

INDIAN INSTITUTE OF SCIENCE EDUCATION AND  
RESEARCH (IISER) PUNE

DOCTORAL THESIS

---

**Search for supersymmetry in events with a  
photon, jets, b-jets, and missing transverse  
momentum**

---

*Author:*  
Vinay HEGDE

*Supervisor:*  
Dr. Seema SHARMA

*A thesis submitted in fulfillment of the requirements  
for the degree of Doctor of Philosophy  
in the*

**Experimental Particle Physics  
Department Physics**

March 26, 2019



*“A quote”*

Author of quote



## *Abstract*

A search for supersymmetry is presented based on events with at least one photon, jets, and large missing transverse momentum produced in proton-proton collisions at a center-of-mass energy of 13 TeV. The data correspond to an integrated luminosity of  $35.9 \text{ fb}^{-1}$  and were recorded at the LHC with the CMS detector in 2016. The analysis characterizes signal-like events by categorizing the data into various signal regions based on the number of jets, the number of b-tagged jets, and the missing transverse momentum. No significant excess of events is observed with respect to the expectations from standard model processes. Limits are placed on the gluino and top squark pair production cross sections using several simplified models of supersymmetric particle production with gauge-mediated supersymmetry breaking. Depending on the model and the mass of the next-to-lightest supersymmetric particle, the production of gluinos with masses as large as 2120 GeV and the production of top squarks with masses as large as 1230 GeV are excluded at 95% confidence level.



## *Acknowledgements*

The acknowledgments and the people to thank go here, don't forget to include your project advisor...



# Contents

<b>Abstract</b>	<b>v</b>
<b>Acknowledgements</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Standard model particles . . . . .	2
1.2 Conservation laws and symmetry . . . . .	3
1.3 Strong interactions . . . . .	3
1.4 Electro-weak interactions . . . . .	4
1.5 The Higgs mechanism . . . . .	4
1.6 Limitations of SM . . . . .	5
1.7 Supersymmetric extension of SM . . . . .	6
1.7.1 R-parity . . . . .	8
1.8 Gauge mediated SUSY breaking . . . . .	8
1.8.1 Phenomenology of GMSB . . . . .	9
1.8.2 Simplified GMSB models . . . . .	9
<b>2 The Experimental Setup</b>	<b>11</b>
2.1 The large hadron collider . . . . .	11
2.2 The compact muon solenoid . . . . .	12
2.2.1 CMS coordinate system . . . . .	13
2.2.2 Superconducting magnet . . . . .	14
2.2.3 Inner tracking system . . . . .	14
Performance of tracker . . . . .	15
2.2.4 Electromagnetic calorimeter . . . . .	16
2.2.5 Hadron calorimeter . . . . .	17
2.2.6 Muon chambers . . . . .	18
2.3 Energy response in long and short fibers of HF calorimeter . . . . .	19
2.3.1 Data and simulation samples . . . . .	20
2.3.2 Energy in long and short fibers of HF . . . . .	21
2.3.3 Effect of Pileup on $R_{S/L}$ . . . . .	23
2.3.4 Studying $R_{S/L}$ at $\sqrt{s} = 13$ TeV and 8 TeV data with 50 ns bunch spacing . . . . .	24
2.3.5 $R_{S/L}$ for 2015 Data and 2016 Data . . . . .	25
2.3.6 Performance of $R_{S/L}$ in 2016 Data . . . . .	27
Stability of $R_{S/L}$ in 2016 Data . . . . .	28
2.3.7 Corrections for the $R_{S/L}$ Based on $\phi$ Symmetry . . . . .	30
2.3.8 Performance of $R_{S/L}$ in data & MC . . . . .	33
$R_{S/L}$ vs $i\eta$ . . . . .	35
$R_{S/L}$ vs $i\phi$ . . . . .	36
2.3.9 Summary of HF studies using $R_{S/L}$ . . . . .	37
2.4 Trigger system . . . . .	37
2.5 Object reconstruction . . . . .	38

2.5.1	Tracking . . . . .	38
2.5.2	Interaction vertices . . . . .	40
2.5.3	Muons . . . . .	40
2.5.4	Electrons and photons . . . . .	40
Photon Identification . . . . .	41	
Electron Identification . . . . .	43	
2.5.5	Jets . . . . .	45
2.5.6	Tagging of b-quark jet . . . . .	47
2.5.7	Missing transverse momentum . . . . .	48
<b>3</b>	<b>SUSY search with photon</b>	<b>51</b>
3.1	Introduction . . . . .	51
3.2	MC simulation samples . . . . .	52
3.3	Triggers . . . . .	52
3.3.1	Signal triggers . . . . .	53
3.3.2	Multijet triggers . . . . .	55
3.4	Event selection . . . . .	55
3.5	SM background and search regions . . . . .	57
3.5.1	Optimization of search bins . . . . .	59
3.6	Background Estimation . . . . .	61
3.6.1	Lost lepton estimation . . . . .	62
3.6.2	Estimation of events with electrons faking photons . . . . .	65
Fake-rate parameterization . . . . .	66	
Data/MC fake rate corrections . . . . .	66	
Fake rate prediction . . . . .	68	
3.6.3	Invisible Z estimation . . . . .	68
3.6.4	$\gamma + \text{jets}$ and QCD multijet estimation . . . . .	70
3.7	Systematic uncertainty for signal models . . . . .	74
<b>4</b>	<b>Results and summary</b>	<b>77</b>
4.1	Results & statistical interpretations . . . . .	77
4.2	Summary . . . . .	81
4.3	Outlook . . . . .	83
4.3.1	Electroweak models . . . . .	83
4.3.2	Sensitivity of the search . . . . .	83
<b>A</b>	<b>Additional tables</b>	<b>85</b>
A.1	Lost lepton + $\tau_{\text{had}}$ predictions . . . . .	85
A.2	Electron fake photon . . . . .	86
A.3	Invisible Z estimation . . . . .	86
A.4	$\gamma + \text{jets}$ and QCD multijet . . . . .	86
<b>B</b>	<b>Data verses MC in control regions</b>	<b>93</b>
B.1	$e/\mu + \gamma$ control region . . . . .	93
B.2	$Z(\ell\ell) + \gamma$ control region . . . . .	94
B.3	Signal acceptance $\times$ efficiency . . . . .	95

# List of Figures

1.1	SM cross section measurements by CMS . . . . .	5
1.2	Top loop diagram contribution to H mass . . . . .	6
1.3	Corrections to H mass from top loop and stop loop. . . . .	6
1.4	BR for bino and neutral wino NLSP decays . . . . .	9
1.5	SMS diagrams . . . . .	10
2.1	CERN accelerator complex . . . . .	11
2.2	CMS detector view . . . . .	13
2.3	Integrated luminosity of data . . . . .	13
2.4	CMS tracker . . . . .	15
2.5	Tracker resolution for muons . . . . .	15
2.6	Track reconstruction efficiency . . . . .	16
2.7	Tracker material budget . . . . .	16
2.8	ECal schematic and resolution . . . . .	17
2.9	Schematic diagram of HCal . . . . .	18
2.10	Muon chambers . . . . .	19
2.11	Muon $p_T$ resolution . . . . .	20
2.12	Channel occupancy for depth 1 (left), and depth 2 (right). . . . .	22
2.13	Total number of rechits (inclusive in $i\eta$ and $i\phi$ ) . . . . .	22
2.14	Distribution of $E_{\text{Short}}$ vs $E_{\text{Long}}$ for HF channels. . . . .	22
2.15	$R_{S/L}$ for $i\eta$ 32 tower with $i\phi$ inclusive for 2015D-50 ns. . . . .	23
2.16	Asymmetric Gaussian fit for $i\eta$ 30 . . . . .	23
2.17	Asymmetric Gaussian fit for $i\eta$ 39 . . . . .	23
2.18	Effect of pileup on $R_{S/L}$ for different $i\eta$ rings (integrated over $i\phi$ channels. . . . .	24
2.19	Pileup comparison for 2015 data and 2012 data . . . . .	24
2.20	Energy in long fibers . . . . .	25
2.21	Energy in short fibers . . . . .	25
2.22	$R_{S/L}$ vs $i\eta$ for 3 different run conditions . . . . .	25
2.23	$R_{S/L}$ vs $i\phi$ plot for $i\eta$ -35 . . . . .	26
2.24	Number of primary vertices before re-weighting . . . . .	26
2.25	Number of primary vertices after re-weighting . . . . .	26
2.26	No. of HF RecHits above 10 GeV . . . . .	27
2.27	HF RecHit energy in long and short fibers . . . . .	27
2.28	Energy in long . . . . .	27
2.29	Energy in short . . . . .	27
2.30	$R_{S/L}$ for 2015 and 2016B data . . . . .	28
2.31	Number of primary vertices (left) for 2016B,C,D,E, and (right) for 2016B,F,G,H. . . . .	28
2.32	Number of RecHits (left) for 2016B,C,D,E, and (right) for 2016B,F,G,H. . . . .	29
2.33	RecHitEnergy distributions for 2016B,C,D,E . . . . .	29
2.34	RecHitEnergy distributions for 2016B,F,G, H . . . . .	29
2.35	$R_{S/L}$ vs $i\eta$ for 2016 with lower $E_{long}$ threshold . . . . .	30

2.36	$R_{S/L}$ vs $i\eta$ for 2016 with higher $E_{long}$ threshold . . . . .	30
2.37	$R_{S/L}$ vs $i\phi$ for 2016B,C,D,E . . . . .	31
2.38	$R_{S/L}$ vs $i\phi$ for 2016B,C,D,E . . . . .	31
2.39	Correction factors as a function of $i\eta, i\phi$ . . . . .	32
2.40	$R_{S/L}$ vs $i\phi$ for $i\eta = 30$ for (left) current detector, and (right) after corrections for 2016B,F,G,H . . . . .	32
2.41	$R_{S/L}$ vs $i\phi$ for $i\eta = 38$ for (left) current detector, and (right) after corrections for 2016B,F,G,H . . . . .	32
2.42	$R_{S/L}$ vs $i\phi$ before and after corrections for 2016B with $E_{long} > 30$ , $E_{short} > 10$ . . . . .	33
2.43	$R_{S/L}$ vs $i\phi$ before and after corrections for 2016B with $E_{long} > 100$ , $E_{short} > 10$ . . . . .	33
2.44	Observed number of interactions in data and MC (Solid blue line- MC before PU re-weighting, dotted blue-after PU re-weighting). . . . .	34
2.45	Number of reconstructed primary vertices in data and MC . . . . .	34
2.46	Number of HF rechits above 10 GeV . . . . .	34
2.47	Distribution of energies in (left) all HF rechits, (middle) long fibers or depth 1, and (right) in short fibers or depth 2. . . . .	34
2.48	$R_{S/L}$ vs $ i\eta $ with $E_{long} > 20, E_{short} > 10$ GeV . . . . .	35
2.49	$R_{S/L}$ vs $ i\eta $ with $E_{long} > 40, E_{short} > 10$ GeV . . . . .	35
2.50	$R_{S/L}$ vs $i\eta$ with $E_{long} > 100, E_{short} > 10$ . . . . .	36
2.51	$R_{S/L}$ vs $i\eta$ with $E_{long} > 50, E_{short} > 50$ . . . . .	36
2.52	$R_{S/L}$ vs $i\phi$ for data and MC . . . . .	36
2.53	SM cross section vs SUSY cross section . . . . .	37
2.54	Schematic of L1T . . . . .	38
2.55	Particle interaction in CMS . . . . .	39
2.56	Loose + pixel seed veto photon efficiency . . . . .	42
2.57	TnP fit for MiniIso pass/fail . . . . .	44
2.58	Mini-isolation efficiency and SF . . . . .	44
2.59	Energy energy composition . . . . .	45
2.60	Application of JEC . . . . .	46
2.61	Jet $p_T$ response with and w/o corrections . . . . .	46
2.62	AK4PF jet energy resolution . . . . .	47
2.63	b-jet showing b-hadron decays . . . . .	47
2.64	CSVv2 output and b-tagging efficiency . . . . .	48
2.65	$p_T^{\text{miss}}$ resolution . . . . .	49
3.1	Efficiency of Photon165 trigger versus $p_T^\gamma$ for all eras (left) and split by eras (right). . . . .	53
3.2	Efficiency of Photon90_PFHT600 trigger versus $p_T^\gamma$ (top) and $H_T^\gamma$ (bottom) for all eras (left) and split by eras (right). . . . .	54
3.3	Efficiency of the OR of the Photon165 and Photon90_PFHT600 triggers versus $H_T^\gamma$ (left) and $p_T^\gamma$ (right). . . . .	54
3.4	Efficiency of the OR of the Photon165 and Photon90_PFHT600 versus $H_T^\gamma$ (left) and $p_T^\gamma$ (right) for events with one tight electron. . . . .	54
3.5	Efficiency of HLT_PFHT800 trigger for eras B-G, and HLT_PFHT900 OR HLT_CaloJet500_NoJetID for era H for zero photon events. . . . .	55

3.6 Distributions of $p_T^\gamma$ , $p_T^{miss}$ , $N_{jets}$ , $N_{b\text{-jets}}$ , and $H_T^\gamma$ after baseline selection. Filled, stacked histograms represent SM backgrounds, taken directly from simulation, open histograms show two examples of T5bbbbZG signals, one with a high mass (1750 GeV) NLSP and one with a low mass (150 GeV) NLSP. The gluino mass for both signals is 1800 GeV. . . . .	58
3.7 Distributions of events after baseline selection. Filled, stacked histograms represent SM backgrounds, taken directly from simulation, open histograms show two examples of T5bbbbZG signals, one with a high mass (1750 GeV) NLSP and one with a low mass (150 GeV) NLSP. The gluino mass for both signals is 1800 GeV. The lowest $p_T^{miss}$ bin is not used for final limit calculations. . . . .	59
3.8 Expected exclusion curves for different types of binning for T5qqqqHG (left) and T5bbbbZG (right). In the legends, subscript $j$ in $N_j^i$ refers to number of jets and superscript $i$ refers to number of b-tagged jets. The red curve corresponds to 63 search bins defined as ( $N_{jets} = 2\text{-}4, 5\text{-}6, \geq 7$ ) $\times$ ( $N_{b\text{-jets}} = 0, 1, \geq 2$ ) $\times$ (7 $p_T^{miss}$ bins). Other curves are obtained by combining these 63 bins in either $N_{jets}$ or $N_{b\text{-jets}}$ or $p_T^{miss}$ or any combination of these variables. The final binning of the analysis is based on orange curve defined as ( $N_{jets} = 2\text{-}4, 5\text{-}6, \geq 7$ ) $\times$ ( $N_{b\text{-jets}} = 0, \geq 1$ ) $\times$ (7 $p_T^{miss}$ bins). . . . .	60
3.9 Pie charts showing BG in low $p_T^{miss}$ . . . . .	61
3.10 $m_T$ distribution for $\mu\gamma$ control region . . . . .	62
3.11 Expected closure of lost electron (left) and lost $\mu + \tau_{had}$ (right) predictions where simulated event yields are treated like data. . . . .	64
3.12 Distribution of $\Delta R(\ell, \gamma)$ for single electron (left) and single muon (right) events. . . . .	65
3.13 $Q_{mult}$ distribution in a jet matched to the electron in DY, W+jets, and $t\bar{t}$ events. Each of these distribution are scaled to unit area. . . . .	66
3.14 Comparison of fake rate prediction using fake rate parameterization and true MC yields using $W/t\bar{t} + (\gamma)$ simulations versus $p_T^{miss}$ (top left), $N_{jets}$ (top right), $N_{b\text{-jets}}$ (bottom left) and search bins (bottom right). The hashed region in the lower panel of bottom right plot shows the total systematic uncertainty. . . . .	67
3.15 Tag & probe fits of the fake rate in 0-1 $Q_{mult}$ bins (top row) and $\geq 2 Q_{mult}$ bins (bottom row). The first and second columns are the the photon region and electron region fits in MC, respectively. The third and fourth columns are the photon region and electron region fits in data, respectively. . . . .	67
3.16 Comparison of $Z(ll)\gamma$ with $Z(\nu\nu)\gamma$ as function of $p_T^\gamma$ (top left), $p_T^{miss}$ (top right), $N_{b\text{-jets}}$ (bottom left) and $N_{jets}$ (bottom right) for $p_T^\gamma > 190$ GeV . . . . .	69
3.17 $\min(\Delta\phi_1, \Delta\phi_2)$ in MC . . . . .	71
3.18 Definitions of various regions used in ABCD methods. Region D is signal region where background is to be estimated or validation region for zero photon sample. . . . .	71
3.19 Comparison of $\gamma + \text{jets}$ and QCD multijet background estimated using low- $\Delta\phi$ control regions taken from MC (blue) to that expected in search regions (cyan). . . . .	72
3.20 In zero photon validation region in MC, comparison of $\gamma + \text{jets}$ and QCD multijet background estimated using low- $\Delta\phi$ control regions taken from MC (blue) to that expected in search regions (cyan). . . . .	73

3.21 Double ratio validation . . . . .	73
4.1 Observed numbers of events and predicted numbers of events from the various SM backgrounds in the 25 signal regions. The categories, denoted by vertical lines, are labeled as $N_j^b$ , where $j$ refers to the number of jets and $b$ refers to the number of b-tagged jets. The numbered bins within each category are the various $p_T^{\text{miss}}$ bins, as defined in Table 4.1. The lower panel shows the ratio of the observed events to the predicted SM background events. The error bars in the lower panel are the quadrature sum of the statistical uncertainty in the observed data and the systematic uncertainty in the predicted backgrounds before the adjustments based on a maximum likelihood fit to data assuming no signal strength. . . . .	81
4.2 Observed and expected 95% CL upper limits for gluino or top squark pair production cross sections for the T5qqqqHG (upper left), T5bbbbZG (upper right), T5tttZG (bottom left), and T6ttZG (bottom right) models. Black lines denote the observed exclusion limit and the uncertainty due to variations of the theoretical prediction of the gluino or top squark pair production cross section. The dashed lines correspond to the region containing 68% of the distribution of the expected exclusion limits under the background-only hypothesis. . . . .	82
4.3 TChiWG and TChiNG SMS digrams . . . . .	83
4.4 Expected UL for EW for SUSY . . . . .	84
B.1 Distribution of $e + \gamma$ events versus $p_T^{\text{miss}}$ (top left), $H_T^\gamma$ (top right), $m_T$ (middle left), $p_{T,e}$ (middle right), $N_{\text{jets}}$ (bottom left), $N_{\text{b-jets}}$ (bottom right). Filled histograms denoted expected yields from MC for various processes. Markers represent data yields. MC event yields are scaled to the same integrated luminosity as data. . . . .	93
B.2 Distribution of $\mu + \gamma$ events versus $p_T^{\text{miss}}$ (top left), $S_T$ (top right), $M_T$ (middle left), $p_{T,\mu}$ (middle right), $N_{\text{jets}}$ (bottom left), $N_{\text{b-jets}}$ (bottom right). Filled histograms denoted expected yields from MC for various processes. Markers represent data yields. MC event yields are scaled to the same integrated luminosity as data. . . . .	94
B.3 Comparison of $(ll)\gamma$ in data with $ll + \gamma$ events in MC as function of $p_T^{\text{miss}}$ (top left), $N_{\text{jets}}$ (top right), $N_{\text{b-jets}}$ (bottom left) and $m_{ll}$ (bottom right) for $p_T^\gamma > 190\text{GeV}$ . Note that $p_T^{\text{miss}}$ is calculated by ignoring the di-lepton pair. For the top left plot of $p_T^{\text{miss}}$ , hashed error bars refer to statistical and EW correction uncertainties listed in Table 3.5. For the remaining 3 plots, error bars are statistical only. . . . .	95
B.4 Acceptance $\times$ efficiency for different mass point corresponding to T5qqqqHG (top left), T5bbbbZG (top right), T5tttZG (bottom left) and T6ttZG (bottom right). . . . .	96

# List of Tables

1.1	Leptons and quarks of SM with their properties . . . . .	2
1.2	Gauge bosons and higgs boson in SM with their properties . . . . .	3
1.3	Chiral and gauge supermultiplets of MSSM . . . . .	7
1.4	Gauge and mass eigen states of MSSM . . . . .	7
2.1	The LHC parameters . . . . .	12
2.2	Characteristics of tracker . . . . .	14
2.3	Collision data used for $R_{S/L}$ studies of HF . . . . .	21
2.4	Photon Identification . . . . .	41
3.1	Signal region and sideband regions . . . . .	59
3.2	Parameterization of transfer factors for lost-e and lost- $\mu + \tau_{\text{had}}$ events. . . . .	63
3.3	fake rate as measured by T&P procedure on Drell-Yan events in data and MC. . . . .	67
3.4	Purity in the $Z\gamma$ control region. . . . .	69
3.5	Electroweak corrections as a function of $p_T^{\text{miss}}$ . . . . .	70
3.6	Double ratio( $\kappa^{\text{MC}}$ ) computed in the one $\gamma$ regions and ratio of high- $\Delta\phi$ to low- $\Delta\phi$ ( $R^{\text{data}}$ ) obtained from data using $100 < p_T^{\text{miss}} < 200$ GeV region with electroweak contribution subtracted. . . . .	72
3.7	Systematic uncertainties for signal samples . . . . .	75
4.1	Predicted and observed event yields for each of the 25 exclusive signal regions. . . . .	80
A.1	Observed events yeilds, average transfer factors and predictions for lost-e and lost- $\mu + \tau_{\text{had}}$ backgrounds in high $\Delta\phi$ regions. . . . .	85
A.2	Observed events yeilds, average transfer factors and predictions for lost-e and lost- $\mu + \tau_{\text{had}}$ backgrounds in the low- $\Delta\Phi$ control region. . . . .	86
A.3	Fake rate parameterization derived from $W/t\bar{t}$ simulations. . . . .	87
A.4	Observed single electron yields, after accounting for lepton scale factors and relative trigger efficiencies factor, in each of the signal regions, the average scale factor due to MC-based fake rates and data/MC fake rate correction factors from T&P measurements, and the corresponding fake-rate predictions with all the uncertainties (statistical and systematic) taken into consideration. . . . .	88
A.5	Observed single electron yields, after accounting for lepton scale factors and relative trigger efficiencies factor, in each of the low- $\Delta\phi$ control regions, the average scale factor due to MC-based fake rates and data/MC fake rate correction factors from T&P measurements, and the corresponding fake-rate predictions with all the uncertainties (statistical and systematic) taken into consideration. . . . .	89
A.6	$Z(\nu\nu) + \gamma$ MC yields and prediction for the signal regions. . . . .	90
A.7	$Z(\nu\nu) + \gamma$ MC yields and prediction for the low- $\Delta\Phi$ control regions. . . . .	91

A.8 Fake $p_T^{miss}$ background control sample statistics and electroweak predictions with uncertainties . . . . .	92
---	----

# List of Abbreviations

<b>CMS</b>	Compact Muon Solenoid
<b>LHC</b>	Large Hadron Collider
<b>SM</b>	Standard Model
<b>SUSY</b>	Supersymmetry
<b>QCD</b>	Quantum Chromodynamics
<b>QFT</b>	Quantum Field Theory



# List of Symbols



*For/Dedicated to/To my...*



## Chapter 1

# Introduction

Many of the natural phenomena taking place in this universe have made human beings to wonder what causes these phenomena. We, the human beings have been trying to find out if there is (or are) any fundamental principle(s) that are responsible for this. Based on today's understanding, there are four fundamental forces:

- Gravitational force: this causes apple to fall and also the planets to revolve around the sun.
- Electromagnetic force: the one which holds atoms together, basic principle in chemistry and almost everything, except gravity, that we come across in daily life is governed by this force.
- Weak force:  $\beta$  decay taking place in the nucleus is because of this force.
- Strong force: this force holds protons and neutrons together in the nucleus.

The matter that we see around us is made up of tiny particles, atoms. Atoms are made up of a nucleus (consisting of protons and neutrons, collectively called nucleons) and electrons. Nucleons have more fundamental particles called quarks. So the matter around us can be thought of as some combination of electrons and quarks. Apart from these, there are also other fundamental particles, some of which are similar to electrons and quarks but not the constituents of matter, some are mediators of the fundamental forces and one is responsible for the particles to have mass. Totally there are 18 such fundamental particles which we know today and four forces. Can we put all these particles and forces into a framework and try to understand the phenomena around us? The attempt that we made so far for this purpose has led to standard model (SM) of physics. Our attempt is not complete because in the framework of SM, we could not fit gravitational force or interaction and also the mediator of gravity, graviton. The graviton is a hypothetical particle, not experimentally detected, is not included in the SM. So the SM consists of 17 fundamental particles and 3 fundamental forces.

The SM has successfully explained many of the experimental results with very high degree of precision. But the SM is not a complete story of the universe, one of the simple reason being, not having gravity included in it. There are many other issues in the SM, such as the inability to explain dark matter and dark energy in the universe, matter anti-matter asymmetry etc, and many theoretical problems. To address many of the issues in the SM, many extensions of SM have been proposed, one of them being the supersymmetry (SUSY). SUSY predicts that for every SM particle, there is a superpartner similar in all the quantum numbers and properties except that the spin quantum number differs by half a unit. Failure to observe SUSY particles with same mass as SM particles (we did not find an SUSY electron, called selectron or sproton, sneutron etc), led to the ideas of breaking of this symmetry.

There are many ways of breaking SUSY, one of them being gauge mediated SUSY breaking (GMSB). In scenarios with GMSB, gravitino (super partner of hypothetical graviton) is the lightest SUSY particle (LSP) and it is stable. **The stability of LSP is a requirement from R-parity conservation which does not allow SUSY particle decay into only SM particles.** LSP is weakly interacting particle and hence its detection is very difficult in an experiment.

The compact muon solenoid (CMS) detector at the large hadron collider (LHC), CERN has a lot of potentials for testing the SM and also to carry out searches for beyond SM particles, such as SUSY particles. At the LHC, proton-proton collisions take place at the center of mass energy of 13 TeV. SUSY particles can be produced in these collisions if they exist and accessible at the LHC. In GMSB models, events with photon, quarks of light or bottom flavor and missing transverse momentum are expected. Missing transverse momentum is the signature of weakly interacting LSPs in the detector.

This thesis is about a SUSY search with at least one photon, light or bottom flavor jets which originate from quarks and large missing transverse momentum using the data collected in the year 2016 at the CMS detector. This chapter 1 gives a very brief introduction to SM, its limitations and supersymmetric extensions of SM and GMSB models. Chapter 2 describes the experimental set up consisting of LHC and the CMS detector. My work on hadron forward calorimeter of CMS is also described in this chapter. The main analysis is discussed in 3. Summary and some possible searches in the future is described in the final chapter 4.

## 1.1 Standard model particles

The SM of particle physics consists of 3 generations of quarks and leptons (Table 1.1), gauge bosons which are mediators of strong, electromagnetic (EM) and weak force, and a scalar higgs boson (Table 1.2).

TABLE 1.1: Leptons and quarks of SM with their properties

Leptons (spin $\frac{1}{2}$ )					
Gen.	Particle	Charge	Mass (MeV)	Interactions <sup>1</sup>	Chiral state
1	$\nu_e$	0	$< 2 \times 10^{-6}$	Weak	$\begin{bmatrix} \nu_e \\ e_L \end{bmatrix}, e_R$
	$e$	$\mp 1$	0.511	EM, Weak	
2	$\nu_\mu$	0	$< 0.19$	Weak	$\begin{bmatrix} \nu_\mu \\ \mu_L \end{bmatrix}, \mu_R$
	$\mu$	$\mp 1$	105.7	EM, Weak	
3	$\nu_\tau$	0	$< 18.2$	Weak	$\begin{bmatrix} \nu_\tau \\ \tau_L \end{bmatrix}, \tau_R$
	$\tau$	$\mp 1$	1777	EM, Weak	
Quarks (spin $\frac{1}{2}$ )					
Gen.	Particle	Charge	Mass (MeV)	Interactions	Chiral state
1	$u$	$\pm \frac{2}{3}$	$\approx 2.2$	Strong, EM, Weak	$\begin{bmatrix} u_L \\ d_L \end{bmatrix}, u_R, d_R$
	$d$	$\mp \frac{1}{3}$	$\approx 4.7$		
2	$c$	$\pm \frac{2}{3}$	$1.275 \times 10^3$		$\begin{bmatrix} c_L \\ s_L \end{bmatrix}, c_R, s_R$
	$s$	$\mp \frac{1}{3}$	$\approx 95$		
3	$t$	$\pm \frac{2}{3}$	$173.1 \times 10^3$		$\begin{bmatrix} t_L \\ b_L \end{bmatrix}, t_R, b_R$
	$b$	$\mp \frac{1}{3}$	$\approx 4.18 \times 10^3$		

<sup>1</sup>All particles with non-zero mass have gravitational interactions. But gravity is not a part of SM.

TABLE 1.2: Gauge bosons and higgs boson in SM with their properties

Particle	Charge	Mass (GeV)	Spin	Force mediation	Mixing fields
g	0	0	1	Strong	g
$\gamma$	0	0	1	EM	$W_3, B$
W	$\mp 1$	80.4	1	Weak	$W_1, W_2$
Z	0	91.2	1	Weak	$W_3, B$
H	0	125.2	0	-	$H_u, H_d$

Categorization of the quarks and leptons into 3 generations is done based on their masses [98] and properties. First generation consists of lightest and the third generation consists of heaviest particles. Neutrinos are grouped according to their interactions and their exact masses are unknown. The masses of SM particles are not predicted by the theory, but they are measured quantities in experiments. All the charged (in this thesis charge always refers to electric charge) particles have EM interactions. Neutral leptons, neutrinos, have only weak interactions. Quarks have additional quantum number, color, which is responsible for strong interaction.

## 1.2 Conservation laws and symmetry

Experiments have played a significant role in development of the SM. Various observations from them lead to framing many laws of nature. Similar to the law of conservation of energy and momentum in basic physics, there are other laws of conservation such as charge, angular momentum, color charge, lepton number within each generation etc. Whenever there is a conservation law, then there is a symmetry associated with it and a symmetry also implies a conservation law. This relation between conservation law and symmetry is given by Noether's theorem [83]. Conservation of energy, linear momentum and angular momentum are connected with symmetry in time translation, space translation and rotation respectively. Gauge transformation invariance in electrodynamics lead to conservation of charge. These symmetries can be represented mathematically using symmetry groups. For example, group representing rotational symmetry is  $SO(3)$ .

## 1.3 Strong interactions

The strong interaction is mediated by gluon and the theory is described by quantum chromodynamics (QCD). QCD is based on the symmetry group  $SU(3)_C$ , where C represents color (red, green or blue), having 3 degrees of freedom and 8 generators or 8 gluon fields. The eight colors of gluon are:

$$R\bar{G}, R\bar{B}, G\bar{R}, G\bar{B}, B\bar{R}, B\bar{G}, \frac{1}{\sqrt{2}}(R\bar{R} - G\bar{G}), \frac{1}{\sqrt{6}}(R\bar{R} + G\bar{G} - 2B\bar{B})$$

$SU(3)_C$  is a non-abelian gauge group which means that the gluons have self interactions. Quarks carry one color charge (red green or blue) and gluons are bi-colored having one of the 8 colors mentioned above. All the observed hadrons and mesons are colorless, in other words, no isolated quarks or gluons exist. The strength of strong interaction increases with distance (or decreases with energy). At very large energy, strong interaction becomes very small and this behavior is named as

asymptotic freedom [61][84]. As two quarks are separated from each other, quark-antiquark pair are created making initial quarks no longer isolated.

## 1.4 Electro-weak interactions

The interaction between charged particles is described by quantum electrodynamics (QED) which is based on the symmetry group  $U(1)$ . The mediator of this interaction is massless spin 1 photon. Unlike QCD, there are no self interactions of the force mediator since photon does not carry any charge.

The neutral leptons, neutrinos, take part only in weak interactions whose description is based on  $SU(2)_L$  symmetry. The mediators of this force are massive W and Z bosons. All the leptons and quarks of SM have weak interactions and neutrinos have only weak interaction. Weak theory is a chiral theory which implies that only left handed particles or right handed antiparticles participate in the interactions. A consequence of chiral theory is that there is no place for right handed neutrinos or left handed anti-neutrinos within SM. In weak interactions, some of the quantum numbers which are conserved in strong and EM interactions, need not be conserved, namely: parity (P), charge conjugation (C) and time reversal (T).

At high energy, weak and EM interactions become indistinguishable and those interactions are described by electro-weak (EW) theory [99][90][59] which obeys  $SU(2) \otimes U(1)$  symmetry. Since the force mediators are spin-1 bosons, the theory must obey gauge invariance. Based on the gauge principle, it is concluded that the observed  $\gamma$ ,  $W^\pm$  and Z are mixtures of  $W_1$ ,  $W_2$ ,  $W_3$ , and  $B$  fields,

$$W^\pm = (W_1 \mp iW_2) / \sqrt{2} \quad (1.1)$$

$$Z = -B \sin \theta_W + W_3 \cos \theta_W \quad (1.2)$$

$$\gamma = B \cos \theta_W + W_3 \sin \theta_W \quad (1.3)$$

where  $\theta_W$  is called the weak mixing angle and  $B$  is such that its predictions are consistent with properties of  $\gamma$  [78].

## 1.5 The Higgs mechanism

Local gauge invariance demands force mediators to be massless. Indeed, this is the case for QCD and QED in which gluon and photon are massless. But the weak force mediators  $W^\pm$  and Z bosons are observed to be massive and were discovered at CERN in 1983 [11][17][12][15]. A spontaneous symmetry breaking mechanism was proposed by Englert, Brout and Higgs [63][50] by which the gauge bosons can acquire mass through the interaction with a new scalar field, called Higgs field. The vacuum expectation value of the Higgs field is nonzero and when the gauge bosons interact with such a field the theory is not invariant under gauge transformations and the gauge bosons become massive. The ratio of masses of  $W^\pm$  and Z boson is given by

$$M_W/M_Z = \cos \theta_W \quad (1.4)$$

The theory predicts a spin 0 quanta (particle) associated with the Higgs field. This particle was discovered at CERN by ATLAS and CMS experiments in 2012 [1][32] and its mass was found to be  $\approx 125$  GeV. The masses of quarks and leptons would have been zero, if there was no Higgs boson. If the interaction of particle is stronger with Higgs field, then its mass is expected to be higher. Top quark is the highest mass

particle in SM and its coupling to Higgs is strongest. The interaction term between fermions and H field is given by  $\mathcal{L}_{Yukawa} = -\lambda_f \bar{\psi} \mathbf{H} \psi$  where  $\psi$  is the Dirac field of the fermion,  $\mathbf{H}$  is the Higgs field and  $\lambda_f$  is Yukawa coupling.

## 1.6 Limitations of SM

With the discovery of the Higgs (H) boson, the SM is a complete theory which describes many of the experimental observations. The expectations from the SM agree with the experimental observations and Fig.1.1 shows production cross sections for various SM processes and the measurements by the CMS experiment. Although H

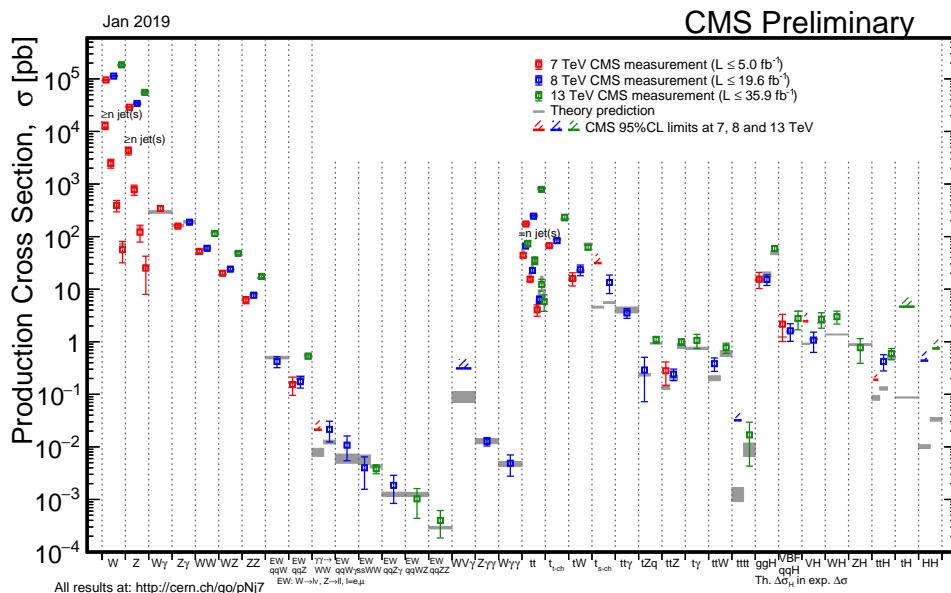


FIGURE 1.1: Summary of the cross section measurements of SM processes by CMS experiment [38].

was the last piece in the SM puzzle and we have discovered it, we are many theoretical issues and experimental results which are not explained by SM. From the theoretical point of view, observed mass of H at 125 GeV itself is an issue and it is called as hierarchy problem which will be discussed later. Some of the unsolved problems within SM framework are:

- Gravitational interactions cannot be explained by SM.
- Matter-antimatter asymmetry: We do not understand why we see more matter in the universe than antimatter today. At the early stages of universe, both of these must have been created in equal amount. There must be a mechanism by which matter started to dominate over antimatter.
- In SM masses of neutrinos is zero, whereas neutrino oscillations indicate that they have nonzero mass.
- The universe comprises of 71.4% of dark energy, 24% of dark matter and only 4.6% is ordinary matter that we know. What is explained by SM is less than 5% of total universe content.
- The strong, EM and weak gauge couplings are functions of energy scale. When these couplings are extrapolated to high energy, we expect all of them to unify at one energy. But this unification does not take place at one point within SM.

- The H boson gets corrections to its mass by loop diagram contributions, the largest contribution coming from top quark loop diagram (Fig.1.2). This contribution is given by

$$\Delta m_H^2 = -\frac{|\lambda_t|^2}{8\pi^2} [\Lambda_{UV}^2 + \dots] \quad (1.5)$$

where  $\lambda_t$  is the top Yukawa coupling ( $\sim 1$ ) and  $\Lambda_{UV}$  is the ultraviolet cutoff scale above which SM is not valid and its value is close to GUT (grand unified theory) scale,  $10^{16}$ GeV or Planck scale,  $> 10^{19}$  GeV. With these large corrections, the mass of H would have been  $> 10^{16}$  GeV and not near 125 GeV. This is called hierarchy problem in SM.

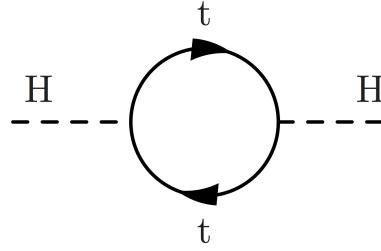


FIGURE 1.2: Top loop diagram contribution to H mass

## 1.7 Supersymmetric extension of SM

To overcome the limitations of SM mentioned in Sec.1.6, we need extensions to SM which can address all (or some) of these issues without contradicting existing observations. One such extension is supersymmetry (SUSY) which can address last three issues mentioned above, namely the dark matter problem, gauge coupling unification and hierarchy problem.

To tackle hierarchy problem, if we can introduce a new term in eqn.1.5 with similar correction but opposite sign, we might be able to get H mass around 125 GeV. Naively speaking, what SUSY does is exactly this - it introduces a scalar partner for every SM fermion and a fermionic partner for bosonic SM particle and hence the corrections from superpartner loop diagrams cancel the corrections from SM particles. Superpartners of SM particles differ in spin by half a unit. For example, superpartner of top quark is top squark(or stop) with spin 0. The superpartners of bosons have spin  $1/2$  and they are named with suffix *ino*, such as gluino, photino etc. Fig.1.3 shows the contributions to H mass from top loop and stop loop. The contribution from second loop diagram is given by eqn.1.6 which has similar form as eqn.1.5 but with an opposite sign.

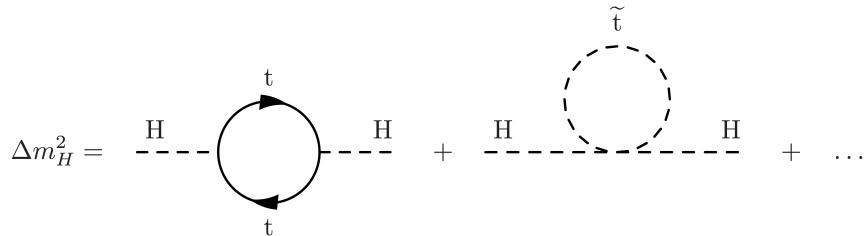


FIGURE 1.3: Corrections to H mass from top loop and stop loop.

$$(\Delta m_H^2)_{\text{SUSY}} = \frac{|\lambda_{\tilde{l}}|^2}{8\pi^2} [\Lambda_{UV}^2 + \dots] \quad (1.6)$$

We are interested in minimal supersymmetric extension of SM (MSSM) which is direct supersymmetrization of SM and has minimum number of new particle states and interactions consistent with phenomenology [14]. Table 1.3 shows fields corresponding to various particles in the MSSM. These superpartners listed are not necessarily the mass eigen states of the theory because there can be mixing of gauginos and higgsinos [79]. The observable mass eigen states are charginos or neutralinos, denoted by  $\tilde{\chi}^{\pm,0}$  and gluino which is not a mixture. Gauge and mass eigen states in MSSM are listed in table 1.4.

TABLE 1.3: Chiral supermultiplets and gauge supermultiplets in the MSSM [79]

	spin 0	spin $\frac{1}{2}$
squarks, quarks (x3 families)	$(\tilde{u}_L \tilde{d}_L)$ $\tilde{u}_R$ $\tilde{d}_R$	$(u_L d_L)$ $u_R$ $d_R$
sleptons, leptons (x3 families)	$(\tilde{\nu} \tilde{e}_L)$ $\tilde{e}_R$	$(\nu e_L)$ $e_R$
Higgs, higgsinos	$(H_u^+ H_u^0)$ $(H_d^0 H_d^-)$	$(\tilde{H}_u^+ \tilde{H}_u^0)$ $(\tilde{H}_d^0 \tilde{H}_d^-)$
	spin $\frac{1}{2}$	spin 1
gluino, gluon winos, W bosons bino, B boson	$\tilde{W}^\pm \tilde{W}^0$ $\tilde{B}^0$	$W^\pm W^0$ $B^0$

TABLE 1.4: Gauge and mass eigen states of MSSM [79][87]

Names	Spin	Gauge eigen states	Mass eigen states
squarks	0	$\tilde{u}_L \tilde{u}_R \tilde{d}_L \tilde{d}_R$	same
		$\tilde{c}_L \tilde{c}_R \tilde{s}_L \tilde{s}_R$	same
		$\tilde{t}_L \tilde{t}_R \tilde{b}_L \tilde{b}_R$	$\tilde{t}_1 \tilde{t}_2 \tilde{b}_1 \tilde{b}_2$
sleptons	0	$\tilde{e}_L \tilde{e}_R \tilde{\nu}_e$	same
		$\tilde{\mu}_L \tilde{\mu}_R \tilde{\nu}_\mu$	same
		$\tilde{\tau}_L \tilde{\tau}_R \tilde{\nu}_\tau$	$\tilde{\tau}_1 \tilde{\tau}_2 \tilde{\nu}_\tau$
neutralinos	$\frac{1}{2}$	$\tilde{B}^0 \tilde{W}^0 \tilde{H}_u^0 \tilde{H}_d^0$	$\tilde{\chi}_1^0 \tilde{\chi}_2^0 \tilde{\chi}_3^0 \tilde{\chi}_4^0$
charginos	$\frac{1}{2}$	$\tilde{W}^\pm \tilde{H}_u^+ \tilde{H}_d^-$	$\tilde{\chi}_1^\pm \tilde{\chi}_2^\pm$
gluino	$\frac{1}{2}$	$\tilde{g}$	same
Higgs bosons	0	$\tilde{H}_u^0 \tilde{H}_d^0 \tilde{H}_u^+ \tilde{H}_d^-$	$h^0 H^0 A^0 H^\pm$
gravitino	3/2	$\tilde{G}$	same

Since we have not observed SUSY particles at the same mass SM counterparts, SUSY is a broken symmetry and the masses of sparticles are larger than SM partners. The effective Lagrangian can be written as

$$\mathcal{L} = \mathcal{L}_{\text{SUSY}} + \mathcal{L}_{\text{soft}} \quad (1.7)$$

The first term on RHS of eqn.1.7 preserves SUSY and contains gauge and Yukawa interactions. The second term breaks SUSY and contains mass terms and coupling parameters which should vanish at very high mass scale at which SUSY is unbroken.

There are multiple ways to break SUSY and some of the popular ways are, gravity mediated or Planck scale mediated SUSY breaking (PMSB), anomaly mediated SUSY braking (AMSB) and gauge mediated SUSY breaking (GMSB). A brief discussion on GMSB is given in section 1.8 to motivate the search for SUSY with photon which is the topic of this thesis.

### 1.7.1 R-parity

A multiplicative quantum number called *R-parity* is introduced for every particle and sparticle to account for very strong experimental bounds on proton lifetime. The proton lifetime is  $> 10^{33}$  years [82] which is larger than the age of the universe. R-parity is defined as

$$P_R = (-1)^{3(B-L)+2s} \quad (1.8)$$

where s is the spin, B and L are baryon number and lepton number respectively. All SM particles have even R-parity,  $P_R = +1$  and SUSY particles have odd R-parity,  $P_R = -1$ . The important consequences of this parity conservation are:

- SUSY particles are always produced in pairs and any SUSY particle decay should involve even number of them. In other words, at any vertex, there should be even number of SUSY particles.
- Lightest supersymmetric particle (LSP) must be stable. If LSP is neutral, then it could serve as dark matter candidate and account for 24% of the universe content [49], thereby solving one of the problems in SM.
- All the non-LSP SUSY particle decay chains must result in odd number (usually 1) of LSPs.

## 1.8 Gauge mediated SUSY breaking

In GMSB scenarios [46, 47, 48, 80, 58, 60], the communication between hidden sector, where SUSY breaking takes place, and the visible MSSM sector (consisting of chiral supermultiplets shown in table 1.4) is via the ordinary gauge interactions. In comparison with other SUSY breaking scenarios, flavor changing processes and new sources of CP violation are naturally suppressed [46] in GMSB. The messengers communicating between MSSM and hidden sector also have  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$  interactions. The soft terms in MSSM come from loop diagrams involving these messengers, whose value is given by

$$m_{soft} \sim \frac{\alpha_a}{4\pi} \frac{\langle F \rangle}{M_{mess}} \quad (1.9)$$

where  $\alpha_a/4\pi$  is loop factor for Feynman diagrams involving gauge interactions,  $F$  relates to the SUSY breaking scale and  $M_{mess}$  is the messenger mass scale [79].

GMSB permits a significantly lower symmetry-breaking scale ( $\langle F \rangle$ ) than, e.g., gravity mediation, and therefore generically predicts that the gravitino ( $\tilde{G}$ ) is the LSP [80, 45, 42] whose value is given by

$$m_{\tilde{G}} \sim \langle F \rangle / M_P \sim \text{keV} \quad (1.10)$$

where  $M_P$  is the Planck scale where gravity is expected to become strong.

### 1.8.1 Phenomenology of GMSB

As mentioned above, gravitino ( $\tilde{G}$ ), superpartner of graviton, is the LSP, it is stable and weakly interacting and results in missing transverse momentum. Next to LSP (NLSP) is either neutralino ( $\tilde{\chi}_1^0$ ) or chargino ( $\tilde{\chi}_1^\pm$ ). The decay modes of NLSP is decided by its nature - bino, wino and higgsino components in it [74].

- For a **bino like** NLSP,  $|M_1| < |\mu|, |M_2|$ , where  $M_1, M_2$  and  $\mu$  are  $U(1)$  gauge mass parameter,  $SU(2)$  gauge mass parameter and higgsino mass parameter respectively. The decay mode of NLSP is  $\gamma/Z + \tilde{G}$  with larger branching ratio (BR) for  $\gamma + \tilde{G}$ . Left plot in Fig.1.4 shows BR for NLSP decay as a function of bino mass [89].

It is worth noting that the coupling of  $\gamma$  with  $\tilde{G}$  is at the tree-level because the gravitino has become massive after SUSY breaking and has goldstino component in it [79].

Experimentally, these kind of scenarios can be targeted using collider searches with  $\gamma\gamma + p_T^{miss}$  or  $\gamma + p_T^{miss}$  final states, where  $p_T^{miss}$  is the magnitude of missing transverse momentum.

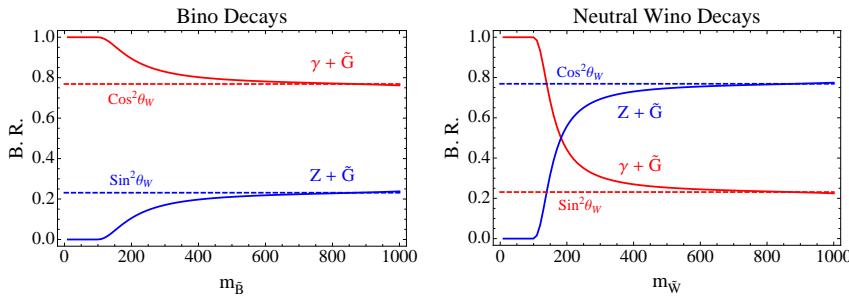


FIGURE 1.4: BR for bino and neutral wino NLSP decays. Plot is taken from Ref.[89]

- For a **wino like** NLSP,  $|M_2| < |\mu|, |M_1|$ . In this case,  $\tilde{\chi}_1^0$  and  $\tilde{\chi}_1^\pm$  are nearly mass degenerate and the decay modes are

$$\begin{aligned}\tilde{\chi}_1^\pm &\rightarrow W^\pm + \tilde{G} \\ \tilde{\chi}_1^0 &\rightarrow Z/\gamma + \tilde{G}\end{aligned}$$

The right plot in Fig.1.4 shows  $\tilde{\chi}_1^0$  decay as a function of wino mass [89]. These scenarios can result in signatures with lepton +  $\gamma + p_T^{miss}$ .

- For a **higgsino like** NLSP,  $|\mu| < M_1, M_2$  and different NLSP decay modes are preferred depending on the value of  $\mu$ .
    - If  $\mu < 0$ , then  $\tilde{\chi}_1^0 \rightarrow H/\gamma + \tilde{G}$  decay is preferred.
    - If  $\mu > 0$ , then  $\tilde{\chi}_1^0 \rightarrow Z/\gamma + \tilde{G}$  decay dominates.
- Models with higgsino like NLSP may result in  $b\bar{b}$ , coming from H decay, and  $\gamma + p_T^{miss}$  final states.

### 1.8.2 Simplified GMSB models

As the name suggests, various simplifications are done to the actual SUSY models and searches are designed based on the simplified model scenarios (SMS) [10, 8, 9, 5, 31]. There are many free parameters in the theory and as a result, there is no

easy way of making a prediction in an experiment. As an example, the nature of NLSP, whether it is a bino, wino, higgsino, or a mixture of these, is not predicted by the theory and hence the signatures expected in proton-proton (p-p) collisions at the LHC is not precisely known. We make certain justifiable assumptions and explicitly state these assumptions. Another advantage in using SMS is that it is easy to re-interpret the results by taking a different SUSY model.

The scenario of a natural SUSY spectrum with GMSB and  $R$ -parity conservation typically manifests as events with multiple jets, at least one photon, and large  $p_T^{miss}$ . Depending on the topology, these jets can arise from either light-flavored quarks ( $u, d, s, c$ ) or  $b$  quarks. We study four simplified models; example diagrams depicting these models are shown in Fig. 1.5. Three models involve gluino pair production (prefixed with T5), and one model involves top squark pair production (prefixed with T6). In the T5qqqqHG model, each gluino decays to a pair of light-flavored quarks ( $q\bar{q}$ ) and a neutralino. The T5bbbbZG and T5ttttZG models are similar to T5qqqqHG, except that the each pair of light-flavored quarks is replaced by a pair of bottom quarks ( $b\bar{b}$ ) or a pair of top quarks ( $t\bar{t}$ ), respectively. In the T5qqqqHG model, the  $\tilde{\chi}_1^0$  decays either to an SM Higgs boson and a  $\tilde{G}$  or to a photon and a  $\tilde{G}$ . The  $\tilde{\chi}_1^0 \rightarrow H \tilde{G}$  branching fraction is assumed to be 50%, and the smallest  $\tilde{\chi}_1^0$  mass considered is 127 GeV. In the T5bbbbZG and T5ttttZG models, the neutralinos decay to  $Z\tilde{G}$  and  $\gamma\tilde{G}$  with equal probability. The T6ttZG model considers top squark pair production, with each top squark decaying into a top quark and a neutralino. The neutralino can then decay with equal probability to a photon and a  $\tilde{G}$  or to a  $Z$  boson and a  $\tilde{G}$ . For the models involving the decay  $\tilde{\chi}_1^0 \rightarrow Z\tilde{G}$ , we probe  $\tilde{\chi}_1^0$  masses down to 10 GeV. All decays of SUSY particles are assumed to be prompt. In all models, the mass  $\tilde{G}$  is fixed to be 1 GeV, to be consistent with other published results of CMS experiment. For the parameter space explored here, the kinematic properties do not depend strongly on the exact value of  $\tilde{G}$ .

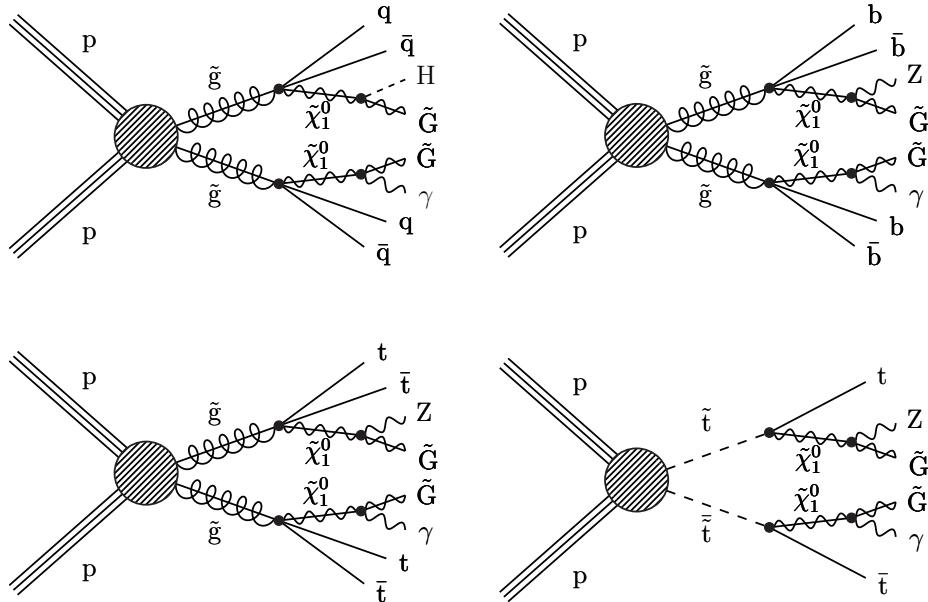


FIGURE 1.5: Example diagrams depicting the simplified models used, which are defined in the text. The top left diagram depicts the T5qqqqHG model, the top right diagram depicts the T5bbbbZG model, the bottom left diagram depicts the T5ttttZG model, and the bottom right depicts the T6ttZG model.

## Chapter 2

# The Experimental Setup

### 2.1 The large hadron collider

The large hadron collider (LHC) is a circular particle accelerator with a 27 km circumference. The LHC is designed to accelerate protons at an energy of 7 TeV [23]. During run 2 (year 2016 to 2018) each beam of protons were accelerated to an energy of 6.5 TeV and the total center of mass energy ( $\sqrt{s}$ ) of collision was 13 TeV. The process of achieving  $\sqrt{s} = 13$  TeV is done using various stages of CERN (European Organization for Nuclear Research) accelerator complex [62] shown in Fig.2.1.

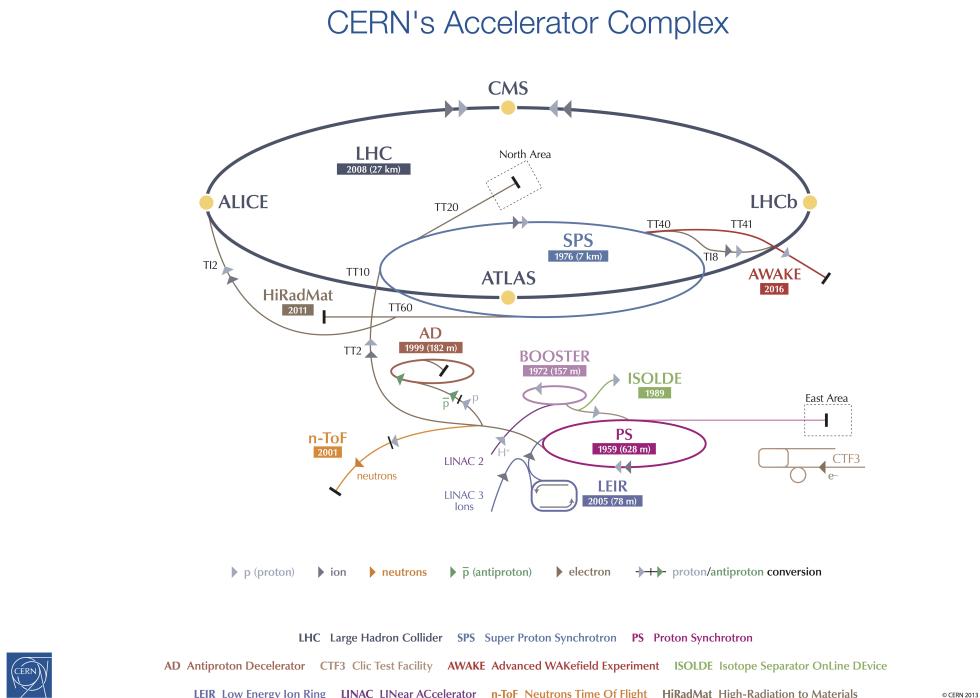


FIGURE 2.1: CERN accelerator complex

The source of protons is a bottle of hydrogen gas which is ionized using electric field. The protons are accelerated to 50 MeV using a linear accelerator (LINAC2). Next stage of acceleration is carried out by proton synchrotron booster (PSB) to reach 1.4 GeV. Proton synchrotron (PS) further accelerates protons to 25 GeV. After this, protons are sent into super proton synchrotron (SPS) to reach energy of 450 GeV and finally they enter into the LHC ring. The LHC also accelerates Pb ions apart from protons, but in the context of this thesis only protons are relevant.

In the LHC, 16 radio frequency (RF) cavities are used to accelerate protons. These RF cavities operate at a frequency of 400 MHz. Protons get accelerated from 450 GeV to 6.5 TeV in around 20 minutes, after passing through the cavities more than  $10^7$  times.

If the colliding beams are of the type particle-antiparticle, magnetic field in one direction can be used to bend the beams in opposite directions. Since LHC collides particle-particle beams, a twin-bore magnetic system is used instead of having two separate rings.

To steer the beam and keep in circular path, 1232 dipole magnets are used. Each of these magnets are about 15m long and made up of Niobium-Titanium (NbTi) superconducting coils. NbTi has critical temperature,  $T_C$  of 10 K and it is operated at 1.9 K. A current of 11.08 kA in these coils generates a magnetic field of 8.3 T.

Quadrupoles are used to focus the beam either horizontally or vertically. They have 4 magnetic poles arranged symmetrically around the beam pipe and also equipped with sextupole, octupole and decapole magnets to correct for small imperfections in the magnetic field.

The beams collide at 4 points where ALICE, ATLAS, CMS and LHCb detectors are located. Before the collision, the beams are made narrower down to 16  $\mu\text{m}$  using a set of quadrupoles magnets. Table 2.1 shows various parameters of the LHC.

TABLE 2.1: The LHC parameters [77]

Quantity	Value
Circumference	26.659 km
Dipole operating temperature	1.9 K (-271.3 °C)
Number of magnets	9593
Number of main dipoles	1232
Number of main quadrupoles	392
Number of RF cavities	8 per direction
Energy of protons (year 2016-2018)	6.5 TeV
Peak magnetic dipole field	7.74 T
Distance between bunches	~ 7.5 m
Peak luminosity (protons)	$\sim 1.2 \times 10^{34} \text{ cm}^{-2} \text{s}^{-1}$
No. of bunches per proton beam (design value)	2808
No. of protons per bunch (at start)	$1.2 \times 10^{11}$
Number of turns per second	11,245
Number of collisions per second	$10^9$

## 2.2 The compact muon solenoid

Surrounding one of the collision points of LHC is the compact muon solenoid (CMS) detector. CMS is a general purpose detector designed to search for new phenomena, test SM and to study properties of higgs boson. It is a cylindrical superconducting solenoid providing 3.8 T axial magnetic field and has inner diameter of 6 m. Detailed description of the CMS detector can be found in [28]. A section of the CMS detector with different parts is shown in Fig. 2.2. The search described in this thesis uses data collected by CMS during the proton-proton (p-p) collision at a  $\sqrt{s} = 13$  TeV in the year 2016. Total integrated luminosity ( $\mathcal{L}$ ) as delivered by LHC and recorded by the CMS detector as a function of time is shown in Fig. 2.3 for the year 2016 [37]. The total data certified as good for physics analysis corresponds to  $\mathcal{L} = 35.9 \text{ fb}^{-1}$ .

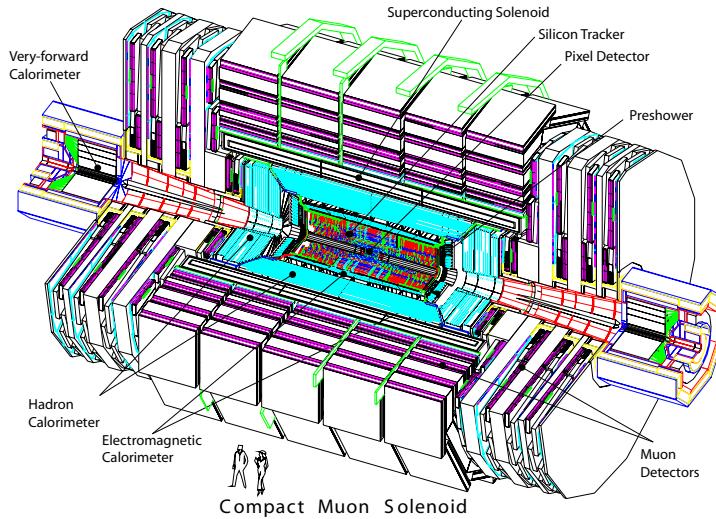


FIGURE 2.2: A view of the CMS detector

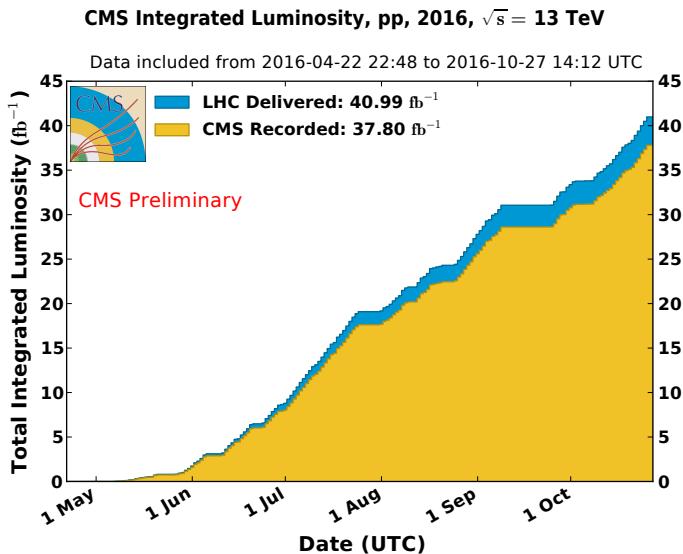


FIGURE 2.3: Total integrated luminosity as delivered by LHC and recorded by CMS as a function of time

### 2.2.1 CMS coordinate system

A right handed cartesian coordinate system adopted by the CMS has origin at the nominal collision point. The  $x$ -axis points towards inner side of the LHC ring,  $y$  points upwards and  $z$  axis is along the beam direction. The azimuthal angle  $\phi$  is measured from  $x$ -axis in  $x - y$  plane and  $r$  is the radial distance. Psudorapidity is defined as  $\eta = -\ln(\tan(\theta/2))$ , where  $\theta$  is the polar angle measured from  $z$ -axis. Psudorapidity is an approximation of rapidity variable defined as  $\frac{1}{2} \ln((E + p_z)/(E - p_z))$  in the limit  $E \approx |\mathbf{p}|$ , where  $E$  and  $\mathbf{p}$  are energy and momentum respectively. Some of

the other quantities which are used in this work are listed below.

$$\text{Transverse momentum, } p_T = \sqrt{p_x^2 + p_y^2} \quad (2.1)$$

$$\text{Distance in } \eta - \phi \text{ plane, } \Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad (2.2)$$

## 2.2.2 Superconducting magnet

The 12.5 m long superconducting magnet in the CMS detector is made up of NbTi coil and capable of providing magnetic field up to 4 T (operated to provide 3.4 T). Some of the goals of LHC physics program is to have a good momentum resolution of muons and other charged particles, tag tau leptons and b-quark jets etc. The magnetic field plays a crucial role in measuring the momentum of charged particles by the tracker and outer muon systems, and also in identifying the charge of particles. The magnetic flux returns through a 10 kilo-ton iron yoke. Total energy stored in the magnet is 2.6 GJ at 4 T magnetic field and the energy stored per unit cold mass is 11.6 kJ/kg which is higher than the values of any of the magnets used in particle physics detectors.

## 2.2.3 Inner tracking system

The tracking system is used to measure the momentum of the charged particles, determine their trajectory and to help in locating primary and secondary vertices. At a luminosity of  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , on an average 1000 particles pass through the tracker every 25 ns. Tracker uses about  $200 \text{ m}^2$  of Si, has length of 5.8 m and diameter of 2.5 m and covers up to  $|\eta| = 2.5$ .

It is composed of inner pixel detector consisting of 3 layers at radii 4.4 cm, 7.3 and 10.2 cm in the barrel region. The other component is the Si strip tracker which extends up to 1.1 m in barrel region with 10 layers in it. On either side of the barrel are the endcaps which have 2 disks in the pixel detector and 3 plus 9 disks in the strip tracker. There are in total 1440 modules in pixel detector with 66 million pixels; 15148 modules in strip tracker with 9.3 million strips. A part of cross sectional schematic view of the tracker [40] is shown in Fig. 2.4 with different modules namely: pixel, inner barrel (TIB), outer barrel (TOB), inner disks (TID) and endcaps (TEC).

The hit position resolution provided by the pixel detector is about  $10 \mu\text{m}$  and  $20-40 \mu\text{m}$  in transverse and longitudinal coordinate respectively and the third coordinate is determined from the sensor plane position. The pixel and some of the strip

TABLE 2.2: Characteristics of subsystems of tracker. Pitch for the strip tracker refers to the distance between neighboring strips

Subsystem	Layers	Location (cm)	Pitch	Position resoln.
Pixel barrel	3 cylindrical	$r : 4.4 - 10.2$	$100 \times 150 \mu\text{m}^2$	$10 \mu\text{m}$ in trans.
Pixel endcap	2 disks	$z : 34.5 - 46.5$		$20 - 40 \mu\text{m}$ longt.
TIB	4 cylindrical	$r : 20 - 55$	$80 - 120 \mu\text{m}$	$13 - 38 \mu\text{m}$ in $r\phi$
TOB	6 cylindrical	$r : 55 - 116$	$122 - 183 \mu\text{m}$	$18 - 47 \mu\text{m}$ in $r\phi$
TID	3 disks	$z : 58 - 124$	$100 - 141 \mu\text{m}$	$13 - 38 \mu\text{m}$ in $r\phi$
TEC	9 disks	$z : 124 - 282$	$97 - 184 \mu\text{m}$	$18 - 47 \mu\text{m}$ in $r\phi$

modules (shown in thick blue lines in Fig.2.4) are capable of providing 3-D hits and the strip modules which provide 2-D hits are shown in thin black line in the figure. The strip modules which provide 3-D hits consist of 2 back-to-back strips in them.

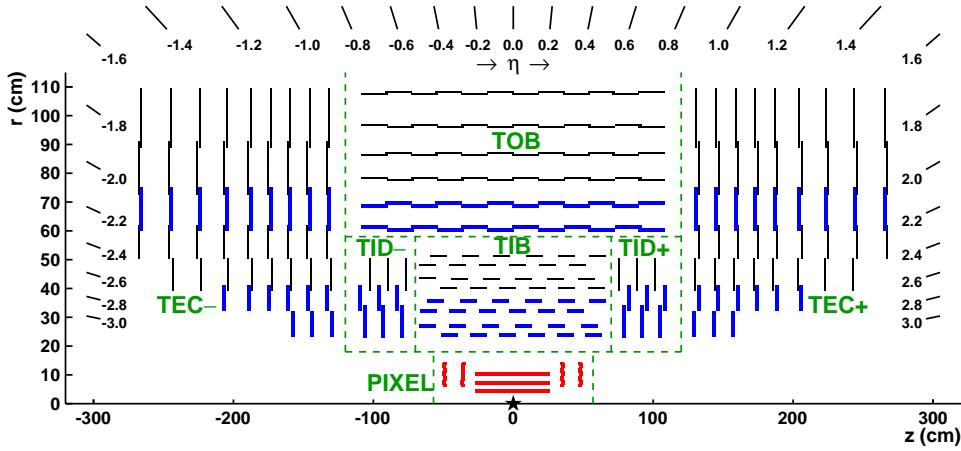


FIGURE 2.4: A schematic view of the CMS tracker in  $r - z$  plane showing different modules. The tracker is symmetric about  $r = 0$  line and the figure shows only the upper part.

### Performance of tracker

Figure 2.5 shows the resolution of various quantities for single muons of  $p_T$  1, 10 and 100 GeV as a function of  $|\eta|$ .

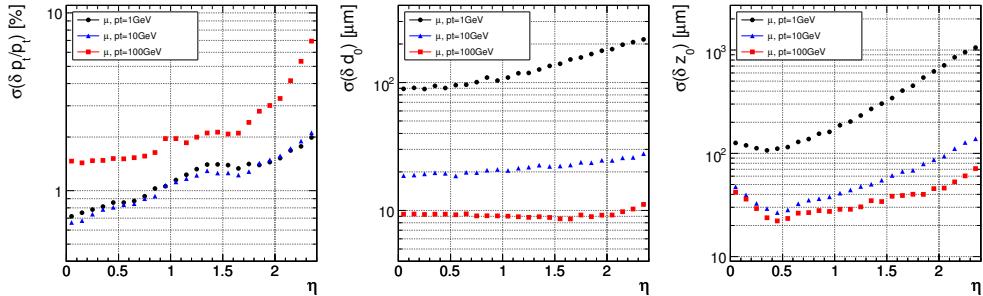


FIGURE 2.5: Resolution of various track parameters for muons with  $p_T = 1$  GeV (black), 10 GeV (blue) and 100 GeV (red). Left plot shows  $p_T$  resolution, middle plot shows transverse impact parameter resolution and right plot shows longitudinal impact parameter resolution.

For 100 GeV muons,  $p_T$  resolution is about 1-2% up to  $|\eta| \approx 1.6$ . For lower  $p_T$  muons, resolution mainly is affected by multiple scattering and for high  $p_T$  this effect is about 20-30%. Multiple scattering of low  $p_T$  muons degrades the impact parameter resolution as well and this effect is reduced for high  $p_T$  muons and hence these have better impact parameter resolution.

Track reconstruction efficiency is about 99% for muons with  $p_T$  in the range 1-100 GeV (left plot in Fig.2.6) and slight degradation for some  $|\eta|$  regions is because of gaps or non-coverage of the tracker. For charged pions (right plot in Fig.2.6) the efficiency is lower because of interactions with the materials of the tracker. Thickness  $t$  of the tracker in terms of radiation length,  $X_0$  (left) and interaction length,  $\lambda_I$  (right) is shown in Fig.2.7 along with the supporting systems and beam pipe contributions using simulation. In an ideal case, the thickness of the tracker should be zero.

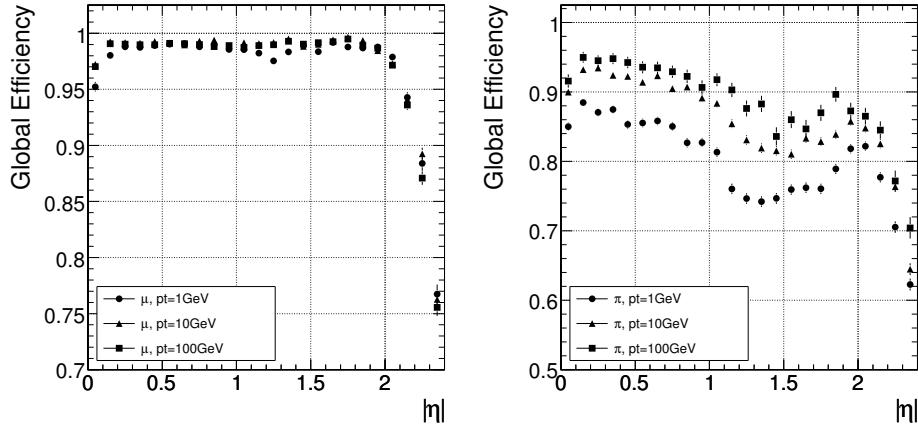


FIGURE 2.6: Global track reconstruction efficiency for muons (left) and pion (right) with  $p_T$  of 1, 10 and 100 GeV.

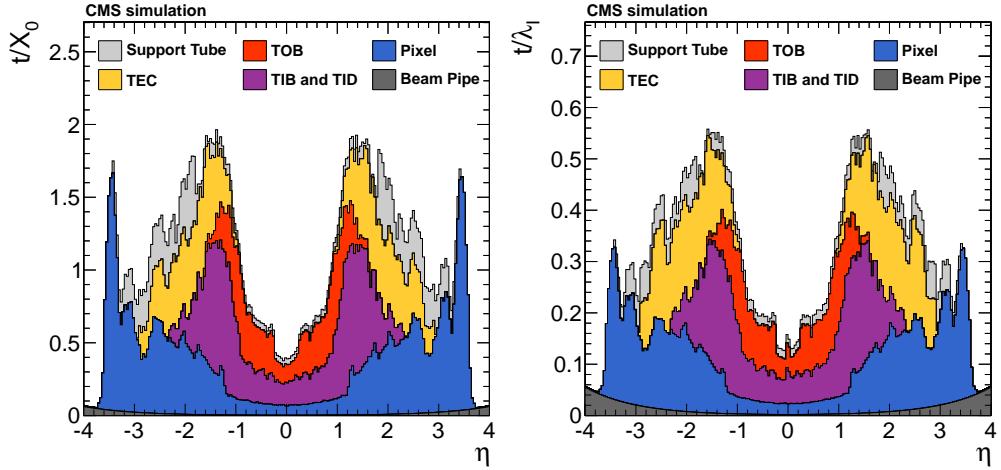


FIGURE 2.7: Thickness of tracker along with beam pipe and supporting systems in terms of radiation length,  $X_0$  (left) and interaction length,  $\lambda_I$  (right).

#### 2.2.4 Electromagnetic calorimeter

The EM calorimeter (ECal) of CMS is a hermetic, compact, granular, homogeneous, radiation tolerant and total absorption calorimeter made up of lead tungstate crystals ( $PbWO_4$ ). It measures energy of photons, electrons and EM component of jets. It has large dynamic range coupled with excellent linearity up to 1 TeV. It also provides triggering information and aids particle identification. The region with  $|\eta| \leq 1.48$  is covered by barrel (EB) and  $1.48 < |\eta| < 3$  region is covered by two endcap (EE) calorimeters. To identify neutral pions ( $\pi^0$ ), a preshower (ES) detector is used in the endcaps with  $1.653 < |\eta| < 2.6$ . The radiation length of EB is  $26X_0$ , EE is  $25X_0$  and that of ES is  $3X_0$ .

Incident electrons and photons produce EM showers which spread laterally over several crystals. Charged particles in these showers produce blue-green scintillation light (420–430 nm) and the amount of light is proportional to the incident particle energy. About 80% of the scintillation light is emitted in 25 ns. The crystals have

radiation length of 0.89 cm and small Molière radius (2.2 cm) which help in containment of shower in smaller volume. The scintillation light is collected by avalanche photo-diodes (APDs) in case of EB and vacuum photo-triodes (VPTs) in case of EE.

Fig.2.8 (left) shows the schematic layout of ECal modules, supermodules and supercrystals. The granularity of crystals in EB is  $0.0174 \times 0.0174$  in  $\eta - \phi$ . Two rows of 5 crystals form a submodule, A supermodule is formed using 4 modules. Almost all the crystal axes are tilted by  $3^\circ$  with respect to the line from nominal interaction point in both  $\eta$  and  $\phi$  directions to avoid any particle passing through the gap between crystals.

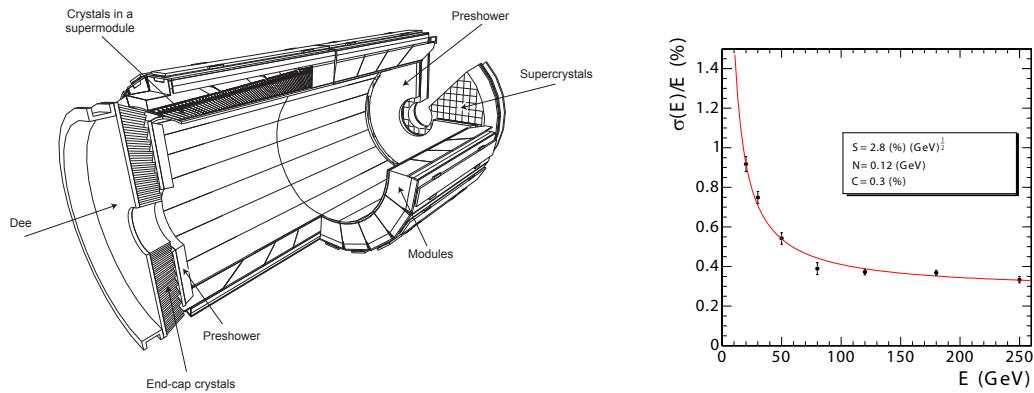


FIGURE 2.8: A schematic of ECal showing different components is shown on the left. Right plot shows resolution of ECal as a function of electron energy using test beam studies.

The energy resolution of ECal is given by the expression,

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (2.3)$$

where  $S$  is the stochastic term,  $N$  is the noise term and  $C$  is the constant term. Typical values after summing energy in  $3 \times 3$  crystals are  $S = 2.8\%$ ,  $N = 0.12$ ,  $C = 0.3\%$ . The resolution as a function of electron energy,  $E$  measured in GeV is shown in Fig.2.8 using test beam studies. The contributors of  $S$  are shower fluctuations and photo-statistics; for  $N$  the contributors are non-uniformity of longitudinal light collection, intercalibration errors and energy leakage; noise from electronics, digitization and pileup contribute to constant term.

A laser monitoring system is used to measure the transparency loss in the crystals which occurs because of radiation damage. This system uses blue laser of wavelength 440 nm which is closer to the scintillation peak and the intensity of the output light from each of the crystals is used to correct the transparency loss and re-calibrate the crystals.

### 2.2.5 Hadron calorimeter

This calorimeter (HCal) is a sampling calorimeter which measures energy of the hadrons and extends up to  $|\eta| = 5.2$ . It is divided into barrel (HB), endcap (HE), forward calorimeter (HF) and outer calorimeter (HO).

The pseudorapidity coverage of HB is  $|\eta| < 1.3$  with a granularity width of 0.087 in both  $\eta$  and  $\phi$ . This sampling calorimeter has alternating brass (70% Cu and 30% Zn) absorber and plastic scintillators. The absorber thickness is  $5.82\lambda_L$  at  $90^\circ$  and

$10.6\lambda_I$  at  $|\eta| = 1.3$ . The ECal in front of HB has  $1.1\lambda_I$ . Fig.2.9 shows a schematic diagram of one quarter slice of HCal. Because of space constraint from the magnet,

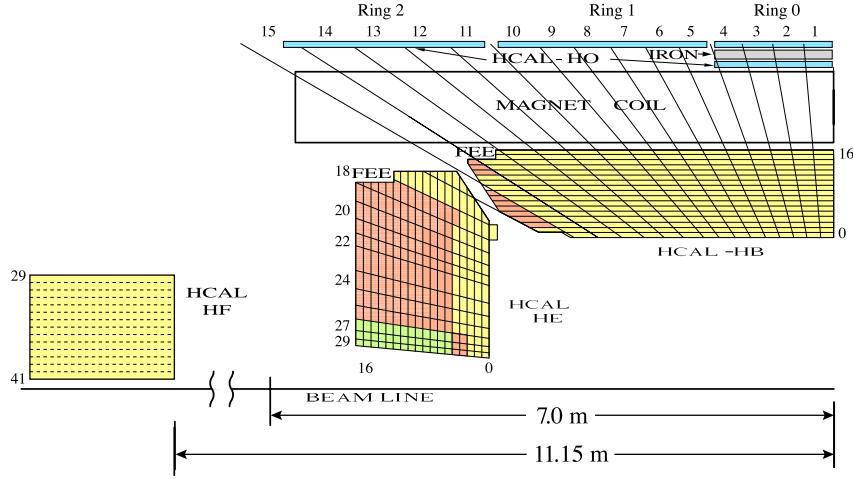


FIGURE 2.9: A schematic diagram of one quarter slice of HCal showing HB, HE, HO and HF along with location of front end electronics (FEE) of HB and HE. The colors represent longitudinal readout scheme.

HB thickness in the low  $|\eta|$  region is kept small, and to record the leakage of hadronic showers, 5 rings of HO are placed outside the magnet. About 5% of all hadrons of energy more than 100 GeV deposit energy in HO.

The scintillation light is collected by the wavelength shifting (WLS) optical fiber which is laid out on the edge of the scintillator tile. WLS fiber is spliced into clear fiber and the light is guided to hybrid photo-diode (HPD) using optical cable. HF uses Cherenkov based quartz fibers instead of scintillators and light from the quartz fibers is fed into photo-multiplier tubes (PMTs). More details about HF is discussed in later part of this chapter.

The signals from HPDs or PMTs are read by charge integration and encoding (QIE) chips. QIE is an ADC which provides almost constant precision over large range by assigning different number of bits to different amounts of charge (which is proportional to energy) collection.

All of the subsystems have LEDs and lasers for the purpose of calibration and monitoring. Except HO, other subsystems are also equipped with radioactive source tubes.

The resolution of HCAL and ECAL setup was determined using test beam setup and it found to be of the form:

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus b \quad (2.4)$$

where  $a$  is stochastic term of value  $0.847 \pm 0.016 \text{ GeV}^{\frac{1}{2}}$  and  $b$  is a constant term of value  $0.074 \pm 0.008$  for HB and HE. For HF,  $a = 1.98 \text{ GeV}^{\frac{1}{2}}$  and  $b = 0.09$  [39].

A study on HF calorimeter performance in 2016 using energy deposits in HF quartz fibers is described in 2.3.

## 2.2.6 Muon chambers

Muons are detected using gas filled muon chambers located outside the magnet. The direction of magnetic field is opposite in direction to that of the field inside the

magnet. The muon chambers have the task of identifying momentum of muons, type of charge and provide triggering information. This system consists of barrel region, up to  $|\eta| = 1.2$ , made up of drift tubes (DT) and two endcaps on either side, covering  $0.9 < |\eta| < 2.4$ , made up of cathode strip chambers (CSC). Resistive plate chambers (RPSc) are mounted in both barrel and endcap and operated in avalanche mode. They are capable of providing excellent timing resolution needed for muon triggers and they cover  $|\eta| < 1.9$ . Fig.2.10 shows a quarter of CMS detector with different muon chambers and their respective locations [92].

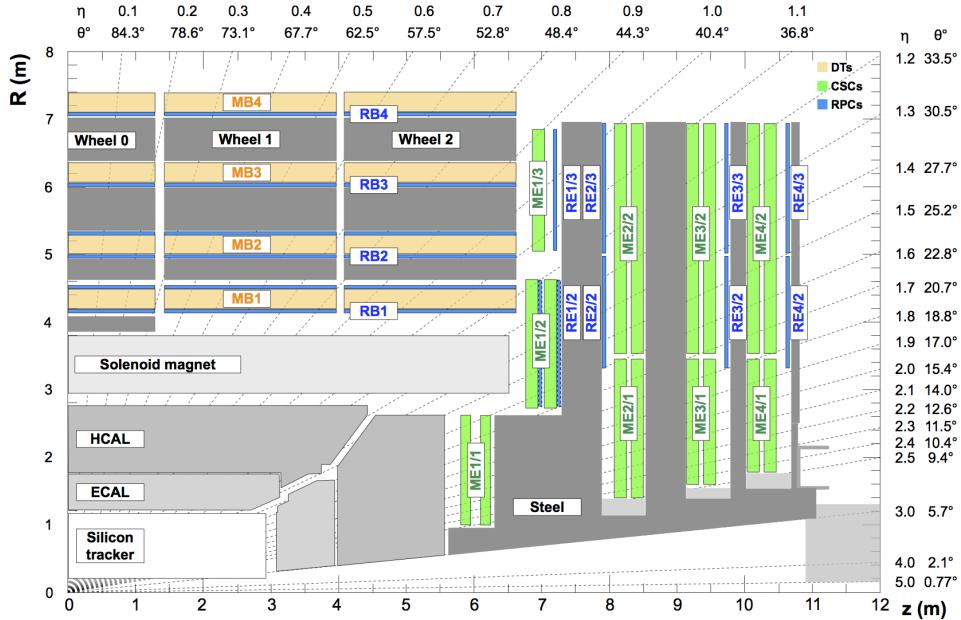


FIGURE 2.10: Schematic diagram of a quarter of CMS detector showing different components of muon chambers. Steel-flux return disks are shown as dark areas, DTs are labelled as MB, CSCs are labelled as ME, and RPCs are labelled as RB and RE. In this labelling *B* refers to barrel and *E* refers to endcap.

As the muon passes through the chambers, it ionizes gas in these chambers. Each of the DTs chambers have drift cells of size  $42 \times 13 \text{ mm}^2$  in transverse direction. The CSCs are operated as standard multi-wire proportional counters. More detailed description of muon system and its performance can be found in Ref.[92, 28, 34].

The  $p_T$  resolution of muons is approximately 1% in barrel and 3% in endcap for  $p_T < 100 \text{ GeV}$  [92]. The combined momentum measurement from tracker and muon system significantly improves the resolution and Fig.2.11 shows  $p_T$  resolution as a function of  $p_T$  of muon.

## 2.3 Energy response in long and short fibers of HF calorimeter

The HF is located at 11.3 m from the interaction point and provides a pseudo-rapidity coverage from  $|\eta| > 3.0$  to  $|\eta| < 5.2$ . Since there is no coverage from the tracker or ECAL in this region, the energy deposited in HF is used to reconstruct forward jets and also to calculate  $p_T^{miss}$ . Hence a stable performance of HF is important both for SM measurements and new physics searches.

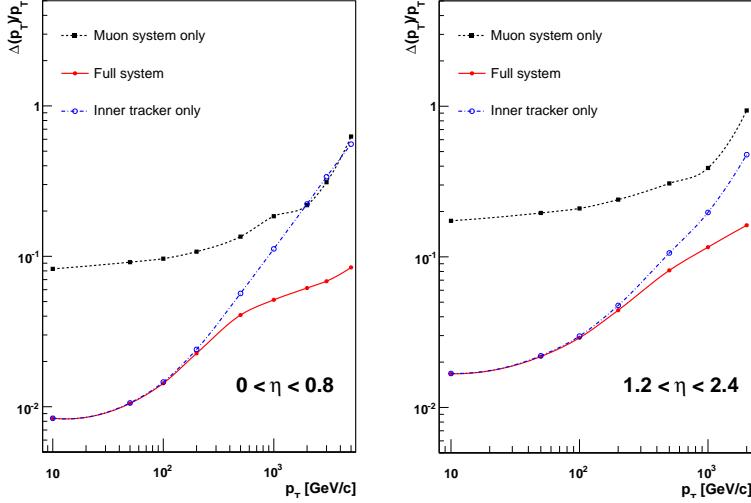


FIGURE 2.11:  $p_T$  resolution of muons as a function of  $p_T$  for  $|\eta| < 0.8$  (left) and  $1.2 < |\eta| < 2.4$  using tracker only, muon system only and combination of both [28].

The quartz fibers are inserted into the steel absorber plates. Half of the fibers run over the full depth of the HF ( $165\text{ cm} \approx 10$  interaction lengths) while the remaining half of the fibers start at a depth of  $22\text{ cm}$  from the front face of the detector. The former are called **long fibers** and latter are called **short fibers**. The energies deposited in long and short fibers are read out separately, and are called  $E_{\text{Long}}$  and  $E_{\text{Short}}$  respectively.

The signal in HF is due to the Cherenkov light produced by charged particles as they traverse through the quartz fibers. The Cherenkov light is emitted only when the particle's velocity is greater than the speed of light in that medium. If the particle has sufficient energy, it can travel faster than light in that medium and produce Cherenkov light. This light is collected by the long and short optical fibers. The recordable signal in calorimeter is due to the electromagnetic and hadronic component of the particle showers. Since electrons or photons result in shorter showers, these result in signal mostly in long fibers. The hadronic showers, however, continue deeper and result in signal in both long and short fibers.

The ratio of energy measured in short and long fibers,  $R_{S/L} = E_{\text{Short}}/E_{\text{Long}}$ , depends on the energy of the incident particle which created the shower and hence on how deeply the shower has penetrated the calorimeter. However, the average  $R_{S/L}$  over a period of time in a given  $\eta$  region is expected to depend on average energy incident on the calorimeter and accelerator run conditions (which determines the pileup, the number of pp interactions). In this section, we describe studies  $R_{S/L}$  for data collected at  $\sqrt{s} = 8\text{ TeV}$  and  $\sqrt{s} = 13\text{ TeV}$ , and also the effect of pileup. Since the average  $R_{S/L}$  for various channels of a given  $i\eta$  ring is expected to be same, this quantity can be used to intercalibrate the short fibers across  $\phi$  while the long fibers are calibrated using  $Z \rightarrow e^+e^-$  events.

### 2.3.1 Data and simulation samples

These studies make use of data collected by CMS detector in the years 2012, 2015 and 2016. Each of these are divided into different parts and they are named by adding a suffix to the year, for example 2012D is a subset of data collected in 2012. The events from these dataset are selected using triggers which are based on hadronic activity,

calculated using sum  $p_T$  of all jets or using highest  $p_T$  jet in the event. The dataset which contains hadronic triggers is called as *JetHT* dataset. Only those data which are certified as good for physics analysis are used. Table 2.3 shows list of datasets used for this study along with pp bunch spacing and integrated luminosity of the dataset. Monte-carlo (MC) simulated sample consists of QCD events generated at leading order (LO) taking  $\sqrt{s} = 13$  TeV using MadGraph generator and hadronization is carried out using Pythia8.

TABLE 2.3: Collision data used for  $R_{S/L}$  studies of HF. The 2012 data were taken at  $\sqrt{s} = 8$  TeV and all other data were taken at  $\sqrt{s} = 13$  TeV.

Data	Bunch spacing (ns)	$\mathcal{L} (\text{pb}^{-1})$
2012D	50	962
2015B	50	40.9
2015C		25.0
2015C		16.3
2015D	25	$1.61 \times 10^3$
2016B		$5.28 \times 10^3$
2016C		789
2016D		$3.28 \times 10^3$
2016E	25	$4.05 \times 10^3$
2016F		$3.11 \times 10^3$
2016G		$7.11 \times 10^3$
2016H		$8.68 \times 10^3$

### 2.3.2 Energy in long and short fibers of HF

The HF starts at  $|\eta|=2.853$  and extends up to  $|\eta|=5.191$ . On both the  $\pm z$  sides, this  $\eta$  range is divided into 13 towers with tower index starting from  $i\eta=29$  and ending at  $i\eta=41$ . Towers  $|i\eta|=29$  to  $|i\eta|=39$  have 36 divisions ( $10^\circ$  each) in  $\phi$  and towers  $|i\eta|=40$  and 41 have 18 divisions ( $20^\circ$  each). In total there are 864 channels. Depth segment with index 1 corresponds to long fibers and depth segment with index 2 refers to short fibers in each of these channels. If the energy in a particular channel is above the noise level, then that channel is considered to have a *rechit* (recorded hit).

All the plots and discussion till the end of section 2.3.3 correspond to 2015C-50ns data.

- Fig. 2.12 shows a typical distribution of number of rechits, with rechit energy  $> 10$  GeV, in each of the channels in depth 1 (long fibers) and depth 2 (short fibers).
- Fig. 2.13 shows the total number of rechits distribution inclusive in  $i\eta$  and  $i\phi$  with rechit energy  $> 10$  GeV.
- Fig. 2.14 shows the energy distribution in short fibers ( $E_{\text{Short}}$ ) vs long fiber ( $E_{\text{Long}}$ ) without any threshold on their recorded energies. This figure shows that there is correlation between  $E_{\text{Long}}$  and  $E_{\text{Short}}$ . So we use  $R_{S/L}$  as a tool for studying the performance of HF.

There two more important points to note from this plot: firstly, there are cases with a large  $E_{\text{Long}}$  while there is almost no energy deposited in corresponding

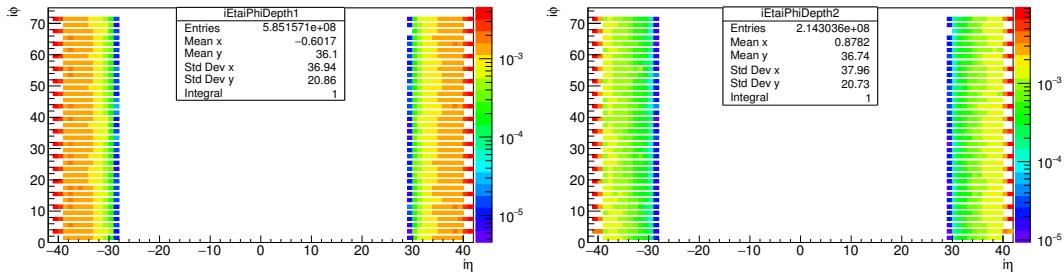


FIGURE 2.12: Channel occupancy for depth 1 (left), and depth 2 (right).

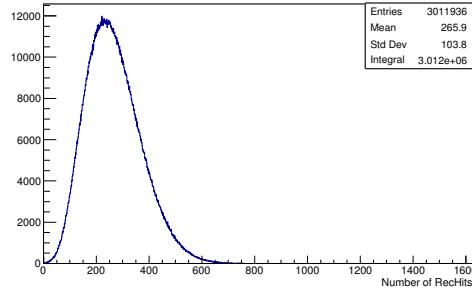


FIGURE 2.13: Total number of rechits (inclusive in  $i\eta$  and  $i\phi$ )

short fiber i.e.  $E_{\text{Short}} < 10 \text{ GeV}$ . This can be due to EM showers. Secondly, there are cases when there is large energy deposited in short fibers while small energies in long fibers ( $E_{\text{Long}} < 30 \text{ GeV}$ ). It is not expected to have large energy deposits in only one of the fibers if the shower originates from hadrons. One of the reasons for this could be that some of the high energy particles directly hit the glass window of PMT and resultant Cherenkov light produced in the glass gives rise to a large signal in only one of the channels. To reject such hits, thresholds are placed on reconstructed energies of rechits,  $E_{\text{Short}} > 10 \text{ GeV}$  and  $E_{\text{Long}} > 40 \text{ GeV}$ . From this point onwards, one can assume that these threshold have been applied on  $E_{\text{Short}}$  and  $E_{\text{Long}}$  unless a different selection is explicitly mentioned.

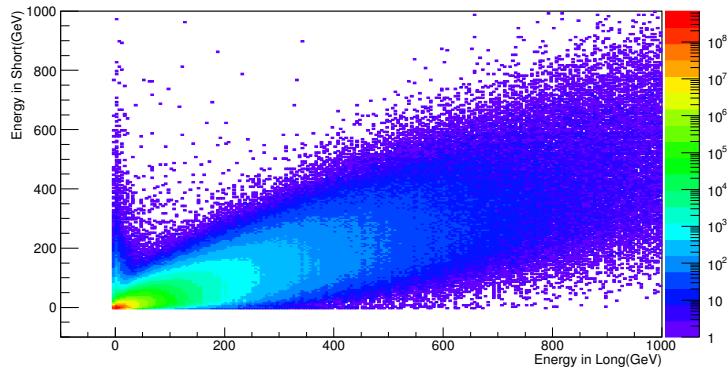


FIGURE 2.14: Distribution of  $E_{\text{Short}}$  vs  $E_{\text{Long}}$  for HF channels.

- With the energy thresholds mentioned above,  $R_{S/L}$  is studied for each  $i\eta$  tower.

The Fig.2.15 shows distribution of the  $R_{S/L}$  for the tower  $i\eta = 32$  integrated over all  $i\phi$  channels. Since the mean value of the distribution is sensitive to the tails, we try to fit it with an asymmetric Gaussian (eqn.2.5) and use the peak value obtained from the fit to indicate the average ratio for a given  $i\eta$  ring.

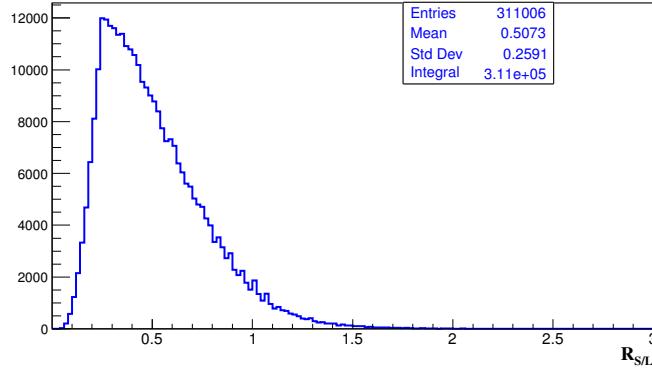


FIGURE 2.15:  $R_{S/L}$  for  $i\eta$  32 tower with  $i\phi$  inclusive for 2015D-50 ns.

$$f(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \left[ 1 + erf \left( \frac{x-\mu}{\sqrt{2}\sigma} \right) \right] \quad (2.5)$$

Fitting the  $R_{S/L}$  with an asymmetric Gaussian works very well for smaller  $i\eta$  channels (Fig. 2.16). For the channels in higher  $i\eta$  regions, fitting does not work well and  $\chi^2/dof$  is very large (Fig. 2.17). Changing the fit range did not improve the fits significantly. So using the peak obtained from the fits cannot be used for studying all the channels. We then use 90% truncated mean of the distribution - starting from the arithmetic peak, bin contents are added on both the sides until the total integral is 90% of the total area under the curve. From this point onward, mean refers to 90% truncated mean.

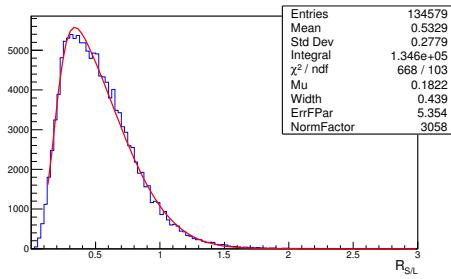


FIGURE 2.16: Asymmetric Gaussian fit for  $i\eta$  30

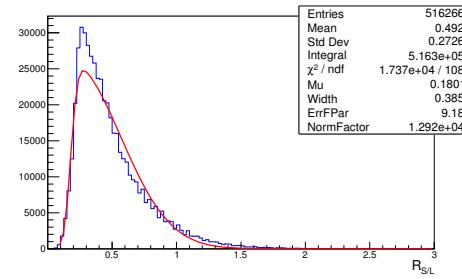


FIGURE 2.17: Asymmetric Gaussian fit for  $i\eta$  39

### 2.3.3 Effect of Pileup on $R_{S/L}$

The  $R_{S/L}$  was studied under different pileup (PU) scenarios. If the number of good primary vertices is less than 9, it is considered as low pileup; between 11 and 16 as medium pileup and above 18 as high pileup.

Pileup is mainly dominated by low energy particles and these particles do not have enough energy to penetrate into the short fibers and hence they deposit energy mainly in long fibers. Hence these particles give lower  $R_{S/L}$  than the

high energy particles. Pileup is added on top of a hard scattered events, and in hard scattered event  $R_{S/L}$  is higher than PU events. Thus with increase of more and more pileup events, the  $R_{S/L}$  is decreasing and that is the observation. With the high pileup, average energy deposits in long fibers is expected to be higher and as a result the  $R_{S/L}$  is expected to be lower as compared to the low pileup scenario. Fig. 2.18 shows that higher pileup means lower ratio and lower pileup means higher ratio. The shaded region corresponds to all events wherein no restriction on pileup is applied.

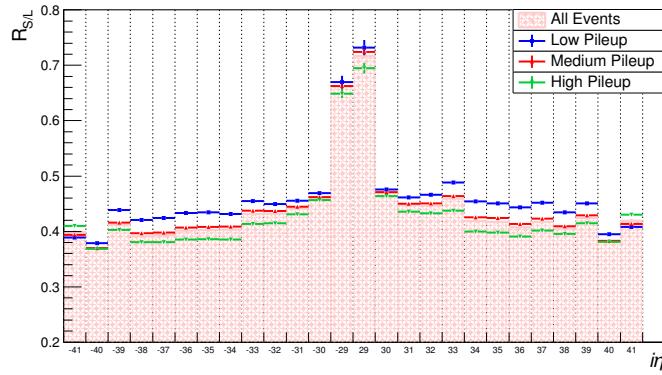


FIGURE 2.18: Effect of pileup on  $R_{S/L}$  for different  $i\eta$  rings (integrated over  $i\phi$  channels).

### 2.3.4 Studying $R_{S/L}$ at $\sqrt{s} = 13$ TeV and 8 TeV data with 50 ns bunch spacing

- In 2015, LHC started 13 TeV collisions with 50 ns bunch spacing. We compared  $R_{S/L}$  for 13 TeV data taken in 2015 and 8 TeV data taken in 2012.
- Used JetHT dataset for 13TeV data and 8 TeV with all the events (no trigger based selection of events) for this study.
- Average pile up was similar (Fig. 2.19) in 8 and 13 TeV datasets.

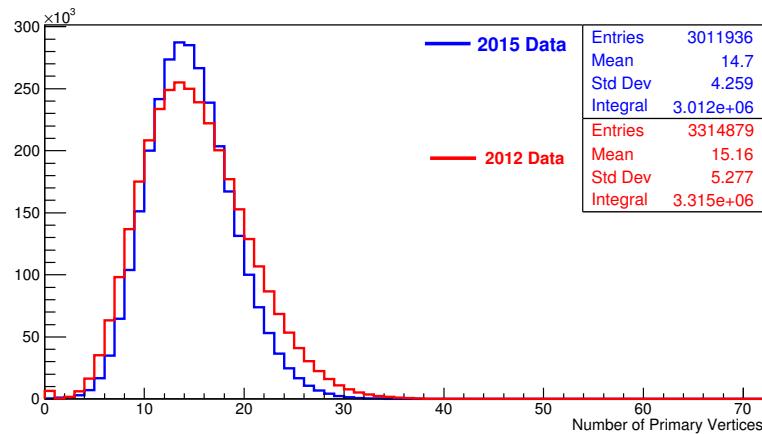


FIGURE 2.19: Pileup comparison for 2015 data and 2012 data

- Fig. 2.20 and 2.21 show the distributions of energies in long and short fibers, for 2012D and 2015C, with  $E_{\text{Long}}, E_{\text{Short}} > 10$  GeV.

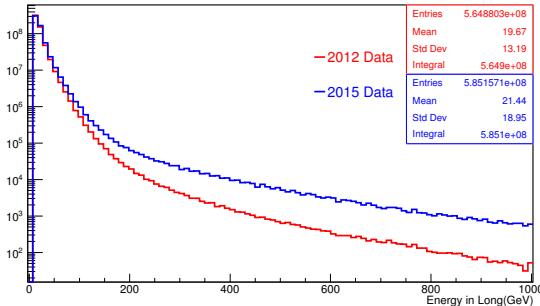


FIGURE 2.20: Energy in long fibers

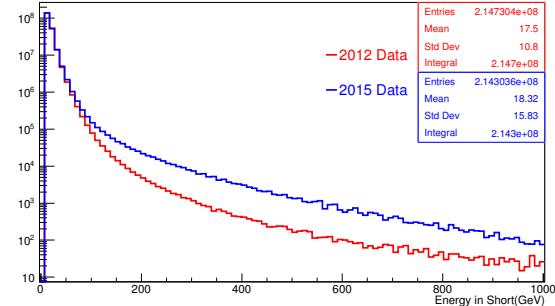


FIGURE 2.21: Energy in short fibers

In 2015 p-p collisions, the  $\sqrt{s}$  was 13 TeV and in 2012 it was only 8 TeV. Because of the increased energy of the collisions, the outgoing particles will have higher energy and hence these particles can penetrate deeper into the calorimeter giving larger energy in short fibers. So the  $R_{S/L}$  is expected to be higher in 13 TeV data as compared to 8 TeV data.  $R_{S/L}$  was determined for different  $i\eta$  towers, with  $i\phi$  inclusive and it was compared for 2012D data, 2015B data and 2015C data (fig. 2.22). For most of the  $i\eta$  towers,  $R_{S/L}$  of 2015 data is higher than that of 2012. With the increase of  $\sqrt{s}$ , average energy of particles is also increasing, consequently the  $R_{S/L}$  is increasing. The reason for differences in 2015 and 2012 data could be because of increased beam energy in p-p collisions and (or) the response of the fibers has changed.  $R_{S/L}$  for  $|i\eta| = 29$  is much higher as compared to other  $|i\eta|$  towers. This is because, this tower is behind  $|i\eta| = 28$  of HE. EM shower energy is already deposited in HE. HF receives only hadronic shower energy and this energy is deposited in both long and short. Hence  $R_{S/L}$  is higher for these towers.  $R_{S/L}$  as a function of  $i\phi$  for each  $i\eta$  was also studied.

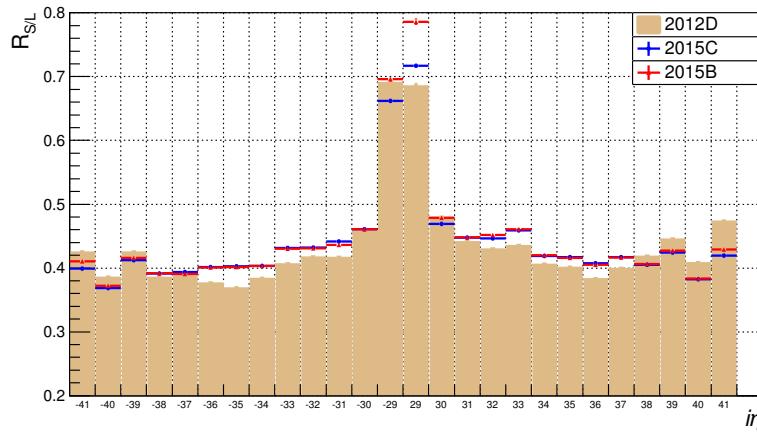
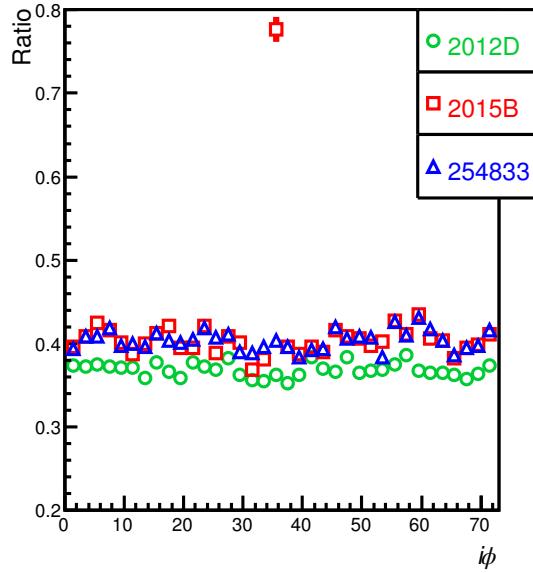
FIGURE 2.22:  $R_{S/L}$  vs  $i\eta$  for 3 different run conditions

Fig. 2.23 shows  $R_{S/L}$  vs  $i\phi$  for  $i\eta = -35$ . From this plot one can see whether a particular channel is problematic or not. In this figure, channel with  $i\phi$  index 35 shows very high ratio in 2015B run. Also in most of the  $i\phi$ s,  $R_{S/L}$  for 2012 data is slightly lower than 2015 data  $R_{S/L}$ .

### 2.3.5 $R_{S/L}$ for 2015 Data and 2016 Data

Data taken in 2015 with 25ns bunch spacing was compared with the 2016 data (25ns bunch spacing) using JetHT dataset and requiring at least one jet with  $p_T > 450$  GeV at trigger level. This comparison is useful to understand the problems or changes in

FIGURE 2.23:  $R_{S/L}$  vs  $i\phi$  plot for  $i\eta = -35$ 

2016 in HF, if any, with respect to the previous data taking.

Pileup conditions for 2015 and 2016B were different (Fig. 2.24). As discussed in sec. 2.3.3, if the pileup is high, then the  $R_{S/L}$  is expected to be smaller since more energy goes into the long fibers.

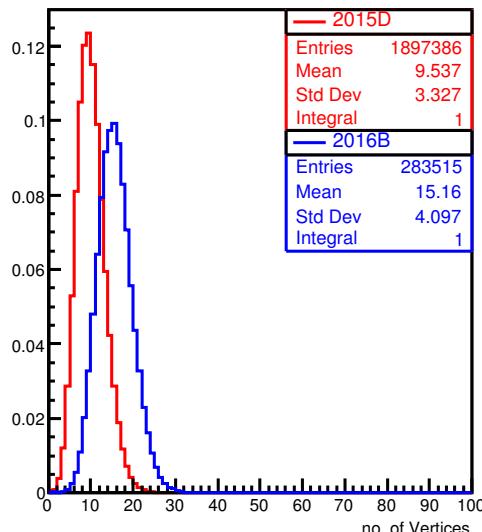


FIGURE 2.24: Number of primary vertices before re-weighting

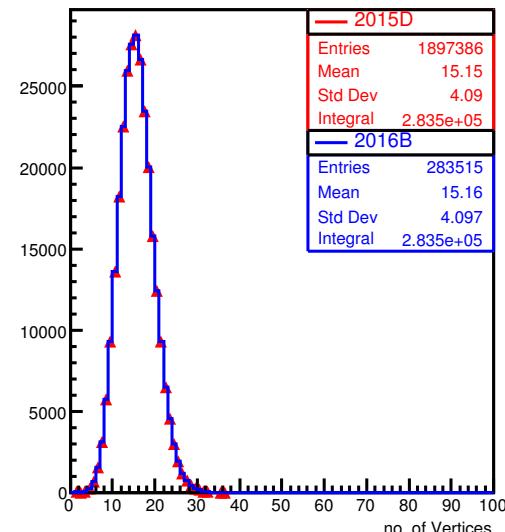


FIGURE 2.25: Number of primary vertices after re-weighting

In order to compare the data taken with these different conditions, events had to be reweighted according to the pileup. In this case, number of primary vertices (PV) distribution in 2015D was re-weighted to match PV distribution of 2016B. This was done by simple division of 2016B histogram by 2015D histogram. The resulting histogram gives the pileup weights for 2015D as a function of different PVs. Fig. 2.25 shows the PV distribution after the reweighing. In Fig. (2.26-2.29) some of the HF parameters such as number of RecHits above 10 GeV, RecHit energy in long and

short fibers for these runs are compared. These runs show very similar behavior with respect to these parameters.

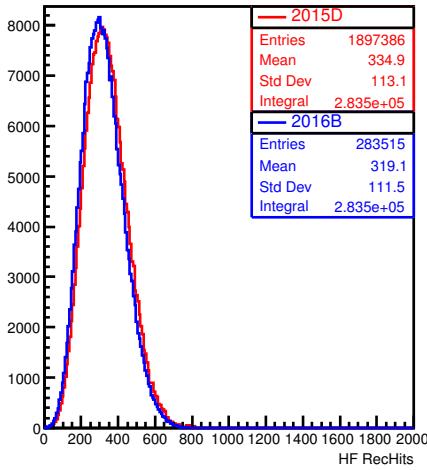


FIGURE 2.26: No. of HF RecHits above 10 GeV

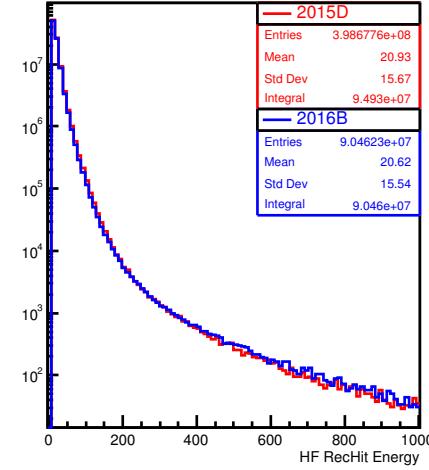


FIGURE 2.27: HF RecHit energy in long and short fibers

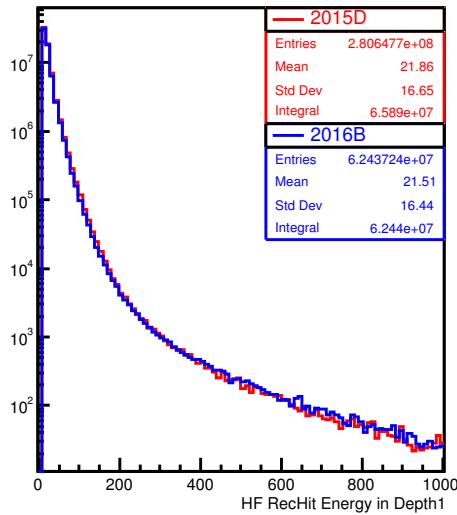


FIGURE 2.28: Energy in long

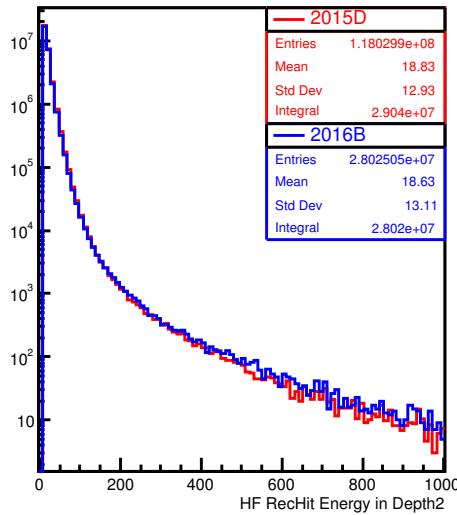
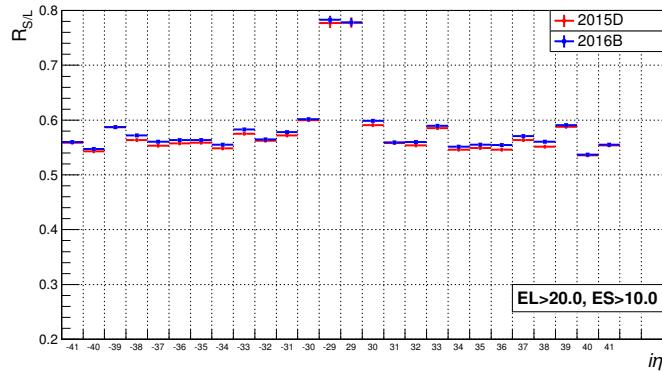


FIGURE 2.29: Energy in short

$R_{S/L}$  as a function of different  $i\eta$  for 2015 data and 2016B data are compared in Fig.2.30. The plot shows that both the datasets have similar ratio and the agreement between these datasets is within 1-2%. It was found that the choice of a different hardonic trigger does not affect  $R_{S/L}$  features seen in this plot.

### 2.3.6 Performance of $R_{S/L}$ in 2016 Data

In this section, different run eras of 2016 data (era B to H) are compared using JetHT dataset and events with at least one jet with  $p_T > 450$  GeV at trigger level. Different run eras of 2016 had different pileup scenarios and 2016B has the smallest pileup among these. This can be clearly seen from the distributions of number of primary

FIGURE 2.30:  $R_{S/L}$  for 2015 and 2016B data

vertices (Fig. 2.31). If the pileup is higher, then the number of rechits are also higher. So larger pileup runs have higher mean number of rechits as shown in Fig. 2.32. All these distributions are normalized to unit area.

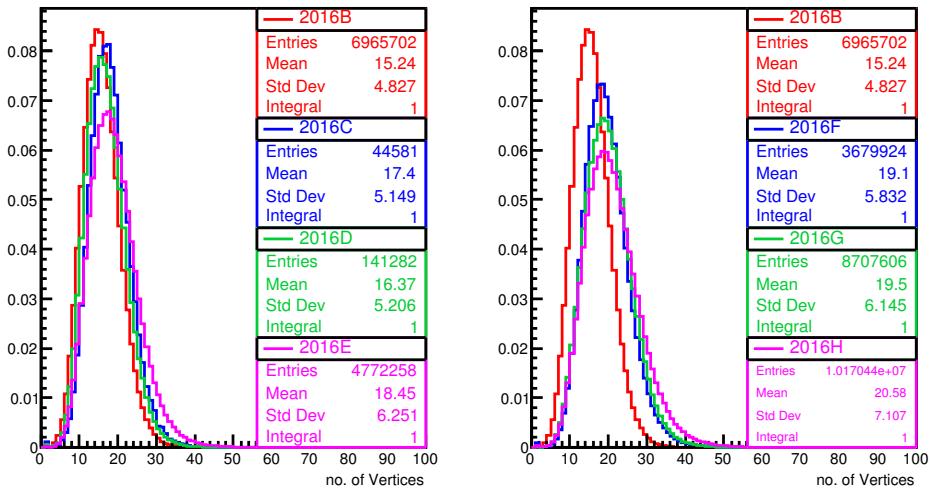


FIGURE 2.31: Number of primary vertices (left) for 2016B,C,D,E, and (right) for 2016B,F,G,H.

RecHit energy distributions for all run eras of 2016 show similar features (Fig. 2.33 - 2.34). Any differences seen in these distributions are because of the different pileup scenarios. In case of run 2016C and D, because of lower statistics (only a part of whole dataset was used), energy distributions show small variations. Run 2016B has lower pileup as compared to others and hence, energy distributions are slightly different from other runs. Runs 2016F to 2016H have very similar pileup and hence their energy distributions show better agreement (Fig. 2.34).

### Stability of $R_{S/L}$ in 2016 Data

In order to check the stability of  $R_{S/L}$  across different  $i\eta$  channels, mean  $R_{S/L}$  as a function of  $i\eta$  is plotted for all the run eras of 2016 (Fig. 2.35). Higher energy thresholds are also used to compare  $R_{S/L}$  of different run eras across various  $i\eta$  channels.

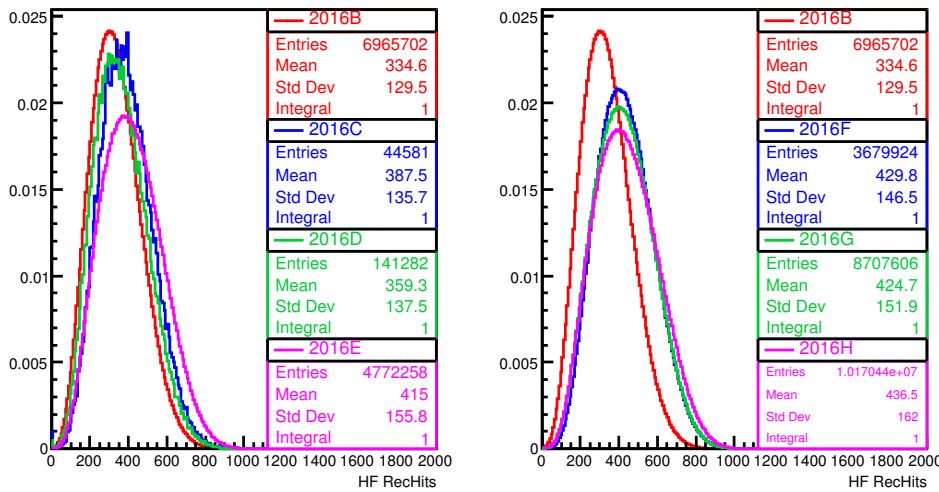


FIGURE 2.32: Number of RecHits (left) for 2016B,C,D,E, and (right) for 2016B,F,G,H.

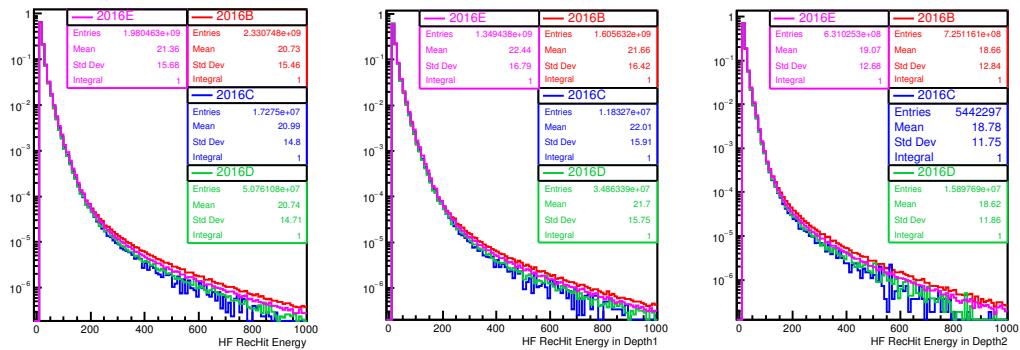


FIGURE 2.33: RecHitEnergy distributions for 2016B,C,D,E

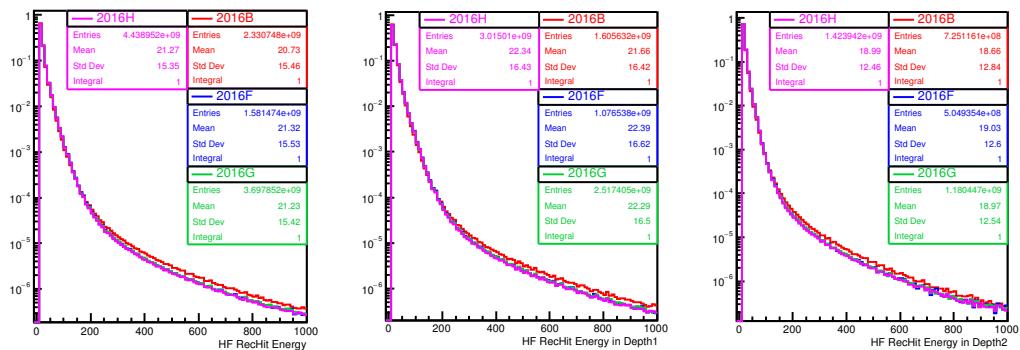
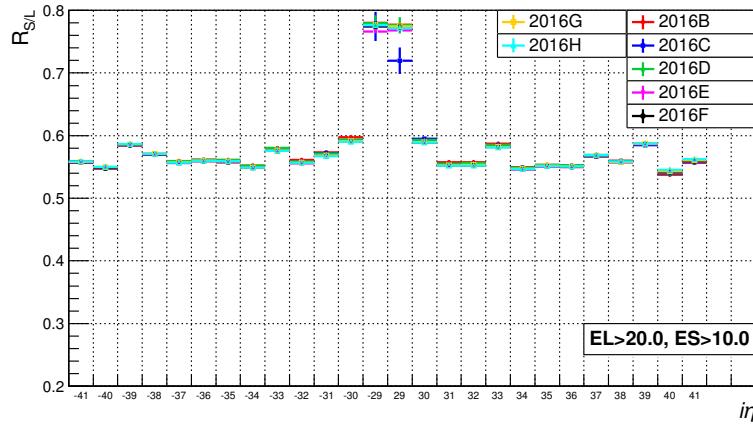
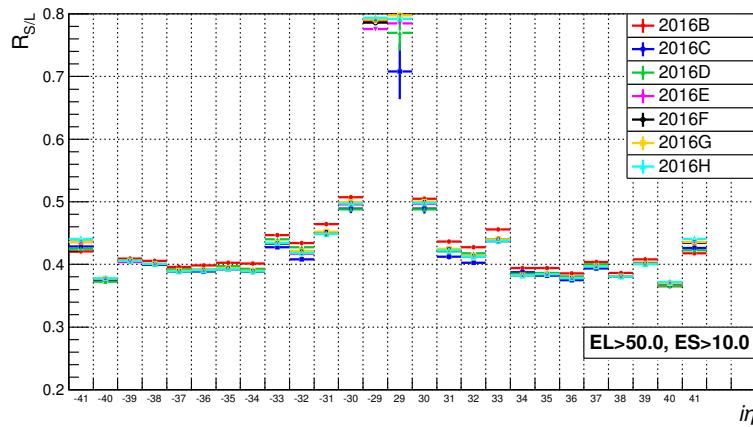


FIGURE 2.34: RecHitEnergy distributions for 2016B,F,G, H

Fig. 2.36 shows  $R_{S/L}$  as a function of  $i\eta$  with  $E_{long} > 50$  GeV and  $E_{short} > 10$  GeV.  $R_{S/L}$  is found to be stable across all these run eras and also for different energy thresholds.

Corresponding to each  $i\eta$  channels, there are 36 divisions in HF. ( $|i\eta| = 40, 41$

FIGURE 2.35:  $R_{S/L}$  vs  $i\eta$  for 2016 with lower  $E_{long}$  thresholdFIGURE 2.36:  $R_{S/L}$  vs  $i\eta$  for 2016 with higher  $E_{long}$  threshold

have 18 divisions).  $R_{S/L}$  as a function of  $i\phi$  for a given  $i\eta$  channel is also studied and a few such plots are shown in Fig.2.37 and Fig.2.38. Across all the channels for all these run eras,  $R_{S/L}$  show consistent features. These plots clearly indicate that  $R_{S/L}$  is stable across different run eras.

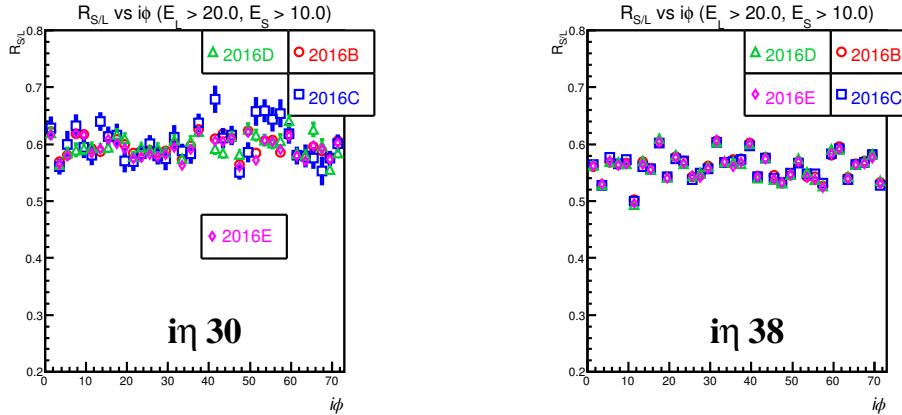
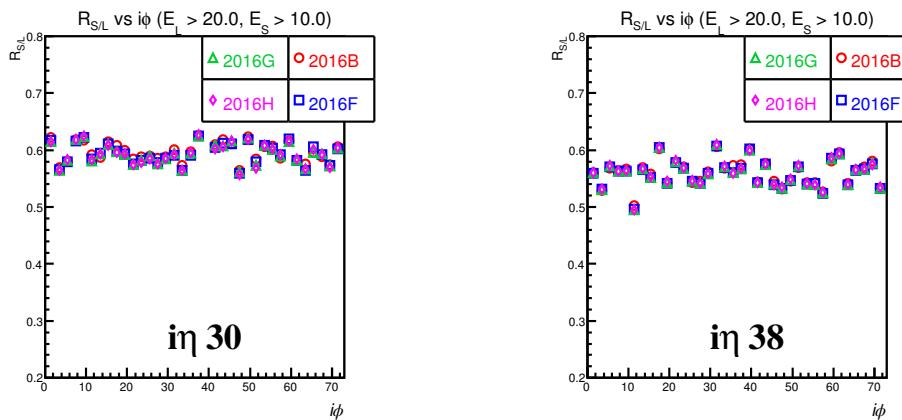
To make sure that there is no bias involved in using the JetHT dataset, comparison of  $R_{S/L}$  for JetHT and a dataset consisting of muons (triggered by a muon with  $p_T > 50$  GeV) is done. These two datasets showed very similar  $R_{S/L}$ .

### 2.3.7 Corrections for the $R_{S/L}$ Based on $\phi$ Symmetry

All the physics processes are symmetric in the azimuthal,  $\phi$  direction ( or the transverse x-y plane). So  $R_{S/L}$  is also expected to be symmetric or flat across all  $\phi$  channels for a given  $\eta$  or  $i\eta$ . Based on this symmetry, corrections are derived for the ratio such that  $R_{S/L}$  is constant across all the  $\phi$  channels for a given  $\eta$ . The corrections are derived as follows:

- Fit the  $R_{S/L}$  vs  $i\phi$  plot of 2016B with line of 0 slope, and intercept  $k$ .
- Correction for a channel is nothing but original  $R_{S/L}$  times the intercept.

$$\text{Correction}(i\eta, i\phi) = k_{i\eta} \times R_{S/L} \quad (2.6)$$

FIGURE 2.37:  $R_{S/L}$  vs  $i\phi$  for 2016B,C,D,EFIGURE 2.38:  $R_{S/L}$  vs  $i\phi$  for 2016B,C,D,E

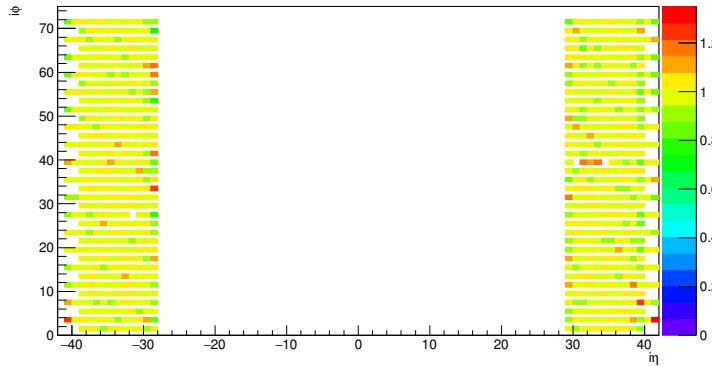
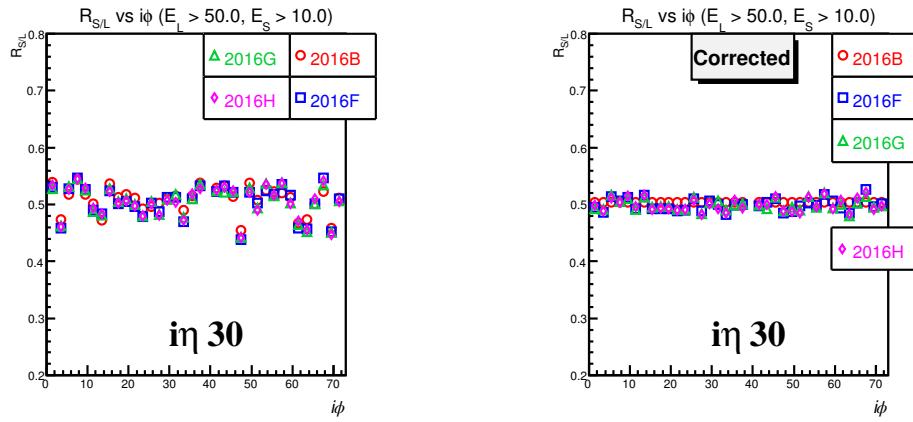
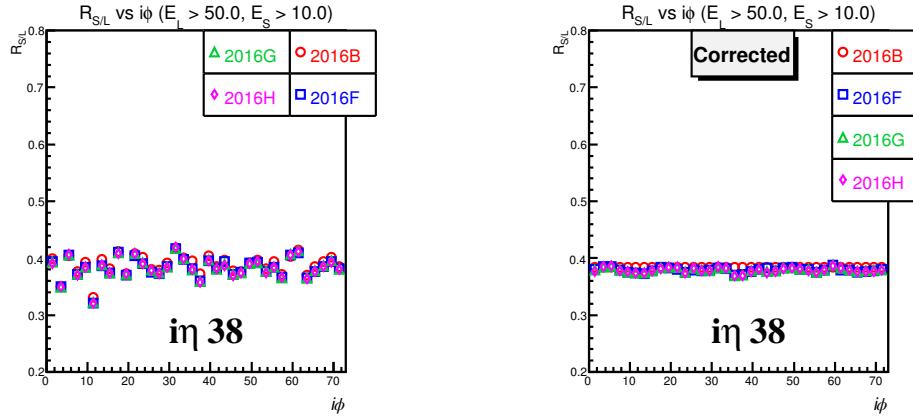
$$\text{Corrected ratio}(i\eta, i\phi) = R_{S/L}^{\text{Corr}} = \text{Correction}(i\eta, i\phi) \times R_{S/L} \quad (2.7)$$

These corrections are obtained for all the channels of HF using 2016B dataset with  $E_{long} > 50 \text{ GeV}$  and  $E_{Short} > 10 \text{ GeV}$ . The corrections factors as a function of  $\eta$  and  $i\phi$  are shown in Fig.2.39. These corrections are applied to 2016B and the remaining 2016 run eras with the same threshold on  $E_{long}$  and  $E_{Short}$ . (The errors on the correction factors are dependent on the error on the mean  $R_{S/L}$  and the fit uncertainty. These errors are less than 1% for almost all the channels, except for  $|i\eta|=29$ )

Fig. 2.41 shows the comparison of ratio plots before correction (left)and after correction (right). Similar features are seen for other run eras i.e, 2016C,D and E.

Corrections derived so far had energy threshold on long as 50 GeV and almost no energy threshold on short energy (10 GeV). Using 50 GeV threshold on long makes the corrections less prone to pileup dependencies. To check the dependency of the corrections on energy thresholds used, these corrections are also used for the ratio plots with different thresholds on long for the same dataset 2016B. Fig. 2.42 and 2.43 show the ratio plots before and after corrections for the energy thresholds  $E_{long} > 30 \text{ GeV}$  and  $E_{long} > 100 \text{ GeV}$ . The corrections are derived from 2016B with  $E_{long} > 50 \text{ GeV}$ . The corrected points are multiplied by a factor of 1.3 so that they can be visualized in the same plots.

It is clear from the Fig.2.41-Fig.2.43 that using the corrections on the ratio, improves the symmetry in  $\phi$ .

FIGURE 2.39: Correction factors as a function of  $i\eta, i\phi$ FIGURE 2.40:  $R_{S/L}$  vs  $i\phi$  for  $i\eta = 30$  for (left) current detector, and (right) after corrections for 2016B,F,G,HFIGURE 2.41:  $R_{S/L}$  vs  $i\phi$  for  $i\eta = 38$  for (left) current detector, and (right) after corrections for 2016B,F,G,H

The long fibers are calibrated using  $Z \rightarrow e^+ + e^-$  events. The plots shown indicate that  $R_{S/L}$  is stable across different runs and it improves the  $\phi$  symmetry. So  $R_{S/L}$  can be used as a tool to inter-calibrate the short fibers.

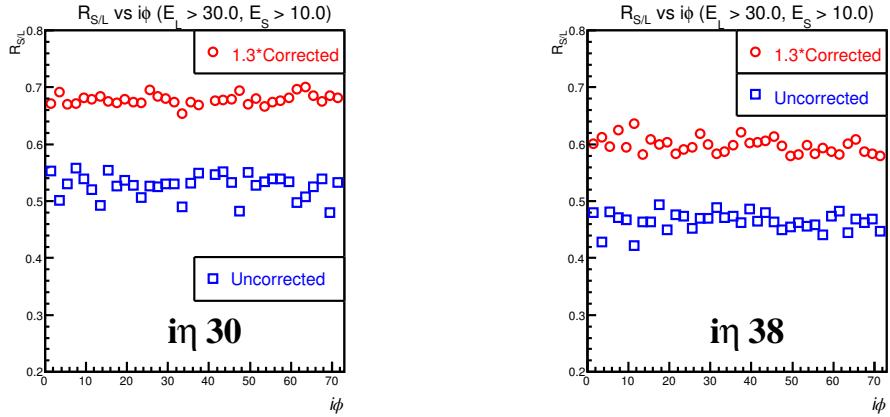


FIGURE 2.42:  $R_{S/L}$  vs  $i\phi$  before and after corrections for 2016B with  $E_{long} > 30, E_{short} > 10$

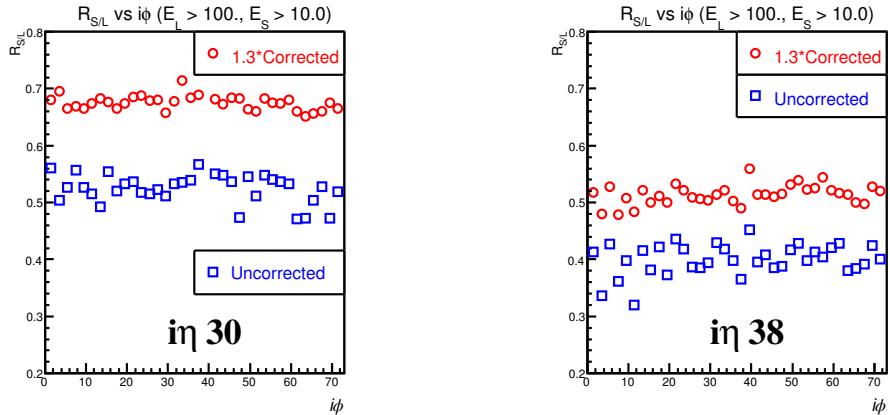


FIGURE 2.43:  $R_{S/L}$  vs  $i\phi$  before and after corrections for 2016B with  $E_{long} > 100, E_{short} > 10$

### 2.3.8 Performance of $R_{S/L}$ in data & MC

In this section comparison of QCD MC sample with data is done. JetHT dataset of 2016E is used for the studies. Trigger used to select events in the data is found to be  $\approx 100\%$  efficient for jet  $p^T > 500$  GeV. There is no trigger requirement for the MC samples. Apart from this selection, the event must contain (in both data and MC) at least one jet with  $p^T > 600$  GeV and within  $|\eta| < 2.4$ . Nominal HF rechit filters are also used.

Data and MC have different number of observed interactions or different pileup (Fig. 2.44). So MC is re-weighted such that the observed number of interactions in data and MC match. Solid blue line corresponds to the observed interactions in MC before PU re-weighting and dotted blue refers to observed interactions in MC after PU re-weighting.

MC is scaled to integrated luminosity of the data ( $= 4.05 \text{ fb}^{-1}$ ). After this scaling of MC, MC had more events (integral) than data (about 20% higher). So MC is scaled down so that the integrals in data and MC match.

Fig. 2.45 shows the number of reconstructed primary vertices and Fig. 2.46 shows the number of HF rechits above 10 GeV (after all MC re-weightings mentioned above).

Fig. 2.47 (left) shows the energy distribution in HF. The energy distributions in

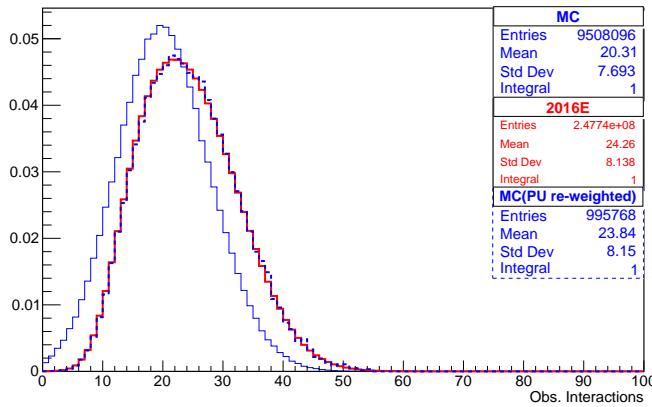


FIGURE 2.44: Observed number of interactions in data and MC (Solid blue line-MC before PU re-weighting, dotted blue-after PU re-weighting).

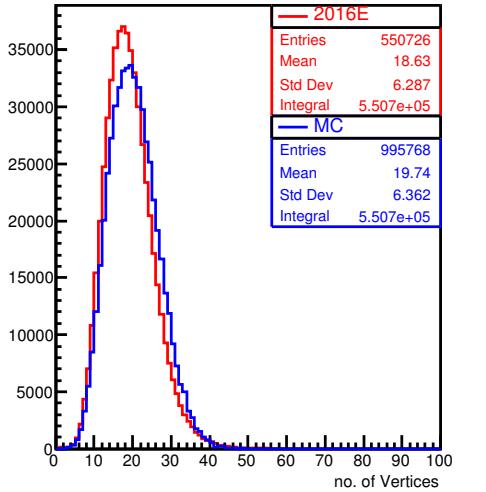


FIGURE 2.45: Number of reconstructed primary vertices in data and MC

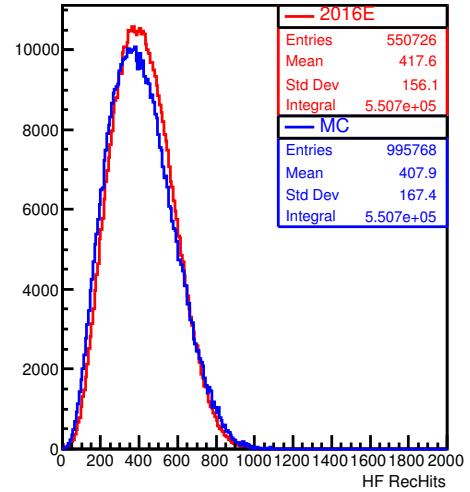


FIGURE 2.46: Number of HF recHits above 10 GeV

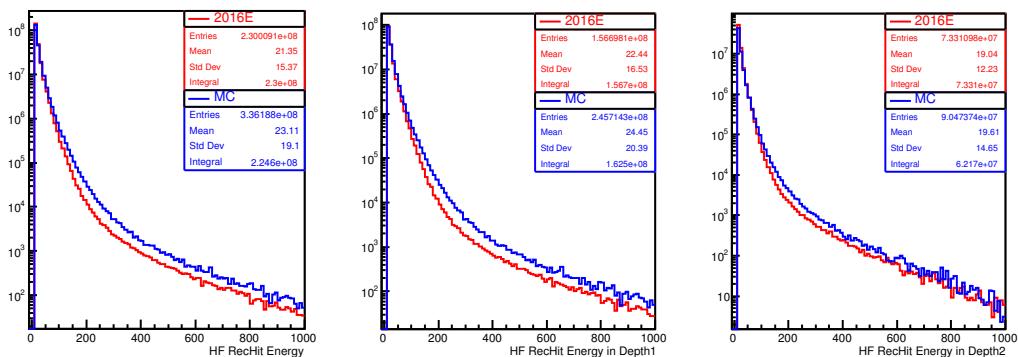


FIGURE 2.47: Distribution of energies in (left) all HF rechits, (middle) long fibers or depth 1, and (right) in short fibers or depth 2.

data and MC show different behavior. Fig. 2.47 (middle) shows energy in long and fig.2.47 (right) shows the energy in short. The energy distributions in long fibers show more discrepancies as compared to the ones in short. In data, energy falls more rapidly than the ones shown by MC.

### $R_{S/L}$ vs $i\eta$

Average  $R_{S/L}$  is determined for data and MC with different energy thresholds on  $E_{long}$  and  $E_{short}$  and they are plotted as a function of  $i\eta$  (Fig. 2.48 - Fig. 2.50). If the energy threshold is low (Fig. 2.48), data and MC show large discrepancies. If higher  $E_{long}$  threshold is used, then the agreement is better (Fig. 2.49). However increasing the  $E_{long}$  threshold to very high values also shows discrepancies (Fig. 2.50). Increasing short energy threshold along with long threshold gives better agreement (Fig. 2.51) between data and MC. The differences are present at long and short energy itself (Fig. 2.47). If data and MC agree for certain thresholds, then it because the differences in long and short cancel to some extent.

As the energy thresholds are varied, shape of the distributions in these plots also change. One reason is that, shape of  $R_{S/L}$  distribution is different for different  $i\eta$  (fig.2.16 and fig.2.17). The other reason is because of the correlation between long and short energies. Any threshold on long or short affects the energy in the other. In fig.2.51, the threshold on long and short are same (50GeV). This energy threshold will not select many of EM showers since  $E_{short}$  is very high. In other words, different thresholds select different EM and hadronic components.

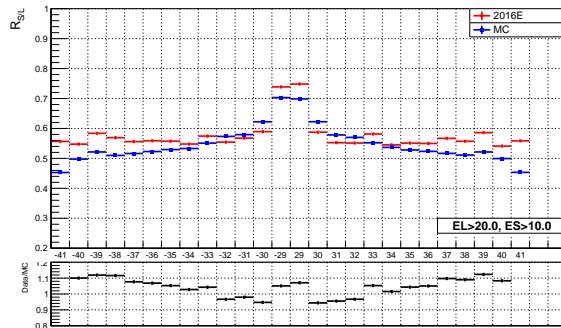


FIGURE 2.48:  $R_{S/L}$  vs  $|i\eta|$  with  $E_{long} > 20, E_{short} > 10$  GeV

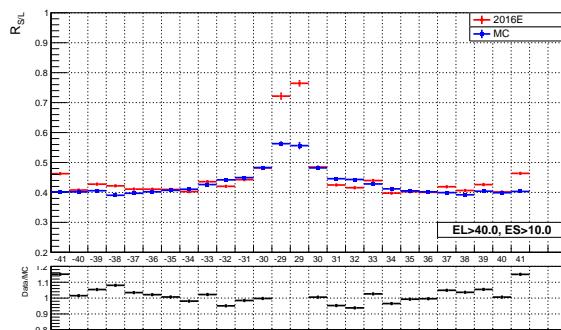
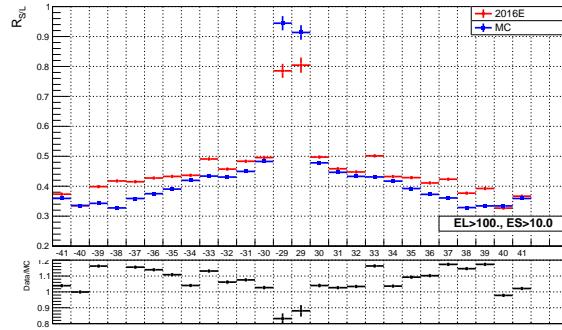
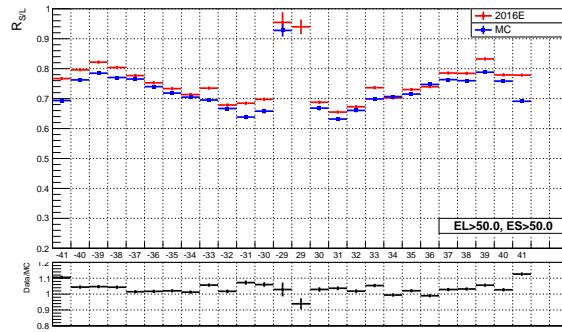
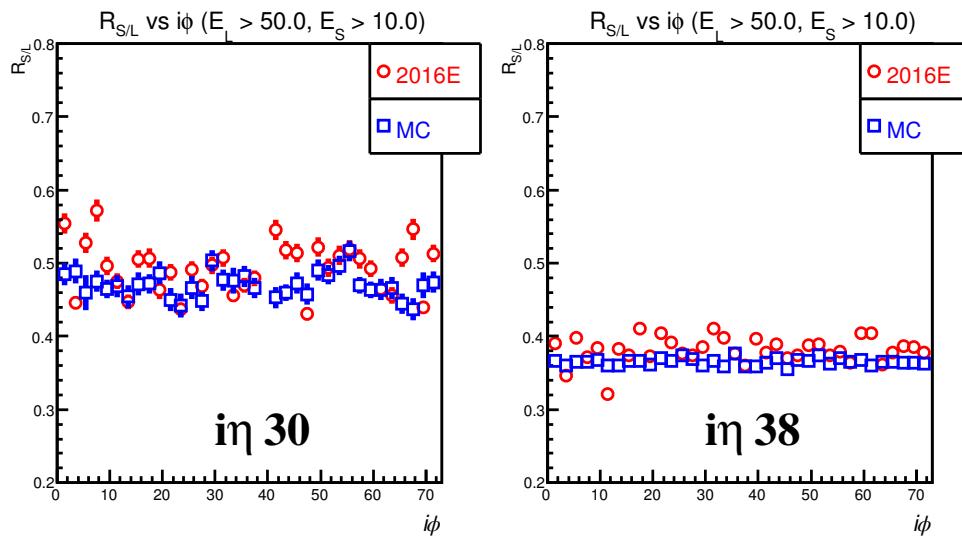


FIGURE 2.49:  $R_{S/L}$  vs  $|i\eta|$  with  $E_{long} > 40, E_{short} > 10$  GeV

FIGURE 2.50:  $R_{S/L}$  vs  $i\eta$  with  $E_{long} > 100, E_{short} > 10$ FIGURE 2.51:  $R_{S/L}$  vs  $i\eta$  with  $E_{long} > 50, E_{short} > 50$ 

### $R_{S/L}$ vs $i\phi$

Since  $R_{S/L}$  vs  $i\eta$  plots for data and MC show differences, it is obvious that  $R_{S/L}$  does not agree for different  $i\phi$ s as well (Fig. 2.52). However, it can be seen that  $R_{S/L}$  is flat (constant across all  $i\phi$ ) for a given  $i\eta$ .

FIGURE 2.52:  $R_{S/L}$  vs  $i\phi$  for data and MC

### 2.3.9 Summary of HF studies using $R_{S/L}$

Study of performance of hadron forward (HF) calorimeter was done using ratio of energy deposited in short to long fibers ( $R_{S/L}$ ) of HF. Data collected by CMS in 2012 with  $\sqrt{s} = 13\text{TeV}$  was compared with 2015 data using  $R_{S/L}$ . For most of the  $i\eta$  towers, 2012 showed lower  $R_{S/L}$  than 2015, which might be because of increased  $\sqrt{s}$  or radiation damage. Affect of pileup on  $R_{S/L}$  was studied and with increased pileup,  $R_{S/L}$  was found to decrease. Comparison of 25ns bunch spacing 2015 data and 2016 data was done and  $R_{S/L}$  showed similar features. Different run eras of 2016 also showed similar  $R_{S/L}$ , which indicates the stability of  $R_{S/L}$  across these run eras. For a given  $i\eta$ , different  $i\phi$ s showed small variations in  $R_{S/L}$ . Corrections were obtained using 2016B to make the  $R_{S/L}$  constant (flat) across these  $\phi$  channels and these corrections improved flatness of  $R_{S/L}$ . Assuming that long fibers are well calibrated within a given  $i\eta$ , these corrections can be used intercalibrate short fibers. Comparison of data and MC was also done and these shows differences in terms of energy distribution in long and short fibers.  $R_{S/L}$  also shows differences for many of the channels. The agreement in terms of  $R_{S/L}$  is better for higher energy thresholds.

## 2.4 Trigger system

In CMS, two levels of triggering system is used to reduce the rate of data flow from various detectors. The first level trigger (L1) is hardware based and the second level is software based (high level or HLT). The trigger system is inevitable since there is several orders of magnitude difference between inelastic pp cross section and other SM and SUSY processes. Figure 2.53 shows the cross section (on left y-axis) for various SM processes, and SUSY processes which are theoretically calculated [22]. Inelastic pp is of no interest to LHC physics and hence these events in pp collision need to be filtered out. There are also limiting factors in terms of the rate of writing into a disk and storage capacity.

The pp collision frequency is 40 MHz. The L1 trigger (L1T) system uses custom hardware made using field programmable gate arrays (FPGAs) and GaAs application-specific integrated circuits (ASICs). The output rate is adjusted depending on the LHC luminosity and it is maintained below 100 kHz which is the upper limit imposed by readout electronics. It selects events containing various candidate objects or a combination of the objects - ionization deposits consistent with muon, or ECAL energy clusters for electrons/photons, or energy deposits in calorimeters expected from  $\tau$ , jet, or missing transverse momentum obtained by vectorially summing the objects and using transverse component. A schematic diagram of L1T is shown in Fig.2.54

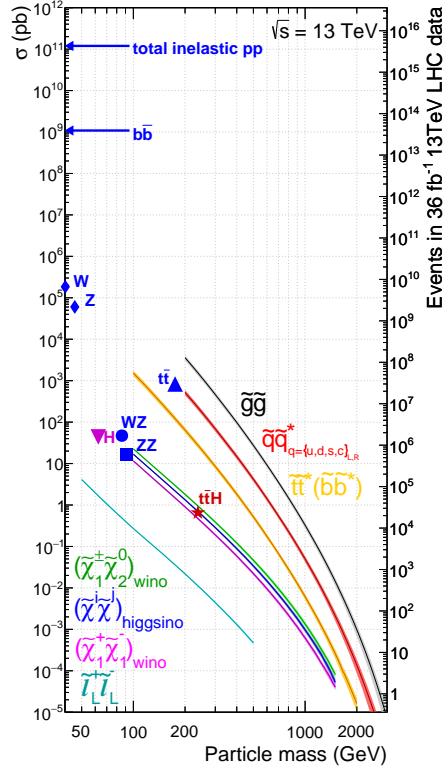


FIGURE 2.53: Comparison of cross section for various SM processes and SUSY processes.

[73]. The data from calorimeters is processed regionally and globally forming re-

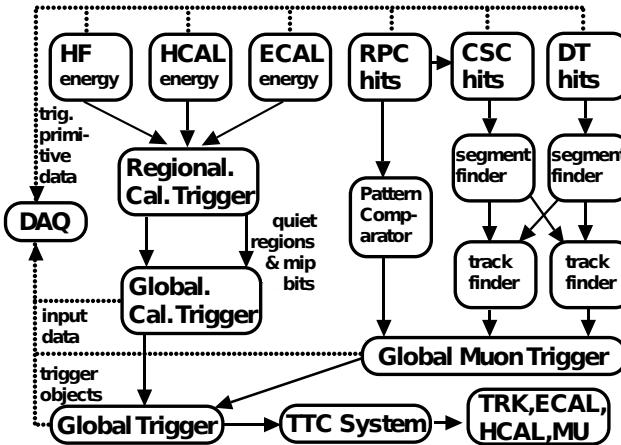


FIGURE 2.54: A schematic diagram of L1T [73].

gional and global calorimeter trigger (RCT and GCT). Global muon trigger (GMT) is formed by taking hits from various muon systems by making use of pattern comparator or segment and track finder. The global trigger (GT) makes the final decision about the event by combining GCT and GMT. This decision is sent to CMS subdetectors via trigger timing and control (TTC). The DAQ (data acquisition system) reads data from subsystems for offline processing.

HLT performs reconstruction and identification of objects similar to offline, but faster and less precise. It makes use of several computers which run various algorithms in a predefined order and increasing complexity. Selections based on the information from calorimeters and muon chambers is carried out first and then the CPU intensive tracking reconstruction. The output data flow from HLT is about 400 Hz and it is used for offline storage.

## 2.5 Object reconstruction

The identification and reconstruction of various physics objects is done using the information from subdetectors using particle-flow (PF) algorithm [91]. Fig.2.55 shows the interaction of different particles in a transverse slice of CMS detector. The charged particles leave signatures in various layers of the inner tracker because of 3.8 T magnetic field provided by the solenoid. Muon, being charged and minimum ionizing particle, bends in magnetic field and gives hits in the tracker and outer muon chambers. Electrons give hits in tracker and deposit almost all their energy in the ECAL. Photons can be identified if there is energy deposit in ECAL, but no track. Hadrons deposit most of the energy in HCAL and distinction of neutral and charged hadron is done by the presence of track.

### 2.5.1 Tracking

Track reconstruction takes place using a combinatorial track finder (CTF) algorithm based on Kalman filtering (KF). Full tracking is an iterative process starting with the easiest tracks and then proceeding to find more difficult tracks. After each iteration, hits compatible with the tracks are removed and this reduces the number of hits to use for the next iteration. There are 4 steps in an iteration.

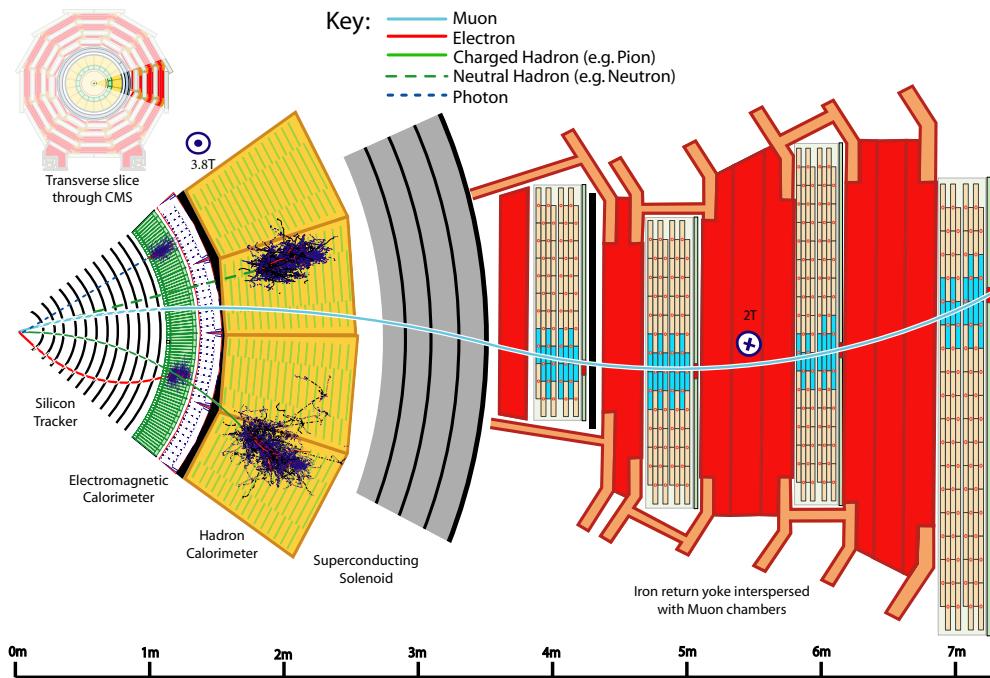


FIGURE 2.55: A sketch showing interaction of particles in a slice of CMS detector

- **Seed generation:** Finding track candidates using 2-3 hits compatible with charged particle trajectory.
- **Track finding:** Extrapolate the trajectory of the seeds to find compatible hits in the other layers using KF technique.
- **Track fitting:** Perform fitting to smooth the trajectory and find track parameters - origin,  $p_T$  and direction.
- **Track selection:** Apply selection criteria on the tracks based on number of layers with a hit, number of layers with missing hits, number of 3D hits, goodness of fit ( $\chi^2$ ) and impact parameters of the track.

For the next iteration, hits used in this iteration are removed and some of the selection criteria are changed in order to reconstruct more difficult tracks which originate because of missing hits in the pixel detector, particle interactions and decays, high  $p_T$  jets with collimated particles etc.

The basic idea in tracking of electrons is extrapolation of tracks to ECAL energy clusters. The iterative tracking mentioned above is efficient for non-radiating electrons. Because of non-negligible tracker material (Fig.2.7), electrons radiate photons. Depending on the energy of the radiated photon, values of  $\chi^2$  and number of hits may vary. In these cases, Gaussian-sum filter (GSF) technique [3] is used instead of KF. GSF fitting allows for sudden change in momentum of the tracks as the electron passes through various layers of tracker. Finally, a requirement is applied on the score of a boosted decision tree (BDT) classifier which uses various parameters of the KF and GSF tracks.

### 2.5.2 Interaction vertices

When the pp bunches collide at the LHC, multiple pp interactions can take place. The goal of vertex reconstruction is to locate the position of such interactions along with the uncertainties associated with it. This process involves 3 steps [29]:

1. selection of tracks: tracks passing quality criteria based on transverse impact parameter relative to the beam spot, number of hits and normalized  $\chi^2$ .
2. clustering of tracks which seem to originate from same vertex: the z-coordinates of the selected tracks measured from center of the beam spot to the point of closest approach are used for clustering. The algorithm should be able to resolve very close vertices, but not split a genuine single vertex into many. Deterministic annealing (DA) algorithm [88] [29] is used for this purpose.
3. fitting the position of each vertex using its associated tracks with the help of adaptive vertex filter [52].

After all interaction vertices are reconstructed, the primary pp interaction vertex is selected as the vertex with the largest  $p_T^2$  sum of all physics objects. The physics objects used in this calculation are produced by a jet-finding algorithm [26, 25] applied to all charged-particle tracks associated to the vertex, plus the corresponding  $p_T^{miss}$  computed from those jets.

### 2.5.3 Muons

Since the calorimeters absorb energy of EM objects and hadrons, hits in muon spectrometer can be used to reconstruct muons. These muons which are reconstructed using muon systems only are called *standalone muons*.

If the standalone muon track is compatible with the track from the inner tracker, a new *global muon* track is fitted to the hits from both the systems.

If at least one muon segment matches with the extrapolated track from inner tracker, corresponding muon is called *tracker muon*.

The performance of global muon is better than standalone or tracker muon since it uses information from inner tracker and muon system.

The momentum of muon is determined using *Tune-P* algorithm [33]. The reconstructed muons are fed into PF algorithm which combines information from other subdetectors of CMS and applies various quality criteria. The identification of muons uses criteria on track fit  $\chi^2$ , number of hits in tracker or muon chambers or both, compatibility of tracker tracks and hits/tracks in muon chambers and also the compatibility with primary vertex. Detailed muon reconstruction procedure and its performance in 2015 and 2016 data is described in ref. [92].

### 2.5.4 Electrons and photons

Electrons and photons interact in similar way in ECAL. The material in the tracker lead to conversion of photons in to  $e^+e^-$  pairs and electrons can emit bremsstrahlung photons, photons again convert etc.

When an electron radiates photon, energy is spread in  $\phi$  direction because of the magnetic field and negligible spread in  $\eta$  direction. The amount of energy radiated depends on  $\eta$  and it varies from 33% to 86% [71].

An electron candidate is seeded from a GSF track, provided the corresponding ECAL cluster is not linked to 3 or more additional tracks. A photon candidate is

seeded from a ECAL energy cluster with transverse energy,  $E_T > 10$  GeV with no link to GSF track.

Two algorithms, *hybdrid* for the barrel (EB) and *multi 5x5* for endcap (EE) [71] are used to collect the energy of the radiated photons in case of electron reconstruction or converted photons in case of photon reconstruction. Clustering is performed on intercalibrated, reconstructed signal amplitudes after taking care of several detector effects. Clustering starts with a seed crystal which is the highest  $E_T$  crystal in its neighborhood. In EB, clusters are centered on the seed crystal and have a fixed width of 5 crystals in  $\eta$ . In the  $\phi$  direction, adjacent strips of 5 crystals are added. Further  $\eta$  aligned crystals may be seeded and added if they lie within  $\phi$  window of  $\pm 17$  crystals. At each step of clustering, predefined thresholds on energy are applied. Clustering in EE is carried out using fixed 5x5 crystal matrices. After identifying the seed crystal, more matrices are added if their centroid lies in  $\eta$  of 0.07 and  $\phi$  distance of about 0.3 radians. These matrices are partially allowed to overlap. If the photons do not convert, then superclusters resulting from these algorithms are simply 5x5 matrices. More details about electron and photon reconstruction and performance can be found in ref. [71, 72].

### Photon Identification

The collection of photons derived from PF algorithm have very loose criteria on their properties. For the purpose of analyses, more tighter requirements are placed on photons. The main idea is to get highest efficiency for prompt isolated photons and reject non-prompt photons which are coming from  $\pi^0$  decays or part of hadronization process. Several properties of PF photons are studied to distinguish prompt and non-prompt photons. Prompt photons generally have smaller energy spread in ECAL and very less energy deposit in HCAL. Non-prompt photons are associated with charged or neutral hadrons nearby and they deposit energy in HCAL and sometimes a part of their energy in ECAL leading to larger ECAL clusters. Table 2.4 shows a list of variables and their values used for selecting photons in this SUSY search. Description of variables used for photon identification:

TABLE 2.4: A list of variables used for identification (ID) of photons in barrel and endcap. Prompt photon properties should not exceed any of the values listed in this table. Bottom 2 rows show the efficiency for signal (prompt photons) and background (non-prompt photons).

Property	Max value in EB	Max value in EE
H/E	0.05	0.05
$\sigma_{\eta\eta\eta}$	0.0102	0.0274
PF $h^\pm$ iso	3.32	1.97
PF $h^0$ iso	$1.92 + 0.014p_T + 0.000019p_T^2$	$11.86 + 0.0139p_T + 0.000025p_T^2$
PF $\gamma$ iso	$0.81 + 0.0053p_T$	$0.83 + 0.0034p_T$
Efficiency	EB (%)	EE (%)
Signal	90	90
Background	16	18

- H/E: Ration of energy in HCAL tower behind ECAL supercluster to energy of supercluster.

- $\sigma_{i\eta i\eta}$ : This is a shower shape variable which is the second moment of the energy distribution along the  $\eta$  coordinate.

$$\sigma_{i\eta i\eta}^2 = \frac{\sum_{i=1}^{5\times 5} w_i (i\eta_i - i\eta_{seed})^2}{\sum_{i=1}^{5\times 5} w_i}, \quad w_i = \max(0, 4.7 + \ln \frac{E_i}{E_{5\times 5}}) \quad (2.8)$$

The summation runs over 5x5 crystal matrix around most energetic crystal in the supercluster.

- PF  $h^\pm$  iso: PF charged hadron isolation variable which is defined as sum of  $p_T$  of all PF charged hadrons which are associated to the primary vertex, not in the footprint of the candidate photon, and within a cone of  $\Delta R = 0.3$ .
- PF  $h^0$  iso: Pileup corrected PF neutral hadron isolation defined as sum of  $E_T$  of all PF neutral hadrons which are not in the footprint of the candidate photon and are within a cone of  $\Delta R = 0.3$ .
- PF  $\gamma$  iso: Pileup corrected PF photon or  $\gamma$  isolation is defined as sum of  $E_T$  of all PF photons which are not in the foot-print of the candidate photon and are within a cone of  $\Delta R = 0.3$ .

The PF  $h^0$  and  $\gamma$  isolations are corrected for pileup contributions using  $\text{Iso}_{\text{corr}} = \text{Iso}_{\text{original}} - \rho \text{EA}$ , where  $\rho$  is the energy density of the event calculated using FastJet package [25] which is a measure of pileup activity in the event per unit area and EA is the effective area.

To reject mis-identification of electrons into photons, an additional criteria called *pixel seed veto* is also used. If the ECAL energy cluster is associated with hit(s) in the pixel detector then the corresponding electron or photon is said to have pixel seed.

This identification criteria shown in table 2.4 with pixel seed veto is called *loose ID + pixel seed veto*. The efficiency for this selection is shown in figure 2.56. This efficiency is determined using simulation with data to simulation corrections applied.

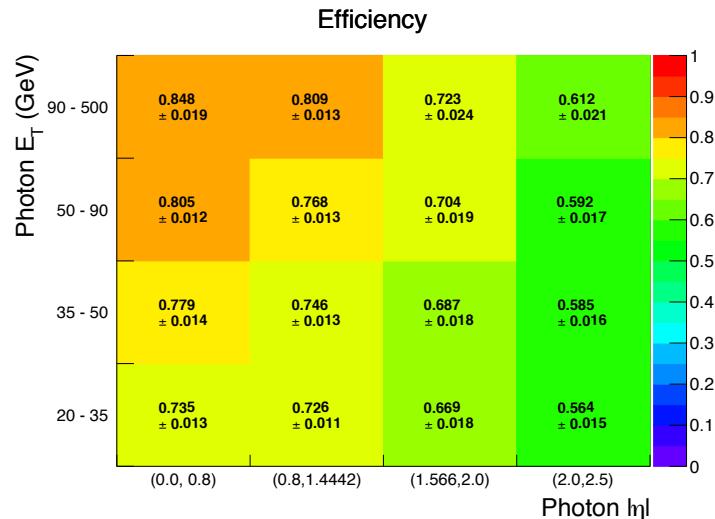


FIGURE 2.56: Efficiency for loose ID + pixel seed veto photons determined from simulation. The uncertainty includes statistical and systematic.

## Electron Identification

Various quality criteria are applied to PF electrons in a similar way as photons. Since electron also has a track, electron ID makes use of track based variables apart from photon ID variables. In this analysis, we use *mini-isolation*,  $I_{mini} < 0.1$  instead of isolation calculated using fixed cone size of  $\Delta R = 0.3$ .

$$I_{mini} = \frac{\sum_R p_T(h^\pm) + \max(0, \sum_R p_T(h^0) + p_T(\gamma) - \rho E A (\frac{R}{0.3})^2)}{p_T(e)} \quad (2.9)$$

The variables used in this equation are same as the ones used for photon and its calculation is similar to the way as fixed cone isolation is calculated, except that the isolation cone size is dependent on  $p_T$  of the electron [86]. The size of the cone varies as:

$$R = \begin{cases} 0.2 & \text{if } p_T < 50 \text{ GeV} \\ 10 \text{ GeV}/p_T & \text{if } 50 < p_T < 200 \text{ GeV} \\ 0.05 & \text{if } p_T > 200 \text{ GeV} \end{cases}$$

In case of boosted  $t\bar{t}$  events involving  $W \rightarrow e\nu$ , b-jet from top decay may be very close to electron. The fixed cone isolation becomes inefficient in identifying these kind of electrons. Since the cone size is smaller for large  $p_T$  electrons in case of mini-isolation, the efficiency for identification of such electrons is high with mini-isolation.

The efficiency of mini-isolation (or any other identification criteria) can be measured using *tag & probe* (TnP) technique both in data and MC. If there are differences in data and MC efficiencies, corrections are applied to MC and these corrections are called as scale factors (SFs). Tag object is a well identified electron with very strict criteria on its properties. This object needs to have  $p_T > 30$  GeV and  $|\eta| < 2.17$  and matched with any of the electron HLT trigger objects ( $\Delta R(\text{Tag}, \text{HLT object}) < 0.3$ ). Probe object is the electron whose efficiency needs to be measured and it has charge opposite to that of tag. To calculate the efficiency of any isolation, we also impose the condition on probe that it passes an ID criteria which does not use isolation variables. In this case we have applied a multivariate technique based electron identification criteria on the probes. The decay of Z boson gives a pair of electrons and one the electrons serves as tag and the other as probe. By using the 4-momenta of electrons, invariant mass of Z boson can be reconstructed. Figure 2.57 shows the invariant mass distribution of tag and probe pairs in data collected in 2016 for events with probe  $p_T \in 35 - 50$  GeV and  $\eta < 0.8$ . The left plot corresponds to events in which probe object passes min-isolation criteria and right plot corresponds to events in which probe fails mini-isolation criteria. This distribution is fitted with a signal function - MC template with Gaussian smearing shown in red, and a background function - CMS shape function shown in blue. The CMS shape function is a product of falling exponential and error function [81]. The efficiency is given by the ratio of number events with passing probe to the total number of events with passing/failing probe. This can be determined by taking the area under the signal function from both of the fits. The efficiency for the above figure is found to be 0.97 with negligible statistical uncertainty. The efficiency in MC can be obtained by using the template or using generator level information.

The whole procedure can be repeated by taking other  $p_T - \eta$  ranges and calculate the efficiency. To evaluate the systematic uncertainty, refitting is done using different signal function (Crystal-Ball with Gaussian smear) and background function (exponential). Systematic uncertainty because of the choice of tag is evaluated by making

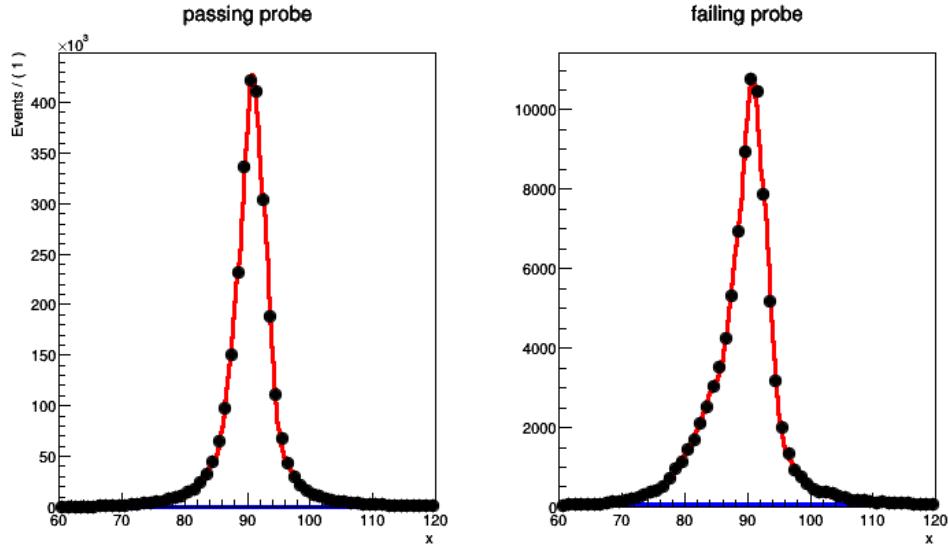


FIGURE 2.57: The left (right) plot shows the invariant mass distribution of tag and probe passing (failing) mini-isolation criteria.

a different tag selection. Figure 2.58 shows the efficiency (top panel) and SFs (bottom panel) for various  $\eta$  regions as a function of  $p_T$ .

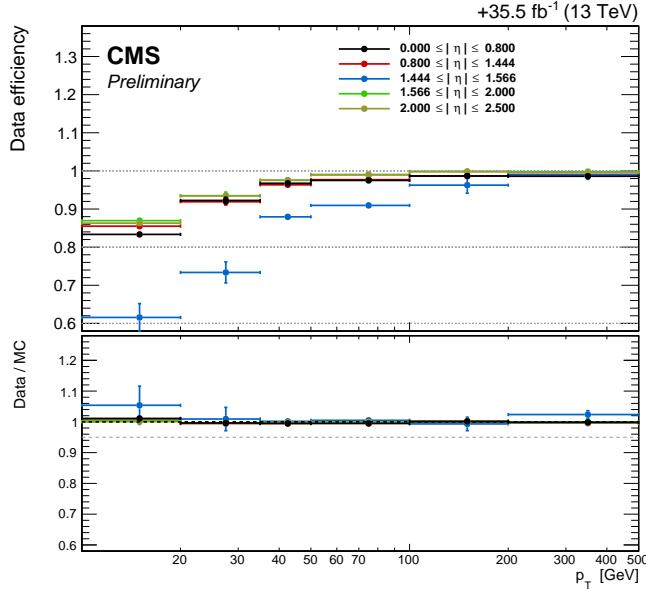


FIGURE 2.58: The efficiency of mini-isolation in data (top panel) as a function of electron  $p_T$  for various  $\eta$  regions and the bottom panel shows the SFs. The error bars include systematic and statistical uncertainties.

These SFs account for any discrepancies in data and MC and they are applied on MC. In a similar way the SFs are derived to account for discrepancies at track reconstruction level, identification, and isolation (as mentioned here for mini-isolation). Wherever needed, these type of corrections are applied to other physics objects as well.

### 2.5.5 Jets

Hadronization of quarks and gluons results in formation of many mesons and baryons which is referred to as a jet, deposits its energy mostly in HCAL. On an average 65% of the total energy of the jet is carried by charged hadrons, 25% by photons and 10% neutral hadrons. Figure 2.59 shows the jet energy fraction as a function of  $p_T$  in data and simulation. The lower panel shows the difference in data and simulation.

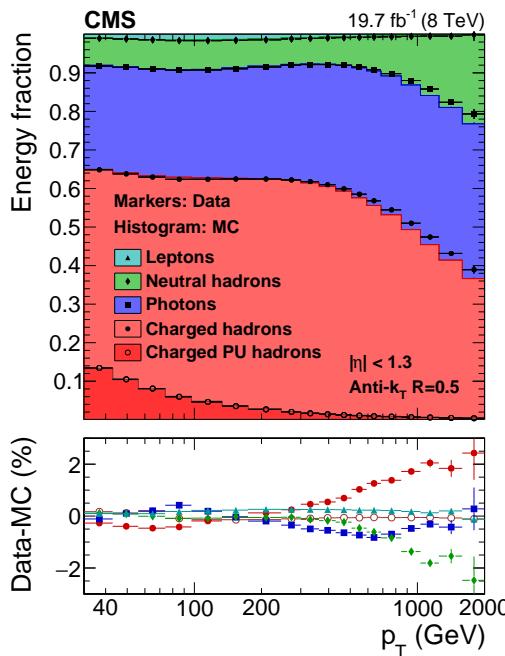


FIGURE 2.59: Energy composition in jets as a function of jet  $p_T$  in data and MC simulation [91].

Jets are reconstructed using PF candidates using anti- $k_T$  algorithm [26, 25] with a size parameter of 0.4. These jets are called as AK4PF jets. They are also subject to pileup mitigation using charged hadron subtraction (CHS). In CHS, charged hadrons associated to vertices apart from primary vertex are removed from PF candidates. To remove remaining energy due to charged hadrons and neutral hadrons, jet area based pileup corrections are applied on an event by event basis. The distance parameters used in jet clustering are:

$$d_{ij} = \min(p_{Ti}^{2n}, p_{Tj}^{2n}) (\Delta_{ij}^2 / R^2) \quad (2.10)$$

$$d_{iB} = p_{Ti}^{2n} \quad (2.11)$$

$$\text{where } \Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 \quad (2.12)$$

The values of  $R$  and  $n$  for AK4 jets are 0.4 and -1 respectively.  $d_{ij}$  is the distance between entities (particles or psudojets)  $i$  and  $j$ ,  $d_{iB}$  is the distance between  $i$  and the beam (B),  $y_i$  is rapidity and  $\phi_i$  is the azimuthal angle of  $i$ th particle. The jet clustering proceeds iteratively by combining *nearest* particles. In the case of anti- $k_T$ , the  $d_{ij}$  between similarly separated soft particles is higher than the  $d_{ij}$  between a hard particle and soft particle. The soft particles in this algorithm do not modify the shape of the jet, but the hard ones do.

The measured energy or the  $p_T$  of the jet needs to be corrected so that its measured 4-momentum is close to parton level 4-momentum. There are several reasons

for imperfect measurement of jet energy such as pileup energy contribution, non-uniformity of detector response to different jet energy etc. The jet energy corrections (JECs) aim to correct the jets and determine the momentum of the initial parton. The

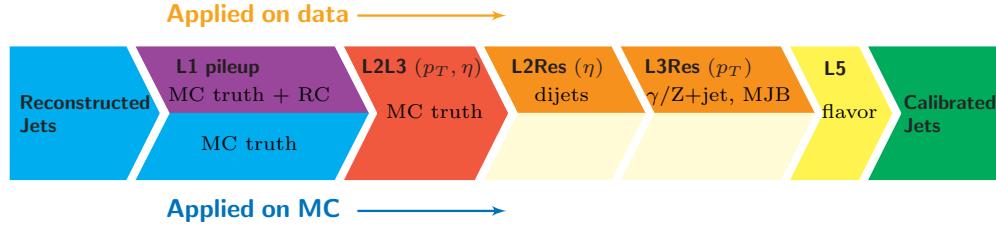


FIGURE 2.60: Pictorial representation of various steps in application of JECs on data and MC simulation [70].

corrections are derived at different levels (L1, L2 and L3) and the output of each level becomes the input for the other, in other words a factorized approach is followed for applying the corrections. Figure 2.60 shows a pictorial representation of different steps in applying JECs both on data and MC simulation. The corrected 4-momenta of the jet can be written as [30]:

$$p_\mu^{corr} = \mathcal{C} p_\mu^{raw} \quad (2.13)$$

$$\text{where } \mathcal{C} = C_{\text{offset}}(p_T^{raw}) \cdot C_{MC}(p'_T, \eta) \cdot C_{rel}(\eta) \cdot C_{abs}(p''_T) \quad (2.14)$$

The offset correction,  $C_{\text{offset}}$  removes extra energy due to pileup and noise, MC calibration factor,  $C_{MC}$  removes bulk of non-uniformity in  $\eta$  and non-linearity in  $p_T$ ,  $C_{rel}$  and  $C_{abs}$  are relative and absolute energy scale corrections which account for small differences in data and MC. In the expression for  $\mathcal{C}$ ,  $p'_T$  is the  $p_T$  of the jet after offset correction,  $p''_T$  is the  $p_T$  of the jet after all previous corrections. Figure 2.61 shows the average jet  $p_T$  response with and without corrections at a few stages of JEC application. The  $p_T$  response is defined as the average value of ratio of measured jet  $p_T$  to particle-level jet  $p_T$ ,  $p_{T,\text{ptcl}}$ .

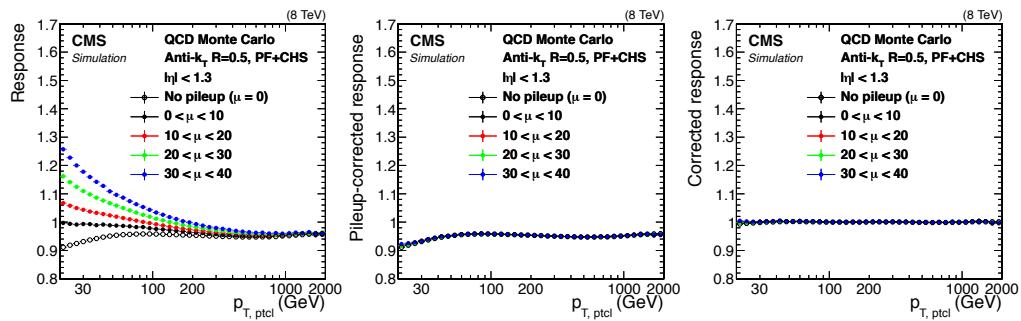


FIGURE 2.61: Average jet  $p_T$  response in QCD MC at different stages of JECs. The left plot is without any corrections, middle plot is for after pileup offset corrections and the right plot is after applying all the JECs.  $\mu$  refers to the average number of pileup interactions per bunch crossing.

The resolution of AK4PF jet as a function of ref jet is shown in 2.62. The uncertainty in the jet energy correction is about 2% for barrel region jets with  $p_T$  of 30 GeV [65].

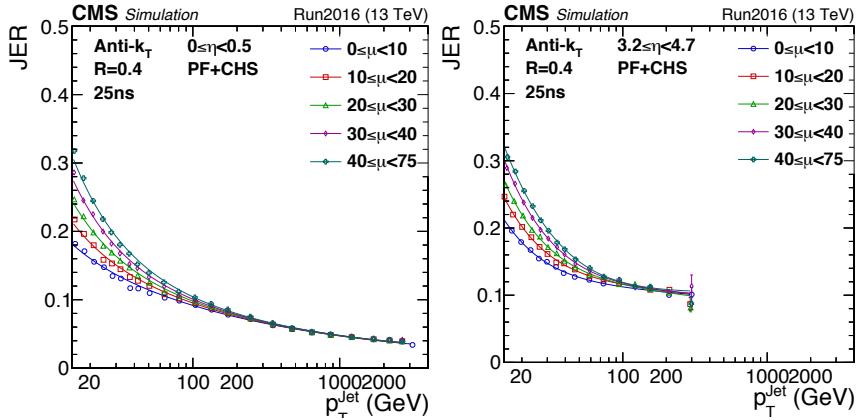


FIGURE 2.62: Resolution of PF jets as a function of  $p_T$  in the barrel and endcap regions [66].

### 2.5.6 Tagging of b-quark jet

The mass of b-quark is about 4.2 GeV and the hadrons formed from b-quark have masses of 5-6 GeV. The lifetime,  $\tau$  of b-hadrons is about 1.5 ps and this corresponds to  $c\tau$  value of 0.45 mm. If the momentum of b-hadron is about 200 GeV, the mean decay length is about 2 cm in the lab frame and this increase in decay length is because of Lorentz-boost. With the high position resolution of tracker, it is possible to identify such decays which take place away from the primary vertex (PV). These decay vertices are called secondary vertices and the impact parameter (IP) for such decays is high. In many cases, b-hadron decays result in a soft lepton. Figure 2.63 illustrates a b-jet containing a soft lepton and a secondary vertex (SV).

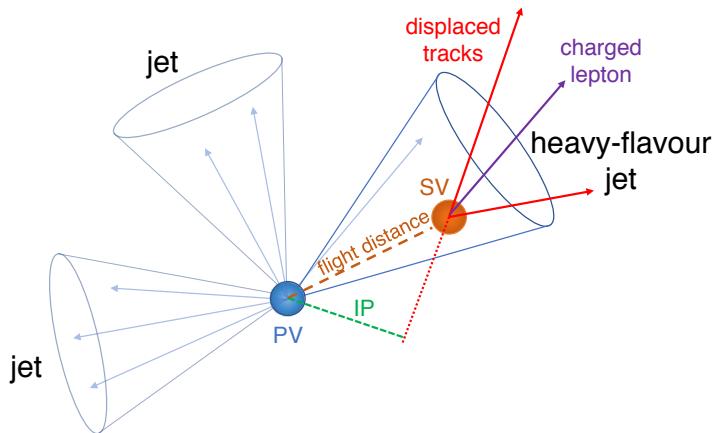


FIGURE 2.63: Illustration of a b-jet showing the decay of a b-hadron and a soft lepton inside the jet [93].

The algorithm used for identifying whether a jet is b-jet or not, is called combined SV version-2 (CSVv2) algorithm [93] which combines information about the displaced tracks and the SVs associated with a jet using a multivariate technique. The output of the CSVv2 algorithm is shown in figure 2.64 (left) for u, d, s or gluon jets, c-jets and b-jets in simulated  $t\bar{t}$  events. A neural network based CSVv2 tagger (DeepCSV) is also used by CMS.

The plot on the right side of figure 2.64 shows the efficiency for b-tagging using CSVv2 and DeepCSV as a function of particle-level jet  $p_T$ . These efficiencies are obtained from simulation after applying data to simulation corrections. The b-tagged jets (or b-jets) used in this search use CSVv2 - medium working point which has a 55% efficiency to correctly identify b-jets with  $p_T \approx 30$  GeV. The corresponding misidentification probabilities are 1.6% for gluon and light-flavor quark jets, and 12% for charm quark jets.

The efficiency of b-tagging is low at low  $p_T$  because of multiple Coulomb interactions and lower Lorentz-boost making the SV closer to PV. At high jet  $p_T$ , tracks in the jets are collimated and the curvature of the tracks are small which may cause close-by tracks to create overlapping hits in the tracker.

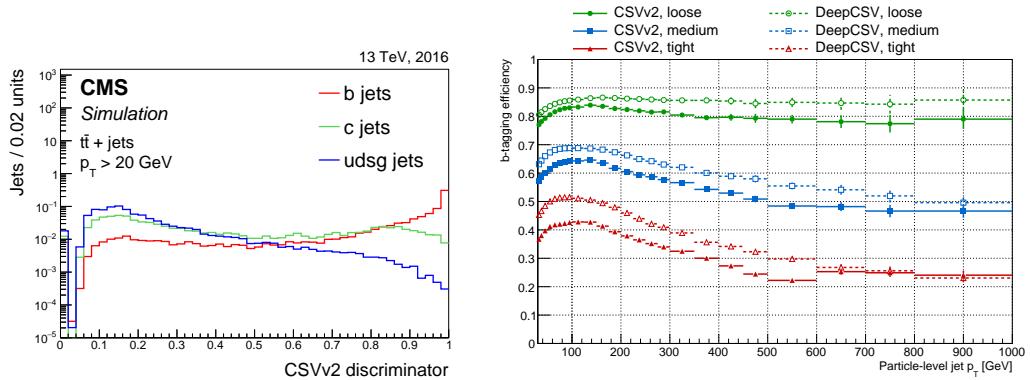


FIGURE 2.64: The output of the CSVv2 algorithm for different flavor jets (left) [93]. The right plot shows the efficiency of b-tagging for CSVv2 and DeepCSV as a function of particle-level jet  $p_T$  [35].

### 2.5.7 Missing transverse momentum

The pp collision takes place in  $z$  direction and hence the momentum before the collision in  $x$  and  $y$  directions, transverse direction, is zero. From momentum conservation, final momentum after the collision in transverse direction should be zero. In CMS, PF algorithm is used to reconstruct various particles, except the neutrinos, and these reconstructed particles are called as PF candidates. If there are  $n$  PF candidates each of them having a transverse momentum vector  $\vec{p}_T^i$ , and a neutrino with transverse momentum vector  $\vec{p}_T^\nu$ , we have

$$0 = \sum_i^n \vec{p}_T^i + \vec{p}_T^\nu \quad (2.15)$$

$$\vec{p}_T^\nu = - \sum_i^n \vec{p}_T^i \quad (2.16)$$

$$p_T^{miss} = |\vec{p}_T^\nu| = \left| - \sum_i^n \vec{p}_T^i \right| \quad (2.17)$$

The transverse momentum of neutrino,  $\vec{p}_T^\nu$  is nothing but the negative vector sum of all PF candidates. Magnitude of this vector is referred to as  $p_T^{miss}$  and it gives the exact  $p_T$  of neutrino if we detect and reconstruct all other particles and measure their transverse meomentum accurately. If there are multiple neutrinos or new particles which do not interact with detector, then their momenta get added vectorially and

the resulting vector behaves in the same way as  $\vec{p}_T^\nu$ , in the context of above equations. If PF algorithm is not able to reconstruct and get accurate measurement of

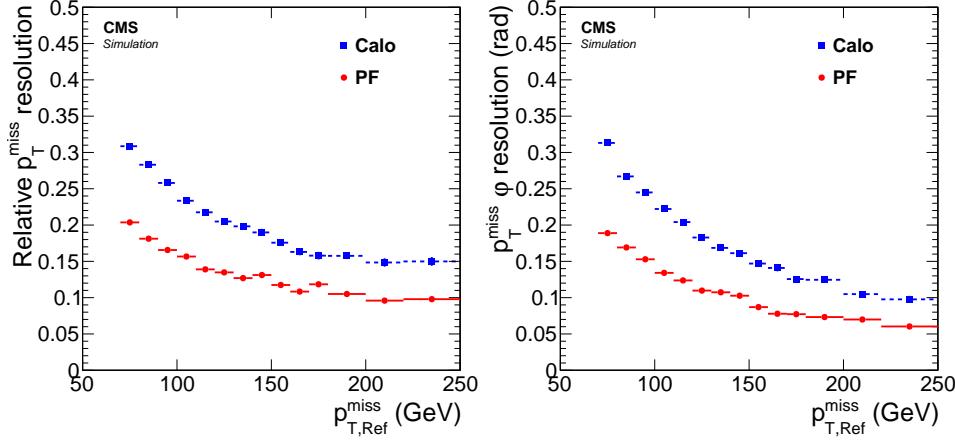


FIGURE 2.65: Relative  $p_T^{\text{miss}}$  resolution (left) and  $p_T^{\text{miss}}\phi$  resolution as a function of  $p_{T,\text{ref}}^{\text{miss}}$  in simulated  $t\bar{t}$  sample [91].

momenta all the particles, either because of detector imperfections or finite resolution while measuring momenta of various objects, we get incorrect measure of  $p_T^{\text{miss}}$ . The amount of *incorrectness* is termed as fake  $p_T^{\text{miss}}$ . As discussed, earlier, jets need various types of corrections to estimate their true 4-momenta. Once the jets are corrected, these corrections need to be propagated to  $p_T^{\text{miss}}$  since corrections on jets do not affect PF candidates. The corrected  $p_T^{\text{miss}}$  is given by

$$\vec{p}_T^{\text{miss, corr}} = \vec{p}_T^{\text{miss}} - \sum_{\text{jets}} (\vec{p}_{T,jet}^{\text{corr}} - \vec{p}_{T,jet}) \quad (2.18)$$

This  $p_T^{\text{miss}}$  is called *type1 corrected*  $p_T^{\text{miss}}$  and it uses all jets with  $p_T > 15$  GeV and having ECAL energy not more than 90% of the total jet energy. If there is a muon in the jet, then 4-momenta of the muon is subtracted from the jet, corrections are applied, 4-momenta of muon is added back to the corrected jet [36]. All the physics objects, electrons, muons, photons, muons and jets (to large extent), play a role in resolution of  $p_T^{\text{miss}}$ . Figure 2.65 shows the relative resolution of  $p_T^{\text{miss}}$  (left) and  $p_T^{\text{miss}}\phi$  (right) as a function of  $p_{T,\text{ref}}^{\text{miss}}$ , where  $p_{T,\text{ref}}^{\text{miss}}$  is calculated using all stable particles, except neutrinos, from the event generator [91].



## Chapter 3

# Search for supersymmetry in events with a photon, jets, b-jets, and missing transverse momentum

### 3.1 Introduction

This chapter describes a search for SUSY in events with one photon, large missing transverse momentum ( $p_T^{miss}$ ), and large hadronic activity. The targeted production mechanism is described by simplified models of gluino and stop production as depicted in Figure 1.5, though the analysis is more generically applicable to other production scenarios. This analysis focuses on SUSY models in which R-parity is conserved forcing the lightest supersymmetric particle (LSP), to be stable weakly interacting, leading to potentially large  $p_T^{miss}$ .

In general, gluinos and squarks will decay to at least one colored SM particles and the LSP through some number of SUSY particles in the form of cascades. Other SM particles, such as vector bosons or Higgs boson, can be produced in these cascades. Massive SM boson decay predominantly to hadronic final states. In particular, the Higgs boson decays predominantly to a pair of b-quarks. In addition to heavy flavor from Z and Higgs bosons, heavy flavor quarks can come from top and bottom squarks. These cases are particularly motivated by naturalness arguments, where squarks at least partially cancel the leading contributions of virtual corrections to the Higgs mass from SM particles.

Given the reasons stated above, final states with high jet and b-jet multiplicities and at least one photon are particularly well-motivated by the GMSB symmetry breaking mechanism and naturalness considerations. This analysis is an extension of existing single photon analysis previously [96], by taking full advantage of  $p_T^{miss}$ , large hadronic activity, and heavy flavor content. The analysis will search for large  $p_T^{miss}$  events in various regions of jet multiplicity and b-jet multiplicity, motivated by several SMS scenarios. The SM backgrounds are estimated from data control regions with as little reliance on simulations as possible. Section 3.2 discusses the SM and signal simulations used in the analysis. The triggers used for both the signal and control regions are described in Section 3.3. Event selections are detailed in Section 3.4. The background estimation techniques are discussed in Section 3.6. Signal systematic uncertainties are described in Section 3.7. Finally, results and a description of our statistical modeling of data and the inferred limits are described in Section 4.1.

## 3.2 MC simulation samples

Monte Carlo simulation is used to design the analysis, to provide input for background estimation methods that use data control regions, and to predict event rates from simplified models. Simulated SM background processes include jets produced through the strong interaction, referred to as quantum chromodynamics (QCD) multijets,  $t\bar{t}$ +jets,  $W$ +jets,  $Z$ +jets,  $\gamma$  + jets,  $t\bar{t}\gamma$ ,  $t\gamma$ , and  $V\gamma$  + jets ( $V = Z, W$ ). The SM background events are generated using the `MADGRAPH5_aMC@NLO` v2.2.2 or v2.3.3 generator [7, 68, 13] at leading order (LO) in perturbative QCD, except  $t\bar{t}\gamma$  and  $t\gamma$ , which are generated at next-to-leading order (NLO). The cross sections used for normalization are computed at NLO or next-to-NLO [7, 43, 56, 55]. The QCD multijets, diboson ( $V\gamma$ ), top quark, and vector boson plus jets events are generated with up to two, two, three, and four additional partons in the matrix element calculations, respectively. Any duplication of events between pairs of related processes—QCD multijets and  $\gamma$  + jets;  $t\bar{t}$ +jets and  $t\bar{t}\gamma$ ;  $W$ +jets and  $W\gamma$  + jets—is removed using generator information.

The NNPDF3.0 [16] LO (NLO) parton distribution functions (PDFs) are used for samples simulated at LO (NLO). Parton showering and hadronization are described using the PYTHIA 8.212 generator [97] with the CUETP8M1 underlying event tune [69]. Partons generated with `MADGRAPH5_aMC@NLO` and PYTHIA that would otherwise be counted twice are removed using the MLM [6] and FxFx [51] matching schemes in LO and NLO samples, respectively.

Signal samples are simulated at LO using the `MADGRAPH5_aMC@NLO` generator and their yields are normalized using NLO plus next-to-leading logarithmic (NLL) cross sections [21, 76, 75, 19, 20]. The decays of gluinos, top squarks, and neutralinos are modeled with PYTHIA.

The detector response to particles produced in the simulated collisions is modeled with the GEANT4 [4] detector simulation package for SM processes. Because of the large number of SUSY signals considered, with various gluino, squark, and neutralino masses, the detector response for these processes is simulated with the CMS fast simulation [2, 57]. The results from the fast simulation generally agree with the results from the full simulation. Where there is disagreement, corrections are applied, most notably a correction of up to 10% to adjust for differences in the modeling of  $p_T^{\text{miss}}$ .

## 3.3 Triggers

The analysis phase space is partly driven by the availability of single photon triggers. We make use of two single photon triggers for both the signal region and all control regions. For the fake-rate background prediction, which relies on extrapolating from single electron events, we also make use of single electron triggers. For validating the multijet backgrounds, where  $p_T^{\text{miss}}$  is dominated by jet mis-measurement, we make use of a zero photon control region, which is triggered using the inclusive  $H_T$  triggers.

These three triggering strategies and studies of their efficiencies are discussed in the following sections.

### 3.3.1 Signal triggers

The signal events for are triggered using a logical OR of HLT\_Photon90\_CaloIdL\_PFHT600 and HLT\_Photon165\_HE10 which require either a photon with  $p_T > 90$  GeV photon and  $H_T > 600$  GeV, or a photon with  $p_T > 165$  GeV respectively. In addition, photons need to have  $H/E < 0.10$ . Both of these triggers are seeded by level-1 triggers that are based on energy deposited in the ECal. The trigger efficiencies are studied using MET dataset by taking OR of inclusive  $p_T^{\text{miss}}$  triggers : HLT\_PFMET\*\_PFMHT\*\_IDTight, where \* corresponds to 90, 100, 110 & 120. Efficiency of the trigger is given by

$$\epsilon = \frac{\text{Events passing photon trigger(s) \& MET triggers}}{\text{Events passing MET triggers}} \quad (3.1)$$

Figure 3.1 shows the efficiency of the Photon165 trigger versus  $p_T^\gamma$  for different data-taking eras and for all 2016 data. No conditions are imposed on kinematic variables, except that the events should have at least one photon with  $p_T > 100$  GeV. This trigger becomes fully efficient for  $p_T^\gamma > 190$  GeV and has a plateau efficiency of 97.6%.

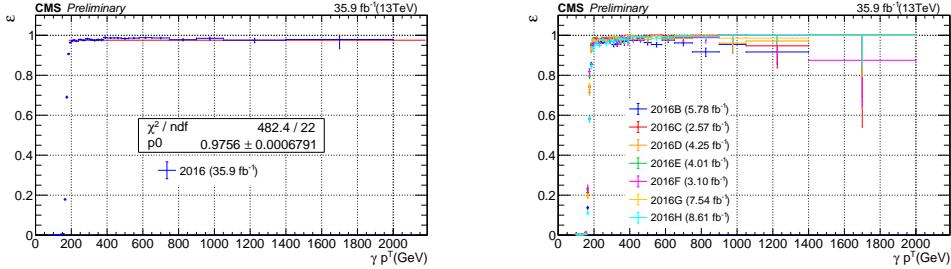


FIGURE 3.1: Efficiency of Photon165 trigger versus  $p_T^\gamma$  for all eras (left) and split by eras (right).

For measuring the efficiency of the Photon90\_PFHT600 trigger as a function of  $p_T^\gamma$ ,  $H_T^\gamma > 900$  GeV is used. To measure the efficiency as a function of  $H_T^\gamma$ ,  $p_T^\gamma > 100$  GeV is used. Figure 3.2 shows the efficiency of the Photon90\_PFHT600 trigger versus both  $p_T^\gamma$ , for events with  $H_T^\gamma > 900$  GeV, and versus  $H_T^\gamma (p_T^\gamma + \sum_{\text{jets}} p_T^{\text{AK4}})$ , for events with  $p_T^\gamma > 100$  GeV. Efficiencies are presented both separated by eras and for all 2016 data combined. The trigger is fully efficient for  $H_T^\gamma > 800$  GeV, but has an inefficiency at high  $H_T^\gamma$ . This inefficiency is ultimately recovered by the Photon165 trigger. The Photon90\_PFHT600 is found to be fully efficiency for  $p_T^\gamma > 100$  GeV for high  $H_T^\gamma$  events. At plateau, the efficiency is > 95% as a function of  $H_T^\gamma$  or  $p_T^\gamma$ .

The combined efficiency from the OR of the Photon165 and Photon90\_PFHT600 triggers is shown in Figure 3.3. This efficiency is measured by considering events with at least 2 jets and the following condition: ( $H_T^\gamma > 500$  &  $p_T^\gamma > 190$ ) OR ( $H_T^\gamma > 800$  &  $p_T^\gamma > 100$ ). The average efficiency is found to be 98% with a negligible statistical uncertainty. The kinematic dependence of the efficiency amounts to a roughly 2% variation and is taken as a systematic uncertainty, where relevant.

The single (tight) electron control region, used for estimating the fake photon background, is also triggered using a logical OR of the Photon165 and Photon90\_PFHT600 triggers. To ensure that there is no trigger bias in our estimations, we measure the trigger efficiency for events with a single tight electron, shown in Figure 3.4.

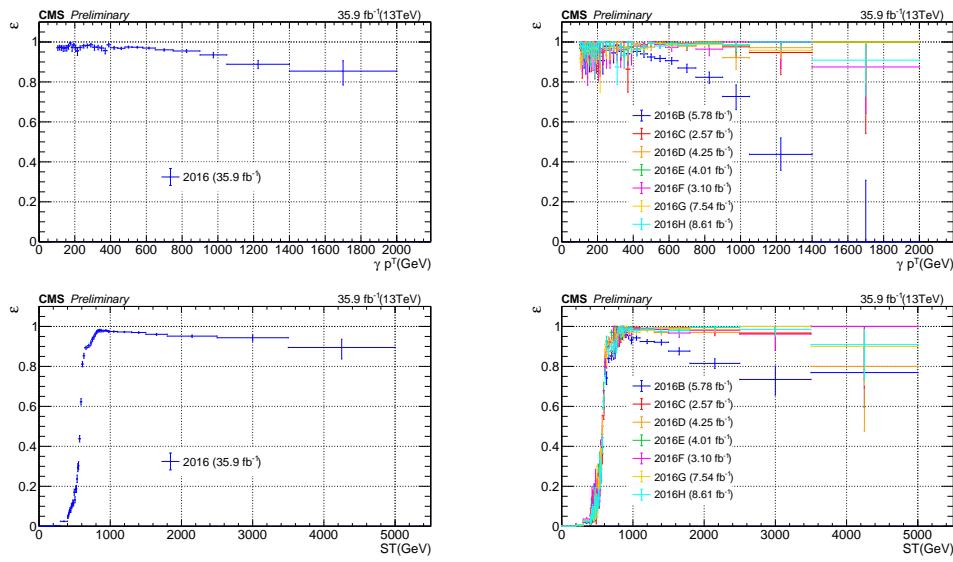


FIGURE 3.2: Efficiency of Photon90\\_PFHT600 trigger versus  $p_T^\gamma$  (top) and  $H_T^\gamma$  (bottom) for all eras (left) and split by eras (right).

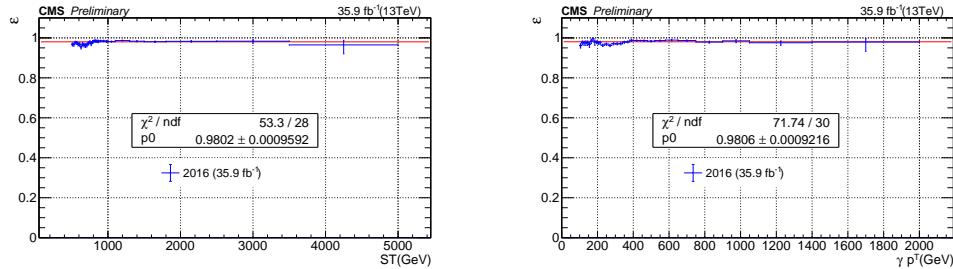


FIGURE 3.3: Efficiency of the OR of the Photon165 and Photon90\\_PFHT600 triggers versus  $H_T^\gamma$  (left) and  $p_T^\gamma$  (right).

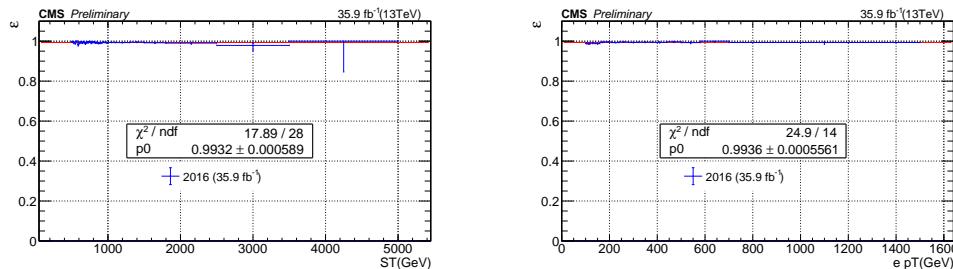


FIGURE 3.4: Efficiency of the OR of the Photon165 and Photon90\\_PFHT600 versus  $H_T^\gamma$  (left) and  $p_T^\gamma$  (right) for events with one tight electron.

### 3.3.2 Multijet triggers

Events with zero photons are used to validate the multijet background predictions. These events are triggered using an OR of the PFHT800 and PFHT900 triggers. The former HLT path is used only for the early data taking, and the later was used throughout 2016 data-taking without any prescaling. These triggers are observed to be inefficient at high  $H_T$ . To recover this inefficiency, a single jet trigger is also included, HLT\_CaloJet500\_NoJetID. The efficiency for these triggers is shown in Figure 3.5, which is measured using events which pass the inclusive PFMET\*\_PFMHT\*\_ triggers. The efficiency is found to plateau for at 98% for  $H_T > 1000$  GeV. The systematic uncertainty on this is taken to be +2%/-1%.

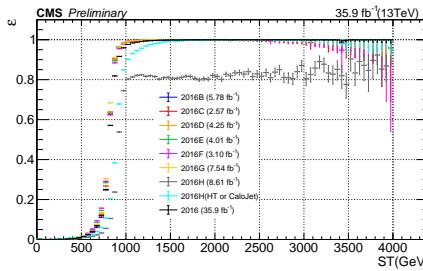


FIGURE 3.5: Efficiency of HLT\_PFHT800 trigger for eras B-G, and HLT\_PFHT900 OR HLT\_CaloJet500\_NoJetID for era H for zero photon events.

## 3.4 Event selection

The search regions for the analysis require large  $p_T^{miss}$ , large  $H_T^\gamma$ , one high energy photon, and no leptons. Jets used in this analysis are reconstructed from charged-hadron subtracted particle-flow (PF) candidates using the anti- $k_T$  algorithm [26] with the distance parameter 0.4. The PF algorithm is used to individually identify and reconstruct all particles produced in the collision (PF candidates); namely charged hadrons, photons, neutral hadrons, muons, and electrons [91]. The following requirements define the baseline selection:

- $H_T^\gamma > 500$  GeV for events with  $p_T^\gamma > 190$  GeV, and  $H_T^\gamma > 800$  GeV for events with  $p_T^\gamma > 100$  GeV, where  $H_T^\gamma = p_T^\gamma + \sum_{\text{jets}} p_T^{\text{AK4}}$ . The photon reconstruction and identification is described in the following.
- Photons: Photons are selected using ID criteria described in 2.5.4. Photon candidates whose supercluster is found to be in the pseudorapidity range of  $1.44 < |\eta| < 1.56$ , which is the EB-EE transition region are not considered. Photons are required to have  $p_T > 100$  GeV and  $|\eta| < 2.5$ .
- AK4PF jets used in this analysis are required to have  $p_T > 30$  GeV,  $|\eta| < 2.4$  and pass the loose jet ID requirements [64]:
  - For jets with  $|\eta| < 2.4$ :
    - neutral hadron fraction  $< 0.99$ ,
    - neutral EM fraction  $< 0.99$ ,
    - number of constituents  $> 1$ ,
    - charged hadron fraction  $> 0$ ,

- charged multiplicity  $> 0$ ,
- charged EM fraction  $< 0.99$

Jets that are matched to isolated photons, within  $\Delta R < 0.3$  are not subject to these requirements.

- Tagging of b-jets:

Although not used for a selection criterion, the number of selected jets satisfying the combined inclusive secondary vertex b-tagging algorithm at the medium working point ( $\text{CSVv2M} > 0.8484$ ) is used as a discriminating variable [93] and is used for deciding which signal region an event belongs to.

- $p_T^{\text{miss}} > 200\text{GeV}$  for signal region and the region  $100 - 200\text{ GeV}$   $p_T^{\text{miss}}$  is used in estimation of  $\gamma + \text{jets}$  background.
- Angular cut: The majority of  $\gamma + \text{jets}$  and QCD multijet events in our high- $p_T^{\text{miss}}$  search region have jets with undermeasured momenta and thus a spurious momentum imbalance. A signature of such an event is a jet closely aligned in direction with the  $p_T^{\text{miss}}$  vector. To suppress this background, we reject all events in which either of the two highest- $p_T$  jets lie within 0.3 radians of the  $p_T^{\text{miss}}$  vector in the azimuthal coordinate i.e.  $\Delta\phi(\vec{p}_T^{\text{miss}}, \vec{p}_T^{\text{jet}}) < 0.3$  for first two leading jets.
- Muon veto: Muon candidates are selected using medium-ID muon selection [92] with the additional requirements:

$$\begin{aligned} d_{xy}(\mu, \text{PV}) &< 0.2 \text{ cm} \\ d_z(\mu, \text{PV}) &< 0.5 \text{ cm} \end{aligned} \quad (3.2)$$

Muon candidates are required to have  $p_T > 10\text{GeV}$  and  $|\eta| < 2.4$ . To distinguish between prompt muons and muons from b-hadron decays, muons are required to satisfy an isolation requirement,  $I_{\text{mini}} < 0.2$ , where  $I_{\text{mini}}$  is the mini-isolation variable, whose description is similar to the one in 2.5.4 for electrons. Any event with a muon satisfying all of the above criteria is vetoed.

- Electron veto:

Electron candidates are selected using very loose criteria on its properties, called veto selection. They are required to have  $p_T > 10\text{GeV}$  and  $|\eta| < 2.5$  and satisfy an isolation requirement of  $I_{\text{mini}} < 0.1$ . Any event with an electron satisfying all of the above criteria is vetoed.

- Isolated track vetoes:

Following the event selection described above, including the muon and electron event vetoes, there is still some background in the search regions from  $t\bar{t}$ , single-top, and  $W+\text{jets}$  events with one  $W \rightarrow \ell\nu$  decay. In about half these background events, the  $W$  boson decays to a  $\tau$  lepton and the  $\tau$  lepton decays hadronically, while in the other half, an electron or muon is not identified or does not satisfy the criteria for an isolated electron or muon candidate given above. To suppress these backgrounds, we reject events with one or more isolated charged track.

The requirements for the definition of an isolated track differ slightly depending on whether the track is identified as leptonic or hadronic by the PF algorithm. For leptonic tracks, we require:

- $p_T > 5\text{GeV}$ ,
- $I_{\text{tk}} < 0.2$ ,

where  $I_{\text{tk}}$  is the scalar  $p_T$  sum of other charged tracks within  $\Delta R \equiv \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.3$  of the primary track, divided by the  $p_T$  value of the primary track. For hadronic tracks, we apply slightly tighter requirements:

- $p_T > 10\text{GeV}$ ,
- $I_{\text{tk}} < 0.1$ .

Isolated tracks are considered only if they satisfy

$$m_T(\text{tk}, p_T^{\text{miss}}) = \sqrt{2p_T^{\text{tk}} p_T^{\text{miss}} (1 - \cos \Delta\phi)} < 100 \text{ GeV}, \quad (3.3)$$

where  $p_T^{\text{tk}}$  is the transverse momentum of the track and  $\Delta\phi$  is the azimuthal separation between the track and  $\vec{p}_T^{\text{miss}}$ .

To reduce the influence of tracks from extraneous pp interactions (pileup), isolated tracks are considered only if their nearest distance of approach along the beam axis to a reconstructed vertex is smaller for the primary event vertex than for any other vertex.

- Event cleaning:

We reject events with a jet that satisfies  $p_T > 30 \text{ GeV}$  and  $|\eta| < 2.4$  if the jet fails the loose jet ID criteria given above. We apply event filters designed to reject events with spurious  $p_T^{\text{miss}}$  signals [36].

In almost all cases there is a jet matched to a photon within  $\Delta R < 0.3$ . Since this jet contains photon plus the activity near it,  $p_T^{\text{jet}} \geq p_T^\gamma$ . Any event which does not have a jet matched to the leading photon or the matching jet  $p_T$  is less than  $p_T^\gamma$  is also rejected.

## 3.5 SM background and search regions

The dominant backgrounds for these high  $p_T^{\text{miss}}$ , single photons events typically involve either fake  $p_T^{\text{miss}}$ , from  $\gamma$ -multijet events, or leptonic decays of weak vector bosons, which produce neutrinos. In electroweak events, photons can either arise from a genuine photon, either prompt or non-prompt, or from an electron faking a photon. While prompt and non-prompt photons are not distinguished in this analysis, fakes from electrons are estimated separately. Since photons can be produced from both simulations of matrix elements and parton shower models, via Pythia, MC simulations are often broken up based on whether or not the matrix element includes photons or not.

Figure 3.6 shows the expected distributions of  $p_T^\gamma$ ,  $p_T^{\text{miss}}$ ,  $N_{\text{jets}}$ ,  $N_{\text{b-jets}}$ , and  $H_T^\gamma$  for SM backgrounds and several signal points after the baseline selection mentioned in previous section.

Figure 3.7 shows the expected distributions of event yields for SM backgrounds and a few signal model points after the baseline selection. The final statistical interpretations are made using 25 independent signal regions. These regions are defined by  $N_{\text{jets}}$ ,  $N_{\text{b-jets}}$ , and  $p_T^{\text{miss}}$  selections. Six groups are defined based on  $N_{\text{jets}}$  and  $N_{\text{b-jets}}$ , they are listed in table 3.1 along with 6 sideband regions which are used for fake  $p_T^{\text{miss}}$  ( $\gamma + \text{jets}$  and QCD multijet) background. The sideband regions are not used for the

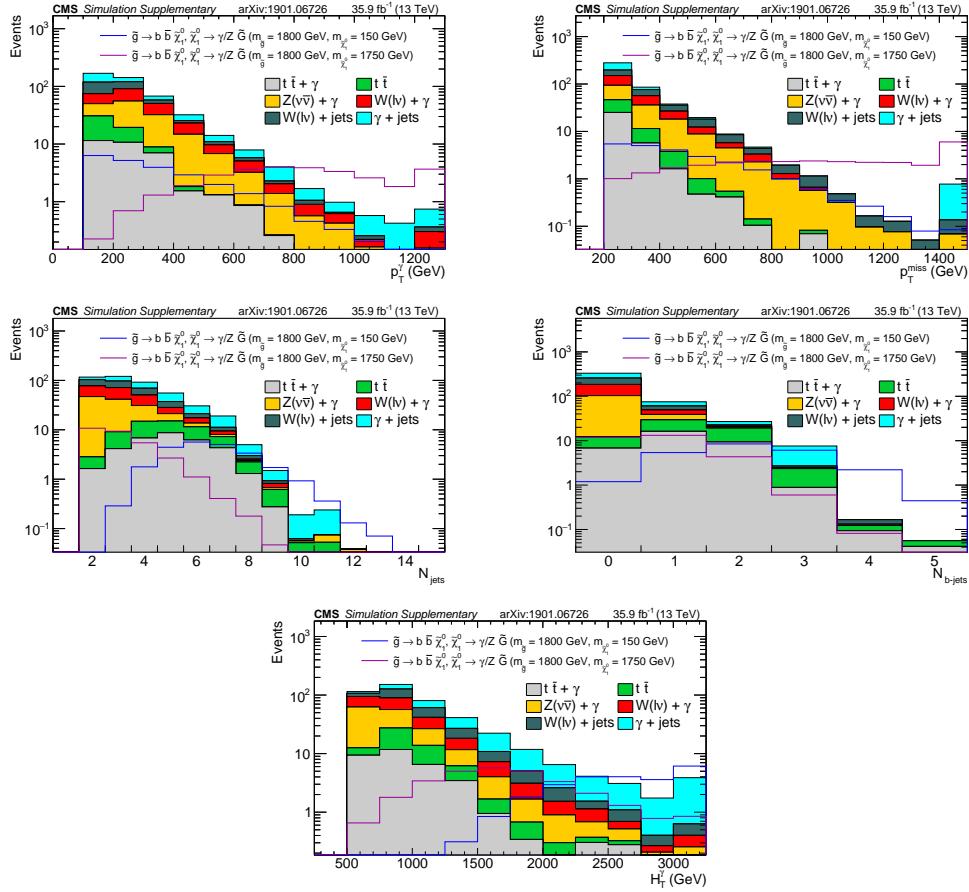


FIGURE 3.6: Distributions of  $p_T^\gamma$ ,  $p_T^{\text{miss}}$ ,  $N_{\text{jets}}$ ,  $N_{\text{b-jets}}$ , and  $H_T^\gamma$  after baseline selection. Filled, stacked histograms represent SM backgrounds, taken directly from simulation, open histograms show two examples of T5bbbbZG signals, one with a high mass (1750 GeV) NLSP and one with a low mass (150 GeV) NLSP. The gluino mass for both signals is 1800 GeV.

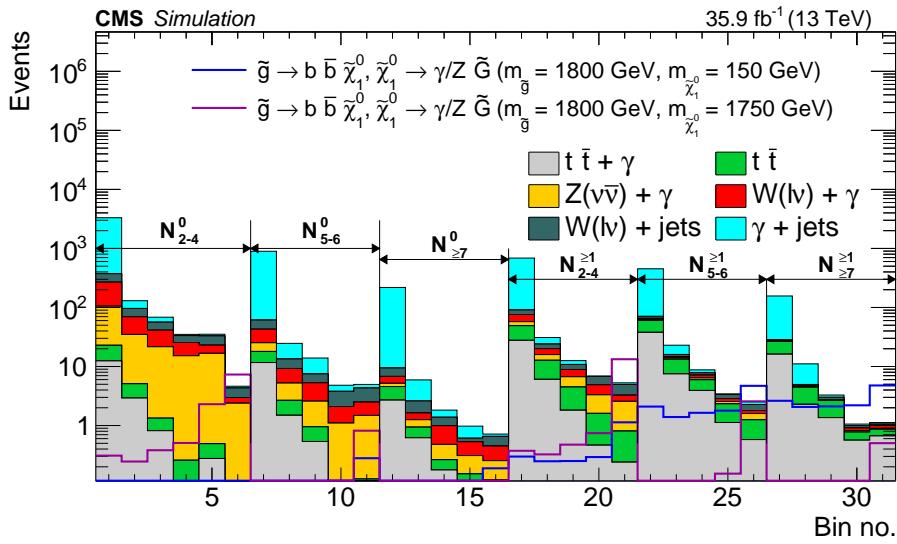


FIGURE 3.7: Distributions of events after baseline selection. Filled, stacked histograms represent SM backgrounds, taken directly from simulation, open histograms show two examples of T5bbbbZG signals, one with a high mass (1750 GeV) NLSP and one with a low mass (150 GeV) NLSP. The gluino mass for both signals is 1800 GeV. The lowest  $p_T^{\text{miss}}$  bin is not used for final limit calculations.

interpretations. Each of these regions,  $N_{\text{jets}}^{\text{b-jets}}$ , is further divided into 5  $p_T^{\text{miss}}$  regions whose boundaries are defined by  $100 < p_T^{\text{miss}} < 200$  (sideband),  $200 < p_T^{\text{miss}} < 270$ ,  $270 < p_T^{\text{miss}} < 350$ ,  $350 < p_T^{\text{miss}} < 450$ , and  $p_T^{\text{miss}} > 450$  GeV. In  $N_{2-4}^0$  the highest  $p_T^{\text{miss}}$  bin is further split such that the highest  $p_T^{\text{miss}}$  region corresponds to  $p_T^{\text{miss}} > 750$  GeV. The first bin in each of the 6 groups in Figure 3.7 corresponds to  $100 \text{ GeV} < p_T^{\text{miss}} < 200 \text{ GeV}$  sideband used for estimation of fake  $p_T^{\text{miss}}$  background.

TABLE 3.1: Signal region and sideband regions

$N_{\text{b-jets}}$	$N_{\text{jets}}$	Notation ( $N_{\text{jets}}^{\text{b-jets}}$ )	Bin number	
			Sideband	Signal region
0	2-4	$N_{2-4}^0$	1	2 - 6
	5-6	$N_{5-6}^0$	7	8 - 11
	$\geq 7$	$N_{\geq 7}^0$	12	13 - 16
$\geq 1$	2-4	$N_{2-4}^{\geq 1}$	17	18 - 21
	5-6	$N_{5-6}^{\geq 1}$	22	23 - 26
	$\geq 7$	$N_{\geq 7}^{\geq 1}$	27	28 - 31

### 3.5.1 Optimization of search bins

The above signal region definitions were optimized by maximizing sensitivity to some of the SMS models shown in Figure 1.5. For a given gluino mass if the mass of NLSP is close to mass of gluino mass, then the event has large  $p_T^{\text{miss}}$  and soft jets leading to low jet multiplicity. On the other hand, if the mass of NLSP is much less than mass of gluino, then there is small  $p_T^{\text{miss}}$ , but large hadronic activity leading to many jets in the final state. To cover both of these scenarios, search region is divided

in bins of  $N_{\text{jets}}$ ,  $N_{\text{b-jets}}$  and  $p_T^{\text{miss}}$ . Adding  $N_{\text{b-jets}}$  binning along with  $N_{\text{jets}}$  and  $p_T^{\text{miss}}$  bins gives higher sensitivity if there are any b quarks in the signal model. So these 3 variables are chosen to obtain higher sensitivity. 3 regions are defined based on  $N_{\text{jets}}$ : 2-4, 5-6 and  $\geq 7$ . These regions can be represented as  $N_{2-4}$ ,  $N_{5-6}$ ,  $N_{\geq 7}$ . Each of these regions is subdivided into 3 regions based on  $N_{\text{b-jets}}$ : 0, 1 and  $\geq 2$  (represented as  $N^0, N^1, N^{\geq 2}$ ). This gives total of 9 regions; all these 9 regions are further divided in  $p_T^{\text{miss}}$  dimension with the following bin boundaries : 100-120, 120-160, 160-200, 200-270, 270-350, 350-450,  $\geq 450$ . In total there are 63 search bins. This binning can give very good sensitivity, but there will not be enough statistics in the control regions to predict the background. In this analysis, single electron +  $\gamma$  control sample (CS), used for lost electron estimation, is one of the smallest control samples. These 63 bins need to be merged in  $N_{\text{jets}}$ ,  $N_{\text{b-jets}}$  or  $p_T^{\text{miss}}$  to avoid too many bins with 0 CS events; without sacrificing too much on sensitivity. Different ways of merging the bins is studied and for each of them expected exclusion curve is plotted in figure 3.8 for T5qqqqHG (left) and T5bbbbZG (right) models.

In the legends, numbers shown in subscript refers to  $N_{\text{jets}}$  and superscript refers to  $N_{\text{b-jets}}$ . These 63 bins are represented as  $(N_{2-4}, N_{5-6}, N_{\geq 7}) \times (N^0, N^1, N^{\geq 2})$  with each of these regions subdivided into 7  $p_T^{\text{miss}}$  bins.

In figure 3.8, red curve (corresponding to 63 bins) gives the best exclusion, but it has many bins with low CS events. Next best exclusion is obtained with orange curve which corresponds to 7  $p_T^{\text{miss}}$  bins  $\times$  3 bins in  $N_{\text{jets}}$   $\times$  2 bins in  $N_{\text{b-jets}}$ . Based on this scenario final search bins are determined. To increase the sensitivity for compressed scenario (mass of gluino is close to mass of NLSP), last  $p_T^{\text{miss}}$  bin in  $N_{2-4}^0$  is further divided into 450-750 and  $\geq 750$  GeV.

These optimization studies are done based on MC samples, by considering MC statistical uncertainties and additional 20% uncertainty for each bin and for each of the MC samples, including signal samples.

The final search region starts at  $p_T^{\text{miss}} \geq 200$  GeV, whereas in these studies, regions  $\geq 100$  GeV are also considered. This does not affect the optimization, since  $100 < p_T^{\text{miss}} < 200$  GeV region is background enriched and does not contribute to sensitivity.

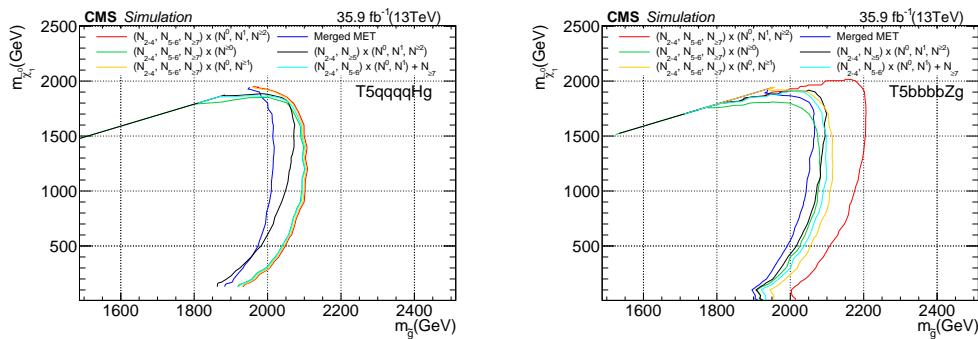


FIGURE 3.8: Expected exclusion curves for different types of binning for T5qqqqHG (left) and T5bbbbZG (right). In the legends, subscript  $j$  in  $N_j^i$  refers to number of jets and superscript  $i$  refers to number of b-tagged jets. The red curve corresponds to 63 search bins defined as  $(N_{\text{jets}} = 2-4, 5-6, \geq 7) \times (N_{\text{b-jets}} = 0, 1, \geq 2) \times (7 p_T^{\text{miss}} \text{ bins})$ . Other curves are obtained by combining these 63 bins in either  $N_{\text{jets}}$  or  $N_{\text{b-jets}}$  or  $p_T^{\text{miss}}$  or any combination of these variables. The final binning of the analysis is based on orange curve defined as  $(N_{\text{jets}} = 2-4, 5-6, \geq 7) \times (N_{\text{b-jets}} = 0, \geq 1) \times (7 p_T^{\text{miss}} \text{ bins})$ .

## 3.6 Background Estimation

In order to have a robust background prediction over all analysis bins, it is necessary to individually estimate the various background components since the composition is expected to vary dramatically in various bins (Figure 3.9). Low  $p_T^{\text{miss}}$  regions have more  $\gamma + \text{jets}$  background, high  $p_T^{\text{miss}}$  regions with low  $N_{\text{jets}}$  or  $N_{\text{b-jets}}$  is dominated by processes involving  $Z$  or  $W$ , and the high  $p_T^{\text{miss}}$ , high  $N_{\text{jets}}$ , high  $N_{\text{b-jets}}$  region is mainly  $t\bar{t}+\text{jets}$  or  $t\bar{t}\gamma$  dominated.

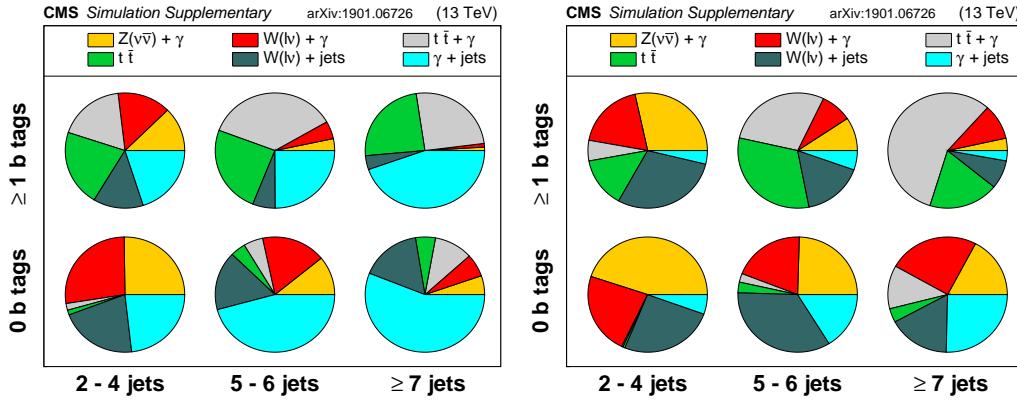


FIGURE 3.9: Pie charts showing relative background composition in various  $N_{\text{jets}}$  and  $N_{\text{b-jets}}$  regions for  $p_T^{\text{miss}}$  of 200 - 350 GeV on left and  $p_T^{\text{miss}} \geq 350$  GeV on right.

There are four main sources of background events: events with a lost lepton (lost  $e/\mu$ ) or a hadronically decaying tau leptons, events in which an electron from a  $W \rightarrow e\nu$  decay fakes a photon,  $Z \rightarrow \nu\bar{\nu}$  events produced in association with a  $\gamma$ , and fake  $p_T^{\text{miss}}$  events.

The lost lepton background arises from events in which the charged lepton from a leptonically decaying  $W$ , produced directly or from the decay of a top quark, fails to be identified by being out of acceptance, or failing the identification or isolation requirements. For example, in events with high  $p_T$  top quarks, the top quark decay products will be collimated, forcing the b-jet to be closer to the charged lepton. In these cases, events are more likely to fail the isolation requirements. These events are estimated by studying events with both a well-identified photon and lepton ( $e/\mu$ ) in both data and simulation.

The hadronic- $\tau$  background arises from events in which  $\tau$  leptons from  $W$  decays decay to mesons and neutrinos, which occurs  $\sim 65\%$  of the time. Due to lepton universality, the fraction of hadronically decaying  $\tau$  events can be estimated from single muon after correcting for reconstruction differences and for the hadronic- $\tau$  branching fraction.

Fake photon events primarily arise from electrons from  $W$  decays faking photons. This can happen when a pixel seed fails to be associated with the photon candidate. Given a fake rate, which relates events with well-identified electrons with events with well-identified photons, the fake photon background can be estimated from a single electron (zero photon) control region. The fake rate will be estimated in MC and data/MC corrections are studied to account for any MC mismodeling.

Invisible  $Z$  decays constitute a major background for low  $N_{\text{b-jets}}$  or  $N_{\text{jets}}$  multiplicity, high  $p_T^{\text{miss}}$  events. These events have a natural counter part,  $Z(\mu\mu)$  and  $Z(ee)$  events, which can be studied in both data and MC to predict the  $Z(\nu\nu)$  contribution.

Fake  $p_T^{miss}$  events occur primarily in events with a single photon and jets, where one of the jets' energy is mismeasured. These events primarily contribute to the low  $p_T^{miss}$  phase space. These events are greatly reduced by requiring events to have a large angle, in the transverse plane, between  $p_T^{miss}$  and the leading and subleading jets,  $\Delta\phi > 0.3$ . There are several control regions that will be used to estimate these events, events with zero photons, events with a photon, but after inverting the  $\Delta\phi$  cuts, finally, the whole method will be anchored to the high  $\Delta\phi$ , photon signal region with low  $p_T^{miss}$ . Region  $\Delta\phi(p_T^{miss}, p_T^{\text{jet}}) > 0.3$  for first two leading jets is defined as high  $\Delta\phi$  and low  $\Delta\phi$  is the region with does not pass this criteria (one of the jets or both of the jets are aligned,  $\Delta\phi < 0.3$ , with  $p_T^{miss}$ ).

### 3.6.1 Lost lepton estimation

The lost lepton background is estimated from single lepton prompt ( $e, \mu$ ) plus photon control regions. The definition of electrons and muons are the same as those used for the signal region veto defined in Section 3.4. Both the control regions require exactly one lepton ( $e, \mu$ ) and veto the opposite flavor lepton ( $\mu, e$ ). In order to reduce the effect of signal contamination from signals with W bosons, for example, events in the control region are also required to satisfy  $m_T = \sqrt{2p_T^\ell p_T^{miss}(1 - \cos(\Delta\phi))} < 100 \text{ GeV}$ . Figure 3.10 shows the distribution of  $m_T$  for  $\mu\gamma$  events for SM processes and representative signal models (signal is scaled up by a factor of 10). For standard model events,  $m_T$  is constrained by the W boson mass, whereas for the signal, missing transverse momentum from the gravitinos is included in  $m_T$ . The region with  $m_T < 100 \text{ GeV}$  provides a background-enriched sample with very small signal contamination.

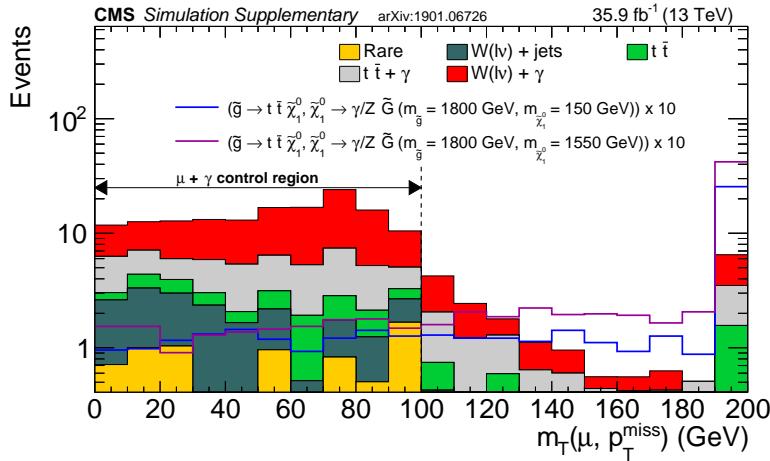


FIGURE 3.10: The transverse mass,  $m_T$  distribution for  $\mu\gamma$  events. Standard model processes are shown as filled histograms, representative signals of T5ttttZG model are shown as solid lines, and the signal is scaled up by a factor of 10 for better visualization.

All other selections for the control region, isolated track vetos,  $H_T^\gamma$ ,  $p_T^{miss}$ ,  $N_{\text{jets}}$ , and  $\Delta\phi$  are the same as the signal region. The control region events are triggered using the same triggers as the signal region and the same trigger efficiency correction/efficiencies are applied.

The method relies on weighting single lepton events. Event weights are derived from simulations for electron and muon control samples separately. The event

TABLE 3.2: Parameterization of transfer factors for lost-e and lost- $\mu + \tau_{\text{had}}$  events.

$N_{\text{jets}}^{\text{b-jets}}$	$p_T^{\text{miss}}$	Event weight			
		High- $\Delta\phi$		Low- $\Delta\phi$	
		Lost-e	Lost- $\mu + \tau_{\text{had}}$	Lost-e	Lost- $\mu + \tau_{\text{had}}$
$N_2^0$	100-150	$0.55 \pm 0.06$	$1.05 \pm 0.07$	$0.23 \pm 0.04$	$0.75 \pm 0.10$
	$\geq 150$	$0.42 \pm 0.04$	$0.84 \pm 0.06$	$0.24 \pm 0.04$	$0.66 \pm 0.10$
$N_3^0$	100-150	$0.30 \pm 0.04$	$0.81 \pm 0.06$	$0.20 \pm 0.04$	$0.59 \pm 0.12$
	$\geq 150$	$0.30 \pm 0.04$	$0.67 \pm 0.06$	$0.21 \pm 0.04$	$0.54 \pm 0.13$
$N_4^0$	100-150	$0.29 \pm 0.06$	$0.62 \pm 0.10$	$0.26 \pm 0.08$	$0.43 \pm 0.12$
	$\geq 150$	$0.64 \pm 0.06$	$1.23 \pm 0.08$	$0.20 \pm 0.04$	$0.67 \pm 0.07$
$N_{5-6}^0$	100-150	$0.30 \pm 0.02$	$0.97 \pm 0.05$	$0.16 \pm 0.03$	$0.60 \pm 0.07$
	$\geq 150$	$0.36 \pm 0.03$	$0.79 \pm 0.05$	$0.11 \pm 0.02$	$0.43 \pm 0.05$
$N_{\geq 6}^0$	100-150	$0.25 \pm 0.03$	$0.64 \pm 0.04$	$0.17 \pm 0.03$	$0.54 \pm 0.07$
	$\geq 150$	$0.33 \pm 0.08$	$0.73 \pm 0.10$	$0.17 \pm 0.05$	$0.59 \pm 0.20$
$N_{2-4}^{\geq 1}$	100-150	$0.51 \pm 0.05$	$0.87 \pm 0.09$	$0.43 \pm 0.06$	$0.98 \pm 0.18$
	$\geq 150$	$0.39 \pm 0.04$	$0.70 \pm 0.07$	$0.21 \pm 0.06$	$0.81 \pm 0.15$
$N_{5-6}^{\geq 1}$	100-150	$0.44 \pm 0.08$	$0.64 \pm 0.09$	$0.27 \pm 0.09$	$0.45 \pm 0.17$
	$\geq 150$	$0.57 \pm 0.05$	$1.11 \pm 0.10$	$0.27 \pm 0.04$	$0.68 \pm 0.10$
$N_{\geq 6}^{\geq 1}$	100-150	$0.35 \pm 0.04$	$0.79 \pm 0.07$	$0.25 \pm 0.04$	$0.54 \pm 0.07$
	$\geq 150$	$0.36 \pm 0.06$	$0.67 \pm 0.10$	$0.25 \pm 0.06$	$0.53 \pm 0.10$

weights are derived according to,

$$N_{\text{lost}-\ell}^{\text{pred}} = N_{\ell}^{\text{data}}(p_T^{\text{miss}}, N_{\text{jets}}, N_{\text{b-jets}}) \cdot T^{\text{MC}}(p_T^{\text{miss}}, N_{\text{jets}}, N_{\text{b-jets}})$$

where  $T^{\text{MC}}(p_T^{\text{miss}}, N_{\text{jets}}, N_{\text{b-jets}}) = N_{\text{lost}-\ell}^{\text{MC}}(p_T^{\text{miss}}, N_{\text{jets}}, N_{\text{b-jets}}) / N_{\ell}^{\text{MC}}(p_T^{\text{miss}}, N_{\text{jets}}, N_{\text{b-jets}})$

(3.4)

$N_{\ell}^{\text{data}}$  and  $T^{\text{MC}}$  are the observed yields in the control region and the simulation-derived event weight for a corresponding search region. For the muon ( $\ell = \mu$ ) transfer factors, we include events with zero e, $\mu$ , but at least one hadronic- $\tau$  in the numerator of the transfer factor.

Table 3.2 shows the event weights for the two lost-lepton control regions. These weights are typically 0.5 (0.75) for the e ( $\mu$ ) control regions.

To test the parameterization of the event weights, the lost lepton method is evaluated on simulated data. Figure 3.11 shows comparisons of the predicted event yield in each of the search regions and the true event yield from simulation for the electron and  $\mu + \tau_{\text{had}}$  events. The event weight parameterization is found to predict the true event yield for each of the search regions with the largest observed deviation corresponding to a 30% discrepancy, which is consistent with expectations from statistical fluctuation only. The uncertainty on the event weights from limited MC statistics is propagated to the final predictions. Any non-closure seen in the tighter  $p_T^{\text{miss}}$  region will be covered by the control sample statistical uncertainty in data.

To model the uncertainty on the prediction due to limited statistics from the single lepton control regions, each prediction will have a prior distribution modeled by a gamma distribution that uses the observed control region statistics and the average transfer factor. The average transfer factor is defined to be the prediction divided by the raw observed control region event yield. In bins where there are no events observed, the average transfer factor is computed based on MC events yields.

Other sources of systematic uncertainty correspond to effects related to:

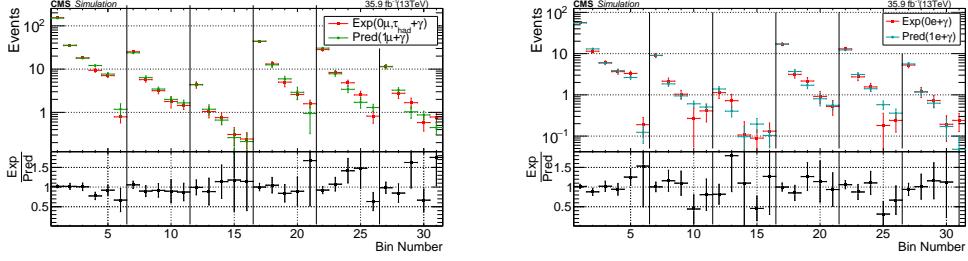


FIGURE 3.11: Expected closure of lost electron (left) and lost  $\mu + \tau_{\text{had}}$  (right) predictions where simulated event yields are treated like data.

- lepton scale factors
- b-tagging scale factors
- PDF and scale uncertainties
- modeling of  $m_T$  in simulations
- modeling of colinear photons

The PDF and scale uncertainties are studied by varying the MC weights according to 101 and 9 weight sets respectively. For each variation the average transfer factor for all signal regions combined is computed. The maximum variation of the lost  $\mu + \tau_{\text{had}}$  average transfer factor is found to be 1.3% and 0.7% for the PDF and scale variations, respectively. The maximum variation of the lost  $\mu\mu$  average TF is found to be 1.3% and 0.7% for the PDF and scale variations respectively. For the lost-e, 4.8% and 1.0% variations were found in TF because of the PDF and scale variations, respectively. The maximum variation of these alternative weights is added as a systematic uncertainty on the lost-lepton and  $\tau_{\text{had}}$  prediction. A single nuisance parameter is used to model this uncertainty, correlating all signal regions and correlating the effect of PDF and scale uncertainties for both lost-e and lost- $\mu + \tau_{\text{had}}$  predictions.

The effect of JEC uncertainties are studied by varying both jet  $p_T$  and  $p_T^{\text{miss}}$ . This variation can cause events to migrate below the minimum  $p_T^{\text{miss}}$  or  $H_T^\gamma$  thresholds, or above the maximum  $m_T$  threshold for single lepton events. No trends are observed versus the various signal region. There is a 2% (0.6%) effect on the lost-e (lost- $\mu + \tau_{\text{had}}$ ) transfer factor, averaged over all signal regions. The 2% effect for the lost-e prediction is applied as a systematic uncertainty. The effect on the lost- $\mu + \tau_{\text{had}}$  prediction is neglected. The same correlation model as used for the PDF and scale uncertainties is applied for the JEC uncertainty systematic.

The modeling of the  $m_T$  distribution is potentially effected by generator level cuts in MC samples, but this effect is found to be less than 1% and is neglected.

The effect of uncertainties on the lepton scale factors are propagated to the lost-e and lost- $\mu + \tau_{\text{had}}$  transfer factors. These are found to have a 2% effect, which is assumed to be correlated among common  $p_T^{\text{miss}}$  bins.

Finally the effect of b-tagging scale factor uncertainties was checked. It was found to have < 1% effect on the lost-lepton transfer factors and is neglected.

Comparisons between simulation and data for single lepton events show a systematic mismodeling of events with small angle between the  $e/\mu$  and the photon. This is due to generator level cuts on the angle between photons and partons, in Madgraph W/ $t\bar{t} + \gamma$  samples, of  $\Delta R < 0.5$ , which is not well modeled by pythia

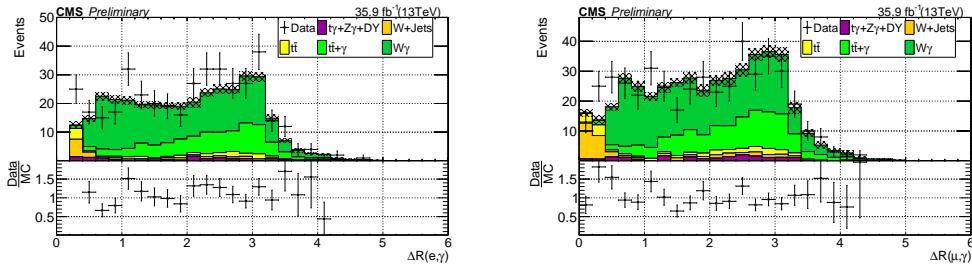


FIGURE 3.12: Distribution of  $\Delta R(\ell, \gamma)$  for single electron (left) and single muon (right) events.

$W/t\bar{t}$ +jets samples. Figure 3.12 shows the  $\Delta R(e, \gamma)$  and  $\Delta R(\mu, \gamma)$  distributions. The region  $\Delta R(e, \gamma) < 0.2$  is excluded due to the photon/electron isolation cuts. To assess the effect of the missing low  $\Delta R$  phase space, the average transfer factor is derived for events with  $\Delta R > 0.5$  and compared to the average transfer factor computed with all events. The effect is found to be a 12% effect on the lost-e transfer factor and < 1% on the lost- $\mu + \tau_{\text{had}}$  transfer factor. Since no trends are found versus various signal regions, a flat 12% uncertainty is applied to lost-e prediction. No systematic uncertainty is applied to the lost- $\mu + \tau_{\text{had}}$  prediction.

The observed event yields and the corresponding predictions are shown in Table A.1 (for high  $\Delta\phi$ ) and Table A.2 (for low  $\Delta\phi$ ) of appendix A. Comparisons of the observed yields and expected MC yields for the  $e + \gamma$  and  $\mu + \gamma$  control regions in Figures B.2 and B.1 of appendix B.

### 3.6.2 Estimation of events with electrons faking photons

The fake-rate estimation predicts the expected rate from  $W(ev)$  events where the electron fakes a photon. The method relies on computing an  $e \rightarrow \gamma$  fake rate, derived from a mixture of simulation and data, and applying the fake rate as an event weight to the single electron control region, in which we veto events with a photon.

The kinematic selection for the single electron control region is exactly the same as the single photon control region except that we require exactly zero photons, exactly one electron with tight ID criteria, and use the tight electron 4-vector in place of the photon 4-vector where relevant.

The crux of this method is the derivation of the fake rate. The fake rate in a given region of phase space is defined to be the ratio of events with photons and no electron with respect to events with exactly one tight electron and zero photons. The fake rate will be derived using simulated  $W/t\bar{t}$  events. However, due to MC modeling of various nuisances relevant to electron reconstruction, an overall difference is to be expected between data and MC fake-rates. To correct for this, we will also measure the fake-rate in data and MC using DY events with a tag-and-probe (T&P) method. These T&P measurements will be used to correct the fake rates derived on  $W/t\bar{t}$  simulations.

The overall prediction is then given by:

$$N_{e \rightarrow \gamma} = N_e^{\text{data}} \cdot f^{tt,W} \cdot \beta^{\text{data}/\text{MC}} \cdot \frac{\epsilon_{SR}^{\text{trig}}}{\epsilon_{CR}^{\text{trig}}}, \quad (3.5)$$

where  $f^{tt,W}$  is the fake-rate derived from  $W/t\bar{t}$  simulation,  $\beta$  represents the T&P corrections factors, and  $\epsilon$  represents the measured trigger efficiency.

There are two important parameterizations, namely,  $f^{tt,W}$  and  $\beta$ .

### Fake-rate parameterization

The fake rate is parameterized as a function of local metrics relevant to electron reconstruction; the electron  $p_T$  and the charged track multiplicity,  $Q_{mult}$  around the electron are found to give the best closure on MC. The metric  $Q_{mult}$  is defined to be the sum of charged PF candidates with the jet matched,  $\Delta R < 0.3$ , to the electron. Since we require electrons have at least  $p_T > 100$  GeV, the jet clustering algorithm almost always clusters the electron into an AK4 jet. On the rare occasions that there is no matched jet, the event is rejected. The distribution of  $Q_{mult}$  is shown in Figure 3.13 for DY, W+jets, and  $t\bar{t}$  events.

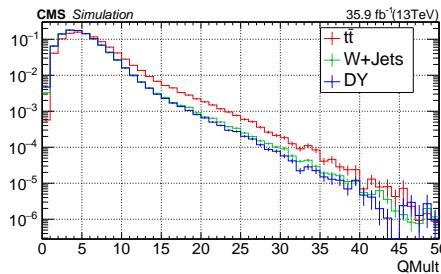


FIGURE 3.13:  $Q_{mult}$  distribution in a jet matched to the electron in DY, W+jets, and  $t\bar{t}$  events. Each of these distribution are scaled to unit area.

When deriving the fake-rate from simulation events are divided into two categories, signal region (SR) events, and control region (CR) events. Signal region events are those with zero electrons found and an EM objects matched,  $\Delta R < 0.2$ , to a gen-level electron. Control region events are those with 1 tight electron ( $p_T > 100$  GeV), exactly zero photons, and  $m_T(e, p_T^{\text{miss}}) < 100$  GeV; the last requirement increases the purity of the control sample and avoids overlap with signal events. After these selections are applied, the fake rate is then defined as  $f^{tt,W} = N_{SR}/N_{CR}$ .

Table A.3 shows the parameterization of the fake rate that is used. To validate this parameterization, the fake rate method is applied to MC and the prediction is compared to the true yields from simulation. Figure 3.14 show the comparison of the prediction (applying the fake rate parameterization to single electron MC events) and true yields; the agreement is found to be within 20%.

### Data/MC fake rate corrections

The data/MC corrections for the fake rate are computed using a T&P method on Drell-Yan events. Z bosons are reconstructed using a tag object and a probe object. The tag object definition is one tight electron with  $p_T > 40$  GeV. The probe can either be one tight electron or one photon with  $p_T > 100$  GeV. Event selections otherwise includes,  $N_{\text{jets}} \geq 2$ ,  $p_T^{\text{miss}} < 100$  GeV, zero isolated tracks (not including the tag/probe objects), and  $\Delta R(\text{tag}, \text{probe}) > 0.2$ . In the case where the probe satisfies the tag criteria, the tag is chosen randomly and the event is counted twice.

Once the photon and electron candidates are formed, the  $m_{ee}$  distribution is fit using a Breit-Wigner convoluted with a Gaussian to model the Z boson peak and a polynomial to model the combinatorial background. The fit is performed on both data and MC in regions of  $Q_{mult}$ , which is the variable for which the discrepancies

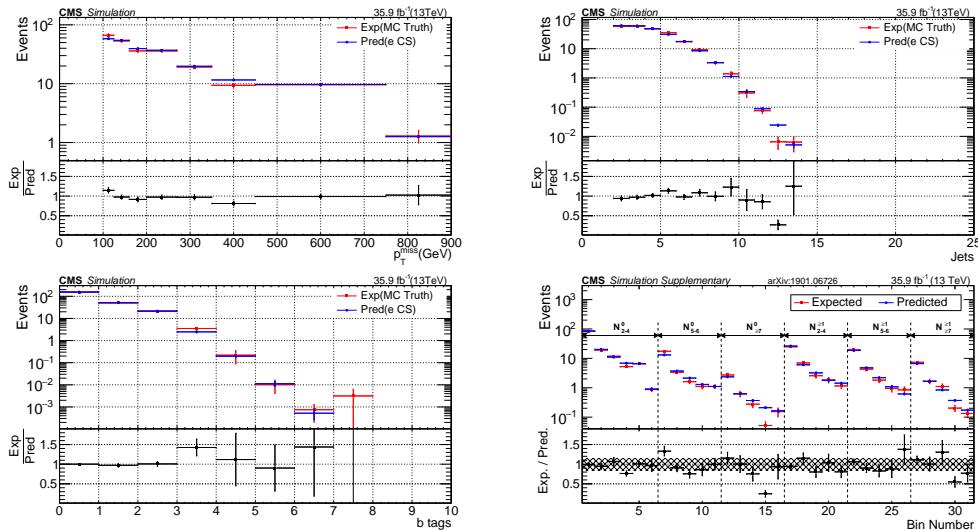


FIGURE 3.14: Comparison of fake rate prediction using fake rate parameterization and true MC yields using  $W/t\bar{t} + (\gamma)$  simulations versus  $p_T^{\text{miss}}$  (top left),  $N_{\text{jets}}$  (top right),  $N_{\text{b-jets}}$  (bottom left) and search bins (bottom right). The hashed region in the lower panel of bottom right plot shows the total systematic uncertainty.

between data and MC are found to be the largest. Table 3.3 shows the fitted fake rate in data and MC, and the corresponding correction factor in each of the  $Q_{\text{mult}}$  bins. Figure 3.15 shows the fit results for each of the fake rate measurements.

TABLE 3.3: fake rate as measured by T&P procedure on Drell-Yan events in data and MC.

$Q_{\text{mult}}$	MC fake rate	data fake rate	scale factor
0-1	$0.014 \pm 0.002$	$0.016 \pm 0.003$	$1.17 \pm 0.27$
$\geq 2$	$0.015 \pm 0.001$	$0.018 \pm 0.002$	$1.21 \pm 0.16$

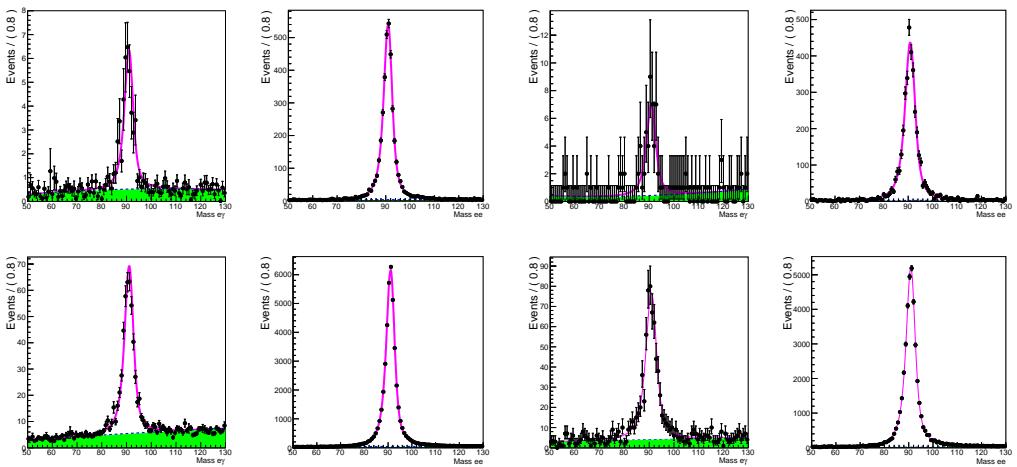


FIGURE 3.15: Tag & probe fits of the fake rate in  $0-1$   $Q_{\text{mult}}$  bins (top row) and  $\geq 2$   $Q_{\text{mult}}$  bins (bottom row). The first and second columns are the the photon region and electron region fits in MC, respectively. The third and fourth columns are the photon region and electron region fits in data, respectively.

### Fake rate prediction

Using the results described above, the prediction for the fake photon background is shown in Table A.4 (for high  $\Delta\phi$ ) and Table A.5 (for low  $\Delta\phi$ ).

### 3.6.3 Invisible Z estimation

The invisible Z background is estimated using a  $Z(\ell\ell)$  plus  $\gamma$  control region. The method relies on the fact that for  $\gamma Z(\nu\nu)$  events, the  $p_T^{miss}$  is roughly equal to the transverse momentum of the Z boson. There is also a small contribution from limited resolution of other objects in the event – primarily jets; this effect of finite resolution means that  $p_{T,Z(\ell\ell)}$  cannot be used directly as a proxy for  $p_T^{miss}$ .

The  $Z(\ell\ell)$  plus  $\gamma$  events are required to have one high- $p_T$  photon and two oppositely charged, same flavor leptons ( $e, \mu$ ) whose invariant mass is consistent with the Z mass,  $80 < m_{\ell\ell} < 100$  GeV. For the purposes of computing event level kinematic variables, the effect of leptons is removed. For  $H_T^\gamma, N_{jets}, N_{b-jets}$ , and  $\Delta\phi$  jets which are matched,  $\Delta R < 0.3$ , to the charged leptons are ignored. For  $p_T^{miss}$  the reconstructed transverse momentum of the Z candidate is added to the  $p_T^{miss}$  vector. This cleaning procedure allow for finite jet resolution to smear the  $p_{T,Z(\ell\ell)}$  and provide a consistent way of modeling the kinematics of neutrino decays. The kinematic selections are then the same as the signal region, but using these “cleaned” variables.

$Z(\ell\ell)$  plus  $\gamma$  events have one primary difference with respect to  $Z(\nu\nu) + \gamma$  events; the former can result from the photon being radiated by charged leptons from the Z boson decay. This effect is small for our phase space,  $p_{T,Z(\ell\ell)} p_T^{miss} > 100$  GeV and  $p_T^\gamma > 100$  GeV. However, these events will typically produce a photon which is roughly colinear with one of the charged leptons in the event. We require that the angle between any electron or muon and the photon be  $\Delta R(e/\mu, \gamma) > 0.2$ . We also require invariant mass of di-lepton pair to be in the range  $80 - 100$  GeV which further suppresses the photon radiation from lepton.

The triggers for the  $\gamma$ -DY control region are the same as the signal region triggers.

Because of the low statistics for  $Z(\ell\ell)+\gamma$  events, the  $p_T^{miss}$  shape will be taken directly from MC. The overall event yield will be measured by scaling the  $Z(\nu\nu) + \gamma$  event yield by a MC transfer factor, TF,  $(N_{\nu\nu+\gamma}^{\text{MC}} / N_{\ell\ell+\gamma}^{\text{MC}})$  that will account for the relative branching ratio,  $\mathcal{B}(\nu\nu)/\mathcal{B}(\ell\ell)$ , and reconstruction efficiencies of the charged leptons. The kinematics of the  $Z(\nu\nu)+\gamma$  and  $Z(\ell\ell)+\gamma$  are very identical and they are shown in figure 3.16 for some of the variables. Because of the nonavailability of the  $Z(\ell\ell)+\gamma$  MC samples with  $p_T^\gamma < 130$  GeV, these transfer factors are derived from high  $p_T^\gamma$  region.

The prediction is given by,

$$N_{\nu\nu+\gamma}^{\text{data}}(\text{SR Bin}) = N_{\ell\ell+\gamma}^{\text{data}} \cdot \beta_{\ell\ell+\gamma} \cdot \left( \frac{N_{\nu\nu+\gamma}^{\text{MC}}}{N_{\ell\ell+\gamma}^{\text{MC}}} \right)^{p_T^\gamma > 190} \cdot \frac{N_{\nu\nu+\gamma}^{\text{MC}}(\text{SR Bin})}{N_{\nu\nu+\gamma}^{\text{MC}}} \quad (3.6)$$

where  $\beta_{\ell\ell+\gamma}$  represents the purity of the  $Z(\ell\ell)+\gamma$  control region and the last term  $N_{\nu\nu+\gamma}^{\text{MC}}(\text{SR Bin})/N_{\nu\nu+\gamma}^{\text{MC}}$  represents a MC template that determines the  $p_T^{miss}, N_{jets}$  shape of the prediction; the template is computed independently for each of the signal regions.  $N_{\ell\ell+\gamma}^{\text{data}}$  represents the observed number of events in data while  $N_{\ell\ell+\gamma}^{\text{MC}}$  represents the expected number of from MC; both are computed inclusively in search bins. As mentioned before, we derive TF in  $p_T^\gamma > 190$  GeV region. There are two reasons why the same TF is applicable even at low photon  $p_T^\gamma$  regions:

- The number of events that lie in the region  $100 < p_T^\gamma < 190 \text{ GeV}$  is a small fraction of the total events (about 15%).
- TF is not dependent on photon  $p_T$  (top left plot in Figure 3.16).

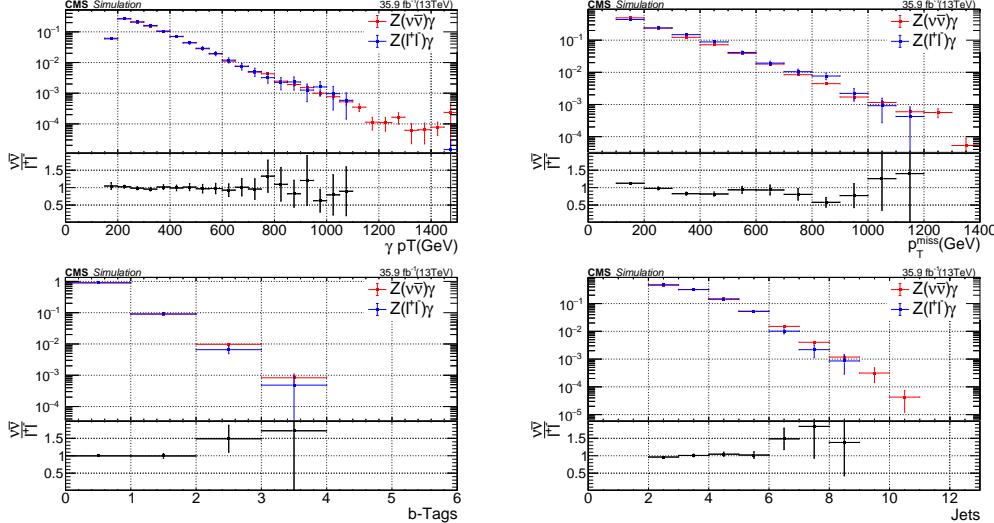


FIGURE 3.16: Comparison of  $Z(l'l')\gamma$  with  $Z(vv)\gamma$  as function of  $p_T^\gamma$  (top left),  $p_T^{\text{miss}}$  (top right),  $N_{\text{b-jets}}$  (bottom left) and  $N_{\text{jets}}$  (bottom right) for  $p_T^\gamma > 190 \text{ GeV}$

The purity of the  $Z(l'l')+\gamma$  control region is determined from a fit to the invariant mass,  $m_{\ell\ell}$ , distribution. This isolates contamination from  $\gamma t\bar{t}$  events, which are expected to have a roughly flat distribution. The purity is checked in three different ways, using MC to predict the number of  $t\bar{t}\gamma$  events, using a fit to the  $m_{\ell\ell}$  distribution, and using the opposite-sign, different-flavor control sample ( $e\mu\gamma$ ). All three predictions give consistent results when calculated within the baseline phase-space, which are reported in Table 3.4. The uncertainties associated with our prediction of the  $Z(l'l')\gamma$  purity is also propagated to the final  $Z\gamma$  prediction.

TABLE 3.4: Purity in the  $Z\gamma$  control region.

$N_{\text{b-jets}}$	Method		
	MC	$m_{\ell\ell}$ fits	$e\mu$ CR
$\geq 0$	$0.978 \pm 0.009$	0.97	$0.97 \pm 0.03$

Because of the very low statistics in low  $\Delta\phi$  region in data, we make use of scale factors that are derived using high  $\Delta\phi$  events to predict events in the low  $\Delta\phi$  region. Scale factor is given by equation 3.7 and it is found to be  $1.11 \pm 0.21$ .

$$SF = \left( \frac{N_{l'l'\gamma}^{\text{data}} \cdot \beta_{l'l'\gamma}}{N_{vv\gamma}^{\text{MC}}} \right)^{p_T^\gamma > 190, \text{high } \Delta\phi} \quad (3.7)$$

Prediction in low  $\Delta\phi$  region is,

$$N_{vv\gamma}^{\text{data}}(\text{SR Bin}) = SF \cdot N_{vv\gamma}^{\text{MC}}(\text{SR Bin}) \quad (3.8)$$

To account for systematic uncertainties associated with the MC modeling of the  $p_T^{miss}$  shape, we apply the full electro-weak corrections obtained from theory calculations [44] as a function of  $p_T^{miss}$ . These uncertainties are listed in table 3.5. they can be as high as 40% in highest  $p_T^{miss}$  bin. These uncertainties are treated as correlated across  $p_T^{miss}$  bins for the limit setting procedure. Uncertainties from b-tag SF are also considered and they are 2% in 0b-tag bins and 6% in  $\geq 1$ b-tag bins. The b-tag SF uncertainty is treated as anti-correlated in 0 and  $\geq 1$  b-tag bins. Final predictions for  $Z(\nu\nu) + \gamma$  are listed in Table A.6 for high  $\Delta\phi$  and in Table A.7 for low  $\Delta\phi$  regions.

TABLE 3.5: Electroweak corrections as a function of  $p_T^{miss}$ .

$p_T^{miss}$ (GeV)	% EW Correction
100-200	8
200-270	18
270-350	20
350-450	25
450-750	35
$\geq 750$	40

### 3.6.4 $\gamma + \text{jets}$ and QCD multijet estimation

Along with the presence of a well identified high  $p_T$  photon in the event, the  $\gamma + \text{jets}$  production contributes to the background in search regions if one of the jets in an event is mis-measured resulting in fake  $p_T^{miss}$  or it contains a jet originating from b-quarks with a semileptonic decay of B mesons. Fluctuations in hadronization of jets can also result in an energetic  $\pi^0$  misidentified as a photon which along with fake  $p_T^{miss}$  originating from other jets in the event can result in a QCD multijet production contribution to the search regions. Fake  $p_T^{miss}$  arising from mis-measurements of jets or semileptonic b-jets is usually aligned with the jet itself. Hence, a most of this background is rejected by angular cuts summarized in section 3.4 i.e.  $\Delta\phi(\vec{p}_T^{miss}, \vec{p}_T^{\text{jet}}) > 0.3$  for first two leading jets. The main contribution to background is due to the  $\gamma + \text{jets}$  processes and QCD multijet contribution is very small. In the method described in this section, the total background due to fake  $p_T^{miss}$  is estimated without further separating the QCD multijet and  $\gamma + \text{jets}$  processes. The control sample to estimate the  $\gamma + \text{jets}$  and QCD multijet background is derived inverting the  $\Delta\phi(p_T^{miss}, \text{jet1})$  and  $\Delta\phi(p_T^{miss}, \text{jet2})$  criteria ( $\min(\Delta\phi_1, \Delta\phi_2) < 0.3$ ) while keeping all the search selections intact except that  $100 < p_T^{miss} < 200$  GeV sideband is also used. Since fake  $p_T^{miss}$  and mis-measured jets are mostly aligned in transverse direction, the inverted  $\Delta\phi$  region provides a high statistics region rich in multijet background. Figure 3.17 shows the  $\min(\Delta\phi_1, \Delta\phi_2)$  distribution in MC with  $p_T^{miss} > 200$  GeV.

The method briefly summarized as follows:

- Using the low  $p_T^{miss}$  sideband, ratio  $R(N_{\text{jets}}, N_{\text{b-jets}}) = \text{high-}\Delta\phi/\text{low-}\Delta\phi$  is determined. This ratio is determined from data sideband. Contribution from electroweak background (lost lepton +  $\tau_{\text{had}}$ , e faking photon and invisible Z) is subtracted from event yields before measuring the ratio. The  $R(N_{\text{jets}}, N_{\text{b-jets}})$  means that the ratio is binned in the bins of jet multiplicity and b-jet multiplicity.
- The number of events obtained in low- $\Delta\phi$  region in  $p_T^{miss} > 200$  GeV region, after subtracting the electroweak backgrounds, is multiplied with ratio to obtained number of events in the search regions.

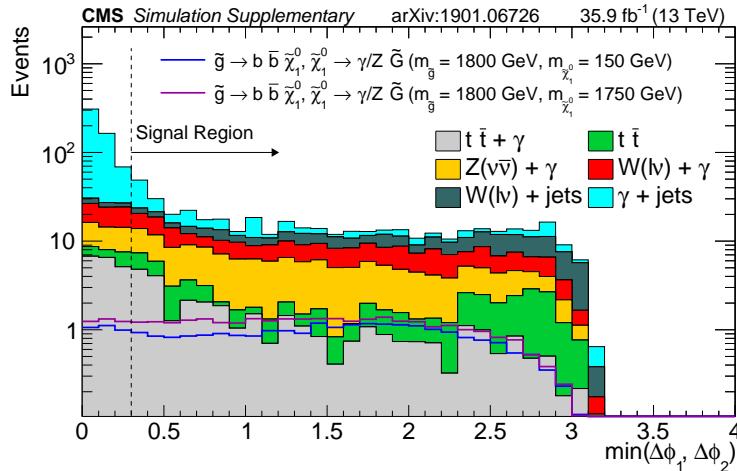


FIGURE 3.17: Distribution of  $\min(\Delta\phi_1, \Delta\phi_2)$  in MC after applying all signal region selections except angular cuts. The region  $\min(\Delta\phi_1, \Delta\phi_2) < 0.3$  serves as a  $\gamma + \text{jets}$  enriched sample and  $\min(\Delta\phi_1, \Delta\phi_2) > 0.3$  is the signal region. In this plot,  $p_T^{\text{miss}} > 200$  GeV selection is applied. The histograms shown as lines represent two signal models which indicate that there is negligible signal contamination.

- Since  $\Delta\phi$  and  $p_T^{\text{miss}}$  are not completely independent, the method is corrected for dependence on  $p_T^{\text{miss}}$  using the MC, using an additional correction factor,  $\kappa(N_{\text{jets}}, N_{\text{b-jets}})$ , where the quantities in bracket shows the binning variables. So in this sense, it is a ABCD method modified to correct for  $\Delta\phi$  and  $p_T^{\text{miss}}$  dependencies using the  $\kappa$  factor.
- The factor  $\kappa$  is validated in data using zero photon events using jet with highest neutral electromagnetic fraction as a proxy to photon.

The boundaries for high- $\Delta\phi$  and low- $\Delta\phi$ , and high  $p_T^{\text{miss}}$  and low  $p_T^{\text{miss}}$  used as ABCD regions are shown in Figure 3.18 (left).

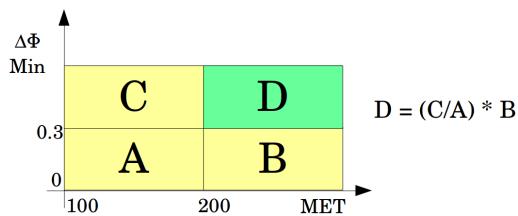


FIGURE 3.18: Definitions of various regions used in ABCD methods. Region D is signal region where background is to be estimated or validation region for zero photon sample.

The ratio of high- $\Delta\phi$  to low- $\Delta\phi$  has a strong dependence as a function of  $p_T^{\text{miss}}$ . This dependence is accounted for by an additional correction factor using ratio, R, in low  $p_T^{\text{miss}}$  and high  $p_T^{\text{miss}}$  regions in MC. That is a factor  $\kappa$  defined as  $\kappa(N_{\text{jets}}, N_{\text{b-jets}}) = R(N_{\text{jets}}, N_{\text{b-jets}}) (\text{high } p_T^{\text{miss}})/R(N_{\text{jets}}, N_{\text{b-jets}})(\text{low } p_T^{\text{miss}})$ . Please note that the ratio R and double ratio,  $\kappa$  are binned in the bins of  $N_{\text{jets}}$  and  $N_{\text{b-jets}}$ .

Using these values of R and  $\kappa$ , the performance of the method is validated using MC simulation. If the numbers R and  $\kappa$  were derived using exact search region definitions, the closure would be one. Hence this validation mainly tests performance

parameterization of these factors. The results of this closure tests are shown in Figure 3.19 as a function of search bins used for this analysis. The number of  $\gamma + \text{jets}$  and QCD multijet background events estimated using the methods (blue) closely reproduces those expected in search regions (cyan). The first bin in each  $N_{\text{jets}}-N_{\text{b-jets}}$  block corresponds to  $100 < p_T^{\text{miss}} < 200$  GeV control regions, where expected and predicted backgrounds exactly match by definition.

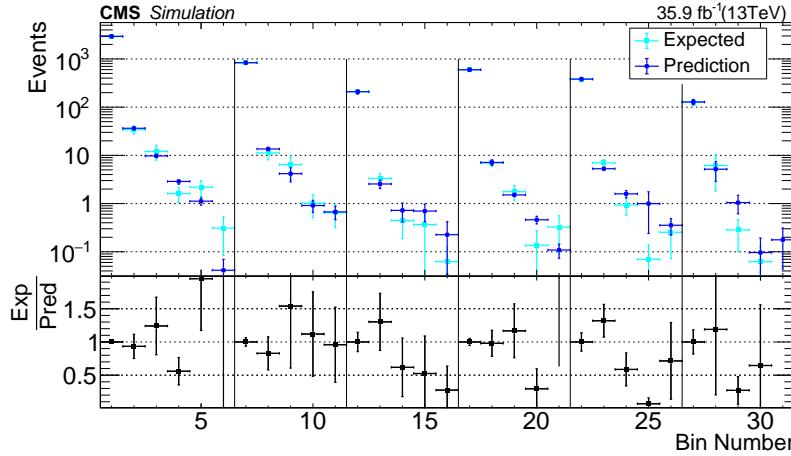


FIGURE 3.19: Comparison of  $\gamma + \text{jets}$  and QCD multijet background estimated using low- $\Delta\phi$  control regions taken from MC (blue) to that expected in search regions (cyan).

TABLE 3.6: Double ratio( $\kappa^{\text{MC}}$ ) computed in the one  $\gamma$  regions and ratio of high- $\Delta\phi$  to low- $\Delta\phi$  ( $R^{\text{data}}$ ) obtained from data using  $100 < p_T^{\text{miss}} < 200$  GeV region with electroweak contribution subtracted.

$N_{\text{b-jets}}$	$N_{\text{jets}}$	$K^{\text{MC}}$	$R^{\text{data}}$
0	2 - 4	$0.29 \pm 0.04$	$0.80 \pm 0.02$
	5 - 6	$0.47 \pm 0.11$	$0.94 \pm 0.05$
	$\geq 7$	$0.40 \pm 0.11$	$1.04 \pm 0.17$
$\geq 1$	2 - 4	$0.24 \pm 0.05$	$0.65 \pm 0.03$
	5 - 6	$0.34 \pm 0.07$	$0.91 \pm 0.09$
	$\geq 7$	$0.47 \pm 0.34$	$1.29 \pm 0.26$

Since the double ratio  $\kappa(N_{\text{jets}}, N_{\text{b-jets}})$  used for correcting R in low and high  $p_T^{\text{miss}}$  regions is obtained from MC, and makes an important component of this method, it is validated using data. To define the data control sample, photon selection is inverted and events with a well identified photon are vetoed. In these zero photon events, the jet with the highest electromagnetic fraction is used as a proxy to the photon, and all baseline selection criteria are applied to the event. The jet used as proxy-photon is removed from the list of jets to avoid any double counting of the objects. With this zero photon event sample, double ratio,  $\kappa(N_{\text{jets}}, N_{\text{b-jets}})$ , is calculated exactly as in MC events.

The zero photon control sample is obtained from the data collected using a combination of inclusive  $H_T$  and single jet triggers as explained in 3.3. The electroweak contribution to this sample is subtracted using event yields from the MC simulation samples. The MC event yields are corrected for the trigger efficiencies. Since these triggers achieve efficiency plateau only for  $H_T > 1$  TeV, the  $H_T > 1$  TeV region

is used for validating the double ratio. The closure of this validation method in MC is shown in Figure 3.20.

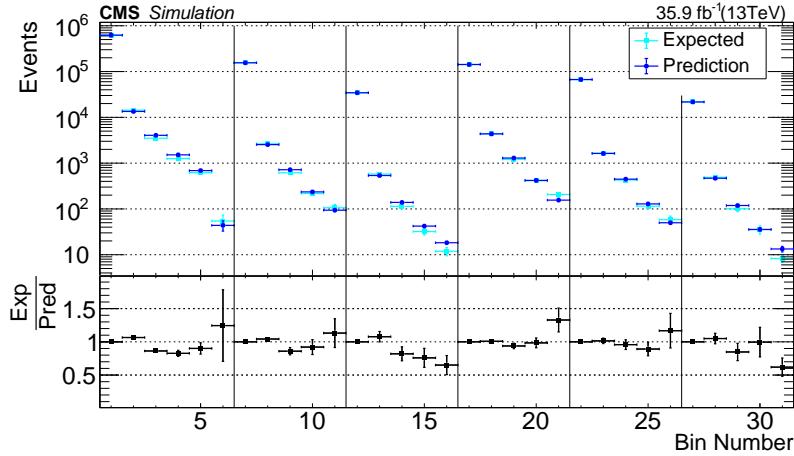


FIGURE 3.20: In zero photon validation region in MC, comparison of  $\gamma + \text{jets}$  and QCD multijet background estimated using low- $\Delta\phi$  control regions taken from MC (blue) to that expected in search regions (cyan).

The values of double ratio,  $\kappa(N_{\text{jets}}, N_{\text{b-jets}})$  obtained from zero photon validation region in data and MC are compared in Figure 3.21.

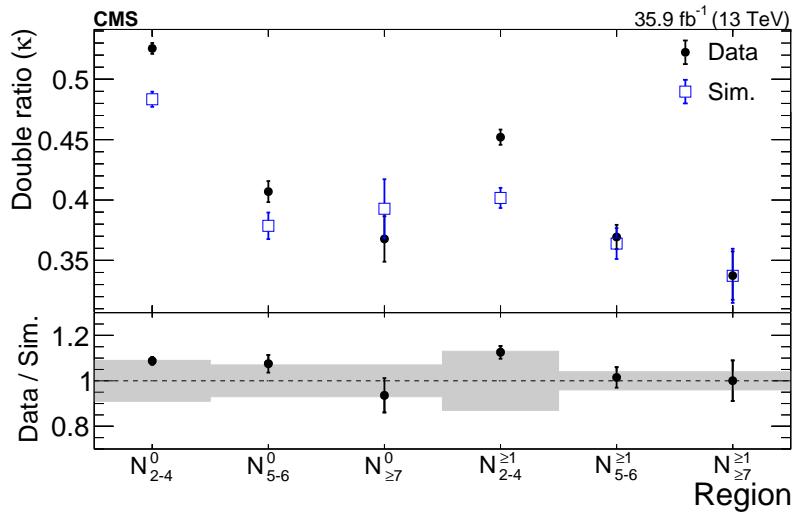


FIGURE 3.21: The double ratio  $\kappa$  in each  $N_{\text{jets}}-N_{\text{b-jets}}$  region for zero-photon events. The filled black circles are the observed  $\kappa$  values after subtracting the electroweak contamination based on simulation. The open blue squares are the  $\kappa$  values computed directly from simulation. The ratio is shown in the bottom panel, where the shaded region corresponds to the systematic uncertainty in the  $\gamma + \text{jets}$  prediction. In the label  $N_j^b$ ,  $j$  refers to the number of jets and  $b$  refers to the number of b-tagged jets.

The uncertainties on predicted  $\gamma + \text{jets}$  and QCD multijet background is dominated by statistical size of event sample in low- $\Delta\phi$  control region with  $p_T^{\text{miss}} > 200$  GeV i.e. the region B, and ranges from 6-100%. These components of systematic uncertainties are taken uncorrelated across all bins. An additional contribution

to this control region is due to uncertainties on predicted electroweak backgrounds which is subtracted from region B. This can range between 10-100%. Since  $R(N_{\text{jets}}, N_{\text{b-jets}}) = \text{high-}\Delta\phi/\text{low-}\Delta\phi$  is defined in the bins of  $N_{\text{jets}}$  and  $N_{\text{b-jets}}$  from low  $p_T^{\text{miss}}$  sideband region in single photon side band, statistical uncertainties on R are taken to be correlated for all bins with same  $N_{\text{jets}}$  and  $N_{\text{b-jets}}$ . The uncertainties on double ratio have two components: (a) difference between  $K_{0\gamma}$  in data and MC, and (b) statistical uncertainty on  $K_{0\gamma}^{\text{MC}}$ . These two contributions are added in quadrature to assign the systematic uncertainties which are also taken fully correlated for all bins with same  $N_{\text{jets}}$  and  $N_{\text{b-jets}}$ . The final predictions of  $\gamma + \text{jets}$  and QCD multijet processes is shown in A.8.

### 3.7 Systematic uncertainty for signal models

We consider a variety of experimental and theoretical systematic uncertainties on the signal rates. They are:

- **Luminosity:** A flat uncertainty of 2.5% is used.
- **Isolated track veto:** A flat uncertainty of 2% is assigned to the T5ttttZG, T5bbbbZG, T5qqqqHG and T6ttZG signal samples to account for any data/MC differences.
- **PF Jet ID:** Some of the PF Jet ID criteria, defined in Section 3.4, are not well-modeled in the fast simulation. Therefore, the jet ID criteria are not applied to the fast simulation samples. The efficiency of the event cleaning cut on the signal samples is expected to be  $>99\%$  from the full simulation, so corrections are not needed. The uncertainty in this correction is taken to be a flat 1% uncertainty.
- **b-tag efficiency:** The b-tagging and mistagging scale factors are functions of the jet  $p_T$  and  $\eta$ . The scale factors are varied by their uncertainties and these variations are propagated as migrations between the different signal bins, with no effect on the overall signal efficiency. The b-tagging, charm-mistagging, and light flavor-mistagging scale factors are varied independently.
- **b-tag FastSim corrections:** The b-tagging and mistagging performance in the fast simulation must be corrected to match the full simulation. Separate correction factors are derived for b-jets, c-jets, and light flavor jets, as functions of the jet  $p_T$  and  $\eta$ . As with the scale factors above, the correction factors for each type of jet are varied independently by their uncertainties and these variations are propagated as migrations between the different signal bins.
- **Jet Energy Corrections:** The jet energy corrections (JECs) are varied using the  $p_T$ - and  $\eta$ -dependent jet energy scale uncertainties, with a separate set of corrections for the fast simulation samples. These variations are propagated into the various jet-dependent variables, including:  $N_{\text{jets}}$ ,  $N_{\text{b-jets}}$ ,  $p_T^{\text{miss}}$ ,  $\Delta\phi(\text{jet}, p_T^{\text{miss}})$  and  $H_T^\gamma$ . The effect is 5% or less.
- **Jet Energy Resolution:** Simulated jet momenta are smeared to match that in data, and the smearing factors are varied according to uncertainties on the jet energy resolution measurements. These variations are propagated into the various jet-dependent variables, including  $N_{\text{jets}}$ ,  $N_{\text{b-jets}}$ ,  $p_T^{\text{miss}}$ ,  $\Delta\phi(\text{jet}, p_T^{\text{miss}})$  and  $H_T^\gamma$ . The overall effect is 2%.

- $p_T^{\text{miss}}$  **Uncertainty**: Mainly special treatment of  $p_T^{\text{miss}}$  modeling in FastSim, and is expected to be important only in compressed regions. The signal yields are obtained using GenMET, summing all visible generator-level particles, and PFMET and average of the GenMET and PFMET is taken as central value. A flat uncertainty equal to one-half the difference between the GenMET and PFMET, fully correlated among  $p_T^{\text{miss}}$  bins is used for limit settings. This procedure results in less than a few percent uncertainty in non-compressed and < 10% uncertainties in highly compressed regions.
- **ISR**: An ISR correction is derived from  $t\bar{t}$  events, with a selection requiring two leptons (electrons or muons) and two b-tagged jets, implying that any other jets in the event arise from ISR. The correction factors are 1.000, 0.920, 0.821, 0.715, 0.662, 0.561, 0.511 for  $N_{\text{jets}}\text{-ISR} = 0, 1, 2, 3, 4, 5, 6+$ . For signal samples, number of ISR jets is the number of jets which are not originating from decay of gluino or vector boson or Higgs or top. The corrections are applied to the simulated signal samples with an additional normalization factor, typically  $\sim 1.15$  (depending on the signal model), to ensure the overall cross section of the sample remains constant. The systematic uncertainty in these corrections is chosen to be half of the deviation from unity for each correction factor. The effect on the yield ranges from 4 – 30%, with the largest effect for compressed signal models. This is the most dominant uncertainty for the signal models which populate high  $N_{\text{jets}}$  bins.

Following theoretical uncertainties are evaluated for each signal model point or explained otherwise.

- **Scales**: The uncertainty is calculated using the envelope of the weights from varying the renormalization and factorization scales,  $\mu_R$  and  $\mu_F$ , by a factor of 2 [24, 27]. The effect on the yield is 2%.
- **PDFs**: The LHC4PDF prescription for the uncertainty on the total cross section is included as  $\pm 1$  sigma bands in the results plots. No additional uncertainty is considered for the uncertainty in the acceptance due to PDFs.

Table 3.7 summarizes the various uncertainties, the range of their impact on signal yields and modeling of the correlations.

TABLE 3.7: Systematic uncertainties for signal samples

Type of uncertainty	Magnitude	Correlation modeling
MC stats	1 – 70%	
JEC, JER	5%, 2%	Uncorr. across all bins
Iso-track veto	2%	
PF jet ID	1%	
b-tag SF	5% in high stat bins	Corr. across bins of same b-tag and anti-corr. across bins of different b-tag
ISR re-weighting $\mu_R$ and $\mu_F$ scales	4 – 30% 2%	Corr. across same $N_{\text{jets}}$ bins
GenMET vs PFMET	< 10%	Corr. across $p_T^{\text{miss}}$ bins
Trigger efficiency	2%	Corr. across all bins

Additionally, the effect of potential signal contamination in the  $e + \gamma$  and  $\mu + \gamma$  control regions is considered. If some events observed in the control regions are in

fact due to a SUSY signal, the corresponding background prediction in the signal region will be too high. This effect is dealt with in the analysis by a corresponding reduction in the effective efficiency for the signal point. For the T5ttttZG signal, the size of this reduction of effective signal efficiency is about 10% coming from  $\mu + \gamma$  control region and 5% coming from  $e + \gamma$  control region. For  $e + \gamma$ , effect of signal contamination is small as compared to  $\mu + \gamma$ , since transfer factor for lost electron is smaller than lost  $\mu + \tau_{\text{had}}$  transfer factor. In the case of T6ttZG model, signal contamination is about 5% for  $\mu + \gamma$  and 2.5% for  $e + \gamma$ . For other SMS that are considered in the analysis signal contamination is negligible.

## Chapter 4

# Results and summary

### 4.1 Results & statistical interpretations

The background predictions described in Section 3.6 are summed together for the final statistical SMS interpretations. Each analysis bin is represented as a individual counting experiment with one signal process and five background processes. Each process is assumed to be a Poisson distribution with some number of prior distribution representing the uncertainty on the corresponding process's rate parameter. The exact details of how these uncertainties are modeled depends on the details of each of the background estimation methods that were used.

In general, each background estimation is a scaling of a Poisson process from a control region,  $N_{\text{bkg}}^{\text{pred}} = N_{\text{CR}}^{\text{pred}} \beta$ , where  $\beta$  is the transfer factor,  $N_{\text{CR}}^{\text{pred}}$  is the expected yield in the control region of the corresponding process, and  $N_{\text{bkg}}^{\text{pred}}$  is the prediction provided as the expected mean of the corresponding process's Poisson distribution for a given signal region. In general,  $N_{\text{CR}}^{\text{pred}}$  is constrained by the observed yields in the control region,  $N_{\text{CR}}^{\text{obs}}$  and there are five such constraints from each of the 5 control regions discussed in Section 3.6. For a simplified counting experiment with only one background process and one signal process, the likelihood for this type of background prediction would be,

$$\mathcal{L}(N_{\text{sig}}^{\text{pred}}, N_{\text{bkg}}^{\text{pred}} | N_{\text{SR}}^{\text{obs}}, N_{\text{CR}}^{\text{obs}}) \propto \left( N_{\text{bkg}}^{\text{pred}} + N_{\text{sig}}^{\text{pred}} \right)^{N_{\text{SR}}^{\text{obs}}} e^{-N_{\text{bkg}}^{\text{pred}}} \cdot \left( \frac{N_{\text{bkg}}^{\text{pred}}}{\beta} \right)^{N_{\text{CR}}^{\text{obs}}} e^{-N_{\text{bkg}}^{\text{pred}}/\beta} \quad (4.1)$$

This likelihood, is equivalent to a Poisson likelihood with a gamma distribution as the prior and can be used to constrain both the predicted background the signal strength in the signal region with the proper modeling of the  $N_{\text{bkg}}^{\text{pred}}$  uncertainty due to limited statistics of the control region observed yield.

As such, we modeling the uncertainty on the prediction due to limited control region statistics as a gamma distribution. This procedure is followed, especially when the observed control region statistics are too small to justifying a Gaussian approximation. While the observed control region events vary considerably, for the signal regions in extreme corners of our phase space the use of a Gamma prior distribution is critical for modeling uncertainties. For uniformity, we generally use it everywhere. The exception to this, as mentioned below, is the  $e+jets$  control region, which is used for estimating the fake-photon background, where we use a Gaussian prior to model the effect of limited control region statistics. Note, the number of observed events in these control regions is always larger than 17 events, for which the  $1\sigma$  intervals of a Poisson distribution and Gaussian distribution still agree to within 10%.

The uncertainty on the transfer factors used to scale the observed event yields in the control regions are modeled with log-normal prior distributions. These priors account for uncertainties typically arising from limited MC statistics, systematic effects due to composition or limited knowledge of object reconstruction efficiencies, or PDF/scale uncertainties. Uncertainties from limited MC statistics are typically uncorrelated across each of the search regions. Other source of systematic uncertainty typically correlate expected yields from various signal regions. Details of the modeling of systematic uncertainties for each of the background estimations is given below.

For the lost-e and lost- $\mu + \tau_{\text{had}}$  predictions, the uncertainties on the prediction are modeled with a gamma distribution, with the shape parameter set to the observed number of events in the corresponding  $\ell - \gamma$  control region and the scale parameter is set to be the average transfer factor, given in Table 3.2. There are also log-normal prior distributions used to model the uncertainties of the transfer factors, which account for:

- limited MC statistics; the corresponding nuisance parameters are fully uncorrelated.
- PDF and scale uncertainties; the corresponding nuisance parameters are fully correlated across signal regions and between lost-e and lost- $\mu + \tau_{\text{had}}$ .
- jet energy scale uncertainties; the corresponding nuisance parameters are fully correlated across all signal regions.
- lepton scale factor uncertainties; the corresponding nuisance parameters are fully correlated across  $p_T^{\text{miss}}$  bins and fully uncorrelated across  $N_{\text{b-jets}}$ ,  $N_{\text{jets}}$  regions;
- uncertainties related to soft/collinear photon modeling; the corresponding nuisance parameters are fully correlated for lost-e and negligible for  $\mu + \tau_{\text{had}}$  prediction.

For the fake-rate estimation of  $W + \gamma$  and  $t\bar{t} + \gamma$  events, the systematic uncertainties from the control region statistics are typically small. All of the uncertainties are modeled with log-normal prior distributions whose uncertainties correspond to:

- the statistical uncertainty from the  $e \rightarrow \gamma$  control regions; the corresponding nuisance parameters are fully uncorrelated across all signal regions.
- fake rate scale factors; the corresponding nuisance parameters are fully correlated across all signal regions.
- uncertainties due to pileup and initial state radiation (ISR) modeling of simulations; the corresponding nuisance parameters are fully correlated across all signal regions.

For the  $Z\gamma$  predictions, the statistical uncertainties from control region are modeled with a gamma prior distribution whose shape parameter is the observed events in the  $Z(\ell\ell)\gamma$  control region and the scale parameter is the purity-corrected MC transfer factor, which in the nomenclature of Equation 3.6 is  $\beta \cdot N_{\nu\nu+\gamma}^{\text{MC}} / N_{\ell\ell+\gamma}^{\text{MC}}$ . Additional log-normal prior distributions are included to account for

- the effect of limited MC statistics on  $N_{X+\gamma}^{\text{MC}}$ ; the corresponding nuisance parameters are fully uncorrelated across all signal regions;

- uncertainties on the purity; the corresponding nuisance parameter fully correlates the effect across all signal regions.
- uncertainties due to b-tag scale factors; the corresponding nuisance parameters are fully correlated across  $p_T^{\text{miss}}$  and  $N_{\text{jets}}$  bins, anti-correlated across  $N_{\text{b-jets}}$  regions.
- uncertainty due to missing higher order corrections; the corresponding nuisance parameters are fully correlated across  $p_T^{\text{miss}}$  bins, uncorrelated across  $N_{\text{b-jets}}, N_{\text{jets}}$  regions.

For the multijet predictions, the systematic uncertainties due to limited control region statistics are modeled with a Gamma prior distribution whose shape parameter is the observed number of events in the low- $\Delta\phi$  control regions and the scale parameter is the product of the high-to-low ratio,  $R_{\text{h/l}}$ , the double ratio correction factors,  $\kappa$ , and the purity,  $\beta$ , obtained from data-driven predictions of the electroweak backgrounds in the low- $\Delta\phi$  control regions. Systematic uncertainties on the knowledge of this combined transfer factor are modeled with log-normal prior distributions and account for

- uncertainties on  $R_{\text{h/l}}$  due to limited number of events in data sideband; the corresponding nuisance parameters are fully correlated across  $p_T^{\text{miss}}$  bins and fully uncorrelated across  $N_{\text{b-jets}}, N_{\text{jets}}$  regions.
- uncertainties on  $\kappa$  based on validations with zero photon events; the corresponding nuisance parameters are fully correlated across  $p_T^{\text{miss}}$  bins and fully uncorrelated across  $N_{\text{b-jets}}$  and  $N_{\text{jets}}$  regions.
- uncertainties on  $\beta$  due to all uncertainties as described above; the corresponding nuisance parameters are fully uncorrelated across all search regions.

All of the uncertainties associated with the signal yields, described in detail in Section 3.7, are modeled with log-normal prior distributions. Nuisance parameters associated with ISR modeling uncertainties are correlated across the various  $N_{\text{jets}}$  bins, but uncorrelated across  $N_{\text{b-jets}}$  and  $p_T^{\text{miss}}$  regions. Nuisance parameters associated with  $p_T^{\text{miss}}$  modeling in simulation are correlated across the various  $p_T^{\text{miss}}$  bins and uncorrelated across  $N_{\text{b-jets}}$  and  $N_{\text{jets}}$  regions. Nuisance parameters associated with b-tagging scale factors are fully correlated across all  $p_T^{\text{miss}}$  and  $N_{\text{jets}}$ , but anti-correlated across the  $N_{\text{b-jets}}$  regions. The nuisance parameters associated with the statistical uncertainties of our simulated samples are fully uncorrelated across all signal regions. A single nuisance parameter is associated with the luminosity uncertainty and fully correlates the effect across all signal regions.

Expected and observed limits are computed after adjusting the central values and uncertainties associated with all nuisance parameters to data by minimizing them with respect to the observed yields while fixing the signal strength to zero.

The predicted background and observed yields are shown in Table 4.1 and Fig. 4.1. The largest deviation is found in bin 2 ( $2 \leq N_{\text{jets}} \leq 4$ ,  $N_{\text{b-jets}} = 0$ , and  $270 < p_T^{\text{miss}} < 350$  GeV), where the background is predicted to be 91 events with 51 events observed. The local significance of this single bin was computed to be around 2 standard deviations below the SM expectation. This calculation does not account for the look-elsewhere effect associated with the use of 25 exclusive signal regions, which is expected to reduce this significance. In general, a large deviation in a single bin is

inconsistent with the expected distributions of events from the signal models considered here. The observations in all other bins are consistent with the SM expectations within one standard deviation.

TABLE 4.1: Predicted and observed event yields for each of the 25 exclusive signal regions.

$N_{\text{jets}}$	$N_{\text{b-jets}}$	$p_T^{\text{miss}}$ [GeV]	Lost e	Lost $\mu + \tau_{\text{had}}$	Misid. $\gamma$	$Z(v\bar{v})\gamma$	$\gamma + \text{jets}$	Total	Data
2-4	0	200-270	$10.5 \pm 2.6$	$31.2 \pm 6.0$	$22.3 \pm 5.4$	$33.6 \pm 8.3$	$60 \pm 11$	$157 \pm 16$	151
2-4	0	270-350	$5.8 \pm 1.8$	$29.6 \pm 5.9$	$11.9 \pm 2.9$	$22.9 \pm 6.0$	$20.5 \pm 4.3$	$91 \pm 10$	51
2-4	0	350-450	$1.68 \pm 0.88$	$13.9 \pm 3.9$	$6.6 \pm 1.6$	$17.0 \pm 5.2$	$4.1 \pm 1.4$	$43.3 \pm 6.8$	50
2-4	0	450-750	$1.98 \pm 0.94$	$8.1 \pm 3.1$	$6.7 \pm 1.5$	$18.1 \pm 7.1$	$2.5 \pm 1.3$	$37.4 \pm 8.0$	33
2-4	0	>750	$0.00^{+0.69}_{-0.00}$	$1.2 \pm 1.2$	$0.79 \pm 0.19$	$2.8 \pm 1.2$	$0.41^{+0.42}_{-0.41}$	$5.2 \pm 1.9$	6
5-6	0	200-270	$1.28 \pm 0.61$	$5.1 \pm 1.9$	$3.53 \pm 0.75$	$3.09 \pm 0.78$	$15.8 \pm 4.8$	$28.8 \pm 5.3$	26
5-6	0	270-350	$2.06 \pm 0.80$	$3.2 \pm 1.5$	$2.39 \pm 0.56$	$1.98 \pm 0.54$	$3.7 \pm 1.8$	$13.3 \pm 2.6$	11
5-6	0	350-450	$0.77 \pm 0.46$	$0.64^{+0.65}_{-0.64}$	$1.26 \pm 0.30$	$1.49 \pm 0.47$	$1.23 \pm 0.97$	$5.4 \pm 1.4$	8
5-6	0	>450	$0.26 \pm 0.26$	$1.9 \pm 1.1$	$1.00 \pm 0.24$	$1.65 \pm 0.65$	$0.07^{+0.52}_{-0.07}$	$4.9 \pm 1.4$	7
$\geq 7$	0	200-270	$0.00^{+0.61}_{-0.00}$	$0.0^{+1.3}_{-0.0}$	$0.72 \pm 0.16$	$0.37 \pm 0.11$	$1.8 \pm 1.2$	$2.9 \pm 1.9$	3
$\geq 7$	0	270-350	$0.34^{+0.35}_{-0.34}$	$1.5 \pm 1.0$	$0.38 \pm 0.10$	$0.24 \pm 0.08$	$1.22 \pm 0.94$	$3.6 \pm 1.5$	3
$\geq 7$	0	350-450	$0.34^{+0.35}_{-0.34}$	$0.73 \pm 0.73$	$0.17 \pm 0.05$	$0.16 \pm 0.07$	$0.07^{+0.50}_{-0.07}$	$1.46 \pm 0.96$	0
$\geq 7$	0	>450	$0.00^{+0.61}_{-0.00}$	$0.0^{+1.3}_{-0.0}$	$0.20 \pm 0.06$	$0.17 \pm 0.08$	$0.00^{+0.75}_{-0.00}$	$0.37^{+1.60}_{-0.37}$	0
2-4	$\geq 1$	200-270	$3.4 \pm 1.5$	$14.5 \pm 4.2$	$7.1 \pm 1.7$	$3.55 \pm 0.89$	$11.3 \pm 3.3$	$39.8 \pm 5.9$	50
2-4	$\geq 1$	270-350	$2.9 \pm 1.4$	$5.6 \pm 2.5$	$3.79 \pm 0.92$	$2.45 \pm 0.65$	$5.7 \pm 1.8$	$20.4 \pm 3.6$	20
2-4	$\geq 1$	350-450	$0.0^{+1.0}_{-0.0}$	$1.1 \pm 1.1$	$2.00 \pm 0.45$	$1.81 \pm 0.55$	$0.59 \pm 0.44$	$5.5 \pm 1.7$	4
2-4	$\geq 1$	>450	$2.3 \pm 1.2$	$4.4 \pm 2.3$	$1.62 \pm 0.38$	$2.14 \pm 0.84$	$0.95 \pm 0.54$	$11.5 \pm 2.8$	8
5-6	$\geq 1$	200-270	$3.5 \pm 1.3$	$2.4 \pm 1.4$	$5.5 \pm 1.2$	$0.76 \pm 0.20$	$7.7 \pm 2.4$	$19.9 \pm 3.3$	21
5-6	$\geq 1$	270-350	$1.06 \pm 0.64$	$4.0 \pm 1.8$	$2.98 \pm 0.63$	$0.49 \pm 0.14$	$2.1 \pm 1.0$	$10.6 \pm 2.3$	15
5-6	$\geq 1$	350-450	$0.71 \pm 0.51$	$2.4 \pm 1.4$	$1.38 \pm 0.29$	$0.32 \pm 0.11$	$0.30^{+0.49}_{-0.30}$	$5.1 \pm 1.6$	6
5-6	$\geq 1$	>450	$0.35^{+0.36}_{-0.35}$	$0.0^{+1.4}_{-0.0}$	$0.67 \pm 0.15$	$0.48 \pm 0.20$	$0.00^{+0.56}_{-0.00}$	$1.5^{+1.6}_{-1.5}$	2
$\geq 7$	$\geq 1$	200-270	$0.72 \pm 0.53$	$2.0 \pm 1.2$	$1.68 \pm 0.37$	$0.13 \pm 0.04$	$5.9 \pm 5.0$	$10.5 \pm 5.1$	12
$\geq 7$	$\geq 1$	270-350	$0.00^{+0.65}_{-0.00}$	$1.33 \pm 0.96$	$0.73 \pm 0.16$	$0.10 \pm 0.04$	$0.0^{+1.1}_{-0.0}$	$2.2 \pm 1.6$	1
$\geq 7$	$\geq 1$	350-450	$0.72 \pm 0.53$	$0.0^{+1.2}_{-0.0}$	$0.44 \pm 0.10$	$0.07 \pm 0.03$	$0.0^{+1.0}_{-0.0}$	$1.2^{+1.7}_{-1.2}$	1
$\geq 7$	$\geq 1$	>450	$0.36^{+0.37}_{-0.36}$	$0.0^{+1.2}_{-0.0}$	$0.23 \pm 0.07$	$0.04 \pm 0.02$	$0.0^{+1.1}_{-0.0}$	$0.6^{+1.7}_{-0.6}$	1

The test statistic  $q_\mu = -2 \ln \mathcal{L}_\mu / \mathcal{L}_{\max}$ , where  $\mathcal{L}_{\max}$  is the maximum likelihood determined by allowing all parameters, including the signal strength, to float, and  $\mathcal{L}_\mu$  is the profiled likelihood. Limits are determined using the asymptotic form of the test statistic [41] in conjunction with the  $CL_s$  criterion [67, 85]. Expected upper limits are derived by varying observed yields according to expectations from the background only hypothesis.

Using the statistical procedure described above, 95% confidence level (CL) upper limits are computed on the signal cross section for each simplified model and each mass hypothesis. Exclusion limits are defined by comparing observed upper limits to the predicted NLO+NLL signal cross section. The signal cross sections are also varied according to theoretical uncertainties to give a  $\pm 1$  standard deviation variation on the observed exclusion contour. The 95% CL cross section limits and exclusion contours for the four models considered, T5qqqqHG, T5bbbbZG, T5tttZG, and T6ttZG, are shown in Fig. 4.2.

Generally, the limits degrade at both high and low  $m_{\tilde{\chi}_1^0}$ . For  $m_{\tilde{\chi}_1^0} \approx m_{\tilde{g}} (m_{\tilde{t}})$ , the quarks from the decay of gluinos (top squarks) have low  $p_T$ . Correspondingly, the  $H_T^\gamma$ ,  $N_{\text{jets}}$ , and  $N_{\text{b-jets}}$  distributions tend toward lower values, reducing the signal efficiency and causing signal events to populate regions with higher background yields. For small  $m_{\tilde{\chi}_1^0}$ , the quarks produced in the decay of gluinos or top squarks have high  $p_T$  but lower  $p_T^{\text{miss}}$  on average. For all models except T5qqqqHG, when the NLSP mass drops below the mass of the Z boson, the kinematics of the NLSP decay require the Z boson to be far off-shell. As the Z boson mass is forced to be lower, the LSP will carry a larger fraction of the momentum of the NLSP, producing larger  $p_T^{\text{miss}}$ . This causes a slight increase in the sensitivity when the NLSP mass is near the Z boson mass. While a similar effect would happen for the T5qqqqHG model, the

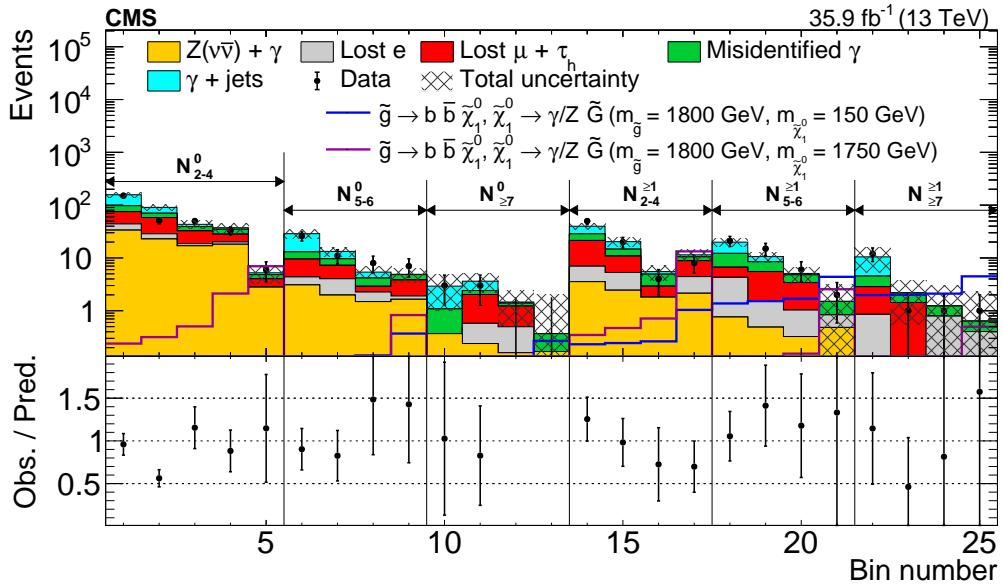


FIGURE 4.1: Observed numbers of events and predicted numbers of events from the various SM backgrounds in the 25 signal regions. The categories, denoted by vertical lines, are labeled as  $N_j^b$ , where  $j$  refers to the number of jets and  $b$  refers to the number of b-tagged jets. The numbered bins within each category are the various  $p_T^{\text{miss}}$  bins, as defined in Table 4.1. The lower panel shows the ratio of the observed events to the predicted SM background events. The error bars in the lower panel are the quadrature sum of the statistical uncertainty in the observed data and the systematic uncertainty in the predicted backgrounds before the adjustments based on a maximum likelihood fit to data assuming no signal strength.

simulation used here does not probe the region of parameter space where the Higgs boson would be forced to have a mass far off-shell.

The features seen in exclusion curves can be understood using acceptance  $\times$  efficiency plots shown in figure B.4 in Appendix B.3.

For moderate  $m_{\tilde{\chi}_1^0}$ , gluino masses as large as 2090, 2120, and 1970 GeV are excluded for the T5qqqqHG, T5bbbbZG, and T5ttttZG models, respectively. Top squark masses as large as 1230 GeV are excluded for the T6ttZG model. For small  $m_{\tilde{\chi}_1^0}$ , gluino masses as large as 1920, 1950, and 1800 GeV are excluded for the T5qqqqHG, T5bbbbZG, and T5ttttZG models, respectively. Top squark masses as large as 1110 GeV are excluded for the T6ttZG model. There is close agreement between the observed and expected limits.

## 4.2 Summary

A search for gluino and top squark pair production is presented, based on a proton-proton collision dataset at a center-of-mass energy of 13 TeV recorded with the CMS detector in 2016. The data correspond to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . Events are required to have at least one isolated photon with transverse momentum  $p_T > 100 \text{ GeV}$ , two jets with  $p_T > 30 \text{ GeV}$  and pseudorapidity  $|\eta| < 2.4$ , and missing transverse momentum  $p_T^{\text{miss}} > 200 \text{ GeV}$ .

The data are categorized into 25 exclusive signal regions based on the number of jets, the number of b-tagged jets, and  $p_T^{\text{miss}}$ . Background yields from the standard

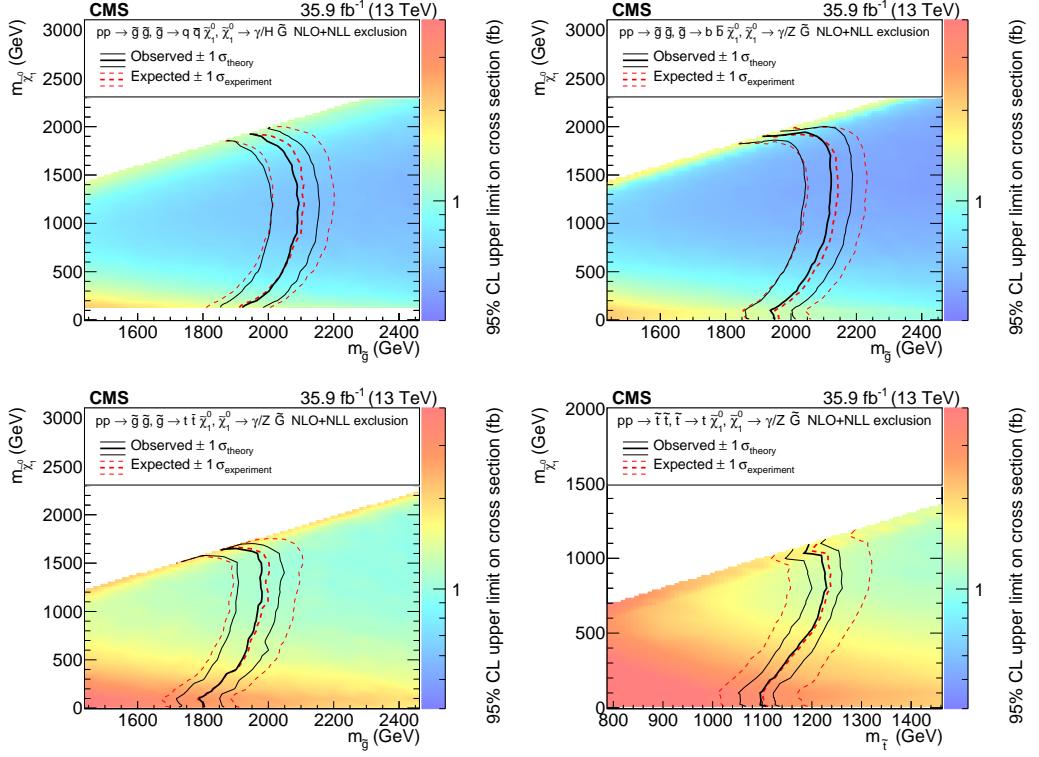


FIGURE 4.2: Observed and expected 95% CL upper limits for gluino or top squark pair production cross sections for the T5qqqqHG (upper left), T5bbbbZG (upper right), T5ttttZG (bottom left), and T6ttZG (bottom right) models. Black lines denote the observed exclusion limit and the uncertainty due to variations of the theoretical prediction of the gluino or top squark pair production cross section. The dashed lines correspond to the region containing 68% of the distribution of the expected exclusion limits under the background-only hypothesis.

model processes are predicted using simulation and data control regions. The observed event yields are found to be consistent with expectations from the standard model processes within the uncertainties.

Results are interpreted in the context of simplified models. Four such models are studied, three of which involve gluino pair production and one of which involves top squark pair production. All models assume a gauge-mediated supersymmetry (SUSY) breaking scenario, in which the lightest SUSY particle is a gravitino ( $\tilde{G}$ ). We consider scenarios in which the gluino decays to a neutralino  $\tilde{\chi}_1^0$  and a pair of light-flavor quarks (T5qqqqHG), bottom quarks (T5bbbbZG), or top quarks (T5ttttZG). In the T5qqqqHG model, the  $\tilde{\chi}_1^0$  decays with equal probability either to a photon and a  $\tilde{G}$  or to a Higgs boson and a  $\tilde{G}$ . In the T5bbbbZG and T5ttttZG models, the  $\tilde{\chi}_1^0$  decays with equal probability either to a photon and a  $\tilde{G}$  or to a Z boson and a  $\tilde{G}$ . In the top squark pair production model (T6ttZG), top squarks decay to a top quark and  $\tilde{\chi}_1^0$ , and the  $\tilde{\chi}_1^0$  decays with equal probability either to a photon and a  $\tilde{G}$  or to a Z boson and a  $\tilde{G}$ .

Using the cross sections for SUSY pair production calculated at next-to-leading order plus next-to-leading logarithmic accuracy, we place 95% confidence level lower limits on the gluino mass as large as 2120 GeV, depending on the model and the  $m_{\tilde{\chi}_1^0}$  value, and limits on the top squark mass as large as 1230 GeV, depending on the  $m_{\tilde{\chi}_1^0}$  value. These results significantly improve upon those from previous searches for SUSY with photons.

## 4.3 Outlook

The main focus of this search was on GMSB scenarios with strong production of gluinos and stops. However, this is a generic search which searches for any beyond SM phenomena resulting in events with photon,  $p_T^{miss}$ ,  $N_{jets}$  and  $N_{b-jets}$  final states. It is also possible to use the results of this search interpret them in terms of any other models. In this section we study the scope of this search by considering EW production of SUSY particles with SMS scenarios.

### 4.3.1 Electroweak models

Two kinds of electroweak (EWK) SMS scenarios, TChiWG and TChiNG, are studied and the expected limits are determined. In TChiWG (Figure 4.3 top), there is production of neutralinos ( $\tilde{\chi}_1^0$ ) and charginos ( $\tilde{\chi}_1^\pm$ ). The decay of  $\tilde{\chi}_1^0$  gives  $\tilde{G} + \gamma$  and the decay of  $\tilde{\chi}_1^\pm$  gives  $W^\pm + \tilde{G}$ . In TChiNG (Figure 4.3 bottom) scenario,  $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_1^0$  are nearly mass degenerate, with  $\tilde{\chi}_1^\pm$  being slightly heavier than  $\tilde{\chi}_1^0$  and its decay gives rise to  $\tilde{\chi}_1^0$  and soft particles which generally do not pass object level requirements. In this case, two production modes,  $\tilde{\chi}_1^\pm\text{-}\tilde{\chi}_1^\pm$  (bottom left) and  $\tilde{\chi}_1^\pm\text{-}\tilde{\chi}_1^0$  (bottom right) are considered. The branching fraction for the decay of  $\tilde{\chi}_1^0$  is 50%  $\tilde{G} + \gamma$ , 25%  $\tilde{G} + H$  and 25%  $\tilde{G} + Z$ . These are the same models as considered by the electrowikino search performed by CMS in Ref. [94]. The cross sections are calculated at NLO+NLL accuracy [54, 53, 18] and are computed in the limit of mass degenerate wino  $\tilde{\chi}_2^0$  and  $\tilde{\chi}_1^\pm$ .

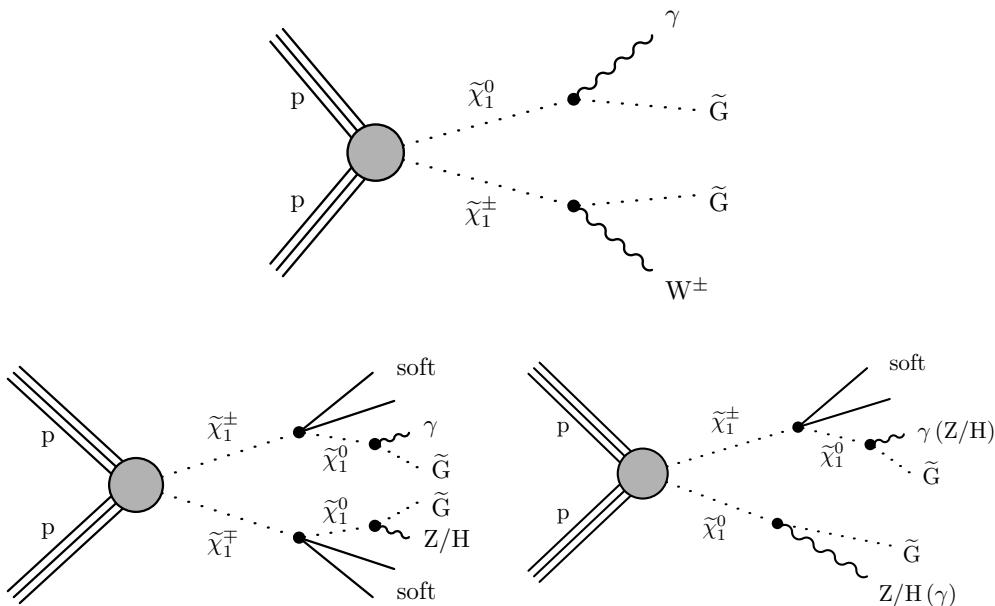


FIGURE 4.3: SMS diagrams for TChiWG (top) and TChiNG (bottom).

### 4.3.2 Sensitivity of the search

The results of the search described in this thesis are interpreted in terms of the above mentioned SMS scenarios and the expected limits on the cross section are compared with Ref. [94]. Figure 4.4 shows the expected upper limits on the cross section for TChiWG scenario (left) and TChiNG (right) scenario using this search.

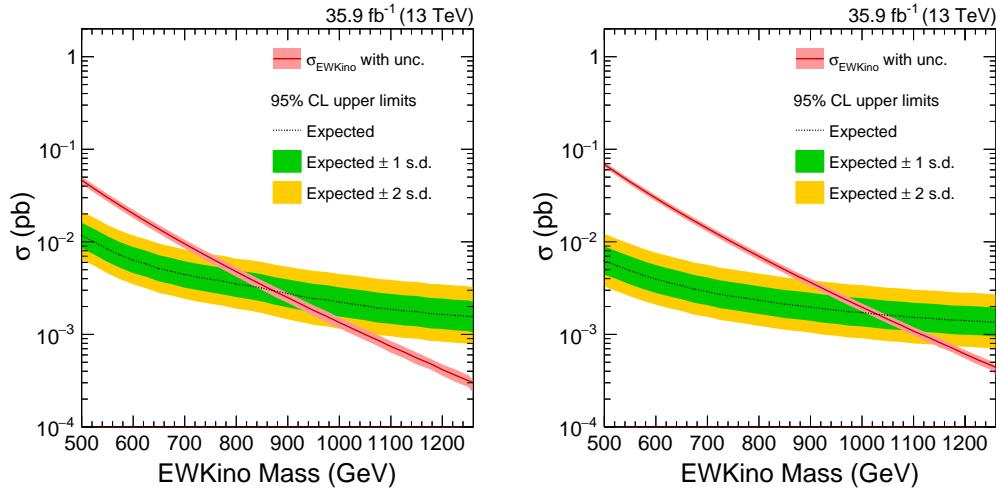


FIGURE 4.4: Expected upper limit on the cross section of TChiWG (left) and TChiNG (right) models.

EWKino masses below 870 GeV for TChiWG case and mass below 1030 GeV for TChiNG case can be excluded. The EW search of CMS [94] has expected EWKino mass exclusion of 920 GeV and 1070 GeV for TChiWG and TChiNG models respectively, which is about 40-50 GeV higher than this search.

It is also worth noting that the search described in this thesis does not consider all sources of systematic uncertainties in EWKino signal models. Only the statistical uncertainty in the MC samples and integrated luminosity uncertainty are considered. If all the sources of systematic uncertainty are considered, the limits in this search are going to degrade to some extent. However, these studies lead to an important conclusion that the search can target a wide range of models although strong production scenarios were the main focus.

One of the possible ways to improve this search and have a good sensitivity for EW models, is to loosen hadronic activity criteria mentioned in section 3.4. At present, the search requires minimum  $H_T^\gamma$  (sum  $p_T$  of jets and photon) of 500 GeV, which is very disadvantageous for EWKino production models. One possible way to reduce the background in low  $H_T^\gamma$  regions, is increasing  $p_T^\gamma$  or  $p_T^{\text{miss}}$  threshold.

During run 2 (year 2016-2018), CMS has collected more than  $135 \text{ fb}^{-1}$  of data and this is going to be an advantage for EW SUSY searches. A combination of several searches with photons is performed by CMS using 2016 data [Add ref for GGM PAS](#). This combination uses photon +  $H_T^\gamma$  search [96], EW search [94], photon + lepton search [95] and diphoton search [Add ref for diphoton PAS](#). Since this search has very good sensitivity for strong SUSY production models, and if one performs further optimization the search for EW SUSY models, this search is going to play a significant role on the path to discovery or to constrain SUSY models.

## Appendix A

# Additional tables

### A.1 Lost lepton + $\tau_{\text{had}}$ predictions

TABLE A.1: Observed events yeilds, average transfer factors and predictions for lost-e and lost- $\mu$  +  $\tau_{\text{had}}$  backgrounds in high  $\Delta\phi$  regions.

Bin	Lost-e			Lost- $\mu$		
	Obs.	Avg. transfer factor	pred.	Obs.	Avg. transfer factor	pred.
1	160	$0.435 \pm 0.073$	$69.6 \pm 12.98$	167	$0.950 \pm 0.066$	$158.7 \pm 16.50$
2	27	$0.385 \pm 0.062$	$10.4 \pm 2.62$	31	$1.008 \pm 0.064$	$31.2 \pm 5.95$
3	14	$0.410 \pm 0.067$	$5.7 \pm 1.79$	29	$1.020 \pm 0.064$	$29.6 \pm 5.80$
4	4	$0.415 \pm 0.068$	$1.7 \pm 0.87$	14	$0.996 \pm 0.064$	$13.9 \pm 3.83$
5	5	$0.392 \pm 0.064$	$2.0 \pm 0.93$	7	$1.154 \pm 0.074$	$8.1 \pm 3.10$
6	0	$0.385 \pm 0.062$	$0.0 \pm 0.69$	1	$1.229 \pm 0.080$	$1.2 \pm 1.23$
7	23	$0.281 \pm 0.052$	$6.5 \pm 1.80$	31	$0.662 \pm 0.057$	$20.5 \pm 4.08$
8	5	$0.254 \pm 0.042$	$1.3 \pm 0.61$	8	$0.643 \pm 0.044$	$5.1 \pm 1.85$
9	8	$0.254 \pm 0.042$	$2.0 \pm 0.79$	5	$0.643 \pm 0.044$	$3.2 \pm 1.45$
10	3	$0.254 \pm 0.042$	$0.8 \pm 0.46$	1	$0.643 \pm 0.044$	$0.6 \pm 0.64$
11	1	$0.254 \pm 0.042$	$0.3 \pm 0.26$	3	$0.643 \pm 0.044$	$1.9 \pm 1.12$
12	7	$0.308 \pm 0.079$	$2.2 \pm 0.99$	10	$0.686 \pm 0.099$	$6.9 \pm 2.39$
13	0	$0.333 \pm 0.088$	$0.0 \pm 0.60$	0	$0.727 \pm 0.100$	$0.0 \pm 1.31$
14	1	$0.333 \pm 0.088$	$0.3 \pm 0.34$	2	$0.727 \pm 0.100$	$1.5 \pm 1.05$
15	1	$0.333 \pm 0.088$	$0.3 \pm 0.34$	1	$0.727 \pm 0.100$	$0.7 \pm 0.73$
16	0	$0.333 \pm 0.088$	$0.0 \pm 0.60$	0	$0.727 \pm 0.100$	$0.0 \pm 1.31$
17	53	$0.526 \pm 0.084$	$27.9 \pm 5.88$	49	$0.962 \pm 0.093$	$47.2 \pm 8.14$
18	6	$0.565 \pm 0.090$	$3.4 \pm 1.49$	13	$1.112 \pm 0.103$	$14.5 \pm 4.23$
19	5	$0.565 \pm 0.090$	$2.8 \pm 1.34$	5	$1.112 \pm 0.103$	$5.6 \pm 2.54$
20	0	$0.565 \pm 0.090$	$0.0 \pm 1.02$	1	$1.112 \pm 0.103$	$1.1 \pm 1.12$
21	4	$0.565 \pm 0.090$	$2.3 \pm 1.19$	4	$1.112 \pm 0.103$	$4.4 \pm 2.26$
22	41	$0.383 \pm 0.066$	$15.7 \pm 3.65$	51	$0.733 \pm 0.072$	$37.4 \pm 6.40$
23	10	$0.349 \pm 0.059$	$3.5 \pm 1.25$	3	$0.794 \pm 0.073$	$2.4 \pm 1.39$
24	3	$0.349 \pm 0.059$	$1.0 \pm 0.63$	5	$0.794 \pm 0.073$	$4.0 \pm 1.81$
25	2	$0.349 \pm 0.059$	$0.7 \pm 0.51$	3	$0.794 \pm 0.073$	$2.4 \pm 1.39$
26	1	$0.349 \pm 0.059$	$0.3 \pm 0.35$	0	$0.794 \pm 0.073$	$0.0 \pm 1.43$
27	16	$0.425 \pm 0.096$	$6.8 \pm 2.29$	16	$0.649 \pm 0.096$	$10.4 \pm 3.02$
28	2	$0.357 \pm 0.077$	$0.7 \pm 0.53$	3	$0.665 \pm 0.098$	$2.0 \pm 1.19$
29	0	$0.357 \pm 0.077$	$0.0 \pm 0.64$	2	$0.665 \pm 0.098$	$1.3 \pm 0.96$
30	2	$0.357 \pm 0.077$	$0.7 \pm 0.53$	0	$0.665 \pm 0.098$	$0.0 \pm 1.20$
31	1	$0.357 \pm 0.077$	$0.4 \pm 0.37$	0	$0.665 \pm 0.098$	$0.0 \pm 1.20$

TABLE A.2: Observed events yeilds, average transfer factors and predictions for lost-e and lost- $\mu$  +  $\tau_{\text{had}}$  backgrounds in the low- $\Delta\Phi$  control region.

Bin	Lost-e			Lost- $\mu$		
	Obs.	Avg. transfer factor	pred.	Obs.	Avg. transfer factor	pred.
1	77	$0.200 \pm 0.045$	$15.4 \pm 3.86$	62	$0.637 \pm 0.092$	$39.5 \pm 7.61$
2	22	$0.162 \pm 0.038$	$3.6 \pm 1.13$	25	$0.566 \pm 0.062$	$14.2 \pm 3.23$
3	11	$0.174 \pm 0.040$	$1.9 \pm 0.73$	7	$0.627 \pm 0.066$	$4.4 \pm 1.72$
4	6	$0.155 \pm 0.036$	$0.9 \pm 0.44$	5	$0.658 \pm 0.068$	$3.3 \pm 1.51$
5	7	$0.172 \pm 0.041$	$1.2 \pm 0.54$	11	$0.586 \pm 0.064$	$6.4 \pm 2.07$
6	0	$0.175 \pm 0.041$	$0.0 \pm 0.31$	0	$0.585 \pm 0.065$	$0.0 \pm 1.05$
7	10	$0.193 \pm 0.044$	$1.9 \pm 0.75$	13	$0.543 \pm 0.110$	$7.1 \pm 2.42$
8	1	$0.165 \pm 0.036$	$0.2 \pm 0.17$	1	$0.542 \pm 0.065$	$0.5 \pm 0.55$
9	2	$0.165 \pm 0.036$	$0.3 \pm 0.24$	3	$0.542 \pm 0.065$	$1.6 \pm 0.96$
10	2	$0.165 \pm 0.036$	$0.3 \pm 0.24$	1	$0.542 \pm 0.065$	$0.5 \pm 0.55$
11	0	$0.165 \pm 0.036$	$0.0 \pm 0.30$	1	$0.542 \pm 0.065$	$0.5 \pm 0.55$
12	1	$0.262 \pm 0.090$	$0.3 \pm 0.28$	4	$0.430 \pm 0.123$	$1.7 \pm 0.99$
13	1	$0.169 \pm 0.055$	$0.2 \pm 0.18$	2	$0.594 \pm 0.198$	$1.2 \pm 0.93$
14	0	$0.169 \pm 0.055$	$0.0 \pm 0.30$	0	$0.594 \pm 0.198$	$0.0 \pm 1.07$
15	1	$0.169 \pm 0.055$	$0.2 \pm 0.18$	1	$0.594 \pm 0.198$	$0.6 \pm 0.63$
16	0	$0.169 \pm 0.055$	$0.0 \pm 0.30$	1	$0.594 \pm 0.198$	$0.6 \pm 0.63$
17	26	$0.388 \pm 0.078$	$10.1 \pm 2.84$	22	$0.856 \pm 0.153$	$18.8 \pm 5.25$
18	4	$0.272 \pm 0.056$	$1.1 \pm 0.59$	4	$0.684 \pm 0.097$	$2.7 \pm 1.42$
19	1	$0.272 \pm 0.056$	$0.3 \pm 0.28$	2	$0.684 \pm 0.097$	$1.4 \pm 0.99$
20	2	$0.272 \pm 0.056$	$0.5 \pm 0.40$	0	$0.684 \pm 0.097$	$0.0 \pm 1.23$
21	0	$0.272 \pm 0.056$	$0.0 \pm 0.49$	2	$0.684 \pm 0.097$	$1.4 \pm 0.99$
22	9	$0.223 \pm 0.061$	$2.0 \pm 0.87$	14	$0.730 \pm 0.130$	$10.2 \pm 3.28$
23	5	$0.250 \pm 0.055$	$1.2 \pm 0.62$	2	$0.536 \pm 0.075$	$1.1 \pm 0.77$
24	1	$0.250 \pm 0.055$	$0.2 \pm 0.26$	1	$0.536 \pm 0.075$	$0.5 \pm 0.54$
25	1	$0.250 \pm 0.055$	$0.2 \pm 0.26$	1	$0.536 \pm 0.075$	$0.5 \pm 0.54$
26	2	$0.250 \pm 0.055$	$0.5 \pm 0.37$	1	$0.536 \pm 0.075$	$0.5 \pm 0.54$
27	3	$0.269 \pm 0.101$	$0.8 \pm 0.56$	4	$0.510 \pm 0.123$	$2.0 \pm 1.13$
28	2	$0.247 \pm 0.072$	$0.5 \pm 0.38$	3	$0.530 \pm 0.103$	$1.6 \pm 0.97$
29	1	$0.247 \pm 0.072$	$0.2 \pm 0.26$	0	$0.530 \pm 0.103$	$0.0 \pm 0.95$
30	1	$0.247 \pm 0.072$	$0.2 \pm 0.26$	0	$0.530 \pm 0.103$	$0.0 \pm 0.95$
31	0	$0.247 \pm 0.072$	$0.0 \pm 0.45$	0	$0.530 \pm 0.103$	$0.0 \pm 0.95$

## A.2 Electron fake photon

## A.3 Invisible Z estimation

## A.4 $\gamma + \text{jets}$ and QCD multijet

TABLE A.3: Fake rate parameterization derived from  $W/t\bar{t}$  simulations.

$p_T$	$Q_{mult}$	fake rate
100 – 120	0 – 1	$0.0261 \pm 0.0029$
	2 – 4	$0.0125 \pm 0.0009$
	4 – 7	$0.0080 \pm 0.0006$
	$\geq 7$	$0.0052 \pm 0.0006$
120 – 140	0 – 1	$0.0208 \pm 0.0028$
	2 – 4	$0.0118 \pm 0.0011$
	4 – 7	$0.0068 \pm 0.0006$
	$\geq 7$	$0.0050 \pm 0.0006$
140 – 160	0 – 1	$0.0263 \pm 0.0036$
	2 – 4	$0.0094 \pm 0.0010$
	4 – 7	$0.0079 \pm 0.0007$
	$\geq 7$	$0.0028 \pm 0.0004$
160 – 180	0 – 1	$0.0243 \pm 0.0041$
	2 – 4	$0.0114 \pm 0.0012$
	4 – 7	$0.0065 \pm 0.0007$
	$\geq 7$	$0.0034 \pm 0.0006$
180 – 200	0 – 1	$0.0194 \pm 0.0038$
	2 – 4	$0.0070 \pm 0.0011$
	4 – 7	$0.0070 \pm 0.0009$
	$\geq 7$	$0.0040 \pm 0.0008$
200 – 230	0 – 1	$0.0177 \pm 0.0030$
	2 – 4	$0.0104 \pm 0.0014$
	4 – 7	$0.0044 \pm 0.0005$
	$\geq 7$	$0.0042 \pm 0.0008$
230 – 260	0 – 1	$0.0235 \pm 0.0052$
	2 – 4	$0.0062 \pm 0.0011$
	4 – 7	$0.0051 \pm 0.0008$
	$\geq 7$	$0.0034 \pm 0.0007$
260 – 300	0 – 1	$0.0293 \pm 0.0079$
	2 – 4	$0.0096 \pm 0.0014$
	4 – 7	$0.0062 \pm 0.0010$
	$\geq 7$	$0.0075 \pm 0.0017$
300 – 380	0 – 1	$0.0224 \pm 0.0052$
	2 – 4	$0.0091 \pm 0.0014$
	4 – 7	$0.0073 \pm 0.0009$
	$\geq 7$	$0.0035 \pm 0.0006$
380 – 500	0 – 1	$0.0232 \pm 0.0048$
	2 – 4	$0.0084 \pm 0.0011$
	4 – 7	$0.0047 \pm 0.0006$
	$\geq 7$	$0.0036 \pm 0.0008$
> 500	0 – 1	$0.0210 \pm 0.0041$
	2 – 4	$0.0102 \pm 0.0015$
	4 – 7	$0.0076 \pm 0.0010$
	$\geq 7$	$0.0047 \pm 0.0014$

TABLE A.4: Observed single electron yields, after accounting for lepton scale factors and relative trigger efficiencies factor, in each of the signal regions, the average scale factor due to MC-based fake rates and data/MC fake rate correction factors from T&P measurements, and the corresponding fake-rate predictions with all the uncertainties (statistical and systematic) taken into consideration.

Bin	Obs. e events	Avg. transfer factor	Prediction
1	11777	$0.0089 \pm 0.0022$	$105.30 \pm 25.45$
2	2426	$0.0092 \pm 0.0022$	$22.27 \pm 5.38$
3	1259	$0.0094 \pm 0.0023$	$11.89 \pm 2.88$
4	699	$0.0094 \pm 0.0022$	$6.59 \pm 1.55$
5	668	$0.0100 \pm 0.0022$	$6.70 \pm 1.49$
6	77	$0.0103 \pm 0.0021$	$0.79 \pm 0.19$
7	1551	$0.0090 \pm 0.0020$	$13.92 \pm 3.18$
8	410	$0.0086 \pm 0.0018$	$3.53 \pm 0.75$
9	249	$0.0096 \pm 0.0022$	$2.39 \pm 0.56$
10	136	$0.0092 \pm 0.0020$	$1.26 \pm 0.30$
11	104	$0.0096 \pm 0.0021$	$1.00 \pm 0.24$
12	255	$0.0087 \pm 0.0018$	$2.21 \pm 0.47$
13	75	$0.0095 \pm 0.0019$	$0.72 \pm 0.16$
14	44	$0.0086 \pm 0.0019$	$0.38 \pm 0.10$
15	20	$0.0084 \pm 0.0016$	$0.17 \pm 0.05$
16	17	$0.0117 \pm 0.0024$	$0.20 \pm 0.06$
17	3879	$0.0084 \pm 0.0019$	$32.42 \pm 7.54$
18	861	$0.0083 \pm 0.0019$	$7.11 \pm 1.69$
19	435	$0.0087 \pm 0.0021$	$3.79 \pm 0.92$
20	232	$0.0086 \pm 0.0019$	$2.00 \pm 0.45$
21	172	$0.0094 \pm 0.0021$	$1.62 \pm 0.38$
22	2322	$0.0085 \pm 0.0018$	$19.80 \pm 4.25$
23	635	$0.0087 \pm 0.0019$	$5.51 \pm 1.21$
24	331	$0.0090 \pm 0.0018$	$2.98 \pm 0.63$
25	169	$0.0082 \pm 0.0016$	$1.38 \pm 0.30$
26	80	$0.0083 \pm 0.0016$	$0.67 \pm 0.15$
27	740	$0.0087 \pm 0.0017$	$6.41 \pm 1.31$
28	184	$0.0092 \pm 0.0019$	$1.68 \pm 0.37$
29	98	$0.0075 \pm 0.0014$	$0.73 \pm 0.16$
30	57	$0.0076 \pm 0.0015$	$0.44 \pm 0.10$
31	26	$0.0089 \pm 0.0018$	$0.23 \pm 0.07$

TABLE A.5: Observed single electron yields, after accounting for lepton scale factors and relative trigger efficiencies factor, in each of the low- $\Delta\phi$  control regions, the average scale factor due to MC-based fake rates and data/MC fake rate correction factors from T&P measurements, and the corresponding fake-rate predictions with all the uncertainties (statistical and systematic) taken into consideration.

Bin	Obs. e events	Avg. transfer factor	Prediction
1	994	$0.0122 \pm 0.0050$	$12.16 \pm 4.97$
2	177	$0.0127 \pm 0.0049$	$2.24 \pm 0.89$
3	100	$0.0117 \pm 0.0044$	$1.17 \pm 0.45$
4	59	$0.0140 \pm 0.0059$	$0.82 \pm 0.37$
5	33	$0.0127 \pm 0.0050$	$0.42 \pm 0.18$
6	3	$0.0154 \pm 0.0053$	$0.05 \pm 0.03$
7	207	$0.0116 \pm 0.0041$	$2.41 \pm 0.87$
8	43	$0.0112 \pm 0.0040$	$0.48 \pm 0.19$
9	14	$0.0071 \pm 0.0023$	$0.10 \pm 0.04$
10	3	$0.0179 \pm 0.0048$	$0.05 \pm 0.03$
11	3	$0.0222 \pm 0.0097$	$0.07 \pm 0.05$
12	36	$0.0106 \pm 0.0036$	$0.38 \pm 0.14$
13	8	$0.0080 \pm 0.0022$	$0.06 \pm 0.03$
14	3	$0.0075 \pm 0.0029$	$0.02 \pm 0.02$
15	2	$0.0185 \pm 0.0053$	$0.04 \pm 0.03$
16	0	$0.0222 \pm 0.0097$	$0.00 \pm 0.04$
17	513	$0.0115 \pm 0.0043$	$5.88 \pm 2.23$
18	132	$0.0120 \pm 0.0044$	$1.58 \pm 0.59$
19	58	$0.0100 \pm 0.0037$	$0.58 \pm 0.23$
20	23	$0.0133 \pm 0.0043$	$0.31 \pm 0.12$
21	26	$0.0106 \pm 0.0041$	$0.28 \pm 0.12$
22	358	$0.0112 \pm 0.0040$	$4.00 \pm 1.46$
23	80	$0.0110 \pm 0.0037$	$0.88 \pm 0.31$
24	35	$0.0092 \pm 0.0030$	$0.32 \pm 0.12$
25	20	$0.0102 \pm 0.0029$	$0.20 \pm 0.07$
26	8	$0.0089 \pm 0.0054$	$0.07 \pm 0.05$
27	106	$0.0117 \pm 0.0039$	$1.24 \pm 0.43$
28	23	$0.0110 \pm 0.0030$	$0.25 \pm 0.09$
29	15	$0.0115 \pm 0.0048$	$0.17 \pm 0.09$
30	6	$0.0115 \pm 0.0028$	$0.07 \pm 0.03$
31	2	$0.0073 \pm 0.0018$	$0.01 \pm 0.01$

TABLE A.6:  $Z(\nu\nu) + \gamma$  MC yields and prediction for the signal regions.

Bin	MC Yield	Prediction
1	$80.90 \pm 0.68$	$92.21 \pm 17.39$
2	$29.46 \pm 0.41$	$33.57 \pm 8.34$
3	$20.06 \pm 0.34$	$22.87 \pm 6.02$
4	$14.90 \pm 0.29$	$16.98 \pm 5.15$
5	$15.92 \pm 0.30$	$18.14 \pm 7.07$
6	$2.46 \pm 0.12$	$2.80 \pm 1.23$
7	$7.30 \pm 0.23$	$8.32 \pm 1.59$
8	$2.71 \pm 0.14$	$3.09 \pm 0.78$
9	$1.74 \pm 0.12$	$1.98 \pm 0.54$
10	$1.31 \pm 0.10$	$1.49 \pm 0.47$
11	$1.44 \pm 0.11$	$1.65 \pm 0.65$
12	$0.70 \pm 0.08$	$0.80 \pm 0.18$
13	$0.32 \pm 0.06$	$0.37 \pm 0.11$
14	$0.21 \pm 0.05$	$0.24 \pm 0.08$
15	$0.14 \pm 0.04$	$0.16 \pm 0.07$
16	$0.15 \pm 0.03$	$0.17 \pm 0.08$
17	$8.43 \pm 0.14$	$9.61 \pm 1.82$
18	$3.12 \pm 0.08$	$3.55 \pm 0.89$
19	$2.15 \pm 0.07$	$2.45 \pm 0.65$
20	$1.59 \pm 0.05$	$1.81 \pm 0.55$
21	$1.88 \pm 0.06$	$2.14 \pm 0.84$
22	$1.63 \pm 0.08$	$1.86 \pm 0.36$
23	$0.67 \pm 0.06$	$0.76 \pm 0.20$
24	$0.43 \pm 0.04$	$0.49 \pm 0.14$
25	$0.28 \pm 0.03$	$0.32 \pm 0.11$
26	$0.42 \pm 0.06$	$0.48 \pm 0.20$
27	$0.19 \pm 0.02$	$0.21 \pm 0.05$
28	$0.12 \pm 0.02$	$0.13 \pm 0.04$
29	$0.09 \pm 0.03$	$0.10 \pm 0.04$
30	$0.06 \pm 0.02$	$0.07 \pm 0.03$
31	$0.04 \pm 0.01$	$0.04 \pm 0.02$

TABLE A.7:  $Z(\nu\nu) + \gamma$  MC yields and prediction for the low- $\Delta\Phi$  control regions.

Bin	MC Yield	Prediction
1	$19.13 \pm 0.32$	$21.28 \pm 4.36$
2	$7.46 \pm 0.22$	$8.29 \pm 2.17$
3	$4.52 \pm 0.17$	$5.03 \pm 1.39$
4	$2.91 \pm 0.14$	$3.24 \pm 1.02$
5	$2.01 \pm 0.12$	$2.24 \pm 0.90$
6	$0.16 \pm 0.03$	$0.18 \pm 0.09$
7	$1.77 \pm 0.14$	$1.97 \pm 0.43$
8	$0.72 \pm 0.09$	$0.80 \pm 0.23$
9	$0.51 \pm 0.08$	$0.57 \pm 0.18$
10	$0.24 \pm 0.05$	$0.26 \pm 0.10$
11	$0.20 \pm 0.04$	$0.23 \pm 0.10$
12	$0.21 \pm 0.05$	$0.23 \pm 0.07$
13	$0.12 \pm 0.04$	$0.14 \pm 0.06$
14	$0.04 \pm 0.01$	$0.04 \pm 0.02$
15	$0.02 \pm 0.01$	$0.03 \pm 0.01$
16	$0.02 \pm 0.01$	$0.03 \pm 0.01$
17	$2.05 \pm 0.07$	$2.28 \pm 0.49$
18	$0.81 \pm 0.05$	$0.90 \pm 0.24$
19	$0.48 \pm 0.03$	$0.53 \pm 0.15$
20	$0.34 \pm 0.03$	$0.38 \pm 0.13$
21	$0.28 \pm 0.04$	$0.31 \pm 0.13$
22	$0.54 \pm 0.08$	$0.60 \pm 0.15$
23	$0.16 \pm 0.03$	$0.17 \pm 0.06$
24	$0.11 \pm 0.03$	$0.12 \pm 0.04$
25	$0.06 \pm 0.01$	$0.06 \pm 0.02$
26	$0.07 \pm 0.02$	$0.08 \pm 0.04$
27	$0.06 \pm 0.01$	$0.06 \pm 0.02$
28	$0.04 \pm 0.02$	$0.04 \pm 0.02$
29	$0.01 \pm 0.01$	$0.02 \pm 0.01$
30	$0.008 \pm 0.003$	$0.009 \pm 0.005$
31	$0.006 \pm 0.003$	$0.007 \pm 0.004$

TABLE A.8: Fake  $p_T^{miss}$  background control sample statistics and electroweak predictions with uncertainties.

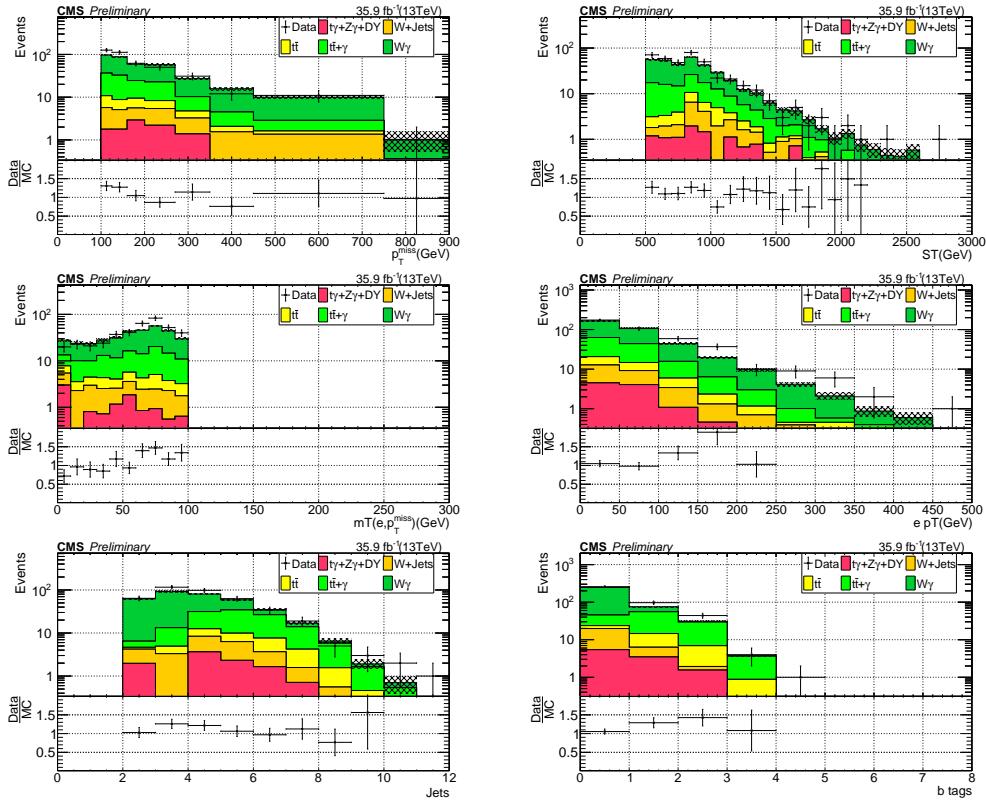
$N_{\text{jets}}^{\text{b-jets}}$	Bin	Low $\Delta\phi$ raw	EW low $\Delta\phi$	$\gamma + \text{jets}$ low $\Delta\phi$	$\gamma + \text{jets}$ high $\Delta\phi$
$N_{2-4}^0$	1	5109	$88 \pm 11$	$5020 \pm 72$	$3994 \pm 112$
	2	286	$28.30 \pm 4.17$	$257.70 \pm 17.42$	$59.79 \pm 11.06$
	3	101	$12.52 \pm 2.38$	$88.48 \pm 10.33$	$20.53 \pm 4.27$
	4	26	$8.29 \pm 1.92$	$17.71 \pm 5.45$	$4.11 \pm 1.45$
	5	21	$10.31 \pm 2.33$	$10.69 \pm 5.14$	$2.48 \pm 1.27$
	6	2	$0.23 \pm 1.10$	$1.77 \pm 1.79$	$0.41 \pm 0.42$
$N_{5-6}^0$	7	710	$13.38 \pm 2.72$	$696.62 \pm 26.78$	$653.69 \pm 44.71$
	8	38	$1.99 \pm 0.64$	$36.01 \pm 6.20$	$15.76 \pm 4.79$
	9	11	$2.63 \pm 1.01$	$8.37 \pm 3.47$	$3.67 \pm 1.77$
	10	4	$1.19 \pm 0.61$	$2.81 \pm 2.09$	$1.23 \pm 0.97$
	11	1	$0.84 \pm 0.63$	$0.16 \pm 1.18$	$0.07 \pm 0.52$
$N_{\geq 7}^0$	12	89	$2.60 \pm 1.04$	$86.40 \pm 9.49$	$89.95 \pm 17.44$
	13	6	$1.56 \pm 0.95$	$4.44 \pm 2.63$	$1.84 \pm 1.24$
	14	3	$0.06 \pm 1.11$	$2.94 \pm 2.06$	$1.22 \pm 0.94$
	15	1	$0.83 \pm 0.65$	$0.17 \pm 1.19$	$0.07 \pm 0.50$
	16	0	$0.62 \pm 0.70$	$0.00 \pm 1.80$	$0.00 \pm 0.75$
$N_{2-4}^{\geq 1}$	17	1320	$37.18 \pm 6.42$	$1282.82 \pm 36.89$	$837.60 \pm 47.74$
	18	78	$6.31 \pm 1.67$	$71.69 \pm 8.99$	$11.29 \pm 3.30$
	19	39	$2.76 \pm 1.06$	$36.24 \pm 6.33$	$5.71 \pm 1.81$
	20	5	$1.24 \pm 1.31$	$3.76 \pm 2.59$	$0.59 \pm 0.44$
	21	8	$1.96 \pm 1.12$	$6.04 \pm 3.04$	$0.95 \pm 0.54$
$N_{5-6}^{\geq 1}$	22	319	$16.85 \pm 3.72$	$302.15 \pm 18.24$	$275.08 \pm 31.23$
	23	28	$3.39 \pm 1.05$	$24.61 \pm 5.39$	$7.72 \pm 2.45$
	24	8	$1.23 \pm 0.61$	$6.77 \pm 2.89$	$2.12 \pm 1.03$
	25	2	$1.05 \pm 0.61$	$0.95 \pm 1.54$	$0.30 \pm 0.49$
	26	1	$1.19 \pm 0.66$	$0.00 \pm 1.80$	$0.00 \pm 0.56$
$N_{\geq 7}^{\geq 1}$	27	61	$4.16 \pm 1.34$	$56.84 \pm 7.92$	$74.10 \pm 18.10$
	28	12	$2.38 \pm 1.05$	$9.62 \pm 3.62$	$5.94 \pm 4.95$
	29	0	$0.44 \pm 0.99$	$0.00 \pm 1.80$	$0.00 \pm 1.11$
	30	0	$0.33 \pm 0.99$	$0.00 \pm 1.80$	$0.00 \pm 1.11$
	31	0	$0.02 \pm 1.05$	$0.00 \pm 1.80$	$0.00 \pm 1.11$

## Appendix B

# Data verses MC in control regions

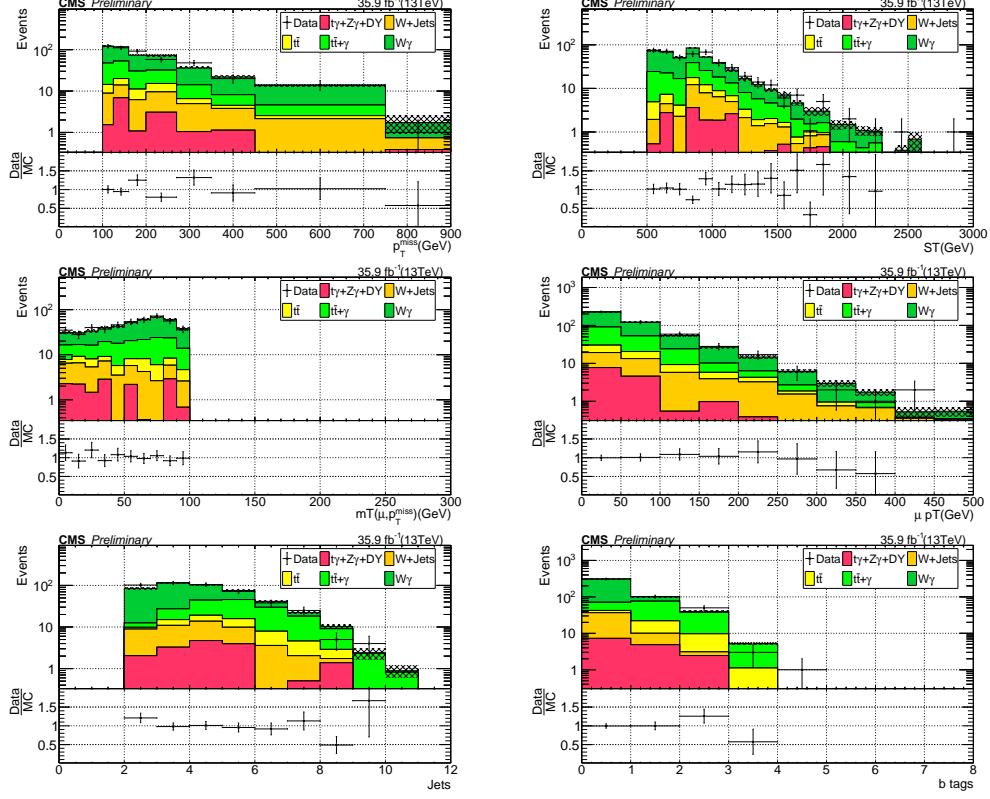
### B.1 $e/\mu + \gamma$ control region

FIGURE B.1: Distribution of  $e + \gamma$  events versus  $p_T^{miss}$  (top left),  $H_T^\gamma$  (top right),  $m_T$  (middle left),  $p_{T,e}$  (middle right),  $N_{\text{jets}}$  (bottom left),  $N_{\text{b-jets}}$  (bottom right). Filled histograms denote expected yields from MC for various processes. Markers represent data yields. MC event yields are scaled to the same integrated luminosity as data.



Comparisons of the observed yields and expected MC yields for the electron and muon control regions in Figures B.2 and B.1. While the agreement between data and MC expectations is not directly required, MC modeling of ratios of events can affect the lost lepton prediction. Thus, data-MC comparisons of the control region represent an important validation of the systematic uncertainties of the transfer factors. MC modeling of the various analysis variables used to define the signal regions are generally found to have good agreement with data. For the  $e + \gamma$  CR the  $p_T$ ,  $p_T^{miss}$ , and  $m_T$  distributions are found to have differing shapes for data and MC. This could

FIGURE B.2: Distribution of  $\mu + \gamma$  events versus  $p_T^{miss}$  (top left),  $S_T$  (top right),  $M_T$  (middle left),  $p_{T,\mu}$  (middle right),  $N_{jets}$  (bottom left),  $N_b$ -jets (bottom right). Filled histograms denote expected yields from MC for various processes. Markers represent data yields. MC event yields are scaled to the same integrated luminosity as data.



be because of some mis-modelling in the electron  $p_T$  and  $p_T^{miss}$  which affect  $m_T$  distribution since  $m_T$  is derived from  $p_T^{miss}$  and electron  $p_T$ . Effect of electron  $p_T$  mis-modeling can have maximum 10% effect on TF. Since the existing uncertainty on TF is 15-25%, this systematic effect(if any) is already covered. In low  $p_T^{miss}$  bins(100-200 GeV), data-MC differences is found to be 20%. The low  $p_T^{miss}$  region is used as a sideband for multijet estimation and it does not correspond to any signal region. Lost electron contribution is very negligible in this sideband and it is dominated by multijet. So these differences are going to have very small effect (1-2% or less) on multijet prediction.

All yields are computed after applying the b-tagging scale factors [93] and lepton scale factors to MC.

## B.2 Z(ℓℓ)+γ control region

A comparison of  $ll + \gamma$  events in data with MC is shown in figure B.3. In top left plot, since di-lepton pair is ignored,  $p_T$  of the pair contributes to  $p_T^{miss}$ . Data agrees well with MC, within the statistical uncertainties.

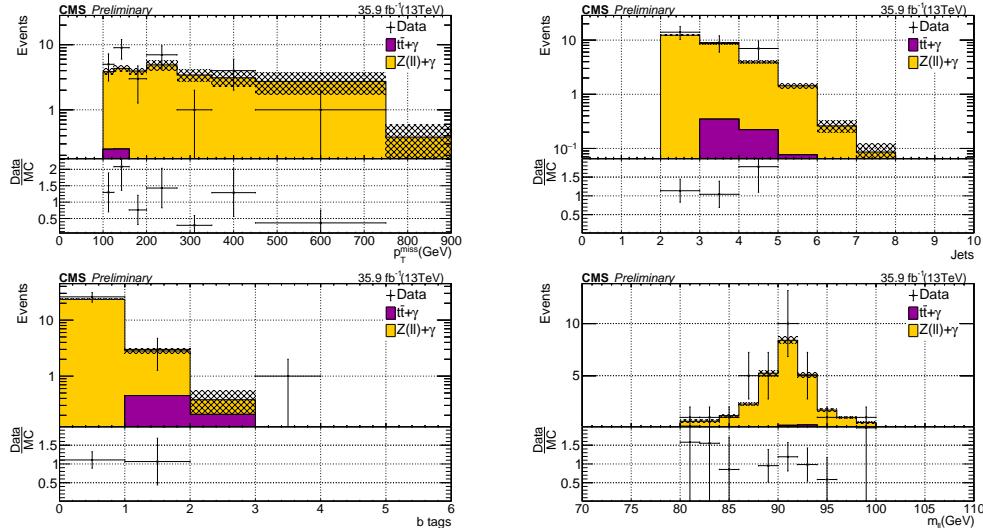


FIGURE B.3: Comparison of  $(ll)\gamma$  in data with  $ll + \gamma$  events in MC as function of  $p_T^{\text{miss}}$  (top left),  $N_{\text{jets}}$  (top right),  $N_{\text{b-jets}}$  (bottom left) and  $m_{ll}$  (bottom right) for  $p_T^\gamma > 190\text{GeV}$ . Note that  $p_T^{\text{miss}}$  is calculated by ignoring the di-lepton pair. For the top left plot of  $p_T^{\text{miss}}$ , hashed error bars refer to statistical and EW correction uncertainties listed in Table 3.5. For the remaining 3 plots, error bars are statistical only.

### B.3 Signal acceptance $\times$ efficiency

Signal acceptance  $\times$  efficiency is studied for different signal models and they are shown in figure B.4. To some extent acceptance  $\times$  efficiency is related to sensitivity of the analysis. Other factors such as background composition in different search bins, signal kinematics and cross section also affect the sensitivity and hence the exclusion limits. These plots are useful to in understanding some of the features seen in exclusion curves- low sensitivity near diagonal and kink for very low NLSP masses etc.

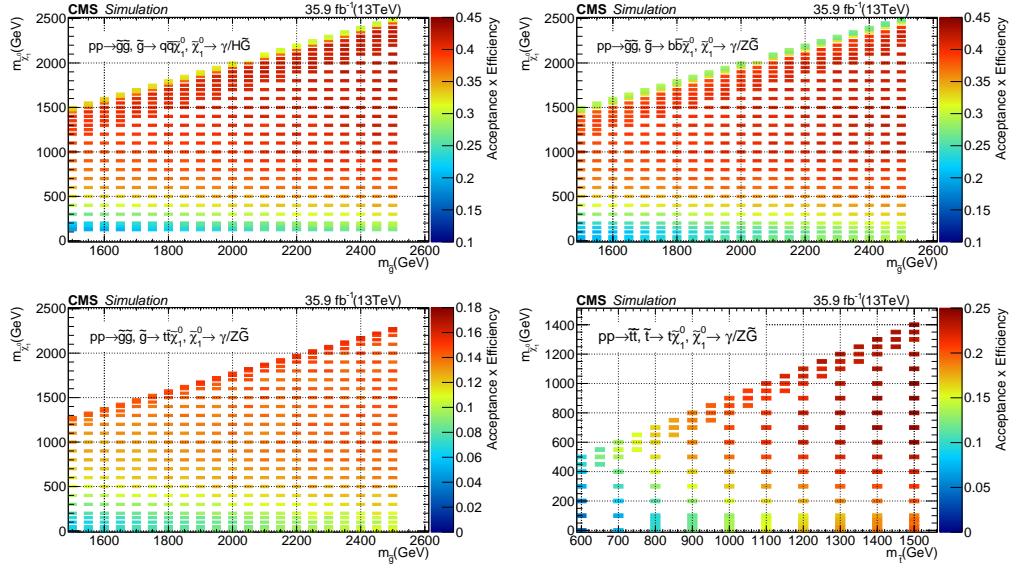


FIGURE B.4: Acceptance  $\times$  efficiency for different mass point corresponding to T5qqqqHG (top left), T5bbbbZG (top right), T5ttttZG (bottom left) and T6ttZG (bottom right).

# Bibliography

- [1] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716 (2012), p. 1. DOI: [10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020). arXiv: [1207.7214 \[hep-ex\]](https://arxiv.org/abs/1207.7214).
- [2] S. Abdullin et al. "The fast simulation of the CMS detector at LHC". In: *J. Phys. Conf. Ser.* 331 (2011), p. 032049. DOI: [10.1088/1742-6596/331/3/032049](https://doi.org/10.1088/1742-6596/331/3/032049).
- [3] W Adam et al. "Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC". In: *Journal of Physics G: Nuclear and Particle Physics* 31.9 (2005), N9–N20. DOI: [10.1088/0954-3899/31/9/n01](https://doi.org/10.1088/0954-3899/31/9/n01). URL: <https://doi.org/10.1088%2F0954-3899%2F31%2F9%2Fn01>.
- [4] S. Agostinelli et al. "GEANT4—a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250. DOI: [10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [5] D. Alves et al. "Simplified models for LHC new physics searches". In: *J. Phys. G* 39 (2012), p. 105005. DOI: [10.1088/0954-3899/39/10/105005](https://doi.org/10.1088/0954-3899/39/10/105005). arXiv: [1105.2838 \[hep-ph\]](https://arxiv.org/abs/1105.2838).
- [6] J. Alwall et al. "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions". In: *Eur. Phys. J. C* 53 (2008), p. 473. DOI: [10.1140/epjc/s10052-007-0490-5](https://doi.org/10.1140/epjc/s10052-007-0490-5). arXiv: [0706.2569 \[hep-ph\]](https://arxiv.org/abs/0706.2569).
- [7] 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". In: *JHEP* 07 (2014), p. 079. DOI: [10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079). arXiv: [1405.0301 \[hep-ph\]](https://arxiv.org/abs/1405.0301).
- [8] J. Alwall, P. C. Schuster, and N. Toro. "Simplified models for a first characterization of new physics at the LHC". In: *Phys. Rev. D* 79 (2009), p. 075020. DOI: [10.1103/PhysRevD.79.075020](https://doi.org/10.1103/PhysRevD.79.075020). arXiv: [0810.3921 \[hep-ph\]](https://arxiv.org/abs/0810.3921).
- [9] J. Alwall et al. "Model-independent jets plus missing energy searches". In: *Phys. Rev. D* 79 (2009), p. 015005. DOI: [10.1103/PhysRevD.79.015005](https://doi.org/10.1103/PhysRevD.79.015005). arXiv: [0809.3264 \[hep-ph\]](https://arxiv.org/abs/0809.3264).
- [10] N. Arkani-Hamed et al. "MAMBOSET: The path from LHC data to the new standard model via on-shell effective theories". 2007.
- [11] G. Arnison et al. "Experimental observation of isolated large transverse energy electrons with associated missing energy at  $\sqrt{s} = 540 \text{ GeV}$ ". In: *Physics Letters B* 122.1 (1983), pp. 103 –116. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(83\)91177-2](https://doi.org/10.1016/0370-2693(83)91177-2). URL: <http://www.sciencedirect.com/science/article/pii/0370269383911772>.

- [12] G. Arnison et al. “Experimental observation of lepton pairs of invariant mass around  $95 \text{ GeV}/c^2$  at the CERN SPS collider”. In: *Physics Letters B* 126.5 (1983), pp. 398–410. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(83\)90188-0](https://doi.org/10.1016/0370-2693(83)90188-0). URL: <http://www.sciencedirect.com/science/article/pii/0370269383901880>.
- [13] P. Artoisenet et al. “Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations”. In: *JHEP* 03 (2013), p. 015. DOI: [10.1007/JHEP03\(2013\)015](https://doi.org/10.1007/JHEP03(2013)015). arXiv: [1212.3460 \[hep-ph\]](https://arxiv.org/abs/1212.3460).
- [14] H. Baer and X. Tata. “The Minimal Supersymmetric Standard Model”. In: *Weak Scale Supersymmetry: From Superfields to Scattering Events*. Cambridge University Press, 2006, 127–189. DOI: [10.1017/CBO9780511617270.009](https://doi.org/10.1017/CBO9780511617270.009).
- [15] P. Bagnaia et al. “Evidence for  $Z^0 \rightarrow e^+e^-$  CERN pp collider”. In: *Physics Letters B* 129.1 (1983), pp. 130–140. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(83\)90744-X](https://doi.org/10.1016/0370-2693(83)90744-X). URL: <http://www.sciencedirect.com/science/article/pii/037026938390744X>.
- [16] R. D. Ball et al. “Parton distributions for the LHC Run II”. In: *JHEP* 04 (2015), p. 040. DOI: [10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040). arXiv: [1410.8849 \[hep-ph\]](https://arxiv.org/abs/1410.8849).
- [17] M. Banner et al. “Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider”. In: *Physics Letters B* 122.5 (1983), pp. 476–485. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(83\)91605-2](https://doi.org/10.1016/0370-2693(83)91605-2). URL: <http://www.sciencedirect.com/science/article/pii/0370269383916052>.
- [18] W. Beenakker et al. “The Production of charginos / neutralinos and sleptons at hadron colliders”. In: *Phys. Rev. Lett.* 83 (1999). [Erratum: *Phys. Rev. Lett.* 100, 029901 (2008)], pp. 3780–3783. DOI: [10.1103/PhysRevLett.100.029901](https://doi.org/10.1103/PhysRevLett.100.029901), [10.1103/PhysRevLett.83.3780](https://doi.org/10.1103/PhysRevLett.83.3780). arXiv: [hep-ph/9906298 \[hep-ph\]](https://arxiv.org/abs/hep-ph/9906298).
- [19] W. Beenakker et al. “Soft-gluon resummation for squark and gluino hadroproduction”. In: *JHEP* 12 (2009), p. 041. DOI: [10.1088/1126-6708/2009/12/041](https://doi.org/10.1088/1126-6708/2009/12/041). arXiv: [0909.4418 \[hep-ph\]](https://arxiv.org/abs/0909.4418).
- [20] W. Beenakker et al. “Squark and gluino hadroproduction”. In: *Int. J. Mod. Phys. A* 26 (2011), p. 2637. DOI: [10.1142/S0217751X11053560](https://doi.org/10.1142/S0217751X11053560). arXiv: [1105.1110 \[hep-ph\]](https://arxiv.org/abs/1105.1110).
- [21] W. Beenakker et al. “Squark and gluino production at hadron colliders”. In: *Nucl. Phys. B* 492 (1997), p. 51. DOI: [10.1016/S0550-3213\(97\)00084-9](https://doi.org/10.1016/S0550-3213(97)00084-9). arXiv: [hep-ph/9610490 \[hep-ph\]](https://arxiv.org/abs/hep-ph/9610490).
- [22] C. Borschensky et al. “Squark and gluino production cross sections in pp collisions at  $\sqrt{s} = 13, 14, 33$  and  $100 \text{ TeV}$ ”. In: *Eur. Phys. J. C* 74 (2014), p. 3174. DOI: [10.1140/epjc/s10052-014-3174-y](https://doi.org/10.1140/epjc/s10052-014-3174-y). arXiv: [1407.5066 \[hep-ph\]](https://arxiv.org/abs/1407.5066).
- [23] O. S. Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. URL: <https://cds.cern.ch/record/782076>.
- [24] M. Cacciari et al. “The  $t\bar{t}$  cross-section at 1.8 TeV and 1.96 TeV: a study of the systematics due to parton densities and scale dependence”. In: *JHEP* 04 (2004), p. 068. DOI: [10.1088/1126-6708/2004/04/068](https://doi.org/10.1088/1126-6708/2004/04/068). arXiv: [hep-ph/0303085 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0303085).
- [25] M. Cacciari, G. P. Salam, and G. Soyez. “FastJet user manual”. In: *Eur. Phys. J. C* 72 (2012), p. 1896. DOI: [10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2). arXiv: [1111.6097 \[hep-ph\]](https://arxiv.org/abs/1111.6097).

- [26] M. Cacciari, G. P. Salam, and G. Soyez. “The anti- $k_T$  jet clustering algorithm”. In: *JHEP* 04 (2008), p. 063. DOI: [10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063). arXiv: [0802.1189 \[hep-ph\]](https://arxiv.org/abs/0802.1189).
- [27] S. Catani et al. “Soft gluon resummation for Higgs boson production at hadron colliders”. In: *JHEP* 07 (2003), p. 028. DOI: [10.1088/1126-6708/2003/07/028](https://doi.org/10.1088/1126-6708/2003/07/028). arXiv: [hep-ph/0306211 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0306211).
- [28] S. Chatrchyan et al. “The CMS experiment at the CERN LHC”. In: *JINST* 3 (2008), S08004. DOI: [10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004).
- [29] S. Chatrchyan et al. “Description and performance of track and primary-vertex reconstruction with the CMS tracker”. In: *JINST* 9.10 (2014), P10009. DOI: [10.1088/1748-0221/9/10/P10009](https://doi.org/10.1088/1748-0221/9/10/P10009). arXiv: [1405.6569 \[physics.ins-det\]](https://arxiv.org/abs/1405.6569).
- [30] S. Chatrchyan et al. “Determination of jet energy calibration and transverse momentum resolution in CMS”. In: *JINST* 6 (2011), P11002. DOI: [10.1088/1748-0221/6/11/P11002](https://doi.org/10.1088/1748-0221/6/11/P11002). arXiv: [1107.4277 \[physics.ins-det\]](https://arxiv.org/abs/1107.4277).
- [31] S. Chatrchyan et al. “Interpretation of searches for supersymmetry with simplified models”. In: *Phys. Rev. D* 88 (2013), p. 052017. DOI: [10.1103/PhysRevD.88.052017](https://doi.org/10.1103/PhysRevD.88.052017). arXiv: [1301.2175 \[hep-ex\]](https://arxiv.org/abs/1301.2175).
- [32] S. Chatrchyan et al. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: [10.1016/j.physlettb.2012.08.021](https://doi.org/10.1016/j.physlettb.2012.08.021). arXiv: [1207.7235 \[hep-ex\]](https://arxiv.org/abs/1207.7235).
- [33] S. Chatrchyan et al. “Performance of CMS muon reconstruction in  $pp$  collision events at  $\sqrt{s} = 7$  TeV”. In: *JINST* 7 (2012), P10002. DOI: [10.1088/1748-0221/7/10/P10002](https://doi.org/10.1088/1748-0221/7/10/P10002). arXiv: [1206.4071 \[physics.ins-det\]](https://arxiv.org/abs/1206.4071).
- [34] S. Chatrchyan et al. “The Performance of the CMS Muon Detector in Proton-Proton Collisions at  $\sqrt{s} = 7$  TeV at the LHC”. In: *JINST* 8 (2013), P11002. DOI: [10.1088/1748-0221/8/11/P11002](https://doi.org/10.1088/1748-0221/8/11/P11002). arXiv: [1306.6905 \[physics.ins-det\]](https://arxiv.org/abs/1306.6905).
- [35] CMS Collaboration. *CMS SUSY Results: Objects Efficiency*. URL: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SUSMoriond2017ObjectsEfficiency>.
- [36] CMS Collaboration. *Performance of missing energy reconstruction in 13 TeV  $pp$  collision data using the CMS detector*. CMS Physics Analysis Summary CMS-PAS-JME-16-004. 2016. URL: <http://cds.cern.ch/record/2205284>.
- [37] CMS Collaboration. *Public CMS Luminosity Information*. URL: [https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2016\\_proton\\_proton\\_collisions](https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2016_proton_proton_collisions).
- [38] CMS Collaboration. “Summaries of CMS cross section measurements”. URL: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined>.
- [39] C. Collaboration. “Performance of the CMS hadron calorimeter with cosmic ray muons and LHC beam data”. In: *Journal of Instrumentation* 5.03 (2010), T03012–T03012. DOI: [10.1088/1748-0221/5/03/t03012](https://doi.org/10.1088/1748-0221/5/03/t03012). URL: <https://doi.org/10.1088%2F1748-0221%2F5%2F03%2FT03012>.
- [40] T. C. Collaboration. “Description and performance of track and primary-vertex reconstruction with the CMS tracker”. In: *Journal of Instrumentation* 9.10 (2014), P10009–P10009. DOI: [10.1088/1748-0221/9/10/p10009](https://doi.org/10.1088/1748-0221/9/10/p10009). URL: <https://doi.org/10.1088%2F1748-0221%2F9%2F10%2Fp10009>.

- [41] G. Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J.* C71 (2011). [Erratum: Eur. Phys. J.C73,2501(2013)], p. 1554. DOI: [10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0), [10.1140/epjc/s10052-013-2501-z](https://doi.org/10.1140/epjc/s10052-013-2501-z). arXiv: [1007.1727 \[physics.data-an\]](https://arxiv.org/abs/1007.1727).
- [42] E. Cremmer et al. "Super-Higgs effect in supergravity with general scalar interactions". In: *Phys. Lett. B* 79 (1978), p. 231. ISSN: 0370-2693. DOI: [10.1016/0370-2693\(78\)90230-7](https://doi.org/10.1016/0370-2693(78)90230-7).
- [43] M. Czakon and A. Mitov. "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders". In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: [10.1016/j.cpc.2014.06.021](https://doi.org/10.1016/j.cpc.2014.06.021). arXiv: [1112.5675 \[hep-ph\]](https://arxiv.org/abs/1112.5675).
- [44] A. Denner et al. "NLO QCD and electroweak corrections to  $Z + \gamma$  production with leptonic Z-boson decays". In: *JHEP* 02 (2016), p. 057. DOI: [10.1007/JHEP02\(2016\)057](https://doi.org/10.1007/JHEP02(2016)057). arXiv: [1510.08742 \[hep-ph\]](https://arxiv.org/abs/1510.08742).
- [45] S. Deser and B. Zumino. "Broken Supersymmetry and Supergravity". In: *Phys. Rev. Lett.* 38 (1977), p. 1433. DOI: [10.1103/PhysRevLett.38.1433](https://doi.org/10.1103/PhysRevLett.38.1433).
- [46] M. Dine and A. E. Nelson. "Dynamical supersymmetry breaking at low-energies". In: *Phys. Rev.* D48 (1993), pp. 1277–1287. DOI: [10.1103/PhysRevD.48.1277](https://doi.org/10.1103/PhysRevD.48.1277). arXiv: [hep-ph/9303230 \[hep-ph\]](https://arxiv.org/abs/hep-ph/9303230).
- [47] M. Dine, A. E. Nelson, and Y. Shirman. "Low-energy dynamical supersymmetry breaking simplified". In: *Phys. Rev.* D51 (1995), pp. 1362–1370. DOI: [10.1103/PhysRevD.51.1362](https://doi.org/10.1103/PhysRevD.51.1362). arXiv: [hep-ph/9408384 \[hep-ph\]](https://arxiv.org/abs/hep-ph/9408384).
- [48] M. Dine et al. "New tools for low-energy dynamical supersymmetry breaking". In: *Phys. Rev.* D53 (1996), pp. 2658–2669. DOI: [10.1103/PhysRevD.53.2658](https://doi.org/10.1103/PhysRevD.53.2658). arXiv: [hep-ph/9507378 \[hep-ph\]](https://arxiv.org/abs/hep-ph/9507378).
- [49] J. Ellis et al. "Supersymmetric relics from the big bang". In: *Nuclear Physics B* 238.2 (1984), pp. 453–476. ISSN: 0550-3213. DOI: [https://doi.org/10.1016/0550-3213\(84\)90461-9](https://doi.org/10.1016/0550-3213(84)90461-9). URL: <http://www.sciencedirect.com/science/article/pii/0550321384904619>.
- [50] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (1964). [,157(1964)], pp. 321–323. DOI: [10.1103/PhysRevLett.13.321](https://doi.org/10.1103/PhysRevLett.13.321).
- [51] R. Frederix and S. Frixione. "Merging meets matching in MC@NLO". In: *JHEP* 12 (2012), p. 061. DOI: [10.1007/JHEP12\(2012\)061](https://doi.org/10.1007/JHEP12(2012)061). arXiv: [1209.6215 \[hep-ph\]](https://arxiv.org/abs/1209.6215).
- [52] R. Fröhwirth, W. Waltenberger, and P. Vanlaer. "Adaptive vertex fitting". In: *J. Phys.* G34 (2007), N343. DOI: [10.1088/0954-3899/34/12/N01](https://doi.org/10.1088/0954-3899/34/12/N01).
- [53] B. Fuks et al. "Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV". In: *JHEP* 10 (2012), p. 081. DOI: [10.1007/JHEP10\(2012\)081](https://doi.org/10.1007/JHEP10(2012)081). arXiv: [1207.2159 \[hep-ph\]](https://arxiv.org/abs/1207.2159).
- [54] B. Fuks et al. "Precision predictions for electroweak superpartner production at hadron colliders with Resummino". In: *Eur. Phys. J.* C73 (2013), p. 2480. DOI: [10.1140/epjc/s10052-013-2480-0](https://doi.org/10.1140/epjc/s10052-013-2480-0). arXiv: [1304.0790 \[hep-ph\]](https://arxiv.org/abs/1304.0790).
- [55] R. Gavin et al. "FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order". In: *Comput. Phys. Commun.* 182 (2011), p. 2388. DOI: [10.1016/j.cpc.2011.06.008](https://doi.org/10.1016/j.cpc.2011.06.008). arXiv: [1011.3540 \[hep-ph\]](https://arxiv.org/abs/1011.3540).
- [56] R. Gavin et al. "W Physics at the LHC with FEWZ 2.1". In: *Comput. Phys. Commun.* 184 (2013), p. 208. DOI: [10.1016/j.cpc.2012.09.005](https://doi.org/10.1016/j.cpc.2012.09.005). arXiv: [1201.5896 \[hep-ph\]](https://arxiv.org/abs/1201.5896).

- [57] A. Giammanco. "The Fast Simulation of the CMS Experiment". In: *J. Phys. Conf. Ser.* 513 (2014), p. 022012. DOI: [10.1088/1742-6596/513/2/022012](https://doi.org/10.1088/1742-6596/513/2/022012).
- [58] G. F. Giudice and R. Rattazzi. "Theories with gauge mediated supersymmetry breaking". In: *Phys. Rept.* 322 (1999), pp. 419–499. DOI: [10.1016/S0370-1573\(99\)00042-3](https://doi.org/10.1016/S0370-1573(99)00042-3). arXiv: [hep-ph/9801271 \[hep-ph\]](https://arxiv.org/abs/hep-ph/9801271).
- [59] S. L. Glashow. "The renormalizability of vector meson interactions". In: *Nucl. Phys.* 10 (1959), pp. 107–117. DOI: [10.1016/0029-5582\(59\)90196-8](https://doi.org/10.1016/0029-5582(59)90196-8).
- [60] P. Grajek, A. Mariotti, and D. Redigolo. "Phenomenology of General Gauge Mediation in light of a 125 GeV Higgs". In: *JHEP* 07 (2013), p. 109. DOI: [10.1007/JHEP07\(2013\)109](https://doi.org/10.1007/JHEP07(2013)109). arXiv: [1303.0870 \[hep-ph\]](https://arxiv.org/abs/1303.0870).
- [61] D. J. Gross and F. Wilczek. "Ultraviolet Behavior of Non-Abelian Gauge Theories". In: *Phys. Rev. Lett.* 30 (26 1973), pp. 1343–1346. DOI: [10.1103/PhysRevLett.30.1343](https://doi.org/10.1103/PhysRevLett.30.1343). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.30.1343>.
- [62] J. Haffner. "The CERN accelerator complex. Complexe des accélérateurs du CERN". In: (2013). General Photo. URL: <https://cds.cern.ch/record/1621894>.
- [63] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964). [,160(1964)], pp. 508–509. DOI: [10.1103/PhysRevLett.13.508](https://doi.org/10.1103/PhysRevLett.13.508).
- [64] *Jet algorithms performance in 13 TeV data*. CMS Physics Analysis Summary CMS-PAS-JME-16-003. 2017. URL: <http://cds.cern.ch/record/2256875>.
- [65] "Jet energy scale and resolution performance with 13 TeV data collected by CMS in 2016". In: (2018). URL: <http://cds.cern.ch/record/2622157>.
- [66] "Jet energy scale and resolution performances with 13TeV data". In: (2016). URL: <https://cds.cern.ch/record/2160347>.
- [67] T. Junk. "Confidence level computation for combining searches with small statistics". In: *Nucl. Instrum. Meth. A* 434 (1999), p. 435. DOI: [10.1016/S0168-9002\(99\)00498-2](https://doi.org/10.1016/S0168-9002(99)00498-2). arXiv: [hep-ex/9902006 \[hep-ex\]](https://arxiv.org/abs/hep-ex/9902006).
- [68] A. Kalogeropoulos and J. Alwall. "The SysCalc code: A tool to derive theoretical systematic uncertainties". 2018.
- [69] V. Khachatryan et al. "Event generator tunes obtained from underlying event and multiparton scattering measurements". In: *Eur. Phys. J. C* 76 (2016), p. 155. DOI: [10.1140/epjc/s10052-016-3988-x](https://doi.org/10.1140/epjc/s10052-016-3988-x). arXiv: [1512.00815 \[hep-ex\]](https://arxiv.org/abs/1512.00815).
- [70] V. Khachatryan et al. "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV". In: *JINST* 12 (2017), P02014. DOI: [10.1088/1748-0221/12/02/P02014](https://doi.org/10.1088/1748-0221/12/02/P02014). arXiv: [1607.03663 \[hep-ex\]](https://arxiv.org/abs/1607.03663).
- [71] V. Khachatryan et al. "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV". In: *JINST* 10 (2015), P06005. DOI: [10.1088/1748-0221/10/06/P06005](https://doi.org/10.1088/1748-0221/10/06/P06005). arXiv: [1502.02701 \[physics.ins-det\]](https://arxiv.org/abs/1502.02701).
- [72] V. Khachatryan et al. "Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV". In: *JINST* 10 (2015), P08010. DOI: [10.1088/1748-0221/10/08/P08010](https://doi.org/10.1088/1748-0221/10/08/P08010). arXiv: [1502.02702 \[physics.ins-det\]](https://arxiv.org/abs/1502.02702).
- [73] V. Khachatryan et al. "The CMS trigger system". In: *JINST* 12 (2017), P01020. DOI: [10.1088/1748-0221/12/01/P01020](https://doi.org/10.1088/1748-0221/12/01/P01020). arXiv: [1609.02366 \[physics.ins-det\]](https://arxiv.org/abs/1609.02366).

- [74] S. Knapen and D. Redigolo. "Gauge mediation at the LHC: status and prospects". In: *JHEP* 01 (2017), p. 135. DOI: [10.1007/JHEP01\(2017\)135](https://doi.org/10.1007/JHEP01(2017)135). arXiv: [1606.07501 \[hep-ph\]](https://arxiv.org/abs/1606.07501).
- [75] A. Kulesza and L. Motyka. "Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC". In: *Phys. Rev. D* 80 (2009), p. 095004. DOI: [10.1103/PhysRevD.80.095004](https://doi.org/10.1103/PhysRevD.80.095004). arXiv: [0905.4749 \[hep-ph\]](https://arxiv.org/abs/0905.4749).
- [76] A. Kulesza and L. Motyka. "Threshold Resummation for Squark-Antisquark and Gluino-Pair Production at the LHC". In: *Phys. Rev. Lett.* 102 (2009), p. 111802. DOI: [10.1103/PhysRevLett.102.111802](https://doi.org/10.1103/PhysRevLett.102.111802). arXiv: [0807.2405 \[hep-ph\]](https://arxiv.org/abs/0807.2405).
- [77] *LHCfaq*. 2017. URL: <https://home.cern/resources/brochure/accelerators/lhc-facts-and-figures>.
- [78] B. Martin and G. Shaw. *Particle Physics*. 3rd ed. John Wiley & Sons.
- [79] S. P. Martin. "A supersymmetry primer". In: *Adv. Ser. Direct. High Energy Phys.* 21 (2010), p. 1. DOI: [10.1142/9789814307505\\_0001](https://doi.org/10.1142/9789814307505_0001). arXiv: [hep-ph/9709356 \[hep-ph\]](https://arxiv.org/abs/hep-ph/9709356).
- [80] P. Meade, N. Seiberg, and D. Shih. "General Gauge Mediation". In: *Prog. Theor. Phys. Suppl.* 177 (2009), p. 143. DOI: [10.1143/PTPS.177.143](https://doi.org/10.1143/PTPS.177.143). arXiv: [0801.3278 \[hep-ph\]](https://arxiv.org/abs/0801.3278).
- [81] R. Meyer and L. W. Feld. "Studies of the misidentification probability of electrons as photons with the CMS experiment in data recorded in 2015". Presented 2016. 2016. URL: <https://cds.cern.ch/record/2239058>.
- [82] H. Nishino et al. "Search for Proton Decay via  $p \rightarrow e^+\pi^0$  and  $p \rightarrow \mu^+\pi^0$  in a Large Water Cherenkov Detector". In: *Phys. Rev. Lett.* 102 (14 2009), p. 141801. DOI: [10.1103/PhysRevLett.102.141801](https://doi.org/10.1103/PhysRevLett.102.141801). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.102.141801>.
- [83] E. Noether. "Invariante Variationsprobleme". ger. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1918 (1918), pp. 235–257. URL: <http://eudml.org/doc/59024>.
- [84] H. D. Politzer. "Reliable Perturbative Results for Strong Interactions?" In: *Phys. Rev. Lett.* 30 (26 1973), pp. 1346–1349. DOI: [10.1103/PhysRevLett.30.1346](https://doi.org/10.1103/PhysRevLett.30.1346). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.30.1346>.
- [85] A. L. Read. "Presentation of search results: the  $CL_s$  technique". In: *J. Phys. G* 28 (2002), p. 2693. DOI: [10.1088/0954-3899/28/10/313](https://doi.org/10.1088/0954-3899/28/10/313).
- [86] K. Rehermann and B. Tweedie. "Efficient identification of boosted semileptonic top quarks at the LHC". In: *JHEP* 03 (2011), p. 059. DOI: [10.1007/JHEP03\(2011\)059](https://doi.org/10.1007/JHEP03(2011)059). arXiv: [1007.2221 \[hep-ph\]](https://arxiv.org/abs/1007.2221).
- [87] C. Rizzi, A. Juste Rozas, and M. P. Casado Lechuga. "Searches for supersymmetric particles in final states with multiple top and bottom quarks with the ATLAS detector". Presented 15 Oct 2018. 2018. URL: [http://cds.cern.ch/record/2646377](https://cds.cern.ch/record/2646377).
- [88] K. Rose. "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems". In: *Proceedings of the IEEE* 86.11 (1998), pp. 2210–2239. ISSN: 0018-9219. DOI: [10.1109/5.726788](https://doi.org/10.1109/5.726788).
- [89] J. T. Ruderman and D. Shih. "General Neutralino NLSPs at the Early LHC". In: *JHEP* 08 (2012), p. 159. DOI: [10.1007/JHEP08\(2012\)159](https://doi.org/10.1007/JHEP08(2012)159). arXiv: [1103.6083 \[hep-ph\]](https://arxiv.org/abs/1103.6083).

- [90] A. Salam and J. C. Ward. “Weak and electromagnetic interactions”. In: *Il Nuovo Cimento* (1955-1965) 11.4 (1959), pp. 568–577. ISSN: 1827-6121. DOI: [10.1007/BF02726525](https://doi.org/10.1007/BF02726525). URL: <https://doi.org/10.1007/BF02726525>.
- [91] A. M. Sirunyan et al. “Particle-flow reconstruction and global event description with the CMS detector”. In: *JINST* 12 (2017), P10003. DOI: [10.1088/1748-0221/12/10/P10003](https://doi.org/10.1088/1748-0221/12/10/P10003). arXiv: [1706.04965 \[physics.ins-det\]](https://arxiv.org/abs/1706.04965).
- [92] A. M. Sirunyan et al. “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at  $\sqrt{s} = 13$  TeV”. In: *JINST* 13.06 (2018), P06015. DOI: [10.1088/1748-0221/13/06/P06015](https://doi.org/10.1088/1748-0221/13/06/P06015). arXiv: [1804.04528 \[physics.ins-det\]](https://arxiv.org/abs/1804.04528).
- [93] A. M. Sirunyan et al. “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”. In: *JINST* 13 (2018), P05011. DOI: [10.1088/1748-0221/13/05/P05011](https://doi.org/10.1088/1748-0221/13/05/P05011). arXiv: [1712.07158 \[physics.ins-det\]](https://arxiv.org/abs/1712.07158).
- [94] A. M. Sirunyan et al. “Search for gauge-mediated supersymmetry in events with at least one photon and missing transverse momentum in pp collisions at  $\sqrt{s} = 13$  TeV”. In: *Phys. Lett. B* 780 (2018), p. 118. DOI: [10.1016/j.physletb.2018.02.045](https://doi.org/10.1016/j.physletb.2018.02.045). arXiv: [1711.08008 \[hep-ex\]](https://arxiv.org/abs/1711.08008).
- [95] A. M. Sirunyan et al. “Search for supersymmetry in events with a photon, a lepton, and missing transverse momentum in proton-proton collisions at  $\sqrt{s} = 13$  TeV”. In: *JHEP* 01 (2019), p. 154. DOI: [10.1007/JHEP01\(2019\)154](https://doi.org/10.1007/JHEP01(2019)154). arXiv: [1812.04066 \[hep-ex\]](https://arxiv.org/abs/1812.04066).
- [96] A. M. Sirunyan et al. “Search for supersymmetry in events with at least one photon, missing transverse momentum, and large transverse event activity in proton-proton collisions at  $\sqrt{s} = 13$  TeV”. In: *JHEP* 12 (2017), p. 142. DOI: [10.1007/JHEP12\(2017\)142](https://doi.org/10.1007/JHEP12(2017)142). arXiv: [1707.06193 \[hep-ex\]](https://arxiv.org/abs/1707.06193).
- [97] T. Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Comput. Phys. Commun.* 191 (2015), p. 159. DOI: [10.1016/j.cpc.2015.01.024](https://doi.org/10.1016/j.cpc.2015.01.024). arXiv: [1410.3012 \[hep-ph\]](https://arxiv.org/abs/1410.3012).
- [98] M. Tanabashi, K. Hagiwara, K. Hikasa, et al. “Review of Particle Physics”. In: *Phys. Rev. D* 98 (3 2018), p. 030001. DOI: [10.1103/PhysRevD.98.030001](https://doi.org/10.1103/PhysRevD.98.030001). URL: <https://link.aps.org/doi/10.1103/PhysRevD.98.030001>.
- [99] S. Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21 1967), pp. 1264–1266. DOI: [10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.19.1264>.