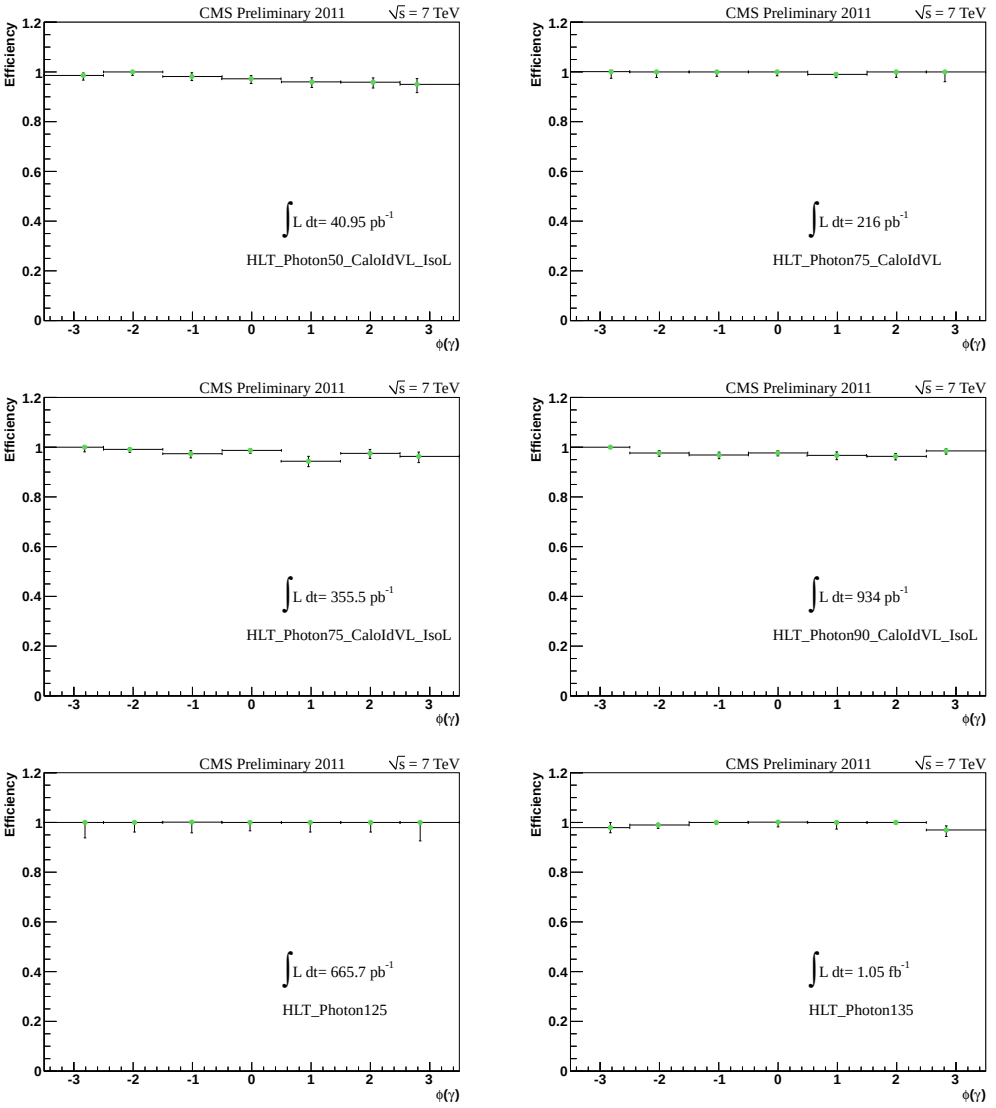

Figure 5.6: Dependancy of single photon HLT efficiencies on probe photon's ϕ by the fitting method.

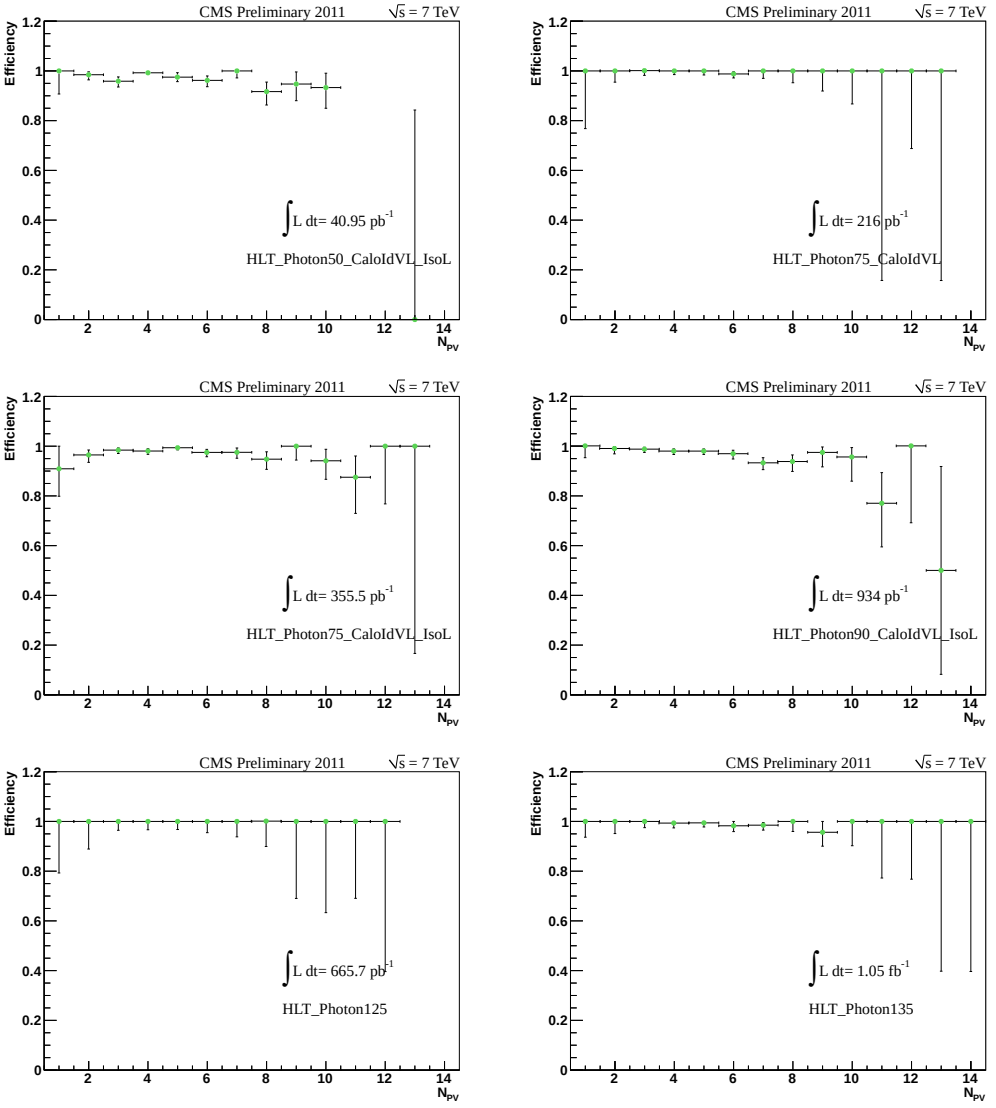

Figure 5.7: Dependancy of single photon HLT efficiencies on event number of primary vertices N_{PV} by the fitting method.

5.3 Reconstruction Efficiency

Given Monte Carlo generated spectrum in terms of photon p_T and η, $G(p_T, \eta)$ and generated spectrum with a matched reconstructed photon, $G_R(p_T, \eta)$, photon reconstruction efficiency (which is also known as geometrical acceptance efficiency in high energy physics experiments) is given by the ratio

$$\epsilon_{RECO} = \frac{G_R(p_T, \eta)}{G(p_T, \eta)}. \tag{5.3}$$

In above equation, ϵ_{RECO} is measured in generated photon spectrum and true photons that are generated pass generation level isolation requirement ($Iso_{GEN} < 5$ GeV) additionally. In this equation, ϵ_{RECO} stands for the probability for a signal photon produced inside the detector geometrical acceptance to be reconstructed by the clustering algorithms.

Efficiency of photon reconstruction (RECO) is measured by using PYTHIA γ+jet simulation samples from Table 5.10 and cross-confirmed by MadGraph γ+jet samples from Table 5.11. Both the samples are *Summer11* production in the CMS experiment with the same tune of *Z2*. PYTHIA samples are binned in $\hat{p}_T$, while MadGraph samples are divided in bins of H_T which is the quantity representing scalar sum of transverse momenta of the jets. Additionally, PYTHIA samples with different bins are weighted when combining for the efficiency measurements as different samples have different number of events processed with different cross sections. Cross section is divided to number of events in order to account for proper weighting of the samples.

Measured photon reconstruction efficiencies are given in Figure 5.8 and Figure 5.9 as a function of generated photon p_T in different generated photon pseudorapidity regions from PYHTIA and Madgraph samples respectively. Efficiency is found higher in inner Barrel ($|\eta| < 0.9$) than other η ranges from both types of simulation samples. Reconstruction efficiencies are also summarized in bins of generated photon p_T and pseudorapidity in Table 5.12 and Table 5.13.

Table 5.10: Monte Carlo PYTHIA photon+jets samples used in efficiency measurement.

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

Table 5.11: Monte Carlo Madgraph photon+jets samples used in efficiency measurement.

MC dataset name	H_T range
MadGraph γ+jet Z2 tune (AODSIM)	
/GJets_TuneZ2_40_HT_100_7TeV-madgraph/Summer11-PU_S4_START42_V11-v1	40-100
/GJets_TuneZ2_100_HT_200_7TeV-madgraph/Summer11-PU_S4_START42_V11-v1	100-200
/GJets_TuneZ2_200_HT_inf_7TeV-madgraph/Summer11-PU_S4_START42_V11-v1	200-∞

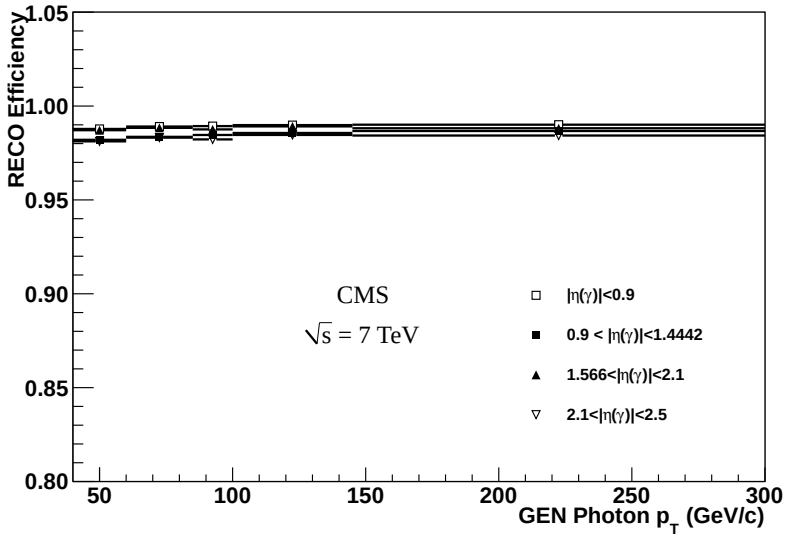

Figure 5.8: PYHTIA Monte Carlo photon reconstruction efficiencies as a function of generated photon p_T in four generated photon η regions.

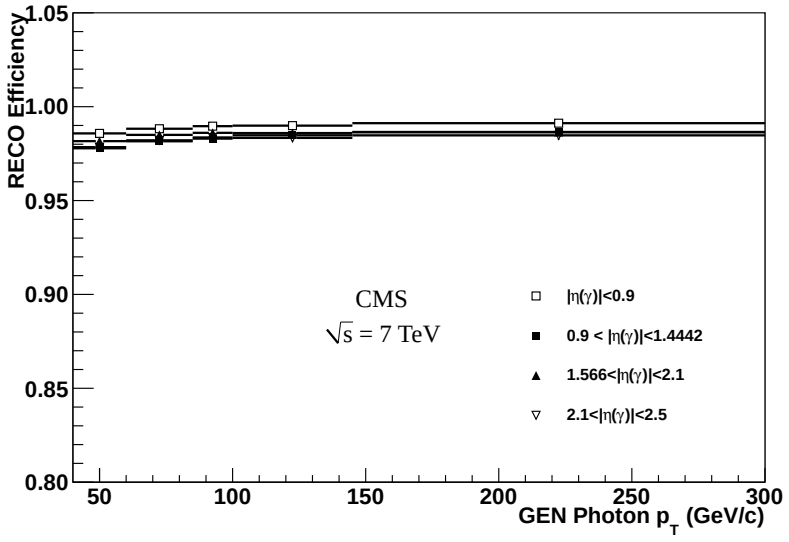

Figure 5.9: MadGraph Monte Carlo photon reconstruction efficiencies as a function of generated photon p_T in four generated photon η regions.

Table 5.12: Efficiency of photon reconstruction in bins of generated photon p_T and η measured from PYTHIA samples.

p_T (GeV)	$\|\eta\| < 0.9$	$0.9 < \|\eta\| < 1.4442$	$1.566 < \|\eta\| < 2.1$	$2.1 < \|\eta\| < 2.5$
40-60	98.78 ± 0.02	98.20 ± 0.04	98.71 ± 0.03	98.12 ± 0.05
60-85	98.90 ± 0.02	98.36 ± 0.03	98.84 ± 0.03	98.32 ± 0.04
85-100	98.93 ± 0.02	98.46 ± 0.03	98.76 ± 0.03	98.22 ± 0.04
100-145	98.98 ± 0.02	98.56 ± 0.03	98.91 ± 0.02	98.46 ± 0.04
145-300	99.00 ± 0.01	98.67 ± 0.02	98.81 ± 0.02	98.43 ± 0.03

Table 5.13: Efficiency of photon reconstruction in bins of generated photon p_T and η measured from MadGraph samples.

p_T (GeV)	$\|\eta\| < 0.9$	$0.9 < \|\eta\| < 1.4442$	$1.566 < \|\eta\| < 2.1$	$2.1 < \|\eta\| < 2.5$
40-60	98.57 ± 0.02	97.78 ± 0.02	98.16 ± 0.02	97.84 ± 0.03
60-85	98.82 ± 0.02	98.16 ± 0.04	98.50 ± 0.03	98.20 ± 0.05
85-100	98.96 ± 0.05	98.30 ± 0.08	98.60 ± 0.07	98.36 ± 0.10
100-145	98.99 ± 0.05	98.47 ± 0.08	98.60 ± 0.08	98.34 ± 0.12
145-300	99.12 ± 0.07	98.65 ± 0.11	98.64 ± 0.13	98.47 ± 0.22

5.4 Identification Efficiency

Identification efficieny is measured in reconstructed spectrum of photon, $R(p_T, \eta)$. Requiring that photon reconstructed spectrum to pass given photon identification (photon ID) selection, $R_{ID}(p_T, \eta)$, photon ID efficiency is computed as

$$\epsilon_{ID} = \frac{R_{ID}(p_T, \eta)}{R(p_T, \eta)}. \tag{5.4}$$

Here, ϵ_{ID} gives the probability for a reconstructed signal photon to pass the photon identification criteria. Photon ID efficiency in this study is measured from both simulation samples (both PYTHIA and MadGraph) and collision data by using Tag and Probe technique. The ratio of data ID efficiencies to MC ID efficiencies are called scale factors which are highly important for correcting cross section measured from data to take MC truth information into account.

5.4.1 Simulation Results

Efficiency of photon identification (ID) is measured by using both PYTHIA γ+jet simulation samples from Table 5.10 and MadGraph γ+jet simulation samples from Table 5.11. At the analysis level, two different identification templates (ID sets, which is named as also signal extraction templates in this thesis) are studied: shower shape template and isolation template. In each template selection, photon p_T is required to be greater than 40 GeV and $H/E < 0.05$ cut is also in place. Shower shape template photon ID selection is given as follows:

- H/E < 0.05
- Tracker Isolation < 2.0 + 0.002*p_T GeV
- ECAL Isolation < 4.2 + 0.012*p_T GeV
- HCAL Isolation < 2.2 + 0.005*p_T GeV
- Shower shape ($\sigma_{i\eta i\eta}$) < 0.020 for ECAL Barrel ($|\eta| < 1.4442$)
- Shower shape ($\sigma_{i\eta i\eta}$) < 0.039 for ECAL Endcap ($1.566 < |\eta| < 2.5$).

For the isolation template, photon candidates are extracted by only cutting on shower shape variable, cut values of the isolation template is listed below:

- 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 ($1.566 < |\eta| < 2.5$).

Monte Carlo photon ID efficiencies from shower shape template selection are demonstrated from PYTHIA MC samples in Figure 5.10 and from Madgraph MC samples in Figure 5.11 as a function reconstructed photon p_T in four photon pseudorapidity regions. Efficiencies are monotonically increasing with increasing photon transverse momentum from PYTHIA samples, while ID efficiency measured from MadGraph samples has a monotnically increasing trend up to photon transverse momentum of 150 GeV and then a decreasing trend to be stable for higher bins. It is also observed that ID efficiency results from Madgraph samples are less reliable for the last photon p_T bin (300 - 500 GeV) due to insufficient number of events in that bin. ID efficiency is the highest in outer barrel according to the both simulation studies.

MC photon ID efficiencies of isolation template selection are shown from PYTHIA MC samples in Figure 5.12 and from Madgraph MC samples in Figure 5.13 as a function reconstructed photon p_T in four photon pseudorapidity regions. In the isolation template, MadGraph samples suffer from low statististics, especially at the highest photon transverse momentum bin of 300-500 GeV. ID efficiencies are found to be increasing monotonically with increasing photon transverse momentum from PYTHIA samples, while this is different in efficiencies measured from MadGraph in the sense that efficiency distributions are almost flat up to 300 GeV and then deacreasing for higher values of transverse momentum.

In photon analyses of the CMS experiment, both shower shape and isolation templates are powerful selections that are used to identify photons for efficiency and cross section studies. Moreover, ID efficiency numbers are reported in bins of reconstructed photon p_T and pseudorapidity in Table 5.14 and Table 5.15 for shower shape template and in Table 5.16 and Table 5.17 for isolation template from both types of MC results.

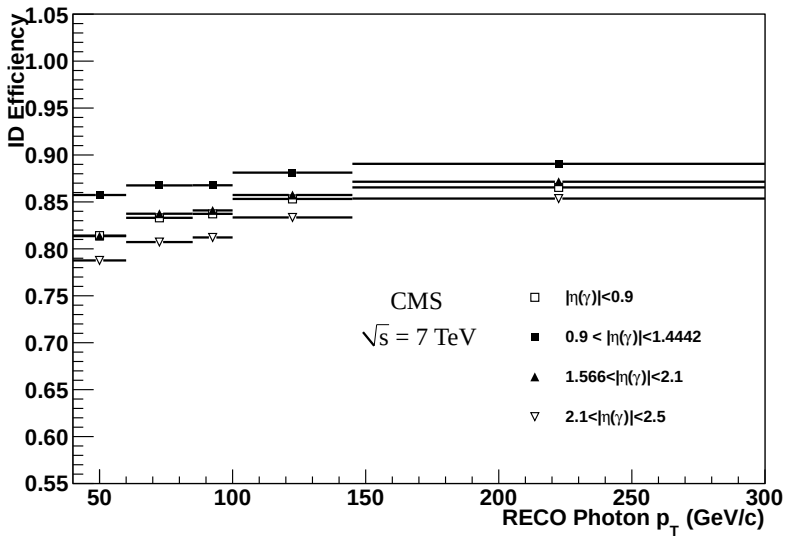

Figure 5.10: Shower shape template: Monte Carlo photon ID efficiencies as a function of reconstructed photon p_T in four photon η regions from PYTHIA samples.

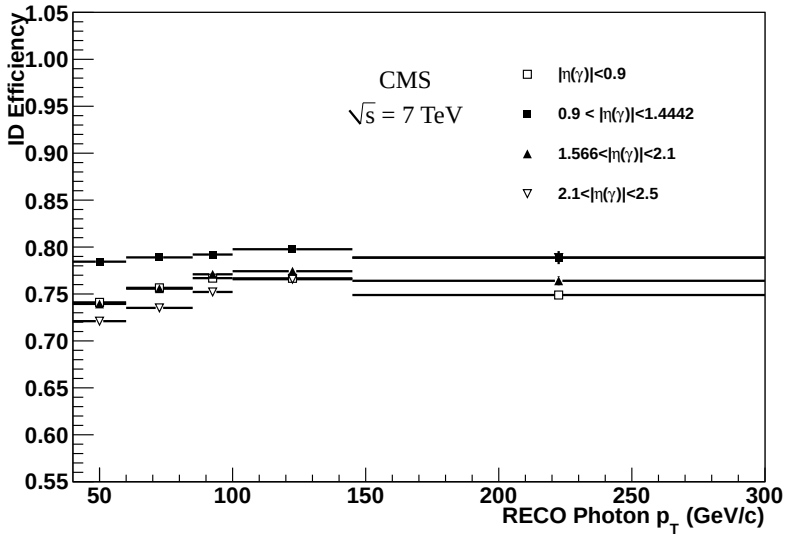

Figure 5.11: Shower shape template: Monte Carlo photon ID efficiencies as a function of reconstructed photon p_T in four photon η regions from Madgraph samples.

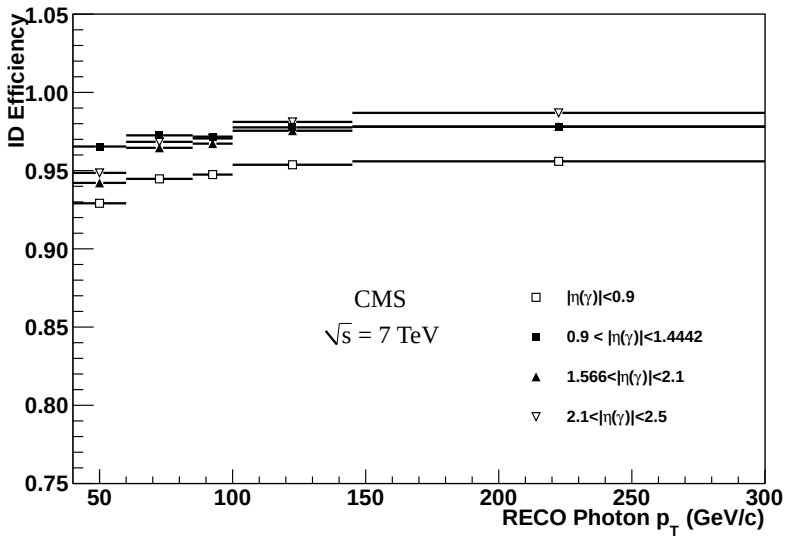

Figure 5.12: Isolation template: Monte Carlo photon ID efficiencies as a function of reconstructed photon p_T in four photon η regions from PYTHIA samples.

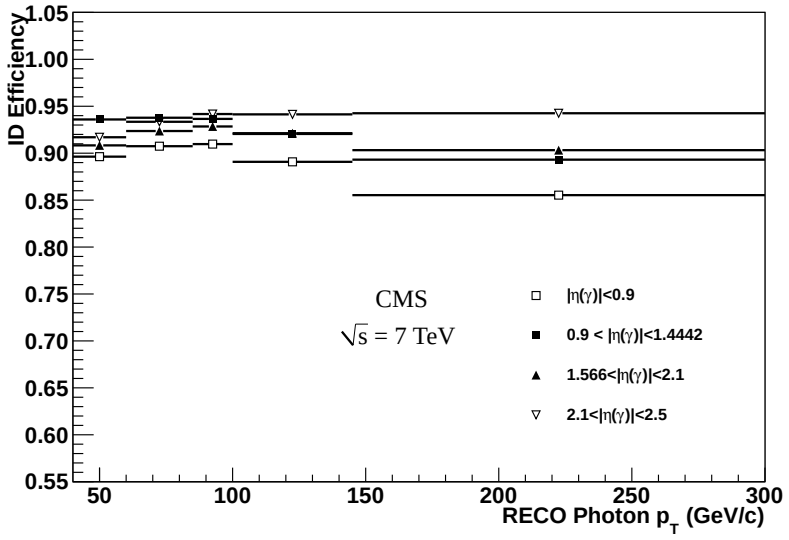

Figure 5.13: Isolation template: Monte Carlo photon ID efficiencies as a function of reconstructed photon p_T in four photon η regions from Madgraph samples.

Table 5.14: Shower shape template photon ID efficiencies in bins of reconstructed photon p_T and η from PYTHIA samples.

p_T (GeV)	$\lvert\eta\rvert < 0.9$	$0.9 < \lvert\eta\rvert < 1.4442$	$1.566 < \lvert\eta\rvert < 2.1$	$2.1 < \lvert\eta\rvert < 2.5$
40-60	81.41 ± 0.08	85.74 ± 0.10	81.36 ± 0.11	78.77 ± 0.14
60-85	83.30 ± 0.07	86.76 ± 0.08	83.74 ± 0.10	80.72 ± 0.12
85-100	83.73 ± 0.07	86.77 ± 0.09	84.10 ± 0.10	81.21 ± 0.13
100-145	85.32 ± 0.06	88.12 ± 0.07	85.74 ± 0.08	83.35 ± 0.11
145-300	86.55 ± 0.04	89.06 ± 0.05	87.17 ± 0.06	85.36 ± 0.10

Table 5.15: Shower shape template photon ID efficiencies in bins of reconstructed photon p_T and η from MadGraph samples.

p_T (GeV)	$\lvert\eta\rvert < 0.9$	$0.9 < \lvert\eta\rvert < 1.4442$	$1.566 < \lvert\eta\rvert < 2.1$	$2.1 < \lvert\eta\rvert < 2.5$
40-60	74.10 ± 0.05	78.44 ± 0.06	73.94 ± 0.07	72.10 ± 0.08
60-85	75.65 ± 0.10	78.90 ± 0.12	75.56 ± 0.12	73.52 ± 0.17
85-100	76.69 ± 0.22	79.20 ± 0.27	77.10 ± 0.30	75.21 ± 0.40
100-145	76.67 ± 0.21	79.77 ± 0.25	77.42 ± 0.28	76.57 ± 0.40
145-300	74.89 ± 0.30	78.87 ± 0.35	76.40 ± 0.47	78.87 ± 0.69

Table 5.16: Isolation template photon ID efficiencies in bins of reconstructed photon p_T and η from PYTHIA samples.

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

Table 5.17: Isolation template photon ID efficiencies in bins of reconstructed photon p_T and η from MadGraph samples.

p_T (GeV)	$\|\eta\| < 0.9$	$0.9 < \|\eta\| < 1.4442$	$1.566 < \|\eta\| < 2.1$	$2.1 < \|\eta\| < 2.5$
40-60	89.63 ± 0.04	93.60 ± 0.04	90.83 ± 0.05	91.70 ± 0.05
60-85	90.74 ± 0.06	93.78 ± 0.06	92.36 ± 0.07	93.35 ± 0.09
85-100	90.97 ± 0.13	93.65 ± 0.12	92.85 ± 0.14	94.17 ± 0.17
100-145	89.08 ± 0.12	92.05 ± 0.11	92.13 ± 0.13	94.14 ± 0.15
145-300	85.53 ± 0.18	89.31 ± 0.18	90.32 ± 0.24	94.25 ± 0.22

5.4.2 T&P Data Results

Photon ID efficiencies are also measured from 2011A data (given previously in Table 5.2) by using Tag and Probe technique. In this measurement, numerator and denominator photons are required to pass *HLT_Photon30_CaloIdVL_IsoL* path and required that photon p_T to be greater than 40 GeV in the acceptance. Numerator photons are required also to pass photon ID selections from shower shape template and isolation template. Photon ID Tag and Probe efficiencies from 2011A data are given in Table 5.18 and Table 5.19.

ID efficiency from data depends much on statistics after cutting by the photon ID and trigger requirements. For that reason, photon transverse momentum bins are reduced by merging into three bins to get as much reliable results as could be. Further, isolation template ID efficiencies are observed to be higher than shower shape ID efficiencies in data which is consistent with what measured from the simulation sanmples that are presented before.

Table 5.18: Efficiency of shower shape template photon ID selecton in bins of reconstructed photon p_T and η from Tag and Probe data.

Photon p_{T} (GeV)	Barrel	Endcap
40-45	89.0 ± 1.3	90.1 ± 3.5
45-50	91.0 ± 1.6	92.3 ± 3.7
50-Inf.	89.4 ± 2.0	86.1 ± 5.3

Table 5.19: Efficiency of isolation template photon ID selecton in bins of reconstructed photon p_T and η from Tag and Probe data.

Photon p_T (GeV)	Barrel	Endcap
40-45	95.1 ± 1.3	92.2 ± 3.2
45-50	94.4 ± 1.2	97.1 ± 2.5
50-Inf.	92.5 ± 1.6	100 ± 2.3

5.5 Sytematic Uncertainities

The main systematic uncertainty assigned for trigger efficiencies comes from the difference between measured HLT efficiencies and 100% efficiency. As non-isolated HLT paths are almost fully efficient in the acceptance which assures their uncertainties are negligible, hence the only uncertainty is given to the isolated HLT paths which are used in the shower shape template selection. Such uncertainty varies in Barrel and Endcap regions. Assigned uncertainties for trigger efficiency are summarized in percentage in Table 5.20.

We also assign systematic uncertainty for Photon ID efficiency results. Here, uncertainty is assigned for the difference between MC ID efficiency results and those measured from data by using the Tag and Probe technique. The uncertainty in this scope varies in different photon pseudorapidity regions and photon E_T ranges. In Table 5.21 and Table 5.22, uncertainty ranges are given in percentage as a function of photon E_T in Barrel and Endcap regions.

Table 5.20: Systematics uncertainties for trigger efficiency in Barrel ($|\eta| < 1.4442$) and Endcap ($1.566 < |\eta| < 2.5$) regions for isolated and non-isolated HLT paths used in the analysis.

Source	Photon E_T(GeV)	Barrel	Endcap
Isolated HLT paths	50-Inf.	0 – 2.2%	0 – 1.2%
Non-isolated HLT paths	75-Inf.	< 0.1%	< 0.1%

Table 5.21: Systematic uncertainty of shower shape template photon ID selection efficiency in bins of reconstructed photon p_T in Barrel ($|\eta| < 1.4442$) and Endcap ($1.566 < |\eta| < 2.5$) regions.

Photon p_T (GeV)	Barrel	Endcap
40-45	4.7%	5.4%
45-50	5.4%	6.2%
50-Inf.	0.1 – 4.6%	0.1 – 5.6%

Table 5.22: Systematic uncertainty of isolation template photon ID selection efficiency in bins of reconstructed photon p_T in Barrel ($|\eta| < 1.4442$) and Endcap ($1.566 < |\eta| < 2.5$) regions.

Photon p_T (GeV)	Barrel	Endcap
40-45	0.1 – 2.5%	0.1 – 3.0%
45-50	0.1 – 1.2%	0.1 – 3.5%
50-Inf.	0.1 – 4.5%	0.1 – 5.0%

CHAPTER 6

DIFFERENTIAL PHOTON-JET CROSS SECTION MEASUREMENT WITH 2011 DATA

Measurement of the triple differential cross section for the γ+jet final state is presented using data collected with the CMS detector from proton-proton collisions at the LHC with $\sqrt{s}$ = 7 TeV. The photons and jets are reconstructed with a maximum pseudorapidity of $|\eta^{\gamma}, \eta^{jet}| < 2.5$ and required to have transverse momenta in the ranges of $40 < p_T^{\gamma} < 300$ GeV and $p_T^{jet} > 30$ GeV, respectively. The differential cross section $d^3\sigma/dp_T^{\gamma}d\eta^{\gamma}d\eta^{jet}$ is determined from measurements using 64 exclusive kinematic regions of photon and jet. The final results are corrected by photon efficiency in detection, photon purity in signal extraction, and unfolding factors in detector reconstruction effects. Results are compared to theoretical predictions provided by the Pythia and the Sherpa event generators.

6.1 Analysis Flow

The analysis of the γ+jet triple differential cross section is presented as a function of photon transverse momentum (p_T^{γ}), photon pseudorapidity (η^{γ}), and jet pseudorapidity (η^{jet}) with 2.19 fb^{-1} of data collected during the 2011 data taking at CMS [53]. The differential cross section is calculated from the Eq. 4.8 (which is given already in Chapter 4).

6.1.1 Data Samples

We use real collision data which is given in Table 5.2 and Monte Carlo simulation samples of Pythia and Madgraph γ+jets which are given in Table 5.10 and in Table 5.11, respectively. For

photon purity studies, Pythia QCD di-jet samples are also used. The Pythia MC samples are generated primarily using PYTHIA version 6.4 to leading order ([27]) for a center-of-mass energy of 7 TeV and CTEQ6L ([77]) is used as parton distribution function (PDF) set.

6.1.2 Event Selection

Events used for this cross section measurement are selected by the two-level trigger system of the CMS detector. Both *Level-1* (L1) trigger and the *High-level* trigger (HLT) are based on energy deposited in the ECAL. The clustering algorithm at L1 is based on the energy in an array of 5x5 crystals in the ECAL, while a sophisticated clustering algorithm is used for the HLT (the reconstructed objects at HLT are called the super-clusters). A threshold of 20 GeV is required for 5x5 crystal energy deposits at L1. Five different HLT paths, corresponding to different E_T thresholds for super-clusters are used in the analysis. They are given in Table 6.1 together with their run ranges, recorded and effective luminosities, number of events passing offline selection and a E_T range for which they are used.

Table 6.1: List of HLT paths used in the analysis. In trigger names, given in first column, numbers 30, 50, 75, 90, and 135 represent supercluster E_T thresholds in GeV. Second column gives the run range during which the trigger was active. Recorded luminosities, and effective luminosities in parentheses, are in third column. Number of events passing offline selection is given in fourth column. Last column gives E_T ranges in GeV for which the triggers are used.

HLT paths	Run range	Rec.(Eff.) (pb^{-1})	Events	p_T range
HLT_Photon30_CaloIdVL_v1-7	160431-173692	2190.8(3.53)	1208139	40-60
HLT_Photon30_CaloIdVL_v1-3 HLT_Photon50_CaloIdVL_v1-4	160431-165087 165088-173692	2190.8(11.67)	1231920	60-85
HLT_Photon75_CaloIdVL_v1-7	160431-173692	2190.8(276.93)	1756132	85-100
HLT_Photon75_CaloIdVL_v1-3 HLT_Photon90_CaloIdVL_v1-4	160431-165087 165088-173692	2190.8(361.80)	1844488	100-145
HLT_Photon75_CaloIdVL_v1-3 HLT_Photon125_v1-2 HLT_Photon135_v1-2	164031-165087 165088-166967 167039-173692	2190.8(2190.6)	1220065	145-300

For photon selection, a pixel seed match veto is applied on reconstructed photon candidates and only the isolation template is used. As discussed previously in Chapter 4, photon candidates are cut on shower shape variable in the isolation template selection which is listed below:

- 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).

In this study, jets with cone radius of $\Delta R = 0.5$ are reconstructed with $anti - k_t$ clustering algorithm ([21]) using the particle flow objects as the input. Jets with a p_T >25 GeV (before applying jet energy corrections) are considered and are required to pass loose selection cuts to remove dedector noise. The following selection variable, also known as *Loose Jet Id*, are used for the selection [75]:

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

As particle flow reconstruction involves summing of ECAL and HCAL towers hence 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 Loose Jet Id requirements are taken into account. In this study, the same jet selection criteria is used both for data and simulation events.

6.1.3 Purity Determination

Photon selection is designed to enhance the fraction of photons in a sample of photon candidates but it can not completely remove background events, such that photon purity is less than 100 %. For that reason the number of photons (N^{γ} from Eq. 4.8) is obtained by the

template method which can be summarized as follows. A variable is chosen that has different distributions for signal (in our case photon) and background events. Templates of signal and background distributions for the chosen variable are obtained from Monte Carlo and/or in a data-driven way. Finally, distribution of photon candidates is fitted with a sum of signal and background templates to obtain the fraction of photons. This procedure is described in detail in Section 4.3.2 of Chapter 4. By template fitting method, the final result of purity estimation is shown in Figure 6.1. The purity of photon candidates are shown to be increasing as with photon transverse momentum ranges.

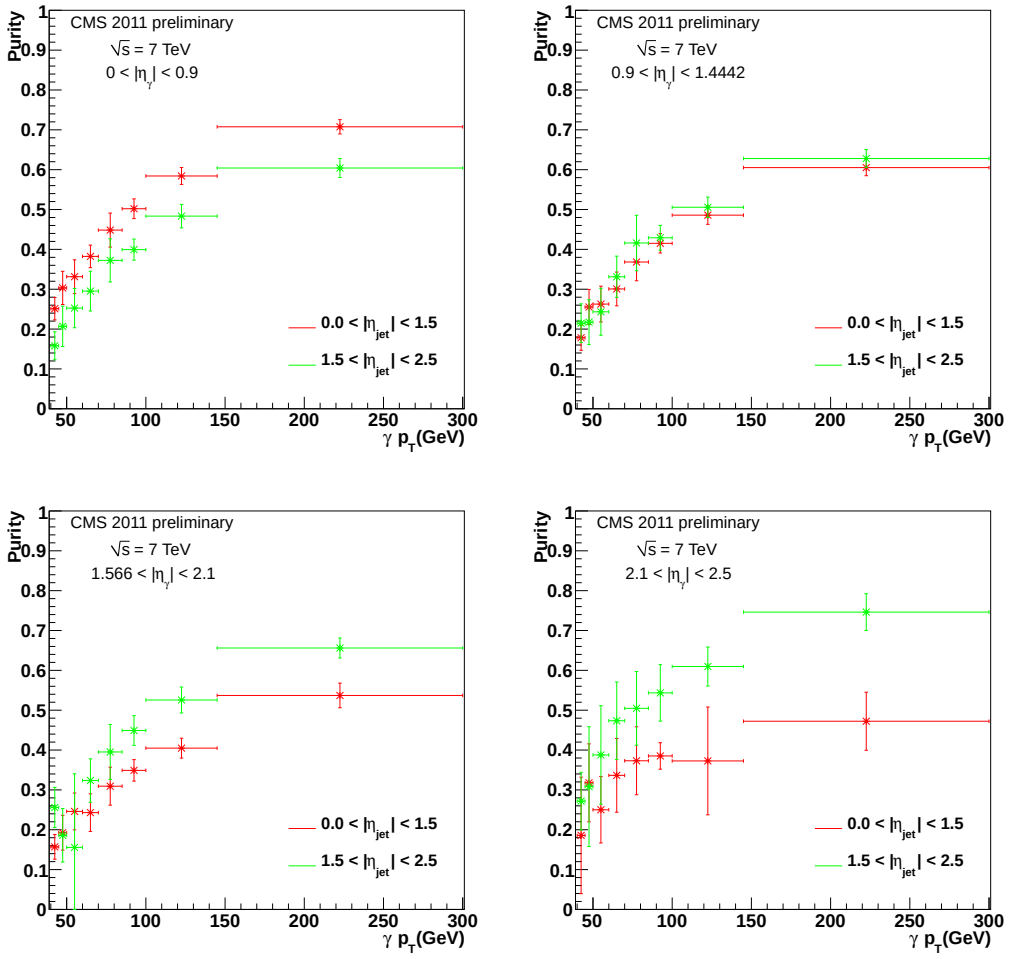

Figure 6.1: Four figures of signal purity versus photon p_T that are corresponding to different photon pseudorapidity regions. In each figure, the red line refers to the events with jet located in the ECAL barrel and the green line refers to the events with jet located in the ECAL endcap.

In purity determination, γ + *jet* Monte Carlo samples are used to generate signal template. These Monte Carlo samples have been generated with a flat plus poisson tail distribution for

the number of pile-up interactions, which does not exactly match the conditions expected for real data recorded by the CMS dedector. The corresponding number of pile-up interactions distributions of data and Monte Carlo are shown as the left plot of Figure 6.2. In order to factorize these effects, MC samples are re-weighted in terms of the number of pile-up interaction from the simulation truth to the target pile-up distribution for data, which is derived by using the per bunch-crossing-per-luminosity section instantaneous luminosity together with the total proton-proton inelastic cross-section to generate pileup distribution, correctly weighted by the per-bunch-crossing-per-lumisection integrated luminosity over the entire data-taking period. The right plot in Figure 6.2 shows the reweighting factor for γ + *jet* Monte Carlo samples.

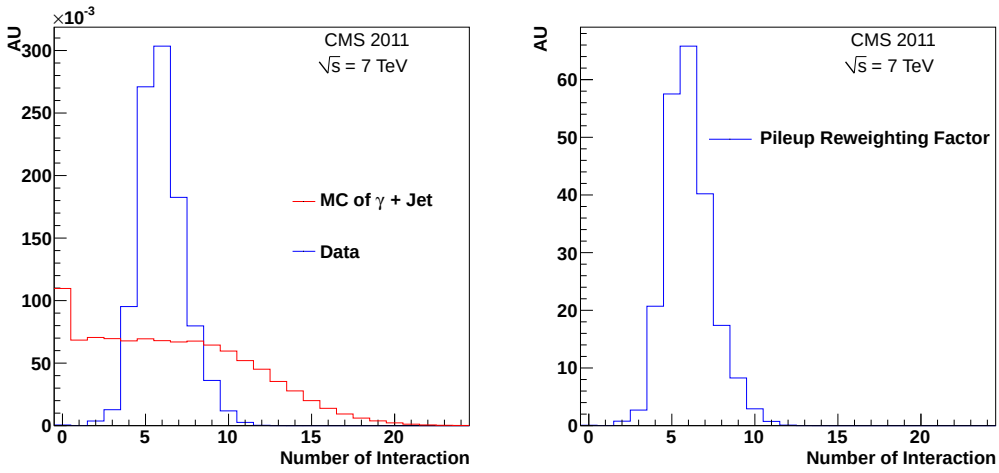

Figure 6.2: The distribution for the number of pileup interactions is shown on the left. Blue line refers to the distribution of data and red line refers to the distribution of Monte Carlo. The re-weighting factor as a function of number of interaction is shown on the right.

6.1.4 Total Efficiency

Photon efficiency results from 2011 data and MC simulation samples wherever possible are given extensively in Chapter 5 of this thesis. The efficiency for detection of photons are factorized into three different terms and measured independently. The trigger selection efficiency is measured with a data-driven method using the electrons from the decay of Z bosons. This efficiency factor is close to 100 %. The reconstruction efficiency is measured from the MC using γ+jet MC samples generated with Pythia. The same sample is used to determine ID,

the efficiency of the photon identification selection for isolation template. An additional correction factor is assigned based on the difference in data and MC from applying a similar selections to electrons in a Z-boson enriched sample. The total efficiency as a function of the transverse momentum of the photon is given in Figure 6.3.

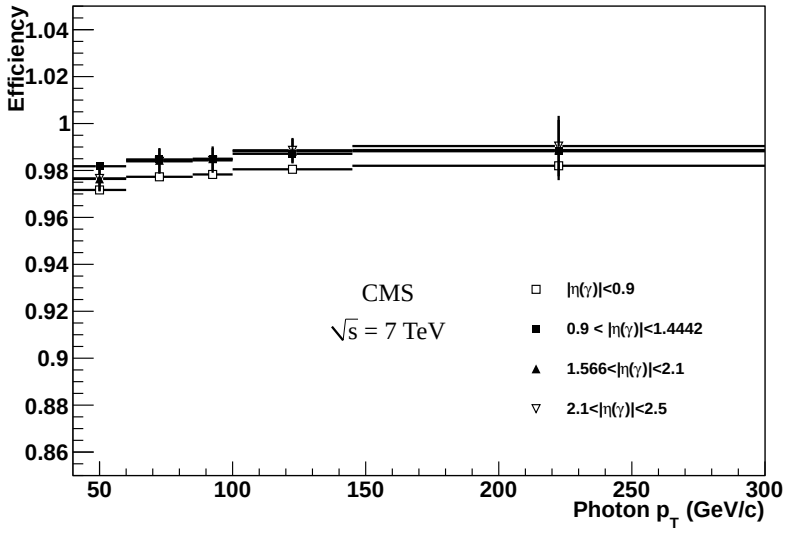

Figure 6.3: Total photon efficiency as a function of photon transverse momenta $p_T(\gamma)$ in four photon pseudorapidity (η) regions. The erros bars correspond to statistical uncertainities only.

6.1.5 Unfolding

Photon yield vaules for cross section calculation are obtained for bins defined by the *measured* values of p_T^γ, η^γ, and η^{jet}. These measured values may slightly differ from the *true* values as a result of detector effects (resolution, calibration etc.). The last step in the analysis is to correct for these effects and to obtain cross sections in bins defined by the true values of p_T^γ, η^γ, and η^{jet}. This procedure is called *unfolding*. Unfolding can only be done with simulation samples where both true and measured values are known. We used the same γ+jet MC samples and selection requirements as for the rest of the analysis. Photons yields defined by the true values of p_T^γ, η^γ, and η^{jet} are found as a function of photons yields defined by measured values. This mapping is then applied to the measured data to get the unfolded cross section results.

The unfolding procedure is performed with a dedicated software package *RooUnfold* ([78]) which is frequently used for such tasks. The package provides several methods for 1D (for

example if cross section would be calculated only as a function of p_T^γ) unfolding. For final results iterative (*Bayesian*) method ([79]) is used which is the only one implemented for 3D unfolding. As a consistency check, the result was compared with a less precise but robust bin-by-bin correction. The bin-by-bin correction only takes the ratio of the measured and true photon yields in each bin as a correction factor and therefore does not depend on measurement.

The same photon transverse momentum and pseudorapidity binnings are used for both true and measured variables and within the acceptance as defined by the experiment. Outside of the experimental acceptance, additional bins were used: for p_T^γ 30-40 GeV and 300-500 GeV; for η^γ 2.5-2.6; for η^{jet} 2.5-2.7.

The correctness of the procedure is verified by dividing MC samples into two statisticaly independent sets of equal size. Only first set is used for obtaining the unfolding correction while the second one is used for testing: test-measured distribution was corrected and compared to the test-truth. This closure test was passed well by iterative method (and of course by bin-by-bin correction which is a trivial result). An example result of the test is shown in Figure 6.4, where the ratio of corrected-measured and true distributions is given as a function of p_T^γ for the most inner photon and jet pseudorapidity bins. For this case the obtained unfolding correction (corrected-measured divided by measured) is given in Figure 6.4.

The uncertainty on unfolding correction can be divided in two parts. First one comes from the limited size of MC samples and is thus statistical in nature. The other source of uncertainty comes from the unavoidable difference between MC (which is used to obtain the unfolding) and data (on which the unfolding is applied).The statistical component of uncertainty is well understood. For iterative method it was calculated by performing pseudo-experiments: measured distributions were assumed Gaussian and varied according to the given statistical error. The effect was that after unfolding relative statistical uncertainty of cross section increased by approximately 15%.

For bin-by-bin correction, the exact statistical uncertainty was calculated separately and for closure tests it was found that the two uncertainties are comparable, with the one for iterative method being slightly higher and slightly increasing with the number of iterations. In this case the statistical component of uncertainty was found to vary from less than 1% to approx. 3% depending on the bin.

Iterative method takes into account the difference between distributions used for obtaining the unfolding correction and distributions on which the unfolding correction is applied. Convergence is monitored by checking the difference between sucessive iterations. For closure tests (where trained and test distributions differ only statistically) only one iteration was sufficient.

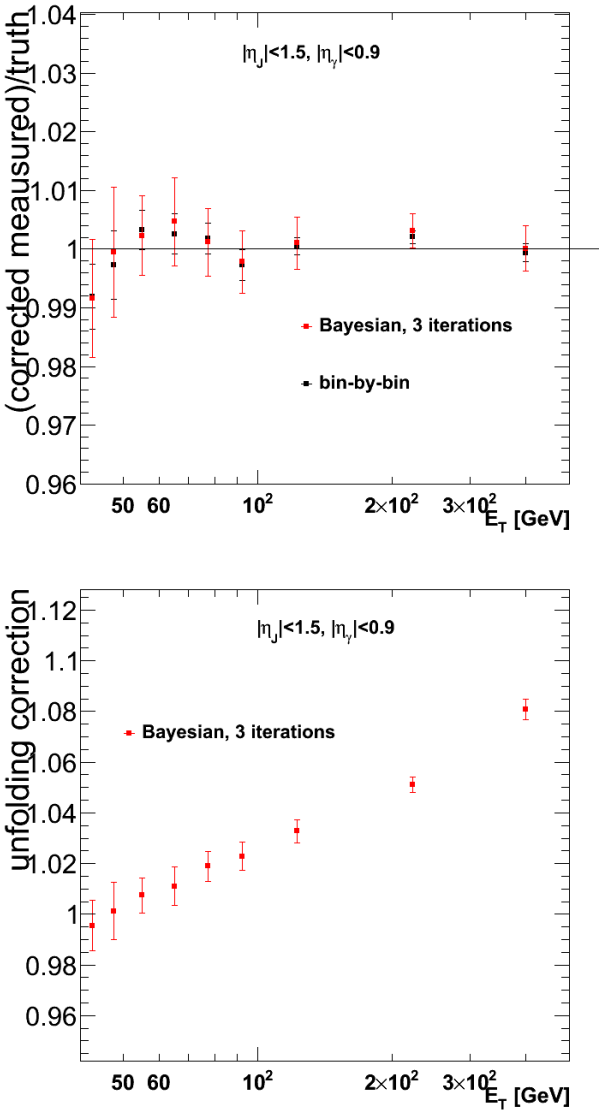

Figure 6.4: The top plot shows unfolding correction test results. The bottom plot shows unfolding correction defined as corrected-train-measured distribution divided by train-measured for bin-by-bin correction, matrix inversion an iterative method with one iteration. For bin-by-bin correction, only statistical uncertainty is shown.

6.2 Theory Predictions

Theoretical predictions for γ+jet processes are obtained from two leading order Monte Carlo generators which are Pythia [27] and Sherpa [80]. Although these generator tools are briefly discussed back in Chapter 2, some outstanding features for cross section measurement need to be emphasized.

Pythia is a traditional leading order MC generator which generates only 2→2 hard processes while any additional jet production comes from the parton showering. The prompt photon processes are produced (Pythia switch MSEL = 10) with a lower $\hat{p_T}$ cut of 20 GeV being applied on the outgoing partons (Pythia switch CKIN(3) = 20). Here, $\hat{p_T}$ is defined as the p_T of the outgoing partons in a 2→2 process in the center-of-mass frame relative to the axis defined along the trajectory of the incoming partons.

Sherpa (Simulation of High Energy Reactions of Particles) can generate hard processes with higher number of outgoing particles in the matrix element. Sherpa calculations include higher order tree level matrix elements. Proper phase space coverage is ensured in Sherpa by a well established jet matching algorithm, known as CKKW merging algorithm [81], which has been extensively validated with previous collider data. Further, the technique has been extended to processes involving photons [82]. The tree level matrix elements of variable photon and QCD parton multiplicity are combined with QCD+QED parton shower. So on top of the leading order matrix elements the parton shower produces interleaved QCD+QED emissions, thus treating the QCD and QED corrections fully democratically. Sherpa provides two kinds of matrix element generators, AMEGIC++ and Comix. Comix is used for the calculation since it is the default large multiplicity matrix element generator for Sherpa. In the Sherpa, events are generated with version 1.2.2. and γ+jet events were generated with upto 3 extra jets. The lower $\hat{p_T}$ cut of 20 GeV is applied on the photon and a minimum separation of $\Delta R = 0.5$ is also required.

After event generation, a similar set of selection criteria were applied on the events obtained from Pythia and Sherpa. The selection criteria are similar to the ones applied to the data sample. The selection criteria excludes the transition gap between barrel and endcap ($1.4442 \leq |\eta^{\gamma}| \leq 1.566$) for photon and includes the following kinematical and geometrical cuts on photon and jet objects:

- $p_T^{\gamma} \geq 20$ GeV,
- $|\eta^{\gamma}| \leq 2.5$,
- $1.4442 \leq |\eta^{\gamma}| \leq 1.566$ is excluded from the measurement,
- $p_T^{jet} \geq 20$ GeV,
- $|\eta^{jet}| \leq 2.5$,
- ΔR (γ,jet) ≥ 0.5,

Additionally, a generator level isolation criteria is also applied on the photon, where hadronic energy around the photon candidate in a cone size of 0.4 is restricted to 5.0 GeV.

6.3 Total Systematic Uncertainity

Total systematic uncertainty comes mainly from efficiency, purity, and unfolding steps in the analysis flow. Differences of variable shapes in simulation and data samples are assigned as systematics. For purity, main systematics come from difference between default and fitting results ([53]). Combined total systematics coming from different steps in the analysis are shown in Figure 6.5 for barrel photons and Figure 6.6 for endcap photons. Systematic uncertainty increases with decreasing p_T^{γ} and for low p_T^{γ} is dominated by the contribution from purity calculation.

6.4 Results

Presented here are the triple differential cross sections for events with γ+jet in the final state. Cross sections with the statistical and systematic uncertainties are given in the Figures 6.7- 6.8 and in Tables 6.2- 6.9, where they are compared to Sherpa and Pythia predictions. Results are shown for four different photon pseudorapidity and for two different jet pseudorapidity regions. Photon and jet transverse momenta are required to be in the ranges of $40 < p_T^{\gamma} < 300$ GeV and $p_T^{jet} > 30$ GeV, respectively.

Cross section results are given for 64 kinematic regions together with photon+jet geometrical orientations. The final cross section results are corrected by taking into account the corrections

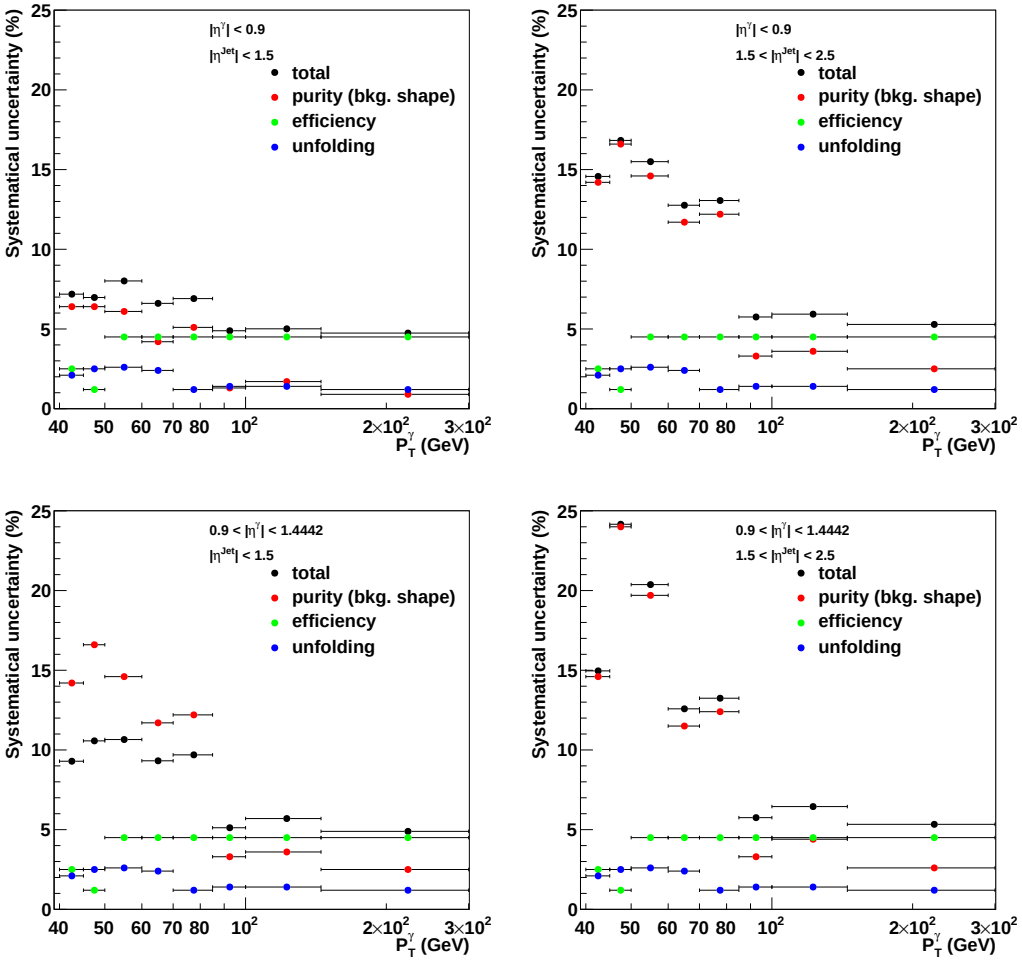

Figure 6.5: Total relative systematic uncertainty with contributions from purity, efficiency, and unfolding calculations as a function of photon transverse momentum for barrel photons.

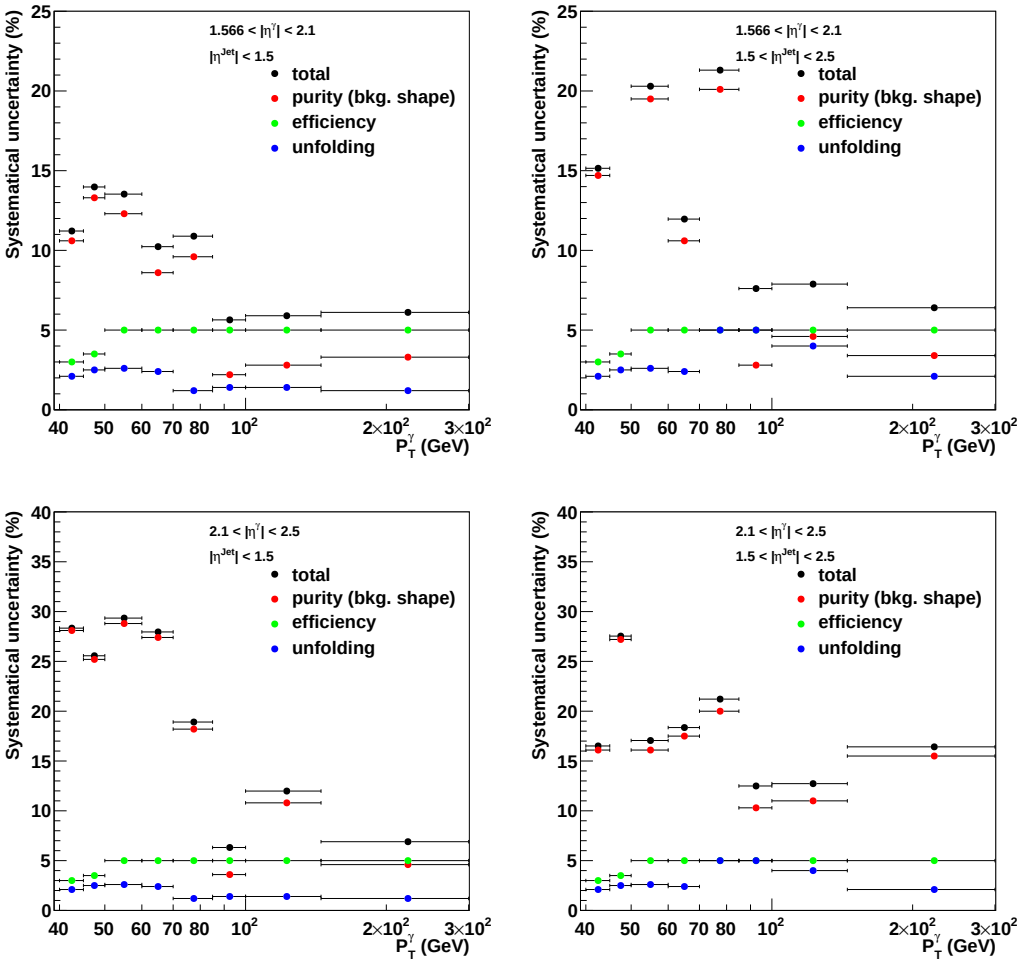

Figure 6.6: Total relative systematic uncertainty with contributions from purity, efficiency, and unfolding calculations as a function of photon transverse momentum for endcap photons.

obtained from photon efficiency in detection, photon purity in signal extraction, and unfolding factors in detector reconstruction effects. Though data and MC results are consistent for all kinematic regions, Sherpa prediction is observed to be in better agreement with the data results.

Table 6.2: Measured triple differential cross section values (pb/GeV) in bins of photon transverse momentum with statistical and systematic uncertainties compared to Sherpa and Pythia predictions ($|\eta^{\gamma}| < 0.9$ and $|\eta^{jet}| < 1.5$).

p_T^{γ} (GeV)	Data (D)	Sherpa (S)	Pythia (P)	D/S	D/P	S/P
40-45	27.4±1.1±2.0	29.9	18.8	0.92±0.08	1.46	1.59
45-50	21.5±0.8±1.5	20.0	12.2	1.08±0.09	1.76	1.64
50-60	10.8±0.4±0.9	10.9	6.63	1.00±0.09	1.63	1.64
60-70	5.20±0.16±0.34	5.10	3.09	1.02±0.08	1.68	1.65
70-85	2.65±0.08±0.18	2.42	1.55	1.10±0.09	1.71	1.56
85-100	1.10±0.01±0.05	1.06	0.664	1.04±0.08	1.66	1.59
100-145	0.350±0.003±0.017	0.367	0.209	0.95±0.07	1.67	1.76
145-300	0.0318±0.0002±0.0015	0.0268	0.0201	1.19±0.14	1.58	1.33

Table 6.3: Measured triple differential cross section values (pb/GeV) in bins of photon transverse momentum with statistical and systematic uncertainties compared to Sherpa and Pythia predictions ($|\eta^{\gamma}| < 0.9$ and $1.5 < |\eta^{jet}| < 2.5$).

p_T^{γ} (GeV)	Data (D)	Sherpa (S)	Pythia (P)	D/S	D/P	S/P
40-45	13.2±1.0±1.9	14.0	6.91	0.94±0.16	1.91	2.02
45-50	9.68±0.88±1.63	8.66	4.70	1.12±0.22	2.06	1.84
50-60	5.17±0.40±0.80	5.38	2.51	0.96±0.17	2.06	2.14
60-70	2.31±0.16±0.28	2.34	1.25	0.99±0.15	1.85	1.87
70-85	1.03±0.06±0.13	1.09	0.512	0.95±0.15	2.01	2.12
85-100	0.459±0.009±0.026	0.478	0.239	0.96±0.12	1.92	2.00
100-145	0.134±0.002±0.007	0.190	0.0647	0.70±0.08	2.07	2.95
145-300	0.0096±0.0001±0.0005	0.00654	0.00544	1.46±0.41	1.76	1.20

Table 6.4: Measured triple differential cross section values (pb/GeV) in bins of photon transverse momentum with statistical and systematic uncertainties compared to Sherpa and Pythia predictions ($0.9 < |\eta^{\gamma}| < 1.4442$ and $|\eta^{jet}| < 1.5$).

p_T^{γ} (GeV)	Data (D)	Sherpa (S)	Pythia (P)	D/S	D/P	S/P
40-45	24.1±1.1±2.2	26.9	15.5	0.90±0.10	1.55	1.74
45-50	19.6±1.1±2.1	16.4	10.7	1.19±0.15	1.83	1.54
50-60	10.5±0.5±1.1	9.65	5.60	1.09±0.13	1.88	1.73
60-70	4.51±0.19±0.39	4.76	2.74	0.95±0.10	1.65	1.74
70-85	2.30±0.10±0.21	2.23	1.25	1.03±0.12	1.84	1.79
85-100	1.03±0.01±0.05	1.09	0.522	0.95±0.09	1.97	2.08
100-145	0.310±0.003±0.016	0.319	0.178	0.97±0.09	1.74	1.79
145-300	0.0278±0.0002±0.0013	0.0285	0.0157	0.98±0.15	1.77	1.81

Table 6.5: Measured triple differential cross section values (pb/GeV) in bins of photon transverse momentum with statistical and systematic uncertainties compared to Sherpa and Pythia predictions ($0.9 < |\eta^{\gamma}| < 1.4442$ and $1.5 < |\eta^{jet}| < 2.5$).

p_T^{γ} (GeV)	Data (D)	Sherpa (S)	Pythia (P)	D/S	D/P	S/P
40-45	18.5±1.2±2.8	14.3	9.84	1.29±0.22	1.88	1.45
45-50	9.19±1.00±2.22	10.6	5.76	0.87±0.24	1.60	1.84
50-60	5.43±0.49±1.11	5.48	3.46	0.99±0.23	1.57	1.58
60-70	2.96±0.18±0.36	2.66	1.71	1.11±0.17	1.73	1.56
70-85	1.36±0.09±0.19	1.18	0.769	1.16±0.20	1.77	1.53
85-100	0.550±0.012±0.032	0.502	0.380	1.10±0.16	1.45	1.32
100-145	0.165±0.003±0.009	0.141	0.123	1.17±0.18	1.34	1.15
145-300	0.0120±0.0002±0.0006	0.0134	0.0082	0.90±0.23	1.47	1.64

Table 6.6: Measured triple differential cross section values (pb/GeV) in bins of photon transverse momentum with statistical and systematic uncertainties compared to Sherpa and Pythia predictions ($1.556 < |\eta^{\gamma}| < 2.1$ and $|\eta^{jet}| < 1.5$).

p_T^{γ} (GeV)	Data (D)	Sherpa (S)	Pythia (P)	D/S	D/P	S/P
40-45	20.8±1.5±2.3	20.6	12.3	1.01±0.14	1.69	1.68
45-50	15.3±1.9±2.1	12.5	7.43	1.22±0.23	2.06	1.68
50-60	9.66±0.93±1.31	7.67	4.37	1.26±0.21	2.21	1.75
60-70	3.86±0.21±0.40	3.65	2.18	1.06±0.13	1.77	1.68
70-85	1.94±0.11±0.20	1.73	0.991	1.12±0.15	1.96	1.74
85-100	0.874±0.014±0.050	0.810	0.515	1.08±0.11	1.70	1.57
100-145	0.256±0.003±0.014	0.210	0.144	1.22±0.14	1.78	1.45
145-300	0.0220±0.0002±0.0012	0.0175	0.0128	1.25±0.23	1.72	1.37

Table 6.7: Measured triple differential cross section values (pb/GeV) in bins of photon transverse momentum with statistical and systematic uncertainties compared to Sherpa and Pythia predictions ($1.556 < |\eta^{\gamma}| < 2.1$ and $1.5 < |\eta^{jet}| < 2.5$).

p_T^{γ} (GeV)	Data (D)	Sherpa (S)	Pythia (P)	D/S	D/P	S/P
40-45	20.6±1.3±3.1	17.5	11.9	1.18±0.20	1.73	1.48
45-50	10.4±1.1±3.3	10.4	8.05	1.00±0.34	1.29	1.29
50-60	7.07±0.68±1.44	5.52	4.56	1.28±0.29	1.55	1.21
60-70	3.20±0.20±0.37	2.82	2.140	1.14±0.17	1.50	1.32
70-85	1.51±0.11±0.34	1.27	0.989	1.19±0.30	1.53	1.28
85-100	0.637±0.013±0.048	0.459	0.400	1.39±0.22	1.59	1.15
100-145	0.190±0.003±0.014	0.187	0.104	1.02±0.15	1.83	1.79
145-300	0.0129±0.0002±0.0008	0.0123	0.0094	1.05±0.28	1.38	1.32

Table 6.8: Measured triple differential cross section values (pb/GeV) in bins of photon transverse momentum with statistical and systematic uncertainties compared to Sherpa and Pythia predictions ($2.1 < |\eta^{\gamma}| < 2.5$ and $|\eta^{jet}| < 1.5$).

p_T^{γ} (GeV)	Data (D)	Sherpa (S)	Pythia (P)	D/S	D/P	S/P
40-45	11.5±1.8±3.3	16.7	8.92	0.69±0.22	1.29	1.88
45-50	6.85±1.35±1.75	11.6	5.62	0.59±0.19	1.22	2.07
50-60	3.36±0.98±0.99	6.24	3.57	0.54±0.22	0.94	1.75
60-70	1.85±0.28±0.33	3.20	1.51	0.58±0.14	1.22	2.12
70-85	1.31±0.12±0.23	1.38	0.658	0.95±0.20	1.99	2.10
85-100	0.474±0.020±0.03	0.694	0.380	0.68±0.09	1.25	1.83
100-145	0.136±0.007±0.018	0.184	0.104	0.74±0.14	1.31	1.77
145-300	0.0069±0.0002±0.0004	0.0098	0.0049	0.70±0.20	1.38	2.00

Table 6.9: Measured triple differential cross section values (pb/GeV) in bins of photon transverse momentum with statistical and systematic uncertainties compared to Sherpa and Pythia predictions ($2.1 < |\eta^{\gamma}| < 2.5$ and $1.5 < |\eta^{jet}| < 2.5$).

p_T^{γ} (GeV)	Data (D)	Sherpa (S)	Pythia (P)	D/S	D/P	S/P
40-45	15.2±2.2±2.5	16.7	12.8	0.91±0.24	1.19	1.31
40-45	15.2±2.2±2.5	16.7	12.8	0.91±0.20	1.19	1.31
45-50	9.32±1.90±2.57	11.1	7.57	0.84±0.29	1.23	1.47
50-60	5.98±0.98±1.02	5.91	4.46	1.01±0.25	1.34	1.32
60-70	2.52±0.39±0.52	2.65	2.02	0.95±0.26	1.25	1.31
70-85	1.33±0.21±0.29	1.24	0.892	1.07±0.31	1.49	1.39
85-100	0.484±0.023±0.060	0.584	0.512	0.83±0.16	0.95	1.14
100-145	0.0957±0.0038±0.0123	0.179	0.133	0.53±0.11	0.72	1.35
145-300	0.0049±0.0002±0.0005	0.0079	0.0037	0.61±0.24	1.32	2.16

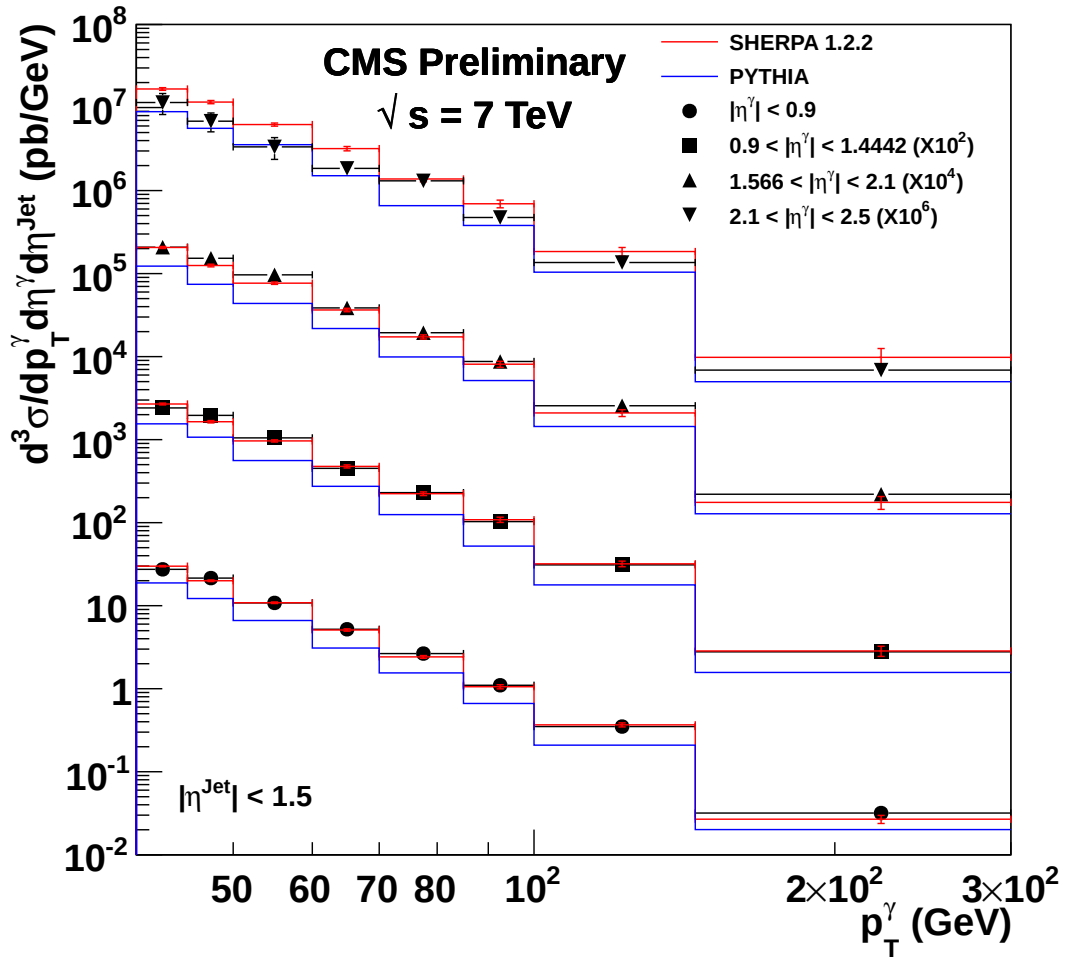

Figure 6.7: Measured photon+jet triple differential cross section as a function of photon p_T (E_T) compared to Pythia and Sherpa predictions. The results are given for four different photon pseudorapidity regions and for jet pseudorapidity region of $|\eta^{jet}| < 1.5$.

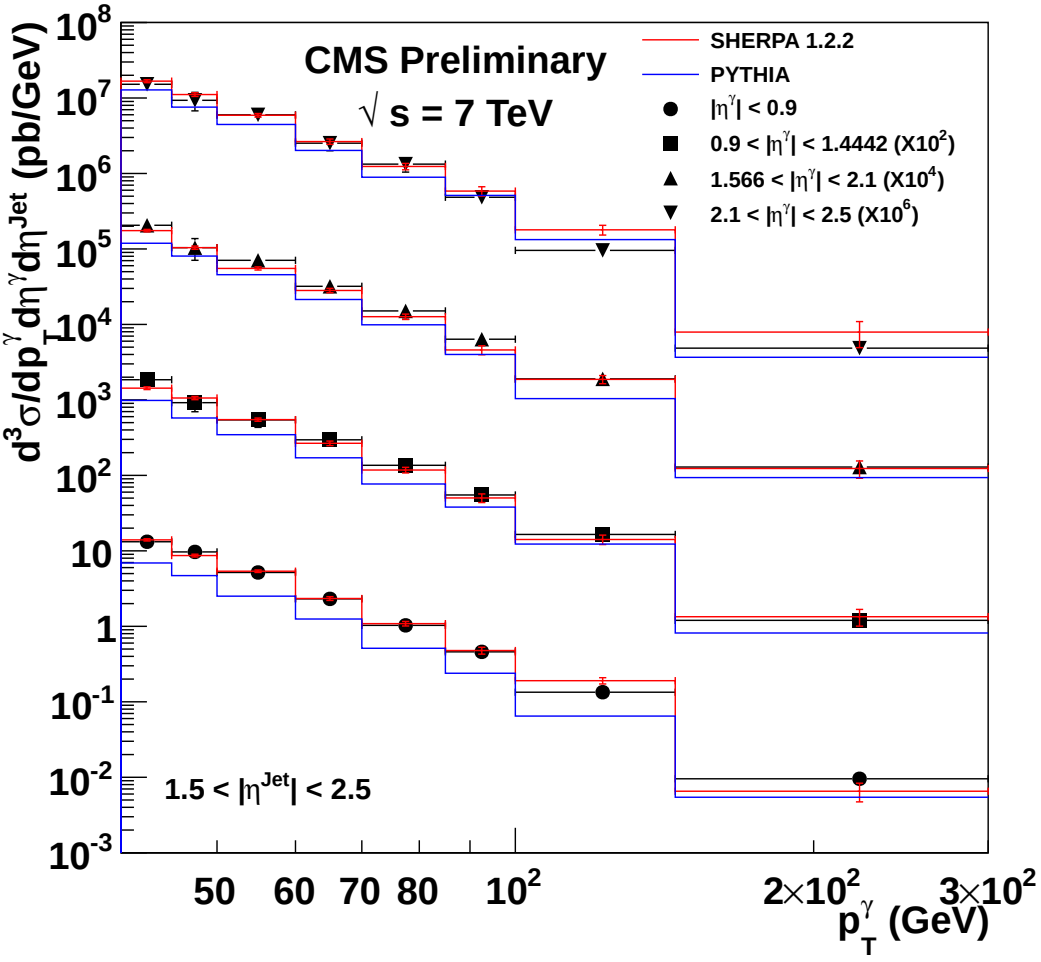

Figure 6.8: Measured photon+jet triple differential cross section as a function of photon p_T (E_T) compared to Pythia and Sherpa predictions. The results are given for four different photon pseudorapidity regions and for jet pseudorapidity region of $1.5 < |\eta^{jet}| < 2.5$.

CHAPTER 7

CONCLUSION

In this thesis, an extensive work of photon efficiency measurements is reported. This work has been carried out for a long term during the author's stay and research at CERN. The results of photon efficiency estimation are well-known in the CMS experiment and well-reported in related CMS meetings for discussions and aggreements of the collaborators. The efficiency of photon triggering is measured from 2.2 /fb collision data recorded by the CMS experiment. The data-driven efficiency measurement is performed with the Tag and Probe technique in which electrons from Z decay are utilized. Reconstruction efficiencies are measured from Monte Carlo samples, while identification efficiencies are measured from both real collision data by Tag and Probe and simulation samples. Monte Carlo efficiency results are cross-confirmed by Pythia and Madgraph simulations.

Photon HLT efficiencies are reported as a function of several variables which are the most relevant for photon spectra analysis such as photon transverse momentum, pseudorapidity, phi angle, number of primary vertices. Photon RECO and ID efficiencies are discussed only in the bins of photon transverse momentum and pseudorapidity. The HLT paths with CaloIdVL are found to be fully efficient, while the paths with CaloIdVL_IsoL are found to be almost 96% efficient in the acceptance. HLT paths without any online selection is observed to be fully efficient. Given that trigger paths both shower shape and isolation selections at the HLT level have inefficiency due to isolation variables, this study yielded an important observation for the CMS collaboration that 2012 single photon HLT paths will no longer be isolated due to less efficiency and higher sensitivity to pile-up.

Single photon in association with jet triple differential cross section results are given in this thesis. This cross section measurement is the first study that has been performed for the CMS

collaboration for photon+jet production in the final state. For this measurements, detailed studies of photon efficiency determination, photon purity extraction, unfolding (deconvolution) for detector effects, and theoretical predictions with Sherpa and Pythia are performed. Data recorded by the CMS experiment in 2011 is used at a center-of-mass energy of 7 TeV which corresponds to 2.2 /bf. Single photon High Level Triggers are used for selection photon object candidates. In photon signal extraction, the so-called isolation template is used in which the sum of photon isolation variables is used for the fitting procedure. For the treatment of Monte Carlo samples, the so-called re-weighting procedure which is common in CMS employed in order to take into account possible pile-up contribution to proton-proton interactions in data.

The results of the prompt photon in association with jet cross section measurement are finally documented in this thesis work. For the final results, different systematic uncertainity constributions are also given as a function of reconstructed photon transverse momentum. Mainly, It has been observed that major systematic errors come from efficiency measurement of photon identification as respect to the difference between real data and simulation samples, purity determination while variable fitting for signal and background shapes, and unfolding the final spectrum due to photon energy scale difference in data and Monte Carlo. The cross section results are found to be consistent between data and theoretical prediction tools Sherpa and Pythia. Sherpa prediction is found to be in better agreement with data results.

The technique used in this thesis for the effciency measurement is first used by the author as contribution to the inclusive prompt photon cross section measurement with the CMS detector in 2011 with the following analysis note and paper:

- K. Ocalan et al., Photon Plus Jet: Single Photon HLT path efficiency measurements for the 2010 dataset, CMS AN-2010/458, 2010.

- CMS Collaboration, K. Ocalan et al., Measurement of the Differential Cross Section for Isolated Prompt Photon Production in pp Collisions at 7 TeV, Phys. Rev. D 84 052011 (2011), arXiv:1108.2044 [hep-ex], 2011.

The cross section measurement results given in this thesis are documented in the form of CMS analysis note as given below:

- K. Ocalan et al., Measurement of the triple differential gamma+jet cross section using 2011 data, CMS AN-2011/331, 2011.

And the results presented in this thesis are going to be published in a respective journal (JHEP) on behalf of the CMS collaboration [83]:

- CMS Collaboration, K. Ocalan et al., Measurement of triple-differential cross section of Gamma+jet production, CMS Physics Analysis Summary (PAS) QCD-11-005 (2012).

As a future outlook, the author plans to continue working on the photon physics analyses in the CMS experiment. Specifically, the author aims to involve in photon, photon+jets, and any foton final state measurements in the CMS experiment with 8 TeV data in his analysis group.

REFERENCES

[1] Particle Data Group (PDG),

http://pdg.lbl.gov/pdg.html, last visited on November 13, 2011.

[2] R. P. Feynman, Mathematical formulation of the quantum theory of electromagnetic interaction, *Phys. Rev.* **80**, 440-457 (1950).

[3] S. Weinberg, A Model of Leptons, *Phys. Rev. Lett.* **19**, 1264-1266 (1967).

[4] A. Salam, Gauge unification of fundamental forces, *Rev. Mod. Phys.* **52**, 525-538 (1980).

[5] S. L. Glashow, Towards a Unified Theory: Threads in a Tapestry, *Rev. Mod. Phys.* **52**, 539-543 (1980).

[6] P. W. Higgs, Broken symmetries, massless particles and gauge fields, *Phys. Lett.* **12**, 132 (1964).

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

[8] D. J. Griffiths, *Introduction to Elementary Particles* (John Wiley and Sons, New York, 1987).

[9] W. N. Cottingham and D. A. Greenwood, *An Introduction to the Standard Model of Particle Physics* (Cambridge University Press, New York, 2007).

[10] B. R. Martin and G. Shaw, *Particle Physics* (John Wiley and Sons, New York, 1997).

[11] K. Gottfried and V. F. Weisskopf, *Concepts of Particle Physics, Volume II* (Oxford University Press, New York, 1986).

[12] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics* (Wiley Text Books, New York, 1984).

[13] H. Fritzsch and M. Gell-Mann, Proceedings of the XVIth International Conference on High energy Physics, Batavia, IL, Vol. 2, p. 135 (1972).

[14] H. Fritzsch, M. Gell-Mann, H. Leutwyler, Advantages of the Color Octet Gluon Picture, *Phys. Lett.* **B** 47, 365 (1973).

[15] R. K. Ellis, W. J. Stirling, B. R. Webber, *QCD and Collider Physics* (Cambridge University Press, New York, 1996).

[16] Y.M. Yao et al., Review of Particle Physics, *J. Phys.* **G33**, 1 (2006).

[17] S. Bethe, Determination of the QCD coupling α_s, *J. Phys. G: Nucl. Part. Phys.* **R27**, 26 (2000).

[18] P. Aurenche et al., Prompt photon production at colliders, *Physical Review D* **42**, 1440-1449 (1990).

[19] Kumar et al., Study of direct photon production at the CERN LHC, *Physical Review D* **67**, 014016 (2003).

[20] P. Gupta et al., Study of Direct Photon plus Jet production in CMS Experiment at $\sqrt{s}$ = 14 TeV, *Eur. Phys. J.* **C53**, 49-58 (2008).

[21] M. Cacciari, G. P. Salam, G. Soyez, The $anti-k_t$ jet clustering algorithm, *JHEP* **04**, 063 (2008).

[22] M. Fontannaz and G. Heinrich, Isolated photon plus jet photoproduction as a tool to constrain the gluon distribution in the proton and the photon, *Eur. Phys. J.* **C34**, 191-199 (2004).

[23] L. Randall and R. Sundrum, A large mass hierarchy from a small extra dimension, *Phys. Rev. Lett.* **83**, 3370-3373 (1999).

[24] N. Arkani-Hamed, S. Dimopoulos, G. Dvali, The hierarchy problem and new dimensions at a millimeter, *Phys. Lett. B* **429**, 263-267 (1991).

[25] M. Dine, W. Fischler, M. Srednicki, Supersymmetric Technicolor, *Nucl. Phys. B* **189**, 575-593 (1981).

[26] S. Catani et al., Cross section of isolated prompt photons in hadron-hadron collisions, *JHEP* **05**, 028 (2001).

[27] T. Sjostrand et al., PYTHIA 6.4 Physics and Manual, *JHEP* **05**, 026 (2006).

[28] B. Anderson et al., Parton Fragmentation and String Dynamics, *Phys. Rept.* **97**, 31-145 (1983).

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

[30] S. Hoeche et al., Hard photon production and matrix-element parton-shower merging, *Phys. Rev.* **D81**, 034026 (2010).

[31] G. Corcella et al., Herwig 6.5, *JHEP* **01**, 010 (2001).

[32] L. Evans and P. Bryant, LHC Machine, *JINST* **3**, S08001 (2008).

[33] CERN, The accelerator complex,

http://public.web.cern.ch/public/en/research/AccelComplex-en.html, last visited on December 02, 2011.

[34] CERN, The LHC dipole magnet,

http://cdsweb.cern.ch/record/40524, last visited on December 20, 2011.

[35] O. Bruening et al., LHC Design Report, CERN-2004-003, CERN (2004).

[36] CMS Collaboration, S. Chatrchyan et al., The CMS experiment at the CERN LHC, *JINST* **3**, S08004 (2008).

[37] CMS Collaboration, G.L. Bayatian et al., CMS Physics Technical Design Report Volume I : Detector Performance and Software, CERN-LHCC-2006-001 (2006).

[38] CMS Collaboration, The Tracker Project Technical Design Report, CERN-LHCC 98-006 (1998).

[39] CMS Collaboration, Addendum to the CMS tracker TDR, CERN-LHCC 2000-016 (2000).

[40] S. Cittolin, CMS Detector posters,

https://cms-docdb.cern.ch/cgi-bin/ PublicDocDB/ShowDocument?docid=4227, last visited on November 13, 2011.

[41] CMS Collaboration, The Electromagnetic Calorimeter Technical Design Report, CERN-LHCC 97-033 (1997).

[42] CMS Collaboration, Addendum to the CMS ECAL technical design report: Changes to the CMS ECAL electronics, CERN-LHCC 2002-027 (2002).

[43] CMS Collaboration, The Hadron Calorimeter Technical Design Report, CERN-LHCC 97-031 (1997).

[44] CMS HCAL Collaboration, S. Chatrchyan et al., Study of Various Photomultiplier Tubes with Muon Beams And Cerenkov Light Produced in Electron Showers, *JINST* **5**, P06002 (2010).

[45] CMS Collaboration, The CMS muon project: Technical Design Report, CERN-LHCC 97-032 (1997).

[46] CMS Collaboration, CMS TriDAS Project, Technical Design Report Volume 1: The Trigger Systems, CERN-LHCC 2000-038 (2000).

[47] W. Adam et al., The CMS high level trigger, *Eur. Phys. J.* **C46**, 605-667 (2006).

[48] CMS Collaboration, G.L. Bayatian et al., The TriDAS project, Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger, CERN-LHCC 2002-026 (2002).

[49] CMS Collaboration, G.L. Bayatian et al., CMS computing: Technical Design Report, CERN-LHCC-2005-023 (2005).

[50] I. Bird et al., LHC Computing Grid Technical Design Report, CERN-LHCC-2005-024 (2005).

[51] CMS Collaboration, Photon reconstruction and identification at $sqrts = 7$ TeV, *CMS Physics Analysis Summary* **EGM-10-005** (2010).

[52] CMS Collaboration, Isolated Photon Reconstruction and Identification at $sqrts = 7$ TeV, *CMS Physics Analysis Summary* **EGM-10-006** (2010).

[53] K. Ocalan et al., Measurement of the triple differential $\gamma + jet$ cross section using 2011 data, *CMS Analysis Note* **CMS-AN-11-31**, (2011).

[54] E. Anassontzis et al., High p(t) Direct Photon Production in pp Collisions, *Z. Phys. C* **13**, 277-289 (1982).

[55] CMOR Collaboration, A. L. S. Angelis et al., Direct Photon Production at the CERN ISR, *Nucl. Phys.* **B327**, 541 (1989).

[56] UA2 Collaboration, J. A. Appel et al., Direct Photon Production at the CERN anti-pp Collider, *Phys. Lett.* **B176**, 239 (1986).

[57] UA1 Collaboration, C. Albajar et al., Direct Photon Production at the CERN proton-antiproton Collider, *Phys. Lett.* **B209**, 385 (1988).

[58] UA6 Collaboration, M. Werlen et al., A New determination of α_s using direct photon production cross-sections in pp and $\overline{p}p$ collisions at S(1/2) = 24.3 GeV, *Phys. Lett.* **B452**, 201-206 (1999).

[59] PHENIX Collaboration, S. S. Adler et al., Midrapidity direct-photon production in p+p collisions at sqrt(s) = 200 GeV, *Phys. Rev.* **D71**, 071102 (2005).

[60] DØ Collaboration, V. M. Abazov et al., Measurement of the isolated photon cross section in $p\overline{p}$ collisions at $\sqrt{s}$ = 1.96 TeV, *Phys. Lett. B* **639**, 151 (2006).

[61] CDF Collaboration, T. Aaltonen et al., Measurement of the Inclusive Isolated Prompt Photon Cross Section in $p\overline{p}$ Collisions at $\sqrt{s}$ = 1.96 TeV using the CDF Detector, *Phys. Rev. D* **80**, 111106 (2009).

[62] ZEUS Collaboration, S. Chekanov et al., Measurement of prompt photons with associated jets in photoproduction at HERA, *Eur. Phys. J.* **C49**, 511-522 (2007).

[63] A. V. Lipatov and N. P. Zotov, Prompt photon photoproduction at HERA in the k(T)-factorization approach, *Phys. Rev. D* **72**, 054002 (2005).

[64] ZEUS Collaboration, S. Chekanov et al., Measurement of isolated photon production in deep inelastic ep scattering, *Phys. Lett. B* **687**, 16-25 (2010).

[65] H1 Collaboration, F. D. Aaron et al., Prompt Photons in Photoproduction at HERA, *Eur. Phys. J.* **C66**, 17-33 (2010).

[66] ATLAS Collaboration, G. Aad et al., Measurement of the inclusive isolated prompt photon cross-section in pp collisions at $\sqrt{s}$ = 7 TeV using 35 pb^{-1} of ATLAS data, *Phys. Lett. B* **706**, 150-167 (2011).

[67] CMS Collaboration, V. Khachatryan et al., Measurement of the Isolated Prompt Photon Production Cross Section in pp Collisions at $\sqrt{s}$ = 7 TeV, *Phys. Rev. Lett.* **106**, 082001 (2011).

[68] CMS Collaboration, S. Chatrchyan et al., Measurement of the differential cross section for isolated prompt photon production in pp collisions at $\sqrt{s}$ = 7 TeV, *Phys. Rev. D* **84**, 052011 (2011).

[69] DØ Collaboration, V. M. Abazov et al., Measurement of the differential cross section for the production of an isolated photon with associated jet in $p\overline{p}$ collisions at $\sqrt{s}$ = 1.96 TeV, *Phys. Lett. B* **666**, 435-445 (2008).

[70] N. Adam et al., Generic Tag and Probe Tool for Measuring Efficiency at CMS with early Data, *CMS Analysis Note* **AN-09-111** (2009).

[71] CMS Collaboration, Measuring Electron Efficiencies at CMS with Early Data, *CMS Physics Analysis Summary* **EGM-07-001** (2008).

[72] Electron Tag and Probe,

https://twiki.cern.ch/twiki/bin/viewauth/CMS/ElectronTagAndProbe, last visited on September 11, 2011.

[73] K. Ocalan et al., Photon Plus Jet: Single Photon HLT path efficiency measurements for the 2010 dataset, *CMS Analysis Note* **AN-10-458** (2010).

[74] J. Berryhill et al., Electron Efficiency Measurements with 2.88 pb^{-1} of pp Collision Data at $\sqrt{s}$ = 7 TeV, *CMS Analysis Summary* **AN-10-349** (2010).

[75] CMS Collaboration, Calorimeter Jet Quality Criteria for the First CMS Collision Data, *CMS Physics Analysis Summary* **JME-09-008** (2009).

[76] CERN, CMS Jet Energy Corrections,

https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookJetEnergyCorrections, last visited on September 29, 2011.

[77] J. Pumplin et al., New Generation of Parton Distributions with Uncertainties from Global QCD Analysis, *JHEP* **0207**, 012 (2002).

[78] T. Adye, RooUnfold: ROOT unfolding framework,

http://hepunx.rl.ac.uk/ adye/software/unfold/RooUnfold.html, last visited on December 23, 2011.

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

[80] T. Gleisberg et al., Event generation with SHERPA 1.1, *JHEP* **0902**, 007 (2009).

[81] S. Hoche et al., QCD matrix elements and truncated showers, *JHEP* **0905**, 053 (2009).

[82] S. Hoche et al., Hard photon production and matrix-element parton-shower merging, *Phys. Rev. D* **81**, 034026 (2010).

[83] CMS Collaboration, K. Ocalan et al., Measurement of triple-differential cross section of Gamma+jet production, *CMS Physics Analysis Summary* **QCD-11-005** (2012).

APPENDIX A

HLT EFFICIENCY DEPENDENCE ON JET MULTIPLICITY

Data-driven HLT efficiency measurement is extended to cover dependancy on jet multiplicity. For this purpose, particle flow jets namely *ak5PFJets* are used as being jet collection [75]. Full jet energy corrections namely *L1FastL2L3* are applied on jets as recommended by the CMS experiment for 2011A data [76]. Jets are required to have $p_T > 25$ GeV within $|\eta| < 2.4$, and pass *Loose PF Jet ID* as below:

- Neutral Hadron Energy Fraction < 0.99,
- Neutral EM Energy Fraction < 0.99,
- Number of Constituents > 1,
- Charged Hadron Energy Fraction > 0,
- Charged EM Energy Fraction < 0.99,
- Charged Multiplicity > 0.

HLT efficiencies of `HLT_Photon75_CaloIdVL` and `HLT_Photon75_CaloIdVL_IsoL` as a function of jet multiplicity are compared for photon $p_T > 85$ GeV in Figure A.1. Non-isolated HLT paths are observed to be almost fully efficient for number of jets ≥ 0, while isolated HLT paths are fully efficient for number of jets ≥ 3. Efficiencies of `HLT_Photon75_CaloIdVL` and `HLT_Photon75_CaloIdVL_IsoL` are reported in Table A.1 and Table A.2, respectively from fitting method of Tag and Probe package in both Barrel ($|\eta| < 1.4442$) and Endcap ($1.566 < |\eta| < 2.5$) regions seperately with photon $p_T > 85$ GeV requirement. In general, isolated and non-isolated trigger paths are fully efficient except some fluctuations caused by less statistics in certain jet multiplicity bins.

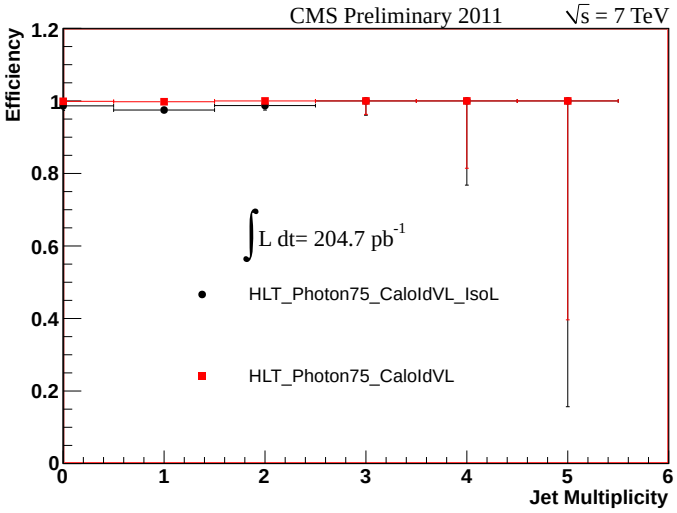

Figure A.1: Dependancy of single photon HLT efficiencies on jet multiplicity measured from 2011A data by using T&P technique.

Table A.1: Efficiency of HLT_Photon75_CaloIdVL as a function of jet multiplicity (upto 4 jets) in Barrel ($|\eta| < 1.4442$) and Endcap ($1.566 < |\eta| < 2.5$) regions.

Jet Multiplicity	Barrel	Endcap
0	100 ± 0.4	99.8 ± 6.3
1	99.7 ± 0.3	100 ± 1.2
2	100 ± 0.7	100 ± 4.2
3	100 ± 2.1	100 ± 24.2
4	100 ± 13.3	100 ± 36.0

Table A.2: Efficiency of HLT_Photon75_CaloIdVL_IsoL as a function of jet multiplicity (upto 4 jets) in Barrel ($|\eta| < 1.4442$) and Endcap ($1.566 < |\eta| < 2.5$) regions.

Jet Multiplicity	Barrel	Endcap
0	99.5 ± 0.5	99.0 ± 4.8
1	98.4 ± 0.7	97.5 ± 3.4
2	98.4 ± 1.3	100 ± 4.5
3	100 ± 2.3	100 ± 24.1
4	100 ± 15.4	100 ± 42.1

APPENDIX B

TRIGGER EFFICIENCIES FROM THE RATIO TECHNIQUE

The efficiency of prescaled photon trigger paths are estimated with a ratio technique (also known as bootstrapping technique), where a low transverse energy E_T threshold trigger with high efficiency is used as a reference for the nominal trigger. In this picture, the efficiency of a higher threshold trigger with a nominal threshold starting from 50 GeV is computed relative to a prescaled sample of lower threshold trigger starting with 30 GeV. The extracted efficencies by taking trigger ratios are summarized below.

- $\frac{HLT\_Photon50\_CaloIdVL}{HLT\_Photon30\_CaloIdVL}$ ratio is shown in Figure B.1. HLT_Photon50_CaloIdVL is fully efficient for 10 GeV above its threshold.

- $\frac{HLT\_Photon75\_CaloIdVL}{HLT\_Photon50\_CaloIdVL}$ ratio is shown in Figure B.2. HLT_Photon75_CaloIdVL is 95% efficient for 10 GeV above its threshold. This path is fully efficient by T&P technique, the difference comes from different prescales applied for the numerator trigger and the denominator trigger.

- $\frac{HLT\_Photon90\_CaloIdVL}{HLT\_Photon75\_CaloIdVL}$ ratio is shown in Figure B.3. HLT_Photon90_CaloIdVL is 97% efficient for 10 GeV above its threshold. The efficiency is close to 100%, since the prescale factors of these two triggers are different.

- $\frac{HLT\_Photon135}{HLT\_Photon90\_CaloIdVL}$ ratio is shown in Figure B.4. The nominal trigger path has a turn-on arund 10 GeV above its threshold. Since HLT_Photon135 path is not prescaled while the denominator one is, efficiency values can not be estimated in a reliable way. T&P results are used for the analysis of this thesis.

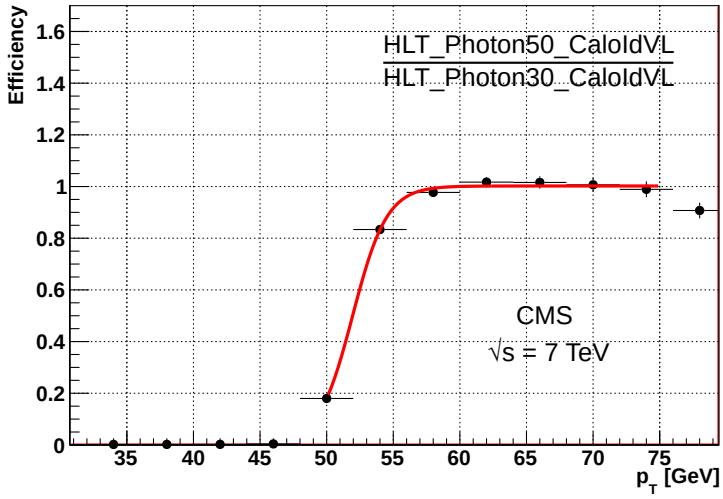

Figure B.1: Eficency of the nominal trigger path HLT_Photon50_CaloIdVL_IsoL relative to HLT_Photon30_CaloIdVL_IsoL. The data points are fitted by an exponential error function.

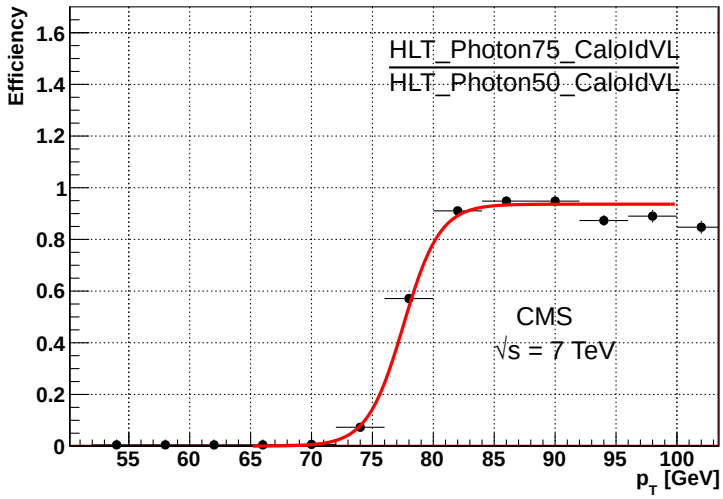

Figure B.2: Eficency of the nominal trigger path HLT_Photon75_CaloIdVL_IsoL relative to HLT_Photon50_CaloIdVL_IsoL. The data points are fitted by an exponential error function.

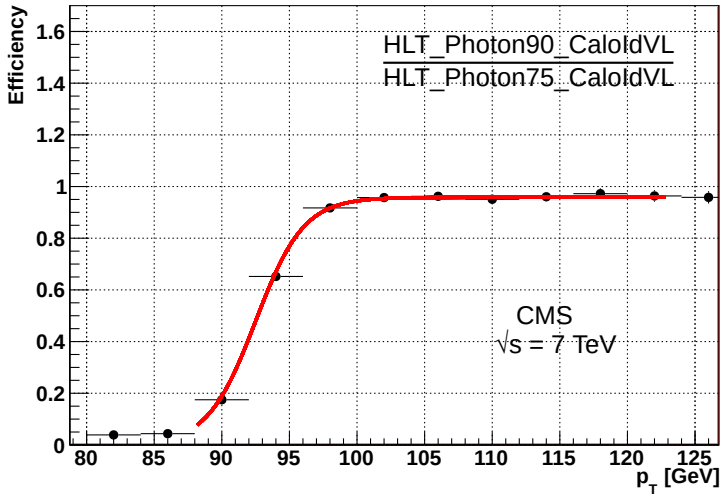

Figure B.3: Eficency of the nominal trigger path HLT_Photon90_CaloIdVL_IsoL relative to HLT_Photon75_CaloIdVL_IsoL. The data points are fitted by an exponential error function.

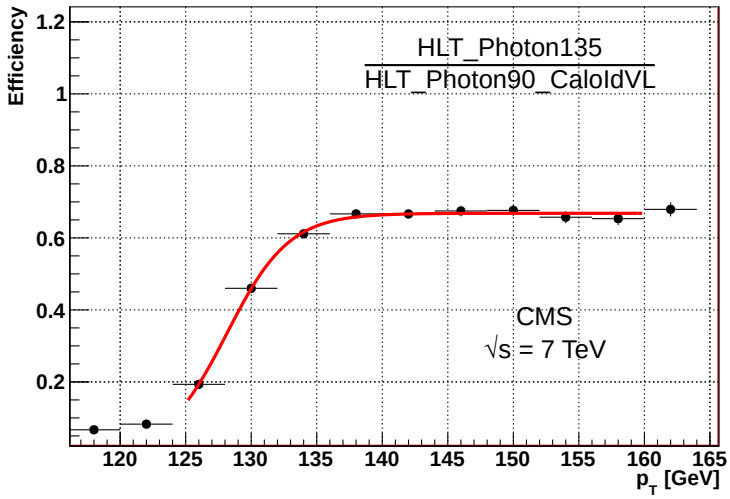

Figure B.4: Eficency of HLT_Photon135 relative to HLT_Photon90_CaloIdVL_IsoL. The data points are fitted by an exponential error function.

APPENDIX C

ID EFFICIENCY DEPENDENCE ON RECONSTRUCTED VERTICES

Shower shape and isolation templates photon identification (ID) efficiencies are investigated in terms of event number of primary vertices (N_{PV}) in 2011A data recorded by the CMS experiment. Shower shape template selection efficiency is observed not to be flat indicating that photon isolation variables (ECAL Iso, HCAL Iso, Track Iso) and thus the shower shape template efficiency is affected by the pile-up (Figure C.1 (left)). Isolation template photon ID selection efficiency is not much dependant on the recunstructed vertices indicating that isolation template ID efficiency is not strongly affected by the pile-up (Figure C.1 (right)). Isolation template is used in cross section measurement of this thesis, so no pile-up reweighting is employed for Monte Carlo samples.

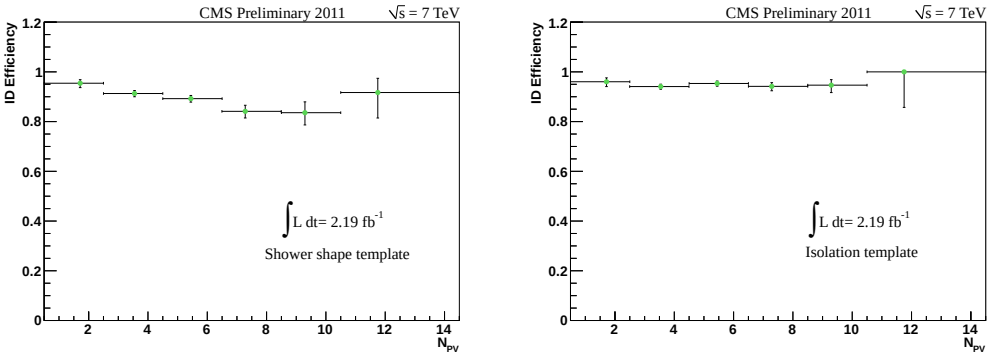

Figure C.1: Dependancy of Shower shape template (left) and Isolation template (right) identification (ID) efficiencies on event number of primary vertices (N_{PV}) variable measured from 2011A collsion data.

CURRICULUM VITAE

Kadir Öcalan

Physics Department, Middle East Technical University, 06800 Ankara, Turkey

CERN - PH/UCM, Mailbox J08410, 51 R-013, 23 Geneva 1211, Switzerland

Education

- Ph.D. in Physics, March 2012, Middle East Technical University, Ankara, Turkey
- M.Sc. in Physics, January 2009, Middle East Technical University, Ankara, Turkey
- B.Sc. in Physics Education, June 2006, Middle East Technical University, Ankara, Turkey

Research Interests

- Experimental High Energy Particle Physics, Collider Photon Physics, Jet Physics, Supersymmetry Searches
- Hadronic and Electromagnetic calorimeters, Trigger and Data acquisition systems
- Compact Muon Solenoid (CMS) Experiment at the Large Hadron Collider (LHC) at CERN

Research Experience

- Gamma+Jet Triple Cross Section Measurement with the CMS Experiment (overall direct contribution), 2011-2012.
- Inclusive Photon Production Cross Section Measurement with the CMS Experiment (direct contribution by photon trigger studies), 2011.

- High Level Trigger (HLT) contact person of CMS QCD Photons Physics analysis group (trigger online-offline monitoring, efficiency, and rate measurements), 2010-2011.

- CMS Hadronic Calorimeter (HCAL) commisioning and Hadronic Forward (HF) sector Test Beam studies, 2008-2011.

- Supersymmetry Searches in Hadronic Channels with the CMS Experiment (direct contribution by Z(nunu) background estimation studies), 2008-2010.

Project Work

1. Turkish Atomic Energy Authority (TAEK), Project No:05K120010, Physics Simulation and Data Analysis for beyond Standard Model and Detector upgrade Studies in CMS, Nov, 2010-Nov, 2013, Researcher.

2. Turkish Atomic Energy Authority (TAEK), CERN-A5.H.2.P1.01/6, Test Beam work at CMS experiment and search for supersymmetric particles, May, 2007-May, 2010, Researcher.

Publications

1. CMS Collaboration, K. Ocalan et al., Measurement of triple-differential cross section of gamma+jet production in pp Collisions at 7 TeV, CMS QCD-11-005, In preperation for JHEP, 2012

2. CMS Collaboration, K. Ocalan et al., Measurement of the Differential Cross Section for Isolated Prompt Photon Production in pp Collisions at 7 TeV, Phys. Rev. D 84 052011 (2011), arXiv:1108.2044 [hep-ex], 2011

3. CMS HCAL Collaboration, K. Ocalan et al., Study of Various Photomultiplier Tubes with Muon Beams And Cerenkov Light Produced in Electron Showers, JINST 5 (2010) P06002, 2010.

4. CMS Collaboration, K. Ocalan et al., Measurement of the charge ratio of atmospheric muons with the CMS detector, Phys.Lett. B692 (2010) 83-104, arXiv:1005.5332 [hep-ex], 2010.

5. CMS Collaboration, K. Ocalan et al., Performance of the CMS Hadron Calorimeter with Cosmic Ray Muons and LHC Beam Data, JINST 5 (2010) T03012, arXiv:0911.4991 [physics.ins-det], 2010.

6. CMS Collaboration, K. Ocalan et al., Identification and Filtering of Uncharacteristic Noise in the CMS Hadron Calorimeter, JINST 5 (2010) T03014, arXiv:0911.4881 [physics.ins-det], 2010.

7. CMS Collaboration, K. Ocalan et al., Performance of CMS Hadron Calorimeter Timing and Synchronization using Test Beam, Cosmic Ray, and LHC Beam Data, JINST 5 (2010) T03013, arXiv:0911.4877 [physics.ins-det], 2010.

8. CMS Collaboration, K. Ocalan et al., Commissioning of the CMS Experiment and the Cosmic Run at Four Tesla, JINST 5 (2010) T03001, arXiv:0911.4845 [physics.ins-det], 2010.

9. CMS Collaboration, K. Ocalan et al., CMS Data Processing Workflows during an Extended Cosmic Ray Run, JINST 5 (2010) T03006, arXiv:0911.4842 [physics.ins-det], 2010.

Internal Notes

1. K. Ocalan et al., Beam halo removal with shower shape template in 2010 data, CMS AN-2011/336, 2011.

2. K. Ocalan et al., Measurement of the triple differential gamma+jet cross section using 2011 data, CMS AN-2011/331, 2011.

3. K. Ocalan et al., Photon Plus Jet: Single Photon HLT path efficiency measurements for the 2011 dataset, CMS AN-2011/314, 2011.

4. K. Ocalan et al., Photon Plus Jet: Single Photon HLT path efficiency measurements for the 2010 dataset, CMS AN-2010/458, 2010.

5. K. Ocalan et al., Invisible Z Background Estimation from Photon+Jets for alphaT Based SUSY Searches, CMS AN-2010/060, 2010.