Search for Supersymmetry in proton-proton collisions at $\sqrt{s} = 13$ TeV with at least one photon, jets and missing transverse momentum

विद्या वाचस्पति की
उपाधि की अपेक्षाओं की आंशिक पूर्ति में प्रस्तुत शोध प्रबंध

A thesis submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy

द्वारा / By
भूमिका कंसल / Bhumika Kansal

पंजीकरण सं. / Registration No.: 20173559

शोध प्रबंध पर्यवेक्षक / Thesis Supervisor: Dr. Seema Sharma

CERN-THESIS-2024-062

IISER PUNE
भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान पुणे
INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE
2023

Dedicated to my late grandfather . . .

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

(Bhumika Kansal)
Roll no. : 20173559

Date: 28/10/2023

Certificate

Certified that the work incorporated in the thesis entitled "Search for Supersymmetry in proton-proton collisions at $\sqrt{s}$ = 13 TeV with at least one photon, jets and missing transverse momentum" submitted by Bhumika Kansal was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institution.

Seema Sharma
(Dr. Seema Sharma)

Date: 28/10/2023

Acknowledgements

I would like to acknowledge and express my deepest gratitude to my supervisor, Dr. Seema Sharma. Her guidance, expertise, and unwavering support throughout researching and writing this thesis have been invaluable. Her insightful feedback, constructive criticism, and encouragement have greatly contributed to the development and completion of this work. I am sincerely grateful for her mentorship and for the opportunities provided to enhance my academic and research skills. I am thankful for her dedication and commitment to my academic growth.

I would also like to thank Dr. Monoranjan Guchait and Dr. Sourabh Dube, members of the Research Advisory Committee (RAC), for their continuous feedback and valuable input that significantly enriched the quality of my work. Additionally, I am deeply grateful to my collaborators, Dr. Vinay Hegde, Dr. Aditee Rane, Dr. Andrew Whitbeck and Dr. Kevin Pedro, for their collaboration and enlightening discussions. Special thanks to Dr. Alex Tapper and other members of the analysis review committee (ARC) at the CMS collaboration, for their invaluable feedback and great suggestions.

Furthermore, I acknowledge the support provided by Dr. Shubham Pandey for his unwavering assistance, constant encouragement, and insightful discussions during the PF calibration work. I want to express my appreciation to the CMS, CERN, and Fermilab collaborations for offering access to their high-quality experimental setup and computing resources, which were essential for the successful execution of the project. The support and the resources provided by 'PARAM Brahma Facility' under the National Supercomputing Mission, Government of India at the Indian Institute of Science Education and Research (IISER) Pune are also gratefully acknowledged. The India-CMS collaboration has provided me an excellent platform to present my work and interact with faculty and fellow students, whom I am thankful to enrich my research experience. I want to acknowledge the financial support from the CSIR and DST government funding agencies throughout my PhD.

I would also like to thank my academic institution, IISER Pune, and its Physics department, including the head of the department and faculty members, for creating a conducive research environment and engaging in enriching discussions. Special thanks go out to Mr. Prabhakar, Mrs. Dhanashree, Mr. Tushar, Mrs. Sayalee, Mrs. Sneha and their team for their invaluable assistance with administrative tasks, and to the IT department at IISER, particularly Mrs. Nisha Kurkure and Mr. Goldi Misra, as well as Mr. Brij Kishore Jashal and his team from TIFR, for providing computing resources and support.

I want to express my gratitude to the Experimental High Energy Physics group at IISER, including Ms. Alpana Sirohi, Dr. Angira Rastogi, Mr. Chirayu Gupta, Mr. Chetan Agrawal, Mr. Arnab Laha, Mr. Kumar Yash and Mr. Prachurjya for the wonderful discussions and enjoyable

birthday celebrations. I want to thank Ms. Alpana for her invaluable assistance with advanced machine learning techniques for developing PF calibration work.

I also want to express my appreciation for the previous Ms. students, such as Ms. Shivani Lomte, Mr. Niramay Gogate, Mr. Nitish Kumar, Mr. Dhruvanshu Parmar, and Mr. Rahul Kumar, for their insightful discussions, particularly during our Saturday meetings. My time during PhD at IISER would not have been completed without the support and camaraderie of Dr. Naveen, Mr. Sarvottam, Dr. Vrinda, Mr. Deepak Roy, Dr. Deepak Sharma, Dr. Sumit, Ms. Manisha, Mr. Umashankar, Dr. Gokul, Mr. Anupam, Mr. Pranab, Dr. Shailendra, Dr. Navita, Dr. Ravi, Dr. Imran, Dr. Charu, Dr. Sneha, Ms. Preeti and all my friends. Their constant support and the cherished moments we've shared have been invaluable.

Lastly, I'd like to acknowledge my parents (Madhu Kansal and Pramod Kansal) for providing unwavering support throughout my PhD journey. I am grateful to my brother, Manish Kansal, whose support was indispensable in completing my PhD. My late grandfather, Mr. Ram Mohan Lal Gupta, was a physics lecturer, and he instilled in me a love for physics from my early years. As such, I dedicate my thesis to my grandfather, who has inspired me to pursue a PhD in physics.

Abstract

The Standard Model (SM) explains three out of four fundamental forces except for gravity. One of the primary limitations of SM is a mass hierarchy problem, and the other is the inability to explain the presence of dark matter in the universe. The discovery of the higgs (H) boson in 2012 at the LHC, has also opened up new avenues for research and fueled further investigations into physics beyond the SM. Supersymmetry (SUSY) model can address discrepancies like the hierarchy problem by providing superpartners to SM particles with a different spin by half unit. It also offers a potential candidate for dark matter.

This thesis presents a search for new physics in final states consisting of at least one photon, multiple jets, and large missing transverse momentum, using proton-proton collision events at a center-of-mass energy of 13 TeV. The data correspond to an integrated luminosity of 137 fb^{-1}, recorded by the CMS experiment at the CERN LHC from 2016 to 2018. The events are divided into mutually exclusive bins characterized by the missing transverse momentum, the number of jets, the number of b-tagged jets, and jets consistent with the presence of hadronically decaying W, Z, or H bosons. The observed data are found to be consistent with the prediction from SM processes. The results are interpreted in the context of simplified models of pair production of supersymmetric particles via strong and electroweak interactions. Depending on the details of the signal models, gluinos and squarks of masses up to 2.35 and 1.43 TeV, respectively, and electroweakinos of masses up to 1.23 TeV are excluded at 95% confidence level. These are the best mass limits to date on electroweakino production with photons in the final state.

Contents

Acknowledgements ix

Abstract xi

1 Introduction 1

2 Standard Model of elementary particles and Supersymmetry 7
2.1 Standard Model . . . 7
2.1.1 Leptons . . . 7
2.1.2 Quarks . . . 8
2.1.3 SM bosons . . . 9
2.2 SM interactions in QFT framework . . . 10
2.2.1 Theoretical formulation of Standard model . . . 10
2.3 Limitations of Standard model . . . 14
2.4 Supersymmetry . . . 16
2.4.1 Motivation for Supersymmetry . . . 17
2.4.2 Minimal Supersymmetry model . . . 18
2.4.3 Gauge-Mediated SUSY breaking . . . 20
2.4.4 Simplified GMSB search models . . . 22

3 Experimental setup and event reconstruction 25
3.1 Large Hadron Collider . . . 25
3.1.1 Luminosity . . . 27
3.1.2 The LHC Run II operations . . . 28
3.2 The CMS detector . . . 29
3.2.1 The Solenoid Magnet . . . 29
3.2.2 Tracking system . . . 31
3.2.3 Electromagnetic calorimeter . . . 32
3.2.4 Hadron calorimeter . . . 34
3.2.5 Muon detectors . . . 35
3.3 Triggers . . . 37
3.4 Event reconstruction . . . 39
3.4.1 Tracking . . . 39
3.4.2 Electrons and Photons . . . 41
3.4.3 Muons . . . 43

3.4.4 Jets . . . 44
3.4.5 b-tagged Jets . . . 46
3.4.6 Missing transverse momentum . . . 48

4 Particle Flow charged hadron calibration 51
4.1 Motivation . . . 51
4.2 Technical details for Ultra Legacy simulation samples . . . 52
4.3 Event selection . . . 53
4.4 Calibration methodology . . . 55
4.4.1 Energy dependent calibrations (Step 1 of calibration) . . . 59
4.4.2 Pseudo rapidity dependent calibrations . . . 61
4.5 Results for UL processing of Run2 (2016-2018) data . . . 64
4.5.1 Calibration parameters for UL processing of 2016-2018 dataset . . . 64
4.5.2 Results - response & resolution as a function of E_t . . . 66
4.5.3 Results - response as a function of η . . . 68
4.5.4 Results - 1D-distributions of raw and corrected responses . . . 68
4.6 Effect of ECAL rechit thresholds on PF hadron cluster energy . . . 68
4.7 Machine learning techniques for PF hadron energy calibration . . . 70
4.7.1 DRN model for PF calibration . . . 72
4.7.2 Results . . . 73
4.7.3 Summary . . . 74

5 A search for supersymmetry with at least one photon, large MET and jets final states 79
5.1 Targeted SUSY signal models and SM backgrounds . . . 79
5.2 Standard model backgrounds . . . 81
5.2.1 tt + jets and W + jets process . . . 81
5.2.2 tt + γ and W + γ process . . . 81
5.2.3 Z($\nu\nu$) + γ + jets process . . . 82
5.2.4 QCD and γ + jets process . . . 82
5.3 Monte Carlo simulation of SM processes . . . 83
5.3.1 MADGRAPH5_AMC@NLO Generator . . . 83
5.3.2 PYTHIA Generator . . . 84
5.3.3 GEANT4 . . . 84
5.3.4 Stitching of simulated SM background processes . . . 85
5.4 Triggers and Efficiency . . . 86
5.4.1 Trigger efficiency for signal region . . . 86
5.4.2 Trigger efficiency for control region . . . 88
5.5 Event selection . . . 89
5.5.1 Baseline selections . . . 89
5.5.2 Event yields after baseline selections . . . 93
5.6 Optimisation studies . . . 93
5.6.1 Sensitivity for strong productions . . . 96
5.6.2 Sensitivity for electroweak productions . . . 97

6 Estimation of standard model backgrounds **101**
6.1 Lost lepton background . . . 101
6.1.1 Transfer factor calculation . . . 103
6.1.2 Data-MC correction factors . . . 105
6.1.3 Closure test . . . 105
6.1.4 Data-MC comparisons in control region . . . 106
6.1.5 Systematics . . . 106
6.1.6 Prediction of lost-lepton events . . . 108
6.2 Misidentification of electrons as photons . . . 110
6.2.1 Misidentification rate calculation . . . 111
6.2.2 Closure test . . . 112
6.2.3 Data-MC correction factor . . . 113
6.2.4 Data-MC comparisons in control region . . . 115
6.2.5 Systematics . . . 119
6.2.6 Prediction of misidentified electron background events . . . 121
6.3 Estimation of $Z(\nu\nu)\gamma$ + jets background . . . 121
6.3.1 Data-MC comparisons in control region . . . 123
6.3.2 Validation of SR and CR events . . . 124
6.3.3 Purity factor . . . 124
6.3.4 Systematics . . . 126
6.3.5 Prediction of Z invisible background events . . . 126
6.4 Estimation of $\gamma + jets$ and QCD multijet backgrounds . . . 127
6.4.1 Methodology . . . 127
6.4.2 R factor using data . . . 128
6.4.3 κ factor using simulations . . . 129
6.4.4 Systematics . . . 130
6.4.5 Prediction of $\gamma + jets$ and QCD multijet background events . . . 130
6.5 Total SM prediction and data observed . . . 132

7 Results **135**
7.1 Signal systematics . . . 135
7.2 Methology for statistical inference from data . . . 137
7.2.1 Likelihood function . . . 139
7.2.2 Test statistics . . . 140
7.3 Results . . . 141
7.4 Exclusion Limits . . . 142
7.4.1 Strong production simplified signal models . . . 142
7.4.2 Electroweakino simplified models . . . 143

8 Summary and outlook **147**

A Results for PF charged Hadron calibration **151**
A.1 Technical details and results for re-processing of Run2 (2016-2018) data . . . 151

A.1.1 List of calibration functions and parameters for all years 152
A.2 Technical details for ECAL ZS sample 153
A.3 Technical details for Run3 sample 153

B Additional tables **155**
B.1 Lost-lepton + τ_{had} predictions 155
B.2 Electron faking photon (e $\rightarrow \gamma$) background predictions 158
B.3 Predictions for Irreducible background from Z($\nu\nu$)+γ 161
B.4 Predictions for QCD+γ + *jets* background 163

C Additional systematics for Trigger efficiency **165**

D Additional checks for fake-rate data-MC correction factors **169**
D.1 Validity of $\alpha(N_{\text{b-tags}})$ yearwise 169

E Appendix E : Different branching fractions for signal samples **171**

List of Figures

1.1 Summary of the experimental cross-section measurements of SM processes by the CMS collaboration. These cross-sections are compared with theoretical predictions [22]. . . . 2
1.2 Overview of lower mass limits of SUSY particles from various simplified SUSY models in the context of GMSB [31]. . . . 4

2.1 Figure showing building block of the SM with their quantum numbers [39]. . . . 8
2.2 Figure showing color charge and gluon exchange [Image courtesy: Wikipedia]. . . . 9
2.3 Diagram depicting the top-loop correction to the m_H^2 term. . . . 15
2.4 Evolution of the inverse of the coupling constants as a function of the logarithm of energy scale (Q or μ), illustrating the concept of gauge coupling unification in the SM [26]. . . . 16
2.5 Diagram super-partners (right) of standard particles (left). [Image credit: DESY at Hamburg [49]] . . . 17
2.6 Diagram depicting the stop loop correction to the m_H^2 term. . . . 17
2.7 Gauge coupling unification in the SUSY Model [26]. . . . 18
2.8 Branching fractions for Bino (left) and neutral wino (right) as a function of mass.[55]. . . . 22
2.9 Cross-sections (in NNLO approximations) for SUSY particle production with respect to mass of corresponding particles in pp collision at LHC with center of mass energy, $\sqrt{s} = 13$ TeV [61]. . . . 23
2.10 Simplified model scenarios for gluino pair production in top two rows and bottom two rows represents simplified models for electroweakino (direct chargino/neutralino) pair productions. . . . 24

3.1 Schematic overview of the injection chain, pre-accelerators, the LHC ring and detectors at the interaction points.[63] . . . 26
3.2 The integrated luminosity recorded at the CMS detector for different years during Run II (left) [67]. On the right side, the number of pileup distributions for all years (2015-2018) are overlaid, along with a grey shaded histogram representing the full Run II [68]. . . . 28
3.3 Sectional view of the CMS detector [69]. . . . 30

3.4 Schematic view for the first quadrant of the CMS tracker in r-z coordinates [72]. Layers with green color represents the inner pixel tracker and orange (blue) color means single (double) sided modules represents outer strip tracker. Single sided modules have only reverse p-n junction sensor whereas double sided modules have two silicon sensors mounted back to back [73]. 31
3.5 Hit efficiency over large range in luminosity for phase 0 (during 2016) and phase 1 (during 2018 upgrades) [75]. 32
3.6 A schematic view showing only one quarter of the ECAL detector [70]. 33
3.7 A longitudnal view (r-z plane) showing HCAL detector during the 2016 LHC Run 2 (right) and phase 1 segmentation of the HE (left) which is upgraded during the 2018 LHC Run 2. The numbers on top and on left refer to the tower numbers. The numbers on the bottom (0–17) indicate the scintillator layers numbers inserted into slots in the absorber [76]. 34
3.8 A schematic r-z view of the CMS detector showing muon system for Run 2 in details. It includes DTs (MB), CSCs (ME), and RPCs (RB or RE), which are described by respective colours in the legend. Here M denotes Muon, B stands for Barrel and E for Endcap [77]. 37
3.9 Overview of the CMS L1 trigger system for Run II [80]. 38
3.10 The CMS transverse slice showing interaction of electron, muons, photons, charged and neutral hadrons with various sub-detectors [82]. 39
3.11 This is pictorial representation of the jet energy correction levels.[88] 45
3.12 The Figure illustrates an outline of techniques employed for pileup (PU) suppression. In the figure, solid lines represent charged PF candidates, while dashed lines represent neutral PF candidates. The PUPPI algorithm applies weights to the PF candidates, which are represented by thin lines in the diagram [92]. 46
3.13 Pictorial representation of the secondary vertex in the heavy flavored jet (left) and 2D IP significance distribution of the most displaced track (right) for different jet flavours. [93]. 2D IP/σ plot shows the higher IP value for b-jets. 46
3.14 DeepCSV discriminator for ak4 jets in $t\bar{t}$ sample having different jet flavours (left). The distributions are normalized to unit area. On the right, mis-identification probability for light quarks with respect to b-tag efficiency [93]. 47
3.15 The distributions of p_T^{miss} is shown after applying event filtering algorithms to one set of events (represented by filled markers), while another set of events did not undergo the event filtering algorithms (represented by open markers) [94]. . 49

4.1 Comparisons of PF jet composition measured as a fraction of energy carried by PF photons, charged hadrons, neutral hadrons, and electrons and muons in data and MC simulation for 2016 (left), 2017 (middle), and 2018 (right) [96] 52
4.2 Composition of particle level jets presented as a fraction of energy carried by various generator level particles produced in the process of hadronization of quarks and gluons [95]. 53

4.3 Distribution of raw energy response, i.e. $(E_t - E^{raw}_{total})/E_t$ for the hadrons with E_t=40−60 GeV (left), and average energy response as a function of E_t (right) for the hadrons which have an associated ECAL cluster, i.e. $E^{raw}_e > 0$. 55
4.4 Average raw energy response i.e. $(E_t - E^{raw}_{total})/E_t$ as a function of E_t (top) and $|\eta|$ (bottom) for EH (left) and H (right) hadrons for Ultra legacy conditions corresponsing to 2016, 2017 & 2018. 56
4.5 The CMS longitudinal layout [101]. 58
4.6 Average energy response after applying energy dependent calibrations as a function of E_t for EH hadrons(left) and for H hadrons(right). 61
4.7 Average energy response after applying energy dependent calibrations as a function of $|\eta|$ for EH hadrons(left) and H hadrons(right). 61
4.8 Calibration parameters for EH hadrons in barrel (left column) and endcap regions (right column): a(E_t) (row 1), b(E_t) (row 2), α(E_t) (row 3) and β(E_t) (row 4). 64
4.9 Calibration parameters for H hadrons in barrel (left column) and endcap regions (right column): c(E_t) (row 1), α(E_t) (row 2) and β(E_t) (row 3). 65
4.10 Final response (left column) & resolution (right column) of **EH hadrons** obtained in the bins of E_t shown for barrel (row 1), endcap within tracker (row 2) and endcap outside tracker (row 3) for 2016, 2017, and 2018. The solid and empty symbols correspond to corrected and raw energy measurements respectively. 66
4.11 Final response (left column) & resolution (right column) of **H hadrons** obtained in the bins of E_t shown for barrel (row 1), endcap within tracker (row 2) and endcap outside tracker (row 3) for 2016, 2017, and 2018. The solid and empty symbols correspond to corrected and raw energy measurements respectively. . . 67
4.12 Final energy response shown in the bins of $|\eta|$ (inclusive in E_t) for EH hadrons (top) and H hadrons (bottom) 68
4.13 1st (3rd) and 2nd (4th) columns represent 1D energy response distributions before and after final calibrations for EH (H) hadrons respectively. Top, middle and bottom rows represent the barrel, endcap within tracker and endcap outside tracker regions. 69
4.14 1D Distribution of corrected energy response, i.e. $(E_t - E^{raw}_{total})/E_t$ of the EH hadrons in the endcap within tracker region (top row) and outside tracker region (bottom row) for "10_0_2" (red) and "10_0_3" (black) samples. 1D energy response distributions in first column represent EH hadrons having E_t = 44-54 GeV and similarly, second column represent E_t = 124-144 GeV. 70
4.15 Comparison of the 1D distributions of the corrected energy response, where "10_0_3" and "10_0_2" calibrations are applied to the "10_0_3" sample. The top-left plot shows the raw energy response distribution, while the top-right plot illustrates the response and resolution as functions of the true energy (E_t) of the EH hadrons. The bottom plot shows the response as a function of $|\eta|$. 71
4.16 A schematic diagram for dynamic reduction network model. [104] 71

4.17 Rechits in x-y (top-left) and x-z (top-right) plane are shown. In bottom plots, 2D distributions of E^{raw}_{total} (left)and E_t (right) with respect to E^{rechit}_{total} are shown. 72
4.18 1D distributions of E_t (left) and weights with respect to E_t (right) are shown. . . 73
4.19 Validation and Training Loss are overlaid with respect to epochs. 73
4.20 1D energy response distribution after calibrations using χ^2 (red), DRN (blue & green) calibrations for EH hadrons(top) and for H hadrons(bottom). These energy responses are in the E_t range of 68–72 GeV for various η regions. 74
4.21 Average energy response (top) and resolution (bottom) after using χ^2 (red), DRN (blue & green) calibration methods, as a function of E_t for EH hadrons. After red dashed line at E_t = 250 GeV, there is an edge effect and that can be ignored for now. 75
4.22 Average energy response (top) and resolution (bottom) after using χ^2 (red), DRN (blue & green) calibration methods, as a function of E_t for EH hadrons. After red dashed line at E_t = 250 GeV, there is an edge effect and that can be ignored for now. 75
4.23 1D-energy response distributions of EH hadrons (on top) for different η bins and inclusive in E_t. Average energy response (bottom) after using χ^2 (red), DRN (blue & green) calibration methods, as a function of $|\eta|$ for EH hadrons. 76
4.24 1D-energy response distributions of H hadrons (on top) for different η bins and inclusive in E_t. Average energy response (bottom) after using χ^2 (red), DRN (blue & green) calibration methods, as a function of $|\eta|$ for H hadrons. 76
4.25 Energy response with respect to E_t and η for EH (left) & H (right) hadrons. 77

5.1 Simplified model scenarios for gluino pair production, T5qqqqHG (top left), T5bbbbZG (top right), T5ttttZG (lower left), and stop pair production T6ttZG (lower right). The models are defined in the text. 80
5.2 Simplified model scenarios for electroweakinos pair production, TChiWG (left), and TChiNG (right). The models are defined in the text. 81
5.3 Standard model background processes having at least one γ, p_T^{miss} and hadronic jets as a final state. 83
5.4 The distribution of minimum $\Delta R(l, \gamma_{reco})$ and $\Delta R(q/g, \gamma_{reco})$ for stitching W + jets and W + γ processes. 86
5.5 MET X MHT trigger efficiency, ϵ_{MET} with respect to p_T^{miss} using single electron dataset for years 2016, 2017, 2018. 87
5.6 The photon trigger efficiencies (ϵ) wrt transverse momenta of photon (p_T^{γ}) (left) and pseudorapidity of photon (η). Green circles shows kinks in trigger efficiencies for both left and right figures. 87
5.7 MET X MHT trigger efficiency in the control region of the QCD multijet process, ϵ_{MET} with respect to p_T^{miss} using JetHT dataset for years 2016, 2017, 2018 and their comparison. 88

5.8 Simulation based SM background contributions in total number of Ak4 jets (left) and soft drop mass for the leading AK8 jet (right) is shown in the top panel. Some signal points corresponding to T5bbbbZg, TChiWG andTChiNG in bottom panel are overlaid on the top of the total background (in shaded). 95
5.9 The definitions and indexing schemes for the SP (left) and EW (right) SRs and low-p_T^{miss} CRs, in the planes of p_T^{miss} , N_{jets} , and $N_{\text{b-tags}}$ (left) and p_T^{miss} , V-tag, and H-tag (right). The gray blocks correspond to the low-p_T^{miss} CRs, the blue blocks to the SP SRs, and the red blocks to the EW SRs. 96
5.10 Simulation based SM background contributions in various search bins is shown in the top panel. Some distributions for some signal points corresponding to T5bbbbZg (in middle panel) and T5qqqqHg (in bottom panel) are overlaid on the top of the total background (in shaded). The numerical values in parentheses in the legend entries for the signal models indicate the gluino and NLSP mass values for strong production. 96
5.11 Simulation based optimisation checks using exclusion curves for the strong production models, T5bbbbZg (left) and T5qqqqHg (right). 97
5.12 Simulation based SM background contributions in various search bins is shown in the top panel. Some distributions for some signal points corresponding to TChiWG (in middle panel) and TChiNG (in bottom panel) are overlaid on the top of the total background (in shaded). In middle and bottom panel, signal distributions corresponding to wino-bino electroweakino pair production are shown. The numerical values in parentheses in the legend entries for the signal models indicate the NLSP mass value for electroweak production. 98
5.13 Simulation based signal acceptance x efficiency for the electroweakino pair production models, TChiWG (on the top) and TChiNG (on the bottom). 99

6.1 Contribution of lost electron (left) and lost muon (right) in SR due to failing acceptance, failing identification and isolation in N_{jets}, N_{b-tag}, V-tag & H-tag bins. On the right, hadronic τ decay events are combined with lost muon SR. In both plots, solid grey colour represents the corresponding CR region. 104
6.2 Relative contributions of events with light lepton(s) or τ_h candidate(s) in the SRs and CRs (upper panel), and the corresponding TFs, along with their statistical uncertainties (lower panel). 104
6.3 Validation of TFs for lost-lepton estimation in the 45 optimised bins. The TF[N_{jets}, N_{b-tag}, W-tag, H-tag] are applied on the MC event yields in CR (in red) & compared with expected MC yields in SR (in blue). 106
6.4 Data-MC comparisons in CR i.e. single lepton(e/μ) + 1γ for lost-lepton estimation. Statistical uncertainty for simulation samples are shown a shaded region in the bottom panel of each plot. 107

6.5 Comparison of nominal TF distrution (in red) and TF after varing cross-section of $t\bar{t}$ /W+ γ by 20% (in blue) in the top panel. Shaded region represents the statistical uncertainty in the bottom panel and black points are showing relativistic differences in both TFs. . . . 107
6.6 Envelope of 100 PDFs (left) and 8 scale (right) uncertainties with respect to the each TF bin. Shaded region represents the statistical uncertainty. . . . 108
6.7 Comparison of nominal TF distrution (in blue) with TF after varying the JEC factors (left) and JER (right) by +1 σ_{sys} is shown in the top panel. Shaded region represents the statistical uncertainty in the bottom panel and black points are showing relativistic differences in TFs. . . . 108
6.8 Final data prediction with total systematic uncertainty as a shaded region and bar line over each bin content represents the statistical uncertainty i.e. zero lepton(e/μ) + atleast-1γ for lost-lepton estimation. . . . 109
6.9 Feynman diagram for W+jets process in which e is misidentified as photon, p_T^{miss} is originating from ν and jets from QCD. . . . 110
6.10 Distribution of the transverse mass between electron and p_T^{miss} for single electron CR in simulation. The histograms for background processes are stacked, and distributions for some signal points corresponding to the strong productions are overlaid. . . . 110
6.11 Kinematic distributions of the electron p_T (left) and Q_{mult} (right) for control and search region in simulation. MC statistical errors are represented by bar on top of each histogram. . . . 112
6.12 A comparison of the number of events with an electron misidentified as a photon in the SRs and the number estimated using the single-electron CRs with simulated samples. The ratio of the expected and predicted event yields in each SR is shown in the lower panel. The shaded region in the lower panel indicates the systematic uncertainties in the predicted number of background events. . . . 113
6.13 Purity factor distributions, $\beta(e,\gamma)$ and $\beta(e,e)$ for data SR and CR respectively in the top panel (left). Misidentification rate distribution in $N_{\text{b-tags}}$ bins for ee data and simulation using full Run II dataset in the top panel (right). The data-MC correction factor, α is shown in the bottom panel (right). . . . 115
6.14 In the left column: f(data) and f(MC) distributions of events containing e^+e^- pairs using data and simulation respectively in the parameterization of different variables. The data-MC correction factor, α' is shown in the bottom panel of the both plots. In the right column: the ratio of f(data) and f(MC) (after reweighted by α) are shown, and error bars are representing α statistical errors which are propagated to given bins. The shaded region in the ratio panel is the systematics related to α. . . . 117
6.15 The comparison of $(1e, 0\gamma)$ event yield from data against expected yield from MC as a function of N_{jets} , p_T^{miss} , p_T^e and $N_{\text{b-tags}}$. Statistic uncertainty for simulation samples are shown a shaded region in the bottom panel of each plot. . . . 118

6.16 Comparison of nominal TF distrution (in blue) with TF after varying the JEC factors (left) and JER (right) by +1 σ_{sys} is shown in the top panel. Shaded region represents the statistical uncertainty in the bottom panel and black points are showing relativistic differences in TFs. 119
6.17 The comparison of W/$t\bar{t}$ like selections and tag and probe selections which are applied on $Z \to ee$ events. 120
6.18 Comparison of f for DY ($Z \to e^+e^-$) & $t\bar{t}$ (left) in the bins of $N_{\text{b-tags}}$, N_{jets} from SP region and V- or H-tag from EW region (bins are inclusive in p_T^{miss}) after using W/$t\bar{t}$ like selections. On the right, there is a comparison of f for DY ($Z \to e^+e^-$) & W+jets (right) in the same bins. 120
6.19 Comparison of α for DY (left) which is calculated using W/$t\bar{t}$ like selections and tag and probe selections in the bins of $N_{\text{b-tags}}$. There is a comparison of f for DY & W+jets (right) in the same bins. 121
6.20 Final data prediction with total systematic uncertainty as a shaded region and bar line over each bin content represents the statistical uncertainty i.e. zero electron + atleast-1γ for misidentified electron estimation. 122
6.21 Feynman diagram for $Z(\nu\nu)\gamma +$ jets process (left) in which photons are coming from radiations, p_T^{miss} is coming from ν and jets from QCD. $Z(\ell^+\ell^-)\gamma +$ jets process (right) is shown in which Z is decaying in to same flavour and opposite charged lepton flavours. 122
6.22 The distributions of the dilepton invariant mass (left) and the magnitude of the dilepton $\vec{p}_T$ plus the $\vec{p}_T^{\text{miss}}$ (right) for $Z(\ell^+\ell^-)\gamma +$ jets events in data and simulation. In the lower panel, the shaded region shows the statistical uncertainty in the simulation. 123
6.23 The distributions of the $Z(\nu\nu)\gamma +$ jets (SR) & $Z(\ell^+\ell^-)\gamma +$ jets (CR) in the bins of p_T^{miss} , N_{jets} , p_T^{γ} and $N_{\text{b-tags}}$. For top-left figure, p_T^{miss} is calculated as the magnitude of the dilepton $\vec{p}_T$ plus the $\vec{p}_T^{\text{miss}}$ (right) for $Z(\ell^+\ell^-)\gamma +$ jets events in simulation. 124
6.24 Purity, $\beta_{\ell\ell\gamma}$, of $\ell\ell + \gamma$ sample for 0 b-tag and > 0 b-tag regions obtained using different methods. 125
6.25 In left column, comparison of $N_{\text{b-tags}}$ distributions for $\ell\ell\gamma$ events in data and simulations both, and for simulations, both DY and non-DY processes are included. In right column, data distributions are reweighted by $\beta_{\ell\ell\gamma}$, which eliminate the non-DY contribution and hence, in simulations only DY process is shown. 126
6.26 Final data prediction with total systematic uncertainty as a shaded region and bar line over each bin content represents the statistical uncertainty i.e. zero electron + atleast-1γ for misidentified electron estimation. 127
6.27 Feynman diagram for QCD multijet and $\gamma + jets$ process (left) in which photons are coming from initial state radiations and jets from QCD. A schematic 2D diagram of ABCD method (right) is shown with p_T^{miss} in x-axis and $\Delta\Phi(p_T^{\text{miss}}, j_1 \,\&\, j_2)|_{min}$ in y-axis. 128

6.28 The event yield in MET sideband regions (A,C) (top panel) and corresponding R-factor distribution (bottom panel) as a function of [N_{jets} , $N_{\text{b-tags}}$, V-tag, H-tag] using data. 129
6.29 Left: R-factors calculated using simulation in low-p_T^{miss} (red) and high-p_T^{miss} (blue) region. The ratio panel shows the κ-factor in [N_{jets} , $N_{\text{b-tags}}$, V-tag, H-tag] bins. Right: κ(MC) re-scaling $\gamma + jets$ contributions (in red) and without rescaling (in blue) 129
6.30 Data-MC comparisons of p_T^{miss} , proxy γ p_T , N_{jets} & $N_{\text{b-tags}}$ in zero photon validation region. 131
6.31 Left: Distributions for K-factors (κ) is calculated using simulation (red) and data (blue) in zero-photon region. The ratio panel shows the K-factor systematics in [N_{jets} , $N_{\text{b-tags}}$, V-tag, H-tag] bins. Right: κ(MC) distributions for zero-gamma validation region and atleast one photon region using simulations. 131
6.32 Final data prediction with total systematic uncertainty as a shaded region and bar line over each bin content represents the statistical uncertainty for $\gamma + jets$ and QCD multijet background estimation. 132
6.33 Top panel: final data prediction with total systematic uncertainty as a shaded region and bar line over each bin content represents the statistical uncertainty for total SM background estimation. Bottom panel: ratio of the observed number of data events and the predicted backgrounds. The error bars represent the statistical uncertainty in the data events, and the shaded band represents the statistical and systematic uncertainties in the predicted background. The p_T^{miss} bin 200–300 GeV is used for the estimation of the $\gamma + jets$ and QCD multijet background. . . 133

7.1 b tagging systematic uncertainty distributions for strong (on the left) and electroweakino pair (on the rigt) production signal processes in 45 optimised bins (including 27 SR SP , 10 SR EW and 8 low-p_T^{miss} CR). The shaded region represents the statistical uncertainties for the middle mass points in both plots. 136
7.2 ISR systematic uncertainty distributions for strong (left) and electroweakino pair (right) production signal processes in 45 optimised bins (including 27 SR SP , 10 SR EW and 8 low-p_T^{miss} CR). The shaded region represents the statistical uncertainties for the middle mass points in both plots. 137
7.3 Distributions of the test statistics (left) and the normal distribution (right) [126]. . 141

7.4 The numbers of predicted background events and observed events in the SRs and low-p_T^{miss} CRs. Predicted lost lepton, fake photon, $Z(\nu\nu)\gamma$ + jets and γ + *jets* +QCD are stacked histograms. The observed numbers of events in data are presented as black points. For illustration, the expected event yields are presented for the signal model T5bbbbZG, for small (blue) and large (purple) differences in the masses of the $\tilde{g}$ and NLSP. Also shown is the expected distribution of events for the signal model TChiWG (red). The numerical values in parentheses in the legend entries for the signal models indicate the $\tilde{g}$ and NLSP mass values in GeV for strong production and the NLSP mass value for electroweak production. The lower panel shows the ratio of the observed number of data events and the predicted backgrounds. The error bars represent the statistical uncertainty in the data events, and the shaded band represents the statistical and systematic uncertainties in the predicted background. The p_T^{miss} bin 200–300 GeV is used for the estimation of the γ + *jets* and QCD multijet background. . . 142

7.5 Comparison of expected and observed exclusion for T5bbbbZG (top left), T5qqqqHG(top right), T5ttttZG (bottom left) & T6ttZG (bottom right) model at 95% confidence level reached by current search. 143

7.6 The expected and observed limits on the electroweakino mass in the TChiWG (upper left), TChiNG (lower left), and TChiNGnn (lower right) models at 95% CL. For the TChiNG model (lower left), scenarios with degenerate charginos and neutralinos leading to the combined process $\tilde{\chi}_1^{\pm}\tilde{\chi}_1^{\mp}$ +$\tilde{\chi}_1^0\tilde{\chi}_2^0$ +($\tilde{\chi}_1^0$ / $\tilde{\chi}_2^0$)$\tilde{\chi}_1^{\pm}$ (red) or the single process $\tilde{\chi}_1^0\tilde{\chi}_2^0$ (blue) are considered. 145

8.1 Overview of recent bounds on lower mass limits of simplified SUSY models with different final states in the context of GMSB.[144] 149

C.1 Comparison of MET trigger efficiency from data and parameterization function . 166

C.2 MET trigger efficiency for different N_{b-tags} & N_{jets} combinations in same MET SR bins using MC bacground samples as a sum of W/tt/Z (source of true MET) only but not QCD & GJets (Fake MET). 166

C.3 MET trigger efficiency for different N_{b-tags} & N_{jets} combinations in same MET SR bins using MC bacground samples as a sum of W/tt/Z (source of true MET) + QCD & GJets (Fake MET). 167

D.1 2016: f(data) and f(MC) (after re-weighted by α) distributions of events containing e^+e^- pairs using data and simulation respectively in the parameterization of p_T (on the top-left), Q_{multi} (on the top-right), N_{jets} (on the bottom-left) & invariant M_Z. In the bottom panel, the ratio f(data) and re-weighted f(MC) are shown, and the shaded region is the systematics related to α. 169

D.2 2017: f(data) and f(MC) (after re-weighted by α) distributions of events containing e^+e^- pairs using data and simulation respectively in the parameterization of p_T (on the top-left), Q_{multi} (on the top-right), N_{jets} (on the bottom-left) & invariant M_Z. In the bottom panel, the ratio f(data) and re-weighted f(MC) are shown, and the shaded region is the systematics related to α. 170
D.3 2018: f(data) and f(MC) (after re-weighted by α) distributions of events containing e^+e^- pairs using data and simulation respectively in the parameterization of p_T (on the top-left), Q_{multi} (on the top-right), N_{jets} (on the bottom-left) & invariant M_Z. In the bottom panel, the ratio f(data) and re-weighted f(MC) are shown, and the shaded region is the systematics related to α. 170

E.1 Expected exclusions for different branching fractions of T5bbbbZg model at 95% confidence level reached by current search. Exclusion curve (in black color) represents mass limits for current scenario, when NLSP $\rightarrow \tilde{G} + \gamma$ or $\tilde{G} + Z$ with 50% BR. Blue curve represents the case when NLSP $\rightarrow \tilde{G} + \gamma$ with 25% BR and magenta curve represents the case when NLSP $\rightarrow \tilde{G} + \gamma$ with 100% BR 171
E.2 Expected exclusions for different branching fractions of TChiNG model at 95% confidence level reached by current search. Blue curve represents the theoretical cross-sections for pp$\rightarrow \tilde{\chi}_1^0\tilde{\chi}_2^0$. Expected limits for $\tilde{\chi}_2^0 \rightarrow \tilde{G} + \gamma$ with 50% BR and $\tilde{\chi}_1^0 \rightarrow \tilde{G} + H\ (or\ Z)$ with 25% (25%) BR is represented as black dotted curve. Expected limits for $\tilde{\chi}_2^0 \rightarrow \tilde{G} + \gamma$ with 100% BR and $\tilde{\chi}_1^0 \rightarrow \tilde{G} + H\ (or\ Z)$ with 50% (50%) BR is represented as magenta dotted curve 172

List of Tables

2.1 Table showing six quarks with their mass, charge and spin. . . . 9
2.2 Table showing details of four fundamental forces in nature. . . . 10
2.3 SUSY particles in MSSM [25]. . . . 19

3.1 Pre-Accelerators for the LHC.[65] . . . 27
3.2 Seeding configuration and targeted tracks of the ten tracking iterations for 2016 Run. In the last column, R is the targeted distance between the track production position and the beam axis [81]. . . . 40
3.3 Cut-based photon identification requirements for the loose working point in the EB and EE. [84] . . . 43
3.4 Cut-based electron identification requirements for the veto working point in the EB and EE. [85] . . . 43

4.1 Detector changes in 2016, 2017 & 2018 [99]. . . . 52
4.2 Eta dependent tracker hits selection used in event selection . . . 54
4.3 Different regions used to obtain various calibration parameters . . . 59
4.4 Definitions of $f(|\eta|)$ used in 2016, 2017 and 2018 . . . 62

5.1 Decay channels of W boson with branching fractions [109]. . . . 82
5.2 Z Boson Decays and Branching Fractions . . . 82
5.3 Available $p_T^{\text{miss}} \times$ MHT triggers and single photon triggers . . . 86
5.4 Available Jet and HT triggers . . . 88
5.5 Summary of the baseline selection criteria used to identify events of interest for this search. . . . 90
5.6 Event yields for all backgrounds after baseline selections using full Run 2 dataset (corresponding to 137 fb^{-1}). . . . 94
5.7 Event yields for T5bbbbZg andTChiWG after baseline selections using full Run 2 dataset (corresponding to 137 fb^{-1}) . . . 94

6.1 Event selection for SR and CR regions of lost electron and lost muon + τ_h estimation. . . . 103
6.2 Event selection for SR and CR regions of fake rate estimation . . . 111
6.3 Detailed object selection for tag and probe in data and MC. In both data & MC, invariant mass between tag and probe should be within 80–100 GeV. . . . 114
6.4 Data-MC correction factors (α) obtained by applying T & P method on Drell-Yan events from data and MC for entire Run II . . . 114

6.5 Event selection for SR and CR regions of Z invisible estimation 123

7.1 Systematic uncertainties for strong production signal MC samples 138
7.2 Systematic uncertainties for electroweak signal MC samples 138
7.3 Systematic uncertainties for background and signal samples (in %). A dash (-) indicates that the source of uncertainty is not applicable or too small. When needed a description of the source is shown in the parenthesis. 139
7.4 Background predictions and observed event yields for SR region (p_T^{miss} >300 GeV). Events yields in low-p_T^{miss} bins (200-300 GeV) are the sidebands used for multijet predictions in other bins. For the background predictions, the uncertainty is stat + sys (added in quadrature). 146

A.1 Dataset of single pion sample for UL 2016, UL 2017 & UL 2018 151
A.2 Technical details used for the production of UL samples. 152
A.3 Calibration functions and parameters after energy dependent calibrations of EH hadrons for UL 2016, UL 2017 & UL 2018 152
A.4 Calibration functions and parameters after energy dependent calibrations of H hadrons for UL 2016, UL 2017 & UL 2018 152
A.5 Calibration functions and parameters after eta dependent calibrations of EH hadrons for UL 2016, UL 2017 & UL 2018 153
A.6 Calibration functions and parameters after eta-dependent calibrations of H hadrons for UL 2016, UL 2017 & UL 2018 153
A.7 Technical details used for the production of 10_0_2 or 10_0_3 sample. 153
A.8 Technical details used for the production of Run3 sample. 154

B.1 Observed events yields in CR, average transfer factors (TFs) and predictions for lost-lepton background in high $\delta\phi$ region 156
B.2 Observed events yields in CR, average transfer factors (TFs) and predictions for lost-lepton background in low $\delta\phi$ region 157
B.3 Fake-rate parameterization derived from W/$t\bar{t}$ simulations for years 2016, 2017, 2018 and for full Run 2. Note that by convention, upper bin boundary value does not belong to the same bin. 158
B.4 Observed single electron yields, the average scale factor due to MC-based fake rates and corresponding fake-rate predictions with statistical uncertainties for high-$\delta\phi$ search regions. 159
B.5 Observed single electron yields, the average scale factor due to MC- based fake rates and corresponding fake-rate predictions with statistical uncertainties for low-$\delta\phi$ control regions. 160
B.6 Z($\nu\nu$) + γ MC yields and prediction for the signal regions. 161
B.7 Z($\nu\nu$) + γ MC yields and prediction for low-$\delta\phi$ control regions. 162
B.8 $\gamma + jets$ + QCD background prediction in high–$\Delta\phi$ & high–MET(≥300 GeV) search region using low–$\Delta\phi$ and high–$\Delta\phi$ & low–MET (200–300 GeV) control region. 164

Chapter 1

Introduction

Particle physics is a fascinating field that seeks to understand the fundamental building blocks of the universe and the forces that govern their interactions [1]. The quest to unravel the secrets of the subatomic world began in the early 20^{th} century. Pioneering experiments and theoretical insights led to the understanding of constituents & structure of atoms, namely electrons, protons, and neutrons [2, 3]. The advent of quantum mechanics and special relativity further revolutionized our understanding of the microscopic realm [4, 5]. Over the century, groundbreaking discoveries have reshaped our understanding of the subatomic world. At the heart of this exploration lies the evolution of Standard Model (SM), a powerful theoretical framework that has successfully described the behaviour of elementary particles and their interactions with remarkable precision.

In the 1950s and 1960s, technological breakthroughs in accelerators and particle detection systems allowed physicists to probe deeper into the subatomic length scales. This led to the discovery of a multitude of new particles, such as kaons (K), sigma (Σ), omega (Ω), etc. raising questions about their nature and interactions [6, 7]. The mid-1970s brought astonishing revelation in the form of quarks, elementary particles that constitute protons, neutrons, and other hadrons. Murray Gell-Mann and George Zweig independently proposed the concept of quarks as fundamental constituents [8, 9]. This discovery led to the development of quantum chromodynamics (QCD), describing the strong nuclear force and the interactions of quarks and gluons [10, 11, 12]. In 1983, the W and Z bosons were experimentally confirmed through heroic efforts by colliding $p\bar{p}$ in super proton synchrotron (SPS) at the CERN, the European organization for nuclear research [13, 14]. The discovery of W and Z bosons supported the electroweak theory and provided crucial support for the SM [15, 16, 17]. The discovery of the top quark by the CDF (Collider Detector at Fermilab) and D0 (DZero) collaborations at Tevatron ($p\bar{p}$ collider), Fermi National Accelerator Laboratory (Fermilab) was announced on March 2, 1995 [18]. The Large Hadron Collider (LHC) at the CERN collides high energy protons. Two multi-purpose detectors are used to study the outcome of collisions: Compact Muon Solenoid (CMS) and A toroidal LHC apparatus (ATLAS) detectors. The Higgs (H) boson, discovered in 2012 by the ATLAS [19] and CMS [20] experimental collaborations, plays a crucial role in providing mass to elementary particles through its interactions. With these discoveries, the predictions of the SM have been validated, making it one of the most successful theoretical frameworks in physics [21]. Figure 1.1 provides a summary of cross-sections of various

standard model processes measured by the CMS experiment compared with their theoretical predictions [22].

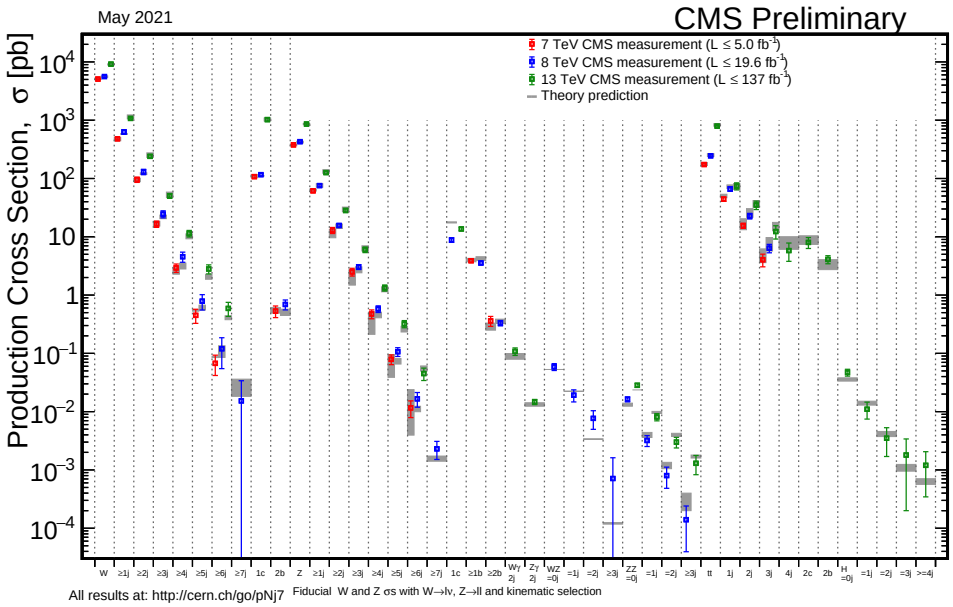

FIGURE 1.1: Summary of the experimental cross-section measurements of SM processes by the CMS collaboration. These cross-sections are compared with theoretical predictions [22].

The SM of particle physics successfully describes interactions among elementary particles via three out of the four fundamental forces: the strong nuclear, electromagnetic, and weak nuclear forces. However, it does not incorporate gravitational interactions which are known to be governed by the principles of general relativity [23], and the absence of a consistent quantum theory of gravity remains one of the major open questions in physics. Moreover, the SM does not provide an explanation for the origin of dark matter as established by astrophysical observations [24]. Another limitation of the SM is the mass hierarchy problem. The experimentally measured mass of the H boson is relatively light, approximately 125 GeV [20]. However, the theoretical calculations in the SM framework predict divergent mass of the H boson due to the contributions of quantum corrections. These large quantum loop corrections to the H boson mass require extremely precise cancellations to explain the observed mass of the H boson, which suggests presence of new physics beyond the SM [25]. Theoretical frameworks such as supersymmetry (SUSY) [25], extra dimensions, extended fermions sectors, Grand Unification Theory (GUTs) [26] and composite higgs models [27] are proposed to address some of these drawbacks of the SM and provide a more natural explanation for the H boson mass. Hence, the ongoing research efforts are focused on exploring new experimental frontiers and developing theoretical models that extend beyond the scope of the SM. This pursuit aims to unveil the mysteries of the universe and deepen our understanding of its fundamental constituents along with physical laws which govern their interactions.

SUSY is one of the most studied extensions of the SM. It predicts superpartners (known as sparticles) of each SM particle having the same quantum number with a difference in spin by half unit. Contributions of these superpartners to the quantum loop correction to the H boson

mass cancel the divergent contributions of the SM particles, and reduce the need for excessive fine tuning of the theory. In R-Parity conserving minimal softly broken SUSY models (MSSM), the lightest supersymmetric particle (LSP) is stable. If the LSP is neutral and weakly interacting, it is a potential dark matter candidate. The lightest neutralino is generally assumed to be the LSP in MSSM models. In Gauge Mediated SUSY breaking (GMSB) models, the gravitino is predicted to be the LSP. The next to the lightest supersymmetric particle (NLSP) then decays into final states containing SM particles and a gravitino. The gravitino (spin $\frac{3}{2}$) is superpartner of the hypothetic graviton (spin 2), the mediator of gravitational interactions, and remains undetected, leading to a signature of the missing transverse momentum (p_T^{miss}) in the reconstructed final states.

If the SUSY particles exist in nature (of up to a TeV mass range), these could be produced at the LHC which collides proton-proton bunches at a center-of-mass energy ($\sqrt{s}$) of 13 TeV. The SUSY production via strong interactions involves the production of gluino or squark pairs as they also carry color charges similar to their SM partners. In electroweakino productions, pairs of charginos and neutralinos are produced. These charginos are a mixture of a charged winos & charged higgsinos, whereas neutralinos are a mixture of binos neutral winos and neutral higgsinos. The SUSY search conducted at the ATLAS and CMS experiments employs a diverse array of final states to comprehensively explore potential signals of new physics. These final states include combinations like leptons + jets + p_T^{miss} , photons + p_T^{miss} , and photons + jets + p_T^{miss} , each exhibiting distinct background compositions arising from the SM processes. This strategic approach effectively addresses various systematic uncertainties, enabling the experiments to achieve sensitivity across a range of SUSY production mechanisms and decay modes, encompassing both strong and electroweak interactions. This approach also ensures that potential signals of new physics are not overlooked and analyzing the data in this manner is crucial to distinguish any potential SUSY signals from the more abundant & well-understood SM background processes. Since the LHC is a hadron collider, copious jet production is expected, leading to more backgrounds in 0-lepton final states. On the other hand, leptonic final states have lower branching ratios but offer cleaner signatures. For example, the leptonic + photons + p_T^{miss} final state provides good sensitivity to the search for sparticles produced via electroweak interactions [28, 29] as it benefits from cleaner backgrounds. Similarly, photons + jets + p_T^{miss} final states offer good sensitivity for strong productions [30] by identifying jets from b quarks. The production cross-section of SUSY particles produced via electroweak interactions is very low and therefore, it is difficult to have good sensitivity for the signals of electroweakinos in the data.

Experimental searches for supersymmetric particles primarily aim to search new physics. However as these sparticles have not been discovered yet, so we generally tend to put lower bounds on the masses of sparticles. Currently, the lower mass bound is in the order of few TeV which might be due to the lower production cross-section of the sparticles, collider constraints (like center-of-mass energy), detector conditions (like its granularity or resolution), etc. Figure 1.2 presents the current bounds on lower mass limits of SUSY particles in various simplified models within the framework of GMSB as of the year 2019.

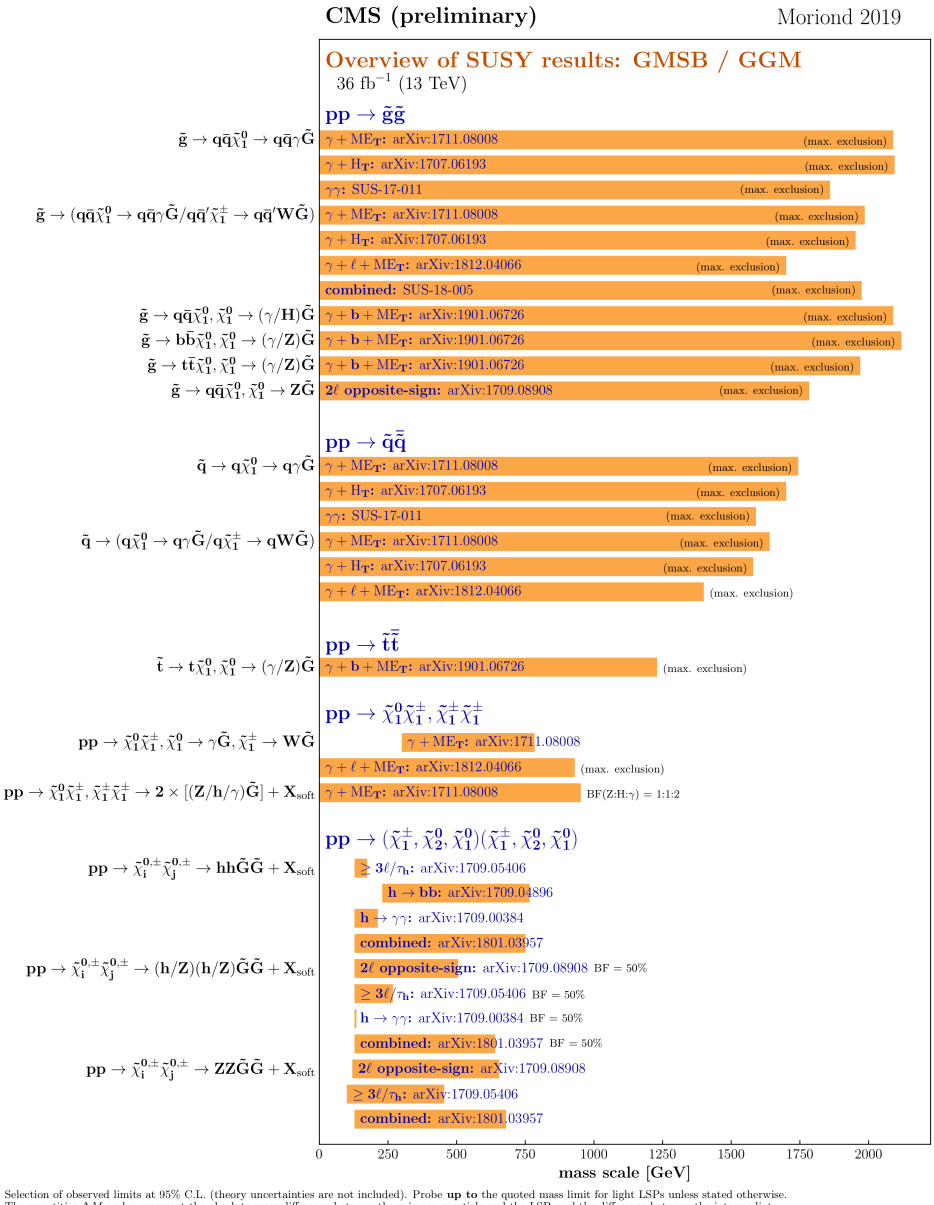

FIGURE 1.2: Overview of lower mass limits of SUSY particles from various simplified SUSY models in the context of GMSB [31].

In this thesis, we are investigating the search for SUSY particles produced via strong and electroweak interactions in the multijets, photon and large p_T^{miss} final states using the data collected at the CMS detector. We specifically focus on GMSB model scenarios with R parity conserved, leading to pair production of SUSY particles. Similar searches have also been reported by the ATLAS Collaboration based on data corresponding to integrated luminosities of 36 fb^{-1} [32, 33] and 139 fb^{-1} [34]. Previous searches by the CMS Collaboration for signatures of squark, gluino, and electroweakino production using final states containing photons with data corresponding to an integrated luminosity of 36 fb^{-1} at $\sqrt{s} = 13$ TeV are documented in refs. [30, 35, 36]. In the present analysis, the search regions are reoptimized based on the availability of four times more data, around 137 fb^{-1}, which can improve the sensitivity for both strong as well as electroweak productions. In addition, using boosted SM boson (W/Z/H) activity to define separate optimised search regions can also increase the sensitivity to analysis for electroweakino productions. This SUSY analysis (described in the chapters 5-7) aims to provide good sensitivity to strong production models as well as electroweakino pair production models, covering maximum phase space for GMSB models. For this, LHC is used as it can produce a wide range of SUSY particles by colliding proton-proton bunches at center-of-mass energy of 13 TeV.

In chapter 2 of this thesis, the theoretical foundation of SM, its limitation and SUSY extention will be explained in details. In chapter 3, we will give the detailed overview of the LHC and CMS experimental setup, object reconstructions from the data collected in the CMS detector, and the particle flow (PF) hadron calibration which is a part of the PF algorithm (used in object reconstruction at the CMS detector) will be explained in chapter 4. We will also discuss our studies and work done on calibrating the PF charge hadron candidates which will be useful for reconstructing jets and very essential element for the analysis. In chapter 5, fundamental details of the SUSY analysis, including targeted simplified models, selections, and optimization studies, will be discussed. Subsequently in chapter 6, we will explain all the background estimations in details. In chapter 7 we will summarize the results and conclusions of this analysis. Finally, summary and the future outlook of the thesis will be discussed in chapter 8. Each chapter thus contributes a distinct layer to the overall narrative, culminating in a cohesive exploration of particle physics theory, experimental methodology, meticulous analysis, and the consequential outcomes.

Chapter 2

Standard Model of elementary particles and Supersymmetry

This chapter initially outlines theoretical framework for describing interactions of the elementary particles, which is popularly known as the Standard Model (SM) of particle physics. I discuss the various components of the SM lagrangian density and the Higgs mechanism responsible for explaining masses of the elementary particles. The subsequent focus is directed towards the SUSY model – an extension beyond the SM – with particular attention given to one of its breaking mechanisms, the GMSB. This chapter stands as an essential part for the rest of the thesis, providing an overview of theoretical motivation and related experimental implications for designing searches for these new particles. For a comprehensive review of the QFT and the SM formalisms, please refer to standard textbooks like [37, 38].

2.1 Standard Model

The SM is based on the quantum field theory (QFT) and it incorporates the properties of elementary particles as known from various phenomena observed in nature and measured experimentally. It describes the particle interactions originating from electromagnetic, weak, and strong forces. It does not include the description of gravitational force at the level of elementary particles. Particle content of the SM is summarized in Figure 2.1. It comprises particles with integral and half-integral spins. The basic building blocks of matter, known as quarks and leptons, are spin-1/2 fermions, and there are three generations of them. These generations follow identical mathematical formulation for various interactions as allowed various symmetries but corresponding particles are heavier in higher generations.

2.1.1 Leptons

There are six leptons in the SM, grouped in pairs as electron and electron neutrino (e, ν_e), muon and muon neutrino (μ, ν_μ), and tau and tau neutrino (τ, ν_τ). The e, μ and τ are charged leptons while the corresponding neutrinos are neutral. The charged leptons can interact via electromagnetic and weak interactions but neutrinos interact only via weak interaction, which makes

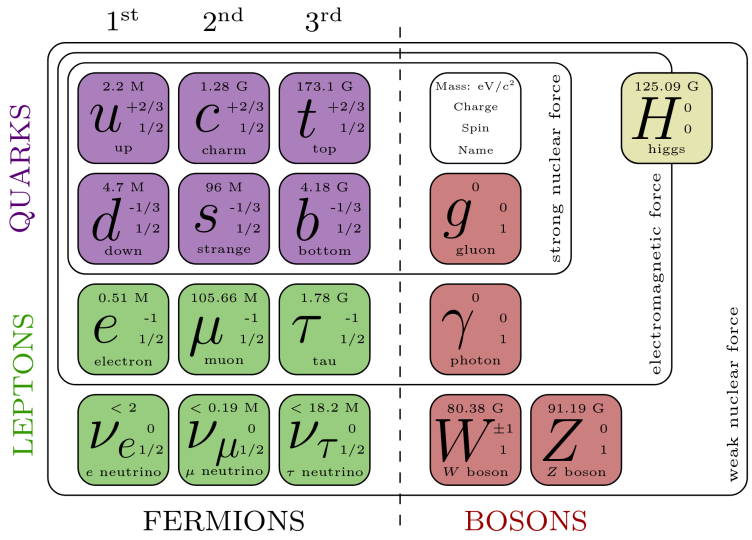

FIGURE 2.1: Figure showing building block of the SM with their quantum numbers [39].

it experimentally difficult to study their properties (as their probability of interacting with detector is very small). A established from the parity violation observed in natural phenomenon such as radioactive β decays, only left handed leptons are known to participate in the weak interactions. Leptons do not experience strong force, which distinguishes them from the remaining fermions of the SM, the quarks.

In the SM, we only consider left-handed neutrinos and right-handed anti-neutrinos, and these are taken to be massless. However, one of the most intriguing observations about neutrinos is the phenomenon of neutrino mixing, also known as flavor oscillation [40]. This phenomenon can arise from the disparity between the masses of neutrinos of different flavors. Consequently, a neutrino of a specific flavor can transform into a different flavor during its propagation through space. The observation of neutrino oscillations has provided strong evidence that neutrinos have mass, contrary to the original assumption of them being massless particles. This is an active field of research investigations presently.

2.1.2 Quarks

There are six types of quarks, also known as flavors, namely up (u), down (d), charm (c), strange (s), top (t), and bottom (b). Quarks carry a fractional electric charge and spin of half unit. They carry a intrinsic quantum number called color charge that allows their interaction to be mediated via the strong force. Charges comes in three types, referred to as "red," "green," and "blue," although these names have nothing to do with the usual colors. Quarks carry one of these three color charges, while antiquarks carry one of the corresponding anticolors ("antired," "antigreen," and "antiblue"). Gluons carry a combination of a color and an anticolor but it will not be colorless. Gluons can have one of eight possible color-anticolor combinations, corresponding to the different combinations of the three primary colors (red, green, and blue) and their anticolors (antired, antigreen, and antiblue). Quarks can interact via strong, electromagnetic and weak interactions. A summary of quantum numbers carried by quarks and their

masses is provided in Table 2.1.

TABLE 2.1: Table showing six quarks with their mass, charge and spin.

Quark	Charge (e)	Isospin (I_3)	Baryon no. (B)	Strangeness (S)	Mass (MeV/c^2)
up (u)	+2/3	1/2	1/3	0	2.2
down (d)	-1/3	-1/2	1/3	0	4.7
charm (c)	+2/3	0	1/3	0	1.28x10^3
strange (s)	-1/3	0	1/3	-1	96
top (t)	+2/3	0	1/3	0	173.1x10^3
bottom (b)	-1/3	0	1/3	0	4.18x10^3

The baryons (bound states of three quarks or antiquarks) and mesons (quark-antiquarks) observed in nature are all color neutral particles i.e. we have not observed isolated color charge in nature or in experiments. The combinations of three colors (or three anticolors) in equal amount (as happens in baryons) or those with equal amount of color-anticolor give "colorless" states (as in mesons). This is illustrated in Figure 2.2 (left and right-top). Strong force between quarks are mediated by gluons by changing their colors as illustrated in Figure 2.2 (right-bottom).

Since gluons themselves also have color, allowing them to interact with other gluons. This unique property makes the strong force different from electromagnetic force, as it becomes non-linear and self-interacting.

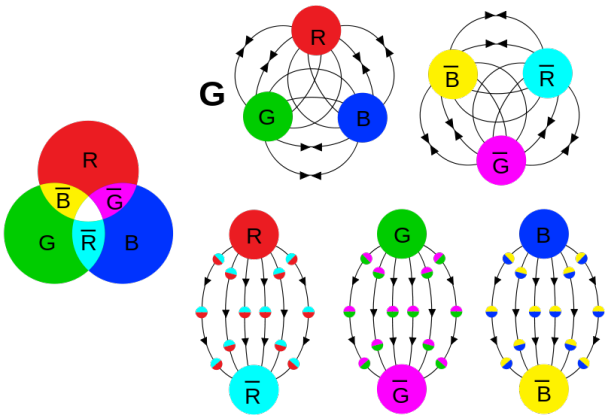

FIGURE 2.2: Figure showing color charge and gluon exchange [Image courtesy: Wikipedia].

2.1.3 SM bosons

Interactions among the SM particles are understood to be mediated by spin-1 bosons (also called vector bosons). The photon and gluons are both massless and they mediate electromagnetic and strong forces, respectively. The weak interactions are mediated by massive vector bosons, denoted as $W^{\pm}$ and Z bosons. The masses of the elementary particles are explained by their interactions with the H, which is a charge neutral particle and carries a spin zero (also

called scalar boson). The strengths of these forces and range of their influence, as we observe around us, are summarized in Table 2.2. It is to be noted that the interaction strengths depend on energies at which these are measured.

TABLE 2.2: Table showing details of four fundamental forces in nature.

Force	Relative Strength	Range	Mediating Particles	Act on
Gravitational	10^{-38}	Infinite	Graviton	All particles with non-zero mass
Weak Nuclear	10^{-5}	10^{-18} m	W and Z bosons	Quarks and leptons
Electromagnetic	10^{-2}	Infinite	Photon	All Charged particles
Strong Nuclear	1	10^{-15} m	Gluons	Quarks

2.2 SM interactions in QFT framework

The QFT is a theoretical framework that unifies quantum mechanics and special relativity to describe elementary particles and their interactions using fields. In QFT, particles such as electrons are considered as manifestations of specific fields, much like how the photon represents the electromagnetic field. The behavior and interactions of these fields are described by a lagrangian density, $\mathscr{L}$, which governs the dynamics of the fields, determining how they evolve in time and space. It encodes the interactions between the fields and incorporates the conservation laws observed in nature. By quantizing these fields and treating them as operators, QFT enables the calculation of physical observables, such as scattering amplitudes. These calculations provide insights into the behavior of particles and the fundamental forces that govern their interactions, and help in making predictions about not yet observed phenomena or measured quantities [38].

One important aspect of QFT is gauge invariance, which ensures that the lagrangian describing the interactions of particle wave functions remains unchanged under certain symmetry transformations. These gauge symmetries reflect the conservation laws observed in nature, such as the conservation of electric charge or weak isospin. Gauge invariance plays a crucial role in maintaining the consistency and predictive power of QFT.

2.2.1 Theoretical formulation of Standard model

The lagrangian of the SM is gauge invariant under the $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry transformations and the terms representing interactions of gauge bosons, fermions and H sector are summarized in the following :

$$\mathscr{L}_{\text{SM}} = \mathscr{L}_{\text{gauge}} + \mathscr{L}_{\text{fermions}} + \mathscr{L}_{\text{Higgs}} + \mathscr{L}_{\text{Yukawa}} \tag{2.1}$$

- **Gauge Boson Terms** ($\mathscr{L}_{\text{gauge}}$): This term describes the interactions of the gauge bosons associated with the three fundamental forces in the SM.

$$\mathscr{L}_{\text{gauge}} = -\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} - \frac{1}{4} W^a_{\mu\nu} W^{a\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \tag{2.2}$$

– $SU(3)_C$ represents the color symmetry of quantum chromodynamics (QCD), describing the strong force that binds quarks together to form protons, neutrons, and other hadrons. The gluon fields are given by:

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f_{abc} G^a_\mu G^a_\nu \tag{2.3}$$

with G^a_ν represent the eight gluon fields, g_s is the gauge coupling constant corresponding to the color charge, and f_{abc} are the structure functions of SU(3). Here, a,b,c are in the range of 1,2...8.

– $SU(2)_L$ represents the weak isospin symmetry, which is responsible for weak interactions involved in processes like beta decay. The weak isospin field $W^a_{\mu\nu}$ are associated with the three weak gauge bosons W^a_μ, where a ranges from 1 to 3.

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g f_{abc} W^a_\mu W^a_\nu \tag{2.4}$$

where, W^a_ν also involve derivatives of the weak gauge bosons, terms related to the weak coupling constant g, and the structure functions f_{abc} of the $SU(2)$ group.

– $U(1)_Y$ represents the hypercharge symmetry, associated with the weak hypercharge, which is a combination of electric charge and the third component of weak isospin. The hypercharge field strength is given by:

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.5}$$

- **Fermion Terms** ($\mathscr{L}_{\text{fermions}}$): This term describes the kinetic and interaction terms for fermions are presented in equation 2.6. In the Electroweak theory, fermions are defined in terms of the left- and right-handed components of χ_L (for isospin doublets) and ψ_R (for isospin singlets) after undergoing $SU(2) \times U(1)$ transformations. For the electrons and neutrino:

$$\chi_L = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \psi_R = e^-_R$$

Quarks also interacts via electroweak interactions, so right and left handed up quarks will be represented as :

$$\chi_L = \begin{pmatrix} u \\ d \end{pmatrix}_L, \psi_R = u_R \text{or } d_R$$

The electroweak lagrangian for fermions, $\mathscr{L}_{\text{fermions}}$ is derived by imposing gauge invariance.

$$\mathscr{L}_{\text{fermions}} = \bar{\chi_L}\gamma^\mu(i\partial_\mu - g\frac{\tau^a}{2}.W^a_\mu - g'\frac{Y_L}{2}B_\mu)\chi_L + \bar{e_R}\gamma^\mu(i\partial_\mu - g'\frac{Y_R}{2}B_\mu)e_R \tag{2.6}$$

where, g' is the weak hypercharge coupling constant and γ^μ are the Dirac gamma matrices.

Simply in QCD, kinetic and interaction terms for quarks in case of $SU(3)_C$ symmetry transformation, can be represented as:

$$\mathscr{L}_{\text{quarks}} = \sum_{k=1}^{3} i\bar{q}_k(\gamma^\mu \partial_\mu - m_k)q_k + \sum_{k=1}^{3} g_s(\bar{q}_k \gamma^\mu T_a q_k)G_\mu^a \tag{2.7}$$

where, first term is kinetic term which involve the derivative of the quark fields (q_k) with respect to spacetime coordinates, γ^μ are the Dirac gamma matrices, and m_k is the quark mass. The summation over k indicates that these terms account for three generations of quarks. Second term represents interaction terms for quarks with gluon as color field strength and T_a matrices represent the color matrices.

- **Higgs Term** ($\mathscr{L}_{\text{Higgs}}$): This term describes the Higgs mechanism, which gives masses to the gauge bosons and fermions. The higgs field is responsible for "breaking" the electroweak symmetry while conserving the gauge invariance, leading to the masses of particles like the W and Z bosons. It involves the concept of spontaneous symmetry breaking (SSB) and lagrangian for a complex scalar (spin 0) field can be written as:

$$\mathscr{L}_{\text{Higgs}} = \left| \left(i\partial_\mu - g\frac{\tau^a}{2}.W_\mu^a - g'\frac{Y}{2}B_\mu \right)\phi \right|^2 + V(\phi) \tag{2.8}$$

where, potential $V(\phi)$ defined as,

$$V(\phi) = -\mu^2 \phi^\dagger \phi - \lambda |\phi^\dagger \phi|^2 \tag{2.9}$$

Complex scalar field ϕ is a $SU(2) \times U(1)$ doublet (choose to arrange fields such that Y = 1):

$$\phi = \frac{1}{2}\begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.10}$$

This is also called a H doublet. Lets consider the case $\mu^2 < 0$ and $\lambda > 0$. We may choose the following vacuum state: $\phi_1 = \phi_2 = \phi_4 = 0$ and $\phi_3 = v$ and also adding $h(x)$ as a quantum fluctuations and describes the real scalar field for H boson, which, up to a phase, yields

$$v^2 = -\frac{\mu^2}{\lambda}, \qquad \phi = \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.11}$$

By inserting the value of ϕ in the interaction term of equation 2.8 and expanding τ^a generators in the form of pauli matrices:

$$\begin{aligned}\frac{1}{8}\left|\left(g\tau^a.W_\mu^a+g'B_\mu\right)\phi\right|^2 &= \frac{1}{8}\left|\begin{pmatrix} gW_\mu^3+g'B_\mu & g(W_\mu^1-iW_\mu^2)\\ g(W_\mu^1-iW_\mu^2) & -gW_\mu^3+g'B_\mu\end{pmatrix}\begin{pmatrix}0\\ v+h(x)\end{pmatrix}\right|^2 \\ &= \frac{1}{8}(v+h(x))^2g^2\left[(W_\mu^1)^2+(W_\mu^2)^2\right] \\ &+ \frac{1}{4}(v+h(x))^2(-gW_\mu^3g'B_\mu)(gW^{3\mu}+g'B^\mu)\end{aligned} \tag{2.12}$$

Since $W^\pm$ are the combinations of $(W_1 \pm iW_2)/\sqrt{2}$ then first term in above equation 2.12, it will give M_W after comparison with $M_W W^+W^-$,

$$\frac{1}{4}(v+h(x))^2g^2W_\mu^+W^{-\mu} \quad => \quad \boxed{M_W=\frac{1}{2}vg} \tag{2.13}$$

Now rewrite the second term of 2.12 as $Z^\mu = (gW_\mu^3 - g'B_\mu)\sqrt{g^2+g'^2}$ and $A^\mu = (gW_\mu^3 - g'B_\mu)\sqrt{g^2+g'^2}$, we will get M_Z after comparing with $\frac{1}{2}M_ZZ^2$ and also verifies that γ is massless.

$$\frac{1}{8}(v+h(x))^2(gW_\mu^3-g'B_\mu)^2+0.(gW_\mu^3+g'B_\mu)^2 \;=> \; \boxed{M_Z=\frac{1}{2}v\sqrt{g^2+g'^2}} \;\&\; \boxed{M_A=0} \tag{2.14}$$

The ratio of g′ and g is defined as $\tan\theta_W$, then masses of W and Z can also be represented in terms of θ_W as in equation 2.15. This relation also shows the inequality between the mass of W and Z.

$$\frac{M_W}{M_Z}=\cos(\theta_W) \tag{2.15}$$

After knowing the M_W & M_Z from experiments, value of g, g' and v can be calculated using fermi constant (G) relation which is,

$$\frac{1}{v^2}=\frac{g^2}{8M_W^2}=\frac{G}{\sqrt{2}} \tag{2.16}$$

- **Yukawa Couplings Term** ($\mathscr{L}_{\text{Yukawa}}$): This term describes the yukawa interactions between the higgs field and the fermions. It is responsible for generating fermion masses through the H mechanism. Lagrangian for electron mass under $SU(2)_L \times U(1)_Y$ symmetry group after pairing χ_L & ψ_R, the H doublet,is shown as

$$\mathscr{L}_{Yukawa}=-G_e\left(\begin{pmatrix}\bar{\nu_e} & \bar{e^-}\end{pmatrix}_L\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix}e_R+\bar{e}_R\begin{pmatrix}\bar{\phi}^+ & \bar{\phi}^0\end{pmatrix}\begin{pmatrix}\nu_e\\ e^-\end{pmatrix}_L\right) \tag{2.17}$$

where G_e denotes the yukawa coupling of the electron and after SSB as shown in equation 2.11:

$$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}$$

$$\mathscr{L}_{\text{Yukawa}} = -\frac{G_e}{\sqrt{2}} v(\bar{e}_R e_L + \bar{e}_L e_R) - \frac{G_e}{\sqrt{2}}(\bar{e}_R e_L + \bar{e}_L e_R)h \tag{2.18}$$

The first expression contains an electron mass term, $M_e = \frac{G_e}{\sqrt{2}} v$. There is no mass term of neutrino in the $\mathscr{L}_{Yukawa}$ because of the absence of right handed neutrino, ν_R. In SM, the Yukawa interactions are the only source of the fermion masses and similarly, mass of quarks can also be derived.

The expanded expression of the SM lagrangian as mentioned by equation 2.1:

$$\begin{aligned} \mathscr{L}_{\text{SM}} = &\left\{ -\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} - \frac{1}{4} W^a_{\mu\nu} W^{a\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \right\} + \left\{ g_s (\bar{q} \gamma^\mu T_a q) G^a_\mu \right\} \\ &\left\{ \bar{\chi}_L \gamma^\mu (i\partial_\mu - g \frac{\tau^a}{2} . W^a_\mu - g' \frac{Y_L}{2} B_\mu) \chi_L + \bar{\psi}_R \gamma^\mu (i\partial_\mu - g' \frac{Y_R}{2} B_\mu) \psi_R \right\} + \\ &\left\{ \left| \left(i\partial_\mu - g \frac{\tau^a}{2} . W^a_\mu - g' \frac{Y}{2} B_\mu \right) \phi \right|^2 + V(\phi) \right\} - \left\{ G_1 \bar{\chi}_L \psi_R - G_2 \bar{\chi}_L \psi_R + h.c. \right\} + \ldots \end{aligned} \tag{2.19}$$

where, first curly bracket represents the interaction of the three SM gauge bosons: $W^\pm$, Z, γ & gluons. All terms in second and third curly brackets represents the kinetic and interaction terms of lepton and quarks with SM gauge bosons. The fourth and fifth curly brackets terms belong to $\mathscr{L}_{\text{Higgs}}$ and $\mathscr{L}_{\text{Yukawa}}$, which describes the coupling of H with gauge bosons and quarks and leptons respectively.

2.3 Limitations of Standard model

The SM of particle physics is a very successful theory that describes the interactions of fundamental particles observed in nature. However, it is not a complete theory, and there are several limitations and open questions that it cannot yet explain. Some of the limitations of the SM are:

- Hierarchy problem: The H boson mass, denoted as m_H is determined by the higgs potential, which relies on the H self-coupling parameter (λ) and the H mass parameter (μ^2). The vacuum expectation value (VEV) of the higgs field is given by $\sqrt{\mu^2/2\lambda}$ as shown in equation 2.11. While the measured value of the H boson mass is around 125 GeV, but theoretically it receives considerable quantum corrections from virtual particles and interactions coupled to the higgs field. These corrections can be expressed as:

$$\Delta m_H^2 = -\frac{\lambda_f^2}{8\pi^2} \Lambda_{\text{UV}}^2 + \ldots \tag{2.20}$$

Here, the correction arises from a loop involving a dirac fermion, f, with mass m_f. Figure 2.3 illustrates the one-loop quantum correction to the H boson mass squared parameter resulting from the top quark (which yields the largest correction).

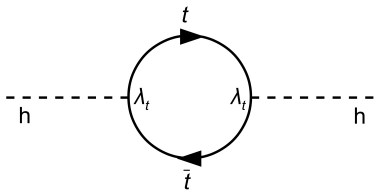

FIGURE 2.3: Diagram depicting the top-loop correction to the m_H^2 term.

The term Λ_{UV} represents the ultraviolet momentum cutoff, which indicates the energy scale at which the theory becomes sensitive to high-energy phenomena. If the ultraviolet cutoff Λ_{UV} is of the order of the planck scale (10^{19} GeV), the quantum correction to μ^2 would be excessively large. This substantial discrepancy is referred as the hierarchy problem, and suggests the presence of new physics beyond the SM that could explain the H boson mass.

- Gauge coupling unification: Gauge coupling unification conceptualises having similar strengths of the fundamental forces i.e. unification of the coupling constants: coupling constants of the weak (g), electromagnetic (g'), and strong (g_s) forces at a large energy scale.The coupling constants like g, g' & g_s are proportional to the α_1, α_2 & α_3. The evolution of the inverse of couplings with energy scale is given by the renormalisation group equations (RGEs):

$$\alpha_i^{-1} = \alpha_i^{-1}(m_Z) + \frac{b_i}{2\pi} \log \frac{Q}{m_Z} \tag{2.21}$$

Here, $\alpha_i^{-1}(m_Z)$ represents the coupling constant calculated at the electroweak scale, and b_i are the beta function coefficients. The beta function coefficients differ for each coupling and depend on the particle content of the theory. They describe how the coupling constants change as the energy scale increases or decreases. The evolution of the coupling constants is illustrated in Figure 2.4. Gauge coupling unification suggests that at very high energy scales, the three coupling constants might converge and become equal, but they do not come to a common value at any scale as depicted in Figure 2.4.

- Including gravity: The SM describes three of the four fundamental forces of nature: the electromagnetic force, the weak force, and the strong force. However, it does not include gravity, which is described by general relativity. The unification of all four forces into a single theory, known as a theory of everything, is a major goal of theoretical physics.

- Dark matter: The SM cannot explain the existence of dark matter, which does not interact with light and other forms of electromagnetic radiation. Dark matter is believed to make up about 85% of the matter in the universe, but its nature is still unknown.

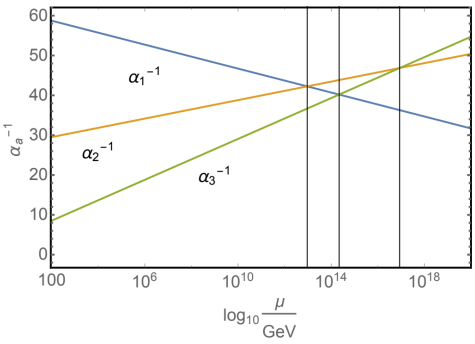

FIGURE 2.4: Evolution of the inverse of the coupling constants as a function of the logarithm of energy scale (Q or μ), illustrating the concept of gauge coupling unification in the SM [26].

- Neutrino masses: The SM assumes that neutrinos are massless particles, but experiments have shown that they have very small masses. The mechanism for giving neutrinos mass is not yet understood.
- Matter-antimatter asymmetry: The SM cannot explain why the universe is made up mostly of matter, with very little antimatter. According to the big bang theory of the origin of universe, the early universe should have produced equal amounts of matter and antimatter, which should have annihilated each other. However, this did not happen, and the reason for this asymmetry is still unknown.

In summary, while the SM is a very successful theory, it has several limitations and open questions that require further study and may require the development of new theories beyond the SM. Several extensions of the SM (SM) of elementary particles attempt to provide an explanation for the origin of dark matter (DM) [41, 42] and to resolve the gauge hierarchy problem [43, 44, 45, 46].

2.4 Supersymmetry

Supersymmetry (SUSY) [47, 48] is a hypothesized symmetry between fermions and bosons, which, when included to extend the SM, predicts a new bosonic (fermionic) superpartner for each SM fermion (boson).

$$\hat{Q}\,|\text{Boson}\rangle \;=\; |\text{Fermion}\rangle\,, \quad \hat{Q}\,|\text{Fermion}\rangle \;=\; |\text{Boson}\rangle \tag{2.22}$$

where the generator of the SUSY algebra, denoted as $\hat{Q}$.

The introduction of SUSY particles, often denoted with a prefix "s" or "super" and represented as tilde over as shown in Figure 2.5. For example, the superpartner of quarks (q) are called squarks ($\tilde{q}$) , while leptons (ℓ) superpartner are called sleptons ($\tilde{\ell}$). Gauge bosons, such as the photon and the gluon, have corresponding fermionic superpartners called gauginos. In supersymmetric extension of the SM, we need two H doublets and corresponding higgsino fields.

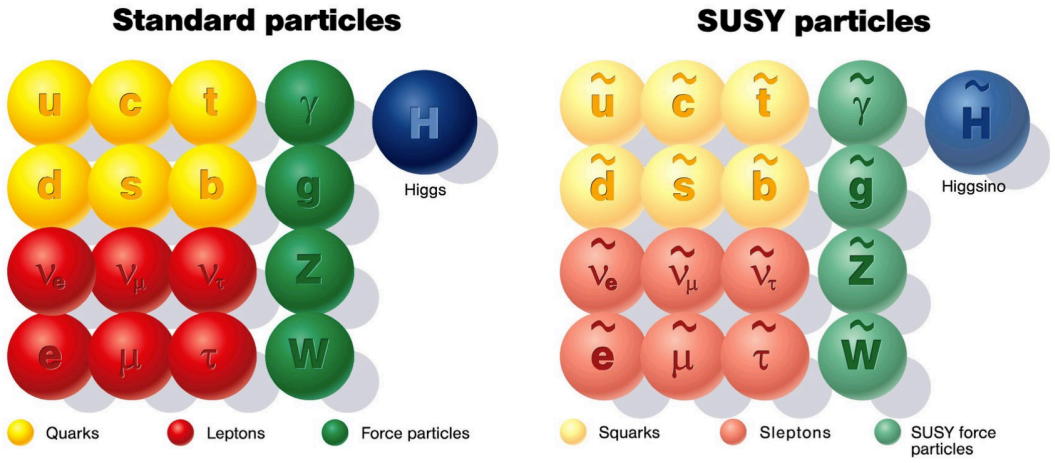

FIGURE 2.5: Diagram super-partners (right) of standard particles (left). [Image credit: DESY at Hamburg [49]]

2.4.1 Motivation for Supersymmetry

The concept of SUSY was developed to address several key issues in particle physics, including the hierarchy problem mentioned in section 2.3. The SUSY provides an elegant solution to the hierarchy problem by introducing complex scalar boson loops (e.g., stop quark) that cancel out the large quantum corrections to the H boson mass arising from fermion loops in the SM (as shown in Figure 2.3). This cancellation stabilizes the H boson mass at a lower energy scale. Figure 2.6 presents a diagram depicting the stop loop correction to the m_H^2 term.

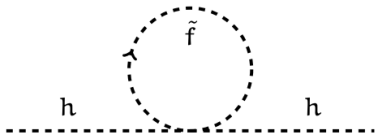

FIGURE 2.6: Diagram depicting the stop loop correction to the m_H^2 term.

The following equation represents the quantum corrections which can avoid quadratic divergences (if $\lambda_s \approx |\lambda_f|^2$) [25]:

$$\Delta m_H^2 \approx \frac{1}{8\pi^2}\left(\lambda_s - |\lambda_f|^2\right)\Lambda_{\mathrm{UV}}^2 + \frac{\lambda_s m_S^2}{8\pi^2}\ln(\Lambda_{UV}/m_S) + .. \tag{2.23}$$

Here, λ_s represents the coupling constant of "stop" associated with the H field, Λ_{UV} is the ultraviolet cutoff scale. Now, H boson mass parameter has only a logarithmic dependence on the scale of new physics. This can reduce the need for fine-tuning of the H boson mass

In SUSY models with R-parity conservation [50], the lightest supersymmetric particle (LSP) is stable and often taken to be neutral and weakly interacting, making it a possible DM candidate. In addition, gauge coupling unification is a natural possibility in supersymmetric theories [51].

An illustrative diagram showcasing the gauge coupling unification of the three forces except gravity using SUSY model as depicted in Figure 2.7.

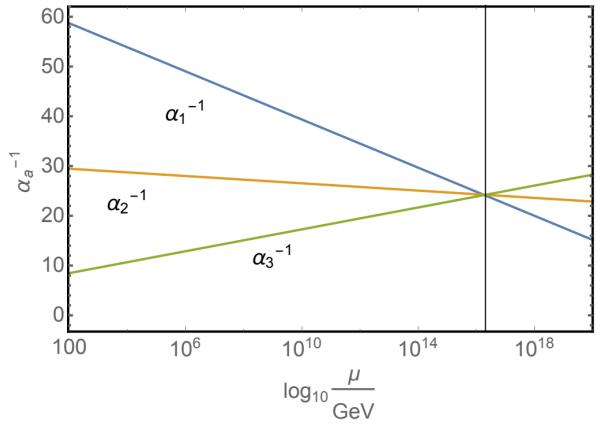

FIGURE 2.7: Gauge coupling unification in the SUSY Model [26].

2.4.2 Minimal Supersymmetry model

The "minimal Supersymmetry model" introduces the fewest possible new particles and interactions to maintain consistency with experimental observations and theoretical requirements. The lagrangian of the MSSM can be expressed as:

$$\mathscr{L}_{\text{MSSM}} = \mathscr{L}_{\text{SUSY}} + \mathscr{L}_{\text{soft}} \tag{2.24}$$

In this equation, the first term corresponds to the SUSY generalization of the lagrangian in the SM, while the second term represents the breaking of SUSY. Following the breaking of SUSY, the MSSM states can mix, giving rise to the physical mass eigenstates. This mixing leads to the emergence of five H bosons after the electroweak symmetry breaking: two CP-even neutral states (h^0 and H^0), a CP-odd neutral state (A^0), and two charged states ($H^{\pm}$). The gaugino and higgsino states combine to form four neutralinos ($\widetilde{\chi}_1^0$, $\widetilde{\chi}_2^0$, $\widetilde{\chi}_3^0$, $\widetilde{\chi}_4^0$) and four charginos ($\widetilde{\chi}_1^{\pm}$, $\widetilde{\chi}_2^{\pm}$), respectively. The specific mixing pattern depends on the chosen mechanism for SUSY breaking. The MSSM particle states, as depicted in gauge and mass eigenstates, are summarized in Table 2.3.

Neutralino/Chargino

The mixing of higgsinos and electroweak gauginos occurs as a result of electroweak symmetry breaking. The combination of neutral higgsinos ($\widetilde{H}_u^0$ and $\widetilde{H}_d^0$) with neutral gauginos ($\widetilde{B}^0$ and $\widetilde{W}^0$) leads to the formation of four mass eigenstates known as neutralinos. Additionally, the mixing of charged higgsinos ($\widetilde{H}_u^+$ and $\widetilde{H}_d^-$) with winos ($\widetilde{W}^+$ and $\widetilde{W}^-$) results in the creation of two mass eigenstates carrying charges ± 1, which are referred to as charginos. There are 4 neutralinos: $\widetilde{\chi}_1^0$, $\widetilde{\chi}_2^0$, $\widetilde{\chi}_3^0$, $\widetilde{\chi}_4^0$ and charginos: $\widetilde{\chi}_1^{\pm}$, $\widetilde{\chi}_2^{\pm}$, and mass of these are arranged in the ascending order.

TABLE 2.3: SUSY particles in MSSM [25].

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

The neutralino mass eigen states are expressed as linear combination of $\psi^0 = (\widetilde{B}^0, \widetilde{W}^0, \widetilde{H}_u^0, \widetilde{H}_d^0)$ as : $\widetilde{\chi}_i^0 = N_{ij}\psi_j$, where N_{ij} is 4x4 matrices. The neutralino mass matrix:

$$M_{\widetilde{\chi}^0} = \begin{pmatrix} M_1 & 0 & -c_\beta s_W m_Z & s_\beta s_W m_Z \\ 0 & M_2 & c_\beta c_W m_Z & -s_\beta c_W m_Z \\ -c_\beta s_W m_Z & c_\beta c_W m_Z & 0 & -\mu \\ s_\beta s_W m_Z & -s_\beta c_W m_Z & -\mu & 0 \end{pmatrix} \tag{2.25}$$

where, M_1,M_2 comes directly from the MSSM soft lagrangian as mass parameter, μ is the higgsino supersymmetric mass parameter and β is the $\tan^{-1}$ of the ratio of two MSSM Higgs vacuum expectation values. s_β, c_β, s_W and c_W represent $\sin\beta$, $\cos\beta$, $\sin\theta_W$ and $\cos\theta_W$.

Similarly, chargino mass matrix is also expressed in 4x4 matrices and its mass eigen states are expressed as linear combination of $\psi^\pm = (\widetilde{W}^+, \widetilde{H}_u^+, \widetilde{W}^-, \widetilde{H}_d^-)$ as : $\widetilde{\chi}_i^\pm = C_{ij}\psi_j^\pm$, where both $M_{\widetilde{\chi}^\pm}$ are expressed in terms of M_2, μ, M_W, s_β and c_β.

R-parity

Proton is stable with a lifetime of $> 10^{34}$ years. Certain interactions mediated by sparticles could lead to the proton decay in MSSM [52], hence violating the baryon and lepton quantum number conservation. To prevent proton instability in MSSM, R-parity conservation is imposed as a discrete symmetry which is defined as $R = (-1)^{3B+L+2S}$, where B represents the baryon number, L represents the lepton number, and S represents the spin of a particle. It is a multiplicative quantum number that distinguishes between ordinary particles (R = +1) and supersymmetric particles (sparticles) (R = -1) [25]. R-parity conservation has several important consequences:

- In high-energy collisions, sparticles can only be produced in pairs, ensuring that R-parity is conserved.
- Each sparticle can decay into odd number of sparticle at each stage of decay chain and eventually decays to LSP.
- The LSP is stable. The lightest neutralino ($\tilde{\chi}_1^0$) is typically taken to be the LSP. Being neutral, weakly interacting and stable, it is a good candidate for dark matter and can significantly contribute to the observed dark matter density in the universe [53].

The conservation of R-parity in the MSSM is often assumed for theoretical and phenomenological reasons. It provides a framework for studying the properties and interactions of sparticles, as well as their implications for cosmology and particle physics experiments.

In general, the masses of the SUSY particles in the MSSM are generated through a process known as soft SUSY breaking. This breaking occurs by introducing terms into the lagrangian that solely involve the SUSY particles, explicitly breaking the symmetry. This procedure introduces approximately 105 new parameters into the theory, including masses, mixing angles, and phases etc. However, assuming the existence of a natural underlying mechanism for SUSY breaking can significantly reduce the number of free parameters. Unlike the Higgs mechanism, a form of spontaneous symmetry breaking is not feasible in the MSSM because none of the fields can have a non-zero v.e.v. that can break SUSY without disrupting gauge invariance. Various methods exist for softly breaking SUSY, with prominent examples being gravity-mediated or Planck scale-mediated SUSY breaking (PMSB), anomaly-mediated SUSY breaking (AMSB), and gauge-mediated SUSY breaking (GMSB) [25]. Section 2.4.3 provides a brief overview of GMSB to underscore the importance of investigating SUSY with photons, which serves as the central theme of this thesis.

2.4.3 Gauge-Mediated SUSY breaking

Gauge-mediated supersymmetry breaking (GMSB) is a mechanism that explains how SUSY can be broken (in a hidden sector) and transmitted to the visible sector (chiral supermultiplets) of the theory through gauge interactions as messengers [54]. These messenger fields interact with the particles in the visible sector through the familiar gauge interactions of the Standard Model, specifically the $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge interactions. The soft mass terms in the MSSM resulting from loop diagrams involving messengers in GMSB, are computed by evaluating the relevant loop integrals and depend on the parameters of the GMSB model , such as the messenger masses (M_{mess}), gauge couplings or loop factor for Feynman diagrams involving gauge interactions ($\alpha_a/4\pi$), and SUSY breaking scale ($< F >$). The soft masses of MSSM particles represented as:

$$M_{\text{soft}} \sim \frac{\alpha_a}{4\pi} \frac{< F >}{M_{\text{mess}}} \tag{2.26}$$

One key feature of GMSB is the possibility of a lower scale for SUSY breaking, characterized by $< F > \sim 10^4$ GeV, if it is roughly comparable to the messenger scale M_{mess}. Consequently, the masses of supersymmetric particles, such as squarks, sleptons, and gauginos, are expected

to be in a TeV (10^3 GeV) range, which can be produced in the pp collisions at the LHC if these exist in nature.

Gravitino/goldstino

The lower scale ($< F > \sim 10^4 - 10^6$ GeV) also leads to a generic prediction in GMSB that the gravitino ($\widetilde{G}$) is the lightest supersymmetric particle (LSP) with spin 3/2 and all other MSSM particles decay with a $\widetilde{G}$ in the final state [25]. When SUSY is spontaneously broken, the $\widetilde{G}$ acquires a mass by absorbing ("eating") the goldstino, which becomes its longitudinal (helicity $\pm 1/2$) components. So, $\widetilde{G}$ inherits the non-gravitational interactions of the goldstino and can also play an important role in collider physics experiments. The mass of the gravitino can be determined by the relationship in some approximations:

$$M_{\widetilde{G}} \sim \frac{< F >}{\sqrt{3} M_{\mathrm{P}}} \tag{2.27}$$

The decay rate of any sparticle ($\widetilde{X}$) into its Standard Model partner (X) plus a goldstino/gravitino ($\widetilde{G}$) can be represented by equation 2.28 using effective lagrangian of goldstino:

$$\Gamma(\widetilde{X} \to X\widetilde{G}) = \frac{m_{\widetilde{X}}^5}{16\pi < F >^2} \left(1 - \frac{m_X^2}{m_{\widetilde{X}}^2}\right)^4 \tag{2.28}$$

For given $m_{\widetilde{X}}$, if $< F >$ is smaller then decay width will be larger. More precisely, if the $m_{\widetilde{X}} > 100$ GeV and scale $< F > \sim 10^6$ GeV, then the decay process $\widetilde{X} \to X\widetilde{G}$ can happen rapidly and can be observed in collider detectors. By comparing equations 2.27 & 2.28, it can be concluded that a very light gravitino is relevant for collider phenomenology.

Decays of Neutralino to Gravitino

As shown by equation 2.28, the decay rate can be faster with involvement of lighter sparticle like $\widetilde{\chi}_1^0$ which is next-to-lightest supersymmetric particle (NLSP) in GMSB. Since $\widetilde{\chi}^0$ is the mixture of bino, wino (or also a admixture of photinos) and neutral higgsinos as discussed in above section 2.4.2, it undergoes decay into a photon and a gravitino with a certain partial width.

$$\Gamma(\widetilde{\chi}_1^0 \to \gamma\widetilde{G}) = 2 \times 10^{-3} \kappa_{1\gamma} \left(\frac{m_{\widetilde{\chi}_1^0}}{100\ \mathrm{GeV}}\right)^5 \left(\frac{< F >}{100\ \mathrm{TeV}}\right)^{-4} \mathrm{eV} \tag{2.29}$$

where $\kappa_{1\gamma}$ is the photino content as $|N_{11}\cos\theta_W + N_{12}\sin\theta_W|^2$, where N_{11} & N_{12} belongs to $M_{\widetilde{\chi}^0}$ matrix. For $m_{\widetilde{\chi}_1^0} > 100$ GeV and $\sqrt{< F >} \sim 10^8 - 10^{12}$ GeV (which correspond to $m_{\widetilde{G}} \sim$ few GeV from equation 2.27), then this decay width is still large enough to allow $\widetilde{\chi}_1^0$ to decay before it has left a collider detector.

Neutralino decay can be categorised completely (under some approximations) into bino, wino and higgsino depending on the M_1, M_2, μ parameters [56, 55, 57].

- Bino like NLSP ($\widetilde{\chi}_1^0$) mass, M_1 is lesser than μ & M_2 and it will mostly decay into $\gamma + \widetilde{G}$ than $Z + \widetilde{G}$ as shown in Figure 2.8 (left).

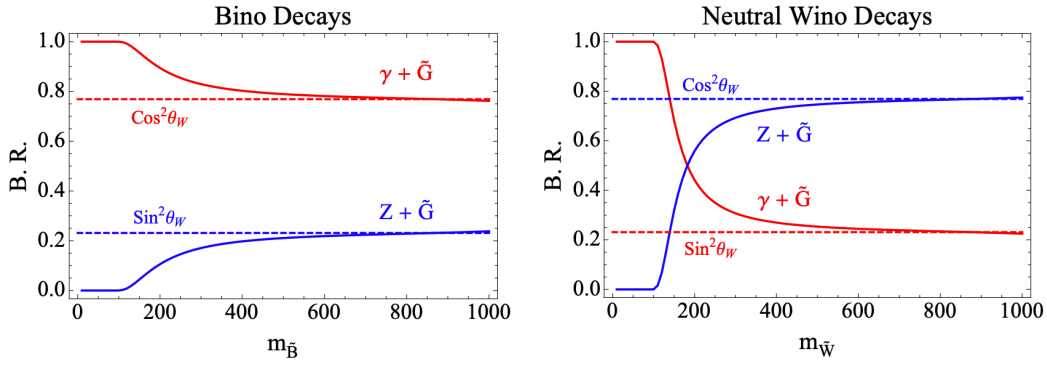

FIGURE 2.8: Branching fractions for Bino (left) and neutral wino (right) as a function of mass.[55].

- Wino like NLSP ($\widetilde{\chi}_1^0$) mass, M_2 is lesser than μ & M_1 and the NLSP neutralino will be nearly degenerate to chargino neutralino. The neutral wino will produce a $Z + \widetilde{G}$ or $\gamma + \widetilde{G}$, as shown in Figure 2.8 (right) and charged wino ($\widetilde{\chi}_1^{\pm}$) will decay into $W^{\pm} + \widetilde{G}$.
- Higgsino like NLSP will mostly decay into Z or H, depending on the sign of μ and tanβ .
  - For μ >0 & low tanβ (≈2), then $\widetilde{\chi}_1^0$ will dominantly decay into $Z/\gamma + \widetilde{G}$.
  - For μ <0 & low tanβ (≈2), then $\widetilde{\chi}_1^0$ will dominantly decay into $H/\gamma + \widetilde{G}$.
  - For moderate to large tanβ, then $\widetilde{\chi}_1^0$ will decay into mixture of $Z/H/\gamma + \widetilde{G}$.

2.4.4 Simplified GMSB search models

In the quest to discover SUSY at high-energy particle colliders like the LHC, physicists have devised various simplified models that explore different regions of the SUSY parameter space [58, 59]. These simplified models are designed to organize the complex task of SUSY searches in a limited number of searches based on final state topologies, to increase the sensitivity of various searches in the data available, and to provide a clearer understanding of specific aspects of the theory

These simplified models typically involve a small number of SUSY particles, such as gluinos, squarks, neutralinos, or charginos, and focus on specific production and decay channels. At the LHC, which is a proton-proton collider, gluinos and squarks are most commonly produced in pairs through processes like gluon-gluon fusion and gluon-quark fusion. This is because gluons and (valence) quarks play a pivotal role in governing the likelihood of these interactions, according to their parton distribution functions (PDFs) with respect to fraction of momenta shared by each parton for a given center-of-mass energy [60]. On the other hand, the production of a chargino-neutralino pair ($pp \rightarrow \widetilde{\chi}_1^{\pm}\widetilde{\chi}_1^{\mp}/\widetilde{\chi}_1^{\pm}\widetilde{\chi}_1^0$) is also possible but occurs less frequently compared to gluinos and squarks, shown in Figure 2.9.

In GMSB, the produced sparticles will decay into neutralinos, which result in final states with $\widetilde{G}$ (LSP). $\widetilde{G}$ is weakly interacting then it will escape the detector and hence missing transverse momentum (p_T^{miss}) will be main signature to infer its presence in a collision event. Depending

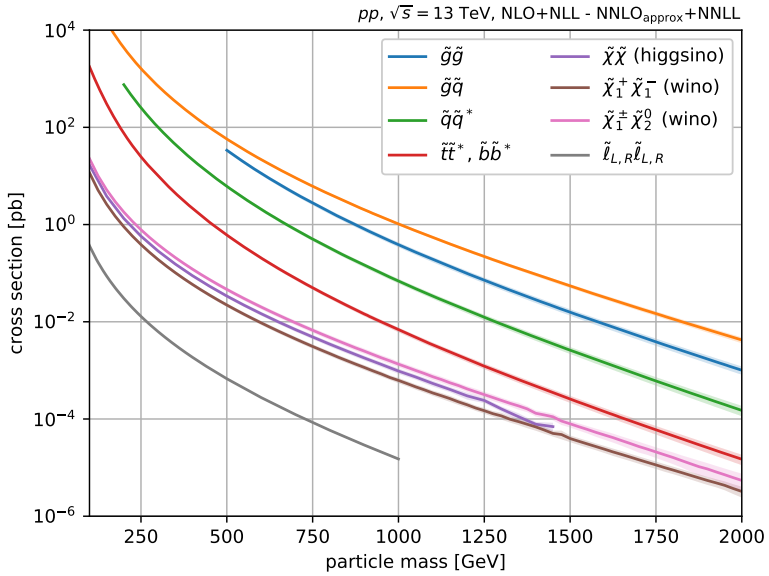

FIGURE 2.9: Cross-sections (in NNLO approximations) for SUSY particle production with respect to mass of corresponding particles in pp collision at LHC with center of mass energy, $\sqrt{s} = 13$ TeV [61].

on the exact nature of the NLSP, the decay chains can results in specific final states. Therefore, the SUSY particle production is investigated in several final states: jets+p_T^{miss} , at least 1ℓ + jets+p_T^{miss} , at least $1\ell + 0$jets+p_T^{miss} , atleast 1γ + jets+p_T^{miss} etc. In this thesis, we present a search for SUSY particles in final states comprising at least 1γ, jets and large p_T^{miss} for both strong productions (gluino and squark pair productions) and electroweakino pair productions (chargino or neutralino pair productions) using pp collision data at $\sqrt{s}$ = 13 TeV collected by the CMS detector. For example, in case of strong productions, gluino or squarks produced in pp collision, will further decay into $\tilde{\chi}_1^0$ or NLSP alongwith SM partners like $b\bar{b}$ with 100% branching fraction (BR), then NLSP will finally decay into two scenarios :

$$\tilde{\chi}_1^0 \to \gamma + \widetilde{G}, \quad \text{with BR} = 50\%$$
$$\tilde{\chi}_1^0 \to Z + \widetilde{G}, \quad \text{with BR} = 50\%$$

In other cases, there can be H instead of Z depending on the nature of NLSP (as discussed in above section 2.4.3) and if there is chargino $\tilde{\chi}_1^{\pm}$ as NLSP then there will be $W^{\pm}$ instead of Z. In all cases, mass of $\widetilde{G}$ is constant at 1 GeV (as discussed in above section 2.4.3) , and mass of $\tilde{\chi}_1^0$ or NLSP & mass of $\tilde{g}$ or $\tilde{q}$ are free parameters. The Feynman-like diagrams for simplified models are shown in Figure 2.10 and each model description will be explained in the chapter 5.

In summary, simplified GMSB search models provide a powerful approach to explore different regions of the SUSY parameter space. By focusing on specific SUSY particles and their production and decay channels, physicists can better understand the signatures of SUSY and analyse the experimental data from high-energy colliders like the LHC. These efforts bring us closer

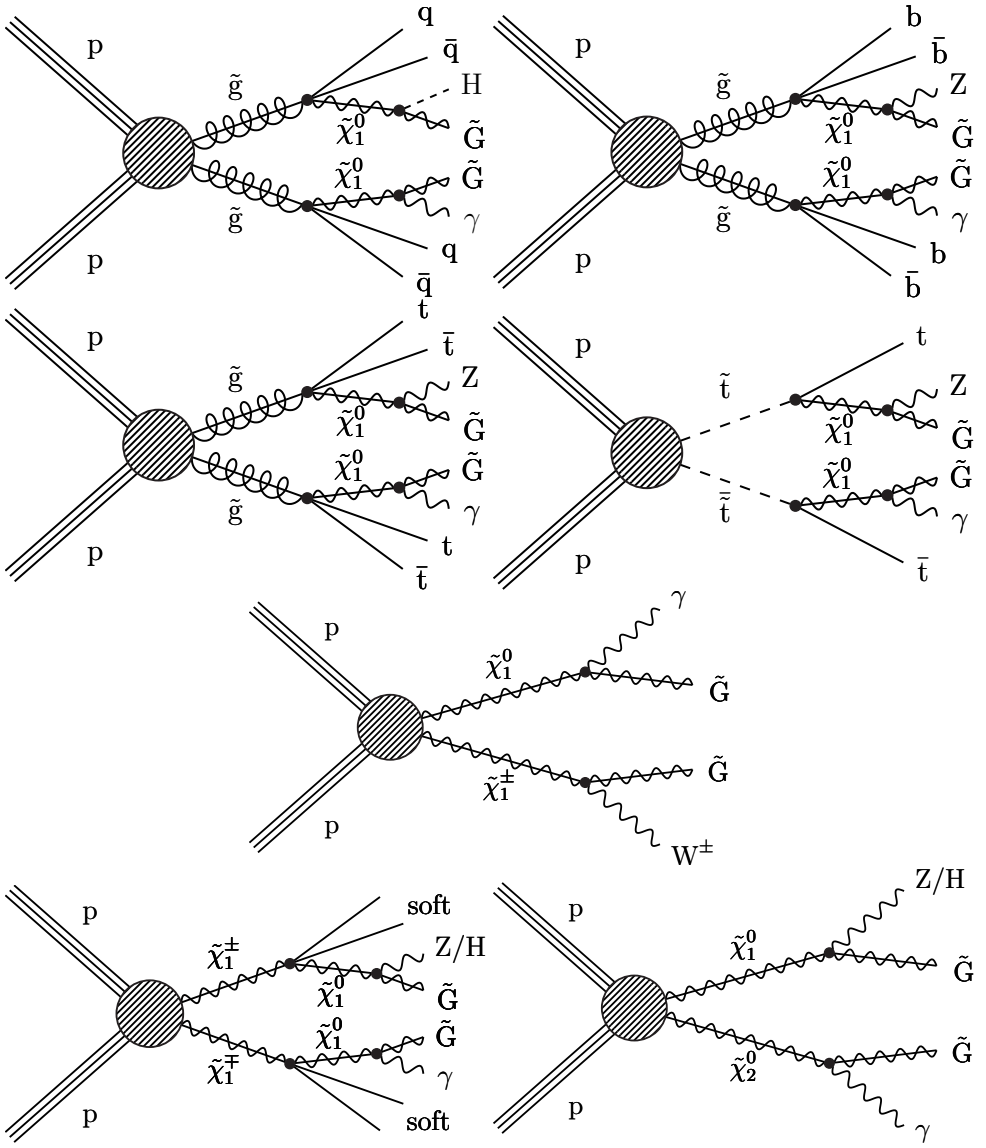

FIGURE 2.10: Simplified model scenarios for gluino pair production in top two rows and bottom two rows represents simplified models for electroweakino (direct chargino/neutralino) pair productions.

to potentially discovering evidence of new physics beyond the Standard Model and gaining a deeper understanding of the fundamental building blocks of the universe.

Chapter 3

Experimental setup and event reconstruction

In this doctoral thesis, we are summarizing a search for new particles which, if they exist in nature, might be produced in proton proton collisions at the LHC. The LHC has been operated at $\sqrt{s}$ of 7 or 8, 13 & 13.6 TeV in Run I, Run II, and Run III respectively. In this chapter we will summarize the LHC operations in general and details of Run II. .

The CMS, ATLAS, ALICE, and LHCb detectors are located at four interaction points of the LHC, to study outcomes of the collisions. ATLAS and CMS are general multi-purpose experiments, ALICE is a specialized detector to examine interactions between heavy nuclei, and LHCb, with a one-sided detector in the forward direction, is a b-physics experiment. This chapter provides a brief introduction to the CMS experimental setup and algorithms to reconstruct the physics objects and interpret the events.

3.1 Large Hadron Collider

The LHC is a particle accelerator, with a circumference of 27 km at the CERN in Switzerland. It is designed to accelerate beams of protons almost to the speed of light and to collide them upto $\sqrt{s}$ = 14 TeV. To achieve upto 7 TeV (tera-electron volts) energy per proton beam at the LHC involves a series of complex steps as shown in the Figure 3.1, involving various advanced technologies and techniques such as RF cavities, dipole and quadrupole magnets [62].

At the first step, protons are generated by stripping hydrogen atoms of their electrons in a device called a duoplasmatron and transferred to linear accelerator (LINAC), which uses radiofrequency cavities (RF cavities) to accelerate the protons. The RF cavities are superconducting cavity system that generates an electromagnetic field and accelerate particles upto 2 MeV when passing by these cavities within appropriate phase. Quadrupole magnets are also needed to focus the protons to keep them tightly packed, and dipole magnets are required to keep the particles in a circular trajectory inside the synchrotron. After reaching the other end of LINAC, protons will acquire energy of 50 MeV with a relative velocity of 0.3c and enter into proton synchrotron booster (PSB). The PSB consists of 4 synchrotron rings, and it delivers the proton bunches to proton synchrotron (PS) after accelerating the proton bunches up to

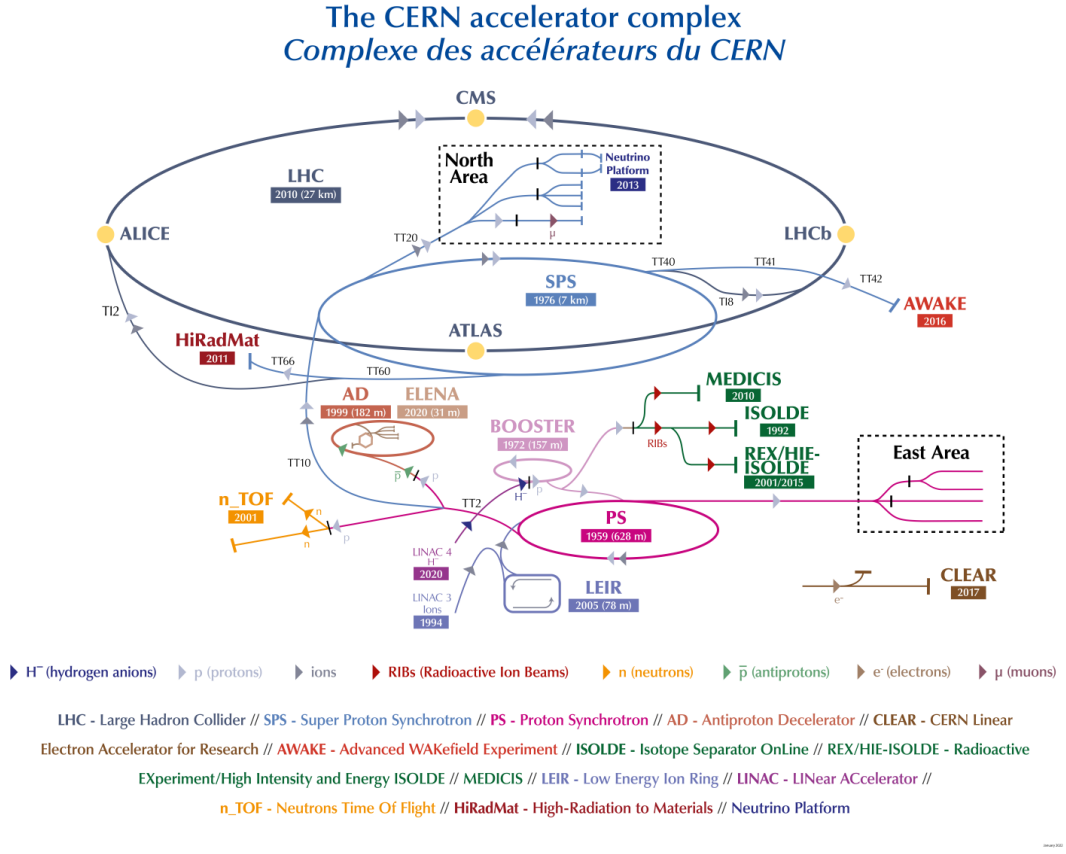

FIGURE 3.1: Schematic overview of the injection chain, pre-accelerators, the LHC ring and detectors at the interaction points.[63]

an energy of 1.4 GeV (0.74c). As the intensity of the proton beam increases, the space charge (repulsion between proton-proton in a bunch) will be a significant issue, so the technique of filling these bunches from PSB to PS has a crucial role in reducing the space charge effect [64]. Next, the protons are accelerated further in the PS, reaching an energy of 26 GeV. The PS has the crucial role of producing the required bunch spacing and obtaining sufficient brightness for the beams by using RF cavity of frequency 40-80 MHz, which are finally delivered to the LHC. In PS, 40 MHz RF cavity is used when bunch length is around 12 ns and have to be shortened to 4 ns to fit the next synchrotron which has 200 MHz system, so 80 MHz RF cavity is used to shorten the bunch length. After this, they are transferred to the super proton synchrotron (SPS), where they are accelerated to an energy of 450 GeV. This is the final step before being injected into the LHC. The LHC's RF cavities (having 400 MHz system) bring the 450 GeV energy of the particles to 7 TeV, which is more than 14 times their injection energy. There are 8 RF cavities in the LHC ring which accelerates the proton bunches by 16 MeV per round. The maximum energy is reached in around 20 minutes, with the bunches passing through the RF cavities more than 10 million times.

There are two rings in the LHC tunnel in which two proton beams travels in opposite direction, directed by powerful superconducting electromagnets that sustain a magnetic field of 8.33 T. These electromagnets employ superconducting cable coils built from Niobium-titanium, kept

at a temperature of -271.3°C using liquid helium.

TABLE 3.1: Pre-Accelerators for the LHC.[65]

Pre-Accelerators	Energy	No. of RF cavity	No. of bunches	Remarks
LINAC 2	50 MeV	6	-	20 μsec pulse length
Proton Synchrotron Booster (PSB)	1.4 GeV	4	6 (per PS cycle)	4 rings and 1-2 bunch/ring
Proton Synchrotron (PS)	26 GeV	10	72	4 ns bunch length and 25 ns bunch spacing Filling with 2 PSB batches of 1.2 sec each
Super Proton Synchrotron (SPS)	450 GeV	16	288	25 ns bunch space Filling with 3-4 PS batches of 3.6 sec each

Overall, to achieve a target energy of up to 14 TeV at the LHC requires a highly sophisticated and precise system of superconducting magnets, RF cavities, pre-accelerators, and collimators, all working together to accelerate and guide the beams of particles around the collider and maximize the number of collisions.

3.1.1 Luminosity

The frequency at which a specific physics process occurs depends on its cross-section at a given center-of-mass energy and the instantaneous luminosity [66], which corresponds to the number of particles interacting within a time interval. To be more precise, the rate of a process, dN_{events}/dt, can be expressed as,

$$\frac{dN_{\text{events}}}{dt} = \mathcal{L}\sigma \tag{3.1}$$

where σ is the cross-section of the process and $\mathcal{L}$ is the instantaneous luminosity depending on the collider properties. The σ depends on the details of the interaction involved, it has units of area, and is generally expressed in units of barns ($cm^{-2}s^{-1}$). For new physics processes, we are targeting, cross-sections are of the the order of fb. Luminosity depends on a numerous factors, including the number of particles in each bunch (N_1 **or** N_2), total number of bunches (N_b), the frequency at which the bunches are collided (ν), the size of the collision area ($4\pi\sigma_x\sigma_y$) and reduction factor (F). This luminosity reduction factor (F) is used to overcome the electric repulsion between the proton bunches of the both pp beam causing beam-beam effect. It depends on total crossing angle (θ_c), transverse & longitudnal bunch length (σ_z).

$$\mathcal{L} = \frac{\nu N_b N_1 N_2}{4\pi\sigma_x\sigma_y}.F \ \ (\text{in units of cm}^{-2}\text{s}^{-1}), \qquad F = \frac{1}{\sqrt{1+(\theta_c\sigma_z/2\sigma_x)^2}} \tag{3.2}$$

where, N_1 or $N_2 = 1.15 \times 10^{11}$, $N_b = 2808$, σ_x or $\sigma_y = 16$ μm, $\sigma_z = 7.5$ cm, $\theta_c = 285$ μrad, for nominal LHC operations.

Beam size or collision area in luminosity formulae can also be expressed in terms of emittance and beta function. Amplitude function or β^* describes the focusing strength of the accelerator magnets and how much the beam is squeezed or spread out in the transverse direction. The ϵ is the transverse emittance, which is a measure of the spread of the particle positions and

momenta in the transverse direction.

$$\mathcal{L} = \frac{\nu N_1 N_2}{4\epsilon\beta^*}.F \quad (\text{in units of } \text{cm}^{-2}\text{s}^{-1}) \tag{3.3}$$

where, $\epsilon = 3.75\ \mu\text{m rad}$, $\beta^* = 0.55$ m

Having more focused beam can achieve higher luminosity but it can pose several challenges to accelerator as well as experiments. In general, space charge effect and beam-beam effect are the limitations for achieving higher instantaneous luminosity. Higher instantaneous luminosity increases the probability of multiple pp interactions in the same bunch crossing, resulting in more final state particles flux into the detectors.

3.1.2 The LHC Run II operations

The LHC has been in operation since 2009 and till now, it is operated at $\sqrt{s}$ of 7 or 8, 13 & 13.6 TeV in Run I, Run II, and Run III respectively. Presently, Run 3 data-taking has been going on, started on 2022 and the LHC is colliding protons with $\sqrt{s}$ of 13.6 TeV. In this chapter we will summarize the LHC operations in general and details of Run II.

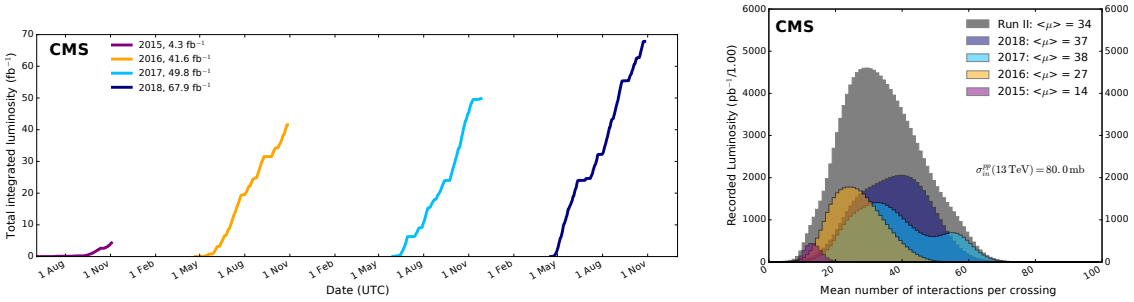

FIGURE 3.2: The integrated luminosity recorded at the CMS detector for different years during Run II (left) [67]. On the right side, the number of pileup distributions for all years (2015-2018) are overlaid, along with a grey shaded histogram representing the full Run II [68].

Run II of the LHC began in 2015 and ended in 2018. During this run, the LHC operated at a center-of-mass energy of 13 TeV, which is the highest energy ever achieved by a particle accelerator. During Run II, more than 150 fb^{-1} data from proton-proton collisions was recorded at the CMS detector as shown in the Figure 3.2 (left). This data is used to study the properties of the H boson and search for new physics beyond the Standard Model.

During collision of proton proton bunches, there will be primary vertex (PV), which is correspond to the hardest scattering in the event, evaluated using the tracker information alone. This primary vertex is required to be within the 24cm ($3\sigma_z$) from the center of the CMS detector in the beam (z) direction. The remaining reconstructed vertices are referred to as pileup vertices and correspond to additional pp interactions in the same bunch crossing. in Figure3.2, integrated luminosity recorded at the CMS detector is increasing with year (left) and higher luminosity will result into more pileup events or mean number of interactions per crossing as shown on the right. As a response to the increasing luminosity and the associated rise in

pileup events, the CMS collaboration has upgraded some of the detector component during the various end-of-year shutdowns during Run II. The CMS detector will be discussed in next section 3.2 in details.

3.2 The CMS detector

The CMS detector [69, 70] exhibits a cylindrical shape, designed to provide maximum 4π coverage, and it is composed of multiple layers of sub-detectors surrounding the beam axis as shown in figure 3.3. The CMS detector coordinate system has its origin at the nominal collision point. The y-axis points vertically upward and the x-axis points radially inward toward the center of the LHC as shown in the figure 3.3. The z-axis points along the beam direction. The azimuthal angle ϕ is measured from the x-axis in the x-y plane and θ is the angle between the x-y plane & z-axis. Pseudorapidity is related to θ as $\eta = \ln[\tan(\theta/2)]$. Particles produced in pp collisions in every bunch crossing traverse the various subdetectors and deposit energy. Given there is a bunch crossing every 25 ns, these detectors are required to have good timing response to collect the signal produced by the particles. The CMS experiment uses different detector technologies for its subdetectors to optimize efficiency and resolutios for different particles.
List of various components of the CMS experiment which are used to identify physics objects :

- A powerful **magnetic field** that assist in distinguishing between charged and neutral particles by bending the charged particle trajectories.
- A **fully silicon tracker** with high efficiency and a fine-grained structure that can accurately determine the position and momentum of charged particles.
- An **ECAL** with high segmentation, which enables the detection of energy deposits from particles such as electrons, photons (also photons from decay of π^0), and some fraction of charged or neutral hadrons.
- A **HCAL** with coarse segmentation where charged and neutral hadrons deposits their energies in the form of hadronic showers. Jets are the experimental signatures of hadrons made from quarks and gluons and will be explained in the section 3.4.4.
- A specialized **muon tracking** system outside the solenoid, to deliver reliable, redundant, and effective identification and reconstruction of muons as they pass through the system.

In the following sections, I give a brief description of various components of the CMS experiment.

3.2.1 The Solenoid Magnet

A superconducting solenoid magnet comprising of a niboium titanium coil stabilized by high-purity Al conductor, is operated at a current of 19.5 kA with zero electrical resistance [70]. The magnet is cooled by the circulation of liquid helium at a temperature of 4K and has a length of 12.5 m and a free-bore radius of 3.15 m, producing a uniform magnetic field (B) of 3.8 T.

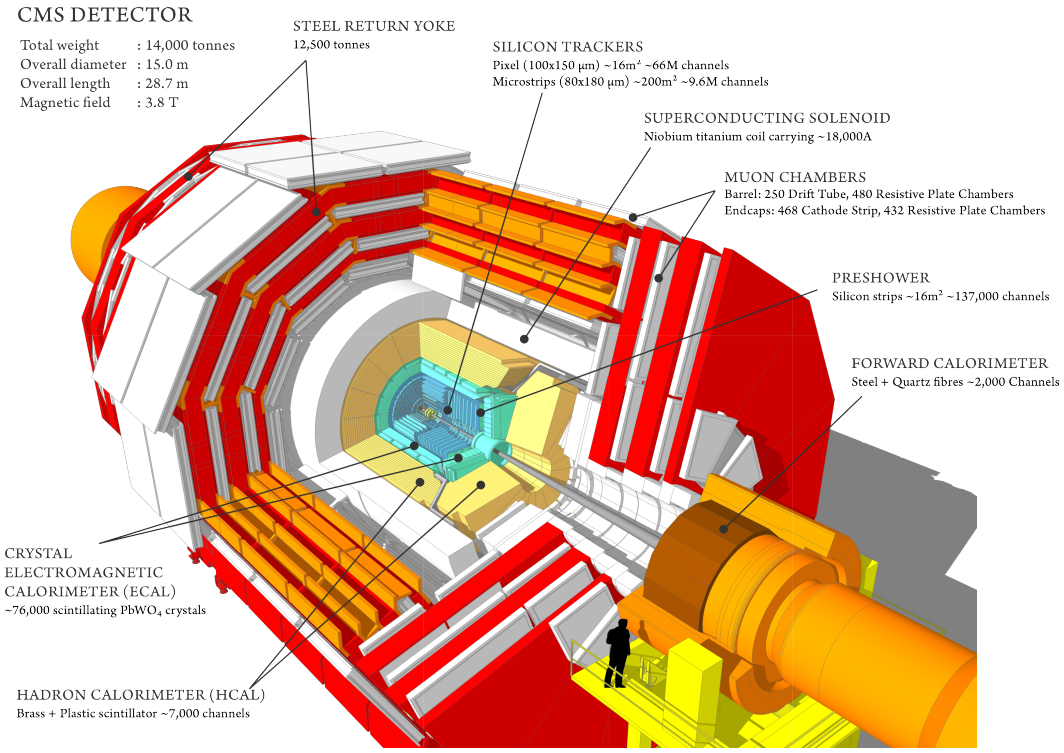

FIGURE 3.3: Sectional view of the CMS detector [69].

The magnet is designed to accommodate the tracker, the ECAL, and the HCAL, minimizing the amount of material in front of the calorimeters and reducing energy losses caused by particles showering in the coil material. The design of the detector heavily relies on the magnetic field configuration, which creates a strong and uniform magnetic field that deflects the path of charged particles depending on their transverse momentum (p_T). The magnet produces a bending power of 4.9 T m to the inner surface of the calorimeter system at normal incidence, providing a strong separation between charged and neutral-particle energy deposits. A charged particle with a p_T of 20 GeV is deflected by 5 cm in the transverse plane at the ECAL surface, which is a sufficient distance to distinguish the energy deposit of a charged particle from that of a photon emitted in the same direction.

$$p_T(GeV/c) = 0.3|q| \times B(T) \times r(m) \tag{3.4}$$

The CMS magnet is surrounded by a yoke made of iron composed of 5 wheels and 2 endcaps, with each endcap comprising three disks, which returns the magnetic flux and ensures a uniform magnetic field within the detector [71]. The iron has a high magnetic permeability, which means that it can easily conduct and return the magnetic flux produced by the solenoid. By completing the magnetic circuit in this way, the yoke ensures that the magnetic field lines remain confined within the detector's active region, resulting in a uniform and precisely controlled magnetic field for accurate particle tracking and measurement.

3.2.2 Tracking system

The CMS Tracker is a silicon-based particle detector that measures the positions and momenta of charged particles as they pass through the detector. The CMS tracker has a longitudinal length of approximately 5.8 m in the beam direction and is composed of two tracking devices utilizing semiconductor technology: the inner pixel and the outer strip trackers as shown in Figure 3.4.

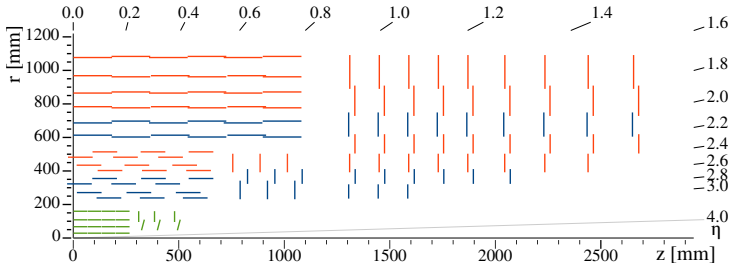

FIGURE 3.4: Schematic view for the first quadrant of the CMS tracker in r-z coordinates [72]. Layers with green color represents the inner pixel tracker and orange (blue) color means single (double) sided modules represents outer strip tracker. Single sided modules have only reverse p-n junction sensor whereas double sided modules have two silicon sensors mounted back to back [73].

During the 2016 LHC period, the innermost region of the barrel had three layers and the endcap had two disks of pixel detectors called the Inner Tracker (IT). During early 2017, the innermost region of the barrel was upgraded to have four layers and the endcap to have three disks covering maximum CMS radius (r) of 200 mm. The Silicon Strip trackers are divided into the inner barrel part (TIB), the inner disks (TID), the outer barrel (TOB) and outer endcaps (TEC). During early 2017, both TIB and TOB systems were upgraded such that they were composed of four and six concentric layer barrel shell structures respectively. The TID system is made of three disk structures on each side, each divided into three concentric rings, while the TEC is made of nine disk structures on each side, each made of four to seven rings. During the 2016 LHC running period, the silicon tracker consisted of 1440 silicon pixel and 15,148 silicon strip detector modules are finely divided into 66 million pixels with dimension of 150x100 μm and 9.6 millions strips in width from 80 to 180 μm. During 2017 and the 2018 LHC running periods, the number of silicon pixel modules increased to 1856 that are finely segmented into 124 million pixels while the number of silicon strip detector modules remained the same [74]. After upgrading trackers, the performance of tracker like tracking and hit efficiency (in Figure 3.5), impact parameter resolution, S/N ratio etc was much better especially in 2017 and 2018 data taking [75].

The tracker is located within a magnetic field of around 3.8 Tesla , which causes the trajectories of charged particles to curve. The charged particles adopt a helical trajectory, allowing for five key tracker parameters - the transverse radius of curvature, azimuthal angle, angle between track and beam line axis, distance between track and reference point in the transverse plane, and distance between track and reference point in the longitudinal plane - to be utilized by the track finding algorithm. By analyzing the curvature of a track, the charge and momentum of

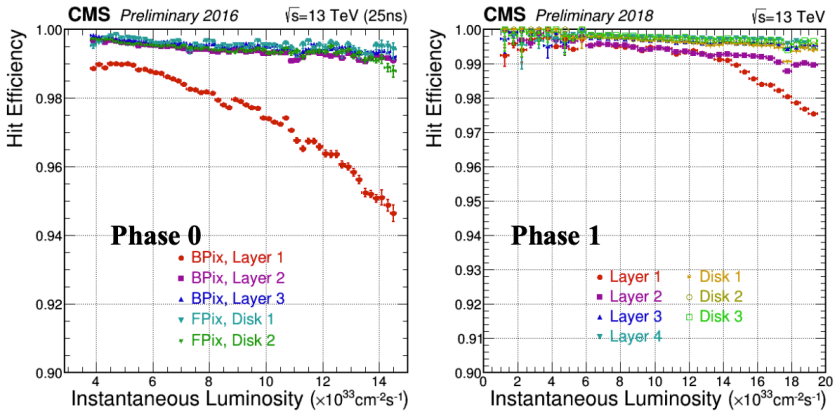

FIGURE 3.5: Hit efficiency over large range in luminosity for phase 0 (during 2016) and phase 1 (during 2018 upgrades) [75].

the particle is determined, as well as the location of the primary vertex where the particle is produced in a collision. To reconstruct primary or secondary vertices, the inner tracker must be situated in close proximity to the beam pipe, which exposes it to substantial radiation damage due to the intense particle flux within the pipe. To minimize energy loss due to multiple scattering, photon conversion, brehmstraalung, and nuclear interaction, the structural support is composed of lightweight material.

The momentum resolution of the CMS tracker depends on several factors, such as the number of hits per track, the detector material and geometry, and the track-fitting algorithm. The momentum resolution is typically expressed in terms of the p_T of charged particles as shown in equation 3.5.

$$\frac{\sigma(p_T)}{p_T} \propto \frac{p_T}{0.3BL^2} \tag{3.5}$$

For p_T values below 10 GeV, the momentum resolution is dominated by multiple scattering and is about 1-2%. For p_T values between 10 and 100 GeV, the resolution is dominated by the intrinsic detector resolution and is about 10-15%. For p_T values above 100 GeV, the resolution is degraded by the increasing amount of energy lost by charged particles before reaching the tracker, and it eventually reaches the energy resolution of the calorimeter.

3.2.3 Electromagnetic calorimeter

The ECAL is a homogeneous calorimeter made up of lead tungstate (PbWO4) [69]. It is composed of two main parts: the barrel (EB) and endcap (EE) regions. The barrel covers the central part of the detector ($|\eta| < 1.479$) having a length of 23 cm ($25.8X_0$) and is made up of 61,200 crystals arranged in a cylindrical shape. The endcaps cover the forward regions ($1.479 < |\eta| < 3$) having a length of 22 cm ($24.7X_0$) and are made up of 14,648 crystals each, arranged in a disk-like shape. The crystals are grouped into 36 supermodules, each covering half the length of the barrel having 20° in ϕ and each crystal inside a supermodule are arranged with longitudnal axis rotated by a 3° angle in both η and ϕ with respect to the interaction point. The

crystal material also has an interaction length of about 16.7 cm, which means that about two-thirds of the hadrons produced in collisions will start showering in the ECAL before entering the hadronic calorimeter (HCAL). This is important for accurately measuring the energy of hadrons, which make up a large fraction of the particles produced in high-energy collisions. The transverse size of the ECAL crystal is matched to the moliere radius of $PbWO_4$, which is 2.2 cm. There is also a preshower device ($3X_0$ of lead radiator) in front of ECAL endcap ($1.653 < |\eta| < 2.6$) to initiate an electromagnetic shower when an e or γ passes through it.

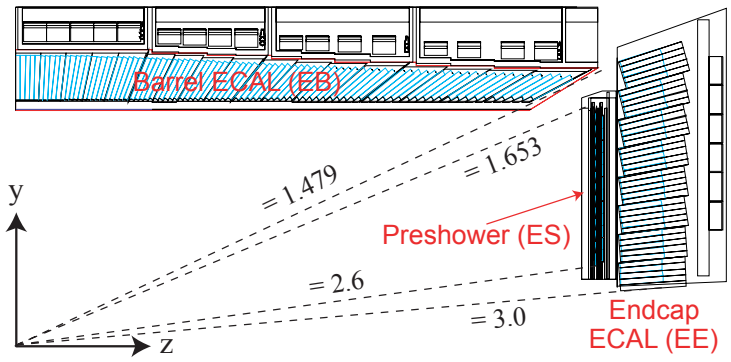

FIGURE 3.6: A schematic view showing only one quarter of the ECAL detector [70].

The ECAL is exposed to a high particle flux, comprising photons, electrons, and hadrons, particularly in the endcap region which is closer to the beam pipe. This exposure can result in radiation damage to the components of the detector, necessitating the use of lead tungstate crystals. These crystals possess a rapid response, with 80% of the emitted light produced within 25 nanoseconds, as well as resistance to radiation damage up to a level of 10 MegaRads. Silicon avlanche photodiodes (APDs) in barrel and vaccum phototriodes (VPTs) in endcap, are used to convert the scintillation light produced by the crystals to electronic signal. These photodiodes can also be damaged by radiation, leading to increased noise levels and decreased efficiency. To address this issue, the ECAL endcap is designed to function at a low temperature (-20°C) to minimize crystal transparency losses. Additionally, the photodetectors are operated at a high voltage to counteract any radiation-induced dark current.

The energy of the particles produced in the crystal is converted into light, which is detected by photo-detectors located behind the crystals. This light is then converted into an electrical signal which is shaped and amplified by a multi-Gain Pre-Amplifier, then digitized by 40-MHz ADC. This digitized signal is then stored after being processed by level 1 (L1) trigger primitives, which are used to identify interesting events that should be saved for further analysis (will be discussed in section 5.4). The energy resolution of the CMS electromagnetic calorimeter (ECAL) can be parameterized as follows:

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \tag{3.6}$$

where σ_E is the energy resolution, E is the energy of the electromagnetic shower, and S, N, and C are constants representing the stochastic, noise, and constant terms, respectively. For

the barrel (endcap) region of the ECAL, the values of these constants are:

$$S = 2.8\%\ \text{GeV}^{-1/2}, \quad N = 12\%\ (25\%)\ \text{GeV}, \quad C = 0.3\% \tag{3.7}$$

3.2.4 Hadron calorimeter

Hadron calorimeter (HCAL) is a sampling detector, which is designed to measure the energy of hadrons as well as p_T^{miss} , produced in the high-energy proton-proton collisions [69]. The HCAL is composed of four main parts: the barrel (HB), endcap (HE), outer (HO) and forward (HF) regions. The HB and HE are located behind the tracker and electromagnetic calorimeter. The barrel within $|\eta| < 1.3$, is restricted radially between the electromagnetic calorimeter (R = 1.77 m) and magnet coil (R = 2.95 m) which limits the material for absorbing hadronic showers. In both HB & HE, there are 18 wedges made from brass and segmented into four ϕ sectors and fully covering the beamline azimuthally with 5° by each ϕ sector. The HB (HE) use brass as an absorber material due to its small nuclear interaction length of 16.4 cm, and plastic scintillator as an active material. Each absorber wedge is interleaved with trays of scintillator tile to form a sampling calorimeter and these scintillators in HB (HE) are longitudinally segmented into 16 (14) η sectors as shown in figure 3.7, each having a dimension of 0.087×0.087 (0.17×0.17) in the $\Delta\eta \times \Delta\phi$ phase space. The effective thickness of the these wedges will vary from 5-10 λ_I depending on $1/\sin(\theta)$, where θ is polar angle.

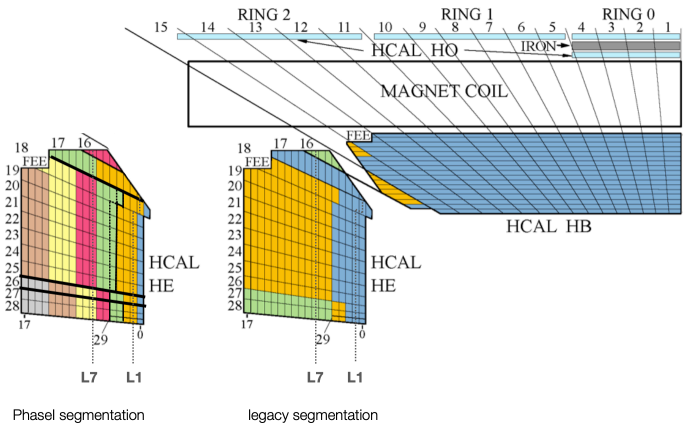

FIGURE 3.7: A longitudnal view (r-z plane) showing HCAL detector during the 2016 LHC Run 2 (right) and phase 1 segmentation of the HE (left) which is upgraded during the 2018 LHC Run 2. The numbers on top and on left refer to the tower numbers. The numbers on the bottom (0–17) indicate the scintillator layers numbers inserted into slots in the absorber [76].

The HO calorimeter or tail catcher is placed outside the solenoid to complement the barrel calorimeter within $|\eta| < 1.26$ for providing sufficient containment of hadronic showers. It roughly maps the layers of HB to make towers of granularity 0.087×0.087 in the $\Delta\eta \times \Delta\phi$. The HO increases the effective thickness of the detector by 2-3 interaction length and hence reduces shower leakage, and improve the p_T^{miss} measurement which is significant for searches of supersymmetric particles. The HF calorimeters are placed at 11.2 m from the interaction

point to extend the coverage down to $|\eta| = 5.2$ using a Cerenkov-based, radiation-hard technology. The HF is required to provide high-precision energy measurements of hadronic jets and missing transverse energy in the forward region which is facing extremely high particle flux. The HO uses the solenoid and iron yoke as absorber material while HF uses steel as the absorber and quartz fibre as the active medium for radiation hardness.

When a hadron passes through the HCAL, it interacts with the atoms in the material and produces a shower of lower-energy particles. These particles in turn produce scintillation light, which is detected by photo-detectors located at the end of each scintillator tile. The amount of scintillation light produced is proportional to the energy of the original hadron, allowing the HCAL to measure the energy of these particles with high precision. Hybrid photodetectors (HPDs) were used in HB and HE in 2016 & 2017 operations, whereas photomultiplier tubes featuring dual output were incorporated in HF providing better discrimination of real hit from muon hitting the PMT window. Subsequently, in 2018, silicon photomultipliers (SiPMs) were integrated into HE, resulting in a photo-detection efficiency that is 2.5 times greater than before. As shown in Figure 3.7 for legacy segmentation, scintillator layers are marked in blue are grouped together as depth 1 and lead to the same photodetector, while the ones in yellow, green, and magenta are combined as depths 2, 3, and 4, respectively. Based on the excellent experience by replacing HPDs with SiPms in HE, more depth segmentation are introduced afterwards (in Run3) to provide better calibration of depth-dependent effects and bring some physics benefits like long-lived particle detection, as shown in fig 3.7 for phase1 segmentation. The energy resolution of the CMS combined calorimeter (ECAL + HCAL) can be parameterized as follows:

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{110\%}{\sqrt{E}}\right)^2 + (9\%)^2 \tag{3.8}$$

, where E is expressed in GeV.

3.2.5 Muon detectors

The CMS magnet has a specialized muon tracking system that is integrated into the return yoke, outside the solenoid. This system has a crucial role of extending the range of tracking measurements by several meters beyond the interaction point to deliver reliable, redundant, and effective identification and reconstruction of muons as they pass through the system, along with the capability to trigger events [69, 77]. The muon detector system at the CMS is divided into four main subsystems:

- **Drift tubes** (DT): The DTs are located in the barrel region of the muon system, covering the pseudorapidity range $|\eta| < 1.2$. There are four layers of DTs, labeled as MB1-MB4 (shown in fig 3.8), arranged in concentric cylinders around the beam line. Each layer consists of 12 chambers, with each chamber containing 60 drift tubes. Each DT consists of a 4.2 m long and 42 mm cylindrical tube filled with a mixture of gases, typically argon (85%) and carbon dioxide (15%). When a muon passes through the tube, it ionizes the gas, creating free electrons and positively charged ions. The electric field in the tube

causes the electrons to drift towards a thin wire at the center of the tube. As the electrons reach the wire, they create a pulse of electrical current that can be detected and used to measure the position of the muon. The spatial resolution of the DTs is around 200 microns, with a time resolution of around 3 ns. The DTs provide good muon momentum resolution, especially for muons with p_T below 100 GeV/c. The efficiency of the DTs is greater than 97% for muons with p_T above 20 GeV/c.

- **Cathode strip chambers** (CSC): The CSCs are located in the endcap regions of the muon system, labeled as ME1-ME4, covering the pseudorapidity range $0.9 < |\eta| < 2.4$. There are 468 CSCs in total, arranged in four layers in each endcap region. Each CSC consists of six layers of positively charged anode wires and cathode strips. The anode wires are held at high voltage and the cathode strips are grounded, creating an electric field that drifts the ionization electrons produced by the passage of a muon towards the strips. The position of the muon can be determined by measuring the electrical signal induced on the cathode strips. The CSCs provide a spatial resolution in between 45 - 145 microns and a time resolution of around 7 ns. They have high efficiency ($> 95\%$) for detecting muons with p_T above 10 GeV/c.

- **Resistive plate chambers** (RPC): RPCs are mounted in both barrel (RB1-RB4) and endcap (RE1-RE4) region. Each RPC consists of two parallel resistive plates separated by a few millimeters of gas, operate in avalanche mode to generates a fast and intense signal that can be detected and used for precise timing measurements (<25 nsec). The RPCs are particularly useful in the high-luminosity environment of the LHC, where collisions occur at a rate of up to 40 million per second. During the transition from Run 1 to Run 2, the additional chambers (ME4/2, RE4/2 & RE4/3) were installed in the ME4 and RE4 regions to increase redundancy within the system, thereby improving its efficiency while simultaneously reducing misidentification rates.

Each subsystem is optimized for a different range of muon energies and angles. The DTs are located in the central region of the muon system and are designed to measure the positions of muons that pass through them. The CSCs are located in the endcap regions of the detector and are designed to provide high-precision muon tracking in the forward and backward directions. The RPCs are located in both the barrel and endcap regions and provide a fast trigger for the CMS detector. During Run II of the LHC (2016-2018), the CMS muon detector system underwent significant upgrades to improve its performance. These upgrades included the installation of additional muon chambers (CSC station ME4/2 and RPC station RE4/2 RE4/3), improvements to the electronics and readout systems, and triggers as part of the CMS L1 trigger upgrade [78]. Role of triggers and their working will be discussed in section 3.3. Overall, the muon detector system at the CMS played a critical role in the success of the Run II physics program, enabling the detection and study of a wide range of physics phenomena.

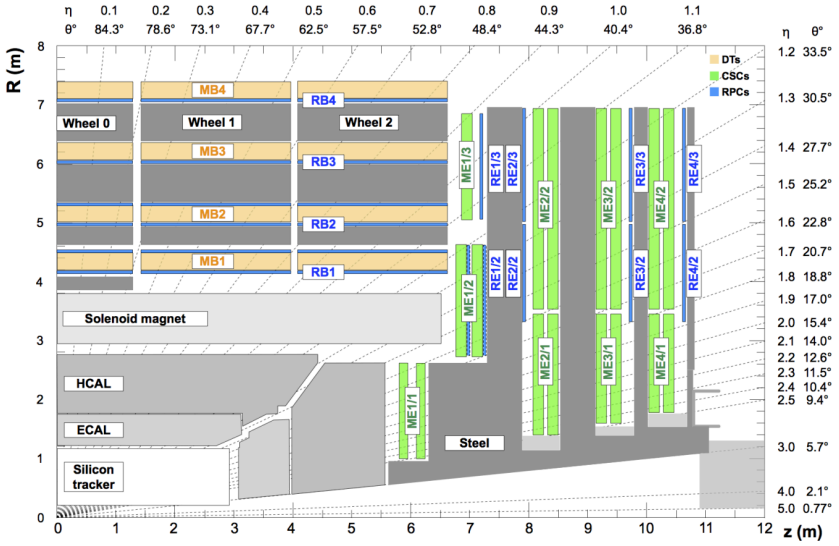

FIGURE 3.8: A schematic r-z view of the CMS detector showing muon system for Run 2 in details. It includes DTs (MB), CSCs (ME), and RPCs (RB or RE), which are described by respective colours in the legend. Here M denotes Muon, B stands for Barrel and E for Endcap [77].

3.3 Triggers

The CMS detector is designed to precisely measure particles like leptons, photons, and jets during proton-proton at the LHC, CERN. These collisions occur with high-energy protons at 13.6 TeV and a luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$. Since there are 10^9 interactions per second at the LHC, only a small portion of collisions are relevant to the CMS physics program. Thus, the trigger systems are required to select these events for offline analysis. The CMS trigger [79] is composed of two levels: the first L1 and the high-level trigger (HLT).

- The **L1 triggers** select events that contain high energy deposits in the calorimeters and/or muon candidates measured in muon systems. The final event selection is based on a programmable menu, and events that pass to the HLT. The L1 thresholds are adjusted during data taking to restrict the output rate to 100 kHz. The L1 trigger system is composed of four major components: the Calorimeter Trigger (CALO), the Muon Trigger (MUON), the Global Trigger (GT), and the Timing Trigger (TT). The CALO trigger receives input from the electromagnetic and hadronic calorimeters, while the MUON trigger receives input from the muon detectors. In the starting of 2016 Run, muon triggers are upgraded such that hit and segment information from various muon chambers passes to the corresponding track finder like BMTF (Barrel Muon Track Finder), OMTF (Overlap Muon Track Finder) and EMTF (Endcap Muon Track Finder), which takes input depending on the $|\eta| < 0.83$, $0.83 < |\eta| < 1.24$ & $|\eta| > 1.24$ respectively. This upgrade improves the trigger rate by 2 factor and also improves the trigger efficiency. In the first layer of calorimeter trigger system, local energy deposits known as trigger primitives (TPs) received from ECAL and HCAL are calibrated and forwarded to the layer 2, where they

are utilized to reconstruct and calibrate various physics objects such as electrons, tau leptons, and jets. Finally, the GT combines the input from the CALO and MUON triggers and decides whether to accept or reject an event. After that, TT provides precise timing information to ensure that the L1 trigger system is synchronized with the LHC beam structure. After passing through the L1 trigger as shown in the schematic diagram 3.9, the data from the sub-detectors is combined by the Data Acquisition (DAQ) system to form a complete event within $4\mu s$. This event data is then passed to the HLT, which uses various algorithms like PF to further analyze the data and make a final decision on whether the event should be stored for further analysis.

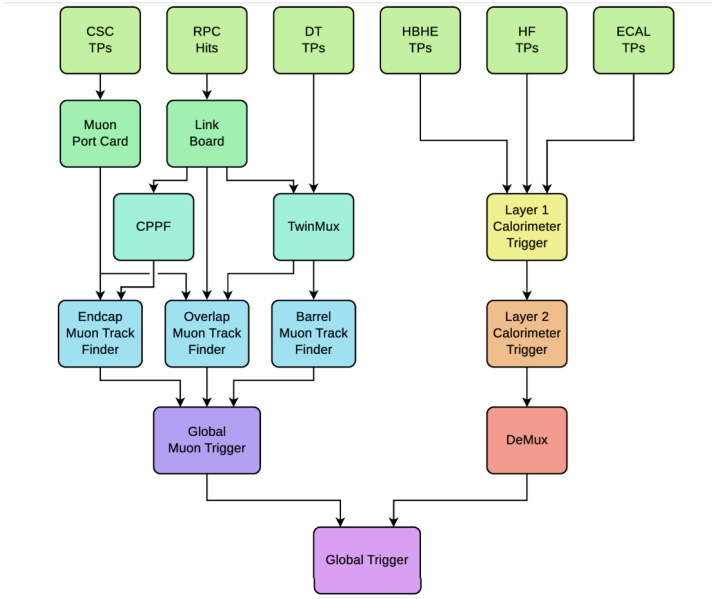

FIGURE 3.9: Overview of the CMS L1 trigger system for Run II [80].

- The **High-Level Trigger** is the second stage of the CMS trigger system, following the Level-1 (L1) trigger. The HLT is responsible for further reducing the event rate from the L1 trigger and selecting the most interesting events for further analysis. During Run II of the LHC, the CMS HLT used a variety of algorithms to identify interesting events based on the kinematic properties of particles produced in collisions. Similar to the offline processing, the HLT performs event selection by reconstructing objects like electrons, muons, and jets for each event. The identified objects are subjected to identification criteria to choose only the events that may have potential interest for the CMS physics analysis program. The algorithms and object performance involved in HLT are similar to offline but not exactly same. The technique of prescaling enables the selection of a fraction of events based on specific criteria, resulting in the saving and recording of only one out of every "N" events onto disk. By altering the prescale scheme during data acquisition, the HLT output rate can be maintained at a near-constant level of about 1 kHz, regardless of the instantaneous luminosity levels generated by the LHC.

3.4 Event reconstruction

Event reconstruction uses particle flow (PF) as a primary algorithm [81], in which information from multiple detectors is combined to identify and measure the energy, momentum and position of each individual particle as shown in Figure 3.10. The PF algorithm in the CMS starts by reconstructing tracks of charged particles in the tracker, which provides precise measurements of the particle trajectory and momentum. The tracks are then extrapolated to the calorimeters, where the energy deposited by the particles in the ECAL and HCAL is measured. Neutral particles, such as photons and neutral hadrons, which do not leave tracks in the tracker, are identified by energy clusters in the calorimeters. The energy deposits from the charged particles are then associated with the charged tracks in the event. Finally, the muon detectors are used to identify muons, which are high-energy charged particles that can penetrate through several layers of the CMS detector. The momentum of the muons is measured using a combination of the tracker and the muon detectors. The PF reconstruction involves dedicated detector-level calibrations for PF hadrons energy cluster as discussed in Chapter 4.

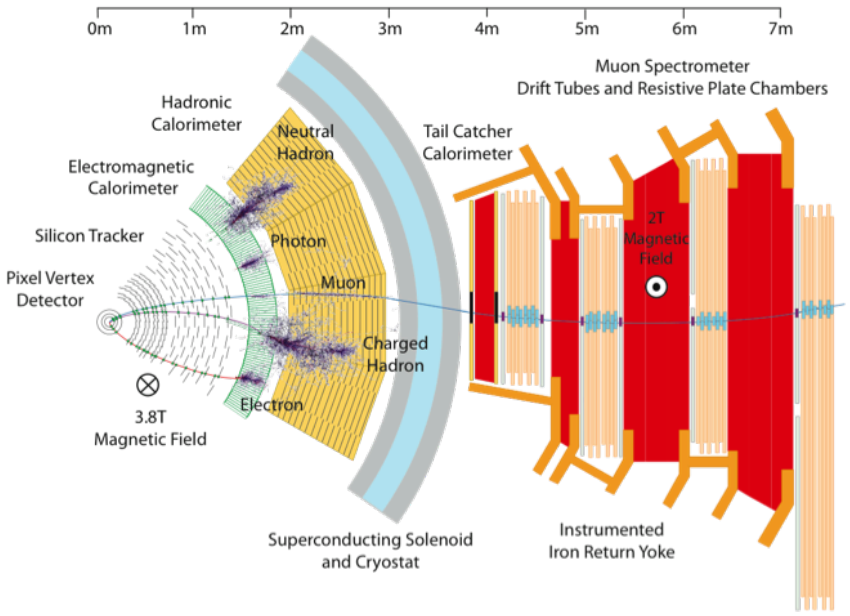

FIGURE 3.10: The CMS transverse slice showing interaction of electron, muons, photons, charged and neutral hadrons with various sub-detectors [82].

3.4.1 Tracking

This is the first step of PF algorithm to determines the charged particle trajectories. The CMS tracker has pixel and strip detector that provides excellent momentum and position resolution as discussed in 3.2.2. A combinational track finder (CTF) algorithm is used and it is based on

kalman filter (KF) for reconstructing the tracks in three stages: initial seed generation, trajectory building and final fitting.

- The **seed generation** algorithm starts by selecting a pair of hits in the pixel detector that are consistent with a straight line trajectory. These two hits are referred to as the seed hits. The algorithm then searches for compatible hits in the first two layers of the silicon strip trackers that are consistent with the expected trajectory of the charged particle. In 2017, the tracker upgrades led to a more efficient seed generation process with the introduction of the "cellular automation (CA)" technique. The CA technique simplifies seed generation by utilizing only pixel hits and identifying triplets or quadruplets from the compatible pairs, checking their compatibility with the beam-spot.
- During the **trajectory building** stage, the initial track candidates are iteratively extended by adding compatible hits at each layer of the tracker. KF algorithm is used to estimate the track's state vector and covariance matrix based on the available hits, and the estimated state vector is propagated to the next layer to look for compatible hits.
- Once all compatible hits have been added to the track, the final set of associated hits are used to perform a final fit to the track. This process is known as **final fitting** and uses a KF or a similar algorithm to optimize the track parameters, such as its momentum, direction, and position at the point of closest approach to the interaction point. In case of electron track, there is a large probability of radiating photon in the tracker material and there be energy loss by the electron while radiating photon, so Gaussian-sum filter (GSF) is used to for selected tracks on top of the χ^2 fit. The GSF can be considered as a nonlinear extension of the KF, consisting of a weighted combination of multiple KFs.

This CTF algorithm is applied iteratively to increase the tracker efficiency and at each step, mis-reconstruction rate is reduced with some quality selection on track seeds, track fit (χ^2) etc as shown in Table 3.2. The hits that are associated with the selected tracks are masked to reduce the likelihood of random hit-to-seed associations in the subsequent iterations.

TABLE 3.2: Seeding configuration and targeted tracks of the ten tracking iterations for 2016 Run. In the last column, R is the targeted distance between the track production position and the beam axis [81].

Iteration	Name	Seeding	Targeted Tracks
1	InitialStep	pixel triplets	prompt, high p_T
2	DetachedTriplet	pixel triplets	from b hadron decays, $R < 5$ cm
3	LowPtTriplet	pixel triplets	prompt, low p_T
4	PixelPair	pixel pairs	recover high p_T
5	MixedTriplet	pixel+strip triplets	displaced, $R < 7$ cm
6	PixelLess	strip triplets/pairs	very displaced, $R < 25$ cm
7	TobTec	strip triplets/pairs	very displaced, $R < 60$ cm
8	JetCoreRegional	pixel+strip pairs	inside high p_T jets
9	MuonSeededInOut	muon-tagged tracks	muons
10	MuonSeededOutIn	muon detectors	muons

3.4.2 Electrons and Photons

The reconstruction of electrons and photons is crucial for many physics analyses, and the accuracy of the reconstruction is critical for the discovery of new particles and interactions beyond the SM. Electrons or photons can interact with material and emit bremsstrahlung photons or convert into an electron-positron pair while interacting with detector material. Electron reconstruction [83] in the CMS starts with the detection of clusters of energy deposits in the electromagnetic calorimeter (ECAL). These clusters are then matched with tracks measured by the tracker to identify electrons, which produce a track in the tracker and a cluster in the ECAL (brief details for these sub-detectors can found at section 3.2.2 & 3.2.3).

Identification

Identifying prompt and isolated electrons and photons, and separating them from background sources, in the CMS detector is a challenging task. For prompt electrons, there are various sources of non-prompt electrons like those overlapping with hadrons like pions and kaons in jets, and secondary electrons that arise from semileptonic decays of b or c quarks. Furthermore, electrons can also be produced in the early stages of photon conversions in the tracker material, which can further contribute to the background, if one of the e^+ or e^- is lost or very soft. The main source of non-prompt photons arises from jets that mainly produce light neutral mesons, such as π^0 or η^0, which promptly decay into two photons. For higher energies, these light mesons are highly boosted, resulting in the two photons from the decay being nearly collinear. This makes it difficult to distinguish them from a single photon incident on the calorimeter.

A boost decision tree (BDT) is utilized to determine if a track should be associated with an ECAL cluster. The BDT considers track information, supercluster observables, and track-cluster matching variables. The electrons or photons are subject to a relaxed requirement on the BDT classifier, where isolated and non-isolated electrons or photons are evaluated with different BDTs using various input variables such as shower-shape, detector-based isolation, and tracker-related variables. Additional selection requirements are placed on E/p (the ratio between electron energy and momentum), H/E, and quantities based on associated generic tracks to eliminate problematic candidates.

A detailed description of some electron and photon identification criteria is given below.

- **Isolation variable**: The isolation variable is used as a discriminator to distinguish prompt photons and electrons, which originate from the primary interaction, from background sources such as jets, where particles tend to cluster around each other. To obtain the isolation variables for electrons and photons, the p_T of charged hadrons (I_{ch}), photons (I_γ), and neutral hadrons (I_n) are added up within a cone of radius ΔR = 0.3 centered on the direction of the particle of interest, excluding the corresponding electron or photon itself. Isolation variables are corrected by subtracting ρA_{eff} from the isolation quantities to mitigate the contribution of pileup in the isolation region. Here, ρ is the median of transverse energy density per unit area in the event, A_{eff} is the area of the isolation region weighted by a factor that accounts for pileup transverse energy density dependence on the object

η. In case of electron identification, relative combined isolation variable used, which is defined as:

$$\frac{I_{combined}}{E_T} = \frac{I_{ch} + max(0, I_n + I_\gamma - \rho A_{eff})}{E_T} \tag{3.9}$$

The radius of the cone for obtaining isolation variable is 0.2 for lepton $p_T < 50$ GeV, 10 GeV p_T for $50 < p_T < 200$ GeV, and 0.05 for $p_T > 200$ GeV. The decreasing radius of the cone with lepton p_T is motivated by the increased collimation of the decay products from the lepton's parent particle with increasing Lorentz boost and helps to retain high lepton isolation efficiency in events with large number of jets or pileup interactions. The isolation variable is required to be less than 0.15 (0.2) for electrons (muons).

- **H/E**: The H/E ratio is calculated by dividing the energy deposited in the HCAL around the super cluster (SC) direction by the energy of the electron or photon candidate in a cone of radius with ΔR = 0.15. The three main sources contributing to the measured hadronic energy of an electromagnetic object are HCAL noise, pileup, and leakage of electron or photon showers through the inter-module cracks. To account for both low and high-energy regions, the H/E selection requirement is in the form $H < X + Y\rho + JE$, where X and Y represent noise and pileup terms, and J is a scaling term for high-energy electrons and photons.

- $\sigma_{i\eta i\eta}$: It is the second moment of the log-weighted distribution of crystal energies along the η direction. It is computed within a 5×5 matrix centered on the crystal with the highest energy deposition in the supercluster. The resulting value is then scaled to units of crystal size. It is expressed as:

$$\sigma_{i\eta i\eta} = \sqrt{\frac{\sum_{5\times5} w_i \cdot (\eta_i - \eta_{5\times5})^2}{\sum_{5\times5} w_i}} \tag{3.10}$$

where η_i represents the pseudorapidity of the ith crystal, $\eta_{5\times5}$ is the pseudorapidity of the crystal with the highest energy deposition in the 5x5 matrix, w_i is the weight assigned to the ith crystal. Weight for each crystal is defined as $max(0, 4.7 + \ln E_i/E_{5\times5})$, so that the crystals in 5×5 matrix have 90% of $E_{5\times5}$. This shower shape variable is used to differentiate between electron and photon showers with $\pi^0 \to \gamma\gamma$ showers.

- The **conversion veto** is a technique used in the context of photon identification to distinguish between genuine photons and electrons that may be misidentified as photons. It aims to reject photons that originate from the conversion of a photon to an electron-positron pair within the detector material. This is achieved by checking the hit in the innermost layer of the pixel detector that is not associated with a reconstructed conversion vertex pointing to the photon cluster in the ECAL. Additionally, this process enhances the purity of identified photons. On the other hand, there is **pixel seed veto** which is required to reject any photon if a pixel detector seed consisting of at least two hits and pointing to the ECAL within some window defined around the photon SC position.The conversion-safe veto is used when electron backgrounds are negligible, whereas the pixel

detector seed veto is applied in the presence of significant electron misidentification as photons.

There are different working points (WP) like identification efficiencies of approximately 70% (tight WP), 80% (medium WP), and 90% (loose WP) for prompt photons or electrons. For identifying photons, cut-based loose working point requirements are shown in Table 3.3. In case of electrons, there is veto WP with 95% identification efficiency in addition to other WPs as shown in Table 3.4 and it is used in particular analysis where events with reconstructed electrons need to be rejected. If there is lower background for either electron or photon, then tighter working point is used and medium or loose WPs are preferred when there is higher background. Due to current limitations in computing resources, running the PF algorithm with multiple e/γ identification requirements is impossible. As a result, a common "loose" identification selection is utilized for electrons and photons (section 4.5 of reference [83]).

TABLE 3.3: Cut-based photon identification requirements for the loose working point in the EB and EE. [84]

Variable	Barrel (loose WP)	Endcap (loose WP)
H/E	< 0.04596	< 0.0590
$\sigma_{i\eta i\eta}$	< 0.0106	< 0.027
I_{ch}	< 1.694 GeV	< 2.089 GeV
I_n	$< 24.032 \text{ GeV} + 0.015E_T + 2.26 \times 10^{-5}E_T^2$	$< 19.722 \text{ GeV} + 0.012E_T + 3.45 \times 10^{-5}E_T^2$
I_γ	$< 2.876 \text{ GeV} + 0.0040E_T$	$< 4.162 \text{ GeV} + 0.0037E_T$

TABLE 3.4: Cut-based electron identification requirements for the veto working point in the EB and EE. [85]

Variable	Barrel (veto WP)	Endcaps (veto WP)
$\sigma_{i\eta i\eta}$	< 0.0126	< 0.0457
$\lvert\Delta\eta_{\text{seed}}\rvert$	< 0.00463	< 0.00814
$\lvert\Delta\phi_{\text{in}}\rvert$	< 0.148 rad	< 0.19 rad
H/E	$< 0.05 + 1.16\frac{\text{GeV}}{E_{\text{SC}}} + 0.032\frac{\rho}{E_{\text{SC}}}$	$< 0.05 + 2.54\frac{\text{GeV}}{E_{\text{SC}}} + 0.183\frac{\rho}{E_{\text{SC}}}$
I_{combined}/E_T	$< 0.198 + 0.51\frac{\text{GeV}}{E_T}$	$< 0.203 + 0.963\frac{\text{GeV}}{E_T}$
$\lvert 1/E - 1/p\rvert$	$< 0.209\frac{\text{GeV}^{-1}}{E}$	$< 0.132\frac{\text{GeV}^{-1}}{E}$
Number of missing hits	≤ 2	≤ 3
Pass conversion veto	Yes	Yes

3.4.3 Muons

The muon reconstruction algorithms [77] use information from different layers of chambers to reconstruct the muons trajectories and determine their momentum and energy. These algorithms during Run II were based on a combination of several different techniques, including the KF and the global muon reconstruction algorithm. These techniques are designed to optimize the reconstruction of muons in different regions of the detector depending on the energy

and momentum of the muons and helps to improve the muon sample's overall purity, which is important for studying rare processes and reducing background in searches for new physics. Here is a detailed overview of the different techniques used for muon identification and reconstruction at the CMS:

- **Segment matching**: The first step in muon identification is to match the hits in the muon system sub-detectors (DT, CSC, and RPC) as described in 3.2.5 to form segments, which are straight-line track segments in the muon system. The segments are then matched to tracks in the silicon tracker, which helps to reduce the amount of background from particles that are not muons.
- **Tracker-muon identification**: Muons can also be identified by matching tracks in the silicon tracker to segments in the muon system. This technique is called "tracker-muon identification". It is used for muons that lose energy as they pass through the detector material and do not leave sufficient hits in the muon system. Tracker-muons must be consistent with the muon hypothesis, meaning their energy loss is consistent with that of a muon passing through the detector material.
- **Standalone-muon identification**: Muons that do not have tracks in the silicon tracker can still be identified as "standalone-muons" by requiring at least two segments in the muon system that are consistent with being produced by a muon. Standalone-muons must be isolated from other particles in the event and have a minimum p_T .
- **Global-muon identification**: The most powerful technique for muon identification is called "global-muon identification" and combines the information from the muon system and the silicon tracker to form a complete track for the muon. The global-muon algorithm requires that the track be consistent with the muon hypothesis, match segments in the muon system, and have hits in the silicon tracker consistent with a muon.
- **Isolation**: To ensure that the identified muons are coming from the proton-proton collisions and not from other sources, such as cosmic rays, they must be isolated from other particles in the event. The isolation is measured by calculating the amount of energy deposited around the muon track in the other CMS sub-detectors and requiring that it be below a certain threshold.

Overall, the muon reconstruction performance at the CMS during Run II is excellent, with high efficiency and good resolution for measuring muon momentum and energy.

3.4.4 Jets

Jets are the collimated sprays of particles, which carry the kinematic information about the initial quarks and gluons that were present in the collision. In the CMS experiment, jet reconstruction is performed by clustering reconstructed PF particle candidates, like charged hadrons (65%), neutral hadrons (10%), and photons (25%). These PF particles are used as input for the jet clustering process using the infrared and collinear safe anti-k_T algorithm [86, 87] with a radius (R) parameter of 0.4 (0.8), referred to as AK4 (AK8) jets. The distance parameter between

two particles (d_{ij}), used in anti-k_T algorithm is defined as follows:

$$d_{ij} = \min\left(p_{Ti}^{-2}, p_{Tj}^{-2}\right)\frac{\Delta_{ij}^2}{R^2} \tag{3.11}$$

$$\text{where, } \Delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2 \tag{3.12}$$

Distance between ith particle and the beam,

$$d_{iB} = p_{Ti}^{-2} \tag{3.13}$$

PF hadronss, electrons & photons are calibrated before clustering into PF Jets. Details for PF charged hadron calibrations are discussed in chapter 4. The energy of the reconstructed jets is corrected to account for detector effects and to ensure that the energy scale is uniform across the detector. These corrections are done at different levels (L1, L2 ,L3) and are derived using data and simulations both as shown in Figure 3.11.

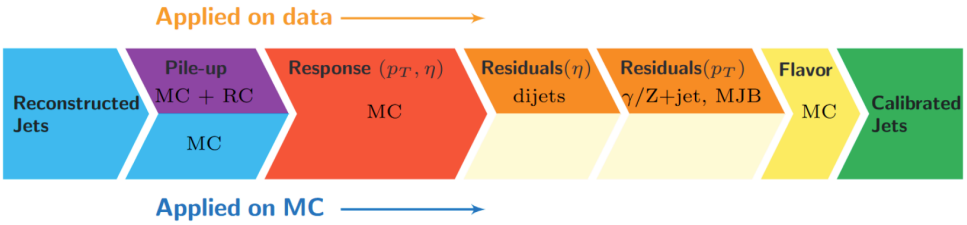

FIGURE 3.11: This is pictorial representation of the jet energy correction levels.[88]

The first level ("L1") of jet energy corrections accounts for "pileup" subtraction and contribution of the pileup to the jet momentum is mitigated. By discarding all the charged-particle tracks associated with pileup vertices and applying an offset correction to mitigate the average contribution of neutral particles [89, 90]. For mitigating the effect of pileup interactions to the AK8 jet momentum, a pileup per particle identification (PUPPI) algorithm [91] is used which makes use of local shape information to distinguish particles originating from hard scatter and pileup interactions.

Second and third levels of corrections ("L2L3") account for the non-uniformity of the detector response across the jet p_T and η ranges, jet energy corrections (JECs) are derived from the simulation in order to make the response of reconstructed jets equal to the particle-level jets on average. Residual differences in response between data and simulation are corrected based on dedicated measurements of the momentum balance in dijet, γ + jet, Z + jet, and multijet events. The jet energy resolution (JER) in the simulation is also modified by smearing the jet p_T, using the scale factors derived from data. Jets potentially dominated by contributions from anomalous detector signals or reconstruction failures are discarded using dedicated jet identification criteria.

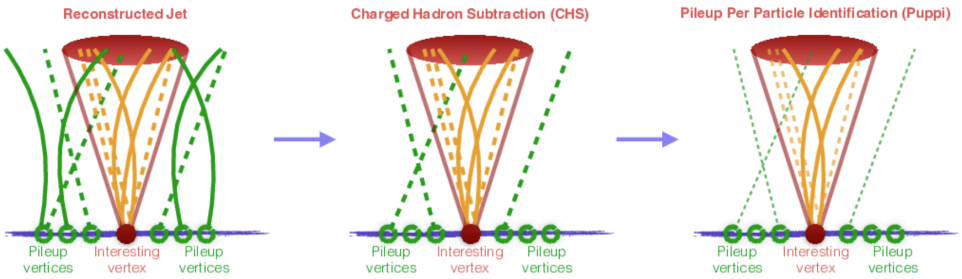

FIGURE 3.12: The Figure illustrates an outline of techniques employed for pileup (PU) suppression. In the figure, solid lines represent charged PF candidates, while dashed lines represent neutral PF candidates. The PUPPI algorithm applies weights to the PF candidates, which are represented by thin lines in the diagram [92].

3.4.5 b-tagged Jets

The identification of b-jets is based on the fact that b-hadrons (like pions) have a relatively long lifetime (≈1.5 ps) compared to other quarks, so they travel a short distance before decaying. B-hadrons with 20 GeV of p_T will have decay length of ≈1.8 mm and it leads to the production of secondary vertex (SV) within the jet as shown in Figure 3.13, which are identified by the CMS tracking system. The displacement of tracks from the primary vertex (PV) is determined by their impact parameter (IP), which quantifies the distance between the primary vertex and the tracks at their closest points of approach. The impact parameter can be described in three-dimensional space (3D) or in the plane transverse to the beam line (2D). Additionally, there is a longitudinal impact parameter defined along the beam line in a single dimension. Heavy-flavor hadrons (produced through the hadronization of b or c quarks) have higher mass and harder fragmentation than light quarks and gluons. As a result, their decay products tend to have higher p_T relative to the jet axis compared to other jet constituents. Approximately 20% (10%) of heavy b (c) hadron decays involve a muon or electron. Hence, charged leptons are used alongside properties of reconstructed secondary vertices or displaced tracks for heavy-flavor jet identification and performance evaluation using experimental data.

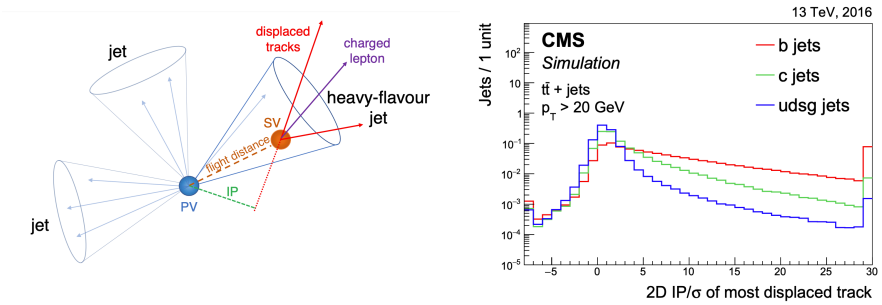

FIGURE 3.13: Pictorial representation of the secondary vertex in the heavy flavored jet (left) and 2D IP significance distribution of the most displaced track (right) for different jet flavours. [93]. 2D IP/σ plot shows the higher IP value for b-jets.

In addition to tracking variables like IP, secondary vertex reconstruction algorithm utilizes the

presence of secondary vertices formed by the decay of B hadrons within a jet. It examines parameters related to the secondary vertex, such as its mass, flight distance, and the number of associated tracks, to determine the b-tagging likelihood. Jets originating from b-quarks are identified by different tagging algorithms [93]:

- Jet probability tagger: This algorithm evaluates the probability that the tracks associated with the secondary vertex originate from the primary vertex. It calculates a likelihood ratio based on the IP significance and the track compatibility with the secondary vertex, providing a measure of the probability that the jet originates from a b-quark.
- DeepCSV tagger: This is advanced combined secondary vertex algorithm based on a deep neural network-based algorithms that exploit a variety of input variables, including track-related information, secondary vertex properties, and additional discriminators. The neural network is trained to distinguish between jets originating from b-quarks and those from other sources, providing a discriminant output for b-tagging. DeepCSV discriminator for tagging b jets in $t\bar{t}$ sample is shown in Figure 3.14 (right).
- Soft lepton and combined tagger: This tagging algorithm utilizes the presence of charged leptons (e.g., muons or electrons) within the jet. It considers the properties of both the secondary vertex and the lepton, such as their kinematic characteristics and the compatibility of the lepton with the jet axis. By combining these pieces of information, it assigns a b-tagging likelihood to the jet.

Among these tagging algorithms, the deepCSV tagger is commonly used for the identification of b-jets. It is applied to the reconstructed AK4 jets due to its favorable b-tagging efficiency with a lower probability of misidentifying light quarks compared to other taggers as shown in Figure 3.14 (right). The efficiency (mis-identification probability) to correctly (wrongly) tag a jet with given flavour is defined as the number of jets of given flavour passing the tagging requirement divided by the total number of jets of that flavour. Different working points such as loose, medium and tight, based on the deepCSV discriminator value are used for the 2016, 2017, and 2018 Run 2 datasets. As an example, deepCSV medium working point corresponds to a b jet tagging efficiency of 65% for jets with $p_T > 30$ GeV with corresponding misidentification probability for gluon and light-quark (charm-quark) jets of 1.6% (13%).

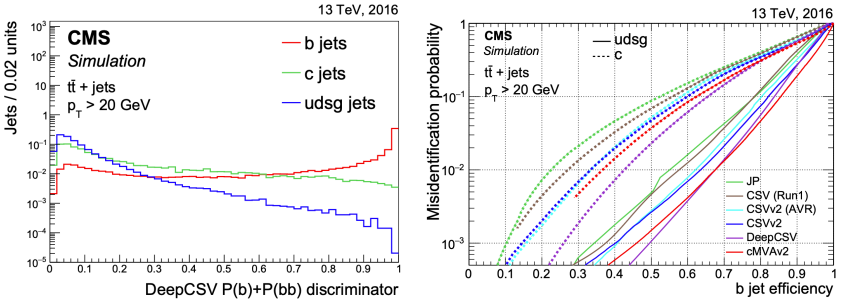

FIGURE 3.14: DeepCSV discriminator for ak4 jets in $t\bar{t}$ sample having different jet flavours (left). The distributions are normalized to unit area. On the right, mis-identification probability for light quarks with respect to b-tag efficiency [93].

3.4.6 Missing transverse momentum

Total initial transverse momentum of two colliding partons is zero and therefore the total vector sum of p_T of final state particles is also expected to be zero. A non-zero value of final state transverse momenta indicate an invisible particle passing through the detector. Thus, the combined p_T of weakly interacting particle like neutrinos or beyond SM particles in the final state, can be inferred by summing the negative $\vec{p}_T$ of all observable particles. The CMS event reconstruction utilizes two PF-based algorithms to determine and incorporate the missing $\vec{p}_T^{miss}$ into the analysis: p_T^{miss} reconstruction algorithm to calculate the $\vec{p}_T^{miss}$ as the negative $\vec{p}_T$ sum of all PF particles. Due to non-linearity of the detector response, PF Jets need to be corrected using JECs and hence p_T^{miss} are calibrated as shown in following equation:

$$\vec{p}_T^{\,miss}(\text{calibrated}) = \vec{p}_T^{\,\text{miss}}(\text{raw}) - \sum_{\text{jets}}(\vec{p}_T^{\,\text{jets}}(\text{corr.}) - \vec{p}_T^{\,\text{jets}}(\text{raw})) \tag{3.14}$$

On top of these corrections, events in which p_T^{miss} is identified to be originating from a mismeasured jet, detector noise, non-functional calorimetric channels, or reconstruction failures are rejected by following dedicated algorithms [94]:

- **HCAL filters**: In HCAL, HPDs and readout box (RBX) electronics are susceptible to the noise and can affect the energy measurements and introduce additional energy deposits that are unrelated to the physical interactions in the detector. The HPD or RBX channels exhibit specific geometrical patterns that can be used to distinguish noise from genuine signals. Anomalous energy deposited can be removed in an event using "noise filtering"and "event filtering" mode is used to remove entire such events from the data set if they are determined to be affected by significant noise. Furthermore, there is an "isolation-based noise filter" which combines energy deposits in the HCAL and ECAL in barrel or endcap regions and compares them with measurements from the tracker.

- **ECAL filters**: Similar to the HCAL filters, the ECAL filters uses the topological and timing information to identify events with abnormal energy deposits or detector noise in the ECAL. Other dedicated "noise filters" are used to reject events with anomalous high p_T^{miss} originating from supercrystals and events where channels with non-functional readout electronics lack information.

- **Beam halo filter**: The beam halo refers to particles, predominantly muons, that are produced by interactions of the circulating proton beams with residual gas in the beam pipe. These muons can mimic genuine p_T^{miss} signals. During Run II, the beam halo filter employs information from both the CSC and the calorimeters to distinguish between genuine physics events with events contaminated by beam halo particles. By removing events with significant contributions from the beam halo, the purity of the p_T^{miss} sample is improved.

- **Reconstruction filters**: The reconstruction filters are algorithms designed to identify and reject events with reconstruction failures or anomalies. These filters target events

where the p_T^{miss} is affected by issues such as non-functional calorimetric channels or misreconstructed objects. They utilize various criteria to identify problematic events based on the characteristics of reconstructed particles, such as their energies, positions, or association with other objects in the detector. By discarding events with reconstruction failures, the quality of the p_T^{miss} reconstruction is enhanced.

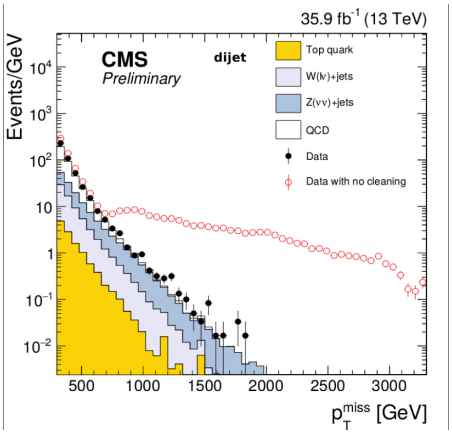

FIGURE 3.15: The distributions of p_T^{miss} is shown after applying event filtering algorithms to one set of events (represented by filled markers), while another set of events did not undergo the event filtering algorithms (represented by open markers) [94].

Chapter 4

Particle Flow charged hadron calibration

As discussed in the previous chapter, PF algorithm is used to reconstruct pp collision events recorded by the CMS detector. The resulting PF candidates are clustered to reconstruct jets and MET. In this chapter, we explain the need for PF hadron calibrations and methodology, and performance achieved through the reprocessing of Run II data (Ultra-legacy). Furthermore, we will investigate the impact of changing ECAL noise thresholds with respect to eta on PF charged hadrons, and also the improvements after applying these calibrations. Finally, this chapter will also touch upon some developmental studies that aim to enhance the efficiency of these calibrations using Graph Neural Networks (GNN) with Run III data and giving promising direction for significant improvements in resolution.

4.1 Motivation

The CMS experiment [69] uses a PF based algorithm, also called global event description, to reconstruct the pp collision events using signal deposited in various sub-detectors. Inputs to PF clustering algorithms are all rechits which qualify predefined criteria on thresholds and other sanity checks, and its outcome is a list of PF particle candidates: electrons, photons, hadrons (charged and neutral), and muons. The initial step involves a link algorithm connecting PF elements (e.g., tracks, calorimeter clusters) from different sub-detectors, aiming to link elements corresponding to the same particle. Successfully connected PF elements form PF blocks, representing collections associated by direct or indirect links through common elements. These PF candidates are used as inputs to jet clustering algorithms and MET reconstructions and provide significantly improved response and resolution for these physics objects over the ones using only calorimetric information [95, 81]. Figure 4.1 shows comparisons of the average fraction of energies carried by various PF particles clustered into jets (PF jet composition) as a function of p_T of jets in collision data and simulated MC samples corresponding to 2016, 2017 and 2018 separately [96]. At the particle level, ≈30% energy of jets is carried by neutral pions (which immediately decay to pairs of photons) and the remaining by charged and neutral hadrons mostly dominated by $\pi^{\pm}$, Fig 4.2. Since the response of calorimeters is similar for hadrons except at very low energy protons and neutrons, pions are a good candidate for the

purpose of PF hadron calibrations. π^0s decaying to photons, are precisely measure by ECAL and charged pions deposit their energy in both the ECAL and the HCAL.

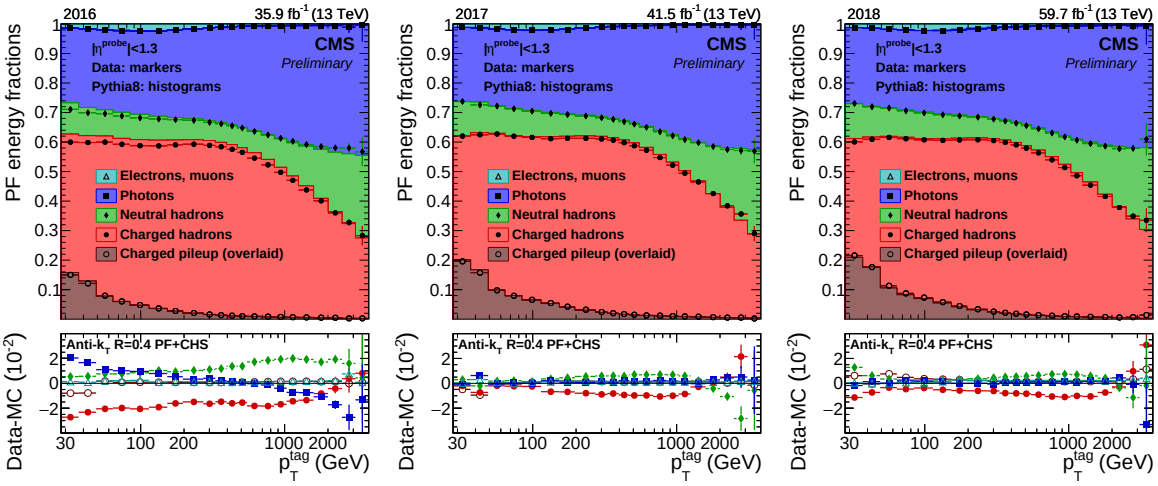

FIGURE 4.1: Comparisons of PF jet composition measured as a fraction of energy carried by PF photons, charged hadrons, neutral hadrons, and electrons and muons in data and MC simulation for 2016 (left), 2017 (middle), and 2018 (right) [96]

Since the PF reconstruction clusters link the energy deposits in calorimeters, dedicated calibrations are needed for each cluster type to ensure the best performance of higher-level physics objects. in this chapter, we describe the calibration of the PF hadron clusters and present the results used as input to the Ultra Legacy (UL) processing of the full Run II dataset , *i.e.* the data collected during the years 2016, 2017 and 2018. The calibration factors are determined using samples of negatively charged pions (π^-) simulated and reconstructed with exact run conditions and detector settings for MC samples used for physics analyses. In addition, one can make dedicated measurements using isolated tracks from the collision data, as is also done for the HCAL in situ calibration as presented in [97] and dedicated validation of GEANT4 physics lists [98]. The PF hadron calibrations based on single π^- are used in the current CMS detector event reconstruction scheme and are described in this chapter.

4.2 Technical details for Ultra Legacy simulation samples

Ultra Legacy refers to the final reprocessing of data collected at the CMS detector in 2016, 2017 & 2018 using the most up to date calibrations and alignment of the subdetectors. Changes (Phase I upgrades) in detectors for these years are shown in Table 4.1 [99].

Year	Phase I upgrades
2016	New μTCA-based L1 trigger
2017	New layer in pixel detector and very forward calorimeter electronics
2018	HCAL endcap photodetectors (HPD replace with SiPMs) and electronics

TABLE 4.1: Detector changes in 2016, 2017 & 2018 [99].

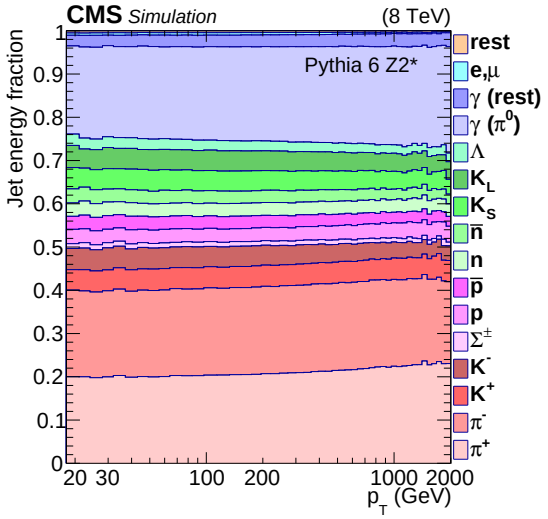

FIGURE 4.2: Composition of particle level jets presented as a fraction of energy carried by various generator level particles produced in the process of hadronization of quarks and gluons [95].

Single π^- events are generated which flat in the energy range from $2 - 500$ GeV in the $|\eta| < 3$. The technical details for this round of calibrations are summarized in Table A.2. We used centrally produced reconstructed samples, also listed in the Table A.1 in appendix A.1.

4.3 Event selection

For the purpose of calibration, a few selection criteria are applied to ensure the particle has not started showering before arriving at the face of ECAL and to reject clusters originating from noise or fake hadrons (or non-isolated hadrons as it may happen in collision events). Events with only one true particle in an event to continue with the further event selections, these true particle informations are contained in PFSimParticles collection. The following event selections are based on the PFSimParticles, PF candidates and PF block elements.

PF charged hadrons are required to satisfy the following criteria:

- Transverse momentum p_T of PF charged hadron ≥ 1.0 GeV.
- Sum of raw ECAL cluster energy and raw HCAL cluster energy ≥ 0.5 GeV.
- Number of tracks associated with the PF block element should be exactly 1 (to ensure the clusters are isolated and not started showering earlier).
- Total momentum (p) and transverse momentum (p_T) of the track should be greater than 1 GeV.
- Number of pixel hits associated with the track should be greater than 2.
- Number of total tracker hits are eta dependent and are summarized in Table 4.2.

$\|\eta\|$ range	Tracker hits
0-1.4	> 14
1.4-1.6	> 17
1.6-2.0	> 20
2.0-2.4	> 17
2.4-2.6	> 10

TABLE 4.2: Eta dependent tracker hits selection used in event selection

When PF candidate is not a PF charged hadron, then the following conditions are used to select the events:

- PF candidate is a neutral hadron
- Angular selection between PF candidate and true particle i.e. dR(PF candidate, true particle) < 0.4.

After the above event selections for a good PF candidate, the following informations for each event are saved for further analysis:

- E_t: true momenta of the particle at the ECAL entrance extrapolated in a magnetic field.
- η & ϕ of the extrapolated true particle at the ECAL entrance.
- E_e^{raw}: raw energy of the ECAL cluster associated with the PF candidate
- E_h^{raw}: raw energy of the HCAL cluster associated with the PF candidate

The total raw energy of the PF hadron candidate refers to the sum of raw energies of the ECAL and HCAL clusters (here, raw refers to reconstructed energies using detector-level calibrations and is explained in more detail in the next section). The total raw energies of PF hadrons then refer to:

$$E_{total}^{raw} = E_e^{raw} + E_h^{raw} \tag{4.1}$$

To determine the calibration constants, we study the response of charged hadrons in the bins of E_t. The 4.3 (left) shows the distribution of response for pions in the bin E_t = 60-64 GeV where the response of raw PF hadrons is defined as $(E_t - E_{total}^{raw})/E_t$. The distribution is fitted with a Gaussian function. We used the Gaussian fitted mean (μ) interchangeably as the average response or response and resolution as Gaussian σ normalized to the μ. The response obtained this way for all bins is shown in Figure 4.3 (right) as a function of E_t using the sample of hadrons for which there is always an associated ECAL cluster, i.e. $E_e^{raw} > 0$. As observed, the raw response is not linear with the true energy with substantial non-linearity in low-energy regions. In case of hadronic showers, some part of the energy is lost in breaking nuclei (nuclear binding energy) and that leads to invisible energy, resulting into non-linearity in energy response. The calibration procedure aims to correct this non-linearity in all bins of E_t and η_t.

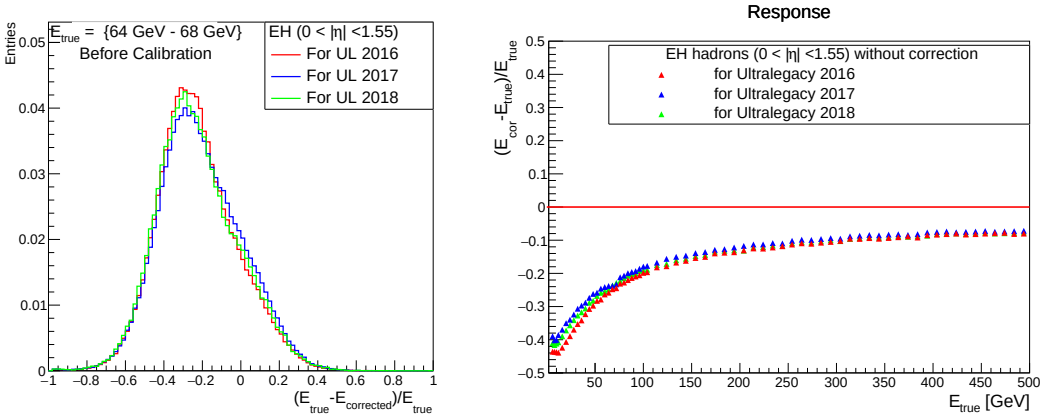

FIGURE 4.3: Distribution of raw energy response, i.e. $(E_t - E^{raw}_{total})/E_t$ for the hadrons with E_t=40−60 GeV (left), and average energy response as a function of E_t (right) for the hadrons which have an associated ECAL cluster, i.e. $E^{raw}_e > 0$.

4.4 Calibration methodology

This section summarizes the calibration procedure using the sample of hadrons selected using the criteria described in Section 4.3. Given the difference in the ECAL and HCAL absorbers, the response to energies deposited by hadrons in these sections is different [100]. This results in a difference in an overall difference in response and resolution for pions which start showering in ECAL, referred to as *EH hadrons*, and those start showering in HCAL, referred to as *H hadrons*, in the following. Hence, the PF cluster calibration is performed on these two types of hadrons separately:

- *EH hadrons* : when hadrons starts showering in ECAL and deposit energy in both ECAL & HCAL clusters. The corresponding PF candidates have linked ECAL and HCAL clusters.
- *H hadrons* : when hadrons starts showering in HCAL and deposit energy in HCAL only. The PF candidates which do not have any ECAL cluster associated with them, and are identified by requiring $E^{raw}_e = 0$.

The calibration procedure starts with the PF charged or neutral hadron candidates, and their associated raw ECAL energies (E^{raw}_e) and raw HCAL energies (E^{raw}_h). Here "raw" refers to the detector level calibrations of the ECAL using $e^{\pm}/\gamma$ and of the HCAL using 40−60 GeV isolated tracks. Total raw energy of the two types of hadrons is defined as:

$$E^{raw}_{EH} = E^{raw}_e + E^{raw}_h, \quad \text{for EH hadrons} \tag{4.2}$$

$$E^{raw}_{H} = E^{raw}_h, \quad \text{for H hadrons} \tag{4.3}$$

Raw energy response of EH and H hadrons reconstructed using 2016, 2017 & 2018 conditions are shown in Fig 4.4 as a function of true energy (top row) and η (bottom row) of hadrons. The corresponding one dimensional distributions for all years and detector region are shown in section 4.5.4. The PF hadron calibration refers to updating (correcting) the E^{raw}_e and E^{raw}_h,

hence E_{EH}^{raw} and E_{H}^{raw}, in a way that the average corrected energy response of PF hadrons is equalized to unity as a function of its true energy and pseudorapidity.

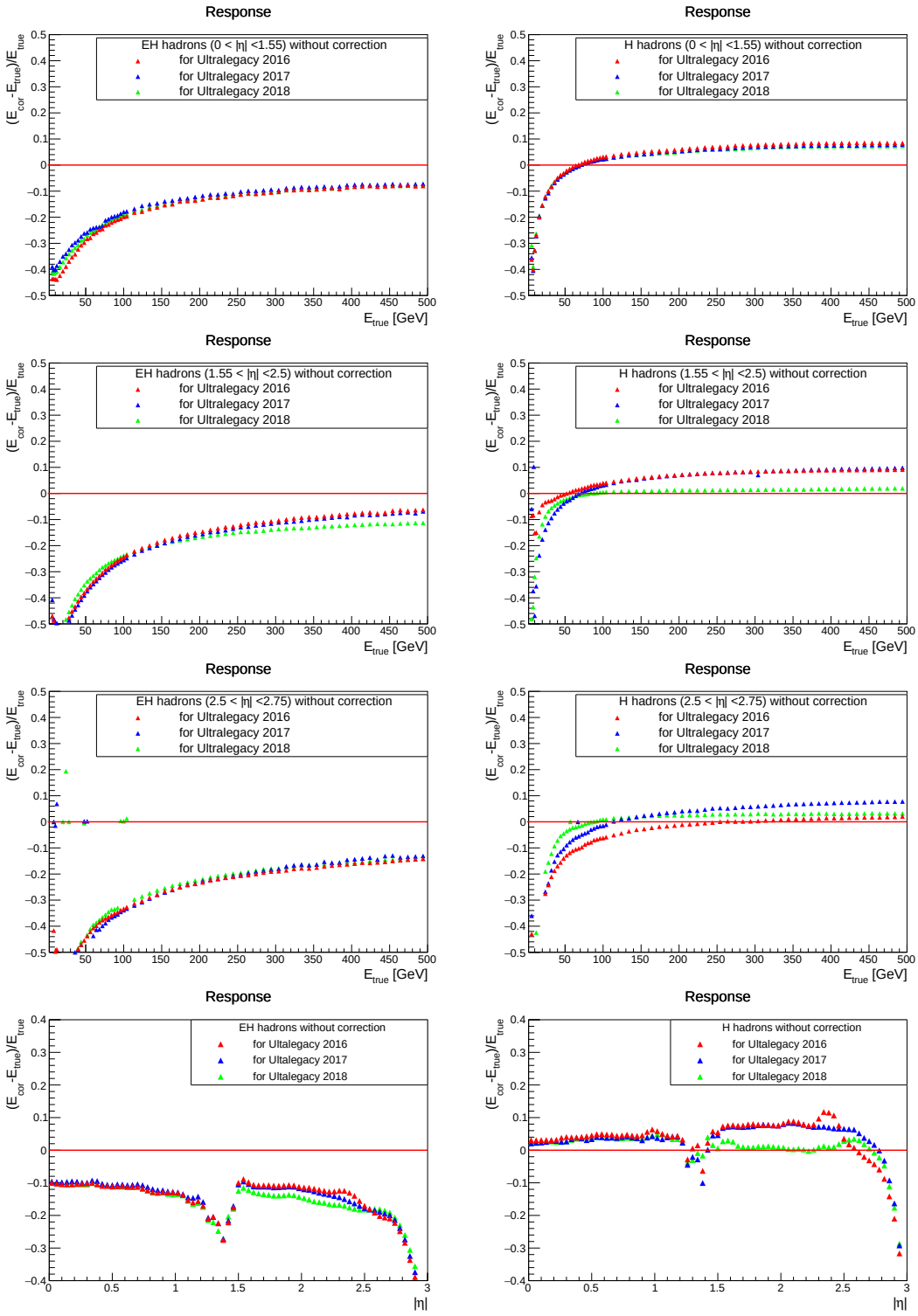

FIGURE 4.4: Average raw energy response i.e. $(E_t - E_{total}^{raw})/E_t$ as a function of E_t (top) and $|\eta|$ (bottom) for EH (left) and H (right) hadrons for Ultra legacy conditions corresponsing to 2016, 2017 & 2018.

The calibration factors are sequentially obtained in two steps, both for *EH hadrons* and *H hadrons*, separately for barrel and endcap regions.

1. **Step 1:** In the first step, energy dependent calibrations are obtained separately for pions depositing their energies in barrel and endcap regions (inclusive in η), E_e^{raw} and E_h^{raw} are obtained to get

$$E_{EH}^{Ecorr} = a(E_t) \times E_e^{raw} + b(E_t) \times E_h^{raw} + O_{EH}, \quad \text{for EH hadrons} \quad (4.4)$$
$$E_H^{Ecorr} = c(E_t) \times E_h^{raw} + O_H, \quad \text{for H hadrons} \quad (4.5)$$

where $a(E_t)$ and $b(E_t)$ are simultaneously determined using a χ^2 method described in Section 4.4.1. Finally, O_{EH} and O_H are two energy offsets, fixed by hand based on previous documentation (default in PF calibration code).

The two offsets are fixed to O_{EH} = 3.5 GeV for EH hadrons, and O_H = 2.5 GeV for H hadrons as described in previous studies [95]. These values are chosen to minimize the dependence of parameters a, b and c on the E_t, and allow to account for the different threshold effects as the HCAL zero suppression and ECAL Selective Readout thresholds, the energy loss in the detector dead material and potential cluster seeding effects. If any of a, b & c parameters in endcap or barrel take negative values less than -0.25, then parameter value will be taken as unity i,e. calibration is not applied.

2. **Step 2:** To account for potential $\eta-$dependence of response, a function $f(|\eta|)$ is introduced in the energy measurements to approximately account for observed features in a given set of conditions, and calibrations are uupdated again as:

$$E_{EH}^{Corr} = (1 + \alpha(E_t) + \beta(E_t)\, f(|\eta|)) \times E_e^{Ecorr} + E_h^{Ecorr}, \quad \text{for EH hadrons} \quad (4.6)$$
$$E_H^{Corr} = (1 + \alpha(E_t) + \beta(E_t)\, f(|\eta|)) \times E_h^{Ecorr}, \quad \text{for H hadrons} \quad (4.7)$$

where, $E_e^{Ecorr} = a(E_t) \times E_e^{raw}$ and $E_h^{Ecorr} = b(E_t) \times E_h^{raw}$ for EH hadrons and $E_h^{Ecorr} = c(E_t) \times E_h^{raw}$ for H hadrons, as referred to in equation 4.4 and equation 4.5, respectively. This minimisation step ignores the offset terms to obtain $\alpha(E_t)$ & $\beta(E_t)$ parameters.

The parameterization function dependent on eta, the $f(|\eta|)$, could be different for different detector regions, and depending on the behaviour of response corrected in the previous step, it is independently chosen for the three pseudorapidities $|\eta|$ regions (Figure 4.5) defined as follows:

- Barrel region: $0 < |\eta| < 1.55$
- Endcap within tracker region: $1.55 < |\eta| < 2.5$
- Endcap outside tracker region: $2.5 < |\eta| < 2.75$

The response is then again equalized as a function of E_t using a χ^2 minimization method to obtain final corrected energies, E_{EH}^{Corr} and E_H^{Corr}, of PF hadron candidates as summarized in Section 4.4.2.

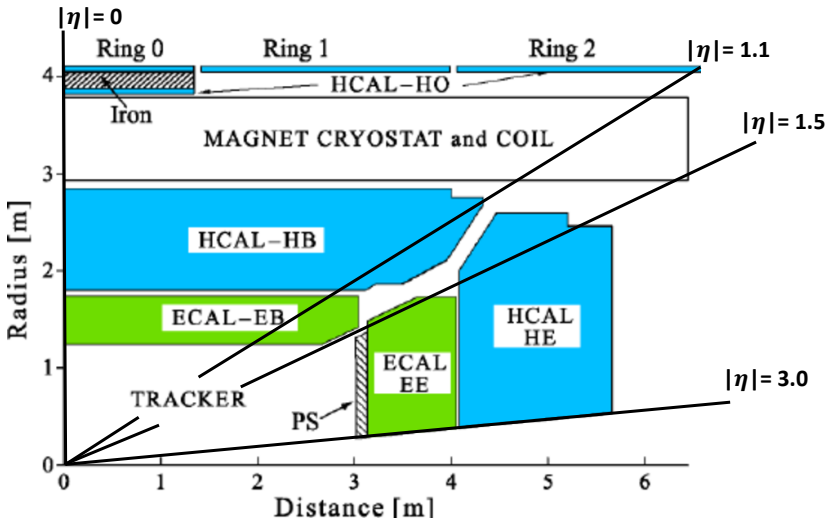

FIGURE 4.5: The CMS longitudinal layout [101].

Table 4.3 summarizes the different sets of independent calibration constants optimized to get the final corrected energies. Based on the equations 4.4, 4.5, 4.6, & 4.7, the final calibrated energies of PF hadrons are obtained in different detector regions as:

$$E_{EH}^{Corr} = (1 + \alpha(E_t) + \beta(E_t)\, f(|\eta|)) \times a(E_t) \times E_e^{raw} + b(E_t) \times E_h^{raw} + O_{EH}, \quad \text{for EH hadrons} \tag{4.8}$$

$$E_H^{Corr} = (1 + \alpha(E_t) + \beta(E_t)\, f(|\eta|)) \times c(E_t) \times E_h^{raw} + O_H, \quad \text{for H hadrons} \tag{4.9}$$

The energy-dependent corrections aim to account for the non-linearity of the calorimeter response to hadron showers in non-compensating calorimeter setups, as in the CMS. In the second step, pseudorapidity dependency is corrected for, if any. The main reasons for the need for η dependent corrections could be that the path length of particles traversing the ECAL and HCAL sections vary with η, the effect of zero suppression may vary as we go higher in η regions, the regions outside tracking may not be as well calibrated at the detector level etc. In both steps , *i.e.* the energy-dependent calibration and the η dependent calibrations, the calibrations are performed using a χ^2 minimization procedure and is described in the Sections 4.4.1 and 4.4.2, respectively.

Energy-dependent corrections	Barrel	Endcap	
EH hadrons	$a(E_t)$, $b(E_t)$	$a(E_t)$, $b(E_t)$	
H hadrons	$c(E_t)$	$c(E_t)$	
η-dependent functions	Barrel	Endcap	
		within tracker	outside tracker
EH hadrons	$f(\lvert\eta\rvert)$	$f(\lvert\eta\rvert)$	$f(\lvert\eta\rvert)$
H hadrons	$f(\lvert\eta\rvert)$	$f(\lvert\eta\rvert)$	$f(\lvert\eta\rvert)$
η-dependent corrections	Barrel	Endcap	
EH hadrons	$\alpha(E_t)$, $\beta(E_t)$	$\alpha(E_t)$, $\beta(E_t)$	
H hadrons	$\alpha(E_t)$, $\beta(E_t)$	$\alpha(E_t)$, $\beta(E_t)$	

TABLE 4.3: Different regions used to obtain various calibration parameters

4.4.1 Energy dependent calibrations (Step 1 of calibration)

This step concerns obtaining the parameters $a(E_t)$, $b(E_t)$ and $c(E_t)$ defined in the equations 4.4 and 4.5. A χ^2 is defined as:

$$\chi^2 = \sum \left[\frac{E_t - E_{EH}^{Ecorr}}{\sigma(E_e^{raw} + E_h^{raw})} \right]^2, \qquad \text{for EH hadrons} \tag{4.10}$$
$$= \sum \left[\frac{E_t - a(E_t) \times E_e^{raw} - b(E_t) \times E_h^{raw} - O_{EH}}{\sigma(E_e^{raw} + E_h^{raw})} \right]^2$$

and

$$\chi^2 = \sum \left[\frac{E_t - E_H^{Ecorr}}{\sigma(E_h^{raw})} \right]^2, \qquad \text{for H hadrons} \tag{4.11}$$
$$= \sum \left[\frac{E_t - c(E_t) \times E_h^{raw} - O_H}{\sigma(E_h^{raw})} \right]^2$$

where the summation is performed over all the hadrons in the H and EH samples in a given bin of E_t. Uncertainties on the measured calorimeter energies, $\sigma(E)$, used in the denominator are given by expected hadron resolution in the barrel and endcap regions as:

$$\sigma(E)^2 = 0.08^2 + 1.04^2 \times max(E, 1.0), \qquad \text{for barrel} \tag{4.12}$$
$$= 0.04^2 + 1.80^2 \times max(E, 1.0), \qquad \text{for endcap} \tag{4.13}$$

Using these uncertainties described in equations 4.12 and 4.13 helps mitigate the effect of tails in raw energies measured in ECAL and HCAL in the χ^2 minimization and permits consistent uncertainties on parameters a, b and c to be obtained by minimizing equations 4.10 and 4.11. These numbers are taken as these were from previous studies [95].

The minimization of two χ^2 , *i.e.* equations 4.10 and 4.11 can be done analytically in each bin of E_t. The result is a set of two parameters, $a(E_t)$ and $b(E_t)$ for EH hadrons, and $c(E_t)$ for H hadrons :

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum \frac{(E_e^{raw})^2}{\sigma^2(E_e^{raw}+E_h^{raw})} & \sum \frac{E_e^{raw}E_h^{raw}}{\sigma^2(E_e^{raw}+E_h^{raw}))} \\ \sum \frac{E_e^{raw}E_h^{raw}}{\sigma^2(E_e^{raw}+E_h^{raw})} & \sum \frac{(E_h^{raw})^2}{\sigma^2(E_e^{raw}+E_h^{raw})} \end{pmatrix}^{-1} \times \begin{pmatrix} \sum \frac{(E_t-O_{EH})E_e^{raw}}{\sigma^2(E_e^{raw}+E_h^{raw})} \\ \sum \frac{(E_t-O_H)E_h^{raw}}{\sigma^2(E_e^{raw}+E_h^{raw})} \end{pmatrix} \tag{4.14}$$

$$c = \frac{\sum \frac{(E_t-O_H)E_h^{raw}}{\sigma^2(E_h^{raw})}}{\sum \frac{(E_h^{raw})^2}{\sigma^2(E_h^{raw})}} \tag{4.15}$$

The parameters obtained using equations 4.14 and 4.15 are summarized in Figure 4.8 for the barrel (left column) and endcap (right column) regions in the bins of E_t (for UL). Details of fitted functions and parameters in different region for all years can be found in Appendix A.

Using the fitted values of $a(E_t)$, $b(E_t)$ and $c(E_t)$, the E_{EH}^{Ecorr} and E_H^{Ecorr} are then obtained using equations 4.4 and 4.5, respectively.

The energy dependent corrections have been derived inclusively for barrel and endcap regions. The energy response of EH and H hadrons are shown in Figure 4.6 for both barrel and endcap regions after energy dependent corrections. However, some η-dependence of response is expected as the path length traversed by the particles in different calorimeter sections and gaps may vary, and zero-suppression thresholds are also η-dependent or impact. The response as a function of η, inclusive in E_t for each η bin is shown in Figure 4.7 (left) and Figure 4.7 (right) for EH and H hadrons respectively. The response decreases almost quadratically with η in the barrel ($|\eta|$ upto 1.5) as well as endcap regions ($|\eta|$ upto 3.0). As can be seen from 4.5, the barrel-endcap transition region for ECAL and HCAL covers $1.1 \leq |\eta| \leq 1.5$. As the particles traverses the gap, there is lesser material for shower to develop and hence decrease in response in this region and subsequently increase as shower starts seeing the endcap region (i.e. the longitudinal path traversed in the regions which are not fully instrumented by the detectors).

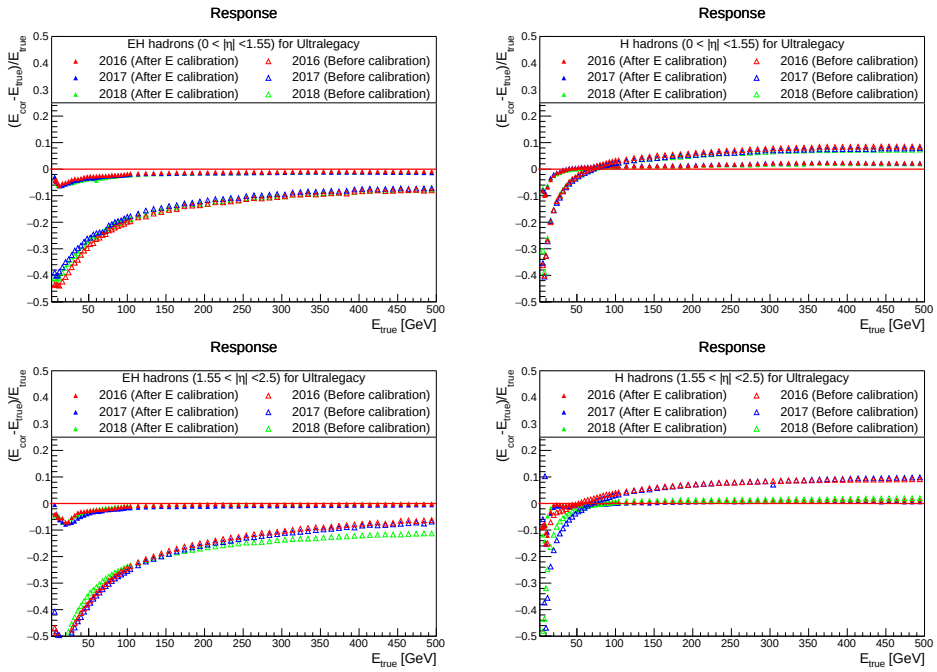

FIGURE 4.6: Average energy response after applying energy dependent calibrations as a function of E_t for EH hadrons(left) and for H hadrons(right).

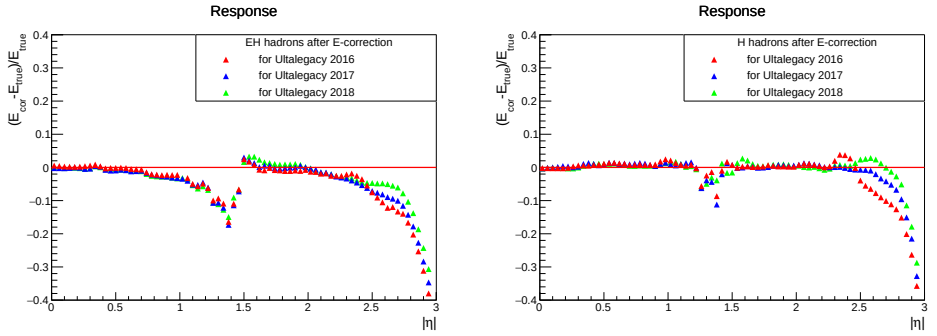

FIGURE 4.7: Average energy response after applying energy dependent calibrations as a function of $|\eta|$ for EH hadrons(left) and H hadrons(right).

4.4.2 Pseudo rapidity dependent calibrations

The η-dependent calibrations are obtained using E_e^{Ecorr} and E_h^{Ecorr} i.e. the corrected energies obtained by applying energy dependent calibrations on E_e^{raw} and E_h^{raw}. In this step, an η-dependence of measured energies is obtained by introducing a function, $f(|\eta|)$, as shown in equations 4.6 and 4.7. During Run 1, the $|\eta|$ dependent energy calibration for EH barrel region were defined as in $|\eta|^2$ for EH barrel region. However, we observed in Run 2 prompt cycles that with different detector configurations, there could be different $|\eta|$ dependencies in barrel and endcap (within and outside tracking) regions. Hence, we started using a parametric $|\eta|$ dependent function with a general form of f($|\eta|$), which can be defined as :

$$f(|\eta|) = (d_3 \times (|\eta| - d_0)^{d_1}) + d_2 \tag{4.16}$$

where d_i's are free parameters. The exact functional form of $f(|\eta|)$ depends on three η regions, name barrel, endcap within and outside tracking coverage, and on the EH and H hadron categories. The exact functional forms used for 2016, 2017 and 2018 UL processing are summarized in Table 4.4.

Functional form of $f(\|\eta\|)$			
2016			
Region	Barrel	Endcap	
		within tracker	outside tracker
EH hadrons	$\|\eta\|^2$	$-0.08 + 0.5\times(\|\eta\|-1.5)$	$0.3 + (\|\eta\|-1.5)^4$
H hadrons	$\|\eta\|^2$	0.08	$(\|\eta\|-1.5)^4$
2017			
	Barrel	Endcap	
		within tracker	outside tracker
EH hadrons	$\|\eta\|^2$	$0.4 + (\|\eta\|-1.5)^4$	$1.0\times(\|\eta\|-2.0)$
H hadrons	$\|\eta\|^2$	0.05	$-0.1 + (\|\eta\|-1.5)^2$
2018			
	Barrel	Endcap	
		within tracker	outside tracker
EH hadrons	$\|\eta\|^2$	$-0.8 + 2.2\times(\|\eta\|-1.5)^2$	$0.04 + (\|\eta\|-1.5)^4$
H hadrons	$\|\eta\|^2$	0.05	$-2.0 + (\|\eta\|-1.5)^2$

TABLE 4.4: Definitions of $f(|\eta|)$ used in 2016, 2017 and 2018

Using the equations 4.8 and 4.9 and the above $f|\eta|$ forms, updated χ^2s are defined for EH and H hadrons as:

$$\chi^2 = \sum \left[\frac{E_t - E_{EH}^{Corr}}{\sigma(E_e^{raw} + E_h^{raw})} \right]^2, \quad \text{for EH hadrons} \tag{4.17}$$

$$= \sum \left[\frac{E_t - (1 + \alpha(E_t) + \beta(E_t) f(|\eta|)) \times E_e^{Ecorr} - E_h^{Ecorr}}{\sigma(E_e^{raw} + E_h^{raw})} \right]^2 \tag{4.18}$$

and

$$\chi^2 = \sum \left[\frac{E_t - E_H^{Corr}}{\sigma(E_h^{raw})} \right]^2, \quad \text{for H hadrons} \tag{4.19}$$

$$= \sum \left[\frac{E_t - (1 + \alpha(E_t) + \beta(E_t) f(|\eta|)) \times E_h^{Ecorr}}{\sigma(E_h^{raw})} \right]^2$$

where summation is performed over all the hadrons in the EH and H samples in a given bin of E_t, and corresponding η value is used to calculate $f(|\eta|)$. Uncertainties on the measured calorimeter energies, $\sigma(E)$, used in the denominator are given by expected hadron resolution in the barrel and endcap regions presented in the equations 4.12 and 4.13.

Calibration parameters $\alpha(E_t)$ and $\beta(E_t)$ are obtained in the bins of true energy of hadron by

minimizing χ^2 in the similar fashion as described for the energy-dependent calibrations in the Section 4.4.1.

$$\begin{pmatrix}\alpha\\ \\ \beta\end{pmatrix} = \begin{pmatrix} \sum \frac{(E_e^{Ecorr})^2}{\sigma^2(E_e^{Ecorr}+E_h^{Ecorr})} & \sum \frac{(E_e^{Ecorr})^2 f(|\eta|)}{\sigma^2(E_e^{Ecorr}+E_h^{Ecorr})} \\ \\ \sum \frac{(E_e^{Ecorr})^2 f(|\eta|)}{\sigma^2(E_e^{Ecorr}+E_h^{Ecorr})} & \sum \frac{(E_e^{Ecorr})^2 (f(|\eta|))^2}{\sigma^2(E_e^{Ecorr}+E_h^{Ecorr})} \end{pmatrix}^{-1} \times \begin{pmatrix} \sum \frac{(E_t - E_e^{Ecorr} - E_h^{Ecorr})E_e^{Ecorr}}{\sigma^2(E_e^{Ecorr}+E_h^{Ecorr})} \\ \\ \sum \frac{(E_t - E_e^{Ecorr} - E_h^{Ecorr})E_e^{Ecorr} f(|\eta|)}{\sigma^2(E_e^{Ecorr}+E_h^{Ecorr})} \end{pmatrix} \tag{4.20}$$

$$\begin{pmatrix}\alpha\\ \\ \beta\end{pmatrix} = \begin{pmatrix} \sum \frac{(E_h^{Ecorr})^2}{\sigma^2(E_h^{Ecorr})} & \sum \frac{(E_h^{Ecorr})^2 f(|\eta|)}{\sigma^2(E_h^{Ecorr})} \\ \\ \sum \frac{(E_h^{Ecorr})^2 f(|\eta|)}{\sigma^2(E_h^{Ecorr})} & \sum \frac{(E_h^{Ecorr})^2 (f(|\eta|))^2}{\sigma^2(E_h^{Ecorr})} \end{pmatrix}^{-1} \times \begin{pmatrix} \sum \frac{(E_t - E_h^{Ecorr})E_h^{Ecorr}}{\sigma^2(E_h^{Ecorr})} \\ \\ \sum \frac{(E_t - E_h^{Ecorr})E_h^{Ecorr} f(|\eta|)}{\sigma^2(E_h^{Ecorr})} \end{pmatrix} \tag{4.21}$$

4.5 Results for UL processing of Run2 (2016-2018) data

4.5.1 Calibration parameters for UL processing of 2016-2018 dataset

The calibration parameters for the EH hadrons for barrel and endcap regions are shown in Figure 4.8, left column and right column respectively. These calibrations are input to the CMS event reconstruction and we have provided the various sets of calibrations which are used for all the publications based on CMS Run2 data and early publication using the first Run3 data.The first two rows correspond to parameters a(E_t) and b(E_t), and last two rows to α(E_t) and β(E_t).

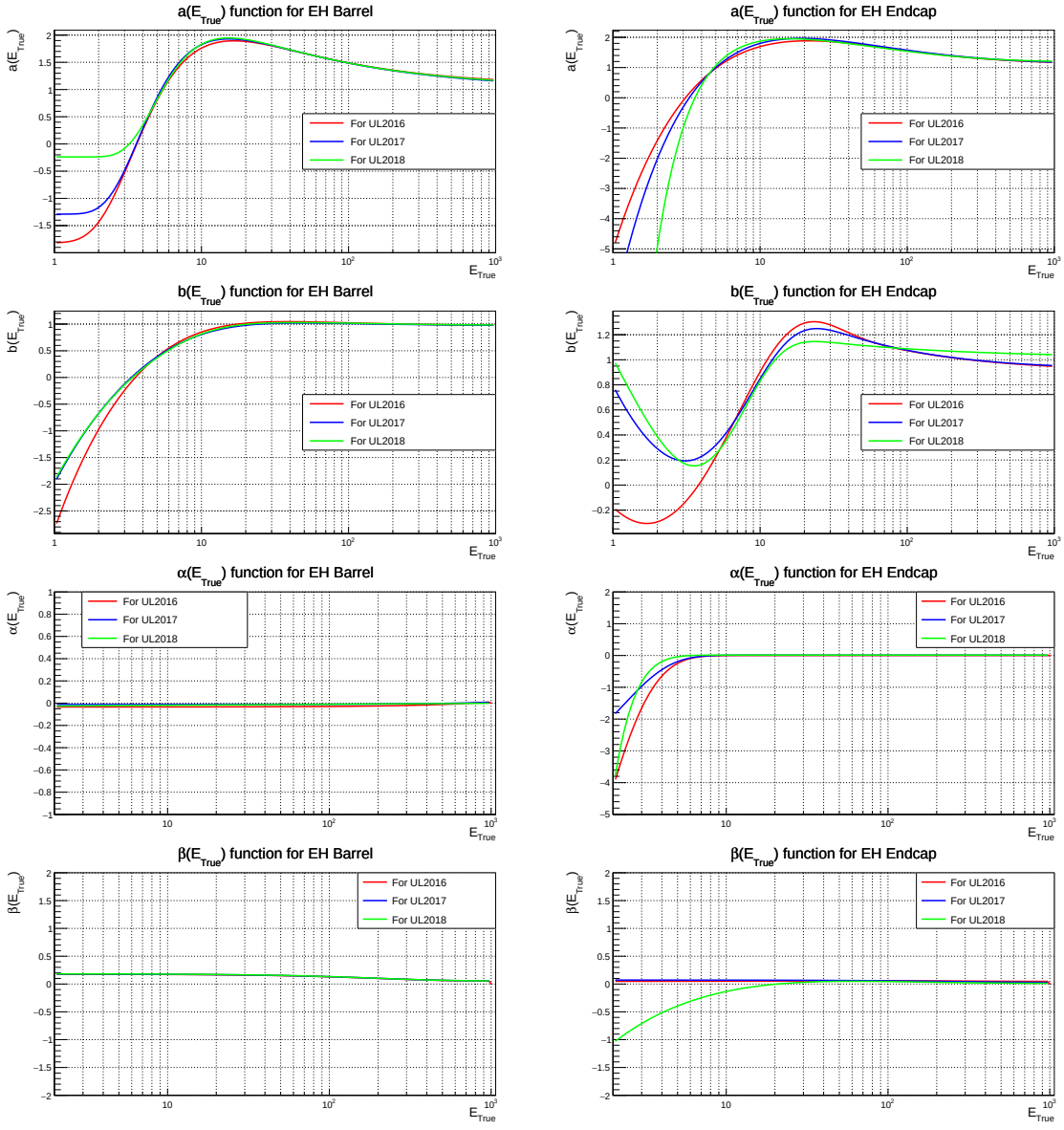

FIGURE 4.8: Calibration parameters for EH hadrons in barrel (left column) and endcap regions (right column): a(E_t) (row 1), b(E_t) (row 2), α(E_t) (row 3) and β(E_t) (row 4).

The calibration parameters for the H hadrons for barrel and endcap regions are shown in Figure 4.9, left column and right column respectively. The first row correspond to parameters

c(E_t), and last two rows correspond to $\alpha(E_t)$ and $\beta(E_t)$.

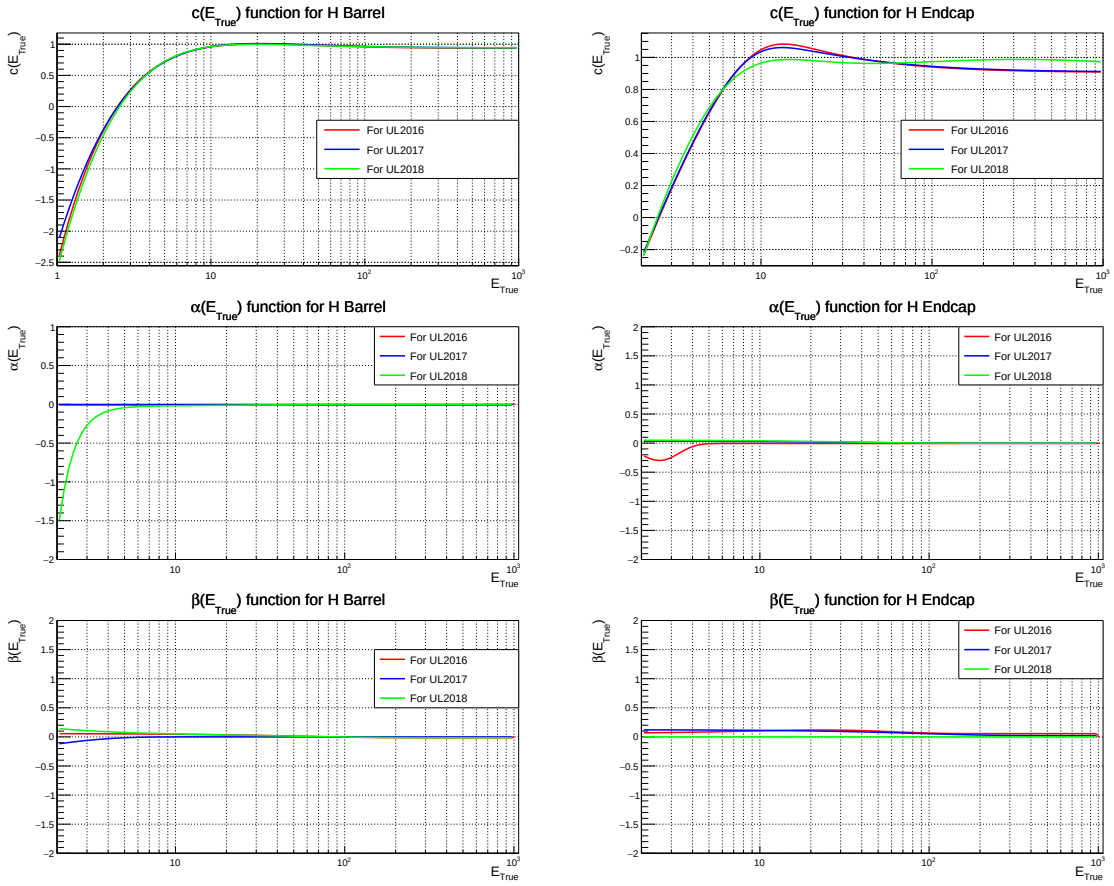

FIGURE 4.9: Calibration parameters for H hadrons in barrel (left column) and endcap regions (right column): c(E_t) (row 1), $\alpha(E_t)$ (row 2) and $\beta(E_t)$ (row 3).

4.5.2 Results - response & resolution as a function of E_t

The final response and resolution of **EH hadrons** obtained after energy dependent and η dependent corrections are shown as a function of E_t in Figure 4.10 left and right columns, respectively. The results are summarized for for the three years along with corresponding raw response and resolution.

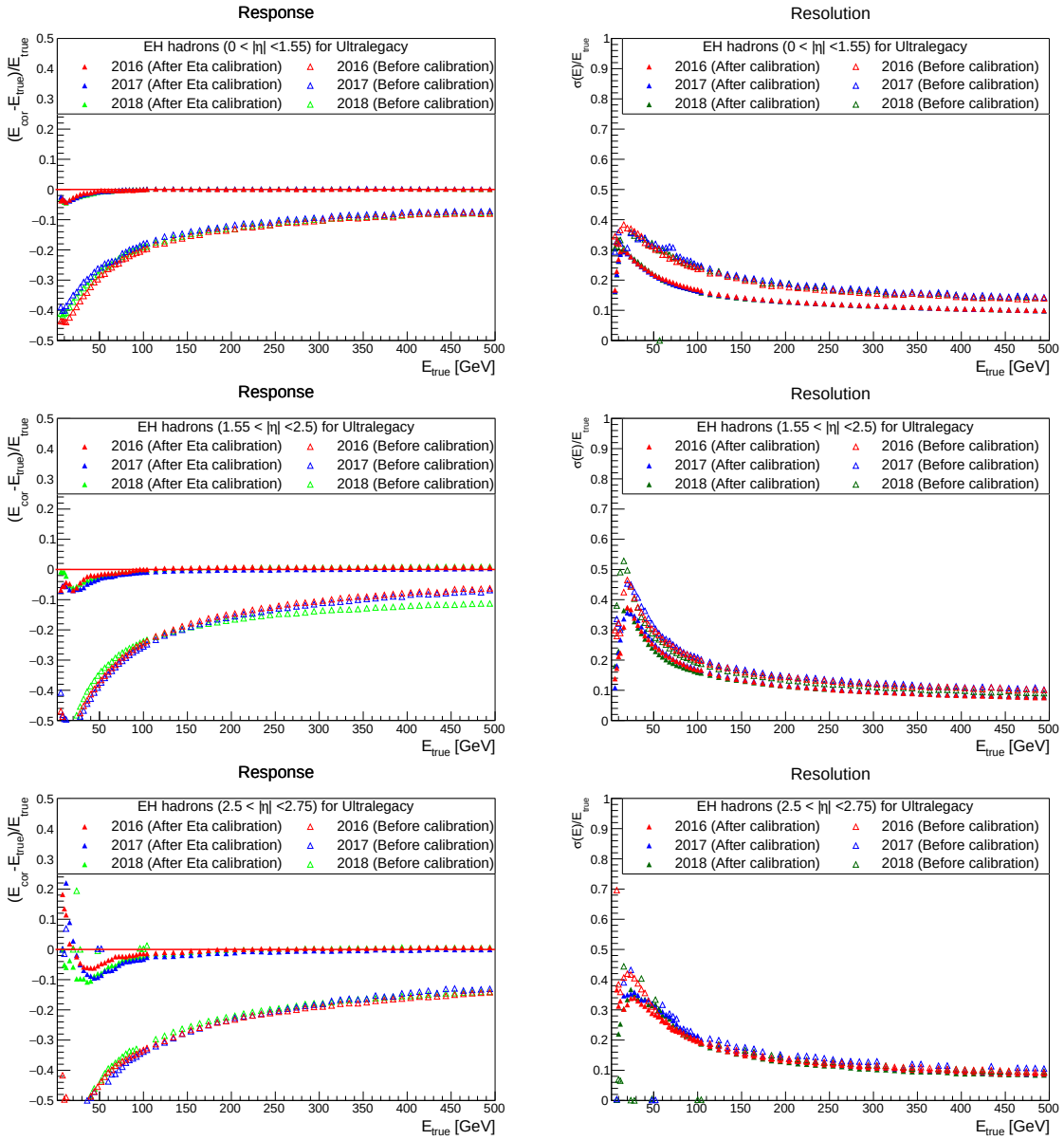

FIGURE 4.10: Final response (left column) & resolution (right column) of **EH hadrons** obtained in the bins of E_t shown for barrel (row 1), endcap within tracker (row 2) and endcap outside tracker (row 3) for 2016, 2017, and 2018. The solid and empty symbols correspond to corrected and raw energy measurements respectively.

The final response and resolution of **H hadrons** obtained after energy dependent and η dependent corrections are shown as a function of E_t in Figure 4.11 left and right columns, respectively. The results are summarized for for the three years along with corresponding raw response and resolution.

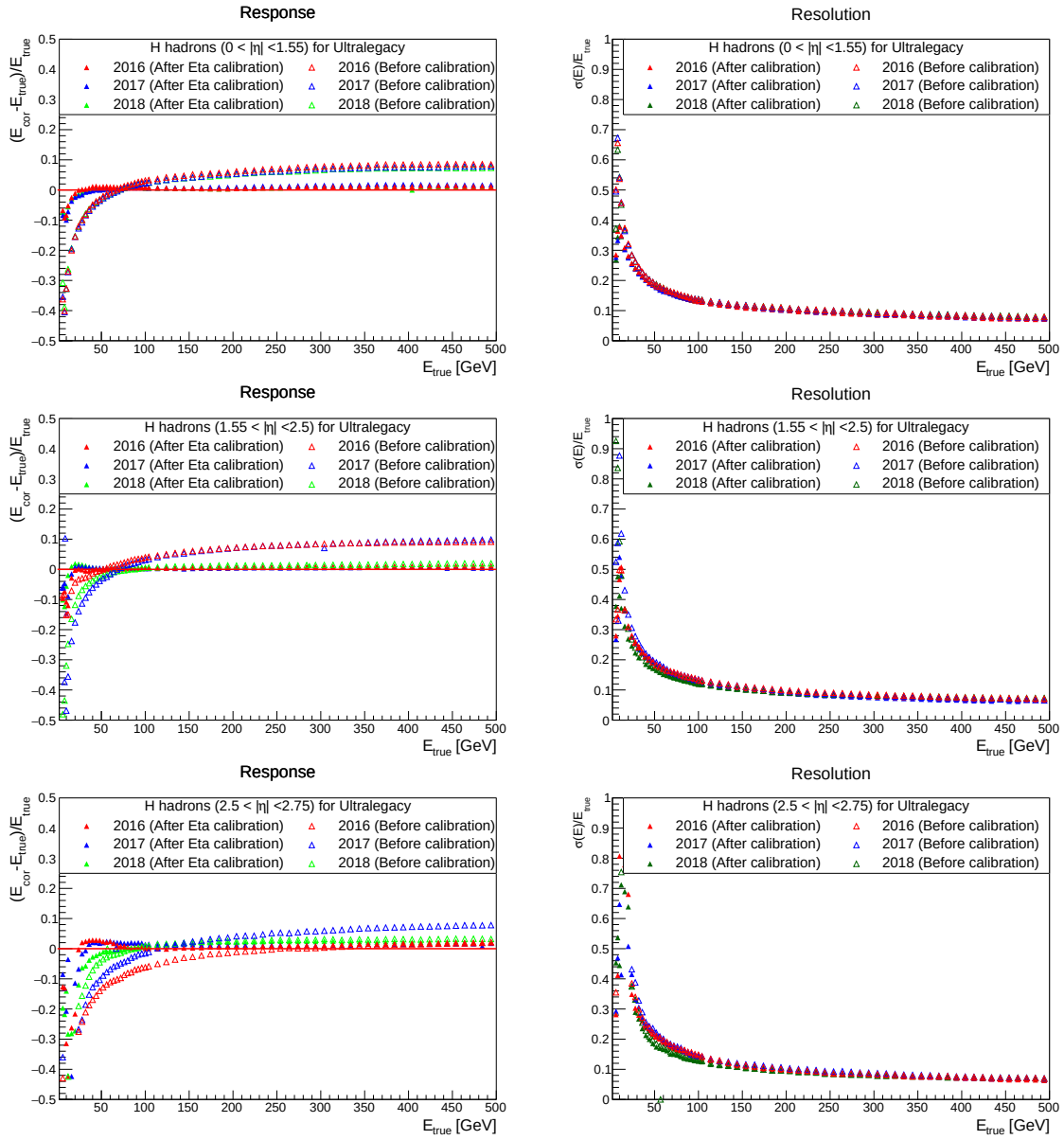

FIGURE 4.11: Final response (left column) & resolution (right column) of **H hadrons** obtained in the bins of E_t shown for barrel (row 1), endcap within tracker (row 2) and endcap outside tracker (row 3) for 2016, 2017, and 2018. The solid and empty symbols correspond to corrected and raw energy measurements respectively.

4.5.3 Results - response as a function of η

The final response obtained after energy dependent and η dependent corrections is shown as a function of η in Figure 4.12 for the three years.

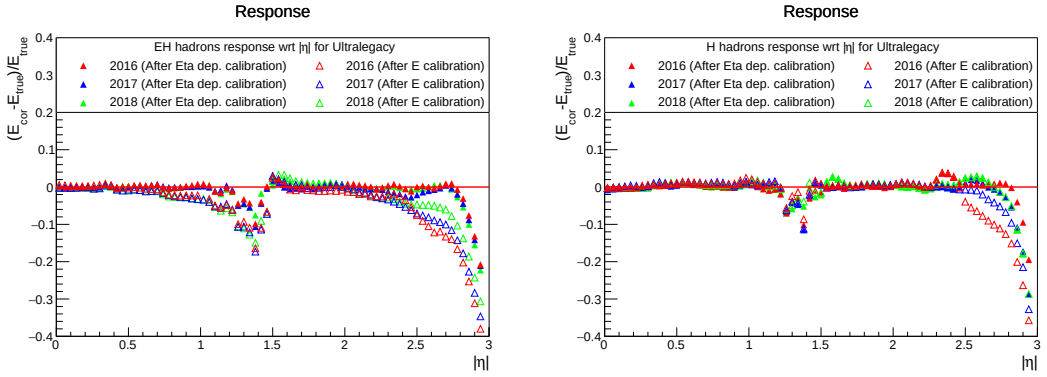

FIGURE 4.12: Final energy response shown in the bins of $|\eta|$ (inclusive in E_t) for EH hadrons (top) and H hadrons (bottom)

4.5.4 Results - 1D-distributions of raw and corrected responses

Distributions of raw and corrected PF hadron energies for a few bins of E_t are shown in the Figures 4.13 for EH and H hadrons in various eta regions. Distributions corresponding to each of 2016, 2017 and 2018 are shown in each canvas.

4.6 Effect of ECAL rechit thresholds on PF hadron cluster energy

The ECAL rechit thresholds refer to the minimum energy requirement applied to the energy deposits recorded in the ECAL crystals for particle reconstruction. These are used to filter out low-energy deposits and reduce noise in the detector. By setting a threshold, only energy deposits with energies above the threshold value are considered "rechits" and used for particle reconstruction, while lower-energy deposits are discarded. This threshold process helps to improve the accuracy and efficiency of particle energy measurements, as it reduces the impact of electronic noise and other sources of background signals. The thresholds can vary between different regions of the ECAL, such as the Barrel and the Endcap. The choice of threshold values depends on the specific needs of the data analysis, the experimental conditions, and the calibration status of the detector.

In 2018, before the reprocessing of Run2 data, the CMS collaboration decided to incorporate a new ECAL PF threshold scenario and earlier scenario (refered as "10_0_2"), ECAL rechit thresholds were 80 MeV in Barrel & 300 MeV in Endcap. In new scenario (refered as "10_0_3"), the threshold for energy deposits in the ECAL (rechit threshold) were set to be equivalent to 1 σ (standard deviation) of the noise in the ECAL, such that:

- Barrel (flat in η) = 0.140 GeV

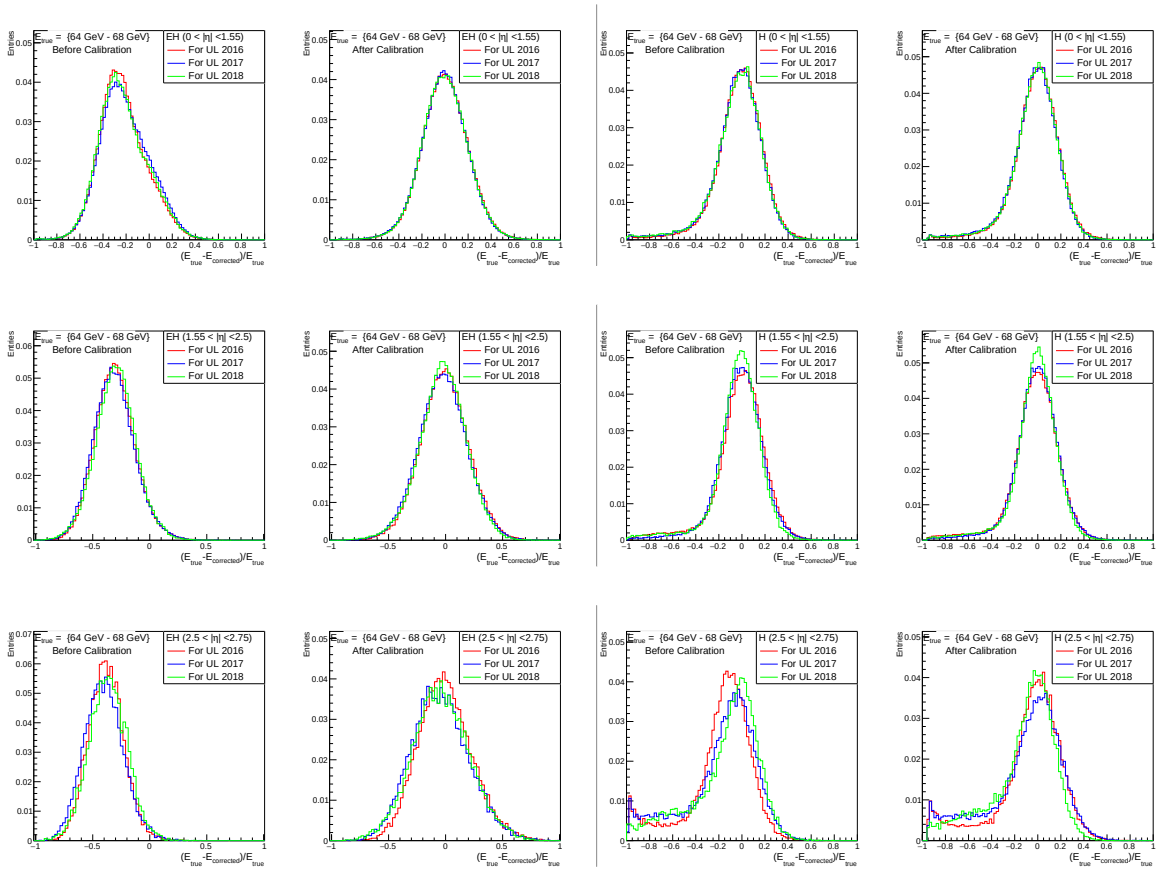

FIGURE 4.13: 1st (3rd) and 2nd (4th) columns represent 1D energy response distributions before and after final calibrations for EH (H) hadrons respectively. Top, middle and bottom rows represent the barrel, endcap within tracker and endcap outside tracker regions.

- Endcap (η dependent) = [0.11, 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.17, 0.18, 0.18, 0.19, 0.19, 0.20, 0.22, 0.23, 0.25, 0.27, 0.29, 0.31, 0.34, 0.36, 0.39, 0.42, 0.45, 0.50, 0.57, 0.68, 0.84, 1.07, 1.40, 1.88, 2.55, 3.47, 4.73, 6.42, 8.65, 11.6, 15.4] GeV

These thresholds are given by JetMET group [102] after checking jet energy responses and resolution with respect to η in different noise scenarios, correspond to 0, 30, 60 fb^{-1}. Since detector is under continous radiation exposure, ECAL transparency decreased by larger. Hence zero suppression thrsholds were increased in new scenario to reject the noise. In these studies, barrel resolution were stable but endcap resolution worsens by factor of 2 (max), and response increases by up to 30% for $|\eta| > 2.2$ with luminosity. Based on increased noise effects in ECAL endcap, we validated the choices of updated ECAL noise thresholds using PF hadrons.

For checking the effect of increasing ECAL noise thresholds, we have compared the 1D raw energy response distributions between "10_0_2" (earlier) and "10_0_3" (after updating ECAL thresholds) scenarios in the endcap region for EH hadrons as shown in Figure 4.14. These both samples will have similar detector configuration but different ECAL noise rechit thresholds as mentioned above. The technical details for both scenarios are mentioned in the Table A.7.

From Figure 4.14, it is evident that the effect of increasing ECAL noise is more pronounced in

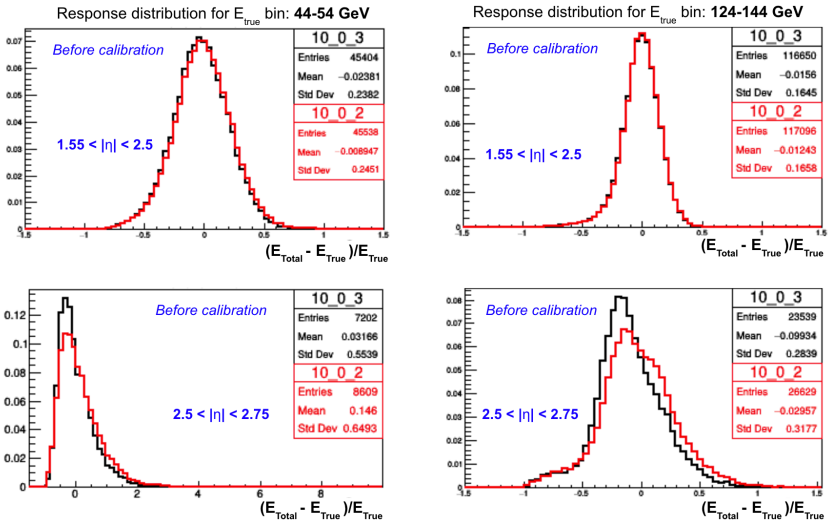

FIGURE 4.14: 1D Distribution of corrected energy response, i.e. $(E_t - E^{raw}_{total})/E_t$ of the EH hadrons in the endcap within tracker region (top row) and outside tracker region (bottom row) for "10_0_2" (red) and "10_0_3" (black) samples. 1D energy response distributions in first column represent EH hadrons having E_t = 44-54 GeV and similarly, second column represent E_t = 124-144 GeV.

the ECAL endcap outside the tracker region (i.e., $|\eta| > 2.5$). The energy response distributions for "10_0_3" show larger deviations from unity, indicating a degradation in the energy response compared to "10_0_2". To address this issue, additional calibrations are derived for the "10_0_3" sample in the endcap region for the EH hadrons. These new calibrations, referred to as "10_0_3" calibrations, aim to mitigate the impact of increased ECAL noise in the endcap and improve the charged hadron energy response in this region. Hence we have derived the calibrations for "10_0_3" samples and the results are shown in Figure 4.15.

4.7 Machine learning techniques for PF hadron energy calibration

PF charged hadron calibration using χ^2 method results in linear response and good resolution but still further improvements can be done for response in transition region and better resolution. For these improvements, we have tried out dynamic reduction network (DRN) method based on graphical neural networks (GNN). The DRN architecture was originally developed for energy regression in high-energy physics, specifically for hadrons [103]. Our group members (Alpana & Nitish) have tried DRN for energy regression of charged pion showers in HGCAL, compared the results with χ^2 and also validated with test beam data [104]. These studies give the motivation to use DRN model for PF charged hadron calibration for the CMS detector Run3. The detailed granularity of the ECAL shows great promise in improving the energy resolution using DRN. This, in turn, leads to more accurate and dependable measurements of charged hadron energies and significantly improves the performance of various physics analyses at the LHC.

DRN has also found application in the energy regression of electrons and photons within the current CMS ECAL [105]. It dynamically reduces noise and unwanted contributions while

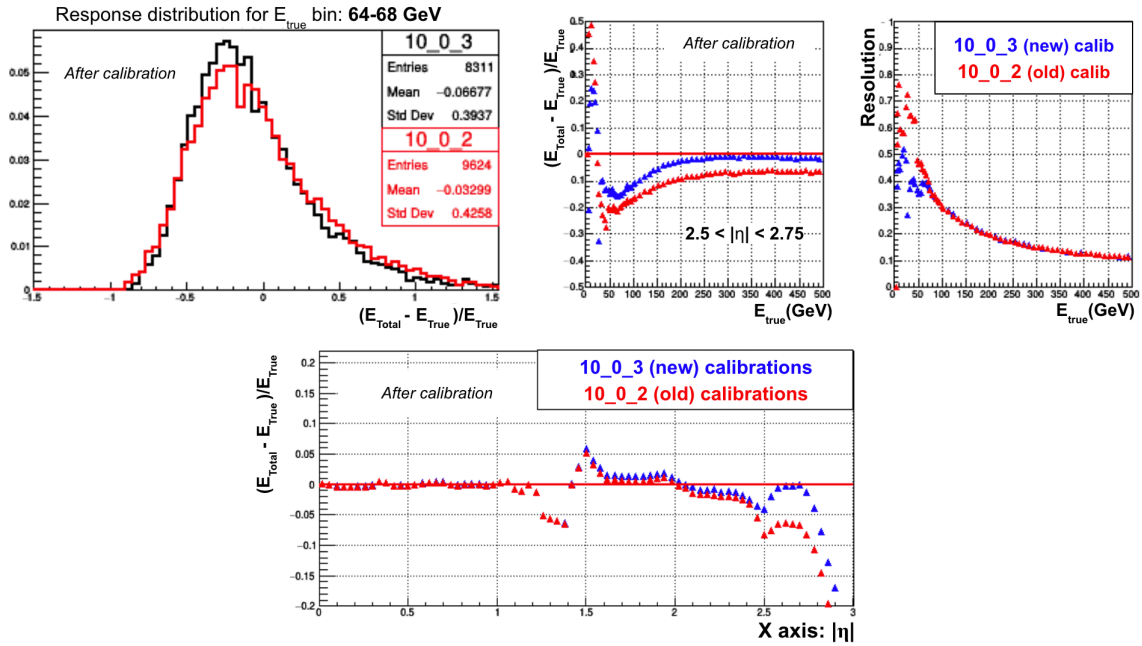

FIGURE 4.15: Comparison of the 1D distributions of the corrected energy response, where "10_0_3" and "10_0_2" calibrations are applied to the "10_0_3" sample. The top-left plot shows the raw energy response distribution, while the top-right plot illustrates the response and resolution as functions of the true energy (E_t) of the EH hadrons. The bottom plot shows the response as a function of $|\eta|$.

preserving crucial information, resulting in more accurate energy measurements for charged hadrons, electrons, and photons. This approach involves the creation of a model based on dynamic graph neural networks, which effectively maps input features into a higher-dimensional latent space. Furthermore, it incorporates clustering and pooling steps to aggregate information. The information flow from input to the required target is shown in Figure 4.16, and more detailed studies are done by our group member (Alpana) in case of π^- energy reconstruction in a high granularity sampling calorimeter (HGCAL) Beam Test prototype can be found at this reference [104].

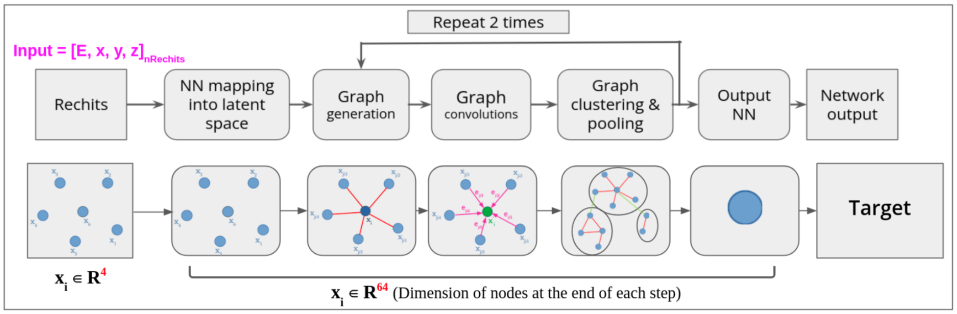

FIGURE 4.16: A schematic diagram for dynamic reduction network model. [104]

4.7.1 DRN model for PF calibration

In the reconstruction of an incident particle, it is represented as a collection of calibrated per-crystal energy deposits known as rechit. Each particle can have from 1 to over 100 rechits assigned to it. To account for various effects that degrade the resolution of the ECAL, corrections are applied on a per-event basis using machine learning regression techniques. PF charged hadron calibration is performed on Run 3 samples using two methodology: χ^2 method and DRN model. Technical details for Run3 samples are mentioned in Appendix Table A.8. The DRN features are similar to reference [104] and some distinctive features are as follows:

- **Inputs**: The DRN architecture takes the coordinates of these rechits in the (x, y, z) directions and the energy (E^{rechit}_{total}) per rechit as input. In top plots of Figure 4.17, rechits in x-y or z-y 2D planes are shown, which also shows the higher granularity in ECAL at x or y nearly around 100 cm. For the selection of relevant rechits per event, there is a matching of each rechit position with corresponding selected PF candidate (as mentioned in section 4.3) in $(\eta - \phi)$ plane using $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, such that $\Delta R < 0.2$ (0.1) for HCAL (ECAL) rechit clusters. In bottom-left plot of Figure 4.17, 2D distribution of E^{raw}_{total} with respect to E^{rechit}_{total} is showing the one-to-one correspondence for most events and on the bottom-right, 2D distribution of E_t with respect to E^{rechit}_{total} are shown.

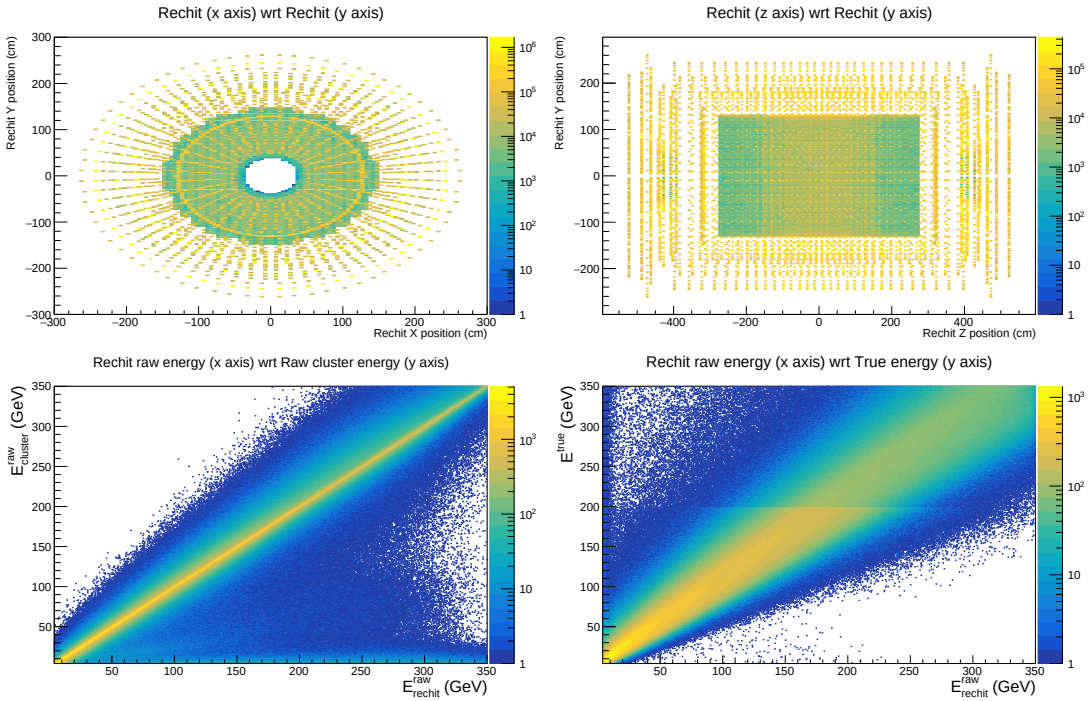

FIGURE 4.17: Rechits in x-y (top-left) and x-z (top-right) plane are shown. In bottom plots, 2D distributions of E^{raw}_{total} (left)and E_t (right) with respect to E^{rechit}_{total} are shown.

- **Target**: This DRN model predicts the energy of PF charged hadron event wise by using E_t as single target. In Figure 4.18, 1D distribution of E_t (left) is showing the uneven behaviour in the E_t after 200 GeV. For this scenario, events after 200 GeV are re-weighted by the factor 1.75 of Figure 4.18 (right).

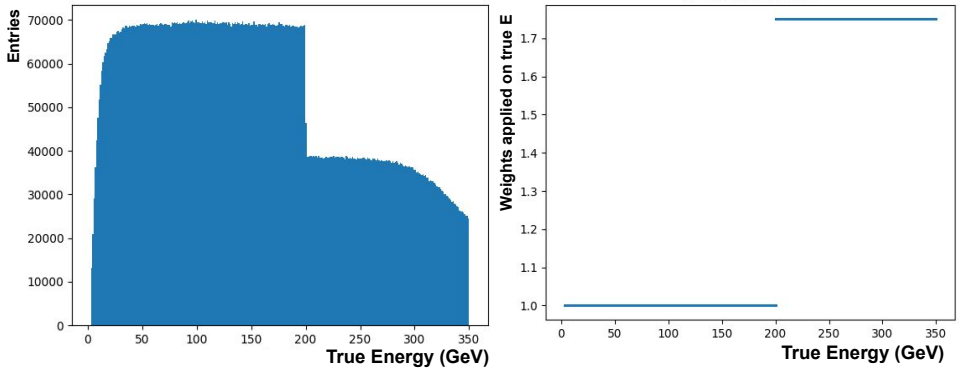

FIGURE 4.18: 1D distributions of E_t (left) and weights with respect to E_t (right) are shown.

Loss function is defined as $(E_t - E_{total}^{predicted})^2/E_t$ in this model, similar to reference [104]. After training (with 80% of total events) and validation (with 20% of total events), the corresponding loss plot for 100 epochs is shown in Figure 4.19.

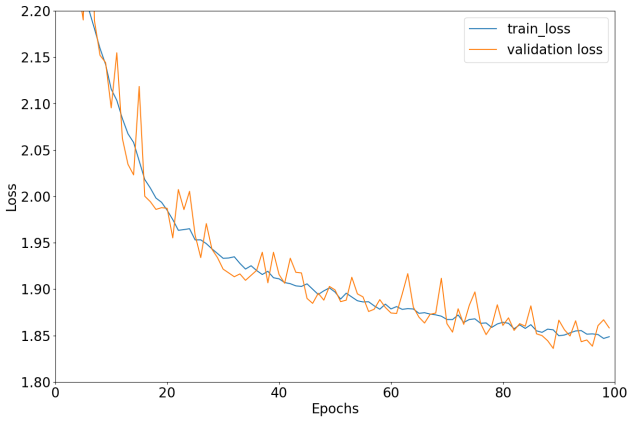

FIGURE 4.19: Validation and Training Loss are overlaid with respect to epochs.

4.7.2 Results

Comparisons between the finally corrected results by χ^2 and DRN method are shown in figures 4.20, 4.21, 4.22, 4.23 & 4.24 which show the good improvements for both EH and H hadrons.

1D energy response distributions

For EH and H hadrons, 1D energy response subsequent to the application of χ^2 calibration (in red), and DRN calibration after training (in blue) & after validation (in green) techniques, are shown. The width of gaussian distribution is relatively lesser for DRN method.

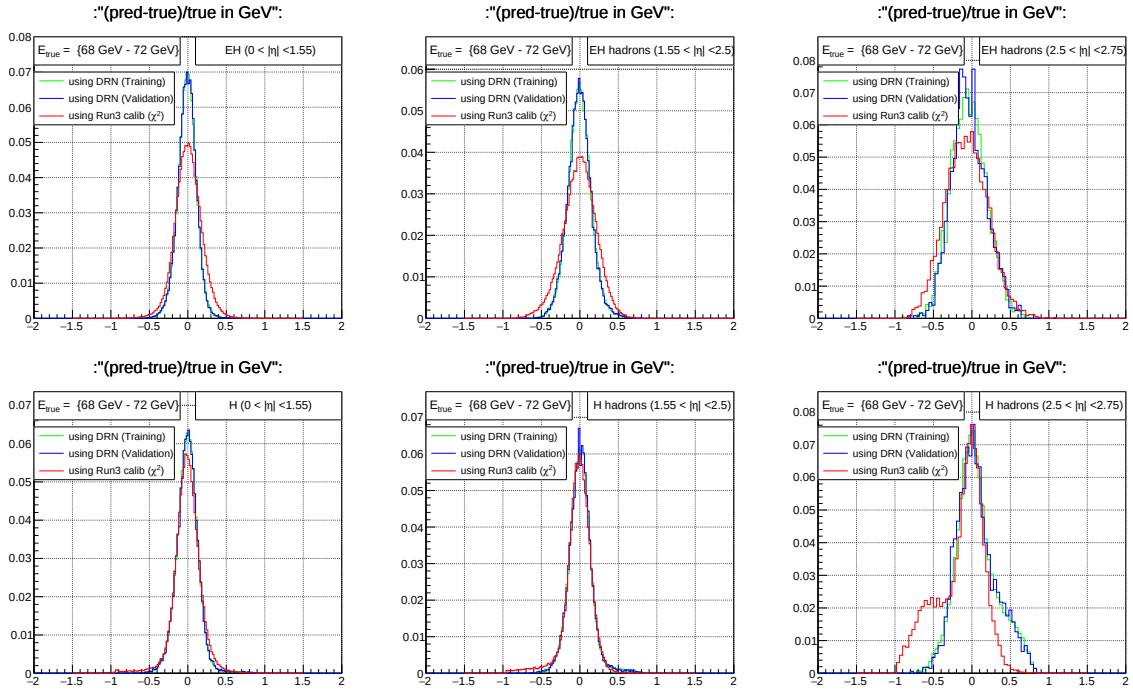

FIGURE 4.20: 1D energy response distribution after calibrations using χ^2 (red), DRN (blue & green) calibrations for EH hadrons(top) and for H hadrons(bottom). These energy responses are in the E_t range of 68–72 GeV for various η regions.

Energy response with respect to E_t

For EH hadrons, energy response (upper panel) and resolution (lower panel) subsequent to the application of χ^2 calibration (in red), and DRN calibration after training (in blue) & after validation (in green) techniques, are shown with respect to true energy (E_t). It is to be noted that any potential edge effects in the energy region beyond $E_t = 250$ GeV are temporarily disregarded for the present analysis. There is good improvement in resolution of EH hadrons by 10-50%.

Similar plots for H hadrons are also shown in Figure 4.22.Performance for the H hadrons in endcaps can be further investigated and could be a potential future direction for improvements.

Energy response with respect to $|\eta|$

For EH and H hadrons, energy response with respect to $|\eta|$ is getting flatter after using DRN especially in transition region ($1.26 < |\eta| < 1.68$) as shown in the figures 4.23 & 4.24.

4.7.3 Summary

The application of the DRN introduces potential advancements in PF hadron calibrations for charged hadrons. The integration of ECAL granularity through rechits as DRN model input showcases resolution improvements ranging from 10% to 50%, surpassing the performance of the conventional χ^2 method. Remarkably, the response of EH or H hadrons concerning

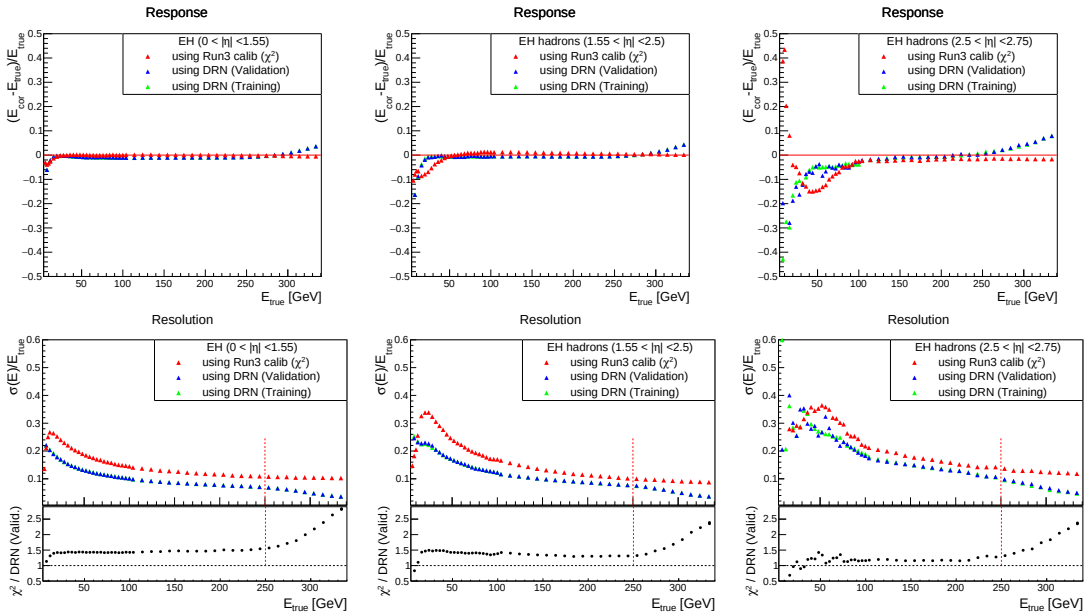

FIGURE 4.21: Average energy response (top) and resolution (bottom) after using χ^2 (red), DRN (blue & green) calibration methods, as a function of E_t for EH hadrons. After red dashed line at E_t = 250 GeV, there is an edge effect and that can be ignored for now.

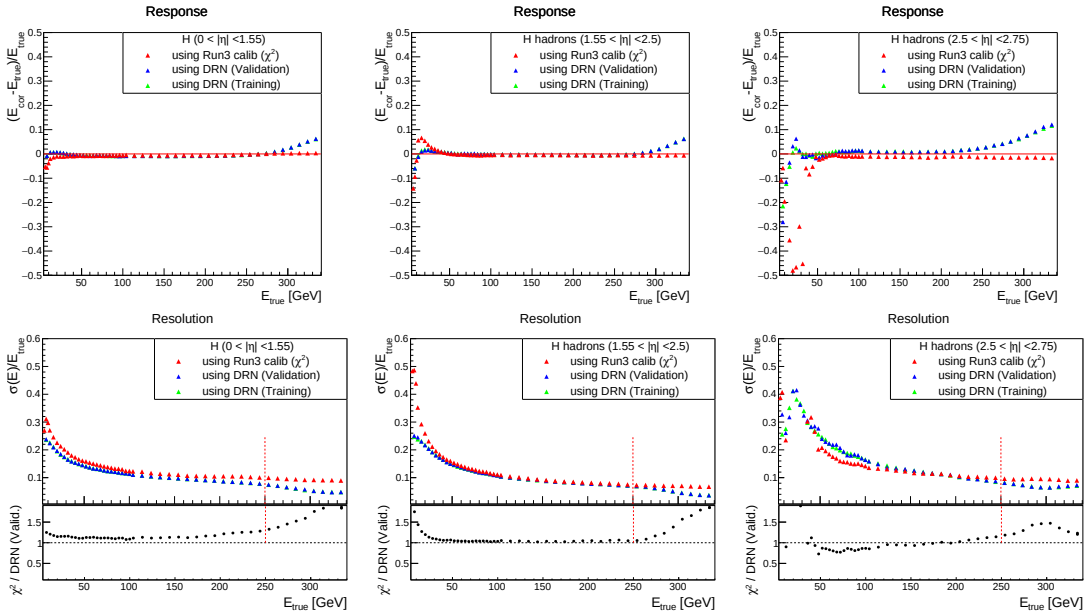

FIGURE 4.22: Average energy response (top) and resolution (bottom) after using χ^2 (red), DRN (blue & green) calibration methods, as a function of E_t for EH hadrons. After red dashed line at E_t = 250 GeV, there is an edge effect and that can be ignored for now.

E_t and η demonstrates remarkable stability around the zero point as shown in Figure4.25. These encouraging findings pave the way for forthcoming efforts geared towards refining the accuracy of the ultimate PF jet reconstruction.

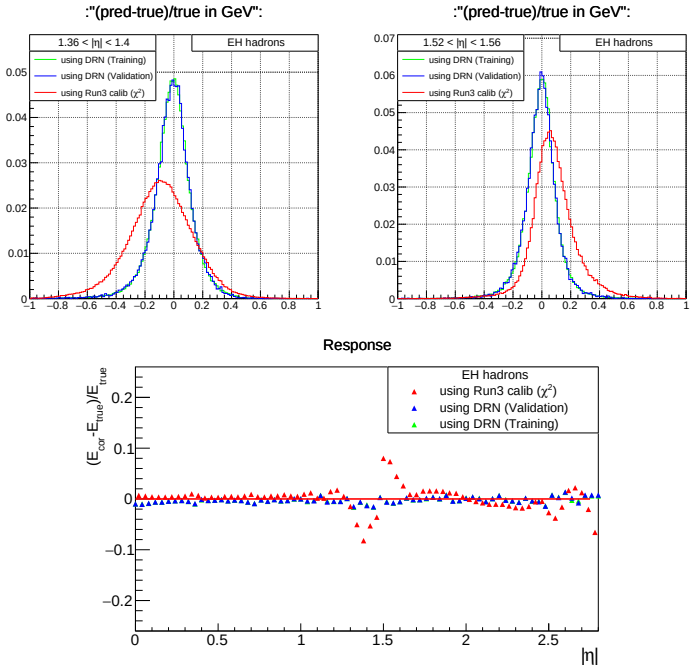

FIGURE 4.23: 1D-energy response distributions of EH hadrons (on top) for different η bins and inclusive in E_t. Average energy response (bottom) after using χ^2 (red), DRN (blue & green) calibration methods, as a function of $|\eta|$ for EH hadrons.

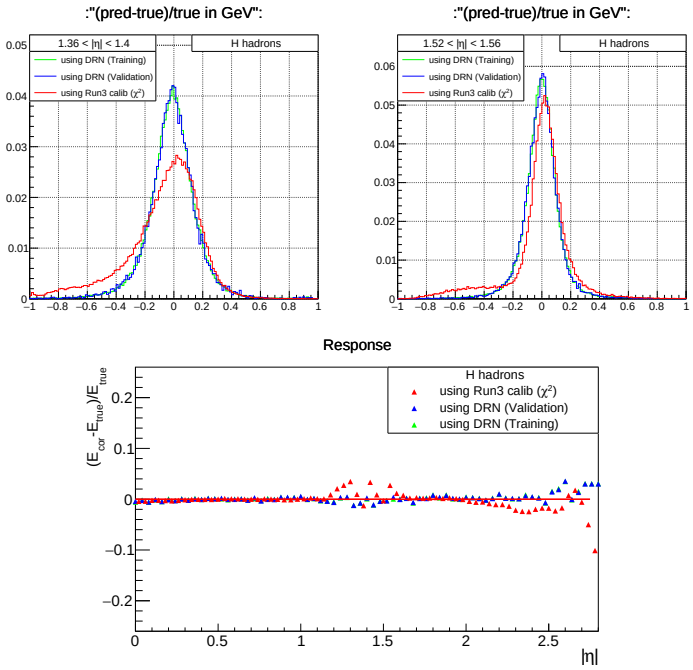

FIGURE 4.24: 1D-energy response distributions of H hadrons (on top) for different η bins and inclusive in E_t. Average energy response (bottom) after using χ^2 (red), DRN (blue & green) calibration methods, as a function of $|\eta|$ for H hadrons.

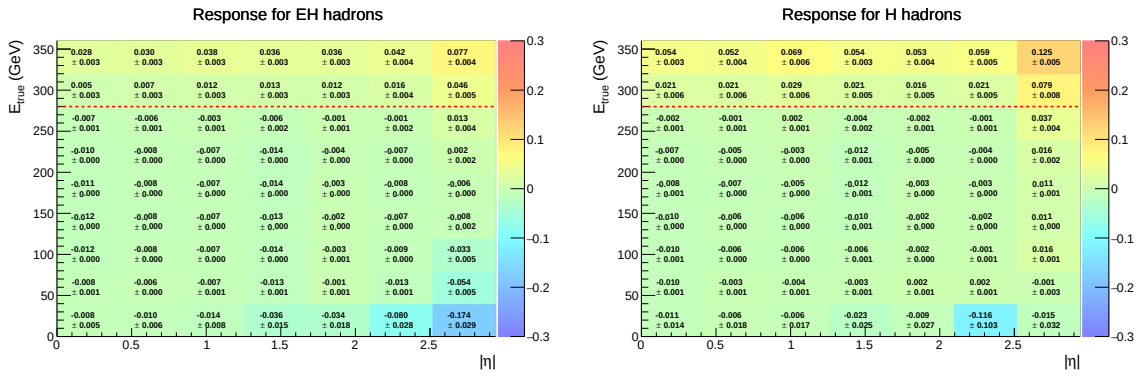

FIGURE 4.25: Energy response with respect to E_t and η for EH (left) & H (right) hadrons.

Chapter 5

A search for supersymmetry with at least one photon, large MET and jets final states

In this chapter, I summarize the salient features of designing a generic search for SUSY particles using the pp collision events at $\sqrt{s}$ = 13 TeV which contain multiple jets, at least one photon and large p_T^{miss} using full Run II (2016-2018) dataset corresponding an integrated luminosity of 137 fb^{-1} collected by the CMS detector. We target particles produced via strong as well as electroweak interactions. Searches for gluinos and squark pair production are generally based on dividing the search regions using number of jets and p_T^{miss} to increase sensitivity over contributions of the SM processes. We extended the search regions using boosted W or Z boson activity to search for electroweakino pair production in this analysis.

The GMSB models, summarized in this chapter, with R parity conservation are used as a guide to optimize the search regions and to interpret the results in terms of lower limits on SUSY particle masses. The gravitino LSP results in p_T^{miss} signature and decays of NLSP gives rise to an LSP, γ/Z/W/H boson, and jets. Then we will be discussing the event selections that can suppress the SM backgrounds and corresponding HLT trigger efficiencies calculations using data. At last, we will give the full overview of optimisation studies and analysis strategy. Relevant theoretical aspects of the GMSB model along with standard model are described in the chapter 2 and insight of the CMS detector with its working is explained in the previous chapters 3 and 4.

5.1 Targeted SUSY signal models and SM backgrounds

We interpret the results of this search in simplified models [58, 59, 106, 107, 108] of SUSY particle production by strong and electroweak interactions. In GMSB scenarios conserving R-parity, the final states typically comprise multiple jets, large p_T^{miss} , and at least one photon, Z, W or H boson.

Strong production involves the production of gluinos or squarks, as they also contain color charges similar to their SM partners. Representative diagrams depicting simplified models of

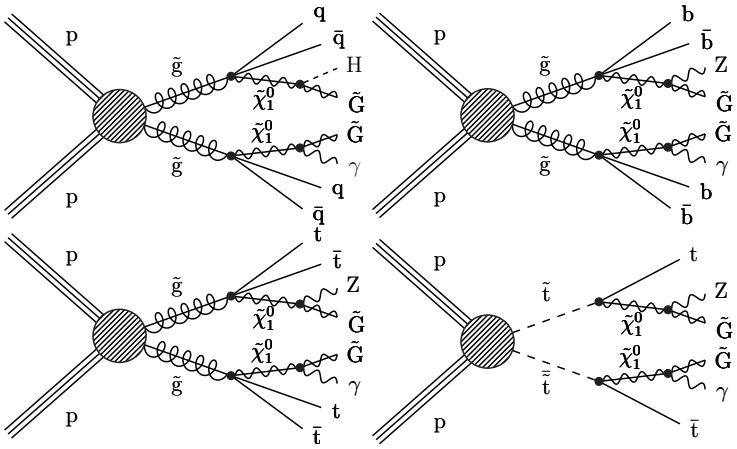

FIGURE 5.1: Simplified model scenarios for gluino pair production, T5qqqqHG (top left), T5bbbbZG (top right), T5ttttZG (lower left), and stop pair production T6ttZG (lower right). The models are defined in the text.

gluino pair production (prefixed with T5) and squark pair production (prefixed with T6) are shown in Figure 5.1. In the T5 models, the gluino decays to a neutralino ($\tilde{\chi}_1^0$) and a pair of light flavored quarks ($q\bar{q}$), bottom quarks ($b\bar{b}$) or top quarks ($t\bar{t}$). The T5qqqqHG is defined by decay of gluinos to $q\bar{q}$ and $\tilde{\chi}_1^0$ followed by $\tilde{\chi}_1^0 \to$ H $\tilde{G}$ or $\tilde{\chi}_1^0 \to \gamma\tilde{G}$ with equal probability. The mass of $\tilde{\chi}_1^0$ is assumed to be 127 GeV or above in this model, and hence the H boson is always on-shell. In the T5bbbbZG and T5ttttZG, $\tilde{g} \to b\bar{b}\,\tilde{\chi}_1^0$ and $\tilde{g} \to t\bar{t}\,\tilde{\chi}_1^0$, respectively, and $\tilde{\chi}_1^0$ decays to Z $\tilde{G}$ or $\gamma\tilde{G}$ with branching fraction for each assumed to be 50%. In the top squark pair production model T6ttZG, each stop decays to a top quark and a $\tilde{\chi}_1^0$ followed by equal branching fractions for the decays $\tilde{\chi}_1^0 \to$ Z $\tilde{G}$ or $\tilde{\chi}_1^0 \to \gamma\tilde{G}$. In the models involving $\tilde{\chi}_1^0 \to$ Z $\tilde{G}$, $\tilde{\chi}_1^0$ masses down to 10 GeV are considered.

In electroweakino pair productions, charginos and neutralinos are produced directly and decay into given final states. These charginos are a mixture of a $W^{\pm}$ and a charged higgsino, whereas neutralinos are a mixture of bino, W^0 and a neutral higgsino. In the TChiWG model, Figure 5.2 (top) branching fractions for $\tilde{\chi}_1^0 \to \gamma\tilde{G}$ and $\tilde{\chi}_1^{\pm} \to$ W$\tilde{G}$ are assumed to be 100%. The TChiNG model comprises the production of $\tilde{\chi}_1^{\pm}\,\tilde{\chi}_1^{\pm}$, $\tilde{\chi}_1^{\pm}\,\tilde{\chi}_1^0$ and $\tilde{\chi}_2^0\,\tilde{\chi}_1^0$ pairs where $\tilde{\chi}_1^{\pm}$, $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^0$ are assumed to be nearly mass degenerate such that the decays of $\tilde{\chi}_1^{\pm}$ and $\tilde{\chi}_2^0$ to $\tilde{\chi}_1^0$ are accompanied by low-momentum particles outside the kinematic acceptance of this analysis. Only $\tilde{\chi}_1^{\pm}\,\tilde{\chi}_1^0$ model is depicted in the Figure 5.2 (bottom left). The $\tilde{\chi}_1^0$ is assumed to decay to $\gamma\tilde{G}$, Z$\tilde{G}$ and H$\tilde{G}$ with branching fractions of 50%, 25% and 25%, respectively. There is another TChiNG model known as TChiNGnn model, having 100% branching fraction for $\tilde{\chi}_1^0 \to$ Z/H$\tilde{G}$ and $\tilde{\chi}_2^0 \to \gamma\tilde{G}$ as shown in the Figure 5.2 (bottom right). For TChiNGnn model, only $\tilde{\chi}_1^0\,\tilde{\chi}_2^0$ process will contribute.

All the decays of SUSY particles are assumed to be prompt. In all the models, the mass of gravitino, $m_{\tilde{G}}$, is fixed to be 1 GeV. The event kinematic properties do not depend strongly on the exact choice of $m_{\tilde{G}}$ in the phase space explored in this analysis.

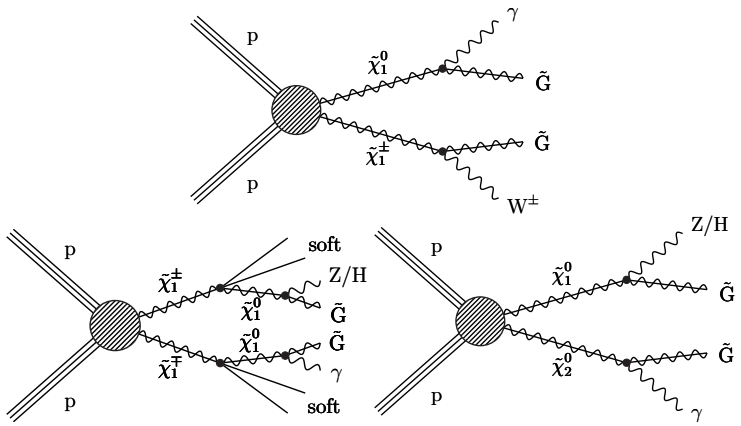

FIGURE 5.2: Simplified model scenarios for electroweakinos pair production, TChiWG (left), and TChiNG (right). The models are defined in the text.

5.2 Standard model backgrounds

Several SM processes with similar final states can imitate signal final states and will therefore be in the background. Some major background processes can contribute to signal regions having at least one γ, large p_T^{miss} , and many hadronic jets.

5.2.1 tt + jets and W + jets process

The W boson decays hadronically or leptonically, with a branching ratio (BR) as shown in Table 5.1. Top quark (t) decays into W boson and b quark with 100% BR and $t\bar{t}$ process also has semi-leptonic decays as there will be two W's. To mimic the signal final states, these backgrounds with leptonic decays have p_T^{miss} from neutrinos, and a photon due to an electron's misidentification as photon, contributing to the signal region as shown in Figure 5.3 (top left). An electron from W decay can fake as photons if some pixel hits associated with ECAL clusters are missed while identifying the electron. This is a major background for this analysis and will be discussed in the section 6.2. In other sceanrios, ISR/FSR jets can also fake as a photon.

5.2.2 tt + γ and W + γ process

Similar to the $t\bar{t}$ /W + jets processes, $t\bar{t}+\gamma$ and $W(\ell\nu)+\gamma$ also decays leptonically or hadronically, but in addition with a gen-level prompt photon. Any overlapping between $t\bar{t}$ /W + jets and $t\bar{t}+\gamma$ /$W(\ell\nu)+\gamma$ is taken care of by applyig some selections as discussed in the section 6.1. To mimic the final states of signals, leptonic decays with neutrinos make up most of the p_T^{miss} contribution. These backgrounds will be highly reduced by vetoing any leptonic activity, and the remaining background will contribute due to the failure in reconstruction, identification, or isolation of leptons in the detector, as shown in Figure 5.3 (top right). It will be discussed in the section 6.1. In case of $W \to \tau\nu_\tau$, taus can decay leptonically (with 35% BR) and hadronically (with 65% BR) 5.1, and hadronic decaying taus can be reconstructed as jets if not rejected by isolated track veto as explained in the section 5.5.1.

TABLE 5.1: Decay channels of W boson with branching fractions [109].

Decay Channel	**Branching Fractions (%)**
$W \to \ell^- \bar{\nu}_\ell$ (leptonic)	-
$e^- \bar{\nu}_e$	10.8
$\mu^- \bar{\nu}_\mu$	10.8
$\tau^- \bar{\nu}_\tau$	10.8
$W \to h^+ h^-$ (hadronic)	33.6
$\tau \to \ell^- \bar{\nu}_\ell \nu_\tau$	35.6
$\tau \to$ hadrons (h)	-
$\tau \to h^- \nu_\tau$	11.5
$\tau \to h^- \pi^0 \nu_\tau$	26.0
$\tau \to h^- \pi^0 \pi^0 \nu_\tau$	9.5
$\tau \to h^- h^+ h^- \nu_\tau$	9.8
$\tau \to h^- h^+ h^- \pi^0 \nu_\tau$	4.8
others	3.3

5.2.3 Z($\nu\nu$) + γ + jets process

The Z boson decay into leptons (with 10% BR), neutrinos (with 20% BR) and hadrons (with 70% BR) as shown in Table 5.2. As compared to other Z+jets decay mode, Z decaying to two neutrinos will give the adequate p_T^{miss} according to the final state's requirement as shown in Figure 5.3 (bottom left). This background is irreducible and has true p_T^{miss} due to the presence of neutrinos in the final state. These events mainly contribute in low N_{jets} bins with no b-tags and in EW SR bins. The background estimation uses $Z(\ell^+\ell^-)\gamma$ + jets CR events from data to correct for the normalization and the shapes of p_T^{miss} and N_{jets} are taken from MC.

TABLE 5.2: Z Boson Decays and Branching Fractions

Decay Channel	Branching Fraction (%)
$Z \to \ell^- \ell^+$ (leptonic)	-
$e^- e^+$	3.365
$\mu^- \mu^+$	3.365
$\tau^- \tau^+$	3.370
$Z \to \nu\bar{\nu}$	20
$Z \to q\bar{q}$ (hadronic)	69.91

5.2.4 QCD and γ + jets process

The QCD and γ+jets events with a well-reconstructed photon can mimic the signal final states, if they also contain large p_T^{miss} resulting from significant mismeasurement of one or more jets or from the presence of neutrinos from the semileptonic decays of heavy-flavor hadrons. The QCD and γ + jets process has contribution from light flavour quark jets and gluon jets as shown in Figure 5.3 (bottom right). QCD and γ + jets will be abundant in $\Delta\phi(\vec{p}_T^{\,\text{jet}}, \vec{p}_T^{\,\text{miss}})$ <0.3 and p_T^{miss} <300 GeV (will be explained in section 6.4).

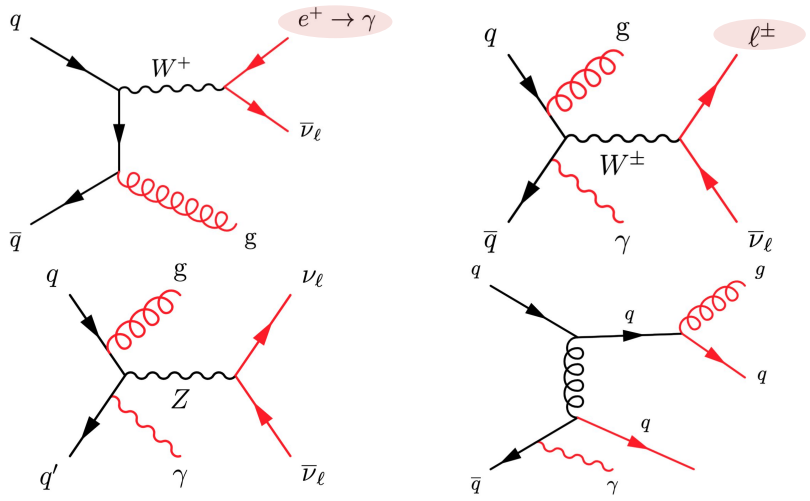

FIGURE 5.3: Standard model background processes having at least one γ, p_T^{miss} and hadronic jets as a final state.

5.3 Monte Carlo simulation of SM processes

For most SUSY searches, the estimation of backgrounds relies on a combination of observed data and monte carlo (MC) events. These monte carlo (MC) events used to derive correction factors and scale factors, to validate various background estimation methods, and to assess systematic uncertainties. We simulate pp collision as MC events at higher energies using various generators like MADGRAPH5_AMC@NLO , PYTHIA , GEANT4 etc. Each generator has its specific role to incorporate physics level and realistic detector effects in the MC simulations. The following tools are often used in conjunction to simulate and analyze high-energy particle collisions, understand the physics processes involved, and compare the results with experimental data.

5.3.1 MADGRAPH5_AMC@NLO Generator

MADGRAPH5_AMC@NLO is used for generating matrix elements (amplitudes) for particle physics processes and simulating MC events. It is primarily used for simulating high-energy particle collisions. Key features of MADGRAPH5_AMC@NLO include:

- Matrix element generation: MADGRAPH5_AMC@NLO generates tree-level matrix elements for specific processes and also incorporates the next-to-leading order (NLO) corrections using the aMC@NLO framework. It calculates the amplitudes for these processes by merging matrix elements at NLO accuracy with parton showering provided by external programs such as PYTHIA .
- Parton distribution functions (PDFs): MADGRAPH5_AMC@NLO interfaces with external PDF libraries to incorporate the effects of parton distribution functions in the simulation. The PDFs are mathematical functions to describe the distribution of fractional momenta of quarks and gluons (collectively known as partons) within proton. These distributions provide information about the probability of finding a parton with a specific momentum fraction within the proton.

- Automatic generation of Feynman diagrams: MADGRAPH5_AMC@NLO employs an automated process for generating Feynman diagrams relevant to the specified process, allowing for efficient calculation of matrix elements.

5.3.2 PYTHIA Generator

It provides a framework for modeling the entire event, including initial state radiation, parton showering, hadronization, and decay processes. Key features of PYTHIA include:

- Parton showering and hadronization: PYTHIA simulates the evolution of partons (quarks and gluons) produced in the initial collision, including parton showering (radiation of additional partons) and hadronization (formation of color-neutral hadrons from partons).
- Particle decays: PYTHIA models the decays of unstable particles, including leptons, quarks, and gauge bosons, based on known branching fractions and decay modes.
- Multiple interactions: PYTHIA incorporates the effects of multiple parton interactions within a single collision event.
- Underlying event and beam remnants: PYTHIA includes modeling of the underlying event, which accounts for additional particle production not directly associated with the primary collision process, as well as remnants of the colliding particles.

5.3.3 GEANT4

GEANT4 is a powerful simulation toolkit widely used in particle and nuclear physics, as well as other fields, to simulate the interactions of particles with the detector material and electronics effects. It provides a comprehensive set of tools for modeling particle transport, electromagnetic interactions, hadronic interactions, and detector response. Key features of GEANT4 include:

- Geometry modeling: GEANT4 allows users to define the geometries of detectors and materials, including complex shapes and hierarchical structures.
- Particle tracking: GEANT4 tracks particles through the defined geometries, simulating their interactions with matter based on physics models.
- Electromagnetic and hadronic physics: GEANT4 provides a range of models to simulate electromagnetic interactions (e.g., ionization, bremsstrahlung, pair production) and hadronic interactions (e.g., elastic scattering, inelastic scattering, particle production) with high accuracy.
- Particle detectors: GEANT4 can simulate the response of particle detectors, including energy deposition, signal generation, and electronic noise.

In this analysis, we utilize the MADGRAPH5_AMC@NLO event generator (version 2.2.2 for 2016 and 2.4.2 for 2017-2018) [110, 111] to simulate the production of various background processes and signal events in proton-proton collisions. The MADGRAPH5_AMC@NLO event generator calculates matrix level elements for the cross-section of the produced particles with the highest momentum in proton-proton collision.

We use MADGRAPH5_AMC@NLO generator at leading order (LO) precision for simulating the production of the $t\bar{t}$ + jets, W + jets, Z($\nu\nu$) + jets, QCD multijet, and W + γ processes. For the $t\bar{t}$ + γ process, we simulate events at next-to-leading order (NLO) precision. The $t\bar{t}$ + γ events are simulated with up to one additional parton at the matrix element level, while the W + γ events have up to two additional partons, the $t\bar{t}$ + jets events up to three additional partons, and the other samples up to four additional partons. The single top quark process is modeled at NLO in perturbative QCD, with MADGRAPH5_AMC@NLO used for s-channel production.

Signal events are generated using the MADGRAPH5_AMC@NLO generator at LO precision, similar to the SM backgrounds, allowing for up to two additional partons at the matrix element level. The decays of gluinos, top squarks, and neutralinos are modeled using PYTHIA 8 [112].

For the background samples, parton showering and hadronization are performed using the PYTHIA 8.212 generator for 2016 with CUETP8M1 [113] and PYTHIA 8.226 (8.230) with the CP5 underlying event tune for 2017 (2018). Signal samples utilize the same underlying event tune as the respective year but use PYTHIA versions 8.226 (2016) and 8.230 (2017-2018) [114].

To avoid double counting of partons generated with MADGRAPH5_AMC@NLO and PYTHIA , we employ the MLM [115] and FxFx [116] matching schemes in LO and NLO samples, respectively. The cross-sections used for normalizing the signal yields are computed at NLO plus next-to-leading logarithmic (NLL) precision. These simulation techniques and event generation procedures are crucial for understanding the expected signal and background contributions in the analysis [117, 118, 119, 120, 121].

5.3.4 Stitching of simulated SM background processes

As we discussed in section 5.2, $t\bar{t}$ /W +jets and $t\bar{t}$ /W + γ are similar processes except prompt photon at MADGRAPH5_AMC@NLO level and using both simultaneously can introduce double counting of events. For $t\bar{t}$ /W + jets and $t\bar{t}$ /W + γ, there is a need to avoid overlapping by using simulation level cuts. There is a cut of $\Delta R(\ell/q, \gamma) < 0.5$ in $t\bar{t}$ /W + γ processes at MADGRAPH5_AMC@NLO level, but no such cuts are not well modelled in $t\bar{t}$ /W + jets sample. Due to the lack of radiated photon at MADGRAPH5_AMC@NLO level in simulated samples, it is difficult to stitch $t\bar{t}$ /W + jets and $t\bar{t}$ /W + γ processes, so we need to take $\Delta R(\ell, \gamma_{\text{reco}}) > 0.5$ and $\Delta R(q/g, \gamma_{\text{reco}}) > 0.5$ as shown in Figure 5.4. For $t\bar{t}$ /W + jets, inverse selections as in the case of $t\bar{t}$ /W + γ processes, are used.

These are not a perfect modelling, so any discrepancies between PYTHIA or MADGRAPH5_AMC@NLO level cross-sections and cuts while stitching is taken care by systematics, which will be discussed in section 6.1.5.

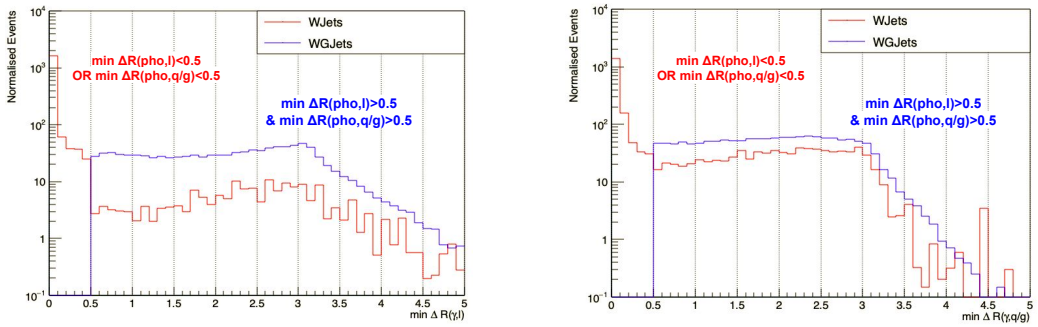

FIGURE 5.4: The distribution of minimum $\Delta R(l, \gamma_{\text{reco}})$ and $\Delta R(q/g, \gamma_{\text{reco}})$ for stitching W + jets and W + γ processes.

5.4 Triggers and Efficiency

There are L1 and HLT level triggers used for selecting or saving events for offline analysis and then event reconstructions within the time scale of 25 ns to 300 μs. L1 triggers are hardware triggers and HLT triggers are mostly software based triggers as discussed in the section 5.4 of chapter 3. There are some un-prescaled offline HLT triggers available that can be used for the analysis, based on atleast one photon, large p_T^{miss} and number of jets (Table 5.3).

TABLE 5.3: Available p_T^{miss} $\times$ MHT triggers and single photon triggers

Year	Triggers
2016	HLT_PFMETx_PFMHTx_IDTight (x = 90, 100, 110, 120)
	HLT_Photon165_HE10 OR HLT_Photon90_CaloIdL_PFHT600
2017	HLT_PFMETx_PFMHTx_IDTight (x = 110, 120, 130, 140)
	HLT_Photon200
2018	HLT_PFMETx_PFMHTx_IDTight (x = 90, 100, 110, 120)
	HLT_Photon200

In above table, x is the p_T^{miss} or missing hadronic transverse momentum (MHT) threshold (in GeV) for available p_T^{miss} $\times$ MHT trigger.

5.4.1 Trigger efficiency for signal region

For event selections (Section 5.5.1), trigger studies are done using inclusive p_T^{miss} triggers (Table 5.3). These p_T^{miss} X MHT trigger efficiency are calculated using single electron datasets as reference which has the electron triggers with lower threshold of p_T like, `HLT_Elex_WPTight_Gsf_v` (x=27, 30, 32 for 2016 and x=27, 32 for 2017, 2018). On single electron dataset, further event selections imposed are $S_T(p_T^{\gamma} + \sum_{jets} p_T^{jets}) > 300$ GeV, at least one photon with $p_T^{\gamma} > 100$ GeV and $\eta^{\gamma} < 2.4$, $N_{jets} \geq 2$ and azimuthal angle separation $\Delta\phi(jet, MET) > 0.3$ for first two leading jets with default p_T^{miss} cleaning applied. Here $\Delta\phi$ condition on leading two jets is to enforce that p_T^{miss} is dominated by true leptonic decay and not due to jet mismeasurement, where the latter is dominant in the QCD multijet processes. The efficiency of p_T^{miss} triggers using single electron events is calculated as follows:

$$\epsilon_{MET} = \frac{\text{Events passing through } p_T^{\text{miss}} \text{ and single electron triggers}}{\text{Events passing through single electron triggers}} \tag{5.1}$$

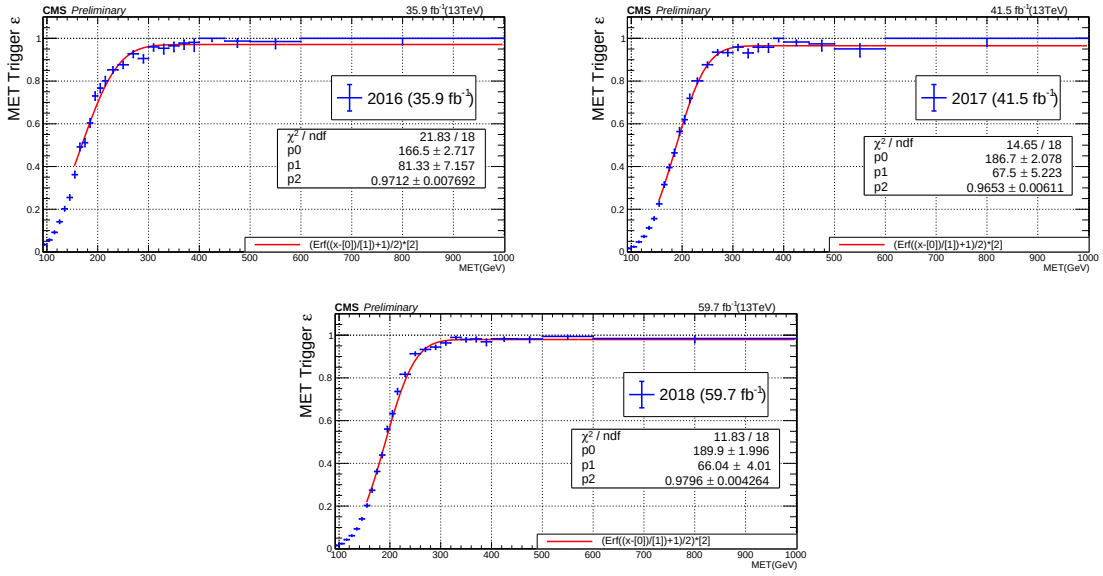

FIGURE 5.5: MET X MHT trigger efficiency, ϵ_{MET} with respect to p_T^{miss} using single electron dataset for years 2016, 2017, 2018.

Figure 5.5 shows that beyond cut off of $p_T^{\text{miss}} > 250$ GeV, the plateau region has efficiency $\epsilon_{MET} > 97\%$ for all three years. Note that trigger efficiency is parameterized as a function of p_T^{miss} and is inclusive in other variables. This efficiency as a function of p_T^{miss} is applied to MC simulated samples.

Issues with available single photon triggers

There is an obvious question why photon triggers have not been used for this analysis. The availability of relevant triggers partly drives the analysis phase space. Initially, single photon triggers for 2017 are investigated for the analysis by calculating trigger efficiencies but found some hiccups even for photon-pT > 200 GeV as shown in Figure 5.6.

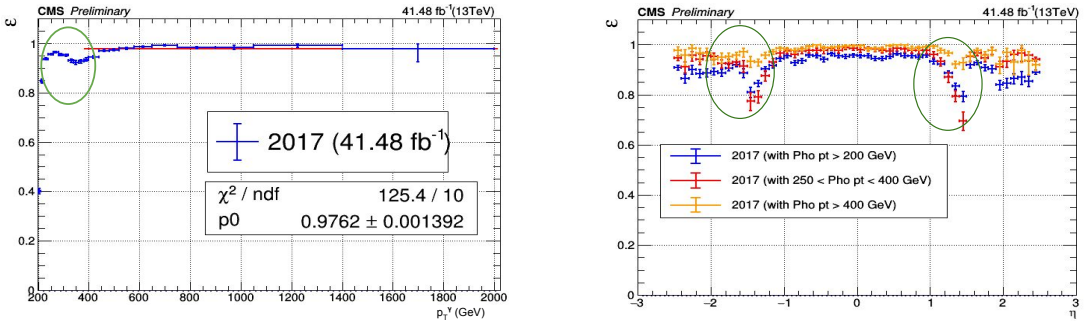

FIGURE 5.6: The photon trigger efficiencies (ϵ) wrt transverse momenta of photon (p_T^{γ}) (left) and pseudorapidity of photon (η). Green circles shows kinks in trigger efficiencies for both left and right figures.

Although single photon triggers would be useful to lower the p_T^{miss} thresholds, in our signal phase space of interest, but we do not loose much by not being able to use these paths.

5.4.2 Trigger efficiency for control region

For calculating the p_T^{miss} X MHT trigger efficiency in the low-$\Delta\phi$ control region of the QCD multijet process, at first JetHT events with significant jet activity and hadronic transverse energy are selected, and corresponding triggers are shown in Table 5.4. S_T, γ, N_{jets} selections are similar to the signal region case, but here azimuthal angle separation $\Delta\phi(jet, MET) < 0.3$ for the first two leading jets. The efficiency of p_T^{miss} triggers using JetHT events is calculated as follows :

$$\epsilon_{MET} = \frac{\text{Events passing through } p_T^{\text{miss}} \text{ and JetHT triggers}}{\text{Events passing through JetHT triggers}} \tag{5.2}$$

TABLE 5.4: Available Jet and HT triggers

Year	Triggers
2016	HLT_PFHTx_v (x=180, 200, 250, 300, 350, 370, 400, 430)
	HLT_PFJetx_v (x = 450, 500, 550)
2017	HLT_PFHTx_v (x=180, 200, 250, 300, 350, 370, 400, 430)
	HLT_PFJetx_v(x = 450, 500, 550)
2018	HLT_PFHTx_v (x=180, 200, 250, 300, 350, 370, 400, 430)
	HLT_PFJetx_v(x = 450, 500, 550)

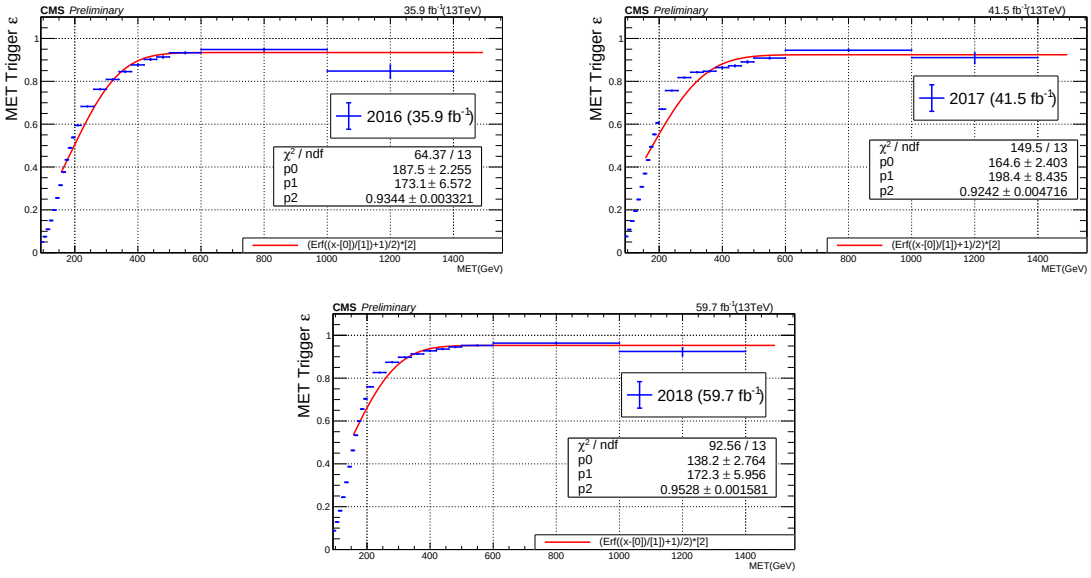

FIGURE 5.7: MET X MHT trigger efficiency in the control region of the QCD multijet process, ϵ_{MET} with respect to p_T^{miss} using JetHT dataset for years 2016, 2017, 2018 and their comparison.

Trigger efficiencies for events with reconstructed p_T^{miss} of at least 200 (300) GeV are found to be 70 (95), 60 (95), and 60 (97)% for data collected in 2016, 2017, and 2018, respectively. In

addition to the p_T^{miss} -triggered data set, single-electron and single-muon data sets are also used for the estimation of the background from electrons misidentified as photons, described in Section 6.2. The estimation of backgrounds primarily relies on observed data, with Monte Carlo (MC) events used to derive correction factors and scale factors, to validate various background estimation methods, and to assess systematic uncertainties. Some additional checks (in Appendix C) are also performed for the validity of these trigger efficiencies, in case if these are not dependent on the variation of the topologies in different samples and other factors.

5.5 Event selection

As shown in Figure 5.1 and 5.2, the final state is expected to contain large p_T^{miss} , a high p_T photons, large number of jets and/or bjets, and zero or more leptons. As this analysis based on final states with zero leptons and hence it will result into different composition of SM processes contributing to backgrounds. This distinct composition aids in both controlling systematic uncertainties and detecting the presence of a new physics signal independently and exclusively in the event of a discovery.

5.5.1 Baseline selections

Events are selected in the signal regions (SRs) using a set of criteria, referred to as the baseline selection, summarized in Table 5.5. The SR events are required to have $p_T^{\text{miss}} > 300$ GeV, and $N_{\text{jets}} \geq 2$, where N_{jets} is the number of AK4 jets with $p_T > 30$ GeV and $|\eta| < 2.4$. The events must also contain at least one photon with $p_T > 100$ GeV and $|\eta| < 2.4$. The photons are required to be separated from the jets by $\Delta R(j, \gamma) > 0.3$. The scalar sum of the p_T of the jets and the photon, denoted as ST, reflects the visible energy scale of the event and is required to be larger than 300 GeV. Significant mismeasurement of jet p_T can lead to high p_T^{miss} , which is generally aligned with the mismeasured jet. To suppress such events in the SRs, we require $\Delta\phi(\vec{p}_T^{\,\text{jet}}, \vec{p}_T^{\,\text{miss}}) > 0.3$ for the two highest p_T jets in the event. As the final state targeted contains only photons and hadronic jets, events with isolated electron and muon candidates are rejected to suppress the SM background arising from leptonic W decays. In order to further reject such SM events, events with an isolated leptonic or hadronic track are rejected. The event samples used for estimation of backgrounds are referred to as control regions (CRs) and are defined to be non-overlapping with the SRs. These are designed to be dominated by SM processes and are expected to have small signal contributions.

Object reconstruction details are discussed in section 3.4 of chapter 3. Following are the additional details of various objects selected for the analysis:

- **Missing transverse energy**: $p_T^{\text{miss}} > 200$ GeV where $p_T^{\text{miss}} = \left|\sum_{\text{pfCandidates}} \vec{p}_T\right|$ is based on trigger choice as explained in Section 5.4.1. Trigger efficiencies for events with reconstructed p_T^{miss} of at least 200 (300) GeV are found to be 70, 60, and 60% (95, 95, and 97%) for data collected in 2016, 2017, and 2018, respectively. The low p_T^{miss} region 200 GeV $< p_T^{\text{miss}} < 300$ GeV is used as QCD and $\gamma + jets$ control region while $p_T^{\text{miss}} > 300$ GeV

TABLE 5.5: Summary of the baseline selection criteria used to identify events of interest for this search.

p_T^{miss}	$>$ 300 GeV for SRs and [200, 300] GeV for low-p_T^{miss} CRs
N_{jets} $(p_T > 30\ GeV, \lvert\eta\rvert < 2.4)$	≥ 2
$(\gamma\ p_T > 100\ GeV, \lvert\eta\rvert < 2.4)$	≥ 1
$\text{ST} = \sum_{\text{jets}} p_T + p_T^{\gamma}$	$>$ 300 GeV
$\Delta\phi(\vec{p}_T^{\text{jet}}, \vec{p}_T^{\text{miss}}) > 0.3$	$>$ 0.3 for two highest p_T jets
Number of leptons (e, μ)	0
Number of isolated tracks	0

is used for signal region. The low p_T^{miss} control region is used for multijet background estimation as described in Section 6.4.

- **Jets**: Jets potentially dominated by contributions from anomalous detector signals or reconstruction failures, so all AK4 jets within $|\eta| < 2.4$ are required to pass the loose jet ID requirements as follows:
  - 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 matched to isolated photons within $\Delta R < 0.3$ are not subject to these requirements. Event should have minimum two jets with p_T^{jet} GeV 30 GeV , $|\eta^{jet}| < 2.4$.

  In addition to AK4 jets, the AK8 jets constructed using anti-kT algorithm with distance parameter 0.8 are used for Electroweak selections. The AK8 jet mass, m_J is reconstructed using the soft-drop algorithm [122], which improves jet mass resolution by removing soft and wide-angle contributions. Requirements are applied to m_J to identify W, Z, or H boson candidates. The same requirement $65 < m_J < 105$ GeV is used for both W and Z bosons, as the difference in the particle masses is smaller than the m_J resolution. Jets that pass this requirement are called V-tags, where V = W or Z. For H bosons, the requirement $105 < m_J < 140$ GeV is applied, and passing jets are called H-tags.

- **Photons**: Photon candidate whose supercluster is found to be in the pseudorapidity range of the ECAL crack ($1.44 < |\eta| < 1.56$) is not considered. Photon candidate is required to have a low H/E, have a transverse shower profile, $\sigma_{i\eta i\eta}$, that is consistent with a prompt photon and be. Isolation sums are computed by summing PF candidates in a

cone centered on the photon candidate with a radius of $R = 0.3$. Identification and isolation requirements are based on the recommended cut based loose working point [85] as shown in Table 3.3. Photons are required to have $p_T \geq 100$ GeV and $|\eta| < 2.4$, excluding the crack in the barrel-endcap transition.

- **Muon veto**: In the given final states, events with an isolated muon are not considered. To apply muon veto during event selection, at the first step, muon candidates are selected using the recommended "Medium Muon" selection [123, 124] along with the additional requirements:

$$\begin{aligned} d_{xy}(\mu, \mathrm{PV}) &< 0.2 \text{ cm} \\ d_z(\mu, \mathrm{PV}) &< 0.5 \text{ cm} \end{aligned} \tag{5.3}$$

  Above conditions verify that the found muon in the event has originated from the primary vertex. Such a muon candidate is also required to have p_T >10 GeV and $|\eta| < 2.4$. Other than prompt muons arising from leptonic decay of heavy bosons, another source of muons are those from b-hadron decays. Muons from this latter case are supposed to have significant hadronic activity around them. As a result, to distinguish between prompt muons from leptonic decay and muons from b-hadron decays, prompt muons are required to satisfy an isolation requirement, $I_{\text{mini}} < 0.2$, where I_{mini} is the mini-isolation variable described in Ref. [125]. For mini-isolation, size of the isolation cone varies with p_T of the lepton. Cone size = 0.2 for $p_T < 10$ Gev; cone size = 10 / p_T for leptons of $p_T \in [50, 200]$ and for $p_T > 200$ GeV, cone size = 0.05.

  Any event with a muon satisfying all of the above criteria is vetoed.

- **Electron veto**: In the given final states , events with an isolated electron are not considered. To apply electron veto during event selection, at the first step, electron candidates are selected using the recommended "Cut Based VETO" selection [85]. Electron candidates are required to have p_T >10 GeV and $|\eta| < 2.5$. Similar to the muon isolation requirement, electron candidates are required to 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**: Electron and muon veto described above significantly helps to suppress the contribution of background events from $t\bar{t}$, single-top, and W+jets processes with one W$\rightarrow l\nu$ decay and having accompanying photon in the event. Even then, in some $W \rightarrow l\nu$ events, particularly when W boson decays to a τ lepton and τ further decays to electron or muon, light leptons may not be identified or may not satisfy the criteria for an isolated electron or muon candidate given above. To suppress these backgrounds, we additionally impose one or more isolated charged track veto conditions. 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 GeV,
- $I_{\text{tk}} < 0.2$,

where I_{tk} is the scalar p_T sum of other charged tracks within $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.3$ 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 GeV,
- $I_{\text{tk}} < 0.1$.

Isolated tracks are considered only if they satisfy

$$m_T(\text{tk}, p_T^{miss}) = \sqrt{2 p_T^{\text{tk}} p_T^{miss} (1 - \cos\Delta\phi)} < 100 \text{ GeV}, \tag{5.4}$$

where p_T^{tk} is the transverse momentum of the track and $\Delta\phi$ is the azimuthal separation between the track and p_T^{miss} . The m_T requirement checks that the isolated track considered under this criteria is potentially from $W \to l\nu$ process.

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.

- **Tagging of b-jets**: All the jets selected in each event are compared against DeepCSV medium working point (DeepCSVM) of b-tagging to tag them as b-jets. For 2016, 2017 and 2018, DeepCSVM had minimum threshold of 0.6321, 0.4941 and 0.4184 respectively. A jet having DeepCSV value above this threshold is tagged as a b-jet. Total number of b-jets in the event decides which signal region an event belongs to.
- **Event cleaning**: While taking proton-proton collision data, spurious p_T^{miss} signals can be originated from a mismeasured jet, detector noise, non-functional calorimetric channels, or reconstruction failures as discussed in the section 3.4.6 of chapter 3. So, there is need to apply event filters designed to reject such events in both simulations and data. The current list includes:
  - globalSuperTightHalo2016Filter
  - HBHENoiseFilter
  - HBHEIsoNoiseFilter
  - eeBadScFilter (only for data)
  - EcalDeadCellTriggerPrimitiveFilter
  - ecalBadCalibReducedFilter (only for 2017, 2018)
  - BadPFMuonFilter

- Good vertex filter (requiring at least one reconstructed vertex satisfying !isFake && $N_{\text{dof}} > 4$ && $|z| < 24$ && $\rho < 2$)
- PFCaloMET ratio filter Uses MET calculated using Particle Flow objects (PFMET) and that before Particle Flow based reconstruction using calorimetry objects (CaloMET). To protect against PF failure reject event if

$$\frac{\text{PFMET}}{\text{CaloMET}} > 5 \tag{5.5}$$

- Corrections: The "ECAL L1 pre-firing issue" led to inefficiency at trigger level for high p_T forward objects. This pre-firing inefficiency was measured in MC and applied on the data as a correction factor.

  In the 2018 data, Run 319077 onward data suffered due to HEM issue, where a small section of the minus side of the HE unexpectedly stopped working. As a result, the usual particle flow algorithm which looks for HCAL energy deposit associated with ECAL deposit to term it as a jet, misidentified additional jets and/or electrons in the disabled sector. To improve the agreement between data and MC, a specially designed wider veto was applied in this disabled region to restrict the reduction in signal efficiency to minimal. This veto is defined as:

  * Veto events with any electron with $p_T > 30$ GeV, -3.0 $< \eta <$ -1.4, and -1.57 $< \phi <$ -0.87.
  * Veto events with any jet with $p_T > 30$ GeV, $\Delta\phi$(jet,MHT) < 0.5, -3.2 $< \eta <$ -1.2 and -1.77 $< \phi <$ -0.67

5.5.2 Event yields after baseline selections

After applying above baseline selections on different backgrounds that can mimic as signals as mentioned in section 5.2, there will be reduction in number of events with each selection depending on the nature of backgrounds as shown in Figure 5.6.

Similarly, these baseline selections are applied on signal samples, but reduction in the total events especially after $p_T^{\gamma} > 100$ GeV is expected to be lesser as shown in Figure 5.7. Since, this analysis is based on at-least one photon with $p_T > 100$ GeV, it might have a impact in rejecting most of the signal events.

5.6 Optimisation studies

The events satisfying the baseline criteria as discussed in section 4.3, are further classified into mutually exclusive SRs to enhance the sensitivity of the analysis to different signal scenarios. We define two sets of SRs called the electroweak (EW) and strong production (SP) SRs to target different types of signal models. The EW SRs include events with $2 \leq N_{\text{jets}} \leq 6$ and at least one V or H-tag as shown in Figure 5.8. These SRs are sensitive to electroweakino models like

TABLE 5.6: Event yields for all backgrounds after baseline selections using full Run 2 dataset (corresponding to 137 fb^{-1}).

Selections	$t\bar{t}+\gamma$	$t\bar{t}$+jets	Z+jets+γ	Wγ	W+jets	γ+QCD jets
No cut	13598.9 ± 63.21	147012 ± 159.3	19130.5 ± 41.78	59051.8 ± 84.76	169048.22 ± 537.452	2126203.2 ± 30222.8
Lepton veto	5747.81 ± 40.98	77600.3 ± 115	19002.4 ± 41.64	26977.4 ± 57.61	116833.08 ± 447.109	2107140.7 ± 30080.6
iso e/$\mu\pi$ track veto	3787.87 ± 33.1	48957.7 ± 90.7	18153.6 ± 40.45	17774.2 ± 46.84	81379.675 ± 382.728	1953802.3 ± 28878.2
$p_T^{\gamma} > 100$ GeV	1189 ± 19	4675 ± 28	4694 ± 13	5138 ± 19	8200 ± 86	2.495e+05 ± 5e+03
$p_T^{miss} > 300$ GeV	48.01 ± 3.9	172.2 ± 5.1	505.5 ± 3.1	283.2 ± 3.8	548.1 ± 12	904.8 ± 47
$N_{jets} > 2$	47.99 ± 3.9	172.1 ± 5.1	503.3 ± 3.1	278.9 ± 3.8	546.9 ± 12	903.5 ± 47
$S_T > 300$ GeV	48.07 ± 3.9	170.5 ± 5	489.8 ± 2.9	275.8 ± 3.8	539.2 ± 11	901.4 ± 47
Trigger eff.	46.86 ± 3.8	165.7 ± 4.9	475.9 ± 2.8	267.9 ± 3.6	524.2 ± 11	875.3 ± 46
Event cleaning	46.7 ± 3.8	164.3 ± 4.9	474.7 ± 2.8	264.4 ± 3.6	520.7 ± 11	392.6 ± 21
$\Delta\phi$(2 lead jets, MET) ≥ 0.3	32.6 ± 3.1	132.7 ± 4.4	429.9 ± 2.6	212.2 ± 3.3	473.3 ± 11	54.9 ± 3.9

TABLE 5.7: Event yields for T5bbbbZg andTChiWG after baseline selections using full Run 2 dataset (corresponding to 137 fb^{-1})

	T5bbbbZg $M(\tilde{g})$ = 2200 GeV		TChiWG	
Selections	$M(\tilde{\chi})$ = 200 GeV	$M(\tilde{\chi})$ = 2100 GeV	$M(\tilde{\chi})$ = 800 GeV	$M(\tilde{\chi})$ = 1200 GeV
No cut	36.38 ± 0.15	41.11 ± 0.2	942.3 ± 1.7	82.82 ± 0.15
Lepton veto	32.17 ± 0.14	35 ± 0.19	710.5 ± 1.5	61.92 ± 0.14
Iso e/μ/π track veto	30.71 ± 0.13	33.48 ± 0.18	688.5 ± 1.5	60.53 ± 0.13
$p_T^{gamma} > 100$ GeV	14.86 ± 0.094	19.31 ± 0.14	513.3 ± 1.3	46.19 ± 0.094
$p_T^{miss} > 300$ GeV	10.45 ± 0.078	18.31 ± 0.14	375.9 ± 1.1	39.77 ± 0.078
$N_{jets} > 2$	10.45 ± 0.078	17.16 ± 0.13	255.5 ± 0.88	24.69 ± 0.078
ST > 300 GeV	10.45 ± 0.078	17.15 ± 0.13	255.2 ± 0.88	24.68 ± 0.078
After Trigger eff.	10.19 ± 0.076	16.72 ± 0.13	248.2 ± 0.86	24.02 ± 0.076
$\Delta\phi$(2 lead jets, MET) ≥ 0.3	8.653 ± 0.069	13.36 ± 0.11	179.7 ± 0.65	16.57 ± 0.069

TChiWg or TChiNg, in which an energetic W/Z or H boson is expected in addition to low N_{jets}. They are also sensitive to gluino production models like T5qqqqZg in scenarios in which the difference between the masses of the gluino and the NLSP is small, resulting in soft jets and a boosted massive vector boson. The SP SRs include any events without a V or H-tag and are sensitive to gluino and squark pair production.

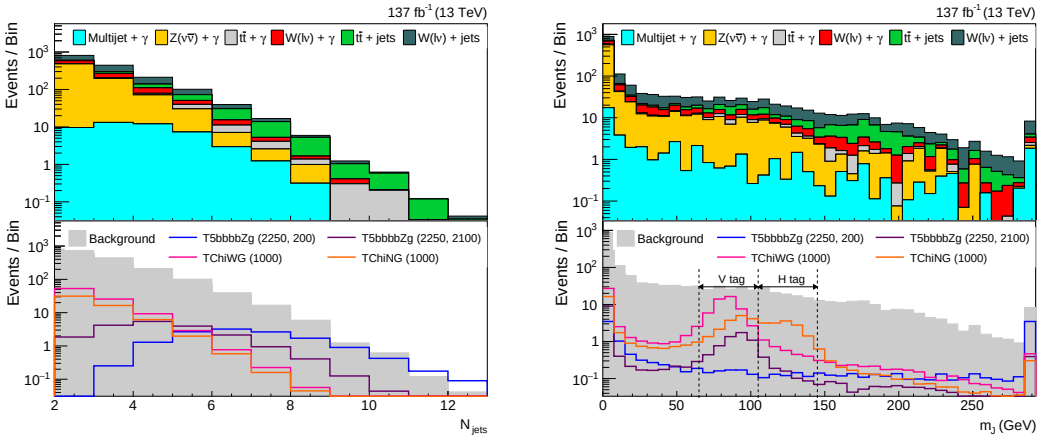

FIGURE 5.8: Simulation based SM background contributions in total number of Ak4 jets (left) and soft drop mass for the leading AK8 jet (right) is shown in the top panel. Some signal points corresponding to T5bbbbZg, TChiWG andTChiNG in bottom panel are overlaid on the top of the total background (in shaded).

Both the SP and EW SRs are binned in p_T^{miss} , with lower bin edges of 300, 370, 450, 600, 750, and $\geq$ 900 GeV as shown in Figure 5.9. For the SP SRs, each of the p_T^{miss} bins is further divided into bins of N_{jets} = 2–4, 5–6, and $\geq$ 7, and $N_{\text{b-tags}}$ = 0 and $\geq$ 1, where $N_{\text{b-tags}}$ is defined as the number of b-tagged jets in the event. For events with $N_{\text{b-tags}}$ = 0 and N_{jets} = 5–6, the two highest p_T^{miss} bins are combined. For events with $N_{\text{b-tags}}$ = 0 and N_{jets} $\geq$ 7, and for all events with $N_{\text{b-tags}}$ $\geq$ 1, the three highest p_T^{miss} bins are combined. The EW p_T^{miss} bins are defined based on the presence of a V or H-tag and the two highest p_T^{miss} bins are combined. Given the variety of signal topologies aim for, MET binning in NJets, Nbtags, V-tag and W-tag are the good choice for the sensitivity (as discussed further in this section). While optimizing the binning, we have kept in consideration to avoid any background CRs with zero event yield and merged the bins as needed.

This scheme results in a total of 27 SP and 10 EW SRs, all statistically independent. In addition, six low-p_T^{miss} CRs are defined with 200 $< p_T^{\text{miss}} <$ 300 GeV in the aforementioned bins of N_{jets} and $N_{\text{b-tags}}$, or V- and H-tags, to be used to estimate the QCD multijet background. The definitions of all the SP and EW SRs and low-p_T^{miss} CRs are summarized in Fig. 5.9. The indexing scheme shown in Fig. 5.9, with bin indices ranging from 1–45, is used to identify the SRs and low-p_T^{miss} CRs in the results presented in the sections 6.1, 6.2, 6.3 and 6.4.

Using the p_T^{miss} optimised binning in SP and EW region as shown in figure 5.9, event distribution is plotted by stacking all described backgrounds from section 5.2 which are survived after all event selections as mentioned in section 4.3.

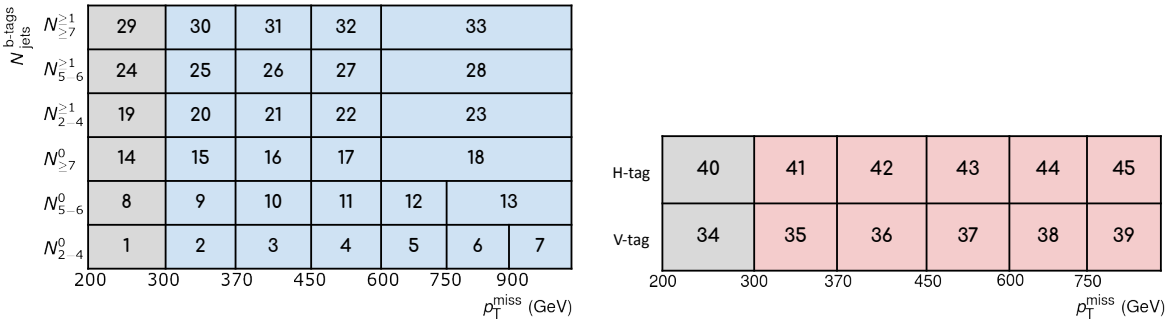

FIGURE 5.9: The definitions and indexing schemes for the SP (left) and EW (right) SRs and low-p_T^{miss} CRs, in the planes of p_T^{miss} , N_{jets} , and $N_{\text{b-tags}}$ (left) and p_T^{miss} , V-tag, and H-tag (right). The gray blocks correspond to the low-p_T^{miss} CRs, the blue blocks to the SP SRs, and the red blocks to the EW SRs.

5.6.1 Sensitivity for strong productions

To check the sensitivity, strong production signal samples like T5bbbbZg and T5qqqqHg with some particular mass points are overlaid on all stacked backgrounds as shown in the middle and bottom panel of figure 5.10. These mass points are: $M_{\tilde{g}} \gg M_{\tilde{\chi}^0_1}$ (in blue), $M_{\tilde{g}} \simeq M_{\tilde{\chi}^0_1}/2$ (in magenta) and $M_{\tilde{g}} \approx M_{\tilde{\chi}^0_1}$ (in pink), which show sensitivity in different search regions. In $M_{\tilde{g}} \gg M_{\tilde{\chi}^0_1}$ scenario, there is large jet activity so bins correspond to N_{jets} = 5–6, and ≥ 7 from SP region are more populated, while bins correspond to N_{jets} = 2–4 from SP region and V– or H–tag from EW region have lesser jet activity with boosted W/Z/H candidate, especially in the case of $M_{\tilde{g}} \approx M_{\tilde{\chi}^0_1}$.

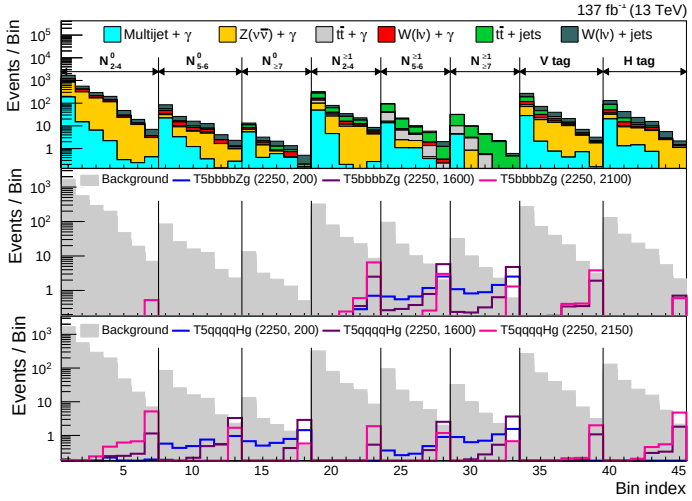

FIGURE 5.10: Simulation based SM background contributions in various search bins is shown in the top panel. Some distributions for some signal points corresponding to T5bbbbZg (in middle panel) and T5qqqqHg (in bottom panel) are overlaid on the top of the total background (in shaded). The numerical values in parentheses in the legend entries for the signal models indicate the gluino and NLSP mass values for strong production.

Separating EW region from baseline selections and dividing into V-tag and H-tag is increasing sensitivity to the final results especially when $M_{\tilde{\chi}^0_1}$ is higher as shown in Figure 5.11 (red or

green curves). By assigning MC statistical + 20% uncertainty, 95% expected limit is evaluated based on Asymptotic approximation [126] as shown in Figure 5.11. Green colour represents V-tag SR bins from EW region added to SP SR bins whereas red colour represents V- and H-tag bins from EW region added to SP SR bins. As shown in Figure 5.11 (left), for T5bbbbZg all sensitivity is coming from V-tag bins, hence exclusion limits with green and red colour are overlapping. In T5qqqqHg case, exclusion limit with red colour is showing good sensitivity as shown in Figure 5.11 (right) because $M_{\tilde{\chi}_1^0}$ is decaying into H with $\tilde{G}$ and also showing more signal events than background in V-tag region as shown in the bottom panel of figure 5.10. There is a exclusion limit for only SP region events shown as blue solid (for 27 SP SR bins) and blue dashed (for 27 SP SR + 6 SP low-p_T^{miss} CR bins) colour lines. Similarly, exclusion limits from SUS-18-002 [30] is also shown as black solid (for 35 fb^{-1})and dotted (for 137 fb^{-1}) lines.

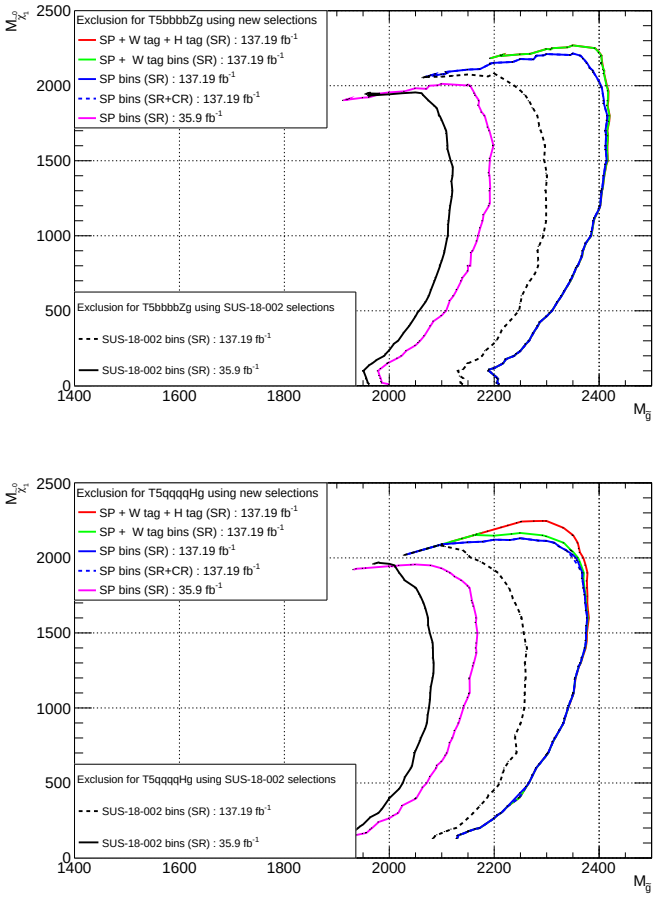

FIGURE 5.11: Simulation based optimisation checks using exclusion curves for the strong production models, T5bbbbZg (left) and T5qqqqHg (right).

5.6.2 Sensitivity for electroweak productions

To check the sensitivity, electroweakino pair production signal samples like TChiWG and TChiNG with some particular mass points are overlaid on all stacked backgrounds as shown in the middle and bottom panel of figure 5.12.

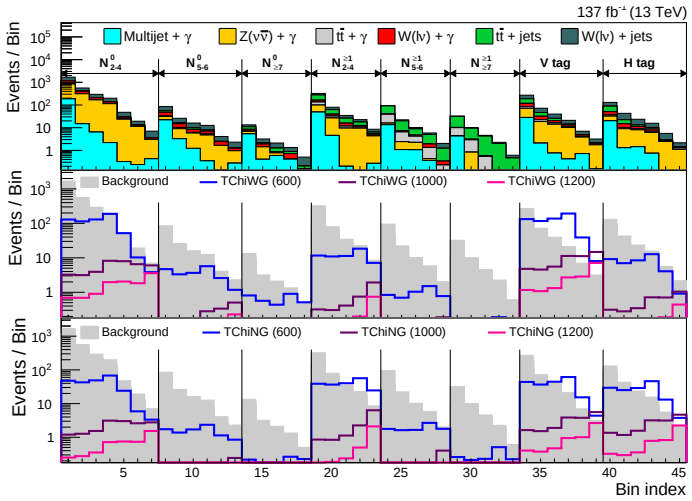

FIGURE 5.12: Simulation based SM background contributions in various search bins is shown in the top panel. Some distributions for some signal points corresponding to TChiWG (in middle panel) and TChiNG (in bottom panel) are overlaid on the top of the total background (in shaded). In middle and bottom panel, signal distributions corresponding to wino-bino electroweakino pair production are shown. The numerical values in parentheses in the legend entries for the signal models indicate the NLSP mass value for electroweak production.

We have also checked the signal acceptance x efficiency in 45 optimised bins after applying all baseline selections for all mass points as shown in Figure 5.13. In the TChiWG scenario, sensitivity is mostly coming from V-tag bins from EW region where as in TChiNG, both V- and H- tag will contribute in the sensitivity because of the presence of boosted W/Z/H candidates. For TChiNG with lower $M_{\tilde{\chi}_1^0}$ like 600 GeV (in blue), there is also a good sensitivity from N_{jets} = 2–4 when $N_{\text{b-tags}} > 0$ because H will deacy into $b\bar{b}$.

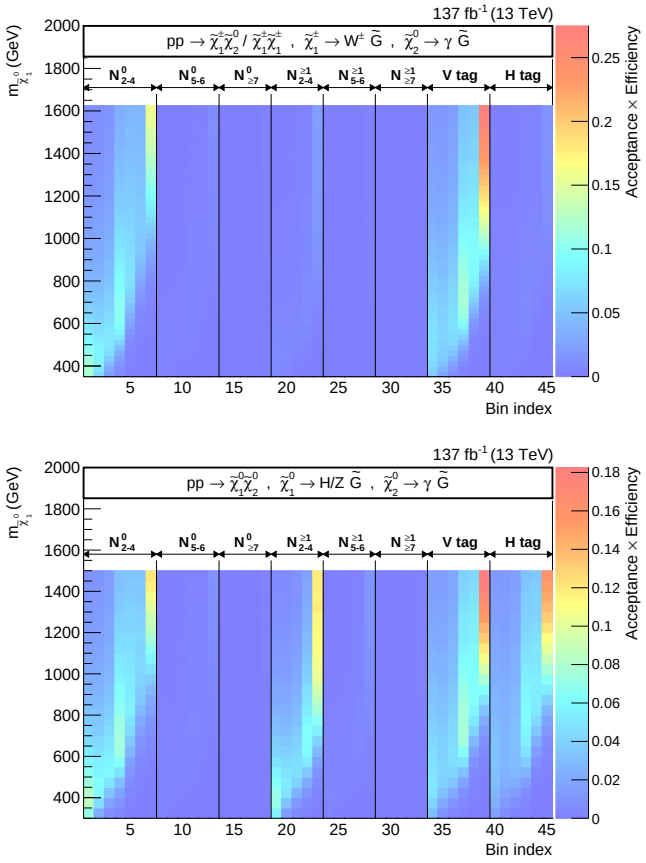

FIGURE 5.13: Simulation based signal acceptance x efficiency for the electroweakino pair production models, TChiWG (on the top) and TChiNG (on the bottom).

Chapter 6

Estimation of standard model backgrounds

There are several SM processes that can result in final states containing at least one high-p_T photon, p_T^{miss} , and multiple jets. The production of Wγ+jets, $t\bar{t}$ γ+jets, W+jets, $t\bar{t}$ +jets, Zγ+jets, γ+jets, and QCD multijet events are the most dominant backgrounds for this search as discussed in section 5.2. In this section, we have divided these backgrounds in four categories: "lost lepton background" originating from semileptonics decays of W/$t\bar{t}$ γ+jets events with the lepton not vetoed, "misidentification of electrons as photons" in W/$t\bar{t}$ +jets processes, "Z invisible" background with Zγ+jets events in which Z decays to a pair of neutrino-antineutrino, and "γ+QCD jets" background with fake p_T^{miss} in otherwise balanced events. To estimate these backgrounds, we use a data-driven approach to have accurate insight from data, especially about the detector run time conditions and physics details which could not be completely accounted for in MC simulation. We use events from data, called control regions (CRs), which are defined to be entirely orthogonal to the search regions. Dedicated CRs are used to estimate contributions from various the SM processes. These CRs are in addition to the low-p_T^{miss} CR as mentioned in the 45 optimised bins in section 5.6. The data-driven methods are assisted by the MC simulated event samples, for example, to obtain transfer factors to convert number of CR events to predicted events in search regions (SRs).

6.1 Lost lepton background

For lost lepton background, the dominant processes are $t\bar{t} + \gamma$ & $W(\ell\nu) + \gamma$ having prompt photon. These events can contribute to background in SRs since these have large p_T^{miss} from the presence of neutrinos originating from the semileptonic decays of W bosons. Although the lepton veto rejects most of such events, leptons like e or μ are considered to be "lost" when it fails reconstruction, identification, or isolation, or if it is outside the detector or kinematic acceptance. Events containing a τ_h (τ decaying hadronically) candidate that is not rejected by the isolated track veto also contribute to the background, as the τ_h candidate is reconstructed as a jet. We estimate both these contributions together as the lost-lepton background, given their similar origins. The CRs used to estimate this background, are collected by the same

p_T^{miss} triggers as used for the SR events except that we invert the lepton veto and isolated track veto for these events. There is one CR corresponding to each SR.

The SR events for lost lepton background estimation are selected using MC simulation processes and have three categories :

- **Lost electron**: In this case, there is no reconstructed electron and muon for the search region but one or two generator level electrons and no generator level muon. In the lost electron case, we do not include dilepton $t\bar{t}$ events which can have e and μ both in the same event. Reconstructed electrons can be lost due to detector acceptance, failing isolation or identification criteria, which can contribute in lost electron SR as shown in Figure 6.1 (left) and in equation 6.1. Any isolated lepton track and reconstructed leptons are vetoed in the search region. It is also necessary to veto hadronic τ events with isolated charged pion track.

$$N_{0e\gamma}^{\text{MC}} = !N^{\text{MC}}(\text{acceptance}) \oplus !N^{\text{MC}}(\text{identification}) \oplus !N^{\text{MC}}(\text{isolation}) \tag{6.1}$$

- **Lost muon**: In this case, for SR, there is no reconstructed muon and electron to have similar phase space as in baseline selection for analysis. At the simulation level, there can be at least one muon, and also including the case with one muon & one electron. Isolated muon track veto is also applied in case muon event has failed muon veto due to muon misidentification or isolation criteria is not satisfied as shown in figure 6.1 (right) and in equation 6.2. To avoid hadronic τ events contribution here, it is necessary to veto events with isolated charged π track.

- **Hadronic τ decay**: For this search region, there is the contribution of τ lepton, but it will decay hadronically. Isolated pion track veto helps to reduce hadronic τ background by vetoing events with isolated charged pi track. On top of that, there will be no electron and muon but only hadronic τ at generator level in this phase space or search region of background.

$$N_{(0\mu+\tau_h)\gamma}^{\text{MC}} = !N^{\text{MC}}(\text{acceptance}) \oplus !N^{\text{MC}}(\text{identification}) \oplus !N^{\text{MC}}(\text{isolation}) \oplus N(\text{hadronic}\tau) \tag{6.2}$$

It is also important to ensure that there is no electron-faking photon event in the SR region. The e faking γ is a different background estimation, and it will be discussed in section 6.2. To avoid overlapping between lost e with e fakes as a photon, events having ΔR between good γ and generator e < 0.2 and p_T^e / p_T^γ between 0.8 & 1.2 should be rejected in lost lepton case. All the SR categories for lost lepton with corresponding CR are briefly summarized in Table 6.1.

The lost lepton background is estimated using the yield in the single-lepton CR, which is then scaled with transfer factors (TFs) to get a zero-lepton yield in the SR. To select single-lepton CR, event selections applied are similar to SR selections except that the lepton-veto requirement is inverted. For "lost electron" case, the CR uses $(1e, 0\mu)$ requirement, while for "lost muon + hadronic τ" case, the CR uses $(0e, 1\mu)$ requirement, as shown in table 6.1. An additional

TABLE 6.1: Event selection for SR and CR regions of lost electron and lost muon + τ_h estimation.

	Lost electron		Lost μ & τ_h	
Selection	SR (0e)	CR (1e)	SR ($0\mu/\tau_h$)	CR (1μ)
No. of e (reco)	0	1	0	0
No. of μ (reco)	0	0	0	1
Isolated tracks	0	Any (e)	0	Any (μ)
p_T^γ (GeV)	> 100	> 100	> 100	> 100
$m_T(\ell, p_T^{miss})$ (GeV)	-	< 100	-	< 100
$\Delta R(reco\ e, \gamma)$	-	> 0.2	-	-
$\Delta R(gen\ e, \gamma)$ OR $p_T^{gen\ e}/p_T^\gamma$	> 0.2 OR $\notin$(0.8,1.2)	-	-	-
p_T^{miss} (GeV)	> 200	> 200	> 200	> 200
ST (GeV)	> 300	> 300	> 300	> 300
No. of jets	≥ 2	≥ 2	≥ 2	≥ 2
$\Delta\Phi(p_T^{miss}, j_1\ \&\ j_2)|_{min}$	> 0.3	> 0.3	> 0.3	> 0.3

requirement for the single-electron CR is an angle between the reconstructed electron and photon, $\Delta R(e, \gamma)$ >0.2. This condition checks that the photon in the event is prompt and not radiated due to electron bremsstrahlung.

The number of events predicted in SR bin i, $N_{0\ell\gamma}^{\text{data}}(i)$, is obtained from the number of events in the corresponding CR, $N_{1\ell\gamma}^{\text{data}}(i)$, as

$$N_{0\ell\gamma}^{\text{data}}(i) = \text{TF}(N_{\text{jets}}, N_{\text{b-tags}} \text{ or V-tag, H-tag}) \cdot N_{1\ell\gamma}^{\text{data}}(i) \tag{6.3}$$

where N_ℓ^{data} and TF are the observed yields in the CR and the simulation-derived TF for a corresponding bin after correction for the data versus MC differences.

6.1.1 Transfer factor calculation

The TFs are defined as the ratio of SR and CR events in the N_{jets} & $N_{\text{b-tags}}$ bins from SP region, V-& H-tag bins from EW regions inclusive in p_T^{miss} . In simulated events, it is easy to tag the event as a CR event or SR event and hence simulated events are used for calculating TFs. The simulated event yields are also corrected for known differences with data in lepton identification, b-jet tagging, and trigger efficiency. Separate TFs are calculated for lost electron and lost muon + hadronic τ as shown in Figure 6.1, which shows the decrement of lepton failing acceptance, identification and isolation events wrt increasing N_{jets} .

In "lost muon + hadronic τ" case, TF in some bins are >1 which are not preferable as these correspond to smaller number of CR events weighted higher to predict larger number of events in SRs. Hence final TFs are calculated after combining SR and CR of lost e, lost μ and hadronic τ events. In addition, there are no such drastic differences in TFs as calculated using simulated

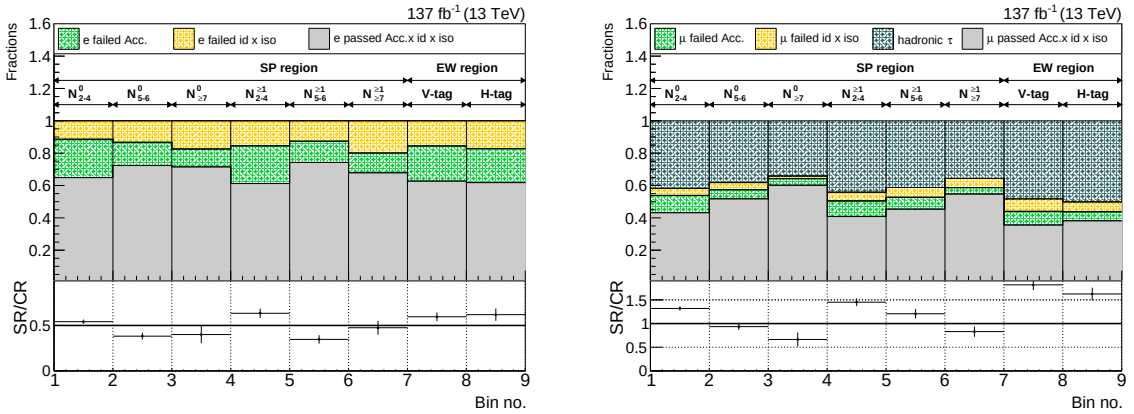

FIGURE 6.1: Contribution of lost electron (left) and lost muon (right) in SR due to failing acceptance, failing identification and isolation in N_{jets}, N_{b-tag}, V-tag & H-tag bins. On the right, hadronic τ decay events are combined with lost muon SR. In both plots, solid grey colour represents the corresponding CR region.

samples for 2016, 2017 & 2018 years as shown in Figure 6.2 (top & bottom-left). Finally, we have calculated the TFs for full Run II datasets as shown in Figure 6.2 (bottom-right). For the estimation, we use the combined TFs/CRs across all years for full Run II datasets.

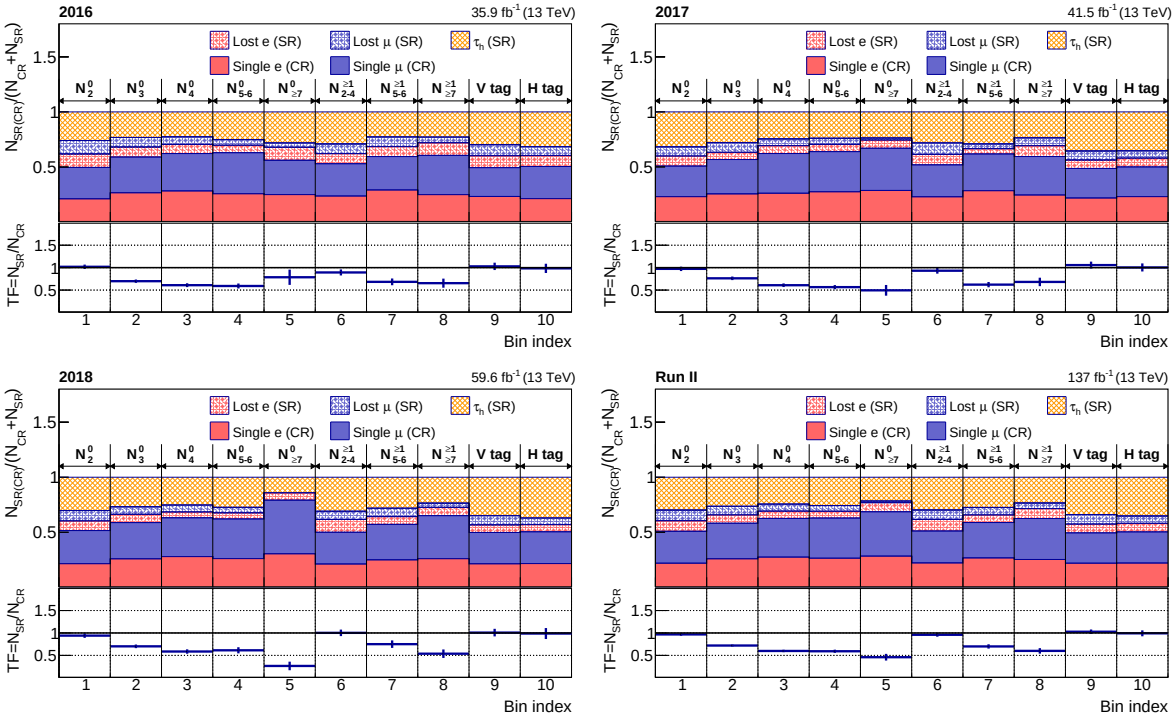

FIGURE 6.2: Relative contributions of events with light lepton(s) or τ_h candidate(s) in the SRs and CRs (upper panel), and the corresponding TFs, along with their statistical uncertainties (lower panel).

6.1.2 Data-MC correction factors

Data-MC correction factors help to correct differences between data and simulation, ensuring that the simulated events can represent the observed data to our best understanding. There are following data-MC scale factors which are applied to simulated samples in SR and CR both before calculating TFs:

- For data, the p_T^{miss} dataset is used with ≈98% trigger efficiency when p_T^{miss} >300 GeV (for details, see section 5.4.1). These trigger efficiency as a parameterization equation for $p_T^{\text{miss}} > 200$ GeV is applied on the SR and CR region simulations. The p_T^{miss} efficiency does not reach plateau as shown in Figure 5.5, hence the efficiency measure from data is used.
- Lepton scale factors are applied to events containing one lepton based on the p_T and η of the lepton. These scale factors correct for differences in the lepton identification, isolation, and reconstruction efficiencies between data and simulation. The scale factors are typically provided as a function of lepton p_T and η.
- B-tag scale factors are applied to events based on the presence and properties of b-tagged jets. These scale factors correct for differences in the b-tagging efficiencies (as discussed in section 3.4.5) between data and simulation. They are typically defined as a function of the jet p_T and η, as well as the jet flavor (b-jet in this case). The event is re-weighted using the scale factors and the corresponding MC b-tagging efficiencies, depending on the number of b-tagged jets in the event.
- Initial State Radiation (ISR) Corrections: ISR corrections are applied to the $t\bar{t}$ +jets simulated sample for the 2016 data only. These corrections account for the effects of additional radiation emitted by the initial-state partons in the hard scattering process. The corrections are typically dependent on the number of jets (N_{jets}) in the event and aim to improve the agreement between data and simulation. ISR corrections based on N_{jets} are applied on $t\bar{t}$ +jets sample but only for 2016 year. No corrections are needed for 2017 and 2018 samples (with TuneCP5 pythia tunes as described in section 5.3) since they show better ISR modelling within 1% uncertainty.

6.1.3 Closure test

We validate the background estimation method and the applicability of these TFs (6.2) for full-Run II dataset in the simulation by using the MC simulated CRs to predict the SR events using equation (6.3). Comparisons of the expected and predicted event yields in the 45 optimised bins including 27 SP SR bins and 10 EW SR bins along-with 8 low-p_T^{miss} bins are presented in Figure 6.3. Low-p_T^{miss} CR bins are also included in the final prediction because it will be used as another CR for γ + QCD multijet background estimation. There is a shaded region in the bottom panel, which correspond to the systematic uncertainty related to the lost lepton estimation (are discussed in section 6.1.5).

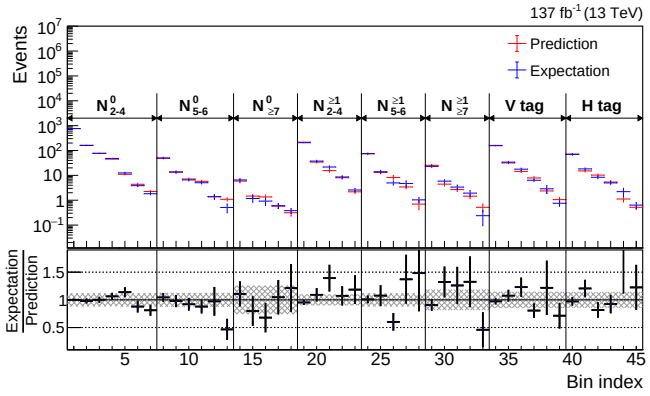

FIGURE 6.3: Validation of TFs for lost-lepton estimation in the 45 optimised bins. The TF[N_{jets}, N_{b-tag}, W-tag, H-tag] are applied on the MC event yields in CR (in red) & compared with expected MC yields in SR (in blue).

6.1.4 Data-MC comparisons in control region

In the CR for both the data and the simulations, we require events with exactly one lepton (either e or μ) and one prompt photon for lost lepton estimation. The CRs used to estimate this background are collected by the same p_T^{miss} triggers as used for the SR events (for details, see section 5.4.1). The MC simulated events for $t\bar{t} + \gamma$ & $W(\ell\nu) + \gamma$ processes, are corrected using various scale factors like lepton scale factors (in CR) and b-jet tag SFs (in both SR and CR) before calculating TFs as mentioned in section 6.1.2. By applying these scale factors, the simulated events in the CR are adjusted to match the observed data as closely as possible as shown in Figure 6.4. It is important to note that the purpose of these data-MC checks is to ensure the sanity of the analysis and not to expect a ratio of one between data and simulation. Since this method is data-driven, the goal is to account for the discrepancies between data and simulation, as the MC cross-sections of the simulated samples cannot be corrected infinitely with NLO (Next-to-Leading Order) corrections.

6.1.5 Systematics

In lost-lepton estimation, the data CR statistical uncertainty is the major source of uncertainty. It is treated as a "gamma uncertainty" while deriving statistical inference. The prediction of the lost-lepton background takes into account several sources of uncertainty that can affect the TFs. Various systematic sources are as follows :

- To account for potential mismodeling of collinear photon radiation from quarks and leptons, the cross-sections of relevant simulated samples are varied by 20% [127] and the maximum effect on the TFs is found to be 4% as shown in the bottom panel of Figure 6.5.

- Possible dependence on the choice of PDFs is evaluated by varying event weights from 100 PDF replicas [128, 129], resulting in an uncertainty of 3% as shown in Figure 6.6. To evaluate the renormalization (μ_R) and factorization (μ_F) scale uncertainties, different weights obtained by varying the scales independently by 0.5 and 2 [130, 131] are used.

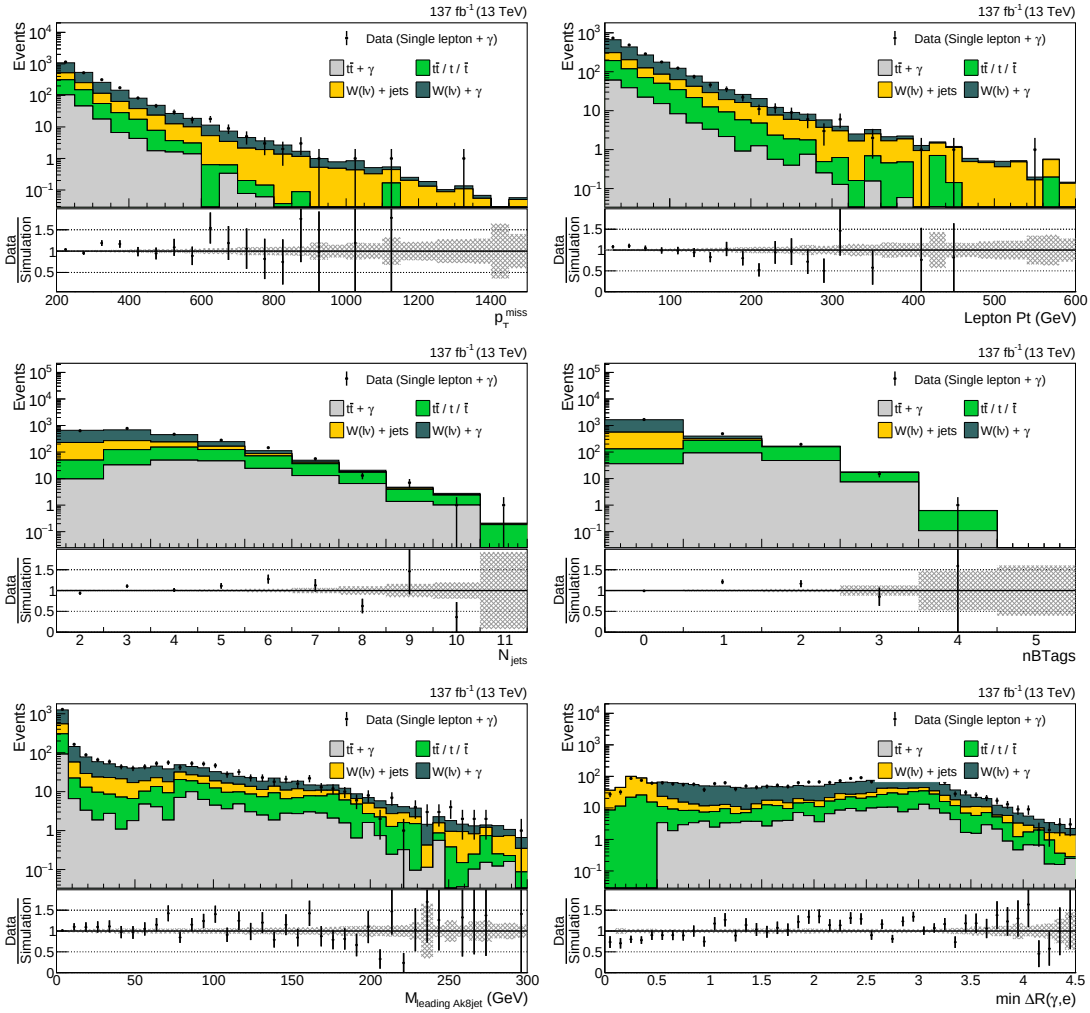

FIGURE 6.4: Data-MC comparisons in CR i.e. single lepton(e/μ) + 1γ for lost-lepton estimation. Statistical uncertainty for simulation samples are shown a shaded region in the bottom panel of each plot.

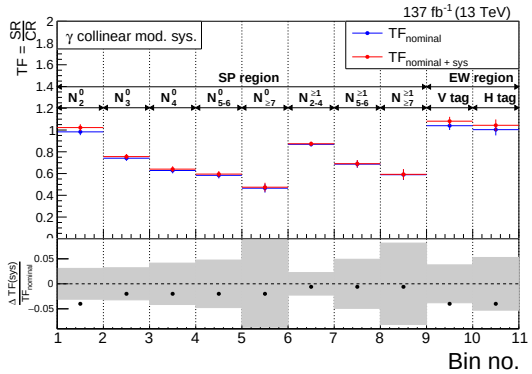

FIGURE 6.5: Comparison of nominal TF distrution (in red) and TF after varing cross-section of $t\bar{t}$ /W+ γ by 20% (in blue) in the top panel. Shaded region represents the statistical uncertainty in the bottom panel and black points are showing relativistic differences in both TFs.

The effect of this uncertainty is found to be 2% on the predicted number of lost-lepton background events.

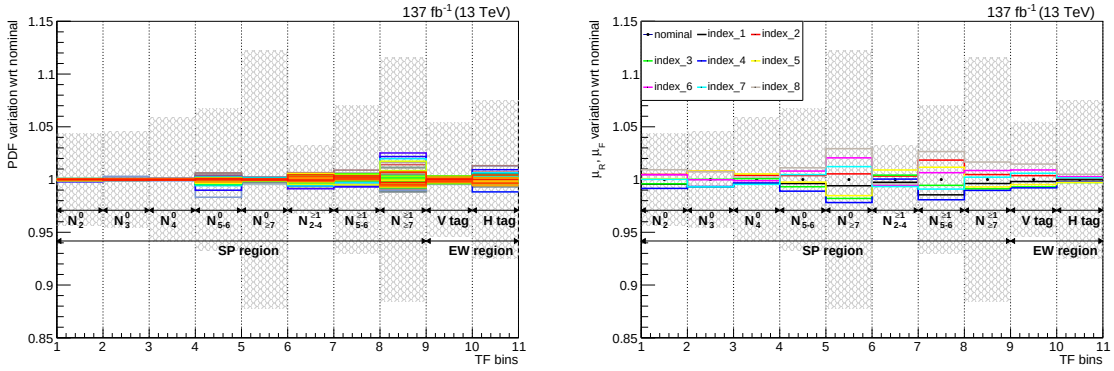

FIGURE 6.6: Envelope of 100 PDFs (left) and 8 scale (right) uncertainties with respect to the each TF bin. Shaded region represents the statistical uncertainty.

- JEC and JER are measured from data and these measurements have uncertainties whose effect on predictions should be accounted for. The systematic uncertainties in the JEC and JER lead to 6% and 6% uncertainties in the lost-lepton background, respectively as shown in Figure 6.7.

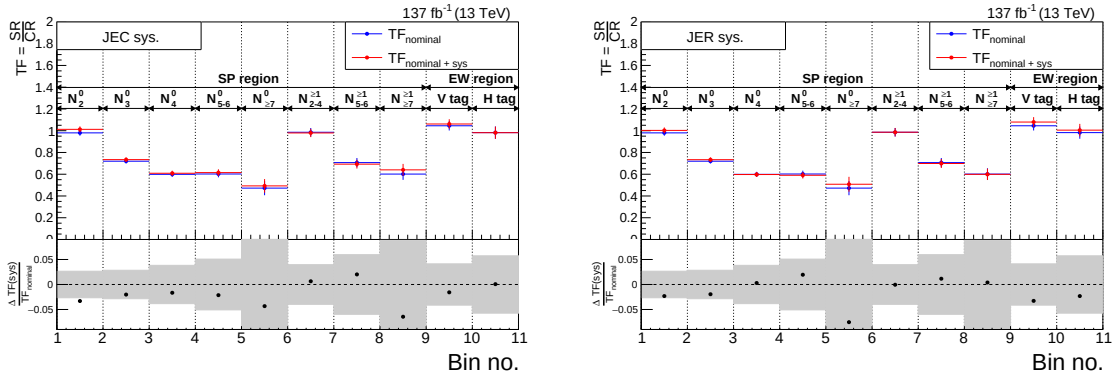

FIGURE 6.7: Comparison of nominal TF distrution (in blue) with TF after varying the JEC factors (left) and JER (right) by +1 σ_{sys} is shown in the top panel. Shaded region represents the statistical uncertainty in the bottom panel and black points are showing relativistic differences in TFs.

- The uncertainties in the lepton identification efficiency & b-tagging efficiency are 0.6 & 0.7% respectively in the lost-lepton background.

The uncertainties in the corrections applied to the simulation are propagated by varying their values and recalculating the predicted event yields in each SR. The variation in the number of predicted event yields with respect to the central values is taken as the respective uncertainty for each correction. The limited size of the simulated samples results in a 2–10% uncertainty in the predicted event yields.

6.1.6 Prediction of lost-lepton events

Final prediction of the lost-lepton background in the data SR is shown in Figure 6.33 along with total uncertainty (systematic and statistical). Table B.1 and B.2 respectively show predicted

lost-lepton background yield as a function of 45 search bins (which includes 27 SP SRs, 10 EW SRs & 8 low-p_T^{miss} CR) in high-$\Delta\phi$ SR and in low-$\Delta\phi$ CR defined above.

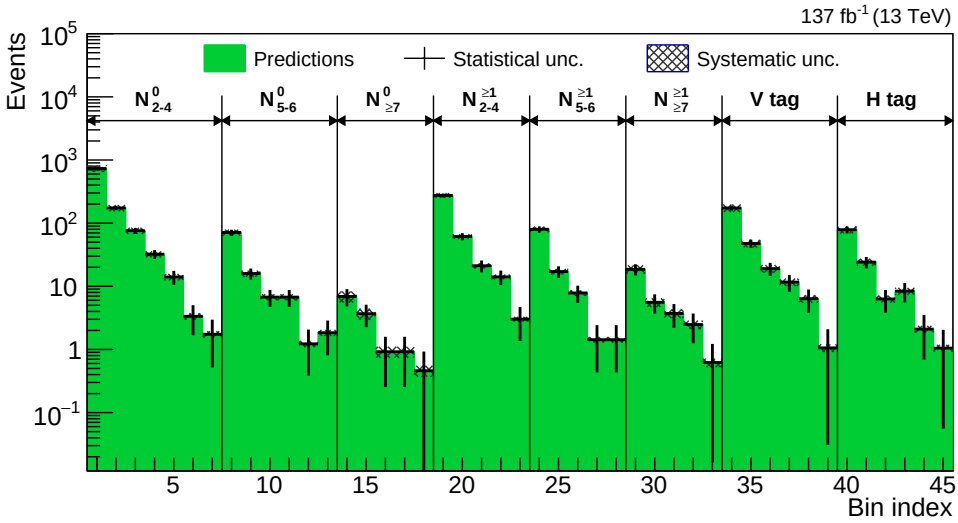

FIGURE 6.8: Final data prediction with total systematic uncertainty as a shaded region and bar line over each bin content represents the statistical uncertainty i.e. zero lepton(e/μ) + atleast-1γ for lost-lepton estimation.

6.2 Misidentification of electrons as photons

If the track associated with an electron is not reconstructed or linked to it, the energy deposited by the electron in the ECAL could potentially be misidentified as a photon. Such electrons are typically present in W+jets, $t\bar{t}$ +jets, and single top quark events in which a W boson decays to an electron and a neutrino as shown in Figure 6.9.

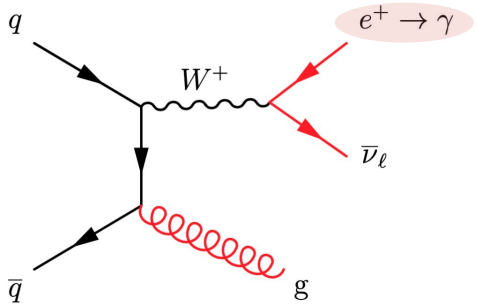

FIGURE 6.9: Feynman diagram for W+jets process in which e is misidentified as photon, p_T^{miss} is originating from ν and jets from QCD.

To estimate the background from electrons misidentified as photons, CRs are defined using events which contain exactly one electron with $p_T > 100$ GeV and no reconstructed photon. Similar to the lost lepton CRs, we require the m_T of the electron and p_T^{miss} to be < 100 GeV because it ensures that the found electron is from W decay, minimizing signal contamination in CR, as shown in Figure 6.10.

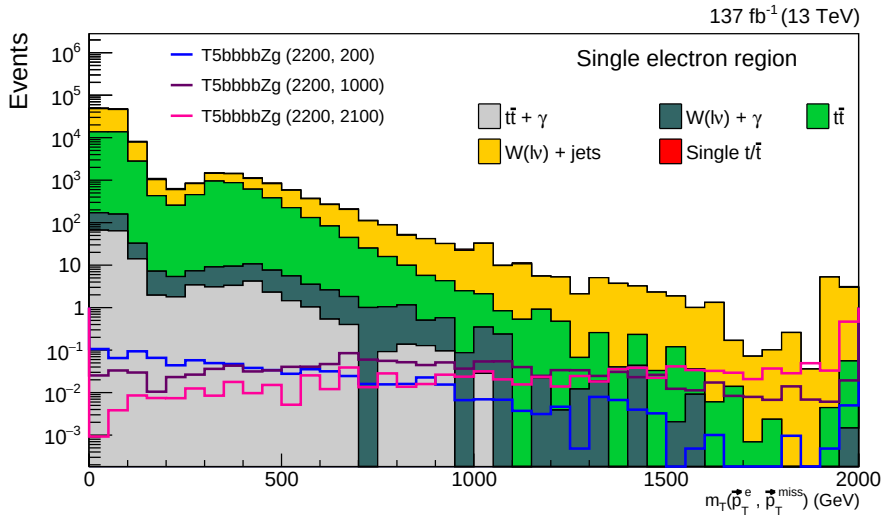

FIGURE 6.10: Distribution of the transverse mass between electron and p_T^{miss} for single electron CR in simulation. The histograms for background processes are stacked, and distributions for some signal points corresponding to the strong productions are overlaid.

In the search region, we estimate the contribution of "an electron misidentified as photon" events where events have zero reconstructed electron, at least one reconstructed photon with $p_T > 100$ GeV and without any pixel seed. In case of simulated processes, angular separation between the reconstructed photon and MADGRAPH5_AMC@NLO generator level electron

from W decay is less than 0.2 and also $p_T^{reco\ \gamma}/p_T^{gen\ e}$ should be within 0.8-1.2. All these selections are summarized in Table 6.2.

TABLE 6.2: Event selection for SR and CR regions of fake rate estimation

Selection	SR (0e)	CR (1e)
No. of e (reco)	0	1
No. of μ (reco)	0	0
Isolated tracks	0	Any (only e)
No. of γ	1	0
p_T^{γ} (GeV)	> 100	-
$m_T(\vec{l}, \vec{p}_T^{\,\text{miss}})$ (GeV)	-	< 100
$\Delta R(gen\ e, \gamma)$	< 0.2	-
p_T^{gene}/p_T^{γ}	(0.8,1.2)	-
p_T^{miss} (GeV)	> 200	> 200
ST (GeV)	> 300	> 300
No. of jets	≥ 2	≥ 2
$\Delta\Phi(p_T^{miss}, j_1\ \&\ j_2)\vert_{min}$	> 0.3	> 0.3

6.2.1 Misidentification rate calculation

The misidentification rate, f, is defined as the ratio of the number of events with a misidentified electron to the number of single reconstructed electron events. The rate f is determined using a sample of simulated W+jets and $t\bar{t}$ +jets events containing $W \to e\nu$ decays. It depends on the kinematic properties of the electron and the presence of jets and other particles near the electron. The activity around the electron is characterized using the charged multiplicity, denoted by Q_{mult} and defined as the number of charged constituents of the closest jet to the electron or photon with $\Delta R(\text{jet}, e \text{ or } \gamma) < 0.3$ as shown in Figure 6.11 (right). In addition to Q_{mult}, misidentification rate, f, is also dependent on p_T of single electron (photon) in CR (SR) as shown in Figure 6.11 (left). By choosing Q_{mult} or e/photon p_T , f captures more details of the proximity of an electron or a photon object with a nearby jet. Fundamentally, the difference between an electron object and a photon object is the track. In proximity of a jet, the chances of a track being misrecunstructed and hence not passing quality checks are higher due to a lot many hits to be fitted. This can results in dependence of electron faking as photon on the environment in electron's immediate neightbourhood.

The values of f are calculated using MC simulations of W/$t\bar{t}$ + jets processes only and these need to be corrected using a data-MC correction factor (α), which will be discussed in 6.2.3. Simulation based f is calculated in the parameterization of Q_{mult} and p_T of single electron (photon) in CR (SR) are shown in Table B.3 for 2016, 2017, 2018 and full Run II dataset. The number of events with the electrons misidentified as photons in the SR bin i, $N_{0e}^{\text{data}}(i)$, is estimated from number of events in corresponding CR, $N_{1e}^{\text{data}}(i)$, using misidentification factors f for full Run II as

$$N_{0e}^{\text{data}}(i) = f(p_T^{e/\gamma}, Q_{\text{mult}}) \cdot \alpha(N_{\text{b-tags}}) \cdot N_{1e}^{\text{data}}(i) \tag{6.4}$$

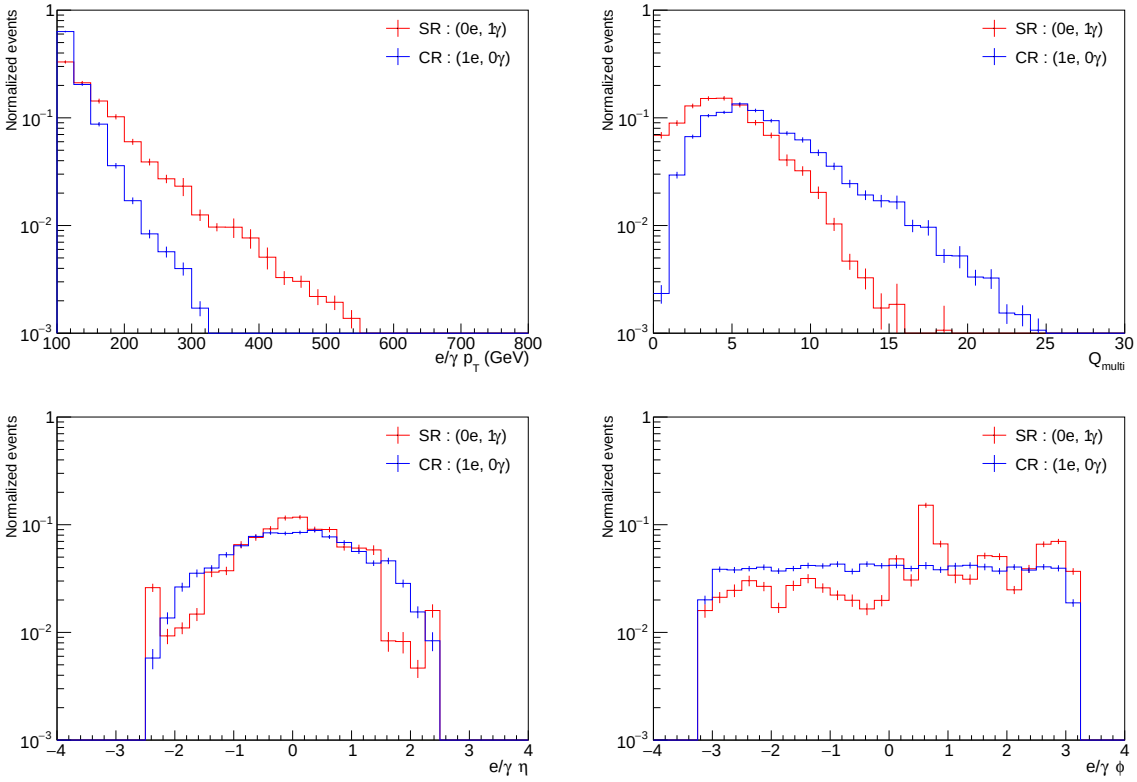

FIGURE 6.11: Kinematic distributions of the electron p_T (left) and Q_{mult} (right) for control and search region in simulation. MC statistical errors are represented by bar on top of each histogram.

For a given CR, f is applied as a per-event weight according to the p_T of the electron and the value of Q_{mult}.

6.2.2 Closure test

A comparison of the number of events expected in the SRs and predicted by the single-electron CRs, with both the SRs and CRs taken from simulated W+jets and $t\bar{t}$ +jets events, is shown in Fig. 6.12. We have not included correction factor (α) for closure test because both expected and predicted yields are from the same MC i.e. same detector conditions and hence correction factor is not needed here.

These misidentification rates are calculated in the parameterization of EM object (e/γ) p_T and Q_{mult}. Then these are applied on 45 optimised p_T^{miss} bins in the N_{jets} , $N_{\text{b-tags}}$, V- and H tag regions. Hence, exact closure of unity in each SR bins is not expected. Except for a few bins that are limited by small number of events in the CRs, the expected and the predicted numbers of events are still consistent. There is a shaded region in the bottom panel, which correspond to the systematic uncertainty related to the lost lepton estimation (will be discussed in section 6.2.5).

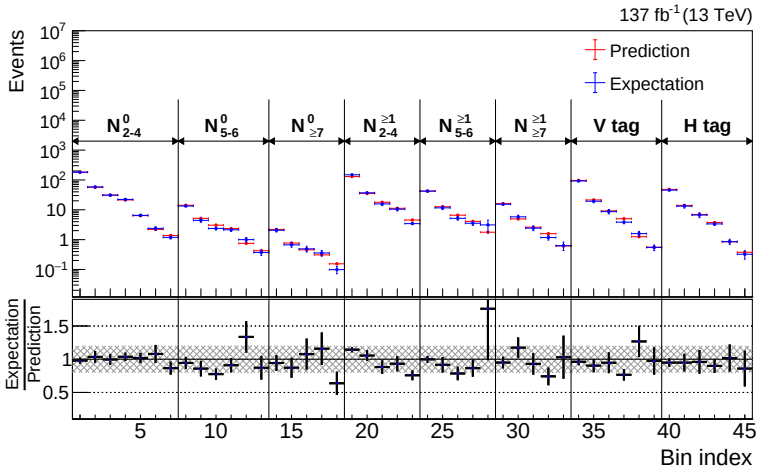

FIGURE 6.12: A comparison of the number of events with an electron misidentified as a photon in the SRs and the number estimated using the single-electron CRs with simulated samples. The ratio of the expected and predicted event yields in each SR is shown in the lower panel. The shaded region in the lower panel indicates the systematic uncertainties in the predicted number of background events.

6.2.3 Data-MC correction factor

A correction factor α is included to account for differences in f between data and simulation. The α is obtained from data and simulated events containing an e^+e^- pair with the "tag and probe" method [132]. The tag electron is required to have $p_T > 40$ GeV and to match within $\Delta R < 0.2$ with a generator-level electron arising from a Z boson in simulation or a trigger-level electron object in data. The probe electron (photon) is selected with requirements similar to the electron in the 1e (0e) CR (SR). The pair of the tag electron and the probe electron or photon is also required to satisfy $\Delta R(\text{tag}, \text{probe}) > 0.2$ and to have an invariant mass within 80–100 GeV.

Possible major backgrounds for α calculation is $t\bar{t}$ having dilepton decay. The purity factor (β) for subtracting the contribution of non-DY process (like $t\bar{t}$) from data containing e^+e^- pair, is calculated as $\beta_{e,\gamma}/\beta_{e,e}$ where $\beta_{e,e/\gamma}$ is the ratio of $e, e/\gamma$ events from DY & DY+$t\bar{t}$ processes in simulations as shown in Figure 6.13. Purity factor is 1 which is expected because any difference in the SR and CR of data will be same and hence it will be canceled out.

In this case, f is defined as the ratio of the number of events with a photon as the probe to the number of events with an electron as the probe. The factor α is determined as the ratio of f measured in data and simulated events, evaluated separately for events with $N_{\text{b-tags}}$ = 0 and $\geq$1. Other variables like N_{jets} , e/γ p_T , Q_{mult} or invariant M_Z are also checked for parameterizing the α, but any discrepancy in α' (calculated in the parameterization of other variables) are covered in the statistical fluctuations due to $\alpha(N_{\text{b-tags}})$. These checks are shown in section 6.2.3. To account for different detector conditions, α is also measured separately for data recorded in 2016, 2017, and 2018, as shown in Table 6.4. The values of weighted α which is calculated for full Run II, vary from 1.9–2.4 as shown in 6.13.

TABLE 6.3: Detailed object selection for tag and probe in data and MC. In both data & MC, invariant mass between tag and probe should be within 80–100 GeV.

	Tag (e)	Probe (e/γ)
Data	Single electron dataset	Single electron dataset
	Triggers : HLT_Ele27_WPTight_Gsf_v (2016) , HLT_Ele35_WPTight_Gsf_v (2017,2018)	Triggers : HLT_Ele27_WPTight_Gsf_v (2016) , HLT_Ele35_WPTight_Gsf_v (2017,2018)
	$p_T^e > 40$ GeV	$p^{e/\gamma_T} > 100$ GeV
	$\Delta R(e_{reco}, e_{HLT}) < 0.2$	$\star$ $p_T^{e,e}$ >100 GeV, $\Delta R(e/\gamma_{reco}, e_{HLT})$ <0.2
	Veto μ, isolated μ/π tracks	Veto μ, isolated μ/π tracks
	MET<200 GeV, ST>200 GeV & $N_{jets} \geq 2$	MET<200 GeV, ST>200 GeV & $N_{jets} \geq 2$
MC	Drell-Yan or DY process ($Z \to e^+e^-$)	Drell-Yan (DY) process ($Z \to e^+e^-$)
	$p_T^e > 40$ GeV	$p_T^{e/\gamma} > 100$ GeV
	$\Delta R(e_{reco}, e_{gen}) < 0.2$	$\Delta R(e_{reco}, e_{gen}) < 0.2$
	$\star p_T^{e,e}$ >100 GeV, $\Delta R(e/\gamma_{reco}, e_{gen})$ <0.2	$\star p_T^{e,e}$ >100 GeV, $\Delta R(e/\gamma_{reco}, e_{gen})$ <0.2
	Veto μ, isolated μ/π tracks	Veto μ, isolated μ/π tracks
	MET<200 GeV, ST>200 GeV & $N_{jets} \geq 2$	MET<200 GeV, ST>200 GeV & $N_{jets} \geq 2$

TABLE 6.4: Data-MC correction factors (α) obtained by applying T & P method on Drell-Yan events from data and MC for entire Run II

Year	$N_{\text{b-tags}}$	$f(\text{MC} : Z \to e^+e^-)$	$f(\text{data})$	α
2016	0	0.012	0.019	1.58 ± 0.064
	≥ 1	0.011	0.022	2.01 ± 0.21
2017	0	0.0093	0.026	2.79 ± 0.094
	≥ 1	0.0094	0.029	3.08 ± 0.28
2018	0	0.014	0.024	1.75 ± 0.053
	≥ 1	0.014	0.030	2.17 ± 0.16
Run 2	0	0.0119	0.0235	1.97 ± 0.039
	≥ 1	0.0117	0.0277	2.36 ± 0.117

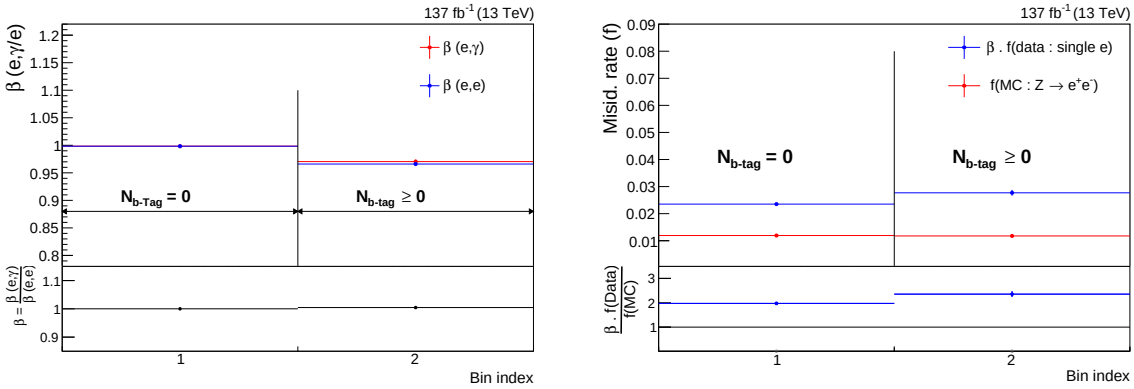

FIGURE 6.13: Purity factor distributions, $\beta(e,\gamma)$ and $\beta(e,e)$ for data SR and CR respectively in the top panel (left). Misidentification rate distribution in $N_{\text{b-tags}}$ bins for *ee* data and simulation using full Run II dataset in the top panel (right). The data-MC correction factor, α is shown in the bottom panel (right).

Additional checks for $\alpha(N_{\text{b-tags}})$

To check the validity of $\alpha(N_{\text{b-tags}})$, there is need to check the dependency of those on other variables related to electrons. There are some dependency after calculating α' in the parameterization of e/γ p_T , Q_{multi}, N_{jets} & invariant M_Z (to differentiate it with α) as shown in the left column of Figure 6.14. In the right column of same figure, f(MC) are re-weighted by $\alpha(N_{\text{b-tags}})$ then any deviation or dependency is covered by statistical uncertainties due to $\alpha(N_{\text{b-tags}})$. Systematic uncertainties related to α are explained in section 6.2.5 and those are represented by shaded region in the bottom panel of Figure 6.14 (right column), which is also covering any discrepancy. Right now, only checks for only full Run 2 are shown, and year-wise details are mentioned in appendix D. Based on these checks and considerations, it is concluded that the parameterization of α in the bins $N_{\text{b-tags}}$ is reasonable for final prediction and related uncertainties..

6.2.4 Data-MC comparisons in control region

In the CR used for the fake rate estimation, both data and simulations follow specific criteria. The CR requires the presence of only one lepton (either an electron or a muon) and no prompt photon. For data, the p_T^{miss} triggers are used and simulated samples are parameterised with trigger efficiency equation with 98% efficiency when $p_T^{\text{miss}} > 300$ GeV. Additional details about this can be found in Section 5.4.1.

The MC simulations, which mainly consist of contributions from $t\bar{t}$ + jets and $W(\ell\nu)$ + jets processes, are adjusted using b-jet tagging scale factors (SFs), JECs, and JERs in the SR. These corrections are applied before calculating the fake rates for Run 2, as explained in Section 6.1.2 of the document.

Figure 6.15 in the document shows a comparison between the event yield from data (1e, 0γ) and the expected yield from MC simulations as a function of different search variables. These data-MC checks and corrections are carried out to ensure the sanity of the analysis and to

account for any discrepancies between data and simulation in the estimation of fake rates. In the W+jets simulated samples (2017/2018), the madgraph "pdfwgt" setting is set to true, rather than false as done for 2016. This results in significantly different boson p_T spectra compared to previous campaigns. The variations in the pdfs can affect the kinematic properties of the boson (W) in the simulation, leading to differences in its p_T or in other distributions as shown in Figure 6.15 . Data-MC corrections like b-jet tagging scale factors (SFs), JECs, and JERs are applied on the MC simulation processes.

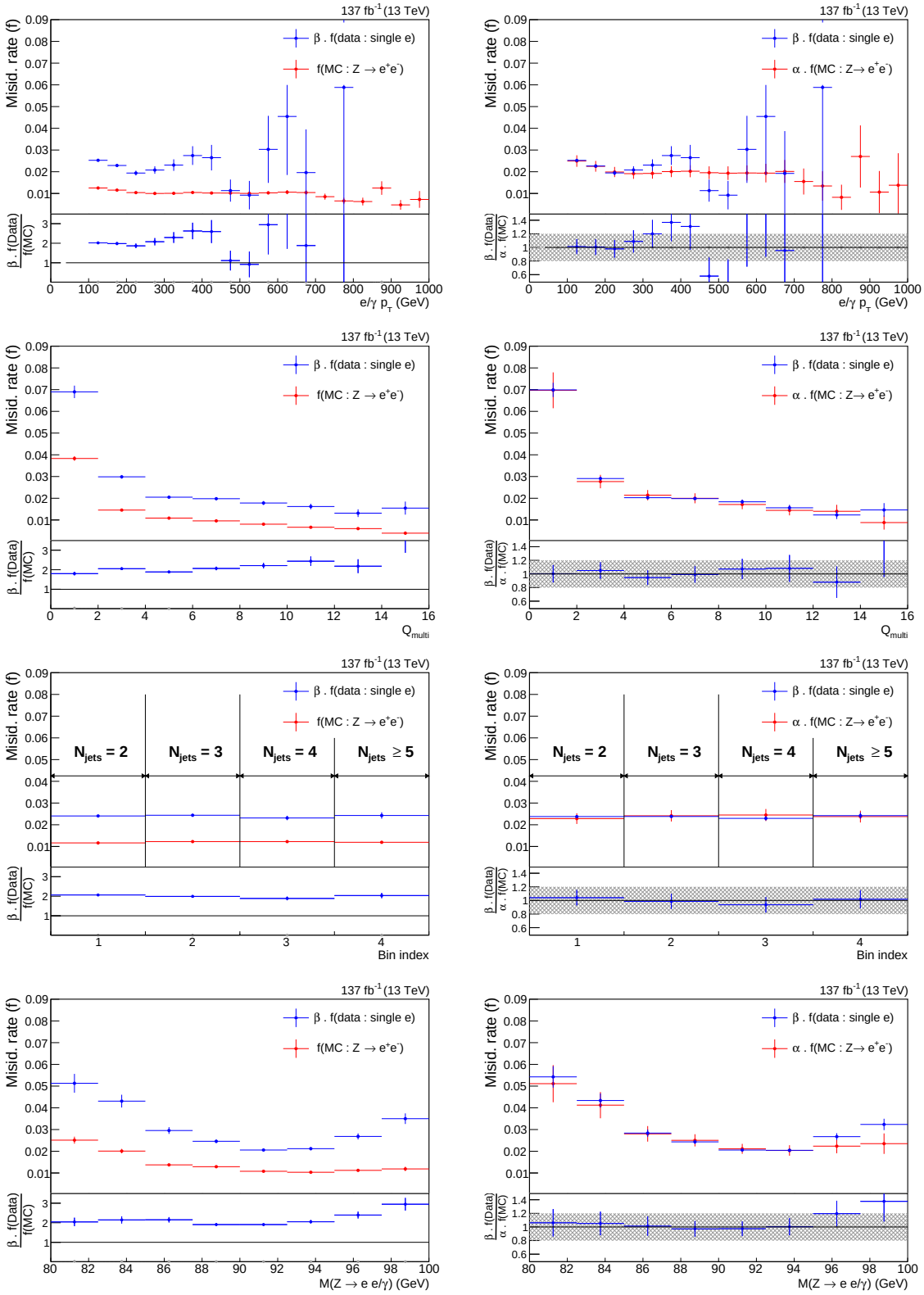

FIGURE 6.14: In the left column: f(data) and f(MC) distributions of events containing e^+e^- pairs using data and simulation respectively in the parameterization of different variables. The data-MC correction factor, α' is shown in the bottom panel of the both plots. In the right column: the ratio of f(data) and f(MC) (after re-weighted by α) are shown, and error bars are representing α statistical errors which are propagated to given bins. The shaded region in the ratio panel is the systematics related to α.

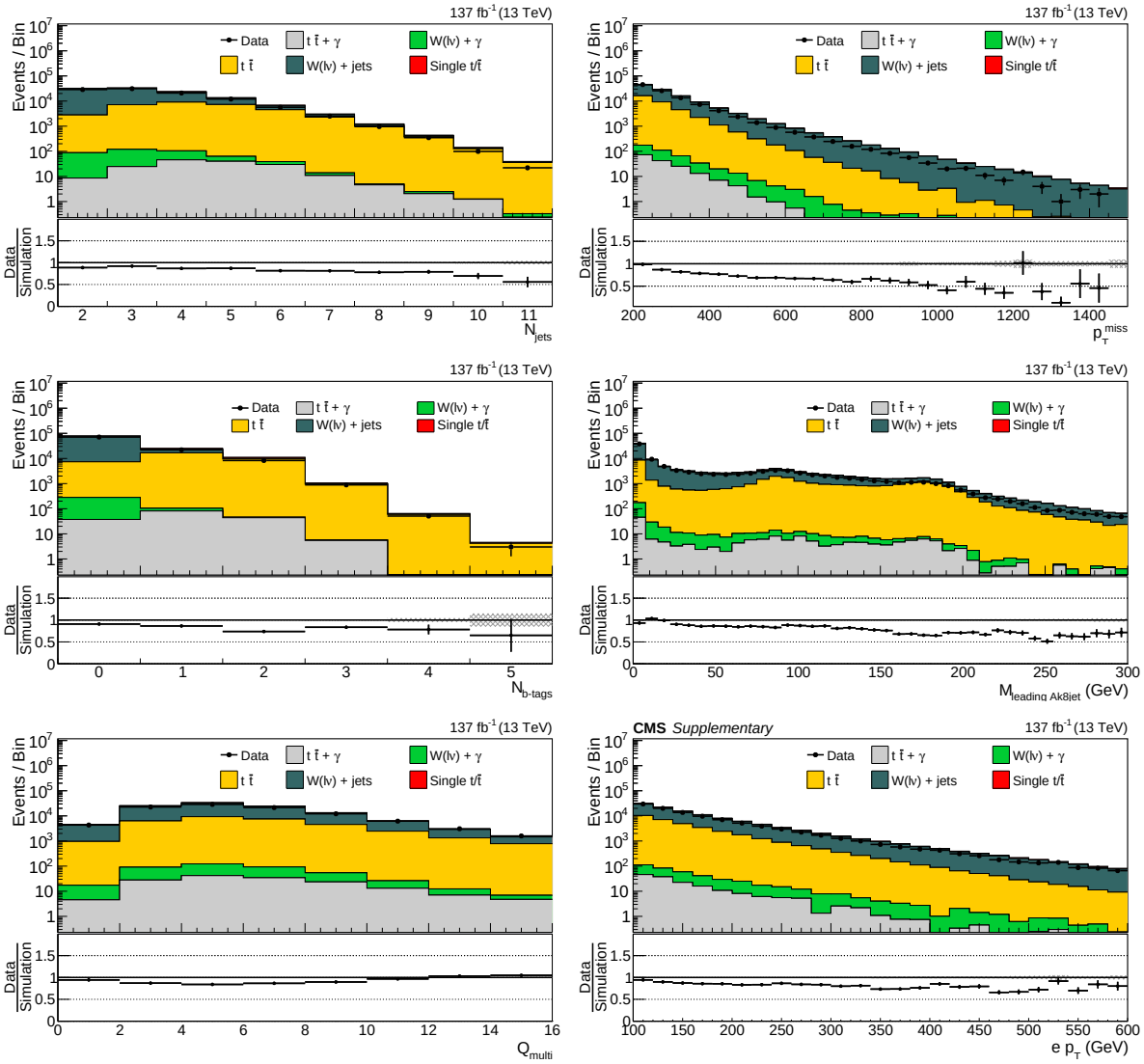

FIGURE 6.15: The comparison of $(1e, 0\gamma)$ event yield from data against expected yield from MC as a function of N_{jets} , p_T^{miss} , p_T^e and $N_{\text{b-tags}}$. Statistic uncertainty for simulation samples are shown a shaded region in the bottom panel of each plot.

6.2.5 Systematics

Similar to the lost lepton estimation as shown in the secion 6.1.5, the uncertainties in the b-jet tagging efficiencies, JECs, and JER smearing are propagated to the simulated events used to determine f, and their effects on the final predicted event yields are found to be <1, 3, and 4%, respectively as shown in Figure 6.16. The uncertainties arising from the limited number of CR and SR events in simulation samples are up to 20%.

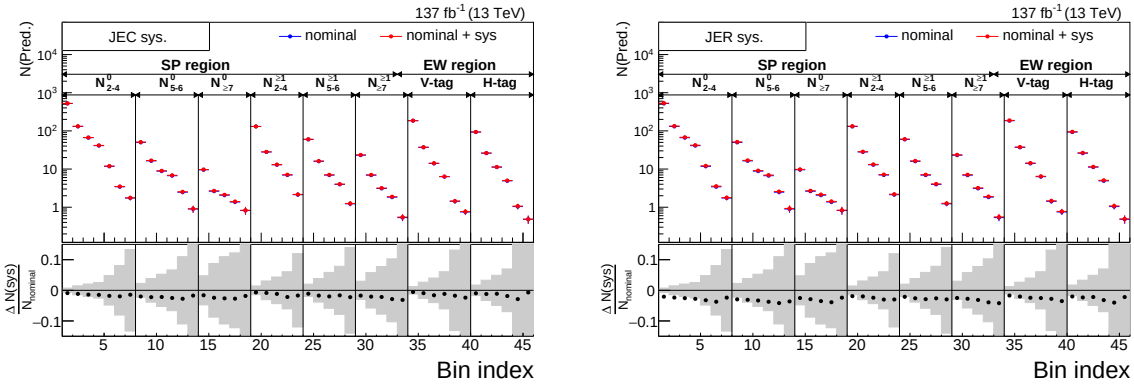

FIGURE 6.16: Comparison of nominal TF distrution (in blue) with TF after varying the JEC factors (left) and JER (right) by +1 σ_{sys} is shown in the top panel. Shaded region represents the statistical uncertainty in the bottom panel and black points are showing relativistic differences in TFs.

Systematics related to misidentification rate

We compared misidentification rates derived from DY samples ($Z \rightarrow e^+e^-$) with W/$t\bar{t}$ +jets by modifying selections for DY samples such that they are similar to W or $t\bar{t}$ processes (tag e is treated as p_T^{miss}). In DY samples, the tagging of e as p_T^{miss} is done using the magnitude of the sum of dielectron ($\vec{p_T}^{Z(ee)}$) and $\vec{p_T}^{\text{miss}}$ greater than 200 GeV, if M_{ee} is in the range of 80-100 GeV following the baseline selections. The difference between W/$t\bar{t}$ like selections and "tag and probe" selections are explained in Figure 6.17.

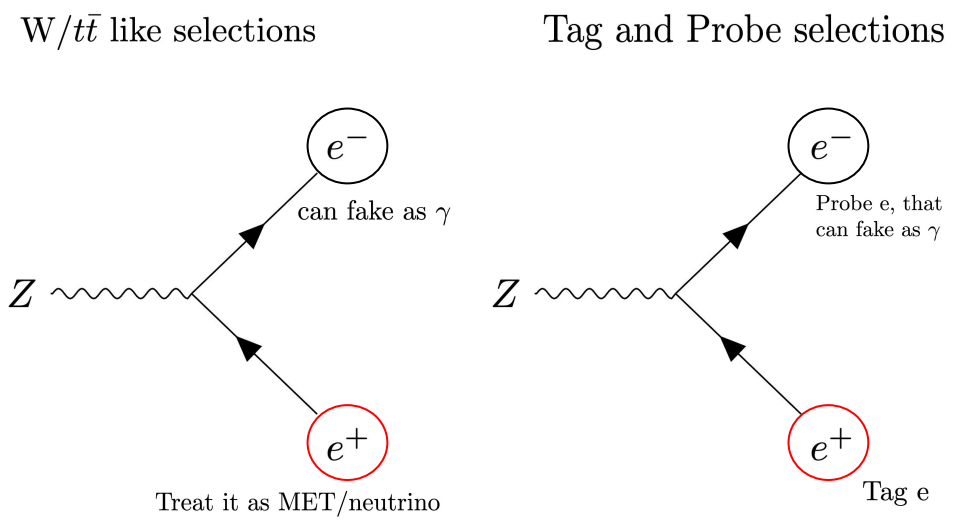

FIGURE 6.17: The comparison of W/$t\bar{t}$ like selections and tag and probe selections which are applied on $Z \to ee$ events.

The f calculated for DY & tt are matching within a variation of up to 20% in various bins as shown in Figure 6.18 (on left), so we varied $t\bar{t}$ composition by 20% conservatively and re-calculate the f(=SR/CR) using tt+W jets MC. The relative systematic uncertainty is around 2.5%(maximum) on f(MC). The f for DY and W($l\nu$) events are matching well after using W/$t\bar{t}$ like selections as shown in 6.18 (on right).

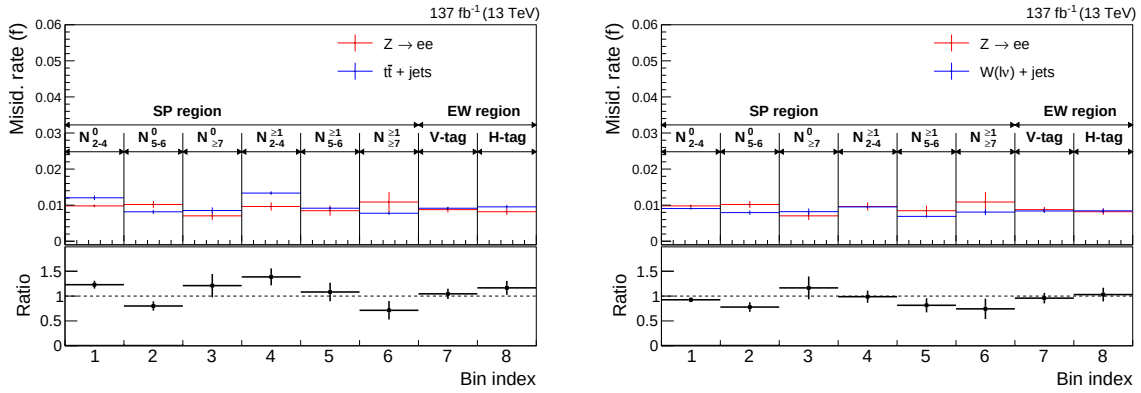

FIGURE 6.18: Comparison of f for DY ($Z \to e^+e^-$) & $t\bar{t}$ (left) in the bins of $N_{\text{b-tags}}$, N_{jets} from SP region and V- or H-tag from EW region (bins are inclusive in p_T^{miss}) after using W/$t\bar{t}$ like selections. On the right, there is a comparison of f for DY ($Z \to e^+e^-$) & W+jets (right) in the same bins.

Systematics related to data-MC correction factor

The overall systematic uncertainty in α from topological differences in $Z \to e^+e^-$ compared to $t\bar{t}$ and W+jets is estimated to be 20% (on left of Figure 6.19). This is based on comparing simulated f values in e^+e^- events for the tag-and-probe and W/$t\bar{t}$ like selections.

Figure 6.19 (on left) represents the differences in the simulated values of f α. However, due to the high statistical uncertainty in the data, we have decided to consider only the differences in the simulated values of f as shown in Figure 6.19 (on right). These differences provide insight into the variations in α due to topological discrepancies between $Z \to e^+e^-$ and $t\bar{t}$ /W+jets processes. The statistical uncertainty refers to the uncertainty arising from limited data samples especially for W/$t\bar{t}$ like selections.

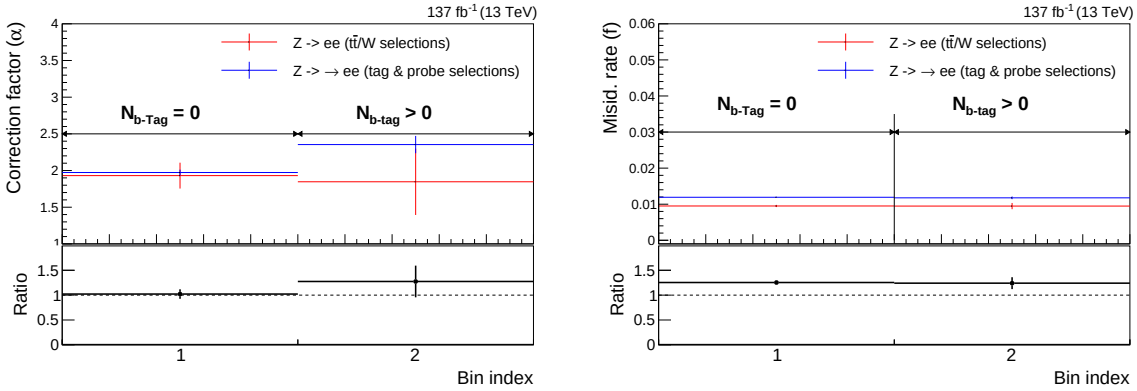

FIGURE 6.19: Comparison of α for DY (left) which is calculated using W/$t\bar{t}$ like selections and tag and probe selections in the bins of $N_{\text{b-tags}}$. There is a comparison of f for DY & W+jets (right) in the same bins.

6.2.6 Prediction of misidentified electron background events

Final prediction of the misidentified electron background in the data SR is shown in Figure 6.20 along with total uncertainty (systematic and statistical). The statistical uncertainty in the single-electron CR in data contributes up to a 20% uncertainty in the predicted number of background events in the SRs. Table B.4 and B.5 show predicted misidentified electron background yield as a function of 45 search bins (which includes 27 SP SRs, 10 EW SRs & 8 low-p_T^{miss} CR) in high-$\Delta\phi$ SR and in low-$\Delta\phi$ CR respectively, as defined above.

6.3 Estimation of $Z(\nu\nu)\gamma$ + jets background

The presence of energetic neutrinos in $Z(\nu\nu)\gamma$ + jets events manifests as large p_T^{miss} and results in a significant background for searches in final states requiring zero leptons, particularly in low-N_{jets} and high-p_T^{miss} SRs as shown in Figure 6.21 (left). We use $Z(\ell^+\ell^-)\gamma$ + jets simulated events as CR (right) to estimate the predicted event yields in the SP SRs, which are defined by N_{jets} , $N_{\text{b-tags}}$, and p_T^{miss} .

To select $Z(\ell^+\ell^-)\gamma$ + jets events, we require a pair of leptons with the same flavor, opposite charge, and invariant mass in the range of 80–100 GeV, depicted in Fig. 6.22 (left). The p_T^{miss} in these events is required to be less than 200 GeV. However, the magnitude of the vector sum of the p_T of the two leptons and the $\vec{p}_T^{\text{miss}}$, shown in Fig. 6.22 (right), must be greater than 300 GeV to mimic the kinematic properties of the $Z(\nu\nu)\gamma$ + jets process in the SRs. We also require

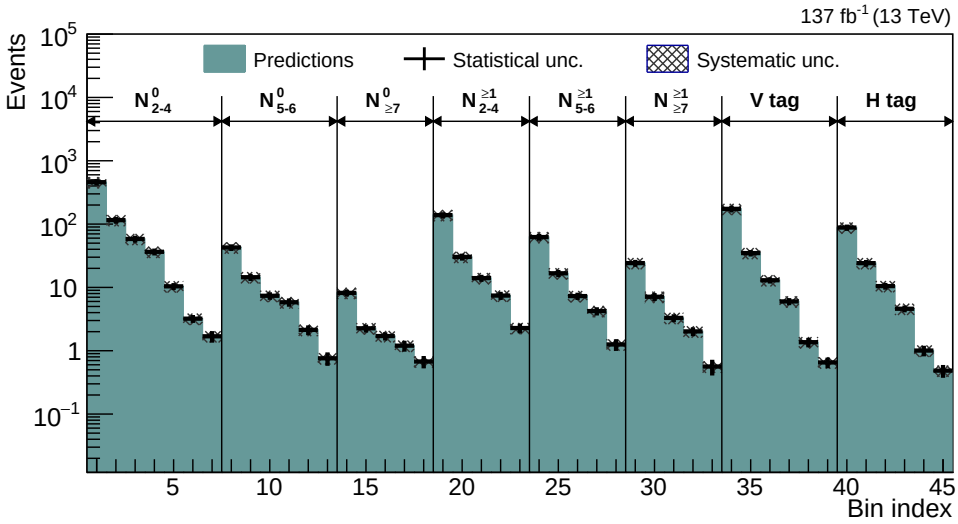

FIGURE 6.20: Final data prediction with total systematic uncertainty as a shaded region and bar line over each bin content represents the statistical uncertainty i.e. zero electron + atleast-1γ for misidentified electron estimation.

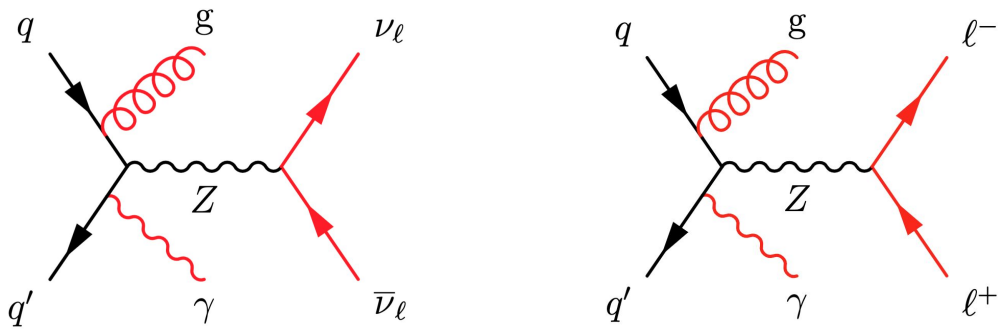

FIGURE 6.21: Feynman diagram for $Z(\nu\nu)\gamma$ + jets process (left) in which photons are coming from radiations, p_T^{miss} is coming from ν and jets from QCD. $Z(\ell^+\ell^-)\gamma$ + jets process (right) is shown in which Z is decaying in to same flavour and opposite charged lepton flavours.

$\Delta R(e,\gamma) > 0.2$ to ensure that the photon is not radiated from one of the leptons. Selection criteria for the events in signal and CR as shown in Table 6.5.

These event yields are adjusted by the ratio of the number of $Z(ee)\gamma + jets$ and $Z(\mu\mu)\gamma + jets$ events measured in data and simulation to account for any potential mismodeling of $Z(\nu\nu)\gamma$ + jets production. The leptonic final states of Z boson decays have limited numbers of events because of their small branching fractions and hence are not directly used to estimate the $Z(\nu\nu)\gamma$ + jets background. The number of events in the SRs is obtained using the expression:

$$N^{\text{data}}_{\ell\ell\gamma}(i) = \left(\frac{N^{\text{data}}_{\ell\ell\gamma} \cdot \beta_{\ell\ell\gamma}}{N^{\text{MC}}_{\ell\ell\gamma}} \right)_j \cdot N^{\text{MC}}_{\nu\nu\gamma}(i), \tag{6.5}$$

where $N^{\text{data}}_{\ell\ell\gamma}(i)$ and $N^{\text{MC}}_{\ell\ell\gamma}(i)$ are the numbers of events predicted in SR i in data and simulation, respectively, and $\beta_{\ell\ell\gamma}$ is a correction factor to account for the contribution of $t\bar{t} + \gamma$ + jets processes to the $Z(\ell^+\ell^-)\gamma$ + jets events in data, it will be discussed in section 6.3.3. The index j =

TABLE 6.5: Event selection for SR and CR regions of Z invisible estimation

Selection	SR ($Z(\nu\nu)\gamma$ + jets)	CR ($Z(\ell^+\ell^-)\gamma$ + jets)
No. of e/μ (reco)	0	2 (same flavour & opp. charge)
Isolated tracks	0	Any 2 e or 2 μ
No. of γ	1	1
p_T^{γ} (GeV)	> 100	> 100
$\Delta R(\gamma, reco\ l)$	-	> 0.2
p_T^{miss} (GeV)	> 200	< 200 & $(\vec{p}_T{}^{Z(\ell\ell)} + \vec{p}_T{}^{\text{miss}}) > 200$, if $80 < M_{\ell\ell} < 100$
ST (GeV)	> 300	> 300
No. of jets	≥ 2	≥ 2
$\Delta\Phi(p_T^{\text{miss}}, j_1 \,\&\, j_2)\|_{\text{min}}$	> 0.3	> 0.3

0 or 1 corresponds to $N_{\text{b-tags}} = 0$ or ≥ 1.

6.3.1 Data-MC comparisons in control region

The numbers of events obtained in data and simulation using these criteria are denoted $N_{\ell\ell\gamma}^{\text{data}}$ and $N_{\ell\ell\gamma}^{\text{MC}}$, respectively. Here, the subscript "$\ell\ell$" refers to either two electrons or two muons with opposite charge. In case of single electron datasets, two electrons with opposite charges are selected and single electron trigger efficiencies are taken to be 100% efficient. Similarly for selecting two muons with opposite charges , single muon datasets are used with 100% efficiency.

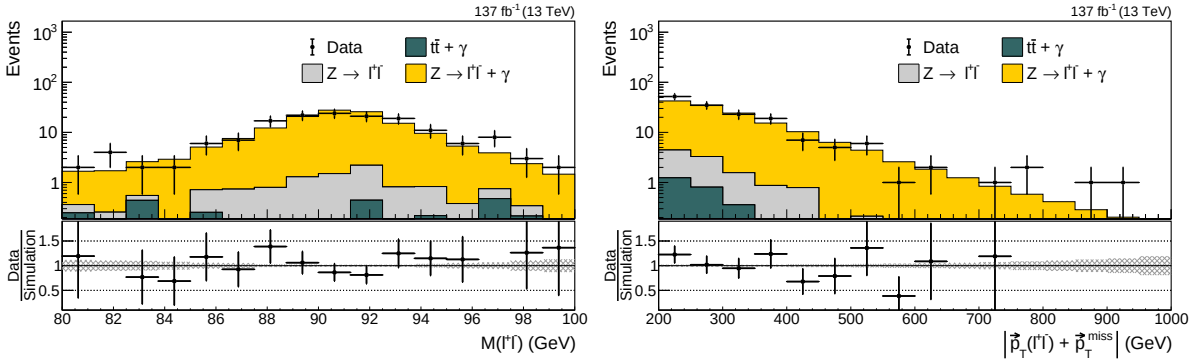

FIGURE 6.22: The distributions of the dilepton invariant mass (left) and the magnitude of the dilepton $\vec{p}_T$ plus the $\vec{p}_T^{\text{miss}}$ (right) for $Z(\ell^+\ell^-)\gamma$ + jets events in data and simulation. In the lower panel, the shaded region shows the statistical uncertainty in the simulation.

6.3.2 Validation of SR and CR events

Validity of the SR and CR region selections are shown as the ratio of normalised events of $Z(\nu\nu)\gamma$ + jets & $Z(\ell^+\ell^-)\gamma$ + jets process as shown in Figure 6.23.

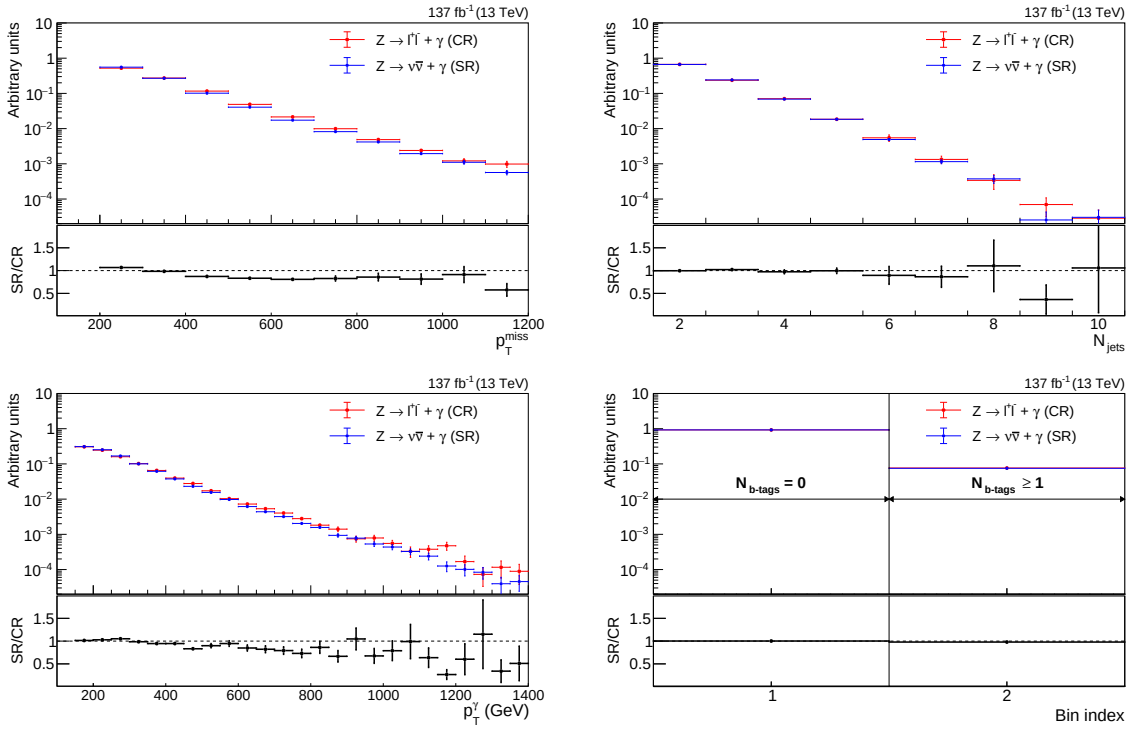

FIGURE 6.23: The distributions of the $Z(\nu\nu)\gamma$ + jets (SR) & $Z(\ell^+\ell^-)\gamma$ + jets (CR) in the bins of p_T^{miss} , N_{jets} , p_T^{γ} and $N_{\text{b-tags}}$. For top-left figure, p_T^{miss} is calculated as the magnitude of the dilepton $\vec{p}_T$ plus the $\vec{p}_T^{\text{miss}}$ (right) for $Z(\ell^+\ell^-)\gamma$ + jets events in simulation.

6.3.3 Purity factor

The dominant contribution is from $Z(\ell^+\ell^-)\gamma$ + jets (DY) events and small fraction of the events have $t\bar{t}+\gamma$ or non-DY process, especially in the higher $N_{\text{b-tags}}$ regions. To account for the differences, we derive purity factor, $\beta_{\ell\ell\gamma}$, as a function of $N_{\text{b-tags}}$. As the contamination of $t\bar{t}+\gamma$ events is expected to be higher in the SRs with $N_{\text{b-tags}}$ $\geq$1, the factor $\beta_{\ell\ell\gamma}$ is derived separately for the regions with and without a b-tagged jet.

$$\beta_{\ell\ell\gamma} = 1 - \frac{N_{t\bar{t}/W\gamma}}{N_{\ell\ell\gamma}} \tag{6.6}$$

We calculate $\beta_{\ell\ell\gamma}$ using two different methods:

- Using $t\bar{t}+\gamma$ simulated events as non-DY process in CR.
- Counting $e\mu$ events in data : In this case, it is assumed that $t\bar{t}+\gamma$ is the dominant non-DY process and the branching fraction for $t\bar{t}$ decaying to ee or $\mu\mu$ is half the $t\bar{t}$ decaying to $e\mu$ events. So, $e\mu$ events should be removed from total $\ell\ell+\gamma$ events to get the number of $Z(\ell^+\ell^-)\gamma$ + jets events.

The contribution of $t\bar{t}+\gamma$ is estimated from simulation and it is found to be statistically compatible with the number of opposite-sign, different-flavor ($e^{\pm}\mu^{\mp}$) events from data. Figure 6.24 shows the purity for 0 b-tag and >0 b-tag regions obtained from these different methods.

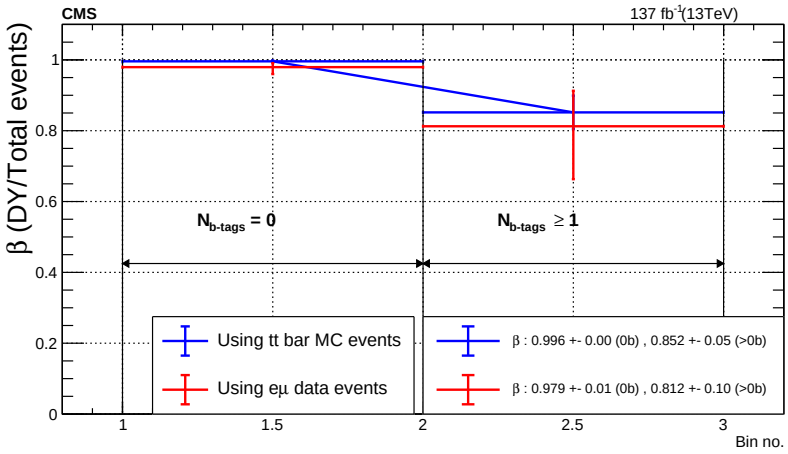

FIGURE 6.24: Purity, $\beta_{\ell\ell\gamma}$, of $\ell\ell+\gamma$ sample for 0 b-tag and >0 b-tag regions obtained using different methods.

The purity factor appears to be influenced by $N_{\text{b-tags}}$, as there will be a greater contribution from $t\bar{t}+\gamma$ and a lower purity factor for the DY process when $N_{\text{b-tags}}$ >0 as shown Figure 6.25 (in the left column). To eliminate contamination from $t\bar{t}+\gamma$ or non-DY processes, these purity factors will be used to re-weight data having the DY process. The ratio of re-weighted DY data with $Z(\ell^{+}\ell^{-})\gamma$ + jets MC events (in the parameterization of $N_{\text{b-tags}}$) will be used to scale $Z(\nu\nu)\gamma$ + jets MC events to get the predictions for Z invisible background. These scale factors are 1.07 ± 0.09 and 1.01 ± 0.28 for $N_{\text{b-tags}}$ = 0 and $\geq$1, respectively as shown in the ratio panel of Figure 6.25 (on the top-right).

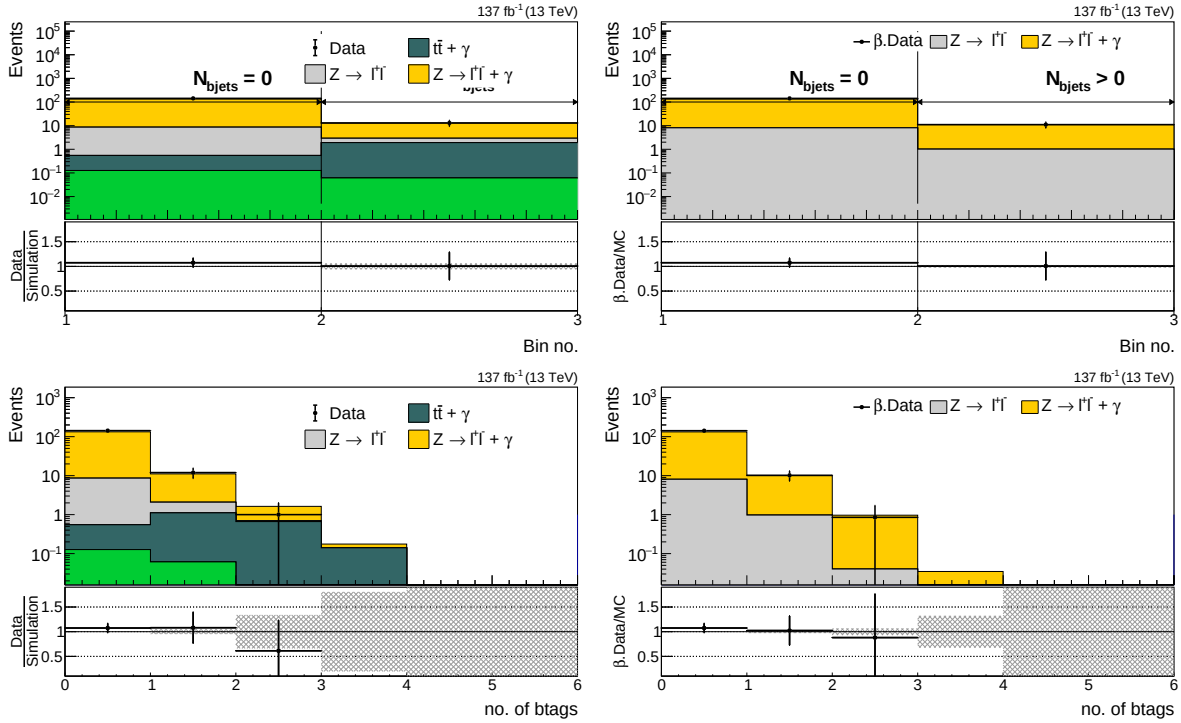

FIGURE 6.25: In left column, comparison of $N_{\text{b-tags}}$ distributions for $\ell\ell\gamma$ events in data and simulations both, and for simulations, both DY and non-DY processes are included. In right column, data distributions are reweighted by $\beta_{\ell\ell\gamma}$, which eliminate the non-DY contribution and hence, in simulations only DY process is shown.

6.3.4 Systematics

For Z invisible estimation, the data CR statistical uncertainty is the major source of uncertainty in the number of predicted background events. Besides, various systematic sources of uncertainty also contribute to the scale factors calculations.

The major systematic sources are as follows :

- Scale factor uncertainty: These uncertainties are propagated to the predicted number of events in the SRs along with the statistical uncertainties from the $Z(\nu\nu)\gamma+$jets simulated samples, which range from 2–70%.
- To account for any mismodeling of photon p_T in the simulation, we apply an additional systematic uncertainty, which ranges 18–40%, depending on the p_T^{miss} bin. This is assessed based on the results presented in Ref. [133].
- Jet Energy correction (JEC) & resolution uncertainty

6.3.5 Prediction of Z invisible background events

Final prediction of the Zinvisible background in the data 45 search region which includes 27 SP SRs, 10 EW SRs & 8 low-p_T^{miss} CR, as shown in Figure 6.26 along with total uncertainty (systematic and statistical). Table B.6 and B.7 respectively show predicted misidentified electron

background yield as a function of 45 search bins in high-$\Delta\phi$ SR and in low-$\Delta\phi$ CR. These predictions are calculated using $Z(\nu\nu)\gamma$ + jets simulated samples after apply scale-factors (which includes data-MC corrections) and in low-$\Delta\phi$ CR, same scale factors are used due to the lack of data-statistics.

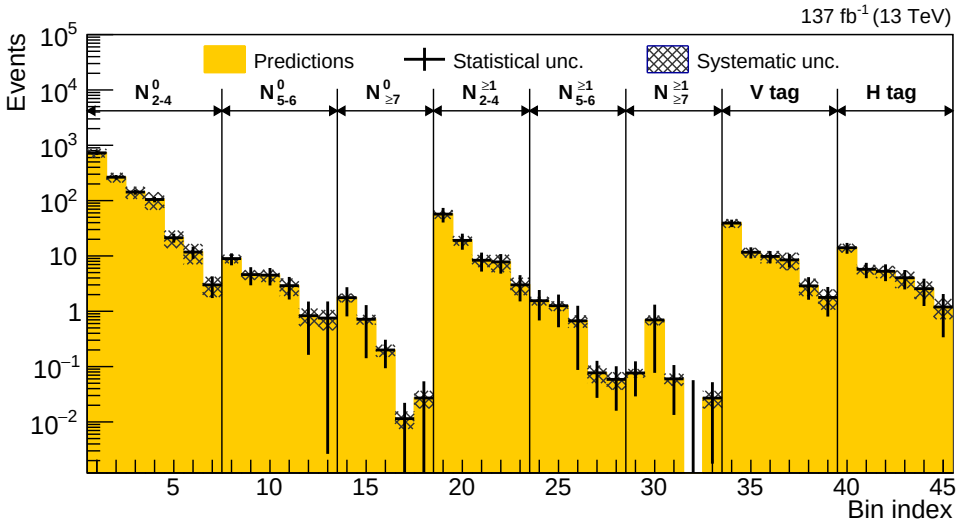

FIGURE 6.26: Final data prediction with total systematic uncertainty as a shaded region and bar line over each bin content represents the statistical uncertainty i.e. zero electron + atleast-1γ for misidentified electron estimation.

6.4 Estimation of γ + *jets* and QCD multijet backgrounds

QCD multijet and γ + *jets* events with a well-reconstructed photon can contribute to the SRs if they also contain large p_T^{miss} resulting from significant mismeasurement of one or more jets or from the presence of neutrinos from the semileptonic decays of heavy-flavor hadrons, as shown in Figure 6.27 (left). These processes can act as a background to this search if one of the following cases is valid:

- One or more jets are mismeasured at the reconstruction level giving rise to fake p_T^{miss} .
- Either of the jets has originated from the hadronization of b-quark, and B-hadron formed after hadronization has decayed semileptonically, giving rise to true p_T^{miss} arising from accompanying neutrino or antineutrino.
- In the QCD process, along with one or more mismeasured jets, an energetic π^0 is misreconstructed as a photon if either of the photons from π^0 accidentally passes the isolation requirement for a photon.

6.4.1 Methodology

We use an "ABCD" method to estimate this background, where the regions A, B, and C are data CRs designed to be non-overlapping with the SRs and the region D corresponds to the SRs, as shown in Figure 6.27 (right). To define these CRs, we use events with 200 $< p_T^{\text{miss}} <$300

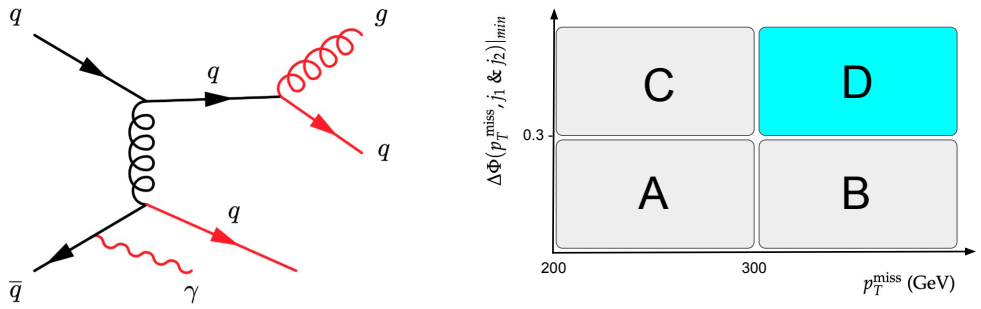

FIGURE 6.27: Feynman diagram for QCD multijet and $\gamma + jets$ process (left) in which photons are coming from initial state radiations and jets from QCD. A schematic 2D diagram of ABCD method (right) is shown with p_T^{miss} in x-axis and $\Delta\Phi(p_T^{miss}, j_1 \,\&\, j_2)|_{min}$ in y-axis.

GeV (low-p_T^{miss}) or $\Delta\Phi(p_T^{miss}, j_1 \,\&\, j_2)|_{min} < 0.3$, indicating at least one of the two leading jets is aligned with the direction of p_T^{miss} (low-$\Delta\phi$). These events are required to satisfy all other baseline selection criteria.

The low-p_T^{miss} CRs (C & A) are used to calculate $R^{data}_{low-p_T^{miss}}$(R-factor) which is discussed in the section 6.4.2, and CR B consists of events with high-p_T^{miss} and low-$\Delta\phi$. These events are divided into the same regions as the SRs, including the p_T^{miss} binning. The predicted number of events in SR i is then:

$$N^{data}_{multijet+\gamma}(i) = R^{data}_{low-p_T^{miss}}(r) \cdot \kappa(r) \cdot N^{data}_{B}(i), \tag{6.7}$$

where the index i refers to the search region bins and r refers to the bins defined as N_{jets} , $N_{b\text{-}tags}$, V-tag and H-tag which are inclusive in p_T^{miss} , as used for the SRs. The factor $\kappa(r)$ is used as a correction determined from the simulation to account for any differences in the ratios of the low-p_T^{miss} regions A and C compared to the high-p_T^{miss} regions B and D. This κ factor is explained in section 6.4.3.

6.4.2 R factor using data

At first, the ratio of the number of QCD plus $\gamma + jets$ events in high-$\Delta\phi$ to low-$\Delta\phi$ (C/A) is obtained as a function of [N_{jets} ,$N_{b\text{-}tags}$] from SP region, and V- and H-tag from EW region using low p_T^{miss} sideband having $p_T^{miss} \in$ [200,300] GeV. This ratio is defined as R-factor ($R^{data}_{low-p_T^{miss}}$), as shown in equation 6.8:

$$R^{data}_{low-p_T^{miss}}(r) = \frac{N^{data}_{high-\Delta\phi}(r)}{N^{data}_{low-\Delta\phi}(r)} \tag{6.8}$$

For calculating R-factors, p_T^{miss} data is used for Run II and QCD multijets plus $\gamma + jets$ simulated samples are re-weighted by p_T^{miss} trigger efficiencies as mentioned in section 5.4.2. As CR (B) and low-p_T^{miss} sideband regions (A & C) are SM background dominated and not contaminated by SUSY signal, the number of QCD and $\gamma + jets$ events in these regions are extracted directly by subtracting the estimated contribution of lost-lepton, fake-photon and Z-invisible backgrounds from total event yield in those regions to get correct values of R-factor (C/A) and data CR yield (B). Estimating the mentioned backgrounds (explained in following sections:

6.1, 6.2 & 6.3) in the low-$\Delta\phi$ region follows an exactly similar strategy as used to estimate SR background contribution, except that corresponding CR, TFs and SFs are derived from low-$\Delta\phi$ regions. The predictions of the electroweak SM background in the low-$\Delta\phi$ region are shown in tables B.2, B.5 & B.7.

The values of $R^{\text{data}}_{\text{low}-p_T^{\text{miss}}}$ are calculated after subtracting the contribution of electroweak SM background from corresponding CRs (A, B & C), as shown in the ratio panel of Figure 6.28.

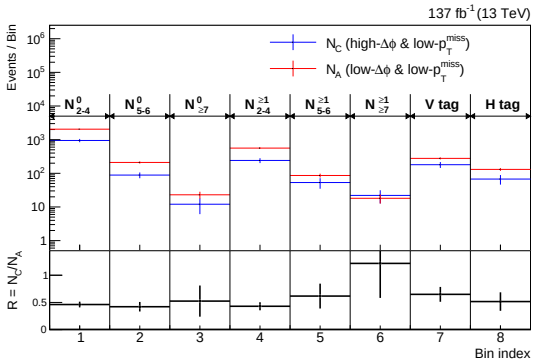

FIGURE 6.28: The event yield in MET sideband regions (A,C) (top panel) and corresponding R-factor distribution (bottom panel) as a function of [N_{jets} , $N_{\text{b-tags}}$, V-tag, H-tag] using data.

6.4.3 κ factor using simulations

κ-factor is an important ingredient for this background estimation and it is obtained from the simulation. To correct any correlation between $\Delta\phi$ & p_T^{miss} , similar ratios $R^{\text{MC}}_{\text{low}-p_T^{\text{miss}}}$ and $R^{\text{MC}}_{\text{high}-p_T^{\text{miss}}}$ are calculated using MC and with requirement $p_T^{\text{miss}} < 300$ GeV and $p_T^{\text{miss}} > 300$ GeV respectively, as shown in equation 6.9. In the given equation, κ-factors are calculated in the bins of N_{jets} and $N_{\text{b-tags}}$ from SP region, and V-tag or H-tag from EW region.

$$\kappa(r) = \left(\frac{N^{\text{MC}}_{\text{high}-\Delta\phi}(r)}{N^{\text{MC}}_{\text{low}-\Delta\phi}(r)}\right)_{\text{high}-p_T^{\text{miss}}} \Big/ \left(\frac{N^{\text{MC}}_{\text{high}-\Delta\phi}(r)}{N^{\text{MC}}_{\text{low}-\Delta\phi}(r)}\right)_{\text{low}-p_T^{\text{miss}}} \tag{6.9}$$

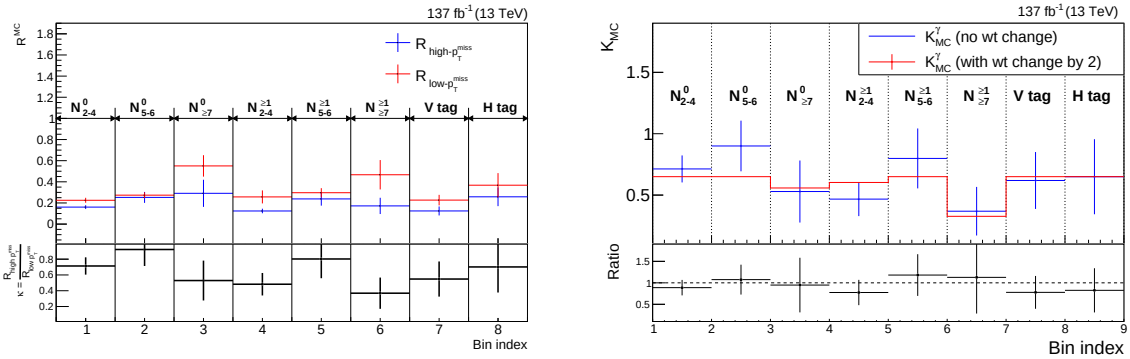

FIGURE 6.29: Left: R-factors calculated using simulation in low-p_T^{miss} (red) and high-p_T^{miss} (blue) region. The ratio panel shows the κ-factor in [N_{jets} , $N_{\text{b-tags}}$, V-tag, H-tag] bins. Right: κ(MC) re-scaling γ + *jets* contributions (in red) and without rescaling (in blue)

Figure 6.29 (left) shows how event yield and corresponding R-factor varies in low-MET sideband and high-p_T^{miss} region as a function of [N_{jets} , $N_{\text{b-tags}}$, V-tag, H-tag] as derived from MC. The κ-factor is applied as a correction factor to $R^{\text{data}}_{\text{low}-p_T^{\text{miss}}}$. To test the stability of MC ratio of κ-factors by varying fractions of $\gamma + jets$ by the factor of 2. We found that any deviation in the κ(MC) is covered by its statistical uncertainty as shown in Figure 6.29 (right). In addition, we also validate these factors in data and simulation samples using events that do not contain a photon, referred to as zero-photon events as discussed in section 6.4.4.

6.4.4 Systematics

For $\gamma + jets$ & QCD multijet estimation, the data CR statistical uncertainty is the major source of uncertainty around 2–100%. The statistical uncertainties in κ and the number of events in low-p_T^{miss} regions vary in the range 10–50%. Besides, κ-factor validation in data and MC contribute to the systematic uncertainty.

κ-factor validation systematic

We validate κ-factor in data and simulation samples using events that do not contain a photon, referred to as zero-photon events. For these validation region, we are using Run II p_T^{miss} datasets and p_T^{miss} trigger efficiencies in low-$\Delta\phi$ & high-$\Delta\phi$ region as mentioned in the section 5.4. We have applied similar baseline selections for atleast-one photon events as mentioned in section 5.5.1 on both data and simulations, but here we used proxy photon instead of real photon. Proxy photon is a leading jet with highest EM energy fraction and its data-MC comparison is shown in Figure 6.30 on top-right.

The zero-photon region is dominated by QCD multijet events, and κ factors in zero-photon region is calculated using ABCD method as mentioned in Figure 6.27 (right). The contributions from the other SM background processes like $t\bar{t}$ +jets, W+jets, $Z(\nu\nu)\gamma +$ jets are small compared to QCD multijets events, so they are estimated from the simulation and subtracted from the zero-photon event yields. For this, we have made data-MC comparisons which clearly shows the large contribution from QCD multijet events as compared to the other SM background processes is shown in Figure 6.30.

In zero-photon validation, R-factor are calculated in low-MET sideband and high-p_T^{miss} region as a function of [N_{jets} , $N_{\text{b-tags}}$, V-tag, H-tag] (similar to the section 6.4.3) but using both data and simulation. Figure 6.31 (on left) shows the differences in the κ calculated from data and simulation and the deviation from unity in the ratio panel gives the estimation of systematics related to k factor calculation. Additionally, we also checked if there is any differences or trend in κ(MC) for zero-photon validation region and at least one photon region as shown in Figure 6.31 (on right). Any differences in κ(MC) are under statistical uncertainty for $\kappa^{1\gamma}_{MC}$.

6.4.5 Prediction of $\gamma + jets$ and QCD multijet background events

Final prediction of the $\gamma + jets$ and QCD multijet background in the data 45 search region which includes 27 SP SRs, 10 EW SRs & 8 low-p_T^{miss} CR, as shown in Figure 6.32 along with

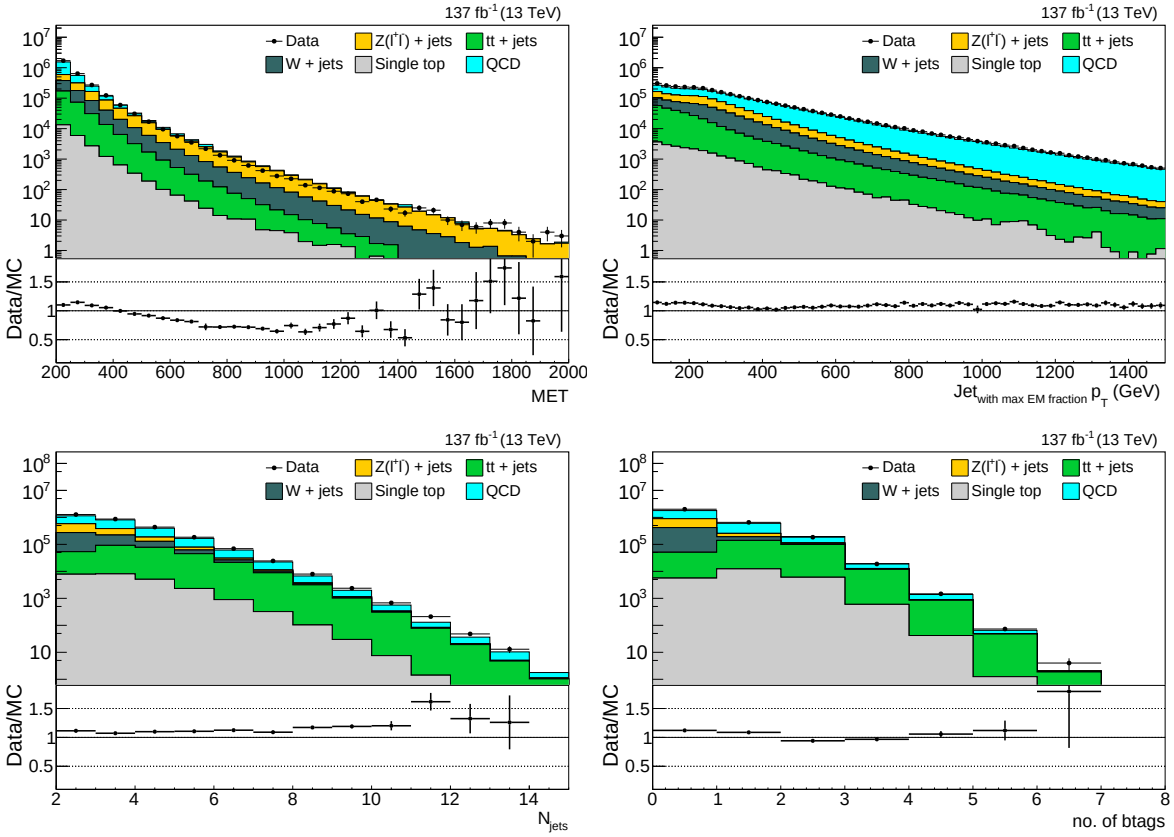

FIGURE 6.30: Data-MC comparisons of p_T^{miss} , proxy γ p_T , N_{jets} & $N_{\text{b-tags}}$ in zero photon validation region.

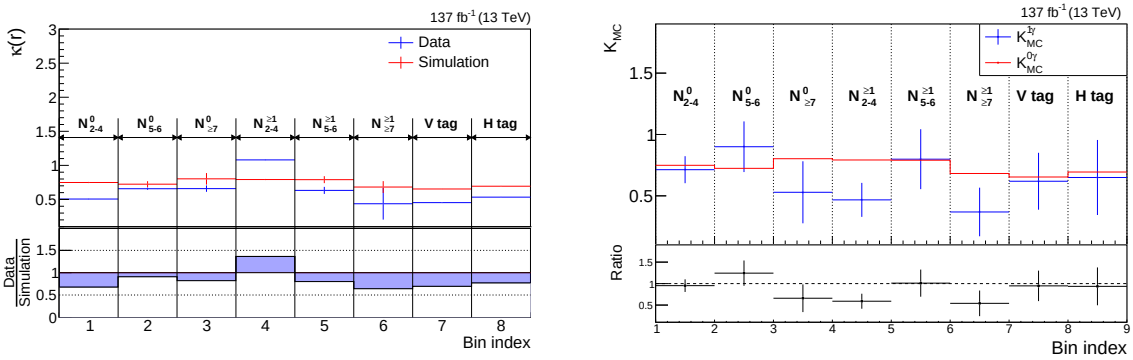

FIGURE 6.31: Left: Distributions for K-factors (κ) is calculated using simulation (red) and data (blue) in zero-photon region. The ratio panel shows the K-factor systematics in [N_{jets} , $N_{\text{b-tags}}$, V-tag, H-tag] bins. Right: κ(MC) distributions for zero-gamma validation region and atleast one photon region using simulations.

total uncertainty (systematic and statistical). Table B.8 show predicted background yield as a function of 45 search bins in high-$\Delta\phi$ SR (D) including low-p_T^{miss} CR (C). These predictions are calculated using B region events from p_T^{miss} full-Run II data after applyling R-factors (corrected by using K-factors).

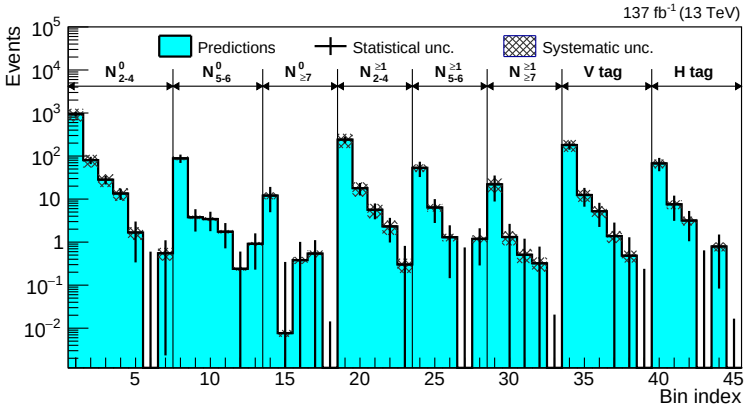

FIGURE 6.32: Final data prediction with total systematic uncertainty as a shaded region and bar line over each bin content represents the statistical uncertainty for $\gamma + jets$ and QCD multijet background estimation.

6.5 Total SM prediction and data observed

After estimating all SM backgrounds, total SM predictions are compared with observed data in the 45 optimised SR bins as shown in Figure 6.33. Four distinct SM backgrounds, namely $\gamma + jets$ and QCD multijet, $Z(\nu\nu)\gamma +$ jets , misidentified electrons, and lost leptons, are stacked over each other and compared with the observed data across these SRs. In most of the SRs, the observed event yields are consistent with the predictions, indicating no significant presence of signal events. The maximum deviation observed is about 2 standard deviations below the prediction, in bin 13 ($5 \leq N_{\text{jets}} \leq 6$, $N_{\text{b-tags}} = 0$, $p_T^{\text{miss}} \geq 750$ GeV), bin 16 ($N_{\text{jets}} \geq 7$, $N_{\text{b-tags}} = 0$, $370 < p_T^{\text{miss}} < 450$ GeV), and bin 44 (H-tag, $600 < p_T^{\text{miss}} < 750$ GeV). The interpretation of these results in the context of searching for massive particles predicted by GMSB SUSY scenarios are presented in the next chapter.

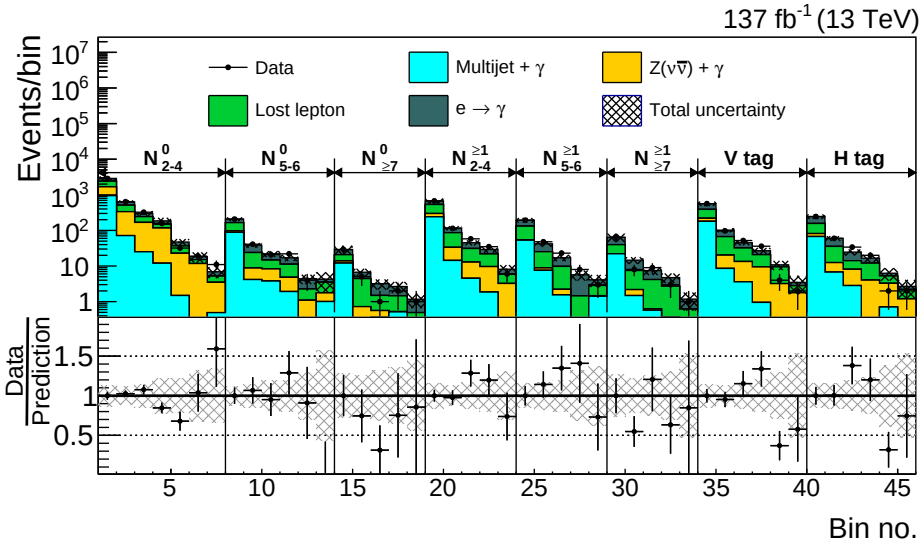

FIGURE 6.33: Top panel: final data prediction with total systematic uncertainty as a shaded region and bar line over each bin content represents the statistical uncertainty for total SM background estimation. Bottom panel: ratio of the observed number of data events and the predicted backgrounds. The error bars represent the statistical uncertainty in the data events, and the shaded band represents the statistical and systematic uncertainties in the predicted background. The p_T^{miss} bin 200–300 GeV is used for the estimation of the $\gamma + jets$ and QCD multijet background.

Chapter 7

Results

This chapter is dedicated to the discussion of procedure for statistical interpretation of the estimated backgrounds, uncertainties on them, and observed data, presented in Chapter 6, in terms of upper limits on production cross-section of pairs of gluinos, squarks and electroweakinos (presented in Chapter 5), which in turn is presented as lower limits on masses of these particles. We discuss the uncertainties on the MC simulated signal yields and statistical tools used for the statistical interpretation as standardized by the CMS collaboration for the purpose.

7.1 Signal systematics

The statistical uncertainty from the limited number of simulated events in the signal samples ranges 0.7–38% when considering all bins; when considering only the most sensitive bins, this uncertainty decreases to 0.1–7%. The following sources of systematic uncertainties are considered for the simulated signal event yields.

- Luminosity: The uncertainties in the total integrated luminosity [134, 135, 136] affect the signal yield in each SR. This effect is taken to be 1.6% for all the SRs and all signal models. Luminosity may depend on trigger paths used and hence we calculate the luminosity explicitly for the data used in this analysis [137]. If we use the json file corresponding to our processed ntuples from the MET datasets, we get 137 fb^{-1} instead of 138 fb^{-1}. So, we believe 137 fb^{-1} is the right number to quote for this paper.
- We reweight the MC simulated signal event samples using the trigger efficiencies calculated in the data as explained in Section 5.4.1 and propagate the respective uncertainties to get the uncertainties on the expected signal event yields in the SRs. The uncertainty in the trigger efficiency for simulated signal events varies from 3–10%, with larger values for SRs with lower p_T^{miss} . These uncertainties are obtained from additional checks (as in Appendix C) which are performed for the validity of the trigger efficiencies, in case if these are not dependent on the variation of the topologies in different samples and other factors.
- The b-jet tagging efficiency and light flavor quark mis-tagging rates in the fast simulation signal samples are corrected with factors derived from the GEANT4 -based simulation

and data. The uncertainties in these corrections are propagated to the final signal yield, and their effect in the SP SRs is up to 10%. Dependency of the b tagging systematic uncertainty on 45 optimised bins are shown in figure 7.1.

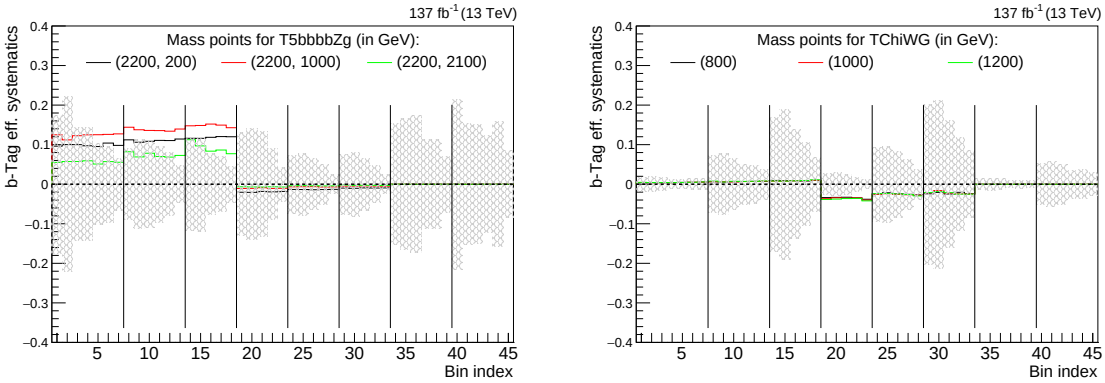

FIGURE 7.1: b tagging systematic uncertainty distributions for strong (on the left) and electroweakino pair (on the rigt) production signal processes in 45 optimised bins (including 27 SR SP , 10 SR EW and 8 low-p_T^{miss} CR). The shaded region represents the statistical uncertainties for the middle mass points in both plots.

- The modelling of initial-state radiation (ISR) in the signal simulation is corrected by applying data-to-simulation correction factors [138]. 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. For strong production, correction factors are 1.000, 0.920, 0.821, 0.715, 0.662, 0.561, 0.511 for N_{jet}^{ISR} = 0, 1, 2, 3, 4, 5, 6+ and these corrections are applied to the simulated signal samples with an additional normalization factor, typically around 1.15 (depending on the signal model), to ensure that the overall cross-section of the sample remains constant. For electroweak samples, correction factors are 1.000, 1.052, 1.179, 1.150, 1.057, 1.000, 0.912, 0.783 for ISR p_T of the pair produced EWKions = 0-50, 50-100, 100-150, 150-200, 200-300, 300-400, 400-500, 500-600, >600 GeV. We use standard CMS SUSY group recommendation [139] to evaluate the uncertainty related to ISR p_T modelling in simulation. These uncertainties have magnitudes of 25% for strong production and 10% for electroweak signal models. Dependency of the ISR systematic uncertainty on 45 optimised bins are shown in figure 7.1.

- The PDF uncertainty, and the μ_R and μ_F scale uncertainties that affect the total production cross-section are treated as theoretical uncertainties. The effects of these on the signal yield are 2% and 5%, respectively [130, 131].

- The jet energy corrections (JECs) are varied using the p_T - and η-dependent jet energy scale uncertainties derived from the simulated and data [90], with a separate set of corrections for the fast simulation samples. These variations are propagated into the various jet-dependent variables, including: N_{jets} , $N_{\text{b-tags}}$, V-tag, H-tag, p_T^{miss} , $\Delta\phi$ (jet, p_T^{miss}) and ST. Similarly, the jet momenta in the MC samples are smeared to match the jet energy resolution (JERs) in data and these smearing factors are varied according to the uncertainties on the jet energy resolution (JERs) measurements. The uncertainties in the JEC

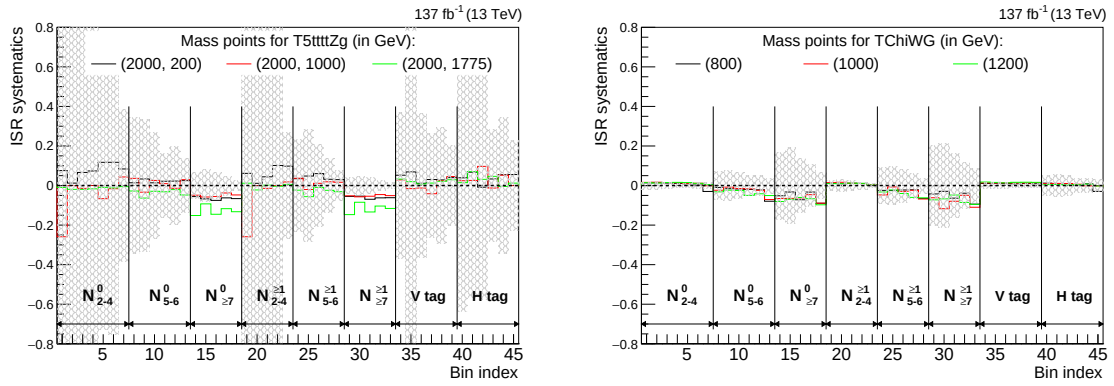

FIGURE 7.2: ISR systematic uncertainty distributions for strong (left) and electroweakino pair (right) production signal processes in 45 optimised bins (including 27 SR SP , 10 SR EW and 8 low-p_T^{miss} CR). The shaded region represents the statistical uncertainties for the middle mass points in both plots.

and JER measurements from data are propagated to individual jets, and their effect on event yields is taken as an uncertainty in the respective SR bins. The uncertainties related to JECs, and JER contribute approximately 2% each.

- The effect of variations in the pileup re-weighting is estimated in a similar way as for JEC and JER and its contribution is also 2% for uncertainty.
- The uncertainties related to the isolated track veto and jet ID modeling in the fast simulation are 2 and 1%, respectively.

Summary of systematic uncertainties for strong productions and electroweakino pair productions is provided in Tables 7.1 and 7.2, respectively. These tables also include bin-wise correlations of the uncertainties. The concept of correlation suggests that uncertainties within the correlated bins tend to vary together in a predictable manner, while uncorrelated bins do not exhibit such interrelated variations.

To gain further insights into the modeling of these correlated uncertainties, reference can be made to Figures 7.1 and 7.2, which illustrate the study of the correlation patterns for these uncertainties. These visual representations help in understanding how and to what extent these uncertainties impact the data analysis, particularly in the context of strong productions and electroweakino pair productions.

Table 7.3 summarizes systematic uncertainties for all background predictions and signal yields. In this table, minimum and maximum uncertainties for signals are taken from detailed systematic tables 7.1 & 7.2.

7.2 Methology for statistical inference from data

The basic idea of statistical inference is to find maximum number of given signal events which can be hidden given the observed data and measured backgrounds along with uncertainties. This can then be converted to an upper limit on signal production cross-section or masses of

TABLE 7.1: Systematic uncertainties for strong production signal MC samples

Uncertainties	T5bbbbZg	T5qqqqHg	T5ttttZg	T6ttZg	correlating modeling
MC−stat	1.4-26 %	1.4-22 %	1.6-36 %	1.6-10 %	uncorrelated
Luminosity	1.6 %	1.6 %	1.6 %	1.6 %	correlated accross all bins
Isotrack veto	0 %	0 %	2 %	2 %	correlated accross all bins
PFJet ID	1 %	1 %	1 %	1 %	correlated accross all bins
JEC	1.0-1.8 %	1.1-1.6 %	0.9-1.9 %	1.7-2.0 %	correlated accross all bins
JER	1.1-1.4 %	0.8-1.4 %	1.2-1.7 %	1.3-1.7 %	correlated accross all bins
Prefiring weight	0-3 %	0.1-1 %	0-2 %	0.1-0.4 %	correlated accross all bins
Pileup reweighting	0.1 %	0.1 %	0.3 %	0.1 %	correlated accross all bins
b−Tag	0-10 %	0-10 %	0-10 %	0-7%	correlated accross 0 b-Tag, >0 b-Tag & EW
ISR reweight	0-25 %	0-25 %	0-20 %	0-20 %	correlated accross N^0_{2-4}, N^0_{5-6}, $N^0_{\geq 7}$, $N^{>0}_{2-4}$, $N^{>0}_{5-6}$, $N^{>0}_{\geq 7}$, V-tag & H-tag region
Scale	1-4 %	0.3-2 %	0.5-5 %	0.5-2 %	correlated accross N^0_{2-4}, N^0_{5-6}, $N^0_{\geq 7}$, $N^{>0}_{2-4}$, $N^{>0}_{5-6}$, $N^{>0}_{\geq 7}$, V-tag & H-tag region
Trigger systematics	3-10 %	3-10 %	3-10 %	3-10 %	10% in p_T^{miss} <300 GeV, 5% in 300≤p_T^{miss} <450 GeV & 3% in p_T^{miss} ≥450 GeV

TABLE 7.2: Systematic uncertainties for electroweak signal MC samples

Uncertainties	TChiWG	TChiNG	correlating modeling
MC−stat	0.7-38 %	0.7-30 %	uncorrelated
Luminosity	1.6 %	1.6 %	correlated accross all bins
PFJet ID	1 %	1 %	correlated accross all bins
JEC	0.9-1.2 %	1.1-1.2 %	correlated accross all bins
JER	1.6-1.7 %	1.7 %	correlated accross all bins
Prefiring weight	0.1-2 %	0-1 %	correlated accross all bins
Pileup reweighting	0.2 %	0.2 %	correlated accross all bins
b−Tag	0-2.3 %	1-4 %	correlated accross 0 b-Tag, >0 b-Tag & EW
ISR reweight	2-10 %	1-10 %	correlated accross N^0_{2-4}, N^0_{5-6}, $N^0_{\geq 7}$, $N^{>0}_{2-4}$, $N^{>0}_{5-6}$, $N^{>0}_{\geq 7}$, V-tag & H-tag region
Scale	1-4 %	0.3-2 %	correlated accross N^0_{2-4}, N^0_{5-6}, $N^0_{\geq 7}$, $N^{>0}_{2-4}$, $N^{>0}_{5-6}$, $N^{>0}_{\geq 7}$, V-tag & H-tag region
Trigger systematics	3-10 %	3-10 %	10% in p_T^{miss} <300 GeV, 5% in 300≤p_T^{miss} <450 GeV & 3% in p_T^{miss} ≥450 GeV

TABLE 7.3: Systematic uncertainties for background and signal samples (in %). A dash (-) indicates that the source of uncertainty is not applicable or too small. When needed a description of the source is shown in the parenthesis.

Source	Lost lepton	Misidentified e	$Z(\nu\nu)\gamma$ + jets	Multijet+γ + *jets*	Signal
Integrated luminosity	-	-	-	-	1.6
Limited number of CR events	3–100	5–20	8–28	2–100	-
Limited number of simulated events	2–10	2–20	2–70	10–50	0.7–38
b tagging	0–1	0–1	-	-	0–10
PDF	3	-	-	-	1–2
μ_R and μ_F scales	2	-	-	-	0.3–5
JEC	0–6	0–3	-	-	1–2
JER	0–6	0–4	-	-	1–2
Pileup	-	-	-	-	0.1-0.3
Trigger efficiency	-	-	-	-	3–10
Collinear γ	4	-	-	-	-
α	-	20	-	-	-
Modeling of γ p_T	-	-	18–40	-	-
κ modeling	-	-	-	10–36	-
low-p_T^{miss} C/A data stat.	-	-	-	10–50	-
Isolated track veto	-	-	-	-	2
Jet ID	-	-	-	-	1

the new physics particles or to find the probability of background fluctuating like signal. This can be done by likelihood estimation using concept of statistics. Given a number of statistically independent SRs used generally in analyses and complexities involved in fitting procedures, the CMS Collaboration has standardized the tools to have consistent conclusions, which can be easily be compared and combined, across the analyses. The package is available as HIGGS COMBINE framework [140]. In this section, I briefly discuss the key statistical concepts and how to obtain exclusion limits.

7.2.1 Likelihood function

The likelihood is a measure of the compatibility between a hypothesis and a given dataset. It is dependent on the data and represents the probability of observing the data for a specific value of a physics parameter, denoted as α.

$$\mathcal{L}(\alpha) \approx \text{p(data}|\alpha) \tag{7.1}$$

This α is dependent on point of interest, μ ($\sigma_{obs.}/\sigma_{SM}$) and nuisance parameters, θ_i. However, likelihood should not be treated as probability distribution for α because because the integral of the likelihood over α does not equal 1. Likelihood function for ith bin is defined as:

$$\mathcal{L}(\mu, \theta_i) = \left(\frac{\mu s_i + b_i}{n_i!}\right)^{n_i} e^{-(\mu s_i + b_i)} F(\theta_i) \tag{7.2}$$

where s_i, b_i & n_i are the signal, background and data yield in the i^{th} bin, respectively. The point of interest, μ represents the signal strength which will be 1 when signal hypothesis and μ=0 in case of background hypothesis. F(θ_i) is the probability distribution associated with subsidiary measurements that help constrain the nuisance parameters. to find the parameter values that maximize the likelihood function. To remove the dependency of likelihood on θ,

profiling method is used in which the value of θ is found such that $\mathcal{L}(\mu, \theta)$ is maximum. This can involve varying the parameter values and comparing the simulated data to the observed data. The parameter values that yield the best agreement or highest likelihood are considered as the maximum likelihood estimates. In the context of the LHC, the profile likelihood ratio is often employed to test a hypothesized value of μ:

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \tag{7.3}$$

where the numerator includes the value of θ, denoted as $\hat{\hat{\theta}}$, that maximizes the likelihood function for the specified value of μ. This value represents the conditional maximum-likelihood estimator of θ, and it is dependent on μ. On the other hand, the denominator represents the maximized likelihood function without any constraints, where $\hat{\mu}$ and $\hat{\theta}$ are the maximum-likelihood estimators of μ and θ, respectively. The inclusion of nuisance parameters in the likelihood analysis results in a broadening of the profile likelihood as a function of μ compared to the scenario where these parameters are fixed. This broadening reflects the impact of systematic uncertainties, leading to a loss of information about μ.

7.2.2 Test statistics

Test statistics play a crucial role in hypothesis testing, parameter estimation, and assessing the significance of experimental results. For discovering a new signal process, one defines the background hypothesis (H_0) which is tested against the signal+background hypothesis (H_1) whereas, when setting limits, the model with signal plus background plays the role of H_0, which is tested against the background-only hypothesis, H_1. Test statistics are used to quantify the level of agreement or discrepancy between observed data and theoretical predictions or hypotheses. It is defined using profile likelihood ration ($\lambda(\mu)$) as:

$$q_\mu = -2\, ln(\lambda(\mu)) \tag{7.4}$$

Higher values of q_μ thus correspond to increasing incompatibility between data and prediction. Test statistics are used to calculate the p-value, and smaller p-values indicate stronger evidence against the H_0 hypothesis. The p-value is a measure of the statistical significance of the observed data under a specific hypothesis. It represents the probability of obtaining data as extreme as or more extreme than the observed data, assuming the H_0 hypothesis is true. It is computed as:

$$p_\mu = \int_{t_{\mu,obs}}^{\infty} f(t_\mu, \mu) dt_\mu \tag{7.5}$$

In the equation, $t_{\mu,obs}$ denotes the value of the test statistic t_μ observed from the data. The $f(t_\mu|\mu)$ represents the probability density function (pdf) of t_μ under the assumption of a specific signal strength μ. The relationship between the p-value, the observed t_μ, and the significance Z is illustrated in Figure 7.3.

If the p-value is less than a specified threshold α (typically $\alpha = 0.05$ correspond to Z=1.64), the

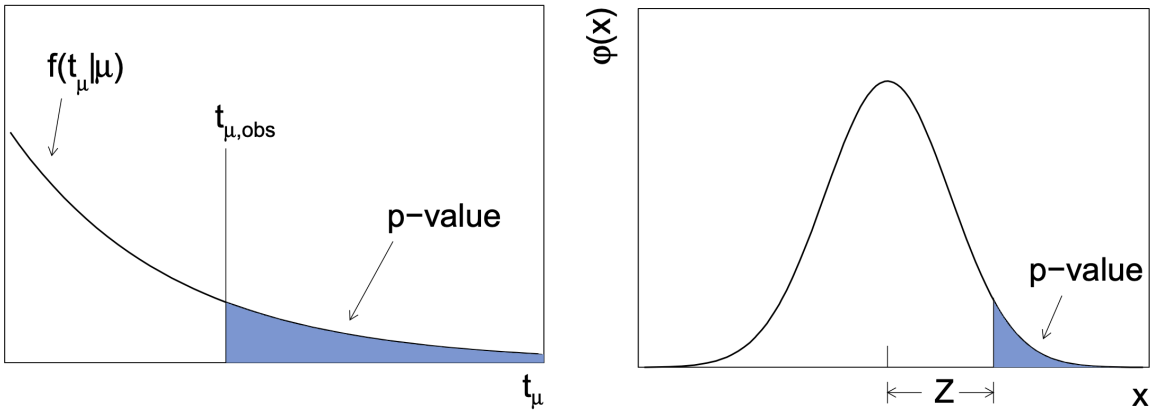

FIGURE 7.3: Distributions of the test statistics (left) and the normal distribution (right) [126].

value of μ is considered excluded at a confidence level (CL) of $1-\alpha$. The collection of points that are not excluded constitutes a confidence interval with CL = $1-\alpha$ (in this case, it will be 95% CL). The endpoints of the interval can be obtained by setting $p_\mu = \alpha$ and solving for μ.

7.3 Results

All the concepts described in section 7.2, are implemented in a coding tool recommended in CMS, the HIGGS COMBINE tool. Using the HIGGS COMBINE tool, we perform a simultaneous maximum likelihood fit to the number of events in the low- and high-$\Delta\phi$ regions to predict the number of SM background events in the SRs. The profile likelihood ratio, q_μ is used as the test statistic to compute limits in the modified frequentist CLs approach [141, 142], employing the asymptotic approximation [126]. The statistic is defined as $q_\mu = -2\ln(L_\mu/L_{max})$, where μ_{Sig} is the SUSY signal strength, L_{max} is the maximum likelihood from varying all parameters including μ_{Sig}, and L_μ is the maximum likelihood for a fixed μ_{Sig}. The observed numbers of events in various CRs are modeled using gamma distributions, which correctly represent the statistical uncertainties. The other systematic uncertainties listed in Table 7.3 are modeled as log-normal constraints in the likelihood.

Additionally, we incorporate low-$\Delta\phi$ control regions as inputs into the HIGGS COMBINE framework. These predictions are derived through the utilization of SR and CR bins, running fits while keeping SR observations masked. So, the results obtained from the CR-only fit under the background-only hypothesis as shown in Fig. 7.4. The numerical values, including the uncertainties in each background prediction, are given in the tables B.1, B.4, B.6 & B.8. In most of the SRs, the observed event yields are consistent with the predictions, indicating no significant presence of signal events. The maximum deviation observed is about 2.0 standard deviations below the prediction, in bin 13 ($5 \leq N_{\text{jets}} \leq 6$, $N_{\text{b-tags}} = 0$, $p_T^{\text{miss}} \geq 750$ GeV), bin 16 ($N_{\text{jets}} \geq 7$, $N_{\text{b-tags}} = 0$, $370 < p_T^{\text{miss}} < 450$ GeV), and bin 44 (H-tag, $600 < p_T^{\text{miss}} < 750$ GeV).

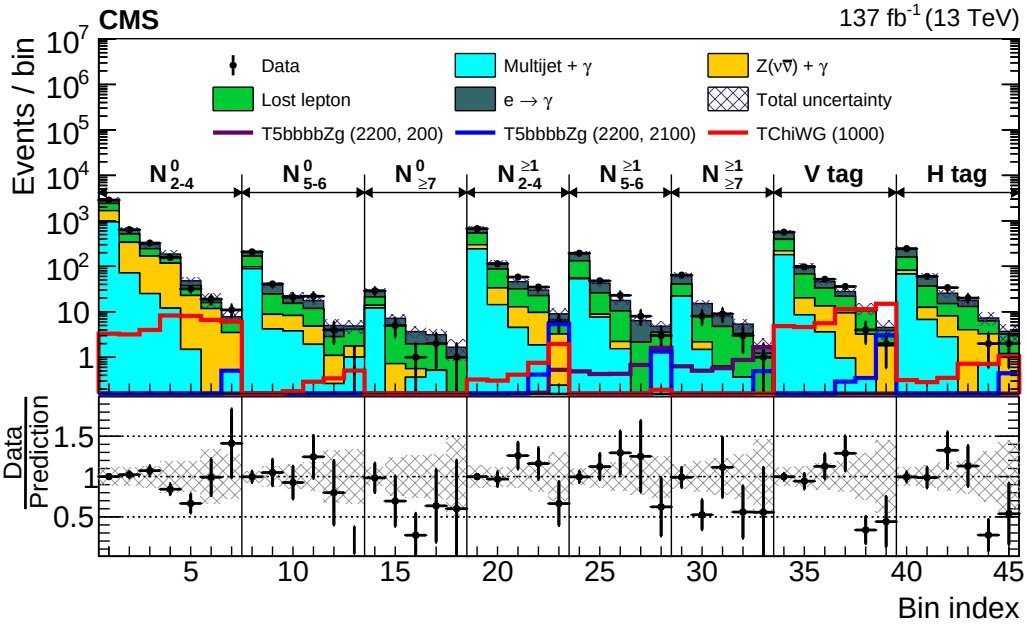

FIGURE 7.4: The numbers of predicted background events and observed events in the SRs and low-p_T^{miss} CRs. Predicted lost lepton, fake photon, $Z(\nu\nu)\gamma$ + jets and $\gamma + jets$ +QCD are stacked histograms. The observed numbers of events in data are presented as black points. For illustration, the expected event yields are presented for the signal model T5bbbbZG, for small (blue) and large (purple) differences in the masses of the $\tilde{g}$ and NLSP. Also shown is the expected distribution of events for the signal model TChiWG (red). The numerical values in parentheses in the legend entries for the signal models indicate the $\tilde{g}$ and NLSP mass values in GeV for strong production and the NLSP mass value for electroweak production. The lower panel shows the ratio of the observed number of data events and the predicted backgrounds. The error bars represent the statistical uncertainty in the data events, and the shaded band represents the statistical and systematic uncertainties in the predicted background. The p_T^{miss} bin 200–300 GeV is used for the estimation of the $\gamma + jets$ and QCD multijet background.

7.4 Exclusion Limits

The measured backgrounds along with their uncertainties and the observed number of events in the SRs are used to determine 95% confidence level (CL) upper limits on the production cross-sections of various SUSY models, discussed in Section 5.1 of chapter 5, using a maximum likelihood fit. We compare these upper limits with theoretical production cross-sections, and determine lower limits on masses of the SUSY particles in specific models and final states, which are excluded by this search.

7.4.1 Strong production simplified signal models

The expected and observed exclusion limits in terms of masses of gluino or stop (in x-axis) and mass of NLSP (in y-axis) involved in given strong production signals model as shown in Figure 7.5 for gluino and squark pair production scenarios.

For the gluino production model with a decay to $b\bar{b}$ and NLSP (T5bbbbZG), the observed (expected) gluino mass exclusion is up to 2.32 (2.27) TeV for small NLSP masses. In the T5qqqqHG and the T5ttttZG models, the observed (expected) upper limits on the gluino masses extend to

2.35 (2.30) and 2.26 (2.25) TeV, respectively. The mass limits degrade for very high and very low NLSP masses. When the NLSP masses are large, the p_T^{miss} is large but the events contain lower hadronic activity (N_{jets}). For the low NLSP masses, hadronic activity is high, but p_T^{miss} is low. In these scenarios, either the signal acceptance is low or the signals populate SRs with larger backgrounds, resulting in a decrease in the sensitivity of the analysis. These features are illustrated in the open histograms shown in Fig. 7.4.

For the T5qqqqHG scenario, the observed limits are stronger than the expected ones because of the small deficit in the observed event yields in the high-N_{jets} and $N_{\text{b-tags}}$ =0 regions and in the high-p_T^{miss} H-tag regions. In the strong production models that involve off-shell Z bosons and very low NLSP masses, the limits are stronger than those from on-shell Z bosons, because the former imparts larger p_T to the gravitinos, leading to larger p_T^{miss} . In the T5qqqqHG model, off-shell H boson decays are not considered. For the top squark pair production model (T6ttZg), where the top squark decays into a top quark and NLSP, the expected mass limit is 1.38 TeV and the observed mass limit is 1.43 TeV. There is an approximately 0.7 standard deviation difference between the expected and observed limits, which comes from signal regions with high p_T^{miss} , high N_{jets} , and $N_{\text{b-tags}} \geq 1$.

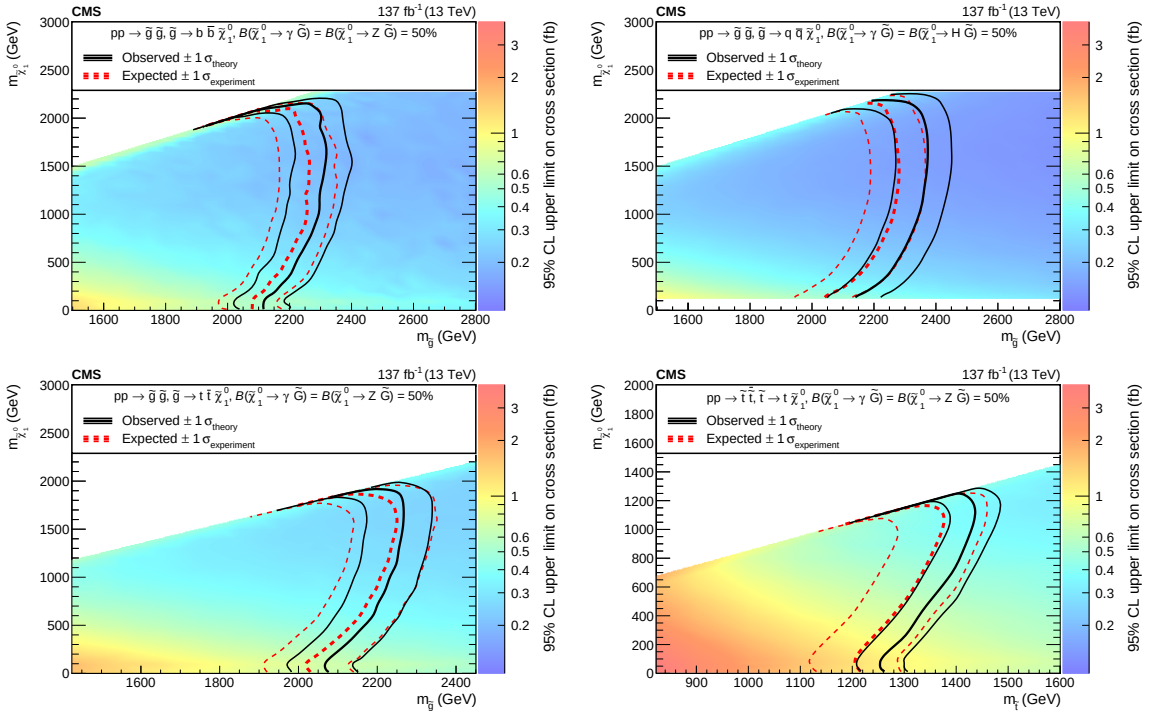

FIGURE 7.5: Comparison of expected and observed exclusion for T5bbbbZG (top left), T5qqqqHG(top right), T5ttttZG (bottom left) & T6ttZG (bottom right) model at 95% confidence level reached by current search.

7.4.2 Electroweakino simplified models

Figure 7.6 presents the exclusion limits (in y-axis) in terms of masses of NLSP (in x-axis) for the production of electroweakino pairs. In electroweak production models, for sufficiently large electroweakino masses, the signal events mostly populate large p_T^{miss} and V- or H-tag

signal regions. In the TChiWG scenario, we observe (expect) the exclusion of $\tilde{\chi}_1^{\pm}\tilde{\chi}_1^0$ masses up to 1.23 (1.17) TeV, assuming wino-like $\tilde{\chi}_1^{\pm}\tilde{\chi}_1^0$ production. In the TChiNG model, the $\tilde{\chi}_1^{\pm}$, $\tilde{\chi}_1^0$, and $\tilde{\chi}_2^0$ are nearly mass degenerate, and we use higgsino-like cross-sections to interpret the results. Electroweakino masses below 1.05 (0.95) TeV are observed (expected) to be excluded, assuming the $\tilde{\chi}_1^{\pm}$ and $\tilde{\chi}_2^0$ decays give rise to $\tilde{\chi}_1^0$ and soft particles. The expected upper limit on the mass is about 60–80 GeV smaller than the observed upper limit because of a deficit in the observed event yields in the highest p_T^{miss} bins of the V- and H-tag SRs. For the lower electroweakino masses, the signal populates moderate p_T^{miss} bins, which have observed event yields higher than the predictions by about one standard deviation. This leads to higher than predicted observed upper limits on the production cross-section. Alternatively, if only the $\tilde{\chi}_1^0\tilde{\chi}_2^0$ process occurs in this model, there is no exclusion in the range of NLSP masses considered in this search, based on the theoretically predicted cross-section. In the TChiNGnn model, with only the $\tilde{\chi}_1^0\tilde{\chi}_2^0$ process, the observed (expected) NLSP mass limit is 0.50 (0.65) TeV. These are the best mass limits to date on electroweakino production with photons in the final state. The observed limit is weaker than expected because of upward fluctuations in several low-N_{jets} $N_{\text{b-tags}} \geq 1$, V-tag, and H-tag SR bins with intermediate p_T^{miss} values.

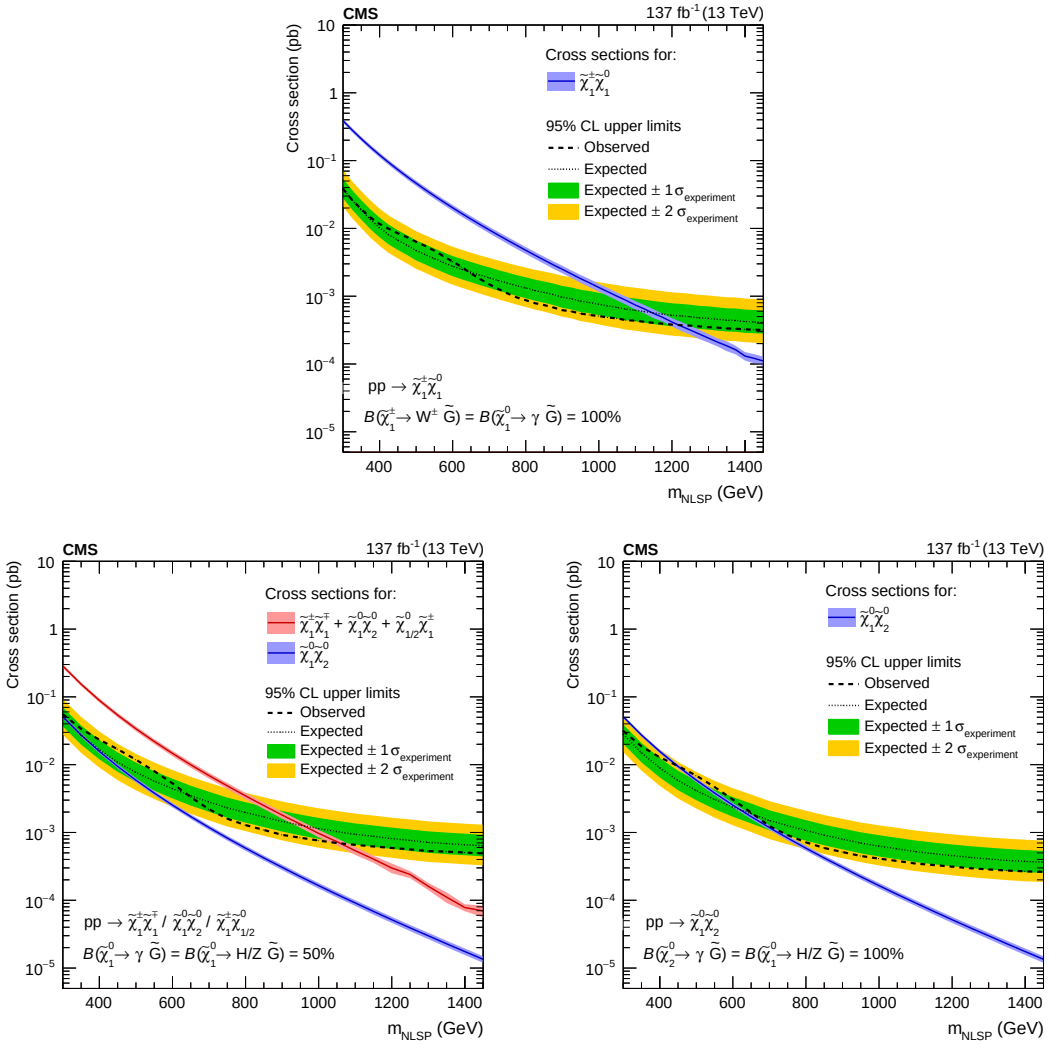

FIGURE 7.6: The expected and observed limits on the electroweakino mass in the TChiWG (upper left), TChiNG (lower left), and TChiNGnn (lower right) models at 95% CL. For the TChiNG model (lower left), scenarios with degenerate charginos and neutralinos leading to the combined process $\tilde{\chi}_1^{\pm}\tilde{\chi}_1^{\mp} + \tilde{\chi}_1^0\tilde{\chi}_2^0 + (\tilde{\chi}_1^0 / \tilde{\chi}_2^0)\tilde{\chi}_1^{\pm}$ (red) or the single process $\tilde{\chi}_1^0\tilde{\chi}_2^0$ (blue) are considered.

TABLE 7.4: Background predictions and observed event yields for SR region (p_T^{miss} >300 GeV). Events yields in low-p_T^{miss} bins (200-300 GeV) are the sidebands used for multijet predictions in other bins. For the background predictions, the uncertainty is stat + sys (added in quadrature).

Bin	MET [GeV]	Lost-lepton background	e → fake γ background	$Z(\nu\nu)\gamma$ + jets background	Multijet + γ background	Total Predictions	Observed
			$2 \leq N_{\text{jets}} \leq 4$, $N_{\text{b-tags}} = 0$				
1	200 – 300	731 ± 63.7	459 ± 96.9	730 ± 146	946 ± 267	2.87e+03 ± 317	2.86e+03
2	300 – 370	174 ± 18.6	116 ± 25.7	265 ± 59	71.5 ± 25.8	626 ± 72.1	641
3	370 – 450	76.5 ± 9.67	58.2 ± 13	143 ± 36.4	25.1 ± 9.04	303 ± 39.6	325
4	450 – 600	33.1 ± 5.54	36.5 ± 7.66	105 ± 34.4	12 ± 4.32	186 ± 35.9	157
5	600 – 750	14.8 ± 3.56	10.4 ± 2.32	21.3 ± 7.32	1.49 ± 0.536	48 ± 8.82	32
6	750 – 900	4.17 ± 1.65	3.23 ± 0.721	11.8 ± 6.4	$0^{+0.0036}_{-0}$	19.2 ± 6.45	19
7	>900	2.59 ± 1.33	1.71 ± 0.503	3 ± 1.73	0.491 ± 0.177	7.79 ± 2.16	11
			$5 \leq N_{\text{jets}} \leq 6$, $N_{\text{b-tags}} = 0$				
8	200 – 300	68.3 ± 7.98	42.7 ± 9.47	8.96 ± 2.62	88.6 ± 7.8	209 ± 14.2	208
9	300 – 370	15.8 ± 3.39	14.5 ± 3.32	4.6 ± 1.71	4.17 ± 0.885	39 ± 4.66	41
10	370 – 450	7 ± 2.24	7.39 ± 1.63	4.49 ± 1.85	3.79 ± 0.803	22.7 ± 3.35	21
11	450 – 600	7 ± 2.1	5.84 ± 1.44	2.9 ± 1.77	1.92 ± 0.408	17.7 ± 3.13	22
12	600 – 750	1.75 ± 1.1	2.15 ± 0.54	0.831 ± 1.01	0.264 ± 0.056	5 ± 1.61	4
13	>750	2.33 ± 1.22	0.779 ± 0.254	0.754 ± 1.01	1.01 ± 0.214	4.87 ± 1.61	0
			$N_{\text{jets}} \geq 7$, $N_{\text{b-tags}} = 0$				
14	200 – 300	7.44 ± 2.01	8.2 ± 1.16	1.77 ± 0.873	12.1 ± 1.94	29.5 ± 3.19	29
15	300 – 370	4.19 ± 1.59	2.28 ± 0.354	0.717 ± 0.58	0.00729 ± 0.0037	7.19 ± 1.7	5
16	370 – 450	1.4 ± 0.883	1.72 ± 0.295	0.2 ± 0.124	0.362 ± 0.184	3.68 ± 0.968	1
17	450 – 600	1.4 ± 0.787	1.22 ± 0.272	0.0115 ± 0.0139	0.515 ± 0.261	3.14 ± 0.857	2
18	>600	0.931 ± 0.786	0.695 ± 0.161	0.0271 ± 0.0397	$0^{+0.00508}_{-0}$	1.66 ± 0.808	1
			$2 \leq N_{\text{jets}} \leq 4$, $N_{\text{b-tags}} \geq 1$				
19	200 – 300	238 ± 22.1	139 ± 29	57 ± 19.6	241 ± 90.6	675 ± 93.7	674
20	300 – 370	53.9 ± 7.91	30.3 ± 6.29	19.1 ± 7.61	14.4 ± 6.6	118 ± 13.9	114
21	370 – 450	19.1 ± 4.01	14 ± 3.04	8.35 ± 3.9	4.52 ± 2.08	46 ± 6.38	58
22	450 – 600	13 ± 3.4	7.45 ± 1.64	7.8 ± 4.08	1.86 ± 0.853	30.1 ± 5.53	35
23	>600	3.47 ± 1.56	2.3 ± 0.579	3 ± 2.31	0.242 ± 0.111	9.02 ± 2.73	6
			$5 \leq N_{\text{jets}} \leq 6$, $N_{\text{b-tags}} \geq 1$				
24	200 – 300	77.6 ± 9.77	62.4 ± 13.6	1.55 ± 1.09	53.2 ± 9.99	195 ± 18.6	194
25	300 – 370	17.2 ± 3.87	16.7 ± 3.73	1.26 ± 0.773	7.57 ± 2.33	42.7 ± 5.91	48
26	370 – 450	8.24 ± 2.46	7.31 ± 1.59	0.672 ± 0.633	1.54 ± 0.475	17.8 ± 3.12	23
27	450 – 600	2.06 ± 1.11	4.25 ± 0.946	0.0772 ± 0.0616	$0^{+0.00308}_{-0}$	6.39 ± 1.46	8
28	>600	2.06 ± 1.15	1.27 ± 0.292	0.0587 ± 0.0452	1.41 ± 0.435	4.81 ± 1.22	3
			$N_{\text{jets}} \geq 7$, $N_{\text{b-tags}} \geq 1$				
29	200 – 300	18.3 ± 4.05	24.2 ± 5.13	0.0767 ± 0.0579	22.1 ± 7.45	64.6 ± 9.96	64
30	300 – 370	5.89 ± 2.02	7.14 ± 1.65	0.697 ± 0.567	1.48 ± 1.02	15.2 ± 2.86	8
31	370 – 450	4.12 ± 1.58	3.31 ± 0.71	0.06 ± 0.0555	0.573 ± 0.395	8.07 ± 1.76	9
32	450 – 600	2.95 ± 1.36	2.04 ± 0.466	$0^{+0.00472}_{-0}$	0.364 ± 0.251	5.36 ± 1.48	3
33	>600	1.18 ± 0.839	0.581 ± 0.223	0.027 ± 0.0283	$0^{+0.00689}_{-0}$	1.8 ± 0.83	1
			EW : V-tag				
34	200 – 300	172 ± 17.4	174 ± 35.1	39.2 ± 8.13	180 ± 50.7	565 ± 62.9	564
35	300 – 370	47.8 ± 8.4	34.9 ± 7.64	11.6 ± 3.59	8.57 ± 4.06	103 ± 12.7	97
36	370 – 450	19.8 ± 4.91	13 ± 2.9	9.8 ± 3.38	3.6 ± 1.71	46.2 ± 7.22	52
37	450 – 600	12.5 ± 3.24	6.02 ± 1.37	8.48 ± 3.55	0.952 ± 0.451	27.9 ± 5.15	36
38	600 – 750	7.28 ± 2.91	1.38 ± 0.338	2.88 ± 1.72	0.334 ± 0.158	11.9 ± 3.36	4
39	>750	2.08 ± 1.37	0.67 ± 0.178	1.78 ± 1.52	$0^{+0.00473}_{-0}$	4.54 ± 2.01	2
			EW : H tag				
40	200 – 300	76.3 ± 10.3	88.2 ± 18.7	14.1 ± 3.69	67.5 ± 15.3	246 ± 27.2	245
41	300 – 370	24.1 ± 5.44	24.2 ± 5.11	5.73 ± 2.2	6.75 ± 2.94	60.7 ± 8.26	60
42	370 – 450	7.02 ± 2.83	10.5 ± 2.23	5.28 ± 2.41	2.83 ± 1.23	25.6 ± 4.49	34
43	450 – 600	9.03 ± 2.98	4.62 ± 1.05	4.02 ± 2.15	$0^{+0.00436}_{-0}$	17.7 ± 3.8	20
44	600 – 750	3.01 ± 1.76	1.02 ± 0.278	2.56 ± 1.34	0.706 ± 0.308	7.3 ± 2.28	2
45	>750	2.01 ± 1.4	0.504 ± 0.144	1.19 ± 1.05	$0^{+0.00436}_{-0}$	3.72 ± 1.66	2

Chapter 8

Summary and outlook

A search for supersymmetry (SUSY) is presented in this thesis, based on a proton-proton collision dataset at a center-of-mass energy of 13 TeV recorded with the CMS detector. The data correspond to an integrated luminosity of 137 fb^{-1}. The CMS experimental setup includes object reconstructions from the data collected in the CMS detector, and the PF hadron calibration which is a part of the PF algorithm (used in object reconstruction at the CMS detector). Within this thesis, we detailed the necessity for PF hadron calibrations and explained the methodology employed, emphasizing the performance gains achieved through the reprocessing of Run II data (Ultra-legacy). Moreover, we undertook an examination of the effects of modifying ECAL noise thresholds concerning η on PF charged hadrons, as well as the notable enhancements following the application of these calibrations. In closing, we also presented developmental studies that aim to augment the efficiency of PF calibrations using Graph Neural Networks (GNN) in conjunction with Run III data. These studies provide a promising direction for achieving substantial improvements in resolution, ushering in a new era of precision in our research.

Furthermore, our research extends into the realm of new physics, particularly within the SUSY search domain. The SUSY search is designed to probe the events with final states containing at least one photon, large p_T^{miss} and jets that may or may not arise from b quarks. These signatures are motivated by models with GMSB, in which the LSP is a gravitino ($\tilde{G}$) and the NLSP is a chargino ($\tilde{\chi}_1^{\pm}$) or neutralino ($\tilde{\chi}_1^0$), collectively called electroweakinos. Several simplified models of strong production of pairs of gluinos and top squarks are considered, with the gluino decaying to a pair of quarks along with an NLSP or the top squark decaying to a top quark and an NLSP; the NLSP then decays to a neutral gauge boson (photon, Z boson, or H boson) and an LSP. Models of pair production of electroweakinos are also considered, with the neutralinos decaying as described above, and the charginos decaying to a W boson and an LSP. The events are divided into mutually exclusive 45 bins characterized by the p_T^{miss} , N_{jets} , $N_{\text{b-tags}}$, and m_J of jets consistent with the presence of hadronically decaying W, Z, or Higgs bosons. This analysis is covering most of the phase space i.e. strong and electroweak GMSB models.

Compared to previous searches, this search achieves increased sensitivity to scenarios with small mass differences between the gluino and the NLSP with dedicated search regions based

on identifying boosted massive bosons. In addition, the search strategy is expanded to provide sensitivity to the production of electroweakino pairs. The observations are consistent with the standard model expectations and 95% confidence level upper limits are set on the production cross-sections of SUSY particles. In the GMSB simplified models, the lower gluino mass limit reaches up to 2.35 TeV for models with $\tilde{g} \to q\bar{q}\tilde{\chi}_1^0$ followed by $\tilde{\chi}_1^0 \to H\tilde{G}$ or $\gamma\tilde{G}$ with equal probability, and the top squark mass limit reaches up to 1.43 TeV for models with $\tilde{t} \to t\tilde{\chi}_1^0$ followed by $\tilde{\chi}_1^0 \to Z\tilde{G}$ or $\gamma\tilde{G}$ with equal probability. These results extend the previous mass limits [35] on gluinos and top squarks by 150–200 GeV. For electroweakino pair production, chargino and neutralino masses up to 1.23 TeV are excluded, assuming wino-like electroweakinos with decays $\tilde{\chi}_1^{\pm} \to W\tilde{G}$ and $\tilde{\chi}_1^0 \to \gamma\tilde{G}$. The higgsino-like electroweakino mass limits reach up to 1.05 TeV for models with $\tilde{\chi}_1^0 \to \gamma\tilde{G}$, $Z\tilde{G}$, or $H\tilde{G}$ with 50, 25, and 25% branching fractions, respectively. These are the best mass limits to date on electroweakino production with photons in the final state. These results have been published in journal of high energy physics (JHEP) [143]. Summary of the maximum excluded SUSY particle mass for various simplified models within the framework of GMSB are shown in the Figure 8.1, including the results of this analysis (labelled as "SUS-21-009"). These are the best mass limits to date on electroweakino production with photons in the final state.

In this thesis, we mainly focused on SUSY search with large p_T^{miss} and higher photon p_T phase space. As a logical progression, we can extend our analysis into a phase space with lower photon p_T while incorporating advanced machine learning techniques. Furthermore, we anticipate harnessing the statistical advantages presented by higher luminosity datasets, notably Run 3 and the High-Luminosity LHC (HL-LHC). Using Run 3 dataset, LHC is set to continue its quest for new physics by delivering an even larger dataset compared to Run 2. This abundance of data will significantly improve our capacity to explore rare processes and conduct new physics searches. Furthermore, the HL-LHC represents a transformative phase in particle physics. Its unprecedented luminosity will facilitate precision physics and empower the exploration of known particles and interactions at an unparalleled level of accuracy. This will enrich our comprehension of fundamental physics. Moreover, the HL-LHC's extraordinary luminosity will unlock the door to the observation of exceptionally rare processes and events, offering a unique opportunity to detect elusive phenomena that have remained hidden under lower luminosities. It will be pivotal in the ongoing search for dark matter and other enigmatic particles, potentially delivering groundbreaking discoveries.

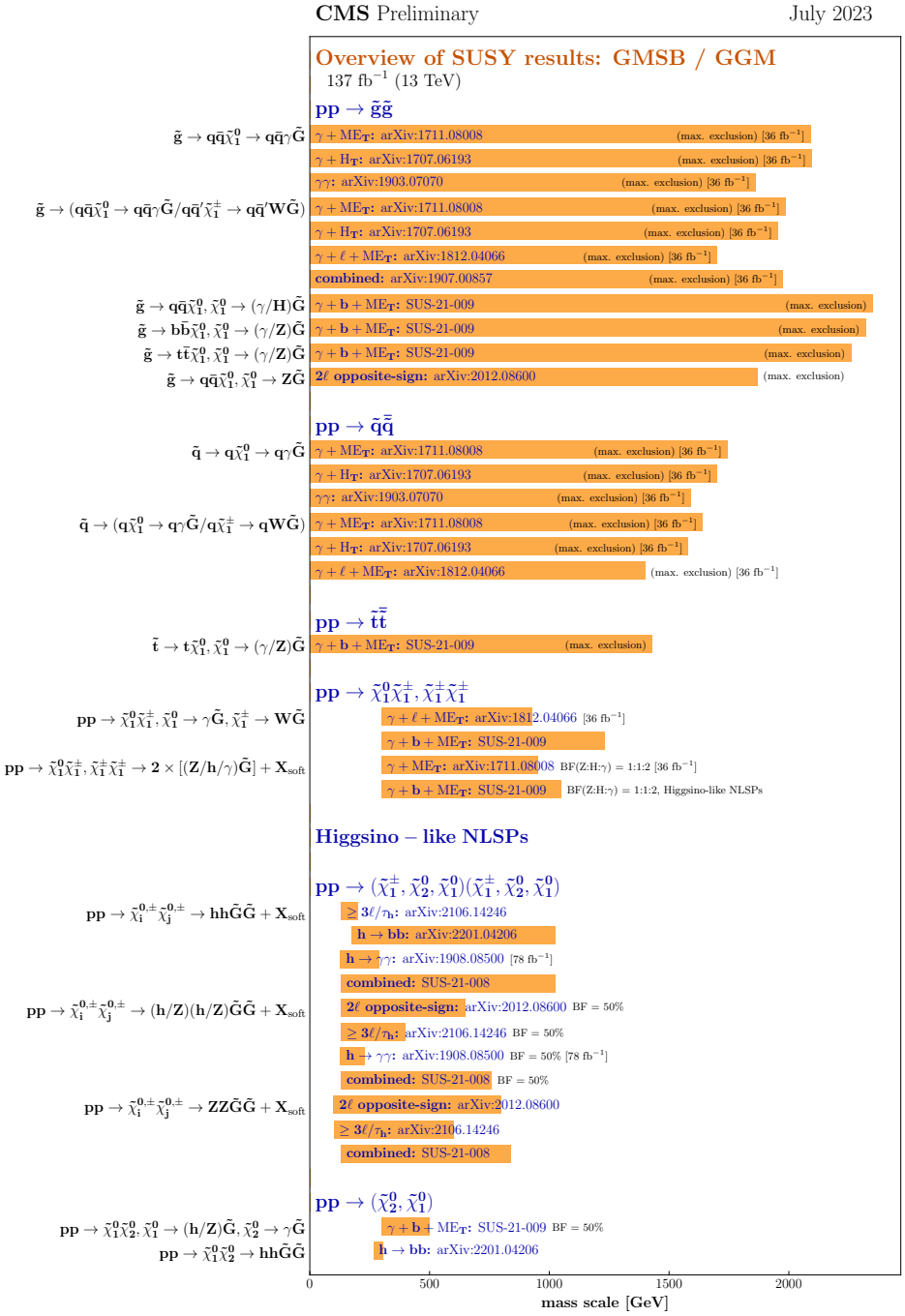

FIGURE 8.1: Overview of recent bounds on lower mass limits of simplified SUSY models with different final states in the context of GMSB.[144]

Appendix A

Results for PF charged Hadron calibration

A.1 Technical details and results for re-processing of Run2 (2016-2018) data

Ultra-legacy (reprocessed Run 2) samples are shown in Table A.1 & Table A.2 containing technical details of those samples. The details of all parameterization functions and their values used in UL samples for all years are mentioned in the tables A.3, A.4, A.5 & A.6 in the section A.1.1.

TABLE A.1: Dataset of single pion sample for UL 2016, UL 2017 & UL 2018

Year	Dataset name	Total Events
2016	/Single_Pion_gun_E_2to200_13TeV_pythia8/RunIIWinter19PFCalibDR-2016ConditionsNoPU_105X_mcRun2_asymptotic_newECALSR_v2-v2/GEN-SIM-RECO	10 M
	/Single_Pion_gun_E_200to500_13TeV_pythia8/RunIIWinter19PFCalibDR-2016ConditionsNoPU_105X_mcRun2_asymptotic_newECALSR_v2-v2/GEN-SIM-RECO	10 M
2017	/Single_Pion_gun_E_2to200_13TeV_pythia8/RunIIWinter19PFCalibDR-2017ConditionsNoPU_105X_mc2017_realistic_v5-v1/GEN-SIM-RECO	10 M
	/Single_Pion_gun_E_200to500_13TeV_pythia8/RunIIWinter19PFCalibDR-2017ConditionsNoPU_105X_mc2017_realistic_v5-v1/GEN-SIM-RECO	10 M
2018	/Single_Pion_gun_E_2to200_13TeV_pythia8/RunIIWinter19PFCalibDR-2018ConditionsNoPU_105X_upgrade2018_realistic_v4-v1/GEN-SIM-RECO	10 M
	/Single_Pion_gun_E_200to500_13TeV_pythia8/RunIIWinter19PFCalibDR-2018ConditionsNoPU_105X_upgrade2018_realistic_v4-v1/GEN-SIM-RECO	10 M

TABLE A.2: Technical details used for the production of UL samples.

	2016	2017	2018
CMSSW version	CMSSW_10_5_0	CMSSW_10_5_0	CMSSW_10_5_0
Global Tag	105X_mcRun2_asymptotic_v2	105X_mc2017_realistic_v5	105X_upgrade2018_realistic_v4
Particle ID	-211 (π^-)	-211 (π^-)	-211 (π^-)
Energy range	2−500 GeV	2−500 GeV	2−500 GeV
$\lvert\eta\rvert$ & $\lvert\phi\rvert$ ranges	$\lvert\eta\rvert < 3, \lvert\phi\rvert < 3.14$	$\lvert\eta\rvert < 3, \lvert\phi\rvert < 3.14$	$\lvert\eta\rvert < 3, \lvert\phi\rvert < 3.14$

A.1.1 List of calibration functions and parameters for all years

Calibrations used for Energy dependent calibrations

Year	EH Barrel	
	a(E_t)	b(E_t)
2016	$[0]+([1]+[2]/x^{[5]})*exp(-x^{[4]}/[3])$ -1.8163,2.8828,7.1228,0.11819,-1.4989,0.5988	$[0]+(([1]+[2]/x^{1/2})*exp(-x^{[6]}/[3]))-([4]*exp(-x^{[7]}/[5]))$ 1.7834,-0.9073,-21.1843,0.6406,0.8203,0.0867,0.1666,-0.8541
2017	$[0]+([1]+[2]/x^{[5]})*exp(-x^{[4]}/[3])$ -1.8163,2.8828,7.1228,0.11819,-1.4989,0.5988	$[0]+(([1]+[2]/x^{1/2})*exp(-x^{[6]}/[3]))-([4]*exp(-x^{[7]}/[5]))$ 1.7495,0.3174,-7.78609,1.44197,0.818853,0.0734628,0.135386,-0.836217
2018	$[0]+([1]+[2]/x^{1/2})*exp(-x^{[4]}/[3])$ -0.2380,1.2558,4.7367,0.0334,-2.0197	$[0]+(([1]+[2]/x^{1/2})*exp(-x^{[6]}/[3]))-([4]*exp(-x^{[7]}/[5]))$ 1.29058,0.506012,-4.8232,3.466,0.3094,0.01268,0.4377,-1.4166
Year	EH Endcap	
	a(E_t)	b(E_t)
2016	$[0]+([1]+[2]/x^{[5]})*exp(-x^{[4]}/[3])$ 1.13723,10.1688,-23.9953,1.23784,0.27871	$[0]+([4]*(x-[5])*exp(-x*[7]))+(([1]+([2]/x^{1/2}))*exp(-x^{[6]}/[3]))$ -2.2433,3.1598,2.5531,9.5396,0.1669,21.7192,-0.2768,0.1172
2017	$[0]+([1]+[2]/x^{[5]})*exp(-x^{[4]}/[3])$ 1.1278,27.6333,-63.9996,0.670415,0.20647	$[0]+([4]*(x-[5])*exp(-x*[7]))+(([1]+([2]/x^{1/2}))*exp(-x^{[6]}/[3]))$ -2.0995,3.4164,2.8106,5.4162,-0.5432,-4.3528,-0.04206,0.242
2018	$[0]+([1]+[2]*x^{[5]})*x^{[4]}*exp(-x^{[6]}/[3])$ 1.18762,2.75791,-78.5781,0.9311,0.4498,-2.0136,0.2925	$[0]+([4]*(x-[5])*exp(-x*[7]))+(([1]+([2]/x^{1/2}))*exp(-x^{[6]}/[3]))$ -2.0058,3.5026,0.6842,7.1749,-1.2409,0.2863,0.00632,0.3425

TABLE A.3: Calibration functions and parameters after energy dependent calibrations of EH hadrons for UL 2016, UL 2017 & UL 2018

Year	H Barrel	H Endcap
	c(E_t)	c(E_t)
2016	$[0]+(([1]+[2]/x^{1/2})*exp(-x^{[6]}/[3]))-([4]*exp(-x^{[7]}/[5]))$ 7.40006,8.34692,-16.4046,6.11434,13.3222,0.684788,0.0159107,-0.584424	$[0]+(([1]+[2]/x^{1/2})*exp(-x^{[6]}/[3]))-([4]*exp(-x^{[7]}/[5]))$ 1.32375, 0.191687, -2.81052, 27.1729, 0.418262, 0.107711, 1.71039, -1.01447
2017	$[0]+(([1]+[2]/x^{1/2})*exp(-x^{[6]}/[3]))-([4]*exp(-x^{[7]}/[5]))$ 3.00232,3.96519,-10.7502,3.2574,4.6885,0.3356,0.03001,-0.6611	$[0]+(([1]+[2]/x^{1/2})*exp(-x^{[6]}/[3]))-([4]*exp(-x^{[7]}/[5]))$ 1.13985, 0.196757, -2.56203, 22.1335, 0.231671, 0.0625535, 1.71307, -0.9870
2018	$[0]+(([1]+[2]/x^{[8]})*exp(-x^{[6]}/[3]))-([4]*exp(-x^{[7]}/[5]))$ 2.17818, -0.127103, -10.6713, 1.23319, 1.23137, 0.163938, 0.159672, -0.939935, 0.766221	$[0]+(([1]+[2]/x^{1/2})*exp(-x^{[6]}/[3]))-([4]*exp(-x^{[7]}/[5]))$ 1.55955, 1.11623, -5.521, 2.67844, 0.921303, 0.0372015, 0.14929, -1.27774

TABLE A.4: Calibration functions and parameters after energy dependent calibrations of H hadrons for UL 2016, UL 2017 & UL 2018

Calibrations used for Eta dependent calibrations

Year	EH Barrel		EH endcap	
	$\alpha(E_t)$	$\beta(E_t)$	$\alpha(E_t)$	$\beta(E_t)$
2016	$[0]+[1]*exp(-x*[3]/[2])$ 340.572 ,-340.606, 6689.73, 0.00088	$[0]+[1]*exp(-x/[2])$ 0.05329, 0.1242, 207.676	$[0]+([1]*x^{[2]}*exp(-x))$ 0.001348, -26.4964, 0.2066	$[0]+[1]*x^{[3]}*exp(-x/[2])$ 0.0476, 0.000573,133.97, 0.7984
2017	$[0]+[1]*exp(-x*[3]/[2])$ 340.583, -340.595, 6192.42 ,0.000373	$[0]+[1]*exp(-x/[2])$ 0.05576, 0.124518, 204.727	$[0]+([1]*x^{[2]}*exp(-x))$ 0.00871, -7.9761, 0.8196	$[0]+[1]*exp(-x/[2])$ 0.0286, 0.0428, 263.774
2018	$[0]+[1]*exp(-x*[3]/[2])$ 57.1792, -46.337, -4.745, -0.00033	$[0]+[1]*exp(-x/[2])$ 0.05633, 0.127, 202.232	$[0]+([1]*x^{[2]}*exp(-x))$ 0.0128452, -87.2358, -1.48798	$[0]+([1]+[2]/x^{1/2})*exp(-x^{[4]}/[3])$ 0.0099, 0.781, -3.737, 2.35, 0.3827

TABLE A.5: Calibration functions and parameters after eta dependent calibrations of EH hadrons for UL 2016, UL 2017 & UL 2018

Year	H Barrel		H endcap	
	$\alpha(E_t)$	$\beta(E_t)$	$\alpha(E_t)$	$\beta(E_t)$
2016	$[0]+[1]*x$ -0.00157472, 0	$[0]+[1]*exp(-x/[2])$ -0.01546, 0.07283, 56.2748	$[0]+[1]*(x^{[3]})*exp(-x/[2])$ -0.006057, -0.7204, 0.1754,14.5679	$[0]+[1]*x*exp(-x/[2])$ 0.05642, 0.007259, 23.1278
2017	$[0]+[1]*x$ -0.006514, 0	$[0]+[1]*exp(-x/[2])$ 0.000578, -0.4919, 1.445	$[0]+[1]*x*exp(-x/[2])$ -0.002671, 0.03767, 42.5313	$[0]+[1]*x*exp(-x/[2])$ 0.02243, 0.1008, 85.7951
2018	$[0]*x^{[3]})+([1]*exp(-x/[2])$ -44.8973, -0.02624, 26.1048, -4.7351	$[0]+(([1]+[2]/x)*exp(-x/[3]))$ -0.01268, 0.05239, 0.2205, 62.9852	$[0]+[1]*x*exp(-x/[2])$ -0.00138, 0.0545, 34.3199	$[0]+[1]*x*exp(-x/[2])$ -0.0001,0,0

TABLE A.6: Calibration functions and parameters after eta-dependent calibrations of H hadrons for UL 2016, UL 2017 & UL 2018

A.2 Technical details for ECAL ZS sample

ECAL ZS sample has incorporated new ECAL PF threshold i.e. rechit threshold equivalent of 1 σ noise and it is referred as "new or 10_0_3". PF hadron calibration version-1, provided earlier to the ECAL ZS sample is referred as "old or 10_0_2". This ols sample corresponds to ECAL rechit thresholds of 80 MeV in Barrel & 300 MeV in Endcap. These both samples will have similar detector configuration but different ECAL noise rechit thresholds.

TABLE A.7: Technical details used for the production of 10_0_2 or 10_0_3 sample.

	10_0_2	10_0_3
CMSSW version	CMSSW_10_0_2	CMSSW_10_0_3
Global Tag	100X_upgrade2018_realistic_v10	100X_upgrade2018_realistic_v10
Particle ID	-211 (π^-)	-211 (π^-)
Energy range	2−500 GeV	2−500 GeV
$\lvert\eta\rvert$ & $\lvert\phi\rvert$ ranges	$\lvert\eta\rvert < 3, \lvert\phi\rvert < 3.14$	$\lvert\eta\rvert < 3, \lvert\phi\rvert < 3.14$

A.3 Technical details for Run3 sample

TABLE A.8: Technical details used for the production of Run3 sample.

	Run 3
CMSSW version	CMSSW_12_6_4
Global Tag	126X_mcRun3_2023_forPU65_v4
Particle ID	-211 (π^-)
Energy range	2−500 GeV
$\|\eta\|$ & $\|\phi\|$ ranges	$\|\eta\| < 3, \|\phi\| < 3.14$

Appendix B

Additional tables

B.1 Lost-lepton + τ_{had} predictions

TABLE B.1: Observed events yields in CR, average transfer factors (TFs) and predictions for lost-lepton background in high $\delta\phi$ region

Bins	$p_T^{miss}(GeV)$	Data CR	Avg. TF	Predictions.
$2 \leq N_{jets} \leq 4$, $N_{b-jets} = 0$	200 - 300	886	0.824 ± 0.0331	730 ± 38.2
	300 - 370	210	0.826 ± 0.0331	173 ± 13.8
	370 - 450	92	0.823 ± 0.0331	75.7 ± 8.46
	450 - 600	42	0.769 ± 0.0331	32.3 ± 5.17
	600 - 750	16	0.87 ± 0.0331	13.9 ± 3.52
	750 - 900	4	0.834 ± 0.0331	3.34 ± 1.67
	$\geq$900	2	0.862 ± 0.0331	1.72 ± 1.22
$5 \leq N_{jets} \leq 6$, $N_{b-jets} = 0$	200 - 300	116	0.583 ± 0.0333	67.7 ± 7.37
	300 - 370	26	0.583 ± 0.0333	15.2 ± 3.1
	370 - 450	11	0.583 ± 0.0333	6.42 ± 1.97
	450 - 600	11	0.583 ± 0.0333	6.42 ± 1.97
	600 - 750	2	0.583 ± 0.0333	1.17 ± 0.828
	$\geq$750	3	0.583 ± 0.0333	1.75 ± 1.02
$N_{jets} \geq 7$, $N_{b-jets} = 0$	200 - 300	15	0.465 ± 0.0489	6.98 ± 1.95
	300 - 370	8	0.465 ± 0.0489	3.72 ± 1.37
	370 - 450	2	0.465 ± 0.0489	0.931 ± 0.665
	450 - 600	2	0.465 ± 0.0489	0.931 ± 0.665
	$\geq$600	1	0.465 ± 0.0489	0.465 ± 0.468
$2 \leq N_{jets} \leq 4$, $N_{b-jets} > 0$	200 - 300	273	0.869 ± 0.0239	237 ± 15.8
	300 - 370	61	0.869 ± 0.0239	53 ± 6.94
	370 - 450	21	0.869 ± 0.0239	18.2 ± 4.01
	450 - 600	14	0.869 ± 0.0239	12.2 ± 3.27
	$\geq$600	3	0.869 ± 0.0239	2.61 ± 1.51
$5 \leq N_{jets} \leq 6$, $N_{b-jets} > 0$	200 - 300	112	0.686 ± 0.041	76.9 ± 8.59
	300 - 370	24	0.686 ± 0.041	16.5 ± 3.5
	370 - 450	11	0.686 ± 0.041	7.55 ± 2.32
	450 - 600	2	0.686 ± 0.041	1.37 ± 0.974
	$\geq$600	2	0.686 ± 0.041	1.37 ± 0.974
$N_{jets} \geq 7$, $N_{b-jets} > 0$	200 - 300	30	0.589 ± 0.0592	17.7 ± 3.68
	300 - 370	9	0.589 ± 0.0592	5.3 ± 1.85
	370 - 450	6	0.589 ± 0.0592	3.53 ± 1.49
	450 - 600	4	0.589 ± 0.0592	2.36 ± 1.2
	$\geq$600	1	0.589 ± 0.0592	0.589 ± 0.592
EW : W tag	200 - 300	164	1.04 ± 0.0483	171 ± 15.5
	300 - 370	45	1.04 ± 0.0483	46.8 ± 7.31
	370 - 450	18	1.04 ± 0.0483	18.7 ± 4.5
	450 - 600	11	1.04 ± 0.0483	11.4 ± 3.49
	600 - 750	6	1.04 ± 0.0483	6.24 ± 2.56
	$\geq$750	1	1.04 ± 0.0483	1.04 ± 1.04
EW : H tag	200 - 300	75	1 ± 0.0648	75.3 ± 9.96
	300 - 370	23	1 ± 0.0648	23.1 ± 5.04
	370 - 450	6	1 ± 0.0648	6.02 ± 2.49
	450 - 600	8	1 ± 0.0648	8.03 ± 2.89
	600 - 750	2	1 ± 0.0648	2.01 ± 1.43
	$\geq$750	1	1 ± 0.0648	1 ± 1.01

TABLE B.2: Observed events yields in CR, average transfer factors (TFs) and predictions for lost-lepton background in low $\delta\phi$ region

Bins	$p_T^{miss}(GeV)$	Data CR	Avg. TF	Pred.
$2 \leq N_{jets} \leq 4$, $N_{b-jets} = 0$	200 - 300	319	0.342 ± 0.0206	109 ± 8.97
	300 - 370	72	0.342 ± 0.0206	24.6 ± 3.26
	370 - 450	40	0.342 ± 0.0206	13.7 ± 2.32
	450 - 600	39	0.342 ± 0.0206	13.4 ± 2.28
	600 - 750	4	0.342 ± 0.0206	1.37 ± 0.69
	750 - 900	5	0.342 ± 0.0206	1.71 ± 0.772
	$\geq$900	1	0.342 ± 0.0206	0.342 ± 0.343
$5 \leq N_{jets} \leq 6$, $N_{b-jets} = 0$	200 - 300	26	0.356 ± 0.0407	9.26 ± 2.1
	300 - 370	16	0.356 ± 0.0407	5.7 ± 1.57
	370 - 450	2	0.356 ± 0.0407	0.713 ± 0.51
	450 - 600	1	0.356 ± 0.0407	0.356 ± 0.359
	600 - 750	0	0.356 ± 0.0407	0 ± 0
	$\geq$750	1	0.356 ± 0.0407	0.356 ± 0.359
$N_{jets} \geq 7$, $N_{b-jets} = 0$	200 - 300	8	0.372 ± 0.115	2.97 ± 1.4
	300 - 370	2	0.372 ± 0.115	0.743 ± 0.574
	370 - 450	4	0.372 ± 0.115	1.49 ± 0.875
	450 - 600	0	0.372 ± 0.115	0 ± 0
	$\geq$600	0	0.372 ± 0.115	0 ± 0
$2 \leq N_{jets} \leq 4$, $N_{b-jets} > 0$	200 - 300	101	0.631 ± 0.0375	63.7 ± 7.38
	300 - 370	20	0.631 ± 0.0375	12.6 ± 2.92
	370 - 450	12	0.631 ± 0.0375	7.57 ± 2.23
	450 - 600	8	0.631 ± 0.0375	5.04 ± 1.81
	$\geq$600	2	0.631 ± 0.0375	1.26 ± 0.895
$5 \leq N_{jets} \leq 6$, $N_{b-jets} > 0$	200 - 300	34	0.527 ± 0.0713	17.9 ± 3.91
	300 - 370	10	0.527 ± 0.0713	5.27 ± 1.81
	370 - 450	2	0.527 ± 0.0713	1.05 ± 0.759
	450 - 600	4	0.527 ± 0.0713	2.11 ± 1.09
	$\geq$600	0	0.527 ± 0.0713	0 ± 0
$N_{jets} \geq 7$, $N_{b-jets} > 0$	200 - 300	15	0.363 ± 0.0825	5.45 ± 1.87
	300 - 370	3	0.363 ± 0.0825	1.09 ± 0.676
	370 - 450	1	0.363 ± 0.0825	0.363 ± 0.372
	450 - 600	0	0.363 ± 0.0825	0 ± 0
	$\geq$600	0	0.363 ± 0.0825	0 ± 0
EW : W tag	200 - 300	46	0.412 ± 0.0355	19 ± 3.24
	300 - 370	19	0.412 ± 0.0355	7.84 ± 1.92
	370 - 450	9	0.412 ± 0.0355	3.71 ± 1.28
	450 - 600	5	0.412 ± 0.0355	2.06 ± 0.939
	600 - 750	2	0.412 ± 0.0355	0.825 ± 0.588
	$\geq$750	1	0.412 ± 0.0355	0.412 ± 0.414
EW : H tag	200 - 300	17	0.34 ± 0.0501	5.78 ± 1.64
	300 - 370	8	0.34 ± 0.0501	2.72 ± 1.04
	370 - 450	4	0.34 ± 0.0501	1.36 ± 0.709
	450 - 600	3	0.34 ± 0.0501	1.02 ± 0.608
	600 - 750	0	0.34 ± 0.0501	0 ± 0
	$\geq$750	0	0.34 ± 0.0501	0 ± 0

B.2 Electron faking photon (e $\rightarrow \gamma$) background predictions

TABLE B.3: Fake-rate parameterization derived from W/$t\bar{t}$ simulations for years 2016, 2017, 2018 and for full Run 2. Note that by convention, upper bin boundary value does not belong to the same bin.

p_T	Q_{mult}	2016	2017	2018	Full Run 2
$100-120$	$0-2$	0.058 ± 0.005	0.027 ± 0.004	0.023 ± 0.003	0.036 ± 0.002
	$2-4$	0.0082 ± 0.0007	0.0094 ± 0.0009	0.015 ± 0.001	0.011 ± 0.000
	$4-7$	0.0062 ± 0.0005	0.007 ± 0.0006	0.0096 ± 0.0008	0.008 ± 0.000
	≥ 7	0.0042 ± 0.0006	0.0039 ± 0.0003	0.0067 ± 0.0008	0.0053 ± 0.0004
$120-140$	$0-2$	0.051 ± 0.006	0.019 ± 0.003	0.042 ± 0.009	0.039 ± 0.004
	$2-4$	0.0089 ± 0.0011	0.0094 ± 0.0013	0.014 ± 0.001	0.011 ± 0.001
	$4-7$	0.0061 ± 0.0007	0.0061 ± 0.0005	0.011 ± 0.001	0.0085 ± 0.0006
	≥ 7	0.0056 ± 0.0009	0.0042 ± 0.0004	0.0059 ± 0.0007	0.0052 ± 0.0004
$140-160$	$0-2$	0.039 ± 0.006	0.023 ± 0.004	0.036 ± 0.008	0.033 ± 0.004
	$2-4$	0.0094 ± 0.0013	0.013 ± 0.002	0.015 ± 0.002	0.013 ± 0.001
	$4-7$	0.0054 ± 0.0007	0.007 ± 0.001	0.011 ± 0.001	0.0084 ± 0.0006
	≥ 7	0.004 ± 0.001	0.005 ± 0.001	0.0066 ± 0.0011	0.0057 ± 0.0007
$160-200$	$0-2$	0.052 ± 0.007	0.021 ± 0.004	0.032 ± 0.006	0.035 ± 0.003
	$2-4$	0.009 ± 0.001	0.012 ± 0.001	0.014 ± 0.002	0.012 ± 0.001
	$4-7$	0.008 ± 0.001	0.0079 ± 0.0007	0.009 ± 0.001	0.0086 ± 0.0005
	≥ 7	0.0039 ± 0.0008	0.005 ± 0.001	0.007 ± 0.001	0.0057 ± 0.0006
$200-240$	$0-2$	0.052 ± 0.009	0.028 ± 0.006	0.025 ± 0.005	0.034 ± 0.004
	$2-4$	0.0093 ± 0.0015	0.0086 ± 0.0012	0.016 ± 0.002	0.012 ± 0.001
	$4-7$	0.0047 ± 0.0008	0.0083 ± 0.0009	0.0089 ± 0.0013	0.0078 ± 0.0007
	≥ 7	0.0053 ± 0.0013	0.0044 ± 0.0007	0.0067 ± 0.0012	0.0056 ± 0.0007
$240-300$	$0-2$	0.034 ± 0.007	0.02 ± 0.01	0.036 ± 0.009	0.031 ± 0.005
	$2-4$	0.0082 ± 0.0014	0.0089 ± 0.0012	0.016 ± 0.003	0.012 ± 0.001
	$4-7$	0.0048 ± 0.0008	0.0071 ± 0.0008	0.0087 ± 0.0011	0.0074 ± 0.0006
	≥ 7	0.0055 ± 0.0016	0.0053 ± 0.0009	0.0083 ± 0.0036	0.0068 ± 0.0018
$300-450$	$0-2$	0.017 ± 0.004	0.023 ± 0.004	0.035 ± 0.006	0.026 ± 0.003
	$2-4$	0.0076 ± 0.001	0.0095 ± 0.001	0.015 ± 0.002	0.011 ± 0.001
	$4-7$	0.0053 ± 0.0007	0.0064 ± 0.0008	0.011 ± 0.002	0.0083 ± 0.0008
	≥ 7	0.004 ± 0.001	0.0059 ± 0.0009	0.0056 ± 0.0008	0.0055 ± 0.0005
$450-600$	$0-2$	0.047 ± 0.012	0.022 ± 0.006	0.025 ± 0.009	0.03 ± 0.01
	$2-4$	0.008 ± 0.002	0.014 ± 0.002	0.02 ± 0.00	0.015 ± 0.002
	$4-7$	0.008 ± 0.002	0.0055 ± 0.0008	0.0098 ± 0.0016	0.0081 ± 0.0009
	≥ 7	0.0017 ± 0.0009	0.0071 ± 0.0019	0.0036 ± 0.0009	0.0045 ± 0.0008
≥ 600	$0-2$	0.017 ± 0.009	0.022 ± 0.006	0.068 ± 0.026	0.039 ± 0.011
	$2-4$	0.0096 ± 0.0028	0.01 ± 0.00	0.02 ± 0.01	0.015 ± 0.002
	$4-7$	0.014 ± 0.007	0.007 ± 0.001	0.009 ± 0.002	0.009 ± 0.002
	≥ 7	0.0023 ± 0.001	0.0042 ± 0.0010	0.0055 ± 0.0019	0.0046 ± 0.0011

TABLE B.4: Observed single electron yields, the average scale factor due to MC-based fake rates and corresponding fake-rate predictions with statistical uncertainties for high-$\delta\phi$ search regions.

Bin	$p_T^{miss}(GeV)$	Obs. e events	Avg. tranfer factors	$N_{Pred.} \pm \Delta N_{Pred.}^{stat}$
$2 \leq N_{jets} \leq 4$, $N_{b-jets} = 0$	200 - 300	26591	0.0172 ± 0.00199	459 ± 53.1
	300 - 370	6629	0.0174 ± 0.0021	116 ± 14
	370 - 450	3363	0.0173 ± 0.00218	58.2 ± 7.4
	450 - 600	2098	0.0174 ± 0.00214	36.5 ± 4.57
	600 - 750	581	0.0179 ± 0.00214	10.4 ± 1.32
	750 - 900	189	0.017 ± 0.00208	3.21 ± 0.458
	$\geq$900	83	0.0204 ± 0.00358	1.69 ± 0.35
$5 \leq N_{jets} \leq 6$, $N_{b-jets} = 0$	200 - 300	2727	0.0157 ± 0.00188	42.7 ± 5.19
	300 - 370	912	0.0159 ± 0.00197	14.5 ± 1.86
	370 - 450	472	0.0156 ± 0.00203	7.38 ± 1.02
	450 - 600	344	0.0169 ± 0.00237	5.83 ± 0.874
	600 - 750	126	0.017 ± 0.00259	2.14 ± 0.377
	$\geq$750	40	0.019 ± 0.00343	0.76 ± 0.182
$N_{jets} \geq 7$, $N_{b-jets} = 0$	200 - 300	582	0.0141 ± 0.00161	8.19 ± 0.997
	300 - 370	156	0.0145 ± 0.00175	2.26 ± 0.328
	370 - 450	116	0.0147 ± 0.00173	1.71 ± 0.255
	450 - 600	78	0.0154 ± 0.00247	1.2 ± 0.236
	$\geq$600	35	0.0193 ± 0.00275	0.676 ± 0.149
$2 \leq N_{jets} \leq 4$, $N_{b-jets} > 0$	200 - 300	6940	0.02 ± 0.00223	139 ± 15.5
	300 - 370	1512	0.0201 ± 0.00228	30.3 ± 3.54
	370 - 450	699	0.0201 ± 0.00225	14 ± 1.66
	450 - 600	385	0.0193 ± 0.00249	7.43 ± 1.03
	$\geq$600	113	0.0202 ± 0.00279	2.28 ± 0.381
$5 \leq N_{jets} \leq 6$, $N_{b-jets} > 0$	200 - 300	3421	0.0182 ± 0.00206	62.4 ± 7.12
	300 - 370	922	0.0181 ± 0.00208	16.7 ± 2
	370 - 450	396	0.0184 ± 0.00205	7.29 ± 0.89
	450 - 600	231	0.0183 ± 0.00227	4.23 ± 0.594
	$\geq$600	74	0.017 ± 0.00282	1.26 ± 0.254
$N_{jets} \geq 7$, $N_{b-jets} > 0$	200 - 300	1435	0.0169 ± 0.00189	24.2 ± 2.78
	300 - 370	423	0.0168 ± 0.00212	7.12 ± 0.96
	370 - 450	198	0.0166 ± 0.00192	3.3 ± 0.447
	450 - 600	114	0.0177 ± 0.00213	2.02 ± 0.308
	$\geq$600	32	0.0176 ± 0.00369	0.563 ± 0.154
EW : W tag	200 - 300	9713	0.0179 ± 0.00196	174 ± 19.1
	300 - 370	1958	0.0178 ± 0.00205	34.8 ± 4.08
	370 - 450	775	0.0167 ± 0.00181	13 ± 1.48
	450 - 600	346	0.0174 ± 0.00202	6.01 ± 0.769
	600 - 750	74	0.0184 ± 0.00231	1.36 ± 0.233
	$\geq$750	40	0.0163 ± 0.00175	0.654 ± 0.125
EW : H tag	200 - 300	4856	0.0182 ± 0.00206	88.2 ± 10.1
	300 - 370	1341	0.018 ± 0.00215	24.2 ± 2.96
	370 - 450	571	0.0184 ± 0.00212	10.5 ± 1.29
	450 - 600	267	0.0172 ± 0.00184	4.6 ± 0.565
	600 - 750	57	0.0177 ± 0.00241	1.01 ± 0.191
	$\geq$750	25	0.0194 ± 0.00217	0.485 ± 0.111

TABLE B.5: Observed single electron yields, the average scale factor due to MC- based fake rates and corresponding fake-rate predictions with statistical uncertainties for low-$\delta\phi$ control regions.

Bins	$p_T^{miss}(GeV)$	Obs. e events	Avg. tranfer factors	$N_{Pred.} \pm \Delta N_{Pred.}^{stat}$
$2 \leq N_{jets} \leq 4$, $N_{b-jets} = 0$	200 - 300	2451	0.0175 ± 0.00585	42.9 ± 14.4
	300 - 370	611	0.0161 ± 0.00548	9.81 ± 3.37
	370 - 450	272	0.018 ± 0.00608	4.88 ± 1.68
	450 - 600	204	0.0181 ± 0.00647	3.7 ± 1.34
	600 - 750	68	0.0182 ± 0.00656	1.24 ± 0.471
	750 - 900	19	0.0205 ± 0.0069	0.389 ± 0.159
	$\geq$900	10	0.0136 ± 0.00342	0.136 ± 0.0549
$5 \leq N_{jets} \leq 6$, $N_{b-jets} = 0$	200 - 300	334	0.0149 ± 0.00489	4.99 ± 1.65
	300 - 370	104	0.0146 ± 0.00453	1.52 ± 0.494
	370 - 450	51	0.0138 ± 0.00367	0.703 ± 0.211
	450 - 600	29	0.0197 ± 0.00698	0.571 ± 0.228
	600 - 750	13	0.0167 ± 0.00719	0.218 ± 0.111
	$\geq$750	2	0.00817 ± 0.0029	0.0163 ± 0.0129
$N_{jets} \geq 7$, $N_{b-jets} = 0$	200 - 300	57	0.0151 ± 0.00465	0.863 ± 0.288
	300 - 370	18	0.0111 ± 0.00435	0.2 ± 0.0914
	370 - 450	15	0.0137 ± 0.00369	0.205 ± 0.0765
	450 - 600	4	0.0237 ± 0.0171	0.0949 ± 0.0834
	$\geq$600	4	0.0126 ± 0.004	0.0505 ± 0.0299
$2 \leq N_{jets} \leq 4$, $N_{b-jets} > 0$	200 - 300	695	0.0169 ± 0.00562	11.8 ± 3.93
	300 - 370	174	0.0172 ± 0.00574	2.98 ± 1.02
	370 - 450	86	0.0166 ± 0.00618	1.43 ± 0.553
	450 - 600	53	0.0191 ± 0.00827	1.01 ± 0.46
	$\geq$600	27	0.0199 ± 0.00712	0.538 ± 0.218
$5 \leq N_{jets} \leq 6$, $N_{b-jets} > 0$	200 - 300	423	0.0155 ± 0.00454	6.56 ± 1.95
	300 - 370	108	0.013 ± 0.00349	1.41 ± 0.401
	370 - 450	59	0.0132 ± 0.00428	0.777 ± 0.272
	450 - 600	30	0.0113 ± 0.00384	0.34 ± 0.131
	$\geq$600	10	0.014 ± 0.00319	0.14 ± 0.0545
$N_{jets} \geq 7$, $N_{b-jets} > 0$	200 - 300	172	0.014 ± 0.00446	2.4 ± 0.788
	300 - 370	38	0.016 ± 0.00511	0.61 ± 0.218
	370 - 450	28	0.0128 ± 0.00392	0.359 ± 0.129
	450 - 600	10	0.0156 ± 0.00444	0.156 ± 0.0663
	$\geq$600	6	0.00987 ± 0.00256	0.0592 ± 0.0287
EW : W tag	200 - 300	666	0.0171 ± 0.00553	11.4 ± 3.71
	300 - 370	172	0.0162 ± 0.00507	2.78 ± 0.897
	370 - 450	87	0.0167 ± 0.00613	1.45 ± 0.556
	450 - 600	53	0.0175 ± 0.00605	0.93 ± 0.345
	600 - 750	13	0.0159 ± 0.00397	0.207 ± 0.0771
	$\geq$750	7	0.0158 ± 0.00604	0.111 ± 0.0595
EW : H tag	200 - 300	261	0.0167 ± 0.00578	4.35 ± 1.53
	300 - 370	81	0.0173 ± 0.0067	1.4 ± 0.564
	370 - 450	33	0.0203 ± 0.0085	0.67 ± 0.304
	450 - 600	43	0.0208 ± 0.00803	0.895 ± 0.371
	600 - 750	4	0.0117 ± 0.00291	0.047 ± 0.0262
	$\geq$750	3	0.0144 ± 0.00287	0.0433 ± 0.0264

B.3 Predictions for Irreducible background from Z($\nu\nu$)+γ

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

Bins	$p_T^{miss}(GeV)$	MC Yield	$N_{Pred.} \pm \Delta N_{Pred.}^{stat}$
$2 \leq N_{jets} \leq 4$, $N_{b-jets} = 0$	200 - 300	680 ± 18.9	730 ± 64.6
	300 - 370	246 ± 11.5	265 ± 25.4
	370 - 450	133 ± 8.46	143 ± 15.1
	450 - 600	97.5 ± 7.37	105 ± 11.8
	600 - 750	19.9 ± 3.23	21.3 ± 3.9
	750 - 900	10.9 ± 2.56	11.8 ± 2.92
	$\geq$900	2.79 ± 1.16	3 ± 1.27
$5 \leq N_{jets} \leq 6$, $N_{b-jets} = 0$	200 - 300	8.34 ± 1.91	8.96 ± 2.18
	300 - 370	4.29 ± 1.5	4.6 ± 1.66
	370 - 450	4.18 ± 1.41	4.49 ± 1.56
	450 - 600	2.7 ± 1.16	2.9 ± 1.27
	600 - 750	0.774 ± 0.619	0.831 ± 0.668
	$\geq$750	0.702 ± 0.697	0.754 ± 0.751
$N_{jets} \geq 7$, $N_{b-jets} = 0$	200 - 300	1.64 ± 0.88	1.77 ± 0.957
	300 - 370	0.667 ± 0.532	0.717 ± 0.575
	370 - 450	0.186 ± 0.0976	0.2 ± 0.106
	450 - 600	0.0107 ± 0.00988	0.0115 ± 0.0107
	$\geq$600	0.0252 ± 0.0252	0.0271 ± 0.0271
$2 \leq N_{jets} \leq 4$, $N_{b-jets} > 0$	200 - 300	56.6 ± 5.36	57 ± 16.8
	300 - 370	19 ± 3.06	19.1 ± 6.15
	370 - 450	8.29 ± 2.09	8.35 ± 3.13
	450 - 600	7.74 ± 2.07	7.8 ± 3.01
	$\geq$600	2.98 ± 1.23	3 ± 1.49
$5 \leq N_{jets} \leq 6$, $N_{b-jets} > 0$	200 - 300	1.54 ± 0.749	1.55 ± 0.869
	300 - 370	1.25 ± 0.652	1.26 ± 0.745
	370 - 450	0.668 ± 0.551	0.672 ± 0.585
	450 - 600	0.0766 ± 0.0448	0.0772 ± 0.05
	$\geq$600	0.0583 ± 0.0392	0.0587 ± 0.0428
$N_{jets} \geq 7$, $N_{b-jets} > 0$	200 - 300	0.0762 ± 0.0424	0.0767 ± 0.0478
	300 - 370	0.692 ± 0.585	0.697 ± 0.62
	370 - 450	0.0596 ± 0.0432	0.06 ± 0.0466
	450 - 600	0 ± 0	0 ± 0
	$\geq$600	0.0268 ± 0.0239	0.027 ± 0.0252
EW : W tag	200 - 300	36.7 ± 4.12	39.2 ± 6.05
	300 - 370	10.9 ± 2.23	11.6 ± 2.67
	370 - 450	9.14 ± 2.12	9.8 ± 2.45
	450 - 600	7.93 ± 2	8.48 ± 2.33
	600 - 750	2.68 ± 1.15	2.88 ± 1.26
	$\geq$750	1.66 ± 0.894	1.78 ± 0.973
EW : H tag	200 - 300	13.2 ± 2.54	14.1 ± 3.12
	300 - 370	5.34 ± 1.58	5.73 ± 1.78
	370 - 450	4.93 ± 1.58	5.28 ± 1.77
	450 - 600	3.79 ± 1.35	4.02 ± 1.54
	600 - 750	2.39 ± 1.21	2.56 ± 1.32
	$\geq$750	1.11 ± 0.791	1.19 ± 0.855

TABLE B.7: Z($\nu\nu$) + γ MC yields and prediction for low-$\delta\phi$ control regions.

Bins	$p_T^{miss}(GeV)$	MC Yield	$N_{Pred.} \pm \Delta N_{Pred.}^{stat}$
$2 \leq N_{jets} \leq 4$, $N_{b-jets} = 0$	200 - 300	99.6 ± 7.17	107 ± 11.8
	300 - 370	32.1 ± 4.06	34.5 ± 5.23
	370 - 450	10.6 ± 2.41	11.4 ± 2.76
	450 - 600	7.26 ± 1.93	7.8 ± 2.17
	600 - 750	2.76 ± 1.2	2.96 ± 1.32
	750 - 900	1.33 ± 0.882	1.43 ± 0.955
	$\geq$900	0.00549 ± 0.00241	0.0059 ± 0.00264
$5 \leq N_{jets} \leq 6$, $N_{b-jets} = 0$	200 - 300	2.18 ± 0.732	2.34 ± 0.81
	300 - 370	0.969 ± 0.627	1.04 ± 0.679
	370 - 450	0.781 ± 0.546	0.839 ± 0.591
	450 - 600	0.118 ± 0.0607	0.127 ± 0.0661
	600 - 750	0.0952 ± 0.0696	0.102 ± 0.0752
	$\geq$750	0.0337 ± 0.0261	0.0362 ± 0.0282
$N_{jets} \geq 7$, $N_{b-jets} = 0$	200 - 300	0.157 ± 0.0838	0.169 ± 0.0912
	300 - 370	0.0284 ± 0.0268	0.0305 ± 0.0289
	370 - 450	0.00394 ± 0.00257	0.00424 ± 0.00278
	450 - 600	0.0476 ± 0.046	0.0511 ± 0.0496
	$\geq$600	0.00375 ± 0.00221	0.00403 ± 0.00239
$2 \leq N_{jets} \leq 4$, $N_{b-jets} > 0$	200 - 300	11.1 ± 2.5	11.1 ± 4
	300 - 370	1.12 ± 0.658	1.13 ± 0.733
	370 - 450	0.197 ± 0.0968	0.198 ± 0.112
	450 - 600	0.973 ± 0.716	0.98 ± 0.771
	$\geq$600	0.0337 ± 0.0295	0.0339 ± 0.0311
$5 \leq N_{jets} \leq 6$, $N_{b-jets} > 0$	200 - 300	0.293 ± 0.106	0.295 ± 0.135
	300 - 370	0.0276 ± 0.0259	0.0278 ± 0.0272
	370 - 450	0.0489 ± 0.0322	0.0492 ± 0.0352
	450 - 600	0.0218 ± 0.0206	0.022 ± 0.0216
	$\geq$600	0.00227 ± 0.00133	0.00229 ± 0.00148
$N_{jets} \geq 7$, $N_{b-jets} > 0$	200 - 300	0.0348 ± 0.034	0.035 ± 0.0356
	300 - 370	0 ± 0	0 ± 0
	370 - 450	0.000795 ± 0.000795	0.000801 ± 0.000831
	450 - 600	0.0335 ± 0.033	0.0338 ± 0.0346
	$\geq$600	0 ± 0	0 ± 0
EW : W tag	200 - 300	3.5 ± 1.25	3.76 ± 1.39
	300 - 370	1.2 ± 0.678	1.29 ± 0.736
	370 - 450	0.662 ± 0.585	0.71 ± 0.631
	450 - 600	0.311 ± 0.162	0.331 ± 0.177
	600 - 750	0.0282 ± 0.0275	0.0303 ± 0.0296
	$\geq$750	0.00874 ± 0.00774	0.00939 ± 0.00835
EW : H tag	200 - 300	1.07 ± 0.506	1.11 ± 0.57
	300 - 370	0.202 ± 0.0991	0.217 ± 0.108
	370 - 450	0.144 ± 0.0839	0.155 ± 0.0911
	450 - 600	0.128 ± 0.0815	0.138 ± 0.0883
	600 - 750	0.00154 ± 0.00154	0.00165 ± 0.00166
	$\geq$750	0.0125 ± 0.0102	0.0133 ± 0.011

B.4 Predictions for QCD+γ + *jets* background

TABLE B.8: $\gamma + jets$ + QCD background prediction in high–$\Delta\phi$ & high–MET($\geq$300 GeV) search region using low–$\Delta\phi$ and high–$\Delta\phi$ & low–MET (200–300 GeV) control region.

Bin	MET	low–$\Delta\phi$			high–$\Delta\phi$ & low-MET (200–300 GeV)			high–$\Delta\phi$ & hig
	[GeV]	Data (Raw)	EWK (Real MET)	γ+jet (fake MET)	Data (Raw)	EWK (Real MET)	γ+jet (fake MET)	Predictio
$2 \leq N_{jets} \leq 4$, $N_{b-jets} = 0$								
1	200–300	2309	259.8 ± 20.13	2049 ± 52.1	2865	1918.92 ± 91.9	946.1 ± 111.7	–
2	300–370	286	68.87 ± 6.938	217.1 ± 18.28	–	–	–	71.53 ± 27
3	370–450	106	29.77 ± 3.914	76.23 ± 11.01	–	–	–	25.11 ± 10
4	450–600	61	24.61 ± 3.377	36.39 ± 8.509	–	–	–	11.99 ± 5.3
5	600–750	10	5.486 ± 1.535	4.514 ± 3.515	–	–	–	1.487 ± 1.2
6	750–900	1	3.404 ± 1.203	0 ± 1.565	–	–	–	0 ± 0
7	$\geq$900	2	0.5085 ± 0.3712	1.491 ± 1.462	–	–	–	0.4914 ± 0.5
$5 \leq N_{jets} \leq 6$, $N_{b-jets} = 0$								
8	200–300	227	16.59 ± 2.756	210.4 ± 15.32	208	119.36 ± 9.27	88.64 ± 19.4	–
9	300–370	19	8.262 ± 1.772	10.74 ± 4.705	–	–	–	4.173 ± 2.2
10	370–450	12	2.255 ± 0.8059	9.745 ± 3.557	–	–	–	3.787 ± 1.8
11	450–600	6	1.054 ± 0.4262	4.946 ± 2.486	–	–	–	1.922 ± 1.1
12	600–750	1	0.32 ± 0.132	0.68 ± 1.009	–	–	–	$0.2643 \pm 0.$
13	$\geq$750	3	0.4088 ± 0.36	2.591 ± 1.769	–	–	–	1.007 ± 0.7
≥ 7 , $N_{b-jets} = 0$								
14	200–300	27	4.005 ± 1.429	23 ± 5.389	29	16.93 ± 2.39	12.07 ± 7.44	–
15	300–370	1	0.9737 ± 0.5815	0.02627 ± 1.157	–	–	–	0.007292 ± 0
16	370–450	3	1.695 ± 0.878	1.305 ± 1.942	–	–	–	0.3623 ± 0.6
17	450–600	2	0.146 ± 0.09601	1.854 ± 1.417	–	–	–	0.5147 ± 0.5
18	$\geq$600	0	0.05452 ± 0.02967	0 ± 0.02967	–	–	–	0 ± 0
$2 \leq N_{jets} \leq 4$, $N_{b-jets} > 0$								
19	200–300	648	86.6 ± 9.179	561.4 ± 27.06	674	432.78 ± 27.77	241.2 ± 97.06	–
20	300–370	86	16.73 ± 3.165	69.27 ± 9.799	–	–	–	14.36 ± 7.4
21	370–450	31	9.194 ± 2.296	21.81 ± 6.023	–	–	–	4.52 ± 2.5
22	450–600	16	7.038 ± 2.017	8.962 ± 4.48	–	–	–	1.858 ± 1.3
23	$\geq$600	3	1.833 ± 0.9211	1.167 ± 1.962	–	–	–	$0.2419 \pm 0.$
$5 \leq N_{jets} \leq 6$, $N_{b-jets} > 0$								
24	200–300	111	24.77 ± 4.331	86.23 ± 11.39	194	140.79 ± 11.61	53.21 ± 23.03	–
25	300–370	22	6.705 ± 1.851	15.29 ± 5.042	–	–	–	7.572 ± 4.6
26	370–450	5	1.88 ± 0.8034	3.12 ± 2.376	–	–	–	1.544 ± 1.4
27	450–600	2	2.469 ± 1.099	0 ± 1.791	–	–	–	0 ± 0
28	$\geq$600	3	0.1422 ± 0.05304	2.858 ± 1.733	–	–	–	1.415 ± 1.1
≥ 7 , $N_{b-jets} > 0$								
29	200–300	26	7.887 ± 2.019	18.11 ± 5.484	64	41.95 ± 4.77	22.05 ± 15.41	–
30	300–370	5	1.699 ± 0.7073	3.301 ± 2.345	–	–	–	1.482 ± 1.6
31	370–450	2	0.7225 ± 0.3928	1.277 ± 1.468	–	–	–	0.5735 ± 0.8
32	450–600	1	0.1894 ± 0.07318	0.8106 ± 1.003	–	–	–	0.3639 ± 0.5
33	$\geq$600	0	0.05923 ± 0.02829	0 ± 0.02829	–	–	–	0 ± 0
EW : W tag								
34	200–300	312	34.12 ± 4.908	277.9 ± 18.33	564	383.74 ± 27.08	180.3 ± 67.37	–
35	300–370	36	11.9 ± 2.221	24.1 ± 6.398	–	–	–	8.568 ± 5.2
36	370–450	16	5.872 ± 1.523	10.13 ± 4.28	–	–	–	3.601 ± 2.4
37	450–600	6	3.323 ± 1.013	2.677 ± 2.651	–	–	–	$0.9518 \pm 1.$
38	600–750	2	1.062 ± 0.5932	0.9382 ± 1.534	–	–	–	0.3336 ± 0.5
39	$\geq$750	0	0.5327 ± 0.4182	0 ± 0.4182	–	–	–	0 ± 0
EW : H tag								
40	200–300	142	11.24 ± 2.221	130.8 ± 12.12	245	177.52 ± 14.52	67.48 ± 27.87	–
41	300–370	23	4.341 ± 1.178	18.66 ± 4.938	–	–	–	6.755 ± 4.5
42	370–450	10	2.185 ± 0.771	7.815 ± 3.255	–	–	–	2.829 ± 2.0
43	450–600	2	2.053 ± 0.7147	1e-14 ± 1.585	–	–	–	0 ± 0
44	600–750	2	0.04864 ± 0.02535	1.951 ± 1.414	–	–	–	0.7064 ± 0.6
45	$\geq$750	0	0.05652 ± 0.02832	0 ± 0.02832	–	–	–	0 ± 0

Appendix C

Additional systematics for Trigger efficiency

We have used the parameterization function as p_T^{miss} trigger effeciency on simulated samples, calculated for different year p_T^{miss} datasets and the corresponding p_T^{miss} triggers as mentioned in table 5.5 & 5.7. There is a need to check if there any deviation between trigger efficiency calculated using parameterization function and MET data set. Here are the following checks:

- Since these efficiencies will be applied to MC samples, so there can also be a dependence of different topology of samples. Any variations in topology by calculating MC trigger efficiencies in WGJets & ttGJets separately is also investigated as shown in figure and around 1-3% relative difference is observed.
- Comparison between trigger efficiency directly from data and parameterization function. We have calculated trigger efficiency for following MET SR bins = [200, 300, 370, 450, 600, $\geq$750] GeV in case of data and parameterization function and compared both for separate years as shown in figure C.1. We found that any difference in trigger efficiency is under 1-4%.
- Variations in trigger efficiency due to different topology in search bins. In current search, optimised bins are in diffent N_{b-tags} = [0,>0] & N_{jets} = [2-4, 5-6, $\geq$7], so we have calculated the MET trigger efficiency using MC background simulations for different N_{b-tags} & N_{jets} combinations in same MET SR bins and compared those as shown in figure C.3, C.2. Here, we found that in case of N_{b-tags} >0, the variation in lower side-band of MET is under 10% and otherwise it will be under 5%.

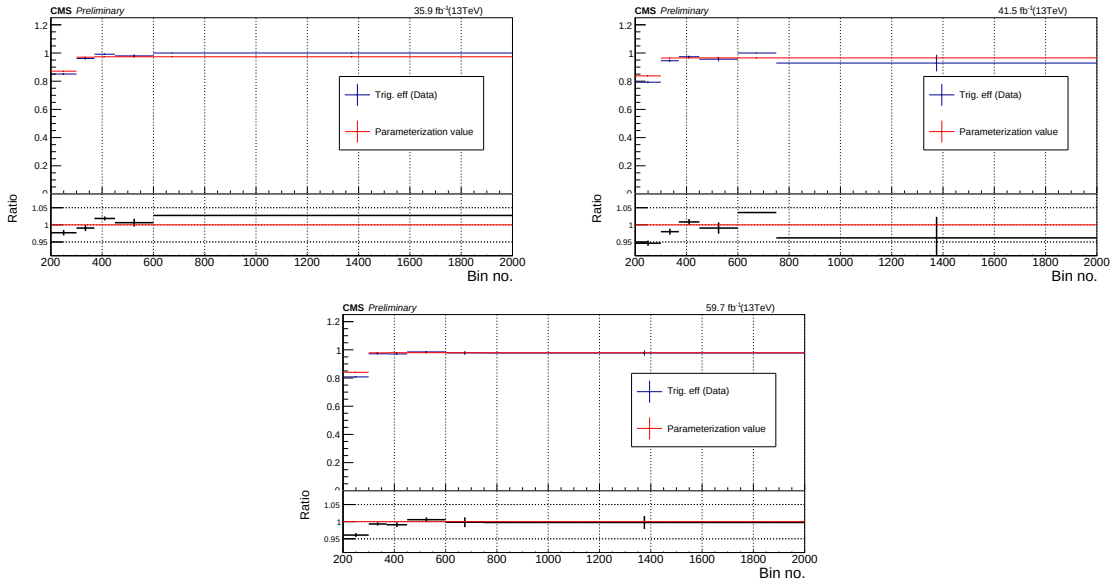

FIGURE C.1: Comparison of MET trigger efficiency from data and parameterization function

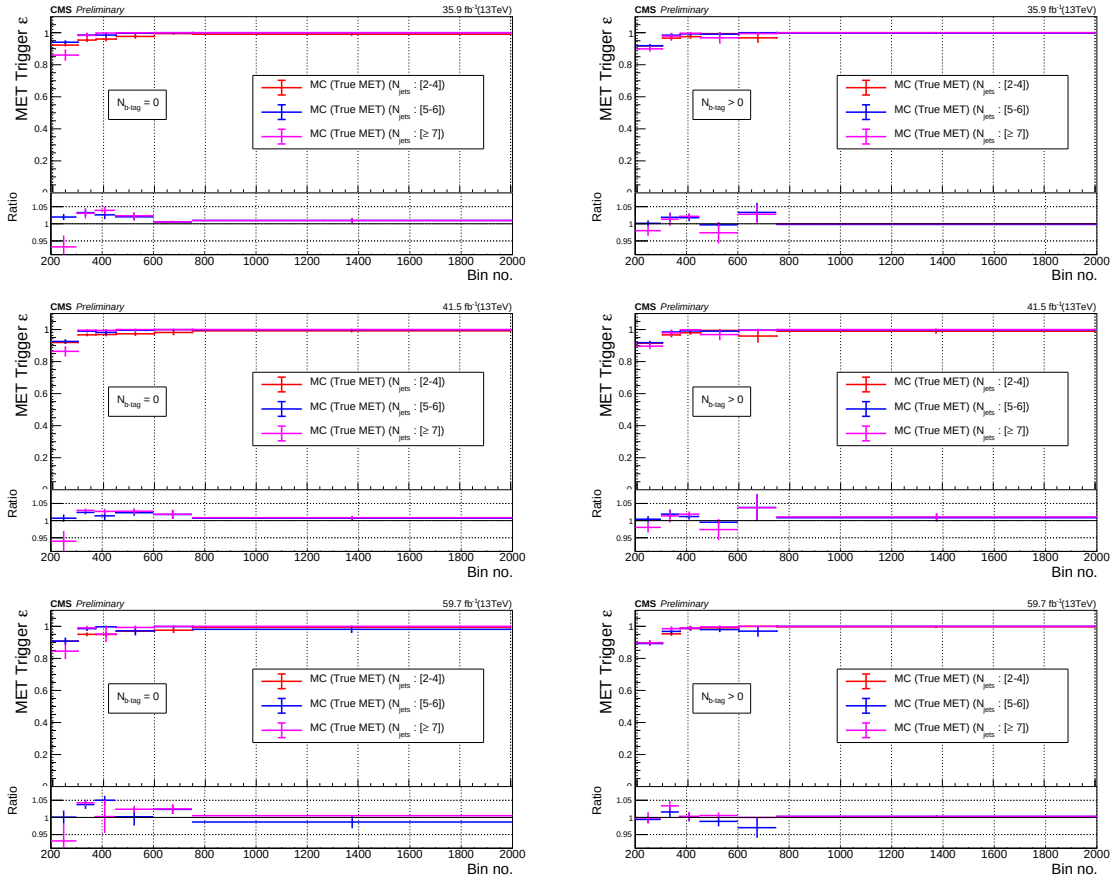

FIGURE C.2: MET trigger efficiency for different N_{b-tags} & N_{jets} combinations in same MET SR bins using MC bacground samples as a sum of W/tt/Z (source of true MET) only but not QCD & GJets (Fake MET).

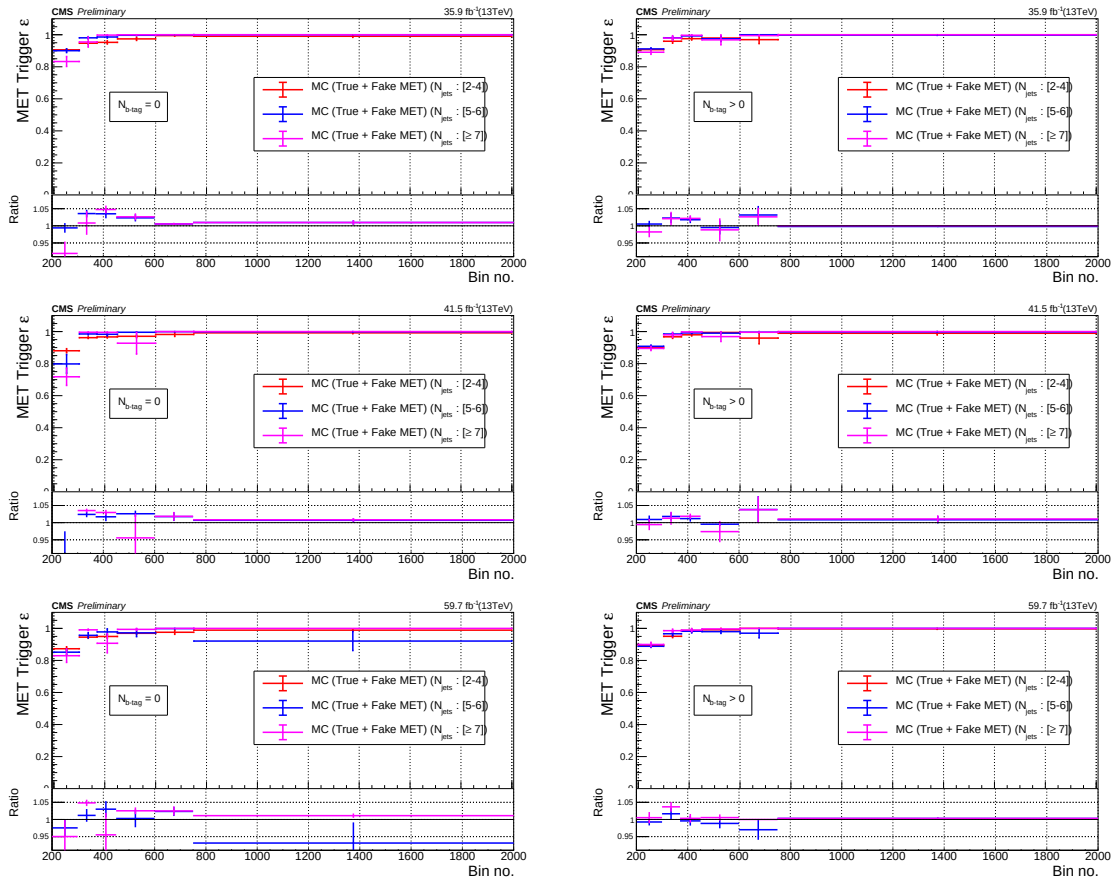

FIGURE C.3: MET trigger efficiency for different N_{b-tags} & N_{jets} combinations in same MET SR bins using MC bacground samples as a sum of W/tt/Z (source of true MET) + QCD & GJets (Fake MET).

Appendix D

Additional checks for fake-rate data-MC correction factors

The misid rate is determined by taking the ratio of the integral of events of type (e^{tag} , γ^{probe}) to the integral of events of type (e^{tag} , e^{probe}) after Z-mass fitting in bins of the $N_{\text{b-tags}}$ =0,≥1. The ratio of the misid. determined in this way in data to that in MC is considered as a correction factor for the entire Run 2 as shown in table 6.4. These correction factors are described in the section 6.2.3 of chapter 6. Correction factors are calculated using different variables (which we are calling α') to check the validity of $\alpha(N_{\text{b-tags}})$ as shown in figure 6.14.

D.1 Validity of $\alpha(N_{\text{b-tags}})$ yearwise

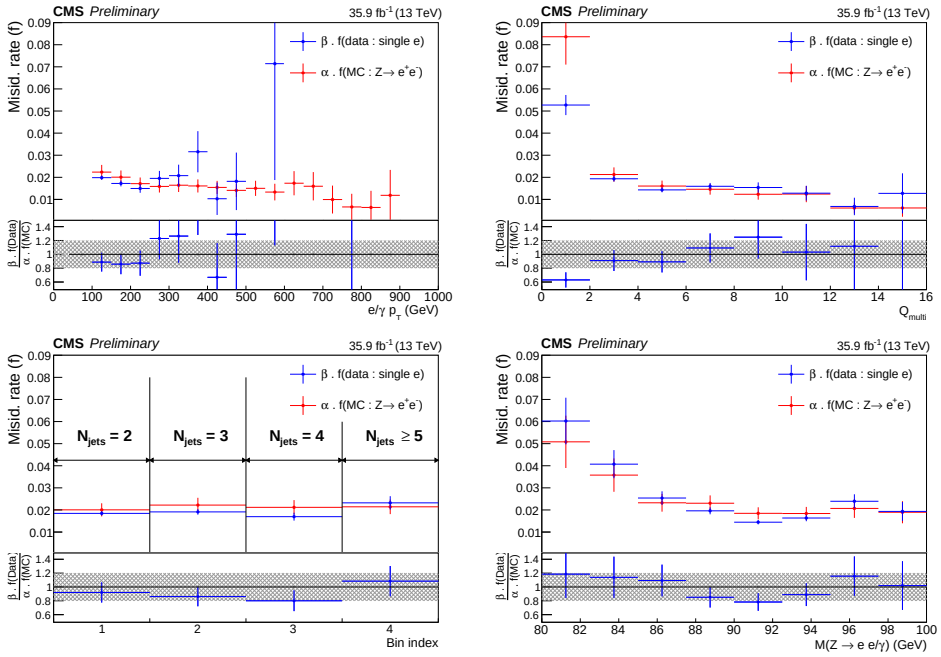

FIGURE D.1: 2016: f(data) and f(MC) (after re-weighted by α) distributions of events containing e^+e^- pairs using data and simulation respectively in the parameterization of p_T (on the top-left), Q_{multi} (on the top-right), N_{jets} (on the bottom-left) & invariant M_Z. In the bottom panel, the ratio f(data) and re-weighted f(MC) are shown, and the shaded region is the systematics related to α.

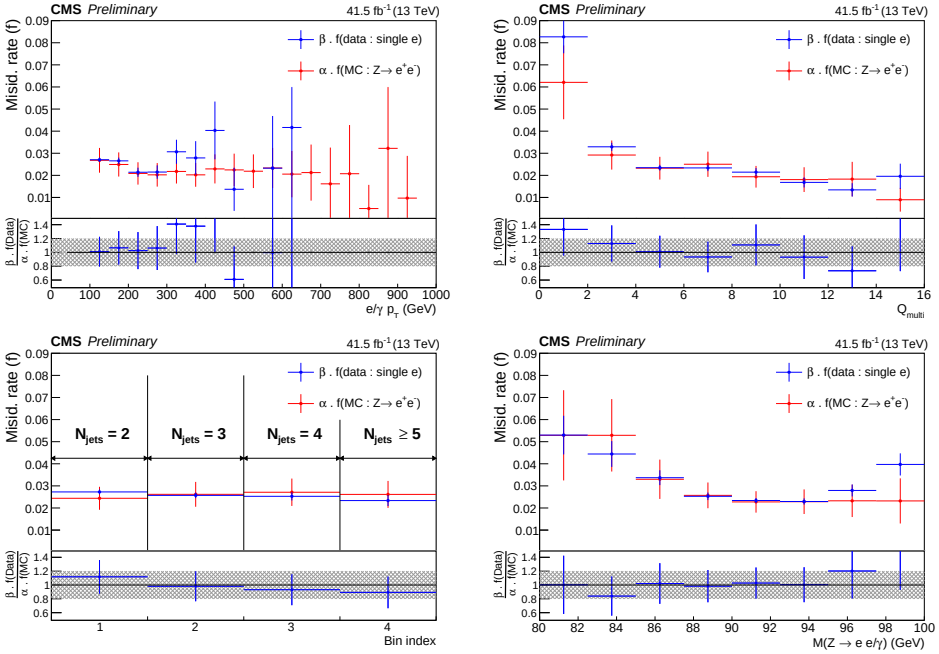

FIGURE D.2: 2017: f(data) and f(MC) (after re-weighted by α) distributions of events containing e^+e^- pairs using data and simulation respectively in the parameterization of p_T (on the top-left), Q_{multi} (on the top-right), N_{jets} (on the bottom-left) & invariant M_Z. In the bottom panel, the ratio f(data) and re-weighted f(MC) are shown, and the shaded region is the systematics related to α.

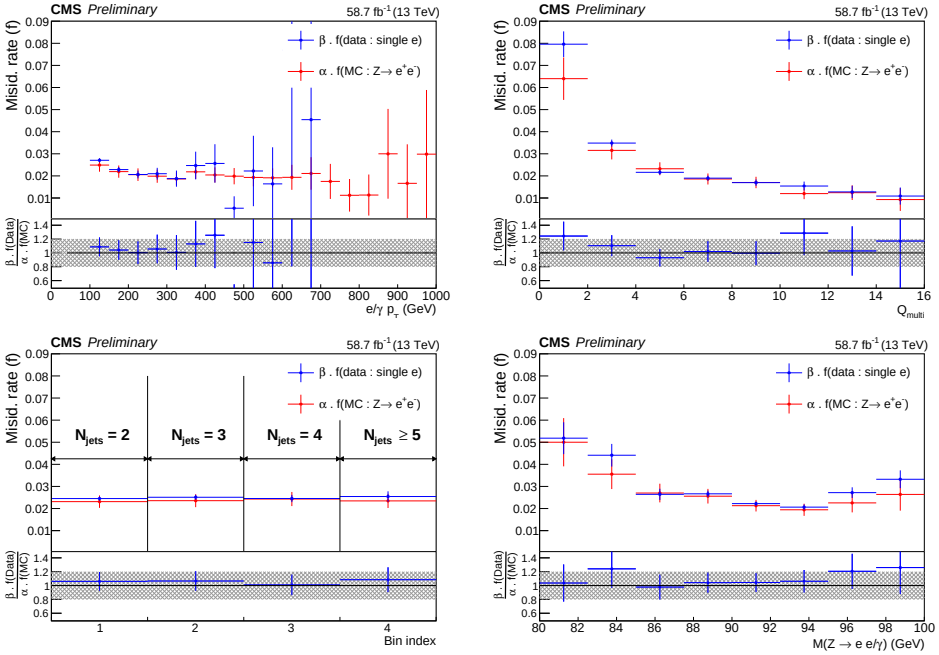

FIGURE D.3: 2018: f(data) and f(MC) (after re-weighted by α) distributions of events containing e^+e^- pairs using data and simulation respectively in the parameterization of p_T (on the top-left), Q_{multi} (on the top-right), N_{jets} (on the bottom-left) & invariant M_Z. In the bottom panel, the ratio f(data) and re-weighted f(MC) are shown, and the shaded region is the systematics related to α.

Appendix E

Appendix E : Different branching fractions for signal samples

For strong production, we have considered different BRs for the signal NLSP ($\tilde{\chi}_1^0$) decays as shown in figure E.1. In current scenario, $\tilde{\chi}_1^0 \rightarrow \tilde{G} + \gamma$ with 50% branching fraction and other $\tilde{\chi}_1^0 \rightarrow \tilde{G} + Z$ with 50% branching fraction but if the branching fraction is changed such as 25% for $\tilde{\chi}_1^0 \rightarrow \tilde{G} + \gamma$ and 75% for $\tilde{\chi}_1^0 \rightarrow \tilde{G} + Z$, then expected limits at 95% CL will change with respect to current scenario.

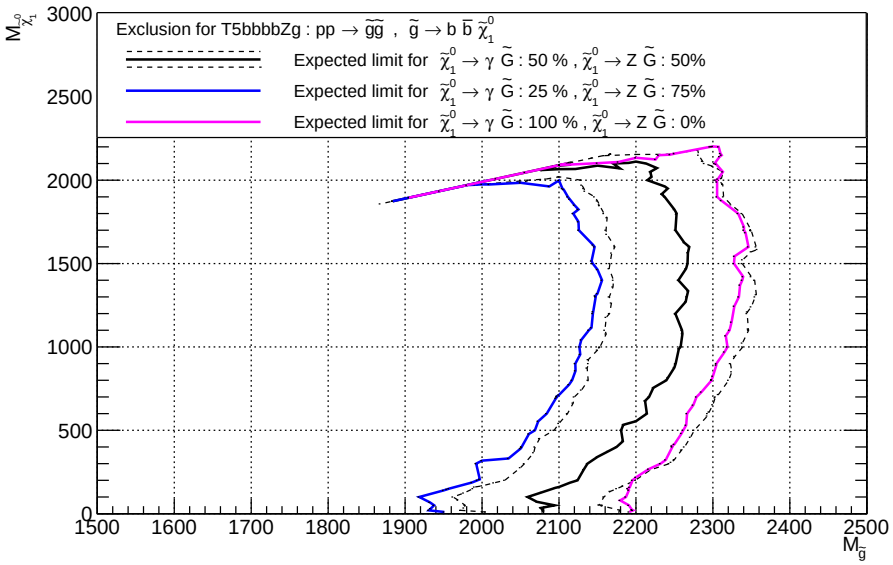

FIGURE E.1: Expected exclusions for different branching fractions of T5bbbbZg model at 95% confidence level reached by current search. Exclusion curve (in black color) represents mass limits for current scenario, when NLSP $\rightarrow \tilde{G} + \gamma$ or $\tilde{G} + Z$ with 50% BR. Blue curve represents the case when NLSP $\rightarrow \tilde{G} + \gamma$ with 25% BR and magenta curve represents the case when NLSP $\rightarrow \tilde{G} + \gamma$ with 100% BR

Similarly, in case of electroweakino pair production, we can choose different branching fractions for neutralinos or NLSP decays and compare the expected linit at 95% CL. In current TChiNG signal sample, branching fraction for $\tilde{\chi}_2^0 \rightarrow \tilde{G} + \gamma$ is assumed to be 50% and $\tilde{\chi}_1^0 \rightarrow \tilde{G} + Z(H)$ is 25 (25)% ($m_{\tilde{\chi}_2^0} \approx m_{\tilde{\chi}_1^0}$) but if we choose branching fraction for $\tilde{\chi}_2^0 \rightarrow \tilde{G} + \gamma$ as 100%

and $\tilde{\chi}_1^0 \rightarrow \tilde{G} + Z(H)$ as 50 (50)%, then expected limits at 95% CL will change with respect to current scenario.

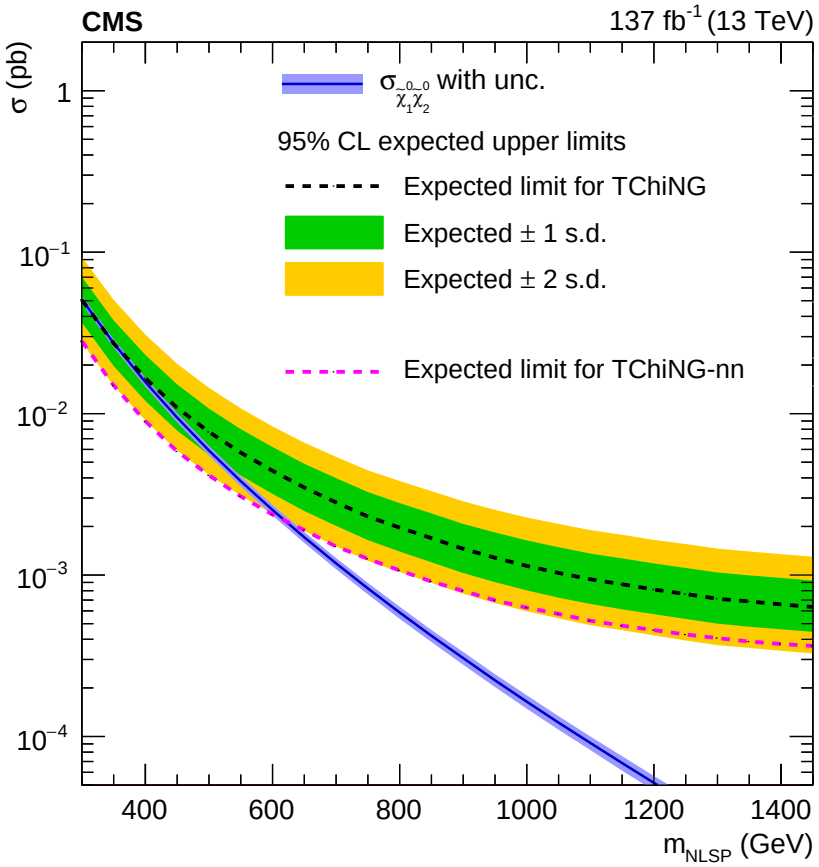

FIGURE E.2: Expected exclusions for different branching fractions of TChiNG model at 95% confidence level reached by current search. Blue curve represents the theoretical cross-sections for pp$\rightarrow \tilde{\chi}_1^0\tilde{\chi}_2^0$. Expected limits for $\tilde{\chi}_2^0 \rightarrow \tilde{G} + \gamma$ with 50% BR and $\tilde{\chi}_1^0 \rightarrow \tilde{G} + H\ (or\ Z)$ with 25% (25%) BR is represented as black dotted curve. Expected limits for $\tilde{\chi}_2^0 \rightarrow \tilde{G} + \gamma$ with 100% BR and $\tilde{\chi}_1^0 \rightarrow \tilde{G} + H\ (or\ Z)$ with 50% (50%) BR is represented as magenta dotted curve

Bibliography

[1] David J. Griffiths. *Introduction to Elementary Particles*. Wiley-VCH, 2008.

[2] Ernest Rutherford. "The Scattering of α and β Particles by Matter and the Structure of the Atom". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 21 (1911), pp. 669–688. DOI: 10.1080/14786440508637080. URL: https://doi.org/10.1080/14786440508637080.

[3] James Chadwick. "Possible Existence of a Neutron". In: *Nature* 129 (1932), p. 312. DOI: 10.1098/rspa.1932.0112.

[4] Albert Einstein. "Zur Elektrodynamik bewegter Körper". In: *Annalen der Physik* 322 (1905), pp. 891–921. DOI: http://dx.doi.org/10.1002/andp.19053221004.

[5] Paul A. M. Dirac. "The Quantum Theory of the Electron". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 117 (1928), pp. 610–624. DOI: 10.1098/rspa.1928.0023.

[6] George D. Rochester and Clifford A. Butler. "Evidence for the Existence of New Unstable Elementary Particles". In: *Nature* 160 (1947), pp. 855–856. DOI: 10.1038/160855a0.

[7] V.E. Barnes et al. "Observation of a Hyperon with Strangeness Minus Three". In: *Physical Review Letters* 12 (1964), pp. 204–206. DOI: 10.1103/PhysRevLett.12.204.

[8] Murray Gell-Mann. "A Schematic Model of Baryons and Mesons". In: *Phys. Lett.* 8 (1964), pp. 214–215. DOI: 10.1016/S0031-9163(64)92001-3.

[9] G. Zweig. "An SU(3) model for strong interaction symmetry and its breaking. Version 2". In: *DEVELOPMENTS IN THE QUARK THEORY OF HADRONS. VOL. 1. 1964 - 1978*. Ed. by D. B. Lichtenberg and Simon Peter Rosen. Feb. 1964, pp. 22–101.

[10] H. David Politzer. "Reliable Perturbative Results for Strong Interactions?" In: *Phys. Rev. Lett.* 30 (26 June 1973), pp. 1346–1349. DOI: 10.1103/PhysRevLett.30.1346.

[11] David J. Gross and Frank Wilczek. "Ultraviolet Behavior of Non-Abelian Gauge Theories". In: *Phys. Rev. Lett.* 30 (26 June 1973), pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343.

[12] H. David Politzer. "Asymptotic Freedom: An Approach to Strong Interactions". In: *Phys. Rept.* 14 (1974), pp. 129–180. DOI: 10.1016/0370-1573(74)90014-3.

[13] G. Arnison et al. "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV". In: *Phys. Lett. B* 122 (1983), pp. 103–116. DOI: 10.1016/0370-2693(83)91177-2.

[14] M. Banner et al. "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider". In: *Phys. Lett. B* 122 (1983), pp. 476–485. DOI: 10.1016/0370-2693(83)91605-2.

[15] Sheldon Lee Glashow. "Partial Symmetries of Weak Interactions". In: *Nuclear Physics* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[16] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[17] Abdus Salam. "Weak and Electromagnetic Interactions". In: *Conf. Proc. C* 680519 (1968), pp. 367–377. DOI: 10.1142/9789812795915_0034.

[18] F. Abe and et al. "Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab". In: *Physical Review Letters* 74.14 (1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626.

[19] ATLAS collaboration. "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), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[20] CMS collaboration. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". In: *Phys. Lett. B* 716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[21] Particle Data Group. *Review of Particle Physics*. https://pdg.lbl.gov/. 2020.

[22] CMS collaboration. "cross section measurements of SMP". https://twiki.cern.ch/twiki/pub/CMSPublic/PhysicsResultsCombined/SigmaNew_v3.pdf. 2021.

[23] Albert Einstein. *Relativity, the Special and the General Theory: A Popular Exposition*. Trans. by Robert W. Lawson. Crown, 1961.

[24] Paul Gorenstein and Wallace Tucker. "Astronomical Signatures of Dark Matter". In: *Advances in High Energy Physics* 2014 (2014), p. 878203. DOI: 10.1155/2014/878203. URL: https://doi.org/10.1155/2014/878203.

[25] S.P. Martin. "A Supersymmetry Primer". In: (2011). DOI: arXiv:hep-ph/9709356.

[26] T. Velasco. "Model Building and Phenomenology in Grand Unified Theories". In: 2015.

[27] Kaustubh Agashe, Roberto Contino, and Alex Pomarol. "The minimal composite Higgs model". In: *Nuclear Physics B* 719.1 (2005), pp. 165–187. ISSN: 0550-3213. DOI: https://doi.org/10.1016/j.nuclphysb.2005.04.035. URL: https://www.sciencedirect.com/science/article/pii/S0550321305003445.

[28] "Search for supersymmetry in events with a photon, a lepton, and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 01 (2019), p. 154. DOI: 10.1007/JHEP01(2019)154. arXiv: 1812.04066 [hep-ex].

[29] Albert M Sirunyan et al. "Combined search for supersymmetry with photons in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Lett. B* 801 (2020), p. 135183. DOI: 10.1016/j.physletb.2019.135183. arXiv: 1907.00857 [hep-ex].

[30] CMS collaboration. "Search for gauge-mediated supersymmetry in events with at least one photon and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Lett. B* 780 (2018), p. 118. DOI: 10.1016/j.physletb.2018.02.045. arXiv: 1711.08008 [hep-ex].

[31] CMS collaboration. "Mass reach for simplified models in the context of gauge-mediated supersymmetry breaking". https://twiki.cern.ch/twiki/bin/viewfile/CMSPublic/PhysicsResultsSUS?filename=barplot_GMSB.pdf;rev=2. 2019.

[32] ATLAS collaboration. "Search for supersymmetry in a final state containing two photons and missing transverse momentum in $\sqrt{s}$ = 13 TeV pp collisions at the LHC using the ATLAS detector". In: *Eur. Phys. J. C* 76 (2016), p. 517. DOI: 10.1140/epjc/s10052-016-4344-x. arXiv: 1606.09150 [hep-ex].

[33] ATLAS collaboration. "Search for photonic signatures of gauge-mediated supersymmetry in 13 TeV pp collisions with the ATLAS detector". In: *Phys. Rev. D* 97 (2018), p. 092006. DOI: 10.1103/PhysRevD.97.092006. arXiv: 1802.03158 [hep-ex].

[34] ATLAS collaboration. "Search for new phenomena in final states with photons, jets and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 07 (2023), p. 021. DOI: 10.1007/JHEP07(2023)021. arXiv: 2206.06012 [hep-ex].

[35] CMS collaboration. "Search for supersymmetry in events with a photon, jets, b-jets, and missing transverse momentum in proton-proton collisions at 13 TeV". In: *Eur. Phys. J. C* 79 (2019), p. 444. DOI: 10.1140/epjc/s10052-019-6926-x. arXiv: 1707.06193 [hep-ex].

[36] CMS collaboration. "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. arXiv: 1707.06193 [hep-ex].

[37] Alan D. Martin Francis Halzen. "Quarks and leptons: introductory course in modern particle physics". In: (1984). URL: http://gen.lib.rus.ec/book/index.php?md5=3bec2275686f95afb768858817bbc5a0.

[38] Michael Edward Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Reading, USA: Addison-Wesley (1995) 842 p. Westview Press, 1995.

[39] Matic Lubej. *standard-model*. https://github.com/mlubej/standard-model. 2021.

[40] B. Pontecorvo. "Neutrino Experiments and the Problem of Conservation of Leptonic Charge". In: *Zh. Eksp. Teor. Fiz.* 53 (1967), pp. 1717–1725.

[41] F. Zwicky. "Die Rotverschiebung von extragalaktischen Nebeln". In: *Helv. Phys. Acta* 6 (1933), p. 110. URL: http://adsabs.harvard.edu/abs/1933AcHPh...6..110Z.

[42] Vera C. Rubin and W. Kent Ford Jr. "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions". In: *Astrophys. J.* 159 (1970), p. 379. DOI: 10.1086/150317.

[43] Riccardo Barbieri and G. F. Giudice. "Upper bounds on supersymmetric particle masses". In: *Nucl. Phys. B* 306 (1988), p. 63. DOI: 10.1016/0550-3213(88)90171-X.

[44] S. Dimopoulos and G. F. Giudice. "Naturalness constraints in supersymmetric theories with nonuniversal soft terms". In: *Phys. Lett. B* 357 (1995), p. 573. DOI: 10.1016/0370-2693(95)00961-J. arXiv: hep-ph/9507282 [hep-ph].

[45] Riccardo Barbieri and Duccio Pappadopulo. "S-particles at their naturalness limits". In: *JHEP* 10 (2009), p. 061. DOI: 10.1088/1126-6708/2009/10/061. arXiv: 0906.4546 [hep-ph].

[46] Michele Papucci, Joshua T. Ruderman, and Andreas Weiler. "Natural SUSY endures". In: *JHEP* 09 (2012), p. 035. DOI: 10.1007/JHEP09(2012)035. arXiv: 1110.6926 [hep-ph].

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

[48] Hans Peter Nilles. "Supersymmetry, supergravity and particle physics". In: *Phys. Rep.* 110 (1984), p. 1. DOI: 10.1016/0370-1573(84)90008-5.

[49] *Overview and frontiers in the hunt for Supersymmetry at the LHC*. https://physikseminar.desy.de/sites2009/site_physikseminar/content/e212/e257140/e266968/e272659/Eifert_slides.pdf.

[50] G. R. Farrar and P. Fayet. "Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry". In: *Phys. Lett. B* 76 (1978), p. 575. DOI: 10.1016/0370-2693(78)90858-4.

[51] David Emmanuel-Costa, Pavel Fileviez Perez, and Ricardo Gonzalez Felipe. "Natural gauge and gravitational coupling unification and the superpartner masses". In: *Phys. Lett. B* 648 (2007), p. 60. DOI: 10.1016/j.physletb.2007.02.061. arXiv: hep-ph/0610178.

[52] R. Barbier et al. "R-parity violating supersymmetry". In: *Phys. Rept.* 420 (2005), pp. 1–202. DOI: 10.1016/j.physrep.2005.08.006. arXiv: hep-ph/0406039 [hep-ph].

[53] John Ellis et al. "Supersymmetric relics from the big bang". In: *Nucl. Phys. B* 238 (1984), pp. 453–476. DOI: 10.1016/0550-3213(84)90461-9. arXiv: 1103.0954 [hep-ph].

[54] Gian F. Giudice and Riccardo Rattazzi. "Theories with gauge-mediated supersymmetry breaking". In: *Phys. Rept.* 322 (1999), pp. 419–499. DOI: 10.1016/S0370-1573(99)00042-3. eprint: hep-ph/9801271.

[55] Joshua T. Ruderman and David Shih. "General Neutralino NLSPs at the Early LHC". In: *Journal of High Energy Physics* 2012.8 (2012), p. 159. DOI: 10.1007/JHEP08(2012)159.

[56] Simon Knapen and Diego Redigolo. "Gauge Mediation at the LHC: Status and Prospects". In: *Journal of High Energy Physics* 2017.1 (2017), p. 135. DOI: 10.1007/JHEP01(2017)135.

[57] Yevgeny Kats et al. "The Status of GMSB after 1/fb at the LHC". In: *Journal of High Energy Physics* 2012.2 (2012), p. 115. DOI: 10.1007/JHEP02(2012)115.

[58] Nima Arkani-Hamed et al. "MARMOSET: The path from LHC data to the new standard model via on-shell effective theories". 2007. arXiv: hep-ph/0703088 [HEP-PH].

[59] Johan Alwall, Phillip C. Schuster, and Natalia 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. arXiv: 0810.3921 [hep-ph].

[60] Lucian Harland–Lang et al. "Parton distributions in the LHC era: MMHT 2014 PDFs". In: *The European Physical Journal. C, Particles and Fields* 75 (2014). URL: https://api.semanticscholar.org/CorpusID:18756197.

[61] LHC SUSY Cross Section Working Group. "Cross sections for various subprocesses at 13 TeV". https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SUSYCrossSections. 2021.

[62] S Baird. *Accelerators for pedestrians; rev. version*. Tech. rep. Geneva: CERN, 2007. URL: http://cds.cern.ch/record/1017689.

[63] Ewa Lopienska. "The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022". In: (2022). URL: https://cds.cern.ch/record/2800984.

[64] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001. URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08001.

[65] *Project to prepare the PS Complex to be a Pre-injector for the LHC*. https://ps-div.web.cern.ch/LHC-PS/LHC-PS.html.

[66] Werner Herr and B Muratori. "Concept of luminosity". In: (2006). DOI: 10.5170/CERN-2006-002.361. URL: https://cds.cern.ch/record/941318.

[67] *Cumulative delivered and recorded luminosity versus time for 2015-2018 (pp data only)*. https://cmslumi.web.cern.ch/publicplots/multiYear/int_lumi_allcumulative_pp_run2.pdf.

[68] *Pile-up distributions for full Run2*. https://cmslumi.web.cern.ch/publicplots/pileup_allYears_run2.png.

[69] The CMS Collaboration. "The CMS experiment at the CERN LHC". In: *JINST* 3 (2008), S08004.

[70] CMS Collaboration. "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software". In: Technical design report. CMS (2006). There is an error on cover due to a technical problem for some items. URL: http://cds.cern.ch/record/922757.

[71] S. Chatrchyan et al. "The CMS experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[72] The CMS Collaboration. In: (). https://twiki.cern.ch/twiki/pub/CMSPublic/DPGResultsTRK/Phase1_Tracker_1Quarter.pdf.

[73] Mathew KilPatrick. *The CMS Tracker: Run 2 Experience and Upgrades*. https://agenda.linearcollider.org/event/7889/contributions/42679/attachments/33787/51850/CMS_Tracker_LCWS18_Matthew_Kilpatrick.pdf.

[74] The CMS Collaboration. "Strategies and performance of the CMS silicon tracker alignment during LHC Run 2". In: *Nucl. Instrum. Methods Phys. Res., A* 1037 (2022), p. 166795. DOI: 10.1016/j.nima.2022.166795. arXiv: 2111.08757. URL: https://cds.cern.ch/record/2790775.

[75] Martina Ressegotti. *Overview of the CMS detector performance at LHC Run 2*. https://indico.cern.ch/event/663474/contributions/3060096/attachments/1681230/2701095/ICNFP2018_Overview_CMS_performance.pdf.

[76] Federico De Guio. *First results from the CMS SiPM-based hadronic endcap calorimeter*. Tech. rep. Geneva: CERN, 2019. DOI: 10.1088/1742-6596/1162/1/012009. URL: https://cds.cern.ch/record/2636475.

[77] CMS Collaboration. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 7$TeV". In: *Journal of Instrumentation* 13.06 (June 2018), P06015. DOI: 10.1088/1748-0221/13/06/P06015. URL: https://dx.doi.org/10.1088/1748-0221/13/06/P06015.

[78] A Tapper and Darin Acosta. *CMS Technical Design Report for the Level-1 Trigger Upgrade*. Tech. rep. 2013. URL: http://cds.cern.ch/record/1556311.

[79] The CMS collaboration. "The CMS trigger system". In: *Journal of Instrumentation* 12.01 (), P01020. DOI: 10.1088/1748-0221/12/01/P01020. URL: https://dx.doi.org/10.1088/1748-0221/12/01/P01020.

[80] The CMS collaboration. "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Journal of Instrumentation* 15.10 (Oct. 2020), P10017. DOI: 10.1088/1748-0221/15/10/P10017. URL: https://dx.doi.org/10.1088/1748-0221/15/10/P10017.

[81] CMS collaboration. "Particle-flow reconstruction and global event description with the CMS detector". In: *JINST* 12.10 (2017), P10003. DOI: 10.1088/1748-0221/12/10/P10003. arXiv: 1706.04965 [physics.ins-det].

[82] Siona Ruth Davis. "Interactive Slice of the CMS detector". In: (2016). URL: https://cds.cern.ch/record/2205172.

[83] The CMS collaboration. "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 16.05 (2021), P05014. DOI: 10.1088/1748-0221/16/05/P05014. URL: https://dx.doi.org/10.1088/1748-0221/16/05/P05014.

[84] CMS E/gamma POG. "Cut Based Photon ID for Run 2". https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedPhotonIdentificationRun2. 2018.

[85] CMS E/gamma POG. "Cut Based Electron ID for Run 2 Legacy". https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2. 2018.

[86] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_T jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ex].

[87] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet user manual". In: *Eur. Phys. J. C* 72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[88] V. Khachatryan, A.M. Sirunyan, et al. "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV". In: *Journal of Instrumentation* 12.02 (Feb. 2017), P02014. DOI: 10.1088/1748-0221/12/02/P02014. URL: https://dx.doi.org/10.1088/1748-0221/12/02/P02014.

[89] Matteo Cacciari and Gavin P. Salam. "Pileup subtraction using jet areas". In: *Phys. Lett. B* 659 (2008), p. 119. DOI: 10.1016/j.physletb.2007.09.077. arXiv: 0707.1378 [hep-ph].

[90] CMS collaboration. "Pileup mitigation at CMS in 13 TeV data". In: *JINST* 15.09 (2020), P09018. DOI: 10.1088/1748-0221/15/09/P09018. arXiv: 2003.00503 [hep-ex].

[91] CMS collaboration. "Pileup mitigation at CMS in 13 TeV data". In: *JINST* 15 (2020), P09018. DOI: 10.1088/1748-0221/15/09/p09018. arXiv: 2003.00503 [hep-ex].

[92] Andrea Malara. "Reconstruction of jets and missing transverse momentum at the CMS experiment: Run 2 and perspective for Run 3". In: *Proceedings of 40th International Conference on High Energy physics — PoS(ICHEP2020)*. Vol. 390. 2021, p. 752. DOI: 10.22323/1.390.0752.

[93] CMS collaboration. "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV". Submitted to *JINST*. 2017. arXiv: 1712.07158 [physics.ins-det].

[94] "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{13}$ TeV using the CMS detector". In: *Journal of Instrumentation* 14.07 (July 2019), P07004. DOI: 10.1088/1748-0221/14/07/P07004. URL: https://dx.doi.org/10.1088/1748-0221/14/07/P07004.

[95] The CMS Collaboration. "Particle-flow reconstruction and global event description with the CMS detector". In: (). https://indico.cern.ch/event/792576/contributions/3403648/attachments/1912620/3160955/20190923_Top2019_ZhenHu.pdf.

[96] Th CMS Collaboration. "Jet energy scale and resolution performance with 13 TeV data collected by CMS in 2016-2018". In: *CMS-DP-2020-019 ; CERN-CMS-DP-2020-019* (). https://cds.cern.ch/record/2715872.

[97] The CMS collaboration. "Calibration of the CMS hadron calorimeters using proton-proton collision data at $\sqrt{s} = 13$ TeV". In: *Journal of Instrumentation* 15.05 (May 2020), P05002. DOI: 10.1088/1748-0221/15/05/P05002. URL: https://dx.doi.org/10.1088/1748-0221/15/05/P05002.

[98] 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.

[99] The CMS Collaboration. "Slides - CMS Phase1 upgrade summary during Run 1". In: *JINST* 12 (2017), P10003.

[100] The CMS Collaboration. "The CMS barrel calorimeter response to particle beams from 2 to 350 GeV/c". In: ().

[101] Sunanda Banerjee. "Calibration of HCAL". In: (2018). https://indico.tifr.res.in/indico/getFile.py/access?resId=0&materialId=slides&confId=6453.

[102] Francesco Pandolf. "10X Jet Performance with Different ECAL Noise". In: (2018). https://indico.cern.ch/event/704940/contributions/2897234/attachments/1602829/2541768/pandolfi_2018_02_19.pdf.

[103] Lindsey Gray, Thomas Klijnsma, and Shamik Ghosh. "A Dynamic Reduction Network for Point Clouds". In: *CoRR* abs/2003.08013 (2020). arXiv: 2003.08013. URL: https://arxiv.org/abs/2003.08013.

[104] "π^- energy reconstruction in HGCAL Beam Test prototype detector using Graph Neural Networks". In: (2022). URL: http://cds.cern.ch/record/2815404.

[105] Simon Rothman. *Calibrating Electrons and Photons in the CMS ECAL using Graph Neural Networks*. Tech. rep. Geneva: CERN, 2021. URL: https://cds.cern.ch/record/2799575.

[106] Johan Alwall et al. "Model-independent jets plus missing energy searches". In: *Phys. Rev. D* 79 (2009), p. 015005. DOI: 10.1103/PhysRevD.79.015005. arXiv: 0809.3264 [hep-ph].

[107] Daniele 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. arXiv: 1105.2838 [hep-ph].

[108] Serguei 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. arXiv: 1301.2175 [hep-ex].

[109] CMS collaboration. "Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s}$=13 TeV". In: *Journal of Instrumentation* 13.10 (Oct. 2018), P10005. DOI: 10.1088/1748-0221/13/10/P10005. URL: https://dx.doi.org/10.1088/1748-0221/13/10/P10005.

[110] 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. arXiv: 1405.0301 [hep-ph].

[111] Johan 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. arXiv: 0706.2569 [hep-ph].

[112] Torbjörn 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. arXiv: 1410.3012 [hep-ph].

[113] CMS collaboration. "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. arXiv: 1512.00815 [hep-ex].

[114] CMS collaboration. "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements". In: *Eur. Phys. J. C* 80 (2020), p. 4. DOI: 10.1140/epjc/s10052-019-7499-4. arXiv: 1903.12179 [hep-ex].

[115] 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. arXiv: 0706.2569 [hep-ph].

[116] R. Frederix and S. Frixione. "Merging meets matching in MC@NLO". In: *JHEP* 12 (2012), p. 061. DOI: 10.1007/JHEP12(2012)061. arXiv: 1209.6215 [hep-ph].

[117] Wim Beenakker et al. "NNLL resummation for squark-antisquark pair production at the LHC". In: *JHEP* 01 (2012), p. 076. DOI: 10.1007/JHEP01(2012)076. arXiv: 1110.2446 [hep-ph].

[118] Wim Beenakker et al. "Towards NNLL resummation: hard matching coefficients for squark and gluino hadroproduction". In: *JHEP* 10 (2013), p. 120. DOI: 10.1007/JHEP10(2013)120. arXiv: 1304.6354 [hep-ph].

[119] Wim Beenakker et al. "NNLL resummation for squark and gluino production at the LHC". In: *JHEP* 12 (2014), p. 023. DOI: 10.1007/JHEP12(2014)023. arXiv: 1404.3134 [hep-ph].

[120] Wim Beenakker et al. "NNLL resummation for stop pair-production at the LHC". In: *JHEP* 05 (2016), p. 153. DOI: 10.1007/JHEP05(2016)153. arXiv: 1601.02954 [hep-ph].

[121] Wim Beenakker et al. "NNLL-fast: predictions for coloured supersymmetric particle production at the LHC with threshold and Coulomb resummation". In: *JHEP* 12 (2016), p. 133. DOI: 10.1007/JHEP12(2016)133. arXiv: 1607.07741 [hep-ph].

[122] Andrew J. Larkoski et al. "Soft Drop". In: *JHEP* 05 (2014), p. 146. DOI: 10.1007/JHEP05(2014)146. arXiv: 1402.2657 [hep-ph].

[123] CMS Muon POG. "Reference muon id, isolation and trigger efficiencies for 2017 data". https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2017. 2018.

[124] CMS Muon POG. "Reference muon id, isolation and trigger efficiencies for 2018 data". https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2018. 2018.

[125] Rohan Bhandari, Adam Dishaw, Ryan Heller, Ana Ovcharova, Jeffrey Richman, Manuel Franco Sevilla, David Stuart, Chris West, Jae Hyeok Yoo. "Search for supersymmetry in pp collisions at $\sqrt{s} = 13$ TeV in the single-lepton final state using the sum of masses of large radius jets". In: *CMS Analysis Note* AN-15-139 (2015). URL: https://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2015/139.

[126] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. C* 71 (2011), p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].

[127] The CMS collaboration. "Measurement of the inclusive and differential $t\bar{t}\,\gamma$ cross sections in the single-lepton channel and EFT interpretation at *sqrts* = 13 TeV". English (US). In: *Journal of High Energy Physics* 2021.12 (Dec. 2021). Publisher Copyright: © 2021, The Author(s). ISSN: 1126-6708. DOI: 10.1007/JHEP12(2021)180.

[128] Richard D. Ball et al. "Parton distributions for the LHC Run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[129] Richard D. Ball et al. "Parton distributions from high-precision collider data". In: *Eur. Phys. J. C* 77 (2017), p. 663. DOI: 10.1140/epjc/s10052-017-5199-5. arXiv: 1706.00428 [hep-ph].

[130] 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. arXiv: hep-ph/0303085 [hep-ph].

[131] Stefano 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. arXiv: hep-ph/0306211 [hep-ph].

[132] CMS collaboration. "Measurements of Inclusive W and Z Cross Sections in pp collisions at $\sqrt{s} = 7$ TeV". In: *JHEP* 01 (2011), p. 080. DOI: 10.1007/JHEP01(2011)080. arXiv: 1012.2466 [hep-ex].

[133] Ansgar 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. arXiv: 1510.08742 [hep-ph].

[134] Albert M. Sirunyan et al. "Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS". In: *Eur. Phys. J. C* 81 (2021), p. 800. DOI: 10.1140/epjc/s10052-021-09538-2. arXiv: 2104.01927 [hep-ex].

[135] CMS Collaboration. *CMS luminosity measurement for the 2017 data-taking period at* $\sqrt{s} =$ *13 TeV*. CMS Physics Analysis Summary CMS-PAS-LUM-17-004. 2018. URL: https://cds.cern.ch/record/2621960.

[136] CMS Collaboration. *CMS luminosity measurement for the 2018 data-taking period at* $\sqrt{s} =$ 13 *TeV*. CMS Physics Analysis Summary CMS-PAS-LUM-18-002. 2019. URL: https://cds.cern.ch/record/2676164.

[137] CMS LUMI POG. "Lumi Recommendations". https://twiki.cern.ch/twiki/bin/viewauth/CMS/TWikiLUM. 2022.

[138] "Search for supersymmetry in proton-proton collisions at 13 TeV in final states with jets and missing transverse momentum". In: *JHEP* 10 (2019), p. 244. DOI: 10.1007/JHEP10(2019)244. arXiv: 1908.04722 [hep-ex].

[139] Manuel Franco Sevilla, Ana Ovcharova. "Preliminary ISR reweighting and uncertainty recommendations for Moriond'17". https://indico.cern.ch/event/592621/contributions/2398559/attachments/1383909/2105089/16-12-05_ana_manuelf_isr.pdf. 2016.

[140] CMS Higgs Combination Group. "Documentation of the RooStats-based statistics tools for Higgs PAG". https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombine 2015.

[141] Thomas 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. arXiv: hep-ex/9902006 [hep-ex].

[142] Alexander L. Read. "Presentation of search results: the CLs technique". In: *J. Phys. G* 28 (2002), p. 2693. DOI: 10.1088/0954-3899/28/10/313.

[143] The CMS Collaboration. "Search for new physics in multijet events with at least one photon and large missing transverse momentum in proton-proton collisions at 13 TeV". In: *JHEP* 2023 (2023), p. 46. DOI: 10.1007/JHEP10(2023)046.

[144] CMS collaboration. "Mass reach for simplified models in the context of gauge-mediated supersymmetry breaking". https://twiki.cern.ch/twiki/pub/CMSPublic/PhysicsResultsSUS/barplot_GMSB.png. 2019.