Top Quark Analysis and Trigger Studies with CMS Run 2 Dataset & Outer Tracker Upgrade for HL-LHC

By

Aloke Kumar Das

Enrollment No : PHYS11201704006

National Institute of Science Education and Research
Bhubaneswar

A thesis submitted to the

Board of Studies in Physical Sciences

In partial fulfillment of requirements

for the Degree of

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE

September, 2022

Homi Bhaba National Institute

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by **Aloke Kumar Das** entitled **Top Quark Analysis and Trigger Studies with CMS Run 2 Dataset & Outer Tracker Upgrade for HL-LHC** and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

Chairman - Prof. Sanjay Kumar Swain

Guide / Convener - Dr. Prolay Kumar Mal

Examiner - Prof. Partha Konar

Member 1 - Prof. Bedangadas Mohanty

Member 2 - Dr. Subhasish Basak

Member 3 - Prof. Subir Sarkar

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to HBNI.
I/We hereby certify that I/we have read this thesis prepared under my/our direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date :

Place : **Signature**

Guide

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Aloke Kumar Das

DECLARATION

I hereby declare that I am the sole author of this thesis in partial fulfillment of the requirements for a postgraduate degree from National Institute of Science Education and Research (NISER). I authorize NISER to lend this thesis to other institutions or individuals for the purpose of scholarly research.

Aloke Kumar Das

List of Publications arising from the thesis

Journal

1. CMS Collaboration, "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**, 12, 180.

2. CMS Collaboration, "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s}$ = 13 TeV", JINST, **2020**, 15, 10017.

Chapters in books and lectures notes

1. "Measurement of the cross section $pp \rightarrow t\bar{t}\gamma$ process at $\sqrt{s}$ = 13 TeV for Run 2 in LHC in semi-leptonic channel", XXIV DAE-BRNS High Energy Physics Symposium 2020, NISER, Jatni, December 2020

Conferences

1. "Measurement of the cross section $pp \rightarrow t\bar{t}\gamma$ process at $\sqrt{s}$ = 13 TeV for Run 2 in LHC in semi-leptonic channel", XXIV DAE-BRNS High Energy Physics Symposium 2020, NISER, Jatni, December 2020

2. "CMS Top Results", NUCLEUS-2021. Nuclear physics and elementary particle physics. Nuclear physics technologies, Sep 20-25, 2021.

Aloke Kumar Das

DEDICATIONS

Dedicated to my parents, family and friends and, of course, all my teachers who have inspired me to come this far.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my parents for their unwavering love, sacrifice, and support throughout my academic journey. Despite the financial challenges that they faced, they worked tirelessly to provide me with the best possible education and opportunities. Their selflessness and determination have inspired me to persevere in the face of adversity and to pursue my dreams. I would also like to extend my heartfelt gratitude to my advisor, Dr. Prolay Kumar Mal, for his unwavering support in my research, patience, motivation, and vast knowledge. His advice was invaluable throughout the research and writing of this thesis. Aside from my advisor, I'd like to thank the other members of my thesis committee for their insightful comments and encouragement, as well as the questions that helped me broaden my research from various perspectives. I'd also like to thank my research collaborator Dr. Daniel Noonan, Dr. Nabin Poudyal, and Dr. Lukas Lechner for their invaluable assistance with the analysis. I also want to thank Mr. Deepak Kumar Pattanaik and Mr. Prafulla Saha for their support in the wire bonding task. I am deeply grateful to my wife to my wife for her unwavering love and support. Her belief in me has been a constant source of motivation and strength throughout this journey. Last but not least, thanks to all my lab mates who always encouraged me and helped me during my hard time.

ABSTRACT

The top quark is the heaviest known elementary particle and plays a special role in the dynamics of fundamental interactions. With the large amount of proton-proton collisions data at the center-of-mass energy 13 TeV at the Large Hadron Collider (LHC), the statistics for the top quark pair production ($t\bar{t}$) also increase compared to previous experiments like Tevatron. This higher statistics allows us to run more precise measurements with lower errors. At the LHC, the pairs of top quarks are predominantly produced through strong interactions - gluon-gluon fusion and quark-antiquark annihilation. In addition, a photon(γ) can originate in the final state through initial and final state radiations involving an additional electroweak vertex. Furthermore, due to the presence of an additional electroweak vertex the charge-asymmetry in $t\bar{t}\gamma$ events is predicted to be enhanced when compared with inclusive top quark pair production. Therefore, studying the top-antitop pair production in association with a photon with LHC data can lead to a thorough scrutiny of the Standard Model (SM) predictions. The first part of this thesis presents the cross sections measurement of $t\bar{t}\gamma$ process using the 137 fb^{-1} of proton-proton collisions data recorded with the Compact Muon Solenoid (CMS) detector during the LHC Run 2. For these measurements, the events selection criteria were set to select an isolated, highly energetic lepton, at least three jets originated from quark hadronization, among which at least one should be b-tagged and have one isolated photon with transverse momentum greater than 20 GeV and pseudo-rapidity $|\eta| < 1.4442$. In addition to cross section measurements, the charge-asymmetry measurements using the reconstructed top quark kinematics also have been performed. In the next part, the performance studies for the CMS trigger system using the Run 2 data have been presented. The last part of this thesis presents the optimization studies related to the ultrasonic wire bonding in connection with the CMS Phase 2 outer tracker upgrade.

Contents

Summary

After the Higgs discovery, the field of experimental particle physics is at a crucial juncture where the particle content of the Standard Model (SM) is complete without any evidence for the new physics through collider experiments. However, it is quite evident from the astrophysical observations and neutrino experiments that the SM cannot be a complete theory; at the most, it can be a low energy approximation of a more complete theory. In such a situation, the new physics signature search can be pursued either through direct searches or by probing the tiny SM predictions through precision measurements. In this thesis, the latter approach has been adapted by pursuing a measurement of the photon associated $t\bar{t}$ pair production process. The CMS recorded dataset corresponding to 137 fb^{-1} of integrated luminosity at the center-of-mass energy ($\sqrt{s}$) of 13 TeV has been utilized to measure the $pp \rightarrow t\bar{t}\gamma$ cross section in semileptonic channel. The inclusive cross section value in a fiducial region defined by $\geq$3 jets and $\geq$1 b-tagged jets, one tight high p_T lepton (e/μ) and exactly one photon with transverse momentum $p_T(\gamma) > 20$ GeV and jet multiplicity greater than three is measured to be $798 \pm 7\ (\text{stat}) \pm 48\ (\text{syst})$ fb, in good agreement with the SM predictions at the next-to-leading order calculations. For a better understanding of the effects originating from the presence of additional photon vertices, the $t\bar{t}$ reconstructed charge asymmetry (A_C) has also been measured, and the inclusive value of the same is measured to be $-0.003 \pm 0.034\ (\text{stat})$. In addition, the differential distribution of charge asymmetry is presented as well.

As a part of detector performance studies, CMS Level-1 trigger rates have been studied, and to monitor the trigger rates during LHC Run 2 and beyond, a graphical representation tool has been developed. Furthermore, the performance efficiencies for electron/photon HLT triggers have been studied using the Run 3 simulation, and the results have been tallied with the ones from Run 2. For HL-LHC Phase 2 tracker upgrade being carried out at NISER, the ultrasonic wire bondings are to be established to connect the silicon sensor bond pads with the ones on front-end hybrids. A detailed optimization procedure for the bond parameters on an automatic bonding machine has been presented in this thesis, along with the quality control checks through wire pull test and wire deformation.

List of Figures

List of Tables

Chapter 1

Introduction

Starting from the times of ancient Greek and Indian philosophy, mankind has always been inquisitive about the building blocks of matter. During our scientific journey over many millennia, many milestone theories, experiments, and subsequent observations have enriched our knowledge on it. Mendeleev's periodic table for atoms, Rutherford's discovery of atomic nuclei, understanding of quantum interactions at the subatomic level and the corresponding interpretations, the discovery of numerous cosmic ray particles, quark models for the hadrons, and last but not the least, the discovery of elementary particles through series of collider experiments are few to name in this long journey. In addition, the astrophysical observations and the cosmological models about the universe have further supplemented our understanding towards the ultimate building blocks of matter. To the best of our understanding till date, over 95% of the universe is constituted by unknown dark matter and dark energy, while only 5% is contributed by the visible matter, popularly called the visible universe. The Standard Model of particle physics summarizes the dynamics of the visible matter, encompassing present theoretical understanding and experimental observations.

1.1 Standard Model of Particle Physics

The fundamental particles constituting the visible matter are characterized by spin, mass, and quantum numbers (charges) which determine the nature of their interactions. The Standard Model of Particle Physics is presently the unique theoretical framework that describes the interactions of all experimentally observed fundamental particles of the visible universe

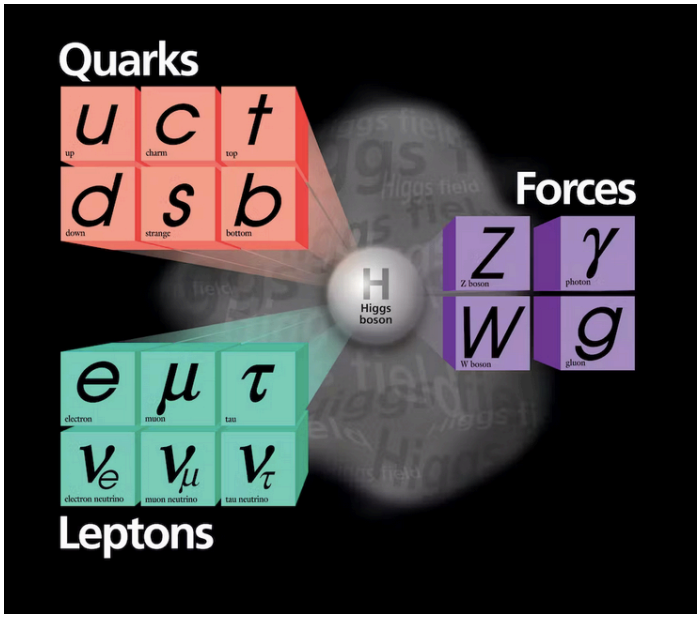


Figure 1.1: The Standard Model of particle physics showing the fundamental building blocks of matter and their interactions.

(see Fig. 1.1). Amongst the two broad categories of SM particles, the fermions (particles with half-integral spin) are classified into three families with identical quantum numbers but have different masses, as observed experimentally. The unstable heavier families subsequently decay into the lightest one, which constitutes the matter around us. Each "generation" consists of four fermions - two quarks and two leptons, each of which reflects their distinctive features through strong and electromagnetic interactions. Here in each family, the two quarks are labeled as "up" and "down" quarks, respectively, based on their 2/3 and -1/3 electromagnetic charges (in units of electron charge). On the other hand, the leptons carry -1 and 0 units of electromagnetic charge and are labeled as charged leptons and neutrinos, respectively. The SM neutrinos only interact through weak interactions and remain neutral to the electromagnetic and strong interactions. While neutrinos being the lightest ones, the masses of the SM particles span over a wide range starting from the sub-eV level neutrino mass to the 1.7×10^{11} eV level top quark mass. Based on their masses, the fermions of different generations are hierarchically separated, and each of them can have

specific chiralities. Under the massless approximation, the fermion chiralities coincide with the corresponding helicities (left-handed and right-handed), and they get conserved. Theoretically, Dirac spinors consisting of two components of different chiralities are required to describe the massive charged fermions, while neutrinos are observed only with the left-handed chirality.

The SM interactions covering strong, electromagnetic and weak forces[1] are characterized by the exchange of four spin 1 vector bosons – the gluons (strong interaction), the photons (electromagnetic interaction), and the W and Z bosons (weak interaction). The photon and the gluons are massless, while the massiveness of the W and Z bosons defines the scale of weak interaction at lower energy (suppressed by powers of $E/M_{Z/W}$). The symmetries related to the Parity P, charge conjugation C, and their combination CP are not respected in weak interaction, while they are all preserved through strong and electromagnetic interactions. In the process, the heavier fermions decay to lighter families through weak interactions, and the arguments are quite relevant at the low energy.

However, above the electroweak scale[2] the fermions with the left-handed chirality components become indistinguishable and are unified into the electroweak doublets. In other words, at the electroweak energy scale, such an "electroweak" symmetry is broken. Similarly, the unification of the electroweak force with the QCD defines the Planck scale, known to be one of the fundamental scales of nature at present. The electroweak symmetry breaking, or more popularly known as spontaneous symmetry breaking, had been proposed [1] [2] [3] to generate masses for the fermions and massive gauge bosons (W

[1] The gravitational force is quite feeble compared to the rest at the length scales of our interest and hence it is safely ignored within the SM

[2] A scale, above which a higher degree of symmetry is expected and the electromagnetic and weak interactions gets unified into the "electroweak" interaction.

and Z) having proportionality with the electroweak scale. With the discovery of W [4] and Z [5] bosons, the experimental evidence for spontaneous symmetry breaking has been affirmed. However, until the 2012 culminating discovery of the Higgs boson by the Large Hadron Collider (LHC) experiments [6] [7] mechanism triggering the spontaneous symmetry breaking was unknown. As established now experimentally, the Higgs mechanism [8] postulates to have a spin 0 field, called Higgs field, and the SM particles interact with the Higgs field to acquire their masses. The Higgs field being a quantum field must posit a physical particle called the Higgs boson, and the mass of a SM particle is directly proportional to its coupling to the Higgs boson. Thus, over the years, the spontaneous electroweak symmetry breaking mechanism has been affirmed through subsequent discoveries of Z, W and finally the Higgs boson. In addition many other experimental observations through cosmic ray and collider experiments have led to the present theoretical formulation of the SM.

While many physicists believe that the SM is quite incomplete or, at the most, a low energy limit of a more fundamental theory, the precision of the SM predictions has been successfully tested at an impressive level of accuracy. With an enormous range of applicability and validity, no significant deviations so far have been observed experimentally. With the unprecedented volume of data being collected by the LHC experiments, the SM predictions may withstand much deeper scrutiny or, even better, may lead to the observation beyond the SM (BSM) physics signature. This thesis is a small attempt to probe such SM precision through one of the processes involving top quarks.

1.2 Top Quark

In 1995, the discovery of the top quark was made by the CDF [9] [10], and DØ experiments [11] [12] , at Fermilab's Tevatron collider. Within the SM, it is the weak-isospin partner of the bottom quark with electric charge Q = +2/3 and T_3 = +1/2. At the LHC, the top quarks get produced through strong and electroweak interactions. Till date, it is the heaviest known elementary particle with a world average value for its mass equal to 173.34 ± 0.76 GeV [13]. The top quark Yukawa coupling to the Higgs boson being close to unity makes it more interesting. It is believed to have a special role in the dynamics of electroweak symmetry breaking. Being heavy, it decays within a time-scale of 5×10^{-25} s which is even shorter than the hadronization time-scale. Therefore, within the SM, the top quark remains to be the only quark that provides the rare opportunity to study a bare quark. This opens up the scope to measure the top quark spin, mass, width, couplings, etc. through its decay products directly.

1.3 Top quark pair production at the LHC and decay

The high energy protons are composite objects consisting of a number of point-like constituents, quarks, and gluons jointly termed as "partons". In general, at the hadron colliders, two constituent partons from the protons collide together to result in high transverse momenta events. As shown in Fig. 1.2, the inclusive $t\bar{t}$ production cross section can be decomposed into two parts applying QCD factorization theorem [14]: partonic cross section ($\sigma_{ij\to t\bar{t}}$), and the product of probabilities ($f_i(x_1, \mu_F)$, $f_j(x_2, \mu_F)$) for the partons (i, or j) to carry certain longitudinal momentum fractions (x_1 or x_2) of the initial hadron momenta at the factorization scale, μ_F. The partonic cross sections are generally calculated using perturbative QCD summed up to certain orders, while the latter represents the long-distance effect through the probability densities, called Parton Distribution Functions (PDF) at the

given energy scale. The PDFs are generally obtained by combining fits on a large number of data points from different experiments. Fig. 1.3 represents the exemplary plots for proton PDFs for two different μ_F values from the NNPDF [15].

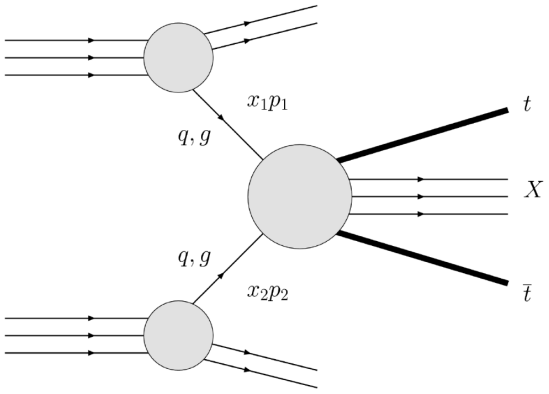


Figure 1.2: Illustration of top quark pair production in proton-proton collision.

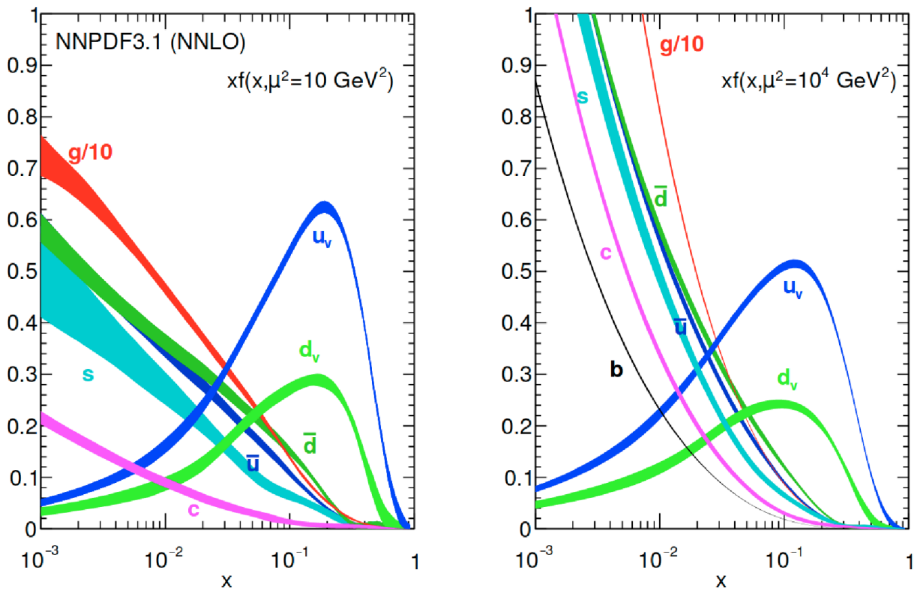


Figure 1.3: Parton distribution functions at two different values of factorization scale factors: $\mu^2 = 10$ GeV^2 (left) and $\mu^2 = 10^4$ GeV^2 (right) from NNPDF collaboration. Plot has been taken from [15]

Thus, the inclusive pp $\rightarrow$ t$\bar{\text{t}}$ production cross section at centre-of-mass energy ($\sqrt{s}$) can

be expressed as,

$$\sigma_{pp\to t\bar{t}}(s, m_t^2) = \sum_{i,j=g,q} \int\int dx_1 dx_2 f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \sigma_{ij\to t\bar{t}}(\hat{s}, m_t^2, \mu_F^2, \mu_R^2), \quad (1.1)$$

where $\hat{s} = x_i x_j s$ is the effective parton-level center-of-mass energy with a threshold value of $4m_t^2$, m_t being the top quark mass. The perturbative QCD term, $\sigma^{ij\to t\bar{t}}$ expresses here the cross section of i-th and j-th partons to produce top-antitop pair, as a function of $\hat{s}$, the coupling constant ($\alpha_s(\mu_F^2)$), and the renormalization scale (μ_R).

At the parton level, the top pair production occurs through two different mechanisms: the quark-antiquark annihilation or $q\bar{q} \to t\bar{t}$, and the gluon-gluon fusion or $gg \to t\bar{t}$. The lowest order Feynman diagrams for these processes are illustrated in Fig. 1.4. In contrast with the $t\bar{t}$ production at Tevatron, [3] the gluon-gluon fusion is the dominant production mechanism at the LHC. Here, the relative fraction for the gluon-gluon fusion can run up to 90% for $\sqrt{s}$=13-14 TeV.

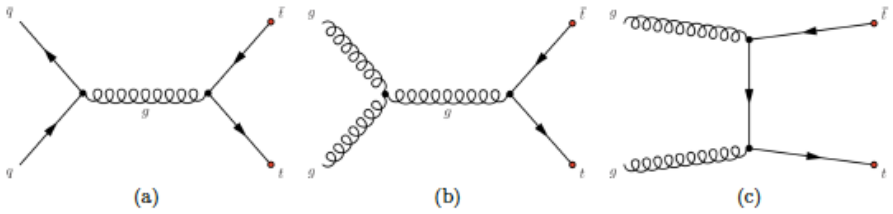


Figure 1.4: Dominant parton-level $t\bar{t}$ pair production modes via the strong interaction (a): quark-antiquark annihilation, and (b)-(c): gluon-gluon fusion.

The inclusive top quark pair production cross section has been calculated at NNLO, including resummation of Next-to-Next-to-Leading-Logarithmic (NNLL) [16] at $\sqrt{s} = 13$ TeV, while most recent CMS measurements [17] lead to 791 ± 25 pb. The latest cross

[3] At Tevatron proton and antiprotons are collided together at a center-of-mass energy of 1.8 and 1.96 TeV respectively during Run 1 and Run 2. Typically 85% of the top pair production is contributed by the $q\bar{q}$ annihilation processes.

section measurements performed at $\sqrt{s} = 13$ TeV are summarized by the LHCTop Working Group [18] in Fig. 1.5.

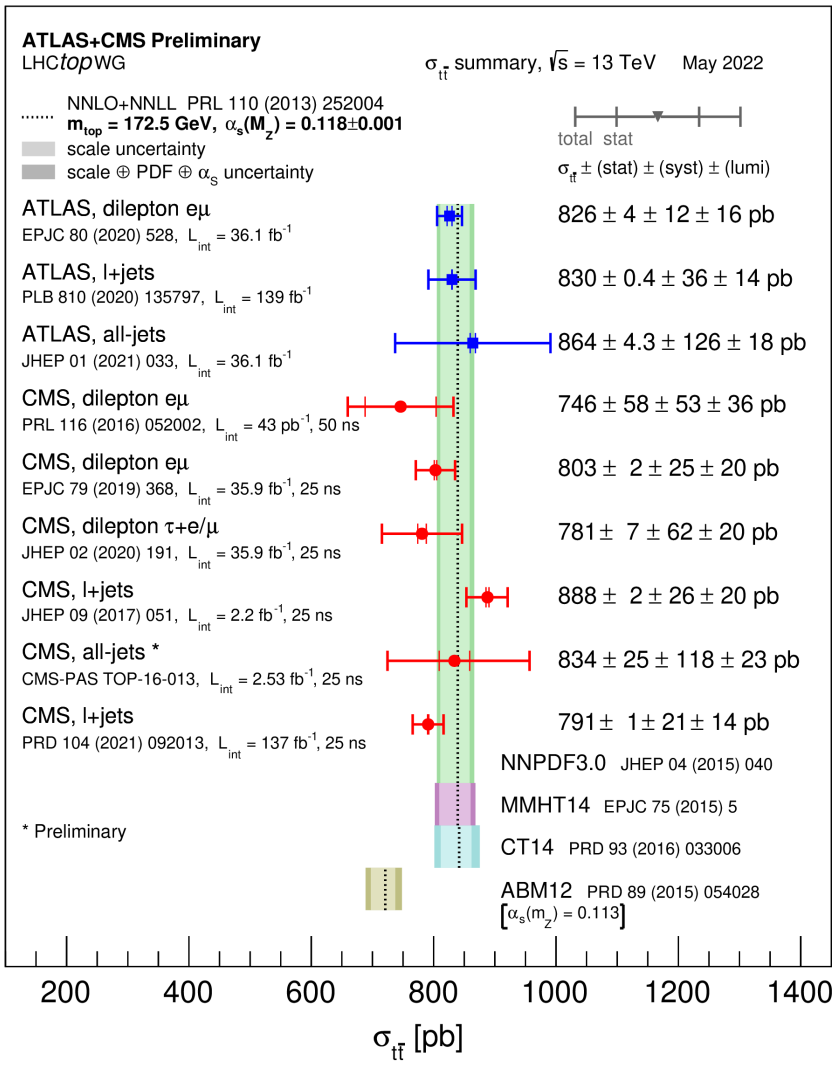


Figure 1.5: Summary of cross section measurements of $t\bar{t}$ production [18]

The decay probability of the top quark into a W boson and a certain down-type quark (d, s or b) is constrained by the corresponding Cabibbo–Kobayashi–Maskawa (CKM) matrix element squared, $|V_{tj}|^2$ with i denoting a down-type quark. As the $|V_{tb}| \gg |V_{ts}|, |V_{td}|$, the top quark decays to a W boson and a b quark with 100% branching ratio. The W boson

decays 33% of the time through charged leptons and neutrino and 76% through hadrons. Therefore, depending on the decay modes of the W boson, final state signatures for the $t\bar{t}$ pair produced events can be categorized into three channels: dileptonic, semileptonic, and all-hadronic. When both W bosons decay to leptons and neutrinos, it falls under the dilepton category with roughly 9% branching ratio (BR). The decay mode in which one of the W bosons decays leptonically and the other decays to hadrons is known as the semileptonic channel. Conversely, if both W bosons decay hadronically, then it results in the all-hadronic or all-jet channel. The branching ratios and different decay modes are shown in Fig. 1.6. This thesis considers only the semileptonic final state signature involving $t\bar{t}$ production as detailed in Chapter 3.

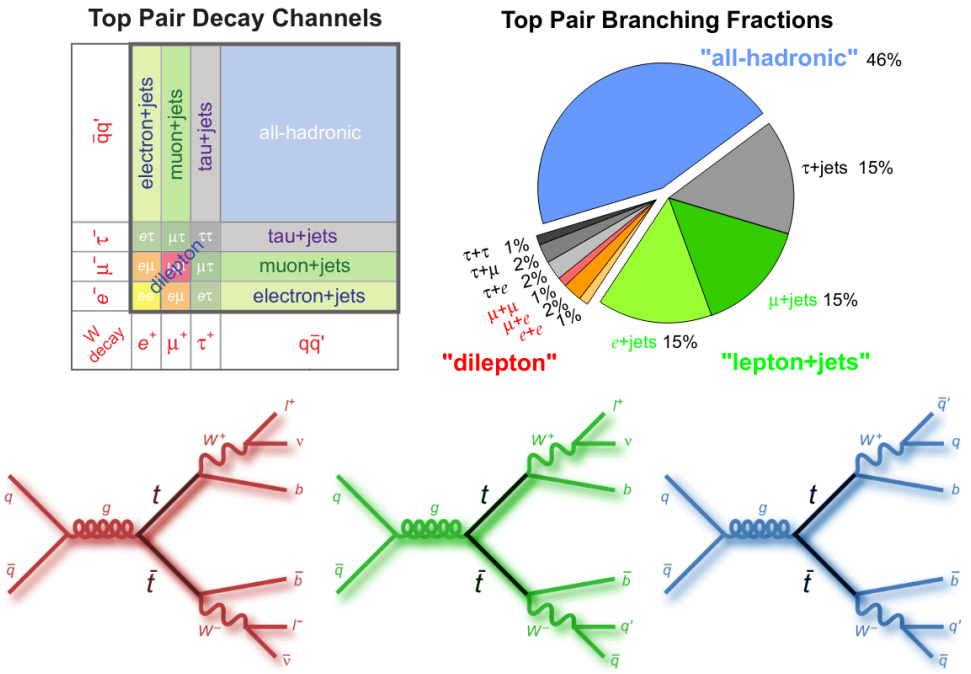


Figure 1.6: A schematic classification of the top quark pair decay channels (upper left) is shown together with distribution of corresponding branching ratios (upper right; numbers are approximate) . In the lower row, examples of LO Feynman diagrams are shown for the $t\bar{t}$ production with the subsequent decays in the dilepton (left), lepton+jets (middle) and all-hadronic (right) channels.

1.4 $t\bar{t}\gamma$ production at the LHC

The associated production of top quark pair with a photon ($t\bar{t}\gamma$) is also a likely scenario at LHC. The photon can get generated through different processes - initial state radiation (ISR) from one of the incoming quarks, or final state radiation (FSR) from top quarks, W bosons, or from the charged leptons. Two of the LO Feynman diagrams of the $t\bar{t}\gamma$ signal process in the semileptonic final state are shown in Fig. 1.7. Here, if the photon is radiated from the initial state quarks, the $t\bar{t}\gamma$ production can only occur through the $q\bar{q}$ annihilation process. In other words, the requirement of initial state radiation (ISR) photon is expected to enhance the probability of $q\bar{q}$ annihilation process over the gg fusion probability. Furthermore, at a given center-of-mass energy, the relative fraction of the two production modes ($gg \rightarrow t\bar{t}$ and $q\bar{q} \rightarrow t\bar{t}$) are driven by the PDFs. Thus, by the measuring the ratio of $t\bar{t}\gamma$ to $t\bar{t}$ events one can indirectly probe the PDFs. Such enhancement of $q\bar{q} \rightarrow t\bar{t}$ mode over the $gg \rightarrow t\bar{t}$ can also lead to the higher value of charge asymmetry as demonstrated in Chapter 3.

Due to the presence of the additional photon vertices, the production cross section for $t\bar{t}\gamma$ gets diminished by several orders of magnitude when compared with the inclusive $t\bar{t}$ production cross section. Therefore, a precise measurement of $t\bar{t}\gamma$ production cross section would also provide thorough scrutiny of the electroweak predictions.

Experimentally, CDF observed the $t\bar{t}\gamma$ production for first time during the Tevatron Run 2 ($\sqrt{s} = 1.96$ TeV) with a measured value of 0.024 ± 0.009 [19] for $\sigma_{t\bar{t}\gamma}/\sigma_{t\bar{t}}$, i.e., the ratio of the production cross sections of $t\bar{t}\gamma$ to $t\bar{t}$. During LHC Run 1, ATLAS observed a significance of 5.3 [20] for the $pp \rightarrow t\bar{t}\gamma$ process at $\sqrt{s} = 7$ TeV, resulting to $\sigma^{fid}_{t\bar{t}\gamma}$ $= 63 \pm 8\,(\text{stat})^{+17}_{-13} \pm 1\,(\text{lumi})$ fb. Subsequently, with the 2012 datasets both ATLAS and CMS had measured the fiducial cross sections in semileptonic channel with the respective values of $139 \pm 7\,(\text{stat}) \pm 17\,(\text{syst})$ fb [21], and $127 \pm 27\,(\text{stat} + \text{syst})$ fb [22] at $\sqrt{s} = 8$

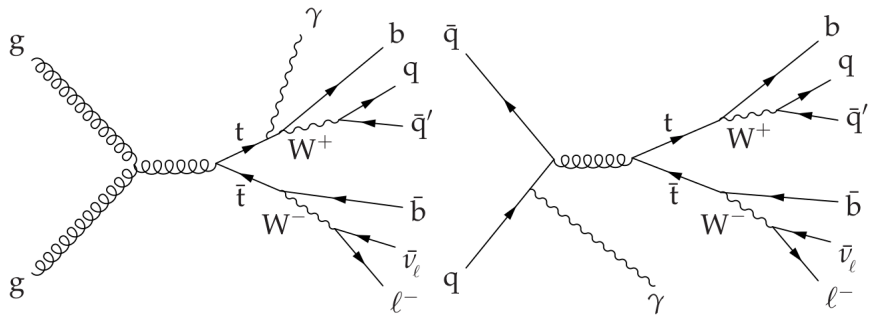


Figure 1.7: Examples of two LO Feynman diagrams for the $t\bar{t}\gamma$ signal process in the semileptonic final state where the $t\bar{t}$ pair is produced via gluon-gluon fusion with a photon emitted from one of the top quarks (left), and via quark-antiquark annihilation with a photon emitted from one of the initial partons (right).

TeV. During LHC Run 2, ATLAS used the 2016 dataset to measure the fiducial cross sections at $\sqrt{s} = 13$ TeV, and for both measured values of the fiducial cross sections are 521 ± 9 (stat) ± 41 (sys) fb and 69 ± 3 (stat) ± 4 (sys) fb [23] respectively for the single-lepton and dilepton channels. All the measurements performed so far by the ATLAS and CMS at different center-of-mass energy are found to be consistent with the theoretical predictions.

To complement the above measurements, a more refined analysis has been performed as a part of this work, as detailed in Chapter 3. Using the full Run 2 dataset have been utilized to measure the fiducial cross section along with the ratio with respect to the inclusive $t\bar{t}$ cross section. Furthermore, the same event selection, a very preliminary measurement for the $t\bar{t}$ charge asymmetry at the reconstruction level has been pursued.

1.5 CMS Trigger Studies

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 Tesla. Within the solenoid volume, there are silicon pixel and strip trackers, electromagnetic calorimeter (ECAL) made of lead tungstate

crystal and brass-scintillator based hadronic calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity (η) coverage provided by the barrel and endcap detectors. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid. In addition, it has a two level trigger system – Level 1 (based on detector hardware and FPGA-based signal processing) and the computing farm based triggering system or High Level Trigger (HLT). During LHC Run 1-3, the Level 1 event filtering is based on the information from the calorimeter and the muon system, while the HLT filtering is based on the reconstruction of the full event involving complex algorithms. The CMS Level 1 trigger system had been heavily upgraded during the technical stop of the LHC between Run 1 and Run 2 and in the initial days of LHC Run 2 operation, it was quite crucial to re-establish the bandwidth usage for different Level 1 triggers. As a part of this thesis work, codes for a Graphic User Interface (GUI) tool have been developed to monitor the rates for different types of inclusive triggers deployed at Level 1 of the triggering system. The transition from the LHC Run 2 to Run 3, would be a crucial one as the instantaneous luminosity would be increased rapidly during Run 3, while the radiation damage would affect the operations of the CMS detector. Here as a part of this thesis work, the efficiency studies related to the electron/photon triggers at the HLT have been performed. In particular, the performance checks and comparison studies for different triggering algorithms/conditions (called HLT filters) have been executed using simulated $Z/\gamma^* \rightarrow e^+e^-$ events, popularly called "tag-and-probe" method.

1.6 Phase 2 Tracker Upgrade & Module Assembly

The CMS detectors have been scheduled through numerous upgrade programmes for optimized performance considering the beam luminosities and radiation damage. During the phase 2 upgrade [24] for High Luminosity LHC (HL-LHC) operation, the Outer Tracker

(OT) of the CMS detector would be replaced completely with silicon detectors consisting of strip-strip (2S) and pixel-strip (PS) modules. The major goal for the Phase 2 upgrade is to cope with a large number of pileup events (up to 200 per bunch crossing) during HL-LHC without any degradation of the CMS physics performance. Here, the upgraded CMS OT modules must be designed for fast data processing along with the capability of track triggering at the hardware level. To match such a stringent requirement, the Phase 2 CMS OT has been segmented optimally to maintain enough granularity, and in the barrel region, the OT would consist of roughly 4464 2S modules. Amongst these, a large number of 2S modules are planned to be assembled, integrated, and tested at NISER, before being commissioned at the CMS cavern in 2026. For the 2S modules consisting of silicon strip sensors and readout ASICs (hybrids), a large number (exactly 4064) of microscopic wire bonding is required to readout the signals from the strips. These wire bondings are carried out through an automatic bonding machine, where the vibrational power at the ultrasonic frequency is applied. Since the bond pads on the sensor and hybrids have different properties (in terms of material and surface smoothness), it is quite crucial to tune the machine parameters for individual surfaces. A systematic optimization procedure for the machine parameters has been presented in this thesis to achieve the wire bonding quality as per the required specifications. The quality of wire bonding has been further cross checked through the pull testing of the wire bonds as well as through wire deformation measurements.

1.7 Thesis Organization

All the aforementioned works performed under the scope of this thesis have been organized through subsequent chapters as outlined below. With a brief introduction to the Standard Model of Particle Physics and the $t\bar{t}\gamma$ production processes at LHC here, Chapter 2, describes the experimental apparatus (the LHC and the CMS detectors), while the measure-

ments related to the $t\bar{t}\gamma$ analysis is covered in Chapter 3. Chapter 4 illustrates the works related to the CMS Trigger performance studies. The HL-LHC CMS upgrade programme, Outer Tracker Module Assembly activities as being performed at NISER are depicted in Chapter 5; as a part of this Ph.D thesis, the ultrasonic wire bonding related to the CMS modules are discussed in great detail in the last few subsections of the same chapter. Finally, Chapter 6 summarizes all the works presented in this dissertation, along with some concluding remarks.

Chapter 2

Experimental Apparatus

2.1 Large Hadron Collider (LHC)

The Large Hadron Collider (LHC) [25] is the most powerful particle accelerator built by mankind to study the fundamental interactions between elementary particles and to discover new physics beyond the Standard Model. It is located at the site of the European Organization for Nuclear Research (CERN) near Geneva, Switzerland. A schematic layout of the LHC along with its components is shown in Fig. 2.1. At the LHC accelerator complex, the protons are accelerated in multiple stages through a chain of linear/circular accelerators – Linear Accelerator (LINAC), Proton Synchrotron Booster (PS Booster), Proton Synchrotron (PS), Super Proton Synchrotron (SPS) and finally, the LHC. Before being injected into the LHC ring, the proton beams are accelerated to the energy of 160 MeV, 2 GeV, 26 GeV, and 450 GeV, respectively, through LINAC4, PS Booster, PS and SPS.

The final step of accelerating proton beams from 450 GeV to the final collision energy, i.e., 6.8 TeV per beam (the maximum beam energy during 2022 operation, although originally designed to be 7 TeV), is performed at the LHC, spanning over the 27-kilometer in circumference. It consists of a large variety of magnets, viz., dipoles, quadrupoles, sextupoles, octupoles, decapoles, etc., to synchronously circulate and focus the proton beams inside the LHC ring in 8 arcs and 8 straight sections. The LHC dipole magnets (a total of 1232) are the most technologically advanced electromagnets using the superconducting material (niobium-titanium, or NbTi) and are operated at a temperature of 1.9 K (i.e., –271.3 degrees of C) to produce intense magnetic field up to 8.33 Tesla. At the LHC, the pro-

ton beams are either circulated in twin aperture (dipoles) or in single aperture (final-focus triplet quadrupoles) high vacuum pipes, while the electromagnetic resonators are used to accelerate and compensate for the synchrotron energy losses. For each beam, eight superconducting radiofrequency (RF) cavities (in groups of four/cryomodule) deliver 2 MV (an accelerating field of 5 MV/m) of energy at 400 MHz.

The oppositely circulating proton beams are collimated and collided at four interaction points that mark the major experiments operated at the LHC: the two general purpose experiments ATLAS [26] and CMS [27], located diagonally opposite sides of the LHC tunnel, the LHCb [28] focussing on the flavor physics and rare decays (such as CP violation studies on the B-meson), and the ALICE [29] designed to investigate heavy-ion collisions and to study the quark-gluon plasma. The additional experiments, LHCf [30], MoEDAL [31] and TOTEM [32] have some special–purpose detectors that are respectively housed near the ATLAS, LHCb, and CMS experiments.

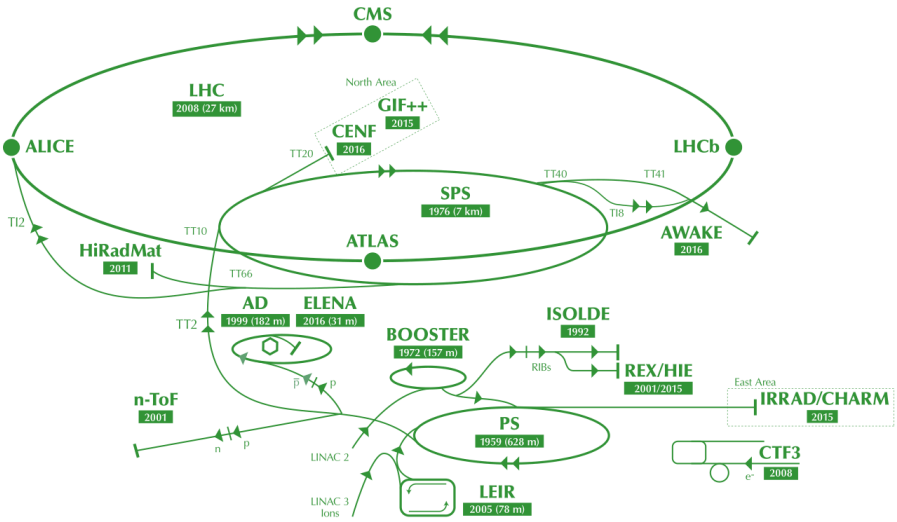


Figure 2.1: Schematic map of LHC and energy scale in different synchrotrons

2.1.1 LHC Parameters

For any collider experiment, the most important matrices are the beam energy and the beam luminosity, the latter being dependent on many parameters related to beam dynamics. At the LHC, the protons are bunched together, consisting of $\sim 10^{11}$ protons per bunch, and there can be up to 2808 proton bunches[1] in the LHC. Typically up to 72 subsequent bunches of protons are injected at the inter-bunch intervals of 25 ns to form a "*bunch train*", and a gap of 12x25 ns is maintained between two subsequent bunch trains. The quadrupole magnets, on the other hand, are utilized to squeeze the bunches vertically or horizontally to enhance the interaction probability at the beam crossing points. Typically, at the interaction points, the proton bunches have a transverse size of $\sim 17\mu$m. Furthermore, the magnets with higher order multipoles correct for the smaller imperfections in the field geometry and stabilize the beam. Overall, the instantaneous luminosity can be written in terms of the accelerator parameters [25] as:

$$\mathcal{L} = \frac{k_b N_1 N_2 f_{rev}}{4\pi\sigma_x^*\sigma_y^*} F = \frac{k_b N_1 N_2 f_{rev}}{4\epsilon\beta^*} F \tag{2.1}$$

where:

- k_b is the number of bunches per ring.
- N_i is the number of protons per bunch.
- f_{rev} is the revolution frequency of bunches.
- ϵ is the transverse beam emittance defines the closeness between protons inside a single a bunch in the phase space defined by position and momentum.

[1] During the 2010-12 operation (Run I) the LHC had been operated with 50 ns bunch spacing (1380 bunches).

- β^* amplitude function at the interaction point, determined by the magnets configuration, represents how much the beam squeezes in a short length.
- The widths of the horizontal and vertical beam profiles at the interaction point are denoted by $\sigma_x^*\sigma_y^*$ .where $\pi\sigma_x^*\sigma_y^* = \epsilon\beta^*$
- F is the symbol used to represent the reduction factor in geometric luminosity caused by the crossing angle at the interaction point.

In collider physics the potential interaction rate (dN_{events}/dt) for any process having a cross section (σ) is in general, directly proportional to the instantaneous luminosity $\mathcal{L}$,

$$\frac{dN_{events}}{dt} = \mathcal{L} \times \sigma \tag{2.2}$$

Therefore, the time-integrated instantaneous luminosity or simply the integrated luminosity can be expressed as

$$\mathcal{L}_{int} = \int \mathcal{L} dt, \tag{2.3}$$

which is a direct measure of the volume of dataset delivered by the LHC.

The first successful run at the LHC took place from March 2010 to February 2012 at a center-of-mass energy of $\sqrt{s}$ = 7 TeV, followed by an increase to $\sqrt{s}$ = 8 TeV in 2012; this operational period constituted the LHC Run 1. After the long technical stop (LS1), the LHC resumed operation at an increased center-of-mass of 13 TeV, marked as the Run 2, and was operational till 2018. Subsequent to another technical stop (LS2), LHC is presently in Run 3 operational mode since early 2022. During the Run 1 and Run 2 LHC has performed exceedingly well to achieve the beam luminosities better than the designed goals, reaching up to a peak instantaneous luminosity of $2 \times 10^{34}\text{cm}^{-2}\text{sec}^{-1}$. Beyond the Run 3, the LHC would undergo a major upgrade to operate at the peak instantaneous luminosity of $5 \times 10^{34}\text{cm}^{-2}\text{sec}^{-1}$ to the onset the High-Luminosity LHC (HL-LHC) era. Tab. 2.1 shows

some of the LHC operational parameters in 2018 in comparison with the designed ones, while Fig. 2.2 shows the schedule of the LHC so far along with the projected luminosities.

Table 2.1: Parameters of the LHC during Run 2 data-taking in 2018, compared to the design parameters [25].

Parameter name	LHC design value	2018 value
Beam energy (TeV)	7.0	6.5
Revolution frequency f_{rev} (kHz)	11.25	11.25
β^* IR 1/5 (cm)	55	30-25
Number of bunches k_b	2808	2556
Number of protons per bunch N_b (10^{11})	1.15	1.1
Bunch spacing (ns)	25	25
Typical emittance (μm)	3.75	1.9
Peak luminosity $\mathcal{L}$ ($10^{34} cm^{-2} s^{-1}$)	1.0	2.1

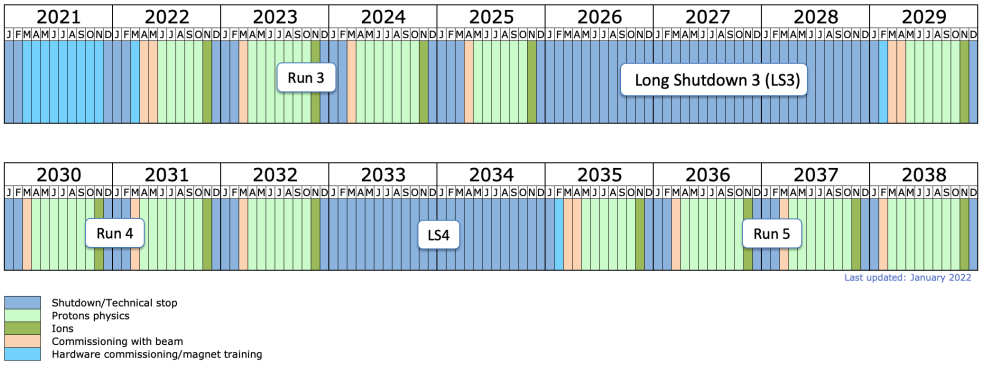


Figure 2.2: Longer term schedule for the LHC operations as in January, 2022 [33]. Note that the LHC Run 3 would continue till 2025, followed by LS3 in 2026.

2.2 The CMS Detector and Physics Object Reconstruction

The Compact Muon Solenoid Detector (CMS) is one of the two general purpose detectors situated at the interaction point 5 (IP5) on the LHC ring. This 14000-tonne giant cylindrical detector has an overall diameter of 15 m, and it is 28.7 m long. At the core of the CMS, it has the most powerful superconducting magnet producing a uniform solenoidal field of 3.8

Tesla almost over the entire volume of the tracker and calorimeters. The high and uniform magnetic field greatly helps to reconstruct the charged particle trajectories as well as to measure their momenta precisely. The return fields from the solenoid are captured through the iron yoke placed in the muon system, where the magnetic field (2 Tesla) helps to remeasure the muon momenta. Different parts of the CMS detector are shown in Fig. 2.3 and are discussed briefly in the next few sub-sections.

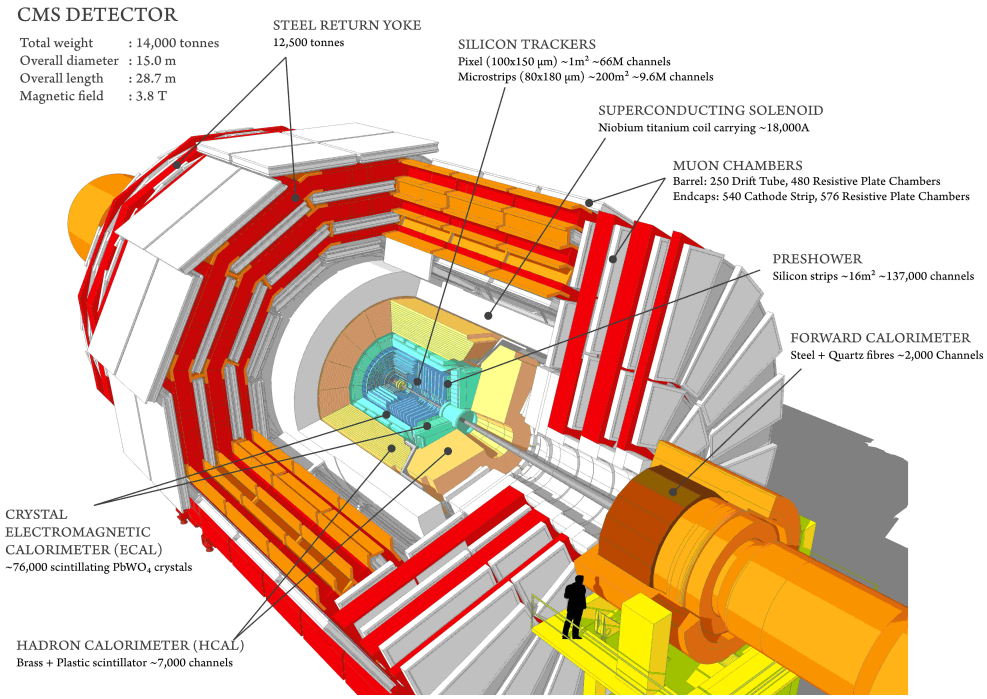


Figure 2.3: Schematic view of CMS detector showing Silicon tracker, electromagnetic and hadronic calorimeters, superconducting solenoid and the muon chambers.

2.2.1 Coordinate System

Considering the cylindrical geometry of the CMS detector, the Z-axis is considered to be the direction for the proton beams, while X-axis is chosen to be directed towards the center of the LHC ring. With the center of the CMS detector as the origin of the coordinate system, the remaining X-axis then points toward the surface to complete a right-handed Cartesian

coordinate system, as shown in Fig. 2.4. In this coordinate system, the charged particles originating from the nominal interaction vertex (i.e., from the origin of the coordinate system) are expected to have helical trajectories where the curvature due to the solenoidal field only occurs in the plane transverse to the beam axis, i.e., in XY-plane. Furthermore, the resultant initial state momenta from the colliding protons in the XY also nullify. Using the 4-momenta of the particle, one can define a Lorentz invariant observable, called the rapidity (y) as

$$y = \frac{1}{2} ln \left(\frac{E + p_Z}{E - p_Z} \right). \tag{2.4}$$

For highly relativistic particles, the rest mass of a particle can be ignored while considering the total energy, i.e., $E \approx p$. Under this approximation, the rapidity can be approximated as pseudorapidity

$$\eta = -ln \tan \left(\frac{\theta}{2} \right), \tag{2.5}$$

Where θ is the polar angle (angle between the particle trajectory and the Z-axis). Any point in the XY-plane can conventionally be identified with the azimuthal angle (ϕ) and the distance from the beam axis (R). Thus, the kinematics of an ultra-relativistic particle can be rewritten as

$$p_x = E \cdot \sin\theta \cos\phi \tag{2.6}$$

$$p_y = E \cdot \sin\theta \sin\phi \tag{2.7}$$

$$p_z = E \cdot \cos\theta \tag{2.8}$$

with $E = p$.

2.2.2 Silicon Tracker

Starting from the interaction point, the silicon tracking detectors [34] make up the innermost layer of the CMS experiment. The primary role of the tracker is to reconstruct the charged

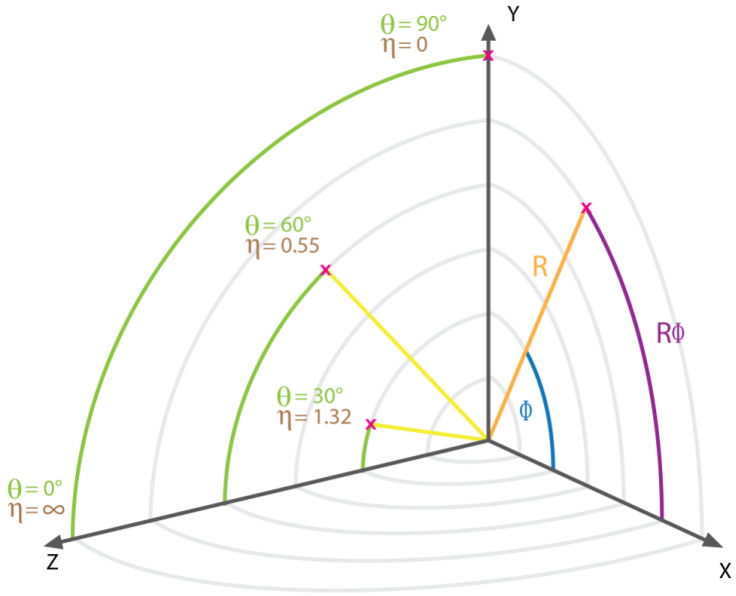


Figure 2.4: Illustration of coordinate systems for the CMS detector.

particle trajectories to measure their momenta. In addition, measurements from the inner tracker play a crucial role in the reconstruction of vertices. The active elements of the CMS tracker are the silicon pixel sensors in the center close to the beam axis and the silicon micro-strip sensors in the outer layers. In total, the silicon sensors cover an area of ∼200 square meters spanning up to a pseudorapidity of 2.5, as shown in Fig. 2.5.

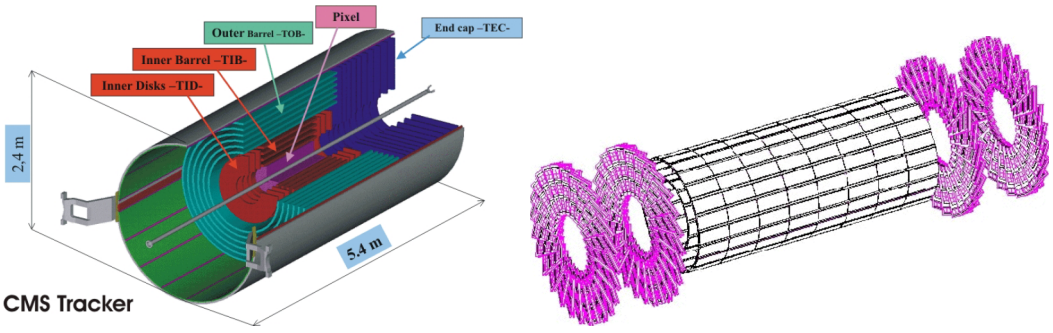


Figure 2.5: Layout of the CMS tracker (left) and the pixel detectors (right)[35].

In the closest vicinity of the beam pipe, 3-D hits are crucially identified through multi-

layer CMS pixel detectors which consist of millions of $100\mu m \times 150\mu m$ granular sensors along with the readout elements. Pixel hits, in particular, drive the precision on vertex position resolution, which in turn can affect the charge particle momenta measurements. However, due to the harsh radiation environment characterized by high track density, the pixel detector is prone to radiation damage. During the extended year-end technical stop in 2016/2017, the CMS Phase-1 pixel upgrade completely replaced the pixel detector with 4-layers of coverage in the barrel and forward regions. The strips are separated by 80 to 122 μm which gives a resolution between 23 μm and 53 μm in the direction perpendicular to the strips. The outer part of the CMS tracker consists of silicon strip detectors to measure one coordinate, i.e., in ϕ-direction. Overall, the CMS strip tracker consists of nearly 9.3 million active elements in a hollow cylinder length of 5.8 m and diameter of 2.4 m. It contains four subsections, as shown in Fig. 2.6.

- Four layers of tracker inner barrel (TIB) confined within diameter of 2.4 m.
- Three tracker inner disks (TID),
- Six layers of tracker outer barrel (TOB)
- Nine tracker endcap disks (TEC)

The CMS tracker has an average track reconstruction efficiency of 94% and 85%, respectively, for charged particles with transverse momenta $p_T > 0.9$ GeV in regions of $|\eta| \leq 0.9$ and $0.9 \leq |\eta| \leq 2.5$ where a combinatorial track finder algorithm was used. For tracks with $1 \leq p_T \leq 10$ GeV in the barrel region, the momentum resolution is approximately 1.5%. Additionally, the precision of determining transverse (longitudinal) impact parameter decreases from 90 μm (150μm) at p_T =1 GeV to 25 μm (45 μm) at p_T = 10 GeV.

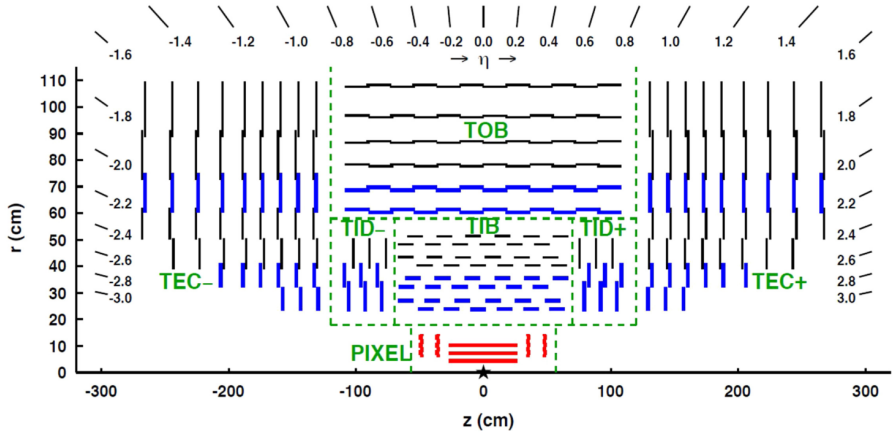


Figure 2.6: RZ-view of the CMS tracker [35]

2.2.3 Electromagnetic Calorimeter

The CMS Electromagnetic Calorimeter (ECAL) is designed to measure the energies of photons and electrons with high accuracy through electromagnetic showers. It is a hermetic homogeneous calorimeter made of 61200 Lead Tungstate ($PbWO_4$) scintillating crystals, each having the dimension of 2.2cm x 2.2cm x 23cm. The crystal dimensions roughly correspond to one Molière radius (R_M) in the lateral direction and over 25 radiation lengths (X_0) in the longitudinal direction. Such design and dimensions of the crystal have been driven by the requirement of 1-2% mass resolution for the $H \rightarrow \gamma\gamma$ peak. The CMS ECAL is divided into barrels (EB) and endcaps (EE) along with preshower detectors (ES), to discriminate the single photon showers from ones from the $\pi^0 \rightarrow \gamma\gamma$. The central region or barrel (EB) of the ECAL has an inner surface located 1.29 meters away from the interaction point, and it provides pseudorapidity coverage of $|\eta| \leq 1.479$. On the other hand, the two endcaps cover the pseudorapidity range of $1.479 \leq |\eta| \leq 3$ and are located at $|z| = 3.15$ meters away from the detector center. Fig. 2.7 shows a longitudinal cross section of the CMS ECAL while its energy resolution [27] can be represented by the following expression

$$\frac{\sigma(E)}{E} = \frac{2.8\%}{\sqrt{E}} \oplus \frac{0.12}{E} \oplus 0.3\% \tag{2.9}$$

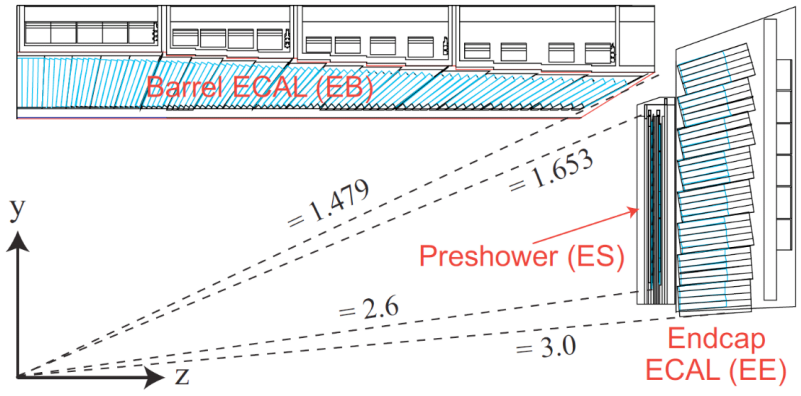


Figure 2.7: A longitudinal view of the of the CMS electromagnetic calorimeter with the respective pseudorapidity coverage [36].

Superclustering in the ECAL

The electron or photon reconstruction in the CMS ECAL is primarily based on the topological clustering algorithms collating the energy deposited in the crystals adjacent to the particle trajectory. Due to the presence of the beam pipe, tracker material and other support structures, early showering for the electrons and photons may occur. For example, electrons traversing through the upstream material of the CMS tracker, and under the 3.8T magnetic field, emit bremsstrahlung photons which would lead to electron-positron pairs. Similarly, the photons originating from the interaction vertex may also get converted into electron-positron pairs in the tracker volume; eventually both electrons and photons may have early showering before reaching to the CMS ECAL. In addition, the effective solenoidal field of 3.8T affects the shower formation to spread over several crystals around the primary ones. The CMS superclustering algorithms [37] aggregate the neighbouring crystal energies considering such effects, and all these clusters collectively form superclusters. Due to the high

magnetic field the low p_T electrons/positrons in the electromagnetic showers would propagate in helical trajectories leaving characteristic "mustache" shapes when projected in the $\Delta\eta$-$\Delta\phi$ plane. It is basically a region delimited by two parabolas in η (based on the η and energy of the cluster), and by a dynamic interval in ϕ considering the transverse energy of the cluster. Thus, purely based on such geometrical shapes of the electromagnetic showers, the mustache superclustering algorithm had been deployed to reconstruct the ECAL showers. The algorithm is optimized to contain 98% of the shower energy through tuning of spatial selection criteria. As illustrated in the Fig. 2.8, when a nearby cluster (passing the energy threshold criteria) falls inside the mustache-like zone around the main seeding cluster, its energy is aggregated. During the LHC Run 2, for all the CMS ECAL-based electron and photon reconstruction these refined superclusters have been utilized to improve the shower energy resolution.

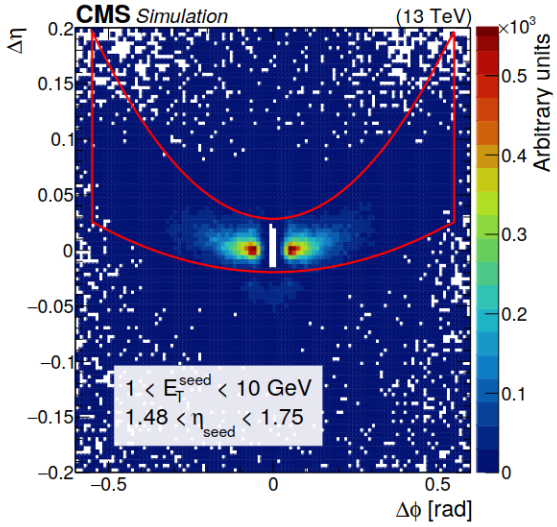


Figure 2.8: Distribution of $\Delta\eta = \eta_{seed-cluster} - \eta_{cluster}$ versus $\Delta\phi = \phi_{seed-cluster} - \phi_{cluster}$ for simulated electrons with $1 < E_T^{seed} < 10$ GeV and $1.48 < \eta_{seed} < 1.75$. The Z axis represents the occupancy of the number of Particle Flow (PF) clusters matched with the simulation around the seed. The red line contains approximately the set of clusters selected by the mustache algorithm [37]

2.2.4 Hadronic Calorimeter

The Hadron Calorimeter (HCAL), split into barrel (HB) and endcaps (HE), measures the energy of charged and neutral hadrons. HCAL can also complement the measurements from ECAL and muon systems for better particle identification of electrons, photons, and muons. It is a sampling calorimeter made of alternating layers of copper alloy absorber plates (5 cm thick in the barrel and 8 cm thick in the endcap) and a plastic scintillator as the active material. In the HB region, the 4 mm thick plastic scintillator tiles are read out using wavelength-shifting (WLS) plastic fibers. Considering the thickness of all the layers, HB depth is about six interaction lengths (λ) which is sufficient to contain most of the hadronic showers. Fig. 2.9 shows the different parts of the CMS HCAL along with the respective pseudorapidity coverage ranges. In addition to HB and HE, the forward hadronic calorimeter (HF) extends the HCAL hermeticity up to a pseudorapidity ($|\eta|$) of 5.2 for better measurement of missing transverse energy. The HF is made of quartz fiber and steel and is located outside the detector wheels. Furthermore, to ensure adequate sampling depth for the entire region of $|\eta| \leq 3.0$, additional scintillator tiles have been instrumented in the Outer Hadronic Calorimeter or HO.

In combination with the ECAL measurements, the HCAL energy resolution [38] for the pions is given by

$$\frac{\sigma(E)}{E} = \frac{84.7\%}{\sqrt{E}} \oplus 7.4\% \tag{2.10}$$

2.2.5 Muon Chamber

The muons can pass through the CMS detectors, and they are identified using the muons stations in the outer part of the CMS detector interleaved with the magnetized iron yokes. There are three different types of detector technologies, viz., drift tubes (DT), cathode strip chambers (CSC), and resistive plate chambers (RPC), utilized in different sections of the

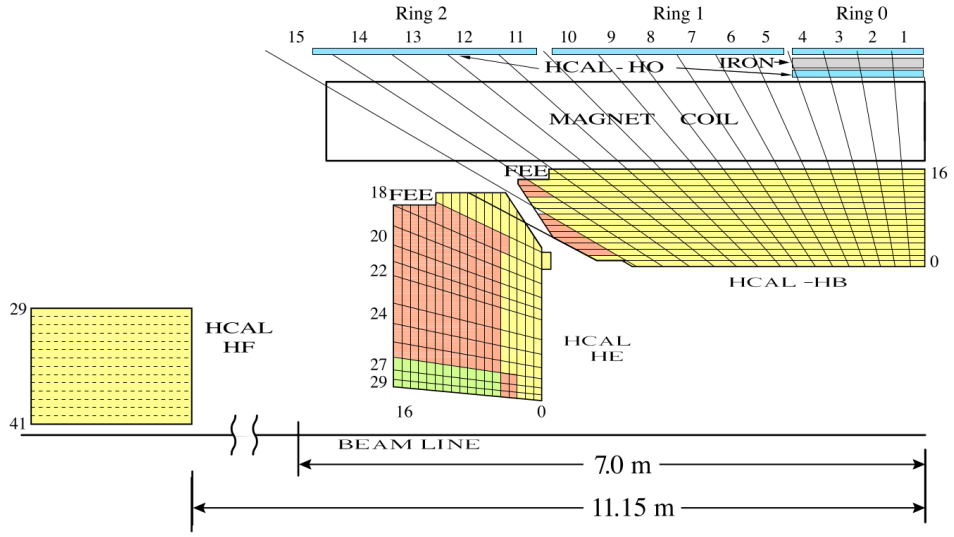


Figure 2.9: A longitudinal view of the CMS hadronic calorimeter depicting the pseudorapidity coverage[27].

CMS muon system. As the polarity of the return magnetic field in the muon system flips (with respect to the solenoidal field direction in the tracker volume), in CMS, the muon momenta are measured very precisely. A cross sectional view of the CMS muon system depicting all the muon stations is shown in Fig. 2.10.

- **Drift tubes (DT)** are 42 mm x 13mm x 2.4m tubes, each of which contains a stretched wire inside a gas volume, which act as an anode at a very high voltage (3600 V). When a muon or any charged particle passes through the gas volume (a mixture of 85% Ar and 15% CO_2 gases), it ionizes the gas atoms, and the electrons "drift." (up to a maximum drift time of 400 ns) towards the anode wire and induce a signal in the anode wire. As in Fig. 2.11, DTs are used to measure the coordinates in one dimension. While combining the measurements from multiple layers of the CMS DTs, a hit resolution better than 100 μm is achieved in the direction perpendicular to the DT wires.

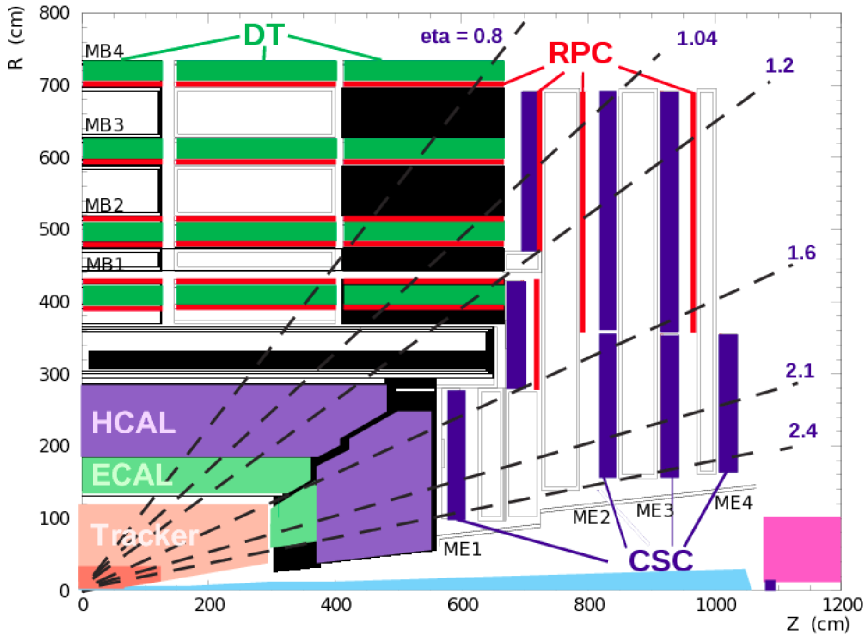


Figure 2.10: cross sectional view of CMS muon system showing the locations of DT, RPC, CSC [39].

- **Cathode strip chambers (CSC)** are used in the endcap region, where the background rates for muons and other charged particles are quite high. CSCs are basically trapezoidal-shaped multiwire proportional chambers where the anode wires (separated by 3.2 mm from each other) are stretched radially with the cathode strips (placed perpendicularly having pitches of 8.4-16 mm) as depicted in Fig. 2.12. Due to the muon passage through the chamber, the ionization in the gas mixture (50% CO_2, 40% Ar, and 10% CF_4) occur, and the electrons drift toward the anode wires. Due to the movement of electrons, the charges are induced to the cathode strips, and therefore reading out the signals from anode wires and cathode strips hit positions in two dimensions are measured in the CSC.

- **Resistive plate chambers (RPC)** are installed in both the barrel and endcap regions. As shown in Fig. 2.11, a narrow layer of appropriate gas mixture (96.2% $C_2H_2F_4$, 3.5

% C_4H_10 and 0.3% SF_6) is maintained between two high resistivity ($10^{10} - 10^{11}$ Ω-cm) bakelite plates were placed parallelly with a separation of 2 mm amongst them. After the passage of charged particles through the RPC, effective charge induction occurs on external strips within a few ns. Because of their fast responses, the RPCs in the CMS muon system is primarily utilized for triggering purpose.

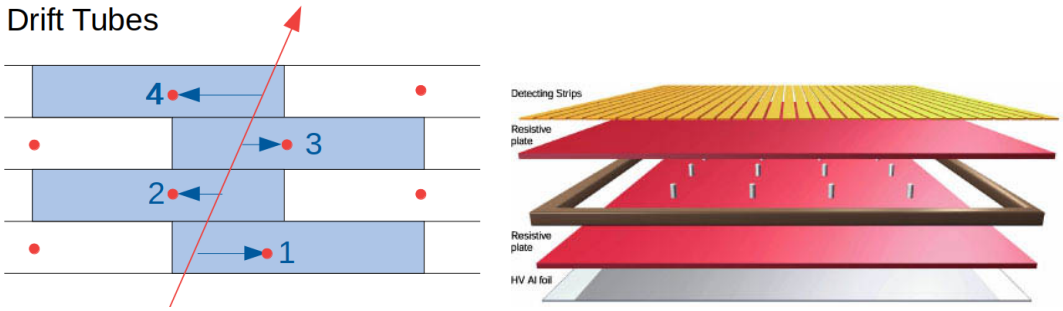


Figure 2.11: Geometry of Drift Tubes (left) and Resistive Plate Chamber (right) [40].

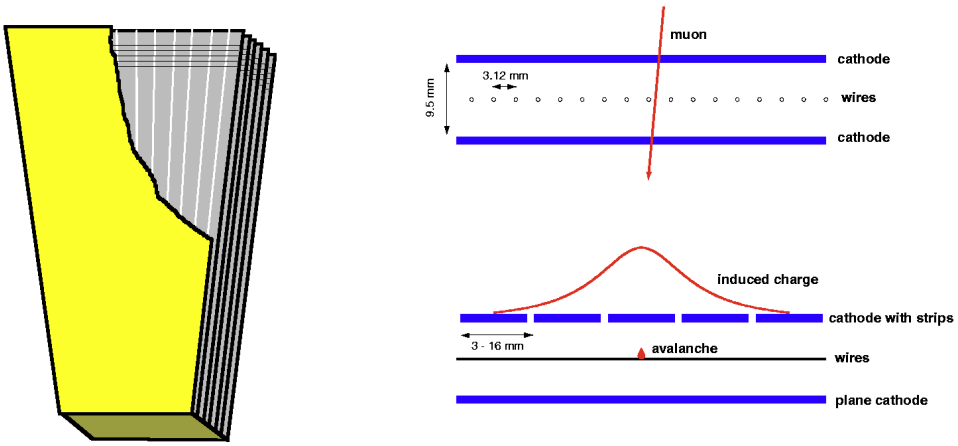


Figure 2.12: CMS Cathode Strip Chambers in the forward regions along with the muon track and induced charge graph [39]

The combination of the above-mentioned three technologies comprises the CMS muon system, which covers the region $|\eta| \leq 2.4$. Overall, the CMS muon system detects muons

with very high efficiency (94-99)%, using the return magnetic field of $\sim$2 T through the iron yokes.

2.2.6 Trigger Systems

The proton-proton collision frequency at the LHC is around 40 MHz, which roughly corresponds to 40 Terabytes of raw data generated per second. Because of the limited data storage capacities, it is quite necessary to save only a small fraction of the collision events that would be "interesting" for offline analyses. Moreover, it is to be ensured that the filtration for these "interesting" events must be performed rapidly and efficiently. The CMS trigger system performs such a real-time filtration of events through two levels of event triggering:

- Level-1 (L1) trigger: It is a hardware level triggering system where the very crude and fast information from calorimeter and muon systems are checked for the predefined trigger logics, while the detailed information from the precision detector components is awaited in the data buffers. The L1 triggers effectively reduce the events rates to 100 kHz for the next level of trigger checks.

- High-level trigger (HLT): It is a software based trigger system where a large number of CPUs are utilized for the preliminary event reconstruction. The reconstructed object kinematics and other characteristic features of the collision events are scrutinized for selecting events at the HLT level. The output events from the HLT are written into the storage tapes for offline analysis at a typical rate of 1 kHz.

2.2.7 Physics Object Reconstruction

Particle Flow Elements

The particle flow algorithm [41] forms the basis of offline event reconstruction. As particles travel through the detector, they generate signals that are digitized and locally reconstructed in the various subdetectors, including the inner tracker, outer tracker, electromagnetic calorimeter, hadronic calorimeter, and muon detector. The reconstructed particle flow (PF) elements are then merged to form candidates for particle-flow-objects such as photons, jets, electrons, muons, missing transverse energy, primary vertices, and secondary vertices.

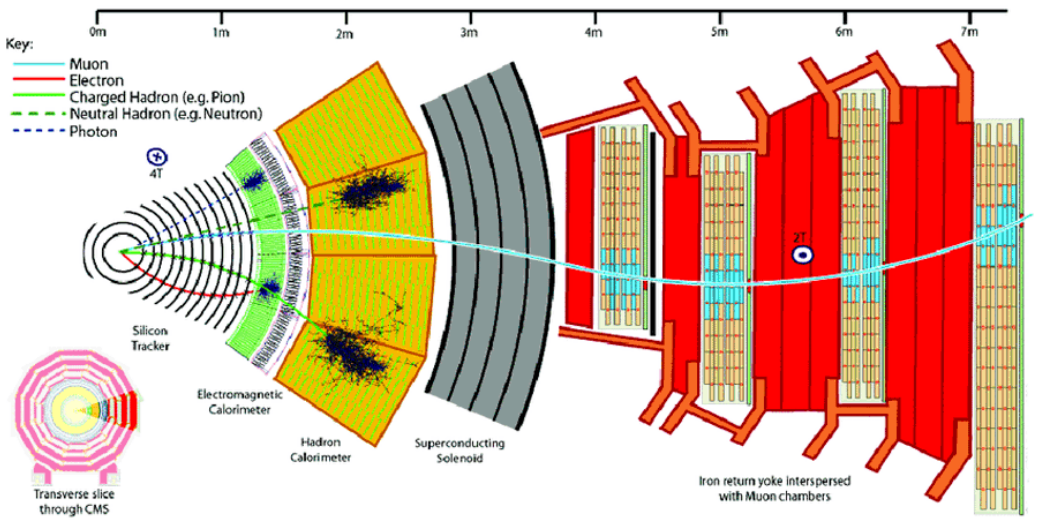


Figure 2.13: Different particles detection in CMS.

PF Algorithm

Particle flow algorithm uses all subdetector information to estimate the track of a stable particle. Hit information from the tracker, cluster information from the calorimeter, and muon detector information are linked by proper quality tests to reconstruct all stable particles.

PF electrons: When an electron passes through the detector material, it can radiate photons via Bremsstrahulung process. If a photon has sufficient energy, it can create e^+e^- pair, which may further produce photons. Thus electrons and photons give similar signatures in the calorimeter. Hence they are reconstructed together. To initiate an electron candidate, a gaussian-sum filter (GSF) track is first employed. Any track can served as seed which satisfies $p_T > 2$ GeV criteria. Similarly for photon candidate, seed ECAL supercluster is selected from tracks with $p_T > 2$ GeV 10 GeV and have no links to the GSF track. For both electron and photon candidates, the ratio of HCAL energy to ECAL energy must be less than 10%, where the total HCAL energy is measured within ΔR of 0.15 from the supercluster position in ECAL. As electrons and positrons may emit a photon while passing through tracker material, the total energy deposited in the ECAL is calculated in a supercluster with small η, but wide ϕ variables. This is because emitted photons travel in the direction of emission, while the magnetic field bends the electron in ϕ. To distinguish between electrons and photons, an isolation criterion is utilized, and various isolation techniques are applied to enhance the reconstruction efficiencies.

PF photons: The PF algorithm employs the cluster's shape and size in the ECAL to detect photons that have a transverse momentum greater than 15 GeV. Energy deposited by the radiating electrons and the converted photons are used for reconstruction. The process begins with a seed crystal, which detects energy exceeding the predefined threshold compared to neighboring crystals. The energy encompassed within a 5 x5 matrix surrounding the seed crystal is then used to reconstruct an unconverted photon. To identify photons, various criteria are utilized, including tracker isolation, ECAL isolation, hadron calorimeter isolation, hadronic to electromagnetic ratio, and R_9. The R9 variable represents the ratio of energy contained in a 3x3 matrix and the supercluster energy, with the R_9 shape distinguishing photons from other objects. For instance, the typical R_9 values for an isolated

photon are higher than π^0 and η meson decay.

Electron/Photon energy correction : Reconstructed energy for electrons and photons are recovered using superclustering algorithm which are susceptible to the losses for various reasons - inter-module gaps, shower leaks, energy loss in the tracker etc. Therefore without any in-situ correction the reconstructed e/gamma energy from the CMS ECAL is expected to be less than the corresponding energy reconstructed in test beam configuration. However, using the $Z \rightarrow ee$ collision events [37] the energies of the electromagnetic showers are corrected as functions of electron p_T, η and R_9. Here the residual energy scale correction is determined through reconstruction of invariant mass from the dielectron system. In CMS such residual corrections (using $Z \rightarrow ee$) are observed to be different in data and Monte Carlo, and hence are applied differently.

PF muons: The CMS experiment has a dedicated muon system that consists of cathode strip chambers (CSC), drift tube (DT) and resistive plate chambers (RPC) placed outside the CMS solenoid magnet. Muon system first identify the standalone muon whereas PF global muons are reconstructed by combining information from the muon system and the tracker. The muon system provides a precise measurement of the muon momentum, while the tracker helps to determine the muon trajectory and identify any associated tracks. In CMS muons with transverse momentum less than 200 GeV, have very good momentum resolution. Very low transverse energy muons might have very less hit in muon system but can well reconstructed using tracker. When an extrapolated track from tracker matches with the muon system hits, it is called tracker muon. The minimum transverse and total momentum for a muon track candidate are 0.5 GeV and 2.5 GeV for identifying it as a tracker muon.

Jets Reconstruction :

The quarks and gluons, when produced through pp collisions or from the decays of other particles, they immediately hadronize, producing a spray of "color"-less neutral and charged particles, known as a "jet". As these particles pass through the CMS detectors, they lead to the formation of showers in the HCAL and occasionally in the ECAL, while the charged particles in the jet would also leave the trajectories in the tracker volume. The goal of jet reconstruction is to combine the showers from these particles to obtain the original energy of the quarks/gluons. CMS adapts a jet reconstruction algorithm called anti-k_T algorithm that relies on the PF candidates seeded by the tracker for better energy resolution of the jet candidates. Nearby charged and neutral hadronic candidates are clustered together using a parameter that is inversely proportional to the square of each candidate's p_T. A distance parameter, d_{ij} (d_{iB}) can be defined between i-th particle and the pseudojet (and the beam) as

$$\begin{aligned} d_{ij} &= min\left(\frac{1}{p_{T_i}^2}, \frac{1}{p_{T_j}^2}\right)\frac{\Delta_{ij}}{R_{ij}} \\ d_{iB} &= \frac{1}{p_{T_i}^2} \end{aligned} \tag{2.11}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, with p_{T_i}, y_i and ϕ_i are respectively the transverse momentum, rapidity, and azimuthal angle of the particle i. The particles with high transverse momentum are clustered first before the low momentum particles since the controlling the distance d_{ij} is estimated using the inverse of the particle transverse momenta, In the $y - \phi$ plane, the jets with high transverse momentum reconstructed through the anti-k_T algorithm exhibit perfect circularity. For nominal jet reconstruction in CMS, the distance parameter is$R$=0.4. Figure 2.14 is an an example of the look of the jets reconstructed by anti-k_T reconstruction. Prior to the reconstruction of jets, any charged hadrons that are associated with a vertex other than the primary or secondary vertex of the event are eliminated from the jet.

This procedure, known as charged hadron subtraction, removes tracks originating from additional pileup collisions. The energies of the reconstructed jets are further corrected using the appropriate collision dataset.

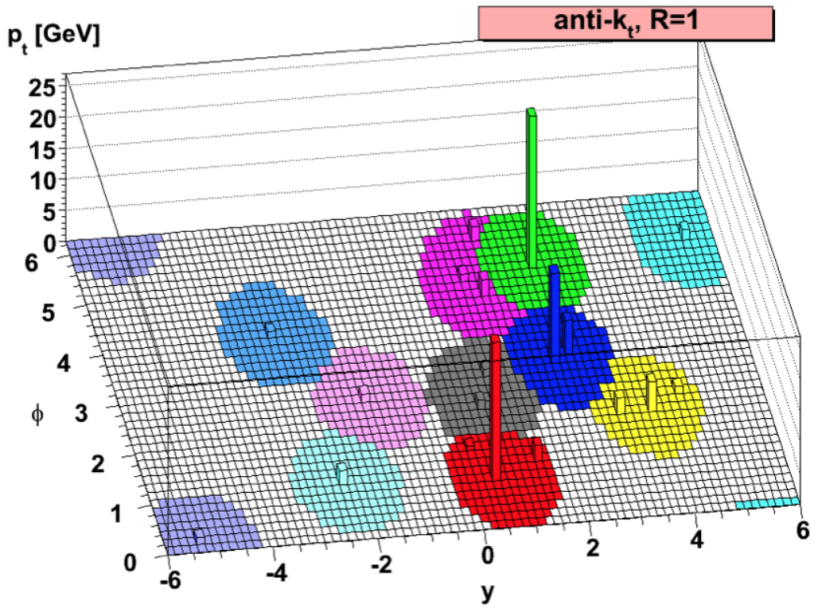


Figure 2.14: Jet reconstruction using anti-k_t clustering algorithm [42]

Inside the spray of particles originating from the heavy quarks (bottom/charm quarks), some particles (e.g; b-, c-hadrons) can travel considerable distance from their interaction point (Primary Vertex, or PV) before they decay. Exploiting this property, the jets from the b/c-quarks can be identified with the secondary vertex reconstruction, and/or impact parameter calculation for the constituent tracks inside the jets and these jets can separated from the light quark jets. To achieve the maximal b-tagging efficiency the observables related to the secondary vertex tagging and the impact parameter tagging are combined together using a deep neural network based methodology, known as the Deep Combined Secondary Vertex (deepCSV) tagging.

2.2.8 Missing Transverse Energy

Particles that go undetected through the CMS detector, such as neutrinos, mainly contribute to the missing transverse energy or MET or $\vec{E}_T^{miss}$. In addition, the miscalibrated objects can also lead to a sizable amount of MET. By definition,

$$\vec{E}_T^{miss} = -\sum_{i \in visible} \vec{P}_T \tag{2.12}$$

This $\vec{E}_T^{miss}$ reconstruction methodology utilizes the "pileup per particle identification" (PUPPI) method [43], and reduces the pileup reliance of the jet and $\vec{E}_T^{miss}$ observables by using local shape information surrounding each PF candidate in the event, event pileup attributes, and tracking information.

Chapter 3

Photon associated $t\bar{t}$ Production

3.1 Cross Section Measurement

In collider experiments, the measurements or searches are performed considering the specific final state signatures. The events from the lepton+jets final states, alternatively referred as semileptonic final states have been used in this thesis to measure the cross section in a specific kinematic phase space constrained with $p_T^{\gamma} \geq 20$ GeV and $|\eta^{\gamma}| \leq 1.44$. The measured value needs to be extrapolated to achieve the cross section value for all channels and full kinematic phase space of the photon (as produced in MC, $p_T \geq 13$ GeV and $|\eta|$ <3.0). So the first cross section is more trustworthy and it is called fiducial cross section, and the later one is inclusive cross section. For inclusive cross section measurement, two acceptances are needed to be considered. One is photon acceptance (A^{γ}) which is defined by the ratio of the events having at least one photon lying in that particular kinematic phase space and the total events generated. While the other acceptance is the top pair acceptance ($A^{t\bar{t}}$) which is calculated from the number of top pair decay events in semileptonic channel out of the total top pair decay events.

$$\sigma_{t\bar{t}+\gamma}^{inclusive} = \sigma_{t\bar{t}+\gamma}(p_T^{\gamma} \geq 13GeV, |\eta| < 3.0) \tag{3.1}$$

$$\sigma_{t\bar{t}+\gamma}(p_T^{\gamma} \geq 20GeV, |\eta| < 1.44) = \sigma_{t\bar{t}+\gamma}(p_T^{\gamma} \geq 13GeV, |\eta| < 3.0) \times A^{\gamma} \tag{3.2}$$

$$\sigma_{t\bar{t}+\gamma}^{fiducial} = \sigma_{t\bar{t}+\gamma}(p_T^{\gamma} \geq 20GeV, |\eta| < 1.44) \times A^{t\bar{t}} \tag{3.3}$$

One can also measure the differential cross section by measuring the distribution of cross section in different p_T^{γ} bins or N_{jets} bins.

$$\frac{1}{N}\frac{d\sigma}{dp_T^{\gamma}} \quad or \quad \frac{1}{N}\frac{d\sigma}{dN_{jets}} \tag{3.4}$$

3.1.1 Monte Carlo Simulation and Dataset

Along with the pp collision data from the detector, we need simulated events to understand it properly. Monte Carlo simulated events are used for measuring signal efficiency, estimating background events, and studying the systematic uncertainties in the results. Any physics process is modeled broadly in three steps, 1. Event generation 2. Parton shower handling, and 3. Detector simulation. In case of pp collision at LHC following steps are followed for event simulation

- Firstly, the cross section of a hard interaction between two incoming partons (coming from two colliding protons) is measured. The calculation is done in a fixed order of the strong coupling constant α_s.
- Next, the parton shower is added to the simulation process. Parton shower is created by the partons present before and after the hard interaction as they can produce quantum chromodynamics (QCD) and quantum electrodynamics (QED) radiation. For example, quarks can radiate a gluon or photon. Gluons can further split into more quarks and gluons, creating a shower of partons.
- The partons are permitted to undergo fragmentation and hadronization, ultimately forming a colorless composite state that rapidly disintegrates into baryons, mesons, leptons, and photons.

- The remnant partons that do not participate in hard interaction can form a colorless bound state. This is called an underlying event and is added to the simulation.
- Colliding proton bunches contain $\sim 10^{11}$ of protons. A good number of soft pp collisions also happen that get incorporated in simulation and called pile up events.
- Finally, all the particles are passed through a CMS detector simulation in the GEANT4 package. In this step the simulated data are stored in the same format as the experimentally collected data.

3.1.1.1 Monte Carlo Simulation

$t\bar{t} + \gamma$ process is simulated at leading order (LO) using the MADGRAPH5_aMC@NLO generator. The top quark mass is set to 172.5 GeV and the NNPDF3.0 parton distribution functions (PDFs) are used in the generation. The parton showering and hadronization have been processed with PYTHIA8.2. The generated photons should have at least 10 GeV transverse momenta and falling in the detector region confined within maximum $|\eta|$ value of 5. Additionally the photon should be well isolated from from any lepton, jet, or other photons with a minimum isolation in $\eta - \phi$ plane defined by ΔR = 0.1 (where, $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\Phi)^2}$). The photons coming from jet hadronization are not included. Photons can be produced as initial state radiation (in the case of $q\bar{q} \rightarrow t\bar{t} + \gamma$) or radiation from top quark and its decay particles. Separate samples have been produced for different $t\bar{t} + \gamma$ event decay channels: semileptonic, dileptonic, and all-hadronic decay. The leading order (LO) cross section for semileptonic, dileptonic, and all-hadronic decay samples are 5.05 pb, 1.495 pb, and 4.149 pb respectively.

In order to include NLO in consideration, an inclusive k-factor defined as the ratio of $\sigma^{NLO}_{t\bar{t}+\gamma}$ and $\sigma^{LO}_{t\bar{t}+\gamma}$ in fiducial phase region is calculated using MADGRAPH5_aMC@NLOv2.6.5. The calculated inclusive value of k-factor is 1.4852, which is multiplied to all LO produc-

tion cross section value of $t\bar{t} + \gamma$.

Background processes relevant to the signal process, $t\bar{t} + \gamma$, are $t\bar{t}$, Single top (separated in channels: t-channel, s-channel, tW channel and tγ+jets), t+ γ, Z+jets , Z+γ, TTV (V=W/Z), W+jets, W+γ, QCD multijets, Diboson (WW/WZ/ZZ) etc. All these background processes have been generated using either POWHEG or MADGRAPH5_a- MC@NLO generator. For all processes, particle showering and hadronization have been handled by PYTHIAv8.212.

Table 3.1: Signal samples used for the $t\bar{t} + \gamma$ process for 2016, 2017, and 2018

Process	cross section (pb)	Dataset
$t\bar{t} + \gamma$	1.495×1.4852	TTGamma_Dilept_TuneCP5_PSweights_13TeV-madgraph-pythia8
	5.056×1.4852	TTGamma_SingleLept_TuneCP5_PSweights_13TeV-madgraph-pythia8
	4.149×1.4852	TTGamma_Hadronic_TuneCP5_PSweights_13TeV-madgraph-pythia8

3.1.1.2 Dataset

We are using CMS full Run 2 dataset. Since this search is on the lepton+jets events looking, Single Electron and Single Muon datasets have been used. We are using a very new lighter format of data which is referred to as NanoAOD format and MC. The 2016 data are processed with the Cert_271036 − 284044_13TeV_ReReco_07Aug2017_Collisions16_JSON golden JSON file that contains the list of luminosity sections with all CMS subdetectors flagged as good. Similarly 2017 and 2018 data are processed with the following golden JSON files. Cert_294927 − 306462_13TeV_EOY2017ReReco_Collisions17_JSON_v1 and Cert_314472 − 325175_13TeV_17SeptEarlyReReco2018ABC_PromptEraD_Collisions18 _JSON.

Table 3.2: Primary dataset from the data-taking period 2016.

Run range	dataset name
272007-275376	Run2016B_ver2-Nano25Oct2019_ver2-v1/NANOAOD
275657-276283	Run2016C-Nano25Oct2019-v1/NANOAOD
276315-276811	Run2016D-Nano25Oct2019-v1/NANOAOD
276831-277420	Run2016E-Nano25Oct2019-v1/NANOAOD
277772-278808	Run2016F-Nano25Oct2019-v1/NANOAOD
278820-280385	Run2016G-Nano25Oct2019-v1/NANOAOD
280919-284068	Run2016H-Nano25Oct2019-v1/NANOAOD

Table 3.3: Primary dataset from the data-taking period 2017.

Run range	dataset name
297047-299329	Run2017B-Nano25Oct2019-v1/NANOAOD
299368-302029	Run2017C-Nano25Oct2019-v1/NANOAOD
302031-302663	Run2017D-Nano25Oct2019-v1/NANOAOD
303824-304797	Run2017E-Nano25Oct2019-v1/NANOAOD
305040-306462	Run2017F-Nano25Oct2019-v1/NANOAOD

Table 3.4: Primary dataset from the data-taking period 2018.

Run range	dataset name
315257-316995	Run2018A-Nano25Oct2019-v1/NANOAOD
317080-319310	Run2018B-Nano25Oct2019-v1/NANOAOD
319337-320065	Run2018C-Nano25Oct2019-v1/NANOAOD
320500-325175	Run2018D-Nano25Oct2019_ver2-v1/NANOAOD

Table 3.5: MC and data production campaign

year	campaign
2016	RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7
2017	RunIIFall17NanoAODv6-PU2017_12Apr2018_Nano25Oct2019_102X_mc2017_realistic_v7
2018	RunIIAutumn18NanoAODv6-Nano25Oct2019_102X_upgrade2018_realistic_v20

3.1.2 Event and Object Selection

Different versions of CMS software were used for event reconstruction for different data-taking year i.e 80x for 2016, 94x for 2017 and 102x for 2018. However, the 102X series of CMSSW has been used for producing the NanoAOD data tier for all three years. The signal

Table 3.6: Simulated MC samples used for 2016 data.

Process	cross section (pb)	Dataset
$t\bar{t}$	87.315	/TTTo2L2Nu_TuneCP5_PSweights_13TeV-powheg-pythia8
	380.095	/TTToHadronic_TuneCP5_PSweights_13TeV-powheg-pythia8
	364.352	/TTToSemiLeptonic_TuneCP5_PSweights_13TeV-powheg-pythia8
W+jets	11775.9345	/W1JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	3839.4345	/W2JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	1165.8108	/W3JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	592.9176	/W4JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
Z+jets	18610	/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	6077.22	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
Single top	3.68064	/ST_s-channel_4f_leptonDecays_TuneCP5_PSweights_13TeV-amcatnlo-pythia8
	136.02	/ST_t-channel_top_4f_InclusiveDecays_TuneCP5_PSweights_13TeV-powheg- pythia8
	80.95	/ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_PSweights_13TeV-powheg- pythia8
	35.85	/ST_tW_top_5f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8
	35.85	/ST_tW_antitop_5f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8
	2.967	/TGJets_TuneCUETP8M1_13TeV_amcatnlo_madspin_pythia8
$W+\gamma$	489	/WGToLNuG_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
$Z+\gamma$	98.3	/ZGToLLG_01J_5f_lowMLL_TuneCP5_13TeV-amcatnloFXFX-pythia8
$t\bar{t}+W$	0.4062	/TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8
	0.2043	/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8 [ext1+ext2]
$t\bar{t}+Z$	0.5297	/TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8
		/TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8
		/TTZToLL_M-1to10_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 [ext1+ext2+ext3]
tZ	0.09418	/tZq_ll_4f_13TeV-amcatnlo-pythia8
WW	49.997	/WWToLNuQQ_13TeV-powheg [+ext1]
	51.723	/WWTo4Q_13TeV-powheg
WZ	3.033	/WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8
	10.71	/WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8
	5.595	/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8
	4.42965	/WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8 [+ext1]
ZZ	4.04	/ZZTo2L2Q_13TeV_powheg_pythia8
		/ZZTo2Q2Nu_13TeV_amcatnloFXFX_madspin_pythia8
		/ZZTo4L_13TeV_powheg_pythia8
VV($2l2$)	11.95	/VVTo2L2Nu_13TeV_amcatnloFXFX_madspin_pythia8 [+ext1]

process final state consists of four jets (two are b-tagged), one tight high p_T lepton, no loose lepton, and one high transverse energy (E_T), isolated photon. Event selection is made in the following steps : 1. Firstly, baseline selection criteria are applied; To satisfy the condition, an event must undergo a trigger, which can either be a single electron or a single muon. Additionally, the event must have a single lepton with high transverse momentum (p_T) that is well isolated and tightly bound with no loosely bound lepton present. Furthermore, the event must contain two or more jets. 2. Secondly, tighter criteria are applied for signal region selection. Where in addition to the baseline requirements, presence of a well defined primary vertex along with at least 4 jets (among which at least one is b-tagged) and one

Table 3.7: Simulated MC samples used for 2017 data.

Process	cross section (pb)	Dataset
$t\bar{t}$	87.315	/TTTo2L2Nu_TuneCP5_PSweights_13TeV-powheg-pythia8
	380.095	/TTToHadronic_TuneCP5_PSweights_13TeV-powheg-pythia8
	364.352	/TTToSemiLeptonic_TuneCP5_PSweights_13TeV-powheg-pythia8
W+jets	11775.9345	/W1JetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8
	3839.4345	/W2JetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8
	1165.8108	/W3JetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8
	592.9176	/W4JetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8
Z+jets	18610	/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 [+ext1]
	6077.22	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 [ext1+ext2]
Single top	3.68064	/ST_s-channel_4f_leptonDecays_TuneCP5_PSweights_13TeV-amcatnlo-pythia8
	136.02	/ST_t-channel_top_4f_InclusiveDecays_TuneCP5_PSweights_13TeV-powheg- pythia8
	80.95	/ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_PSweights_13TeV-powheg- pythia8
	35.85	/ST_tW_top_5f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8
	35.85	/ST_tW_antitop_5f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8
	2.967	/TGJets_TuneCUETP8M1_13TeV_amcatnlo_madspin_pythia8
$W+\gamma$	489	/WGToLNuG_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
$Z+\gamma$	98.3	/ZGToLLG_01J_5f_lowMLL_TuneCP5_13TeV-amcatnloFXFX-pythia8
$t\bar{t}+W$	0.4062	/TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8
	0.2043	/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8
$t\bar{t}+Z$	0.5297	/TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8 [+ext1]
		/TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8
		/TTZToLL_M-1to10_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
tZ	0.09418	/tZq_ll_4f_13TeV-amcatnlo-pythia8
WW	49.997	/WWToLNuQQ_NNPDF31_TuneCP5_PSweights_13TeV-powheg-pythia8
	51.723	/WWTo4Q_NNPDF31_TuneCP5_PSweights_13TeV-powheg-pythia8
WZ	3.033	/WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8
	10.71	/WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8
	5.595	/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8
	4.42965	/WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8
ZZ	4.04	/ZZTo2L2Q_13TeV_powheg_pythia8
		/ZZTo2Q2Nu_13TeV_amcatnloFXFX_madspin_pythia8
		/ZZTo4L_13TeV_powheg_pythia8 [+ext1+ext2]
VV(2l2)	11.95	/VVTo2L2Nu_13TeV_amcatnloFXFX_madspin_pythia8

highly energetic photon are required. Along with the above mentioned selection criteria, different selection criteria are applied for different control regions.

3.1.2.1 Triggering criteria

In CMS there are triggers with lower thresholds which are prescaled with some factor to cope up with their higher triggering rates. But in this analysis only unprescaled triggers have been utilized, as recommended in the TopTriggerTwiki [44]. Triggers are listed in Table 3.11. To select the e+jets events, event should have one well islolated electron passing the HLT criteria with different thresholds in different data-taking year- 27, 32 and 32 GeV

Table 3.8: Simulated MC samples used for 2018 data.

Process	cross section (pb)	Dataset
$t\bar{t}$	87.315	/TTTo2L2Nu_TuneCP5_PSweights_13TeV-powheg-pythia8
	380.095	/TTToHadronic_TuneCP5_PSweights_13TeV-powheg-pythia8
	364.352	/TTToSemiLeptonic_TuneCP5_PSweights_13TeV-powheg-pythia8
W+jets	11775.9345	/W1JetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8
	3839.4345	/W2JetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8
	1165.8108	/W3JetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8
	592.9176	/W4JetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8
Z+jets	18610	/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 [+ext1]
	6077.22	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 [ext1+ext2]
Single top	3.68064	/ST_s-channel_4f_leptonDecays_TuneCP5_PSweights_13TeV-amcatnlo-pythia8
	136.02	/ST_t-channel_top_4f_InclusiveDecays_TuneCP5_PSweights_13TeV-powheg- pythia8
	80.95	/ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_PSweights_13TeV-powheg- pythia8
	35.85	/ST_tW_top_5f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8
	35.85	/ST_tW_antitop_5f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8
	2.967	/TGJets_TuneCUETP8M1_13TeV_amcatnlo_madspin_pythia8
$W+\gamma$	489	/WGToLNuG_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
$Z+\gamma$	98.3	/ZGToLLG_01J_5f_lowMLL_TuneCP5_13TeV-amcatnloFXFX-pythia8
$t\bar{t}+W$	0.4062	/TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8
	0.2043	/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8
$t\bar{t}+Z$	0.5297	/TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8 [+ext1]
		/TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8
		/TTZToLL_M-1to10_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
tZ	0.09418	/tZq_ll_4f_13TeV-amcatnlo-pythia8
WW	49.997	/WWToLNuQQ_NNPDF31_TuneCP5_PSweights_13TeV-powheg-pythia8
	51.723	/WWTo4Q_NNPDF31_TuneCP5_PSweights_13TeV-powheg-pythia8
WZ	3.033	/WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8
	10.71	/WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8
	5.595	/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8
	4.42965	/WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8 [+ext1]
ZZ	4.04	/ZZTo2L2Q_13TeV_powheg_pythia8
		/ZZTo2Q2Nu_13TeV_amcatnloFXFX_madspin_pythia8
		/ZZTo4L_13TeV_powheg_pythia8
VV($2l2$)	11.95	/VVTo2L2Nu_13TeV_amcatnloFXFX_madspin_pythia8

for 2016, 207 and 2018 respectively. Similarly for μ+jets event selection those transverse momentum thresholds are 24, 27 and 24 respectively. For both the electron and muon case, the lepton should be confined with $|\eta| \leq 2.4$.

Table 3.9: MC simulated samples for QCD multijet events, used for validation for the year 2016.

Process	cross section (pb)	Dataset
QCD (μ enriched)	2960198.40	/QCD_Pt-20to30_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8
	1652471.46	/QCD_Pt-30to50_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8
	437504.10	/QCD_Pt-50to80_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8
	106033.66	/QCD_Pt-80to120_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8
	25190.51	/QCD_Pt-120to170_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8
	8654.49	/QCD_Pt-170to300_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8
	797.35	/QCD_Pt-300to470_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8
	79.03	/QCD_Pt-470to600_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8
	25.10	/QCD_Pt-600to800_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8
	4.71	/QCD_Pt-800to1000_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8
	1.62	/QCD_Pt-1000toInf_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8
QCD ($e\gamma$ riched)	5352960	/QCD_Pt-20to30_EMEnriched_TuneCUETP8M1_13TeV_pythia8
	9928000	/QCD_Pt-30to50_EMEnriched_TuneCUETP8M1_13TeV_pythia8
	2890800	/QCD_Pt-50to80_EMEnriched_TuneCUETP8M1_13TeV_pythia8
	2890800	/QCD_Pt-50to80_EMEnriched_TuneCUETP8M1_13TeV_pythia8
	350000	/QCD_Pt-80to120_EMEnriched_TuneCUETP8M1_13TeV_pythia8
	62964	/QCD_Pt-120to170_EMEnriched_TuneCUETP8M1_13TeV_pythia8
	18810	/QCD_Pt-170to300_EMEnriched_TuneCUETP8M1_13TeV_pythia8
	1350	/QCD_Pt-300toInf_EMEnriched_TuneCUETP8M1_13TeV_pythia8
QCD (bcToE)	328999.93	/QCD_Pt_20to30_bcToE_TuneCUETP8M1_13TeV_pythia8
	405623.40	/QCD_Pt_30to80_bcToE_TuneCUETP8M1_13TeV_pythia8
	38104.43	/QCD_Pt_80to170_bcToE_TuneCUETP8M1_13TeV_pythia8
	2635.81	/QCD_Pt_170to250_bcToE_TuneCUETP8M1_13TeV_pythia8
	711.92	/QCD_Pt_250toInf_bcToE_TuneCUETP8M1_13TeV_pythia8
γ+jets	20730	/GJets_HT-40To100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	9226	/GJets_HT-100To200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	2300	/GJets_HT-200To400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	274.40	/GJets_HT-400To600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	93.38	/GJets_HT-600ToInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8

3.1.2.2 Object selection

3.1.2.2.1 Primary vertex selection To reconstruct primary vertices, track candidates that are accurately reconstructed and meet the following criteria are utilized.

- Minimum number of degrees of freedom > 4.
- Longitudinal vertex position, $|z| < 24$ cm.
- Track transverse position, $d_0 < 2$ cm.

The primary pp interaction vertex is selected with criteria of giving the largest value of

Table 3.10: MC simulated samples for QCD multijet events, used for validation for the year 2017 and 2018.

Process	cross section (pb)	Dataset
QCD (μ enriched)	2960198.40	/QCD_Pt-20to30_MuEnrichedPt5_TuneCP5_13TeV_pythia8
	1652471.46	/QCD_Pt-30to50_MuEnrichedPt5_TuneCP5_13TeV_pythia8
	437504.10	/QCD_Pt-50to80_MuEnrichedPt5_TuneCP5_13TeV_pythia8
	106033.66	/QCD_Pt-80to120_MuEnrichedPt5_TuneCP5_13TeV_pythia8
	25190.51	/QCD_Pt-120to170_MuEnrichedPt5_TuneCP5_13TeV_pythia8
	8654.49	/QCD_Pt-170to300_MuEnrichedPt5_TuneCP5_13TeV_pythia8
	797.35	/QCD_Pt-300to470_MuEnrichedPt5_TuneCP5_13TeV_pythia8
	79.03	/QCD_Pt-470to600_MuEnrichedPt5_TuneCP5_13TeV_pythia8
	25.10	/QCD_Pt-600to800_MuEnrichedPt5_TuneCP5_13TeV_pythia8
	4.71	/QCD_Pt-800to1000_MuEnrichedPt5_TuneCP5_13TeV_pythia8
	1.62	/QCD_Pt-1000toInf_MuEnrichedPt5_TuneCP5_13TeV_pythia8
QCD ($e\gamma$ riched)	5352960	/QCD_Pt-20to30_EMEnriched_TuneCP5_13TeV_pythia8
	9928000	/QCD_Pt-30to50_EMEnriched_TuneCP5_13TeV_pythia8
	2890800	/QCD_Pt-50to80_EMEnriched_TuneCP5_13TeV_pythia8
	2890800	/QCD_Pt-50to80_EMEnriched_TuneCP5_13TeV_pythia8
	350000	/QCD_Pt-80to120_EMEnriched_TuneCP5_13TeV_pythia8
	62964	/QCD_Pt-120to170_EMEnriched_TuneCP5_13TeV_ pythia8
	18810	/QCD_Pt-170to300_EMEnriched_TuneCP5_13TeV_pythia8
	1350	/QCD_Pt-300toInf_EMEnriched_TuneCP5_13TeV_pythia8
QCD (bcToE)	328999.93	/QCD_Pt_20to30_bcToE_TuneCP5_13TeV_pythia8
	405623.40	/QCD_Pt_30to80_bcToE_TuneCP5_13TeV_pythia8
	38104.43	/QCD_Pt_80to170_bcToE_TuneCP5_13TeV_pythia8
	2635.81	/QCD_Pt_170to250_bcToE_TuneCP5_13TeV_pythia8
	711.92	/QCD_Pt_250toInf_bcToE_TuneCP5_13TeV_pythia8
γ+jets	20730	/GJets_HT-40To100_TuneCP5_13TeV-madgraphMLM-pythia8
	9226	/GJets_HT-100To200_TuneCP51_13TeV-madgraphMLM-pythia8
	2300	/GJets_HT-200To400_TuneCP5_13TeV-madgraphMLM-pythia8
	274.40	/GJets_HT-400To600_TuneCP5_13TeV-madgraphMLM-pythia8
	93.38	/GJets_HT-600ToInf_TuneCP5_13TeV-madgraphMLM-pythia8

Table 3.11: Trigger requirements from the TopTriggerTwiki

year	Primary Dataset	trigger	comment
2016	SingleElectron	HLT_Ele27_WPTight_Gsf_v*	
	SingleMuon	HLT_Iso(Tk)Mu24_v*	
2017	SingleElectron	HLT_Ele32_WPTight_Gsf_L1DoubleEG_v*	L1 SingleEG seed required
	SingleMuon	HLT_IsoMu27_v*	
2018	EGamma	HLT_Ele32_WPTight_Gsf_v*	
	SingleMuon	HLT_IsoMu24_v*	

summed P_T^2 of all physics objects. We require events to have at least one good first primary vertex.

3.1.2.2.2 Electron For selecting an electron, the baseline criteria are the "cut based" tight ID [45] selection with a minimum transverse momentum of 35 GeV and within $|\eta| < 2.4$. The electrons falling in the gap between the electromagnetic barrel (EB) and electromagnetic endcap, defined by $1.4442 < |\eta_{SC}| < 1.566$ are not considered here. Additionally, the electron should have 0.05 (0.1) minimum impact parameter, d_{xy} value in EB (EE) in x-y plane, while d_z i.e., the impact parameter in z-direction should not have values exceeding 0.1 (0.2) in EB (EE). For vetoing extra leptons, the POG[1] veto IDs are used with $p_T \geq 15$ GeV and $|\eta| \geq 2.4$. VetoID for loose leptons as recommended by POG are used to veto any extra lepton from the events. These requirements suppress dileptonic $t\bar{t} + \gamma$, Z+γ and Z+jets events. Table 3.12 shows a elaborated criteria of electron ID [46] in CMS.

Table 3.12: Electron ID requirements

Requirements	Tight	Veto
Full 5x5 $\sigma_{i\eta i\eta} \leq$	0.0104	0.0128
abs($\Delta\eta seed$)	0.00353	0.00523
abs(dphiIn) $\leq$	0.0499	0.159
C_0	0.026	0.356
relIsoWithEA $\leq$	0.0361	0.175
abs(1/E -1/p) $\leq$	0.0278	0.299
expected missing inner hits $\leq$	1	2
pass conversion veto	yes	yes

3.1.2.2.3 Muon For muon selection, the required baseline criterion is the tight POG ID [47], along with a transverse momentum above 30 GeV and within $|\eta| < 2.4$. Since the cut-based ID criteria for muons do not include the relative isolation condition, it is applied as an additional criterion for muon selection. Here, the muon should have minimum relative

[1]POG=Physics Objcet Group, monitors the standard reconstruction criteria of a physics object at CMS

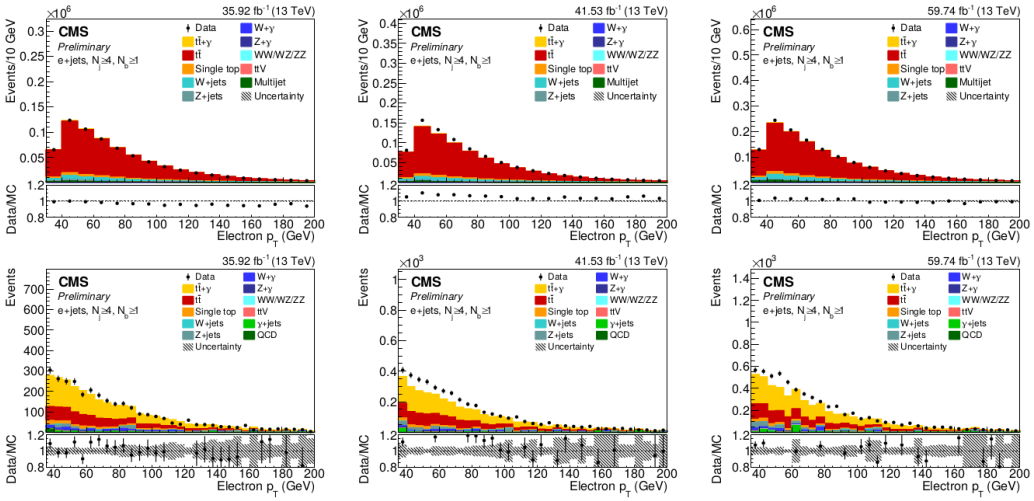


Figure 3.1: p_T distribution of electron different sample in $N_{jet} \geq 4$ & $N_{b-jet} \geq 1$ (top) selection only and with a photon selection (bottom) in e+jets events for 2016, 2017 and 2018 data taking year.

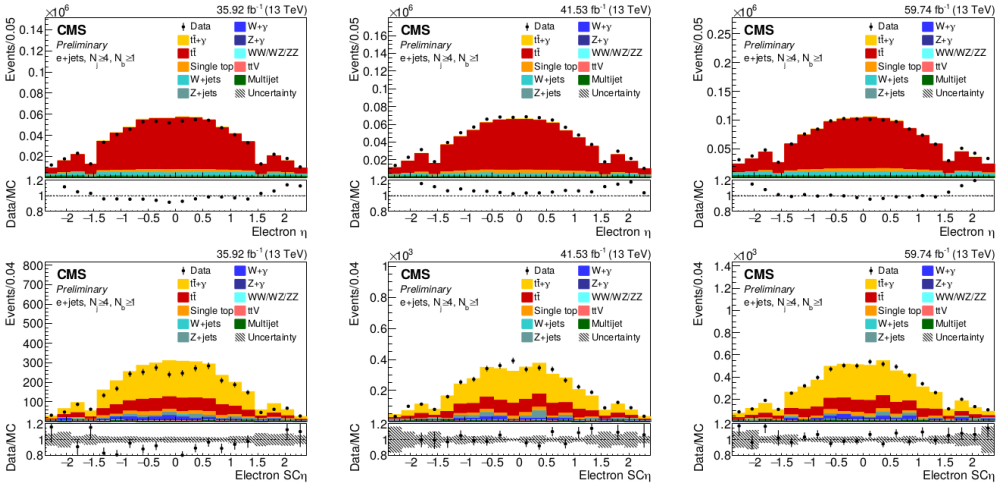


Figure 3.2: Super cluster η distribution of electron in different samples in $N_{jet} \geq 4$ & $N_{b-jet} \geq 1$ (top) selection only and with a photon selection (bottom) in e+jets events for 2016, 2017 and 2018 data taking year.

isolation of 0.15. The ratio of muon transverse momentum to that of all particles within a cone with radius $R \leq 0.4$ is used to calculate the relative isolation variable (I_{rel}) of the muon.

$$I_{rel} = \frac{1}{p_T^{\mu}} \times \left(\sum_{ch.had.} p_T^{PU} + \max \left[0, \sum_{pho} p_T + \sum_{neu.had.} p_T - 0.5 \times \sum_{ch.had.} p_T^{PU} \right] \right) \quad (3.5)$$

where the summations run over the charged hadrons, neutral hadrons and photons. All the selected muons are required to satisfy the relative isolation less than 0.15. To reject the events with an additional loose muon along with this above discussed tight muon, the POG veto ID criteria are utilized with minimum transverse momentum of 15 GeV and absolute pseudo-rapidity of maximum 2.4. Furthermore the I_{rel} is increased to 0.25 for the loose muons. Vetoing the loose muon events suppresses the dimuonic events originating from $t\bar{t} + \gamma$, Z+γ and Z+jets processes. Table 3.14 and 3.13 show the full list of criteria of muon ID [47] in CMS.

Table 3.13: Muon ID requirements

Requirements	Tight	LooseID
Reconstructed Global Muon	yes	-
PF Muon ID	yes	yes
Global or Tracker Muon	-	yes
value of normalized χ^2 in the fit for the tracks of global muons	$<$10	-
number of muon chamber hits taken in the track fit of global muons	> 0	-
Muon segments in at least two muon stations	> 1	-
d_{xy} (cm)	< 0.2	-
d_z (cm)	< 0.5	-
number of pixel hits	$>$0	-
tracker layer number containing hits	> 5	-

3.1.2.2.4 Photon The photons passing POG MediumID [48] selection are considered here along with the requirement a transverse energy greater than 20 GeV. The photon under consideration must be confined withing the barrel region only, i.e., $|\eta| \leq 1.4442$. The signal events consisting exactly one photon (passing the MediumID criteria) are selected while, the events with additional photons passing POG VetoID are rejected. The transverse energy and pseudo-rapidity thresholds for VetoID are similar to those of MediumID. However, the

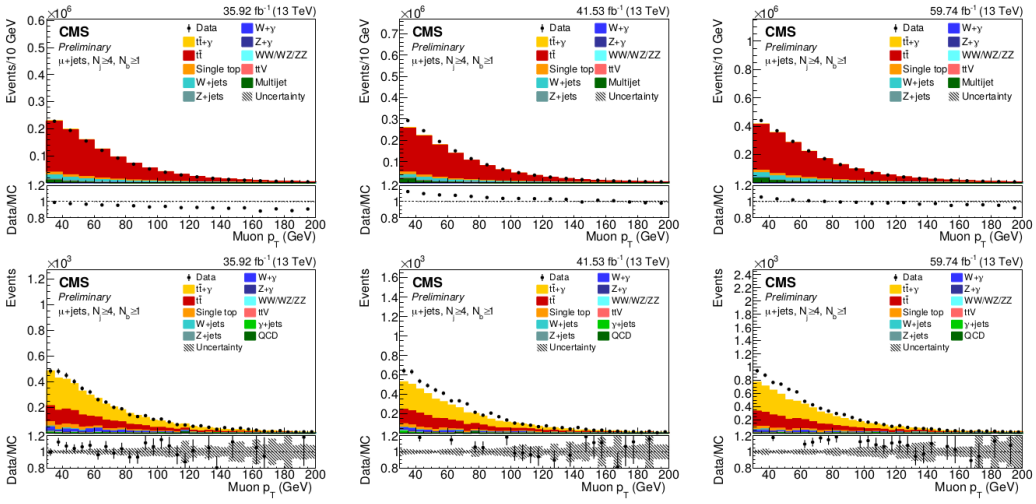


Figure 3.3: p_T distribution of muon in different samples in $N_{jet} \geq 4$ & $N_{b-jet} \geq 1$ (top) selection only and with a photon selection (bottom) in μ+jets events for 2016, 2017 and 2018 data taking year.

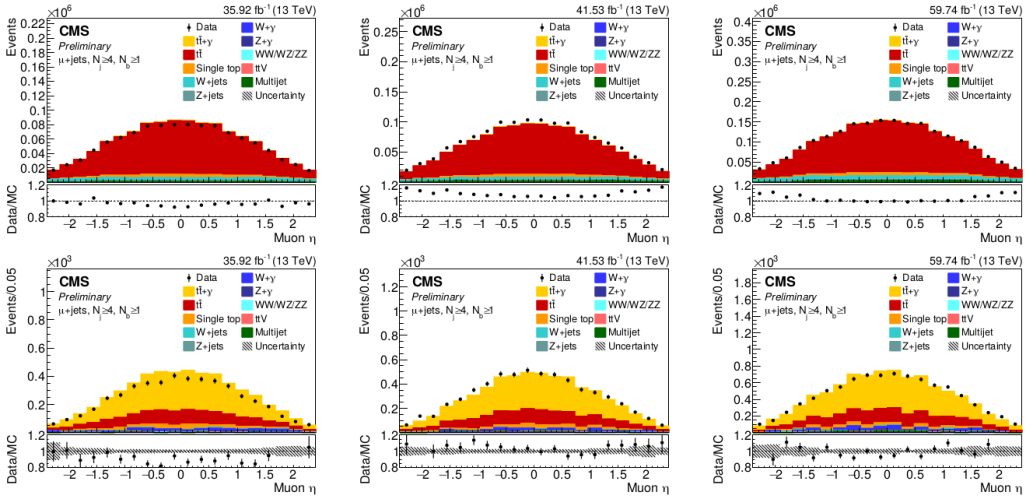


Figure 3.4: η distribution of muon in different samples in $N_{jet} \geq 4$ & $N_{b-jet} \geq 1$ (top) selection only and with a photon selection (bottom) in μ+jets events for 2016, 2017 and 2018 data taking year.

vetoID comprises of MediumID criteria without any restriction on $\sigma_{i\eta i\eta}$, charge isolation and the requirement of a track seed in the pixel detector. Table 3.15 and 3.16 shows the

Table 3.14: Electrons [45] and Muons [47] identification criteria based on the cut-based tight ID selection.

Observable	Lepton Selection			
	e		μ	
	Nominal ID	Veto ID	Nominal ID	Veto ID
p_T (GeV)	≥ 35	≥ 15	≥ 30	≥ 15
$\|\eta\|$	≤ 2.4	≤ 2.4	≤ 2.4	≤ 2.4
$I_r^{0.4}el$	-	-	≤ 0.15	≤ 0.25
$\|d_{xy}\|(cm)$	≤ 0.05 EB (0.1 EC)	$\leq 0.05EB(0.1EC)$		
$\|d_z\|(cm)$	≤ 0.1 EB (0.2 EC)	$\leq 0.1EB(0.2EC)$		
ID	cut-based tight ID	cut-based veto ID	cut-based tight ID	PF global μ or tracker μ

detailed criteria of photon selection [45] in CMS.

Table 3.15: The requirements for photon ID selection.

Variables	Values
H/E	< 0.027
$\sigma_{i\eta i\eta}$	<0.01015
ρ-corr charged hadron isolation	< 0.141
ρ-corr neutral hadron isolation	$1.189 + 0.01512 \times p_T(\gamma) + 2.259 \times 10^{-05} \times p_T(\gamma)^2$
ρ-corr photon hadron isolation	$2.08 + 0.004017 \times p_T(\gamma)$
has pixel seed veto	yes

Table 3.16: Final requirements for photon identification

Observables	Photon Selection	
	Nominal ID	Veto ID
p_T (GeV)	≥ 20	≥ 20
$\|\eta\|$	≤ 1.4442 (barrel)	≤ 1.4442 (barrel)
	cut-based medium ID	cut-based medium ID
	no pixel seed	no pixel seed
		w/o requirements on $\sigma_{i\eta i\eta}$ and chg.Iso

3.1.2.2.5 Jet It is quite important to reconstruct jets efficiently as the signal events consist at least four jets. This analysis uses AK4PFCHS[2] jets that pass the ID requirements listed in Table 3.18 with $p_T \geq 30$ GeV and $|\eta| < 2.4$. The CMS collaboration recommended

[2]AK4 PF CHS = Anti-K_t algorithm with cone radius 0.4, particle flow, charged hadrons subtraction

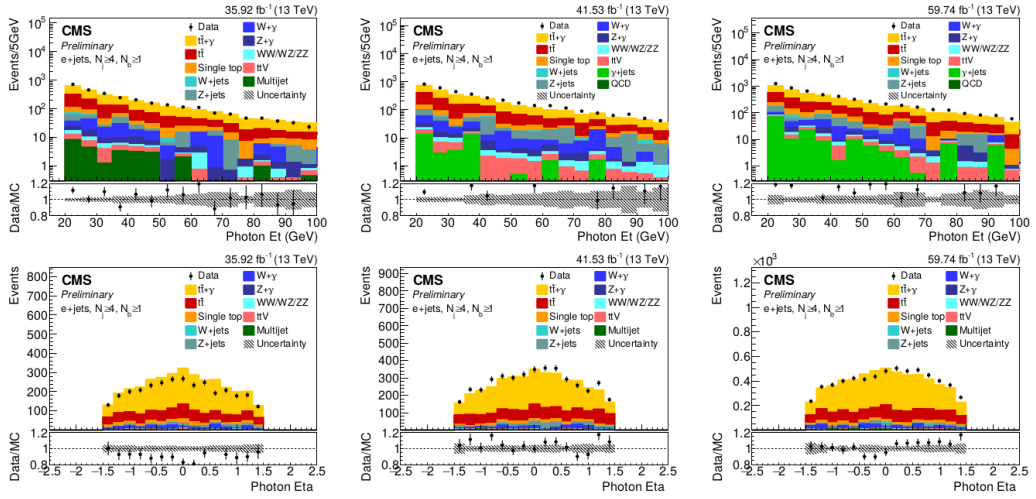


Figure 3.5: Distribution of E_T (top) and η (bottom) values of leading photon in $N_{jet} \geq 4$ & $N_{b-jet} \geq 1$ selection in e+jets events for 2016, 2017, and 2018 data.

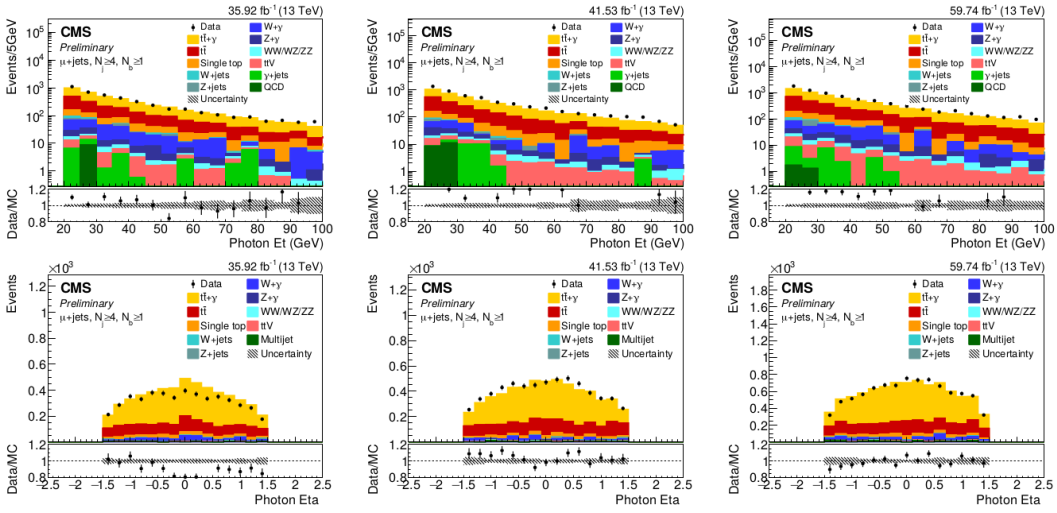


Figure 3.6: Distribution of E_T (top) and η (bottom) values of leading photon in $N_{jet} \geq 4$ & $N_{b-jet} \geq 1$ search region in μ+jets events for 2016, 2017, and 2018 data.

jet energy corrections (Summer16_07Aug2017_V11, Fall17_17Nov2017_V32, and Autumn18_V19) are utilized to correct for the jet energy scale, and these adjustments propagated to correct for the Missing Transverse Energy (MET).

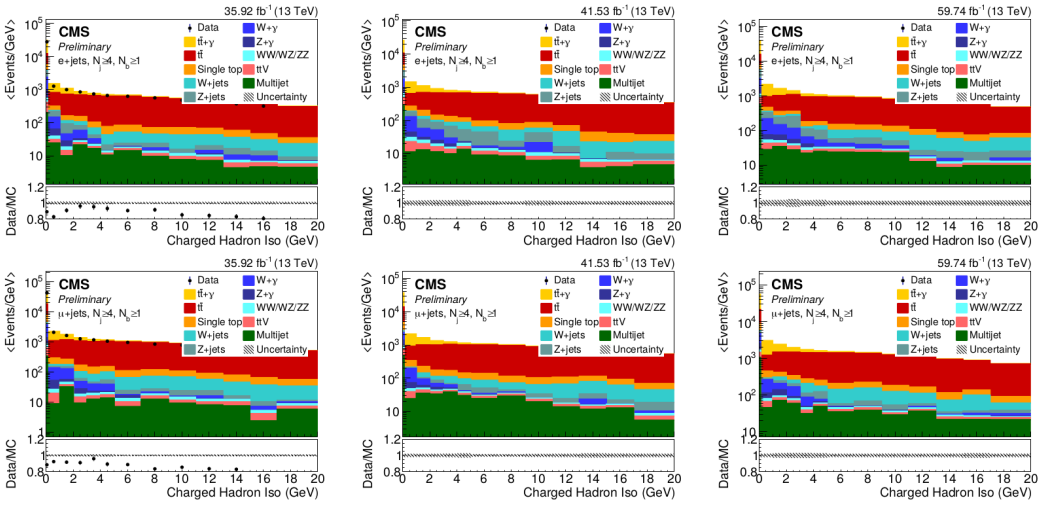


Figure 3.7: Charged hadron isolation (ChIso) distribution of the most energetic photon, in e+jets (top) and μ+jets (bottom) events in $N_{jet} \geq 4$ & $N_{b-jet} \geq 1$ search region for 2016, 2017, and 2018 data.

3.1.2.2.6 b-tagging of Jet DeepCSV discriminator has been utilized to tag a jet originating from the bottom quark hadronization. Different values of discriminator score have been used for different data-taking years. Here, the BTV[3] POG [49] recommended values are 0.6321, 0.4941, and 0.4184 for the years 2016, 2017, and 2018, respectively.

Table 3.17: Loose and tight ID criteria for the jets

Variables	Loose Jet	Tight Jet
Neutral hadron fraction	$< 0.99 < 0.99$	
Neutral EM fraction	< 0.99	< 0.90
Number of constituents	> 1	> 1
Charged hadron fraction	$> 0 > 0$	
Charged multiplicity	$> 0 > 0$	
Charged EM fraction	$< 0.99 < 0.99$	

3.1.2.2.7 Missing Energy Adjustments due to jet energy corrections are applied to the missing transverse energy (MET) through the particle flow algorithms. The signal event

[3]BTV = B-Tagging and Vertexing

Table 3.18: Details of jet selection thresholds.

Observables	Nominal ID
p_T (GeV)	≥ 30
$\|\eta\|$	≤ 2.4
ID	loose PF ID (2016), tight PF ID (2017/2018)
b-tag (DeepCSV)	≥ 0.6321 / 0.4941 / 0.4184 (2016/2017/2018)

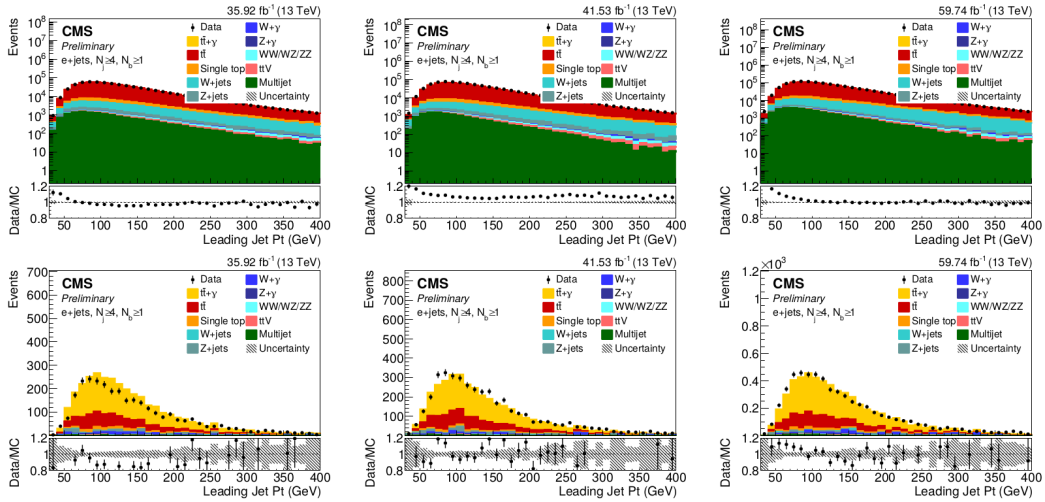


Figure 3.8: p_T distribution of most energetic jet in different samples in $N_{jet} \geq 4$ & $N_{b-jet} \geq 1$ (top) selection only and with a photon selection (bottom) in e+jets events for 2016, 2017 and 2018 data taking year.

selection does not have any explicit requirement for MET. However, it is essential for leptonically decaying W boson transverse mass reconstruction which is the key variable for QCD multijet background estimation. The uncertainties involving the jet energy corrections, can lead to a significant change in MET calculations.

3.1.3 Object Cleaning

Since the signal event signature consists of isolated objects, it is crucial to require the reconstructed objects to be well-isolated from one another. If a tight lepton is found within a

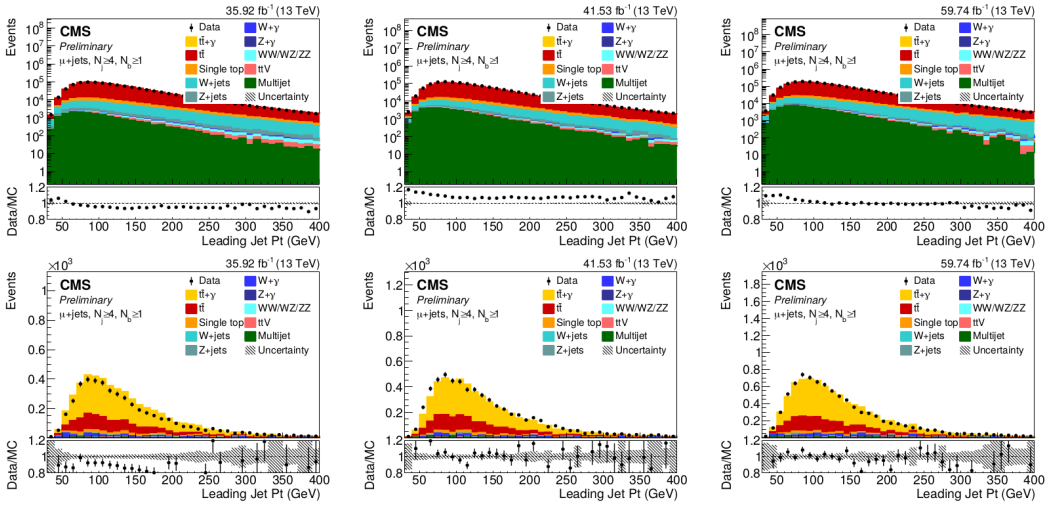


Figure 3.9: p_T distribution of most energetic jet in different samples in $N_{jet} \geq 4$ & $N_{b-jet} \geq 1$ (top) selection only and with a photon selection (bottom) in μ+jets events for 2016, 2017 and 2018 data taking year.

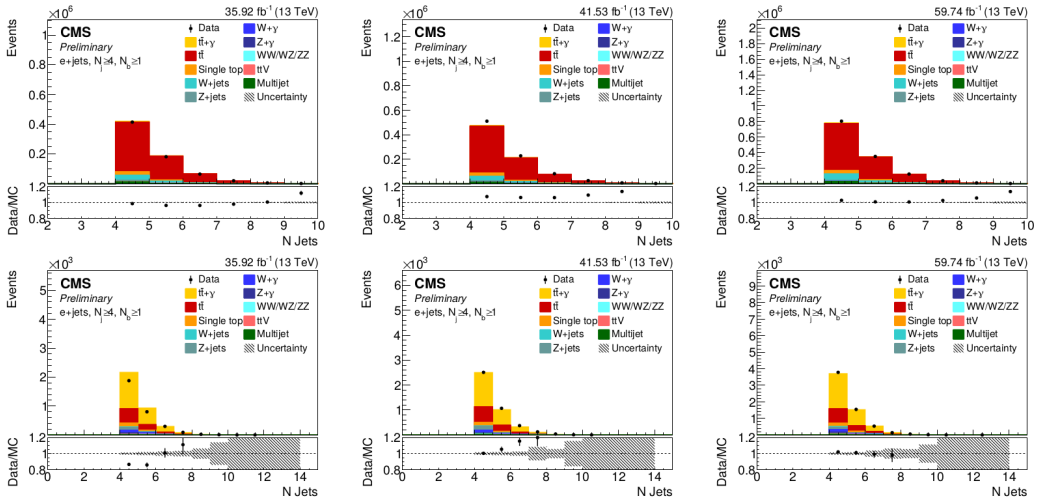


Figure 3.10: Distribution of the number of jets present in different samples in $N_{jet} \geq 4$ & $N_{b-jet} \geq 1$ (top) selection only and with a photon selection (bottom) in e+jets events for 2016, 2017 and 2018 data taking year.

distance of $\Delta R < 0.4$, photons are not considered, which reduces the probability of selecting a photon that has been emitted by a lepton. Similarly, to exclude jets that are clustered

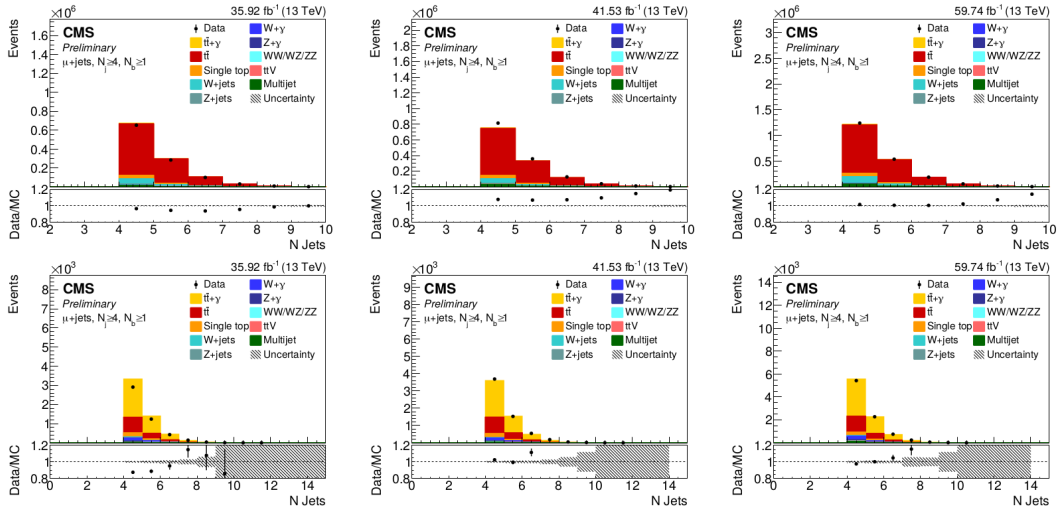


Figure 3.11: Distribution of the number of jets present in different samples in $N_{jet} \geq 4$ & $N_{b-jet} \geq 1$ (top) selection only and with a photon selection (bottom) in μ+jets events for 2016, 2017 and 2018 data taking year.

around tight leptons, the same ΔR criteria are used to check the distance between the jets and the tight leptons. To eliminate jets that are clustered around photons, a tighter criterion of $\Delta R < 0.1$ is applied to the jets and photons.

3.1.4 Photon Categorization

The origin of the reconstructed photon is needed to be checked in order to identify the signal photons in the simulated the events. Generator level particle in the close proximity of the reconstructed photon is considered to be the origin of the reconstructed photon. Based on the types of generator photons, the reconstructed photons are classified into different categories. Close proximity criteria is accomplished by a $\Delta R < 0.3$ requirement. Additionally, the matched particle must have the transverse momentum value within 50% of the reconstructed photon E_T. In cases when there are no generator level particles nearby the reconstructed photon some additional checks are applied to further classify them. Definition of different categories have been discussed below.

- Genuine photon:

  - When a generator photon originating from either lepton, boson or quarks, falls within the close proximity of the reconstructed photon.

- Misidentified electron:

  - When a generator level electron is matched to the reconstructed photon.

- Hadronic photon:

  - When the reconstructed photon is matched with a generator level photon coming from a hadronic source such as neutral pion decay.

- Hadronic fake (PU photon and fake photons):

  - PU photon : no generator particle without status requirement is found within ΔR $<$0.3.

  - Fake photon: any photon other than above mentioned category.

The "prompt photons are the genuine photons and the misidentified electrons, where as the term "non-prompt photons" comprises the last three contributions, i.e., hadronic photons, PU photons, and fake photons. The term is introduced, because these contributions are estimated together.

3.1.5 Object and Event Level Corrections

In order to address biases in detector response and for the discrepancies in physical observables between simulation and data, various corrections are applied to several observables in both data and simulation. These corrections are necessary to ensure that the simulation accurately models the important aspects of the signal and background processes. Although the background is often determined using data-driven methods, the following corrections are still applied.

3.1.5.1 Pileup Reweighting

The pileup profile for the Monte Carlo simulated events, are reweighted to match the actual pileup profile in data. Total inelastic cross section of proton-proton collision is assumed to be 69.2 mb [50]. The estimation of distribution of the actual number of proton-proton interactions in the data is made based on the instantaneous luminosity per bunch crossing for each luminosity section. An uncertainty of $\pm$ 4.6% on this cross section is considered due to the uncertainties associated with the pileup models. The simulated samples are reweighted considering the number of interactions in individual data periods. Through the application of pileup reweighting techniques, additional corrections for muons, electrons, jets, and b-tagged jets have been determined.

3.1.5.2 Correction due to L1 prefiring

During the 2017 LHC operations, there was a drift in the time alignment of the endcap ECAL readout electronics, which severely affected the data recorded in "F" period. Due to this misalignment in timing the trigger primitives formed by the energy deposits in ECAL may be reconstructed as originating from previous bunch crossing. This leads to an inadequacy in the decision made by L1 trigger, since consecutive bunch crossings [51] are not accepted to be triggered in CMS. Detailed study reveals that this prefiring issue [52] has an impact only on the events with high transverse momentum jets within the region of $2.4 < |\eta| < 3.0$. Although in this analysis no physics objects are reconstructed in such forward region directly, some additional objects might have impact on those events. Using a unique set of triggered events known as "un-prefirable" events, prefiring inefficiency is quantified by the JME POG[4]. All of the jets in the event are used to account for the event efficiency factor ($\epsilon_j = 1-\mathcal{P}$), which is applied to the MC simulation as follows

[4]Jet and missing energy physics object group

$$\epsilon = \prod_{j \in jets} 1 - \mathcal{P}_{prefire}(j) \tag{3.6}$$

A detail explanation of the L1 trigger prefiring can be found at [53].

3.1.5.3 HEM issue and correction

On September 13, 2018, during a routine maintenance operation, an electrical short circuit occurred in the power supply system of the one end of the hadronic endcap calorimeter (HEM) [54]. Events for run numbers following 319077 that contain an electron or photon with $p_T \geq 15$ GeV, $-3 < \eta < -1.4$, and $-1.57 < \phi < -0.87$ are excluded in order to combat the HEM effect. This eliminates deceptive e/γ objects generated based on the ECAL deposit in that affected direction. Events containing an object that meets these requirements are weighted in the 2018 simulation by the luminosity fraction (0.3518) of the relevant data collection period.

3.1.5.4 Scale factors on objects

The simulated objects are calibrated/recontructed considering certain in-situ operational conditions, which may not exactly represent the reality. For this particular reason, there are always residual differences in terms of reconstructed efficiencies between collision data and simulation and appropriate correction factors are applied accordingly.

- b-tagging: As the MC does not reproduce completely the b-tagging performance in data, the officially recommended (by BTV POG) data-to-MC scale factors [49] are applied.

- Photon efficiency: The MC samples are reweighted to account for the efficiency in photon identification, as well as for the efficiency of the Pixel Seed Veto (PSV), according to recommendations of the EGamma POG [55]. The ”102X” 2D ID efficiency map for 2018 data, the ”94X” 2D map for 2017, and the ”94X” map for 2016

are used to calculate the efficiency correction scale factor. Similarly "ScalingFactors_80X_Summer16" and "PixelSeed_ScaleFactors_2017" were used for The PSV scale factor for 2016 and 2017, where as for 2018, the Pixel Seed Veto scale factors were remeasured, and the results are in line with the scale factors that were centrally provided. The identification uncertainty has an significant impact on measured cross section.

- Electron efficiency: The tight ID efficiency map [56] produced with same version of CMS software structure as used in photon efficiency map estimation, have been used for electron efficiency correction. For electron, the efficiency difference arises due to trigger, as well as identification and isolation.

- Photon energy scale uncertainties: EGM provides residual corrections to scale the data to the MC and smear the MC to the resolution in data. The systematic uncertainties for the electron and photon objects are applied as prescribed in Ref [57]. In contrast to the identification efficiency systematics, the $p_T(\gamma)$ scale uncertainties have a negligible impact on the measured cross section.

- Muon efficiency: Similar to the electron, muon reconstruction also has efficiency difference between collisions data and simulation due to discrepancy in triggers, identification, and isolation processes. The correction factors have been used as centrally recommended by MUON POG [58, 59]. ID, isolation, and trigger efficiency are factorized into the 2D scale factors. Muon ID uncertainties lead to an additional uncertainty of 0.5% on cross section value [60].

Systematic uncertainties originating for all scale factor determination techniques are taken into consideration and are appropritely transmitted to the final cross section measurement.

3.1.6 Overlap removal among simulated samples

Separate samples are used for the $t\bar{t} + \gamma$ process and the inclusive $t\bar{t}$ production. The inclusive $t\bar{t}$ sample contains, in principle, events where a shower photon is radiated at high energy and at a large angle. This phase space is already covered by the $t\bar{t} + \gamma$ sample, and thus, the overlap must be removed by vetoing events in the $t\bar{t}$ sample that fall into the $t\bar{t} + \gamma$ phase space.

The algorithm followed in this analysis was used for the 8 TeV analysis in Ref[61] but adjusted for the change in generator thresholds. The overlap removal discards events from the $t\bar{t}$ sample, which contain a generated photon satisfying $p_T > 10$ GeV, $|\eta| < 5.0$, and which does not contain a meson in its generator ancestry. The latter is technically accomplished by limiting the paternity chain of the produced photon to quarks, gluons, leptons, or fundamental SM bosons. The kinematic thresholds and the requirement on the parentage ensure coverage of this phase space by the $t\bar{t} + \gamma$ sample. Furthermore, the generated photon must be isolated, i.e., separated by $\Delta R > 0.1$ from other generated particles with $p_T > 5$ GeV, excluding neutrinos and photons. In very rare cases, the photon yields a lepton pair at the generator level such that such products are also removed from the isolation requirement. The specific ΔR value matches the generator settings in the simulation. The portion of phase space with photons originating during the hadronization of jets, or having parentage including pions or other mesons, is thus covered by the $t\bar{t}$ sample. In turn, in the $t\bar{t} + \gamma$ sample, the presence of at least one such photon is required. For the purpose of eliminating events that conform to the W+γ or Z+γ sample criteria, a similar overlap reduction procedure is performed to the W +jets and Z+jets samples. In these situations, the definition of the generator level cuts for the photon in the overlap removal corresponds to the simulation settings of the W+γ or Z+γ samples. Events are excluded from W+jets and Z+jets sample if any there are any photons with minimum p_T threshold of 15 GeV and $|\eta|$ confinement

within 2.6, satisfying the leptonic, quark or boson parentage and ΔR within 0.1. The single top quark t-channel sample undergoes an overlap removal process, where events already present in the tγ sample are eliminated. According to the generator settings, the generator level cuts such as $p_T > 10$ GeV, $|\eta| < 2.6$ are aligned with the phase space of the single top quark plus photon sample. Moreover, the isolation requirement for the sample is specified as $\Delta R < 0.05$. Single top events are not eliminated by taking into account photons from the decay products of the top quark during the removal process since the tγ sample does not consist of such photons. All cuts are summarized in Table 3.19

Table 3.19: List of cuts used for removing the overlapping events between simulated samples.

cut/sample	$t\bar{t}+\gamma/t\bar{t}$	$W+\gamma/W+jets$	$t\gamma/single-t(t-channel)$	$\gamma+jets/QCD$	
$p_T(\gamma^{gen})(GeV)$	≥ 10	≥ 15	≥ 15	≥ 10	≥ 25
$\lvert\eta\rvert(\gamma^{gen})$	≤ 5.0	≤ 2.6	≤ 2.6	≤ 2.6	≤ 2.6
ΔR cone size	0.1	0.05	0.05	0.05	0.4

3.1.7 Different selection and control regions for signal and background

After the PF algorithm selects the physics objects such as e, μ, γ and jets, we apply subsequent criteria to ensure the correct reconstruction of these objects. Along with the lepton trigger criteria suitable to our phase space, events must pass the MET filter as recommended by the JETMET POG [62]. High cross section of inclusive top pair production resulting high statistics enable us to apply very tighter criteria to select the events with high purity. We look for exactly one tight lepton and at least 3 jets, and at least one b-tagged jet for baseline event selection as listed in Table 3.20. But top quark pair production in association with photon is significantly lower than inclusive top pair production. Thus the identification criteria on photon is medium ID as discussed in Sec 3.1.2.2.4. Depending on number of jets requirement signal region is splitted in two bin - SR3 ($N_{jet} = 3$) and SR4P ($N_{jet} \geq 4$). Fig-

ure 3.12 illustrates that the additional control regions are determined based on the baseline selection and may have less strict criteria for the number of jets and b-tag multiplicities. The subsequent sections outline the specific roles and any further requirements for these control regions.

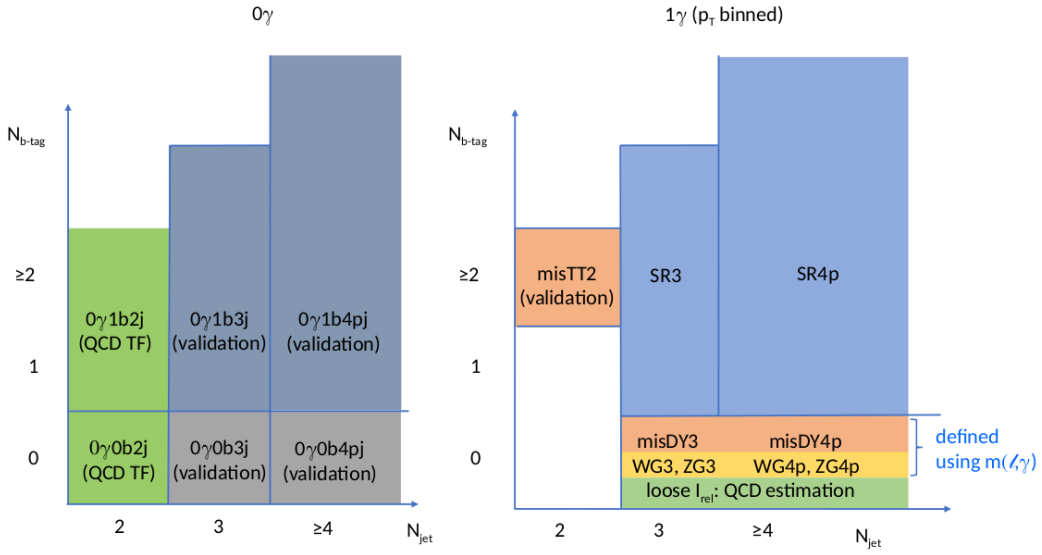


Figure 3.12: Selection criteria for different signal and control regions.

3.1.8 Particle level fiducial phase space definition

The stable particles are subjected to an event selection procedure akin to that used to determine the reconstruction level after the event generation, parton showering, and hadronization but before the detector simulation. This event selection is referred to as particle level selection. In the Table 3.22 the definitions of particle level objects are listed.

The fiducial phase space is then defined as $N_{\gamma}^{fid} = 1$, $N_{\ell}^{fid} = 1$, $N_{jet}^{fid} \geq 3$, and $N_{b-jet}^{fid} \geq 1$. The photon is required to be matched to a generator photon with status 1 within $\Delta R \leq 0.1$.

Table 3.20: Baseline event selection criteria.

Trigger Selection	Table 3.12
PV	leading PV satisfies requirements in sec 3.3.2
E_T miss filters	Flag HBHENoiseFilter
	Flag HBHENoiseIsoFilter
	Flag goodVertices
	Flag globalTightHalo2016Filter
	Flag EcalDeadCellTriggerPrimitiveFilter
	Flag BadPFMuonFilter
	ecalBadCalibReducedMINIAODFilter (2017 and 2018)
N_l	=1
$N_{l,veto}$	=1 (the selected lepton also passes the veto electron selection)
N_{jet}	≥ 3
N_{b-jet}	-
N_γ	=1
$N_{\gamma,veto}$	= 1 (the selected photon also passes the veto photon selection)

3.1.9 Background Estimation

3.1.9.1 Summary of the Constituent Elements of the Background

In Figure 3.13 the distribution of p_T for photons is depicted, with the left panel displaying the photons classified by simulated samples and the right panel showing the photons classified by photon categories as discussed in Section 3.1.4. The right panel illustrates a higher contribution of genuine photons in comparison to the contribution by the $t\bar{t} + \gamma$ signal sample in right panel. Hence, it is clear that additional processes like $t\bar{t}$, tWγ and tγ are also contributing to the genuine photons category. The normalization of background processes determined by performing maximum likelihood fittings, where as shape templates are taken from simulation as well as by data-driven method for some cases.

3.1.9.2 Fit procedure

In the final cross section measurement, signal and control regions are fitted together using the RooStats-based statistics tool provided by the Higgs PAG [63]. All the information needed to perform the fit is stored in datacards. For each signal point, a datacard is con-

Table 3.21: Epitome of the different signal and control regions.

Region	N_l	N_{jet}	N_{b-jet}	N_γ	$Z-Window/M_3$ slices	comment
DY3	2	3	0	0 $\lvert m(l,l)-m_Z\rvert < 10$		
DY3p	2	≥3	0	0 $\lvert m(l,l)-m_Z\rvert < 10$		
DY34p	2	≥4	0	0 $\lvert m(l,l)-m_Z\rvert < 10$		
0γ1b2j	1	2	1	0		QCD transfer factor
0γ1b3j	1	3	1	0		validation region
0γ1b3pj	1	≥ 3	1	0		validation region
0γ1b4pj	1	≥ 4	1	0		validation region
0γ0b2j	1	2	0	0		QCD transfer factor
0γ0b3j	1	3	0	0		validation region
0γ0b3pj	1	≥ 3	0	0		validation region
0γ0b4pj	1	≥ 4	0	0		validation region
ZG3	1	3	0	1	$m(e,\gamma) < m_Z - 10$ $m(\mu,\gamma) < m_Z$	
ZG3p	1	≥3	0	1	$m(e,\gamma) < m_Z - 10$ $m(\mu,\gamma) < m_Z$	
ZG4p	1	≥4	0	1	$m(e,\gamma) < m_Z - 10$ $m(\mu,\gamma) < m_Z$	
WG3	1	3	0	1	$m(e,\gamma) > m_Z + 10$ $m(\mu,\gamma) > m_Z$	
WG3p	1	≥3	0	1	$m(e,\gamma) > m_Z + 10$ $m(\mu,\gamma) > m_Z$	
WG4p	1	≥4	0	1	$m(e,\gamma) > m_Z + 10$ $m(\mu,\gamma) > m_Z$	
misDY3	1	3	0	1 $\lvert m(e,\gamma)-m_Z\rvert < 10$	e-channel only	
misDY3p	1	≥3	0	1 $\lvert m(e,\gamma)-m_Z\rvert < 10$	e-channel only	
misDY34p	1	≥4	0	1 $\lvert m(e,\gamma)-m_Z\rvert < 10$	e-channel only	
VGmis3	1	3	0	1		WG3+ZG3+misDY3
VGmis3p	1	≥3	0	1		WG3p+ZG3p+misDY3p
VGmis4p	1	≥4	0	1		WG4p+ZG4p+misDY4p
misTT2	1	2	2	1		validation region
SR3	1	3	≥1	1	further binned in M_3	
SR3p	1	≥3	≥1	1	further binned in M_3	
SR4p	1	≥4	≥1	1	further binned in M_3	

Table 3.22: The fiducial phase space definition by the kinematics of different physics objects.

Cut	gen-Photon	gen-Electron	gen-Muon	gen-Jet	gen-b-Jet
$p_T(\gamma^{gen})(GeV)$	≥ 20	≥ 35	≥ 30	≥ 30	≥ 30
$\lvert\eta\rvert$	≤ 1.4442	≤ 2.4	≤ 2.4	≤ 2.4	≤ 2.4
|pdgID|	22	11	13		
status	1	1	1		
other	no meson mother isloated min $\Delta R(\gamma,\ell) > 0.4$	no meson mother	no meson mother	min $\Delta R(jets,\ell) > 0.4$ min $\Delta R(jets,\gamma) > 0.1$	|partonFlavour|=5 min $\Delta R(b-jets,\ell) > 0.4$ min $\Delta R(b-jets,\gamma) > 0.1$

structed containing the data, signal, and background yields in each of the signal and control regions. The systematic variations, as described in this section and in Section 3.1.10, are

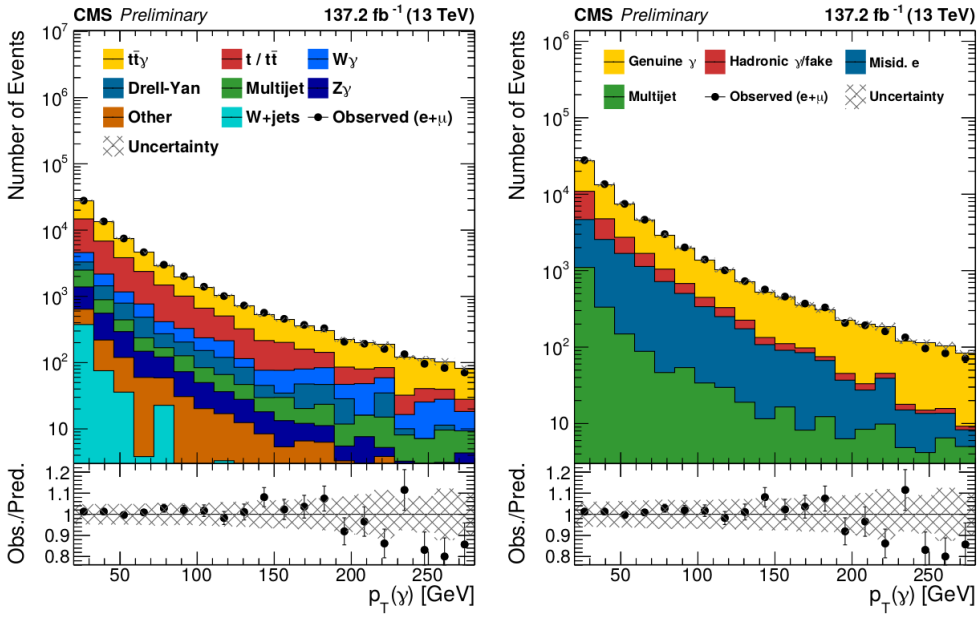


Figure 3.13: Photon p_T distribution for the combined lepton channels (e+μ) in the SR3p baseline selection for RunII categorized in terms of simulated samples (left) and in photon categories (right). The normalization is postfit.

Table 3.23: The overview of the contributions of the backgrounds in the signal region $N_{jet} \geq 3$, $N_{b-jet} \geq 1$ combined in single e and single μ channels in full Run 2 .

RunII, $N_{jet} \geq 3, e + \mu$ *channel*

Sample	events	γ	misiD e	had γ/fake/ PU γ
$t\bar{t}\gamma$	26560.28	26383.36	51.31	125.62 (63.7%/22.8%,13.5%)
$t(\gamma)/tW(\gamma)/t\bar{t}$	18322.05	4203.38	4228.86	9889.81 (74.2%/10.1%/15.7%)
$W\gamma$	3139.58	3133.38	0.00	6.21 (100.0%/ 0.0%/0.0%)
$Z\gamma$	1573.24	1548.85	12.42	11.97 (10.7%/72.7%/18.7%)
W+jets	534.07	0.00	0.08	533.99 (43.8%/4.5%/51.7%)
Drell-Yan	1230.87	125.91	911.17	193.79 (49.9%/18.2%/31.8%)
Multijet (MC)	1147.99	952.56	3.25	192.18 (25.3%/0.6%/74.1%)
Other	630.21	480.63	70.20	79.38 (71.2%/10.4%/18.4%)
MC total	52662.31	36353.03	5278.20	11031.09 (71.2%/10.1%/18.8%)

added as nuisance parameters to the datacard. The statistical uncertainties associated with the signal and background samples are treated as independent nuisance parameters for each

signal and control region, without any correlation.

The test statistic for the inclusive cross section measurement is the profiled likelihood ratio, $q(r) = -2 \ln L(r, \hat{\theta}_r)/L(\hat{r}, \hat{\theta})$, where $\hat{\theta}_r$ reflects the nuisance parameter values that maximise the likelihood function for signal strength r. A multidimensional fit is used to extract the observed cross section of the signal process as well as the associated uncertainties. The quantities that simultaneously maximize the number L are $\hat{r}$ and $\hat{\theta}$. For obtaining the final result, this fitting procedure is used to obtain SF on important backgrounds in-situ. In order to gain confidence, we perform fits in various combinations of control regions to establish the stability of the statistical model.

3.1.9.3 Estimating QCD Multijet Background Using Data-Driven Method

QCD is a minor background in the final signal region of this analysis but has a significant impact while extracting the normalization scale factor (SF) for backgrounds like Z+jets, Z/W+γ and electrons misidentified as photons. Since above mentioned backgrounds are estimated in different control regions where QCD events comes in consideration. As QCD is hard to simulate properly and also available simulated data has very low statistics, it needs to be estimated by a Data Driven technique. The strategy is to get the QCD shape from a QCD enriched region, achieved by relaxing the b-tag requirements along with reversing the relative isolation cut for leptons (high-I_{rel}). All non-QCD Monte Carlos are subtracted from data in the QCD control region (see Figure 3.14 left), which gives the shape of QCD events. Normalization (QCD transfer factor: TF) is extracted by fitting the distribution of W boson transverse mass (M_T) to data along with all other Monte Carlos in the signal region (see Figure 3.14 right). M_T is calculated from the following formula

$$m_T(W) = \sqrt{2 p_T^{\ell} p_T^{miss} [1 - cos(\Delta\phi_{\ell}, \vec{p}_T^{miss})]} \tag{3.7}$$

Where ℓ represents the lepton considered in the event. Fittings are performed at lower jet selections, specifically for $N_{jet} = 2, N_{\gamma} = 0, N_{b-jet} = 0$ and $N_{jet} = 2, N_{\gamma} = 0, N_{b-jet} \geq 1$

where higher QCD statistics is expected to give more accurate fit result. The QCD transfer factor is then applied to higher N_{jet} and N_{γ} selections incorporating an N_{jet}-dependent correction factor $\kappa_{MC}^{QCD}(N_{jet}, N_{b-jet})$. This correction factor is taken from simulation and accounts for a linear N_{jet} dependence of the transfer factor for $N_{b-jet} \geq 1$, although for $N_{b-jet} = 0$ this correction factor is 1.

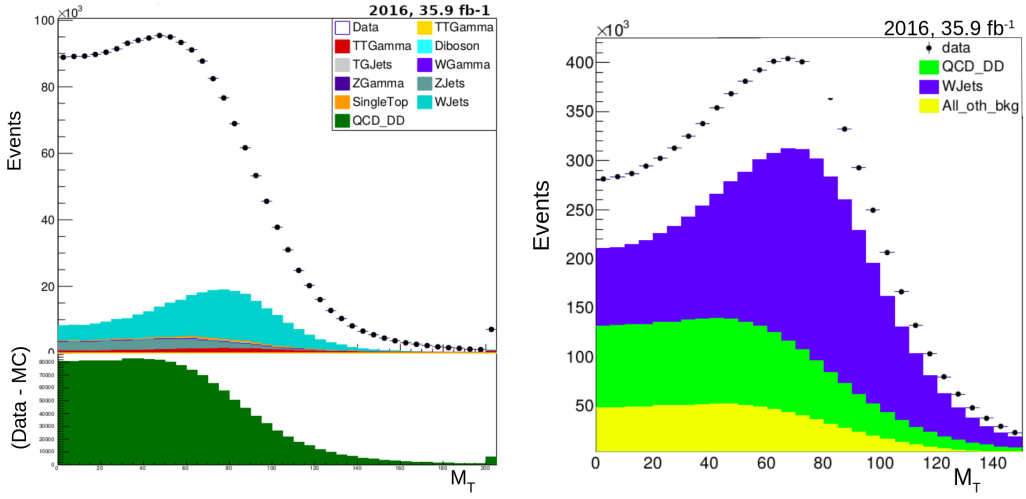


Figure 3.14: QCD multijet process shape template extraction in control region (left), normalization factor extraction by fitting in tight lepton selection region.

Table 3.24: p_T and η binning as used for the estimation of the QCD multijet events

$p_T(\ell)$ bins	0	45	65	80	100	120	∞
$\eta(\ell)$ bins	0	1.479	1.7	2.1	2.4		

Thus, in each bin $1 \leq i \leq N_{QCD}$ according to $p_T(\ell)$ and $\eta(\ell)$ (as shown in Table 3.24), the electroweak simulation (denoted by "subtr.") is subtracted as extracted from the fit to obtain the transfer factor

$$TF_i(N_{b-jet}) = \frac{\text{Data}_{\text{subtr.}}(N_{jet} = 2, N_{b-jet}, \text{low-}I_{rel})}{\text{Data}_{\text{subtr.}}(N_{jet} = 2, N_{b-jet} = 0, \text{high-}I_{rel})} \tag{3.8}$$

$$\kappa_{MC}^{QCD}(N_{jet}, N_{b-jet}) = \frac{\frac{\text{QCD}_{MC}(,N_{b-jet},\text{low-}I_{rel})}{\text{QCD}_{MC}(N_{jet},N_{b-jet}=0,\text{high-}I_{rel})}}{\frac{\text{QCD}_{MC}(N_{jet}=2,N_{b-jet},\text{low-}I_{rel})}{\text{QCD}_{MC}(N_{jet}=2,N_{b-jet}=0,\text{high-}I_{rel})}} \tag{3.9}$$

Finally after summing up the lepton-binned QCD contributions, the estimate thus becomes

$$\text{QCD}^{\text{predicted}} = \sum_{i=1}^{N_{\text{QCD}}} \text{Data}_{\text{subtr.}}(N_{jet}, N_{b-jet} = 0, \text{high-}I_{rel}, \text{bin-}i) \cdot TF_i(N_{b-jet}) \cdot \kappa^{\text{QCD}}_{\text{MC},i}(N_{jet}, N_{b-jet}). \tag{3.10}$$

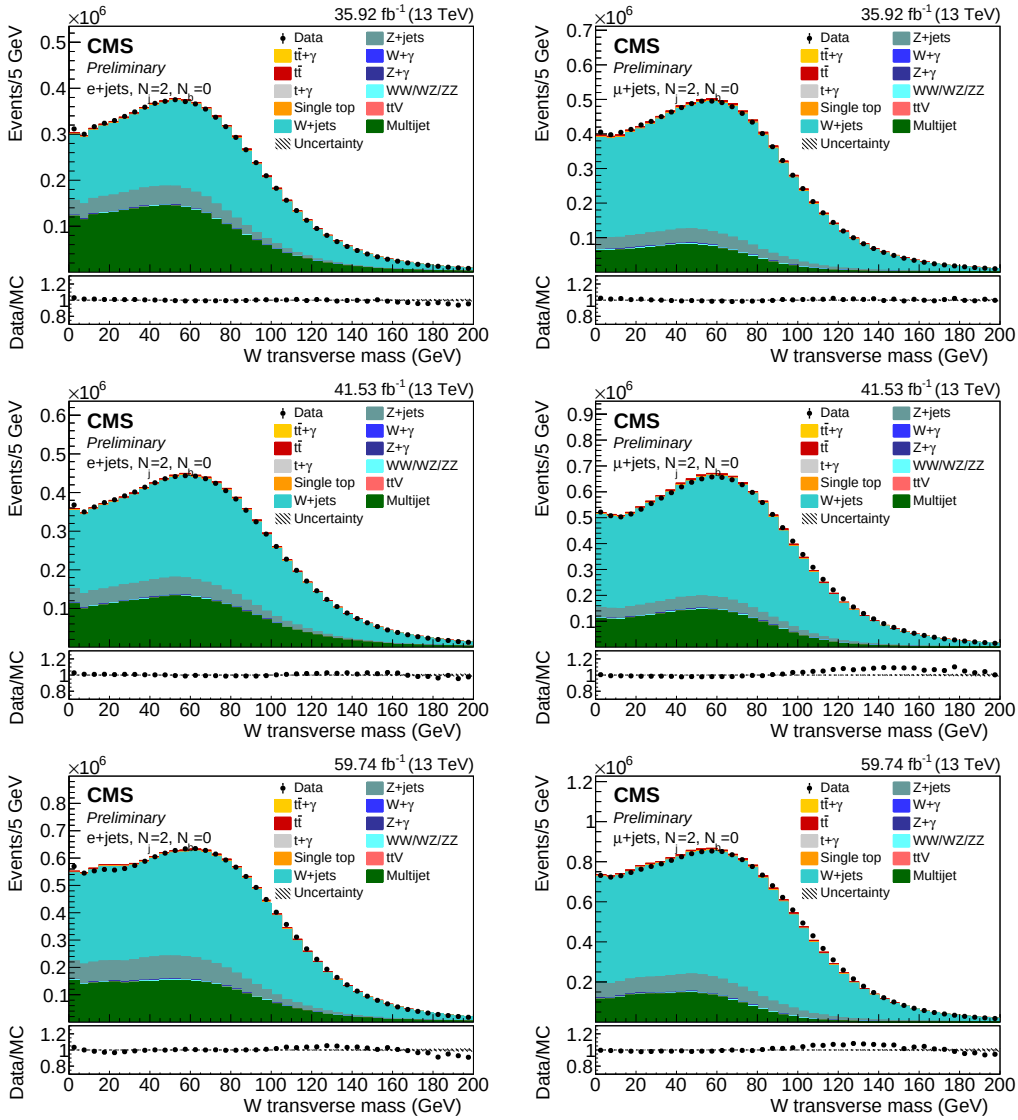


Figure 3.15: Data-MC comparison plots after the fit in zero photon and $N_{jet} = 2$, $N_{b-jet} = 0$ selection. The e+jets ($\mu + jets$) events are shown in left (right) where as the different years starting from 2016 to 2018 have been present from top to bottom in same order.

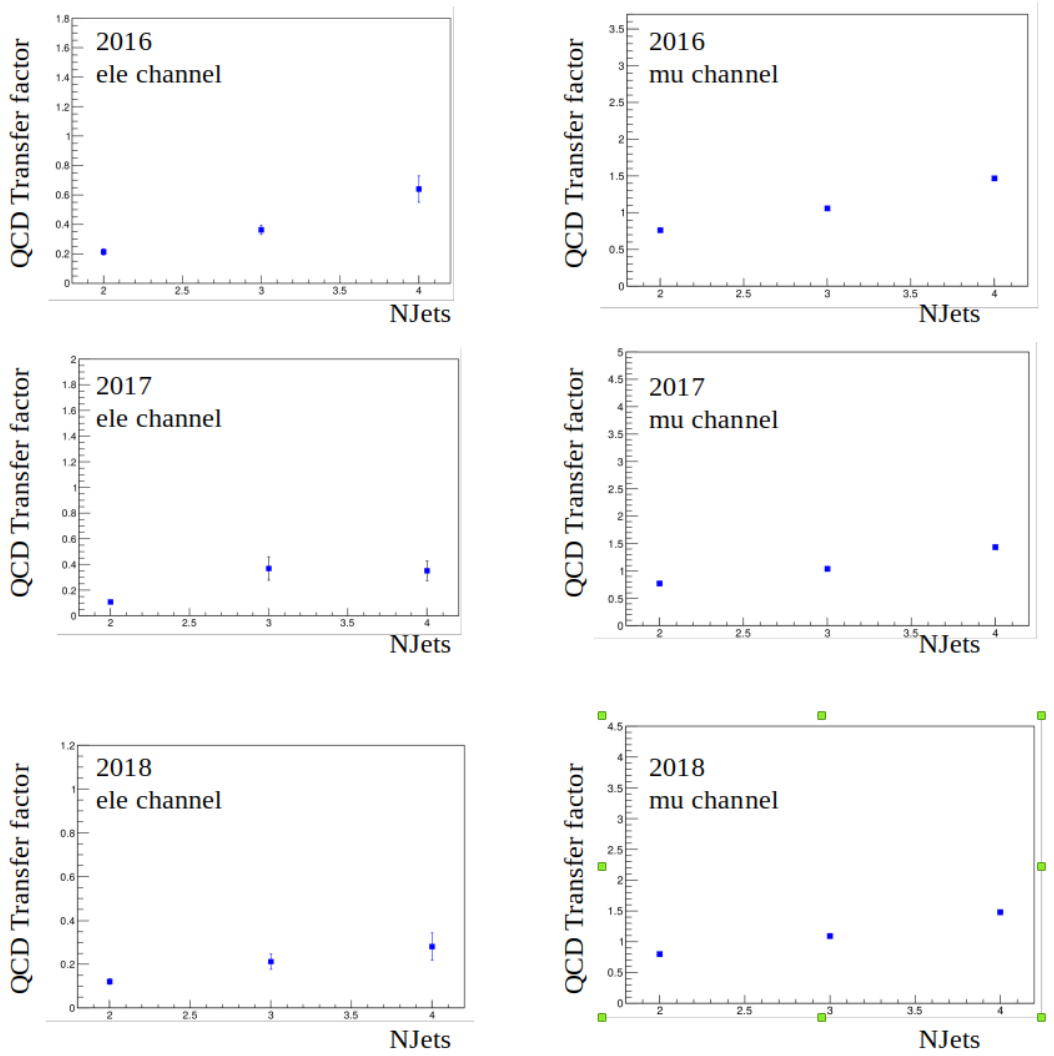


Figure 3.16: Stability of QCD transfer factors calculated from simulation for different numbers of jet selection.

3.1.9.4 Z + Jets correction factor extraction

As the electrons which get misidentified as photon gives very much similar distribution like genuine isolated photon So proper correction factor is needed for misidentified electrons. But most of the misidentified electrons comes from Z+jets process. So any kind of mismodeling of Z+jets process can affect our result. So First we extract the correction factor on Z+jets sample cross section in the $N_{jet} = 3$ and $N_{jet} \geq 4$ regions by normalizing it to data on

the Z peak, i.e. for $|m(\ell,\ell) - m_Z| < 10\text{GeV}$ in the DY3 and DY4p region. A comparison of the ee invariant mass shape is shown in Figure 3.17

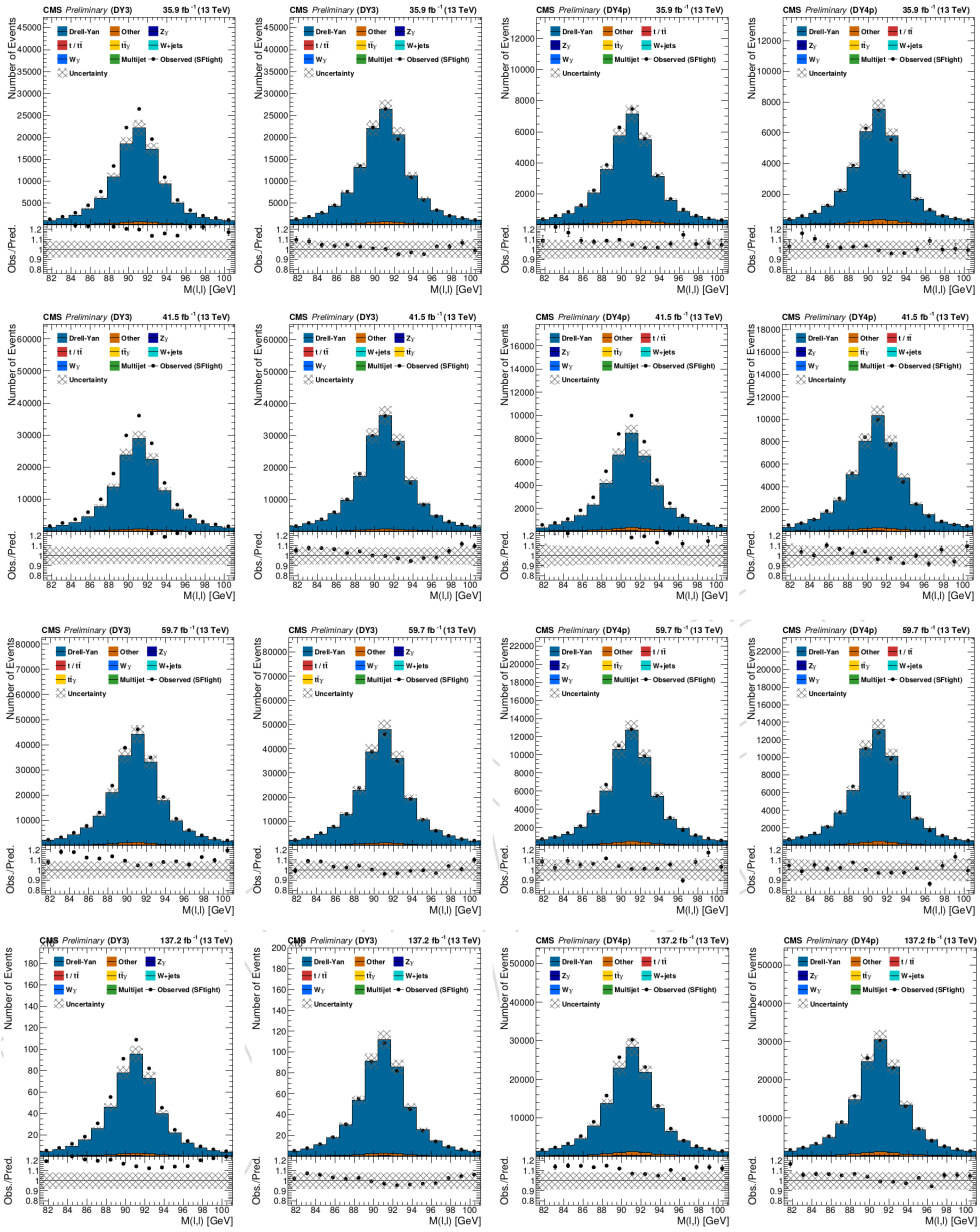


Figure 3.17: Fit of the Z mass peak in the DY3 (left) and DY4p (right) control region in a same-flavor selection (ee, $\mu\mu$). Simulation is normalized with prefit (left) and pos-tfit scale factors (right). From top to bottom, the distributions are for 2016, 2017, 2018, and Run II

3.1.9.5 Misidentified electron and W+γ, Z+γ background extraction

The background from electrons that are mis-identified as photons is measured in control regions with a requirement of $|m(e,\gamma) - m_Z| < 10$ GeV, misDY3 and misDY4p as shown in Table 3.21. This selects Z+jets events with $Z \to ee$, where one of the electrons is misidentified as a photon. We use the Z peak region to extract a scale factor (misIDSF) for the misidentified electron contribution. It is not appropriate to use the ratios of data to simulation in the $m(e,\gamma)$ peak region directly to extract misIDSF, because there are normalization uncertainties on other processes as well.

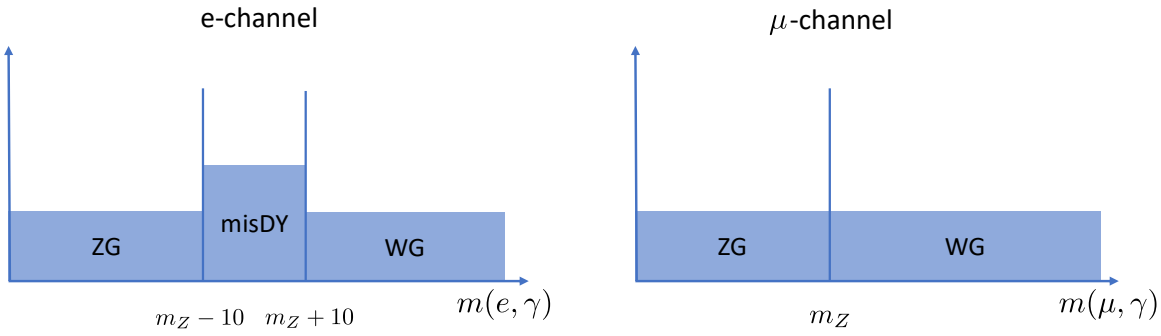


Figure 3.18: Sketch of the ZG, WG, and misDY control region classified with different values of $m(\ell,\gamma)$. The left image represnts the e+jets events and right image shows μ+jets events.

Because the V+γ processes (W+γ, Z+γ) also contribute prompt photons to the $m(e,\gamma)$ peak, sidebands at high and low values of $m(e,\gamma)$ are used to normalize these. From the pretfit distributions in Figure 3.19 (first and third column), it is evident that the normalization of the W+γ and Z+γ contributions affects the misIDSF measurement, because it sizably contributes under the $m(e,\gamma)$ peak region. These processes dominantly contribute to the events with well-identified prompt photons and leptons. Their contribution is constrained outside the peak region by including the ZG3 (ZG4p) and WG3 (WG4p) control regions for $N_{jet} = 3$ ($N_{jet} \geq 4$). We additionally bin the ZG3, ZG4p, WG3, and WG4p regions coarsely in $p_T(\gamma)$ using the thresholds 20, 65, 160, and ∞. This choice avoids bins

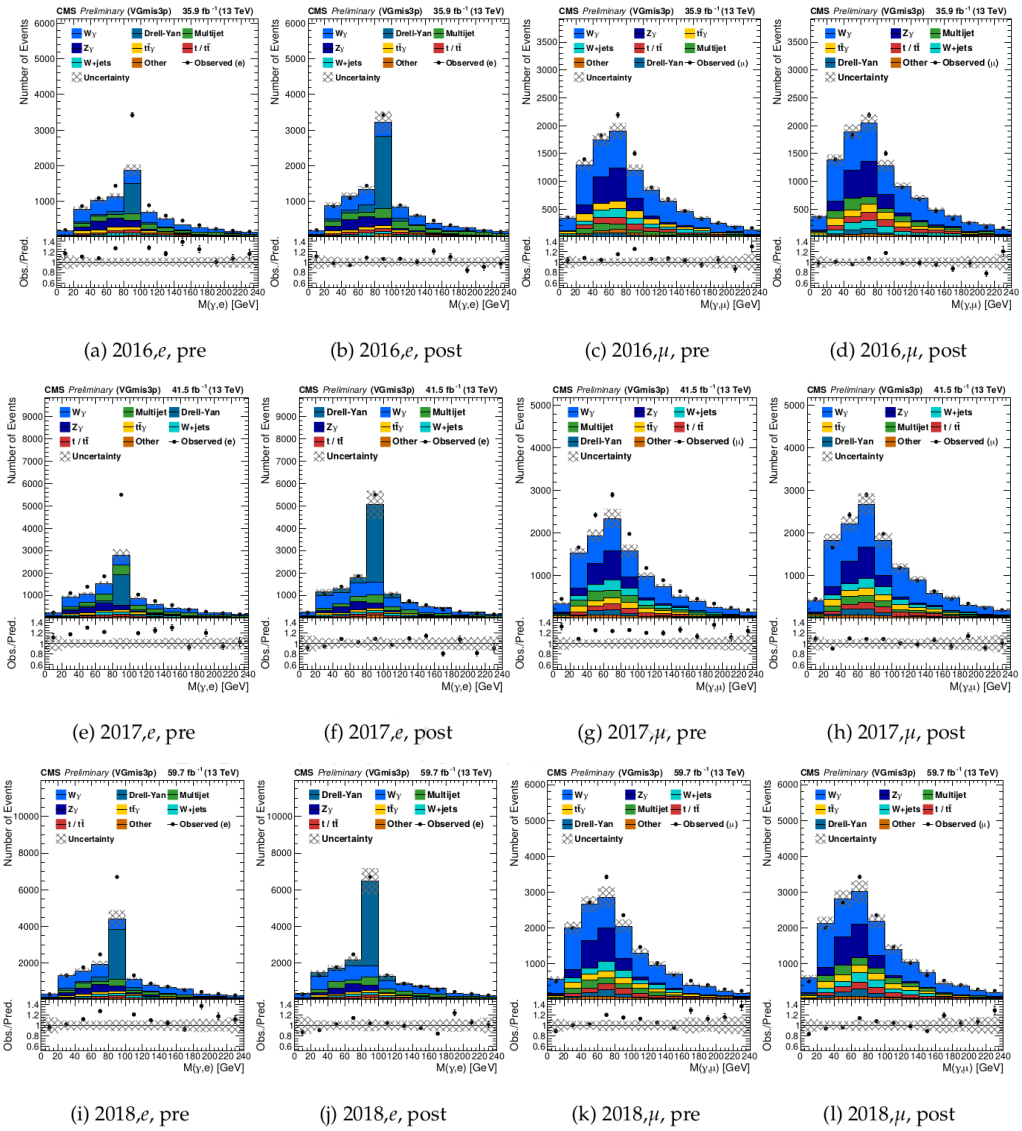
with low event counts even in the partial 2016 result. The naming convention of these regions merely reflects the observation that the Z+γ contribution tends to populate low values of m(e, γ), because in many events the photon is radiated off a lepton with high angular separation. The different relative fractions of W+γ and Z+γ in WG and ZG control regions helps to decorrelate the corresponding SF measurements.

Moreover, there is a QCD contribution in the e channel. In order to further help decorrelate the uncertainties in the QCD multijet prediction from the W+γ and Z+γ normalization, we include the μ channel as well. It does not contain the misidentified electron contribution and the QCD contribution is also smaller, but it provides separate information on the V+γ normalization. Because the probability for a μ to fake a photon is negligible, there is no peak at around the Z mass in m(μ, γ) (see Figure 3.19 c,g,k) and we separate the WG from the ZG region in the μ channel arbitrarily at the Z mass. The distribution of m(ℓ, γ) without application of scale factors is shown in Figure 3.19 a,e,i for the N_{jet} = 3 and $N_{jet} \geq 4$ regions, respectively.

Next, the stability of the extracted scale factors is assessed while progressively combining signal regions and data-taking periods. A table with the extracted scale factors for different jet multiplicities and for different eras is shown in Table 3.25. The extracted scale factors generally agree within uncertainties. We combine the data-taking periods for each jet multiplicity and list the results in Table 3.26.

3.1.9.6 Non-prompt Photon Background Estimation

The hadronic, fake, and small contribution from PU photons is collectively denoted by "non-prompt photons" as described in Sec. 3.1.4. Event yield of the "non-prompt photons" distribution is estimated in the normalization region (LsHc) defined by low $\sigma_{i\eta i\eta}(\gamma)$ and high charged relative isolation $I_{chg}(\gamma)$ where "non-prompt photons" are dominating. Further a normalization correction factor (r_{iso}^{Hs}) is multiplied to get the proper normalization in the



(a) 2016,e, pre (b) 2016,e, post (c) 2016,μ, pre (d) 2016,μ, post

(e) 2017,e, pre (f) 2017,e, post (g) 2017,μ, pre (h) 2017,μ, post

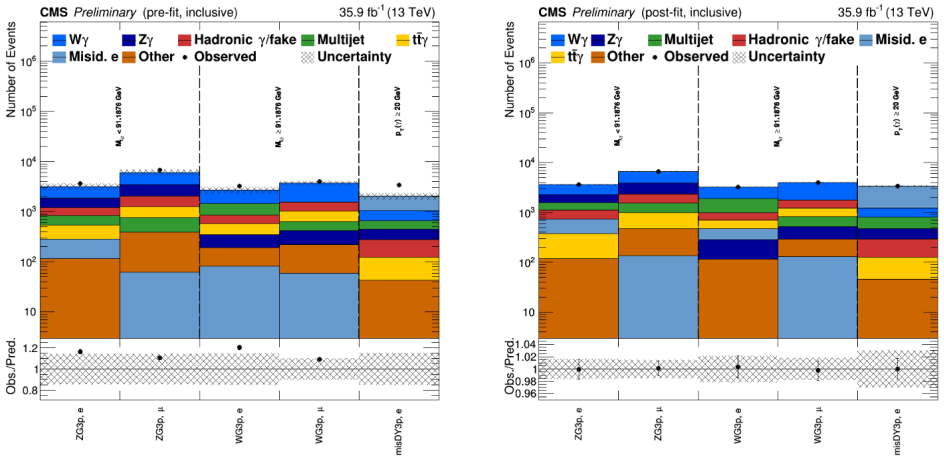
(i) 2018,e, pre (j) 2018,e, post (k) 2018,μ, pre (l) 2018,μ, post

Figure 3.19: The ZG3p, WG3p, and misDY3p control regions. The distributions come in pairs with pretfit (first and third column) and pos-tfit scale factors (second and fourth column) for the e channel (left two columns) and μ channel (right two columns) for 2016 (top), 2017 (middle) and 2018 (bottom).

signal region i.e LsLc defined by low $\sigma_{i\eta i\eta}(\gamma)$ and low charged relative isolation $I_{chg}(\gamma)$.

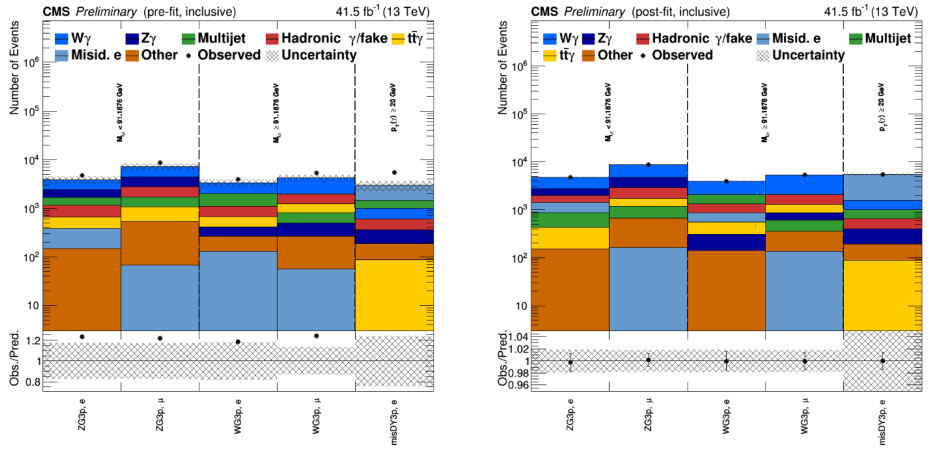
$$r_{iso}^{Hs} = \frac{\text{data}_{\text{HsLc}} - \sum \text{MC}^{\text{prompt}}_{\text{HsLc}}}{\text{data}_{\text{HsHc}} - \sum \text{MC}^{\text{prompt}}_{\text{HsHc}}} \tag{3.11}$$

This estimation procedure can be carried out separately for any signal region, e.g. dif-



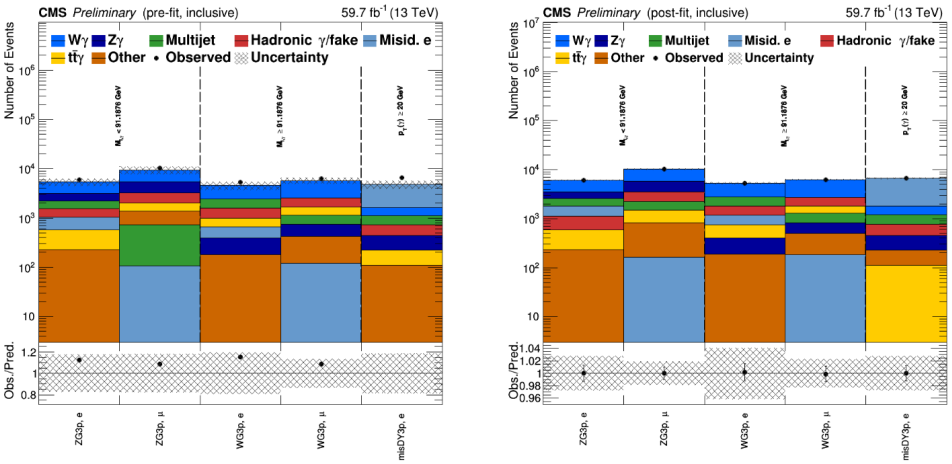
(a) 2016, pre

(b) 2016, post



(c) 2017, pre

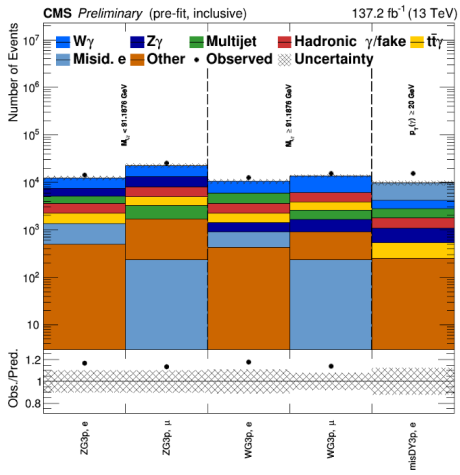
(d) 2017, post



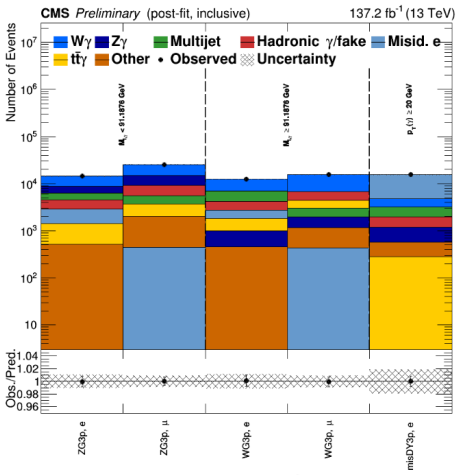
(e) 2018, pre

(f) 2018, post

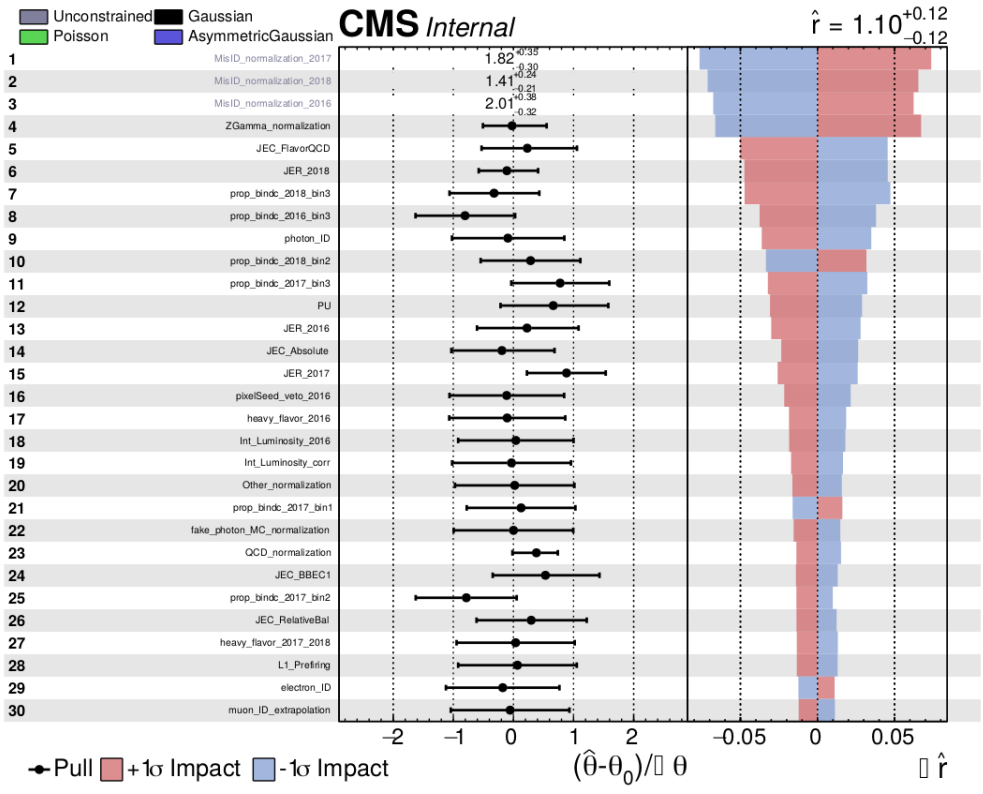
Figure 3.20: pretfit (left) and pos-tfit (right) region plots of the W+γ SF fits for $N_{jet} \geq 3$ for 2016 (top), 2017 (middle), and 2018 (bottom)



(a) RunII, pre



(b) RunII, post



(c) RunII, impact

Figure 3.21: pretfit (top left), pos-tfit (top right) region plots and impact plot (bottom) of the W+γ SF fits for $N_{jet} \geq 3$ for RunII. The $W+\gamma$ SF is used as POI.

Table 3.25: Extracted scale factors for the contribution from electrons misidentified as photons (misIDSF) and the Z+γ, W+γ, multijets and Z+jets processes in the $N_{b-jet} = 0$ bins.

	N_{jet}	**misIDSF**	**Zγ SF**	**Wγ SF**	**multijets SF**	**DY+jets SF**
2016	**2j**	2.3 ± 0.27	0.88 ± 0.1	1.16 ± 0.08	1.22 ± 0.18	1.18 ± 0.06
	3j	2.3 ± 0.32	1.03 ± 0.16	1.05 ± 0.13	1.42 ± 0.24	1.2 ± 0.07
	4j	1.88 ± 0.33	1.15 ± 0.27	1.23 ± 0.18	1.33 ± 0.44	1.08 ± 0.07
	5j	1.54 ± 0.87	1.03 ± 0.29	1.36 ± 0.34	1.03 ± 0.44	1.01 ± 0.09
	$\geq$2j	2.26 ± 0.27	0.93 ± 0.09	1.13 ± 0.08	1.33 ± 0.14	1.18 ± 0.07
	$\geq$3j	2.11 ± 0.45	1.06 ± 0.21	1.04 ± 0.16	1.47 ± 0.21	1.17 ± 0.14
	$\geq$4j	1.92 ± 0.32	1.19 ± 0.26	1.26 ± 0.18	1.35 ± 0.39	1.06 ± 0.08
2017	**2j**	2.6 ± 0.34	0.96 ± 0.1	1.18 ± 0.08	0.97 ± 0.17	1.25 ± 0.09
	3j	2.3 ± 0.34	1.06 ± 0.19	1.29 ± 0.14	0.89 ± 0.23	1.26 ± 0.11
	4j	2.99 ± 0.65	1.04 ± 0.28	1.61 ± 0.24	0.61 ± 0.33	1.22 ± 0.12
	5j	2.22 ± 1.07	1.06 ± 0.29	2.11 ± 0.5	0.86 ± 0.38	1.28 ± 0.16
	$\geq$2j	2.59 ± 0.34	0.98 ± 0.1	1.24 ± 0.08	0.94 ± 0.17	1.25 ± 0.1
	$\geq$3j	2.27 ± 0.55	1.05 ± 0.23	1.33 ± 0.20	0.85 ± 0.2	1.25 ± 0.27
	$\geq$4j	2.83 ± 0.6	1.08 ± 0.28	1.71 ± 0.25	0.71 ± 0.32	1.23 ± 0.13
2018	**2j**	1.56 ± 0.18	0.93 ± 0.09	1.11 ± 0.08	0.94 ± 0.15	1.11 ± 0.08
	3j	1.49 ± 0.19	0.99 ± 0.16	1.08 ± 0.14	1.29 ± 0.37	1.09 ± 0.09
	4j	1.74 ± 0.28	1.06 ± 0.25	1.09 ± 0.2	1.1 ± 0.43	1.02 ± 0.1
	5j	1.17 ± 0.53	1.04 ± 0.29	1.46 ± 0.37	1.04 ± 0.52	1.07 ± 0.12
	$\geq$2j	1.56 ± 0.18	0.95 ± 0.09	1.11 ± 0.08	1.03 ± 0.17	1.1 ± 0.08
	$\geq$3j	1.46 ± 0.28	1.0 ± 0.2	1.06 ± 0.19	1.2 ± 0.46	1.07 ± 0.21
	$\geq$4j	1.7 ± 0.27	1.09 ± 0.27	1.28 ± 0.2	1.0 ± 0.45	1.04 ± 0.11

ferentially in the $p_T(\gamma)$ bins. The $I_{chg}(\gamma)$ threshold of 1.141 that separates the Lc and the Hc regions and the $\sigma_{i\eta i\eta}(\gamma) < 0.01015$ and $\sigma_{i\eta i\eta}(\gamma > 0.011$ requirement that define the Ls and Hs regions are generically sketched in Figure 3.22.

Along with r_{iso}^{Hs} a second order correction factor (κ_{MC}) is applied to this ABCD method. This second order correction comes into the picture due to the residual correlation of the $\sigma_{i\eta i\eta}$ shower shape variable with the amount of activity in the vicinity parameterized by I_{chg} of the photon candidate. In Figure 3.23, it is seen that the $\sigma_{i\eta i\eta}$ shape gradually evolves with I_{chg}. The shape difference is corrected for by this simulation based factor κ_{MC}.

Table 3.26: Extracted scale factors for the contribution from electrons misidentified as photons (misIDSF) and the Z+γ, W+γ, multijets and DY+jets processes.

	N_{jet}	**misIDSF 16/17/18**			Z +γ **SF**	W +γ**SF**	**multijets SF**	**DY+jets SF**
RunII	**2j**	2.27 ± 0.26	2.48 ± 0.29	1.59 ± 0.17	0.92 ± 0.07	1.14 ± 0.06	1.11 ± 0.14	1.18 ± 0.06
	3j	2.32 ± 0.31	2.10 ± 0.29	1.52 ± 0.18	1.03 ± 0.12	1.13 ± 0.09	1.24 ± 0.2	1.18 ± 0.07
	4j	1.88 ± 0.33	2.53 ± 0.52	1.77 ± 0.27	1.12 ± 0.23	1.28 ± 0.14	0.87 ± 0.23	1.09 ± 0.07
	5j	1.40 ± 0.84	1.83 ± 0.88	1.33 ± 0.50	1.16 ± 0.32	1.5 ± 0.27	1.02 ± 0.37	1.06 ± 0.09
	≥2j	2.27 ± 0.26	2.41 ± 0.29	1.60 ± 0.17	0.95 ± 0.07	1.15 ± 0.06	1.18 ± 0.13	1.18 ± 0.06
	≥3j	2.01 ± 0.38	1.82 ± 0.35	1.41 ± 0.24	0.99 ± 0.16	1.1 ± 0.12	1.16 ± 0.21	1.16 ± 0.14
	≥4j	1.94 ± 0.32	2.36 ± 0.47	1.69 ± 0.26	1.22 ± 0.21	1.36 ± 0.14	0.96 ± 0.23	1.08 ± 0.08
	3j$\otimes$ **≥4j**	1.87 ± 0.23	1.68 ± 0.21	1.33 ± 0.15	0.86 ± 0.09	1.02 ± 0.08	0.99 ± 0.15	1.39 ± 0.05

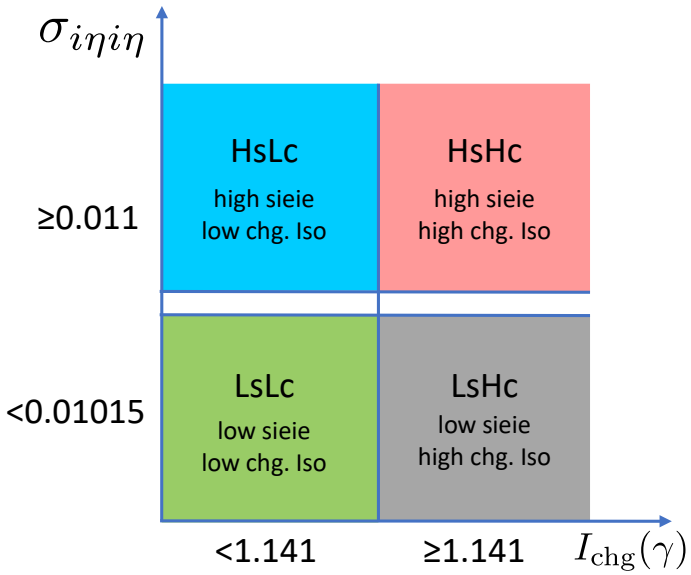


Figure 3.22: Sketch of the $I_{chg}(\gamma)$ and $\sigma_{i\eta i\eta}(\gamma)$ thresholds in the non-prompt photon estimation.

$$\text{data-driven non-prompt } \gamma = \underbrace{\underbrace{\left(\text{data}_{\text{ LsHc}} - \sum \text{MC}^{\text{ prompt}}_{\text{ LsHc}}\right)}_{\text{non-prompt } \gamma \text{ in high chg.Iso}} \times \underbrace{\frac{\text{data}_{\text{ HsLc}} - \sum \text{MC}^{\text{ prompt}}_{\text{ HsLc}}}{\text{data}_{\text{ HsHc}} - \sum \text{MC}^{\text{ prompt}}_{\text{ HsHc}}}}_{r^{\text{HS}}_{\text{iso}}\text{: data normalization}}}_{\text{ABCD like estimation for LsLc}} \times \underbrace{\frac{\frac{\sum \text{MC}^{\text{ non-prompt}}_{\text{ LsLc}}}{\sum \text{MC}^{\text{ non-prompt}}_{\text{ LsHc}}}}{\frac{\sum \text{MC}^{\text{ non-prompt}}_{\text{ HsLc}}}{\sum \text{MC}^{\text{ non-prompt}}_{\text{ HsHc}}}}}_{\kappa_{\text{MC}}\text{: MC shape correction}}$$

Figure 3.24 shows the shapes whose shape difference (double-ratio) corresponds to κ_{MC} for different years and lepton flavor.

Table 3.27: Results of the hadronic and fake background estimation for bins in $p_T(\gamma)$, for both lepton flavors, for the $N_{jet} = 3$ and $N_{jet} \geq 4$ regions in 2016.

SR3 2016												
	e channel						**μ channel**					
Region	fakes high chgIso	r_{iso}^{HS}	κ_{MC}	**DD-fakes**	**MC-fakes**	**DD-Fakes/MC-Fakes**	fakes high chgIso	r_{iso}^{HS}	κ_{MC}	DD-fakes	MC-fakes	**DD-Fakes/MC-Fakes**
inclusive	4058.82 ± 77.66	0.13 ± 0.00	1.11 ± 0.03	$\mathbf{579.25 \pm 26.68}$	$\mathbf{588.97 \pm 12.51}$	$\mathbf{0.98 \pm 0.05}$	6365.73 ± 96.78	0.13 ± 0.00	1.09 ± 0.02	$\mathbf{905.29 \pm 33.58}$	$\mathbf{953.40 \pm 16.21}$	$\mathbf{0.95 \pm 0.04}$
$20 \leq p_T(\gamma) < 35$	2567.36 ± 60.82	0.15 ± 0.01	1.01 ± 0.03	$\mathbf{377.85 \pm 22.44}$	$\mathbf{360.16 \pm 10.05}$	$\mathbf{1.05 \pm 0.07}$	4114.58 ± 75.85	0.14 ± 0.00	0.98 ± 0.03	$\mathbf{568.31 \pm 27.82}$	$\mathbf{587.60 \pm 14.05}$	$\mathbf{0.97 \pm 0.05}$
$35 \leq p_T(\gamma) < 50$	831.53 ± 34.34	0.13 ± 0.01	1.13 ± 0.06	$\mathbf{119.43 \pm 12.01}$	$\mathbf{118.42 \pm 4.29}$	$\mathbf{1.01 \pm 0.11}$	1243.84 ± 43.35	0.12 ± 0.01	1.16 ± 0.05	$\mathbf{178.36 \pm 14.59}$	$\mathbf{200.74 \pm 6.41}$	$\mathbf{0.89 \pm 0.08}$
$50 \leq p_T(\gamma) < 65$	322.76 ± 22.93	0.09 ± 0.01	1.27 ± 0.12	$\mathbf{37.84 \pm 6.38}$	$\mathbf{54.29 \pm 4.48}$	$\mathbf{0.70 \pm 0.13}$	474.17 ± 28.23	0.13 ± 0.01	1.13 ± 0.07	$\mathbf{69.93 \pm 8.20}$	$\mathbf{79.63 \pm 2.89}$	$\mathbf{0.88 \pm 0.11}$
$65 \leq p_T(\gamma) < 80$	140.99 ± 15.09	0.11 ± 0.02	1.41 ± 0.16	$\mathbf{21.57 \pm 4.69}$	$\mathbf{28.72 \pm 3.47}$	$\mathbf{0.75 \pm 0.19}$	216.76 ± 18.99	0.11 ± 0.01	1.50 ± 0.14	$\mathbf{36.23 \pm 6.30}$	$\mathbf{43.04 \pm 3.30}$	$\mathbf{0.84 \pm 0.16}$
$80 \leq p_T(\gamma) < 120$	146.53 ± 15.83	0.10 ± 0.01	1.11 ± 0.13	$\mathbf{15.70 \pm 3.28}$	$\mathbf{18.91 \pm 1.32}$	$\mathbf{0.83 \pm 0.18}$	219.32 ± 19.42	0.11 ± 0.01	1.30 ± 0.11	$\mathbf{31.07 \pm 5.04}$	$\mathbf{32.31 \pm 1.98}$	$\mathbf{0.96 \pm 0.17}$
$120 \leq p_T(\gamma) < 160$	25.50 ± 9.05	0.12 ± 0.03	1.54 ± 0.34	$\mathbf{4.72 \pm 2.21}$	$\mathbf{5.30 \pm 0.74}$	$\mathbf{0.89 \pm 0.44}$	61.28 ± 10.55	0.08 ± 0.02	1.50 ± 0.24	$\mathbf{6.90 \pm 2.19}$	$\mathbf{6.96 \pm 0.83}$	$\mathbf{0.99 \pm 0.34}$
$160 \leq p_T(\gamma) < 200$	3.67 ± 6.33	0.03 ± 0.02	3.30 ± 2.00	$\mathbf{0.41 \pm 0.79}$	$\mathbf{2.81 \pm 1.63}$	$\mathbf{0.15 \pm 0.29}$	28.84 ± 6.49	0.08 ± 0.03	0.80 ± 0.30	$\mathbf{1.77 \pm 0.97}$	$\mathbf{1.55 \pm 0.38}$	$\mathbf{1.14 \pm 0.69}$
$200 \leq p_T(\gamma) < 260$	13.78 ± 4.30	0.04 ± 0.03	0.51 ± 0.48	$\mathbf{0.27 \pm 0.34}$	$\mathbf{0.27 \pm 0.23}$	$\mathbf{1.01 \pm 1.52}$	5.70 ± 5.37	0.12 ± 0.04	1.50 ± 0.83	$\mathbf{1.01 \pm 1.16}$	$\mathbf{1.16 \pm 0.54}$	$\mathbf{0.87 \pm 1.08}$
$260 \leq p_T(\gamma) < 320$	0.83 ± 1.53	0.15 ± 0.09	1.03 ± 1.04	$\mathbf{0.13 \pm 0.28}$	$\mathbf{0.05 \pm 0.04}$	$\mathbf{2.68 \pm 6.33}$	0.10 ± 0.10	0.09 ± 0.07	1.49 ± 1.19	$\mathbf{0.01 \pm 0.02}$	$\mathbf{0.30 \pm 0.18}$	$\mathbf{0.05 \pm 0.07}$
$320 \leq p_T(\gamma)$	5.87 ± 2.65	0.01 ± 0.01	2.03 ± 2.37	$\mathbf{0.08 \pm 0.13}$	$\mathbf{0.05 \pm 0.05}$	$\mathbf{1.59 \pm 3.03}$	1.60 ± 3.61	0.05 ± 0.05	1.49 ± 1.10	$\mathbf{0.12 \pm 0.31}$	$\mathbf{0.12 \pm 0.07}$	$\mathbf{0.99 \pm 2.64}$

SR4p 2016												
	e channel						**μ channel**					
Region	fakes high chgIso	r_{iso}^{HS}	κ_{MC}	**DD-fakes**	**MC-fakes**	**DD-Fakes/MC-Fakes**	fakes high chgIso	r_{iso}^{HS}	κ_{MC}	DD-fakes	MC-fakes	**DD-Fakes/MC-Fakes**
inclusive	3965.53 ± 74.29	0.12 ± 0.00	1.18 ± 0.03	$\mathbf{558.40 \pm 27.65}$	$\mathbf{521.95 \pm 18.13}$	$\mathbf{1.07 \pm 0.06}$	5950.88 ± 91.45	0.12 ± 0.00	1.18 ± 0.03	$\mathbf{861.27 \pm 32.73}$	$\mathbf{816.83 \pm 15.79}$	$\mathbf{1.05 \pm 0.04}$
$20 \leq p_T(\gamma) < 35$	2565.69 ± 58.55	0.14 ± 0.01	1.11 ± 0.04	$\mathbf{391.80 \pm 24.43}$	$\mathbf{341.80 \pm 16.57}$	$\mathbf{1.15 \pm 0.09}$	3813.68 ± 71.81	0.14 ± 0.00	1.06 ± 0.03	$\mathbf{583.99 \pm 27.80}$	$\mathbf{527.95 \pm 13.34}$	$\mathbf{1.11 \pm 0.06}$
$35 \leq p_T(\gamma) < 50$	754.41 ± 32.25	0.11 ± 0.01	1.23 ± 0.07	$\mathbf{102.59 \pm 11.08}$	$\mathbf{97.89 \pm 3.91}$	$\mathbf{1.05 \pm 0.12}$	1124.60 ± 39.76	0.10 ± 0.01	1.21 ± 0.06	$\mathbf{140.70 \pm 12.82}$	$\mathbf{152.89 \pm 5.73}$	$\mathbf{0.92 \pm 0.09}$
$50 \leq p_T(\gamma) < 65$	265.36 ± 20.78	0.10 ± 0.01	1.07 ± 0.16	$\mathbf{27.65 \pm 5.84}$	$\mathbf{36.09 \pm 5.01}$	$\mathbf{0.77 \pm 0.19}$	446.85 ± 26.50	0.10 ± 0.01	1.37 ± 0.10	$\mathbf{64.01 \pm 8.58}$	$\mathbf{66.40 \pm 3.97}$	$\mathbf{0.96 \pm 0.14}$
$65 \leq p_T(\gamma) < 80$	156.24 ± 15.24	0.14 ± 0.02	1.27 ± 0.14	$\mathbf{28.68 \pm 5.90}$	$\mathbf{18.62 \pm 1.44}$	$\mathbf{1.54 \pm 0.34}$	239.88 ± 19.38	0.11 ± 0.01	1.57 ± 0.20	$\mathbf{40.96 \pm 7.95}$	$\mathbf{31.50 \pm 3.54}$	$\mathbf{1.30 \pm 0.29}$
$80 \leq p_T(\gamma) < 120$	145.02 ± 15.93	0.06 ± 0.01	1.22 ± 0.12	$\mathbf{10.91 \pm 2.74}$	$\mathbf{16.80 \pm 1.32}$	$\mathbf{0.65 \pm 0.17}$	228.67 ± 18.82	0.08 ± 0.01	1.11 ± 0.10	$\mathbf{19.64 \pm 3.50}$	$\mathbf{23.32 \pm 1.54}$	$\mathbf{0.84 \pm 0.16}$
$120 \leq p_T(\gamma) < 160$	43.72 ± 8.61	0.06 ± 0.02	2.33 ± 0.87	$\mathbf{5.99 \pm 3.02}$	$\mathbf{8.81 \pm 3.11}$	$\mathbf{0.68 \pm 0.42}$	41.02 ± 9.72	0.07 ± 0.01	1.43 ± 0.22	$\mathbf{4.13 \pm 1.46}$	$\mathbf{7.39 \pm 0.95}$	$\mathbf{0.56 \pm 0.21}$
$160 \leq p_T(\gamma) < 200$	16.84 ± 5.27	0.02 ± 0.01	0.70 ± 0.34	$\mathbf{0.22 \pm 0.21}$	$\mathbf{0.98 \pm 0.31}$	$\mathbf{0.23 \pm 0.23}$	26.47 ± 6.59	0.07 ± 0.02	1.22 ± 0.32	$\mathbf{2.39 \pm 1.16}$	$\mathbf{2.03 \pm 0.42}$	$\mathbf{1.18 \pm 0.62}$
$200 \leq p_T(\gamma) < 260$	10.85 ± 4.18	0.04 ± 0.02	1.64 ± 0.75	$\mathbf{0.68 \pm 0.59}$	$\mathbf{0.75 \pm 0.26}$	$\mathbf{0.91 \pm 0.85}$	9.87 ± 4.80	0.07 ± 0.03	8.95 ± 5.65	$\mathbf{6.43 \pm 5.66}$	$\mathbf{4.84 \pm 2.59}$	$\mathbf{1.33 \pm 1.37}$
$260 \leq p_T(\gamma) < 320$	4.91 ± 2.87	0.00 ± 0.00	0.89 ± 0.71	$\mathbf{0.01 \pm 0.02}$	$\mathbf{0.19 \pm 0.14}$	$\mathbf{0.08 \pm 0.12}$	11.49 ± 4.32	0.06 ± 0.04	2.86 ± 1.85	$\mathbf{1.98 \pm 2.08}$	$\mathbf{0.33 \pm 0.17}$	$\mathbf{5.95 \pm 6.97}$
$320 \leq p_T(\gamma)$	2.49 ± 2.51	0.07 ± 0.08	0.29 ± 0.19	$\mathbf{0.05 \pm 0.08}$	$\mathbf{0.02 \pm 0.01}$	$\mathbf{2.09 \pm 3.51}$	8.35 ± 3.64	0.00 ± 0.00	1.98 ± 1.71	$\mathbf{0.04 \pm 0.06}$	$\mathbf{0.19 \pm 0.14}$	$\mathbf{0.21 \pm 0.33}$

Table 3.28: Results of the hadronic and fake background estimation for bins in $p_T(\gamma)$, for both lepton flavors, for the $N_{jet} = 3$ and $N_{jet} \geq 4$ regions in 2017.

SR3 2017												
	e channel						**μ channel**					
Region	fakes high chgIso	$r_{\text{iso}}^{\text{HS}}$	κ_{MC}	**DD-fakes**	**MC-fakes**	**DD-Fakes/MC-Fakes**	fakes high chgIso	$r_{\text{iso}}^{\text{HS}}$	κ_{MC}	DD-fakes	MC-fakes	**DD-Fakes/MC-Fakes**
inclusive	5018.72 $\pm$ 95.48	0.12 $\pm$ 0.00	1.36 $\pm$ 0.05	**788.98 $\pm$ 44.58**	**887.53 $\pm$ 47.87**	**0.89 $\pm$ 0.07**	7241.72 $\pm$ 107.51	0.11 $\pm$ 0.00	1.28 $\pm$ 0.04	**1034.67 $\pm$ 46.22**	**1064.44 $\pm$ 34.33**	**0.97 $\pm$ 0.05**
$20 \leq p_T(\gamma) < 35$	2984.14 $\pm$ 71.31	0.13 $\pm$ 0.01	1.32 $\pm$ 0.07	**525.98 $\pm$ 38.84**	**620.51 $\pm$ 46.56**	**0.85 $\pm$ 0.09**	4475.36 $\pm$ 83.41	0.12 $\pm$ 0.00	1.16 $\pm$ 0.05	**638.91 $\pm$ 36.82**	**661.61 $\pm$ 31.24**	**0.97 $\pm$ 0.07**
$35 \leq p_T(\gamma) < 50$	1038.02 $\pm$ 41.65	0.10 $\pm$ 0.01	1.35 $\pm$ 0.11	**133.48 $\pm$ 17.28**	**137.94 $\pm$ 8.48**	**0.97 $\pm$ 0.14**	1505.27 $\pm$ 47.45	0.11 $\pm$ 0.01	1.35 $\pm$ 0.09	**213.17 $\pm$ 20.87**	**207.37 $\pm$ 9.26**	**1.03 $\pm$ 0.11**
$50 \leq p_T(\gamma) < 65$	413.05 $\pm$ 36.47	0.12 $\pm$ 0.01	1.37 $\pm$ 0.15	**67.07 $\pm$ 11.99**	**55.89 $\pm$ 4.21**	**1.20 $\pm$ 0.23**	619.42 $\pm$ 32.96	0.11 $\pm$ 0.01	1.38 $\pm$ 0.15	**94.72 $\pm$ 14.73**	**90.66 $\pm$ 8.24**	**1.04 $\pm$ 0.19**
$65 \leq p_T(\gamma) < 80$	239.21 $\pm$ 20.45	0.11 $\pm$ 0.02	1.46 $\pm$ 0.23	**37.13 $\pm$ 9.30**	**34.54 $\pm$ 4.11**	**1.08 $\pm$ 0.30**	272.27 $\pm$ 22.62	0.08 $\pm$ 0.01	1.89 $\pm$ 0.24	**43.17 $\pm$ 9.07**	**47.24 $\pm$ 4.58**	**0.91 $\pm$ 0.21**
$80 \leq p_T(\gamma) < 120$	209.84 $\pm$ 18.86	0.06 $\pm$ 0.01	1.69 $\pm$ 0.29	**22.83 $\pm$ 6.14**	**29.55 $\pm$ 3.72**	**0.77 $\pm$ 0.23**	260.34 $\pm$ 22.37	0.08 $\pm$ 0.01	1.67 $\pm$ 0.23	**34.15 $\pm$ 7.20**	**42.18 $\pm$ 4.28**	**0.81 $\pm$ 0.19**
$120 \leq p_T(\gamma) < 160$	75.13 $\pm$ 10.27	0.08 $\pm$ 0.02	1.01 $\pm$ 0.30	**5.96 $\pm$ 2.46**	**5.85 $\pm$ 1.42**	**1.02 $\pm$ 0.49**	65.15 $\pm$ 10.87	0.06 $\pm$ 0.02	1.37 $\pm$ 0.35	**4.96 $\pm$ 2.04**	**8.27 $\pm$ 1.67**	**0.60 $\pm$ 0.28**
$160 \leq p_T(\gamma) < 200$	31.89 $\pm$ 6.69	0.05 $\pm$ 0.03	1.18 $\pm$ 0.64	**1.89 $\pm$ 1.47**	**1.40 $\pm$ 0.62**	**1.35 $\pm$ 1.21**	27.57 $\pm$ 7.61	0.10 $\pm$ 0.03	2.94 $\pm$ 1.61	**8.11 $\pm$ 5.59**	**5.94 $\pm$ 2.51**	**1.37 $\pm$ 1.10**
$200 \leq p_T(\gamma) < 260$	16.94 $\pm$ 5.14	0.01 $\pm$ 0.02	2.12 $\pm$ 1.54	**0.39 $\pm$ 0.86**	**1.48 $\pm$ 0.73**	**0.27 $\pm$ 0.60**	11.98 $\pm$ 7.07	0.02 $\pm$ 0.02	2.00 $\pm$ 1.44	**0.40 $\pm$ 0.56**	**1.15 $\pm$ 0.62**	**0.34 $\pm$ 0.52**
$260 \leq p_T(\gamma) < 320$	5.43 $\pm$ 3.06	0.01 $\pm$ 0.01	2.89 $\pm$ 3.08	**0.09 $\pm$ 0.14**	**0.36 $\pm$ 0.26**	**0.24 $\pm$ 0.42**	0.70 $\pm$ 3.92	0.23 $\pm$ 0.13	0.03 $\pm$ 0.02	**0.00 $\pm$ 0.03**	**0.01 $\pm$ 0.00**	**0.44 $\pm$ 2.48**
$320 \leq p_T(\gamma)$	5.07 $\pm$ 2.66	0.01 $\pm$ 0.01	24.08 $\pm$ 23.21	**1.22 $\pm$ 1.85**	**0.01 $\pm$ 0.00**	**180.51 $\pm$ 280.18**	3.66 $\pm$ 3.13	0.11 $\pm$ 0.12	0.07 $\pm$ 0.06	**0.03 $\pm$ 0.05**	**0.01 $\pm$ 0.00**	**2.46 $\pm$ 4.00**

SR4p 2017												
	e channel						**μ channel**					
Region	fakes high chgIso	$r_{\text{iso}}^{\text{HS}}$	κ_{MC}	**DD-fakes**	**MC-fakes**	**DD-Fakes/MC-Fakes**	fakes high chgIso	$r_{\text{iso}}^{\text{HS}}$	κ_{MC}	DD-fakes	MC-fakes	**DD-Fakes/MC-Fakes**
inclusive	4535.95 $\pm$ 89.66	0.11 $\pm$ 0.00	1.42 $\pm$ 0.05	**719.89 $\pm$ 39.01**	**762.50 $\pm$ 41.12**	**0.94 $\pm$ 0.07**	6249.00 $\pm$ 103.72	0.10 $\pm$ 0.00	1.39 $\pm$ 0.04	**875.46 $\pm$ 40.53**	**931.00 $\pm$ 21.37**	**0.94 $\pm$ 0.05**
$20 \leq p_T(\gamma) < 35$	2748.03 $\pm$ 70.97	0.13 $\pm$ 0.01	1.30 $\pm$ 0.06	**460.87 $\pm$ 30.82**	**537.85 $\pm$ 39.72**	**0.86 $\pm$ 0.09**	3870.62 $\pm$ 80.86	0.11 $\pm$ 0.00	1.29 $\pm$ 0.05	**553.51 $\pm$ 32.58**	**600.28 $\pm$ 17.83**	**0.92 $\pm$ 0.06**
$35 \leq p_T(\gamma) < 50$	876.10 $\pm$ 37.35	0.10 $\pm$ 0.01	1.43 $\pm$ 0.13	**120.90 $\pm$ 15.98**	**111.00 $\pm$ 7.62**	**1.09 $\pm$ 0.16**	1200.72 $\pm$ 44.44	0.10 $\pm$ 0.01	1.48 $\pm$ 0.11	**170.84 $\pm$ 18.72**	**174.79 $\pm$ 9.15**	**0.98 $\pm$ 0.12**
$50 \leq p_T(\gamma) < 65$	356.08 $\pm$ 24.69	0.10 $\pm$ 0.02	1.67 $\pm$ 0.21	**59.47 $\pm$ 12.83**	**44.67 $\pm$ 4.12**	**1.33 $\pm$ 0.31**	521.23 $\pm$ 30.63	0.08 $\pm$ 0.01	1.48 $\pm$ 0.14	**64.29 $\pm$ 10.06**	**73.78 $\pm$ 4.93**	**0.87 $\pm$ 0.15**
$65 \leq p_T(\gamma) < 80$	215.57 $\pm$ 19.68	0.06 $\pm$ 0.01	1.60 $\pm$ 0.26	**20.05 $\pm$ 5.48**	**27.05 $\pm$ 3.48**	**0.74 $\pm$ 0.22**	239.74 $\pm$ 21.52	0.09 $\pm$ 0.01	1.50 $\pm$ 0.22	**32.35 $\pm$ 7.05**	**32.17 $\pm$ 3.36**	**1.01 $\pm$ 0.24**
$80 \leq p_T(\gamma) < 120$	206.61 $\pm$ 19.60	0.09 $\pm$ 0.01	2.16 $\pm$ 0.41	**39.55 $\pm$ 10.18**	**29.98 $\pm$ 4.73**	**1.32 $\pm$ 0.40**	287.11 $\pm$ 22.78	0.08 $\pm$ 0.01	1.86 $\pm$ 0.24	**43.41 $\pm$ 8.65**	**35.51 $\pm$ 3.43**	**1.22 $\pm$ 0.27**
$120 \leq p_T(\gamma) < 160$	72.79 $\pm$ 10.69	0.07 $\pm$ 0.02	2.30 $\pm$ 0.67	**11.61 $\pm$ 4.82**	**8.06 $\pm$ 1.75**	**1.44 $\pm$ 0.67**	71.86 $\pm$ 13.26	0.06 $\pm$ 0.02	2.62 $\pm$ 0.73	**12.12 $\pm$ 5.14**	**11.89 $\pm$ 2.67**	**1.02 $\pm$ 0.49**
$160 \leq p_T(\gamma) < 200$	26.79 $\pm$ 7.65	0.09 $\pm$ 0.03	1.41 $\pm$ 0.68	**3.44 $\pm$ 2.27**	**2.12 $\pm$ 0.77**	**1.62 $\pm$ 1.22**	34.09 $\pm$ 8.36	0.08 $\pm$ 0.03	1.09 $\pm$ 0.54	**3.15 $\pm$ 2.02**	**1.52 $\pm$ 0.61**	**2.08 $\pm$ 1.57**
$200 \leq p_T(\gamma) < 260$	18.78 $\pm$ 5.31	0.05 $\pm$ 0.03	3.87 $\pm$ 2.48	**3.75 $\pm$ 3.37**	**1.59 $\pm$ 0.71**	**2.35 $\pm$ 2.36**	12.18 $\pm$ 7.28	0.01 $\pm$ 0.01	2.48 $\pm$ 1.81	**0.30 $\pm$ 0.42**	**1.02 $\pm$ 0.54**	**0.29 $\pm$ 0.44**
$260 \leq p_T(\gamma) < 320$	9.37 $\pm$ 3.79	0.13 $\pm$ 0.10	0.01 $\pm$ 0.02	**0.02 $\pm$ 0.03**	**0.00 $\pm$ 0.01**	**3.90 $\pm$ 7.67**	4.51 $\pm$ 3.40	0.00 $\pm$ 0.00	0.02 $\pm$ 0.01	**0.00 $\pm$ 0.00**	**0.01 $\pm$ 0.01**	**0.03 $\pm$ 0.04**
$320 \leq p_T(\gamma)$	5.82 $\pm$ 3.17	0.01 $\pm$ 0.01	1.00 $\pm$ 0.00	**0.05 $\pm$ 0.06**	**0.17 $\pm$ 0.23**	**0.31 $\pm$ 0.55**	6.95 $\pm$ 3.80	0.01 $\pm$ 0.01	0.39 $\pm$ 0.53	**0.02 $\pm$ 0.04**	**0.03 $\pm$ 0.02**	**0.60 $\pm$ 1.14**

Table 3.29: Results of the hadronic and fake background estimation for bins in $p_T(\gamma)$, for both lepton flavors, for the $N_{jet} = 3$ and $N_{jet} \geq 4$ regions in 2018.

SR3 2018												
	e channel						**μ channel**					
Region	fakes high chgIso	$r_{\text{iso}}^{\text{HS}}$	κ_{MC}	**DD-fakes**	**MC-fakes**	**DD-Fakes/MC-Fakes**	fakes high chgIso	$r_{\text{iso}}^{\text{HS}}$	κ_{MC}	DD-fakes	MC-fakes	**DD-Fakes/MC-Fakes**
inclusive	7416.52 ± 107.45	0.11 ± 0.00	1.37 ± 0.05	**1095.09 ± 52.92**	**1039.90 ± 26.08**	**1.05 ± 0.06**	11571.58 ± 130.16	0.11 ± 0.00	1.37 ± 0.04	**1792.28 ± 73.56**	**1628.02 ± 35.27**	**1.10 ± 0.05**
$20 \leq p_T(\gamma) < 35$	4544.91 ± 80.25	0.12 ± 0.00	1.32 ± 0.07	**721.42 ± 47.61**	**626.17 ± 22.41**	**1.15 ± 0.09**	7118.97 ± 99.50	0.13 ± 0.00	1.33 ± 0.06	**1188.03 ± 65.29**	**979.22 ± 26.58**	**1.21 ± 0.07**
$35 \leq p_T(\gamma) < 50$	1482.65 ± 49.25	0.11 ± 0.01	1.32 ± 0.09	**208.26 ± 21.09**	**203.98 ± 8.86**	**1.02 ± 0.11**	2408.82 ± 58.14	0.10 ± 0.01	1.30 ± 0.09	**327.15 ± 30.99**	**329.13 ± 17.62**	**0.99 ± 0.11**
$50 \leq p_T(\gamma) < 65$	673.10 ± 32.79	0.11 ± 0.01	1.20 ± 0.11	**84.91 ± 12.36**	**92.59 ± 7.98**	**0.92 ± 0.16**	996.75 ± 41.71	0.10 ± 0.01	1.48 ± 0.11	**144.34 ± 16.53**	**154.07 ± 7.99**	**0.94 ± 0.12**
$65 \leq p_T(\gamma) < 80$	264.16 ± 25.26	0.08 ± 0.01	1.62 ± 0.19	**35.08 ± 7.82**	**52.81 ± 3.02**	**0.66 ± 0.15**	442.55 ± 26.66	0.10 ± 0.01	1.38 ± 0.11	**58.49 ± 8.19**	**72.46 ± 3.43**	**0.81 ± 0.12**
$80 \leq p_T(\gamma) < 120$	304.87 ± 25.02	0.08 ± 0.01	1.80 ± 0.22	**44.96 ± 8.55**	**47.69 ± 4.87**	**0.94 ± 0.20**	422.42 ± 27.37	0.09 ± 0.01	1.61 ± 0.29	**58.31 ± 12.52**	**72.71 ± 12.15**	**0.80 ± 0.22**
$120 \leq p_T(\gamma) < 160$	80.61 ± 12.64	0.05 ± 0.01	1.78 ± 0.33	**6.98 ± 2.47**	**12.32 ± 1.56**	**0.57 ± 0.21**	127.92 ± 14.19	0.09 ± 0.01	1.46 ± 0.24	**16.00 ± 4.15**	**14.59 ± 1.54**	**1.10 ± 0.31**
$160 \leq p_T(\gamma) < 200$	30.65 ± 8.57	0.02 ± 0.01	1.06 ± 0.39	**0.73 ± 0.55**	**2.51 ± 0.63**	**0.29 ± 0.23**	23.04 ± 12.57	0.09 ± 0.02	1.55 ± 0.50	**3.04 ± 2.09**	**3.48 ± 0.91**	**0.87 ± 0.64**
$200 \leq p_T(\gamma) < 260$	21.94 ± 5.96	0.05 ± 0.03	1.92 ± 1.02	**2.09 ± 1.73**	**1.30 ± 0.48**	**1.61 ± 1.46**	14.01 ± 8.62	0.03 ± 0.02	1.82 ± 0.78	**0.79 ± 0.76**	**1.61 ± 0.47**	**0.49 ± 0.50**
$260 \leq p_T(\gamma) < 320$	3.29 ± 7.73	0.09 ± 0.07	1.12 ± 0.84	**0.34 ± 0.89**	**0.22 ± 0.13**	**1.53 ± 4.07**	9.02 ± 3.98	0.00 ± 0.00	2.20 ± 2.13	**0.07 ± 0.10**	**0.71 ± 0.49**	**0.10 ± 0.16**
$320 \leq p_T(\gamma)$	10.34 ± 3.62	0.01 ± 0.01	6.84 ± 7.04	**0.59 ± 0.89**	**0.32 ± 0.21**	**1.85 ± 3.04**	8.07 ± 3.68	0.05 ± 0.05	0.51 ± 0.53	**0.19 ± 0.30**	**0.03 ± 0.02**	**5.59 ± 9.33**

SR4p 2018												
	e channel						**μ channel**					
Region	fakes high chgIso	$r_{\text{iso}}^{\text{HS}}$	κ_{MC}	**DD-fakes**	**MC-fakes**	**DD-Fakes/MC-Fakes**	fakes high chgIso	$r_{\text{iso}}^{\text{HS}}$	κ_{MC}	DD-fakes	MC-fakes	**DD-Fakes/MC-Fakes**
inclusive	6598.45 ± 103.08	0.10 ± 0.00	1.39 ± 0.04	**878.21 ± 40.45**	**917.64 ± 59.13**	**0.96 ± 0.08**	10210.80 ± 121.39	0.10 ± 0.00	1.41 ± 0.03	**1462.90 ± 51.40**	**1302.00 ± 20.19**	**1.12 ± 0.04**
$20 \leq p_T(\gamma) < 35$	4089.70 ± 77.22	0.11 ± 0.00	1.29 ± 0.05	**574.11 ± 33.90**	**585.67 ± 58.16**	**0.98 ± 0.11**	6208.93 ± 91.36	0.12 ± 0.00	1.32 ± 0.04	**952.99 ± 42.40**	**802.78 ± 16.52**	**1.19 ± 0.06**
$35 \leq p_T(\gamma) < 50$	1233.99 ± 46.10	0.09 ± 0.01	1.32 ± 0.08	**141.29 ± 14.87**	**160.13 ± 6.72**	**0.88 ± 0.10**	1991.98 ± 55.87	0.08 ± 0.01	1.36 ± 0.07	**225.55 ± 19.46**	**244.92 ± 8.69**	**0.92 ± 0.09**
$50 \leq p_T(\gamma) < 65$	522.50 ± 30.73	0.09 ± 0.01	1.68 ± 0.16	**80.42 ± 12.31**	**78.53 ± 5.32**	**1.02 ± 0.17**	875.83 ± 35.30	0.10 ± 0.01	1.60 ± 0.10	**138.98 ± 15.12**	**113.62 ± 4.62**	**1.22 ± 0.14**
$65 \leq p_T(\gamma) < 80$	240.22 ± 24.38	0.07 ± 0.01	1.34 ± 0.15	**22.82 ± 4.94**	**34.31 ± 2.68**	**0.67 ± 0.15**	415.61 ± 28.15	0.08 ± 0.01	1.58 ± 0.19	**52.35 ± 9.49**	**56.09 ± 3.16**	**0.93 ± 0.18**
$80 \leq p_T(\gamma) < 120$	326.66 ± 24.54	0.07 ± 0.01	1.75 ± 0.24	**38.19 ± 8.12**	**39.70 ± 4.23**	**0.96 ± 0.23**	489.34 ± 27.92	0.07 ± 0.01	1.85 ± 0.19	**65.75 ± 10.68**	**60.64 ± 4.83**	**1.08 ± 0.20**
$120 \leq p_T(\gamma) < 160$	96.68 ± 14.11	0.04 ± 0.02	2.12 ± 0.40	**8.81 ± 4.28**	**10.28 ± 1.37**	**0.86 ± 0.43**	127.31 ± 16.10	0.08 ± 0.01	2.17 ± 0.35	**21.83 ± 5.82**	**16.14 ± 1.74**	**1.35 ± 0.39**
$160 \leq p_T(\gamma) < 200$	47.92 ± 8.78	0.08 ± 0.02	2.06 ± 0.65	**7.50 ± 3.64**	**3.06 ± 0.75**	**2.45 ± 1.33**	42.22 ± 8.96	0.02 ± 0.01	1.67 ± 0.47	**1.74 ± 1.00**	**4.53 ± 0.90**	**0.38 ± 0.23**
$200 \leq p_T(\gamma) < 260$	24.73 ± 6.66	0.07 ± 0.03	3.07 ± 1.58	**5.65 ± 4.08**	**1.69 ± 0.55**	**3.34 ± 2.65**	35.02 ± 7.76	0.03 ± 0.02	1.58 ± 0.59	**1.53 ± 1.10**	**2.53 ± 0.66**	**0.60 ± 0.46**
$260 \leq p_T(\gamma) < 320$	6.81 ± 7.58	0.00 ± 0.00	0.63 ± 0.54	**0.01 ± 0.02**	**0.46 ± 0.31**	**0.03 ± 0.06**	18.82 ± 5.42	0.12 ± 0.07	1.64 ± 1.18	**3.81 ± 3.60**	**0.49 ± 0.21**	**7.72 ± 7.97**
$320 \leq p_T(\gamma)$	9.24 ± 4.33	0.01 ± 0.01	15.87 ± 17.61	**0.81 ± 1.29**	**3.82 ± 3.61**	**0.21 ± 0.39**	5.74 ± 4.67	0.00 ± 0.00	1.15 ± 1.34	**0.02 ± 0.04**	**0.26 ± 0.24**	**0.08 ± 0.16**

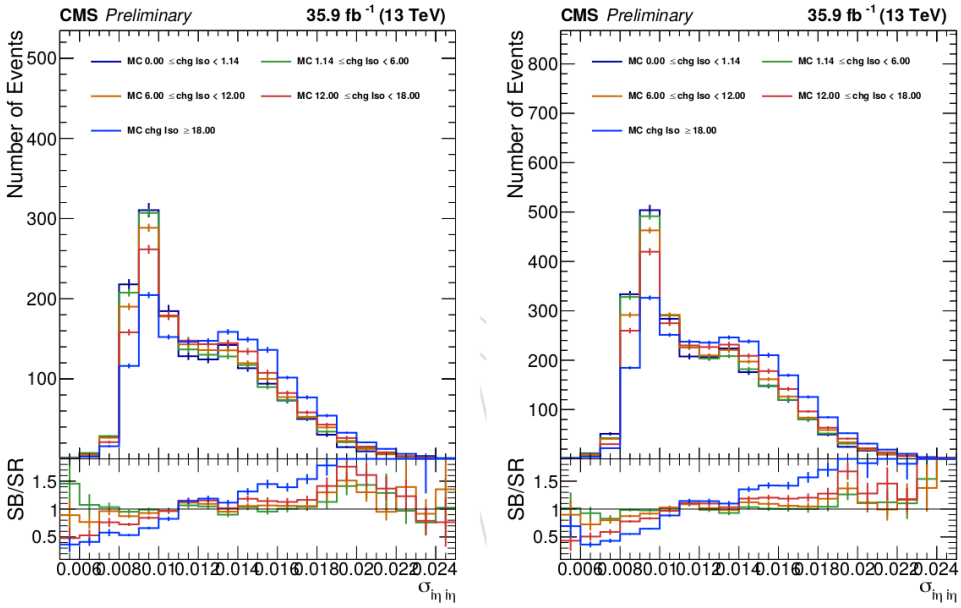


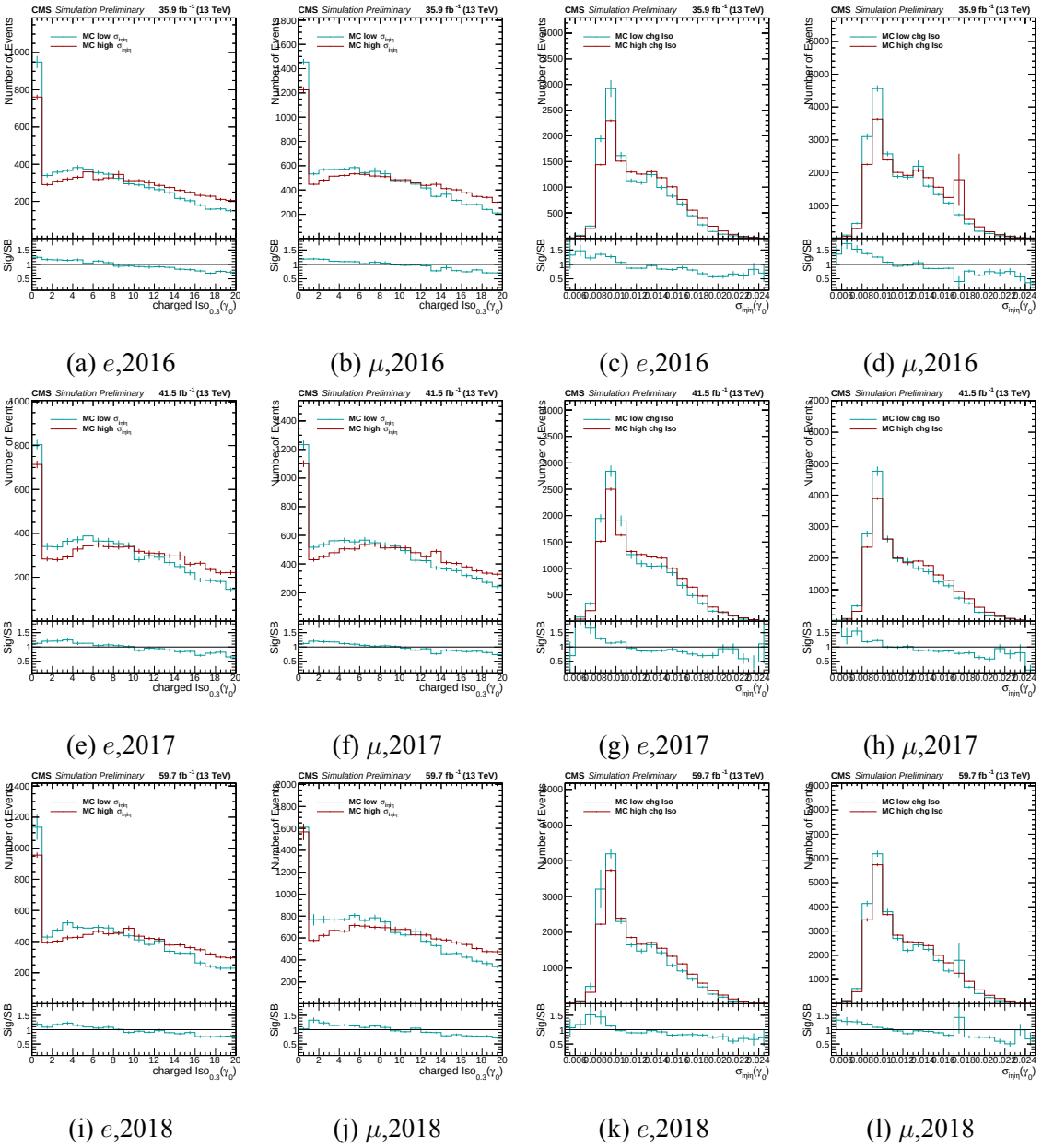
Figure 3.23: Shape comparison in 2016 simulation of the $\sigma_{i\eta i\eta}$ observable in the e-channel (right) and the μ-channel (left) for non-prompt γ in slices of mI_{chg}.

3.1.10 Systematic uncertainties

There are some theoretical uncertainties as well as some experimental uncertainties which affect the signal efficiency and the background normalization. Systematic uncertainties generally stem from various sources such as inaccurate modeling and simulation, uncertainties in the input parameters, miscalibration of the detector, uncertainties associated with corrections, and theoretical limitations. These uncertainties are correlated over the data-taking period. All such systematic uncertainties are summarized in the Table 3.30.

3.1.10.1 Experimental Uncertainties

- **Luminosity uncertainty:** For 2016, 2017, and 2018, the integrated luminosity is 35.92 fb^{-1}, 41.53 fb^{-1}, and 59.74 fb^{-1}respectively. The uncertainty on the integrated luminosity is estimated to be 1.2% for 2016, 2.3% for 2017 and 2.5% for 2018. The total uncertainty is split in different sources partially correlated across years following



(a) e,2016 (b) μ,2016 (c) e,2016 (d) μ,2016

(e) e,2017 (f) μ,2017 (g) e,2017 (h) μ,2017

(i) e,2018 (j) μ,2018 (k) e,2018 (l) μ,2018

Figure 3.24: Shape comparison of the charged isolation (left two rows) and the $\sigma_{i\eta i\eta}$ variable (right two rows) of the photon categories that form the non-prompt contribution (hadronic γ, fake γ and PU γ) in the SR4p region for the years 2016 (top), 2017 (middle) and 2018 (bottom).

the recommendations from the luminosity POG.

- **Pileup uncertainty:** An inelastic cross section of 69.2 mb [64] is utilized to determine the mean pileup for all three data-taking years. The uncertainty assigned to the number of pileup events in simulation is obtained by changing the inelastic pp cross section, which is used to estimate the pileup in data, within its uncertainty of $\pm$ 4.6%. The uncertainty of pileup is indicated by the difference between the nominal and reweighed distributions, which is entirely correlated over the years.

- **Trigger efficiency:** The uncertainties in the trigger efficiency are propagated to the final results. The trigger efficiency uncertainty arises from the scale factors used for matching the trigger selection efficiencies between simulation and observed data. As per the "tag-and-probe" measurement [65] [66] an uncertainty of up to 0.5% is stipulated to the simulated event yields.

- **Lepton efficiency uncertainty:** The uncertainties in the trigger and lepton identification and isolation efficiencies in simulation are estimated by varying data-to-simulation scale factors by their uncertainties. As suggested by the TOP EGM in Ref.[67] and MUOPOG in Ref. [68], the recommended uncertainties on the scale factors of electron and muon trigger, identification, and isolation are adopted. The identification and isolation uncertainties are considered separately.

- **Photon efficiency uncertainty:** It is expected that identification and pixel seed veto ID will have an impact on the scale factors for photon efficiency. The uncertainties associated with those variables are adopted in our study as in Ref. [67] and taken from EGM POG.

- **Jet energy scale and resolution:** Jet energy scale and resolution The uncertainty due to the limited knowledge of the jet energy sclae (JES) and jet energy resolution (JER)

is determined by varying the scale and resolution within the uncertainties in bins of p_T and η, typically by a few percent. The JES and JER uncertainties are propagated to MET. The JES uncertainty sources are separated in multiple components, that are correlated or uncorrelated across years following the recommendations of the JME POG [69] and using a reduced set of JEC sources. The JER uncertainty is uncorrelated across years.

- **b-tagging uncertainty:** The uncertainties that arise from b-tagging efficiency and misidentification rate are evaluated by adjusting the data-to-simulation scale factors of b and c jets, as well as light-flavor jets, within their respective uncertainties. These uncertainties vary with the p_T and η of the jet as recommended by BtagPOG Ref. [49].

- **L1 prefiring:** In the forward region of the endcap indicated by $|\eta|$ >2.4, a progressive shift in the input timing of the ECAL L1 trigger was noticed between the data collection periods of 2016 and 2017. This reduced the L1 trigger efficiency. This effect was corrected for the events with jets with $2.4 < |\eta| > 3.0$ and $p_T > 100$ GeV using an unbiased data sample. The measurement of p_T^{miss} observable may be impacted by the restriction on directly accessing any rebuilt item in the aforementioned η region. When this correction factor is systematically varied for the impacted objects by 20%, the uncertainty in the anticipated yields is discovered to be between 0.3 and 0.9 percent.

3.1.10.2 Theoretical

- **Color reconnection uncertainty:** The concept of color reconnection refers to the way in which the color flow between the partons is rearranged when they interact. It may result in significantly large uncertainties in simulated events. The primary model for the color reconnection utilized in simulations is based on the multi-parton inter-

action (MPI) model. Alternative models, such as the QCD-inspired (CR1) and gluon move (CR2) models, are also being used in several simulation. To estimate the uncertainties associated with color reconnection, the largest deviation between the nominal MPI model and the alternative models is taken as a measure of the uncertainty.

- **ISR/FSR uncertainty:** To address the uncertainty in initial(ISR) and final (FSR) state radiations, the parton shower scales employed in the simulation of the ISR and FSR, is varied up and down by a factor of two. These variations are motivated by the uncertainties in the parton shower tuning, and correlated through the years.

- **Parton distribution function uncertainty (PDF):** The uncertainty from the choice of PDFs is determined by reweighting the simulated $t\bar{t}\gamma$ events. The variations are correlated across years.

- **Matrix element (ME) scale uncertainty:** The uncertainty in the modeling of the hard-production processes is assessed by independently varying renormalization (μ_R) and factorization (μ_F) scales in the POWHEG sample by a factor of 2 and 0.5, which is included as a shape uncertainty in the fit. The value of μ_R and μ_F are varied to get a total of seven combinations, which are accounted for in the higher order correction and are included as a shape uncertainty in the fit.

3.1.10.3 Background

- The change of the TF with N_{jet} for various N_{b-jet} accounts for 50% of the uncertainty in the normalization of the QCD multijet component. For the contributions to the N_{b-jet} = 0 and $\geq$1 yields, independent uncertainties are taken into account. These result in an uncertainty of 0.9% in the measured inclusive cross section and have a meaningful effect only in the LM3, LM4p, HM3, and HM4p control zones.

- The $tW\gamma$ background component accounts for no more than 3.3 percent of total event yield in signal regions denoted as SR3 and SR4p and is predicted by the tW sample, which is simulated with POWHEG at next-to-leading order precision. This results in a 1.6 percent uncertainty in the inclusive cross section in fiducial phase space.

- Z+jets normalization uncertainty: An uncertainty of 5% is determined for the Z+jets process cross section in control regions with zero b-tagged jet and zero γ, while an uncertainty of 8% is applied for control or signal regions that have one or more b-tagged jets.

- The non-prompt photon prediction uncertainty is based on the modelling of the $I_{chg}(\gamma)$ distribution for various $\sigma_{i\eta i\eta}(\gamma)$ criteria, and it results in an uncertainty of 1.8% in the inclusive cross section.

- The normalisation of the $Z\gamma$($W\gamma$) background in the $N_{jet} \geq 4$ signal and control areas is given a 40 (20)% uncertainty.

- Up to 8% of the expected background yields are unknown due to the component with misidentified electrons, which has a 1.8% influence on the inclusive cross section.

- Additional uncertainties below 1% are caused by the $t/t\bar{t}$ normalization uncertainty of 5% and the uncertainties in the normalization of other tiny background components.

Table 3.30: List of different uncertainty sources with their statistical and systematic components. Sources are enlisted in the first column, where as second column represents their correlation over the three data-taking periods. Third and fourth column represent statistical uncertainty in the simulated event yields before the fit and uncertainty on final cross section measurement respectively

	Source	Correlation	Uncertainty [%] yield	$\sigma(t\bar{t}\gamma)$
Experimental	Integrated luminosity	partial	2.3-2.5	1.8
	Pileup	100%	0.5-2.0	<0.5
	Trigger efficiency	—	< 0.5	< 0.5
	Electron reconstruction and identification	100%	0.2-1.7	<0.5
	Muon reconstruction and identification	partial	0.5-0.7	<0.7
	Photon reconstruction and identification	100%	0.4–1.4	1.1
	p_T (e) and p_T (γ) reconstruction	100%	0.1–1.2	<0.5
	JES	partial	1.0–4.1	1.9
	JER	—	0.4–1.6	0.6
	b tagging	100% (2017/2018)	0.8–1.6	1.1
	L1 prefiring	100% (2016/2017)	0.3–0.9	<0.5
Theoretical	Tune	100%	0.1–1.9	< 0.5
	Color and reconnection	100%	0.4–3.6	<0.5
	ISR/FSR	100%	1.0–5.6	1.9
	PDF	100%	< 0.5	<0.5
	ME scales μ_R,μ_F	100%	0.4–4.7	< 0.5
Background	Multijet normalization	100%	1.3–6.5	0.9
	Non-prompt photon background	100%	1.2–2.7	1.8
	Misidentified e	—	2.5–8.0	1.8
	$Z\gamma$ normalization	100%	0.6–2.5	0.5
	$W\gamma$ normalization	100%	1.0–3.5	2.3
	DY normalization	100%	0.1–1.1	1.0
	$t/t\bar{t}$ normalization	100%	1.0–1.9	0.8
	$tW\gamma$ modeling	100%	1.6–4.4	1.6
	"Other" bkg. normalization	100%	0.3–1.0	<0.5
	Total systematic uncertainty			6.0
	Statistical uncertainty			0.9
	Total			6.0

3.1.11 Results

Finally, binned likelihood fit has been done with M_3 and $I_{chg}(\gamma)$ variable simultaneously to finally extract scale factor (r) on our theory of $t\bar{t} + \gamma$ cross section where r $= \frac{N^{obs}_{t\bar{t}+\gamma}}{N^{th}_{t\bar{t}+\gamma}}$. Figures 3.25 and 3.26 display the observed data as well as the anticipated signal and background event yields extracted from the likelihood fit to all signal and control regions. These values account for the correlation of the systematic uncertainty by summing the contributions from the three data collection periods. According to their postfit impact on the measured inclusive cross section, the major systematic uncertainties are ranked in Figure 3.27l impact.

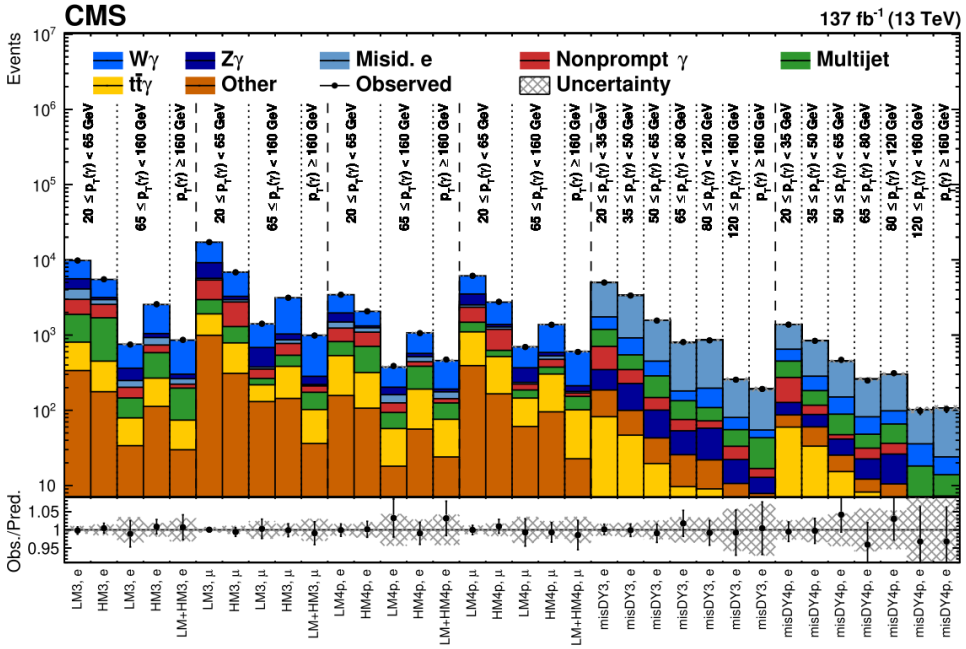


Figure 3.25: Event yield after fitting to data in the different control regions i.e, HM3, HM4p, LM3, LM4p, misDY3, and misDY4p. Postfit values have been used for the nuisance parameters. The ratio between the observed and predicted event yield has been shown in the lower panel. Hatched band represents the systematic uncertainties after the fit.

Within the fiducial phase space, the Combine inclusive cross section of $t\bar{t} + \gamma$ process

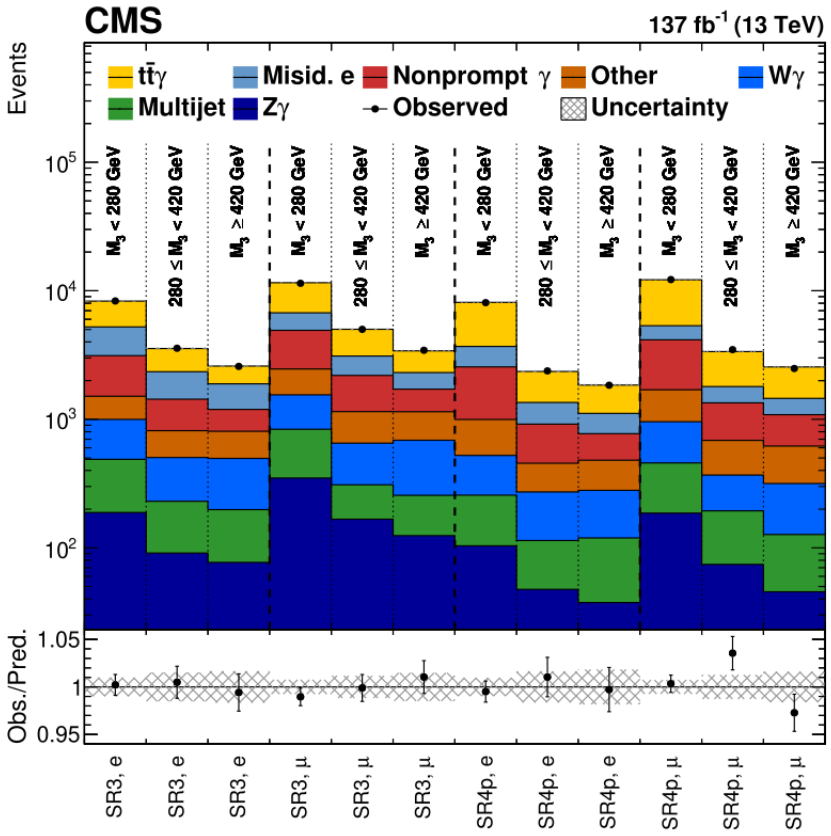


Figure 3.26: Event yield after fitting to data in the different signal regions i.e, SR3 and SR4p. Postfit values have been used for the nuisance parameters. The ratio between the observed and predicted event yield has been shown in the lower panel. Hatched band represents the systematic uncertainties after the fit.

in the $N_{jet} = 3$ and ≥ 4 selection is measured to be

$$\sigma_{t\bar{t}+\gamma} = 798 \pm 7(stat) \pm 48(syst)\text{fb} \tag{3.12}$$

is in line with the prediction made by the standard model i.e, $\sigma^{NLO}_{t\bar{t}+\gamma} = 773 \pm 135$ fb (see table 3.31). The signal normalization modifier is measured to be

$$r = 1.032 \pm 0.009(stat) \pm 0.062(syst). \tag{3.13}$$

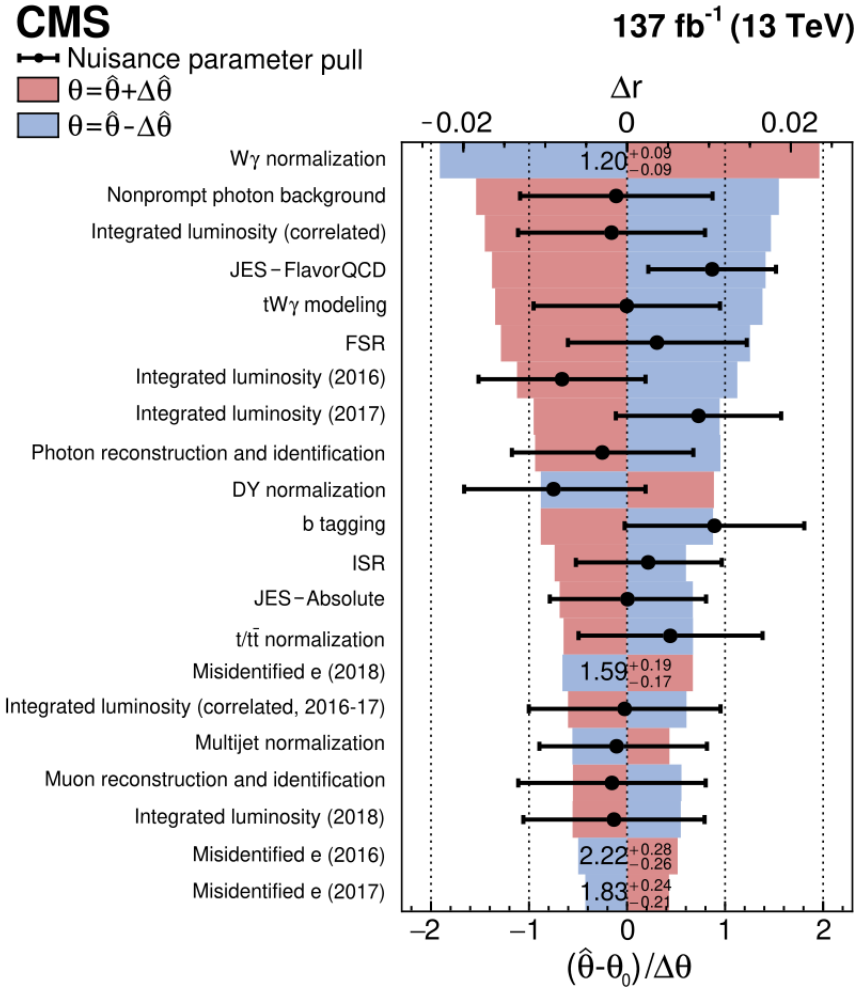


Figure 3.27: Ranking of the impact of the different systematic uncertainties on the inclusive cross section extraction by profile likelihood fit. Impacts are indicated by red and blue bands. And the postfit values of the nuisance parameters are indicated by black dots. Furthermore, the Wγ normalization scale factor and misidentified electron scale factor are shown in numerical values.

3.1.12 Summary and Outlook

Using data from proton-proton collisions at $\sqrt{s}$ = 13 TeV, which corresponds to a total integrated luminosity of 137 fb^{-1}, and collected with the CMS detector at the LHC, a precise cross section measurement for the production of top quark pairs in association with a photon

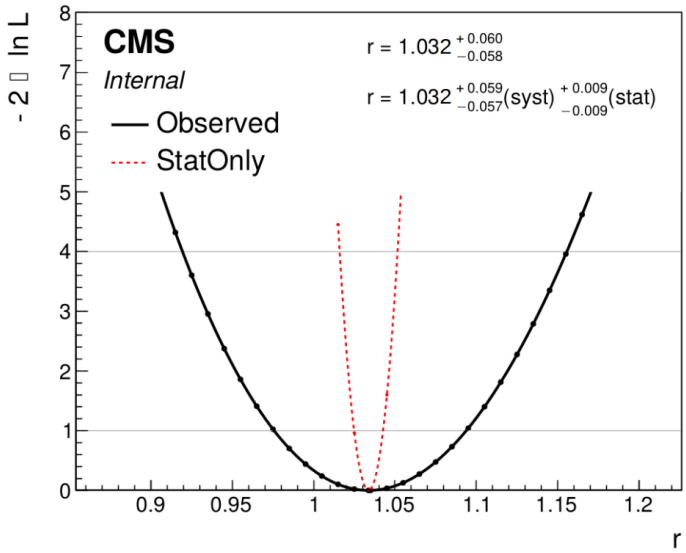


Figure 3.28: Negative log Likelihood value distribution. Minimum value corresponds to the measured signal strength.

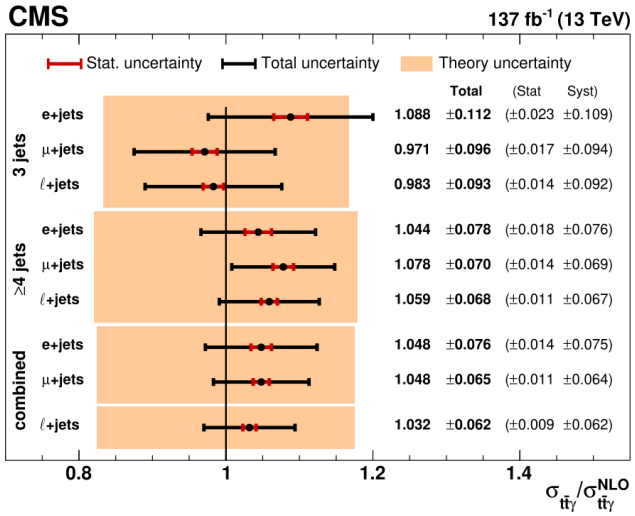


Figure 3.29: Uncertainty on signal strength parameter in combined N_{jet}=3 and $N_{jet} \geq 4$ region in full Run 2. Black dots are the ratio of measured and NLO predicted cross section of $t\bar{t}\gamma$ signal process in different signal regions defined by the different number of jets and lepton flavor selection. Whereas the black line represents total uncertainty, theoretical uncertainty in the NLO prediction is shown in the yellow band.

Table 3.31: Cutflow for the $t\bar{t} + \gamma$ signal in the 2016 period. Numbers are cross sections in fb. the fiducial requirements Cuts are applied sequentially, except for lines with indentation. Those requirements are dropped in the lines that follow. Line 09 gives the fiducial cross section for $N_{jet} \geq 3$.

	index	requirement	cross section (fb)			
			$t\bar{t}$+γ total	$t\bar{t}$+γ (0ℓ)	$t\bar{t}$+γ (1ℓ)	$t\bar{t}$+γ (2ℓ)
fiducial phase space	00	total	15891.64	6162.10	7509.17	2220.37
	01	overlap removal	15263.92	5960.83	7191.65	2111.44
	02	$N_{\gamma}^{\text{fid.}} = 1, p_T(\gamma) > 20\ GeV$	5863.40	2461.30	2678.02	724.08
	03	$N_{\ell}^{\text{fid.}} == 1$	1225.14	0.45	893.07	331.62
	04	$N_{e}^{\text{fid.}} = 1, N_{\mu}^{\text{fid.}} = 0$	569.09	0.31	414.72	154.06
	05	$N_{e}^{\text{fid.}} = 0, N_{\mu}^{\text{fid.}} = 1$	656.07	0.14	478.36	177.57
	06	$N_{jet}^{\text{fid.}} \geq 1$	1211.02	0.45	888.83	321.74
	07	$N_{jet}^{\text{fid.}} \geq 2$	1098.57	0.43	841.25	256.90
	08	$N_{jet}^{\text{fid.}} \geq 3$	792.33	0.40	659.60	132.34
	09	$N_{b-jet}^{\text{fid.}} \geq 1$	773.35	0.38	642.02	130.96
	10	$N_{jet}^{\text{fid.}} \geq 4$	414.54	0.32	366.20	48.02

has been presented in this chapter. It is the first outcome of measurements made utilizing data from 13 TeV for the $t\bar{t}\gamma$ final state by the CMS Collaboration. Events with precisely three and four or more jets, among which at least one is b-tagged, have been used in the analysis in the single-lepton channel. Based on the data, estimates are made for background components with misidentified electrons, photons produced during the hadronization of jets, the multijet component, and prompt photons from the $W\gamma$ and $Z\gamma$ processes. The measured value $798 \pm 7\ (stat) \pm 48\ (syst)$ fb for the inclusive cross section is in line with standard model NLO prediction.

3.2 Charge Asymmetry Measurement

3.2.1 Charge asymmetry in top-pair production

In the context of top quark pair production, the charge asymmetry refers to an anisotropy in the directions of the quark and antiquark in the final state with respect to the incoming partons. Let us consider the process $q\bar{q} \rightarrow Q\overline{Q}$, where $Q\overline{Q}$ refers to a pair of heavy quarks, such as top or bottom quarks, while q represents the incoming quarks from the protons. We limit our discussions here to the top quarks i.e., Q refers to top quark only. As illustrated in Fig. 3.30, if Q is produced at an angle θ with respect to q, the differential charge asymmetry can be defined as

$$A_C^{Q\bar{Q}}(cos\theta) = \frac{N_Q(cos\theta) - N_{\bar{Q}}(cos\theta)}{N_Q(cos\theta) + N_{\bar{Q}}(cos\theta)}, \tag{3.14}$$

with

$$N_Q(cos\theta) = \frac{d\sigma_{Q\bar{Q}}}{d\Omega(cos\theta)} \tag{3.15}$$

and

$$N_{\bar{Q}}(cos\theta) = \frac{d\sigma_{\bar{Q}Q}}{d\Omega(cos\theta)}. \tag{3.16}$$

Here, $N_Q(cos\theta)$ denotes the production rate of $Q\overline{Q}$, where Q is produced within an angular interval $(\theta, \theta + d\theta)$, and $N_{\overline{Q}}(cos\theta)$ represents the production rate when directions of Q and $\overline{Q}$ are swapped in the final state.

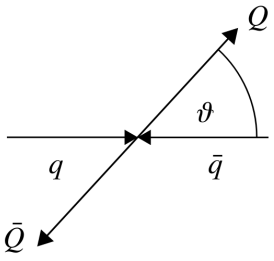


Figure 3.30: Illustration of production angle θ in $q\bar{q} \rightarrow Q\overline{Q}$ process in the $q\bar{q}$ pair rest frame

In the standard model no charge asymmetry is expected at LO in top-pair production. But $q\bar{q} \rightarrow t\bar{t}$ (quark-antiquark annihilation) has small asymmetry due to interference of higher order feynman diagrams as shown in Figure 3.31. Since the charge asymmetry requires charge-asymmetric initial state, there is no asymmetry in the dominant process $gg \rightarrow t\bar{t}$. The combined measurement by the ATALS and CMS experiment at $\sqrt{s}$ =8 TeV observed to be 0.5% [70].

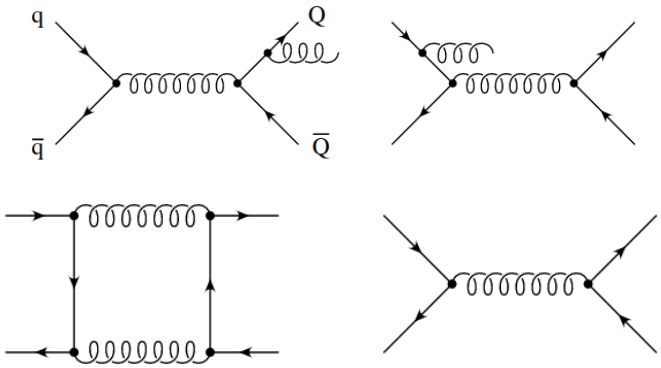


Figure 3.31: Higher order Feynman diagrams of top pair production, which leads to charge asymmetry. Above two diagrams show the gluon radiation from final (left) and initial quark (right).

3.2.2 Charge asymmetry in $t\bar{t} + \gamma$ events

In comparison with $t\bar{t}$, the $t\bar{t} + \gamma$ process consists of an extra photon vertex and the same can enhance charge asymmetry. Through the requirement of this photon, the selection of events with ISR photons increases which inherently selects the $q\bar{q} \rightarrow t\bar{t} + \gamma$ production. The other possibility is that the photon couples differently with u (up quark) and d (down quark), resulting to a difference between $u\bar{u} \rightarrow t\bar{t} + \gamma$ and $d\bar{d} \rightarrow t\bar{t} + \gamma$ cross section. Thus it deduce the possibility of cancellation of opposite contribution to charge asymmetry from these two process. Genuine tree level asymmetry in $q\bar{q} \rightarrow t\bar{t} + \gamma$ process is about 12% [71]. Various sensitive variables like $\Delta|\eta| = |\eta_t| - |\eta_{\bar{t}}|$, $\Delta|y| = |y_t| - |y_{\bar{t}}|$ are considered for charge

asymmetry calculations. For $t\bar{t}+\gamma$ asymmetry can be defined as $A_C = \frac{N^+ - N^-}{N^+ + N^-}$ for $|y_\gamma| > 1$, where N^+, N^- are the numbers of events with positive and negative values for $\Delta|\eta|$ or $\Delta|y|$ as appropriate.

3.2.3 Analysis strategy

For charge asymmetry measurement same dataset and the MC modeling have been used as in the cross section analysis discussed in the previous section. Also, the same procedures have been repeated for the following steps: 1. Overlap removal between samples 2. DY+jets background scale factor estimation 3. Misidentified electron, $W+\gamma$, $Z+\gamma$ scale factor estimation 4. Multijet background estimation. Here the reconstruction of the top and anti-top quark from their decay products is essential to estimate the charge asymmetry values at the reconstruction level.

3.2.4 Top quark reconstruction

Depending on the W boson decay mode, the top quark can be categorized into two types; one is leptonic, and another one is hadronic. The leptonic top can be reconstructed from lepton, neutrino, and b-jet. The hadronic top can be reconstructed from three jets (one is b-tagged). **Step 1:** Firstly, neutrino four momenta are reconstructed. Transverse components

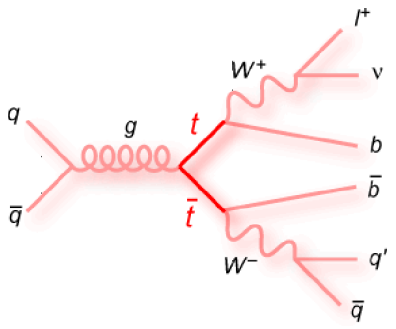


Figure 3.32: Illustration of final state particles available for reconstruction of the top quarks

are taken from p_T^{miss}. Z component ($P_{z,\nu}$) is solved from four momenta conservation at $W\ell\nu$

vertex.

$$P^{\pm}_{z,\nu} = \frac{\Lambda P_{z,\ell}}{E^2_\ell} \pm \sqrt{\frac{(\Lambda P_{z,\ell})^2}{(E^2_\ell - P^2_{z,\ell})^2} - \frac{E^2_\ell P^2_{T,\ell} - \Lambda^2}{E^2_\ell - P^2_{z,\ell}}} \tag{3.17}$$

$$where \Lambda = \frac{M^2_W}{2} + cos(\Delta\phi) P_{T,\ell} P_{z,\ell} + P_{z,\nu} P_{z,\nu} \tag{3.18}$$

Step 2: Now one need to find proper combination of jets and b-jets and neutrino $P_{z,\nu}$ solutions.χ^2 value is calculated by constraining W and Top mass. Minimum χ^2 value gives the correct combination.

$$\chi^2 = \frac{(M^{obs}_{th} - M_t)^2}{(\sigma^2_{j1} + \sigma^2_{j2} + \sigma^2_{bh})} + \frac{(M^{obs}_{t\ell} - M_t)^2}{(\sigma^2_\ell + \sigma^2_\nu + \sigma^2_{bl})} + \frac{(M^{obs}_{Wh} - M_W)^2}{(\sigma^2_{j1} + \sigma^2_{j2})} \tag{3.19}$$

The first term corresponds to the hadronic top, the middle term corresponds to the leptonic top, and the last term corresponds to the hadronic W boson.

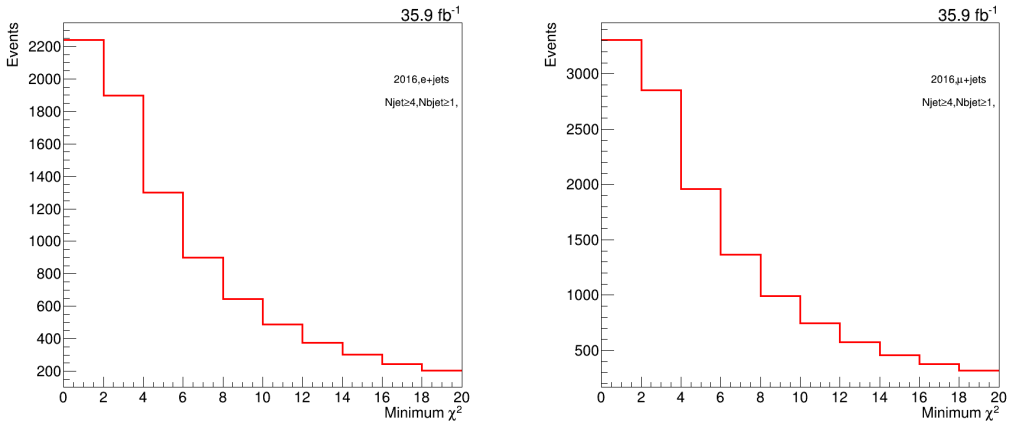


Figure 3.33: Minimum χ^2 value in 2016 signal sample in $N_{jet} \geq 4$ and $N_{b-jet} \geq 1$ selection in e+jets channel (left) and μ+jets (right)

3.2.5 Top quark kinematic variables

Depending of the decay mode of the W boson originating from the top quarks the events are categorized as hadronic and leptonic events. The mass distributions of the reconstructed top quark with hadronic and leptonic decays are shown in Fig. 3.34 (Fig. 3.35) for the e+jets

(μ+jets) channel. As mentioned in Sec. 3.2.2, the charge asymmetry in top pair production can be measured using several kinematic variables such as rapidity/pseudorapidity of the top/anti-top quarks, etc. The distribution of top and anti-top quark rapidity and their differences have been shown in Fig.3.36 (Fig.3.37) for e+jets (μ+jets) channel.

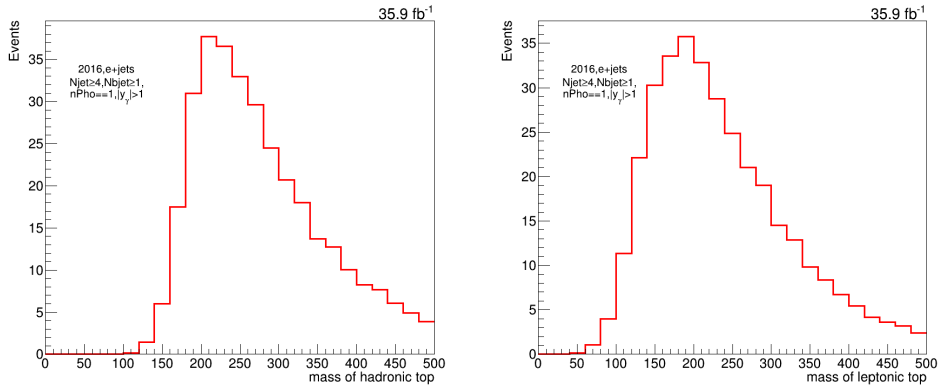


Figure 3.34: Reconstructed top mass in e+jets channel, hadronic top (left) leptonic top (right)

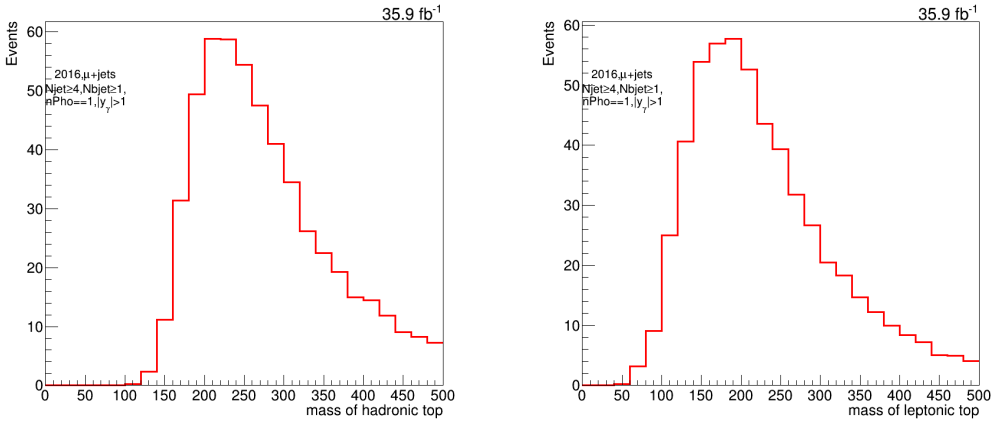


Figure 3.35: Reconstructed top mass in μ+jets channel, hadronic top (left) leptonic top (right)

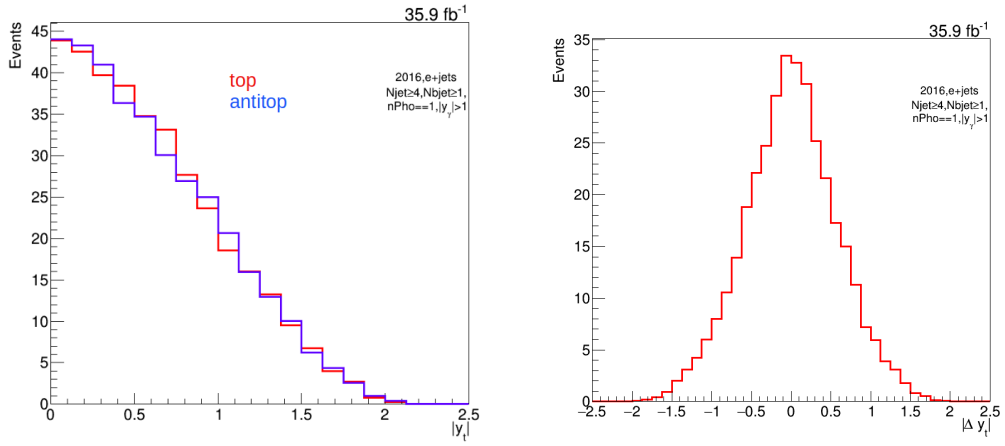


Figure 3.36: Rapidity of top and anti-top quark (left), distribution of the rapidity difference between top and anti-top quark in e+jets channel.

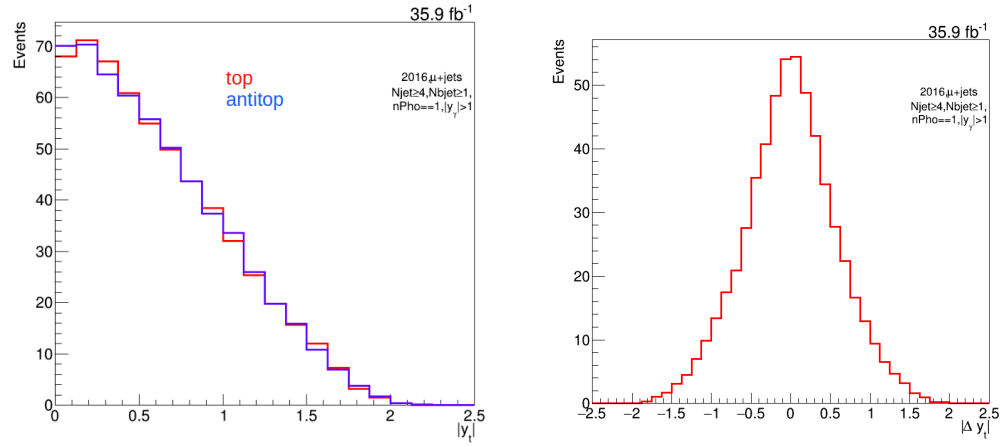


Figure 3.37: Rapidity of top and anti-top quarks (left), distribution of the rapidity difference between top and anti-top quark in μ channel.

3.2.6 Charge asymmetry results

After the top reconstruction, charge asymmetry has been calculated at the reconstruction level. The measured value of inclusive charge asymmetry (A_C) at the reconstruction level corresponds to -0.003$\pm$0.034 (stat), with the requirement of a photon in high rapidity ($|y_\gamma| > 1$) region. The differential distribution of the charge asymmetry in $t\bar{t}\gamma$ Monte Carlo simulation

has been shown in the Fig. 3.38. However, the unfolding of reconstruction level distribution into particle and parton level distributions are not shown here.

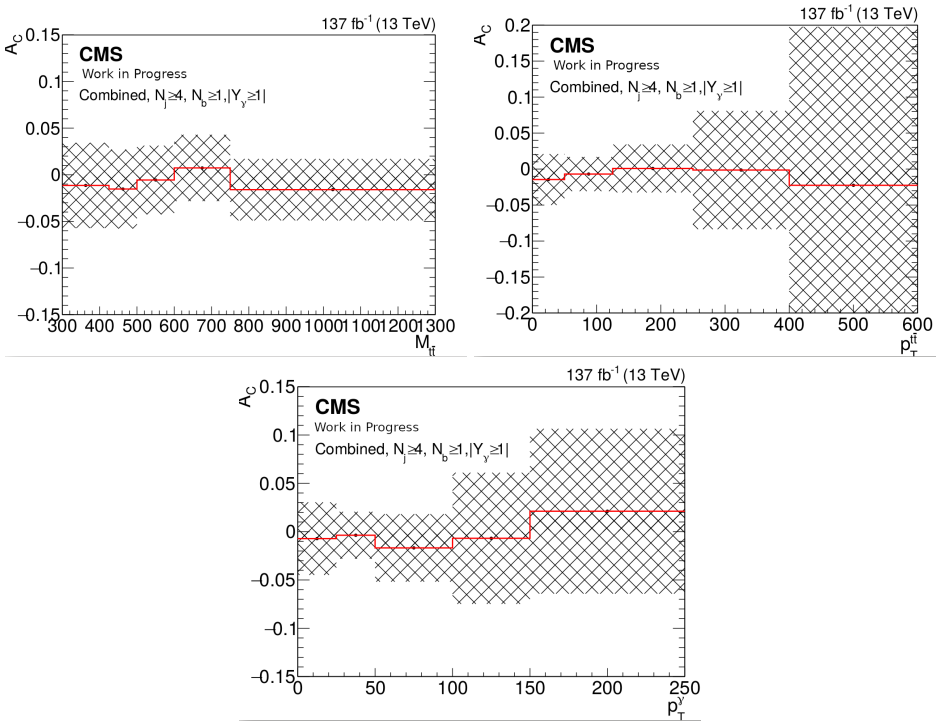


Figure 3.38: The differential distributions of charge asymmetry in $t\bar{t}\gamma$ MC events as function of $M_{t\bar{t}}$ (upper left),$p_T^{t\bar{t}}$ (upper right) and p_T^{γ} (bottom)

Chapter 4

CMS Trigger Studies

4.1 CMS Trigger System

In hadron colliders, the Trigger and Data Acquisition system plays a crucial role in reducing the data size to a manageable volume. Starting from the LHC Run 2, proton-proton collisions have been occurring at a frequency of 40 MHz, while at the nominal instantaneous luminosity of 10^{34} cm^{-2}sec^{-1} the proton-proton bunch-crossings lead to 20 (on average) inelastic collisions with approximately 1 MB event size. Clearly, such volumes of the dataset are beyond the manageable storage capability of any collider experiment and are of little interest. Moreover, a vast range of cross sections for a variety of physics processes[1] demand efficient designing of the trigger system without losing the sensitivity for the new physics searches. As illustrated in Fig. 4.1, a limited subset of all the proton-proton collisions[2] consisting of the events of interest are filtered through the trigger system and are stored for offline analysis.

In the CMS trigger system, the data reduction and event filtering process from the input collision rate of 10^9 Hz to a rate of 10^3 Hz has been implemented through two stages, and a schematic diagram of the CMS trigger system is shown in Fig. 4.2.

- Level-1 (L1) triggers are actually based on custom electronics using very coarsely segmented data from calorimeter and muon detectors. The decision-making proce-

[1] The dominant QCD processes have a cross section of O (10 μb-10 mb), relatively less abundant electroweak physics processes, and the rare BSM processes e.g., scalar Leptoquarks (LQ) have respective cross sections of O (10 nb-pb), and O($\leq$ 10 fb).

[2] Typically at a rate of 100-1000 kHz rate for CMS.

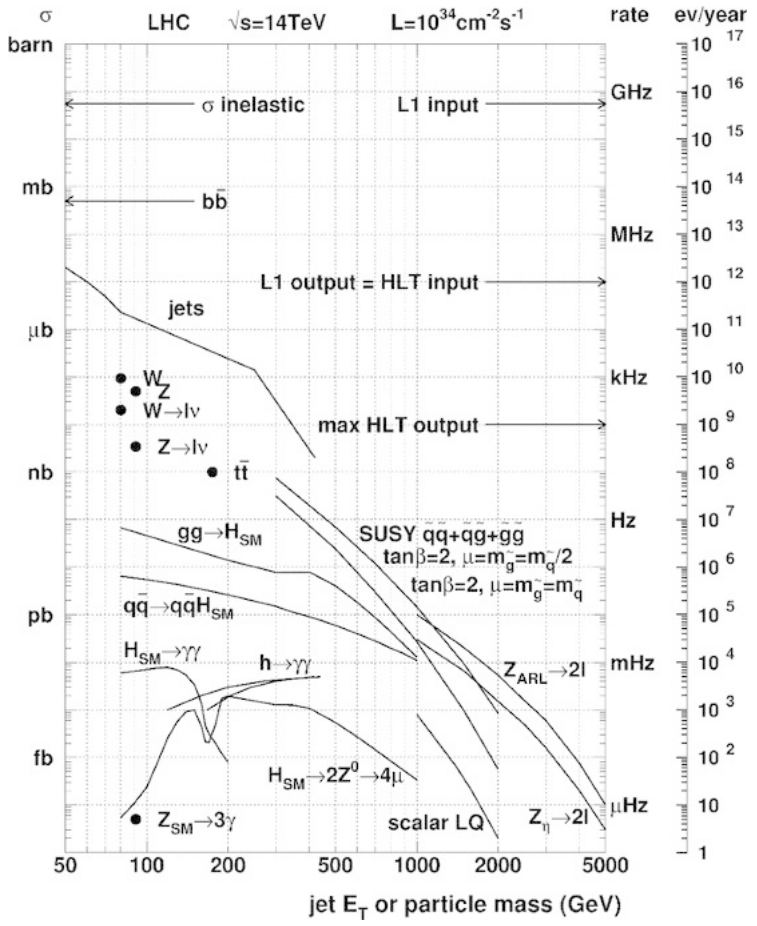


Figure 4.1: Inclusive proton-proton cross sections at the LHC for different physics processes. The typical interaction rates, input, and output rates of the CMS trigger system at a nominal luminosity of $10^{34}\ cm^{-2}sec^{-1}$ are also shown [72].

dure within the L1 is organized through three subsystems: the L1 calorimeter trigger, the L1 muon trigger, and the L1 global trigger. Here, the regional calorimeter triggers are initiated for the electron, photon, tau, and jet candidates based on ECAL, HCAL, and HF signals; in addition, the total transverse energy in an event and the total missing energy vector are also estimated by the L1 calorimeter trigger. The muon triggers, on the other hand, are initiated through the signals from three muon detector subsystems: DT in the barrel, the CSC in the endcap, and the RPCs in both barrel and endcap. The candidate muon tracks from individual subsystems are then sorted for

correlation (e.g., between DT/CSC and RPC tracks) and are validated for the muon sign. The L1 muon trigger also attempts to correlate the muon track candidates for any calorimeter-based isolation. Overall, the CMS L1 trigger latency is limited to 3.2 μs (equivalent to 128 bunch crossings at 25 ns interval) due to finite signal propagation time from the detector front ends to the underground cavern housing the trigger electronics. Within this span of time, the L1 global trigger supervises the data collection from the detector front ends, L1 decision making to discard a large fraction of uninteresting events, and transmission of L1 decisions to pass the selected events to the HLT.

- The CMS High level triggers (HLT) system has been implemented through a single processor farm (referred to as the "Event Filter Farm") which connects the subdetector readout elements through a high performance readout network; the event management system controls the flow of the dataset. It is a software based filtering system that uses a lightweight version of the CMS offline reconstruction software and checks for more precise object and/or event kinematics. In addition, having full access to the information used in L1, at HLT the events are scrutinized for further combinations and other topological calculations on the digital list of L1 objects. Due to its intricate selection algorithms related to the event reconstruction refinement and physics sophistication, the CMS HLT system requires a total processing time of up to $\sim$1 s, with a typical event output rate of 1 kHz.

4.2 L1 Bandwidth for Different Physics Objects

As mentioned earlier, the interaction rates are directly proportional to the instantaneous luminosity and collision energy. In addition, the CMS trigger performance depends on several other beam parameters – the number of colliding bunches and even the structure of the fill-

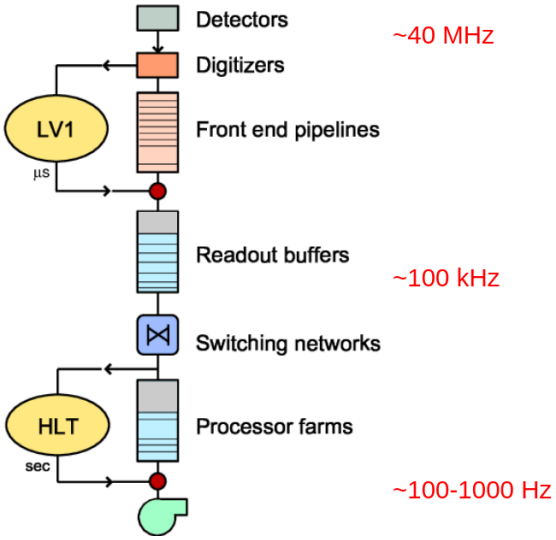


Figure 4.2: Data flow through the CMS Trigger/DAQ system: the HLT Triggers provide all filtering after the Level-1 decision is made. [73]

ing scheme as they introduce different dead times for the detector electronics. During Run 1 operations, the LHC collision energies had been gradually increased from 7 TeV (2011) to 8 TeV (2012), while the peak instantaneous luminosity was always below $1.0 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$; the nominal bunch crossing time was 50 ns. The LS1 allowed heavy upgradation of the LHC luminosity through an enhanced number of circulating proton bunches and the reduced interval between nominal bunch crossing time. In addition, the preventive measures against plausible helium leaks and quenching and the electron cloud mitigation allowed the proton-proton collisions to occur at $\sqrt{s}$=13 TeV during LHC Run 2 (2015-18); the peak luminosity even surpassed the design luminosity of $1.0 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$ leading to an average > 30 pileup events per bunch crossing. The peak instantaneous luminosities as recorded by the CMS during LHC Run 1-3 are shown in the Fig. 4.3.

To safeguard against such harsh operational conditions and to ensure efficient performance of the CMS trigger and data acquisition, the CMS L1 trigger system has been up-

graded heavily during LS1. Therefore, it had been quite crucial to cross-check the L1 performance using the Run 2 dataset and remeasure the bandwidth usage by different classes of L1 trigger objects. As detailed in the subsections below, this thesis work also consists of the development of Graphic User Interface (GUI) based monitoring tools. Few snapshots based on those tools, along with the rate summary, have been shown in the Run 2 L1 trigger performance publication[74].

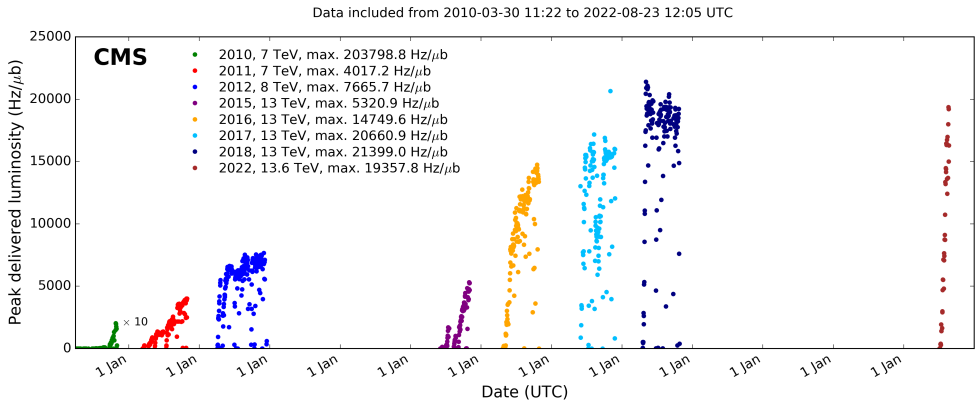


Figure 4.3: The peak instantaneous luminosity delivered by the LHC during all operations of the LHC so far; note that luminosities for the Run 3 (started in 2022) is also displayed [75].

4.2.1 L1 Trigger Logics or Seeds

The peak instantaneous luminosity delivered by the LHC during all operations of the LHC so far; note that luminosities for the Run 3 (started in 2022) is also displayed. At L1, the sums of calorimeter cell energies from the ECAL, HCAL, and HF form the trigger tower energy using the Trigger Primitive Generator (TPG) circuits. From the TPG, the information for candidate towers is transmitted to the Regional Calorimeter Trigger (RCT), which looks for the electron, photon, taus, and jet candidates. These candidates, along with the sums of transverse energy, are passed to the Global Calorimeter Trigger (GCT). The

GCT performs object sorting and calculates the missing transverse energy before passing all the information to the L1 Global Trigger. Similarly, predefined trigger logic set for individual muon tracks in the RPC, DT, and CSC. The Global Muon Trigger (GMT) sorts all the tracks and correlates the CSC and DT tracks with the RPC ones. It also determines whether the muon candidates are isolated based on the quiet calorimeter towers in the $\eta - \phi$ grid as obtained from the RCT.

Overall, such object selections are executed through one or more predefined trigger logics, called "seeds" loaded into the trigger electronics. Such seeds can be defined for selecting single or multiple objects of different types, e.g., jets, electron/photon, taus, muons, scalar and vector sum of transverse energy, etc., and with varying criteria (energy thresholds and detector acceptances). The L1 "Trigger Menu" is an optimized union of all such "trigger seeds" serving the broad physics objectives of the CMS. For individual seeds, a crucial adjustment for the object thresholds and/or acceptance is needed based on the instantaneous luminosity and other operational conditions so that the overall L1 output rate is within the bandwidth of 100 kHz. For Some of the L1 seeds (as listed in Tab. 4.1 along with their kinematic criteria), the triggering rate as a function of the pileup events (directly proportional to the instantaneous luminosity) is shown in Fig. 4.4. In order to improve the physics acceptance for certain low-p_T physics and sometimes for detector performance and rate studies, the individual L1 seed thresholds/acceptances are sometimes set quite low/wide; however, in such cases, the seeds are "prescaled"[3] appropriately so that the overall L1 rate remains within the L1 bandwidth.

[3] A mechanism/logic set into the hardware or through software so that only a certain type of events passing the trigger criteria are selected only at the required frequency. For example, a prescale of 50 for an arbitrary seed "A" implies that once in every 50 occurrences, the events passing requirements of "A" would be passed/selected by the trigger system.

Table 4.1: Detailed Criteria of L1 trigger seeds used in figure 4.4

L1 Seed	Criteria
L1_SingleLooseIsoEG28er2p5	Single loosely isolated e/γ with $E_T > 28$ GeV and $\|\eta\| < 2.5$
L1_DoubleIsoTau32er2p1	Double isolated τ with $E_T > 32$ GeV and $\|\eta\| < 2.1$
L1_SingleMu22	Single muon with $p_T > 22$ GeV
L1_DoubleEG_25_12_er2p5	Double e/γ with $E_T > 25$ GeV, 12 GeV and $\|\eta\| < 2.5$
L1_DoubleMu_15_7	Double muon with $p_T > 15$ GeV, 7 GeV
L1_ETMHF100	$E_T^{miss} > 100$ GeV
L1_SingleJet180	Single jet with $E_T > 180$ GeV
L1_DoubleJet150er2p5	Double jet with $E_T > 150$ GeV and $\|\eta\| < 2.5$

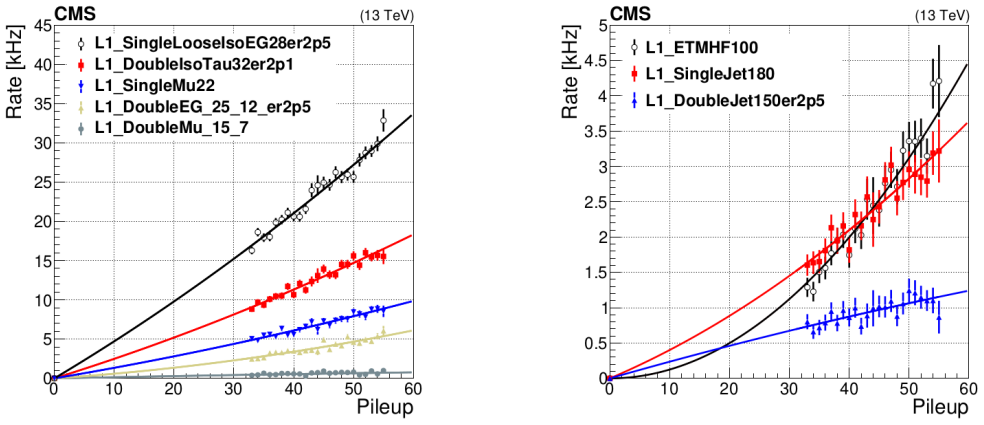


Figure 4.4: L1 firing rates for some of the seeds (listed in Tab. 4.1) as functions of number of pileup events.

4.2.2 L1 Trigger Rates

During the LS1, the L1 trigger system had been upgraded with newer FPGAs and ASICs, along with faster interconnections [76]. With the onset of Run 2, the trigger thresholds for different physics objects, e.g., jets, electrons/photons, muons, tau leptons, etc., were gradually adjusted to accommodate data-flow within the available bandwidth. Therefore, it had been quite mandatory to cross-check bandwidth usage by different L1 objects routinely to ensure non-detrimental impact on the physics acceptance of the CMS. The trigger rates and fraction usages are routinely monitored using a C++-based GUI, and the same has been

shown in Figs. 4.5-4.7, for three typical runs during LHC Run 2. The datasets used for the trigger rate estimation are based on an unbiased L1 trigger that only requires the reference for a proton-proton bunch-crossing, called zero-bias trigger. The trigger rate tables containing the name and the rate of triggers seeds, are used as input to the above-mentioned GUI. The GUI categorizes the seeds for different physics objects by checking their names and produce a pie chart for visualization of the rates consisting of those physics objects. Based on these rate fractions presented in the pie chart, over the different data-taking periods, the L1 triggering thresholds are adjusted to maintain the total L1 output trigger rate at a level below 100 kHz.

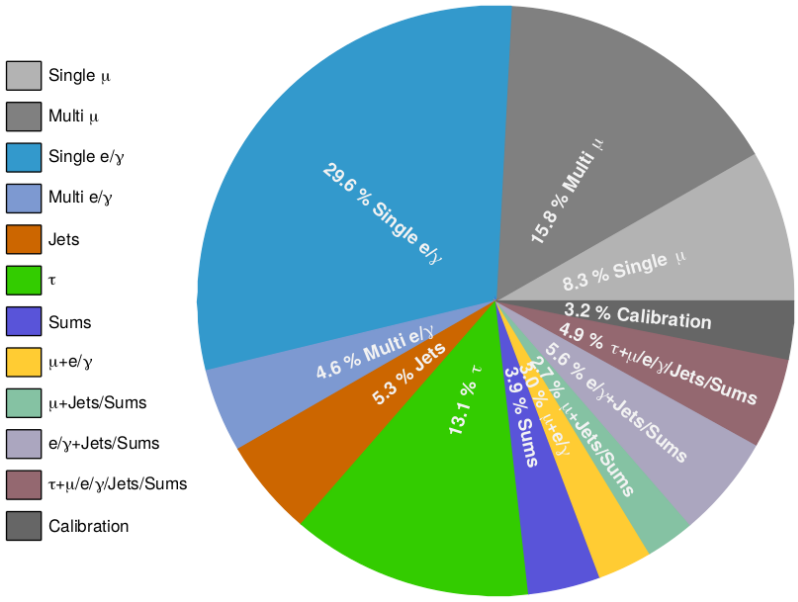


Figure 4.5: Rate fractions shared by different physics object at L1 for run 323940.

4.3 Efficiency for $e\gamma$ HLT Conditions

As mentioned earlier, at the HLT, complete electronic readout data from all the subdetectors, along with the L1 trigger objects information, are available for event reconstruction. Indi-

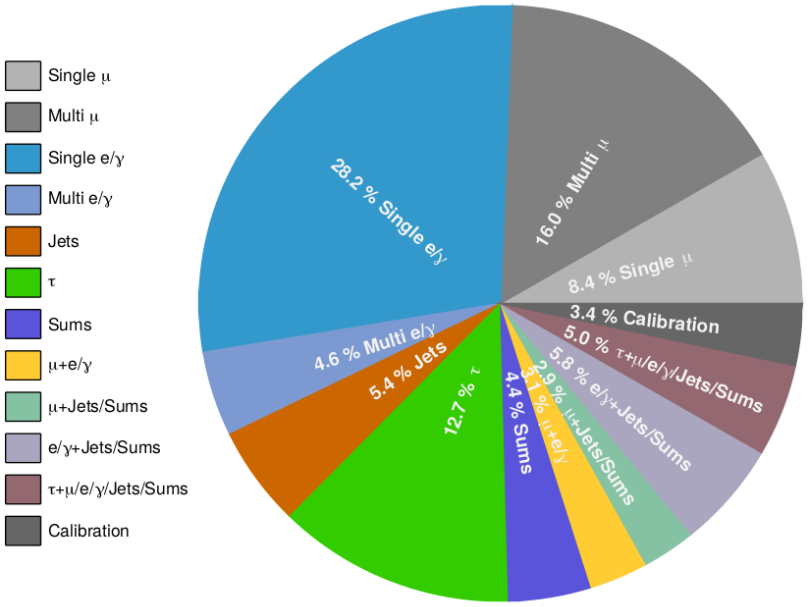


Figure 4.6: Rate fractions shared by different physics objects at L1 for run 324077.

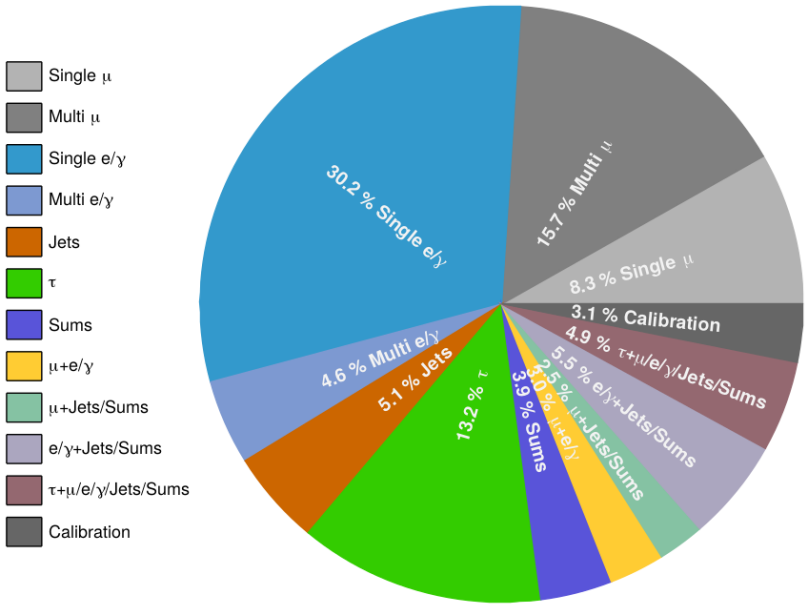


Figure 4.7: Rate fractions shared by different physics object at L1 for run 325022.

vidual trigger candidate objects like electrons, photons, muons, jets, and finally, the missing transverse energy (MET) are defined through sequential algorithms combining multiple subdetector's information. Such sequential algorithms to build the HLT candidate objects are known as HLT "filters". Finally, one or more HLT filters for different candidate objects are combined together into an HLT "path" to target specific event signature(s). Similar to L1, the HLT in CMS consists of an extensive menu list with 600 different paths to target a broad range of physics signatures, respecting the overall output bandwidth of 100-1000 Hz.

At the HLT, an electron candidate reconstruction starts from the ECAL super-cluster [4] in a region specified by the L1 Trigger. The super-cluster is first required to fall within the precision physics fiducial region of the ECAL ($|\eta| < 2.5$ but excluding the region of $1.4442 \leq |\eta| \leq 1.5660$). The reconstructed super-cluster in the calorimeter is then checked for consistency with the hits in the pixel detector. The candidate electron is then looked for a matching track seeded by the pixel hits consistent with the original electron candidate super-cluster. Further checks between η coordinates of the super-cluster and the extrapolated track, H/E (i.e., the super-cluster energy ratio in HCAL to ECAL), etc. are performed to ensure good quality electron reconstruction at the HLT. All of these criterion are checked with sequential filters of the corresponding HLT trigger path. The present study has been aimed as a Run 3 preparatory efficiency cross-checks for the HLT electron/photon (e/γ) filters, using the real electrons selected by the Tag and Probe methodology.

[4] In its flight from the interaction point to the ECAL, the electron emits bremstrahlung radiation and due to the strong solenoidal magnetic field they subsequently lead to a spray of energy in the ϕ-direction. The overall electron energy therefore requires to consider all such cluster energies and is combinedly called "super-cluster".

4.3.1 Tag and Probe Method

In the collisions dataset, it is quite difficult to identify the electromagnetic clusters originating from the real electron/photon on a cluster-by-cluster basis. However, the profile of electromagnetic clusters from true electrons can be ensured by using the ones from the $pp \rightarrow Z \rightarrow e^+e^-$ events. In $Z \rightarrow e^+e^-$ events, the electron-positron pairs are expected to have an invariant mass consistent with the Z-boson resonance. Thus, a selection consisting of exactly two oppositely charged electrons, with their invariant mass within the Z-boson mass window (10 GeV around the Z-boson peak), can lead to a significantly pure sample of electrons. Typically in a tag and probe method, one of the clusters (either from electron or positron) is selected, passing very stringent offline selection criteria for the electron, called "tag", while the other electron in the event, called "probe" must have the invariant mass (with the "tag" electron) within the Z-boson mass window. All the efficiencies are usually checked using the probe electrons as the stringent selection criteria are not applied explicitly.

In this particular study, the efficiency for a HLT e/γ path, called "HLT_Ele32_WPTight_Gsf" consisting of selection criteria for a tight electron reconstructed using Gaussian sum filter (GSF) method with minimum p_T of 32 GeV, has been measured. This trigger path is a sequential combination of thirteen filter algorithms, such as (1) hltEGL1SingleEGOrFilter, and (2) hltEG32L1SingleEGOrEtFilter used for clustering HLT algorithm around L1 seed. Then (3) hltEle32WPTightClusterShapeFilter, (4) hltEle32WPTightHEFilter, (5) hltEle32WPTightEcalIsoFilter and (6) hltEle32WPTightHcalIsoFilter are used for utilizing the calori-meter informations to reconstruct an well isolated trigger candidate. Pixel matching is achieved by the following filters: (7) hltEle32WPTightPixelMatchFilter, (8) hltEle32WPTightPMS2Filter to distinguish between electron and photon. Finally (9) hltEle32WPTightGsfOneOEMinusOneOPFilter, (10) hltEle32WPTightGsfMissingHitsFilter, (11)

hltEle32WPTightGsfDetaFilter, (12) hltEle32WPTightGsfDphiFilter and (13) hltEle32WP-TightGsfTrackIsoFilter are used to reconstruct an e/γ object with Gaussian sum filter (GSF) method. The efficiency of these filters have been checked using the probe electrons, where the tag electrons with $p_T > 37$ GeV and $|\eta| < 2.5$ are required to pass the HLT_Ele32_WPT-ight_Gsf condition. Here, the efficiency of the i-th filter is defined as

$$\frac{\text{Events with probe passing 1st to i-th filter}}{\text{Events with probe passing 1st to (i-1)-th filter}}. \tag{4.1}$$

To illustrate further, let us consider the "HLT_Ele32_WP-Tight_Gsf" trigger. If N_{all} is number of probes passing all the criteria up to the 10-th filter (hltEle32WPTightGsfMissin-gHitsFilter), and N_{pass} (out of the N_{all} probes) probes pass the next filter criteria (11-th filter i.e hltEle32WPTightGsfDetaFilter), then the ratio ($\frac{N_{pass}}{N_{all}}$) represents the efficiency of the 11-th filter. The efficiencies are measured here as functions of the number of pileup events, transverse momenta (p_T), and pseudorapidity (η) of the HLT electrons, as shown in Fig. 4.8. The results are compared between 2018 (Run 2) and 2021 (Run 3)[5] simulations to understand any unexpected behavior of the HLT electron/photon filtering. Similarly, in Fig. 4.9 to Fig. 4.12 comparison between 2018 (Run 2) and 2021 (Run 3) efficiencies have been presented for the constituent filters of "HLT_Ele32_WP-Tight_Gsf".

To summarize, the efficiency studies for several e/γ HLT trigger algorithms have been performed with the tag-and-probe method using Z $\rightarrow$ ee MC events. These studies have been performed as functions of number of pile-up vertices, and the (p_T, η) of the e/γ objects. Such methodology has been practiced regularly for several e/γ HLT algorithms and no degradation in terms of e/γ HLT trigger efficiencies have been observed. In addition, the methodology provides further scope to optimize the triggering algorithm for Run 3 operations.

[5]LHC Run 3 was originally scheduled in 2021, but finally started in 2022.

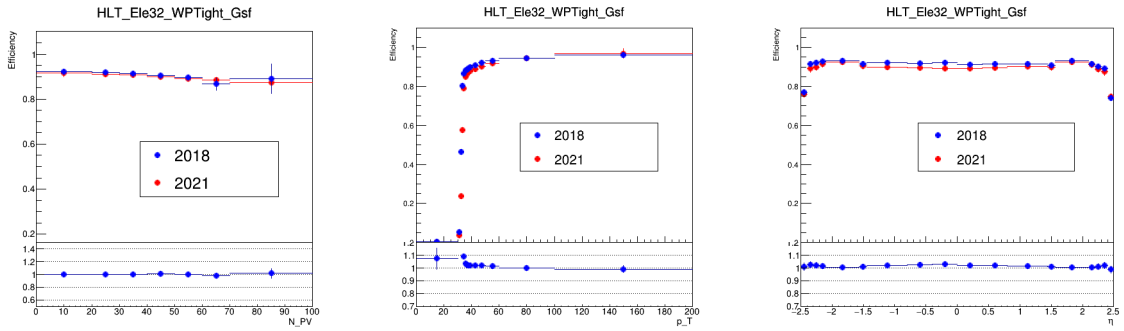


Figure 4.8: Full trigger HLT_Ele32_WPTight_Gsf efficiency as function of number of pileup (left), p_T of e/γ object (middle), η of e/γ object (right).

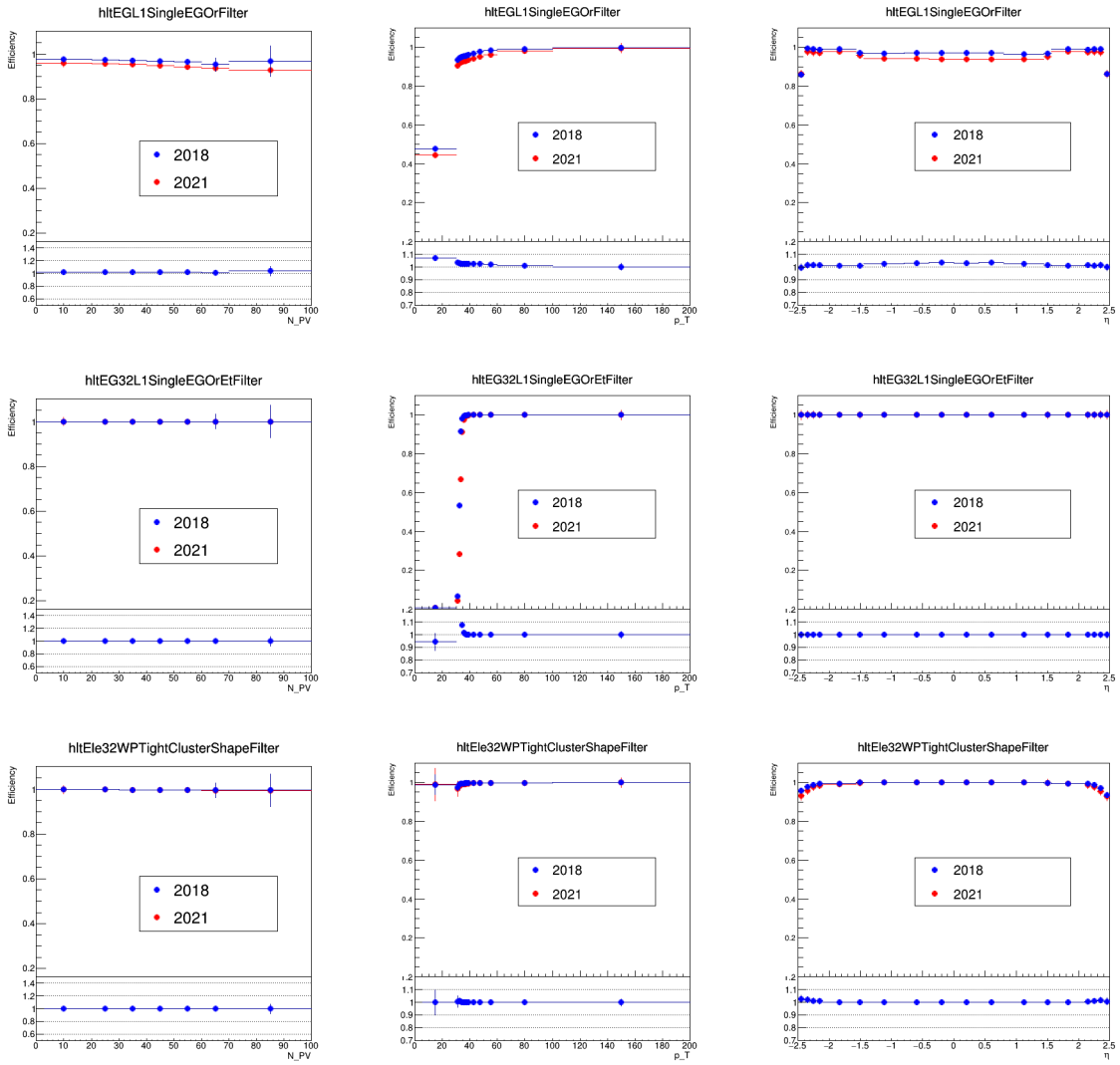


Figure 4.9: Efficiency of the filters of the trigger HLT_Ele32_WPTight_Gsf as function of number of pileup (left column), p_T of e/γ object (middle column), η of e/γ object (right column). Efficiency of hltEGL1SingleEGOrFilter (top row), hltEG32L1SingleEGOrEtFilter (middle row), and hltEle32WPTightClusterShapeFilter (bottom row) have been shown here.

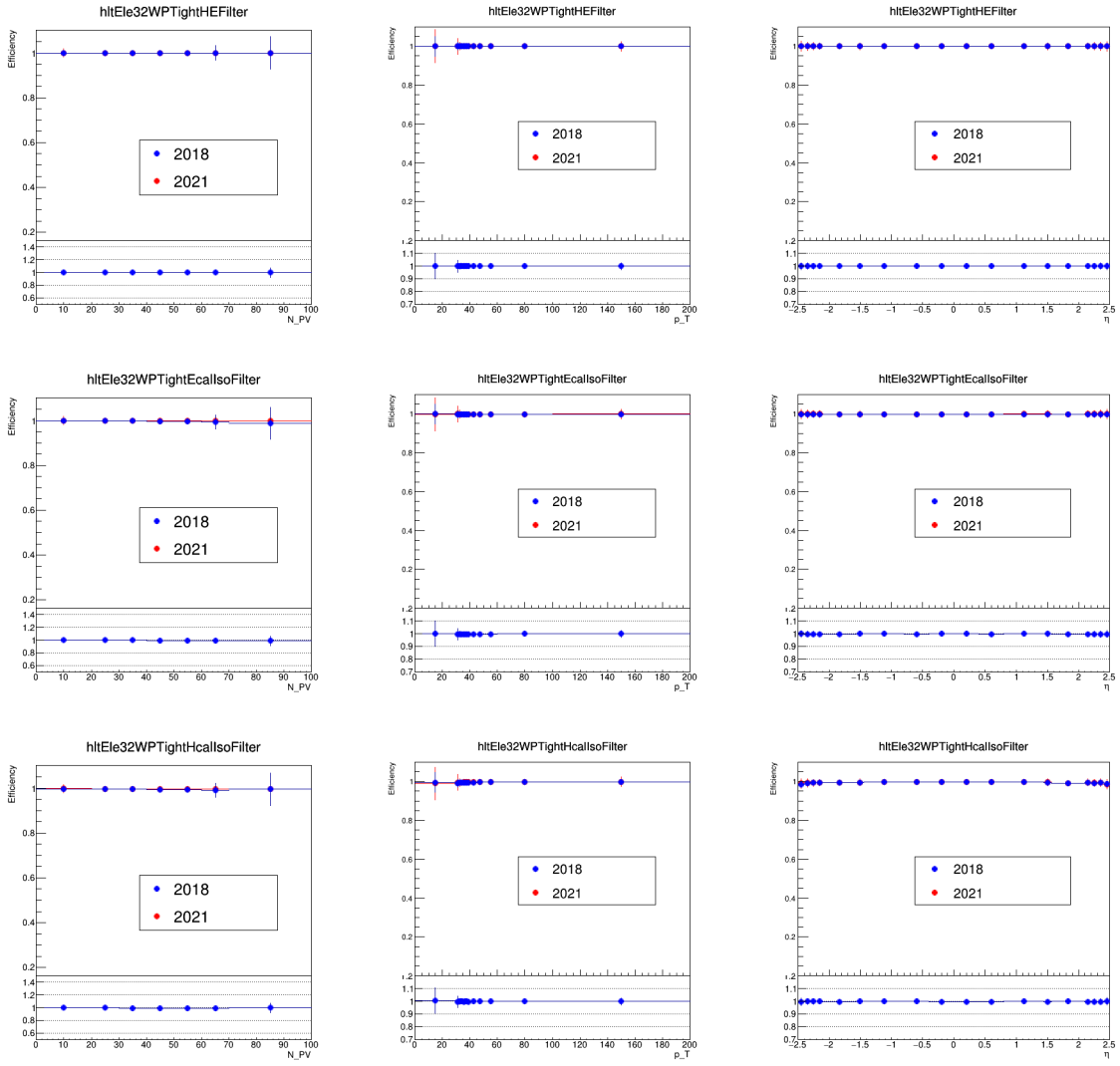


Figure 4.10: Efficiency of the filters of the trigger HLT_Ele32_WPTight_Gsf as function of number of pileup (left column), p_T of e/γ object (middle column), η of e/γ object (right column). Efficiency of hltEle32WPTightHEFilter (top row), hltEle32WPTightEcalIsoFilter (middle row), and hltEle32WPTightHcalIsoFilter (bottom row)

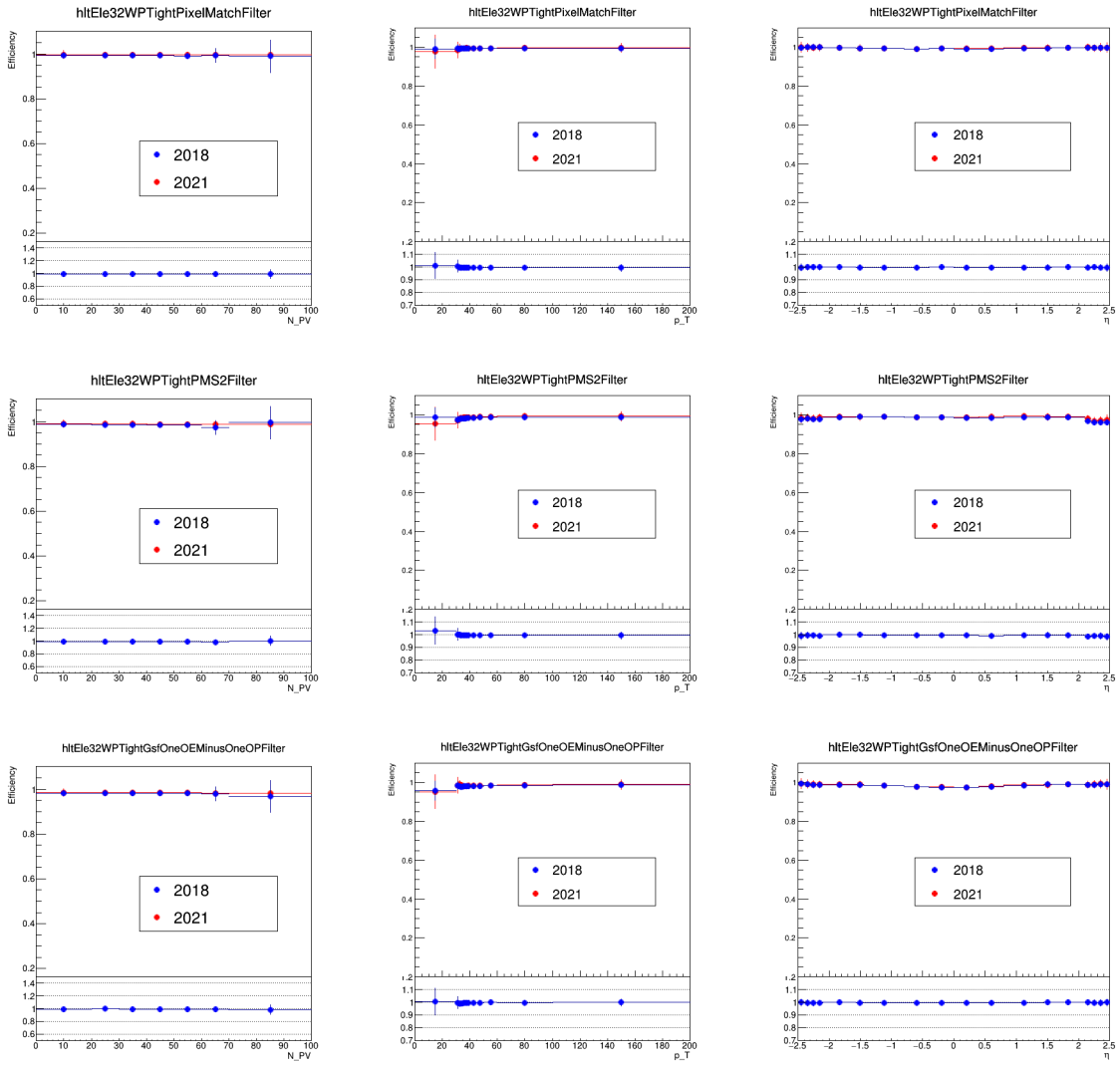


Figure 4.11: Efficiency of the filters of the trigger HLT_Ele32_WPTight_Gsf as function of number of pileup (left column), p_T of e/γ object (middle column), η of e/γ object (right column). Efficiency of hltEle32WPTightPixelMatchFilter (top row), hltEle32WPTightPMS2Filter (middle row), and hltEle32WPTightGsfOneOEMinusOneOPFilter (bottom row).

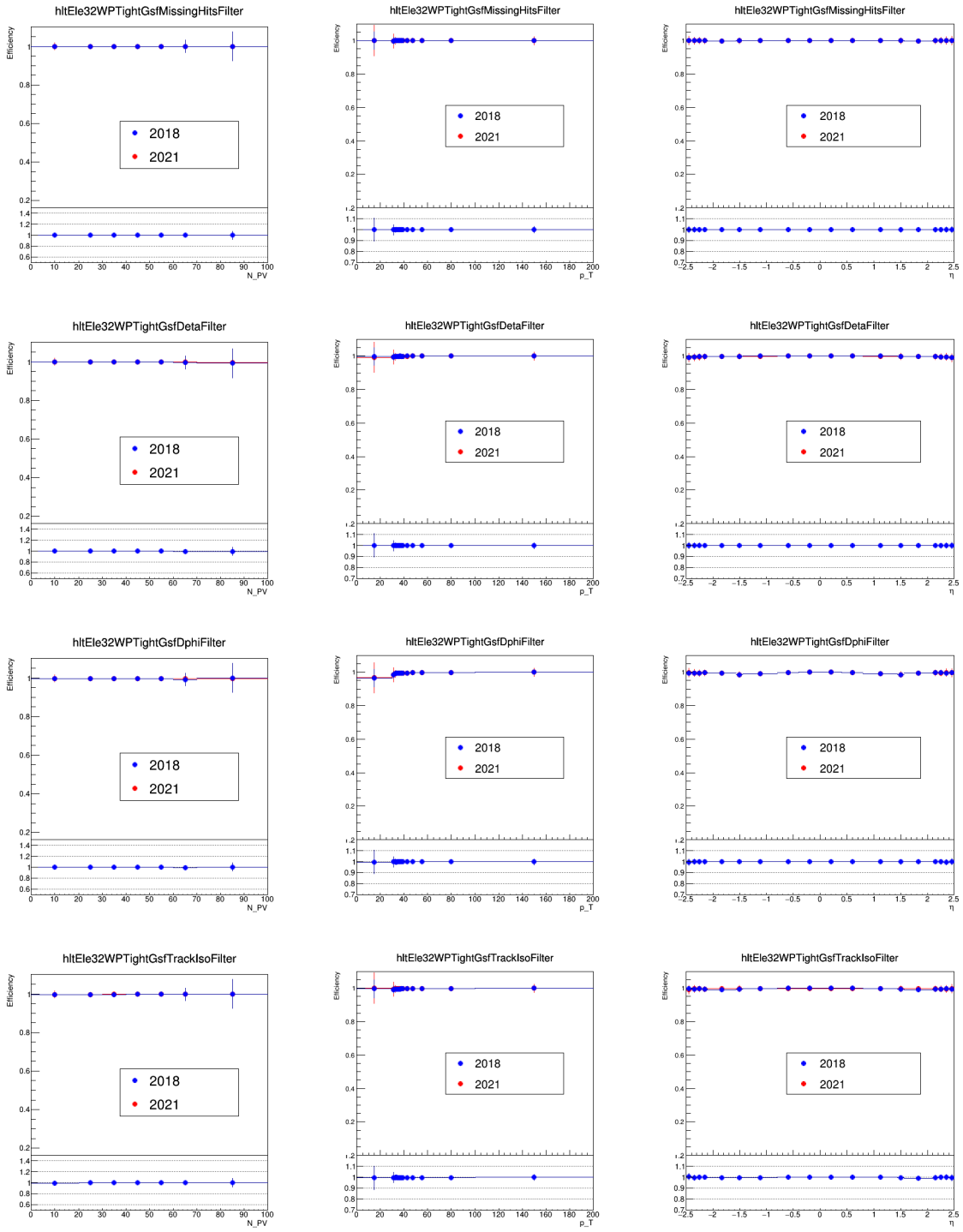


Figure 4.12: Efficiency of the filters of the trigger HLT_Ele32_WPTight_Gsf as function of number of pileup (left), p_T of e/γ object (middle), η of e/γ object (right). Efficiency of hltEle32WPTightGsfMissingHitsFilter (first row), hltEle32WPTightGsfDetaFilter (second row), hltEle32WPTightGsfDetaFilter (third row), and hltEle32WPTightGsfTrackIsoFilter (row) have been shown here

Chapter 5

Phase 2 Outer Tracker Upgrade

5.1 Phase 2 Upgrade of CMS

Despite the fact that the LHC experiments have been running for more than a decade now, some of the key fundamental physics questions still remain unanswered. Both the ATLAS and CMS experiments have completed the long-sought SM Higgs boson discovery through LHC Run 1, followed by precise coupling measurements with Run 2 dataset, but without any definite signature for the physics beyond the SM (BSM). The model specific new physics searches, Super Symmetry (SUSY), extra-dimensions, etc. also resulted in negation so far, constraining the mass limits on the BSM particles at the TeV scale. The other possible avenues to search for the BSM physics are pursued through the study of rare SM decays, e.g., $B_s \rightarrow \mu^+\mu^-$, Falvour Changing Neutral Current (FCNC) decays of the top quark, etc. However, precision SM measurements, rare decay studies, along with the refined new physics searches require much larger datasets even considering the present center-of-mass energy ($\sqrt{s} = 13 - 14$ TeV).

The LHC has already been planned for stage-wise upgradation, even before its operation started, in terms of its energy (up to 14 TeV), and the luminosity, as shown in Fig. 5.1. After the glorious Run 1 operations in 2011 and 2012 delivering nearly 25 fb^{-1} (individually to CMS/ATLAS) of luminosity at $\sqrt{s}$=7 and 8 TeV, respectively, the LHC went into the Long Shutdown period 1 (LS1). During the LS1, several modifications and risk protection mechanisms for the LHC magnets were performed to enable it to run at $\sqrt{s}$=13 TeV. In addition,

several other renovations work in the LHC accelerator chain, i.e., the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS), allowed the LHC to surpass the designed peak luminosity of 1×10^{34} cm^{-2}s^{-1} with bunch spacing time of 25 ns. During Run 2, the LHC performance had been splendid; it had delivered 140 fb^{-1} luminosity to each of the experiments. After Run 2 in 2019, LS2 started, where the PS is upgraded again to operate with a new schema of bunch trains, and the improvements with the injector chain would deliver very bright bunches (high intensity and low emittance). All these improvements and modifications are expected to drive the peak luminosity to 2×10^{34} cm^{-2}s^{-1} during Run 3, and by the end of 2025 the LHC is expected to deliver 300 fb^{-1} of pp collisions ($\sqrt{s}$=13.6 TeV) dataset to each of CMS and ATLAS.

The period that follows LS3 where the LHC luminosity would be heavily increased is referred to as the High Luminosity LHC (HL-LHC) or Phase 2. At the end of Run 3, the quadrupole focusing magnets at the ATLAS and CMS collision regions would be replaced as they would be close to the end of their lives due to radiation exposure. In addition, the low-β quadrupole triplets would be replaced. Also, the crab cavities would optimize the overlap of bunches in the interaction region. However, the LS3, originally envisaged to start in 2024, is delayed to 2026 due to global pandemics and the associated constraints. Overall, during Run 4 (from 2029) the expected operating scenario is to level the instantaneous luminosity at $5 - 7.5 \times 10^{34}$ cm^{-2}s^{-1} with 250 fb^{-1} pp collisions per year over a decade of operation at $\sqrt{s}$=13.6-14 TeV.

With the increase of the LHC luminosity, the CMS detector is exposed to a harsher radiation environment leading to detector damage as well as an increase in the data volumes. During Run 2, the maximum number of pileup events, i.e., the number of collisions per bunch crossing, already reached to a value of > 70 (See Fig. 5.2), while for HL-LHC it is projected to be 140-200. In such an unprecedented scenario, the design integrated luminos-

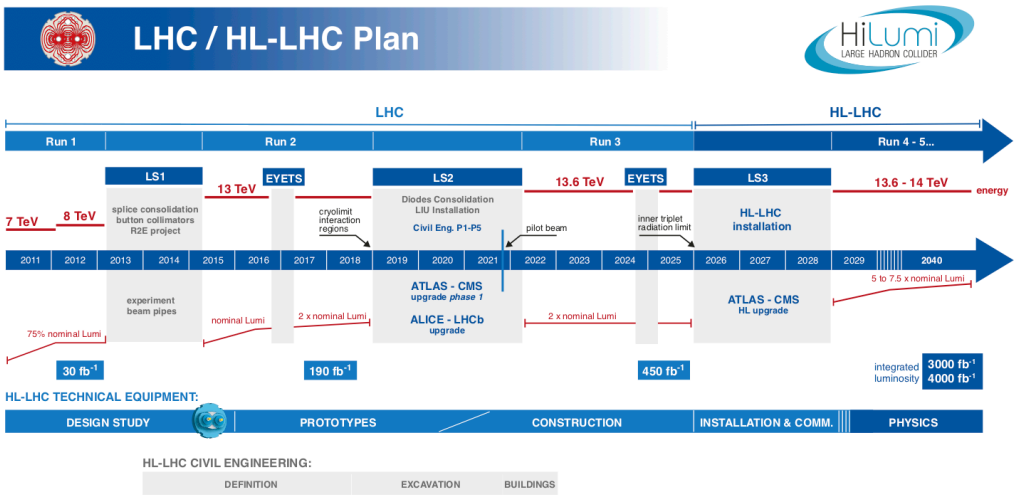


Figure 5.1: The latest LHC upgrade schedule along with the projected peak and delivered integrated luminosities as updated in January, 2022 [33].

ity of 3000 fb^{-1} would be equivalent to a 1 MeV neutron affluence of 2.3×10^{16} n_{eq}/cm^2 at the heart of the CMS detector. Therefore, to maintain its operational performance without deteriorating the physics goals, the CMS detectors have also been going through several major upgrades during the LS2 and LS3, respectively, referred to as Phase 1 and Phase 2 upgrades. While the CMS Phase 1 [77] consisted of upgrades related to Pixel detectors, trigger hardware, and Hadron calorimeter. The CMS Phase 2 upgrade [78] involves a much more extensive detector upgrade programme for the outer tracker, calorimeter, muon systems, and Trigger system. In addition, a completely new minimum ionizing particles (MIP) Timing Detector, called MTD [79] has been planned for installation inside the CMS detector for the additional capability to handle the pile up events more efficiently.

5.1.1 Tracker Upgrade

The present CMS tracker, called Phase 0 tracker, consists of two parts – an Inner Tracker (IT) based on silicon pixel modules and an Outer Tracker (OT) consisting of modules made

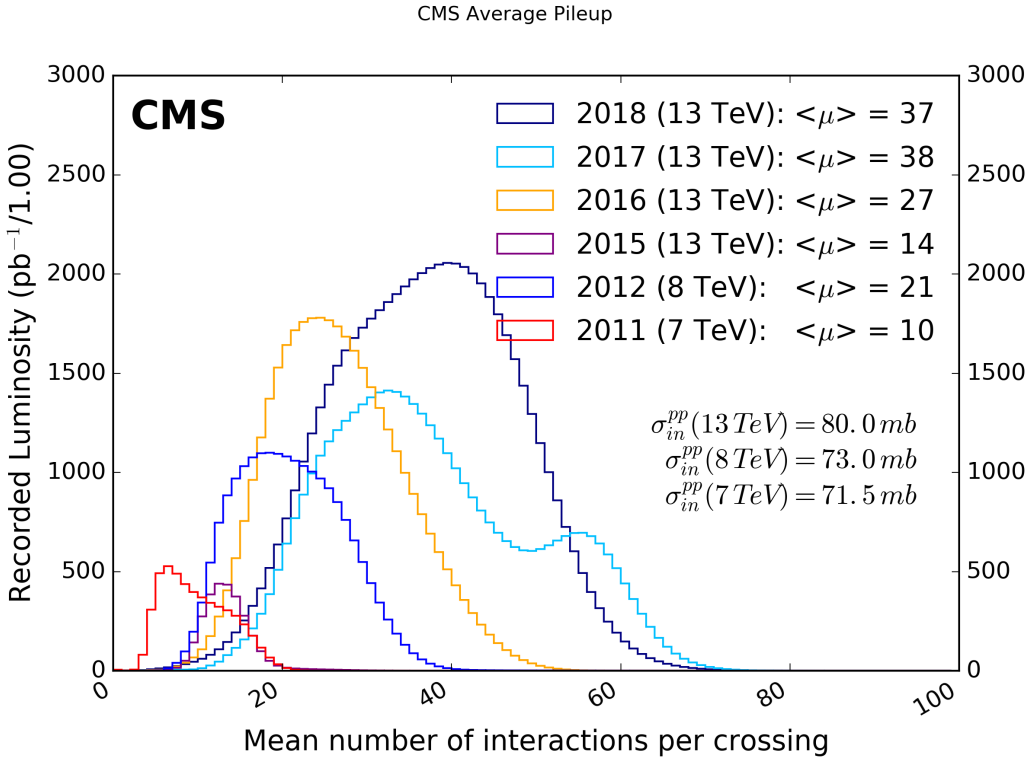


Figure 5.2: The average number of pile up events as recorded by CMS individually for all years of LHC operations so far (Run 1 and Run 2) [80]

with silicon strip detectors. During the 2016/17 extended year-end technical stop (EYETS), the original pixel detector was replaced. The present strip tracker is still operational, although it was originally designed to sustain its performance up to a luminosity exposure of 500 fb^{-1}. However, it has been demonstrated that after exposure to 1000 fb^{-1} of integrated luminosity, the Phase 0 double-sided strip modules would not be operational due to the increase of the sensor depletion voltage and large leakage current. Therefore, it is completely inevitable that the CMS outer tracker needs a complete replacement during the Phase 2 upgrade to enrich the CMS physics program.

Considering various aspects like radiation tolerance, increased granularity, reduced ma-

terial budget, robust pattern recognition, and better two-track separability, extended acceptance, the Phase 2 OT has been designed to contain silicon detector modules with strip and macro-pixel sensors, while IT would consist of silicon pixel modules. Most importantly, the OT design has been driven by the requirements for track-triggering capability at Level-1 (L1)[1] with binary readout. As shown in Fig. 5.3, the Outer Tracker is composed of six cylindrical layers in the barrel region complemented by the five double disks on each side. Basically, two types of modules are used in the OT: PS (pixel-strip) modules composed of a strip and a macro-pixel sensor and 2S (strip-strip) modules composed of two strip sensors. The three sub-detectors are distinguished as the Tracker Barrel with PS modules, or TBPS, the Tracker Barrel with 2S modules, or TB2S, and the Tracker Endcap Double-Discs, or TEDD.

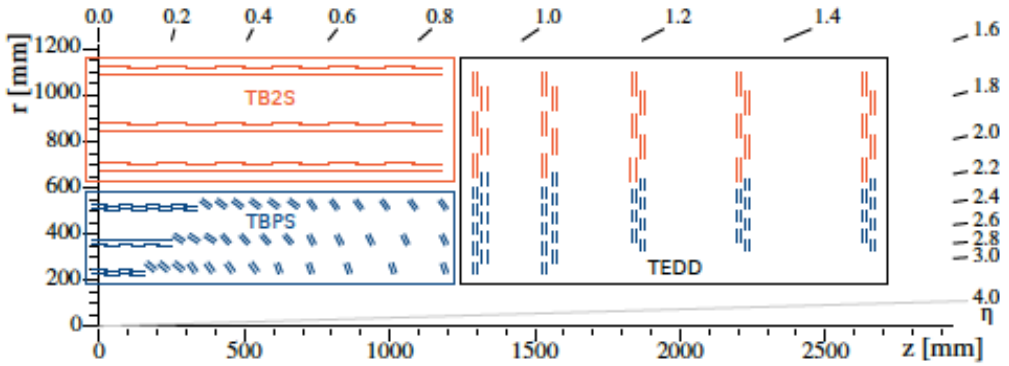


Figure 5.3: One quarter of the CMS Phase 2 Outer Tracker in r-z view: Blue (red) lines represent PS (2S) modules while the three sub-detectors, named TBPS, TB2S, and TEDD, are indicated [81].

While IT with four layers of pixel detectors will help in excellent vertex identification by providing three-dimensional hit coordinates, the OT (with 2S and PS modules) would have long enough depth for precision momentum measurement for the charged particles with $p_T > 2$ GeV. The PS modules are instrumented with two sensors – a strip sensor with

[1] L1 tracking requires the OT track reconstruction within 4 μs of a stipulated time.

2×960 strips having a pitch of 100 μ, and a pixelated sensor with 32×960 macro-pixels. The 2S modules in OT barrel and the endcap disks would consist of two identical strip sensors, each with 2×1016 strips with a pitch of 90 μ and a length of 5 cm. The schematics, along with their geometry and connections to front-end (FE) hybrids, are shown in Fig. 5.4. In the outermost barrel region (TB2S), the 2S modules would be mounted on "ladder" structures (see Fig. 5.5, left) where the consecutive modules are mounted on opposite sides of the ladder in ϕ-direction, and the consecutive ladders are staggered in r-direction (Fig. 5.5, right). Here, the ladder covers half of the barrel, but the overlap between the two ladders ensures hermetic coverage at z = 0. In the central part of the TBPS, the PS modules would be mounted on the "planks" with a radial staggering of the modules by alternating them on the two sides of the plank (Fig. 5.5, left), while consecutive planks along ϕ-direction are located at different radii. A single plank in z-direction would cover the entire length, and with the odd number of modules, any gap at z = 0 is avoided. In the endcaps, the PS and 2S modules are mounted on the TEDD double-discs respectively in inner (r$<$ 60 cm) and outer (r$>$ 60 cm) parts (See Fig. 5.7). For hermetic coverage, z-staggered modules are mounted through the D-shaped parts called the "dees". The consecutive rings along r-direction are mounted on the two discs of the pair, ensuring radial overlap.

5.1.1.1 L1 Tracking and the p_T Module concept

It is almost mandatory to have a restructured trigger system to select the events more efficiently using improved object selection to handle enormously high pileup events during HL-LHC operations. Therefore, Phase 2 planning for the CMS L1 trigger system is meticulously performed to accommodate a longer latency time (12.5 μs) incorporating the track association to the objects; the L1 output rate would be increased to 750 kHz. The front-end electronics of the Phase 2 OT are thus needed to be capable of processing information within the 4 μs of stipulated time. For such functionality, the data reduction locally in the

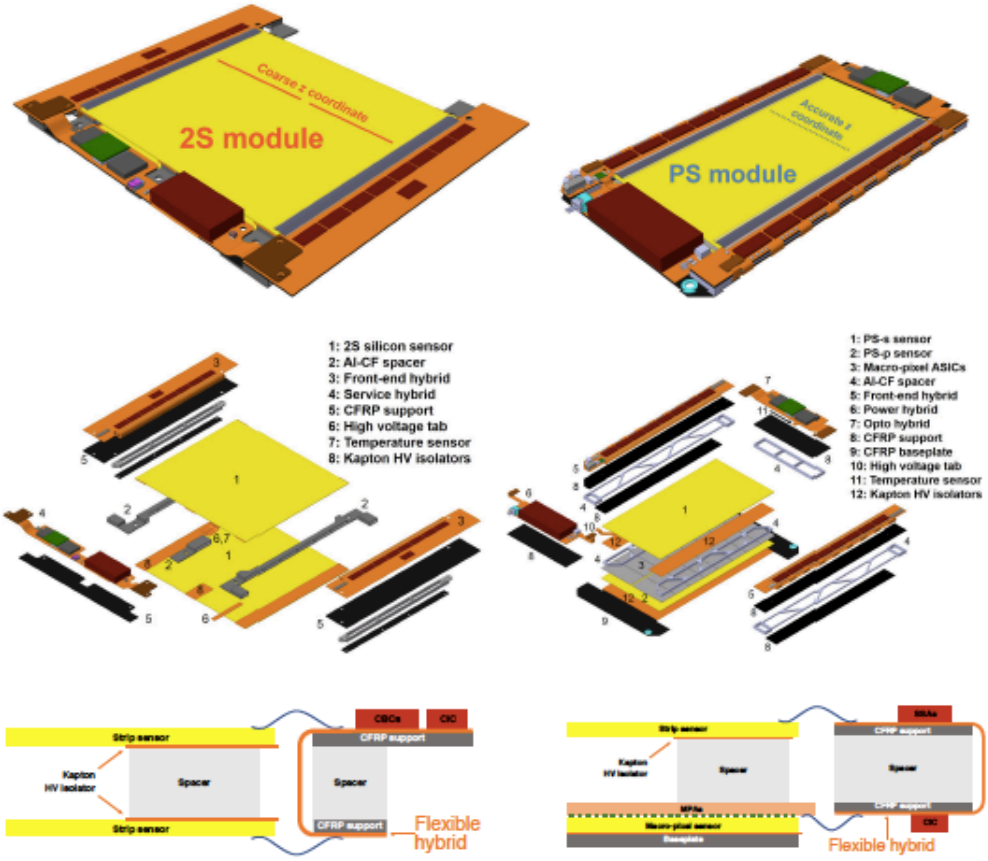


Figure 5.4: 2S(left) and PS (right) Module's schematic view. [81]

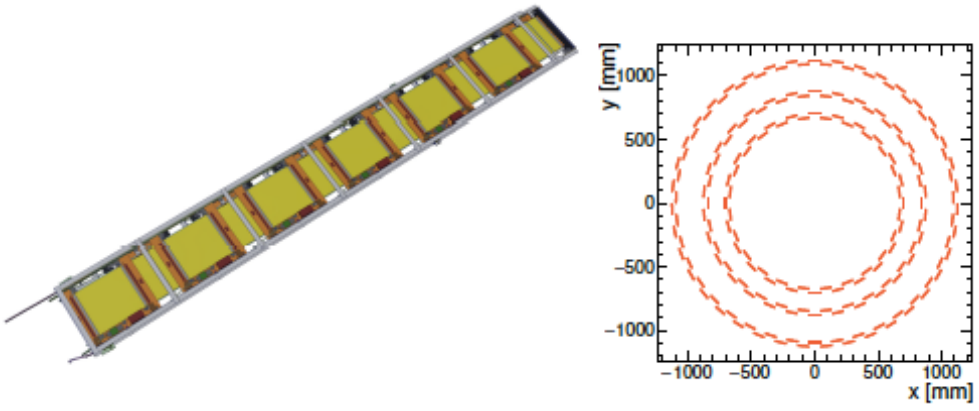


Figure 5.5: Left: model of a TB2S ladder, housing twelve 2S modules. Right: x-y view of the TB2S, showing the staggering of neighbouring ladders [81].

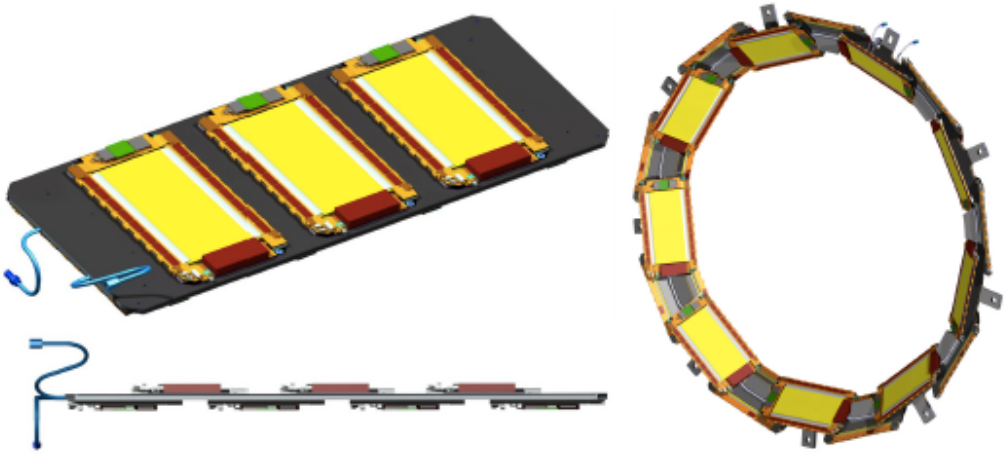


Figure 5.6: The schematics of the TBPS support structures: central section "planks" (left) and the rings (right). A sideview of the plank is also shown in the lower left picture [81].

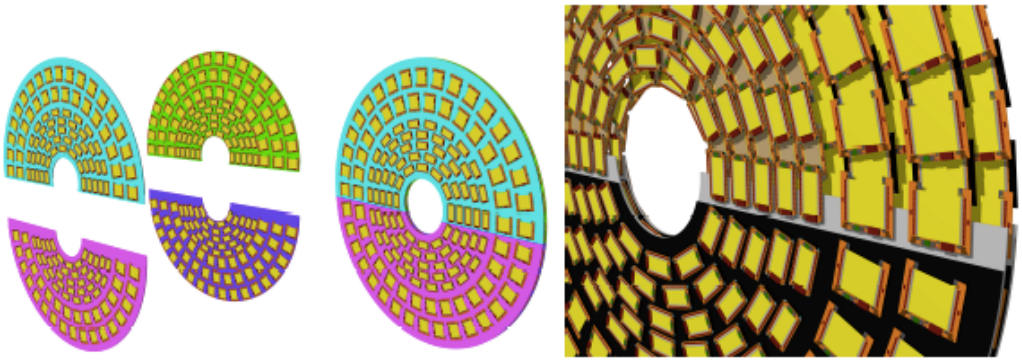


Figure 5.7: Sketch of four dees (left) forming a double-disc (centre), and drawing of a part of a TEDD double-disc (right), illustrating the overlap of modules in ϕ and z. The upper two dee support structures are removed in order to show all layers of modules [81].

front-end electronics are being incorporated through the "p_T modules"[82], where the signals from the tracks below a certain p_T threshold are rejected. The adapted strategy exploits the bending angle (within a module) of the charged particle trajectory under the 3.8 T of CMS magnet, and the correlation of hits from the module sensors within an acceptance window is utilized as the trigger criterion. In other words, within each module, two single-sided closely-spaced sensors are read out by a common set of front-end ASICs correlating

the signals to form the hit pairs, referred to as "stubs" compatible with particles above a p_T threshold (Fig. 5.8). For transmission of the stubs data at 40 MHz rate, a track p_T-threshold of around 2 GeV is found to be the optimal choice here; however, the p_T thresholds can be tuned to a certain level through programming of the readout chip settings. Furthermore, it is the resolution requirement of the p_T modules in different regions that mandated the strip pitches for the OT sensors (strips or macro-pixels).

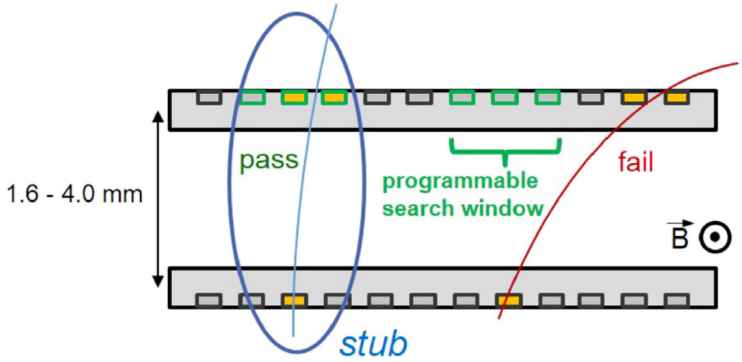


Figure 5.8: Illustration of "p_T module" concept[81].

5.1.1.2 2S Module

As mentioned above, the Phase 2 2S modules consist of two identical silicon strip sensors (each consisting of 2×1016 strips with a pitch of 90 μ) separated by a few mm; based on the gaps between two sensors, there are two versions of the 2S modules called 1.8mm (in TB2S) and 4.0mm (TEDD) modules. As shown in Fig. 5.9, the inter-sensor gaps in 2S modules consist of spacers or bridges (made of aluminum carbon-fiber composite or Al-CF[2]), Kapton strips[3] as HV-isolators. In addition, different types of thermally conductive

[2]This material has both high thermal conductivity and low coefficient of thermal expansion (CTE). This allows for good heat conduction from the sensor and hybrids as well as for low stress on the glue joints between sensors, spacers, and CFRP (carbon fibre reinforced polymer) parts at -25^0 C.

[3]The Kapton MT polyimide film of 25 μ thickness provides the HV isolation of 1000V between sensors and Al-CF spacers and it has good thermal conductivity.

adhesives[4] are used to glue them together.

The FE hybrid (two for each module) is made up of eight CMS Binary Chips (CBC), each with 127 readout channels and overall consists of integrated circuitry of preamplifiers, postamplifiers and comparators. The FE hybrids are supported by carbon fiber reinforced polymer (CFRP) and are folded around Al-CF spacers so that the top and bottom sensors can be connected. The CFRP provides the primary heat conduction path for hybrid cooling through the bridges, and the fold-over part of the hybrid allows for optimal positioning between sensors and hybrids for the ease of wire bonding. Here ultrasonic wire bonding is considered to establish the high reliability connections between the sensor and FE hybrids, as well as for the backplane bias circuit. Furthermore, all wire bonds are required to be encapsulated to reduce risk of handling damage and damage due to possible resonant vibrations in the magnetic field. The signals produced in individual strips are read out by the CBCs in the FE hybrid on both sides of the module are collected by the Concentrator Integrated Circuit (CIC) on the service hybrid (SE). The CIC then sends the data to the counting house through the optical VTRx+ (Versatile TRansceiver plus) optoelectronic transceiver, after serialization through the LpGBT (Low-power Gigabit Transceiver). In addition, the SE distributes power to the electronics through the DC-DC converters to generate appropriate voltages.

5.2 2S Module Wire Bonding

As an integral part of the 2S module assembly, the ultrasonic wire bonding and the subsequent quality control are to be performed at the Module Assembly Center at NISER

[4]The Phase 2 Tracker would be operational at -25^0 C, and the cooling tubes would be connected to the bridges. Therefore, to maintain a uniform temperature for the sensors, efficient heat transfer through the adhesives plays a critical role here.

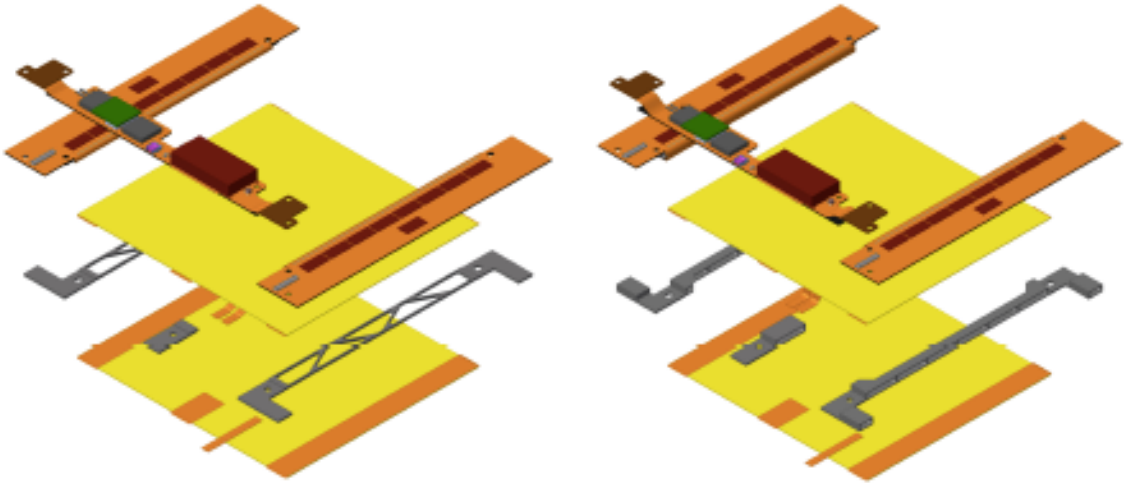


Figure 5.9: Exploded views of the 1.8mm (left) and 4.0mm (right) 2S module variants. From top to bottom: the three hybrids, the top sensor, the bridges, and the bottom sensor. The orange strips on the sensors are the Kapton strips used for HV isolation (between the bridges and the sensors) and flex cables to deliver the bias voltage and/or to carry a temperature sensor [81].

(MPC@NISER). In the subsequent sections of this chapter, a brief introduction of the wire bonding techniques, along with the actual adaptation, optimization, and execution of 2S module wire bonding and pull testing as performed at the MPC@NISER, have been described. Here, the Phase 2 tracker activities at NISER are solely focused on the TB2S modules and ladders, and hence all the wire bonding fixtures and bond parameter optimization are fine-tuned to the specific variant of the 2S modules, i.e., with 1.8 mm gaping between the strip sensors.

5.2.1 Ultrasonic Wire Bonding

In the electronics industry, the ultrasonic joining between two surfaces/substrates, including dissimilar ones, is a well established method being practiced for decades now. It is basically a multi-step welding technique where the induced oscillating shear between the faying surfaces effectively forms the metallurgical bond, hence the formation of electrical connec-

tions. The ultrasonic power (US power) and pressure are applied by a fine tool, called *wedge* to execute the welding process between the connecting wire and the bond surface, while two categories of wire bonding, *wedge-wedge* and *ball-wedge* bondings, are routinely practiced in the semiconductor industry. Typically, in "ball bonding" applications, the thermocompression (TC) techniques create an intermetallic bond while thermosonic (TS) method adds ultrasonic energy into the bonding process. Before bonding occurs, a spark from a "electronic flame off" or EFO beneath the capillary first creates a "free air ball". When the capillary meets the bond pad's surface and provides stress and ultrasonic power to the ball for a certain period of time, the free air ball is then deformed. As a result, the metallization of the bond pad and wire interact, forming the intermetallic bond. Generally speaking, copper wires that have undergone "ball bonding" with gold are better suited for fine pitch applications with pitches of 40 microns or less and enable quicker rates of roughly 5 to 12+ wires per second.

The wire and the bond pad are joined together by the wedge bonding technique, which uses ultrasonic pressure and energy. Wedge bonding, which is similar to ball bonding and employs gold wire, uses temperatures up to 150^0C and is also known as thermosonic (TS) bonding. Low temperature bonding, also known as ambient temperature bonding, is the most common method for wedge bonding. In this method, the interconnections are created using aluminum wire. Because the initial bond formed during the welding process, whether hot or cold, is a wedge bond rather than a ball bond, the process of joining wires is known as wedge-wedge bonding.

Wedge bonding is preferable for much finer pitch applications of 40 microns or less for its advantage due to the absence of the ball on the first bond. Bonding speeds using aluminium or gold wires are generally in the range of 3-6 wires per second, which is considerably slower than a ball bonder.

For CMS Phase 2 OT upgrade related to the wire bonding of 2S modules, fine-pitch wedge-wedge bonding is carried out with aluminum (with 1% silicon) fine-wires (1 mil diameter), and hence the remaining discussions below are completely focused on it. The ultrasonic bond formation process between the workpiece (wire) and substrate (bond pad) progresses through multiple stages as described in Fig. 5.10. In the first phase, called the Pre-Deformation phase the wedge tool applies a static touchdown force (F_{TD}) to the workpiece to create an initial contact area. Next, in the Cleaning phase, the ultrasonic vibration (x_W) and the bond normal force (F_{bn}), sometimes also called "bond force" (can have a different value than the (F_{TD}) to detach the oxide layers and other contamination from the faying surfaces. In the third phase (Deformation Phase), high plastic deformation of the workpiece occurs, formatting the interface region with the substrate. Here, the F_{bn}-values are not changed significantly, and the reduction of contact roughness leads to the increased contact area. In the final phase of Interdiffusion, the material flow between workpiece and substrate occurs without melting the materials and is induced by the oscillating shear stress and plastic strain in the interface. Effectively, the material flow leads to an inter-metallic connection between the workpiece and substrate; thus, two dissimilar metal surfaces with different melting temperatures can be bonded together without being molten.

5.2.2 F&K Delvotec 64000 G5HS

As can be seen in Fig. 5.11, at NISER an $F\&K$ Delvotec [84] 64000 G5HS automatic fine-wire bonding machine has been procured for the production of large number of 2S modules for the CMS Phase 2 OT upgrade. The machine performs wedge-wedge wire bonding with the capability of applying ultrasonic power in a range of 40 to 160 kHz, having an oscillation amplitude of 1-2 μ. The fine-wire (suitable for aluminum or gold wires of diameter 17-75 μ) automatic bonding process over a large region of 254mm x 153mm (in x-y) can

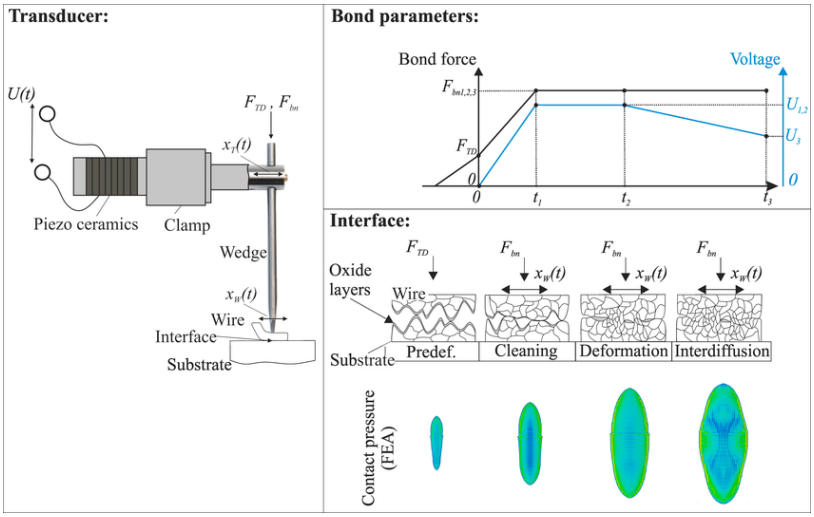


Figure 5.10: On the left, one can see an ultrasonic transducer for wire bonding that is powered by an oscillating voltage U(t) and a wedge that clamps the wire using the bond normal force Fbn. The wire is stimulated by the wedge's amplitude xW while the transducer's amplitude xT (t) excite the wedge to a bending oscillation (t). On the right: Trajectories of the key bond parameters (bond force Fbn and voltage U(t)) during the bond's lifespan, as well as the varying interface conditions throughout the course of the bond's four phases. Results from Finite Element Analysis (FEA) show that during bond formation, the contact area grows and the contact pressure distribution shifts [83].

be executed with the help of pattern recognition software, while the machine has a bonding range of 40mm in z-direction. At present, for 2S module assembly work, the machine is operated with a Tungsten Carbide wedge, *4WFV4-1820-W7C-F00*[5] suitable for fine-pitch applications. To handle the challenges of fine-pitch wedge bonding (i.e., to avoid contact between the wedge and adjacent wires), this specific concave wedge tool has additional vertical side relief (VSR) VSRh of 6 mils (see Fig. 5.12), where the ultrasonic power is transmitted through a transducer as shown in Fig. 5.13. The wire bonding for the 2S modules is mandated to be executed with a very specific *semi-hard* aluminum wire with 1% Si is referred to as Al-Si wire because of its mechanical properties, as detailed in Tab. 5.1. Overall, the machine is capable of executing 2-3 wire bonds per second with a standard 45^0

[5]This specific wedge tool is manufactured by Micro Point Pro [85].

wire feed of 25 μ diameter Al-Si wire.

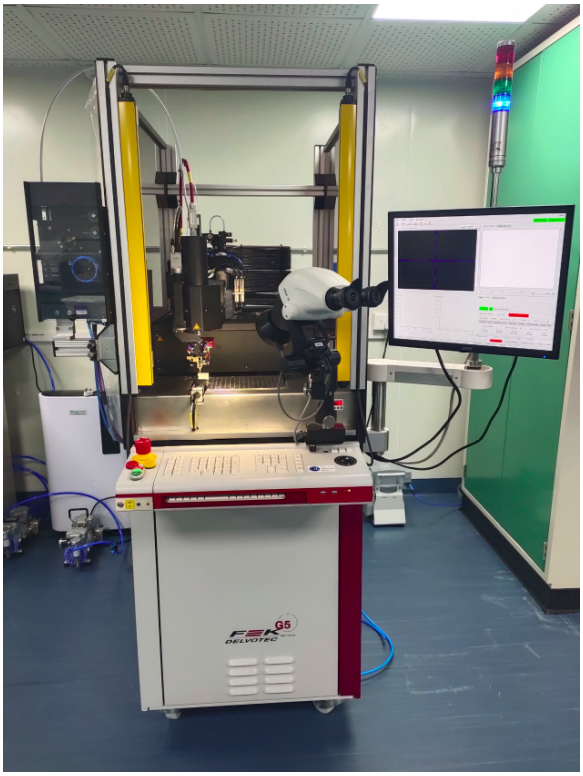
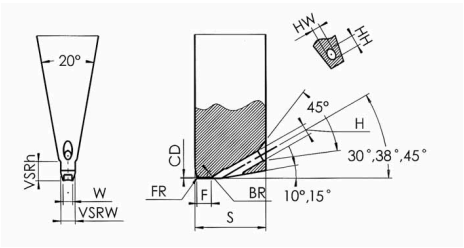


Figure 5.11: F&K Delvotec 64000 HSG5 wire bonding machine at the MP@NISER installed since October, 2019.



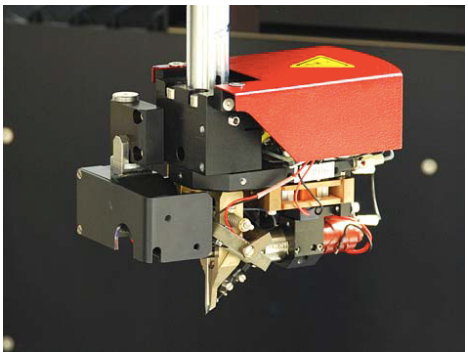


Figure 5.12: Wedge tool picture (left) demonstrating two basic features: the height (VSRh) and the width (VSRw); the vertical side relief (VSR) is intended to increase the clearance between the wedge and the adjacent wire. Actual photograph of the same (right) on the F&K Delvotec wire bonder.

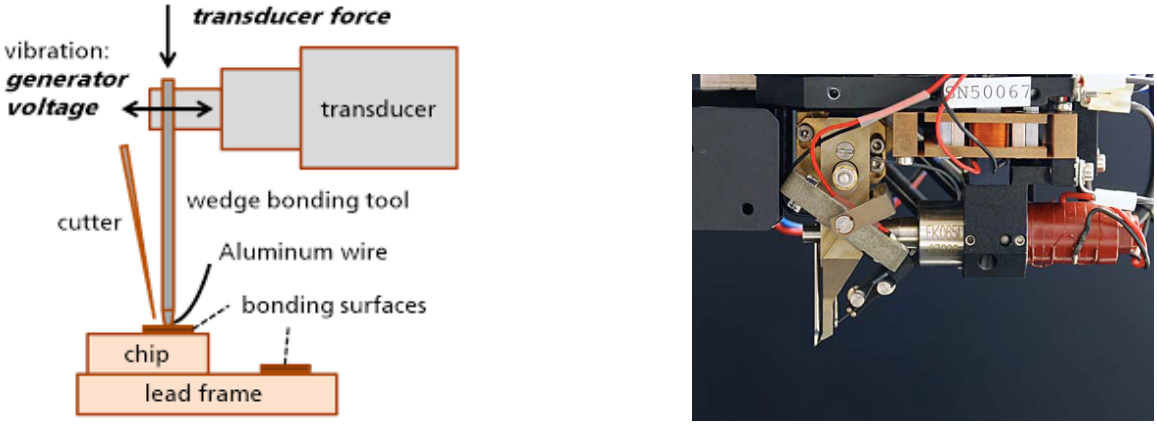


Figure 5.13: Schematics of a typical ultrasonic transducer (left) and its actual photograph on the F&K Delvotec wire bonder (right).

Metal	Diameter	Breaking Load	Elongation
TABW (Al-1%Si)	25$\pm$1 μ	13-15 g	0.5-4.5%

Table 5.1: Specification of the semi-hard Al-Si wire manufactured by Tanaka Kikinzoku, Japan.

For the wire bonding application in 2S modules, there are two different types: the bonding between the bond pads of the hybrid and the sensors, and the ones for the HV tail bond pad and the sensor backplane. In either of the cases, the machine executes the operation in the steps as described in Fig. 5.14, where the actual wire connection between two bond pads, called it Loop, is defined through a geometrical shape (see Fig. 5.15), which can be defined and/or controlled through the software interface of the machine. For example, Fig. 5.15 shows a typical loop geometry, where the wedge moves vertically upward through a distance, called *Z-Loop Presign*. During the upward motion, the wedge also moves in the direction of the second bond pad, called (*XY Loop Height Factor*), until the required *Loop Height* is achieved.

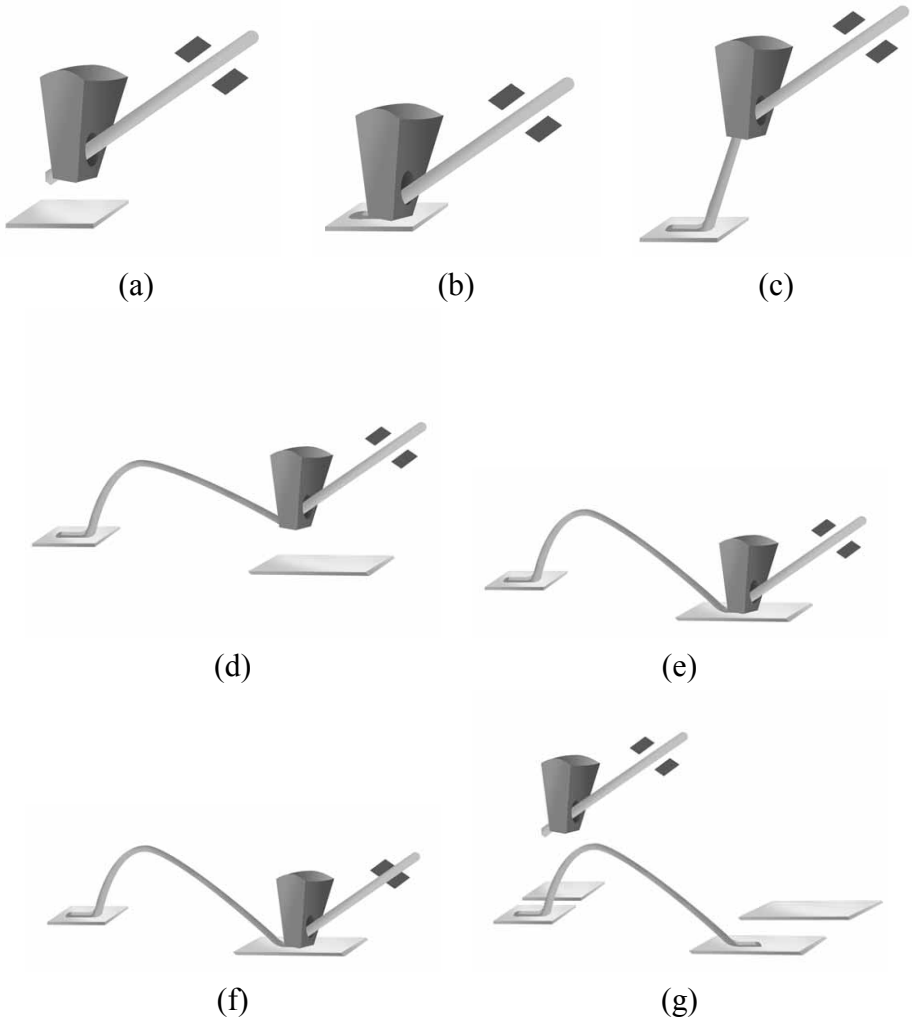


Figure 5.14: A wire is bonded between two bond pads in the following order: (a) The machine precisely targets and aligns the wedge with the targeted bond pad as the wire emerges from the hole beneath the wedge foot. (b) The wedge is lowered, and the foot deforms the wire while transmitting force and ultrasonic vibrations via the wedge. (c) The wedge rises over the pad following the deformation of the first bond. The clamp's aperture enables the wire to pass through. d) With the clamps still open, the wedge tool advances to the location of the second bond; the free feeding of the wire through the wedge hole produces the loop formation, which is dependent on the machine bondhead motion profile (most common are Square and Triangle). (e) The wedge moves down toward the second bond pad, forcing the wire into the lead with the foot while force and ultrasonic energy are used to form the second bond. f) At the completion of the second bond, the wire clamps retract, drawing the wire and forcing it to break at its weakest point; a clean termination of the wire at this point is essential for maintaining uniformity in tail length. (g) The bond head elevates the wedge to its starting height, and the clamps force the wire into the hole beneath the foot. The wedge is prepared for a new cycle and a new tail is created as a result [85].

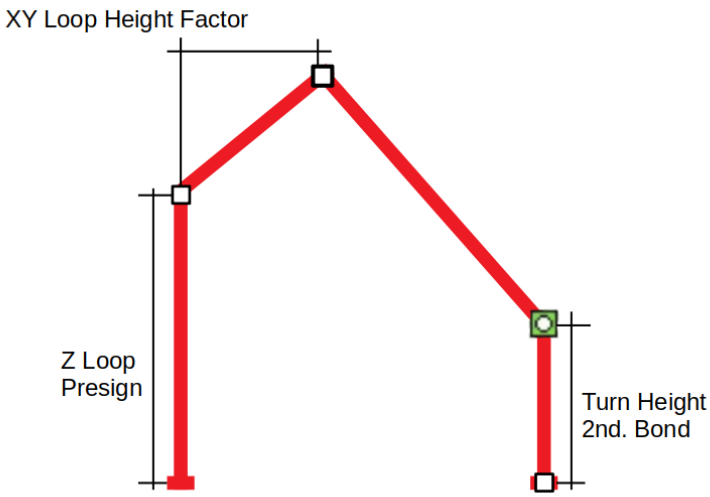


Figure 5.15: Geometry of the wire loop as can be defined through the control software of the wire bonding machine.

The wire is pressed and deformed by the wedge tool during the wire bonding process, resulting in a lower height and proportionately bigger breadth, and the wire is then bonded with its "deformed shape" (see Fig. 5.16). The bonder can record the final deformation after the specified bond time and can track the deformation speed over time. Thus can monitor the deformation in the Z direction, typically in a time-resolved manner, and it offers an excellent quality control tool. A set of typical deformation curves for the first and second bonding during the 2S module wire bonding operation is shown in Fig. 5.17. This deformation curve is utilized to modify the ultrasonic power higher or lower in a more sophisticated version created and implemented by F&K Delvotec if the deformation run ahead or lags behind a preprogrammed standard curve. It is a patented control programme known as "*Bond Process Control (BPC) [86]*" that enables automatic fine-tuning of the wire bonding process to lower error rates and is anticipated to lead to higher bonding quality (and perhaps the lifetime) of the detector module.

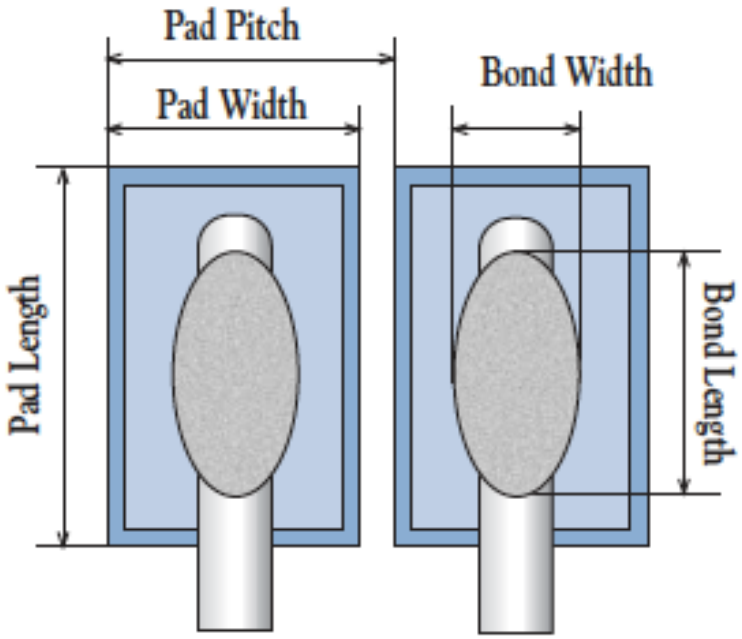


Figure 5.16: Deformation of the wire in bonding direction as well as the lateral direction, respectively denoted as bond length and bond width; the bond pitch is also shown here.

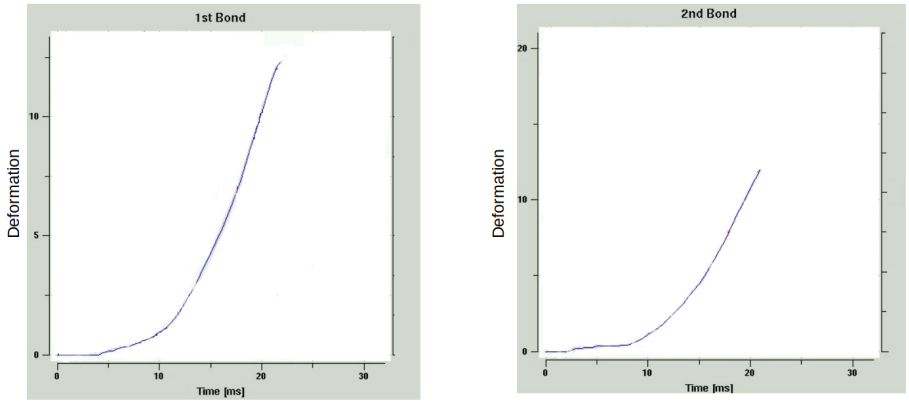


Figure 5.17: Typical wire deformation curve in z-direction at different times within of bonding time period for first (left) and second (right) bonds.

5.3 Wire Pull Test

The quality of the wire bonding for the 2S modules is needed to be ensured at the highest possible level because of the following obvious reasons.

- the modules would not be accessible frequently for any replacement/repair.
- the breakage of bonds would lead to noisy channels and hence the detector efficiency would be lower.
- the bonded wire should be able to withstand the forces under the strong magnetic field
- the bonding should withstand thermal contraction at the operating temperature.

To ensure such requirements, there are two commonly used methods adapted apart from optimizing the bonding process: pull test i.e., to pull the wire, and visual inspection of the wire deformation. The pull test procedure can be classified in two categories:

- **Destructive pull test:** In this type of test, a test hook is carefully first placed under the bonded wire. The hook is now continuously pulled upwards (with a predefined velocity) until the wire breaks, and the hook stops moving further. The corresponding force and the height[6] is noted down. Also, it becomes important to note the position of the wire breakage, and they are classified into different types, as shown in Fig. 5.18. If the whole wire lifts off from the surface, it is called *Lift-off*, while the failure mode is called *heel breaks* if the wire cuts exactly at the heel of the bond. Other than these two cases, if wire breaks far away from the bonding position (usually in the region of the hook position), it is called *wire break*.
- **Non-destructive pull test:** Similar to the destructive case the test hook is carefully placed under the wire. But instead of breaking the wire with increasing force, a predefined force is applied for a predefined time. If the wire is able to withstand the force, it passes the test.

[6]The force needed to break the wire is called pull force. And the height is called pull height

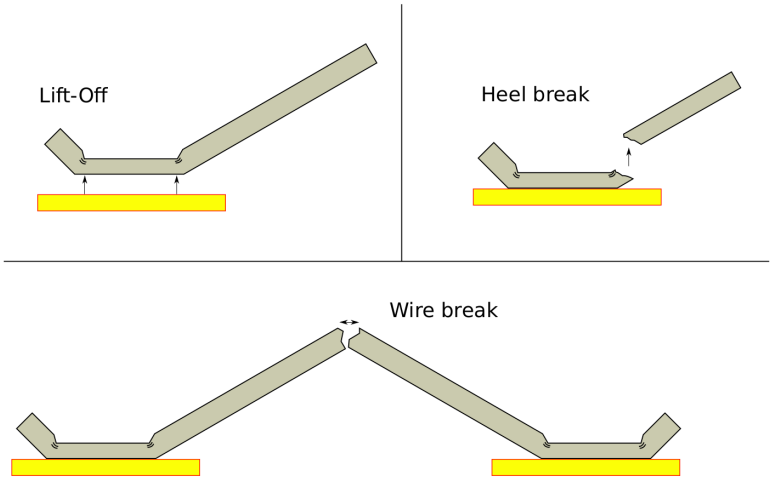


Figure 5.18: Different types of failure modes for destructive pull tests.

At NISER, a Nordson-Dage 4000 Plus series pull tester has been commissioned and operational for the purpose of routine pull testing. The machine is equipped with a cartridge to apply up to 100 gm force on wire bonds (see Fig. 5.19). Although the force, **F** is the total force applied on the wire, it can have two effective components (F_1 and F_2) applied to the two bond pads. The values of F_1 and F_2 depend on the position of the test hook along the wire and the wire loop shape. Therefore, two correction factors, K_1 and K_2, have to be applied on **F** to get the effective force components. Fig. 5.20 represents the two dimensional geometry of the loop shape and all the variables needed to calculate the correction factors.

$$F_1 = K_1 \times F \tag{5.1}$$

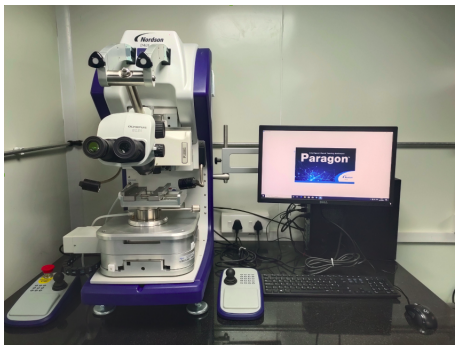
$$F_2 = K_2 \times F. \tag{5.2}$$

One can use the loop height H, test hook distances from two bonds i.e d_1 and d_2, and bond surface height difference (Δ) to calculate the angles, α and β. By using the Eqs. 5.3 one can easily calculate the correction factors and in this chapter all the correction factors related to

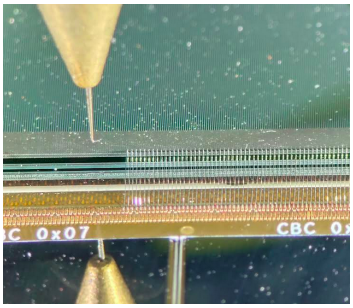
the pull testing results are performed using the same equation.

$$k_1^s = \frac{cos\beta}{sin(\alpha+\beta)} \tag{5.3}$$

$$k_2^s = \frac{cos\alpha}{sin(\alpha+\beta)} \tag{5.4}$$



(a) Pull testing machine



(b) Close up view of pull testing hook and wires

Figure 5.19: Photograph of Nordson-Dage 4000 Plus pull tester.

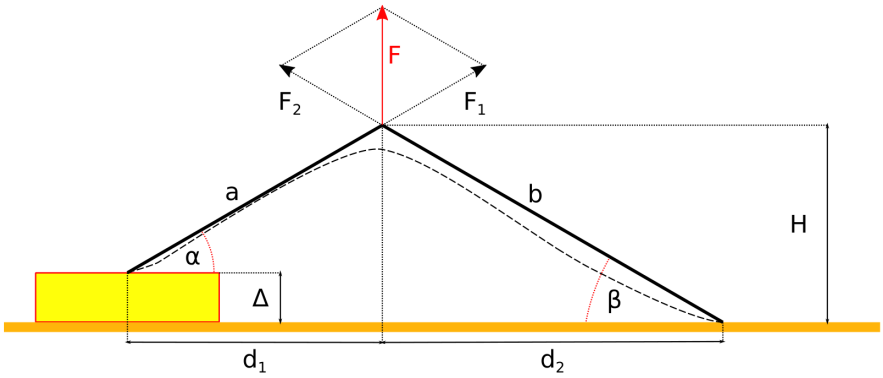


Figure 5.20: Sketch illustrating the calculation of the correction factors for pull tests.

The wire bonder manufacturer, F&K Delvotec Bondtechnik GmbH has derived a more accurate formula through the simulation considering the 3D configuration of the forces.

$$k_1^D = \frac{1}{e^{0.3\times(\alpha+\beta-(sgn(2\alpha-\beta)+1)(2\alpha-\beta))}sin\beta + sin\alpha}, \tag{5.5}$$

$$k_2^D = \frac{1}{e^{0.3\times(\alpha+\beta-(sgn(2\alpha-\beta)+1)(2\alpha-\beta))}sin\alpha + sin\beta}. \tag{5.6}$$

Considering the elongation factor, ϵ provided by the wire manufacture, the following relation can be established:

$$\epsilon L = \hat{L} = a + b = \sqrt{d_1^2 + (H-\Delta)^2} + \sqrt{d_2^2 + (H)^2}. \tag{5.7}$$

Here, H is measured by the pull tester and one can calculate α and β:

$$\alpha = tan^{-1}\left(\frac{H-\Delta}{d_1}\right) \tag{5.8}$$

$$\beta = tan^{-1}\left(\frac{H}{d_2}\right) \tag{5.9}$$

5.3.1 Visual Inspection

Although the pull test is a good quality check for wire bonds, it is always recommended to perform a visual inspection under the inspection microscope to get a better understanding of the bond failure. Since the dust particles and grease on the bond surface may damage and contaminate the surface, which may degrade the bond quality, it is always good to check the bond surface using an inspection microscope in the vicinity of the bond contacts. Through this inspection, one can measure the wire deformation *B*, bond tail length *T*, and the bond contact asymmetry, B_{max}/B_{min} (see Fig. 5.21). F&K Delvotec Bondtechnik GmbH has recommended some benchmark values for these parameters as tabulated in Tab. 5.2. Furthermore, the loop shape and any non-uniformity can be checked through visual inspection.

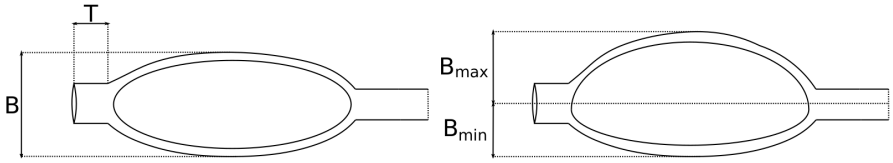


Figure 5.21: Illustration of the deformation parameters to be measured through the inspection microscope.

Deformation B	$\leq 2 \times$ WD
Asymmetry B_{max}/B_{min}	≤ 1.25
Tail length T	$\leq 3 \times$ WD

Table 5.2: F& K Delvotec Bondtechnik GmbH recommended benchmark parameters for visual inspection. WD represents the diameter of the wire.

5.4 Optimization of Bond Parameters

The quality of wire bonding can be controlled and operated through many parameters that can be set through the wire bonder software. However, most of the key parameters are listed below:

- Ultrasonic Power (USP unit)[7]
- Bond Normal Force, F_{bn} (centi-Newton or cN) – may be set to vary from the starting point to the end point
- Touch Down Force, F_{TD} (cN)
- Loop Percent Height (%)
- Bond Time (milli-seconds or ms)

It is to be noted that apart from the quality and cleanliness of the bond pads, the bond quality and especially bonding failure rates are quite sensitive to the ambient conditions, viz., temperature and humidity. Since the 2S modules involve two different surfaces, i.e., the bond pads on top of the silicon sensors and the bond pads on the FE hybrids, the wire bonding parameters for these surfaces have been optimized separately.

[7]The unit is proportional to the ultrasonic frequency but not remain undisclosed by the manufacturer.

5.4.1 Wire Bonding on Silicon Sensor Bond Pads

The silicon sensors for the 2S modules have been manufactured by Hamamatsu Photonics K.K. (HPK), Japan, having embedded bond pads with a high level of cleanliness. To optimize the bond parameters for the sensor bond pads, the extra sets of bond pads at the middle of the 2S sensors (see Fig. 5.22) have been utilized. The optimization presented here is based on the variation of ultrasonic power, only while all other parameters are fixed e.g., $F_{bn} = 25$, $F_{TD} = 20$, and the Loop %Height=130. The wire deformation measurements and the destructive pull test results for these wire bondings on the sensor bond pads are summarized in Tab. 5.3, while the distribution of pull strength and the trends for the loop height (at the wire/bond break-point) are shown in Fig. 5.23. It is obvious that the best quality of bonding is achieved for the sensor bond pads with a USP of 50-55. Hence, for the sensor-hybrid bondings in the 2S dummy module (see Sec.5.5), the USP is fixed at 55 for the sensor bond pads along with other parameters ($F_{bn} = 25$, $F_{TD} = 20$, and the Loop %Height=130).

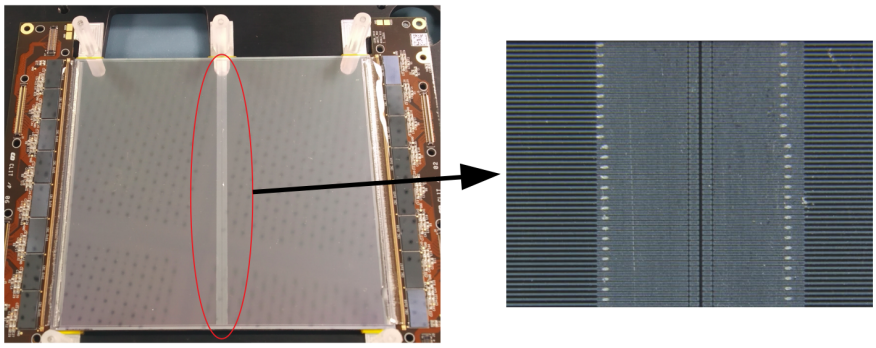


Figure 5.22: Middle area of a 2S module sensor containing the bond pads used for the bond parameter optimization.

US power	Deformation in microscope (μm)		Pull force	
	1^{st} bond	2^{nd} bond	Mean (gm)	Standard deviation (gm)
45	36	36	7.419	xx
50	41	41	6.988	xx
55	44	44	6.794	xx

Table 5.3: Bond quality check for the sensor-sensor bonding for different USP with $F_{bn} = 25$, $F_{TD} = 20$, and the Loop %Height=130.

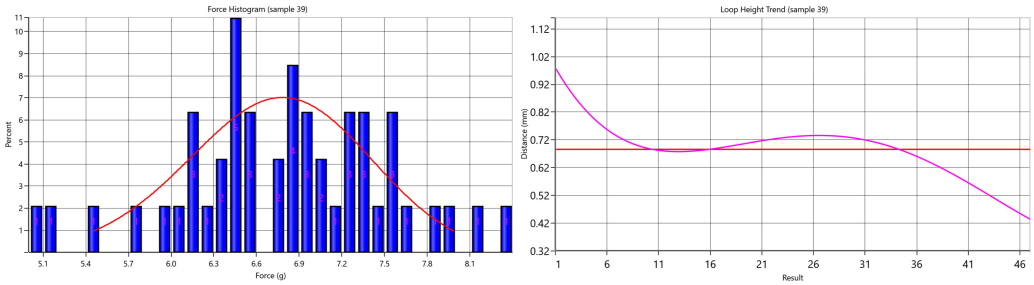


Figure 5.23: Pull force results for sensor-sensor wire bonding on silicon sensor bond pads where ultrasonic power is set at 55 units.

5.4.2 Wire Bonding on Gold-plated Bond Pads on PCB

Similar to the silicon sensor bond pads, the bonding parameters are to be optimized on the FE hybrid bond pads. But, there have not been any hybrid quality bond pads available with rows of the bond pads developed on top of the PCB. Therefore, the bond optimization features in terms of varying the bond parameters have been tested on the specially manufactured gold-plated PCB[8] In the following section, the general optimization procedure of bond parameters has been discussed, while the most important bond parameters, USP, F_{bn}, F_{TD}, and the loop %Height are varied, keeping the bond length constant at 4000 μm.

1. **USP** Optimization : All other parameters are kept fixed with the USP is varied

[8]The plates consist of a material called electroless nickel immersion gold which is specified to be a chemically applied layer of nickel of 3 up to 6 μm with a top coating of at least 50nm of gold. The PCB has been manufactured in Germany in collaboration with the RWTH Aachen University.

over a range of 60 to 80 units, in steps of 5 units. For each set of parameters the bond quality results are listed in the Tab. 5.4. Here are the values for other parameters are kept at $F_{bn} = 25$, $F_{TD} = 20$, and the Loop %Height=130.

USP	Deformation in microscope (μm)		Pull force	
	1^{st} bond	2^{nd} bond	Mean (gm)	Standard Deviation (gm)
60	44.17	42.00	6.23	0.59
65	43.00	45.00	5.97	0.37
70	45.00	48.00	6.08	0.32
75	47.00	48.00	5.69	0.36
80	50.00	51.00	5.06	0.27

Table 5.4: Bond test results for different USP on the gold-plated PCB with $F_{bn} = 25$, $F_{TD} = 20$, and the Loop %Height=130.

2.**Bond Force (F_{bn})** Optimization: Similar to US power optimization, keeping rest of the parameters are fixed, F_{bn} is varied from 22 to 30 with a step size of 2. The values for the rest of the parameters are kept fixed at USP=65, $F_{TD} = 20$, and the Loop %Height=130 and the corresponding deformation and pull test results are shown in Tab. 5.5.

F_{bn}	Deformation in microscope (μm)		Pull force	
	1^{st} bond	2^{nd} bond	Mean (gm)	Standard Deviation (gm)
22	46	46	6.025	0.461
24	47	48	5.541	0.409
26	45	45	5.562	0.357
28	45	46.5	5.612	0.357
30	46	47.5	5.872	0.615

Table 5.5: Bond test results for different F_{bn} on the gold-plated PCB with USP=65, F_{TD}=20, and the Loop %Height=130.

3. **Touch Down Force (F_{TD})** optimization: F_{TD} is varied between 18 and 24 with USP=65, F_{bn}=24, and Loop %Height = 130% and the corresponding pull test results are shown in Tab. 5.6.

4. **Loop % Height** optimization: Loop % Height is varied from 110% to 150% while

F_{TD}	Deformation in microscope (μm)		Pull force	
	1^{st} bond	2^{nd} bond	Mean (gm)	Standard Deviation (gm)
18	45	45	5.07	0.52
20	46	46	4.75	0.85
22	46.5	46.5	4.83	0.62
24	47	47	5.83	0.75

Table 5.6: Bond test results for different F_{TD} on the gold-plated PCB with USP=65, F_{bn}=24, and the Loop %Height=130.

USP=65, F_{bn}=24, and F_{TD}=20. For this set of bond parameters, the corresponding pull test results can be found in Tab. 5.7.

% Height	Deformation in microscope (μm)		Pull force	
	1^{st} bond	2^{nd} bond	Mean (gm)	Standard Deviation (gm)
120	45	45	4.32	0.41
130	45	45	6.81	0.40
140	45	45	7.68	0.36
150	45	45	6.10	1.10

Table 5.7: Bond test results in different Loop percent Height

On the gold-plated PCB, with a typical set of bond parameters, the bonding strength test results along with the wire deformation are shown in Fig. 5.24. Based on the above pull test results and trends, the bond parameters for wire bonding on the FE hybrid bond pads (of the 2S modules) are set, as discussed in the next subsection.

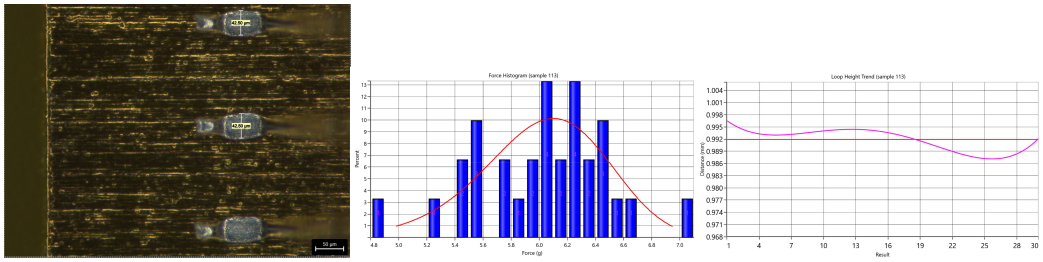


Figure 5.24: Typical Wire deformation view under the inspection microscope (left), the pull test results (center) and the loop height (right) with the bond parameters of USP=60, $F_{bn} = 25$, $F_{TD} = 20$, and the Loop %Height=130.

5.5 2S module bonding

With the optimized parameters from the previous two sub-sections, the wire bonding for the 2S modules consisting of a total 4064 wire bonding (2032 in each sensor), are further fine-tuned by analyzing the pull test results from dummy 2S modules[9]. It is to be noted here that in a 2S module, the strip pitch is 90 μm, but there are two types of wire bondings to be executed in terms of bond length as shown schematically in Fig. 5.25. For safer operation, small wires are bonded first, followed by longer ones. Furthermore, each 2S module requires the bonding to be executed on two sides, i.e., on the two sensors of the module and two different support fixtures/jigs required for these operations. Both the jigs (see Fig. 5.26) have been designed very meticulously to support the FE hybrids and the silicon sensors during the bonding operation, while Tab. 5.8 lists the required specifications to be achieved for the 2S modules.

Variable	Thresholds
Mean value of Pull strength	$\geq$ 8 gm
RMS value of Pull strength T	$\leq$ 10% of Mean
Break type T	$\leq$ 20% lift off

Table 5.8: Required specification for the 2S wire bonds.

[9]It is exactly identical to a fully functional 2S module in terms of all mechanical features (dimension, bond pads, etc.), except the fact that sensors there don't have proper crystal orientation.

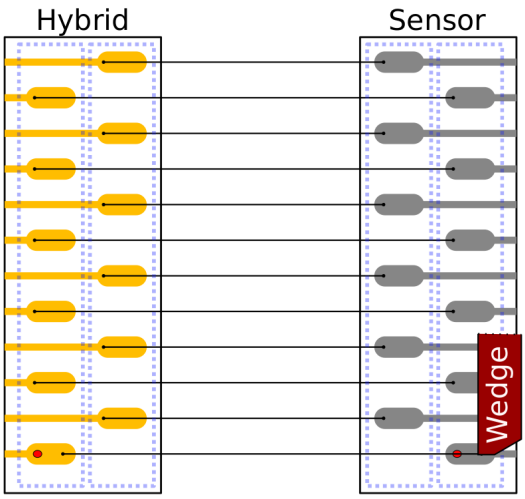


Figure 5.25: The schematics of long and short wire bonding connections in a 2S module: the wire bonding is generally executed from the FE hybrid bond pad to the sensor bond pad.

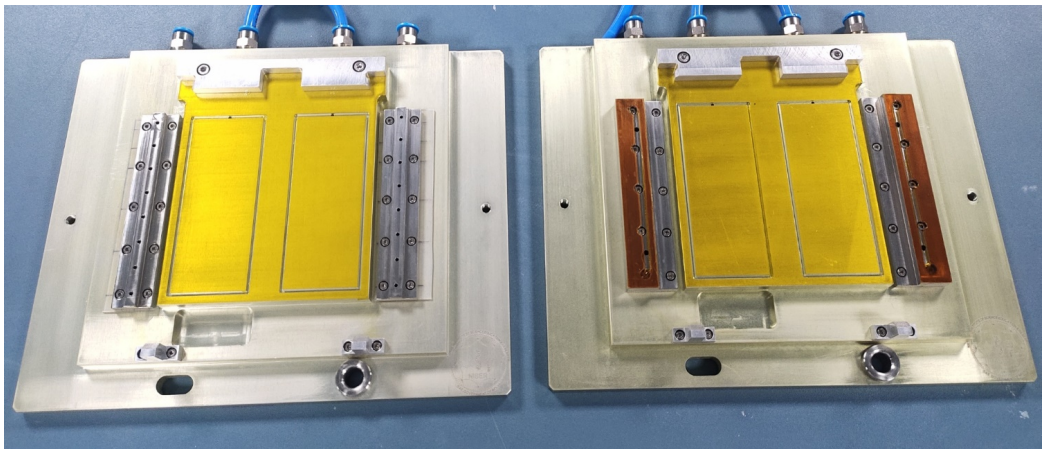


Figure 5.26: Wire bonding jig for the bottom sensor (left), and for the top sensor (right).

Due to the differences in the surface types of the silicon sensor bond pad and hybrid bond pads, the parameters for the first (FE hybrid) and second (sensor) bonding has to be different. For the sensor bond pads, they are driven by the parameters as concluded in Sec. 5.4.1. However, for the hybrid bond pads, the initial set of parameters are set from the results in Sec. 5.4.2 with further fine-tuning with dummy 2S module wire bonding. The

USP as usual varied in steps of 5 units on the FE hybrid bond pads between 50 and 80 units, and the USP on the sensor bond pads is kept at 55 units with all other parameters fixed. The results of the wire deformation and pull test results (including the correction factors using the Eqn. 5.3) are shown in Tab. 5.9. The distribution of the destructive pull test results, along with the bond breakage point loop height trend are shown in Fig. 5.27. In addition, a closeup view of the wire bondings on one of the dummy 2S modules (built at MPC@NISER) through high resolution microscope is shown in Fig. 5.28.

US power	Pull force		Absolute Loop Height (H) (μm)	Corrected Force (gm)
	Mean (gm)	Sigma (gm)		
50	7.699	1.28	717	8.926
55	7.826	0.644	717	9.073
60	7.619	0.570	748	8.536
65	6.779	0.35	769	7.429
70	6.221	0.30	739	7.492
75	5.913	0.47	712	6.895
80	5.611	0.36	799	5.967

Table 5.9: Pull test results for the sensor-hybrid wire bond on a 2S dummy module where USP on the FE hybrid bond pad is varied.

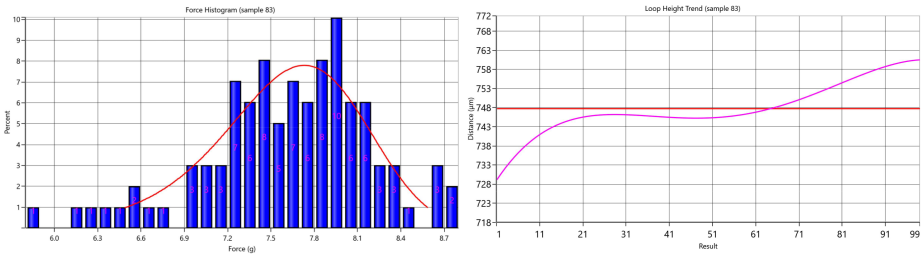


Figure 5.27: Distribution of the destructive pull test strength and loop height with optimized USP settings of 55 and 60 units on sensor and FE hybrid bond pads. Although the results for the FE bond pads with the usage of USP 55 and 60 satisfy the required specifications as mentioned in Tab. 5.8, the USP 60 has been considered for the functional module bonding due to its lesser standard deviation.

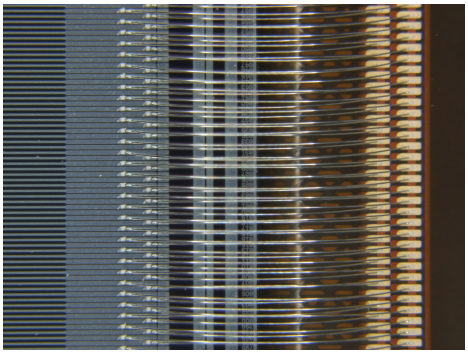


Figure 5.28: A close up view of the wire bondings on a dummy 2S module assembled at MPC@NISER.

It is to be noted here that the sensor backplane and the HV connectors are also required to be wire bonded for each of the 2S sensors before getting them glued together into a module. Here, the bond parameters are also kept identical to the optimized values of USP for the sensor backplane and the bond pad on the HV tail, i.e., 55 and 60 units, respectively. A closeup view of the wire bonds for the HV tail, along with the jig to hold the sensor for this particular wire bonding operation, are shown in Fig. 5.29. Finally, the picture of a functional 2S module as assembled at the MPC@NISER is presented in Fig. 5.30. On the functional module, both types of wire bondings (sensor-FE hybrid and sensor-HV tail) operations have been carried out using the optimized bond parameters as summarized in Tab. 5.10.

Bond Parameters	Optimized values	
	Sensor bond pad	FE hybrid bond pad
Ultrasonic Power	55	60
Touch Down Force (F_{TD})	20	20
Bond Force (F_{bn})	24	24
Percentage Loop Height	130	130

Table 5.10: Final set of optimized bond parameters used for wire bonding operation on the functional module.

Figure 5.29: A close up view of the HV tail wire bondings on a 2S sensor as performed at MPC@NISER (left); HV tail wire bonding jig to support the 2S sensor (right).

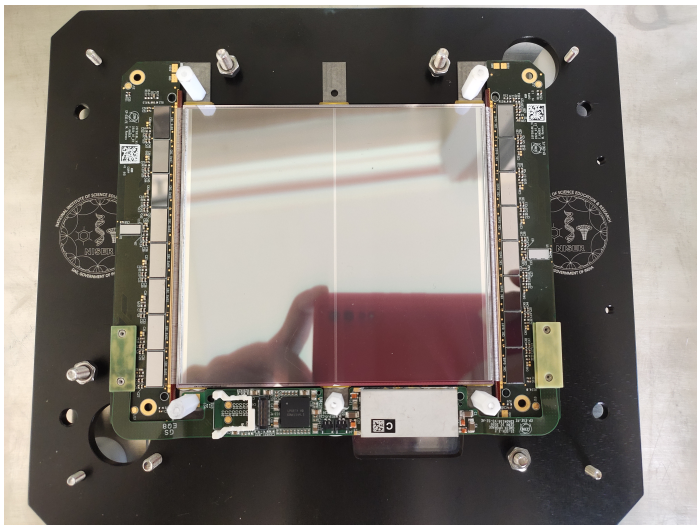


Figure 5.30: Fully wire bonded functional 2S module prepared at NISER.

Chapter 6

Summary and Conclusions

Top pair production in association with photon is an interesting physics process with a sizeable cross section at the LHC center-of-mass energy of 13 TeV. A comprehensive cross section measurement for the pp $\rightarrow$ $t\bar{t}\gamma$ process in the semileptonic channel using 137 fb^{-1} of CMS dataset has been presented in this thesis. The semileptonic events filtered through the electron/muon triggers, with additional offline requirements of a high p_T and well isolated lepton (electron or muon), with $\geq$3 jets and $\geq$1 b-tagged jets, and exactly one high p_T photon. The contributions from various background processes viz., inclusive top production processes (pair-production and single top), W/Z+γ, Drell-Yan, QCD multijets, etc. have been meticulously estimated through data-driven and Monte Carlo (MC) simulated events. The events are categorized based on the final state photons originating from genuine, misidentified, hadronic, and fake/pileup category photons. The residual differences between the data and Monte Carlo modeling are corrected following the routine recommendations of CMS Top Quark Physics Group, while the data-driven background estimation have been validated on different collisions datasets, defined through orthogonal object selection and/or kinematic phase-space. Finally, the results are extracted in terms of inclusive and differential cross sections in a fiducial region with photon transverse momentum $p_T(\gamma) > 20$ GeV with three or more jets. The measured inclusive cross section of 798 ± 7 (stat) $\pm$ 48 (syst) fb is in good agreement with the values predicted by the QCD calculations corrected at NLO.

Furthermore, as illustrated in the Chapter 3, the presence of a additional photon in the $t\bar{t}$ final state can plausibly enhance the $q\bar{q} \rightarrow t\bar{t}$ production modes, otherwise swamped by the $gg \rightarrow t\bar{t}$ modes at 13 TeV. Therefore, an interpretation on the $t\bar{t}$ charge asymmetry (A_C) at the reconstruction level has been performed using the same events as in the cross section measurement. The kinematics of the top and antitop quarks have been reconstructed and the differential A_C measurements are presented here as functions of invariant mass of top pair system ($M_{t\bar{t}}$), transverse momentum of top pair system ($p_T^{t\bar{t}}$), and rapidity of the photon ($|y_\gamma|$).

This thesis also includes the works towards the performance studies related to CMS trigger system as described in the Chapter 4. In particular, during the LHC Run 2 with the increasing numbers of pile up events, it had been quite crucial to monitor and adjust the bandwidth usage by different trigger objects at L1. In addition, a detailed MC efficiency study has been presented here for the modified e/γ HLT triggers, presently deployed during the LHC Run 3.

A significant amount of work presented in this thesis consists of the detector hardware activities related to the CMS Outer Tracker (OT) Upgrade aimed for the HL-LHC. A brief overview of the high-granularity strip-strip (2S) module along with the assembly works being carried out at NISER are covered in Chapter 5. Here, establishment of wire connections through an automatic ultrasonic wire bonder have always been a tough task as they involve two sets of dissimilar surfaces. In addition, lack of technical expertise within the country on the operational aspects of such a state-of-the-art machine had posed the extra challenges. However, through online consultation with the bond experts (F&K Delvotec engineers and physicists at CERN and RWTH Aachen University) various machine parameters have been optimized for the dissimilar bond pad surfaces of the 2S modules. In the process, sev-

eral bond parameters for ultrasonic wire bonding, viz., ultrasonic power, touchdown force, bond force, wire loop height, bonding time etc. have been tuned to result in best quality wire bonding required for the CMS Phase 2 OT modules. Successful quality assurance checks through wire pull test and the wire deformation measurements have been performed on the prototype of the 2S modules (assembled at NISER) to tally with the stringent wire bonding requirements as set by the CMS Phase 2 OT Modules Working Group. Finally, with the optimized bond parameters as presented in this thesis a couple of functional modules have been wire bonded successfully at NISER.

Bibliography

[1] Sheldon L. Glashow, "Partial-symmetries of weak interactions," *Nuclear Physics*, vol. 22, no. 4, pp. 579–588, 1961.

[2] Steven Weinberg, "A Model of Leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 1967.

[3] Abdus Salam, "Weak and Electromagnetic Interactions," *Proceedings of the Nobel Symposium Held at Lerum*, vol. 680519, pp. 367–377, 1968.

[4] J.J. Aubert et al. , "The ratio of the nucleon structure functions F2N for iron and deuterium," *Physics Letters B*, vol. 123, no. 3, pp. 275–278, 1983.

[5] G. Arnison et al., "Experimental observation of lepton pairs of invariant mass around 95 GeV/c2 at the CERN SPS collider," *Physics Letters B*, vol. 126, no. 5, pp. 398–410, 1983.

[6] 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*, vol. 716, no. 1, pp. 1–29, 2012.

[7] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Physics Letters B*, vol. 716, no. 1, pp. 30–61, 2012.

[8] Peter W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, 1964.

[9] CDF Collaboration, "The CDF-II detector: Technical design report," tech. rep., 1996.

[10] CDF Collaboration, "Observation of top quark production in $\bar{p}p$ collisions," *Phys. Rev. Lett.*, vol. 74, pp. 2626–2631, 1995.

[11] DØ Collaboration, "The DO Detector," *FERMILAB-PUB-93-179-E, D0-1808*, 1993.

[12] DØ Collaboration , "Observation of the Top Quark," *Phys. Rev. Lett.*, vol. 74, pp. 2632–2637, Apr 1995.

[13] ATLAS, CDF, CMS, and DØ Collaboration, "First combination of Tevatron and LHC measurements of the top-quark mass," 3 2014.

[14] Collins, John C. and Soper, Davison E. and Sterman, George F., "Factorization of Hard Processes in QCD," *Adv. Ser. Direct. High Energy Phys.*, vol. 5, pp. 1–91, 1989.

[15] NNPDF Collaboration, "NNPDF collaboration homepage," `http://nnpdf.mi.infn.it`.

[16] Michal Czakon et. al., "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$," *Phys. Rev. Lett.*, vol. 110, p. 252004, 2013.

[17] CMS Collaboration, "Measurement of differential $t\bar{t}$ production cross sections in the full kinematic range using lepton + jets events from proton-proton collisions at $\sqrt{s} = 13$ TeV," *Phys. Rev. D*, vol. 104, p. 092013, Nov 2021.

[18] "Top quark cross-section results by top working group," `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWG`.

[19] CDF Collaboration, "Evidence for $t\bar{t}\gamma$ production and measurement of $\sigma_{t\bar{t}\gamma}/\sigma_{t\bar{t}}$," *Phys. Rev. D*, vol. 84, p. 031104, Aug 2011.

[20] ATLAS Collaboration, "Observation of top-quark pair production in association with a photon and measurement of the $t\bar{t}\gamma$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector," *Phys. Rev. D*, vol. 91, p. 072007, Apr 2015.

[21] ATLAS Collaboration, "Measurement of the $t\bar{t}\gamma$ production cross section in proton-proton collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector," *Journal of High Energy Physics*, vol. 2017, nov 2017.

[22] CMS Collaboration, "Measurement of the semileptonic $t\bar{t}\gamma$ production cross section in pp collisions at $\sqrt{s}$ = 8 TeV," *Journal of High Energy Physics*, vol. 2017, oct 2017.

[23] ATLAS Collaboration, "Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}\gamma$ production in leptonic final states at $\sqrt{s}$ = 13 TeV in ATLAS," *The European Physical Journal C*, vol. 79, may 2019.

[24] CMS collaboration, "Technical proposal for the upgrade of the CMS detector through 2020," 2011.

[25] Lyndon Evans and Philip Bryant, "Lhc machine," *Journal of Instrumentation*, vol. 3, p. S08001, 2008.

[26] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider," *Journal of Instrumentation*, vol. 3, p. 31, aug 2008.

[27] CMS Collaboration, "The CMS Experiment at the CERN LHC," *Journal of Instrumentation*, vol. 3, no. S08004, p. 31, 2008.

[28] L. Collaboration, "The LHCb Detector at the LHC," *Journal of Instrumentation*, vol. 3, p. S08005, 2008.

[29] ALICE Collaboration , "The ALICE experiment at the CERN LHC," *Journal of Instrumentation*, vol. 3, p. 31, aug 2008.

[30] Adriani, O. and others, "The LHCf detector at the CERN Large Hadron Collider," *Journal of Instrumentation*, vol. 3, p. S08006, 2008.

[31] Pinfold, James and others, "Technical Design Report of the MoEDAL Experiment," *CERN-LHCC-2009-006, MoEDAL-TDR-001*, 6 2009.

[32] Berardi, V. and others, "TOTEM: Technical design report. Total cross section, elastic scattering and diffraction dissociation at the Large Hadron Collider at CERN," *CERN-LHCC-2004-002, TOTEM-TDR-001*, 1 2004.

[33] "LHC Performance Workshop," `https://indico.cern.ch/event/1097716/`, Jan 22-27 2022.

[34] A. Affolder, "The CMS Silicon Strip Tracker: Design and Production Status," *Nuclear Physics B - Proceedings Supplements*, vol. 150, pp. 118–123, 2006. Proceedings of the 9th Topical Seminar on Innovative Particle and Radiation Detectors.

[35] CMS Collaboration, "The CMS tracker system project: Technical Design Report," 1997.

[36] Clerbaux, Barbara and Mahmoud, T. and Marage, Pierre and Elgammal, Sherif, "Saturation and Energy Corrections for TeV Electrons and Photons," 12 2006.

[37] The CMS collaboration , "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC," *Journal of Instrumentation*, vol. 16, p. P05014, may 2021.

[38] Francesca Cavallari, "Performance of calorimeters at the LHC," *Journal of Physics: Conference Series*, vol. 293, p. 012001, apr 2011.

[39] CMS Collaboration, "CMS Muon system reference," `https://twiki.cern.ch/twiki/bin/view/CMS/MuonWikiHome`.

[40] CMS Collaboration, "CMS detector website reference for muon detection," `https://cms.cern/detector/detecting-muons/resistive-plate-chambers`.

[41] Perrotta, Andrea, "CMS event reconstruction status in Run 2," *The European Physical Journal C*, vol. 214, p. 02015, 2019.

[42] Matteo Cacciari and Gavin P. Salam and Gregory Soyez, "The anti-kt jet clustering algorithm," *Journal of High Energy Physics*, vol. 2008, p. 063, apr 2008.

[43] D. Bertolini, P. Harris, M. Low and N. Tran, "Pileup Per Particle Identification," *Journal of High Energy Physics*, vol. 10, p. 059, 2014.

[44] CMS Collaboration, "Trigger application as recommended by the Top Physics group," `https://twiki.cern.ch/twiki/bin/view/CMS/TopTrigger`, 2016-2018.

[45] CMS Collaboration, "Cut-based electron ID recommendations for Run II," `https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2`, 2016-2018.

[46] CMS Collaboration, "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC," *Journal of Instrumentation*, vol. 16, p. P05014, may 2021.

[47] CMS Collaboration, "Cut-based muon ID recommendations for Run II," `https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2`, 2016-2018.

[48] CMS Collaboration, "Cut-based photon ID recommendations for Run II," `https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedPhotonIdentificationRun2#Cut_Ba-sed_Photon_ID_for_Run_2`, 2016-2018.

[49] CMS Collaboration, "Methods to apply BTagging efficiency scale factors," `https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMethods#1a_Event_reweight-ing_using_scale`, 2016-2018.

[50] CMS Collaboration, "Utilities for accessing pile-up information for data," `https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupJSONFileforData`, May 2021.

[51] Khachatryan, Vardan and others, "The CMS trigger system," *Journal of Instrumentation*, vol. 12, no. 01, p. P01020, 2017.

[52] CMS Collaboration, "Measurement of the inclusive and differential $\mathrm{t\bar{t}}+\gamma$ cross section and EFT interpretation in the single lepton channel at $\sqrt{s} = 13$ TeV," *CMS-PAS-TOP-18-010*, 2021.

[53] CMS Collaboration, "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV," *Journal of Instrumentation*, vol. 15, no. 10, p. P10017, 2020.

[54] CMS Collaboration, "Search for disappearing tracks in proton-proton collisions at $\sqrt{s}$=13TeV," *Physics Letters B*, vol. 806, p. 135502, 2020.

[55] CMS Collaboration, "102X/94X/80X photon efficiency scalefactors," `https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun2#Photon_efficien-cies_and_scale_fa`, 2016-2018.

[56] CMS Collaboration, "102X/94X/80X electron efficiency scalefactors," `https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun2#Electron_efficien-cies_and_scale`, 2016-2018.

[57] CMS Collaboration, "Energy scale and smear corrections for electrons and photons," `https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun2#Electron_efficien-cies_and_scale`, May 2021.

[58] CMS Collaboration, "Muon reference efficiencies for 2016," `https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2016LegacyRereco`, 2016.

[59] CMS Collaboration, "Muon reference efficiencies for 2018," `https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2018`, 2018.

[60] A. C. Maestro, S. Folgueras, J. R. González et al., "Lepton efficiency uncertainties in TOP analyses,CMS Analysis Note CMS AN-2018/210," *CERN LHC*, 2018.

[61] CMS Collaboration, "Measurement of the semileptonic $t\bar{t}+\gamma$ production cross section in pp collisions at $\sqrt{s} = 8$ TeV," *Journal of High Eergy physics*, vol. 10, p. 006, 2017.

[62] CMS Collaboration, "Missing transverse energy optional filters for Run II," `https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2#MiniAO-D_805`.

[63] CMS Collaboration, "documentation of the roostats -based statistics tools for higgs pag," `http://cms-analysis.github.io/higgsanalysis-combinedlimit/`.

[64] CMS Collaboration, "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV," *Journal of High Energy Physics*, vol. 07, p. 161, 2018.

[65] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV," *Journal of Instrumentation*, vol. 10, no. 06, p. P06005, 2015.

[66] CMS Collaboration, "Performance of CMS Muon Reconstruction in pp Collision Events at $\sqrt{s} = 7$ TeV," *Journal of Instrumentation*, vol. 7, p. P10002, 2012.

[67] CMS Collaboration, "CMS Top EGamma Coordination(Run2)," `https://twiki.cern.ch/twiki/bin/view/CMS/TopEGM`, 2020.

[68] CMS Collaboration, "CMS Top Muon Information for Analysis (Run2)," `https://twiki.cern.ch/twiki/bin/view/CMS/TopMUO`, 2020.

[69] CMS Collaboration, "CMS TOP JetMET Analysis," `https://twiki.cern.ch/twiki/bin/view/CMS/TopJME`, 2016-2018.

[70] ATLAS and CMS Collaboration, "Combination of inclusive and differential $t\bar{t}$ charge asymmetry measurements using ATLAS and CMS data at $\sqrt{s} = 7$ and 8 TeV," *Journal of High Energy Physics*, vol. 04, p. 033, 2018.

[71] J. Bergnera, M. Schulze, "The top quark charge asymmetry in $t\bar{t}\gamma$ production at the LHC," *The European Physical Journal C*, p. 79:189, 2019.

[72] Manfred Jeitler, "Trigger Concepts at HL-LHC," *11th Terascale Detector Workshop 2018*.

[73] CMS Collaboration, "Commissioning of CMS HLT," *Journal of Instrumentation*, vol. 5, p. T03005, 03 2010.

[74] CMS Collaboration, "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV," *Journal of Instrumentation*, vol. 15, no. 10, p. P10017, 2020.

[75] CMS Collaboration, "Cms luminosity public results," `https://cmslumi.web.cern.ch/publicplots/multiYear/peak_lumi_pp.png`.

[76] CMS Collaboration, "CMS Technical Design Report for the Level-1 Trigger Upgrade," *CERN-LHCC-2013-011, CMS-TDR-12, CMS-TDR-012*, June 2013.

[77] CMS Collaboration, "Technical Proposal for the Upgrade of the CMS Detector through 2020," *CERN-LHCC-2011-006, LHCC-P-004, CMS-UG-TP-1*, June 2011.

[78] CMS Collaboration, "Technical Proposal for the Phase-II Upgrade of the CMS Detector," *CERN-LHCC-2015-010, LHCC-P-008, CMS-TDR-15-02*, June 2015.

[79] CMS Collaboration, "A MIP Timing Detector for the CMS Phase-2 Upgrade," *CERN-LHCC-2019-003, CMS-TDR-020*, 2019.

[80] CMS Collaboration, "CMS pileup public results," `https://cmslumi.web.cern.ch/publicplots/pileup_allYears.png`.

[81] CMS Collaboration, "The Phase-2 Upgrade of the CMS Tracker," *CERN-LHCC-2017-009, CMS-TDR-014*, June 2017.

[82] Foudas, C. and Rose, A. and Jones, J. and Hall, G., "A Study for a tracking trigger at first level for CMS at LHC," in *11th Workshop on Electronics for LHC and Future Experiments (LECC 2005)*, p. 90, 10 2005.

[83] R. Schemmel, T. Hemsel, C. Dymel, M. Hunstig, M. Brökelmann, and W. Sextro, "Using complex multi-dimensional vibration trajectories in ultrasonic bonding and welding," *Sensors and Actuators A: Physical*, vol. 295, pp. 653–662, 2019.

[84] F&K Delvotec, "F&K Delvotec Manufacturer's webpage," `https://www.fkdelvotec.com/en/`.

[85] Micro Point Pro - Tools, "Manufacturer's webpage," `https://mpptools.com`.

[86] F&K Delvotec Bondtechnik, "Bondprozesskontrolle," `https://www.fkdelvotec.com/fileadmin/media_fuk/dokumente/Bondprozesskontrolle.pdf`, August 2017.