Energy measurement of hadron showers in the CMS HGCAL using Graph Neural Networks and a new physics search with photons at the CMS experiment

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

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

द्वारा / By

अल्पना / Alpana

पंजीकरण सं. / Registration No.: 20193693

शोध प्रबंध पर्यवेक्षक / Thesis Supervisor: Prof. Seema Sharma

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

2025

Dedicated to my grandmother

Declaration by Student

Name of Student: Alpana

Reg. No.: 20193693

Thesis Supervisor(s): Prof. Seema Sharma

Department: Physics

Date of joining program: 01 August 2019

Date of Pre-Synopsis Seminar: 04 December 2024

Title of Thesis: Energy measurement of hadron showers in the CMS HGCAL using Graph Neural Networks and a new physics search with photons at the CMS experiment

I declare that this written submission represents my idea in my own words and where others' ideas have been included; I have adequately cited and referenced the original sources. I declare that I have acknowledged collaborative work and discussions wherever such work has been included. 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.

The work reported in this thesis is the original work done by me under the guidance of Prof. Seema Sharma

Date: 30 January 2025

Signature of the student:

CERTIFICATE

Certified that the work incorporated in the thesis entitled “**Energy measurement of hadron showers in the CMS HGCAL using Graph Neural Networks and a new physics search with photons at the CMS experiment**” Submitted by **Alpana** 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.

Prof. Seema Sharma

Date: 30-01-2025

Acknowledgments

As I am completing my Ph.D., I find myself reflecting on the journey that has brought me here. The aspiration to earn a Ph.D. and contribute to scientific research for the betterment of the world has been a dream of mine since the 8th grade, inspired by Dr. APJ Abdul Kalam. While I always knew what I wanted to pursue, I was uncertain about the path to get there. The Ph.D. position at IISER Pune provided me with the opportunity to turn this vision into reality, shaping my journey in ways that I could not have foreseen.

First and foremost, I am profoundly grateful to my supervisor, Prof. Seema Sharma, for her invaluable mentorship, which has been the pillar of my Ph.D. journey, not only in advancing my research, but also in shaping my identity as a physicist. She has played a pivotal role in teaching me how to conduct rigorous research, fostering in me the patience, analytical mindset, and perseverance necessary to navigate the complexities of experimental high-energy physics. Throughout these years, her unwavering belief in my abilities, even in challenging times, has motivated me to work on diverse projects, including the validation of HGCAL back-end (BE) readout systems and supersymmetry searches. With her constructive criticism and insightful feedback, she has ensured that I am equipped with the essential skills to grow as an independent researcher. I have been deeply inspired by her dedication to academia and I aspire to carry forward the same commitment and passion in my own journey as a physicist.

As members of my Research Advisory Committee (RAC), Prof. Bhas Bapat and Dr. Jyothsna Komaragiri have played a crucial role in guiding my research. Their deep understanding of physics, through insightful discussions and expert advice, has significantly enriched my academic growth, research experience, and clarity of my work.

My heartfelt thanks go to my collaborators in CMS HGCAL and its BE-DAQ & Detector Physics Group, as well as CMS SUS-PAG, for their generous guidance and encouragement throughout the various stages of my research. I had the pleasure of closely working with brilliant physicists including Andre David, Andrew Gilbert, Paul Dauncey, Pedro Silva, Roger Rusack, and Rajdeep Chatterjee. Their insights and discussions have played a crucial role in refining my work and my growth as a physicist. I sincerely appreciate the insightful discussions with Lindsey Gray and Simon Rothman, which enhanced my understanding of the DRN architecture. I am grateful to Marco Rovere for giving me the opportunity to contribute to the HGCAL reconstruction efforts. I'm also thankful to Bhargav Joshi for the chance to collaborate on making CMS data public.

During my visit to CERN, I had the incredible opportunity to participate in the HGCAL beam test, an experience that was both enriching and unforgettable. I am deeply grateful to my colleagues Milos Vojinovic, Raghunandan Shukla, Arnaud Steen, Stavros Mallios, and Martim Rosado, whose support and guidance were pivotal in enhancing my understanding of the HGCAL BE system, and I greatly value the insights they shared. Discussions on general physics with Prof. Sunanda Banerjee have been deeply insightful and enlightening.

I also thank Bhumika Kansal, Andrew Askew, and Samuel Bein for their valuable insights that enriched the analysis presented in this thesis. I am grateful to Kevin Pedro and Purbita Rahman for their assistance in obtaining the necessary datasets for this analysis.

I am sincerely grateful to Dr. Shubham Pandey, who guided me from my first steps in understanding calorimeters to the point where I now look toward postdoctoral opportunities. His invaluable discussions and constant support have not only deepened my knowledge, but also ignited my passion for detector physics. I am deeply grateful to Dr. Bhumika Kansal for her guidance in shaping my understanding of new physics searches and statistical physics. I feel privileged to have worked alongside an amazing team of colleagues who turned this into an enjoyable journey, including Dr. Aditee Rane, Niramay, Shivani, Nukul, Rahul, Dhruvanshu, Nitish, Chirayu, Tanay, Chetan, Somanko, and Bapi. Big thanks to Shivani and Nukul for making my CERN visit such a blast. I am grateful for all the fun memories I had with Sumit and Akash in the IISER corridors.

I would like to extend my sincere gratitude to Nisha Kurkure and Goldi Mishra for their dedicated and prompt assistance with the IISER NSM facilities and IISER-CMS cluster. I am also grateful for our administrative staff, Prabhakar Anagare and Dhanashree Sheth who alleviated bureaucratic challenges throughout my tenure at IISER Pune. I also appreciate support of Mr. Nilesh and Prashant from the UG laboratory for their assistance in the EHEP lab. I also acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota and IISER NSM facilities for providing resources that contributed to this thesis. I gratefully acknowledge the support of the Council of Scientific and Industrial Research (CSIR), the Department of Science and Technology (DST), the Department of Atomic Energy (DAE), and the India CMS collaboration through the Ph.D. fellowship and CMS grant.

I owe a great deal to my friends at IISER for making this journey beautiful and peaceful. I could not have done it without my people, including my fellow women in science, Bhawakshi, Cavya, Rukhsana, and Sarika, and my fellow men in science, Deepak, Nahid, Rishu, Sandeep and Umashish. My heartfelt thanks go to my dear friends Ashish, Chandraprakash, Kunal, Kush, Milos, Pinki, Ramesh, and Ritesh, whose unwavering support has been a source of strength, bringing warmth and happiness.

Above all, I am forever grateful to my dearest family. Their constant belief in me and unconditional support have been my greatest source of strength. Their presence has been a constant reminder of the love and resilience that have carried me through every step of this journey.
I am eternally grateful to Kuldeep, the one who has been my rock throughout this journey. Your constant support turned it into reality. You have supported me through all the ups and downs, protecting and encouraging me to push through every hurdle. Thank you for always being there.
"Finally, here's a big shout out to Mummy, Papa, Ritu, Divya, Shakti, Shivam, Dadi, Chacha, Maai, Baba, and Kuldeep. I love you to the moon and back!"

- Alpana

Contents

Abstract **xiii**

1 Introduction **1**

2 Theoretical and experimental aspects of particle physics **7**
2.1 Elementary particles . . . 7
2.2 Standard model of elementary particles . . . 9
2.2.1 Quantum electrodynamics . . . 9
2.2.2 Quantum chromodynamics as non abelian gauge theory . . . 10
2.2.3 Electroweak interactions . . . 11
2.2.4 Higgs mechanism . . . 12
2.2.5 Shortcomings of Standard Model . . . 14
2.3 Supersymmetry . . . 14
2.3.1 Advantageous features of Supersymmetry . . . 15
2.3.2 Minimal Supersymmetric extension of the Standard Model . . . 16
2.4 Supersymmetry models examined in this thesis . . . 17
2.4.1 Phenomenological Minimal Supersymmetric Standard Model (pMSSM) 17
2.4.2 Gauge mediated symmetry breaking (GMSB) model . . . 18
2.5 Simplified SUSY models at LHC . . . 20
2.6 Particle interaction with matter . . . 22
2.6.1 Electromagnetic interactions of particles with material . . . 22
Charged particles . . . 22
Radiation . . . 23
2.6.2 Interaction of photons with matter . . . 25
2.6.3 Energy loss via strong nuclear interactions . . . 26
2.7 Particle cascades in the material . . . 27
2.7.1 Electromagnetic showers . . . 28
2.7.2 Hadronic showers . . . 30
2.8 Particle detectors . . . 32
2.8.1 Ionization detectors . . . 32
2.8.2 Scintillator detectors . . . 33
2.9 Calorimeters . . . 34
2.9.1 Performance metrics for calorimeters . . . 36
2.10 Event simulations in high energy physics . . . 39
2.10.1 Overview of detector simulation using GEANT4 . . . 39
Modeling electromagnetic showers . . . 40
Modeling of hadronic showers . . . 40
Cascade models . . . 40
2.10.2 Event simulation in GEANT4 . . . 42
2.11 Summary . . . 43

3 Experimental setups **45**
3.1 The Large Hadron Collider (LHC) . . . 45
3.1.1 Luminosity and pileup interactions . . . 46
3.1.2 High luminosity phase of LHC (HL-LHC) . . . 48
3.2 The compact muon solenoid (CMS) experiment and event reconstruction 48
3.2.1 Silicon tracker . . . 49
3.2.2 The Electromagnetic calorimeter (ECAL) . . . 50
3.2.3 The Hadron calorimeter (HCAL) . . . 51
3.2.4 Muon system . . . 51
3.2.5 Triggers and data acquisition system . . . 52
3.2.6 Event reconstruction . . . 53
3.3 CMS detector upgrades along the LHC . . . 54
3.3.1 Phase-1 upgrades of CMS . . . 54
3.3.2 Phase-2 upgrade of CMS . . . 57
3.4 The CMS High Granularity Calorimeter . . . 60
3.4.1 Electromagnetic and hadronic sections . . . 61
3.4.2 Silicon sensor modules and scintillator-SiPM tileboards . . . 62
3.4.3 Expected physics performance of HGCAL . . . 65

4 Advanced machine learning algorithms **73**
4.1 Building blocks of machine learning . . . 74
4.1.1 Neural networks: extending linear models . . . 76
4.2 Graph Neural Networks (GNNs) . . . 78
4.2.1 Graph Clustering . . . 81
4.2.2 Pooling Operations in GNNs . . . 82
4.3 Dynamic reduction network . . . 82
4.4 Summary . . . 85

5 Energy reconstruction of charged pions in HGCAL using DRN **87**
5.1 Event sample specifications and cleansing procedures . . . 88
5.2 Energy reconstruction of pions using classical methods . . . 89
5.2.1 Energy reconstruction of charged pions using a fixed weight method . . . 91
5.2.2 Energy reconstruction of charged pions using weighted summation (WS) method . . . 93
5.3 Dynamic reduction network for charged pion energy reconstruction . 95
5.3.1 Optimization of dynamic reduction network architecture . . . 96
Impact of dropping the AHCAL layers on DRN's performance 98
Impact of training sample statistics on the performance . . . 99
Optimizing number of aggregation layers . . . 100
Optimizing number of layers in input net . . . 101
Optimization of learning rate . . . 102
5.3.2 Optimizing the input features and target . . . 104
Optimizing training target . . . 104
5.4 Performance of Dynamic reduction network for charged pion energy reconstruction . . . 108
5.4.1 Inference on shuffled order hits on DRN trained with ordered hits . . . 116
5.5 DRN performance with rechit energies in MIPs . . . 117
5.6 Performance of DRN on charged pion data collected in beam test 2018 experiments at CERN . . . 122

5.6.1 Event categorization based on pion shower start location 126
5.7 Understanding improvements in energy resolution obtained using DRN 131
5.7.1 Assessing DRN performance based on the energy deposited in the absorbers 131
5.7.2 Compensating the fluctuations in EM component 140
5.8 Energy calibration of CLUE clusters 144
5.9 Expected performance using the full HGCAL geometry 149
5.10 Summary 153

6 Testing and validating the HGCAL readout system 155
6.1 The HGCAL electronics readout system 155
6.1.1 Front-end (FE) readout system 156
HGCROC - the HGCAL ReadOut Chip 157
Endcap Concentrator for DAQ (ECON-D) 159
The lpGBTs and VTRX+ ASICs 160
6.1.2 Back-end (BE) readout system 161
6.1.3 HGCAL buffer emulator 164
6.2 Modeling occupancy differences in different geometry versions 168
6.2.1 Methodology for HDM $\leftrightarrow$ LDM occupancy transformation . . . 169
Uniformly extrapolating HDM $\leftrightarrow$ LDM occupancy 171
6.3 Pseudo event generation for enhancing the event statistics 173
6.3.1 Multivariate Gaussian distributions 173
6.3.2 Addressing the challenges involved in HGCAL pseudo event generation 174
6.3.3 An optimized method for Pseudo Event Generation in HGCAL 178
6.4 Effect of data rates and system settings on buffer occupancy and data loss 181
6.4.1 Impact of high L1A trigger rates on data losses 182
6.4.2 Buffers performance with artificially increased data rates . . . 183
6.5 Throttling to avoid truncation 192
6.5.1 Dead time introduced by linearly increasing data rates 194
6.5.2 Dead time introduced with increasing data rates non-linearly . 195
6.6 Testing of integrated HGCAL readout system with beam test experiment at CERN 199
6.6.1 Description of beam test setup 200
6.6.2 Milestones accomplished 202
6.6.3 Alternative approaches to examining raw data 202
6.6.4 Inspecting the raw data with finer details 202
Alignment of L1A time stamps in the HGCAL readout chain . . 205
6.7 Summary 210

7 New physics searches in multijet events with at least a photon and high transverse momentum imbalance 211
7.1 Targeted physics scenarios & simplified SUSY models 211
7.1.1 Background SM processes 214
7.1.2 MC event simulation & dataset generator 215
7.2 Composition of signal and background 217
7.2.1 Baseline objects and event selections 217
7.2.2 Optimization of search region bins 221
Optimization for pMSSM motivated models 226
7.3 Background estimations 230

7.3.1 Estimation of lost lepton background . . . 230
Correction of MC to match data and their comparisons in CR . 234
Transfer factor calculation . . . 234
7.3.2 Predictions of lost lepton background in data & sources of uncertainties . . . 238
7.3.3 Estimation e-faking-photon background . . . 243
Comparison of single electron CR in Data & MC . . . 244
Transfer factor calculation . . . 245
Data-MC correction factor . . . 249
7.3.4 Predictions of electron faking photon background in data & sources of uncertainties . . . 253
7.3.5 Estimation of Z-invisible background . . . 256
7.3.6 Predictions of Z-invisible background & sources of uncertainties 259
7.4 Results & interpretations . . . 263
7.4.1 Expected exclusion limits using data driven background estimation . . . 265
7.5 Future aspects & outlook . . . 267
7.5.1 Boosted Decision Trees . . . 267
7.5.2 Model parameters and its training . . . 268
7.5.3 Updated exclusion limits with BDT . . . 268
7.6 Summary . . . 270

8 Summary and outlook **275**

A Appendix A - charged pion energy reconstruction supporting **279**
A.1 Reconstruction of charged pions . . . 279

B Appendix B - New physics search using photons and $p_{\mathrm{T}}^{\mathrm{miss}}$ in multijet events **297**
B.1 Background estimation . . . 297

List of Tables

2.1 Properties of materials relevant to electromagnetic and hadron interactions [1]. 27

3.1 Different parameters of LHC and HL-LHC [2]. 48

5.1 The above table lists the energy resolution obtained for different pion energies using various methods in Data. 107
5.2 The list of epochs used to evaluate different DRNs performance along with listing down the minimum validation loss. 109
5.3 Stochastic and constant terms and the uncertainties on these using different energy reconstruction method for simulated dataset. 125
5.4 Stochastic and constant terms and the uncertainties on these using different energy reconstruction method for pion data recorded with beam test setup. 125

7.1 Cutflows for all backgrounds after baseline selections. 220
7.2 Cutflows for T5bbbbZG & TChiWG after baseline selections. 220

Abstract

The CMS experiment at the LHC has been rigorously testing the Standard Model in proton-proton collision data for more than a decade now. As the LHC progresses through Run-3 at a center-of-mass energy of 13.6 TeV, the CMS collaboration is also preparing for the High-Luminosity LHC (HL-LHC) phase to operate with more intense beams in the upcoming decade. To meet the challenges of high radiation doses and more particle flux, current CMS endcap calorimeters will be replaced by the High-Granularity Calorimeter (HGCAL), a state-of-the-art detector that features six million read-out channels capable of providing highly detailed 5-dimensional measurements: energy, position, and timing. These capabilities open new frontiers in the reconstruction of particle showers with unprecedented resolution.

In this thesis, a novel dynamic reduction network (DRN) based on graphical neural networks (GNN) is used to significantly improve energy resolution of charged pions in the HGCAL. The improvements from the algorithm are validated on negatively charged pion data collected with the HGCAL beam test experiments at the CERN facilities. Further insight into the additional information learned by the GNN is obtained using GEANT4 based detector simulation. As the HGCAL approaches its preproduction phase, it is imperative to evaluate the efficacy of the finalized design, with a particular focus on the readout components, through both system tests and beam tests to guarantee operational integrity and seamless data transmission without loss. This thesis presents key results from software based studies of the HGCAL readout electronics and data handling and from the beam test of complete near-final HGCAL readout system during the August-September 2023 experiments at the CERN H4 beam test facility.

I also present a new methodology for searching for new physics produced via strong and electroweak supersymmetry scenarios in Run 2 data using soft photons as a probe. Lowering the photon momentum threshold increases the background tenfold and reduces signal sensitivity when using the jet multiplicity and MET as primary criteria based on previous searches. By reevaluating events with photon momentum and the scalar sum of all jets and photon momentum, signal sensitivity improves across strong and electroweakino productions.

Chapter 1

Introduction

Even though Mendeleev put together the atoms, the then indivisible entities of matter, as early as in the 1860s, the real insight into their internal structure came only with the discovery of electrons in 1897 by J.J. Thomson and the nucleus in 1911 by Ernest Rutherford. Furthermore, it was the discovery of neutrons in 1932 by James Chadwick that completed the picture of the atom's fundamental constituents. Subatomic particles like muons and pions were discovered in cosmic ray experiments in the 1930s and 1940s, revealing a new layer of complexity in the particle world. By the mid-20th century, we learned that even protons and neutrons are made of smaller particles called quarks, marking the birth of particle physics as a separate field of study. Particle physics helps us understand the atomic and subatomic world, tracing back to the Big Bang to explain many features of the universe we see today.

The quest for finding and understanding the most fundamental constituents of universe which are irreducible into smaller units, also referred to as elementary particles, along with the nature of interactions among them is the core of the ongoing research in the particle physics. The primary interactions among elementary particles are governed by four fundamental forces: gravitational, electromagnetic, strong nuclear, and weak nuclear forces. Except gravity, the knowledge of the other three interactions is encapsulated in the Standard Model of elementary particle physics that explains the basic particles and their interactions. As of our understanding today, matter is made of fermions while bosons mediate the forces or interactions among particles.

While the SM effectively accounts for fundamental forces, it does not resolve several crucial issues, including the particle nature of dark matter, origin of matter-antimatter asymmetry observed in the universe, quantum nature of gravity, to name a few. Filling these gaps necessitates exact SM property measurements and exploration for beyond Standard Model (BSM) phenomena. This requires the production of heavy SM particles, such as the W/Z bosons (each approximately 80-90 times the proton mass), the Higgs boson (125 times the proton mass), and the top quark (175 times the proton mass). These particles can be produced in collisions of very high energy e^+e^- or proton-proton (pp).

The LHC [3] at CERN is presently the world's largest particle accelerator and collider. It is a circular accelerator with a circumference of 27 km, situated 100 m underground at the borders of Switzerland and France. It accelerates two proton beams traveling in opposite directions to an energy of up to 6.5 TeV each, and makes them collide at four locations which are instrumented with CMS [4], ATLAS [5], ALICE [6], and LHCb [7] experiments. These experiment aims to deepen our understanding of the fundamental constituents of universe and test predictions of various theories of elementary particles. The ATLAS and the CMS are general-purpose detectors designed to explore a broad spectrum of physics processes for precision measurements of SM

and searching for new physics motivated by several extensions to SM or BSM theories. The ALICE [6] detector investigates the physics of strong interactions under extreme energy densities, where quark-gluon plasma emerges through heavy-ion collisions at the LHC. The LHCb [7] is specifically designed to study the decays of particles with "beauty" (b) quarks. It facilitates precise studies of matter-antimatter asymmetry via measuring CP violation in the beauty quark sector to ascertain the dominance of matter in the universe.

The CMS experiment [4] is built around a large solenoid magnet producing a magnetic field of up to 4 Tesla. It consists of several subdetectors for measuring the momentum and position of photons, electrons, muons, and hadrons resulting from pp collisions. Neutrinos, which interact via weak interaction, do not leave any visible signature in the detector and their presence in a collision event is inferred from an imbalance in total visible momenta in direction transverse to the beam direction. The CMS experiment employs a two-level trigger system (Level-1 and High-level trigger) to select events of interest in realtime, reducing the event rates for permanent storage to acceptable levels. The reconstruction of an event in CMS detector starts with identifying & reconstructing particles using a Particle Flow (PF) algorithm [8] that integrates data from various CMS detectors.

Over a decade ago, the Higgs boson, the final missing piece of the SM, was discovered by the CMS and ATLAS experiments at the Large Hadron Collider (LHC) [9, 10]. This pivotal discovery substantiated the electroweak symmetry breaking mechanism, attributing mass to fundamental particles.
The CMS has actively pursued a wide-ranging program to validate SM predictions. Crucial measurements consists of production cross section of Higgs boson via different modes such as gluon-gluon fusion (ggF), vector boson fusion (VBF), and associated production with vector bosons (VH) or top quarks (ttH), measurements of different decay modes and branching ratios. Measured values of these Higgs boson properties aligns closely with the SM predictions as shown in Fig. 1.1.
Besides the Higgs-related research, CMS has conducted precise determinations of production cross sections across diverse SM processes. These include Electroweak processes and the production of W and Z bosons, both inclusively and differential in number of jets and other kinematic variables. Additionally, there are analyses of single top and top quark pair production, which are critical for scrutinizing SM parameters. Furthermore, CMS examines QCD processes such as strong interaction phenomena at unprecedented energy levels.
Theoretical extensions of the SM, like Supersymmetry (SUSY), extra-dimensions or those with additional symmetries, aim to provide solutions to these issues. Potential answers to these issues include SM extensions like Supersymmetry (SUSY) and theories involving extra dimensions. A fermion (bosonic) superpartner is added to each bosonic (fermion) SM particle in SUSY, a theoretical extension of the SM that assumes a symmetry between fermions and bosons [12].
SUSY offers a stable dark matter candidate in the form of the lightest supersymmetric particle (LSP), which is widely believed to interact weakly, assuming that R-parity [13] is conserved. It is therefore a compelling contender for dark matter [14].

In the search for SUSY at the LHC, simplified models are often used to interpret experimental data to avoid dealing with over 120 free parameters of SUSY. These simplified models typically focus on key SUSY particles, such as squarks (the superpartners of quarks), gluinos (the superpartners of gluons), or sleptons (the superpartners of

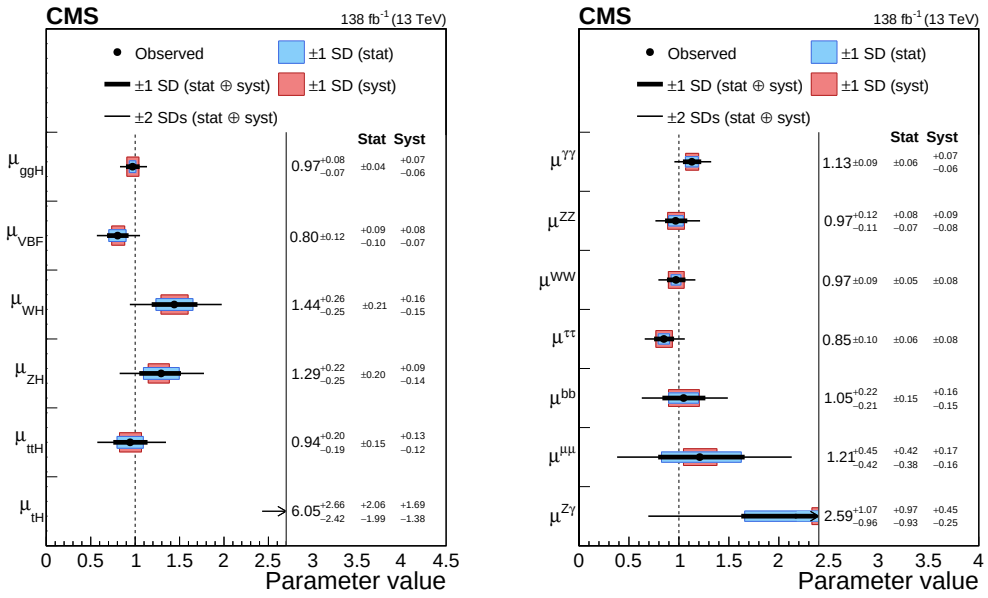
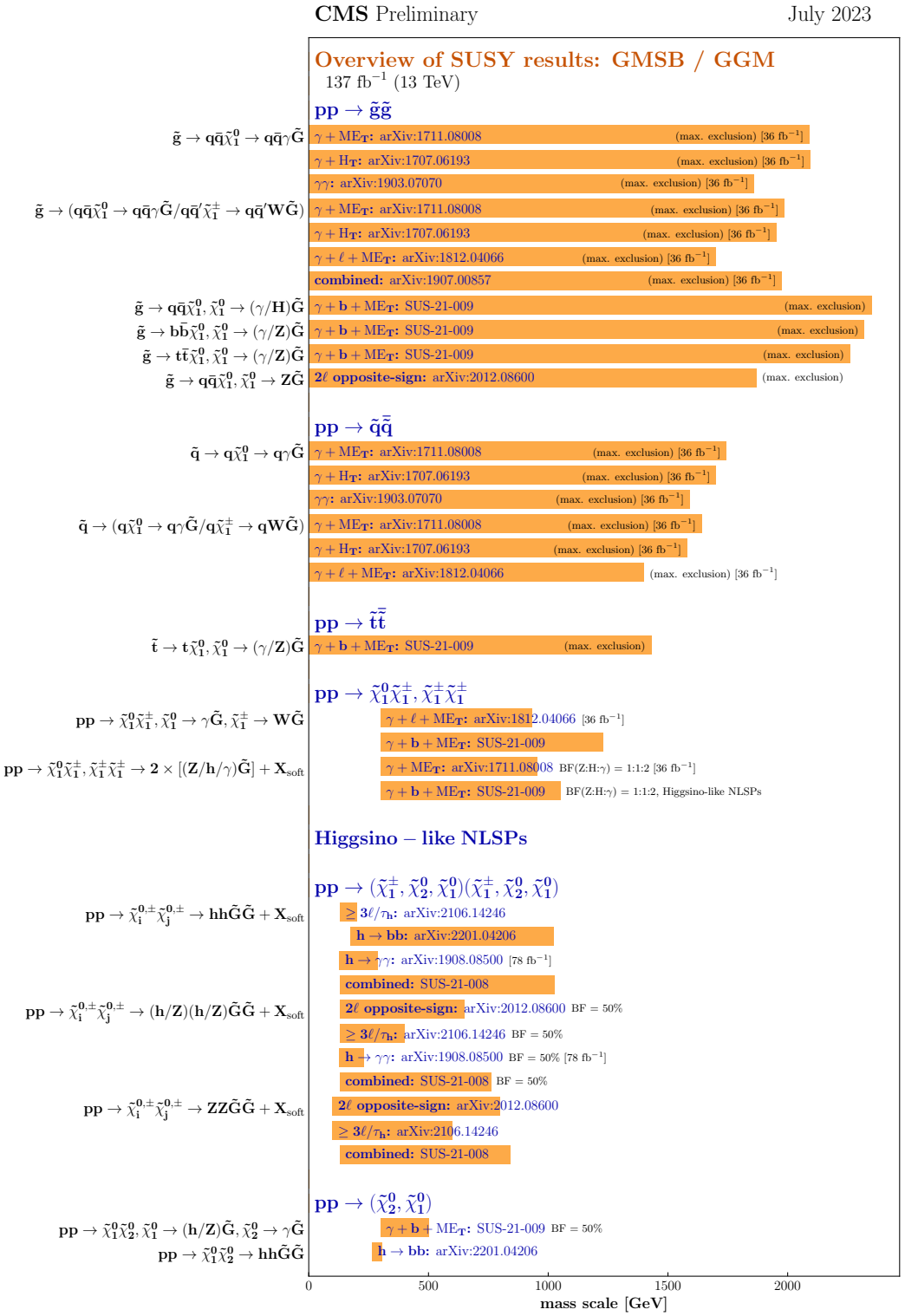


Figure 1.1: Comparison of observed and SM predictions for both production modes (left) and decay channels (right). Signal strength parameters are derived for (left) distinct production modes μ_i, under the assumption BF=$(BF)_{SM}$, and (right) decay channels μ_f, assuming $\sigma^i = (\sigma^i)_{SM}$. The dashed vertical line at one denotes the SM values for μ_i and μ_f [11].

leptons). These models predict various final states observable in detectors through particle production. For instance, strong interactions through gluinos and squarks typically produce multijets with notable transverse momentum imbalance ($p_{\mathrm{T}}^{\mathrm{miss}}$), sometimes accompanied by leptons and/or photons. Electroweakino processes often lead to many final-state leptons from W, Z, and H boson decays. Algorithms are employed to detect these boosted massive bosons via their hadronic decay modes to identify possible signatures of SUSY particles.

To avoid dealing with large background at the same time, these searches are often conducted by focusing on specific final states such as zero or more photons and leptons, and multijets events with MET. Historically, a multitude of investigations have concentrated on these final states, driven by supersymmetry scenarios involving both strong and electroweakino production. Fig. 1.2 shows the mass range for different superpartners excluded by different SUSY searches conducted in CMS experiment. Gluino and $\tilde{\chi}_1^0$ masses up to 2 TeV and 1 TeV, respectively, have been excluded.

Our focus is on the exploration of SUSY through the utilization of soft photons as probes in multijet events characterized by high missing transverse energy (MET). Historically, this phase space has been investigated using photons with high transverse momentum p_T [16], avoiding unmanageable background with lowering momentum thresholds for photons. The interest to reduce these thresholds is motivated by the phenomenologically Minimal Supersymmetric Standard Model (pMSSM), wherein next-to-lightest supersymmetric particles (NLSPs), (wino or bino like), decay into soft photons. Furthermore, this methodology raises the potential to enhance sensitivity to SUSY particle production motivated by Gauge-Mediated Symmetry Breaking (GMSB) [14] in both strong and electroweak scenarios.



Selection of observed limits at 95% C.L. (theory uncertainties are not included). Probe **up to** the quoted mass limit for light LSPs unless stated otherwise. The quantities ΔM and x represent the absolute mass difference between the primary sparticle and the LSP, and the difference between the intermediate sparticle and the LSP relative to ΔM, respectively, unless indicated otherwise.

Figure 1.2: Excluded mass range of different sparticles for simplified models within GMSB scenarios. [15]

The ongoing operational phase of the LHC is expected to conclude next year with collecting pp collision data equivalent to a total of about 450 fb^{-1} luminosity. Over more than a decade, CMS has examined pp collisions to test the SUSY but has not observed any evidence. The considerable volume of data accumulated from LHC Run 3 facilitates the development of novel methodologies that aim to effectively control the background in these investigations, thus enhancing signal sensitivity.

The CMS experiment has collected data corresponding to an integrated luminosity of 140 fb^{-1} during Run 2 operations (2016-2018) and several results on SM measurements and new physics searches have been presented [17]. More than 190 fb^{-1} data is already collected in Run 3 which started in 2018 and is expected collect another 100 fb^{-1} by end of 2025. Despite thorough searches, no concrete evidence for BSM physics has been found. Investigations into rare SM processes, like Higgs self-coupling, remain incomplete but could provide insights into novel physics.
The LHC will be upgraded for high luminosity operations (HL-LHC) [2], which is supposed to commence around year 2029 after a major long shutdown LS3. It aims to accumulate an integrated luminosity of 3000 fb^{-1} over a period of ten years with the instantaneous luminosity ranging from $5 \times 10^{34} cm^{-2} s^{-1}$ to $7.5 \times 10^{34} cm^{-2} s^{-1}$, enabling detailed studies of known physics phenomenon and potential discovery of new physics.
The physics motivation drives behind the HL-LHC encompass a wide range of factors. Di-Higgs generation, the scattering of longitudinal vector bosons, rare decay channels will be investigated, along with precise measurements of Higgs boson couplings. HL-LHC era could reveal new resonances or persistent particles predicted by various extensions of the SM, such as supersymmetry (SUSY), extra spatial dimensions, or strongly interacting dynamics, and could also help finding dark matter. Finding previously unreachable parameter spaces, like compressed supersymmetric partner mass spectra, will be possible.
Beam configurations required to achieve higher integrated luminosity will result in higher number of interactions per bunch crossing (pileup), which is expected to average 140 and could reach up to 200. To contrast, the Run 2 and Run 3 average pileup range are less than 40 and 60 respectively. To address these challenges, the CMS Collaboration is working on extensive upgrades [18] of its subdetectors. These include complete replacement of tracking system, end-cap calorimeters, electronics system of barrel calorimeters, and muon systems. Additionally, new technologies are being integrated, such as Gas Electron Multiplier (GEM) detectors for muon chambers and Minimum Ionizing Particle (MIP) timing detectors.
The collaboration has opted for a high-granularity calorimeter (HGCAL) to replace endcap electromagnetic and hadron calorimeters, $1.5 < |\eta| < 3.0$. It is a sampling calorimeter based on silicon sensors and scintillator tiles directly read out by silicon photomultipliers as layers of active materials sandwiched between layers of lead or stainless steel depending on the expected radiation dose of the detector region. For the first time, silicon is being used in the calorimetry at such large scale. The high granularity of the HGCAL presents a significant challenge in designing a readout system capable of handling the large data rates generated.
Chapter 2 provides an overview of SM, its successes and limitations, and SUSY as its potential extensions. The chapter elaborates on particle interactions in proton-proton collisions, the decay of short-lived particles, and the experimental techniques used to determine the four-momenta of these particles, with an emphasis on using

calorimeters to measure the energy of electrons, photons, and hadrons by fully stopping them.
Chapter 3 initially provides a brief overview of the LHC, detailing the experimental setups involved. We then encapsulate brief description of various subdetectors of CMS along with the methods of event reconstruction at CMS. Following this, attention is directed towards the CMS detector upgrades for Phase-1 and Phase-2, with an emphasis on HGCAL.
Chapter 4 explores foundational machine learning and neural network principles. As this thesis focuses on reconstructing charged-pion energies in HGCAL utilizing dynamic reduction network (DRN), an introduction to graph convolutional networks and DGCNNs is included along with an outline of the dynamic reduction network.
Chapter 5 presents the reconstruction of charged pions with a dynamic reduction network (DRN) based on GNN. This method significantly improves energy resolution compared to the weighted sum (WS) approach, validated by detailed GEANT4 simulations and 2018 beam test data. The chapter assesses the algorithm's performance in a calorimeter similar to the HGCAL detector, highlighting the DRN's adaptability and reliability over traditional methods in collider environments.
Chapter 6 provides an overview of HGCAL readout electronics system and studies conducted to understand the implication of data rates on configuring the readout system based on predictions from CMS event reconstruction software (CMSSW) implementation of the system. A short related study of data collected in beam test experiment in August-September 2023 is also presented which aimed to validate the full readout chain of the HGCAL.
The chapter 7 discusses a search for new physics within specific SUSY scenarios in pp collision events with energy levels of 13 TeV. This search focuses on events featuring at least one photon and large amounts of missing energy in multijet events, utilizing data from Run 2 (2016-2018), with an integrated luminosity of 137 fb^{-1}. This study significantly narrows the search phase space to explore new unexplored models using a refined approach that redefines search regions based on insights from the comprehensive dataset. This chapter details targeted physics scenarios, optimization of search regions, background estimation methods, and interprets the results concerning the limits on masses of SUSY particles.
Chapter 8 provides the summary of different studies presented in this thesis and their impact on the growth of the particle physics.

Studies and results presented in chapters 5, 6 and 7 are completely performed by me as a part of my doctoral thesis, with a due acknowledgment to the efforts from lots of CMS collaborators. In addition, to analysis of data and simulated data, I participated in several exercises in system tests in laboratory at IISER-Pune and at CERN getting hands on experience of working with detection and readout systems. A comprehensive glossary of acronyms and technical terms used throughout the thesis has been included at the end of the document for the reader's convenience. By addressing critical challenges in detector technology and analysis techniques, this work aligns with the overarching goals of experimental particle physics which are to test the limits of the Standard Model and uncover the fundamental principles governing our universe.

Chapter 2

Theoretical and experimental aspects of particle physics

This chapter outlines the various particles that make up the matter and carrier particles that mediate the interactions among them along with their theoretical description in the SM. This is followed by a short discussion on its successes and its limitation in explaining certain observed phenomena. We discuss potential extension of the SM, such as Supersymmetry (SUSY), alongside highlights from the searches for new particles conducted at the Large Hadron Collider (LHC) aimed at uncovering the unknown aspects of nature.

Experiments at colliders are designed to measure the SM parameters as well as discovery of the beyond SM (BSM) phenomena. In the second part of this chapter, we summarize interactions of the particles which are produced in the proton-proton (pp) collisions or from the decay of short lived particles produced in pp collisions with the detector materials. We discuss the experimental techniques employed to ascertain the four-momenta of these particles. In particular, the use of calorimeters to measure the energy of electrons, photons, and hadrons by fully stopping them is elaborated upon.

2.1 Elementary particles

Elementary particles are classified into two main categories: fermions, which constitute the matter, and gauge bosons, which mediate the various fundamental forces. Fermions follow Pauli exclusion principle and have half-integral spin. Fermions are further subdivided into two groups according to the forces they interact with: quarks and leptons, each having three distinct generation of particles. Gauge bosons follow Bose statistics and have integral spin. Each particle possesses distinct attributes, including mass, charge, and spin, which determine its interactions with other particles. Each of them also has a corresponding antiparticle with the same properties except an opposite charge. Neutral particles can be their own antiparticles (gluons and photons) or have distinct antiparticles (neutrinos). Elementary particles known as of today and encompassed in the SM are listed in Fig. 2.1.

Fermions - quarks & leptons

Six flavors of quarks are organized into three generations in order of increasing masses as up (u) & down (d), charm (c) & strange (s), and top (t) & bottom (b), carrying an electric charged of +2/3 & -1/3, respectively. The c, s, b, and t quarks have their own intrinsic quantum numbers, namely, charmness ($C = 1$), strangeness ($S = 1$), bottomness ($b = 1$), and topness ($t = 1$), respectively. Quarks and gluons carry an intrinsic quantum number called color charge which allows them to interact via strong forces [19]. The color charge comes in three flavors - red, green, and blue. Each quark (or anti-quark) exists with one specific color (or anti-color) quantum state and gluons carry a color and an anti-color simultaneously. Quarks can interact via

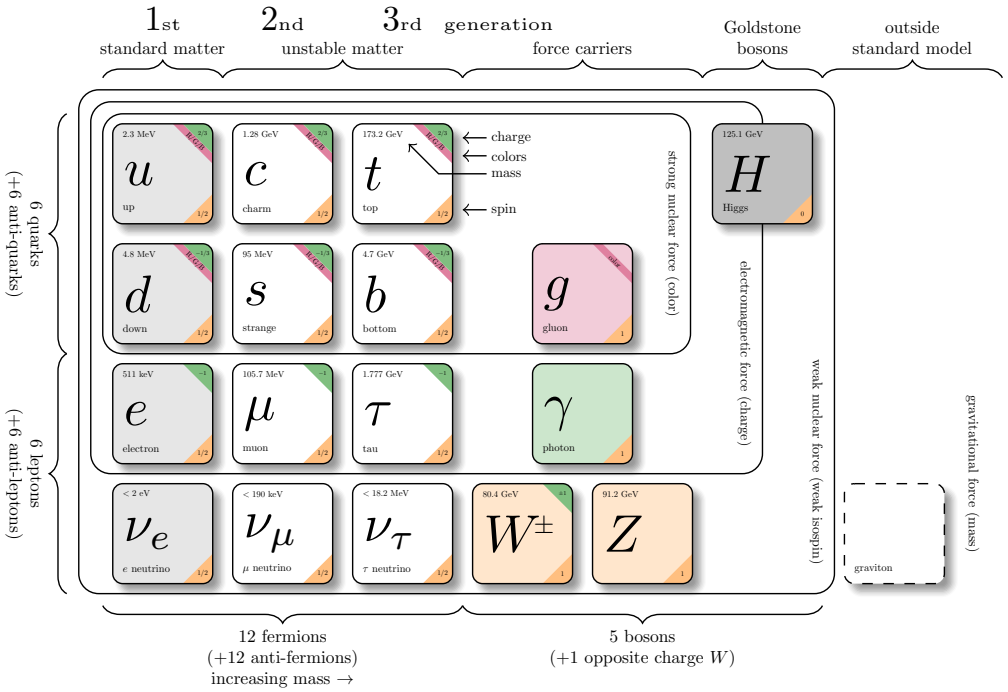


Figure 2.1: Table of elementary particles in the Standard Model, comprising quarks, leptons, and gauge bosons.[20]

strong, electromagnetic, and weak interactions, as they possess an electric charge as well.

Strong force coupling (strength) varies with distance or energy scale. It is weakly coupled at high energies (short distance) and strongly coupled at low energies (high distance). Thus, it shows asymptotic freedom letting quarks move freely within the bound states they form. Potential energy between a pair of quarks or quark-antiquark increase with distance between them such that production of new quark-antiquark pairs becomes more favorable, referred to as **color confinement**. Hence, quarks can not be isolated and exist only in colorless bound states called hadrons, except the top quark. Hadrons can be made of three quarks, called baryons e.g. protons and neutrons or a quark and an antiquark, called mesons e.g. pions and kaons. Except protons, hadrons are usually unstable. The top quark, on the other hand, decays before forming bound states because of its substantial mass. Furthermore, quarks are characterized by a baryon number of 1/3. All the quantum numbers are conserved throughout interactions, except for flavor, which is distinctively violated by weak interactions.

Analogous to quarks, there are six leptons within the SM organized into three generations - the electron and the electron neutrino (e, ν_e), the muon and the muon neutrino (μ, ν_μ), and the tau and the tau neutrino (τ, ν_τ). Each lepton has flavor quantum number, electron number (L_e), muon number (L_μ) and tau number (L_τ) to distinguish between generations of leptons, and these are conserved within the formulation of SM. Charged leptons (e,μ,τ) engage in electromagnetic and weak interactions, whereas neutrinos are limited to weak interactions only given they are electrically neutral. This characteristic makes them particularly challenging to investigate experimentally, as their interaction probability with detectors is notably

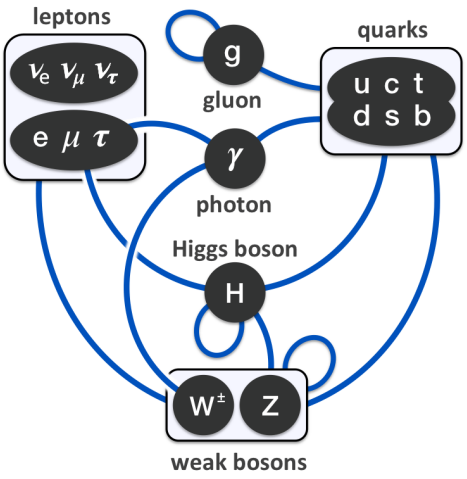


Figure 2.2: Elementary particles in the Standard Model are categorized within dark ovals, with blue arcs indicating interactions. `Image courtesy - Wikipedia`

low. The leptons do not engage in the strong force. All the leptons interact with the different gauge bosons in a similar way aside from mass-related differences. This is referred as lepton universality. Neutrinos are massless within the scope of SM but recent observations of neutrino oscillations indicate these have small but non-zero mass.

Bosons

Gauge bosons with spin-1, referred to as force mediators, facilitate the fundamental forces of nature. Photons (γ) mediate electromagnetic force, $W^{\pm}$ and Z^0 bosons mediate weak force, and gluons mediate the strong force. Photons and gluons are massless. The SM fermions and W/Z bosons acquire mass via interactions with the Higgs boson (spin-0). Photons and gluons are electrically neutral whereas gluons are colored objects carrying a combination of color and anti-color charges and comes in eight type. Fig. 2.2 summarizes the interactions among the elementary particles.

2.2 Standard model of elementary particles

Quantum Field Theory (QFT) [21, 22] forms the foundation of the SM of particle physics, integrating quantum mechanics with special relativity and classical field theory. In QFT, fields are fundamental and represent natural degrees of freedom, with particles as quantized excitations of these fields. Interactions among them considered as operators and their dynamics are mathematically encapsulated in Lagrangian. The SM is fundamentally based on three essential gauge symmetries, represented as $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ with C, L, and Y as color charge, weak isospin, and hypercharge, respectively. In the following, natural units ($c = \hbar = 1$) are used along with Einstein's summation convention, indicating that repeated indices in a term are summed. Greek letters serve as indices for spacetime. Also, "Lagrangian density" and "Lagrangian" may be used interchangeably.

2.2.1 Quantum electrodynamics

The quantum electrodynamics (QED) [23, 22] represents the theory of interaction of lights with the matter, essentially the electromagnetic fields. Lagrangian for free

electromagnetic field can be written as

$$\mathcal{L}_{\text{free}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{2.1}$$

where $F_{\mu\nu}$ is the strength of electromagnetic field and defined as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{2.2}$$

where A_μ is a massless four-vector field. To ensure that this field describes massless photons which have only two polarization states, we note that one degree of freedom is made redundant via the A_0 component of A_μ which is not dynamical. In addition, Lagrangian in equation 2.1 is invariant under the following transformation of vector field A_μ:

$$A_\mu\,(x) \rightarrow A_\mu\,(x) + \partial_\mu \lambda(x) \tag{2.3}$$

where λ is any function of x and expected to dies off quickly as $x \rightarrow \infty$. This is referred as $U(1)$ gauge symmetry. A quantum field theory to couple electromagnetic forces with matter or fermions is constructed using Dirac Lagrangian density ($\mathcal{L}_{\text{Dirac}}$), along with the condition of $U(1)$ gauge symmetry, defined as

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \tag{2.4}$$

Here, ψ is a fermion spinor with mass m, γ^μ refers to Dirac matrices, and $\bar{\psi}$ is $\psi^\dagger \gamma^0$, with $\psi^\dagger$ being the Hermitian conjugate of ψ.
Then the $\mathcal{L}$ for theory of electromagnetism coupling to fermions, can be written as

$$\mathcal{L}_{QED} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - e\bar{\psi}\gamma^\mu A_\mu \psi \tag{2.5}$$

where e is the charge. To make this Lagrangian invariant under the gauge symmetry from equation 2.3, a covariant derivative is defined as the following:

$$D_\mu = \partial_\mu \psi + ieA_\mu \psi \tag{2.6}$$

Then the Lagrangian becomes,

$$\mathcal{L}_{QED} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi \tag{2.7}$$

which is invariant under the following two simultaneous gauge transformations,

$$A_\mu \rightarrow A_\mu + \partial\lambda \quad \textit{and} \quad \psi \rightarrow \psi + e^{-ie\lambda}\psi \tag{2.8}$$

where λ serves as the generator of the $U(1)$ symmetry group.
It is noteworthy that a mass term for photons, such as $m^2 A_\mu A^\mu$ spoils the gauge invariance of the Lagrangian, hence can not be included. This gauge symmetry also leads to the conservation of the electric charge.

2.2.2 Quantum chromodynamics as non abelian gauge theory

Quantum chromodynamics (QCD) theory [22, 19] is a theory of strong force among quarks mediated by gluons characterized by $SU(3)$ group which has eight generators corresponding to eight distinct gluon mediators. It is a non-abelian gauge theory as its gauge fields do not commute, leading to self-interactions of the gluons. The QCD

Lagrangian can be written as the following:

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{2} Tr G_{\mu\nu} G^{\mu\nu} + \sum_{i=1}^{6} \bar{q}_i \left(i g \gamma^\mu D_\mu - m_i \right) q_i \tag{2.9}$$

where q represents a quark field and $G_{\mu\nu}$ are the gauge field strength which can be written in terms of gluon fields G_ν as

$$G_{\mu\nu} = \partial_\mu G_\nu - \partial_\nu G_\mu - i g_s [G_\mu, G_\nu] \tag{2.10}$$

$$G_{\mu\nu} = \partial_\mu G_\nu - \partial_\nu G_\mu - i g_s f^{abc} G_\mu G_\nu \tag{2.11}$$

where f^{abc} are SU(3) structure constants. The D_μ is the covariant derivative for quarks and defined as

$$D_\mu q^a = \partial_\mu q^a - i g_s (G_\mu)^a_b q^b \tag{2.12}$$

where g_s denotes the strong interaction coupling constant and index a (=1,2,3) denotes the color charge. With these definitions of G_μ and D_μ, we obtain three point and four point vertices for gluon field interactions in the QCD Lagrangian described in equation 2.9. Similarly to QED, there is no mass term added in the $\mathcal{L}_{\mathcal{QCD}}$ to maintain gauge symmetry, leaving the gluons massless.

2.2.3 Electroweak interactions

Non-conservation of parity in weak interactions leads to the inference that left-handed and right-handed fermions interact differently. Fermions are defined as left-handed (right-handed) based on their spin aligned in direction opposite (along) of their motion. Only left-handed fermions and right-handed anti-fermions participate in the weak interactions.

In the SM, the left-handed (right-handed) fermions are represented as $SU(2)$ doublet (singlet) of weak isospin charge T_3 equal $\pm 1/2$ (0). Electromagnetic interactions do not differentiate between right-handed and left-handed fermions. Hence, an unified electromagnetic and weak interactions framework can be structured with $\text{SU(2)}_L \otimes \text{U(1)}_Y$ symmetry group [24].

Doublet or singlet fields for the first generation of lepton family are represented as

$$\psi_1(x) = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L \tag{2.13}$$

$$\psi_2(x) = \nu_R \tag{2.14}$$

$$\psi_3(x) = e_R \tag{2.15}$$

The unified electroweak Lagrangian is formed as the following:

$$L_{\text{EW}} = -\frac{1}{4} W^i_{\mu\nu} W^{i\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} + \sum_{j=1}^{3} \bar{\psi}_j(x) i \gamma^\mu D_\mu \psi_j(x) \tag{2.16}$$

$\psi_i(x)$ denote the left- and right-handed fermion spinors and W^i and B are SU(2)_L and U(1)_Y gauge field strengths, respectively defined as following:

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g \epsilon^{ijk} W^i_\mu W^j_\nu \tag{2.17}$$

$$B_\mu \nu = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.18}$$

where g denotes the coupling constant for electroweak interactions, ϵ^{ijk} refers to the structure constant of the $SU(2)$ group. The Lagrangian specified in equation 2.16 retains its invariance under a gauge transformation wherein B remains unchanged and W undergoes a covariant transformation as follows:

$$B_\mu\nu \rightarrow B_\mu\nu \quad and \quad W_\mu\nu \rightarrow U_L W_\mu\nu U_L^\dagger \tag{2.19}$$

and the covariant derivative D_μ as presented in equation 2.16 is given by:

$$D_\mu = \partial_\mu - ig'\frac{Y^i}{2}B_\mu(x) - igW_\mu^a \tag{2.20}$$

The parameters g' and g denote the coupling strengths between fermions and these gauge fields. The combined SU(2) $\otimes$ U(1) symmetry ensures the conservation of weak hypercharge, defined as following;

$$Y = 2(Q - T_3) \tag{2.21}$$

where Y represents the weak hypercharge, Q denotes the electric charge.
The imposition of gauge symmetry prohibits a mass terms in the Lagrangian inherently leading to the massless gauge bosons, contrary to experimental observations [25, 26]. Similarly, the existence of fermionic masses is prohibited, as they would entail communication between the left- and right-handed fields, which possess distinct transformation properties, leading to explicit violation of gauge symmetry. Masses of these particles are incorporated via spontaneously symmetry breaking.

2.2.4 Higgs mechanism

The SM incorporates a scalar field, the Brout–Englert–Higgs (or Higgs) field, which breaks electroweak symmetry, granting mass to bosons and fermions [27, 28]. The $SU(2)_L \times U(1)_Y$ symmetry is spontaneously broken to $U(1)_{\text{em}}$ via the Higgs mechanism. The Higgs field ϕ is a complex scalar (spin-0) field represented by an $SU(2)_L$ doublet with hypercharge $Y = 1$ and defined as

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}$$

where ϕ_1, ϕ_2, ϕ_3, and ϕ_4 are real scalar fields.
The Lagrangian of the Higgs field is written as

$$L_{\text{Higgs}} = D_\mu\phi^\dagger D^\mu\phi - V(\phi) \tag{2.22}$$

where D_μ is the covariant derivative of $SU(2)_L \times U(1)_Y$ (see equation 2.20) and $V(\phi)$ is Higgs potential, defined as

$$V(\phi) = -\mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \tag{2.23}$$

The pictorial representation of the V(ϕ) for cases with $\mu^2 < 0$ and $\mu^2 > 0$ is shown in Fig. 2.3. For $\mu^2 > 0$, the potential has a nonzero vacuum expectation value (VEV):

$$\langle\phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}, \quad v = \sqrt{\frac{-\mu^2}{\lambda}}$$

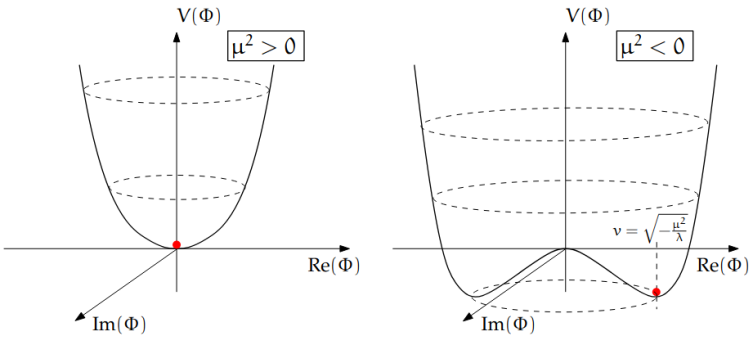


Figure 2.3: Representation of Higgs potential for different values of μ^2 [29].

To study the vacuum excitations, we introduce a quantum fluctuation $H(x)$ into ϕ_0 and expand it near its minimum:

$$\phi_0 = \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{2.24}$$

Inserting this form of ϕ_0 into the Lagrangian of equation 2.22, allows the Higgs potential to be expressed as:

$$V(H) = \frac{\mu^2}{2}H^2 + \frac{\lambda}{3}H^3 + \frac{\lambda}{4}H^4 \tag{2.25}$$

This elucidates mass of the Higgs boson is

$$m_H = \sqrt{2\mu^2} = \sqrt{2\lambda v^2} \tag{2.26}$$

Along with equation 2.20, the first term of equation 2.22 leads to mass term for gauge bosons

$$\mathcal{L}_{\text{mass}} = \frac{1}{4}v^2\left[g^2(W^1_\mu W^{1\mu} + W^2_\mu W^{2\mu}) + (gW^3_\mu - g'B_\mu)^2\right]. \tag{2.27}$$

With the linear combination of W^a_μ and B_μ, one can define the following gauge fields:

- A massless photon A_μ: $A_\mu = \cos\theta_W B_\mu + \sin\theta_W W^3_\mu$.
- A massive Z boson: $Z_\mu = -\sin\theta_W B_\mu + \cos\theta_W W^3_\mu$, with mass $M_Z = \frac{gv}{2\cos\theta_W}$.
- Massive $W^\pm$ bosons: $W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu)$, with mass $M_W = \frac{gv}{2}$.

Here, θ_W is the Weinberg angle, defined by $\tan\theta_W = \frac{g'}{g}$. The weak mixing angle θ_W is a free parameter, which is determined by precisely measuring the masses of the W and Z bosons.

Mass terms of the fermions via their interactions with Higgs field are given as $\mathcal{L}_{Yukawa}$. For quarks Yukawa interactions are given as

$$\mathcal{L}_{Yukawa} = -y^d\bar{Q}_L H d_R - y^u\bar{Q}_L\tilde{H}u_R + h.c.. \tag{2.28}$$

where y^u and y^d are Yukawa couplings for the down and up quarks, respectively. Similarly this term can be written for other massive fermions.

The Higgs boson is capable of self-interaction due to the presence of cubic and quartic terms of $V(H)$ defined in equation 2.25 and its coupling strength λ, is to be measured from the experiment.
With this, the standard model of electroweak interactions is represented by

$$L_{\text{SM}}^{EW} = L_{\text{EW}} + L_{\text{Higgs}} + L_{\text{Yukawa}} \tag{2.29}$$

where L_{EW} is for gauge and fermion interactions (equation 2.16), L_{Higgs} is for Higgs (equation 2.22), and L_{Yukawa} is for fermionic interactions with Higgs field.

2.2.5 Shortcomings of Standard Model

Although the SM predictions have been confirmed with data with very high precision, it comes with its limitations, mentioned briefly in the following.

- The SM does not provide a theoretical expression for the mass of the Higgs boson. However, it does incorporate loop corrections that entail quantum modifications arising from virtual particles. These corrections are divergent and extend to the Planck scale (10^{19}GeV). Experimentally, the mass of the Higgs boson has been measured to be approximately 125 GeV [9, 10]. The SM does not account for this discrepancy between the theoretical predictions and the experimental findings.
- Astronomical observations have confirmed that dark matter and dark energy constitute approximately 95% of the universe [30]. However, the SM does not provide an explanation for these phenomena.
- As of the present understanding, matter and antimatter should have been generated in equivalent amount during the Big Bang. However, observational evidence indicates a predominance of matter in the universe. The degree of CP violation accounted for within the SM is inadequate to elucidate this asymmetry, suggesting the necessity for new physics beyond the current framework.
- Neutrinos are considered to be massless in the SM framework. However, the observation of neutrino flavor oscillations [31] in recent experiments suggest that they have very small but nonzero masses.
- The SM does not incorporate a mechanism for gravitational interactions among the elementary particles. A successful quantum field theory of the gravity has not yet been established.

While the SM remains a fundamental pillar of contemporary physics, its limitations underscore the need for a more comprehensive theoretical framework. Experimental endeavors, such as those conducted at the LHC and prospective future collider facilities, strive to identify new phenomena that can address these gaps and advance our understanding of the fundamental laws of nature. Theoretical extensions of the SM, such as supersymmetry (SUSY), extradimensional theories, or models incorporating additional symmetries, aim to offer resolutions to several of these issues.

2.3 Supersymmetry

Supersymmetry (SUSY) [12, 32] represents a fundamental symmetry in space-time, delineating a relationship between fermions and bosons. Using the simplified picture and generator $\mathcal{Q}$, fermions can be converted into their bosonic superpartners,

$$\mathcal{Q}|Fermion\rangle = |Boson\rangle \tag{2.30}$$

Figure 2.4: One-loop quantum corrections to the Higgs squared mass parameter m_H^2 from a Dirac fermion f (left) and for a scalar S (right). [12]

The generator $\mathcal{Q}$ follows the following commutation relation with the four momentum generator (P_μ) of spacetime translation,

$$[\mathcal{Q}, P_\mu] = [P_\mu, \mathcal{Q}^\dagger] = 0 \tag{2.31}$$

Based on the above relation, each superpartner is expected to have mass equivalent to its partner along with other quantum numbers [33]. So far, we have not observed any bosonic partner of fermions (such as e μ, u & d quarks). This means that supersymmetry is a broken symmetry.

2.3.1 Advantageous features of Supersymmetry

SUSY is compelling due to various phenomenological reasons near the electroweak scale, complemented by strong technical justifications for its applicability across different energy scales. Some of them are listed below as majorly motivated by this course [33].

In Section 2.2.5, we briefly discussed the divergent corrections to the Higgs mass. Additional contribution to the correction in the Higgs mass through one-loop diagrams from SUSY scalar (fermion) particles cancels the corresponding divergent contributions from the SM fermion (scalar), as illustrated in figure 2.4. This is mathematically represented as

$$\Delta\, m_H^2 \approx \frac{1}{8\pi^2}(\lambda_S - \lambda_f^2)\Lambda_{UV}^2 + \frac{\lambda_S\, m_S^2}{8\pi^2} ln(\frac{\Lambda_{UV}}{m_S}) + ... \tag{2.32}$$

where $\lambda's$ represents coupling of fermions and scalars to Higgs field and Λ_{UV} is the ultraviolet cut off scale. If $\lambda_S \approx \lambda_f^2$, corrections to the Higgs mass depend only logarithmically on the scale of new physics, reducing the fine tuning for the Higgs mass. Hence, Supersymmetry intrinsically provides a solution.

The coupling to the three fundamental force couplings (g) evolve with energy scales according to the Renormalization Group (RG) equation,

$$\frac{4\pi}{g_i^2(\mu)} = \frac{b_i}{2\pi} ln \frac{\mu}{\Lambda_i} \quad with \quad i = 1, 2, 3 \tag{2.33}$$

suggesting that they unified at a very large energy scale such as 10^{15} GeV. If there are no new particles existing between electroweak and Planck scales, then these couplings only match approximately as shown in Fig 2.5 with dotted black lines. Within the SUSY framework, the unification of gauge couplings within uncertainties can be achieved [34].

The SM does not include a mechanism for elucidating dark matter. In many SUSY, the lightest supersymmetric particle (LSP), known as the neutralino, typically emerges as stable, neutral, and weakly interacting, thus presenting itself as a viable dark matter

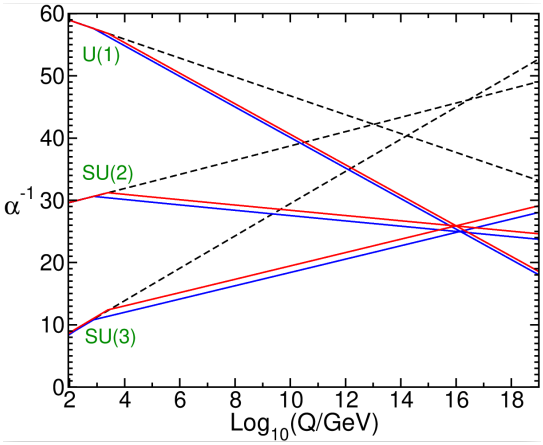


Figure 2.5: Examination of two-loop renormalization group behavior for inverse gauge couplings $\alpha_a^{-1}(Q)$ within the SM (dashed) versus the MSSM (solid). In MSSM analysis, sparticle mass thresholds range from 750 GeV to 2.5 TeV, with $\alpha_3(m_Z)$ ranging between 0.117 and 0.120. [12]

candidate. Several SUSY models, referred to as supergravity (SUGRA) models, invoke a spin 3/2 particle known as the gravitino, indicating a fundamental connection between local SUSY and General Relativity. This also presents a promising framework for the unification of all known interactions.

2.3.2 Minimal Supersymmetric extension of the Standard Model

The *Minimal Supersymmetric Standard Model* (MSSM) [12, 29] extends the SM by requiring a superpartner for each SM particle differing by spin-1/2, referred as sparticles. Superpartner interacts via electroweak and strong interactions with the same strength as their SM counterparts. The superpartner of quarks, gluon are referred squarks and sgluons, respectively. The leptons has corresponding superpartners named as sleptons. Each gauge boson, namely the photon, W, and Z, has a corresponding gaugino named as photino, wino and bino, respectively.

In the MSSM, the Higgs sector is expanded to include two Higgs doublets, H_u and H_d resulting in five Higgs bosons: two CP-even neutral Higgs bosons (h^0 and H^0), one CP-odd neutral Higgs boson (A^0), and a pair of charged Higgs bosons ($H^{\pm}$). These Higgs bosons arise from the mixing of the Higgs fields after electroweak symmetry breaking. The SM Higgs boson, which gives mass to the fermions, is a linear combination of the neutral components of these two Higgs doublets. The two-doublet structure ensures correct mass generation for up-type and down-type quarks and charged leptons. It also cancels potential gauge anomalies and is necessary for proper symmetry breaking in SUSY, maintaining model consistency.

The gauginos mix with the Higgsinos, the superpartners of the Higgs bosons, to form the mass eigenstates of neutralinos and charginos. There are four neutralinos ($\tilde{\chi}_1^0, \tilde{\chi}_2^0, \tilde{\chi}_3^0, \tilde{\chi}_4^0$) and two charginos ($\tilde{\chi}_1^{\pm}, \tilde{\chi}_2^{\pm}$) in the MSSM, which are the mass eigenstates resulting from the mixing of the gauginos and higgsinos. The particle content of MSSM is shown in Fig. 2.6.

R-parity

In the SM, baryon number (B) conservation prevents proton decay, as the proton is the lightest baryon. However, in MSSM, there is no fundamental symmetry to enforce the separate conservation of baryon and lepton numbers (L). Consequently, processes such as proton decay ($p \to e^+\pi^0$) are theoretically allowed, contradicting

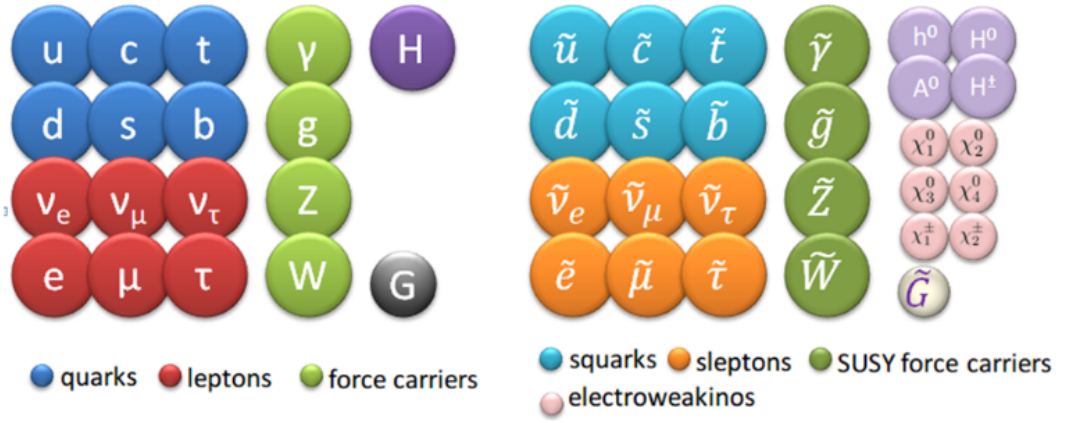


Figure 2.6: Different particles within the MSSM framework. [35]

experimental observations of proton stability with a decay limit of 10^{33} years.
To prevent such decays in the MSSM, a new quantum number, *R-parity* ($P_R = (-1)^{3(B-L)+2s}$), is introduced, where B, L and s represent baryon number, lepton number, and spin, respectively. SM particles have even P_R, while their supersymmetric (SUSY) partners have odd P_R.
The key consequences of conserving R-parity [13] include:

- SUSY particles are produced in pairs during proton-proton collisions.
- Decay chain of each SUSY particles must always include an odd number of SUSY particles.

2.4 Supersymmetry models examined in this thesis

SUSY is considered broken due to the absence of superpartners at expected mass scales, suggesting a breaking mechanism at the natural energy scale. SUSY particle masses in MSSM arise from adding terms to the Lagrangian that explicitly break supersymmetry, known as soft SUSY breaking. This can occur spontaneously, where the theory is inherently supersymmetric with a scalar field acquiring a vacuum expectation value, or explicitly, where the Lagrangian is modified. The added terms have a positive mass dimension, making them irrelevant in the ultraviolet (UV) context. Including the SUSY breaking terms, we have a total of 105 theory parameters in the MSSM scenarios. Methods of SUSY breaking such as gravity-mediated, anomaly-mediated, and gauge-mediated SUSY breaking (e.g., GMSB) reduce these parameters while maintaining gauge invariance. There are also the models, such as phenomenological MSSM (pMSSM) implementing constraints from astrophysics and particle physics experiments which significantly reduce the free parameters of the MSSM to be explored. For the SUSY search presented in this thesis, we have considered scenarios inspired by GMSB model and pMSSM for optimizing the search strategy and interpretation of our results. These are briefly discussed in the following.

2.4.1 Phenomenological Minimal Supersymmetric Standard Model (pMSSM)

The pMSSM is a simplified framework derived from general MSSM by applying a set of well-motivated assumptions that reduce the number of free parameters to 19 [36].

- It assumes absence of new sources of CP violation beyond those present in the SM. This implies that the only CP-violating phase considered is the one already present in the SM, which eliminates complex phases in the supersymmetric parameters that could otherwise introduce additional CP violation.

- The sfermion mass matrices are aligned with the corresponding fermion masses, which effectively reduces the number of parameters related to flavor mixing, restricting the flavor-changing neutral currents (FCNCs) to those present in the SM.
- It assumes the degeneracy of the first and second generations of sfermions and requires to have identical soft supersymmetry-breaking masses and trilinear couplings.

The pMSSM parameters include three mass parameters of the gauginos: M_1 (bino mass), M_2 (wino mass) and M_3 (gluino mass). The Higgs and Higgsino sector is characterized by three parameters: μ, the Higgsino mass parameter; $\tan\beta$, the ratio of the vacuum expectation values of the two Higgs doublets; and m_A, the mass of the pseudo-scalar Higgs boson. The sfermion sector includes parameters for the first and second generations, which are $m_{\tilde{Q}_{1,2}}$ (left-handed squarks), $m_{\tilde{u}_{R_{1,2}}}$ (right-handed up-type squarks), $m_{\tilde{d}_{R_{1,2}}}$ (right-handed down-type squarks), $m_{\tilde{L}_{1,2}}$ (left-handed sleptons) and $m_{\tilde{e}_{R_{1,2}}}$ (right-handed sleptons). For the third generation, the parameters include $m_{\tilde{Q}_3}$, $m_{\tilde{t}_R}$ (right-handed stop), $m_{\tilde{b}_R}$ (right-handed sbottom), $m_{\tilde{L}_3}$, and $m_{\tilde{\tau}_R}$ (right-handed stau). Finally, the model includes three trilinear coupling parameters: A_t for the top quark, A_b for the bottom quark, and A_τ for the tau lepton.
This 19-parameter framework retains the essential phenomenological features of the MSSM while being sufficiently constrained to allow for meaningful analyses of supersymmetry phenomenology. For a detailed discussion, refer to [36].

2.4.2 Gauge mediated symmetry breaking (GMSB) model

The GMSB [14] is a theoretical framework in which SUSY breaking is mediated by messenger fields, which couple to the SM fields through gauge interactions. This results in a suppression of flavor-changing neutral currents and a natural solution to the SUSY flavor problem. The SUSY breaking parameters encompass the masses of the messengers (M_{mess}), the gauge couplings or loop factor for Feynman diagrams incorporating gauge interactions ($\alpha_a/4\pi$), and the SUSY breaking scale ($< F >$), related to the soft mass term in the MSSM as

$$M_{soft} \approx \frac{\alpha_a < F >}{4\pi M_{mess}} \tag{2.34}$$

A significant characteristic of GMSB is its potential for a reduced SUSY breaking scale, denoted by $< F > \sim 10^4$ GeV, assuming this value is roughly comparable to the messenger mass scale, M_{mess}. Consequently, the masses of supersymmetric particles such as squarks, sleptons, and gauginos are expected to be in the range of a TeV, potentially detectable through production in pp collisions at the LHC, provided they exist in nature. These models are also characterized by distinct signatures at hadron colliders such as the production of weakly interacting particles like the gravitino.

Gravitinos : Gravitinos [12] ($\tilde{G}$), the spin-$\frac{3}{2}$ superpartners of gravitons, play a pivotal role in GMSB models. Their mass is directly linked to the supersymmetry-breaking scale (F) and is given by:

$$m_{\tilde{G}} \sim \frac{F}{\sqrt{3}M_P} \tag{2.35}$$

where M_P is the reduced Planck mass. In GMSB scenarios, with F typically in the range of 10^4 to 10^6 GeV, gravitinos acquire masses from a few eV to keV, making them

LSP and a strong dark matter candidate. Gravitino interactions are predominantly gravitational, leading to weak couplings. As a result, sparticles decay into their SM counterparts and a gravitino with relatively long lifetimes. The decay width of a sparticle ($\tilde{X}$) into a gravitino and a SM particle (X) is approximated by:

$$\Gamma(\tilde{X} \to X\tilde{G}) \sim \frac{m_{\tilde{X}}^5}{16\pi F^2}, \tag{2.36}$$

where $m_{\tilde{X}}$ is the sparticle mass. For sparticle masses near 100 to 1 TeV scale, this corresponds to lifetimes on the order of 10^{-2} seconds, allowing detection of decay products within collider experiments. GMSB frameworks, gravitinos, as LSP, act as the ultimate decay products of all sparticles. The process of acquiring mass involves absorbing the goldstino, which provides helicity $\pm\frac{1}{2}$ components, and results in both gravitational and inherited MSSM interactions. This unique nature makes gravitinos essential for both collider phenomenology and cosmological considerations.

Decay of Neutralino into Gravitino In GMSB models, the next-to-lightest supersymmetric particle (NLSP), often the lightest neutralino ($\tilde{\chi}_1^0$), plays a crucial role in collider phenomenology. The decay of $\tilde{\chi}_1^0$ into a photon (γ) and a gravitino ($\tilde{G}$) is characterized by the partial decay width:

$$\Gamma\left(\tilde{\chi}_1^0 \to \gamma\tilde{G}\right) = 2 \times 10\,\kappa_{1\gamma}\left(\frac{m_{\tilde{\chi}_1^0}}{100\,\text{GeV}}\right)^5\left(\frac{\langle F\rangle}{100\,\text{TeV}}\right)^{-4}\text{eV}, \tag{2.37}$$

where $\kappa_{1\gamma}$ represents the photino fraction of $\tilde{\chi}_1^0$, expressed as:

$$\kappa_{1\gamma} = |N_{11}\cos\theta_W + N_{12}\sin\theta_W|^2 \tag{2.38}$$

with N_{11} and N_{12} being elements of the neutralino mixing matrix $M_{\tilde{\chi}^0}$, and θ_W is the weak mixing angle. For $\tilde{\chi}_1^0$ masses exceeding 100 GeV and a supersymmetry-breaking scale $\langle F\rangle$ between 10^8 and 10^{12} GeV (corresponding to a gravitino mass $m_{\tilde{G}}$ of a few GeV), the decay width remains sufficiently large to ensure that $\tilde{\chi}_1^0$ decays within the detector.

The decay properties of $\tilde{\chi}_1^0$ depend on its composition, which is a mixture of bino, wino, and higgsino states [37, 38, 39]. This composition is determined by the parameters M_1, M_2, and μ:

- **Bino-like NLSP**: If $M_1 < \mu, M_2$, the neutralino is predominantly bino. In this case, $\tilde{\chi}_1^0$ decays primarily into $\gamma + \tilde{G}$.
- **Wino-like NLSP**: For $M_2 < \mu, M_1$, the neutralino is predominantly wino and closely related to the chargino. The neutral wino decays into $Z + \tilde{G}$, while the charged wino ($\tilde{\chi}^\pm$) decays into $W^\pm + \tilde{G}$.
- **Higgsino-like NLSP**: If $\mu < M_1, M_2$, the neutralino is primarily higgsino. Its decay channels depend on the sign of μ and the value of $\tan\beta$:
  - For $\mu > 0$ and low $\tan\beta$ (~ 2), $\tilde{\chi}_1^0$ decays into $Z/\gamma + \tilde{G}$.
  - For $\mu < 0$ and low $\tan\beta$, the decay predominantly involves $H/\gamma + \tilde{G}$.
  - At moderate to high $\tan\beta$, the decay channels include a combination of $Z/H/\gamma + \tilde{G}$.

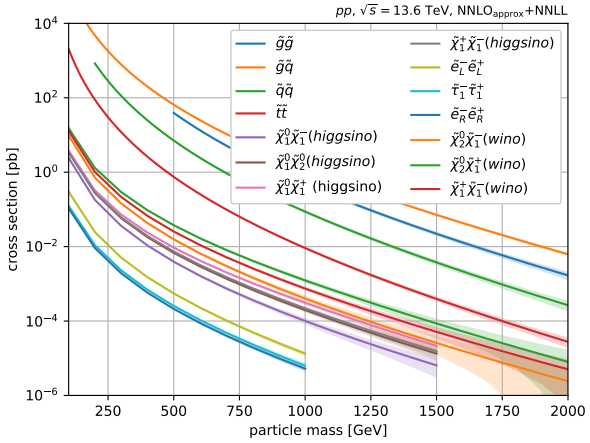


Figure 2.7: Cross sections for various subprocesses at 13.6 TeV with analyzed simplified topologies. [43]

These decay patterns provide distinct experimental signatures, enabling the identification of the NLSP's nature and the underlying supersymmetric parameters in collider experiments.

2.5 Simplified SUSY models at LHC

The phenomenology of SUSY is governed by the mechanism and scale of SUSY breaking, which influence the masses, mass hierarchies, and decay patterns of SUSY particles. The precise mechanism of SUSY breaking remains unidentified. Due to unknown interaction couplings and masses, predicting exact decay modes is challenging. Conducting a phenomenological analysis of SUSY searches with all parameters unrestrained is impractical.

SUSY particles decay through multiple modes, creating a variety of signatures in the detector. Simplifying the search by focusing on specific decay modes or small sections of phase space enables easier interpretation of the results. This motivates simplified models spectra (SMS), generically referred to as simplified models [40, 41, 42]. For each simplified model, a pair of SUSY particles are produced which decay to specified SUSY particles with predefined branching fractions. Other than the SUSY particles involved in the production and decays, all other particles are assumed to be inaccessible at the LHC (masses assumed to be 100 TeV).

Models typically considered in simplified analyses include the pair production of gluinos ($\tilde{g}$) or squarks ($\tilde{q}$), electroweakinos (both neutralinos($\widetilde{\chi}^0$) and charginos ($\widetilde{\chi}^{\pm}$)), and sleptons ($\tilde{l}$) and their cross sections at $\sqrt{13}TeV$ are shown in Fig. 2.7. The LHC's gluon-rich environment enhances the production of strongly interacting SUSY particles, such as gluinos ($\tilde{g}$) and squarks, particularly stops ($\tilde{t}$) and sbottoms ($\tilde{b}$). The charginos ($\widetilde{\chi}^{\pm}$) and neutralinos ($\widetilde{\chi}^0$) are produced via electroweak interactions with a smaller relative cross section owing to their couplings.

Events within SUSY frameworks are characterized by complex final states and significant missing transverse momentum ($p_{\mathrm{T}}^{\mathrm{miss}}$) due to the undetected LSP in R-parity-conserving scenarios.

Based on the particle production in these models, different final states can arise in the detector. For example, strong production channels involving gluinos and squarks often result in multijet events with significant imbalance in transverse momentum ($p_{\mathrm{T}}^{\mathrm{miss}}$) with and without a number of leptons and/or photons. The electroweakino

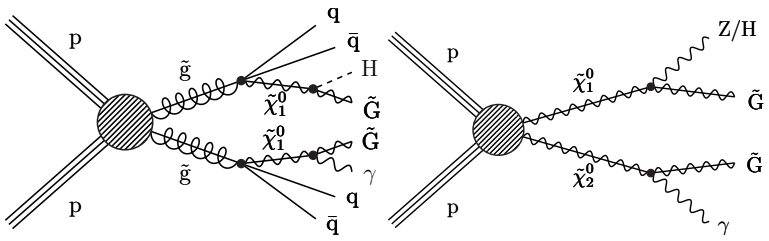


Figure 2.8: Simplified model of Gluino pair strong production (T5bbbbZG) (left) and direct production of neutralinos(right).

signatures are generally characterized by large number of leptons in the final state originating from decays of W, Z and H bosons. Boosted massive bosons can also be identified in their hadronic decay channels using dedicated algorithms. In GMSB, depending upon the $\tilde{\chi}_2^0$ compositions as discussed in Section 2.4.2, final state can be composed of different particles containing γ, W, Z, and H bosons along with a gravitino. Fig. 2.8 illustrates two of the simplified models with gluino pair production (left) and electroweakino production (right) considered in the SUSY search described in Chapter 7. We also considered the pMSSM scenarios in which soft photons can emerge as a part of the final state, especially when the mass difference between NLSP ($\tilde{\chi}_2^0$) and $\tilde{\chi}_1^0$ is small, enhancing the probability of soft photons as a decay product. Details of all the simplified models and relevant information are provided in the same chapter.

2.6 Particle interaction with matter

In pp collisions, most SM or hypothetical new physics particles produced are unstable and transform into lighter particles, stable enough to reach the detectors. Particles that are detected and reconstructed by the experiments positioned at the collision points include electrons, photons, muons, charged and neutral hadrons. Quarks and gluons manifest themselves as particle showers in the detector, primarily comprising charged and neutral hadrons, known as jets. Using these particles, reconstruction of massive unstable particles such as W/Z, and H bosons, top quarks is done. Understanding particle interaction with the matter are crucial to understand their behavior as well as to quantify their unwanted effects to accurately reconstruct the event. These particles can engage with materials via electromagnetic, weak, or strong interactions. Based on these specific type of particle-material interaction, a range of detection methodologies is employed in these experiments. This section offers an overview of the mechanisms by which these particles lose energy in materials and the various experimental methods employed for their identification.

2.6.1 Electromagnetic interactions of particles with material

When charged particles or photons pass through the material, they interact with electrons and nuclei of the atoms via Coulomb interactions and lose their energies.

Charged particles

Energy loss by charged particles per unit length while traversing a material via Coulomb interactions with atomic electrons, resulting into the ionization, or excitation of electrons to higher states. The energy loss as a function of depth traverse through the material is governed by Bethe-Bloch equation [1]. For a relativistic particle with charge z, velocity $\beta = v/c$, and Lorentz factor γ, the energy loss is given by:

$$-\frac{dE}{dx} = \frac{4\pi\alpha^2\hbar^2}{m_e} \cdot \frac{z^2}{\beta^2} \cdot \rho N_A \cdot \frac{Z}{A} \cdot \left[\ln\left(\frac{2m_e c^2 \beta^2 \gamma^2}{I} \right) - \beta^2 - \frac{\delta}{2} \right] \tag{2.39}$$

where Z and A are the atomic and mass numbers, m_e is the electron mass, N_A is Avogadro's number, ρ denotes the medium's density, the I represents the mean ionization potential of the material, adjusted by δ to address reduced ionization energy loss at high energies, particularly in dense media. These adjustments enhance the precision of energy loss predictions. Figure 2.9 depicts the average energy dissipation rate normalized with density (in units of $\mathrm{MeVg^{-1}cm^2}$) across different materials, based on $\beta\gamma$ on the X-axis, in terms of particle momentum. There are quite few features extracted from the Bethe-Bloch expression:

- The primary dissipation of energy occurs via collisions with the electrons.
- The energy losses are independent of mass of the incoming particle.
- The rate of ionization energy loss is greater for slower particles suggested by the $1/\beta^2$ factor.
- Additionally, energy loss is directly proportional to the density of the medium. Generally, dE/dx is presented normalized to the material density and has the units of $\mathrm{MeVg^{-1}cm^2}$. Presented this way, the energy loss per unit length is similar for different materials.
- The greater the electric charge of the incoming particle, the more significant is the energy loss.

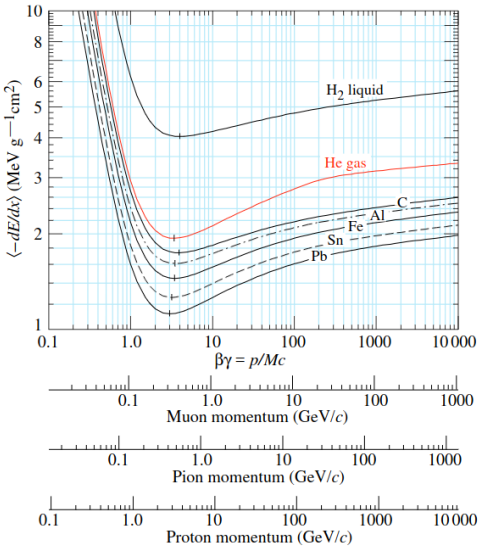


Figure 2.9: The mean energy loss (dE/dX) as a function of $\beta\gamma$ in different materials. X-axis scale is represented as a function of momentum for different particles [1].

- At very high energies, where β approaches 1, the logarithmic component on the right-hand side contributes to an increase in ionization energy loss, referred to as the 'relativistic rise'.
- The ionization energy loss is minimum around a $\beta\gamma \approx 3-4$ and has a value between $1-2$MeV g^{-1}cm^2 in all materials for singly charged particles. These particles are known as "minimum-ionizing particles" (MIPs).

Radiation

In the presence of an electric field of the nucleus, high-energy charged particles are accelerated or decelerated, resulting in emission of electromagnetic radiation, called Bremsstrahlung. The Feynman diagram is shown in Fig. 2.10 (left). The probability of bremsstrahlung radiation for a particle with mass m is given by:

$$\sigma \propto \frac{Z^2}{m^2} \tag{2.40}$$

The energy loss by particle with energy E due to bremsstrahlung is defined as

$$\frac{-dE}{dx} \propto \frac{Z^2}{m^2} E \tag{2.41}$$

where Z is the atomic number of the material. Both the cross section and the energy loss dE/dx are proportional to Z^2 and inversely proportional to the mass (m^2) of the particle. The photon energy distribution exhibits a decrease proportional to $1/E$. The traversing particle deviation from its path is influenced by the angle and energy of the emitted photon, which depend on the Coulomb field intensity related to Z.
From equations 2.40 and 2.41, for electrons at high energies, bremsstrahlung is particularly significant because of their low mass. Conversely, the cross section of bremsstrahlung for muons is approximately 40,000 times smaller than that for electrons.

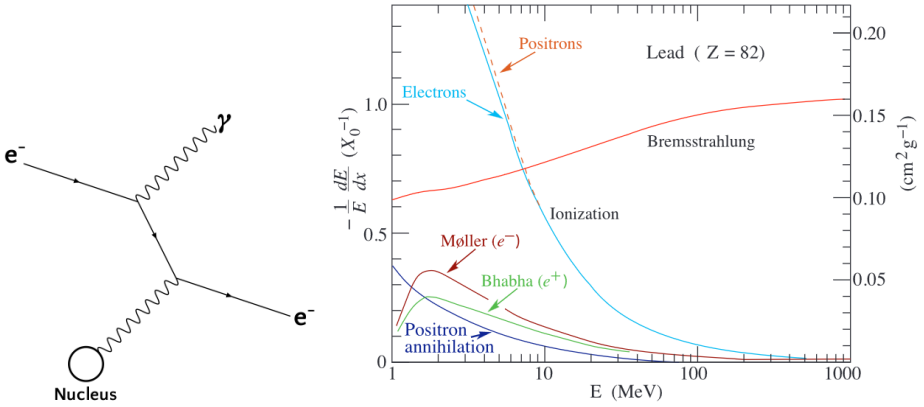


Figure 2.10: Feynman diagram on the left illustrates the Bremsstrahlung mechanism. On the right, the plot shows the fractional energy loss per radiation length in lead as a function of the electron's energy. [1]

Radiation length (X_0) and critical energy (E_c) are two key parameters of the material, crucial for detector design, and they are defined as,

Radiation Length (X_0) : It is the mean distance over which a high energy charged particle, such as an electron, reduces its energy by a factor of $1/e$ through radiation. It is a material dependent quantity depending on the atomic number Z of the material,

$$X_0 \propto \frac{A}{Z^2} \tag{2.42}$$

Critical Energy (E_c) : It is defined as the energy at which energy loss by a charged particle due to radiation becomes equal to its energy loss due to ionization. For electrons, it is approximately given by [20]:

$$E_c \approx \frac{610\,\text{MeV}}{Z + 1.24}, \tag{2.43}$$

For Uranium, E_c is ∼9 MeV, while for Carbon it is ∼95 MeV. Hence, E_c is small for high Z materials. Consequently, particles traversing in high Z material experience radiation losses until the later stages of its trajectory.

Fig. 2.10 (right) shows the mean energy loss per radiation length in lead as a function of incident electron energy through different processes. At energies above 100 MeV, the electrons (& positrons) mainly loses their energy via radiation. At very low energies, Møller Scattering for electrons and Bhabha Scattering and annihilation for positrons also play a negligible role in energy deposition.

Muons with energies around $(10 - 100)$ GeV typically behave as Minimum Ionizing Particles (MIPs) within the detector as they do not significantly radiate owing to their large mass and they do not participate in strong interactions. Therefore, substantial quantities of absorber materials are essential to effectively shield against muon penetration.

In addition, if the incident particle speed exceeds the speed of light in a medium, the particle can emit photons and lose its energy. This phenomenon is called Cherenkov radiation. The radiation is emitted at a characteristic angle, which depends on the speed of the particle and the refractive index of the medium. This angle, θ, is given by the relation:

$$cos\theta = \frac{c}{vn} \tag{2.44}$$

where c is the speed of light in a vacuum, v is the particle's speed, and n is the refractive index of the medium. This radiation can be utilized to identify particles and measure their speed.

2.6.2 Interaction of photons with matter

High energy photons lose their energy in the material mainly via photoelectric effect, compton scattering, and pair production.

Photoelectric effect

In this process, a photon is absorbed by an atom, resulting in the ejection of a photoelectron. The photon energy must exceed the material's work function ($W_p \approx$ a few eV) to eject this electron. The excess photon energy is transferred to the photoelectron as kinetic energy.
The cross section for the photoelectric effect is highly dependent on the atomic number Z, scaling approximately as Z^5, and inversely on the photon energy E_γ: Hence, the photoelectric effect is dominant in high-Z materials such as lead (Pb) or uranium (U) for low-energy photons. As the photon energy increases, the importance of this process diminishes rapidly as shown in Fig. 2.11.

Compton scattering

In this, a photon transfers part of its energy and momentum to an atomic electron, ejecting the electron. The cross-section for this process depends on Z and its energy E_γ as

$$\sigma \propto \frac{ln(E_\gamma).Z}{E_\gamma} \tag{2.45}$$

While the photoelectric effect dominates at low energies, Compton scattering becomes significant at the photon energies of a MeV to few tens of MeV as shown in Fig. 2.11.

Pair production

At photon energies exceeding twice the electron rest mass ($2m_ec^2 \approx 1.02\,\text{MeV}$), pair production becomes possible where it interacts with the electromagnetic field of a nucleus or an electron and creates an electron-positron pair. Pair production can not occur in a vacuum because momentum conservation requires a recoil object.
For energies beyond $2m_e$, the probability of pair production is directly proportional to the atomic mass (A), and inversely proportional to the radiation length (X_0) of the material. Using a material with higher atomic numbers reduces the energy threshold for pair production, making it the primary way photons dissipate energy into electron-positron pairs. Mean free path of photons before completely disappearing is defined as,

$$L_{mean} \approx \frac{9}{7} X_0 \tag{2.46}$$

Consequently, a photon traverses an extra radiation length than that of an electron when traversing the same distance within the material. Electrons lose energy continuously through ionization and radiation, whereas photons can penetrate matter without interacting and then vanish through pair production or photo electric effect or get scattered away via compton scattering.

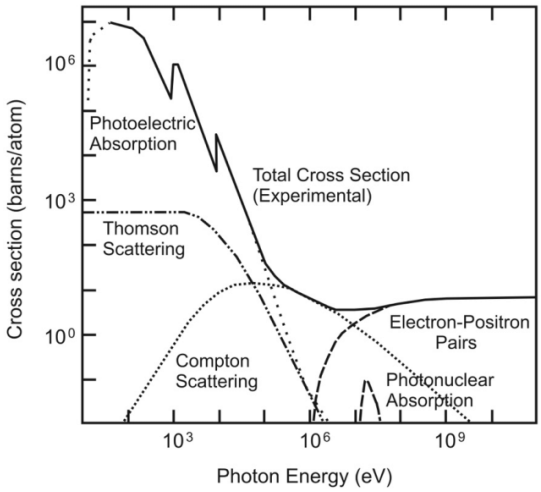


Figure 2.11: Dominant photon interaction processes as a function of incident photon energy [44].

2.6.3 Energy loss via strong nuclear interactions

High energy hadrons interact with nuclei mediated by the strong nuclear force and can result into elastic or inelastic scattering. In elastic interactions, the incident hadron scatters off the nucleus, transferring part of its kinetic energy while continuing along a deflected trajectory. The deflection angle depends on factors such as energy transfer and the impact parameter, making the process inherently stochastic. Inelastic scattering or nuclear spallation is more complex and involves the production of secondary particles. These processes occur in two stages [45]:

- **Fast intra-nuclear cascade :** The incident hadron collides with a nucleus, transferring energy to nucleons, which then interact with other nucleons in a rapid cascade. During this, pions and other unstable hadrons may be created if enough energy is transferred. Some nucleons may escape the nucleus if they overcome the nuclear binding energy, while others get caught and distribute their kinetic energy to other nucleons. The energy spent overcoming the binding energy is lost and does not contribute to the detector signal. This contribution can be up to 30-40% of the energy fraction involving strong interactions.
- **Nuclear evaporation :** Nucleons that lack sufficient energy to escape excite the nucleus. As the nucleus deexcites, it emits particles such as neutrons, photons, and alpha particles. Heavy nuclei, such as uranium, may also undergo fission during this stage. The probability of producing spallation products (Z_i, A_i) is defined as approximately

$$\sigma(Z_i, A_i) \sim exp(-P(A_T - A_i)) \times exp(-R|Z_i - SA_i + TAMA2_i|3/2) \quad (2.47)$$

  where P, R, S, T are the parameters, A_T is the atomic mass of the material, (Z_i, A_i) are the atomic number and mass of the spallation products.

A large variety of spallation reactions ($\sim$300) are possible and they are of similar probabilities. The maximum cross section for a single reaction is not more than 2% while most of the reactions contribute in the order of 0.1% of the total spallation cross section [45].

During evaporation, more neutrons are produced than protons, because protons need to overcome the Coulomb barrier. The production of neutrons is more in high Z materials, such as lead, and the typical energy is of the order of 10 MeV. Energetic

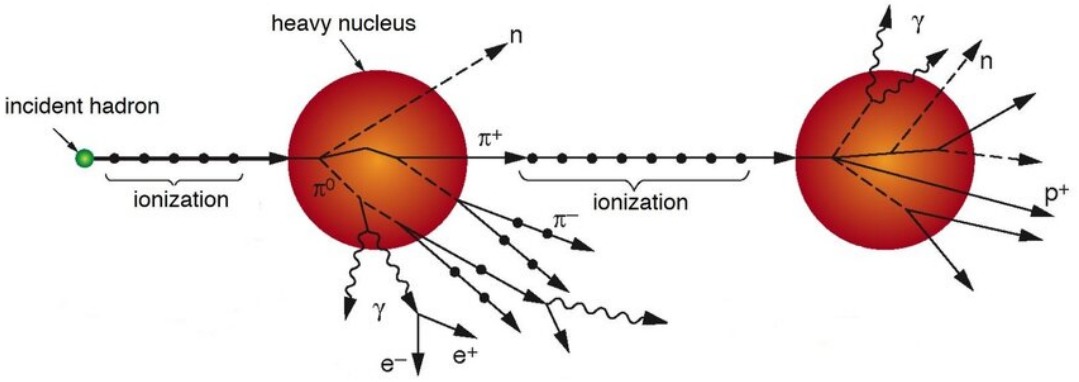


Figure 2.12: Illustration of hadronic interactions with material [46].

evaporate nucleons can further escalate the spallation by creating more neutrons. Neutrons can also be absorbed by nuclei, forming intermediate states that can emit γ -rays, charged particles (e.g., protons or alpha particles), or additional neutrons. Low-Z materials, such as hydrogen, are effective in slowing neutrons rapidly as neutrons collide elastically with heavier nuclei without losing much energy.
Nuclear interaction length λ_{int} is defined as the mean distance a hadron travels in a material before interacting with a probability of 63%. It is expressed as:

$$\lambda_{\text{int}} \propto A^{1/3}, \tag{2.48}$$

where A is the atomic mass of the material. Although it is typically measured for protons at 200 GeV, λ_{int} also applies to other hadrons, with differences in σ_{total} affecting interaction probabilities. Table 2.1 lists the values of λ_{int}, X_0, E_c and R_m (Moilere radius, defined in the next sections) for different materials.

Table 2.1: Properties of materials relevant to electromagnetic and hadron interactions [1].

Material	Atomic mass (A)	Atomic number (Z)	λ_{int}(cm)	X_0(cm)	E_c(MeV)	R_m(cm)
Lead (Pb)	207.2	82	17.59	0.56	7.43	1.60
Iron (Fe)	55.84	26	16.77	1.75	21.68	1.71
Uranium (U)	238.0	92	11.03	0.31	6.65	1.00
Copper (Cu)	63.54	29	15.32	1.43	19.42	1.56
Tungsten (W)	183.84	74	9.94	0.35	7.97	0.93

2.7 Particle cascades in the material

Particles produced in high-energy collisions, such as those occurring at the LHC, are very energetic and do not immediately dissipate their energy in an electromagnetic or strong interaction. Instead, they create a cascade of secondary particles, which in turn generate lower-energy particles that discharge their energy through various mechanisms as described in previous section. The resulting showers from these particles are divided into two categories based on the interactions: electromagnetic

(EM) showers and hadronic (Had) showers. Understanding the distinct characteristics of these showers is crucial for selecting appropriate detectors. Most of the information discussed here is adopted from [45].

2.7.1 Electromagnetic showers

High-energy electrons and photons initiate cascades of secondary particles when they interact with a material. These shower particles continue to interact with the material, generating an additional cascade of particles. This process, illustrated in Fig. 2.13 (left), leads to a rapid increase in the number of particles, with each interaction approximately doubling the shower multiplicity. Hence, relatively less material is required for full containing the electromagnetic showers due to their high particle multiplicity.

This particle multiplication continues until the energy of the photons and electrons drops below the pair production threshold and the critical energy E_c, respectively. Beyond these energies, low-energy processes such as the photoelectric effect and Compton scattering for photons and ionization for electrons take over. This electromagnetic (EM) shower develops both longitudinally and laterally.

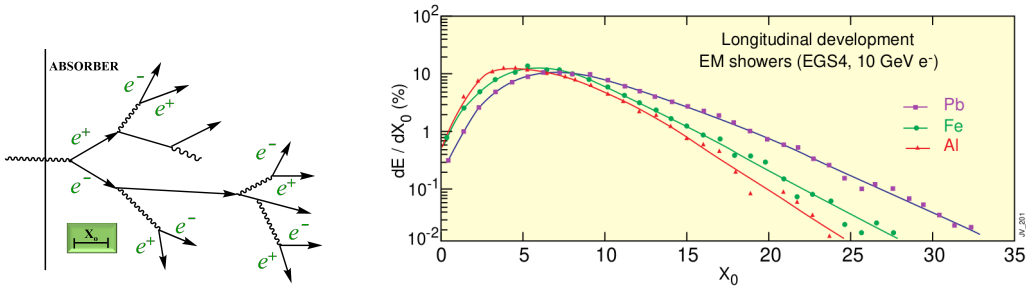


Figure 2.13: (Left) Schematic of electromagnetic shower development. (Right) Longitudinal energy deposition profiles of EM showers for different materials (see text for more details) [47].

Figure 2.13 (right) shows the longitudinal energy deposition profile of EM showers in lead (Pb), iron (Fe) and aluminum (Al). The energy deposition initially increases, reaching a maximum, and then decreases exponentially with depth. The differences between materials arise from their atomic properties; for instance, higher-Z materials (such as Pb) sustain particle multiplication to lower energies and exhibit a slower decline in particle production beyond the shower maximum. For example, a high-energy electron produces approximately three times more positrons in lead than in aluminum, requiring more radiation lengths (X_0) of lead to achieve 99% containment. Consequently, when considering *cm*, the depth of lead required to contain EM showers is significantly less than that of iron or aluminum. This reduction is attributed to the fact that lead has a radiation length that is at least twice as small and approximately 0.56 *cm* (see Table 2.1 for additional materials).

The depth of the shower maximum increases logarithmically with the energy of the incident particle, as shown in Fig. 2.14 (left). This logarithmic relationship implies that doubling the energy of the incoming particle does not double the material required for containment. For example, about 28 X_0 is sufficient to contain more than 99% of the energy from a 1 TeV electron-initiated shower.

The lateral spread of electromagnetic showers shown in Fig 2.14 (right), is influenced by two main processes: multiple scattering of electrons and positrons which causes them to deviate from the shower axis and isotropic processes such as compton scattering and the photoelectric effect which also contribute to the transverse spread. The first process dominates in the early stages of the shower, while the second

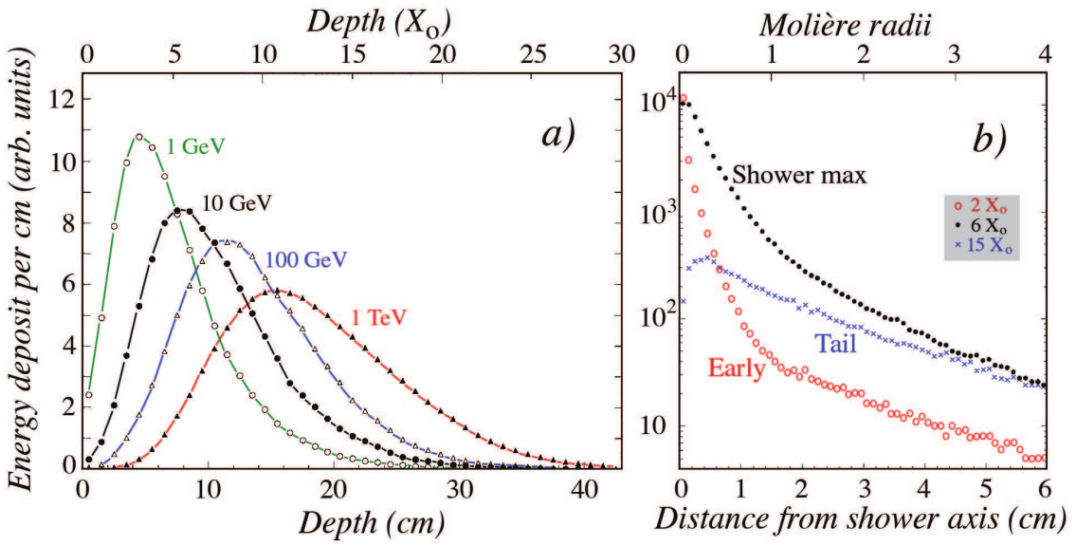


Figure 2.14: Longitudinal energy profiles of electromagnetic showers with different energies (left). The energy distribution profiles from 10 GeV electron showers in copper at different penetration depths (right) [45].

becomes significant beyond the shower maximum. Both processes have characteristic exponential scales. The Molière radius (ρ_M) defines the radius of a cylinder that contains 90% of the lateral shower energy. When expressed in g/cm^2, ρ_M scales as A/Z, in contrast X_0 scales as A/Z^2. Unlike longitudinal containment (scales with X_0), the lateral containment of electromagnetic showers shows minimal material dependence and is energy independent.

The relationship between energy and the required calorimeter thickness to contain electromagnetic showers is illustrated in Fig. 2.15. This implies that a sufficiently long cylinder will contain the same fraction of energy from showers initiated by 1 GeV and 1 TeV particles,

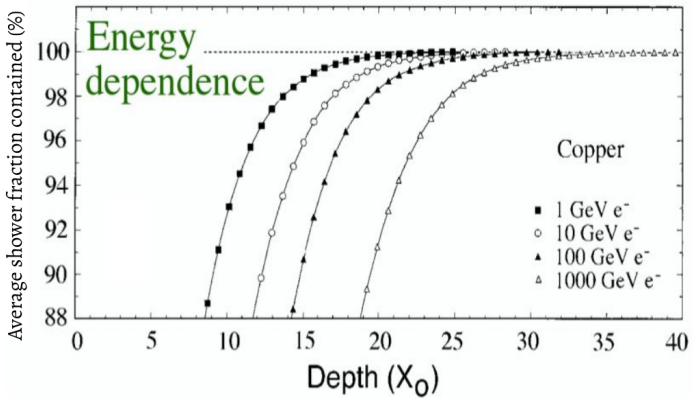


Figure 2.15: The average energy fraction retained in a material block with infinite transverse dimensions as a function of absorber thickness for different electron energy levels. [45].

High-energy electrons and photons interact differently with material. Electrons radiate bremsstrahlung photons almost immediately upon entering the material, potentially emitting thousands of photons within a few millimeters. In contrast, photons may not interact immediately, retaining their energy until they undergo pair production. When pair production occurs early, photons can deposit as much or more

energy than electrons over the same distance. Hence, the material depth to contain electron or photon induced showers is sightly different.

2.7.2 Hadronic showers

Unlike EM showers, hadronic showers are significantly more complex due to the strong interaction and the wide variety of nuclear processes involved. As discussed in Sect. 2.6.3, hadrons lose their energy through ionization as well as through strong interactions. A hadron produces secondary hadrons interacting with the nucleus which can further engage in such nuclear interactions, initiating a cascade of particle multiplication similar to an electromagnetic shower (EM) as shown in Fig. 2.16. This process continues until the energy of the particles falls below the threshold (typically around 1 GeV) required to produce additional secondary lightest hadrons such as pions [45]. Hadronic showers have fewer shower particles than electromagnetic showers because the energy needed to produce new particles in EM showers is about 10 MeV in comparison to 1 GeV required to produce hadrons. Predominantly charged and neutral pions ($\sim$ 90%) are produced as secondaries, as they are the lightest hadrons [45].

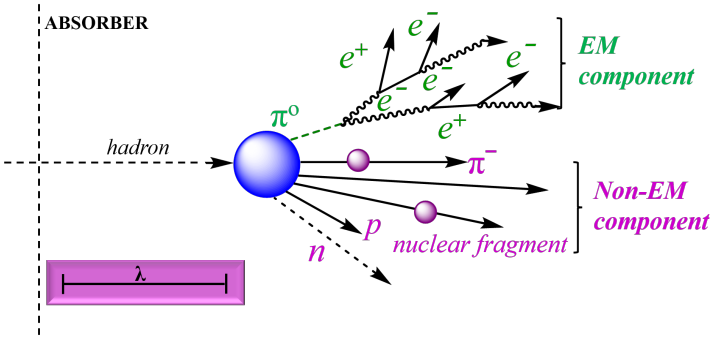


Figure 2.16: Schematic representation of a hadronic interaction in a shower.

Based on the nature of secondary hadrons produced, hadronic showers have two components: electromagnetic and hadronic. The π^0 and η^0 mesons produced during hadronic interactions immediately decay into photons, which deposit their energy through EM showers, as discussed in Section 2.7.1. The fraction of energy carried by the EM component (f_{em}) varies event by event and typically increases logarithmically with the energy of the incident hadron, ranging from about 30% to 50% and is shown in Fig. 2.17. The f_{em}'s dependence on energy can be parametrized as,

$$\langle f_{\text{em}} \rangle = 1 - \left[\frac{E}{E_0}\right]^{k-1} \tag{2.49}$$

where E is the incident hadron's energy, E_0 is the average energy needed for pion production ($E_0 \approx 0.7$ and 1.3 GeV in copper and lead, respectively), and k_1 relates to the average pion multiplicity in the cascade ($k \approx 0.82$ for both materials).

The non-EM component originates from nuclear interactions where hadrons are ejected from the nuclei. The nucleon binding energy does not generate signals in the calorimeter, leading to invisible energy, unlike EM showers where energy is fully deposited. This invisible fraction can account for up to 40% of the total shower energy [45]. These nucleons deposit energy through ionization, with protons typically carrying 50 to 100 MeV. Ionizing particles contribute to $\sim$56% of the deposited energy, while neutrons contribute about 10% per GeV. The non-EM component is

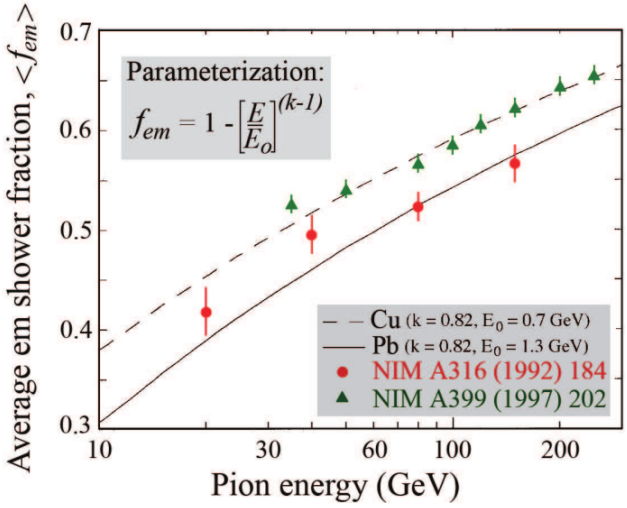


Figure 2.17: The variation of f_{em} with respect to the energy of hadrons in hadronic showers [45].

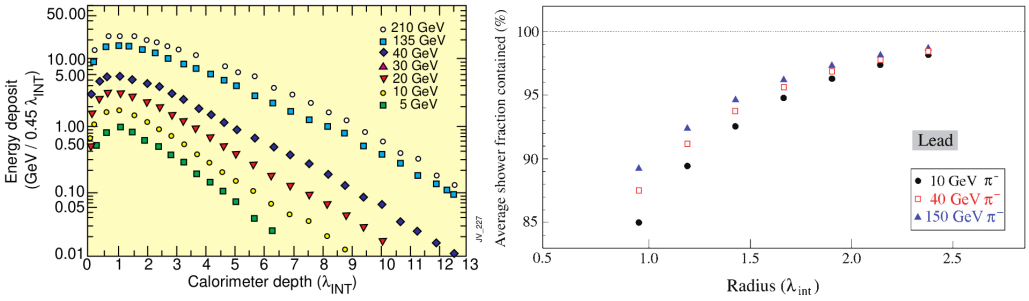


Figure 2.18: Left : Longitudinal profile of hadronic showers of different energies in Lead. Right : The average shower fraction contained in lateral as a function of λ_{int}, shown for different energies [45].

characterized by significant energy fluctuations and a broader energy deposition profile compared to the EM component. Hadronic showers exhibit large event-to-event fluctuations, due to the formation of neutral pions as the shower develops. These neutral pions can be produced at different points throughout the shower with varying energies. Figure 5.9 of Chapter 5 illustrates this situation, showing three distinct pion showers with the same energy within the HGCAL detector, but with dramatic variation in the spatial patterns of the energy deposits in the detector.

The nuclear interaction length (λ_{int}) characterized the development of the hadronic shower in the material. Fig. 2.18 (left) shows the longitudinal profile of hadronic showers of different energies in lead (Pb). The depth at which the energy deposition profile of a hadronic shower peaks is determined by the EM component and is largely independent of the incident hadron's energy, in contrast to logarithmic dependence in the case of EM showers (see Fig. 2.14). This peak typically occurs approximately one nuclear interaction length (λ_{int}) from the shower initiation point. The longitudinal profile of hadronic showers can be described as a combination of an initial peak from the EM component and an exponentially decaying contribution from the non-EM component. Similarly to the EM shower, the material depth necessary to contain hadronic showers escalates logarithmically with increasing energy.

Hadronic showers also exhibit a wider lateral spread compared to EM showers, as

shown in Fig. 2.18 (right). This is due to the frequent multiple scattering of secondary hadrons, often at large angles relative to the shower axis. The figure 2.18 (right) compares the lateral spread of hadronic showers as a function of hadron energy, indicating that the lateral spread reduced as we increase the energy. This is attributed to the fact that the EM showers generated by neutral pions tend to develop close to the shower axis and the average total fraction (f_{em}) increases with energy.

2.8 Particle detectors

Building on the information of how the particle interacts with the material as discussed in the previous sections, in this section we delve into the details of particle detectors and their detecting mechanism. Particle detectors play a crucial role in experimental particle physics, enabling the identification and measurement of various properties such as charge, mass, and energy of particles as they interact with detector materials. To determine particle properties, interactions with the detector material must produce measurable signals.

Detector material and sizes are optimized depending on the particle type, charge and energy range of interest. In complex high energy experiments, generally a combination of detection concepts and technologies are employed to cover a large range of energy and momentum, position measurements and identification of particle types. The following sections explore key detection methods such as ionization, scintillation, and the detectors based on these phenomena. We discuss the calorimeters at length as these are the main focus of this thesis work (Chapter 5).

2.8.1 Ionization detectors

The basic principle of a semiconductor and gaseous detectors are based on ionization. Semiconductors have an energy band gap between their conduction and valence bands, smaller than insulators but larger than conductors, allowing free electron movement in the conduction band while valence band electrons are bound to atoms. Silicon (Si) and germanium (Ge) are the most widely used semiconductors. The silicon is more preferable for large high-energy particle detectors as it is cheaper and industry protocols of large scale production of silicon sensors are well established. At room temperature, thermal fluctuations generate noise that obscures signals produced by ionizing particles in intrinsic semiconductors. This is overcome by doping intrinsic semiconductors with trivalent or pentavalent atoms, creating extrinsic semiconductors. The conductivity of doped semiconductors is higher than that of intrinsic ones, with electrons and holes as the major charge carriers in n-type and p-type materials, respectively.

The simplest semiconductor device is the p-n junction. At the junction, most charge carriers diffuse into the opposite region and recombine, leaving behind a region with a net charge: positive in the n-type region and negative in the p-type region. This charge separation creates an electric field and a potential difference, forming a depletion zone devoid of any free charge carriers. Such a p-n junction can be operated in both forward and reverse bias.

For particle detection, the p-n junction in reverse bias is used that enhances the electric field and widens the depletion zone. Fig. 2.19 (left) shows a cartoon schematic design of a semiconductor detector operated in reverse bias mode. There is an almost negligible leakage current ($\sim nA$). However, when ionizing radiation traverses the depletion zone, it generates electron-hole pairs along its path. These pairs are rapidly separated by the electric field, producing a measurable current proportional to the

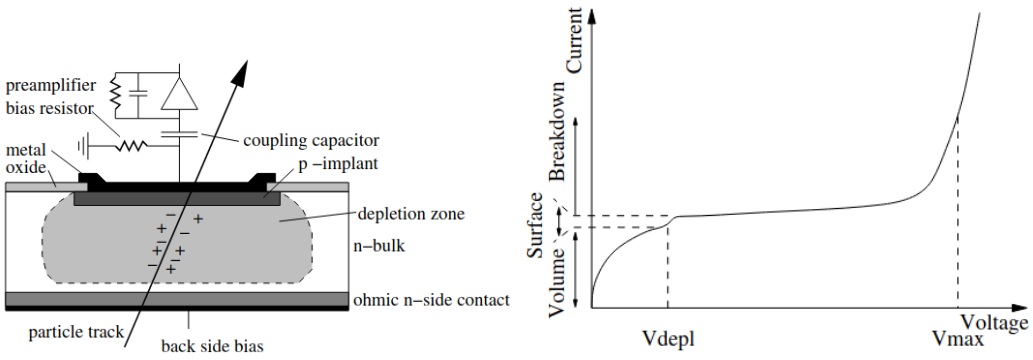


Figure 2.19: Left: Illustration of a charged particle ionizing and creating the e-h pairs in the silicon sensor. Right: The behavior of dark current (when no signal is present) versus the voltage applied between p-n junction [48].

energy deposited by the radiation that can be amplified and processed via read-out electronics. Fig. 2.19 (right) illustrates the behavior of dark current in a semiconductor detector as a function of the applied voltage between the p and n types. The optimal operational condition for a semiconductor detector is achieved when the dark current exhibits constancy as a function of the depletion voltage.

In modern high-energy physics experiments, such as CMS [4], ATLAS, and LHCb at the LHC, silicon detectors are used in tracking devices, placed closest to the interaction vertices owing to their radiation hardness and compactness. The CMS experiment has opted to use silicon sensors in EM and hadron endcap calorimeters for HL-LHC operations (Chapter 3).

In gaseous detectors, incident particles ionize the gas atoms and resultant electrons and ions are collected to create an electronic signal. Energy required to produce one electron-ion pair generally ranges between 30 and 40 eV, which is 10 times higher than that of silicon detectors, $\sim$ 3 eV to create a e-h pair. Gaseous detectors offer good timing and spatial resolution, and these are cost effective. However, these detectors have limited radiation hardness as compared to silicon and scintillators. In addition, maintaining the gaseous detector is a complex task. Drift tubes, Cathode strip chambers, and resistive plate chambers are employed in the CMS experiment to track and measure the momentum and position of high energy muons.

2.8.2 Scintillator detectors

Scintillator detectors use a scintillation mechanism [49] where visible or UV photons are emitted when charged particles traverse certain materials. These materials, known as scintillators, exhibit this behavior due to their energy band structure such that when a charged particle interacts with it, the molecules are excited to high energy levels. Instead of returning directly to the ground states, the molecules first pass through a quasi-stable intermediate state, dissipating energy as lattice vibrations, before finally emitting photons with wavelengths in the visible or UV range. The entire lifetime of this process is around a few tens of nanoseconds. The energy required to produce a single scintillation photon varies depending on the material, ranging from 10 to 100 electron volts. For per MeV energy transfer, the number of scintillation photons can go up to 30 to 60 k.

Fig. 2.20 illustrates the basic construction of scintillator detectors. Scintillator photons are converted to an electronic signal using a photodetector. Signal generated via scintillation is very small; hence, photomultiplier tubes (PMTs) are commonly used. A PMT consists of several key components: a photocathode, focusing and accelerating electrodes, and a series of dynodes. Photocathode converts scintillation photons into

electrical signals by generating photoelectrons in response to these photons. These electrons are then focused and accelerated toward the dynode chain. Each dynode amplifies the signal by producing multiple secondary electrons (typically around ten per incident electron). With about six dynodes in a typical PMT, the overall amplification reaches $\sim 10^6$ per photoelectron, producing a robust electrical pulse proportional to the energy deposited in the scintillator. This amplified signal is used for energy measurements or counting incident radiation. These photodetectors are optimized in a way to match spectral characteristics of scintillation photons to ensure efficient detection of scintillation photons leading to precise energy measurements of particles.

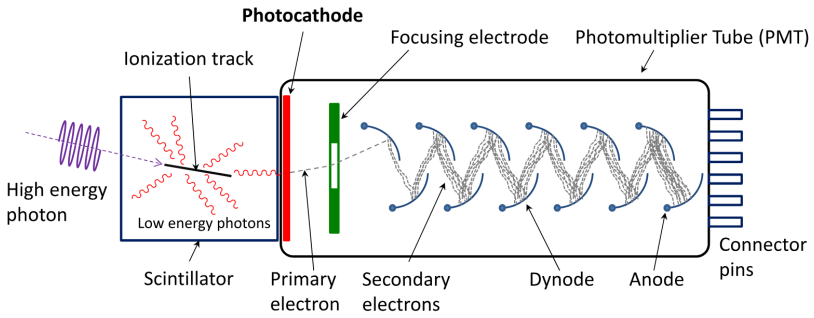


Figure 2.20: Illustration of the scintillator detector : An incoming high-energy photon produces scintillation photons, which are then transformed into photoelectrons and subsequently amplified. Image Credit - Wikipedia

Scintillators play a vital role in nuclear and particle physics experiments, as they can measure energies ranging from keV to GeV. Scintillator properties, such as light output, peak emission wavelength, and decay time, are crucial to select suitable photodetectors. Their fast response times ($\sim ns$) are ideal for triggering systems that decide whether to retain or discard an event data in high-energy physics experiments. In the CMS experiment, hybrid photodiodes (HPDs) and silicon photomultipliers (SiPMs) are employed in the hadron calorimeter (HCAL) to readout signal from plastic scintillators, whereas avalanche photodiodes (APDs) and vacuum phototubes (VPTs) are utilized in the electromagnetic calorimeter (ECAL) to read out lead-tungstate ($PbWO_4$) crystals.

2.9 Calorimeters

Calorimeters are a type of detector, designed to measure the energy and position of charged and neutral particles by completely absorbing their energy. The energy measurement of neutral particles is only possible through the calorimeters, and hence these are very crucial for collider physics such as LHC at CERN. A key advantage of calorimeters is that their energy resolution improves as the energy of the particle increases. In contrast to the trackers where the resolution for momentum measurement worsens with its energy. In such scenarios, calorimeters provide crucial handle to measure energy of these particles, especially.

Calorimeter systems are engineered to fully contain the particle showers initiated by electrons, photons, and hadrons. These showers are cascades of secondary particles produced when the primary particle interacts with the detector material. Dense materials are preferable to build compact calorimeters. Various detection mechanisms including scintillation, ionization, Cherenkov radiation can be used. The choice of

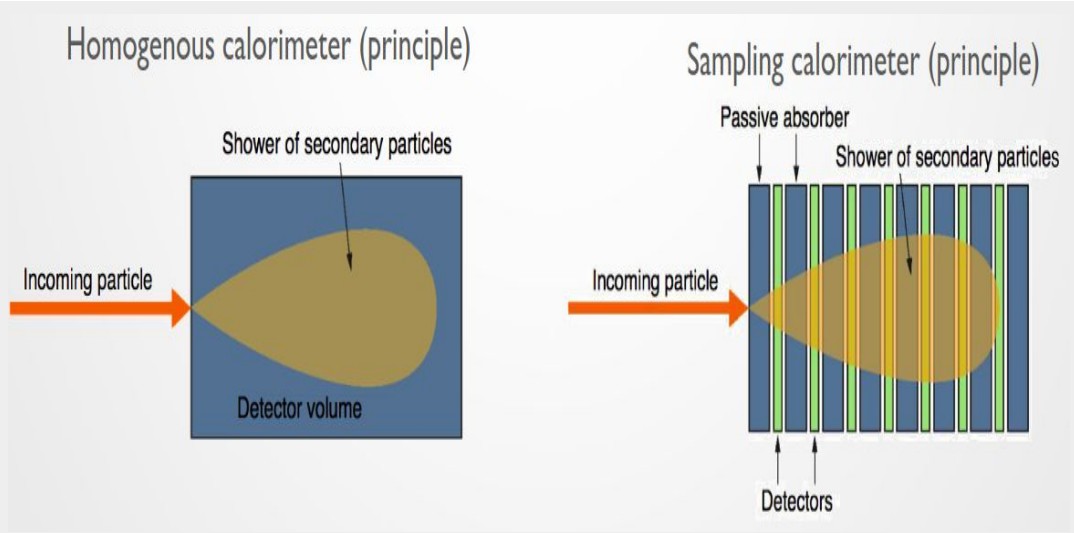


Figure 2.21: Illustration of basic principle behind the homogeneous (left) and sampling calorimeters (right) [50].

material and design of the calorimeters is entirely driven by the nature of the particle showers.

The material in which the incident particle initiates and develops the shower is called the **absorber medium**, while the material that converts the energy of the shower particles into a measurable signal is known as the **active medium**. Depending on the configuration of the absorber and active mediums, calorimeters are broadly classified into two categories: **homogeneous calorimeters** and **sampling calorimeters**, as described below.

Homogeneous calorimeters

Homogeneous calorimeters consist of a single material that serves dual purposes: absorbing the energy of the particle and producing a signal proportional to the energy loss of the particle, as illustrated in the left of Fig. 2.21. Examples include scintillator detectors from materials such as liquefied argon (LAr) and lead tungstate ($PbWO_4$). Given these are high Z materials, these calorimeters effectively measure electromagnetic particle energy (electrons and photons) due to their compact showers. Homogeneous calorimeters offer excellent energy resolution for electromagnetic showers, making them indispensable for precision measurements in particle physics experiments.

Sampling calorimeters

Sampling calorimeters consist of alternating layers of absorber material (passive layers) and active material. The absorber material induces the development of particle showers, while the active layers sample the energy deposition at various stages of the longitudinal development of the shower, as depicted in Fig. 2.21 (right). Each active layer provides a snapshot of the shower's development, and these snapshots are combined to reconstruct the complete shower profile. Various geometrical configurations for active components can be implemented, such as fibers or tiles, although the most common arrangement involves active layers interspersed between layers of absorber. High-Z-Materials, such as iron (Fe), lead (Pb), or uranium (U), are typically used as absorbers. As shown in Table 2.1, λ_{int} is much larger than X_0, requiring thicker calorimeters to fully contain hadronic showers as compared to EM showers

with same incident particle energy. Thus, sampling calorimeters are especially useful for detecting hadronic showers, which have a much larger longitudinal spread than electromagnetic showers. Consequently, they are often employed as **hadronic calorimeters**. The benefit of employing sampling calorimeters lies in the ability to independently select the absorber and active material optimally. In addition, sampling calorimeters allow for examination of the longitudinal development of the shower.
The performance of a sampling calorimeter is characterized by the **sampling fraction (S.F.)**, defined as:

$$\text{S.F.} = \frac{E_{\text{active}}}{E_{\text{active}} + E_{\text{passive}}} \tag{2.50}$$

Here, E_{active} represents the energy deposited in the active layers, and E_{passive} denotes the energy deposited in the passive layers. Typically, the sampling fraction is less than a few percent. Consequently energy resolution of sampling calorimeters is much worse than the homogeneous calorimeters as explained in the following.

2.9.1 Performance metrics for calorimeters

The main purpose of a calorimeter is to measure the energy of the incident particles. Two critical figures of merit for calorimeters are the **energy response** and the **energy resolution**.

- **Energy response** refers to the total signal generated by the calorimeter in response to the energy deposited by an incident particle.
- **Energy resolution** represents fluctuations in the energy response of the calorimeter, indicating its precision in energy measurement.

The following text comprehensively discusses both of these metrics for sampling and homogeneous calorimeters, as well as their application to electromagnetic and hadronic showers.

Energy response of calorimeters

The energy response of a calorimeter is significantly dependent on whether the particle induces an electromagnetic or hadronic shower. The calorimeter response to EM showers is linear with incident energy since all shower energy is converted to signal within the detector. Consequently, the energy response to electromagnetic showers is inherently linear in all calorimeter types. Fig. 2.23 (left) illustrates the linearity of a sampling ECAL as a part of the CMS high-granularity calorimeters (HGCAL).
However, deviations from this linearity can occur due to instrumental effects such as saturation of photodetectors at high energy levels or shower leakage beyond the calorimeter boundaries. These issues can be mitigated through measures such as lowering the photomultiplier tube (PMT) gain or ensuring sufficient calorimeter depth. Assessing the linearity of the calorimeter's response to electromagnetic showers offers further insight into comprehending the specific limitations of the detector.
However, the situation is complex in the case of hadronic showers. As discussed in Section 2.7.2, hadronic showers have two distinct components (EM and non-EM), differing in energy deposition patterns and particle multiplicity. The EM component (mainly from $\pi^0 \rightarrow 2\gamma$), deposits its entire energy within the calorimeter and generates an equivalent signal in the calorimeter. In contrast, the Non-EM component includes an invisible energy fraction due to nuclear breakup in the calorimeter medium, which does not contribute to the signal.
This difference leads to a distinct calorimetric response to hadronic showers compared to pure EM showers. If a calorimeter responds with energy e to a 50 GeV

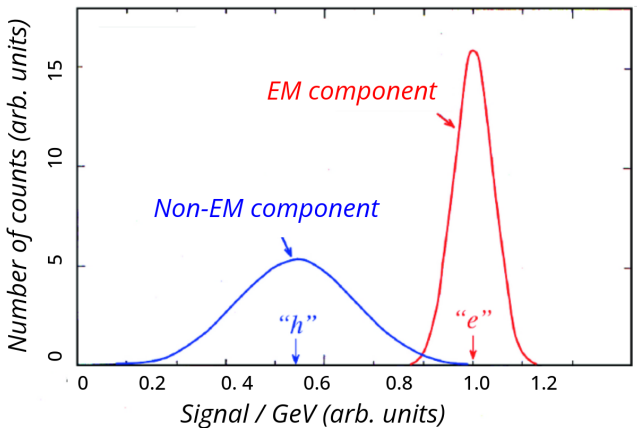


Figure 2.22: The calorimeter's response to both electromagnetic and non-electromagnetic components of hadronic showers is depicted. The width of the non-electromagnetic component reflects the energy deposition fluctuations caused by hadrons [51].

electron and π to a 50 GeV pion, the hadronic response can be expressed as:

$$\pi = f_{\text{em}} \cdot e + (1 - f_{\text{em}}) \cdot h \tag{2.51}$$

Here:

- f_{em} is the electromagnetic fraction of the hadronic shower.
- e is the calorimeter response to the electromagnetic component.
- h is the calorimeter response to the hadronic component.

Rearranging the equation, we derive the following relations:

$$\frac{e}{h} = \frac{1 - f_{\text{em}}}{\pi/e - f_{\text{em}}} \tag{2.52}$$

The ratio e/h is a fundamental property of the calorimeter that influences its response to hadronic showers. Based on the e/h ratio, calorimeters can be classified as follows:

- **Compensating calorimeter:** $e/h = 1$
- **Under-compensating calorimeter:** $e/h > 1$
- **Over-compensating calorimeter:** $e/h < 1$

The e/h ratio cannot be directly measured but can be inferred using the e/π ratio, which compares the calorimeter response to electromagnetic and hadronic showers (e.g. electrons versus charged pions). The calorimeters discussed in this thesis are mostly under-compensating, which means their response to EM showers is higher as compared to hadronic showers. Fig. 2.22 illustrate the undercompensating calorimeter's response to the EM and non-EM components of hadronic showers.
Also, the EM component varies with the incident hadron energy as shown in Fig. 2.17. This leads to a non-linear response of calorimeters for hadronic showers. Hence, a software compensation is necessary to obtain a linear response for hadronic showers.
An initially non-compensating calorimeter can be adjusted to achieve a compensatory state, ensuring that its electromagnetic (e) and hadronic (h) responses are equal. This is accomplished by:

- **Reducing electromagnetic response (*e*):** Using high-Z absorbers increase the probability of photon interactions in the absorbers and limits its interaction in the active medium due to the Z^5 dependence of the photoelectric effect. Only if the electrons generated in this process are produced near the edge of the absorbing layer do they contribute to the signal. Hence, overall response of calorimeters is reduced for the EM component.
- **Enhancing hadronic response (*h*):** Materials abundant in hydrogen, such as plastic scintillators, enhance neutron sensitivity via $n - p$ scattering interactions. In this active medium, the recoiled protons generate signals, thereby increasing h. Absorbers such as ^{238}U contribute additional energy via fission induced by neutrons, compensating for invisible energy losses.
- One can also opt for dual-readout calorimeters to distinguish between electromagnetic and hadronic signals and measure these individually event-by-event. This involves changes in the detector geometry and configuration. These approaches improve energy resolution and event reconstruction accuracy.

Energy resolution of calorimeters

Energy resolution quantifies the precision of energy measurements and is generally represented as relative resolution (σ/μ), where σ is the uncertainty on energy measured with mean μ. The energy resolution can be parametrized with respect to energy of incident particle energy,

$$\frac{\sigma_E}{E} = \sqrt{\frac{S^2}{E} + \frac{N^2}{E^2} + C^2} \tag{2.53}$$

where

- S is the **stochastic term**, representing stochastic fluctuations, the most fundamental source of fluctuation in the energy response, arising from the statistical nature of particle multiplicity in the shower. The particle multiplicity (n) in the shower follows a Poisson distribution with a standard deviation proportional to $\sqrt{n}$. Consequently, the relative resolution scales as $1/\sqrt{n}$. Due to the lower particle multiplicity and larger event-to-event fluctuations in hadronic showers, their resolution is significantly worse than that of electromagnetic showers. Stochastic fluctuations are irreducible and dominate the energy resolution in the intermediate energy range (10 to 100 GeV).
- N is the **noise term**, representing noise-induced fluctuations, that are introduced by the electronics of the detector and particles from pileup events. These fluctuations are most significant at lower energies (below a few GeV) and diminish at higher energies.
- C is the **constant term**, representing fluctuations due to detector imperfections, such as insufficient depth causing shower leakage or calibration inaccuracies between cells, contribute to fluctuations in the energy response. These effects become dominant at very high energies (hundreds of GeV and above).

Each of these factors independently contributes to the overall energy resolution of the calorimeter. These terms are added in quadrature, as they are independent of each other. In sampling calorimeters, only a fraction of the shower energy contributes to the recorded signal. This introduces additional fluctuations in the sampled energy, leading to a worse resolution than homogeneous calorimeters. For a first approximation, the fluctuation due to sampling can be incorporated into the stochastic term

S. This makes the stochastic term larger for sampling calorimeters, reflecting their inherently poorer energy resolution.

Selecting sampling calorimeters for EM showers is feasible, although at the cost of compromising resolution. The sampling fraction must be meticulously optimized to reduce the energy resolution, which is typically on the order of 5%.

Fig. 2.23 (right) shows resolution of the CMS high granularity sampling EM calorimeter designated for HL-LHC. Over a range of beam energies of 20 to 300 GeV, it has stochastic and constant energy resolution terms have been recorded as 22.1 and 0.58, respectively [52][53] in contrast to 2.83 and 0.26, respectively for the current homogeneous ECAL based on $PbWO_4$.

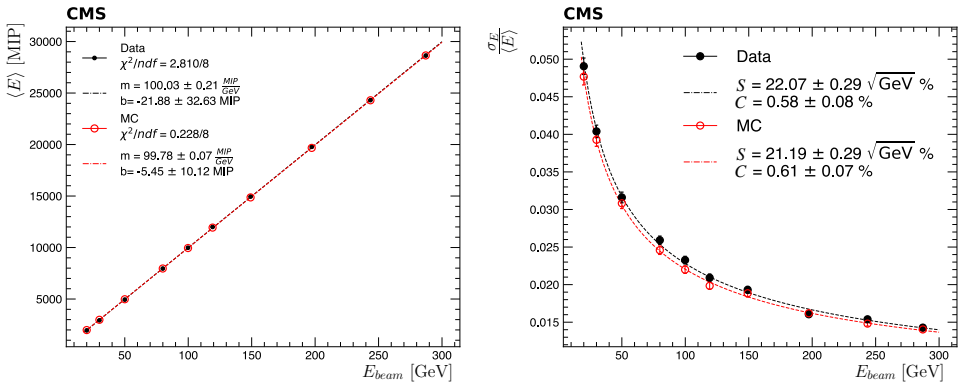


Figure 2.23: Left: The mean measured energy is plotted against the beam energy for both experimental data and simulation outcomes. Right: The relative energy resolution of the measured energy is illustrated for both experimental data and simulation results [52].

2.10 Event simulations in high energy physics

The Monte Carlo [54] event simulations are pivotal in HEP data analysis and they drive accuracy of SM measurements and new physics searches. For large-scale detectors like CMS, modeling the detector geometry, interaction of particles with detector materials, and simulating the detector response is the most challenging and computationally demanding task.

Event simulation is a multi-step process that begins with modeling the fundamental physics interactions using tools like MADGRAPH. MADGRAPH [55, 56] generates precise matrix elements for specific processes. These events are then passed on to PYTHIA [57], which simulates subsequent showering, hadronization, and decays of the parton, effectively modeling the complex evolution from initial partons to stable particles. Finally, GEANT4 [58] is employed to simulate the passage of these particles through the detector material, accounting for interactions with the detector components to produce realistic signals. This process is repeated for billions of events to generate simulated datasets.

For large-scale detectors like CMS, modeling the detector geometry, interaction of particles with detector materials, and simulating the detector response is the most challenging and computationally demanding task.

2.10.1 Overview of detector simulation using GEANT4

The CMS experiment employs the GEANT4 simulation toolkit [58] for a detailed simulation of the detector and simulation effects. The primary particle content of the event either can be plugged in from external sources or internal particle guns can be used. GEANT4 also, enables particle transport simulation, construction of detector models with various shapes and materials, and point navigation within them. It

accounts for physical interactions to generate secondary particles and records data through hits for detector simulation. It also supports visualizing geometry and particle paths, with application interaction via a customizable terminal or interface.
To start the simulation, one needs to define the detector geometry. GEANT4 uses a modular approach in which the detector is composed of various volumes. The largest volume, called the *world volume*, contains all other volumes. Each volume is defined by its shape, size, and material. Sensitive and passive detector regions are defined within the geometry to record energy deposition by particles, which is essential for energy reconstruction. Using these details, the complete detector geometry can be constructed.
As we understood in Section 2.6, the material detector plays a crucial role in determining the energy depositions of different particles. Moreover, two of the key parameters of EM and hadronic showers, such as radiation length and interaction length are highly dependent on its atomic mass (A) and atomic number (Z). GEANT4 includes a database of predefined materials and allows users to define custom materials such as alloys [59].
GEANT4 provides a comprehensive list of particles and interaction models that users can configure based on their requirements. Particles are propagated through the detector material and undergo interactions determined by their type and energy. Each particle is tracked until its energy falls below a threshold, at which point it is terminated. GEANT4 supports separate models to simulate electromagnetic and hadronic interactions.

Modeling electromagnetic showers

Electromagnetic interactions include processes like multiple Coulomb scattering, ionization, bremsstrahlung, and pair production. These interactions are well-understood and have been extensively validated against experimental data, showing agreement within a few percent [60]. Other processes like Rayleigh scattering, etc are also included but play a negligible role in our studies. Electromagnetic (EM) interactions in GEANT4 are managed through the `G4VProcess` interface, with three specialized classes: `G4VEnergyLossProcess`, `G4VMultipleScattering`, and `G4VEmProcess`. These handle energy loss, scattering, and post-step/rest interactions, respectively. EM processes are based on models implemented via the `G4VEmModel` interface, which operate within defined energy ranges and regions.

Modeling of hadronic showers

Hadronic shower models in GEANT4 describe the interactions of hadrons with nuclei in different energy regimes. GEANT4 provides two types of models for simulating hadron cascades: parametric models and models based on theoretical concepts. These models operate on the principle that at low energies, individual nucleons in a nucleus must be considered, while at high energies, the quark structure of the projectile and a single target nucleon are crucial. Given the complexity of hadron interactions, currently, no hadronic shower models can describe shower observables with the same precision as electromagnetic shower models.

Cascade models

Cascade models operate in the energy range of a few hundred MeV to a few GeV, where the de Broglie wavelength of the projectile is comparable to the nucleon spacing, but still larger than the quark substructure scale. These models track the

projectile through the nucleus, modeling interactions with nucleons based on nucleon densities and cross sections. Secondary particles generated in these interactions are similarly tracked, potentially undergoing further interactions, absorption, or escape. The schematic of this process is illustrated in Fig. 2.24.

GEANT4 includes two implementations of the cascade model: the Bertini and binary cascade models. The Bertini model [61] approximates the nucleus as spherical shells of constant nucleon density with Fermi-gas momentum distributions. It calculates nucleon momenta, reaction types, and products directly during collisions. A pre-equilibrium model describes evaporation using exciton configurations, while de-excitation involves Fermi-breakup, fission, and evaporation models.

The binary cascade [62] model represents nucleons as discrete entities with positions sampled from a Wood-Saxon distribution [63] for heavy nuclei or a harmonic oscillator model for light nuclei. Nucleon momenta are randomly assigned up to the Fermi momentum. Interactions transfer the initial scattering state to products through intermediate steps, generating hadronic resonances. Decays are governed by experimental branching ratios, producing secondary particles. Unlike the Bertini model, it relies on external models for pre-equilibrium and de-excitation processes.

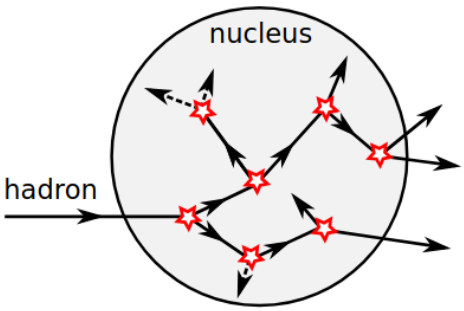


Figure 2.24: Illustrated operational concept of cascade models realized in GEANT4 [64].

Parton string models

Parton string models describe high energy hadron-nucleus interactions, typically above 5 GeV, where quark-level dynamics dominate. The nucleus is modeled using protons and neutrons with positions derived from harmonic oscillator or Wood-Saxon potentials, depending on nuclear size. Nucleons are assigned momenta up to the Fermi limit. The projectile interacts with a single nucleon, with the interaction determined by the impact parameter, center-of-mass energy, and cross sections.

The schematic representation of this model is shown in Fig. 2.25. A string forms between quarks of the projectile and target nucleon, characterized by its four-momentum and endpoints. The string stretches and fragments into quark-antiquark or diquark-antidiquark pairs, forming new hadrons. This fragmentation continues until the string's energy is exhausted.

GEANT4 implements two parton string models: the Fritiof model [65] and the quark-gluon string model [66]. The Fritiof model handles diffractive interactions via momentum exchange, while the quark-gluon string model employs pomerons for scattering. The models differ in string formation and fragmentation functions.

Post-fragmentation, the excited nucleus transitions to equilibrium through cascade or precompound models. Deexcitation is handled using precompound and phase-space models, ensuring accurate simulation of the final nuclear state.

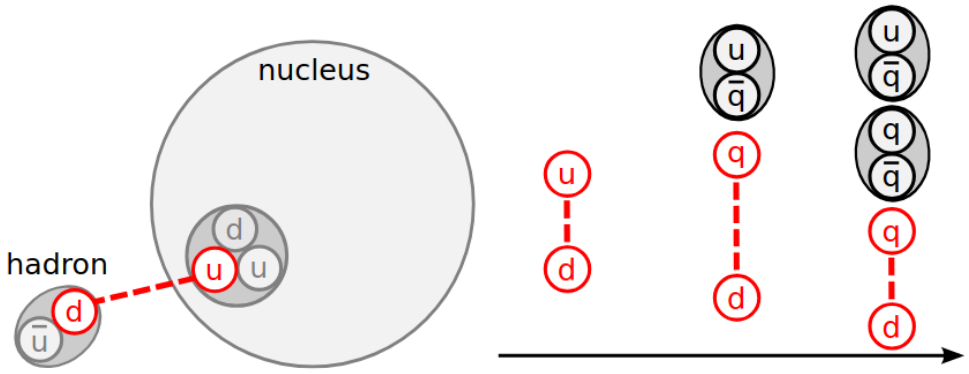


Figure 2.25: Left: The development of strings between the constituent quarks of an incoming hadron and those within a target nucleus. Right: The fragmentation of gluon strings through the generation of quark-antiquark pairs, succeeding which, hadronisation occurs [64].

Furthermore, GEANT4's capabilities are expanded to incorporate high precision modeling for low-energy neutrons, allowing for more precise forecasts of energy deposits from neutrons further down the cascade. These high-precision neutron models rely on a complete neutron library that includes cross sections, angular distributions, and final state data.

Physics lists

A single model may not accurately capture hadronic interactions across all energy regimes. The physics models are merged to create a "physics list". The physics lists utilize different models with a specified validity region and little overlap between them. To ensure a smooth transition, the physics model is picked randomly and equally across the overlap zone. Two of such physics lists are used in the analysis presented in Chapter 5.

- **FTFP_BERT_EMN:** applies the Bertini cascade model for energies under 12 GeV and the Friitof model for energies above 3 GeV. The EMN is an electromagnetic physics model that focuses on handling multiple scattering, which is crucial for calorimetry sampling.
- **QGSP_FTFP_BERT_EMN:** uses the Bertini cascade model for energies up to 8 GeV, the Fritiof model for 3-25 GeV, and the QGS model for energies over 12 GeV.

2.10.2 Event simulation in GEANT4

Having constructed the detector geometry and the list of particle interaction models in GEANT4, the required number of events can be generated. GEANT4 provides two manager classes to handle the simulation: `G4RunManager` and `G4EventManager`. The `G4RunManager` encapsulates all the necessary information needed to simulate events, such as geometry and physics. The `G4EventManager` encapsulates the representation and structure of the event. A hierarchy of processes, called `RunAction`, `EventAction`, and `StepAction`, are invoked while simulating the events. The simulation process is briefly described below.

For each event, particles are generated using the inbuilt GEANT4 event generator or an external event generator interfaced with GEANT4, such as PYTHIA [57]. Each particle is tracked along its path, referred to as `G4Track`. The `G4Track` is propagated in steps, called `G4Step`. At each `G4Step`, a stepping action is invoked, where the

step is propagated to the next step. The step length can be adjusted on the basis of resource consumption and required precision. At each step:

- The stepping action evaluates the current state of the step, such as kinetic energy and particle type.
- All physical processes applicable to the particle at that step are invoked.
- Secondary particles and their `G4Track` instances are created, and their information (e.g., kinetic energy, particle type, position) is stored.
- The kinetic energy of the primary particle is updated on the basis of interactions.

If the particle traverses through the sensitive material of the detector, a *hit* is created and the following associated `G4Step` information is saved:

- Energy deposited by the step,
- Position of the step,
- Time-stamp of the step,
- Geometrical information of the detector.

This hit is called a `simHit` in GEANT4 terminology.
The geometrical information of the detector and the deposited energy are crucial for event reconstruction. Based on geometric information, `simHit` can be mapped to a particular detector sensor, and the total `simHit` energy recorded can be converted into a signal generated by the sensor. For each event, every `G4Track` is propagated and tracked inside the detector until the particle comes to rest, or it crosses the boundaries of the geometry of the detector. The `simHit` information is stored for each event and used for event reconstruction. The simulation process ends at the end of the run. Using GEANT4 's inbuilt `simWatcher` method, additional `G4Step` information can be stored as per user requirements, such as particle type, energy, and the number of secondaries generated. This process is described in more detail in Chapter 5.

2.11 Summary

This chapter explores the Standard Model (SM) of particle physics and its shortcomings, before introducing the Supersymmetry as a possible expansion. The chapter discusses fundamental aspects of particle interactions, electromagnetic and hadronic showers, and the function of detectors and calorimeters in such measurements. In addition, the chapter discusses the use of GEANT4 simulations to predict particle behavior and detector response.

Chapter 3

Experimental setups

In the investigations presented in this thesis, we use Run 2 pp collision data recorded with the CMS experiment to search for SUSY like signatures and data collected in 2018 beam test experiment with prototype HGCAL detector to validate a machine learning based approach to reconstruct charged pion showers. In this chapter, we start with a brief introduction to the LHC and discuss these experimental setups. We summarize the event reconstruction using the different CMS subsystems along with their description. This is followed by a discussion of Phase-1 and Phase-2 CMS detector upgrades with a specific focus on the HGCAL. The HGCAL readout electronics system is described in Chapter 6.

3.1 The Large Hadron Collider (LHC)

The LHC, CERN [3], is the most powerful particle accelerator in the world. It features a 27-kilometer ring located approximately 100 meters underground, which stretches the border between France and Switzerland. It accelerates the particles (proton or heavy ions) bunches and collides them at four interaction points, as illustrated in Fig. 3.1. Several distinct experiments are strategically placed at these interaction points. Among its four primary detectors, ATLAS [5] and CMS [4] are general purpose experiments, while ALICE [6] and LHCb [7] are specialized for conducting heavy-ion and b-physics research, respectively. The LHC achieves a center-of-mass energy of 13 (13.6) TeV in proton-proton collisions during its Run 2 (Run 3) operations.

Proton beams, organized into bunches containing approximately 1.15×10^{11} protons, collide every 25 nanoseconds. Fig. 3.1 shows the CERN accelerator complex, highlighting the positions of the series of accelerators used to achieve a tremendous amount of energy and as well as different particle detectors. The proton production process begins with hydrogen atoms, which are stripped of their electrons using a duoplasmatron, yielding protons. These protons are accelerated through a series of accelerators before entering the LHC ring. The linear accelerator (LINAC) initially boosts protons to 50 MeV using superconducting radiofrequency (RF) cavities that generate electromagnetic fields, each boosting proton energy by 2 MeV. Subsequently, the protons pass through Proton Synchrotron Booster (PSB) and Proton Synchrotron (PS), achieving energies of 1.4 GeV and 26 GeV, respectively, simultaneously reducing the bunch size. The beam's trajectory and focus are controlled by dipole and quadrupole magnets, respectively. In PS, the RF cavity operates at 40-80 MHz. The PS optimizes beam brightness and adjusts the spacing of the bunch (essentially creating bunches), preparing protons with a bunch length of approximately 4 *ns* for further acceleration in the super-proton synchrotron (SPS), where they reach 450 GeV with the help of RF cavities operating at 200 MHz.

The SPS and PS have cycling times of 21.6 seconds and 3.6 seconds, respectively, resulting in an LHC filling time of approximately 4 minutes per beam. An additional setup time is required for injection, including 4 SPS cycles for pilot bunches and

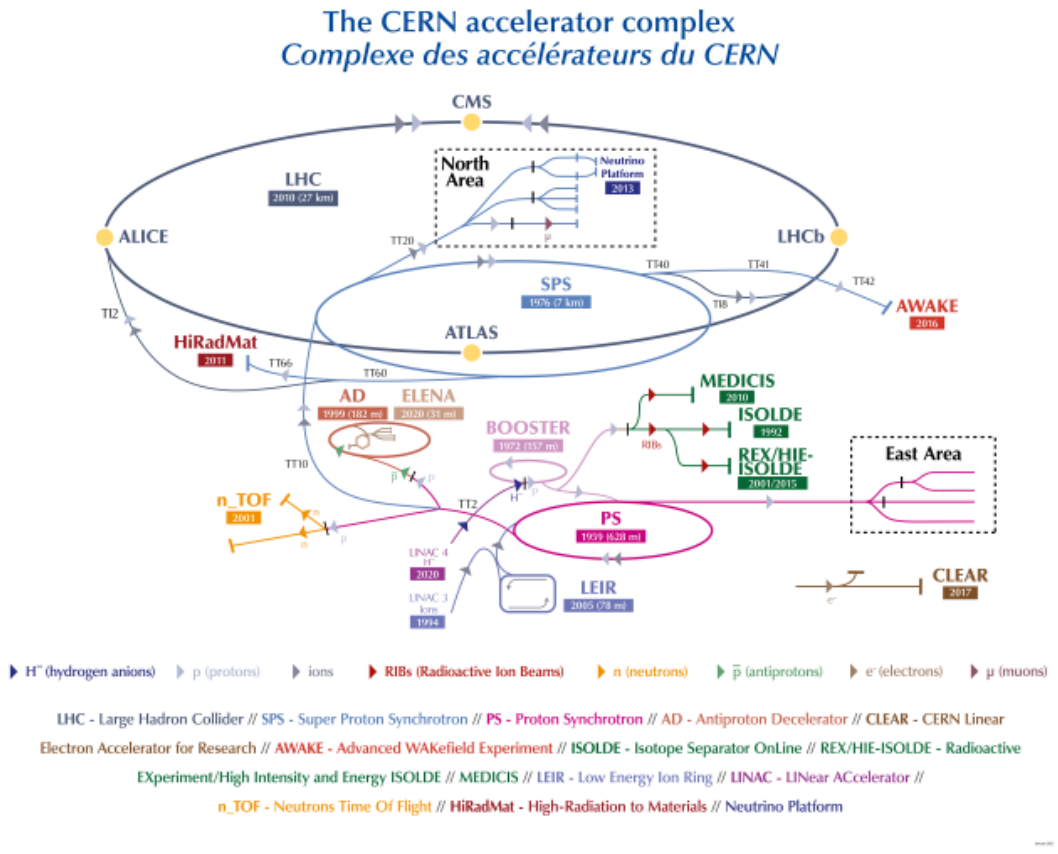


Figure 3.1: CERN accelerator complex. (Image Credit CERN)

nominal intensity. The total minimum injection time for the LHC is approximately 16 minutes.

Finally, the LHC's RF cavities, operating at 400 MHz, accelerate the protons to their peak energy of approximately 6.8 TeV over approximately 20 minutes. Each pass through these RF cavities increases the proton's energy by 16 MeV, resulting into approximately 10 million turns. Out of total 3560 bunches with a spacing of 25 ns each, only 2808 bunches are filled at a time to allow proton bunch injection to maintain the collision rate at the interaction point (IP) and safe dumping of the proton beam in case of an emergency. The LHC tunnel has two rings where two proton beams flow in opposing directions. They are guided by massive superconducting electromagnets with an 8.33 T magnetic field. These electromagnets use superconducting cable coils made of niobium-titanium, which are held at -271.3°C with liquid helium.

3.1.1 Luminosity and pileup interactions

The frequency of occurrence of a specific physics process is determined by its cross section at a given center-of-mass energy and the instantaneous luminosity. The process rate can be expressed as;

$$\frac{dN_{\text{events}}}{dt} = L_{inst} \cdot \sigma \tag{3.1}$$

where σ is the process cross section, and L_{inst} represents the instantaneous luminosity, which depends on the collider characteristics. The cross section, σ, is determined by the details of the interaction and is measured in units of barns (1 barn =

10^{-24}cm^{-2}).The total events can be calculated as;

$$N_{total} = L_{int}\sigma \tag{3.2}$$

where L_{int} is the integrated luminosity, defined as $\int L_{inst}.dt$.
The instantaneous luminosity [67] depends on several factors, including the number of particles per bunch (N_b), the total number of bunches (n_b), the collision frequency (ν), the transverse size of the collision area ($4\pi\sigma_x\sigma_y$), and a reduction factor (F) as per the following equation;

$$L = \frac{N_b^2 n_b \nu \gamma_r}{4\pi\epsilon_n\beta^*}.F \tag{3.3}$$

where γ_r is the Lorentz boost factor. The ϵ^* and β^* are the normalized transverse emittance and beta function at the crossing point. The reduction factor (F) accounts for the beam-beam effect caused by the electric repulsion between the proton bunches. It depends on the total crossing angle (θ_c), as well as the transverse and longitudinal bunch dimensions (σ_{xy} and σ_z) as per the following equation;

$$F = \frac{1}{\sqrt{1 + (\frac{\theta_c\sigma_z}{\sigma_{xy}})^2}} \tag{3.4}$$

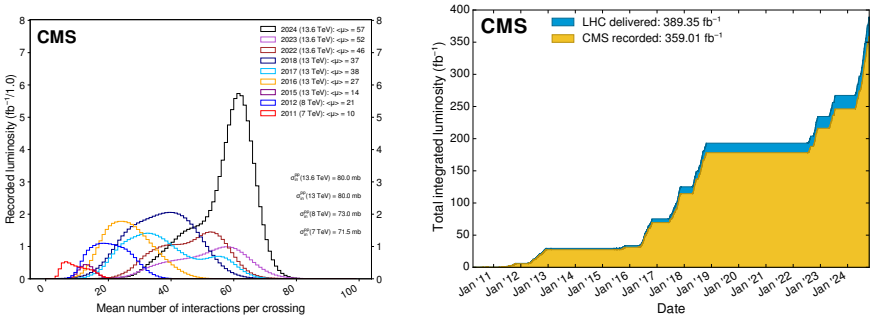


Figure 3.2: Left: Number interactions per event crossing (pileup). Right: The cumulative luminosity delivered and recorded [68].

Currently, LHC is in its Run 3 operation phase colliding pp bunches with an instantaneous luminosity of $\approx 2.5 \times 10^{34}\text{cm}^{-2}\,\text{s}^{-1}$. In a pp bunch crossing, there are multiple pp interactions possible. The majority of these interactions are gentle, with minimal momentum transfer, resulting in soft particles. Hard interactions with high momentum transfer can result in heavy SM or BSM particles, which are of interest. These soft interactions are called "pileup". Substituting the total inelastic cross section ($\sigma^{pp} \approx 80$ mb) and the instantaneous luminosity (L_{inst}) with a time interval between two bunches (ΔT) of 25 ns into equation 3.1, we calculate an average pileup of approximately 40. Fig. 3.2 (left) shows distribution of number of pileup interactions in data collected over the various years. Throughout Run 3, the average pileup has intensified to about 57.
Fig. 3.2 (right) demonstrates the total data delivered by LHC (in cyan) and the total data recorded in the CMS experiment over the LHC operations from Run 1 to Run 3. During its Run 2 phase, LHC delivered around 150 fb^{-1} of data at 13 TeV energy and with 40 average pileup. Currently, the LHC is in its Run-3 phase and has delivered more than 250 fb^{-1} of data.

3.1.2 High luminosity phase of LHC (HL-LHC)

Table 3.1: Different parameters of LHC and HL-LHC [2].

Parameter	LHC	HL-LHC
Number of bunches	2808	2760
Beam current (A)	0.58	1.1
Minimum β^* (m)	0.55	0.15
Peak luminosity (fb)	1.18	17
Leveled luminosity	-	5
Average pileup	60	200

The HL-LHC [2] is a major and ambitious upgrade to the LHC, designed to increase its luminosity by a factor of five and deliver up to ten times more than the current integrated luminosity over a decade of running. This upgrade will ensure higher statistics of interesting events for the same cross-section enabling detailed studies of known SM processes and rare productions including beyond SM. For instance, throughout Run 1, Run 2, and Run 3 phases of the LHC, approximately $O(M)$ Higgs bosons are generated over a ten-year period of its operation. During HL-LHC, this number is expected to be produced per year of operations.
The HL-LHC is scheduled to commence its operations by the end of 2029, following the installation of new components during the third long shutdown (LS3) starting in 2026.
Table 3.1 summarizes the different parameters for LHC and HL-LHC, including the increase in the number of pileup interactions. Section 3.3.2 delves into the repercussions and obstacles that come with HL-LHC for particle detectors.

3.2 The compact muon solenoid (CMS) experiment and event reconstruction

The primary aims of the CMS experiment [4] are to probe a wide range of physics, including measurements of the properties of the Higgs boson, studies of top quarks, and electroweak symmetry breaking, while concurrently exploring realms of physics that extend beyond the Standard Model (BSM).
To accommodate such a large physics program, the CMS detector was built to meet specifications regarding the reconstruction of the particles reaching the detector. It is equipped with detecting technologies to provide precise measurements of electrons, photons, and muons along with reconstruction of jets originating from the presence of quarks and gluons. Hermetic coverage ensures the correct calculation of the missing transverse energy ($p_{\mathrm{T}}^{\mathrm{miss}}$) to infer the presence of neutrinos. These characteristics ensure the CMS detector has excellent mass resolution for dilepton and diphoton events in Higgs boson searches, along with precise reconstruction of missing transverse energy and tracking for new physics investigations.

The CMS detector is based on a cylindrical design, providing almost complete 4π coverage as shown in Fig. 3.3. Multiple detection techniques are implemented in the CMS experiment with each sub-detector optimized for efficiently measuring specific particles. Tracking systems are placed closest to the beampipe followed by the electromagnetic and hadron calorimeters. These detectors are immersed in 3.8 T magnetic field provided via superconducting solenoid magnet. Surrounding these are the return yoke, which incorporates muon detector systems.

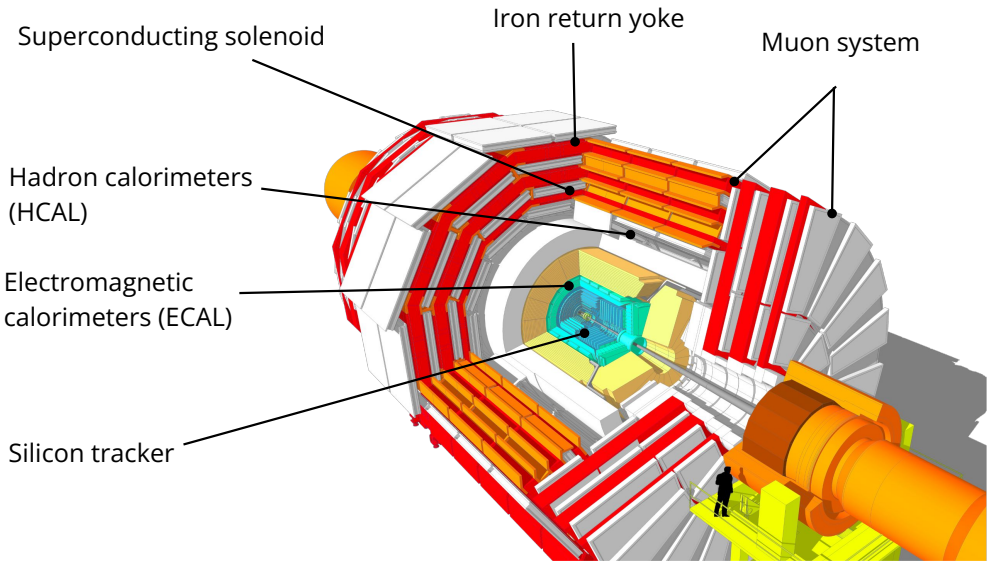


Figure 3.3: Schematic view of CMS detector highlighting different subsystems [69].

The detector coordinate system is defined with its origin at the nominal collision point. The y-axis points vertically upward, the x-axis points radially inward toward the LHC center, and the z-axis aligns with the beam direction. The azimuthal angle ϕ is measured in the x-y plane from the x-axis, while the polar angle θ is measured from the z-axis. Pseudorapidity, η, is related to θ as $\eta = -\ln[\tan(\theta/2)]$.

3.2.1 Silicon tracker

The CMS tracker [70] is a silicon-based particle detector designed to measure trajectories and momenta of charged particles and identify primary and secondary production vertices in pp collisions at the LHC. The CMS tracker spans a longitudinal length of 5.8 m and provides pseudo-rapidity coverage of $|\eta| < 2.5$. The lightweight support structure of the tracker reduces the potential for multiple scattering and ensures that the particles do not lose significant energy in the tracker prior to entering the calorimeters.

The tracker consists of two main components: the inner pixel detector (within $|eta| < 1.2$) and the outer strip tracker (beyond 1.2 and within $|\eta| < 1.2$). The pixel detector is located closest to the beamline, with layers positioned at radii ranging from 4.4 to 10.2 cm. The proximity of the detector to the beamline subjects it to particle fluxes reaching 600 million particles per cm^2 per second. Keeping this in consideration, silicon pixels are finely segmented, each measuring approximately $100 \times 150\,\mu\text{m}^2$ with a thickness of 300 μm, providing a position resolution of around 10 μm. The strip tracker extends further out, covering larger radii and featuring silicon strips with pitches varying between 80 μm and 180 μm, depending on their distance from the beamline.

The silicon tracker is composed of multiple layers: four pixel layers and ten (twelve) strip layers in the barrel (endcap) region. The silicon sensors on the outer tracker are thicker (500 μm) compared to those on the inner tracker (320 μm) to optimize the signal-to-noise ratio while minimizing read-out channel requirements.

The momentum resolution for charged particle is directly proportional to its momentum p_{T},

$$\frac{dp_T}{p_{\text{T}}} \propto \frac{p_{\text{T}}}{0.3BL^2} \tag{3.5}$$

where B is the magnetic field strength and L is the length of lever-arm. The tracker achieves a transverse momentum resolution of about 1–2% for $p_T < 10\,\text{GeV}$, mainly limited by multiple scattering. For p_T between 10 and 100 GeV, the resolution is dominated by the intrinsic detector resolution and is approximately 10-15%. At high p_T, tracker momentum performance degrades as particle trajectories become nearly linear due to minimal curvature in the magnetic field.
In total, the CMS tracker includes 1,856 silicon pixel modules with 124 million pixels and 15,148 silicon strip modules with 9.6 million strips.

3.2.2 The Electromagnetic calorimeter (ECAL)

The ECAL [71] is a homogeneous calorimeter constructed using lead tungstate ($PbWO_4$) scintillating crystals. The principal objective of the ECAL is to measure and identify electrons and photons by utilizing tracker information, noting the absence of tracks for neutral particles. A certain fraction of energy is also deposited in ECAL by hadrons, which together with the energy deposited in HCAL is eventually used later to reconstruct charged and neutral hadrons (see Section 3.2.6). $PbWO_4$ crystals are chosen for their fast scintillation response, with 80% of the emitted light produced within 25 ns. In addition, they exhibit excellent radiation tolerance, capable of withstanding doses of up to 10 Mrad.
It is divided into two main components: the barrel (EB) and the end cap (EE), covering the pseudorapidity ranges of $|\eta| < 1.479$ for the barrel and $1.479 < |\eta| < 3.0$ for the end caps. To detect the scintillation light, the ECAL employs avalanche photodiodes (APDs) in the barrel and vacuum phototriodes (VPTs) in the endcaps. Although VPTs have lower quantum efficiency compared to APDs, they are more resistant to radiation damage, making them suitable for the high-radiation environment of the endcaps. To mitigate radiation-induced transparency losses in the crystals and noise in the photodetectors, the endcaps are operated at a temperature of -20°C.
The barrel section of the ECAL consists of 61,200 crystals arranged in a cylindrical configuration. Each crystal has dimensions of $2.2 \times 2.2 \times 23$ cm^3. The transverse size of these crystals is derived from the Molière radius of $PbWO_4$ (see table 2.1). Its length in the longitudinal directions corresponds to $25X_0$ in radiation length to ensure ∼98% containment of electromagnetic showers. The crystals are grouped into 36 supermodules, each spanning 20° in ϕ, and are tilted by 3° in both η and ϕ to optimize coverage. The endcap region contains two disks, each with 7,324 crystals of dimensions $2.9 \times 2.9 \times 22$ cm^3.

The energy resolution of the ECAL for photons and electrons is characterized by the equation:

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E}} \oplus \frac{12\%}{E} \oplus 0.3\%, \tag{3.6}$$

The ECAL is also capable of providing timing information with an approximate resolution of 150 ps.

A preshower detector is installed in front of the end caps, covering the pseudorapidity range $1.653 < |\eta| < 2.6$. This device consists of two lead absorbers with a total thickness of $3X_0$, followed by silicon micro-strip sensors with a strip pitch of 1.9 mm. The preshower system provides fine transverse granularity, which is crucial for distinguishing prompt photons from those originating from neutral-pion decays.

3.2.3 The Hadron calorimeter (HCAL)

The HCAL [72] is a sampling calorimeter designed to measure energy of hadrons. It is divided into four main sections: barrel (HB), endcap (HE), outer (HO), and forward (HF) regions.

Positioned behind the tracker and the electromagnetic calorimeter (ECAL), the HB and HE provide coverage for $|\eta| < 1.3$ and $1.3 < |\eta| < 3.0$, respectively. The HB and HE employ brass as the absorber material, owing to its short nuclear interaction length and its immunity to magnetic fields; concurrently, plastic scintillators are utilized as the active layers. The signal generator in the scintillator by traversing particles is transferred to photodetectors via wavelength-shifting (WLS) fibers. Hybrid photodiodes (HPDs) were initially utilized in the barrel, and silicon photomultipliers (SiPMs) were used in the endcap. During Phase 1 upgrade, the HPDs of the barrel are replaced with SiPMs.The scintillator tiles are divided into $\Delta\eta \times \Delta\phi$ units, with dimensions of 0.087×0.087 in the barrel and 0.17×0.17 in the endcap. The thickness of the absorber layers varies, resulting in an effective interaction length of 6 to 10 λ_{int} based on $|\eta|$.

The HCAL also includes the outer calorimeter (HO), which is located outside the solenoid and complements the barrel calorimeter for $|\eta| < 1.26$. By increasing the effective thickness of the detector by an additional 2–3 λ_{int}, the HO reduces hadronic shower leakage, improving the energy resolution for hadrons, and improving missing transverse energy measurements, crucial for new-physics searches, such as supersymmetry.

The HCAL energy resolution for hadronic showers is parameterized as:

$$\frac{\sigma_E}{E} = \frac{110\%}{\sqrt{E}} \oplus 9\%, \tag{3.7}$$

where E is expressed in GeV.

In the very forward region, the HF extends the pseudorapidity coverage to $|\eta| < 5.0$. HF uses steel as the absorber material and quartz fibers as the active medium to ensure radiation hardness. It makes use of Cherenkov radiation to measure the energy of the traversing particle. Two types of fibers; long fibers covering the entire longitudinal depth and short fibers starting from 22 cm downstream of the HF's front face, enable the separation of electromagnetic and hadronic components of showers. The HF's location, approximately 11 meters from the interaction point, allows it to handle the high particle flux in the forward region. In total, the HCAL comprises 9072 readout channels, with 2592 channels in the barrel and endcap regions and 3888 channels in the forward region.

With its hermetic coverage the CMS calorimeter system provide robust measurements of jets and $p_{\text{T}}^{\text{miss}}$ which are discovery tools at the LHC.

3.2.4 Muon system

The CMS muon system [73] is designed to identify, reconstruct, and trigger on muons passing through the detector. It makes use of three different different detectors working on the principle of multiwire proportional chambers: Drift Tubes (DT) in the barrel region ($|\eta| < 1.2$), Cathode Strip Chambers (CSC) in the endcap region ($1.0 < |\eta| < 2.4$) where the magnetic field is intense, and Resistive Plate Chambers (RPC) are in both the barrel and endcap region. The subsystems are optimized for different muon energy ranges and angles. DTs measure the positions of muons in

the central region, CSCs provide precise tracking in the endcaps, and RPCs offer fast triggering, although their position resolution is less precise.

- **Drift tubes (DT):** consist of four layers arranged in concentric cylinders around the beam line. Each layer contains 12 chambers with 60 drift tubes. The tubes are filled with a gas mixture of 85% argon and 15% carbon dioxide. DTs provide a spatial resolution of 200 μm and a time resolution of 3 ns, with an efficiency greater than 97% for muons with $p_T > 20$ GeV/c.
- **Cathode strip chambers (CSCs):** consists of six layers of anode wires and cathode strips. The CSCs offer a spatial resolution of 45–145 μm and a time resolution of 7 ns, with an efficiency exceeding 95% for muons with $p_T > 10$ GeV/c.
- **Resistive plate chambers (RPC):** consists of a parallel plate counter, with two electrodes made of high resistivity plastic. RPCs operate in an avalanche mode, generating fast signals with a timing resolution of $\sim$1 ns. RPCs can also operate efficiently at high flux rates. The stripes align parallel to the DTs in the barrel and the radial strips of the CSCs. The hit strip pattern produces radial CSC strips. The hit strip pattern provides crucial information on the muon momentum.

Each muon station has 12 (barrel) or 6 (endcap) measuring planes, with a position resolution of approximately 100 μm. The muon tracking system is integrated into the CMS magnet's return yoke, extending the tracking coverage beyond the interaction point.
The momentum is determined via the trajectory's sagitta. The combination of the data from the inner tracker and the muon stations facilitates outstanding momentum resolution for muons across a broad range of angles. A global track is reconstructed by combining hits from the inner tracker and muon stations, allowing for precise momentum measurement, with a resolution of 1% (10%) at 50 GeV (1 TeV).

3.2.5 Triggers and data acquisition system

With its $\sim$ 140 million channels designed to read out every 25 nanoseconds, CMS generates 40 terabytes (TB) data per second. It is not possible to store and process all events. Furthermore, majority of these events are soft multijet events and interesting events with hard interactions are very rare. Therefore, the CMS trigger system [74] selects a small fraction of events relevant to the physics program for offline analysis. The trigger system reduces the event rates from 40 MHz to $\sim$1 kHz (during Run 2 and 5 kHz during Run 3 LHC operations) through a two-stage process consisting of the Level-1 (L1) Trigger and the High-Level Trigger (HLT).

The **Level-1 Trigger (L1T)** is a hardware-based system that processes coarsely segmented data from calorimeters and muon systems for preset patterns or physics objects such as e/γ or muons with high p_{T} and aims to maximize interesting events in triggered events. The L1 system comprises the Calorimeter Trigger (CALO), Muon Trigger (MUON), Global Trigger (GT), and Timing Trigger (TT). The CALO trigger processes energy deposits (trigger primitives) from the electromagnetic and hadronic calorimeters to reconstruct physics objects such as EGamma, taus, and jets. The MUON trigger uses input from muon chambers and track finders to identify muon candidates based on their pseudorapidity ($|\eta|$) ranges. The GT integrates input from the CALO and MUON triggers to make the final L1 decision. The TT ensures synchronization of the L1 system timing with the LHC beam structure. The L1 trigger system reduces the event rate from 40 MHz to 100 kHz and forwards the selected events to HLT. Events are stored or rejected within 3.4 μs while all subdetector data is stored

in a pipeline. During this period, data must reach L1T processing units, finalize a decision, and communicate the result to FE electronics. This limits the processing time within the L1T system to about 1 μs.

The **high-level trigger** is a software-based system running on a CPU farm located on the surface away from the detector. It performs detailed event reconstruction using inputs from all sub-detectors. Advanced algorithms are employed to identify physics objects and apply further selection criteria. Events passing the HLT are permanently stored for the further analysis.
Upon passing both trigger systems, an event is stored with the help of a Data Acquisition (DAQ) system for permanent analysis. The DAQ system is also responsible for transferring data to the L1T system at 100 kHz. The DAQ system [75] utilizes optical links with extensive bandwidth, electronics based on Field Programmable Gate Arrays (FPGA) and a resilient network that includes high-bandwidth optical links and VME or μ-TCA-based telecommunication architectures to efficiently handle a high volume of data. The Data Acquisition (DAQ) system manages data rates reaching several terabytes each second (with the Level-1 trigger), ultimately achieving a recorded data storage and analysis rate of approximately 1GB/s given the average event size of CMS is $\sim$1 MB.

3.2.6 Event reconstruction

The particles of interest such as W/Z, H, top quarks or hypothetical particles predicted by BSM theories are unstable and instantaneously decay before reaching the detector. Their presence is inferred with the signature of their decay particles in different subdetectors. These are primarily electron, photon, muon, and jets (experimental signature of q/g). Event reconstruction in the CMS detector is centered around the Particle Flow (PF) algorithm [8], which integrates information from multiple subdetectors to identify and measure energy of these individual particles in an optimized way.

The first step in the PF algorithm is to reconstruct the tracks. It is performed iteratively to extract the best-quality tracks using hits in the silicon tracker [76]. This method achieves high efficiency and reduced misreconstruction by applying quality cuts on seeds and tracks. The tracker employs a COMBINATORIAL TRACK FINDER (CTF) algorithm based on KALMAN FILTER (KF). Initial seeds are created from pairs of hits in the pixel detector that follow a straight-line trajectory. Advanced algorithms, such as Cellular Automaton (CA), improve efficiency by identifying compatible triplets or quadruplets of pixel hits. These tracks are then iteratively extended by layering compatible hits, with the KF predicting the state vector and covariance matrix at each step. After associating all compatible hits, a final fit optimizes track parameters such as momentum and position. For electrons, a Gaussian-Sum Filter (GSF) accounts for energy loss due to bremsstrahlung radiation.
These tracks are crucial for reconstructing the primary vertex via a DETERMINISTIC ANNEALING ALGORITHM that clusters them by their z coordinate proximity to the beam spot [77]. The vertex position is determined by the adaptive vertex fit that provides high precision, achieving resolutions of approximately 20 μm transversely and 30 μm longitudinally. Accurate vertex identification is essential to mitigate pileup effects and identify displaced vertices from decaying particles such as b-hadrons.

Electromagnetic particles ($e^{\pm}$, γ) deposit energy in the ECAL, whereas hadrons produce showers that extend into the HCAL. Electrons and positrons (charged hadrons)

are signified by an ECAL (ECAL+HCAL or HCAL) linked track, with charge determined by the direction of its bending. Neutral particles, such as photons and neutral hadrons, do not leave tracks. Photons are identified as isolated ECAL clusters without HCAL counterparts, while neutral hadrons are reconstructed from HCAL clusters, optionally linked to ECAL clusters.
The energy scales of the ECAL and HCAL clusters are set by their respective calibrations. ECAL calibration uses isolated electrons, while HCAL calibration relies on isolated charged hadrons with well-measured track momenta. Since hadrons exhibit non-linear energy responses, cluster calibration is essential for accurate energy assignment. This calibration is critical for the precise reconstruction of hadrons and eventually jets and $p_{\mathrm{T}}^{\mathrm{miss}}$.

Muons are reconstructed by combining information from the tracker and muon detectors. Their ability to penetrate multiple detector layers allows for precise momentum determination, critical for many SM and BSM physics. Given the redundant information available, there can be many algorithms to identify muons. In particular, GLOBAL-MUONS are reconstructed with the combined track of the tracker and muon systems. Tracks must match muon system segments and exhibit hits consistent with a muon trajectory.
PF candidates, charged hadrons, neutral hadrons, electrons and photons, are clustered into jets using jet clustering algorithms. Momenta of jets are measured by clustering like PF candidates, charged hadrons, neutral hadrons, electrons and photons, using AK4 and AK8 algorithms corresponding to a radius parameter R of 0.4 and 0.8, respectively. The vector sum of all PF candidates is used to reconstruct the magnitude and direction of MET ($p_{\mathrm{T}}^{\mathrm{miss}}$).
More specific and relevant details of physics objects reconstruction are discussed in Chapter 7.

3.3 CMS detector upgrades along the LHC

To keep up the performance of various physics objects, several component of CMS detector have been upgraded during end of the year and long shutdowns (LS-1 and LS-2). They are referred to as Phase-1 upgrades. During the upcoming LS3, the detector subsystems will be replaced or upgraded to meet the requirements of HL-LHC. These are referred to as Phase-2 upgrades. In this section we briefly summarized the various upgrades undertaken and planned.

3.3.1 Phase-1 upgrades of CMS

Phase-1 upgrade of CMS [78] includes the enhancements to the tracker, hadron calorimeter (HCAL), muon system, trigger, and data acquisition (DAQ) system which are briefly described below.

Pixel tracker : The operation of the pixel detector, which is designed for a luminosity of 1×10^{34} cm^{-2} s^{-1}, is affected by significant data losses in the readout chip.
To resolve this, a sophisticated readout chip is developed to avoid data loss. Furthermore, by moving electronic boards and connections outside the tracking region and using an ultra lightweight support structure and CO_2 cooling, the amount of material used was reduced. Even with the high occupancy challenges of the LHC upgrade, the existing tracking performance levels will be maintained by adding a third series of forward disks and a fourth barrel layer with a radius of 16 cm. Reusing existing fibers and cables is made easier by the development of high bandwidth readout circuits, connections, and DC-DC power converters [78].

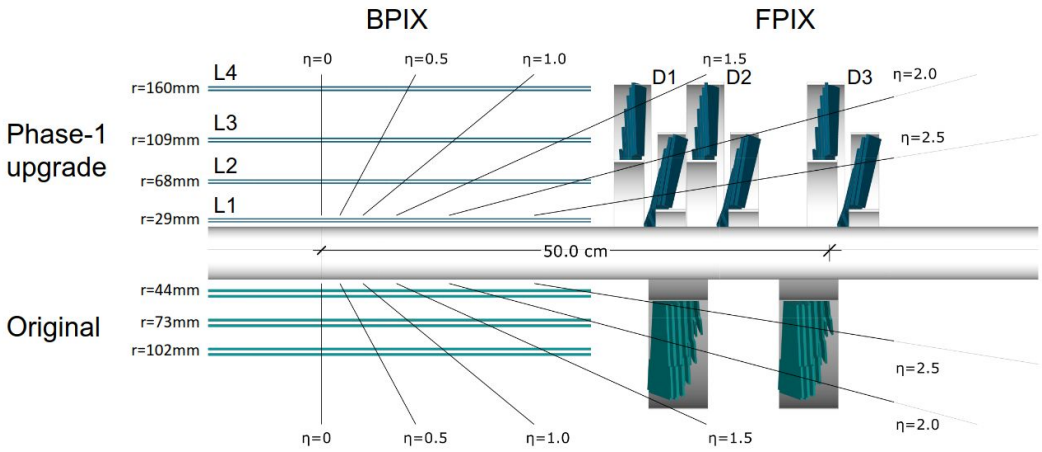


Figure 3.4: The structural configuration of the CMS Phase-1 pixel detector, shown in a longitudinal perspective, is compared with that of the initial detector's arrangement. [79]

Together, the upgrade of the pixel system−such as a longer lever arm, a smaller innermost radius, and a lower mass− greatly improve the current system's tracking, vertexing, and b−jet recognition capabilities. Fig. 3.4 illustrate the comparison of upgraded pixel tracker with old pixel tracker, highlighting the different between the two.

Hadron calorimeter : In the HB/HE/HO part of the HCAL, Hybrid photodetectors (HPDs) are replaced with Silicon Photomultipliers (SiPMs), offering better quantum efficiency, higher gain, and immunity to magnetic fields. SiPMs operates at much lower voltage as compared to HPDs and also immune to high-voltage breakdowns. The SiPMs allow for the more longitudinal segmentation and timing measurements which can be used to mitigate radiation damage effects and improve the energy reconstruction.
Figure 3.5 illustrates the comparison of the longitudinal segmentation between the phase-1 and phase-0 HCAL.

Muon system : Pileup nearing CMS design limits challenges the Muon System in triggering high transverse momentum muons, crucial for detecting significant electroweak interactions. Following upgrades were required to handle the challenges of increased pileup and luminosity,

- The Cathode Strip Chambers (CSCs) are enhanced with a fourth layer of chambers (ME4/2) and upgraded electronics, including new Digital Front End Boards (DCFEBs) for improved strip-level readout, preserving performance in the forward region ($2.1 < |\eta| < 2.5$).
- Barrel Drift Tubes (DTs) undergo maintenance-focused upgrades, such as replacing Theta Trigger Boards with FPGA-based solutions and relocating Sector Collector boards to radiation-safe areas.
- Endcap Resistive Plate Chambers (RPCs) receive a fourth layer to extend coverage up to $|\eta| = 1.6$, with R&D exploring detector technologies for higher η regions, ensuring robust performance at high luminosities.

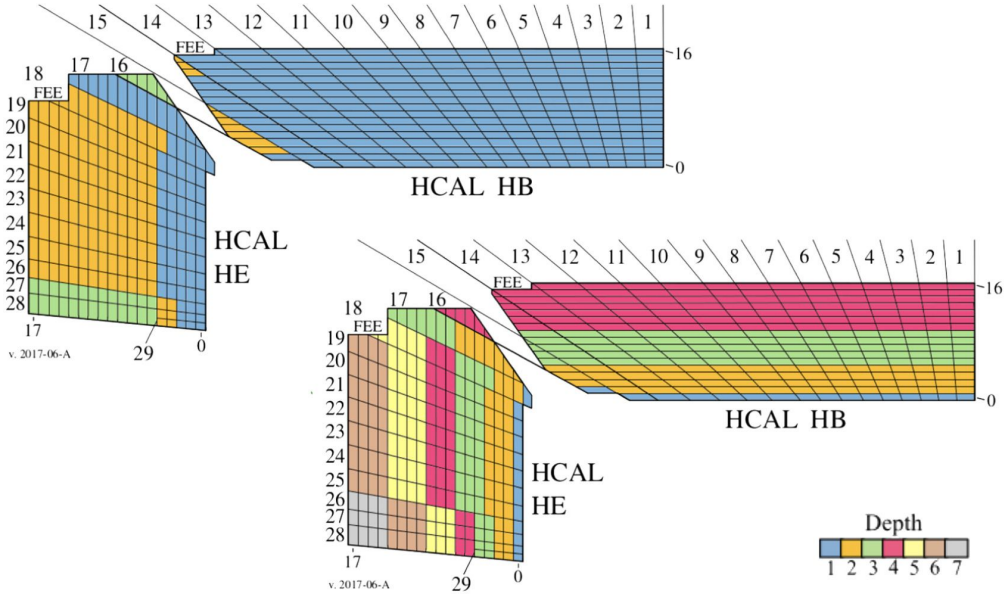


Figure 3.5: HCAL segmentation in Phase-0 (upper section) and Phase-1 (lower section) configuration [80].

- Gas Electron Multiplier (GEM) detectors, known for their thin design and ability to operate in high particle flux environments, were installed in the endcap region ($1.6 < |\eta| < 2.2$) during LS-2.

Figure 3.6 shows the CMS muon spectrometer upgrades.

Trigger and DAQ system : The trigger system was upgraded to address increased detector occupancy and background rates. The full granularity of the calorimeter towers (ECAL + HCAL) was utilized, and the GEM detector information was integrated to improve muon triggering. Advanced algorithms were developed and implemented on custom hardware to enhance efficiency while maintaining a data rate of 100 kHz.
The DAQ system was transitioned from VME-based technology to µ-TCA technology, which is widely used in telecommunications and military applications. This upgrade provided faster and more reliable data transfer, ensuring robust performance under higher event rates and as well as increasing the bandwidth by a factor of 2 to 5.

These comprehensive upgrades, as detailed in the [78], have significantly enhanced the detector's capability to operate efficiently and maintain high data quality under the challenging conditions of Phase-1 LHC.

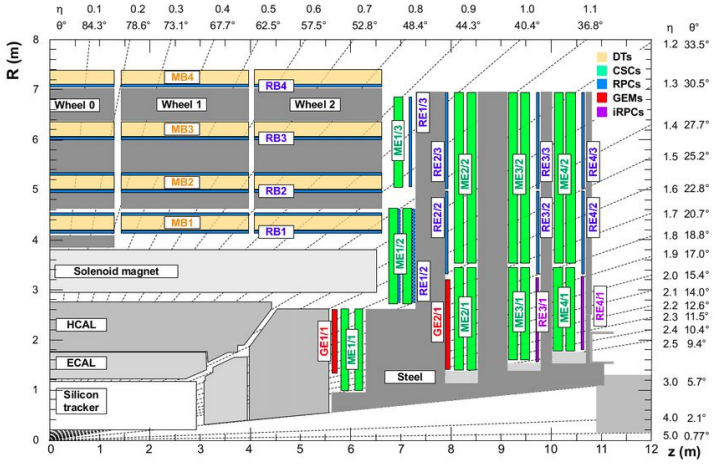


Figure 3.6: The CMS muon spectrometer upgrade project features, highlighted in red (purple), the proposed installations of GEMs and RPCs stations [81].

3.3.2 Phase-2 upgrade of CMS

The primary obstacles during Phase-2 operations for CMS detector [18] are the extremely high "pileup" resulting from the high instantaneous luminosity and radiation damage to the CMS detector caused by integrated luminosity in HL-LHC.

The annual ionization radiation dose during Phase-2 will be equivalent to the total dose accumulated over the entire Phase-1 operation, with forward regions of the detector experiencing doses exceeding 1 MGy as shown in Fig. 3.7 (left) obtained using FLUKA simulation. Radiation damage affects detector components in different ways. In semiconductor devices, non-ionizing radiation creates Frenkel pair defects, affecting charge collection efficiency. Damage to the semiconductor bulk requires higher reverse bias voltages for full depletion width, increasing power consumption. Eventually, damage severely degrades silicon sensors, compromising signal generation and physics performance. In scintillators, radiation damage affects light propagation, creating absorption bands and color centers, reducing light output and transparency. Phosphorescence after irradiation increases noise levels and degrades detector performance.

At HL-LHC, on average, 140 or 200 number of pileup interactions are expected as compared to 60 in the current Run 3 LHC operations. Fig. 3.7 (right) illustrates a simulated event display of the order of 100 pileup interactions occurring under HL-LHC conditions. Such large pileup makes it difficult to accurately reconstruct EM and hadronic showers with overlapping contributions from the neighboring particles.

To address these challenges, the CMS Collaboration is working on extensive upgrades [18] of its subdetectors. These include complete replacement of tracking system, end-cap calorimeters, and electronics system of barrel calorimeters and muon systems. Additionally, new technologies are being integrated, such as GEM (Gas Electron Multiplier) detectors for muon chambers and MIP (Minimum Ionizing Particle) timing detectors.

Tracker upgrades

The Phase-2 tracker [82] will feature a completely redesigned silicon tracker with improved radiation tolerance and four times more granularity. The outer tracker will achieve higher granularity by shortening silicon sensor strips without significantly changing their pitch. The inner tracker will use silicon pixel detectors with improved

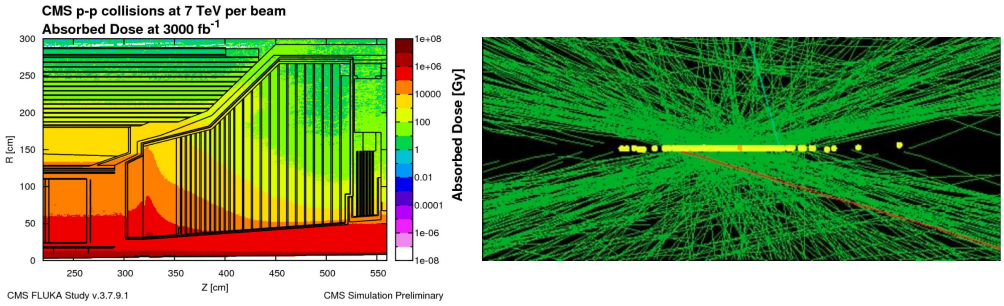


Figure 3.7: Left: The two-dimensional maps in the radial (r) and longitudinal (z) coordinates are shown for radiation dose accumulated in the CMS detector after an integrated luminosity of 3000 fb^{-1}, simulated using the FLUKA program. Right: A model showing multiple p-p interaction vertices (yellow points) during HL-LHC operation, with green lines representing soft particle trajectories from bunch-crossing events [18].

electronics to reduce data loss. For the first time, silicon tracker information will

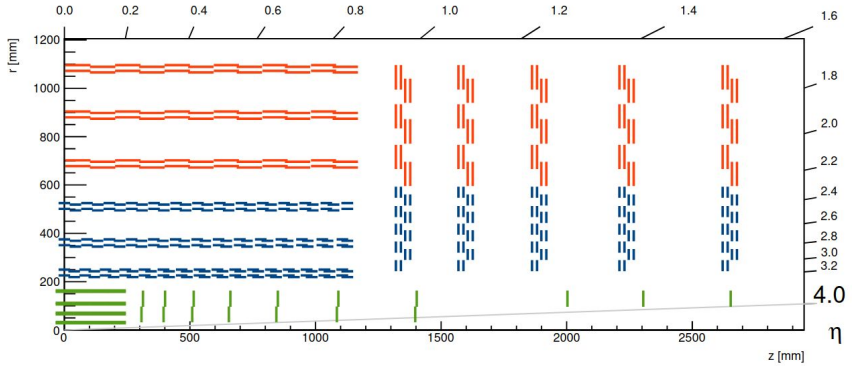


Figure 3.8: Illustration of a quarter section of the Tracker. Blue and red lines represent the Outer Tracker, with blue for PS modules and red for 2S modules (see text for details). The green highlights the Inner Pixel detector and its forward extension [82].

be used in making L1 trigger decision. The outer tracker will have modules for transmitting track stubs to the Level-1 trigger, enhancing track-based triggering. This upgrade reduces tracker material by at least $1X_0$ to minimize scattering and improve performance, ensuring effective background rejection at the start of event selection. Smaller pixels and thinner sensors in the pixel system will enhance track separation and resolution, improving b-tagging, τ-hadronic decay efficiency, and track reconstruction in boosted jets. The coverage will be expanded to near $|\eta| = 4$ with up to 10 additional pixel disks in each forward region. Fig. 3.8 illustrate the silicon tracker structure for Phase-2, also highlighting the stub modules.

Barrel calorimeters

During the HL-LHC, information from each ECAL crystal will be provided to the Level-1 calorimeter trigger with the help of upgraded Front End (FE) card and off-detector electronics. The timing precision of ECAL will be improved up to 30 ps from the current 150 ps. Spikes, or abnormal signals caused by particles striking the APD directly, are eliminated using a redesigned Very Front End (VFE) boards. APD aging effects are lessened by cooling to 9°C (from 18°C in Run 3).

The hadron barrel calorimeter back end (BE) will employ the same ATCA boards for control and status monitoring, data readout, trigger primitive generation, and clock delivery to the FE in order to bring the systems into harmony.

Muon system upgrades

In order to ensure redundant coverage for the four CSC stations in the $1.5 \leq |\eta| \geq 2.4$ region, which is vital for muon detection, these stations will be upgraded with new chambers designed to accommodate higher rate capacities [83]. This enhancement is in alignment with the original CMS design and aims to sustain robust L1 muon trigger acceptance. The first two stations will use GEM chambers for better position resolution, particularly in high magnetic fields. The last two stations will utilize low-resistivity RPCs for better timing despite lower granularity. Additionally, adding a GEM station behind new endcap calorimeters is recommended to extend muon detection to $|\eta| \approx 3$.

MIP timing detector

New MIP timing detectors (MTD) [84, 85] will be installed in barrel and endcape regions to precisely measure time of arrival of particles. The Barrel Timing Layer (BTL) contains LYSO:Ce crystals of size 12 coupled with 4×4 mm^2 SiPMs, while the Endcap Timing Layer (ETL) uses silicon-based Low-Gain-Avalanche-Diodes (LGAD) of size 3 mm^2. Figure 3.9 presents a schematic representation of the planned MTD structure, including the BTL and ETL zones, as configured within the GEANT4 simulation of the CMS detector.

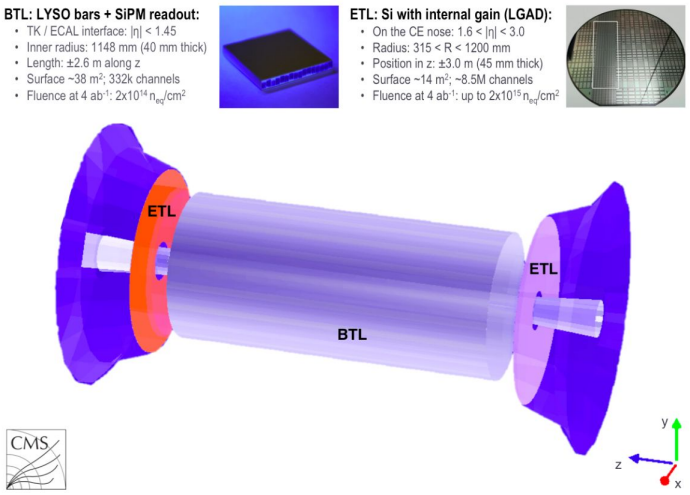


Figure 3.9: GEANT4 simulated schematic representation of Mip Timing Detector (MTD). The setup includes a barrel layer (grey cylinder) at the tracker-ECAL junction and two silicon timing layers (orange and light violet discs) before the endcap calorimeter. [84].

Trigger and DAQ upgrades

The L1 trigger's latency is set to increase to 12.5 μs for Phase-2 operation, allowing for hardware track reconstruction and matching to muons and calorimeter information, enabling more sophisticated algorithms for event selection [86]. This change will require upgrades to readout electronics in existing sub-detectors. The proposed

L1-trigger acceptance rate is 500 kHz for beam conditions yielding 140 PU, while 750 kHz is needed for 200 PU.
The Data Acquisition (DAQ) system [87] will be upgraded to increase bandwidth and computing power to accommodate larger event sizes and L1-trigger rates at high PU. This will increase the bandwidth and computing power requirements by 10(15) and 15(30) factors, within Phase-2's expected network and computing technology capabilities.

Endcap calorimeters

Given the large center-of-mass energy of the pp collisions at the LHC, number density of outgoing secondary particles in higher η regions is much higher than in the lower η regions [88]. Fig. 3.10 (left) shows ECAL crystals in highest η regions have already lost up to 90% of their transparency and these will not maintain acceptable level of performance for physics analysis. Expected degradation in response of HCAL towers in endcap is shown in Fig. 3.10 (right) with an accumulated integrated luminosity of 1000 fb^{-1}.

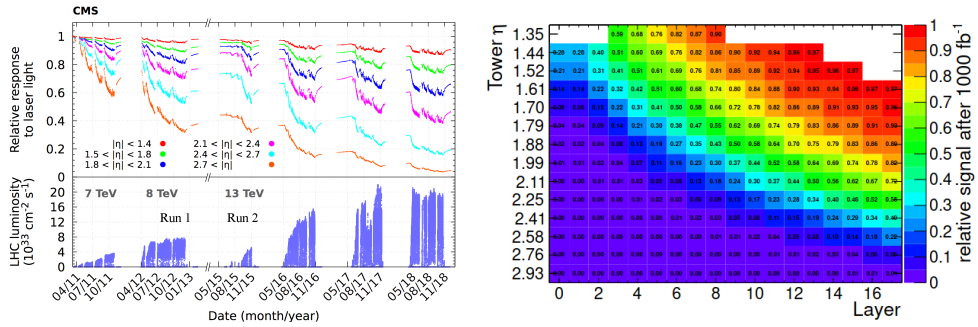


Figure 3.10: The left plot is showing the relative response to laser light as a function of accumulated integrated luminosity till Run 2 [89]. The right plot is showing the expected attenuation of scintillator light signals at HCAL endcap after integrated luminosity exposure of up to 1000 fb^{-1} [18].

The endcap calorimeters ($1.5 < |\eta| < 3.0$) will be replaced with a high-granularity calorimeter (HGCAL) based on silicon and scintillators directly readout by SiPM technologies. More details are discussed in the following.

3.4 The CMS High Granularity Calorimeter

The objective of the HGCAL is to precisely reconstruct jets, MET, b-jet tagging, and hadronically decaying tau leptons to achieve the physics goals of the HL-LHC. This necessitates the efficiency in identifying electrons, photons, and charged and neutral hadrons within the endcap calorimeters. The endcap calorimeter needs to be upgraded with:

- Radiation hard technologies for detector and readout electronics components to withstand harsh radiation environment while maintaining a good signal to noise (S/N) ratio for Minimum ionizing particles (MIP) throughout the HL-LHC operations for a decade.
- High lateral granularity is essential for reconstructing particle showers that originate from nearby particles while also maintaining the compactness of EM showers. This is necessary to effectively reject particles that arise from pile-up interactions.

These objectives are achieved for the HGCAL by opting for sampling calorimeter design with active-material silicon sensors in high-radiation regions and scintillator tiles directly read out by SiPMs in lower-radiation regions, as shown in Fig. 3.11. It uses Pb and PbW or stainless steel as absorber material. The total area covered by silicon sensors is around 620 m^2, while scintillator tiles cover a total of $\sim$ 320 m^2 [90]. Each of the two endcap calorimeters weighs about 230 tonnes and will consume 125 kW per endcap [91]. The whole system must operate at -30^{0}C to minimize the impact of radiation damage on noise levels. The HGCAL corresponds to 6M silicon and 240k scintillator readout channels posing an unprecedented challenge of data handling at the detector front-end (FE) and back-end (BE). In contrast, we have approximately 14 k ECAL crystals and 2.3 k readout channels in HCAL in the current CMS endcap calorimeter. The detector channels are equipped to also provide the time of arrival of the signal with a resolution of 50 ps, thus providing a 5D measurement of the development of the particle shower, namely energy, spatial coordinates, and time.

The HGCAL is expected to generate a data throughput of approximately 100 Tb/s at 40 MHz and approximately 40 Tb/s at 750 kHz corresponding to the L1A trigger rates, designated for offline storage.To emphasize the scale of these high data rates, consider CMS Silicon Tracker for HL-LHC equipped with approximately 2.2 billion readout channels [92] that generates data rates that are roughly half of what is expected from HGCAL [93]. This difference arises because, unlike the binary "yes" or "no" information from the tracker, HGCAL transmits 5D information (energy, position, and time) per channel.

For energy measurements and offline processing of the data, signals from all HGCAL layers (including CE-E and CE-H) are readout. For constructing L1 trigger primitives, only the alternate layers of CE-E and all layers of CE-H are readout.

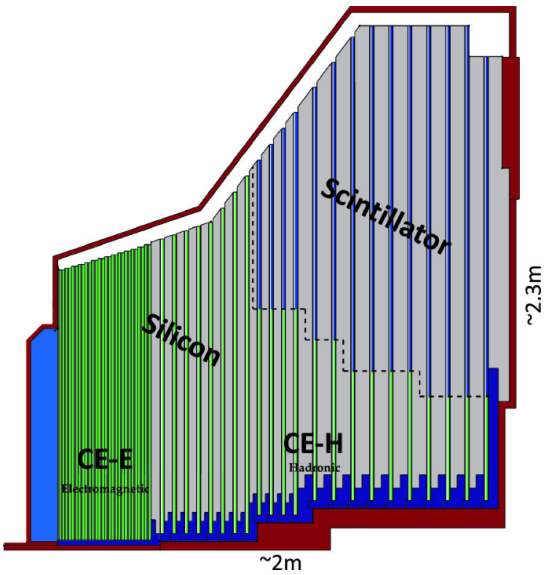


Figure 3.11: The cross sectional view of HGCAL detector highlighting its Silicon and Scintillator based CE-E and CE-H sections [94].

3.4.1 Electromagnetic and hadronic sections

The CMS HGCAL is a sampling calorimeter, comprising a 37 cm electromagnetic section (CE-E) followed by a 157 cm deep hadron section (CE-H) with combined depth enough to contain electromagnetic and hadronic showers of single particles up to the TeV energy scale.

Electromagnetic section (CE-E) :
The longitudinal sampling of CE-E is optimized to maintain the current best resolution achieved for H $\rightarrow \gamma\gamma$ even though the resolution of a single photon is worse than that achieved in the current ECAL based on PbWO4. CE-E consists of 26 sampling layers with a depth of approximately 26 X_0 or 1.3 λ_{int}, equipped with silicon as active material. An example layer from CE-E is shown in the middle plot of Fig. 3.12. Lead (Pb) with X_0= 0.5612 cm, & λ_{int}= 17.59 cm is used as absorber material in CE-E. Its Moliere radius of 2 cm ensures compact electromagnetic showers with minimal transverse spread, helping to separate showers due to nearby particles. The Cu and CuW (1.4 mm) plates used to provide support structure and cooling for silicon modules contribute to additional absorber material. The fine longitudinal and transverse granularity along with timing information helps in pile-up rejection as well as maintaining a good electromagnetic shower resolution.

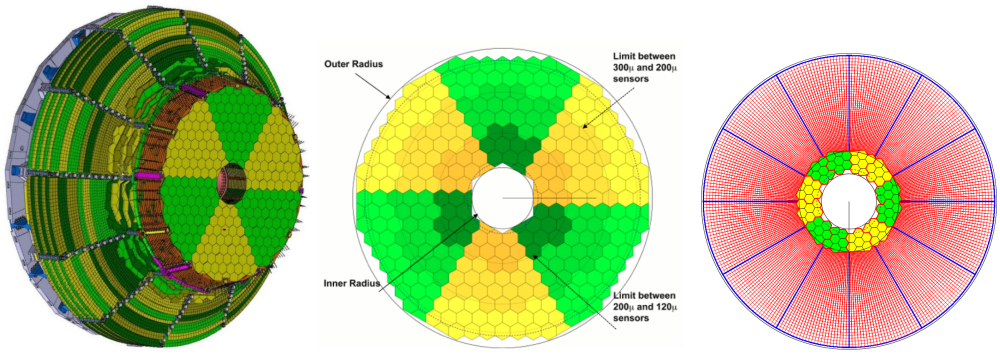


Figure 3.12: The above is showing the image of 3D view of 1 HGCAL endcap (left most), a sampling layer of CE-E (middle) and that for CE-H (right most) [94].

Hadronic section (CE-H) :
The CE-H section has a total of 21 sampling layers with a depth of 8.5 λ_{int}. Seven sampling layers are equipped with silicon as the only active material as in CE-E layers and 14 are mixed layers using silicon in regions close to beam pipes and scintillators otherwise, as shown in Fig. 3.12 (right). The absorber material used for CE-H is stainless steel (X_0= 1.75 cm & λ_{int}= 16.77 cm). The supporting material for silicon modules (Cu & CuW) do not significantly contribute to the amount of absorbers in CE-H, as their thickness in terms of $\lambda_{int}$is very minimal compared to Steel.
The total depth of HGCAL that includes CE-E and CE-H and neutron moderator to reduce neutron fluence is around 10.7 λ_{int}.

3.4.2 Silicon sensor modules and scintillator-SiPM tileboards

Silicon sensors and modules :
Thin silicon sensors, widely used in tracking detectors close to the pp collision point, are well established for their radiation hardness and sustained physics performance [95]. For HGCAL, n-on-p silicon sensors, which have an n-type layer on a p-type bulk, are used, as charge collection efficiency is less sensitive to the impact of radiation damage in the bulk [96] even after exposure to fluences of the order of 10^{16} n_{eq}/cm^2, approximately 50% higher than expected at the HL-LHC [94]. HGCAL silicon sensors use 8$^{'}$ circular wafer technology, which is manufactured in hexagonal shapes with truncated tips to keep the waster and dead regions of the material minimal.
Each silicon sensor is further divided into small independent hexagonal cells that can be readout independently. Silicon sensors with three different active thicknesses

are used, including 300 μm, 200 μm, and 120 μm, corresponding to regions with increasing level of radiation dose. The cell size is determined by a trade-off between the inherent noise of the sensor cell, which scales with the total cell capacitance, and the sensor's thickness and pad gap. Consequently, the sensor cell areas are chosen as 0.52 cm^2 for 120 μm and 1.1 cm^2 for 200 μm & 300 μm thick sensors. These correspond to 436 cells (high density sensors) and 192 cells (low density sensors), and are referred to as high density modules (HDM) and low density modules (LDM), respectively. These silicon sensors are operated above the full depletion voltage to achieve optimal charge collection efficiency. The operating voltage is designed to vary between 200 V and 800 V to maintain charge collection efficiency through the HL-LHC phase. Each sensor stacked along with its mechanical structure and printed electronics circuit board is referred to as a silicon module. Most modules are full hexagons ('F'), but to fill the edges where full modules cannot fit, irregularly shaped modules, called 'partials', are used. Fig. 3.13 shows various layers of stacked silicon module [94]. The

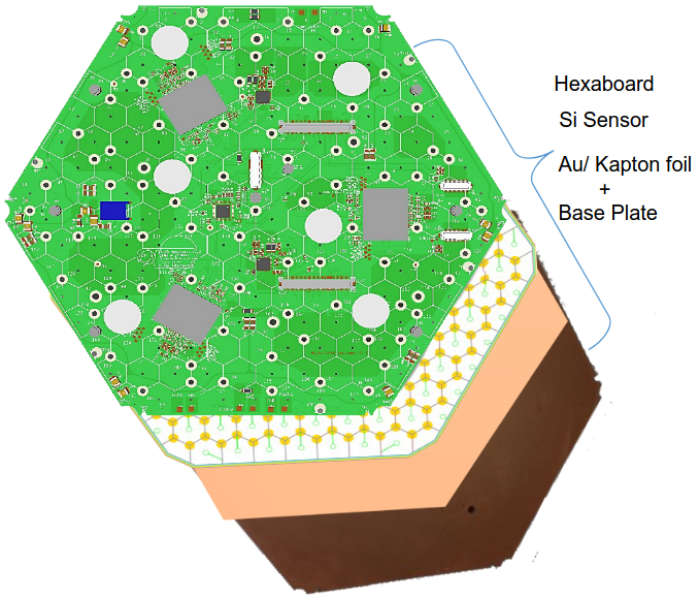


Figure 3.13: This image is showing the different components stacked to make a silicon module [96].

description and purpose of each component follows:

- Baseplate : The CE-E base plate is made from a tungsten-copper (WCu) composite for good thermal conductivity and reduced thermal expansion. In CE-H, carbon fiber is used for its high thermal conductivity.
- Kapton-Gold Sheet : A 105 μm thick sheet provides electrical insulation between the base plate and the sensor, with a gold layer for high-voltage connection.
- Silicon Sensor : The core component of the module, responsible for charge detection.
- Hexaboards Printed Circuit Board (PCB) : Connects sensor to HGCROC and provides bias voltage. House readout chips for signal processing and data digitization. More details of other ASICs hosted on Hexaboards are described here 6.1.

The CE-E consists of alternating absorber and silicon active layers, with hexagonal sensor modules mounted on both sides of a cooled copper plate. The CE-H is made of alternating absorber and active layers.

Plastic scintillators and Silicon photo multipliers (SiPMs) :
Plastic scintillator tiles of area 4−30 cm^2 are used as an active detector component in the low-fluence region of CE-H. These scintillator with SiPMs are placed on a PCB and are referred to as tile boards as shown in Fig. 3.14.
SiPM readout, already used in the HCAL Phase-I upgrade, provides high gain (up

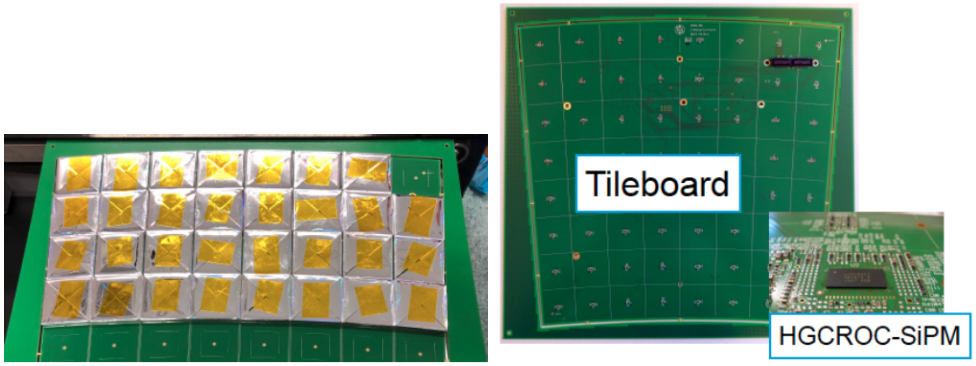


Figure 3.14: This image is showing the tileboards to be used in scintillator sections [97].

to 10^4) and more than 30% photo detection efficiency. Like silicon sensors, the dark current of SiPM increases significantly with radiation exposure, and the rate at which pixels fire is linear with radiation fluence. The MIP signal is inversely proportional to the square root of the SiPM tile area, with smaller tiles yielding more photoelectrons. Therefore, smaller tiles are used in areas of high radiation fluence, arranged in an r-ϕ grid with equal radial and azimuthal widths at the center of the tiles. A total of ∼4k scintillator tileboards will be used in the HGCAL.

HGCAL mechanical layout :
Silicon sensors modules and scintillator tileboards along with a service motherboard that provides power and electronics are arranged in double-sided (single-sided) cassettes of size 60° (30°) for the CE-E (CE-H) section [94], as shown in Fig. 3.15 (left) (Fig. 3.15 (right)) . These modules are serviced by a motherboard that provides power and electronics, which forms part of an assembly called a cassette as shown in the left of Fig. 3.15. An HGCAL layer is formed after putting six of such cassettes together. The HGCAL is designed to have a 120° azimuthal symmetry in each of the identical

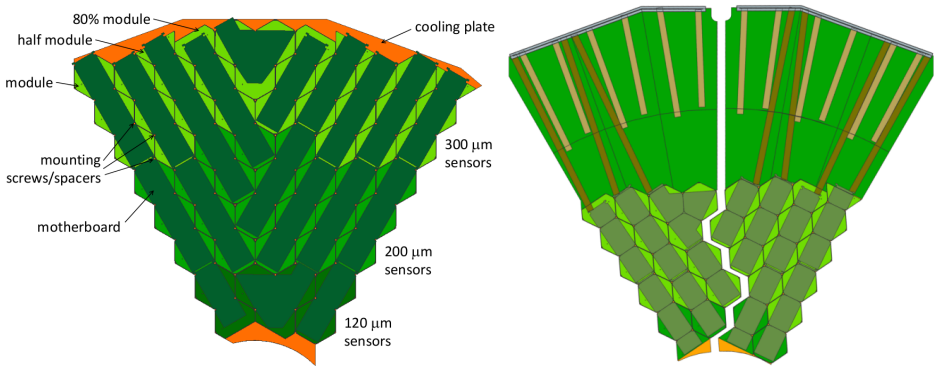


Figure 3.15: An example 60° cassette and two 30° mixed silicon and scintillator cassettes of CE-E and CE-H are shown from left to right respectively. [94].

end caps, resulting in 6 identical 120° sectors. Each silicon module and tileboard can be identified based on identical coordinates (u, v) in 6 sectors and are shown in

Fig. 3.16.

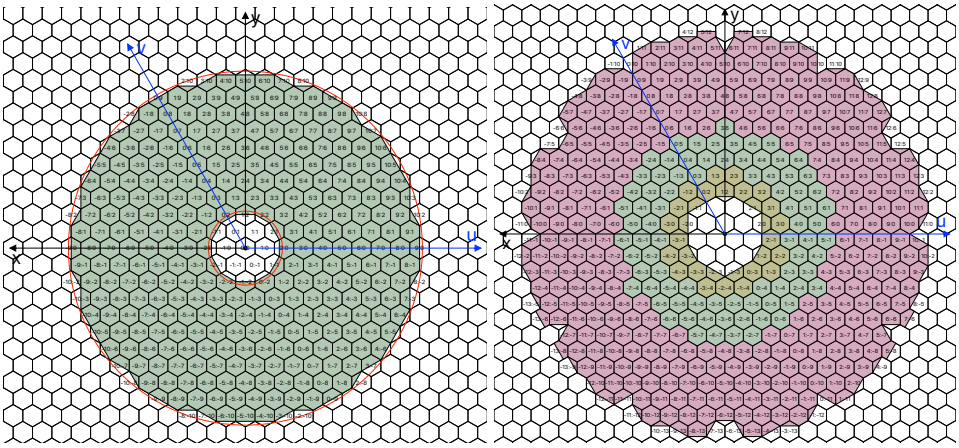


Figure 3.16: The above figure illustrates the (u,v) coordinate systems as well as highlighting the three sectors per endcap in HGCAL layer-1 and layer-35.

3.4.3 Expected physics performance of HGCAL

HGCAL is designed for dense radiation environments with high pileup, using fine granularity and precise timing for 5D imaging calorimetry. It improves the separation, reconstruction, and identification of physics objects like e, γ, μ, τ, jet and $p_{\mathrm{T}}^{\mathrm{miss}}$, enabling precise measurements critical for various physics studies. These improvements enhance the discovery potential for SUSY particles and provide stronger constraints on their parameter spaces, especially for searches like presented in this thesis (Chapter 7).

One of the most ambitious aim of HL-LHC is to measure Higgs coupling to all SM particles with minimum uncertainties, as well measurement of its self coupling which is predicted by SM but has not been observed yet.

The Higgs boson production via vector-boson-fusion (VBF) process can be identified by its two forward energetic jets and huge gap in pseudorapidity during the event. Similarly for quark originated jets or VBF jets, the ability of the HGCAL to analyze radial and longitudinal shower profiles facilitates the differentiation between pileup jets and VBF/quark jets as shown in Fig. 3.17. Pileup jets consist of lower-energy particles resulting from multiple interactions. Hence, pileup jets have broader radial profiles in contrast to quark jets due to their higher energy density.

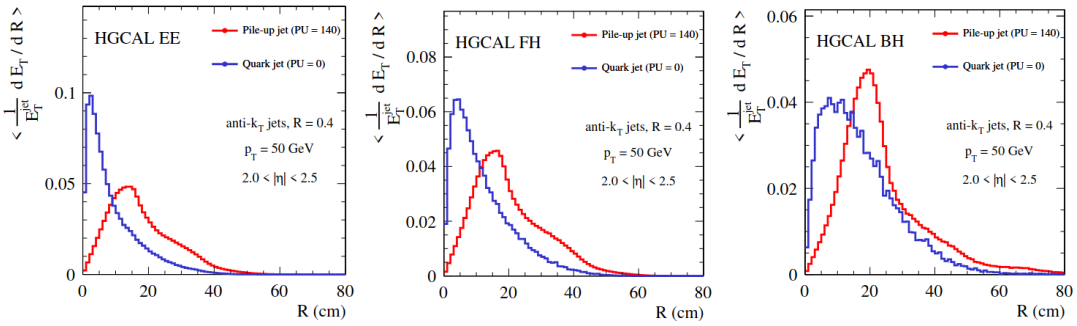


Figure 3.17: Comparison of radial profiles of pileup originated jets (red) and quark originated jets (blue) across the longitudinal depth of HGCAL [94].

In high-luminosity environments, separating jets from hard scatterings amidst pileup jets poses a significant challenge. HGCAL addresses this with its precise timing resolution ((< 30) ps for high-energy clusters), which associates energy deposits with their correct primary vertices, effectively mitigating pileup contributions. Figure 3.18 illustrates a VBF event characterized by the simultaneous occurrence of a photon and a VBF jet within a single quadrant. The figure shows event displays of both with (right) and without (left) a timing selection. Using timing information, pileup is significantly reduced, resulting in improved clarity of energy clusters associated with the photon and the VBF jet. These enhancements significantly improve the sensitivity of analyses focused on VBF Higgs generation.

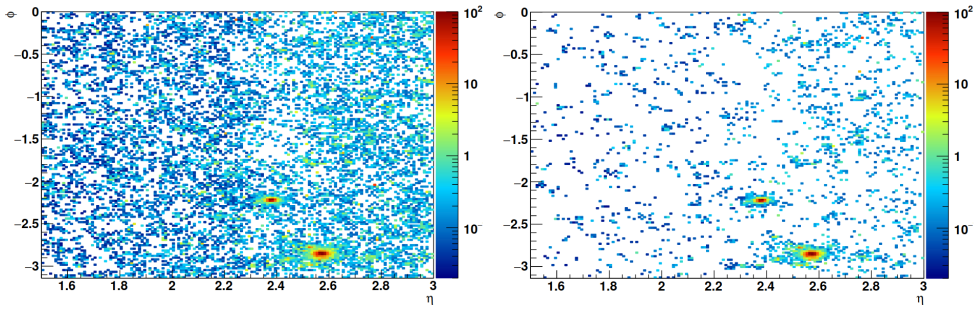


Figure 3.18: Depicted is a singular event illustrating the VBF process with the decay H→2γ, where one photon and a pair of VBF jets are confined to one quadrant of HGCAL, projected onto its outer surface [94].

More physics studies illustrating the benefit of HGCAL can be found in [94].

HGCAL beam test prototype

The CMS HGCAL collaboration has conducted comprehensive beam test experiments in October 2018 at H2 area of CERN's SPS facility. These experiments aimed to validate the performance of silicon sensor prototypes and readout electronics under conditions replicating the operational environment of the final detector. A full longitudinal section of HGCAL was tested with the positrons, negative charged pions and muon beams with momenta ranging from 20 to 300 GeV. The primary goals included ensuring synchronized data acquisition across multiple detector units, maintaining pedestal & noise levels throughout the run, and demonstrating HGCAL's capability to measure electromagnetic and hadronic showers originated from positrons and charged pions. The dataset collected served as a reference for comparison with GEANT4-based simulations, providing critical insights for development of detector simulation.

At CERN, the SPS accelerates protons to 450 GeV, which are the divided and configured to direct at several fixed target stations by varying the bending angle of proton beams. Parallel beam-lines deliver the secondary particle beam formed during the encounter to the experimental halls. H2 beam-line is one such beam line where proton is directed onto a 500 mm Beryllium target (X_0 = 35 cm, λ_{int} = 42 cm) and secondary beam of particles are produced. The fixed target material is required to be unbiased with respect to both electromagnetic and hadronic showers, implying that its radiation length and interaction length must be comparable. The momentum selection, beam collimation, and particle type refinement are achieved using a series of bending magnets, collimators, and particle identification devices such as Cherenkov detectors. The beam purity is further enhanced using secondary targets and absorbers, ensuring specific particle types dominate the beam composition. For example, 40 cm thick Copper block or 4 mm thick lead block can be used to absorb soft hadrons and electrons, respectively. Muon beams are generated in pion or kaons decay.

The H2 beam-line delivered particles with momenta ranging from 10 to 400 GeV/c with an uncertainty margin of 2%, to the experimental hall. The beam underwent additional collimation and focusing to reduce its spot size and divergence, optimizing conditions for prototype evaluation. Each 30-second SPS spill produced a particle count between 10^3 and 10^7, enabling extensive data collection during the beam test campaign. The beam test experiment at H2 beamline utilized various detectors, including Cerenkov detectors for particle identification, four delay wire chambers (DWCs), micro-channel plate (MCP) detectors for timing reference, and two scintillators.The DWC tracking information provided crucial data for data analysis, such as selecting one silicon cell per layer through which MIP-like particles pass. Two micro-channel plate (MCP) detectors were placed in front of CE-E to provide timing reference. Two scintillators were placed in front of CE-E to generate coincidence signals, and another larger scintillator was placed behind CE-H to generate veto-triggers for e+ runs. The experiment aimed to improve the efficiency of the detectors used in the experiment.

Prototype geometry and Data Acquisition (DAQ)

The 2018 HGCAL beam test prototype [98] consisted of two primary sections: the silicon-based electromagnetic calorimeter (CE-E) and the hadronic calorimeter (CE-H). These were complemented by the CALICE AHCAL prototype, a scintillator-based sampling calorimeter. Fig. 3.19 shows the mounted HGCAL beam test prototype along with CALICE AHCAL.

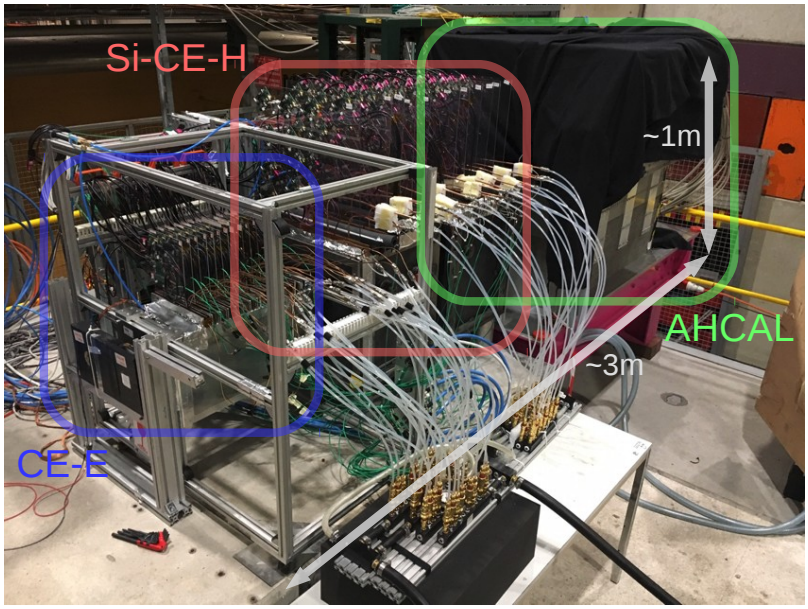


Figure 3.19: A picture of the CE-E, CE-H, and AHCAL prototypes setup mounted on a concrete platform in the H2 experimental area of the CERN SPS.

Each section represents a sampling calorimeter, composed of several sampling layers of segmented active material and absorber material. The different parameters of three sections are discussed as following;

- **CE-E section :** Silicon sensors with an area of 1.1 cm^2 are combined with Cu/CuW and Pb absorbers, forming 28 sampling layers with one module per layer as shown in left most image in Fig. 3.20. Each module has lateral dimensions of 15 cm $\times$ 15 cm. The absorber layers consist of 6 mm thick copper cooling plates, 1.2 mm thick copper-tungsten baseplates, 4.9 mm thick lead plates, and 300 μm thick stainless steel cladding. The total thickness of CE-E section is 1.4 λ_{int}.
- **CE-H section :** Silicon sensors of the same size as to CE-E are paired with Cu/CuW and steel absorbers, creating 12 sampling layers. Copper cooling plate (6 mm thick) do not contribute significantly as the absorber in this section. This Copper cooling plate is sandwiched between 4 cm thick steel absorber plates. Each layer typically contains seven modules arranged in a daisy structure as shown in Fig. 3.20 (middle), except for the last three layers,each of which instrumented with only single module. The lateral dimensions of these modules are 37 cm $\times$ 37 cm. The total depth of CE-H section is 3.4 λ_{int}
- **AHCAL section :** It features scintillators coupled to SiPMs, each with dimensions of 3 $\times$ 3 $\times$ 0.3 cm^3, interleaved with steel absorbers of 17 mm thickness across 39 sampling layers. The lateral dimensions of the CALICE AHCAL modules are 72 cm $\times$ 72 cm. With a little indentation in the middle that accommodates the SiPMs positioned on the PCBs, the scintillator tiles are injection-molded from polystyrene. The SiPMs are of the Hamamatsu type S13360-1325PE, measuring 1.3$\times$1.3 mm^2 and containing 2668 pixels [99]. A picture of two scintillator tiles with SiPMs installed on a base unit is displayed in Fig. 3.20 (right). The total depth of AHCAL is 4.4 λ_{int}

Each silicon module contained 128 hexagonal cells with independent readout for

amplitude and timing data. This makes a total 3.5 k and 8.5 k channels to be readout independently from CE-E and CE-H, respectively. The CALICE AHCAL was equipped with 22 k readout channels.

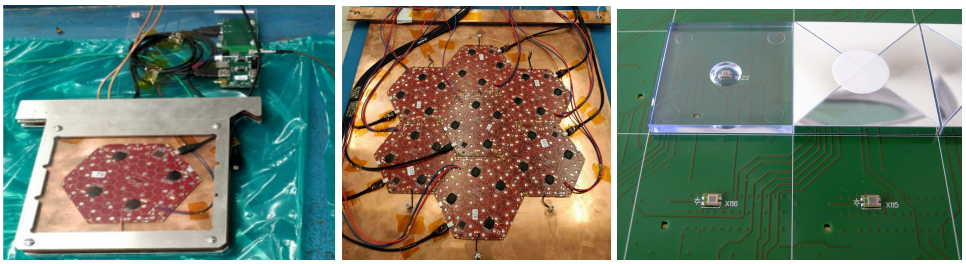


Figure 3.20: A picture of the CE-E, CE-H, and AHCAL prototypes setup mounted on a concrete platform in the H2 experimental area of the CERN SPS.

Signals from hexagonal silicon cells in the HGCAL prototype are processed by four SKIROC2-CMS ASICs [100] mounted on the PCB positioned on top of silicon sensor. These ASICs are connected to the silicon sensor via wire bonds. The SKIROC2 chips does the signal processing, which recorded amplitude and timing for each detector cell. The signal produced in the sensros ranges from a few femtocoulombs (fC) to about 10 picocoulombs (pC), representing energy deposited by a MIP like or the core shower particles. The SKIROC2-CMS ASIC employs two amplification modes: high-gain and low-gain, as well as a time-over-threshold (ToT) measurement to handle this wide dynamic signal generated. It also gives a time of arrival (ToA) measurement. Signals are sampled at 40 MHz for 13 samples which are digitized with a 12-bit analog-to-digital converter (ADC) and saved for offline processing. ToT and ToA are measured using a time-to-digital converter (TDC) with a resolution of 25 ps per time bin[101].

Apart from the electronics to digitize signals from silicon sensors, HGCAL beam test was also equipped with the back-end DAQ system [101] using custom front-end electronics to handle data aggregation, processing, and transmission. The system handled high data rates with precise synchronization, real-time monitoring, and data integrity checks. Key features include a trigger and timing system for event alignment.

For the AHCAL, 144 SiPMs are readout together by four SPIROC2E ASICs [99] with each SiPM collecting light from an individual scintillator tile. It allows charge measurements in high-gain and low-gain modes, handling signals from 160 fC to 320 pC with a SiPM gain of 10^6. The SPIROC2E ASIC also performs time measurements using a TDC. Signals are digitized internally via ADCs, with only two measurements per channel digitized for further analysis.

Simulation of beam test setup and Event reconstruction

The detector geometry was simulated using GEANT4 [58] version 10.4.3, with corresponding absorber and active material in CE-E, CE-H, and AHCAL sections. The event simulation discussed in Section 2.10 is repeated here with updated absorber, and active material and respective layer thicknesses taken from the beam test experiment to accurately simulate detector geometry. Figure 3.21 presents a schematic representation of the HGCAL and AHCAL prototype detectors utilized in the beam test experiment. The silicon sensors (for CE-E and CE-H) and scintillator tiles (AHCAL) were made to be "sensitive" elements, recording and saving energy from shower

particles as "simHits" with their corresponding "detectorID" representing the location of these cells.

Beam optics and detector materials in the H2 beamline influenced single particle beams reaching the calorimeter setup. GEANT4's G4BEAMLINE package simulates these beamline effects. These data serve as input for the subsequent phase of the simulation process, namely, the HGCAL beam test setup simulation. The particle gun is placed at z = 0 cm, and all particles arriving at the HGCAL front are shot into the detector setup with their corresponding species, momentum, and coordinates.

The `simHits` generated from GEANT4 need to be reconstructed into a format similar to the data for one-to-one comparison. To convert the energy of simulated hits from GeV to MIP, a hard-coded factor is used for silicon. To account for electronics noise, the converted energy is smeared with a Gaussian distribution with $\mu = 0$ and $\sigma = 1/6$ MIP., resulting in a shift of 4% for 300 and 200 μm thick silicon.

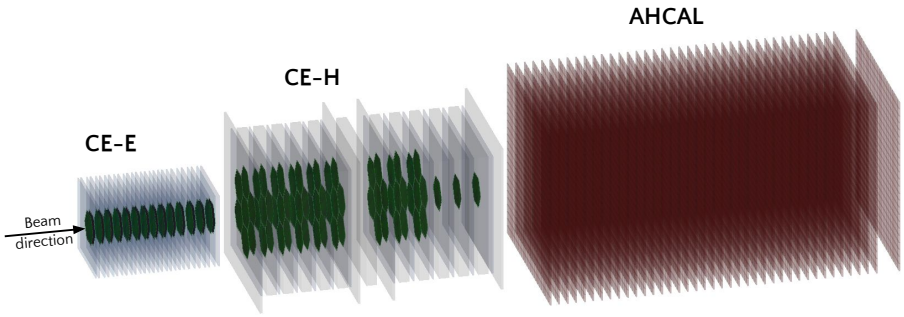


Figure 3.21: Simulated geometry of the CE-E, CE-H and AHCAL detector prototypes depicting the position of active layers and absorbing material.

The simulated event samples are produced using the GEANT4 toolkit implemented in the CMS software. The `FTFP_BERT_EMN` and `QGSP_FTFP_BERT_EMN` physics lists are used to model the development of hadronic showers and compare & understand data-MC differences. More detailed investigation of these comparisons is documented in this thesis [102].

The data for offline analysis include the signal amplitude, ToT, and ToA for each hit, along with the physical locations and electronic ID of each cell. The process involves converting these raw data to energies measured by calorimeters, event reconstruction, and calibration of the detectors.

For each silicon cell, thirteen samples are acquired at intervals of 25 ns, resulting in an amplitude waveform as a function of time. For each silicon cell, pedestal noise and common mode noise are subtracted from the ADC counts in each time sample before the signal extraction via fitting the waveform. The corresponding high gain or low gain amplitude is extracted after fitting the waveform as a function of time. The signal corresponding to each cell is converted to equivalent high gain ADC counts using a gain linearization procedure [53]. The linearized ADC value is then converted into the corresponding number of minimum ionizing particles (MIPs) using 200 GeV muons. This method involves fitting the ADC distribution within a cell using a Landau function convoluted with a Gaussian function, where the most probable value serves as the ADC to MIP calibration factor. Compared to MIPs, 200 GeV muons deposit more energy but still provide a reliable calibration tool for detectors. Ideally, each sensor cell is calibrated using muons, but during beam

testing, the muon beam's limited transverse size covered only some sensor cells across different sampling layers. Similar way, the ADC to MIP calibration was obtained for ≈85% of cells in CE-E and CE-H. Similar procedure is followed for AHCAL. This is elaborated in more details in [99, 102]. In simulation, the energy distribution of 200 GeV muons is fitted with the same functional form as used in the data, and the maximum value of the fitted function is used to convert the energy deposited in each cell to the number of MIPs. Gaussian smearing is applied to the simulated MIP spectra before the fitting procedure. Each channel's digitized signal is represented into energy values expressed in the number of minimum ionizing particles (MIPs).
The channel-to-channel response in the simulation is equalized for a given thickness of the silicon cells or scintillator tiles. For the AHCAL, a threshold of 0.5 MIP is applied in the reconstruction, well above the self-trigger threshold for all channels. Data and simulation were found to be in good agreement.

The 2018 test beam experiments for the CMS HGCAL demonstrated the feasibility of silicon-based calorimetry for high-granularity applications. The synchronized data acquisition system [101], effective calibration techniques, and agreement between experimental and simulated data [52] underscored the readiness of the HGCAL technology for deployment in the CMS detector. The agreement of energy scales in data and simulation to better than 3% (9.5%) for electromagnetic and hadronic showers [98].
Data recorded with this beam test setup and its simulation has been used to develop and validate the machine learning based reconstruction of charged pions (See Chapter 5).

Chapter 4

Advanced machine learning algorithms

In experimental particle physics, machine learning (ML) has become an essential tool, especially for experiments such as CMS at the LHC that produce large and intricate datasets. ML algorithms offer strong solutions for a variety of tasks, such as data processing, event reconstruction, physics analysis, and detector optimization, in light of the growing amounts of data and the difficulty of separating rare signals from overpowering backgrounds. For classification and regression problems, conventional techniques, such as decision trees and deep neural networks (DNNs), have been extensively employed. Despite challenges such as computation, interpretability, and biases, modern methods like graph neural networks are revolutionizing particle tracking, calorimeter data interpretation, detector simulations, and data augmentation in particle physics.

The CMS experiment uses calorimeters to measure the energy of showers generated by electron, photon, charged, and neutral hadrons by completely stopping these particles. These showers have significant event-to-event variations with sparse and irregular energy deposits, meaning that many detector cells show no signal. The CMS HGCAL provides a unique opportunity to deploy such advanced ML algorithms to benefit from its highly granular information corresponding to signal generated by traversing particles. Unlike traditional calorimeters, this richer information enables ML algorithms to utilize spatial and energy information to accurately reconstruct hadronic and electromagnetic showers. Traditional ML models such as MLPs [103] and BDTs struggle with irregular datasets, as they depend on predefined input features and fixed input sizes, which are not optimal for capturing all relevant properties of these showers.

Modern ML techniques excel in feature learning and surpass traditional models that rely on manual feature engineering, especially with raw detector data such as calorimeter hits [104, 105, 106, 107]. Although promising, CNNs face challenges with fixed-size inputs, requiring zero padding or truncation, and difficulty in representing data across multiple detector layers with different granularity. Future ML advancements aim to overcome these issues by incorporating domain-specific knowledge, boosting model robustness, and expanding real-time applications, which will enhance particle reconstruction and deepen our understanding of fundamental physics. Graph neural network is one of such modern algorithm equipped to handle large variations and irregular data.

A substantial component of this thesis work is devoted to the reconstruction of charged-pion energies in the HGCAL using an advanced machine learning approach, notably based on the dynamic graph convolutional neural network (DGCNN). In this chapter, the basic building blocks of machine learning are discussed in Section 4.1, giving a brief overview of neural networks as well. Section 4.2 presents the basic

structure of graph convolutional networks and DGCNNs. Section 4.3 summarizes the dynamic reduction network used to reconstruct hadronic showers in HGCAL.

4.1 Building blocks of machine learning

Machine learning algorithms are used to identify patterns and make decisions from data, while improving their performance with the experience gained along with exposure to more data. These models are categorized into supervised, unsupervised, and reinforcement learning. In unsupervised learning, hidden patterns within unlabeled data are discovered, such as in clustering tasks. Reinforcement learning involves agents learning optimal decision making by optimizing reward criteria. Supervised learning is further divided into tasks such as classification and regression and is discussed in the remainder of the chapter.

A fundamental instance of supervised machine learning algorithms is linear regression, which establishes the connection between a dependent variable and one or more independent variables. Suppose there exist a dataset of X_i (input features) and a variable Y_i is dependent on this X_i linearly such as,

$$Y_i = A_0 + A_1 X_i \tag{4.1}$$

where A_0 and A_1 are the unknown biased and weights (parameters). Fig. 4.1 (left) illustrates the relation between Y_i and X_i. The task of the algorithm is to find the

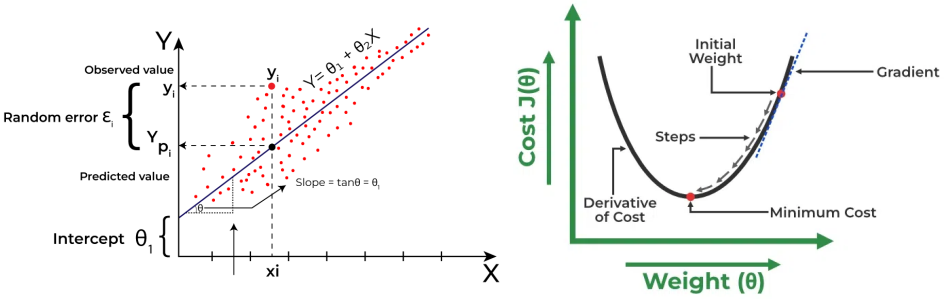


Figure 4.1: Left; correlation between X and Ys. Right; Cost or loss values as a functions of θ. The A's are represented as θ's. [108]

best-fit line, provided that the error between predicted and known Y_i is minimal. This precision is ensured through a loss function J(Y, $f_A(X)$), which assesses the divergence between the actual results and those predicted by the model. A frequently used loss or cost function for regression tasks is the mean squared error (MSE), defined as

$$L = \frac{1}{n}\sum_{i=1}^{n}(y_{pred} - \hat{y}_{true})^2 \tag{4.2}$$

The training process aims to minimize the average loss over the dataset, expressed as:

$$\mathcal{J}(A) = \frac{1}{N}\sum_{i=1}^{N} J(Y_i, f_A(X_i)). \tag{4.3}$$

One can implement a least squares fitting or train the model using an optimization algorithm such as gradient descents [109] to obtain the best fit. During gradient descent, the idea is to start with randomly initialized values for A's and iteratively modify these to reduce the cost function of the model till its minima are reached as

shown in Fig. 4.1 (left).
Linear regression is an easy-to-understand and implement algorithm and works well as long as a linear relation among the input and output variables. The coefficients indicate how the dependent variable changes per unit increase in the independent variable, offering relationship insights. Before we train any such models, it is a good idea to divide your data sets into, mutually exclusive, training and validation data sets. The idea of the model is to learn using the training data set and validate its performance on the validation data set that it did not see during the training.
In more detail of the descent of the gradient, for a cost function $\mathcal{J}(A)$, the gradient is defined as the derivative of the cost function with respect to these parameters. Based on this $\nabla_A \mathcal{J}(A)$, the parameters are updated for next iteration in the opposite direction of gradient, as;

$$A \leftarrow A - \eta \nabla_A \mathcal{J}(A), \tag{4.4}$$

where η is the learning rate, controlling step size to find a (local) minimum. When configured to extremes, either very high or very low, the algorithm may bypass the optimal solution entirely or require an excessive number of iterations to achieve convergence. So using a good learning rate is crucial as it can converge a model to garbage or a state-of-the-art, as shown in Fig 4.2, where epoch represents the iterative repetition of the whole process. One can opt for a constant learning rate or adaptive learning rate (changing it during the training), or cyclic learning rate where it fluctuates between the minimum and maximum set for learning rate as per the requirement of the problem at hand.

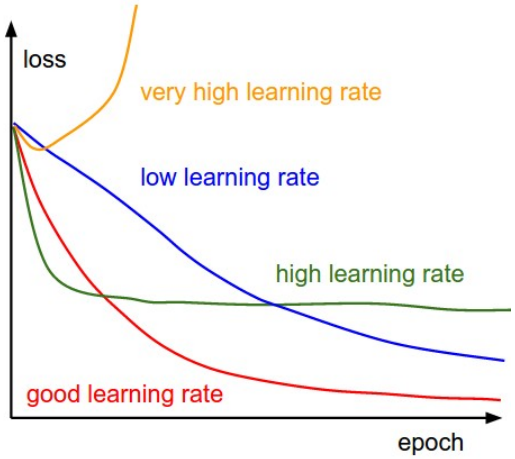


Figure 4.2: An illustration shows that low learning rates lead to gradual progress, while higher rates cause rapid improvements but can get stuck at non-optimal loss values (green line) [110].

The selection of a learning rate, the dataset utilized for gradient computation, and the timing of parameter updates are contingent upon the specific problem and the dataset size. Gradient descent algorithms are classified into three main types, distinguished primarily by the **batch size** [109]. This parameter determines the number of samples considered for the gradient calculation in each iteration.

- **Batch Gradient Descent :** The method determines the mean gradient of the cost function across all training samples. While it ensures reaching the global minimum, it can be resource-intensive and time-consuming when handling extensive datasets.

- **Stochastic Gradient Descent (SGD) :** This uses a single random training data point to calculate gradient descent and update the parameters in opposite directions. Its key advantage is the introduction of randomness, essentially helping it to escape local minima. Although it ensures computationally efficiency faster convergence, it may lead to large fluctuations which can be avoided by using learning rate schedulers, or mini batch SGD. The advanced alternatives like *Momentum*, *RMSprop*, and *Adam* are frequently utilized as well.
- **Mini-Batch Gradient Descent :** It computes the average gradient over a mini-batch, combining the advantages of both approaches to balance computational efficiency and update stability.

The mini-batch gradient descent method is frequently employed in neural network training. This technique, while ensuring convergence to a loss function minimum, does not necessarily reach the global minimum. That is where the Adam (adaptive moment) optimization algorithm [111] helps the ML model paving its way toward global minima.
Adam optimization begins with initiating two trackers: one for the exponential moving average of the gradient (first moment) and another for the squared values (second moment), focusing on the average variability of the gradient to identify changes in speed, both initialized at zero. In the descent stage of the gradient, the algorithm assesses the gradient by comparing the predicted and actual outcomes. Bias Correction is applied to the initial estimates to boost reliability. Lastly, the Parameter Adjustment stage uses these refined estimates to update the network parameters for improved performance. It modifies the classic weight update formula as follows :

$$A_{t+1} = A_t - \frac{\eta}{\sqrt{v_t} + \epsilon} \cdot m_t \tag{4.5}$$

where A_{t+1} is the subsequent weight, A_t denotes the current weight, η is the learning rate, m_t represents the first-order momentum, v_t is the second-order momentum, and ϵ is the constant for numerical stability. The first-order momentum is described by :

$$m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot \frac{\partial J}{\partial A_t} \tag{4.6}$$

Like a ball rolling over rough terrain, this momentum helps navigate local minima and bypass saddle points, improving convergence speed.

4.1.1 Neural networks: extending linear models

As the complexity of the problem at hand increases, nonlinear models become more efficient than simple linear regression models. Neural networks are one such kind of model designed to mimic the complex functionality of human brain via neurons, are capable of learning complex structure data and identify patterns. The basic unit is a neuron that receives inputs, governed by weight and activation functions. The connections between neurons are presented by weights (w) derived the strength between the same. The similarity between the neuron and a neural network is shown in Fig. 4.3.

A single-layer perceptron representing a single input feature in data, operates as follows:

$$\mathbf{h} = \sigma(\mathbf{w}\mathbf{x} + \mathbf{b}), \tag{4.7}$$

where $\sigma(x)$ is an activation function and can be defined as the Rectified Linear Unit (ReLU) $\max(0, x)$ function, introducing non-linearity into the model. The output

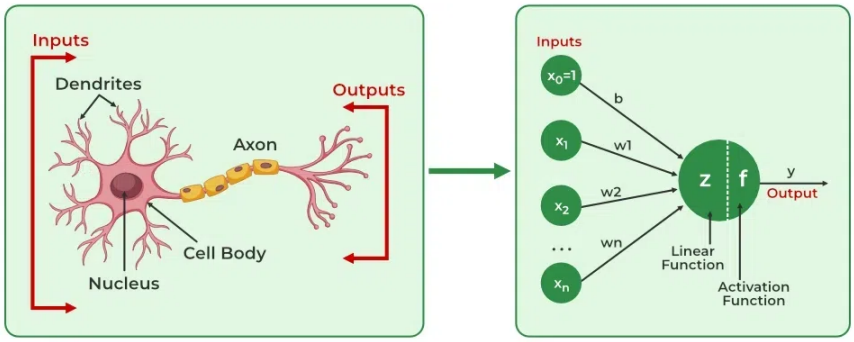


Figure 4.3: The figure illustrates the juxtaposition of a biological neuron with an artificial neuron, elucidating the manner in which inputs are acquired and processed to produce outputs in both systems [112].

layer undergoes a comparable transformation, providing predictions necessary for classification or regression analyses.
Training for neural networks begins by initializing the weights of each neuron with random values. Analogously to linear regression, the process of optimizing neural networks requires the minimization of a loss function, employing gradient descent methods. A neural network (NN) consists of an input layer which receives the input features, hidden layers equipped with information transfer function and activation function as mentioned earlier, and an output layer forms the output of the networks. Fig. 4.4 illustrates the general working of the neural network highlighting forward and backward propagation. In the next step, the input features are passed from the input layers, through hidden layers, to the output layers with a linear transformation of each neuron based on its connections to other neurons via a function z. This information is passed through an activation function f.
After forward propagation, the output of the NN is obtained, and the loss function is calculated. Based on this, gradient descent is calculated and weights are updated in the backward propagation. This whole process is repeated for many iterations (epochs) while minimizing the loss function. The number of hidden layers is one of the hyper-parameters and can be tuned as per the problem at hand.

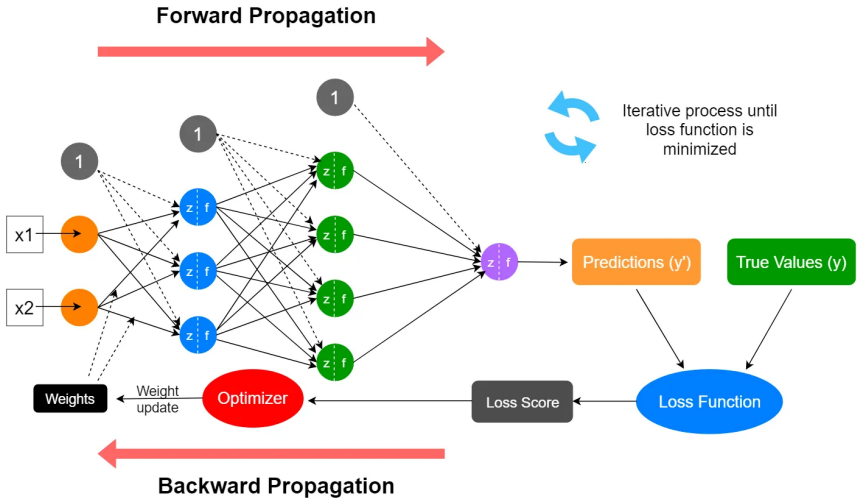
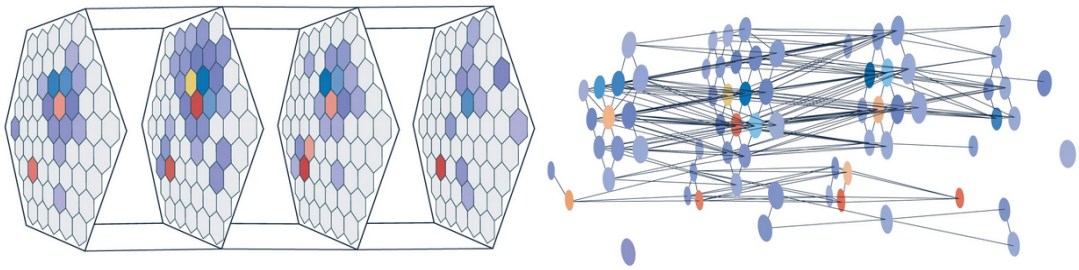


Figure 4.4: The learning mechanism of a neural network, emphasizing the forward and backward propagation processes [113].



(a) (b)

Figure 4.5: The depiction of edge representation and object condensation in multilayer calorimetry: (a) point clouds of signal generated in different cells as particles passes through the detector losing its energy via interaction with detector material. (b) Edges serve as representations for the associations between these energy deposits converting the point clouds to graph representation. [114]

To exploit neural networks for tasks dealing with low-level features, like those found in image data comprising two-dimensional arrays of pixel values across varied channels, Convolutional Neural Networks (CNNs) [105] are utilized. Within this operational framework, each image is processed via a convolution operation with a kernel matrix populated with trainable parameters. This approach facilitates the abstraction of information from individual pixels while encoding adjacent contextual details, enabling the network to discern spatial patterns. However, in our situation, the calorimeters feature different levels of granularity such as HGCAL compared to current endcap ECAL and HCAL of CMS, and the stochastic nature of particle showers along with their irregular margins result in sparse imagery. Since CNNs utilize zero-padding to standardize input grids, this approach enlarges the input size, rendering it inefficient for our specific application. The sparse and scattered nature of rechit sets, ranging from isolated hits to dense clusters, makes graph-based models a natural choice.

4.2 Graph Neural Networks (GNNs)

Graphs, represented as G(V,E), consist of vertices (V) and edges (E) that define relationships between nodes. Geometric deep learning (GDL) extends deep learning (DL) methodologies to non-Euclidean spaces, with graph neural networks (GNNs) as a pivotal framework for analyzing interconnected data. Given their versatility, GNNs are particularly suitable for studying complex systems, such as data from LHC [114]. Modern particle detectors such as HGCAL, with their high granularity and intricate interactions of hadrons, demand advanced data representations [115]. By treating detector hits as nodes and their relationships as edges, graphs encapsulate geometric and relational information, enabling GNNs to process these data effectively. Recent advancements in GNN methodologies, coupled with high-performance computational resources, have allowed the direct use of rechits as inputs for machine learning models. This approach bypasses traditional reliance on engineered high-level features, paving the way for more accurate and efficient particle reconstruction. Fig. 4.5 illustrates (a) point clouds created by signals in various detector cells as particles traverse the detector, dissipating energy through interactions with the material and (b) the edges represent the connections among these energy deposits, transforming the point clouds into a graph representation.

There exist various types of graph neural networks (GNNs), including graph convolutional networks (GCNs) [116], graph autoencoder networks (GAENs), recurring

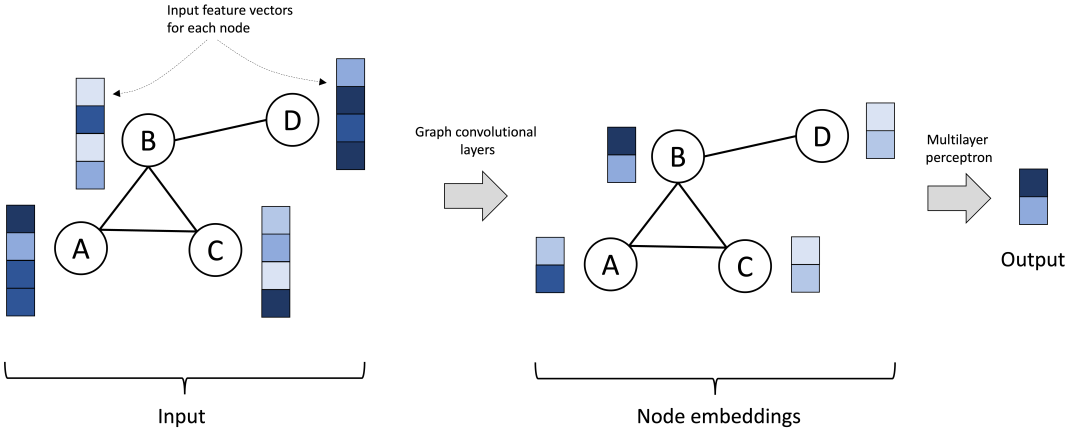


Figure 4.6: Diagrammatic representation of the fundamental architecture of Graph convolutional networks (GCNs), elucidating the operation of the graph convolutional (message-passing) layer. [117]

graph neural networks (RGNNs), and gated graph neural networks (GGNNs). The GCNs are one of the popular ones among these and are similar to Convolutional Neural Networks (CNNs) as they extract features from adjacent nodes in a graph.

Graph convolutions extend the concept of CNNs to non-Euclidean domains, enabling the processing of graph-structured data. A GCN takes graphs along with a set of feature vectors associated with each node. The construction of the graph is done by linking each node to its k-nearest neighbors [118] within a higher-dimensional feature space. Formally, for a given point cloud $X = \{x_1, x_2, \ldots, x_n\} \in \mathbb{R}^F$, a directed graph $G = (V, E)$ is calculated, where $V = \{1, \ldots, n\}$ are denoted as vertices and $E \subset V \times V$ as edges. In its most fundamental implementation, G is formulated as the undirected k-nearest neighbor (k-NN) graph of X within $\mathbb{R}^F$. This graph configuration includes self-loops, which ensures that each node maintains a connection with itself [115].
The input graphs are passed through multiple graph convolutional layers, which progressively modify the feature vectors for each node. The resulting graph comprises output vectors linked to each node, which may differ in dimension from the initial input vectors. This process is illustrated in Fig. 4.6 and is taken from [117]. Graph convolutions differ in their handling of node and edge data. Key types include spectral-based, spatial-based, attention-based convolutions, and message-passing neural networks (MPNNs). Edge convolutions, a Spatial-Based Convolution in GCNs, are designed for graph-structured data. They use node and edge features to dynamically represent relationships, adapt to irregular graph structures, and effectively integrate node and edge information.
The principal objective of EdgeConv is to consolidate information from adjacent nodes, channeling it into the central node, thereby fostering an understanding of the proximal context. Following the construction of the graph, EdgeConv is executed by defining edge characteristics $e_{ij} = h_\theta(x_i, x_j)$, employing h_θ as a nonlinear function characterized by learnable parameters θ. This function commonly utilizes a fully connected neural network (FCNN) for implementation [115] and can be mathematically expressed as:

$$\mathbf{h}_i^{(k)} = \sigma\left(\mathbf{W}^{(k)} \cdot \mathrm{AGG}\left(\{\mathbf{h}_j^{(k-1)} : i \in \mathcal{N}(i)\}\right)\right), \tag{4.8}$$

where $\mathbf{h}_i^{(k)}$ is the representation of node i at layer k, $\mathcal{N}(i)$ represents the neighbors of

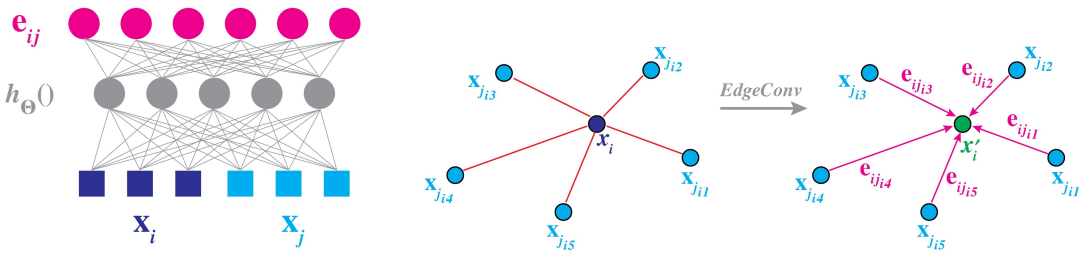


Figure 4.7: Left: The derivation of an edge feature (e_{ij}) (top), based on a pair of points, x_i and x_j (bottom), utilizing the function $h_\theta()$ defined in text. Right: The EdgeConv mechanism computes the result by aggregating the edge features of all edges emanating from each connected vertex. [115]

node i, $\mathbf{W}^{(k)}$ is the learnable weight matrix, and σ is a non-linear activation function. The AGG is an aggregation function such as mean or maximum grouping. This structure enables GNNs to model interactions and correlations between particles.
A symmetric aggregation operation, such as summation or maximization, is also applied on the edge features correlated with all edges that converge at each vertex:

$$x_i' = \bigoplus_{j:(i,j)\in E} h_\theta(x_i, x_j) \tag{4.9}$$

Through this process, the central node x_i is enriched with data from its k-nearest neighbors via a procedure termed message passing. These fully connected layers, referred to as the message passing layers. The edge features for a singular message passing layer are delineated as:

$$e'_{ijm} = \mathrm{ReLU}(\theta_m \cdot (x_j - x_i) + \phi_m \cdot x_i) \tag{4.10}$$

wherein ReLU functions as the rectified linear unit activation, with θ_m and ϕ_m representing trainable parameters.
As more message passing layers are incorporated, the complexity of edge features increases. Ultimately, a node's updated attributes are derived by aggregating the edge features of its proximal nodes:

$$x'_{im} = \max_{j:(i,j)\in E} e'_{ijm} \tag{4.11}$$

The resulting output is a M-dimensional point cloud, preserving the initial graph's node count. The whole process is shown in Fig. 4.7 taken from [115].
For node-level tasks, the model outputs vectors representing the probability that each node belongs to different classes. The output of convolutional layers can be used for this task. In contrast, for graph-level tasks, the model focuses on the entire graph by utilizing a neural network to process the collective per-node vectors and produce a single output vector, which can be used, for example, for graph classification or regressions.

Utilizing edge convolutions, a dynamic graph CNN (DGCNN) is constructed in which the input graph is not static but is reconstructed at each layer considering the edge features. The dynamically evolving graph G(l) = (V(l), E(l)) is delineated at the l^{th} layer. The architecture of DGCNNs is adept at reconstructing the graph at each layer, thus achieving permutation and translation invariance. More discussion on

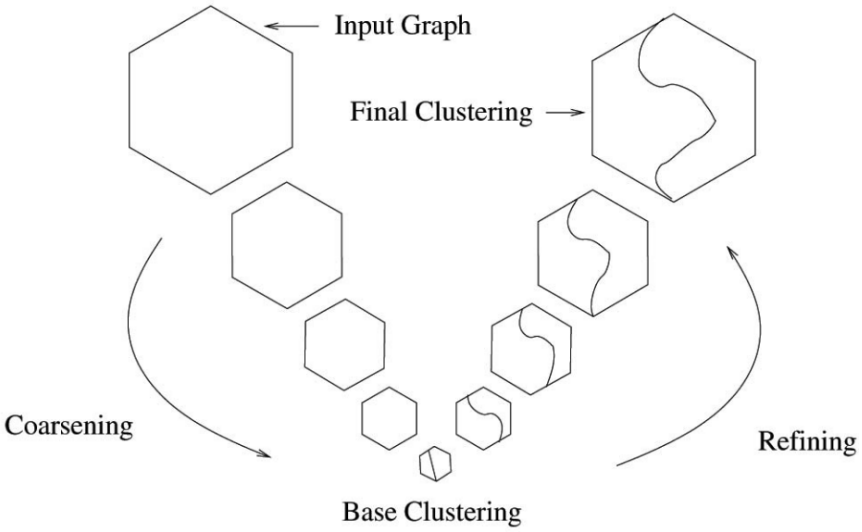


Figure 4.8: Illustration of multi-level algorithm used in graclus clustering highlighting the coarsening and refining of graph [120]. The process commences by aggregating points based on the maximum number of adjacent nodes from a top-down perspective. Conversely, during the bottom-up phase, points are segmented back, with cluster assignments determined by the "popularity" or the quantity of neighboring nodes.

this topic is available in [115].

4.2.1 Graph Clustering

Node clustering in GNNs is vital to uncover structural and connectivity patterns of data, improve information aggregation, feature transmission, and network learning. Graph clustering can be viewed as an indexing of nodes. Clustering also helps reduce computational complexity and highlight important relationships within the graph. Traditional methods K-Means clustering and Spectral clustering use graph Laplacian eigenvalues for node embedding and partitioning. Algorithms DBSCAN and HDBSCAN find dense regions and address non-linearity, while Louvain and Graph Cuts focus on modularity and reducing edge cuts for community detection. Advanced techniques such as, Hierarchical clustering and Mean Shift clustering offer multi-resolution views and detect complex clusters, improving GNN efficiency and scalability.

Within the spectrum of clustering algorithms, the method described in **Graclus clustering** is widely used to coarsen the graphs. This technique utilizes a greedy strategy to minimize normalized cut, thereby reducing the graph size while maintaining its structural attributes. Furthermore, Graclus is an efficient clustering tool utilized for partitioning of unlabeled data through graphical representations. It facilitates the conversion of data into either a simple or a weighted graph format. In simple graphs, edges possess unitary weight, whereas in weighted graphs, edge weights vary and are determined via approaches such as cooccurrence statistics. Notably, this method obviates the need for eigenvector computation, rendering it particularly suitable for application to large graphs. The overview structure of the multilevel algorithm implemented in graclus is shown in Fig. 4.8 taken from [119].

The multilevel algorithm [119] involves three main stages. At first, the graph is coarsened to create a smaller representation iteratively. The nodes in each level are combined to create supernodes in the next level, with the edge weights being the sum of the original edge weights. Then, this base graph is clustered by selecting random vertices and expanding regions around them to form clusters. During this process, a

random node is selected and subsequently connected to either a random neighbor or the neighbor that exhibits the greatest edge weight to form a cluster. In instances where a node lacks a neighbor that is not yet clustered, it forms a cluster comprising only itself. Consequently, the resultant number of nodes is generally reduced by approximately fifty percent, although this reduction may be less pronounced when a node must form a cluster independently. And finally, this clustering is refined as the graph is progressively uncoarsened. By connecting these objectives with the weighted kernel k-means formulation, the algorithm applies an iterative k-means-like approach to achieve optimization. This technique underpins the Graclus clustering tool [120], which efficiently computes objectives like normalized cuts and ratio associations without relying on eigenvector calculations, making it well suited for large-scale graph clustering tasks.

4.2.2 Pooling Operations in GNNs

Pooling operations in GNNs aim to reduce graph size while preserving essential information, akin to downsampling in CNNs. This is particularly important in particle physics applications, where the detector data can be extremely large and sparse. Global pooling aggregates node features across the entire graph using operations such as mean, maximum, or sum. Hierarchical pooling iteratively coarsens the graph by merging nodes or edges, with methods such as Graclus pooling, which uses graph clustering algorithms, and DiffPool, which learns a soft assignment of nodes to clusters, enabling differentiable pooling. Another approach is top-k grouping, which selects the top k nodes based on their scores derived from the learning parameters. These pooling techniques facilitate hierarchical representation learning and improve computational efficiency.

EdgePool [121], is a novel graph pooling layer for GNNs that enhances hierarchical representation learning by contracting edges to merge nodes. This method involves selecting edges based on learned scores, merging the connected nodes, and updating the graph structure accordingly. By focusing on edge contractions, EdgePool effectively captures local graph structures and maintains important connectivity patterns, leading to improved performance in both node and graph classification tasks. Traditional pooling methods often face challenges such as fixed pooling sizes and the potential loss of important structural information. EdgePool addresses these issues by dynamically learning which edges to contract, allowing for adaptive pooling that preserves crucial graph topology. This approach not only enhances the flexibility of GNN architectures, but also contributes to more accurate and efficient graph analysis. An exemplary illustration is shown in Fig. 4.9 that highlights the grouping of nodes and edges in clusters.

For a complete understanding of EdgePool and its applications, please refer to the full paper [121]

4.3 Dynamic reduction network

The HGCAL provides an intricate portrayal of the longitudinal and lateral progression of the shower due to its high resolution granularity in both dimensions. Employing machine learning techniques enables the systematic exploitation of these four-dimensional data to deduce the energy and additional characteristics of the incoming hadron. These showers cause substantial variation in the spatial distribution of energy deposits within the detector cells event-by-event and are sparsely arranged, with channels often lacking measurable signals, as discussed in Section 2.7.2. A more fitting method to depict these sparse, high-dimensional, and variable-geometry

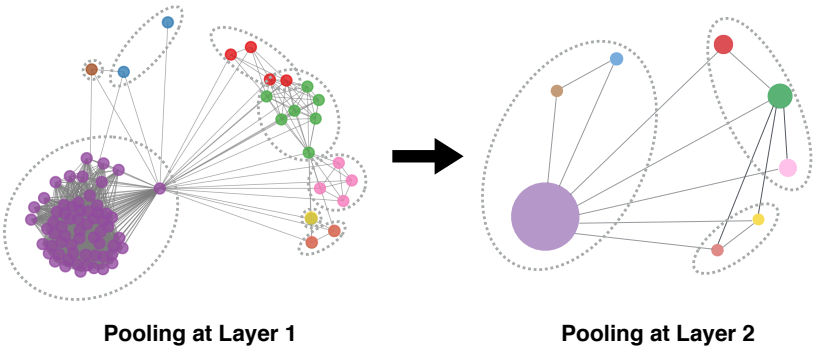


Figure 4.9: Two-layer hierarchical clustering is shown. Second-layer nodes represented by color denote first-layer clusters represented by dotted lines. [122]

input features is through dynamic graphs. Based on DGCNN [115], a novel architecture named dynamic reduction network (DRN) [123] is defined with two specific differences; graph construction and convolutions are performed within a higher-dimensional latent space, and a graph clustering and pooling step is introduced to iteratively consolidate information across the graph. The core idea behind DRN is to dynamically learn and adapt the structure of the input data by employing a series of graph-based operations that refine both local and global features. Unlike traditional approaches that rely on predefined graph structures, DRN uses dynamic graph construction, which allows it to identify and model important relationships within the data on the fly.

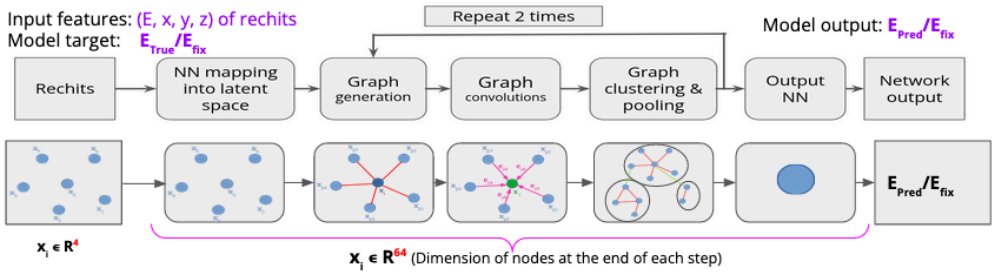


Figure 4.10: This diagram shows a GNN-based dynamic reduction network. Rechits are labeled as x_i, with red lines indicating their latent space connections, refined by graph convolution in magenta. In the graph clustering stage, clusters are small circles with blue dot nodes inside. Final predictions are made through the output neural network.

The collection of hits recorded in the detector corresponding to each pion object is considered as a point cloud of hits. The point clouds of hits along with respective input features are mapped to a high-dimensional latent space. The constructing graphs by connecting proximate hits is done in this high-dimensional latent space. Fig. 4.10 showcases the network architecture utilized in this study. The conversion of spatial coordinates and energy data from detector readout channels into the required output is described in the following steps.

1. The attributes of the detector channels, identified as input features, are normalized against predefined reference values to confine most of their range between zero and one, allowing for the presence of outliers. These normalized features are then transformed into a 64-dimensional latent space using a fully connected

neural network (FCNN). This network, implemented as a multilayer perceptron (MLP) encoder, comprises three layers interconnected by exponential linear unit (ELU) activation functions. The number of hidden layers in FCNN is one of the hyper parameters to be optimized (see Section 5.3 of Chapter 5. Transferring the input features to high-dimensional latent space via FCNN further enhances DRN's feature extraction capabilities and enables learning hierarchical features that are crucial for tasks like regression.

2. Advanced feature extraction is performed through GCNs by iteratively executing the following three steps:

   (a) **Graph construction:** Within the 64-dimensional latent space, each point in the point cloud is connected to its 16 nearest neighbors (including itself) to form graphs using k-nearest algorithm. The number of dimensions in latent space and selection of the number of adjacent nodes to be interconnected are hyperparameter, which may be adjusted based on the complexity of the problem.

   (b) **Graph convolution:** The constructed graphs are processed using the "EdgeConv" operation [124]. This involves multilayer fully connected networks that compute "messages" along the graph edges, which are then aggregated at each node to update its feature representation. The number of such layers is another important hyperparameter, which may be adjusted based on the complexity of the problem.

   (c) **Graph clustering and pooling:** After graph convolution, the *graclus* clustering algorithm [120] is applied to cluster graphs, followed by grouping, which is a key component of this model. Details of the clustering algorithm and pooling are discussed in Sections 4.2.1 and 4.2.2. During this phase, the updated node features are weighted by distance, and clusters are merged using a max-pooling strategy to form a unified representation. The Graclus clustering algorithm incrementally merges connected nodes within a graph cluster in pairs, optimizing for the maximum edge weight in the newly formed node.

   Each iteration of these three steps, graph construction, convolution, and clustering with grouping, reduces the number of nodes in the graph by half. By iteratively refining the representation of data through these pooling layers, the network progressively distills the most critical features, enabling efficient downstream processing. The use of dynamic pooling not only preserves the structural integrity, but also enhances the network's ability to generalize across different datasets. The edge pooling dynamically reduces the input data size while retaining essential information, hence the model is termed a dynamic reduction network (DRN).

3. The final set of reduced nodes is subjected to global maximum pooling and processed through a fully connected neural network with ELU activation at the encoder level. This process produces the final output, which in this context corresponds to the reconstructed energy or the classification score. The number of such layers is another important hyperparameter, which may be adjusted based on the complexity of the problem.

In the context of the charged pion reconstruction in HGCAL, each event contains several rechits, each described by four features: energy (E) and spatial coordinates (x,

y, z). The first layer of the model, a FCNN, processes each rechit independently and transforms its 4 input features into 64 outputs- a learned combination of the inputs. This creates a 64-dimensional latent representation for each rechit. In the next step, each rechit is treated as a node in a graph, and edges are formed between nearby rechits—based on the geometric distance calculated using standard euclidean distance in 64-dimensional latent space. In the message passing layer, each node updates its features based on its neighbors and connecting edge's features that are calculated using edge convolutional operations, allowing the model to learn information about the neighboring node's features. Once obtained, we update the features of i^{th} node by taking the maximum of the corresponding elements across all features from all the edges connecting to that node.

A clustering algorithm then groups nodes similar in their features, followed by pooling to reduce the number of nodes by aggregating features within each cluster. Finally, a global max pooling extracts the most significant 64 features across all clusters. These are passed through another FCNN to predict a single value that is the energy of the incoming particle.

The combination of dynamic graph construction, adaptive pooling, and FCNN-based feature extraction makes DRN a powerful tool for analyzing complex data structures. This architecture highlights the potential of integrating dynamic graph-based methods with deep learning frameworks to tackle challenging problems in diverse fields. These layers seamlessly integrate with dynamic pooling operations, ensuring that the extracted features are both spatially coherent and semantically rich.

4.4 Summary

This chapter provides a succinct overview of various machine learning models and their applications in experimental particle physics. It emphasizes the role of graph clustering techniques in GNNs in identifying structural patterns, reducing computational complexity, and enhancing the feature learning. Key methods like K-Means, spectral clustering, and Graclus are noted for their efficiency, with Graclus being particularly effective for large-scale tasks. The strengths and weaknesses of different neural networks are highlighted, with linear regression being interpretable but prone to overfitting, and neural networks being adept at modeling complex relationships but requiring large datasets and posing interpretability challenges. GNNs and DGCNNs provide a strong handling for irregular data, but are limited by computational demands and sensitivity to graph structures. The chapter also discusses methods like DRN, which are suitable for specific tasks such as energy reconstruction of hadronic showers.

Chapter 5

Energy reconstruction of charged pions in HGCAL using DRN

In collider physics, $p_{\mathrm{T}}^{\mathrm{miss}}$ is one of the commonly used search variables, especially for new physics searches, and it is highly dependent on the uncertainties on the visible energy measurements, essentially derived by uncertainties in jet energy measurements. To pave the way for precision measurements, an improvement in the single-particle resolution is required.

As discussed earlier (Chapter 2), when hadrons interact with a calorimeter detector, several nuclear processes occur as they dissipate their energy, leading to significant event-by-event fluctuations. These fluctuations manifest themselves in the types and number of shower particles, the spatial distribution of energy deposits, and the portion of energy lost to nuclear binding. Such variations depend on the energy of the incident hadron. The energy reconstruction of these showers is further complicated by energy leakage due to incomplete detector coverage or losses from missing channels. The complexities of hadronic showers result in an inferior jet energy resolution of calorimeters.

Due to the complexity of reconstructing hadronic shower energy, traditional methods like the weighted summation (WS) algorithm tend to be intricate, rely on arbitrary parametrization, and do not completely address the aforementioned issues.
With its high granularity, HGCAL provides a detailed 5-dimensional view of hadron showers, enabling the use of advanced machine learning techniques to improve the accuracy of energy and hadron reconstruction. During a beam test at CERN in 2018 (as described in Section 3.4.3), the prototype had around 34,000 independent readout channels, offering a highly detailed view of hadronic showers. The data collected from this prototype has been carefully analyzed and understood, providing a strong foundation for the development and validation of these advanced ML algorithms.

The graph neural networks (GNNs) are particularly well suited for this task in HGCAL due to the substantial variation in the spatial distribution of the few hundred active cells among the tens of thousands available, which changes significantly from one event to another. In addition, GNNs are more suitable for handling longitudinal granularity than CNNs.

This chapter presents an algorithm based on a GNN-based dynamic reduction network (DRN) designed to reconstruct the energy of charged pion showers in non-compensating calorimeters. This method achieves significant improvements in energy resolution, surpassing the weighted sum (WS) approach. These results have been

validated using detailed GEANT4 simulations and supported by experimental beam test data from 2018. The chapter also evaluates the algorithm's performance in a calorimeter with a geometry similar to the upcoming HGCAL detector. It highlights the advancements made by the DRN compared to traditional methods, showing its adaptability and reliability in practical collider environments.
A comprehensive overview of these findings has been documented in a CMS internal note, which is accessible at [125]. An essential summary of these studies has been published in JINST [126].

5.1 Event sample specifications and cleansing procedures

A sample of hadron showers produced by negatively charged pions, uniformly sampled in the energy range from 10 to 350 GeV, is generated using a GEANT4 based simulation of the beam test setup. A total of 5 M events are generated. The simulated set-up and physics list used are identical to that described in Section 3.4.3 and used in the studies presented in [98], with the exception that the pions are aimed down from a distance of 33 m from the beam test set-up, rather than the 600 m distance used in the full beam line simulation. The event selection criteria to reject the events where the particles lose a large fraction of their energy in the beam line or rejection of the recorded hits due to spurious noise are identical to those used in the data [98] and are summarized here briefly:

- Channel mask : Information from the faulty channels with spurious noise has been masked for further analysis
- Noise rejection : Channels with signal less than the noise thresholds of ~ 0.5 MIP are rejected.
- Track quality cut : Events with good quality tracks having recorded hits in at least 3 out of 4 wire chambers and $\chi^2/\texttt{n.d.f} < 10$ to avoid passing of multiple particles produced in the interactions of primary beam particles with beam line elements.
- Track window cut : Using the extrapolated track from DWC to the first CE-E layer of the HGCAL prototype, a spatial cut of a 4×4 cm^2 window is implemented for lower beam energies, while a 2×2 cm^2 window is employed for higher beam energies to exclude pions that are out of acceptance and ensure similar beam profiles in data and simulation, avoiding discrepancies due to particles arriving at large angles.
- Muon veto : This criterion is devised to mitigate the potential contamination by muons generated at the target or through in-flight decay of pions. Muons with energies of a few hundred GeV behave as MIPs, depositing only a small amount of energy and exhibiting a narrower lateral spread compared to showering particles such as pions. Hence, events are discarded if the total energy within CE-E reaches $<$100, the total energy within CE-H is $<$60, and the lateral spread of the shower, defined as the ratio of the maximum energy deposited in a cell relative to its adjacent 5×5 cells, is $>$0.8 [102].
- Preshower rejection : The beam-line elements in front of the HGCAL beam test setup contribute to an effective material between secondary beam production and the experimental setup. There is a probability of pions initiating their shower before reaching the calorimeter, leading to incomplete containment and poor energy resolution measurements for the pions. A shower startfinder algorithm [98] is used to reject events in which showers begin within the first two layers of the CE-E, due to information loss from a masked chip in the first layer. Consequently, this cut has maximum influence.

The information needed to implement track window cut was not saved in the simulated sample with flat energy; hence the cut in the track window cut has been applied to the data but not simulated events. Its impact is expected to be negligible in the simulated dataset. Effect of sequentially applying these event level rejections on simulated events originally uniformly distributed in incident energy is shown in Fig. 5.1. We are left with approximately 4.1 M events after applying the above selection criteria. The events remaining after applying the described selection criteria are employed in the analyses presented within this chapter, encompassing the demonstration of classical methods and the training of the DRNs, except where otherwise specified.

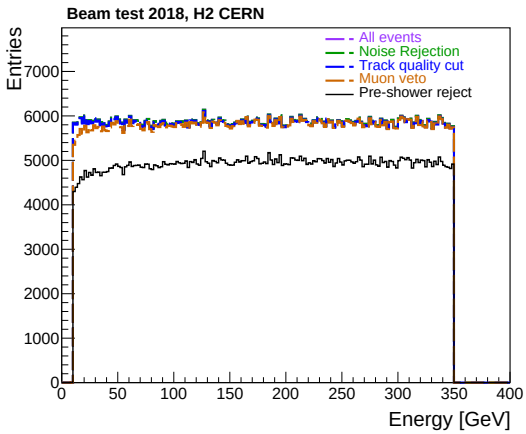


Figure 5.1: Energy distribution of simulated pion sample after sequentially applying event rejection criteria described in the text.

5.2 Energy reconstruction of pions using classical methods

For the studies presented here, we have considered information only from every 4^{th} layer of AHCAL, effectively reducing the sampling layers of AHCAL from 39 to 10, to approximate the number of sampling layers in the scintillator section of the CMS-HGCAL geometry. The signal in each cell or channel of the detector, together with their positions, is referred to as *reconstructed hits* or *rechits*. The beam test set-up is equipped with approximately 34 k number of read-out channels providing the granular information of showers generated by particles traversing the detector. The distributions of the number of rechits with a signal measured above the noise thresholds in the CE-E, CE-H and AHCAL sections, as well as the total number of rechits across all three sections (denoted as "All"), are shown in Fig. 5.2 for simulated 100 and 200 GeV π^-.

The Fig. 5.3 displays the x and y coordinates (top row), x versus depth z (bottom row) of each rechits for 100 and 200 GeV π^- highlighting the transverse and longitudinal attributes of beam test geometry. The plots x versus depth z is reflection of the fact that the last three layers of CE-H are equipped with a single module and the AHCAL is down-scaled with only 10 layers for this study. Cell-to-cell variation in response is made uniform by using 200 GeV muons for ADC-to-MIP conversion factors, as summarized in Section 3.4.3.The energy deposited in each detector cell is then presented as equivalent to the number of MIPs passing through it. Given that this is a sampling calorimeter, the signal deposited in the active layer reflects the number of shower particles at the respective depth and hence the energy deposited in the absorbers. Two traditional methodologies for reconstructing incident pion energy have been developed, accurately determining the energy deposition in the respective detector segments utilizing varying absorbers and detector technologies,

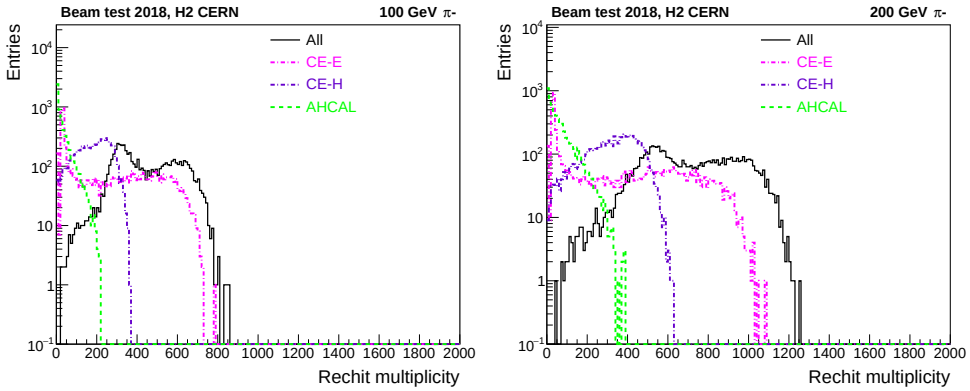


Figure 5.2: Distribution of number of rechits across different detector sections of beam test setup, as well as the total number of rechits across all three sections (denoted as "All") or 100 and 200 GeV incident pion energy.

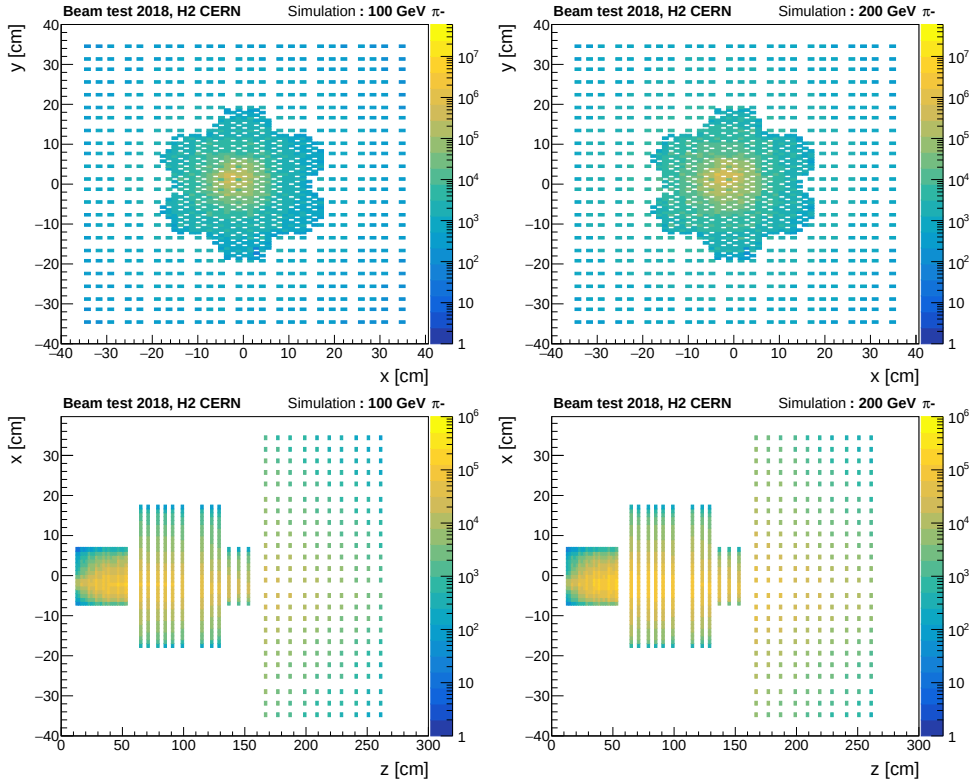


Figure 5.3: Distribution of x vs y coordinates of rechits (top row) and x vs z coordinates of rechits of 100 GeV (left) and 200 GeV (right) pions. Here, Z-axis represents the frequency of events.

for the AHCAL featuring 39 sampling layers, as detailed in [98] and extensively examined in Thesis [102]. These methodologies have been recalibrated for a prototype featuring only 10 AHCAL layers for the purposes of these studies and are concisely discussed in the subsequent sections.

- Using fixed weights 5.2.1
- Using energy dependent weights 5.2.2

5.2.1 Energy reconstruction of charged pions using a fixed weight method

The HGCAL being a sampling calorimeter, the signal deposited in the active layer reflects the number of shower particles at the respective depth, hence the energy deposited in the absorbers. We can choose an energy scale, that is, MIP-to-GeV conversion factor or *weight* such that the sum of all rechit energies equals a given incident energy, say 50 GeV, on average as presented in [98]. The same conversion can be used for all other energies. However, the energy scales of different detector sections can vary depending on non-identical on the absorber material and sampling fractions.

Total pion energy can be reconstructed using fixed weight, E_{fix}, as:

$$E_{fix}\,(GeV) = \alpha^{fix} \times E_{MIPs}^{CE-E} + \beta^{fix} \times (E_{MIPs}^{CE-H} + \delta^{fix} \times E_{MIPs}^{AHCAL}) \tag{5.1}$$

where α^{fix}, β^{fix} are MIP-to-GeV conversion factors for the electromagnetic and hadronic sections, respectively and δ^{fix} is a relative factor between CE-H and AHCAL.

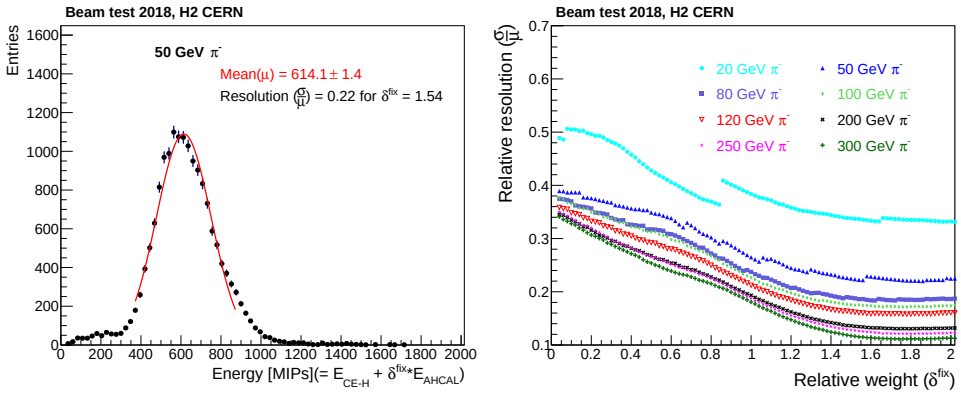


Figure 5.4: The energy resolution obtained for different pion energies after fitting the energy ($E^{CE-H} + \delta^{fix} \times E^{AHCAL}$) distribution with varying the δ^{fix} with a step of 0.02 is demonstrated (left). On right, the energy distribution obtained with that value of δ^{fix} which minimizes the resolution.

Since CE-E is optimized for measuring electromagnetic showers, 50 GeV e^+ data is used to obtain α^{fix} and it is found to be approximately 10.5 MeV per MIP [52]. We used 50 GeV pi^- events that do not initiate a hadron shower in CE-E (i.e., MIPs in CE-E or referred to as CE-H pions) to obtain β^{fix} after fixing δ^{fix}, the relative weight factor to account for differences in sampling fractions and transverse dimensions of CE-H and AHCAL although the absorbers are the same. For each value of δ^{fix} scanned from 0 to 2 with a step size of 0.02, the energy distribution of combined CE-H and AHCAL energies in units of MIPs after weighting down the AHCAL by δ^{fix} is fitted with a Gaussian in the core, as shown in Fig. 5.4 (left). The μ and σ of the fitted Gaussian function are used to define the response and resolution as μ/E_{true} and σ/μ, respectively.

Based on Fig. 5.4 (right) showing resolution as a function of δ^{fix}, we obtain δ^{fix}=1.5 as the value minimizing energy resolution for different pion energies. Using this value, we obtain 80 MeV per MIP for for $E_{MIPs}^{CE-H} + \delta^{fix} \times E_{MIPs}^{AHCAL}$ for pions not showering in CE-E.

The energy measured in each cell in units of GeV using fixed weights across different sections of beam test setup, the energy versus the depth z are shown in the top and

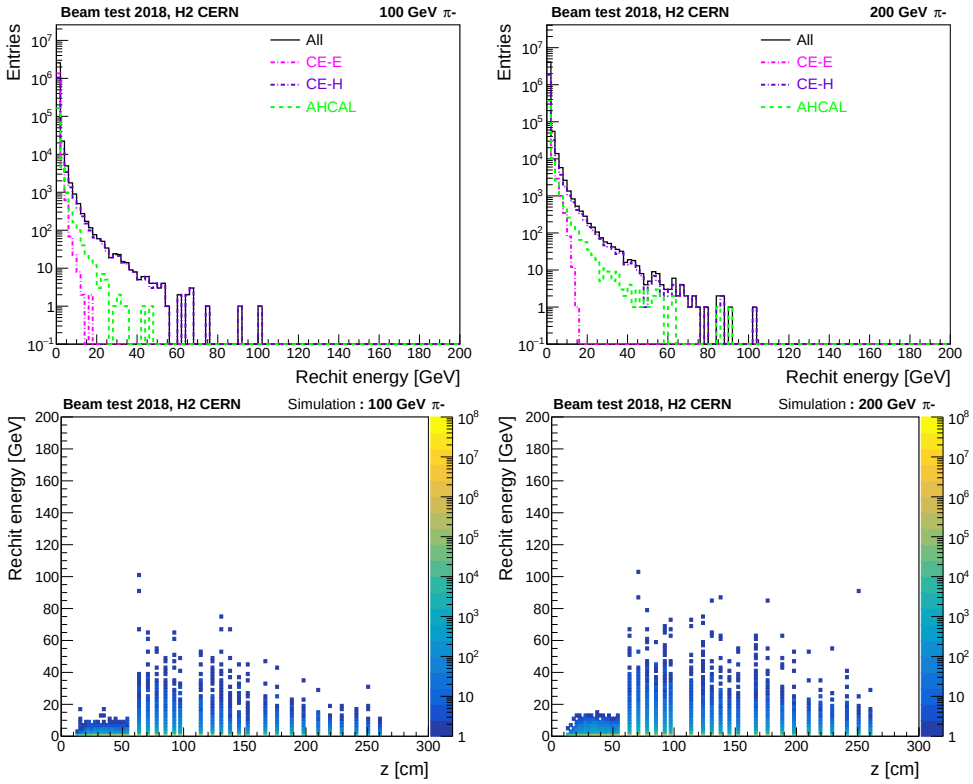


Figure 5.5: The distribution of rechit energies (top) and its distribution as a function of depth z (bottom) for 100(left) and 200(right) GeV pions. The rechit energy is represented in units of GeV which is calibrated using fixed weights.

bottom row of Fig. 5.5, respectively for 100 and 200 GeV π^-. The rechit energy presented in this unit is used as input to train the DRN discussed in Section 5.3.1 is the unit for the rechit energy used as input during its training.

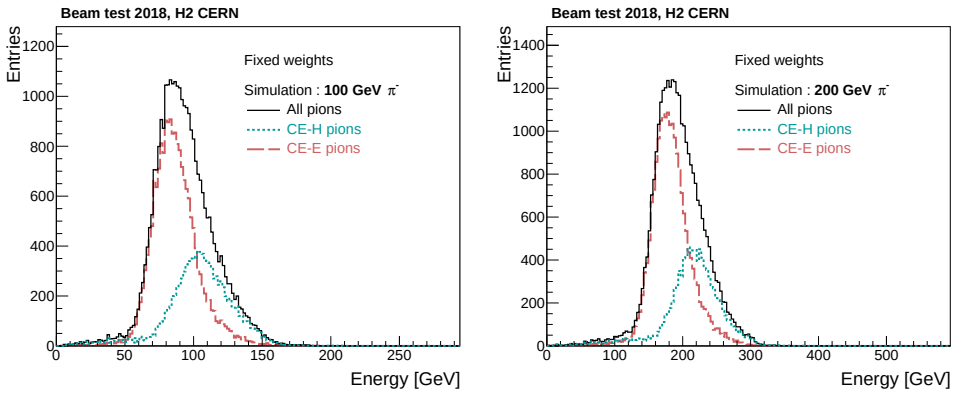


Figure 5.6: Distributions of reconstructed energies for simulated 100 GeV (left) and 200 GeV (right) pions in different event categories including "All events", "CE-E pions" and "CE-H pions" using fixed weights.

The total energy of the pions that start showering in CE-E (*CE-E pions*) or CE-H (*CE-H pions*) and "All events" are shown in Fig. 5.6 for 100 (left) and 200 (right) GeV pions. Details on these pion shower location-based event categorizations are added in Section 5.6.1.

These energy distributions are fitted with a Gaussian in the core and μ and σ parameters are extracted. We define energy response and relative energy resolution as μ/E_{true} and σ/μ, respectively. This is the procedure used to obtain the resolution and response from one-dimensional distributions throughout the studies presented in this chapter. As expected, the response and resolution of the reconstructed energy are very different for the two populations when using fixed weights (see top and bottom right of Fig. 5.8).

The HGCAL is a non-compensating sampling calorimeter. Its response to positrons is higher than that for pions and because the total electromagnetic component of hadronic showers increases with incident hadron energy. Consequently a non-linearity of calorimeter response is expected as shown in black points of Fig. 5.8 (right). This shows that the E_{fix} is a sub-optimal estimate of the particle's true energy and conventional energy dependent weighted summation method is presented in the following section.

5.2.2 Energy reconstruction of charged pions using weighted summation (WS) method

A conventional weighted summation method is employed, aimed to achieve linearity in response by deriving incident energy dependent weights to combined energy measured in the three sections of the HGCAL. We implement a χ^2 optimization method similar to Ref. [98] and optimize it for downscaled AHCAL with 10 layers. The energy reconstructed using the minimization χ^2 after combining the energies measured in the CE-E, CE-H, and AHCAL sections using energy-dependent weights is shown in equation 5.2.

$$E_{corr} = \alpha(E_{true})E_{fix}^{CE-E} + \beta(E_{true})E_{fix}^{CE-H} + \gamma(E_{true})E_{fix}^{AHCAL} \tag{5.2}$$

where α, β, γ are energy dependent weights and E_{fix}^{CE-E}, E_{fix}^{CE-H} and E_{fix}^{AHCAL} are total reconstructed energies using fixed weights in CE-E, CE-H and AHCAL sections. The weights α, β, γ are determined after minimizing a χ^2 constructing for i^{th} energy point as shown in the equation 5.3.

$$\chi^2 = \sum_i \frac{(E_{true} - E_{corr}^i)^2}{\sigma^2(E_{fix}^i)} \tag{5.3}$$

where $\sigma^2(E_{fix}^i)$ in the denominator on right hand side is a formative evaluation of the uncertainty in E_{fix}^i. Detailed explanation of different aspects of reconstructing energy of pions using χ^2 optimization are discussed in Thesis [102]. The distributions for α, β, γ obtained after minimizing a χ^2 for CE-E and CE-H pions are shown as a function of energy and are shown in the Fig. A.1 and Fig. A.2 of Appendix A.1, respectively.

Distributions of energy reconstructed using the WS method for 100 (left) and 200 (right) GeV simulated pions for CE-E, CE-H and inclusive pion samples are shown in Fig. 5.7. The WS method responds to events where the pion either starts showering in CE-E or CE-H similarly.

The energy spectrum is partitioned into 85 equal 4 GeV bins to predict energy distributions. These distributions are fitted with a Gaussian to extract the parameters μ and σ. The energy response and the relative energy resolution are defined as μ/E_{true} and σ/μ. Fig. 5.8 shows the comparisons of energy resolution and response

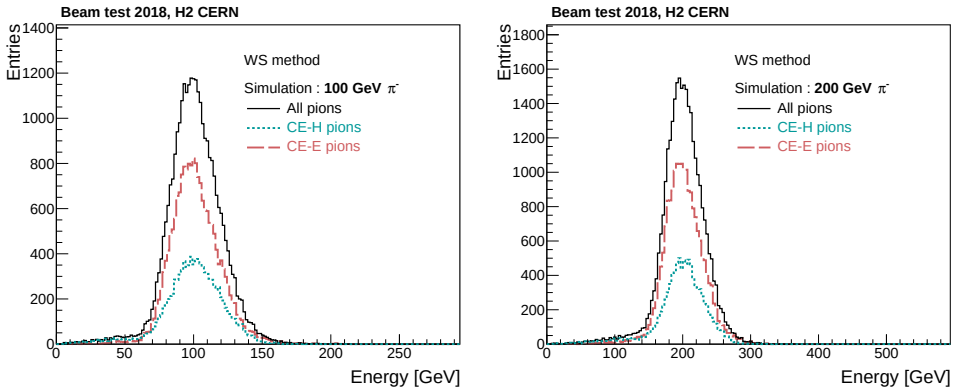


Figure 5.7: Reconstructed energy distributions for different event categories including "All events", "CE-E pions" and "CE-H pions" using WS method for simulated pions of energies 100 (left) and 200 GeV (right).

obtained using fixed weights and the WS method for CE-E and CE-H pions for these 85 bins. The energy dependent weights restore the linearity of calorimeter response to hadrons, but the energy resolution obtained does not improve much. That can be attributed to the fact that this method does not use the detailed shower information available with the unprecedented transverse and longitudinal granularity of the HGCAL and optimized to restore average response to unity, hence not correcting for event-to-event fluctuations.

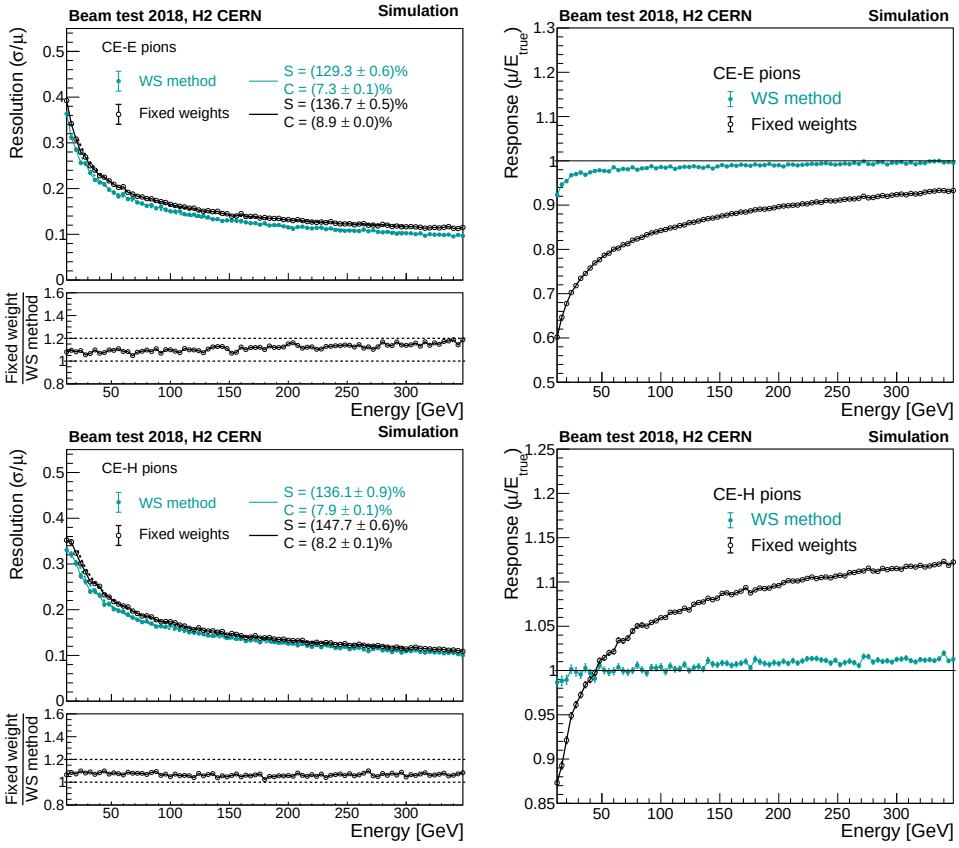


Figure 5.8: The energy resolution (left) and response (right) for CE-E pions (top) and CE-H pions (bottom) using fixed weights and WS method as functions of pion energy.

5.3 Dynamic reduction network for charged pion energy reconstruction

The unprecedented high longitudinal and transverse granularity of the HGCAL makes it possible to record details of development of hadron showers in space and time and learn from shower-to-shower differences (fluctuations). The availability of such granular information makes the use of a DRN to mitigate the impact of these fluctuations feasible. Following attributes of rechits passing the noise thresholds belonging to pion showers simulated in the beam test detector setup are provided as inputs to the DRN:

- Energy of each rechit either in units of MIPs or GeV (obtained using fixed weights 5.2.1)
- Position of each sensor cell : x, y coordinates
- Depth at which the cell is positioned : z coordinates

This set of information either can be provided all at once for DRN to learn or can be used in different combination to understand the impact of various information on DRN's performance as you will see later in this section.

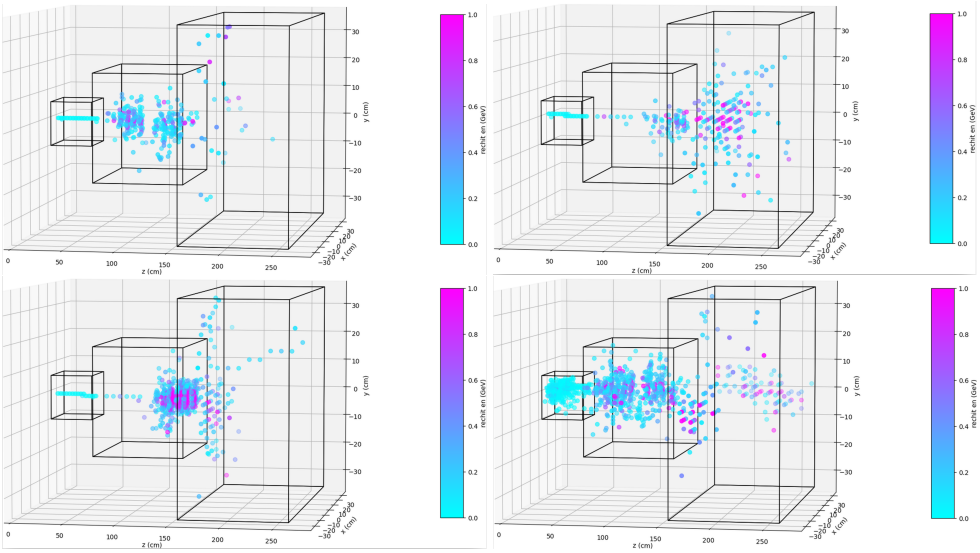


Figure 5.9: Displays of showers within the beam test geometry for 100 GeV (top) and 300 GeV (bottom) pions.

Figs. 5.9 show the 3D projection of rechits for individual showers due to pions of 100 (top) and 300 (bottom) GeV energies, respectively. Such event-by-event point clouds are provided as input to the DRN allowing it to learn complex energy and position correlations inherent to hadronic showers.

In the context of supervised learning, a target needs to be provided to the model being trained so that it learns from input features and is able to predict that target. In a regression problem such as the reconstruction of the pion shower energy, we can use E_{true} as the target or a response like $\frac{E_{true}}{E_{fix}}$. A loss function is used to ensure that the model also learns the accuracy of its prediction. It has either a very large value if predictions are way off or a small value in case predictions are close to the target. The purpose of loss functions can be understood in terms of measuring the error, guiding optimizations, and evaluating the model. Lower loss values indicate better model performance. Based on the concern whether we want to improve outliers or consider

a large varying range of predictions, one can either go for mean squared loss (MSE) as defined in equation 5.4 or use the absolute of relative error (*absloss*) as defined in equation 5.5.

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{5.4}$$

$$absloss = \| \frac{(E_{true} - E_{pred})^2}{(E_{true} + 0.01)} \| \tag{5.5}$$

The fact that loss function *absloss* focuses more on the relative error rather than the absolute error, makes it easier for the model to compare performances across different scales. Hence, using *absloss* as the loss function for the energy regression of pions is a preferred choice.

5.3.1 Optimization of dynamic reduction network architecture

In addition to optimizing the input features and target data provided to the model, it is crucial to tune a set of hyper-parameters to achieve optimal performance. Given that DRN [115] has a complex architecture that incorporates various types of Fully Convolutional Neural Networks (FCNNs), each component should be optimized individually. Few of such hyper-parameters specific to DRN are listed below;

- Number of input layers
- Number of aggregation layers
- Dimensions of latent space
- Number of output layer
- Number of message passing layers
- Learning rate
- Optimization of batch size

A thorough optimization of the hyperparameters is done and presented in this section. Based on this observation, the best performing model is selected for the results presented in the following sections. The model's performance will be quantified using the following performance metrics, designed with the specific problem context in mind:

- Training time : how much time it takes to train the DRN per epoch ?
- How the model performs on validation & training ? Performance should be identical on both dataset.
- Over-training : over-training occurs when, with an increasing number of epochs , the training loss continues to decrease while the validation loss starts to increase. This pattern indicates that the model is over-fitting, meaning it is becoming too specialized to the training dataset and is not generalizing enough to perform well on the validation dataset.
- Energy resolution and energy response as a function of incident pion energy : idea is to minimize the resolution and obtain a linear energy response.

We implemented the DRN model in PyTorch geometric [127]. The model is trained on a set of $\sim$4 M simulated pion showers, as explained earlier in 5.1, generated with a flat incident energy distribution. Approximately 20% events are reserved for validation of models using the network parameters derived in the training procedure performed on the remaining 80% events. The choice of training and validation batch sizes is majorly governed by the RAM available on the GPU processor. Because optimizing the performance of the model in terms of GPU usage and time it takes to train has

not been the aim of this study, hence, the batch sizes vary based on the GPU nodes specifics and available RAM.
For training the model, we are using NVIDIA V100 and A100 GPU nodes. With given granular information and training the model on such large datasets, one can use effective batch size. In these studies, we use gradient accumulation to effectively increase the batch size without needing a proportional amount of RAM. This technique allows us to simulate a larger batch size by accumulating gradients over several smaller batches before performing a weight update. To implement this, one has to decide the effective batch size based on the available RAM, and the number of batches it can handle. Then, we modify the training loop so that instead of updating model parameters after each epoch, the gradients are accumulated over several mini-batches and perform a single update of the model's parameters.
In addition to that, to fit the reasonable batch size for events with such high rechit multiplicity per event on NIVIDIA V100 GPU nodes with only 16 GB RAM on board, we make use of data parallelization with the help of DataParallel class of PyTorch [128]. With that we can run a job over multiple GPU nodes while sharing the RAM available on each node, and this makes training the DRN faster by at least an order of N if using N nodes for data parallelization. For example, to train a DRN architecture with 2 aggregation layers, 3 input, and 2 output layers with 5M events on a single GPU nodes, it takes approximately 90-150 minutes per epoch. Using the DataParallel class to parallelize the training on 4 GPU nodes, the training time is reduced to 20 to 40 minutes per epoch.

We divide the flat energy range into equal bins of 4 GeV, making a total of 85 such energy bins, and obtain the predicted energy distributions for each bin. The 4 GeV bin width is chosen intuitively to present the results. We present response and resolution of pions obtained in similar way as discussed in Section 5.2.1 for the following variations of hyper-parameters, input features and target in the rest of the section:

- Impact of dropping the AHCAL layers on resolution and response
- Impact of statistics on the performance
- Optimizing number of aggregation layers
- Optimizing number of layers in input net
- Optimizing learning rate
- Optimizing the input features and target

For the various models examined within this section, both the training and validation loss have been consistently observed to decrease until reaching a minimum, where they remain congruent with one another, unless indicated otherwise. Similarly, the performance of these models on the training and validation datasets exhibits considerable similarity. To maintain brevity and clarity in the discussion, such comparisons are not elaborated upon in this section unless explicitly mentioned otherwise. These aspects are thoroughly presented and analyzed for the ultimate optimized DRN models used for final results in Section 5.4.

Also, we observe edge effects in energy resolution and response for energies below 20 GeV and above 300 GeV. These effects are understood to be artifacts of using a flat sample with sharp edges at 10 and 350 GeV. This issue can be mitigated by employing a flat sample within the range of interest but with Gaussian edges extending beyond the relevant energies, as demonstrated in the thesis by Nitish [129] . To avoid these edge effects, we are showing the results with 20 to 300 GeV unless stated otherwise. In the figure legends of this section, "Training" refers to simulated dataset used

for training the models & "Validation" or "Simulation" refers to the dataset used to validate the models.

Impact of dropping the AHCAL layers on DRN's performance

Since we use information only from every 4^{th} layer of the AHCAL, reducing the number of sampling layers to 10 to realistically match the HGCAL scintillator sections, we explicitly checked its impact on the energy resolution and response of charged pions. Two DRNs with identical architecture are trained on AHCAL with 39 samplings layers and AHCAL with only 10 sampling layers using the flat energy pion sample in beam test geometry. The following model parameters are identically used for both trainings:

- Target : $\frac{E_{true}}{E_{fix}}$
- loss function : absloss
- Input features : (E,x,y,z) of rechits
- Aggregation layers : 2
- Input net layers : 3
- output net layers : 2
- Message passing layer : 2
- Epoch : 90
- learning rate : 10^{-4}, constant

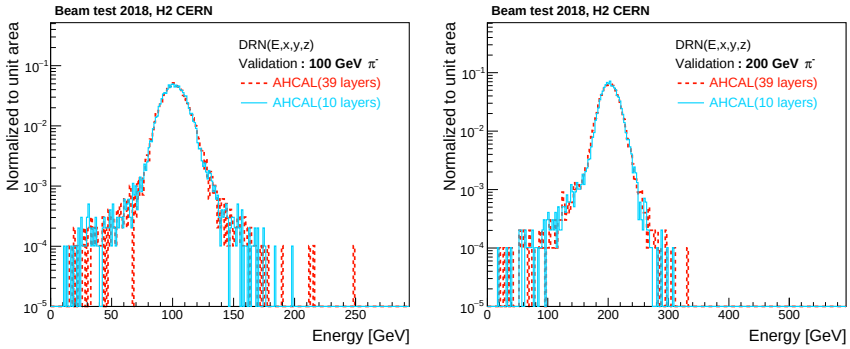


Figure 5.10: Distribution of energy reconstructed with 39 AHCAL layers and 10 AHCAL layers for incident pions of 100 GeV (left) and 200 GeV (right) in validation dataset.

Comparisons of energies reconstructed with 39 AHCAL layers and 10 AHCAL layers are shown in Fig. 5.10 for incident pions of 100 (left) and 200 GeV (right) in validation dataset. The DRN is able to reproduce energy distributions using information from lesser number of AHCAL layers to match those with all 39 AHCAL layers.
Fig. 5.11 compares the energy resolution and response on the validation data set obtained using DRNs trained with all responses from 39 sampling layers of AHCAL versus with 10 sampling layers. The resolution worsens at most by $\sim 5\%$ for higher energy pions when the provided information is reduced from 39 to 10 AHCAL layers, and we believe this is mainly attributed to the increased fluctuations in the inputs. There is no significance difference in the response. This exercise indicates that the DRN's performance is almost unchanged with respect to the information from AHCAL. For rest of the studies shown in this chapter, we will be using the information from only 10 layers of AHCAL, to stay close to the final HGCAL geometry.

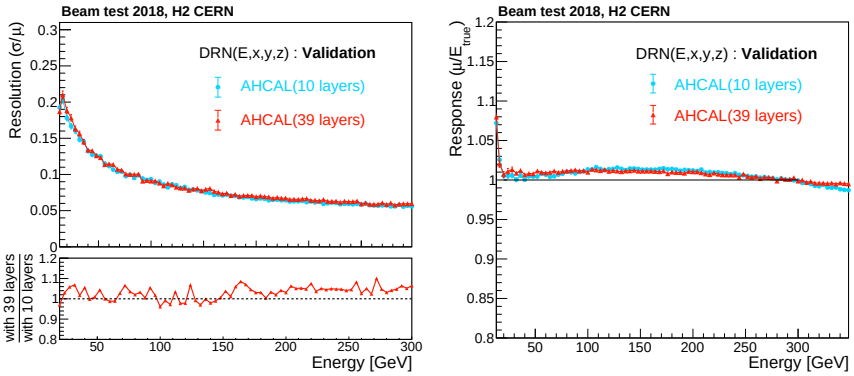


Figure 5.11: Energy resolution (left) and response (right) as functions of incident pion energy for DRNs with 39 versus 10 sampling layers in AHCAL on validation dataset. The left bottom panel compares resolutions obtained with 39 versus 10 sampling layers.

Impact of training sample statistics on the performance

As mentioned in the section 5.3.1, it takes approximately 150 minutes to train DRN for 1 epoch (without parallelization), and we need to train for at least 100 epochs for optimal performance and stable results. This is attributed to the fact that even with the simplest architecture that includes 2 aggregation layers, we have approximately 63 k parameters to optimize. Complexity is further increased with highly granular information available for the hadron showers, making it unrealistically slow to train and test the model while doing multi-dimensional optimization of different hyperparameters. To reduce the training time, we checked the resolution and response of the pion energy obtained by training a DRN with the same architecture as mentioned in Section 5.3.1 but using only ∼1 M events.

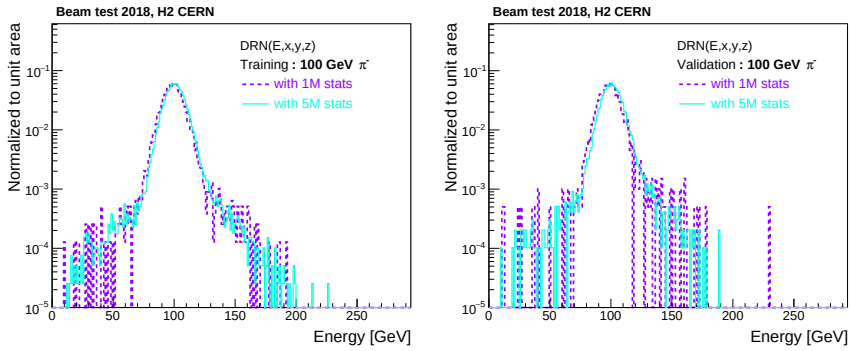


Figure 5.12: Reconstructed energy distributions for 100 GeV pions in training (left) and validation (right) dataset obtained by training the DRN with 5 M and with 1 M events.

Comparisons of reconstructed energy distributions for 100 GeV pions in training (left) and validation (right) data sets obtained by training the DRN with 5 M events versus 1 M events are presented in Fig. 5.12. Shapes of the distributions are very similar. Comparisons of resolution (response) obtained from training using 5M and 1M events are shown in Fig. 5.13 for validation datasets, respectively. The bottom panel shows the ratio of resolution obtained from 1M events to 5M events and indicates a worsening of resolution by 5-10% across the full range by reducing the statistics.
Training a DRN for 100 epochs takes approximately 15 hours to 2 days, depending on whether we use 1 M or 5 M events, with effective batch sizes and data parallelization in place. Given that there is no drastic worsening in resolution or response by using

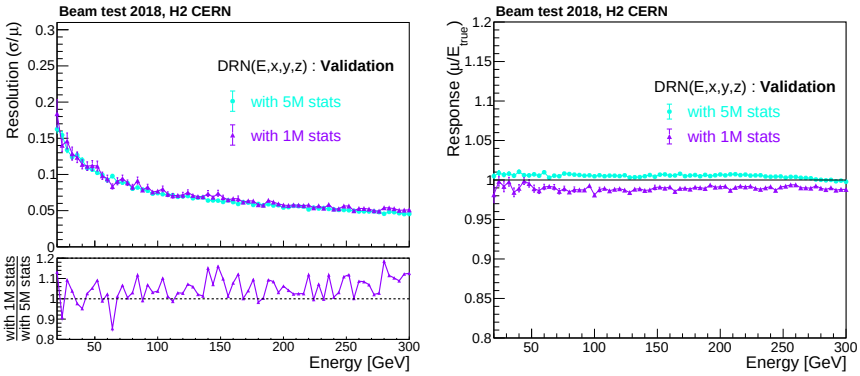


Figure 5.13: Energy resolution (left) and response (right) comparisons from DRN trained on 5M and 1M events for validation datasets. Bottom panel shows the resolution ratio of 1M to 5M event statistics.

a sample of 1 M pions, we use 1 M dataset to check the impact of various hyper-parameter optimization studies. These results indicate that DRN stabilizes the energy resolution much earlier than the energy response. In addition, the uncertainties in the energy resolution obtained with the validation dataset increase when we use less statistics, leading to more fluctuations in the energy response and resolution. For this reason, we will be quoting the results from DRN based on both training and validation data sets in this section.

Optimizing number of aggregation layers

As described in Section 4.3, the aggregation step in the DRN essentially combines and reduces data effectively, and ensures leveraging important features from different layers of the network, leading to a more robust and efficient model. Increasing the number of aggregation layers by twice in the DRN approximately doubles the number of parameters to be optimized, which invariably leads to an increase in time and sample statistics required to train the model reliably. Hence, an optimization of the number of aggregation layers is important to achieve an optimal balance between desired physics performance, computing time, and avoiding over-training.

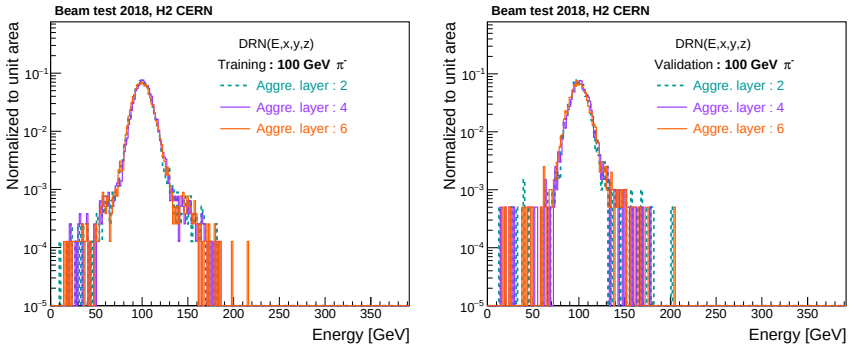


Figure 5.14: Distributions of reconstructed energy for pions of 100 GeV in training (left) and validation (right) datasets using DRNs with different aggregation layers.

We train three DRN models keeping all model parameters listed below unchanged except the number of aggregation layers and check its impact on energy resolution and response of pions and on time take for training the model.

- Target : $\frac{E_{true}}{E_{fix}}$

- loss function : absloss
- Input features : (E,x,y,z) of rechits
- **Number of aggregation layers : 2 or 4 or 6**
- Input net layers : 4
- output net layers : 2
- Message passing layer : 2
- Epoch : 100 - 140
- learning rate : 10^{-4}, constant

The comparative analysis of the reconstructed energy distributions for 100 GeV pions utilizing the three models is illustrated in the figure on the right (left) side of 5.14, corresponding to the validation (training) dataset. The bulk of distributions for the three cases look similar. The energy response and resolution obtained from the three models, with a number of 2, 4 & 6 are compared in Fig. 5.15. The bottom panel on the left canvases shows the ratio of resolution with respect to the model with two aggregation layers. The models with 2 & 4 aggregation layers show similar resolution (within a few %) while the former is slightly preferable, as it seems to provide more stable response. Slightly worse resolution and response predicted by the model with 6 aggregation layers may be attributed to a much larger number of parameters to be optimized, resulting in higher complexity of the model. Larger statistical fluctuations in the results from validation dataset are attributed to limited statistics, approximately 200 k events in the range of 10 to 350 GeV incident pion energies. Based on these studies, we have chosen DRN with 2 aggregation layers for the rest of the pion studies presented in this chapter.

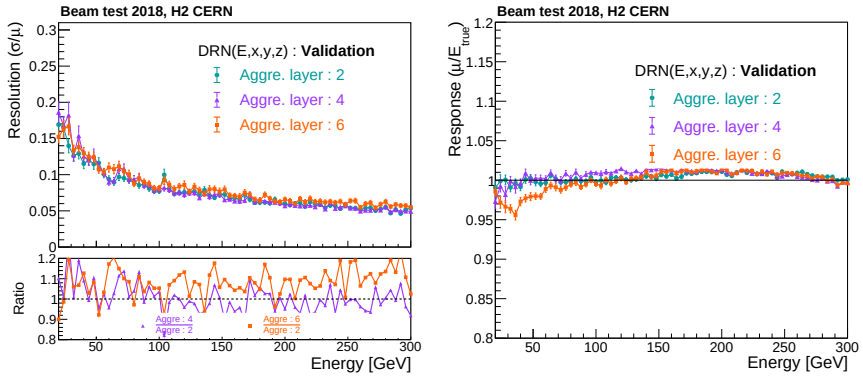


Figure 5.15: Energy resolution (left) and response (right) versus incident pion energy for DRN models with 2, 4, and 6 aggregation layers on validation datasets. The ratio panel on left shows resolution ratios using 4 and 6 layers compared to 2 layers.

Optimizing number of layers in input net

Number of input layers in any kind of neural network essentially drives its performance and its ability to learn and represent complex patterns while adding nonlinearity to the model. To further optimize the model architecture, we vary the number of layers in the input net and evaluated DRN's performance with different input net layers while keeping the other parameters same :

- Target : $\frac{E_{true}}{E_{fix}}$
- loss function : absloss
- Input features : (E,x,y,z) of rechits

- Aggregation layers : 2
- **Input net layers : 4 or 3**
- output net layers : 2
- Message passing layer : 2
- Epoch : 100 - 140
- learning rate : 10^{-4}, constant
- statistics : 1 M (80:20 split between training and validation dataset)

The shape comparisons for reconstructed energy distributions using DRNs with different number of layers in input for 100 GeV (left) and 200 GeV (right) pions are shown in Fig. A.3 of Appendix A.1. The overall shape is reproduced with reducing down the number of layer from 4 to 3. Energy resolution and response obtained from the model with 4 and 3 input layers are shown in the Fig. 5.16 for validation dataset. The ratio panel on the left plots is indicating the improvement or worsening of the resolution upon using DRN with 4 layer versus 3 layers in the input net. These results indicate that we do not benefit much from using larger number of input layers and the model gives optimal results for 3 input net layers.

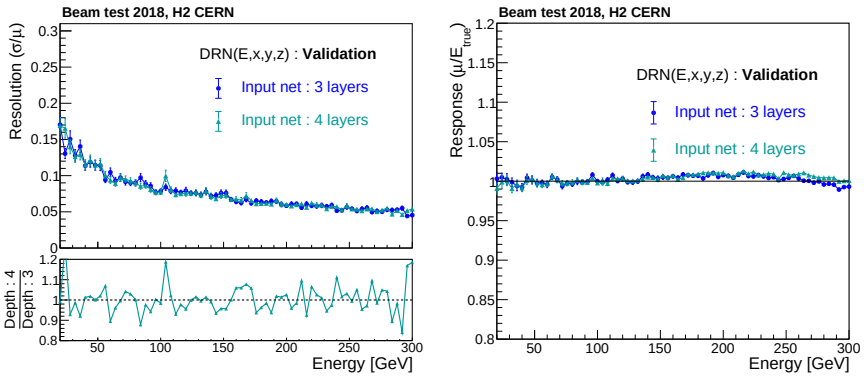


Figure 5.16: Energy resolution (left) and the energy response (right) as a function of incident pion energy obtained on validation dataset using DRNs with different number of input layers.

Optimization of learning rate

The learning rate is a hyperparameter that influences the training dynamics of DRN models, and correct tuning of the learning rate is essential to achieve effective and efficient training. It determines the speed at which the model's parameters (weights) are adjusted in response to the calculated gradient during training. The choice of learning rate is crucial to avoid overshooting the minima of loss function and resulting into unstable results (high learning rate) or results into slow convergence of model, more computing power consuming (low learning rate). The results summarized so far are based on a DRN with a constant learning of 10^{-4}.

There are different ways available to find the optimal learning rate for model training. The choice is either to start with a small learning rate to ensure stable training or use learning rate schedulers to adjust the learning rate based on the training progress. One can also opt for keeping the learning rate constant throughout the training. To quantify whether or not the set learning rate or schedule is optimal for the purpose, we observe the training and validation loss curves. Depending on whether the training loss decreases but the validation loss increases or if the loss fluctuates significantly, one may need to adjust the learning rate.

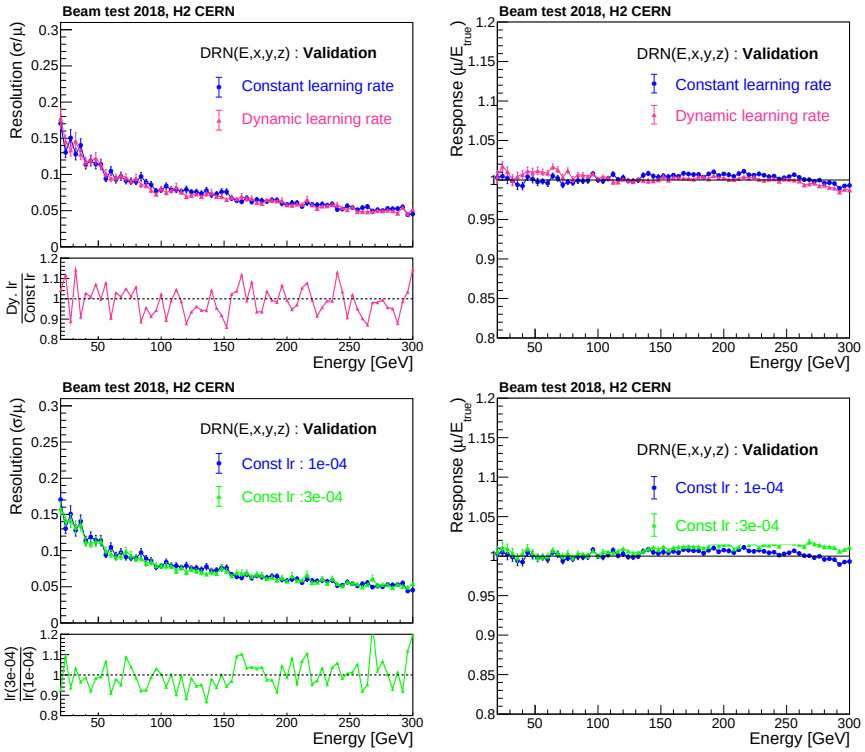


Figure 5.17: Energy resolution (left) and energy response (right) versus incident pion energy for DRNs trained with a constant versus dynamic learning rate (top row) and 3.10^{-4} vs 10^{-4} constant learning rates (bottom row) for validation datasets. The bottom left panel compares the resolution ratio of dynamic to constant and 3.10^{-4} vs 10^{-4} learning rate.

Keeping the following parameters unchanged except for the learning rate, we trained DRN keeping the learning rate constant versus varying the learning rate using a cyclic scheduler. The latter is a technique where the learning rate is varied within a predefined range throughout the training rather than fixing it or monotonically increasing it.

- target : $\frac{E_{true}}{E_{fix}}$
- loss function : absloss
- input features : (E,x,y,z) of rechits
- aggregation layers : 2
- input net layers : 3
- output net layers : 2
- message passing layer : 2
- epoch : 100 - 140
- learning rate : 10^{-4} or 3×10^{-4} constant or Cyclic learning rate scheduler.
- statistics : 1 M (80:20 split between training and validation dataset)

Fig. A.4 of Appendix A.1 shows that the models trained with constant learning rate or dynamic learning rate predict similar distributions of reconstructed pion energies. Fig. 5.17 (top row) illustrates the energy resolution and response as a function of incident pion energy estimated for validation dataset using DRN trained with constant and dynamic learning rates. The performance in terms of energy resolution is very similar.Since the learning rate has to do with model's stable performance, it is clearly

visible in the energy response obtained with constant vs. dynamic learning rate. We obtained much stable energy response if we use constant learning rate.
The energy distributions obtained by training the DRNs with a constant learning rate of 10^{-4} or 3×10^{-4} are shown in Fig. A.5 of Appendix A.1 for 100 and 200 GeV pions on training (upper row) and validation (bottom row) datasets. The shapes are reproduced overall with different DRNs. The bottom row of Fig. 5.17 shows the energy resolution and response as functions of the incident pion energy for validation datasets, estimated using DRNs trained with various constant learning rates. The ratio panel illustrates the improvement in energy resolution achieved with these different constant learning rates.
The performance in terms of energy resolution is quite similar across different learning rates, indicating that the DRN stabilizes the resolution well before the energy response. However, the learning rate impacts the stability of the model performance, as evidenced by the energy response. We observed that a constant learning rate of 10^{-4} leads to a more stable energy response compared to 3×10^{-4}.

5.3.2 Optimizing the input features and target

In this section, we test DRN with different targets, input features, and different combinations of input features provided to it to learn the details of the pion showers. DRN with the following architecture has been used to conduct these tests:

- Target : $\frac{E_{true}}{E_{fix}}$ or E_{true}
- loss function : absloss
- **Input features : (E, x, y, z) of rechits or (E, x, y) or (E,x,y,λ_{int})**
- Aggregation layers : 2
- Input net layers : 3
- output net layers : 2
- Message passing layer : 2
- Epoch : 100 - 140
- learning rate : 10^{-4}
- statistics : 1 M (80:20 split between training and validation dataset)

Naming convention used in the text for different DRNs based on input features used to train, namely:

- DRN(E) - energy
- DRN(E,z) - energy and z coordinate (depth)
- DRN(E,x,y,z) - energy, z coordinate (depth) and x & y coordinates (transverse position)
- DRN (E,x,y) - energy, x and y coordinates

Optimizing training target

Fig. A.6 of Appendix A.1 shows the shape comparisons for the reconstructed energy distributions in training and validation datasets for 20, 100 and 200 GeV pions using DRN with a ratio of E_{true} to E_{fix} as target and DRN using E_{true} as target.
Fig. 5.18 shows the energy resolution and response as a function of incident pion energy estimated for the validation data set using DRN trained using E_{true}/E_{fix} and E_{true} as the targets. The bottom panel shows the ratio of energy resolution obtained using E_{true} as a target to that of $\frac{E_{true}}{E_{fix}}$ as target. Using $\frac{E_{true}}{E_{fix}}$ as the target, we observe a more than 10% improvement in resolution for low-energy pions. The energy response is also stable, especially at lower incident energies, indicating that the ratio $\frac{E_{true}}{E_{fix}}$ as

the target is a good choice. It essentially provides a measure of how much correction is needed to align E_{fix} with the actual truth values.

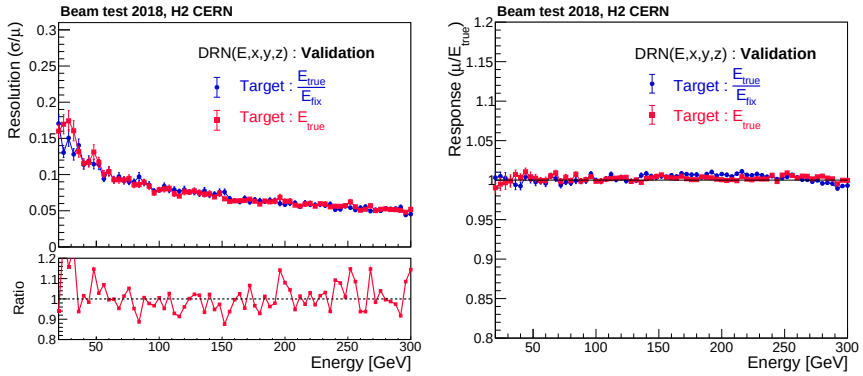


Figure 5.18: The left plot shows energy resolution, while the right plot illustrates energy response, both as functions of incident pion energy for validation datasets using DRN(E,x,y,z) with ratio E_{true} to E_{fix} as the target, and DRN with E_{true} as the target. The bottom panel on the left displays the ratio of the latter mode to the former.

Optimizing input features

We train a DRN with the architecture described above and $\frac{E_{true}}{E_{fix}}$ as a target with different inputs as follows:

- DRN(E) trained with 5M events and 100 epochs, as also shown in Section 5.3, using rechit energies as input features
- DRN(E,x,y) with 1M events and 100 epochs, using rechit energies and their transverse position x & y as input features
- DRN(E,x,y,z) with 5M events and 100 epochs, as also shown in Section 5.3, using rechit energies and their transverse position x & y as well as longitudinal locations z as input features
- DRN(E,x,y,λ_{int}) with 1M events and 100 epochs, using rechit energies and their transverse position x & y and λ_{int} instead of z position as input features.

Fig A.7 of Appendix A.1 shows comparisons of energy distributions reconstructed with DRN(E), DRN(E,x,y) and DRN(E,x,y,z) for 100 GeV and 200 GeV pions. The respective energy resolution and response are presented in the top row of Fig. 5.19. The bottom panel of the left plot shows the ratio of energy resolution obtained with DRN(E) and DRN(E,x,y) with respect to that obtained with DRN(E,x,y,z). It is observed that omitting depth information (z) while providing only coordinates (x, y) does not significantly improve the DRN's ability to capture detailed features of the shower. This is because the coordinates of x and y vary between different subsections and their overlapping regions become more distinct when information about z is included, as illustrated in Figure 5.3.

We trained the DRN model replacing the z coordinate of rechits with the interaction length (λ_{int} while keeping the other parameters the same. The models were trained for at least 100 epochs. Fig. A.8 of Appendix A.1 compares the energy reconstructed using DRN(E,x,y,z) and DRN(E,x,y,λ_{int}) for 100 and 200 GeV pions. The overall shape looks similar, with small differences in the bulk part of the distributions.

Fig. 5.19 (bottom row) shows the energy resolution and response as a function of incident pion energy, estimated for validation datasets using DRN(E,x,y,z) and

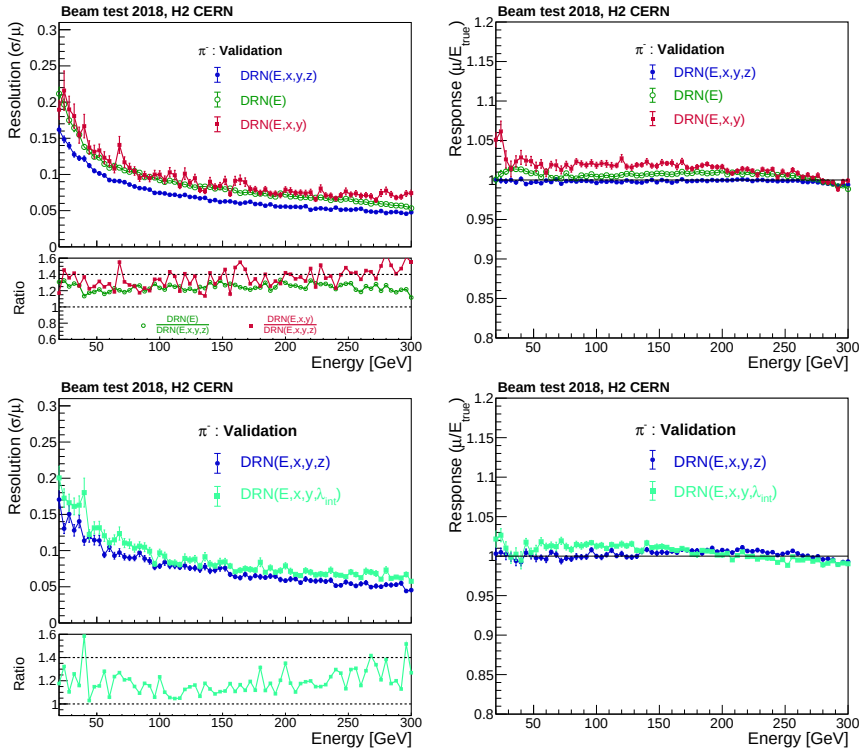


Figure 5.19: Energy resolution (left) and response (right) as functions of pion energy for validation datasets using DRN(E,x,y,z), DRN(E), and DRN(E,x,y) (top row) and using DRN(E,x,y,z), DRN(E), and DRN(E,x,y,λ) (bottom row). The bottom panel shows the energy resolution ratio of DRN(E) and DRN(E,x,y) or DRN(E,x,y,λ) compared to DRN(E,x,y,z).

DRN(E,x,y,λ_{int}). We observe that using z instead of λ_{int} improves energy resolution by up to 20%, whereas the energy response obtained is more linear. However, it is possible that the model state for DRN(E,x,y,λ_{int}) has not yet been achieved optimally and needs more iterations of epochs to achieve the optimal state.

The final results show limited sensitivity to changes in most hyperparameters. This is partly because the variations explored were relatively small, and these parameters mainly affect computational aspects rather than physics performance. This corroborates with the fact that the richness of information in HGCAL allows the model to learn complex shower patterns effectively, reducing the need to fine-tuning hyperparameters.

Test	Model architecture							Conclusion
	Input net	Aggregation layers	Output net	Learning rate	Target	Input features	Statistics	
Layers in AHCAL (39 vs 10)	3	2	2	10^{-04} (const)	$\frac{E_{true}}{E_{fix}}$	(E, x, y, z)	5M	AHCAL with layer-10
5M vs 1M statistics	3	2	2	10^{-04} (const)	$\frac{E_{true}}{E_{fix}}$	(E, x, y, z)	5M (1M)	Use 1M to optimize
Aggregation layers (2 vs 4 vs 6)	3	2, 4, 6	2	10^{-04} (const)	$\frac{E_{true}}{E_{fix}}$	(E, x, y, z)	1M	Use 2 aggregation layers
Input net 4 vs 3	4 (3)	2	2	10^{-04} (const)	$\frac{E_{true}}{E_{fix}}$	(E, x, y, z)	1M	Use 3 layers in input net
Learning rate (const. vs dynamic)	3	2	2	10^{-04} (const) vs dynamic	$\frac{E_{true}}{E_{fix}}$	(E, x, y, z)	1M	Use constant learning
Learning rate 10^{-04} vs 3.10^{-04}	3	2	2	10^{-04} (3.10^{-04})	$\frac{E_{true}}{E_{fix}}$	(E, x, y, z)	1M	Use constant learning of 10^{-04}
Target $\frac{E_{true}}{E_{fix}}$ vs E_{true}	3	2	2	10^{-04}	$\frac{E_{true}}{E_{fix}}$ (E_{true})	(E, x, y, z)	1M	Use $\frac{E_{true}}{E_{fix}}$
λ_{int} vs z	3	2	2	10^{-04}	$\frac{E_{true}}{E_{fix}}$	(E, x, y, z) or (E, x, y, λ_{int})	1M	Use z

Table 5.1: The above table lists the energy resolution obtained for different pion energies using various methods in Data.

5.4 Performance of Dynamic reduction network for charged pion energy reconstruction

Based on the detailed optimization as described in section 5.3.1, we opt for the following set of model parameters to test the effect of using three different sets of features as input features.

- Loss function = $\| \frac{(E_{true} - E_{pred})^2}{(E_{true}+0.01)} \|$
- Input layers : 3
- Aggregation layers : 2
- message passing layers : 2
- Output layers : 2
- Train batch size : 800
- Valid batch size : 400
- Constant learning rate : 0.0001
- Number of GPU nodes : 4
- Training time : 20 or 30 or 40 minutes per epoch
- Total number of epochs trained: 150 or 186 or 136
- Target : $\frac{E_{true}}{E_{fix}}$
- List of energy (E), energy and longitudinal position (E, z), and energy and both lateral and longitudinal positions (E, x, y, z) of each detector channel or rechit are provided to train three models, namely:
  - DRN(E) - energy
  - DRN(E,z) - energy and z coordinate (depth)
  - DRN(E,x,y,z) - energy, z coordinate (depth) and x & y coordinates (transverse position)

The DRN(E), DRN(E,z) and DRN(E,x,y,z) models have been trained for a total of 150, 180 and 136 epochs, respectively, until the value of loss function evaluated on the validation dataset stopped improving in subsequent epochs. Figure 5.20 shows

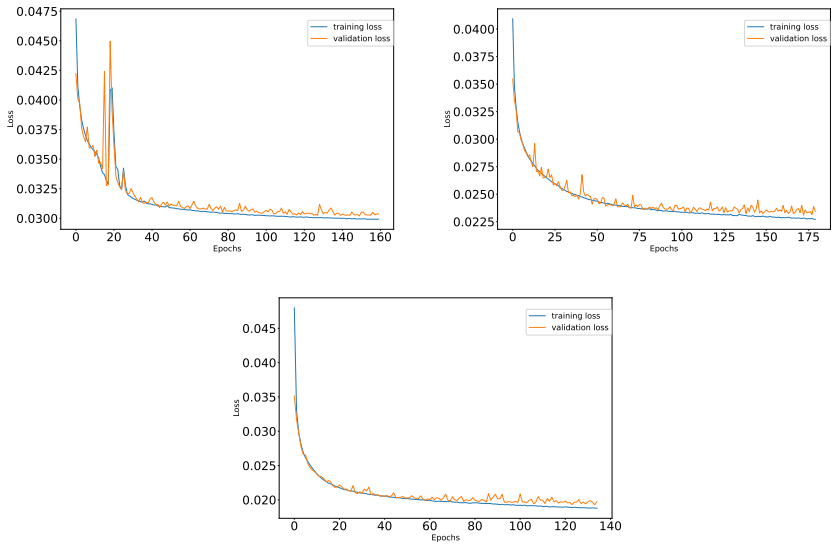


Figure 5.20: Training and validation loss over epochs for each DRN model: top-right for DRN(E,x,y,z), top-left for DRN(E,z), and bottom for DRN(E). The losses generally align, with no clear indication of over-training.

the variation in training and validation loss as a function of epoch. The training and validation loss agrees very well for the initial set of epochs, but validation loss starts deviating away from the training loss indicating the possible over-training. To evaluate the performance of the model, we use the learned training weights for the epoch with minimum validation loss. To avoid the possibility of picking up the epoch with overtrained DRN, we have evaluated DRN's performances on three epochs with best three minimum validation loss as listed in the table 5.2 and compared the DRN's performance on these set of epochs and picked the best performing so far.

S.No.	DRN(E)		DRN(E,z)		DRN(E,x,y,z)	
	Epoch	Min. loss	Epoch	Min. loss	Epoch	Min. loss
I	95	0.032	74	0.0236	79	0.0197
II	110	0.0304	94	0.0238	100	0.0196
III	150	0.0302	114	0.0230	135	0.0193

Table 5.2: The list of epochs used to evaluate different DRNs performance along with listing down the minimum validation loss.

Training against $\frac{E_{true}}{E_{fix}}$ as the target, the predictions of the models are obtained as $\frac{E_{pred}}{E_{fix}}$. The top row of Figure 5.21 displays distributions of the training target and target predicted by the DRN at a few different epochs for DRN(E), DRN(E,z) and DRN(E,x,y,z), from left to right respectively. We obtain the predicted shower energy by multiplying the predicted target by E_{fix}. Figure 5.21 (bottom row) shows the predicted energy or E_{reco} at different epochs for DRN(E), DRN(E,z) and DRN(E,x,y,z) from left to right, respectively. The gray dotted curve indicates the E_{true} for the sample pions used for these studies. The target provided to DRN varies from very close to 0 up to 50 or so, hence the predictions from the DRN also follow a similar distribution. We put a hard limit on the predicted target for it not to be more than 3 to avoid any large corrections predicted by the DRN. It has been found that putting this hard limit does not affect the DRN's performance in terms of energy resolution and response.

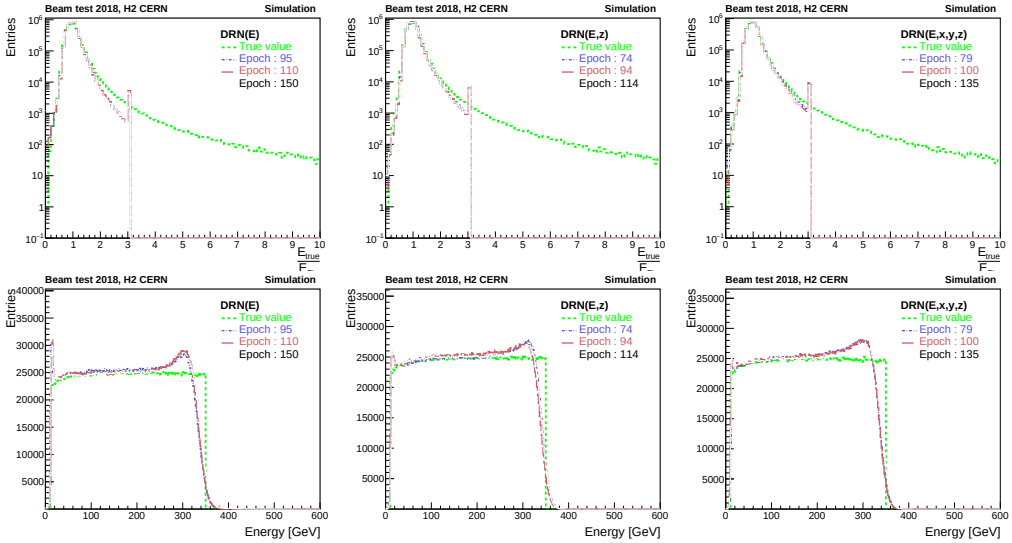


Figure 5.21: The shape of predictions obtained from DRNs (top row), and the shape of predicted energy for flat sample (bottom row), for DRN(E), DRN(E,z) and DRN(E,x,y,z) models from left to right, respectively.

The figure 5.22, shows comparisons of the energy distributions predicted by the three DRNs for the training (top) and validation (bottom) datasets at different epochs for 100 GeV pions. The validation data set is referred to as "Simulation" in the legends of the plots for the rest of the section. The general shapes do not change significantly with respect to different epochs. Similar comparisons for 200 GeV pions are shown in Fig. A.9 of Appendix A.1.

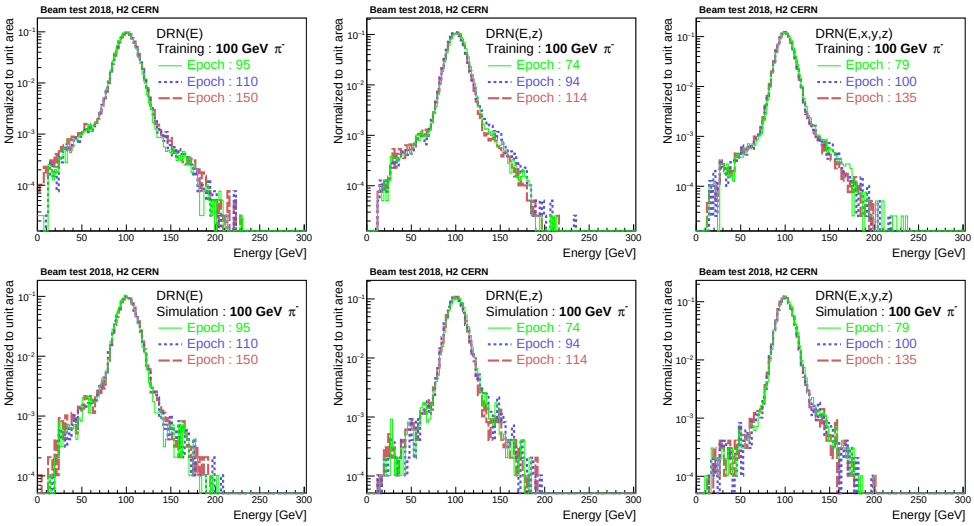


Figure 5.22: Reconstructed energy distributions evaluating DRN's performance at different epochs using DRN(E) (left), DRN(E,z) (middle) and DRN(E,x,y,z) (right) on training (top) and validation or referred as simulation (bottom) dataset for 100 GeV pions. The shape agrees overall except quite different tails.

These energy distributions are fitted with a Gaussian function approximately in the range of one standard deviation around the mean measured energy, and μ and σ parameters from the fitted Gaussian are extracted to define relative resolution as $\frac{\mu}{\sigma}$ and energy response as $\frac{\mu}{E_{true}}$, as shown in Fig. 5.23 for the inference made at 135^{th} epoch. Similar plots showing Gaussian fits along with estimated energy resolution for other energy points using DRN(E,x,y,z) are added in Figure A.10 of the appendix A.1.

Comparisons of energy resolution and response as a function of incident pion energy as obtained from DRN(E), DNR(E,z) and DRN(E,x,y,z), evaluated at three different epochs in validation data sets, are shown in Figure 5.25 (from left to right). Interestingly, the resolution obtained at different epochs for three DRNs is almost similar, indicating that resolution is minimized and stable even if we pick the different epoch to evaluate DRN's predictions. However, the energy response obtained fluctuates more with inferring on different epochs. Based on these observations, we have selected **110^{th}, 114^{th} and 135^{th} epochs for DRN(E), DRN(E,z) and DRN(E,x,y,z) as the best epoch, respectively**. For the rest of the studies presented in this chapter, these models will be used as a reference, and best performing models and performance with validation dataset (referred to as Simulation in legend) will be used as reference.

Fig. 5.24 shows event-by-event comparison of predicted energy versus the true incident pion energy in flat energy sample using different DRNs at the best epochs is shown in the Fig. 5.24.

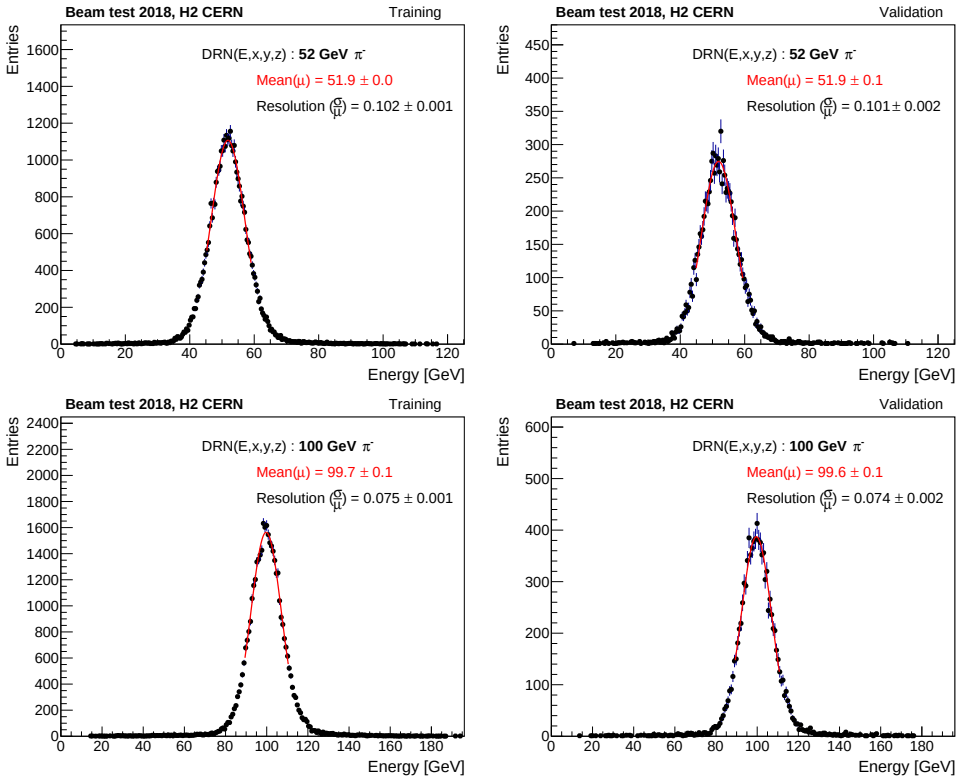


Figure 5.23: Energy distributions reconstructed by DRN(E,x,y,z) along with the Gaussian fitted to extract σ and μ, shown in legend.

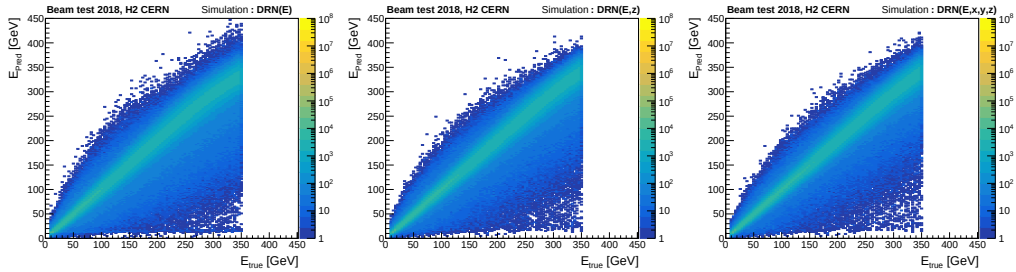


Figure 5.24: The event-by-event comparison of the true pion energy versus the reconstructed using different DRNs evaluated at the best epochs for DRN(E), DRN(E,z) and DRN(E,x,y,z) from left to right respectively.

Performance of DRN on training and validation datasets

The predicted energies significantly deviate from the true energy values for energies less than 20 GeV and higher than 300 GeV, resulting in a response deviating from unity and resolution deviating from expected calorimetric behavior. This effect has been understood as an artifact of using a flat sample with sharp edges at 10 and 350 GeV. It can be resolved if we use a flat sample in the interest energy range but have Gaussian edges that extend beyond the interest energies as shown in the thesis [129]. To avoid the energy range affected by edge effects, we use the energy range of 20 to 300 in the rest of the studies presented in this chapter.

The distributions of the total pion energy predicted in the training and validation dataset by the three models are compared in Fig. 5.26 for 100 GeV pions. The energy distributions obtained with the validation dataset are reproducing the overall shapes of those obtained with the training dataset without hinting towards any bias. The

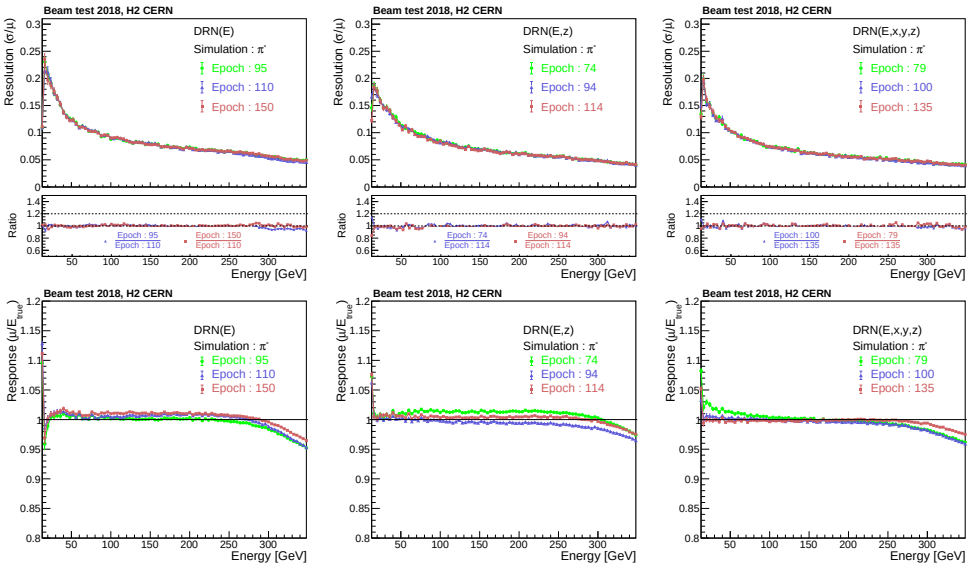


Figure 5.25: Energy resolution (top) and response (bottom) versus incident pion energies for various DRNs are compared at different epochs on the validation dataset. Epochs 110*th*, 114*th*, and 135*th* are chosen for DRN(E), DRN(E,z), and DRN(E,x,y,z), respectively, as the best for further studies.

Figure A.11 Appendix A.1 demonstrates the performance of DRNs on the validation and training data set for 200 GeV pions.

The response and resolution as a function of the incident pion energy obtained with the validation data set is well matched to those obtained in the training dataset, as shown in Fig. 5.27 for DRN(E) (left), DRN(E, z) (middle) and DRN(E, x, y, z) (right).

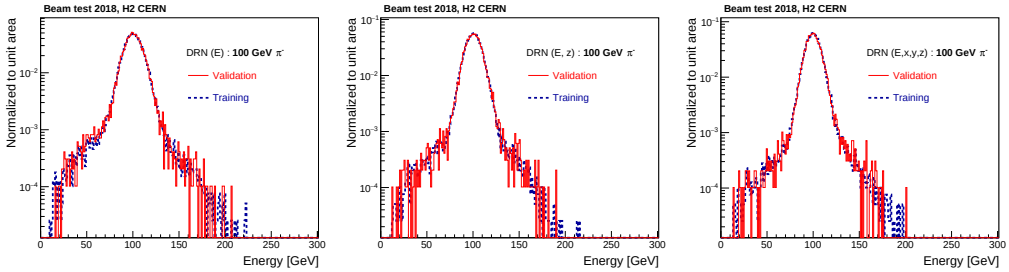


Figure 5.26: Comparison of distribution of energy obtained in training and validation datasets for the models DRN(E) (left), DRN(E,z) (middle), and DRN(E,x,y,z) (right) for pions with incident energy of 100 GeV.

Comparison of pion energy reconstruction using DRN and WS methods

Fig. 5.28 demonstrates the comparison of reconstructed energy distributions predicted on the validation dataset using the DRN(E) and WS method for 100 (left) and 200 (right) GeV pions. The overall bulk as well as the tails of the distribution shrink using DRN(E) as compared to WS method.

Comparisons of the energy distributions obtained for different incident pion energies using the three models are shown in Fig. 5.29. When DRN is presented with spatial information about the shower, we observe that the bulk part of the distribution shrinks as well high and low end tails are reduced. Similar comparison for other energy points with validation and training data set has been added in Figs. A.12 and

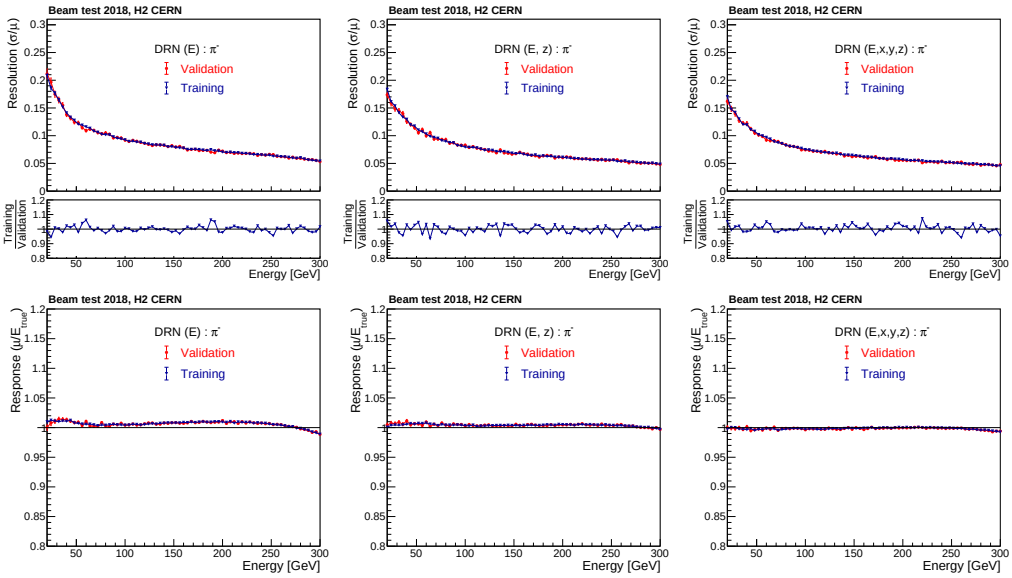


Figure 5.27: Comparison of resolution (upper row) and response (bottom row) obtained in training and validation datasets for the models DRN(E) (left), DRN(E,z) (middle), and DRN(E,x,y,z) (right) as function of incident pion energy.

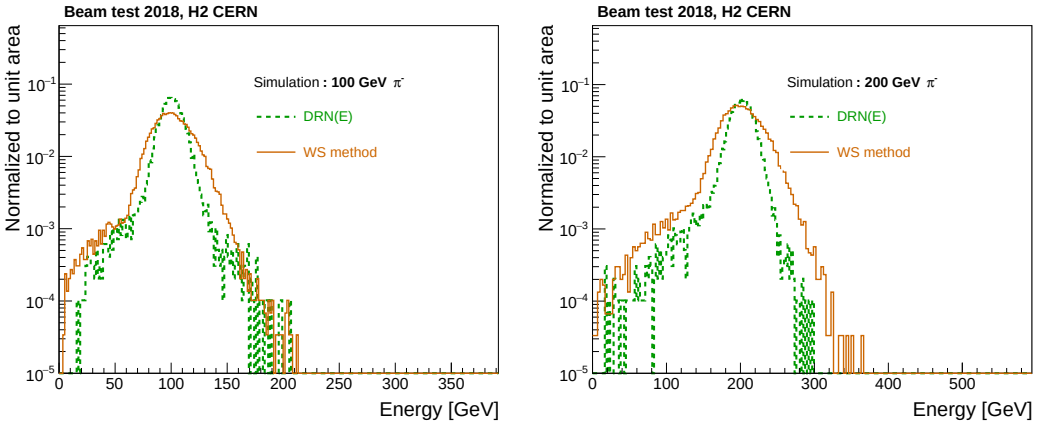


Figure 5.28: Comparisons of energy distributions predicted using different DRN(E) and WS method 100, and 200 GeV pions for validation datasets.

A.13 of the appendix A.1.

The energy resolution as a function of the incident pion energy obtained from the three models and the WS method are shown in Fig. 5.30 (top). The middle plot of Fig. 5.30 shows the resolution ratio obtained using the DRN(E), DRN(E,z) and the WS method to DRN(E,x,y,z).
When compared with WS method, we observed almost a factor of two improvement in energy resolution obtained using DRN(E) trained only on the rechit energy information. This improvement in the resolution can be understand as that DRN is able to learn and accurately compensate for the event-by-event fluctuations of pion showers with the help of granular information available via HGCAL beam test setup, whereas the current WS method is not equipped to do this. Presenting the DRN with rechits longitudinal and transverse positions, improves the energy resolution further up to 20-30%. These sequential improvements can be interpreted as an indication that the network has learn how the shower spreads in the calorimeter, and is able to

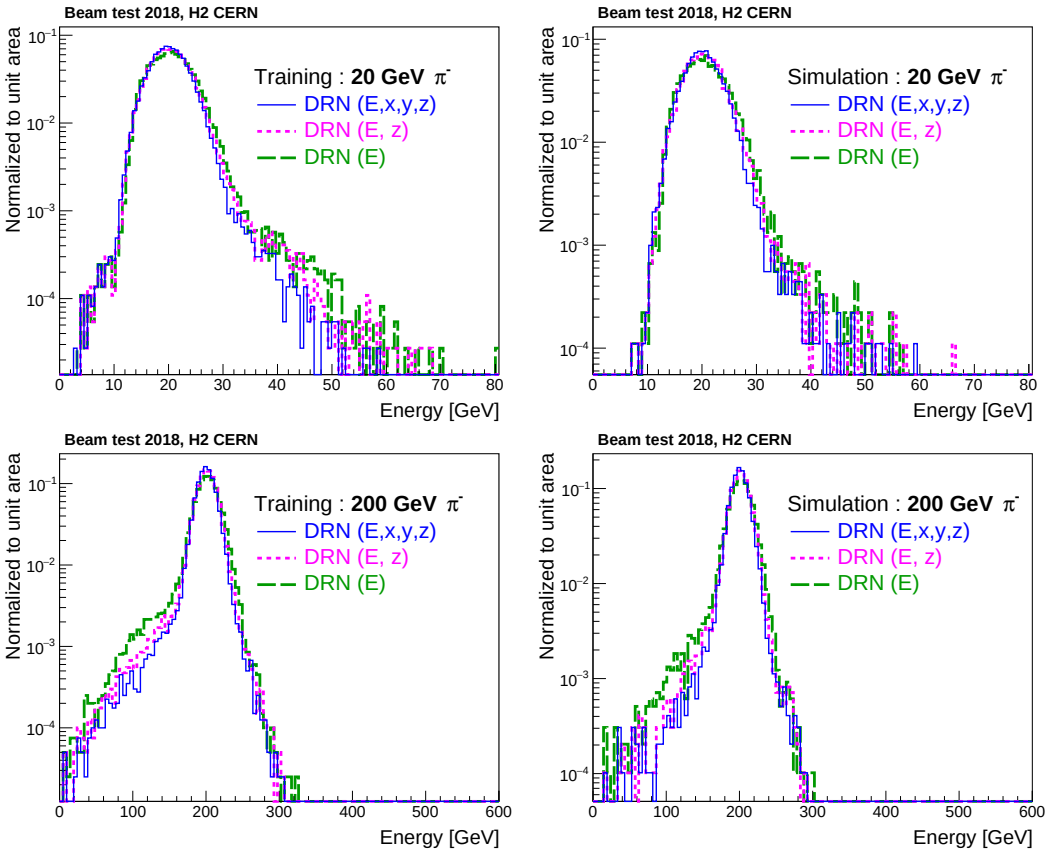


Figure 5.29: Distributions of energies predicted using different DRN models for 20 GeV (up) (middle) and 200 GeV (lower) pions on training (left) & validation referred as "Simulation" (right) datasets.

partially correct for the energy leakage of the detector. There is more than a factor of two improvement in the resolution obtained with DRN(E,x,y,z) with respect to WS method. The energy response obtained with three models as a function of incident pion energy is shown in Figure 5.30 (bottom) and it is close to true energy of pions. Linearity of energy response as a function of incident pion energy is observed for all the three models as well as with WS method.

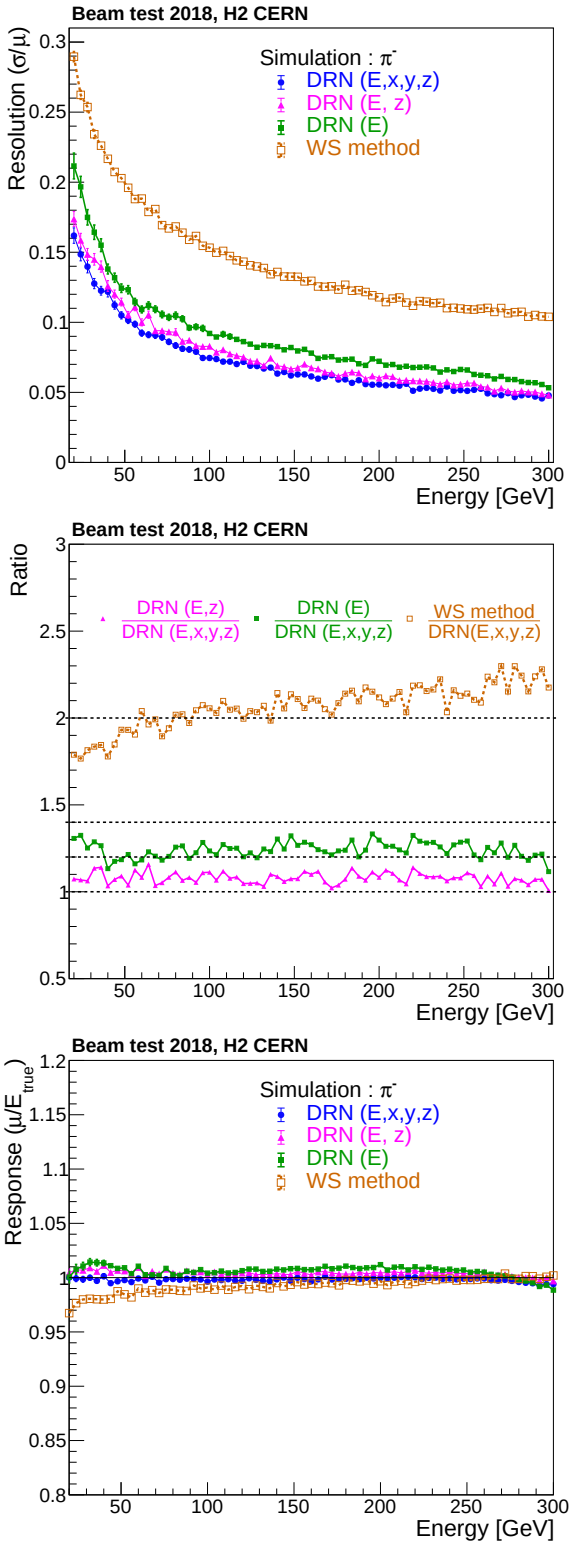


Figure 5.30: Resolution (top), ratio of resolution to DRN(E,x,y,z) (middle), and response (bottom) as function of pion energy obtained from the three models trained with different features in simulation.

5.4.1 Inference on shuffled order hits on DRN trained with ordered hits

The rechits used in training are provided as a list of rechits from CE-E, CE-H, and AHCAL in a sequence. To ensure that the significant improvement in DRN resolution as compared to the WS method is not influenced by the ordering of the rechits (hence learning about the depth in the detector), we infer DRN(E) from Section 5.3 with shuffled order of rechit provided per event and compare the performance of DRN(E) when the ordered rechits are provided. The latter has been used to train the model. Permutation invariance ensures that the model generalizes well to graphs of varying node orderings, making it robust and flexible. It is a fundamental property of GNNs. It ensures the model's output is consistent regardless of how nodes and edges are ordered. However, the DRN architecture is complex and comprises different parts including input, output net, and aggregation step. Hence, we explicitly checked it.
Figure 5.31 presents the energy resolution and response as a function of the incident energy of the pions, comparing scenarios where ordered shuffled rechits are entered into the DRN versus ordered rechits in case of DRN(E), DRN(E,z) and DRN(E,x,y,z). The results remain unchanged, highlighting the fact that DRN is equipped with permutation-invariant quality including the aggregation step with pooling and clustering. We conclude that the improved energy resolution using DRN(E) is unaffected by the input order of rechits. Figure A.14 in Appendix A.1 includes shape com-

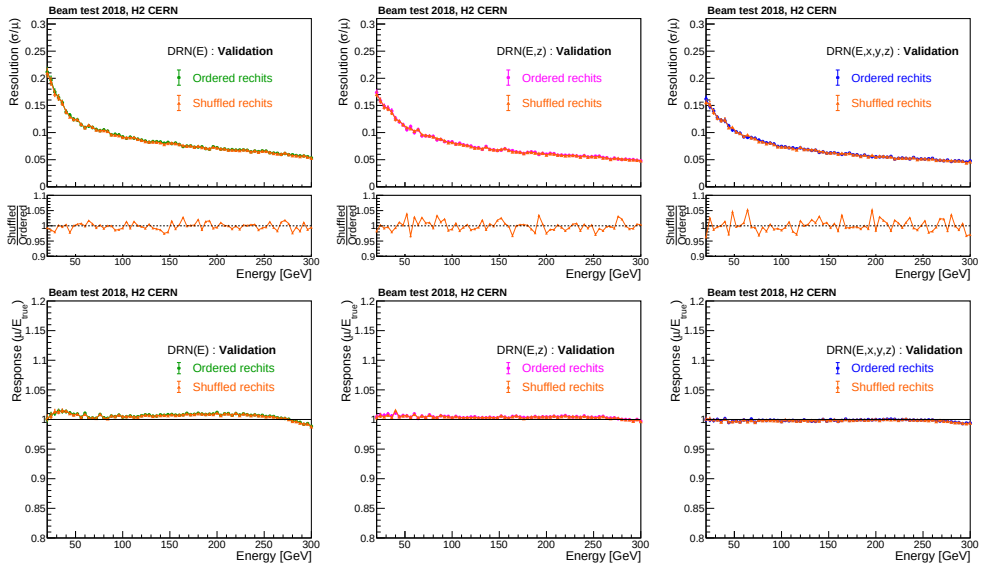


Figure 5.31: Energy resolution (top) and the energy response (bottom) obtained with inferring DRNs on shuffled order of rechits are compared with the DRNs inferred on ordered rechits. The bottom panel shows the ratio of resolutions obtained with inferring on shuffled rechits to that with ordered rechits.

parisons of energy distributions, illustrating DRN's performance in two cases on a validation dataset for 100 and 200 GeV pions.

5.5 DRN performance with rechit energies in MIPs

Rechit energies used for DRN training have been provided in units of GeV obtained by using MIP-to-GeV conversion factors based on 50 GeV positrons and 50 GeV pions for CE-E and CE-H+ AHCAL, respectively, as discussed in section 3.4.3. To asses the extent to which DRN might be biased toward the calibration units for rechit energy, we trained DRNs using rechit energy expressed in MIPs instead of GeV. This approach aims to determine whether the DRN's performance is influenced by the low-level scale calibration of detector cells (rechits) and whether it is susceptible to potential calibration errors.

The first row of Figure 5.32 displays the distributions of rechit energies, expressed in MIPs, for three different sections of the calorimeter: CE-E, CE-H, and AHCAL. The second row shows how these distributions vary with depth in the detector, represented by the z coordinate. Compared with the distributions presented in Fig. 5.5, it is

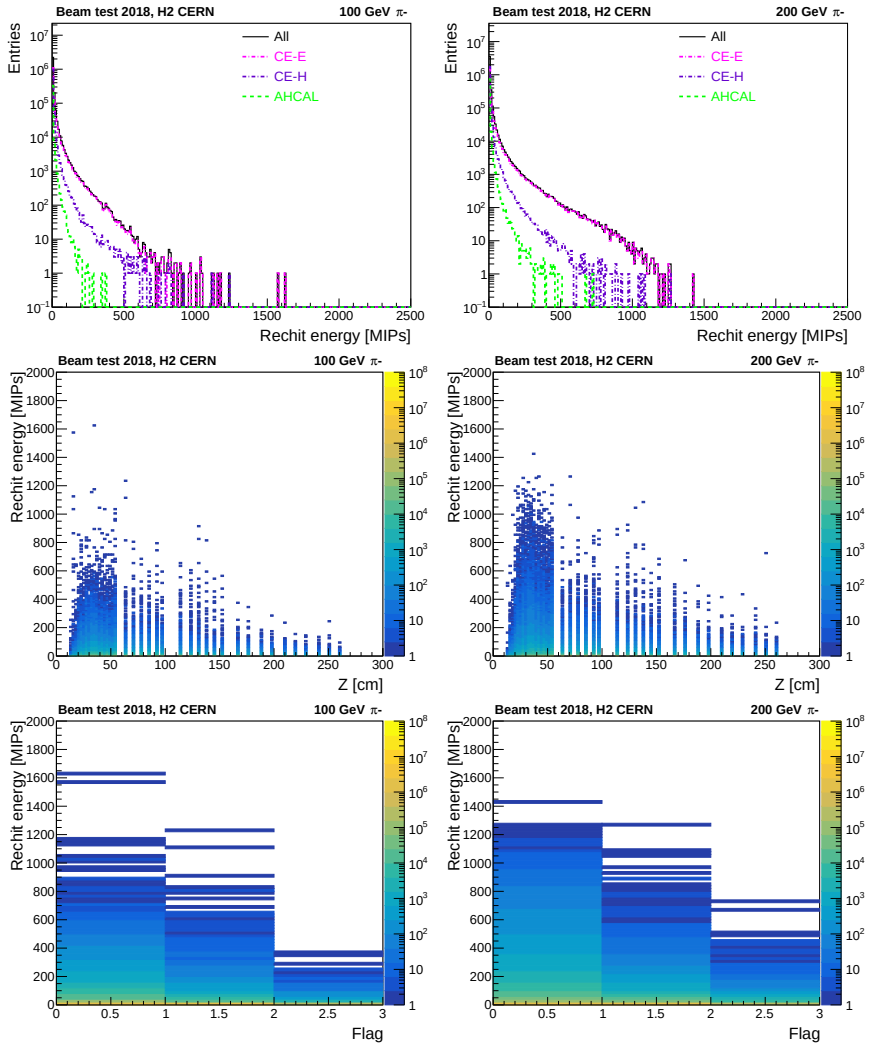


Figure 5.32: Top row: Rechit energy distributions in MIPs, middle row: Rechit energies vs. detector depth, bottom row: Rechit energies with flags indicating location ('0' for CE-E, '1' for CE-H, '2' for AHCAL) for 100 and 200 GeV pions in simulation.

evident that rechit energies distributions expressed in GeV are quite different across the three CE-E, CE-H and AHCAL sections and much lower in CE-E compared to the other two sections. This is also visible from the much larger MIP-to-GeV conversion

in the electromagnetic section.
The CE-E, with its 28 sampling layers within a depth of $\lambda \sim 1.4$ and Pb/Cu/CuW as

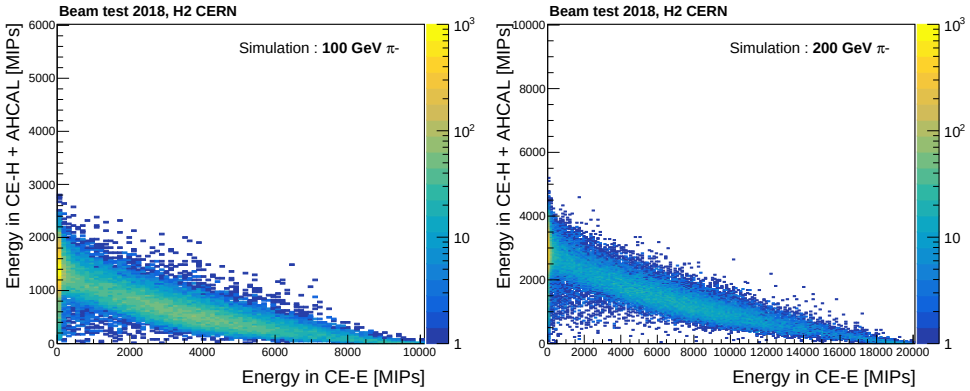


Figure 5.33: Energy measured in electromagnetic and hadronic sections in units of MIPs for 100, and 200 GeV incident pions energies in Simulation.

the absorbers, is more efficient in full-containing electromagnetic showers originating from π^0produced during the first hadronic interaction [102]. Since the probability of the first hadronic interaction occurring in CE-E is around 70% (see Section 5.6.1), a substantial amount of energy in terms of MIPs is recorded in CE-E, as shown in Figure 5.33. However, this does not necessarily translate into a correspondingly high energy deposition in CE-E given it has depth of only 1.4 λ.

To evaluate the DRN's ability to correct for relative miscalibration across different sections of the calorimeter, which vary in absorber materials and thicknesses, we trained three DRNs (as DRN(E), DRN(E,z) and DRN(E,x,y,z)) with the same optimized architecture as described in section 5.3 but rechit energy provided in MIPs instead of GeV. The models were at least trained for 100-150 epochs and their performance was inferred on the epoch with minimum validation loss.

Figures 5.34 show the reconstructed energy using DRN(E), DRN(E,z) and DRN(E, x, y,z) from left to right respectively, for 100 and 200 GeV pions in simulation (validation). These figures compare the shape of reconstructed energy distributions when the DRNs are trained on rechit energy in MIPs versus GeV. When DRNs are presented with spatial information, for example DRN(E,z) and DRN(E,xy,z), nearly identical energy distributions shapes are reproduced with rechit energy is in either MIPs or GeV. In contrast to the above, for DRN trained solely with the energy of rechits, training it with rechit energy is in either MIPs or GeV results in large differences in the overall shape and bulk of the energy distributions.
The top, middle and bottom rows of Figure 5.35 show the energy resolution and response obtained when training on rechit energy in MIPs versus GeV for DRN(E), DRN(E,z), and DRN(E,x,y,z), respectively. The performance of DRN(E,z) and DRN(E, x, y, z) is essentially unchanged when trained on the rechit energy in MIPs versus GeV. This highlights that when DRNs are provided with spatial information for each rechit, the specifics of rechit energy calibration become redundant, resulting in identical performance.
In contrast, the energy resolution and response obtained with DRN(E) are significantly worse when trained on rechit energy in MIPs. This is because, when rechit energy is provided in GeV, DRN(E) can infer information about the three different sections, including their absorbers and sampling fractions, through the MIP-to-GeV

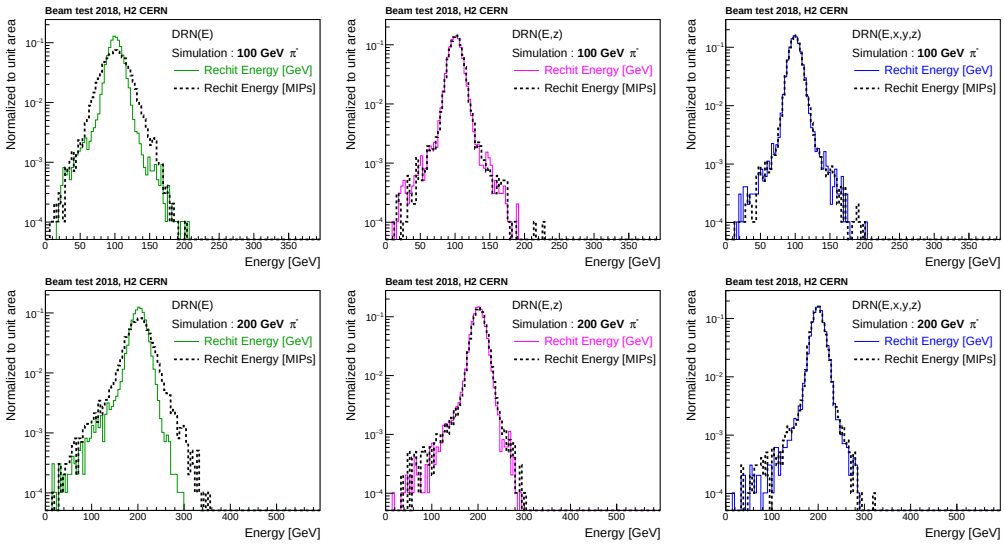


Figure 5.34: Energy distributions reconstructed with DRN(E) (left), DRN(E,z) (middle) & DRN(E,x,y,z) (right) trained on rechit energy in MIPs versus GeV at 100 and 200 GeV pions from the validation dataset.

calibration factors, particularly the large factor for CE-E. However, when energy is presented in MIPs, the DRN is misled by the higher MIP counts in CE-E, which is due entirely to its geometry and π^0showers as explained earlier in this section, leading to poorer overall performance. In contrast, DRN(E,z) and DRN(E,x,y,z) utilize spatial information, allowing the DRN to categorize these MIPs according to the different sections and account for the respective fluctuations and differences.

This also explains why DRN(E) when presented with only rechit energy in GeV, improves energy resolution by more than a factor of 1.5 with respect to the WS method. This improvement is attributed to the ability of DRN to learn encoded information about absorbers and different detector sections to some extent. To explicitly test this hypothesis, we trained a second model, DRN(E[MIPs], flag), using rechit energy in MIPs with an additional flag indicating the section: CE-E, CE-H, or AHCAL, in which the rechit was located, without providing the z-information as in DRN(E,z). The rechit energy distributions after categorize the rechits based on this flag is shown in the third row of Fig. 5.32. Distributions of energy reconstructed using the DRN(E [GeV]), DRN(E [MIPs]) and DRN(E[MIPs],flag) are shown in Fig. 5.36 for 52, 100, 200 and 300 GeV pions in simulation (validation dataset). The overall shape of the reconstructed energy distributions using DRN(E[GeV]) is reproduced by DRN(E[MIPs],flag) with some minor improvements in the tails compared with DRN(E[GeV]).

The resolution curves obtained with these models are shown in Fig. 5.37. When adding the flag that indicates the position of the rechit with respect to different sections, the DRN trained on the rechits in MIPs reproduces similar performance with respect to the DRN(E[GeV]). This demonstrates that the large improvement in resolution with DRN(E[GeV]) over the WS method is attributed to the fact that the information of the different sections details is encoded in the distribution of rechit energies obtained using detector level calibration. This encoded information helps DRN(E) to learn and compensate the event-by-event fluctuations more accurately when presented rechit energy in MIPs. The reason DRN(E[MIPs],flag) model performs slightly better than DRN(E) model is because the information of three different

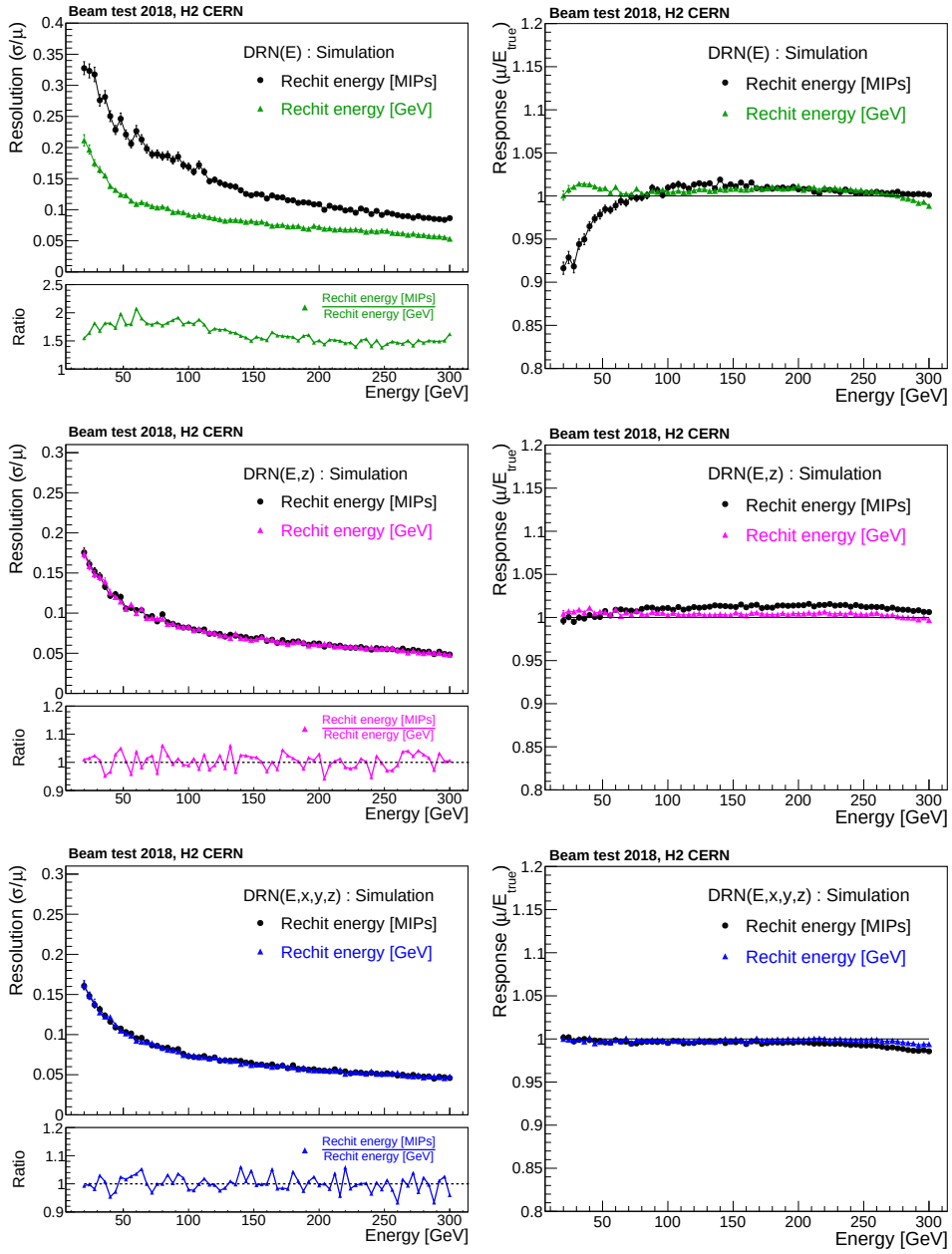


Figure 5.35: The energy resolution (left) and response (right) obtained with DRN(E) (top), DRN(E,z) (middle) and DRN(E,x,y,z) (bottom) trained with rechit energies in GeV versus MIPs are compared in validation dataset (simulation).

sections is provided explicitly to DRN(E[MIPs],flag), whereas DRN(E[GeV] is learning it from the encoded details of rechit energy distribution. In conclusion, DRN has the capability to correct calibration errors, making absolute calibration unnecessary when spatial information of the showers is included during model training.

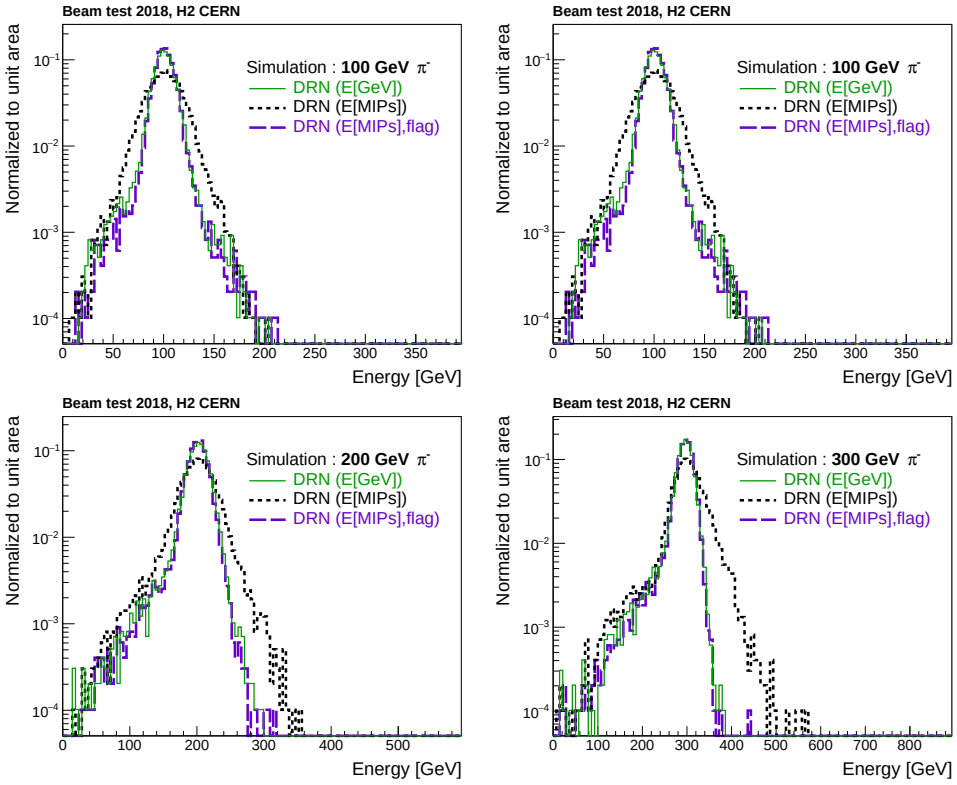


Figure 5.36: The energy reconstructed for 52, 100, 200 and 300 GeV pions using DRN(E[GeV]), DRN(E[MIPs]) and DRN(E[MIPs],flag) in simulation (validation dataset).

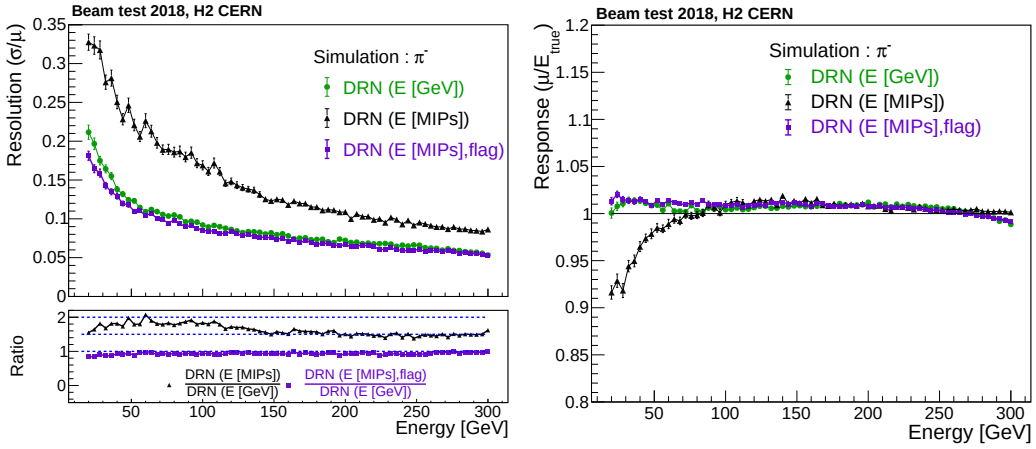


Figure 5.37: Resolution curves obtained with the models DRN(E), DRN(E[MIPs]), DRN(E[MIPs],Flag) are displayed above with validation dataset. By including a flag indicating in which subsection of the calorimeter the rechit belongs to, the resolution obtained with DRN(E) is recovered.

5.6 Performance of DRN on charged pion data collected in beam test 2018 experiments at CERN

We inferred the DRN model described in Section 5.4 trained on a simulated sample of negatively charged pions generated with a uniform momentum spectrum on data collected from a negatively charged pion beam at the CERN SPS in 2018. During the 2018 beam test experiment, approximately 60–100 k event samples were recorded in the setup described in section 3.4.3, for the pions with beam energy of 20, 50, 80, 100, 120, 200, 250 and 300 GeV. These pion beams were derived from the secondary particles produced by bombarding 450 GeV protons from SPS on a target ∼600 m away from the HGCAL prototype. Passing through the beam line elements, there is a finite probability that the original pions can interact with the material on their way. This may result in the production of secondary particles that can also pass through the HGCAL beam test setup and the prototype detector. In addition, some pions can decay in flight with daughter muons entering the detector. Thus, the purity achieved for the pion beam was 80–90% with the remaining consisting mainly of muons, as shown in [102] based on a detailed simulation of the beam line. To select a clean sample of pions that did not interact upstream, several event cleaning selections were used. The details are outlined in Section 5.3. Approximately 40 k−80 k events are available after event selection for further analysis.

The performance of DRNs with data trained on different input features, as described in Section 5.3 is compared with the performance achieved with the simulation data set simulated using the `FTFP_BERT_EMN` physics list. The dataset used for validating the DRN's performance while training 5.3 (validation dataset) are used to compare various quantities in the data and is referred to as "Simulation" in the following. The beam test simulation was done in details except the electronics readout and digitization simulation. As described in [52] and [98], the energy scales were found to be different. The energy scale obtained in the simulation were scaled down by 9.5% for CE-H and AHCAL based on the response of 50 GeV π^-, and by 3. 5% for CE-E based on the response of e^+ according to the studies described in [52] and [98] to match the respective energy scales in data. We have trained the DRN on simulation without correcting for these scales. Hence, the energy scale in the data has been scaled up by the same factors for comparison.

Figures 5.38 and 5.39 show the distributions of the reconstructed energy in the data and the simulation for different energy points using DRN(E,x,y,z), DRN(E,z) and DRN(E), respectively. Overall distributions of energy reconstructed with the DRN trained on simulation are well produced in data.

The event-by-event comparison of the energy reconstructed for 200 GeV pions using the DRNs (X-axis) and WS method (Y-axis) in simulation (left column) and the data (right column) are shown in Figure 5.40. The first row of the figure for DRN(E,x,y,z), second for DRN(E,z), and third row shows the same for DRN(E). Although the bulk of these distributions look very similar in data and simulation, there are differences in tails indicating the differences between data and simulation in terms of fine details of input features.

The Figure 5.41 is comparing the energy reconstructed using different DRNs in data for 200 GeV pions. The first row of the figure compares DRN(E,x,y,z) with DRN(E,z), the second row compares DRN(E,x,y,z) with DRN(E), and the third row compares DRN(E,z) with DRN(E). As in the studies described in the previous sections, the energy distributions shown in Fig. 5.38 are fitted with a Gaussian function iteratively to extract the parameters σ and μ. The change in values of σ and μ is withing 2−3% with respect to those obtained by fitting the Gaussian just once. Energy resolution as a

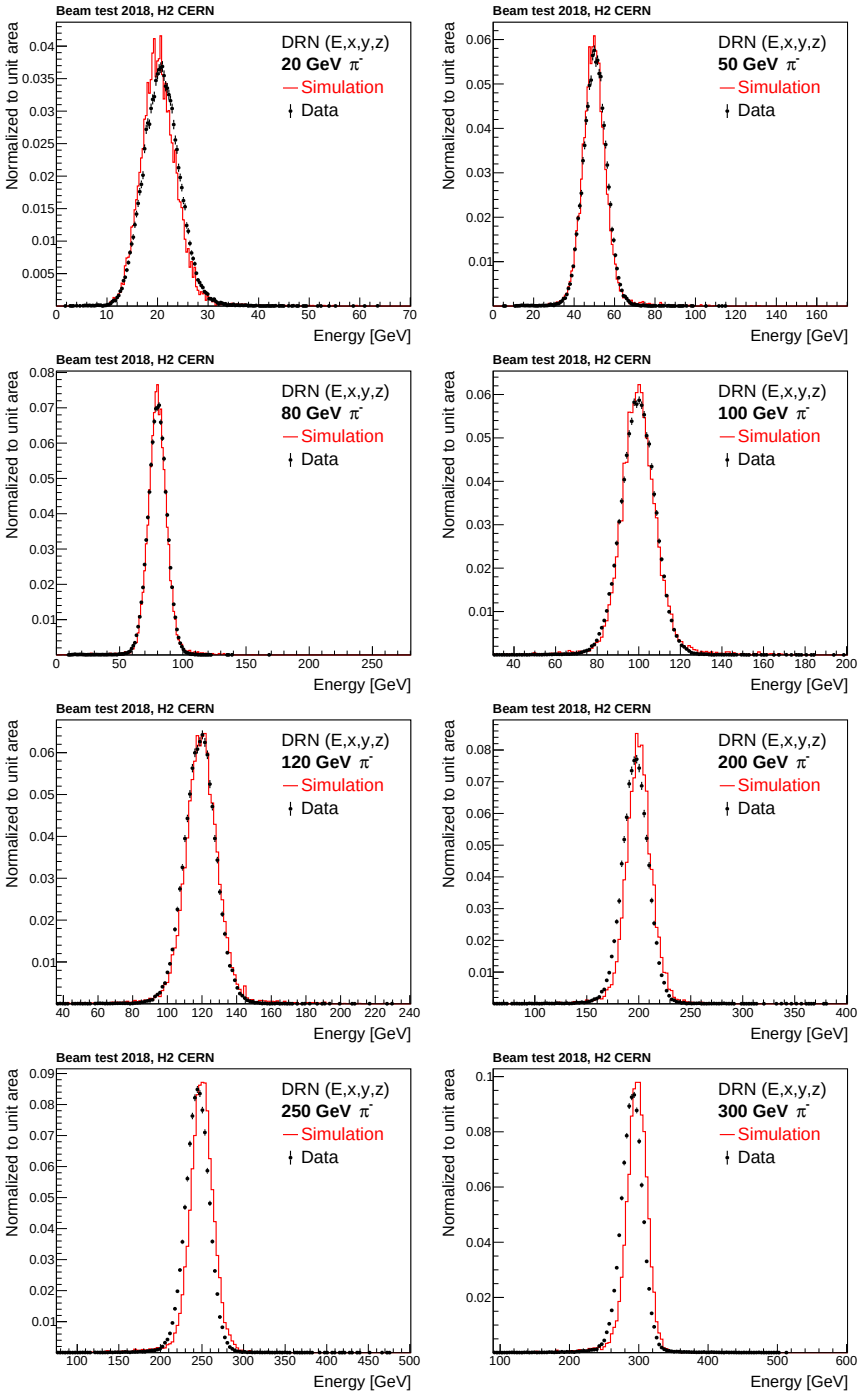


Figure 5.38: Distribution of energy predicted by the DRN(E,x,y,z) on hadron showers generated by pions of true energy 20, 50, 80, 100, 120, 200, 250 and 300 GeV in beam test HGCAL prototype detector and its simulation. The bulk is matching well in simulation and data within few-%.

function of incident pion energy for data and simulation obtained using DRN(E,x,y,z), DRN(E,z) and DRN(E) are presented in the left column of Figure 5.42 top row, middle row and bottom row, respectively. The corresponding response values are shown in the right column of Figure 5.42.

The WS method, which has been used as a reference currently in use by the CMS Collaboration, is identical to the one detailed in [98] and has been reevaluated for the

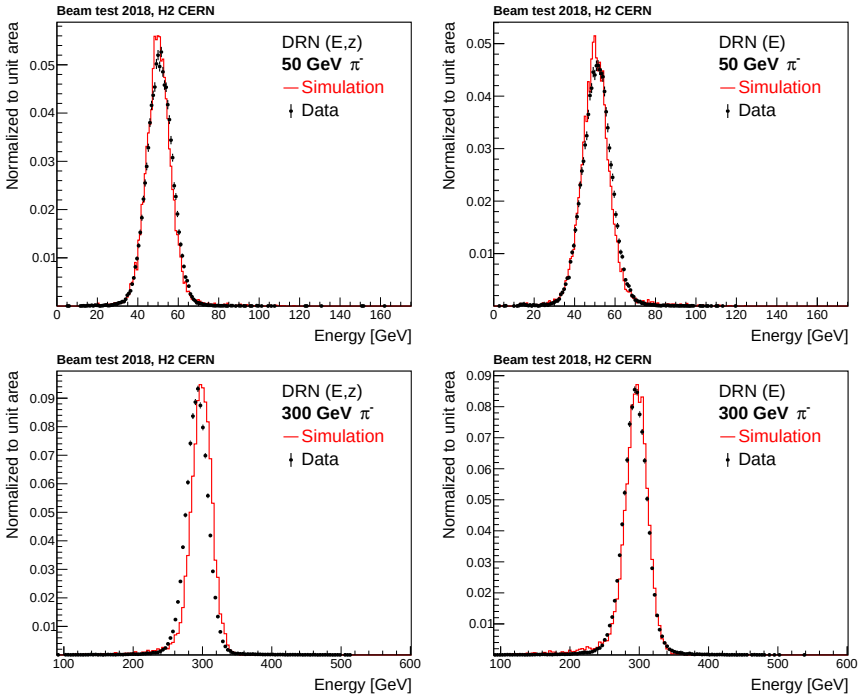


Figure 5.39: Distribution of energy predicted by the DRN(E,z) (left) and DRN(E) (right) model on hadron showers generated by pions of true energy 50 and 300 GeV in beam test HGCAL prototype detector and its simulation.

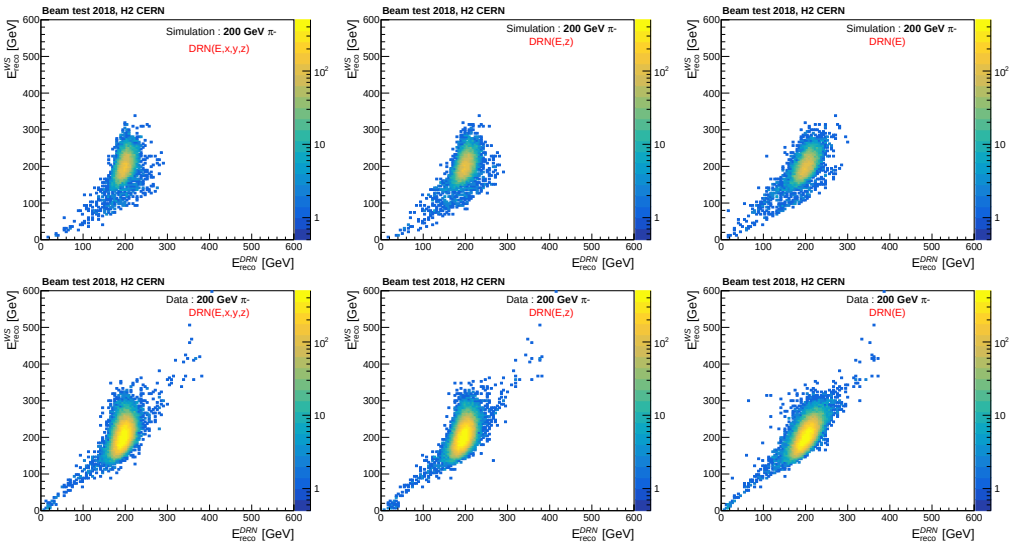


Figure 5.40: The event by event comparisons of energy reconstructed using WS method (Y-axis) and DRN(E,x,y,z), DRN(E,z) and DRN(E) in simulation (top) and data (bottom) for 200 GeV pions is illustrated above in the first, second and third row, respectively.

configuration with the reduced number of AHCAL layers utilized in our studies. The DRN model shows similar performance on charged pion showers collected in the beam test experiments and simulated using GEANT4, and we consistently observe approximately two fold improvement in the resolution of pion showers with DRN reconstruction compared to the WS method in the beam test energy range. We also show an enlarged version of the ratio panels from the left plot of Figures 5.42 in Figure 5.43. This improvement increases with higher incident pion energies, as DRN

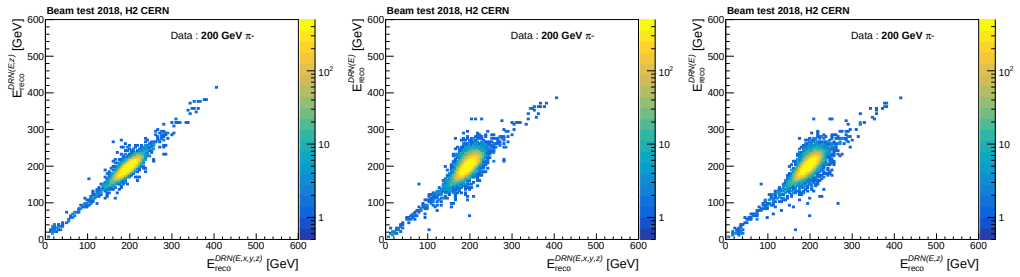


Figure 5.41: Event by event comparisons of the energy reconstructed for 200 GeV pions in data using DRN(E,x,y,z) with DRN(E,z), DRN(E,x,y,z) with DRN(E), and DRN(E,z) with DRN(E) in the first, second and third rows.

seems to benefit from more available information to learn from and compensate for, coupled with a smaller relative impact of transverse leakage at higher energies. The observed improvement in energy resolution obtained using DRN in data and simulation is consistent within a few percent. Specifically, the energy resolution with the DRN improved from about 20% (12%) with the WS method to 10% (6%) for 50 GeV (200 GeV) pions.

The resolution curves obtained for the data and the simulation using DRN and WS methods are fitted with the functional form $\sqrt{S^2/E + C^2}$, where S and C are referred to as stochastic and constant terms of resolution, respectively. We obtained the best stochastic and constant terms of $75.7 \pm 0.4\%$ ($72.3 \pm 0.8\%$) and $2.8 \pm 0.1\%$ ($2.4 \pm 0.1\%$), respectively, for pion showers measured by the HGCAL prototype in the data (simulation) using DRN(E,x,y,z).

Tables 5.3 and 5.4 show comparisons of S and C using DRN(E,x,y,z), DRN(E,z), DRN(E) and WS method.

Method	S(%)	Δ S(%)	C(%)	Δ C(%)
WS	134.9	2.1	7.3	0.2
DRN(E)	91.2	1.0	2.4	0.2
DRN(E,z)	78.1	0.8	2.4	0.2
DRN(E,x,y,z)	72.3	0.8	2.4	0.1

Table 5.3: Stochastic and constant terms and the uncertainties on these using different energy reconstruction method for simulated dataset.

Method	S(%)	Δ S(%)	C(%)	Δ C(%)
WS	132.5	0.9	8.5	0.1
DRN(E)	96.5	0.5	1.6	0.2
DRN(E,z)	84.5	0.4	2	0.1
DRN(E,x,y,z)	75.7	0.4	2.8	0.1

Table 5.4: Stochastic and constant terms and the uncertainties on these using different energy reconstruction method for pion data recorded with beam test setup.

The energy response obtained for the data is also within a few % of unity. This little deviation of the response is due to differences in the details of the longitudinal and transverse development of the shower in the simulation and data. The different energy scales in the data and the simulation have been corrected on average. The simulation has not been tuned to match the fine details in the data. The DRN is presented with granular information event-by-event and is more capable of picking

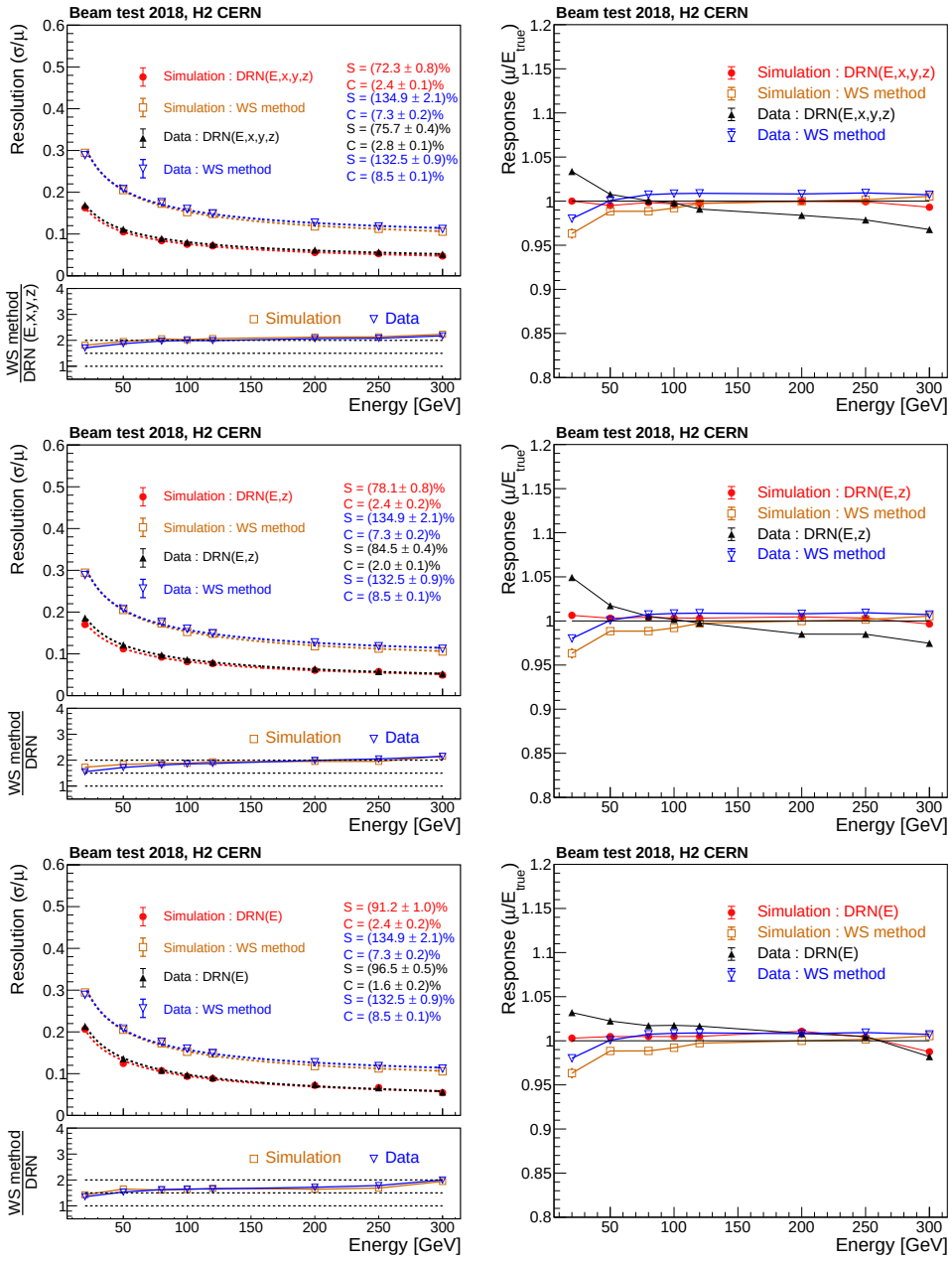


Figure 5.42: Energy resolution (left) and response (right) of pion showers reconstructed in beam test data and simulation using WS method and DRN(E,x,y,z) (top), DRN(E,z) (middle) and DRN(E) (bottom) models. The resolution curve is fitted with the functional form $\sqrt{S^2/E + C^2}$ where S and C are the stochastic and constant term governing the resolution for any calorimeter.

up these differences to these scales. The potential causes and solutions for these discrepancies have not been explored within the scope of this study. Typically, such data and simulation differences are addressed by applying correction factors, which are determined in situ from the data, along with their associated measurement uncertainties.

5.6.1 Event categorization based on pion shower start location

When a pion enters the beam test setup, it travels a certain distance before it goes under the hard interaction with the nucleus of the material and starts the shower.

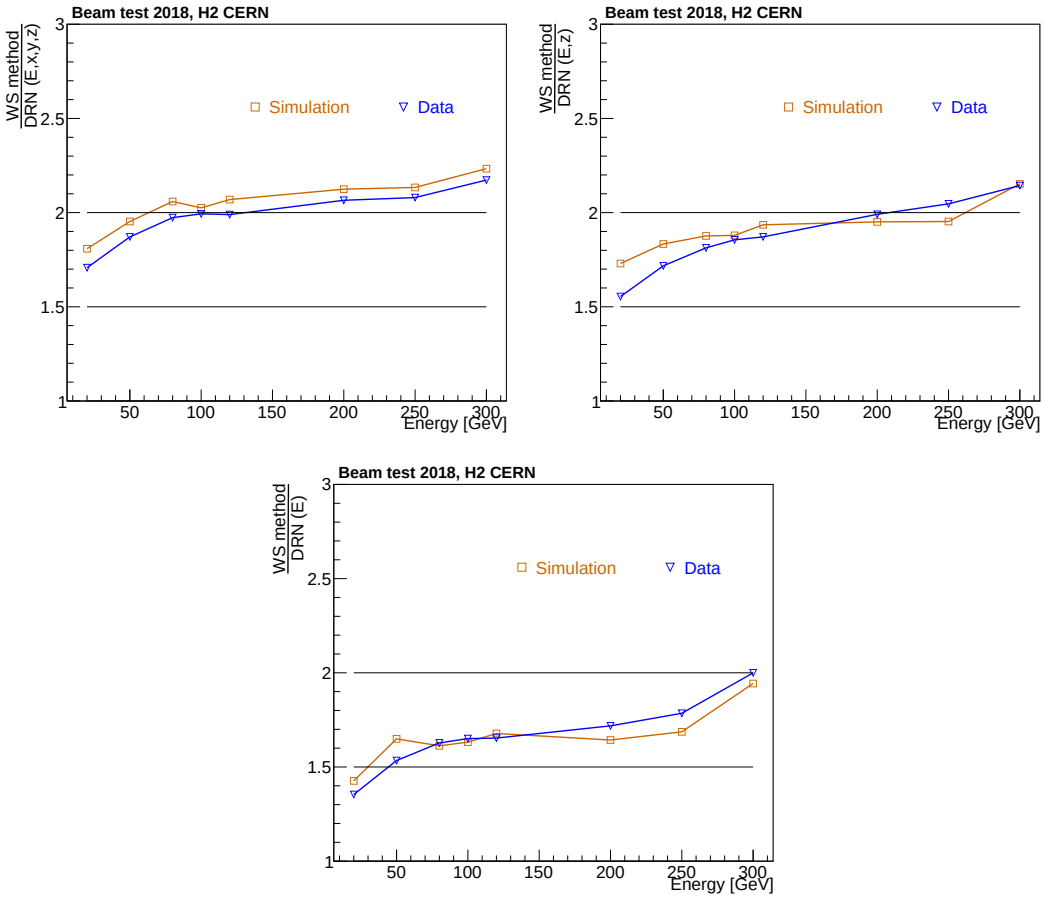


Figure 5.43: The ratio of energy resolutions obtained using different DRNs compared to the WS method in both simulation and data is shown. We observed a twofold improvement in energy resolution when using DRN(E,x,y,z) for pion energy reconstruction.

Until then, pions only interact like MIPs with the material, so very small energy is deposited inside the calorimeters. How much distance a pion will travel before initiating the shower is highly depending upon the λ^{π}_{int} of the material is probabilistic as per the following equation (5.6),

$$N = N_0 \times \exp \frac{-z}{\lambda^{\pi}_{int}} \tag{5.6}$$

where N refers to the number of hadrons which survives i.e. didn't undergo a hadronic interaction, N_0 is the initial number of incident hadrons and λ^{π}_{int} is the pion interaction length of the material. Approximately $\sim$ 70 (75)% of the total events where the pion will start showering in CE-E and the remaining will start shower in the CE-H & beyond as CE-E length is equivalent to 1.19 (1.4) in terms of λ^{π}_{int} (λ_{int}).
We are making use of a shower start finder algorithm as described and developed in Thesis [102] for the beam test setup, developed to achieve optimal reconstruction of hadrons including energy and shower development and better understanding of hadronic showers. The algorithm ascertains the most probable location of a particle shower within the CMS HGCAL by taking advantage of the detector's high granularity and by conducting an analysis of the spatial energy distributions across the calorimeter layers. It employs weighted averaging and iterative refinement to precisely determine the shower's position.

For algorithm optimization, three discriminating variables that are the energy sum within a circle of 10 cm around the center of gravity (COG), the transverse spread measured as R_i, and the number of reconstructed hits in this 10 cm circle. The COG is used and calculated to account for deviations in the track impact position due to scattering.
The transverse energy spread is quantified using a lateral energy ratio, R_i, which helps distinguish showering particles from MIP-like particles, averaged over three consecutive layers and defined as the following,

$$R_i = \frac{\sum_{j=i}^{i+2} En^j_{2cm}}{\sum_{j=i}^{i+2} En^j_{10cm}} \tag{5.7}$$

where the numerator and denominator is the energy summed in a circle of 2 cm and 10 cm, respectively, around the COG in three consecutive layers.
The thresholds for hit multiplicity and R_i are set to greater than 3 and less than 0.96, respectively, and are consistent across all beam energies. However, the energy-deposition threshold is energy-dependent to ensure uniform performance. The thresholds are optimized and extracted by comparing the energy deposited by muons as shown in Fig. 5.44 and taken from the thesis [102]. The shower start (SS) algorithm

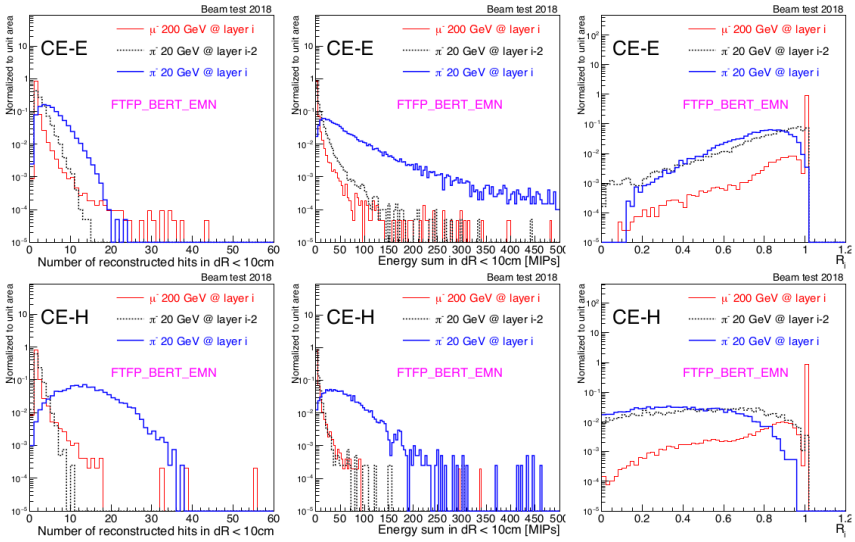


Figure 5.44: The left column shows reconstructed hits, the middle displays energy comparisons, and the right illustrates lateral energy ratios. The top row refers to CE-E, and the bottom to CE-H, contrasting muons with 20 GeV pions at the start layer of the shower and the one before [102].

is found to be $> 90\%$ (95%) efficient in correctly identifying the shower start location within 2(1) layers in CE-E (CE-H). Events where the algorithm does not find the SS location are rejected.

The events are categorized into the following categories based on the location of shower start -

- CE-E pions: pions that start showering in any of the CE-E layers.
- CE-H pions: pions that start showering in any of the CE-H layers.
- All pions: pions that start showering in any of the HGCAL layers.

These categories are motivated by the use of different absorbers: CE-E uses Pb/Cu/CuW, while CE-H and AHCAL use steel, leading to variable e/h [102]. The CE-E pions encounter multiple absorbers, resulting in additional fluctuations. It is crucial to verify whether DRN is learning these absorber details and showering categories.

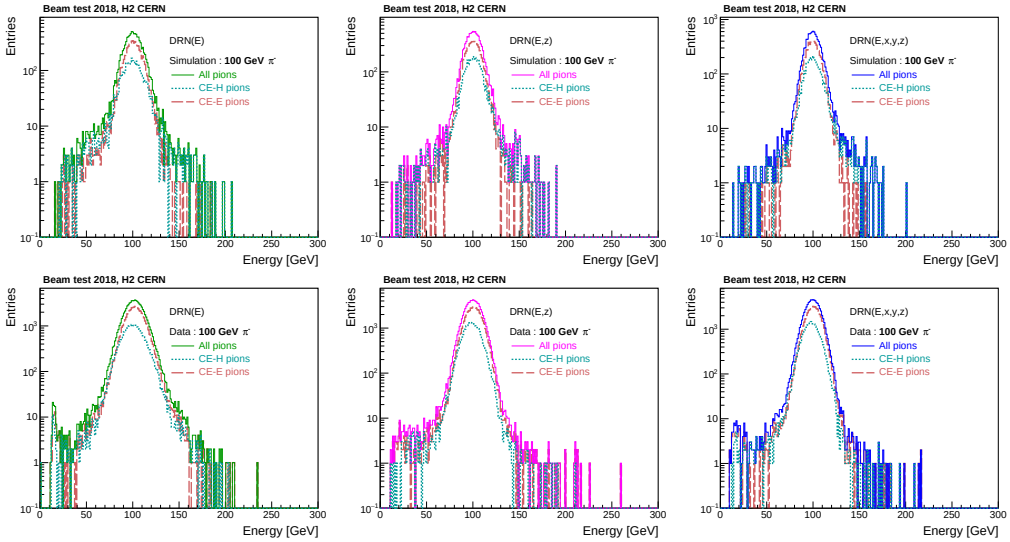


Figure 5.45: The energy distributions for different event categories are compared reconstructed using the DRN(E) on Simulation (left column) and Data (right column) for 100 GeV pions. "CE-H pions" have more long tails.

The reconstructed energy distributions for 100 GeV pions using DRN(E,x,y,z), DRN(E,z) and DRN(E) for simulation (top) and data (bottom) are compared in different event categories devised based on the start layer of the shower for different pion energies and are shown in Figures 5.45 from right to left, respectively. Note that the simulation is represented by the same data set that was used to validate the performance of DRN while training as described in Section 5.3.

The performance on "All pions" represents the inclusive sample of both CE-E and CE-H pions. The bulk of distributions in three categories is very similar. Notably, the performance for "CE-H pions" exhibits more pronounced high and low E_{reco} tails compared to those of "CE-E pions". This difference arises because events where the pion begins showering later in the calorimeter tend to experience greater longitudinal leakage, especially at higher energies. In contrast, pions that begin to shower in the early layers encounter a higher calorimeter depth, reducing the extent of leakage.
Comparisons of energy resolution and response in different categories for data (simulation) using DRN(E), DRN(E,z) and DRN(E,x,y,z) are shown in figure 5.46 (Figure A.16). The energy resolution for CE-E pions is worse than that of the inclusive category by ∼10% in contrast to CE-E pions, which show better resolution than the inclusive ones. The energy resolution for CE-H pions is better than that of CE-E pions for higher energies. This improvement is due to the greater longitudinal shower containment for CE-E pions, since they traverse approximately 1.4λ more of the detector compared to CE-H pions, and also the fact that CE-E is optimized to measure EM showers & effectively captures early produced π^0's. For DRN(E,x,y,z), the stochastic term derived for these three categories is comparable as given in legends of Fig. 5.46, although the constant terms exhibit some differences. In contrast, these stochastic terms differ by ∼10% with the WS method [102]. This highlights an

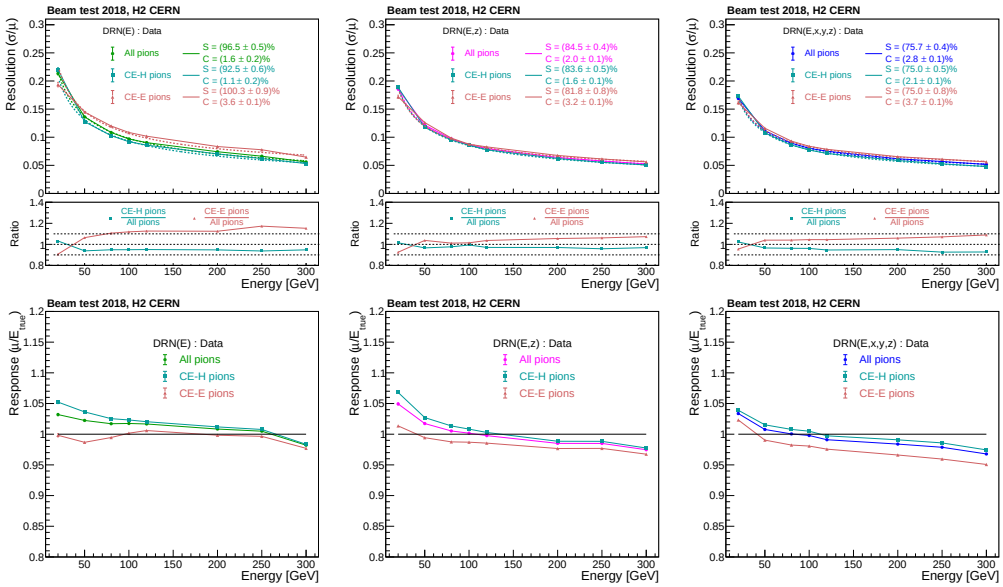


Figure 5.46: The energy resolution (top) and response (bottom) for different event categories is compared and shown above for data using DRN(E), DRN(E,z) and DRN(E,x,y,z) from top to bottom row, respectively. The ratio panel shows the ratio of resolution for "CE-E pions" and "CE-H pions" categories to "All pions" category.

important aspect of DRN's capability to discern distinct event populations within the pion event samples.

5.7 Understanding improvements in energy resolution obtained using DRN

The high fluctuations in hadronic showers as discussed in Section 3.4.3, contribute to dramatic variation in the spatial patterns of energy deposits in the detector and results in the showers initiated by the pions of the same true incident energy very different from each other. Variations in the EM component manifest differently in events, highlighting intrinsic fluctuations. Such non-Gaussian fluctuations are the origin of dramatic presence of hadronic showers, and each event looks very different from the other even with the same true incident energy of pions as shown in figure 5.9.

The substantial improvement in energy resolution using the DRN is likely due to its ability to compensate for event-to-event fluctuations in hadronic showers using the 3D images of showers made possible with a fine granularity of the HGCAL detector. In order to understand which aspects of the shower development DRN is able to compensate for, we performed several studies based on the GEANT4 truth information. In this section, we summarize the following studies :

- Effect of total energy deposited in the absorbers on energy predicted by DRN and WS.
- Effect of total energy leakage from the detector on the energy predicted by DRN and WS.
- Energy carried by π^0or the total EM component and the energy predicted by DRN and WS.

5.7.1 Assessing DRN performance based on the energy deposited in the absorbers

In a sampling calorimeter, while the shower primarily develops and deposits energy in passive material or absorbers, only a small fraction of the energy deposited in active materials is read out as a measurable signal. The signal recorded in the active materials samples the progress of a shower and reflects the energy deposited in the absorbers. Generally only $1-2\%$ of total energy deposited is recorded in active materials, and fluctuations in the sampling fraction contribute a large amount to the worsening of resolution [45].

In the detector geometries in which the shower is fully contained in transverse and longitudinal directions, the energy deposited in the absorbers is affected by variations in the nuclear processes & particles produced, especially the fluctuations in the EM component of the pion showers, in the course of the evolution of the shower. The true incident pion energy can be written as

$$E_{true} = E_{abs} + E_{active} + E_{invisible} \tag{5.8}$$

where E_{abs} is total energy deposited in the absorbers, E_{active} is the energy deposited in the active layers and $E_{invisible}$ is the total energy used up to break up the nuclei.

The HGCAL beam test prototype detector is not large enough to fully contain pion showers in the transverse and longitudinal directions. A certain amount of pion showers will develop outside the detector and cannot be compensated by inferring the signal generated in the active material. Layers equipped with only one silicon module contribute more to the energy leaked in transverse leakage. For example, the CE-E section and the last three layers of the CE-H sections are composed of 1 modules only. This specific portion of the shower is called the energy of the shower that leaked out of the detector. This potential of leaked energy in the transverse and

longitudinal direction essentially add more to the intrinsic fluctuations in the pion showers heightening complexity. Hence, the definition of E_{true} as in equation(5.8) is updated as:

$$E_{true} = E_{abs} + E_{leak} + E_{active} + E_{invisible} \tag{5.9}$$

where E_{leak} on the right hand side of equation (5.9) is the total energy is leaked out of the HGCAL prototype detector used in beam test experiment. The visible component of pion showers, the recordable energy, are redefined as,

$$E_{visible} = E_{abs} + E_{leak} + E_{active} \tag{5.10}$$

To obtain additional information, such as the energy deposited in the absorbers, the number of secondary π^0throughout the showers, and the energy carried by these secondaries, we utilize an inbuilt method of GEANT4 called *simWatcher*. For the studies we conducted to understand what DRN is learning by checking the reconstructed energy as a function of various truth information, we generated ∼45k simulated events for pions with energies (20, 50, 80, 100, 120, 200, 250, 300) using the 33 m beam test simulation setup as described in Section 3.4.3. The showers are modeled using the `FTFP_BERT_EMN` physics list which has also been used to train the model. The total energy deposited in the absorbers is obtained after summing up all the recorded `simHits` across the passive material of the three subsections (CE-E, CE-H and AHCAL) of the beam test setup as defined in the equation 5.11.

$$E_{abs} = \sum_{i}^{N_{passive}^{CE-E}} E_i^{\texttt{simHits}} + \sum_{i}^{N_{passive}^{CE-H}} E_i^{\texttt{simHits}} + \sum_{i}^{N_{passive}^{AHCAL}} E_i^{\texttt{simHits}} \tag{5.11}$$

On the right hand side of the above equation, $N_{passive}^{CE-E}$, $N_{passive}^{CE-H}$ and $N_{passive}^{AHCAL}$ are the total number of `simHits` generated in the passive materials of CE-E and CE-H and AHCAL sections of beam test setup whereas, E_i is the energy of the i^{th} `simHit` in the same event.

If a step in *G4Track* is the last step within a defined physical volume and does not proceed to the next physical volume or go through any corresponding post-step, it is considered to be outside the physical geometry and the corresponding `simHits` is referred to as the energy leaked out of the detector. The energy leaked out of the specific section can be calculated according to equation 5.12, where E_{leak}^{A} is calculated for each section represented by A and summing up all the `simHits` outside of the "A" section.

$$E_{leak}^{A} = \sum_{i}^{N_A} E_i^{\texttt{simHits}} \tag{5.12}$$

where, A can be CE-E or CE-H or AHCAL or outside the setup in the above equation. Thus, the total energy leaked out of the beam test setup can be written as in terms of mutually exclusive components as per 5.13,

$$E_{leak} = E_{leak}^{CE-E} + E_{leak}^{CE-H} + E_{leak}^{AHCAL} + E_{leak}^{longi} + E_{leak}^{<CE-E} \tag{5.13}$$

where,

- E_{leak}^{CE-E} : The energy leaked out of the CE-E in the transverse direction and energy was lost in the space between the CE-E and CE-H sections.
- E_{leak}^{CE-H} : The energy leaked out of the CE-H in the transverse direction and was lost in the space between the CE-H and AHCAL sections.

- E_{leak}^{AHCAL} : The energy leaked out of the AHCAL in the transverse direction.
- E_{leak}^{longi} : The energy leaked outside the AHCAL in the longitudinal direction.
- $E_{leak}^{<CE-E}$: The energy lost before entering into the HGCAL.

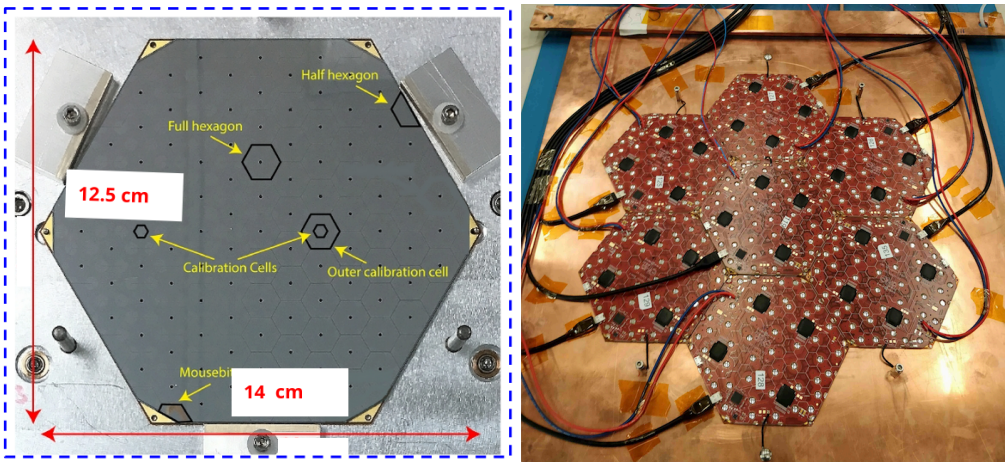


Figure 5.47: A layer of CE-E (right) and CE-H (left) from the beam test setup, highlighting the sizes of the absorber and active plates.

The true absorber energy is defined as the total energy deposited inside the absorber plate region that overlaps with the active material and everything outside this overlapped region is called leakage even though `simHits` were recorded in the absorber plates. The defined absorber energy (E_{abs}) in equation 5.11 and the leakage energy (E_{leak}) in equation 5.13 are correct for the calorimeters where the size of the active material and passive material are the same and the alignment is perfect. But this is not the case with the beam test setup we are studying here. The size of the absorber plates and active material for different compartment is as following:

- For CE-E :
  - Active layer radius : 6.18 cm
  - Absorber layer radius : 8 cm
- For CE-H :
  - Active layer radius : 18.5 cm
  - Absorber layer radius : 19.5 cm
- For AHCAL :
  - Absorber layer size (height $\times$ width) : (72 cm $\times$ 72 cm)
  - Active layer size (height $\times$ width) : (72 cm $\times$ 72 cm)

The above differences are illustrated for CE-E and CE-H in the figure 5.47. Thus, the absorber calculated using equation 5.11 has a small certain fraction that should be counted as the leaked energy as per equation 5.13. Since the sizes of the absorber and active layers are not significantly different, E_{abs} and E_{leak} provide a reasonable approximation of the true energy absorbed and leakage, respectively. The E_{abs} and E_{leak} defined in this manner are mutually exclusive and effectively capture the intrinsic shower-to-shower fluctuations even though these are not precisely calculated. For the studies presented in this section, we will use the approximate total absorber energy and total energy leakage, indicated as E_{abs} and E_{leak} and calculated using equation 5.11 and equation 5.13, respectively.

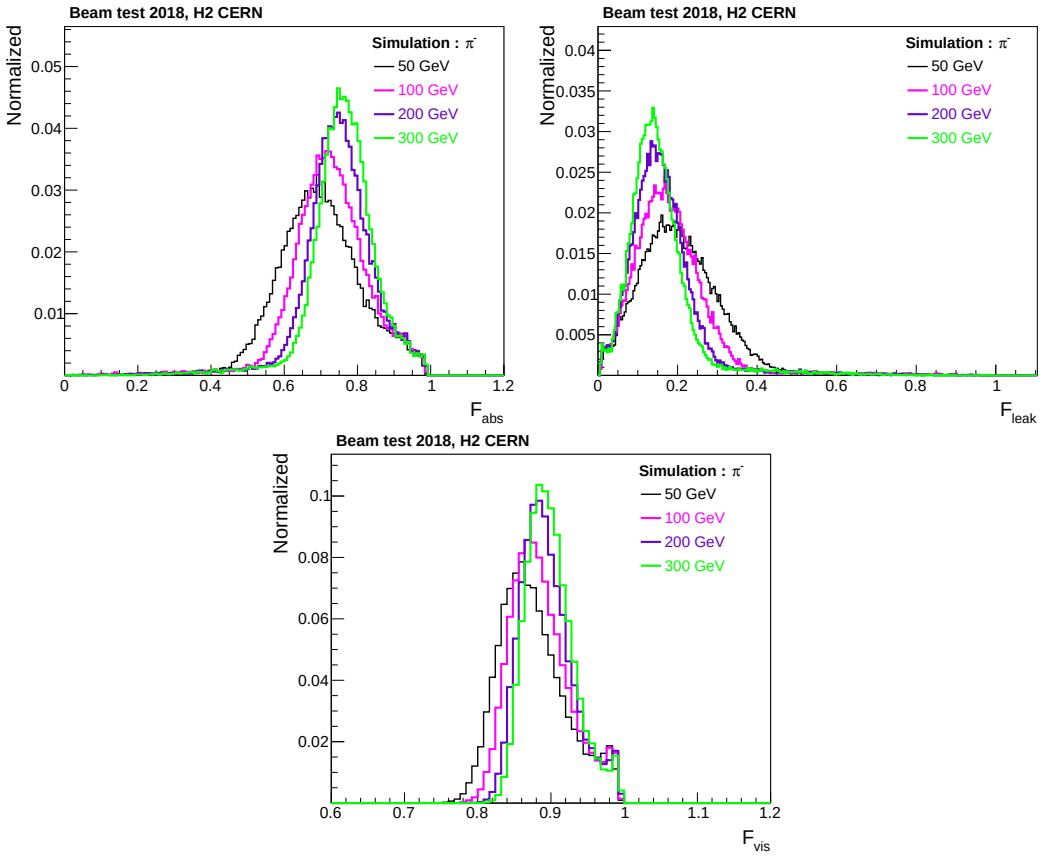


Figure 5.48: F_{abs}, F_{leak} and F_{visi} for incident pion energies of (50, 100, 200, 300) GeV. These are calculated as per the equations 5.14, 5.15 and 5.16 respectively with the help of additional truth information added in the simulation.

We calculate the total absorber energy fraction (F_{abs}), the total leaked energy fraction (F_{leak}) and the total visible energy fraction (F_{vis}) as defined in the equations 5.14, 5.15 and 5.16 respectively after putting in the values for E_{abs} and E_{leak}.

$$F_{abs} = \frac{E_{abs}}{E^{\pi^-}_{true}} \tag{5.14}$$

$$F_{leak} = \frac{E_{leak}}{E^{\pi^-}_{true}} \tag{5.15}$$

$$F_{vis} = F_{abs} + F_{leak} \tag{5.16}$$

The F_{abs}, F_{leak} and F_{vis} for a few incident pion energies are shown in figure 5.48 from left to right, respectively. The mean of F_{abs}(and F_{vis}) increasing with energy is highlighting the non-linear energy response for pion showers. The nonzero variance in F_{abs}, F_{leak} and F_{visi} is essentially the result of fluctuations in the EM component of pion showers. The value of F_{vis} is always < 1 demonstrating the nonzero fraction being used to break up the atomic nuclei. On average, F_{leak} decreases with increased incident pion energy. This reduction is due to the decreased transverse spread of showers at higher incident energies, allowing the secondary particles to penetrate deeper into the calorimeter.

The distributions of different components of F_{leak} are shown in Figure 5.49 (left). As CE-E and CE-H have lower transverse coverage than the expected transverse

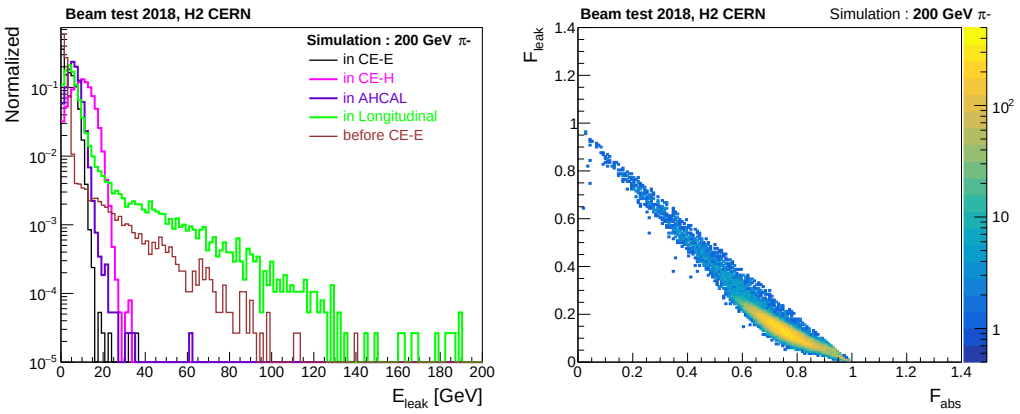


Figure 5.49: The energy leaked out of the different sections of beam test setup for different incident pion energies (left), F_{abs} vs F_{leak} (right) for 200 GeV pions are shown. See text for F_{abs} and F_{leak} definition.

spread of hadron showers, E_{leak}^{CE-E} and E_{leak}^{CE-H} show leakage of energy for most of the events and not too many tails as the shower is developing. In contrast, the energy leaked in the longitudinal direction, beyond the AHCAL features more pronounced high-end tails but much smaller mean values consistent with expectations from later starting showers. With an increase in incident energy, due to the more penetrating pion showers, there is a greater leakage of energy in the longitudinal direction as compared to the CE-E and CE-H sections.

The figure 5.49 (right) illustrates the variation of F_{abs} and F_{leak} as a function of different incident pion energies, and these show the expected anticorrelation. In an ideal geometry with no leakage, the spread on the y axis will not be present, and the pion showers recorded in the beam test will only have intrinsic shower fluctuations. Mostly events are with F_{abs} up to ∼60-80% and we have the tail of events with very high leakage up to 40%.

The energy reconstructed using the WS method and DRN(E,x,y,z) as a function of E_{abs} is shown in figure 5.50. The bulk of E_{reco} using the WS method varies linearly as a function of E_{abs} for different incident pion energies. The WS method is entirely based on the total reconstructed energy using fixed weights (E_{fix}) in the CE-E, CE-H, and AHCAL sections. When applied to the rechits recorded in the active material across different sections, these fixed weights reflect the shower energy deposited in the absorbers. As a result, the event-by-event differences intrinsic to the hadronic shower evolution are not accounted for by the WS method. The energy reconstructed for the events with negligible energy deposited in the absorbers is a long way off from the true energy in the case of the WS method. These are events in which the EM component in the hadron showers fluctuated very low, the leakage was large, or the incident particle did not cross the full detector.

On the other hand, DRN correctly predicts the energy even with extreme high and low end fluctuations in E_{abs}. The bulk of the total reconstructed energy is close to the true energy and does not appear to vary with respect to E_{abs}. This performance is largely attributed to the fact that DRN has access to highly granular information for each event, enabling it to learn and adapt to the complexity of different intrinsic shower fluctuations.

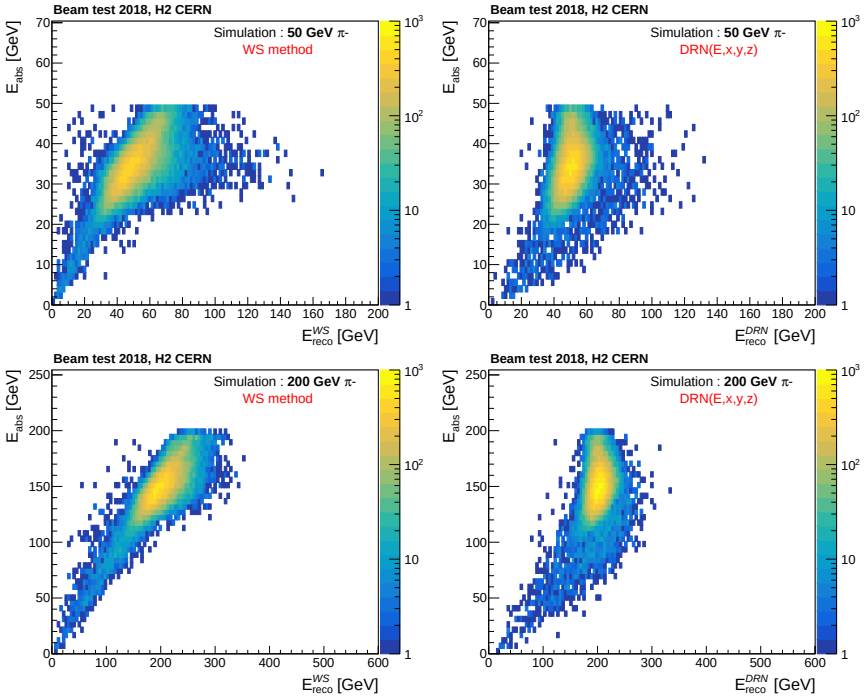


Figure 5.50: Reconstructed energies using WS method (left) and DRN(E,x,y,z) (right), as a function of E_{abs} for 50 (top) and 200 (bottom) GeV pions. The E_{reco} using WS method is showing a linear dependence on the E_{abs} and DRN is able to compensate and reconstruct correctly even for the smaller values of E_{abs}.

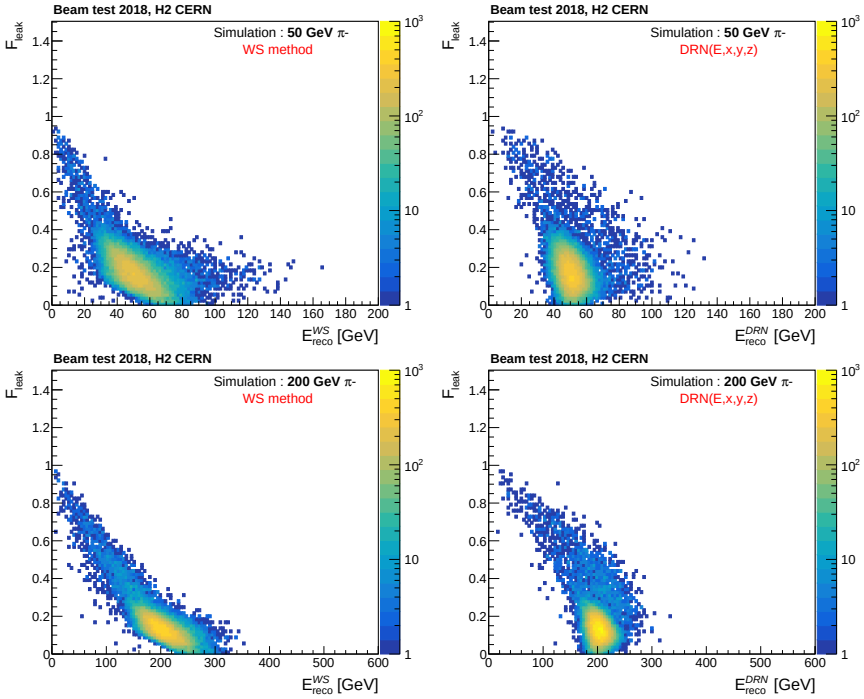


Figure 5.51: Reconstructed energy using WS method (left) and DRN(E,x,y,z) (right), as a function of F_{leak} for 50 and 200 GeV pions. As F_{leak} increases, using WS method results in E_{reco} that reduces, and diverges from the true incident pion energy. Conversely, the bulk of E_{reco} distribution with DRN remains stable relative to F_{leak}.

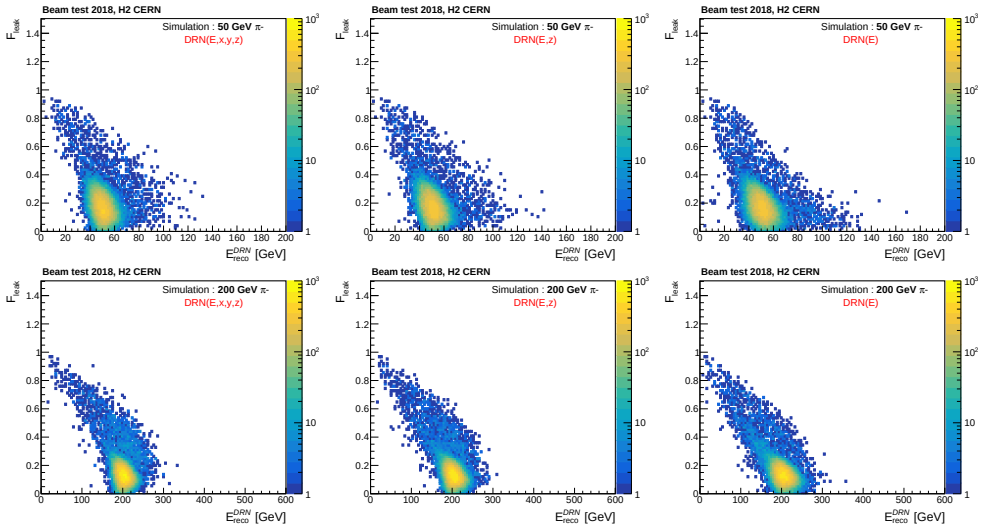


Figure 5.52: Reconstructed energy as a function of F_{leak} for DRN(E,x,y,z), DRN(E,z) and DRN(E) from left to right, respectively, for 50 and 200 GeV pions. DRN(E) struggles with energy prediction for high F_{leak} due to lack of spatial data in contrast to DRN(E,x,y,z) and DRN(E,z) utilizing spatial information.

Figure 5.51 shows how E_{reco} using the WS method (left) and the DRN method (right) vary as a function of F_{leak}. The WS method lacks spatial information and is effective in relieving the total energy deposited in the difference sections. Events with high F_{leak} are also events with very low E_{abs} as shown in Figure 5.49 (right). Thus, the WS method predicts E_{reco} far away from the truth energy. In contrast, DRN integrates spatial information that helps it to learn and extrapolate the leaked energy out of the detector. DRN seems to be differentiating events with high and low F_{leak} and accurately compensating for it. DRN also predicts E_{reco} deviation from the true energy for approximately 1-2% fraction of events with extreme F_{leak}. These events are with very small number of rechits recorded in the beam test setup, thus DRN does not have enough information available to learn and compensate for these events. E_{reco} using DRN(E,x,y,z), DRN(E,z) and DRN(E) varying with F_{leak} for incident energy of 50 and 200 GeV pions are shown in figure 5.52 from left to right, respectively. High F_{leak}, less energy is deposited in the absorbers and active material. DRN(E), lacking spatial information, struggles to accurately predict the energy of events with high F_{leak}. In contrast, DRN(E,x,y,z) and DRN(E,z) incorporate spatial information, leading to better performance overall and especially in regions with very high F_{leak}.

To further quantified the DRN's performance evaluation, the events are categorized into the following four different categories designed based on F_{abs} as defined in equation 5.14,

- F_{abs} within $\mu \pm \sigma$
- $F_{abs} > \mu + \sigma$
- $F_{abs} < \mu - \sigma$ and $> \mu - 2\sigma$
- $F_{abs} < \mu - 2\sigma$

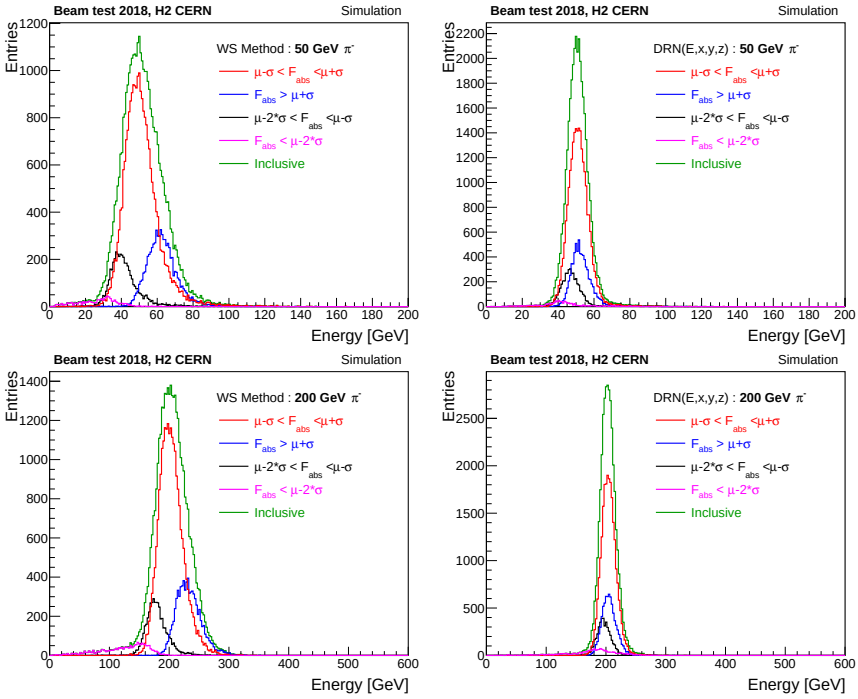


Figure 5.53: The reconstructed energy using DRN(E,x,y,z) (right) and WS method (left) is compared for different event categories defined based on F_{abs} in the above plots for 50 and 200 GeV pions. "Inclusive" line represents the energy predictions for all events.

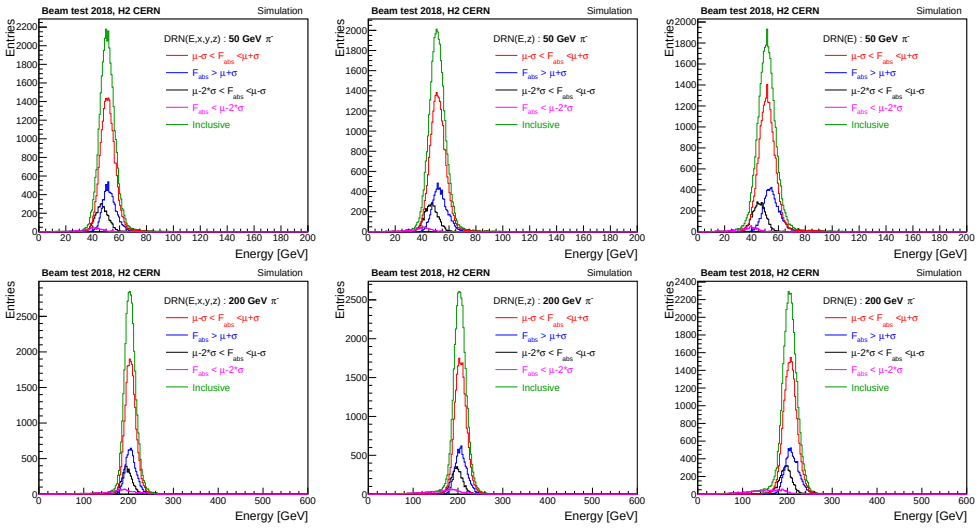


Figure 5.54: The reconstructed energy using DRN(E,x,y,z), DRN(E,z) and DRN(E) in the right most, middle and left columns, respectively, is compared for different event categories defined based on F_{abs} in the above plots for 50 and 200 GeV pions. "Inclusive" line represents the energy predictions for all events.

The figure 5.53 shows E_{reco} for different event categories devised based on F_{abs} including the plot for all the events referred to as "inclusive", for the DRN and WS method. In the 1σ bins around the mean fractional absorber energy, the DRN predictions exhibit significantly less variation compared to the WS method. These distributions clearly demonstrate how the DRN positions events with drastically different fluctuation levels much closer to the true value, whereas the WS method remains more sensitive to these fluctuations. The bins with the lowest F_{abs}, primarily contribute to the tails in the DRN case. The figure 5.54 presents the comparison of DRN(E,x,y,z), DRN(E,z), and DRN(E,z) from left to right, respectively. As more input features are provided for the DRN to learn from, it becomes increasingly capable of accurately positioning events with extreme fluctuations in F_{abs}.In summary, the DRN demonstrates the ability to capture the intricate details of the shower when provided with granular spatial information. Using transverse information on the development of the shower, the DRN attempts to estimate the fraction of energy lost, which will be more clearly illustrated in the subsequent analysis.

5.7.2 Compensating the fluctuations in EM component

To understand the details about the EM component of hadronic showers and how DRN is able to compensate for it, we have extracted the information of all the secondaries created in hard nuclear interactions (including *HadInelastic, HadCapture, HadFission, HadAtRest* processes and not *hIoni* processes in GEANT4 framework) using the GEANT4's *simWatcher* method. Information on how many secondaries are produced with their respective kinetic energies, position coordinates, and particle IDs is saved. We then consider the neutral pions created in secondary interactions throughout the shower developed, including those produced as the decay product of η^0. The total fraction of energy carried by $\pi^0$is calculated for each event as defined in equation 5.17 and is referred to as the total electromagnetic component of the hadronic shower per event.

$$F_{\pi^0} = \frac{\sum_i E^i_{\pi^0}}{E^{\pi^-}_{true}} \tag{5.17}$$

On the right hand side of the above equation, $E^i_{\pi^0}$ is the energy carried by i^{th} $\pi^0$and is summed over all the π^0produced throughout the hard interaction in a given pion showers for an event. The $E^{\pi^-}_{true}$ is the true energy of the pion. Figure 5.55 shows the energy fraction (F_{π^0}) carried by π^0in pion showers of different energies, highlighting event-to-event fluctuations in pion showers with the same energy. The general increase in F_{π^0} with energy reveals the root cause of the nonlinear calorimeter response to pion showers.

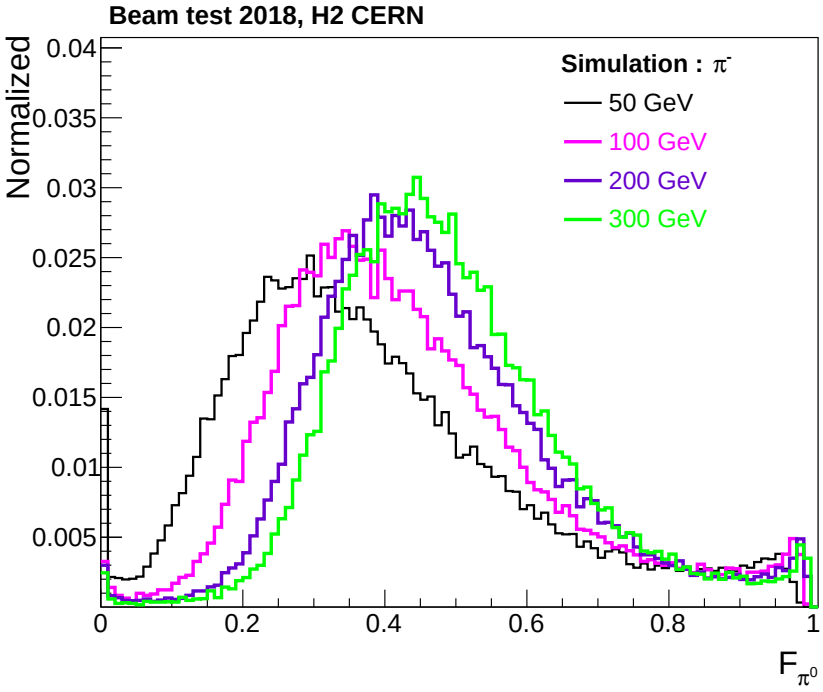


Figure 5.55: Distribution of fraction of EM component defined as F_{π^0} in hadron showers of different energy pions. This has been calculated using beam test simulation with *FTFP_BERT_EMN* physics list.

The left column of Figure 5.56 shows the event-by-event variation in energy reconstructed using the DRN and WS methods as a function of F_{π^0} for different pion energies. In the right column, the average profile of E_{reco} is plotted against F_{π^0}. The energy reconstruction of the WS method tends to fluctuate along with variations in

F_{π^0}, while the DRN method remains relatively stable, even in the presence of extreme fluctuations in F_{π^0}. Similar observation has been found for the other energy points as shown in figures A.17 and A.18 in the Appendix A.1.

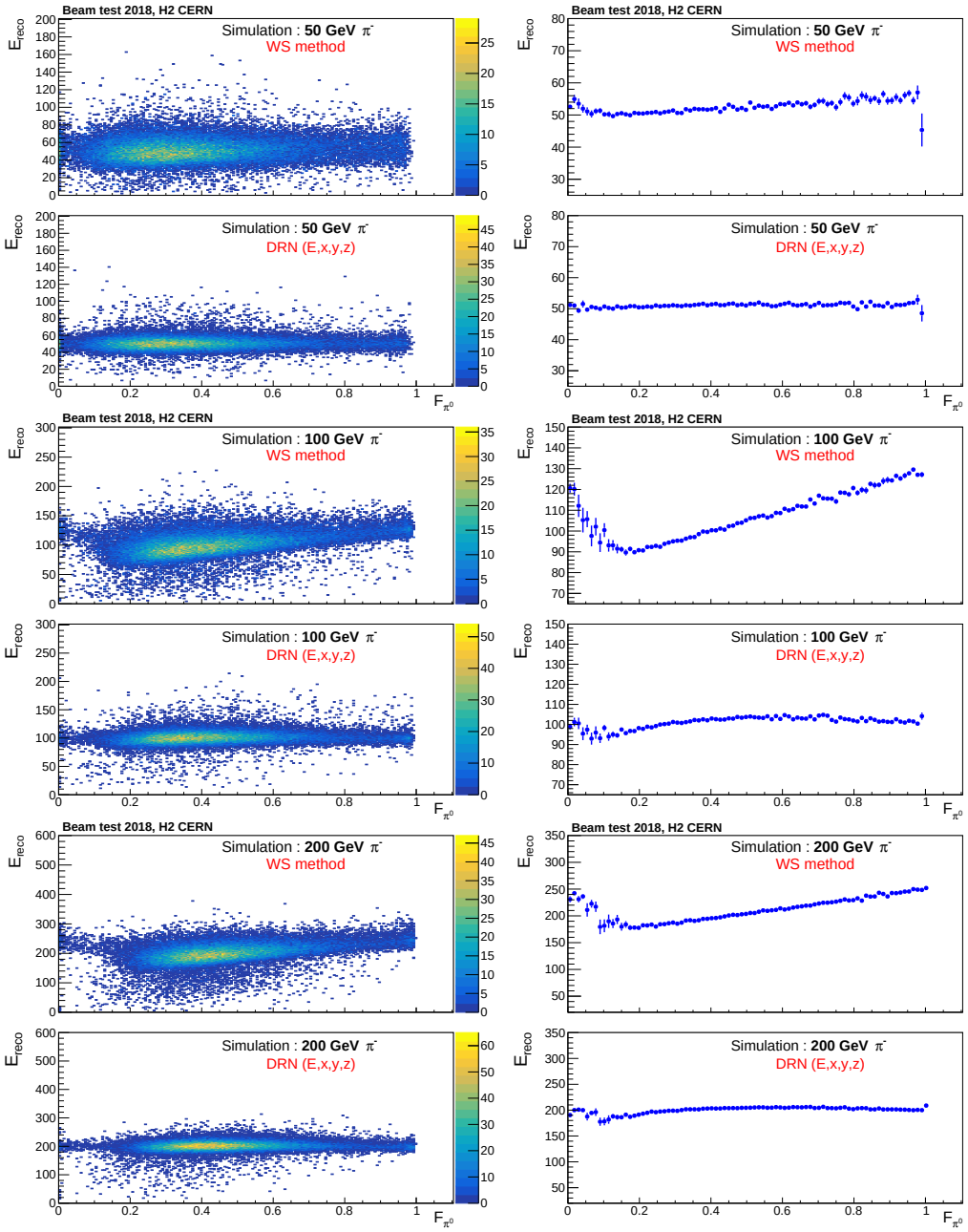


Figure 5.56: Energy reconstructed using WS method (top panel) and DRN(E,x,y,z) model (lower panel) for pions of energy 50 GeV, 100 GeV and 200 GeV as a function of fraction of energy carried by π^0s in a given shower (left). The right column shows the profile plots created after calculating the mean of reconstructed energy for different bins in F_{π^0}.

For hadron showers created by pions of given incident energy, we categorize the events into the following 4 event samples based on F_{π^0},

- F_{π^0} within $\mu \pm \sigma/2$
- $F_{\pi^0} > \mu + \sigma/2$
- $F_{\pi^0} < \mu - \sigma/2$ and $> \mu - \sigma$
- $F_{\pi^0} < \mu - \sigma$

The energy reconstructed using the DRN and WS method for different event categories is shown in Figure 5.57 for pions with different energy. The "Inclusive" stands for all the events generated using this simulation. The prediction of the WS method deviates from the true energy as F_{π^0} deviates from its average values. In contrast to this, DRN can position events with extreme fluctuations in F_{π^0}, close to the true energy values. The variation in rechits across events represent fluctuations in energy deposition patterns of hadronic showers. These hits, containing a complete set of information, are provided to the DRN on an event-by-event basis, allowing it to learn these fluctuations and accurately predict the true energy, even in cases where the pion behaves more like a minimum ionizing particle (MIP) or when F_{π^0} is very high. In contrast, the WS method lacks this detailed information to accurately predict the energy even for such extreme cases. As a result, when F_{π^0} deviates from the expected average values, the response of the WS method shifts similarly, as illustrated in figure 5.56. The plots for other remaining energy points are shown in Figure A.19 of Appendix A.1.

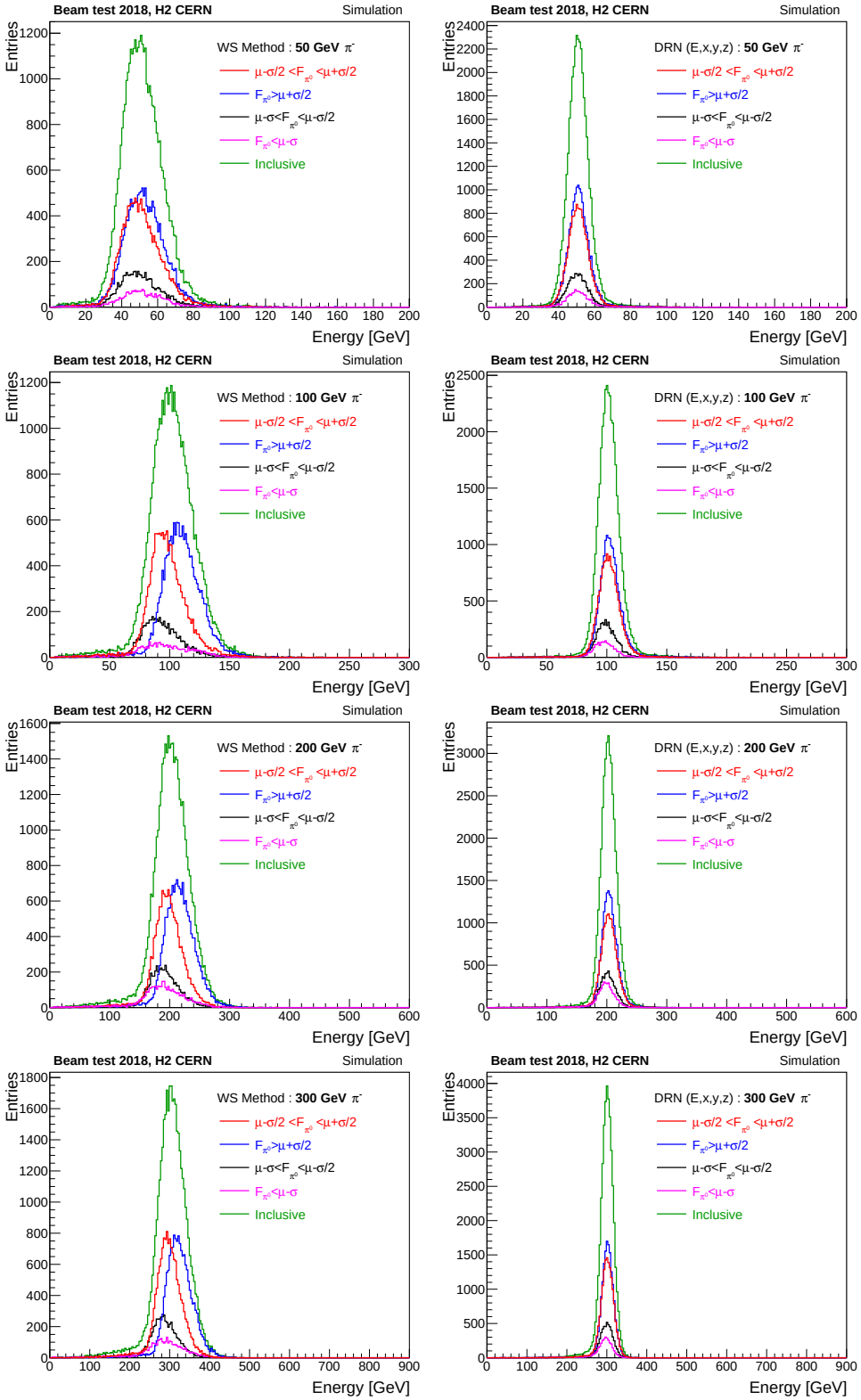


Figure 5.57: This figure illustrates performance of DRN and WS method in terms of reconstructing the energy for different events falling into different categories for 50, 100, 200 and 300 GeV pions. "Inclusive" line represents the energy predictions for all events.

5.8 Energy calibration of CLUE clusters

The high PU environment, resulting from average 200 pp collisions in the same bunch crossing expected at the HL-LHC, poses an unequaled challenge for both the online and offline event reconstruction. This is further complicated by the fact that the electromagnetic and hadronic showers resulting from neighboring particles are likely to overlap, especially in the forward regions of the detector. The HGCAL, with its high transverse and longitudinal granularity, also presents a great challenge, as the conventional algorithms used for clustering the energy belonging to the same mother particle are likely to blow in memory and computational time. In addition, combinatorial backgrounds can be huge in such a situation and will add to the complexity of the problem. The machine learning clustering algorithm that can be fully parallelized as well as run on heterogeneous platforms like combinations of CPUs and GPUs is needed. The Iterative CLustering (TICL) [130] framework for HGCAL reconstruction is being developed to achieve these goals and is designed to work in the CMSSW software. TICL workflow starts with clustering the rechits in each layer of the HGCAL (2D clusters), classifying them as seed clusters and outliers followed by a procedure for pattern recognition and linking across the whole detector. Given the large number of rechits per layer, it is computationally more beneficial to first make the 2D clusters using a CLUstering of Energy (CLUE) algorithm [131]. The CLUE is an algorithm based on energy density inspired by clustering using Fast Search and Finding Density Peak (CFSFDP) [132] that is capable of clustering rechits into arbitrary shapes of clusters. The algorithm starts with spatial partitioning of the data in a fixed grid format. Unlike CFSFDP [132] in which one has to go over all houses to investigate the neighborhood of a given house (denoted N_d), only those situated in the houses affected by the square window around it are used in CLUE, as illustrated in Figure 5.58.

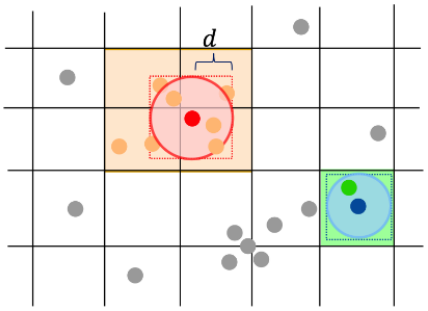


Figure 5.58: A fixed grid based on parameter d is defined for each 2D points such that finding the highest neighbour is fast as compared to other algorithms. For local density calculation and querying for neighbourhood, all the points in the bins touched by square window [x $\pm$ d_i], [y $\pm$ d_i] are considered as shown green and blue points in the above image [131].

After creating a grid and defining the neighborhood of a given point, the algorithm proceeds in the following steps:

- The local density at the i-th point, ρ_i is calculated as equation 5.18 where W_j is the energy of the point and χ_{ij} is a convolution kernel that can be a flat, Gaussian, or exponential function based on specific problems.
- In the second step, the nearest point with the highest density is found using equation 5.19 where N (i) are all points in the grid box. Minimum separation

using equation 5.20 where d_m is $\max(\delta_o, \delta_c)$ where δ_c or δ_o is the minimum spatial separation for a point to be seed or outlier, respectively.

- In the third step, a hit is promoted to seed if $\rho_i > \rho_c$ and $\delta_i > \delta_c$ or reduced to outliers if $\rho_i < \rho_c$ and $\delta_i < \delta_c$ where ρ_c is the minimum or maximum threshold in ρ_i.
- In the fourth and final step, the cluster is created combining the follower points with the seed. Outlier rechits and their descendant followers are not assigned to any cluster.

Various steps of this process of clustering the hits are illustrated in Figure 5.59. The position of each cluster is recalculated with the logarithmic weighted position of the energy deposits. The cluster energy is the total local density of all the rechits in that cluster.

$$\rho_i = \sum_{j:j\in N_{d_c}} \chi(d_{ij}) W_j \tag{5.18}$$

$$nh_i = argmin_{j\in N'_{d_m}(i)} d_{ij}, \;\; if |N'_{d_m}(i)| \neq 0, \;\; else \;\; nh_i = 0 \tag{5.19}$$

$$\delta_i = d_{inh_i} \;\; if |N'_{d_m}(i)| \neq 0, \;\; else \;\; \delta_i = \infty \tag{5.20}$$

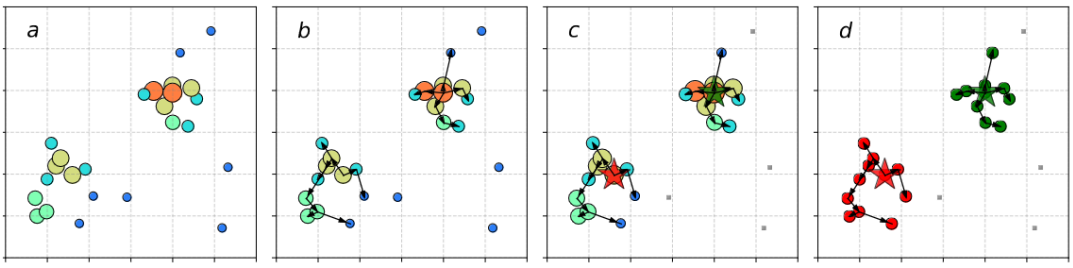


Figure 5.59: This image demonstrates the step wise process of CLUE algorithm. In step (a) local density ρ is calculated and represented by color and size of each point. In step (b), pointing arrows towards other point represents nearest higher points after calculating separation δ. Seed and outliers are selected based on δ and ρ in step (c) and nearest neighbor points are clustered with the seed in step (d) [131].

Hence, there are four key parameters in the algorithm, $\mathbf{d_c}$, $\boldsymbol{\rho_c}$, $\boldsymbol{\delta_c}$ and $\boldsymbol{\delta_o}$ which must be optimized for a given calorimeter and particle type. These are generally chosen on the basis of the transverse shower size or the lateral granularity of the detector. The CLUE optimized for hadronic and electromagnetic showers with full HGCAL geometry is tweaked for the beam test geometry, and the following parameters are used:

- ρ_c = 4 MIPs in each of the three subdetectors.
- δ_c = 1.3 cm for CE-E and CE-H and $\delta_c = (3\sqrt{2} + 0.1)$ cm for AHCAL.
- δ_o is set to $2\delta_c$.
- To calculate the local density for each rechit, κ_{ij} is set to 1 if i = j and 0.5 if i≠j and calculated within the $x_i \pm \delta_c$ and $y_i \pm \delta_c$, where x_i and y_i are the coordinates x and y for that rechit.
- The search for the nearest higher points is done in $x_i \pm 2\delta_c$ and $y_i \pm 2\delta_c$.

The CLUE algorithm takes n rechits and creates k number of 2D clusters (also called layer-wise clusters) with E, x, y, and z information and reject some rechits as outliers. In terms of 2D clusters, this reduces the granularity of information available by a factor of k/n on an average which corresponds to approximately 30−50% in our current settings, and removing the outlier rechits reduces the granularity by 10%. The distributions of the number of clustered rechits versus the number of all rechits in showers produced by different energies π^- are shown in Figure 5.60 (left column). Distributions of number of 2D CLUE clustered rechits versus the number of clusters in showers produced by different energies π^- are shown in Figure 5.60 (right column).

The cluster multiplicity as a function of layers is shown in figures A.20 and A.21 for 20 and 300 GeV pions. The number of clusters in any of the first 9 layers (29-37th) of CE-H is higher than in any layer of CE-E and last three layers of CE-H given only one single module available in these layers and the lateral granularity in terms of cell size is the same. The number of clusters in any AHCAL layers is also less as compared to CE-H given less activity in AHCAL and we are only taking 10 layers out of 39 sampling layers.

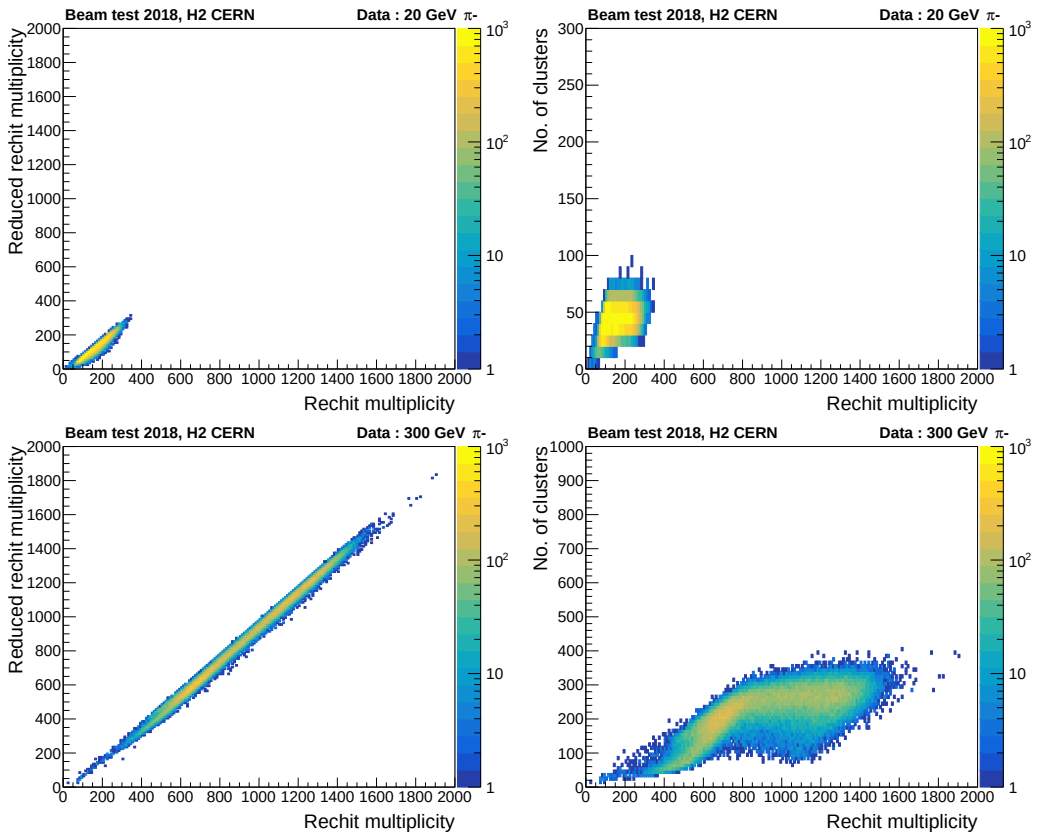


Figure 5.60: The left and right column shows the total number of rechits (X-axis) vs the rechits selected to create the clusters (Y-axis) and total number of rechits(X-axis) vs the total number of clusters created (Y-axis) respectively. Reduction in granularity up to 40-30% and up to $\sim$ 50% if we reject rechits that are not considered for clustering or consider only clusters respectively.

To understand the impact of information reduction due to the clustering process on the energy reconstruction of π^- showers, we performed two sets of DRN training using the (E, x, y, z), (E, z) and E of (a) all clustered rechits and (b) 2D CLUE clusters as

input features (i.e., a total of six sets of trainings for 150-200 epochs). The distributions of the pion energies reconstructed by DRN(E,x,y,z) model using clustered rechits (CLUE 2D rechits) and 2D clusters as input are compared with those reconstructed using all rechits in Figures 5.61. Although the bulk of distributions are similar, there is a broadening of distributions when using CLUE 2D rechits or CLUE 2D clusters as inputs.

Figure 5.62 shows the energy resolution as a function of the true energy for the three DRN models in simulation. For DRN(E,x,y,z), the resolution degrades up to 20% when using 2D clusters as input, and by 10% when using rechits used to create these clusters. In case of DRN(E,z) and DRN(E), the resolution degrades by 20% for both training with 2D clusters and rechits used to create these clusters. The performance of the DRNs trained on the evaluated data clusters is similar to that of simulation and shown in figure 5.63. Even with the transverse granularity information reduced by 50%, provided as CLUE 2D clusters, the DRN is able to compensate for hadron shower fluctuations and maintains resolution within 15-20% of the best obtained using all rechits in these studies in simulation as well as beam test data.

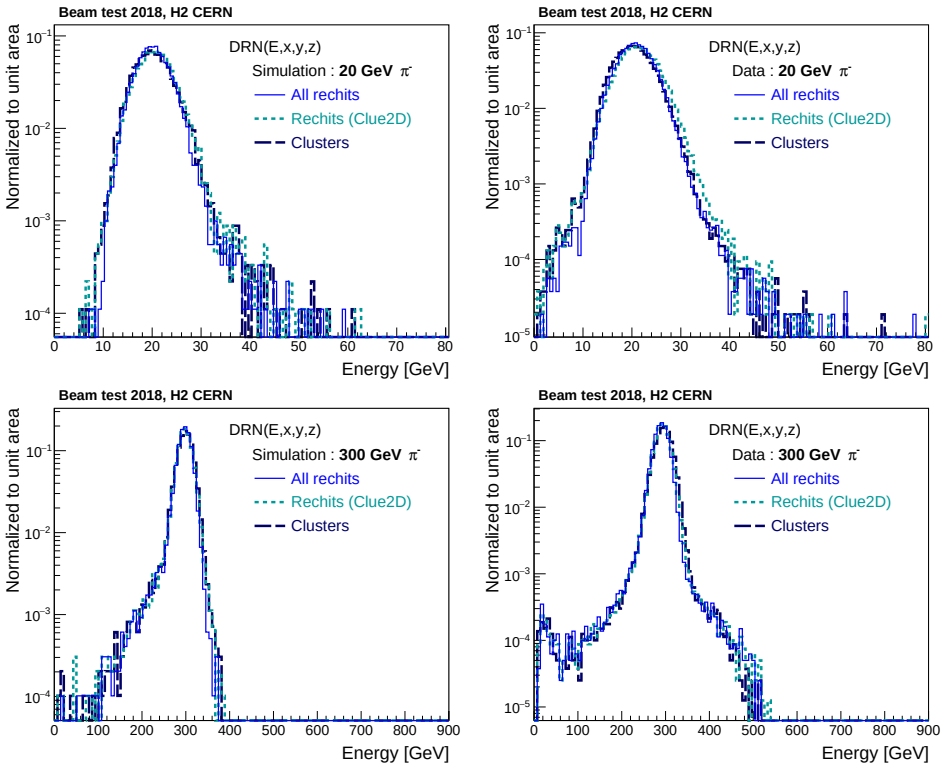


Figure 5.61: Reconstructed energy distributions using DRNs trained on clusters, rechits used to create cluster and all rechits for different energies ranging from 20 and 300 GeV are shown for simulation (left) and data (right). DRN was using four input features that is energy and x, y, z position of rechits/cluster (DRN(E,x,y,z)).

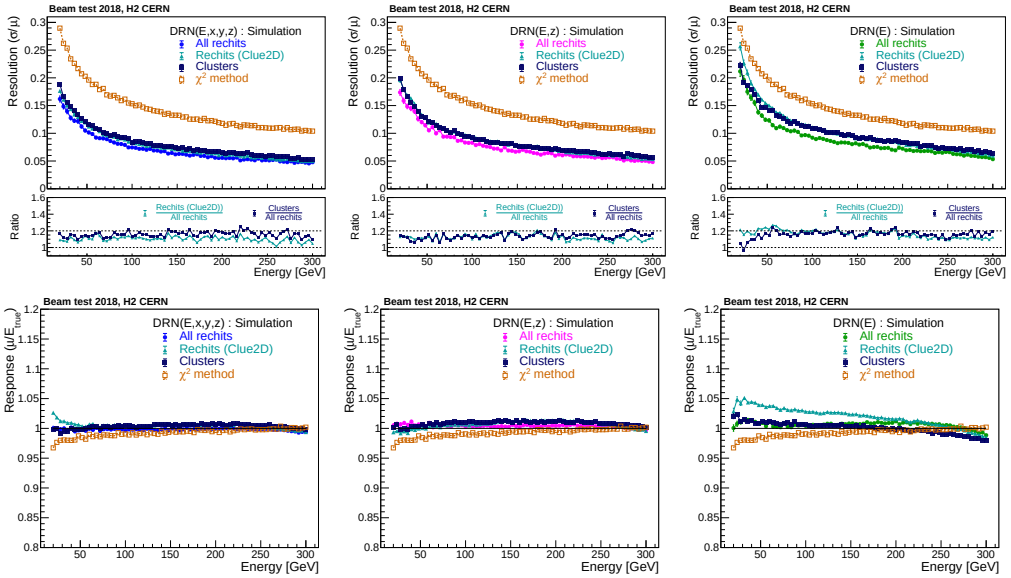


Figure 5.62: Resolution (top) and response (bottom) as a function of energy predicted in simulation by DRN(E), DRN(E,z) and DRN(E x y z) trained on different inputs are shown from left to right, respectively.

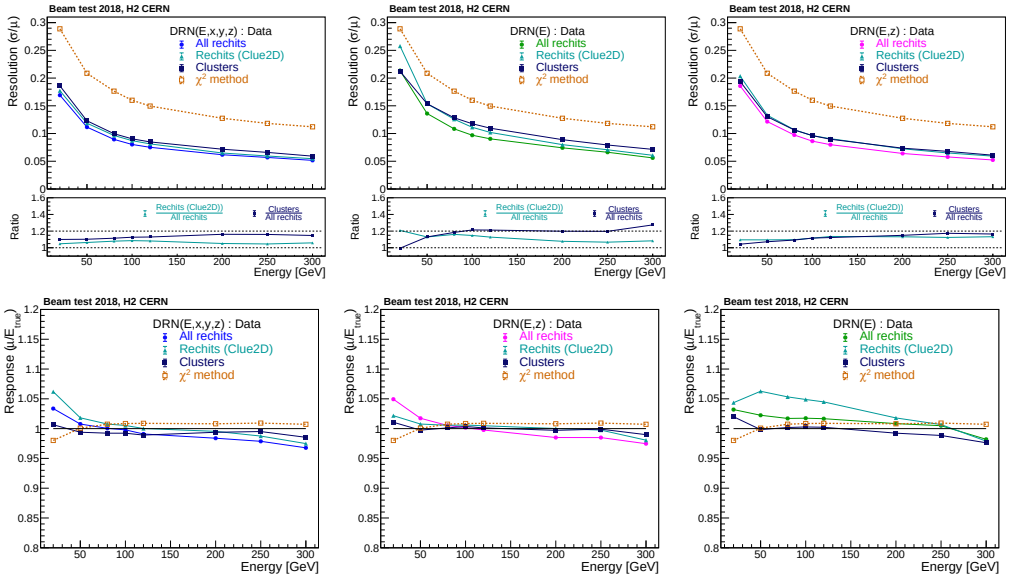


Figure 5.63: Resolution (top) and response (bottom) as a function of energy predicted in data by DRN(E), DRN(E,z) and DRN(E x y z) trained on different inputs are shown from left to right, respectively.

5.9 Expected performance using the full HGCAL geometry

With the HGCAL prototype detector used in the beam test experiment, we could establish the significant improvement in simulation and data. We also performed several studies to understand the aspects of the hadron shower which the DRN is able to learn and compensate for event-to-event fluctuations in energy deposits. However, we know that the prototype detector is not large enough to completely contain the pion showers in the transverse direction and the leakage was further enhanced in CE-H since the last three layers could be instrumented only with one module each. However, evolution of the shower continued in these layers as the absorber plates were larger (but only a limited signal could be recorded).

To get a realistic idea of the final hadron shower energy resolution which could be achieved by the CMS HGCAL [94], we trained the same DRN on a sample of charged pions generated with the full HGCAL geometry [133] simulated using GEANT4 [58]. This test setup does not include simulation of the other CMS detector components. It includes the latest configuration of three subsections CE-E, CE-H (fine) and CE-H (coarse) having 26, 11, and 10 sampling layers, respectively. The total depth of the calorimeter is $\approx 10\lambda_{Int}$. The detector simulation set-up is configured with the silicon modules having three different thicknesses 100, 200, and 300 μm and cell sizes of 0.5 cm^2 and 1 cm^2. This simulation uses 100 μm Si modules as a proxy to 120 μm, which will be used in the final HGCAL geometry. The size of the scintillator cell is defined as a square in $\eta - \phi$ space with 1^0 or 1.25^0 side, which makes the cell area slightly depending on the position of z and of the order of 30 cm^2.

In the digitization step, the energy deposited by shower particles in the active material is converted into equivalent number of MIPs by applying hard-coded factors for silicon modules with different thickness and for scintillators determined with simulated 200 GeV muons. For silicon modules with thicknesses 100, 200 and 300 μm, we used 3.65, 1.82 and 1.21 as MeV to MIP conversion factor, respectively, whereas for all scintillator cells 2.04 has been used [134]. The converted energy is smeared with Gaussian by 2% to emulate the effects of electronic noise. We simulated a sample of charged-pion showers with energies distributed uniformly from 10 to 350 GeV using this geometry setup. A total of 9k pions showers per energy point has been generated using the fixed energy particle gun which is symmetric in ϕ and η coordinate of particle gun has been fixed at 2.0 . The η=2 has been chosen such that pion showers pass through all the three subsections, including the last few layers with only scintillators as the active material, providing a good approximation to the beam test prototype geometry. The magnetic field was set to zero. The `QGSP_FTFP_BERT_EMN` physics list has been used to simulate pion showers in this geometry, whereas `FTFP_BERT_EMN` has been used for the beam test simulation set-up.

All rechits that are passing a noise threshold of 0.5 MIP and are within dR radius of 0.5 around the expected flight direction of pions and defined as $\sqrt{(d\eta^2 + d\phi^2)}$ are selected to train the DRNs. The rechit energy is converted to GeV using the fixed energy weights estimated separately for all three sections following the procedure described in 5.3. The MIP to MeV scale for CE-E is set to 9.8 MeV per MIP using 50 GeV positrons and set to 58.3 and 84 MeV per MIP for CE-H (fine) and CE-H (coarse), respectively, using 50 GeV pions that do not initiate a hadronic shower in CE-E. We calculate the total energy reconstructed using a fixed weight (E_{fix}) after putting all these factors in the equation 5.1. The models DRN(E), DRN(E,z) and DRN(E,x,y,z) are trained up to $100-150$ epochs using energy and spatial coordinates of rechits.

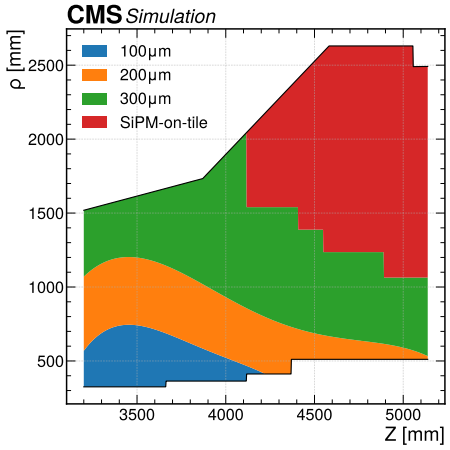


Figure 5.64: Schematic of standalone geometry of expected HGCAL configuration showing the boundaries between sections and Si thickness implemented in simulation.

Respective epochs with minimum validation losses (131, 96 and 131th epoch for DRN(E,x,y,z), DRN(E,z) and DRN(E), respectively) are used to infer predicted energies and assess the performance of DRNs for the energy resolution and response, shown in Fig. 5.65. The energy resolution using DRN(E,x,y,z) and DRN(E,z) is very similar, indicating that the addition of transverse shower information does not provide any additional learning beyond what is already obtained from rechit energies and depth information as in case of DRN(E,z). We would like to remind here that up to 15 % improvement in energy resolution was observed in the beam test setup by adding transverse shower information on top of rechit energies and depth information as shown in Fig. 5.30. This observation can be understood as the difference in the transverse coverage for the pion shower in both the full HGCAL and the beam test geometry. For full HGCAL, full shower information is available without any leakage and the rechit energy distribution is enough for model to compensate for energy deposit patterns.In case beam test setup, it helps DRN to learn the details about the energy leaked out of the detector. This also suggests that having transverse shower information will be beneficial in more realistic scenarios where energy showers in the calorimetry are likely to overlap due to high pile-up scenarios of HL-LHC, and certain fraction of shower information is not available for the model to learn. This is reflected in more detail in Section 5.8.

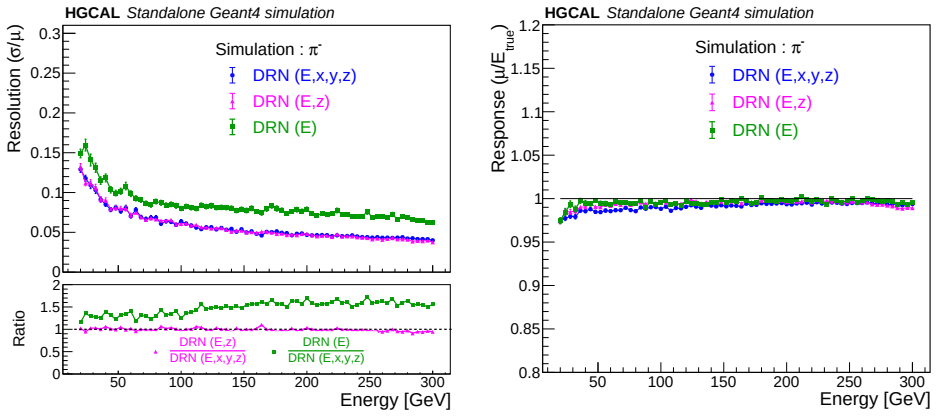


Figure 5.65: Resolution (left) and response (right) predicted by DRN(E), DRN(E,z) and DRN(E x y z) in the full HGCAL setup for simulated validation dataset.

The total number of rechits for different energies in full HGCAL and beam test geometry is shown in Fig. A.22, showing more rechit multiplicity with full HGCAL geometry. Comparisons of the reconstructed energy distributions of pions with 20 and 200 GeV energies for the validation dataset in the full HGCAL and beam test setups using different DRNs are shown in Fig. 5.66. The DRNs trained with the full HGCAL inputs improves the bulk of the distribution as well as reduces the tails overall as compared to the beam test geometry. The energy resolution and response as a function of incident pion energy of 20–300 GeV using different DRNs for validation dataset are shown in Fig. 5.67.

We observe an improvement of up to 30% in energy resolution predicted by DRN(E,x,y,z) at lower energies and up to 20% at higher energies. Similarly for DRN(E,z) improvements in energy resolution vary from 40% to 20% for low and high energies, respectively. For DRN(E), it improves at the lower energies but performs very similarly at higher energies. The stochastic term improves from 89 to 58%, 78 to 53%, 72 to 52% using full HGCAL geometry compared to the beam test setup for DRN(E), DRN(E,z) and DRN(E,x,y,z), respectively. This significant improvement in the energy resolution going from the beam test geometry to the full HGCAL geometry can be understood based on the transverse containment of the shower in both geometry. In the beam test, there was energy leakage, which was more significant for low-energy pion showers because of the greater transverse spread for low-energy showers in the detector. The DRN was extrapolating the available information to account for this, which is not the case for the full HGCAL.

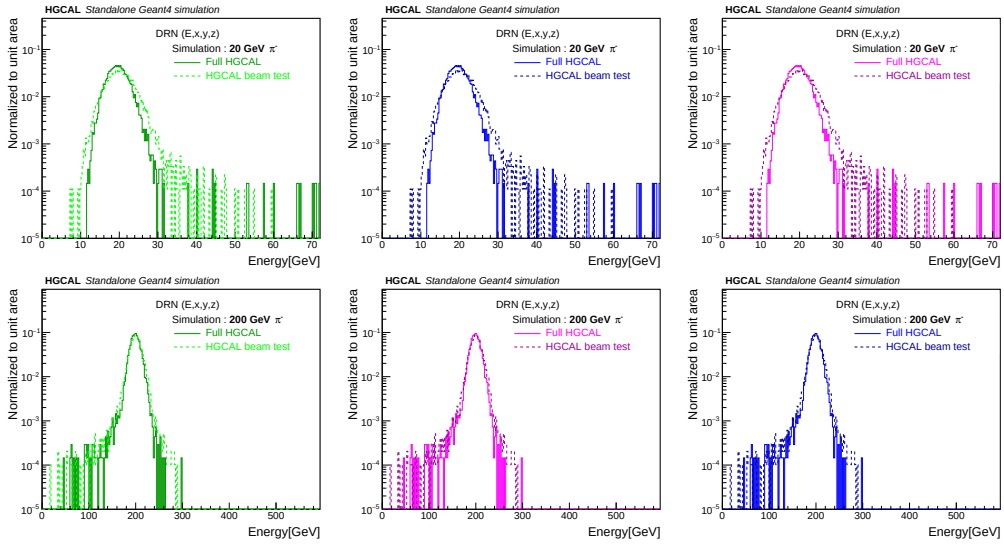


Figure 5.66: Reconstructed energy distributions for pions with energies 20 GeV (top) and 200 GeV (bottom) using DRN(E), DRN(E,z) and DRN(E x y z) (from left to right) on validation simulated dataset in full HGCAL and beam test setups.

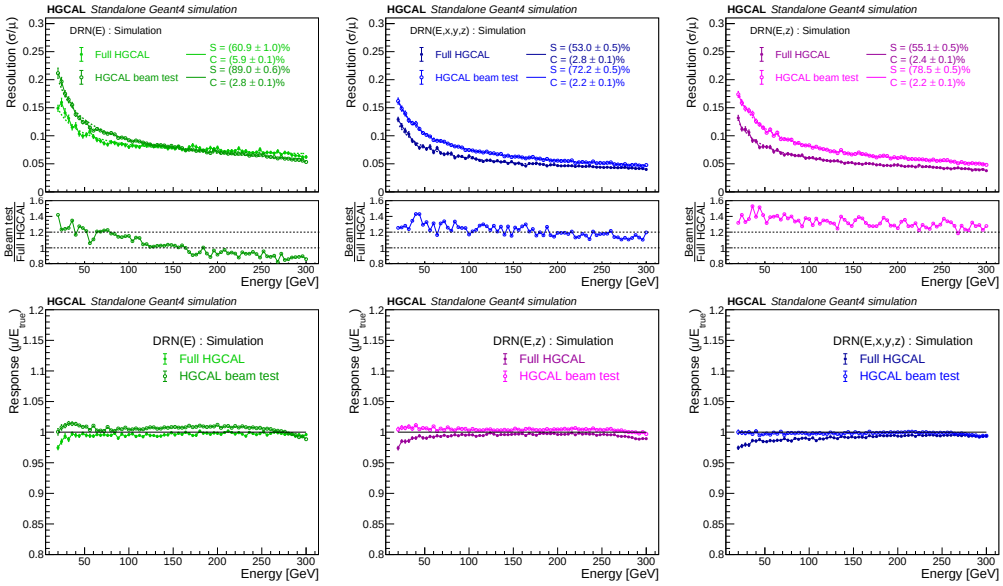


Figure 5.67: Comparison of energy resolution (top) and response (bottom) in full HGCAL and TB setups using DRN(E), DRN(E,z) and DRN(E x y z) (from left to right) for validation dataset.

5.10 Summary

In this study, we carefully examined the abilities of a dynamic reduction network (DRN) based on graphical neural networks (GNN). Our goal was to improve the reconstruction of hadron showers from charged pions using a prototype of the CMS HGCAL detector during a beam test at CERN SPS in October 2018. Our results show that the DRN improves the hadron energy resolution by a factor of two compared to the traditional method, which only sums the energy deposited in the detector's readout channels. We also found that using both transverse and longitudinal spatial information, along with signal amplitudes, makes detailed rechit energy calibration unnecessary. Additionally, simulations with a full endcap geometry show that the DRN achieves an energy resolution with a stochastic term of 53%. This is the first report on the performance of a machine learning-based energy reconstruction algorithm for hadron showers in the CMS HGCAL prototype, using data from a beam test.

Chapter 6

Testing and validating the HGCAL readout system

Following upon Chapter 3, this chapter delves into the critical aspects of the design and optimization of the HGCAL readout electronic system, which is capable of managing the enormous data from more than 6 million channels. The CMS HGCAL collaboration has made extensive efforts in system development and iterative improvements. The near-final readout system, shaped by such numerous system tests and beam tests, is presented along with a detailed account of the challenges encountered and the solutions implemented to address them.
For testing and validating the readout system, two complementary approaches are briefly discussed: simulation-based studies and beam tests using real particles. The correction method to model occupancies of different type of modules with mismatched geometry is also developed and discussed in this chapter. For the simulation-based studies, a method to create high-statistics pseudo-events using a multivariate Gaussian distribution and quantile transformation to maintain complex correlations among detector modules is discussed. This chapter also summarizes studies of the measurement of deadtime and data losses within the readout system. A comprehensive overview of these findings has been documented in a CMS internal note, which is accessible at [135].

6.1 The HGCAL electronics readout system

The electronics machinery that collects the signals generated by silicon sensors and scintillator tiles and process them for final storage is collectively termed the HGCAL readout system. Signal pulse height typically varies from a few millivolts (mV) to a few hundreds of mV, and it needs to be amplified, shaped and digitized using an on-detector or front-end (FE) system to avoid introducing any additional noise to analogue signal while transmitting. Since these are on-detector, the FE components are strictly required to be radiation hard. The digitized signal is further formatted, analyzed, and prepared for offline storage using an off-detector or back-end (BE) system. A schematic of the HGCAL readout system with FE and BE components is shown in Fig. 6.1. Mainly three kinds of information (data) exchange happen between FE and BE:

- **Fast controls** include time specific signals (commands) for distributing LHC system clock, level-1 trigger accept (L1A) and orbit synchronization signals, initiating reset procedures etc. [136]. The number of commands that can be sent in a single bunch crossing (BX) are limited by the 320 Mbps encoding of fast control streams.

- **Slow control** stream including FE configuration, gain configurations, zero suppression, etc. These need not be time specific and are encoded as 40 Mbps data streams [137].
- **Acquisition of DAQ and TPG data**: The TPG data are coarser information, nominally the digital sum of 2×2 or 3×3 cells for LD and HD modules, respectively [138] and is sent to the central CMS trigger system to make the real-time trigger decision whether or not to store a given event. The DAQ data are the data from each readout channel that pass the noise or zero suppression threshold and are permanently stored for offline analysis.

An overview of FE and BE components and their functionalities are shown in Fig. 6.1. The FE reads, digitizes, concentrates the data corresponding to energy & timing measurements from the cells of a given module and sends it off to BE via high-speed electronic links. It also distributes clock, control, and configuration commands. The BE processes the data and communicates with the central DAQ to obtain a final decision on whether to keep the event or reject it permanently.

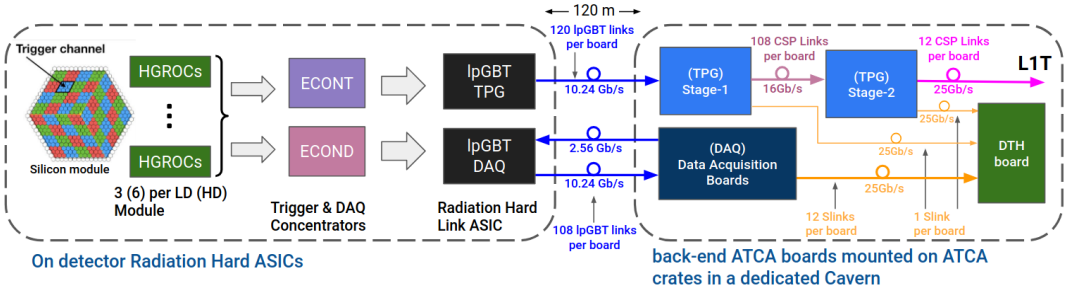


Figure 6.1: Schematic diagram showing the key components of the readout system of HGCAL [139].

Optimizing the HGCAL readout is a challenging task, not only because of a large number of channels, but also a large variation in module types and channel occupancies depending on location of modules. This can result in bandwidth saturation if the aggregation of information across layers or cassettes is not optimized carefully. We summarize the components and aspects of FE & BE which are relevant for emulator-based studies which is the focus of this chapter. A similar system is also used for scintillator tile boards FE but adapted for differences in SiPM high gains and channel multiplicities. Details are presented in [97].

6.1.1 Front-end (FE) readout system

A schematic of the FE electronics system and its main components designed for silicon modules is shown in Fig. 6.2. Hexaboard PCBs host HGCROCs (read, shape, and digitize signal pulses, prepare DAQ and TPG dataformats), ECON-Ds and ECON-Ts to concentrate the digitized data and Rafael ASICs which are a fanout chip for providing clock and fast control data stream to the HGCROC and ECON ASICs [140]. The Hexaboard connects the sensors to HGCROC and the motherboard for data transfer and control signals and provides low voltage to power various ASICs and bias voltage to silicon sensors [141]. Up to 11 variants of the Hexaboard design are needed to accommodate variations in the shapes and channel densities of silicon modules [141]. Data transmission "engines" host lpGBTs and VTRx+ ASICs for serializing and de-serializing data and send it off the detector to BE. Wagons are passive boards that connect Hexaboards to engines.

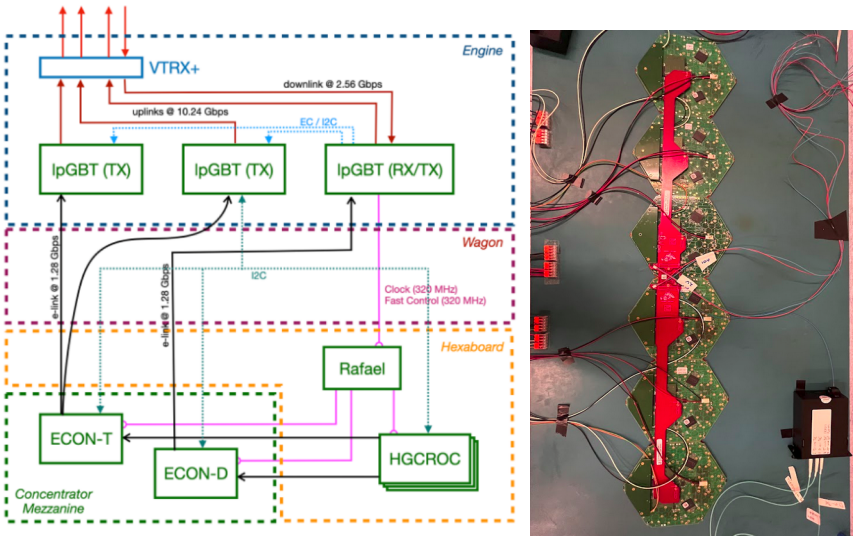


Figure 6.2: A schematic of main components of radiation hard FE electronic system of silicon modules [142] (left) and actual module train extended in horizontal direction.

HGCROC - the HGCAL ReadOut Chip

The HGCROC is designed based on a radiation-hard 130 nm CMOS technology to tolerate radiation levels of up to 200 Mrad or equivalent fluence of 1.10^{16} neq/cm^2. Its schematic diagram is presented in Fig. 6.3. A preamplifier (PA) amplifies the signal generated by charge collected from the silicon sensor and converts it to a voltage signal with the best possible signal-to-noise (S/N) ratio and high gain for MIP-like signals [143]. Fast and slow shapers ensure short delays and peaking times of the order of 25 ns. Digitization in a large dynamic range of ≈ 0.3 fC to about 10 pC corresponding to the charge generated in a single cell by a MIP and energetic shower particles, respectively, is achieved employing two measurements [143]. For input charges equivalent to or less than 100 MIPs, a measurement of the peak height of the pulse is provided in units of digitized counts using a 10-bit ADC. For higher charges, a Time-over-Threshold (ToT) measurement proportional to the pulse height is provided along with a flag indicating its presence. An 11 bit digital-to-analog converter (DAC) and two injection capacitors (0.5 pF and 8 pF) are also included in the HGCROC for in situ intergain calibration between ADC and ToT measurement. The HGCROC also provides the time of arrival (ToA) of the signal with a resolution of $\approx$ 50 ps for charges above 12 fC. All of this information is made available at every 25 ns. In addition to all these advanced features, the HGCROC is optimized to have low noise levels and low power consumption less than 15 mW per channel.

For each silicon cell, every event per 25 ns, the HGCROC prepares a 32-bit word providing number of ADC counts in current and previous BX, ADC(t) and ADC(t-1), respectively, TOT and TOA values. Only three of these digitized values are saved depending on predefined user configurable thresholds as summarized in Table 6.4 [143]. Two flags Tp and Tc, one bit each, indicating TOT-in-progress and TOT-Complete are also saved. Although these measurements are completed at different clock cycles, a latency manager ensures that they are aligned to be available within the same clock cycle for a specific bunch crossing (BX) [143]. At this point, the data are split into two separate processing paths, TPG and DAQ. A simplified compressed TPG input as digital sum of 2×2 or 3×3 cells, called trigger cells, for LDM and HDM respectively [138] while the DAQ data is stored in RAM-1 buffer for

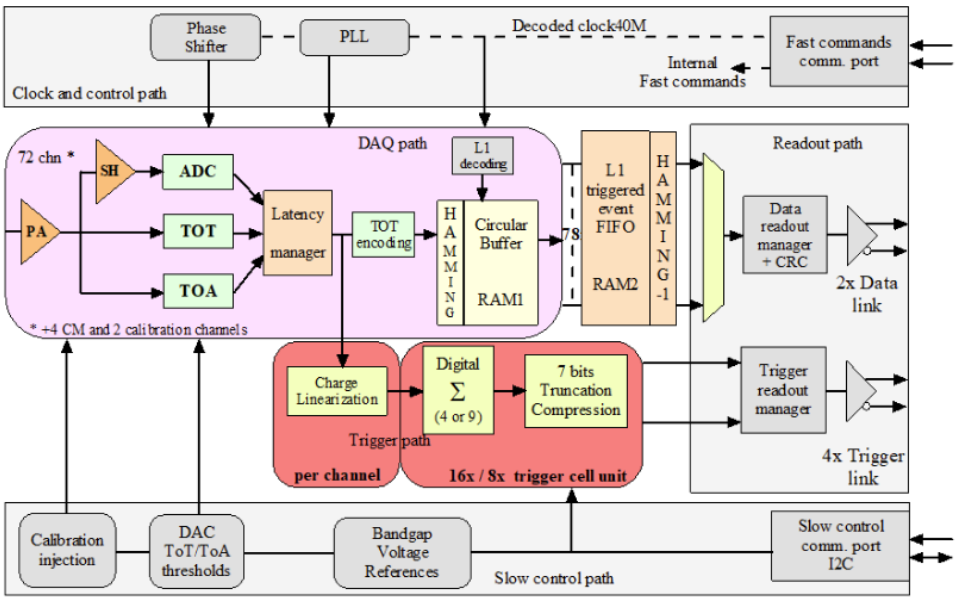


Figure 6.3: The different components of the HGCROC chip are shown in the above schematic diagram [143].

the next 512 BX.

	ADC (t-1)	ADC (t)	TOT	TOA	Charge collection	Data type
1	x	x		x (=0)	Q < TOA_thr AN	Normal
2	x	x		x	Q < TOT_thr AN	Normal
3	x		x	x	Q > TOT_thr AN	Normal
4		x	x	x		Characterization

Figure 6.4: Information to be filled in DAQ data path [143].

Each HGCROC is made up of two symmetrical parts and can read up to a total of 78 channels, 39 channels per halfROC. Of these 39 channels per halfROC, 36 are full hexagonal readout cells, 1 common mode, and 2 are calibration channels (typically smaller size than normal readout cells). The formatted DAQ data from a halfROC are shown in Fig. 6.5. Each halfROC generates DAQ data of 41 32-bit-long words, namely one word for each of 39 channels, DAQ header, and Cyclic Redundancy Check (CRC). The header includes a 4 bit header and a 4 bit trailer. It also includes the BX number (12 bits), event number (6 bits), orbit number (3 bits), and Hamming error codes. The DAQ data are held in the HGCROC RAM-1 buffer for the following 512 BX, which, when receiving an L1A fast control command from TCDS2, is transferred to RAM-2 which acts like a FIFO memory. Three internal counters, the bunch crossing counter (BC), the orbit counter (OC), and the event counter (EC), are attached to the data in the RAM-1 $\rightarrow$ RAM-2 transfer. The estimated time latency for the whole process from signal collection to saving the event packet in RAM-1 is expected to be around 16 BX [144].

Each HGCROC transfers data from RAM-2 to ECON-D over two 1.28 Gbps receiver elinks (eRx) as shown in Fig. 6.3, at a rate of one 32-bit word per BX per elink, taking a total of 41 BX to completely transfer data from one halfROC. To read a full-size HD,

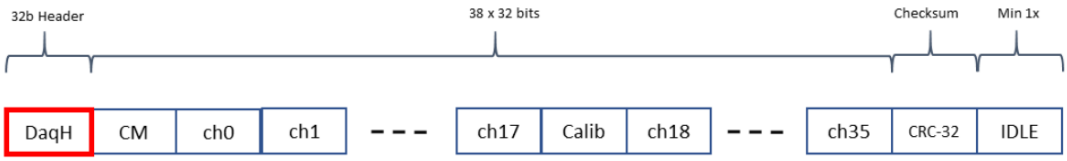


Figure 6.5: The diagram above provides a general overview of a readout system. The blocks illustrate the various components of the system, while the lower cartoon depicts the actual electronic circuitry [143].

(LD) silicon module, a total of 6, (3) HGCROCs is needed, and therefore a total of 12,(6) elinks are available. RAM-2 cannot store more than 33 events at any instant, restricting the trigger rate for CMS to not be greater than 1 MHz.

Endcap Concentrator for DAQ (ECON-D)

The ECON-D first aligns the digitized data received from HGCROC RAM-2 over the elinks and consolidates data received from multiple HGCROCs. This is essential to manage the vast number of channels and reduce the complexity of the data before transmitting them to the BE system. An important task of ECON-D is to apply zero suppression (ZS) using the user-configured algorithms designed to correct for common-mode noise and pileup from previous BX. In the ZS, the data from cells with signal below a certain threshold are not sent further, thus rejecting noise, as well as a significant reduction in the data volume which can otherwise exhaust the DAQ system. It happens on the channel data processor, CDP, or CP processor. ECON-D also serves as an interface between FE and BE electronics systems and allows configuring FE using control signals sent from BE. An L1A issued by TCDS2 is sent to all parts of the system, namely HGCROC, ECON-D, and BE. Upon receiving an L1A, the ECON-D pushes the current Event, BX and Orbit numbers into the L1A FIFO. It checks the half HGCROC headers for different event counters, few of which are mentioned in section 6.6 and more details in [145] and [137], and it reports any mismatches found with each other or their status at ECON-D which are important for system validation.

Using a Formatter functionality, the ECON-D assembles in an event packet the data from the channels to be kept after ZS and event header containing information about event, BX, Orbit (E/B/O) counters, various error flags, and the packet length (in units of number of 32-bit words). The formatted information in a *standard* data format event packet is illustrated in Fig. 6.6 and is used for normal data acquisition when ZS is applied and the buffer is not full. A *pass-through* data format has similar header information but includes all the raw data received from the HGCROC without processing any information. The packet should be 235 (472) 32-bit words in this case for LDM (HDM) modules. If the buffer is full not to be able to accommodate one more full packet, a *truncated* data format as in Fig. 6.7, with zero information from channels while keeping standard header (total three 32-bit words), is used to prevent incomplete packets which would de-synchronize the readout stream. The space for this is always reserved by the buffer control clock. For the studies presented in this chapter, the truncation threshold is set at 2600 words. When ECON-D expects data, but HGCROC event header information is not received in the expected BX, a *missing* data format is sent.

Up to 6 elink transceivers (eTx) with 1.28 Gbps can output event packets from the ECON-D to lpGBT (more details in the next section), transmitting one 32-bit word per elink. The event packet extends over hundreds of BX, depending on the variants

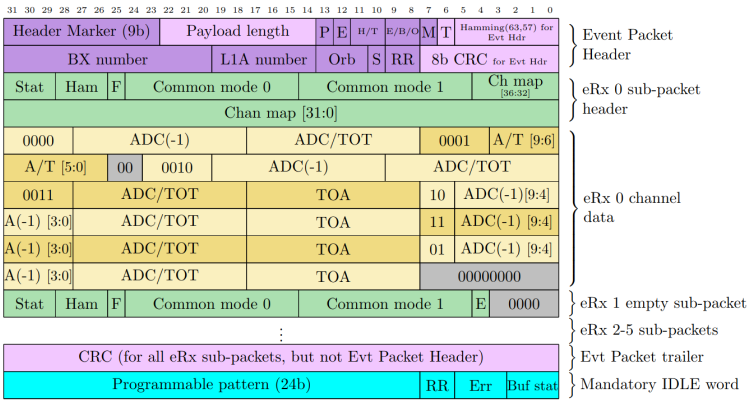


Figure 6.6: The above is the standard data format from the ECON-D [145].

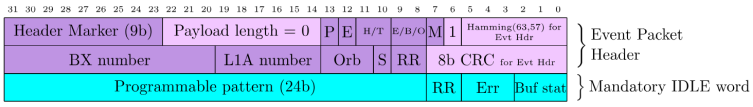


Figure 6.7: Truncated data format from ECON-D is shown above [145].

of the modules and the data rates. For example, a LDM with 192 readout channels in nonzero suppression mode, the event packet consists of 241 32-bit words. With 3 eTx connected, the event packet transmits over 80 BXs to lpGBT. If an L1A arrives during this time, ECON-D can process and store the next event in its output buffer, provided that there is enough buffer space. Each ECON-D has an estimated latency of 20 BX. The corresponding trigger data are handled and processed by Endcap Concentrator for Trigger (ECON-T). Both ECON-T and ECON-D ASICs are hosted on the Concentrator Mezzanine PCB for silicon modules, as shown in Fig. 6.2. This mezzanine interfaces with the Hexaboards, and any interaction with the Concentrator Mezzanine must occur through its connection to the Hexaboards.

The lpGBTs and VTRX+ ASICs

Event packets prepared by ECON-Ds are sent via elinks to low-power gigabit transmission optical fibers (lpGBT) [146] as shown in Fig. 6.2. The lpGBT is an advanced ASIC developed by CERN for high-speed (10.24 Gbps) data communication with a fixed and deterministic latency while maintaining low power consumption. At the lpGBT, data streams from multiple modules are merged, de-serialized, formatted, error protected with error detector and correction, and re-serialized. The combined data (both DAQ and trigger) are then transmitted through the High Speed Line Driver interface to the Versatile Link+ (VTRx+) [147]. Both lpGBT and VTRX+ are hosted on an Engine PCB that can be connected to a maximum of 7 modules at a time via Wagon PCBs; see Fig 6.2. With a wide variety in terms of geometry and data rates, the number of eTx connected to each module shall be optimized according to the module position, possible foreseen data rates, and number of channels available. FE mapping of elinks from ECON-Ds to lpGBTs is designed to minimize elinks and adhere to hardware constraints, ensuring that each lpGBT manages similar data rates and minimizes data loss.

The Engine PCB provides distribution of clock, fast and slow control signals, reset and enable signals to up to 7 modules, serialization and transmission of trigger and DAQ data, as well as measurements of configurable parameters [142]. The Wagon

PCB is a passive board connecting Hexaboards to engine board and routing path for e-links coming from lpGBTs. Address variations in shapes, channel density and data rates, 2 LD Engine, 51 LD Wagon, 2 HD Engine, 9 HD Wagon has been designed after thoroughly optimizing the purpose. The FE structure that integrates all of these various components in variable numbers of Hexaboards, wagon PCBs, and engine PCBs is known as an FE train. The Fig. 6.8 is showing the picture of the LD engine, one of the LD wagons, the DC-DC module to power the engine, and a FE train equipped with all these ASICs.

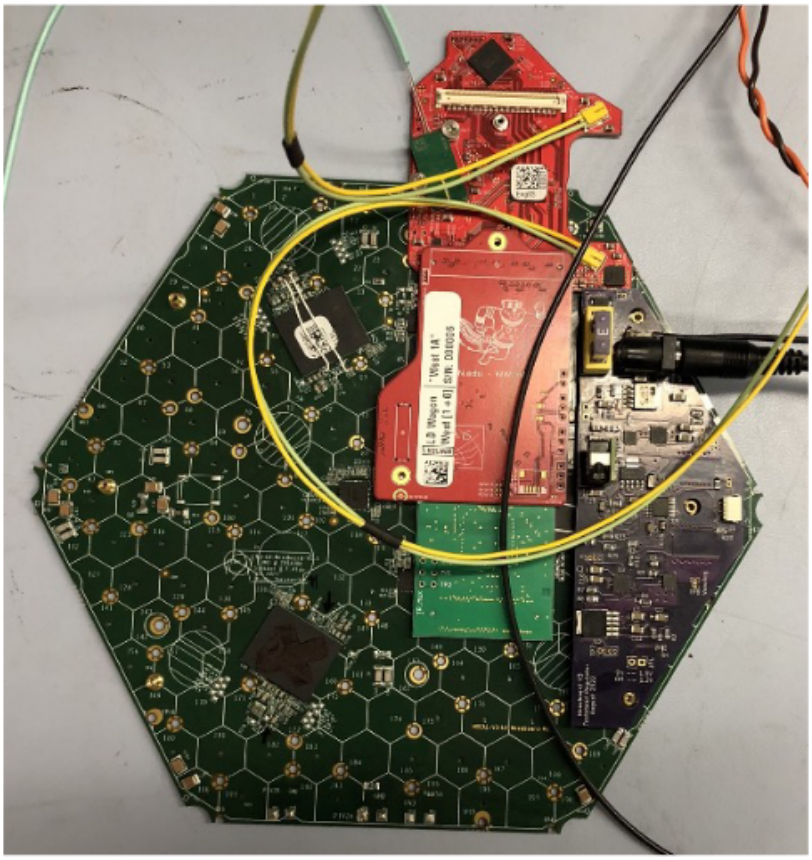


Figure 6.8: The FE train equipped with LD engine, LD wagon connector and DC-DC mezzanine [96].

The VTRX+ interfaces with optical fibers to handle high-bandwidth data transfer. At the input of BE, these functions are inverted to retreat the original packet. The output buffer is designed to have 12 SRAMs each with a depth of 256 words and 32b wide. Hence, at any instant, not more than 3072 32-bit words can be accommodated in the output buffer. There is a dedicated buffer status register of 3 bits indicating the currently occupied space in the output buffer which are sent along packet header. The user can configure 7 buffer status codes set through 7 programmable bin boundaries. A total of approximately 120 k HGCROCs, 30 k ECON-Ds and ECON-Ts will send DAQ and trigger data on approximately 10 k lpGBTs.

6.1.2 Back-end (BE) readout system

The HGCAL BE readout system is a high-bandwidth, low-latency specialized computer hardware that uses Xilinx VU13P FPGAs (field-programmable gate arrays) for off-detector data processing. It comprises a Data Acquisition System (DAQ), a Trigger Primitive Generation System (TPG), and a Detector Control and Safety System (DCS),

and their communication routes are shown in Fig. 6.9. These are mounted on custom Advanced Telecommunications Computing Architecture (ATCA) Serenity boards housed in CMS BE racks ≈100 m away from the FE electronics. In addition, a Detector Timing Hub (DTH) system on each Serenity board distributes timing and control signals to the DAQ BE, which passes them on to the rest of the HGCAL electronics.

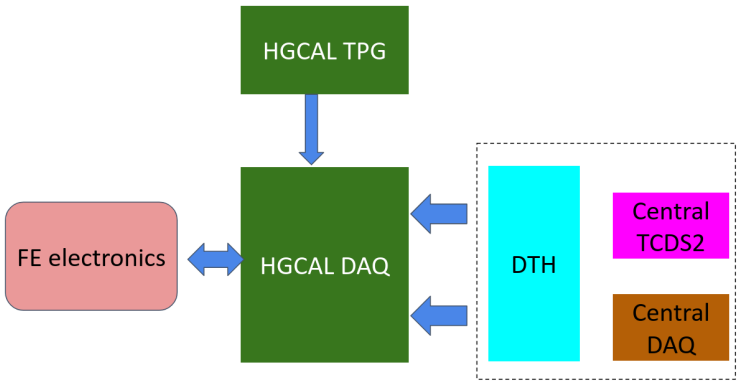


Figure 6.9: Schematic showing the interfaces of different systems with BE DAQ system [94].

A central CMS Timing, Control, and Distribution System (TCDS2) distributes 40 MHz LHC clock to the HGCAL BE and FE system and generates the Level-1 Accept (L1A) trigger which initiates the readout process in all CMS detector subsystems. The TCDS2 issues L1A and fast commands every 25 ns with precise timing requirements to synchronize different detector and electronics parts, manage resets, and clear buffers. The typical distribution latency is around 1 μs, except for L1A, which has a 10 μs latency [137].

The L1A is received in all parts of the HGCAL readout system, including HGCROC, ECON-D, and DAQ. While HGCROC is preparing to send the corresponding event information, ECON-D and DAQ are prepared to receive that information and look for the respective programmed header pattern at the input. Each BE DAQ receives these data, serializes, de-serializes, detects, and builds the event packet, and moves these event data to the central DAQ via SlinkRocket [148] high-speed link protocols developed specifically for CMS BE readout systems [149]. The BE DAQ also manages the clock and fast control signals from the TCDS2 (fast control) and configures the FE electronics (slow control), as well as sending reset requests and throttle signals to TCDS2. These control signals are distributed to the FE through lpGBT downlinks at 2.5 Gbps. The BE DAQ must guarantee that no data are lost in the process of data transmission from FE $\rightarrow$ BE $\rightarrow$ cDAQ.

To handle data rates from approximately 10 k lpGBT links per DAQ and TPG data, a total of 96 DAQ boards (108 lpGBT links per board) and 84 boards at TPG stage 1 (120 lpGBT links per board) and 108 boards at TPG stage 2 will be used for the final HGCAL [137].

DAQ system

The VU13P FPGA which hosts the DAQ firmware is a high-performance, power-efficient chip from Xilinx's Virtex UltraScale+ family offering large logic capacity, high-speed I/O and supports high-bandwidth memory. An overview of the DAQ architecture and functionality is shown in Fig. 6.10.

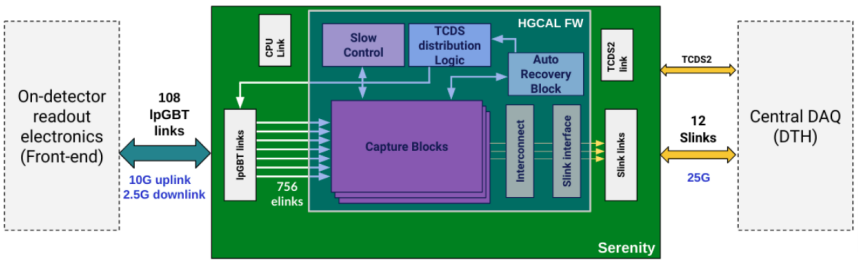


Figure 6.10: Schematic highlighting the key parts of DAQ firmware designed for HGCAL [149].

The DAQ-FPGA takes data from 108 lpGBT links connected. Each lpGBT link can take input from 7 elinks with incoming stream of 224-bit data bus consisting of one 32-bit word corresponding to one of 7 elinks connecting ECON-D ASICs to the lpGBT. Depending on position, geometry, and data rates seen by module, an ECON-D may utilize between 1 and 6 elinks, and multiple ECON-Ds can be connected to a single lpGBT. A single lpGBT can manage data from either one ECON-D using all 6 elinks or from several ECON-Ds with varying numbers of elinks. Fig. 6.11 shows an example of one such possible combination of ECON-D to lpGBT links.

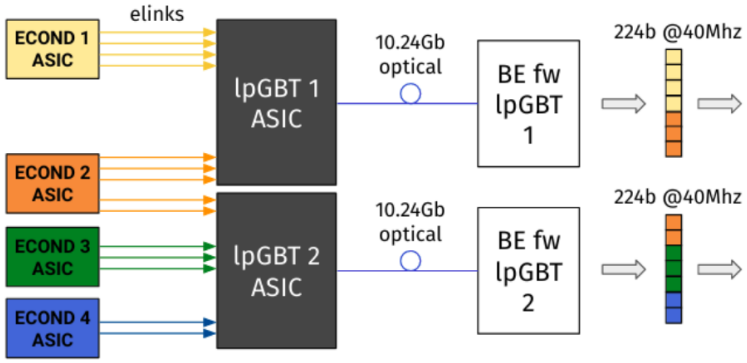


Figure 6.11: An illustration of ECON-D to lpGBT mapping. [149].

The main block of the DAQ firmware is the capture block, which processes tasks such as packet detection, buffering, and event construction of the ECON-D data. Each DAQ-FPGA contains 54 Capture Blocks, with each one processing data received from 2 lpGBT ASICs. Since a lpGBT ASIC can be linked up to 6 ECON-Ds simultaneously as discussed in Sec. 6.1.1, the DAQ FPGA is designed to handle and process data from $6 \times 2 \times 54 = 648$ ECON-Ds. To handle the cases where large data rates are seen by modules, ECON-Ds might distribute their data across two lpGBTs. Hence, a detailed pairing of lpGBTs for data transmission is necessary at the DAQ level. The input to the DAQ is the event packet in 32-bit words. The DAQ saves this information in the main buffer as a 64-bit word corresponding to the two 32-bit ECON-D words after detecting and assembling the packet. It also adds 3-bit status codes for all ECON-D packets and the local event number, BX, and orbit counter values.

The data packets received at the DAQ pass the CRC check, and they are forwarded to the capture block in the DAQ firmware for further validation. The capture block performs various checks such as verifying event identification counters for inconsistencies, e.g., BX counter mismatches between FE and BE or event ID mismatches, where the received packet does not correspond to the L1A issued by the BE (from TCDS2). These validations ensure the system's reliability, confirming that the event packet corresponds to the correct L1A and that the system is functioning correctly. If

no discrepancies are detected, the header information of the event packet is updated accordingly. Depending on whether the DAQ is in pass-through or veto mode, it either forwards the event packet with a flag or sends only the headers. The event packet is then transmitted to the Slink and subsequently to the DAQ-PC via 10 Gbps links, or routed through the DTH to the DAQ-PC. More details about these status codes and the significance of L1A counter values are added in Section 6.6. Once the event packet is formed for all the ECON-Ds sending their data via lpGBT links connected to DAQ, it stores all this data in the respective event buffers for ECON-D which are read by Slinks.
If the CRC check fails, the packet is discarded and a timeout packet is sent to the Slink/PC. This timeout packet consists of three 128-bit words: a Slink header and a DAQ header indicating that the DAQ did not receive the event packet within the programmable time, either due to non-arrival from the FE or a CRC failure.
Twelve Slinks are connected to transfer data to DTH board. Efficiently mapping these 54 event buffers corresponding to capture blocks to the 12 available Slinks presents a challenge due to the limited routing capabilities of the FPGA [149]. This requires a dynamic load balancing optimizing solution to handle all scenarios while satisfying timing constraints.

TPG system

The HGCAL TPG system is designed to process trigger primitives and interface with the central CMS Level-1 Trigger system at the LHC clock rate of 40 MHz. The L1 trigger system decides whether to capture a full event based on these inputs. The TPG system operates in two stages:

- TPG Stage-1: This stage receives raw data from the front-end, identifies trigger cells, and computes individual module sums to partially construct tower energies.
- TPG Stage-2: This stage performs clustering of trigger cells into three-dimensional cluster objects, calculating their positions, energies, and shapes.

The system efficiently processes and forwards these trigger primitives efficiently, enabling the Level-1 trigger to make informed decisions on event capture. More details on this can be found in [137].

In this chapter, validation of the readout system is presented based on (a) simulated collision events using a full-scale emulator of the HGCAL FE-BE system and (b) testing the full readout chain in beam test experiment and the test stand of the system in a laboratory.

6.1.3 HGCAL buffer emulator

As explained in the previous section, the HGCAL BE readout system is a complex readout system which needs to handle not just a large amount of data but also deal with large variations in data volume in different sections of the detector depending on incoming particle flux and geometric locations. This makes thorough validation of system design, number of various components and configurations imperative before the various parameters are frozen for the final production. To make these studies, we developed a C++ based software emulator (i.e. hardware simulation) of the readout system that takes simulated pp collision events as the input and models the responses of various components in the readout chain, mainly focusing on FE⇔BE chain. The

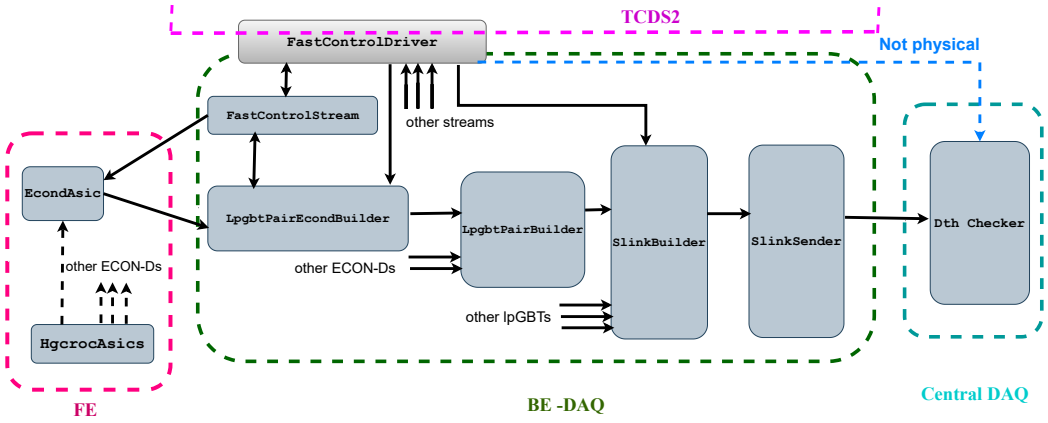


Figure 6.12: Schematic illustrating the different entities of the HGCAL buffer model emulator.

studies presented in this section are based on $t\bar{t}$ pair production events simulated with 200 PU and noise at 3000 fb^{-1} conditions using CMS software (CMSSW) as input. The objectives of emulator-based studies are to estimate & validate resource usage, buffer occupancies, deadtime & data losses in the readout chain and identify their sources, and to study the impact on buffer performance for choices of system parameters such as number of links and data transmission speed, FPGA clock speeds, etc.

The emulator is based on a total of 8 entity classes represent FE & BE components such as fast timing and control (TCDS2), FE HGCROC and ECON-D, BE FPGA and DTH receiver. Fig. 6.12 is shows its overall structure, and connections represent interfaces between these entity classes. It provides the following mappings and connections in the FE and DAQ components:

- FE - ECON-D and HGCROC ASICs
- TCDS2 - Distribution of fast control streams
- BE DAQ - DAQ firmware specifics are not included. The interface from lpGBT to Slink is included.
- Central DAQ via DTH - storing the information for offline analysis with the help of this class.

Connection indicated as "Not physical" does not exist in the actual readout chain but is added to the emulator to complete the picture.

The basic idea of emulator studies is to model the collection of data in each of the slinks. Event data building starts with preparing the event packet for all ECON-Ds (32-bit words) connected to a lpGBT (72-bit words) via elinks, after which the packet is formed for all lpGBTs on a Slink (128-bit words). Each Slink is designed as an independent DAQ system unit that receives input from multiple lpGBT links. An interface to `HgcrocAsics` allows input of simulated hit numbers which are correlated across all ECON-Ds within the same event. The `ECON-DAsic` class determines the size of the data packet for each L1A trigger and includes the flexibility to switch to different hits values for the calibration and non-zero suppression (NZS) modes. Fast control-specific classes have been added to handle synchronization, send L1A triggers, and manage resynchronization. The reset signal as well as the L1A trigger throttling are sent via fast control streams. Fast control-specific classes have been added to handle synchronization, send L1A triggers, and manage resynchronization, resets, and L1A trigger throttling. `FastControlDriver` class incorporates a

realistic HL-LHC orbit structure, including the abort gap (e.g. during non-collision periods), and only realistic combinations of fast commands are supported.

The numbers and connections between each ECON-D to lpGBT (FE mapping) and lpGBT to Slink (BE mapping) may vary across the HGCAL, as discussed in Section 6.1.2. We define this realistic mapping of FE and BE through a class `Definer`, which takes this mapping file as a configurable parameter. This allows us to study the impact of different mapping choices on buffer occupancies. Buffer limits for different ASICs are also implemented and flagged in data packets, enabling throttling through the `ECON-DBeBuilder` → `FastControlStream` → `FastControlDriver` path.

For each clock tick, each component updates its state based on clock edges (typically, the rising edge for most hardware designs). The transitions between different states across the entities are also done at the designed clock speed for that entity. Hardware time latency and delays are added by hand in the emulator to accurately simulate the data transmission flow from the FE to BE. Our major aim is to keep a tab on the different buffers and their occupancies indicated by the metadata added to emulate the ASICs response. The emulator is designed with minimal hard coded configurable parameters such that different parameters and their impact on buffers can be studied. For example, one can decide the L1A trigger rate, trigger type, or which mapping to implement, etc. depending on the studies to be performed.

The L1A triggers are issued from the `FastControlDriver` class. These can be issued in a random or periodic way. In a random way, the trigger is issued for a BX if the random uniform number generated is higher than the probability of L1A calculated according to the designed trigger rate and the bunch structure. In the periodic method, L1A triggers are issued on equidistant BXs such that the average rate is as per the user-configured trigger rate. `FastControlDriver` class also keeps a count of how many throttle signals were sent or HGCROC buffer overflow, how many L1As received and sent. Certain entities, such as `FastControlDriver`, include a `Analyser` class that can be enabled at the start of the program. This class calls a method at the end of processing for each clock tick of the associated entity. Performance of the configuration emulated is measured in terms of different variables e.g. the total size of the event, the number of packets dropped due to limited buffer efficacy. To prevent analysis from adding more challenges computational along with the core algorithms, these `Analyser` classes containing several histograms as different validation checks can be toggled on or off as needed.

In the emulator, we simulate the entities for a complete 120° HGCAL sector. We have approximately ≈ 5 k silicon modules per sector. It takes approximately 11 hours to run the emulator for 1 M BXs. Each module is identified based on its layer (l) and transverse (u and v) coordinates as shown in Fig. 3.16 and we define an unique `luv` variable to identify the position of module in the HGCAL as per the following equation,

$$luv = 256 \times l + 16 \times u + v \tag{6.1}$$

Higher the luv number, more depth the module is positioned at. Due to the symmetry ϕ, we can multiply it by 6 to obtain the total estimated data rates.

The focus of the studies presented in this chapter is on the silicon-specific readout details and estimation. The scintillator tile-boards have not been discussed at length. However, short discussions on key differences are summarized in the text.

Building data packets at emulator

To simply simulated physics events plugin, the digitized signal obtained from the simulation for each cell is converted to one of the 10-bit, 20-bit or 30-bit information depending on the energy deposited and these are defined below:

- 10-bit : ADC or TOT information (10-bit).
- 20-bit : ADC or TOT (10-bit), and energy deposited from previous BX, ADC (BX-1) (10-bit).
- 30-bit : ADC or TOT (10-bit), ADC (BX-1) (10-bit) and TOA (10-bit).

A cell with this signal information along with its position coordinates is referred to as a hit. Module occupancy is defined as the total number of such hits. We define occupancies *n10*, n20 and *n30* for a module as the total number of cells with 10-bit, 20-bit or 30-bit information as:

- *n10* : this represents the total number of 10-bit hits per module. Each of *n10* hits are actually of 16 bits, including 10 bits for ADC or TOT, 4 bits reserved for code, and 2 extra bits.
- *n20* : this represents the total number of 20-bit hits per module. Each of *n20* hits has 24 bits including 20 bits for ADC (TOT) + ADC (BX-1) and 4 reserved bits.
- *n30* : this represents the total number of 30-bit hits per module. Each of *n30* hits has 32 bits including 30 bits for ADC (TOT) + ADC (BX-1) + TOA and 2 bits for the TOT presence code.

These information obtained from simulation is fed as the input to the emulator via `HGCHitProvider` class. This class also applies necessary corrections to the hits which are discussed later on in this section.
To emulate that the packet formatter at ECON-D as discussed in Section 6.1.1, we combined these different *n10*, *n20*, *n30* hits to format the ECON-D data packet for that module in units of 32-bit words. This is done using the class `ECON-DAsicHitCounts`. This step converts hits into number of words using Equation 6.2 [150];

$$N^{i}_{words} = n^{i}_{half} \times (2 + \frac{(16 \times n10^{i} + 24 \times n20^{i} + 32 \times n30^{i} + 31)}{32}) \tag{6.2}$$

where, N^{i}_{words} is the total number of 32-bit words for i^{th} module, n^{i}_{half} is the number of halfHGCROCs (see 6.1.1) connected to that module, $n10^{i}$, $n20^{i}$ and $n30^{i}$ are the number of *n10*, *n20* and *n30* hits for i^{th} module. Summing up these N^{i}_{words} over all modules in the HGCAL gives the estimated data rates per event of the HGCAL. Such 32-bit words are then formed into 64 or 128 bits of words. The `ECON-DAsic` class has the ability to scale the *n10*, *n20*, *n30* hits according to different modules.

6.2 Modeling occupancy differences in different geometry versions

The number of detector cells or *hit occupancies* with one of the 10-bits, 20-bits or 30-bits of digitized information about signal measured per cell for each module are referred to as $n10$, $n20$ and $n30$, respectively and total occupancy of a module is defined as

$$\begin{aligned} n &= n10 + n20 + n30 \\ n &\leq \text{total number of cells in the given module} \end{aligned} \tag{6.3}$$

The higher the occupancy, the more data is transmitted from that module. As described in the previous section, the BE emulator takes these hit occupancies as input. To ensure system performance in a realistic collision environment, $t\bar{t}$ pair production events overlayed over 200 pileup interactions simulated for detector response and digitization effects using `GEANT4` included in the `CMSSW` framework are used for studies presented in this section.

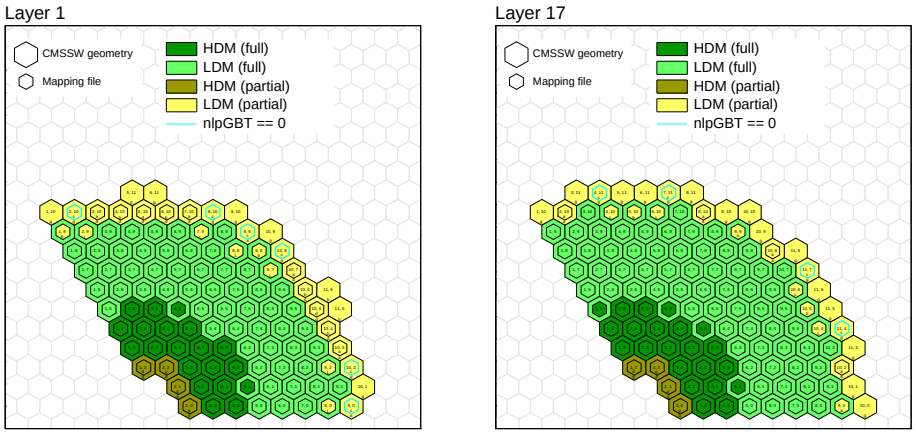


Figure 6.13: Comparison of module types and locations in CMSSW reference (larger hexagons) and a baseline updated (smaller hexagons) HGCAL geometries. The number in the hexagons are the (u,v) coordinates for that module.

The HGCAL geometry used in this simulation is referred to as *v11*. However, the detailed detector geometry, i.e. configurations of different module types and locations, have been evolving continuously based on experiences gained through optimization studies for physics performances, as well as with detailed testing of prototype versions of various ASICs and other components in laboratory setups.

Various versions of updated geometry can differ in the number of HDMs, LDMs, and partial modules and their locations, as illustrated in Fig. 6.13 for a few HGCAL layers. There can also be additional or missing modules in various geometry versions. The larger and smaller hexagons represent modules in the CMSSW reference and updated baseline geometries, respectively. The dark green (light green) and brown (yellow) hexagons represent HDM full and partial modules, respectively. The geometries considered here include 50 sampling layers, not the final design. These HDM $\rightarrow$ LDM discrepancies exists for the first 38 layers of HGCAL.

The hit occupancy per module depends on

- Number of channels in module. For example, an HDM (432 channels) will have more than twice the hit occupancy as compared to an LDM (192 channels) at the same position in the detector.

- Distance from the interaction point. Particle densities are much higher closer to interaction point, hence more cells of the modules situated close to interaction point or beam pipe are readout as shown on left and middle distributions in Fig. 6.21 for two modules positioned very close and far away from the interaction point.

Hence, we end up underestimating (if HDM $\rightarrow$ LDM) or overestimating (if LDM $\rightarrow$ HDM) the data rates. To keep up with these changes in geometry closely and in time, we developed a model to extrapolate hit occupancies obtained for module from $t\bar{t}$ events as reference to those expected from updates in HGCAL geometry, called *target geometry* . This method is discussed in the following text. Only full-shape modules are considered for this extrapolation. The study excludes partial-shaped modules due to insufficient data and varying occupancy slopes. 300 μm sensors, usually far from the interaction points ($R > 1000$ mm) or in specific transition regions, are not the focus of the study. The analysis targets the first 38 layers of HGCAL, excluding layers lacking adequate HDM module representation.

6.2.1 Methodology for HDM $\leftrightarrow$ LDM occupancy transformation

Variation in hit occupancy for a module as a function of its distance from the interaction point (R) varies smoothly for each layer, as shown in 6.14 for HDM (black markers) and LDM (green markers) for CMSSW reference.

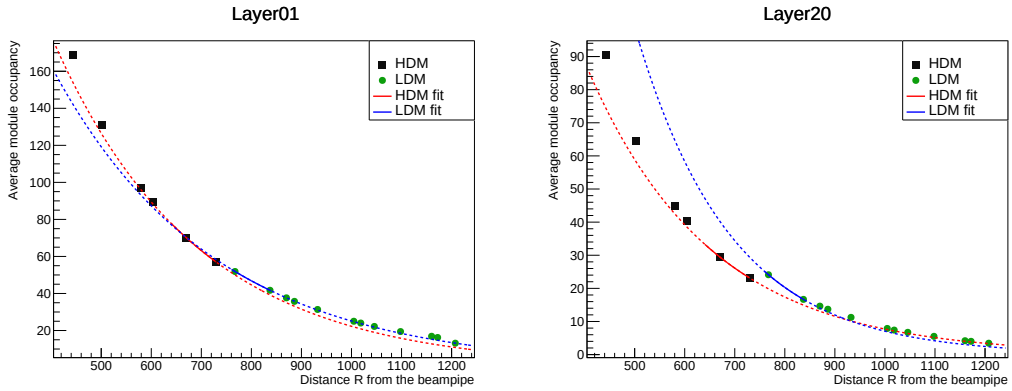


Figure 6.14: Average occupancy (Y axis) of modules as a function of its distance from the interaction point (X axis) across different layers of HGCAL. Only the modules with difference in geometries are shown.

However, the two curves follow different slopes. Both curves are separately fitted with an exponential function with two parameters `exp(p0+p1x)`. Range of fitting is limited to HDM $\leftrightarrow$ LDM transition region which is where most of the discrepancies are present. The behavior of the fitted function in the very high and low R regions is not considered in the context of this exercise. These functions can now be used to evaluate the average occupancies of modules at a given R and the same function can be used to correct for the occupancies of HDM and LDM modules by extrapolating in the neighborhood of their range. The overall occupancy decreases far from the interaction point and the difference in occupancy between HDM and LDM decreases. As outer radii region see less radiation fluence as compared to inner radii ones, hence sparse true energy deposits and signal become comparable to the noise.

A summary of the average module occupancy obtained by the fitted functions as a function of the module position represented as *luv* variable defined in equation 6.1 for the LDM and HDM modules is shown in Fig 6.15 (top left). The ratio panel shows the difference between the two occupancies for LDM normalized with the statistical

error on the reference occupancy. The two values match closely over the full range of *luv*s. For reference, the HDM occupancies evaluated from the fitted function are presented in the same figure. An interesting observation is that for very large R, the predicted occupancies for HDM are smaller than for LDM, which is counterintuitive given the larger channel density. This is due to increasing noise dominance with distance from the interaction point where lower noise levels in HDMs as compared to LDMs become evident.

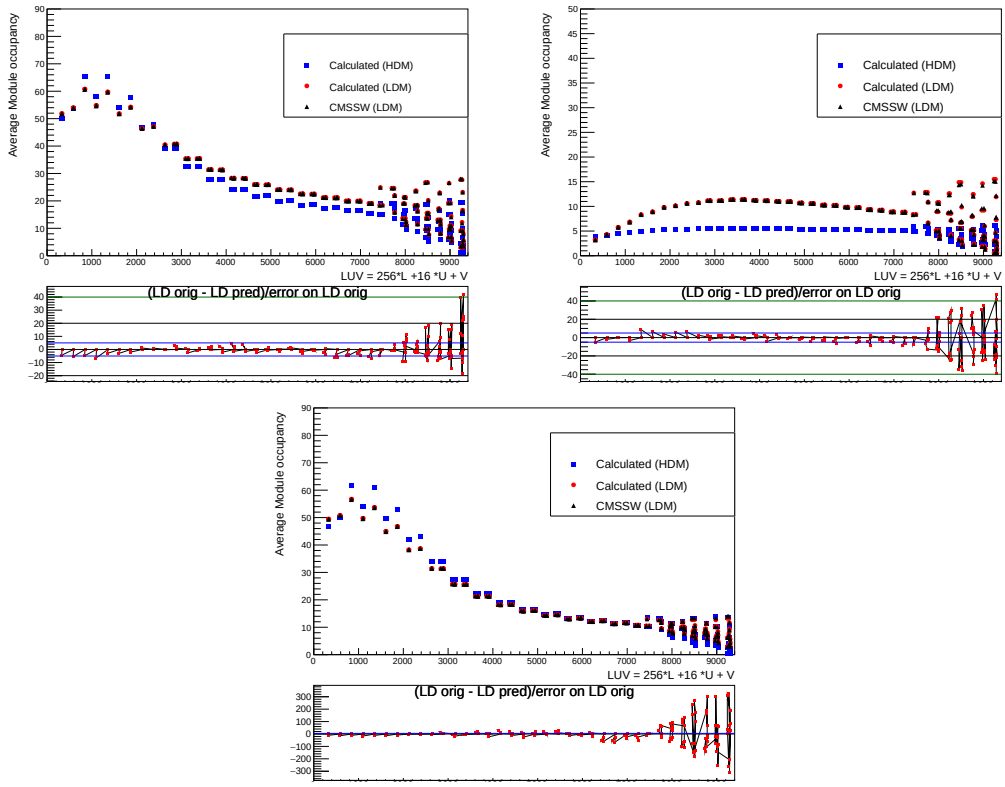


Figure 6.15: The above figure illustrates the comparison of calculated occupancies (Y-axis) for LDM with that obtained from CMSSW for total signal (top left), only noise (top right) and signal (bottom). X axis is indicating its position in luv variable as defined in equation 6.1.

To verify and understand this feature, we studied the effect of noise using a neutrino gun event sample simulated without any pileup. Since neutrinos do not leave a measurable signal in the detector, occupancy in modules is dominated by noise contributions. The average module occupancy (average noise) as a function of R is fitted with the same procedure as described above. The top right plot of Fig. 6.15 illustrates the mean noise per module with its position, indicated by the *luv* variable 6.1, for both the LDM and HDM modules. The ratio panel shows that the discrepancy in LDM noise obtained from the fitted model versus CMSSW, adjusted by the statistical uncertainty on the reference occupancy, indicates a consistent correlation throughout the full *luv* range. The silicon modules' average noise-induced occupancy remains fairly constant throughout the HGCAL, with only minor deviations at the final layers.

In real experiments, the measured signal is a combination of true signal overlaid over some noise contribution. Noise cannot be fully eliminated or subtracted. Hence, we follow a probability-based approach to assess the occupancy behaviors in the presence of only signal. We define p_n (p_s) as the probability of a noise (signal) hit

when there is no signal (noise), such that

$$p_n = \frac{N_n}{N_T} \qquad P_{hit} = \frac{N}{N_T} \tag{6.4}$$

where N_n and N are the average hits obtained from the neutrino and $t\bar{t}$ samples, respectively and N_T are the total number of cells in the module. The probability of no hit ($P_{!hits}$) in a cell can be defined as;

$$P_{!hit} = (1 - p_n)(1 - p_s) \tag{6.5}$$

$$P_{hit} = 1 - (1 - p_n)(1 - p_s) \tag{6.6}$$

Solving for p_s

$$p_s = (P_{hit} - p_n)/(1 - p_n) \tag{6.7}$$

Hence the average signal occupancy (N_s) after subtracting noise is:

$$N_s = N_T \times p_s = (N - N_n)/(1 - p_n) \tag{6.8}$$

Thus obtained average signal occupancies curves as a function of distance R for different modules are fitted in the same manner as discussed in the beginning of this section. The lower plot in Fig. 6.15 provides an overview of the occupancy results derived from the fitted curves for HDM and LDM, along with that for LDM from CMSSW. The LDM signal predicted using the fits aligns closely with that obtained from CMSSW, remaining consistent within the margin of uncertainty. It is observed that the signal recorded in HDM and LDM exhibit notable similarity in the subsequent layers.

Uniformly extrapolating HDM ↔ LDM occupancy

Instead of going through 100's of fit parameters, we can use analytical approach and use an unanimous function to transform HDM ↔ LDM for a given HDM (LDM). Three such methods are described below to scale up or down the occupancies obtained from `CMSSW` *V11*, N, to that from target geometry, N';

- Fixed Occupancy : In this model, we assume the fraction of hits above the threshold is identical in both density types, representing hits from localized multi-MIP groups, and occupancy transforms as

$$N^{'} = \frac{192 \times N}{432} \tag{6.9}$$

- Fixed Number : This model is very optimistic. It assumes the number of hits above the threshold is the same, corresponding to hits from single MIPs only, resulting in

$$N^{'} = min(N, 192) \tag{6.10}$$

- Fixed Density : Here, we assume that MIPs have a uniform probability density proportional to the area. For uniform density in an event with average λ number of MIPs per unit area, probability of no hit for a cell with area A is given by

$$P = e^{-A\lambda} \tag{6.11}$$

Using a threshold occupancy N_T = 192 for LDM and 432 for HDM and N as reference CMSSW occupancy, we can define λ as;

$$e^{-A\lambda} = \frac{N_T - N}{N_T} \rightarrow A\lambda = ln[\frac{N_T}{N_T - N}] \tag{6.12}$$

So, to transform to the occupancy N′ for modules with different density & area A′,

$$e^{-A^{'}\lambda} = \frac{N_T^{'} - N^{'}}{N_T^{'}} \tag{6.13}$$

Putting this back in equation 6.12, we obtain,

$$N^{'} = \frac{N_T^{'} \times N_T^{'} \times N(1 - [(N_T - N)/N_T])}{N_T} \tag{6.14}$$

Figure 6.16 presents the predicted occupancies when employing the three models to transform LDM occupancies to HDM, positioned equidistantly from the interaction point as well the average signal occupancies obtained using the fits. The fixed number method yields a more pessimistic estimate, while the fixed occupancy approach appears to be overly optimistic and does not align with actual conditions. In contrast, the fixed density method is regarded as the most realistic approach as the signal calculated according to Equation 6.8, after removing noise contributions, aligns with the trend of the fixed density model. Therefore, the fixed density methodology is

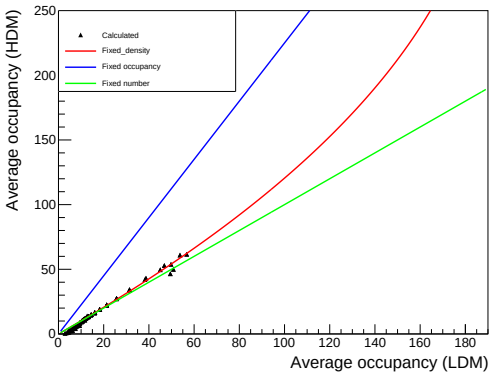


Figure 6.16: The transformation of occupancies from LDM to HDM predicted by different analytic approaches and obtained using the fits is shown.

employed to convert LDM occupancies from CMSSW to HDM (and vice-versa), facilitating an in-depth analysis of average data rates and buffer occupancies.

6.3 Pseudo event generation for enhancing the event statistics

It has already been mentioned in earlier sections that the flux of particles traversing the detector depends on the distance from the interaction point and direction with respect to the beam. There are further complications such as high energy jets produced in pp collisions can cause high instantaneous occupancies and increased data rates in neighboring modules. If these modules share the same lpGBTs for data transmission, it could result in data losses which manifest as fake missing transverse energy (MET), which is a new physics (e.g. dark-matter) signature. While the available buffer depths of readout ASICs, total number of links and bandwidth per link, mapping of lpGBT links to ECON-D have been optimized to balance average data rates observed by the ECON-D or module, the current design choices should be rigorously tested against such rare scenarios. We need to run the software-based emulator, as detailed in section 6.1, for extended durations and with high statistical samples to capture rare high-occupancy events ($< 10^{-7}$). To avoid complete detailed simulations of a few billion collision events, we employ a method using multivariate Gaussian random numbers [151] to generate pseudo-events based on empirical occupancy distributions (ED) obtained from CMSSW $t\bar{t}$ events while preserving correlations among the energy deposits in neighboring modules. This section describes the method for generating these pseudo-events and validating the methodology. We are using the latest HGCAL geometry implemented with 47 sampling layers in `CMSSW` for these studies.

6.3.1 Multivariate Gaussian distributions

A multivariate Gaussian distribution in N dimensions [151], also known as a multivariate normal distribution, extends the concept of a one-dimensional normal distribution to higher dimensions and can be defined according to equation 6.15.

$$P\left(X_N = x\right) = \frac{exp(-\frac{1}{2}\left(x-m\right)^T S^{-1}(x-m))}{(2\pi)^{\frac{N}{2}}\sqrt{det(S)}} \tag{6.15}$$

where x is an N-dimensional vector of the random variable, m is the vector of length N containing mean values and S is the covariance matrix. It represents a continuous probability distribution for a vector of random variables each of which follows a normal distribution. The dependencies and interactions between these variables are captured by a covariance matrix, which encodes both the variances of individual variables and the correlations among them.

To generate a set of random samples jointly following the X_N probabilility distribution, a vector $\mathbf{r}$ of independent standard normal samples like (r_1 ,r_2 ,r_3 ,....) is generated and required correlations are introduced by multiplying this vector by a lower triangular matrix A, constructed so that the covariance matrix S satisfies the equation 6.16 [151].

$$S = AA^T \tag{6.16}$$

Finally, the means of the resulting vector are adjusted by adding the mean vector m. This process gives the k^{th} sample x_k as

$$x_k = A\, r_k \, + \, m_k \tag{6.17}$$

where x is the resulting vector of correlated random variables, r is the vector of independent standard normal random variables and m is the vector containing the mean values [151]. Using equation 6.17, we can generate correlated random variables

x_k based on uncorrelated normal random variables r_k. If we consider A to be the lower triangular composition of correlation matrix $S^{'}$, the equation 6.17 can be rewritten as per the equation below,

$$x_k = A\, r_k \tag{6.18}$$

With the correlation matrix, we do not need to add the mean of samples.

6.3.2 Addressing the challenges involved in HGCAL pseudo event generation

Using the method described in section 6.3.1 for generating pseudo events for HGCAL with proper correlations among module occupancies is nontrivial because of large number of cells, non-Gaussian occupancy distributions, which may have long tails, and computationally demanding.

Correlation patterns among HGCAL modules

Electromagnetic and hadronic showers propagate through multiple detector cells, which may belong to different modules, spreading in both longitudinal and transverse directions. As a result, the occupancies of neighboring modules are correlated within the same event, leading to correlated data rates among these modules. Capturing a large number of correlations with approximately 5 k modules arranged in 120° sectors, each with three different occupancy formats (n10, *n20* and *n30*), requires a total of ≈15 k variables to model and therefore, constructing a 15 k × 15 k correlation matrix.

The correlation (ρ_{xy}) is a normalized measurement of the degree to which the modules are correlated and is defined as

$$\rho_{xy} = \frac{Cov(x,y)}{\sigma_x \sigma_y} \quad \text{where} \tag{6.19}$$

$$Cov(x,y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) \tag{6.20}$$

Here, x_i and y_i are occupancies in ith event, $\overline{x}$ and $\overline{y}$ are the average occupancies for modules x and y, respectively and n is the total number of event samples used. We will be using ρ_{xy} to quantify the correlations between the modules. It can take values between -1 and 1. The higher value indicates a strong variation of the occupancies with respect to each other.

The plots on the left and right of Fig. 6.17 show absolute correlations for a module (l,u,v = 1,5,3) with its other neighboring modules within layer 1 and with modules of layer 30 indicating a high correlation with its neighboring modules both in the transverse and longitudinal directions. The *n10* and *n30* formats also have a strong correlation, and *n20* is weakly correlated, attributed to their definition as in Section 6.1.3.

The 5 k×5 k correlation matrices of the correlation of all modules among themselves are shown in Fig. 6.18 for the *n10*, *n20* and *n30* formats. Only the entries with correlation >10% are shown in these figures. We can observe that the correlation is non-negligible for most of the modules, and these correlations need to be preserved while generating the pseudo events. The lower triangular matrix representation of this correlation matrix is used to correlate the random samples generated according to Equation 6.18.

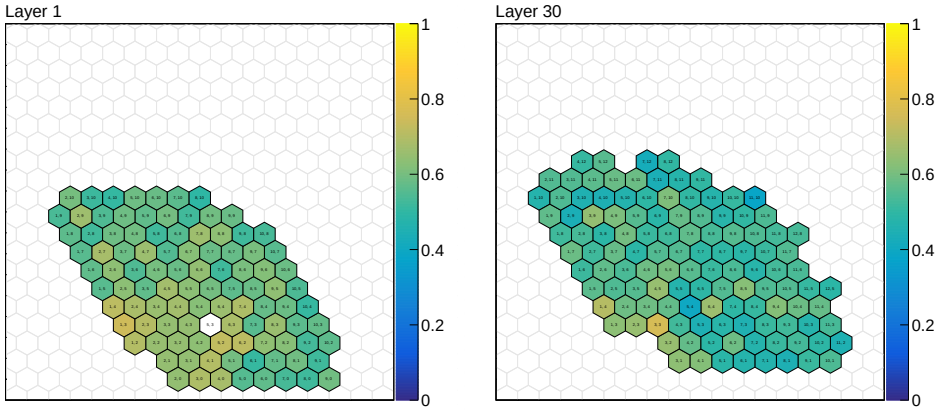


Figure 6.17: Correlations for a module (l,u,v = 1,5,3) with neighboring modules in layer 1 (left) and with layer 30 (right).

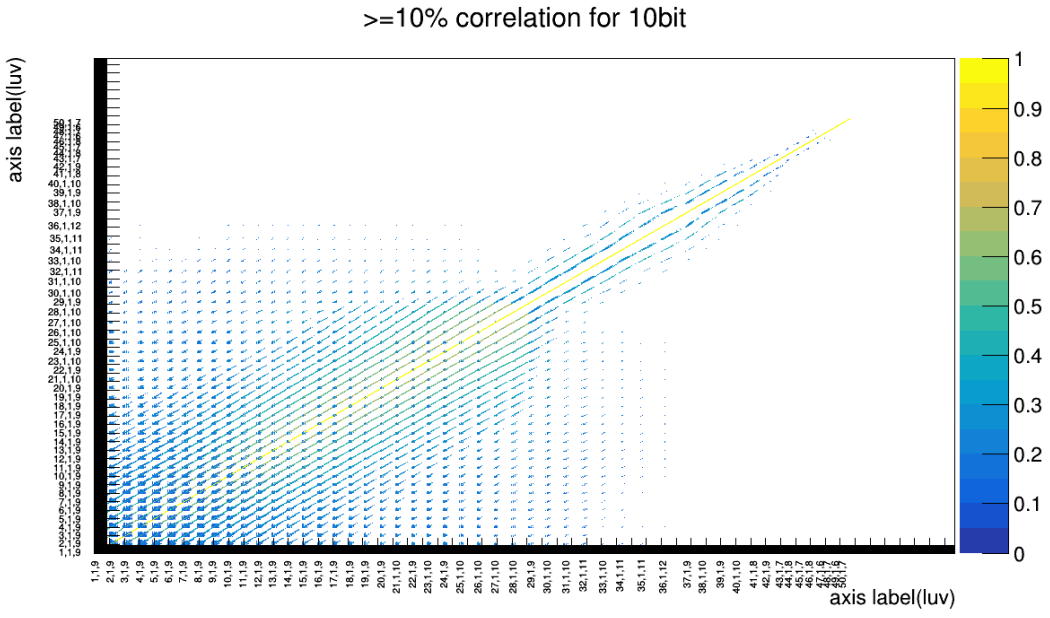


Figure 6.18: A15 k×15 k correlation matrix for module, displaying only elements with at least 10% correlation.

Quantile mapping for shaping empirical distributions into normalized distributions

Quantile mapping is an effective method to transform a numerical input variable to a different data distribution like Gaussian or uniform. It transforms the input to a desired distribution, regardless of whether the probability distribution function (PDF) is known or not. It uses a cumulative distribution function (CDF), which is defined as the total probability that a random variable takes on a value less than or equal to a specified value. The quantile function returns the value corresponding to a given probability, essentially mapping the probabilities back to their associated data values. An example of how quantile mapping is performed to transform a model to observed data using PDF is shown in Fig. 6.19 (left) and using CDF in Fig. 6.19 (right). It is a commonly employed method for Monte Carlo (MC) data corrections.
We use `QuantileTransformer` from the `scikit.learn` [152] to morph occupancy distributions into normalized Gaussian distributions based on the CDF for the number of quantiles to be computed. The idea is to correlate the random samples according to equation 6.18 in the normalized Gaussian phase space obtained after morphing occupancy distributions obtained from CMSSW using Quantile transformation (QT).

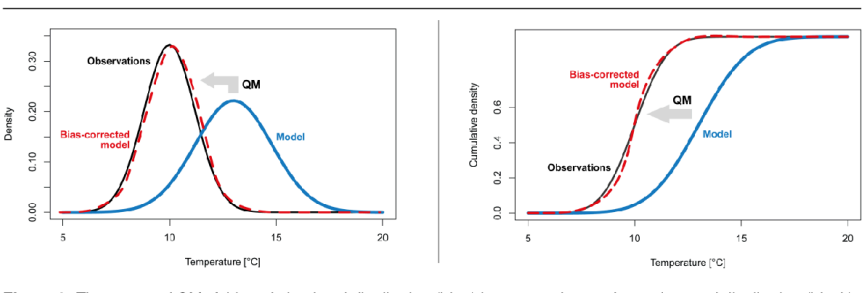


Figure 6.19: This example shows how a biased simulated distribution (blue) is adjusted to match the observed distribution (black) using PDF (left) and CDF (right). The corrected distribution (dashed red) closely resembles the observed distribution, although exact matching might be hindered by sampling uncertainty or QM implementation specifics [153].

The QT is a nonlinear transformation. The plots in figure 6.20 show the correlations of the module at (1,3) with the module at (10,5) in terms of occupancies: from CMSSW (left), after quantile transformation to Gaussian (middle), and after inverse QT back to the original space (right). These similar correlation levels indicate that QT preserves the original correlations, as well the original distributions can be obtained using inverse QT.

Function-based tail population for empirical distributions

The occupancy distributions obtained from a statistically limited event sample generated using CMSSW do not always represent potential events in the tails, and hence, these tails do get populated when sampling a large number of random samples from them. To overcome this limitation of reference samples, we fit suitable parametric functions to the occupancy distributions obtained from CMSSW and the random samples from these fitted functions. To model Gaussian and non-Gaussian (or Poisson) components of the occupancy distributions across the HGCAL, we fit all the distributions individually with a combination of the Poisson, Gaussian and Landau

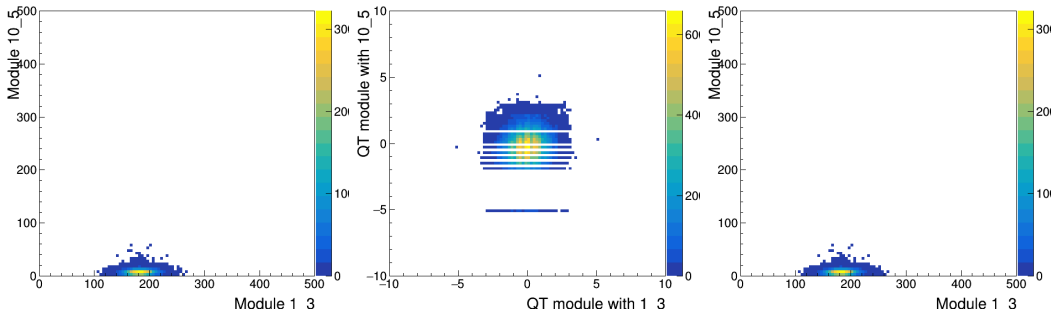


Figure 6.20: The plots show the correlations of the module at (1,3) with module at (10,5) in terms of occupancies: from CMSSW (left), after quantile transformation to Gaussian (middle), and after inverse QT back to the original space (right).

functions:

$$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} + \frac{\lambda^k e^{-\lambda}}{k!} + \frac{1}{\sigma_L} L\left(\frac{x-\mu_L}{\sigma_L}\right) \tag{6.21}$$

where x is the occupancy. The μ and σ are mean and sigma parameters for the Gaussian function, the λ and k are the parameters for the Poisson function, and the μ_L and σ_L are the most probable and scale parameters for Landau function. The Landau

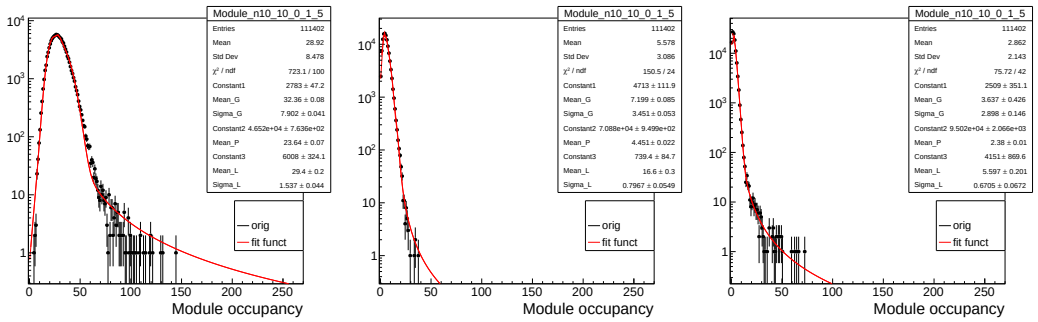


Figure 6.21: The occupancy distributions for different modules along with the different parameters obtained after fitting it as per equation 6.21 are shown.

function is mainly included to allow longer tails representing larger occupancy which are not easily represented by Gaussian PDFs. This procedure is carried out for each of the distributions of *n10*, n20 and *n30* formats per module, and one such example for a module is shown in Fig. 6.21, indicating that the combination of functions effectively fits various occupancy distributions, regardless of their shape.

By sampling random numbers from these fitted functions, we achieve the desired population in the tails along with high statistics. In some modules, the Landau distribution more accurately represents empirical distributions (EDs), leading to non-dying tails. When we sample large numbers of random values from these fits, we sometimes get events where a module's occupancy exceeds its total channels, creating nonphysical situations. To address this, we limit the sampled random numbers to the range [0, N_T], where N_T is the number of channels in the module. This method ensures that we have high uncorrelated samples statistics with well-populated tails across the three formats. To create realistic event scenarios, we need to correlate these samples using the method outlined in Section 6.3.1.

6.3.3 An optimized method for Pseudo Event Generation in HGCAL

In addition to addressing the above challenges, the method of generating pseudo-events needs to be optimized from the system memory perspective. The correlation matrix has dimensions of 15 k × 15 k and matrix multiplication dimensions would be 15 k × 1 M to generate 1 M events. Such large matrix multiplications need to be done in batches as doing it all at once is limited by the available CPU memory. These calculations are time-consuming. In addition, the quantile transformation and its inverse should be applied simultaneously to avoid creating and saving a large quantile transformation matrix 15 k × 1M. Hence, the procedure to generate large statistics pseudo-event needs to be optimized further, keeping in consideration the available CPU memory limitations. We summarize the procedure we used to generate N correlated pseudo-events for the studies presented here:

1. Calculate the lower triangular component A of the correlation matrix S' using the Cholesky decomposition of a matrix from scipy.

2. Generate N random numbers following normalized Gaussian distributions. Correlate these random numbers by multiplying them by the lower triangular component A of the correlation matrix, as done in equation 6.18 and save these outputs on disk. This step reduces the memory consumption in the matrix multiplication, however, it increase the storage needed to save these intermediate files.

3. A set of random number variables ($r_k^{'}$) are sampled from the fitted function obtained from section 6.3.2 for each module and the three formats, separately.

4. These random numbers ($r_k^{'}$) are distributed into normalized Gaussian phase space (r_k) using quantile transformation as discussed in section 6.3.2. The inverse quantile transformation is applied to the correlated Gaussian random numbers obtained from step (2) listed above, and referred to as x_k. These x_k are the correlated pseudo-events.

For our studies presented here, we generated (N = 1 M pseudo events starting with 110 k events generated using CMSSW. The most time-consuming step is (2), followed by step (4).

Comparisons of occupancy distributions obtained using the method above with those obtained from the CMSSW for the *n10*, *n20* and *n30* formats for one module are shown in Figs. 6.22, from left to right, respectively. We observe that the bulk distributions agree and events are added in the tails which are not original present.

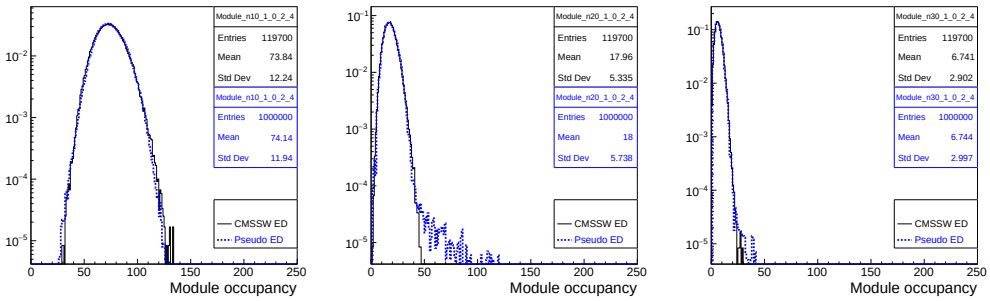


Figure 6.22: Shapes of *n10* occupancy distributions from CMSSW & from Pseudo event generation.

Fig. 6.23 and Fig. 6.24 show average occupancies and standard deviation obtained from pseudo-events versus those obtained from simulation, respectively, for *n10*, *n20*, and *n30* formats. The average occupancies are similar for the two event samples. The deviations observed in standard deviation are attributed to the fact that we use the Landau function to fit the empirical distributions. This results in extended tails which contribute to the standard deviation, especially for the modules positions away from the interaction point.

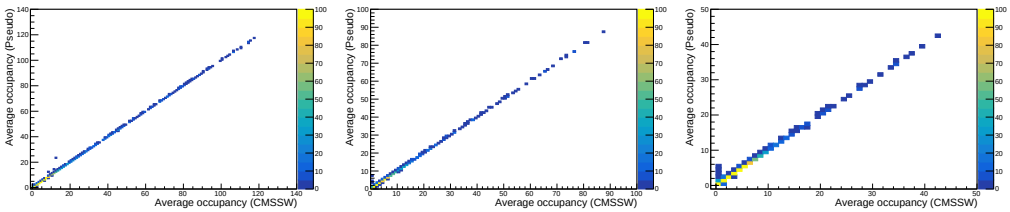


Figure 6.23: The comparison of the average occupancy obtained from pseudo events (Y axis) and CMSSW (X axis) for *n10*, *n20*, *n30* formats shown from left to right, respectively.

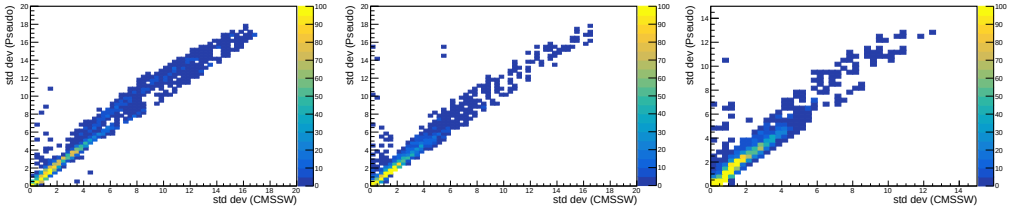


Figure 6.24: The comparison standard deviation of occupancy distributions obtained from pseudo events (Y axis) and CMSSW (X axis) for *n10*, *n20*, *n30* formats shown from left to right, respectively.

The emulator, as discussed in section 6.1, is employed to register 1000 events at a frequency of 750 kHz using both CMSSW simulated $t\bar{t}$ events and pseudo-events. The aggregate event size is determined by summing the data size corresponding to the *n10*, *n20*, and *n30* occupancies of all modules in 120°, in accordance with Equation 6.2. A comparison of these sizes for CMSSW events, pseudo-correlated and uncorrelated events, is illustrated in Fig. 6.25. The uncorrelated pseudo-events are derived from random numbers sampled from the fitted function as delineated in step (3). On average, the mean event size is consistent across the three formats; however, the standard deviation (std dev) exhibits variation. Predictably, the std dev for the non-correlated events is minimal, approaching zero. The std dev of the event size distribution is more pronounced for pseudo-events than for CMSSW events. This observation can be attributed to intentional extended occupancy tails in pseudo-events, which may lead to variations in correlations among the modules as per Equation 6.20.

A method for generating correlated pseudo-events with high occupancy tails based on simulated events using CMSSW has been developed and validated. The correlation among the modules are preserved with some differences, which are not the concern for underlying interests of stress testing the readout electronics.

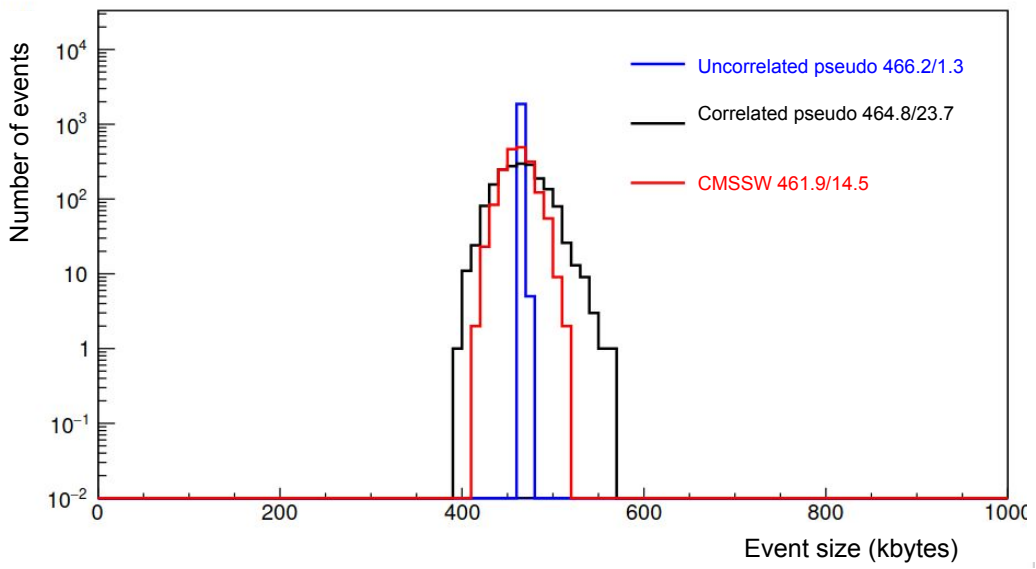


Figure 6.25: The plots display the total event size per 120° after running the emulator with CMSSW events, and both uncorrelated and correlated pseudo-events. While the mean event size is consistent across datasets, the standard deviations differ.

6.4 Effect of data rates and system settings on buffer occupancy and data loss

This section summarizes the detailed validation of HGCAL readout system using the HGCAL readout emulator described in Section 6.1.3 with $t\bar{t}$ events generated in CMSSW & pseudo-events. We emulate the HGCAL readout chain and buffer model, as discussed in Section 6.1.3, for N BXs with a fixed L1 acceptance (L1A) rate or L1 trigger rate. The L1As can be issued randomly or periodically. An event from the input data set is picked for a BX and if the L1A is set to be true, the `HgcrocAsics` class sends the current event information to the `ECON-DAsic` specific classes. The `Lpgbt` specific classes model the collection of data by an lpGBT from all the ECON-Ds connected to it. The emulator implements realistic constraints due to the limited buffer in each ASIC and the latency involved in transmitting data from one part to another. The `Analyser` classes corresponding to different entities save such useful information in the histograms which can be analyzed later.
The following quantities help benchmarking the impact of different system parameters on overall data rates and potential data losses:

- Event size (kilobytes or kB) : the total amount of data collected per event across all Slinks in a 120° sector.
- Maximum ECON-D buffer occupancy (number of 32-bit words) : refers to the maximum state reached by an ECON-D buffer during an emulator run. This distribution gives a quick overview of how many ECON-D buffers are nearly full, indicating possible data losses. The ECON-D drops the incoming packets if the buffer occupancy exceeds 2600 32-bit words.
- Number of dropped packets per event : the total number of data packets lost during transmission from FE to BE.
- Number of ECON-Ds truncating : when an ECON-D buffer occupancy exceeds 2600 words, it drops the next event entirely, sending header words to the BE to indicate the packet loss due to buffer limitations. However, complete data from rest of the ECON-Ds arrive at the BE, meaning only part of the full event packet is lost, not the entire event. This partial data loss is referred to as ECON-D truncation, resulting in incomplete events. Tracking the number and location of ECON-Ds that experience truncation helps quantify the extent of data loss.
- Fraction of truncated events : refers to the fraction of events in which at least one ECON-D experiences truncation and helps to assess how often incomplete events are recorded in cDAQ.

In this section, we summarize the impact of different system parameters and FE data rates on the buffer occupancies and event data rates of the HGCAL. The emulator is run for 100 k or 1 M BX with a random L1A trigger rate of 750 kHz unless otherwise stated. We used $\approx$ 100 k CMSSW simulated events with the latest HGCAL geometry with 47 sampling layers and 1 M pseudo-events. No provisions are made to throttle or stop the L1A trigger to prevent buffer overflow during these runs. The emulator is run in zero suppression mode. Discrepancies in the occupancies due to mismatch of CMSSW geometry and the target geometry are corrected by using the fixed density method discussed in Section 6.2.1.
Comparisons of average packet size per module, total event size and maximum buffer occupancy for ECON-Ds during emulator runs obtained from CMSSW events, uncorrelated pseudo events, and correlated pseudo events are shown in Fig. 6.26 from left to right, respectively. The uncorrelated pseudo-events refer to the occupancies for each module randomly sampled from the fit functions as discussed in Section 6.3.2.

The average packet size per module (in terms of 32-bit words) and the average total event size match well in the three cases. However, the RMS of total event size is higher for correlated pseudo-events compared to CMSSW events, which is mostly attributed to the fact that the former have higher occupancy tails by construction. Almost zero width of total event size is as expected in the uncorrelated pseudo events sample. The higher tails modeled in the pseudo events as compared to CMSSW events are also an explanation for the higher maximum ECON-D buffer occupancy shown in Fig. 6.26 (right). We explicitly verified by comparing the occupancy distributions in CMSSW and pseudo-events for those modules that show significantly different maximum buffer occupancies as shown in Fig. 6.27.

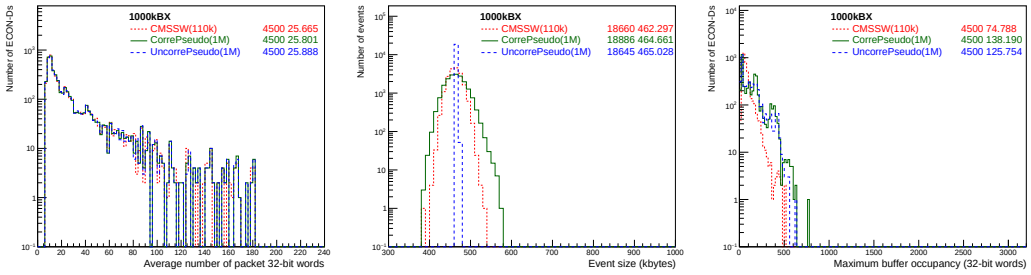


Figure 6.26: Comparison of CMSSW and pseudo events (correlated and uncorrelated); average packet size in 32-bits of words (left), total event size (middle) and maximum buffer occupancy for ECON-Ds (right).

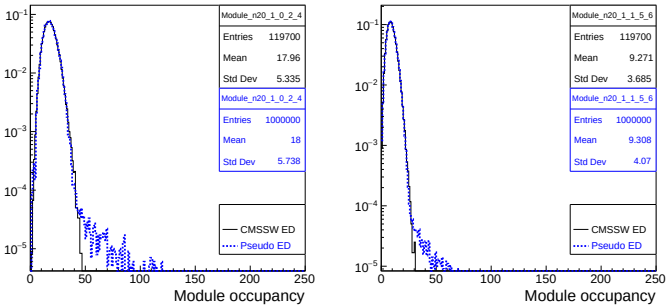


Figure 6.27: Module occupancy in CMSSW and pseudo events. These are the modules where the corresponding ECON-D buffers have higher occupancy with pseudo events compared to CMSSW events.

The above checks confirm that the pseudo-events closely represent the results expected from the CMSSW events or on the conservative side of slightly overestimating the data rates, which is what we need to put conservative limits on the system tolerance. Note that for the remainder of this discussion, the "pseudo events" refer to the *correlated occupancies per module per event* obtained using the prescription discussed in the previous section.

6.4.1 Impact of high L1A trigger rates on data losses

The CMS trigger system for the HL-LHC is designed to operate at an average L1A rate of 750 kHz, but this rate can vary depending on the LHC bunch structure, leading to higher local fluctuations. To assess the limitations of the current system in handling higher trigger rates, we run the emulator by varying the L1A trigger rate from 600 kHz to 1175 kHz while keeping other system parameters constant throughout. "Random" trigger method is used for this exercise. As the trigger rate increases, the total number

of events collected increases, but the average event size should remain unchanged since module occupancies are constant.

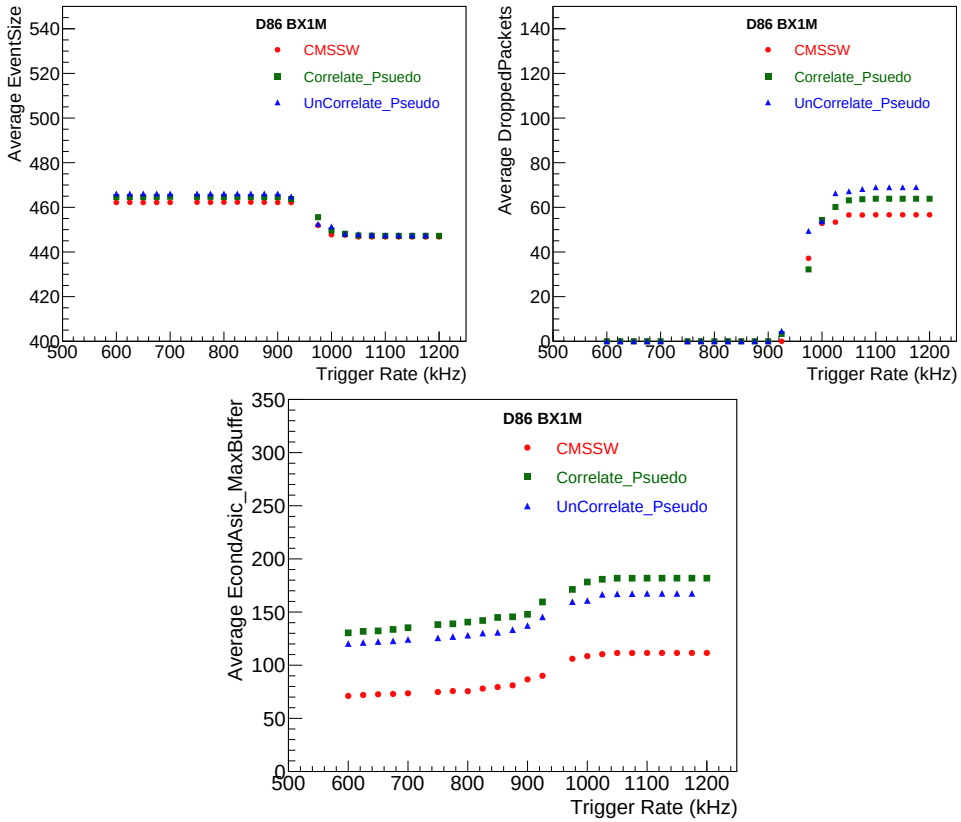


Figure 6.28: The average event size (top left), average maximum buffer occupancy (top right) and average number of dropped packets (bottom) as a function of trigger rate are shown. The plots are also comparing the numbers obtained with CMSSW and pseudo events.

Fig. 6.28 shows the average event size (top left), average dropped packets (top right), and maximum buffer occupancy (bottom) as a function of the trigger rate for both CMSSW and pseudo-events. We observe that the mean event size remains steady up to an L1A trigger rate of 900 kHz but begins to drop beyond that. Twofold reasons : The HGCROC can store only up to 33 events at a time (∼975 kHz), and any additional L1A signals during this period result in dropped events. SECON-Dly, ECON-D buffer limitations begin to affect data, causing packet truncation. As the trigger rate increases, the maximum ECON-D buffer occupancies increase, as shown in Fig. 6.28 (bottom). More event packets are lost at the HGCROC & ECON-D level, reflected in the average packet dropped per event shown in Fig. 6.28 (top right). No data loss is observed up to 900 kHz. Beyond 900 kHz, the number of dropped packets by HGCROCs and ECON-Ds increases and this effect appears in the average dropped packets.

6.4.2 Buffers performance with artificially increased data rates

The CMSSW simulation for HGCAL pseudo-event generation is inaccurate due to changes in geometry, imperfect scintillator modeling, and detector aging approximations, leading to increased noise and occupancy rates. To understand the limitations of the readout design in coping with such situations with higher data rates, we artificially increase the data rates and emulate the buffers. Higher data rates can reveal bottlenecks and inefficiencies in data handling, processing, and transmission through

various stages of the readout chain. Module occupancies obtained from CMSSW or pseudo-events are scaled before creating the corresponding ECON-D data packet and implemented in the `ECON-DAsic` class of emulator. We can scale up one of the *n10*, *n20*, or *n30* occupancy formats at a time or scale up to total occupancies. There is also a provision to scale only certain types of modules. In this section, we summarize the impact of increased data rates on the performance of readout using the following scaling scenarios:

- Scale the data rates by a constant factor for all modules at once (Section 6.4.2).
- Add a variable fraction number of cells more to the occupancies for all modules at once (Section 6.4.2).
- Add 10 % of number of cells to only the modules positioned in a layer at a time (Section 6.4.2).

Scaling the nominal data rates for each module with a constant factor

We run the emulator for 100 k BXs at a L1A trigger rate of 750 kHz with the module occupancies, either one of *n10*, *n20*, *n30* at a time or all together, scaled up in steps of 10 % for each event. The legend within the summarized plots specifies the occupancy formats that are adjusted by a constant scaling factor. Scaling the module occupancies on an event-by-event basis results in corresponding increase in the average occupancies by the same factor. Please note that if a module occupancy obtained from CMSSW is 0 then it stays 0.

The average event size obtained with scaling *n10* (top left), *n20* (top right), *n30* (bottom left) and nhits (bottom right) by different scaling factors is shown in Fig. 6.29.

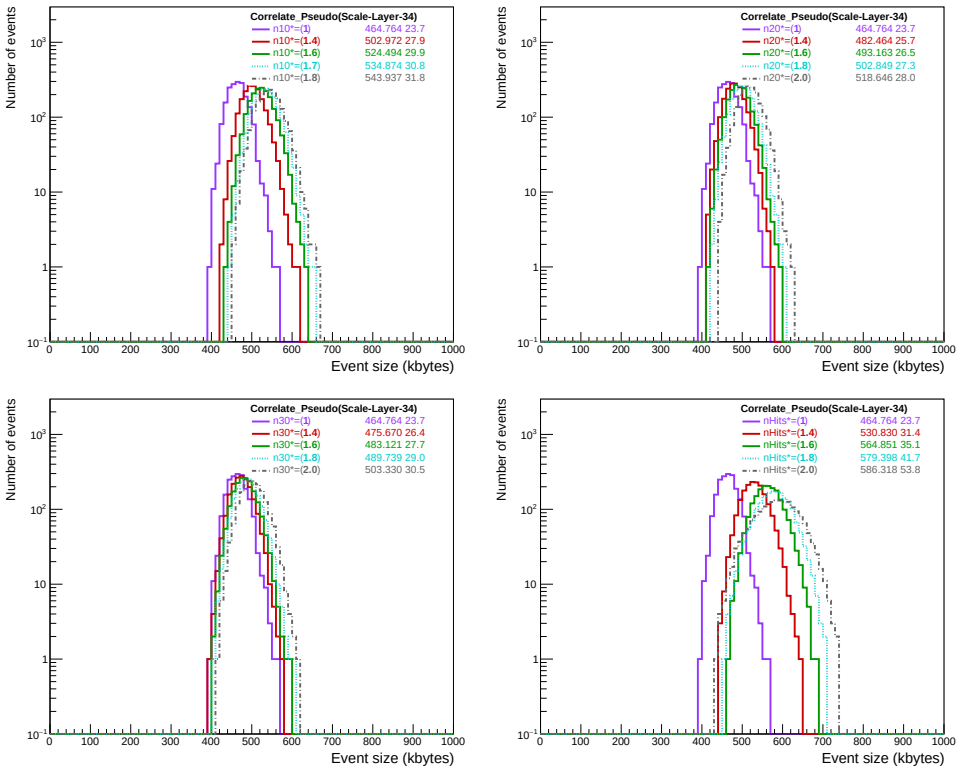


Figure 6.29: The average event size obtained after scaling only *n10* (top left), *n20* (top right), *n30* hits (bottom left), and all three at once (bottom right). The emulator was run using pseudo events and different scaling factor.

Scaling up the *n10* hits leads to a greater increase in event size compared to scaling only the *n20* or *n30* hits as *n10* hits are more probable and modules have high occupancy in terms of *n10* hits, as discussed in section 6.2. The increase in event size does not scale at the same rate as the scale factor.

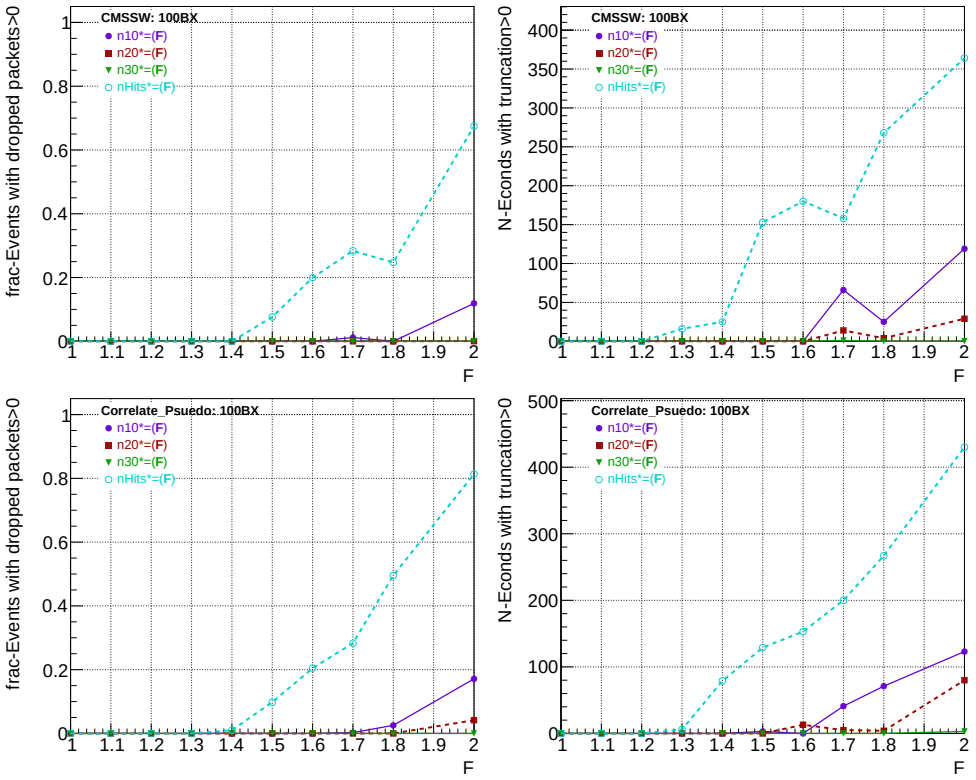


Figure 6.30: The left column shows the percentage of events with dropped packets or incomplete events, and the right column shows the average number of ECON-Ds truncated per event as a function of scale factor for *n10*, *n20*, and *n30* individually. Data is from CMSSW (top row) and pseudo events (bottom row).

The possible data losses with the increase in occupancies at the FE are entirely attributed to the ECON-D buffer limitations. HGCROC buffer limitation does not come into picture as they only are affected by increase in trigger rates.

To quantify the impact of increased data rates, Fig. 6.30 summarize the fraction of incomplete events or events with packets dropped (left column) and number of ECON-Ds with truncation (right column) as a function of scale factor. The upper (lower) row corresponds to the results obtained using CMSSW (correlated pseudo events). As depicted in the left column of Figure 6.30, $\approx 70\,\%$ increase in nhits leads to a 30 % of dropped event packets. However, when exclusively scaling *n10* hits, data loss becomes evident solely beyond a scaling factor of 70 %. The right column illustrates a maximum truncation of $\approx$ 10% for ECON-Ds when all hits are scaled by a factor exceeding 70 %. In contrast, scaling only *n10* hits results in a non-zero truncation observed only after reaching a 70 % increase.

The existing configuration linking the FE to the BE is capable of handling the data rates from the HGCAL-FE without any data loss, even in the event of a 20% increase in occupancy levels.

To understand the origin of data losses when scaling hits in pseudo-events, we identified the ECON-Ds with non-zero truncation with respective *luv*s (left most),

average packet size (middle), and the corresponding number of connected elinks (right most) are shown in Fig 6.31 in case of scaling by 30% (top row), and 80 % (bottom row), respectively. A higher value of *luv* indicates layers located further from the interaction point. Initially, ECON-Ds with high average data rates positioned near the interaction point show truncation. However, as the scaling factor increases, ECON-Ds with lower data rates & positioned further away from the interaction points also start to truncate, as seen in Fig. 6.31. Typically connected to multiple elinks, these ECON-Ds eventually truncate even when linked to just one elink as data rates rise.

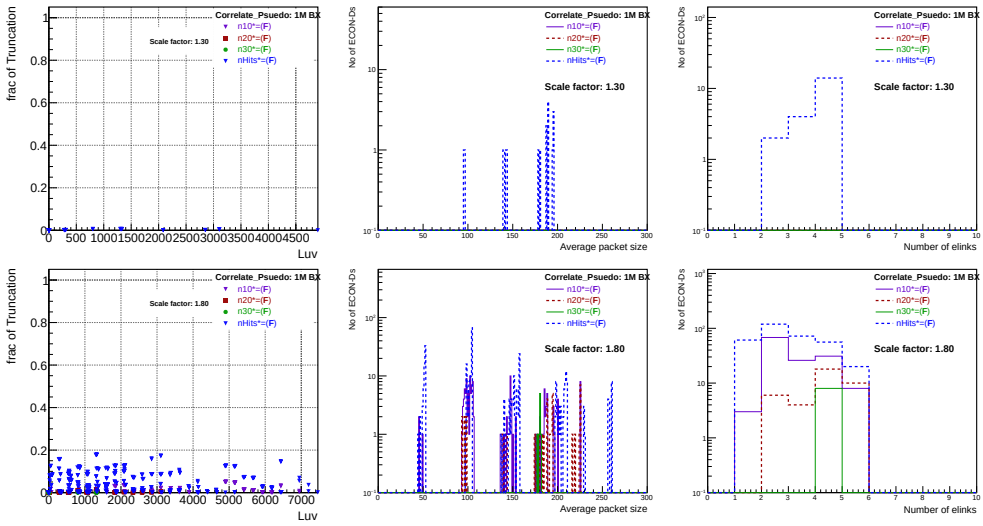


Figure 6.31: The ECON-Ds featuring non-zero truncation, represented by their respective *luv*s (left), mean packet size (center), and associated count of active elinks (right), are illustrated for a scaling factor of 30% (upper row) and 80 % (lower row), respectively.

Non-linearly scaling the occupancies

Any additional signal because of noise fluctuations exceeding the zero-suppression thresholds or malfunction of cells resulting in consistently higher noise adds an extra 32-bit word of data per affected cell. If this happens in multiple cells and modules, the data rates will exceed the nominal rates. To simulate these scenarios in our emulator studies, we add a fraction of nCells (number of cells) to the hit occupancies *n10*, *n20* and *n30* for each module in every event. Instead of scaling the data rates by a constant factor, we use a nonlinear scaling method that depends on the number of cells in the module, independent of its occupancy. Such scaling results in more hits for HDM (432 channels) compared to an LDM (192 channels).

We run the emulator by adding F % of nCells to *n10*, *n20*, and *n30*, for one format at a time where F is varied from 0 to 10 in steps of 1. The emulator is run for 1 M BX with a random L1A trigger rate of 750 kHz, keeping all other parameters unchanged. The legend within the summarized plots specifies the occupancy formats that are adjusted by a constant scaling factor.

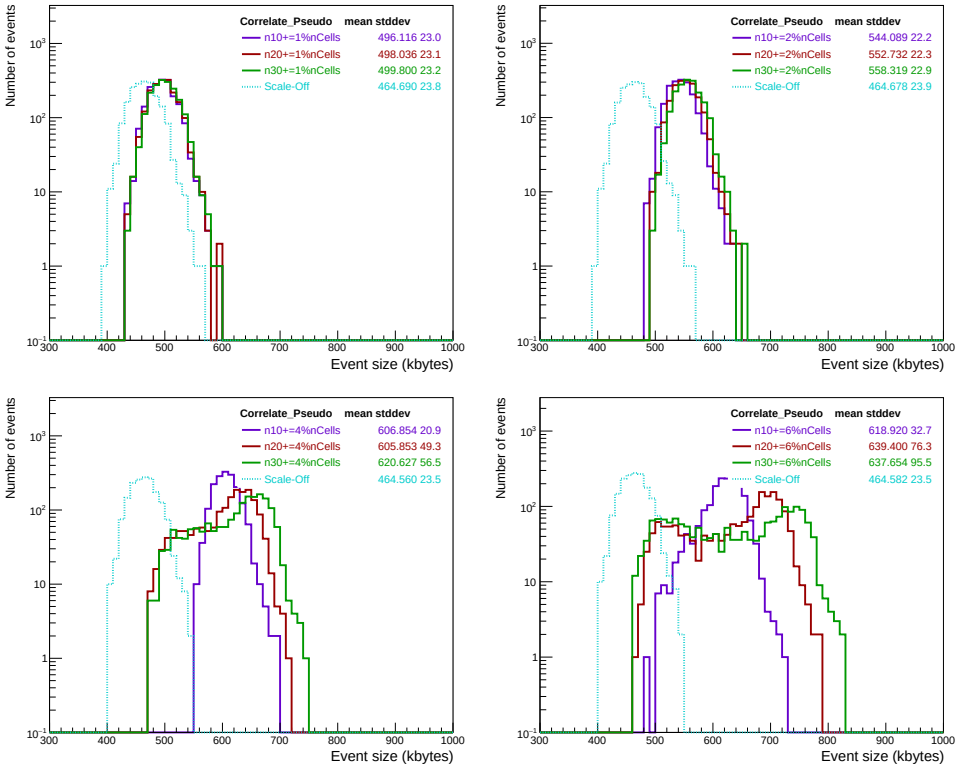


Figure 6.32: Event size distributions for 1%, 2%, 4%, and 6% increments of nCells (from top left to bottom right) added individually to *n10*, *n20*, and *n30* from pseudo events are shown. The dotted Cyan line represents the default distribution without scaling. Mean and standard deviation are in the two rightmost columns of the plot legends.

Adding one *n30* hit, 32 bits are added to the data packet, whereas adding 1 more *n10* and *n20* hits results in additional 12 and 24 bits to the data packet, respectively (see Section 6.1.3 for more details). Hence, the mean distribution increases when F is added to *n30* hits. As discussed earlier, larger data rates result in busier buffers for a large number of ECON-Ds.

The event size distributions obtained by adding F to *n10*, *n20*, and *n30* hits from pseudo events are displayed in Fig. 6.32. Integrating a mere 1% of nCells into *n30* hits across all modules triggers a remarkable $\sim$7 % surge in the average event size (see

top left of Fig. 6.32). In contrast, increasing all hit counts by 70 % resulted in a slight 7 % increase in the average event size (see bottom right of Fig. 6.29).
As we increase F, the event size distribution takes on a strange shape due to the significant number of event packets lost, which is a nonlinear process as shown in Fig. 6.32. At a certain point, most event packets are dropped, causing the average event size collected in cDAQ to remain relatively low as a function of F. However, high fluctuations in noise that induce *n20* or *n30* hits are quite rare, making these scenarios unlikely. In contrast, only *n10* is affected by increased noise and defective channels; and adding F to *n10* does not lead to drastic data losses, indicating that the system is stable against such changes.

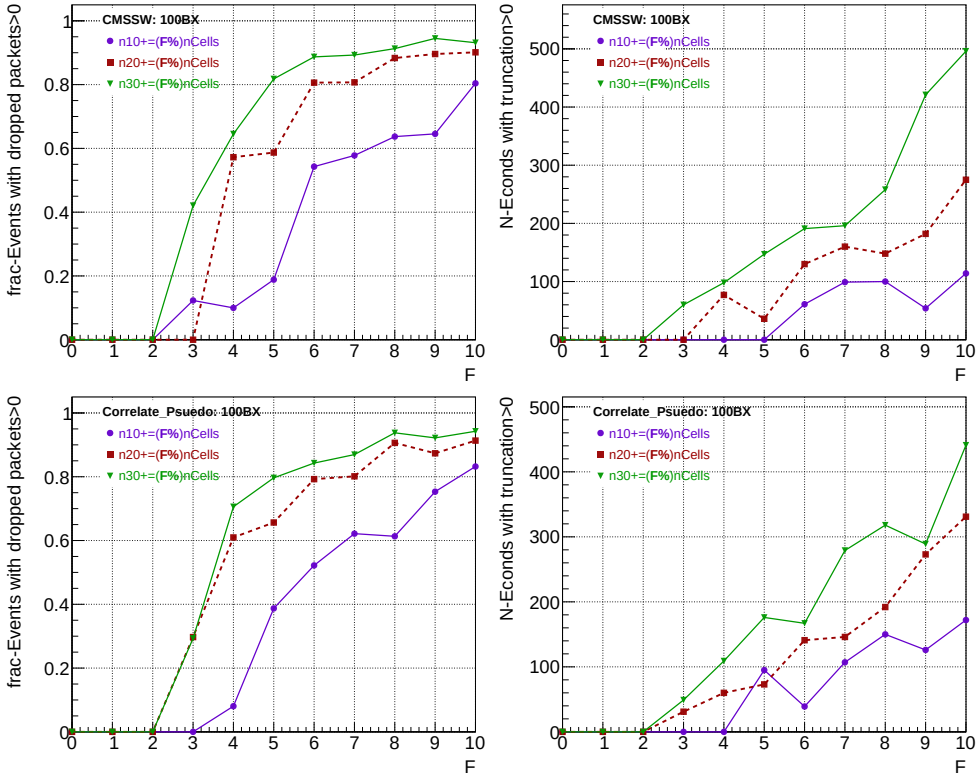


Figure 6.33: The fraction of events with at least one dropped packet or incomplete event (left column) and the average number of ECON-Ds showing truncation per event (right column) are displayed above when F% (X-axis) of nCells have been added individually at a time, to *n10*, *n20* and *n30* obtained from CMSSW (top row) and pseudo events (bottom row).

Fig. 6.33 shows the fraction of events with at least an ECON-D truncating or incomplete event (left column) and the number of ECON-D truncating in an event as a function of different F values, individual addition of nCells to *n10*, *n20*, and *n30*, based on data from CMSSW (top row) and pseudo-events (bottom row). The zigzag pattern arises from the use of random triggers and event selection, along with the limited runtime of the emulator.

For values of F that reach 2 %, we observed an almost negligible occurrence of data loss and truncations. This underscores the exceptional robustness of our current readout design in the face of fluctuating conditions. Beyond this, the fraction of incomplete events increases and becomes close to 1 beyond 6 %, indicating that

all events received in cDAQ are incomplete. The number of ECON-D that show a truncation also increases with F. Besides 8 %, almost 30 % of ECON-Ds show truncation. The observations do not change whether we use CMSSW or pseudo-events.

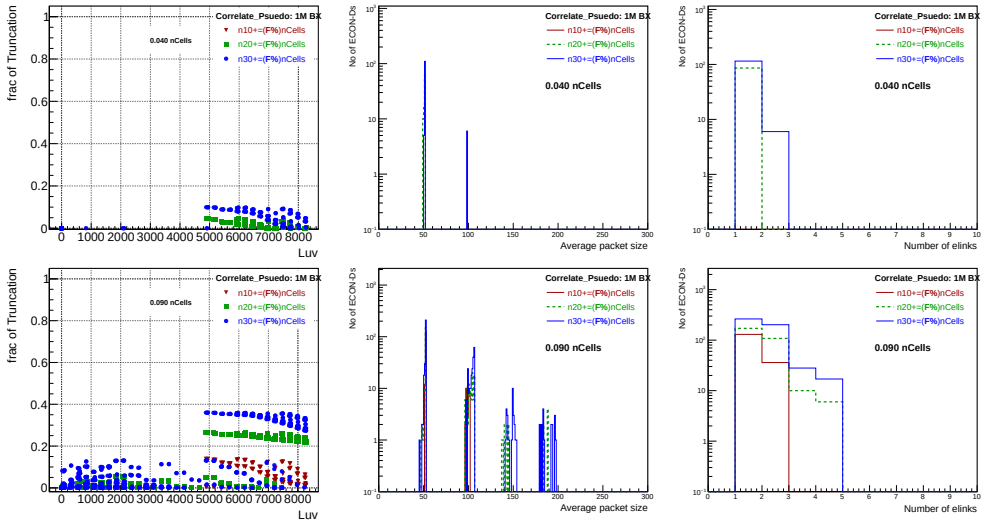


Figure 6.34: The ECON-Ds featuring non-zero truncation, represented by their respective *luv*s (left), mean packet size (center), and associated count of active elinks (right), are illustrated for a fraction F of 4% (top row) and 9 % (bottom row) of nCells, respectively.

To identify ECON-Ds leading to data losses when adding more to their occupancies, we identified ECON-Ds with nonzero truncation with respective *luv*s (left most), average packet size (middle) and the corresponding number of connected elinks (right most) are shown in Fig. 6.34 for adding 4% (top row) and 9% (bottom row) of nCells, respectively. Unlike the scenario where hits are scaled by a constant factor, ECON-Ds with small average data rates, connected to 1 elinks and positioned away from the interaction point show truncation. However, as F increases, the ECON-Ds with higher data rates, connected with more elinks & positioned near the interaction points also begin to truncate.

Stress testing for one layer at a time

Simultaneously scaling up hit rates for all modules subjects the system to scenarios far more extreme than realistic conditions. In contrast, increasing hit rates for a specific layer at a time reflects potential real-world issues such as noise, malfunctioning modules, or other failures. This situation becomes particularly problematic when modules along the same radial direction see a sudden increase in data rates, resulting in truncated data packets due to sharing their data over same lpGBTs connected, potentially leading to false MET-like signatures. To mitigate this risk, we perform a layer-by-layer analysis in this section, incrementally increasing data rates to observe how data losses due to truncation evolve.

The emulator is run for 1 M BX and at a random L1A trigger rate of 750 kHz with adding F% of Cells to the occupancies of modules of a layer at the same time (instead of all layers simultaneously). We check the impact of varying F in steps of 1% in the range 1% to 10% on variables characterised by data losses. Since data losses were observed only after adding 10% of nCells, we summarise the results corresponding to F=10%.

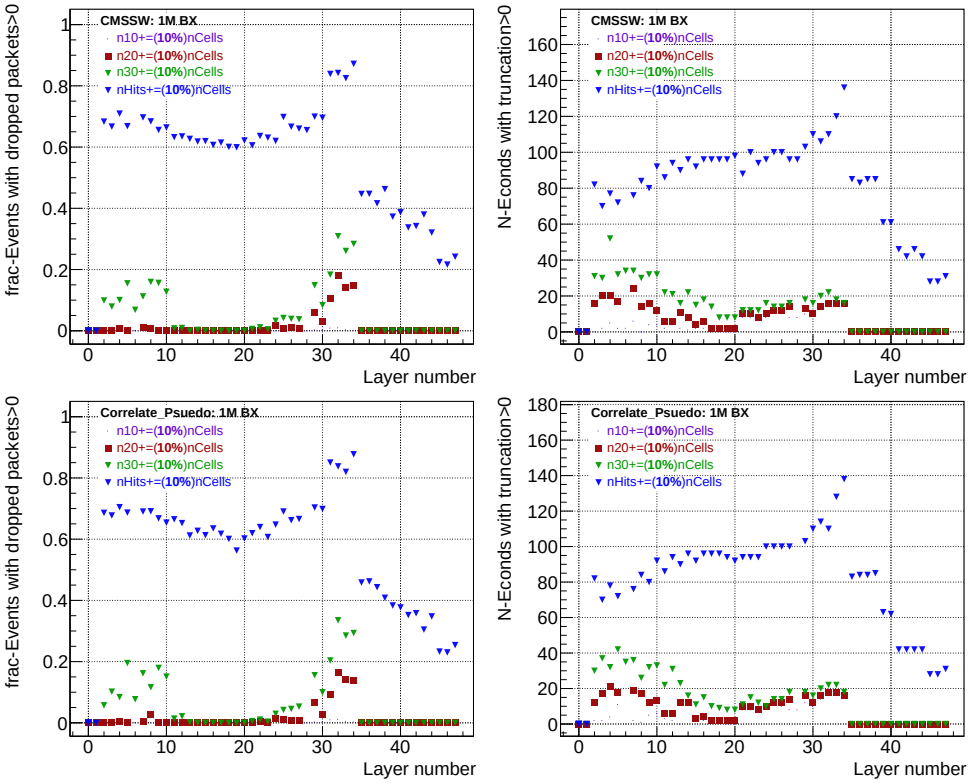


Figure 6.35: The Y-axis shows the fraction of events with dropped packets (left) and average truncated ECON-Ds per event (right). The X-axis represents layer numbers with 10% of nCells added to *n10*, *n20*, *n30*, and nhits for these layers, using CMSSW (top) and pseudo-events (bottom).

Figure 6.35 illustrates the fraction of incomplete events (left column) and the average number of ECON-Ds per event that exhibits truncation (right column), respectively, as a function of the layer number, where 10% nCells (comprising *n10*, *n20* and *n30*) are added while keeping the remaining unchanged. When we add 10% of nCells to all formats for the first 34 layers of HGCAL, ∼60% incomplete events are recorded in cDAQ, after which we observe a sudden drop in the fraction of incomplete events. The number of ECON-Ds that truncate packets also increases (up to 5%) initially and suddenly drops beyond the 34^{th} layer. Higher-level truncation in the first 9 layers can be understood as the maxima of electromagnetic showers that will likely occur in these. Observations from the CMSSW events (top row) and pseudo-events (lower row) are consistent.

To identify ECON-Ds leading to data losses when adding more to their occupancies, we identified ECON-Ds with nonzero truncation with respective *luv*s (left most), average packet size (middle) and the corresponding number of connected elinks (right most) are shown in Fig. 6.36 for layer 5 (top row) and layer 7 (bottom row) and for adding 4% (top row) and 9% (bottom row) of nCells, respectively and in Fig. 6.37 for layer 26 (top row) and 34 (bottom row). In contrast to the case where hit rates are uniformly scaled or nonlinearly scaled, the average packet size for these ECON-D modules varies from 50 to 200 or so, with typically only 1 or 2 connected elinks. Modules located farther from the interaction point are the first to experience truncation as these are mostly connected with one elink at most and the degree of truncation increasing as F rises.

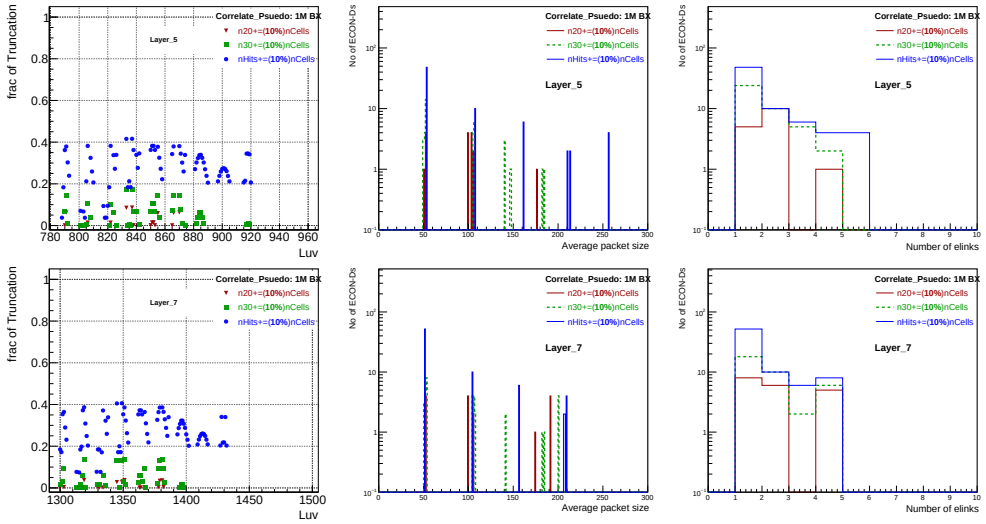


Figure 6.36: The ECON-Ds featuring non-zero truncation, represented by their respective *luv*s (left), mean packet size (center), and associated count of active elinks (right), are illustrated when 10% of nCells is added to layer 5 (top row) and 7 (bottom row).

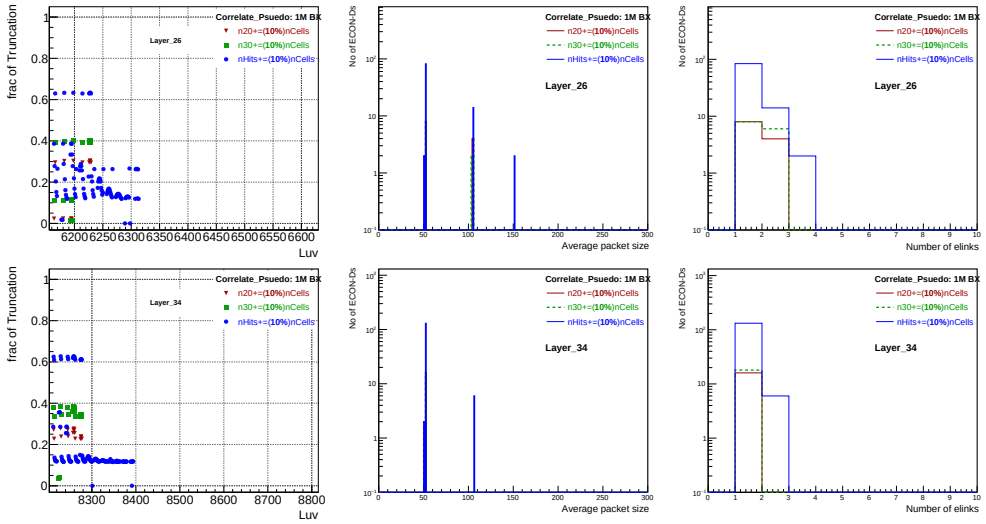


Figure 6.37: The ECON-Ds featuring non-zero truncation, represented by their respective *luv*s (left), mean packet size (center), and associated count of active elinks (right), are illustrated when 10% of nCells is added to layer 26 (top row) and 34 (bottom row).

The system can manage a 20 to 30% occupancy increase with minimal data loss. Layer-wise validation of ECON-D to lpGBT mapping shows the system's optimization and reduced data loss risk.

6.5 Throttling to avoid truncation

The BE DAQ has a provision to stop the incoming L1As if any of the relevant buffers is going to be full or full and may lead to data losses for that event. This process is referred to as *throttling of L1As*. Data losses, presented in Section 6.4, as function of module occupancies, data rates and trigger rates refer to the scenario when BE does not activate the L1A throttling.

The buffers relevant to data truncation and throttling are the ECON-D main buffer, the HGCROC RAM-2 buffer, and the DAQ and Slink buffers. As discussed in section 6.1.1, the RAM2 of HGCROC can buffer at most 33 L1A's at a time. If the L1As rate exceeds "950 kHz", more L1As arrive before the first L1A has not been transmitted to the ECON-D and HGCROC overwrite the first L1A with the next one. This leads to losing that event information completely and causing HGCROC to be out of sync with rest of the system. This should be avoided at all cost. Hence at the BE, we emulate the HGCROC buffer occupancy and throttle the L1As before it overflows (refer to thesis [137] for more technical details and methods). The size of Slink and DAQ buffers are not defined as of now and depends on the available resources in FPGA. So the impact of these buffer limitations is not studied yet.

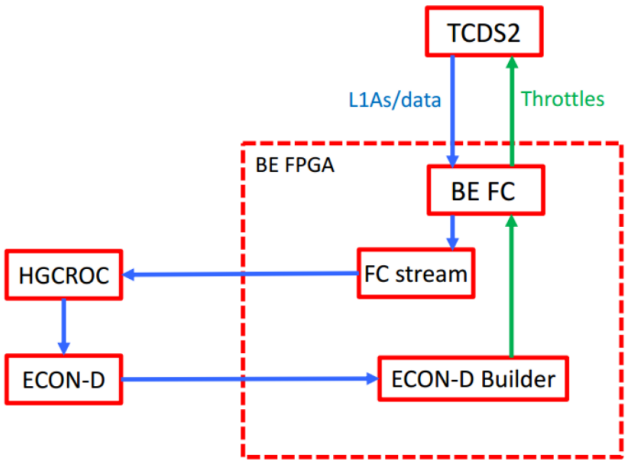


Figure 6.38: Above schematic shows how L1As and throttling signal travels through the readout system.

As discussed in section 6.1.1, when the ECON-D buffer occupancy exceeds 2600 32-bit words, it drops the next event completely and sends a truncated data format containing event header information (Fig. 6.7 truncated data format) to the BE. This leads to incomplete events with missing information from the corresponding modules. The ECON-D also regularly sends the buffer status bits to BE along with each L1A packet, thus, the BE can keep a tap on the buffer occupancy. When buffer occupancy exceeds a predefined threshold for BE, called *throttling thresholds*, BE can send a throttling signal to the TCDS2 system to stop incoming L1A. This results in a loss of events for the CMS HGCAL and affects the functionality of CMS overall, leading to a larger overall detector dead time. Hence, we need to make a choice at the BE level between collecting incomplete events and lowering the rate of complete events increasing the dead-time of the system.

We need to assess the severity of ECON-D truncation and determine the level of inaccuracies introduced in the offline reconstruction. For instance, if multiple ECON-Ds in the same angular direction experience truncation in an event, this could result

in false MET signatures. Hence, it is essential to evaluate how frequently and how many ECON-Ds are affected by truncation as we vary the trigger rates or occupancy rates.

Reflecting on the characteristics of the throttle process, there is a latency between the BE issuing a throttle command and when it actually takes effect. Fig. 6.38 illustrates the signal flow from the BE to the HGCROC, with each block in the chain contributing an additional delay. The total latency is currently estimated to be around 135 BXs, or approximately 3.4 μs, though this value may change with future adjustments of the system configuration and parameters. During this delay, ECON-D may initiate data truncation before the throttling takes effect if the throttling threshold is set to be a very close truncation threshold. If we set the throttling threshold very far from truncation, then it will result in more dead-time for the system. As a result, throttling and truncation are interconnected but act in opposing directions $--$ minimizing dead time increases the likelihood of truncation. Therefore, this becomes an optimization problem, where both the ECON-D truncation threshold and the BE throttle signal threshold must be carefully tuned for optimal performance.

At the BE, the throttling threshold can either be as *a single threshold* or a set of two thresholds referred as *a double threshold*. The schematic in Fig. 6.39 shows the impact of single and double thresholds on total time when throttling is "ON". Using a single threshold to turn "ON" the throttling can result into more dead-time or throttling time as it is sensitive to even small fluctuations in the occupancies. Hence, a double threshold system is chosen to turn "ON" and "OFF" throttling. The higher and lower thresholds are called *high and low water marks*, respectively. The throttle is turned *ON* when the ECON-D buffer occupancy crosses the high water mark and is not turned *OFF* until the buffer occupancy drops below the low water mark. These thresholds are set in units of 32-bit words.

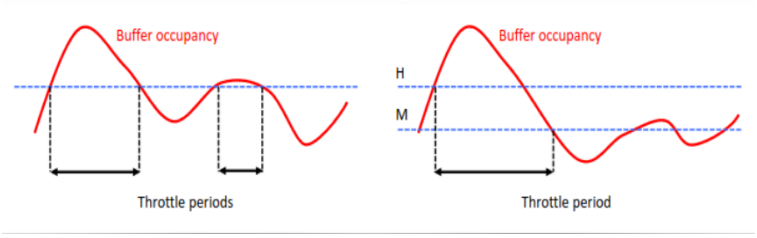


Figure 6.39: Above schematic shows the impact on throttling time if we have single throttling threshold or double throttling thresholds.

Following are some of the common terminologies used in the later test,

- **Total dead time**: total number of BXs during which throttling signal was active due to overflow buffers of HGCROC, ECON-D, or BE DAQ.
- **Fraction of time system is busy**: fraction of total run time during which the throttle signal was active due to overflow buffers from HGCROC, ECON-D, and BE DAQ.
- **Fraction of events received**: the ratio of L1A sent by HGCROC to the total L1A issued by TCDS2.
- **Fraction of incomplete events** or **Fraction of events with at least 1 ECON-D truncating**: events with at least 1 ECON-D truncation with respect to total events received in cDAQ.

The values of high and low water marks need to be scanned to compute the possible dead-times as a function of different trigger rates and data rates. Statistical fluctuations across different emulator runs due to randomised L1As and events limit the accuracy of such calculations. To minimize statistical fluctuations in the accuracy of the calculated quantities listed above, we run the emulator 10 times for 500k BX for each point and compute the average of quantities while keeping a track of the standard error. The error in computed quantities is the average standard error on each point divided by the square root of the number of emulator runs. We are using pseudo-events with 1 M statistics to run the emulator.

6.5.1 Dead time introduced by linearly increasing data rates

As summarized in Section 6.4.2, data losses quantified as the fraction of incomplete events, without using the throttle signal, are maximum when all nhit formats (*n10*, n20, n30 data formats) are scaled at the same time. In this section, we scale the nhits (hence, average occupancy) by a constant factor in the range $1 - 2$ or 0 to 200 %. For each scale factor value, the high (low) water marks are varied in the range of $1850 - 2550$ ($1450 - 2150$) 32-bit words and are represented as a fraction of truncation threshold (2600 32-bit words). For example, if the high water mark is set at 1850 then it is represented as *i.e.* 70% of truncation threshold, and if set to 2550 *i.e.* 98% of the truncation threshold and similarly for low watermark. Additionally, the difference between low and high thresholds is maintained at 400 in all studies. The emulator was run 10 times for 500 k BX at the L1A rate of 750 kHz with the periodic trigger method for each threshold setting.

For these scaling of nhits, Fig. 6.40 (left) shows the fraction of time system is busy due to ECON-D buffer limitations compared versus the fraction of events that have at least one ECON-D truncating for different values of high water mark presented as a fraction of truncation threshold (set at 2600 words). Fig. 6.40 (right) is a zoomed version of Fig. 6.40 (left), zoomed closer to more realistic scaling, *e.g.* 10% or 20%. The system is observed to be busier when the throttling threshold is significantly lower than the truncation threshold. Conversely, a higher fraction of events are truncated when the throttling and truncation thresholds are closely aligned, due to the latency between triggering and activating the throttling mechanism. The ECON-D throttling can cause the system to be busy for a maximum of 20% of the total runtime while the fraction of incomplete events ranges from 2% to 8% when the hits are scaled by factor 1.6 . Increasing the data rates by a factor of 1.6 and turning *ON* the throttling , the number of incomplete events is much less than 10% with a minimum dead time of 10% as compared to 20% (as without throttling (see Fig. 6.30.

Fig. 6.41 shows the distributions of number of ECON-Ds truncating in an event with different scaled data amounts. At max, only 15 out of 4500 ECON-Ds are truncating at a time, making only 0.3 % fraction of information lost from the HGCAL, indicating data losses are not significant. In contrast, when throttling was *OFF*, almost 100 ECON-Ds out of 4500 truncate, making approximately 2 % fraction of information lost from the HGCAL, at the same time in an event as shown in Section 6.4.2. These are mostly the ECON-Ds positioned at the inner radii region, foreseeing high average occupancies and connected with more than 2 elinks as discussed in Fig. 6.31 of Section 6.4.2. This concludes that HGCAL readout system is reliable even with the drastic change in data rates of up to 30 % and BE can avoid throttling L1A signals without loosing much of the information from the events as data losses are not significant.

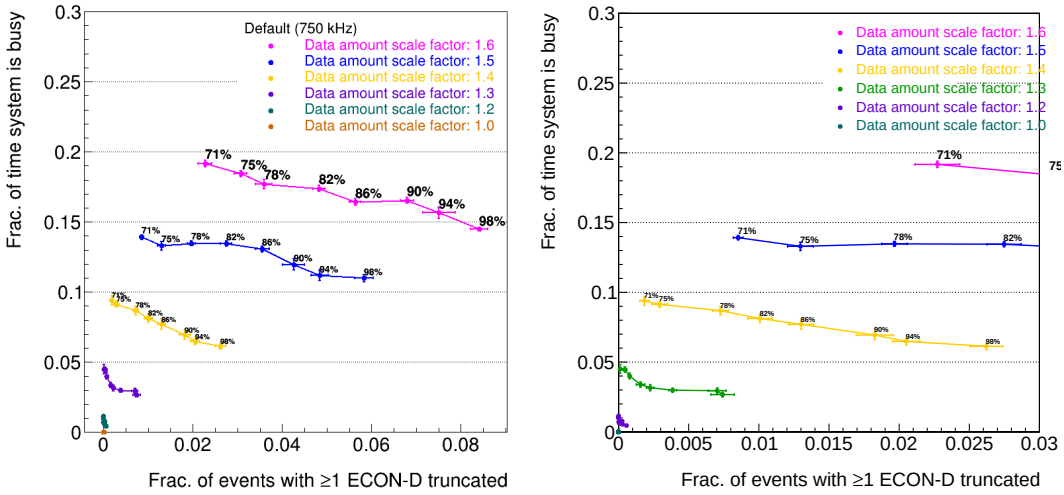


Figure 6.40: The left plot compares system deadtime from ECON-D buffer constraints with event truncation rates across different truncation threshold fractions. The right plot zooms in, emphasizing lower scaling factors. Each point represents the high water marks set at certain fraction of truncation threshold (2600 32-bit words). For each scale factor value, the high (low) water marks are varied in the range of $1850 - 2550$ ($1450 - 2150$) 32-bit words maintaining the difference between two at 400.

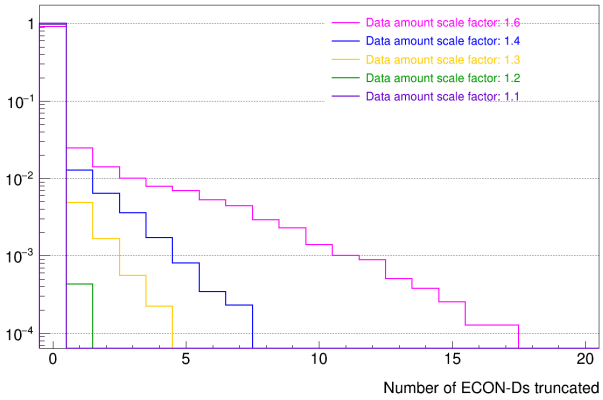


Figure 6.41: Distributions of ECON-Ds truncating per event for varying scaled data across all occupancy formats (nhits) and modules. High and low watermarks are 2550 and 2150.

6.5.2 Dead time introduced with increasing data rates non-linearly

The *n10*, *n20*, and *n30* for each module are scaled based on the number of cells in each module (and not by a constant factor), regardless of occupancy, as discussed in Section 6.4.2. The emulator run details are same as discussed in section 6.5.1. The regular trigger method at 750 kHz is used.

The fraction of time the system is busy due to ECON-D buffer limitations versus the fraction of events that have at least one ECON-D truncation for various scaling of *n10*, *n20*, *n30* and nhits are presented for different watermarks (as a fraction of truncation threshold) and are shown in Fig. 6.42 (left), 6.43 (left), 6.44 (left) and 6.45 (left), respectively. The right plots on the same figures provide a zoomed-in view of the same, highlighting more realistic scenarios shown only in the case of scaling *n10*, *n20* and *n30*. The pattern of busy time versus the fraction of events with at least one ECON-D truncation (incomplete events) remains consistent with both increasing as data rates increase. The system experiences more busy time or incomplete events

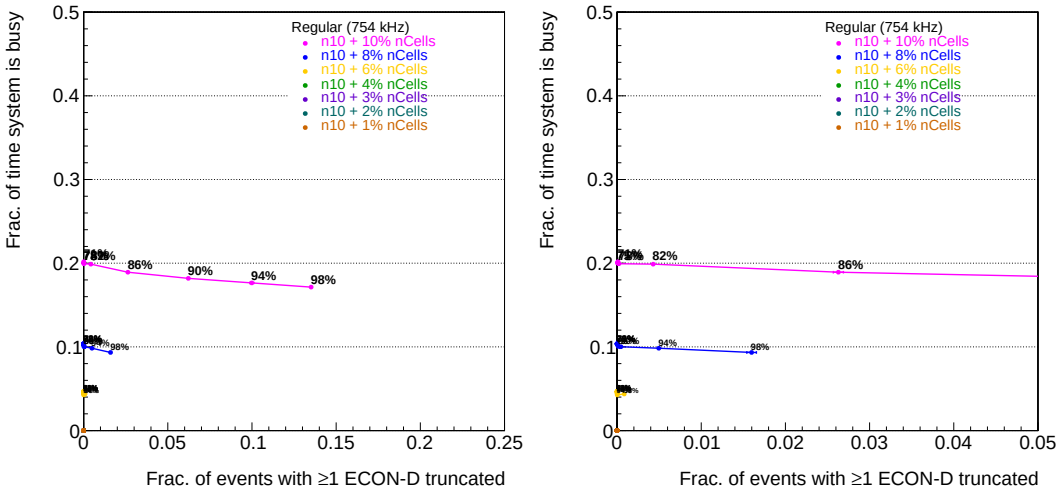


Figure 6.42: The left plot shows the busy time fraction due to ECON-D buffer limits for events with at least one ECON-D truncation, across various watermark levels and added nCells to *n10* occupancy. The right plot is a zoomed version of right plot.

when changes are made to the *n30* hit occupancy as expected and discussed in Section 6.4.2. When the throttle is *ON* and we increase the amount of data by adding 6% nCells to *n30*, the maximum busy time (corresponding to high water mark of *throttling threshold* equal to 1850) is around 20% and with a maximum fraction of truncated events (corresponding to high water of *throttling threshold* equal to 2550) is less than 20%. In contrast when throttle is *OFF*, the fraction of incomplete events increase up to 80% as shown in Fig. 6.33. Figures 6.46 shows the distributions of the number of ECON-D with truncation adding different fractions of nCells to nhits when the high water mark is set at 1850 *i.e.* 70% of truncation threshold (right), and 2550 *i.e.* 98% of the truncation threshold (left), respectively. When throttling was *OFF*, almost 400 ECON-Ds out of 4500 truncate, making approximately 10 % fraction of information lost from the HGCAL, at the same time in an event as shown in Section 6.4.2. In contrast, at max with adding 6 % of nCells to nhits, only 100 (50) out of 4500 ECON-Ds are truncating at a time, making only 2(1) % fraction of information lost from the HGCAL, indicating data losses are not that significant even in such extreme increase in data rates. These are mostly the ECON-Ds positioned at the outer radii region, foreseeing small average occupancies and connected with 1 elinks as discussed in Fig. 6.34 of Section 6.4.2.

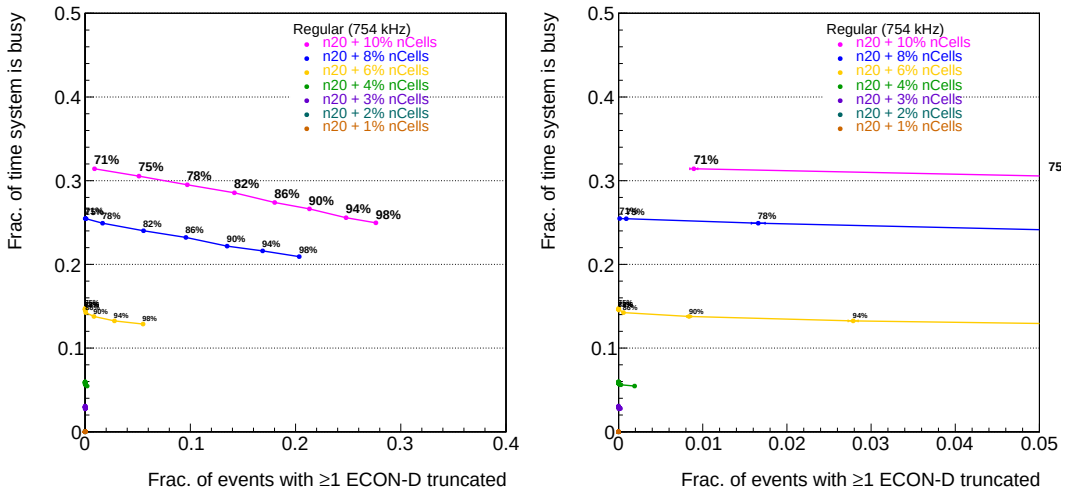


Figure 6.43: Time fraction when the system is busy due to ECON-D buffer limits and the fraction of events with ECON-D truncation across different watermarks and nCells fractions added to *n30* for all modules (left). Right focuses on smaller factors with a zoomed-in view.

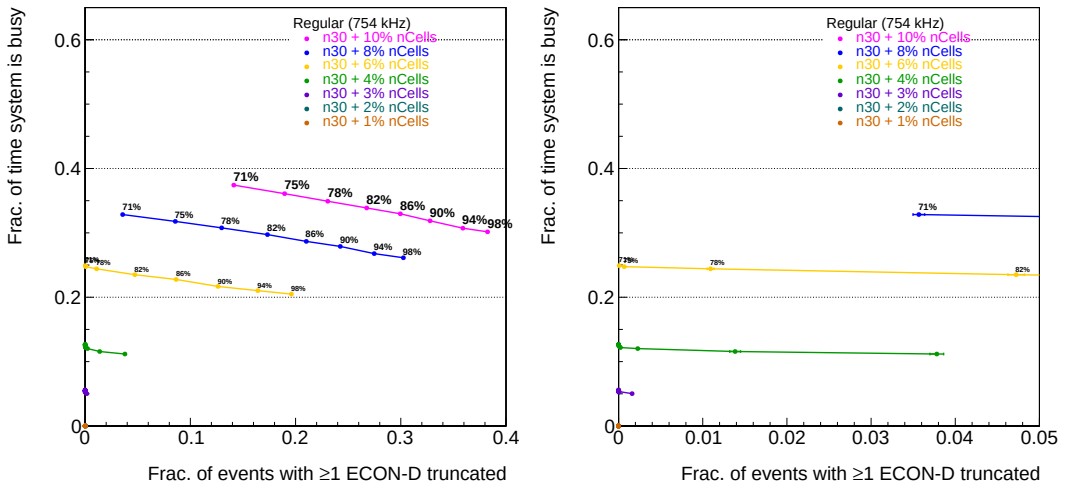


Figure 6.44: The left plot illustrates the time fraction when the system is busy due to ECON-D buffer limits and the fraction of events with ECON-D truncation across different watermarks and nCells fractions added to *n30* for all modules. The right plot focuses on smaller factors with a zoomed-in view.

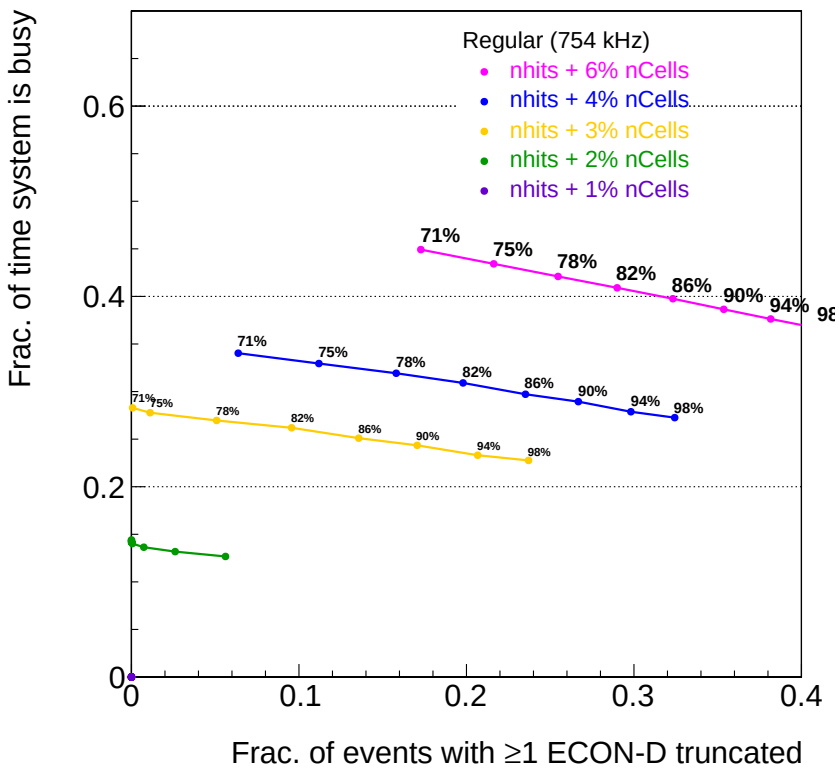


Figure 6.45: The left plot compares the system's busy time due to ECON-D buffer limits against events with at least one truncation, using different watermarks as fractions of the truncation threshold and varying nCells added across all occupancy formats.

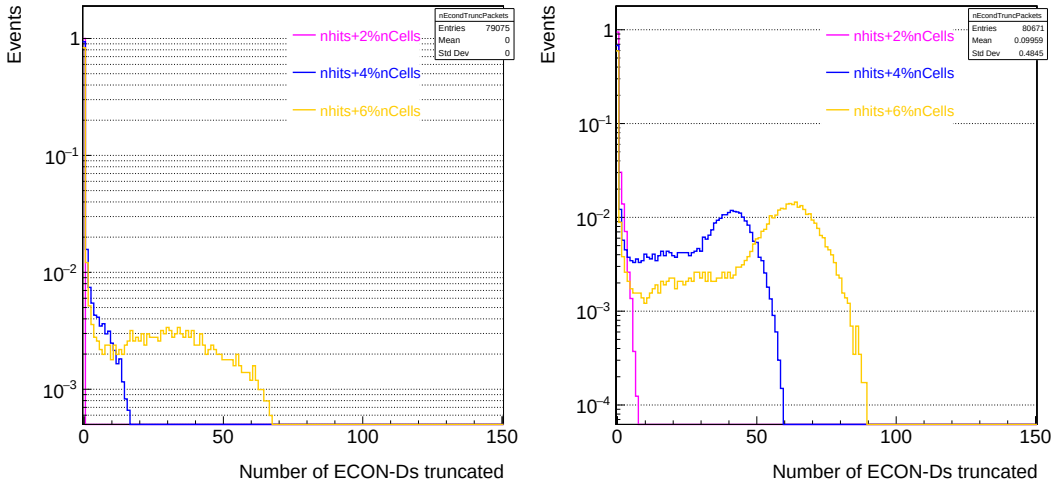


Figure 6.46: Distributions of ECON-D numbers with event truncation for different fractions of nCells added to nhits across all modules for High and low watermarks are 1850 (2550) and 1450 (2150) shown in left and right, respectively.

6.6 Testing of integrated HGCAL readout system with beam test experiment at CERN

To bring this advanced detector from concept to reality, significant progress is being made on multiple fronts. The ongoing activities include the mass production of both absorber and active components, as well as the finalization of the readout chain, which encompasses readout ASICs and BE electronics. Moreover, novel algorithms, such as graph neural networks (GNNs), are being developed to optimize particle reconstruction by harnessing the high granularity of the detector.

Each ASIC design is nearly complete, having been extensively tested, including thousands of register validations. Most design flaws are corrected, but pre-production testing as part of the integrated readout system in labs and real particle settings (e.g., beam tests) is crucial. In recent years, numerous beam tests have been conducted to provide a thorough characterization and validation of the detector and readout components, several of which are referenced in this document.

Given the complexity of the HGCAL readout system, as discussed in Section 3.4, the system can be tested in two ways. Vertically, involving data transfers and control across the FE and BE, or horizontally, by scaling the system in parallel, thereby increasing the number of lpGBT fibers, optical links, and bandwidth required between FE and BE. At CERN, the HGCAL collaboration has successfully tested a vertical chain, comprising two single FE module Trains, an optical FE-to-BE interface, and a Serenity BE board. These efforts are paving the way for final validation before large-scale production and integration into the CMS detector.

The beam test, conducted between August and September 2023, aimed to test and

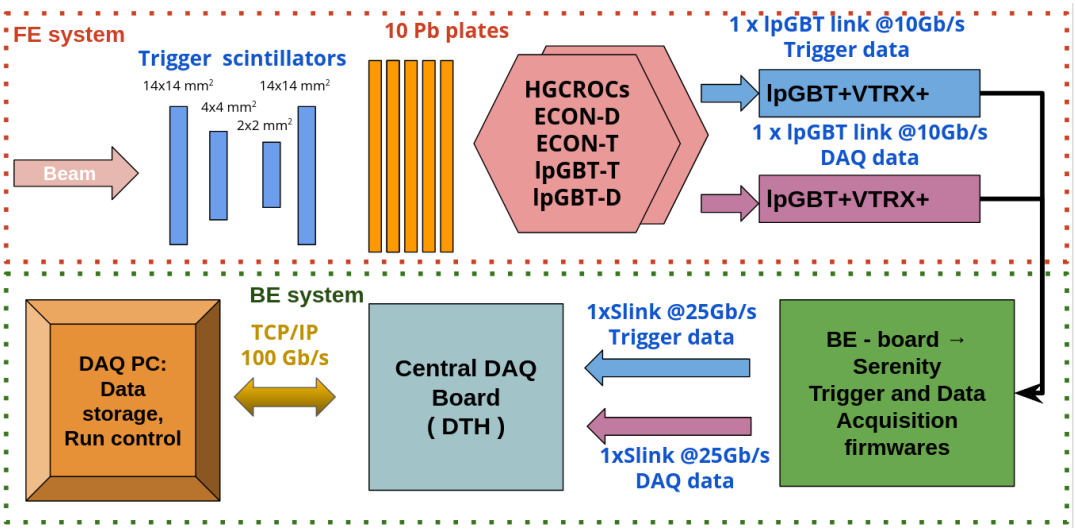


Figure 6.47: Schematic of setup used for September 2023 beam test with HGCAL readout chain.

characterize several silicon modules of varying densities (HD and LD) and shapes (full and partial), while also integrating different components of the readout system in their near-final design. During the August phase of the test, the ECON-D and ECON-T ASICs were not yet available, so ZCU (a Xilinx development board) based emulators were used to simulate their responses, which are not discussed in the present studies and details can be checked in [137]. I contributed to the setting up the system, lab tests, data taking and data analysis of the 2023 September beam test. During this beam test, the ECON ASICs were available. Also, the TPG stage-2 system was not yet available during the test, and the DAQ firmware used was an early,

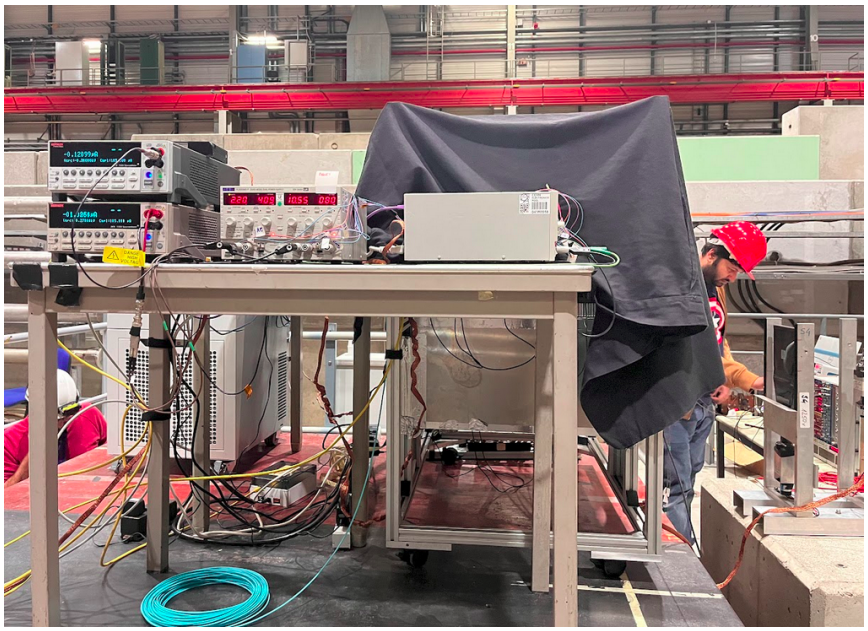


Figure 6.48: Picture of the HGCAL beam test setup at H4 beamline facility of SPS.

small-scale version requiring further development for full-scale deployment. This section focuses on the setup used in the September test and the detailed investigation of the raw data analysis.

6.6.1 Description of beam test setup

Fig. 6.47 shows the schematic illustration of the beam test setup, highlighting the data exchange from FE to BE system. Fig. 6.48 shows a picture of the setup taken during beam test. The beam test was conducted in the H4 beam test facility area of the SPS North area beam. The primary proton beam of 450 GeV energy extracted from SPS is bombarded on the Beryllium target and secondary particles of different types with different energies are produced as discussed in Section 3.4.3. The particle rates in the secondary beams generated in such a way are of the order of 100 kHz. The different energy particles obtained are incident on the experimental setup.

FE Trains

Two single LD module trains V3 are placed back to back in the FE, called the east and west trains. Each train is equipped with a silicon sensor having 192 individual sensor cells, which is mounted on the LD Hexaboards PCB along with the copper base plate and the kapton sheet as explained in Section 3.4. Three V3 HGCROC ASICs and one Rafael ASIC per FE train are hosted on the Hexaboards PCB. Each of the Hexaboardss is connected to two ECON ASICs (ECON-T (ECON-T-P1) and ECON-D (ECON-D-P1) ASICs) via concentrator mezzanine. The V3 LD Engine board designed for LD modules hosts the (DAQ, TPG) lpGBT ASICs and VTRX+, and is linked to Hexaboards via LD single-module Wagon board. DC-DC mezzanines and LDO modules are also placed on top of the Hexaboards to provide stable voltage supply for the Engine board and HGCROCs, respectively. The FE is attached to a copper base plate that features water cooling provided by a chiller to prevent overheating. The entire LD silicon trains are enclosed in a dark box to minimize noise from surrounding light, and a dry air source is utilized within the box to keep

humidity levels constant.
The physics aim of the beam test does not involve measurement of the reconstruction of particle showers but rather is to characterize the sensor cells including signal and noise measurements, setting up the calibration across ADC to TOT measurements, and common-mode noise studies. To maximize the characterization of cells for the listed measurements, a series of 10 absorber plates, each made up of 0.3 mm Fe and 4.9 mm Pb, are placed in front of the FE trains, creating a total depth of approximately 5 X_0. This configuration causes electrons or positrons to start showering before entering the HGCAL beam test setup. Thus, the corresponding signal is detected in multiple neighboring cells at the same time. In contrast, for pions, this depth is not enough to start showering before entering in the HGCAL beam test setup. Hence, most of the time, pions act like MIPs.

BE setup

At the BE, Serenity board with V1.1 is used where both DAQ data packet processing and buffering blocks, TPG unpackers and TC processors, FE configuration (Slow control) and Fast controls are hosted. The 4 lpGBTs (10 Gbit/s) of the engine boards (2 for each module) from the FE train are connected to the Serenity board through a 80 m long fiber. The firmware version used in the beam test as the DAQ is called miniDAQ. The beam test firmware integrates key components from the miniDAQ, which handles ECON-D packet validation and buffering, and Stage-1 trigger systems such as data unpackers and TC processors. The TCDS2 emulator provides both internal and external triggers, along with fast commands and calibration sequences. The DAQ and TPG paths are supported by a 25 Gb/s Slink interface, with Slink and lpGBT link implementations included in the Serenity board firmware [139].
To understand more about the data readout paths, two separate readout data paths were tested during beam test, both using the Slink, one with two 25 Gbit/s optical links through the DTH p1-v2 board via a 100 Gbit/s TCP/IP link, while the sECONDary path used two 10 Gbit/s Ethernet links directly to the DAQ PC (acting central DAQ). The raw data from Serenity are transmitted to the DTH board via Slink (25 Gbit/s) and stored on disk for offline analysis by the DAQ PC, which also manages the run control and configuration. All BE-specific boards and the DAQ PC were located in the control area of the H4 beamline.

Trigger details

At the Serenity board, we use a firmware block to set up the triggers (using external or self triggers) and identifying particles crossing the front-end for L1A generation in the BE . A set of 4 scintillators with different sizes (S1 and S4 are of $14 \times 14\,\mathrm{mm}^2$ S2 of $4 \times 4\,\mathrm{mm}^2$, and S3 of $2 \times 2\,\mathrm{mm}^2$ sizes) were placed in front of the HGCAL beam test setup, out of which only 3 are powered up and we are taking coincidence to create external trigger from only two with the help of a standard Nuclear Instrumentation Module (NIM) crate. An additional lpGBT, independent of the HGCAL front-end hardware, was used to sample the signal and transmit it via VTRx+ to the Serenity back-end through the fiber optic interface (not shown explicitly in the schematic in Fig. 6.47). In parallel, the ECON-T trigger was generated and tested against the external trigger, testing its self-triggered functionality. This is based on rather simple conditions on the energy deposited in the trigger cells (TC) and obtained from ECON-Ts. TPG-stage2 system was not implemented in the TPG firmware during this beam

test. For offline purposes, the information from both external and self-triggers was saved to cross-check the self-trigger's performance with external more stable trigger.

6.6.2 Milestones accomplished

We have recorded more than a billion events acquired using a self-triggering trigger and an external trigger based on a scintillator at an average trigger rate of 100 kHz. Different trigger input algorithms, including the Super Trigger Cell algorithm and the Best Choice algorithm, were tested. Additionally, DAQ data were acquired while running ECON-D in both pass-through mode and standard zero suppression mode, where a simple adjustable threshold was applied to reduce noise. The DAQ and Clock Distribution Hub (DTH) board, a central CMS board, was tested for the first time, and a bug causing the DTH to "freeze" and "stall" data acquisition was found in its firmware and reported to the experts. Data Quality Monitoring (DQM) was implemented for real-time online data analysis. The system reliably captured data during long 8-hour acquisition runs, recording up to 100 million events per run without human intervention. Different slow and fast control mechanisms were also tested.

6.6.3 Alternative approaches to examining raw data

From the readout electronics point of view, our goal is to acquire data more reliably without human intervention while minimizing data losses. We are particularly focused on the stability of the system during extended runs. From a system validation perspective, we are interested in,

- Events that exhibit mismatches among different identification counters,
- Instances of data corruption on links.
- Investigate the specific events that led to these errors.

6.6.4 Inspecting the raw data with finer details

For each L1A fast command, the components of the vertical chain (HGCROC $\rightarrow$ ECON-D $\rightarrow$ miniDAQ $\rightarrow$ Slink $\rightarrow$ DTH $\rightarrow$ DAQ-PC) receive packets from the previous component. At each such stage, different data quality and consistency checks including L1A timestamp, status or quality flags, and protective bits are evaluated and added in the corresponding Header and Trailer (H/T) words. In this section, we are more motivated in understanding the meta-data added by different ASICs and what it says about the overall system's functioning.

L1A time stamps

Each of HGCROCs, ECON-Ds, miniDAQ and Slinks have their internal event identification counters such as event counter, BX counter and Orbit counter defined as the following,

- Event counter : It is incremented every time whenever L1A Fast control (FC) command is received. The number of bits assigned for Event counter in ECON-D, miniDAQ and Slink is 6 bits, 6 bits, and 32 bits, respectively. The event counter keeps incrementing until the maximum size for assigned bits is achieved and rolls over to zero beyond that. For ECON-D and miniDAQ, it counts from 0 to 63, whereas for Slink, it counts from 0 to 4294967295.
- BX counter : It represents a time window of 25 ns or the bunch collisions. It varies from a minimum of 1 to a maximum of 3564. A 12-bit register is assigned for this across different components.

- Orbit counter : One Orbit consists of 3564 BXs. It increments by 1 when the BX counter reaches its maximum value and rolls over to 1. It serves as a sECONDary event tag and validates the proper event synchronization. The number of bits assigned for Orbit counter register in ECON-D, miniDAQ and Slink is 3 bits, 3 bits, and 32 bits, respectively. For ECON-D and miniDAQ, it counts from 0 to 7, whereas for Slink, it counts from 0 to 4294967295.

The BX and Orbit counters keep incrementing irrespective of whether L1A is received or not. The L1A time stamps provide a way to test the synchronization of different parts of the system. Once L1A is received, except HGCROC, each part of this chain (HGCROC $\rightarrow$ ECON-D $\rightarrow$ miniDAQ $\rightarrow$ Slink $\rightarrow$ DTH) expects an incoming packet with the corresponding event, BX, and orbit counter values for which L1A was generated. The importance and validation of time stamps across the different parts are discussed towards the end of this section.

From the perspective of HGCROC $\rightarrow$ ECON-D $\rightarrow$ miniDAQ

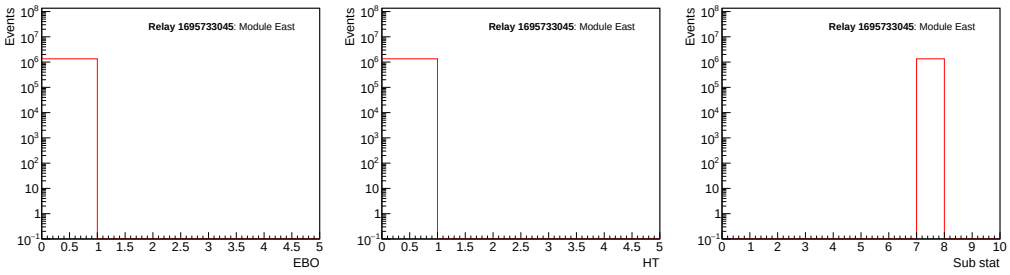


Figure 6.49: EBO field, HT field and sub stat bit field extracted from the ECON-D header for all events recorded in an electron of 100 GeV during beam test are shown.

The ECON-D header includes a 9 bit user-programmable pattern to signal the start of an incoming packet and fields to show the results of various checks performed. Two key fields, H/T (Header/Trailer) and E/B/O (Event/BX/Orbit), provide a 2-bit status for the reconstruction of the HGCROC Event H/T and E/B/O numbers across all active eRx links. A "0" status indicates perfect agreement among the eRx links or HGCROCs. Fig. 6.49 shows the values of EBO (left) and HT (right) during the physics run in the beam test indicating that all eRx are in perfect agreement.

ECON-D can be set to transmit the HGCROC's event counter values instead of its own and includes a flag (referred to as *M*) to indicate whether these transmitted values (HGCROC or its own) match the ECON-D's E/B/O values. This configuration was used during beam test runs to monitor the E/B/O values of HGCROC. Additionally, ECON-D performs a consistency check between the transmitted E/B/O values and the E/B/O for all eRx links. The 3-bit sub-stat status code indicates agreement: MSB for H/T match, middle bit for E/B/O match, and LSB for half-ROC CRC check. Fig. 6.49 (right) shows these bit values in a test beam run, confirming the consistency of eRx in the E/B/O values and the correct CRC check.

The payload length reflects the size of the packet, excluding the ECON-D header, and all of this information is forwarded to the miniDAQ system. Fig. 6.50 shows distribution of payload length in units of 32-bit words extracted from the ECON-D header, when ECON-D is in pass-through mode (left) and when zero suppression is applied (right). Reduced payload length indicates that zero suppression is working as intended. This has been checked across all the runs recorded during the beam test

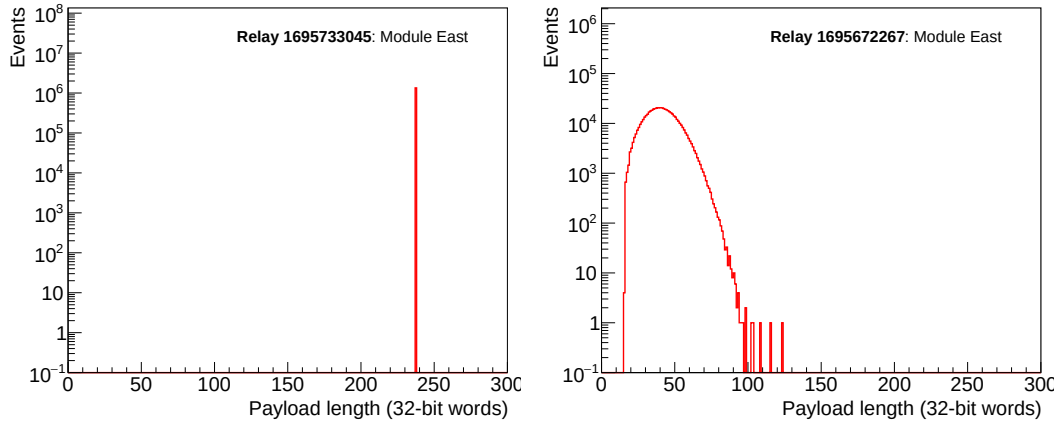


Figure 6.50: The payload length field values listed at the ECON-D header for all events recorded during a run when ECON-D is in pass through mode (left) and when in zero suppression mode (right).

experiment, and we have found it to be consistent. Details of individual event stamps as well as the *M* flag is be discussed later in the text.

From the perspective of ECON-D → miniDAQ → Slink

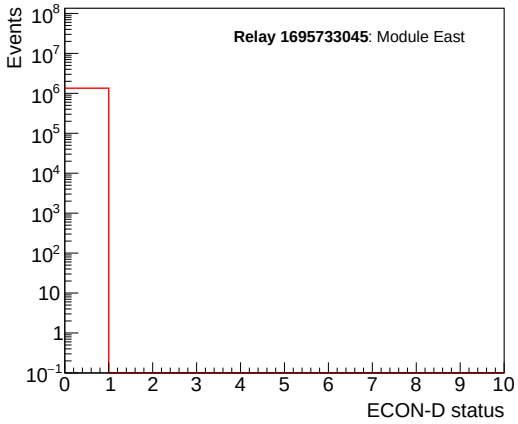


Figure 6.51: Value of ECON-D status bits field of the BE miniDAQ header for all events recorded in a run during beam test.

The miniDAQ puts two 32-bit words from ECON-D as on 64-bit word in its main buffer to be outputted to Slink, with padding if needed. A miniDAQ packet includes a 3-bit status field that represents the quality of the received ECON-D packets, with eight possible values. These values indicate whether a packet was detected, discarded due to exceeding the maximum size, or rejected due to mismatched E/B/O values. Additionally, the status field reflects the result of a CRC check on the ECON-D packet. One of the status options also accounts for scenarios where the event builder state machine timed out while waiting for a packet or when a buffer overflow occurred [149]. Fig. 6.51 shows the ECON-D status bits for all recorded events during the beam test. They are "0" for all events, indicating perfect packet reception at BE miniDAQ from ECON-D. All CRC checks at miniDAQ were also successful throughout the runs recorded in beam test.

From the perspective of miniDAQ → Slink → DTH

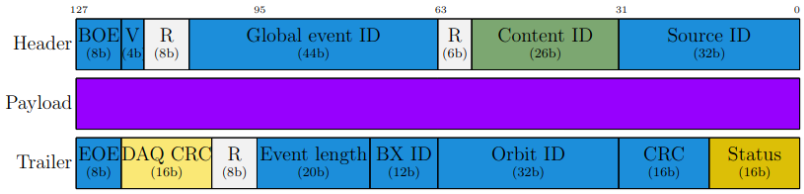


Figure 6.52: The slink packet format indicating different fields in its header and trailer [148].

The 128 bit wide data bus is used in the Slink core [148], hence the Slink readout firmware puts two 64-bit words into one 128-bit words with padding if needed. The Slink event data format consists of payload data that begin with a 128-bit header (beginning of event, BOE) and ends with a 128-bit trailer (end of event, EOE). Figure 6.52 illustrates the specifics of the Slink packet format, more details can be found at [148]. Slink packet header contains a 32-bit wide global event ID field that counts the total L1A issued by TCDS2 and is taken as a reference throughout the beam test. It also consists of a field named Emulation flag (E) indicating the source of L1A.
A good Slink packet always consists of a fixed programmed header and trailer pattern. During offline analysis, the Slink event counter was tested for packet drops, where it would increase by more than one instead of increasing by one. No such issues were observed, confirming that no packet loss was confirmed through miniDAQ → Slink,→ DTH. At various stages of the readout chain (HGCROC → ECON-D → miniDAQ → Slink → DTH), the Cyclic Redundancy Check (CRC) is performed to detect potential errors caused by noise or bit loss and to ensure that the data remain uncorrupted. No such error was detected at any point on the chain throughout the beam test.

Alignment of L1A time stamps in the HGCAL readout chain

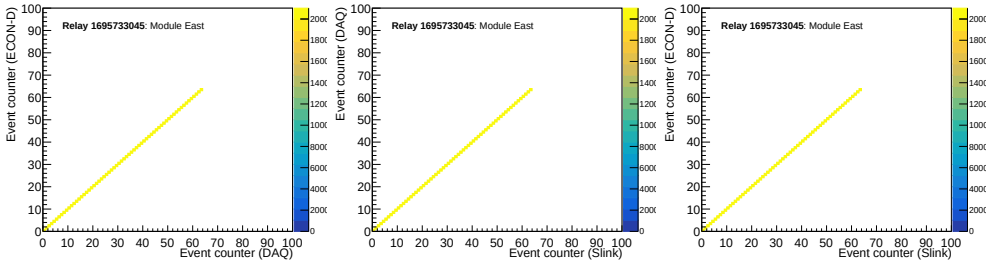


Figure 6.53: The plots compare Event counter values for ECON-D and miniDAQ (left), miniDAQ and Slink (middle), and ECON-D and Slink (right) for the east train.

We need to ensure that event ID counters across the system remain consistent. Misaligned timestamps can lead to incorrect event-trigger matching, impacting offline reconstruction and degrading physics performance, especially in new physics searches. Unsynchronized ASICs can compromise the reliability of data between CMS and LHC systems. This section will delve into details of E/B/O counters across the readout chain. During the beam test, the ECON-D was configured to output E/B/O values corresponding to HGCROC. Therefore, the ECON-D and HGCROC values, especially in the figures, are interchangeable unless otherwise specified.

Fig. 6.53 shows that the event counters are matching among ECON-D (HGCROC), miniDAQ and Slink for one of the runs recorded during beam test. It has been verified across all the runs recorded in the beam test. While extracting the Slink event counter (32-bit counter) to compare with ECON-D and miniDAQ event counter (3-bit), we take the modulus of the global event ID from Slink with 64. This ensures that the counter cycles between 0 and 63 as the event ID increases.

Fig. 6.54 illustrates the BX and the orbit counter for different components in ideal

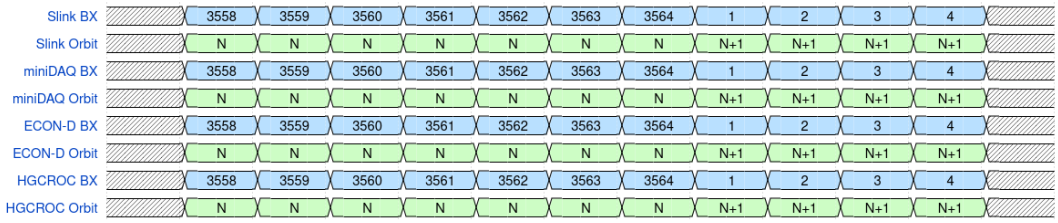


Figure 6.54: The above wave diagram shows BX and Orbit counters represented by N, in sync for Slink, miniDAQ, ECON-D and HGCROC in ideal scenarios.

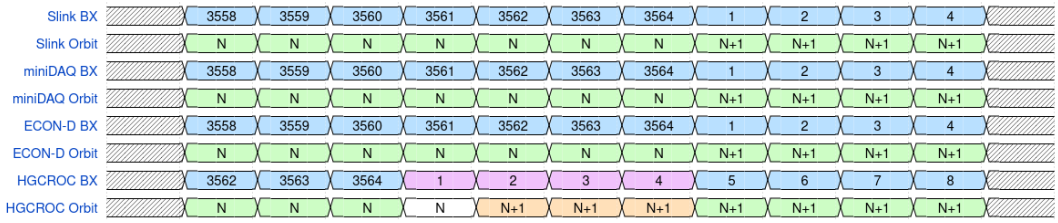


Figure 6.55: The wave diagram above shows the BX and Orbit counters for Slink, miniDAQ, ECON-D, and HGCROC, highlighting the artifacts in the BX and Orbit counters caused by HGCROC V3 design flaws, indicated by magenta and empty boxes. Here N is the current Orbit count and $N + 1$ is the incremented Orbit count.

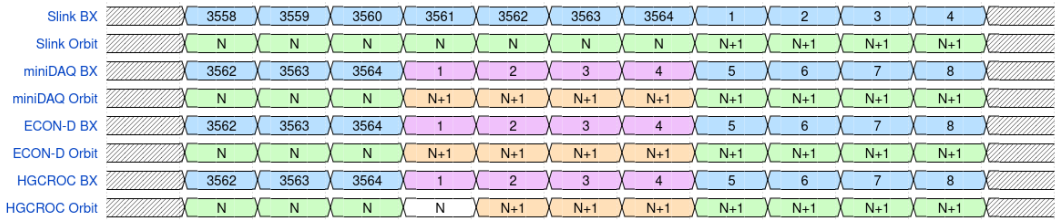


Figure 6.56: The wave diagram above shows the BX and Orbit counters for Slink, miniDAQ, ECON-D, and HGCROC in the actual test setup, where miniDAQ and ECON-D's BX counters have an offset of 4 to align with HGCROC's BX counter. The empty box highlights that the Orbit counter of HGCROC does not increment correctly at the BX rollover.

scenarios. The HGCROC V3 ASIC had two design flaws. The first, identified before the beam test, was the BX counter being 4 BX ahead, as seen in Fig. 6.55. During testing, it was found that the orbit counter incremented 1 BX late after the BX counter reset from 3564 to 0, as depicted in Fig. 6.55. These issues were reported and resolved in HGCROC V3b.

The purpose of the beam test was to investigate potential mismatches among timestamps after correcting for known discrepancies. To achieve this, an offset of "4" was

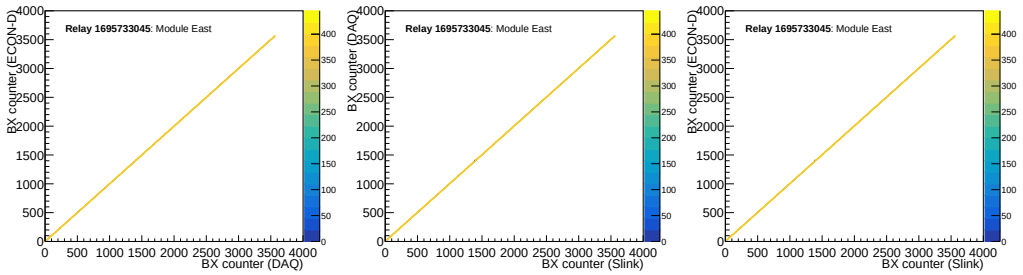


Figure 6.57: The plots compare BX counter values for ECON-D and miniDAQ (left), miniDAQ and Slink (middle), and ECON-D and Slink (right) for the east train. The Slink's BX counter is always 4 ahead of that from ECON-D (HGCROC) and miniDAQ.

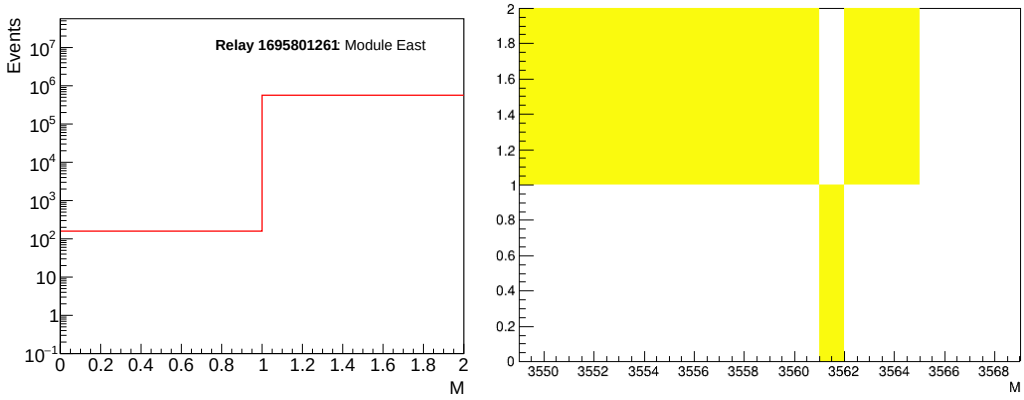


Figure 6.58: The plot shows the "M" bit from the ECON-D header, representing the agreement between HGCROC and ECON-D during the beam test. About 100 events occur at BX "3561" without agreement between ECON-D and HGCROCs.

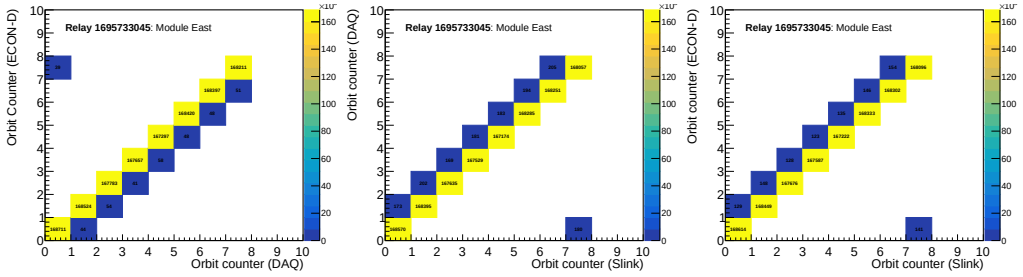


Figure 6.59: The plots compare Orbit counter values for ECON-D and miniDAQ (right), miniDAQ and Slink (middle), and ECON-D and Slink (left) for the east and west trains in the first and sECON-D columns, respectively.

applied to the BX counters of both the miniDAQ and ECON-D, ensuring synchronization across HGCAL-specific systems. Figure 6.56 presents a wave view of the BX and Orbit counters for the Slink, miniDAQ, ECON-D, and HGCROC, with manual offsets applied to the BX counters of the miniDAQ and ECON-D. The Slink counters were left unchanged and served as the true reference.

Figure 6.57 illustrates the alignment of the BX counters among ECON-D (HGCROC), miniDAQ, and Slink (with the offset of 4), based on one of the recorded runs during the beam test. This synchronization has been consistently verified across all runs conducted during the beam test.

The HGCROC orbit counter's inability to roll over affects all beam test runs, impact-

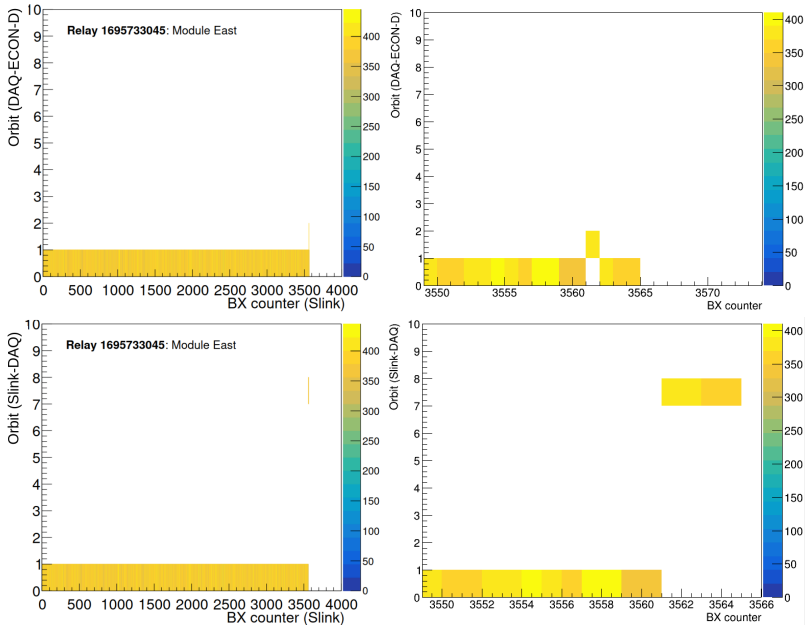


Figure 6.60: The difference between the Orbit counters versus the Slink BX where L1A is issued, for miniDAQ and ECON-D (top left), for Slink and miniDAQ (bottom left) are shown. Zoomed in version of the left plots are added on right side. (with a different Y axis label).

ing the readout chain. Within an orbit at 3561 BX (Beam Crossings) from Slink, discrepancies arise between the ECON-D and HGCROC orbit counters, as the HGCROC counter fails to roll over to the next orbit as shown in Fig. 6.56. The distribution of the field M shown in the plot on the left of Fig. 6.58, indicates that ECON-D and HGCROC are aligned in most cases. However, there is a small fraction where mismatches occur. These mismatches have been confirmed to correspond to events with a Slink BX counter value of "3561," as shown in the right plot of Fig. 6.58.

Consistency of Orbit counters among HGCROC → ECON-D → miniDAQ → Slink chain Orbit counter mismatches occur across different components when L1As (Level-1 Accepts) are issued at BX values beyond "3560," due to the HGCROC bug. Fig. 6.59 shows the comparison of orbit counter values for ECON-D and miniDAQ (top row), miniDAQ and Slink (middle row) and ECON-D and Slink (bottom row) for the east and west trains in the first and sECON-D columns, respectively. A significant fraction of events exhibit mismatches across the orbit counters of the different components. On closer inspection, these mismatches appear consistent; the orbit counter for ECON-D and HGCROC is always one count behind that of miniDAQ, while the Slink Orbit Counter lags behind both miniDAQ and ECON-D (HGCROC) by one count.
The Fig. 6.60 highlights the differences between Orbit counters versus Slink BX where L1As are issued. The top left plot compares miniDAQ and ECON-D, while the bottom left plot compares Slink and miniDAQ. The zoomed-in versions of these distributions on the right detail where the mismatches occur. Specifically, the mismatch between the Orbit counters of miniDAQ and ECON-D (HGCROC) happens at Slink BX "3561," where the HGCROC Orbit counter fails to roll over, lagging behind the miniDAQ Orbit counter. This occurs once per orbit when an L1A is issued at BX "3561". For the Slink versus miniDAQ comparison, the Orbit counter mismatch is due to the offset of 4 applied to the miniDAQ BX counter to align it with the HGCROC BX counter. As

a result, the Slink Orbit counter will always trail that of miniDAQ for these 4 BXs, as demonstrated in Fig. 6.60 (bottom right). Additionally, since the HGCROC orbit counter rolls over to the next orbit with a delay of 1 BX, the Slink Orbit counter will lag behind the ECON-D (reflecting HGCROC values) by only 3 BXs after "3561."

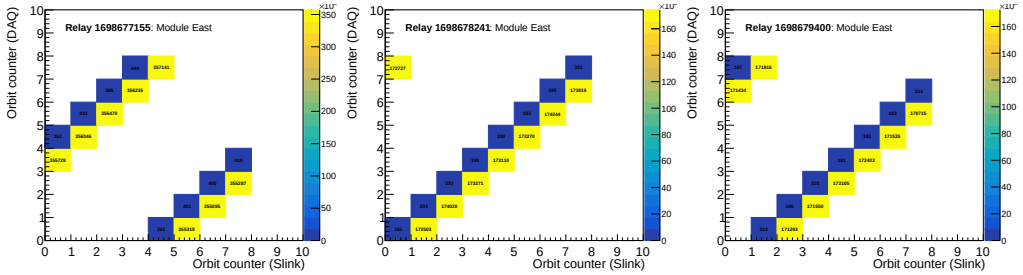


Figure 6.61: The plots compare Orbit counter values for miniDAQ and Slink for the east trains across three different relays recorded back to back after initialising the system in between.

Apart from the above-listed features, the miniDAQ orbit counter was changing from run to run, as observed in Fig. 6.61, indicating that the orbit counters between miniDAQ and Slink are out of sync by different offsets from run to run. This data shown in Fig. 6.61, was acquired with the internal software trigger at 100 kHz across multiple runs with reloading the system in between. Reloading the same system with the same firmware used in the test beam. The miniDAQ and its counters were not separated, which contributed to the synchronization problem. Before the acquisition of data begins, the event identification counters (event, orbit, BX) are reset and should only be reset with fast control (FC) commands irrespective of the system state. Every time miniDAQ enters reset mode, the counters are also reset, causing miniDAQ to go out of synchronization.

To resolve this, it is necessary to separate the miniDAQ resets and the counters in the firmware. Multiple sets of acquired data were taken with updated firmware while all other parameters were kept unchanged. The consistent pattern across multiple runs is obtained, as shown in Fig. 6.62.

In addition to all the studies listed above, a separate detailed check has been per-

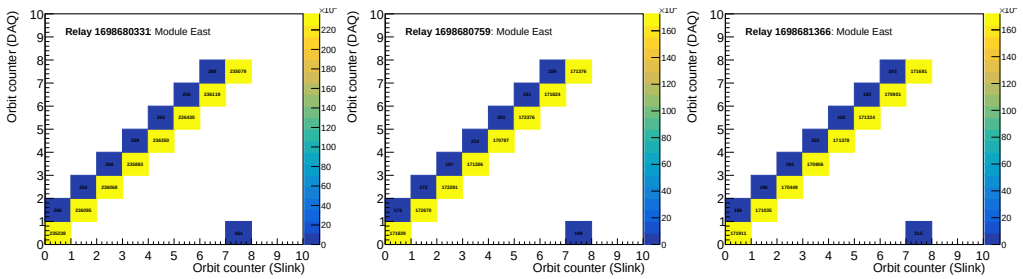


Figure 6.62: The plots compare Orbit counter values for miniDAQ and Slink for the east trains across three different relays recorded with updated firmware after taking the Orbit counters out of reset, and re-initialising the system in between.

formed at much higher trigger rates (up to 750 kHz) and different error rates along with systematic checks of data losses and bit errors are measured and discussed in length at [137].

6.7 Summary

In this chapter, various methods and tests were developed and validated to ensure the robustness of the HGCAL readout system. A correction technique was devised for module occupancies with mismatched geometries to accurately estimate the data rates. A method for generating correlated pseudo-events with high occupancy tails was successfully designed and tested. Emulator-based studies tests confirmed the system's ability to handle increased occupancies without significant data loss, while layer-wise validation ensured optimal mapping between ECON-D and lpGBT. Despite the observed ECON-D truncation and HGCROC throttling with high data rates and trigger rates above 900 kHz, respectively , no data loss occurred at the design rate and the dead time of the system remained manageable. Additionally, the final physics performance met expectations, with recent beam tests in 2024 validating fixes to HGCROC V3 ASIC design flaws and ensuring synchronized timestamps across components.

Chapter 7

New physics searches in multijet events with at least a photon and high transverse momentum imbalance

As discussed Chapter 2, searching for evidence of BSM physics is a prime mission of the LHC physics program. In this chapter, we present a search for new physics, guided by a few SUSY scenarios, in pp collision events at $\sqrt{s}$=13 TeV, characterized by at least one photon and substantial $p_{\mathrm{T}}^{\mathrm{miss}}$ in multijet events. We utilize full Run 2 collected during 2016-2018 corresponding to integrated luminosity of 137 fb^{-1}. The previous related search used $p_{\mathrm{T}}^{\gamma} > 100$ GeV [16] and Gluino masses of up to 2 TeV and electroweakino masses of up to 1 TeV have been probed. We use $p_{\mathrm{T}}^{\gamma} >$40 GeV, substantially reducing it to target phase space uncovered & new models not yet explored.

I also revisited the strategy used in recently published searches & propose to redefine the search regions using the newly found insight of the phase space available in this large dataset, which is explained in the following. The motivation and discussion on targeted physics scenarios for this search is done in Section 7.1 and optimization of SRs is discussed in Section 7.2. Methods to estimate backgrounds are discussed in Section 7.3. The interpretation of our results in the context of limits on masses of SUSY particles is discussed in Section 7.4.

7.1 Targeted physics scenarios & simplified SUSY models

We considered three SUSY scenarios to explore by lowering the p_{T}^{γ} and summarized in the following:

- The observed limit for different simplified models motivated by GMSB in previous search with high p_{T}^{γ} is worse for large mass differences between $\tilde{\chi}_1^0$ and $\tilde{g}$ ($m_{\tilde{g}}$-$m_{\tilde{\chi}_1^0}$) than the small mass differences, as shown in Fig. 7.1. Small $\tilde{\chi}_1^0$ masses lead to final state with softer photons.
- The pMSSM framework operates without any specific assumptions as discussed in Chapter 2 and this flexibility allows for diverse particle compositions for the LSP ($\tilde{\chi}_1^0$), which can be bino, Higgsino, or wino-like. In such scenarios, softer photons can emerge as part of the final state, especially when the mass difference between NLSP ($\tilde{\chi}_2^0$) and $\tilde{\chi}_1^0$ is small.
- Thus far, studies on SUSY utilizing photon signatures have largely focused on tree-level decays that produce high-energy photons. However, theoretical studies indicate that decays can also manifest via kinematical enhancement, where

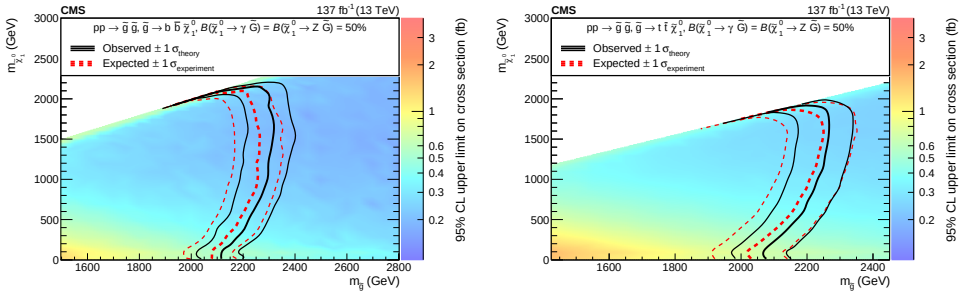


Figure 7.1: The observed and expected exclusion at 95% confidence level for T5bbbbZG (left) and T5ttttZG (right) achieved in previous round of analysis [16] with high p_T^γ. The exclusion limits are worse for smaller NLSP masses given Gluino mass.

a mass-degenerate neutralino spectrum triggers loop-level decays, resulting in low-energy photons [154].

Within the scope of this thesis, simplified models driven by GMSB used in [16] have been used to optimize search regions and determine analysis methodologies. The sensitivity of these optimized search regions for pMSSM-motivated models is discussed in Section 7.2.2. That for electroweakino models is elaborated in section 7.5 as future outlook for this analysis. The subsequent text elaborates on simplified models used in this analysis as well as the underlying assumptions made while generating event samples.

Gluinos and squarks carry color charges similar to their SM counterparts and are produced in strong interactions. We have considered four simplified models of strong production shown in Fig. 7.2. The gluino pair production model names are prefixed with T5 and stop ($\tilde{t}$) pair production using T6. In all the models, $\tilde{\chi}_1^0$ is a next-to-LSP (NLSP) which decays to a gravitino ($\widetilde{G}$) accompanied by one of the γ, Z or H boson assuming a 50% branching ratio (BR). In T5qqqqHG model, each gluino decays to a pair of light quarks and $\tilde{\chi}_1^0$ followed by $\tilde{\chi}_1^0 \to \gamma/\mathrm{H}$ along with $\widetilde{G}$. The minimum $\tilde{\chi}_1^0$ mass is set to 127 GeV ensuring Higgs is on-shell. In T5bbbbZG and T5ttttZG, the gluinos decay to a pair of b- and t-quarks along with a $\tilde{\chi}_1^0$ which decays to $\widetilde{G}$ & γ or Z , each with 50% BR. The T6ttZG model refers to pair production of top-squarks each decaying to t-quarks & $\tilde{\chi}_1^0$ which subsequently decays to a γ or Z along with a $\widetilde{G}$. In all these models, a wide range of $\tilde{\chi}_1^0$ masses is considered, extending down to as low as 10 GeV. The gravitino mass is fixed at 1 GeV in all the models based on GMSB. All SM particles are allowed to decay to any possible final-state particles according to their nominal BRs, offering a broad range of potential final-decay modes and final states.

Figure 7.3 illustrates simplified models for production of neutralino and chargino pairs, together called electroweakinos, designated as TChiWG and TChiNG. The masses of neutralinos and charginos are assumed to be nearly degenerate. The TChiNG model includes a variety of electroweakino production involving a nearly degenerate set of chargino and neutralino states, namely $\widetilde{\chi}_1^\pm\widetilde{\chi}_1^\mp$, $\widetilde{\chi}_1^\pm\widetilde{\chi}_1^0$, $\widetilde{\chi}_1^\pm\widetilde{\chi}_2^0$, and $\widetilde{\chi}_1^0\widetilde{\chi}_2^0$. The $\widetilde{\chi}_1^0$ is assumed to decay into gravitino with photons($\gamma\tilde{G}$), Z bosons ($Z\tilde{G}$), and Higgs bosons ($H\tilde{G}$) with branching ratios of 50%, 25% and 25%, respectively. In the TChiWG model, the branching ratios for neutralinos and charginos that decay into a $\gamma\tilde{G}$ and $W\widetilde{G}$ are set to 100%.

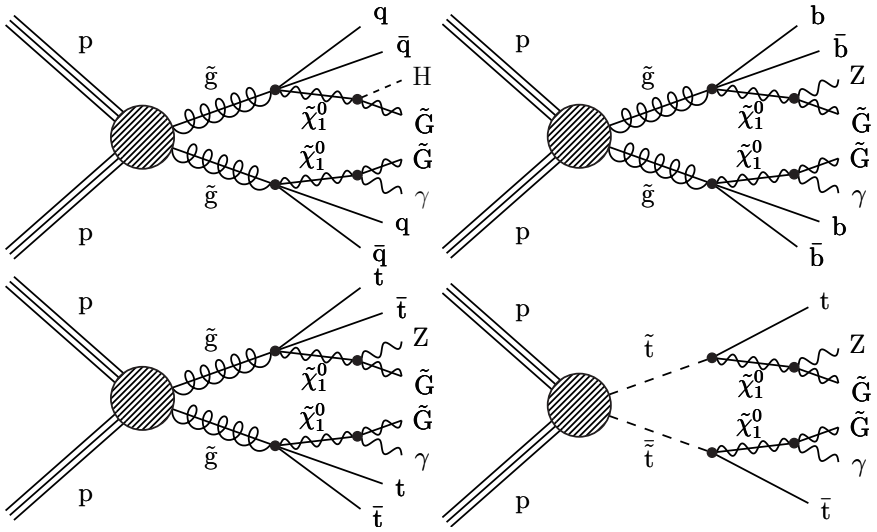


Figure 7.2: The strong production simplified SUSY models, T5qqqqHG (top left), T5bbbbZG (top right), T5ttttZG (bottom left), and T6ttZG (bottom right).

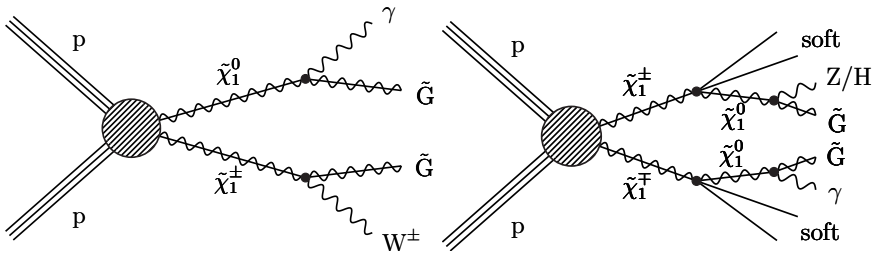


Figure 7.3: Simplified electroweak production SUSY models, TChiWG (left) and TChiNG (right).

Event samples for all these simplified models are generated using Fast Simulation and the reconstruction in CMSSW version 9_4_9 for 2016, 2017 and 10_2_*X* for 2018. The exact same samples have been used for the interpretation in [16].

T5gg model motivated by pMSSM scenarios is shown in Fig. 7.4. Event samples for this model are privately generated (thanks to Purbita) and are used to check the sensitivity of this analysis. In these models, $\tilde{\chi}_2^0$ is the NLSP and $\tilde{\chi}_1^0$ is the LSP (no $\widetilde{\mathrm{G}}$). It is a gluino pair production model in which each gluino decays to a pair of quarks along with the $\tilde{\chi}_2^0$ which eventually decays to a photon and a $\tilde{\chi}_1^0$ with a branching ratio 100%. Kinematics of final states largely depends on mass differences between $\tilde{g}$ & $\tilde{\chi}_2^0$ ($m_{\tilde{g}}$-$m_{\tilde{\chi}_2^0}$) and $\tilde{\chi}_2^0$ & $\tilde{\chi}_1^0$ ($m_{\tilde{\chi}_2^0}$-$m_{\tilde{\chi}_1^0}$). For our studies the gluino mass is fixed at 2.2 TeV, while the mass of $\tilde{\chi}_2^0$ is varied between 10 GeV and 1800 GeV. For each $\tilde{\chi}_2^0$ mass point, the parameter Δm is maintained consistently between $\tilde{\chi}_2^0$ and $\tilde{\chi}_1^0$. Subsequently, three distinct values for Δm are also examined: 10, 30, and 100.

In all these models, presence of $\widetilde{\mathrm{G}}$ or a pair of $\tilde{\chi}_1^0$ in final states, results in signature of large $p_{\mathrm{T}}^{\mathrm{miss}}$ as both these particles are electrically neutral & weakly interacting, not leaving a recordable signal in detector. We are interested in final states containing at least one photon which are also accompanied with quark & gluon jets and other bosons. The similar final states can arise from the SM processes as well. We veto the events containing an electron or muon which also helps in reducing the background due to SM process containing W (direct production or via top quark decay) or Z

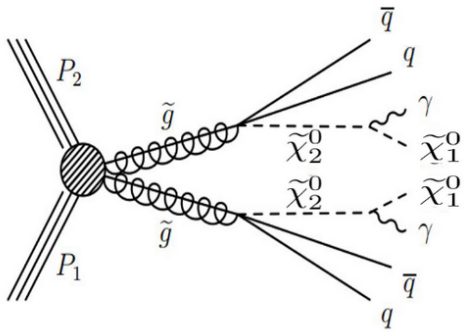


Figure 7.4: The T5gg model, inspired by scenarios within the pMSSM, illustrating the production of gluino pairs.

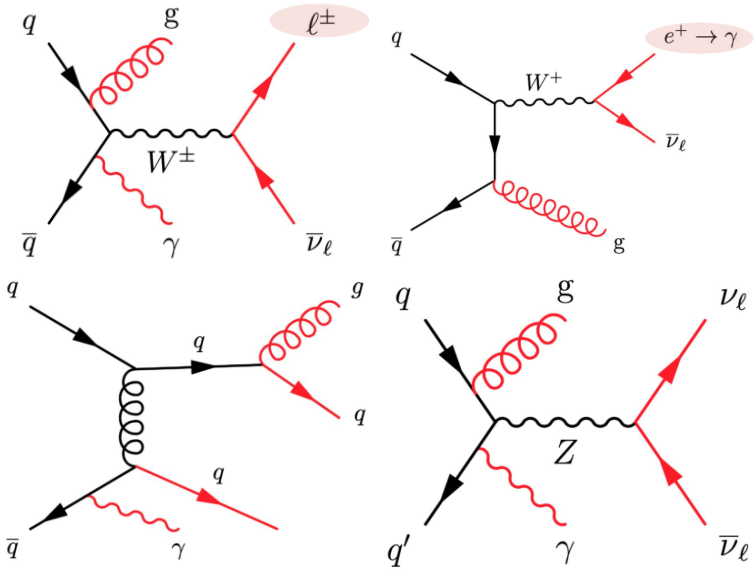


Figure 7.5: Standard Model background processes that feature at least one photon, $p_{\mathrm{T}}^{\mathrm{miss}}$ and large no of jets are shown above with $\mathrm{t\bar{t}}$+jets (top right), $\mathrm{t\bar{t}}\gamma$+jets (top left), QCD+γ+jets (bottom left) and $\mathrm{Z}(\nu\bar{\nu}) + \gamma$ (bottom right).

events as explained in the following section.

7.1.1 Background SM processes

The SM processes with a good reconstructed photon, high $p_{\mathrm{T}}^{\mathrm{miss}}$, multiple jets, and no lepton in the final states are background for this search, as events originated from these and the signal of interest are indistinguishable. Based on whether the background is different at the truth level but looks like our signal due to reconstruction effects or it is indistinguishable from our signal at the truth, backgrounds are classified as reducible or irreducible, respectively. The reducible can be mitigated by improving event selection criteria or reconstruction at the detector level, but no such thing can be made for irreducible backgrounds. Dominant background due to SM processes are discussed in the following. The following text discusses these in more detail.

$\mathrm{Z}(\nu\bar{\nu}) + \gamma$ process : The Z boson can decay to a $q\bar{q}$ pair with BR$\simeq$ 67 %, to a l^+l^- pair with BR$\simeq$9% and a $\nu\bar{\nu}$ pair with BR$\simeq$20% [155]. High $p_{\mathrm{T}}^{\mathrm{miss}}$ and zero lepton

requirements reject events with hadronic and charged lepton decay modes of Z boson. The only contributing decay mode to the SR is its invisible decay to the $\nu\bar{\nu}$ pair resulting in a true high $p_{\mathrm{T}}^{\mathrm{miss}}$ signature in the final states, and photons can be produced directly at the vertex or radiated from a quark, as shown in Fig. 7.5. This is an irreducible background. This background strongly dominates in the low-jet multiplicity. To measure this background, we use the $\mathrm{Z}(\nu\bar{\nu}) + \gamma$ MC events corrected for the MC data differences derived using dilepton events using $\mathrm{Z} \to \ell\ell\gamma$ in MC and the single-electron and muon trigger datasets in the data.

$\mathrm{t\bar{t}}$+jets and W+jets processes : The W boson can decay to a $q\bar{q}$ or $l\nu$ pair with BR≈67% and 33%, respectively. Top quark decays to a W boson and a b-quark with ∼100% BR, resulting in two W boson final states in $\mathrm{t\bar{t}}$+jets process. Hence, the $\mathrm{t\bar{t}}$+jets events can have fully hadronic (44%), single leptonic (44%) and dileptonic (11%) final states depending on W boson decays in the event. Presence of ν in final state is a source of true $p_{\mathrm{T}}^{\mathrm{miss}}$.
Most of the W+jets and $\mathrm{t\bar{t}}$+jets events are rejected by prompt photon requirement in the final state. However, the electron in such decay modes is mis-reconstructed as a photon because the associated track for an ECAL cluster was not reconstructed, then the final state of these processes will be high $p_{\mathrm{T}}^{\mathrm{miss}}$ due to neutrinos, electron faking as photon and multiple jets. These events will contribute to our SR as shown in Fig. 7.5. Several event selections are applied to reduce this background as discussed in Section 7.2.1. To measure this background, we are using the average transfer factors obtained from MC, the data-MC correction factor estimated using Tag & Probe (TnP) method with di-electron events in $\mathrm{Z} \to \ell\ell\gamma$ (for MC) and single-electron trigger data sets.

$\mathrm{t\bar{t}}\gamma$+jets and Wγ + jets processes : In Wγ + jets ($\mathrm{t\bar{t}}\gamma$+jets), there is a prompt photon at the truth level.Most of these events are rejected by lepton veto requirement. The leptons of W decay can be lost if e or μ are outside detector acceptance or failed identification or isolation criteria or there is a τ decaying hadronically (with BR≃65%), giving true $p_{\mathrm{T}}^{\mathrm{miss}}$ and jet signature in the final state. We are using the average transfer factor from MC to measure this background which are applied to single e and μ events in data. The TFs are calculated for data-MC differences.

QCD multijet and γ+jets : In QCD multijet and γ+jets processes, the photon can be radiated from quarks or promptly produced at vertex. These events contribute to the SRs only if true $p_{\mathrm{T}}^{\mathrm{miss}}$ is present due to semileptonic decay of heavy-flavor (B or C) hadrons or fake $p_{\mathrm{T}}^{\mathrm{miss}}$ if jet energy is mis-measured. Fake $p_{\mathrm{T}}^{\mathrm{miss}}$ contributed more in events with zero b-tagged jets and for events with more b-tagged jets, true $p_{\mathrm{T}}^{\mathrm{miss}}$ from semileptonic decay contributes more. The $p_{\mathrm{T}}^{\mathrm{miss}}$ will be aligned to the direction of jets, hence a $\Delta\phi$ between jets and $p_{\mathrm{T}}^{\mathrm{miss}}$ will be small, and we reject such events to reduce the contribution from these background. More details of such event selection will be discussed in Section 7.2.1. For the studies presented here, we have used background event yields from the MC simulated events. Several data driven methods such as Rebalance & Smearing (R&S) and ABCD method can give more robust measurement of this background.

7.1.2 MC event simulation & dataset generator

We simulate the production of these processes in pp collisions using sophisticated tools like `MADGRAPH 5_aMC@NLO`, PYTHIA, and GEANT4. These tools are designed

to model specific aspects, from parton-level matrix element calculations to simulation of realistic detector effects. MC simulations are crucial for deriving correction and scale factors, validating techniques for background estimation, and evaluating systematic uncertainties.
In this study, MADGRAPH5_aMC@NLO serves as the main simulation tool for both signal and background events. For background samples, we have used version 2.6.5 for all three years. Leading-order (LO) simulations are used for processes such as $t\bar{t}$+jets, W+jets, $Z \to \nu\bar{\nu}$, QCD, and γ+jets, while $Z(\nu\bar{\nu}) + \gamma$ and $t\bar{t}\gamma$+jets processes are simulated in the next-to-leading order (NLO). In generating background samples, PYTHIA 8 is used for parton shower and hadronization processes, applying the CP5 underlying event tunes consistently across all years. We used MLM and FxFx matching schemes in LO and NLO samples to avoid double counting of partons generated both in MADGRAPH5_aMC@NLO and PYTHIA. Signal events are generated using MADGRAPH5_aMC@NLO with leading order (LO) precision, incorporating up to two additional partons in matrix element calculations. The decays of SUSY particles, such as gluinos, top squarks, and neutralinos, are simulated with PYTHIA 8, adopting version 8.212 for the 2016 dataset and 8.230 for data spanning 2017 to 2018. Signal normalization coefficients are determined from cross sections computed in the next-to-leading order (NLO). Signal yields are normalized using cross sections calculated at NLO with next-to-leading logarithmic (NLL) precision. This extensive simulation framework is essential for accurately modeling signal and background contributions, thereby allowing for an accurate comparison with experimental data and strengthening the reliability of the analysis.
We have used both SM MC samples and data processed with Ultra-Legacy (UL) reconstruction.
The $t\bar{t}\gamma$+jets and $t\bar{t}$+jets correspond to the same tree-level process; however, $t\bar{t}\gamma$+jets incorporates an additional photon vertex. During simulation of $t\bar{t}$+jets, events associated with $t\bar{t}\gamma$+jets are generated consistent with their cross section, resulting in the replication of events within the MC datasets $t\bar{t}$+jets and $t\bar{t}\gamma$+jets. At the MADGRAPH5_aMC@NLO level, a $\Delta R(\ell/q, \gamma) < 0.5$ cut is implemented for $t\bar{t}\gamma$+jets or $W\gamma$ + jets processes. Nevertheless, such cuts are inadequately modeled in the $t\bar{t}$+jets or W+jets datasets. The lack of radiated photons at the MADGRAPH5_aMC@NLO level within the simulated data sets complicates the integration of the $t\bar{t}$+jets (W+jets) and $t\bar{t}\gamma$+jets ($W\gamma$ + jets) processes. To mitigate this, specific selection criteria $\Delta R(\ell, \gamma_{\text{reco}}) > 0.5$ and $\Delta R(q/g, \gamma_{\text{reco}}) > 0.5$ are applied to $t\bar{t}\gamma$+jets ($W\gamma$ + jets) and inversely to $t\bar{t}$+jets (W+jets). This method effectively minimizes redundancy between these two types of datasets. This approach is consequently applied to additional samples as follows: $Z \to \nu\bar{\nu}$ & $Z(\nu\bar{\nu}) + \gamma$, W+jets & $W\gamma$ + jets, QCD & γ+jets, and $Z \to \ell\ell\gamma$ & Z+jets. This method was developed and well tested and discussed in detail in [156].

Data collected with MET triggers have been used to define designated signal regions and control regions. Trigger efficiencies were determined according to [156] for Ultra-Legacy (UL) samples [Thanks to Purbita]. In summary, the computation of $p_{\text{T}}^{\text{miss}}$ trigger efficiencies references single electron datasets, capitalizing on electron triggers that exhibit reduced p_T thresholds (HLT_Elex_WPTight_Gsf_v), specifically with $x = 27, 30, 32$ for the year 2016, and $x = 27, 32$ applicable to the years 2017 and 2018.Further selection criteria applied to the single electron dataset include:

- $S_{\text{T}}(p_{\text{T}}^{\gamma} + \sum p_{\text{T}}^{\text{jets}}) > 300$ GeV,
- At least one photon with $p_{\text{T}}^{\gamma} > 40$ GeV and $\eta^{\gamma} < 2.4$,
- $N_{\text{jets}} \geq 2$,

- Azimuthal angle separation $\Delta\phi(\text{jet}, p_{\mathrm{T}}^{\mathrm{miss}}) > 0.3$ for the first two leading jets, with default $p_{\mathrm{T}}^{\mathrm{miss}}$ cleaning applied.

The $\Delta\phi$ criterion imposed on the leading jets is implemented to ensure dominance by genuine leptonic decay, thus mitigating the influence of jet mis-measurement, which is frequently encountered in QCD multijet processes.

7.2 Composition of signal and background

The event selections are devised in such a way that we target a wide variety of SUSY models with our signal region while minimizing the inclusion of SM backgrounds. Hence, one should not opt for a fine-tuned set of event selections based on particular signal models. The following section discusses the details of the objects selected in an event and the set of baseline selection applied to minimize the background in the search regions.

7.2.1 Baseline objects and event selections

Apart from the initial reconstruction discussion in Section 3.2.6, each object is identified using a set of fine-tuned selection criteria to obtain good efficiency and minimal fake reconstruction. In signal regions, we are interested in selecting good reconstructed photons, jets, b-tagged jets, and $p_{\mathrm{T}}^{\mathrm{miss}}$ in the event and leptons to veto events. Their identification is discussed briefly in the following text.

Identification of photons : Based on the properties of the ECAL supercluster, several reconstruction IDs with different working points to be used according to the specific need of the analysis are available for photons. In this analysis, we have used the CMS POG recommended `loose` working point (with $\sim 90\,\%$ efficiency) of cut-based identification (ID) which uses criteria of very small values of H/E and its transverse shower spread ($\sigma_{i\eta\, i\eta}$). In addition, a strong isolation criterion is applied based on the sum of PF candidates with a cone of $\Delta R = 0.3$ centered around the photon candidate. The superclusters to be found in the ECAL crack region ($1.44 < |\eta| < 1.56$) are not considered for photon reconstruction. In addition to the above, we select photons that are within the tracker acceptance ($|\eta| < 2.4$) and with $p_{\mathrm{T}} > 40\,\mathrm{GeV}$. To reduce the contamination by false photons in our search region, the photon candidates associated with hit in any of the pixel layers but no final track reconstructed are not considered.

Good jets selection : AK4 jets with η within 2.4 and $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ are considered in this analysis. The jets that match with good photon within $\Delta R < 0.3$ are removed from the good jet collection. To avoid spurious jets that arise due to detector noise or reconstruction issues, AK4 jets passing a loose ID criterion are considered. This loose ID [157] ensures that the energy fraction of charged hadrons is > 0, of the contribution of neutral hadron contribution <0.99, of charged & neutral electromagnetic (EM) < 0, and with more than 1 multiplicity of charge particles. Event with $N_{\mathrm{jets}} \geq 2$ are considered, else rejected.

Angular separation of jets and $\boldsymbol{p_{\mathrm{T}}^{\mathrm{miss}}}$: The fake $p_{\mathrm{T}}^{\mathrm{miss}}$ that arises in the QCD processes are expected to be aligned with the mismeasured jets. To mitigate this contribution, we require that the $\Delta\phi$ between the $p_{\mathrm{T}}^{\mathrm{miss}}$ and the first two leading jets be greater than 0.3.

Total visible energy (S_{T}) : The scalar p_{T} sum of all jets and photons selected with above mentioned criteria, ($\sum_i p_{\mathrm{T}}^{jet} + p_{\mathrm{T}}^{\gamma}$) within an event is defined as the total visible

energy (S_T). Events with $S_T > 300$ GeV are used, since the SUSY particles targeted in this study possess masses of up to 1 TeV. Consequently, S_T for the signal is expected to predominate on the higher side.

Electron veto : Electrons should pass the CMS POG recommended `VETO` working point criteria of cut based ID with ∼95% efficiency for electrons. Other than the shower shape variables, conversion veto ensures that converted photons in tracker are not mis-reconstructed as electron. The electron candidate should have $p_T >$ 10 GeV and is within $|\eta|$ <2.5 . Alongside prompt electrons that result from the leptonic decay of heavy bosons, electrons may also emerge from the decay of b-hadrons with significant hadronic activity around them. Electrons should satisfy an isolation criterion of $I_{mini} < 0.1$, where I_{mini} is defined as in [158]. This ensures efficient distinguishing of prompt muons as compared to those originating from b-hadron decays. Events containing such electron candidates are excluded from the search region and incorporated into the control region for background estimations.

Muon veto : Muon candidates should satisfy POG recommended `Medium Muon` selection with additional condition on the muon track's distance from the primary vertex (PV), d(μ,PV) should be within 0.2 (0.4) cm in the transverse and longitudinal direction, respectively. This ensures that the selected muon is a prompt muon. A good muon should have $p_T > 10$ GeV and $|\eta|$ <2.5 . Similarly to the electron, we apply a minimum isolation criterion on the muons where I_{mini} is defined in [158]. The size of the isolation cone in mini-isolation depends on the transverse momentum (p_T) of the muon. For muons with $p_T < 10$ GeV, the cone size is fixed at 0.2. In the range $p_T \in [50, 200]$ GeV, the cone size is determined by 10 / p_T. For $p_T > 200$ GeV, the cone size is set to 0.05. Events containing such muon candidates are excluded from the search region and incorporated into the control region for background estimations (see the later sections).

Isolated tracks vetoes : This criteria acts as a complement to the standalone lepton veto, aimed at rejecting events where any lepton-induced event does not satisfy the lepton veto due to errors in lepton identification or does not meet isolation criteria. Additionally, it contributes to reducing the background from hadronic tau decays by filtering out events with an isolated charged-pion trajectory. To comply, the lepton track must fulfill the conditions $p_T > 5$ GeV and $I_{tk} < 0.2$, whereas the hadronic track, predominantly a pion, must conform to $p_T > 10$ GeV and $I_{tk} < 0.1$. Track isolation I_{tk} is defined as the ratio between the sum of scalar p_T charged tracks located within $\Delta R < 0.3$ of the primary track and the p_T of the primary track.

MET : p_T^{miss} is defined as the negative vector sum of all PF candidates p_T ($\sum_{PFcandidates} \vec{p}_T$). Events with $p_T^{miss} > 200$ GeV are selected for this analysis. This selection is derived from the signal of interest and trigger efficiencies. It was noted trigger efficiency are >97 % for p_T^{miss} >300 GeV and goes upto 75% for p_T^{miss} >200 GeV.

Tagging of b-jets : We are using the deep machine learning version of the combined secondary vertex (DeepCSV) [159] algorithm to tag a good AK4 jet as b-tags. It uses a variety of variables associated with the secondary vertices (SV) associated with the jet and tracks originating from secondary vertices (SV), such as the transverse energy of the tracks associated with the SV, the angular separation relative to the jet axis, the track p_T and its decay length. The POG recommended medium working point [160]

with ∼70% efficiency to tag b-jets with 1% misidentification probability of light jets, is used for this search. The DeepCSV tagger score should be more than 0.60, 0.58, 0.45, and 0.41 for the datasets `2016preVFP`, `2016postVFP`, `2017 and 2018`, respectively, for a jet to be tagged as bjet. As signal models of interest are enriched with final states of bjets, dividing the search bins by the number of b-tagged jets (N_{bjets}) helps in obtain better signal sensitivity.

MET filters : Before any physics analysis, it is crucial to understand the detector data to prevent misinterpretations of false signals [161]. Data Quality Monitoring (DQM) certified datasets can still exhibit anomalous MET events, stemming from detector issues such as detector noise, nonfunctional calorimetric channels, etc., rather than actual physical phenomena. Such events can give false signatures of new physics events. To avoid this, we used the following recommended MET filters [162].

- globalSuperTightHalo2016Filter : To mitigate the impact of high-energy muons resulting from the interaction of incident protons with the gas within the beam pipe and its associated materials, which cannot be effectively halted by the shielding.
- HBHENoiseFilter & HBHEIsoNoiseFilter : To eliminate the spurious $p_{\text{T}}^{\text{miss}}$ produced in the HCAL barrel and endcap as a result of magnetic-field-dependent noise in hybrid photo detectors (HPDs), ion feedback noise in HPDs, and collective readout noise across all HPDs.
- eeBadScFilter : apply only in data
- EcalDeadCellTriggerPrimitiveFilter : This filter flags and removes events in which the trigger primitive value approaches saturation limit (127.5 GeV) in channels without data links and are likely leading to fake $p_{\text{T}}^{\text{miss}}$.
- ecalBadCalibReducedFilter : Exclude events in which $p_{\text{T}}^{\text{miss}}$ emerges as a consequence of ECAL cell miscalibration induced by intense radiation exposure. This procedure has been used for the years 2017 and 2018.
- BadPFMuonFilter : Exclude events that feature muons which do not satisfy the Particle Flow (PF) requirement or exhibit a substantial uncertainty in their p_{T}.
- Good vertex filter : Ensure that the event has a well reconstructed primary vertex.
- PFCaloMET : The ratio of MET derived from PF objects to that reconstructed using calorimeter objects (CaloMET) should remain within a reasonable range and not fall beyond 5.

During Run 2 operations of LHC, several detector issues were observed that could not be fixed online and should be taken care of during offline analysis if the analysis is sensitive to those. Some of them are listed here, and more details can be found in [163].

- L1-prefiring issue : Throughout the 2016 and 2017 data collection period, the timing offset for the ECAL was improperly incorporated into the L1 trigger primitives (mostly ECAL noise related). Consequently, EGamma objects with $\eta > 2.0$ were linked to earlier bunch crossings, leading to "self-veto" occurrences due to L1 trigger constraints. This pre-firing inefficiency was evaluated using MC simulations and corrected in the data accordingly.
- HEM-veto : This is applied to mitigate the effect of two-disk failure in the HCAL endcap in terms of increased fake leptons and under measured jet

energy contributing to $p_{\mathrm{T}}^{\mathrm{miss}}$. Events were vetoed if any of the electrons (jets) satisfy the conditions given in equation 7.1 (7.2).

$$p_{\mathrm{T}}^{elec} > 30\,\mathrm{GeV}\ \&\ -3.0 < \eta^{elec} < -1.4\ \&\ -1.57 < \phi^{elec} < -0.87 \tag{7.1}$$

$$p_{\mathrm{T}}^{jet} > 30\,\mathrm{GeV}\ \&\ -3.2 < \eta^{jet} < -1.2\ \&\ -1.77 < \phi^{jes} < -0.67\ \&\ \Delta\phi(jet, p_{\mathrm{T}}^{\mathrm{miss}}) < 0.5 \tag{7.2}$$

These filters and cleaning procedures are more important for data and have minimal impact on event yields in MC simulations. The cutflows for background and signal events, including various baseline selections, are detailed in tables 7.1 and 7.2.

Table 7.1: Cutflows for all backgrounds after baseline selections.

Selections	Wγ	$t\bar{t}+\gamma$	W+jets	$t\bar{t}$+jets	Z($\nu\bar{\nu}$)	γ+QCD jets
No cut	887653	134585	1.27e+07	4.73e+06	1.74e+07	2.22e+07
Lepton veto	546141	84232.5	6.36e+06	1.80e+06	8.16e+06	9.96e+06
iso e/$\mu\pi$ track veto	398636	63796.6	4.44e+06	1.10e+06	5.54e+06	6.64e+06
$p_T^{\gamma} > 40$ GeV	253204	24656.7	345686	87760.9	4.33e+05	521207.8
$p_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV	2251.28	405.824	5910.73	2760.63	8.67e+03	11431.99
$N_{\mathrm{jets}} > 2$	1334.61	388.926	3820.8	2672.89	6.49e+03	9166.58
$S_T > 300$ GeV	1163.14	364.348	3431.78	2453.59	5.89e+03	8338.96
Event cleaning	1147.7	359.46	3390.05	2421.29	5.81e+03	8232.63
$\Delta\phi$(2 lead jets,$p_{\mathrm{T}}^{\mathrm{miss}}$)$\geq$0.3	992.726	283.909	3142.62	2131.56	5.27e+03	7405.74
remove overlap	881.11	256.119	2419.91	1618.98	4.04e+03	5657.87
Trigger eff.	752.213	215.421	2,075.35	1,347.98	3,423.33	4,771.31

Table 7.2: Cutflows for T5bbbbZG & TChiWG after baseline selections.

	T5bbbbZG $M(\tilde{g})$ = 2200 GeV		T6ttZG $M(\tilde{t})$ = 1300 GeV	
Selections	$M(\tilde{\chi})$ = 10 GeV	$M(\tilde{\chi})$ = 1000 GeV	$M(\tilde{\chi})$ = 10 GeV	$M(\tilde{\chi})$ = 1000 GeV
No cut	49.2948	49.2948	116.659	116.659
Lepton veto	59.2842	63.4292	91.1801	91.833
iso e/$\mu\pi$ track veto	57.3009	62.3743	85.164	86.1583
$p_T^{\gamma} > 40$ GeV	35.5968	42.7823	54.8791	56.8435
$p_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV	31.8309	42.4574	50.0317	56.265
$N_{\mathrm{jets}} > 2$	29.4572	41.0919	45.6362	53.7493
$S_T > 300$ GeV	33.3581	42.5834	51.2636	55.6529
Event cleaning	33.3581	42.5834	51.2636	55.6529
$\Delta\phi$(2 lead jets,$p_{\mathrm{T}}^{\mathrm{miss}}$)$\geq$0.3	30.7146	40.8257	47.0528	53.3508
remove overlap	29.4256	39.7135	45.6897	51.8649
Trigger eff.	11.3439	18.3553	16.8603	23.2175

7.2.2 Optimization of search region bins

With all baseline selections, Fig. 7.6 shows the signal and background composition as functions of different kinematic variables. The solid fill histograms represent different SM processes and the other non-filled histograms show T5bbbbZG signal points with different $m_{\tilde{\chi}_1^0}$ given fixed $m_{\tilde{g}}$ at 2.2 TeV. Both SM backgrounds and signal events are normalized to $137 fb^{-1}$ and are weighted with their respective cross sections. The background composition changes depending on the kinematic phase space. Different SM backgrounds are dominant in different phase space, such as $Z(\nu\bar{\nu}) + \gamma$ in low N_{jets}, QCD+γ+jets in low $p_{\text{T}}^{\text{miss}}$ and high S_{T}, $t\bar{t}\gamma$+jets and $t\bar{t}$+jets in high N_{jets}.
In case of signal, $p_{\text{T}}^{\text{miss}}$ and N_{jets} (top row of Fig. 7.6) do not discriminate between the different $m_{\tilde{\chi}_1^0}$ signal points, although these variables do discriminate background and signal in the high $p_{\text{T}}^{\text{miss}}$ and high N_{jets} regions. The bottom row of Fig. 7.6 shows the variables S_{T} and p_{T}^{γ}. At 2.2 TeV $m_{\tilde{g}}$, if the mass splitting (Δm) between $\tilde{\chi}_1^0$ and $\tilde{g}$ is high, resulting in a small $m_{\tilde{\chi}_1^0}$ and more available energy for $b\bar{b}$ pair, the corresponding signatures in detector are, high ST & high N_{jets}, and low p_{T}^{γ}. This is clearly demonstrated in the blue line in Fig. 7.6. The relation between N_{jets} and S_{T} can be tricky as for the same S_{T} one can have 10 N_{jets} with a smaller p_{T} or 4 N_{jets} with high p_{T}, and their respective probability depends on the strong coupling. On the other hand, with small Δm with high $m_{\tilde{\chi}_1^0}$, and small fraction of $m_{\tilde{g}}$ available for $b\bar{b}$ pair, the signatures in the detector are low S_{T} and high p_{T}^{γ} shown in the red and magenta line in Fig. 7.6. In addition to $p_{\text{T}}^{\text{miss}}$ and N_{jets}, the distribution of S_{T} and p_{T}^{γ} effectively separates the signal and background, as well as the small $m_{\tilde{\chi}_1^0}$ from the high $m_{\tilde{\chi}_1^0}$.

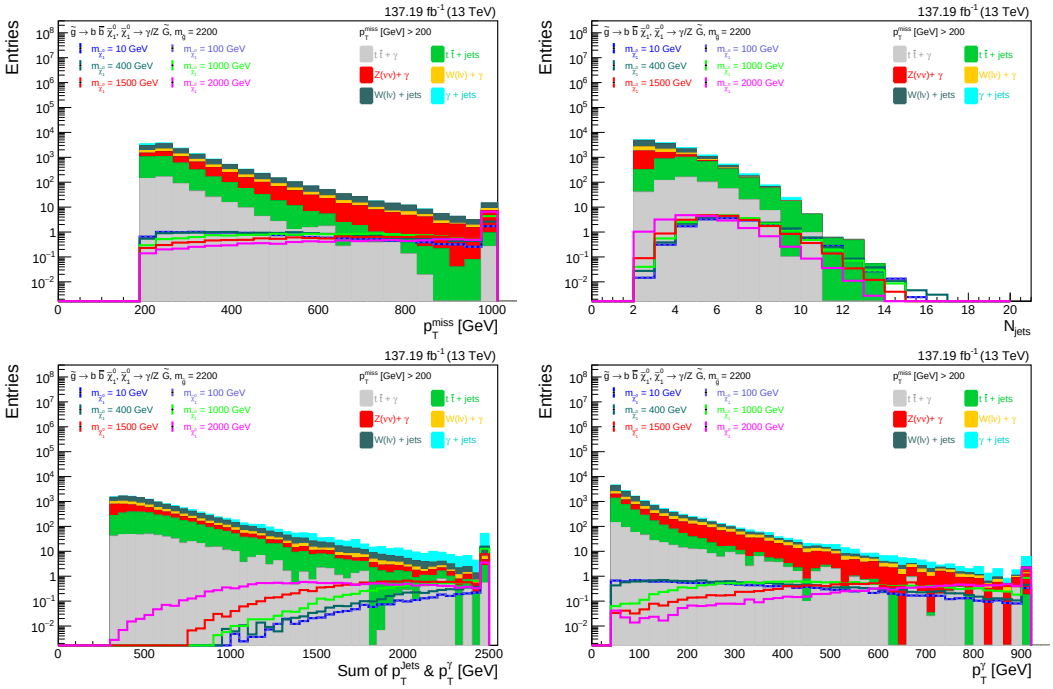


Figure 7.6: Distribution of $p_{\text{T}}^{\text{miss}}$ (top left, N_{jets} (top right), S_{T} (bottom left) and p_{T}^{γ} (bottom right) for SM backgrounds and different SUSY signal topologies after applying baseline selections.

To understand more of these correlations between kinematics, the distributions of $p_{\text{T}}^{\text{miss}}$ versus p_{T}^{γ} (top row),S_{T} versus p_{T}^{γ} (middle row), and S_{T} versus N_{jets} (bottom row)

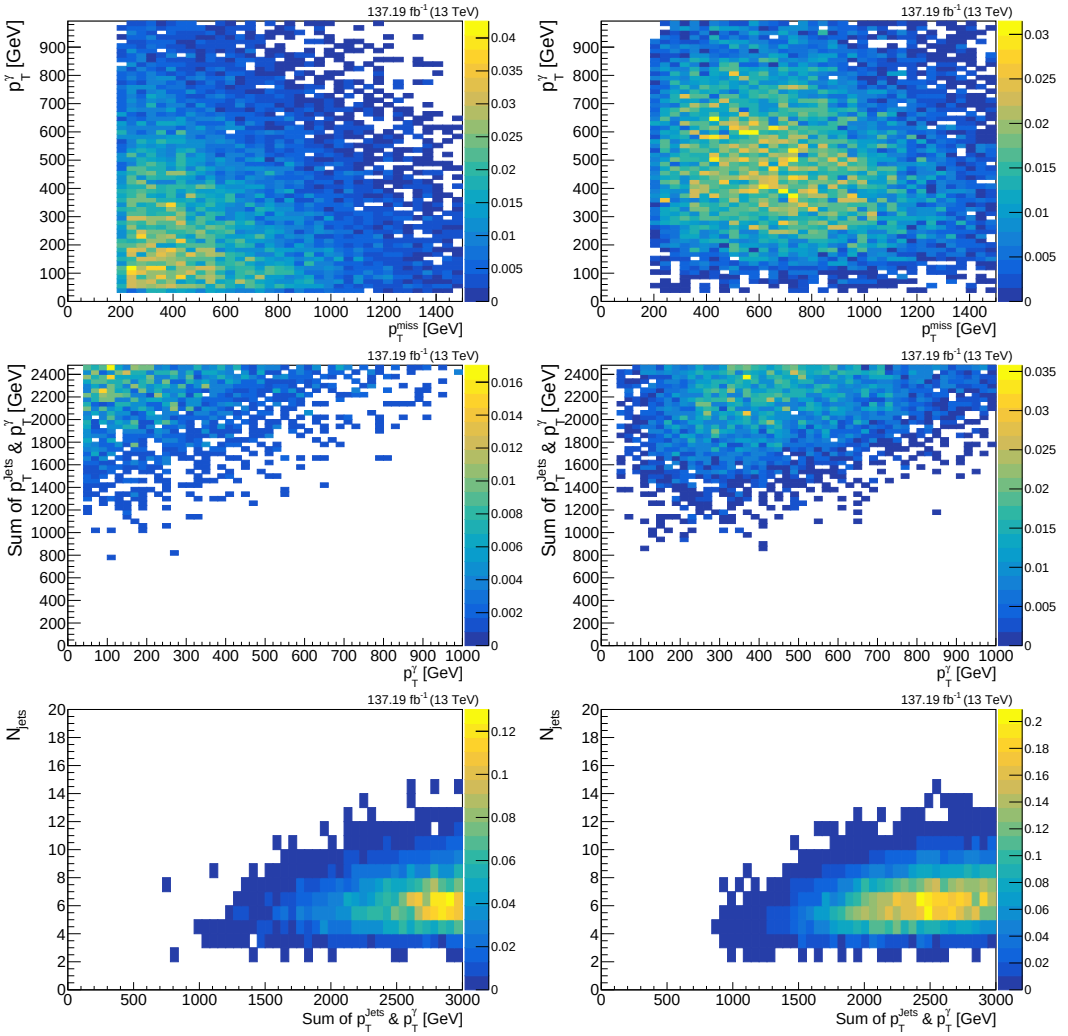


Figure 7.7: The distributions of $p_{\mathrm{T}}^{\mathrm{miss}}$ versus p_{T}^{γ} (top row) and S_{T} versus p_{T}^{γ} (middle row), and S_{T} versus N_{jets} (bottom row) with $m_{\tilde{\chi}_1^0} = 10\,\mathrm{GeV}$ in the first column (indicating high mass splitting) and $m_{\tilde{\chi}_1^0} = 1000\,\mathrm{GeV}$ in the second column (indicating low mass splitting) are shown.

for $m_{\tilde{\chi}_1^0} = 10\,\mathrm{GeV}$ (first column - high mass splitting) and $m_{\tilde{\chi}_1^0} = 1000\,\mathrm{GeV}$ (second column - low mass splitting) are shown in Fig. 7.7. Both p_{T}^{γ} and $p_{\mathrm{T}}^{\mathrm{miss}}$ move to the higher side with increase in $m_{\tilde{\chi}_1^0}$ and no direct correlation is observed. This helps in the case of lowering p_{T}^{γ} but still keeping the search region focused on high $p_{\mathrm{T}}^{\mathrm{miss}}$. On the other hand, S_{T} versus p_{T}^{γ}, softer p_{T}^{γ} with high S_{T} and otherwise.

In order to enhance the sensitivity of detecting the signal relative to the background, the search regions are typically subdivided into a multitude of smaller regions. These subdivisions are systematically formulated based on distinct kinematic variables, particularly those that offer greater discriminatory capability between signal and background.

In the previous round of searches based on similar final states, high p_{T}^{γ}, $p_{\mathrm{T}}^{\mathrm{miss}}$, N_{jets} and N_{bjets} were used as search variables. Such strong production motivated bins (SP bins) boundaries [16] used previously are shown in Fig. 7.8 (left). The plot on the left and right of Fig. 7.9 shows the event distribution from different background and

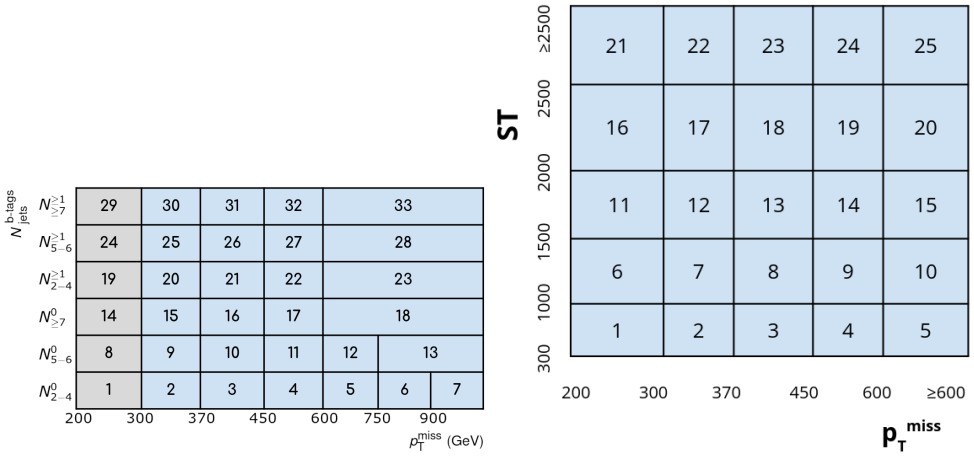


Figure 7.8: Definition of the different bin indexing based on Strong Production (SP) motivated search bins, as derived from (p_T^{miss}, N_{jets} & N_{bjets}) (left), and the newly established search bins within (p_T^{miss}, S_T) (right).

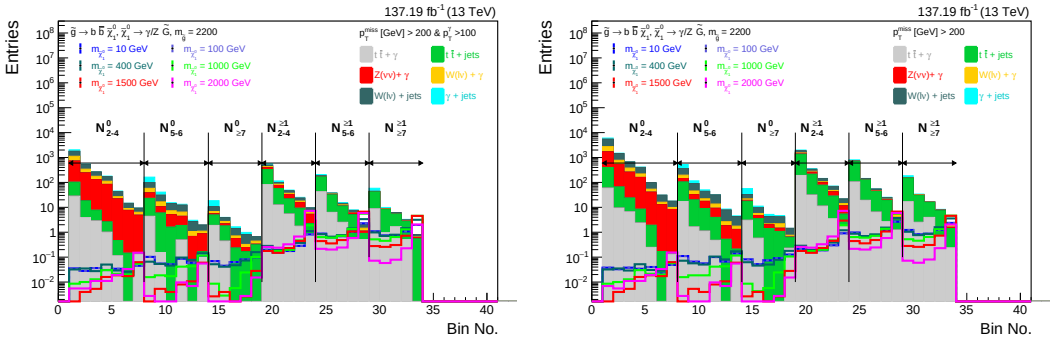


Figure 7.9: The distributions of background and signal events in SP bins with p_T^{γ} >100 GeV (left) and p_T^{γ} >40 GeV (right).

T5bbbbZG signal with p_T^{γ} >100 GeV and p_T^{γ} >40 GeV, respectively in the SR bins of Fig. 7.8 (left). By reducing p_T^{γ}, the signal yields within the search region increase by up to 30%; however, the background increases exponentially, leading to a tenfold increase. The search bins with high p_T^{miss} and high N_{jets} are the bins with the highest sensitivity to the signal bin when p_T^{γ} >100 GeV is used. As we reduce p_T^{γ} to 40 GeV, the same bins are dominated by more background, losing the sensitivity to the signal. This concludes that the binning in N_{jets} is not optimal for the background-dominated phase space.

Based on observations from Fig. 7.6 and 7.7, 25 SR bins are designed based on S_T and p_T^{miss} in an event and their specific boundaries and definition are shown in Fig. 7.8 (right). These 25 SR bins are further divided into 2 more SRs based on N_{bjets} (0, ≥1), making a total of 50 SRs, called SR bins v1 in the rest of this chapter. The events are further divided according to the values of p_T^{γ} (40 to 100 GeV, ≥100 GeV), making a total of 100 SRs, called SR bins v2 in the rest of this chapter. The signal and background compositions as functions of the SRs v1 and v2 are shown in the left and right plots of Fig. 7.10, respectively. In contrast to SP bins, most signal sensitivity is observed in high S_T, p_T^{miss} & N_{bjets} ≥1 bins.

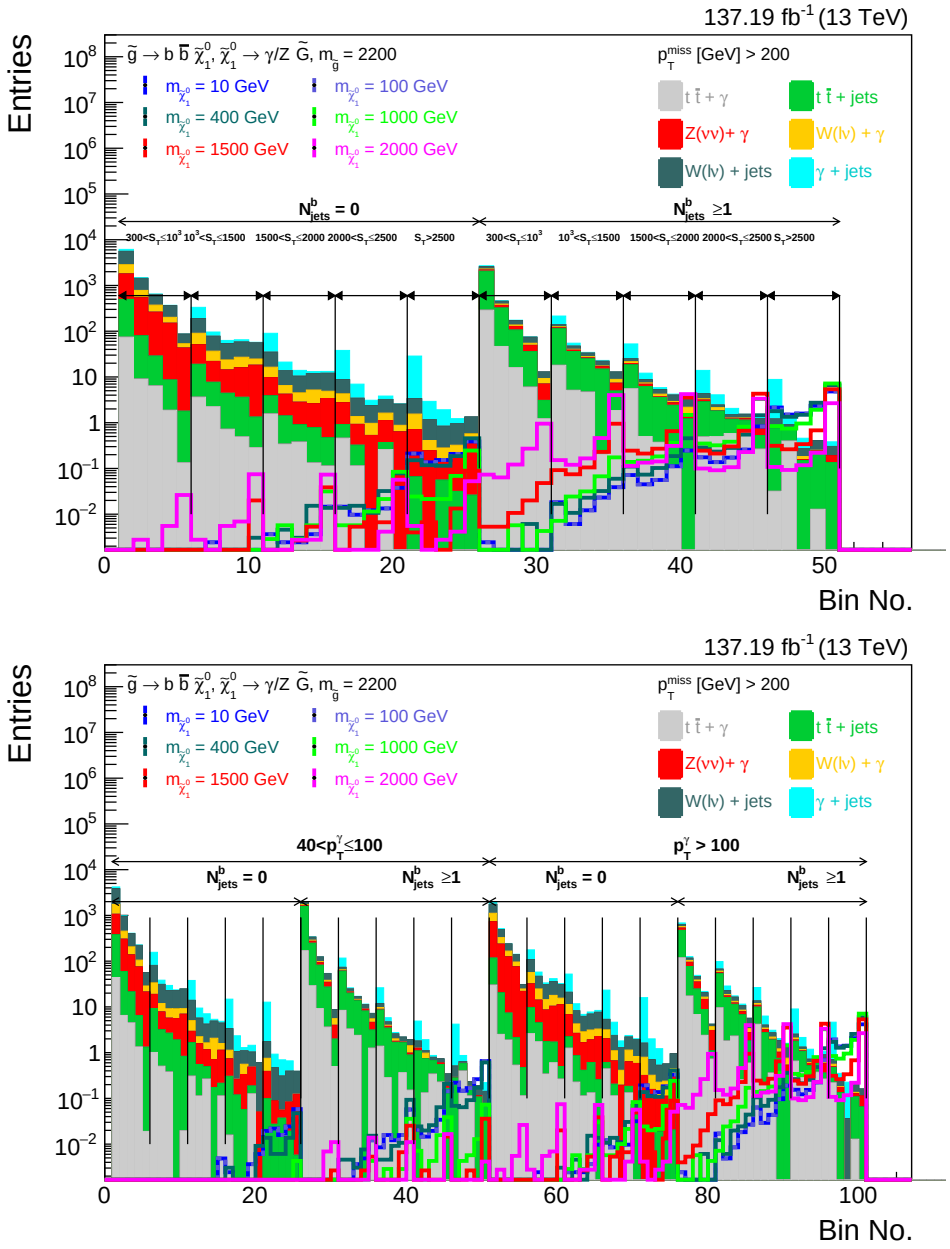


Figure 7.10: The distributions of background and signal events in new SRs based on ($p_{\mathrm{T}}^{\mathrm{miss}}$, S_{T}, N_{bjets}) (top) & ($p_{\mathrm{T}}^{\mathrm{miss}}$, S_{T}, N_{bjets}, & p_{T}^{γ}) (bottom) with p_{T}^{γ} >40 GeV.

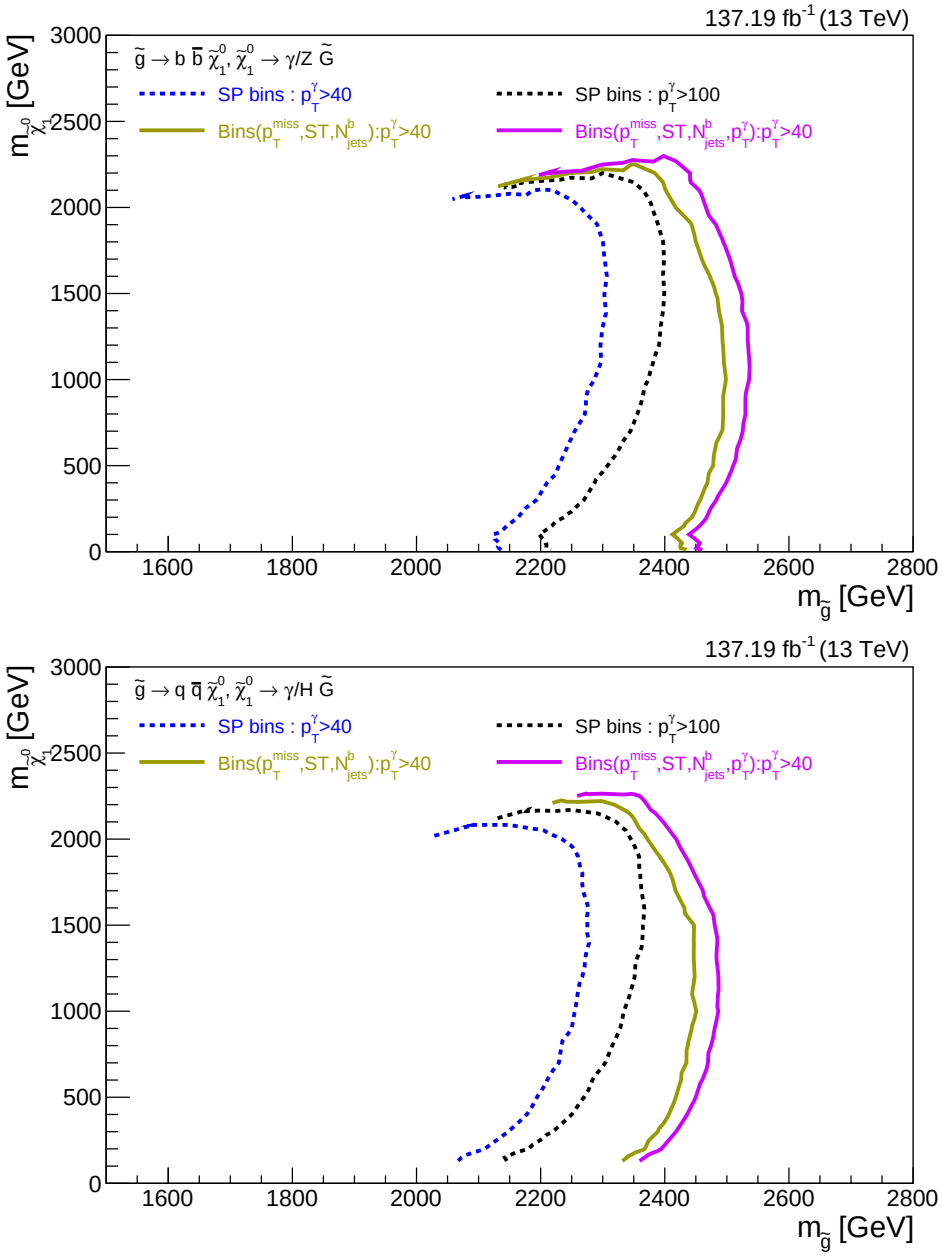


Figure 7.11: The expected exclusion limits at 95% confidence interval (CL) for T5bbbbZG (top) and T5qqqqHG (bottom) models with different SRs. The new SR bins in Magenta line shows improvement in mass exclusion by 200 GeV in both the models.

The expected exclusion with different SRs is calculated using the asymptotic approximation at 95% confidence interval (CL) [164] with 20% systematic uncertainty added to the statistical uncertainty. Fig. 7.11 shows the expected exclusion limits at 95% CL with different SRs for the T5bbbbZG signal model. The black and blue dotted curves represent the expected limits obtained with p_{T}^{γ} >100 GeV and p_{T}^{γ} >40 GeV using SP bins, respectively. Lowering p_{T}^{γ} results in expected limits worse by up to 100 GeV. The expected exclusion limits as determined by the novel 50 SRs are depicted by the yellow curve, whereas those with the 100 SRs are represented by the magenta curve. Up to 200 GeV improvements for small $m_{\tilde{\chi}_1^0}$ are obtained with a new 100 SR based on $p_{\mathrm{T}}^{\mathrm{miss}}$, S_{T}, N_{bjets}, & p_{T}^{γ}.
Fig. 7.11 (right) displays the expected exclusion for the T5qqqqHG model with various SRs. In this scenario, too, the introduction of new 4D SRs similarly enhances the expected limits, specially for large mass difference in $\tilde{g}$ & $\tilde{\chi}_1^0$.
For the T5tttZG model, there is more exclusion (up to 200 GeV) for small $m_{\tilde{\chi}_1^0}$ masses with new SRs, but at high $m_{\tilde{\chi}_1^0}$ masses the exclusion is worse compared to SP bins with high p_{T}^{γ} phase space, as shown in Fig. 7.12 (left). In case of the T6ttZG model, we observed a ∼100 GeV improvement at low $m_{\tilde{\chi}_1^0}$ but a worse exclusion for high $m_{\tilde{\chi}_1^0}$ as shown to the right of Fig. 7.12. This worse exclusion is attributed to the fact that for high $m_{\tilde{\chi}_1^0}$, S_{T} becomes softer and N_{jets} offer better discrimination with respect to background. Detailed examination of this scenario is provided in Section 7.5, where we have enhanced exclusions for T5ttttZG at high $m_{\tilde{\chi}_1^0}$ by employing boosted decision tree (BDT) scores for baseline selection alongside new SRs.

Optimization for pMSSM motivated models

Fig. 7.13 shows the correlation of $p_{\mathrm{T}}^{\mathrm{miss}}$ & p_{T}^{γ} (left most), S_{T} & $p_{\mathrm{T}}^{\mathrm{miss}}$ (middle), and p_{T}^{γ} & S_{T} (right most) for the T5gg model with $m_{\tilde{g}}$ =2.2 TeV, $m_{\tilde{\chi}_2^0}$ =100 GeV and Δm ($m_{\tilde{\chi}_2^0}$, $m_{\tilde{\chi}_1^0}$) =10 GeV. These show low p_{T}^{γ} with high $p_{\mathrm{T}}^{\mathrm{miss}}$ and S_{T} signatures.
The Δm ($m_{\tilde{\chi}_2^0}$, $m_{\tilde{\chi}_1^0}$) derives the kinematics of final state, especially in terms of p_{T}^{γ} and $p_{\mathrm{T}}^{\mathrm{miss}}$. The different composition of the T5gg signal for different $m_{\tilde{\chi}_2^0}$ with Δm =10 GeV (first column) and Δm =100 GeV (second column) with respect to background is shown in Fig. 7.14. The distribution of S_{T} (top row) shifts towards low values as Δm increases, whereas the p_{T}^{γ} distribution (bottom row) is more prominent towards low value events with high Δm.
These T5gg signal points are allocated across the 100 SRs structured according to ($p_{\mathrm{T}}^{\mathrm{miss}}$, S_{T}, N_{bjets}, p_{T}^{γ}), as shown for Δm =10 GeV (left) and 100 GeV (right) in Fig. 7.15. The most sensitive regions are those with high values of $p_{\mathrm{T}}^{\mathrm{miss}}$ and S_{T}, although there is also a noticeable presence of signal events in areas characterized by lower values of S_{T} and $p_{\mathrm{T}}^{\mathrm{miss}}$ for Δm =100 GeV.
For the T5gg model, the expected exclusion limits as a function of $m_{\tilde{\chi}_2^0}$ are shown in Fig. 7.16 for Δm =10, 30, and 100 GeV from right to left, respectively. The region in phase space characterized by a small Δm ($\tilde{\chi}_2^0$, $\tilde{\chi}_1^0$) poses significant challenges, as it results in soft photons. The SP bins with large p_{T}^{γ} do not effectively capture these signals, depicted as the black lines in the middle and right plots. In contrast, the newly introduced SR bins exhibit improved sensitivity to these challenging regions of phase space, imposing tighter limits. Preliminary investigations indicate that with

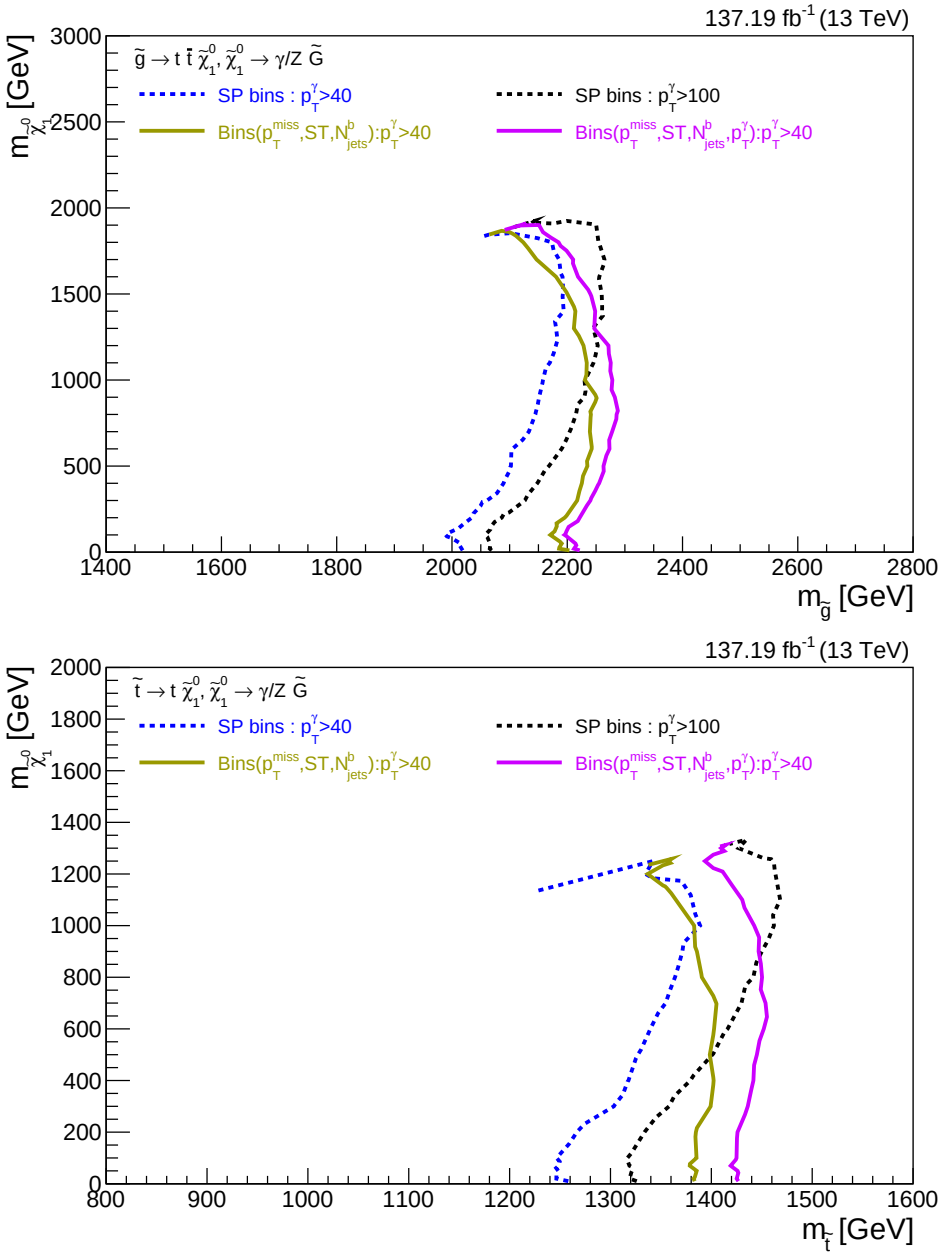


Figure 7.12: The expected exclusion limits at 95% confidence interval (CL) for T5ttttZG (top) and T6ttZG (bottom) models with different SRs. The new SR bins in Magenta line shows improvement in mass exclusion in small $\tilde{\chi}_1^0$ masses.

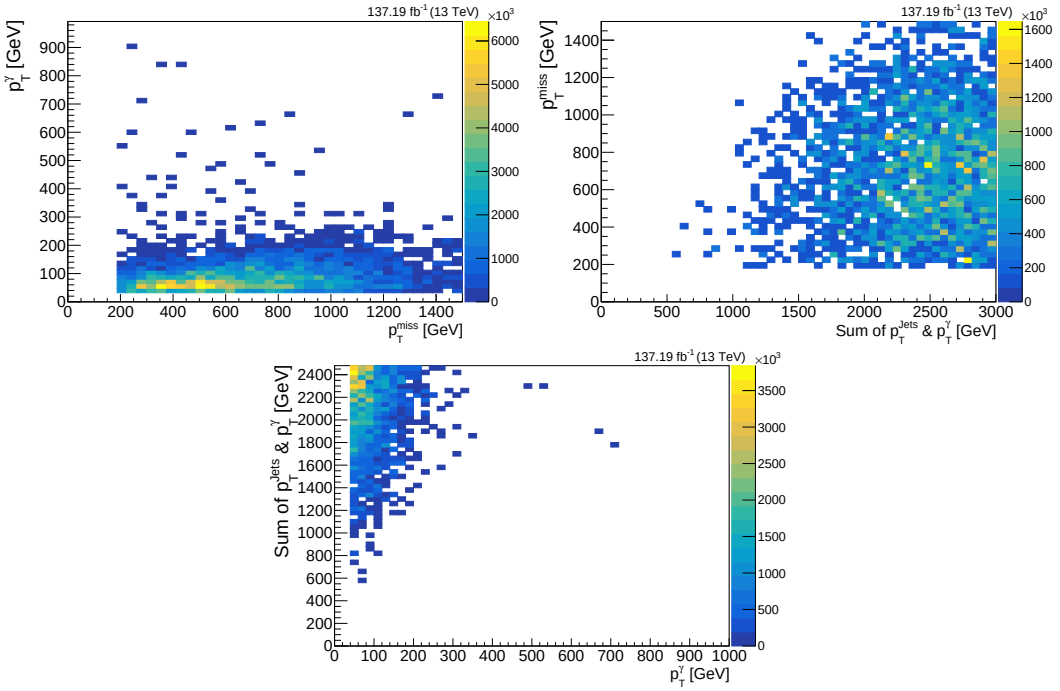


Figure 7.13: The distributions of $p_{\mathrm{T}}^{\mathrm{miss}}$ versus p_{T}^{γ} (top), S_{T} versus $p_{\mathrm{T}}^{\mathrm{miss}}$ (middle), and S_{T} versus p_{T}^{γ} (bottom), with $m_{\tilde{\chi}_2^0} = 100\,\mathrm{GeV}$ and $\Delta m = 10\,\mathrm{GeV}$ in T5gg model.

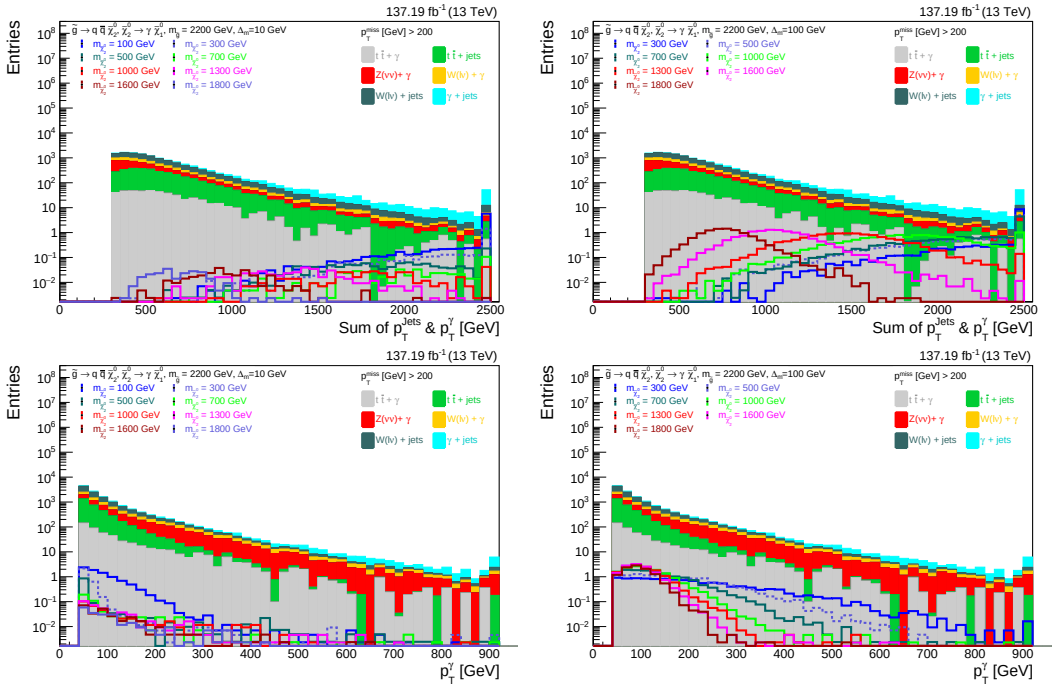


Figure 7.14: The top and bottom rows illustrate the distribution of S_{T} and p_{T}^{γ}, respectively, for SM backgrounds and a selection of T5gg motivated $m_{\tilde{\chi}_2^0}$ signal points, with Δm set to 10 & 100 GeV, respectively.

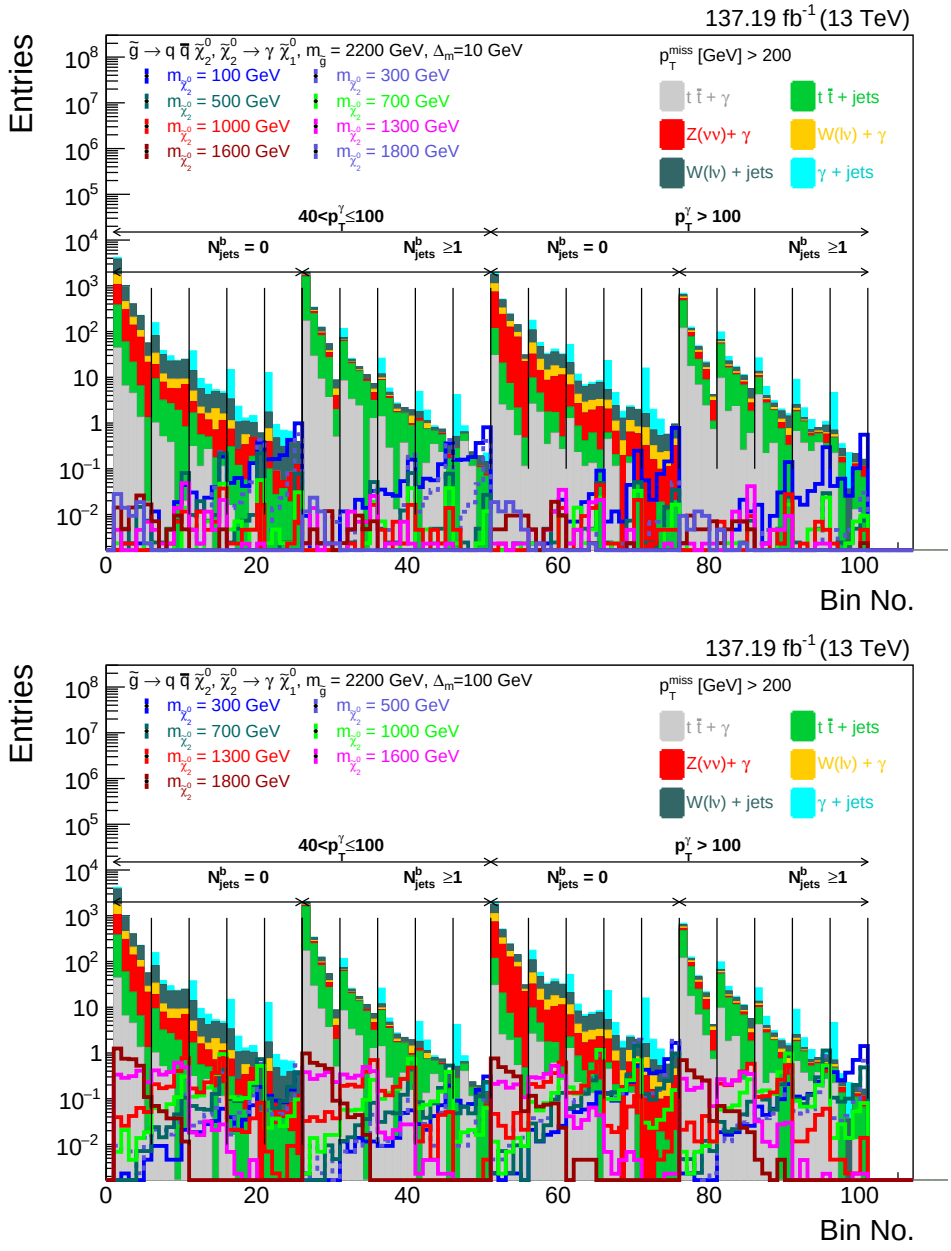


Figure 7.15: The distributions of background and a selection of T5gg motivated $m_{\tilde{\chi}_2^0}$ signal points, with Δm set to 10 (top) & 100 GeV (bottom), respectively is shown in new SRs based on (p_T^{miss}, S_T, N_{bjets} & p_T^γ) with p_T^γ >40 GeV.

the updated SRs, achieving sensitivity to such highly compressed phase spaces is feasible.

7.3 Background estimations

As explained in Section 7.1.1, multiple processes such as $t\bar{t}$+jets, W+jets, $t\bar{t}\gamma$+jets, $W\gamma$ + jets, $Z(\nu\bar{\nu}) + \gamma$, and QCD multijet +γ+jets can serve as background in the signal region. Since the recorded data set is a combination of SM processes and potential new physics process, it is not feasible to distinguish them directly. Although the selections of the baseline events discussed in 7.2.1 reduce the contributions of these background sources in the SR, even then these background sources have a significant contribution to the SR. These should be accurately measured to infer presence or absence of signal in the data. The data driven approach is preferred over taking these backgrounds from MC simulation directly as these are not perfect and are susceptible to mis-modeling, limited accuracy of cross sections, which may not accurately reflect data in terms of normalization, as well as shape.

To measure the backgrounds, we design a phase space that is orthogonal to, yet very similar to, the SR and devoid of any signal, referred to as control region (CR). As CR is used to provide corrections or translate to SR, CR is kinematically chosen to be very similar to phase space with SR to avoid large contributions of systematic uncertainties. We can then translate the results of the CR event from the data to the SR in the data with the help of the CR $\rightarrow$ SR transfer factors calculated using MC where MC is corrected for the differences between MC and the data. The validation of these methods is performed by comparing expected and predicted yields in the same MC simulated event sample.

Based on which of the detector inefficiencies are contributing to raising certain backgrounds and to simplify their measurements, we have divided reducible backgrounds into the following three categories.

- Lost lepton (LL) background
- Fake photon from electron background
- Fake MET background

That leaves the other irreducible background, invisible background ($Z(\nu\bar{\nu}) + \gamma$ process). The rest of the sections discuss the methods, their validation, and respective predictions in detail.

7.3.1 Estimation of lost lepton background

Our SR consists of at least a good photon (p_{T}^{γ} >40 GeV), multijets, and $p_{\mathrm{T}}^{\mathrm{miss}}$ without leptons in the final state. There are SM processes with almost similar final states except that leptons are present. We measured the contribution to the signal region from events in which leptons were present at the truth level but were not reconstructed. The $t\bar{t}\gamma$+jets and $W\gamma$ + jets are the main contributing processes.

If e or μ from the decay of the W boson in the $t\bar{t}\gamma$+jets ($W\gamma$ + jets) process is not identified because it was outside the detector acceptance range, or it was not isolated or identified according to the reconstructed ID requirement, then the event is indistinguishable from the signal and cannot be mitigated with e (μ) veto. On the other hand, τ from W bosons decays to leptons $\sim$35% and the rest to hadrons. In the initial scenario, the e & μ veto is effective provided that the leptons are accurately identified. For hadronic decays, isolated track vetoes are implemented to diminish these occurrences. However, residual events contribute to the signal region. So we divide the lost lepton background's contribution to SR into three mutually exclusive categories; lost electron, lost μ and τ-hadronic (targeting its hadronic decay).

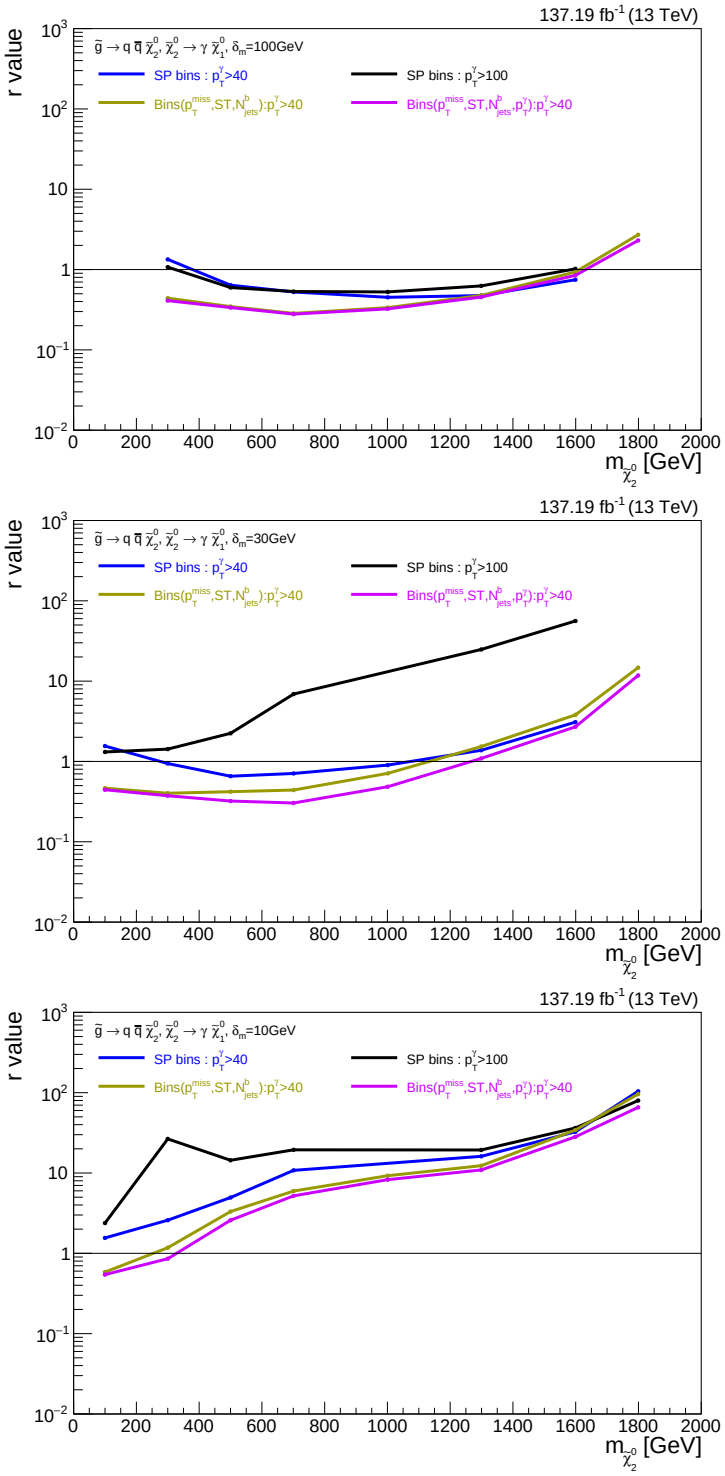


Figure 7.16: The expected exclusion limits at 95% confidence interval (CL) for different $m_{\tilde{\chi}_2^0}$ for T5gg with Δm = 10 GeV (top), 30 GeV (middle) and 100 GeV (bottom) with different SRs. The new SR bins in Magenta line shows improvement in mass exclusion.

When selecting the lost-electron SR, there should be at least one electron and no muon at the generator level. The dileptonic decay of $t\bar{t}$ resulting in eμ is excluded from consideration, while the scenario in which eτ is produced from τ decaying hadronically is taken into account in the lost electron SR. Similarly, muon SR requires at least one μ at the generator level, and in the case of τ hadronic, no muon or electron should be present at the generator level. To avoid overlap between the fake photon from the electron background in this SR, ΔR (reco γ, gen e)$>$ 0.2 and the p_{T} ratio of the photon and the generator electron ($p_{\mathrm{T}}^{\gamma}/p_{\mathrm{T}}^{gen-e}$) should be $>$ 1.2 or $<$ 0.8. In addition to the above criteria, three SRs should satisfy all the baseline selections discussed in Sect. 7.2.1.

Two mutually exclusive CRs are defined by changing the lepton requirements in SR; (1 e, $\geq$ 1 γ) CR to measure the contribution due to the lost electron, (1μ, $\geq$ 1 γ) CR to measure the contribution due to the lost μ and τ-hadronic. The e (μ) in CR should satisfy the requirements described for good e (μ) in Section 7.2.1. The e (μ) should be within $|\eta| < 2.5$ with respective $p_{\mathrm{T}} >$10 GeV and should satisfy POG recommended VETO (medium) ID along with additional minimum isolation requirement. To ensure that reconstructed photon is prompt and not a bremsstrahlung photon, ΔR (reco e (μ), reco γ)$>$ 0.2). To avoid the contribution of the signal in CR, an additional condition on transverse mass (M_{T}) of (M_{T} (e (μ),$p_{\mathrm{T}}^{\mathrm{miss}}$)$<$ 100 GeV) is applied for events in CR. In single e or μ CR, μ or e are vetoed, respectively. Events with the isolated pion tracks are also vetoed along with isolated μ (or e) tracks in e (or μ) CRs.

The contribution from different processes is combined after weighing the events with respective cross section and full Run 2 luminosity. The average transfer factor (TF) method is used to measure this background. The "Average" TF represents the mean ratio of lost (e,μ, τ-hadronic) to identified (e,μ) lepton events, independent of lepton kinematics. The TF for an i^{th} bin (TF_i^{MC})can be defined as per the following equations using the truth-level information in MC to tag an event as lepton found or lost based on the generator level information,

$$TF_i^{MC}(H_{\mathrm{T}}, p_{\mathrm{T}}^{\gamma}, N_{\mathrm{bjets}}) = \frac{N_{0l+1\gamma}^{MC}(i)}{N_{1l+1\gamma}^{MC}(i)} \tag{7.3}$$

where $N_{0l+1\gamma}^{MC}$ (i) is the total event yields in (lost e + lost μ + τ*hadronic* SR and $N_{1l+1\gamma}^{MC}$ (i) in single lepton CR (combined muon and electron) for the ith bin. Contributions are added in MC (majorly from t$\bar{\mathrm{t}}\gamma$+jets & Wγ + jets and secondary from t$\bar{\mathrm{t}}$+jets & W+jets). The predictions for data in j^{th} SR bin can be obtained by translating CR events to SR using TF obtained from equation 7.3,

$$N_{0l+1\gamma}^{Data}(j) = TF_i^{MC}(H_{\mathrm{T}}, p_{\mathrm{T}}^{\gamma}, N_{\mathrm{bjets}}) \times N_{1l+1\gamma}^{Data}(j) \tag{7.4}$$

where $N_{0l+1\gamma}^{Data}$ (j) are the predicted event yields for lost lepton background in the data and $N_{1l+1\gamma}^{Data}$ (j) are the event yield in single lepton CR for j^{th} SR bin. We are using MET trigger dataset to obtain single lepton CR in data.

We define H_{T} as the scalar sum of all jets p_{T}. Fig. 7.17 shows the distribution of H_{T} (left column), p_{T}^{γ} (right column) in CR (red curve) and SR (blue curve) for the lost electron case (top row) and lost muon + τ*hadronic* using MC with complete Run 2 luminosity with the bottom panel of the plots representing the ratio of SR events

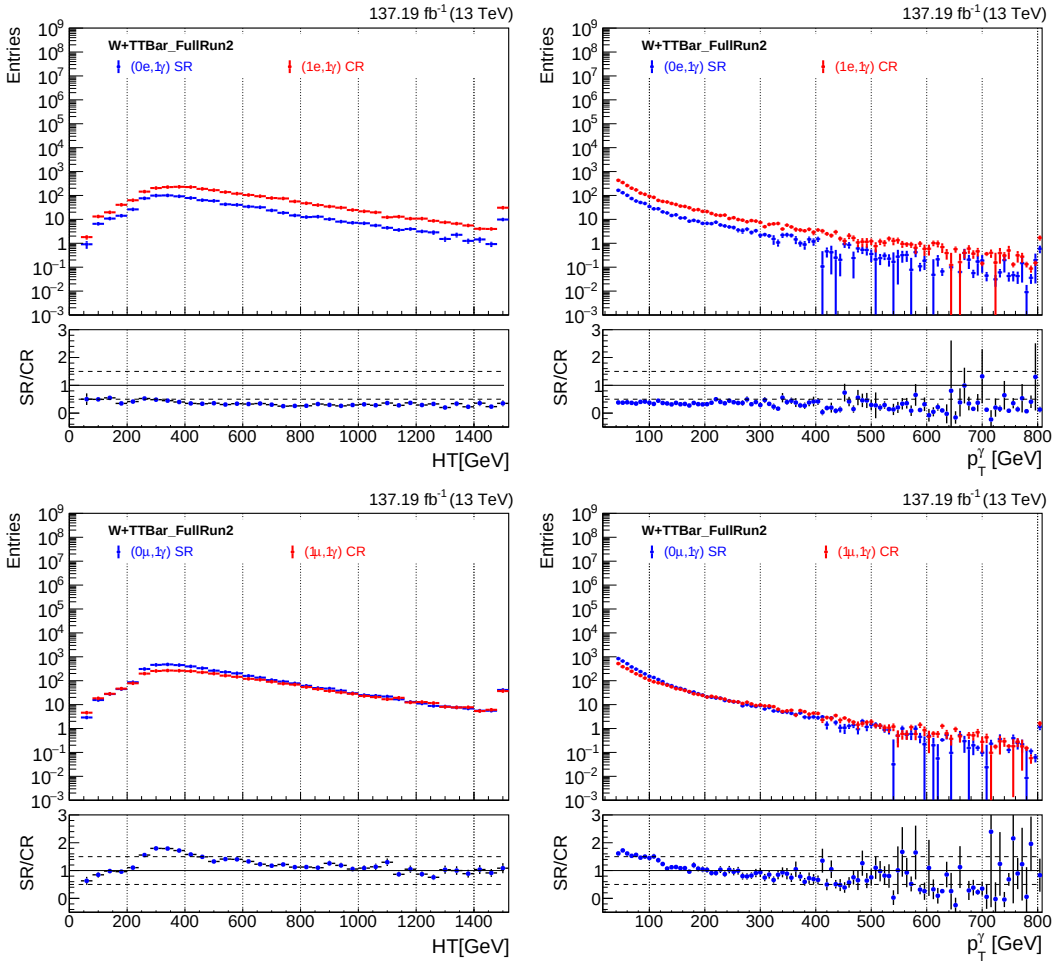


Figure 7.17: Distributions of H_{T} (left column), p_{T}^{γ} (right column)are shown in CR (red curve) and SR (blue curve) for lost electron case (top row) and lost muon + $\tau - hadronic$ using MC with complete Run 2 luminosity. The bottom panel of the plots represents the ratio of SR events yields to CR in different bins.

to CR in different bins. MC here represents combined contribution from $t\bar{t}\gamma$+jets & $W\gamma$ + jets. Some interesting trends are observed in the bottom panel of the plots for the variables H_T and p_T^γ in the case of lost muon + τ hadronic SR in contrast to lost electron where no such trend is present. The upward trend in the ratio of SR to CR for $p_T^\gamma < 100\,\text{GeV}$ is attributed to τ *hadronic* SR as more non-prompt photon contributes in this region. For the scenario involving H_T, the ratio of SR to CR is markedly elevated when H_T is within the range of 200 to 400. This phenomenon is attributed to the contribution of $\tau - hadronic$ SR, which is reconstructed as jets, thus contributing to H_T through its p_T. We also compared the general shape in SR and CR in terms of other distributions such as p_T^{miss}, N_{jets}, & N_{bjets}, it is found to be similar.

Correction of MC to match data and their comparisons in CR

We are using the MET trigger data set in this analysis and corresponding trigger efficiencies are calculated and applied on MC. MC simulations are not perfect, and there are discrepancies between observed data and MC due to mismodeling or limited accuracy of the calculations. Hence, data-MC correction factors are employed to rectify such discrepancies between experimental observations and simulated datasets, thus allowing the simulated events to be as representative as possible of the observed data. On the other hand, observed data are more prone to detector-level noises, including in-functionality of certain detector parts or its readout components. The majority of such data cleaning have been discussed in length in Section 7.2.1 including noise filters to avoid spurious p_T^{miss}, HEM-veto issue in 2018, and L1 trigger pre-firing in 2016 and 2017.
Following are the list of corrections are taken into account before calculating TF,

- Efficiencies of p_T^{miss} triggers parameterized as a function of p_T^{miss} and applied on MC.
- Weights are applied to MC events to account for event pile-up based on observed data.
- Jet energy corrections are applied to MC events to account for discrepancies in the energy scale between the experimental data and MC simulations.
- B-tagging scale factors (SF) are applied to MC simulations to account for discrepancies in b-tagging efficiencies between data and MC simulations. The events are subsequently reweighed according to the number of jets tagged with b, the scale factors and efficiencies, and are parameterized as a function of their p_T and η.
- The lepton efficiency scale factor, intended to align discrepancies in efficiencies between data and MC simulations.

To ensure the robustness of this background estimation, Fig. 7.18 shows the comparison of single lepton CR (1 electron + 1 muon) in MC and data for complete Run 2 luminosity as a function of search variables (p_T^{miss} (top left), S_T (top right), p_T^γ (bottom left) and N_{jets} (bottom right)). The bottom panel of the plots represents the data with respect to the MC ratio. The agreement between the data and the MC is mostly 10 to 50%. Exact match between data & MC is not crucial for us. This method, being data-driven, does not rely on simulations completely.

Transfer factor calculation

Learning from the observation of Fig. 7.17, the TFs for the lost lepton backgrounds are parameterized as a function of H_T, p_T^γ and N_{bjets}. A total of 18 mutually exclusive

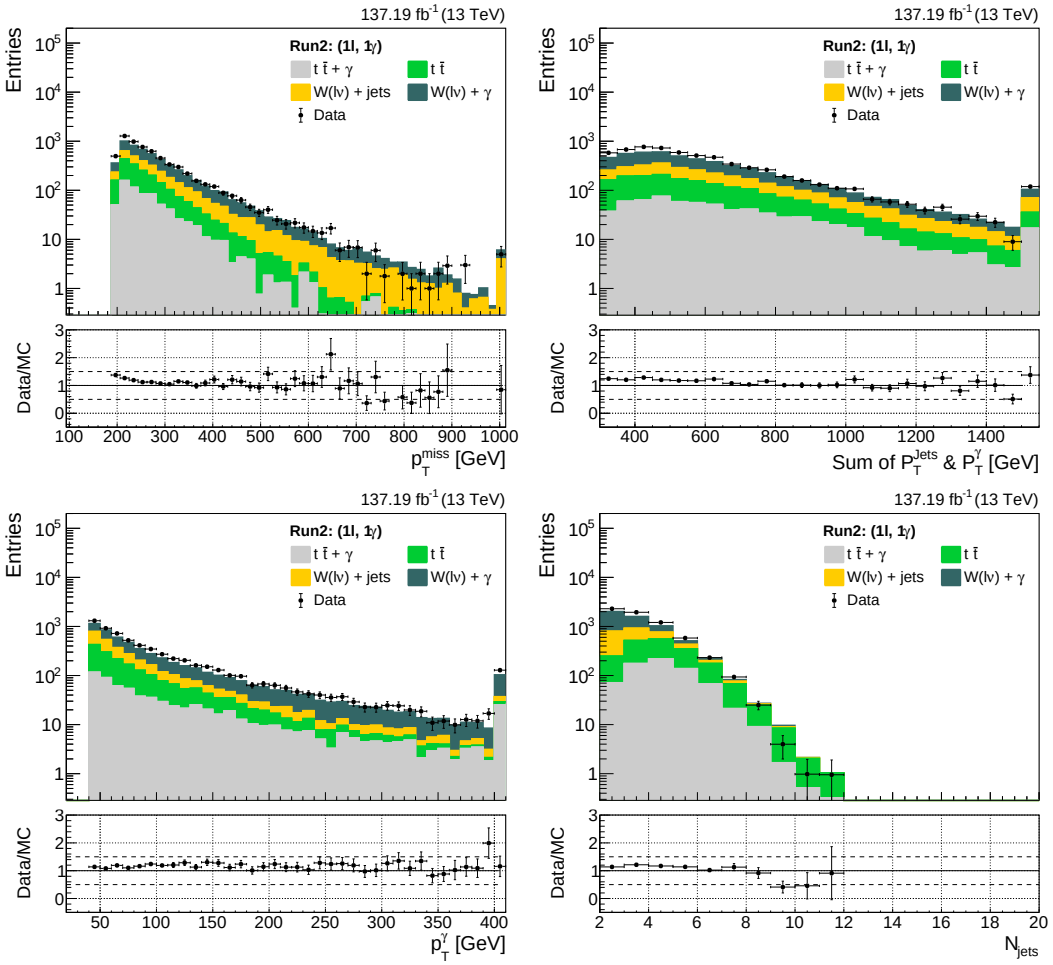


Figure 7.18: The above figures illustrate a comparison between MC and experimental recorded data for the single lepton control region, which includes one electron and one muon, across the entire Run 2 luminosity. This comparison is shown as a function of various search variables: $p_{\mathrm{T}}^{\mathrm{miss}}$ (top left), S_{T} (top right), p_{T}^{γ} (bottom left), and N_{jets} (bottom right). The panel displaying ratios provides the relationship between observed data and Monte Carlo predictions.

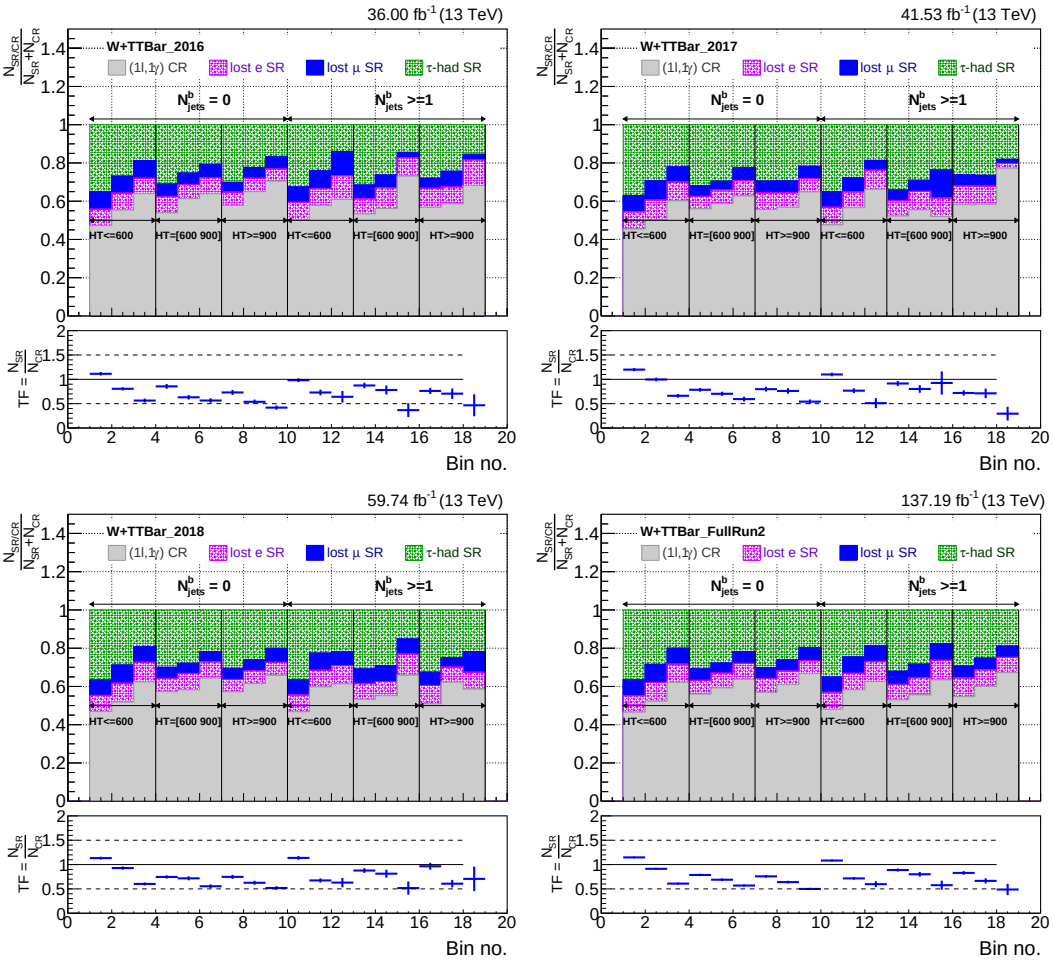


Figure 7.19: The fractional distributions of lost electrons, lost muons, and *tauhadronic* SRs along with the total CR, which combines single electron and muon CR, are depicted as a function of different TF bins for datasets from the years 2016 (top left), 2017 (top right), 2018 (bottom left), and Run 2 (bottom right). The bottom panel of the plots illustrates the transfer function derived for each respective bin.

TF bins are devised with 3 bins in H_T ($H_T \leq 600\,\text{GeV}$, $600\,\text{GeV} < H_T \leq 900\,\text{GeV}$, $H_T \geq 900\,\text{GeV}$), 3 bins in p_T^γ ($p_T^\gamma \leq 100\,\text{GeV}$, $100\,\text{GeV} < p_T^\gamma \leq 250\,\text{GeV}$, $p_T^\gamma \geq 250\,\text{GeV}$) and 2 bins in N_{bjets} ($N_{\text{bjets}} = 0$, $N_{\text{bjets}} \geq 1$).

Fig. 7.19 shows the distributions of individual fractions of lost electrons, lost muons, and τ hadronic SRs and total CR (combined single electron and muon CR) as a function of different TF bins for data sets from the year 2016 (top left), 2017 (top right), 2018 (bottom left) and Run 2 (bottom right). The contribution of *tau* hadronic is the highest among the three SRs and the maximum for the $H_T \leq 600$ and the lowest bins in p_T^γ and gradually decreases with increasing p_T^γ and H_T. In contrast, the contribution of lost electron and muon to each bin is almost constant regardless of the values of H_T and p_T^γ except very high H_T and non-zero N_{bjets}.

The bottom panel of the plots in Fig. 7.19 indicates the TF obtained for each bin. TFs are the highest for $H_T \leq 600$ and the lowest bins in p_T^γ irrespective of N_{bjets}. A gradual decrease in TF is observed as a function of p_T^γ for a given H_T bin. The overall TF values drop with high H_T as expected in Fig. 7.17. We have used the combined TF for Run 2 luminosity to evaluate the predictions in the data according to the equation 7.4.

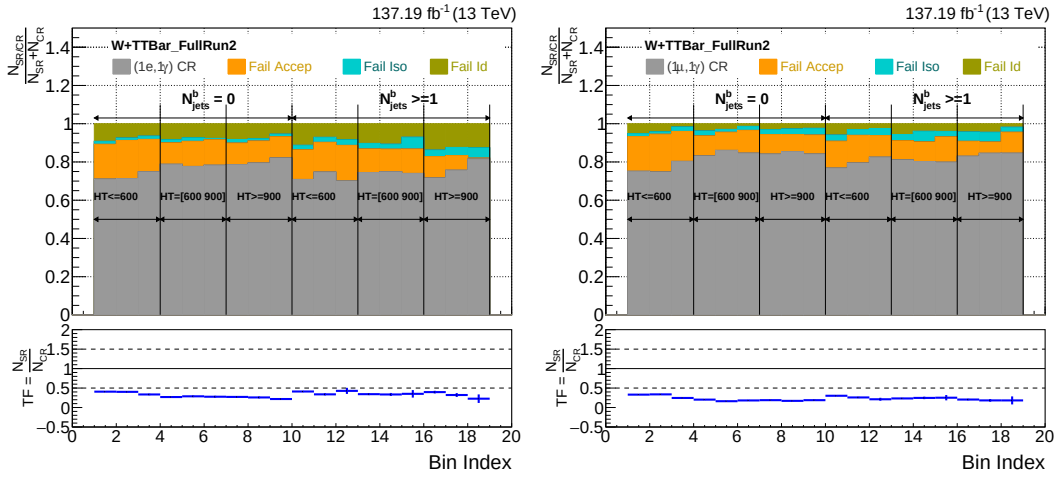


Figure 7.20: Distribution of individual normalized fraction of the contributions of these three categories in the lost electron SR (left) and the lost muon SR (right) as a function of different TF bins.

The lost electron and muon SRs are further categorized on the basis of failed criteria that introduce these events to the SRs. These are failed acceptance criteria, failed isolation, and failed identification criteria. Fig. 7.20 shows the individual normalized fraction of the contributions of these three categories in the lost electron SR (left) and the lost muon SR (right) as a function of different TF bins. Most of the lost electrons or lost muon SRs arise due to failure of the detector acceptance criteria (up to 20%) and for these events we completely rely on MC simulations, as there is no information about such events in the data.

Closure test

The validation of the average TF method is done using the complete Run 2 MC event sample. For this, we use equation 7.4 and get the predicted SR yields in MC with CR as well from MC. These yields are then compared with the number in the box that we obtained with MC. For validation, we used 100 search region bins designed using $p_{\mathrm{T}}^{\mathrm{miss}}$, S_{T},p_{T}^{γ}, and N_{bjets} as discussed in Section 7.2.
Fig. 7.21 shows the comparison of expected SR directly taken from MC (blue points) and predicted SR (red points) in MC using the TF shown on the bottom right of Fig. 7.19. The bottom panel of the plots represents the expected and predicted SR ratio. Each vertical block in this plot shows the $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution in a S_{T} block (with a fixed value S_{T}). The overall closure looks good as expected and the predicted values are in great agreement up to 10-20% and within the statistical uncertainties except for very high S_{T} and $p_{\mathrm{T}}^{\mathrm{miss}}$ bins.
Fig. 7.22 shows the comparisons of expected and predicted event yields in $p_{\mathrm{T}}^{\mathrm{miss}}$ (top right), S_{T} (top left), N_{bjets} (bottom left) and p_{T}^{γ} (bottom right) in MC with the complete Run 2 dataset. The bottom panel of the plots represents the expected to predicted ratio. The closure test as a function of different variables is also good, showing no biases towards any particular variable.

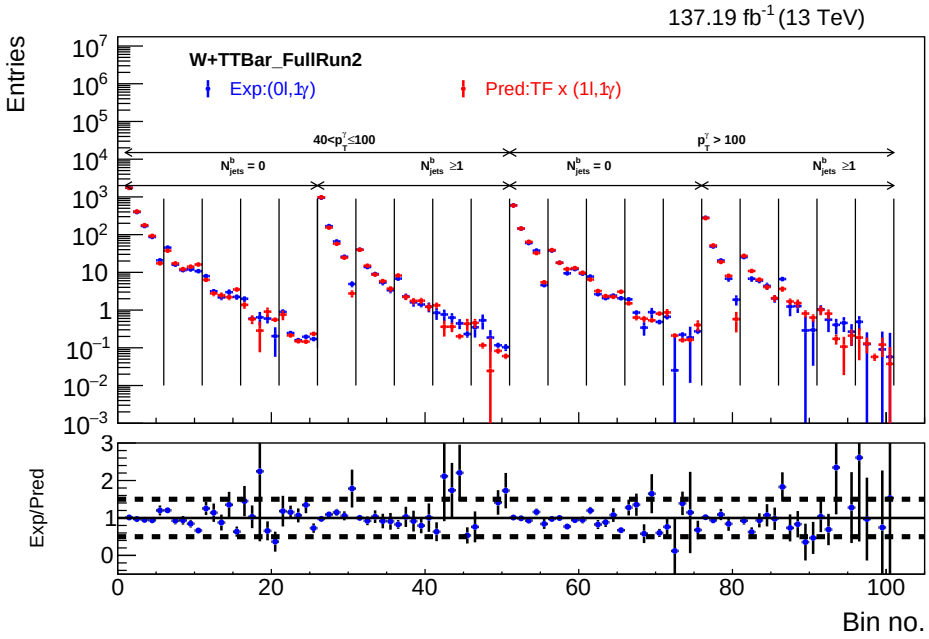


Figure 7.21: The comparison between the anticipated SR derived from MC (blue markers) and the SR predicted using equation 7.4 (red markers) is illustrated. The bottom panel of the plots delineates the ratio between the expected and predicted signal regions.

7.3.2 Predictions of lost lepton background in data & sources of uncertainties

Systematic uncertainties, distinct from statistical uncertainties, arise from experimental conditions, theoretical models, or assumptions made during analysis. These uncertainties can impact either the normalization or the shape of distributions and are broadly categorized into theoretical and experimental systematics.

The primary source of uncertainty in LL predictions is the statistical uncertainty of the data CR. When estimating the lost-lepton background, multiple uncertainty sources impacting the TFs are considered. Uncertainties from various systematic sources are measured and discussed as follows:

- The dependence on the choice of parton distribution functions (PDFs) [165, 166] is evaluated by varying event weights across 100 PDF replicas, resulting in a maximum of 3% variation in TF as shown in Fig 7.23 (left). This maximum value is taken as systematics for predictions in data.
- Renormalization (μ_R) and factorization (μ_F) scale uncertainties [167, 168] are assessed by independently varying the scales by factors of 0.5 and 2. This leads to a maximum of 10% variation in TF as shown in Fig. 7.23 (right). This maximum value is taken as systematics for predictions in data.
- To address possible mis-modeling of collinear photon radiation from quarks and leptons [169], cross-sections of relevant simulated samples are varied by 20%. The maximum impact on TFs is found to be 4% as shown in the bottom panel of Fig. 7.24 and corresponding maximum effect on predictions is $\sim$3%.
- Impact of Jet energy corrections (JEC) and jet energy resolution (JER) uncertainties, derived from data, is also assessed by varying these by $+\ 1\sigma$ and shown in left and right plots of Fig. 7.25, respectively. For varying these uncertainties by $-\ 1\ \sigma$, the similar comparison is added in Fig. B.1 of Appendix B.1. The variation of 0-10% (0-3%) & 0-8% (0-5%) in TFs is obtained with varying JEC

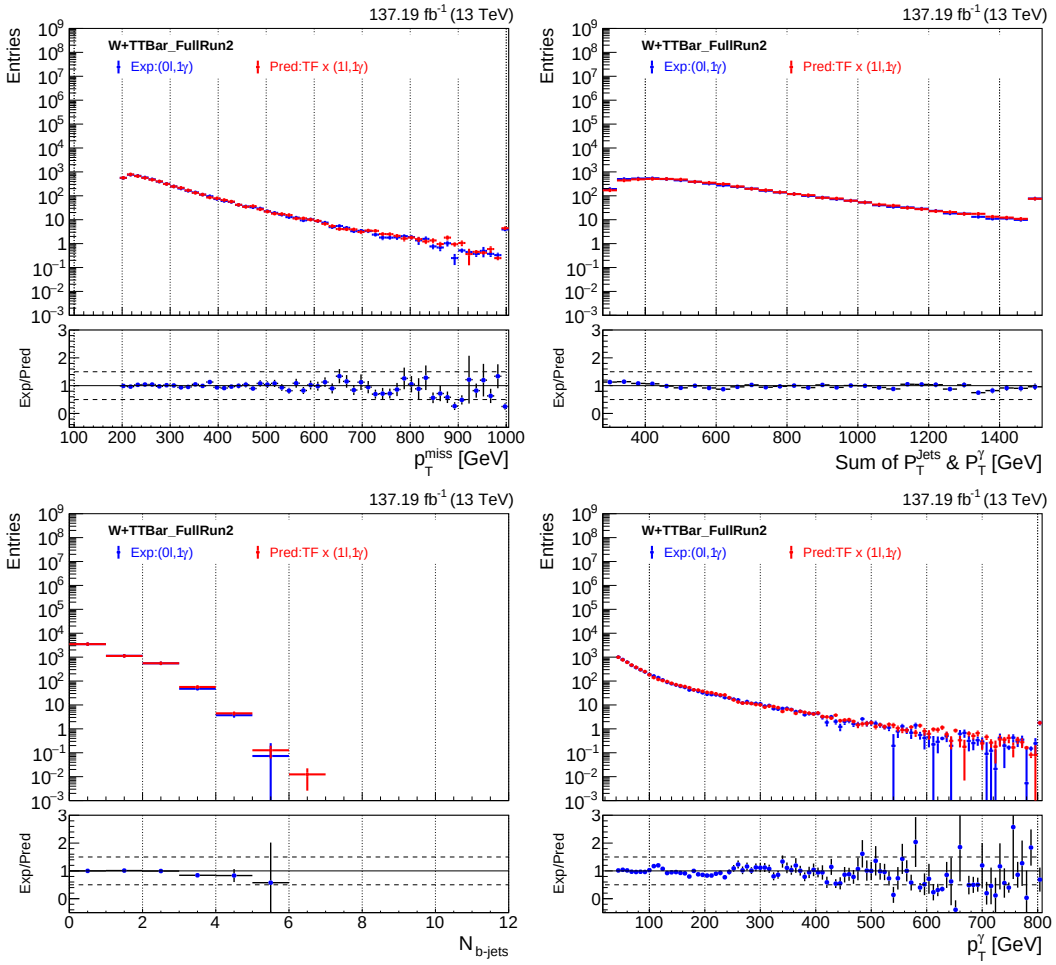


Figure 7.22: Expected and predicted event yields are compared in terms of $p_{\mathrm{T}}^{\mathrm{miss}}$ (top right), S_{T} (top left), N_{bjets} (bottom left), and p_{T}^{γ} (bottom right) in MC using the full Run 2 dataset are compared. The bottom panel of the plots shows the expected-to-predicted ratio.

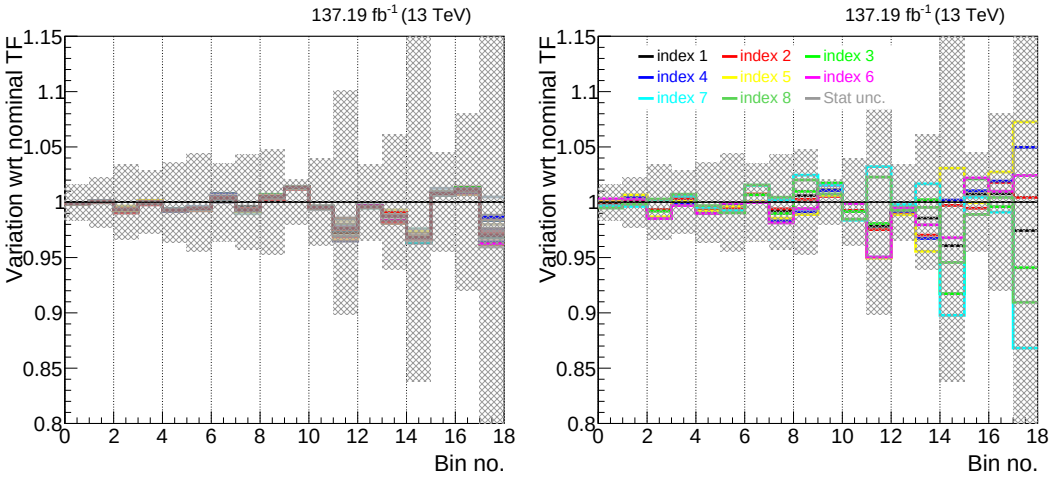


Figure 7.23: Variation in TF (Y-axis) with 100 different PDFs (left) and 8 scale (right) uncertainties corresponding to each TF bin (X-axis). The shaded area indicates the statistical uncertainty.

and JER uncertainties by +1σ(-1σ), respectively. These variations of TFs with respect to JEC/JER, introduce a maximum of 7% uncertainty in the predictions.

- Uncertainties in scale factor for b-tagging efficiencies are evaluated by varying the SF by $+\ 1\sigma$ and shown in Fig. 7.26 (left). For varying this uncertainties by $-\ 1\ \sigma$, the similar comparison is added in Fig. B.2 (left) of Appendix B.1. A maximum of 4% (5%) variation in TF is observed and corresponding maximum variation in predictions is around 1.5%(1%) is observed when uncertainty on the scale factors are varied by $+1\sigma$ (-1σ).
- The bottom panel in Fig. 7.26 (right) and Fig. B.2 (left) shows the TFs with $+\ 1\sigma$ and $-\ 1\sigma$ variation in pileup weights, resulting into a total of 3% and 5% variation in TFs, introducing $\sim$3% and 1.5% uncertainty on the predictions, respectively.

Corrections applied to simulations are varied to propagate uncertainties. The resulting variations in predicted event yields in each signal region are taken as systematic uncertainties for the respective corrections.

Fig 7.27 shows distributions of the predicted electron faking photon background in data using equation 7.7 in 100 SRs devised based on $p_{\mathrm{T}}^{\mathrm{miss}}$, S_{T}, N_{bjets} & p_{T}^{γ}. The black points represents the statistical uncertainty. The gray shaded region denotes the aggregate of systematic uncertainties that were quantified and previously addressed in this section.

Within the scope of thesis, individual evaluation of these systematics is not done and we chose to incorporate a 20% systematic uncertainty in addition to statistical uncertainties while evaluating projected exclusion limits discussed in Section 7.4.

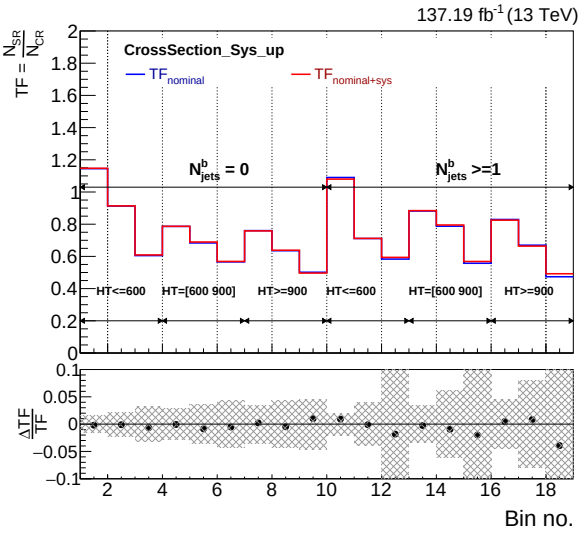


Figure 7.24: The top panel compare the nominal TF distribution (red) with the TF obtained with a 20% increment in the $t\bar{t}\gamma$+jets & $W\gamma$ + jets cross section (blue). The shaded area indicates the statistical uncertainty in the bottom panel, while the black points illustrate the relativistic differences between both TFs defined as the ratio of $(TF_{nominal} - TF_{nominal+sys})$ to $TF_{nominal}$.

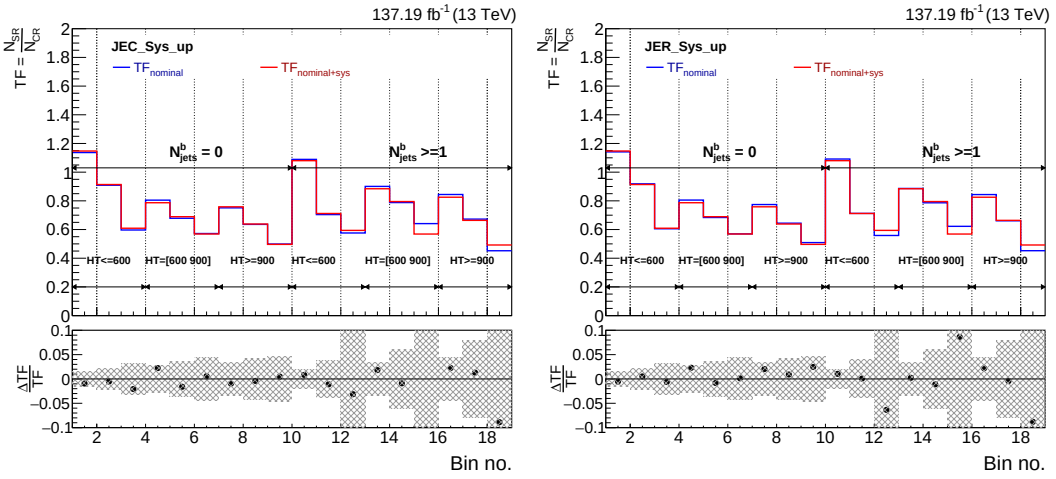


Figure 7.25: Top panel: the nominal TF distribution (blue) with the TF (red) obtained with adjusting JEC factors (left) and JER (right) by $+1\ \sigma$ uncertainty. Bottom panel: shaded region represents statistical uncertainty, with black points showing TFs' relative discrepancies defined as the ratio of $(TF_{nominal} - TF_{nominal+sys})$ to $TF_{nominal}$.

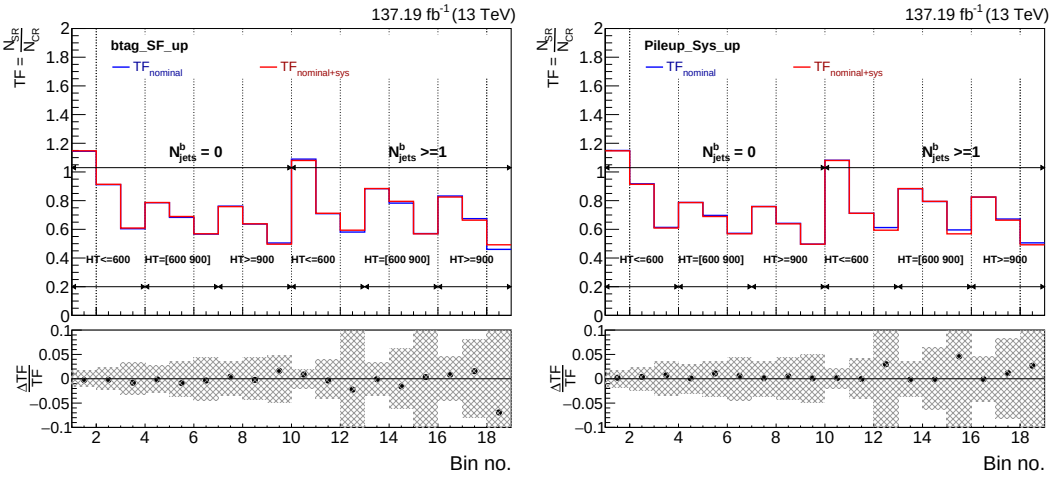


Figure 7.26: Top panel: the nominal TF distribution (blue) with the TF (red) obtained with adjusting btag SF (left) and pileup weights (right) by $+1\ \sigma$ uncertainty. Bottom panel: shaded region represents statistical uncertainty, with black points showing TFs' relative discrepancies defined as the ratio of $(TF_{nominal} - TF_{nominal+sys})$ to $TF_{nominal}$.

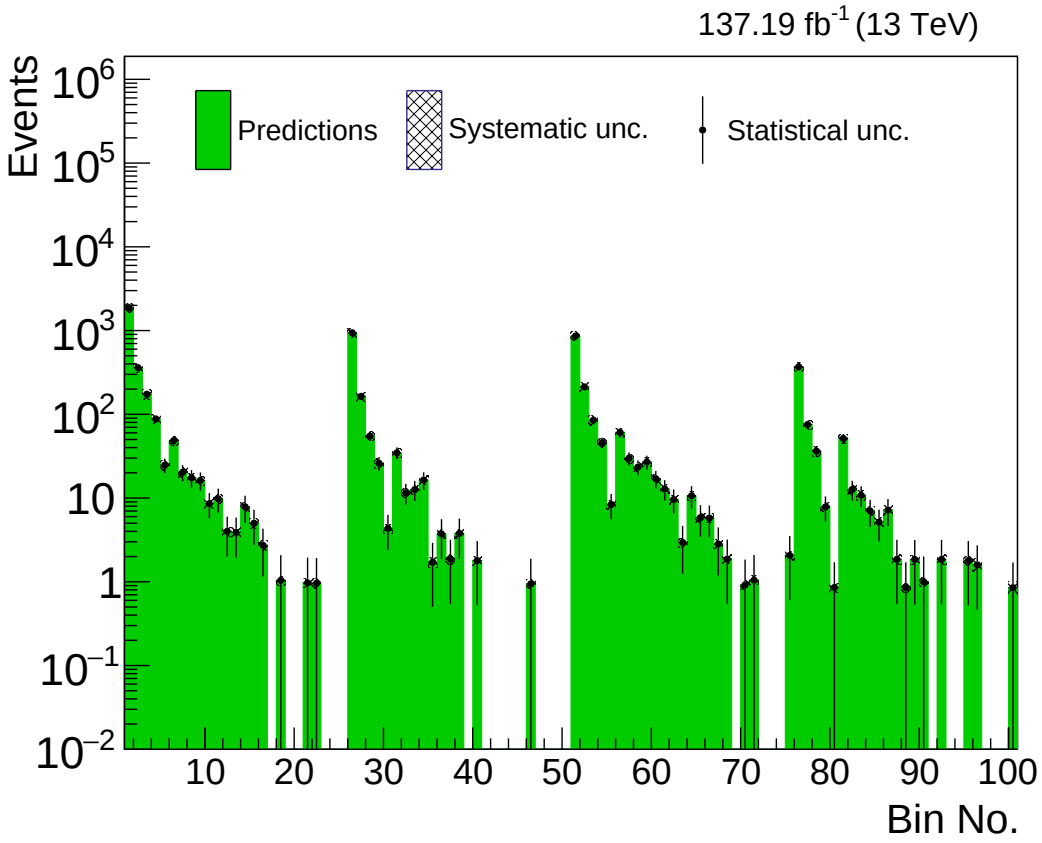


Figure 7.27: The predicted background events in the SRs for Lost lepton background. The black points indicate statistical uncertainty and gray shaded region represents systematic uncertainty.

7.3.3 Estimation e-faking-photon background

In this category of background, we estimate the contribution from the events which do not have a prompt photon at generator level, instead there is an electron which is misreconstructed as photon. The dominating contribution are from $t\bar{t}$+jets & W+jets SM processes with the W boson decaying to $e\nu$. Single-top production also has a small contribution to this background. We have defined a single-electron CR with no reconstructed photon (1e, 0γ) to translate to the prediction in the SR having a reconstructed fake photon (0e, 1γ). To ensure that reconstructed photon in SR is indeed generator level electron (gen-e) from the W-boson mis-reconstructed as photon, the ΔR between reconstructed photon and gen-e should be less than 0.2 and ratio of their respective p_{T} ($p_{\mathrm{T}}^{\gamma}/p_{\mathrm{T}}^{gen-e}$) should be within 0.8 to 1.2. Except this, the fake photon SR is exactly similar to the signal region for the analysis selected using the baseline selections discussed in Section 7.2.1. In single-electron CR, events with a nonzero number of muons, isolated muon tracks, and isolated pion tracks are excluded. The reconstructed electron should satisfy the criteria discussed for a good electron veto in Section 7.2.1 and its p_{T} should be > 40 GeV. In order to prevent signal contribution in CR, a condition of M_{T} (reco electron, $p_{\mathrm{T}}^{\mathrm{miss}}$) < 100 GeV is implemented, analogously to the criteria used for single-lepton CR in the context of LL.

Transfer factors are computed using MC simulations, with an aggregated contribution from $t\bar{t}$+jets, W+jets, and single top quark production using MC simulated event sample corresponding to full Run 2 luminosity and are defined as the following:

$$TF(p_{\mathrm{T}}^{EM-obj}, q-multi)\ (i) = \frac{N_{0e+1\gamma}^{MC}\ (i)}{N_{1e+0\gamma}^{MC}\ (i)} \tag{7.5}$$

Where $N_{0e+1\gamma}^{MC}\ (i)$ and $N_{1e+0\gamma}^{MC}\ (i)$ are the SR and CR event yields from MC in i^{th} TF bin. The p_{T}^{EM-obj} is p_{T} of an electromagnetic (EM) object; an electron in CS and photon in SR, in the event, and the charge multiplicity is defined as the multiplicity of charged particles in a jet in the vicinity of the EM object within $\Delta R < 0.3$. Both variables are used to parameterize TF and more details are discussed in Section 7.3.3.
The predicted event yield in MC for j^{th} SR bin thus can be obtained as the following;

$$N_{0e+1\gamma}^{MC}\ (j) = TF(p_{\mathrm{T}}^{EM-obj}, q-multi) \times N_{1e+0\gamma}^{MC}\ (j) \tag{7.6}$$

where $N_{0e+1\gamma}^{MC}\ (j)$ and $N_{1e+0\gamma}^{MC}\ (j)$ are the predicted and expected event yields using MC in j^{th} SR bin.
A data-MC correction factor α (also called scale factor) is calculated to take into account the differences in the misidentification rate between data and MC (see Section 7.3.3). The predicted event yields in data for j^{th} SR bin is computed as per the following equation;

$$N_{0e+1\gamma}^{Data}\ (j) = TF(p_{\mathrm{T}}^{EM-obj}, q-multi) \times \alpha(N_{\mathrm{bjets}}) \times N_{1e+0\gamma}^{Data}\ (j) \tag{7.7}$$

where $N_{0e+1\gamma}^{Data}\ (j)$ and $N_{1e+0\gamma}^{Data}\ (j)$ are the predicted and expected event yields using MC in j^{th} SR bin.

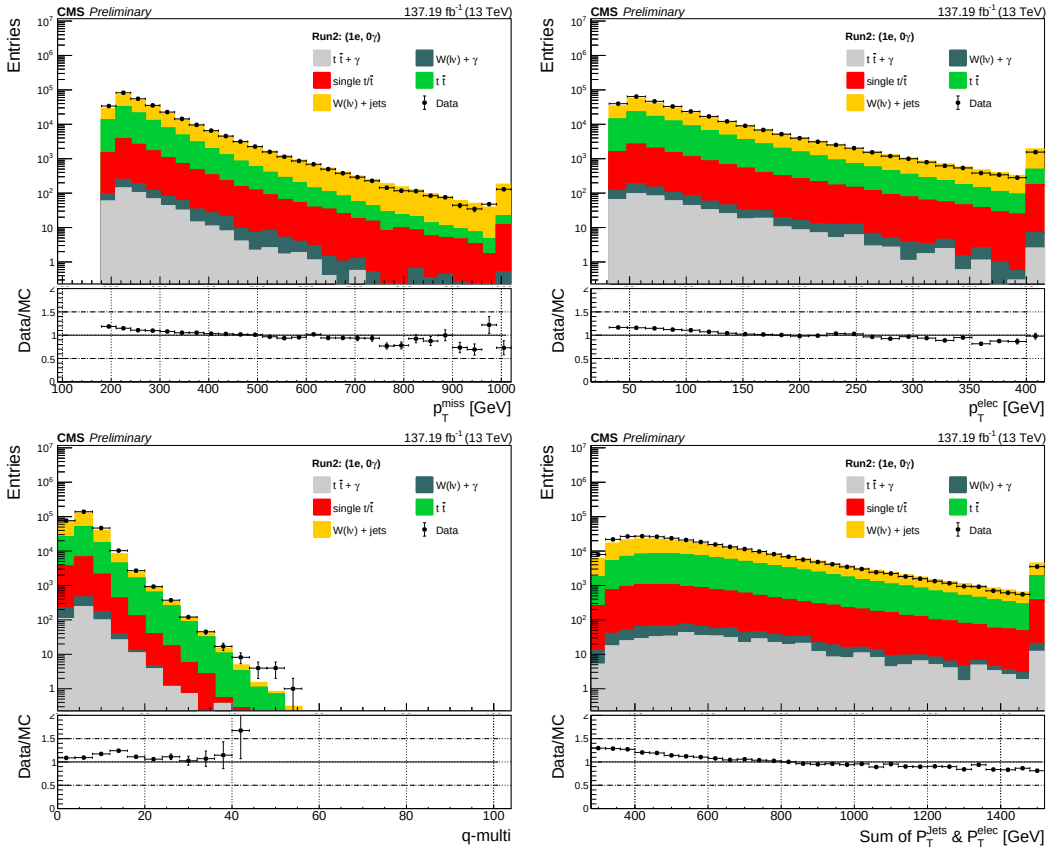


Figure 7.28: $p_{\mathrm{T}}^{\mathrm{miss}}$ (top left), p_{T}^{elec} (top right), q-multiplicity (bottom left), and S_{T} (bottom right) for single electron CR are compared in MC (filled histograms) and Data (black point) with complete Run 2 luminosity.

Comparison of single electron CR in Data & MC

The full Run 2 $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger datasets are used for SR in case of data and the corresponding trigger efficiencies are parameterized as a function of $p_{\mathrm{T}}^{\mathrm{miss}}$ and applied on MC simulations. The b-tag SF and lepton SFs are not taken into account.
In addition to corrections to MC and cleaning in data detailed in Sections 7.2.1 and 7.3.1, we also apply supplementary data cleaning measures to address complications associated with pixel tracker layers, denoted as "pixel-veto". In the 2017 and 2018 data, a large region of low hit efficiency was observed attributed to High Voltage (HV) power supply issues for two pixel layers within the same sector. Hence, the misidentification of an electron as fake has more probability in these certain $\eta\&\phi$ regions. Based on the studies done in the search [170], we have vetoed the events if EM object (electron (photon) in CR (SR) lies within the following regions;

$$\begin{aligned} For\ \ 2017 = 0 < \eta < 1.5\ \ and\ \ \phi > 2.7 \\ and\ \ for\ \ 2018 = 0 < \eta < 1.5\ \ and\ \ 0.4 < \phi < 0.8 \end{aligned} \tag{7.8}$$

Before calculating the TF using MC simulations of $\mathrm{t\bar{t}}$+jets & W+jets, MC has been corrected for b-jet tagging SFs, JECs and JERs. To evaluate the robustness of the estimation method, the single-electron control region, derived from the full Run 2 luminosity in both the MC simulations and the data, is compared and presented as $p_{\mathrm{T}}^{\mathrm{miss}}$ (top left), p_{T}^{elec} (top right), q-multiplicity (bottom left), and S_{T} (bottom right) in

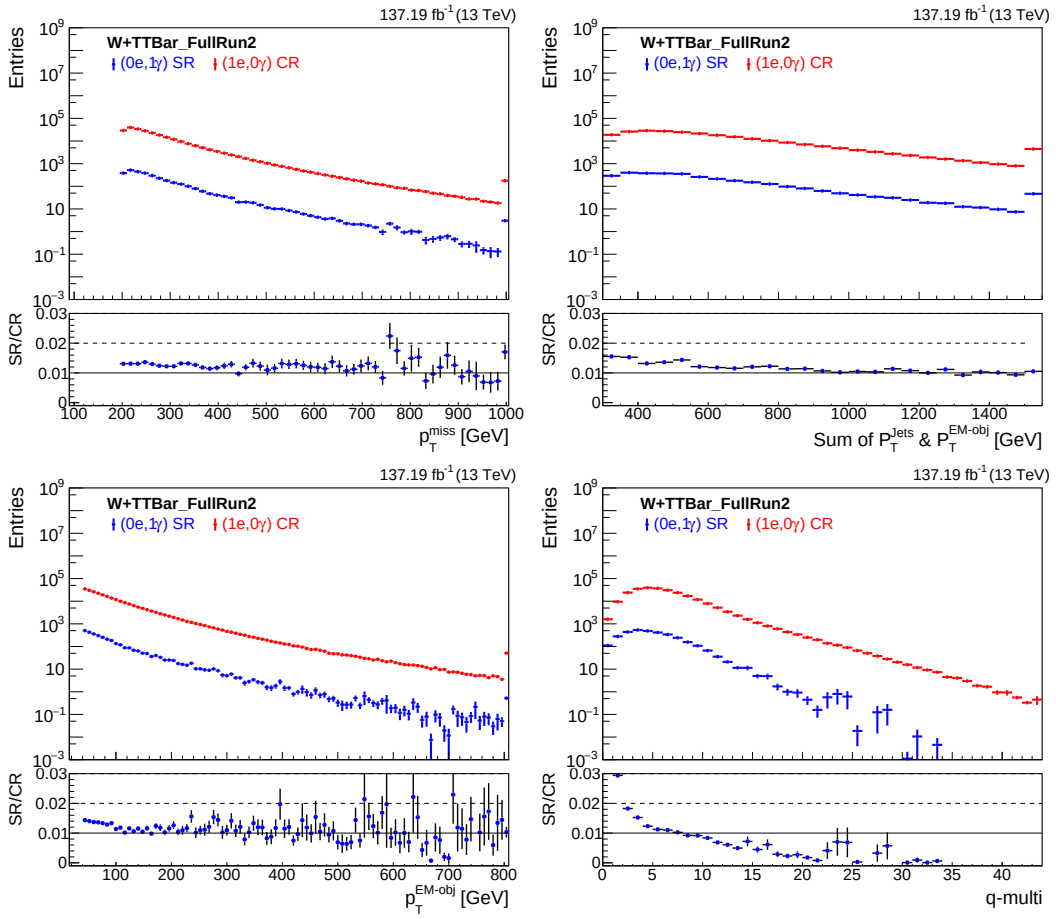


Figure 7.29: $p_{\mathrm{T}}^{\mathrm{miss}}$ (top left), S_{T} (top right), p_{T}^{EM-obj} (bottom left) and the q-multiplicity (bottom right) in single electron CR and photon SR in MC with bottom panel representing the ratio of SR to CR.

Fig. 7.28. The bottom panel of the plots indicates the data to the MC ratio. In the bulk, data-MC both agree well with some 20-30% deviation of the higher end tails. This shows that the CR exhibits substantial similarity between the data and the MC simulations. However we note that these MC distributions are not used directly for background predictions.

Transfer factor calculation

The CR is chosen to be as close as possible to SR especially in terms of event topologies. The shape of different variables; $p_{\mathrm{T}}^{\mathrm{miss}}$ (top left), S_{T} (top right), p_{T}^{EM-obj} (bottom left) and the q-multiplicity (bottom right), in CR and SR using MC are compared in Fig. 7.29. Contributions from $t\bar{t}$+jets, W+jets are combined for the full Run 2 luminosity in MC. The bottom panel of the plots indicates the ratio of SR to CR. The misidentification rate shows a higher-order dependence on the p_{T} of the EM object and the charge multiplicity. The charge multiplicity is an approximate variable that indicates the activity around the EM object.

In order to capture these variations in misidentification rate as a function of p_{T} of EM objects and q-multiplicity, we have established 44 bins with 4 bins in q-multiplicity and 11 bins in p_{T}^{EM-obj}, to compute individual TF and are shown in Fig. 7.30 for 2016 (top left), 2017 (top right), 2018 (bottom left) and full Run 2 (bottom right). High TF

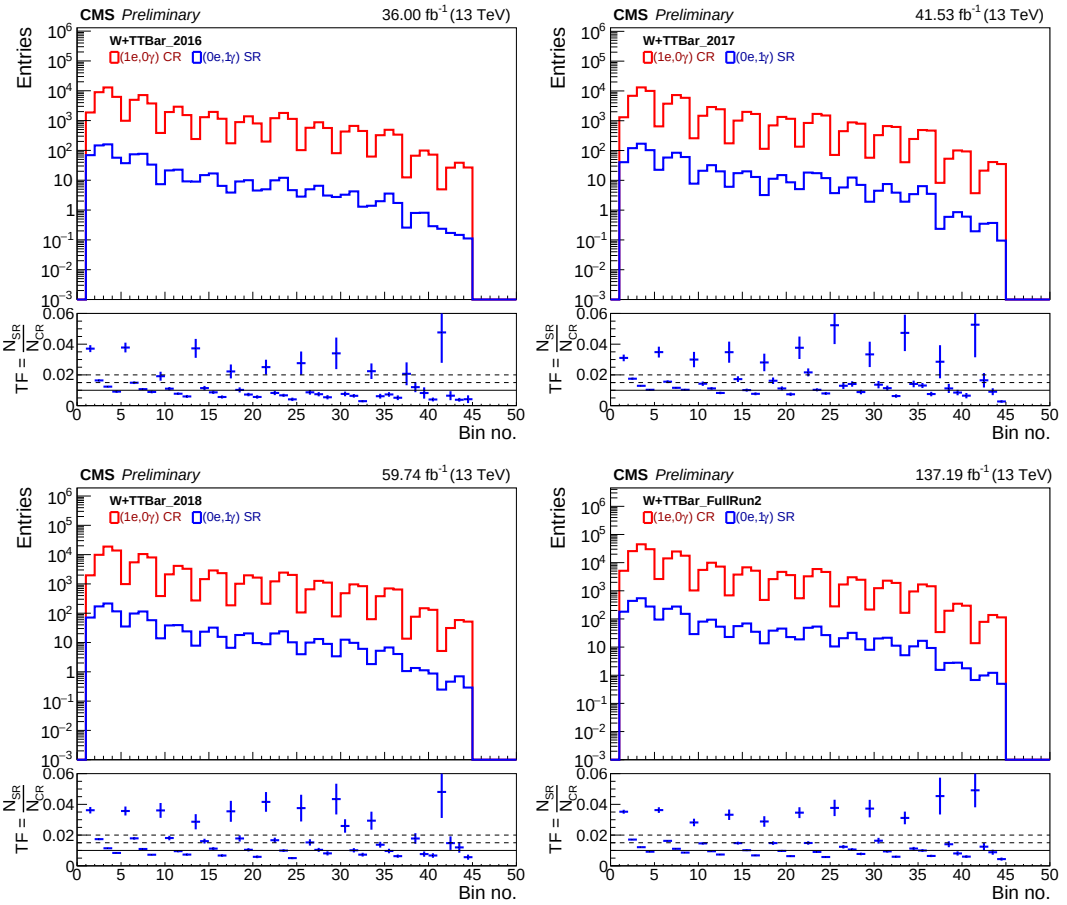


Figure 7.30: CR (red curve) and SR (blue curve) (upper panel) and parametrization of TF (bottom panel)as a function of different bins in p_{T}^{EM-obj} and q-multiplicity for 2016 (top left), 2017 (top right), 2018 (bottom left) and full Run 2 (bottom right).

obtained for TF bins with a small number of q-multiplicity and p_{T}^{EM-obj}.

We have also tried using coarser bins based on these two variables and using different combinations of variables such as N_{jets}, N_{bjets}, p_{T}^{EM-obj}. We obtained the best closure test with the 44 bins presented here in the text.

Closure test

We put the expected CR of MC and the TF obtained with complete Run 2 luminosity (shown to the bottom right of Fig. 7.30) in equation 7.6 to obtain the predicted SR yields in MC. These yields are then compared with the number in the box that we obtained with MC (expected SR). For validation, we used 100 search region bins designed using $p_{\mathrm{T}}^{\mathrm{miss}}$, S_{T},p_{T}^{γ}, and N_{bjets} as discussed in Section 7.2.
Fig. 7.31 shows the comparison of expected SR (blue points) and predicted SR (red points) in MC. The bottom panel of the plots represents the expected and predicted SR ratio. Each vertical block in this plot shows the $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution in a S_{T} block (with a fixed value S_{T}). The overall closure looks good as expected and the predicted values are in great agreement within 10-20% and the statistical uncertainties.
Fig. 7.32 shows the comparisons of expected and predicted event yields as a function of p_{T}^{EM-obj} (top right), q-multiplicity (top left), η^{EM-obj} (bottom left) and ϕ^{EM-obj}

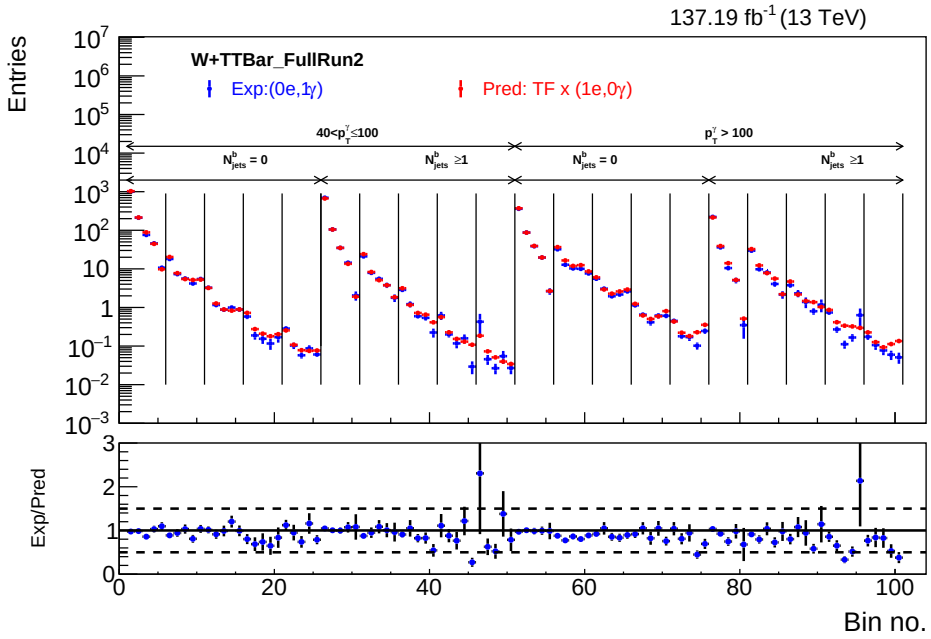


Figure 7.31: Expected SR (blue markers) are compared to those predicted by equation 7.6 (red markers) in MC with Run-2 luminosity. The lower panel shows the expected-to-predicted ratio.

(bottom right) in MC with the full Run 2 dataset. The bottom panel of the plots represents the expected to predicted ratio. As expected, the closure test in terms of p_{T}^{EM-obj} and q-multiplicity is good, especially in bulk. For high $\eta\&\phi$, the ratio of expected to predicted deviates up to and beyond 50%. However, these are the regions of very low statistics and are not expected to effect final results.

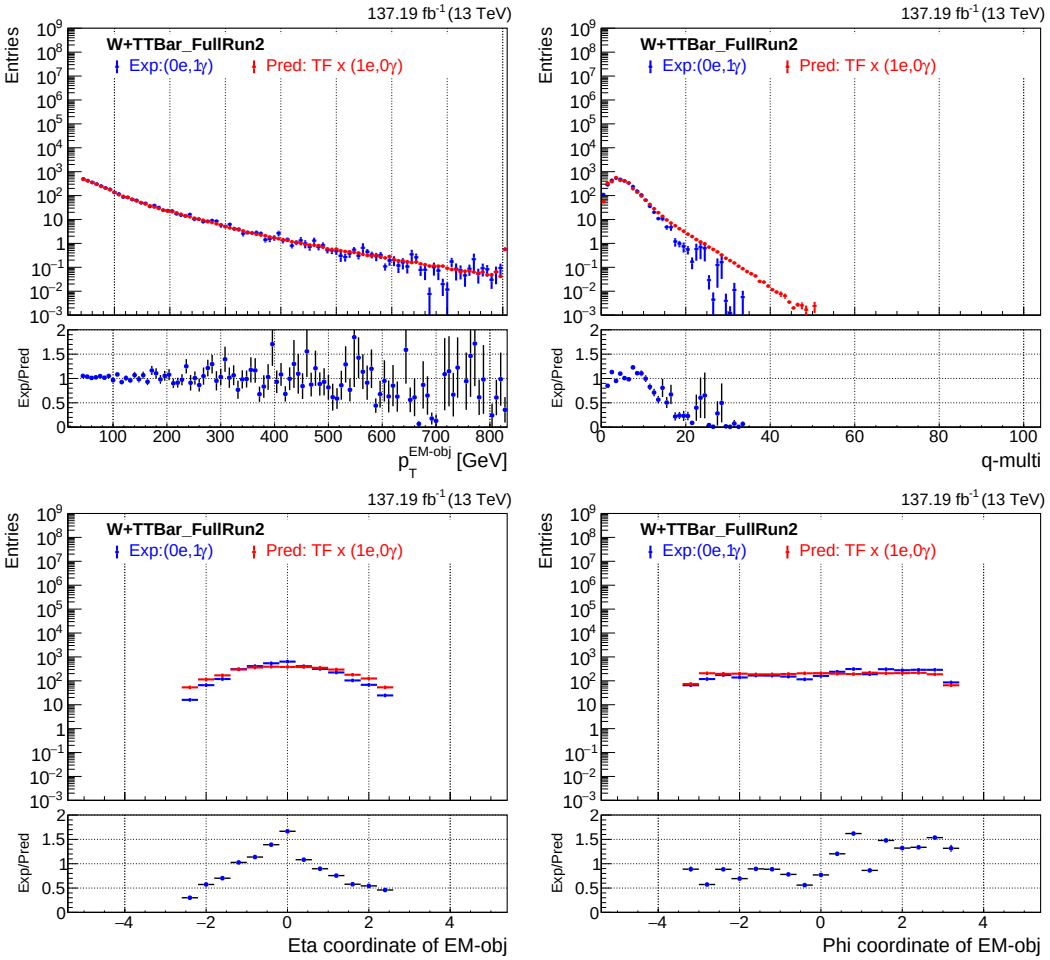


Figure 7.32: Expected and predicted event yields are compared in terms of p_{T}^{EM-obj} (top right), q-multiplicity (top left), η^{EM-obj} (bottom left), and ϕ^{EM-obj} (bottom right) distributions using MC with the full Run 2 dataset. The bottom panel of the plots shows the expected-to-predicted ratio.

Data-MC correction factor

The probability for an electron to be misreconstructed as a photon depends on several factors including event topologies, activity around the electron candidates, and essentially detector efficiencies. As MC simulations are an approximate of what happens in a pp collision, and average run conditions of each year's data taking, the electron efficiencies can be different than those in the data. Even though we do take this difference into account via using electron identification scale factors derived as a function of its p_T and η, the differences in probability of its mis-reconstruction as photons are still not accounted for. Hence, the TF obtained using MC representing this probability should be corrected to realistic scenarios in the data before evaluating the predictions in the data. This correction is referred to as the data-MC correction factor and should be calculated as an independent validation region. We implement Tag & Probe (TnP) method to estimate this factor using di-electron events of the DY ($Z \to \ell\ell\gamma$) process in MC and selected from $HLT_Ele27_WPTight_Gsf_v$ `(2016)`, $HLT_Ele35_WPTight_Gsf_v(2017, 2018)$ in the single-electron dataset (2016, 2017) and the Egamma dataset (2018) for data.
The TnP method is widely used to measure efficiencies in data & MC. It uses resonances that decay to two particles such as the J/ψ particle or Z boson. One of the decay particle candidates passing tight selections is referred to as "Tag" and another particle passing a loose criterion and essentially representing a set of unbiased candidate is referred to as "Probe". The invariant mass of both the tag and probe should be close to the resonance peak.

The tag electron should satisfy the POG recommended `tight` ID and its p_T (p_T^{tag-e}) should be > 40 GeV. The ΔR between the tag and the generator-level (gen-level) electron (in MC) or the high-level trigger electron (HLT) (in data) should be less than 0.3. Additional electrons may either be accurately reconstructed as electrons (denoted as probe e) or incorrectly reconstructed as photons (referred to as probe γ), thereby delineating two distinct scenarios. The probe e(γ) should satisfy the conditions for CR (SR) discussed in Section 7.3.3. The probe e(γ) should be within $\Delta R < 0.2$ with respect to the electron at the gen level in MC (or the HLT electron in the data). In instances where the second HLT electron is absent in the data, this condition is disregarded for those specific events. A supplementary criterion of $\Delta R > 0.2$ is imposed between the tag and the probe to confirm that these objects are distinct particles. The invariant mass of the candidate tag and probe ($m_{\ell\ell}$) must be within the range of 80 to 100 GeV, which confirms that both the tag and probe are consistent with the Z boson resonance. In addition to the above conditions on the tag and probe candidates, we select events with $N_{\text{jets}} \geq 2$, $p_T^{\text{miss}} < 200$ GeV with $\Delta\phi$ (Jets,p_T^{miss})>0.3, and S_T ($=\sum^{jets} p_T + p_T^{tag} + p_T^{probe}$)$> 200$ GeV. Events with non-zero number of reconstructed muons or non-zero number of isolated muon and pion tracks are rejected for this study.
Fig. 7.33 shows the distributions of p_T^{probe} (top row) and N_{bjets} (bottom row) in the events in which the probe is identified as an electron (left column) and as a photon (right column). The MC simulation distribution include small contribution from $Z(\ell\ell)\gamma + jets$, $t\bar{t}$+jets, and W+jets in addition to the $Z \to \ell\ell\gamma$. The ratio panel represents the data-MC ratio in different bins. The data and MC are in good agreement when the probe is an electron. In the case where the probe is γ, the data-MC ratio is high and goes up to 2.5. Non-zero contribution from the $t\bar{t}$+jets process in non-zero N_{bjets} bins as expected.

In both MC and data, the misidentification rate is defined as the ratio of the total

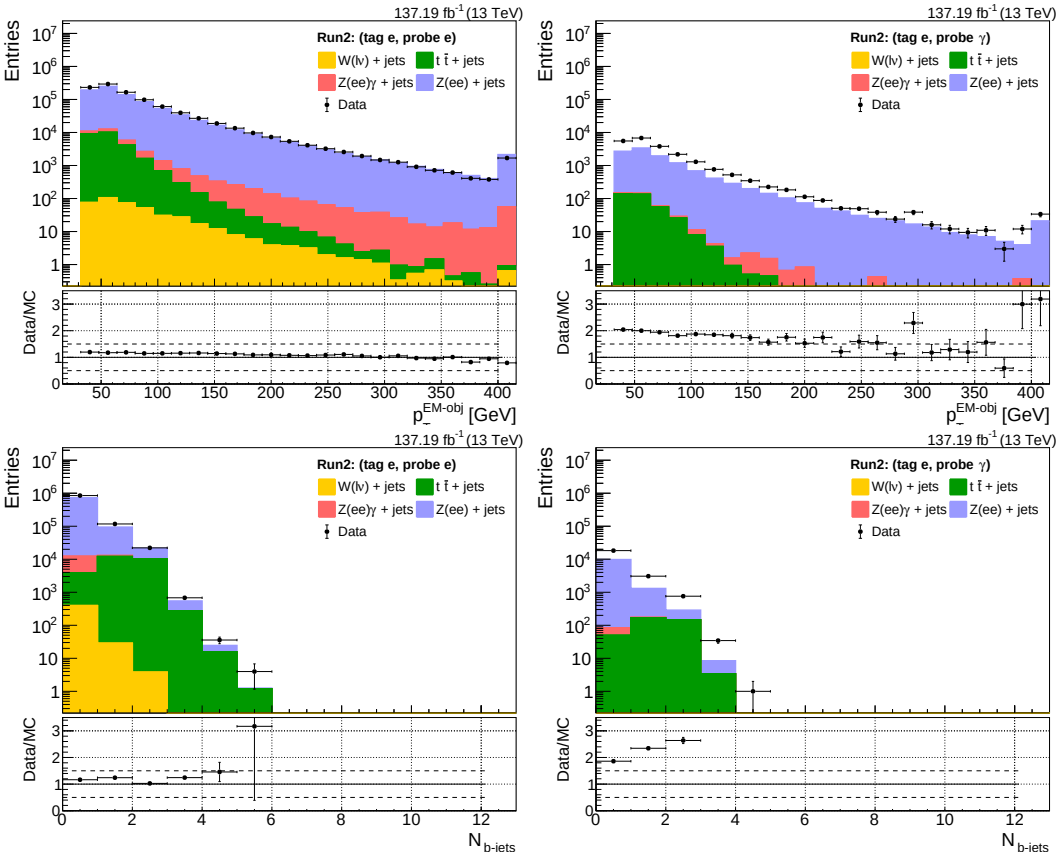


Figure 7.33: For categories where the probe is an electron (first column) and γ (second column), MC and data event distributions are shown as functions of p_{T}^{probe} (top) and N_{bjets} (bottom), with the data-MC ratio across bins.

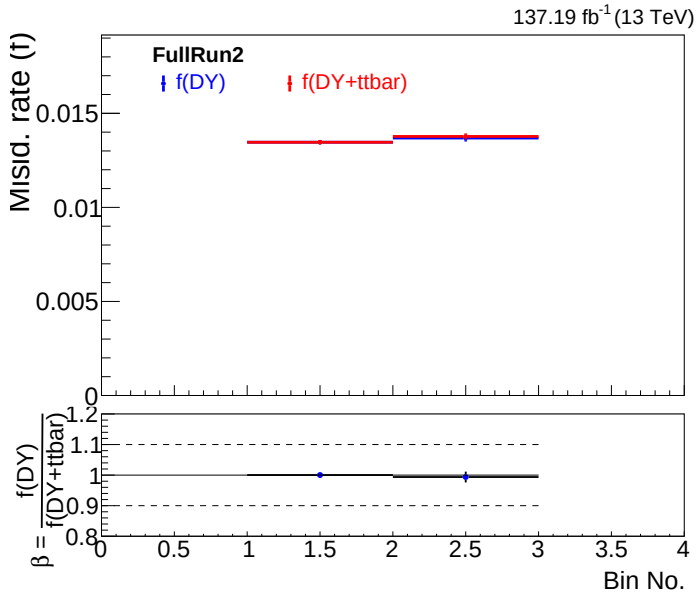


Figure 7.34: Upper panel: misidentification rate from simulations considering only the DY ($Z \to \ell\ell\gamma$) process and from those including both DY ($Z \to \ell\ell\gamma$) and non-DY ($t\bar{t}$+jets) processes. Bottom panel: purity factor (β) in N_{bjets} bins.

events with the probe as γ to the total events with the probe as an electron, as defined below.

$$f^{MC/Data} = \frac{N^{MC/Data}(1\ tag\ e, 1\ probe\ \gamma)}{N^{MC/Data}(1\ tag\ e, 1\ probe\ e)} \tag{7.9}$$

where $N^{MC/Data}(1\ tag\ e, 1\ probe\ \gamma)$ and $N^{MC/Data}(1\ tag\ e, 1\ probe\ e)$ is the total number of events with probe as γ and electron respectively, in MC and data.
To account for the influence of non-DY processes on the correction factor, a purity factor (β) is determined as the following,

$$\beta^{MC}(N_{\text{bjets}}) = \frac{f^{Z\to\ell\ell\gamma}}{f^{Z\to\ell\ell\gamma \& t\bar{t}+\text{jets}}} \tag{7.10}$$

in which $f^{Z\to\ell\ell\gamma}$ and $f^{Z\to\ell\ell\gamma \& t\bar{t}+\text{jets}}$ denote the misidentification rate derived from simulations that consider solely the DY ($Z \to \ell\ell\gamma$) process, and from those that include both DY ($Z \to \ell\ell\gamma$) and non-DY ($t\bar{t}$+jets) processes, respectively. The parameter β is evaluated as a function of N_{bjets} and is observed to be nearly unity under the full Run 2 luminosity, as illustrated in Fig. 7.34. This evaluation is subsequently applied as a weighting factor to di-electron data events across all three years, contingent upon the N_{bjets} multiplicity.
Hence, the data-MC correction factor (α) can be derived as per the following equation,

$$\alpha^{N_{\text{bjets}}} = \beta^{MC}(N_{\text{bjets}}) \times \frac{f^{Data}(N_{\text{bjets}})}{f^{MC}(N_{\text{bjets}})} \tag{7.11}$$

where f^{Data} and f^{MC} are the misidentification rate obtained for data and MC, respectively using equation 7.9.The α is computed and implemented yearwise, taking into account the dynamic alterations in detector conditions over time.
Thus obtained α is shown in the bottom panels of the distributions shown in Fig. 7.35

as a function of different event kinematic variables such as N_{bjets} (top left), N_{jets} (top right), q-mulitplicity (bottom left) and p_{T}^{probe} (bottom right). We do not observe a strong correlations of α with different variables. The calculated α as a function of N_{bjets} (=0, $\geq$1) varies from 1.5 to 2. The $Z \to \ell\ell\gamma$ events exhibit topologies analogous

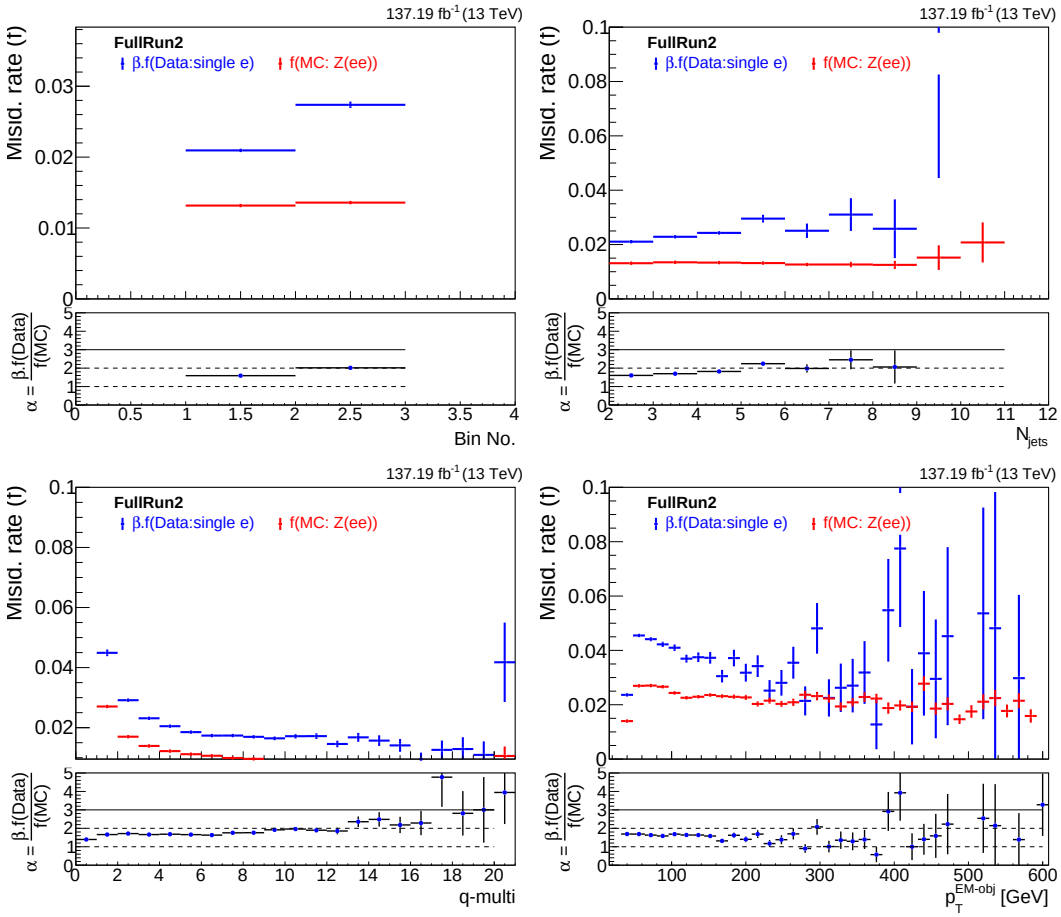


Figure 7.35: The upper panel displays the misidentification rate with blue points for data and red for MC, while the lower panel shows α, as a function of N_{bjets} (top left), N_{jets} (top right), q-multiplicity (bottom left), and p_{T}^{probe} (bottom right).

to W+jets events if a "Tag" electron is incorporated into the $p_{\text{T}}^{\text{miss}}$. However, variations in event topologies may arise in the context of $t\bar{t}$+jets. One way to check this topology bias is by treating tag electron as neutrino and making the resulting event topology of $Z \to \ell\ell\gamma$ events resemble $t\bar{t}$+jets or W+jets events. Thus obtained fake rate using $Z \to \ell\ell\gamma$ samples should be comparable with those obtained from $t\bar{t}$+jets and W+jets. The resultant difference can be added as systematic. In DY samples, the tagging of e as $p_{\text{T}}^{\text{miss}}$ is implemented after selecting events with the magnitude of sum of tag electron and MET vector greater than 200 GeV, if the invariant mass of both the tag and probe objects is in the range of 80-100 GeV following the baseline selections. Thus obtained fake rates for DY process are compared with that from $t\bar{t}$+jets (left) and W+jets (right) and shown in Fig. 7.36. The fake rates from different processes are matching with 10% variations. Similarly, recalculation of data-MC correction factor α with tag electron as $p_{\text{T}}^{\text{miss}}$ is done. This is an independent way to validate the correction factor obtained using T&P method. To validate the T&P method used to calculate α, we calculated α in a similar way considering the tag electron as neutrino

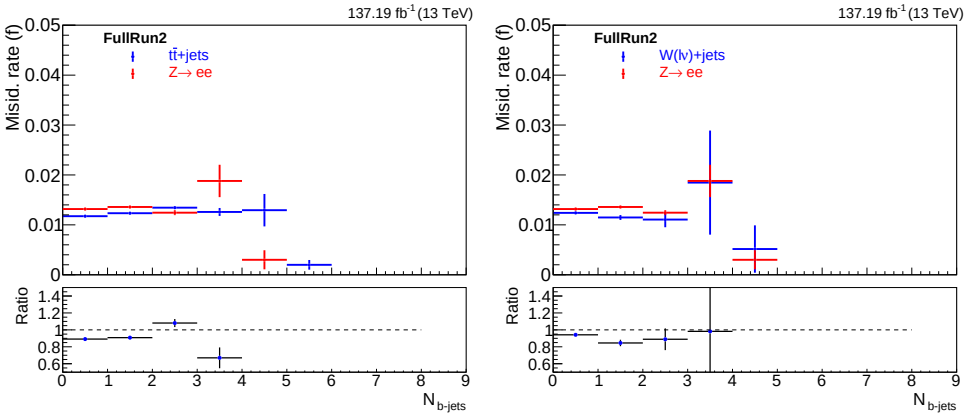


Figure 7.36: A comparison of fake rate for DY process $(Z \to e^+e^-)$ and $t\bar{t}$ is presented on the left as a function of the variable N_{bjets}, using W/$t\bar{t}$ like selections. On the right, a similar comparison is shown for DY $(Z \to e^+e^-)$ and W+jets.

in both data and MC and calculate misidentification rates and α according to equations 7.9 and 7.11, respectively. Fig. 7.37 displays the comparison of fake rate (left) and correction factor (α) (right) obtained using T&P method (blue) and using W/$t\bar{t}$ like selections (red). A maximum variation of 3% and 20% is observed in fake rate and α, respectively and this is added as systematics on the correction factor.

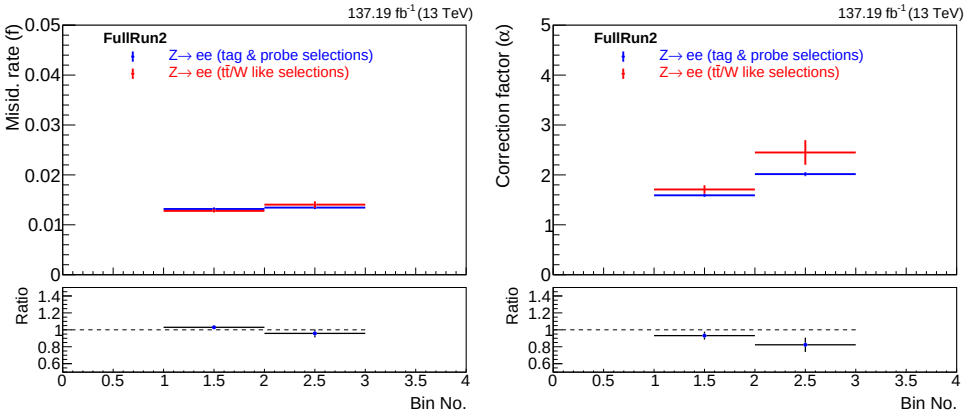


Figure 7.37: A comparison of fake rate obtained using Tag and probe method (blue line) and using W/$t\bar{t}$ like selections (red line) for DY process is shown on left. On the right, comparison of scale factor obtained with different methods is shown.

7.3.4 Predictions of electron faking photon background in data & sources of uncertainties

Similar to lost lepton backgrounds, systematics due to the uncertainty in scale factors calculations, jet energy corrections & resolution uncertainties, and uncertainties in the scale factor for b-tagging efficiencies, also contribute to the uncertainties in the measurements of electron faking photon background. We have studied the impact of these individual uncertainties on the predictions for fake photon background:

- Impact of jet energy corrections (JEC) and jet energy resolution (JER) uncertainties, derived from data, is also assessed by varying these by $+1\sigma$ and shown in top left and top right plots of Fig. 7.38, respectively. For varying these uncertainties by -1σ, similar comparison is added in top left and top right plot in Fig. B.3 of Appendix B.1. The variations of 0-5 (0-10)% and 0-10 (0-7)% in TF

are obtained for JEC and JER, respectively when varied by $+1\sigma$ (-1σ). These variations of TF with respect to JEC/JER, introduce a 0-5 (0-4)% uncertainty in the predictions.

- Uncertainties due to scale factor for b-tagging efficiencies are evaluated by varying them by $+1\sigma$ and -1σ, and corresponding TFs are compared in bottom plot of Fig. 7.38 and Fig. B.3, respectively. A maximum of 5% variation in TF is observed and corresponding variation in predictions varies in 0-15%.
- In addition to the above, uncertainties from differences in correction factor obtained using T&P method and the SF calculation methods (e.g., tag-and-probe vs. tag electron as MET), and potential biases from parametrization of SFs with different kinematic variables.

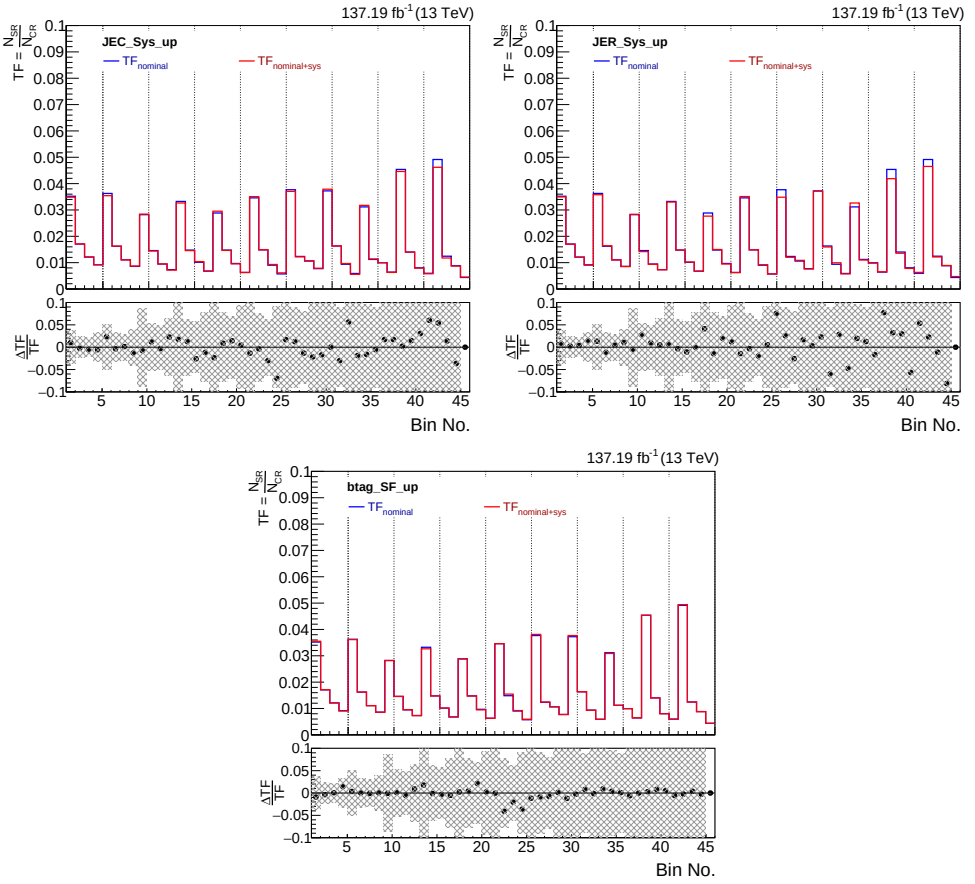


Figure 7.38: Comparison of nominal TF (in blue) with TF after varying the JEC factors (top left) and JER (top right) by $+1\sigma_{sys}$, and varying the scale factor of b-tagging (bottom) are shown in the top panel. Shaded region represents the statistical uncertainty in the bottom panel and black points are showing relativistic differences in the TF defined as the ratio of $(TF_{nominal} - TF_{nominal+sys})$ to $TF_{nominal}$.

Fig 7.39 shows distributions of the predicted electron faking photon background in data using equation 7.7 in 100 SRs devised based on $p_{\mathrm{T}}^{\mathrm{miss}}$, S_{T}, N_{bjets} & p_{T}^{γ}. The black points represents the statistical uncertainty. Within the scope of thesis, we chose to incorporate a 20% systematic uncertainty in addition to statistical uncertainties while evaluating projected exclusion limits discussed in Section 7.4.

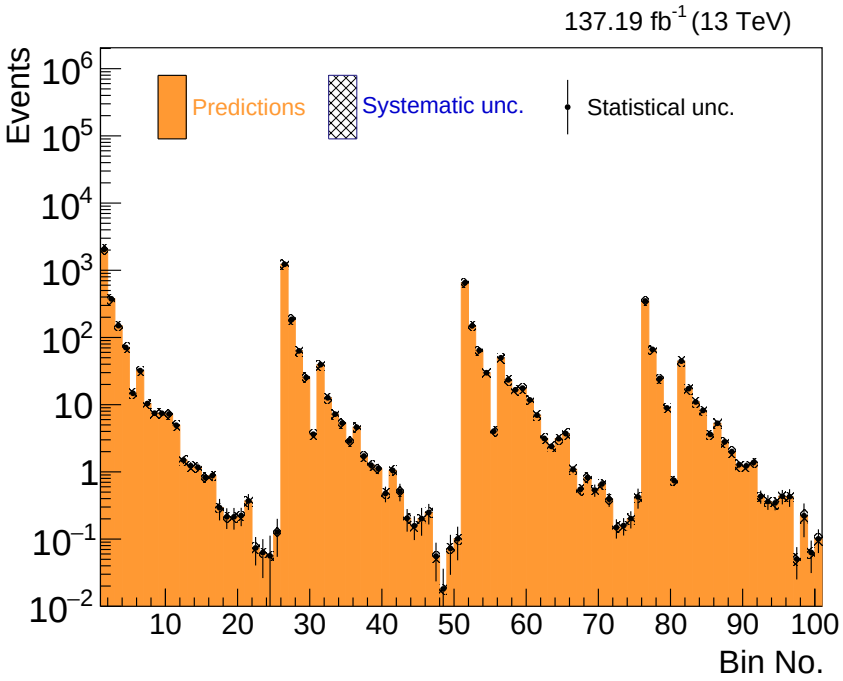


Figure 7.39: The predicted background events in the SRs for electron faking photon background. The black points indicate statistical uncertainty and shaded region represents systematic uncertainty.

7.3.5 Estimation of Z-invisible background

The pair $\nu\bar{\nu}$ in $Z(\nu\bar{\nu})\gamma + jets$ is a source of true and high p_T^{miss} signature on the detector with low N_{jets}. Due to the exceedingly small branching ratio of the Z boson's leptonic decay (∼10%), $Z(\ell\ell)\gamma + jets$ in MC (dilepton events in the data) possesses limited statistical significance. Consequently, $Z(\ell\ell)\gamma + jets$ in MC (dilepton events in data), events such as CR are not utilized for prediction purposes, as is typical with the average transfer factor method. In lieu of this approach, we employ $Z(\ell\ell)\gamma + jets$ events to obtain the data-MC correction factor (α) and utilize $Z(\nu\bar{\nu})\gamma + jets$ directly to derive predictions for the data in various SR (S_T, p_T^{miss}, N_{bjets}, p_T^{γ}) bins, weighted by this correction factor.

A dilepton control region is defined using the presence of two leptons that share the same flavor but possess opposite charges, according to the criteria for well-reconstructed electrons and muons as outlined in Section 7.2.1. The invariant mass of the dilepton pair must reside in the range of 80 to 100 GeV. Events characterized by p_T^{miss} <200 GeV are selected, and $p_T^{miss}{}'$ (with the two leptons' p_T added to p_T^{miss}) > 200 GeV, the resulting event topology is expected to resemble $Z(\nu\bar{\nu})\gamma + jets$. To ensure that the selected reconstructed photon is a prompt photon and not radiated from the electron, an additional criteria of ΔR (γ, e)>0.2. In MC, we use the combined contributions from $Z(\ell\ell)\gamma + jets$ and $Z \rightarrow \ell\ell\gamma$ +jets with secondary contribution from $t\bar{t}\gamma$+jets and in data, we make use of the single electron and muon trigger data sets across full Run 2. The rest of SR selections are also applied except veto of isolated electron and muon tracks.
In MC, it is important to validate that $Z(\ell\ell)\gamma + jets$ events in CR represent topologies similar to $Z(\nu\bar{\nu})\gamma + jets$ events. This comparative analysis of shapes is illustrated in Fig. 7.40, which features p_T^{miss} ($p_T^{miss}{}'$ in $Z \rightarrow \ell\ell\gamma$ events) on the top left, p_T^{γ} on the top right, S_T on the bottom left and N_{bjets} on the bottom right. The blue curve represents the cumulative contributions from the combined events of $Z(\nu\bar{\nu})\gamma + jets$ and $Z \rightarrow \nu\bar{\nu}$ + jets (SR), while the red curve illustrates the combined events of $Z(\ell\ell)\gamma + jets$ and $Z \rightarrow \ell\ell\gamma$ + jets (CR). The overall event topologies can indeed be accurately reproduced by employing $Z(\ell\ell)\gamma + jets$ events.
The data and MC comparisons in CR are shown in Fig. 7.41 as in the distributions N_{bjets} (top left), $m_{\ell\ell}$ (top right), p_T^{γ} (bottom left) and S_T (bottom right). The bottom panel represents the ratio of data to MC in different bins, which is essentially the correction factor that we want to calculate. There is a nonzero contribution from non-DY processes, especially in nonzero N_{bjets} bins, which should be taken into account before calculating the data-MC correction factor.
Thus the predictions in data ($N_{\nu\bar{\nu}}^{data}$ (i) for invisible background in i^{th} SR bins are,

$$N_{\nu\bar{\nu}}^{data}\ (i) = \alpha(N_{bjets}) \times N_{\nu\bar{\nu}}^{MC}\ (i) \tag{7.12}$$

where $N_{\nu\bar{\nu}}^{MC}$ are the event yields for i^{th} SR bins in MC and α is the data-MC correction factor parametrized in terms of N_{bjets} and defined as per the following equation,

$$\alpha(N_{bjets}) = \frac{\beta_{ll+\gamma}(N_{bjets}) \times N_{ll\gamma}^{Data}}{N_{ll\gamma}^{MC}} \tag{7.13}$$

where $N_{ll\gamma}^{Data}$ and $N_{ll\gamma}^{MC}$ are the total number of event yields from data and MC, respectively, in the N_{bjets} bins and $\beta_{ll+\gamma}$ is a purity factor derived to take into account the non-DY process contribution in Z-mass window and can be evaluated using eμ

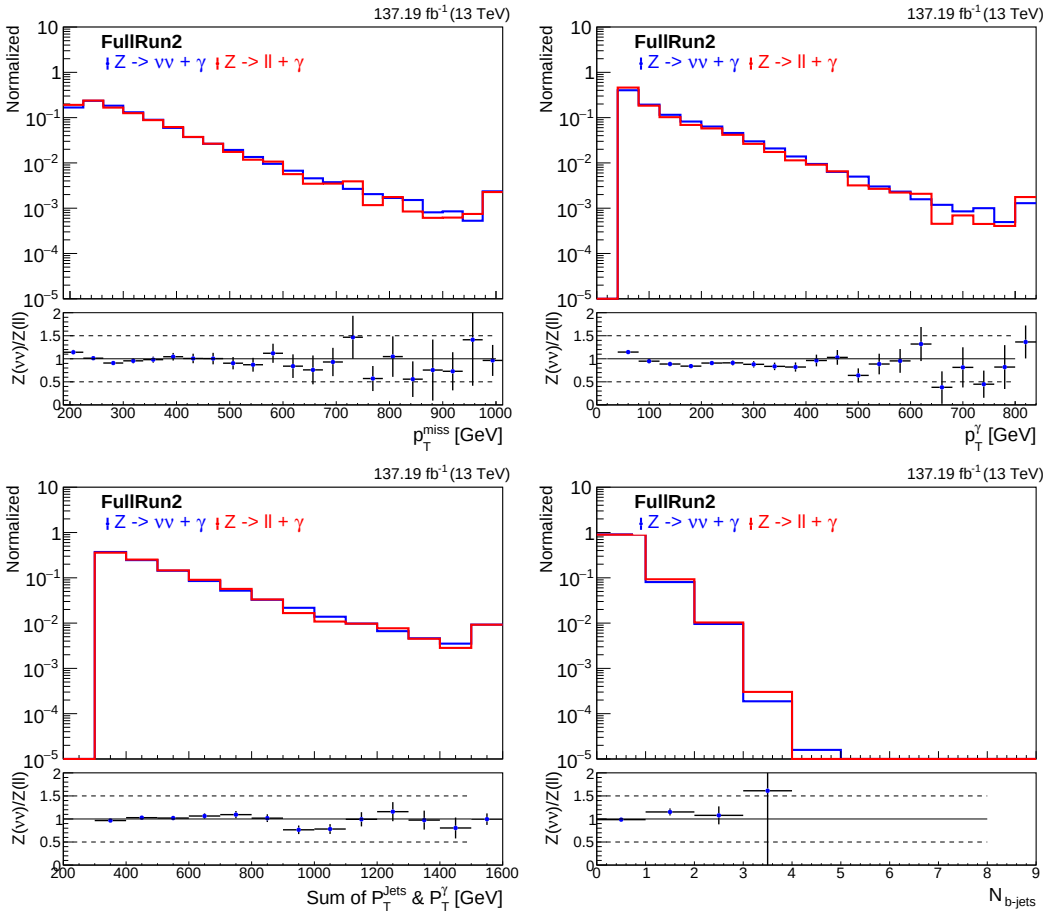


Figure 7.40: The $p_{\mathrm{T}}^{\mathrm{miss}}$ in the top left, p_{T}^{γ} in the top right, S_{T} in the bottom left, and N_{bjets} in the bottom right distributions in $\mathrm{Z}(\nu\bar{\nu})\gamma + jets$ SR and $\mathrm{Z}(\ell\ell)\gamma + jets$ CR.

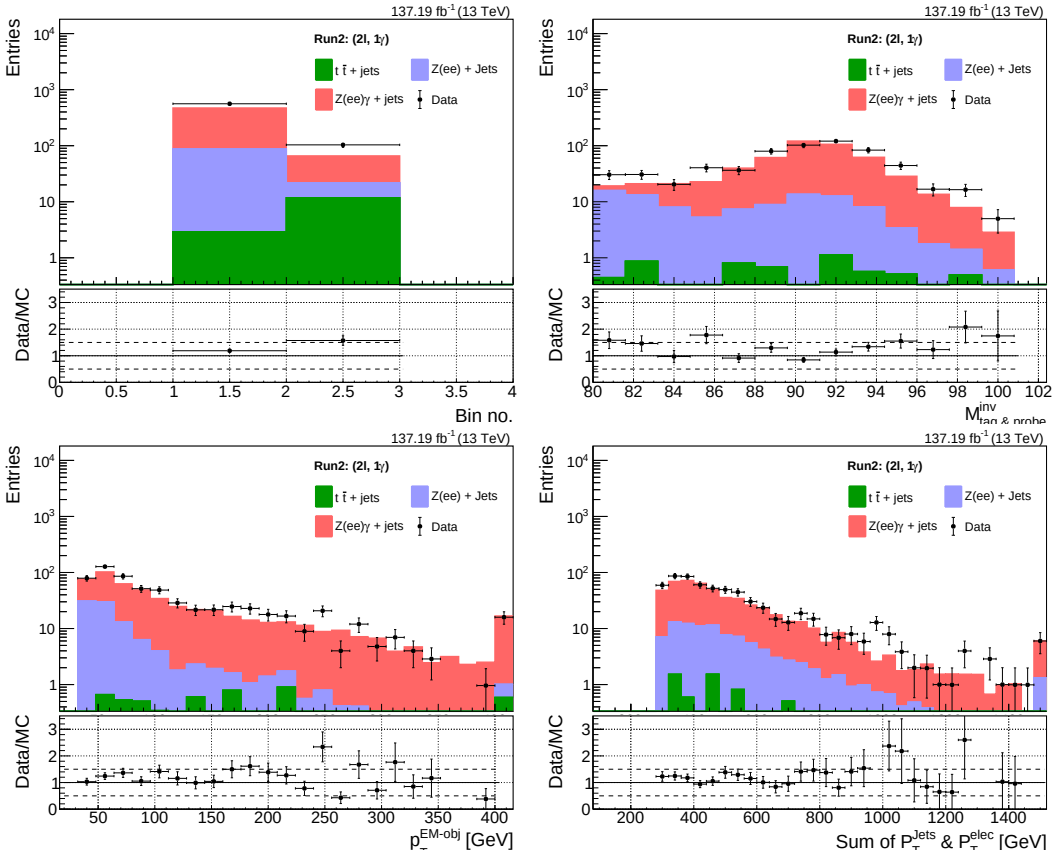


Figure 7.41: Data-MC comparisons in dilepton control region in terms of N_{bjets} (top left), $m_{\ell\ell}$ (top right), p_{T}^{γ} (bottom left) and S_{T} (bottom right) with bottom panel as the data to MC ratio. The non-DY contributions are not mitigated from data yet.

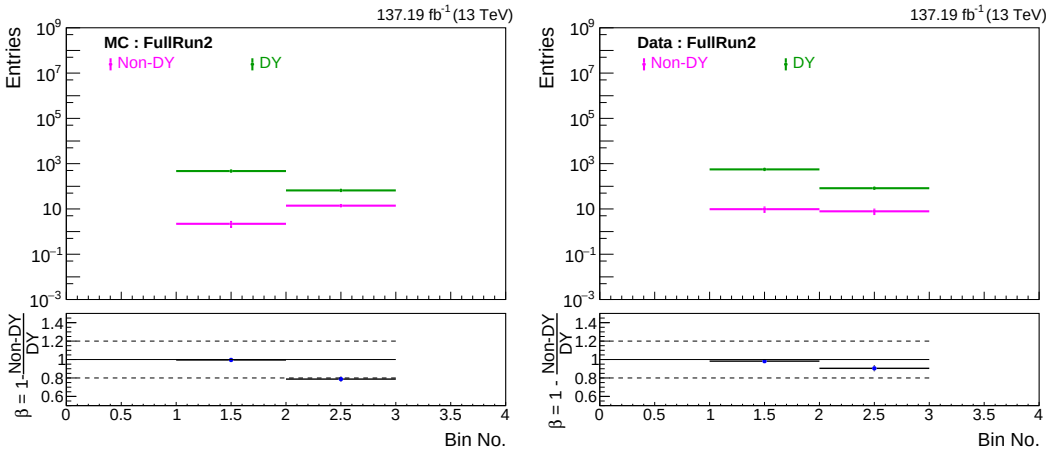


Figure 7.42: Purity factor obtained in N_{bjets} bins and shown in bottom panel using MC (left) and data (right) with full Run 2 luminosity.

events in data or MC as defined in the following equation,

$$\beta_{ll+\gamma} = 1 - \frac{N^{MC/Data}_{e\mu+\gamma}}{N^{MC/Data}_{ll+\gamma}} \tag{7.14}$$

where $N^{MC/Data}_{e\mu+\gamma}$ and $N^{MC/Data}_{ll+\gamma}$ are the total number of events with eμ+γ and ll+γ, respectively, in MC (t$\bar{\text{t}}\gamma$+jets) or data.
Fig. 7.42 shows the contributions of DY (Z $\rightarrow$ $\ell\ell\gamma$) in green points and non-DY (t$\bar{\text{t}}$+jets) in magenta points in N_{bjets} bins (=0,$\geq$1) obtained using MC (left) and data (right). The bottom panel shows thus obtained β defined per equation 7.14. The uncertainties associated with β when derived from MC are notably smaller compared to those obtained from the data. Consequently, these values are utilized to determine the data-MC correction factor, following the adjustment of the data events in the dilepton control region with the purity factor referenced in N_{bjets}.
After putting β in equation 7.13, the α can be calculated. Thus, the obtained α is shown in the bottom panel of Fig. 7.43 where blue points in the upper panels are the total event yields in CR using MC and black points are those for the data.
Using this correction factor, we can get predictions in the data for $Z(\nu\bar{\nu})\gamma + jets$ background in SR.

7.3.6 Predictions of Z-invisible background & sources of uncertainties

Data CR statistical uncertainty predominantly contributes to the uncertainty in the predicted number of background events for Z-invisible background estimation. Besides, the systematic uncertainties including b-tagging scale factors, jet energy corrections, pileup modeling, also contribute. Also note, the lepton SFs to MC are not applied in this version of analysis. We have evaluated the impact of different systematics on the final predictions listed as below;

- Impact of Jet energy corrections (JEC) and jet energy resolution (JER) uncertainties, derived from data, is also assessed by varying these by $+1\sigma$ and -1σ and their impact on obtained scale factor is shown in top left and top right plots of Fig. 7.44 and Fig. B.4, respectively. The variation of 0-10% and 0-2% in α is obtained for JEC and JER, respectively. These variations of α with respect to JEC and JER, introduce a 0-14% and 0-7 % uncertainty in the predictions, respectively.

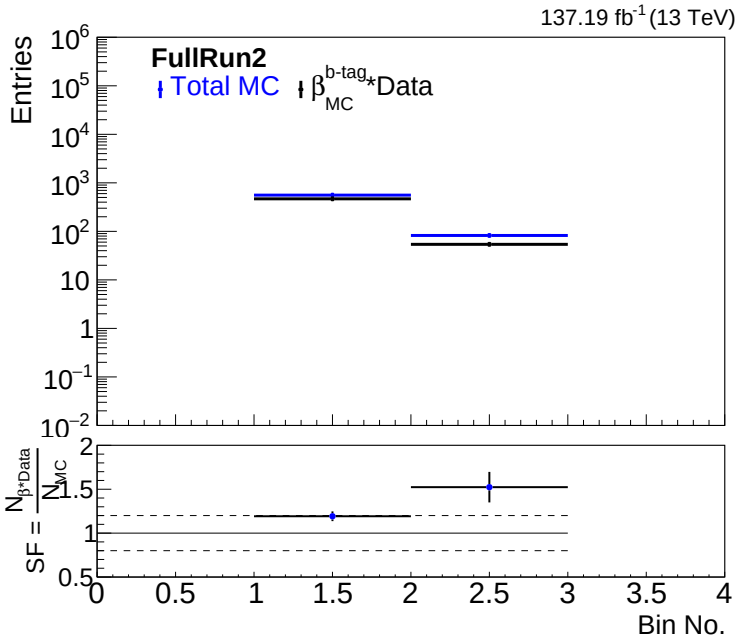


Figure 7.43: Data-MC correction factor (α) in N_{bjets} bins is shown in the bottom panel.

- Uncertainties in the pile up weights and scale factor for b-tagging efficiencies are evaluated by varying them by $+1\sigma$ and $+1\sigma$, and their impact on scale factor is shown in bottom left and right plot of Fig. 7.44 and Fig. B.4, respectively. A maximum of 5% variation in α is observed and corresponding variation in predictions is around 0-10%.
- The uncertainty of α are incorporated into the predicted event yields in the SRs, along with statistical uncertainties from the simulated $Z(\nu\bar{\nu})\gamma + jets$ samples which results into a variation of 10% in final predictions.

Fig 7.45 shows the distributions of the predicted Z invisible background in the data using equation 7.12 in 100 SRs devised based on $p_{\text{T}}^{\text{miss}}$, S_{T}, N_{bjets} & p_{T}^{γ}. The black points represents the statistical uncertainty and shaded region represents systematic uncertainty. Within the scope of thesis, we chose to incorporate a 20% systematic uncertainty in addition to statistical uncertainties while evaluating projected exclusion limits discussed in Section 7.4.

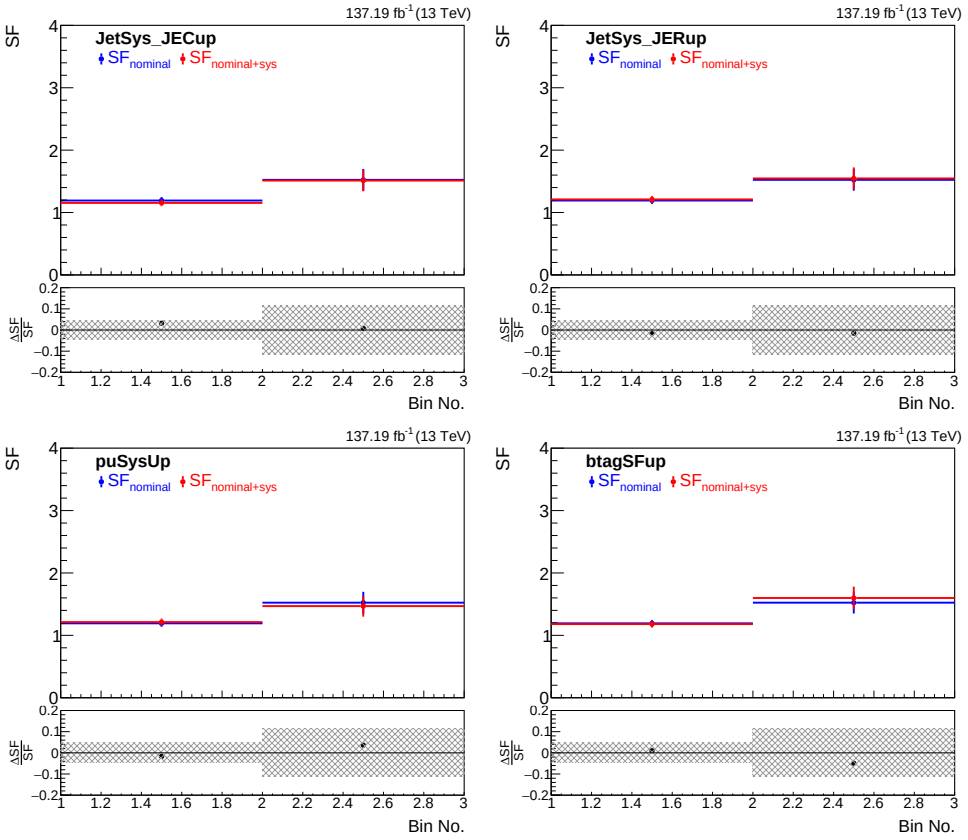


Figure 7.44: Comparison of nominal α or SF (in blue) with α (SF) after varying the JEC factors (top left) and JER (top right) by $\pm 1\sigma_{sys}$, and varying the pile up weights (bottom left) and scale factor of b-tagging (bottom right) are shown in the top panel. Shaded region represents the statistical uncertainty in the bottom panel and black points are showing relativistic differences in the α defined as the ratio of $(\alpha_{nominal} - \alpha_{nominal+sys})$ to $\alpha_{nominal}$.

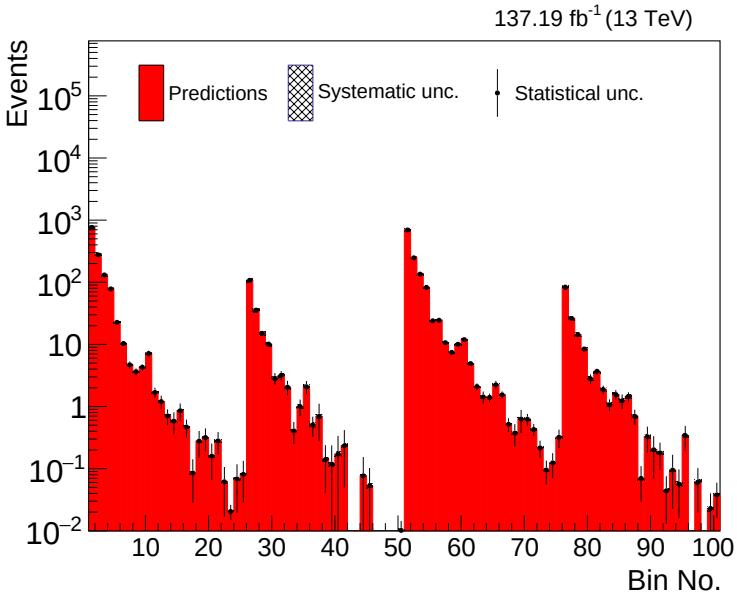


Figure 7.45: The predicted background events in the SRs for $Z(\nu\bar{\nu}) + \gamma$. The black points indicate statistical uncertainty.

7.4 Results & interpretations

The nature of collision events and the detector response is intrinsically stochastic. Statistical methods are used to infer meaningful results from the data. In searches for new physics like SUSY, the objective is to detect new signal events or, in their absence, establish upper limits on the process production cross section. Upper limits are derived as the maximum number of given signal events hidden in the observed data. The *Asymptotic Confidence Level (CLs) method* [171, 172]," based on a modified frequentist approach, provides the means for setting these limits.

Given the division of the search regions into 100 SRs, the procedure of fitting each SR, as well as the calculation of likelihoods and CLs, is a complex and time-intensive endeavor. For analyses involving multiple statistically independent SRs and the complexities of fitting processes, the CMS Collaboration has unified the tools to ensure consistent and comparable outcomes in different studies, named the *HIGGS COMBINE framework* [173]. This tool extracts the signal strength from the fits and converts it to cross sections multiplied by branching ratios or upper limits at 95% confidence. It can perform goodness-of-fit tests, fit diagnostics, and identify systematic uncertainties as nuisance parameters that affect the parameters of interest.

This section describes the methodology for setting such limits, incorporating statistical techniques and systematic uncertainties that are included in the *HIGGS COMBINE framework* and used in this analysis.

Likelihood function

The predictions for both signal and background yields involve various uncertainties as discussed in Section 7.3. These are managed by incorporating nuisance parameters, θ, so that the expectations for signal and background are expressed as functions: s(θ) and b(θ). The notation in the rest of the text is adopted from [174]. The likelihood function quantifies the consistency between the observed data and the model prediction. For a counting experiment with independent bins, the likelihood function $\mathcal{L}(\mu)$ [174] is given by:

$$\mathcal{L}(data|\mu,\theta) = Poisson(data|\mu.s(\theta) + b(\theta)).p(\tilde{\theta}|\theta) \tag{7.15}$$

where μ is the signal strength parameter and "data" is the experimental observation. And, $Poisson(data|\mu.s(\theta) + b(\theta))$ is defined as;

$$Poisson(data|\mu.s(\theta) + b(\theta)) = \prod_i \frac{(\mu S_i + B_i)^{N_i}}{N_i!}.e^{-\mu S_i - B_i} \tag{7.16}$$

Equation 7.15 can be written as,

$$\mathcal{L}(\mu) = \prod_i \frac{(\mu S_i + B_i)^{N_i} e^{-(\mu S_i + B_i)}}{N_i!} \tag{7.17}$$

where S_i and B_i are the expected signal and background events in bin i. N_i is the observed number of events in bin i. The signal strength μ scales the signal prediction across all bins.

Systematic uncertainties are incorporated via nuisance parameters θ, modifying S_i and B_i. The full likelihood becomes:

$$\mathcal{L}(\mu,\theta) = \prod_i \frac{(\mu S_i(\theta) + B_i(\theta))^{N_i} e^{-(\mu S_i(\theta) + S_i(\theta))}}{N_i!} \cdot \prod_j \pi_j(\tilde{\theta}_j|\theta_j), \tag{7.18}$$

where π_j are priors for nuisance parameters, typically Gaussian or log-normal distributions.

Test statistic

A statistical test aims to determine how well observed data supports a null hypothesis, by evaluating its fit to both background-only and signal-plus-background models. The strength of the signal (μ) is evaluated by maximizing the likelihood function based on observed events. Nuisance parameter influences are assessed by profiling the negative logarithmic likelihood (NLL) around its minimum. A likelihood scan on the signal strength is executed for discrete nuisance parameters. The profile likelihood ratio, a widely recognized statistic for such hypothesis testing, is defined as follows:

$$q(\mu) = -2 \ln \frac{\mathcal{L}(\mu, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\mu}, \hat{\theta})}, \tag{7.19}$$

where $\hat{\mu}$ and $\hat{\theta}$ are the maximum likelihood estimates (MLEs) of μ and θ, and $\hat{\hat{\theta}}$ are the profiled nuisance parameters for a fixed μ. Under Wilks' theorem, $q(\mu)$ asymptotically follows a χ^2 distribution, enabling confidence intervals and hypothesis tests.

The procedure for calculating limits involves several sequential steps. First, the observed value of the test statistic, $\tilde{q}_\mu^{\text{obs}}$, is determined for the signal strength parameter μ under test. This serves as a basis for evaluating the compatibility of the observed data with the signal hypothesis. The nuisance parameters $\hat{\theta}_0^{\text{obs}}$ and $\hat{\theta}_\mu^{\text{obs}}$ are then fitted to the data by maximizing the likelihood function for the background-only and signal-plus-background hypotheses, respectively.

Subsequently, toy Monte Carlo simulations are used to generate pseudodatasets, which are used to construct probability density functions (PDFs) $f(\tilde{q}_\mu|\mu, \hat{\theta}_\mu^{\text{obs}})$ and $f(\tilde{q}_\mu|0, \hat{\theta}_0^{\text{obs}})$. These PDFs correspond to the signal plus background hypothesis with strength μ and the background-only hypothesis ($\mu = 0$). For generating the pseudodatasets, the nuisance parameters are fixed to their best-fit values obtained from the observed data, but are allowed to vary during the fitting process.
Using the constructed PDFs, two p values are calculated: p_μ, which quantifies the probability of observing a test statistic at least as extreme as $\tilde{q}_\mu^{\text{obs}}$ under the signal-plus-background hypothesis, and $1 - p_b$, which represents the probability of observing a test statistic at least as extreme as $\tilde{q}_\mu^{\text{obs}}$ under the background-only hypothesis. These p-values are defined as:

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

Pseudo-experiments simulate distributions of $q(\mu)$ for both hypotheses. In high-statistics environments, asymptotic formulas simplify this process while drastically lowering computational cost. The confidence level CL_s is calculated as a ratio of these two probabilities:

$$CL_s(\mu) = \frac{p_\mu}{(1 - p_b)} \tag{7.20}$$

If $CL_s \leq \alpha$ when $\mu = 1$, the signal hypothesis is rejected with a confidence level

of $(1-\alpha)$. The CL_s approach has conservative restrictions, resulting in a higher confidence level than $(1-\alpha)$.

Finally, to find the 95% confidence level upper limit on μ, represented by $\mu_{95\%\text{CL}}$, the value of μ is modified iteratively until $CL_s = 0.05$ (see equation 7.20). This number represents the signal intensity parameter's upper limit.

This statistical framework provides a robust methodology to test new physics and set exclusion limits in collider experiments. The interplay of likelihood functions, nuisance parameters, and test statistics ensures that the results are both precise and interpretable.

7.4.1 Expected exclusion limits using data driven background estimation

Using the measured lost lepton, fake photon and Z-invisble backgrounds as discussed in Section 7.3 and simulated QCD multijet + γ jets events from MC, and corresponding statistical and 20% systematic uncertainties in the SRs, we establish 95% CL upper limits on production cross-sections for different simplified strong production SUSY models detailed in Section 7.1. The limits, determined by maximum likelihood fitting, are compared with theoretical cross section predictions, allowing us to establish minimum mass thresholds for SUSY particles expected to be excluded by this analysis in specific models and final states.

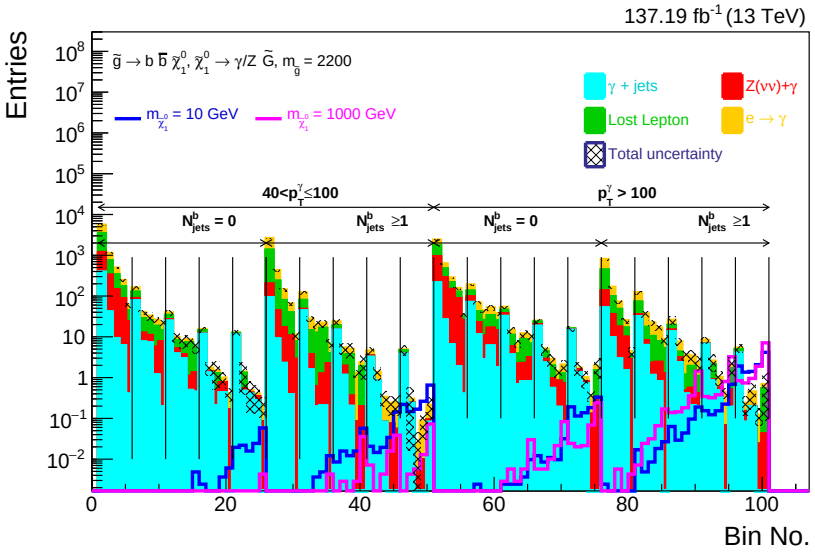


Figure 7.46: Background events in the SRs comprise stacked histograms of lost lepton, fake photon, and $Z(\nu\bar{\nu}) + \gamma$ predictions in data, as well as simulated QCD multijet + γ+jets. The gray-shaded area indicates statistical uncertainty including a 20% systematic error. The open histograms illustrate the T5bbbbZG signal distribution within SRs for $m_{\tilde{g}}$ =2.2 TeV, comparing $m_{\tilde{\chi}^0_1}$ = 10 GeV (blue) and = 1000 GeV (magenta).

Fig. 7.46 shows lost lepton, fake photon, and $Z(\nu\bar{\nu}) + \gamma$ predicted backgrounds, as well as simulated QCD multijet + γ+jets in SRs along with open histograms showing T5bbbbZG signal distribution within SRs for $m_{\tilde{g}}$ =2.2 TeV, comparing $m_{\tilde{\chi}^0_1}$ = 10 GeV (blue) and = 1000 GeV (magenta). Fig. 7.47 illustrates the projected exclusion limits with 95% CL, for gluino ($\tilde{g}$) or stop ($\tilde{t}$) masses (X-axis) and NLSP ($\tilde{\chi}^0_1$) mass (Y-axis) in T5bbbbZG (top left), T5qqqqHG (top right), T5ttttZG (bottom left) and T6ttZG (bottom right) simplified models. We expect up to ~2.4 TeV gluino mass exclusion corresponding to small $\tilde{\chi}^0_1$ masses, in T5bbbbZG & T5qqqqHG models which is 200 GeV more than achieved with high p_T^γ (with >100 GeV) analysis done in past. In

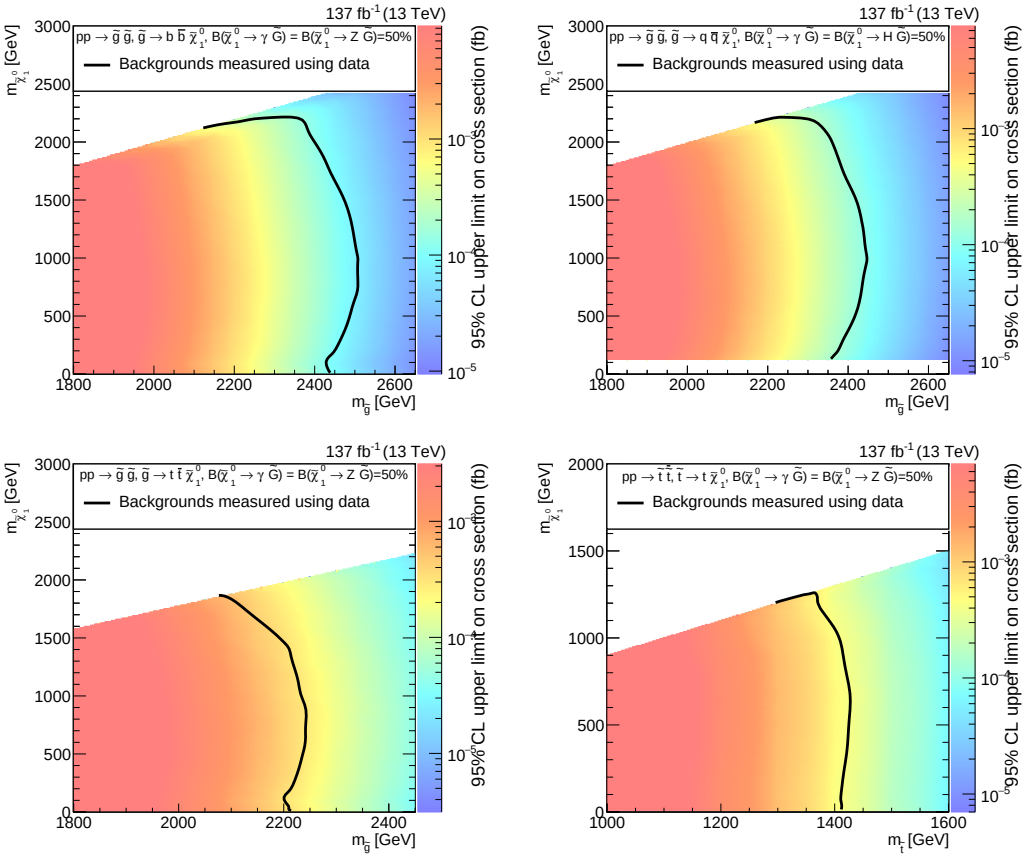


Figure 7.47: Comparison of expected exclusion in data with 95% CL, for T5bbbbZG (top left), T5qqqqHG (top right), T5ttttZG (bottom left) & T6ttZG (bottom right) model.

case of T5ttttZG and T6ttZG simplified models, expected exclusion for gluino and stop masses is around 2.2 TeV and 1.4 TeV, respectively, with small $\tilde{\chi}_1^0$ mass. This is 100 GeV higher than achieved with high p_{T}^{γ} analysis [16] in small $\tilde{\chi}_1^0$ mass. However, for high $m_{\tilde{\chi}_1^0}$, high p_{T}^{γ} (with >100 GeV) analysis excludes more as it used N_{jets} as search variables.

With these expected exclusions, this analysis may provide the best limits obtained for masses of gluino and stop quarks with the full Run 2 luminosity.

7.5 Future aspects & outlook

In Section 7.2, the new devised SRs improved the expected exclusion limits for T5bbbbZG, T5qqqqHG for all the $\tilde{\chi}_1^0$ mass regions given $\tilde{g}$ mass as shown in Fig. 7.11. In contrast, for T5ttttZG and T6ttZG models, the improvements are seen in regions with smaller $\tilde{\chi}_1^0$ masses but the exclusion limits are worse for high $\tilde{\chi}_1^0$ masses compared to the SR bins in ($p_{\mathrm{T}}^{\mathrm{miss}}$, N_{jets}, N_{bjets}) with high p_{T}^{γ} >100 GeV as shown in Figure 7.12. The signal model with top quarks in the final states has a high N_{jets} multiplicity and provides more discrimination against background. Also, for electroweakino motivated signal models represents challenging phase space and sensitivity for these models might not come just with devising simple 4D SRs. In high p_{T}^{γ} analysis, excellent sensitivity and up to 1 TeV mass exclusion was obtained by using two dedicated search regions designed based on boosted W and H region. In this section, we explore the possibility of enhancing exclusions for these complex and crucial phase space motivated from several signal scenarios with the help of advanced machine learning techniques such as Boosted Decision Trees (BDTs), which are powerful algorithms commonly used in physics analysis for classification. This physics search uses multiple variables to differentiate the signal from the background in various phase space regions. In HEP analysis, multivariate discriminants use ML methods to improve separation beyond single variables. The BDT method, within the TMVA[175] and ROOT[176] frameworks, offers tools for model tuning, training, evaluation, and analysis, along with a graphical interface for plot generation.

7.5.1 Boosted Decision Trees

A decision tree is a binary classifier that decides whether an event belongs to the signal or background category based on splitting the input data using a set of input variables. The choice of variables is based on their expected discrimination power, often guided by a combination of physics insight and trial and error. The division process divides training events into two subsets, each processed using the same algorithm, until the tree structure is complete. The split identifies the optimal variable and the cut value at each node defined with a purity p representing the fraction of signal events between all events in the node. The child node with $p = 1$ indicates that it contains only signal events and $p = 0$ for the child node with only background events. Beyond this, no further separation is made. Perfect separation is rare because of the overlap of the signal and background distributions. Thus, the split maximized separation, measured by the Gini index $p(1 - p)$ or the cross-entropy, is chosen.
Sequential splits are created, concluding at leaves that signify the final classification of an event when a stop condition is reached, determined by the tree-growing method's adjustable parameters. The hyperparameters include the count of trees (ntrees), the limit of split nodes (maxDepth), and the variable bin quantity (nCuts).

In the repeated learning process known as "boosting [175]," a sequence of algorithms is applied to reweighted training data, giving higher weights to instances that were incorrectly identified in previous iterations. Typically, it teaches several weak learners, like decision trees, in which the training of subsequent learners is influenced by the performance of previous ones. The boosted prediction (y (x)) is mathematically represented by the final model, which aggregates the output of all classifiers.

$$y_{Boost}(x) = \sum_{i}^{N_{DT}} ln(\alpha_i).h_i(x) \tag{7.21}$$

with α_i representing the boost weight from the misclassification rate of the prior tree, h(x) is the output of a classifier, and the sum encompasses all classifiers.
Hence, BDTs enhance accuracy by training successive trees to address prior errors, focusing on misclassified cases. This ensemble method decreases bias and variance, effectively handling complex, high-dimensional datasets commonly used in particle physics for tasks like event classification and background rejection.

7.5.2 Model parameters and its training

An individual BDT with 200 trees and a maxDepth of 2 is trained for each simplified signal model against the combined SM background considered in this model (see Section 2.3.1). Events that satisfy all preliminary baseline selections (except $p_{\mathrm{T}}^{\mathrm{miss}}$ >200) discussed in Section 7.2.1 are passed through the BDT. Both signal and background are treated on the same footing, indicating that no weights representing their respective cross sections were applied to the events. The training and testing data set split is 50%. Following variables were provided to BDT to learn signal and background kinematics;

- Jet multiplicity (N_{jets}),
- Missing transverse momentum ($p_{\mathrm{T}}^{\mathrm{miss}}$),
- Transverse momentum of photon (p_{T}^{γ}),
- Total visible energy (S_{T}),
- B-tagged jets multiplicity (N_{bjets}),
- $d\phi$ between two leading jets and $p_{\mathrm{T}}^{\mathrm{miss}}$ ($d\Phi(p_{\mathrm{T}}^{\mathrm{miss}}, Jets_{1,2}^{lead})$),
- $d\phi$ between $p_{\mathrm{T}}^{\mathrm{miss}}$ and γ ($d\phi(p_{\mathrm{T}}^{\mathrm{miss}}, \gamma)$),
- Transverse mass of γ and $p_{\mathrm{T}}^{\mathrm{miss}}$ ($M_t(p_{\mathrm{T}}^{\mathrm{miss}}, \gamma)$),
- Transverse momentum (p_{T}) of two leading jets
- $d\phi$ between two leading jets and γ ($d\Phi(\gamma, Jets_{1,2}^{lead})$),

Fig. 7.48 shows the shape comparisons of 12 of 13 input variables for the T5ttttZG signal model (blue) against the SM background (red). In fact, these variables are providing a certain phase space where the background and signal discrimination is evident.
Fig. 7.49 shows the 13$\times$ 13 correlation matrix between input variables for the T5ttttZG signal (left) and the background (right). These correlations essentially help BDT to discriminate between signal and background.

7.5.3 Updated exclusion limits with BDT

The effectiveness of a BDT is typically assessed using an ROC curve, which shows background rejection $1 - \epsilon_{\mathrm{bg}}$ versus signal efficiency ϵ_{sig} in a range of 0 to 1. The AUC measures this curve, with larger values signifying improved separation. An AUC of 1 represents ideal discrimination, while an AUC of 0.5 suggests randomness.

In this study, we used the mass region excluded by BDTs as a performance matrix. Once trained, BDT provides a score for each event based on its kinematic properties defined by input variables. Positive or close to 1 score, then event is classified as a signal, and if close to -1 or less than zero, then background. Fig. 7.50 shows the distribution of the BDT scores of different BDTs trained on same SM backgrounds but with different simplified signal models; T5ttttZG (top left), T6tttZG (top right), TChiWG (bottom left) and TChiNG (bottom right). In these plots, the signal and background events are weighted by their respective cross section and full Run 2 luminosity. Signal models derived from strong production mechanisms are proficiently identified as signal events by the BDT due to the markedly distinct profiles of the input kinematic

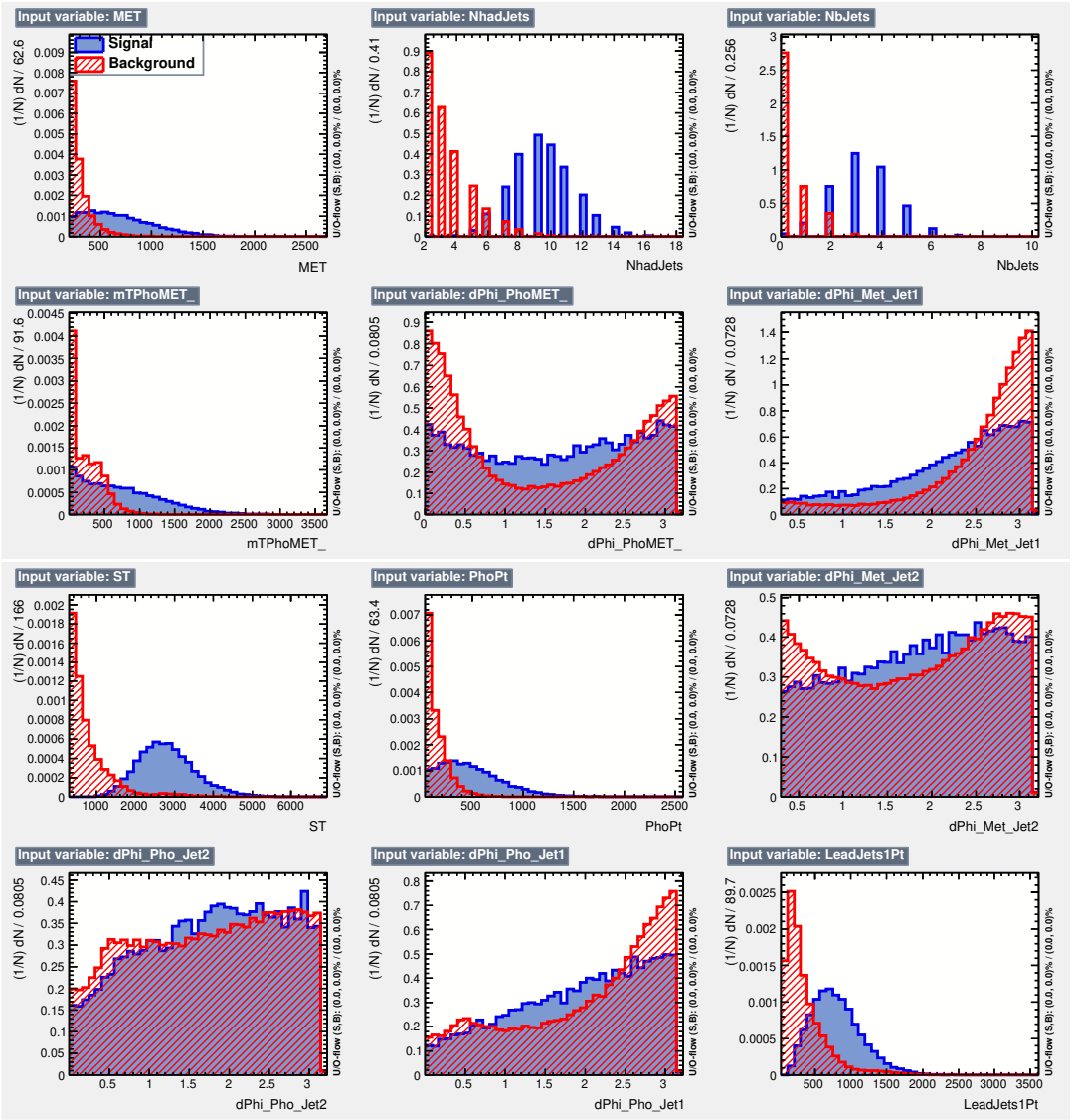


Figure 7.48: The shape comparisons of 12 out of 13 input variables for the T5ttttZG signal model (blue) against the SM background (red).

variables, as demonstrated for certain cases in Sec 7.2. Conversely, for signals inspired by electroweakino processes, such as TChiWG and TChiNG, the differentiation is less pronounced, thereby limiting the BDT's capacity to readily classify these events as signals. For the studies presented in this section, we have opted for rejecting events with a BDT score < 0.

Fig. 7.51 shows the impact of the BDT score criteria (>0) on the exclusion limits obtained with new SRs (defined in Sec. 7.2 for T5bbbbZG (left) and T5qqqqHG (right). The performance with and without BDT criteria is very similar for both models. This is attributed to the fact that new SRs($p_{\mathrm{T}}^{\mathrm{miss}}$, S_{T},N_{bjets} and p_{T}^{γ}) is already able to achieve good sensitivity given the distinct shapes of these variables in signal as compared to background as shown in Fig. 7.6.

Fig. 7.52 demonstrate the impact of using a selection on the BDT score, on the exclusion limits obtained for T5ttttZG (left) and T6ttZG (right). In case of T5ttttZG, using

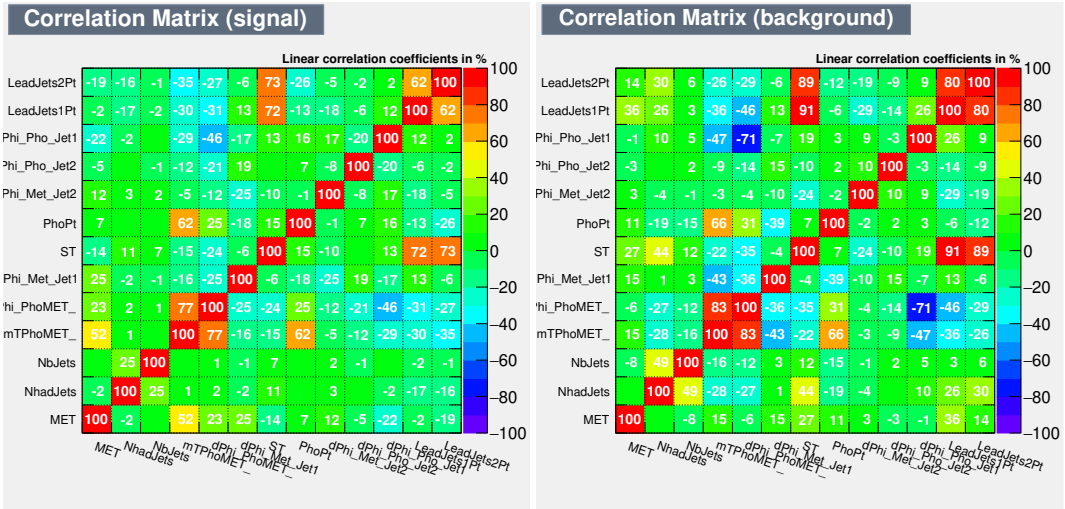


Figure 7.49: Correlation among input variables for signal (left) and background (right).

a BDT helps to improve exclusion in high $\tilde{\chi}_1^0$ masses. However, in case of T6ttZG, with the same SR defined based on ($p_{\mathrm{T}}^{\mathrm{miss}}$, S_{T}, N_{bjets} and p_{T}^{γ}) does not help in high masses $\tilde{\chi}_1^0$ even using BDT. A new set of SR bins very similar to SP bins from high p_{T}^{γ} analysis but dividing events further based on S_{T} and p_{T}^{γ} are tried. The green curve on the right of Fig. 7.52 shows the exclusion limits obtained for these new SRs and BDT. Inclusion of N_{jets} actually helps the limit improve in the high $\tilde{\chi}_1^0$ mass region as illustrated by green curve in Fig. 7.52 (right).

For electroweakino models, we have picked 50 SRs (out of new 100 SRs with merging p_{T}^{γ} regions) designed based on $p_{\mathrm{T}}^{\mathrm{miss}}$, N_{bjets} and S_{T}. Fig. 7.53 shows the distributions of signal and background events as a function of these SRs for TChiWG (left) and TChiNG (right).

Fig. 7.54 shows the comparison of exclusion limits obtained with high p_{T}^{γ} analysis (black curve), for p_{T}^{γ} >40 (blue curve) using SP bins defined in Section 7.2. The yellow solid and dotted curves represent the exclusion limits obtained in 50 SRs without and with a BDT score of >0, respectively. With the help of BDT, we observe that higher mass can be probed, resulting in tighter upper limits on electroweakino production cross section.

7.6 Summary

The phase space involving soft photons presents both intriguing opportunities and significant challenges. It is demonstrated that revisiting signal and background kinematics and by introducing new search variables to divide search regions, the analysis achieves better exclusion limits compared to previous studies. Background estimation has been conducted thoroughly, and predicted exclusion limits for specific mass ranges are obtained. We are able to probe 200 GeV more gluino masses with small $\tilde{\chi}_1^0$ masses in T5bbbbZG and T5qqqqHG simplified models with the help of S_{T} and p_{T}^{γ} based SRs. The gluino mass exclusion for T5ttttZG and T6ttZG is improved by 100 GeV for small $\tilde{\chi}_1^0$ masses. Future work will focus on incorporating a Boosted Decision Tree (BDT) analysis to enhance the sensitivity of the current search regions. Optimizing the BDT for electroweak (EW) signals is expected to achieve comparable or superior exclusion limits. Additionally, the photon transverse momentum

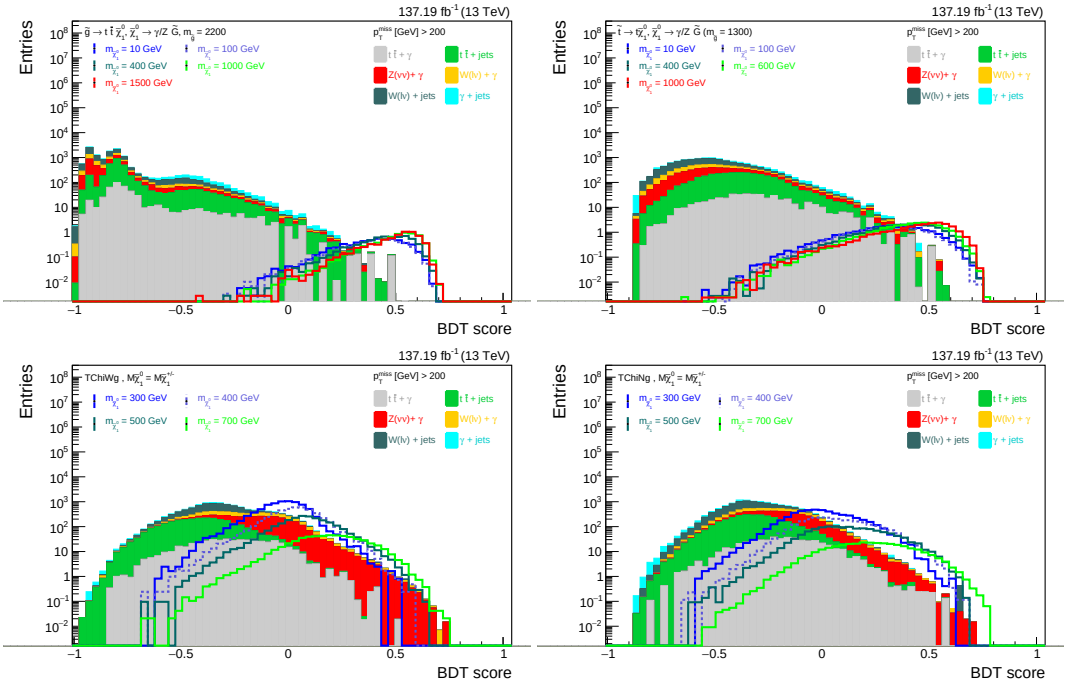


Figure 7.50: BDT score distribution in different BDTs trained on SM backgrounds and different simplified signal models: T5ttttZG (top left), T6tttZG (top right), TChiWG (bottom left) and TChiNG (bottom right). Signal and background events are cross section weighted and scaled to full Run 2 luminosity.

threshold can be lowered further, enabling a more comprehensive exploration of the parameter space. Efforts will also concentrate on developing and optimizing a unified BDT that performs effectively across all signal models, further enhancing the analysis' sensitivity. While the simplified SUSY models serve as effective benchmarks to illustrate specific final states and guide search strategies, they do not capture the full complexity of realistic SUSY scenarios, including variations in production modes, mass spectra, and decay chains. As such, the interpretation of results remains model-dependent and limited in generality. The analysis provides a foundation that can be extended to more complete models in future work.

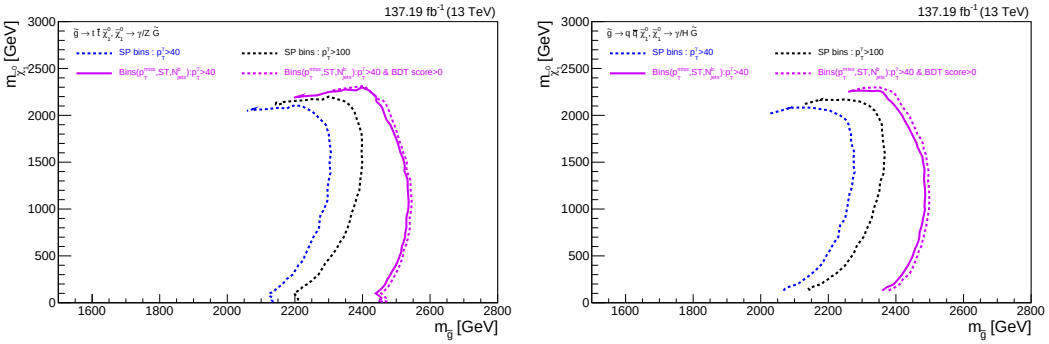


Figure 7.51: The expected exclusion curve, represented by the dotted magenta line for including events with BDT score (> 0) and the solid magenta line for those without such preselection, is shown for T5bbbbZG (left) and T5qqqqHG (right) for comparison.

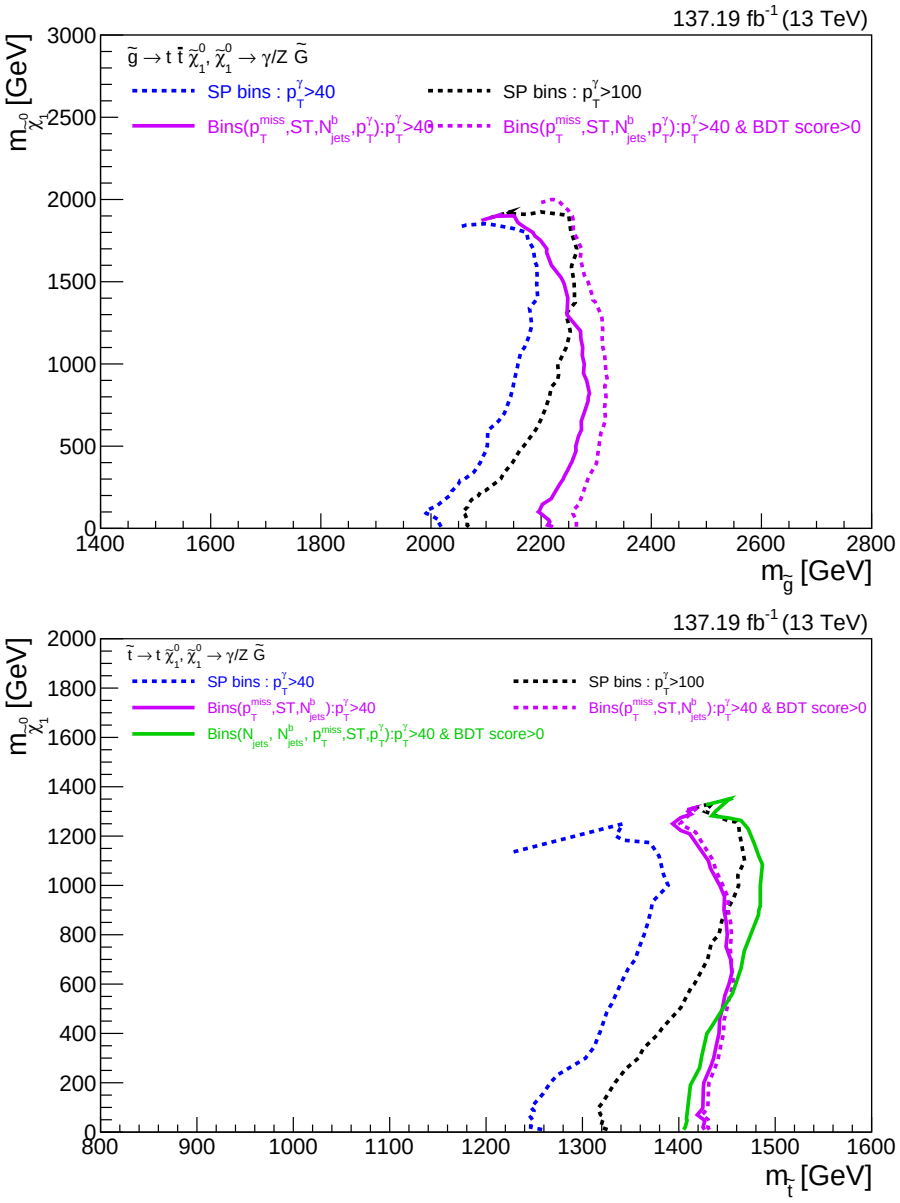


Figure 7.52: The expected exclusion curve, represented by the dotted magenta line for including events with BDT score (> 0) and the solid magenta line for those without such preselection, is shown for T5ttttZG (top) and T6ttZG (bottom) for comparison. Additionally, limits obtained with using N_{jets} as one of search variables is shown in green curve for T6ttZG.

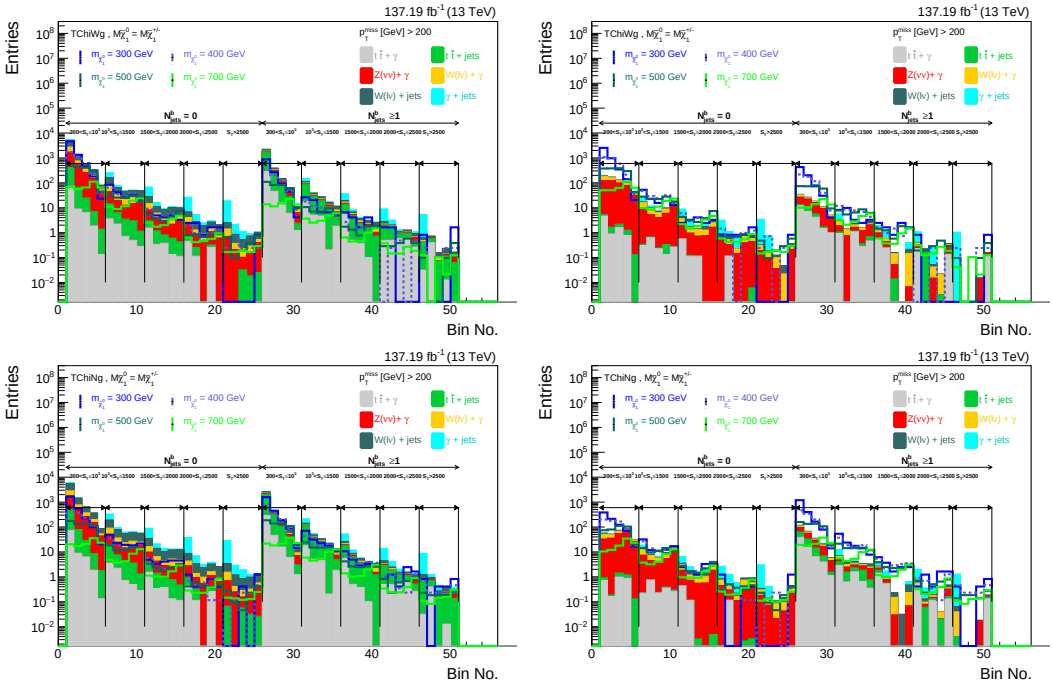


Figure 7.53: Distributions of signal and background events in SRs defined in ($p_{\mathrm{T}}^{\mathrm{miss}}$, S_{T} and N_{bjets}) without BDT (left) and with BDT preselection (right) shown for TChiWG (top) and TChiNG (bottom).

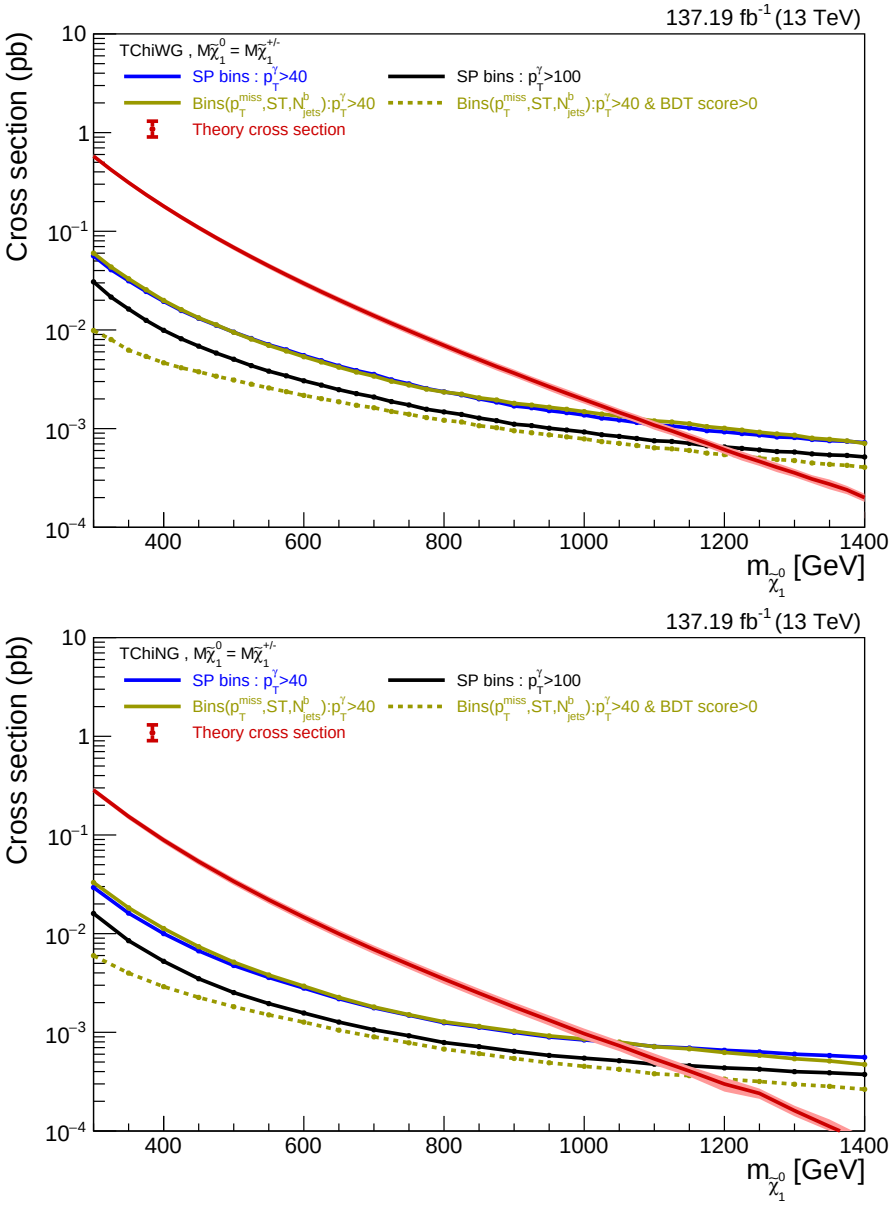


Figure 7.54: The dotted magenta line for events with BDT score (> 0) and the solid magenta line without preselection illustrate the expected exclusion curve for TChiWG (top) and TChiNG (bottom).

Chapter 8

Summary and outlook

At the end of Run 3, it is projected that the CMS has accumulated $> 450\ fb^{-1}$ of data. The LHC is set to enter its high luminosity (HL-LHC) phase, starting in 2029, with the goal of accumulating 3000 fb^{-1} over ten years, enabling detailed physics studies and the potential discovery of new physics. The CMS is preparing for the HL-LHC with comprehensive detector upgrades, including the complete replacement of certain existing detectors and their associated readout systems. To address high radiation doses and increased particle densities and to effectively trigger and reconstruct VBF jets and merged jets during HL-LHC, current CMS endcap calorimeters will be replaced by the High-Granularity Calorimeter (HGCAL), a state-of-the-art detector that features six million read-out channels capable of providing highly detailed 5D measurements: energy, position, and timing. For the first time silicon has been used in the calorimeters at such a large scale. The calorimeter's large-scale imaging requires significant R&D, component testing, and efficient management of the substantial data rates from the HGCAL. Furthermore, understanding the response of HGCAL to different particles derives their precise reconstruction utilizing the highly granular information from the HGCAL represents a substantial challenge to undertake.

In the CMS experiment at the LHC, the precise reconstruction of hadronic showers in calorimeters is key to jet reconstruction. The precision in measuring jet energy affects the precision of the $p_{\mathrm{T}}^{\mathrm{miss}}$ measurements. Many new physics searches, especially those inspired by SUSY, rely on jet and $p_{\mathrm{T}}^{\mathrm{miss}}$ measurements. Improving calorimeter performance to single particle showers will improve reconstruction of the jet and $p_{\mathrm{T}}^{\mathrm{miss}}$, leading to better sensitivity to probe difficult areas of phase space. In contemplation of these long-term objectives, this thesis outlines methodologies designed to enhance the precision of forthcoming particle physics experiments.

Hadronic showers consist of large event-to-event fluctuations resulting from the production of neutral pions instantaneously decaying into photons, forming an EM shower (Electromagnetic (EM) component). This leads to a poorer calorimeter resolution for hadronic showers. Chapter 5 demonstrated the capabilities of a novel dynamic reduction network (DRN) based on graphical neural networks (GNN) for reconstructing hadron showers from charged pions in the HGCAL beam test prototype commissioned for 2018 beam test conducted at CERN. The performance of this method is compared against a classical weight summation (WS) method making use of χ^2-minimization. This classical method is a part of the PF hadron calibration in the CMS. Energy resolution of charged pions is improved by twofold upon using DRN as compared to WS method as DRN is able to capture & compensate for event-by-event fluctuations of hadronic showers using high granular information from the detector. The performance of DRN is also inferred on the real data corresponding to

charged pions collected with HGCAL beam test and it is found to be similar to its performance on simulation. This highlight the robustness of the DRN to minimize the data-simulation differences.

To understand these improvements, the energy deposited in the absorber materials representing fluctuations in energy deposit patterns, and energy carried by secondary neutral pions representing EM component of hadronic showers, are extracted for each events using truth information from GEANT4. On inspecting DRN's performance as a function of these truth information, it is observed that DRN is able to compensate these fluctuations in energy deposit patterns essentially reflecting fluctuations in EM component of hadronic showers. DRN's performance is unchanged with respect to the energy calibration unit used to train it as long as spatial coordinates of the rechits are provided. By leveraging transverse and longitudinal spatial information alongside signal amplitudes, the DRN simplifies the detailed rechit energy calibration for over 6 M readout channels.
DRN's performance degrades by 15-20% upon using CLUE rechits clusters reducing granularity up to 50%, instead of full rechit usage. Consequently, despite potential inaccuracies in clustering and associating the whole of the energy from particle showers, DRN is still capable of providing enhanced performance.

Simulations with the full endcap geometry further validated the DRN's potential to enhance energy resolution, marking a significant milestone as the first machine learning-driven energy reconstruction algorithm applied to hadron showers in the CMS HGCAL prototype. With the full endcap geometry, a stochastic term of 53% is obtained. This is just a tip of the iceberg and full potential of DRN in terms of hadron calibration in CMS for HL-LHC is yet to be seen.

Additionally, DRN-based end-to-end calibration of electromagnetic and hadronic showers in the final CMS detector—already showing promise in barrel calorimeters. Future prospects involve the extension of the DRN's application to the reconstruction of EM and hadronic showers-eventually jets within the CMS at the HL-LHC. In doing so, the approach aims to integrate HGCAL's capabilities with the comprehensive granular readout provided by the barrel calorimeters during HL-LHC. A simplified version of the DRN can be integrated at the trigger level to enhance jet performance and ultimately improve the efficiency of selecting desirable events. With these, DRN can eventually become the key to improve the jet energy & $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution at the complex environment of the hadron colliders.

Chapter 6 presents a detailed summary of the HGCAL readout electronics system consisting of different ASICs in different places across the full chain. These ASICs design underwent rigorous validation through various methods and lab tests. As the HGCAL approaches its pre-production phase, it is imperative to evaluate the efficacy of the finalised design, with a particular focus on the readout components, via both system and beam tests to guarantee operational integrity and seamless data transmission without loss. I participated in one of the first system integration tests in September 2023, playing a key role in the development of the real-time raw data analysis framework and collaborating with DAQ experts to resolve readout chain errors, if any.

To determine whether the final read-out system can effectively manage HGCAL data rates, we developed a C^{++} based emulator, which simulates buffers in various

ASICs of the read-out design and integrates the mapping of front-end (FE) to back-end (BE) optical links. Its main goal was to estimate BE data rates from FE and assess buffer performance under high data rate conditions. On average, no data loss was found, but rare cases demanded extensive statistical analysis. A method to create high-statistics pseudo-events using a multivariate Gaussian distribution and quantile transformation to maintain complex correlations among detector modules, utilizing scipy and numpy was developed. Managing the correlation matrix and its matrix multiplications, due to their $O(10^4)$ dimensions, presented challenges that were addressed by optimizing RAM.
Irretrievable data loss in the readout system can result in false MET signatures, especially if the loss is uniform across detector modules in one direction. To address this problem, optimization of when the BE should stop accepting triggers to avoid data losses is performed and summarized in Chapter 6. BE can stop the incoming trigger based on the ASIC buffer status, affecting the overall CMS read-out process. The emulator is run at increased FE data rates to assess maximum data losses (buffer overflow) and dead time (halted triggers). Such studies confirmed the system's resilience to high occupancies without significant data loss, and layer-wise validation ensured optimal mapping between ECON-D and lpGBT. Despite challenges like ECON-D truncation at high data rates and HGCROC throttling above 900 kHz trigger rates, no data loss was observed at the design rate, and the system's dead time remained manageable.
Looking ahead, these developments establish a solid foundation for minimizing data losses and mitigating false MET signatures. DRN-like models could further recover information lost in the ECON-D, enhancing data quality and robustness. Together, these efforts provide a comprehensive framework for the HGCAL readout system's reliability, improvements in its reconstruction.

Although no direct SUSY evidence has been discovered, some complex phase-space regions are still unexplored, such as SUSY signatures with soft photons. New scenarios like pMSSM with wino & bino like LSP production, as well as radiative decays of the NLSP, produce soft photons. Chapter 7 explores SUSY searches based on the pMSSM and GMSB motivated simplified models in challenging phase-space regions characterized by soft photons, focusing on improving exclusion limits and probing previously unexplored scenarios. The investigation defines the search region with at least one photon momentum (p_T) over 40 GeV, MET over 200 GeV, and multiple jets. Lowering the photon momentum threshold increases the background tenfold and reduces signal sensitivity when using the jet multiplicity and MET as primary criteria based on previous searches. By reassessing events with photon momentum and the scalar sum of all jets and photon momentum, signal sensitivity improves across strong & electroweakino, and pMSSM scenarios. Major SM backgrounds are estimated using data driven methodologies and validated in this phase space, and prediction in data are estimated along with some systematic measurements and data-driven limit predictions in place. For the T5bbbbZG and T5qqqqHG simplified models , we are able to increase sensitivity by 200 GeV as compared to the published results. A comprehensive methodology has been developed and tested to lower the photon's momentum further up to 20 GeV using boosted decision trees (BDT).
Future efforts will target comprehensive searches with Run-3 data, incorporating more stringent BDT-based optimizations for electroweak (EW) signals is expected to achieve tighter exclusion limits. Additionally, the photon transverse momentum threshold can be lowered further, enabling a more comprehensive exploration of the parameter space. Efforts will also concentrate on developing and optimizing a

unified BDT that performs effectively across all signal models, further enhancing the analysis' sensitivity. At the HL-LHC, the wealth of data and improved precision will enable systematic studies and deeper exploration of electroweak SUSY scenarios.

The pursuit of new physics is ongoing, and with refined tools and methodologies to interpret pp collisions, we can uncover deeper insights into the fundamental particles and their role in shaping our universe. This thesis highlights key areas where future efforts in experimental particle physics are needed, providing insights into the capabilities of state-of-the-art detectors. I wish to inspire readers to explore the exciting future of particle physics and contribute to advancements in precision measurements.

Appendix A

Appendix A - charged pion energy reconstruction supporting

A.1 Reconstruction of charged pions

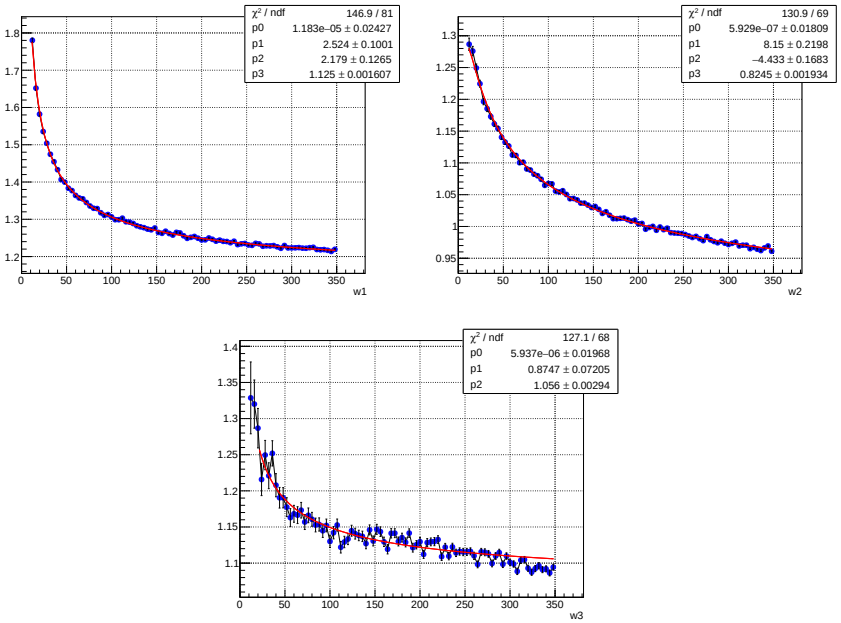


Figure A.1: The energy dependent weights α, β, γ shown on X axis and referred as w1, w2, and w3 respectively, extracted as a function of energy using χ^2 minimization for pions start showering in CE-E (CE-E pions) are shown above. A parametric function has been fitted to extract these weights for any energy values.

This section contains additional plots for testing DRN's performance on data and simulation.

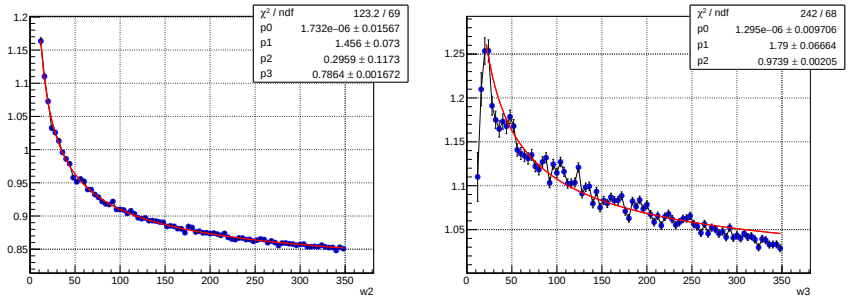


Figure A.2: The energy dependent weights α, β, γ extracted as a function of energy using χ^2 minimization for pions start showering in CE-H (CE-H pions) are shown above. A parametric function has been fitted to extract these weights for any energy values.

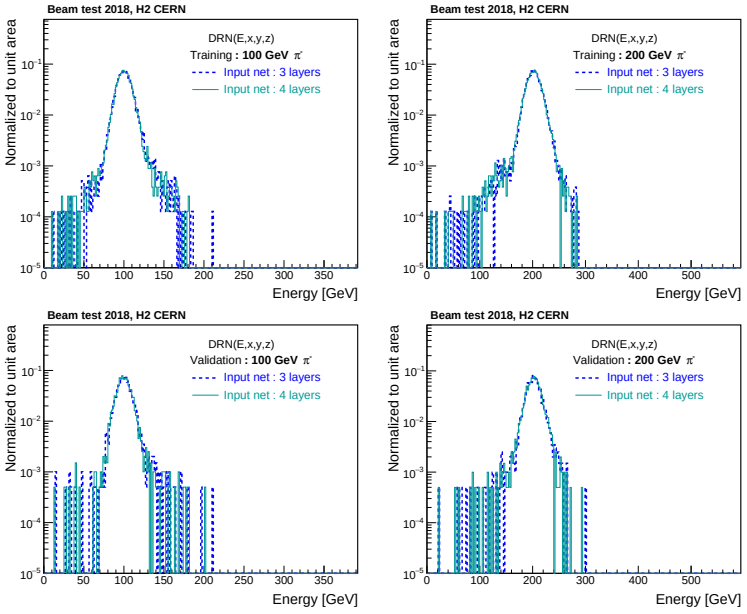


Figure A.3: Comparisons of reconstructed energy distributions for 100 (left) and 200 (right) GeV pions using DRN trained with 4 layers or 3 layers in its input net for training (upper row) and validation (bottom row) datasets.

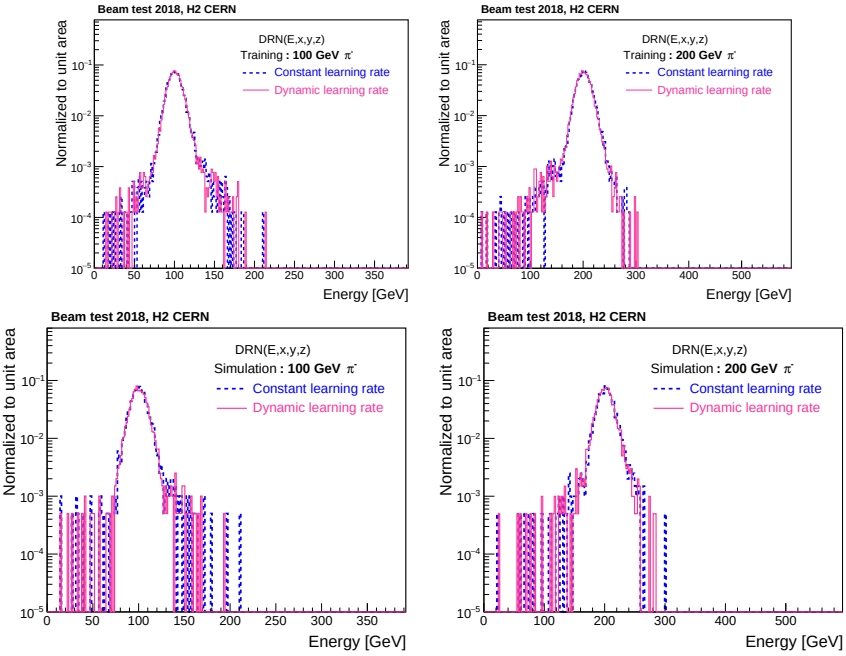


Figure A.4: Comparison of reconstructed energy distributions for 100 (left) and 200 (right) GeV pions with constant vs cyclic learning rates during DRN training in validation dataset.

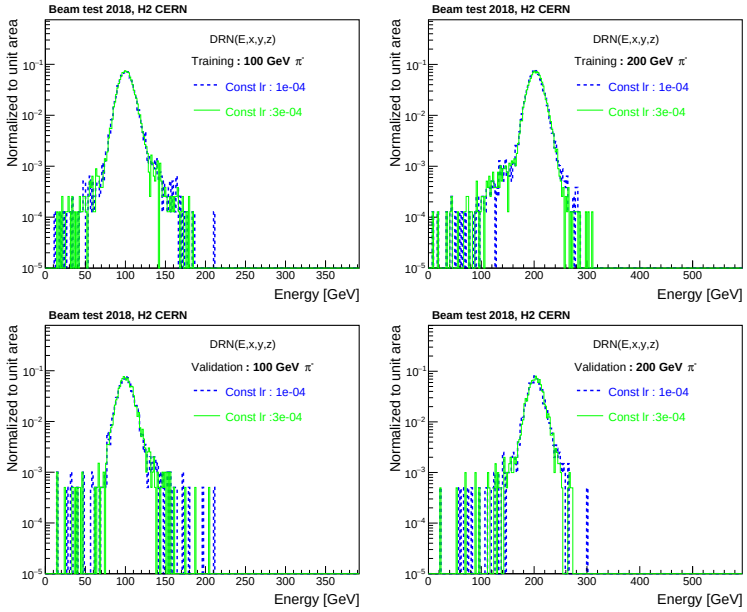


Figure A.5: Reconstructed energy distributions for 100 and 200 GeV pions using different constant learning rates when training DRNs are shown. The plots display results for training (top row) and validation (bottom row) datasets.

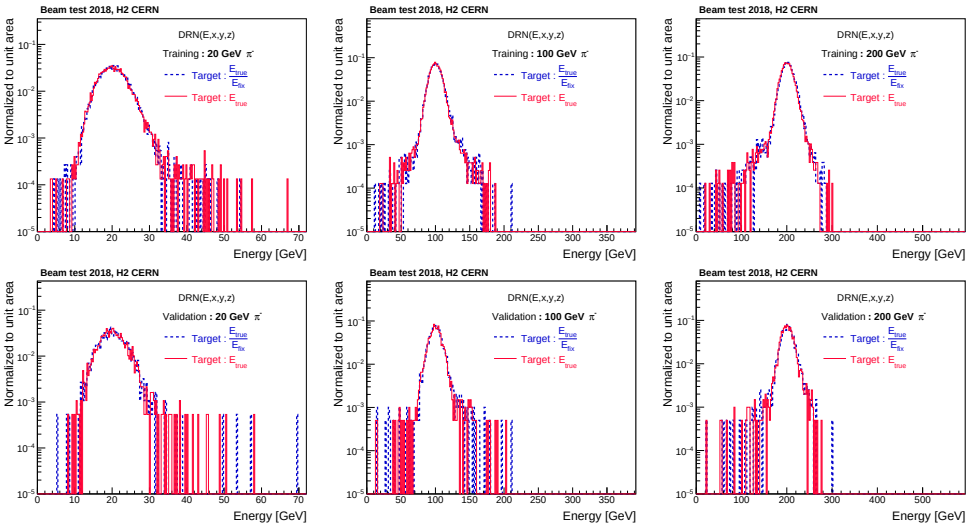


Figure A.6: The reconstructed energy distributions for 20, 100 and 200 GeV pions obtained for training (upper row) and validation (bottom row) datasets with DRN(E,x,y,z) but with different targets used are shown.

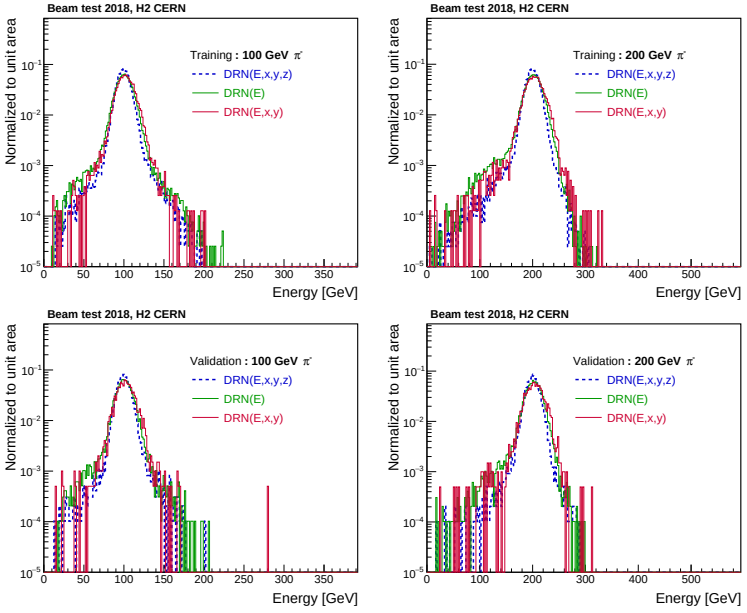


Figure A.7: The reconstructed energy distributions for 100 and 200 GeV pions obtained for training (upper row) and validation (bottom row) datasets with DRN(E,x,y,z), DRN(E) and DRN(E,x,y) are shown.

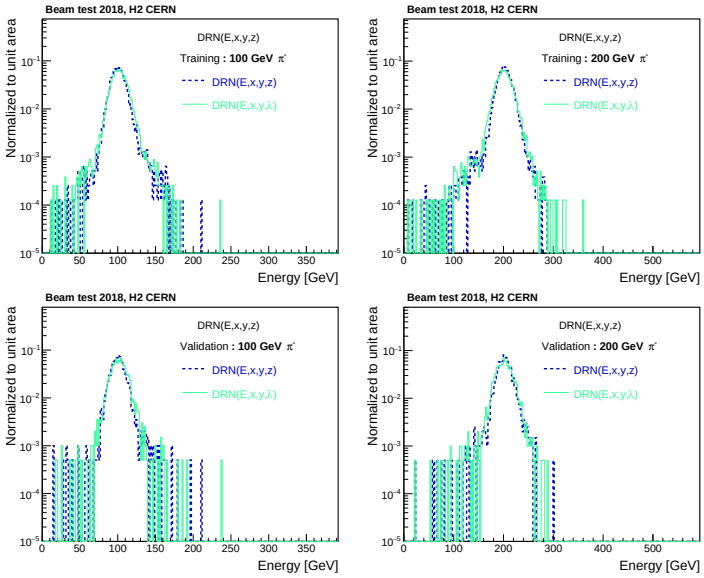


Figure A.8: Reconstructed energy distributions for 100 and 200 GeV pions obtained for training (top) and validation (bottom) datasets with DRN(E,x,y,z) and DRN(E,x,y,λ_{int}) are shown.

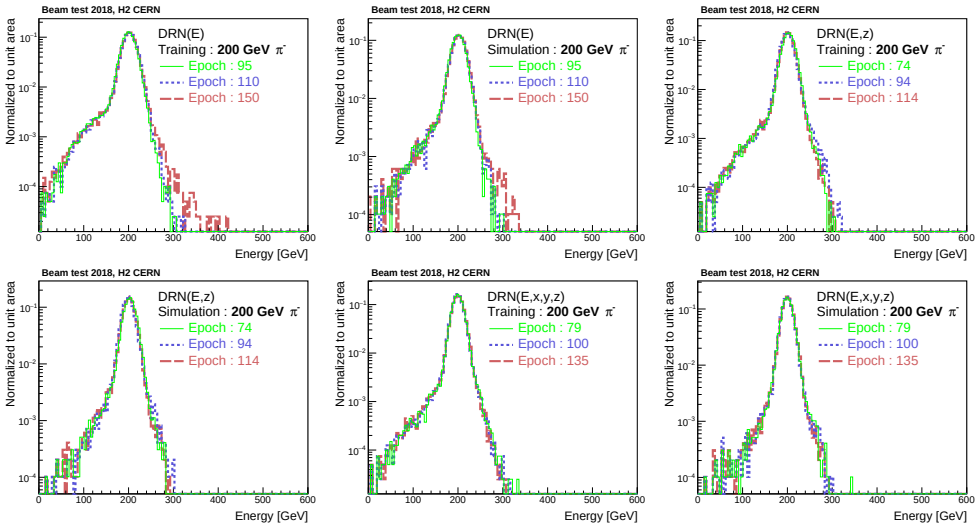


Figure A.9: Reconstructed energy distributions evaluating DRN's performance at different epochs using DRN(E) (top row), DRN (E,z) (middle) and DRN (E,x,y,z) (bottom) on training on left and validation or referred as simulation on right dataset for 100 GeV pions. The shape agrees overall except quite different tails.

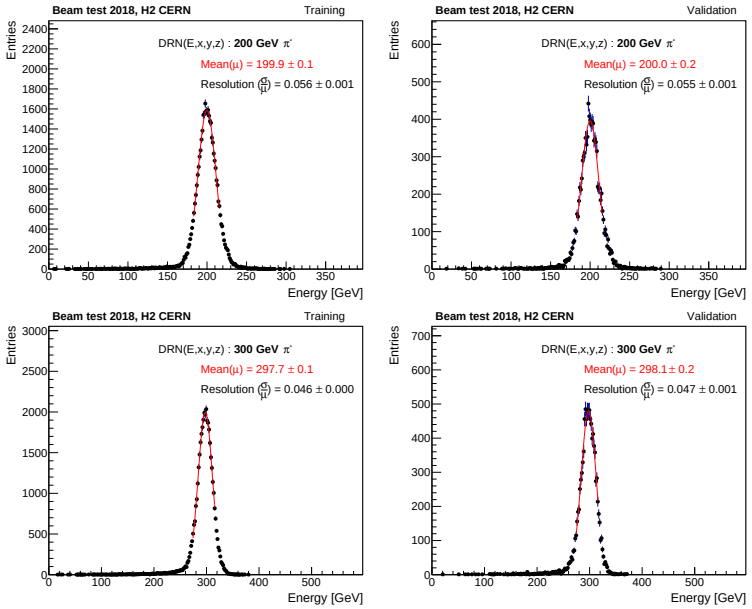


Figure A.10: Energy distributions reconstructed by DRN(E,x,y,z) along with the Gaussian fitted to extract σ and μ, shown in legend.

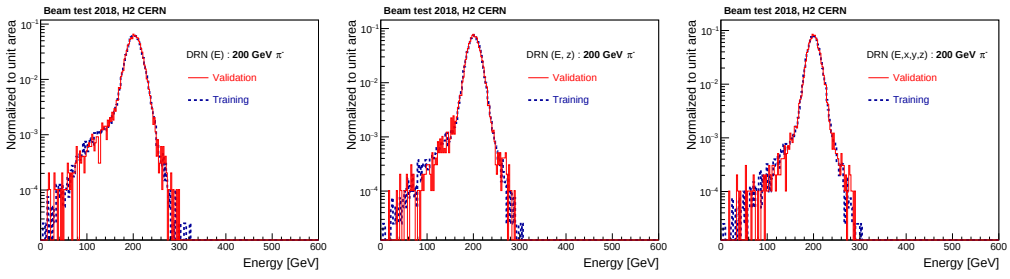


Figure A.11: Comparison of distribution of energy obtained in training and validation datasets for the models DRN (E) (left), DRN (E,z) (middle), and DRN (E,x,y,z) (right) for pions with incident energy of 100 and 200 GeV.

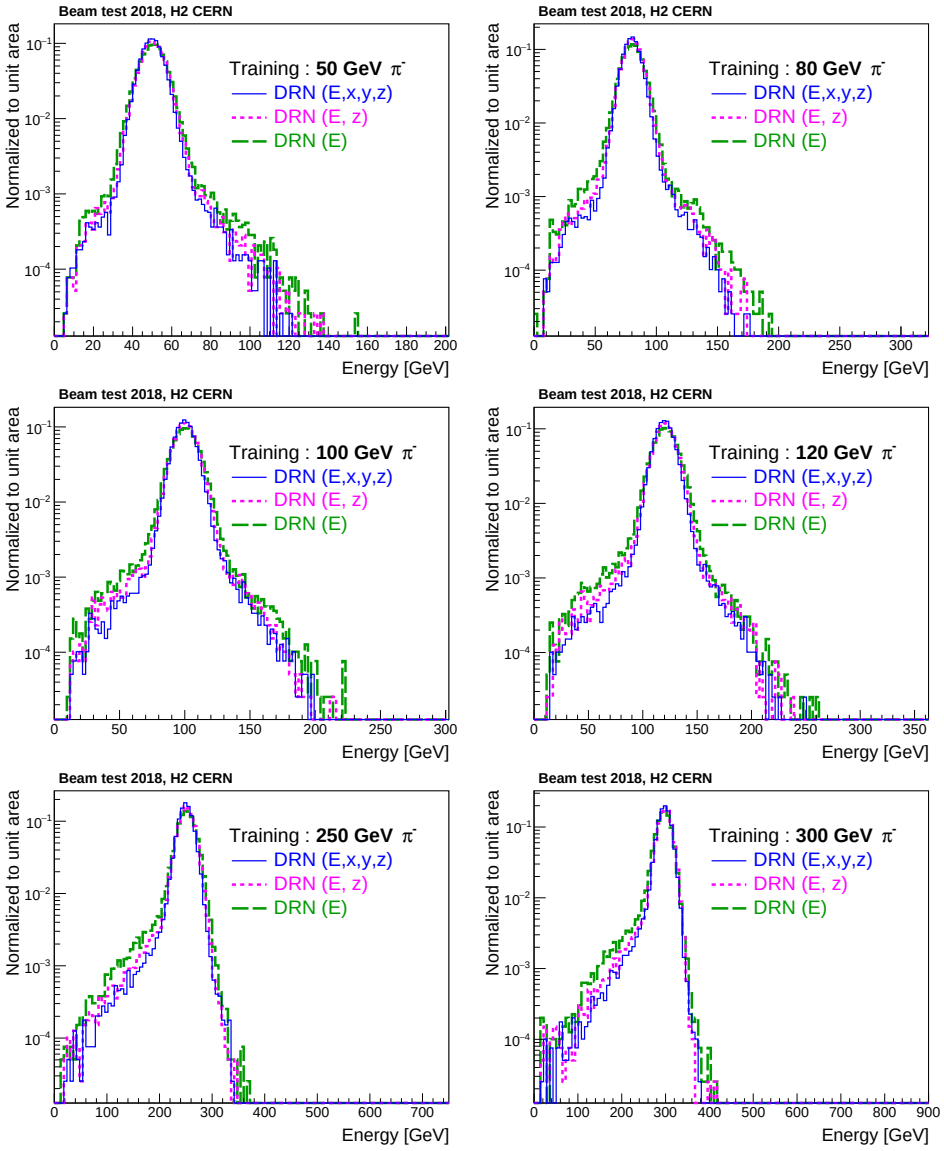


Figure A.12: Distributions of energies predicted using different DRN models for 50, 80, 100, 120, 250 and 300 GeV pions on training datasets.

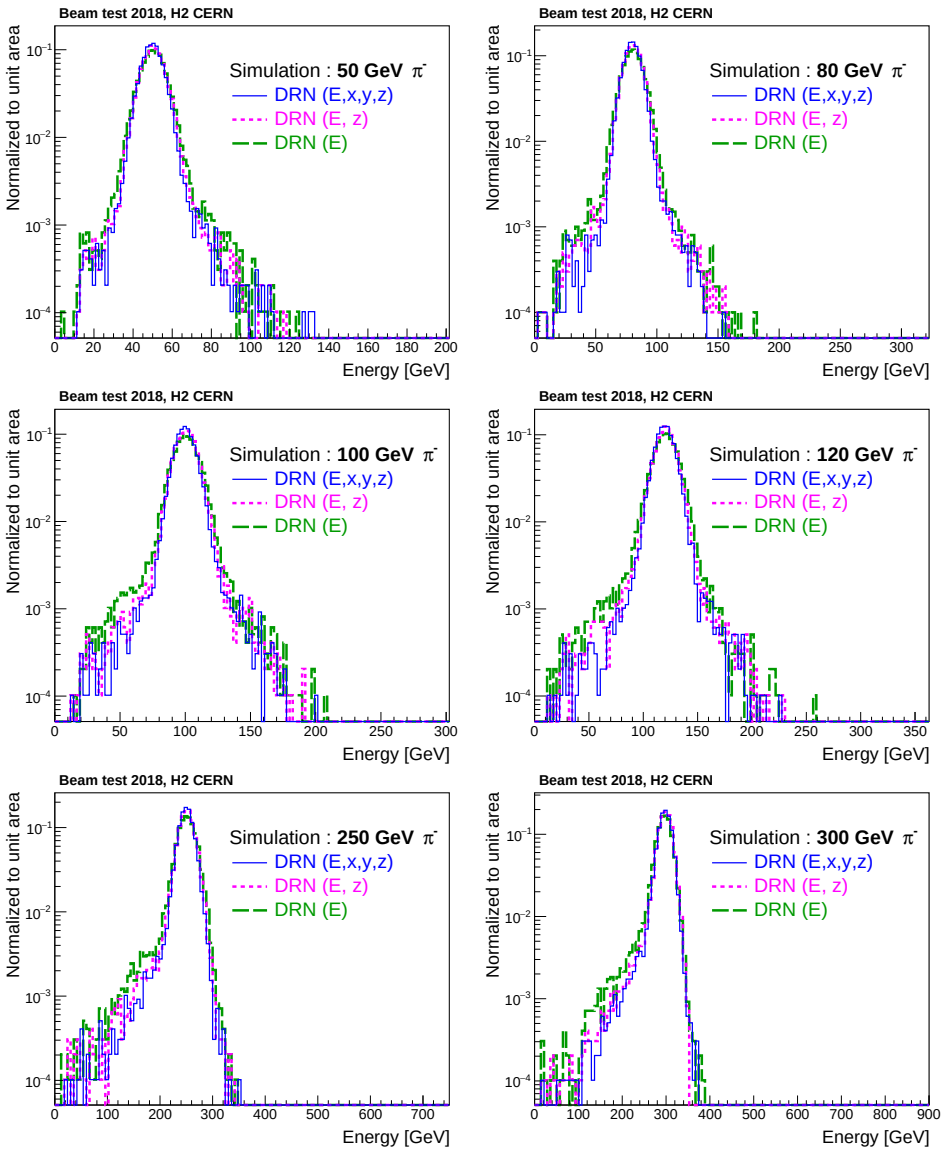


Figure A.13: Distributions of energies predicted using different DRN models for 50, 80, 100, 120, 250 and 300 GeV pions on validation datasets.

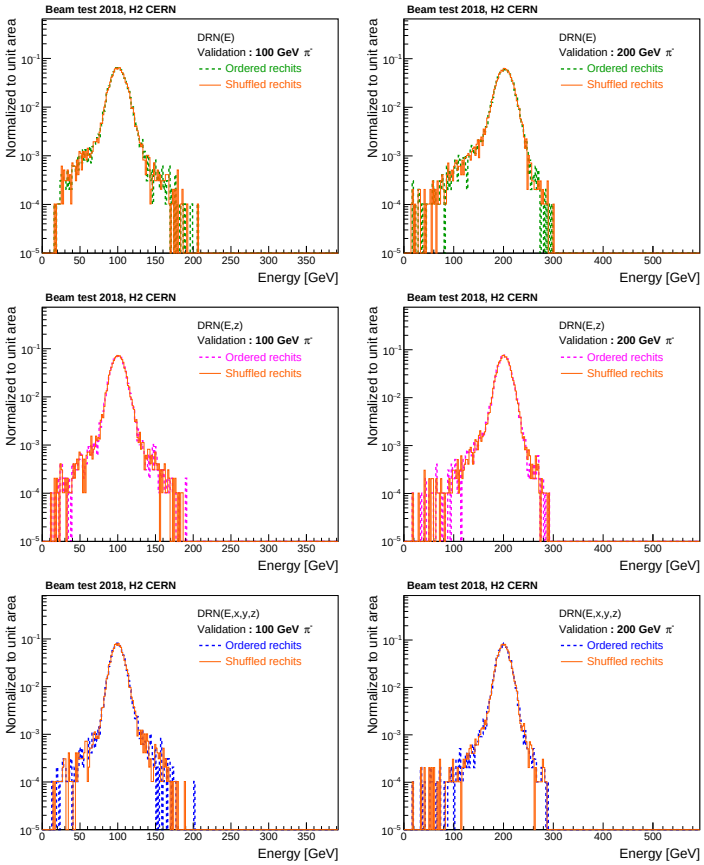


Figure A.14: The reconstructed energy distributions are compared for DRNs inferred on shuffled rechits versus inferred on ordered rechits for 100 and 200 GeV pions in validation dataset. The overall shape comes out to be identical.

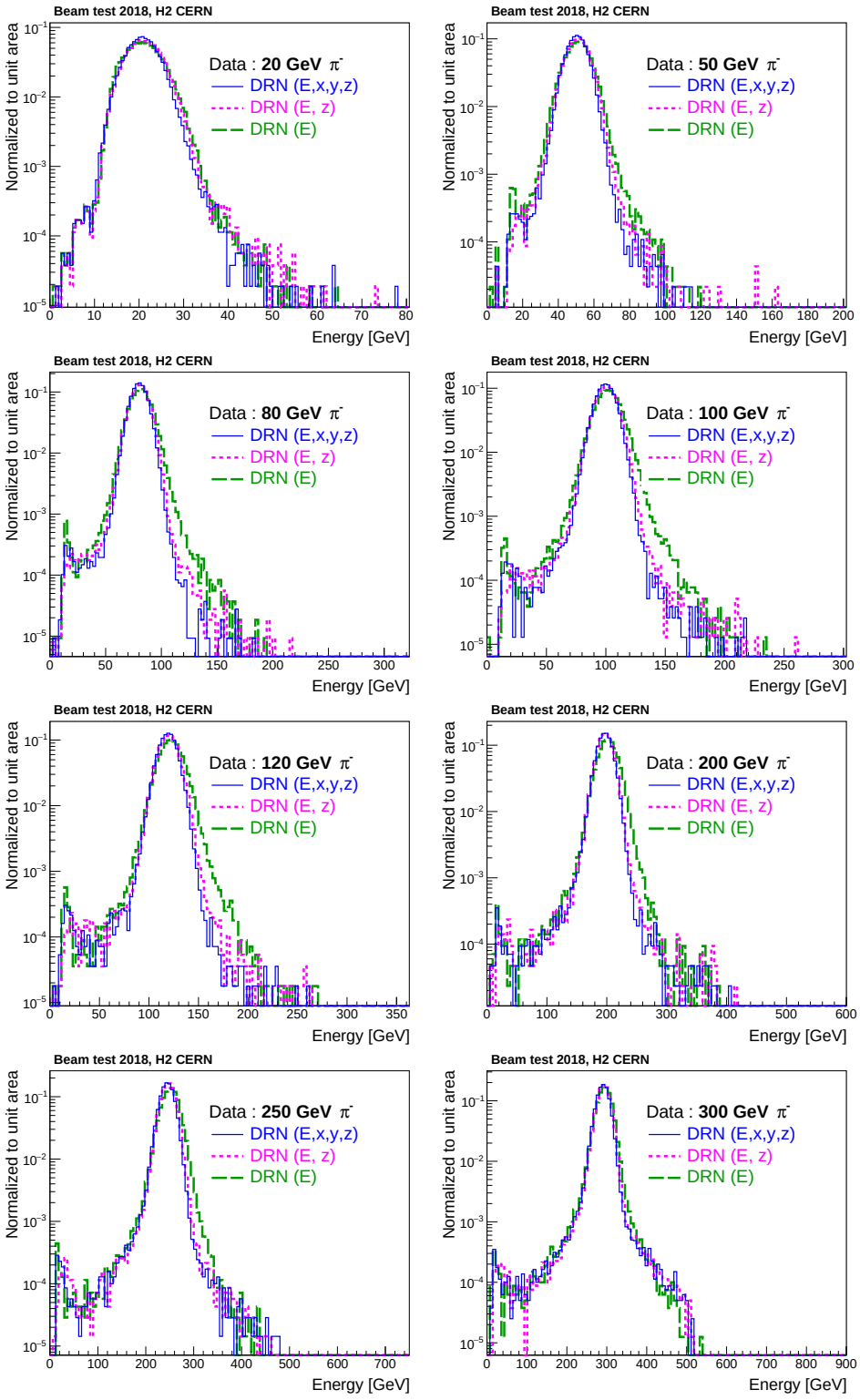


Figure A.15: Distributions of energies predicted using different DRN models for 20, 50, 80, 100, 120, 200, 250 and 300 GeV pions on data recorded in the beam test setup.

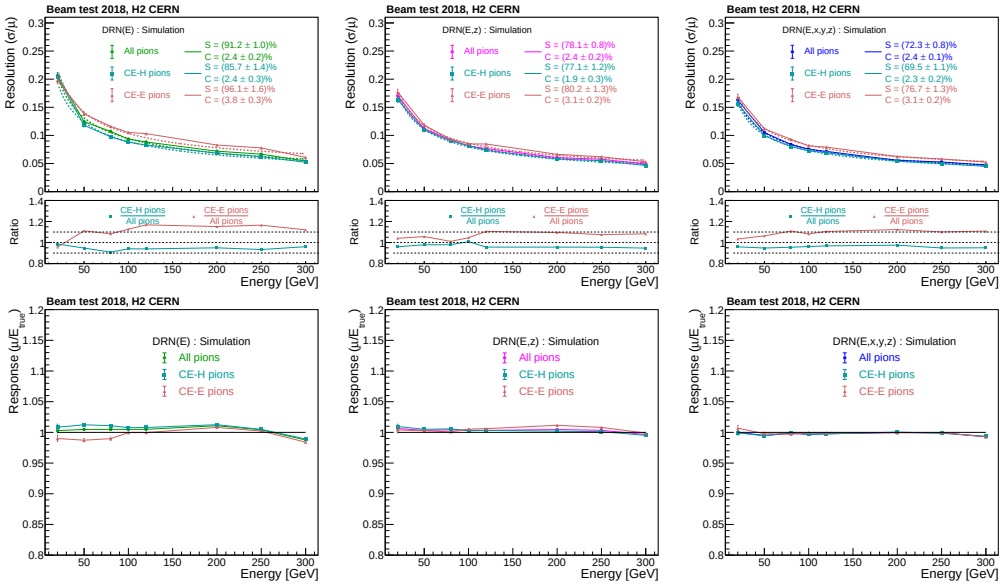


Figure A.16: The energy resolution (top) and response (bottom) for different event categories is compared and shown above for Simulation using DRN(E), DRN(E,z) and DRN(E,x,y,z) from top to bottom row, respectively. The ratio panel shows the ratio of resolution for "CE-E pions" and "CE-H pions" categories to "All pions" category.

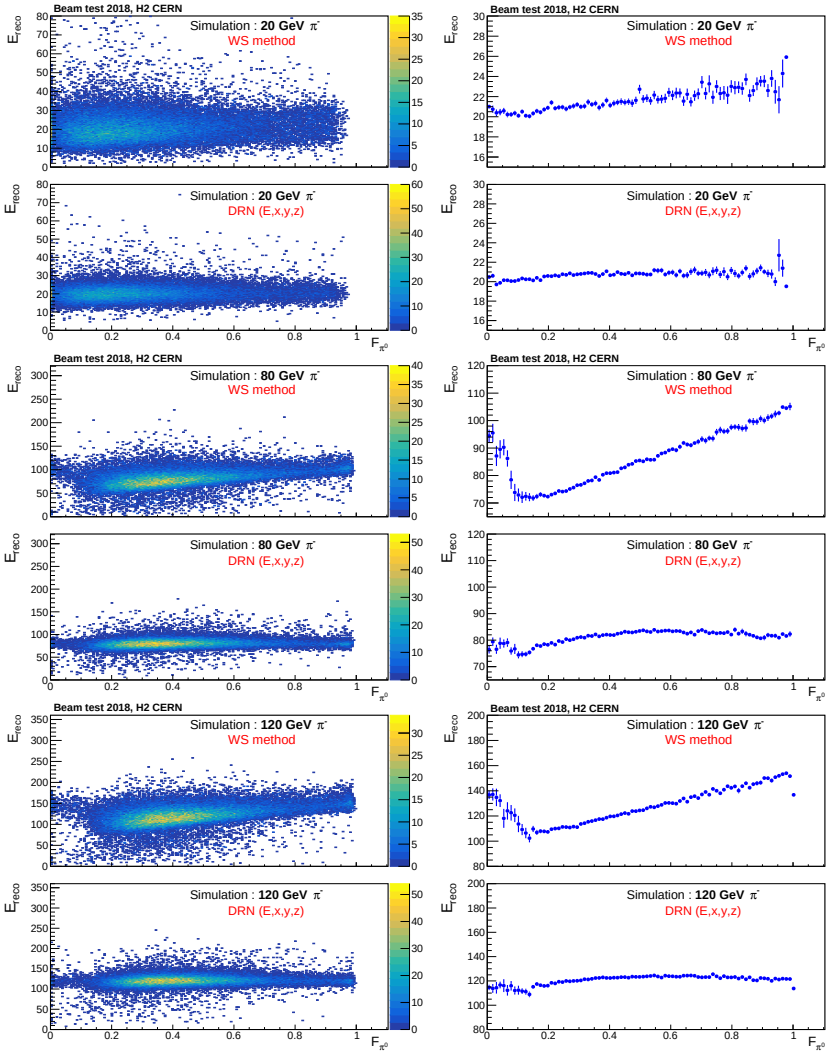


Figure A.17: Energy reconstructed using WS method (top panel) and DRN(E,x,y,z) model (lower panel) for pions of energy 20, 80 and 120 GeV (in the left column) as a function of fraction of energy carried by π^0s in a given shower is shown. The right column shows the profile plots created after calculating the mean of reconstructed energy for different bins in F_{π^0}.

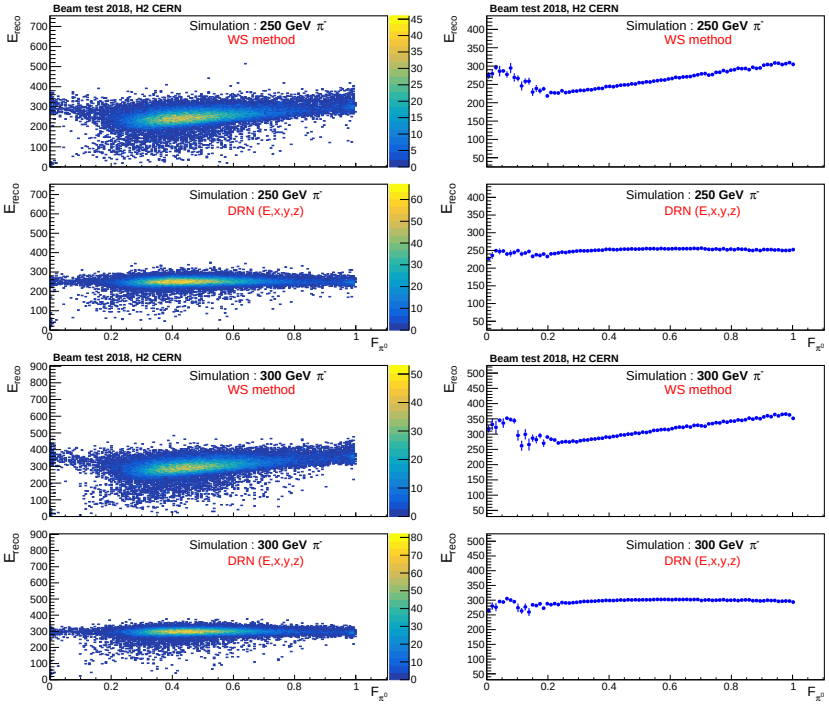


Figure A.18: Energy reconstructed using WS method (top panel) and DRN(E,x,y,z) model (lower panel) for pions of energy 250 and 300 GeV (in the left column) as a function of fraction of energy carried by π^0s in a given shower is shown. The right column shows the profile plots created after calculating the mean of reconstructed energy for different bins in F_{π^0}.

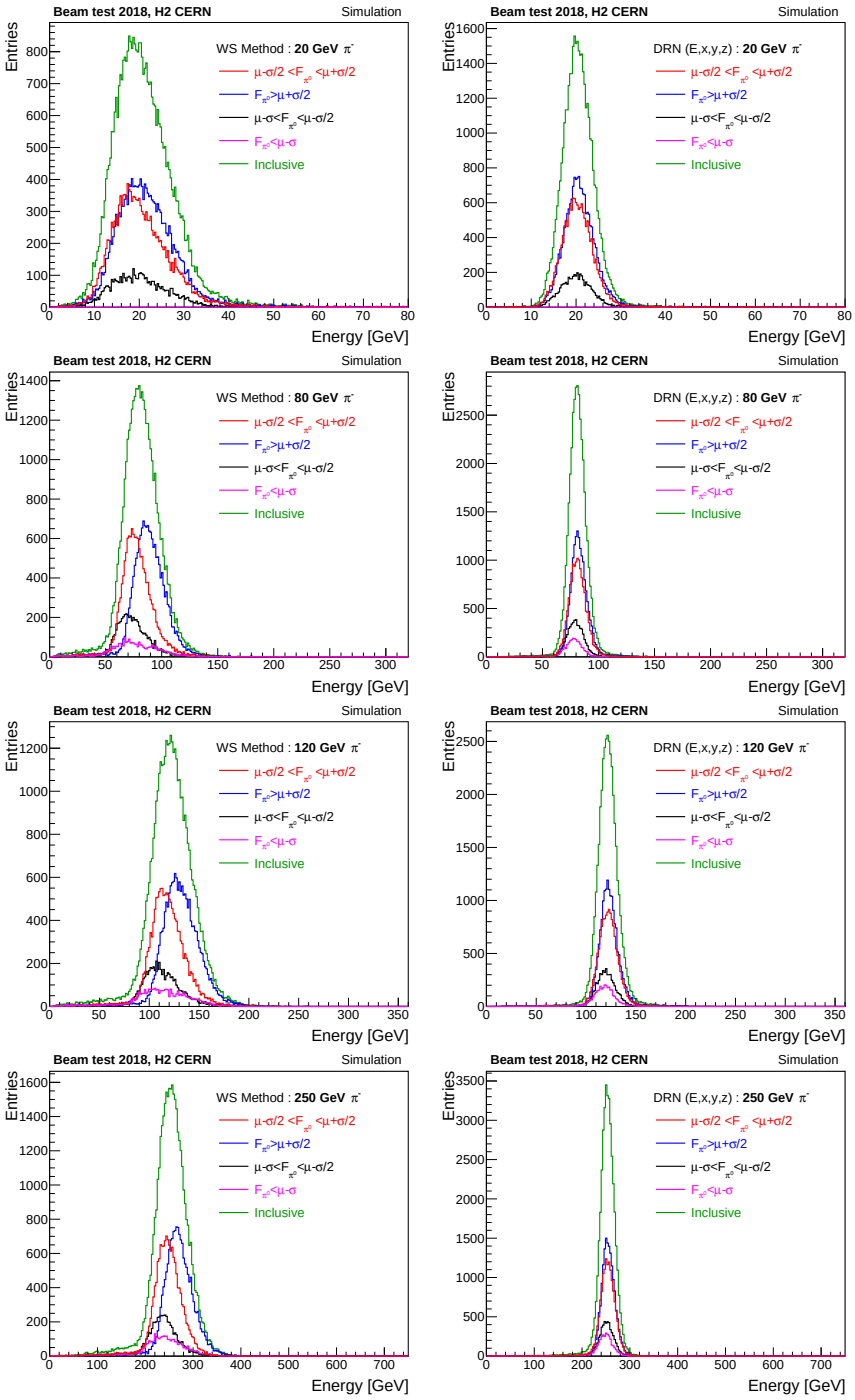


Figure A.19: This figure illustrates performance of DRN and WS method in terms of reconstructing the energy for different events falling into different categories for 20, 80, 120 and 250 GeV pions.

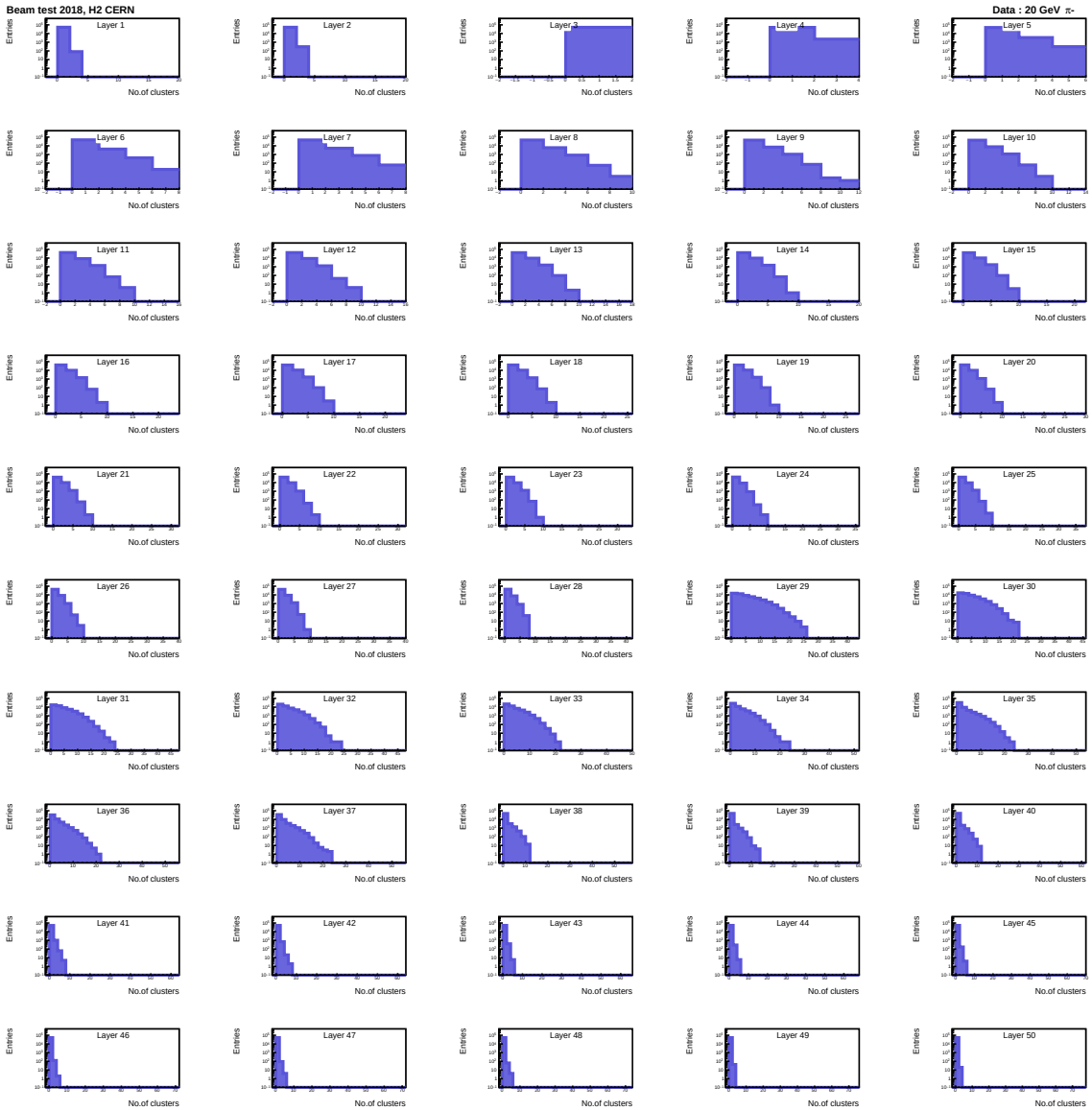


Figure A.20: Number of clusters per layer for 20 GeV pions in beam test data.

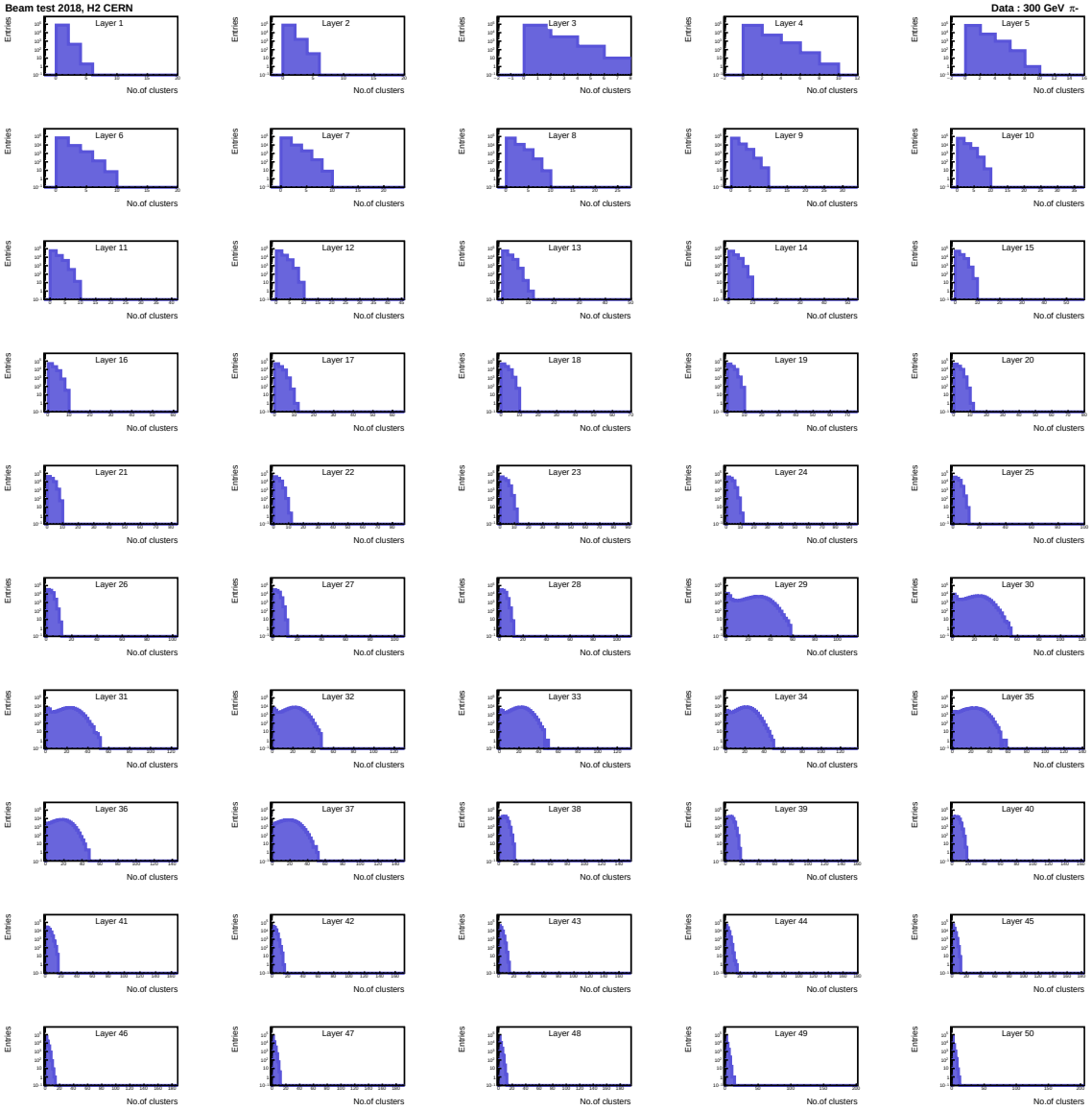


Figure A.21: Number of clusters per layer for 300 GeV pions in beam test data.

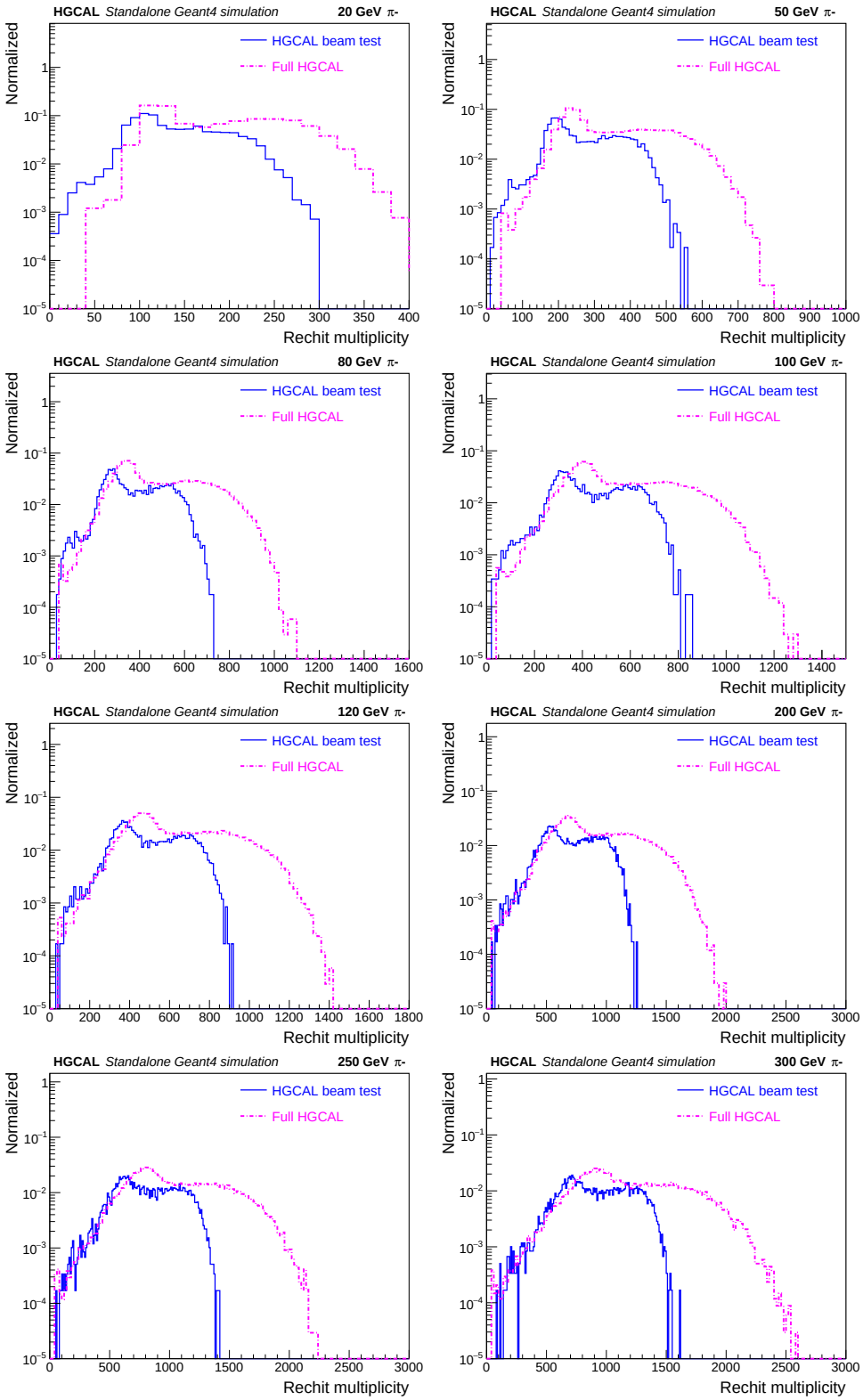


Figure A.22: Comparison of total number of rechits recorded in the full HGCAL geometry vs the beam test geometry for different pion energies. For the full HGCAL, all the rechits within dR cone of 0.5 and passing noise threshold criteria are considered here.

Appendix B

Appendix B - New physics search using photons and $p_{\mathrm{T}}^{\mathrm{miss}}$ in multijet events

B.1 Background estimation

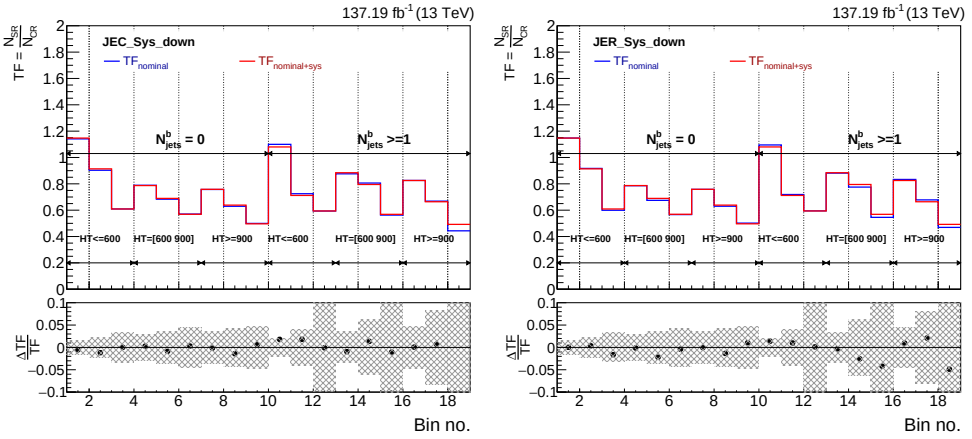


Figure B.1: Top panel: the nominal TF distribution (blue) with the TF (red) obtained with adjusting JEC factors (left) and JER (right) by -1 σ uncertainty. Bottom panel: shaded region represents statistical uncertainty, with black points showing TFs' relative discrepancies defined as the ratio of $(TF_{nominal} - TF_{nominal+sys})$ to $TF_{nominal}$.

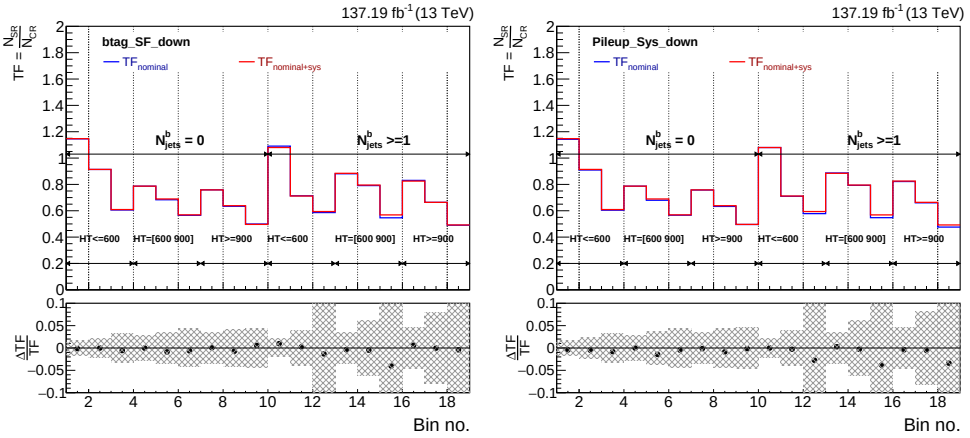


Figure B.2: Top panel: the nominal TF distribution (blue) with the TF (red) obtained with adjusting btag SF (left) and pileup weights (right) by -1 σ uncertainty. Bottom panel: shaded region represents statistical uncertainty, with black points showing TFs' relative discrepancies defined as the ratio of $(TF_{nominal} - TF_{nominal+sys})$ to $TF_{nominal}$.

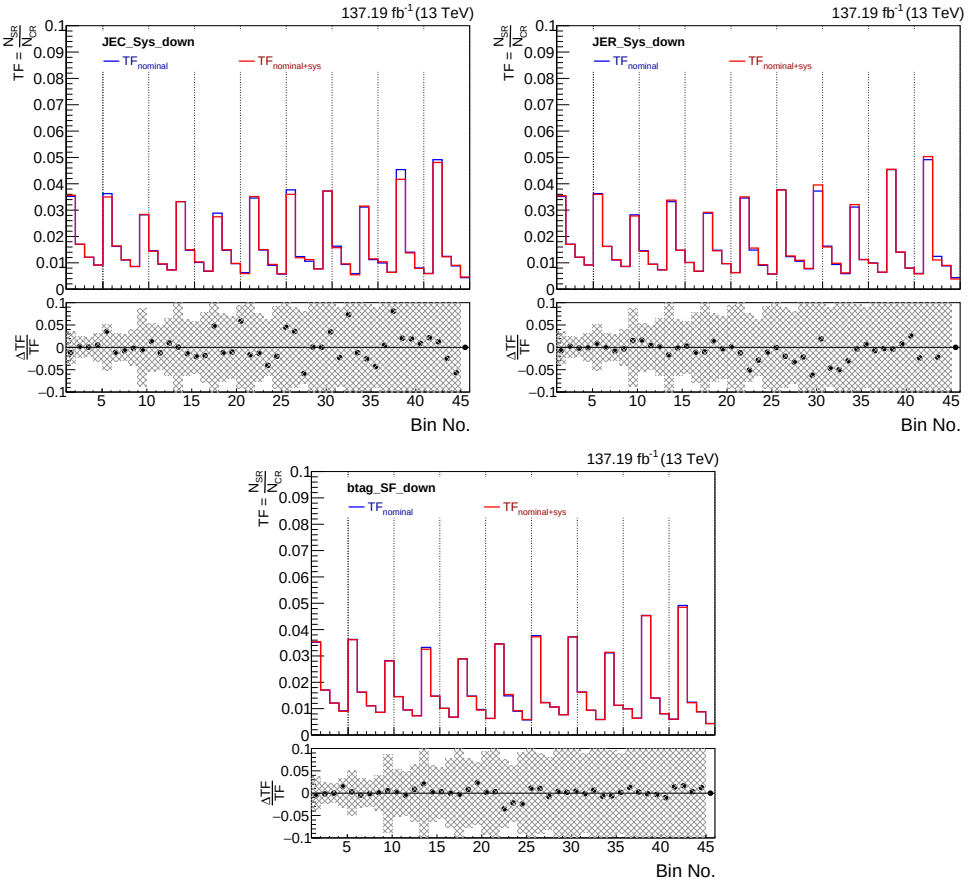


Figure B.3: Comparison of nominal TF (in blue) with TF after varying the JEC factors (top left) and JER (top right) by $-1\sigma_{sys}$, and varying the scale factor of b-tagging (bottom) are shown in the top panel. Shaded region represents the statistical uncertainty in the bottom panel and black points are showing relativistic differences in the TF defined as the ratio of $(TF_{nominal} - TF_{nominal+sys})$ to $TF_{nominal}$.

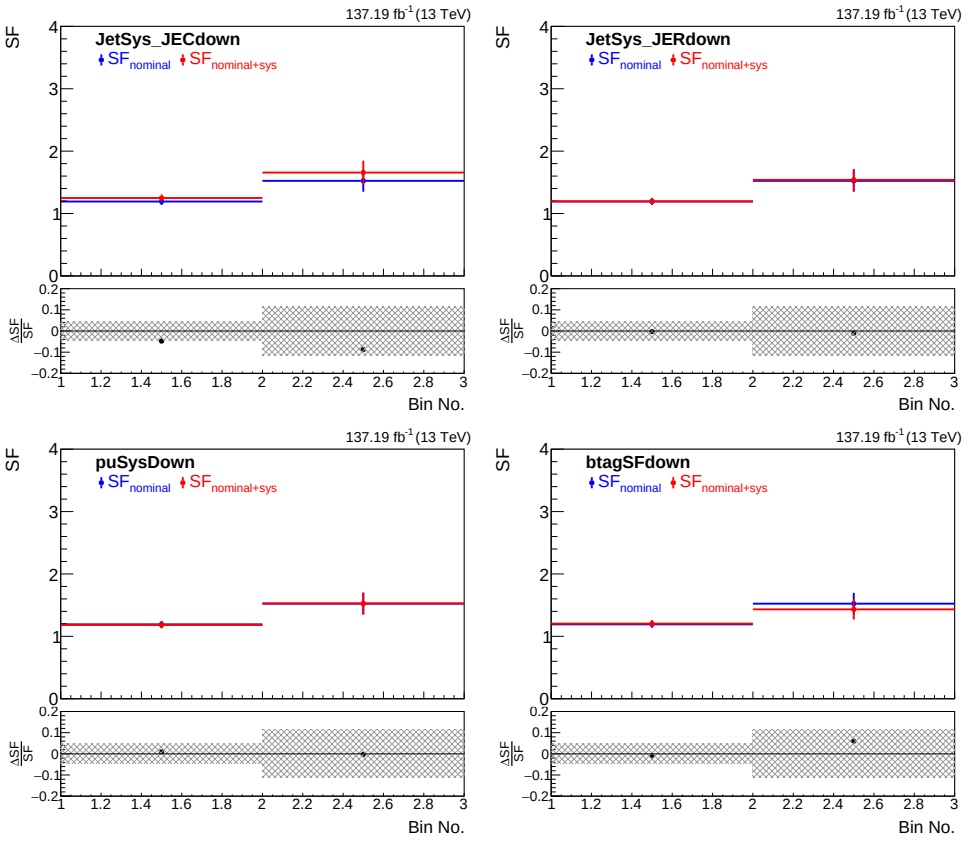


Figure B.4: Comparison of nominal α or SF (in blue) with α after varying the JEC factors (top left) and JER (top right) by $-1\sigma_{sys}$, and varying the pile up weights (bottom left) and scale factor of b-tagging (bottom right) are shown in the top panel. Shaded region represents the statistical uncertainty in the bottom panel and black points are showing relativistic differences in the α or SF defined as the ratio of $(\alpha_{nominal} - \alpha_{nominal+sys})$ to $\alpha_{nominal}$.

List of Acronyms

ADC *Analogue to Digital Converter — Converts analog signals from cells into digital form for processing.*

AK4 *Anti-k_T algorithm with R=0.4 — Jet clustering algorithm with radius parameter 0.4.*

ALICE *A Large Ion Collider Experiment — LHC experiment focused on heavy-ion collisions.*

ASIC *Application Specific Integrated Circuit — Custom-designed microchip for a specific function, used in readout electronics.*

ATCA *Advanced Telecommunications Computing Architecture — A standardized hardware platform used for modular electronics systems in HEP.*

ATLAS *A Toroidal LHC Apparatus*

BDT *Boosted Decision Tree — A machine learning algorithm used for classification, common in particle identification tasks.*

BE *Back-End — Refers to the electronics and systems located off-detector, typically for data handling and control.*

BR *Branching Ratio — Probability of a particle decaying in a specific channel.*

BX *Bunch Crossing — The interval (25 ns) between consecutive proton bunches in the LHC.*

CALICE AHCAL *CALICE Analogue Hadronic Calorimeter — High-resolution sampling calorimeter prototype for future colliders.*

cDAQ *Central DAQ — The main data acquisition system of the CMS detector.*

CE-E *Calorimeter Endcap - Electromagnetic — The electromagnetic section of the endcap calorimeter in CMS.*

CE-H *Calorimeter Endcap - Hadronic — The hadronic section of the endcap calorimeter in CMS.*

CERN *European Organisation for Nuclear Research*

CL *Confidence Level — Statistical measure of significance.*

CLUE *Clustering for Light-weight Unsupervised Energy reconstruction — Fast clustering algorithm developed for HGCAL.*

CMS *Compact Muon Solenoid*

CMS POG *CMS Physics Object Group — CMS subgroups focused on object reconstruction (e.g., photons, jets).*

CMSSW *CMS Software Framework — The main software framework used for CMS simulation, reconstruction, and analysis.*

CNN *Convolutional Neural Network — Widely used deep learning model for image and pattern recognition.*

CR *Control Region — Region in data orthogonal to signal region and used to estimate background processes.*

CRC *Cyclic Redundancy Check — An error-detection technique used in digital communications.*

CSC *Cathode Strip Chamber — A type of gaseous muon detector used in the CMS endcap.*

DAQ *Data Acquisition — The system responsible for collecting data from the detector.*

DCS *Detector Control System — Manages and monitors the operating parameters of the detector.*

DNN *Deep Neural Network — Multi-layered neural network used in complex pattern recognition.*

DQM *Data Quality Monitoring — Framework for online and offline monitoring of CMS data integrity.*

DRN *Dynamic Reduction Network*

DT *Drift Tube — A gaseous detector used in the CMS barrel muon system.*

DTH *DAQ and TCDS2 Hub — Hub module for handling data and timing control in Phase-2 electronics.*

ECAL *Electromagnetic Calorimeter — Measures energy of electrons and photons with high precision.*

ECON-D *DAQ elink Concentrator — A concentrator ASIC that aggregates data from several eLinks.*

ECON-T *TPG elink Concentrator — ASIC for aggregating trigger primitives for fast decision logic.*

elinks *Electrical Links — High-speed serial data lines used to transmit detector data.*

eRx *elink Receiver*

eTx *elink Transmitter*

FC *Fast Control — High-priority control commands sent to the front-end.*

FCNN *Fully Connected Neural Network — Basic form of neural network where each layer is fully connected to the next.*

FE *Front-End — Electronics positioned close to the detector, performing digitization and preliminary signal processing and sending data to Back-End.*

FPGA *Field Programmable Gate Array — Reconfigurable chip used for custom digital logic tasks.*

FR *Fake Rate — Probability that an object is misidentified as another.*

Gbps *Gigabits per second — A unit of data transfer rate equal to* 10^9 *bits per second.*

GCNN *Graph Convolutional Neural Network — A type of GNN with convolution layers designed for graphs.*

GEANT4 *GEometry ANd Tracking version 4 — Toolkit for simulating particle interactions with matter.*

GeV *Giga electron Volt — Unit of energy commonly used in particle physics.*

GMSB *Gauge Mediated Supersymmetry Breaking — SUSY breaking scenario predicting long-lived particles.*

GNN *Graph Neural Network — Machine learning model that operates on graph-structured data.*

HCAL *Hadronic Calorimeter — Measures energy of hadrons such as protons and neutrons.*

HD *High Density*

HGCAL *High Granularity Calorimeter*

HGCROC *HGCAL Readout Chip — Front-end ASIC developed for the readout of the HGCAL detector, performing digitization and preliminary signal processing.*

HL-LHC *High Luminosity LHC — The upcoming upgrade of the LHC aiming at a 10× increase in integrated luminosity.*

HLT *High-Level Trigger*

ID *Identification — Object-level selection, e.g., photon or electron ID criteria.*

L1A *Level-1 Accept — Signal indicating that the Level-1 trigger has selected an event.*

LD *Low Density*

LHC *Large Hadron Collider — The world's most powerful particle accelerator located at CERN.*

LHCb *Large Hadron Collider beauty — LHC experiment focused on studying b-hadrons.*

LINAC *Linear Accelerator*

LO *Leading Order — First term in a perturbative expansion in quantum field theory calculations.*

lpGBT *Low Power Gigabit Transceiver — Radiation-hard transceiver ASIC used in high-speed data links.*

LSP *Lightest Supersymmetric Particle — A stable particle in many SUSY models, candidate for dark matter.*

Mbps *Megabits per second — A unit of data transfer rate equal to* 10^6 *bits per second.*

MC *Monte Carlo — Simulated events generated using physics models.*

MET *Missing Transverse Energy — Vector sum imbalance of transverse momentum; indicates undetected particles in collider experiments.*

MeV *Mega electron Volt — Unit of energy;* 10^6 *eV.*

MTD *MIP Timing Detector — A new CMS subsystem for precision timing of minimum ionizing particles.*

NIM *Nuclear Instrumentation Module — Standard for fast electronics used in nuclear and particle physics.*

NLL *Next-to-Leading Logarithmic — Logarithmic correction to QCD calculations at high orders.*

NLO *Next-to-Leading Order — Second term in perturbative expansion, includes additional precision.*

NLSP *Next-to-Lightest Supersymmetric Particle — Decays to LSP; important in SUSY phenomenology.*

NN *Neural Network — Machine learning model inspired by biological neural systems.*

NZS *Non-Zero-Suppression — Refers to data readout where even non-significant channels are included.*

Pb *Lead*

PCB *Printed Circuit Board — The baseboard that holds and connects electronic components.*

PDF *Parton Distribution Function — Describes the momentum distribution of quarks and gluons in a proton.*

pMSSM *Phenomenological MSSM — A simplified version of the Minimal Supersymmetric Standard Model.*

PS *Proton Synchrotron — A circular accelerator that pre-accelerates protons for injection into the LHC.*

PSB *Proton Synchrotron Booster — Pre-accelerates particles before they enter the PS.*

PU *Pileup — Multiple proton-proton interactions occurring in the same bunch crossing.*

RAM *Random Access Memory — Volatile memory used for temporary data storage in electronics.*

RPC *Resistive Plate Chamber — Gaseous detector used in the CMS muon system.*

SC *Slow Control — Configuration and monitoring commands for system control.*

SF *Scale Factor*

SPS *Super Proton Synchrotron — An intermediate accelerator stage before the LHC.*

SR *Signal Region — Region of phase space enriched in potential new physics signal.*

ST *Total Visible Energy*

SUSY *Supersymmetry — Theoretical framework extending the Standard Model.*

TB *Test Beam*

TC *Trigger Cell — Group of channels combined for triggering in HGCAL.*

TCDS2 *Phase-2 Timing and Control Distribution System — Provides timing, clock, and control signals in Phase-2 CMS.*

TDC *Time Digital Converter*

TF *Transfer Factor*

TICL *The Iterative CLustering algorithm — Algorithm for track-shower reconstruction in HGCAL.*

ToA *Time of Arrival — The recorded time at which a particle hit occurs.*

ToT *Time Over Threshold — Duration a signal stays above a predefined threshold.*

TPG *Trigger Primitive Generator — Creates initial trigger objects based on detector input.*

VTRx+ *Versatile Link+ Transceiver — Optical transceiver module used in CMS Phase-2.*

WS *Weighted Summation*

Bibliography

[1] S. Navas et al. Review of particle physics. *Phys. Rev. D*, 110(3):030001, 2024. doi: 10.1103/PhysRevD.110.030001.

[2] Andrea Dainese, Michelangelo Mangano, Andreas B Meyer, Aleandro Nisati, Gavin Salam, and Mika Anton Vesterinen. Report on the Physics at the HL-LHC, and Perspectives for the HE-LHC. Technical report, CERN, Geneva, Switzerland, 2019. URL https://cds.cern.ch/record/2703572.

[3] Lyndon Evans and Philip Bryant. Lhc machine. *JINST*, 3(S08001), 2008.

[4] The CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 3 (S08004), 2008.

[5] The ATLAS Collaboration. The atlas experiment at the cern large hadron collider. *Journal of Instrumentation*, 3(08):S08003, 08 2008. doi: 10.1088/1748-0221/3/08/S08003. URL https://dx.doi.org/10.1088/1748-0221/3/08/S08003.

[6] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008. doi: 10.1088/1748-0221/3/08/S08002.

[7] The LHCb collaboration. The LHCb Detector at the LHC. volume 3, page S08005, 2008. doi: 10.1088/1748-0221/3/08/S08005. URL https://cds.cern.ch/record/1129809. Also published by CERN Geneva in 2010.

[8] A.M. Sirunyan *et al.* CMS Collaboration. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(P10003), 2017.

[9] The CMS collaboration. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30–61, 2012. ISSN 0370-2693. doi: https://doi.org/10.1016/j.physletb.2012.08.021. URL https://www.sciencedirect.com/science/article/pii/S0370269312008581.

[10] The ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29, 2012. ISSN 0370-2693. doi: https://doi.org/10.1016/j.physletb.2012.08.020. URL https://www.sciencedirect.com/science/article/pii/S037026931200857X.

[11] The CMS collaboration. A portrait of the higgs boson by the cms experiment ten years after the discovery. *Nature*, 607(7917). ISSN 1476-4687. doi: 10.1038/s41586-022-04892-x. URL http://dx.doi.org/10.1038/s41586-022-04892-x.

[12] Stephen P. Martin. A Supersymmetry primer. *Adv. Ser. Direct. High Energy Phys.*, 18:1–98, 1998. doi: 10.1142/9789812839657_0001.

[13] Glennys R. Farrar and Pierre Fayet. Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry. *Physics Letters B*, 76(5):575–579, 1978. ISSN 0370-2693. doi: https://doi.org/10.1016/0370-2693(78)90858-4. URL https://www.sciencedirect.com/science/article/pii/0370269378908584.

[14] G.F. Giudice and R. Rattazzi. Theories with gauge-mediated supersymmetry breaking. *Physics Reports*, 322(6):419–499, 1999. ISSN 0370-1573. doi: https://doi.org/10.1016/S0370-1573(99)00042-3. URL https://www.sciencedirect.com/science/article/pii/S0370157399000423.

[15] CMS twiki. URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS#Run_2_Summary_Plots_13_TeV. CMS SUS Physics Results.

[16] 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. *JHEP*, 10(046), 2023. doi: https://doi.org/10.1007/JHEP10(2023)046.

[17] The CMS Collaboration. The cms publications, . URL https://cms-results.web.cern.ch/cms-results/public-results/publications/. Accessed on January 24, 2025.

[18] The CMS Collaboration. Technical Proposal for the Phase-II Upgrade of the CMS Detector. *CERN-LHCC-2015-010 ; LHCC-P-008 ; CMS-TDR-15-02*, 2015. URL https://cds.cern.ch/record/2020886.

[19] Alan D. Martin and Francis Halzen. *Quarks and Leptons: Introductory Course in Modern Particle Physics*. John Wiley & Sons, 1984.

[20] M. Tanabashi *et al. (Particle Data Group)*. Review of Particle Physics. *Phys. Rev. D*, 98(030001), 2018.

[21] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Addison-Wesley, 1995.

[22] David Tong. Quantum field theory, lecture notes. URL https://www.damtp.cam.ac.uk/user/tong/qft/. Accessed: 2025-01-10.

[23] David J. Griffiths. *Introduction to Elementary Particles*. John Wiley & Sons, 2020.

[24] *The Glashow-Salam-Weinberg Electroweak Theory*, chapter Chapter 13, pages 316–336. . doi: 10.1142/9789813270039_0014. URL https://worldscientific.com/doi/abs/10.1142/9789813270039_0014.

[25] G. Arnison and *et.al* A. Astbury. Experimental observation of isolated large transverse energy electrons with associated missing energy at s=540 gev. *Physics Letters B*, 122(1):103–116, 1983. ISSN 0370-2693. doi: https://doi.org/10.1016/0370-2693(83)91177-2. URL https://www.sciencedirect.com/science/article/pii/0370269383911772.

[26] G. Arnison and *et al.* A. Astbury. Experimental observation of lepton pairs of invariant mass around 95 gev/c2 at the cern sps collider. *Physics Letters B*, 126 (5):398–410, 1983. ISSN 0370-2693. doi: https://doi.org/10.1016/0370-2693(83)90188-0. URL https://www.sciencedirect.com/science/article/pii/0370269383901880.

[27] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, 1964. doi: 10.1103/PhysRevLett.13.508. URL https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[28] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, 1964. doi: 10.1103/PhysRevLett.13.321. URL https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[29] Marco Valente. *Supersymmetric Beasts and Where to Find Them: From Novel Hadronic Reconstruction Methods to Search Results in Large Jet Multiplicity Final States at the ATLAS Experiment*. PhD thesis, Geneva U., 2020. URL https://cds.cern.ch/record/2721059. Presented 24 Apr 2020.

[30] Edvige Corbelli and Paolo Salucci. The extended rotation curve and the dark matter halo of m33. *Monthly Notices of the Royal Astronomical Society*, 311(2): 441–447, 01 2000. ISSN 0035-8711. doi: 10.1046/j.1365-8711.2000.03075.x. URL https://doi.org/10.1046/j.1365-8711.2000.03075.x.

[31] Y. Fukuda and *et.al* Hayakawa, T. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, 1998. doi: 10.1103/PhysRevLett.81.1562. URL https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[32] H.P. Nilles. Supersymmetry, supergravity and particle physics. *Physics Reports*, 110(1):1–162, 1984. ISSN 0370-1573. doi: https://doi.org/10.1016/0370-1573(84)90008-5. URL https://www.sciencedirect.com/science/article/pii/0370157384900085.

[33] Lectures on Supersymmetry by Matteo Bertolini, SISSA. URL https://www.sissa.it/tpp/phdsection/OnlineResources/6/susycourse.pdf. June 09, 2022.

[34] David Emmanuel-Costa, Pavel Fileviez Pérez, and Ricardo González Felipe. Natural gauge and gravitational coupling unification and the superpartner masses. *Physics Letters B*, 648(1):60–63, 2007. ISSN 0370-2693. doi: https://doi.org/10.1016/j.physletb.2007.02.061. URL https://www.sciencedirect.com/science/article/pii/S0370269307002298.

[35] Zinonas Zinonos. *Search for Supersymmetry with the ATLAS detector at the LHC*. PhD thesis, Technical University of Munich, 2020. URL https://cds.cern.ch/record/2804387. Presented 15 Feb 2022.

[36] Shehu S. AbdusSalam, Benjamin C. Allanach, Fernando Quevedo, Farhan Feroz, and Mike Hobson. Fitting the phenomenological mssm. *Physical Review D*, 81(9), 05 2010. ISSN 1550-2368. doi: 10.1103/physrevd.81.095012. URL http://dx.doi.org/10.1103/PhysRevD.81.095012.

[37] Ruderman J.T. Shih D. General neutralino nlsps at the early lhc. *Journal of High Energy Physics*, 2012(8), 2012. doi: 10.1007/JHEP08(2012)159. URL https://link.springer.com/article/10.1007/JHEP08(2012)159.

[38] Knapen S. Redigolo D. Gauge mediation at the lhc: status and prospects. *Journal of High Energy Physics*, 2012(2), 2012. doi: 10.1007/JHEP02(2012)115. URL https://link.springer.com/article/10.1007/JHEP02(2012)115.

[39] Kats Y. Meade P. Reece M. *et al.* The status of gmsb after 1/fb at the lhc. *Journal of High Energy Physics*, 2017(1), 2017. doi: 10.1007/JHEP01(2017)135. URL https://link.springer.com/article/10.1007/JHEP01(2017)135.

[40] Johan Alwall, Philip C. Schuster, and Natalia Toro. Simplified models for a first characterization of new physics at the lhc. *Physical Review D*, 79(7), 04 2009. ISSN 1550-2368. doi: 10.1103/physrevd.79.075020. URL http://dx.doi.org/10.1103/PhysRevD.79.075020.

[41] Alves Daniele and Arkani-Hamed *et al.* Simplified models for lhc new physics searches. *Journal of Physics G: Nuclear and Particle Physics*, 39(10):105005, sep 2012. doi: 10.1088/0954-3899/39/10/105005. URL https://dx.doi.org/10.1088/0954-3899/39/10/105005.

[42] Chatrchyan S. and Khachatryan V. *et al.* Interpretation of searches for supersymmetry with simplified models. *Physical Review D*, 88(5), 09 2013. ISSN 1550-2368. doi: 10.1103/physrevd.88.052017. URL http://dx.doi.org/10.1103/PhysRevD.88.052017.

[43] CERN twiki. URL https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SUSYCrossSections. LHC SUSY Cross Section Working Group.

[44] C.F. Hermanns. *X-ray Absorption Studies of Metalloporphyrin Molecules on Surfaces: Electronic Interactions, Magnetic Coupling, and Chemical Switches*. 2013. URL https://books.google.co.in/books?id=-wSD0AEACAAJ.

[45] R. Wigmans. *Calorimetry: Energy Measurement in Particle Physics*. International Series of Monographs on Physics. Oxford University Press, 2017. ISBN 9780198786351.

[46] Patrycja Dyrcz. *Radiological characterization of low- and intermediate level (LL/IL) radioactive waste*. PhD thesis, 12 2022.

[47] Tejinder S Virdee. Experimental techniques. 1999. doi: 10.5170/CERN-1999-004.347. URL https://cds.cern.ch/record/454176.

[48] Leonardo Rossi, Peter Fischer, Tilman Rohe, and Norbert Wermes. *Pixel Detectors: From Fundamentals to Applications*. Particle Acceleration and Detection. Springer Berlin, Heidelberg, 2006. ISBN 978-3-540-28332-4. doi: 10.1007/3-540-28333-1. URL https://doi.org/10.1007/3-540-28333-1. Springer-Verlag Berlin Heidelberg 2006.

[49] L. C. Andrews, R. L. Phillips, C. Y. Hopen, and M. A. Al-Habash. Theory of optical scintillation. *Journal of the Optical Society of America A*, 16(6):1417, 1999. doi: 10.1364/josaa.16.001417. URL https://doi.org/10.1364/josaa.16.001417.

[50] T. Bergauer. Calorimeters. URL https://www.oeaw.ac.at/fileadmin/Institute/HEPHY/PDF/ausbildung/teilchendetektoren/VO-7-2017-Calorimeters_v2.pdf. Accessed: 2025-01-13.

[51] Sehwook Lee, Michele Livan, and Richard Wigmans. On the limits of the hadronic energy resolution of calorimeters. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated*

Equipment, 882:148–157, 2018. ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2017.10.087. URL https://www.sciencedirect.com/science/article/pii/S0168900217311762.

[52] CMS HGCAL Collaboration. Response of a CMS HGCAL silicon-pad electromagnetic calorimeter prototype to 20-300 GeV positrons. *JINST*, 17(P05022), 2021.

[53] CMS HGCAL Collaboration. Construction and commissioning of CMS CE prototype silicon modules. *JINST*, 16(T04002), 2021.

[54] Christopher Z Mooney. *Monte carlo simulation*. Number 116. Sage, 1997.

[55] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014 (7), 06 2014. ISSN 1029-8479. doi: 10.1007/jhep07(2014)079. URL http://dx.doi.org/10.1007/JHEP07(2014)079.

[56] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H.-S. Shao, and M. Zaro. The automation of next-to-leading order electroweak calculations. *Journal of High Energy Physics*, 2018(7), 06 2018. ISSN 1029-8479. doi: 10.1007/jhep07(2018)185. URL http://dx.doi.org/10.1007/JHEP07(2018)185.

[57] *et.al* Christian Bierlich. A comprehensive guide to the physics and usage of pythia 8.3, 2022. URL https://arxiv.org/abs/2203.11601.

[58] S. Agostinelli *et al.* GEANT4 - a simulation toolkit. *NIMA*, 506(2):250–303, 2003. URL https://doi.org/10.1016/S0168-9002(03)01368-8.

[59] CERN. Physics reference manual. URL https://geant4.web.cern.ch/documentation/dev/prm_html/PhysicsReferenceManual/. Accessed: 2025-01-13.

[60] Vladimir IVANCHENKO and Apostolakis *et.al*. Recent improvements in geant4 electromagnetic physics models and interfaces. *Progress in NUCLEAR SCIENCE and TECHNOLOGY*, 2:898–903, 10 2011. doi: 10.15669/pnst.2.898.

[61] Aatos Heikkinen, Nikita Stepanov, and J. P. Wellisch. Bertini intra-nuclear cascade implementation in geant4, 2003. URL https://arxiv.org/abs/nucl-th/0306008.

[62] G. Folger, VN Ivanchenko, and Johannes Wellisch. The binary cascade: Nucleon nuclear reactions. *The European Physical Journal A - Hadrons and Nuclei*, 21:407–417, 09 2004. doi: 10.1140/epja/i2003-10219-7.

[63] Kaushick Parui and Arun Minj. Study of the woods-saxon potential and the interactions inside the nucleon core. 01 2018.

[64] Low-energetic hadron interactions in a highly granular calorimeter, 2022. URL www-library.desy.de/preparch/desy/thesis/desy-thesis-11-048.pdf.

[65] G. Folger and J. P. Wellisch. String parton models in geant4, 2003. URL https://arxiv.org/abs/nucl-th/0306007.

[66] N. S. Amelin, E. F. Staubo, and L. P. Csernai. Semihard collisions in Monte Carlo quark gluon string model. *Phys. Rev. D*, 46:4873–4881, 1992. doi: 10.1103/PhysRevD.46.4873.

[67] Werner Herr and B Muratori. Concept of luminosity. 2006. doi: 10.5170/CERN-2006-002.361. URL https://cds.cern.ch/record/941318.

[68] The CMS collaboration. Public cms luminosity information, 2024. URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#Full_summary_proton_proton_colli. Accessed on 13 Jan 2025.

[69] CMS Data Preservation & Open Access group. Cms open data workshop lessons. URL https://cms-opendata-workshop.github.io/workshop2024-lesson-cms-detector/images.html. Accessed on 14 January, 2025.

[70] Guido Tonelli. The silicon tracking system of cms. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 386(1):129–137, 1997. ISSN 0168-9002. doi: https://doi.org/10.1016/S0168-9002(96)01106-0. URL https://www.sciencedirect.com/science/article/pii/S0168900296011060. Proceedings of the 5th International Workshop on Vertex Detectors.

[71] G. Franzoni. Performance of cms ecal with first lhc data. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 628(1):90–94, 2011. ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2010.06.291. URL https://www.sciencedirect.com/science/article/pii/S0168900210014816. VCI 2010.

[72] CMS Collaboration. Performance of cms hadron calorimeter timing and synchronization using test beam, cosmic ray, and lhc beam data. *Journal of Instrumentation*, 5(03):T03013–T03013, 03 2010. ISSN 1748-0221. doi: 10.1088/1748-0221/5/03/t03013. URL http://dx.doi.org/10.1088/1748-0221/5/03/T03013.

[73] The CMS collaboration. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical design report. CMS. CERN, Geneva, 2006. URL https://cds.cern.ch/record/922757. There is an error on cover due to a technical problem for some items.

[74] S. Dasu et al. CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems. 12 2000.

[75] CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger. 12 2002.

[76] Paolo Azzurri. Track reconstruction performance in cms. *Nuclear Physics B - Proceedings Supplements*, 197(1):275–278, 12 2009. ISSN 0920-5632. doi: 10.1016/j.nuclphysbps.2009.10.084. URL http://dx.doi.org/10.1016/j.nuclphysbps.2009.10.084.

[77] The CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the cms tracker. *Journal of Instrumentation*, 9(10):P10009–P10009, 10 2014. ISSN 1748-0221. doi: 10.1088/1748-0221/9/10/p10009. URL http://dx.doi.org/10.1088/1748-0221/9/10/P10009.

[78] CMS Collaboration. Technical proposal for the upgrade of the CMS detector through 2020. Technical report, 2011. URL https://cds.cern.ch/record/1355706.

[79] W. Adam and *et.al* Bergauer T. The CMS Phase-1 Pixel Detector Upgrade. *JINST*, 16(02):P02027, 2021. doi: 10.1088/1748-0221/16/02/P02027. URL https://cds.cern.ch/record/2748381.

[80] Candan Isik. Phase 1 upgrade of the cms hadron barrel calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1042:167389, 2022. ISSN 0168-9002. doi: https://doi.org/10.1016/j.nima.2022.167389. URL https://www.sciencedirect.com/science/article/pii/S0168900222006933.

[81] Lagarde F. and Gouzevitch M. *et.al*. High rate, fast timing glass rpc for the high η cms muon detectors. *Journal of Instrumentation*, 11, 06 2016. doi: 10.1088/1748-0221/11/09/C09006.

[82] The CMS collaboration. The Phase-2 Upgrade of the CMS Tracker. Technical report, CERN, Geneva, 2017. URL https://cds.cern.ch/record/2272264.

[83] The CMS collaboration. The Phase-2 Upgrade of the CMS Muon Detectors. Technical report, CERN, Geneva, 2017. URL https://cds.cern.ch/record/2283189. This is the final version, approved by the LHCC.

[84] Collaboration CMS. A MIP Timing Detector for the CMS Phase-2 Upgrade. Technical report, CERN, Geneva, 2019. URL https://cds.cern.ch/record/2667167.

[85] Cristián H. Peña and on behalf of the CMS Collaboration. Precision timing with the cms mip timing detector. *Journal of Physics: Conference Series*, 1162 (1):012035, 01 2019. doi: 10.1088/1742-6596/1162/1/012035. URL https://dx.doi.org/10.1088/1742-6596/1162/1/012035.

[86] The CMS collaboration. The Phase-2 Upgrade of the CMS L1 Trigger Interim Technical Design Report. Technical report, CERN, Geneva, 2017. URL https://cds.cern.ch/record/2283192. This is the CMS Interim TDR devoted to the upgrade of the CMS L1 trigger in view of the HL-LHC running, as approved by the LHCC.

[87] Badaro Gilbert and Behrens Ulf *et.al*. The Phase-2 Upgrade of the CMS Data Acquisition. Technical report, CERN, Geneva, 2021. URL https://cds.cern.ch/record/2797509.

[88] Dan Green. *High Pt Physics at Hadron Colliders*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, Cambridge, UK.

[89] A. Hayrapetyan and *et al.* Tumasyan. Performance of the cms electromagnetic calorimeter in pp collisions at $\sqrt{s}$= 13 tev. *Journal of Instrumentation*, 19(09): P09004, 09 2024. ISSN 1748-0221. doi: 10.1088/1748-0221/19/09/p09004. URL http://dx.doi.org/10.1088/1748-0221/19/09/P09004.

[90] T.French (CERN) on behalf of the CMS HGCAL collaboration. An overview of the cms high granularity calorimeter, and the engineering challenges in its construction, 2024. URL https://indico.cern.ch/event/1416680/contributions/5954483/attachments/2856535/4997949/An%20overview%20of%20the%20CMS%20High%20Granularity%20Calorimeter%20CALOR24%20v3.pdf.

[91] A. David on behalf of the CMS Collaboration. Performance of cms high-granularity calorimeter for hl-lhc, 2011. URL https://agenda.infn.it/event/28874/contributions/169080/. Parallel talk.

[92] Alessandro La Rosa. *The Upgrade of the CMS Tracker at HL-LHC*. JPS, 2021. doi: 10.7566/JPSCP.34.010006. URL https://journals.jps.jp/doi/abs/10.7566/JPSCP.34.010006.

[93] CMS Collaboration. The Phase-2 Upgrade of the CMS Data Acquisition and High Level Trigger. Technical report, CERN, Geneva, 2021. URL https://cds.cern.ch/record/2759072. This is the final version of the document, approved by the LHCC.

[94] The CMS Collaboration. The Phase-2 Upgrade of the CMS Endcap Calorimeter. *CERN-LHCC-2017-023. LHCC-P-008. CMS-TDR-019*, 2017.

[95] A. Oblakowska-Mucha. Radiation Hard Silicon Particle Detectors for Phase-II LHC Trackers. *JINST*, 12(C02054), 2017.

[96] Paul Aspell. Hgcal electronics system overview, 2023. URL https://indico.cern.ch/event/1325034/contributions/5575786/attachments/2748899/4785693/HGCAL%20Electronics%20System%20Overview%20-%20ESR1.pdf. [Accessed: 20-Sept-2024].

[97] for HGCAL SiPM-on-tile group Ted Kolberg. Ce esr1: Sipm-on-tile system overview, motherboards, cable assemblies, 2023. URL https://indico.cern.ch/event/1325034/contributions/5595275/attachments/2749680/4786929/ESR_tkolberg_overview.pdf.

[98] CMS HGCAL Collaboration. Performance of the CMS High Granularity Calorimeter prototype to charged pion beams of 20–300 GeV/c. *JINST*, 18 (P08014), 2023.

[99] F. Sefkow and F. Simon. A highly granular SiPM-on-tile calorimeter prototype. *J. Phys. Conf. Ser.*, 1162(012012), 2019. doi: arXiv:1808.09281.

[100] J. Borg *et al.* Skiroc2-CMS an ASIC for testing CMS HGCAL. *JINST*, 12(C02019), 2017. doi: 10.1088/1748-0221/12/02/C02019. URL https://dx.doi.org/10.1088/1748-0221/12/02/C02019.

[101] CMS HGCAL Collaboration. The DAQ system of the 12,000 channel CMS high granularity calorimeter prototype. *JINST*, 16(T04001), 2021.

[102] Shubham Pandey. *Performance of High Granularity Calorimeter prototypes for the CMS HL-LHC upgrade in beam test experiments at CERN*. PhD thesis, Indian Institute of Science Education and Research (IISER) Pune, Department of Physics, India, IISER, Pune, 2022.

[103] Martín Abadi. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. 2015. URL http://tensorflow.org/. Software available from tensorflow.org.

[104] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks. *Commun. ACM*, 60 (6):84–90, 05 2017. ISSN 0001-0782. doi: 10.1145/3065386. URL https://doi.org/10.1145/3065386.

[105] A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M.D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa, and P. Vahle. A convolutional neural network neutrino event classifier. *Journal of Instrumentation*, 11(09):P09001, 09 2016. doi: 10.1088/1748-0221/11/09/P09001. URL https://dx.doi.org/10.1088/1748-0221/11/09/P09001.

[106] MicroBooNE Collaboration. Deep neural network for pixel-level electromagnetic particle identification in the MicroBooNE liquid argon time projection chamber. *Phys. Rev. D*, 99:092001, 05 2019. doi: 10.1103/PhysRevD.99.092001. URL https://link.aps.org/doi/10.1103/PhysRevD.99.092001.

[107] CMS Collaboration. Search for Exotic Higgs Boson Decays $H \to \mathcal{A}\mathcal{A} \to 4\gamma$ with Events Containing Two Merged Diphotons in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. Lett.*, 131:101801, 09 2023. doi: 10.1103/PhysRevLett.131.101801. URL https://link.aps.org/doi/10.1103/PhysRevLett.131.101801.

[108] Geeks for Geeks. Linear regression in machine learning, . URL https://www.geeksforgeeks.org/ml-linear-regression/.

[109] Sebastian Ruder. An overview of gradient descent optimization algorithms. *CoRR*, abs/1609.04747, 2016. URL http://arxiv.org/abs/1609.04747.

[110] Cnns for visual recognition, . URL http://cs231n.github.io/neural-networks-3/.

[111] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017. URL https://arxiv.org/abs/1412.6980.

[112] Geeks for Geeks. What is a neural network?, . URL https://www.geeksforgeeks.org/neural-networks-a-beginners-guide/.

[113] Rukshan Pramoditha. Overview of a neural network's learning process. URL https://medium.com/data-science-365/overview-of-a-neural-networks-learning-process-61690a502fa.

[114] DeZoort Gage Battaglia Peter W Biscarat Catherine Vlimant Jean-Roch. Graph neural networks at the large hadron collider. *Nature Reviews Physics*, 2023. doi: 10.1038/s42254-023-00569-0. URL https://doi.org/10.1038/s42254-023-00569-0.

[115] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, and Justin M. Solomon. Dynamic Graph CNN for Learning on Point Clouds. 2019. doi: https://doi.org/10.48550/arXiv.1801.07829.

[116] Gentle introduction to graph neural and graph convolutional networks, . URL https://www.avenga.com/magazine/graph-neural-networks-and-graph-convolutional-networks/.

[117] Matthew N. Bernstein. Graph convolutional neural networks. 2023. URL https://mbernste.github.io/posts/gcn/.

[118] Larissa C. Shimomura, Rafael Seidi Oyamada, Marcos R. Vieira, and Daniel S. Kaster. A survey on graph-based methods for similarity searches in metric spaces. *Information Systems*, 95:101507, 2021. ISSN 0306-4379. doi: https://doi.org/10.1016/j.is.2020.101507. URL https://www.sciencedirect.com/science/article/pii/S0306437920300181.

[119] Inderjit S Dhillon, Yuqiang Guan, and Brian Kulis. Weighted graph cuts without eigenvectors: A multilevel approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(11):1944–1957, 2007.

[120] Bas O. Fagginger Auer and Rob H. Bisseling. A GPU Algorithm for Greedy Graph Matching. *Springer-Verlag, Berlin, Heidelberg*, pages 108–119, 2012.

[121] Rex Ying, Jiaxuan You, Christopher Morris, Xiang Ren, William L. Hamilton, and Jure Leskovec. Hierarchical graph representation learning with differentiable pooling, 2019. URL https://arxiv.org/abs/1806.08804.

[122] Rex Ying, Jiaxuan You, Christopher Morris, Xiang Ren, William L. Hamilton, and Jure Leskovec. Hierarchical graph representation learning with differentiable pooling. *ArXiv*, abs/1806.08804, 2018. URL https://api.semanticscholar.org/CorpusID:49420315.

[123] Lindsey Gray *et al.* A dymanic reduction network for point clouds. *arXiv:2003.08013v1*, 2020.

[124] Frederik Diehl. Edge Contraction Pooling for Graph Neural Networks. *arXiv:1905.10990*, 2019.

[125] Alpana Alpana *et al.* Hgcal pions energy reconstruction using graphical neural networks. *CMS internal detector note*, 2022. URL https://cms.cern.ch/iCMS/user/noteinfo?cmsnoteid=CMS%20DN-2022/006.

[126] Aamir M. *et al.* Using graph neural networks to reconstruct charged pion showers in the cms high granularity calorimeter. *Journal of Instrumentation*, 19 (11):P11025, nov 2024. doi: 10.1088/1748-0221/19/11/P11025. URL https://dx.doi.org/10.1088/1748-0221/19/11/P11025.

[127] Jan Eric Lenssen Matthias Fey. Fast graph representation learning with pytorch geometric. *arXiv*, 2019.

[128] .

[129] NITISH KUMAR K V. *Semiparametric graph neural networks for energy regression of hadron showers in the CMS High Granularity Calorimeter*. PhD thesis. URL http://dr.iiserpune.ac.in:8080/xmlui/handle/123456789/7880.

[130] Felice Pantaleo *et al.* The Iterative Clustering framework for the CMS HGCAL Reconstruction. *Journal of Physics : Conference series*, 2438(012096), 2023.

[131] Marco Rovere *et al.* CLUE: A Fast Parallel Clustering Algorithm for High Granularity Calorimeters in High Energy Physics. *arXiv:2001.09761*, 2020.

[132] Alessandro Laio Alex Rodriguez. Clustering by fast search and find of density peaks. *https://www.science.org/doi/epdf/10.1126/science.1242072*, pages 1492–1496, 2014.

[133] A.-M. Magnan *et al.* Geant4 simulation of a Si-based sampling calorimeter. *https://github.com/pfs/PFCal/tree/master/PFCalEE*. URL https://github.com/pfs/PFCal/tree/master/PFCalEE.

[134] P.Ferreira da Silva, B.Alves, A.M. Magnan, and C.Seez. Investigations on the impact of re-optimization of the hgcal longitudinal structure on hadron energy resolution. *CMS internal detector note*, 2021. URL https://cms.cern.ch/iCMS/user/noteinfo?cmsnoteid=CMS%20DN-2021/013.

[135] Alpana Alpana *et al.* Performance studies of hgcal readout buffers - truncation vs deadtime. *CMS internal detector note*, 2024. URL https://cms.cern.ch/iCMS/user/noteinfo?cmsnoteid=CMS%20DN-2024/026.

[136] J. Mans. FAST CONTROL FOR THE CMS ENDCAP CALORIMETER. *Internal CMS note*, 2023. URL https://edms.cern.ch/ui/file/2044200/4.4/fast-control-spec_V4_8.pdf.

[137] Milos Vojinovic. *Design and implementation of on-detector and off-detector electronics for the high-luminosity upgrade of the CMS endcap calorimeter system within LHC at CERN*. PhD thesis, Imperial Coll., London, 2024. URL https://cds.cern.ch/record/2918629. Presented 07 Oct 2024.

[138] CMS Collaboration. The Phase-2 Upgrade of the CMS Level-1 Trigger. Technical report, CERN, Geneva, 2020. URL https://cds.cern.ch/record/2714892. Final version.

[139] on behalf of the CMS HGCAL collaboration S. Mallios. Overview of the front-end electronics of cms hgcal, 2024. URL https://indico.cern.ch/event/1416680/contributions/5954503/attachments/2856548/4997954/CALOR2024_smallios.pdf.

[140] F.A. Khan on behalf of the CMS collaboration. Design and performance optimisation of the Hexaboards for CMS HGCAL silicon sensor readout electronisc. *JINST*, 18(C03015). doi: 10.1088/1748-0221/18/03/C03015. URL https://iopscience.iop.org/article/10.1088/1748-0221/18/03/C03015.

[141] Hafiza Ayesha Ahmed. Hexaboards – the family of readout boards for hgcal silicon modules, 2023. URL https://indico.cern.ch/event/1254144/contributions/5307543/attachments/2613464/4516510/Hexaboards_CMS2023.pdf.

[142] on behalf of the CMS HGCAL collaboration Nadja Strobbe. Ld engines & wagon, 2023. URL https://indico.cern.ch/event/1325036/contributions/5595447/attachments/2750521/4787837/ESR_engines_wagons_20231113.pdf.

[143] M. Noy (et al.) P. Aspell. Working document on Specification of HGCROC3. *Internal CMS note*, 2022. URL https://edms.cern.ch/ui/file/2324379/1/HGCROC3_Spec_Working_Document_v2.0.pdf.

[144] on behalf of the CMS collaboration Damien Thienpont. Hgcroc: the front-end readout asics for the cms high granularity calorimeter, 2022. URL https://indico.cern.ch/event/981823/contributions/4271242/attachments/2253760/3824229/TIPP21_HGCROC.pdf.

[145] Danny Noonan James Hirschauer, James Hoff. DAQ CONCENTRATOR ASIC (ECON-D) FOR THE CMS HL-LHC ENDCAP CALORIMETER. *Internal CMS note*, 2022. URL https://edms.cern.ch/ui/file/2599620/1/ECON-D_specification_working_doc_v1.3_08jul2022.pdf.

[146] Paulo RODRIGUES SIMOES MOREIRA. The LpGBT Project, Status and Overview. ACES 2016 - Fifth Common ATLAS CMS Electronics Workshop for LHC Upgrades. 2016. URL https://cds.cern.ch/record/2137809.

[147] Jan Troska, Alexander Brandon-Bravo, Stephane Detraz, Andrea Kraxner, Lauri Olanterä, Carmelo Scarcella, Christophe Sigaud, Csaba Soos, and Francois Vasey. The VTRx+, an optical link module for data transmission at HL-LHC. *PoS*, TWEPP-17:048, 2017. doi: 10.22323/1.313.0048. URL https://cds.cern.ch/record/2312396.

[148] Christoph Schwick Dominique Gigi, Jeroen Hegeman. The Phase-2 CMS SLinkRocket DAQ link Protocol and format description. *Internal CMS note*, 2022. URL https://edms.cern.ch/ui/file/2502737/2/cms_phase2_slinkrocket.pdf.

[149] HGCAL BE Group. HGCAL back end DAQ firmware. *Internal CMS note*, 2022. URL https://docs.google.com/document/d/1HtrDfQ7ZlckcuLTkP_LrK5srN_jrwCcLhrM2Vyt8Rgc/edit.

[150] Paul Dauncey. Event data size, 2021. URL https://indico.cern.ch/event/991178/contributions/4188990/attachments/2175171/3706244/210121.pdf.

[151] Multivariate gaussian random number generation targeting reconfigurable hardware. URL https://www.doc.ic.ac.uk/~wl/papers/08/trets08dt.pdf.

[152] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[153] Exploring quantile mapping as a tool to produce user-tailored climate scenarios for switzerland. 2018. URL https://api.semanticscholar.org/CorpusID:211566777.

[154] S. Kulkarni. Susy new ideas - pushing the boundaries. URL https://indico.cern.ch/event/828287/contributions/3546370/attachments/1910401/3156974/Kulkarni_SUSY_workshop_CMS_internal.pdf. Accessed on September 18, 2019.

[155] S. Navas et al. Review of particle physics. *Phys. Rev. D*, 110(3):030001, 2024. doi: 10.1103/PhysRevD.110.030001.

[156] Bhumika Kansal. *Search for Supersymmetry in proton-proton collisions at $\sqrt{s}$ = 13 TeV with at least one photon, jets and missing transverse momentum*. PhD thesis, IISER, Pune, 2024. URL https://cds.cern.ch/record/2897937. Presented 28 Oct 2023.

[157] The CMS Collaboration. Jet identification for the 13 tev ul data, . URL https://twiki.cern.ch/twiki/bin/view/CMS/JetID13TeVUL. Accessed on January 22, 2025.

[158] 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. *CMS Analysis Note*. AN-15-139 (2015).

[159] The CMS Collaboration. Identification of heavy-flavour jets with the cms detector in pp collisions at 13 tev. *JINST*, 13(05):P05011, 05 2018. doi: 10.1088/1748-0221/13/05/P05011. URL https://dx.doi.org/10.1088/1748-0221/13/05/P05011.

[160] The CMS Collaboration. Cms b-tagging & vertexing, . URL https://btv-wiki.docs.cern.ch/ScaleFactors/UL2016preVFP/#ak4-b-tagging. Accessed on December 25, 2024.

[161] A.M. Sirunyan et al. Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s}$ = 13 tev using the cms detector. *JINST*, 14(07):P07004, 07 2019. doi: 10.1088/1748-0221/14/07/P07004. URL https://dx.doi.org/10.1088/1748-0221/14/07/P07004.

[162] The CMS Collaboration. Noise filter recommendations for run ii & run iii, . URL https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2#2018_2017_data_and_MC_UL. Accessed on December 25, 2024.

[163] Aditee Prabhakar Rane. *Search for Supersymmetry in proton-proton collisions at $\sqrt{s}$ = 13 TeV with jets, b-jets and missing transverse momentum*. PhD thesis, IISER, Pune, 2020. URL https://cds.cern.ch/record/2740850.

[164] Gross E. et al Cowan G., Cranmer K. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71(1554), 01 201. doi: 10.1140/epjc/s10052-011-1554-0. URL https://doi.org/10.1140/epjc/s10052-011-1554-0.

[165] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, José I. Latorre, Juan Rojo, and Maria Ubiali. Parton distributions for the lhc run ii. *Journal of High Energy Physics*, 2015(4), 04 2015. ISSN 1029-8479. doi: 10.1007/jhep04(2015)040. URL http://dx.doi.org/10.1007/JHEP04(2015)040.

[166] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Luigi Del Debbio, Stefano Forte, Patrick Groth-Merrild, Alberto Guffanti, Nathan P. Hartland, Zahari Kassabov, José I. Latorre, Emanuele R. Nocera, Juan Rojo, Luca Rottoli, Emma

Slade, and Maria Ubiali. Parton distributions from high-precision collider data: Nnpdf collaboration. *The European Physical Journal C*, 77(10), 10 2017. ISSN 1434-6052. doi: 10.1140/epjc/s10052-017-5199-5. URL http://dx.doi.org/10.1140/epjc/s10052-017-5199-5.

[167] Stefano Catani, Daniel de Florian, Massimiliano Grazzini, and Paolo Nason. Soft-gluon resummation for higgs boson production at hadron colliders. *Journal of High Energy Physics*, 2003(07):028–028, 07 2003. ISSN 1029-8479. doi: 10.1088/1126-6708/2003/07/028. URL http://dx.doi.org/10.1088/1126-6708/2003/07/028.

[168] M Cacciari, S Frixione, M.L Mangano, P Nason, and G Ridolfi. The tbar-t cross-section at 1.8 and 1.96 tev: a study of the systematics due to parton densities and scale dependence. *Journal of High Energy Physics*, 2004(04):068–068, 04 2004. ISSN 1029-8479. doi: 10.1088/1126-6708/2004/04/068. URL http://dx.doi.org/10.1088/1126-6708/2004/04/068.

[169] Tumasyan A. and Adam W. *et al.* Measurement of the inclusive and differential $t\bar{t}\gamma$ cross sections in the single-lepton channel and eft interpretation at $\sqrt{s}$ = 13 tev. *JHEP*, 2021. doi: 10.1007/JHEP12(2021)180.

[170] A.M. Sirunyan and A. Tumasyan *et al.* Search for disappearing tracks in proton-proton collisions at s=13tev. *Physics Letters B*, 806:135502, 2020. ISSN 0370-2693. doi: https://doi.org/10.1016/j.physletb.2020.135502. URL https://www.sciencedirect.com/science/article/pii/S0370269320303063.

[171] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011. doi: 10.1140/epjc/s10052-011-1554-0. URL https://doi.org/10.1140/epjc/s10052-011-1554-0.

[172] A L Read. Presentation of search results: the cls technique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693, 09 2002. doi: 10.1088/0954-3899/28/10/313. URL https://dx.doi.org/10.1088/0954-3899/28/10/313.

[173] Aram Hayrapetyan et al. The CMS statistical analysis and combination tool: COMBINE. *Comput. Softw. Big Sci.*, 8:19, 2024. doi: 10.1007/s41781-024-00121-4.

[174] CERN. Procedure for the LHC Higgs boson search combination in Summer 2011. Technical report, CERN, Geneva, 2011. URL https://cds.cern.ch/record/1379837.

[175] Andreas Hocker et al. TMVA - Toolkit for Multivariate Data Analysis. 3 2007.

[176] R. Brun and F. Rademakers. ROOT: An object oriented data analysis framework. *Nucl. Instrum. Meth. A*, 389:81–86, 1997. doi: 10.1016/S0168-9002(97)00048-X.