Carnegie Mellon University
MELLON COLLEGE OF SCIENCE

THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN THE FIELD OF PHYSICS

TITLE: "Search for supersymmetry using events with a photon plus lepton and missing transverse momentum in proton-proton collisions at sqrt(s) = 13 TeV with the CMS detector."

PRESENTED BY: Menglei Sun

ACCEPTED BY THE DEPARTMENT OF PHYSICS

Manfred Paulini	9/21/18
MANFRED PAULINI, CHAIR PROFESSOR	DATE

Scott Dodelson	9/21/18
SCOTT DODELSON, DEPT HEAD	DATE

APPROVED BY THE COLLEGE COUNCIL

Rebecca Doerge	9/27/18
REBECCA DOERGE, DEAN	DATE

Search for supersymmetry using events with a photon plus lepton and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV with the CMS detector

by

Menglei Sun

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
at
Carnegie Mellon University
Department of Physics
Pittsburgh, Pennsylvania

Advised by Professor Manfred Paulini

September 20, 2018

Abstract

Results of a search for new physics in events with a photon, an electron or muon, and large missing transverse momentum (p_T^{miss}) is presented. The study is based on a sample of proton-proton collisions at $\sqrt{s} = 13$ TeV corresponding to an integrated luminosity of 35.9 fb^{-1} collected with the CMS detector in 2016. Many models of new physics predict events with significant p_T^{miss} in addition to electroweak gauge bosons. Models of supersymmetry (SUSY) with gauge-mediated supersymmetry breaking naturally yield events with photons in the final state. Searches for events with both a photon and a lepton enhanced the sensitivity to electroweak production of supersymmetric particles. No significant excess above the standard model background is observed in the signal region. We interpret the results of our search in the context of SUSY with gauge-mediated supersymmetry breaking as well as simplified SUSY models.

Acknowledgements

First of all I would like to express my sincere gratitude to my advisor, Manfred Paulini, for guiding and inspiring me during my graduate studies. Manfred has always been very supportive of students trying new ideas. He also encourages us to attend conferences and collaborate with other researchers. Thanks to Manfred, I have got the opportunity to participate in many interesting research topics, including the study of the Phase II upgrade of the CMS detector and the search for supersymmetry.

I would like to thank the conveners of the SUSY photon subgroup: Andrew Askew, Si Xie, Rishi Patel, and Marc Weinberg. Their comments and suggestions to this thesis research have been very helpful in pushing the analysis forward. I would like to thank particularly Marc, who is also a co-author of this analysis. I benefit a lot from his wide knowledge on the supersymmetry search and CMS detector. It's a pleasure to work with him.

Working in a large group like the ECAL subdetector group is both enjoyable and challenging. I would like to express my appreciation to the leaders of the ECAL group, Dave Barney and David Petyt, for their helping and supporting on ECAL related work. My gratitude also goes to other members of the Carnegie Mellon group, Tanmay Mudholkar and Michael Andrews, whose dedicated work on the ECAL and DQM systems are crucial for the successful data taking of the CMS experiment.

I would also like to thank Yutaro Iiyama, who has helped me a lot on academic work. When I first came to CERN in 2014, I was just a beginner of the CMS, feeling nervous about presenting my work in meetings and taking detector-on-call shifts. Yutaro taught me a lot of useful things about the ECAL detector and the DQM system. With his help, I gradually became the on-call expert of the ECAL and DQM system. He is also very responsive when I have questions about the physics analysis.

Doing research in CERN was an exciting experience, and the life in Geneva became even more wonderful when I met a lot of fabulous friends there. To my dear friends Mingjian Lu and Peng Jiang, who were also my housemates, I would like to thank them for the discussions we ever had and all the joy they brought to my life. It was a tragedy that we lost them in an accident. I will always miss them. Luckily, I still have the support and accompany from other friends. Here I would like to thank my friends, Hua Wei, Xinmei Niu, Haonan Lu, Qi Zend, and Liangjing Zhu, for their support and help. I also benefit a lot from the discussion of physics problems with these friends.

My deepest gratitude to all my family, particularly to my beloved husband, Jingkun Gao, for his love and encouragement; and to my parents, for their support and understanding.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

What is the world made of? This is a question that has been asked by humans for thousands of years. The idea that all matter is composed of elementary particles dates from at least the 5th century BC. The development of the atomic theory at the start of the 19th century showed that all materials are made of atoms, which were thought to be the fundamental particles. However, the discovery of the electron in 1896, the nucleus in 1911 and the neutron in 1931 revealed that atoms are made of smaller sub-atomic particles. To better understand the structure of matter, physicists need to look deeper and study the most elementary particles. Particle physics is a branch of the physics that aims to understand the fundamental nature of matter and energy, the interactions between them, and to apply that knowledge to better understand the origin and structure of the universe. Throughout the 1950s and 1960s, hundreds of sub-atomic particles were discovered in particle physics experiments. In examining and organizing these particles, a standard model of particle physics has emerged.

The standard model is our current theory that best describes the subatomic world. According to the standard model, matter is built up from a set of spin-1/2 particles, called fermions. Forces between matter result from the exchange of spin-1 particles, called bosons. In addition, the standard model predicts a spin-0 particle, called the

Higgs boson, which explains the origin of particle mass. Since its formulation in the 1970s, the standard model has stood up to many experimental tests. The Higgs boson, which is the final piece of the standard model, has been discovered in 2012 by the ATLAS and CMS experiments. This discovery marks the triumph of the standard model.

Despite its great success, the standard model still leaves open many unanswered questions. What is the cause of the significant asymmetry between matter and anti-matter? Why is the Higgs mass so small while it is sensitive to the ultraviolet physics through radiative corrections? Moreover, a variety of astronomical and cosmological experiments provide strong evidence of the existence of dark matter and dark energy. It turns out that the "normal" matter only accounts for about 5% of the universe. So what is the nature of the dark matter and dark energy that fill up the rest of the universe? To solve these problems, many new physics models are explored, each extending the standard model using a different mechanism. The supersymmetric extension of the standard model is an appealing theory that offers solutions to these problems. It proposes a symmetry between fermions and bosons and doubles the number of particles in the standard model. The new set of particles predicted by the Symmetry Model helps stabilize the mass of the Higgs boson, and the lightest supersymmetric particle provides a candidate for dark matter.

The Large Hadron Collider (LHC) at CERN was built to search for the Higgs boson and new physics beyond the standard model. In 2015, the LHC started its second run with an increased center-of-mass energy of 13 TeV, providing proton-proton collisions every 25 ns. This powerful machine provides us the opportunities to probe physics at the TeV scale. It hosts seven detectors, each designed for certain kinds of researches. The Compact Muon Solenoid (CMS) is a general-purpose detector sitting in one of the LHC collision points. The essential design of the CMS is the use of a 3.8 T superconducting magnet that can provide a large bending power and enable a precise momentum measurement. This thesis presents an analysis using the proton-proton collision data collected by the CMS detector in 2016.

The analysis presented in this thesis searches for supersymmetry using events with a photon plus lepton and large transverse missing momentum. Events with photons are typical signatures of supersymmetry scenarios with gauge-mediated SUSY breaking. Final states with an additional lepton suppress the standard model backgrounds and enhance the sensitivity to the electroweak production of supersymmetric particles. The remaining backgrounds from standard model processes are estimated using both simulation and data-driven methods.

This thesis is organized as follows: Chapter 2 gives a brief introduction to the standard model and its supersymmetric extension. Chapter 3 describes the LHC and the CMS detector, while the reconstruction of data is given in Chapter 4. Chapter 5 describes the data used in this analysis, and Chapter 6 presents the selection of the data. Background estimations are discussed in detailed in Chapter 7. The results of the search and its interpretations are presented in Chapter 8. Finally, the thesis is summarized in Chapter 9.

Chapter 2

Theory and Motivation

2.1 The Standard Model

The standard model (SM) of particle physics is a well-tested and successful theory that describes the known elementary particles and their interactions. It is built within a framework known as Quantum Field Theory (QFT), which provides the mathematical tools to describe the subatomic particles using theories that are consistent with both quantum mechanics and special relativity. In QFT, particles are represented by excitations of quantum fields. The dynamics and interactions of the particles are governed by the Lagrangian density $\mathcal{L}$.

Throughout this thesis we use the natural units, defined by:

$$c = \hbar = 1, \tag{2.1}$$

where $\hbar$ is the reduced Planck constant and c is the speed of light. Using this convention, mass, energy, and momentum can be expressed with the same unit: eV. For heavy particles, MeV ($= 10^6$ eV) and GeV ($= 10^9$ eV) are used.

The SM consists of a set of matter particles and force carriers, as well as the Higgs boson which gives mass to all fundamental particles. Table 2.1 summarises the SM particles and their properties. These particles are classified into two groups: fermions with half-integer spin, and bosons with integer spin. Fermions are further grouped into three generations, with masses increasing from one generation to the next.

Table 2.1: Summary of the particle content of the standard model, along with their properties.

Standard Model particles									
Fermion: spin 1/2									
	1st generation			2nd generation			3rd generation		
	particle	charge	mass	particle	charge	mass	particle	charge	mass
Lepton	e	± 1	0.511 MeV	μ	± 1	0.105 GeV	τ	± 1	1.7768 GeV
	ν_e	0	< 2.2 eV	ν_μ	0	< 1.7 MeV	ν_τ	0	< 15.5 MeV
Quark	u	±2/3	2.4 MeV	c	±2/3	1.275 GeV	t	±2/3	172.44 GeV
	d	∓1/3	4.8 MeV	s	∓1/3	95 MeV	b	∓1/3	4.18 GeV
Gauge Boson: spin 1									
Partile	Interaction			Charge			Mass (GeV)		
γ	EM			0			0		
W	Weak			± 1			80.39		
Z	Weak			0			91.19		
g	Strong			0			0		
Higgs Boson: spin 0									
H	-			0			125.09		

Fundamental fermions are the basic building blocks of matter. They can be classified into two types: leptons that undergo only the electromagnetic (EM) and weak interactions, and quarks that in addition participate in the strong interactions. Each generation of fermions consists of a charged lepton (e, μ, τ) with one unit of electric charge, a neutral neutrino (ν_e, ν_μ, ν_τ), an up-type quark (u, c, t) with +2/3 electric charge, and a down-type quark (d, s, b) with $-1/3$ electric charge. Neutrinos are almost massless and interact only via the weak force. Each of these fermions has an associated anti-particle with the same mass but opposite charge.

There are four fundamental forces of nature: the strong interaction, the weak interaction, the electromagnetic force, and gravity, which we do not consider any further. Each of these forces is mediated by corresponding gauge bosons. The SM is a gauge quantum field theory based on the symmetry group

$$SU(3) \otimes SU(2) \otimes U(1)_Y, \tag{2.2}$$

where Y is the weak hypercharge. $SU(2)\otimes U(1)_Y$ is the symmetry of the electroweak interaction, and the $SU(3)$ group describes the symmetry of the strong interaction. The $SU(2)$ group has three gauge bosons W_i^μ, $i = 1,2,3$, and the $U(1)_Y$ group has one gauge boson B^μ. The linear combination of these gauge bosons form the physical gauge particles: photon (γ), $W^\pm$ and Z. The photon is a massless vector boson which has two polarizations, and constitutes the force carrier of the electromagnetic interaction. The $W^\pm$ and Z bosons are massive vector bosons that mediate the weak force. The gauge boson of the $SU(3)$ group is the gluon (g), which is a massless particle that carries color charge and couples to quarks and other gluons.

The Lagrangian of the SM can be decomposed into four terms:

$$\mathcal{L} = \mathcal{L}_{Gauge} + \mathcal{L}_{kin} + \mathcal{L}_Y + \mathcal{L}_H. \tag{2.3}$$

The first term is the kinetic term of the SM gauge bosons:

$$\mathcal{L}_{Gauge} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} - \frac{1}{4}G^A_{\mu\nu}G^{A\mu\nu}, \tag{2.4}$$

where $a = 1,2,3$; $A = 1, \ldots ,8$. The field strengths are:

$$\begin{aligned} B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \\ F^a_{\mu\nu} &= \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g_2\epsilon^{abc}W^b_\mu W^c_\nu \\ G^A_{\mu\nu} &= \partial_\mu G^A_\nu - \partial_\nu G^A_\nu + g_s f^{ABC}G^B_\mu G^C_\nu, \end{aligned} \tag{2.5}$$

where g_2 is the weak coupling constant and g_s is the strong coupling constant. The Levi-Civita tensor ϵ^{abc} and the Gell-Mann tensor f^{ABC} are the structure constants of the $SU(2)$ and $SU(3)$ group, respectively.

The second term of Eq. 2.3 describes the fermion fields and their gauge interactions. Fermion fields are classified into left-chiral and right-chiral states according to their chirality under Lorentz transformations. The left-chiral states are $SU(2)$

doublets, while the right-chiral states are singlets:

$$\begin{aligned}
\text{Lepton doublet}: &\quad L = \begin{pmatrix} \nu \\ e \end{pmatrix}_L \\
\text{Lepton singlet}: &\quad \overline{\ell}_R \\
\text{Quark doublet}: &\quad Q = \begin{pmatrix} u \\ d \end{pmatrix}_L \\
\text{Quark singlet}: &\quad \overline{u}_R, \overline{d}_R.
\end{aligned} \tag{2.6}$$

The $\mathcal{L}_{kin}$ term has the form:

$$\mathcal{L}_{kin} = i\sum_{j=1}^{3}\left(L_j^\dagger \overline{\sigma}^\mu D_\mu L_j + \bar{\ell}_j^\dagger \overline{\sigma}^\mu D_\mu \bar{\ell}_j + Q_j^\dagger \overline{\sigma}^\mu D_\mu Q_j + \bar{u}_j^\dagger \overline{\sigma}^\mu D_\mu \bar{u}_j + \bar{d}_j^\dagger \overline{\sigma}^\mu D_\mu \bar{d}_j\right), \tag{2.7}$$

where j is the generation index. D_μ is the covariant derivative, which has the form:

$$D_\mu = (\partial_\mu + i\frac{g_1}{2}YB_\mu + i\frac{g_2}{2}\sigma_a W_\mu^a + i\frac{g_s}{2}\lambda_A G_\mu^A). \tag{2.8}$$

λ_A are the eight Gell-Mann matrices that generate the SU(3) group.

The gauge bosons are required to be massless in order to preserve gauge invariance. This is the case for gluons and photons. However, the observed $W^\pm$ and Z bosons are massive particles, which implies that the electroweak symmetry is broken. A mechanism named "spontaneous symmetry breaking"[4, 5, 6] is introduced to give masses to the $W^\pm$ and Z bosons, while preserving the renormalizability of the theory. The basic idea is that the ground state of the theory is not invariant under symmetry transformation. This is achieved by adding a complex $SU(2)$ doublet scalar field, which is known as the Higgs field, to the SM Lagrangian.

The Lagrangian of the Higgs field reads as

$$\mathcal{L} = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi), \tag{2.9}$$

with the potential

$$V(\phi) = \mu^2\phi^*\phi + \lambda(\phi^*\phi)^2. \tag{2.10}$$

Provided $\mu^2 < 0$ this potential has minimum at

$$v = \sqrt{-\mu^2/\lambda}. \tag{2.11}$$

There is an infinite number of ground states lying along a ring described by $\phi^*\phi = -\mu^2/\lambda$. The system making a choice of the ground state will spontaneously break the symmetry, since a gauge transformation will take the system to a different vacuum state. Adopting the unitary gauge, we choose the ground state to be along the real axis of the lower component of the scalar doublet. The Higgs field can then be expanded around the vacuum as:

$$\langle\phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+H \end{pmatrix}. \tag{2.12}$$

Substituting the Higgs field to the kinetic term $(D_\mu\phi)^\dagger(D^\mu\phi)$, we find:

$$(D_\mu\phi)^\dagger(D^\mu\phi) = \frac{1}{2}(\partial_\mu H)^2 + \frac{g_2^2(v+H)^2}{4}W^{+\mu}W^-_\mu + \frac{(v+H)^2}{8}(g_2W^3_\mu - g_1B_\mu)^2. \tag{2.13}$$

Defining

$$\begin{aligned} Z_\mu &= \cos\theta_{\mathrm{W}}W^3_\mu - \sin\theta_{\mathrm{W}}B_\mu, \\ A_\mu &= \cos\theta_{\mathrm{W}}B_\mu + \sin\theta_{\mathrm{W}}W^3_\mu, \end{aligned} \tag{2.14}$$

with the weak mixing angle (Weinberg angle) θ_W defined as:

$$\tan\theta_W \equiv \frac{g_1}{g_2}, \tag{2.15}$$

the kinematic term can be rewritten as:

$$(D_\mu\phi)^\dagger(D^\mu\phi) = \frac{1}{2}(\partial_\mu H)^2 + \frac{g_2^2(v+H)^2}{4}W^{+\mu}W^-_\mu + \frac{(v+H)^2g_2^2}{8cos^2\theta_W}Z_\mu Z^\mu. \tag{2.16}$$

We recognize the second and third terms of Eq. 2.16 as the mass term of the W and Z bosons. On the other hand, leptons and quarks gain mass through a Yukawa coupling to the Higgs particle:

$$\mathcal{L}_Y = (-Y_e\bar{L}\phi l_R - Y_d\bar{Q}Hd_R - Y_u\bar{Q}\sigma_2\phi^\dagger u_R + h.c.), \tag{2.17}$$

which is the third term of Eq. 2.3.

An extensive set of experiments has been carried out in the past few decades to test the SM. The W and Z bosons were discovered in 1983 at CERN [7, 8, 9, 10]. The top quark, which is the heavist quark, were observed in 1995 at Fermilab by the CDF and D0 experiment [11, 12]. In 2012, the CMS and ATLAS experiment at CERN announced the observation of a boson with a mass around 125 GeV [13, 14], which was later confirmed to be the SM Higgs bosons. This discovery completed the last missing piece of the SM.

The predictions of the SM agree well with the observed data in a wide range of measurements. For example, Figure 2.1 shows a summary of the cross-sections of various SM processes measurements made by the CMS experiment. Excellent agreement is found between the SM predictions and measured results over many orders of magnitude.

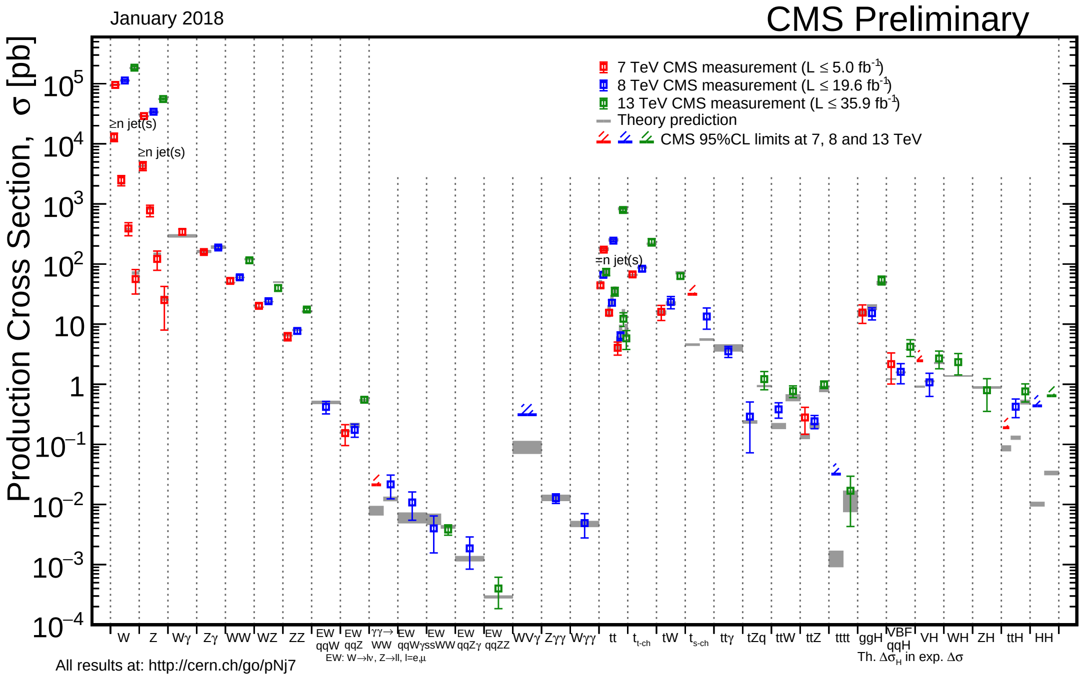


Figure 2.1: Compilation of different SM predictions compared to the corresponding CMS cross section measurements. Where the theory values appear to be absent they actually agree with the data within the experimental uncertainties.

2.2 Problems of the Standard Model

Although the SM is remarkably successful at explaining a wide range of phenomena, it does leave many questions unanswered. One major problem is the omission of gravity - the most familiar fundamental force. Phenomena at large scale can be well described by general relativity, however, there is not yet a satisfactory theory, which can incorporate the gravity at the subatomic scale. Therefore the SM is believed to be a low-energy approximation of a theory existing at the 10^{19} GeV Planck scale (M_{Pl}), the energy at which the gravitational force is comparable to the other forces and can no longer be ignored.

Another outstanding problem of the SM is the lack of an explanation for dark matter (DM) that makes up about 25% of the mass-energy in the universe and is not luminous. The dark matter was discovered through the measurement of velocity dispersion and rotation curves of galaxies [15], whose results indicated that the luminous matter cannot explain the observed curves on its own and some invisible dark matter is needed to account for the observation. Measurements show that roughly 80% of the matter in the universe is composed of dark matter. Despite the extensive experimental efforts of searching for DM, we still know little about the nature of DM and its interactions with normal matter as described by the SM.

Besides the lack of particle candidates for DM, there are other observations that the SM cannot explain. One of these mysteries is the large discrepancy between the electroweak and the Planck scale, which is known as the “hierarchy problem”. The Higgs mass is measured to be 126 GeV, 17 orders of magnitude lower than the Planck scale. The physical mass of the Higgs boson is a combination of the tree-level bare mass and high order corrections. Because the loop corrections of the Higgs mass contain quadratic divergences, its bare mass has to cancel the large corrections with an accuracy up to 34 digits leaving a small mass at electroweak scale. This unnatural cancellation is called “fine tuning problem”, and suggests the appearance of new physics which can stabilize the Higgs mass in a different way.

Due to these issues, the SM is considered to be an incomplete theory. Many new theories beyond the SM have been introduced to incorporate the unexplained features in the SM. Some of the popular extensions of the SM include supersymmetry, composite Higgs models, and large extra dimensions. In this thesis we will focus on the supersymmetric extension of the SM.

2.3 Supersymmetry

Supersymmetry (SUSY) [16, 17, 18, 19, 20, 21, 22] is a favored extension of the SM that provides solutions for many of the problems of the SM. It unifies the description of forces and matter by inducing a symmetric transformation between bosons and fermions. The supersymmetry transformations are generated by an operator $\mathcal{Q}$, which is called the supercharge. The operator $\mathcal{Q}$ acts on a state, turning a fermionic component into a bosonic component, and vice verse:

$$\mathcal{Q}\,|\mathrm{Boson}\rangle = |\mathrm{Fermion}\rangle\,, \quad \mathcal{Q}\,|\mathrm{Fermion}\rangle = |\mathrm{Boson}\rangle\,. \tag{2.18}$$

The supercharge itself is a spinor, and its hermitian conjugate $\mathcal{Q}^\dagger$ is also a symmetry generator. They both satisfy the following anti-commutation relations:

$$\{\mathcal{Q}_a, \mathcal{Q}^\dagger_{\dot{a}}\} = (\sigma^\mu)_{a\dot{a}} P_\mu, \tag{2.19}$$

$$\{\mathcal{Q}_a, \mathcal{Q}_b\} = 0, \quad \{\mathcal{Q}^\dagger_{\dot{a}}, \mathcal{Q}^\dagger_{\dot{b}}\} = 0, \tag{2.20}$$

$$[\mathcal{Q}_a, P_\mu] = 0, \quad [\mathcal{Q}^\dagger_{\dot{a}}, P_\mu] = 0, \tag{2.21}$$

where $P_\mu = i\partial_\mu$ is the momentum operator. The first anti-commutation relation, Eq. (2.19), indicates that the product of two supersymmetric transformations results in a translation in space-time, i.e. supersymmetry is a space-time symmetry. This suggests the extension of Minkowski space-time with two anti-commuting coordinates to form the superspace.

2.3.1 The Minimal Supersymmetric Standard Model

SUSY doubles the number of particles by pairing each SM particle with a SUSY partner - each fermion has a bosonic partner and each gauge boson has a fermionic partner. Each SM particle and its superpartner have the same quantum number except for the spin. The SM fields and superpartners are grouped together into supermultiplets. With this general concept of SUSY, one can build models with any number of supermultiplets. It is appropriate to first consider the simplest version of SUSY models, which contains the minimum number of SUSY particles and new interactions. Such a model is referred to as the Minimal Supersymmetric Standard Model (MSSM).

In the MSSM, all the SM fermions are taken to be the fermionic components of chiral supermultiplets. Each of them is paired with a spin-0 scalar partner, whose name is obtained by adding a prefix "*s*" to the name of the corresponding SM particle. For example, the scalar partner of an electron is called a selectron, and the scalar quark is called squark. The SM gauge fields are taken to be the members of vector supermultiplets and are paired with spin-1/2 superpartners. The naming convention for the spin-1/2 superpartners is to add "*ino*" after the name of the SM gauge bosons. So for B, W and gluons, the superpartners are Bino, Winos, and gluinos. Different from the SM, in the MSSM two Higgs doublets are required to break the electroweak symmetry. The scalar Higgs fields are paired with spin-1/2 partners, which are named Higgsinos, forming the chiral multiplets. Table 2.2 lists the particle content of the MSSM.

Table 2.2: Supermultiplets in the Minimal Supersymmetric Standard Model.

MSSM particles				
	SUSY particle fields		Particle fields	
	name	symbol	name	symbol
sfermion/femion	slepton	$(\tilde{\nu}_\ell, \tilde{\ell})_L$ $\tilde{\ell}_R$	lepton	$(\nu_\ell, \ell)_L$ ℓ_R
	squark	$(\tilde{u}, \tilde{d})_L$ $\tilde{u}_R$ $\tilde{d}_R$	quark	$(u, d)_L$ u_R d_R
gaugino/gauge boson	Bino	$\tilde{B}$	B boson	B
	Wino	$\tilde{W}^\pm$, $\tilde{W}^0$	W boson	$W^\pm$, W^0
	Gluino	$\tilde{g}$	Gluon	g
Higgsino/Higgs	Higgsino	$(\tilde{H}_d^0, \tilde{H}_d^-)$ $(\tilde{H}_u^+, \tilde{H}_u^0)$	Higgs	(H_d^0, H_d^-) (H_u^+, H_u^0)

Constructing a supersymmetric Lagrangian in the 4-dimensional space-time can be very difficult and tedious. The introduction of superspace greatly simplifies the calculations. Superspace extends the usual space-time coordinates with two additional fermionic (or Grassmannian) coordinates. Points in superspace thus have coordinates

$$(x^\mu, \theta, \theta^\dagger), \tag{2.22}$$

where x^μ are the regular 4-dimensional space-time coordinates and θ, $\theta^\dagger$ are anti-commuting Grassmann variables. Supermultiplets are described by functions over the superspace, which are referred to as superfields. A chiral superfield can be expressed by an expansion in Grassmann variables as

$$\Phi(x,\theta,\theta^\dagger) = \phi(x) + \theta\psi(x) + \frac{1}{2}\theta\theta F(x) + (\text{space-time derivatives acting on } \phi \text{ and } \psi), \tag{2.23}$$

where ϕ is a complex scalar field, ψ is a left-chiral spinor, and $F(x)$ is an auxiliary field. The auxiliary field is introduced to ensure that the number of bosonic and fermionic degrees of freedom match with each other for both on-shell and off-shell cases.

A general vector superfield fixed by the Wess-Zumino gauge [23] can be expressed as:

$$V = -\theta\sigma^\mu\bar{\sigma}V_\mu(x) + i\theta^2\bar{\theta}\bar{\lambda}(x) - i\bar{\theta}^2\theta\lambda(x) + \frac{1}{2}\theta^2\bar{\theta}^2 D(x). \tag{2.24}$$

Here, $V_\mu(x)$ is the gauge field, $\lambda(x)$ and $\bar{\lambda}(x)$ are gauginos, and $D(x)$ is the auxiliary field.

Then the supersymmetric Lagrangian can be conveniently written as:

$$\mathcal{L} = \int d^4\theta \Phi_i^\dagger e^{gV}\Phi_i + \int d^2\theta(\frac{1}{4}W_\alpha^a W^{\alpha a} + h.c.) + \int d^2\theta(W(\Phi) + h.c.), \tag{2.25}$$

where $W(\Phi)$ is the super potential of the chiral fields. The superpotential of the MSSM has the form:

$$W = l^i\Phi_i + \frac{1}{2}m^{ij}\Phi_i\Phi_j + \frac{1}{6}y^{ijk}\Phi_i\Phi_j\Phi_k. \tag{2.26}$$

The quantity $W^{\alpha a}$ in Eq. (2.25) is the field-strength of the vector superfield,

$$W_\alpha = -\frac{1}{4}\bar{D}^\alpha\bar{D}_\alpha D_\alpha V, \tag{2.27}$$

where D_α is the covariant derivatives in superspace.

SUSY provides a solution to the hierarchy problem. In the loop corrections to the Higgs self energy, the contributions from fermions and bosons cancel with each other, leaving a light Higgs mass at the electroweak scale. This is illustrated in Fig. 2.2.

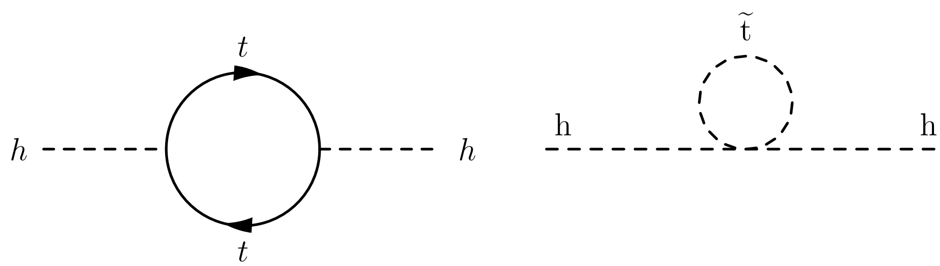


Figure 2.2: Contributions of SM and SUSY loop corrections to the mass of Higgs.

In the MSSM, a new symmetry called R-parity is introduced, where the R quantum number is defined as

$$R = (-1)^{3(B-L)+2S}, \tag{2.28}$$

where B is the baryon number, L is the lepton number, and S is the spin. All SM particles have even R-parity, whereas all supersymmetric particles have odd R-parity. If R-parity is conserved, the decays of lightest SUSY particle (LSP) into SM particles are forbidden, which implies that the LSP is stable and can serve as a dark matter candidate. For DM which can be weakly interacting massive particles (WIMPs), the interactions between the DM and normal particles are so weak that the DM will escape the detector, resulting in a momentum imbalance in the transverse plane. This signature is a crucial element for all the searches for SUSY at particle colliders.

2.3.2 Supersymmetry Breaking

As mentioned above, the quadratic divergence of the Higgs mass arising from SM loop corrections can be exactly cancelled by the contribution from SUSY partners when the masses of all states in a supermultiplet are degenerate. However, any superpartners with the same mass of leptons or light quarks should have been observed. Therefore supersymmetry must be a broken symmetry at the low energy scale if it is realized in nature. In order to maintain the good ultraviolet behaviour of supersymmetry despite a mass splitting between the SM particles and their SUSY partners, a soft symmetry breaking is considered. The idea is that supersymmetry is unbroken at some high energy scale at which the exact SUSY is preserved. At the low energy scale, supersymmetry breaking takes place, allowing the SUSY particles to obtain heavier masses than their SM partners. This can be achieved by introducing soft breaking terms into the Lagrangian. The MSSM soft-breaking terms are:

$$\begin{aligned} \mathcal{L}_{soft} &= -\tfrac{1}{2}(M_1\tilde{B}\tilde{B} + M_2\tilde{W}\tilde{W} + M_3\tilde{g}\tilde{g}) \\ &\quad -(a_u\tilde{Q}H_u\bar{\tilde{u}} + a_d\tilde{Q}H_d\bar{\tilde{d}} + a_e\tilde{L}H_e\bar{\tilde{d}}) \\ &\quad -m_{H_u}^2|H_u|^2 - m_{H_d}^2|H_d|^2 - bH_uH_d \\ &\quad -M_Q^2|\tilde{q}_L|^2 - M_U^2|\tilde{u}_R|^2 - M_D^2|\tilde{u}_D|^2 - M_L^2|\tilde{l}_L|^2 - M_E^2|\tilde{l}_R|^2. \end{aligned} \tag{2.29}$$

The coefficient of each term is called the soft SUSY breaking parameter. The appearance of the soft breaking terms introduces 105 more independent parameters. Assuming these breaking terms originate from the same mechanism, the scale of the SUSY breaking can be denoted as $m_{\rm SUSY}$:

$$\begin{aligned} M_{1,2,3}, a_{u,d,e} &\sim m_{\rm SUSY}, \\ M^2_{Q,U,D,L,E,H_u,H_d}, b &\sim m^2_{\rm SUSY}. \end{aligned} \tag{2.30}$$

To understand the origin of these breaking terms, we can consider the soft supersymmetry breaking as an effective description of the spontaneous supersymmetry breaking, i.e. the Lagrangian is invariant under supersymmetric transformations but the vacuum state is not. The spontaneous supersymmetry breaking happens when the supercharges do not annihilate the vacuum, i.e.,

$$\mathcal{Q}\,|0\rangle \neq 0. \tag{2.31}$$

Using the commutation relation Eq. (2.19) of the supercharges, we can write the Hamiltonian as

$$H = P^0 = \frac{1}{4}(\mathcal{Q}_1\bar{\mathcal{Q}}_{\dot{1}} + \bar{\mathcal{Q}}_{\dot{1}}\mathcal{Q}_1 + \mathcal{Q}_2\bar{\mathcal{Q}}_{\dot{2}} + \bar{\mathcal{Q}}_{\dot{2}}\mathcal{Q}_2), \tag{2.32}$$

which has positive vacuum expectation value (VEV) when SUSY is spontaneous broken, i.e.

$$\langle 0|H|0\rangle > 0. \tag{2.33}$$

SUSY breaking within the MSSM is not easy, because it predicts SUSY particles that are lighter than their SM partners. The way around this problem is to assume the existence of a hidden sector that is uncharged under the SM gauge group. The SUSY breaking originates in the hidden sector and communicate to the visible MSSM sector by a set of messenger fields. The messenger sector transmits the SUSY breaking via loop corrections, allowing the SUSY particles to become massive. There are several well studied mechanisms for mediating the SUSY breaking to the visible sector, including gravity mediation, gauge mediation and anomaly mediation. The search presented in this thesis is motivated by the gauge mediated SUSY breaking (GMSB) [24, 25, 26].

The GMSB mechanism assumes that the messenger sector couples to the visible sector via flavor-blind gauge interactions. Suppose the hidden sector has some supermultiplet S which has VEV $\langle F\rangle$, where F is the auxiliary field of S. A set of messenger fields $\Phi_I, \bar{\Phi}_I$ couple to the hidden sector via an interaction term:

$$W = \sum_I y_I S \Phi_I \bar{\Phi}_I. \tag{2.34}$$

If $\langle F\rangle$ is non-zero, mass splittings are generated in the messenger sector. The messenger fields are charged under the SM gauge. Therefore the SUSY breaking is mediated from the hidden sector to the visible sector through gauge interaction. A scheme of the gauge mediation is shown in Figure 2.3.

The gauginos become heavier due to the radiative corrections from the messenger particle loops, as illustrated in the Feynman diagram of Figure 2.4.

Assuming that the messenger fields have mass M_{mess}, we can integrate out the

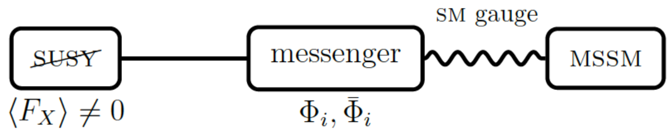


Figure 2.3: Sketch of SUSY breaking in the hidden sector that is mediated to the MSSM sector via the messenger fields.

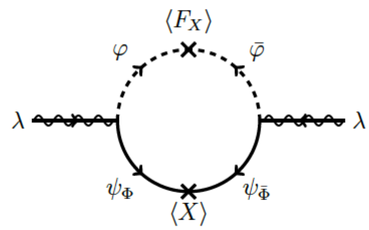


Figure 2.4: Radiative corrections to the mass of SUSY particles.

messenger fields to obtain the effective SUSY breaking. The resulting soft mass is proportional to

$$M_{soft} \sim \frac{\langle F \rangle}{M_{mess}}. \tag{2.35}$$

If the M_{mess} and $\sqrt{\langle F \rangle}$ are of the same order, the SUSY breaking can be realized in a scale as low as the electroweak scale. Requiring gravity effects to be negligible, one can also impose an 10^{15} GeV upper bound on the scale of M_{mess}.

When the electroweak symmetry is broken, only the SU(3) and U(1) gauge remains unbroken. Similar to the mixing of the gauge bosons in the SM, the Bino, Winos and higgsinos will mix to form mass eigenstates. The mixing of two neutral gauginos $(\widetilde{B}, \widetilde{W}^0)$ and two neutral higgsinos $(\widetilde{h}_d^0, \widetilde{h}_u^0)$ will give rise to four neutral mass eigenstates called neutralinos, denoted as $\widetilde{\chi}_{1,2,3,4}^0$. The mixing mass matrix for the neutralinos is:

$$M_N = \begin{pmatrix} M_1 & 0 & -M_Z sin\beta S_W & M_Z cos\beta S_W \\ 0 & M_2 & M_Z cos\beta C_W & -M_Z sin\beta C_W \\ -M_Z cos\beta C_W & M_Z sin\beta C_W & 0 & \mu \\ M_Z sin\beta S_W & -M_Z cos\beta S_W & \mu & 0 \end{pmatrix}. \tag{2.36}$$

Similarly, the charged gaugino-higgsino mixing will give rise to four charged mass eigenstates. The mixing matrix is:

$$M_C = \begin{pmatrix} M_2 & \sqrt{2} M_W sin\beta \\ \sqrt{2} M_W cos\beta & \mu \end{pmatrix}. \tag{2.37}$$

The resulting mass eigenstates, $\widetilde{\chi}_1^{\pm}, \widetilde{\chi}_2^{\pm}$, are called the charginos.

2.3.3 Phenomenology of General Gauge Mediation Supersymmetry

According to the Goldstone theorem, for every global symmetry with spontaneous breaking there exist a massless Nambu-Goldstone boson. Similar to the goldstone boson, spontaneous SUSY breaking gives rise to a massless goldstino. When SUSY is imposed as a local symmetry, the goldstino is absorbed by the gravitino, becoming the longitudinal component of the gravitino. The gravitino mass scales as:

$$m_{2/3} \sim \langle F \rangle / M_{PL}. \tag{2.38}$$

The SUSY breaking scale could be very low in gauge-mediated models, making the gravitino to have a mass roughly in the 1 eV to 1 GeV range. Therefore, one important consequence of GMSB is that the gravitino is the LSP. All the heavier SUSY particles will eventually decay to the gravitino, either directly or through a cascade decay chain. Therefore, the phenomenology of GMSB is mainly determined by the nature of the next-to-lightest SUSY particle (NLSP).

The decay length of the NLSP is proportional to $\langle F \rangle^2$. Depending on the scale of the messenger mass, the decay of the NLSP can be prompt, long-lived, or very long-lived outside the detector. In this search, we will focus on the prompt scenario: the NLSP immediately decay to its SM partner and a gravitino. In the general gauge-mediated (GGM) models, the NLSP could be the lightest neutralino. The general neutralino NLSP scenarios lead to very interesting signatures involving photons, W, Z or Higgs bosons in the final states.

The neutralino NLSP is a mixture of the neutral gauginos and higgsinos. Depending on the relative hierarchy among the soft masses, the NLSP can be one of three cases:

- Bino-Like: if $|M_1| < |\mu|, |M_2|$, the neutralino NLSP is mostly a Bino. It will dominantly decay through the $\tilde{\chi}_1^0 \to \gamma + \tilde{G}$ channel. The typical signature of Bino-like neutralino pair production is large transverse missing momentum plus a pair of photons in the final states.

- Wino-like: if $|M_2| < |\mu|, |M_1|$, the neutralino NLSP is dominated by the Wino component. Since the wino multiplet consists of both charged ($\tilde{W}^{\pm}$) and neutral ($\tilde{W}^0$) gauginos, the lightest chargino ($\tilde{\chi}_1^{\pm}$) can be as light as the neutralino NLSP. In this case, the $\tilde{\chi}_1^0$ and $\tilde{\chi}_1^{\pm}$ are nearly mass degenerate, and are called "co-NLSP". The wino-like $\tilde{\chi}_1^0$ can decay into $\gamma + \tilde{G}$ or $Z + \tilde{G}$. The branching fraction of the wino-like $\tilde{\chi}_1^0$ decay is shown in Figure 2.5. On the other hand, the $\tilde{\chi}_1^{\pm}$ decays to the $W^{\pm} + \tilde{G}$ final states, where the W boson can decay hadroniclly or leptonically. Depending on the decay mode of the NLSP and the W/Z bosons, the final states can contain photons or leptons plus large transverse momentum.

- Higgsino-like: if $\mu < |M_1|, |M_2|$, the NLSP is higgsino-like. The decay mode of the NLSP varies with the mass parameters and mixing angles. In the case where the $h + \tilde{G}$ decay mode is preferred, one can use events containing Higgs bosons to probe the production of higgsino-like NLSP and enhance the search sensitivity.

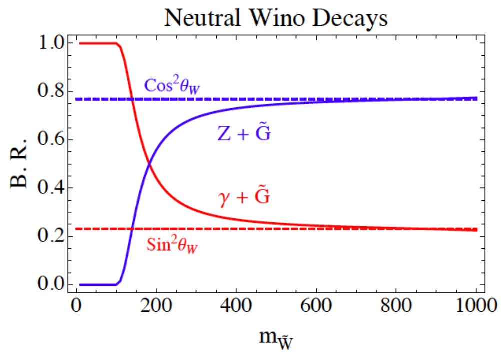


Figure 2.5: Branching ratio of a wino-like neutralino NLSP as a function of the neutralino mass.

In this thesis we present a search for GGM SUSY signatures involving the production of wino-like $\tilde{\chi}_1^0\tilde{\chi}_1^{\pm}$ pairs in final states with a photon, a lepton (either e or μ) and significant missing momentum. Photons (leptons) in the final state are the result of neutralino (chargino) decays to a gravitino and a photon (W boson), $\tilde{\chi}_1^0 \to \gamma + \tilde{G}$ ($\tilde{\chi}_1^{\pm} \to W^{\pm} + \tilde{G}$, with $W^{\pm} \to l^{\pm} + \nu$). The additional lepton in the final state not only suppresses backgrounds from QCD processes, but also offers the unique opportunity to probe the nature of the neutralino NLSP.

2.3.4 Signal Scenarios

There are many different manifestations of SUSY, each with different particle contents and mechanisms for SUSY breaking. However, many of these models predict a similar phenomenology, which inspired the formulation of simplified models [27]. In this analysis we consider both the simplified models and GGM senarios to interpret the search results.

Simplified Models of Supersymmetry

The simplified models use an effective Lagrangian to describe the new particles and their decays. The masses of the particles and the branching ratio of their decay modes can be tuned directly. In this thesis, only the production of a pair of primary particles is considered for the simplified models. Figure 2.6 shows the cross sections for the production of $\tilde{g}\tilde{g}$, $\tilde{q}\tilde{\bar{q}}$ and $\tilde{\chi}_1^0\tilde{\chi}_1^{\pm}$ pairs as a function of the primary sparticle mass. Each primary particle undergos a direct decay or a cascade decay through a SUSY particle to finally decay into LSP.

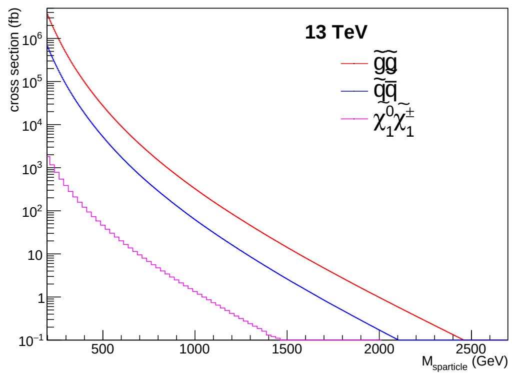


Figure 2.6: Cross sections for producing various pairs of sparticles at $\sqrt{s} = 13$ TeV.

The simplified models considered in this analysis include the following three processes:

- **TChiWG**: This model assumes the direct production of $\tilde{\chi}_1^0/\tilde{\chi}_1^{\pm}$ pairs. The $\tilde{\chi}_1^0$ and $\tilde{\chi}_1^{\pm}$ are taken to be mass degenerate. The $\tilde{\chi}_1^0$ undergos direct decay to a photon and LSP, while the $\tilde{\chi}_1^{\pm}$ decays to $W^{\pm}$ plus LSP. This model corresponds to an electroweak production mechanism of $\tilde{\chi}_1^0/\tilde{\chi}_1^{\pm}$.

- **T5WG**: This model is a simplified version of strong gluino pair production in which each gluino decays to a quark-antiquark pair and an intermediate neutralino or chargino. The mass of the neutralino and chargino are set to be the same. A 50% branching ratio is assumed for the gluino to neutralino/chargino decay, resulting again in a photon plus lepton final state.

- **T6WG**: Similar to the T5WG model, the strong production of a pair of squarks is assumed in the T6WG model. The squark decays to an SM quark plus an intermediate neutralino or chargino. The branching ratio for the squark to neutralino/chargino is also assumed to be 50%.

These processes are illustrated in the Feynman diagrams in Figure 2.7.

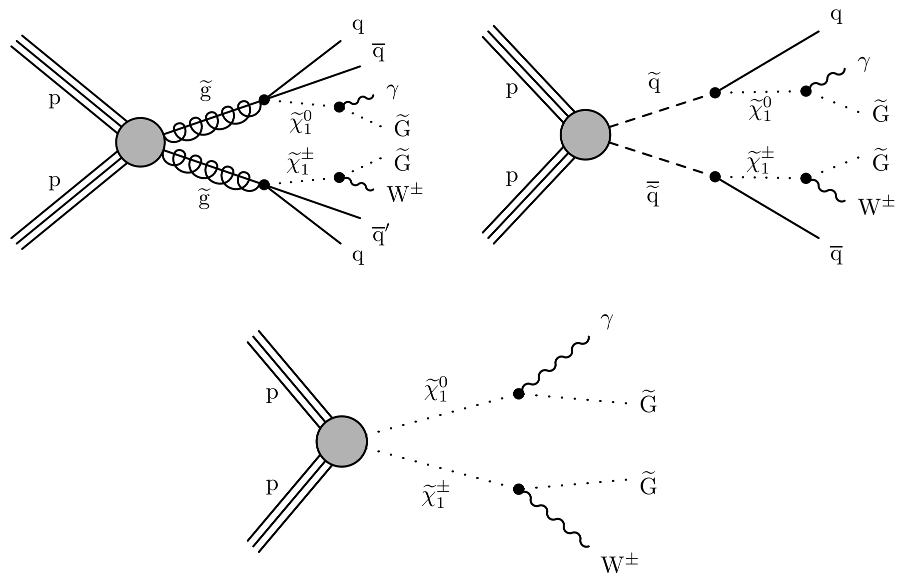


Figure 2.7: Feynman diagrams showing the production and decay modes of the signal models T5Wg (top left), T6Wg (top right), and TChiWg (bottom) considered in this analysis.

GGM Model

The masses and properties of SUSY particles in GGM models are controlled by the following 8 mass parameters:

$$M_1, M_2, M_3, \mu, m_Q^2, m_U^2, m_L^2 \text{ and } M_{mess}, \tag{2.39}$$

with M_1, M_2 and M_3 being the gaugino mass parameters, μ being the higgsino mass parameter, m_Q^2, m_U^2 and m_L^2 being the mass scales for squarks and sleptons, respectively, and M_{mess} being the messenger scale. In this thesis, we consider a bench mark model where the messenger scale is set to be 10^{15} GeV and the squarks and sleptons are set to be very heavy, so that the GGM phase space reduces to a 2D plane of two gaugino mass parameters. The gravitino mass is fixed to be 10 eV.

The following model is considered:

- GGM model: M_1, M_2 parameter scan. In this model, M_3 and μ are set to 8 TeV, allowing a scan over the M_1 and M_2 evaluated at the messenger scale. In the case when $M_1 < M2$, the NLSP is the lightest neutralino χ_1^0 which is dominated by the Bino component, while for $M_1 > M2$, the χ_1^0 and $\chi_1^{\pm}$ are almost mass degenerate and the χ_1^0 is dominated by the wino component.

2.3.5 Previous Results on GMSB SUSY Searches

Searches for GMSB signatures have been performed for many years at multiple colliders, including the LEP, Tevatron and LHC. Various analyses using events with one or more photons have been conducted to search for neutralino NLSP, but no evidence of the GMSB has been found so far.

The large e^+e^-collider (LEP) at CERN operated from 1989 to 2000 with a centre-of-mass energy of up to 209 GeV. Searches for neutralino NLSP in events with two acoplanar photons and missing energy were performed by the LEP experiments ALEPH, DELPHI, L3, and OPAL [28, 29, 30, 31]. Limits on the neutralino mass were set at 99 GeV using the combined data of the four experiments. Searches for neutralino NLSP with nonprompt decays were also performed using photons not pointing toward the interaction point. Results of these searches lead to a neutralino lower limit of 55 GeV within the minimal GMSB framework.

The Tevatron was a $p\bar{p}$ collider at Fermilab, operating at 1.8 TeV during its first phase and 1.96 TeV during the second phase. Searches for acoplanar photons with large missing transverse energy were performed by both CDF and D0 experiments [32, 33] at the Tevatron. No excess of events was observed over the backgrounds. The results were interpreted within the "Snowmass slope SPS 8" benchmark GMSB model, and limits on the neutralino mass were set up to 138 GeV.

The sensitivity to SUSY signals is largely improved when the Large Hadron Collider (LHC) started operation in 2009 with a collision energy of 7 TeV and increased to 8 TeV in 2012. Event with diphoton and large missing transverse energy is still an important signature for the production of SUSY particles with a decay chain proceeding through a binolike NLSP. Such searches were performed by both the ATLAS and CMS experiments using the 7 TeV and 8 TeV data [34, 35, 36, 37]. Lower limits of 1.3 TeV were set on the masses of gluinos. The LHC got a series of upgrades upgrade in 2013-2014 and restarted in 2015 with a center-of-mass energy of 13 TeV and an increased luminosity. The search for GMSB signals with diphotons is performed again using the 13 TeV data, extending previous limits by up to 850 GeV [38].

Besides the searches for bino-like NLSP, a variety of analyses have also been conducted to explore other neutralino NLSP scenarios. For the wino co-NLSP scenario, searches are performed using events with one photon, missing transverse energy, and either large hadronic activity or an addition lepton, targeting at strongly and weakly produced SUSY particles, respectively. Searches in the photon-lepton channel have been performed by both ATLAS and CMS using 8 TeV pp collision data [39, 40]. In the context of direct production of NLSP states, winos are excluded up to 370 GeV. This thesis improves the sensitivity of the previous CMS result obtained at 8 TeV.

Chapter 3

Experimental Apparatus

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [41], located at CERN on the France-Switzerland border, is currently the world's largest particle accelerator and collider. It is installed in a 27 km long near-circular tunnel that was constructed in the 1980s to host the Large Electron-Positron (LEP) machine. The LHC uses twin-bore magnets to allow two proton beams to travel in opposite directions. The main superconducting magnets of the LHC operate at a temperature below 2K, which is maintained by a superfluid helium cryogenic system, and generates a peak dipole field of 8.33 T. The LHC is designed to provides a peak luminosity of $L = 10^{34}$ cm^{-2}s^{-1} and a maximum center of mass energy of $\sqrt{s}$ =14 TeV for proton-proton collisions.

The protons injected to the LHC are boosted by a chain of accelerators, each of which increases the energy of the protons to a certain scale and the beams can also be delivered to other experiments at low energy. The configuration of the CERN accelerator complex is shown in Fig. 3.1. At the first stage, the protons extracted from a bottle of hydrogen gas are fed into the LINAC 2, a linear accelerator, and are boosted to the energy of 50 MeV. The beams are then injected to the Proton

Synchrotron Booster, followed by the Proton Synchrotron (PS), which brings the protons to 25 GeV. The Super Proton Synchrotron (SPS) then accelerates the proton beams to 450 GeV and injects them into the LHC. The LHC is the last element in the accelerator complex, where protons are stored and accelerated to the designed energy.

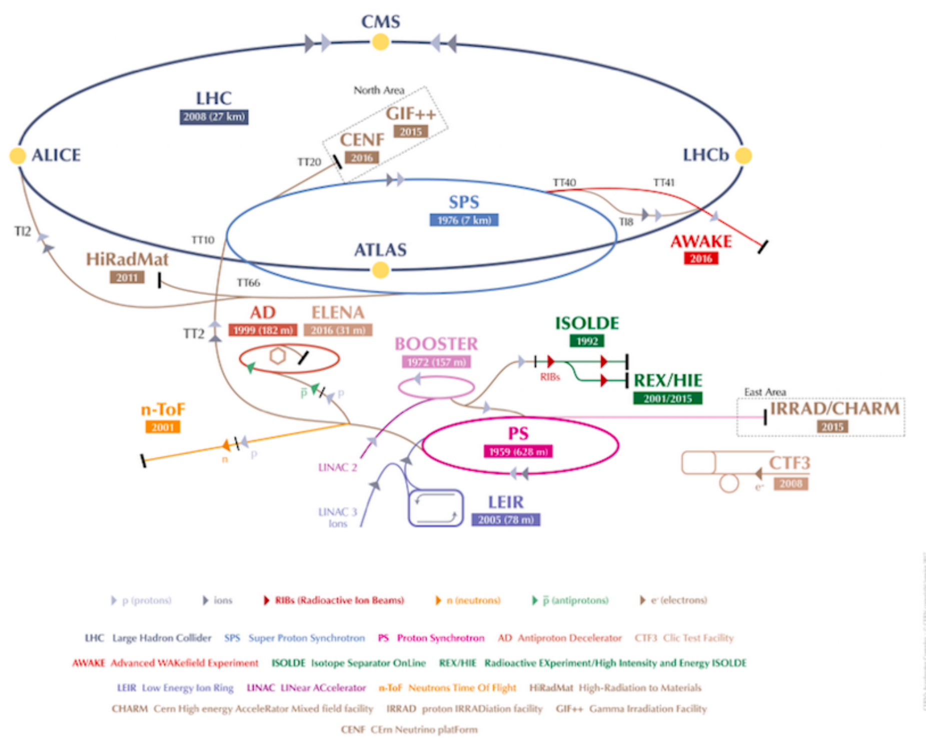


Figure 3.1: Scheme of the CERN Accelerator Complex.

The two high-energy proton beams circulate in the LHC at close to the speed of light and collide with each other at four interaction points (IP), corresponding to the location of the four detectors: CMS, ATLAS, ALICE and LHCb. The CMS and ATLAS are general-purpose detectors, which investigate a wide range of physics from the precise measurements of the SM to searches for new physics. The ALICE is a heavy-ion detector, focusing on the study of strong interactions. The LHCb experiment specializes in b quark physics, and is a single arm detector which mainly collects forward particles. In addition, there are three smaller experiments on the LHC: TOTEM (Total, elastic and diffractive cross-section measurement), LHCf ((Large Hadron Collider forward) and MoEdal (Monopole and Exotics Detector). This thesis will focus on the results of the CMS experiments.

3.2 The CMS Detector

The Compact Muon Solenoid (CMS) experiment is a hermetic, general-purpose detector located at the collision point 5 of the LHC. It is 21.6 m long, 15 m high, 15 m in diameter, and weighs about 12,500 tons. The physics goal is to investigate a wide range of phenomena including the study of the Higgs mechanism and searches for unknown particles beyond the SM. The cross-section of the interesting physics process is typically 5-10 order of magnitude smaller than the total cross-section, making it crucial to measure the momentum and time of the particles with high resolution. To meet the physics goals, the detector is designed to fulfill the following requirements:

- good muon identification and high momentum resolution;
- good electromagnetic energy resolution;
- efficient tracking and accurate momentum measurements for charged particles;
- good missing transverse momentum resolution, requiring hadron calorimeters with a large hermetic geometric coverage.

To satisfy these requirements, a 13 m long, 6 m inner diameter superconducting solenoid is built to generate a magnetic field of 4 T, providing large bending power to precisely measure the momentum of charged particles. The coil is cooled by liquid helium to operate at a temperature of 4.5 K. The flux is returned through a steel yoke consisting of five barrel wheels and four end-cap disks at each end.

Figure 3.2 shows an overview of the layout of the CMS detector. Close to the interaction point is a high granular three-layer silicon pixel detector which improves the track and vertex reconstruction, followed by a silicon strip tracker. Surrounding it is a lead-tungstate crystal electromagnetic calorimeter (ECAL) and a brass/scintillator sampling hadron calorimeter (HCAL). The tracker and calorimeter are compact enough to fit into the bore of the magnet coil. Outside the solenoid are the muon stations, embedded in the return yoke.

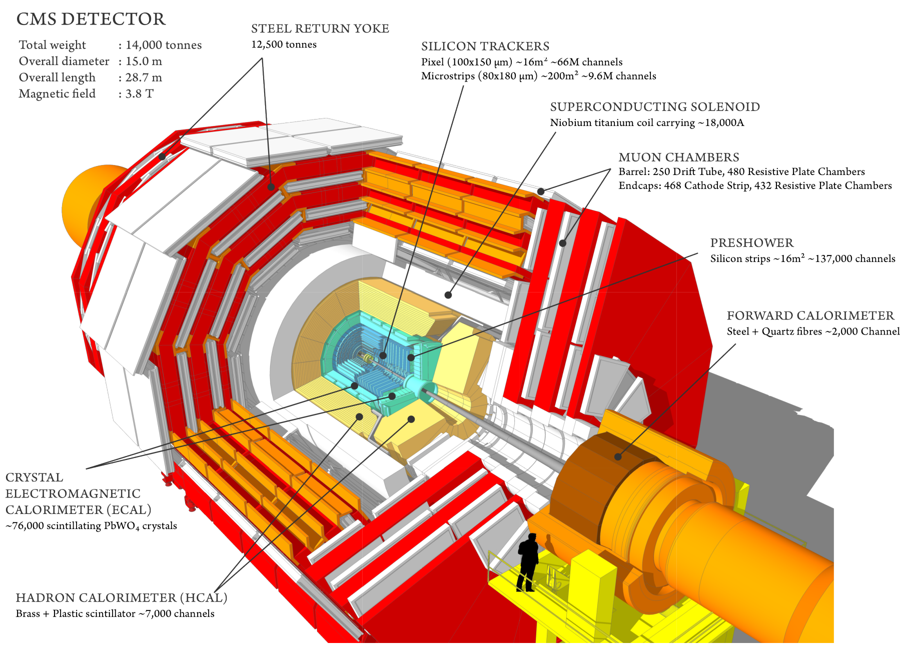


Figure 3.2: A perspective view of the CMS detector.

The CMS uses a coordinate system with the origin centered at the interaction point. The positive y-axis is chosen to be the vertically upward direction, and the x-axis points toward the center of the LHC. Therefore the z-axis is pointing along the anti-clockwise beam direction. Cylindrical coordinates (r,ϕ) are often used to describe the transverse plane, with ϕ being the azimuthal angle measured from the x-axis and r being the radial coordinate in the plane. The momentum and energy in the transverse plane, denoted as p_T and E_T, are usually used as kinematic variables for physics objects, because the transverse components are invariant under Lorentz boosts in the z-direction. The polar angle θ is measured from the z-axis, and the pseudorapidity is defined as:

$$\eta = -\ln \tan(\theta/2). \tag{3.1}$$

3.2.1 The Inner Tracking System

The tracking system is designed to provide a precise and efficient reconstruction of the trajectories of charged particles. At the same time, accurate measurements of secondary vertices and impact parameters are necessary for the estimation of the positions of primary vertices as well as the identification of heavy flavors. At the designed luminosity of the LHC, about 1000 particles are expected to hit the tracker at each bunch crossing. To keep the occupancy at around 1% in such challenging operation conditions, pixelated detectors and micro-strip detectors are used at the small radii and large radii regions, respectively. The sensors are solely made of radiation tolerant silicons. With a sensitive area of about 200 m^2, the CMS tracker is so far the largest silicon tracker.

The tracker has a cylindrical volume of 5.8 m in length and 2.5 m in diameter, embedded in the homogeneous magnetic field of 3.8 T. A schematic view of the CMS tracker is shown in Fig. 3.3. In the barrel region, the pixel detector consists of three cylindrical layers at radii between 4.4 cm and 10.2 cm and the strip tracker consists of ten layers extending to 110 cm in radius. Both subdetectors are complemented by disks on each side, extending to 290 cm in z and covering a pseudorapidity range up to $|\eta| < 2.5$.

The pixel detector consists of three barrel layers located at radii of 4.4, 7.3 and 10.2 cm, closed by two disks on each side at z = $\pm$ 34.5 and $\pm$ 46.5 cm. It contains a total number of about 66 million silicon pixels with a cell size of 100 $\times$150 μm^2. It provides good track resolution in both the r-ϕ and z directions, allowing a three-dimensional (3-D) vertex reconstruction. It is essential for the identification of heavy

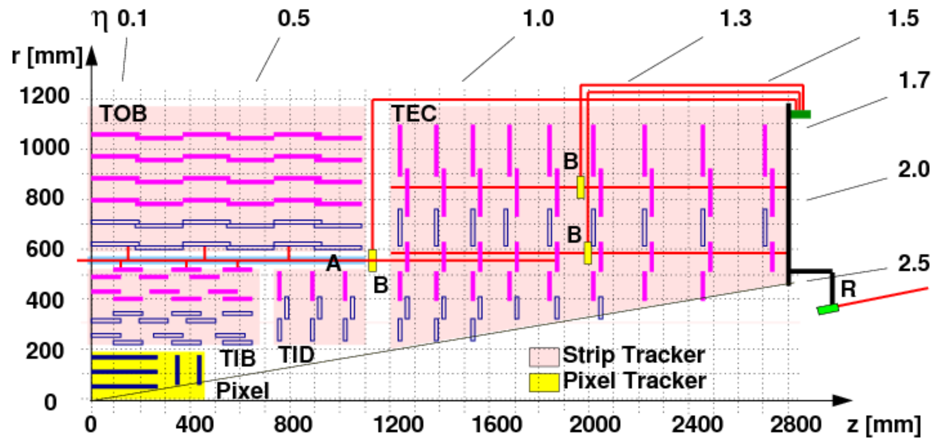


Figure 3.3: Layout of the CMS inner tracking system in an r-z view, showing the pixel detector and the silicon strip tracker.

flavors and forming seeds for the track finding.

Surrounded the pixel system is the strip tracker. It is composed of four subsystems. The Tracker Inner Barrel (TIB) and Disk (TID) extend to a radius of 55 cm, and are composed of four barrel layers and three endcap disks on each side. The sensors used by the TIB/TID are 320 μm thick silicon micro-strips. The strips in the barrel are parallel to the beam pipe and have a pitch of 80 μm on layer 1,2 and 120 μm on layer 3,4, providing a position resolution of 23-35 μm. The disks use wedge-shaped sensors, with a pitch varies between 100 μm and 141 μm. The innermost two layers of the TIB and the first two rings of the TID also carry a second module, mounted back-to-back with a stereo angle of 100 mrad, in order to measure the z (r) coordinate in the barrel (disks).

The Tracker Outer Barrel (TOB) covers the radius from 55 cm to 116cm and consists of six barrel layers, with the innermost two layers having double-modules. The sensors are made of 500 μm thick micro-strips, with pitches vary from 183 μm to 122 μm. The TOB covers z $<$ 118 cm. The rest of the tracker volume is occupied by the Tracker Endcaps (TEC), composed of 9 disks on each side. 300 μm silicons are used on the inner fours rings, and 500 μm thick strips are used in the outer rings.

Figure 3.4 shows the resolution of the pixel detectors during the Run-II operation. A 8-10 μm in $r - \phi$ and 20-25 μm in z resolution is achieved in the pixel barrel and a 15-20 μm resolution is achieved in the forward pixel detector. Figure 3.5 shows the resolution of the strip detectors. A 20-30 μm in TIB and 20-45 μm TOB hit resolution is achieved during Run-II operation.

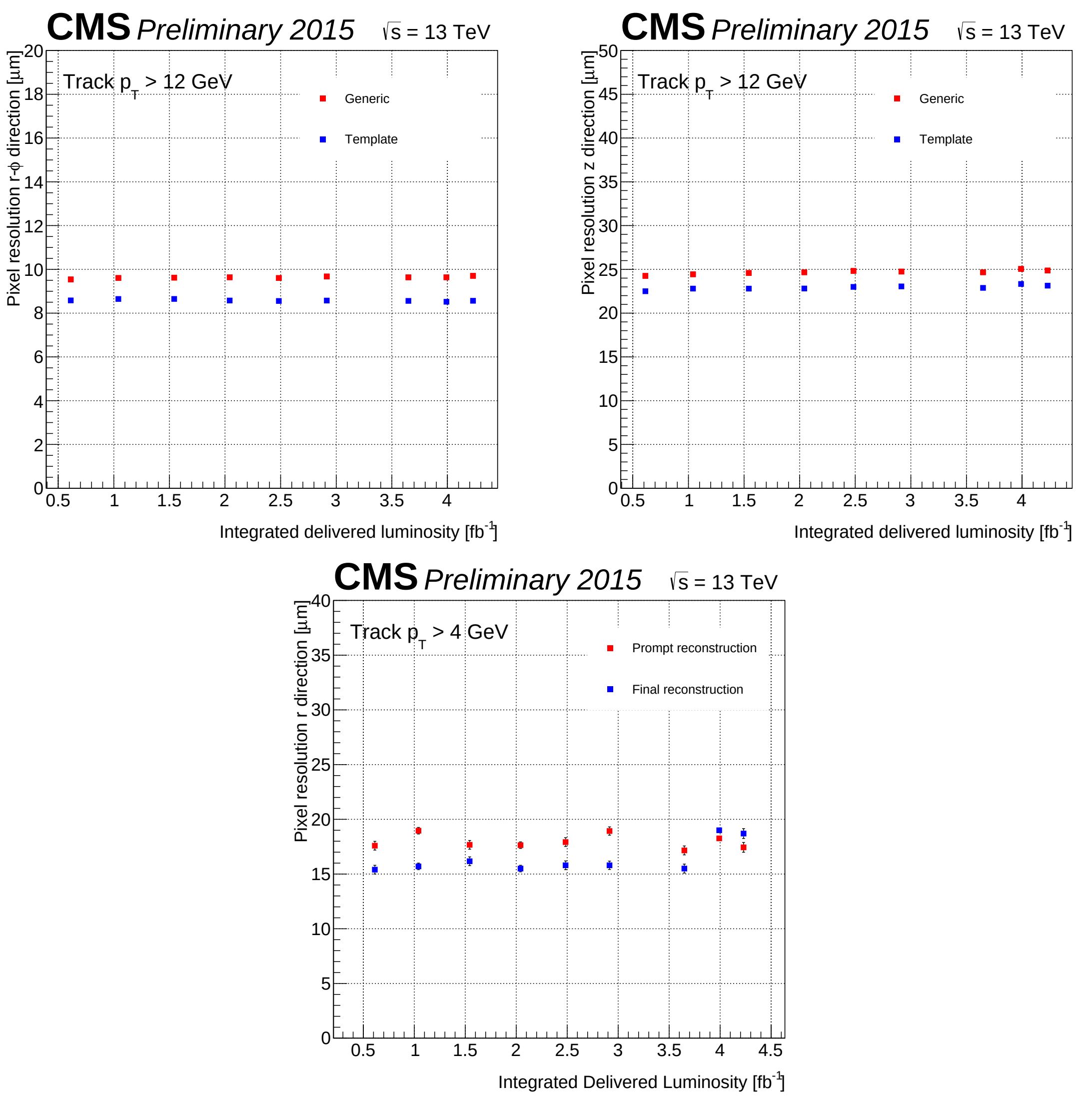


Figure 3.4: Evolution of the BPix resolution in r-ϕ direction (top left), BPix resolution in z (top right) direction, and the forward pixel resolution as a function of integrated luminosity in LHC Run II. Reprint from [1].

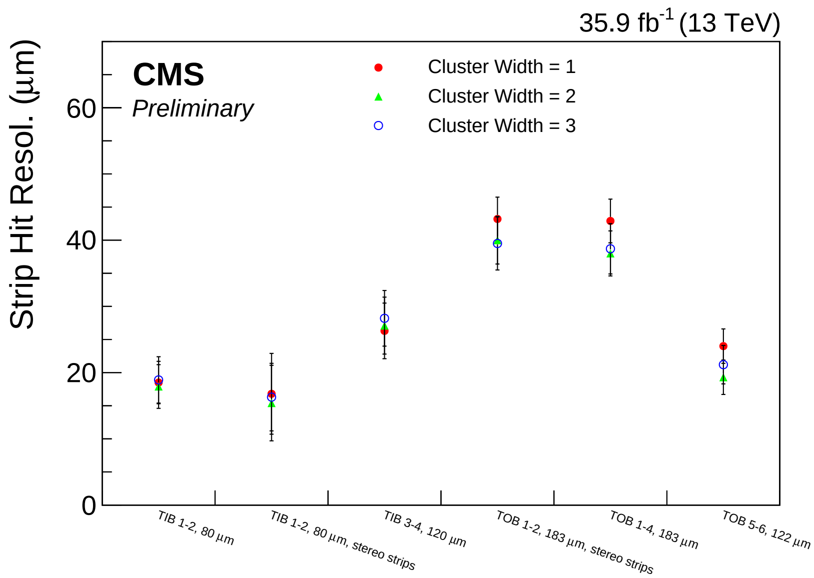


Figure 3.5: Strip Hit resolution for different cluster widths expressed in units of number of strips.

3.2.2 Electromagnetic Calorimeter

The CMS electromagnetic calorimeter (ECAL) is a hermetic, homogeneous detector consisting of 61,200 $PbWO_4$ crystals in the barrel (EB), and 7324 crystals in each of the two endcap sections (EE). The $PbWO_4$ crystals have a density of 8.28 g/cm^3, radiation length of 0.89 cm and Molière radius of 2.2 cm. These characteristics make them the appropriate material for a compact calorimeter. The scintillation light emitted by the crystals are blue-green color, with a maximum wavelength at 420-430 nm. Therefore blue laser can be used to monitor the transparency and response of the crystals.

The scintillation light is collected by avalanche photodiodes (APDs) in the barrel section and vacuum phototriods (VPT) in the endcaps. Each barrel crystal has one pair of APDs glued to its back face, and each endcap crystal has only one APT mounted. The photodetectors are depleted by a custom high voltage (HV) power supply which can precisely control the bias voltage and allow the gain of the photodetectors to be stable.

A schematic drawing of the layout of the ECAL is shown in Figure 3.6. The EB uses 230 mm long crystals, corresponding to a radiation length of 25.8 X_0, while the length of the crystals in the EE is 220 mm (24.7 X_0). The crystals in the EB are grouped into 36 supermodules (SM), 18 on each side of the interaction point, and provide a granularity of 360-fold in ϕ and (2×85)-fold in η, covering the pseudorapidity range up to $|\eta| < 1.479$. The EE is composed of two Dees, each consists of 3662 crystals, and extends the coverage to $|\eta|$ =3.0. In addition, a preshower (ES) detector made of lead absorbers and silicon strip sensors is placed in front of EE to improve the identification of closely spaced photons from π^0 decays.

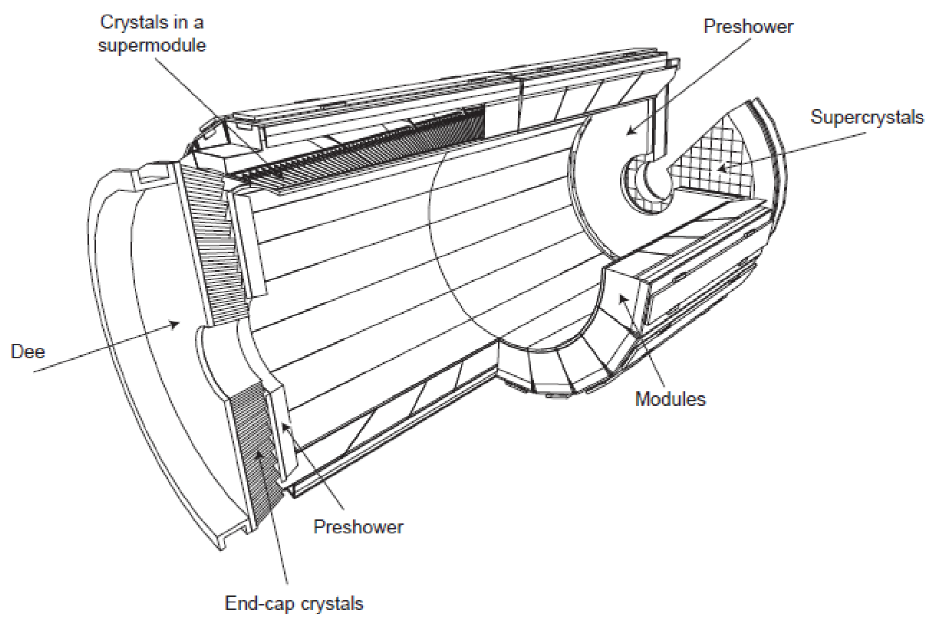


Figure 3.6: Layout of the CMS Electromagnetic Calorimeter (ECAL).

The signals from the photodetectors are shaped and amplified by the Multi Gain Pre-Amplifier (MGPA) with gains of 1, 6 and 12. If the signal is saturated, the read-out will switch to a lower gain. Each of the output signals of the MGPA is digitized by a 12-bit analog-to-digital converter (ADC) running at 40MHz and a set of 10 consecutive samples is recorded for amplitude reconstruction. The signals from a trigger tower (5×5 crystals) in the EB or a super-crystal in the EE are read out together by the on-detector electronics and are stored in pipelines during the Level-1 trigger latency. Once a Level-1 trigger is received, the data will be sent out to the off-detector electronics.

Radiation can cause a degradation in the crystal transparency due to the formation of color centers. The response of the VPT also varies under irradiation. In absence of irradiation, the transparency loss can partially recover through spontaneous annealing. These response changes are measured and corrected using a laser-based light monitoring system during the LHC operation. Blue and green laser light pulses are injected via optical fibers to the crystals and off-detector silicon PN photodiodes. The relative responses are then normalized to the measurements at the start of the run to derive correction coefficients. Figure 3.7 shows the evolution of the ECAL response for the Run I and Run II 2016 data taking periods.

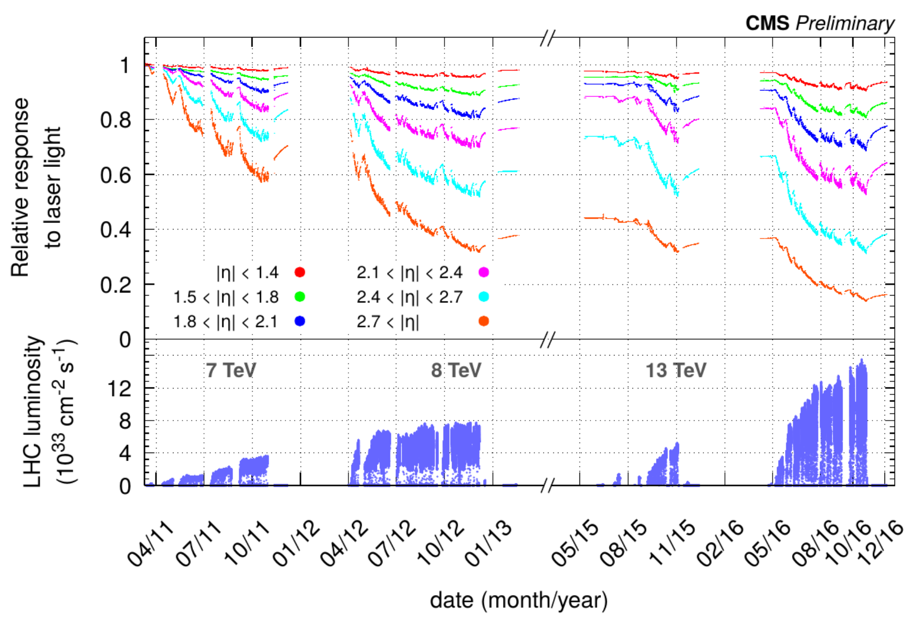


Figure 3.7: Evolution of the ECAL relative response as a function of data taking period.

The ECAL has been operating stably throughout the 2015 and 2016 LHC Run II operation. The analysis using 2.5 fb^{-1} collision data collected in 2015 shows that a relative energy resolution between 1.4% and 3% for electrons is achieved in the EB, and 3-4% for EE [42].

3.2.3 Hadron Calorimeter

The hadron calorimeter (HCAL) is designed to absorb and measure the energy of hadrons. In addition, it is important for the indirect measurement of weakly interacting particles such as neutrinos which escape the detector and result in imbalanced momentum. It contains four parts: a barrel (HB) with pseudorapidity range $|\eta| < 1.3$, two endcaps (HE) covering $1.3 < |\eta| < 3.0$, an outer calorimeter (HO) outside the solenoid to catch the shower tails, and two forward calorimeters (HF) positioned at either end of CMS. Figure 3.8 shows the location of the HCAL in CMS detector.

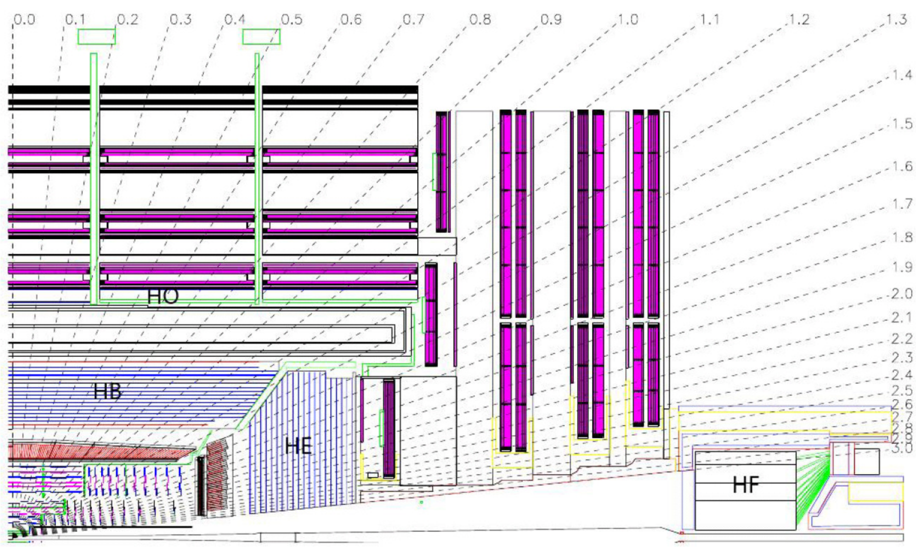


Figure 3.8: HCAL tower segmentation in the r-z plane.

The HB and HE are the major sub-detector of the HCAL. They are sampling calorimeter made of alternating layers of brass absorber plates and plastic scintillator tiles. The HB consists of 17 layers of absorber/scintillator, corresponding to a total interactions length of 5.39 at $|\eta| = 0$ and 10.3 at $|\eta| = 1.3$. Each HE has 18 scintillator and 17 absorber layers, where the absorber is designed to minimize the cracks between HB and HE. When particles deposit energy in the HCAL, the active materials are excited by ionizing radiation and generate scintillation light, which is read out by wavelength shifting fibers.

The optical signals are amplified and converted into electronic signals by multi-channel Hybrid Photodiodes (HPDs). The output analogue signals are then digitized by the charge-integrator and encoder (QIE) chips. After digitization, the signals are delivered to the trigger/read-out (HTR) board of the backend electronics located in the service cavern. In the HTR, trigger primitives are generated and sent to the Regional Calorimeter trigger. The HTR also buffers the readout data and transfers it to the data acquisition system when a Level 1 acceptance decision is received.

Because of the geometric constraints of the solenoid coil, the EB plus HB doesn't provide enough materials to fully contain the hadron showers. To ensure adequate sampling depth in the $|\eta| < 1.3$ region, the HO is placed outside the solenoid to detect any late showers or leakage from HB. The HO contains one layer of scintillator and utilizes the solenoid coil as absorber, extending the total interaction length of the calorimeter system to be at least 11.8 λ_I.

In the forward region, the HF calorimeter extends the coverage up to $|\eta| < 5.2$. The design of the HF is very challenging due to the extremely high radiation dose near the beampipe. Therefore the HF is made up of radiation hard quartz fibers embedded within a 165-cm-long steel absorber. Signals are generated when charged particles above the threshold (0.2 MeV for electrons, 400 MeV for protons) emit Cherenkov lights in the quartz. The lights are transported to photomultiplier tubes (PMT) for readout. The HF tower occupancy can be used for CMS luminosity measurements.

3.2.4 The Muon System

The detection of muon is an important task of CMS, as suggested by the name of the experiment. Many interesting physics processes, including the decay of Higgs boson to ZZ and potential production of new particles, may involve muons in the final states. Muons are less affected by radiative losses because they are 200 times heavier than electrons. Most of the muons can penetrate the calorimeter system with little interaction. Therefore the sub-system to detect the muons are placed in the outmost part of the CMS where other particles except neutrinos are expected to be fully absorbed.

The CMS muon system is designed to precisely measure the momentum and charge of muons over a wide kinematic range. The detectors sit in between the segments of iron return yoke of the CMS solenoid, where the magnetic field is below 2 T, as shown in the Figure 3.9. Because of the large coverage area outside the solenoid, the materials of the muon system have to be cost-effective and robust. Thus gas-ionization detectors are chosen to measure the muons. The muon system uses three different technologies: drift tubes (DT) in the barrel section, cathode strip chambers (CSC) in the endcaps, and resistive plate chambers (RPC) in both the barrel and endcaps to provide fast triggers.

Drift Tubes

The muon rate and radiation dose in the central region are relatively low, thus the drift tubes are used in the barrel to cover the pseudorapidity region up to $|\eta| < 1.2$. The DT is a closed-cell wire chamber which consists of a thin wire at high positive voltage (anode) within a gas-filled volume. When charged particles pass through the chamber, gas atoms will be ionized and the resulting electrons will drift towards the anode wire. The hit position of the muons can be reconstructed by measuring the drift time.

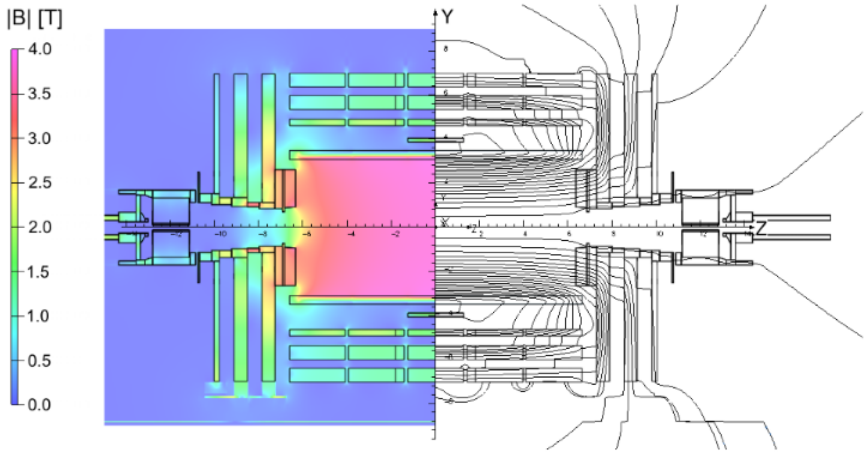


Figure 3.9: Map of the magnet field (left) and field lines (right) predicted for a longitudinal section of the CMS detector by a magnetic field model at a central magnetic flux density of 3.8 T. Each field line represents a magnetic flux increment of 6 Wb

The basic element of the DT is a rectangular drift cell with a transverse cross-section of $42 \times 13\ mm^2$. A gold-plated stainless-steel wire is suspended in the center of the chamber and operates at +3600 V. Four parallel layers of such drift tubes are staggered by half a cell to form a superlayer (SL). SLs with wires along the beam direction can provide measurements for r-ϕ coordinates, while orthogonal SLs can measure the r-z coordinates. The barrel muon detector consists of four concentric DT stations interleaved with iron yokes, as shown in Figure 3.10. Each of the inner three stations, labeled as MB1, MB2, MB3, has 60 chambers which consist of two parallel SLs and one perpendicular SL. The outer station has 70 chambers, however, each chamber has only an r-ϕ SL.

The single-hit spatial resolution of the DT tube is averagely 260 μm in the r-ϕ direction. When segments of the whole chamber are fitted together, the resolution is improved to be 90 μm.

Cathode Strip Chambers

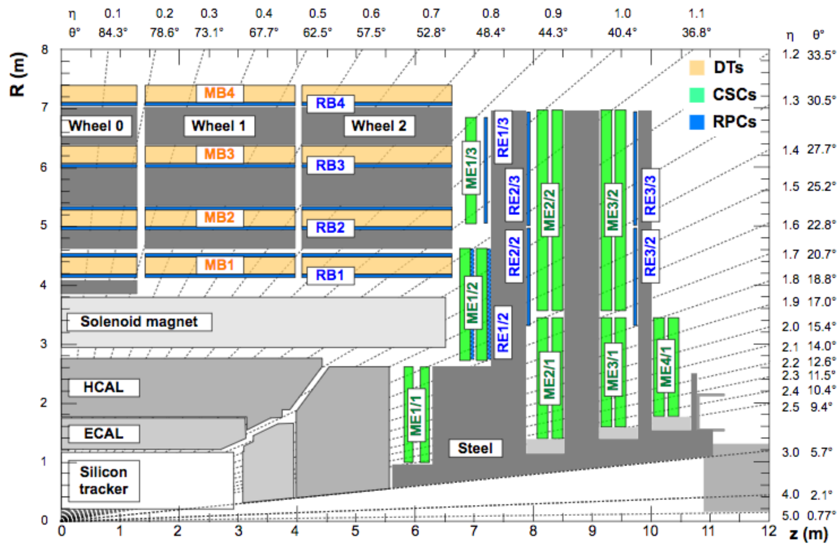


Figure 3.10: An R-z cross section of a quadrant of the CMS detector with the axis parallel to the beam (z) running horizontally and radius (R) increasing upward. The interaction point is at the lower left corner. Shown are the locations of the various muon stations and the steel disks (dark grey areas). The 4 drift tube (DT, in light orange) stations are labeled MB (muon barrel) and the cathode strip chambers (CSC, in green) are labeled ME (muon endcap). Resistive plate chambers (RPC, in blue) are in both the barrel and the endcaps of CMS, where they are labeled RB and RE, respectively.

In the endcap region, the muon rate is higher and the magnetic field is non-uniform. Thus the cathode strip chambers are chosen to measure the muons in the endcaps since they have a faster response time. The CSC is another type of multiwire chamber which consists of thin parallel anode wires crossed with negatively-charged cathode strips. All chambers are filled with a gas mixture of 50% CO_2, 40% Ar, and 10% CF_4. When charged particles pass through, they ionize the gas and produce free electrons, which move to the anode wires creating an avalanche pulse. Positive ions move to the cathode strips and induce a pulse so that a pair of position coordinates can be read out from one individual chamber. Six such chambers form a CSC module, allowing a precise measurement of muon tracks.

Each endcap has 4 CSC stations, containing a total of 468 chambers. The chambers are mounted on the surface of endcap disks in a way that the cathode strips run radially and the anode wires measure the r coordinate. The full CSCs provide a coverage of the pseudorapidity from 0.9 to 2.4. A reconstruction efficiency of 96-99% is achieved for muons except in the gap regions.

Resistive Plate Chambers

The gaseous detectors with wires can provide excellent spatial resolution, however, the time resolution is relatively poor because of the fluctuations in drift time. The time resolution can be improved using detectors with intense electric field so that the amplification of the signals can immediately start after primary ionization. To ensure unambiguous identification of the collision bunch at the highest LHC luminosities, a dedicated trigger detector made of resistive plate chambers (RPC) is implemented to complement the DTs and CSCs.

The RPCs are double-gap parallel-plate detectors, each gap consists of two plates made of high resistive plastic material. The outer surfaces of the plates are coated with conductive graphite paint, and a voltage of 9.6 kV is applied, allowing the RPCs to operate at avalanche mode. The use of high resistive electrode limits the discharge of the chamber to a small region so that the deadtime of the full detector is negligible and a good efficiency is ensured.

The RPCs are installed in both the barrel and endcap region, covering a pseudorapidity region $|\eta| < 2.1$. There are four RPC stations in the barrel, of which the innermost two stations consist of 2 layers of RPCs with DT sandwiched. In the endcap region, there were 3 RPC stations in LHC Run I. During the first long shutdown (LS1), an additional 144 RPCs are installed to the fourth disk. This improved the muon reconstruction efficiency in the range $1.2 < |\eta| < 1.8$.

The RPC system operated smoothly during the LHC Run II with an average efficiency of 94%. A time resolution of 3 ns or better is achieved for all 3 systems [43, 44].

3.3 CMS Trigger System

The LHC provides proton-proton collisions every 25 ns, leading to an event rate of 40 MHz. The raw data per event is about 1 MB, which means 40 TB data is produced every second. This large amount of data is far beyond the ability of the machine to store and process the events. Moreover, the production rate of the interesting physics processes are a few orders of magnitude lower than the total pp collisions rates. Therefore, a data filtering system is needed to drastically reduce the event rate while keeping the interesting events recorded. This task is performed by the CMS trigger system. The CMS trigger system reduces the event rate in two steps: the Level-1 trigger (L1) uses hardware electronics to reduce the data rate to 100 kHz, and the high level trigger (HLT) uses software to future reduce the rate to $\sim$ 1 kHz.

Level-1 Trigger

The L1 trigger receives coarsely segmented data from the sub-detectors and performs rough reconstruction of the particles using programmable electronics (FPGA, ASICs). It consists of local and global components. The scheme of the L1 system is shown in Figure 3.11. In the local trigger level, energy deposits in the calorimeters and hits in the muon stations are summed with reduced resolution to generate trigger primitives (TP). The TPs are fed into two trigger processing layers for object identification and global energy summation, and the resulting objects are sent to the global trigger (GT) for the final L1 decision.

The calorimeter trigger was upgraded during LHC LS1 to adopt a time-multiplexed

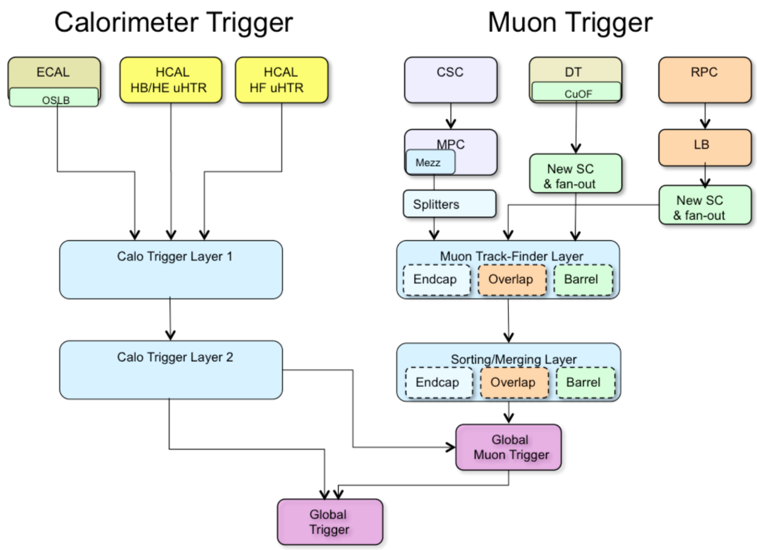


Figure 3.11: Dataflow for the L1 trigger. Reprint from [2].

system that enables the data from the entire calorimeter to be processed in a single FPGA at full granularity [45]. It consists of two layers, of which the first layer is designed to pre-process the TPs from ECAL and HCAL and format the data for layer 2. The second layer is the main processing layer that performs the object identification and energy reconstruction. e/γ candidates are reconstructed by dynamic clustering algorithms around local energy maxima above a fixed threshold. Trigger level energy calibration factors are applied to both the TPs and e/γ objects in order to make the summed energy to be similar to the offline reconstructed energy. Jet candidates are formed by taking 9×9 TTs sliding windows around the seed towers. Other global quantities, such as the transverse missing momentum, are computed using the full granularity.

The muon trigger system was also upgraded during LS1 to meet the increased luminosity of LHC Run II. Trigger level muon tracks are reconstructed in the track finder (TF) layer which is composed of three separate components: barrel muon track finder for the DT and RPC in $|\eta| < 0.83$, the endcap muon track finder for RPC and

CSC in $|\eta| > 1.24$, and the overlap muon track finder covering the three subdetectors in region $0.83 < |\eta| < 1.24$. Candidates from the TFs will be propagated to the global muon trigger (GMT) for sorting and isolation computing. The selected muons will be sent to the global trigger (GT) along with the information from the calorimeter trigger.

GT is the last stage of the L1 trigger. L1 decision is made using information received from the calorimeter and muon triggers. The GT has been upgraded with state-of-the-art FPGAs on Advanced Mezzanine Cards in a MicroTCA crate. The upgraded processors can compute different objects and higher level physics quantities, such as invariant mass, with high efficiency.

High Level Trigger

The HLT is a software trigger running on a computing farm which consists of more than 10,000 CPU cores. The basic unit of the HLT menu is called a trigger path, which consists of a series of reconstruction sequences and event filters. Event selections of each trigger path start from requiring the presence of one or more L1 objects, which are called L1 seeds. When the L1 seeds pass the kinematic selections of the trigger path, object reconstruction will be performed sequentially with increased complexity so that simple and fast algorithms can run first in the trigger path. A list of filters are implemented in the trigger path to decide if an object/event should be tentatively accepted or rejected. There are more than 400 trigger paths implemented in HLT menu in Run II to cover a wide range of physics signatures. An event is only recorded if it is accepted by at least one trigger path.

The HLT paths are grouped into several output datasets called primary data sets. These data sets include data streams for both physics analysis and detector operations. The RAW data of each event is about 1 MB, while the storage manager can temporarily store the data on disk with a 2GB/s rate. Therefore the HLT rate is limited to 1000 Hz. In Run II, the HLT reconstruction sequences are implemented with the Particle Flow algorithms, which improves the energy resolution and pile-up mitigation in the increased luminosity conditions.

3.4 The Data Quality Monitoring and Certification System

The Data Quality Monitoring (DQM) system [46] is a histogram-based software toolkit that evaluates and records the data quality of each sub-detector. The primary goal of the DQM system is to identify problems of each sub-detector and guarantee the quality of physics data. The DQM has an online instance as well as an offline instance. The online DQM receives a subset of the CMS data and performs a real-time analysis on these data in order to give immediate feedback about the detector and trigger status. The offline DQM, on the other hand, takes the full dataset as input and performs a more precise reconstruction and evaluation of the data than in the online instance. The offline DQM is used in the monitoring of the reconstruction quality, validation of simulated data, and test new CMS software releases.

Figure 3.12 shows the workflow of the DQM. The subset of raw data containing both the physics events and calibration data are delivered to the DQM farm for unpacking and reconstruction. Each sub-detector has a set of DQM code that checks the quality of the data relevant to it and record the analysis results in a ROOT file. The quality check is performed every lumi section, and the ROOT files are delivered to the graphical user interface (GUI) for view. GUI, the central component of the DQM, is a website for browsing data quality histograms. The content in GUI is organized in workspaces depending on the scope and ranging from high-level summaries to shift views to expert areas.

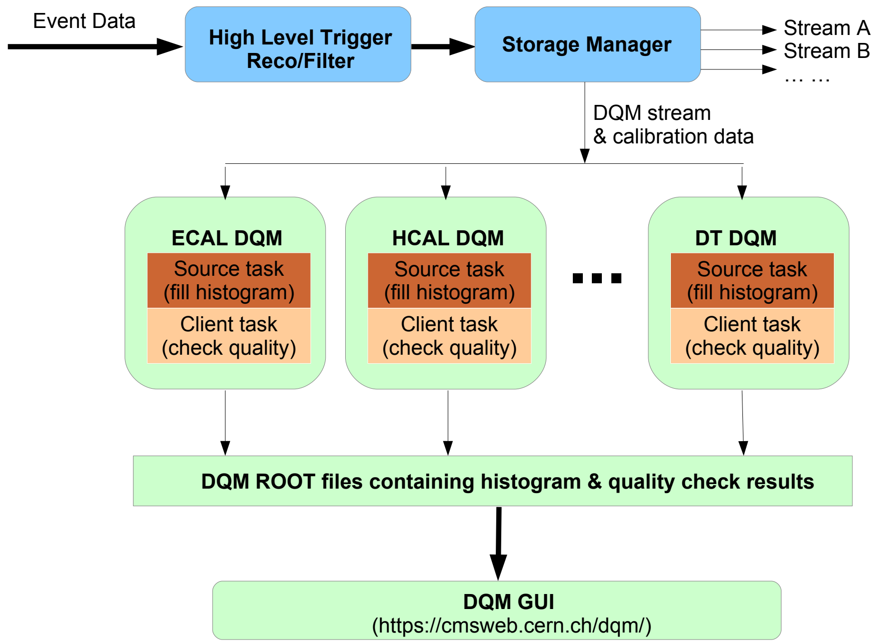


Figure 3.12: Workflow of the data quality monitoring system.

A snapshot of the ECAL DQM GUI is shown in Figure 3.13. It has access to both low level detector information, such as channel pedestals and noise, and high level data quality, like reconstructed hit timing. The values of pedestals and noise of the ECAL channels can be extracted from the DQM histogram and recorded into the CMS database, which are used for overall detector goodness check. Besides the detector quality check, the DQM is also useful in a variety of analysis. For instance, the timing histograms in the ECAL DQM can be used to align the timing of different ECAL components.

The DQM is in addition used in the certification of data good for physics analysis. This process starts with the physicists checking the data quality of each lumi section using the offline DQM. For each sub-detector a single boolean flag is used to describe the final quality result. Once the quality check is done, the physicists will fill the basic run information along with the quality results in the run registry system, which

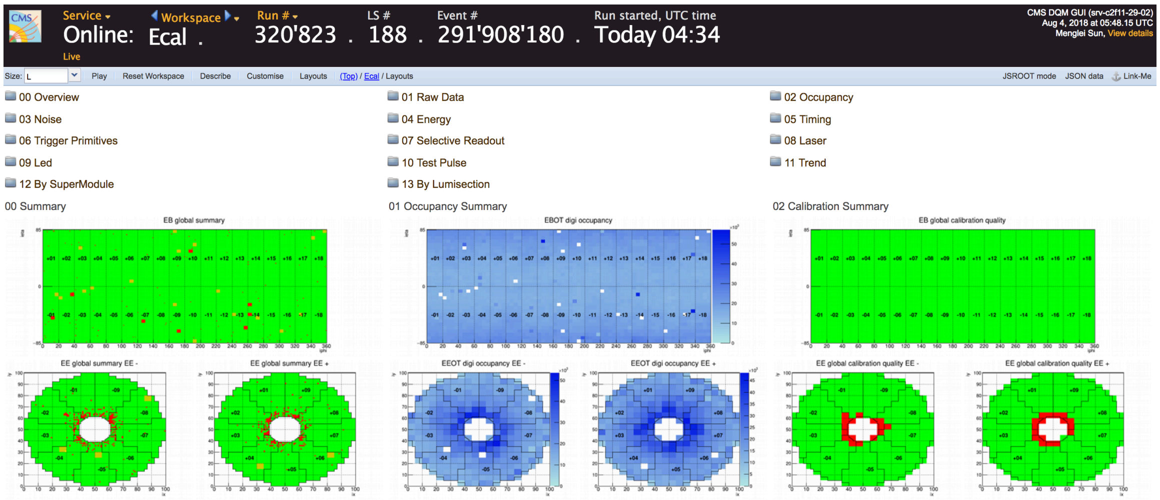


Figure 3.13: Snapshot of the ECAL DQM.

is a dataset that bookkeeps the certification result. If the data of all sub-detectors in a lumi section have good quality, this lumi section is certified to be good and its information is recorded in a JSON file which is used to select data for analysis.

Chapter 4

Event Reconstruction

Physics objects in CMS are reconstructed using the particle-flow (PF) algorithm [3], which identifies each particle through an optimized combination of signals of all subdetectors. The PF candidates are classified as photons, charged hadrons, neutral hadrons, electrons, or muons. Photons are reconstructed by clustering energy deposits in the ECAL. Tracks are first formed using hits in the inner tracker and then extrapolated to the calorimeter and muon chambers. If a track is geometrically close to ECAL and HCAL clusters within a certain distance, they are linked together to form a charged hadron. ECAL and HCAL clusters without associated tracks are identified as neutral hadrons. Muons are formed if the tracks can be associated with segments in muon stations. Electrons are reconstructed by associating tracks to ECAL clusters, and their energies are determined from a combination of the track momentums, calorimeter energy and a sum of all bremsstrahlung photons. Fig. 4.1 illustrates the response of different types of PF candidates in CMS. The algorithms for the reconstruction of different PF elements are detailed in the following sections.

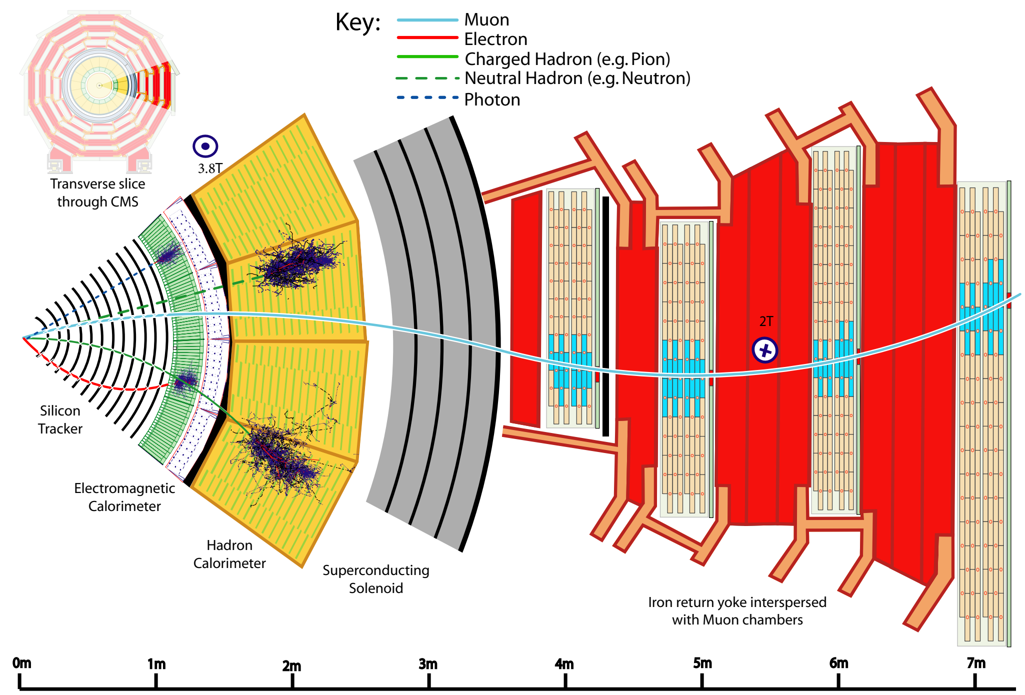


Figure 4.1: A sketch of the responses of different types of particles in a transverse slice of the CMS detector. The muon and the charged pion are positively charged, and the electron is negatively charged. Reprint from [3].

4.1 Tracks

The tracker is the subdetector that is closest to the collision point of the LHC. An accurate reconstruction of tracks is essential for the identification of charged particles and measurement of momentums. The vertexes of the pp collisions can be reconstructed by taking the intersection points of the tracks. A correct estimation of the positions of interaction vertexes is essential for both object identification and pileup mitigation. Therefore, the track-finding algorithms must be able to fully exploit the capabilities of the tracker and reconstruct the tracks with high efficiency and accuracy.

A Combinatorial Track Finder (CTF) based on extended Kalman filter (KF) is used for track reconstruction [47]. The CTF reconstruction sequence is performed in several iterations, with the initial interactions targeting at easy-identifiable tracks whereas later interactions dealing with more complex situations. Hits associated with the tracks in each iteration are removed from later iterations in order to reduce the complexity for reconstructing more difficult tracks. The CTF generates track seeds in the inner layers of the tracker and constructs the tracks outwards. Each CTF iteration proceeds in four steps:

- Seed generation. The track seeds are a set of hits whose patterns are compatible with charged-particle trajectories. In the first few iterations, triplets of pixel hits or mixed pixel pairs with additional vertex constraint are used as seeds. In later iterations, matched hits from the double-layer strip detector are also used to form the seeds.
- Track finding. The seed trajectories are extrapolated outwards and inwards to search for other tracker hits compatible with the expected path of a charged particle. The algorithm is based on Kalman filter method.
- Track fitting. A final fit to the track trajectories is performed to smooth the tracks and provide the best possible estimation of the track parameters. The χ^2 value is saved for each track.
- Track selection. Tracks are selected with a list of criteria, including requirements on the minimum number of layers with associated hits, requirements on the maximum number of layers containing no associated hits, upper bounds on χ^2, and thresholds on other track parameters. These selections help remove the fake tracks.

The CTF algorithms reconstruct tracks with high efficiency over the pseudorapidity range of the tracker. An average reconstruction efficiency of 94% in the barrel and 85% in higher $|\eta|$ regions is achieved for charged particles of $p_T > 0.9$ GeV in $t\bar{t}$ MC events.

The primary vertex, which is defined as the vertex with the highest $\sum p_T^2$ of associated tracks, is reconstructed by clustering the prompt tracks and performing fits to determine their common vertex. Tracks entering the primary vertex finding are required to satisfy cuts on the transverse impact parameter, the number of track hits and the normalized χ^2. The achieved vertex spatial resolution is 10-12 μm.

4.2 Photons

Photons are reconstructed from clusters of energy deposits in the ECAL [48]. The pulse from each crystal is digitized by a 12 bit ADC running at 40MHz and a set of 10 consecutive samples is recorded for amplitude reconstruction. In the high luminosity conditions, the pulse contains not only the signals from the current bunch collision but also the contributions from previous pileup bunches. To mitigate the out of time pileup, a multi-fit algorithm is used to reconstruct the signal amplitudes by estimating the in-time signal amplitude with up to 9 out-of-time amplitudes. Figure 4.2 shows examples of fitted pulses for simulated events in the EB and the EE with 20 average pileup interactions.

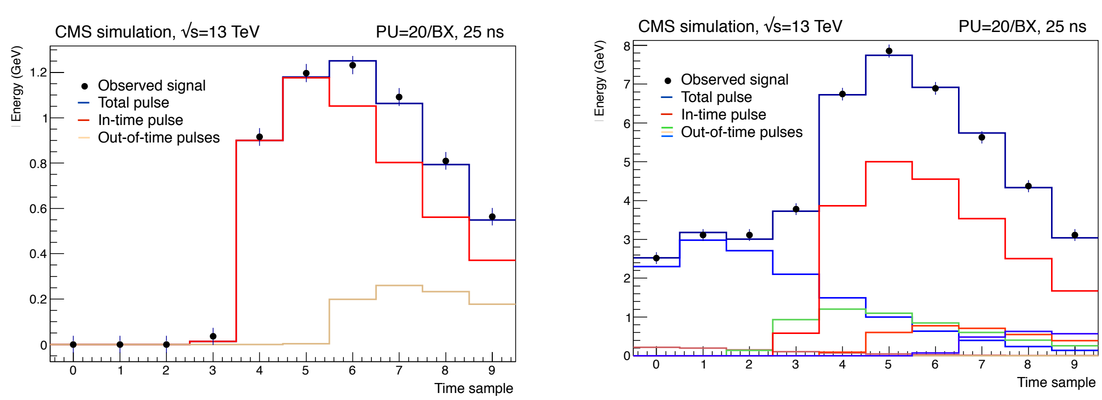


Figure 4.2: Examples of fitted pulses for signals in the EB (left) and in the EE (right) with 20 pileup interactions.

Energy of the photons usually spread over several crystals. Calorimeter hits deposited by the same electromagnetic shower are collected by the clustering algorithm. First, cluster seeds are formed by taking local maxima of energy deposits above a given threshold. Second, adjacent cells with an energy above the noise cut are added to the cluster seeds to form basic clusters. In the final step, basic clusters belonging to the same object are merged together to form a supercluster, which is extended in ϕ, to recover the radiated energy. The shower energy can be estimated as:

$$E_{e,\gamma} = F_{e,\gamma} \cdot [G \cdot \sum_i S_i(t) \cdot C_i \cdot A_i + E_{ES}], \tag{4.1}$$

where the sum runs over all clustered crystals. The quantity A_i is the pulse amplitude, which is converted to a GeV scale by multiplying the calibration factor G, and $S_i(t)$ is a correction term that accounts for the time variations of the channel response with C_i being the inter-calibration factor. For showers in the EE the energy measured by the preshower (E_{ES}) is added. Finally, the energy correction term $F_{e,\gamma}$ is applied to take into account geometry and upstream material effects, as well as the difference between electron and photon showers.

In-situ ECAL calibrations with physics events are applied to the EM clusters to improve the energy resolution. There are two steps of calibrations, inter-calibration and energy scale calibration, applied on top of the responses corrections of single ECAL channels. The purpose of inter-calibration is to equalize the variations in channel response due to different crystal light-yield and photo-detector gains. The inter-calibration factors are measured using multiple methods, including the use of the azimuthal symmetry of energy deposit in minimum bias events, the diphoton invariant mass of π^0 and η^0 decays, the E/p ratio of electrons from W and Z decays, and the mass peak of Z $\rightarrow e^+e^-$ decays. The combined coefficient is obtained by taking the mean value of individual corrections weighted by their respective precisions. The energy scale calibration, on the other hand, is applied to convert the digitized signal to GeV. The absolute ADC to GeV scale factor is determined separately in EB and EE using the Z $\rightarrow e^+e^-$ invariant mass peak, matching the reconstructed mass in data to that in simulation.

Any object that deposits a significant fraction of energy in the ECAL can be reconstructed as a photon. Therefore, electrons and electromagnetically rich jets are all registered in the photon collection. To distinguish the real photons from other objects, photon candidates are required to pass a list of shower shape quality cuts and not to have a matching track seed from the tracker. Details of photon identification criteria is described in Section 6.2.

4.3 Electrons

Electrons are reconstructed by associating tracks reconstructed in the silicon tracker to ECAL superclusters [49]. A dedicated track reconstruction algorithm is implemented for electrons because they have large radiative losses in the tracker than other charged particles. The electron track reconstruction also has the seeding, track finding and track fitting steps. In the seeding step, both ECAL-based and tracker-based seeding are considered. The track building starts with track reconstructions using the KF algorithm, with the track hits collected up to the ECAL. This method works well for electrons with negligible radiative loss. In the case of large bremsstrahlung, the KF reconstruction will fail and a dedicated Gaussian sum filter (GSF) is applied. The GSF models the energy loss in each tracker layer by a mixture of Gaussian distributions. The use of GSF improves the momentum resolution of electrons.

An electron is identified when the GSF track can be associated with an ECAL cluster. For the ECAL-seeded electrons, the SCs are propagated inwards to the inner tracker layers under the assumption that the energy-weighted positions of the clusters are on the trajectory of the electrons. The ECAL-predicted hits are then matched geometrically to the GSF tracks with the following requirements:

- $|\Delta\eta| < 0.02$, with $|\Delta\eta|$ being the distance in η between the SC and track η extrapolated to the position of closest approach to the SC
- $|\Delta\phi| < 0.15$, with $|\Delta\phi|$ being the distance in ϕ

For tracker-seeded electrons, the matching between the tracks and clusters are evaluated by the Multivariate Analysis (MVA) technique.

The overall reconstruction efficiency is $\sim 93\%$ for electrons from Z decay. For collision data, the efficiency is measured to be 88-98% in the barrel and 90-96% in the endcaps in the p_T range from 10 to 100 GeV.

4.4 Jets

Jet is a collimated spray of hadrons as a result of the hadronisation of quarks or gluons. Depending on the fraction of final particles produced by hadronization, a jet may contain hadrons, non-isolated electrons and collinear photons. Thus multiple sub-detectors can have energy deposits from a jet. A combination of these signals is performed with jet algorithms [50] to form a reconstructed jet.

The jets used in this analysis are reconstructed with the anti-k_T algorithms [51]. In this algorithm, each entity i (particles, pseudojets) is assigned a distance to the beam (B):

$$d_{iB} = k_{ti}^{2p}, \tag{4.2}$$

where k_{ti} is the transverse momentum of particle i. Here we use $p = -1$, and this is why this algorithm is named as "anti-k_T". The distance between two entities is defined as:

$$d_{ij} = min(k_{ti}^{2p}, k_{tj}^{2p})\frac{\Delta_{ij}^2}{R^2}, \tag{4.3}$$

where $\Delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$, and R is the radius of the reconstruction cone. In CMS Run II, a R = 0.4 cone is used for normal jets.

The anti-k_T is a sequential clustering algorithm which proceeds in an iterative way. In each iteration, it computes the d_{iB} and d_{ij} of all entities, and compares them to find the smallest quantity. If d_{ij} is the smallest one, then the ith and jth object are merged into one. Otherwise, if d_{iB} is the smallest one, then the ith object is called a jet and gets removed from the list. These steps are repeated until all particles are clustered into jets.

4.5 Muons

Muons leave hits in both the inner tracker and muon stations. The tracker can measure the momentum of muons with high precision, while the muon stations can detect the muons with higher purity because other particles are absorbed in the calorimeter system. Muons can be reconstructed in an inside-out, outside-in or standalone way in CMS [52]. Based on the reconstruction method, the muons can be classified into three types:

- Standalone Muon. Hits in the muon chambers are grouped to form standalone tracks, which provides an initial estimate of the muon tracks and can be used as seeds.
- Global Muon. A track reconstructed in the inner tracker is assigned to a standalone muon if they propagate to the same position onto a common surface.
- Tracker Muon. All reconstructed tracks are considered as muons and are extrapolated to the muon chambers. If the track can match at least one muon segment, it is recognized as a tracker muon.

These types are not mutually exclusive. The muon identification criteria usually require the muon to be two or more of these types.

4.6 Missing Transverse Momentum

Particles that do not interact with the detector materials, such as neutrinos, usually escape the detector, carrying part of the energy of the final states. In the initial state, the total momentum in the transverse plane is almost zero. The presence of invisible particles, therefore, results in an imbalance in the transverse plane, which is named as missing transverse momentum, denoted as p_T^{miss}.

The raw p_T^{miss} is calculated as the negative of the vector sum of the transverse momentum of all reconstructed objects in an event [53]:

$$p_{T,raw}^{miss} = -\sum_{i}^{N_{particles}} p_{T,i}.$$

p_T^{miss} is a high-level variable reconstructed using all visible particles in an event. It is sensitive to anomalous signals in the detector. Such anomalous signals include unphysical energy deposits in the calorimeter, noise of detector, beam-induced backgrounds, and poorly instrumented regions of the detector. Dedicated event filters are designed to identify and remove these anomalies.

Chapter 5

Datasets and Triggers

5.1 Datasets

This search is based on a data sample corresponding to an integrated luminosity of 35.9 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV produced by the LHC in 2016. The analysis is performed in both the $e + \gamma$ and $\mu + \gamma$ channels. Events in each channel are selected from the corresponding primary dataset with dedicated lepton plus photon triggers. Monte-Carlo (MC) simulated events are also used for various purpose.

5.1.1 Data

Reconstructed data sets of the CMS are further skimmed by event filters to keep only the information needed by physics analysis. These samples with reduced file size are called miniAOD [54]. Events of this analysis are selected from the miniAOD datasets produced with the CMS software framework, CMSSW_8_0_26_patch1. A total of 35.9 fb^{-1} of data was certified to be good for physics analysis. The "golden JSON" file Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt, which provides a list of run numbers and lumi section numbers of data with good quality, is used to select the certified datasets.

The candidate $e\gamma$ events are selected out of the DoubleEG primary dataset with the diphoton trigger: HLT_Diphoton30_18_R9Id_OR_IsoCaloId_AND_HE_R9Id_Mass90. The $\mu\gamma$ events are selected from the MuonEG dataset, triggered by either the HLT_Mu17_Photon30_CaloIdL_L1Iso or HLT_Mu38NoFilterNoVtx_Photon38_CaloIdL. In addition, SingleMuon and SingleElectron datasets are used in this analysis for various backgrounds estimation and efficiency measurements. The datasets used in this analysis are summarized in Table 5.1.

5.1.2 Simulation

MC simulated events are used in this analysis to study the search strategy, model the SUSY signal yields, estimate the background contributions, and validate the analysis methods. All the MC simulation samples are listed in Table 5.2, along with the cross-sections calculated at the same perturbative order as the generator.

The SM processes are generated with MadGraph_aMC@NLO 2.3.3 [55] in the CMS Moriond17 campaign, except for the W+γ samples which are generated with MadGraph in leading order (LO). To correct the W+γ cross-section, a constant NNLO k-factor of 1.34 is applied [56, 57]. All samples use the NNPDF 3.0 [58] parton distribution functions (PDFs). The hard scattering processes are interfaced to Pythia 8.2 [59] with the CUETP8M1 generator tune [60] for simulation of parton showers, hadronization and decays. Detector simulations are performed with a Geant4 [61] based package which has a detailed description of the CMS detector. Simulated events are digitized and reconstructed by the same algorithms that are used for real collision data.

The pile-up contribution is also considered in MC samples by superimposing min-

Table 5.1: Datasets used in this analysis.

Primary Dataset	Samples
DoubleEG	/DoubleEG/Run2016B-03Feb2017_ver2-v2/MINIAOD /DoubleEG/Run2016C-03Feb2017-v1/MINIAOD /DoubleEG/Run2016D-03Feb2017-v1/MINIAOD /DoubleEG/Run2016E-03Feb2017-v1/MINIAOD /DoubleEG/Run2016F-03Feb2017-v1/MINIAOD /DoubleEG/Run2016G-03Feb2017-v1/MINIAOD /DoubleEG/Run2016H-03Feb2017_ver*-v1/MINIAOD
MuonEG	/MuonEG/Run2016B-03Feb2017_ver2-v2/MINIAOD /MuonEG/Run2016C-03Feb2017-v1/MINIAOD /MuonEG/Run2016D-03Feb2017-v1/MINIAOD /MuonEG/Run2016E-03Feb2017-v1/MINIAOD /MuonEG/Run2016F-03Feb2017-v1/MINIAOD /MuonEG/Run2016G-03Feb2017-v1/MINIAOD /MuonEG/Run2016H-03Feb2017_ver*-v1/MINIAOD
SingleElectron	/SingleElectron/Run2016B-03Feb2017_ver2-v2/MINIAOD /SingleElectron/Run2016C-03Feb2017-v1/MINIAOD /SingleElectron/Run2016D-03Feb2017-v1/MINIAOD /SingleElectron/Run2016E-03Feb2017-v1/MINIAOD /SingleElectron/Run2016F-03Feb2017-v1/MINIAOD /SingleElectron/Run2016G-03Feb2017-v1/MINIAOD /SingleElectron/Run2016H-03Feb2017_ver*-v1/MINIAOD
SingleMuon	/SingleMuon/Run2016B-03Feb2017_ver2-v2/MINIAOD /SingleMuon/Run2016C-03Feb2017-v1/MINIAOD /SingleMuon/Run2016D-03Feb2017-v1/MINIAOD /SingleMuon/Run2016E-03Feb2017-v1/MINIAOD /SingleMuon/Run2016F-03Feb2017-v1/MINIAOD /SingleMuon/Run2016G-03Feb2017-v1/MINIAOD /SingleMuon/Run2016H-03Feb2017_ver*-v1/MINIAOD

Table 5.2: List of the MC samples used for signal and SM background processes, with their cross-sections and corresponding equivalent integrated luminosity. [13TeV] stands for TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8, and [S16] stands for RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6.

Process	Samples	σ(pb)	$\mathcal{L}_{eff}$fb^{-1})
SUSY	/SMS-T5Wg_[13TeV]/RunIISpring16MiniAODv2-PUSpring16Fast_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	-	-
	/SMS-TChiWG_[13TeV]/RunIISpring16MiniAODv2-PUSpring16Fast_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	-	-
$W\gamma$	/WGToLNuG_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/[S16]-v1/MINIAODSIM	337	18.11
	/WGJets_MonoPhoton_PtG-40to130_TuneCUETP8M1_13TeV-madgraph/[S16]-*/MINIAODSIM	12.5	409
	/WGJets_MonoPhoton_PtG-130_TuneCUETP8M1_13TeV-madgraph/[S16]-*/MINIAODSIM	0.64	3.6e+03
$Z\gamma$	/ZGTo2LG_[13TeV]/[S16]_ext1-v1/MINIAODSIM	117.864	121.9
$t\bar{t}\gamma$	/TTGJets_[13TeV]-madspin-pythia8/[S16]_ext1-v1/MINIAODSIM	3.697	1.32e+03
$t\bar{t}$	/TTJets_[13TeV]-madgraphMLM-pythia8/[S16]-v1/MINIAODSIM	502.2	20.19
WW	/WW_TuneCUETP8M1_13TeV-pythia8/[S16]_ext1-v1/MINIAODSIM	63.21	110.53
WZ	/WZ_TuneCUETP8M1_13TeV-pythia8/[S16]_ext1-v1/MINIAODSIM	22.82	131.28
$WW\gamma$	/WWG_[13TeV]/[S16]_ext1-v1/MINIAODSIM	0.2147	4.65e+03
$WZ\gamma$	/WZG_[13TeV]/[S16]-v1/MINIAODSIM	0.04123	2.42e+04
DY	/DYJetsToLL_M-50_[13TeV]/[S16]_ext2-v1/MINIAODSIM	5670	21.5

imum bias events on the simulated events. To mitigate the possible mismodeling of pile-up profile in simulations, the MC events are reweighted such that the number of vertices has the same distribution as measured in data, as shown in Figure 5.1.

5.2 Trigger

A diphoton trigger, HLT_Diphoton30_18_R9Id_OR_IsoCaloId_AND_HE_R9Id_Mass90_v*, is used to collect the signal candidate events for the $e\gamma$ channel. Each HLT diphoton path is seeded from the logic OR of a list of L1 EG triggers. For the $\mu\gamma$ channel, a muon-photon trigger, HLT_Mu17_Photon30_CaloIdL_L1Iso_v*, is used for events selection. HLT_Mu38NoFilterNoVtx_Photon38_CaloIdL is used in addition with logic OR to compensate for possible trigger inefficiencies caused by the isolation requirement in L1 EG seed. A summary of the signal HLT paths with their L1 seeds is shown in Table 5.3. The measurement of trigger efficiencies is described in subsequent sections.

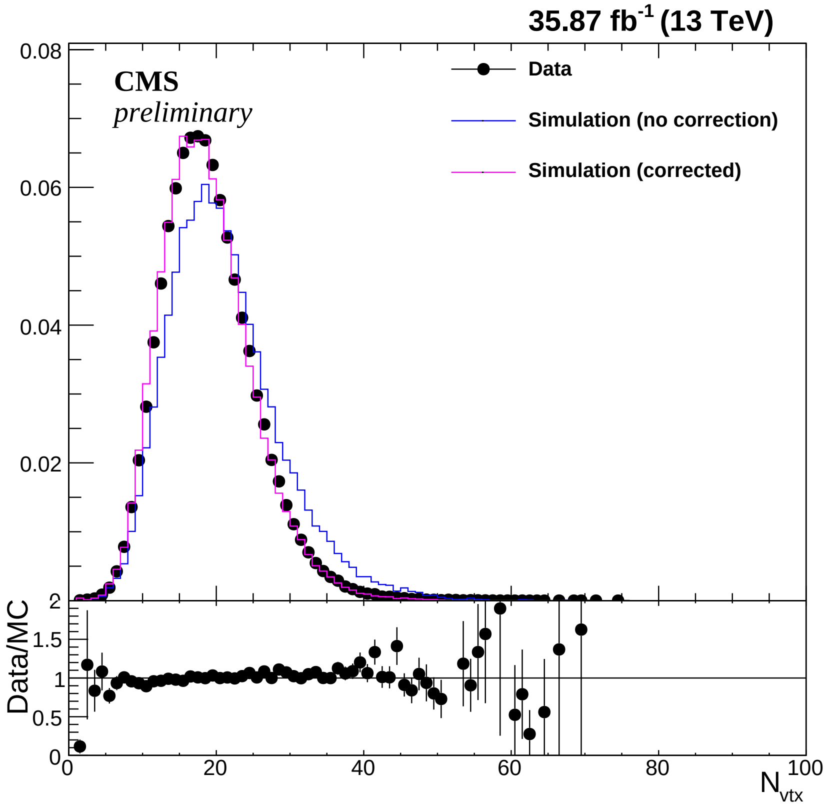


Figure 5.1: The distribution of the number of reconstructed vertices for data (black dots) and simulation (solid lines). The simulated events before (blue) and after (magenta) the pile-up reweighting are shown. The lower panel shows the ratio between the reweighted simulation and data.

Table 5.3: List of the HLT used for signal selections.

Channel	HLT	L1 seed
$e\gamma$	HLT_Diphoton30_18_R9Id_OR_IsoCaloId_AND_HE_R9Id_Mass90_v*	L1_SingleEG* OR L1_SingleIsoEG* OR L1_DoubleEG_*
$\mu\gamma$	HLT_Mu17_Photon30_CaloIdL_L1Iso_v*	L1_Mu5IsoEG18 OR L1_Mu5IsoEG20
	HLT_Mu38NoFilterNoVtx_Photon38_CaloIdL	L1_Mu5EG20 OR L1_Mu20EG15

5.2.1 $e\gamma$ Trigger Performance

The diphoton HLT path doesn't veto events if one or more photons have associated tracks. Therefore this path allows the $e + \gamma$ events to be triggered. The diphoton HLT is seeded from a combination of various L1 EG triggers to achieve an optimal L1 efficiency. The leading leg of the HLT is then reconstructed around the L1 seed and is required to pass the following quality filters:

- $E_T > 30$ GeV, $|\eta| < 2.5$
- $R_{9(5\times5)} > 0.5(0.8)$ in EB(EE), where $R_{9(5\times5)}$ is the energy sum of 3×3 crystals centered on the cluster seed divided by that of 5×5 crystals.
- H/E < 0.12 (0.1) in EB(EE)
- Either $\sigma_{i\eta i\eta} < 0.015$ (0.035) in EB(EE), ECAL isolation $< 6.0 + 0.012E_T$ OR $R_{9(5\times5)} > 0.85$ (0.9)

If the leading leg passes these kinematic and shower quality filters, the entire ECAL data will be reconstructed. In the unseeded step, at least two clusters are required to have $E_T > 18$ GeV and pass the unseeded filters, which have the same criteria with the seeded leg and in addition a cut on the track isolation. The final trigger requires the diphoton mass to be greater than 90 GeV. The total efficiency is given by:

$$\epsilon_{diphoton} = \epsilon_{sub,sub} \cdot \epsilon_{lead,sub} \cdot \epsilon_{lead,lead},$$

where $\epsilon_{lead,lead}$ is the efficiency of matching the leading photon to the seeded leg object, $\epsilon_{lead,sub}$ is the efficiency for the leading photon to pass the unseeded filters, and $\epsilon_{sub,sub}$ is the efficiency of matching the sub-leading photon to an unseeded object.

Because photons and electrons have similar ECAL cluster shapes, the diphoton trigger efficiency can be measured with a tag-and-probe method using the electron pairs from Z decays. Events are selected from the SingleElectron dataset using HLT_Ele27_eta2p1_WPLoose_Gsf, where the tag electron is required to match the single electron trigger leg and the probe object remains unbiased. The tag is a tightly selected electron with $p_T > 30$ GeV. The probe is a photon passing the full offline selections except that the probe is allowed to have an associated pixel seed. The tag and probe pair must have an invariant mass between 80 GeV and 100 GeV. The total number of selected probe objects is used as the denominator for the efficiency calculation, whereas the number of probes which can match leading leg objects of the diphoton HLT within $\Delta R < 0.3$ is used as the numerator.

The tag and probe pairs are used a second time for the sub-leading leg efficiency measurement. Since the sub-leading leg only gets reconstructed when a leading leg exists, the tag is required to match both the single electron trigger leg and the diphoton leading leg. The probe in this measurement is defined as a medium working point electron that passes all the offline selections described in section 6.3. Out of such denominator electrons, those that have the sub-leading leg matched are counted as numerators. The trigger efficiency of probe photons and electrons are shown in Figure 5.2.

In simulation, the identity of an object can be checked by geometrically matching it to a generated particle. This method is called MC truth-matching. Thus the trigger efficiency of simulated sample is derived by selecting the photon and electron with truth-matching and counting the number of objects which match the trigger legs. The efficiency measured in MC is then compared with that in data, and their ratio, $R = \epsilon^{data}/\epsilon^{MC}$ is used as the data-to-simulation trigger efficiency scale factor (ESF). Figure 5.3 shows the ESF of the photon and electron as a function of p_T and η. Systematic uncertainties of the ESF are calculated from the tag-and-probe fits to the Z-peak.

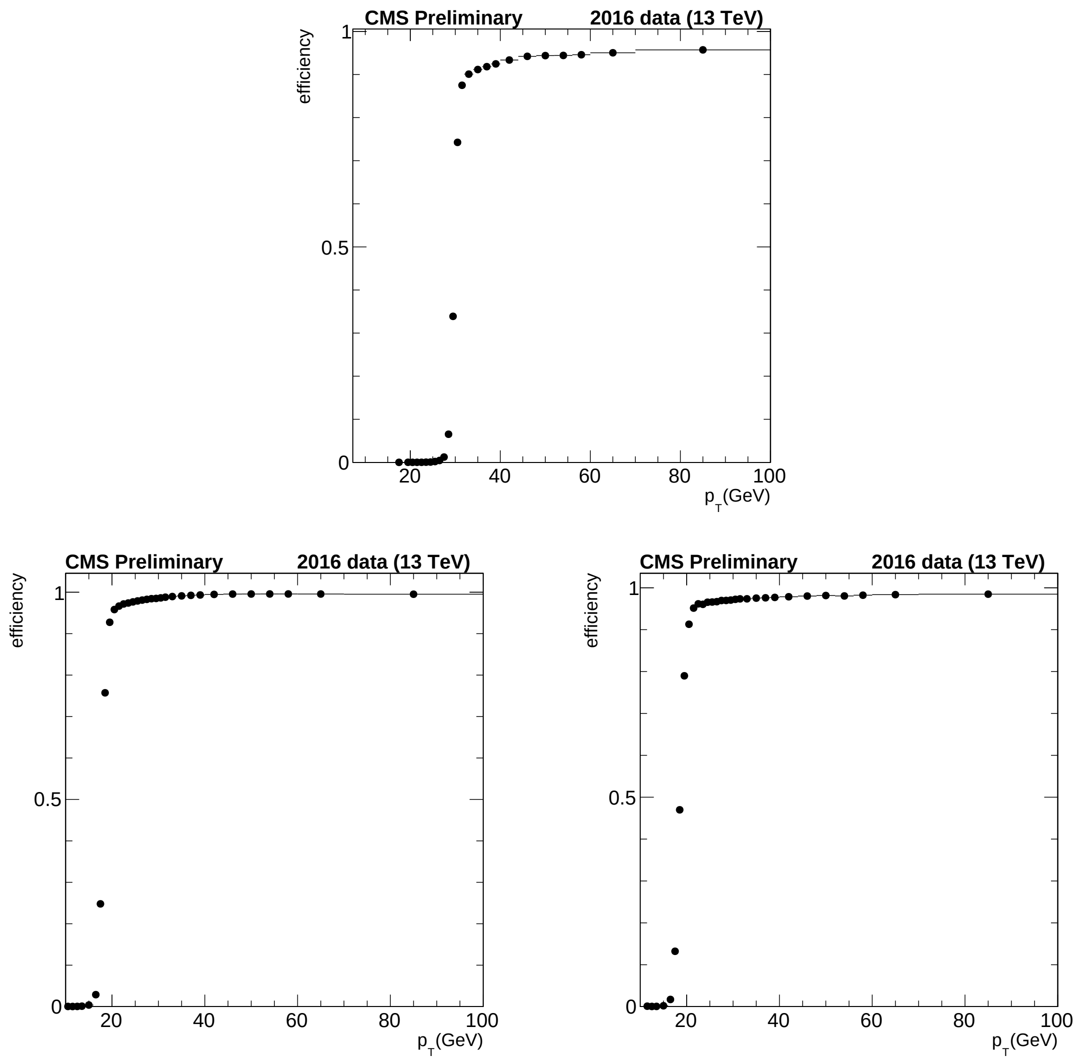


Figure 5.2: Trigger efficiencies for $e\gamma$ channel. Top: leading leg efficiency, bottom left: sub-leading leg efficiency (EB), bottom right: sub-leading leg efficiency (EE).

5.2.2 $\mu\gamma$ Trigger Performance

The HLT used in $\mu\gamma$ channel is a combination of two muon-photon cross triggers. The main HLT is HLT_Mu17_Photon30_CaloIdL_L1Iso, which is based on L1_Mu5_IsoEG20/L1_Mu5_IsoEG18 L1 seeds. Because of the presence of isolation requirements on L1 EG object, the L1 efficiency has a slower turn-on curve than non-isolated EG triggers. Thus a second trigger, HLT_Mu38NoFilterNoVtx_Photon38_

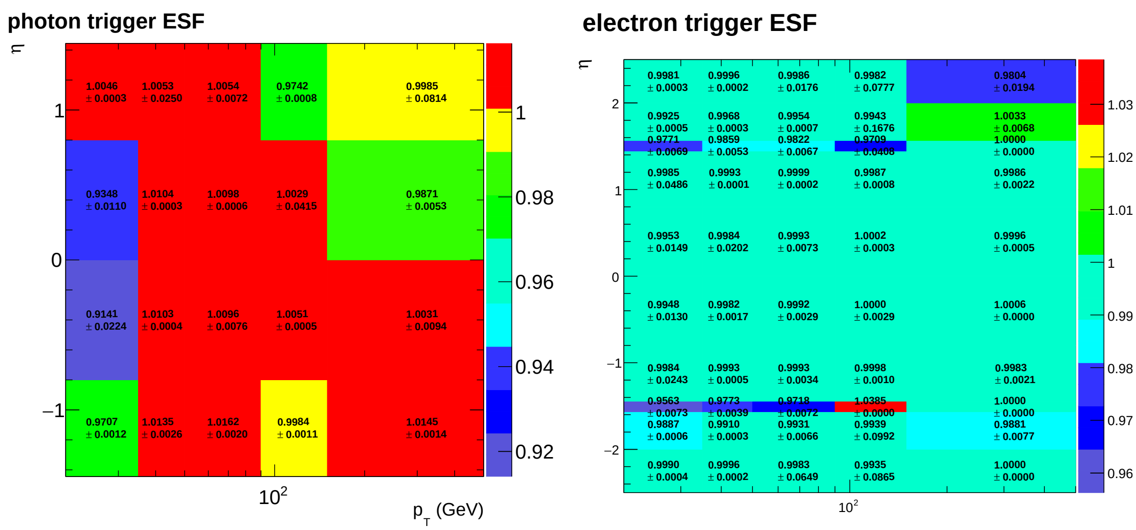


Figure 5.3: Trigger efficiency scale factors for the photon (left) and electron (right).

CaloIdL which has no L1 isolation requirements, is used in logic OR to improve the efficiency.

The $\mu\gamma$ channel trigger efficiency is measured using $Z \to \mu\mu\gamma$ events selected from SingleMuon dataset. HLT_MuIso24 and HLT_MuTrkIso24 are used in a logic OR to select events. Again, a tag-and-probe method is applied to get pure muon and photon objects. The tag is defined as a medium working-point muon which fires one or both of the single muon triggers. A second muon and a photon are selected as probes. The probe muon and photon must pass the corresponding offline selections and the $\mu\mu\gamma$ three-body mass must be between 60 GeV and 120 GeV. To ensure that the photon comes from FSR, the muon pair invariant mass is required to be less than 80 GeV. Out of the muon-photon probe pairs, those ones that have both objects matching the trigger objects are selected as the numerator. Figure 5.4 shows the muon-photon trigger efficiency as a function of photon p_T and muon p_T.

The efficiency measurements based on FSR events suffer from low statistics at high p_T. To cross check the results obtained from $Z \to \mu\mu\gamma$, we separately measure

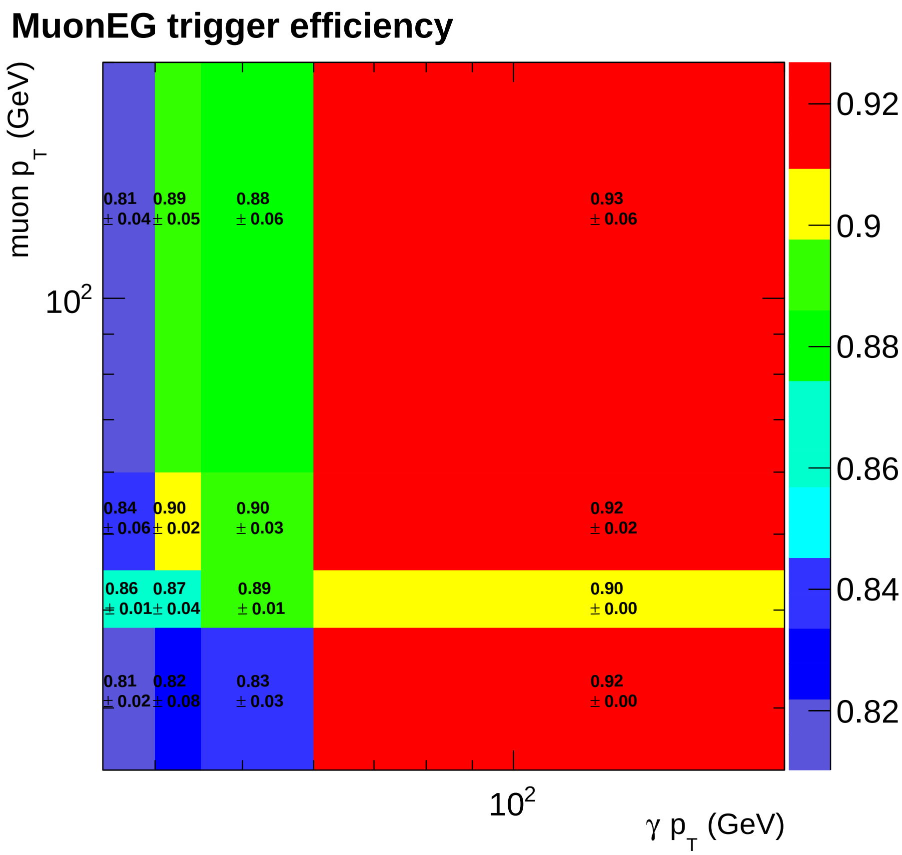


Figure 5.4: $\mu\gamma$ trigger efficiency as a function of photon and muon p_T

the muon and photon trigger efficiencies in SinglePhoton and SingleMuon dataset. The discrepancies between these two methods are used as the systematic uncertainties of the $\mu\gamma$ trigger ESF. Figure 5.5 shows the muon-photon ESF, along with the uncertainties.

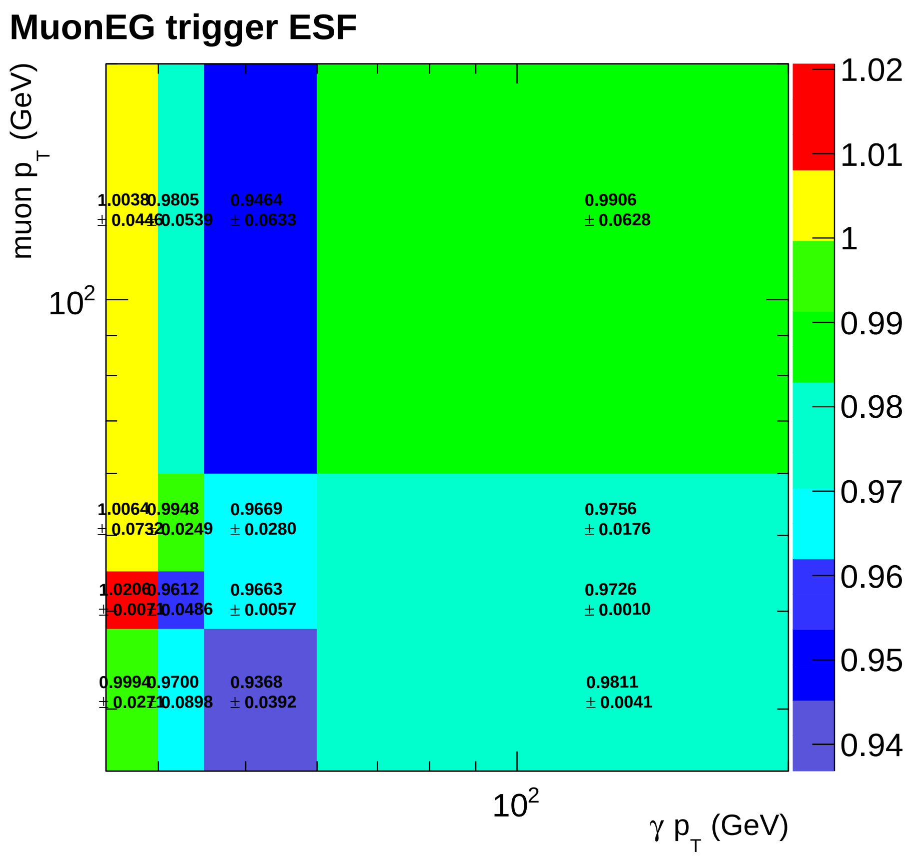


Figure 5.5: $\mu\gamma$ simulation-to-data trigger efficiency scale factor as a function of photon and muon p_T

Chapter 6

Object Definitions and Event Selection

6.1 Definition of Particle Identification Variables

A description of variables used for particle identification (ID) and selection is given here:

Calorimeter shower shape variables:

- R_9: a sum of energy of the 3×3 crystals centred on the most energetic crystal in the supercluster divided by that of 5×5 crystals.

- $\sigma_{i\eta i\eta}$: lateral width of the electromagnetic shower in η direction.
  It is calculated as: $\sigma_{i\eta i\eta} = \sqrt{\sum_i^{5\times5}[(\eta_i^{\text{cryst.in cluster}} \times 0.0175 + \eta^{\text{seed cryst.}} - \bar{\eta}_{5\times5})^2 \times w_i / \sum_i^{5\times5} w_i]}$,
  where $w_i = 4.2 + ln\frac{E_i}{E_{5\times5}}$.

Isolation variables:

- H/E: the ratio of energy deposited in the HCAL tower behind the ECAL supercluster and energy of the supercluster.

- PF charged isolation $\text{Iso}_{h^\pm}$: transverse energy sum of all PF charged hadrons falling inside a cone of size ΔR= 0.3 around the candidate particle. Contributions from pileup are estimated by multiplying the median density of pileup contamination (ρ) per event with the effective area of the cone. The pileup contributions are then subtracted from the isolation. This is the so-called "ρ correction" method.

- PF neutral isolation Iso_{h^0}: transverse energy sum of all PF neutral hadrons falling inside a cone of size ΔR= 0.3 around the candidate particle. ρ correction is applied to mitigate the pileup effect.

- PF photon isolation Iso_{pho}: transverse energy sum of all PF photons reconstructed inside a cone of size ΔR= 0.3 around the candidate particle. The footprint of the candidate itself has been removed to avoid double counting. ρ correction is applied to mitigate the pileup effect.

Electron ID variables:

- $|\Delta\eta_{in}| = |\eta_{SC} - \eta_{in}^{\text{extrap}}|$: the distance between SC energy-weighted η (η_{SC}) and the track η extrapolated from the tracker to the ECAL ($\eta_{in}^{\text{extrap}}$).

- $|\Delta\phi_{in}| = |\phi_{SC} - \phi_{in}^{\text{extrap}}|$: the distance between SC and track extrapolated position in ϕ

- $|\frac{1}{E} - \frac{1}{p}|$: the difference between the SC energy (E) and track momentum (p).

Muon ID variables:

- $\chi^2/d.o.f.$: χ^2 of the track normalized to the degree of freedom in fitting

- χ^2 of the kick finder: the χ^2 between the inward and outward states of a track in each layer is calculated to look for kicks, and the largest χ^2 per track is used as the χ^2 for kick finder.
- $|\mathrm{d_{xy}}|$: transverse impact parameter relative to the beam-spot position
- $|\mathrm{d_z}|$: longitudinal impact parameter relative to the beam-spot position

6.2 Photon

The default format of the photons in miniAOD datasets is "slimmedPhotons", which keeps high level physics objects. e/γ energy scales, which are described in 4.2, are applied to the photons to optimize the energy reconstruction and resolution.

Photons with $p_T > 35$ GeV reconstructed in the ECAL barrel ($|\eta| < 1.4442$) region are used. They are required to match the trigger objects of the HLT within $\Delta R < 0.3$. To ensure that photon candidates pass the R_9 filter in the $e\gamma$ trigger, the photons are in addition required to have $R_9 > 0.5$. Out of such preselected photons, those fulfilling a set of ID criteria are selected. The selection criteria usually require the ID variables of a particle to either above some thresholds or below certain upper bounds, which are called "cuts". Photons in this search are selected using the loose working point ID, which includes the following cuts:

- $\mathrm{H/E} < 0.0597$
- $\sigma_{i\eta i\eta} < 0.01031$
- $\mathrm{Iso}_h^{\pm} < 1.295$
- $\mathrm{Iso}_h^0 < 10.910 + 0.0148 \cdot p_T + 0.000017 \cdot p_T^2$
- $\mathrm{Iso}_{pho} < 3.630 + 0.0047 \cdot p_T$

To suppress the $e \to \gamma$ fake objects, candidate photons are required not to be associated with a pixel track seed. In addition, a photon is rejected if an electron is close to it within $\Delta R < 0.02$. Such a veto helps remove the very rare cases where the ECAL clusters fail to match pixel seeds but electrons still get reconstructed by ECAL-driven tracking algorithm.

The identification and pixel veto efficiencies are measured for both data and simulation by the CMS EGM group and are given in [62]. The data-to-simulation efficiency scale factors (ESF) are applied on MC samples to correct the simulation response. Fig. 6.1 shows the 2D map of photon scale factors along with the uncertainty of each $\eta - p_T$ bin.

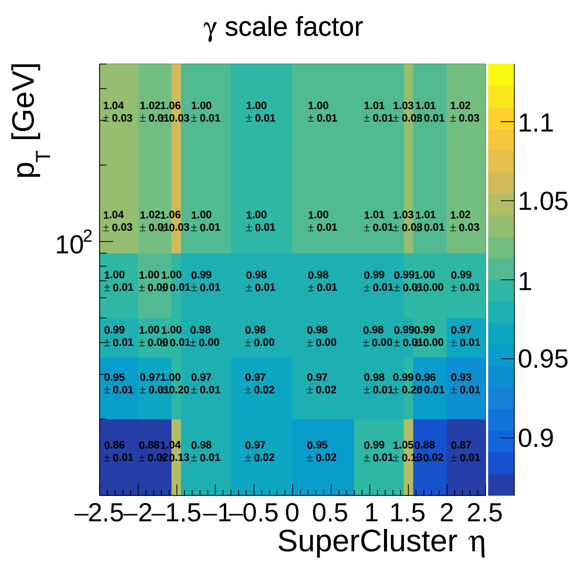


Figure 6.1: data-to-simulation scale factors for photons.

6.3 Electron

Electrons with $p_T > 25$ GeV reconstructed in the pesudorapidity range $|\eta| < 2.5$ are used. To ensure a good acceptance efficiency, objects falling in the barrel-endcap region $1.442 < |\eta| < 1.56$ are rejected.

Electrons passing the kinematic cuts are required to match the sub-leading leg of the di-photon trigger. To mimic the R_9 filters in trigger, a $R_9 > 0.5(0.8)$ preselection cut is applied on the electrons in EB(EE). Candidate electrons are then identified using cut-based medium working point ID, including the following selections:

- $\sigma_{i\eta i\eta} < 0.00998$(EB), 0.0298(EE)
- $\Delta\eta_{in} < 0.00311$(EB), 0.00609(EE)
- $\Delta\phi_{in} < 0.103$(EB), 0.045(EE)
- H/E < 0.253(EB), 0.0878(EE)
- $|\frac{1}{E} - \frac{1}{p}| < 0.0129$
- At most one expected missing hit in inner tracker layers
- Pass conversion veto

The conversion veto is applied to reject the electrons that arise from conversions of photons. When photon conversions happen inside the tracker, the tracks of the resulting electrons are more likely to begin later than prompt electrons, and missing hits are therefore present in the first few layers of the inner tracker. In addition, conversions are identified by fitting track pairs to a common vertex and searching for nearby partner tracks. The conversion algorithm combines all these methods to veto the photon conversions.

The relative isolation cut is removed from the medium ID selections. Instead, we use the mini-Isolation [63], which is defined as the ratio of the energy sum in a cone to the p_T of the electron. The cone size is p_T-dependent: R = 10 GeV/min(max($p_T(e)$, 50 GeV), 200 GeV), resulting in a radius of 0.05 for low-p_T electrons and 0.2 for high-p_T ones. The use of mini-Isolation ensures a good acceptance of lepton even in

Lorentz-boosted topologies. In this analysis, we required the mini-Isolation of the electron to be smaller than 0.1.

The efficiencies of the electron identification and mini-Isolation filters are measured by EGM group and are given in [62]. Fig. 6.2 shows the data-to-simulation ESF of the electron ID and mini-Isolation.

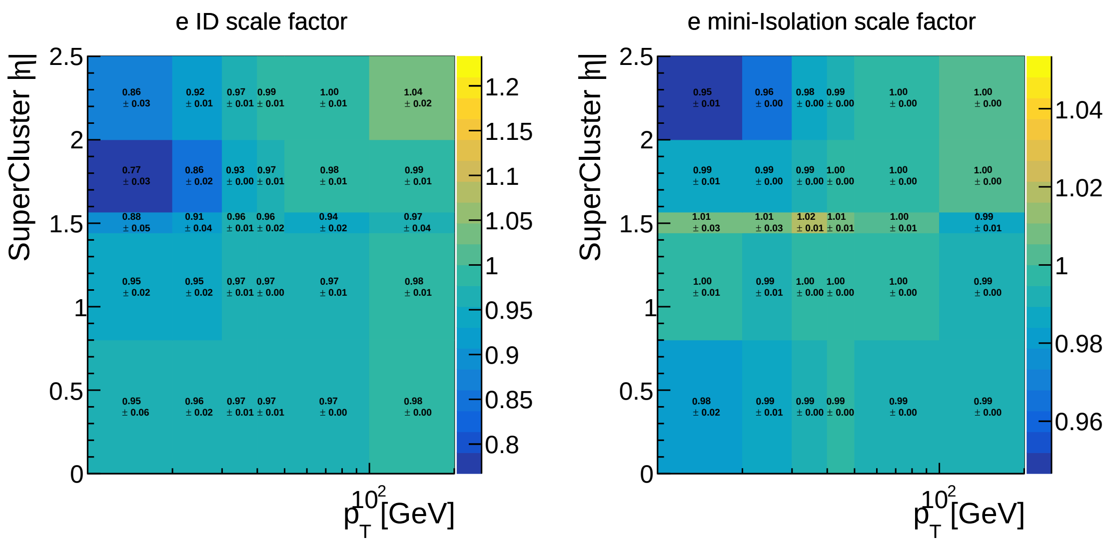


Figure 6.2: data-to-simulation scale factors for electrons.

6.4 Muon

Muons with $p_T > 25$ GeV in the pesudorapidity range $|\eta| < 2.4$ are used. Matching to the trigger leg is required for a muon to enter the candidate collection. The muon is identified using the standard medium ID, defined as:

- Is PF muon
- Is also reconstructed as a global-muon or as a tracker-muon
- Fraction of valid tracker hits > 0.8

- Either satisfies the good global muon criteria:
  - is global muon
  - normalized χ^2 of the global track is less than 3
  - χ^2 of the position of the standalone muon and associated tracker tracks is less than 12
  - χ^2 of the kick finder is less than 20
  - segment compatibility between the tracker tracks and muon chambers is greater than 0.303

  or has a segment compatibility greater than 0.451

The Muon is in addition required to pass the impact parameter and isolation cuts:

- $|d_{xy}| < 0.05$ cm, $|d_z| < 0.1$ cm.
- mini-Isolation < 0.2.

Both the identification efficiency and isolation efficiency are measured by CMS SUSY group and are given in [62]. Fig. 6.3 shows the ESF used for muon data-to-simulation corrections.

If more than one candidate object are identified in an event, the one with the highest p_T is used.

6.5 Jet and HT

ak4PFJets jets, which are reconstructed using the anti-k_T algorithm with a radii of 0.4, are used in this search. Jets with $p_T > 30$ GeV and $|\eta| < 2.5$ are selected. The raw energies of the jets are first corrected to take account the contributions from

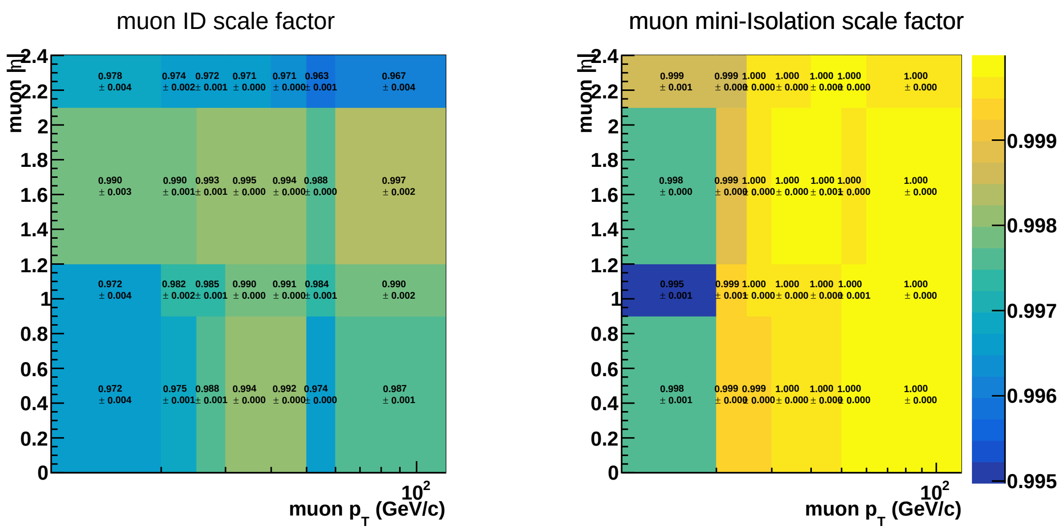


Figure 6.3: data-to-simulation scale factors for muons.

pileup. Then, jet energy corrections [64] derived from the simulation are applied to scale the reconstructed energy. Finally, the jet energy is further corrected using di-jet events.

To avoid double counting, jets that overlap with photon or lepton candidate with $\Delta R < 0.4$ are not considered. The cone size of 0.4 is chosen because the ak4 algorithm uses R = 0.4 to cluster the jets.

H_T is a variable that quantifies the hadronic activity of an event. It is defined as the scalar sum of p_T of all selected jets, as:

$$H_T = \sum_{i}^{N_{jets}} |p_{T,i}|.$$

Events from electroweak production usually have small H_T, whereas those from hadronic production tend to have larger H_T.

6.6 Missing Transverse Momentum

The type-1 corrections are applied to the raw p_T^{miss}. Such corrections propagate the jet energy corrections to p_T^{miss}.

A set of p_T^{miss} Filters recommended by the JetMET group are applied to clean events with large fake p_T^{miss}, such as detector noise, cosmic rays and beam halos. Events are rejected if they fail the following filters:

- primary vertex filter:
  At least one good vertex is required to be present in each event. A vertex is considered to be good if it has at least 5 degrees of freedom and its distance from the interaction point is less than 24 cm in z direction and 2cm in the x-y plane.
- beam halo filter
- HBHE noise filter
- HBHEiso noise filter
- ECAL TP filter
- Bad PF Muon Filter
- Bad Charged hadron filter
- EE badSC noise filter
- badMuons flag
- duplicateMuons flag

6.7 Photon FSR

Photons emitted in vector boson (W or Z) decays or radiated off the leptons can be energetic and isolated. Such events are called final state radiation (FSR) and can mimic the SUSY signals if large p_T^{miss} is also present. Simulation shows that the separation between the photon and the lepton is typically smaller in final states events (FSR) than in SUSY signal events. Therefore, three additional cuts are designed to suppress the FSR contributions:

- The candidate photon must be well separated from the candidate electron by $\Delta R > 0.8$
- No leptons are reconstructed close to the candidate photon within $\Delta R < 0.3$
- In the $e\gamma$ channel, $M_{e\gamma} > M_Z + 10\text{GeV}$

The last cut is applied on top of the 90 GeV invariant mass filter of the $e\gamma$ HLT to further reject $Z \to ee$ events, where one of the electrons is misidentified as a photon.

Chapter 7

Background Estimation

The SM backgrounds of this search can be classified into three categories. The main background is the production of W or Z bosons in association with a photon, denoted as Vγ. In particular, neutrinos from leptonic decays of W bosons escape the detector and result in genuine p_T^{miss}.

The second category contains misidentified objects, including fake photons from misreconstructed electrons or jets, and misidentified leptons from non-prompt productions. Such backgrounds are estimated by scaling control samples with misidentification rate that is estimated in the $p_T^{\text{miss}} < 70$ GeV control region.

The rare, but non-negligible backgrounds, are associated productions of multiboson or top quarks plus a photon. These backgrounds are estimated using simulation.

7.1 The Misidentification of Electrons as Photons

The misidentification of electrons as photons may occur if the matching between ECAL clusters and pixel track seeds fails during reconstruction. Such events can arise from Drell-Yan di-electron productions, where the photon candidates are in fact misreconstructed electrons. In addition, $t\bar{t}$ and WW events with ee or $e\mu$ in the final states also enter the signal selection if one electron fakes the photon. This background can be estimated by constructing an electron enriched control sample and scaling it by a transfer factor R, where R is the rate of fake photons arising from electron objects. The control sample is formed by inverting the pixel veto and electron veto requirements of the signal photon candidates while keeping other selections unchanged. Having a pixel seed matched to the photon, we in fact replace the photon with a well reconstructed electron.

The misidentification rate R is measured using a tag-and-probe method in the $p_T^{\text{miss}} < 70$ GeV control region. We select di-electron from Z decays in a similar way with the trigger efficiency measurement. Events triggered by HLT_Ele27_eta2p1_WPLoose_Gsf in the SingleElectron dataset are used. The tag object is a well reconstructed electron with $p_T > 30$ GeV and matches within $\Delta R < 0.3$ to the single electron trigger leg. Then a photon passing all the offline selections excluding the pixel seed veto and electron veto is selected as the probe. To measure the misidentification rate in a wider range, the p_T threshold of the probe is lowered to 30 GeV. The invariant mass of the tag and probe pair is required to be consistent with the Z boson mass. Since the pixel seed is not checked for the probe, it is also possible for the probe to be registered as a tag. To avoid any bias in the selection of probes, all combinations of the tag-and-probe pairs are counted. If the probe can additionally pass the pixel seed veto and electron veto, this event will enter the numerator sample, denoted as $e\gamma$. Those fail either pixel veto or electron veto will enter the denominator sample, which is labeled as ee in this thesis. The transfer factor is defined as $R = N^{e\gamma}/N^{ee}$, which is the misidentification rate of electrons as photons.

To determine the number of $Z \to ee$ events in the numerator and denominator samples, a fit to the tag-and-probe invariant mass in the range [60 GeV, 120 GeV] is performed using the RooFit [65] package. The fit model consists of two probability density functions: one describes the signal ($Z \to ee$) shape and the other models the background. We choose the Breit-Wigner function to describe the shape of the signal. To account for the radiation loss and resolution effect, the signal function is convoluted with a double-side crystal ball density function. The background shape is obtained from a μ+probe template selected using the SingleMuon dataset, under the assumption of lepton flavor symmetry. To form the template, we keep the probe definition unchanged and replace the tag electron with a medium working point muon that fired the HLT_IsoMu24 trigger. The distribution of the invariant mass of the μ and the probe is then smoothed using a Gaussian kernel estimator. Figure 7.1 shows the fitting results of the denominator samples in different p_T bins, while Figure 7.2 shows that of the numerator samples.

Systematic uncertainties of the fit majorly come from mismodeling of the signal and background shapes. To evaluate this effect, the invariant mass distributions are fitted again using alternative shapes. We first perform the fit with the signal model replaced by a $Z \to ee$ template selecting from the Drell-Yan simulation and convoluting with Gaussian distributions. We then change the background shape to an exponential function and fix the signal model to be the nominal Breit-Wigner. The choices of fitting functions are summarized in Table 7.1. The fake rates evaluated using altered shapes are compared with the nominal values, and their difference is used as the systematic uncertainties.

Table 7.1: Summary of the functions used to fit the tag-and-probe invariant mass. Nominal shapes in the table are used to determine the central values of the number of events in the numerator and denominator samples. Other functions are used for the measurement of systematic uncertainties.

Signal Model	Background Model	fake rate ($p_T^\gamma > 30$ GeV)
Nominal Shapes		
Breit-Wigner convoluting with crystal ball	μ + probe template	2.29%
Alter Signal Shapes		
Drell-Yan simulation convoluting with Gaussian	μ + probe template	2.26%
Alter Background Shapes		
Breit-Wigner convoluting with crystal ball	exponential function	2.43%

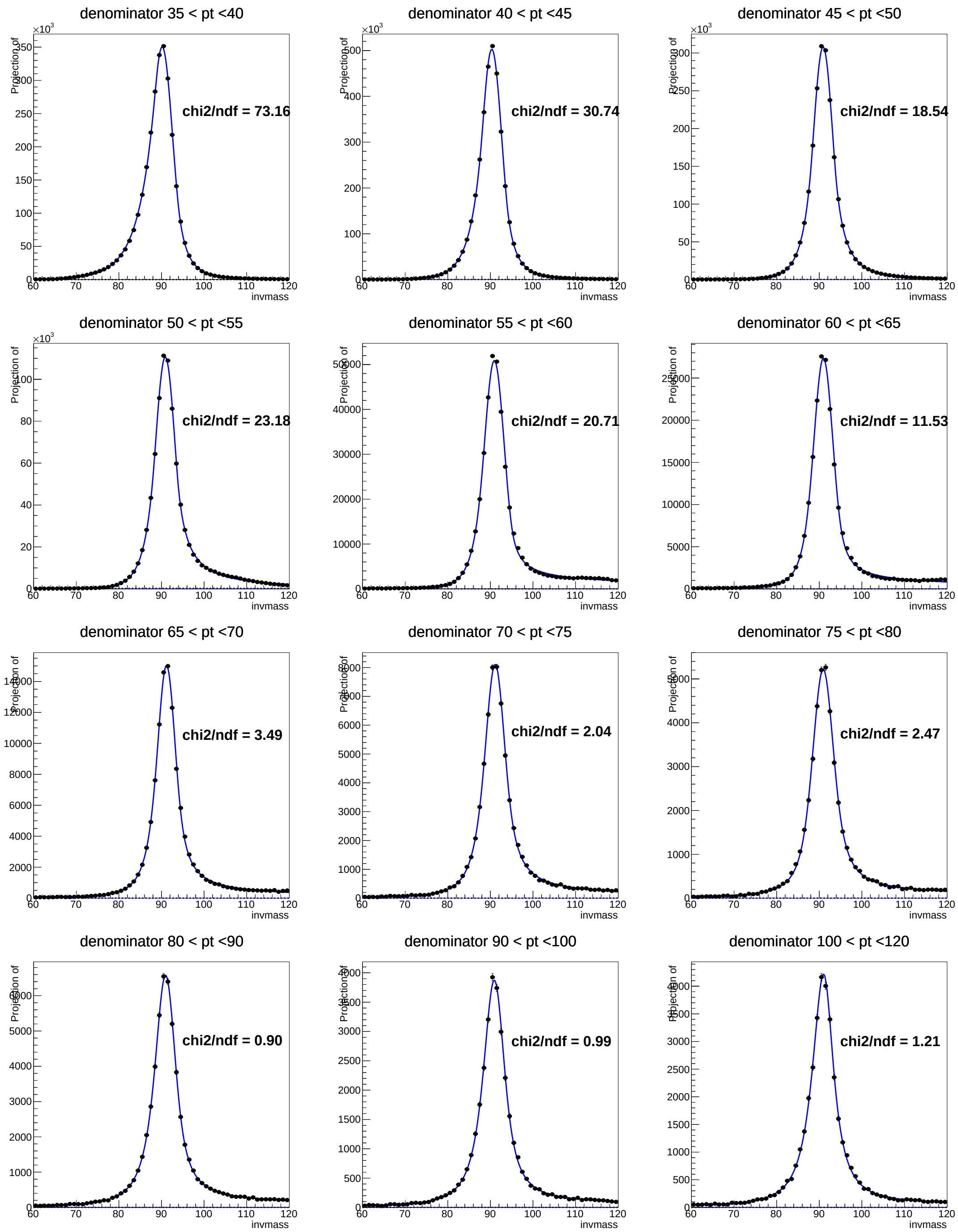


Figure 7.1: Fits to the tag-and-probe invariant mass of the denominator samples, which are defined as the samples with probe electrons matching to pixel track seeds. This fit is performed in various p_T bins. The black dots, the blue solid lines, and the blue dashed lines represent the data, the signal+background model, and the background-only model, respectively.

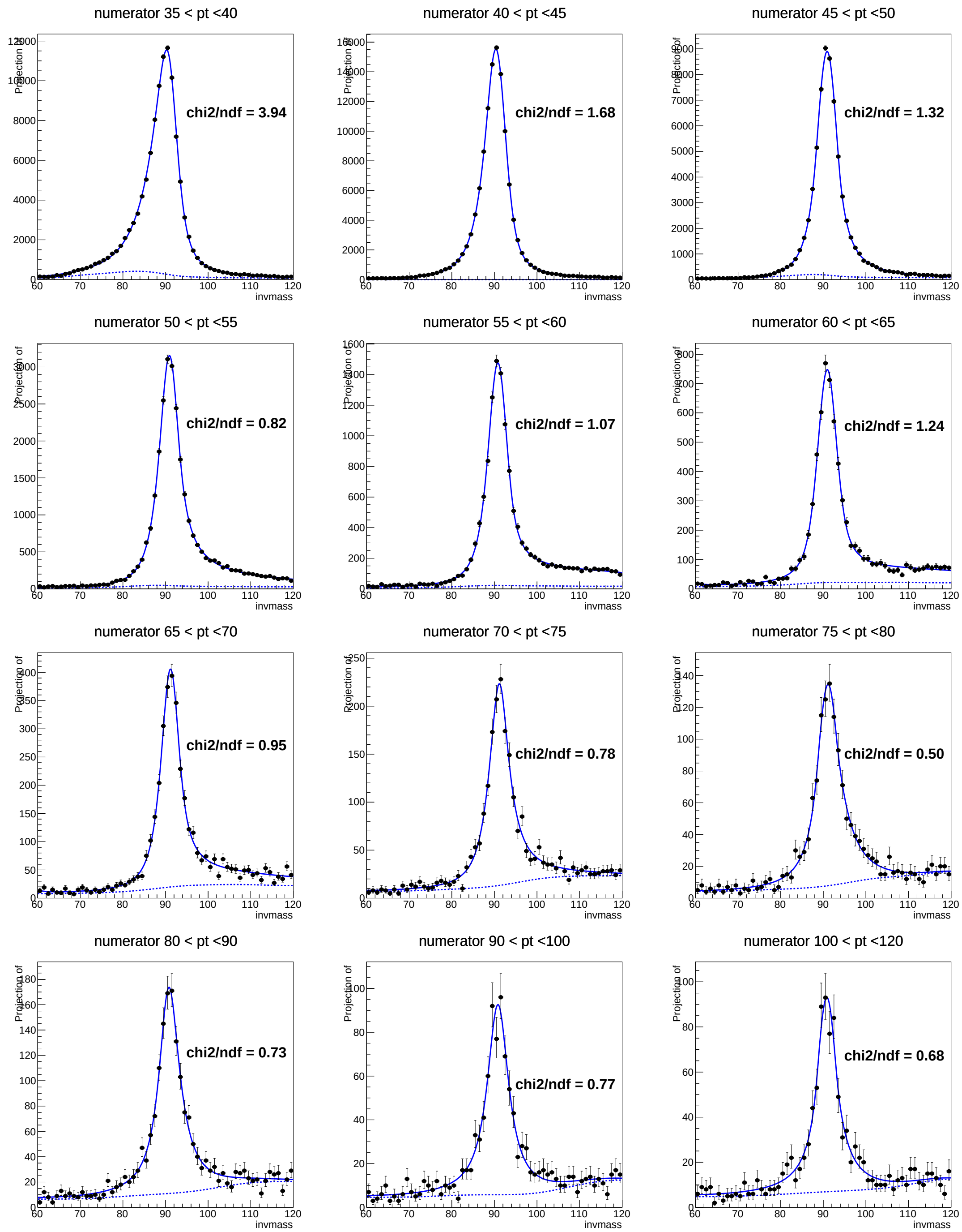


Figure 7.2: Fits to the tag-and-probe invariant mass of the numerator samples, which are defined as the samples with probe electrons failing to be matched with pixel track seeds. This fit is performed in various p_T bins. The black dots, the blue solid lines, and the blue dashed lines represent the data, the signal+background model, and the background-only model, respectively.

Once the normalization and parameters of the signal functions are determined from the fit, the N^{ee} and $N^{e\gamma}$ values are given by the integrals of signal shapes between 80 GeV and 101 GeV in the corresponding samples. The global fake rate for electrons with $p_T > 30$ GeV is:

$$f = (2.29 \pm 0.14)\%. \tag{7.1}$$

The electron misidentification rate measured in Z decay events will be applied to estimate the number and shape of fake photons in signal region. However, the kinematic distributions of signal photon candidates may differ from the ones from Z decay, and result in a different fake rate. To correctly extrapolate the misidentification rate to signal region, we considered its dependence on the following three variables: the transverse momentum of probe (p_T); the pseudorapidity of the probe (η); and the number of vertices in the event (N_{vtx}). The chosen of p_T and η is motivated by the tracker efficiency dependence on these two variables. N_{vtx}, on the other hand, is used to model the pileup effect. These dependencies are shown in Figure 7.3.

The dependencies of the misidentification rate on p_T and N_{vtx} can be described by parametrized functions. By studying the fake rate in simulations, we determined that the functional forms:

$$\begin{aligned} f(p_T) &= (a + b \cdot p_T)^{-\alpha} \\ f(N_{vtx}) &= c + \beta \cdot N_{vtx} \end{aligned}$$

can be used to model the fake rate dependencies on these two variables. The param-

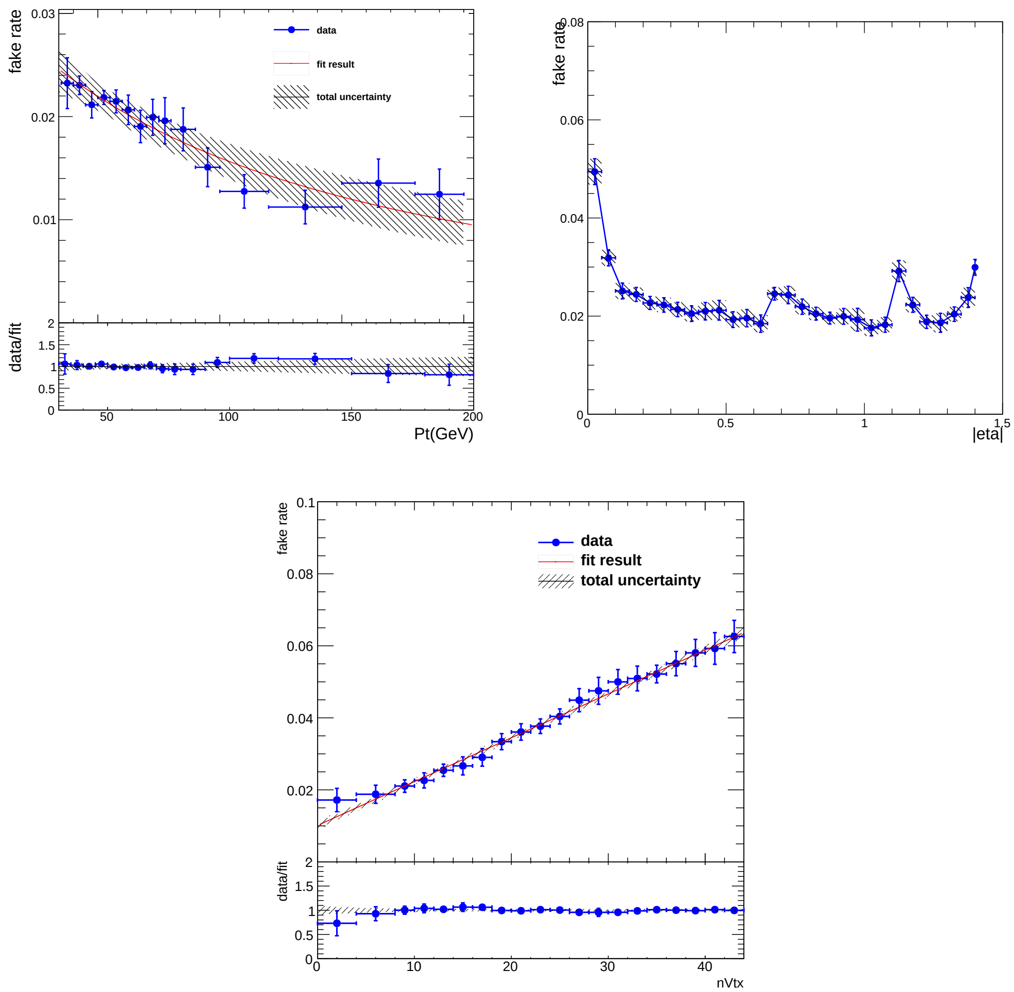


Figure 7.3: The electron misidentification rate as a function of p_T (top left), η (top right) and N_{vtx} (bottom), along with the systematic uncertainties.

eters obtained by fitting the functions to fake rate distributions are:

$$\begin{aligned} a &= 6.4 \pm 0.8 \\ b &= (2.9 \pm 0.8) \cdot 10^{-2} \\ \alpha &= 1.8 \pm 0.1 \\ c &= (5.9 \pm 1.1) \cdot 10^{-3} \\ \beta &= (9.2 \pm 0.4) \cdot 10^{-4} \end{aligned}$$

Assuming that the fake rate depends on each variable independently, the combined fake rate can be expressed as a function of (p_T, η, N_{vtx}) with the following form

$$f(p_T, \eta, N_{vtx}) = N \cdot f(p_T) \cdot f(N_{vtx}) \cdot f(\eta), \tag{7.2}$$

where N is a constant, and f(η) is the dependence of pseudorapidity, whose value is given by the fake rate of the corresponding η bin.

To fix the constant N, we applied the misidentification rate to the denominator sample and required the predicted number of fake photons to be the same as $N^{e\gamma}$. The resulting N value is:

$$N = (1.83 \pm 0.11) \cdot 10^3. \tag{7.3}$$

The electron-fake-photon rate calculated above is applied to the electron-enriched control sample to predict the contribution from misidentified electrons. Uncertainties of the misidentified electron background are evaluated with toy Monti-Carlo experiments. First, the parameters of the fake rate functions are pulled from a multi-variant Gaussian distribution with means given by the nominal fitting result, and covariance matrix given by the fitting errors. The ranges of the pulled parameters are restricted to be one sigma around the nominal values. The normalization factor N is recalculated using the new parameters. The toy MC is generated 1000 times to obtain distributions of the background variations. The one sigma band is constructed using the toy results.

This background estimation method is tested on simulated DY, WW, WZ and $t\bar{t}$ events. The misidentification rate derived from the DY samples using tag and probe method is compared to the true fake rate calculated by truth-matching the photons to generated electrons, and their difference are considered as a source of systematic uncertainty. The control events selected from the mixed simulation samples are scaled by the misidentification rate to predict the number of fake photons in the samples. As shown in Figure 7.4 and Figure 7.5, the predicted distributions show good agreement with the real misidentified electrons.

7.2 The Misidentification of Hadrons as Photons

Jets with electromagnetic fluctuations can deposit a significant fraction of energy in the ECAL and mimic photon signals. In particular, such misidentification occurs when the jet energy is mostly carried by a π^0 or η which decays into two nearly collinear photons.

The jet-to-photon misidentification rate is low, but the cross-section of QCD process is large, so the misidentified jets still have non-negligible contributions to the backgrounds. Such backgrounds are difficult to estimate using simulation because of the large statistics requirement and the possible mis-modeling of hadron fragmentation and hadronization. Therefore a data-driven method is deployed to predict this background. First, a control-to-fake transfer factor, defined as the ratio of estimated number of fake photons to the number of jet control objects, is derived in a low p_T^{miss} control region. This transfer factor will then be applied to a control sample with e/γ enriched jets to predict the jet-to-photon fakes in the high p_T^{miss} signal region.

We compute the transfer factor in a $p_T^{\mathrm{miss}} < 70$ GeV control region. To calculate the

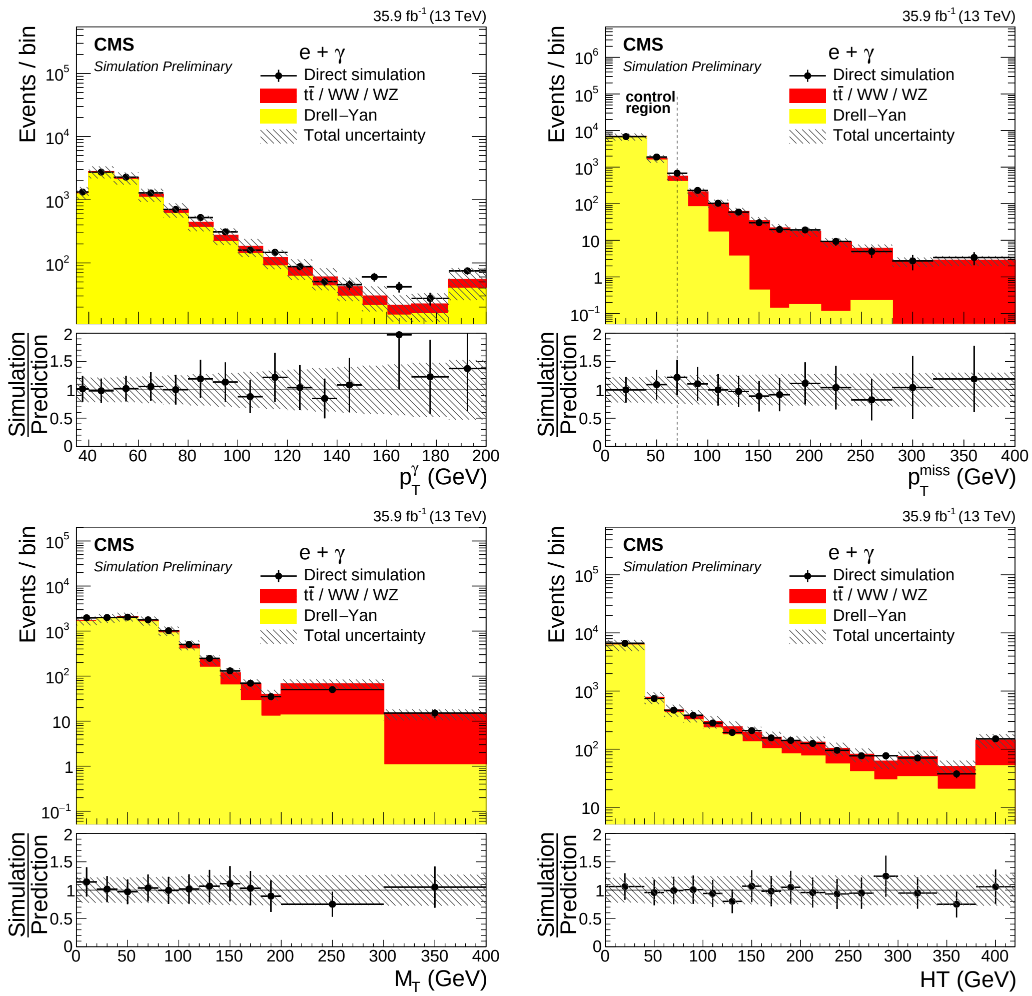


Figure 7.4: Simulation closure test for electron misidentification estimation in the $e\gamma$ channel. The misidentification rate derived from Drell-Yan sample is applied to a combination of Drell-Yan, $t\bar{t}$ and WW samples. Top left: photon p_T; top right: $p_T^{\rm miss}$; bottom left: M_T; bottom right: HT.

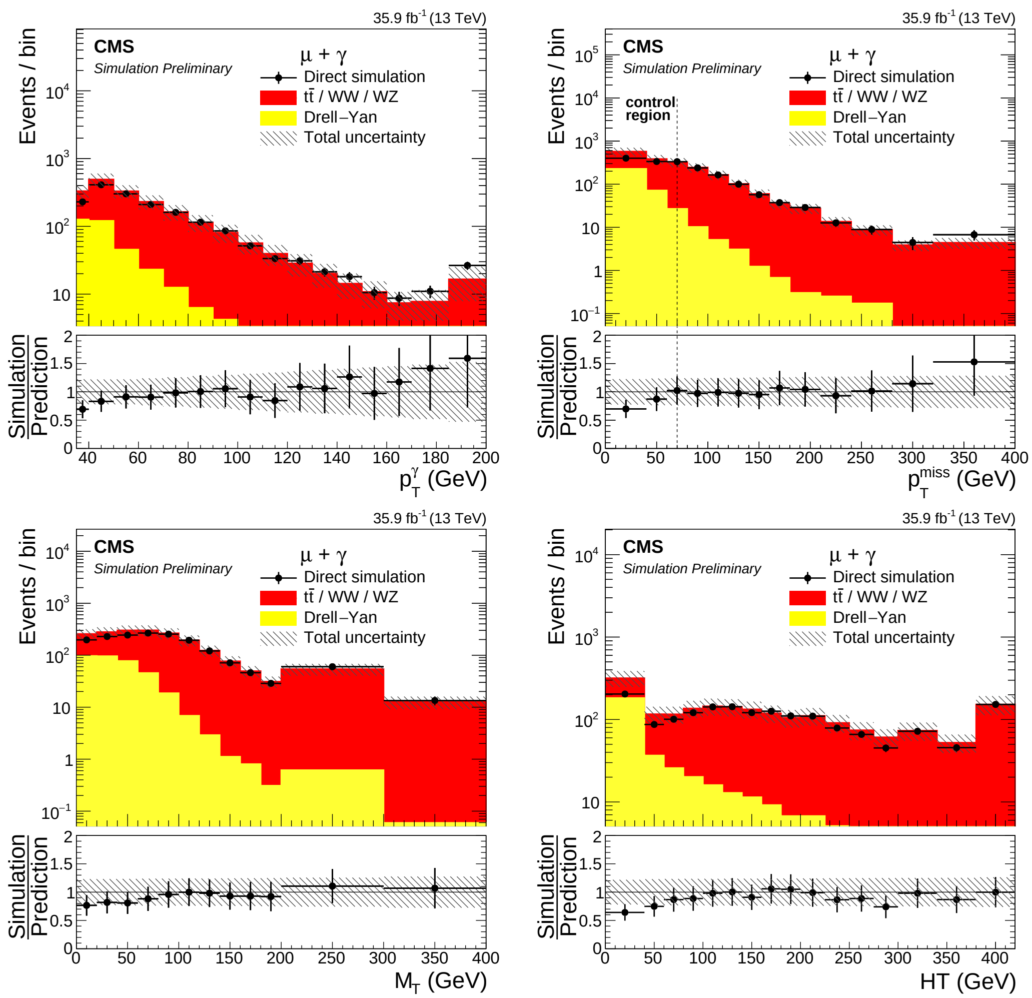


Figure 7.5: Simulation closure test for electron misidentification estimation in the $\mu\gamma$ channel. The misidentification rate derived from Drell-Yan sample is applied to a combination of Drell-Yan, $t\bar{t}$ and WW samples. Top left: photon p_T; top right: p_T^{miss}; bottom left: M_T; bottom right: HT.

ratio between the fake photons and jet control samples, we first need to determine the fraction of hadrons within the candidate photons. Compared to the prompt photons, the fake photons tend to have wider shower shapes and more hadronic activities. Therefore we can use variables like $\sigma_{i\eta i\eta}$ and $Iso_{h^\pm}$ to discriminate between true photons and fakes. Figure 7.6 shows the $Iso_{h^\pm}$ distribution in the GJet simulation samples.

Because the fake and true components of the photon candidates have different $Iso_{h^\pm}$ shapes, we can determine the contribution of fake photons by a template fit to the isolation distribution of the photon objects. This fit is performed using the $p_T^{\rm miss} < 70$ GeV samples formed in the DoubleEG and MuonEG dataset for the $e\gamma$ and $\mu\gamma$ channel respectively. First, a well identified lepton passing all the offline selections including trigger matching is required. A photon is then selected with all identification criteria excluding the $\sigma_{i\eta i\eta}$ and $Iso_{h^\pm}$ cut. Trigger leg matching is also required for the photon in order to keep the same phase space as the signal photons. Out of such photons, those passing the shower shape cut, i.e. $\sigma_{i\eta i\eta} < 0.01031$, are used as the fit target, and the $Iso_{h^\pm}$ distribution of the photons falling in the sideband $0.01031 < \sigma_{i\eta i\eta} < 0.015$ forms the hadronic template. The true photon template, on the other hand, is selected from the GJet simulation with the photon truth-matching to a generator level prompt photon. The fit to the $Iso_{h^\pm}$ distribution is performed using the ROOFIT package.

Once the normalization of the hadronic template is determined, an integral between $0 < Iso_{h^\pm} < 1.295$ is performed on the template to estimate the number of fake photons in the candidate samples. The fraction of fake photons, f, is defined as:

$$f = \frac{N_{fake}}{N_{sig}}, \tag{7.4}$$

where N_{fake} is the estimated number of fake photons passing $Iso_{h^\pm} < 1.295$, and N_{sig} is the number of events passing all the offline selections.

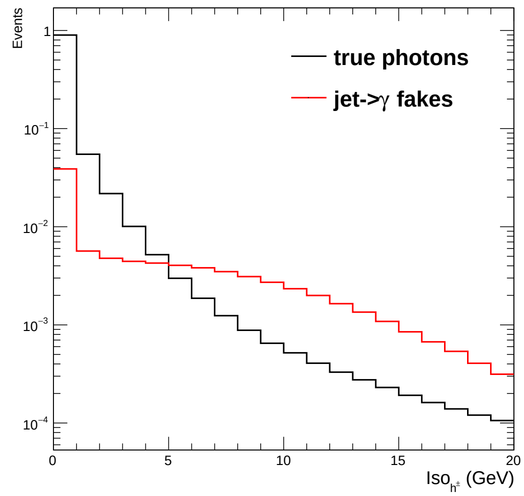


Figure 7.6: The normalized $Iso_{h^{\pm}}$ distributions of the true (black line) and fake (red line) photons in GJet sample.

Though most of the true photons have narrow shower shape, a small fraction of the true photons still fall in the $\sigma_{i\eta i\eta}$ sideband and therefore contaminate the hadronic template. Having true photons in the hadronic template will result in overestimating the fraction of fakes. To correct the hadronic $Iso_{h^{\pm}}$ template shape, an iterative method is deployed to remove the true photon contaminations.

This method scales the normalized true photon $Iso_{h^{\pm}}$ shape taken from the GJet MC by a factor $N_{\text{true-in-sb}}$, and subtract this shape from the raw hadronic template. The factor, $N_{\text{true-in-sb}}$, is calculated as:

$$N_{\text{true-in-sb}} = N_{sig} \cdot (1 - f_i) \cdot \frac{p_{sb}}{p_{sig}}, \tag{7.5}$$

where f_i is the fraction of fakes estimated from the template fit in the ith iteration, p_{sig} and p_{sb} are the probability of a simulated true photon falling into the signal region and sideband region, respectively, as illustrated in Figure 7.7. This procedure is iteratively repeated several times until the calculated f_i converges, i.e. $|f_i - f_{i+1}| < 0.001 f_i$. This method is validated using the GJet simulation samples and shows a considerable improvement on the fake fraction estimation. The validation results are shown in Figure 7.8 .

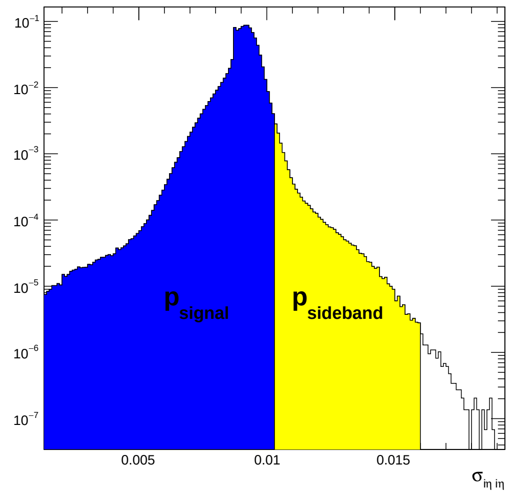


Figure 7.7: The $\sigma_{i\eta i\eta}$ distribution of the true photons in GJet sample. Most of the true photons fall in signal region (blue), however, there is still a small fraction of true photons enter the sideband region (yellow).

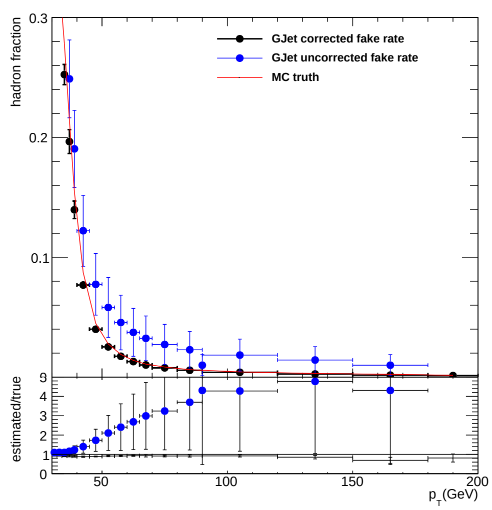


Figure 7.8: The hadron fractions estimated in GJet sample. The estimation with iteration corrections is shown in black dots, and the one without such correction is shown in blue.

The $Iso_{h^{\pm}}$ fit is performed in various p_T bins on the data. The fitting results for the $e\gamma$ channel are shown in Figure 7.9. The calculated fraction of fakes as a function of photon p_T are shown in Figure 7.10. The fractions of hadrons are lower in the $\mu\gamma$ channel than those in the $e\gamma$ channel. This is because the MuonEG HLT trigger is seeded on isolated L1 objects, which have tighter isolation requirements than in the $e\gamma$ trigger.

Once the fraction of fakes are determined, the number of fakes in the candidate photons can be calculated and used as the numerator sample for the transfer factor calculation. The denominator sample of the transfer factor is a jet-enriched control sample. This sample is formed by replacing the candidate photon with a control object, which fulfills the following criteria:

- pass H/E, Iso_{pho}, Iso_{h^0} cut in the loose-ID
- pass the pixel veto, electron veto and FSR veto
- $\sigma_{i\eta i\eta} > 0.01031$
  or $1.29 < Iso_{h^{\pm}} < 15$ GeV

The transfer factor is then determined in the control region $p_T^{\text{miss}} < 70$ GeV by calculating the ratio of numerator events to the denominators.

It is useful to parameterize the p_T dependence of the transfer factors using analytical functions so that the result measured in low pt regions can be interpolated to the high p_T region where statistics are limited. We choose to use a sum of two exponential functions to fit the photon p_T spectrum. The fitting results are shown in Figure 7.11.

From the fits, we obtained the function forms for the $e\gamma$ channel:

$$R_{fake/control} = \frac{1.43 \times 10^4 \cdot e^{-0.069 \cdot p_T} + 5.11 \times 10^2 \cdot e^{-0.030 \cdot p_T}}{4.08 \times 10^4 \cdot e^{-0.055 \cdot p_T} + 5.51 \times 10^2 \cdot e^{-0.019 \cdot p_T}}, \tag{7.6}$$

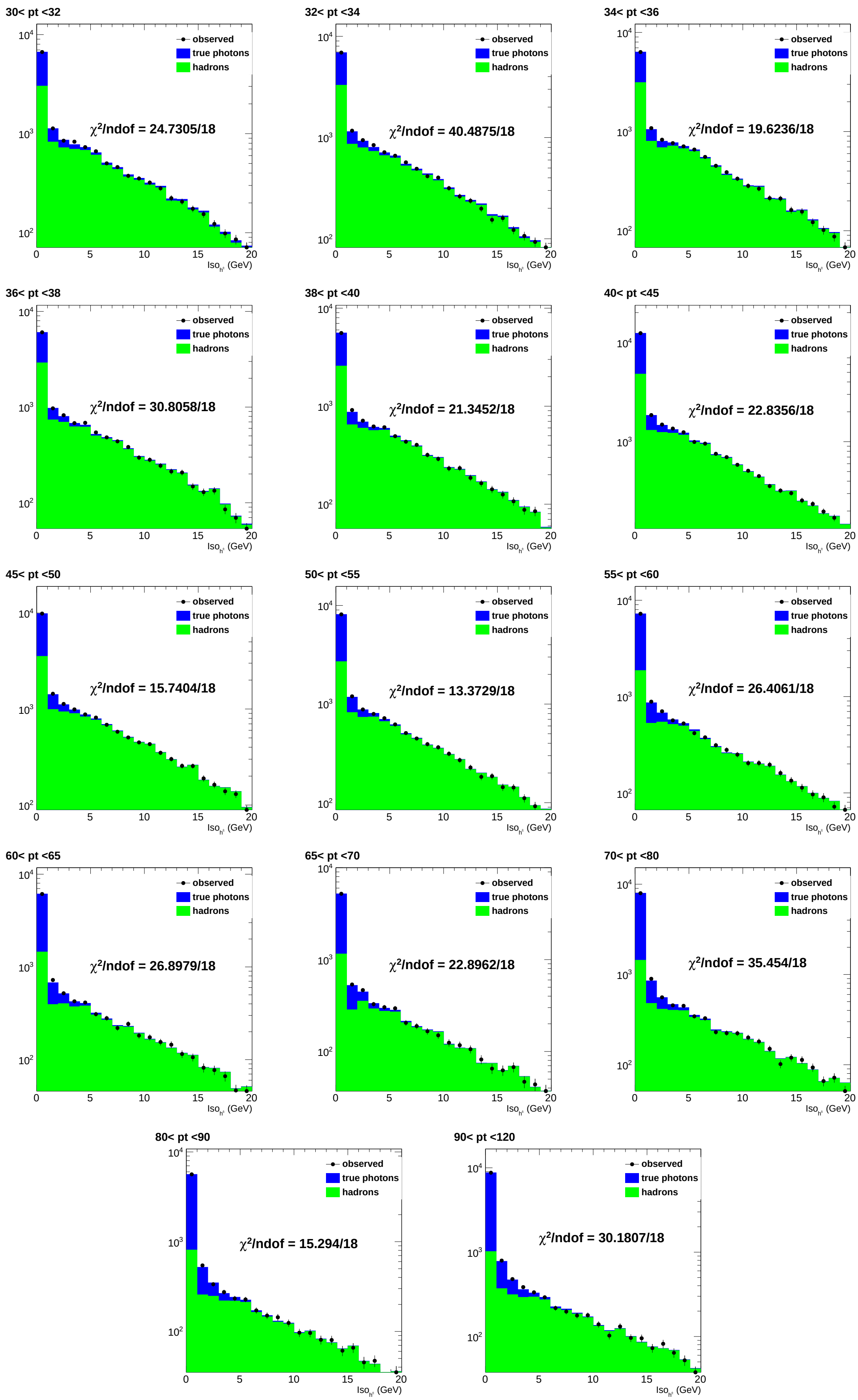


Figure 7.9: Template fits for hadron fraction measurements in $e\gamma$ channel.

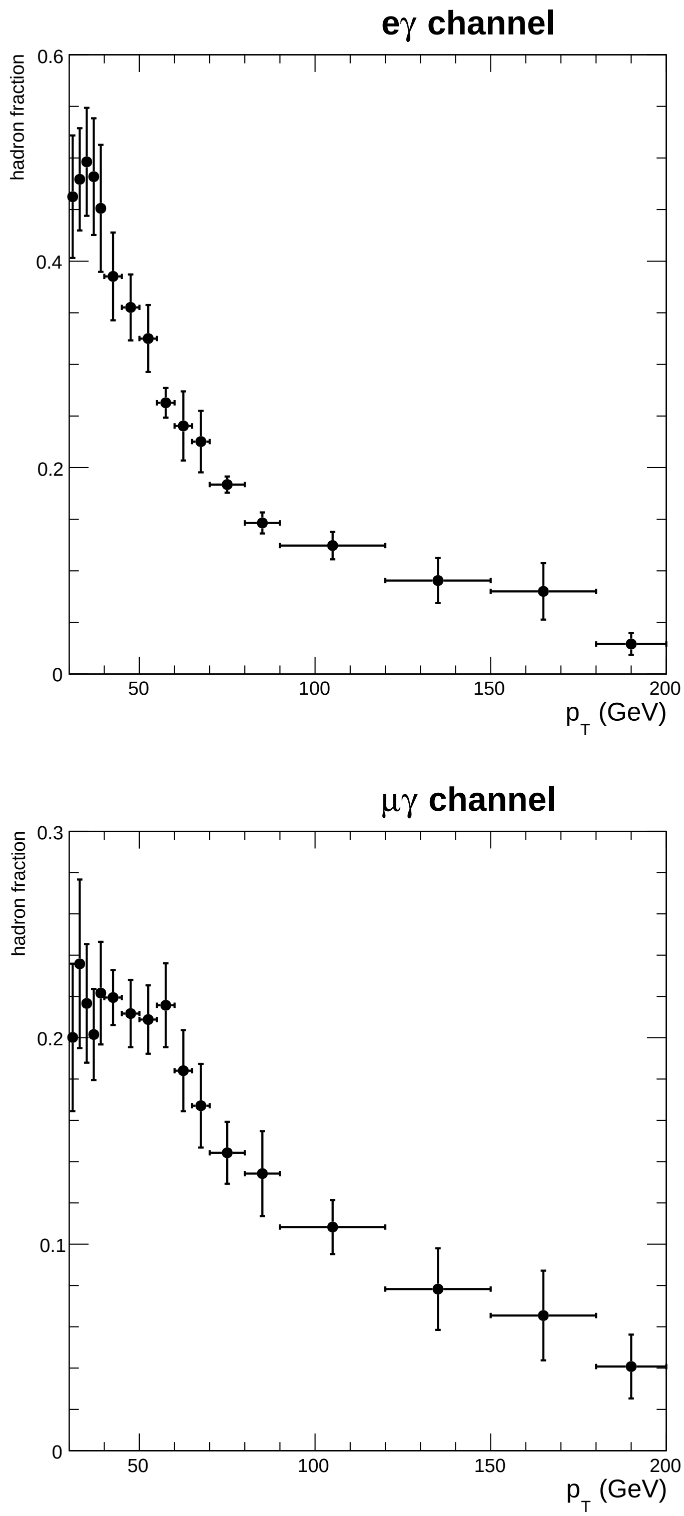
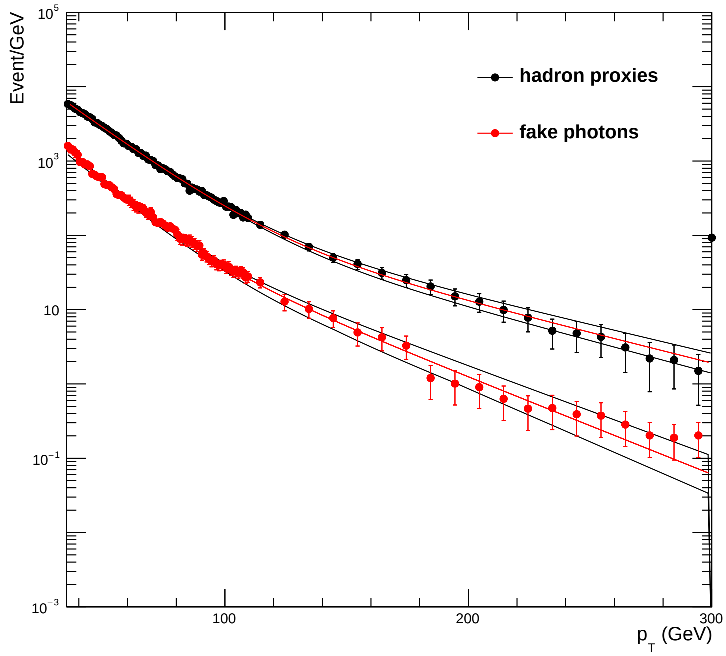


Figure 7.10: The hadron fractions estimated in DoubleEG (top) and MuonEG (bottom) dataset.



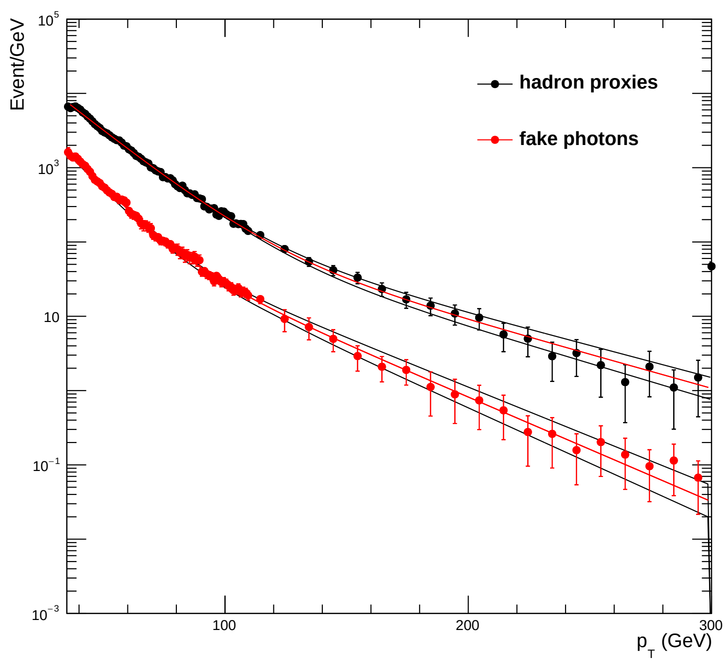


Figure 7.11: Function fits for jet-fake-photon transfer factor measurements in $e\gamma$ (top) and $\mu\gamma$ (bottom) channel. The p_T distributions of the control objects and fake photons are shown in black and red dots respectively. The one sigma band of the fitting are shown in black lines.

and $\mu\gamma$ channel:

$$R_{fake/control} = \frac{2.39 \times 10^4 \cdot e^{-0.082 \cdot p_T} + 3.33 \times 10^2 \cdot e^{-0.030 \cdot p_T}}{6.01 \times 10^4 \cdot e^{-0.059 \cdot p_T} + 5.90 \times 10^2 \cdot e^{-0.021 \cdot p_T}}. \tag{7.7}$$

The choice of the $\sigma_{i\eta i\eta}$ sideband is identified as the major source of systematic uncertainties of the fraction of fakes. Since the $Iso_{h^\pm}$ and $\sigma_{i\eta i\eta}$ are correlated, inaccuracy on the fake fraction estimation is expected even though the iterative method is applied. This uncertainty can be assessed by varying the sideband definition and repeating the fit with the modified template. We scan the lower bound of the sideband from 0.01031 to 0.0112 in a step of 0.00005, and scan the upper bound from 0.0140 to 0.0185 in a step of 0.0005. The scan result in one of the p_T bin is shown in Figure 7.12. The full variation is taken as the systematic uncertainty.

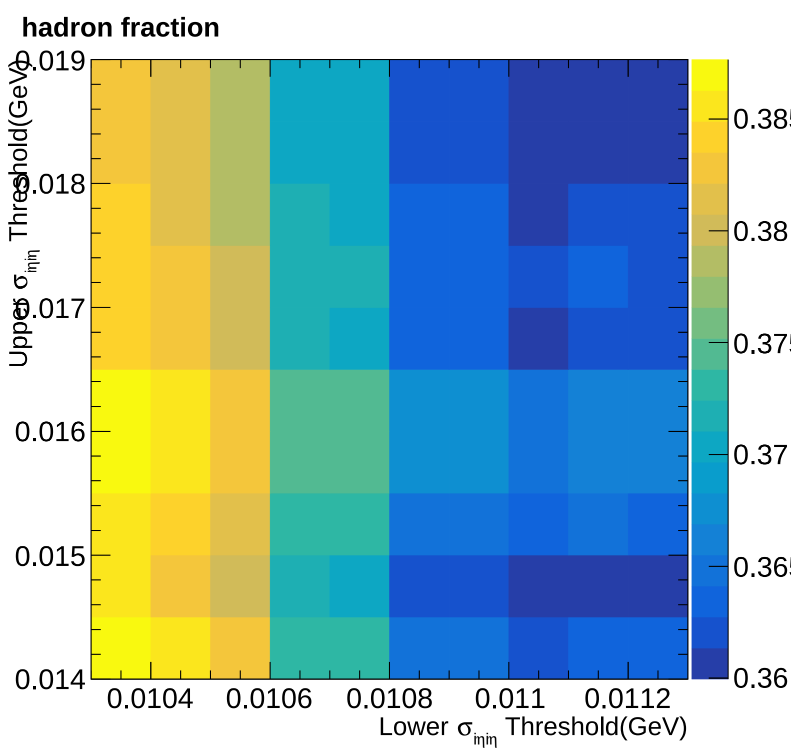


Figure 7.12: 2D distribution of the fraction of hadrons for photons with 50 GeV $< p_T <$ 55 GeV. The full variation will be used as systematic uncertainties of the template fit.

The systematic uncertainties on the fraction of fakes get propagated to the functional form of the fake photons via the p_T spectrum fitting. To form the 1 σ error band of the p_T distribution, we performed toy MC experiments with the parameters and errors from the p_T fitting. Assuming the fitting parameters have gaussian distributions around their nominal values, we construct a multi-variant Gaussian PDF using the nominal fitting values and their covariant matrix. 1000 sets of parameters are then generated from this Gaussian distribution. The 1 σ band is formed by wrapping these toy distributions, as shown in 7.11.

7.3 The Misidentification of Hadrons as Leptons

Events with jets plus prompt photons also enter the signal region if the jets get misidentified as leptons. In this analysis, we consider all leptons that do not originate from a $W^{\pm}$ or Z^0 as fakes. This includes leptons from heavy-flavour decays, misidentified jets, light-meson decays, and electrons from photon conversions. Studies on simulated QCD events show that the fake muons mostly come from heavy flavor quarks, while the fake electrons predominantly come from light flavor jets with significant electromagnetic components.

We estimate the contribution of fake leptons in a manner analogous to the estimation of fake photons: select a control sample of events enriched in the fake leptons, and then use a scale factor to extrapolate this sample to the background in the signal region. The selection of the control sample is close to that of the signal candidate. We require the control sample to contain one candidate photon, one fake lepton control object, and no candidate leptons.

The scale factor for the fake lepton control sample is derived in a $40 < p_T^{\text{miss}} < 70$ GeV control region using a template fit to the $\Delta\phi\,(\ell, p_T^{\text{miss}})$ distribution. After removing the contribution of fake photons, events in the control region are dominated by $W\gamma/Z\gamma$ processes and fake lepton contributions. p_T^{miss} in the fake lepton events usually come from mismeasured objects. Thus the $\Delta\phi\,(\ell, p_T^{\text{miss}})$ shape of the fake lepton background is different with that of the $W\gamma/Z\gamma$ processes. This feature allows us to perform a two templates fit to the $\Delta\phi\,(\ell, p_T^{\text{miss}})$ of the data and simultaneously determine the normalization for the fake lepton control sample and $W\gamma/Z\gamma$ simulated samples. The normalized control sample then provides the estimate of the background contribution from events with fake leptons in the signal region. The template fit is described in details in Section 7.4.

7.3.1 Fake Electron Control Sample

We find from simulations of QCD events that fake electrons commonly result from light flavor jets which shower significantly in the ECAL and get misidentified as electrons. These fake electrons have some characteristic features like large energy sum around the objects, broader cluster shapes in η, and less consistent match between the reconstructed track momentum and the corresponding cluster energy. We therefore impose the following requirements on the electron objects of our control sample:

- Electron $p_T > 25$GeV, $|\eta| < 1.442$ or $1.56 < |\eta| < 2.5$.
- Pass the medium H/E, $|\frac{1}{E} - \frac{1}{p}|$, nMissHits and conversion veto cuts
- Fail any of the $\sigma_{i\eta i\eta}$, $|\Delta\eta|$, $|\Delta\phi|$ and mini-isolation cuts.
- mini-isolation < 0.4

From simulation, we find the selection has negligible contamination from prompt electrons.

To check the consistency of the background modeling, we compare events in data with control electrons to events in QCD simulation with electrons matched to fakes.

To ensure sufficient statistics in the QCD simulation, we replace the requirement that the event contains a loose photon with the requirement that it contains a loose jet, as described in Section 6.5, and we insist that $\Delta R\,(\ell,\,\text{jet}) > 0.8$. The distributions of several kinematic variables are compared between the two samples, as shown in Fig. 7.13. Reasonable agreements are obtained in all distributions. The $\Delta\phi\,(\ell, p_T^{\text{miss}})$ distribution is used in the template fit to determine the normalization factor of the control sample.

7.3.2 Fake Muon Control Sample

Simulation indicates that, in contrast to the electron case, muon fakes tend to be the product of real, non-prompt muons from the decay of heavy-flavor quarks. Because these are real muons, the distributions of their primary ID variables tend to be indistinguishable from those of prompt muons. The primary handle for distinguishing these objects is thus the mini-isolation variable, which tends to be greater for objects produced in the decays of heavy quarks. We therefore construct muon control sample by requiring candidate muons to pass the following selection criteria:

- Muon $p_T > 25\text{GeV}$, $|\eta| < 2.4$.
- Passes cut-based medium ID.
- $0.2 <$ mini-isolation < 0.4.

From simulation we conclude that the muon control sample has less than 5% contamination from prompt muons.

We again check the consistency of the background modeling by comparing the kinematic distributions of data events with control muons to simulated events with fake muons. We switch the requirement of a loose photon in the event to the require-

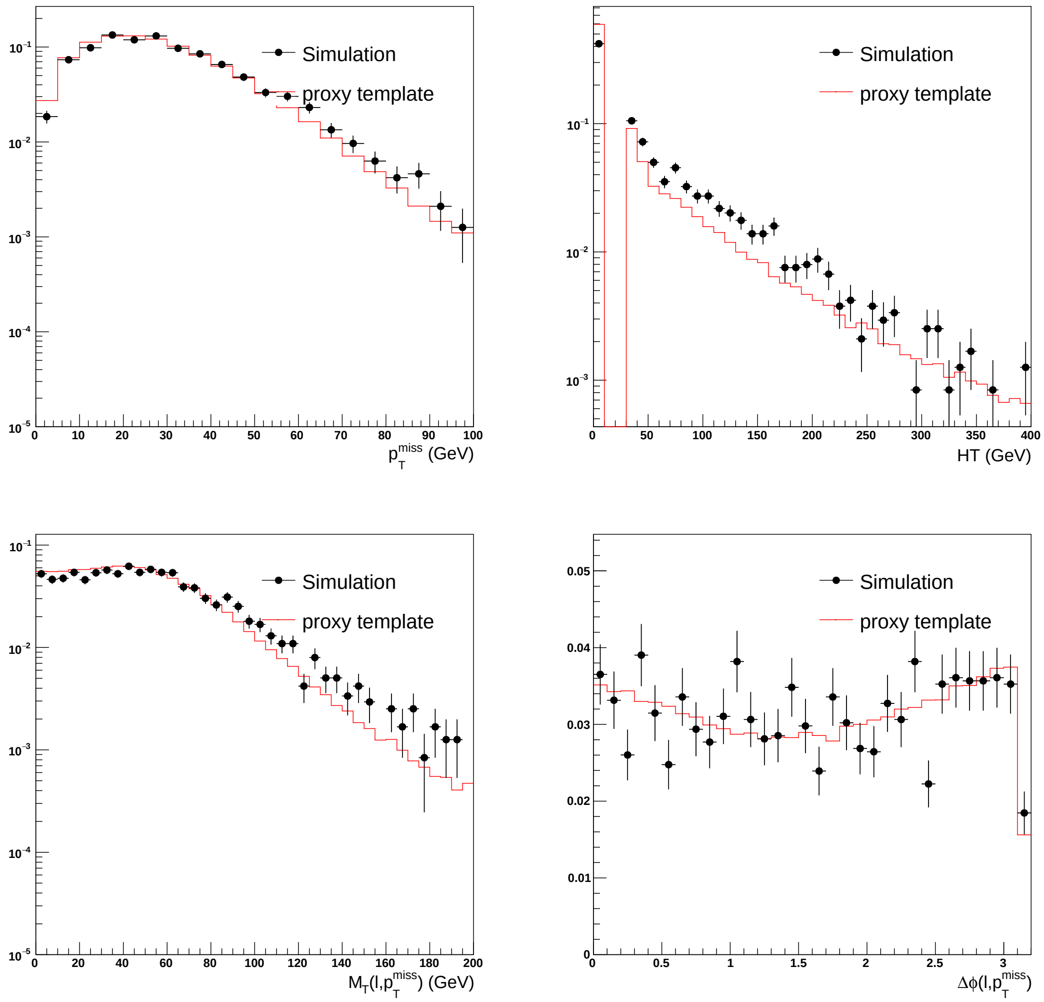


Figure 7.13: Comparison of various kinematic distributions between simulated QCD events with a fake electron and events in data with a control electron. The distributions are area-normalized to facilitate a comparison of their shapes.

ment of a loose jet, which is well separated from the lepton candidate ($\Delta R\,(\ell, \text{jet}) >$ 0.8). The distributions of these variables are shown in Fig. 7.14.

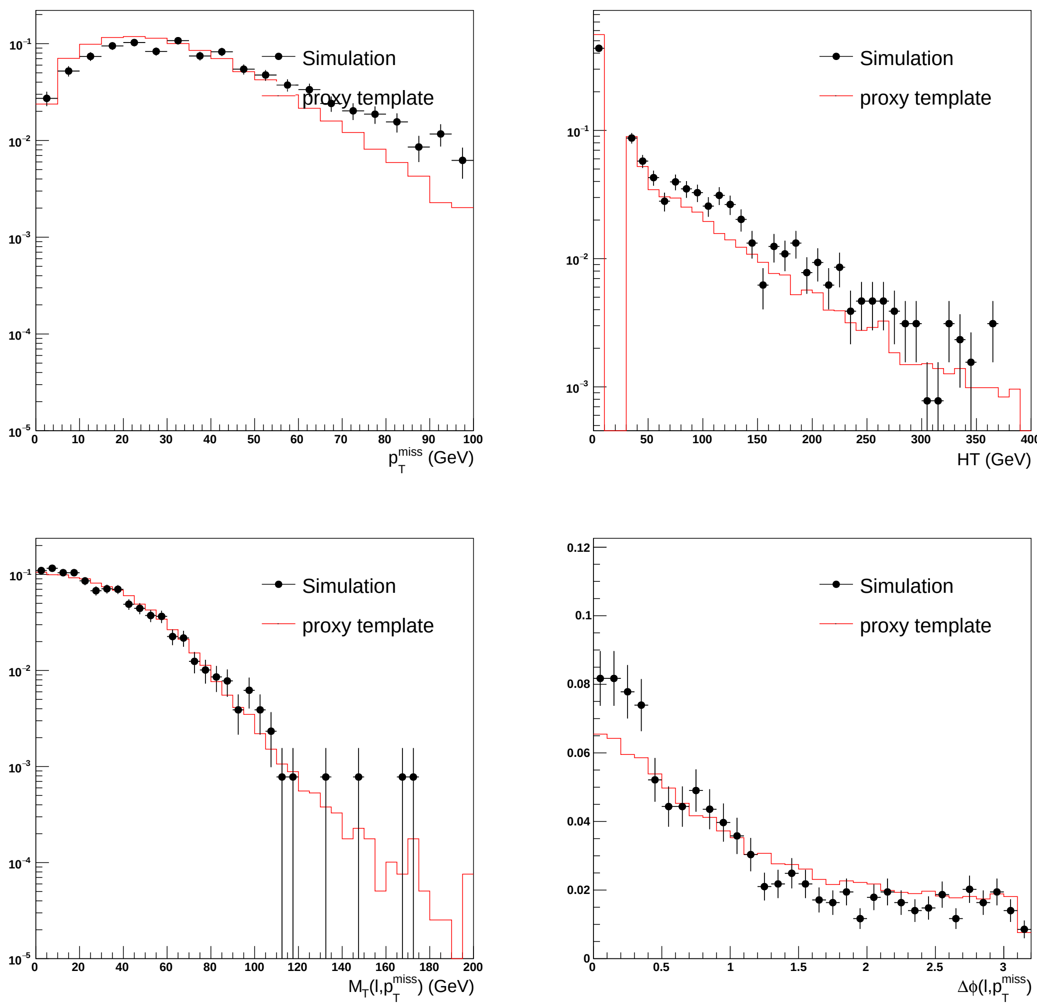


Figure 7.14: Comparison of various kinematic distributions between simulated QCD events with a fake muon and events in data with a control muon. The distributions are area-normalized to facilitate a comparison of their shapes.

7.3.3 Corrections on Proxy Samples

Because the mini-isolation is computed using a cone size depending on the momentum of the lepton, the mini-isolation cuts could have non-linear efficiencies. Thus the

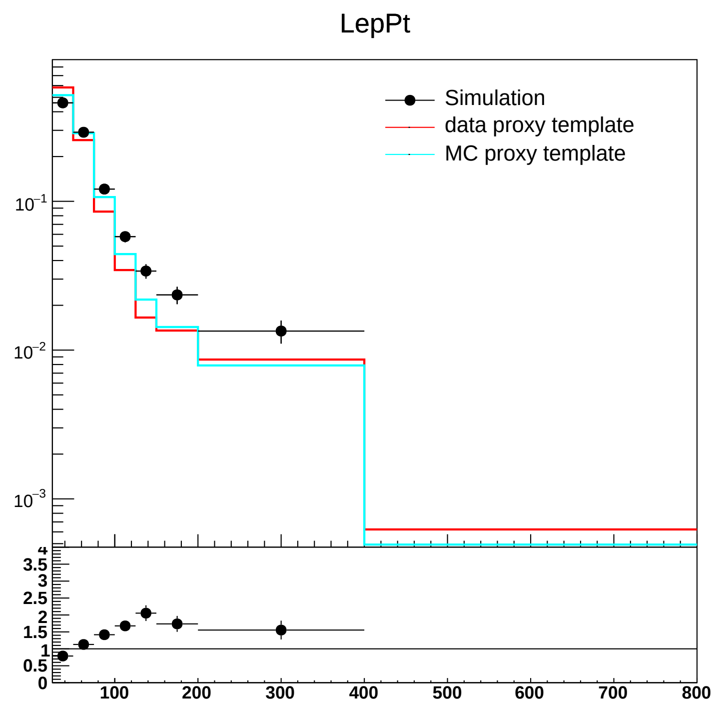
p_T distribution of the fake lepton background may not be perfectly modeled by the control sample. By comparing the p_T distribution of the control objects and simulated QCD events, we find disagreements in the fake electron p_T shapes. Though the lepton p_T is not directly used in defining the signal bins, a mismodeled lepton p_T could still affect the determination of the $V\gamma$ and fake lepton scale factors. Therefore, we use the ratios between the simulated events and fake electron control sample as correction factors to reweight the fake electron control samples. For the fake muons, the shape of the control sample is more consistent with the QCD simulations, moreover, the rate for a jet to be misidentified as a muon is much smaller than that of an electron. Thus we won't correct the muon control samples. Figure 7.15 shows the p_T distributions and the MC-to-data ratios. Table 7.2 summarize the correction factors we used for the electron control reweighting.

Table 7.2: Correction factors used in electron control reweighting.

p_T (GeV)	e control corrections
25-50	0.79± 0.02
50-75	1.12± 0.04
75-100	1.41± 0.08
100-125	1.67± 0.14
125-150	2.05± 0.23
150-200	1.73± 0.23
200-400	1.55± 0.27
> 400	1.0

7.4 $W\gamma$ and $Z\gamma$ Background

The standard model production of W or Z boson in association with a photon is the major background of this search. Since the cross-section measurement of this process at 13 TeV is not published yet, and the calculated cross-section for WW and WZ processes show deviations from measured results, this background is estimated by scaling the simulated samples by a factor derived in the control region using a template fit.



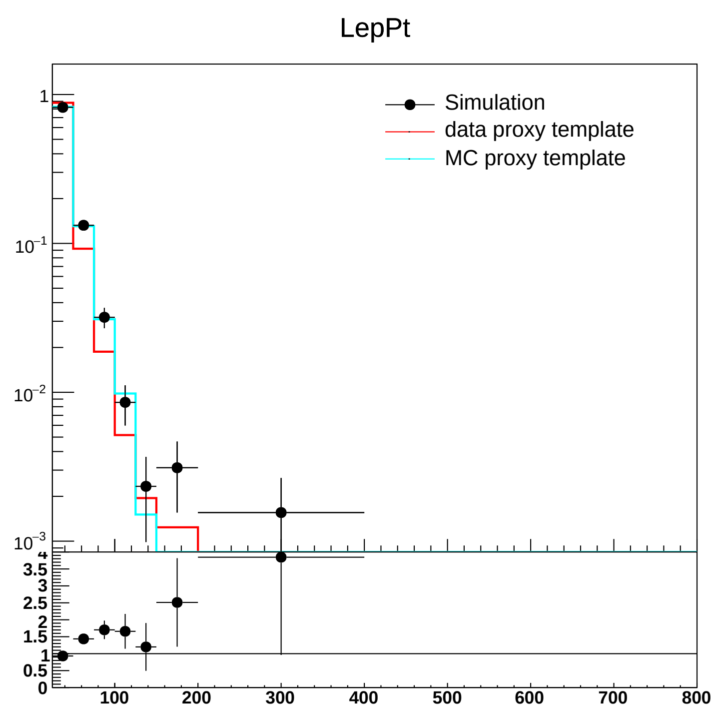


Figure 7.15: Comparison of lepton p_T distributions between the QCD simulation and fake lepton control samples. Top: $e\gamma$ events, bottom: $\mu\gamma$ events. The ratio can be used as correction factors.

The $W\gamma$ sample is formed by mixing the inclusive $WGToLNuG$ sample with two p_T-binned samples: WGJets_MonoPhoton_PtG-40to130 and WGJets_MonoPhoton_PtG-130. The former one is truncated at 50 GeV and the high-statistics p_T-binned samples are used for p_T above 50 GeV. All of the three $W\gamma$ samples are generated with MadGraph at leading order. To account for higher order corrections, a constant NNLO k-factor of 1.34 is applied. Figure 7.16 shows that a smooth photon p_T distribution can be obtained via this mixing procedure. For the $Z\gamma$ process, we use an inclusive NLO sample with $p_T > 10$ GeV. Because the $Z\gamma$ sample has a $M_{ll} > 30$ GeV cut at matrix-element level, we use part of the Drell-Yan samples to supplement the events with $M_{ll} < 30$ GeV. All of these simulated samples are then normalized to 35.9 fb^{-1} and mixed to form the $V\gamma$ template.

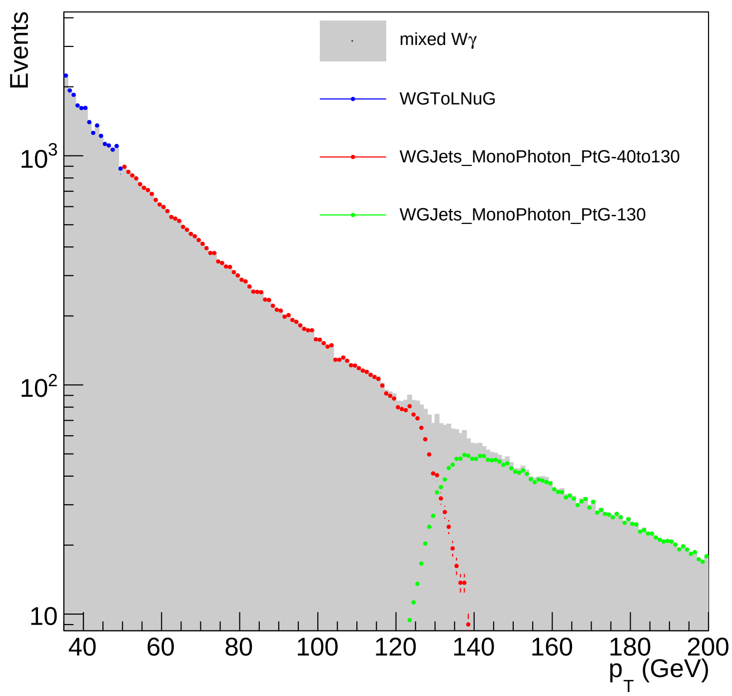


Figure 7.16: Photon p_T distributions in individual $W\gamma$ samples and their combination.

The initial state radiation (ISR) can boost the total transverse energy of the hard scattering and affect kinematic quantities like object p_T or p_T^{miss}. Studies using $Z\gamma \to \mu\mu\gamma$ events show that there is a systematic discrepancy between the data and simulated events. To improve the modeling of the $V\gamma$ samples, we assign each simulated event a correction factor according to its ISR p_T. The ISR p_T is calculated using the ISR-jet counting algorithm. The correction factors are derived from the analysis of $Z\gamma \to \mu\mu\gamma$ events.

The same MuonEG dataset and HLT_MuEG triggers are used to select events with one photon and exactly two isolated muons. The leading muon and the photon should pass the criteria deifned in Section 6.2, and the sub-leading muon is required to have $p_T > 15$ GeV. In addition, the di-muon mass is required to be between 80 and 100 GeV. The ISR p_T is deifned as the vector sum of all jets which have $p_T > 30$ GeV and $|\eta| < 2.5$. To avoid counting the decays products of the Z boson, the jets are cleaned from isolated leptons. Contributions of fake photons are estimated and removed using the same methods described in Section 7.2. Other rare backgrounds, including $t\bar{t}\gamma$, WWG and WZG are estimated using MC sampels.

Figure 7.17 shows the comparision between the ISR p_T spectrum of the data and simulation. The simulation over-predicts the number of events at high pt by about 35%. The data-to-MC ratio in coarse pt bins are taken as the correction factors, as shown in Table 7.3. We reweight the WG and ZG samples event-by-event using the ISR correction factors while keeping the total normalization unchanged. The ISR reweighting simultaneously improves the modeling of multiple variables, including photon p_T, p_T^{miss} and M_T, as shown in Figure 7.18.

The corrected $V\gamma$ distributions are normalized with scale factors derived in the control region to estimate the backgrounds in the signal region. Because the angular distance between the signal lepton and p_T^{miss} has different shapes in the $V\gamma$ and misidentified object background, we choose to use $\Delta\phi(\ell, p_T^{\text{miss}})$ as the distribution for the template fit. For the fake lepton background, the missing transverse momentum is

ISR p_T (GeV)	Correction $\pm$ stat. unc. $\pm$ syst. unc.
0-50	1.015 $\pm$ 0.018 $\pm$ 0.04
50-100	1.110 $\pm$ 0.015 $\pm$ 0.03
100-150	0.845 $\pm$ 0.023 $\pm$ 0.04
150-200	0.715 $\pm$ 0.035 $\pm$ 0.07
200-250	0.730 $\pm$ 0.056 $\pm$ 0.09
250-300	0.732 $\pm$ 0.086 $\pm$ 0.08
$>$ 300	0.642 $\pm$ 0.069 $\pm$ 0.08

Table 7.3: Correction factors for the ISR reweighting.

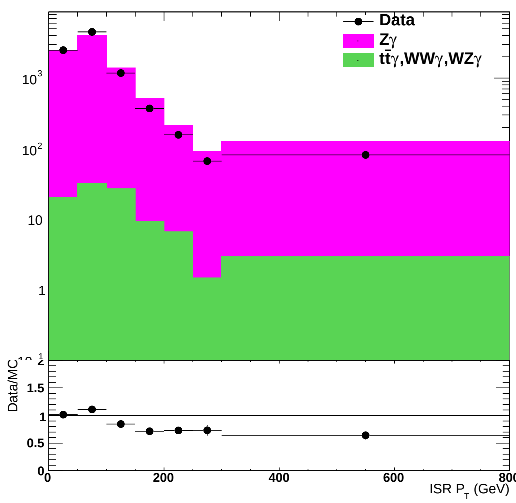


Figure 7.17: The electron misidentification rate as a function of p_T (top left), η (top right) and N_{vtx} (bottom), along with the systematic uncertainties.

typically caused by mismeasured object and tends to be aligned with the lepton. The $W\gamma$ events, on the other hand, has a neutrino and thus contains genuine p_T^{miss}. Figure 7.19 shows the distribution of these two sources. Fitting the two $\Delta\phi$ templates to the distributions of the data will simultaneously determine the normalization factors for the $V\gamma$ and fake leptons backgrounds.

The template fitting is performed in the control region, defined by 40 GeV $<$ p_T^{miss} $<$ 70 GeV. The lower cut is applied to suppress $Z\gamma$ events. Figure 7.20 shows the normalized p_T^{miss} distribution of the $W\gamma$ and $Z\gamma$ samples. By applying a 40 GeV

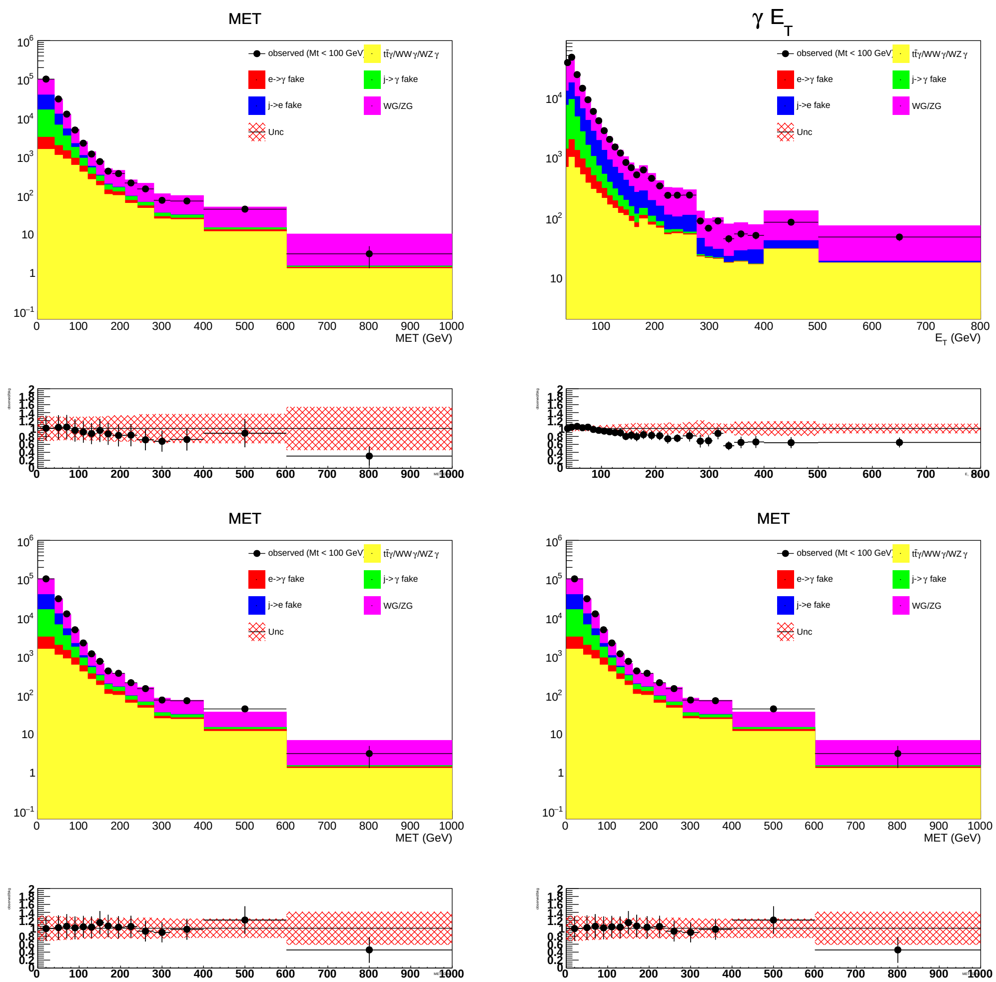


Figure 7.18: p_T^{miss} (left) and photon p_T distributions of the data and background predictions in the validation region. Top: no ISR reweighting applied, bottom: with ISR corrections.

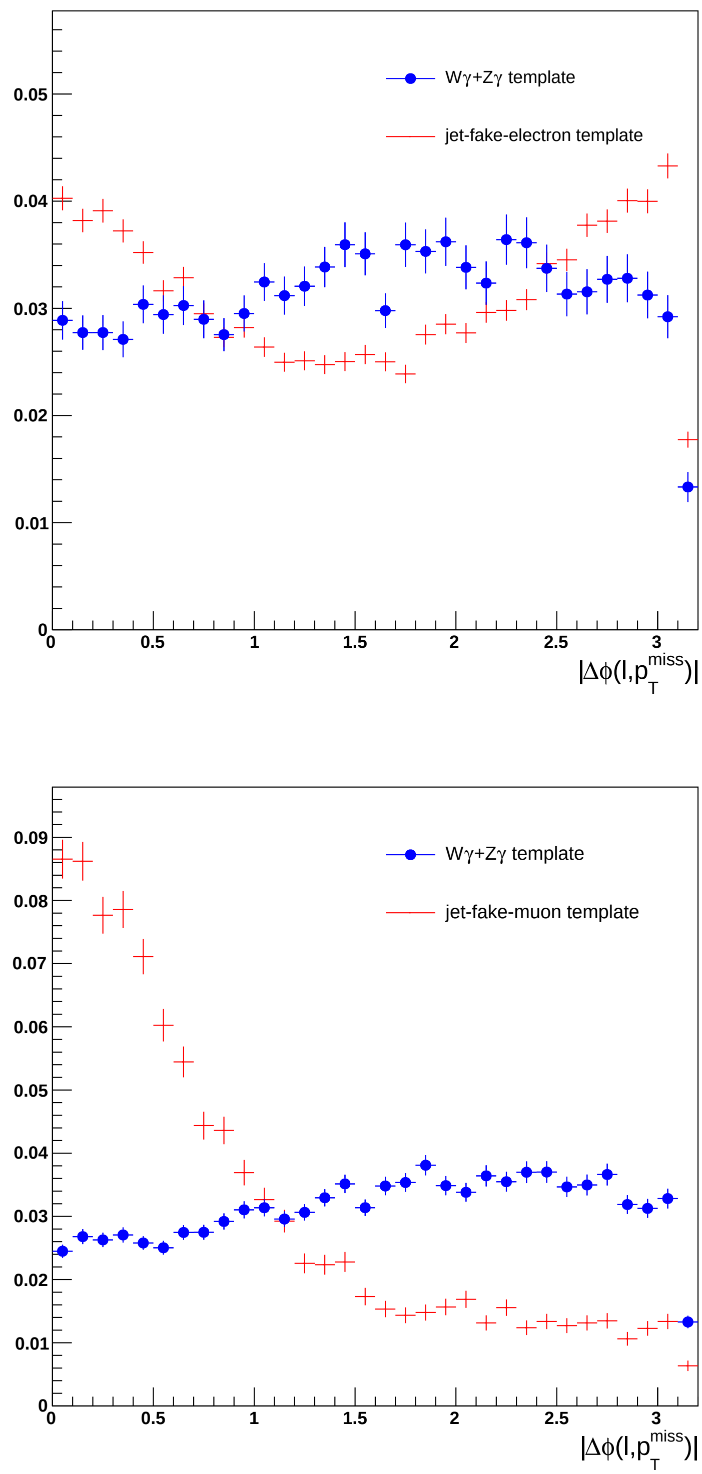


Figure 7.19: $\Delta\phi\,(\ell, p_T^{\mathrm{miss}})$ distribution of the $V\gamma$ sample (blue) and fake lepton template(red) in the $e\gamma$ channel (top) and $\mu\gamma$ channel (bottom).

p_T^{miss} cut, we select a control sample dominated by $W\gamma$ events.

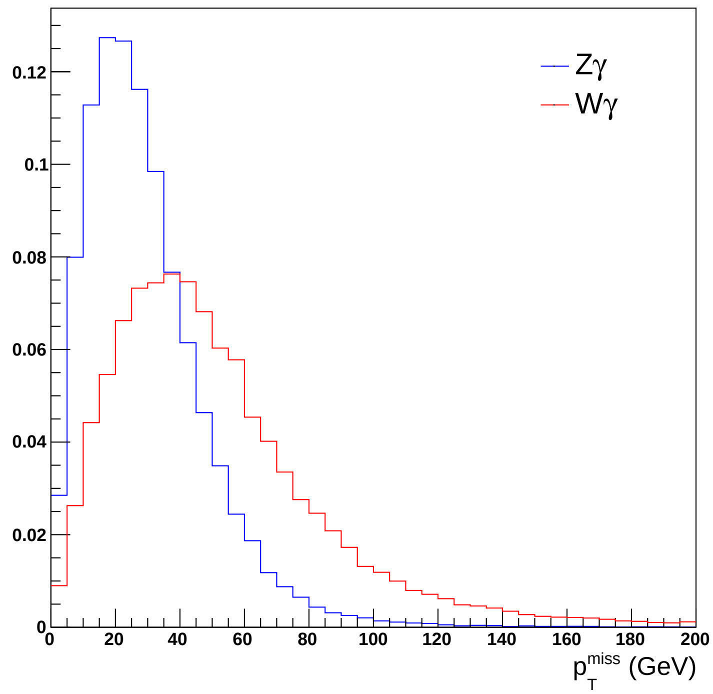


Figure 7.20: p_T^{miss} distributions of the $W\gamma$ (red) and $Z\gamma$ (blue) sample. The 40 GeV cut will remove a large fraction of the $Z\gamma$ contributions.

The target distribution of this template fit is the $\Delta\phi(\ell, p_T^{\text{miss}})$ shape of the data in the control region, with fake photon and rare EWK backgrounds subtracted. Two templates are then selected, one from the $V\gamma$ sample and the other one from the fake lepton samples. The fit is performed with the ROOFIT package using the binned maximum likelihood estimator.

Once the fraction of $V\gamma$ events in the control region is determined, the scale factor for the $V\gamma$ sample can be calculated as:

$$a_{V\gamma} = \frac{N_{con} \times f_{V\gamma}}{N_{V\gamma}}, \tag{7.8}$$

where N_{con} is the total number of events in the control sample, $f_{V\gamma}$ is the fraction of $V\gamma$ events derived from the template fit, and $N_{V\gamma}$ is the number of events in the $V\gamma$

template. After removing the $V\gamma$ events, the remaining events in the control sample are the contribution of fake leptons.

To study the possible p_T dependence of the $V\gamma$ scale factors, we perform the template fit in four lepton p_T bins: (25-50, 50-70, 70-100, >=100) GeV. Figure 7.21 shows that the fluctuation of the scale factors in different p_T bins are within systematic uncertainties. Since no obvious p_T dependence is observed and the statistics at high p_T is limited, we decide to perform the template fit in the full p_T range and use a global $V\gamma$ scale factor. Figure 7.22 shows the result of the fit.

The resulting scale factors for the $e\gamma$ channel are:

$$\begin{aligned} a_{V\gamma(e)} &= 1.17 \pm 0.23 \\ a_{fake(e)} &= 0.24 \pm 0.05 \end{aligned}$$

and for the $\mu\gamma$ channel are:

$$\begin{aligned} a_{V\gamma(\mu)} &= 1.33 \pm 0.26 \\ a_{fake(\mu)} &= 0.62 \pm 0.12 \end{aligned}$$

To check how the normalization factors evolve, the template fit is performed again in an intermediate p_T^{miss} region: [70-120] GeV, as shown in Figure 7.23. The resulting scale factors,

$$\begin{aligned} a'_{V\gamma(e)} &= 1.15 \pm 0.09(\text{stat. unc.}) \\ a'_{fake(e)} &= 0.26 \pm 0.05(\text{stat. unc.}) \\ a'_{V\gamma(\mu)} &= 1.20 \pm 0.16(\text{stat. unc.}) \\ a'_{fake(\mu)} &= 0.75 \pm 0.1(\text{stat. unc.}), \end{aligned}$$

agree well with the values derived in the low p_T^{miss} (40-70 GeV) region. Therefore, the normalization obtained in control region can be used to predict the background in signal region.

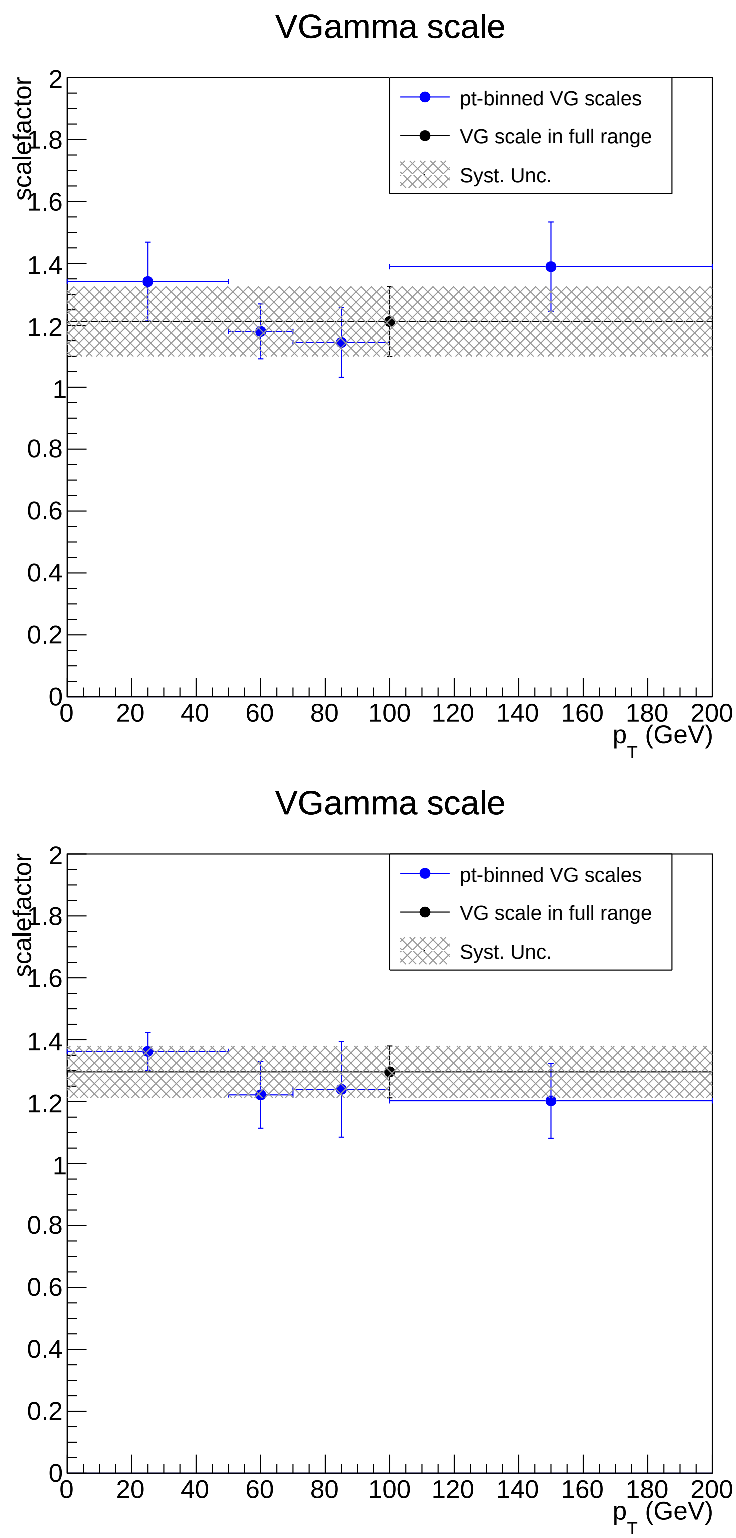


Figure 7.21: The $V\gamma$ scale factors as a function of lepton p_T. The blue dots denote the p_T binned scale factors, and the black dots denote the scale in full p_T range. Top: $e\gamma$ channel, bottom: $\mu\gamma$ channel.

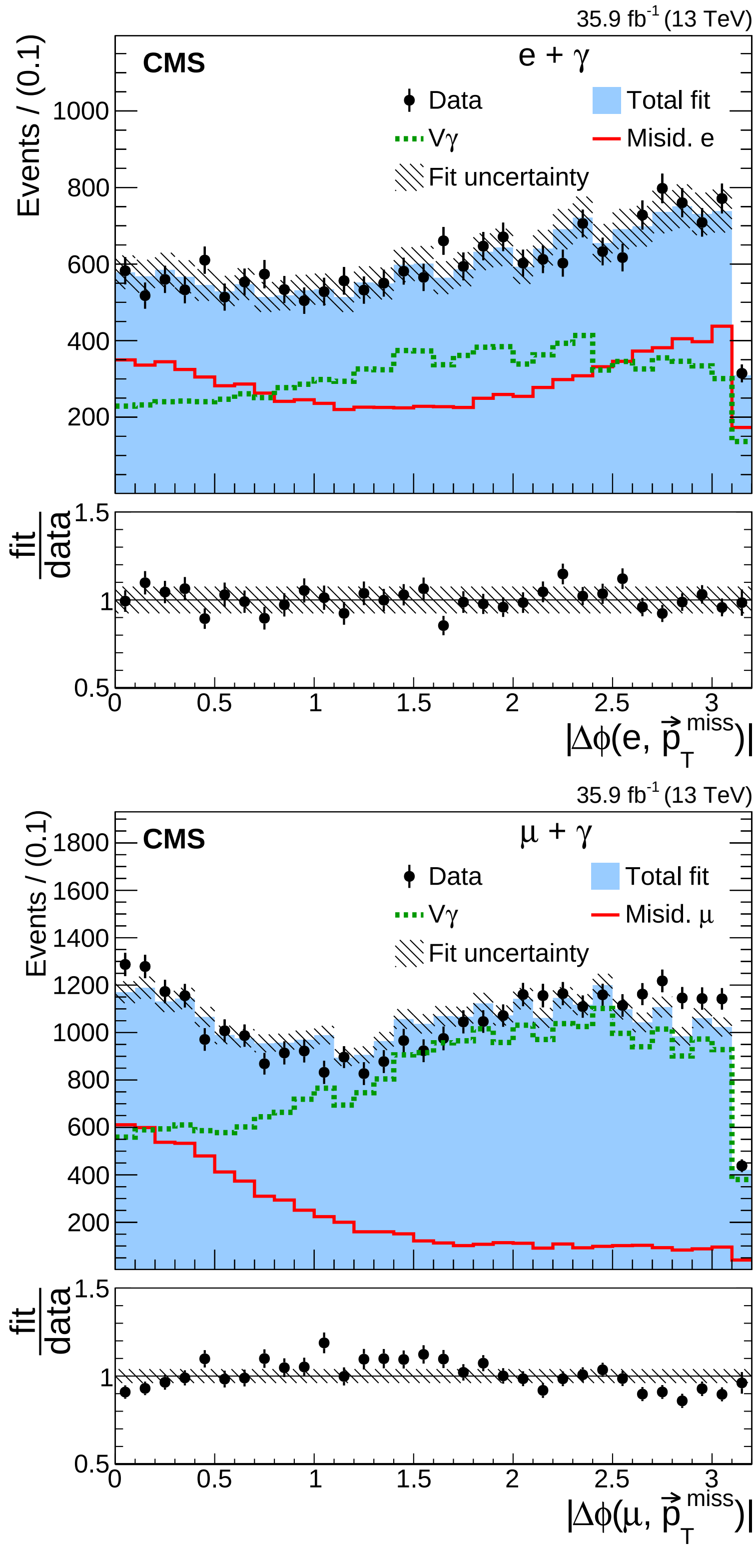


Figure 7.22: Results of the $\Delta\phi(\ell, p_T^{\mathrm{miss}})$ template fit in the $e\gamma$ channel (top) and $\mu\gamma$ channel (bottom).

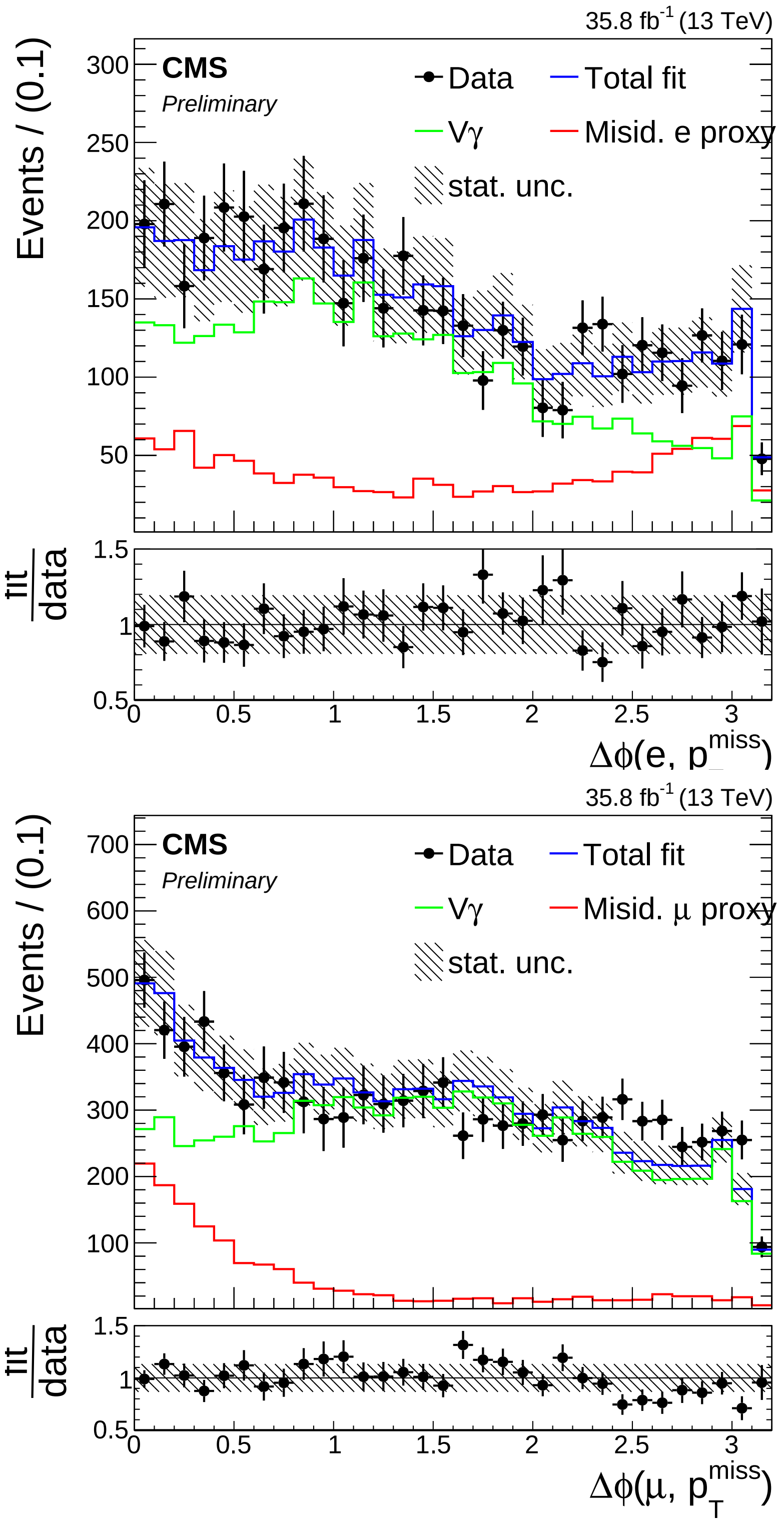


Figure 7.23: Validation of the $\Delta\phi(\ell, p_T^{\rm miss})$ template fit in the intermediate $p_T^{\rm miss}$ region for the $e\gamma$ channel (top) and $\mu\gamma$ channel (bottom).

The major systematic uncertainties of this estimation come from the shapes of the fitting target and data-to-simulation ESF applied to the Vγ samples. To evaluate the size of the uncertainty, we use toy MC to vary the Vγ distribution and the contribution from the fake photon background, and repeat the fit 1000 times. Figure 7.24 shows the distribution of the scale factors of these 1000 toy tests. This method revealed a 20% systematic uncertainty on the scale factor.

7.5 The Rare EWK Background

The final contribution to the background comes from EW processes with multiple bosons or top quarks associated with a photon, including $WW\gamma$, $WZ\gamma$ and $t\bar{t}\gamma$, referred to as rare EW processes. This background is estimated using simulated samples, scaled with cross-section at NLO precision. To evaluate the goodness of simulation, a dedicated rare EWK validation region is used to check the background estimation. The rare EWK validation region is defined as

- $M_T < 100$ GeV,
- has at least one b-jet.

The requirement of ≥ 1 b-jet enhanced the component of $t\bar{t}\gamma$ process, which is the major contribution of the rare EWK background. The predicted background is compared to the selected data in this region. As shown in Figure 7.25 and 7.26, good agreement is obtained between the data and estimation.

7.6 Systematic Uncertainties

There are several sources of systematic uncertainties that can affect the predicted backgrounds and signal expectation. Table 7.4 summarizes the sources and relative

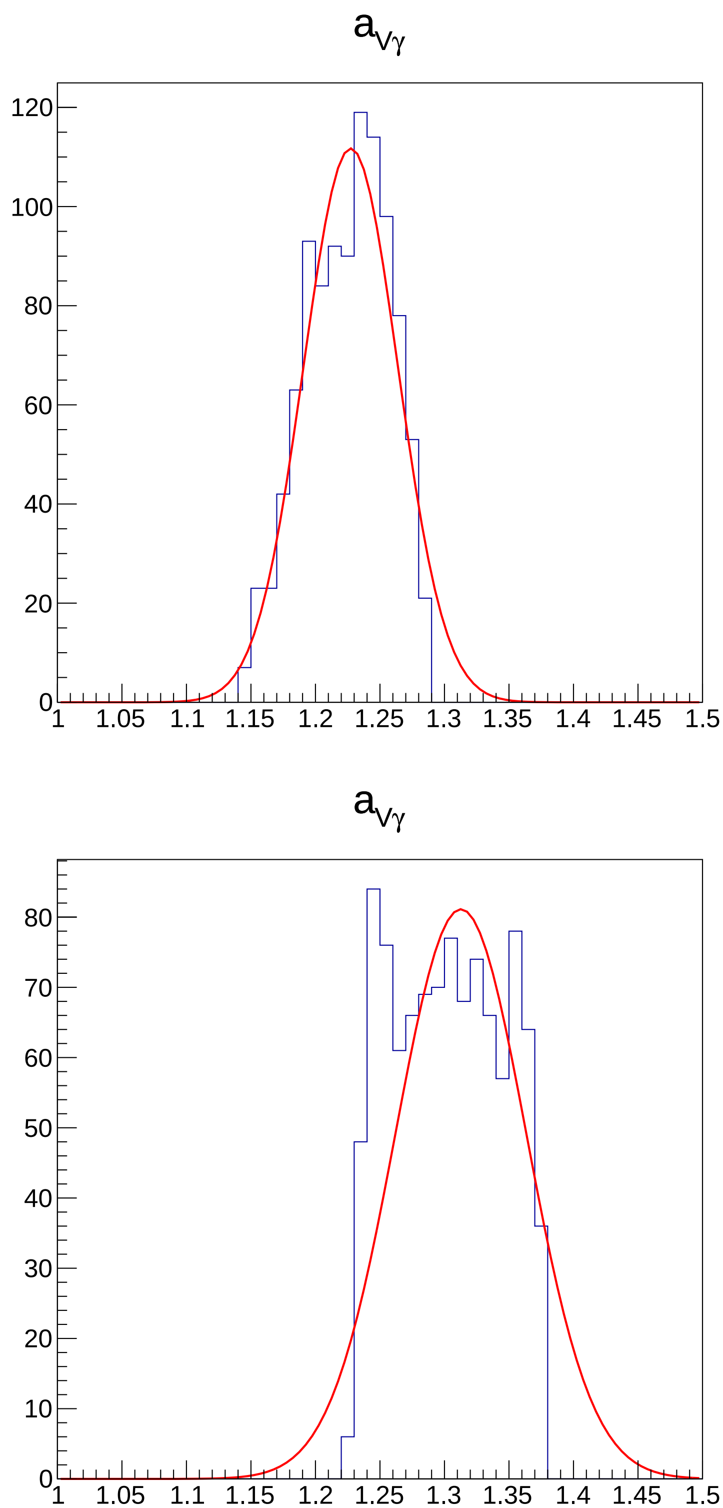


Figure 7.24: Distribution of the scale factors derived from 1000 toy MC experiments. Top: $e\gamma$ channel, bottom: $\mu\gamma$ channel.

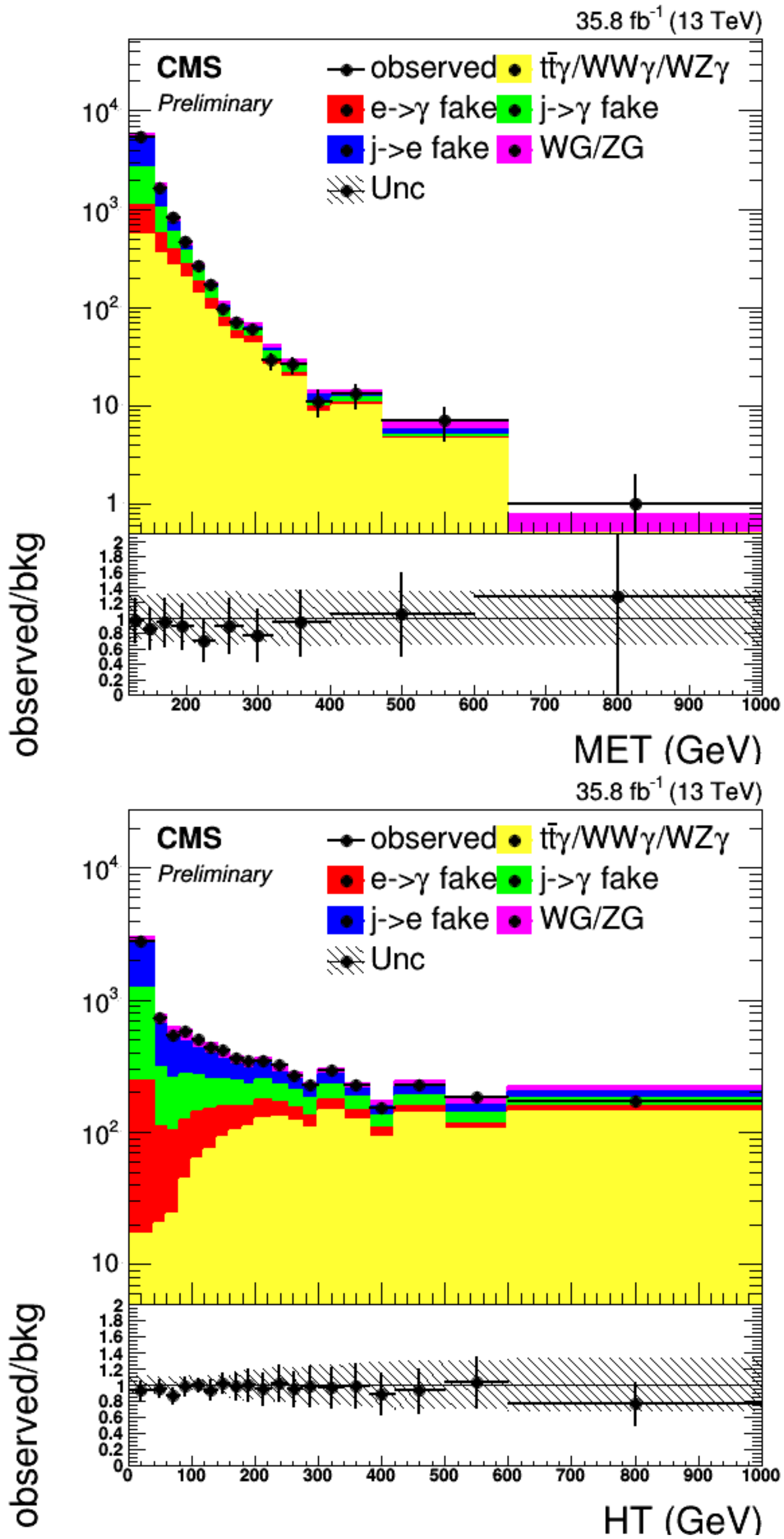


Figure 7.25: p_T^{miss} (top) and H_T (bottom) distributions of the $e\gamma$ channel in the validation region defined as ≥ 1 b-jet, $M_T(\ell, p_T^{\text{miss}}) < 100$ GeV.

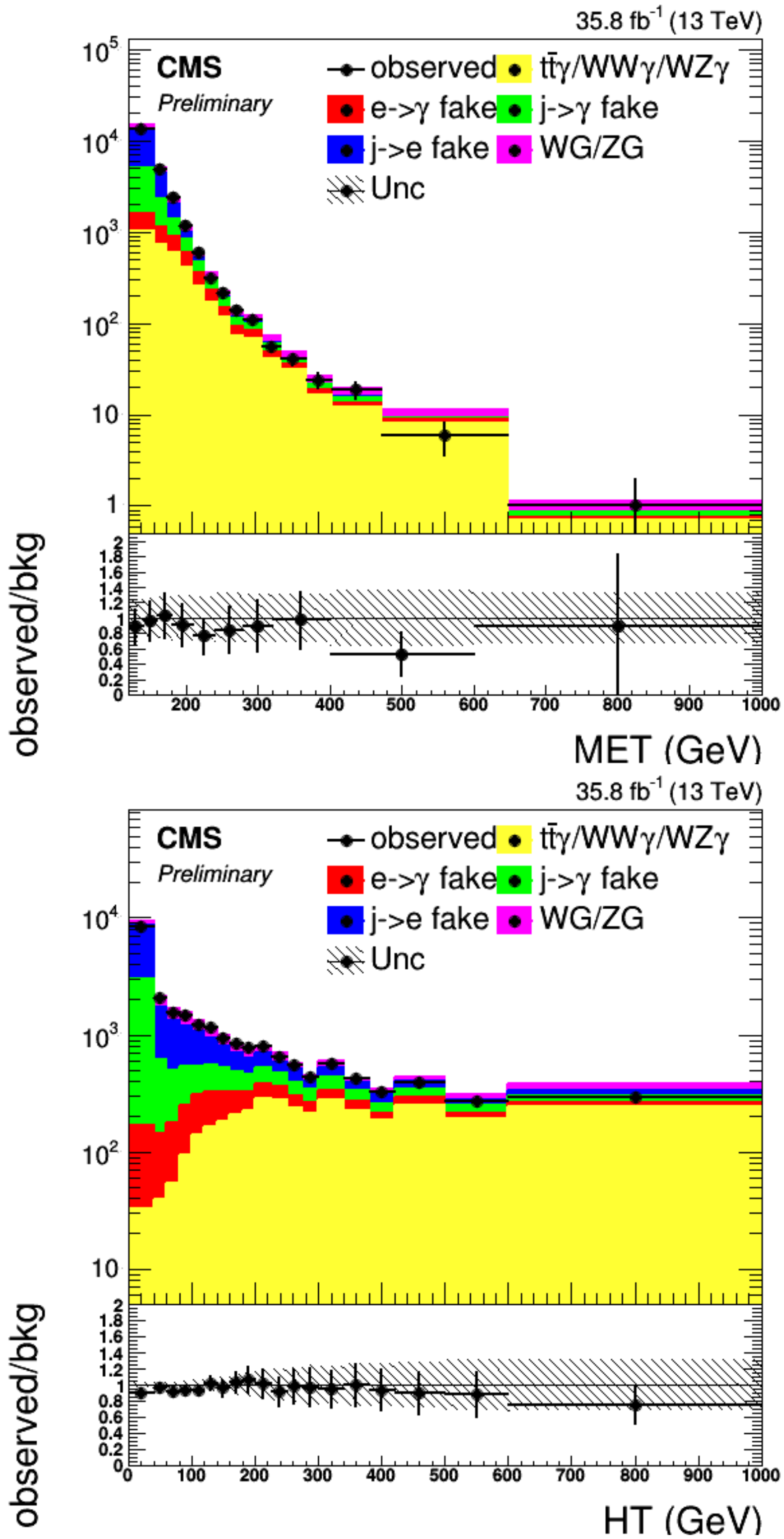


Figure 7.26: p_T^{miss} (top) and H_T (bottom) distributions of the $\mu\gamma$ channel in the validation region defined as ≥ 1 b-jet, $M_T(\ell, p_T^{\mathrm{miss}}) < 100$ GeV.

values of systematic uncertainties. If the relative uncertainties differ significantly in different kinematic regions because of limited number of events available in the control sample or simulation, the minimum and maximum values of the relative uncertainty is shown. The main sources of systematic uncertainties are the scale factors derived from the $\Delta\phi(\ell, p_T^{\text{miss}})$ template fit and the cross sections used to normalize the rare EWK simulated samples. The subdominant systematic uncertainties come from the ISR reweighting of the $V\gamma$ samples, and modeling of the misidentified photos. The systematic uncertainties considered for this analysis is described in the following.

Jet Energy Scale Potential discrepancies between the jet energy scale (JES) measured in data and in simulation should be considered as systematic uncertainties for MC samples. JES affects not only the variables of jets, such as H_T, but also the p_T^{miss} and M_T through the type-1 corrections described in Section 6.6. Since the search bins are defined in terms of H_T and p_T^{miss}, changes in these variables can induce a migration of events in the signal region. This uncertainty impacts the $V\gamma$ and rare EWK backgrounds as well as the SUSY signal yields. The JES uncertainty is evaluated by shifting the correction factors by $\pm1\sigma$ and recalculating all affected variables. For each search bin, the relative change in event yields due to the $\pm1\sigma$ shift in JES is used as the systematic uncertainty. JES uncertainties in all search bins and processes are fully correlated.

Efficiency Scale Factors Efficiency scale factors (ESF) for object identification and trigger are applied to account for the difference between data and simulation, as described in Section 5.2 and 6. To assess the effect of ESF uncertainties, the scale factors assigned to each simulated event are scaled by $\pm1\sigma$. The relative variation of event yields in each search bin is assigned as the ESF systematic uncertainty.

Photon Misidentification Rate For the fake photon backgrounds, systematic uncertainties are obtained using toy MC methods, as described in Section 7.1 and 7.2. For each search bin, the 1σ variation of predicted background due to the changes in photon misidentification rate is used as the systematic uncertainty. The number of events of the fake object control samples is limited, especially in the high p_T^{γ} and high p_T^{miss} bins which contains very few fake object backgrounds. Therefore we use gamma distribution and estimated transfer factors to model the statistical uncertainties of these backgrounds. This allows us to assign uncertainties on the bins which have zero event in the control sample.

ISR Corrections Uncertainties caused by the ISR reweighting are considered. We use the corrected distributions as the central values for the $V\gamma$ background, and take the full corrections as the systematic uncertainties. For the SUSY signal samples, the same set of ISR weights are assigned to the sum of transverse momentum of the initial SUSY particle system. The uncorrected distributions are used as the nominal expectations for the SUSY signals, and the difference between corrected and uncorrected event yields are taken as systematic uncertainties.

$V\gamma$ **normalization scale** The uncertainty of the normalization scale factors derived from the $\Delta\phi(\ell, p_T^{\text{miss}})$ template fit is one of the major uncertainties of this analysis. As described in Section 7.4, the uncertainty in normalization scales due to the $V\gamma$ shape is obtained by varying the ESF, PDF and renormalization scales of the $V\gamma$ template. Uncertainties on the normalization magnitude are evaluated by shifting the number of events subtracted from the fit target by 1σ. An overall 20% uncertainty is assigned to the normalization scale factors, and this uncertainty is anti-correlated between the $V\gamma$ and misidentified lepton backgrounds.

Rare EWK Cross-section For the rare background, a 50% uncertainty on the cross-section is assumed, covering the difference between calculated cross-sections and latest CMS measurements [66, 67]. This is one of the major uncertainties of this analysis.

Integrated Luminosity Follow the recommendation of the luminosity study group, we assign a 2.5% uncertainty on the integrated luminosity [68]. This uncertainty is applied to the rare EWK background and SUSY signal yields. For the $V\gamma$ sample, uncertainty caused by luminosity has been absorbed to the normalization scale factors.

PDF and Renormalization Scale The uncertainties of the background estimation due to the choices of PDF and renormalization and factorization scales have been absorbed to the normalization scale factors of the $V\gamma$ and rare EWK background. However, for the SUSY signal samples, additional uncertainties should be considered. The main effect of PDFs is on the cross-section of the SUSY signal, and its remaining effect on the acceptance of SUSY events is very small. Therefore, only the variations of acceptance caused by different renormalization scales are considered. To estimate this uncertainty, renormalization scales are shifted upward and downward by a factor of two with respect to their nominal values and the corresponding change in each search bin is taken as the systematic uncertainty.

FastSim $p_T^{\mathbf{miss}}$ The SUSY samples are simulated using the FastSim tools of CMS. To access the uncertainty due to potential mis-modeling of the p_T^{miss} shape, the analysis is performed using both the PF-p_T^{miss} and generator level p_T^{miss}. The uncertainty is obtained by taking one-half the difference between the acceptance of these two methods.

7.7 Validation of the Background Prediction

The $M_T(\ell, p_T^{\text{miss}}) < 100$ GeV region is dominated by background events. Studies with TChiWG and T5WG samples show that this region has negligible signal contamination. Therefore, we propose to use $M_T < 100$ GeV as a validation region to test the background prediction, especially the modeling of p_T^{miss} distribution. Figure 7.27,7.28, and 7.29 show the comparison between the data and background prediction in the validation region. A good agreement is obtained in both channels.

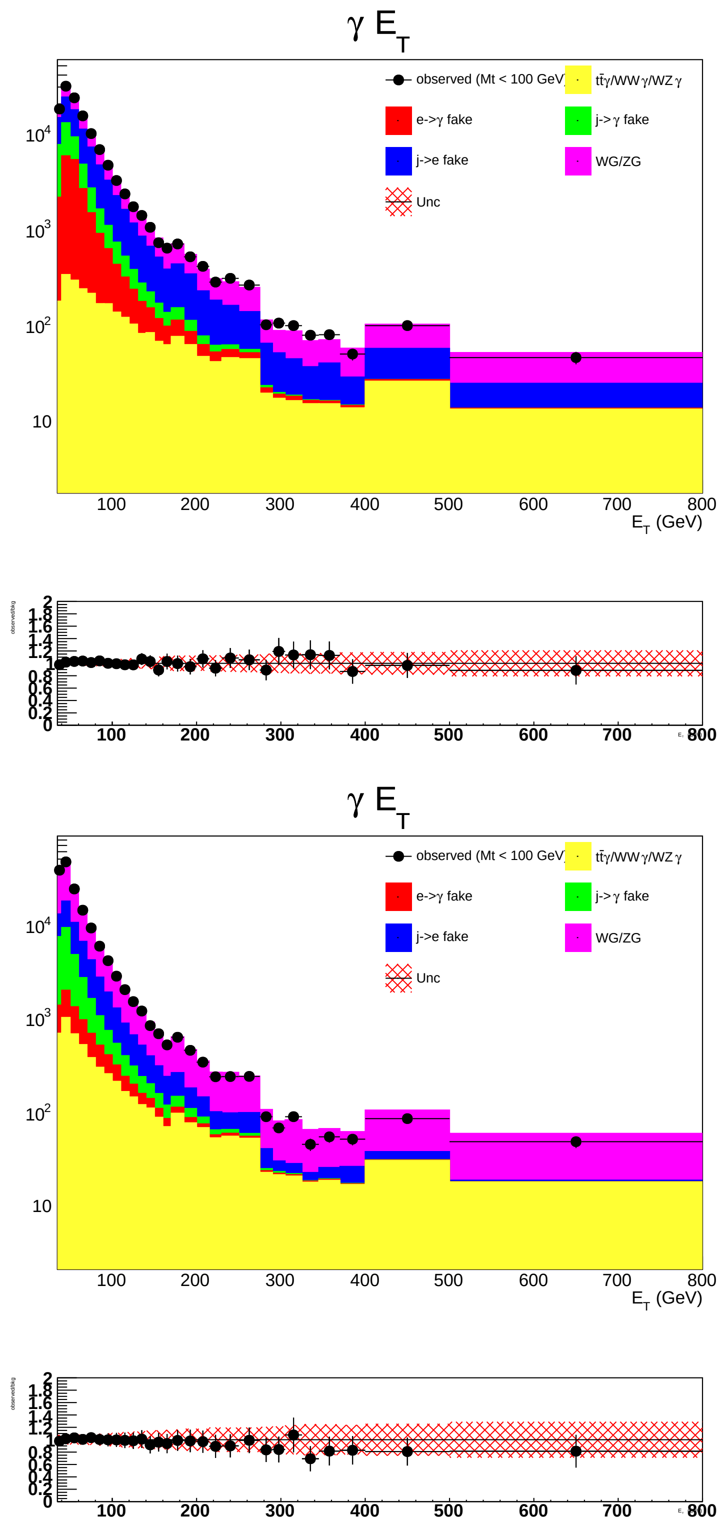


Figure 7.27: Photon p_T distributions in the $e\gamma$ channel (top) and the $\mu\gamma$ channel (bottom).

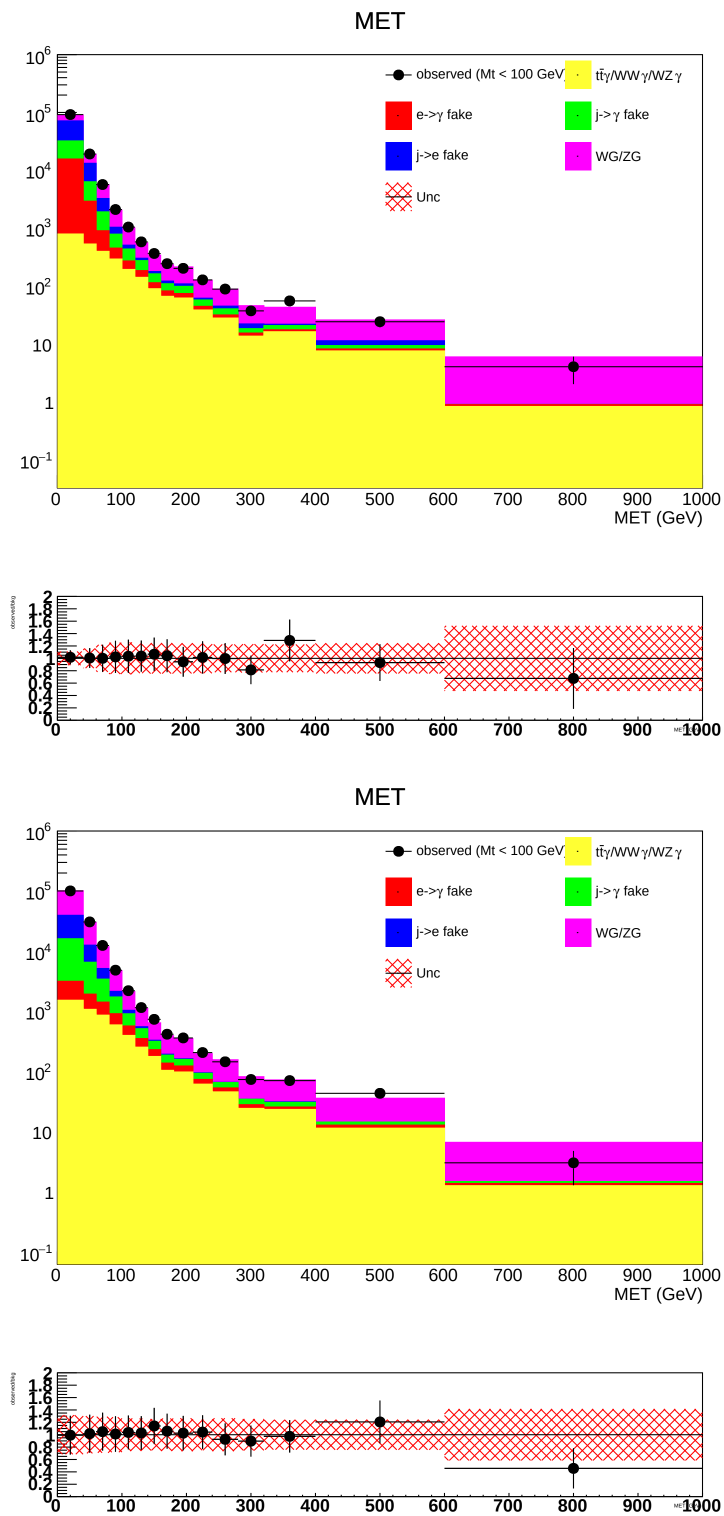


Figure 7.28: p_T^{miss} distributions in the $e\gamma$ channel (top) and the $\mu\gamma$ channel (bottom).

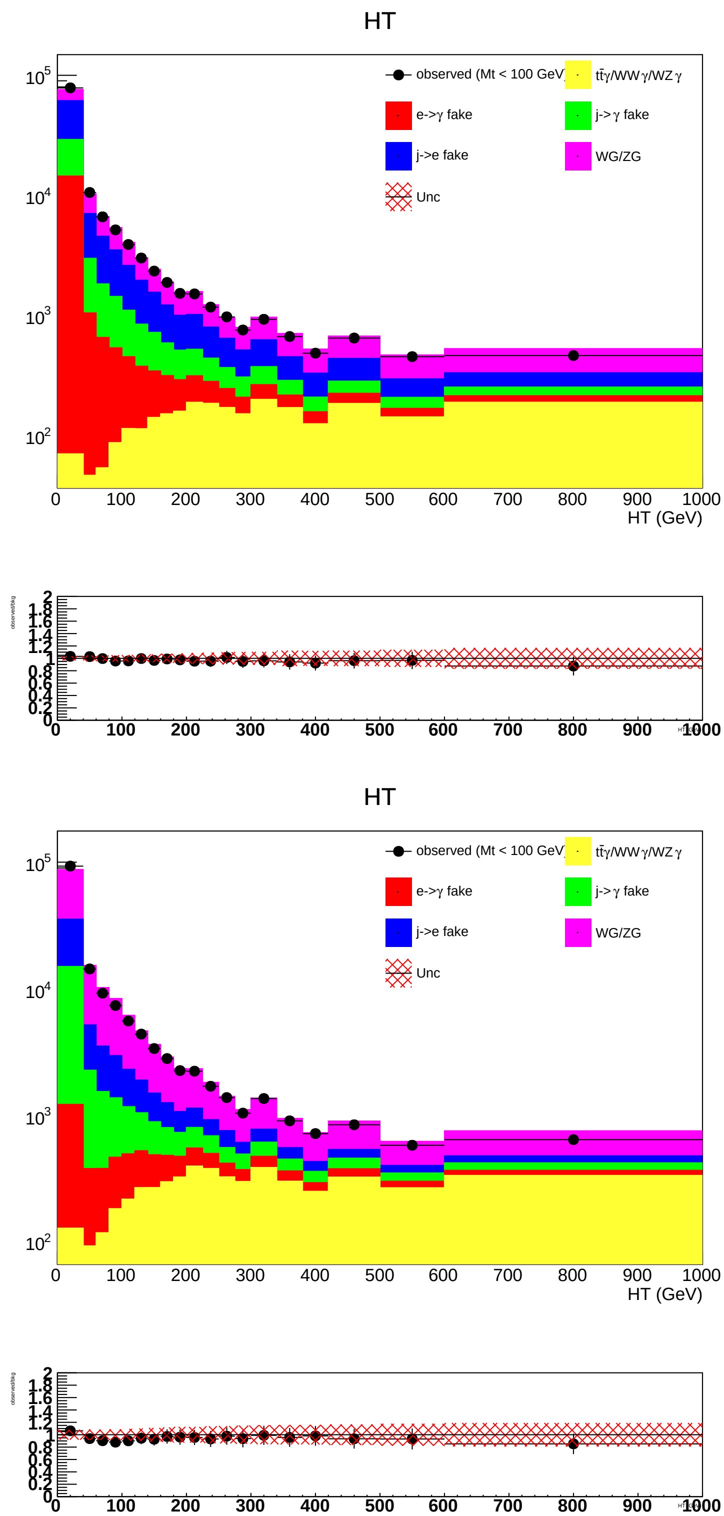


Figure 7.29: HT distributions in the $e\gamma$ channel (top) and the $\mu\gamma$ channel (bottom).

Table 7.4: Summary of the systematic uncertainties affecting the SUSY signals and SM background estimates, given in percent.

Source of uncertainty	Process	Relative uncertainty (%)
Jet energy scale	Vγ, rare EW	0 – 22.5
Normalization scale	Vγ, jet$\rightarrow \ell$ misid.	20
Cross section	rare EW	50
Ident. and trigger efficiency	Vγ, rare EW	1.3 – 6.5
e $\rightarrow \gamma$	e $\rightarrow \gamma$ misid.	8.0 – 50.5
Jet$\rightarrow \gamma$ shape	jet$\rightarrow \gamma$ misid.	8.1 – 56.1
Misid. lepton shape	jet$\rightarrow \ell$ misid.	0 – 42.4
ISR corrections	Vγ	2.6 – 57.8
Integrated luminosity	rare EW	2.5
Cross section, PDF	SUSY signal	4.3 – 36.8
Jet energy scale	SUSY signal	0 – 10
Ident. and trigger efficiency	SUSY signal	4
Pileup uncertainty	SUSY signal	2 – 10
Integrated luminosity	SUSY signal	2.5
ISR corrections	SUSY signal	0 – 31.8
Renormalization/factorization scales	SUSY signal	0 – 10.2
Fast simulation p_T^{miss} modelling	SUSY signal	0 – 30.5

Chapter 8

Results and Interpretations

8.1 Results

The background estimation is performed in a $p_T^{\text{miss}} < 70$ GeV control region and validated in the $M_T(\ell, p_T^{\text{miss}}) < 100$ GeV validation region. The high p_T^{miss} signal region was kept blinded until all methods were established.

The signal region is defined as $p_T^{\text{miss}} > 120$ GeV and $M_T(\ell, p_T^{\text{miss}}) > 100$ GeV. To improve the sensitivity to SUSY signals, the signal region is divided into p_T^{miss}, $\mathrm{H_T}$ and photon p_T bins. The binning is optimized using T5Wg expected limits, as shown in Figure 8.1. For the optimization, three bins in p_T^{miss} (120-200,200-400,$>$ 400 GeV) and three bins in $\mathrm{H_T}$ (0-100, 100-400, $>$ 400 GeV) are used as the baseline binning. The performance of other binning strategies, such as adding p_T^{γ} bins, are compared to the baseline binning. By choosing the bins that give the optimal exclusion limit, we decide to use 18 bins for each channel: (120 GeV $< p_T^{\text{miss}} <$ 200 GeV, 200 GeV $< p_T^{\text{miss}} <$ 400 GeV, $p_T^{\text{miss}} >$ 400 GeV) $\times$ (0 $< \mathrm{H_T} <$ 100 GeV, 100 GeV $< \mathrm{H_T} <$ 400 GeV, $\mathrm{H_T} >$ 400 GeV) $\times$ (35 GeV $< \gamma p_T <$ 200 GeV, $\gamma p_T >$ 200 GeV).

Figure 8.2 shows the unblinded data distributions of the photon p_T^{γ}, p_T^{miss} and

$\mathrm{H_T}$, along with the background predictions. Two signal distributions, one from the TChiWg with $M_{\tilde{\chi}_1^{\pm}}$ = 800 GeV and the other one from the T5Wg with $M_{\tilde{g}}$ =1700 GeV, $M_{\tilde{\chi}_1^{\pm}}$ = 1000 GeV, are also overlayed on the plots. The event yields and SM backgrounds for each search bin are shown in Figure 8.3, and the corresponding numbers are given in Table 8.1.

The data are consistent with the estimated SM backgrounds within the uncertainties through the search region. The 22^{nd} and 36^{th} bins, which are both in $e\gamma$ channel, show excesses above prediction with local significances of 2.3 and 1.2 standard deviations, respectively. In the corresponding regions of the $\mu\gamma$ channel, the data agree with the SM backgrounds. Thus, we conclude that no significant excess of events above the SM expectation is observed.

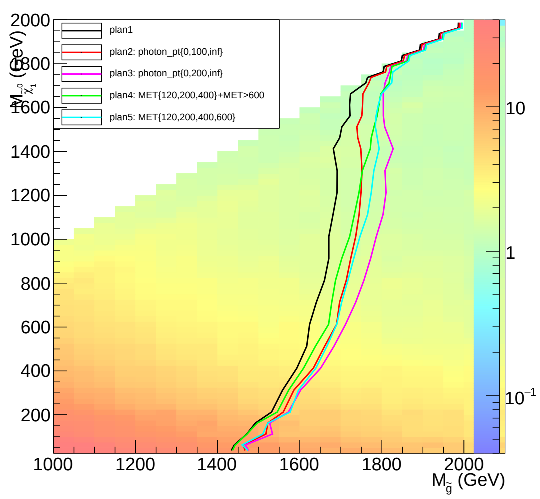


Figure 8.1: Optimization of the signal bins. Baseline binning (black): p_T^{miss}\{120-200, 200-400, > 400\}GeV × $\mathrm{H_T}$\{ <100, 100-400, > 400\}GeV. Plan 2 (red): p_T^{miss}\{120-200, 200-400, > 400\}GeV × $\mathrm{H_T}$\{ <100, 100-400, > 400\}GeV × p_T^{γ} \{< 100, > 100\}GeV. Plan 3 (magenta): p_T^{miss}\{120-200, 200-400, > 400\}GeV × $\mathrm{H_T}$\{ <100, 100-400, > 400\}GeV × p_T^{γ} \{< 200, > 200\}GeV. Plan 4 (green): p_T^{miss}\{120-200, 200-400, 400-600\}GeV × $\mathrm{H_T}$\{ <100, 100-400, > 400\}GeV. + \{p_T^{miss} > 600 GeV\}. Plan 5 (cyan): p_T^{miss}\{120-200, 200-400, 400-600\}GeV × $\mathrm{H_T}$\{ <100, 100-400, > 400\}GeV. + \{p_T^{miss} > 600 GeV\}× $\mathrm{H_T}$\{ < 400, > 400\}GeV.

Bin	channel	p_T^{miss}	$\mathrm{H_T}$	p_T^{γ}	$e \to \gamma$ fakes	jet$\to \gamma$ fakes	jet$\to l$ fakes	$W\gamma/Z\gamma$	rare	SM background	Data
1	$\mu\gamma$	120-200	<100	<200	46.920 ± 3.905	44.067 ± 4.611	3.100 ± 1.519	183.161 ± 45.284	37.238 ± 19.051	314.486 ± 49.522	309
2	$\mu\gamma$	120-200	100-400	<200	140.845 ±11.790	52.269 ± 5.214	7.440 ± 2.613	85.740 ± 27.722	176.442 ± 89.401	462.736 ± 94.661	494
3	$\mu\gamma$	120-200	>400	<200	21.039 ± 2.085	12.701 ± 1.808	2.480 ± 1.336	11.665 ± 4.425	50.288 ± 25.599	98.173 ± 26.409	85
4	$\mu\gamma$	200-400	<100	<200	2.043 ± 0.279	3.217 ± 0.780	0.620 ± 0.632	17.609 ± 4.476	2.837 ± 1.603	26.326 ± 4.867	32
5	$\mu\gamma$	200-400	100-400	<200	17.600 ± 1.654	6.547 ± 1.193	1.240 ± 0.911	8.522 ± 3.794	26.959 ± 13.802	60.869 ± 14.502	64
6	$\mu\gamma$	200-400	>400	<200	9.351 ± 1.028	4.561 ± 0.987	0.620 ± 0.632	5.246 ± 2.922	25.095 ± 12.899	44.873 ± 13.325	45
7	$\mu\gamma$	>400	<100	<200	0.116 ± 0.055	0.108 ± 0.113	0.000 ± 0.000	0.931 ± 0.217	0.031 ± 0.023	1.187 ± 0.252	1
8	$\mu\gamma$	>400	100-400	<200	0.108 ± 0.043	0.000 ± 0.000	1.240 ± 0.912	0.169 ± 0.100	1.352 ± 0.997	2.869 ± 1.355	1
9	$\mu\gamma$	>400	>400	<200	0.686 ± 0.133	0.000 ± 0.000	0.620 ± 0.632	0.598 ± 0.350	3.358 ± 1.897	5.262 ± 2.034	5
10	$\mu\gamma$	120-200	<100	>200	0.248 ± 0.096	0.096 ± 0.090	0.000 ± 0.000	4.294 ± 2.207	1.673 ± 0.868	6.312 ± 2.374	12
11	$\mu\gamma$	120-200	100-400	>200	0.715 ± 0.243	0.270 ± 0.192	0.000 ± 0.000	6.938 ± 2.776	13.164 ± 6.688	21.086 ± 7.219	23
12	$\mu\gamma$	120-200	>400	>200	0.580 ± 0.209	0.265 ± 0.182	0.000 ± 0.000	6.255 ± 2.378	8.149 ± 4.251	15.248 ± 4.885	20
13	$\mu\gamma$	200-400	<100	>200	0.081 ± 0.037	0.438 ± 0.258	0.000 ± 0.000	3.398 ± 1.736	0.916 ± 0.473	4.833 ± 1.818	4
14	$\mu\gamma$	200-400	100-400	>200	0.142 ± 0.061	0.223 ± 0.161	0.000 ± 0.000	1.948 ± 1.004	5.952 ± 3.017	8.265 ± 3.185	12
15	$\mu\gamma$	200-400	>400	>200	0.302 ± 0.115	0.118 ± 0.103	0.000 ± 0.000	1.607 ± 0.889	3.347 ± 1.736	5.374 ± 1.957	7
16	$\mu\gamma$	>400	<100	>200	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.679 ± 0.398	0.051 ± 0.036	0.730 ± 0.397	1
17	$\mu\gamma$	>400	100-400	>200	0.010 ± 0.008	0.057 ± 0.064	0.000 ± 0.000	0.208 ± 0.126	0.348 ± 0.185	0.623 ± 0.234	1
18	$\mu\gamma$	>400	>400	>200	0.025 ± 0.015	0.000 ± 0.000	0.000 ± 0.000	0.178 ± 0.113	0.326 ± 0.174	0.529 ± 0.209	0
19	$e\gamma$	120-200	<100	<200	30.691 ± 2.748	23.562 ± 2.985	17.260 ± 4.153	72.967 ± 10.793	19.145 ± 9.884	163.626 ± 15.745	153
20	$e\gamma$	120-200	100-400	<200	79.791 ± 7.376	33.203 ± 3.721	14.430 ± 4.280	35.227 ± 12.349	94.304 ± 47.654	256.955 ± 50.100	277
21	$e\gamma$	120-200	>400	<200	13.083 ± 1.442	10.713 ± 1.785	5.778 ± 1.880	12.160 ± 5.883	37.216 ± 18.867	78.950 ± 19.984	67
22	$e\gamma$	200-400	<100	<200	2.485 ± 0.326	1.478 ± 0.590	2.314 ± 1.340	8.903 ± 1.876	2.129 ± 1.154	17.309 ± 2.664	32
23	$e\gamma$	200-400	100-400	<200	11.659 ± 1.149	8.026 ± 1.514	2.298 ± 1.023	7.126 ± 2.859	20.846 ± 10.636	49.956 ± 11.223	46
24	$e\gamma$	200-400	>400	<200	6.479 ± 0.787	2.465 ± 0.772	0.336 ± 0.365	2.511 ± 1.375	16.372 ± 8.417	28.164 ± 8.607	32
25	$e\gamma$	>400	<100	<200	0.097 ± 0.043	0.184 ± 0.185	0.000 ± 0.000	0.997 ± 0.190	0.023 ± 0.018	1.302 ± 0.269	1
26	$e\gamma$	>400	100-400	<200	0.160 ± 0.067	0.389 ± 0.279	0.336 ± 0.352	0.111 ± 0.053	0.165 ± 0.091	1.160 ± 0.466	1
27	$e\gamma$	>400	>400	<200	0.635 ± 0.132	0.490 ± 0.308	0.000 ± 0.000	0.682 ± 0.410	0.997 ± 0.545	2.804 ± 0.759	4
28	$e\gamma$	120-200	<100	>200	0.262 ± 0.100	0.232 ± 0.184	0.412 ± 0.450	3.376 ± 1.581	1.645 ± 1.001	5.927 ± 1.936	10
29	$e\gamma$	120-200	100-400	>200	0.644 ± 0.221	1.393 ± 0.750	0.977 ± 0.557	5.808 ± 2.266	12.792 ± 6.472	21.614 ± 6.924	21
30	$e\gamma$	120-200	>400	>200	0.678 ± 0.239	0.116 ± 0.132	0.535 ± 0.422	3.979 ± 1.501	6.368 ± 3.286	11.677 ± 3.647	14
31	$e\gamma$	200-400	<100	>200	0.113 ± 0.057	0.347 ± 0.243	0.000 ± 0.000	2.320 ± 1.142	1.355 ± 0.808	4.135 ± 1.421	6
32	$e\gamma$	200-400	100-400	>200	0.140 ± 0.056	0.467 ± 0.317	0.187 ± 0.195	1.953 ± 0.764	6.129 ± 3.115	8.876 ± 3.230	9
33	$e\gamma$	200-400	>400	>200	0.317 ± 0.120	0.116 ± 0.131	0.187 ± 0.195	1.237 ± 0.632	3.266 ± 1.659	5.123 ± 1.795	4
34	$e\gamma$	>400	<100	>200	0.009 ± 0.008	0.000 ± 0.000	0.000 ± 0.000	0.341 ± 0.191	0.044 ± 0.031	0.394 ± 0.194	0
35	$e\gamma$	>400	100-400	>200	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.219 ± 0.131	0.312 ± 0.169	0.532 ± 0.213	1
36	$e\gamma$	>400	>400	>200	0.030 ± 0.018	0.000 ± 0.000	0.000 ± 0.000	0.157 ± 0.096	0.660 ± 0.493	0.848 ± 0.502	3

Table 8.1: Observed number of events and predicted SM backgrounds in search bins. The systematic uncertainties of each background components are added in quadrature.

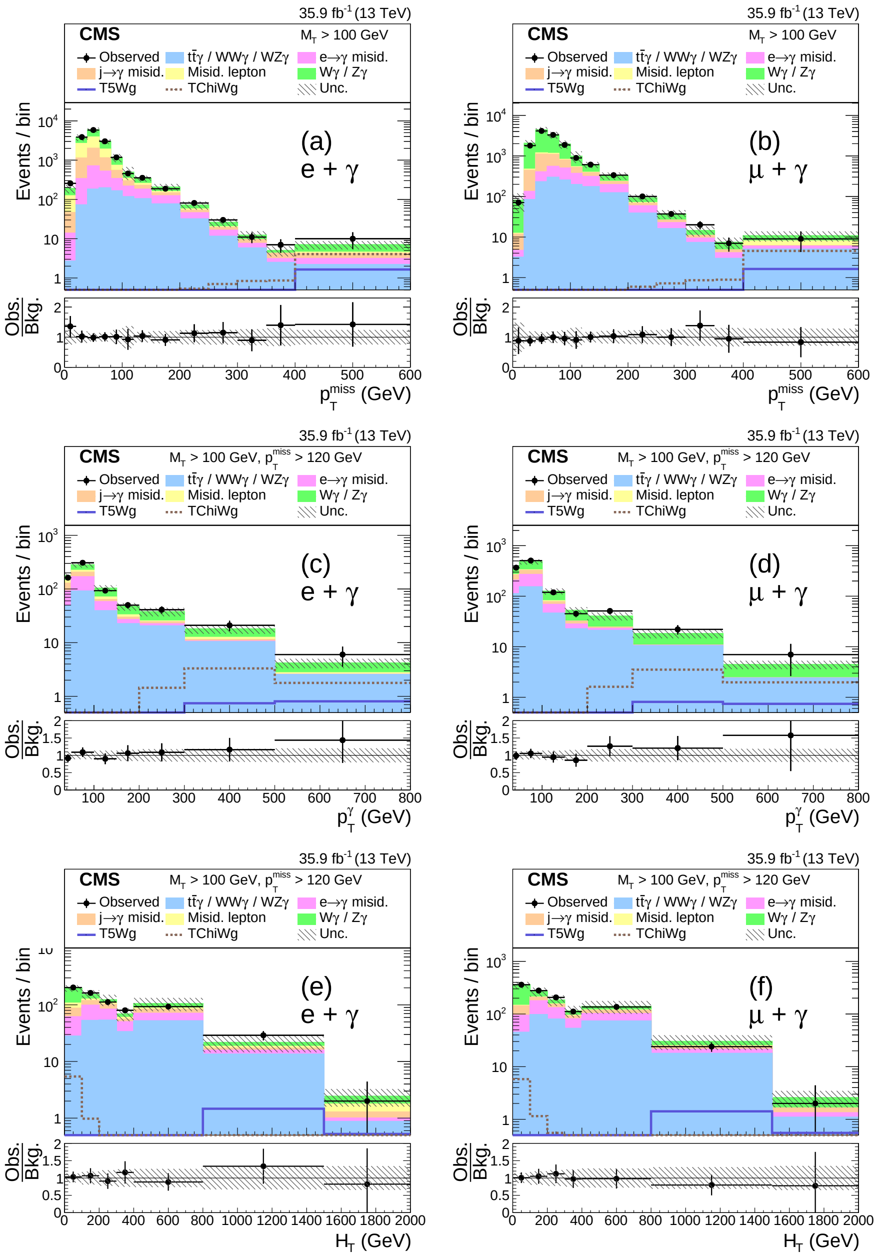


Figure 8.2: Distributions of p_T^{miss}(a, b), p_T^{γ} (c, d), and $\mathrm{H_T}$ (e, f) from data (points) and estimated SM predictions (stacked histograms) for the $e\gamma$ (left) and $\mu\gamma$ (right) channels.

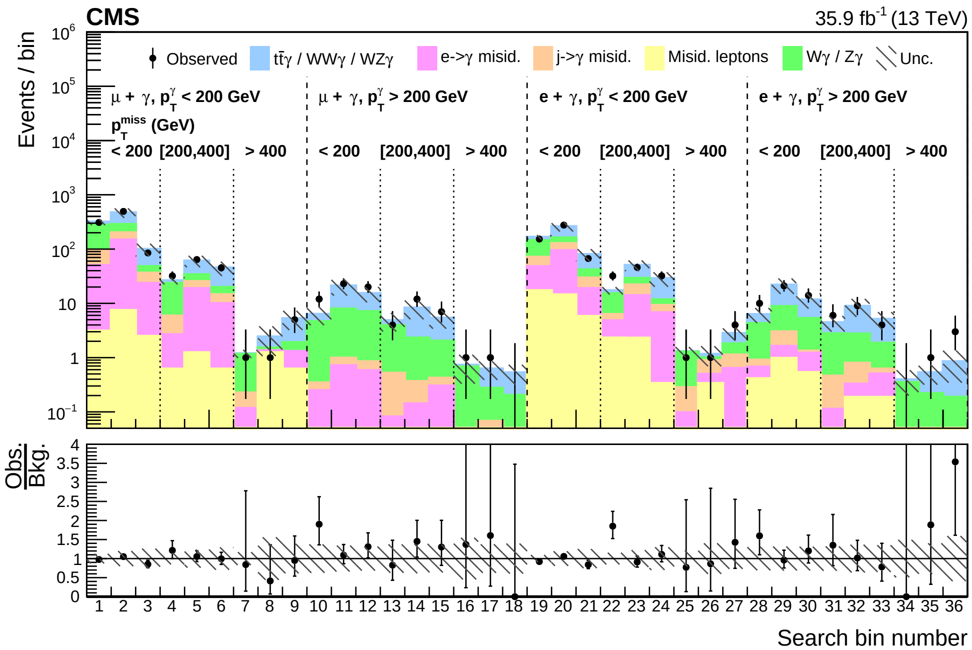


Figure 8.3: Event yields and stacked background predictions as a function of search bin numbers.

8.2 Limit setting procedure

The results are interpreted in terms of upper limits on the cross-sections of SUSY models. The 95% confidence level (CL) upper limits are obtained using a modified frequentist (CL_s) method [69, 70, 71]. In this section, the expected SUSY event yields is denoted as s, while the number of SM background events is denoted as b. A variable known as the signal strength modifier, denoted as μ, is introduced to scale the yields of the SUSY process, so that the total number of events is modified as $\mu \cdot s + b$. $\mu = 0$ and b events corresponds to the background-only hypothesis, and $\mu = 1$ and $s + b$ events corresponds to the signal+background hypothesis. Uncertainties of the backgrounds and signal yields are handled by introducing a set of nuisance parameters θ. The expected number of events in the ith bin become functions of the nuisance parameters:

$$n_i = \mu \cdot s_i(\theta) + b_i(\theta). \tag{8.1}$$

A likelihood function is then constructed as:

$$\begin{aligned} \mathcal{L}(\text{data}|\mu,\theta) &= \text{Poisson}(\text{data}|\mu\cdot s(\theta)+b(\theta))\cdot p(\tilde{\theta}|\theta) \\ &= \prod_{i=1}^{N}\frac{(\mu\cdot s_i+b_i)^{n_i}}{n_i!}e^{-(\mu\cdot s_i+b_i)}\cdot p(\tilde{\theta}|\theta), \end{aligned}, \tag{8.2}$$

where $p(\theta|\tilde{\theta})$ is the PDFs of the uncertainties and $\tilde{\theta}$ denotes the default values of the nuisance parameters.

To test the background-only and signal+background hypotheses, a test statistic $\tilde{q}_\mu$ based on the profile likelihood ratio is constructed as:

$$\tilde{q}_\mu = -2ln\frac{\mathcal{L}(\text{data}|\mu,\hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu},\hat{\theta})}, \text{with } 0\leq\hat{\mu}\leq\mu, \tag{8.3}$$

where $\hat{\mu}$ and $\hat{\theta}$ refer to the global maximum of the likelihood, and $\hat{\theta}_\mu$ refers to the conditional maximum estimator of θ given μ and "data". In this analysis, the asymptotic form of the profiled likelihood [72] is used to speed up the computations.

The next step of the procedure is to construct the distributions of the test statistic associated with the signal+background and background-only hypotheses:

$$f(\tilde{q}_\mu|\mu,\hat{\theta}_\mu), f(\tilde{q}_\mu|0,\hat{\theta}_0). \tag{8.4}$$

Then the p value for the signal+background hypothesis can be constructed as:

$$p_\mu = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs})|\text{signal+background}) = \int_{\tilde{q}_\mu^{obs}}^{\infty} f(\tilde{q}_\mu|\mu,\hat{\theta}_\mu)d\tilde{q}_\mu, \tag{8.5}$$

and that for the background-only hypothesis can be defined as:

$$1-p_b = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs})|\text{background-only}) = \int_{\tilde{q}_\mu^{obs}}^{\infty} f(\tilde{q}_\mu|0,\hat{\theta}_0)d\tilde{q}_\mu. \tag{8.6}$$

The ratio of these two p values is defined to be the CL$_s$, as:

$$\text{CL}_\mu(\mu) = \frac{p_\mu}{1-p_b}. \tag{8.7}$$

If, for $\mu = 1$, CL$_s \leq 0.05$, we state that the SUSY model considered is excluded at 95% CL.

8.3 Interpretations

The results are interpreted in terms of 95% CL upper limits on the cross-sections of three different simplified models: TChiWg, T5Wg and T6Wg, and the GGM model, which are described in Chapter 2.

The TChiWg model is initiated by the direct pair production of NLSP particles $\tilde{\chi}_1^0$ and $\tilde{\chi}_1^{\pm}$, which decay to $W^{\pm}\tilde{G}$ and $\gamma\tilde{G}$, respectively. The masses of $\tilde{\chi}_1^0$ and $\tilde{\chi}_1^{\pm}$ are taken to be degenerate and are scanned in the range 300-1200 GeV in the TChiWg model. Figure 8.4 shows the limits for the TChiWg models as a function of the NLSP masses, along with the theoretical cross-section of the $\tilde{\chi}_1^0/\tilde{\chi}_1^{\pm}$ pair production. The crossing point between the theoretical cross-section and the observed limits denotes the excluded mass. This search excludes NLSP masses up to 900 GeV, extending the current best limit by about 120 GeV [73].

The T5Wg (T6Wg) model assumes the strong production of $\tilde{g}$ ($\tilde{q}$) pairs, which decay to $\tilde{\chi}_1^0$ or $\tilde{\chi}_1^{\pm}$ with a branching fraction of 50%. Figure 8.5 shows the upper limits on the cross-section $\times$ 50% branching fraction for the T5Wg (left) and T6Wg (right) models. The black and red curves in Figure 8.5 denote the observed and expected exclusion contours. For both the T5Wg and T6Wg models, the observed limits are slightly lower than the expected limits due to the excess of data in several search bins. Since the excess is insignificant, the observed limits still overlap with the expected limits within uncertainties. This search excludes the gluino (squark) mass up to 1700 (1400) GeV in the T5Wg (T6Wg) scenarios.

Figure 8.6 shows the exclusion limits for the GGM model in terms of the model parameters: M1 and M2. For M1 = 1500 GeV, the model points below M2 $<$ 1200 GeV are exluded, which corresponds to the production of $\tilde{\chi}_1^0$ with mass around 700 GeV and $\tilde{\chi}_1^{\pm}$ with mass around 1100 GeV.

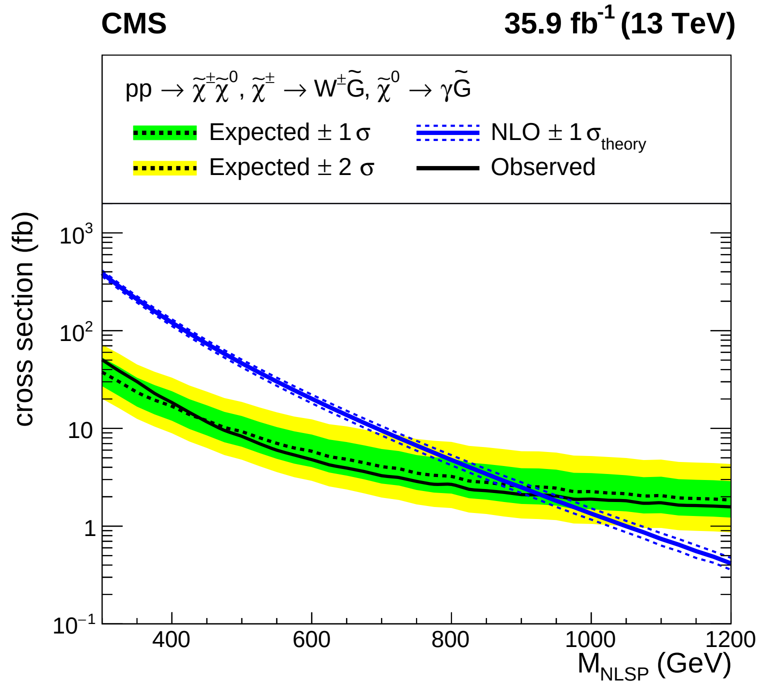


Figure 8.4: Observed and expected upper limits on the cross-sections of the TChiWg model, together with the theoretical cross-sections.

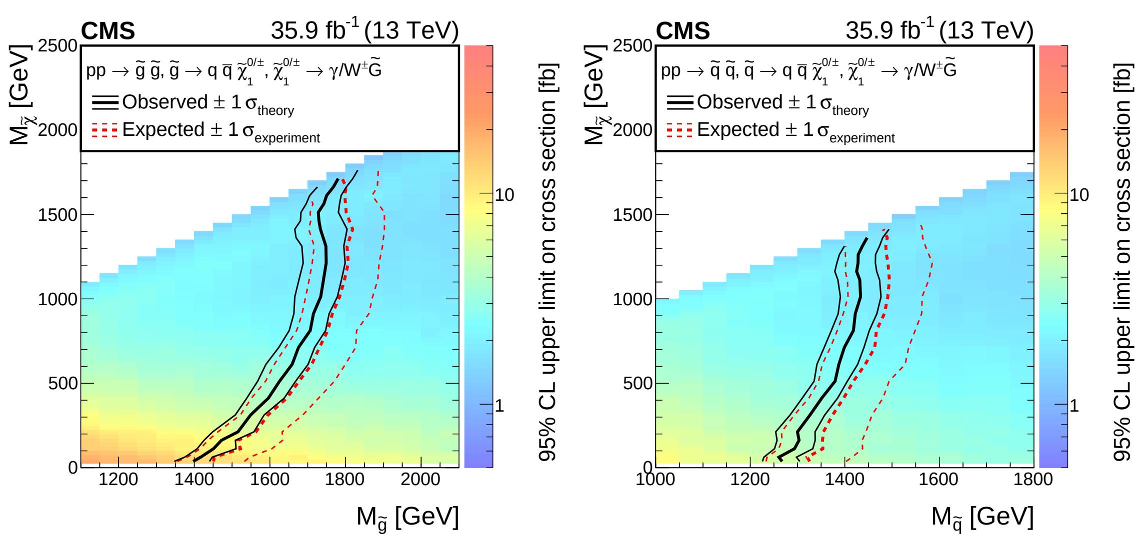


Figure 8.5: Observed and expected exclusion limits for T5Wg (left) and T6Wg (right) models. The upper limits are set on the cross-section $\times$ BR, where BR is the branching fractions to the neutralino/chargino states. For simplicity, BR is assumed to be 50%.

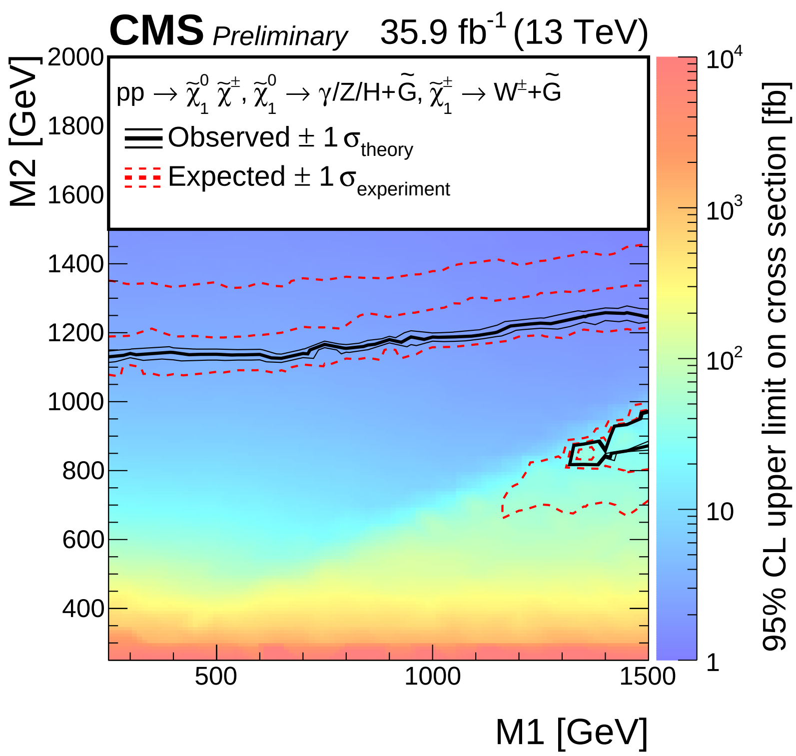


Figure 8.6: Exclusion limits for GGM models as a function of the M1 and M2 masses.

Chapter 9

Conclusions

A search for SUSY with general gauge mediation in events with at least one lepton, one photon and large missing transverse momentum using proton-proton collisions at $\sqrt{s} = 13$ TeV has been presented. The data sample, collected in 2016 with the CMS detector at the CERN LHC, corresponds to an integrated luminosity of 35.9 fb^{-1}.

The search selects events online using diphoton and photon+muon triggers, and applies an offline selection using kinematic requirements on photon, lepton and missing transverse momentum. Signal candidate events in the $p_T^{\rm miss} > 120$ GeV and $M_T > 100$ GeV search region are counted in multiple bins of p_T^{γ}, H_T, and $p_T^{\rm miss}$. The estimation of the SM backgrounds are performed using data-driven methods and MC simulations, while the estimation methods are verified in a dedicated validation region.

No significant excess above the SM background is observed in the signal region. The results are interpreted in GGM models as well as simplified models motivated by gauge-mediated supersymmetry breaking. GGM scans are performed in an M1 and M2 parameter space where model points up to M2 = 1200 GeV are excluded. For strong production simplified models, gluino production up to 1700 GeV and squarks up to 1400 GeV can be excluded at 95% confidence level. Final states with an additional lepton enhance the sensitivity to electroweak production of SUSY particles. For the TChiWg model, where $\tilde{\chi}_1^0\tilde{\chi}_1^{\pm}$ are pair produced, a NLSP of mass of up to 900 GeV is excluded, extending the current best limit by about 120 GeV.

The future of this search appears very promising. With more data, the sensitivity to the SUSY signals can be enhanced and the estimation of the SM backgrounds can be improved. A possible future improvement of this search is to use data-driven method to estimate the contributions from $t\bar{t}\gamma$, $WW\gamma$, and $WZ\gamma$ in a control region formed by one photon and two leptons. Such methods would greatly benefit from the increasing amount of data. Up to today, the LHC has delivered a total of 136.9 fb^{-1} of proton-proton collisions at $\sqrt{s}$ = 13 TeV. The Phase I upgrade of the LHC is scheduled after the data-taking in 2018, and the LHC will restart with ultimate design luminosity after this upgrade. A full exploration of these data will bring us closer to new physics.

Bibliography

[1] CMS Collaboration, "Operational experience with the CMS pixel detector in LHC Run II", *Journal of Instrumentation* **11** (2016) C12057. ix, 32

[2] CMS Collaboration, "CMS Technical Design Report for the Level-1 Trigger Upgrade", Technical Report CERN-LHCC-2013-011. CMS-TDR-12, 2013. ix, 43

[3] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965. ix, 48, 49

[4] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Phys. Rev. Lett.* **13** (1964) 508–509, doi:10.1103/PhysRevLett.13.508. 7

[5] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", *Phys. Rev. Lett.* **13** (1964) 321–323, doi:10.1103/PhysRevLett.13.321. 7

[6] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles", *Phys. Rev. Lett.* **13** (1964) 585–587, doi:10.1103/PhysRevLett.13.585. 7

[7] UA1 Collaboration, "Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV", *Physics*

Letters B **122** (1983) 103 – 116,
doi:https://doi.org/10.1016/0370-2693(83)91177-2. 9

[8] UA2 Collaboration, "Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN $\bar{p}p$ collider", *Physics Letters B* **122** (1983) 476 – 485,
doi:https://doi.org/10.1016/0370-2693(83)91605-2. 9

[9] UA1 Collaboration, "Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider", *Physics Letters B* **126** (1983) 398 – 410, doi:https://doi.org/10.1016/0370-2693(83)90188-0. 9

[10] UA2 Collaboration, "Evidence for $Z^0 \to e^+e^-$ at the CERN $\bar{p}p$ collider", *Physics Letters B* **129** (1983) 130 – 140,
doi:https://doi.org/10.1016/0370-2693(83)90744-X. 9

[11] CDF Collaboration, "Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab", *Phys. Rev. Lett.* **74** (1995) 2626–2631, doi:10.1103/PhysRevLett.74.2626. 9

[12] D0 Collaboration, "Observation of the Top Quark", *Phys. Rev. Lett.* **74** (1995) 2632–2637, doi:10.1103/PhysRevLett.74.2632. 9

[13] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Physics Letters B* **716** (2012) 30 – 61,
doi:https://doi.org/10.1016/j.physletb.2012.08.021. 9

[14] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Physics Letters B* **716** (2012) 1 – 29,
doi:https://doi.org/10.1016/j.physletb.2012.08.020. 9

[15] V. C. Rubin and J. Ford, W. Kent, "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions", *Astrophysical Journal* **159** (1970) 379, doi:10.1086/150317. 11

[16] P. Ramond, "Dual theory for free fermions", *Phys. Rev. D* **3** (1971) 2415, doi:10.1103/PhysRevD.3.2415. 11

[17] A. Neveu and J. H. Schwarz, "Factorizable dual model of pions", *Nucl. Phys. B* **31** (1971) 86, doi:10.1016/0550-3213(71)90448-2. 11

[18] D. V. Volkov and V. P. Akulov, "Possible universal neutrino interaction", *JETP Lett.* **16** (1972) 438. 11

[19] J. Wess and B. Zumino, "A lagrangian model invariant under supergauge transformations", *Phys. Lett. B* **49** (1974) 52, doi:10.1016/0370-2693(74)90578-4. 11

[20] J. Wess and B. Zumino, "Supergauge transformations in four dimensions", *Nucl. Phys. B* **70** (1974) 39, doi:10.1016/0550-3213(74)90355-1. 11

[21] P. Fayet, "Supergauge invariant extension of the Higgs mechanism and a model for the electron and its neutrino", *Nucl. Phys. B* **90** (1975) 104, doi:10.1016/0550-3213(75)90636-7. 11

[22] H. P. Nilles, "Supersymmetry, supergravity and particle physics", *Phys. Rep.* **110** (1984) 1, doi:10.1016/0370-1573(84)90008-5. 11

[23] J. B. Julius Wess, "Supersymmetry and Supergravity". Princeton University Press, 1992. 14

[24] N. S. P. Meade and D. Shih, "General gauge mediation", *Prog. Theor. Phys. Suppl.* **177** (2009) 143, doi:10.1143/PTPS.177.143, arXiv:0801.3278. 17

[25] A. M. P. Grajek and D. Redigolo, “Phenomenology of general gauge mediation in light of a 125 GeV Higgs”, *JHEP* **07** (2013) 109, doi:10.1007/JHEP07(2013)109. 17

[26] G. F. Giudice and R. Rattazzi, “Gauge-mediated supersymmetry breaking”, *Perspectives on Supersymmetry* (World Scientific, Singapore, 1998) p. 355. 17

[27] CMS Collaboration, “Interpretation of searches for supersymmetry with simplified models”, *Phys. Rev. D* **88** (2013) 052017, doi:10.1103/PhysRevD.88.052017. 21

[28] A. Heister et al., “Search for gauge mediated SUSY breaking topologies in ee collisions at centre-of-mass energies up to 209 GeV”, *The European Physical Journal C - Particles and Fields* **25** (2002) 339–351, doi:10.1007/s10052-002-1005-z. 24

[29] M. Gataullin, S. Rosier, L. Xia, and H. Yang, “Searches for Gauge-Mediated SUSY Breaking Topologies with the L3 Detector at LEP”, *AIP Conference Proceedings* **903** (2007) 217–220, doi:10.1063/1.2735164, arXiv:https://aip.scitation.org/doi/pdf/10.1063/1.2735164. 24

[30] DELPHI Collaboration, “Photon events with missing energy at $\sqrt{s} = 183$ to 189 GeV”, *The European Physical Journal C - Particles and Fields* **17** (2000) 53–65, doi:10.1007/s100520000434. 24

[31] OPAL Collaboration, “Searches for Gauge-Mediated Supersymmetry Breaking topologies in e^+e^- collisions at centre-of-mass energies up to $\sqrt{s} = 209$ GeV”, *The European Physical Journal C - Particles and Fields* **46** (2006) 307–341, doi:10.1140/epjc/s2006-02524-8. 24

[32] CDF Collaboration, “Search for anomalous production of diphoton events with missing transverse energy at CDF and limits on gauge-mediated

supersymmetry-breaking models", *Phys. Rev. D* **71** (2005) 031104, doi:10.1103/PhysRevD.71.031104. 25

[33] D0 Collaboration, "Search for supersymmetry in di-photon final states at s=1.96 TeV", *Physics Letters B* **659** (2008) 856 – 863, doi:https://doi.org/10.1016/j.physletb.2007.12.006. 25

[34] ATLAS Collaboration, "Search for diphoton events with large missing transverse momentum in 7 TeV proton-proton collision data with the ATLAS detector", *Physics Letters B* **718** (2012) 411 – 430, doi:https://doi.org/10.1016/j.physletb.2012.10.069. 25

[35] CMS Collaboration, "Search for new physics in events with photons, jets, and missing transverse energy in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **03** (2013) 111, doi:10.1007/JHEP03(2013)111, arXiv:1211.4784. 25

[36] CMS Collaboration, "Search for supersymmetry with photons in pp collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev. D* **92** (2015) 072006, doi:10.1103/PhysRevD.92.072006. 25

[37] ATLAS Collaboration, "Search for photonic signatures of gauge-mediated supersymmetry in 8 TeV pp collisions with the ATLAS detector", *Phys. Rev. D* **92** (2015) 072001, doi:10.1103/PhysRevD.92.072001. 25

[38] ATLAS Collaboration, "Search for photonic signatures of gauge-mediated supersymmetry in 13 TeV pp collisions with the ATLAS detector", *Phys. Rev. D* **97** (2018) 092006, doi:10.1103/PhysRevD.97.092006. 25

[39] CMS Collaboration, "Search for supersymmetry in events with a photon, a lepton, and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV", *Physics Letters B* **757** (2016) 6 – 31, doi:https://doi.org/10.1016/j.physletb.2016.03.039. 25

[40] Y. Iiyama, "Search for Supersymmetry in pp Collisions at $\sqrt{s} = 8$ TeV with a Photon, Lepton, and Missing Transverse Energy", *Dissertations.* (2015) Paper 594. 25

[41] R. Voss and A. Breskin, eds., "The CERN Large Hadron Collider, accelerator and experiments". CERN, 2009. 26

[42] CMS Collaboration, "Achieving the optimal performance of the CMS ECAL in Run II", Technical Report CMS-CR-2016-325, CERN, Geneva, 2016. 36

[43] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", *JINST* **8** (2013) P11002, `doi:10.1088/1748-0221/8/11/P11002`, `arXiv:1306.6905`. 42

[44] C. Battilana, "The CMS muon system: status and upgrades for LHC Run-2 and performance of muon reconstruction with 13 TeV data", *Journal of Instrumentation* **12** (2017) C01048. 42

[45] A. Zabiet al., "The CMS Level-1 Calorimeter Trigger for the LHC Run II", *Journal of Instrumentation* **12** (2017) C01065. 43

[46] Federico De Guio and the CMS Collaboration, "The CMS data quality monitoring software: experience and future prospects", *Journal of Physics: Conference Series* **513** (2014) 032024. 45

[47] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *Journal of Instrumentation* **9** (2014) P10009. 49

[48] CMS, "Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Journal of Instrumentation* **10** (2015) P08010. 51

[49] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Journal of Instrumentation* **10** (2015) P06005. 53

[50] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *The European Physical Journal C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2. 54

[51] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-kt jet clustering algorithm", *Journal of High Energy Physics* **2008** (2008) 063. 54

[52] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", *Journal of Instrumentation* **7** (2012) P10002. 55

[53] CMS Collaboration, "Missing transverse energy performance of the CMS detector", *JINST* **6** (2011), no. arXiv:1106.5048. CMS-JME-10-009. CERN-PH-EP-2011-051, P09001. 56 p. 56

[54] G. Petrucciani, A. Rizzi, and C. Vuosalo, "Mini-AOD: A New Analysis Data Format for CMS", *Journal of Physics: Conference Series* **664** (2015) 072052. 57

[55] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301. 58

[56] CMS Collaboration, "Search for dark matter in monophoton final state at 13 TeV", CMS AN 2016/078. 58

[57] G. B. et al., "Production of Drell-Yan lepton pairs in hadron collisions Transverse-momentum resummation at next-to-next-to-leading logarithmic

accuracy", *Phys. Lett. B* **696** (2011) 207–213, doi:10.1016/j.physletb.2010.12.024, arXiv:1007.2351. 58

[58] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849. 58

[59] T. Sjöstrand et al., "An Introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159–177, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012. 58

[60] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J.* **C76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815. 58

[61] GEANT4 Collaboration, "GEANT4: A Simulation toolkit", *Nucl. Instrum. Meth.* **A506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8. 58

[62] CMS Collaboration, "twiki:SUSY Leptons selection and Data/MC Scale Factors", technical report. 71, 73, 74

[63] K. Rehermann and B. Tweedie, "Efficient identification of boosted semileptonic top quarks at the LHC", *JHEP* **03** (2011) 059, doi:doi:10.1007/JHEP03(2011)059, arXiv:1007.2221. 72

[64] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663. 75

[65] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling", *eConf* **C0303241** (2003) MOLT007, arXiv:physics/0306116. [186(2003)]. 79

[66] CMS Collaboration, "Measurement of the semileptonic $t\bar{t}+\gamma$ production cross section in pp collisions at $\sqrt{s}=8$ TeV", *JHEP* **10** (2017) 006, doi:10.1007/JHEP10(2017)006, arXiv:1706.08128. 119

[67] CMS Collaboration, "Measurement of the $t\bar{t}$ production cross section using events in the eμ final state in pp collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **77** (2017) 172, doi:10.1140/epjc/s10052-017-4718-8, arXiv:1611.04040. 119

[68] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", Technical Report CMS-PAS-LUM-17-001, CERN, Geneva, 2017. 119

[69] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006. 129

[70] A. L. Read, "Presentation of search results: The CL_s technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313. 129

[71] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group Collaboration, "Procedure for the LHC Higgs boson search combination in Summer 2011", Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, 2011. 129

[72] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *The European Physical Journal C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0. 130

[73] CMS Collaboration, "Search for gauge-mediated supersymmetry in events with at least one photon and missing transverse momentum in pp collisions at $\sqrt{s} =$ 13 TeV", *Phys. Lett. B* **780** (2018) 118, doi:doi:https://doi.org/10.1016/j.physletb.2018.02.045, arXiv:1711.08008. 131