January 2021

Measurement Of The Cross Section Of Top Quark Pairs Produced In Association With A Photon In Lepton + Jets Events At √s = 13 Tev With Full Runii Cms Data

Nabin Poudyal
Wayne State University

MEASUREMENT OF THE CROSS SECTION OF TOP QUARK PAIRS PRODUCED IN ASSOCIATION WITH A PHOTON IN LEPTON + JETS EVENTS AT $\sqrt{s}$ = 13 TEV WITH FULL RUNII CMS DATA

by

NABIN POUDYAL

Ph.D DISSERTATION

Submitted to the Graduate School,

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2021

MAJOR: PHYSICS

Approved By:

_____

Advisor Date

_____

_____

_____

DEDICATION

THIS DISSERTATION IS DEDICATED TO MY PARENTS.

For their endless love, support and encouragement.

ACKNOWLEDGEMENTS

I would like to extend my deepest gratitude to my supervisor, Dr. Robert Harr, for your patience, guidance, and support. I have benefited greatly from your wealth of knowledge and meticulous editing. I am also grateful to my committee members, Dr. Paul Karchin, Dr. David Cinabro, and Dr. Sarah Brownlee. I am especially grateful to my research mentor Dr. Daniel Noonan. Without his support and guidance, it would not be possible. I would like to thank Physics Department, Wayne State Graduate school, Fermilab, and University Research Associate for their financial support for my graduate study. Special thanks to Dr. Lorenzo Uplegger and Dr. Anadi Canepa for their guidance and support during the CMS Tracker upgrade program at Fermilab. I would like to thank Aloke Das and Lukas Lechner who helped me during my research. Special thanks to my colleagues, classmates, friends, and family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION AND MOTIVATION

1.1 Standard Model of Particle Physics

The standard model of particle physics (SM) is a framework that organizes and synthesizes the results of particle physics experiments into a concise prescription for producing quantitative predictions. The framework is based on quantum field theory and symmetry groups. The initial form of the SM was the work of Glashow, Salam, and Weinberg [1]. Other theorists provided important support to the theory and experimental provided evidence for additional particles that added to the framework (c [2, 3], τ [4], b [5] and t [6, 7]) or filled in predictions ($W^{\pm}$ [8], Z^0 [9], and, H [10]). The latter provided much of the critical support for the SM over competing proposals. The SM can be viewed a set of fundamental particles, their properties (mass, spin, and couplings), and the rules for interactions among them. The particles of the SM are summarized in Figure 1. They come in two main groups: bosons, particles of integer spin; and fermions, particles of half integer spin.

Fermions are divided into quarks and leptons of 3 generations as shown in Figure 1. In the SM, fermions interact among one another by exchanging bosons. There are 4 types of gauge bosons with spin 1, hence called vector bosons; g, γ, Z^0, and $W^{\pm}$. The electromagnetic, weak, and strong interactions are mediated by γ, Z^0 and $W^{\pm}$, and g, respectively. Furthermore, $W^{\pm}$ mediate the weak interaction by acting on left handed particles and right handed antiparticles. The massive W boson is very short lived with a mean lifetime of about 3×10^{-25} s [11]. Gluons mediates the strong force that confine the quarks inside mesons or baryons. However, in very high energy interaction, quarks and gluons interact

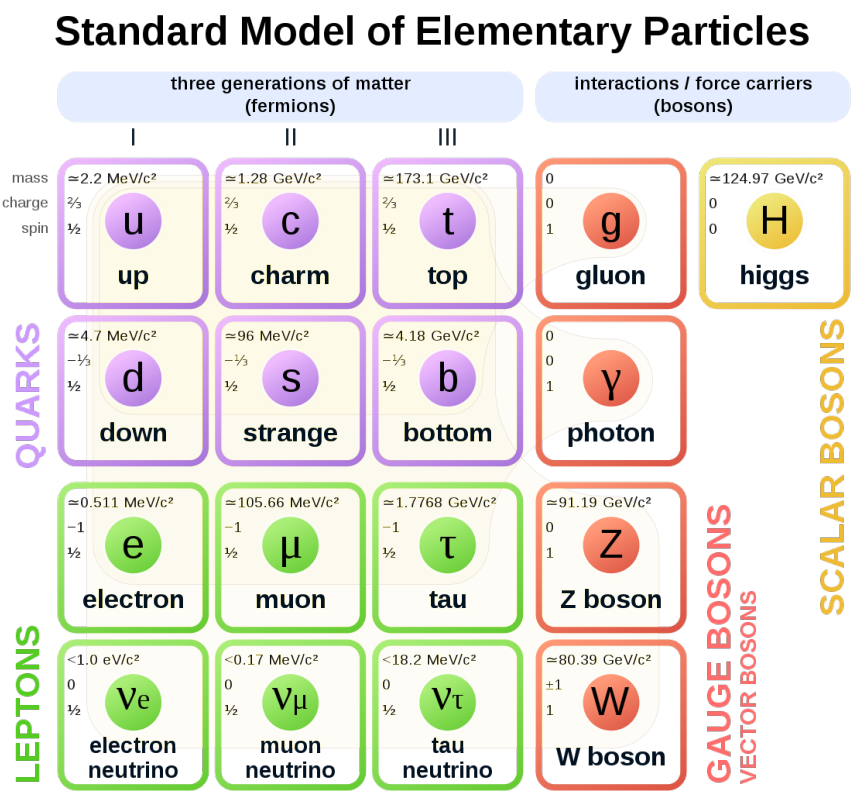


Figure 1: Fundamental elementary particles with their masses, charges, spin values [12].

weakly and form a guark-gluon plasma. This phenomenon is known as the asymptotic freedom of the strong interaction. In addition to the force carrying particles, we have six quarks (u, d, c, s, t, b) and six leptons (ν_e, ν_μ, ν_τ, e, μ τ). Each quark or lepton has a corresponding antiquark or anti-lepton respectively. All the ordinary matter and antimatter in the universe can be described as the combination of quarks, anti-quarks, leptons and anti-leptons. The quarks carry color charge, electric charge, and weak isospin, so they interact via the strong electromagnetic and weak interactions. The three leptons e, μ and τ are charged and interact electromagnetically and weakly whereas ν_e, ν_μ and ν_τ are neutral and hence interact via the weak force only. Figure 1 shows the organization of 3 generations of fundamental particles with their principle properties. The interaction between

fermions and a photon can be described by the following electromagnetic Lagrangian

$$\mathcal{L}_{\text{QED}} = \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{photon}} + \mathcal{L}_{\text{interaction}} \tag{1.1}$$

$$= \bar{\psi}(i\partial\!\!\!/ + m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - q\bar{\psi}\gamma^{\mu}A_{\mu}\psi. \tag{1.2}$$

The first term describes the kinematics of the spin-$\frac{1}{2}$ fermion field represented by ψ and $\bar{\psi}$. The second term describes the kinematics of the photon field A_{μ}, where $F_{\mu\nu}$ is the electromagnetic field tensor defined as $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$. The third term describes the interaction between the photon and fermion fields.

The $\mathrm{t\bar{t}}\gamma$ vertex probes a coupling between photons and top quarks. The lepton interacts via the weak interaction and are divided into left-handed doublet $\binom{\nu_l}{l^-}_L$ and right-handed singlet l_R^-. The left-handed doublet carry weak isospin. The doublet has eigenvalues $I_3 = -\frac{1}{2}$ for charged lepton and $I_3 = \frac{1}{2}$ for neutral leptons respectively. However the right-handed singlet has eigenvalues $I = 0$. All the neutrinos are massless in the SM. However, the existence of neutrino oscillations from one flavor to another indicates that neutrinos are not massless [13]. The quarks interact via the electromagnetic, weak and strong forces. These are also divided into doublet and singlet states. In addition, the cross-family coupling between the quarks exists and is representd by the CKM matrix [14]. The mass eigenstates of quarks are given by $\binom{q_{\mathrm{u}}}{q_{\mathrm{d}}'}_L$, $q_{\mathrm{u},R}$, and $q_{\mathrm{d},R}$ where the left-handed doublets the up quarks have $I_3 = \frac{1}{2}$ and down quarks have $I_3 = -\frac{1}{2}$. Similarly, right-handed singlets have weak isospin as $I = 0$.

1.2 The Top Quark

The discovery of the bottom quark (b) at Fermilab in 1977 [5] motivated the search for its weak isospin partner, the top quark (t), as needed by the SM. The top quark was discovered by the CDF and D0 experiments [15] at the Tevatron at Fermilab in 1995. The top quark is the heaviest elementary particle. It is the only particle that decays into a real W boson and b quark. It has a very short life time insufficient for it to hadronize. The study of the top quark offers an opportunity to understand and test the perturbative and non-perturbative strong interaction. The Yukawa couplings of the top quark to the Higgs boson is of order unity, and hence the top quark has a special role in testing the Higgs sector. The precise measurement of top quark properties probes interactions at the electroweak scale, symmetry breaking, and the physics beyond the SM. The physics of the top quark is an interesting area for validation of the SM.

1.2.1 Production, Decay and Properties of the Top Quark in Brief

Top quarks are predominantly created in pairs at the LHC via the strong interaction. The LHC is sometimes called a top quark factory [16]. The top pairs are produced via gluon-gluon fusion ($gg \rightarrow g \rightarrow \mathrm{t\bar{t}}$) as shown in Figure 4 (left) and quark-antiquark annihilation ($q\bar{q} \rightarrow g \rightarrow \mathrm{t\bar{t}}$) as shown in Figure 4 (right) in the pp collsion. Almost 90% of the $\mathrm{t\bar{t}}$ are produced via gluon-gluon fusion at $\sqrt{s} = 13$ TeVcollision energy. The cross section of $\mathrm{t\bar{t}}$ production can be given as a function of center of mass energy ($\sqrt{s}$) and top mass (m_t)

as

$$\sigma^{pp\to t\bar{t}}(\sqrt{s}, m_t) = \sum_{i,j=q,g} \int dx_i dx_j f_i(x_i, \mu_f^2) f_i(x_j, \mu_f^2) \cdot \hat{\sigma}^{ij\to t\bar{t}}(m_t, \sqrt{\hat{s}}, x_i, x_j, \alpha_s(\mu_R^2), \mu_R^2), \tag{1.3}$$

where $f_i(x_i, \mu_f^2)$ are the parton distribution functions (PDF) for parton i and (x_i) estimates the fraction of the proton momentum. The PDF functions are calculated using a factorization scale μ_f which separates non-perturbative from perturbative QCD. These are experimentally determined as explained in [17]. The perturbative QCD term, $\hat{\sigma}^{ij\to t\bar{t}}$, expresses the cross section as a function of the center of mass energy, the top-quark mass, coupling constant α_s, and renormalisation scale μ_R.

The SM top quark decays into a W-boson and b quark essentially $100\,\%$ of the time; the decay into s and d quarks are strongly suppressed by the CKM factors. The W-boson decays leptonically into a charged lepton and the corresponding neutrino around $33\,\%$ of the time. It decays hadronically to quarks the remaining $76\,\%$ of the time. Based on the types of W decays, the top pair events are divided into 3 categories; dileptonic, hadronic and semileptonic. When both of the W-bosons decay leptonically, the top decay is categorized as dileptonic as in $t\bar{t} \to W^+bW^-\bar{b} \to l^+\bar{\nu}_l\bar{b}l^-\nu_l b$. This occurs $9\,\%$ of the time. When both of the W-bosons decay into quark anti-quark pairs, it is categorized as a hadronic decay as in $t\bar{t} \to W^+bW^-\bar{b} \to q\bar{q}q'\bar{q}'\bar{b}b$. This occurs $46\,\%$ of the time. When one of the W-boson decays leptonically and the other decays hadronically, the event is categorized as semileptonic as in $t\bar{t} \to W^+bW^-\bar{b} \to l^+\bar{\nu}_l q\bar{q}\bar{b}b$. This occurs $45\,\%$ of the time. A breakdown of the branching ratios of $t\bar{t}$ decays is shown in Figure 2. Figure 3 shows a schematic diagram of

the semi-leptonic decay of a top pair at the LHC.

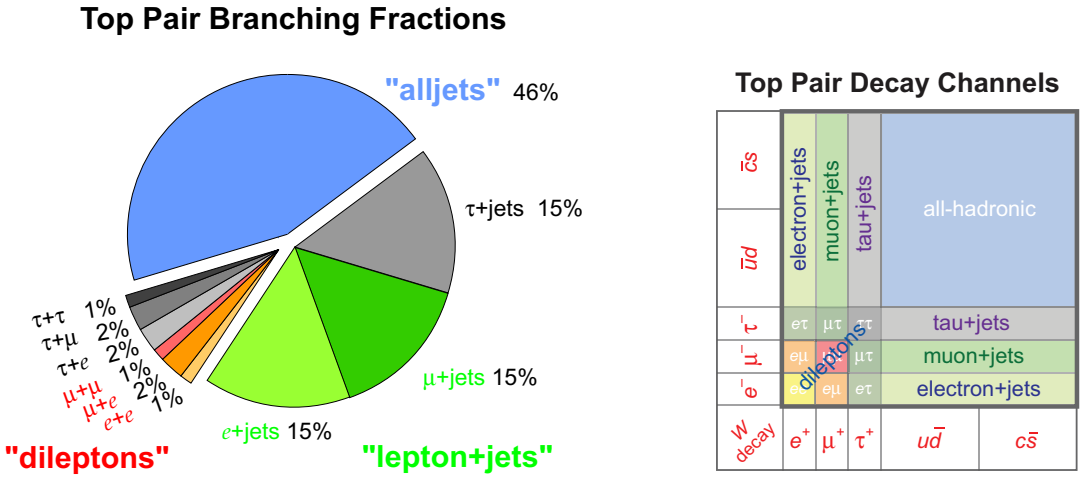


Figure 2: Branching ratio (left) and decay channels (right) for $t\bar{t}$ decay [18].

1.3 $t\bar{t}\gamma$ Production at the LHC

A cross section measurement of a top quark pair produced with a prompt photon probes the electromagnetic coupling of the top quark, $t\bar{t}\gamma$vertex. The SM Lagrangian of interaction between a top quark and photon is $-eQ_t t\bar{t}\gamma_\mu A_\mu$. The Lagrangian for the interaction of a top quark with a photon is given by

$$\mathcal{L}_{t\bar{t}\gamma} = -eQ_t t\bar{t}\gamma_\mu A_\mu - e\bar{t}\frac{-i\sigma^{\mu\nu}q_\nu}{m_t}(d_V^\gamma + id_A^\gamma\gamma_5)tA_\mu \tag{1.4}$$

where, e is the proton charge, m_t is the top quark mass, Q_t is the top quark electric charge, d_V^γ is the electric dipole moment, d_A^γ is the magnetic dipole moment, the first term is tree-level contribution in the SM, and the second term is the contribution from the first order loop correction.

Any anomalous top quark coupling can modify the cross section and affect the P_T spectrum

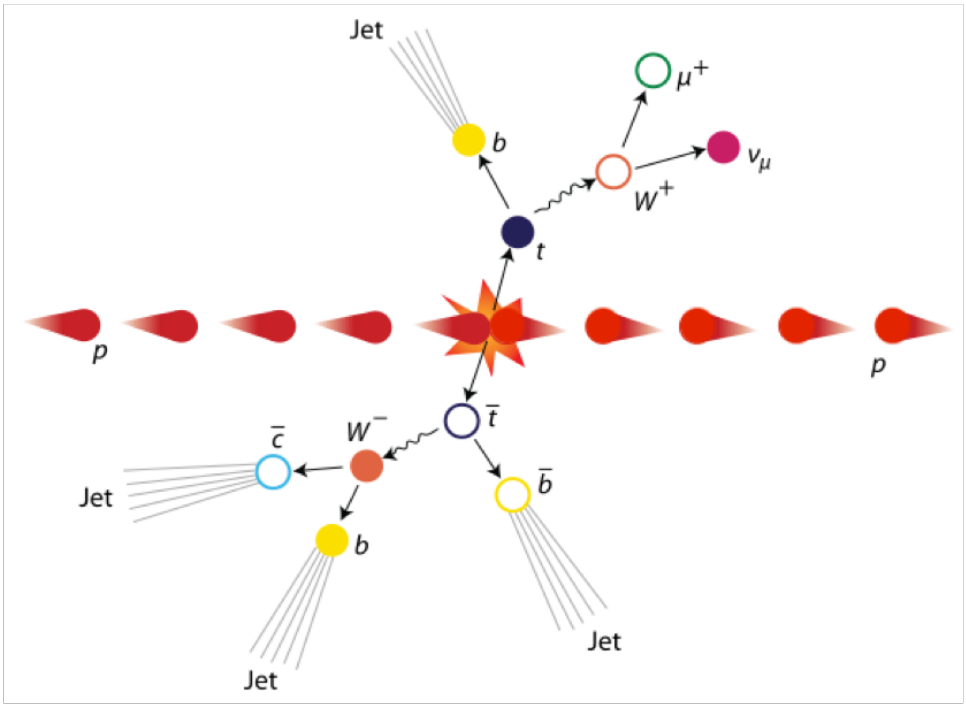


Figure 3: A schematic diagram of semi-leptonic decay of a $t\bar{t}$ event in pp collision at the LHC [19].

of the photon [20]. The photon in the $t\bar{t}\gamma$ signal can originate from an incoming quark, directly from one of the top quarks, or from a decay product of the top quarks. In $q\bar{q}$

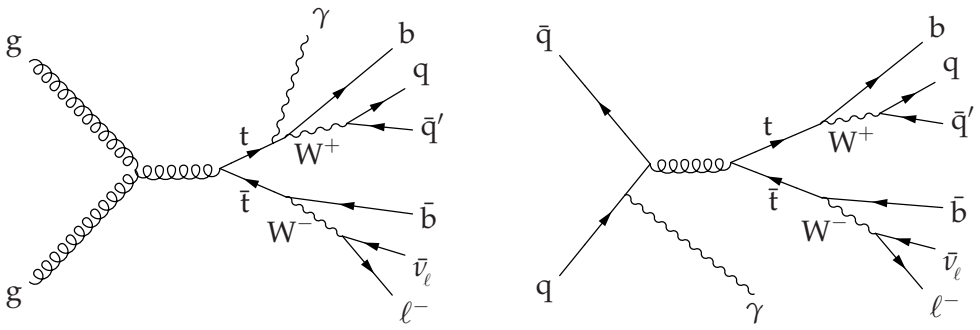


Figure 4: Two LO Feynman diagrams of the $t\bar{t}\gamma$ signal process in the semileptonic final state. The $t\bar{t}$ pair is produced via gluon-gluon fusion (left) where the photon is radiated from the top quark and via quark-antiquark annihilation (right) where the photon is radiated as a ISR from an incoming quark.

production (Figure 4 (right)) the incoming quarks may emit a photon known as initial

state radiation (ISR). The decay products include W-bosons, leptons, and b-quarks. The $t\bar{t}\gamma$ process can be represented with the same diagram as $t\bar{t}$ production with an additional photon. Figure 4 shows leading order (LO) $t\bar{t}\gamma$ production mechanisms via gluon-gluon fusion (left) and quark anti-quark annihilation (right). The contribution of the signal photon radiated from the incoming quarks is about $10\,\%$ of the total $t\bar{t}\gamma$ production. A theoretical

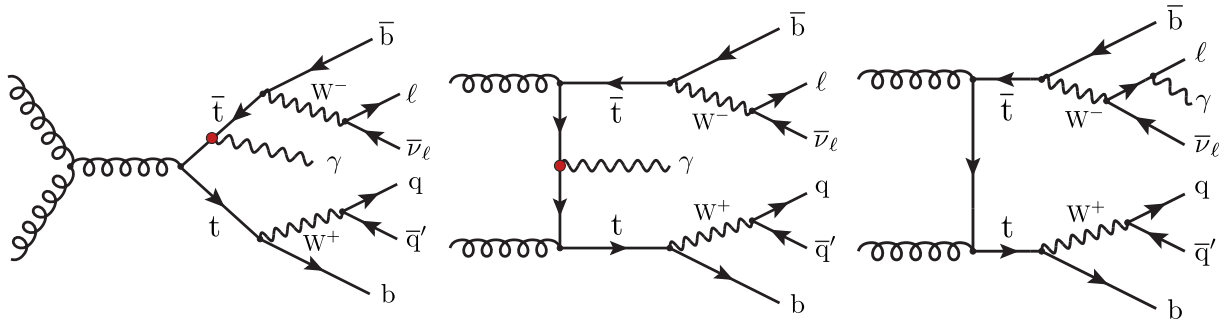


Figure 5: Some LO Feynman diagrams for the $t\bar{t}\gamma$ process in the single lepton channel where the high energy photon originates from the top quark (left, middle), or is emitted from a lepton (right). The $t\bar{t}\gamma$ interaction is indicated by a red circle in the left and middle diagrams.

prediction of the $t\bar{t}\gamma$ production cross section at leading order (LO) and next-to-leading order (NLO) with the center of mass energy of 14 TeV is presented in [21] with the QCD correction at NLO. The NLO calculation includes the emission of a gluon, quark, and loop correction from both processes. Figure 6 shows the gluon emissions (top left and top right), quark emission (bottom left) and a loop correction (bottom right) in NLO calculation.

1.4 Previous $t\bar{t}\gamma$ Measurements

The CDF experiment was the first to find evidence of $t\bar{t}\gamma$ production using 6.0 fb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV [24]. They performed a search requiring a lepton, a photon (with $E_{\mathrm{T}} \geq 10$GeV), MET, and M_3 (invariant mass of the three highest $\mathrm{P_T}$ jets) in an event. They estimated the cross section value of $\sigma_{t\bar{t}\gamma}$ = 0.18 $\pm$ 0.08 pb and cross section ratio

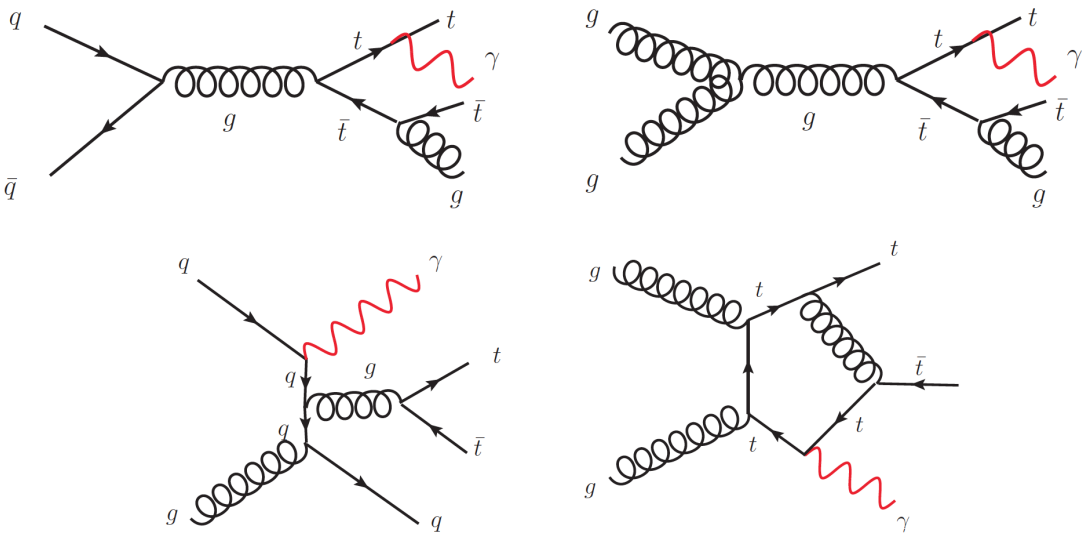


Figure 6: Real gluon emissions (top left and top right), real quark emission (bottom left) and the loop correction (bottom right) in the NLO calculation [22].

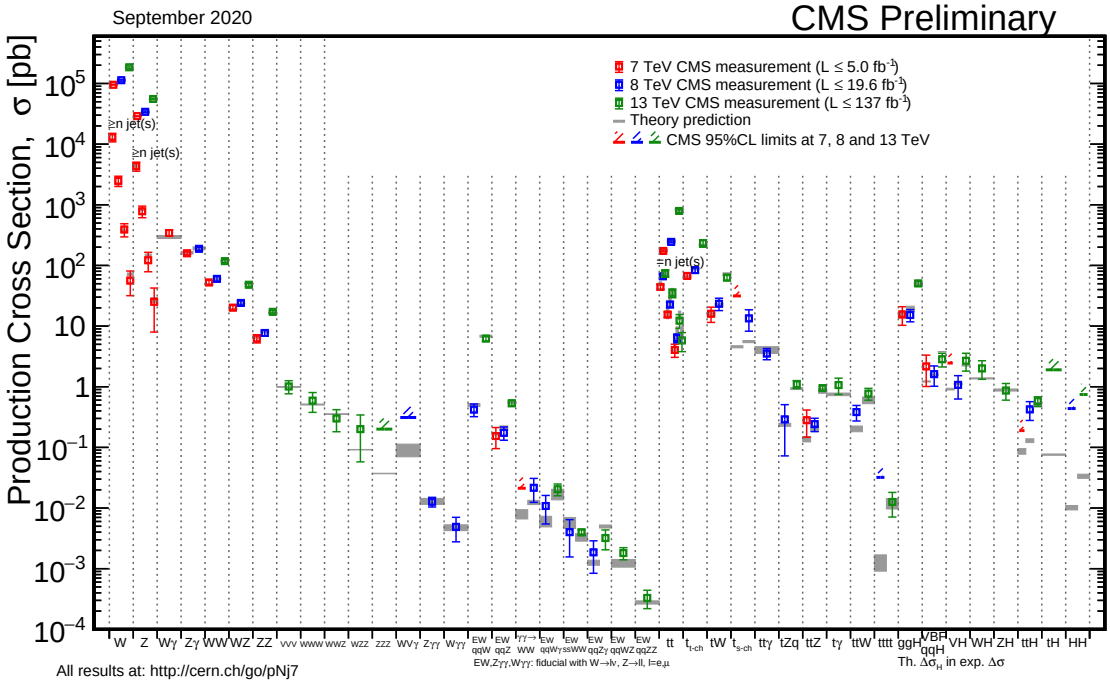


Figure 7: The production cross section of various physics processes measured by CMS for 7 TeV, 8 TeV, and 13 TeV center of mass energies [23].

$\sigma_{t\bar{t}\gamma}/\sigma_{t\bar{t}}$ = 0.024 $\pm$ 0.009. In the ratio measurement, some of the systematics and experimental uncertainties of $t\bar{t}\gamma$ and $t\bar{t}$ processes cancel out resulting in reduced uncertainties.

ATLAS performed a $t\bar{t}\gamma$ measurement in pp collisions at $\sqrt{s} = 7$ TeVusing of 4.6 fb^{-1} of data in a single lepton channel requiring large transverse momentum of the lepton, and a photon with $E_T \geq 20$ GeV [25]. The production of $t\bar{t}\gamma$ events was observed with a significance of 5.3σ in the fiducial region. The fiducial cross section was found to be $\sigma_{t\bar{t}\gamma} \times$ BR $= 64 \pm 8$(stat) $^{+17}_{-13}$(syst) $\pm$ 1(lumi) fb per lepton flavor. It is consistent with the leading order (LO) SM prediction. ATLAS also performed another analysis at $\sqrt{s} = 8$ TeV in the semi-leptonic channel with an integrated luminosity of 20.2 fb^{-1} [26]. The fiducial cross section was found to be $\sigma_{t\bar{t}\gamma} = 139 \pm 7$(stat.) $\pm$ 17(syst.) fb. ATLAS measured the $t\bar{t}\gamma$ cross section again at $\sqrt{s} = 13$ TeVwith a total integrated luminosity of 36.1 fb^{-1} [27]. Machine learning (neural network) techniques were applied to separate signal from background. The fiducial cross sections of $\sigma^{fid.}_{t\bar{t}\gamma} = 521 \pm 9(\text{stat}) \pm 41(\text{syst})$fb and $\sigma^{fid.}_{t\bar{t}\gamma} = 69 \pm 3(\text{stat}) \pm 4(\text{syst})$fb were found for the semileptonic and dileptonic channels, respectively.

CMS performed an analysis at $\sqrt{s} = 8$ TeV center of mass energy using data corresponding to an integrated luminosity of 19.7 fb^{-1} in the semileptonic channel [28] . The ratio of $t\bar{t}\gamma$ to $t\bar{t}$ cross sections was found to be $\sigma_{t\bar{t}\gamma}/\sigma_{t\bar{t}} = (5.2 \pm 1.1) \times 10^{-4}$. The $t\bar{t}\gamma$ cross section was calculated using the measured ratio and the $t\bar{t}$ cross section in the semi-leptonic channel and found to be $\sigma_{t\bar{t}\gamma} \times$ BR $= 515 \pm 108$(stat+syst) fb per channel. Very recently CMS measured the $t\bar{t}\gamma$ cross section at $\sqrt{s} = 13$ TeVwith a total integrated luminosity of 137 fb^{-1} [28]. The fiducial inclusive cross section was found to be $\sigma^{fid.}_{t\bar{t}\gamma} = 800 \pm 46(\text{stat}) \pm 7(\text{syst})$fb with a photon of $P_T > 20$GeV. The result is in good agreement with the SM prediction. All the above measurement results are summarized

as a function of luminosity, and center of mass energy in Figure 8. The x-axis is the ratio of the measured cross section to the NLO calculation. The y-axis represents the total integrated luminosity.

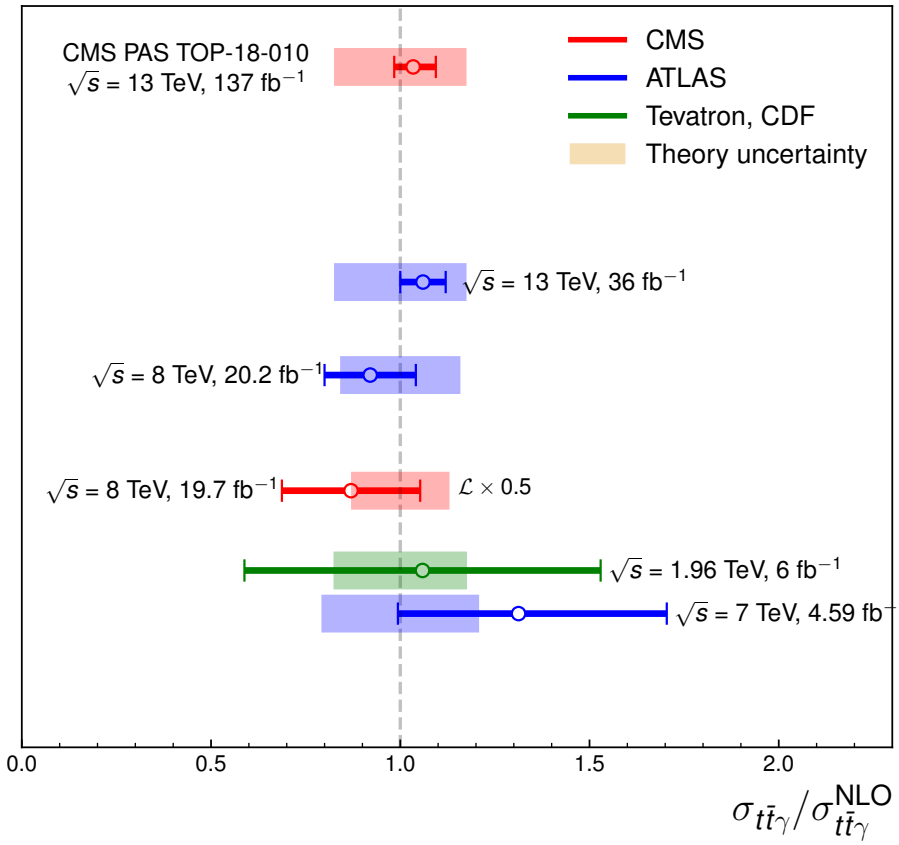


Figure 8: Summary of the previous $t\bar{t}\gamma$ cross section measurements performed at the Tevatron and LHC (CMS and ATLAS), scaled and compared to the respective theoretical predictions. The shaded regions represent the theoretical uncertainties in each analysis. The luminosity for the CMS measurement is scaled by 0.5 to avoid overlapping text.

1.5 Motivation

The top quark has unique properties that make its study of particular interest. It decays before it can form bound states, unlike the other quarks, so it can be studied as an unbound

quark [29]. It is the most massive particle known and therefore is particularly sensitive to the effects of possible, new physics that only manifest at very high energies. The study of $\mathrm{t\bar{t}}\gamma$ production is sensitive to new physics.

The $\mathrm{t\bar{t}}\gamma$ cross section is sensitive to the electric charge of the top quark and hence can be used to test models with a modified charge or higher moments. As the Wtb vertex occurs in beyond the standard model (BSM) searches, $\mathrm{t\bar{t}}\gamma$ production is sensitive to BSM effects such as magnetic and electric dipole moments [30]. Similarly, if the $\mathrm{P_T}$ spectrum of photons in $\mathrm{t\bar{t}}\gamma$ production is different from the SM prediction, then it could be a sign of new physics via anomalous top quark couplings [31]. And $\mathrm{t\bar{t}}\gamma$ events are background to Higgs physics, SUSY searches, and rare physics searches. The determination of a precise cross section for $\mathrm{t\bar{t}}\gamma$ production reduces the uncertainty for other analysis.

The analysis is carried out in the semileptonic (electron and muon) decay channels. Backgrounds are constrained by extensive studies in control regions. We extract the $\mathrm{t\bar{t}}\gamma$ cross section with a simultaneous maximum likelihood fit of control regions with the signal region. The simultaneous fit controls for backgrounds (for instance $\mathrm{t\bar{t}}$with additional photons, $V + \gamma$, etc) in the fit model.

CHAPTER 2 THE LHC AND CMS EXPERIMENT

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a circular synchrotron that collides beams of high energy protons (or heavy ions) circulating in opposite directions inside parallel vacuum pipes [32]. It began operations in 2009, and has produced large numbers of proton-proton (pp) collisions at center of mass energies ($\sqrt{s}$) of 2.36, 7, 8, and 13 TeV. The study of such high energy collisions is expected to illuminate the understanding of electroweak symmetry breaking via the Higgs mechanism and potential new physics beyond the standard model. For this purpose, seven different experiments use LHC collisions. They are CMS, ATLAS, ALICE, LHCb, TOTEM, LHCf and MoEDAL. The CMS experiment is dedicated to study a wide range of possible physics topics [33].

The series of accelerators used to supply protons to the LHC is displayed in Figure 9. Protons are produced by breaking apart hydrogen gas (H_2) molecules with an applied electric field. The protons are injected into the Proton Synchrotron Booster and accelerated to 25 GeV energy. The protons are transferred from the Proton Synchroton to the Super Proton Synchrotron and accelerated to an energy of 450 GeV. These protons are injected into the LHC beam pipe where they are accelerated to an energy of 6.5 TeV before collision. The bunches of protons are kept in circular orbit by superconducting dipole magnets and the bunches are separated by 7.5 m (or 25 ns). Each proton beam has 2802 bunches and each bunch has about 1.15×10^{11} protons. Bunches cross the interaction point 40 million times per second and in each bunch crossing there is up to about 40 collisions. This results in production of about 1 billion collisions per second [34].

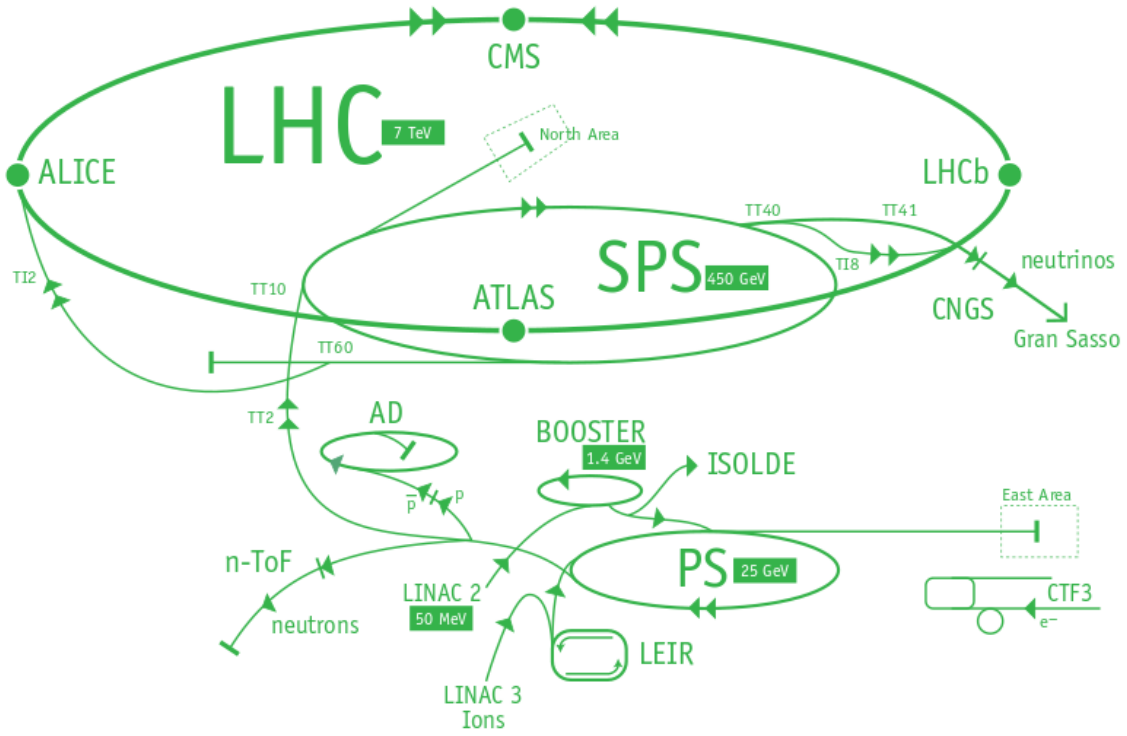


Figure 9: Layout of accelerators and major experiments at the LHC [35].

2.1.1 Luminosity

The instantaneous luminosity $\mathcal{L}_{inst}$ is the number of collisions that are produced per unit time and per unit area and defined as

$$\mathcal{L}_{inst} = f\frac{n_1 n_2}{4\epsilon\beta^*}\sigma, \tag{2.1}$$

where n_1 and n_2 are the number of particles per bunch, f is the frequency of bunch crossings. ϵ is the transverse emittance and β^* is the amplitude function [36]. The emittance of a beam is defined as how close two protons in a bunch are in position-momentum phase space and the amplitude function represents how much the beam squeezes in a short length. These parameters are related with the size of the proton bunches in the transverse

direction and given as $\epsilon\beta^* = \pi\sigma_x\sigma_y$, where σ_x and σ_y are the Gaussian transverse beam profiles in the horizontal (bend) and vertical directions [36]. The typical average values for these parameters are $\epsilon = 3.75\,\text{mm}\ \mu\text{rad}$ and β^* 0.55 m, however these values were tuned during different data taking periods. The rate of the number of events, N, for a process with a cross section σ is

$$\frac{dN}{dt} = \sigma\mathcal{L}_{inst} \tag{2.2}$$

where $\mathcal{L}_{inst}$ is the instantaneous luminosity. The integral of the instantaneous luminosity over time is called integrated luminosity

$$L_{int} = \int \mathcal{L}_{inst}\dot{dt}. \tag{2.3}$$

Figure 10 shows the integrated luminosity delivered to and recorded by CMS in 2016, 2017 and 2018. A total of 137 fb^{-1} data is used in the $\text{t}\bar{\text{t}}\gamma$ cross section measurement.

2.2 The Compact Muon Solenoid (CMS) Detector

The Compact Muon Solenoid (CMS) detector is a general purpose detector at the LHC. The aim of the CMS experiment is to study particle physics at the TeV energy scale, search for physics beyond the standard model such as super symmetry or extra dimensions, and study the properties of the newly discovered Higgs boson. Every 25 nanoseconds debris from the collisions of protons are measured in detail by the CMS detector. The CMS detector is 21 meters long, 15 meters in diameter, and weighs about 14,000 tonnes [38]. It is located in a cavern in Cessy, France near the border with Geneva, Switzerland. The central

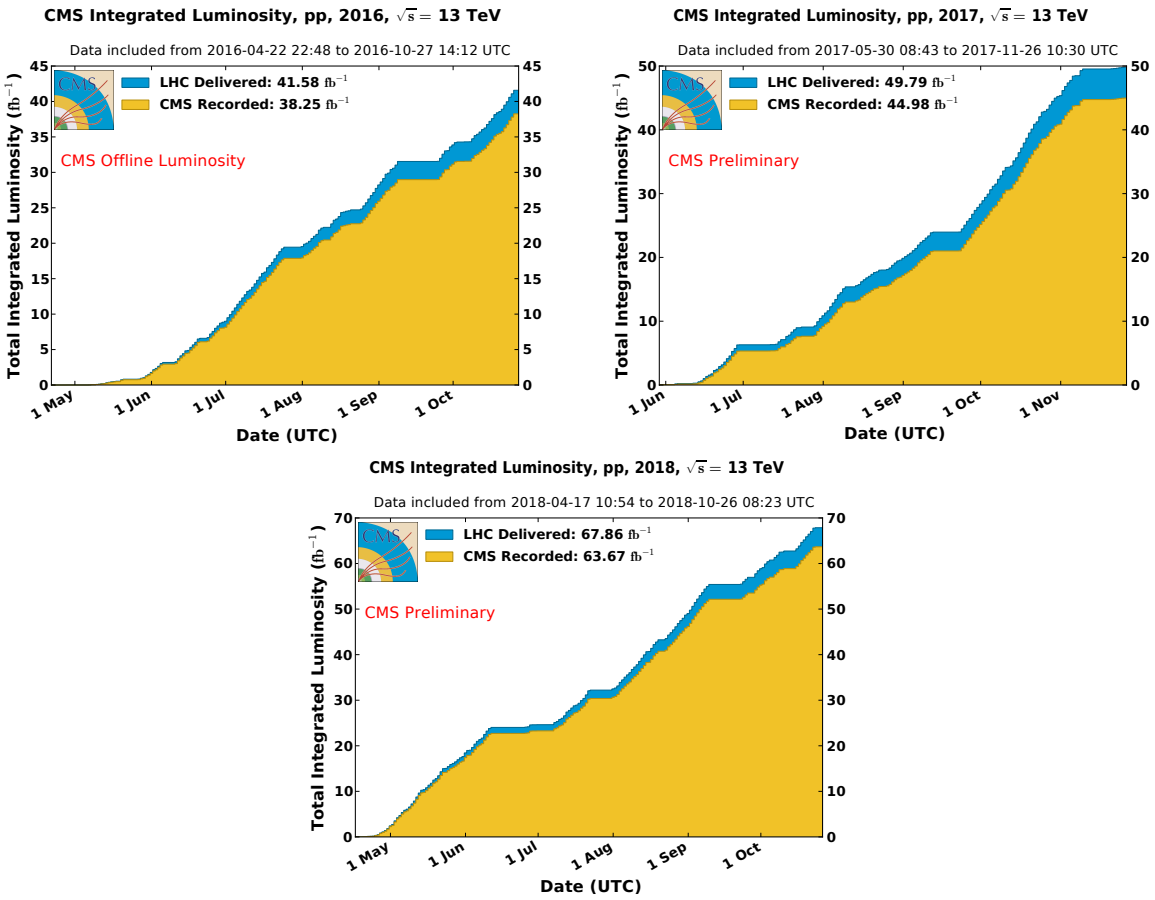


Figure 10: The integrated luminosity profiles for 2016 (left), 2017 (top right), and 2018 (bottom) data taking periods. The blue is the total delivered and the yellow is the total recorded by CMS as a function of a time [37].

element of the CMS detector is a superconducting solenoid. The solenoid creates a homogeneous axial magnetic field of 3.8 T which bends the paths of charged particles emerging from pp collisions. The detector contains subsystems designed to measure the energy, momentum, and trajectory of final state particles such as photons, electrons, muons, and hadrons. These subdetectors are arranged in concentric layers around the incoming proton beams, referred to as the beamline. The protons collide in a small region at the center of the detector known as the interaction point. A schematic view of the CMS detector with all subdetectors is shown in Figure 11. The innermost detector layer is a silicon-based tracker.

It is surrounded by a scintillating crystal electromagnetic calorimeter, which is itself surrounded with a hadronic calorimeter. The tracker and the calorimeters are located inside the solenoid magnetic. The muon detector is situated inside the iron yoke on the outside of the magnet.

The tracking volume is a cylinder of 5.8 m length and 2.6 m diameter. Three (in 2016) or four (in 2017, 2018) layers of silicon pixel detectors are placed close to the interaction point and surrounding them are 10 layers of silicon strip tracking detector. The pixel detector is used to give a precise measurement of charged-particle positions near the collision, needed to determine the impact parameter of charged-particle tracks and the position of secondary vertices. Strip layers provide the required granularity and precision to measure the charged-particle positions over a large volume necessary for an accurate determination of momentum.

The electromagnetic calorimeter (ECAL) uses lead tungsten (PbWO4) crystals corresponding to about 25.8 radiation lengths and covering the region $|\eta| \leq 3$. By using the position and magnitude of the energy depositions, the ECAL measures the kinematic properties of electrons and photons.

The ECAL detector is surrounded by the hadronic calorimeter (HCAL). It is a brass-scintillator sampling calorimeter of about 10 nuclear interaction lengths thickness. It measures the energy of hadrons.

The last layer of the CMS detector is the muon detector which is located outside the solenoid. It identifies muons and measures their momenta and positions. CMS uses 3 different types of muon sub detectors: drift tubes (DT), cathode strip chambers (CSC) and resistive plate chambers (RPC). The DT and the CSC measure the trajectory in the barrel

and end cap respectively. The RPC provides fast signals for the muon trigger system.

A schematic view of the CMS detector is shown in Figure 11. The barrel wheels and endcap are shown in their retracted positions and the elements within the solenoid are cut away to allow the central detector structures to been seen. The details of the CMS subdetectors are explained in the following sections.

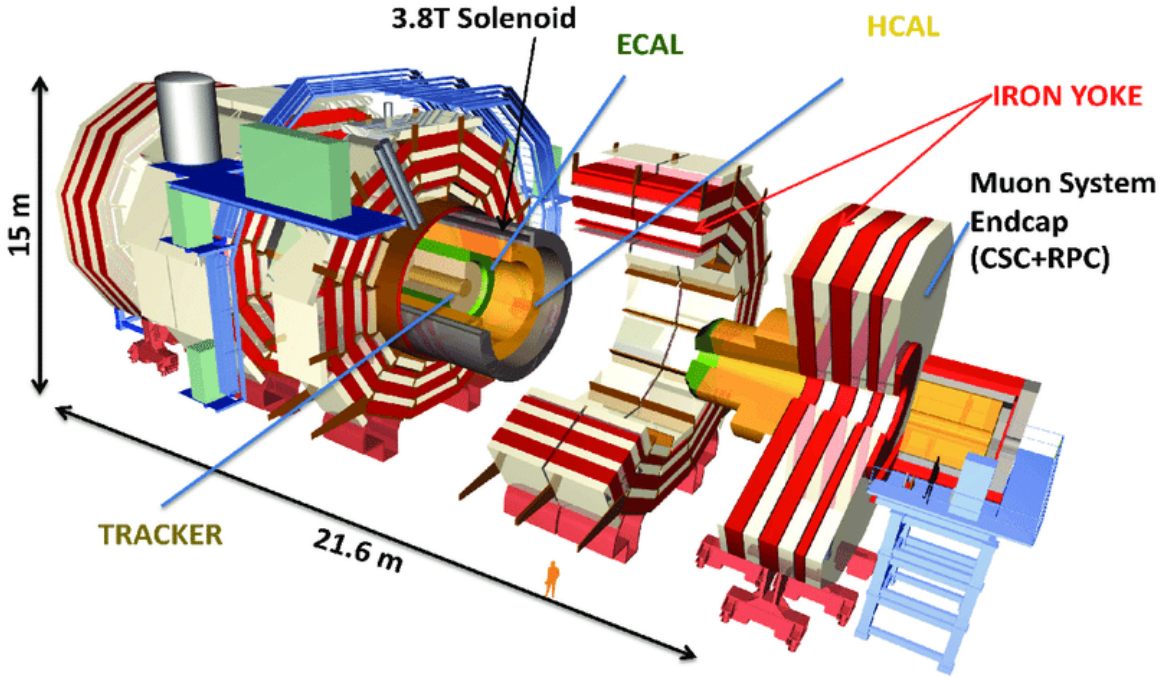


Figure 11: A schematic diagram of the CMS detector [39]. The barrel wheels and endcap are shown in their retracted positions and the elements within the solenoid are cut away to allow the central detector structures to been seen.

2.2.1 Coordinate System

CMS uses a coordinate system whose origin is at the center (interaction point) of the CMS detector. The x-axis points horizontally towards the center of the LHC ring, the y-axis points vertically upward and the z-axis points along the direction of one of the proton beams as shown in Figure 12. It is convenient to use cylindrical coordinates R, ϕ, and polar angle θ. The variables R, θ, and ϕ are the radial distance in the x-y plane, polar

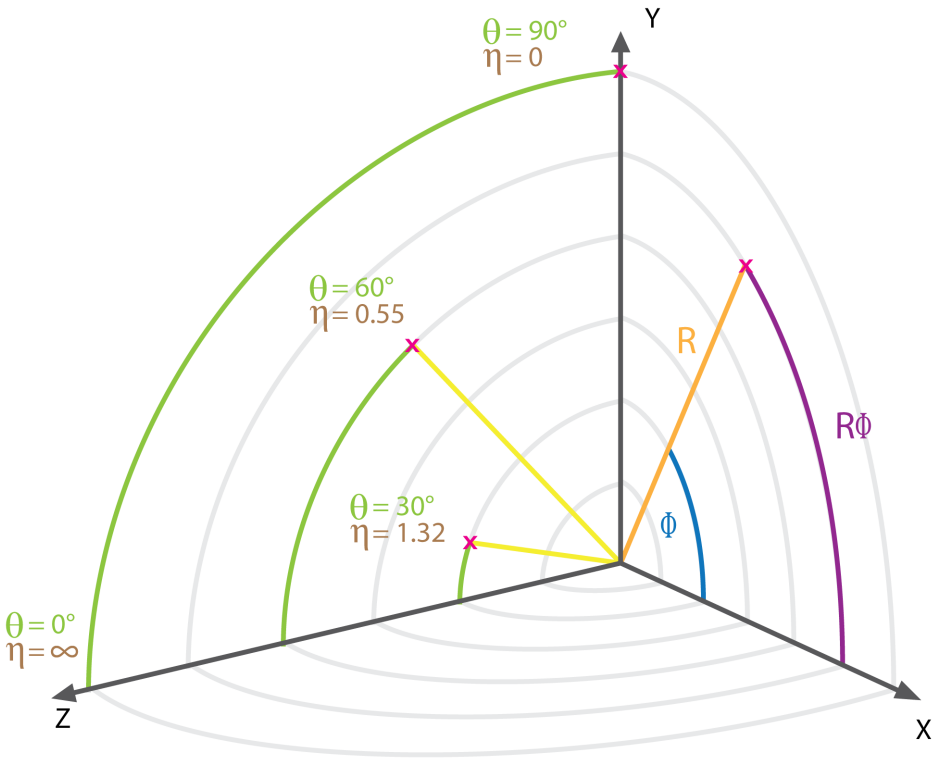


Figure 12: Coordinate systems for CMS detector [40]

angle, and azimuthal angle of the coordinate system respectively. The collision of partons within the protons can be asymmetric such that the center of mass (CoM) of the collision is boosted along the beam direction (z-axis), but polar angle differences ($\Delta\theta$) are not invariant under a boost. Two other useful quantities that behave better under a relativistic boost are rapidity (y) and pseudorapidity (η). The rapidity is defined as

$$y = \frac{1}{2} \ln \frac{E + p_L}{E - p_L} \tag{2.4}$$

where E and P_z are the energy and longitudinal momentum of the particle. It is difficult to measure the rapidity precisely without knowing the particle mass, however for ultra-

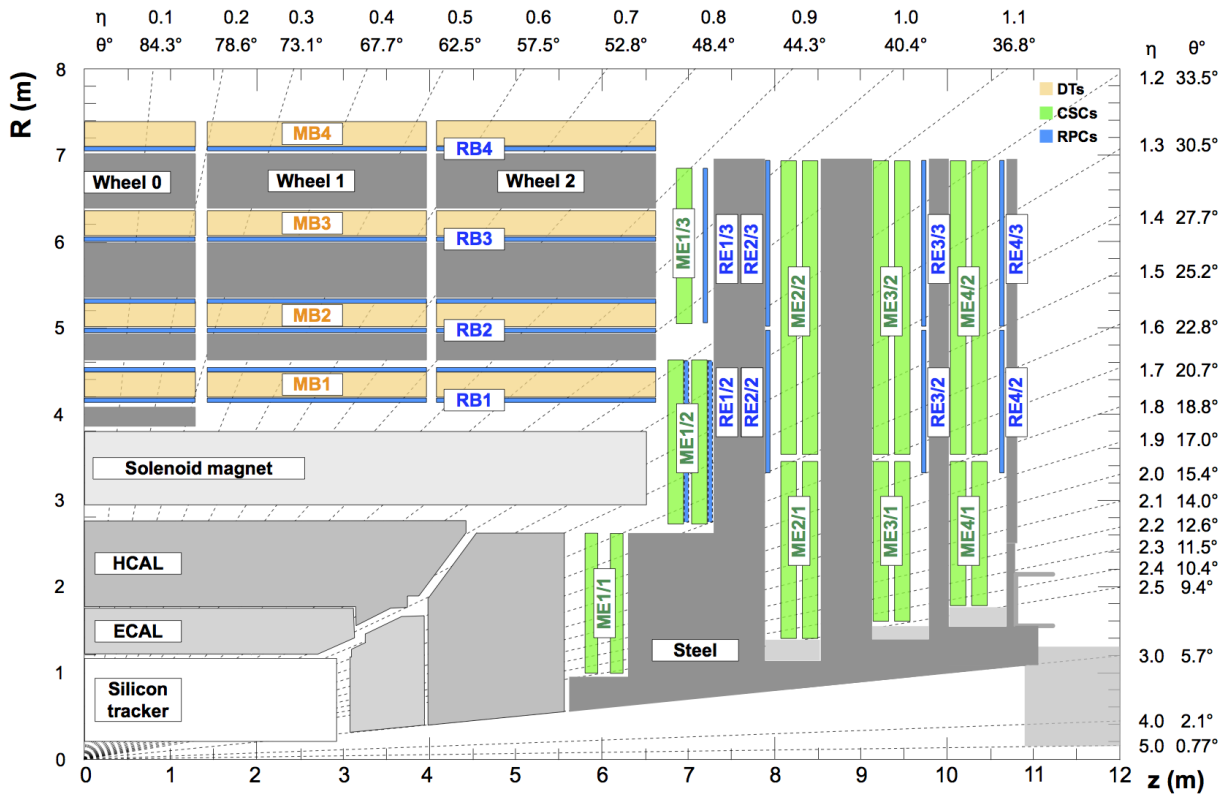


Figure 13: The CMS detector layout displaying the position in terms of both θ and η [41]

relativistic particle, the rapidity can be approximated with pseudorapidity

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right) = \sinh^{-1}\left(\frac{P_z}{P_T}\right) \tag{2.5}$$

which only depends on the direction of a particle. Pseudorapidity of $\eta = 0$ corresponds to particles in the transverse plane, $\theta = 90^0$, whereas large values correspond to particles approaching the beamline as in Figure 13. The transverse energy and momentum are given by $E_T = E \sin\theta$ and $P_T = P \sin\theta$ respectively. The momentum of a particle is measured in

terms of transverse momentum, azimuthal angle and pseudorapidity

$$\begin{aligned} p_\mu &= (E, p_x, p_y, p_z) \\ &= (m_T \cosh y, p_T \sin\phi, p_T \cos\phi, m_T \sinh y) \end{aligned} \tag{2.6}$$

where, $m_T = \sqrt{p_T^2 + m^2}$ is the transverse mass of a particle. The spatial components of particle momentum are

$$\begin{aligned} p_x &= p_T \cos\phi \\ p_y &= p_T \sin\phi, \\ \text{and } \ p_z &= p_T \sinh\eta. \end{aligned} \tag{2.7}$$

The subdetectors of the CMS system are explained briefly in the following sections. Figure 14 shows a radial section of the CMS detector and examples of how particles behave when traversing the subdetectors [38].

2.2.2 The Tracker

The tracker system is the innermost part of the CMS detector [43, 44]. It is designed to provide a precise and efficient measurement of the trajectories of charged particles emerging from the pp collisions and a precise reconstruction of secondary vertices from particle decays. A detector technology featuring high granularity and fast response is required to identify the trajectories from the same brunch crossing. The tracker must be designed for radiation tolerance. A minimum amount of material should be used to limit the multiple scattering, bremsstrahlung, photon conversion and nuclear interactions in

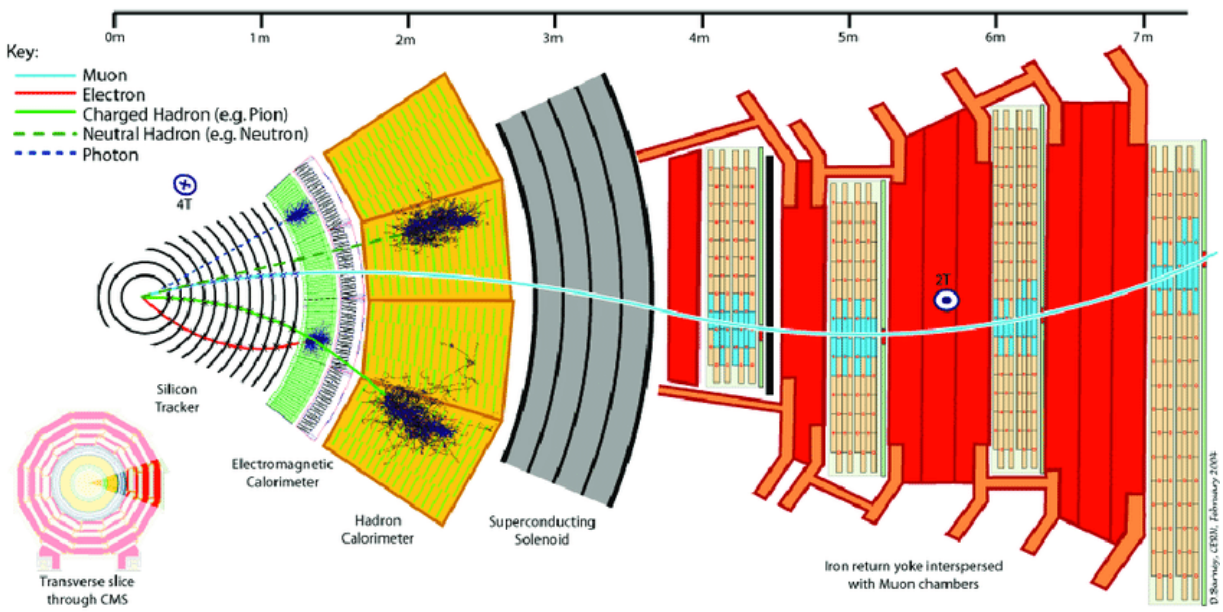


Figure 14: A radial section of CMS detector [42]

the tracker. CMS chose silicon detectors for the tracker [45]. The tracker consists of two main components: a pixel detector which covers the range of R = 4 cm to R = 15 cm and $|Z| \leq 49$ cm , and a strip detector which covers the range of R = 25 cm to R = 110 cm and $|z| \leq 280$ cm .

The pixel detector has 66 million active elements with a total surface area of about 1m^2. The active elements are 100 μm $\times$ 150 μm pixels which are oriented in the azimuthal direction with $100\ \mu$m pitch in the barrel and in the radial direction in the disks. In 2016, it has 3 barrel layers at radii of 4.4, 7.3 and 10.2 cm and 4 endcap layers at distances of $z = \pm 35.5$ and ± 48.5 cm from the interaction point. It provides efficient three-hit coverage in the region of pseudorapidity $|\eta| \leq 2.2$ and two-hit coverage in the region $|\eta| \leq 2.5$. The pixel detector produces a 3-D measurement on the path of a charged particle with a (hit) resolution between (10-20) μm . At the end of 2016, the pixel tracker was upgraded and

its current configuration was used for data taking in 2017 and 2018 [45]. The upgraded version of the pixel tracker has four layers of 124 million pixels in total, corresponding to total active area of $2\,\mathrm{m}^2$ in the barrel. The inclusion of a fourth layer with a minimal radius improved the pattern recognition and track reconstruction.

Figure 15 shows an rz-view of the CMS tracking detector in the 2016 configuration. The strip tracker has 9.3 million active elements with a total surface area of $198\,\mathrm{m}^2$. It is a hollow cylinder of length 5.8 m and diameter 2.4 m with 4 layers of inner barrel, 6 layers of outer barrel and 3 inner disks and 9 endcap disks. The strip tracker is located at radii between 0.2 to 1.2 m from the z axis and consists of three components. The tracker inner barrel and disks (TIB/TID) are located from 20 cm to 55 cm in radius and are composed of four barrel layers and three disks at each end as shown in Figure 16. The TIB/TID measures R-ϕ positions (maximum of four) of a trajectory using 320 μm thick silicon microstrip sensors. In the barrel, these strips are oriented parallel to the beam axis whereas in the disks, they are radial. The strip pitch is 80 μm in the inner pair of TIB layers and 120 μm in the outer pair of TIB layers. In the TID, the mean pitch varies between (100 - 141) μm . The TIB/TID are enclosed by the tracker outer barrel (TOB) from radius of 55 cm to 116 cm . The TOB consists of six barrel layers of 500 μm thick microstrip sensors with strip pitches of 183 μm in the first four layers and 122 μm in the last pair of layers.

The third component of the tracker is the tracker endcaps (TEC) which are located in the regions $124 \leq |z| \leq 280$ cm and $22.0 \leq R \leq 113.5$ cm . Each TEC is composed of nine disks with up to seven rings of radial-strip silicon detectors. The sensor thicknesses are thin (320 μm) in the inner four rings and thick (500 μm) in the outer three rings. The inner two layers of the TIB and TOB, the inner two rings of the TID and TEC, and the fifth

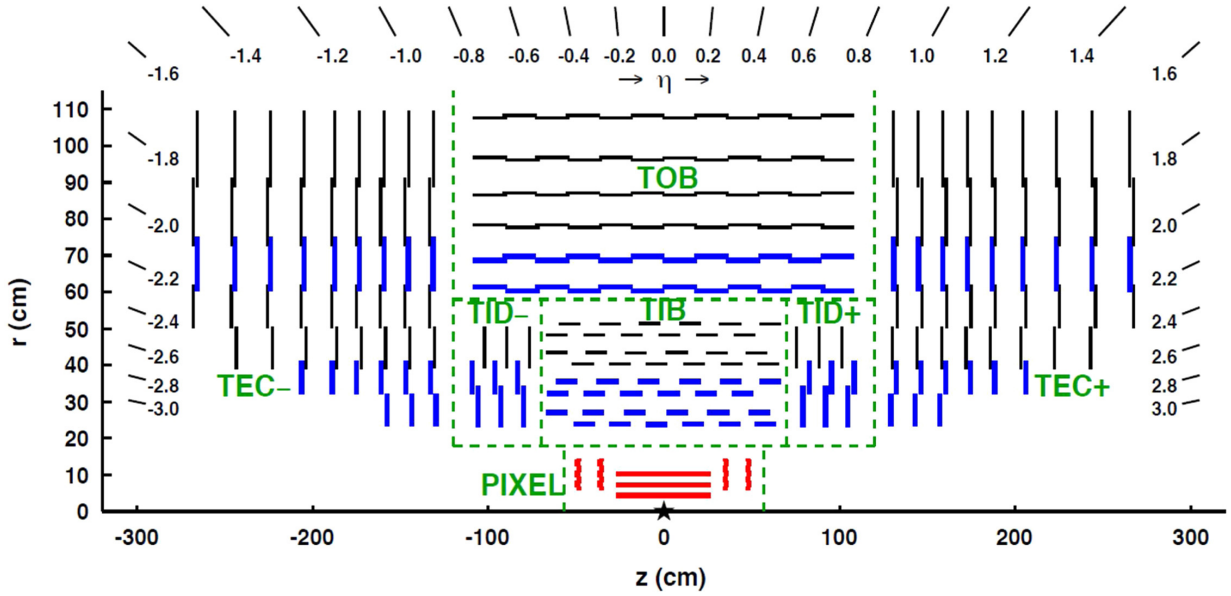


Figure 15: A r-z view of a tracker slice. Pixel modules are shown in red around the z = 0. The single-sided strip modules are shown in black whereas strip stereo modules are shown in blue [46].

ring of the TEC include a second microstrip detector module that is mounted back-to-back at a stereo angle of 100 mrad enabling an estimate of the orthogonal coordinate. A r-z view of the tracker slice is shown in Figure 15.

As particles travel through the tracker, they leave ionization signals (hits) that are amplified and detected by readout electronics. We measure the curvature of a particle track in the solenoidal magnetic field to determine the momentum of the particle with excellent resolution. The helical trajectory of a charged particle traveling through the tracker is described by the five parameters; curvature (k, inverse of radius), azimuthal angle (ϕ), pseudorapidity (η), transverse impact parameter (d_0) and longitudinal impact parameter (z_0). The pixel detector is used online for fast tracking, for primary vertex reconstruction, electron and photon distinction, muon reconstruction, tau identification,

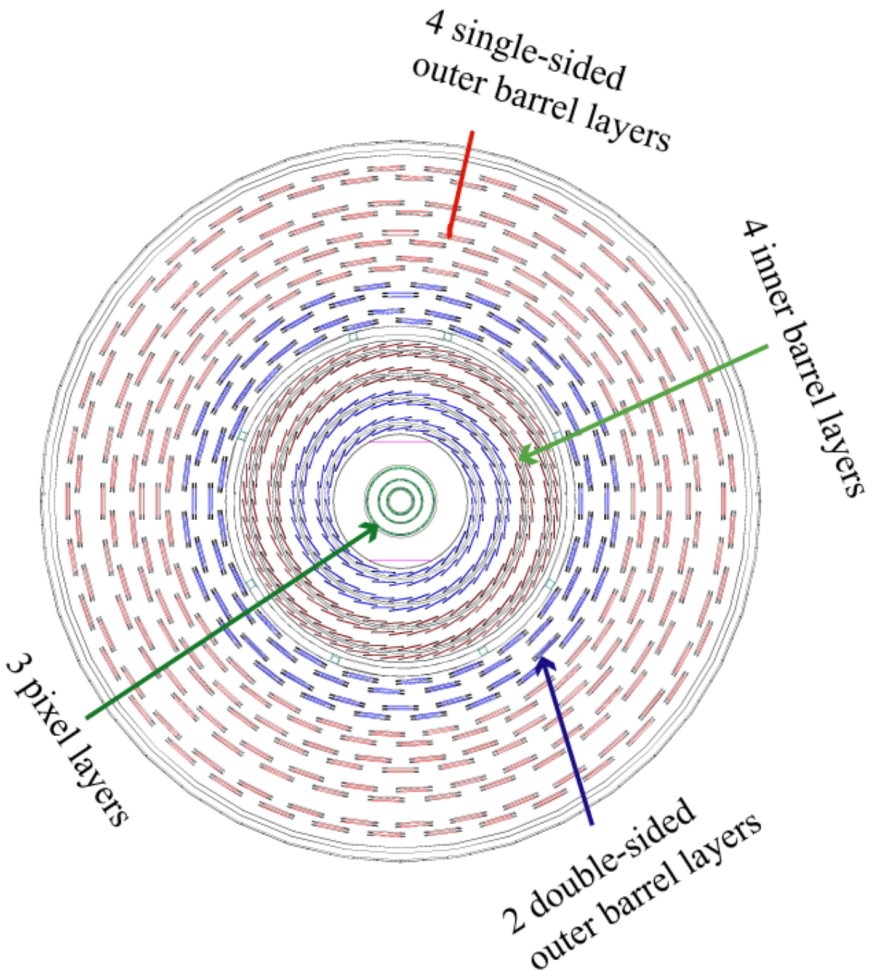


Figure 16: CMS tracker barrel layers in plane perpendicular to the beam. Three pixel layers are shown in green around the beam pipe. Four inner barrel layers, 2 double-sided outer barrel layers, and 4 single-sided outer barrel layers are shown by arrows [47].

and b-tagging. For offline use, the pixel detector forms high quality seeds for the track reconstruction algorithm. A combinatorial tracker finder algorithm is used to find the trajectory of charged particles [48].

The average track reconstruction efficiency for a charged particles with transverse momenta $P_T > 0.9$ GeV is $94\,\%$ and $85\,\%$ for pseudorapidities of $|\eta| \leq 0.9$ and $0.9 \leq |\eta| \leq 2.5$, respectively. The momentum resolution of tracks with $1 \leq P_T \leq 10$ GeV is approximately $1.5\,\%$ in the barrel region. The precision of transverse (longitudinal) impact parameter

determination decreases from 90 μm (150 μm) at P_T = 1 GeV to 25 μm (45 μm) at P_T = 10 GeV.

2.2.3 Electromagnetic Calorimeter

Information about the energies of the particles produced in each collision is collected by the CMS calorimeters; electromagnetic (ECAL) and hadronic (HCAL) calorimeters. The ECAL is the innermost of the two and measures the energy of electrons and photons by stopping them with high Z atoms. It is a hermetic detector divided into barrels and endcaps along with preshower detectors. The inner surface of the central region or barrel (EB) is at a distance of 1.29 m from the interaction point with a pseudorapidity coverage of $|\eta| \leq 1.479$. The two endcaps cover the pseudorapidity range of $1.479 \leq |\eta| \leq 3$ and are located at $|z|$ = 3.15 m from the detector center. The preshower detector (ES) is placed in front of the two endcaps to discriminate a single photon from a pair of photons coming from neutral pion decays. The ECAL is located inside the solenoid tho reduce the amount of material electrons and photons will traverse and improve the energy resolution. It is a homogeneous calorimeter made of 61200 lead tungstate ($PbWO_4$) scintillating crystals of about 25 radiation length to contain the showers of photons and electrons produced in pp collisions providing both their position and energy [49, 50]. Figure 17 shows a longitudinal section of the electromagnetic calorimeter. The thickness of all the layers of the central barrel (endcap) are 25.8 (24.7) radiations lengths [51].

2.2.4 Hadronic Calorimeter

The hadronic calorimeter (HCAL) is a sampling calorimeter made of alternating layers of brass plates and plastic scintillator tiles. The brass plates serve as the passive absorber and plastic scintillator is the active material. The HCAL measures a particle's position,

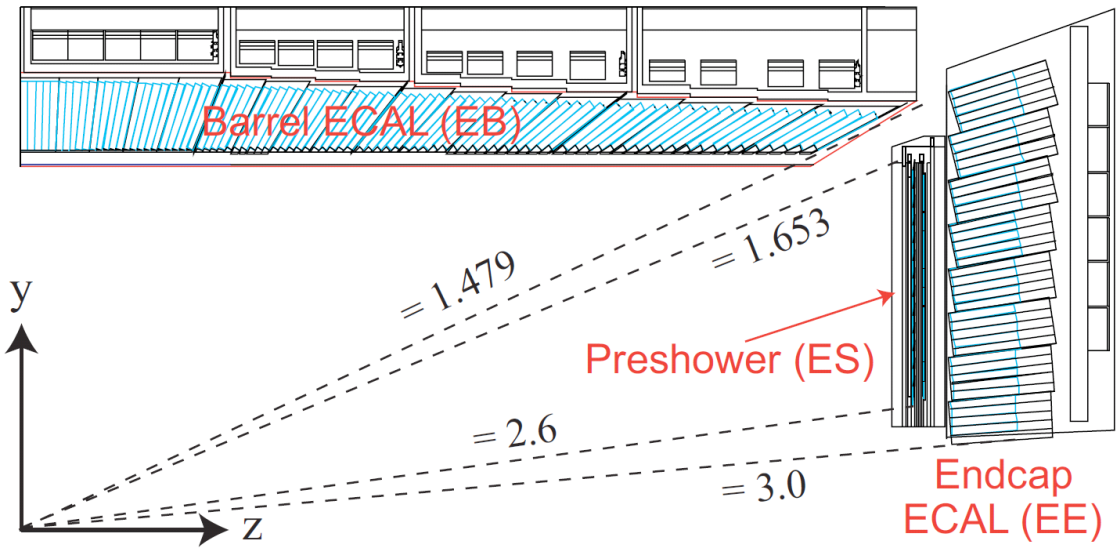


Figure 17: A longitudinal view of the CMS electromagnetic calorimeter with the respective pseudorapidity coverage [42]

energy and arrival time using a fluorescent scintillator material that produces a rapid light pulse when an ionizing particle passed through. The light is collected using optical fibers and delivered to readout boxes where photosensors and amplifiers integrate the signal and send it to the trigger and data acquisition systems. The total amount of light produced by a shower is a measure of the energy of the particles that initiated that shower. The thickness of all the layers of the central barrel is about six interaction lengths which is sufficient to contain most of a hadronic shower [51].

A longitudinal section of the hadronic calorimeter is shown in Fig. 18. The hadronic calorimeter is divided into inner and outer barrel and endcaps. The barrels cover the pseudorapidity range $-1.4 < \eta < 1.4$ and endcaps cover the pseudorapidity range $1.3 < |\eta| < 3$. The inner surface of the inner barrel starts from the outer surface of the ECAL and extends up to the inner surface of the magnetic coil. The surface of the outer hadronic calorimeter starts from the outer surface of the magnetic solenoid and extends to the inner

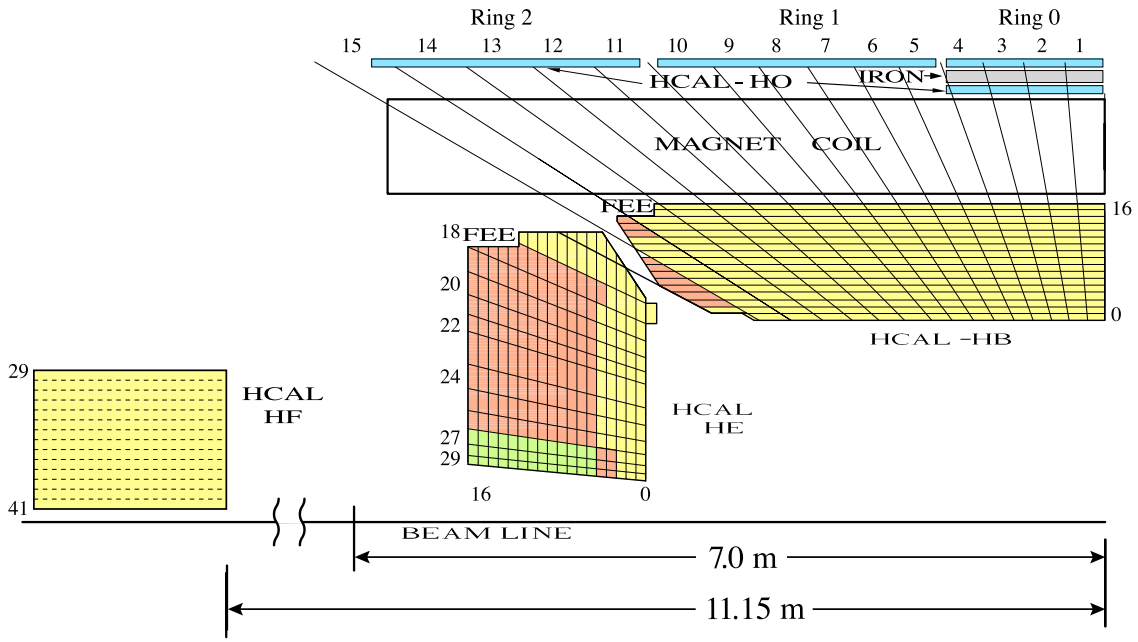


Figure 18: A quarter of the the CMS hadronic calorimeter with its four sub systems; HCAL barrel, endcaps, and forward. [53]

surface of the muon detector.

The hadron forward calorimeter (HF) is made of quartz fiber and steel [52]. It is located outside the detector wheels and covers pseudorapidies of $3.0 < |\eta| < 5.2$. When a particle passes through the fiber, it emits Cerenkov light. Photo-multipliers are used to register the Cerenkov light and generate the signal. By measuring the energy deposition, HF separates showers generated by electrons and photons from those generated by hadrons.

2.3 The Muon System

The muon system is located outside the solenoid as muons with momenta above 5 GeV can penetrate several meters of iron. The muon system is inserted among the layers of the steel magnetic flux-return yoke. A cross-section view of the barrel muon system is shown in Figure 19. A muon leaves a curved trajectory in four layers of muon detectors. There are three subsystems used: the drift tube chambers, resistive plate chambers and cathode strip

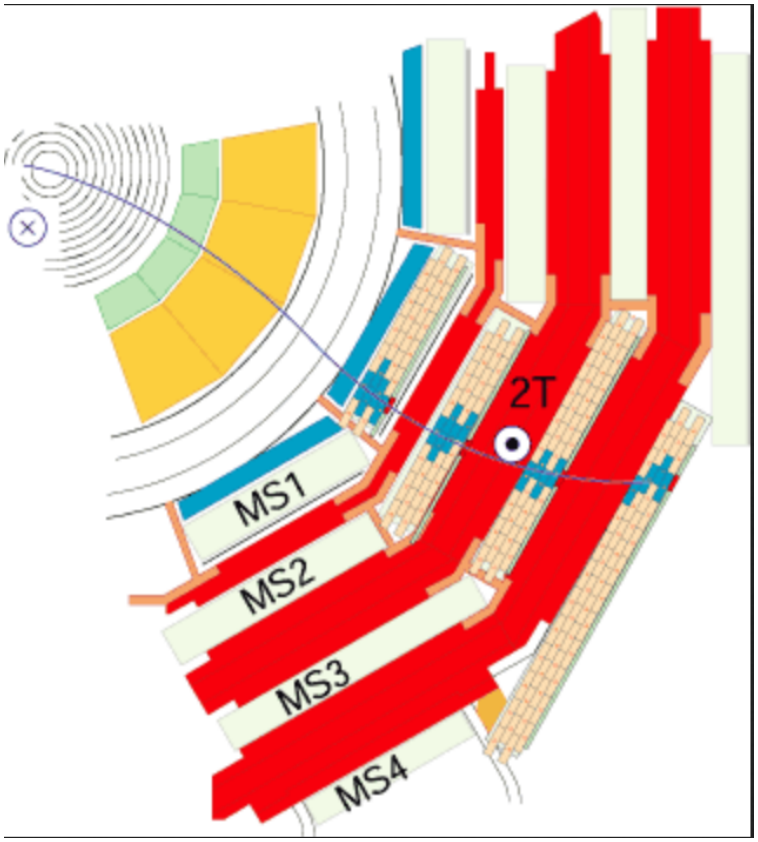


Figure 19: A cross section view of the CMS barrel Muon system. A muon leaves a curved trajectory in four layers of muon detectors [53].

chambers as shown in Figure 20. The combination of all these systems creates a highly efficient muon detector that covers $|\eta| \leq 2.4$ with a hit resolution of 200 - 350 μm in the DT, 40 - 150 μm in the CSC, and 0.8 - 1.4 cm in the RPC, with a time resolution of about 3 ns. The efficiency for muon reconstruction is from $(94 - 99)\,\%$ when using all the chamber information.

Drift Tubes The drift tube (DT) chambers are located in the barrel wheels and cover pseudorapidity of $|\eta| \leq 1.2$. The DT is constructed of drift tube cells. When a charged particle traverses the cell volume, it generates a signal (hit) in its anode wire. The signal is amplified and discriminated for time digitization [54]. Hence, the position of the charged

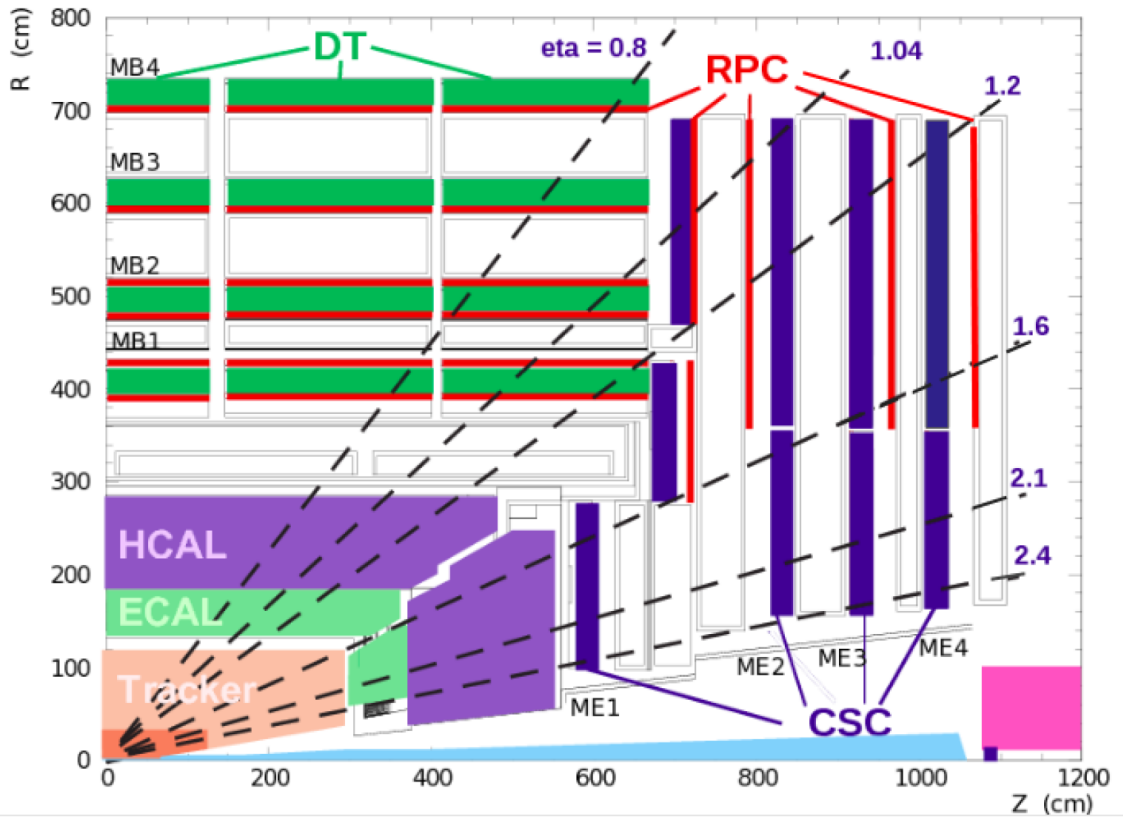


Figure 20: An r-z view of a quadrant of the CMS detector where the interaction point is at the lower left corner. The muon subsystems: drift tubes (DT), resistive plate chambers (RPC), and cathode strip chambers (CSC) in green, red, and blue respectively. The grey areas represent the solenoid and the steel flux-return disks. [53]

particle can be related to the time measurement The DT cell has a central anode wire at a high voltage (3600 V), two electrodes (cathodes) on the sides at -1800 V, and two electrodes above and below the wires at +1800 V. The gas mixture is $85\,\%$ Ar and $15\,\%$ CO_2, slightly over pressured by 2-10 mbar with a saturated electron drift velocity of about 55 μm $/ns$. The maximum drift time is almost 400 ns with a typical cell size of 42 mm .

Cathode Strip Chambers The cathode strip chambers (CSC) are installed in the end-cap region where muon rates and backgrounds are higher and the magnetic field is non-uniform and strong. As the drift length is short, the CSC has fast response time. Hence

CSCs are finely segmented to tolerate the non-uniformity of the magnetic field. The CSCs have four stations at each endcap located in a direction perpendicular to the beam and designated as ME$\pm$1 to ME$\pm$4. ME1 is composed of three rings of chambers in the radial direction and the others are composed of two. In the inner rings of stations 2, 3, and 4, each CSC chamber subtends a ϕ angle of 20 degrees, whereas the other chambers subtend an angle of 10 degrees (18 and 36 chambers in each ring, respectively). The outermost chambers are about 3.4 m long and 1.5 m wide. All chambers use a gas mixture of $50\,\%$ CO_2, $40\,\%$ Ar, and $10\,\%$ CF_4.

Resistive Plate Chambers The resistive plate chambers (RPC) are installed in both the barrel and endcap regions. The RPCs are double-gap chambers operated in avalanche mode and are primarily designed to provide timing information for the muon trigger. They provide a fast, independent trigger with a lower P_T threshold over a large portion of the pseudorapidity range ($|\eta| \leq 1.9$). An RPC gap is made by two parallel bakelite plates separated by 2 mm and filled with a gas mixture of $96.2\,\%$ $C_2H_2F_4$, $3.5\,\%$ C_4H_10 and $0.3\,\%$ SF_6. A charged particle passing through the RPC produces an avalanche of electrons in the gap between the two plates. This charge induces a signal on an external strip readout plane to identify muons from collision events with a precision of a few ns.

2.4 The CMS Trigger System

The pp collisions occurs at an average rate of 40 MHz. However, only a very small fraction of collisions with hard interactions are useful for further study. The rate of event collection is lowered by the CMS trigger system and triggered events are saved on tape for offline analysis. The trigger system function"s in two levels: the level 1 trigger (L1), which

is hardware-based, and the high level trigger (HLT), which is software-based.

2.4.1 The L1 Trigger System

The L1 trigger system takes data from the CMS subdetectors and selects events at a rate of 100 kHz from pp collisions. The architecture of the L1 trigger system is shown in Fig. 21. It is capable of generating a decision to save or discard an event within $4\,\mu$s. The full information from local, regional and global trigger components of the muon trigger system and the calorimeter trigger system are used in the L1 decision. Once the L1 selects the event, then data from the full CMS detector is extracted for high level trigger processing.

The L1 muon trigger first forms segments of muon tracks using the hits from the CSC and DT. Then, a track finder algorithm is used to reconstruct the path of muon candidates and measure the transverse energies. The RPC provides trigger information with timing resolution that help the CSC and DT systems in finding tracks. The hits on the RPC are used for finding muon candidates in different chambers, then eliminating duplicates and selecting those with the best quality and the largest transverse momentum. For each candidate the P_T, charge, and position in (ϕ, η) are estimated. All the muon candidates (with minimum P_T of 10 GeV) from the Barrel Muon Track Finder (BMTF), the endcap muon track finder (EMTF), and the Overlap Muon Track Finder (OMTF) are fed into the global muon trigger (GMT). The GMT merges candidates, removes duplicates, estimates an optimized momentum and ranks the candidates according to the quality of the tracks and momentum. Finally, the eight best candidates are sent to the L1 global trigger (GT).

In the ECAL, arrays of 5×5 cells covering 0.087 in η and 0.087 in ϕ are used to measure the energy depositions from electrons and photons. This energy is used to generate the trigger for electrons and photons. In the HCAL, a granularity 0.087 in η and 0.087 rad in

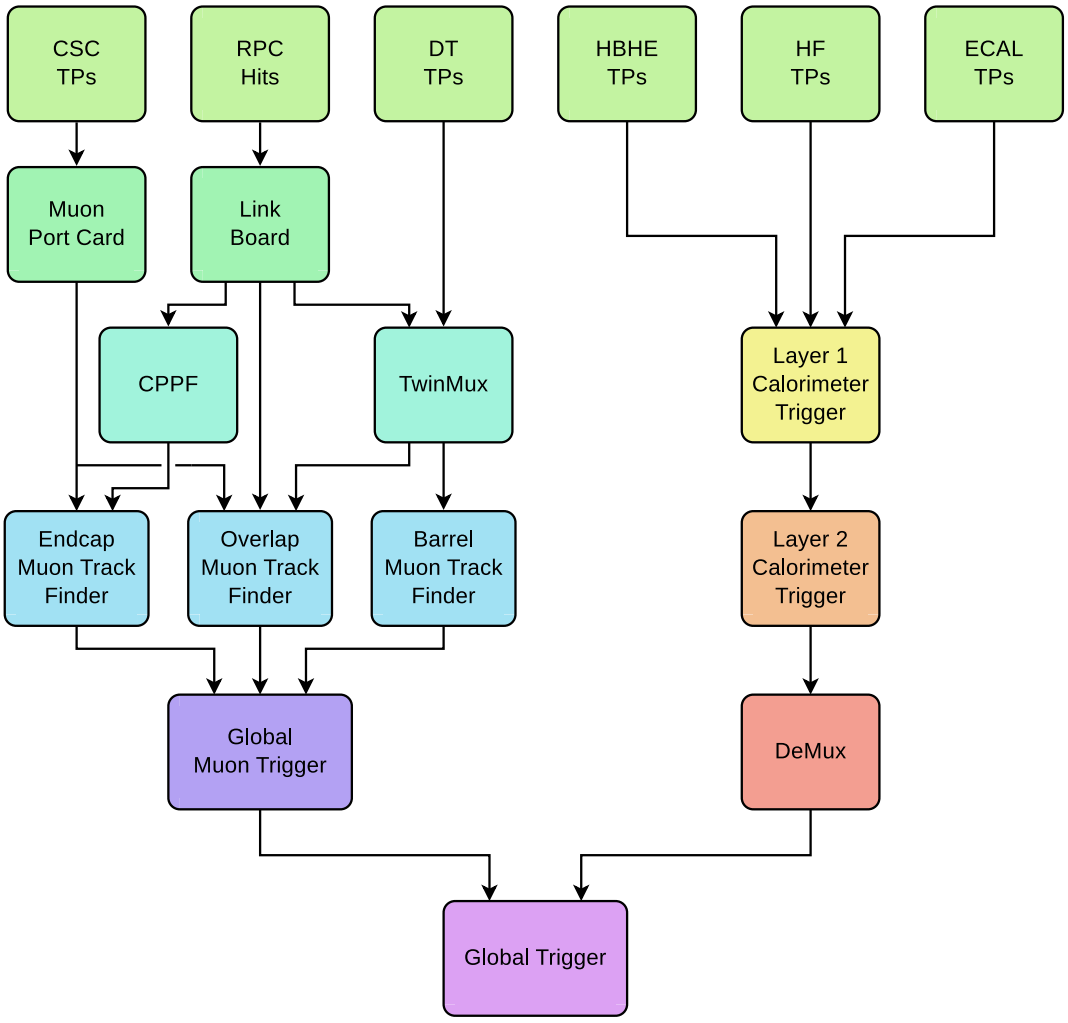


Figure 21: The schematic architecture of the level 1 trigger [55].

ϕ in the barrel, and 0.17 in η and 0.17 rad in ϕ in the endcaps is used. The information from the ECAL, HCAL and HCAL Forward (HF) is fed to the layer 1 calorimeter trigger. The layer 2 calorimeter computes other global quantities, such as the total transverse energy, the transverse energy of the jets alone, the missing transverse energy (MET) and the jet-based equivalent of the MET. It estimates the pileup contribution of an event and subtracts the energy in calorimeter objects.

Finally, the global trigger collects all the candidates and forwards them to the high level

trigger (HLT) if the selection criteria satisfy a trigger path. Each trigger path requires an object or a combination of objects to pass the selection criteria. These physics objects are fed to the global trigger from the global muon trigger and calorimeter trigger.

2.4.2 The High Level Trigger System

The HLT trigger system must have large acceptance for physics signal while keeping the output rate and CPU time under control [56]. To keep the output rate low, algorithms with high reconstruction efficiency, good identification and quality selection of main physics objects are used. To keep CPU time low, we use a modular structure of the trigger path, a sequence of reconstruction and filtering blocks of increasing complexity. The trigger path runs the faster algorithm first and filters their products. If the filter fails, the rest of the paths are skipped. The HLT selects events at the rate of about 1 kHz for offline storage.

The HLT filtering is carried out by a computer farm that uses sophisticated algorithms to reduce the event rate from 100 kHz to 1 kHz. The algorithms are more closely related to the offline reconstruction criteria and use physics object candidates. It also uses the tracker information which was not used in the L1 trigger. The overall output rate of the trigger system is adjusted by prescaling the number of events that pass the selection criteria of specific algorithms [57]. The event data are stored locally on disk and eventually transferred to the CMS Tier-0 computing center for offline processing and permanent storage.

2.5 The CMS computing system

The raw data are stored on tape as backup, as well as on disk for processing into analysis form. The amount of information is large enough to require petabytes of storage capacity per year. The LHC uses the World-wide LHC Computing Grid (WLCG) to overcome

the storage requirement. It spreads the work and resources necessary into different tiers. The CMS offline grid computing system is arranged in four different layers or tiers shown schematically in Figure 22. Tier-0 is connected to the output of the trigger system and the online data acquisition system of CMS and performs a prompt reconstruction of the raw data to create datasets with the physics objects. These data are distributed to the seven Tier-1 centers, and these centers process, reconstruct, and calibrate the data. Tier-2 centers receive data from Tier-1 and skim the data to reduce the size. They also provide storage and computing capacity for analysis and simulation of Monte Carlo (MC) events. Tier-3 centers are small capacity and resources for data analysis at individual institutes.

The data analysis is performed using the CMS software framework (CMSSW). The CMSSW framework is written in object-oriented C++ and python code. It is equipped with all the information needed by the simulation, calibration and alignment, and reconstruction modules that process event data. In a computing sense, an event is a C++ object container for all raw and reconstructed information for a particular collision event. [58]

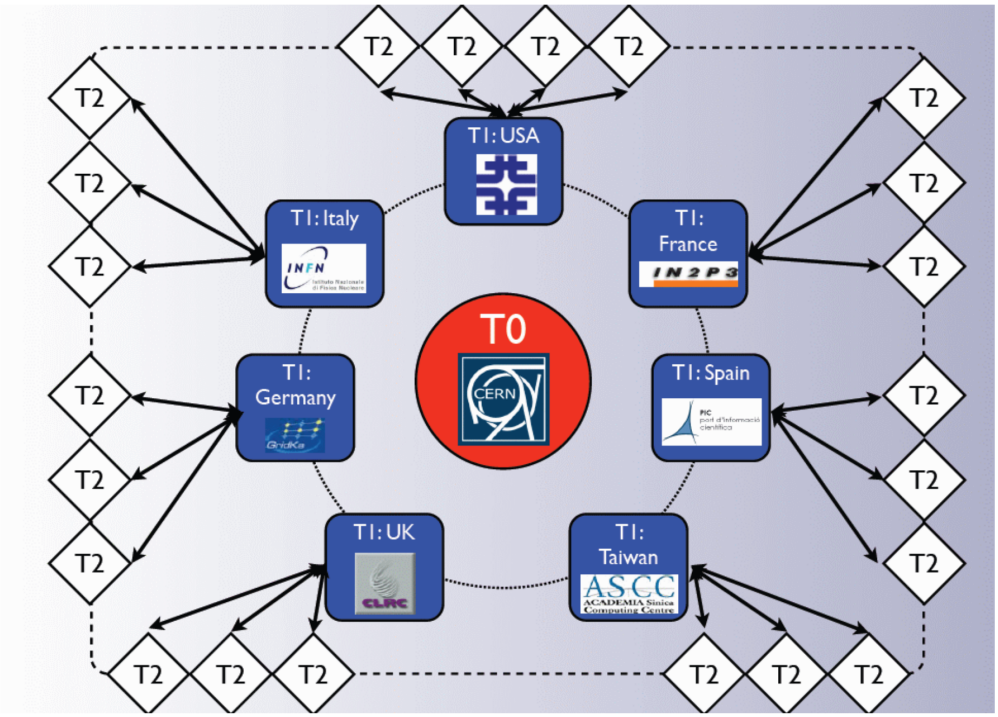


Figure 22: The CMS computing system with different tier layers [59].

CHAPTER 3 MONTE-CARLO SIMULATION AND DATASET

We use samples of simulated data to determine the acceptance and efficiency of the signal process, estimate backgrounds to the signal, and calculate systematic uncertainties of the result. The simulated samples are generated using the SM predictions for processes such as $\mathrm{t\bar{t}}$, $W + \gamma$, $Z + \gamma$, and others, as well as the $\mathrm{t\bar{t}}\gamma$ signal. The physics processes are modeled by event generators such as MADGRAPH5_aMC@NLO [60], PYTHIA 8 [61], and POWHEG [62]. A pp collision simulation using Sherpa [63] is shown schematically in Figure 23. The steps for generating a simulated event are:

- We calculate the cross section of a hard interaction between two incoming partons for a physics process at LO or NLO. The calculation is done using perturbation theory at a fixed order of α_{s} . The hard interaction includes creation of short lived resonances and their decay products.

- The partons present before and after the hard interaction can produce QED and QCD radiation as long as they conserve energy and momentum. For example, an incoming quark can split into a quark and a gluon. The quark and gluon can further split into more quarks and gluons creating a shower of partons. The parton shower interaction is added to the simulation process.

- The partons are allowed to fragment and hadronize into a colorless bound state which quickly decay into baryons, mesons, leptons and photons.

- The colliding protons leave remnants which form colorless bound states perhaps after multiparton interactions. This is called the underlying event and added to the simulation.

- A random number of additional soft pp interactions are generated and added to the events to incorporate the pile-up.

- Finally, the particles are propagated through a GEANT4 simulation of the CMS detector. GEANT4 simulates physics process and the detector response. These simulated datasets contain raw data in the same format as recorded collision events.

3.1 Description of Dataset

3.1.1 Observed Dataset

We use the full Run-II dataset collected at the center of mass energy $\sqrt{s} = 13\,\text{TeV}$ corresponding to an integrated luminosity of 137 fb^{-1} of pp collisions. Events are selected from the SingleElectron and SingleMuon primary datasets. Events are selected from the centrally produced nanoAOD version 6 (NANOAOD V6). For reference, the nanoAOD files of the primary data set are listed in Table 1 for 2016. The 2016 data are processed with

Table 1: Primary data set for the 2016 data period

Run range	data set 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`

the

`Cert_271036-284044_13TeV_ReReco_07Aug2017_Collisions16_JSON`

golden JSON file where all CMS subdetectors are flagged as good. The nanoAOD files in

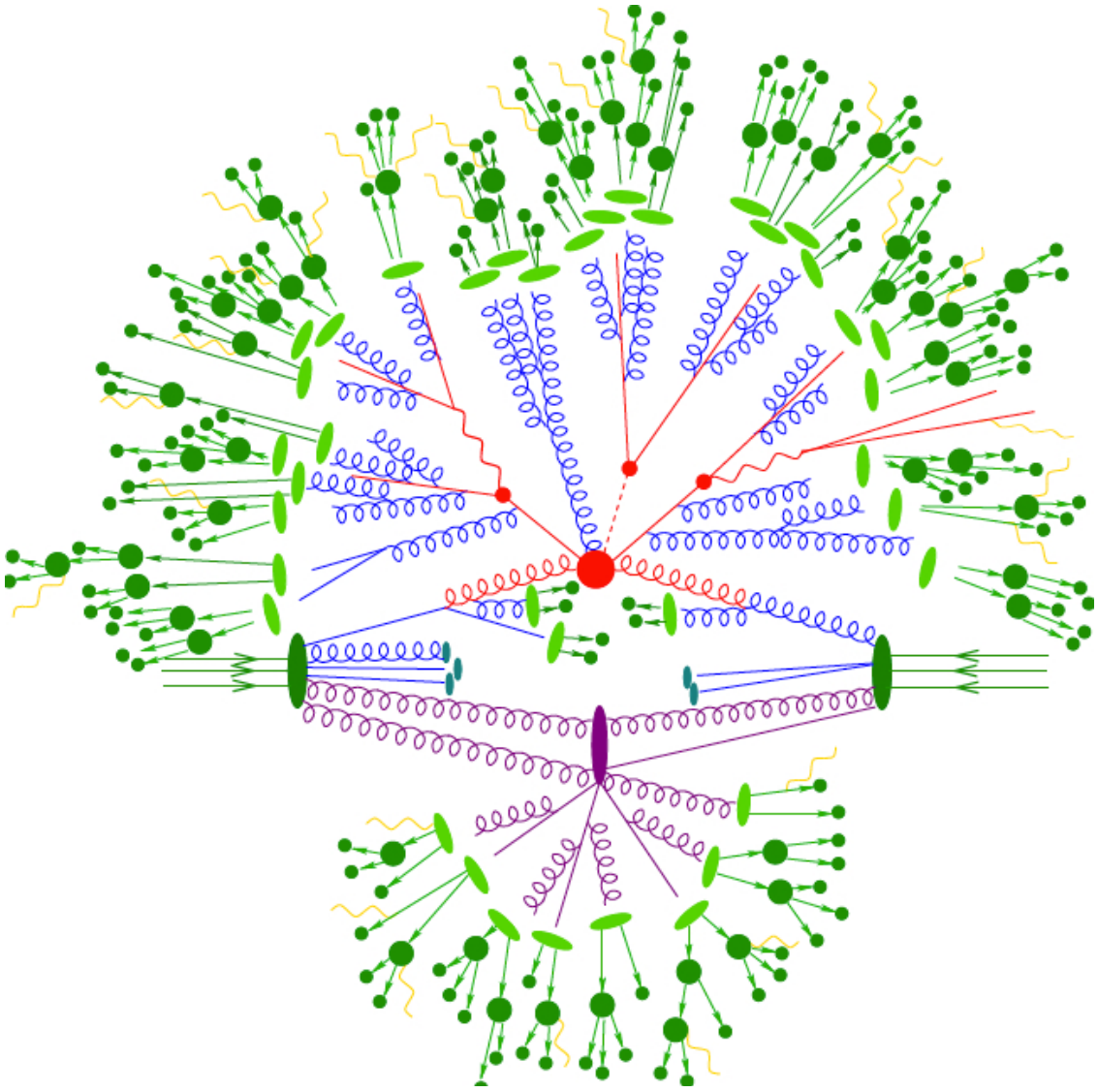


Figure 23: An example of event generation for a pp collision at the LHC. It shows the various processes from the hard interaction (red) of incoming partons, parton shower (blue), fragmentation and hadronization (light green) and decay (green), beam remnant and underlying event (purple), and emission of gamma ray (yellow) from the particles [64].

the 2017 primary data set are listed in Table 2.

Table 2: Primary data set for the 2017 data period

Run range	data set 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`

The 2017 data are processed with the

`Cert_294927-306462_13TeV_EOY2017ReReco_Collisions17_JSON_v1`

golden JSON file where all CMS subdetectors are flagged as good.

For 2018, events are selected from EGamma and SingleMuon primary data sets. The nanoAOD files used for primary data set are listed in Table 3 for 2018.

Table 3: Primary data set for the 2018 data period

Run range	data set 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`

The 2018 data are processed with

`Cert_314472-325175_13TeV_17SeptEarlyReReco2018ABC_PromptEraD_Collisions18_JSON`

golden JSON file where all CMS subdetectors are flagged as good.

3.1.2 Simulated Signal and Background Samples

All simulated samples (backgrounds and signal) are centrally produced by the CMS during nanoAOD version 6 campaigns (`NANOAODv6`) as shown in Table 4.

Table 4: 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`

The $\mathrm{t\bar{t}}\gamma$ signal sample is generated at leading order (LO) using a 2→7 process, $q\bar{q}(g\bar{g}) \to \ell^{\pm}\nu bq\bar{q}\bar{b}\gamma$, with the MADGRAPH5_aMC@NLO generator and then processed with PYTHIA 8.2 and TuneCP5 for parton showering and hadronization. The generated photon is required to have $\mathrm{P_T} \geq 10$ GeV, $|\eta| \leq 5.0$, and a minimum separation of $\Delta\mathrm{R} = 0.1$, where where, ΔR is $\Delta\mathrm{R} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, from any lepton, jet, or other photon. The photon can be produced as initial state radiation (in the case of quark anti-quark annihilation) or from the top quark or its daughter particles. Three separate samples are produced for $\mathrm{t\bar{t}}\gamma$ via different decay channels: dileptonic decay, semileptonic decay, and fully hadronic decay. The leading order (LO) production cross sections for dileptonic, semileptonic and hadronic channels are extracted from MC simulation files and 1.495 pb, 5.056 pb, and 4.149 pb respectively.

For the signal samples, an inclusive k-factor (defined as the ratio of $\sigma^{NLO}_{\mathrm{t\bar{t}}\gamma}$ to $\sigma^{NO}_{\mathrm{t\bar{t}}\gamma}$) is calculated using MADGRAPH5_aMC@NLO v2.6.5.

The LO and NLO cross sections are computed in the fiducial phase space region, resulting in an inclusive NLO k-factor of 1.4852. Finally, the k-factors are multiplied by the LO cross sections calculated in the production samples, to define NLO cross sections for use in scaling the $\mathrm{t\bar{t}}\gamma$ simulation. Photons from top decay products are not simulated in the NLO sample, and therefore the overlap removal with the $\mathrm{t\bar{t}}$ NLO POWHEG sample is

modified such that this contribution is included again in the NLO case. One of the major backgrounds is $\mathrm{t\bar{t}}$with a nonprompt photon. The $\mathrm{t\bar{t}}$sample is generated at NLO by POWHEG v2 and then combined with PYTHIA 8 with the TuneCP5 [65] tune to simulate the parton showers and hadronization, assuming a top quark mass, m_t, of 172.5 GeV .

The background processes W+jets $\rightarrow \ell\nu$+jets and $Z \rightarrow \ell^+\ell^-$+jets are simulated using MADGRAPH5_aMC@NLO (MLM) [66] along with PYTHIA 8 and TuneCUETP8M1 tune to simulate the parton shower and hadronization. These samples are normalized to their NNLO cross-section. The single top quark processes are separated into the t-channel, s-channel, tW channel and TGJets (top quark, γ +jets). The t-channel and tW channel samples are generated with POWHEG v1 and then combined with PYTHIA 8 and TuneCP5 tune to simulate the parton shower and hadronization. The s-channel is generated using MADGRAPH5_aMC@NLO and PYTHIA 8. The samples are normalized using the NLO cross-section. The $\mathrm{t\bar{t}}$ +V processes are generated by MADGRAPH5_aMC@NLO (MLM) and showering with PYTHIA 8. Diboson (WW/WZ/ZZ) samples are generated with PYTHIA 8 [65]. The QCD multijet processes are generated with PYTHIA 8.226 (8.230) the 2016 (2017, 2018) data-taking periods. The simulated QCD samples are used to estimate the Z+jets cross section in the signal region. Eventually, all the simulated QCD samples are replaced by a data driven estimation. The details of data driven QCD estimation are explained in Section 6.4. Double counting of the partons generated with MADGRAPH5_aMC@NLO and PYTHIA are removed using the MLM [67] and the FxFx [68] matching schemes for LO and NLO samples, respectively. The simulated signal and background samples for 2016, 2017 and 2018 are listed in Table 5 and Tables 6-10, respectively. The references for the CMS cross section measurements of all the samples can be found at [69].

Table 5: Signal samples for the $t\bar{t}\gamma$ process for all three years.

Process	x-sec (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

3.1.3 Event Overlap Removal in Simulated Samples

The $t\bar{t}\gamma$ signal is generated at leading order (LO) as a $2 \rightarrow 7$ process including all decay channels (dileptonic, semi-leptonic, and hadronic) with MADGRAPH5_aMC@NLO. The signal events must have a photon radiated from a top quark, the intermediate W bosons, the W decay products, the b quarks, or radiation from initial state quarks (for quark-antiquark annihilation). These are called prompt photons and are required to satisfy $P_T \geq 10$ GeV and pseudorapidity $|\eta| \leq 5$. There are no P_T cuts for the lepton but it must be within pseudorapidity of $|\eta| \leq 5$. It is required that the angular separation between the photon and any of the six final state particles (with $P_T \leq 5$ GeV) is $\Delta R > 0.1$ where, ΔR is

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{3.1}$$

On the other hand, no photons are simulated at the matrix element [1] (ME) level in the $t\bar{t}$ process but initial and final state radiation may produce photons in the samples. In such cases, we remove the $t\bar{t}$events if they contain a photon that passes the photon requirements within the phase space of the $t\bar{t}\gamma$ signal. We also require that the photon is produced by quarks, leptons, or fundamental SM bosons but not from mesons. The portion of phase

[1]The matrix element method provides a direct connection between the underlying theory of particle physics processes and detector-level physical observables [70].

space with photons originating during the hadronization of jets, or hadronic activity is covered by the $\mathrm{t\bar{t}}$ sample.

A similar overlap removal process is applied to Drell-Yan and W + jets samples, in order to remove double counting in $Z + \gamma$ and $W + \gamma$ samples. Events are removed from Drell-Yan and W + jets if they contain photons with $\mathrm{P_T} \geq 15$ GeV, $|\eta| \leq 2.6$, or with $\Delta\mathrm{R} < 0.1$ requirements. The specific isolation (ΔR) value is used to match with the generator settings in the simulation. The isolation requirement for $W + \gamma$ ($Z + \gamma$) is applied with $\Delta R > 0.1$ (0.05). Finally, overlap removal is applied to the single top samples (via t-channel) to remove events from the single-top + jets which are already included in the t+γ+jets sample. The generator level cuts are matched to the phase space of the single top quark plus photon sample: $\mathrm{P_T} \geq 10$ GeV, $|\eta| \leq 2.6$, and isolation of $\Delta\mathrm{R} < 0.05$. As the tγ sample does not include photons originating from the decay products of the top quark, such photons are not counted for the removal of single-t events.

All selection criteria for overlap removal are summarized in Table 11.

Table 6: Background MC samples for the year 2016.

Process	x-sec (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
ZJets	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
Zgamma	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		/ZZTo2L2Q_13TeV_powheg_pythia8
	4.04	/ZZTo2Q2Nu_13TeV_amcatnloFXFX_madspin_pythia8
		/ZZTo4L_13TeV_powheg_pythia8
VV($2\ell 2\nu$)	11.95	/VVTo2L2Nu_13TeV_amcatnloFXFX_madspin_pythia8 [+ext1]

Table 7: Background MC samples for the year 2017.

Process	x-sec (pb)	Dataset
$t\bar{t}$	87.315	/TTTo2L2Nu_TuneCP5_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
ZJets	18610	/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8 [+ext1]
	6077.22	/DYJetsToLL_M-50_TuneCP5_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_TuneCP5_13TeV_amcatnlo_madspin_pythia8
$W + \gamma$	463.9*1.295	/WGToLNuG_TuneCP5_13TeV-madgraphMLM-pythia8
Zgamma	105.4	/ZGToLLG_01J_5f_lowMLL_TuneCP5_13TeV-amcatnloFXFX-pythia8
$t\bar{t} + W$	0.4062	/TTWJetsToQQ_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8
	0.2043	/TTWJetsToLNu_TuneCP5_PSweights_13TeV-amcatnloFXFX-madspin-pythia8
$t\bar{t} + Z$	0.5297	/TTZToQQ_TuneCP5_13TeV-amcatnlo-pythia8 [+ext1]
		/TTZToLLNuNu_M-10_TuneCP5_PSweights_13TeV-amcatnlo-pythia8
		/TTZToLL_M-1to10_TuneCP5_13TeV-amcatnlo-pythia8
tZ	0.09418	/tZq_ll_4f_ckm_NLO_TuneCP5_PSweights_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_v2
	10.71	/WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8
	5.595	/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8
	4.42965	/WZTo3LNu_13TeV-powheg-pythia8
ZZ		/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8
	4.04	/ZZTo2Q2Nu_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8
		/ZZTo4L_13TeV_powheg_pythia8 [+ext1+ext2]
VV($2\ell 2\nu$)	11.95	/VVTo2L2Nu_13TeV_amcatnloFXFX_madspin_pythia8

Table 8: Background MC samples for the year 2018.

Process	x-sec (pb)	Dataset
$t\bar{t}$	87.315	/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8
	380.095	/TTToHadronic_TuneCP5_13TeV-powheg-pythia8
	364.352	/TTToSemiLeptonic_TuneCP5_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
ZJets	18610	/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8 [+ext1]
	6077.22	/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8
Single top	3.68064	/ST_s-channel_4f_leptonDecays_TuneCP5_PSweights_13TeV-amcatnlo-pythia8 [ext1]
	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 [ext1]
	35.85	/ST_tW_antitop_5f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8 [ext1]
	2.967	/TGJets_TuneCP5_13TeV_amcatnlo_madspin_pythia8
$W+\gamma$	463.9*1.295	/WGToLNuG_TuneCP5_13TeV-madgraphMLM-pythia8
Zgamma	105.4	/ZGToLLG_01J_5f_lowMLL_TuneCP5_13TeV-amcatnloFXFX-pythia8
$t\bar{t}+W$	0.4062	/TTWJetsToQQ_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8
	0.2043	/TTWJetsToLNu_TuneCP5_PSweights_13TeV-amcatnloFXFX-madspin-pythia8
$t\bar{t}+Z$	0.5297	/TTZToQQ_TuneCP5_13TeV-amcatnlo-pythia8 [+ext1]
		/TTZToLLNuNu_M-10_TuneCP5_PSweights_13TeV-amcatnlo-pythia8
		/TTZToLL_M-1to10_TuneCP5_13TeV-amcatnlo-pythia8
tZ	0.09418	/tZq_ll_4f_ckm_NLO_TuneCP5_13TeV-madgraph-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_TuneCP5_13TeV-amcatnloFXFX-pythia8 [+ext1]
ZZ		/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8
	4.04	/ZZTo2Q2Nu_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8
		/ZZTo4L_TuneCP5_13TeV-amcatnloFXFX-pythia8
VV($2\ell 2\nu$)	11.95	/VVTo2L2Nu_13TeV_amcatnloFXFX_madspin_pythia8

Table 9: QCD MC samples used for additional checks for the year 2016.

Process	x-sec (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 (EG enriched)	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

Table 10: QCD MC samples used for additional checks for the year 2017 and 2018.

Process	x-sec (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 (em enriched)	5352960	/QCD_Pt-20to30_EMEnriched_TuneCP5_13TeV_pythia8
	9928000	/QCD_Pt-30to50_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_TuneCP5_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 11: Summary of the cuts for the overlap removal of simulated samples.

	samples				
Cuts	$t\bar{t}\gamma$ / $t\bar{t}$	$W+\gamma$ / $W+$jets	$Z+\gamma$ / $Z+$jets	tγ / single-t (t-channel)	γ+jets / QCD
P_T (γ^{gen}) (GeV)	≥ 10	≥ 15	≥ 15	≥ 10	≥ 25
$\lvert\eta\rvert$ (γ^{gen})	≤ 5.0	≤ 2.6	≤ 2.6	≤ 2.6	≤ 2.5
ΔR cone size	0.1	0.05	0.05	0.05	0.4

CHAPTER 4 EVENT RECONSTRUCTION AND DATA TIER

4.1 Trigger

At the LHC, proton bunches collide at a rate of 40 MHz i.e., every 25ns, but only about 1kHz of data are registered for analysis. The upper limit of 1 kHz is determined by the capacity to save the data. CMS uses a two level trigger system, level 1 (L1) and high level (HLT) that selects events at a rate of 1kHZ. The L1 trigger is hardware based and uses calorimeter and muon information to reduce the rate from 40MHz to 100kHz. It makes a decision to accept an event within 4μs. Once the event is selected by L1, full readout of the CMS detector takes place. The HLT is a computer farm running sophisticated algorithms that select events at a rate of 1kHz i.e., 1 event every 260 ms. It is composed of hundreds of HLT trigger paths which are sequences of reconstruction and filtering modules [71].

The HLT electron: The Gaussian-sum filter track algorithm is used to reconstruct the track information of the electron in the HLT. The track information is then combined with an ECAL energy deposition to reconstruct the HLT electron. The online electron reconstruction efficiency in the HLT is very high ($\geq 99\%$) [71]. To reduce the rate of jets and photons misidentified as electrons, tight identification requirements are applied. The tight identification selection helps in lowering the P_T threshold value in trigger.

The HLT muon: At HLT, muons are reconstructed using the information from both muon subdetectors and tracker. The reconstruction efficiency of muons is greater than 99% [71] and computed using muon P_T and η.

4.2 Event Reconstruction and the Particle Flow Algorithm

The particles produced in a collision propagate through the detector, interacting with the detector material. The signals left by the interactions are used to reconstruct the event, including energies and positions of jets and photons, and momentum and trajectories of electrons and muons. This is known as event reconstruction. An example of event reconstruction of a top-quark pair is shown in Figure 24. Offline event reconstruction is based on the particle flow algorithm [72]. When a particle traverses the detector it produces signals which are digitized and locally reconstructed in the subdetectors (inner tracker, outer tracker, electromagnetic calorimeter, hadronic calorimeter, and muon detector). The reconstructed particle flow (PF) elements are combined to form particle-flow-object candidates used in this analysis: photons, jets, electrons, muons, missing transverse energy, primary vertices, and secondary vertices.

Particle flow uses the information from tracks formed from tracker and muon chamber signals and the information of energy deposition in the ECAL and HCAL to estimate the trajectory of a stable particle in an event. All the stable particles are reconstructed by linking the responses from the appropriate subdetectors. Composite objects like jets, taus, and MET are also reconstructed from the particle flow candidates. For example, when an electron candidate passes through the inner and outer tracker it leaves signals in different layers. These hits are used to create track candidates and the tracks are used to create vertices. The electron, a charged particle, will follow a curved path in the tracker due to the presence of the magnetic field. The photon will travel straight. They both deposit energy and get absorbed in the ECAL. However, charged and neutral hadrons can traverse

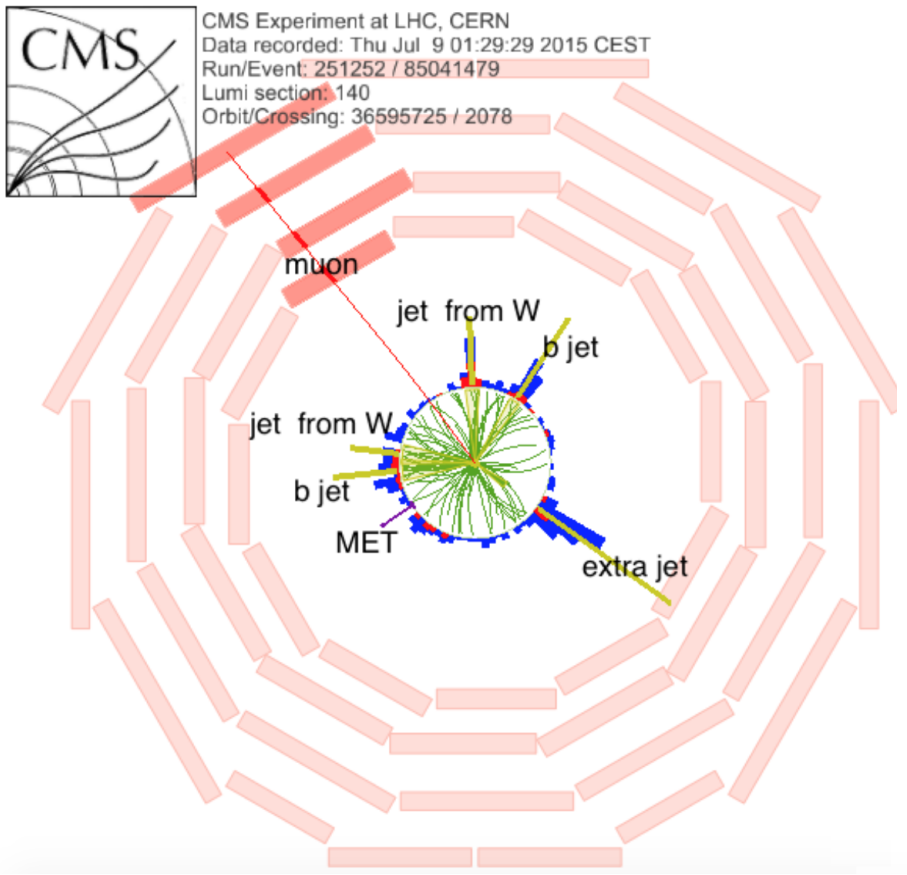


Figure 24: Event reconstruction of a top-quark pair in the xy view. The event has one isolated muon and five hadronic jets. Two of the jets pass the tight threshold on the b-tagging discriminant and are interpreted as originating from the b quarks from top quark decay. Two of the others form a hadronically-decaying W boson [73].

the ECAL and deposit energy in the HCAL. Similar to electrons and photons, charged hadrons bend in the magnetic field but neutral hadrons do not. The muon chambers are used to detect muons since they can penetrated the ECAL and HCAL without being stopped whereas neutrinos leave the detector undetected. The subdetector information is linked together by the particle flow (PF) algorithm to create a collection of physics objects: electron, muon, photon, and jets are the PF objects used in this analysis [51].

4.2.1 Tracking

The reconstruction of the track made by a charged particle is based on a Kalman filter algorithm. First, a track seed is generated from hits[2] in the pixel tracker. Using the Kalman filter track extrapolation, more hits are added to the current track candidate from successive tracker layers [74]. New track candidates are created in the process if multiple hits match with the extrapolated track. Once all the hits are included in the track candidate from all layers in the pixel and strip tracker, any duplicate tracks are removed, tracks with the most missing hits are removed, and tracks with the highest χ^2 are removed. The charged particle properties such as production vertex, P_T , and charge are determined from final track fitting. The process is done recursively with additional track quality requirements such as minimum $P_T > 0.9$ GeV, χ^2 requirement, and primary vertex distance. Once the track is accepted then all the hits associated with the tracks are removed from the collection and the search starts again for additional tracks [75].

4.2.2 Clustering in the Calorimeters

The clustering algorithm takes the values of energy depositions, positions, indices to neighbors, and index from subdetector and uses them to generate the seed selection. The seed cluster must have more energy than its neighbors and an energy greater than the threshold, E_{th}. The threshold energy is 0.23 GeV and 0.8 GeV in the ECAL and HCAL, respectively. Then we determine the topological cluster by selecting the set of cells having a common side or corner with the seed. Sometimes it is possible to have several seeds in the same topological cluster. Then we calculate the fraction of energy contributed to the

[2]"hits" are the signals in the tracking detector

cluster using particle flow rechits [3] information. The sharing of the energy is calculated iteratively a maximum of 50 times to find the final cluster energy and position [77].

4.2.3 Linking Algorithm

When a final state particle emerges from a pp collision, it traverses the detector creating PF elements in different subdetectors. The particle flow algorithm links those elements to reconstruct a particle. If two elements are found to be linked, a quality test is done by finding a distance between the elements. Then the algorithm produces PF blocks of elements from common linked elements. Finally particle flow objects are identified and reconstructed as explained below.

4.2.4 Particle Flow Objects

This physics analysis is based on the counting of well identified physics objects such as muons, electrons, jets, b-tagged jets, photons, and MET. In this section we will describe briefly how we identify those objects using the particle flow algorithm. The algorithm uses the PF blocks as input and performs reconstruction and identification in the following order.

Particle flow muons: PF muons are identified using global muon and tracker muon properties. The muons are identified in three different ways in the CMS detector. First, the hits in the drift tube (DT) and cathode strip chambers (CSC) are used to form a segment of a track. This track is used as a seed to match with a possible muon path using all the hits from DT, CSC and resistive plate chambers (RPC). The muon candidate identified in this way is called a particle flow standalone muon. Signals from muon chamber and tracker

[3] Raw data from the detector are unpacked offline into integer-based objects called “digis.” The information stored in the digis is processed to produce a collection of objects called "rechits" [76]

tracks are aligned to reconstruct the full path of a muon. The muon identified in such case is called a particle flow global muon. A muon with $P_T < 200$ GeV will have a very good momentum resolution. If the muon candidate has P_T less than 10 GeV, it may have few hits in the muon chambers yet be well reconstructed in the tracker. When there is a match between muon system hits and an extrapolated track from a tracker, it is called a tracker muon. To identify a tracker muon, the transverse momentum of the track should be greater than 0.5 GeV and total momentum should be greater than 2.5 GeV.

As the objects are not exclusive, a muon candidate can be classified as both tracker muon and global muon. If both match to the same track, then they are classified as the same object. It is possible that an energetic charged hadron may penetrate to the muon chambers and be classified as a muon candidate. To avoid the misidentification, the particle flow algorithm checks if there exists on ECAL or HCAL energy cluster that aligns with the muon track. Once the muons are identified, the clusters are removed from the particle flow blocks.

Particle flow electrons: When an electron traverses material, it can emit Bremsstrahulung photons and, if sufficiently energetic a photon can convert into an e^+e^- pair. These electron-positron pairs can again emit photons. Hence, electrons and isolated photons show similar properties in the tracker and ECAL and are reconstructed together in the particle flow algorithm. At first, a gaussian-sum filter (GSF) track is used to seed an electron candidate. The GSF track corresponds to the ECAL cluster which is linked with zero, one or two tracks. Any track with $P_T > 2$ GeV is used as a seed. A photon candidate is seeded from an the ECAL supercluster with transverse energy larger than 10 GeV and no links

to GSF track. For electron and photon candidates, the ratio of HCAL energy and 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 can emit photon when traversing tracker material, the total energy deposited in the ECAL is calculated in a supercluster with small η but wide ϕ variables as the emitted photons travel in the direction emitted while the magnetic field bends the electron in ϕ. To separate electrons from photons, an isolation criteria is used. There are several isolation techniques applied to improve the reconstruction efficiencies [78].

Particle flow photon: Photon reconstruction primarily relies on the ECAL using a pseudorapidity range of $|\eta| \leq 3.0$ with a small gap between barrel and endcap around $|\eta| = 1.479$. The particle flow algorithm uses the shape and size of the cluster in the ECAL to identify photons with $P_T >$ 15 GeV. The photon reconstruction consists of sequential steps as explained below

First the individual channel in the ECAL is calibrated using

1. the ϕ symmetry of the energy deposited by the pileup and underlying event,
2. the invariant mass of the decays of $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$, and
3. the momentum measured by the tracker for isolated electrons from W and Z boson decays.

The energy is collected from radiating electrons and converted photons to reconstruct a cluster. A cluster is built starting from a seed crystal which registers an the energy greater than neighboring crystals and exceeds a predefined threshold. The energy contained in a

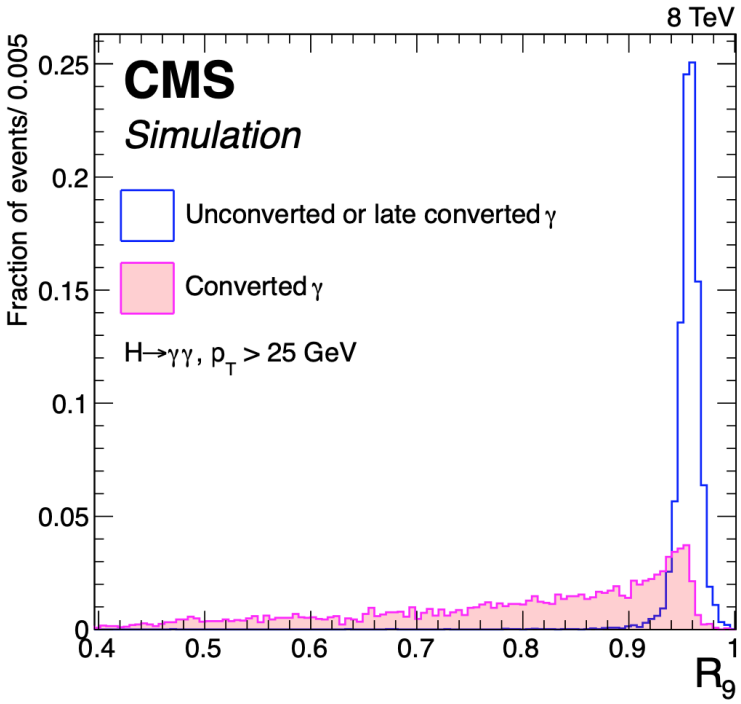


Figure 25: The distribution of R_9 for photons in the ECAL barrel. The red histogram represents the photons that converted in the tracker while the blue histogram represents the photons that did not convert or convert later before reaching the ECAL. The figure is taken from [79].

5 x 5 matrix around the seed crystal is used to reconstruct an unconverted photon. Several criteria are used to identify photons such as tracker isolation, ECAL isolation, hadron calorimeter isolation, hadronic to electromagnetic ratio and R_9. The R_9 variable is the ratio of the energy contained in 3x3 matrix and the supercluster energy, and behaves as shown in Figure 25. For example, R_9 values for π^0 and η meson decay will be lower than an isolated photon.

The raw energy E_{raw} obtained from the supercluster must be corrected to estimate the true energy, E_{true}. The correction values are determined as a function of η, E_T and R_9. In the analysis, we apply photon E_T smearing and scaling corrections. Photon energy scaling is applied to adjust the number of electrons from the di-electron decay of the Z boson.

The mismatch in the energy resolution between data and simulation is accommodate by smearing. A random Gaussian smearing is added to the photon energy resolution to smear the photon energy assuming the same degradation in resolution between data and MC events for photons.

Particle flow jets: Quarks and gluons hadronize to form colorless hadrons. High energy quarks and gluons, like those from a hard interaction, tend to hadronize into a narrow beam of particles called a jet. Once the isolated muons, electrons, and photons are reconstructed they are removed from the particle flow blocks. The blocks consist of the remaining hadrons, leptons, and photons. Jets are reconstructed using the information from the tracker, ECAL, and HCAL as they leave tracks while traversing the tracker material and deposit energy in the ECAL and HCAL. The charged hadrons leave tracks in the tracker, and deposited energy in both the ECAL and HCAL. First we identify the neutral hadrons and photons and remove them from the PF blocks. Then the remaining charged hadrons are reconstructed. CMS uses a sequential clustering algorithm to reconstruct jets as it is collinear safe, infrared safe and fast [80]. The anti-k_t jet clustering algorithm, an idealized cone algorithm, is used to reconstruct jets. For each pair of the particles i, j, and the distance d_{ij} is

$$d_{ij} = \min\left(\frac{1}{p_{Ti}^2}, \frac{1}{p_{Tj}^2}\right)\frac{\Delta R_{ij}^2}{R_{ij}}$$
$$d_{iB} = \frac{1}{p_{Ti}^2} \tag{4.1}$$

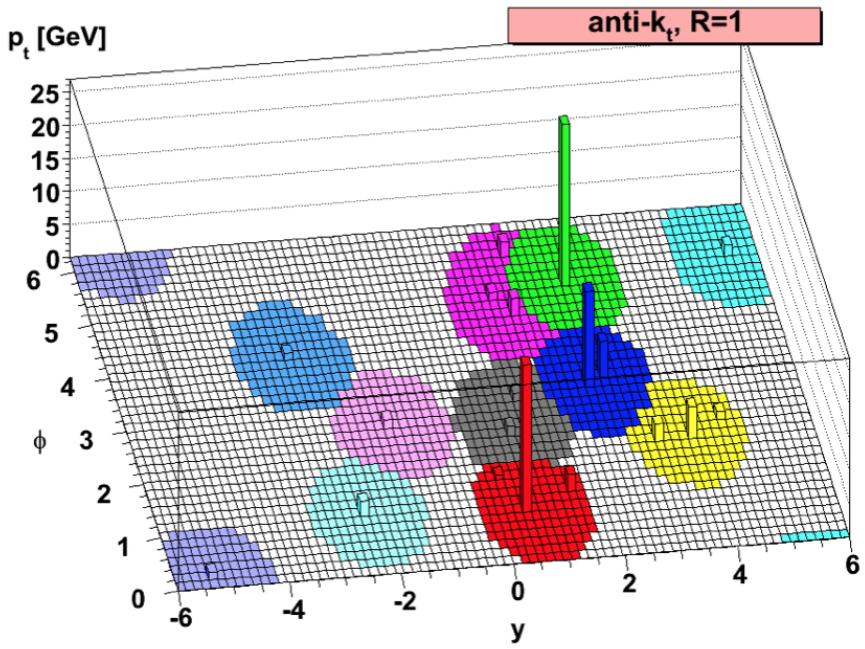


Figure 26: Reconstruction of jets in an event using the anti-k_t algorithm. The high P_T jets are circular and the boundary is defined by the respective hardness among neighboring jets. [80]

where, d_{iB} is distance of closest approach between particle i and the beamline. As the distance d_{ij} is calculated using the inverse of the particle transverse momenta, the high transverse momentum particles are used in clustering at first, then the soft particles are clustered together with nearby hard particles rather than nearby soft particles. The hard jets (jets with high P_T) reconstructed by the anti-k_t algorithm are exactly circular in the $y - \phi$ plane [81]. CMS uses a cone of radius R=0.4 for anti-k_t jet reconstruction. An examples of anti-k_t jet reconstruction is shown in Figure 26. Before jet reconstruction, any charged hadrons that match with a vertex other than the primary vertex or a secondary vertex of the event are removed from the jet. This is called charged hadron subtraction and it removes tracks from particles originating from other collisions. Once the particle flow jets are reconstructed it is necessary to apply jet energy corrections on data to remove

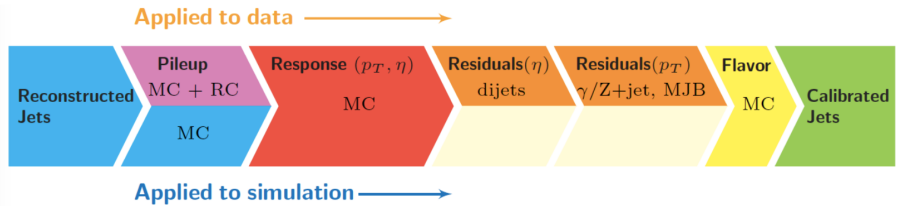


Figure 27: The corrections for a reconstructed jet are applied sequentially to MC and data to get the final calibrated jets. The sketch is taken from [77].

contamination and on simulation to match with the observed resolution in the data.

JET energy correction: To match the measured jet energy with the true parton energy, we apply jet energy corrections. CMS uses a factorized approach to apply jet energy corrections for both MC and data at each level sequentially as outlined in Figure 27. The L1 pile-up correction removes the energy contributed by pile-up events in a jet. The correction is applied to both data and simulation and is determined from simulation of QCD dijet events with and without pileup. The offset information is extracted as a function of ρ, jet area, jet η, and jet P_T. The corrected transverse momentum of the jet is

$$p_{T,j}^{L1,corr} = p_{T,j}^{uncorr} - PU_{\text{offset}}(p_T^{reco}, \eta, \rho, A). \tag{4.2}$$

where ρ is energy density. After the L1 pileup correction, we correct the jets for uniform response over jet P_T and jet η, called the L2L3 response correction. The response correction is determined by comparing the reconstructed P_T to the particle level P_T of a jet in dijet QCD simulation. The jet energy response scale factor (S_f) is defined as

$$S_f = \frac{p_{T,j}^{L1,corr}}{p_{T,GenJet}^{L1,corr}}. \tag{4.3}$$

Next we correct the residual differences between jet response in data and MC. Part is an η dependent correction and part is P_T dependent. The residual correction is applied only to data.

Jet energy resolution(JER): The uncertainty of a measured jet energy in the data is higher than in the MC samples. The jet energy in MC is smeared to match with the jet energy resolution in data [82].

4.2.5 Higher Level Observable

In pp collisions, a jet can originate from b, c, s, u, and d quarks and are classified into b jets, c jets, and light jets. A b jet contains at least one bottom-hadron, c jets contains at least one charm-hadron and no bottom hardron, and the rest of the jets are light jets [83]. CMS uses a b-tagging algorithm to identify jets originating from a b-quark. The b-quark can be distinguished from the light and c-quarks by its larger mass (4.2 GeV) and typical lifetime of 1.5 psec. Due to the 1.5 psec lifetime, b hadrons travel a few millimeters from the primary vertex before decaying. The decay vertex of the b hadron is called the secondary vertex. Hence the b jet originates from the primary vertex and has a secondary vertex within the jet. This is used by the b-tagging algorithm to infer the presence of a b-jet in the CMS detector. Figure 28 shows schematically the signature of a b jet in a pp collision.

The sum of all transverse momenta in a pp collision is zero due to conservation of momentum. But due to the presence of invisible and undetectable particles, there can

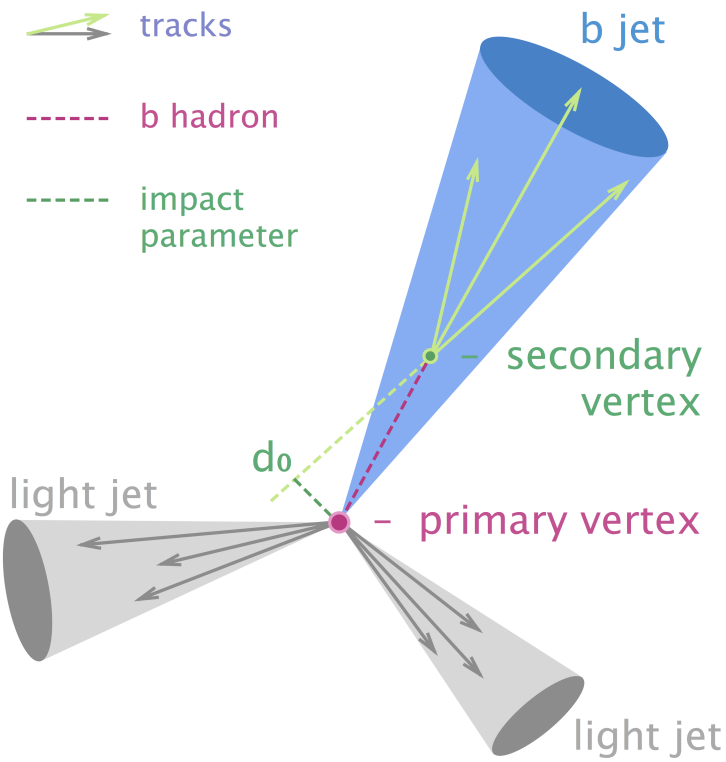


Figure 28: The signature of b-quark jet that originates from the primary vertex. It has a secondary vertex with transverse impact d_0 and a decay length indicated by the red dashed line. The picture is taken from [84]

exist an imbalance in the sum of P_T of all visible particles which is known as missing transverse energy (MET). MET is calculated [85] as

$$\vec{\not{E}}_T = -\sum_{i=1}^{n} \vec{p}_{T,i} \tag{4.4}$$

where n is the total number of reconstructed particles and $\vec{p}_{T,i}$ is the transverse momentum of the i^{th} final state particle. The computation of MET depends on the application of jet energy correction (JEC) during PF object reconstruction. If an event has an usually large missing P_T, the PF algorithm is applied for a post processing such as checking

for cosmic muons, poorly reconstructed muons or overlapping muon and neutral hadron misreconstructed as a charged hadron.

The lepton and photon isolation criteria are applied during their reconstruction. We apply the PF relative isolation (I_{rel}) to the PF objects to have strict criteria which are often required in physics analysis. The PF relative isolation is defined as the ratio of the sum of the P_T of all the particles around the lepton within a cone to the P_T of the lepton. The charged particles from pile-up are removed but neutral particles cannot be removed directly. So we estimate the ratio of charged hadron to neutral hadron (plus photon) P_T in inelastic collisions to estimate the neutral contribution from pileup. This is known as the $\Delta\beta$ correction. The PF relative isolation is

$$I^{l}_{rel} = \frac{\sum p_T^{charged} + \max(0, \sum p_T^{neutral} + \sum p_T^{\gamma} - \Delta\beta \sum_{pileup} p_T^{charged})}{p^{l}_{T}} \tag{4.5}$$

where p_T^{γ} is the transverse momentum of a photon, $p_T^{charged}$ is the transverse momemtum of a charged particle, $p_T^{neutral}$ is the transverse momentum of a neutral particle, and $\Delta\beta$ is beta correction. The lepton and photon reconstruction algorithm requires a basic level of isolation. A stricter relative isolation (I_{rel}) requirement is applied to leptons and photons selected for the analysis.

4.3 CMS Data Tier:

The raw data recorded by the CMS data acquisition system undergoes successive stages of processing to refine the data, apply calibrations, and create high level physics objects for analysis purposes. Event information in the simulation and reconstruction chain is

grouped into what is called a data tier [86]. A data set may consist of multiple data tiers such as analysis object data (AOD), miniAOD, and nanoAOD.

There are several data tiers in CMS. The first is zero suppressed raw data, the standard raw data event format. The size of the RAW data is about 1MB per event. The second data tier is the reconstructed (RECO) data which contains the detailed information about the reconstructed physics objects and has a size of about 3 MB per event. The analysis object data (AOD) format contains physics objects used in analyses (400-500kB per event).

MiniAOD is a high level data tier which has a size of 30-50 kB per event. MiniAOD format is produced for real and simulated events. It has a very large flexibility as it includes all observed particles. It allows us to do reclustering of jets or re-calibration of jet energies and helps to design new identification techniques. The nanoAOD data format has a size of 1-2kB per event. It consists of all physics objects, trigger information, generator information, event weight, cleaning flags, primary and secondary vertex, isolated tracks, boosted jets, and their substructures. It contains the top level information typically used in a physics analysis [87]. We skimmed nanoAODs to produce ntuples of selected events with a format specialized for our analysis.

4.3.1 The production of ntuples from the nanoAOD

The Monte Carlo events are generated with a pile-up distribution, however the distribution may differ from that of data. We determine the average number of pile-up in data and using Poisson statistics and adjust the number in the corresponding simulation to match. Next, we select only the events that passed the HLT triggers listed in Table 12 and recommended Run 2 MET filters by JETMETPOG [88]. The MET filter checks if an event has passed the following filter requirements

1. A good primary vertex filter: The definition of a good primary vertex as described in Section 5.2.

2. Halo beam filter: When a proton interacts with gas or beampipe material it induces particles traveling parallel with the beam but at large radius, up to 5m, are called halo particles. For example, a high energy halo muon can interact in the calorimeter and effect jet+met, photon+met and trigger efficiency.

3. HCAL barrel and endcap (HBHE) noise filter and HBHE isolated noise filter: These filters are applied to remove events that have noise signals originating from HCAL.

4. ECAL trigger primitive filter: When the energy based on the trigger primitives (TP) is near the saturation energy[4], the recovered energy is likely underestimated, leading to high MET events. The ECAL trigger primitive filter is used to tag these events and remove them from analysis.

5. Bad particle flow muon filter: Events where a muon is low quality and therefore not declared a PF muon, and this non-PF muon makes it into the PF-MET calculation as a charged hadron candidate, are removed.

All the selected events from the nanoAOD are saved in an ntuple files containing a selected subset of information to further compress the data.

[4]The trigger primitive (TP) information saturates above 127.5 GeV [89].

CHAPTER 5 EVENT AND OBJECT SELECTION

Events are reconstructed using the CMS software (CMSSW); the 80X series of CMSSW is used for the year 2016, 94X for the year 2017, and 102X for the year 2018. However, the nanoAOD data tier is produced with 102X in all years. Signal region events are required to pass the trigger (either single electron or single muon), to have a well defined primary vertex, with exactly one high P_T and well isolated lepton (electron or muon), with $\geq$4 jets and $\geq$ 1 b-tagged jet, and exactly one high P_T photon. Control region events are required to pass the trigger (either single electron or single muon), to have a well defined primary vertex, with exactly one high P_T and well isolated lepton (electron or muon), with any number of jets and/or any number of b-tagged jets, and with or without a photon. The following trigger selection, object definitions, and event selections are applied.

5.1 Trigger Selection

Unprescaled trigger with the lowest thresholds are used in both data and simulation, for all data taking periods as listed on the `TopTriggerTwiki` [90]. Each event must pass either the single electron trigger or single muon trigger for different data periods. For the electron channel, the high level trigger selection requires an electron with $P_T \geq$ 27 GeV, $P_T \geq$ 32 GeV, and $P_T \geq$ 32 GeV for 2016, 2017, and 2018 data respectively, and within $|\eta| \leq 2.4$. For the muon channel, the high level trigger selection requires an isolated muon with $P_T \geq$ 24 GeV, $P_T \geq$ 27 GeV, and $P_T \geq$ 24 GeV for 2016, 2017, and 2018 data respectively, and within $|\eta| \leq 2.4$. A detailed object selection criteria are explained in Section 5.2. The list of triggers for 2016, 2017, and 2018 data taking eras is shown in Table 12.

Table 12: Trigger requirements from the `TopTriggerTwiki` [90].

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*`	

5.2 Object Selection

This precision analysis is based on the requirement of well-identified physics objects such as jets, b-tagged jets, and leptons (electrons or muons). The events are selected on the basis of such objects in the signal and control regions. A details of the object selection are discussed below.

5.2.1 Primary Vertex

The vertex with the largest scalar value of summed physics-object P_T^2 is called the primary vertex of the event. Primary vertices are reconstructed by clustering tracks together and performing fits to determine the maximum likelihood that they originated from a common vertex. These tracks must satisfy

- minimum number of degrees of freedom > 4,
- longitudinal vertex position $|z| < 24$ cm, and
- track transverse position $d_0 < 2$ cm.

Tracks that are pointing to other vertices are considered as pile-up.

5.2.2 Electrons

The baseline for the electron selection is the "cut-based" tight ID [91], with transverse momentum $P_T \geq 35$ GeV and $|\eta| \leq 2.4$. The relative isolation requirement is a part of the cut-based ID, so separate isolation criteria are not necessary for electrons. Electrons must not be in the EB-EE [5] ECAL gap ($|\eta|_{SC} \leq 1.4442$ or $|\eta|_{SC} \geq 1.566$) and must satisfy the impact parameter requirements $d_{xy} \leq 0.05$ (0.1) in EB (EE); $d_z \leq 0.1$ (0.2) in EB (EE). The POG [6] veto IDs are used with $P_T \geq 15$GeV and $|\eta| \geq 2.4$ to veto any extra lepton in the event. This requirement suppresses dileptonic $t\bar{t}\gamma$, $Z + \gamma$ and Z+jets events. A detailed CMS electron ID selection requirements are given in Table 13 for 2016 data. The requirements for 2017 and 2018 can be found at [91].

Table 13: Electron ID requirements

Requirements	Tight	Veto
Full 5x5 $\sigma_{i\eta i\eta} \leq$	0.0104	0.0128
abs($\Delta\eta$ seed) $\leq$	0.00353	0.00523
abs(dPhiIn) $\leq$	0.0499	0.159
C_0 [7]	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

Figures 29 and 30 show electron kinematics plots for 2016, 2017, and 2018 data taking periods. A detailed selection criteria are displayed in Table 15.

[5] EB = Electromagnetic Barrel, EE = Electromagnetic Endcap

[6] POG = Physics Object Group

[7] H/E < C_0 + 1.12/E + 0.0368*rho/E.

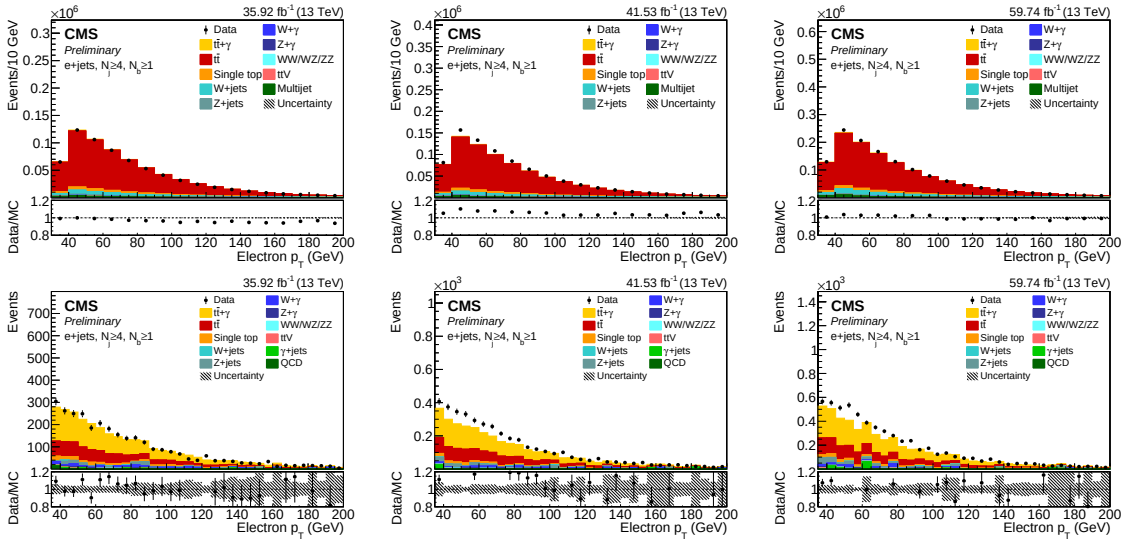


Figure 29: Electron P_T distributions for $t\bar{t}$ (top) and $t\bar{t}\gamma$ (bottom) selections in $e+$jets for 2016, 2017, and 2018 data.

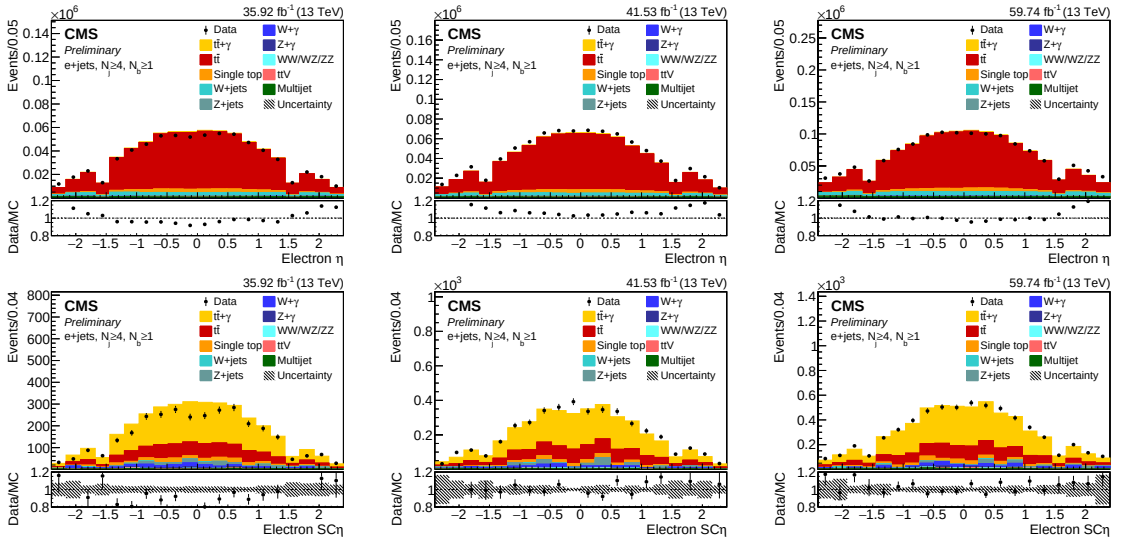


Figure 30: Electron super cluster η distributions for $t\bar{t}$ (top) and $t\bar{t}\gamma$ (bottom) selections in $e+$jets for 2016, 2017, and 2018 data.

5.2.3 Muons

The baseline for the muon selection is the tight POG ID [92] with a transverse momentum requirement of $\mathrm{P_T} \geq 30$ GeV and $|\eta| \leq 2.4$. The relative isolation requirement is not a part of the cut-based ID for muons, and is implemented separately using Particle Flow (PF) relative isolation with a cone size of 0.4 with $\Delta\beta$ corrections ($I_{rel}^{0.4}$) for PU mitigation. The muon relative isolation variable is defined as ratio of a sum of transverse energies (momenta) of all particles in a cone of $\Delta\mathrm{R} \leq 0.4$

$$I_{rel} = \frac{\Sigma \mathrm{P_T}\,(\text{ch. had. from PV}) + max(0, \Sigma E_{\mathrm{T}}\,(\text{neu. had.}) + \Sigma E_{\mathrm{T}}\,(\text{pho}) - 0.5 * \Sigma \mathrm{P_T}\,(\text{ch. had. from PU}))}{\mathrm{P_T}\,(\mu)} \tag{5.1}$$

where, ch. had. means charged hadrons, neu. had. means neutral hadron, and pho means photon. All the selected muons must satisfy the relative isolation cut of 0.15. The POG veto ID requirements with $\mathrm{P_T} \geq 15$ GeV and $|\eta| \leq 2.4$ are applied. The $I_{rel}^{0.4}$ cut for loose muons is relaxed to 0.25 to veto events with an extra muon. This requirement suppresses the dimuonic $\mathrm{t\bar{t}}\gamma$, $Z + \gamma$ and Z+jets events. A detail selection criteria are provided in Table 15. A detailed muon ID selection requirements are listed in Table 14 for 2016 data. The similar tables for 2017 and 2018 can be found at [92].

5.2.4 Photons

Photons must pass the POG MediumID cut for photons [93]. In addition, $\mathrm{P_T} \geq 20$ GeV and $|\eta| \leq 1.4442$ are required. The POG veto ID requirement of $E_{\mathrm{T}} \geq 20$ GeV and $|\eta| \leq 1.4442$ are applied. The veto ID includes a medium photon ID without requirements on $\sigma_{i\eta i\eta}$ and charge isolation and have no track seed in the pixel detector. A detailed selection criteria are listed in Table 17. A detail medium photon ID selection requirements in the barrel region are listed in Table 16 for 2016, 2017, and 2018 data [91].

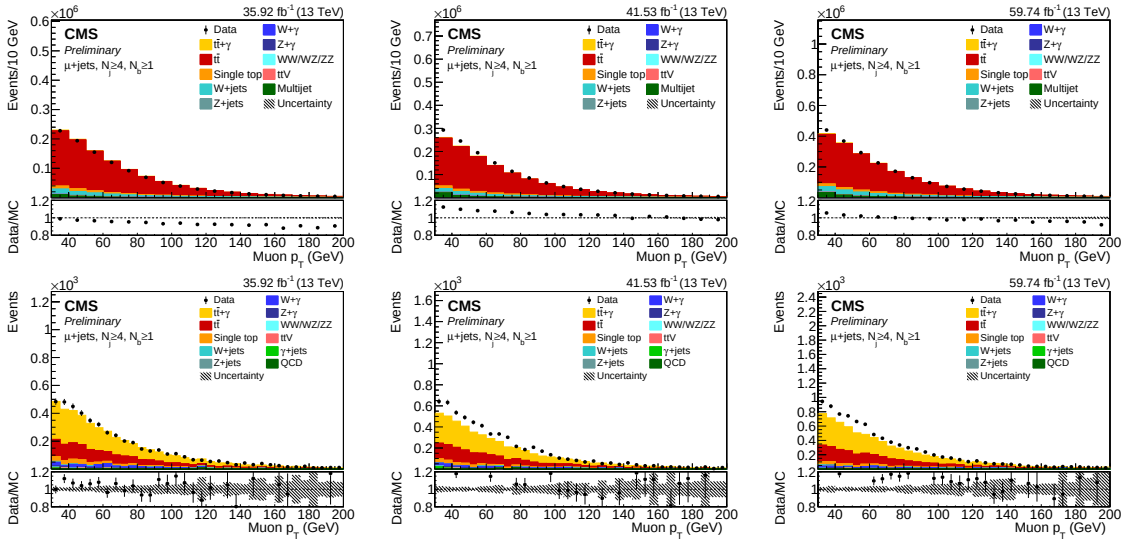


Figure 31: Muon P_T distribution for $t\bar{t}$ (top) and $t\bar{t}\gamma$ (bottom) selections in $e+$jets for 2016 (left), 2017 (center), and 2018 (right) data.

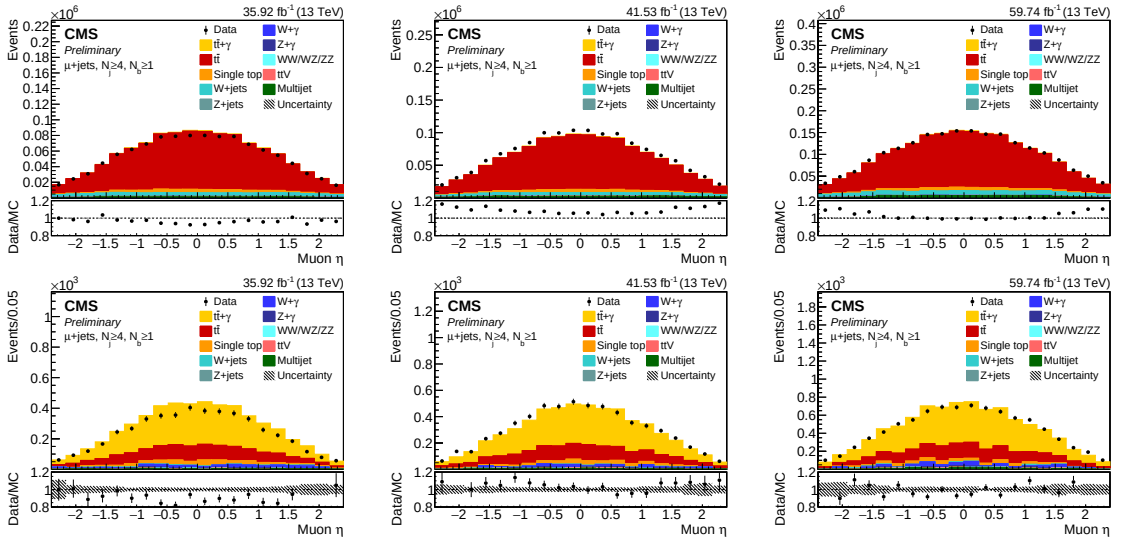


Figure 32: Muon η distribution for $t\bar{t}$ (top) and $t\bar{t}\gamma$ (bottom) selections in $e+$jets for 2016 (left), 2017 (center), and 2018 (right) data.

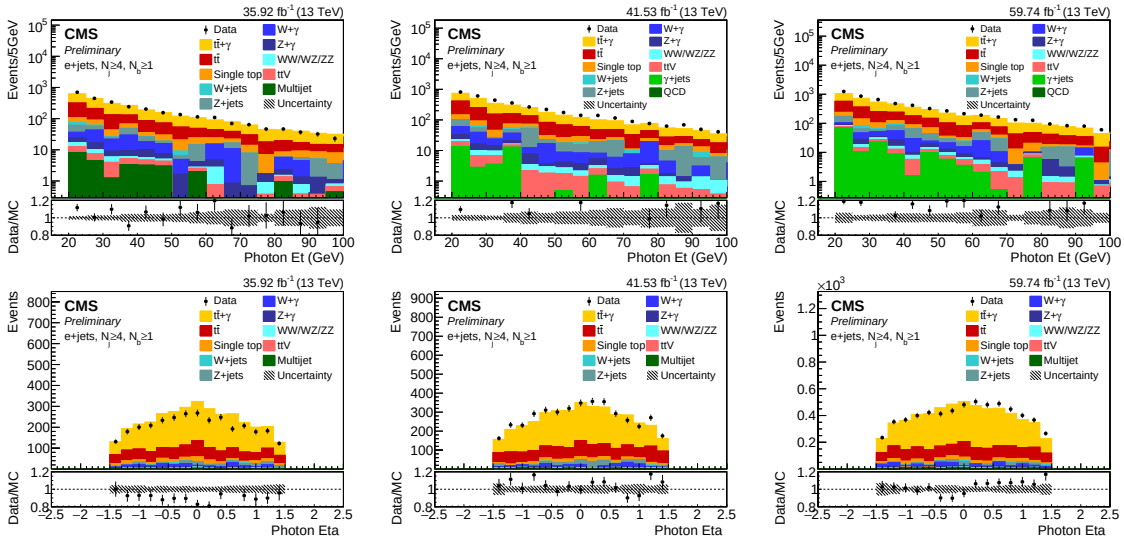


Figure 33: Leading photon E_{T} (top) and η (bottom) distribution for signal region in $e+$jets for 2016, 2017, and 2018 data.

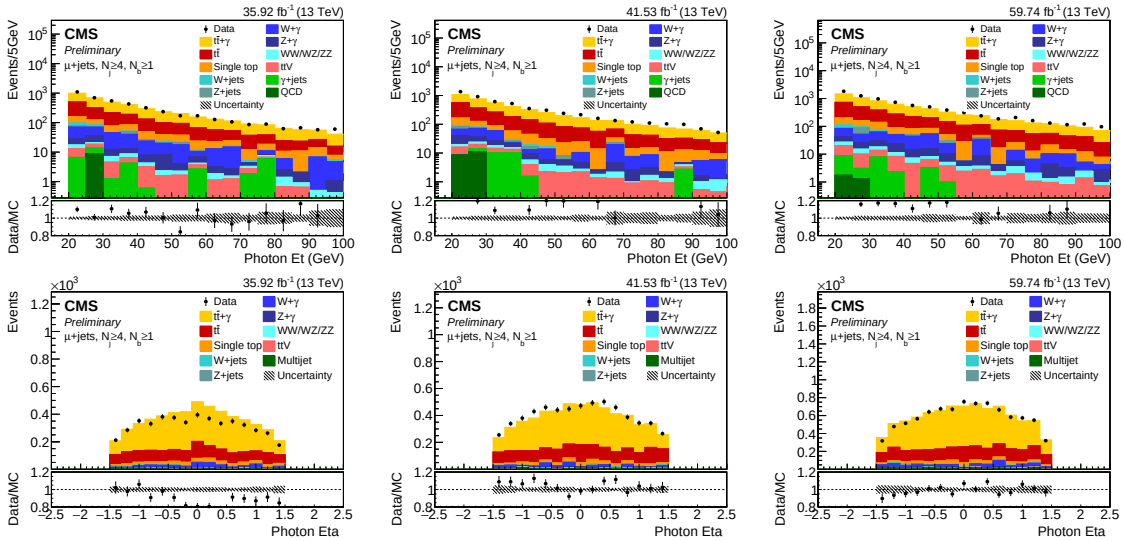


Figure 34: Leading photon E_{T} (top) and η (bottom) distribution for signal region in $\mu+$jets for 2016, 2017, and 2018 data.

Table 14: Muon ID selection criteria

Selection criteria	TightID	LooseID
Reconstructed Global Muon	yes	-
PF Muon ID	yes	yes
Global or Tracker Muon	-	yes
normalized χ^2 of the global-muon track fit	< 10	-
muon-chamber hit included in the global-muon track fit	> 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	-
Cut on number of tracker layers with hits	> 5	-

Table 15: Lepton identification requirement based on the cut-based tight IDs for electrons [91] and muons [92].

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

5.2.5 Jets and b-tagging

Several jet multiplicity requirements are used in the analysis. In the signal region, there must be at least four jets out of which at least one must be b-tagged [8]. The jets must have at least two constituents, where at least one must be charged. The electromagnetic and hadronic energy fractions must both be smaller than 99%, to reduce misidentified jets from photons and pions respectively.

[8] "b-tagging" is defined as tagging of b jets arising from the process of hadronization of b quarks

Table 16: The detail of photon ID requirements

Variables	values
H/E	< 0.027
$\sigma_{i\eta i\eta}$	< 0.01015
ρ-corr charged hadron isoltion	< 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 isolation	$2.08 + 0.004017 \times p_T(\gamma)$
has Pixel seed veto	yes

Table 17: Photon identification requirements for the cut-based medium ID [93] .

Observable	Photon Selection	
	Nominal ID	Veto ID
P_T (GeV)	$\geq$ 20	$\geq$ 20
$\lvert\eta\rvert$	$\leq$ 1.4442 (barrel)	$\leq$ 1.4442 (barrel)
	cut-based medium ID	cut-based medium ID
	no pixelSeed	w/o requirements on $\sigma_{i\eta i\eta}$ and chg.Iso
		no pixelSeed

We use (standard) AK4 PF CHS[9] jets that pass the ID requirements listed in Table 18 with $P_T \geq 30$ GeV and $|\eta| \leq 2.4$. The difference between loose ID and tight ID jets are described in Table 19

Jets identified as b-tagged must pass the same criteria as the other jets and additionally pass the medium working point cut of the Deep CSV discriminator of greater than 0.6321, 0.4941, and 0.4184 for the years 2016, 2017, and 2018, respectively, as specified by the BTV POG [94]. Details and DeepCSV thresholds for the other years are summarized in Table 18. The jet energy scale is corrected using the centrally produced jet energy corrections (`Summer16_07Aug2017_V11`, `Fall17_17Nov2017_V32`, and `Autumn18_V19`) and these corrections are propagated to MET (type-1 corrections).

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

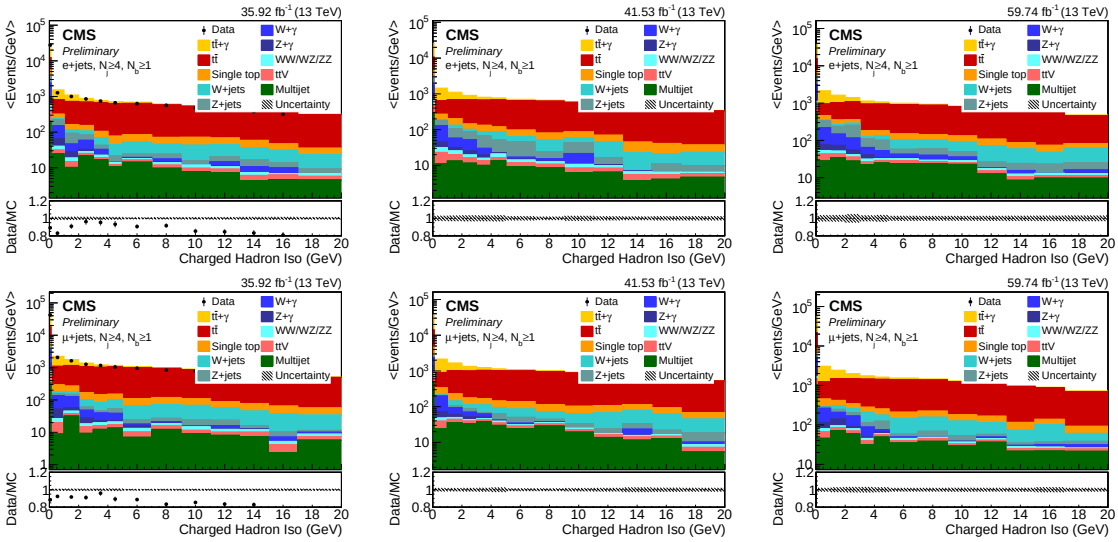


Figure 35: Leading photon charged hadron isolation (ChIso) distribution in $e+$jets (top) and $\mu+$jets (bottom) for signal region for 2016, 2017, and 2018 data.

Table 18: Jet identification criteria and selection thresholds.

Observable	Nominal ID
P_T (GeV)	≥ 30
$\lvert\eta\rvert$	≤ 2.4
ID	loose PF ID (2016), tight PF ID (2017/2018)
b-Tag (DeepCSV)	≥ 0.6321 / 0.4941 / 0.4184 (2016/2017/2018)

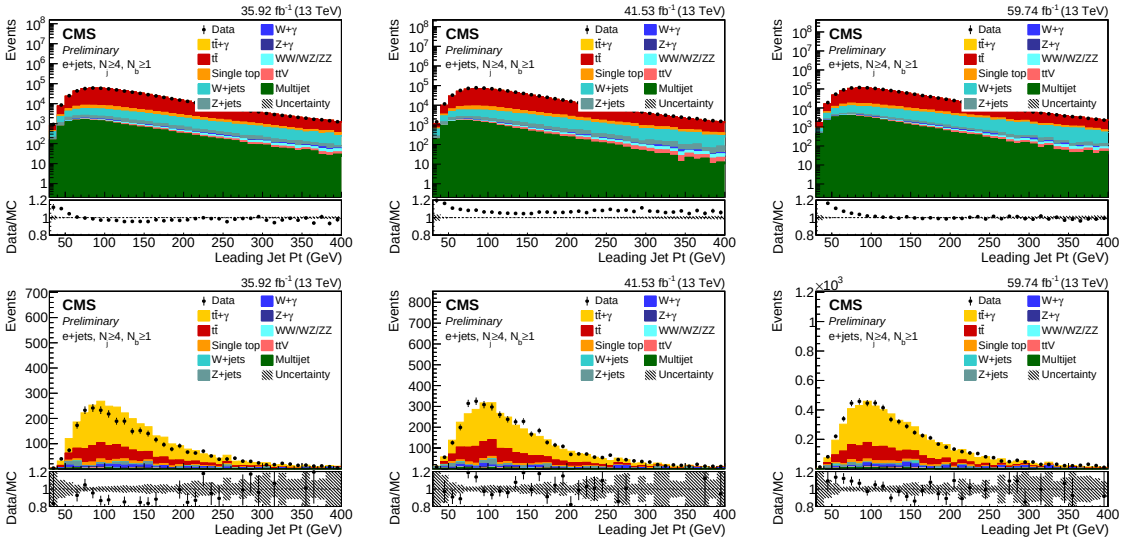


Figure 36: Jet P_T distributions for $t\bar{t}$ (top) and $t\bar{t}\gamma$ (bottom) selections in $e+$jets for 2016, 2017, and 2018 data.

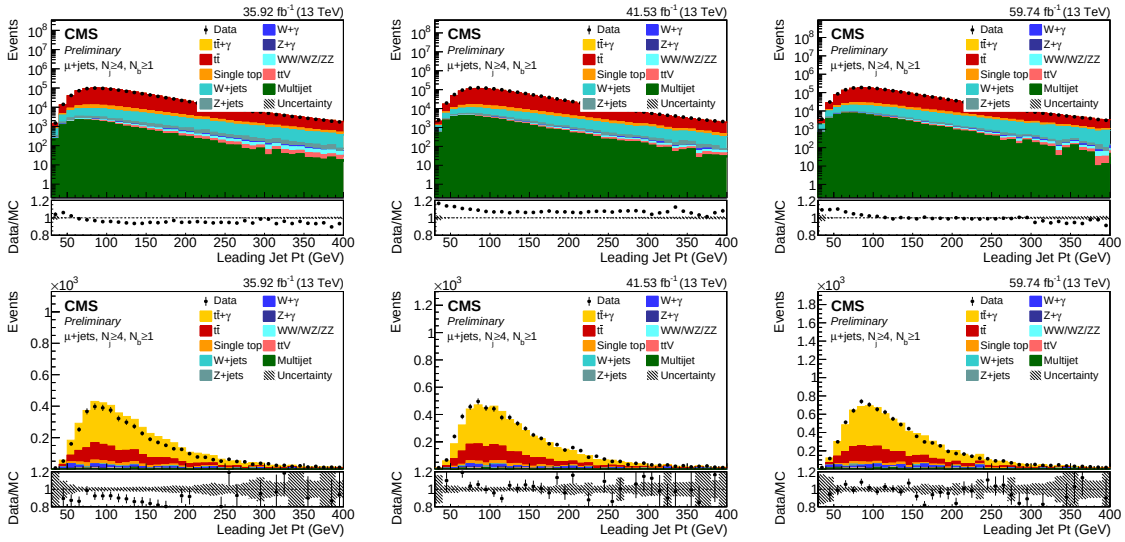


Figure 37: Jet P_T distributions for $t\bar{t}$ (top) and $t\bar{t}\gamma$ (bottom) selections in $\mu +$ jets for 2016, 2017, and 2018 data.

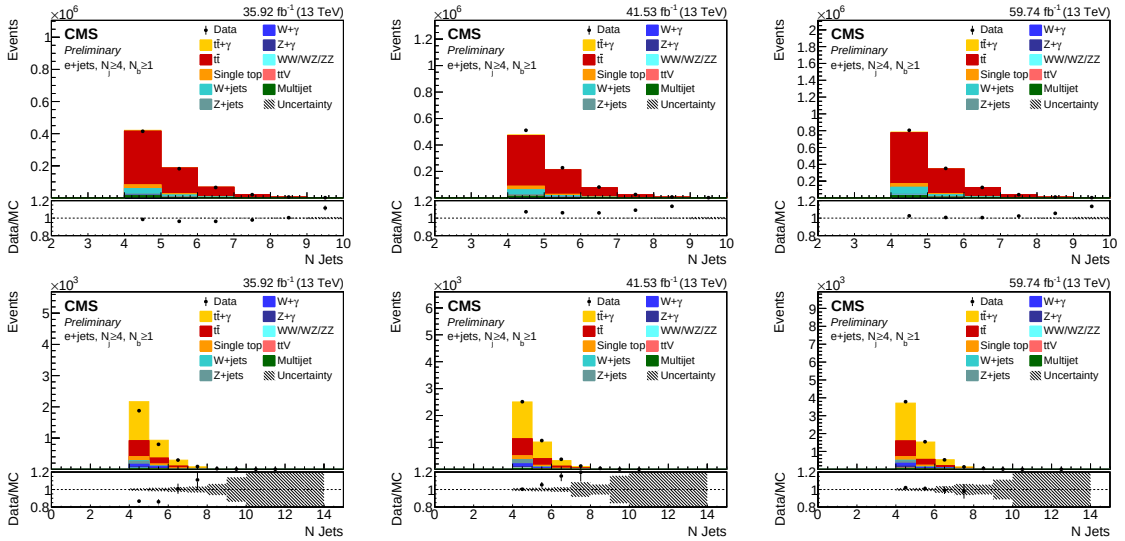


Figure 38: Leading jet η distributions for $t\bar{t}$ (top) and $t\bar{t}\gamma$ (bottom) event selections in $e +$ jets for 2016, 2017, and 2018 data.

Table 19: Jet loose and tight ID requirements for $|\eta| \leq 2.7$

Variables	Loose Jet	tight jet
Neutral hadron fraction	< 0.99	< 0.90
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

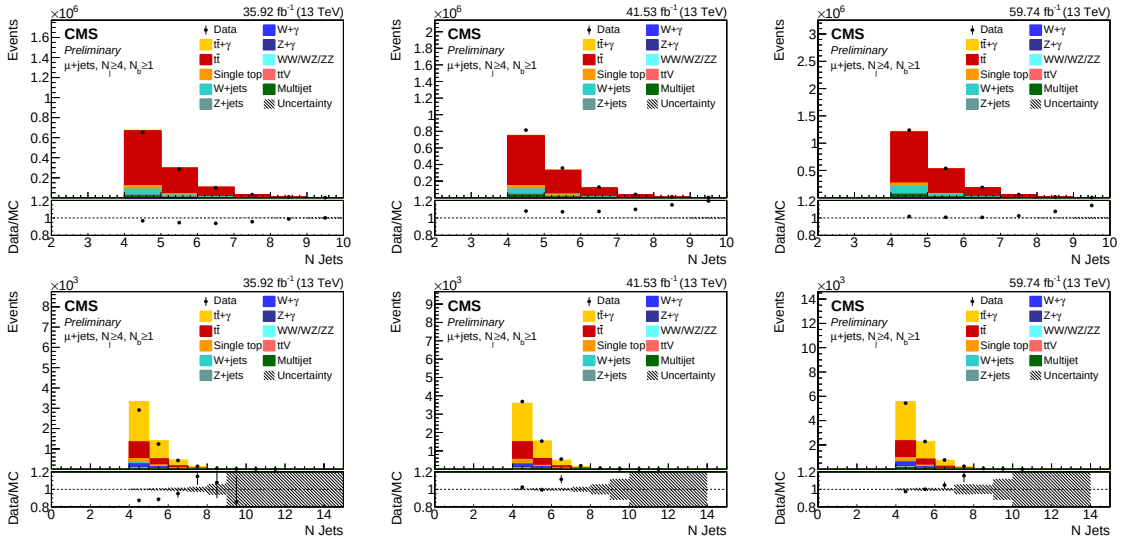


Figure 39: Leading jet η distributions for $\mathrm{t\bar{t}}$ (top) and $\mathrm{t\bar{t}}\gamma$ (bottom) event selections in $\mu+$jets for 2016, 2017, and 2018 data.

5.2.6 Missing Energy

Standard PF missing transverse energy (MET) is used and corrected by type-1 corrections. We used the MET observable to form m_{T} in the QCD estimation procedure in this analysis. However, there is not a minimum MET requirement for signal event selection. There exist a large MET variation due to the jet energy uncertainties. For reference, more plots comparing data and simulation are included in Appendix E.

5.2.7 Object Disambiguation

The electron, muon, photon, and jet object collections are not necessarily exclusive. The following cleaning procedure is applied sequentially to the objects satisfying the above ID criteria.

1. Photons are removed if there is a tight lepton within $\Delta\mathrm{R} < 0.4$. This requirement removes photons radiated from leptons.

2. Jets are removed if there is a tight lepton within $\Delta\mathrm{R} < 0.4$. This requirement removes jets that are clustered from prompt leptons.

3. Jets are removed if there is a photon within $\Delta\mathrm{R} < 0.1$. This requirement removes jets that are clustered from prompt photons.

5.3 Corrections at the Object and Event Level

We apply corrections to several observables in both data and simulation. This is essential to correct for biases in detector response, mismodeling of physical observables in simulation, and to ensure that the simulation correctly models the key features of the signal and background processes. In most cases, the background is determined in a data driven manner, nevertheless the following corrections are applied.

5.3.1 Pileup Reweighting

Monte Carlo events are generated according to an assumed pileup profile and are reweighted to match the pileup profile in data. The distribution of the true number of interactions in data is estimated using the instantaneous luminosity per bunch crossing for each luminosity section and assuming a total pp inelastic cross section of 69.2 mb [95].

A variation of $\pm 4.6\%$ on this cross section is used to estimate the uncertainties due to the pile-up modeling. The true number of interactions in simulation is taken from each simulated sample and reweighted to the estimated distribution in data separately in each data taking period. All the experimental corrections for muons, electrons, jets, and b-jets have been derived applying only the official pileup reweighting, hence they should already take into account the effect of residual differences in pileup between data and simulation.

5.3.2 L1 EGM Prefiring Correction in 2016 and 2017

The time alignment of the endcap ECAL readout electronics drifted during the data-taking, reaching a severe condition in 2017 Run F. The L1 trigger system of CMS has been designed to forbid triggering events in consecutive bunch crossings [89]. The trigger primitives generated by the ECAL deposits and reconstructed as belonging to the previous bunch crossings are therefore able to generate an inefficiency in the L1 trigger decision. It is found that L1 prefiring issues effect events with high-P_T jets with $2.4 < |\eta| < 3.0$. We don't use reconstructed objects in that η region directly, however additional objects from the event may be affected by this issue [96]. Prefiring inefficiency is measured using a special set of triggered events called "un-prefirable" events. This inefficiency has been measured by the JME POG[10] as a function of both photon and jet P_T in the forward η region of endcap, $2.5 < |\eta| < 3.0$. The event efficiency factor ($\varepsilon_j = 1 - \mathcal{P}$) is applied to the MC simulation and it is accounted using all jets in the event, as

$$\varepsilon = \prod_{j \in \text{jets}} 1. - \mathcal{P}_{\text{prefire}}(j)$$

[10] Jet and missing energy physics object group

A detail explanation of the L1 trigger prefiring can be found at [97].

5.3.3 HEM Correction in 2018

In 2018, one end of the hadronic endcap calorimeter (HEM) lost power during the data-taking period [98]. To correct the HEM effect, we veto events for run numbers after 319077 if there is an electron or photon with $P_T \geq 15$, $-3 < \eta < -1.4$, and $-1.57 < \phi < -0.87$. This removes spurious EGM objects formed from the ECAL deposit of hadronic activity in the absence of the corresponding HCAL signal in the affected region. In 2018 simulation, events with an object that satisfy these criteria are weighted by the luminosity fraction (0.3518) of the affected data taking period.

5.3.4 Object Level Scale Factors

The reconstruction efficiency of the reconstructed objects is not perfectly simulated. Suitable scale factors are applied to account for the respective differences.

- b-tagging: The simulated efficiencies for the identification of b-quark jets and for the misidentification of c-quark, light-quark or gluon jets of the DeepCSV tagger are corrected using scale factors provided by the BTV POG and applied using method 1a [94].

- Photon efficiency: Corrections of the simulated efficiencies for photon identification and pixel seed veto [99] for the medium EGM POG WP are applied to simulated samples. Specifically, the `Fall18V2` 102X 2D ID efficiency map is used for 2018 data, the 94X 2D map for the year 2017 and the 94X map for the 2016 legacy rereco. The PixelSeed Veto SF is applied for each year as well: `ScalingFactors_80X_Summer16` and `PixelSeed_ScaleFactors_2017` (inclusive in R9) [99]. For 2018 the measurement of the PixelSeed Veto scale factors were recreated and are in good agreement

with the centrally provided scale factors (`HasPix_2018`).

- Electron efficiency: Corrections of the simulated efficiencies for electron trigger, identification, and isolation [100] for the tight EGM POG WP are applied to simulated samples. Again, the `Fall18V2` 102X 2D efficiency maps for reco and ID are used for 2018 data, the 94X 2D reco and ID maps are used for the year 2017, and the 94X reco and ID maps are used for the 2016 legacy rereco.

- Photon and 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. [101]. In contrast to the identification efficiency systematics, the $p_T(\gamma)$ scale uncertainties have negligible impact on the measured cross section.

- Muon efficiency: The simulated efficiencies for the trigger, identification, and isolation of muons are corrected in simulated samples using the MUO POG efficiencies [102–104]. The 2D scale factors are factorized in ID, isolation, and trigger efficiency. An extra uncertainty of 0.5% is added to cover extrapolation effects of muon ID uncertainties [105].

All scale factors are provided with systematic uncertainties which are taken into account and propagated to the predicted yields.

CHAPTER 6 ANALYSIS STRATEGY

The measurement is performed in $e+$ jets and $\mu+$ jets channels, requiring one isolated lepton (electron or muon), at least four jets out of which at least one must be b-tagged and the presence of an isolated hard photon with $\mathrm{P_T} \geq 20$ GeV in the final state, the selection for the inclusive $\mathrm{t\bar{t}}\gamma$ event candidates. However, the signal region is contaminated with irreducible backgrounds such as semi-leptonically decaying $\mathrm{t\bar{t}}$ events with an isolated photons or a nonprompt photons. For instance, the isolated photon may arise from a misidentified electron or from any other processes (or fake process) that is not from $\mathrm{t\bar{t}}$ vertex (such as misidentified photon, a Bremsstrahlung photon or photon from hadronic activity). Other backgrounds arise from the production of a leptonically decaying W/Z + γ with jets in the final state. Figure 40 shows how the major backgrounds are constrained during the extraction of the $\mathrm{t\bar{t}}\gamma$ cross section value. The final state selection includes top and non-top events with genuine and nonprompt photons. To differentiate top vs non top events, the M_3 observable is used. The M_3 observable is the invariant mass of 3 jets (out of which at least one must be b-tagged) with the highest summed $\mathrm{P_T}$. Shape of distribution of the M_3 observable for top events has a peak at the top mass (172.5 GeV). Conversely, non-top events have a flatter distribution throughout as shown in Figure 40 [left box]. To differentiate events with a genuine from a non prompt[11] photon, the photon charged hadron isolation (ChIso) variable is used. ChIso is difference between photon $\mathrm{P_T}$ and the sum of the transverse momenta of all the charged hadrons that lies within a cone of radius $\Delta\mathrm{R} \leq 0.4$ about the photon. The events with genuine[12] photon have a negative exponen-

[11] For this analysis, nonprompt represents photons from hadronic activities, fake photons, or photons from PU events.

[12] For this analysis, genuine represents the prompt isolated photon and misidentified electrons.

tial shape as shown in the left plot in bottom box in Figure 40. On the other hand, events with hadronic or fake photons have a wider shape with little dip and then slight bump as shown in the right plot in the bottom box in Figure 40.

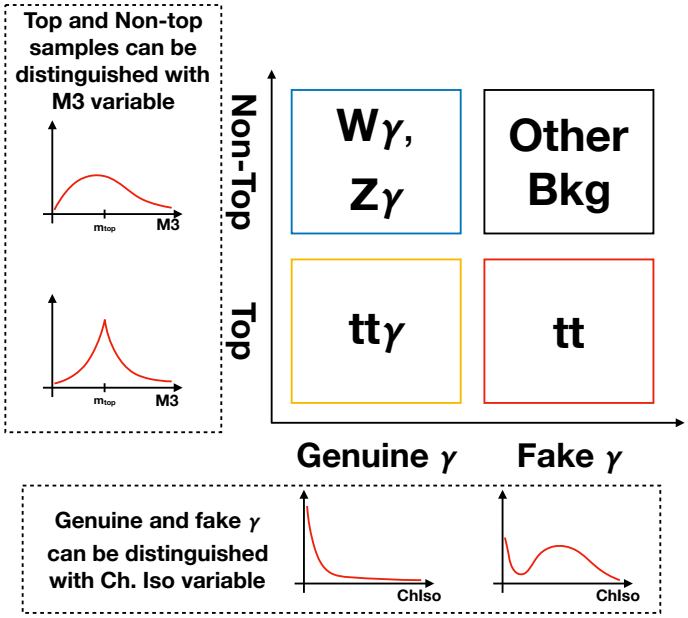


Figure 40: Analysis strategy for the separation of $t\bar{t}\gamma$ from backgrounds.

Before explaining the fit procedures, it is worth while to explore the basic properties of photon $p_T(\gamma)$ to understand the major backgrounds. In Figure 41 we show the $p_T(\gamma)$ spectrum in terms of simulated samples (left) and in photon categories (right). The genuine-γ contribution in Figure 41 (right) is larger than the signal contribution ($t\bar{t}\gamma$ in Figure 41 (left)), because a contribution from $t\bar{t}$, tWγ and tγ is included in the latter.

First, we normalize the $Z+$jets cross section by fitting a data in dilepton selection region satisfying all $t\bar{t}$ selection criteria. The detail explanation is found in subsection 6.5.1. Although QCD is a minor background in our final signal region, it has significant effects while we extract Z/W+γ, misidentified electrons and $Z+$jets . A QCD data driven tem-

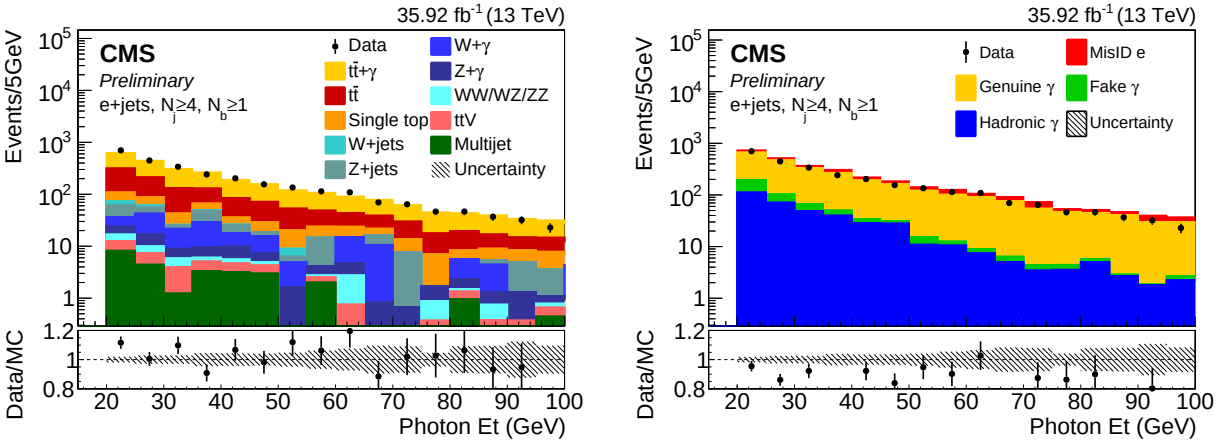


Figure 41: Photon P_T distribution for $e+$ jets channel in the signal region, $N_j \geq 4, N_b \geq 1$, for 2016 data categorized in terms of simulated samples (left) and in photon categories(right).

plate is extracted from side band region of high relative isolation distribution of M_T with 0-btag control region. The detail explanation is found in subsection 6.4. Hadronic and misidentified photons are called nonprompt photons and they are one of the major backgrounds. Data driven estimation is done using the side band region of the data distribution of the $\sigma_{i\eta i\eta}$ variable. It is explained in detail in the later section 6.6. Electrons misidentified as photons, and scale factors for $W+\gamma$ and $Z+\gamma$ are estimated together using the invariant mass distribution of lepton (electron or muon) plus photon, explained in detail in Section 6.5. The estimation of background components and how they are used in the signal region is sketched in Figure 42.

6.1 Signal and Control Regions

Once the PF algorithm identifies the physics objects (e, μ, γ, jets), we impose additional criteria to make sure that the objects are reconstructed correctly. If strict identification criteria are applied then the selection efficiency of an object decreases accordingly. Thus, tighter selection criteria are used if we have sufficient statistics or when we need high

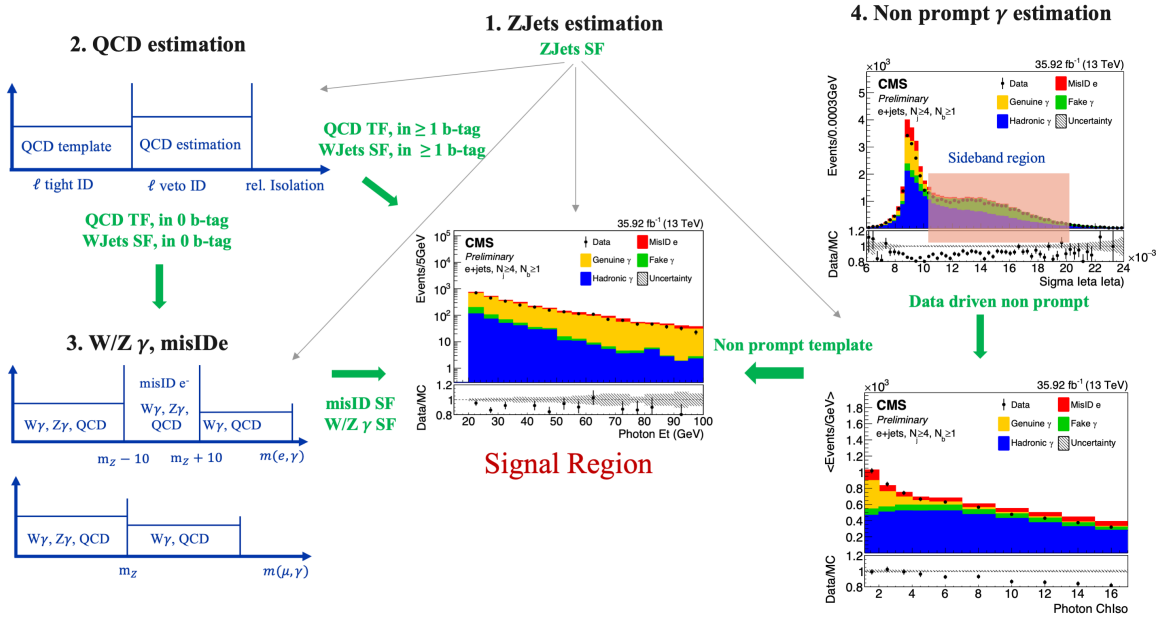


Figure 42: A sketch for estimation of major backgrounds and application in the signal region.

purity objects. As the $\mathrm{t\bar{t}}$ production cross section is high, we require tight lepton and jets, and b-tagged jet identification for event selection. But the production of a photon in association with $\mathrm{t\bar{t}}$ is low, so we require medium ID photon selection. In the signal region, events are selected with one high $\mathrm{P_T}$, well isolated lepton with at least 4 jets out of which at least one is b-tagged, and a high $\mathrm{P_T}$ photon.

6.1.1 Reconstruction Level Event Selection

Events that pass the trigger selection are further required to pass the MET filters as recommended by the JETMET POG [88]. A single lepton (electron or muon) satisfying the identification criteria in Secs. 5.2.3 5.2.2 is required. If there is an additional lepton in an event other than tight lepton, the event is discarded. The requirements of tight and veto lepton are shown in Table 15. A single photon satisfying the criteria of Sec. 5.2.4 is required. A requirement on the number of jets of $N_{\text{jets}} \geq 4$ among which at least one

must be b-tagged, $N_{\text{b-jets}} \geq 1$, defines the baseline selection as summarized in Table 20. The other control regions are defined with respect to the baseline selection and can have looser requirements on jet and b-tag multiplicities as shown in Fig. 43. Other requirements and the roles of these control regions are defined in the following sections.

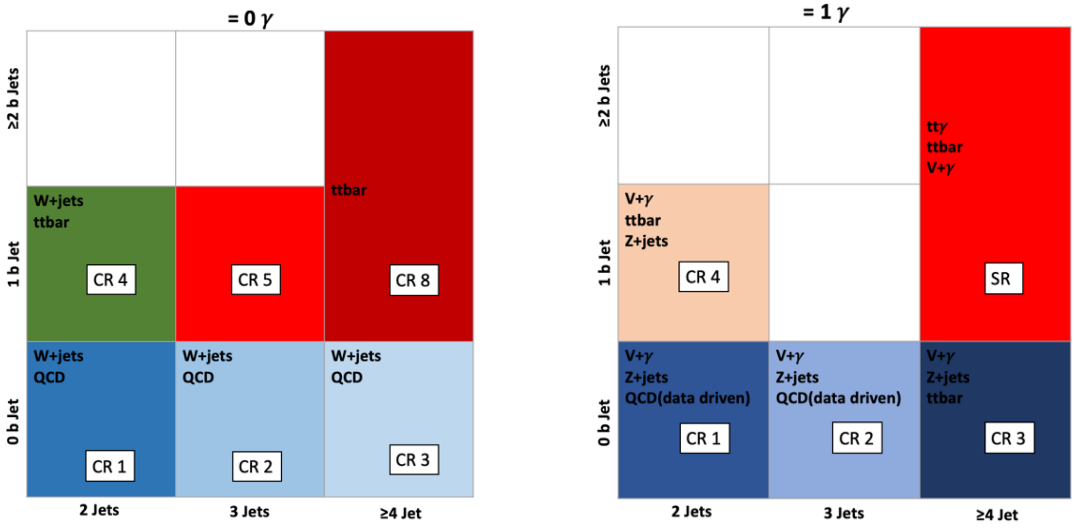


Figure 43: Definition of control and signal regions for 0γ (left) and 1γ (right) with different jets and b jets requirements.

Table 20: Summary of the baseline selection.

Trigger selection	Table 12
PV	leading PV satisfies requirements in Sec. 5.2.1
MET filters	Flag_HBHENoiseFilter Flag_HBHENoiseIsoFilter Flag_goodVertices Flag_globalTightHalo2016Filter Flag_EcalDeadCellTriggerPrimitiveFilter Flag_BadPFMuonFilter ecalBadCalibReducedMINIAODFilter (2017 and 2018)
N_{ℓ}	$= 1$
$N_{\ell,veto}$	$= 1$ (the selected lepton also passes the veto electron selection)
N_{jet}	≥ 4
N_{b-jet}	≥ 1
N_{γ}	$= 1$
$N_{\gamma,veto}$	$= 1$ (the selected photon also passes the veto photon selection)

6.2 Fit Strategy

The $t\bar{t}\gamma$ cross section measurement requires estimation of $t\bar{t}\gamma$ events with a well isolated photon mainly from $t\bar{t}$ events with nonprompt photon, $t\bar{t}$ events with isolated photon, $W + \gamma$and $Z + \gamma$ with isolated photon, and Single top, $W +$ jets , and $Z +$ jets [13] with isolated photon. The $t\bar{t}\gamma$ cross section extraction is outlined in Figure **??**. The signal region is dominated by the $t\bar{t}\gamma$ but there are significant backgrounds from $t\bar{t}$ (with both isolated and nonprompt photons), $W + \gamma$, and a collection of other processes. The estimation of the nonprompt photon contribution comes from the ChIso distribution. The M_3 distribution in the zero photon control region is used to confirm the $t\bar{t}$ contribution, and the lepton+photon invariant mass of the zero b-tag control region is used to estimate a number of backgrounds including $W + \gamma$ and $Z + \gamma$. Any sample that has misidentified electrons is scaled with the misID SF [14]. The QCD data driven template were estimated and were used in the signal and control regions. The ChIso distribution is split into isolated and nonprompt templates to have a better handle on nonprompt $t\bar{t}$ in the signal region. The zero photon M_3 distribution is introduced to estimate the $t\bar{t}$ normalization in the ChIso control region and signal region. Furthermore, the zero b-tagged control region is used to have a better handle on $W + \gamma$ and $Z + \gamma$ normalization in all the control and signal regions. Finally, a simultaneous binned maximum likelihood fit of the signal region and three control regions is done to extract the parameter of interest. A detailed discussion of the fit method can be found in Section 6.7.

[13] They are combined into a single template called other_1γ isolated

[14] misID SF = misidentified electron SF

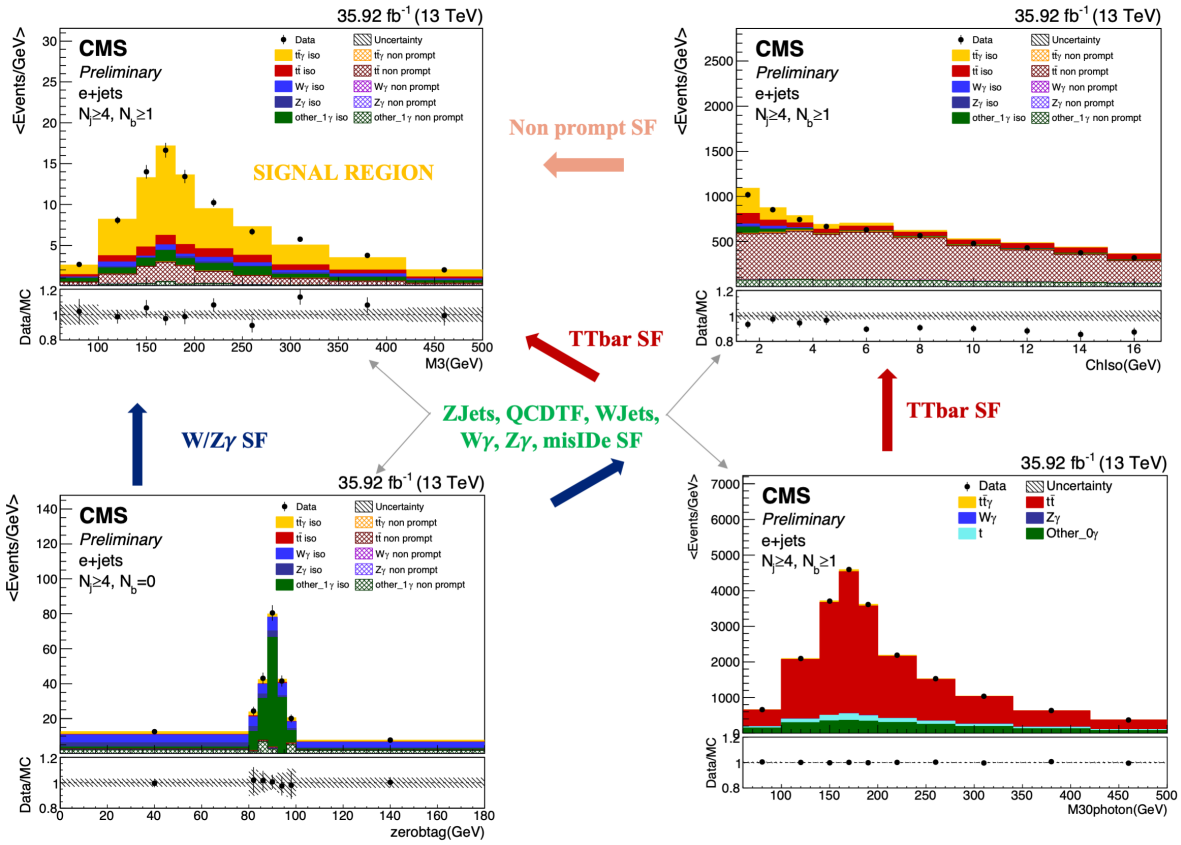


Figure 44: Analysis strategy and fit method sketch for extraction of signal.

6.3 Photon Categories and Background Overview

6.3.1 Photon Categories

The $\mathrm{t\bar{t}\gamma}$ cross section measurement requires differentiation of $\mathrm{t\bar{t}\gamma}$ events with a well isolated photon from $\mathrm{t\bar{t}}$ events with a nonprompt photon. The events in simulation are categorized based on the origin of the reconstructed photon. The categorization is performed by matching the reconstructed photon with generator level particles. The reconstructed photon is matched to a generator particle with stable status, if the $\mathrm{P_T}$ is within 50% of the reconstructed photon and if the angular separation, ΔR, is within value of 0.3. The minimum value of ΔR is taken to match the reconstructed photon with the generator par-

ticle. If the generator matching fails to match the reconstructed photon, additional checks are performed to categorize the photon. In case of the decay of a neutral pion into two photons with similar P_T , the matching algorithm fails if both generated photons are reconstructed as one with higher P_T then the sum of both. Then, neither of the P_T values of the generated photon is within 50% of the reconstructed photon. The reconstructed photon is categorized as a hadronic photon, if a generator pion and a generator photon without a requirement on stable status are found within a cone of $\Delta R < 0.3$. If no generator particle without status requirement is found within a cone of $\Delta R < 0.3$, the reconstructed photon is categorized as a pileup photon. Based on the nature of the matched generator particle, four categories are established:

1. Genuine photon:
    - the reconstructed photon is matched to a generator photon, and
    - the generator photon originates from a lepton, a boson or a quark
2. Misidentified electron:
    - the reconstructed photon is matched to a generator electron
3. Hadronic photon:
    - the reconstructed photon is matched to a generator photon, and the generator photon originates from a hadronic source, or
    - the reconstructed photon is not matched to a generator particle, and a generator neutral pion and a generator photon without status requirement are found within $\Delta R < 0.3$

4. Misidentified or pile-up photon:

    - a photon from PU event when there is no generator particle without status requirement is found within $\Delta R < 0.3$

The term ”non-prompt photons” is short for hadronic photons, and misidentified or pile-up photons, whereas the term ”isolated photons” stands for genuine and misidentified electrons. The terms are introduced, because respective contributions are estimated together for each. Figure 45 shows the distribution of simulated events by photon category such as genuine, misidentified electron, hadronic, and hadronic pile-up photon for both $e+$ jets channel and $\mu+$ jets channel for Monte-Carlo simulation for 2016, 2017 and 2018.

Tables 21 to 26, shows the event yields for each samples classified by photon categories in the signal region ($N_j \geq 4, N_b \geq 1$) with the percentage values for both $e+$ jets (left) and $\mu+$ jets (right) final states for 2016 (top), 2017 (middle) and 2017 (bottom) dataset.

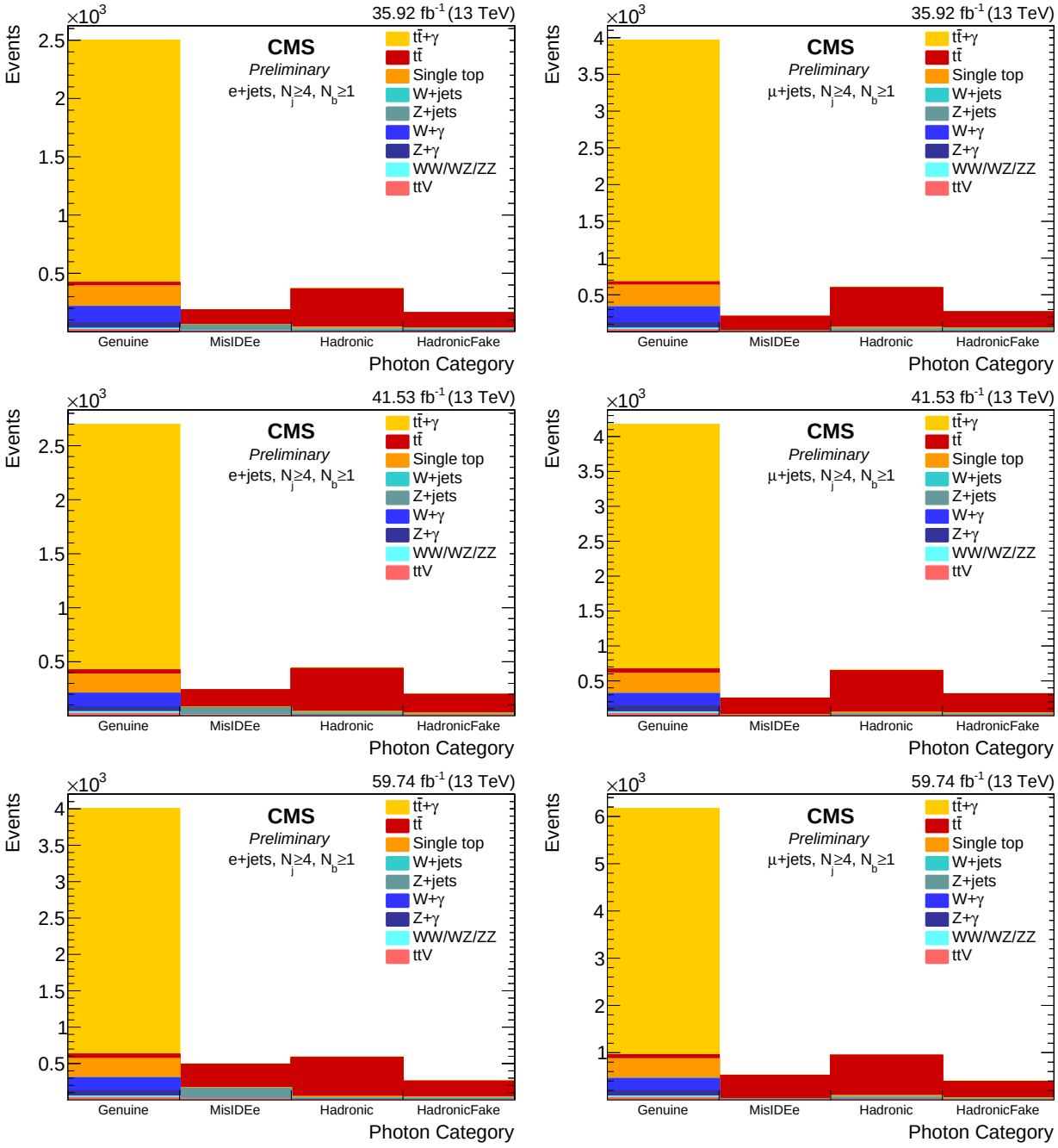


Figure 45: The distribution of genuine, misidentified electron, hadronic, and pile-up photons in the signal region ($N_j \geq 4, N_b \geq 1$) classified by samples in $e+$jets (left) and $\mu+$jets (right) for 2016 (top), 2017 (middle), and 2017 (bottom).

Table 21: Number of events in the $e+$jets final state in each of the 2016 MC simulated processes in the signal region($N_j \geq 4, N_b \geq 1$), classified into the four event categories based on the origin of the photon. The last row and column show the percent of events in simulation coming from each event category and process, respectively. Uncertainties are derived from the MC statistics.

Sample	GenuinePhoton	MisIDEle	Hadronic	Pile-up	Total	Percent
TTGamma	1982.5 ± 7.7	5.8 ± 0.2	5.8 ± 0.4	4.2 ± 0.3	1998.3 ± 7.7	60.55 ± 1.08
TTbar	25.7 ± 1.1	263.8 ± 2.4	311.8 ± 5.6	128.4 ± 3.7	729.7 ± 7.2	22.11 ± 0.44
SingleTop	154.8 ± 8.7	9.1 ± 1.0	13.9 ± 1.8	7.7 ± 1.3	185.5 ± 9.0	5.62 ± 0.29
WJets	0.0 ± 0.0	0.0 ± 0.0	14.5 ± 3.4	10.8 ± 4.4	25.3 ± 5.6	0.77 ± 0.17
ZJets	1.7 ± 1.7	109.0 ± 8.4	4.4 ± 2.6	5.5 ± 2.9	120.6 ± 9.4	3.65 ± 0.29
WGamma	138.6 ± 20.9	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	138.6 ± 20.9	4.20 ± 0.64
ZGamma	39.6 ± 4.1	1.4 ± 0.5	0.0 ± 0.0	1.2 ± 0.6	42.2 ± 4.1	1.28 ± 0.13
Diboson	9.9 ± 1.1	5.3 ± 0.6	0.7 ± 0.3	0.9 ± 0.4	16.8 ± 1.4	0.51 ± 0.04
TTV	13.2 ± 0.8	5.3 ± 0.2	1.9 ± 0.3	0.5 ± 5.3	20.9 ± 5.4	0.63 ± 0.16
QCD_DD	0.0 ± 0.0	0.0 ± 0.0	22.2 ± 5.3	0.0 ± 0.0	22.2 ± 5.3	0.67 ± 0.16
MC	2366.1 ± 24.4	399.6 ± 8.8	375.2 ± 7.3	159.2 ± 6.6	3300.1 ± 28.7	
Percent	71.70 ± 0.97	12.11 ± 0.29	11.37 ± 0.24	4.82 ± 0.20	100.00	100.00
Data					2910.2	

Table 22: Number of events in the $\mu+$jets final state in each of the 2016 MC simulated processes in the signal region($N_j \geq 4, N_b \geq 1$), classified into the four event categories based on the origin of the photon. The last row and column show the percent of events in simulation coming from each event category and process, respectively. Uncertainties are derived from the MC statistics.

Sample	GenuinePhoton	MisIDEle	Hadronic	Pile-up	Total	Percent
TTGamma	3147.1 ± 9.8	6.3 ± 0.2	9.2 ± 0.5	5.7 ± 0.4	3168.2 ± 9.8	62.29 ± 0.89
TTbar	41.4 ± 1.4	407.9 ± 3.0	510.3 ± 7.3	214.1 ± 4.9	1173.6 ± 9.4	23.07 ± 0.37
SingleTop	264.3 ± 11.1	12.5 ± 1.2	20.0 ± 2.2	13.3 ± 1.7	310.2 ± 11.5	6.10 ± 0.24
WJets	0.0 ± 0.0	0.0 ± 0.0	17.5 ± 3.8	24.9 ± 5.0	42.4 ± 6.3	0.83 ± 0.12
ZJets	7.1 ± 3.4	4.0 ± 1.8	9.2 ± 3.8	6.1 ± 3.2	26.3 ± 6.3	0.52 ± 0.12
WGamma	218.0 ± 26.3	0.0 ± 0.0	3.6 ± 3.6	0.0 ± 0.0	221.6 ± 26.5	4.36 ± 0.53
ZGamma	56.2 ± 5.1	0.0 ± 0.0	0.3 ± 0.3	0.4 ± 0.3	56.9 ± 5.1	1.12 ± 0.10
Diboson	16.7 ± 1.7	0.3 ± 0.2	1.7 ± 0.4	0.4 ± 0.2	19.1 ± 1.8	0.37 ± 0.04
TTV	20.7 ± 1.0	5.0 ± 0.3	3.6 ± 0.4	1.0 ± 8.8	30.3 ± 8.9	0.60 ± 0.18
QCD_DD	0.0 ± 0.0	0.0 ± 0.0	37.8 ± 8.8	0.0 ± 0.0	37.8 ± 8.8	0.74 ± 0.17
MC	3771.5 ± 30.9	436.0 ± 3.7	613.1 ± 10.0	265.9 ± 7.9	5086.4 ± 35.9	
Percent	74.15 ± 0.80	8.57 ± 0.09	12.05 ± 0.21	5.23 ± 0.16	100.00	100.00
Data					4506.2	

6.3.2 Overview of Background Components

Figure 46 shows the E_{T} distribution of photons classified by the simulated samples (left) and by the photon categories(right). The genuine photon contribution in Figure 46 (right) is larger than the photon contribution by signal sample $\mathrm{t\bar{t}}\gamma$ in 46 (left). There are

Table 23: Number of events in the $e+$jets final state in each of the 2017 MC simulated processes in the signal region($N_j \geq 4, N_b \geq 1$), classified into the four event categories based on the origin of the photon. The last row and column show the percent of events in simulation coming from each event category and process, respectively. Uncertainties are derived from the MC statistics.

Sample	GenuinePhoton	MisIDEle	Hadronic	Pile-up	Total	Percent
TTGamma	2175.7 ± 8.1	6.8 ± 0.2	6.9 ± 0.5	4.7 ± 0.4	2194.1 ± 8.1	58.26 ± 0.97
TTbar	34.5 ± 1.3	342.2 ± 2.8	377.9 ± 6.4	167.4 ± 4.5	921.9 ± 8.4	24.48 ± 0.46
SingleTop	154.2 ± 8.6	10.6 ± 0.9	13.7 ± 1.5	11.1 ± 1.2	189.6 ± 8.9	5.03 ± 0.25
WJets	0.0 ± 0.0	0.0 ± 0.0	13.4 ± 6.8	8.1 ± 4.4	21.5 ± 8.1	0.57 ± 0.22
ZJets	0.0 ± 0.0	162.5 ± 14.2	4.6 ± 3.4	1.7 ± 1.7	168.9 ± 14.7	4.48 ± 0.40
WGamma	145.9 ± 25.6	0.0 ± 0.0	3.6 ± 3.6	0.0 ± 0.0	149.5 ± 25.8	3.97 ± 0.69
ZGamma	31.9 ± 4.1	1.0 ± 0.3	0.0 ± 0.0	0.7 ± 0.5	33.6 ± 4.1	0.89 ± 0.11
Diboson	14.1 ± 1.4	4.9 ± 0.3	1.0 ± 0.4	0.6 ± 0.2	20.7 ± 1.5	0.55 ± 0.04
TTV	17.6 ± 0.6	7.5 ± 0.2	2.8 ± 0.2	0.8 ± 5.1	28.8 ± 5.2	0.76 ± 0.14
QCD_DD	0.0 ± 0.0	0.0 ± 0.0	37.3 ± 5.1	0.0 ± 0.0	37.3 ± 5.1	0.99 ± 0.14
MC	2573.8 ± 28.6	535.6 ± 14.5	461.2 ± 10.7	195.2 ± 6.6	3765.8 ± 35.2	
Percent	68.35 ± 0.99	14.22 ± 0.41	12.25 ± 0.31	5.18 ± 0.18	100.00	100.00
Data						

Table 24: Number of events in the $\mu+$jets final state in each of the 2017 MC simulated processes in the signal region($N_j \geq 4, N_b \geq 1$), classified into the four event categories based on the origin of the photon. The last row and column show the percent of events in simulation coming from each event category and process, respectively. Uncertainties are derived from the MC statistics.

Sample	GenuinePhoton	MisIDEle	Hadronic	Pile-up	Total	Percent
TTGamma	3364.2 ± 10.5	8.1 ± 0.3	9.0 ± 0.5	6.5 ± 0.5	3387.8 ± 10.5	61.69 ± 0.85
TTbar	54.0 ± 1.7	510.4 ± 3.6	571.3 ± 8.2	260.3 ± 5.8	1396.0 ± 10.8	25.42 ± 0.40
SingleTop	253.2 ± 11.1	15.7 ± 1.1	20.4 ± 1.9	17.9 ± 1.7	307.1 ± 11.5	5.59 ± 0.22
WJets	0.0 ± 0.0	0.0 ± 0.0	6.9 ± 4.0	6.7 ± 3.6	13.5 ± 5.4	0.25 ± 0.10
ZJets	9.2 ± 5.3	0.0 ± 0.0	12.5 ± 6.2	8.7 ± 4.5	30.4 ± 9.3	0.55 ± 0.17
WGamma	168.3 ± 28.5	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	168.3 ± 28.5	3.06 ± 0.52
ZGamma	68.2 ± 5.6	0.0 ± 0.0	0.3 ± 0.3	0.0 ± 0.0	68.5 ± 5.6	1.25 ± 0.10
Diboson	19.1 ± 1.7	0.5 ± 0.1	1.4 ± 0.5	0.9 ± 0.3	21.8 ± 1.8	0.40 ± 0.03
TTV	26.8 ± 0.8	5.6 ± 0.2	4.4 ± 0.3	1.0 ± 12.0	37.8 ± 12.1	0.69 ± 0.22
QCD_DD	0.0 ± 0.0	0.0 ± 0.0	60.6 ± 12.0	0.0 ± 0.0	60.6 ± 12.0	1.10 ± 0.22
MC	3962.9 ± 33.3	540.4 ± 3.8	686.8 ± 11.2	302.0 ± 8.4	5492.1 ± 40.1	
Percent	72.16 ± 0.80	9.84 ± 0.10	12.51 ± 0.22	5.50 ± 0.16	100.00	100.00
Data						

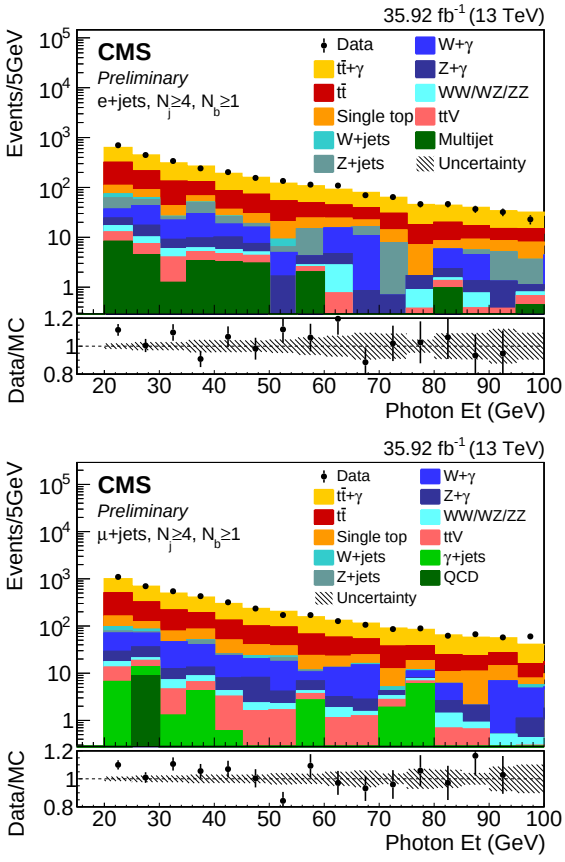
Table 25: Number of events in the $e+$jets final state in each of the 2018 MC simulated processes in the signal region($N_j \geq 4, N_b \geq 1$), classified into the four event categories based on the origin of the photon. The last row and column show the percent of events in simulation coming from each event category and process, respectively. Uncertainties are derived from the MC statistics.

Sample	GenuinePhoton	MisIDEle	Hadronic	Pile-up	Total	Percent
TTGamma	3412.3 ± 12.8	10.3 ± 0.5	9.7 ± 0.7	6.8 ± 0.5	3439.1 ± 12.8	58.73 ± 0.80
TTbar	58.9 ± 2.3	466.2 ± 5.0	541.4 ± 10.0	223.8 ± 6.5	1290.3 ± 13.2	22.04 ± 0.37
SingleTop	253.4 ± 13.1	13.5 ± 1.4	18.7 ± 2.1	12.2 ± 1.7	297.8 ± 13.5	5.09 ± 0.24
WJets	0.0 ± 0.0	0.0 ± 0.0	22.4 ± 14.8	33.6 ± 12.3	56.0 ± 19.3	0.96 ± 0.33
ZJets	4.1 ± 4.1	297.1 ± 31.0	6.2 ± 6.2	14.0 ± 8.1	321.4 ± 32.9	5.49 ± 0.57
WGamma	219.9 ± 45.8	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	219.9 ± 45.8	3.76 ± 0.78
ZGamma	73.3 ± 8.7	0.9 ± 0.6	0.0 ± 0.0	1.1 ± 1.1	75.3 ± 8.8	1.29 ± 0.15
Diboson	21.0 ± 1.6	7.8 ± 0.5	1.1 ± 0.3	1.1 ± 0.4	31.1 ± 1.8	0.53 ± 0.03
TTV	28.0 ± 1.5	11.8 ± 0.5	3.2 ± 0.5	0.8 ± 10.6	43.8 ± 10.8	0.75 ± 0.18
QCD_DD	0.0 ± 0.0	0.0 ± 0.0	80.7 ± 10.6	0.0 ± 0.0	80.7 ± 10.6	1.38 ± 0.18
MC	4070.9 ± 50.4	807.6 ± 31.5	683.4 ± 19.1	293.5 ± 16.3	5855.4 ± 66.2	
Percent	69.52 ± 1.17	13.79 ± 0.56	11.67 ± 0.35	5.01 ± 0.28	100.00	100.00
Data						

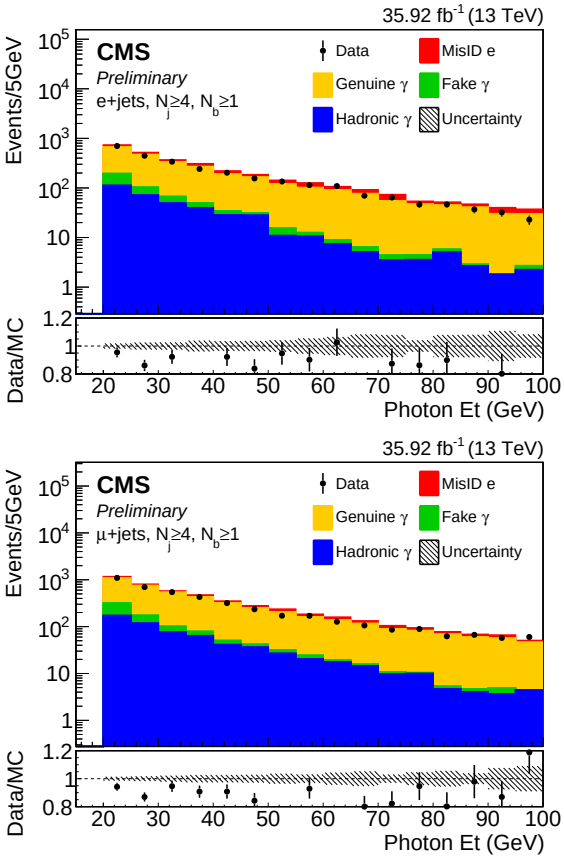
Table 26: Number of events in the $\mu+$jets final state in each of the 2018 MC simulated processes in the signal region($N_j \geq 4, N_b \geq 1$), classified into the four event categories based on the origin of the photon. The last row and column show the percent of events in simulation coming from each event category and process, respectively. Uncertainties are derived from the MC statistics.

Sample	GenuinePhoton	MisIDEle	Hadronic	Pile-up	Total	Percent
TTGamma	5168.5 ± 16.0	11.2 ± 0.5	14.9 ± 0.9	7.9 ± 0.6	5202.5 ± 16.1	61.40 ± 0.69
TTbar	80.3 ± 2.6	725.0 ± 6.5	839.7 ± 12.9	351.7 ± 8.3	1996.6 ± 16.9	23.57 ± 0.32
SingleTop	376.8 ± 16.8	16.2 ± 1.6	30.3 ± 2.9	20.3 ± 2.1	443.7 ± 17.3	5.24 ± 0.21
WJets	0.0 ± 0.0	0.0 ± 0.0	20.9 ± 10.5	13.0 ± 7.5	33.9 ± 12.9	0.40 ± 0.15
ZJets	25.6 ± 11.0	6.2 ± 4.1	37.1 ± 28.6	0.0 ± 0.0	68.9 ± 30.9	0.81 ± 0.36
WGamma	376.7 ± 61.5	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	376.7 ± 61.5	4.45 ± 0.73
ZGamma	113.0 ± 11.1	0.0 ± 0.0	0.0 ± 0.0	-0.1 ± 0.1	112.9 ± 11.1	1.33 ± 0.13
Diboson	30.2 ± 1.9	0.6 ± 0.3	2.3 ± 0.6	1.1 ± 0.3	34.1 ± 2.1	0.40 ± 0.02
TTV	43.1 ± 1.9	6.9 ± 0.5	6.8 ± 0.7	1.3 ± 19.9	58.1 ± 20.0	0.69 ± 0.24
QCD_DD	0.0 ± 0.0	0.0 ± 0.0	145.1 ± 19.9	0.0 ± 0.0	145.1 ± 19.9	1.71 ± 0.24
MC	6214.2 ± 67.7	766.0 ± 7.9	1097.2 ± 33.2	395.3 ± 11.4	8472.7 ± 81.7	
Percent	73.34 ± 1.07	9.04 ± 0.13	12.95 ± 0.41	4.67 ± 0.14	100.00	100.00
Data						

genuine photons contribution from other processes such as $\mathrm{t\bar{t}}$, $\mathrm{tW}\gamma$ and $\mathrm{t}\gamma$ in photon category plots. All background processes are scaled by theirs respective scale factors obtained from the maximum likelihood fittings. The overview of the background contributions in the signal region $N_j \geq 4, N_b \geq 1$ for both $e+$jets and $\mu+$jets channels in 2016, 2017, and 2018 data are tabulated in Table from 21 to 26. This analysis does binned template fits using the Higgs Combine Tool [106] to get normalization factors for the background processes as well as the signal strength.



(a) 2016, $e+$jets , classified by MC samples



(b) 2016, $\mu+$jets , classified by photon categories

Figure 46: Photon E_{T} distribution for $e+$jets (left) and $\mu+$jets (right) channels in the signal region $N_j \geq 4, N_b \geq 1$ for 2016 dataset classified in simulated samples(left) and in photon category (right).

6.4 Data Driven Estimation of QCD background

QCD is a minor background in our final signal region but it has significant effects while we extract Z+jets, Z/W+γ, and electrons misidentified as photon, because for those estimates are made in separate control regions with a lower number of jets. For the QCD multijet background we use a data driven estimation. First we select a QCD enriched region by relaxing the b-tag requirements. In the electron channel we invert the electron relative isolation cut in the 0 b-tag control region. In the muon channel we look at a control region with 0 b-jets and the muon relative isolation reversed to RelIso $\geq$ 0.15. We subtract all non QCD Monte Carlo from data as seen in Figure 47. But to use this QCD data driven template in our signal region we need to scale it properly. A fitting method is used to get the scale factor, called the QCD transfer factor, where data and all other MC histograms are from the tight lepton isolation region except QCD as seen in Figure 47. We do fitting in two selections : (1) $N_{jet} = 2, N_{\gamma} = 0, N_{b-jet} = 0$ and (2) $N_{jet} = 2, N_{\gamma} = 0, N_{b-jet} \geq 1$, where QCD statics are higher and expected to give a better fit result. The QCD transfer factor applied to higher N_{jet} and N_{γ} selections with an N_{jet} dependence correction factor. This correction factor is taken from simulation and accounts for a linear N_{Jet} dependence fo the transfer factor for $N_{b-jet} \geq 1$, although for $N_{b-jet} = 0$ this correction factor is 1.

Table 27: Binning choices for estimation of the QCD template

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

$$(S_{data})_{bin} = (\mu_1)_{bin} \times (S_{QCD})_{bin} + \mu_2 \times (S_{W+Jets})_{bin} + (\sum_{all\ bkg} S_{bkg})_{bin} \tag{6.1}$$

where, bin $\in$ [1, 24],S_{data}) is the data distribution from sideband region. For fitting we use the M_T^W distribution, split in 24 bins on the basis of different p_T and η of leptons as listed in Table 27. All bins are fitted simultaneously. QCD and W+Jets normalization are allowed to float. As a result we get 24 different QCD transfer factors along with one global W+Jets scale factor in each $e+$ jets and $\mu+$ jets channels.

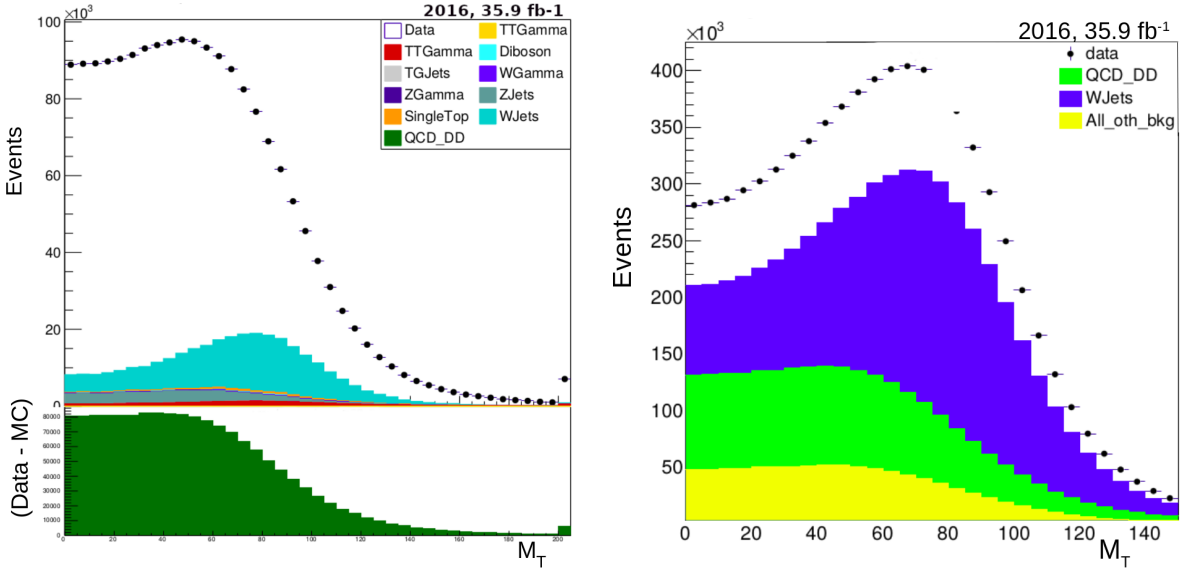


Figure 47: left: QCD (Data -all MC) in QCD CR for $e+$ jets channel. right: All histograms to be fitted in the signal region for $e+$ jets channel

6.5 Estimation of Backgrounds from Misidentified Electron, $W+\gamma$, and $Z+\gamma$ Processes

6.5.1 Estimation of $Z+$ jets Scale Factors

Dilepton events that satisfy the $t\bar{t}$ selection criteria are selected with two same flavor leptons and opposite charges that pass all the lepton selection requirements as explained in the Section 5.2. The correct estimation of Z+jets is important for the misidentified electron scale factors. The templates with the dilepton invariant mass is created for the binned shape analysis in both $ee+$ jets and $\mu\mu+$ jets +jets channels for 2016, 2017, and

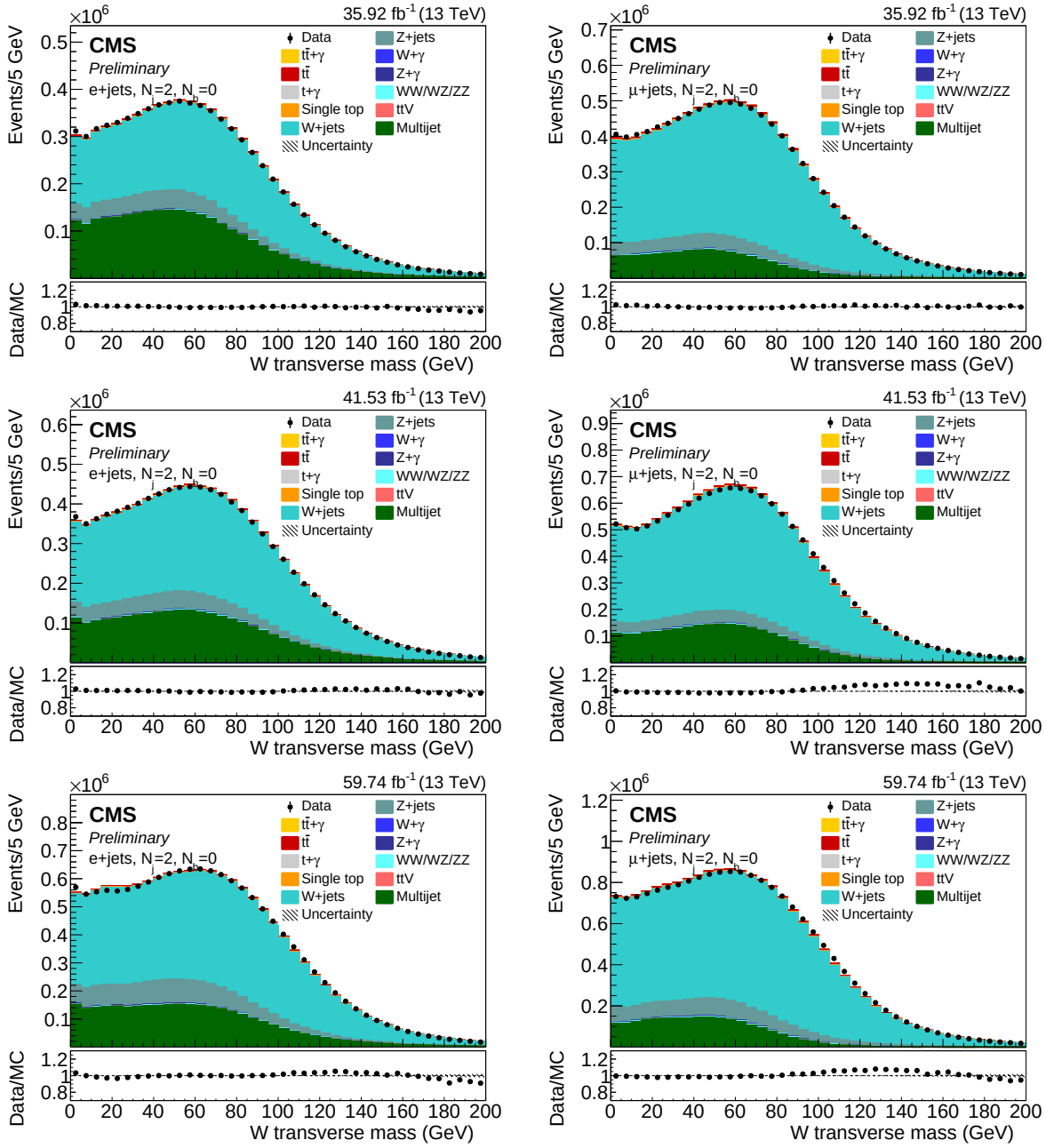


Figure 48: post fit plot in $N_j = 2, N_b = 0$ with with 0 photon . left $e+$jets channel for 2016 (top), 2017 (middle), and 2018 (bottom). right $\mu+$jets channel for 2016 (top), 2017 (middle), and 2018 (bottom)

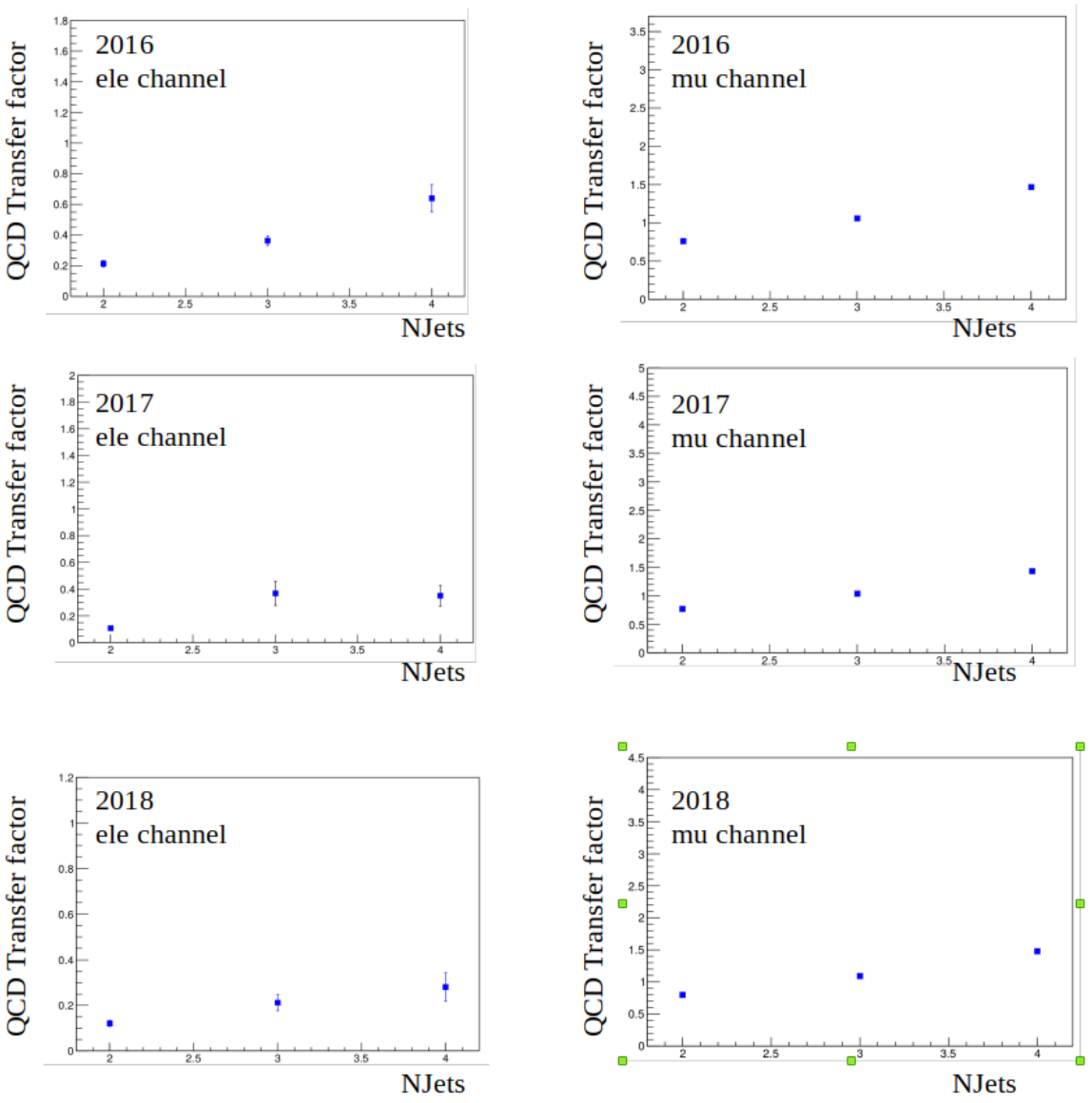


Figure 49: QCD transfer factors as a function of jet multiplicity which is estimated from MC simulation

2018 simulation era. The templates have the $Z+$jets sample as the signal process and all the other samples merged as the background processes as shown in the left plots in Figure 50. The dilepton invariant mass prefit plots are shown in Figure 51 on the left for both $ee+$jets (top) and $\mu\mu+$jets (bottom). A simultaneous maximum likelihood fit to data is performed on the binned shaped distributions in $ee+$jets and $\mu\mu+$jets final state to extract the $Z+$jets scale factors. Backgrounds are fixed while the signal is allowed to float freely during the fit. The average $Z+$jets scale factor is obtained from a control region with $N_j \geq 2, N_b = 0$ and the other control regions are used to validate the obtained scale factor as listed in Table 28.

We find a scale factor of $1.09 \pm 0.027(\text{syst}) \pm 0.002(\text{stat})$, $1.13 \pm 0.027(\text{syst}) \pm 0.002(\text{stat})$, and $1.37 \pm 0.027(\text{syst}) \pm 0.002(\text{stat})$ in simultaneous fitting of $ee+$jets and $\mu\mu+$jets for 2016, 2017, and 2018 data respectively, as shown in Table 28. The systematic and statistical uncertainties of the post fit is less than 2.7% and 0.2%, respectively. Figure 52 is a graphical representation of Table 28. A comparison of both $e+$jets and $\mu+$jets channels and 2016, 2017, and 2018 dataset for control regions $N_\ell = 2, N_\gamma = 0$, $N_j \geq 2, N_b = 0$ and $N_\ell = 2, N_\gamma = 0$, $N_j \geq 4, N_b = 0$ is displayed in Appendix B. The mismodelling of the $Z+$jets process for higher jet multiplicity are studied and scaled by the average scale factor in later steps of the analysis.

We used the HiggsCombineTools to measure the best-fit signal strength and and its uncertainties, broken down into systematics and statistical components. The explicit scan of the profile maximum log-likelihood of the parameter of interest is used as it shows any unexpected features in the shape of the likelihood curve. Figure **??** shows the scan of

$Z+$jets scale factor for Asimov[15] dataset and observed dataset for a control region with $N_j \geq 2, N_b = 0$ for 2016 (top), 2017 (middle) and 2017 (bottom) era respectively. The classification of systematics and statistical uncertainties are also shown. The intersection points on the horizontal lines at 1 and 4 with the curve represent the $\pm 1\sigma$ and $\pm 2\sigma$ confidence interval around the best-fit value respectively.

To study the effects of systematics and statistical uncertainties on the signal strength measurement, we generate the "impact" plot. It shows the shift in the signal strength with respect to the best-fit value that is induced if a nuisance parameter is shifted by its $\pm 1\sigma$ post-fit uncertainty values. The correlation coefficients between the signal strength and the nuisance parameters are calculated in the plot. The impact plots shows the postfit pulls and constraints on nuisance parameters for $Z+$jets scale factor fit to Asimov dataset (left) and observed dataset (right) for 2016 (top), 2017 (middle) and 2017 (bottom) respectively in Figure 53.

Table 28: The extraction of $Z+$jets scale factors for several control regions with different number of jets and btagged jets multiplicity by simultaneous fitting of $ee+$jets and $\mu\mu+$jets for year 2016, 2017 and 2018 dataset. The average scale factors for different years of dataset are given in the first row in a control region with $N_j \geq 2, N_b = 0$.

Control Regions	$Z+$jets SF 2016	$Z+$jets SF 2017	$Z+$jets SF 2018
$N_j \geq 2, N_b = 0$	$1.09^{+0.03}_{-0.03}$	$1.16^{+0.04}_{-0.04}$	$1.14^{+0.03}_{-0.03}$
$N_j = 2, N_b = 0$	$1.10^{+0.03}_{-0.03}$	$1.17^{+0.04}_{-0.04}$	$1.14^{+0.03}_{-0.03}$
$N_j = 3, N_b = 0$	$1.12^{+0.03}_{-0.03}$	$1.15^{+0.04}_{-0.04}$	$1.12^{+0.03}_{-0.03}$
$N_j \geq 4, N_b = 0$	$1.00^{+0.03}_{-0.03}$	$1.11^{+0.04}_{-0.04}$	$1.09^{+0.03}_{-0.03}$
$N_j = 2, N_b = 1$	$1.14^{+0.05}_{-0.05}$	$1.23^{+0.05}_{-0.05}$	$1.22^{+0.06}_{-0.06}$
$N_j = 3, N_b = 1$	$1.11^{+0.05}_{-0.05}$	$1.33^{+0.07}_{-0.07}$	$1.32^{+0.07}_{-0.07}$
$N_j = 2, N_b = 2$	$1.19^{+0.07}_{-0.06}$	$1.16^{+0.08}_{-0.08}$	$1.19^{+0.09}_{-0.08}$
$N_j = 3, N_b \geq 2$	$1.22^{+0.09}_{-0.08}$	$1.24^{+0.10}_{-0.10}$	$1.24^{+0.10}_{-0.10}$

[15] Asimov data is defined as a number of toy data exactly equal to the number of expected events.

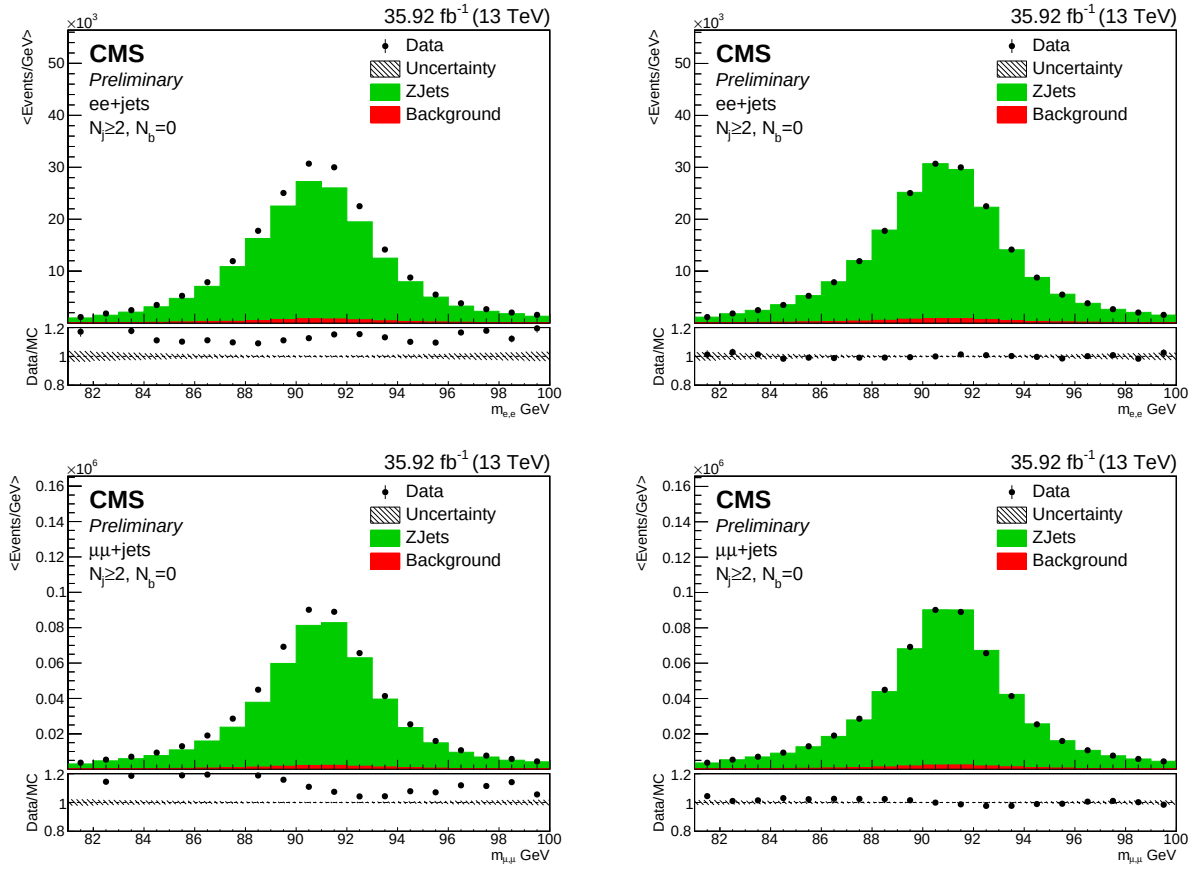


Figure 50: The prefit and postfit templates in a control region with $N_\ell = 2, N_\gamma = 0$, $N_j \geq 2, N_b = 0$ in $ee +$ jets channel(left) and $\mu\mu +$ jets channel(right) for 2016 dataset.

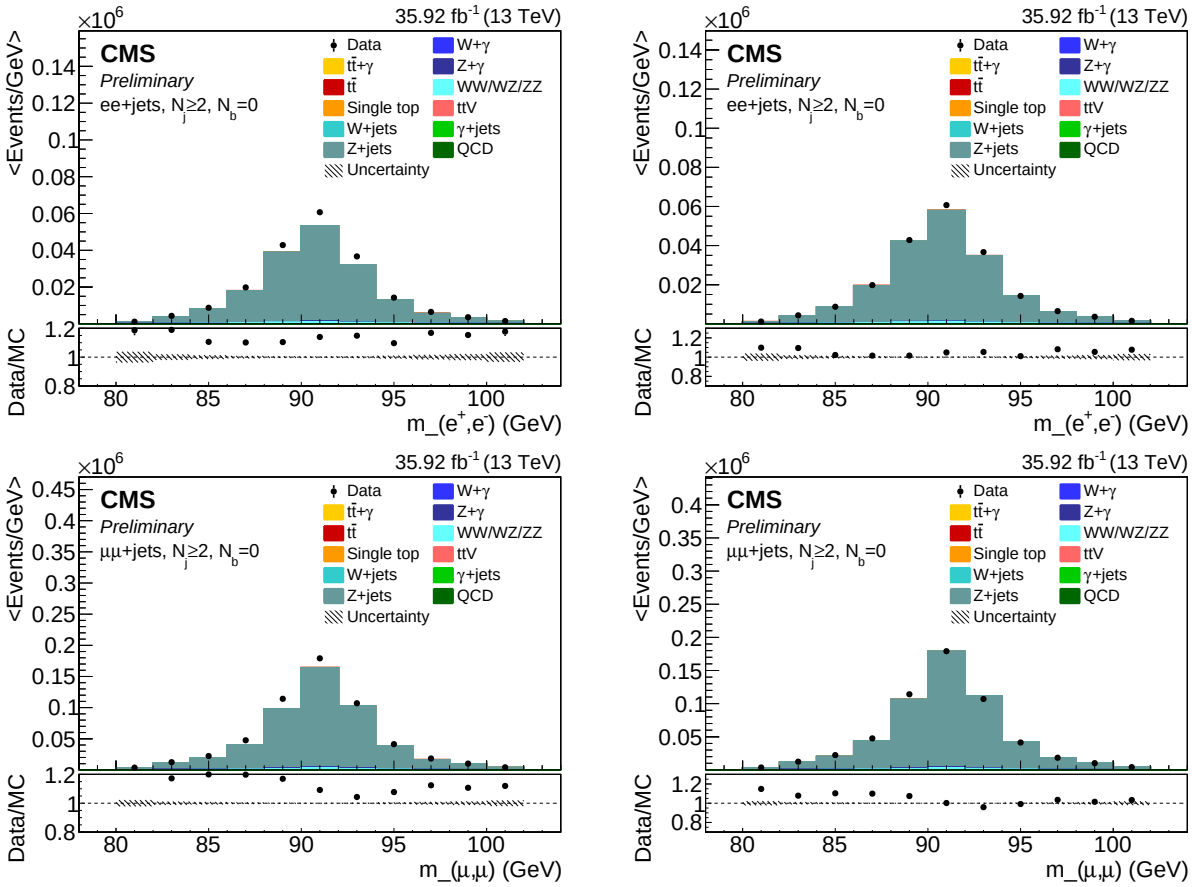


Figure 51: The prefit(left) and postfit(right) plots of an invariant mass distribution of dilepton samples in a control region with $N_\ell = 2, N_\gamma = 0$, $N_j \geq 2, N_b = 0$ in the diejets channel(top) and the $\mu\mu$ + jets channel(bottom) for 2016 dataset.

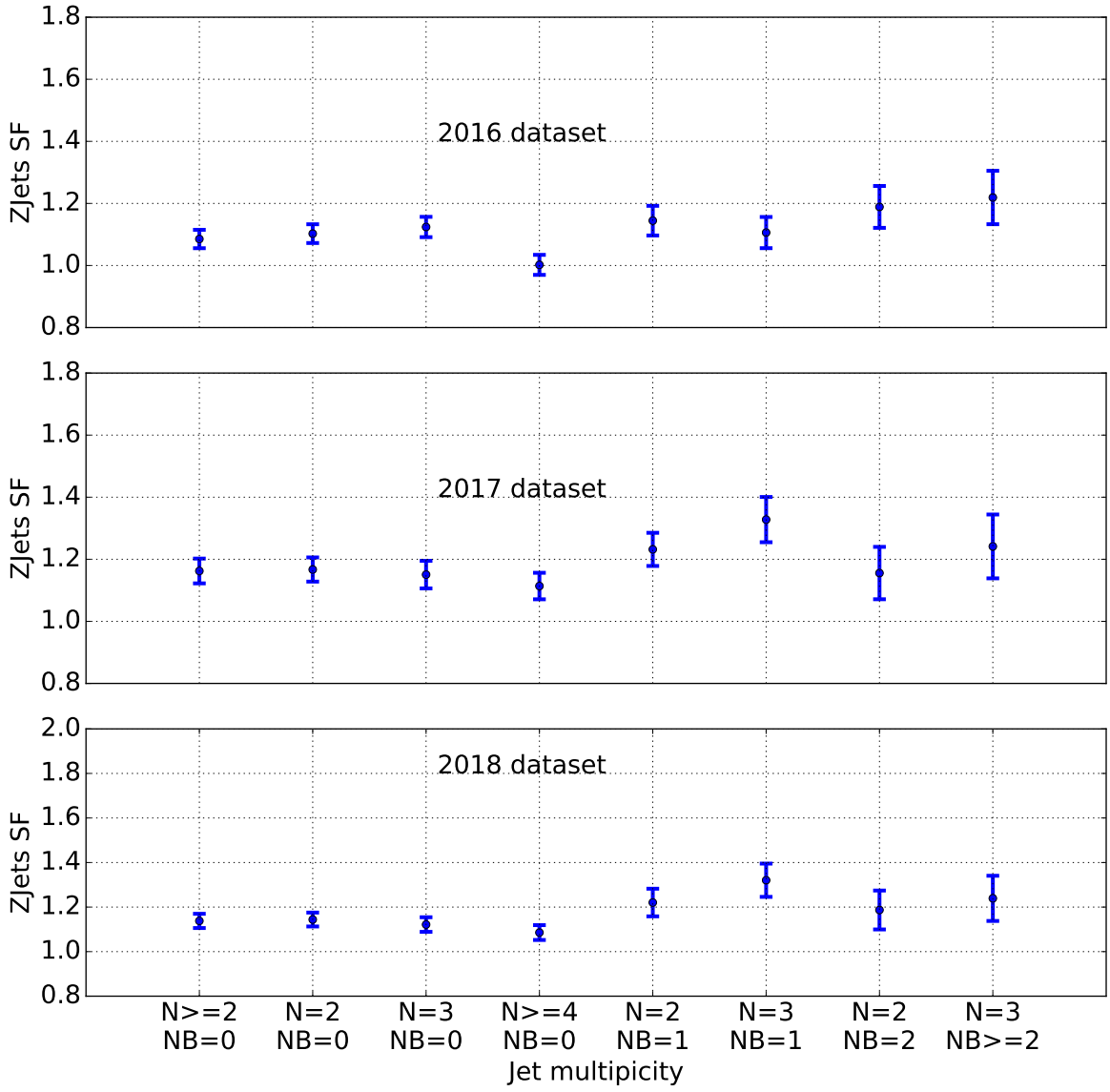


Figure 52: The comparison of the $Z+$jets scale factors from control regions with different number of jets and b-tagged jets for 2016, 2017, 2018. This is a graphical representation of Table 28.

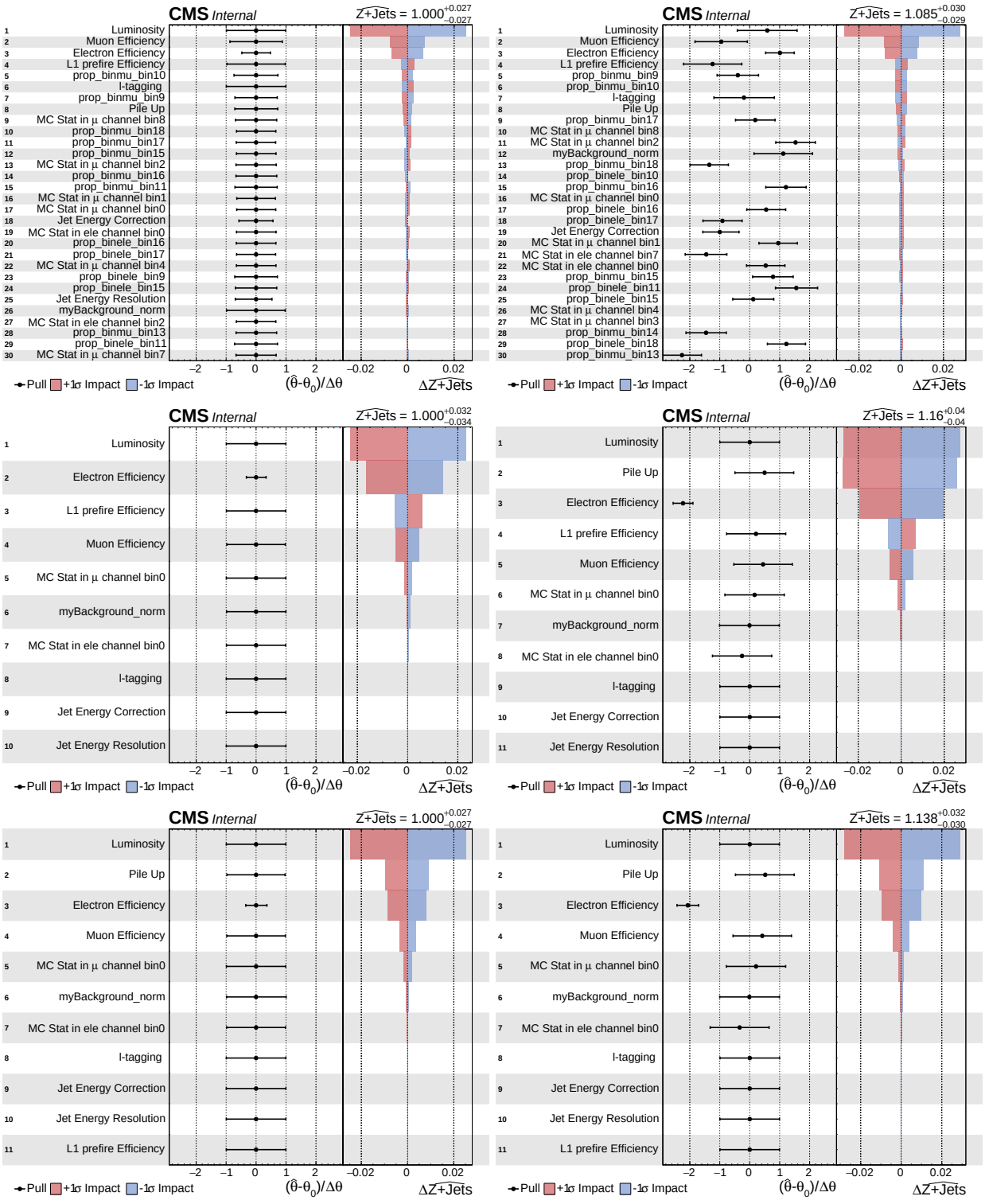


Figure 53: The impact plot shows postfit pulls and constraints on nuisance parameters for the Z + jets scale factor fit to Asimov data (left) and data (right) for 2016 (top), 2017 (middle), and 2017 (bottom), respectively.

6.5.2 Estimation of Misidentified Electron, $Z + \gamma$, and $W + \gamma$ Scale Factors

When an electron loses its track information, it can be identified as a photon. It may be well isolated and have the same shape as a genuine photon in the charged hadron isolation distribution. This background is measured in a control region with with 1 photon , $N_j \geq 2, N_b = 0$ with a $|m(e,\gamma) - m_Z| < 10$ where, one of the electrons is misidentified as a photon, as the invariant mass $m(e,\gamma)$ makes a Z boson mass peak in $e +$ jets channel. However, other processes will have a wider distribution and effect the misidentified electron scale factor (misIDE SF) estimation as shown in Figure 54(left). Such other processes are $W + \gamma$ and $Z + \gamma$ which contribute a well isolated electron and a prompt photon to the $|m(e,\gamma)|$ peak. Figure 63(left) shows the sizable contributions from those processes under the Z boson mass peak. Hence it is necessary to estimate them before we extract the misIDE SF. The $Z + \gamma$ contribution tends to populate the lower band of the Z mass peak which helps to de-correlate the $Z+\gamma$ scale factor from $W+\gamma$ scale factor. So, the invariant mass of the photon and the electron is required to lie in a mass range of 0-180 GeV to enable the background components to be extracted in the fit. In the $e +$ jets channel, there is a contribution from QCD in the Z mass region. The use of the $m(\mu,\gamma)$ control region from the $\mu +$ jets channel helps in to separate the QCD multijet prediction from $W + \gamma$ and $Z + \gamma$ scale factors as it does not contain both the misidentified electron and the QCD contribution. We separate the $W + \gamma$ from the $Z + \gamma$ region in the $\mu +$ jets channel at the Z mass.

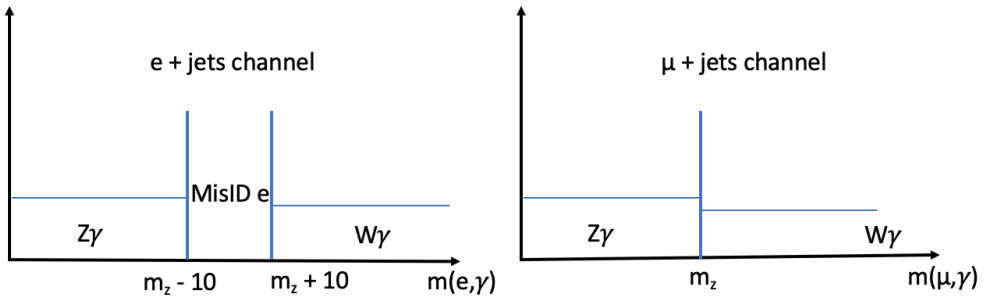


Figure 54: Sketch of the $Z+\gamma$, $W+\gamma$, and misidentified electrons control region in terms of $m(\ell,\gamma)$ the $e+$jets (left) and $\mu+$jets (right) channels.

Process	Gen Photon	misIDEle Photon	Fake Photon	Had Photon
Zγ				
Wγ				
Background (tt, ttγ, V+jets, all)				

Figure 55: The misIDEle template consist of the shape of the distributions in the misidentified electron category. The $Z+\gamma$, $W+\gamma$, and other templates consist of the shape of the events in genuine, hadronic and pile-up photon categories.

The fit templates consist of all the events in the misidentified electron category whereas $Z+\gamma$, $W+\gamma$, and other samples in genuine, hadronic and pile-up photons categories, considered as a non-misidentified electron category, as shown in Figure 55. The template distributions of $m(\ell,\gamma)$ before the fit are shown in Figure 62 (left) for $e+$jets and $\mu+$jets . We used the QCD data driven template as described in Section 6.4 and photons from the QCD template are categorized as hadronic photons to create the misidentification fit templates. The $Z+$jets scale factors are applied to all $Z+$jets events in the template before the extraction of misidentified electron scale factor (misIDSF), $W+\gamma$ scale factor (WGSF), and $Z+\gamma$ scale factor (ZGSF). We used the HiggsCombineTools [107] to measure the best-fit misidentified electron, $Z+\gamma$, and $W+\gamma$ scale factors and their uncertainties

for control region with 1 photon and $N_j \geq 2, N_b = 0$, $N_j = 2, N_b = 0$, $N_j = 3, N_b = 0$, and $N_j \geq 4, N_b = 0$. Details about the HiggsCombineTool can be found in Appendix A. The fit results to these regions in 2016, 2017, 2018 are shown in Table 29 where the parameter of interest (POIs) are misidentified electron SF, $Z + \gamma$ SF, and $W + \gamma$ SF. Figure 56 merely shows the graphical display of the Table 29 data. The signal region requires the presence of a b-tagged jet. The b-tagging SF uncertainties cover the component of these processes that enter the SR with a b-tag requirement via a light jet that is misidentified as a b-tagged jet. The template prefit and postfit plots are shown in Figure 62 for with 1 photon , $N_j \geq 2, N_b = 0$ in $e +$ jets and $\mu +$ jets channels for 2016 dataset. The corresponding profile maximum likelihood scan and impact plots are shown in Figure 57 to 60 respectively for the control region with with 1 photon, $N_j \geq 2, N_b = 0$ for 2016, 2017 and 2018. The full detail plots for each year and each channel can be found in Appendix B.

Table 29: MisIDEle, $Z + \gamma$ and $W + \gamma$ SF for different control regions for 2016, 2017 and 2018

Control Regions	MisIDEleSF	$Z + \gamma$ SF	$W + \gamma$ SF
$N_j \geq 2, N_b = 0$	$2.24^{+0.18}_{-0.17}$	$0.82^{+0.18}_{-0.12}$	$1.15^{+0.18}_{-0.10}$
$N_j = 2, N_b = 0$	$2.21^{+0.18}_{-0.17}$	$0.78^{+0.18}_{-0.13}$	$1.18^{+0.18}_{-0.11}$
$N_j = 3, N_b = 0$	$2.42^{+0.27}_{-0.24}$	$0.80^{+0.27}_{-0.20}$	$1.19^{+0.27}_{-0.14}$
$N_j \geq 4, N_b = 0$	$2.42^{+0.37}_{-0.32}$	$1.17^{+0.37}_{-0.49}$	$1.25^{+0.37}_{-0.21}$
$N_j \geq 2, N_b = 0$	$2.31^{+0.23}_{-0.19}$	$0.87^{+0.23}_{-0.09}$	$1.19^{+0.23}_{-0.10}$
$N_j = 2, N_b = 0$	$2.35^{+0.29}_{-0.20}$	$0.83^{+0.29}_{-0.09}$	$1.17^{+0.29}_{-0.10}$
$N_j = 3, N_b = 0$	$2.28^{+0.25}_{-0.22}$	$1.03^{+0.25}_{-0.21}$	$1.24^{+0.25}_{-0.13}$
$N_j \geq 4, N_b = 0$	$2.49^{+0.41}_{-0.37}$	$0.94^{+0.41}_{-0.67}$	$1.56^{+0.41}_{-0.27}$
$N_j \geq 2, N_b = 0$	$1.63^{+0.10}_{-0.09}$	$0.84^{+0.10}_{-0.08}$	$1.08^{+0.10}_{-0.09}$
$N_j = 2, N_b = 0$	$1.66^{+0.10}_{-0.10}$	$0.81^{+0.10}_{-0.09}$	$1.08^{+0.10}_{-0.09}$
$N_j = 3, N_b = 0$	$1.64^{+0.13}_{-0.13}$	$0.95^{+0.13}_{-0.21}$	$1.11^{+0.13}_{-0.13}$
$N_j \geq 4, N_b = 0$	$1.68^{+0.24}_{-0.21}$	$0.61^{+0.24}_{-0.58}$	$1.32^{+0.24}_{-0.24}$

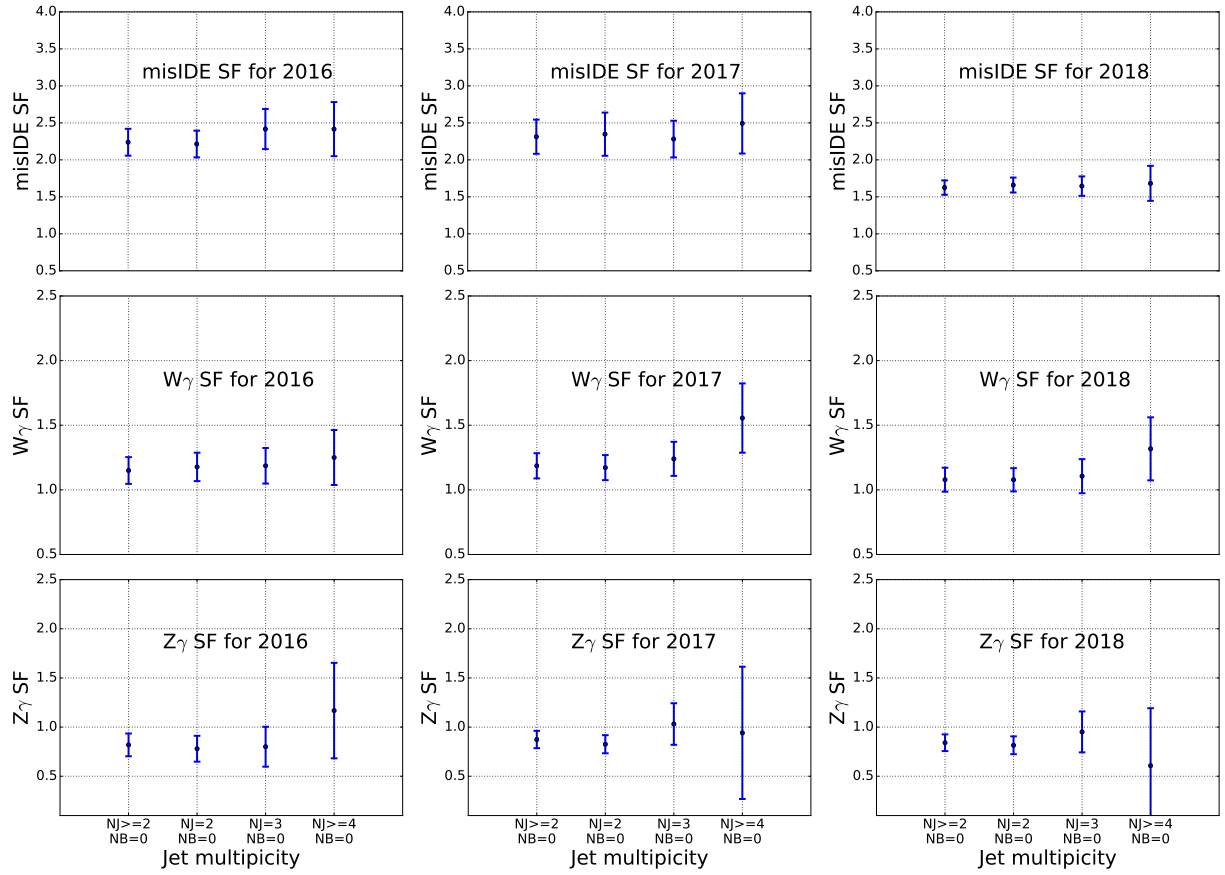


Figure 56: The misidentified electron, $W+\gamma$, and $Z+\gamma$ scale factors comparison for several control regions accross the 2016, 2017, and 2018 dataset from the simultaneous fitting of $e+$ jets and $\mu+$ jets channels.

6.6 data driven Estimation of Non Prompt Photons

In the signal region, the M_3 distribution has contributions from isolated and nonprompt photons. We use a ChIso distribution separated by isolated and nonprompt photon categories to have a better handle on them. We rely on MC simulation for the isolated photons templates. However, a data driven method is used to extract the templates for the nonprompt photons to avoid any mismodeling in simulation and also to mitigate systematic uncertainties. The nonprompt photons dominate the side band region of the $\sigma_{i\eta i\eta}$ distribution, and hence the sideband is used to produce templates from the data dis-

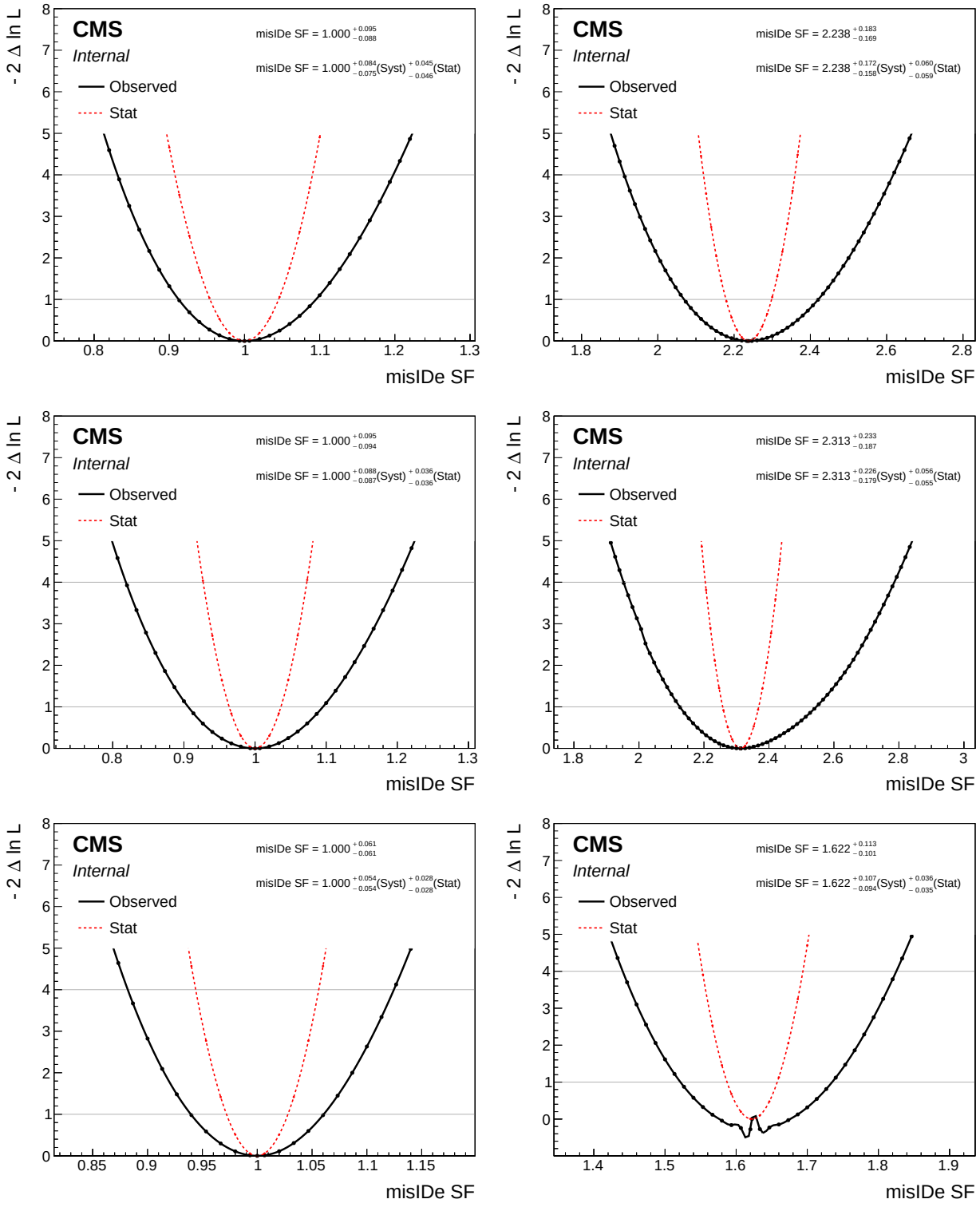


Figure 57: The profiled maximum log-likelihood scan of misidentified electron estimator for Asimov dataset (left) and observed dataset (right) for the control region with 1 photon, $N_j \geq 2, N_b = 0$ for 2016 (top), 2017 (middle) and 2017 (bottom).

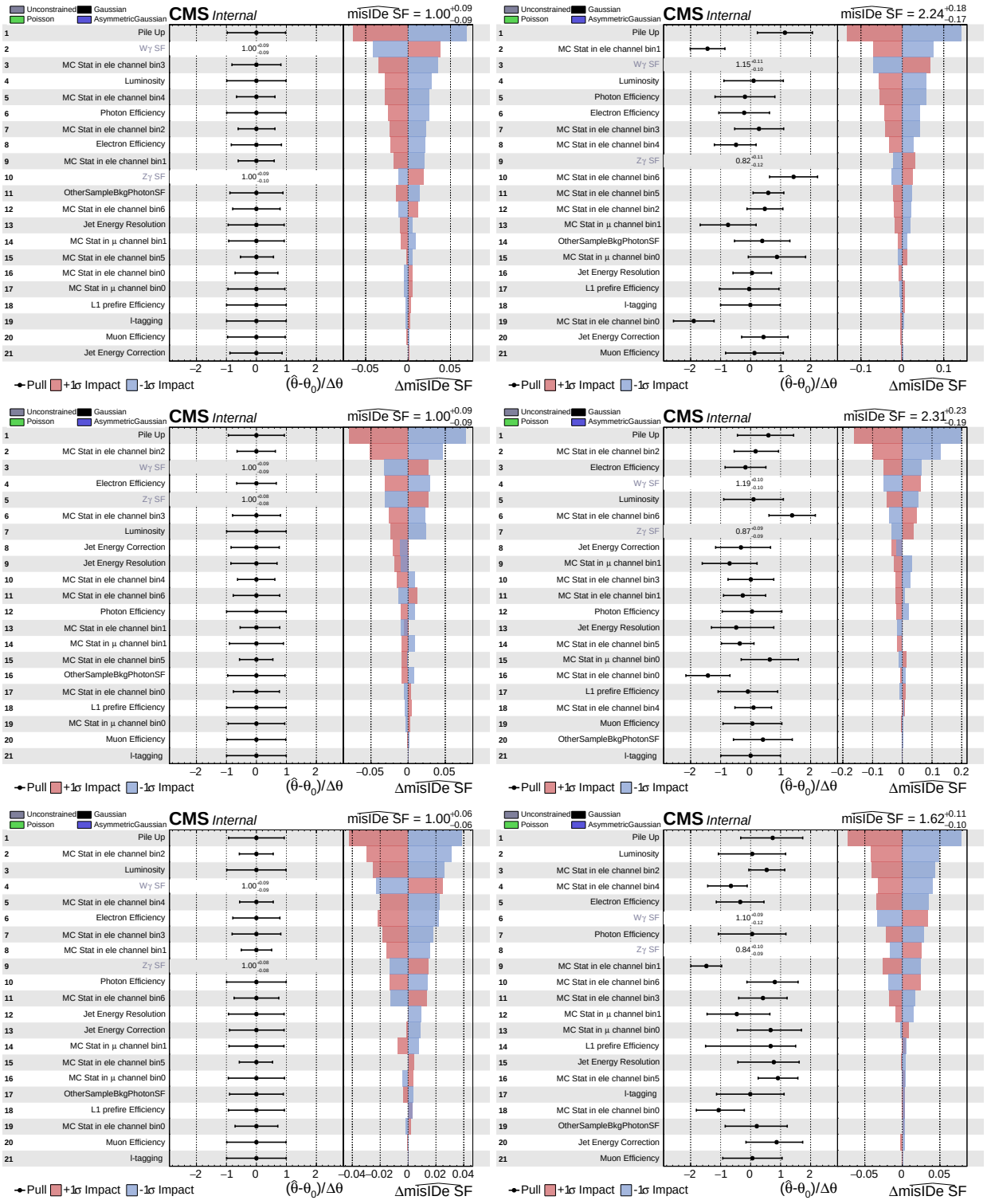


Figure 58: The impact plot that shows postfit pulls and constraints on nuisance parameters for misidentified electron scale factors fit to Asimov dataset (left) and observed dataset (right) for 2016 (top), 2017 (middle) and 2017 (bottom) respectively in the control region with 1 photon , $N_j \geq 2, N_b = 0$.

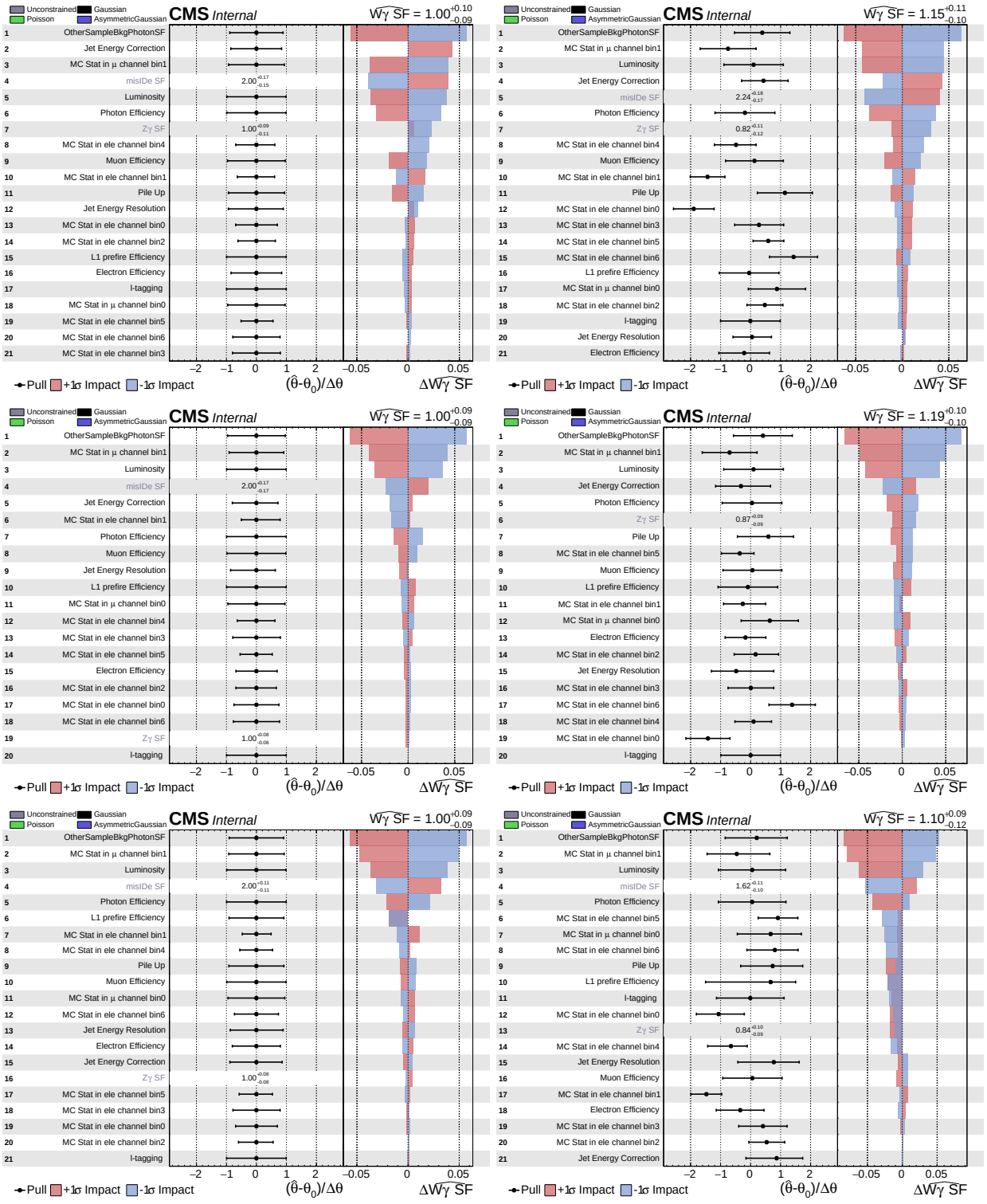


Figure 59: The impact plot that shows postfit pulls and constraints on nuisance parameters for parameter of interest $W + \gamma$ SF fit to Asimov dataset (left) and observed dataset (right) for 2016 (top), 2017 (middle) and 2017 (bottom) respectively in the control region with 1 photon , $N_j \geq 2, N_b = 0$.

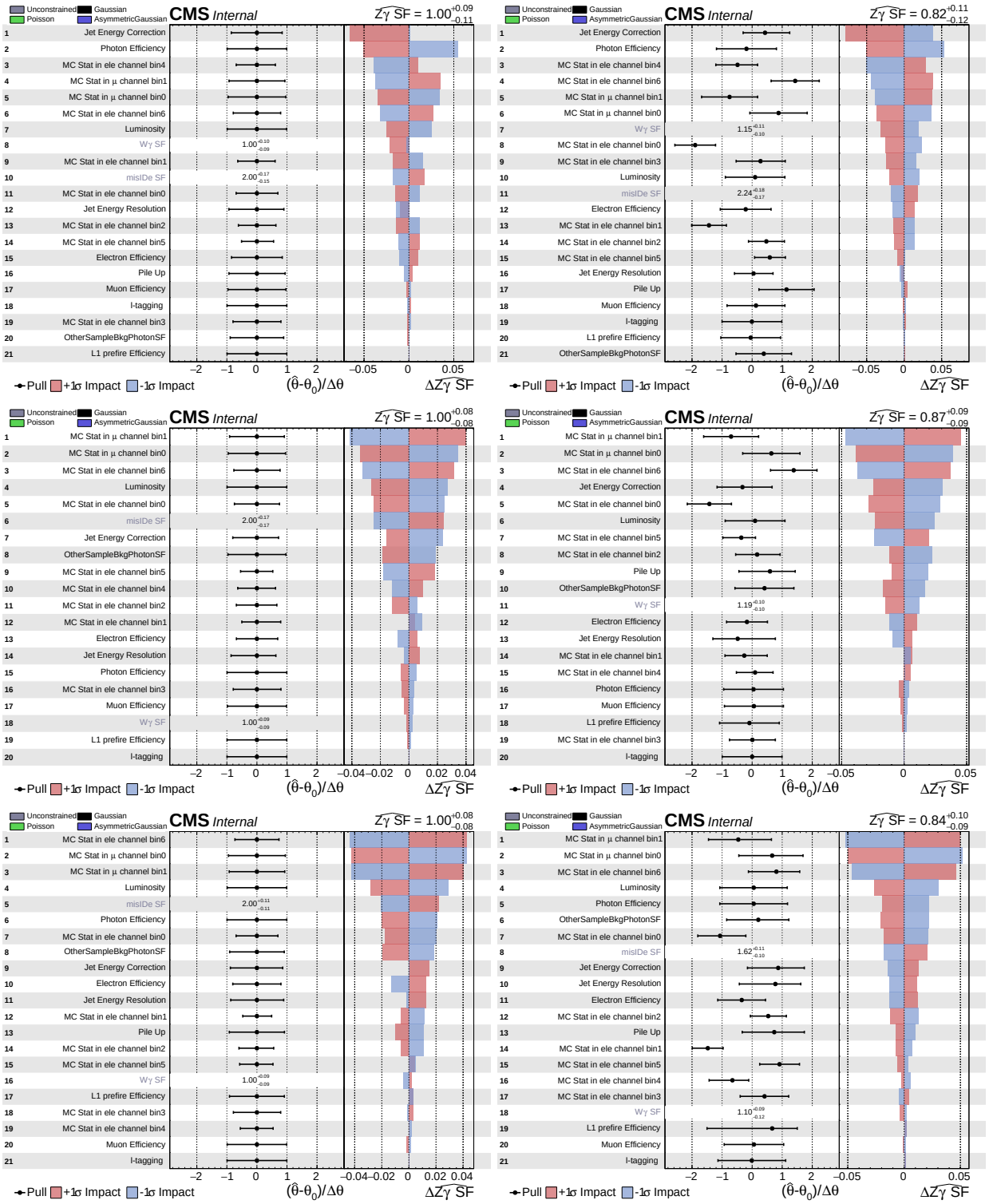


Figure 60: The impact plot that shows postfit pulls and constraints on nuisance parameters for parameter of interest $Z + \gamma$ SF fit to Asimov dataset (left) and observed dataset (right) for 2016 (top), 2017 (middle) and 2017 (bottom) respectively in the control region with 1 photon , $N_j \geq 2, N_b = 0$.

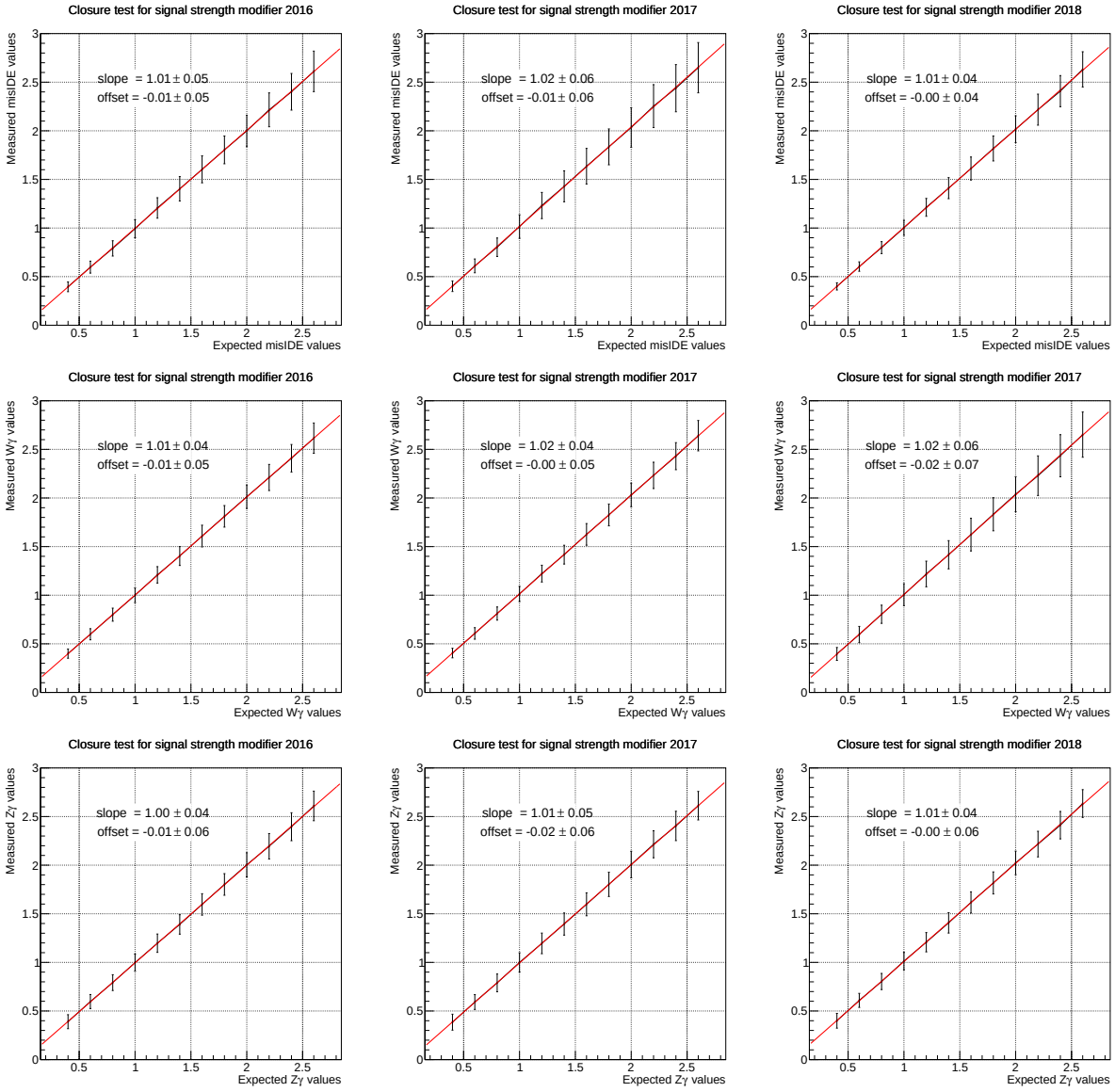
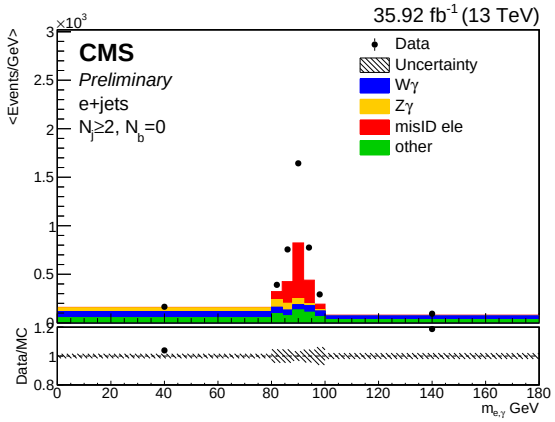
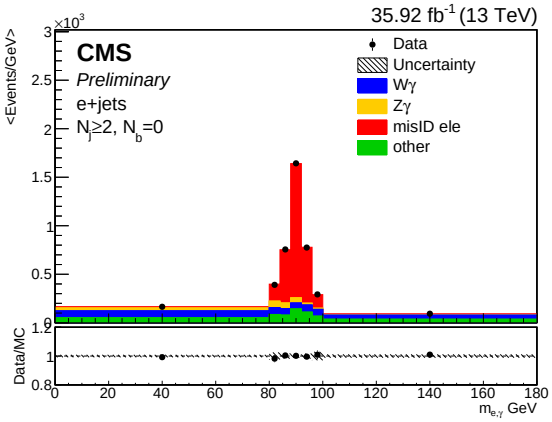


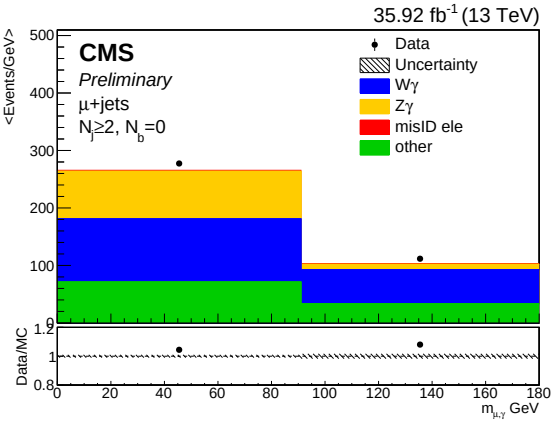
Figure 61: The bias test for misidentified electron scale factor(top), $W + \gamma$ SF(middle) and $Z + \gamma$(bottom) SF with 500 toy data for 2016(left), 2017 (middle) and 2018(right) respectively in the control region with 1 photon , $N_j \geq 2, N_b = 0$.


35.92 fb-1 (13 TeV)
CMS
Preliminary
e+jets
Nj≥2, Nb=0
Data
Uncertainty
Wγ
Zγ
misID ele
other
<Events/GeV>
Data/MC
me,γ GeV

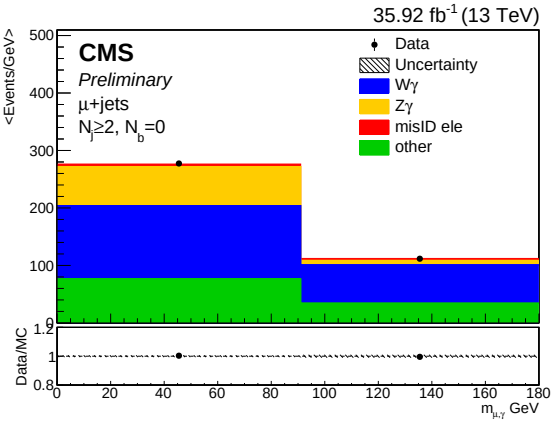
(a)


35.92 fb-1 (13 TeV)
CMS
Preliminary
e+jets
Nj≥2, Nb=0
Data
Uncertainty
Wγ
Zγ
misID ele
other
<Events/GeV>
Data/MC
me,γ GeV

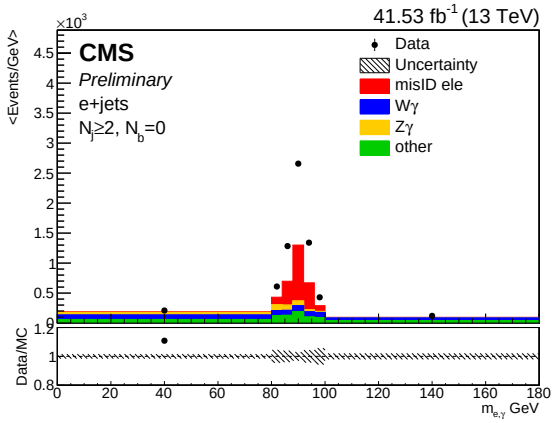
(b)


35.92 fb-1 (13 TeV)
CMS
Preliminary
μ+jets
Nj≥2, Nb=0
Data
Uncertainty
Wγ
Zγ
misID ele
other
<Events/GeV>
Data/MC
mμ,γ GeV

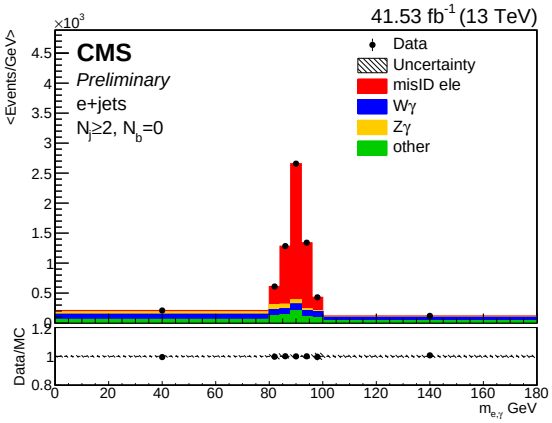
(c)


35.92 fb-1 (13 TeV)
CMS
Preliminary
μ+jets
Nj≥2, Nb=0
Data
Uncertainty
Wγ
Zγ
misID ele
other
<Events/GeV>
Data/MC
mμ,γ GeV

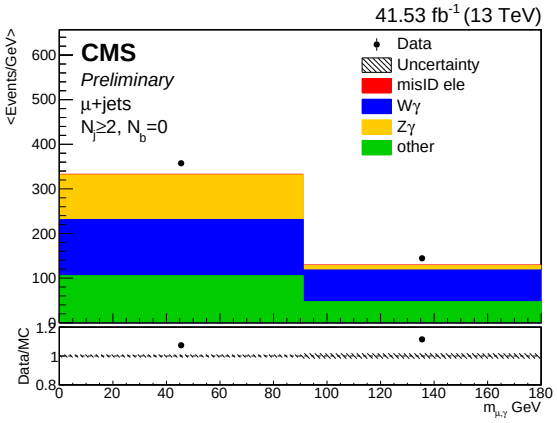
(d)


41.53 fb-1 (13 TeV)
CMS
Preliminary
e+jets
Nj≥2, Nb=0
Data
Uncertainty
misID ele
Wγ
Zγ
other
<Events/GeV>
Data/MC
me,γ GeV

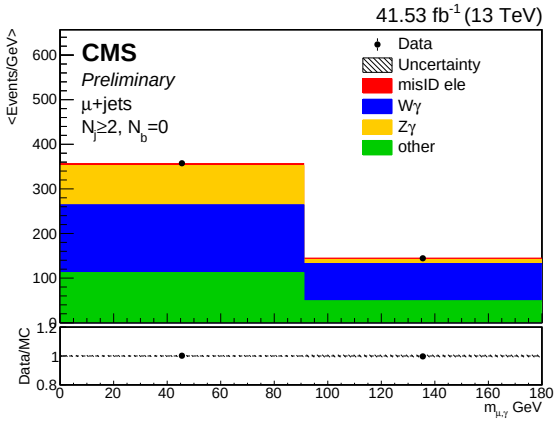
(e)


41.53 fb-1 (13 TeV)
CMS
Preliminary
e+jets
Nj≥2, Nb=0
Data
Uncertainty
misID ele
Wγ
Zγ
other
<Events/GeV>
Data/MC
me,γ GeV

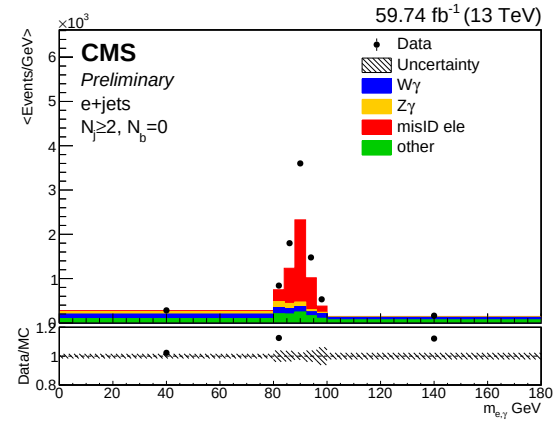
(f)



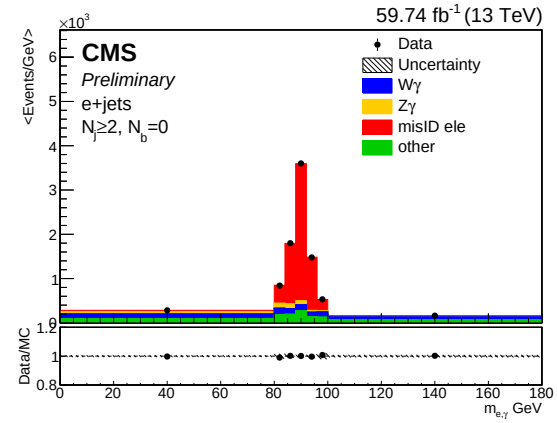
(g)



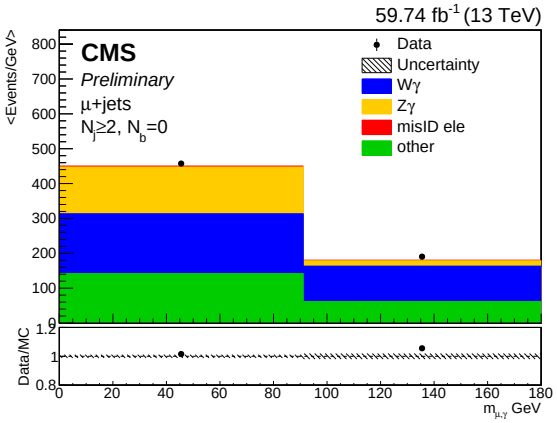
(h)



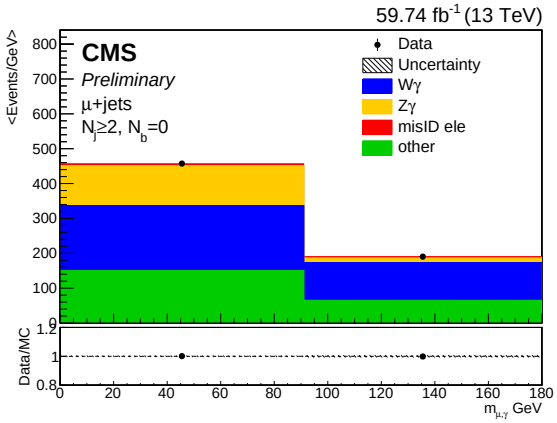
(i)



(j)

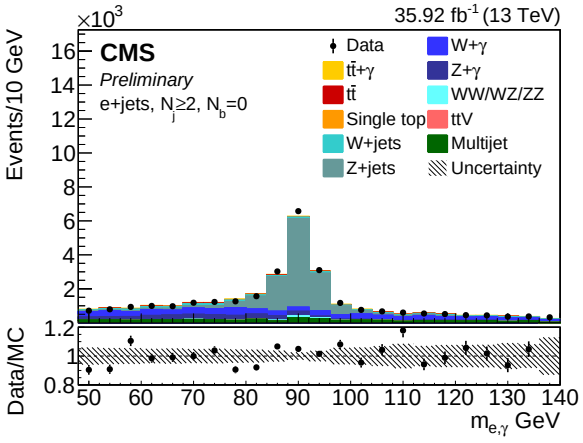


(k)

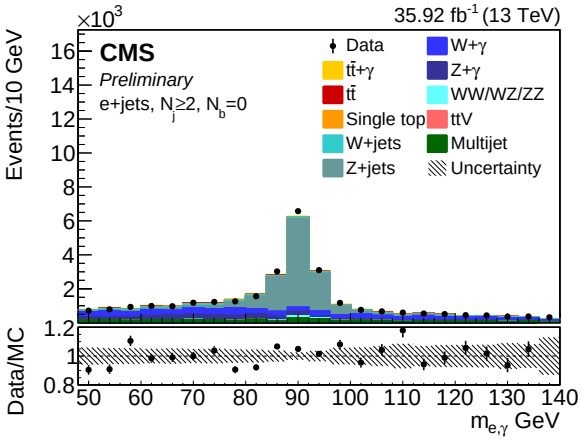


(l)

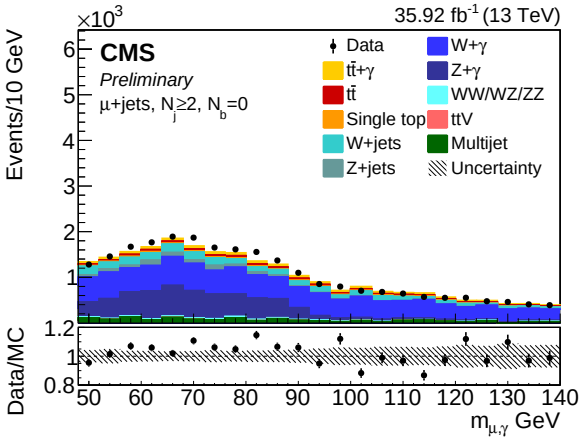
Figure 62: The prefit (left) and postfit (right) templates in a control region with with 1 photon , $N_j \geq 2, N_b = 0$ in the $e+$jets channel (a,b,e,f,i,j) and the $\mu+$jets channel (c,d,g,h,k,l).


×10^3
35.92 fb^-1 (13 TeV)
CMS
Preliminary
e+jets, N_j≥2, N_b=0
Data
W+γ
tt̄+γ
Z+γ
tt̄
WW/WZ/ZZ
Single top
ttV
W+jets
Multijet
Z+jets
Uncertainty
Events/10 GeV
Data/MC
m_e,γ GeV

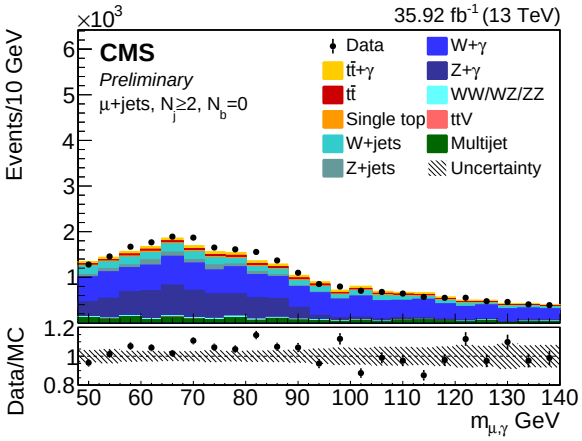
(a)


×10^3
35.92 fb^-1 (13 TeV)
CMS
Preliminary
e+jets, N_j≥2, N_b=0
Data
W+γ
tt̄+γ
Z+γ
tt̄
WW/WZ/ZZ
Single top
ttV
W+jets
Multijet
Z+jets
Uncertainty
Events/10 GeV
Data/MC
m_e,γ GeV

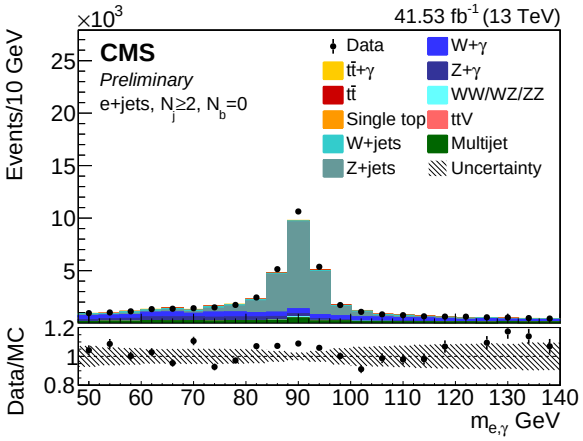
(b)


×10^3
35.92 fb^-1 (13 TeV)
CMS
Preliminary
μ+jets, N_j≥2, N_b=0
Data
W+γ
tt̄+γ
Z+γ
tt̄
WW/WZ/ZZ
Single top
ttV
W+jets
Multijet
Z+jets
Uncertainty
Events/10 GeV
Data/MC
m_μ,γ GeV

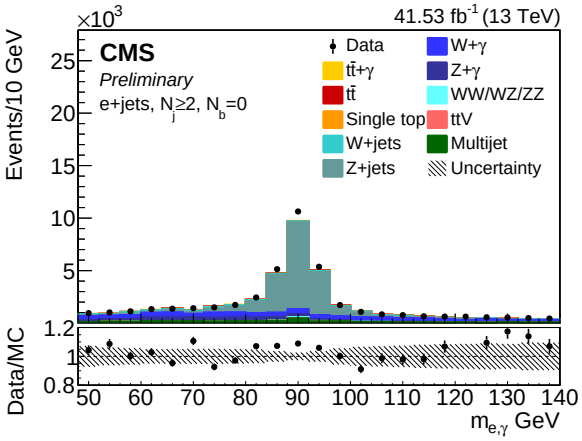
(c)


×10^3
35.92 fb^-1 (13 TeV)
CMS
Preliminary
μ+jets, N_j≥2, N_b=0
Data
W+γ
tt̄+γ
Z+γ
tt̄
WW/WZ/ZZ
Single top
ttV
W+jets
Multijet
Z+jets
Uncertainty
Events/10 GeV
Data/MC
m_μ,γ GeV

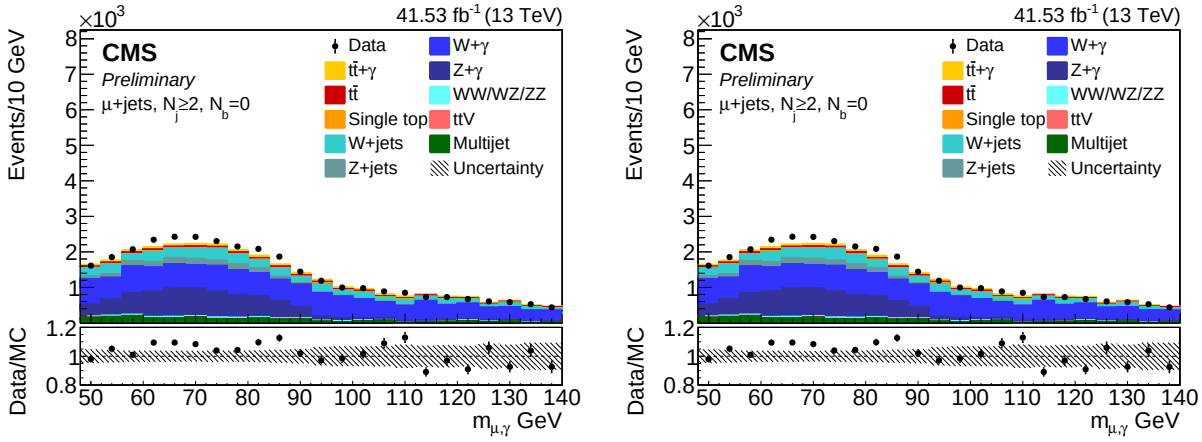
(d)


×10^3
41.53 fb^-1 (13 TeV)
CMS
Preliminary
e+jets, N_j≥2, N_b=0
Data
W+γ
tt̄+γ
Z+γ
tt̄
WW/WZ/ZZ
Single top
ttV
W+jets
Multijet
Z+jets
Uncertainty
Events/10 GeV
Data/MC
m_e,γ GeV

(e)

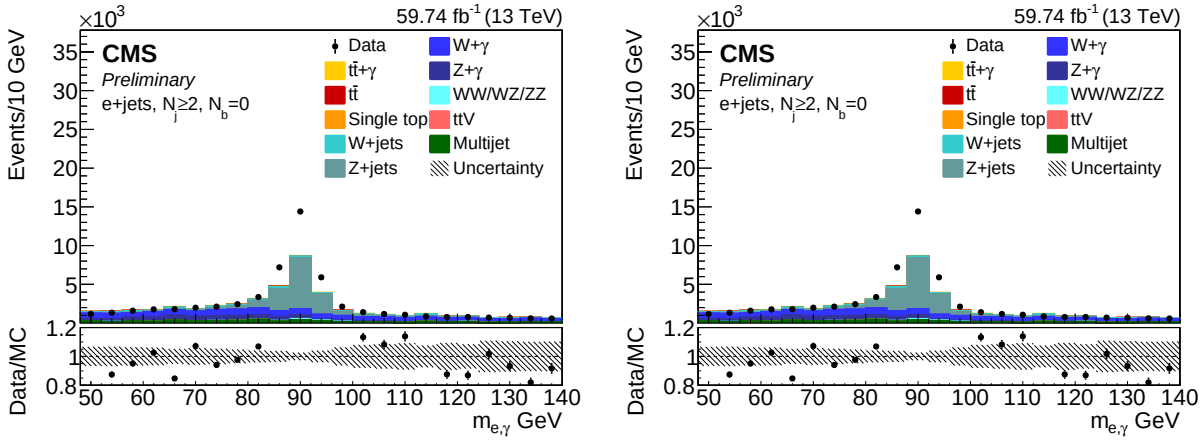

×10^3
41.53 fb^-1 (13 TeV)
CMS
Preliminary
e+jets, N_j≥2, N_b=0
Data
W+γ
tt̄+γ
Z+γ
tt̄
WW/WZ/ZZ
Single top
ttV
W+jets
Multijet
Z+jets
Uncertainty
Events/10 GeV
Data/MC
m_e,γ GeV

(f)



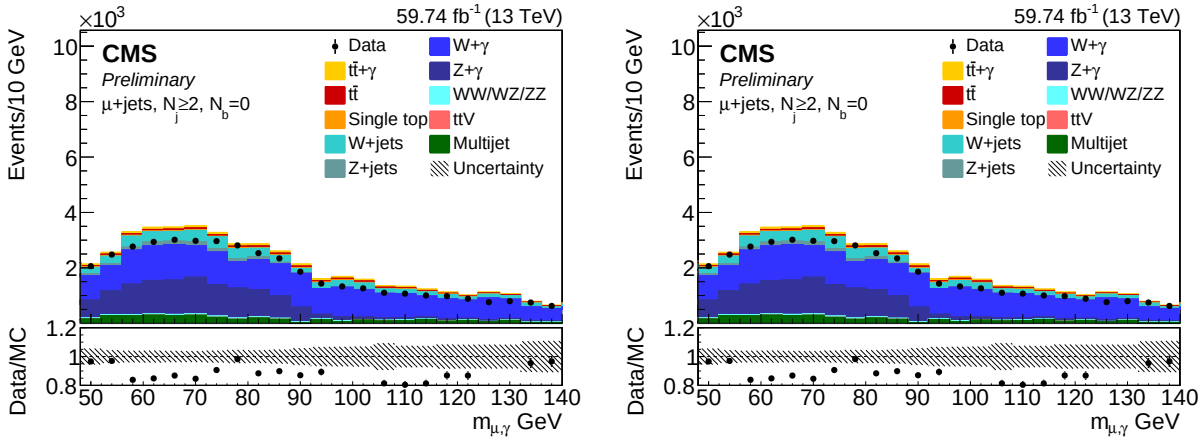
(g)

(h)



(i)

(j)



(k)

(l)

Figure 63: The prefit and postfit plots in a control region with with 1 photon, $N_j \geq 2, N_b = 0$ in the $e+$jets channel (a,b,e,f,i,j)) and the $\mu+$jets channel (c,d,g,h,k,l).

tribution and normalized to the different nonprompt processes in the simulation.

The ChIso variable in the sideband region of the $\sigma_{i\eta i\eta}$ is dominated by nonprompt photons as shown in Figures 64, 65, and 66. We take the shape of the nonprompt photon templates from the data in the $\sigma_{i\eta i\eta}$ sideband region and normalize it to the different processes in simulation. The detail plots for each year and each channel can be found in Appendix F.

We compare the different windows of the $\sigma_{i\eta i\eta}$ in the sideband region to check whether the shape depends on the $\sigma_{i\eta i\eta}$ cut or not. The shape of nonprompt photon is studied in three different categories; nonprompt photon shape in the signal region, nonprompt photon shape in the $\sigma_{i\eta i\eta}$ sideband in data, and nonprompt photons in MC from sideband as shown in Figures 67. The shapes of the templates are more or less similar, the small variation is taken into account as a systematic uncertainty in the final fit.

6.7 Extraction of $\mathrm{t\bar{t}\gamma}$ Cross Section

We measured two observables, signal strength parameter ($\mu = \frac{\sigma_{obs}}{\sigma_{SM,NLO}}$) and ratio ($R_\gamma = \frac{\sigma_{\mathrm{t\bar{t}\gamma}}}{\sigma_{\mathrm{t\bar{t}}}}$). First, the inclusive $\mathrm{t\bar{t}\gamma}$ cross section is obtained by multiplying the μ with $\sigma_{SM,NLO}$. The R_γ allows for the cancellation of factors and their associated systematic uncertainties, such as luminosity, ISR/FSR systematics, trigger efficiency, etc. The $\sigma_{\mathrm{t\bar{t}\gamma}}$ cross section is obtained by multiplying the measured value of R_γ with a recent $\mathrm{t\bar{t}}$ cross section measurement.

The RooStats-based statistics tool provided by the Higgs PAG [107] is used to create a physics model for the analysis. To measure the production cross section of $\mathrm{t\bar{t}\gamma}$, we define a model that predicts different expected yield for the signal and fit the parameters of this

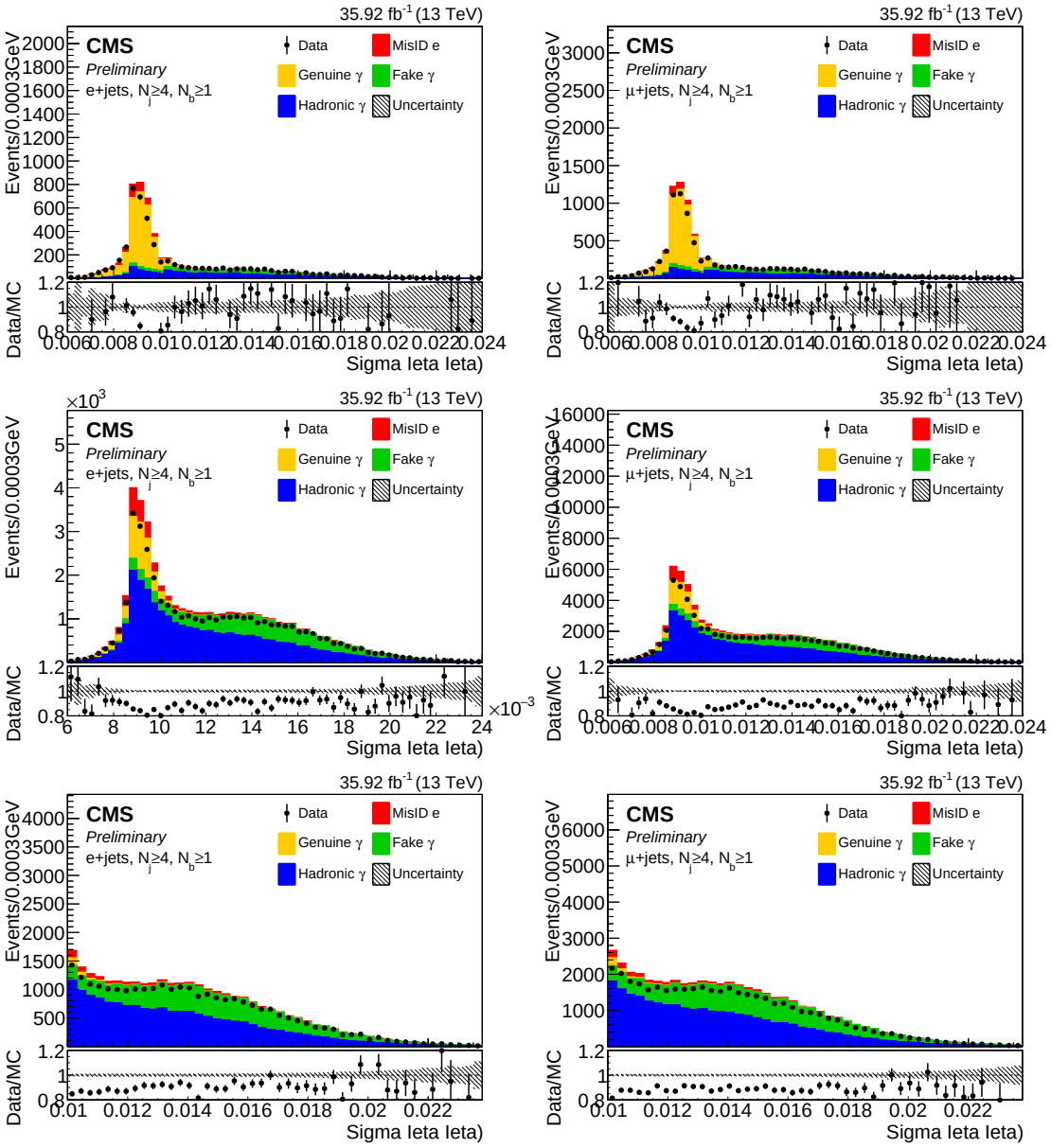


Figure 64: The distribution of $\sigma_{i\eta i\eta}$ variable (top), with no cut in ChIso (middle), and focused on the sideband in $e+$jets (left) and $\mu+$jets (right) for 2016. The sideband region is $0.011 < \sigma_{i\eta i\eta} < 0.02$ in barrel.

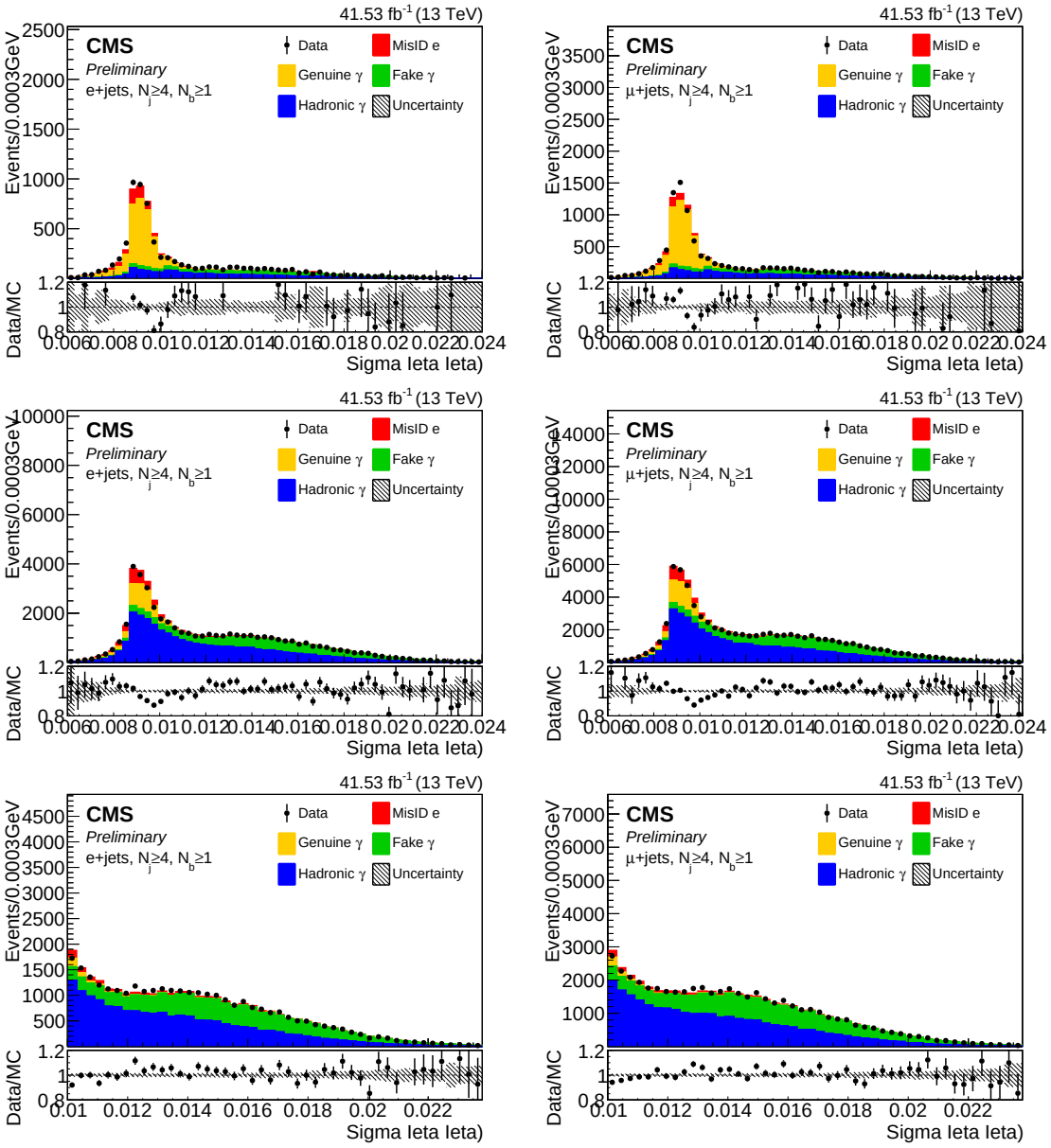


Figure 65: The distribution of $\sigma_{i\eta i\eta}$ variable (top), with no cut in ChIso (middle), and focused on the sideband in $e+$jets (left) and $\mu+$jets (right) for 2017. The sideband region is $0.011 < \sigma_{i\eta i\eta} < 0.02$ in barrel.

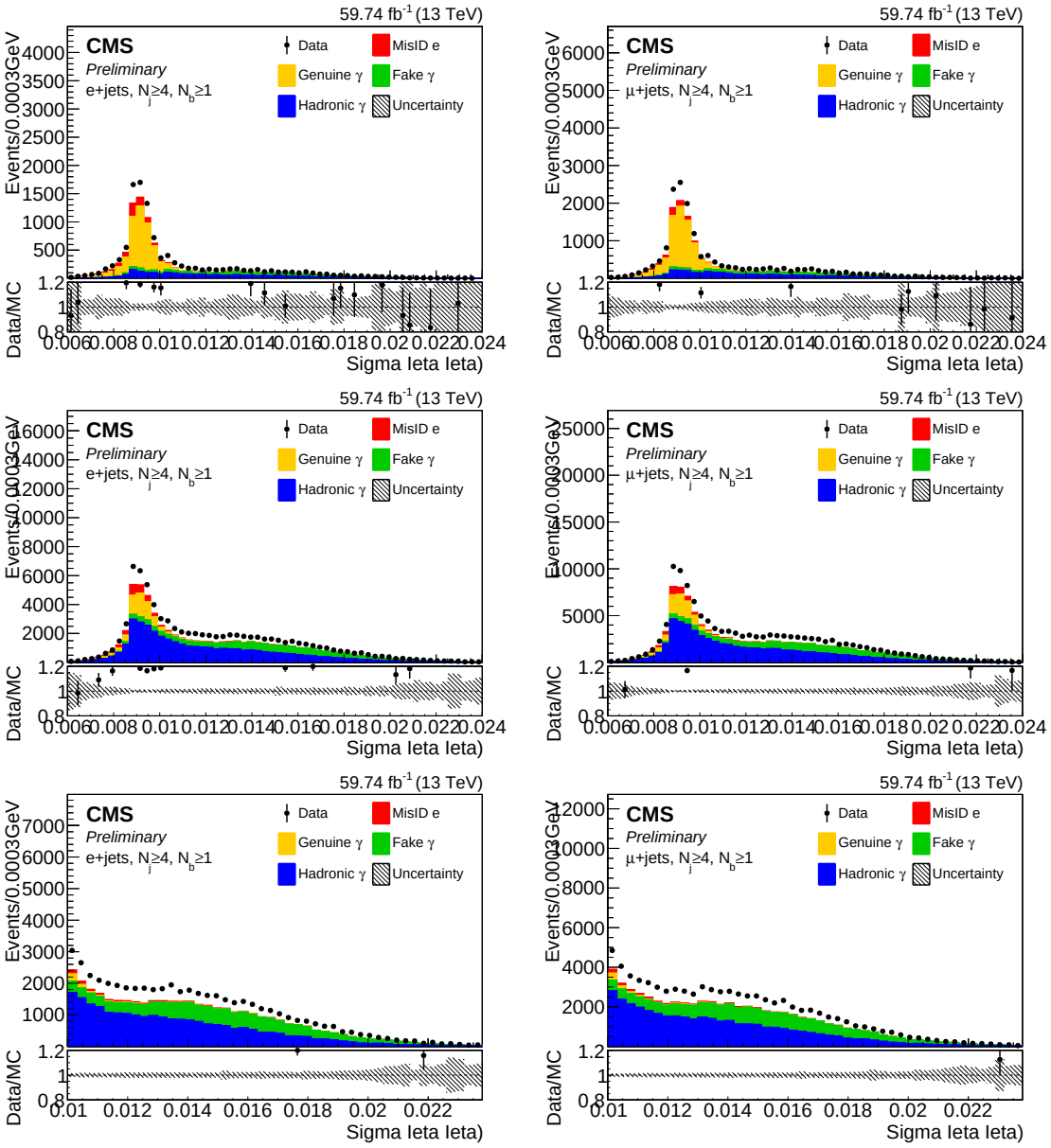


Figure 66: The distribution of $\sigma_{i\eta i\eta}$ variable (top), with no cut in ChIso (middle), and focused on the sideband in $e+$jets (left) and $\mu+$jets (right) for 2018. The sideband region is $0.011 < \sigma_{i\eta i\eta} < 0.02$ in barrel.

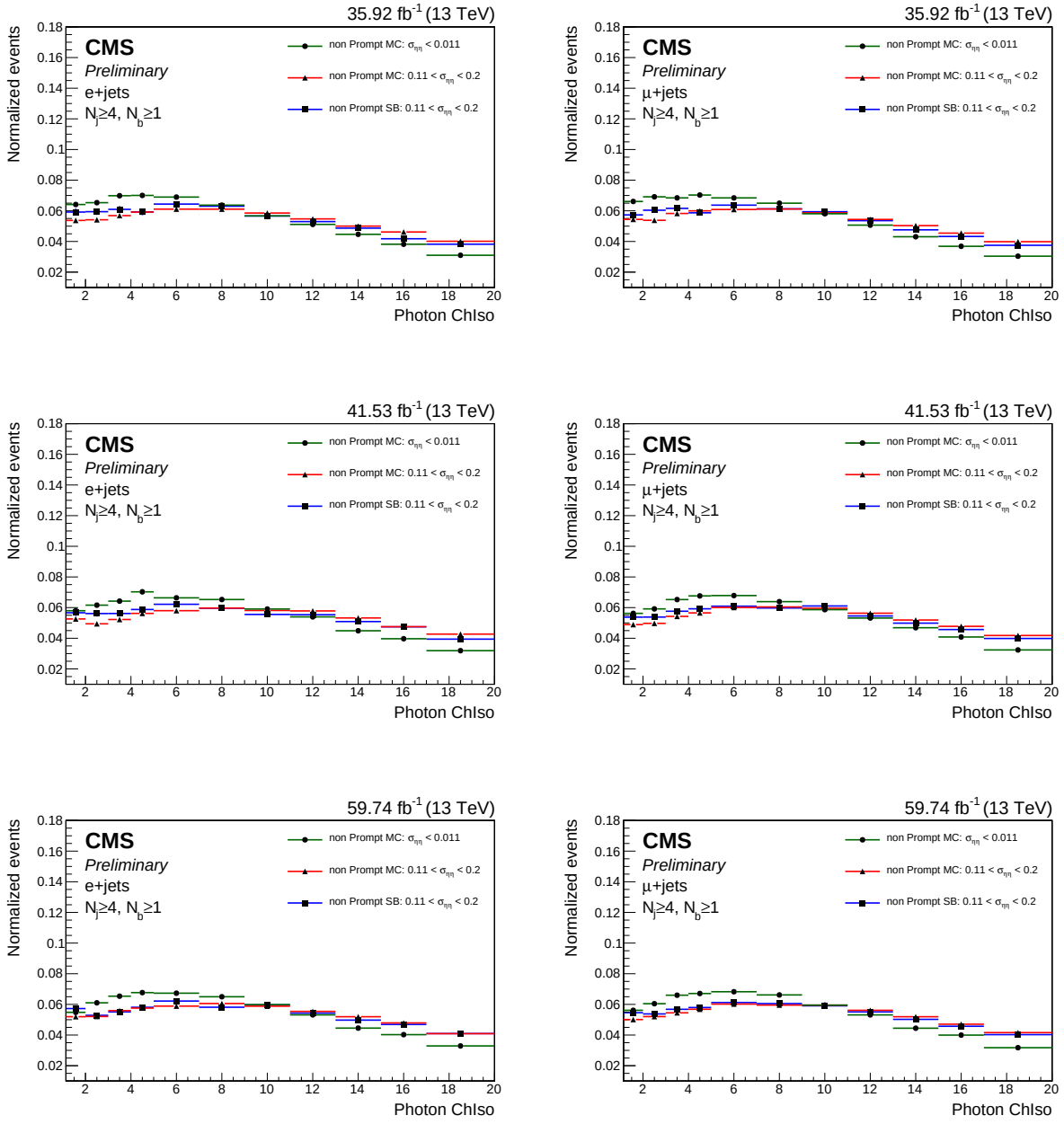


Figure 67: Comparison of nonprompt photon shape in the signal region, nonprompt photon shape in $\sigma_{i\eta i\eta}$ sideband in data, nonprompt photons in MC from sideband for $e+$jets (left) and $\mu+$jets (right) for 2016 (top), 2017 (middle), and 2018 (bottom) data.

model to the data. The model is a function of the observed data, signal strength, and expected signal and background rates where they are expressed as a function of nuisance parameters. For the binned shape template, a physics model is the product of models for each bin. A datacard contains the information about the shape of observed and expected events for various physical processes and their systematic variations. The datacard defines the likelihood function used in the HiggsCombineTool. The test statistic is the profile likelihood ratio, $q(r) = -2\ln L(r,\hat{\theta}_{\mathrm{r}})/L(\hat{r},\hat{\theta})$, where $\hat{\theta}_{\mathrm{r}}$ are the values of the nuisance parameters that maximize the likelihood function for signal strength r. The function $q(r)$ is used to determine the uncertainties on the fitted value of r. The $q(r)$ can be treated as a χ^2 with one degree of freedom which means 68% confidence interval (CI) for r is defined by $q(r) \leq 1$.

A simultaneous binned maximum likelihood fit is done to extract the signal strength parameter value on $\mathrm{t\bar{t}}\gamma$ cross section using the two distributions: photon charge hadron isolation (ChIso) and M_3. The shape of these distributions are used to discriminate the signal from the background in the analysis. The expected shape of the ChIso distributions is distinguishable for genuine photons and for nonprompt photons. Similarly, the expected shape of the M_3 distribution distinguishes top from non-top events as shown in Figure 40.

Photon charged hadron isolation (ChIso) is the sum of the transverse momenta of all the charged hadrons that lies within a cone of radius $\Delta\mathrm{R} \leq 0.4$ of the photon. Prompt photons, originating from physics processes such as $\mathrm{t\bar{t}}\gamma$, $W+\gamma$, etc, will have fewer nearby particles, whereas nonprompt photons, originating from hadronic activity, will be surrounded by more particles. The variable the M_3 is an invariant mass of the 3 jets have the highest vector summed $\mathrm{P_T}$. Hence, the simultaneous maximum likelihood fit of ChIso, distinguishing genuine and non-prompt photons, and M_3 can differentiate the signal ($\mathrm{t\bar{t}}\gamma$with genuine

photon) from all other background processes such as $\mathrm{t\bar{t}}$ with non-prompt photons; $W + \gamma$ and $Z + \gamma$ with genuine photons; and SingleTop, $V +$ jets , Diboson, $\mathrm{t\bar{t}}$ V, and QCD with non-prompt photons.

CHAPTER 7 SOURCES OF SYSTEMATIC UNCERTAINTIES

7.1 Systematic Uncertainties

The precise measurement of the cross section is affected by numerous uncertainties arising from theoretical mismodeling, experimental bias, statistical fluctuations, and background estimation of the measurement. The theoretical systematic uncertainties are obtained by changing the corresponding values of the sources. For example, we vary the renormalization (μ_R) and factorization scale (μ_F) during the MC simulation to get the systematic effect of higher order contribution to the ME calculation. Then, we did the maximum log-likelihood fit to estimate the corresponding fit values due to the shift of the renormalization and factorization scale in comparison to the nominal value. The difference will be the uncertainty on measured POI due to the uncertainty. [108]

The typical origins of systematic uncertainties are imperfect modeling and simulation, uncertainties in parameters used, miscalibration of the detector, uncertainties associated with the corrections, and theoretical limitations. We divide systematic uncertainties into three different categories: experimental uncertainties related to uncertainties of the physics object selection efficiencies and uncertainties of their calibrations; theoretical uncertainties in normalization of the simulated processes, modeling of parton shower, hadronization, and underlying events; and uncertainties related to background estimations. We followed the "Top systematic uncertainties for Run 2" recommendation from Ref. [109] for estimation of experimental and theoretical uncertainties. These uncertainties are correlated year by year and shown in Table 30 Table 31 summaries the systematic uncertainties in the inclusive measurement discussed in Sec. 8.

7.1.1 Experimental Uncertainties

- **Luminosity uncertainty:** The integrated luminosity is 35.92 fb^{-1}, 41.53 fb^{-1}, and 59.74 fb^{-1} for 2016, 2017, and 2018 respectively. Uncertainties of 2.5$\%$, 2.3%, and 2.5% on the luminosity are used to calculate the systematic uncertainties for 2016, 2017, and 2018 respectively [110–112].

- **Pile-up uncertainty:** The average pile-up is calculated using an inelastic cross section of 69.2 mb for all 3 years. The uncertainty in the pileup estimation is affected by the uncertainty in the inelastic cross section and the uncertainty in the luminosity. The luminosity uncertainty is smaller than the minimum bias cross section uncertainties. The cross section is varied up and down by 4.6% to produce the re-weight distributions. The difference between the reweighed and nominal distributions represent the pile-up uncertainty. It is fully correlated across the years.

- **b-tagging uncertainty:** The analysis uses the medium b-tagging combined secondary vertex version 2 working point (WP) of 0.8484. A b-tagging scale factors is applied on an event by event basis as a function of jet P_T and η. We introduce a variation of the b-tagging uncertainty varied by $\pm\sigma$ according to the BtagPOG[16] recommendation. The variations are performed by recalculating the event probability using all jets and their true flavor (such as light and heavy flavor) using method_1a of Ref. [94]

- **Lepton efficiency uncertainty:** A combination of lepton trigger, identification and isolation scale factors makes up the lepton efficiency in this analysis. The uncertain-

[16] BtagPOG is the physics object group (POG) expert on btagging and vertexing [113].

ties on scale factors for electron and muon trigger, identification and isolation are taken as recommended by the TOP EGM of Ref. [114] and MUOPOG of Ref.[115].

- **Photon efficiency uncertainty:** The photon efficiency scale factors are calculated to take into account effects from identification and pixel seed veto ID. The uncertainties on those variables are taken from EGM POG and adopted in our analysis as in Ref. [114].

- **Jet energy correction uncertainty:** The jet energy measured in the detector is corrected to match the jet energy at the particle level simulations.The jet energy scale is classified into different sources: `Absolute`, `Absolute_201*`, `BBEC1`, `BBEC1_201*`, `EC2`, `EC2_201*`, `HF`, `HF_201*`, `RelativeSample_201*`, `RelativeBal`, `FlavorQCD`. We follow the POG recommendations about their year-by-year correlations [116]. The differences in the distributions are obtained by rescaling the jet momenta in the simulation.

- **Jet energy resolution uncertainty:** Separate uncertainties in the jet energy resolution are introduced. The uncertainties are uncorrelated among the years. The differences in the distributions are obtained by rescaling the jet momenta in the simulation. The JER parameters are `Summer16_25nsV1_MC`, `Fall17_V3_MC`, and `Autumn18_V7_MC`.

- **Photon and electron scaling uncertainty:** The EGM POG provides residual corrections to scale the data to the MC and systematic uncertainties for each year of Run 2. The scale uncertainties correct for slew-rate in ECAL electronics which affects the pulse shape of the reconstructed electrons and photons which encounter

the gain-switch turning point. The nominal scale value is applied to data only and the uncertainties are used to calculate the up and down systematic uncertainties in simulation. The scaling factors are implemented before object selections.

- **Photon and electron smearing uncertainty:** The EGM POG provides residual corrections to smear the MC to the resolution matched in data and systematic uncertainties for each year of Run 2. The smearing correction has two components: rho and phi. It is a guassian with a mean of 1 and sigma of the given uncertainty.

7.1.2 Theoretical Uncertainties

- **Factorization and Renormalization scales uncertainty:** The effect of higher order contributions to the ME calculation are estimated per event by varying the renormalization (μ_R) and factorization scale (μ_F).

  * first fix μ_R and vary μ_F by 2 and 1/2 (for up and down uncertainties respectively).
  * then fix μ_F and vary μ_R by 2 and 1/2
  * the combination where both μ_F, μ_R are either up or down.

  A total of seven combinations are accounted for in the higher order correction and are included as a shape uncertainty in the fit. Since we use the same method of ME calculation, the uncertainties are fully correlated among the years.

- **Parton distribution function uncertainty (PDF):** We use 100 replicas generated with NNPDF3.1 and construct an envelope from the acceptances using the corresponding event weights. On the $\mathrm{t\bar{t}}\gamma$ sample, the PDF uncertainties affect the shape

while keeping the normalization constant on the $\mathrm{t\bar{t}}$ background PDF normalization. The uncertainties are correlated across years.

- **Parton Shower (PS) uncertainties:** The sources of uncertainties that aries from the modelling of perturbative and non perturbative QCD effects are studied in parton shower MC generators. First a single parton shower model approach is considered and all the possible aspects are modeled such as ISR, FSR, and ME-PS matching. Although underlying event uncertainties and color reconnection uncertainties are not purely parton shower, for convenience they are included under the PS category. PYTHIA 8 is used as a default MC which means energy scales and efficiencies (i.e., isolation efficiency, jet energy scale, and b-tagging) are calibrated using PYTHIA 8 as the nominal MC.

- **ISR/FSR uncertainty:** The value of the strong coupling constant α_{s} at scale Q^2 has corresponding uncertainties which are evaluated by varying the renormalization scale for QCD emissions in initial and final state radiation by a factor of 2 and 0.5. These variation effectively vary the values of α_{s} and contain NLO compensation that preserve the soft-gluon limits.

- **Color reconnection uncertainty:** The nominal color reconnection model is based on multi-parton interaction (MPI). The other models are QCD-inspired (CR1) and gluon move (CR2). The largest difference between these variations and nominal is used to assess the uncertainties in color reconnection.

- **Statistical uncertainty:** The Barlow-Beeston ”light” bin-by-bin nuisances [117] are

used to account for the statistical uncertainties in simulated event yields above an event count of 20 events. Below that, individual Poissonian nuisances separately for each background component are included.

7.1.3 Background Estimation Uncertainties

- **Drell Yan normalization uncertainty:** The DY cross section uncertainty is measured in 0 b-tag and 0 γ control regions and found to be 5% where as an 8% uncertainty is applied for control or signal regions with greater than or equal to 1 btag.

- **QCD normalization uncertainty:** We estimated two QCD normalization uncertainties for 0 b-tag and $\geq$1 b-tag control regions and found both to be 50%. The uncertainty accounts for mismodeling of the QCD data driven template.

- **Non prompt photon estimation uncertainty:** The data driven templates are created from the $\sigma_{i\eta i\eta}$ sideband region of the photon charged hadron isolation (ChIso) variable. To incorporate the uncertainties in the template creation, a bin by bin difference of the data driven template from the nonprompt photon template in simulation is used to estimate the up and down uncertainties.

- **Misidentified electron uncertainty:** The uncertainty of 20% from the misID fit is used to make up and down templates which are incorporated in the signal region fit as a nuisance parameters. The additional 10% uncertainties are added in signal region to account for a possible differences in Drell-Yan and $\mathrm{t\bar{t}}$ dominated regions.

- **$W + \gamma$and $Z + \gamma$ SF estimation uncertainty:** The cross section uncertainties on $W + \gamma$, and $Z + \gamma$ are 19% and 21% respectively. These uncertainties are applied

on signal region fitting as a gaussian constraint with a $\pm 1\sigma$. These uncertainties are estimated from the simultaneous fitting of the invariant mass distribution of photon and lepton in $e+$ jets and $\mu+$ jets channels.

- **Other uncertainties:** The "other" samples mainly consists of multiboson processes and a normalization uncertainty of 30% is used.

- **Additional uncertainties on $\mathrm{t\bar{t}}$ processes:** The $\mathrm{t\bar{t}}$ background process has a unit ($\sigma = 1$) gaussian constraint applied on them. The cross section uncertainty on $\mathrm{t\bar{t}}$ is 5.5%

Table 30: List of experimental and background estimation related uncertainties and their correlations among years.

	Source	16/17 (%)	16/18 (%)	17/18 (%)	Comments
Experimental	Luminosity	21	29	30	Using Lumi POG recomm., correlation from Ref. [118].
	Pileup	100	100	100	4.6 % on σ(MB)=69.2 mb
	Jet energy correction	mult.	mult.	mult.	reduced sources (6, JME POG)
	Jet energy resolution	0	0	0	JME POG, uncorrelated
	Electron trigger, reco, and id	100	100	100	EGM recommendations
	Muon trigger, reco, and id	100	100	100	MUO recommendations
	Photon and pixel seed veto id	100	100	100	EGM recommendations
	B-tagging	0	0	100	BTV contact recommendations
	EGamma energy scale/resolution	100	100	100	EGM recommendations
Theoretical	L1 prefiring	100	-	-	only in 2016 and 2017
	Tune	100	100	100	using CP5 tune for all years
	color reconnection	100	100	100	erdOn, QCD based, gluon Move
	ISR	100	100	100	using same α_S for all years
	FSR	100	100	100	using same α_S for all years
	Q2	100	100	100	using same α_S for all years
	PDF	100	100	100	using same PDF set for all years
	ME-PS matching(h_{damp})	100	100	100	
Background	QCD normalization	100	100	100	50 %, see Sec. 6.4.
	$\mathrm{t\bar{t}}$ normalization	100	100	100	5.5 % on $\mathrm{t\bar{t}}$ MC
	misidentified electron	0	0	0	free parameter, extract SF
	Zγ normalization	100	100	100	free parameter, extract SF
	Wγ normalization	100	100	100	free parameter, extract SF
	"Other" bkg. normalization	100	100	100	30 % on "other"

Table 31: Summary of the sources of systematic uncertainties. The first column lists the source of the uncertainty. The last column gives the corresponding systematic uncertainty in the $t\bar{t}\gamma$ cross section using the fit result.

Nuisance Parameter	Total uncertainty on $t\bar{t}\gamma$	Relative Uncertainty (%)
Photon Efficiency	+0.0241/-0.0244	1.85
One γ other bkg normalization	+0.0135/-0.0132	1.03
tt normalization	+0.0121/-0.0121	0.92
ISR	+0.0103/-0.0111	0.85
Mis-identified ele shape and norm	+0.0106/-0.0104	0.81
Tune	+0.0001/-0.0087	0.66
Shape of DataDriven templates	+0.0083/-0.0086	0.66
Pile Up	+0.0071/-0.0077	0.58
Renorm/Fact Scale	+0.0073/-0.0076	0.58
hadronic photon normalization	+0.0071/-0.0073	0.56
Jet Energy Resolution 1γ 2017	+0.0062/-0.0073	0.55
lumi 161718	+0.0072/-0.0072	0.55
b tagging 2016	-0.0000/-0.0063	0.48
lumi 2016	+0.0061/-0.0061	0.47
Zero γ other bkg normalization	+0.0047/-0.0047	0.36
FSR	+0.0044/-0.0040	0.33
Jet Energy Resolution 0γ 2016	+0.0041/-0.0042	0.32
Single top	+0.0042/-0.0041	0.32
lumi 2017	+0.0038/-0.0040	0.31
l tagging 2016	+0.0039/-0.0039	0.30
Jet Energy Correction 1γ	+0.0022/-0.0030	0.23
Photon Energy Scale 2016	+0.0029/-0.0012	0.22
Zγ normalization	+0.0025/-0.0028	0.21
Wγ normalization	+0.0022/-0.0026	0.19
Jet Energy Resolution 1γ 2018	+0.0023/-0.0025	0.19
Photon Energy Smearing	+0.0005/-0.0025	0.19
L1 prefire Efficiency	+0.0024/-0.0025	0.19
Color Reconnection	-0.0003/-0.0020	0.15
Photon Energy Scale 2018	+0.0018/-0.0015	0.13
Photon Energy Scale 2017	-0.0014/-0.0017	0.13
l tagging 2018	+0.0016/-0.0016	0.12
l tagging 2017	+0.0015/-0.0015	0.11
b tagging 2017	+0.0013/-0.0014	0.11
lumi 2018	+0.0013/-0.0014	0.11
PDF	+0.0012/-0.0012	0.09
Jet Energy Resolution 0γ 2018	+0.0010/-0.0011	0.08
Jet Energy Resolution 0γ 2017	+0.0009/-0.0011	0.08
Electron Efficiency	+0.0010/-0.0011	0.08
Jet Energy Resolution 1γ 2016	+0.0008/-0.0010	0.08
Muon Efficiency	+0.0009/-0.0010	0.08

CHAPTER 8 RESULT

The strategy to obtain the cross-section measurements is to maximize the profile log-likelihood in a set of signal and control regions. For the inclusive cross section measurement, we perform a simultaneous maximum likelihood fit in the signal region variable M_3 with photon $P_T > 20$ GeV along with three different control region variables: a) charged hadron isolation of photon in $N_j \geq 4, N_b \geq 1$ region, b) M_3 distribution with zero photon control region in $N_j \geq 4, N_b \geq 1$, and c) invariant mass distribution of lepton and photon in $N_j \geq 4, N_b \geq 1$ region. The fits are done separately in e channel and μ channel for each year and eventually done simultaneously for all years combined.

Initially, the fit is performed only on MC data (Asimov and toys), to show the effect of the systematic uncertainties. The Asimov fits are studied for each year separately in e channel , μ channel , $e + \mu$ channel , and for e channel , μ channel , $e + \mu$ channel for all years combined. In this test, we fit the simulation to itself, and check that the injected signal strength is reproduced. A closure test is performed on the inclusive cross section measurement with toy data. The impacts of the nuisances in the inclusive fit for each year and for the full Run II data set are analyzed. Likelihood scans of a parameter of interest (POI, such as a signal strength or a ratio) are scrutinized.

8.1 Checks with Toy Data

To consider various uncertainties of the measurement, hundreds of nuisance parameters (NP) are introduced in the likelihood fit. They may affect the normalization or/and shape of the observable distribution differently. The constraints of a nuisance parameter and its impact on the parameter of interest (POI) are the main concerns for a precise

measurement. For example, the post-fit uncertainty for a systematic source may be much smaller than its initial estimation in a measurement. In other words, the corresponding nuisance parameter turns out to be constrained in the fit. Then the measurement may be biased as the fit doesn't consider the full uncertainty. In the analysis, approximately 65% of the uncertainty is due to systematic sources and the rest from statistics. So we care which systematic sources have large contributions to the uncertainty of the POI. When a systematic is constrained post-fit, then investigation is necessary. We give particular scrutiny to results which contradict established models, previous measurements, or conventional wisdom [119].

8.1.1 Check for $t\bar{t}\gamma$ Signal Strength

To check the sensitivity to the $t\bar{t}\gamma$ signal and to make sure systematics behave as expected the fit is first done using Asimov data. The Asimov data is defined as data with the total number of expected events in the MC. For shape analyses with templates, the Asimov data set will be constructed as a histogram using the same binning defined for the analysis. Figures 68-72 show the effect of the various systematic uncertainties in Asimov data constructed under a signal strength hypothesis of $\mu = 1$ for $e+$jets and $\mu+$jets channels.

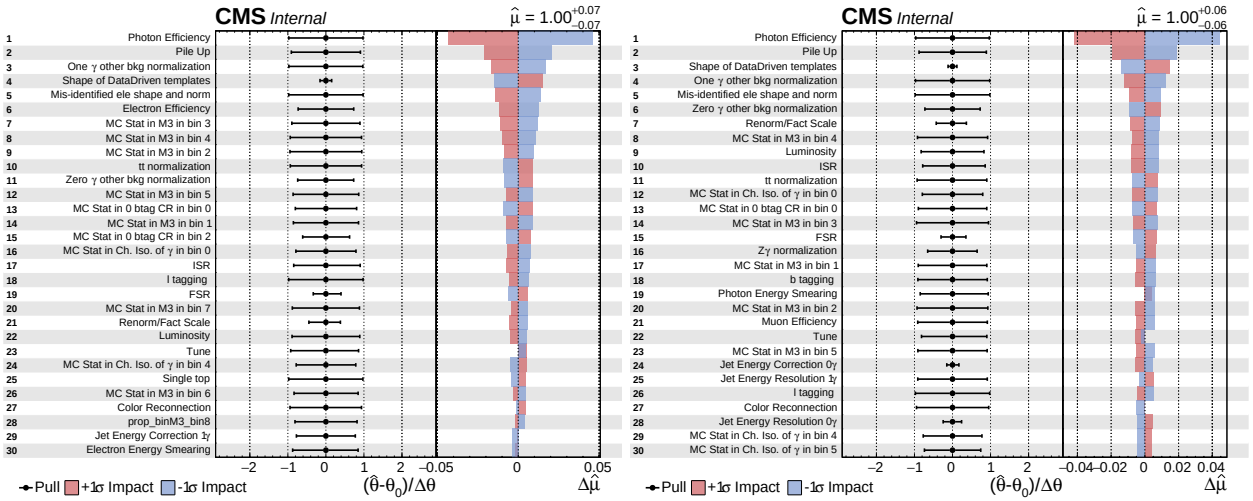


Figure 68: The impact plot that shows postfit pulls and constraints on nuisance parameters for $t\bar{t}\gamma$ fit in $e+$jets (left) and $\mu+$jets (right) separately to Asimov data for 2016.

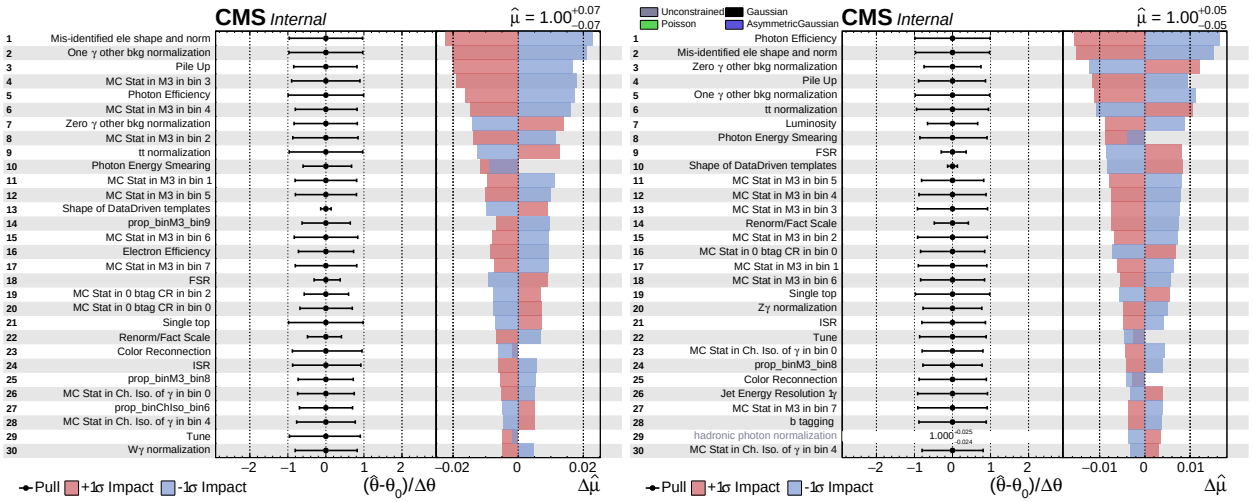


Figure 69: The impact plot that shows postfit pulls and constraints on nuisance parameters for $t\bar{t}\gamma$ fit in $e+$jets (left) and $\mu+$jets (right) separately to Asimov data for 2017.

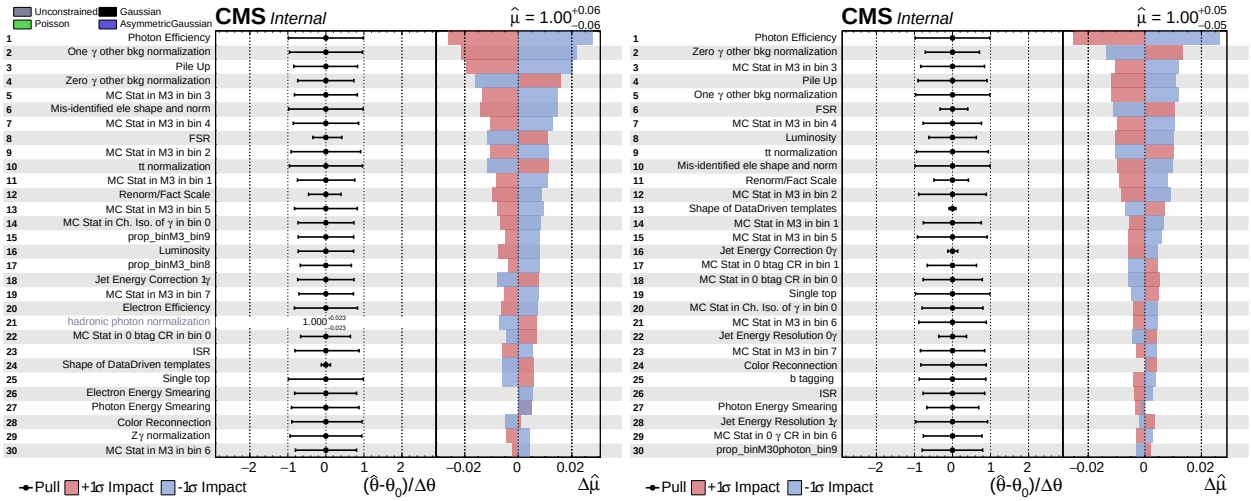


Figure 70: The impact plot that shows postfit pulls and constraints on nuisance parameters for $\mathrm{t\bar{t}}\gamma$ fit in $e+$jets (left) and $\mu+$jets (right) separately to Asimov data for 2016.

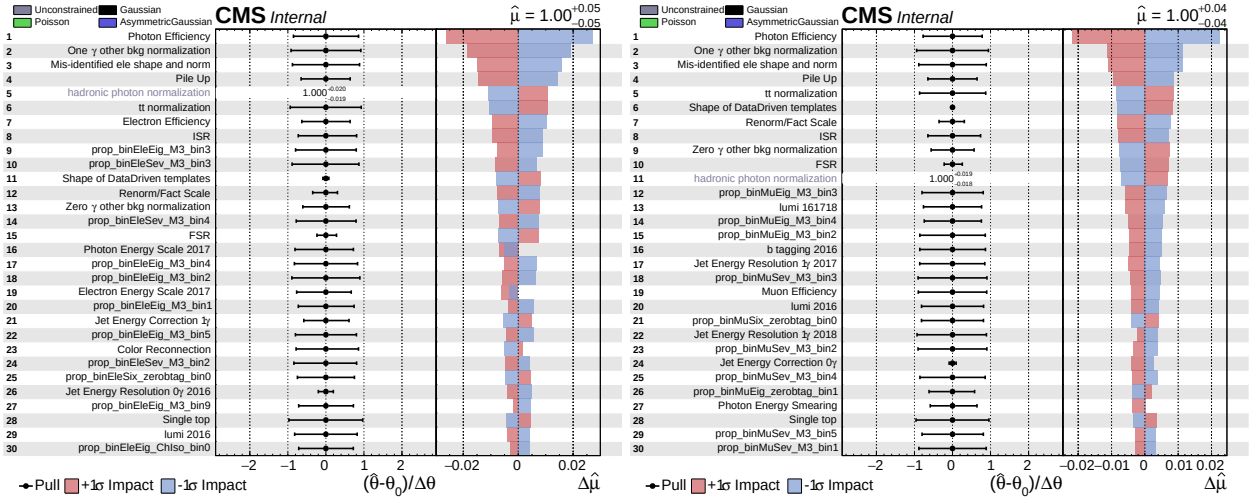


Figure 71: The impact plot that shows postfit pulls and constraints on nuisance parameters for $\mathrm{t\bar{t}}\gamma$ fit in $e+$jets (left) and $\mu+$jets (right) separately to Asimov data for full Run 2.

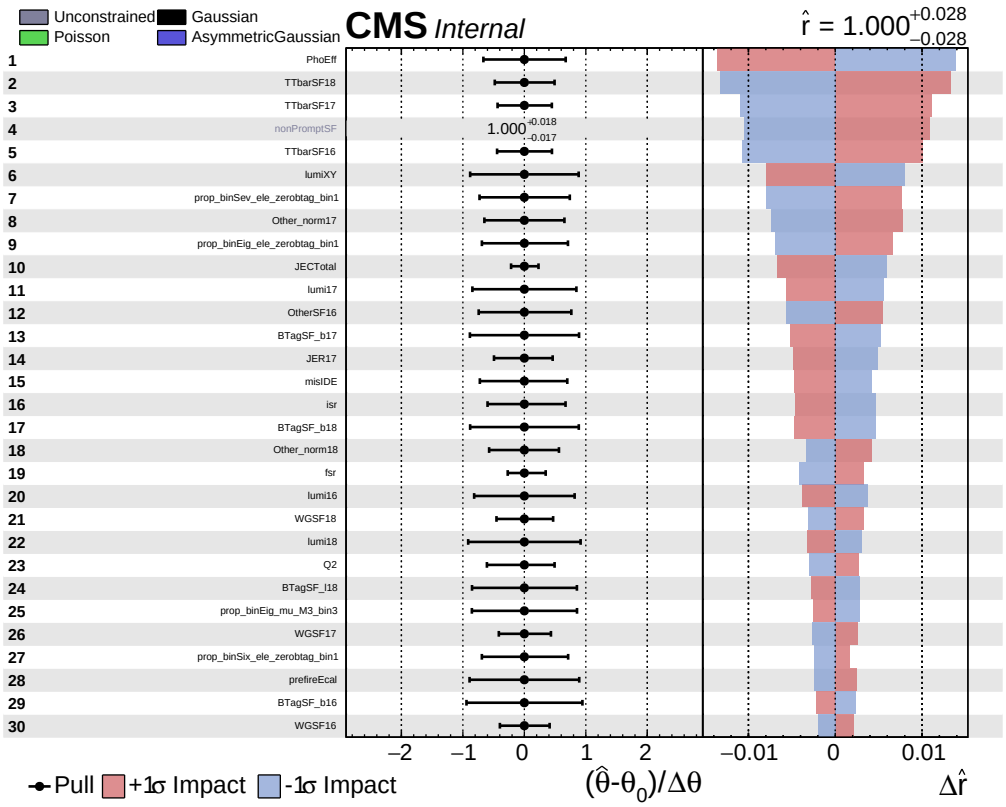


Figure 72: The impact plot that shows postfit pulls and constraints on nuisance parameters for $\mathrm{t\bar{t}\gamma}$ fit to Asimov data for full Run 2.

To check if the fit is biased by any of these parameters the fit was performed assuming a variety of other signal strength and nonprompt photon correction factors values while producing toy data. We set the signal strength parameter equal and run the fit with 500 toy datasets to get the measured values as shown in Figure 73 (top left). The comparison of the expected signal strength (used for the creation of the toy data) and the signal strength parameter measured as a result of the fits is shown in Figures 73-76 along with the comparison of the expected nonprompt photon correction factor and the measured nonprompt photon correction factor extracted from the fit.

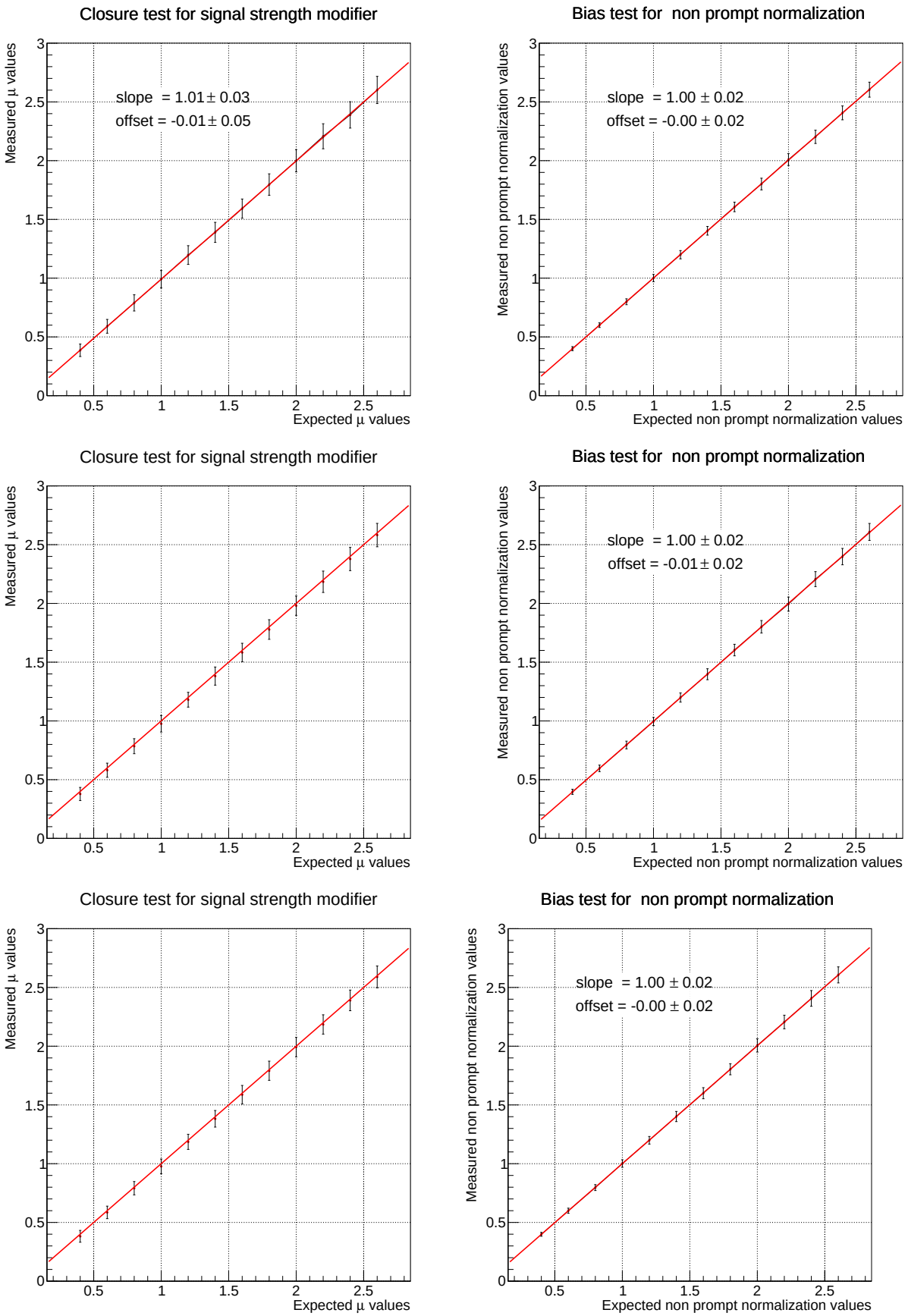


Figure 73: Closure test: $e+$jets ($t\bar{t}\gamma$ (left) and non prompt (right)) for 2016 (top), 2017 (middle), and 2018 (bottom).

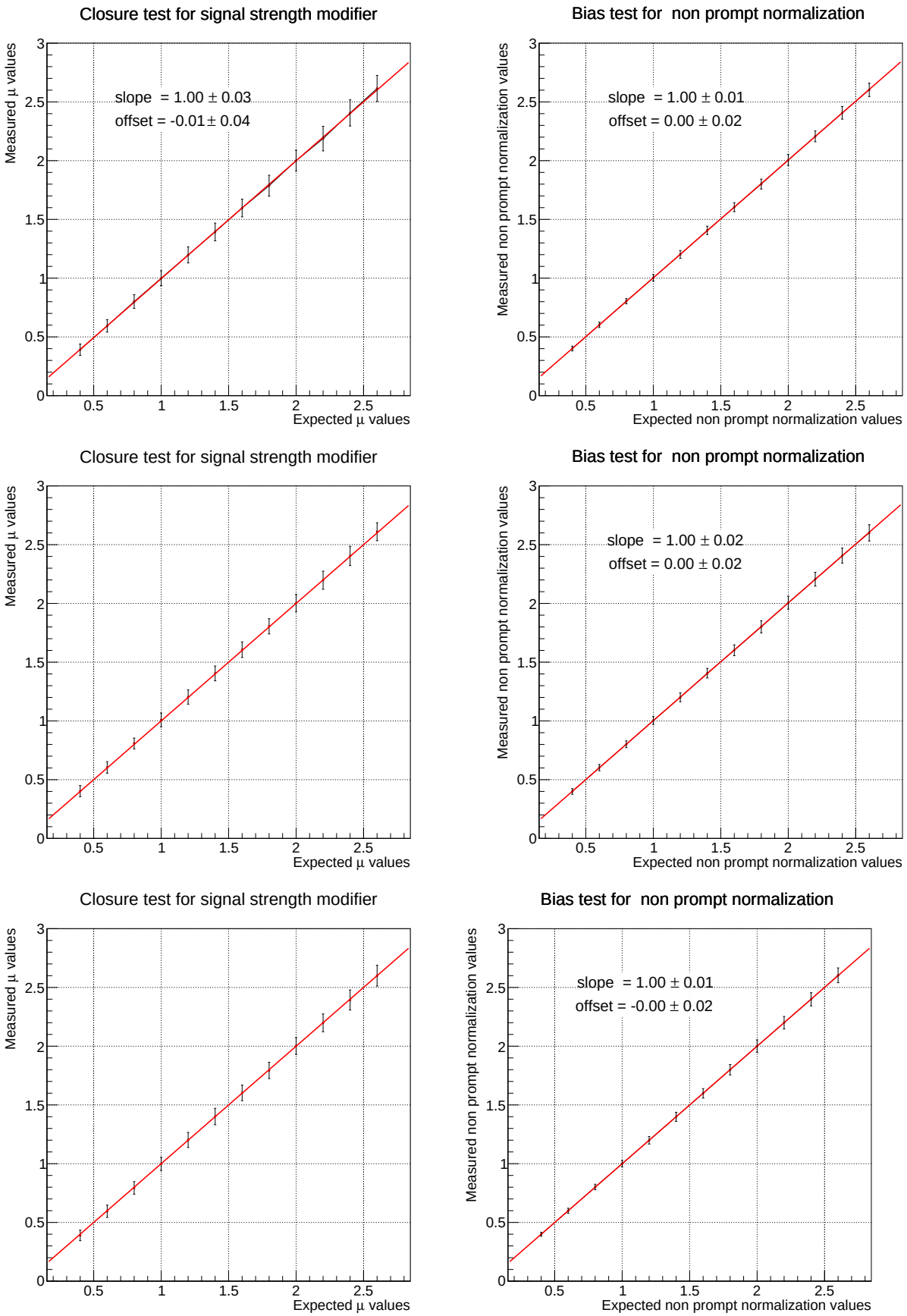


Figure 74: Closure test: μ + jets ($t\bar{t}\gamma$ (left) and non prompt (right)) for 2016 (top), 2017 (middle), and 2018 (bottom).

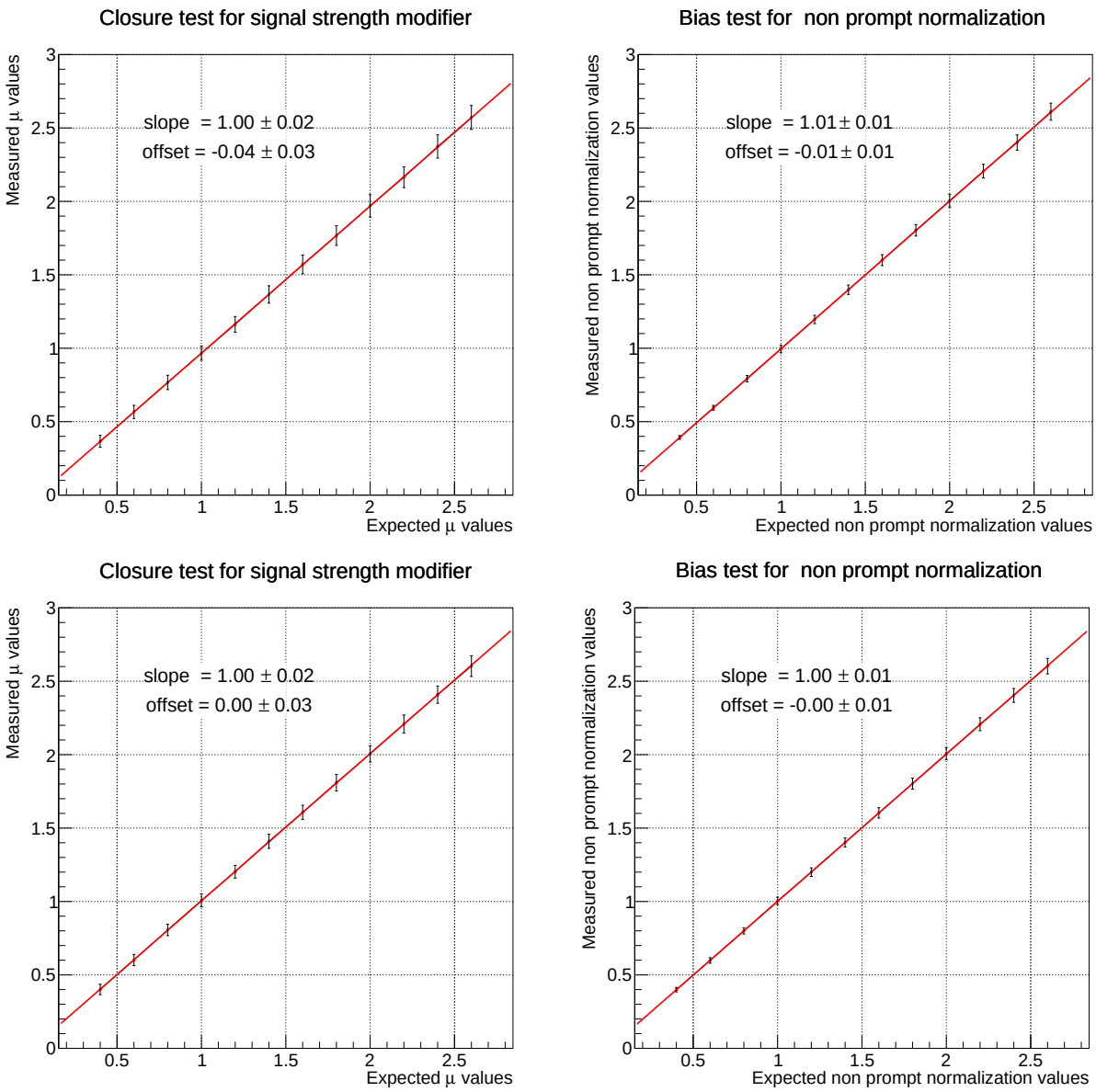


Figure 75: Closure test: $e+$jets and $\mu+$jets ($\mathrm{t\bar{t}}\gamma$ (left) and non prompt (right)) for full Run 2.

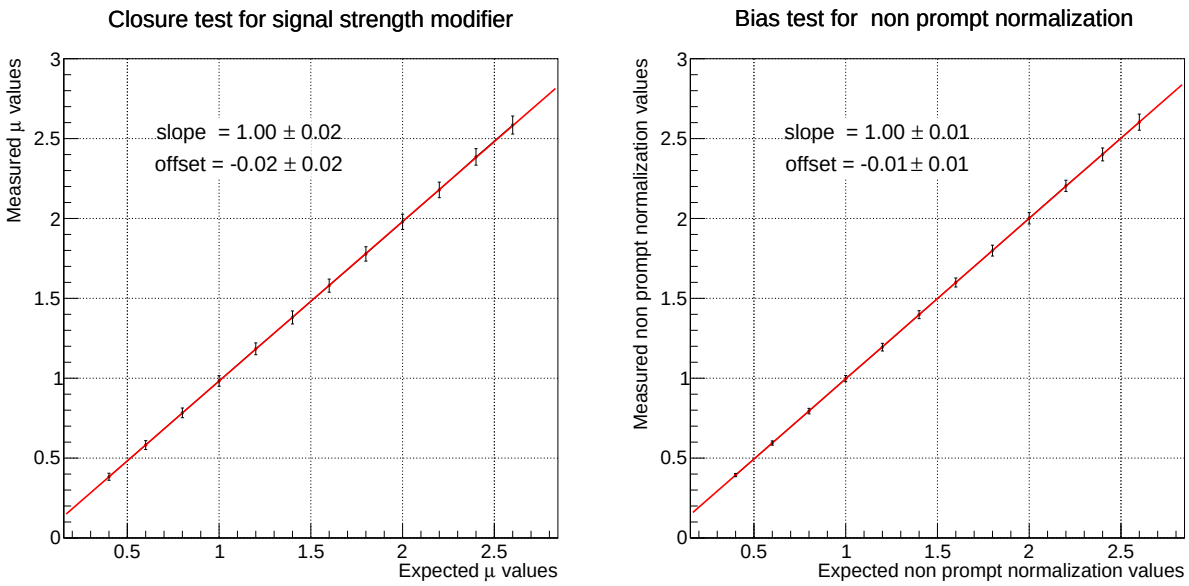


Figure 76: Closure test: ($t\bar{t}\gamma$ (top) and non prompt (bottom)) for full Run 2.

The explicit scan of the profile maximum log-likelihood (minimum of the negative log-likelihood [NNL]) of the parameter of interest is used as it shows any unexpected features in the shape of the likelihood curve. A well behaved fit will have a quadratic shape near the minimum of the NLL. The width of the parabola is the uncertainty on the POI. When all the nuisance parameters are frozen in the minimization, the width of the parabola is due to statistical uncertainty only. When the nuisance parameters are allowed to float, we get the combined systematic and statistical uncertainty.

We used the HiggsCombineTools [106] to measure the best-fit signal strength and its uncertainties as shown in Figures 77-79.

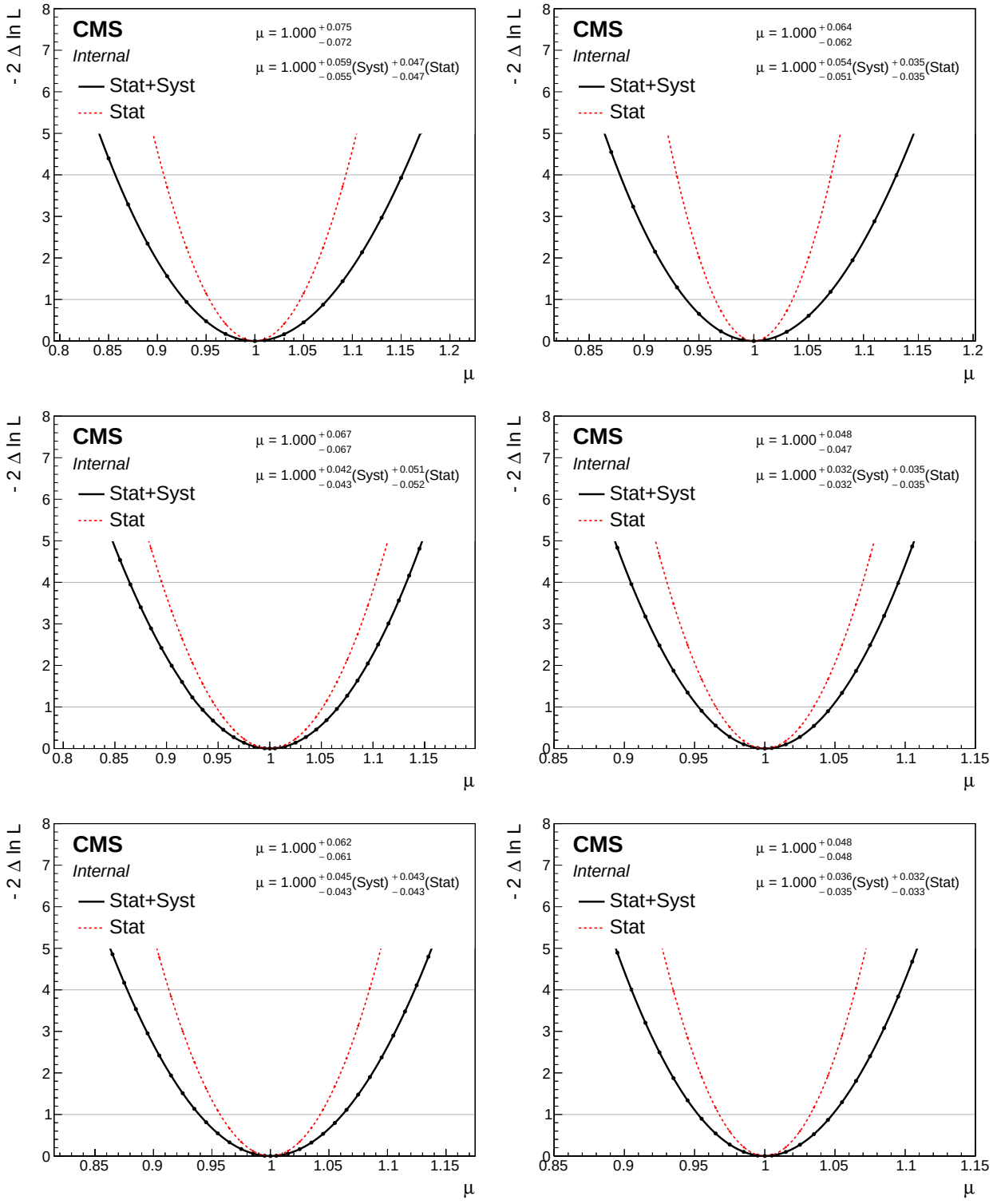


Figure 77: Double negative log likelihood distribution for $t\bar{t}\gamma$ in $e+$jets (left) and $\mu+$jets (right) channels with Asimov fit 2016 (top), 2017 (middle), and 2018 (bottom) data taking periods.

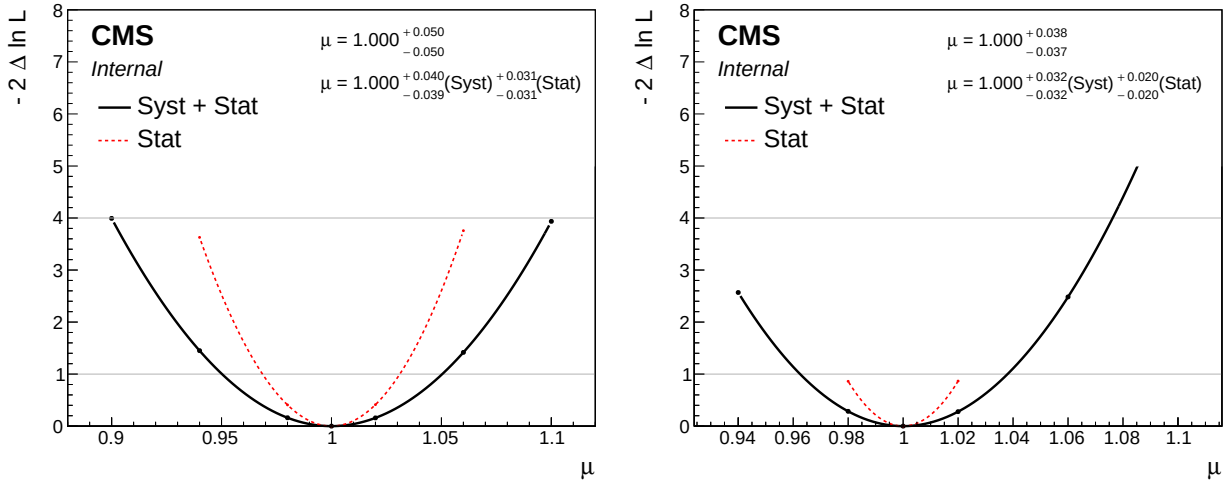


Figure 78: Double negative log likelihood distribution for $\mathrm{t\bar{t}}\gamma$ in $e+$jets (left) and $\mu+$jets (right) channels with Asimov fit for full Run 2 data.

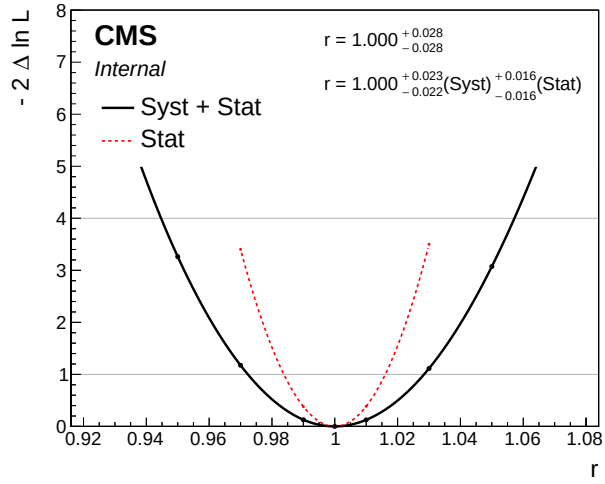


Figure 79: Double negative log likelihood distribution for $\mathrm{t\bar{t}}\gamma$ with Asimov fit for full Run 2 data.

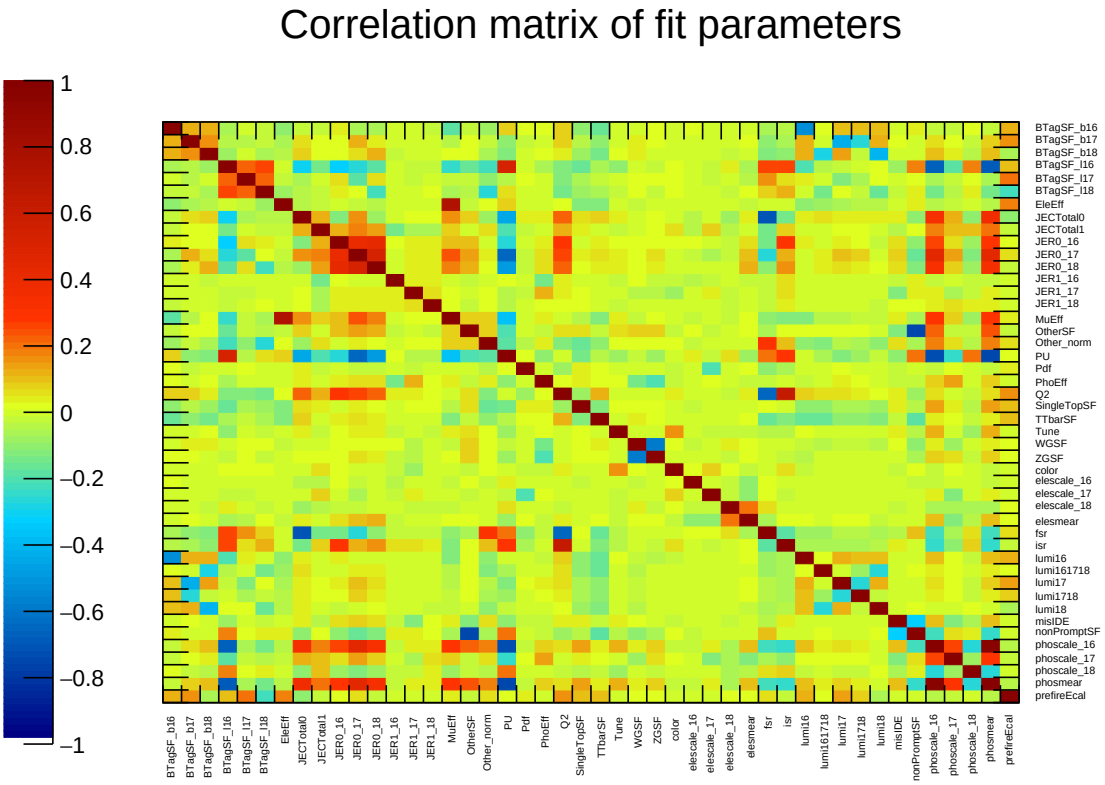


Figure 80: Double negative log likelihood distribution for $t\bar{t}\gamma$ with Asimov fit for full Run 2 data.

The correlation matrices of signal strength with systematic and statistical uncertainties are shown in Figure 140 with toy MC.

8.1.2 Check for $R_\gamma = \frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$

To check the sensitivity to the R_γ and to make sure systematics behave as expected the fit is first done using Asimov data. Figures 81-85 show the effect of the various systematic uncertainties in Asimov data constructed under a signal strength hypothesis of R_γ =0.02 for $e+$jets and $\mu+$jets channels.

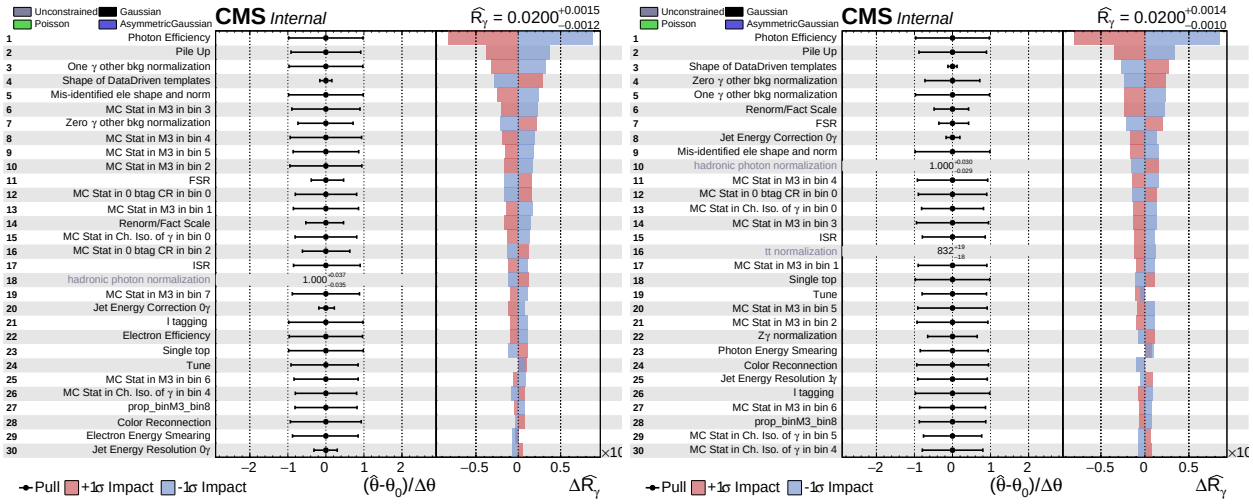


Figure 81: The impact plot that shows postfit pulls and constraints on nuisance parameters for R_γ fit in $e+$jets (left) and $\mu+$jets (right) separately to Asimov data for 2016.

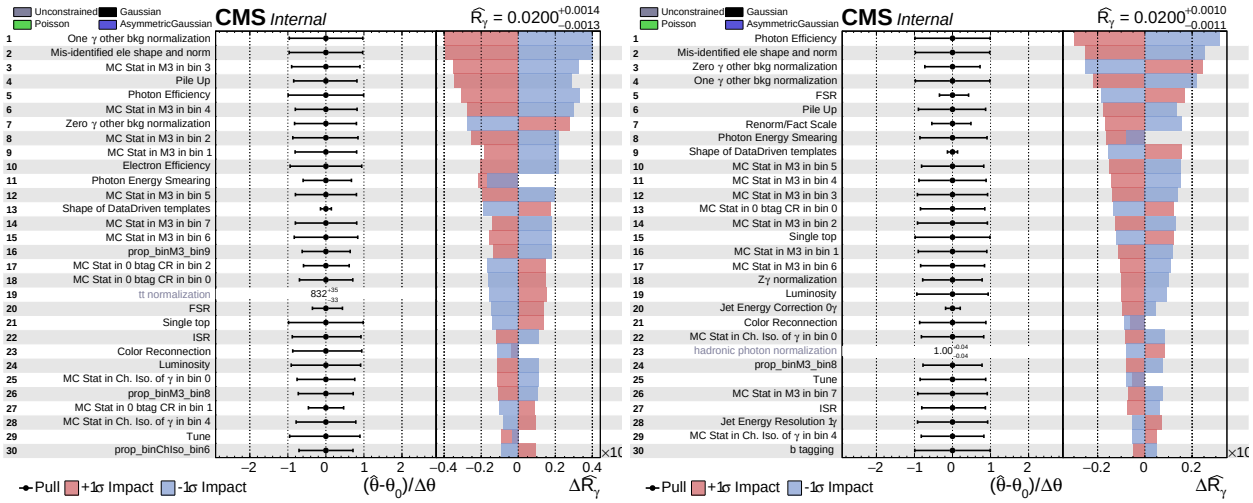


Figure 82: The impact plot that shows postfit pulls and constraints on nuisance parameters for R_γ fit in $e+$jets (left) and $\mu+$jets (right) separately to Asimov data for 2017.

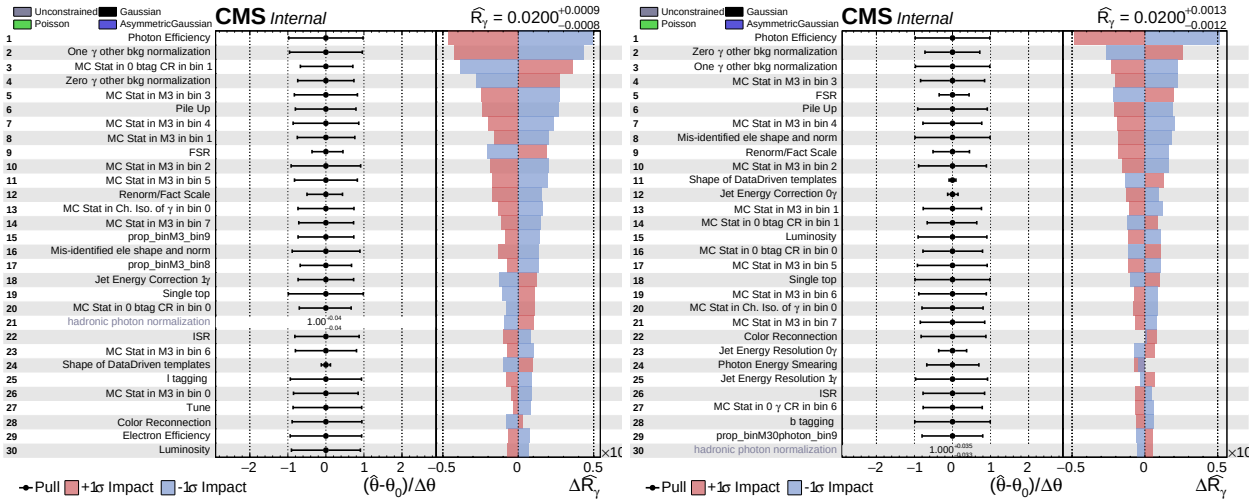


Figure 83: The impact plot that shows postfit pulls and constraints on nuisance parameters for R_γ fit in $e+$jets (left) and $\mu+$jets (right) separately to Asimov data for 2016.

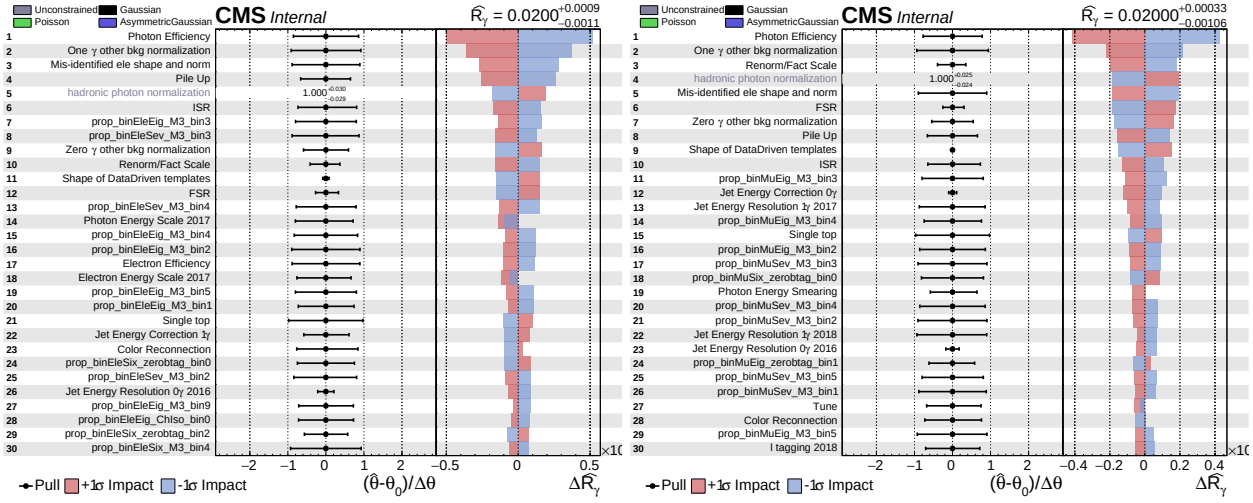


Figure 84: The impact plot that shows postfit pulls and constraints on nuisance parameters for R_γ fit in $e+$jets (left) and $\mu+$jets (right) separately to Asimov data for full Run 2.

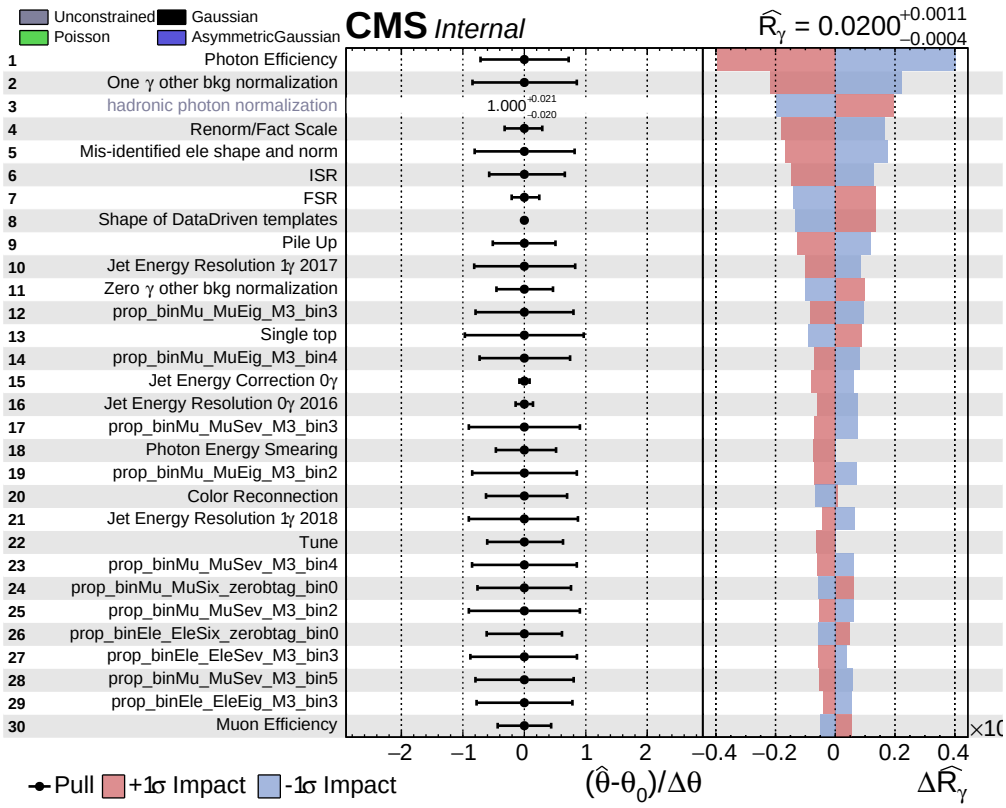


Figure 85: The impact plot that shows postfit pulls and constraints on nuisance parameters for R_γ fit to Asimov data for full Run 2.

To check if the fit is biased by any of these parameters the fit was performed assuming a variety of other R_γ , $\mathrm{t\bar{t}}$ cross section and nonprompt photon correction factors values while producing toy data. The comparison of the expected R_γ (used for the creation of the toy data) and the R_γ measured as a result of the fits is shown in Figures 86-89 along with the comparison of the expected nonprompt photon correction factor and the measured nonprompt photon correction factor as a result of the fit.

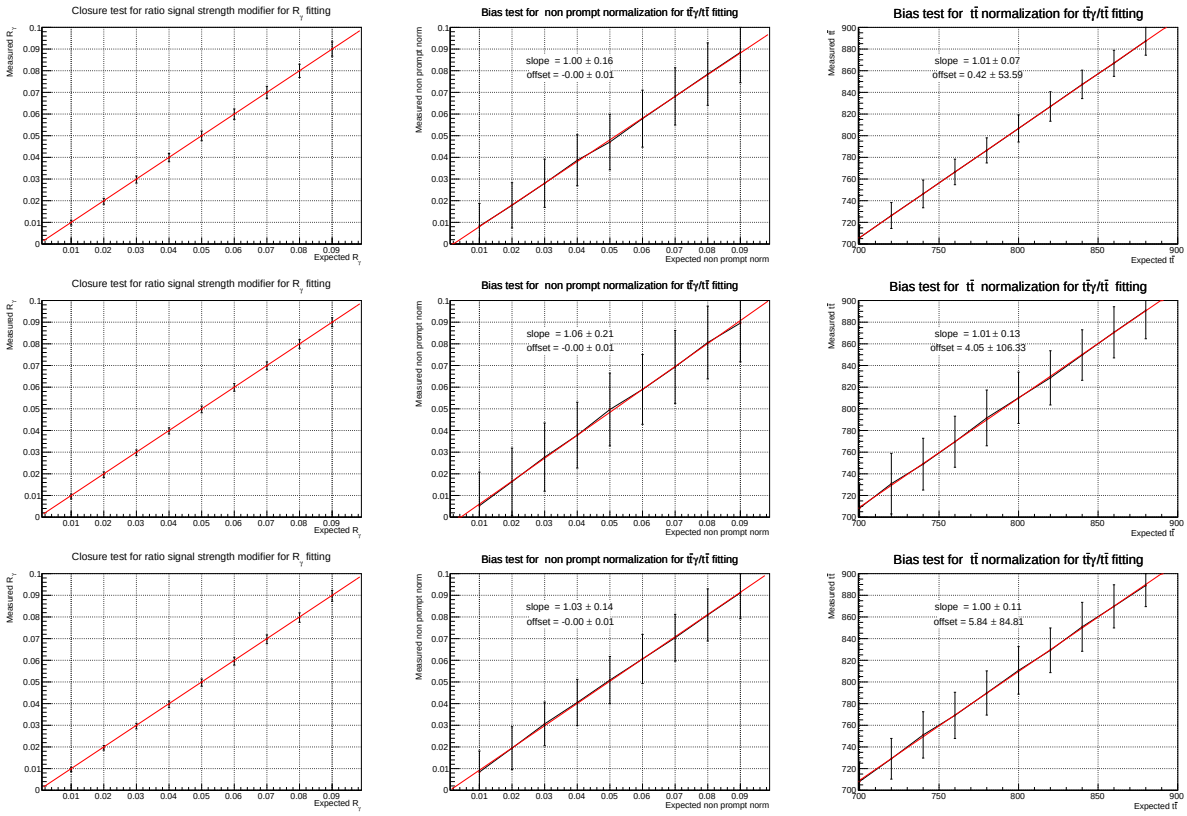


Figure 86: Closure test: $e+$jets (R_γ , non prompt, and $\mathrm{t\bar{t}}$) for 2016 (top), 2017 (middle), and 2018 (bottom).

We used the HiggsCombineTools [106] to measure the best-fit R_γ and and its uncertainties as shown in Figures 77-79.

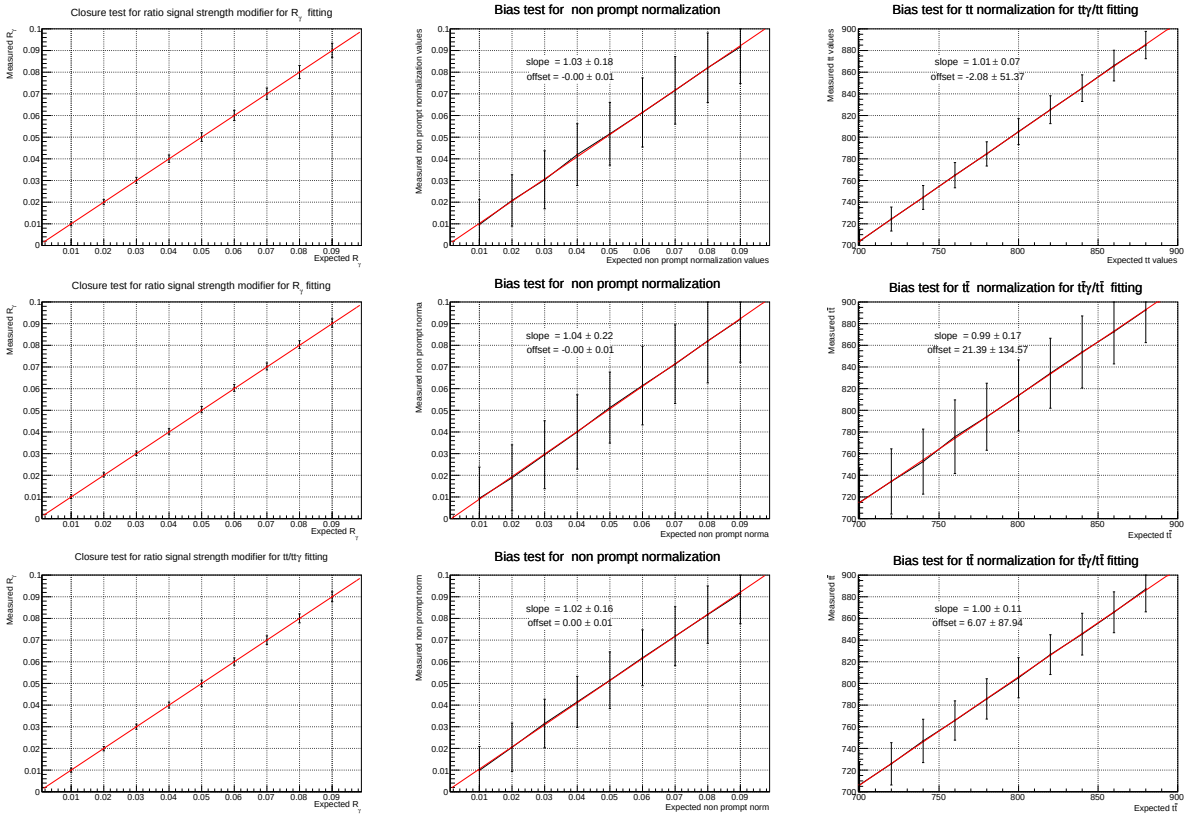


Figure 87: Closure test: $\mu + \text{jets}$ (R_γ (left), non prompt (middle), and $\mathrm{t\bar{t}}$ (right)) for 2016 (top), 2017 (middle), and 2018 (bottom).

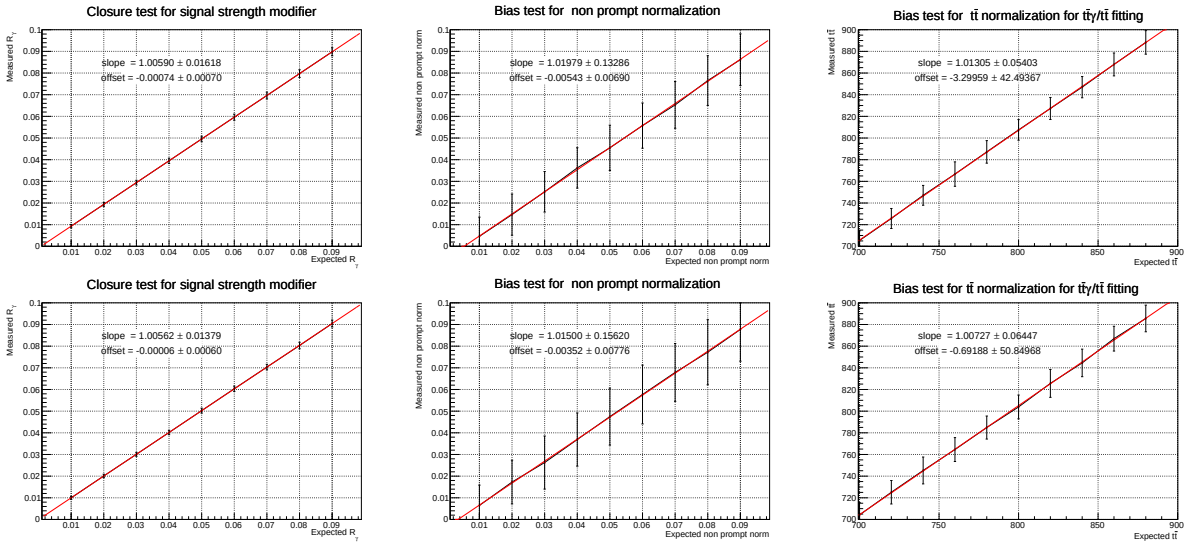


Figure 88: Closure test: $e + \text{jets}$ (top) and $\mu + \text{jets}$ (bottom) (R_γ (left), non prompt (middle), and $\mathrm{t\bar{t}}$ (right)) for full Run 2.

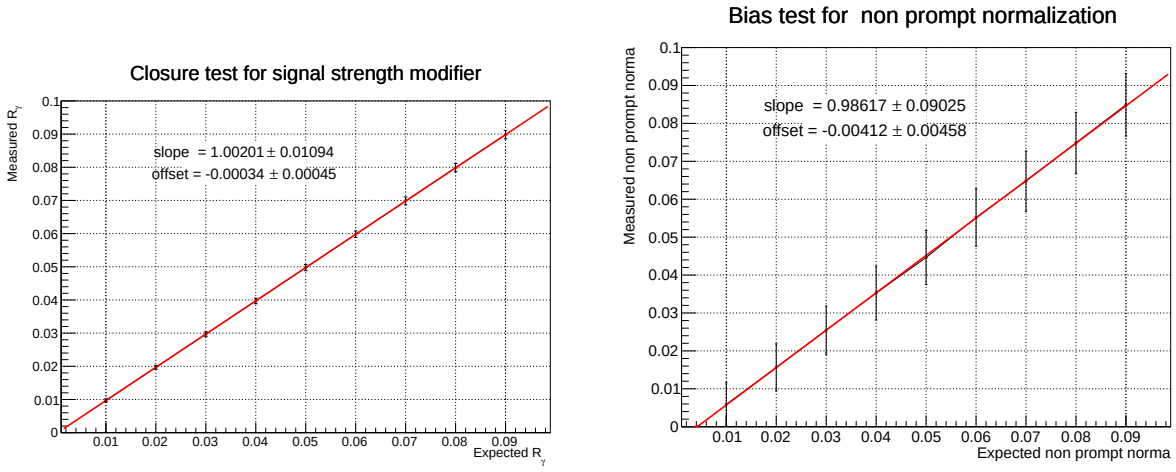


Figure 89: Closure test: (R_γ , non prompt, and $\mathrm{t\bar{t}}$) for full Run 2.

The correlation matrices of R_γ with systematic and statistical uncertainties are shown in Figure 140 with toy MC.

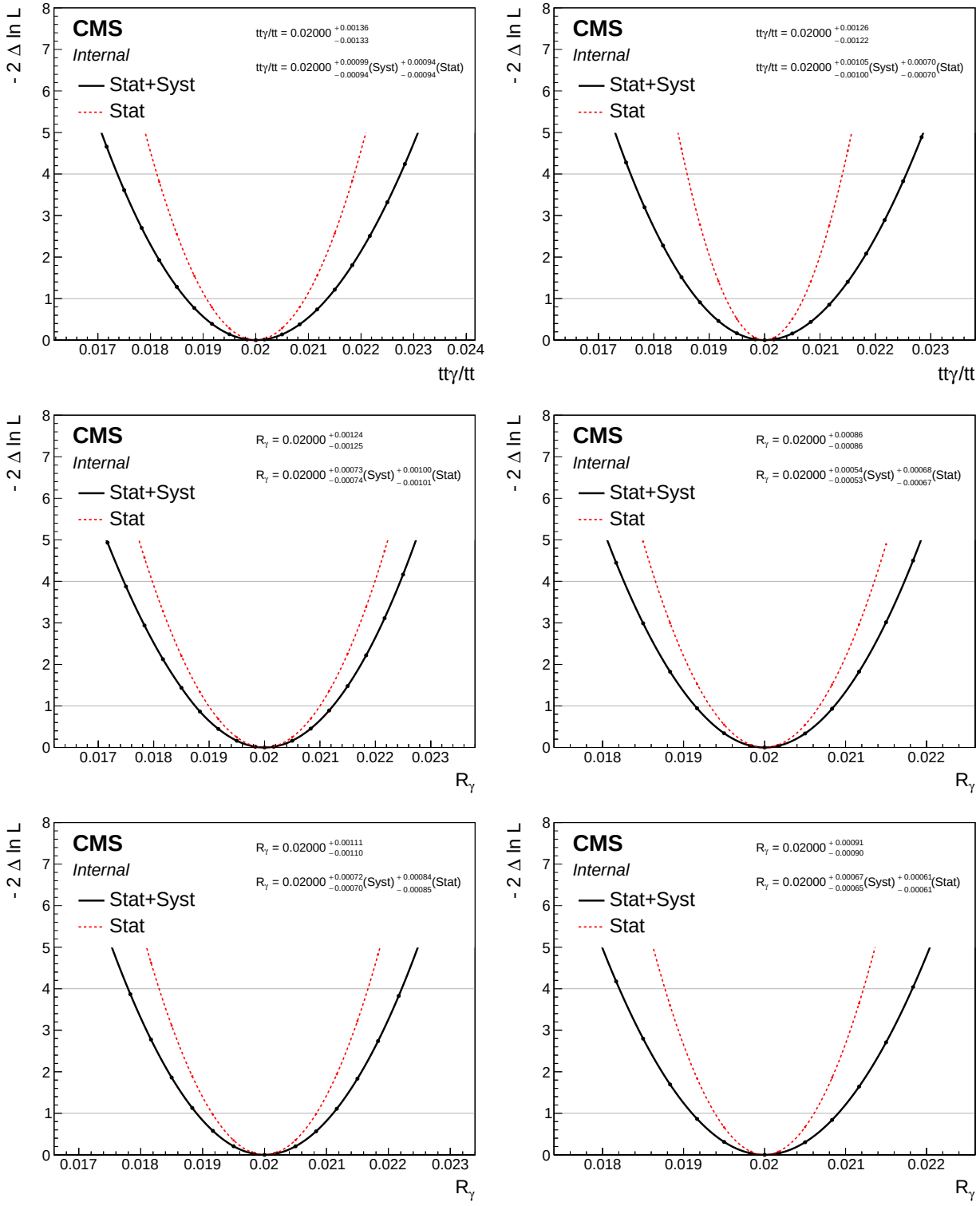


Figure 90: Double negative log likelihood distribution for R_γ in $e+$jets (left) and $\mu+$jets (right) channels with Asimov fit 2016 (top), 2017 (middle), and 2018 (bottom) data taking periods.

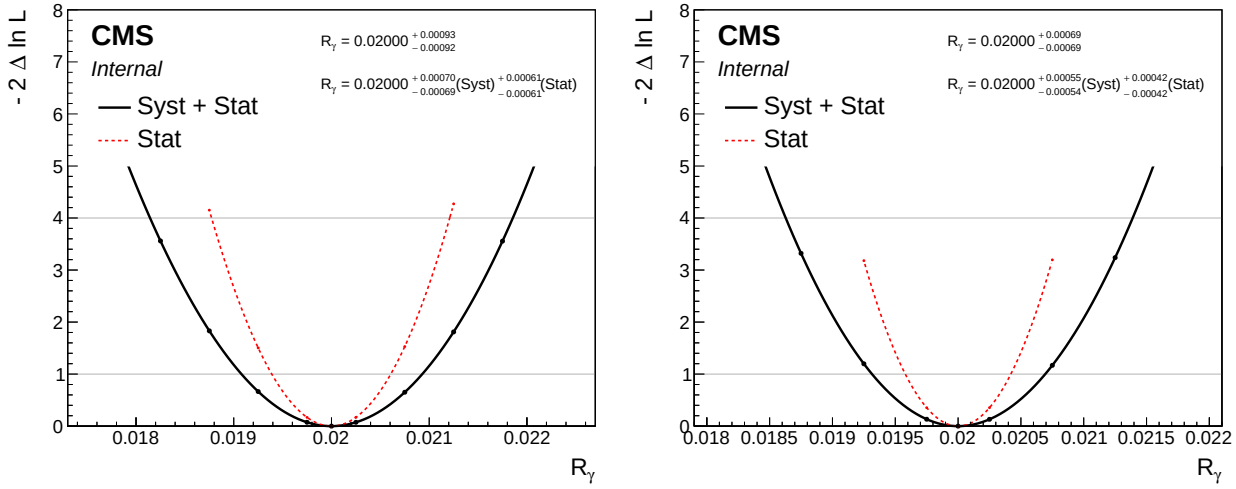


Figure 91: Double negative log likelihood distribution for R_γ in $e+$jets (left) and $\mu+$jets (right) channels with Asimov fit for full Run 2 data.

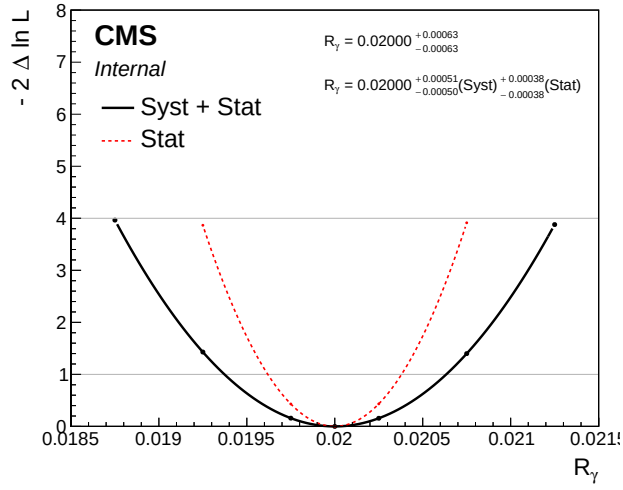


Figure 92: Double negative log likelihood distribution for R_γ with Asimov fit for full Run 2 data.

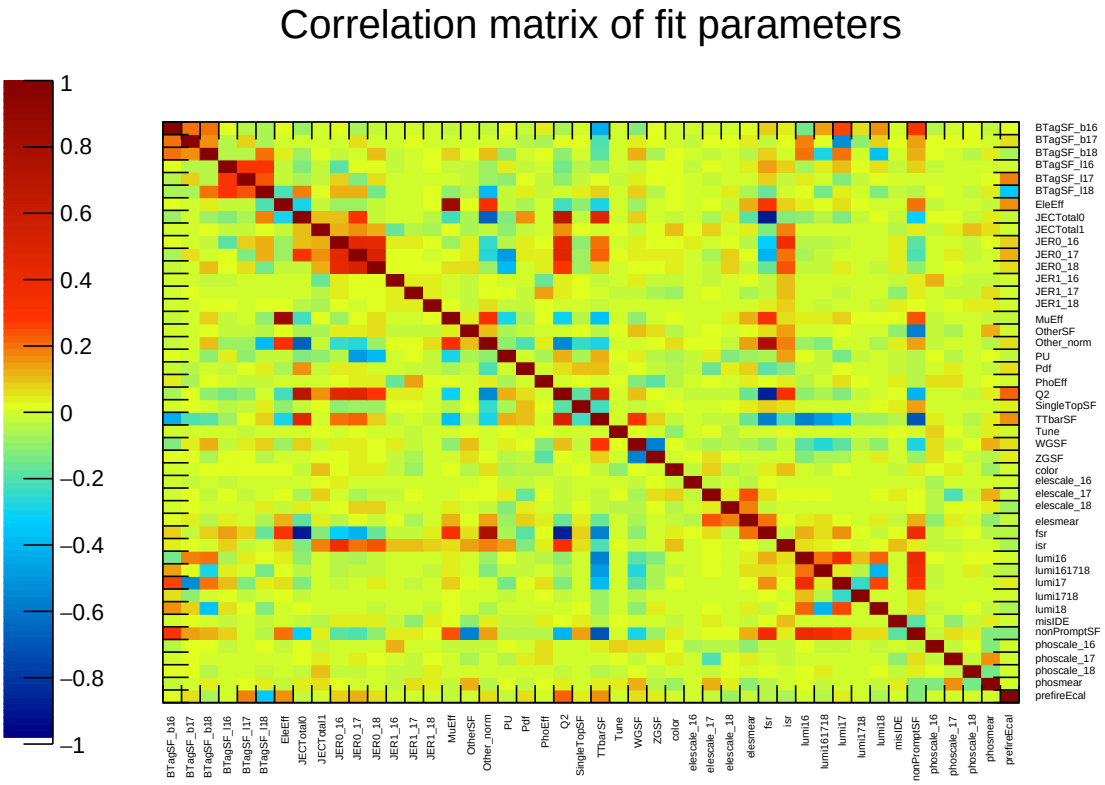


Figure 93: Double negative log likelihood distribution for R_γ with Asimov fit for full Run 2 data.

8.2 Inclusive and Fiducial cross-section

The $\mathrm{t\bar{t}\gamma}$ cross section can be measured inclusively for specific kinematics of the photon. The simulation of the $\mathrm{t\bar{t}\gamma}$ process is performed with $\mathrm{P_T}\,(\gamma) \geq 10$ GeV, $|\eta| < 5.0$. However, the cross section is measured and reported in the kinematic phase space visible within the selection cuts applied in this analysis, $\mathrm{P_T}\,(\gamma) \geq 20$ GeV, $|\eta| < 1.4441$. The inclusive cross section measured is related to the total cross section of the full MC simulation sample by the photon acceptance A^γ through

$$\sigma_{\mathrm{t\bar{t}\gamma}}\left(\mathrm{P_T}\left(\gamma\right) \geq 20\ \mathrm{GeV}, |\eta| < 1.4441\right) = \sigma_{\mathrm{t\bar{t}\gamma}}\left(\mathrm{P_T}\left(\gamma\right) \geq 10\ \mathrm{GeV}, |\eta| < 5.0\right) \times A^{\gamma} \tag{8.1}$$

where, $\sigma_{\mathrm{t\bar{t}\gamma}}$ is the cross section for the fiducial requirements specified.

The photon acceptance can be measured from simulation, using generator level photons. Photons are required to be prompt, final state particles, with $\mathrm{P_T} > 20$ GeV and $|\eta| < 1.44$. The acceptance is defined as $A^{\gamma} = N_{\gamma}^{gen}/N_{total}^{gen}$, where N_{total}^{gen} is the total number of generated $\mathrm{t\bar{t}\gamma}$ events in all top decay channels, and N_{γ}^{gen} is the total number of events with the photon in fiducial the phase space. The photon acceptance values are found to be $A^{\gamma} = (38.56 \pm 0.02)\%$, $(43.05 \pm 0.01)\%$, and $(38.54 \pm 0.20)\%$ for 2016, 2017, 2018 respectively. The weighted average photon acceptance for full Run 2 is $(40.05 \pm 0.02)\%$.

The fiducial cross section is measured in the visible phase space of the final particles in $\mathrm{t\bar{t}\gamma}$ event where the top quark pair decays leptonically. The fiducial phase space is defined by the presence of exactly one lepton (either an electron or muon), at least four jets of which one originates from a b quark decay, and a photon. Electrons are required to be prompt final state particles with $\mathrm{P_T} > 35$ GeV and $|\eta| < 2.4$. Muons are required to be prompt final state particles with $\mathrm{P_T} > 30$ GeV and $|\eta| < 2.4$. Jets are clustered from generator particles with the anti-k_t algorithm of cone size 0.4, and must have $\mathrm{P_T} > 30$ GeV with $|\eta| < 2.4$. The cross section in this fiducial phase space is related to the inclusive $\mathrm{t\bar{t}\gamma}$ cross section via the $\mathrm{t\bar{t}}$ acceptance.

In Table 32, we define the particle level objects that are used to define the fiducial phase space. The fiducial phase space is then defined as ${N_{\gamma}}^{\text{fid.}} = 1$, ${N_{\ell}}^{\text{fid.}} = 1$,

Table 32: Definition of objects used to define the fiducial phase space.

Selection criteria	gen-Photon
P_T (GeV)	$\geq$20
$\|\eta\|$	$\leq$1.4442
|pdgID|	22
status	1

$N_{jet}{}^{\text{fid.}} \geq 4$ and $N_{b-jet}{}^{\text{fid.}} \geq 1$. The total $\mathrm{t\bar{t}}\gamma$ acceptance can be measured from simulation where photons are required to be prompt, final state particles, with $P_T(\gamma) > 20$ GeV and $|\eta|(\gamma) < 1.44$ along with a $\mathrm{t\bar{t}}$ pair decaying semi-leptonically. The acceptance is defined as $A^{\mathrm{t\bar{t}}\gamma} = N^{selected}_{\mathrm{t\bar{t}}\gamma}/N^{gen}_{total}$, where the N^{gen}_{total} is the total number of generated $\mathrm{t\bar{t}}\gamma$ events in all top decay channels, and $N^{selected}_{\mathrm{t\bar{t}}\gamma}$ is the total number of events in the signal phase space. The acceptance values are found to be $A^{\mathrm{t\bar{t}}\gamma} =$ (6.28 $\pm$ 0.01))%, (6.05 $\pm$ 0.01))%, and (7.98 $\pm$ 0.01))% for 2016, 2017, 2018 respectively. The weighted average acceptance for full Run 2 is (6.86 $\pm$ 0.01))%.

The fiducial cross section in terms of $\mathrm{t\bar{t}}\gamma$ acceptance is defined as

$$\sigma_{\mathrm{t\bar{t}}\gamma}(P_T(\gamma) \geq 20 \text{ GeV}, |\eta| < 1.4441) = \sigma_{\mathrm{t\bar{t}}\gamma}(P_T(\gamma) \geq 10 \text{ GeV}, |\eta| < 5.0) \times A^{\mathrm{t\bar{t}}\gamma}. \quad (8.2)$$

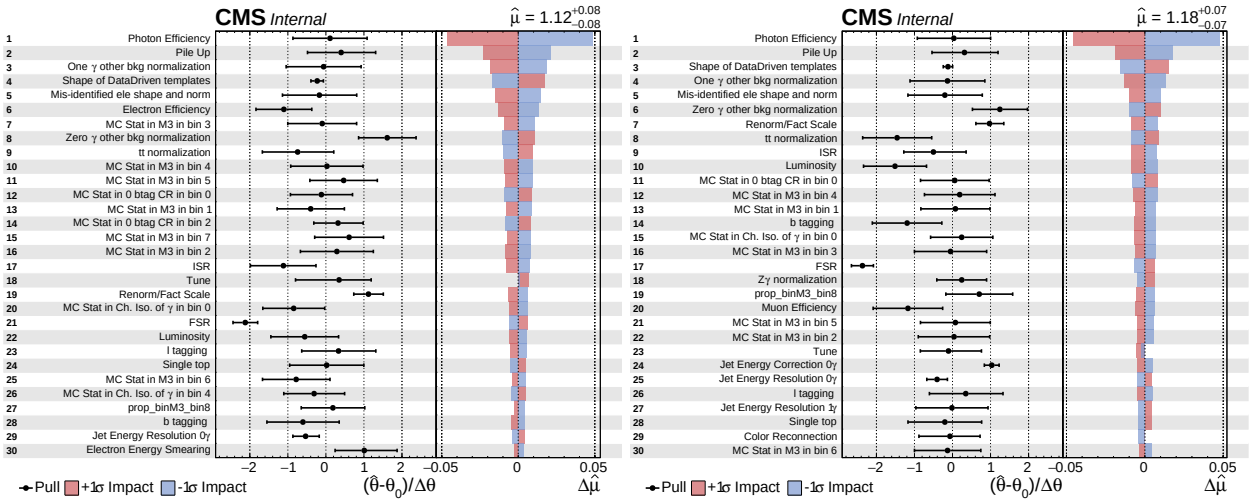


Figure 94: The impact plot that shows postfit pulls and constraints on nuisance parameters for $\mathrm{t\bar{t}}\gamma$ fit in $e+$jets (left) and $\mu+$jets (right) separately to observed data for 2016

8.3 Fit to Data

8.3.1 $\mathrm{t\bar{t}}\gamma$ Cross Section from Signal Strength

First we do a fit to extract the signal strength parameter $\mu = \frac{\sigma^{\mathrm{meas}}_{\mathrm{t\bar{t}}\gamma}}{\sigma^{\mathrm{NLO}}_{\mathrm{t\bar{t}}\gamma}}$. Figures 94 - 112 show the pulls on nuisance parameters and their impacts on μ, and maximum likelihood scan for 2016, 2017, 2018, and for all years combined. The extraction of the signal strength along with other important postfit scale factors is shown in Tables 41-43 for 2016, 2017, and 2018 data in $e+$jets and $\mu+$jets channels.

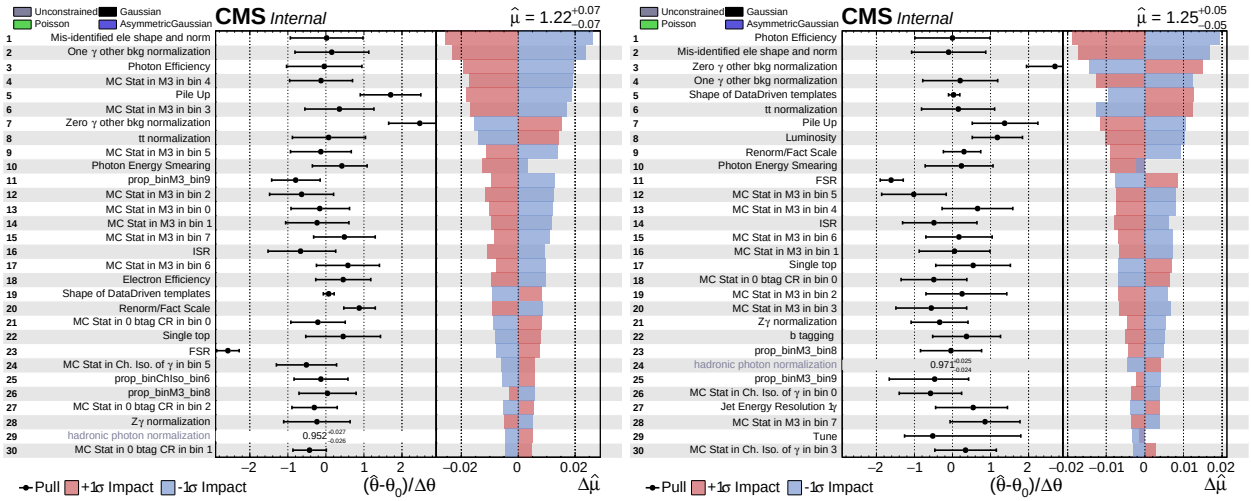


Figure 95: The impact plot that shows postfit pulls and constraints on nuisance parameters for $\mathrm{t\bar{t}}\gamma$ fit in $e+$jets (left) and $\mu+$jets (right) separately to observed data for 2017

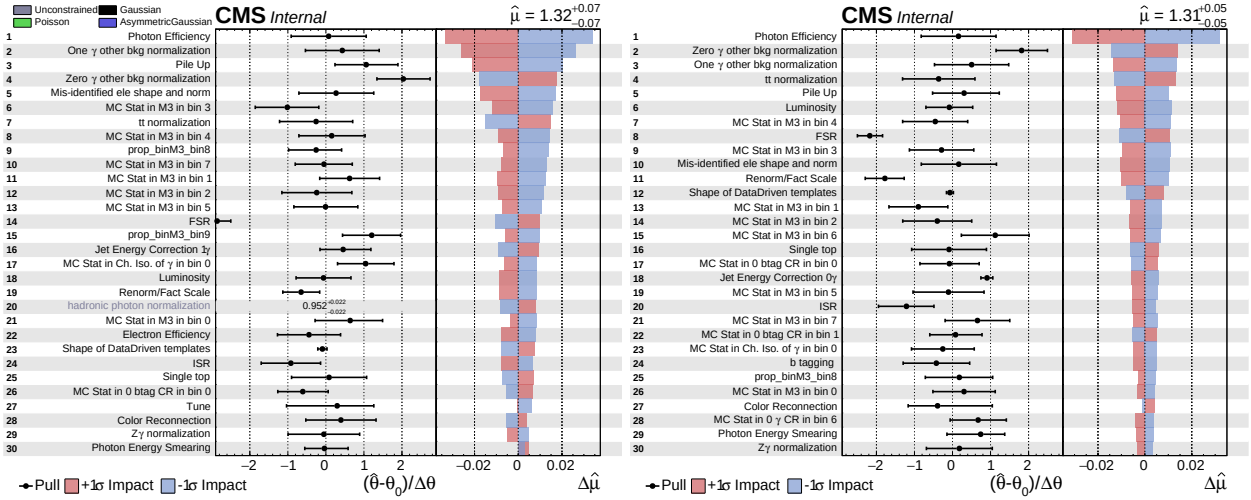


Figure 96: The impact plot that shows postfit pulls and constraints on nuisance parameters for $\mathrm{t\bar{t}}\gamma$ fit in $e+$jets (left) and $\mu+$jets (right) separately to observed data for 2016

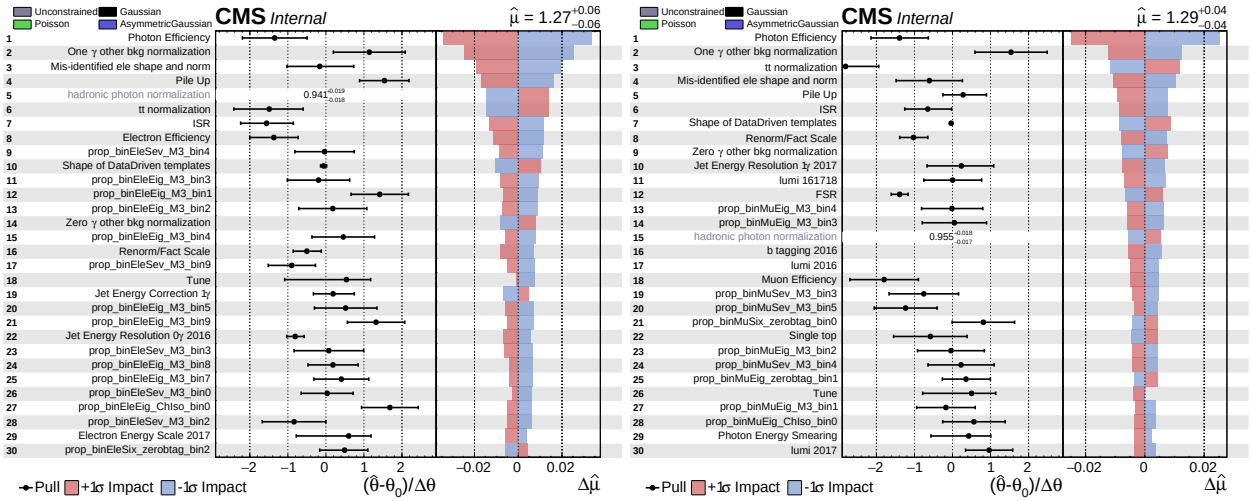


Figure 97: The impact plot that shows postfit pulls and constraints on nuisance parameters for $\mathrm{t\bar{t}}\gamma$ fit in $e+$jets (left) and $\mu+$jets (right) separately to observed data for full Run 2

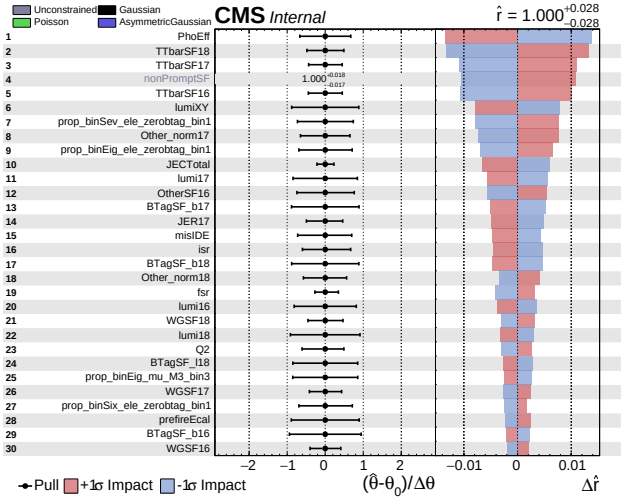


Figure 98: The impact plot that shows postfit pulls and constraints on nuisance parameters for $\mathrm{t\bar{t}}\gamma$ fit to observed data for full Run 2

The detail values of extraction of the scale factors can be found at Appendix C. The full Run 2 pre-fit and post-fit Data/MC agreement plots in the signal region for M_3 and ChIso variables are shown in Figure 103. The pre-fit and post-fit Data/MC agreement plots in in zero btag (top) and zero photon (bottom) control regions are shown in Figure 104.

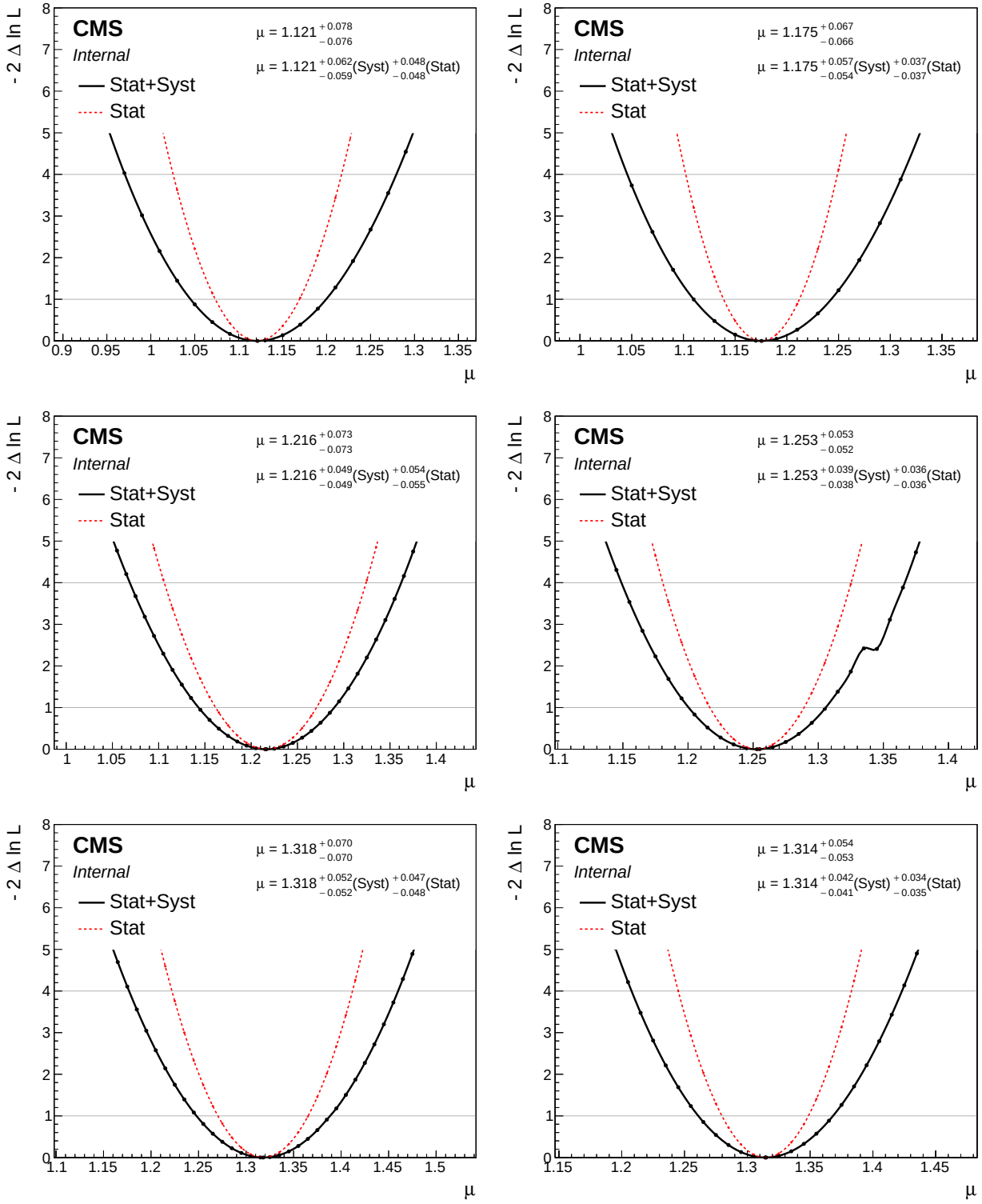


Figure 99: Double negative log likelihood distribution for $t\bar{t}\gamma$ in $e+$jets (left) and $\mu+$jets (right) channels with observed data fit 2016 (top), 2017 (middle), and 2018 (bottom) data taking periods.

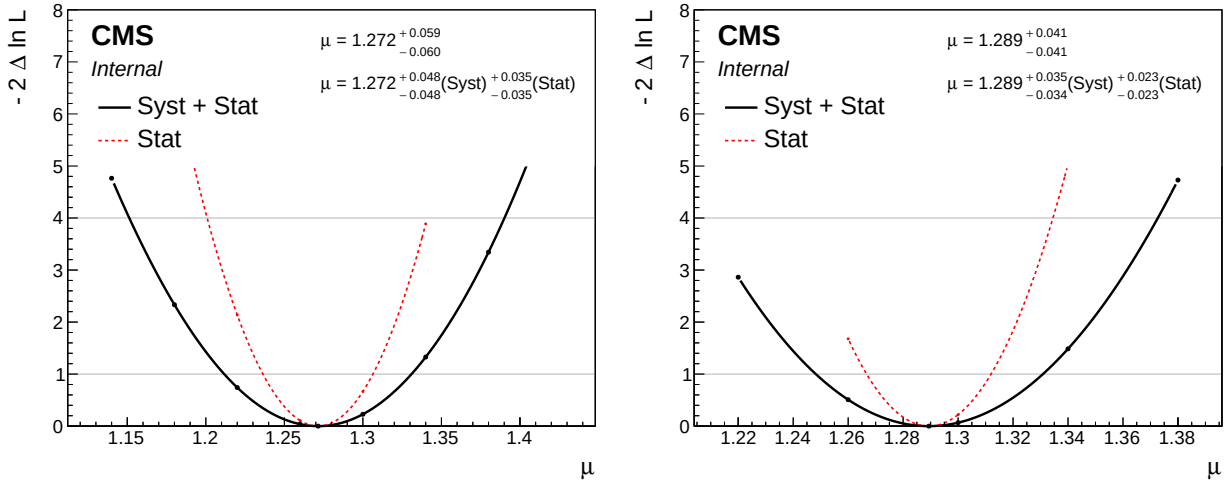


Figure 100: Double negative log likelihood distribution for $t\bar{t}\gamma$ in $e+$jets (left) and $\mu+$jets (right) channels with observed data fit for full Run 2 data.

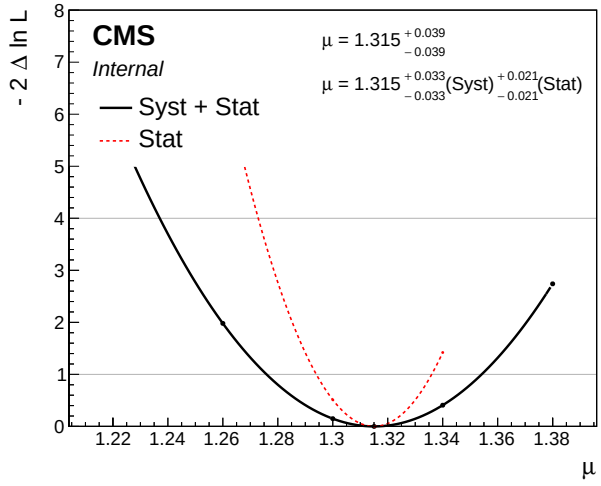


Figure 101: Double negative log likelihood distribution for $t\bar{t}\gamma$ with observed data fit for full Run 2 data.

Table 33: The signal strength (μ) extraction along with other important scale factors for $e+$jets and $\mu+$jets channels for 2016 data.

Year	channel	$t\bar{t}\gamma$ SF	nonprompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2016	$e+$jets	$1.12^{+0.08}_{-0.08}$	$0.89^{+0.08}_{-0.02}$	$0.99^{+0.01}_{-0.01}$	$1.02^{+0.09}_{-0.09}$	$1.05^{+0.28}_{-0.28}$	$0.99^{+0.10}_{-0.10}$	$1.16^{+0.08}_{-0.08}$
2016	$\mu+$jets	$1.18^{+0.07}_{-0.07}$	$0.87^{+0.07}_{-0.02}$	$0.99^{+0.01}_{-0.01}$	$1.03^{+0.08}_{-0.08}$	$1.10^{+0.26}_{-0.26}$	$0.99^{+0.10}_{-0.10}$	$1.12^{+0.07}_{-0.07}$

Table 34: The signal strength (μ) extraction along with other important scale factors for $e+$jets and $\mu+$jets channels for 2017 data.

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2017	$e+$jets	$1.22^{+0.07}_{-0.07}$	$0.95^{+0.07}_{-0.03}$	$1.00^{+0.01}_{-0.01}$	$0.95^{+0.08}_{-0.08}$	$0.91^{+0.35}_{-0.35}$	$1.02^{+0.10}_{-0.10}$	$1.25^{+0.08}_{-0.08}$
2017	$\mu+$jets	$1.25^{+0.05}_{-0.05}$	$0.97^{+0.05}_{-0.02}$	$1.00^{+0.01}_{-0.01}$	$0.93^{+0.07}_{-0.07}$	$0.86^{+0.30}_{-0.30}$	$1.02^{+0.10}_{-0.10}$	$1.27^{+0.08}_{-0.08}$

Table 35: The signal strength (μ) extraction along with other important scale factors for $e+$jets and $\mu+$jets channels for 2018 data.

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2018	$e+$jets	$1.32^{+0.07}_{-0.07}$	$0.95^{+0.07}_{-0.02}$	$1.00^{+0.01}_{-0.01}$	$0.95^{+0.08}_{-0.08}$	$0.98^{+0.37}_{-0.37}$	$1.04^{+0.10}_{-0.10}$	$1.20^{+0.07}_{-0.07}$
2018	$\mu+$jets	$1.31^{+0.05}_{-0.05}$	$0.97^{+0.05}_{-0.02}$	$1.00^{+0.01}_{-0.01}$	$0.98^{+0.07}_{-0.07}$	$1.07^{+0.34}_{-0.34}$	$1.05^{+0.10}_{-0.10}$	$1.18^{+0.07}_{-0.07}$

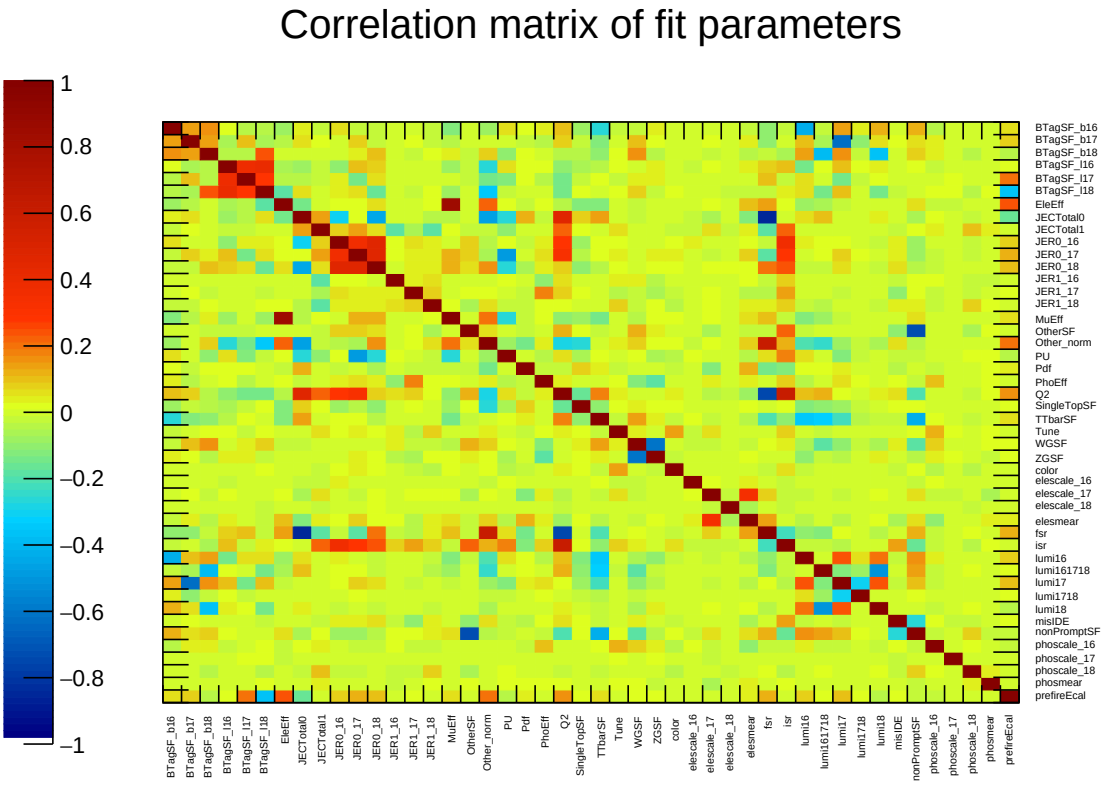


Figure 102: Double negative log likelihood distribution for $t\bar{t}\gamma$ with data fit for full Run 2 data.

8.3.2 Fit to Ratio ($R_\gamma = \frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$)

The ratio of the $t\bar{t}\gamma$ cross section to the $t\bar{t}$ cross section is used to study anomalous couplings of the top quark in phenomenological studies. An advantage of measuring a ratio as compared to measuring the $t\bar{t}\gamma$ cross section directly is that some of the common sources of systematic uncertainties cancel out.

Fit Function First, the constraint on the normalization of the $t\bar{t}$ cross section is removed, allowing the normalization of the $t\bar{t}$ cross section to float. In order to fit directly to the cross section, rather than to a signal strength parameter, the templates for the $t\bar{t}\gamma$ production and $t\bar{t}$ production are re-normalized by the inverse of their theoretical cross sections. This

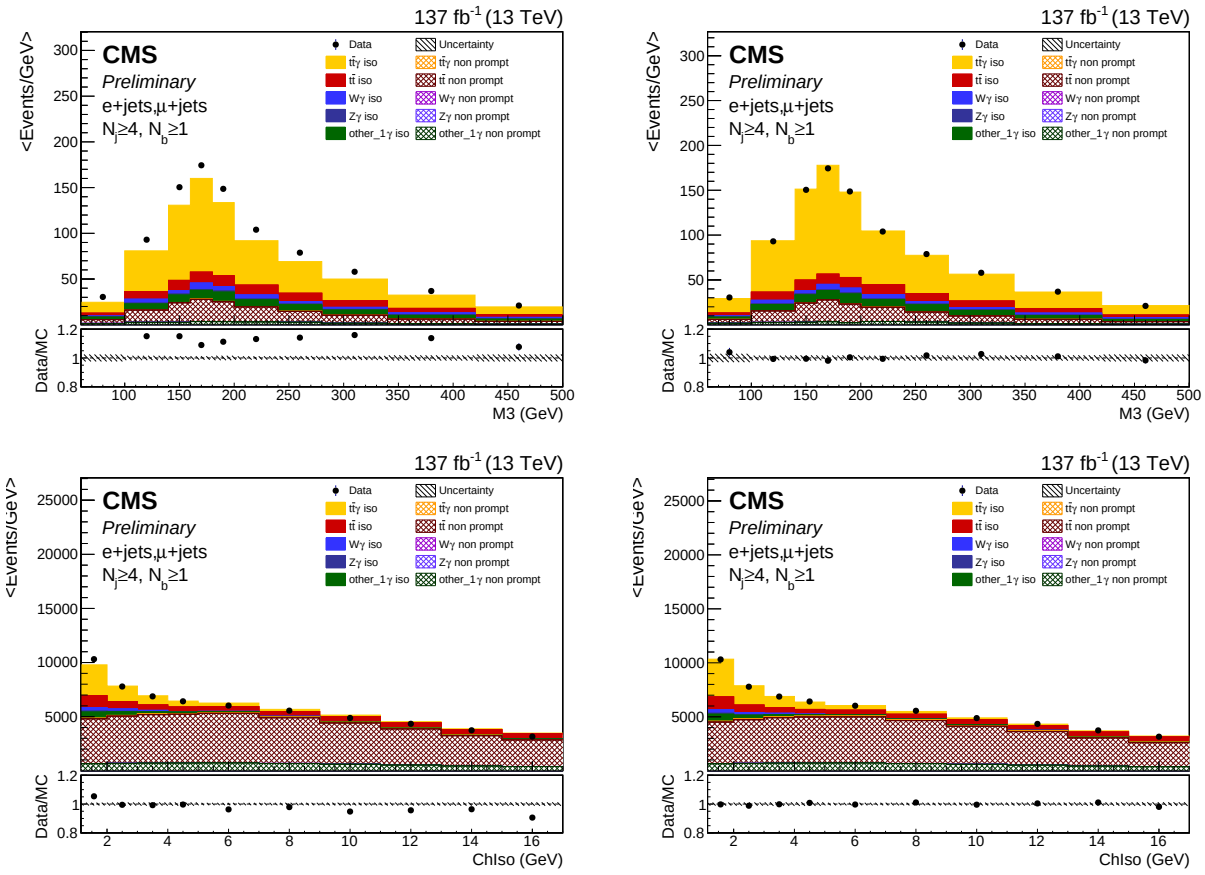


Figure 103: Pre-fit (left) and post-fit (right) plots for data MC agreement in the signal region for M_3 (top) and ChIso (bottom) with all years combined.

effectively removes the normalization that was put in when the samples were scaled to an integrated luminosity. To start out, we have a fit function, with free floating parameters for both the $t\bar{t}\gamma$ and $t\bar{t}$ scale of:

$$f(\mu_{t\bar{t}\gamma}, \mu_{t\bar{t}}) = \mu_{t\bar{t}\gamma} \cdot N^{\text{expected}}_{t\bar{t}\gamma} + \mu_{t\bar{t}} \cdot N^{\text{expected}}_{t\bar{t}} + N_{\text{bkg}} \tag{8.3}$$

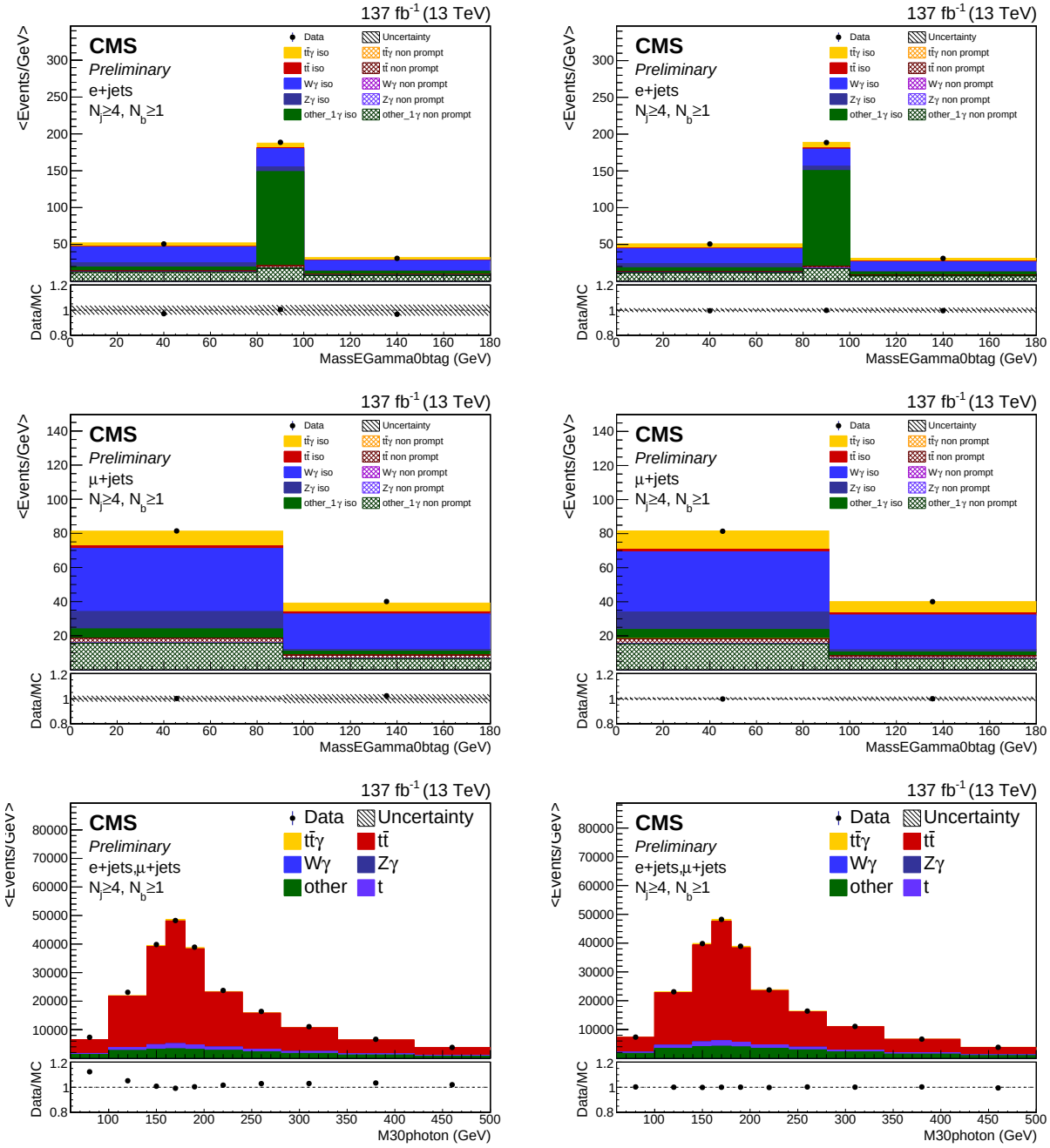


Figure 104: Pre-fit (left) and post-fit (right) plots for data MC agreement in zero btag (top) and zero photon (bottom) control regions with all years combined.

The number of events in $\mathrm{t\bar{t}\gamma}$ and $\mathrm{t\bar{t}}$ can be expanded to:

$$\begin{aligned} N^{\text{expected}}_{\mathrm{t\bar{t}\gamma}} &= \sigma^{\text{theory}}_{\mathrm{t\bar{t}\gamma}} \cdot \epsilon_{\mathrm{t\bar{t}\gamma}} \cdot A_{\mathrm{t\bar{t}\gamma}} \cdot \mathrm{L}\,, \\ N^{\text{expected}}_{\mathrm{t\bar{t}}} &= \sigma^{\text{theory}}_{\mathrm{t\bar{t}}} \cdot \epsilon_{\mathrm{t\bar{t}}} \cdot A_{\mathrm{t\bar{t}}} \cdot \mathrm{L}\,. \end{aligned} \tag{8.4}$$

Substituting back into the fit function we get:

$$f(\mu_{\mathrm{t\bar{t}\gamma}}\,, \mu_{\mathrm{t\bar{t}}}\,) = \mu_{\mathrm{t\bar{t}\gamma}} \cdot \sigma^{\text{theory}}_{\mathrm{t\bar{t}\gamma}} \cdot \epsilon_{\mathrm{t\bar{t}\gamma}} \cdot A_{\mathrm{t\bar{t}\gamma}} \cdot \mathrm{L} + \mu_{\mathrm{t\bar{t}}} \cdot \sigma^{\text{theory}}_{\mathrm{t\bar{t}}} \cdot \epsilon_{\mathrm{t\bar{t}}} \cdot A_{\mathrm{t\bar{t}}} \cdot \mathrm{L} + N_{\text{bkg}} \tag{8.5}$$

The observed cross section is equal to the theoretical value multiplied by the signal strength parameter (μ):

$$\begin{aligned} \sigma^{\text{obs.}}_{\mathrm{t\bar{t}\gamma}} &= \mu_{\mathrm{t\bar{t}\gamma}} \cdot \sigma^{\text{theory}}_{\mathrm{t\bar{t}\gamma}} \\ \sigma^{\text{obs.}}_{\mathrm{t\bar{t}}} &= \mu_{\mathrm{t\bar{t}}} \cdot \sigma^{\text{theory}}_{\mathrm{t\bar{t}}} \\ \epsilon_{\mathrm{t\bar{t}\gamma}} \cdot A_{\mathrm{t\bar{t}\gamma}} \cdot \mathrm{L} &= N^{\text{expected}}_{\mathrm{t\bar{t}\gamma}} / \sigma^{\text{theory}}_{\mathrm{t\bar{t}\gamma}} \\ \epsilon_{\mathrm{t\bar{t}}} \cdot A_{\mathrm{t\bar{t}}} \cdot \mathrm{L} &= N^{\text{expected}}_{\mathrm{t\bar{t}}} / \sigma^{\text{theory}}_{\mathrm{t\bar{t}}} \end{aligned} \tag{8.6}$$

Thus, by scaling the number of events by the theoretical cross-section, we get a fit function in which the parameters are directly the measured cross sections.

$$f(\sigma^{\text{obs.}}_{\mathrm{t\bar{t}\gamma}}\,, \sigma^{\text{obs.}}_{\mathrm{t\bar{t}}}\,) = \sigma^{\text{obs.}}_{\mathrm{t\bar{t}\gamma}} \cdot (N^{\text{expected}}_{\mathrm{t\bar{t}\gamma}} / \sigma^{\text{theory}}_{\mathrm{t\bar{t}\gamma}}) + \sigma^{\text{obs.}}_{\mathrm{t\bar{t}}} \cdot (N^{\text{expected}}_{\mathrm{t\bar{t}}} / \sigma^{\text{theory}}_{\mathrm{t\bar{t}}}) + N_{\text{bkg}} \tag{8.7}$$

Finally, if the $\mathrm{t\bar{t}\gamma}$ portion is multiplied by $\sigma^{\text{obs.}}_{\mathrm{t\bar{t}}} / \sigma^{\text{obs.}}_{\mathrm{t\bar{t}}} = 1$, the terms can be rearranged into the cross section ratio $R = \sigma^{\text{obs.}}_{\mathrm{t\bar{t}\gamma}} / \sigma^{\text{obs.}}_{\mathrm{t\bar{t}}}$. Thus, by scaling both the $\mathrm{t\bar{t}}$ and $\mathrm{t\bar{t}\gamma}$ portions of

the fit function by $\sigma_{\mathrm{t\bar{t}}}^{\text{obs.}}$, the fit function can be redefined in terms of $\sigma_{\mathrm{t\bar{t}}\gamma}^{\text{obs.}}$ and the ratio R

$$f(R_\gamma\,,\sigma_{\mathrm{t\bar{t}}}^{\text{obs.}}) = R_\gamma\ \cdot \sigma_{\mathrm{t\bar{t}}}^{\text{obs.}}\ \cdot (N_{\mathrm{t\bar{t}}\gamma}^{\text{expected}}/\sigma_{\mathrm{t\bar{t}}\gamma}^{\text{theory}}) + \sigma_{\mathrm{t\bar{t}}}^{\text{obs.}}\ \cdot (N_{\mathrm{t\bar{t}}}^{\text{expected}}/\sigma_{\mathrm{t\bar{t}}}^{\text{theory}}) + N_{\text{bkg}} \tag{8.8}$$

Finally we can calculate the R_γ directly from Higgs Combine Tool.

$$R = \frac{\sigma_{\mathrm{t\bar{t}}\gamma\,,incl}^{meas}(\mathrm{P_T}\,(\gamma) \geq 20\ \mathrm{GeV}\,,\eta < 1.44)}{\sigma_{\mathrm{t\bar{t}}\,,incl}^{meas}} \tag{8.9}$$

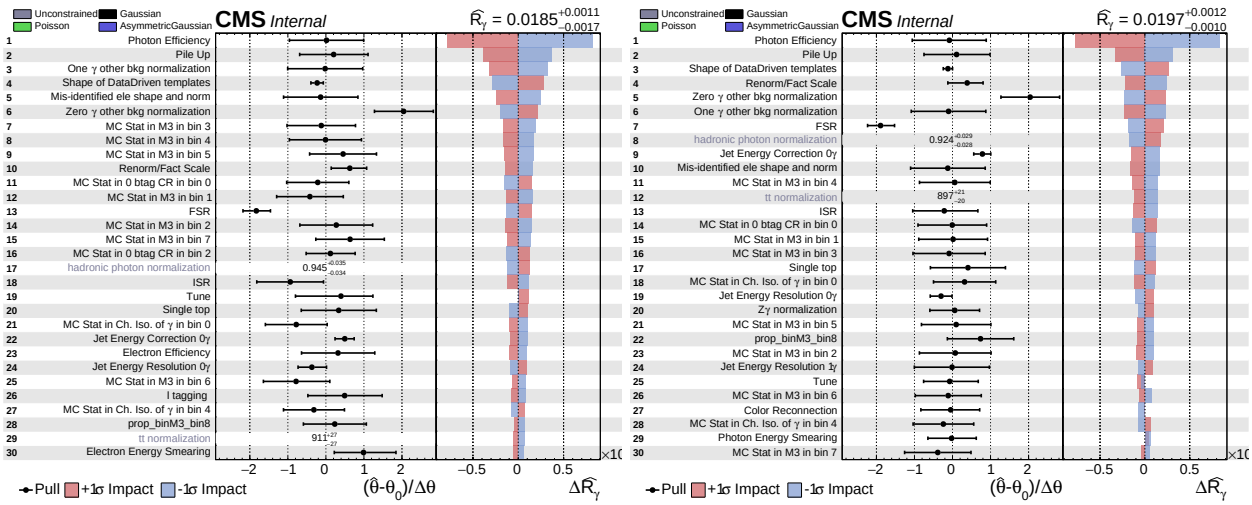


Figure 105: The impact plot that shows postfit pulls and constraints on nuisance parameters for R_γ fit in $e+$jets (left) and $\mu+$jets (right) separately to 2016 data.

Figures 105 - 109 show the pulls on nuisance parameters and their impacts on μ, and the maximum likelihood scans for 2016, 2017, 2018, and all years combined. The extracted values for R_γ, nonprompt photon correction factor, and measured $\mathrm{t\bar{t}}$cross section in the signal phase space are shown in Tables 44 - 46 for the $e+$jets and the $\mu+$jets channels respectively for 2016, 2017, and 2018 data. A description of the implementation in the HiggscombineTool is in Appendix D. The correlation plots can be found in Appendix G.

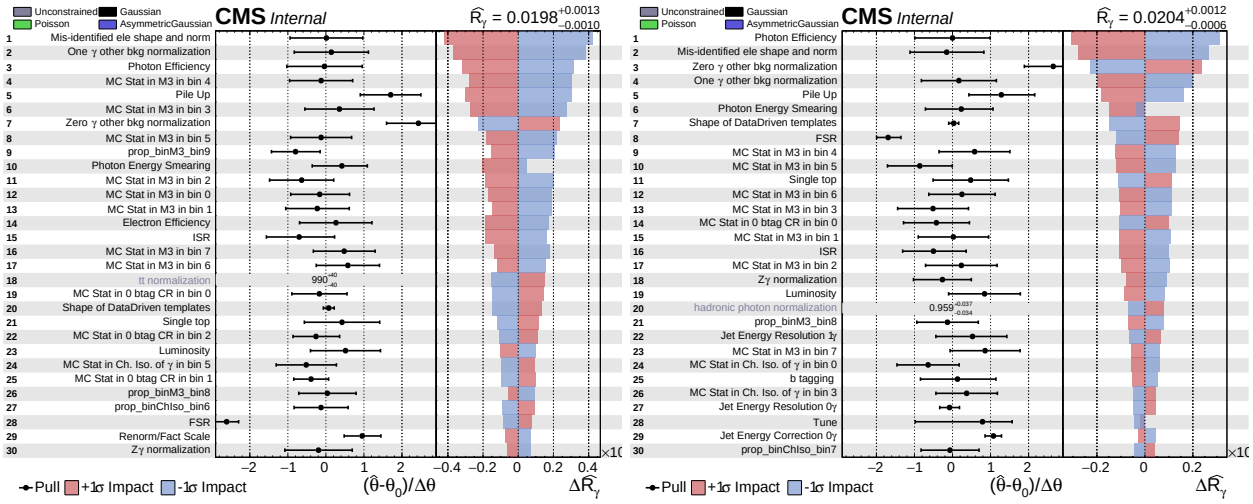


Figure 106: The impact plot that shows postfit pulls and constraints on nuisance parameters for R_γ fit in $e+$jets (left) and $\mu+$jets (right) separately to 2017 data.

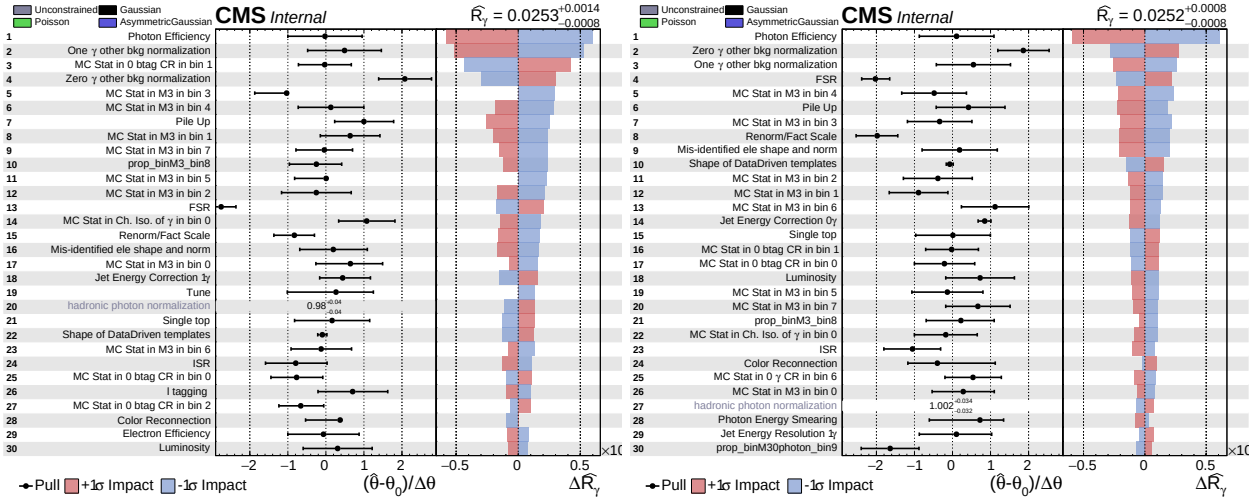


Figure 107: The impact plot that shows postfit pulls and constraints on nuisance parameters for R_γ fit in $e+$jets (left) and $\mu+$jets (right) separately to 2018 data.

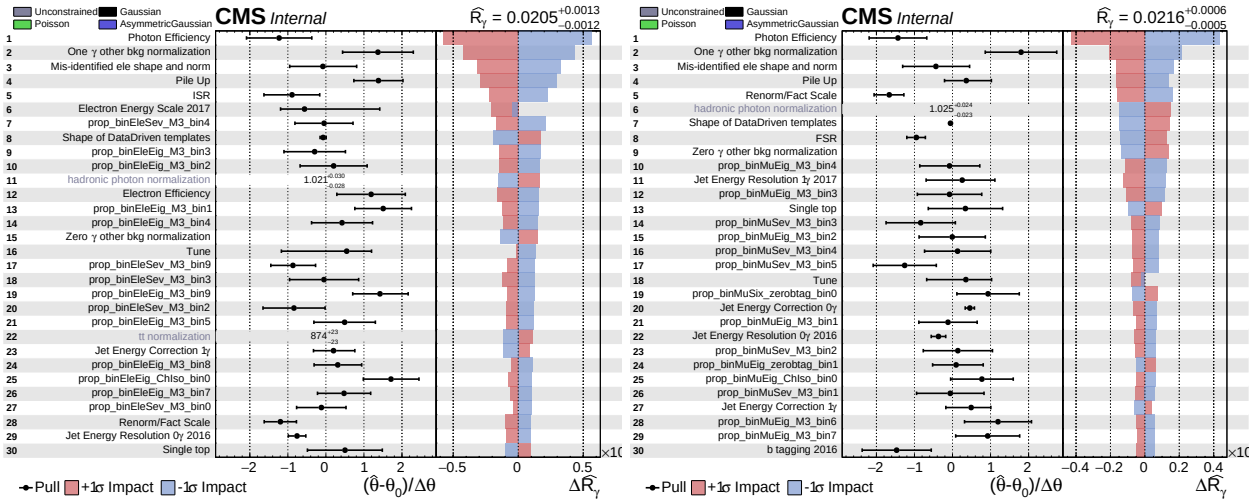


Figure 108: The impact plot that shows postfit pulls and constraints on nuisance parameters for R_γ fit in $e+$jets (left) and $\mu+$jets (right) separately to full Run 2 data.

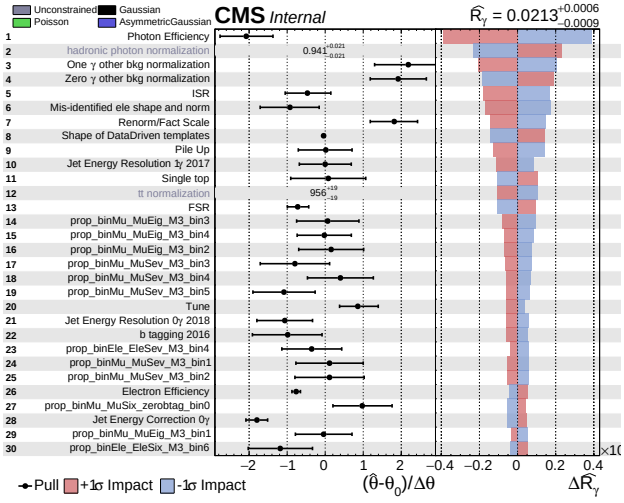


Figure 109: The impact plot that shows postfit pulls and constraints on nuisance parameters for R_γ fit to full Run 2 data.

8.4 Summary Result for Full RunII

Finally, we extracted the signal strength (μ), defined as $\sigma^{\text{measured}}_{\text{t}\bar{\text{t}}\gamma}/\sigma^{\text{NLO}}_{\text{t}\bar{\text{t}}\gamma}$, for $e+$jets , $\mu+$jets , and $l+$jets channels for full Run 2 and shown in Table 39. We found photon acceptances, A^{γ}, (38.56 $\pm$ 0.02)%, (43.05 $\pm$ 0.02)%, and (38.54 $\pm$ 0.01)% for 2016, 2017, 2018 respectively in the signal phase space defined by Equation 8.2. The weighted average for full Run 2 is (40.05 $\pm$ 0.01)%. The $\text{t}\bar{\text{t}}\gamma$ acceptances, $A^{\text{t}\bar{\text{t}}\gamma}$, for fiducial phase space defined as in Table 32 are found to be (6.28 $\pm$ 0.01)%, (6.05 $\pm$ 0.01)%, and (7.98 $\pm$ 0.01)% for 2016, 2017, 2018 respectively. The weighted average for full Run 2 is (6.86 $\pm$ 0.01)%.

The measured inclusive cross section is $3.81 \pm 0.15\,(\text{syst.}) \pm 0.10\,(\text{stat.})$ pb in e channel , $3.87 \pm 0.11\,(\text{syst.}) \pm 0.07\,(\text{stat.})$ pb in μ channel , and $3.96 \pm 0.10\,(\text{syst.}) \pm 0.06\,(\text{stat.})$ pb in $e+\mu$ channel , respectively for full RunII data. The fiducial semileptonic cross section is $0.65 \pm 0.03\,(\text{syst.}) \pm 0.02\,(\text{stat.})$ pb in e channel , $0.66 \pm 0.02\,(\text{syst.}) \pm 0.01\,(\text{stat.})$ pb in μ channel , and $0.68 \pm 0.02\,(\text{syst.}) \pm 0.01\,(\text{stat.})$ pb in $e+\mu$ channel , respectively.

Finally we extract the R_{γ} (as defined in Equation 8.9). The extracted values for R_{γ} , nonprompt photon correction factor, and measured $\text{t}\bar{\text{t}}$ cross section in the signal phase space are shown in Table 40 for e channel , μ channel , and $e+\mu$ channel , respectively for all years combined. The measurement of the ratio of $\text{t}\bar{\text{t}}\gamma$ to $\text{t}\bar{\text{t}}$ is $0.02055 \pm 0.00099\,(\text{syst.}) \pm 0.00099\,(\text{stat.})$ in e channel , $0.02156 \pm 0.00068\,(\text{syst.}) \pm 0.00068\,(\text{stat.})$ in μ channel , and $0.02203 \pm 0.00064\,(\text{syst.}) \pm 0.00064\,(\text{stat.})$ in $e+\mu$ channel , respectively.

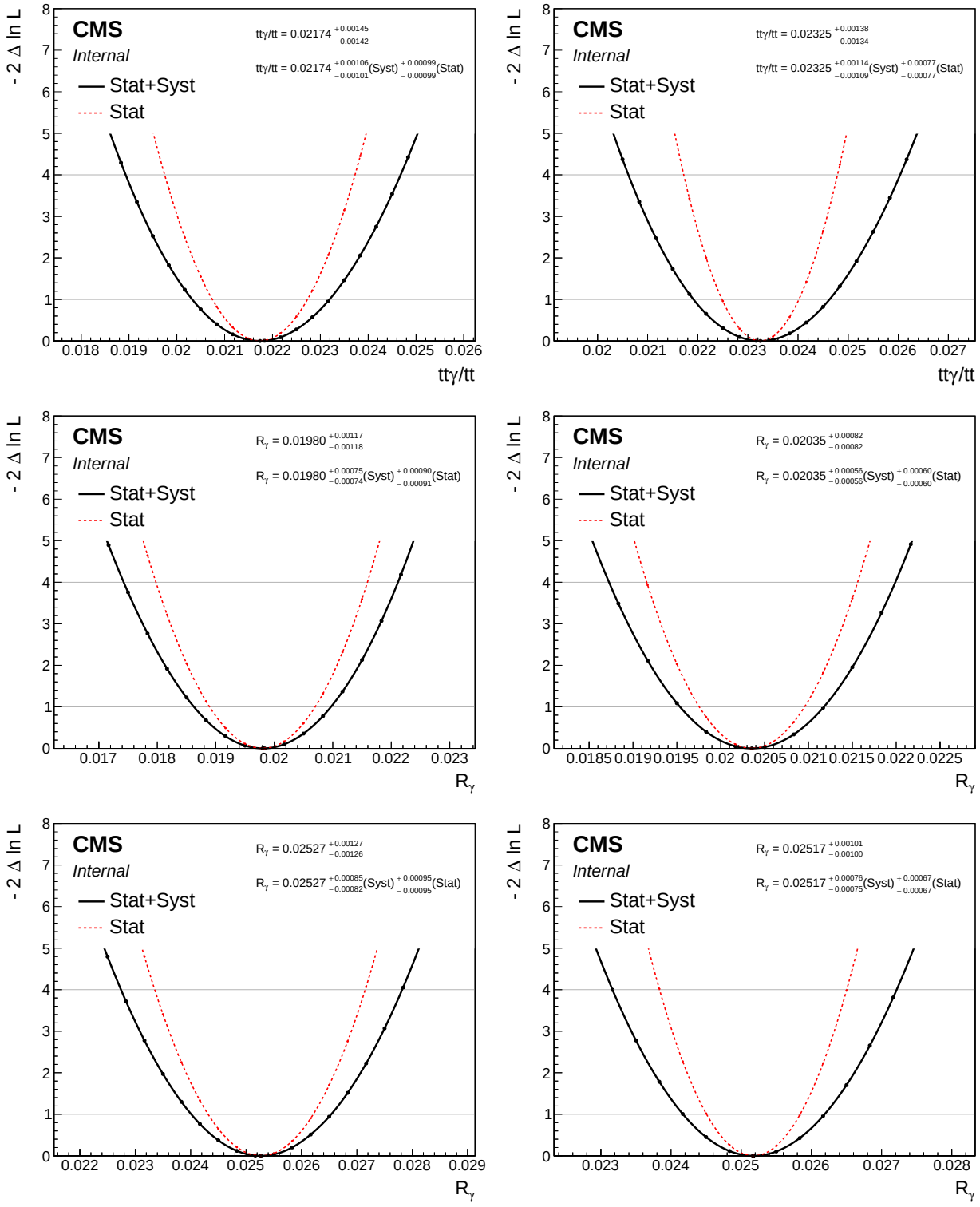


Figure 110: Double negative log likelihood distribution for R_γ in $e+$jets (left) and $\mu+$jets (right) channels for 2016 (top), 2017 (middle), and 2018 (bottom) data taking periods.

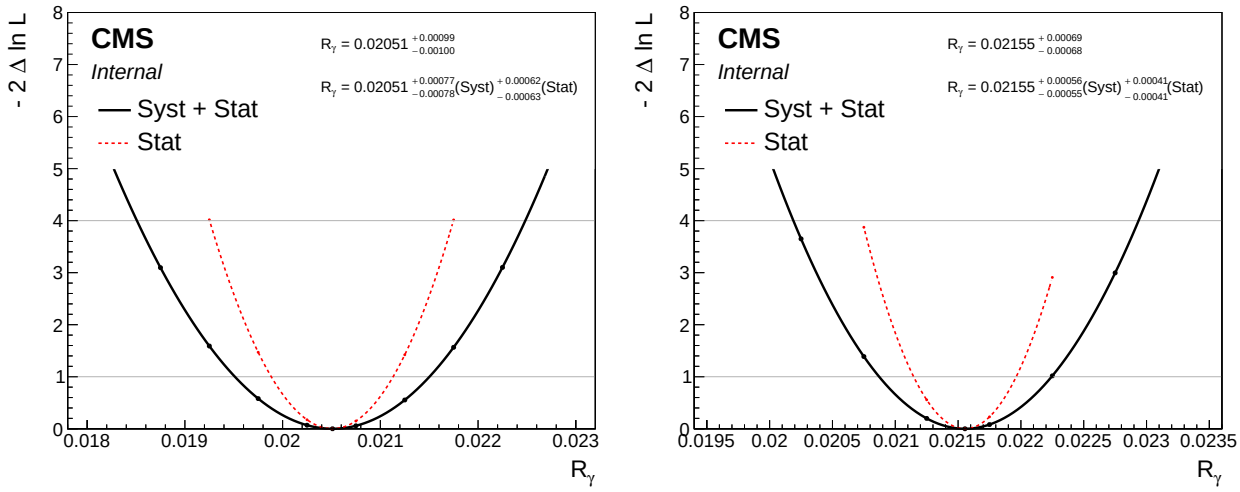


Figure 111: Double negative log likelihood distribution for R_γ in $e+$jets (left) and $\mu+$jets (right) channels for full Run 2 data.

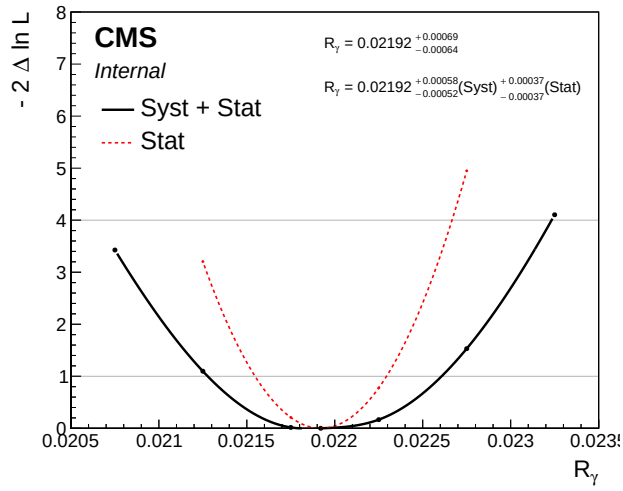


Figure 112: Double negative log likelihood distribution for R_γ for full Run 2 data.

Table 36: The ratio (R_γ) extraction along with other important scale factors for $e+$jets and $\mu+$jets channels for 2016 data.

Year	channel	R_γ	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2016	$e+$jets	$0.01848^{+0.00130}_{-0.00126}$	$0.95^{+0.00}_{-0.03}$	$911.43^{+26.19}_{-26.19}$	$1.00^{+0.09}_{-0.09}$	$0.97^{+0.28}_{-0.28}$	$1.00^{+0.10}_{-0.10}$	$1.20^{+0.08}_{-0.08}$
2016	$\mu+$jets	$0.01972^{+0.00115}_{-0.00110}$	$0.92^{+0.00}_{-0.03}$	$897.33^{+20.42}_{-20.42}$	$1.01^{+0.08}_{-0.08}$	$1.02^{+0.26}_{-0.26}$	$0.99^{+0.10}_{-0.10}$	$1.20^{+0.08}_{-0.08}$

Table 37: The ratio (R_γ) extraction along with other important scale factors for $e+$jets and $\mu+$jets channels for 2017 data.

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2017	$e+$jets	$0.01980^{+0.00116}_{-0.00116}$	$0.94^{+0.00}_{-0.04}$	$991.10^{+35.47}_{-35.47}$	$0.96^{+0.09}_{-0.09}$	$0.92^{+0.35}_{-0.35}$	$1.01^{+0.10}_{-0.10}$	$1.24^{+0.08}_{-0.08}$
2017	$\mu+$jets	$0.02035^{+0.00082}_{-0.00082}$	$0.96^{+0.00}_{-0.03}$	$995.03^{+27.68}_{-27.68}$	$0.93^{+0.07}_{-0.07}$	$0.89^{+0.30}_{-0.30}$	$1.02^{+0.10}_{-0.10}$	$1.26^{+0.08}_{-0.08}$

Table 38: The ratio (R_γ) extraction along with other important scale factors for $e+$jets and $\mu+$jets channels for 2018 data.

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2018	$e+$jets	$0.02527^{+0.00125}_{-0.00125}$	$0.98^{+0.00}_{-0.04}$	$805.80^{+29.46}_{-29.46}$	$0.93^{+0.09}_{-0.09}$	$0.92^{+0.38}_{-0.38}$	$1.05^{+0.10}_{-0.10}$	$1.21^{+0.07}_{-0.07}$
2018	$\mu+$jets	$0.02517^{+0.00100}_{-0.00100}$	$1.00^{+0.00}_{-0.03}$	$798.46^{+25.42}_{-25.42}$	$0.96^{+0.08}_{-0.08}$	$1.03^{+0.35}_{-0.35}$	$1.06^{+0.10}_{-0.10}$	$1.19^{+0.07}_{-0.07}$

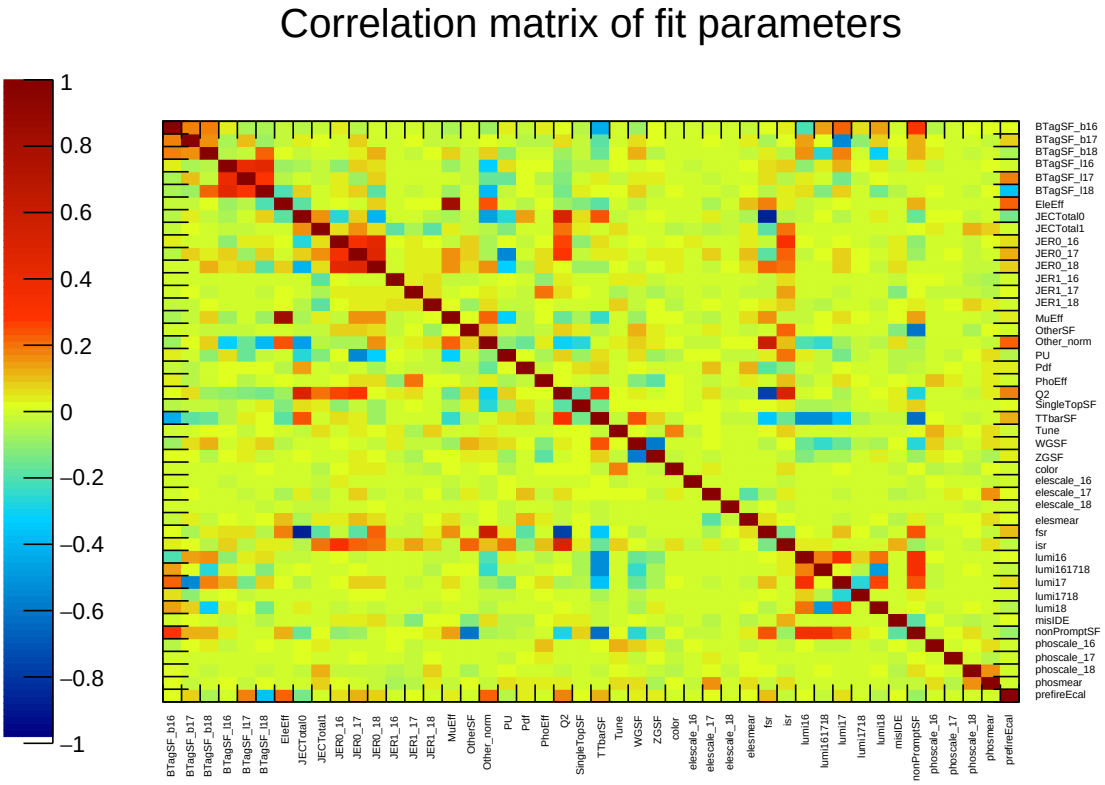


Figure 113: Double negative log likelihood distribution for R_γ with full Run 2 data.

Table 39: Summary of the signal strength extraction for all years combined

Full Run 2	$t\bar{t}\gamma$ SF	nonPrompt SF
e + jets	$1.27^{+0.06}_{-0.06}$	$0.94^{+0.06}_{-0.02}$
μ + jets	$1.29^{+0.04}_{-0.04}$	$0.96^{+0.04}_{-0.02}$
l + jets	$1.32^{+0.04}_{-0.04}$	$0.95^{+0.04}_{-0.02}$

Table 40: Summary of the ratio (R_γ) extraction for all years combined

Full Run 2	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF	$t\bar{t}$ cross section (pb)
e + jets	$0.02055^{+0.00099}_{-0.00099}$	$1.02^{+0.00}_{-0.03}$	$873.69^{+22.80}_{-22.80}$
μ + jets	$0.02156^{+0.00068}_{-0.00068}$	$1.03^{+0.00}_{-0.02}$	$870.28^{+15.39}_{-15.39}$
l + jets	$0.02203^{+0.00064}_{-0.00064}$	$1.01^{+0.00}_{-0.02}$	$876.95^{+13.64}_{-13.64}$

CHAPTER 9 CONCLUSION AND OUTLOOK

We make a step towards the direct quantification of the electromagnetic interaction of the top quark using the CMS detector at the LHC. We present fiducial and inclusive cross section measurements of the semileptonic $pp \rightarrow \mathrm{t\bar{t}}\gamma$ process using the full Run 2 data with a corresponding integrated luminosity of 137 fb^{-1} collected by the CMS detector at the center of mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$. We used the centrally produced primary and Monte-Carlo data set for signal and backgrounds. To normalize the $\mathrm{t\bar{t}}\gamma$ signal sample to NLO level, we calculated the k-factor using MADGRAPH5_aMC@NLO . Events are required to pass the HLT trigger (either single electron or single muon trigger), to have a well defined primary vertex, with a high $\mathrm{P_T}$ and well isolated lepton (electron or muon), with $\geq$4 jets and $\geq$ 1 b-tagged jet, and exactly one high $\mathrm{P_T}$ photon. All the selected events and objects are corrected according to the recommendation provided by CMS Top Quark Physics Group. All the major backgrounds are estimated using data driven techniques.

The analysis strategy is separate the events into genuine, misidentified, hadronic, and fake/pile-up category based on the origin of photon. A data driven estimation of hadronic and fake photons is performed using a photon isolation variable ($\sigma_{i\eta i\eta}$) in the signal region. A data driven estimation of misidentified electrons is performed along with $W+\gamma$ and $Z+\gamma$ using the invariant mass of lepton and photon in a control region with no b-jet. The effects of misidentified electrons, $W+\gamma$, and $Z+\gamma$ are carried to the signal region. A data driven QCD multijet background is estimated in a control region with no b-jet by inverting lepton isolation. The QCD template is used to replace the MC simulation in signal region.

Experimental, theoretical, and data driven estimation uncertainties are implemented in

a fit as recommended by TOP Systematic Uncertainties (Run 2). The maximum likelihood fit to Asimov and toy MC are performed to check the following: closure test of parameter of interest, diagnostics on uncertainty model, correlation between the parameter of interest and nuisances, and profile maximum log-likelihood features of fit parameters. The hadronic photon normalization and $\mathrm{t\bar{t}}$ cross section normalization enters as free-floating parameters in the fit. The cross section of $\mathrm{t\bar{t}}\gamma$ and the ratio $\frac{\sigma_{\mathrm{t\bar{t}}\gamma}}{\sigma_{\mathrm{t\bar{t}}}}$ are extracted using a maximum likelihood fit to data in $e+$ jets and $\mu+$ jets channels for each of the data taking periods 2016, 2017, and 2018 separately, and data from both the channels and all years combined to extract the cross section and ratio for full Run 2 data. The measurement of the ratio of $\mathrm{t\bar{t}}\gamma$ to $\mathrm{t\bar{t}}$ is $0.02203 \pm 0.00064\,(\text{syst.}) \pm 0.00064\,(\text{stat.})$ in $e+\mu$ channel and the measured inclusive cross section is $3.96 \pm 0.10\,(\text{syst.}) \pm 0.06\,(\text{stat.})$ pb in $e+\mu$ channel for full RunII data. The total inclusive cross section with a photon of transverse momentum $\geq$ 20 GeV agrees with the NLO SM predictions within uncertainties.

APPENDIX A

A simple Higgs combine model: Let n be the observed number of events, b be the number of background events from Monte Carlo (MC) simulation and s be the number of signal events to be determined in a counting experiments. Let δ and Δ be the MC statistical uncertainty and systematic uncertainties. The likelihood function in this model is

$$\mathcal{L}(s,\alpha,\gamma) = P(n|s+\gamma b+\alpha\Delta) \times P(m|\gamma m) \times G(\alpha|0,1)\ ,$$

where s is the parameter of interest, α and γ are two nuisance parameters; $m \equiv \frac{b^2}{\delta^2}$ is a constant; $P(n|\lambda) = \frac{\lambda^n}{n!}e^{-\lambda}$ is the Poisson distribution function with the expectation value λ, $G(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ is the Gaussian distribution function with mean μ and standard deviation σ. The log likelihood function is then

$$\ln\mathcal{L} = n\ln(s+\gamma b+\alpha\Delta) - (s+\gamma b+\alpha\Delta) - \frac{\alpha^2}{2} + m\ln\gamma - \gamma m\ .$$

Maximizing the likelihood function leads to the following estimation of the best-fit values.

$$\hat{s} = n-b\ , \quad \hat{\alpha} = 0\ , \quad \hat{\gamma} = 1\ .$$

Let $\mathbf{V}$ be the covariance matrix of the fitting parameters, the inverse of its estimation at the best-fit values,

$$(\mathbf{V}^{-1})_{ij} = -\frac{\partial^2\ln\mathcal{L}(\hat{\theta}_i,\hat{\theta}_j)}{\partial\theta_i\partial\theta_j}$$

where θ_is denote the parameters $(\theta_1, \theta_2, \theta_3) = (s, \alpha, \gamma)$. The inverse of $\mathbf{V}$ is

$$\mathbf{V}^{-1} = \frac{1}{n} \begin{pmatrix} 1 & \Delta & b \\ \Delta & \Delta^2 + n & \Delta b \\ b & \Delta b & b^2 + n\frac{b^2}{\delta^2} \end{pmatrix} ,$$

with the determinant $\det |\mathbf{V}^{-1}| = \frac{b^2}{n\delta^2}$. The diagonal elements of $\mathbf{V}$ give the uncertainty of the fitting parameters [119].

$$\hat{\sigma}_s = \sqrt{\mathbf{V}_{11}} = \sqrt{n + \Delta^2 + \delta^2} \, , \quad \hat{\sigma}_\alpha = \sqrt{\mathbf{V}_{22}} = 1 \, , \quad \hat{\sigma}_\gamma = \sqrt{\mathbf{V}_{33}} = \frac{\delta}{b} \, .$$

A higgs combine Ratio model: The default model of combine is just to include a scale factor (with a range of 0 to 20) that multiplies whatever template is declared as signal.

$$S_{\text{data}} = \mu' S_{\text{signal}} + \Sigma S_{\text{bkg}}$$

where μ' is just a correction factor combine needs to apply to S_{signal} to get the best fit. Typically, when we normalize the signal template to the number of events we expect to see from a predicted cross section, μ can be interpreted as how much we need to correct the predicted cross section, i.e., $\mu = \sigma_{\text{measured}} / \sigma_{\text{theory}}$. However, if we do a new fit where we scale the S_{signal} by some constant, the meaning of the fit parameter μ of this fit would be different

$$S_{\text{data}} = \mu'.(C.S_{\text{signal}}) + \Sigma S_{\text{bkg}}$$
$$\mu' = \frac{\mu}{C}$$

where, μ' is just the initial μ from the first fit over the constant factor C. Now, replace C with $1/\sigma_{\text{theory}}$.

$$S_{\text{data}} = \mu'.(\frac{S_{\text{signal}}}{\sigma_{\text{theory}}}) + \Sigma S_{\text{bkg}}$$

$$\begin{aligned}\mu' &= \mu.\sigma_{\text{theory}} \\ &= (\frac{\sigma_{\text{measured}}}{\sigma_{\text{theory}}}).\sigma_{\text{theory}} \\ &= \sigma_{\text{measured}}\end{aligned}$$

This result in fitting to a cross section. We also need to fit for ratio $R = \mathrm{t\bar{t}\gamma}\,/\mathrm{t\bar{t}}$. First, we allow $\mathrm{t\bar{t}}$ to float and scale it with $\mu_{\mathrm{t\bar{t}}}$.

$$S_{\text{data}} = \mu_{\text{t}}\bar{\text{t}}\gamma \,.(S_{\text{t}\bar{\text{t}}\gamma}/\sigma^{\text{t}\bar{\text{t}}\gamma}_{\text{theory}}) + \mu_{\text{t}}\bar{\text{t}} \,.(S_{\text{t}\bar{\text{t}}}/\sigma^{\text{t}\bar{\text{t}}}_{\text{theory}})\Sigma S_{\text{bkg}}$$

Next, we scale $S_{\text{t}}\bar{\text{t}}\gamma$ by $\mu_{\,\text{t}\bar{\text{t}}}$.

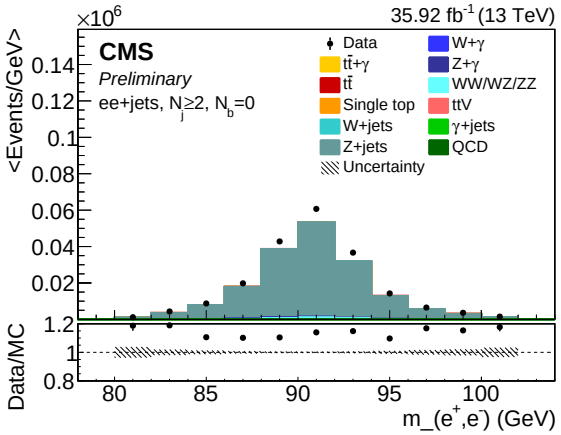
$$S_{\text{data}} = \mu''.\mu_{\text{t}}\bar{\text{t}} \,.(S_{\text{t}\bar{\text{t}}\gamma}/\sigma^{\text{t}\bar{\text{t}}\gamma}_{\text{theory}}) + \mu_{\text{t}}\bar{\text{t}} \,.(S_{\text{t}\bar{\text{t}}}/\sigma^{\text{t}\bar{\text{t}}}_{\text{theory}})\Sigma S_{\text{bkg}}$$

which means signal template is scaled by $\mu_{\,\text{t}\bar{\text{t}}}$ over the $\text{t}\bar{\text{t}}\gamma$ theory cross section and $\mu^{''}$ will be the ratio of $\text{t}\bar{\text{t}}\gamma$ to $\text{t}\bar{\text{t}}$ cross section.

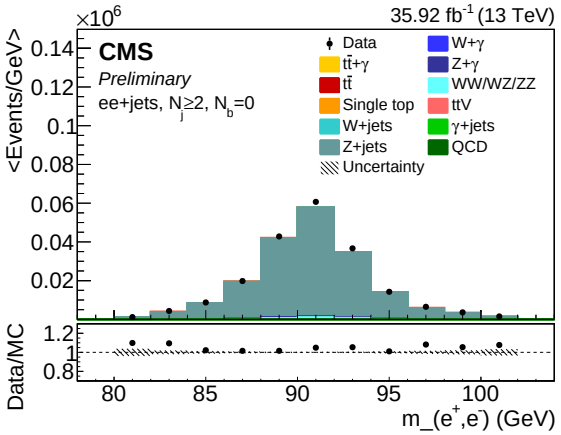
$$\begin{aligned}
\mu'' &= \mu.(\frac{\sigma^{\text{t}\bar{\text{t}}\gamma}_{\text{theory}}}{\sigma^{\text{t}\bar{\text{t}}}_{\text{measured}}}) \\
&= (\frac{\sigma^{\text{t}\bar{\text{t}}\gamma}_{\text{measured}}}{\sigma^{\text{t}\bar{\text{t}}\gamma}_{\text{theory}}}).(\frac{\sigma^{\text{t}\bar{\text{t}}\gamma}_{\text{theory}}}{\sigma^{\text{t}\bar{\text{t}}}_{\text{measured}}}) \\
&= \frac{\sigma^{\text{t}\bar{\text{t}}\gamma}_{\text{measured}}}{\sigma^{\text{t}\bar{\text{t}}}_{\text{measured}}}
\end{aligned}$$

For the combine implementation, we introduce three parameters a) a scale factor of $1/\sigma_{\text{t}}\bar{\text{t}}\gamma$ which is multiplied to $\text{t}\bar{\text{t}}\gamma$template and freeze it b) a scale factor of $1/\sigma_{\text{t}}\bar{\text{t}}$ which is multiplied to $\text{t}\bar{\text{t}}$ template and freeze it c) a parameter that multiplies both $\text{t}\bar{\text{t}}$ and $\text{t}\bar{\text{t}}\gamma$ templates and let it float.

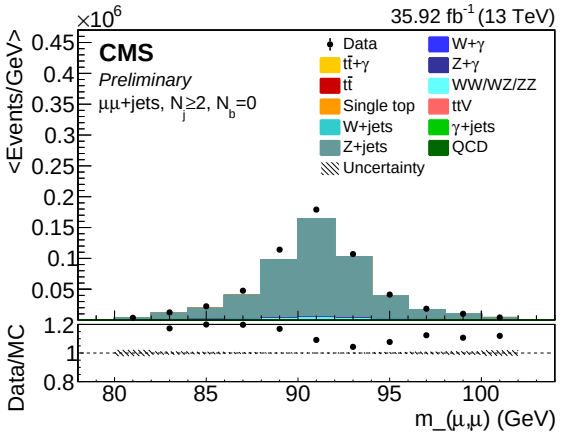
APPENDIX B



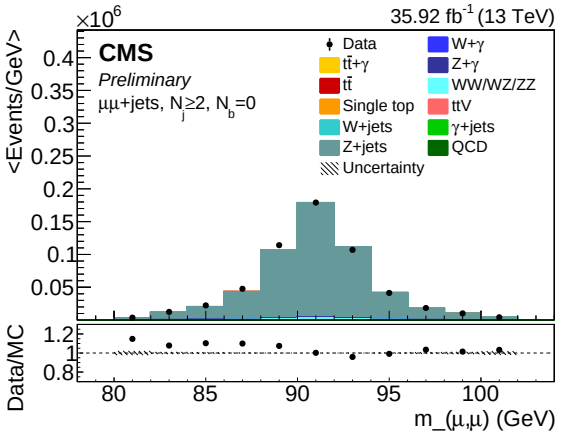
(a) prefit, ee + jets, 2016



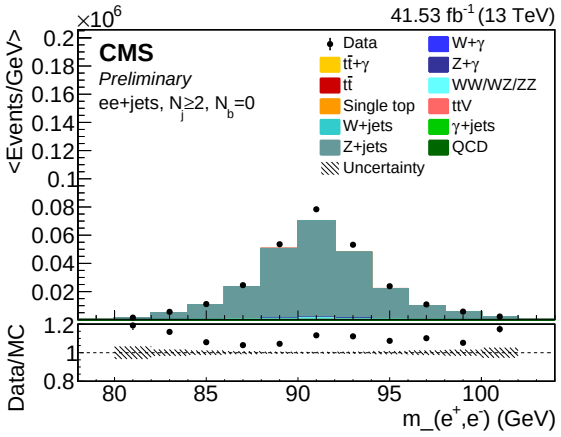
(b) postfit, ee + jets, 2016



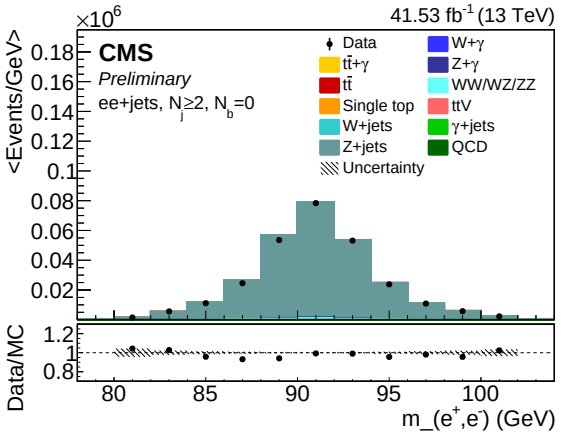
(c) prefit, $\mu\mu$ + jets, 2016



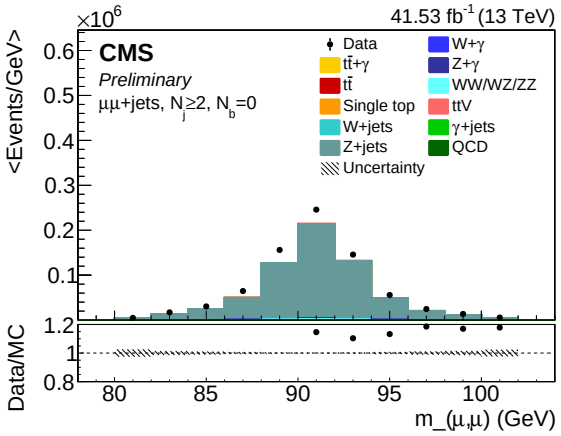
(d) postfit, $\mu\mu$ + jets, 2016



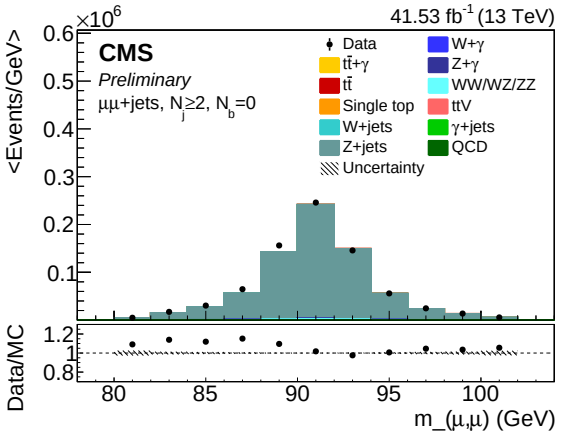
(e) prefit, ee + jets, 2017



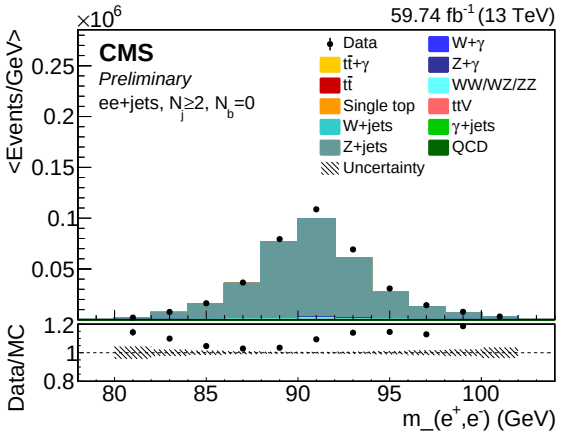
(f) postfit, ee + jets, 2017



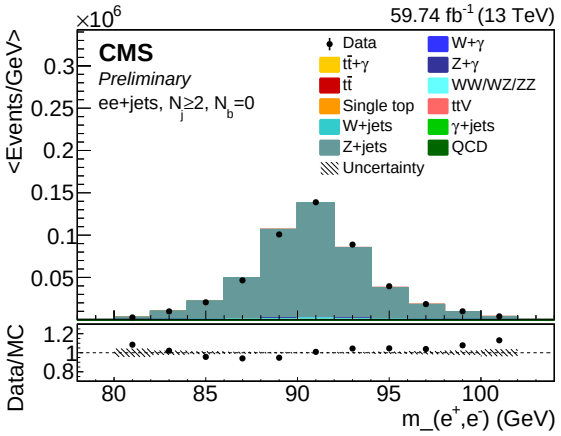
(g) prefit,$\mu\mu$ + jets , 2017



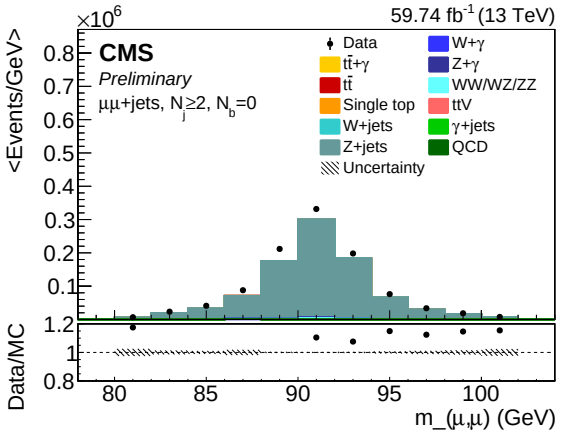
(h) postfit, $\mu\mu$ + jets , 2017



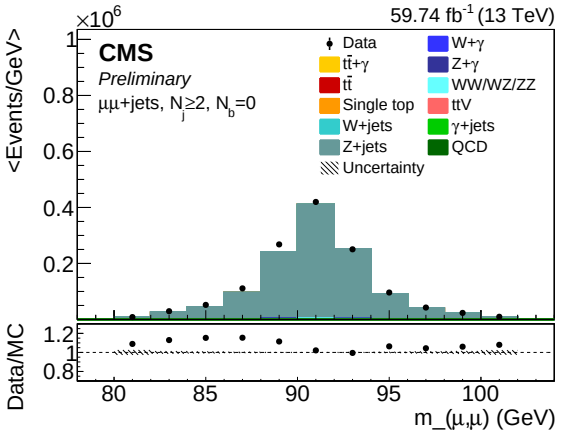
(i) prefit, ee + jets , 2018



(j) postfit, ee + jets , 2018

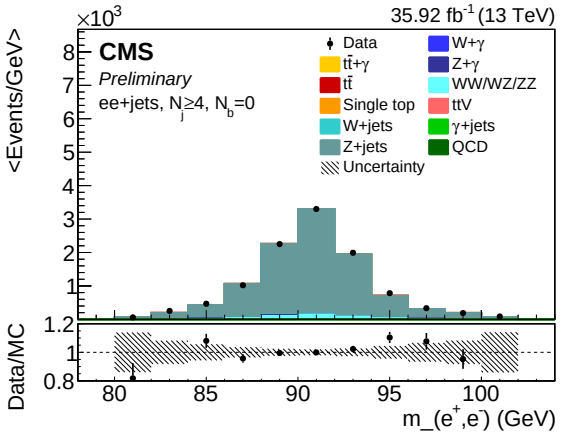


(k) prefit, $\mu\mu$ + jets , 2018

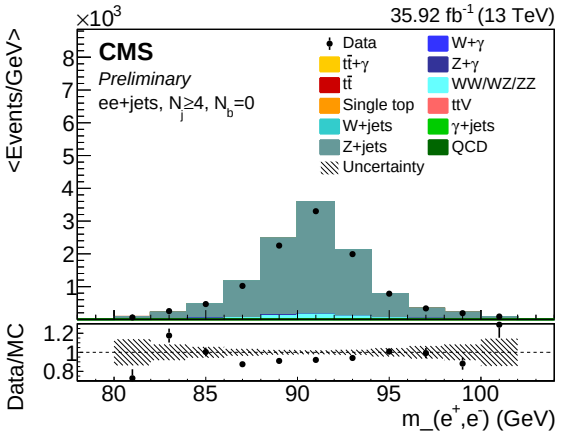


(l) postfit, $\mu\mu$ + jets , 2018

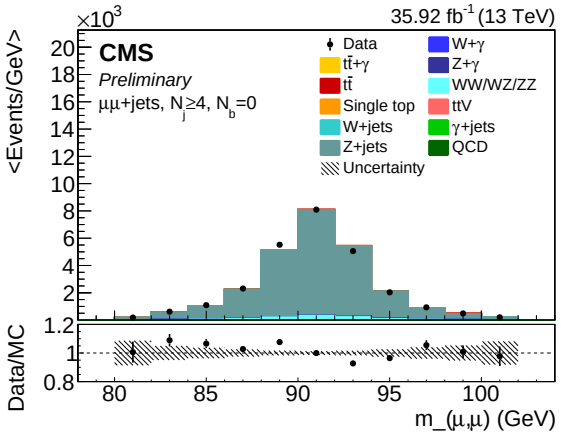
Figure 114: The prefit(left) and postfit(right) plots of an invariant mass distribution of dilepton samples in a control region with $N_\ell = 2, N_\gamma = 0$, $N_j \geq 2, N_b = 0$ in the ee + jets channel and the $\mu\mu$ + jets channel for 2016, 2017, and 2018 dataset.



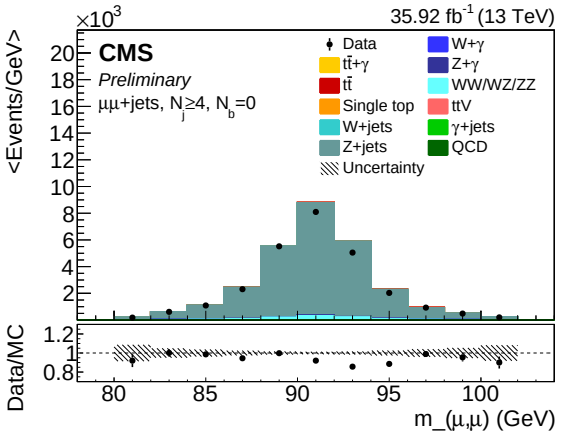
(a) prefit, ee + jets, 2016



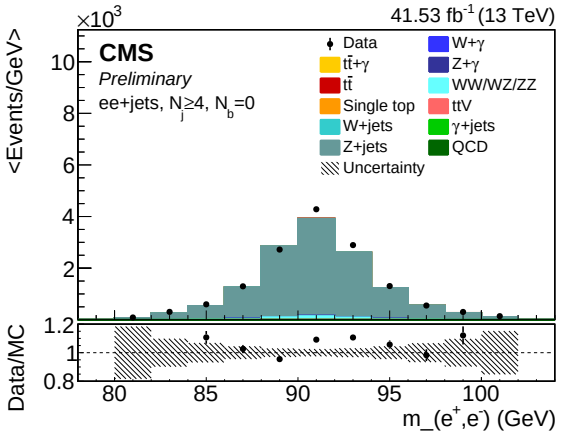
(b) postfit, ee + jets, 2016



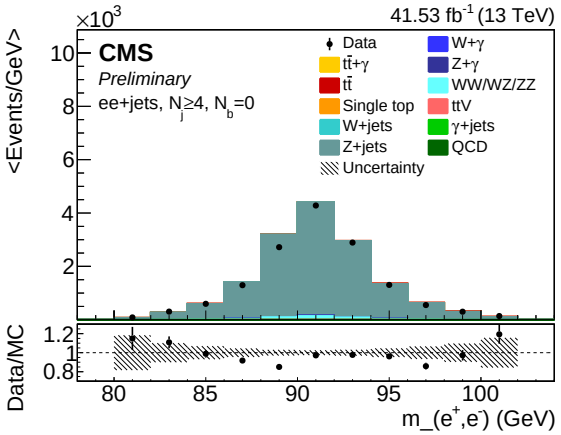
(c) prefit, $\mu\mu$ + jets, 2016



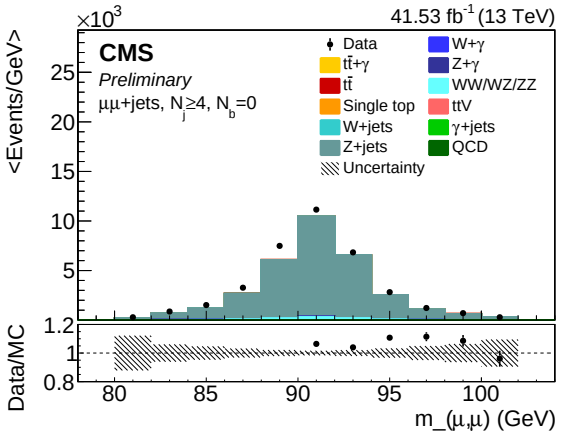
(d) postfit, $\mu\mu$ + jets, 2016



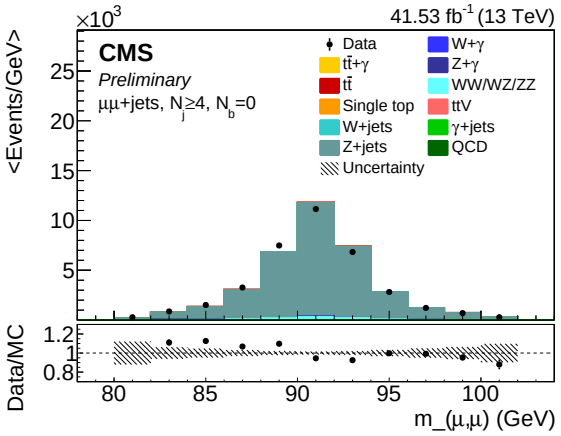
(e) prefit, ee + jets, 2017



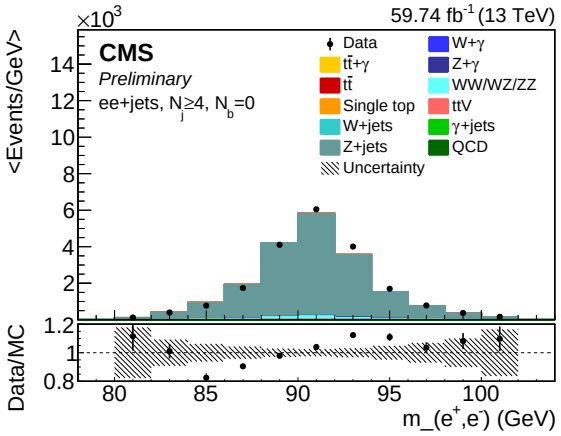
(f) postfit, ee + jets, 2017



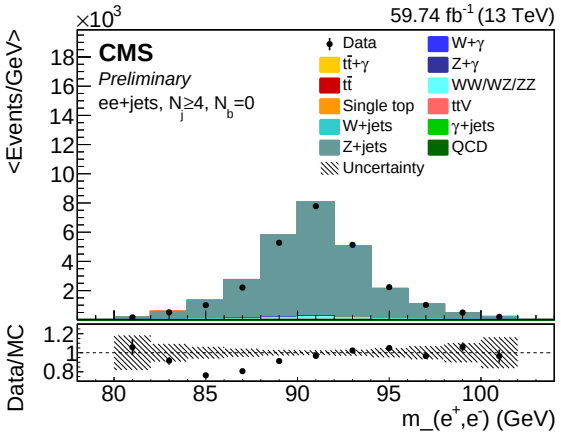
(a) prefit,$\mu\mu$ + jets , 2017



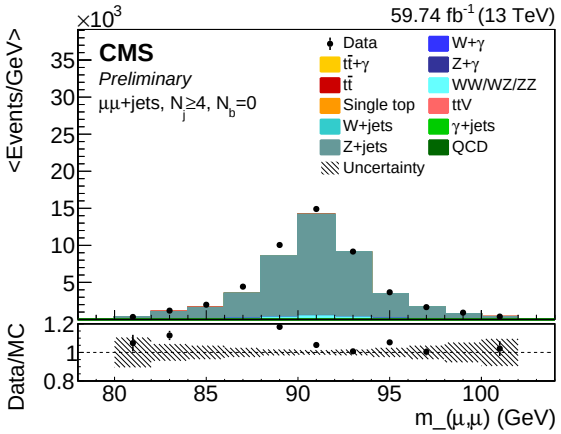
(b) postfit, $\mu\mu$ + jets , 2017



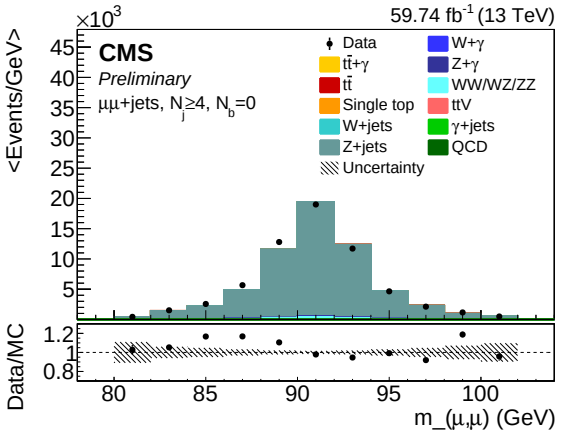
(c) prefit, ee + jets , 2018



(d) postfit, ee + jets , 2018



(e) prefit, $\mu\mu$ + jets , 2018



(f) postfit, $\mu\mu$ + jets , 2018

Figure 116: The prefit(left) and postfit(right) plots of an invariant mass distribution of dilepton samples in a control region with $N_\ell = 2, N_\gamma = 0$, $N_j \geq 4, N_b = 0$ in the ee + jets channel and the $\mu\mu$ + jets channel for 2016, 2017, and 2018 dataset.

Table 41: The signal strength (μ) extraction along with other important scale factors for $e+$jets and $\mu+$jets channels for 2016 data taking.

Prefit:

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2016	$e+$jets	free floating	free floating	$1 \pm 5\%$	$1 \pm 19\%$	$1 \pm 21\%$	$1 \pm 10\%$	$1 \pm 10\%$
2016	$\mu+$jets	free floating	free floating	$1 \pm 5\%$	$1 \pm 19\%$	$1 \pm 21\%$	$1 \pm 10\%$	$1 \pm 10\%$

Postfit Asimov: Expected $t\bar{t}\gamma$ and nonPrompt SF = 1

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2016	$e+$jets	$1.00^{+0.07}_{-0.07}$	$1.00^{+0.01}_{-0.01}$	$1.00^{+0.00}_{-0.00}$	$1.00^{+0.03}_{-0.03}$	$1.00^{+0.09}_{-0.09}$	$1.00^{+0.04}_{-0.04}$	$1.00^{+0.01}_{-0.01}$
2016	$\mu+$jets	$1.00^{+0.06}_{-0.06}$	$1.00^{+0.06}_{-0.03}$	$1.00^{+0.01}_{-0.01}$	$1.00^{+0.08}_{-0.08}$	$1.00^{+0.26}_{-0.26}$	$1.00^{+0.11}_{-0.11}$	$1.00^{+0.08}_{-0.08}$

Postfit Toy: Expected $t\bar{t}\gamma$ and nonPrompt SF = 1 with 500 toys

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2016	$e+$jets	$0.99^{+0.07}_{-0.07}$	$1.00^{+0.04}_{-0.04}$	$1.00^{+0.05}_{-0.05}$	$1.09^{+0.33}_{-0.33}$	$0.90^{+0.25}_{-0.25}$	$1.00^{+0.03}_{-0.03}$	$1.00^{+0.07}_{-0.07}$
2016	$\mu+$jets	$1.00^{+0.07}_{-0.07}$	$1.00^{+0.03}_{-0.03}$	$1.00^{+0.05}_{-0.05}$	$1.07^{+0.36}_{-0.36}$	$0.86^{+0.27}_{-0.27}$	$1.00^{+0.02}_{-0.02}$	$1.00^{+0.06}_{-0.06}$

Postfit Data:

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2016	$e+$jets	$1.12^{+0.08}_{-0.08}$	$0.89^{+0.08}_{-0.02}$	$0.99^{+0.01}_{-0.01}$	$1.02^{+0.09}_{-0.09}$	$1.05^{+0.28}_{-0.28}$	$0.99^{+0.10}_{-0.10}$	$1.16^{+0.08}_{-0.08}$
2016	$\mu+$jets	$1.18^{+0.07}_{-0.07}$	$0.87^{+0.07}_{-0.02}$	$0.99^{+0.01}_{-0.01}$	$1.03^{+0.08}_{-0.08}$	$1.10^{+0.26}_{-0.26}$	$0.99^{+0.10}_{-0.10}$	$1.12^{+0.07}_{-0.07}$

Table 42: The signal strength (μ) extraction along with other important scale factors for $e+$jets and $\mu+$jets channels for 2017 data taking.

Prefit:

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2017	$e+$jets	free floating	free floating	$1 \pm 5\%$	$1 \pm 19\%$	$1 \pm 21\%$	$1 \pm 10\%$	$1 \pm 10\%$
2017	$\mu+$jets	free floating	free floating	$1 \pm 5\%$	$1 \pm 19\%$	$1 \pm 21\%$	$1 \pm 10\%$	$1 \pm 10\%$

Postfit Asimov: Expected $t\bar{t}\gamma$ and nonPrompt SF = 1

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2017	$e+$jets	$1.00^{+0.02}_{-0.02}$	$1.00^{+0.01}_{-0.01}$	$1.00^{+0.00}_{-0.00}$	$1.00^{+0.03}_{-0.03}$	$1.00^{+0.14}_{-0.14}$	$1.00^{+0.04}_{-0.04}$	$1.00^{+0.01}_{-0.01}$
2017	$\mu+$jets	$1.00^{+0.06}_{-0.06}$	$1.00^{+0.06}_{-0.06}$	$1.00^{+0.04}_{-0.04}$	$1.00^{+0.11}_{-0.11}$	$1.00^{+0.34}_{-0.34}$	$1.00^{+0.14}_{-0.14}$	$1.00^{+0.08}_{-0.08}$

Postfit Toy: Expected $t\bar{t}\gamma$ and nonPrompt SF = 1 with 500 toys

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2017	$e+$jets	$0.98^{+0.07}_{-0.07}$	$1.00^{+0.04}_{-0.04}$	$1.00^{+0.01}_{-0.01}$	$1.07^{+0.37}_{-0.37}$	$0.92^{+0.27}_{-0.27}$	$1.02^{+0.02}_{-0.02}$	$0.99^{+0.06}_{-0.06}$
2017	$\mu+$jets	$1.01^{+0.05}_{-0.05}$	$1.00^{+0.03}_{-0.03}$	$1.00^{+0.01}_{-0.01}$	$1.01^{+0.36}_{-0.36}$	$0.97^{+0.39}_{-0.39}$	$1.00^{+0.02}_{-0.02}$	$1.00^{+0.07}_{-0.07}$

Postfit Data:

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2017	$e+$jets	$1.22^{+0.07}_{-0.07}$	$0.95^{+0.07}_{-0.03}$	$1.00^{+0.01}_{-0.01}$	$0.95^{+0.08}_{-0.08}$	$0.91^{+0.35}_{-0.35}$	$1.02^{+0.10}_{-0.10}$	$1.25^{+0.08}_{-0.08}$
2017	$\mu+$jets	$1.25^{+0.05}_{-0.05}$	$0.97^{+0.05}_{-0.02}$	$1.00^{+0.01}_{-0.01}$	$0.93^{+0.07}_{-0.07}$	$0.86^{+0.30}_{-0.30}$	$1.02^{+0.10}_{-0.10}$	$1.27^{+0.08}_{-0.08}$

Table 43: The signal strength (μ) extraction along with other important scale factors for $e+$jets and $\mu+$jets channels for 2018 data taking.

Prefit:

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2018	$e+$jets	free floating	free floating	$1 \pm 5\%$	$1 \pm 19\%$	$1 \pm 21\%$	$1 \pm 10\%$	$1 \pm 10\%$
2018	$\mu+$jets	free floating	free floating	$1 \pm 5\%$	$1 \pm 19\%$	$1 \pm 21\%$	$1 \pm 10\%$	$1 \pm 10\%$

Postfit Asimov: Expected $t\bar{t}\gamma$ and nonPrompt SF = 1

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2018	$e+$jets	$1.00^{+0.06}_{-0.06}$	$1.00^{+0.02}_{-0.02}$	$1.00^{+0.01}_{-0.01}$	$1.00^{+0.07}_{-0.07}$	$1.00^{+0.36}_{-0.36}$	$1.00^{+0.09}_{-0.09}$	$1.00^{+0.03}_{-0.03}$
2018	$\mu+$jets	$1.00^{+0.05}_{-0.05}$	$1.00^{+0.05}_{-0.02}$	$1.00^{+0.01}_{-0.01}$	$1.00^{+0.07}_{-0.07}$	$1.00^{+0.34}_{-0.34}$	$1.00^{+0.10}_{-0.10}$	$1.00^{+0.07}_{-0.07}$

Postfit Toy: Expected $t\bar{t}\gamma$ and nonPrompt SF = 1 with 500 toys

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2018	$e+$jets	$0.99^{+0.06}_{-0.06}$	$1.00^{+0.03}_{-0.03}$	$1.00^{+0.01}_{-0.01}$	$1.05^{+0.40}_{-0.40}$	$0.92^{+0.16}_{-0.16}$	$1.02^{+0.03}_{-0.03}$	$1.00^{+0.07}_{-0.07}$
2018	$\mu+$jets	$1.00^{+0.06}_{-0.06}$	$1.00^{+0.03}_{-0.03}$	$1.00^{+0.01}_{-0.01}$	$1.01^{+0.37}_{-0.37}$	$0.96^{+0.30}_{-0.30}$	$1.01^{+0.02}_{-0.02}$	$1.00^{+0.07}_{-0.07}$

Postfit Data:

Year	channel	$t\bar{t}\gamma$ SF	nonPrompt SF	TTbar SF	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2018	$e+$jets	$1.32^{+0.07}_{-0.07}$	$0.95^{+0.07}_{-0.02}$	$1.00^{+0.01}_{-0.01}$	$0.95^{+0.08}_{-0.08}$	$0.98^{+0.37}_{-0.37}$	$1.04^{+0.10}_{-0.10}$	$1.20^{+0.07}_{-0.07}$
2018	$\mu+$jets	$1.31^{+0.05}_{-0.05}$	$0.97^{+0.05}_{-0.02}$	$1.00^{+0.01}_{-0.01}$	$0.98^{+0.07}_{-0.07}$	$1.07^{+0.34}_{-0.34}$	$1.05^{+0.10}_{-0.10}$	$1.18^{+0.07}_{-0.07}$

Table 44: The ratio (R) extraction along with other important scale factors for $e+$jets and $\mu+$jets channels for 2016 data taking.

Prefit:

Year	channel	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2016	$e+$jets	free floating	free floating	1 ± 831	$1 \pm 19\%$	$1 \pm 21\%$	$1 \pm 10\%$	$1 \pm 10\%$
2016	$\mu+$jets	free floating	free floating	1 ± 831	$1 \pm 19\%$	$1 \pm 21\%$	$1 \pm 10\%$	$1 \pm 10\%$

Postfit Asimov: Expected $t\bar{t}\gamma$ and nonPrompt SF = 1

Year	channel	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2016	$e+$jets	$0.02000^{+0.00137}_{-0.00137}$	$1.00^{+0.04}_{-0.04}$	$831.75^{+23.72}_{-23.72}$	$1.00^{+0.09}_{-0.09}$	$1.00^{+0.28}_{-0.28}$	$1.00^{+0.10}_{-0.10}$	$1.00^{+0.07}_{-0.07}$
2016	$\mu+$jets	$0.02000^{+0.00120}_{-0.00115}$	$1.00^{+0.00}_{-0.03}$	$831.74^{+18.14}_{-18.14}$	$1.00^{+0.08}_{-0.08}$	$1.00^{+0.26}_{-0.26}$	$1.00^{+0.10}_{-0.10}$	$1.00^{+0.07}_{-0.07}$

Postfit Toy: Expected $\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$ and nonPrompt SF = 1 with 500 toys

Year	channel	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2016	$e+$jets	$0.01981^{+0.00139}_{-0.00139}$	$1.00^{+0.04}_{-0.04}$	$830.39^{+23.40}_{-23.40}$	$1.06^{+0.36}_{-0.36}$	$0.97^{+0.40}_{-0.40}$	$1.01^{+0.02}_{-0.02}$	$1.00^{+0.06}_{-0.06}$
2016	$\mu+$jets	$0.01998^{+0.00131}_{-0.00131}$	$1.00^{+0.03}_{-0.03}$	$830.85^{+17.21}_{-17.21}$	$1.08^{+0.38}_{-0.38}$	$0.97^{+0.45}_{-0.45}$	$1.00^{+0.02}_{-0.02}$	$1.00^{+0.07}_{-0.07}$

Postfit Data:

Year	channel	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2016	$e+$jets	$0.01848^{+0.00130}_{-0.00126}$	$0.95^{+0.00}_{-0.03}$	$911.43^{+26.19}_{-26.19}$	$1.00^{+0.09}_{-0.09}$	$0.97^{+0.28}_{-0.28}$	$1.00^{+0.10}_{-0.10}$	$1.20^{+0.08}_{-0.08}$
2016	$\mu+$jets	$0.01972^{+0.00115}_{-0.00110}$	$0.92^{+0.00}_{-0.03}$	$897.33^{+20.42}_{-20.42}$	$1.01^{+0.08}_{-0.08}$	$1.02^{+0.26}_{-0.26}$	$0.99^{+0.10}_{-0.10}$	$1.20^{+0.08}_{-0.08}$

Table 45: The ratio (R) extraction along with other important scale factors for $e+$jets and $\mu+$jets channels for 2017 data taking.

Prefit:

Year	channel	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2017	$e+$jets	free floating	free floating	831	$1 \pm 19\%$	$1 \pm 21\%$	$1 \pm 10\%$	$1 \pm 10\%$
2017	$\mu+$jets	free floating	free floating	831	$1 \pm 19\%$	$1 \pm 21\%$	$1 \pm 10\%$	$1 \pm 10\%$

Postfit Asimov: Expected $t\bar{t}\gamma$ and nonPrompt SF = 1

Year	channel	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2017	$e+$jets	$0.02000^{+0.00123}_{-0.00125}$	$1.00^{+0.04}_{-0.04}$	$831.79^{+33.27}_{-33.27}$	$1.00^{+0.08}_{-0.08}$	$1.00^{+0.36}_{-0.36}$	$1.00^{+0.10}_{-0.10}$	$1.00^{+0.08}_{-0.08}$
2017	$\mu+$jets	$0.02000^{+0.00086}_{-0.00085}$	$1.00^{+0.00}_{-0.04}$	$831.79^{+25.89}_{-25.89}$	$1.00^{+0.07}_{-0.07}$	$1.00^{+0.31}_{-0.31}$	$1.00^{+0.10}_{-0.10}$	$1.00^{+0.07}_{-0.07}$

Postfit Toy: Expected $\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$ and nonPrompt SF = 1 with 500 toys

Year	channel	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2017	$e+$jets	$0.01964^{+0.00127}_{-0.00127}$	$1.00^{+0.05}_{-0.05}$	$826.43^{+31.56}_{-31.56}$	$1.06^{+0.37}_{-0.37}$	$0.90^{+0.25}_{-0.25}$	$1.02^{+0.02}_{-0.02}$	$1.00^{+0.05}_{-0.05}$
2017	$\mu+$jets	$0.02014^{+0.00106}_{-0.00106}$	$1.00^{+0.04}_{-0.04}$	$830.61^{+24.45}_{-24.45}$	$1.03^{+0.38}_{-0.38}$	$0.96^{+0.38}_{-0.38}$	$1.00^{+0.02}_{-0.02}$	$1.00^{+0.07}_{-0.07}$

Postfit Data:

Year	channel	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2017	$e+$jets	$0.01980^{+0.00116}_{-0.00116}$	$0.94^{+0.00}_{-0.04}$	$991.10^{+35.47}_{-35.47}$	$0.96^{+0.09}_{-0.09}$	$0.92^{+0.35}_{-0.35}$	$1.01^{+0.10}_{-0.10}$	$1.24^{+0.08}_{-0.08}$
2017	$\mu+$jets	$0.02035^{+0.00082}_{-0.00082}$	$0.96^{+0.00}_{-0.03}$	$995.03^{+27.68}_{-27.68}$	$0.93^{+0.07}_{-0.07}$	$0.89^{+0.30}_{-0.30}$	$1.02^{+0.10}_{-0.10}$	$1.26^{+0.08}_{-0.08}$

Table 46: The ratio (R) extraction along with other important scale factors for $e+$jets and $\mu+$jets channels for 2018 data taking.

Prefit:

Year	channel	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2018	$e+$jets	free floating	free floating	831	$1 \pm 19\%$	$1 \pm 21\%$	$1 \pm 10\%$	$1 \pm 10\%$
2018	$\mu+$jets	free floating	free floating	831	$1 \pm 19\%$	$1 \pm 21\%$	$1 \pm 10\%$	$1 \pm 10\%$

Postfit Asimov: Expected $t\bar{t}\gamma$ and nonPrompt SF = 1

Year	channel	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2018	$e+$jets	$0.02000^{+0.00110}_{-0.00109}$	$1.00^{+0.04}_{-0.04}$	$831.74^{+29.96}_{-29.96}$	$1.00^{+0.08}_{-0.08}$	$1.00^{+0.38}_{-0.38}$	$1.00^{+0.10}_{-0.10}$	$1.00^{+0.07}_{-0.07}$
2018	$\mu+$jets	$0.02000^{+0.00090}_{-0.00090}$	$1.00^{+0.00}_{-0.03}$	$831.72^{+26.38}_{-26.38}$	$1.00^{+0.08}_{-0.08}$	$1.00^{+0.35}_{-0.35}$	$1.00^{+0.10}_{-0.10}$	$1.00^{+0.07}_{-0.07}$

Postfit Toy: Expected $\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$ and nonPrompt SF = 1 with 500 toys

Year	channel	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2018	$e+$jets	$0.01974^{+0.00122}_{-0.00122}$	$1.00^{+0.05}_{-0.05}$	$826.75^{+30.33}_{-30.33}$	$1.06^{+0.36}_{-0.36}$	$0.92^{+0.15}_{-0.15}$	$1.02^{+0.03}_{-0.03}$	$1.00^{+0.07}_{-0.07}$
2018	$\mu+$jets	$0.01996^{+0.00111}_{-0.00111}$	$1.00^{+0.04}_{-0.04}$	$830.50^{+26.82}_{-26.82}$	$1.05^{+0.38}_{-0.38}$	$0.95^{+0.28}_{-0.28}$	$1.01^{+0.02}_{-0.02}$	$1.00^{+0.07}_{-0.07}$

Postfit Data:

Year	channel	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF	TTbar XS	$W+\gamma$ SF	$Z+\gamma$ SF	Other_1γ SF	Other_0γ SF
2018	$e+$jets	$0.02527^{+0.00125}_{-0.00125}$	$0.98^{+0.00}_{-0.04}$	$805.80^{+29.46}_{-29.46}$	$0.93^{+0.09}_{-0.09}$	$0.92^{+0.38}_{-0.38}$	$1.05^{+0.10}_{-0.10}$	$1.21^{+0.07}_{-0.07}$
2018	$\mu+$jets	$0.02517^{+0.00100}_{-0.00100}$	$1.00^{+0.00}_{-0.03}$	$798.46^{+25.42}_{-25.42}$	$0.96^{+0.08}_{-0.08}$	$1.03^{+0.35}_{-0.35}$	$1.06^{+0.10}_{-0.10}$	$1.19^{+0.07}_{-0.07}$

Table 47: Summary of the signal strength extraction for full RunII data

Postfit Asimov:

Full RunII	$t\bar{t}\gamma$ SF	nonPrompt SF
$e+$jets	$1.00^{+0.04}_{-0.04}$	$1.00^{+0.01}_{-0.01}$
$\mu+$jets	$1.00^{+0.03}_{-0.03}$	$1.00^{+0.03}_{-0.01}$
$l+$jets	$1.00^{+0.02}_{-0.02}$	$1.00^{+0.02}_{-0.01}$

Postfit Toy:

Full RunII	$t\bar{t}\gamma$ SF	nonPrompt SF
$e+$jets	$0.97^{+0.05}_{-0.05}$	$0.99^{+0.02}_{-0.02}$
$\mu+$jets	$1.00^{+0.04}_{-0.04}$	$1.00^{+0.03}_{-0.03}$
$l+$jets	$0.98^{+0.03}_{-0.03}$	$1.00^{+0.02}_{-0.02}$

Postfit Data:

Full RunII	$t\bar{t}\gamma$ SF	nonPrompt SF
$e+$jets	$1.27^{+0.06}_{-0.06}$	$0.94^{+0.06}_{-0.02}$
$\mu+$jets	$1.29^{+0.04}_{-0.04}$	$0.96^{+0.04}_{-0.02}$
$l+$jets	$1.32^{+0.04}_{-0.04}$	$0.95^{+0.04}_{-0.02}$

Postfit Toy:

Full RunII	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF
$e+$jets	$0.01941^{+0.00095}_{-0.00095}$	$1.00^{+0.03}_{-0.03}$
$\mu+$jets	$0.02009^{+0.00083}_{-0.00083}$	$1.00^{+0.03}_{-0.03}$
$l+$jets	$0.01976^{+0.00068}_{-0.00068}$	$1.00^{+0.02}_{-0.02}$

Postfit Data:

Full RunII	$\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$	nonPrompt SF	$t\bar{t}$ cross section (pb)
$e+$jets	$0.02055^{+0.00099}_{-0.00099}$	$1.02^{+0.00}_{-0.03}$	$873.69^{+22.80}_{-22.80}$
$\mu+$jets	$0.02156^{+0.00068}_{-0.00068}$	$1.03^{+0.00}_{-0.02}$	$870.28^{+15.39}_{-15.39}$
$l+$jets	$0.02203^{+0.00064}_{-0.00064}$	$1.01^{+0.00}_{-0.02}$	$876.95^{+13.64}_{-13.64}$

APPENDIX D

The technical implementation for fitting to ratio ($R = \frac{t\bar{t}\gamma}{t\bar{t}}$ is done combine input datacards by adding the following scripts. The rescaling of the $t\bar{t}\gamma$ and $t\bar{t}$ templates is done by scaling with $1/(\sigma^{\text{theory},pb}_{t\bar{t}\gamma})$ and $1/(\sigma^{\text{theory},pb}_{t\bar{t}})$, respectively, which are frozen to their particular values. A rate parameter used to scale both the $t\bar{t}\gamma$ and $t\bar{t}$ templates.

```
renormTTGamma rateParam * *TTGamma* 0.06292616746918506
nuisance edit freeze renormTTGamma ifexists
renormTTbar   rateParam * *TTbar*   0.0012022698855439068
nuisance edit freeze renormTTbar   ifexists
TTbar_SF rateParam *   *TT* 831.76
```

APPENDIX E

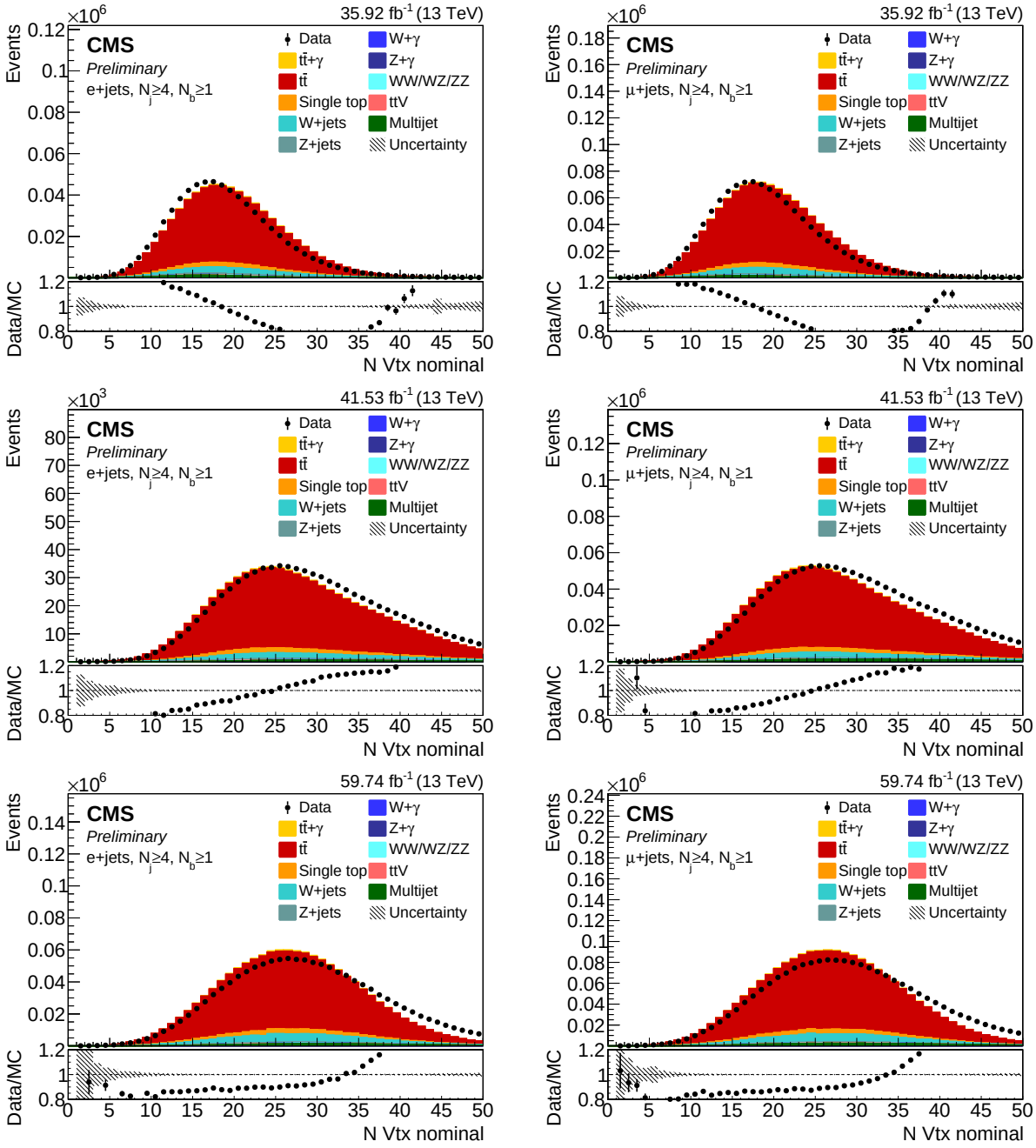


Figure 117: The distribution of the number of vertices in 0 γ control region with $N_j \geq 4, N_b \geq 1$ in $e+$jets and $\mu+$jets for 2016(top), 2017(middle), and 2018(bottom).

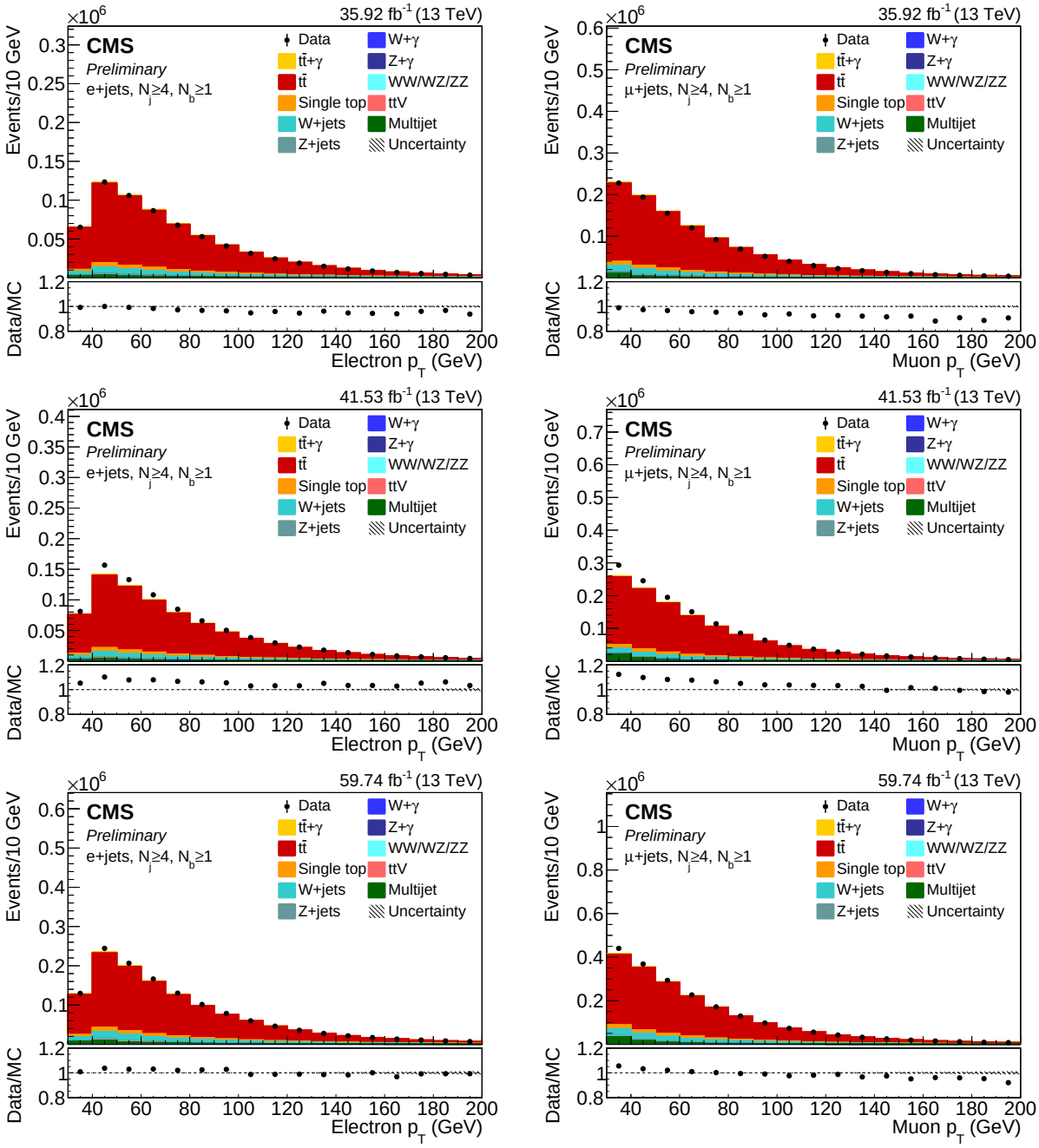


Figure 118: The electron(L) and muon(R) P_T distribution in 0 γ control region with $N_j \geq 4, N_b \geq 1$ in $e+$jets and $\mu+$jets for 2016(top), 2017(middle), and 2018(bottom).

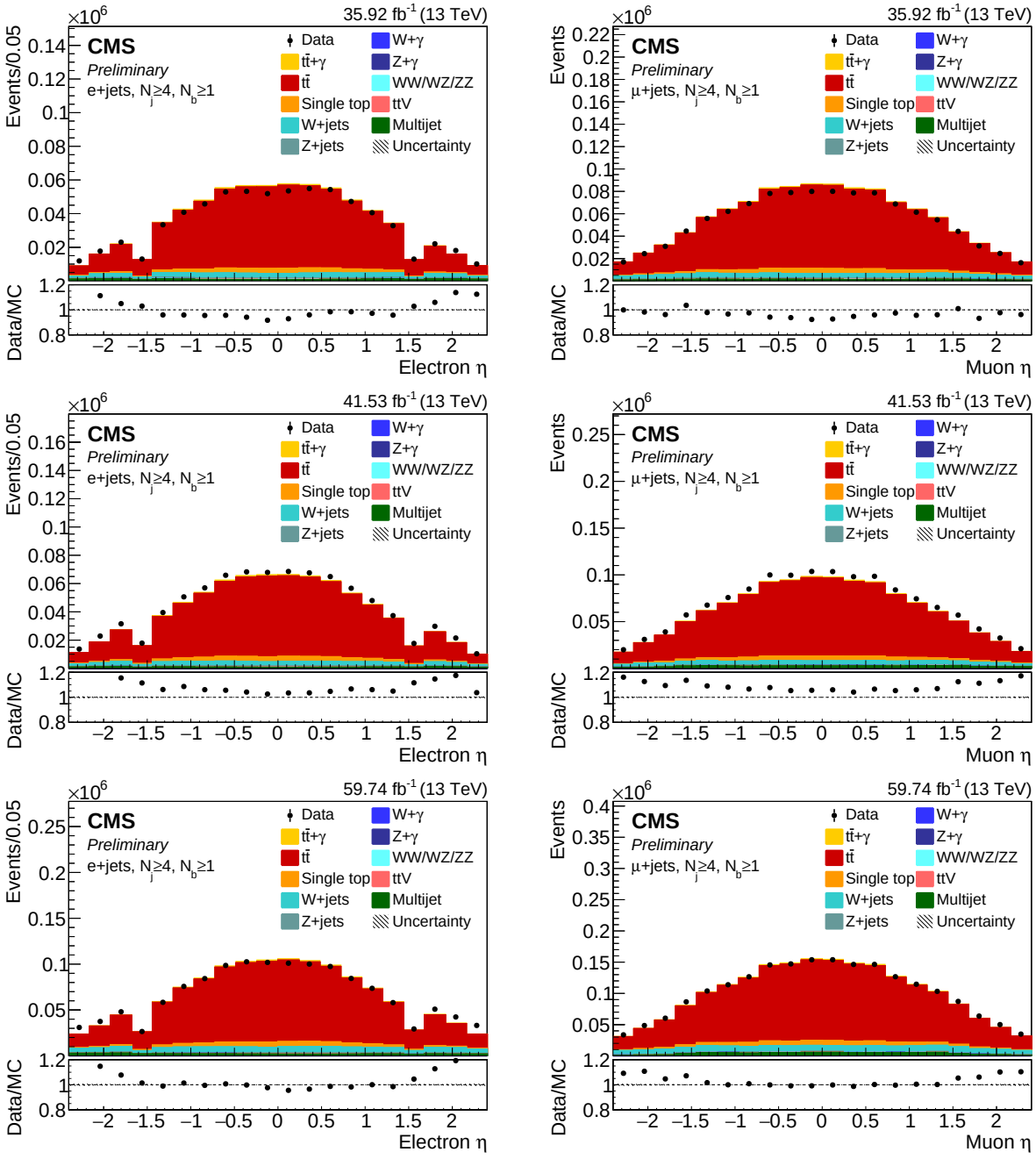


Figure 119: The electron(L) and muon(R) η distribution in 0 γ control region with $N_j \geq 4, N_b \geq 1$ in $e+$jets and $\mu+$jets for 2016(top), 2017(middle), and 2018(bottom).

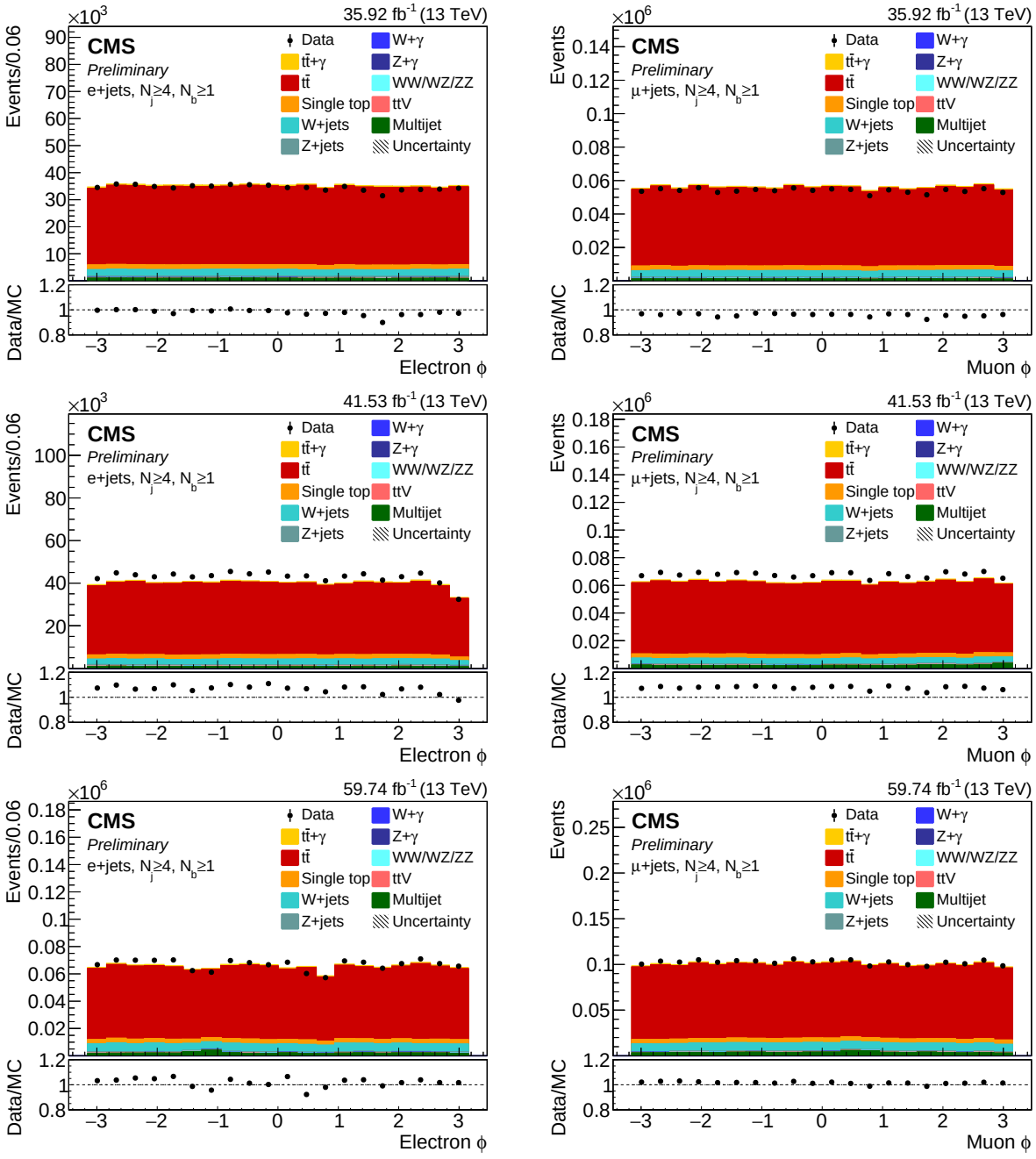


Figure 120: The electron(L) and muon(R) ϕ distribution in 0 γ control region with $N_j \geq 4, N_b \geq 1$ in $e+$jets and $\mu+$jets for 2016(top), 2017(middle), and 2018(bottom).

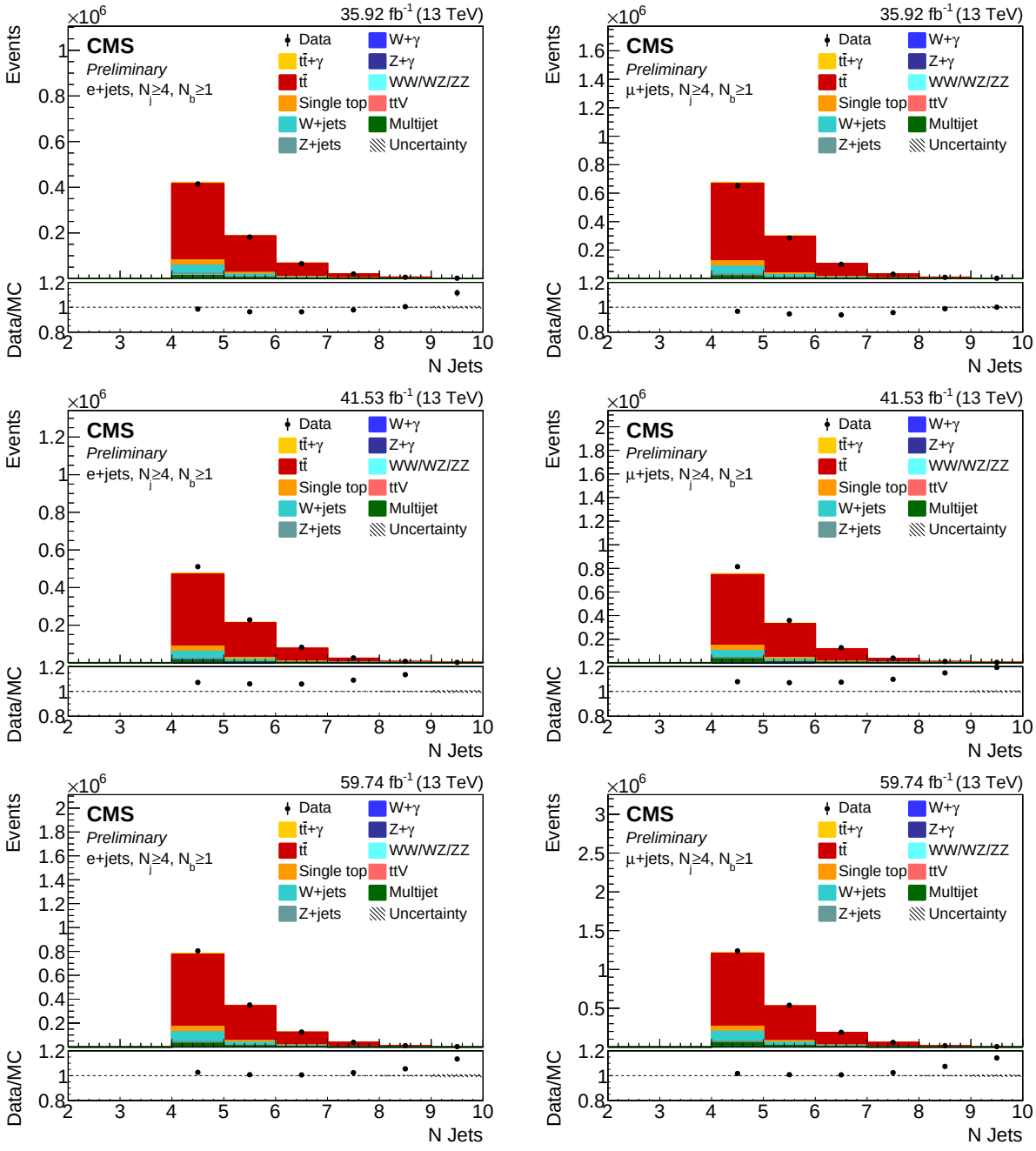


Figure 121: The number of jets distribution in 0 γ control region with $N_j \geq 4, N_b \geq 1$ in $e+$jets and $\mu+$jets for 2016(top), 2017(middle), and 2018(bottom).

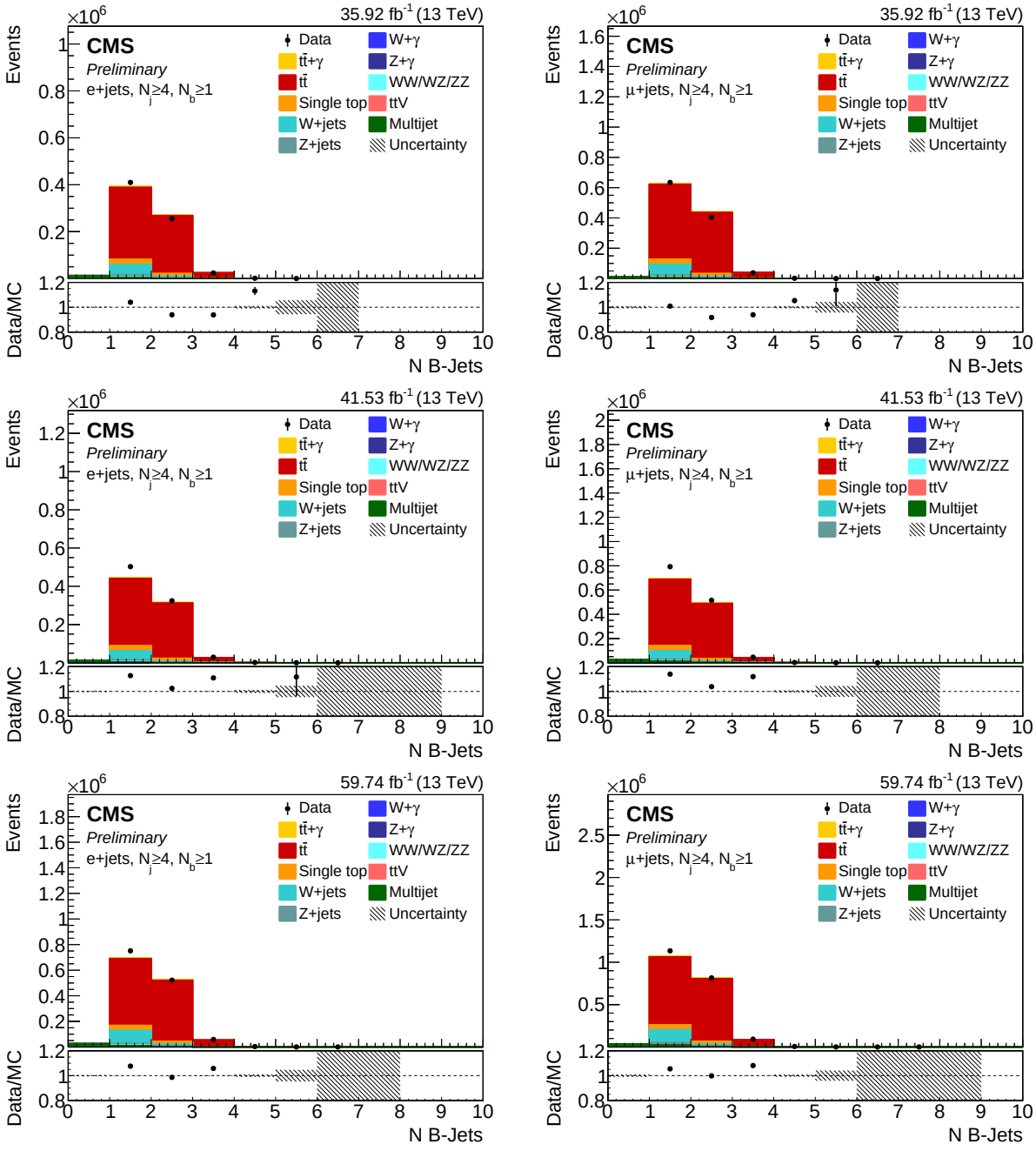


Figure 122: The number of b tagged jets distribution in 0 γ control region with $N_j \geq 4, N_b \geq 1$ in $e+$jets and $\mu+$jets for 2016(top), 2017(middle), and 2018(bottom).

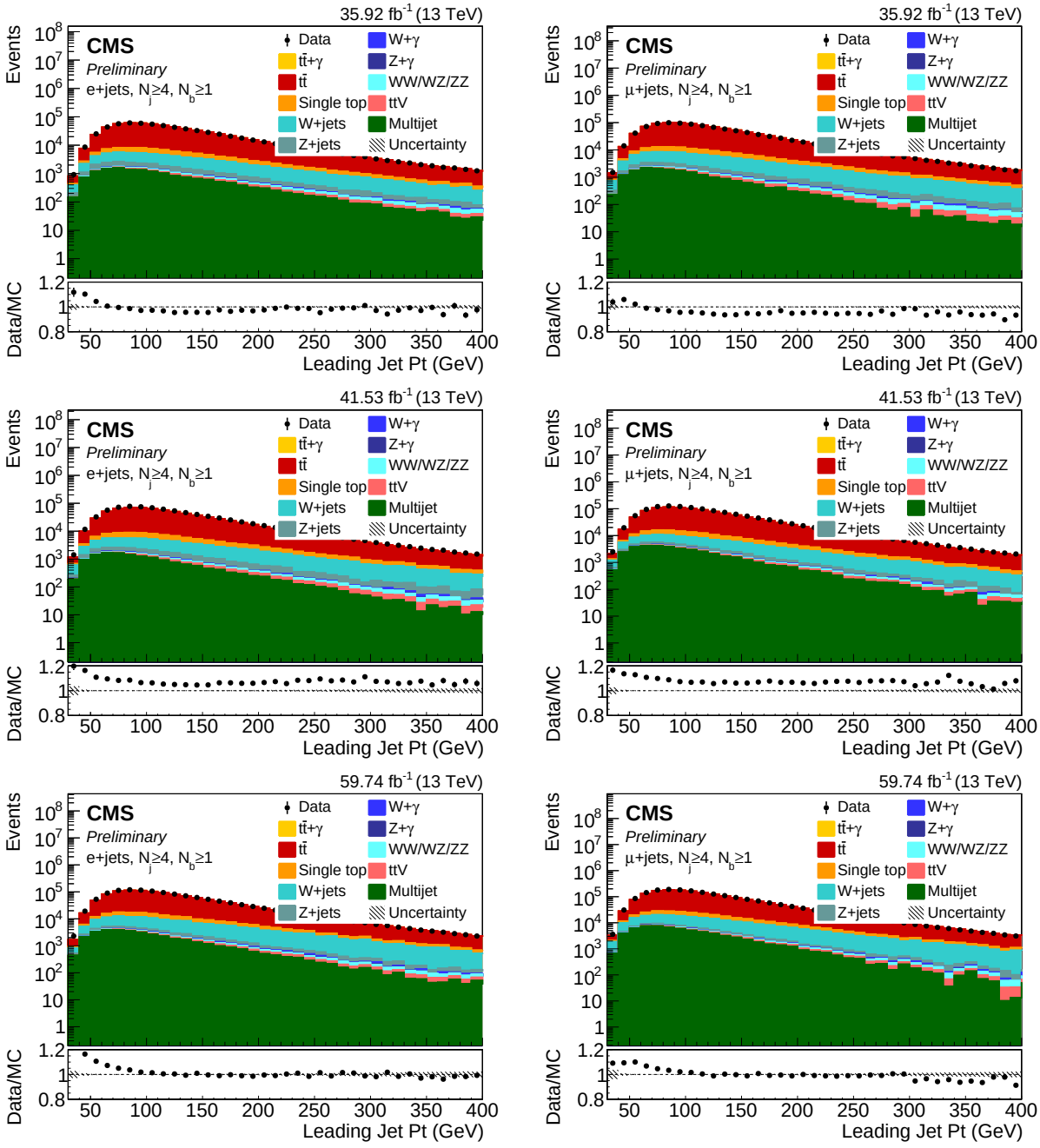


Figure 123: The P_T distribution of the leading jet in 0 γ control region with $N_j \geq 4, N_b \geq 1$ in $e+$jets and $\mu+$jets for 2016(top), 2017(middle), and 2018(bottom).

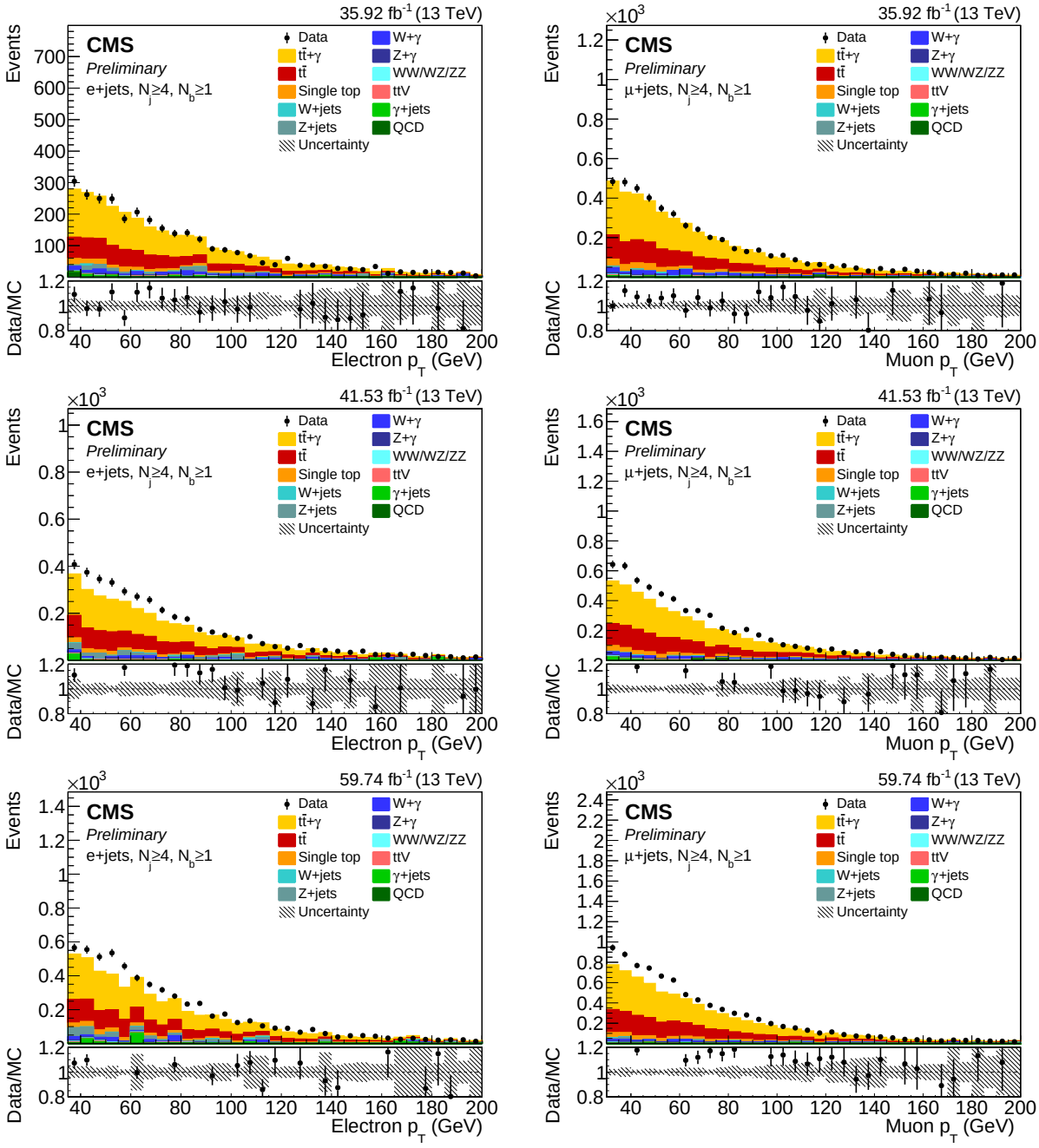


Figure 124: The electron(L) and muon(R) P_T distribution in the signal region ($N_j \geq 4, N_b \geq 1$) in $e+$jets and $\mu+$jets for 2016(top), 2017(middle), and 2018(bottom).

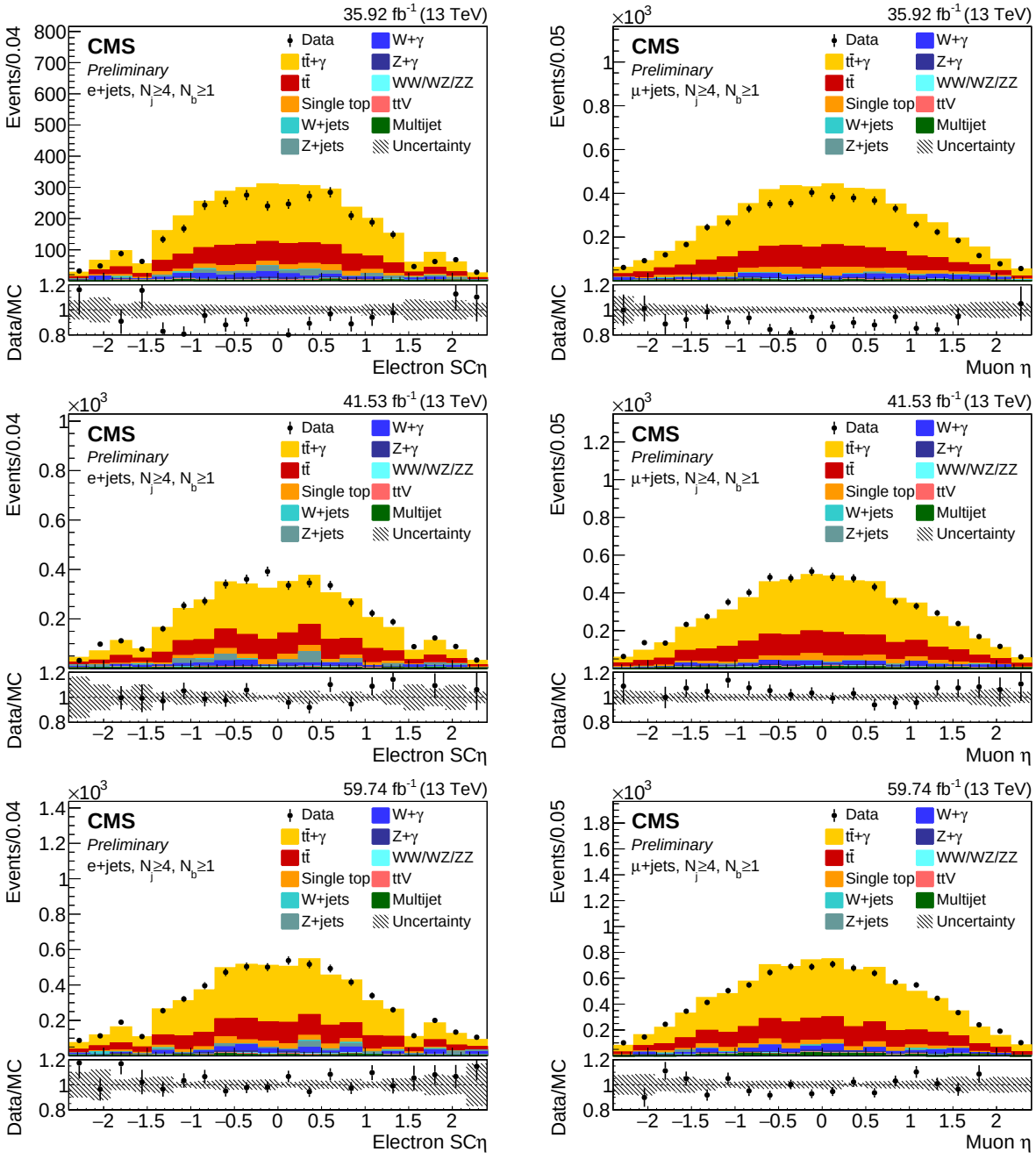


Figure 125: The electron(L) and muon(R) η distribution in the signal region ($N_j \geq 4, N_b \geq 1$) in $e+$jets and $\mu+$jets for 2016(top), 2017(middle), and 2018(bottom).

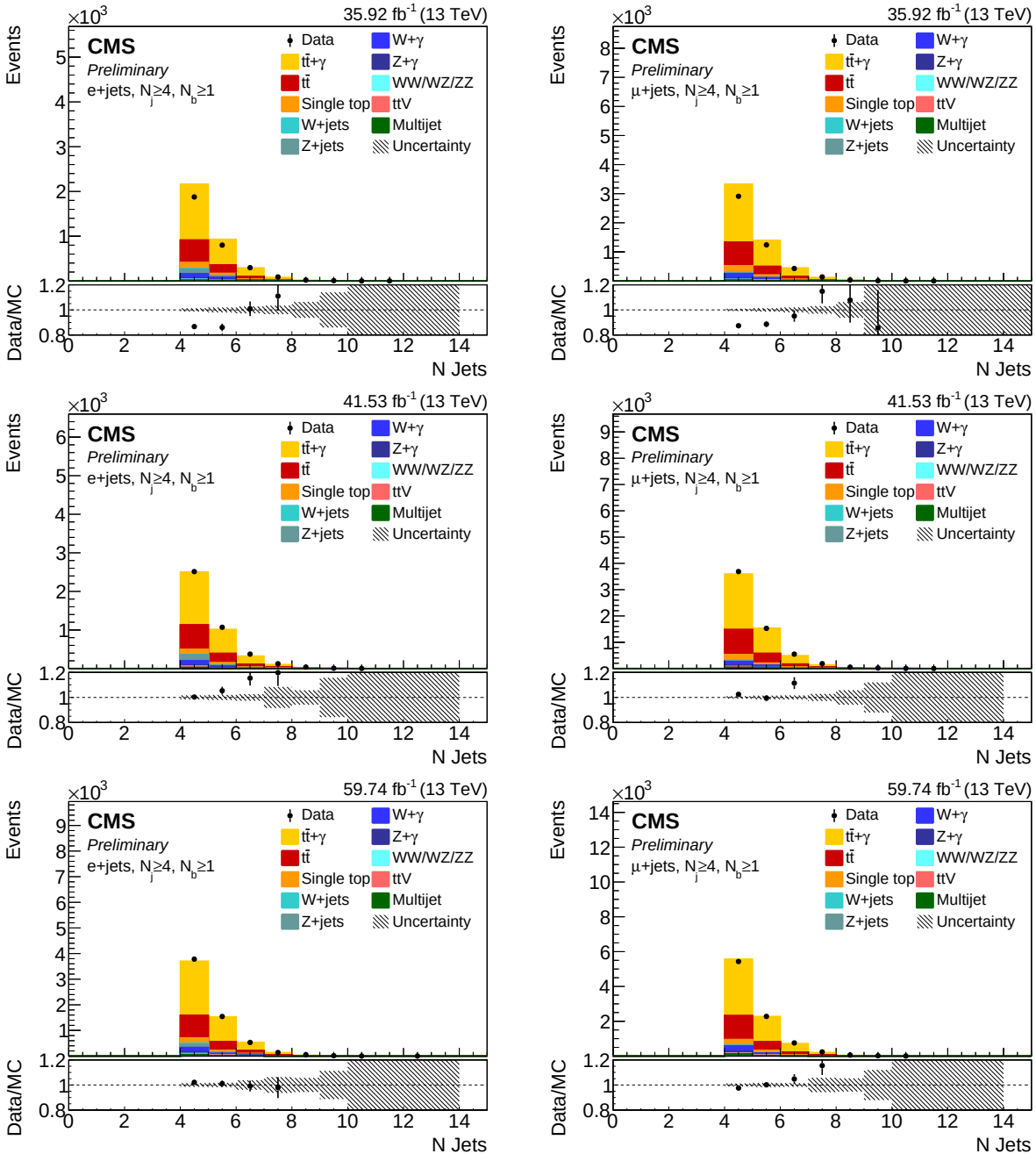


Figure 126: The distribution of number of jets in the signal region ($N_j \geq 4, N_b \geq 1$) in $e+$jets and $\mu+$jets for 2016(top), 2017(middle), and 2018(bottom).

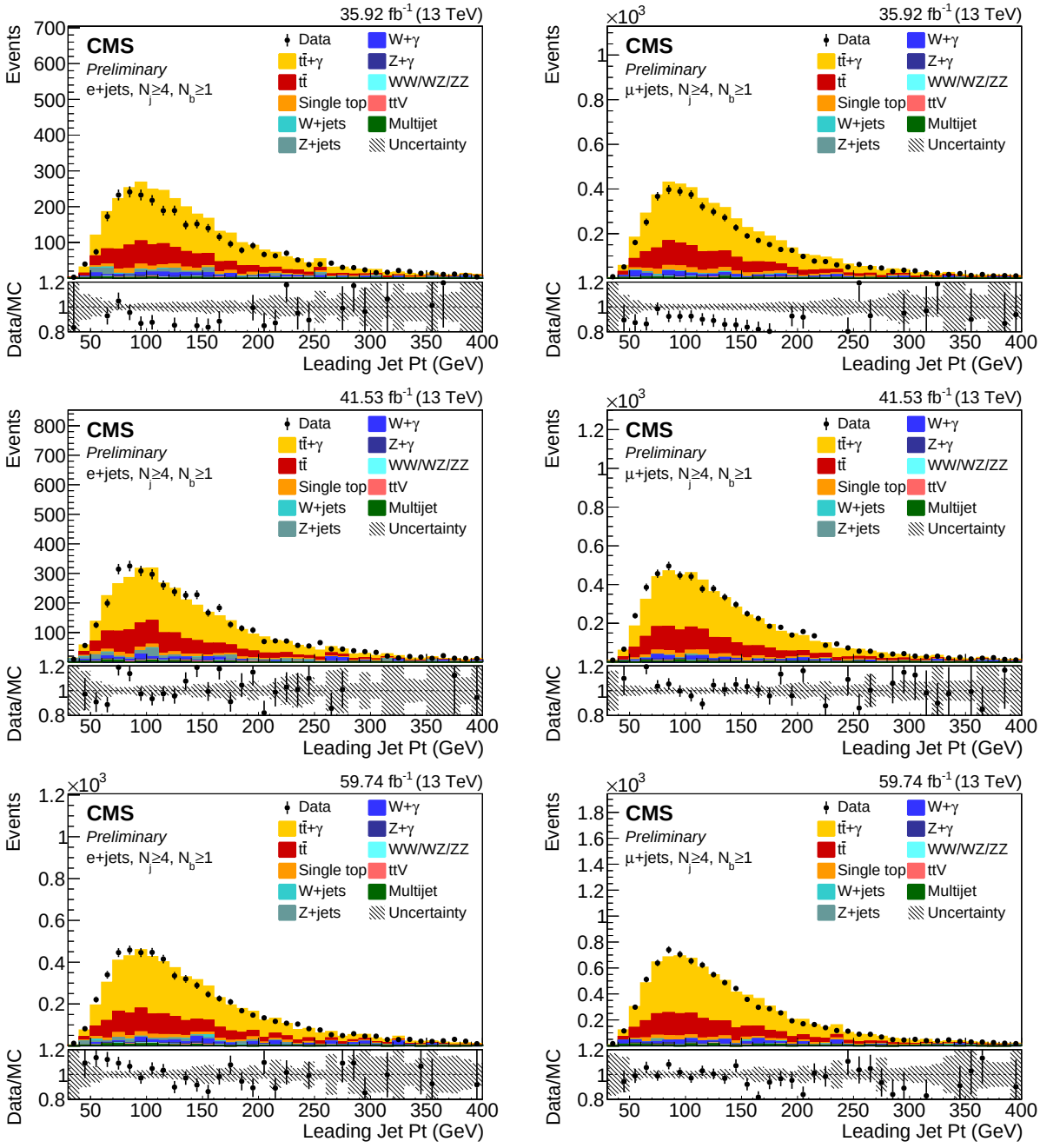


Figure 127: The leading jet P_T distribution in the signal region ($N_j \geq 4, N_b \geq 1$) in $e+$jets and $\mu+$jets for 2016(top), 2017(middle), and 2018(bottom).

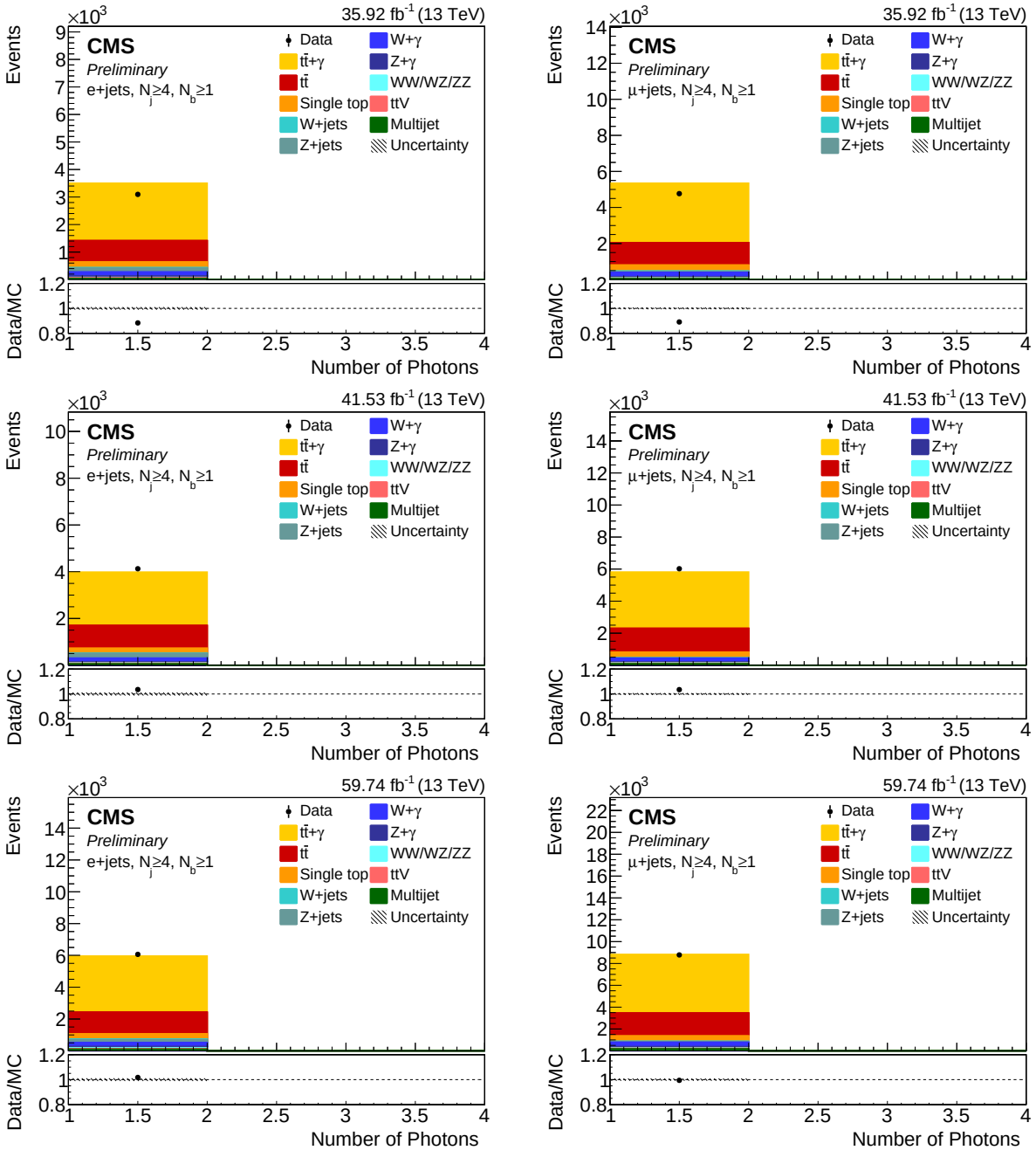


Figure 128: The distribution of number of photons in the signal region ($N_j \geq 4, N_b \geq 1$) in $e+$jets and $\mu+$jets for 2016(top), 2017(middle), and 2018(bottom).

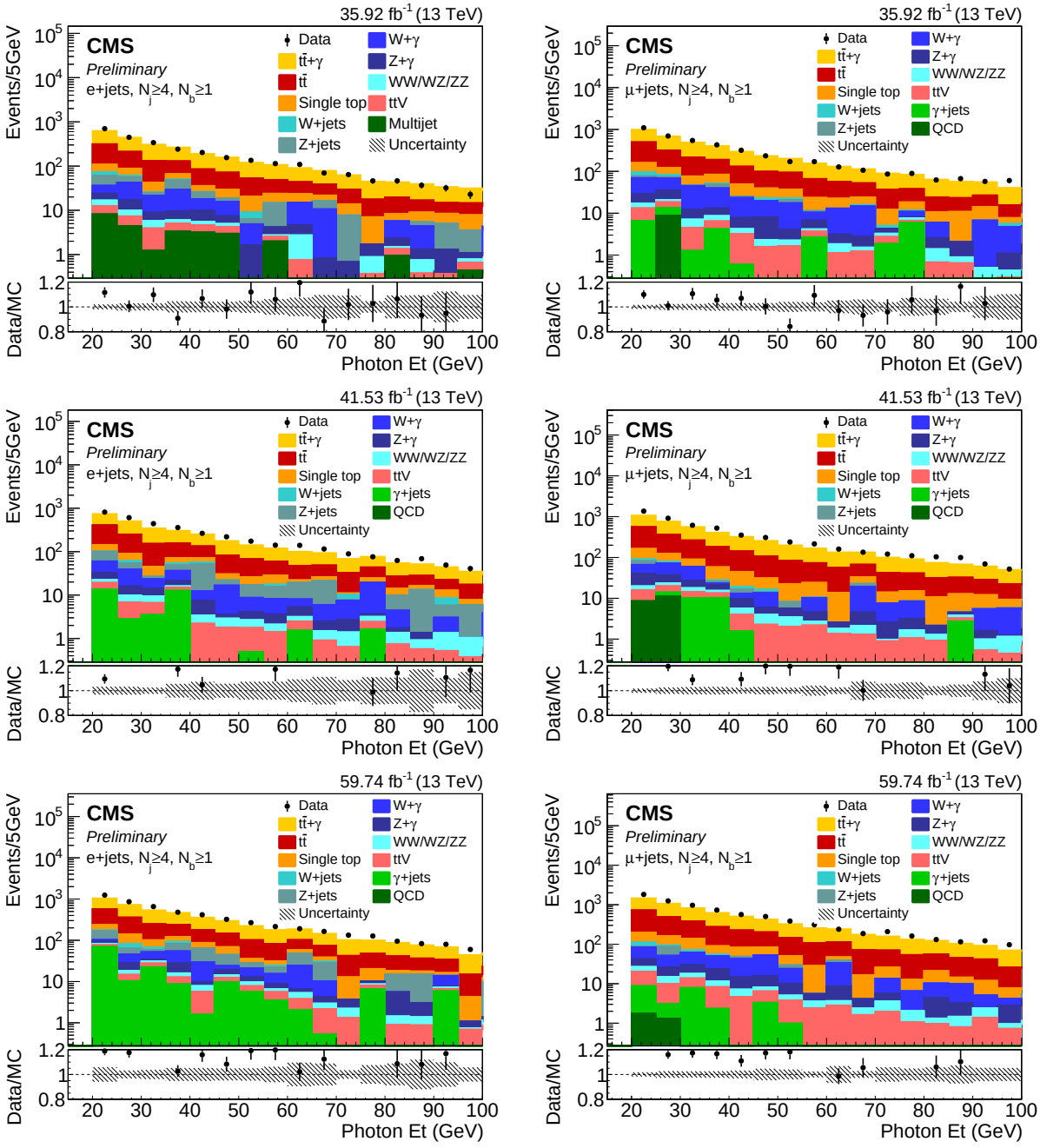


Figure 129: The P_T distribution of the leading photon in the signal region ($N_j \geq 4, N_b \geq 1$) in $e+$jets and $\mu+$jets for 2016(top), 2017(middle), and 2018(bottom).

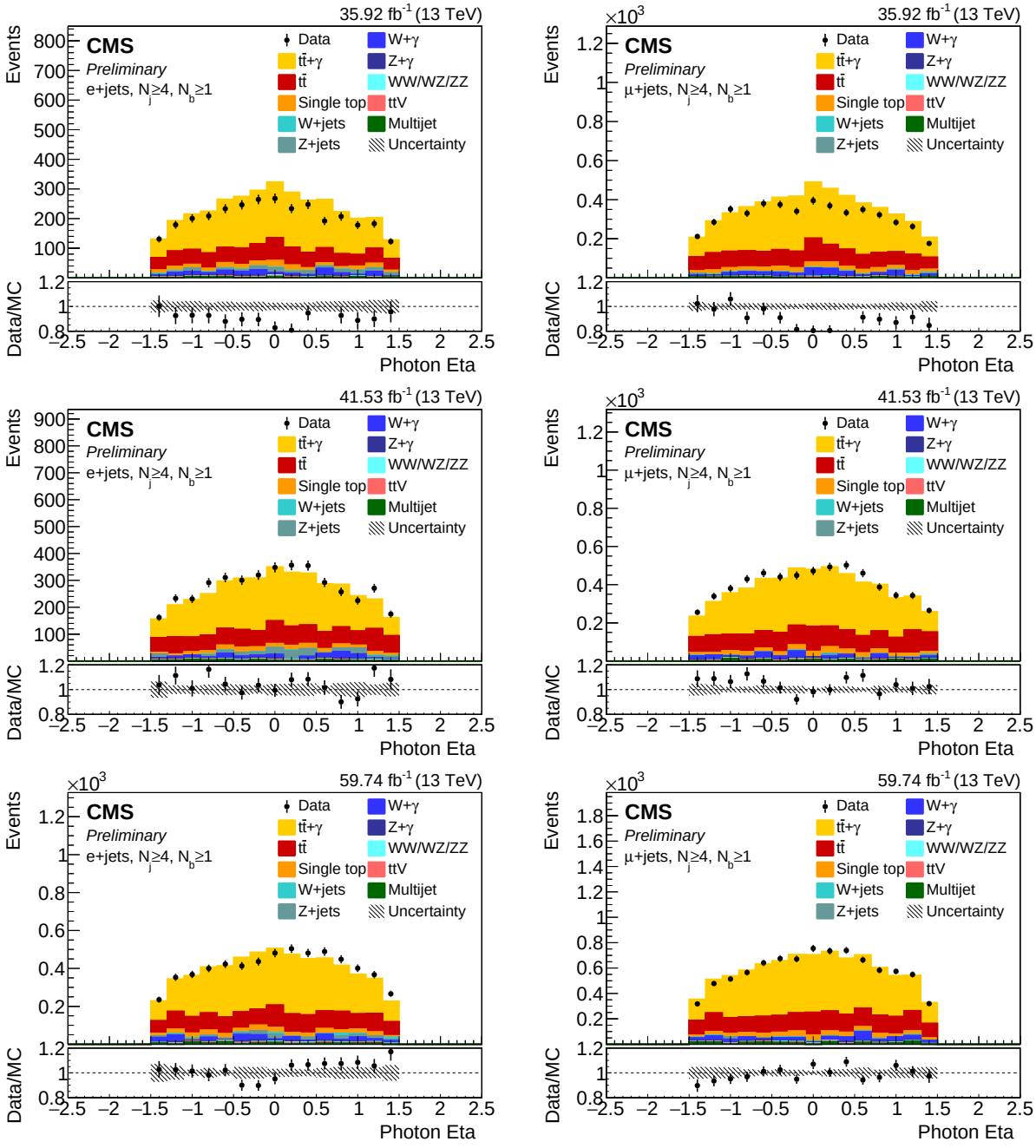


Figure 130: The η distribution of the leading photon in the signal region ($N_j \geq 4, N_b \geq 1$) in $e+$jets and $\mu+$jets for 2016(top), 2017(middle), and 2018(bottom).

APPENDIX F

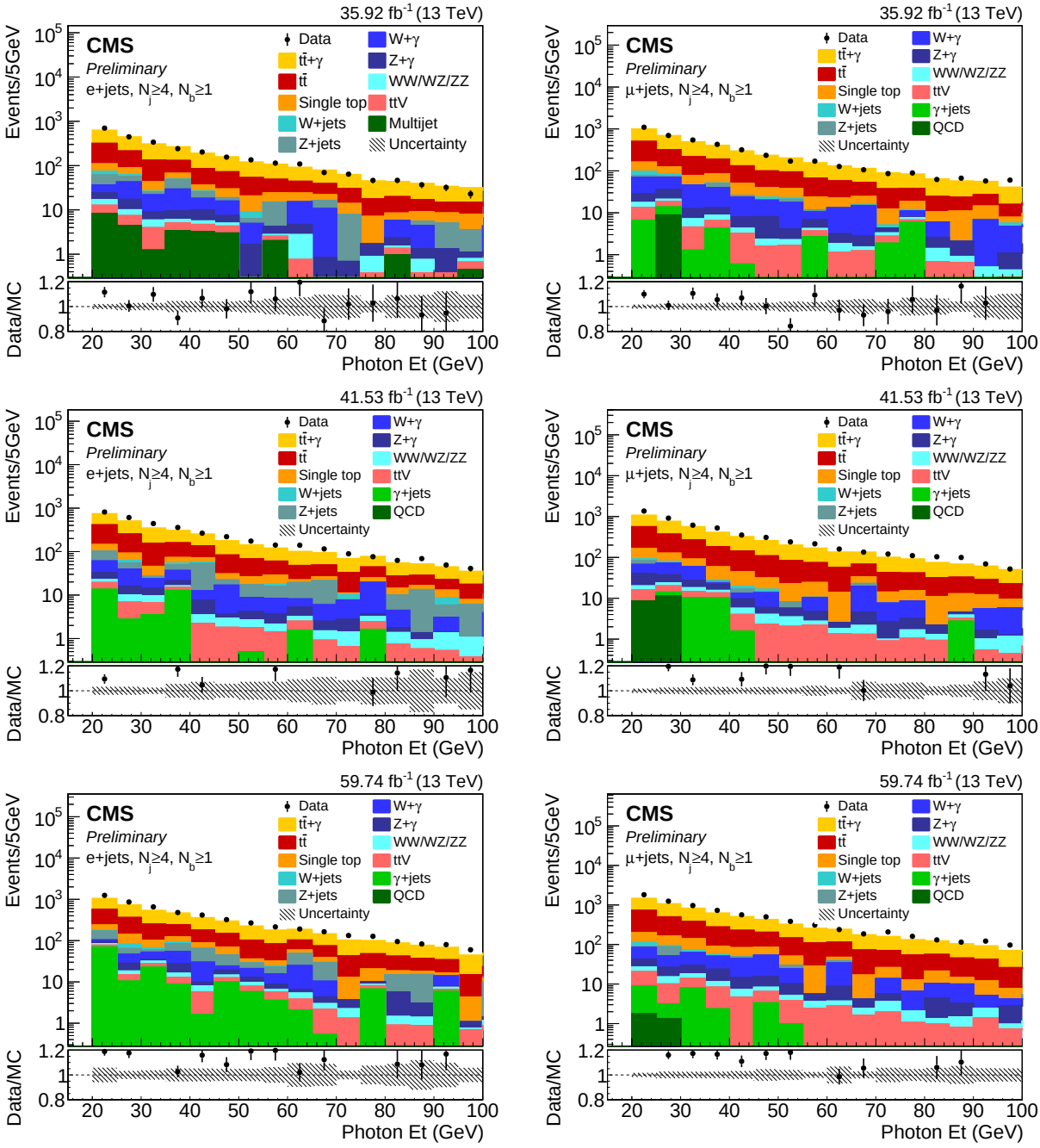


Figure 131: The distribution of the photon P_T in the signal region ($N_j \geq 4, N_b \geq 1$) split by samples in $e+$ jets (L) and $\mu+$ jets (R) for 2016(top), 2017(middle), and 2018(bottom).

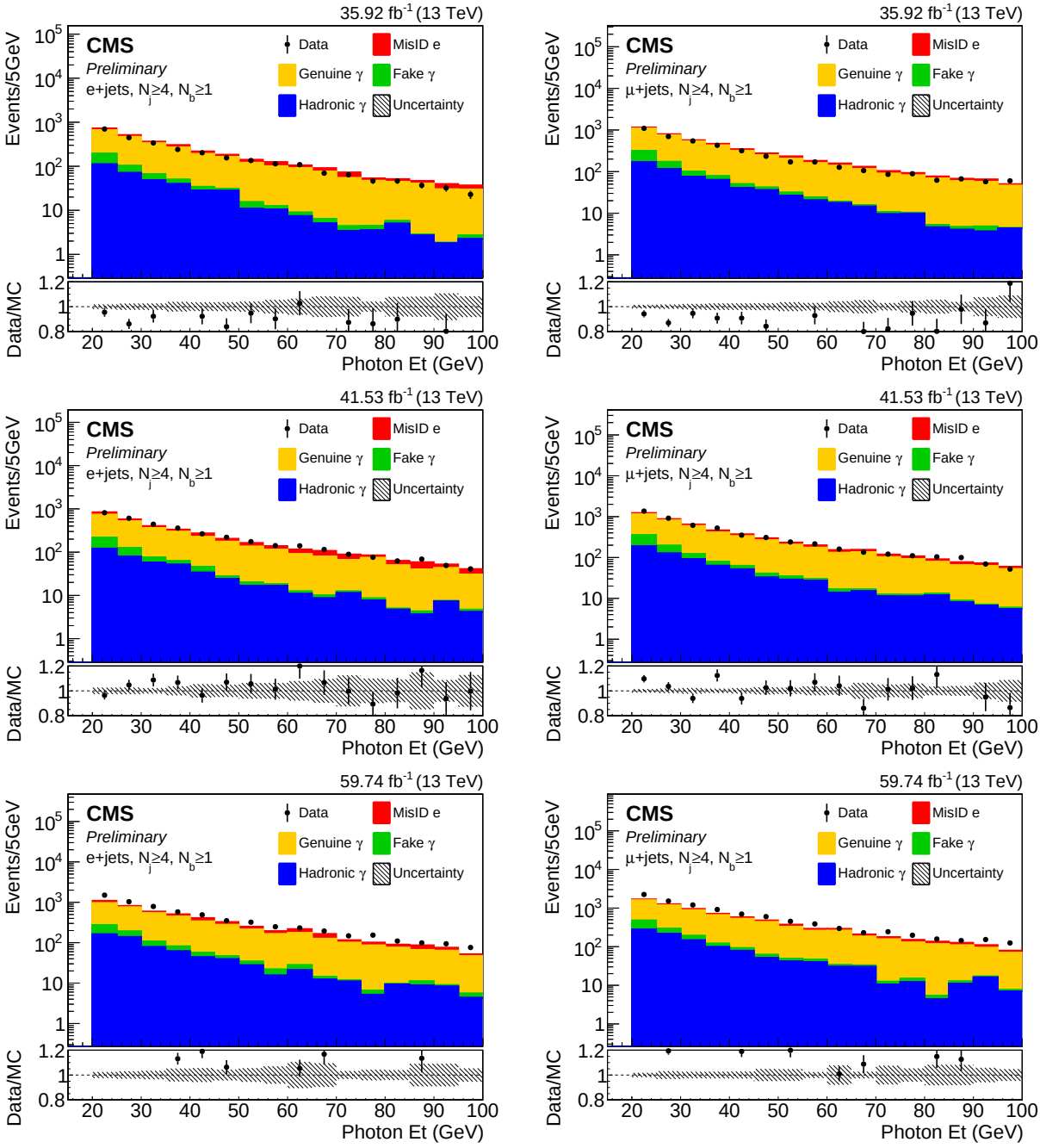


Figure 132: The distribution of the photon P_T in the signal region ($N_j \geq 4, N_b \geq 1$) split by photon category in $e+$jets (L) and $\mu+$jets (R) for 2016(top), 2017(middle), and 2018(bottom).

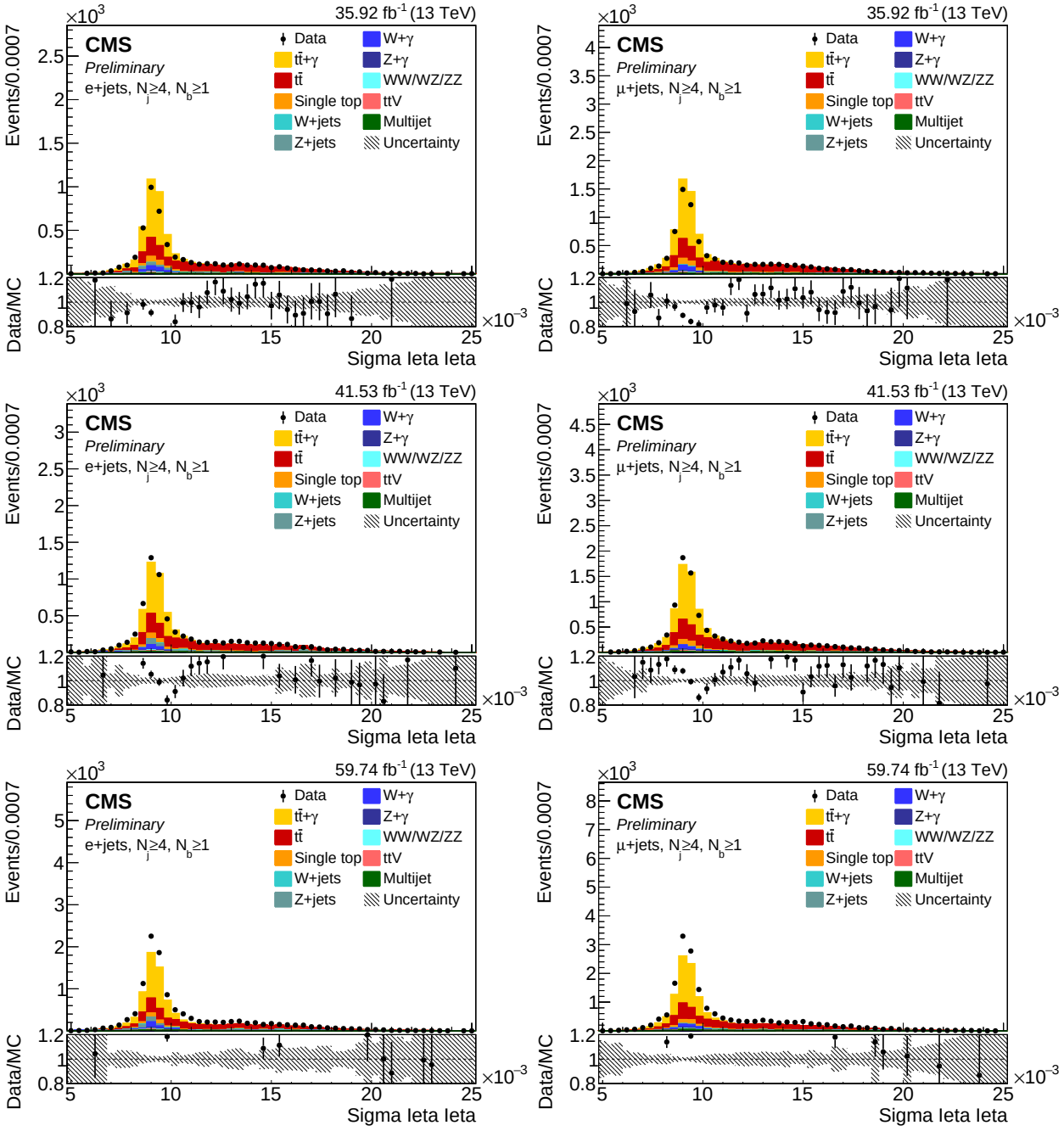


Figure 133: The distribution of the $\sigma_{i\eta i\eta}$ with no cut in the signal region ($N_j \geq 4, N_b \geq 1$) in $e+$jets (L) and $\mu+$jets (R) for 2016(top), 2017(middle), and 2018(bottom). The sideband region is defined as $0.011 \leq \sigma_{i\eta i\eta} \leq 0.021$.

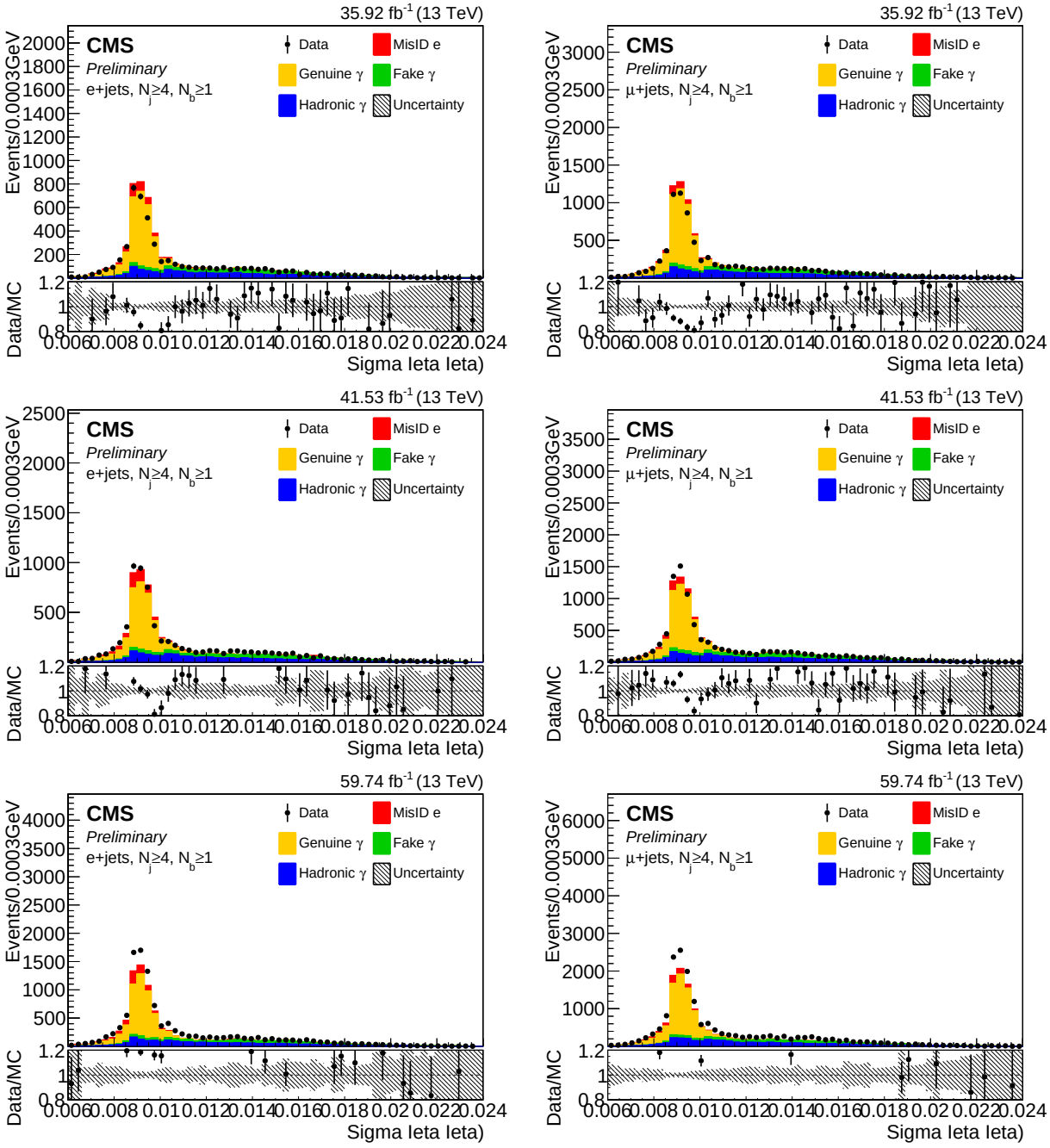


Figure 134: The distribution of the $\sigma_{i\eta i\eta}$ with no cut in the signal region ($N_j \geq 4, N_b \geq 1$) split by photon category in $e+$jets (L) and $\mu+$jets (R) for 2016(top), 2017(middle), and 2018(bottom). The sideband region is defined as $0.011 \leq \sigma_{i\eta i\eta} \leq 0.021$.

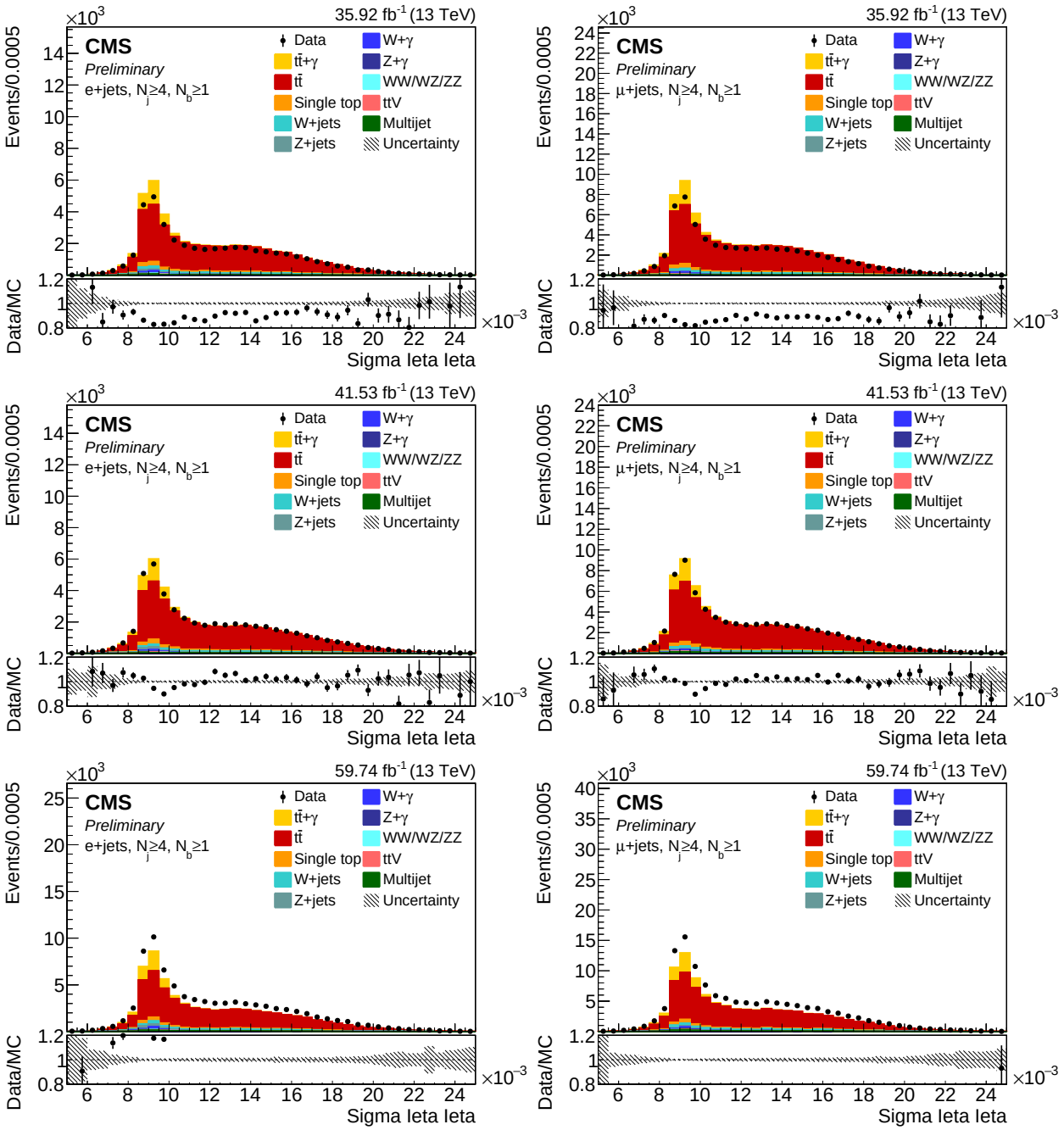


Figure 135: The distribution of the $\sigma_{i\eta i\eta}$ with no cut in Charge hadron isolation in the signal region ($N_j \geq 4, N_b \geq 1$) in $e+$jets (L) and $\mu+$jets (R) for 2016(top), 2017(middle), and 2018(bottom). The sideband region is defined as $0.011 \leq \sigma_{i\eta i\eta} \leq 0.021$.

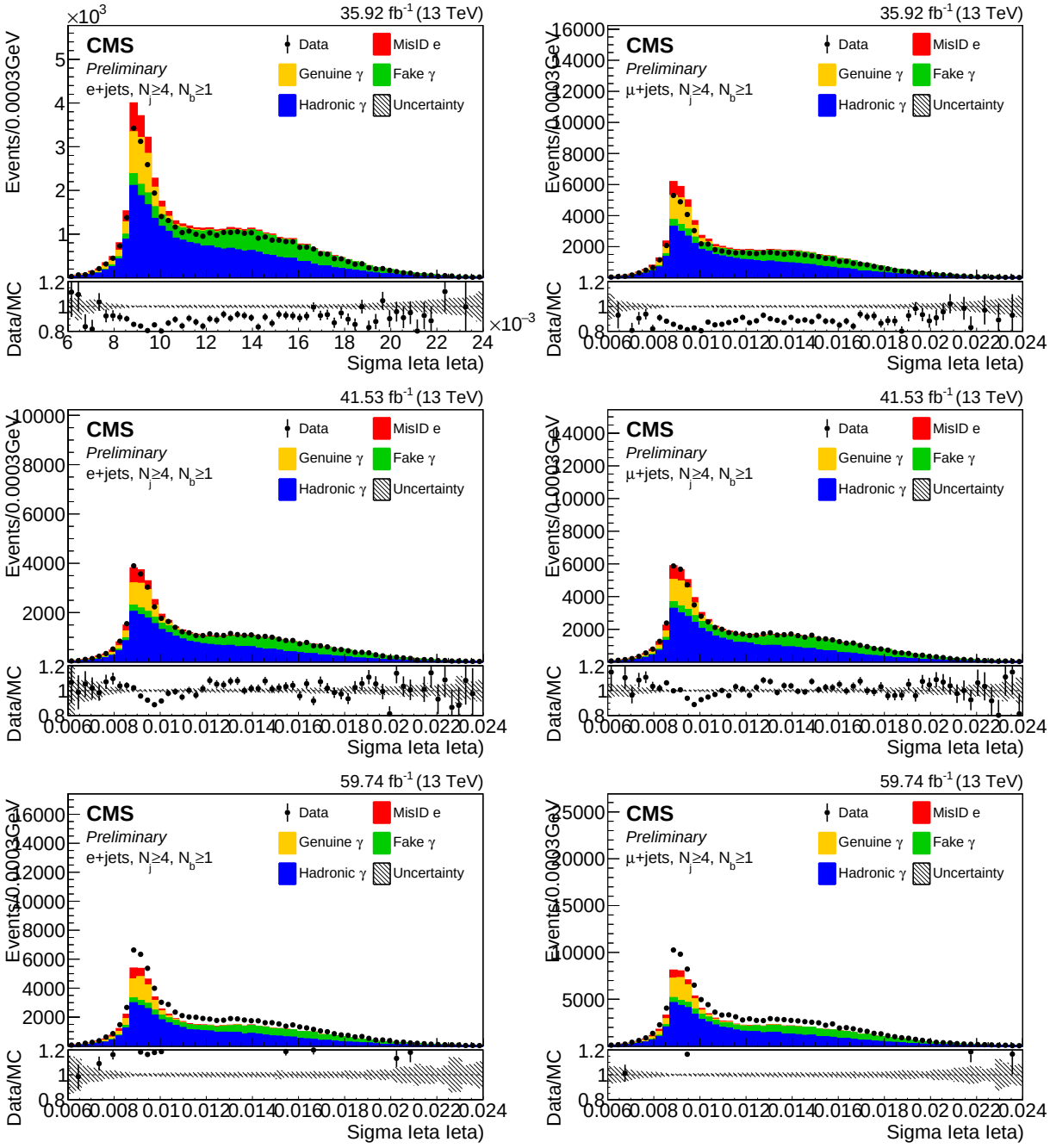


Figure 136: The distribution of the $\sigma_{i\eta i\eta}$ with no cut in Charge hadron isolation in the signal region ($N_j \geq 4, N_b \geq 1$) split by photon category in $e+$jets (L) and $\mu+$jets (R) for 2016(top), 2017(middle), and 2018(bottom). The sideband region is defined as $0.011 \leq \sigma_{i\eta i\eta} \leq 0.021$.

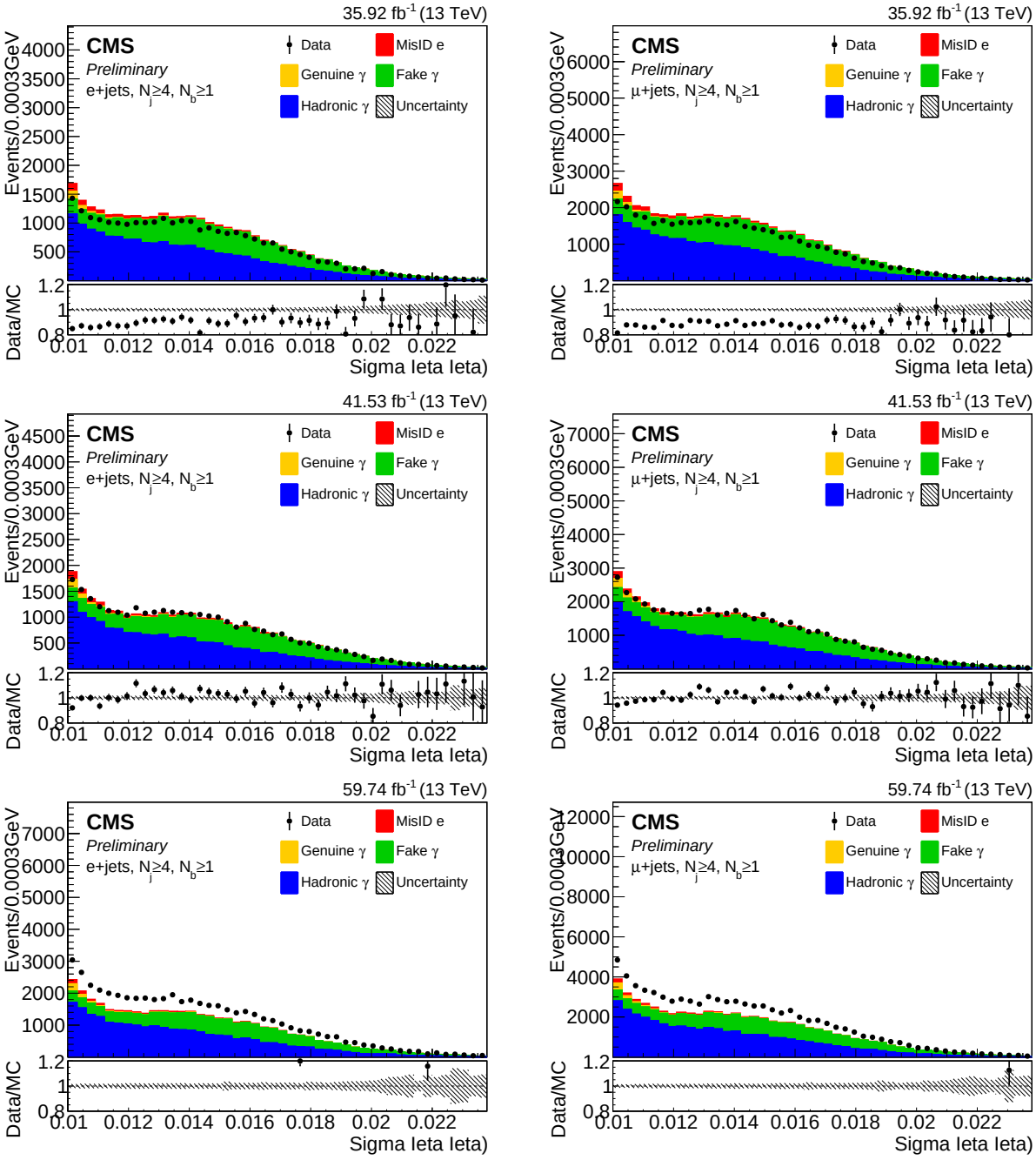


Figure 137: The distribution of the focused $\sigma_{i\eta i\eta}$ with no cut in Charge hadron isolation in the signal region ($N_j \geq 4, N_b \geq 1$) split by photon category in $e+$jets (L) and $\mu+$jets (R) for 2016(top), 2017(middle), and 2018(bottom). The sideband region is defined as $0.011 \leq \sigma_{i\eta i\eta} \leq 0.021$.

APPENDIX G

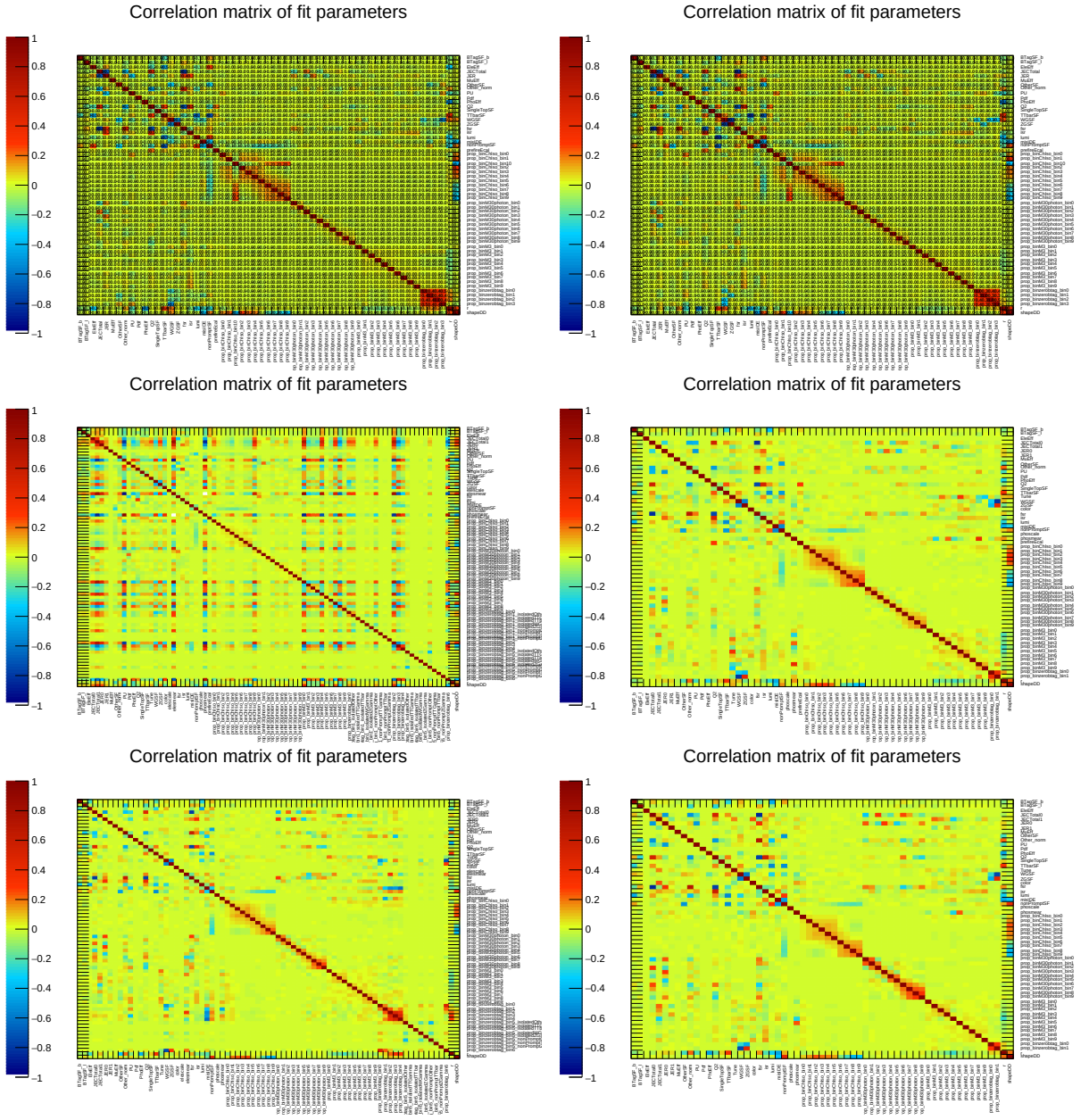


Figure 138: Double Negative log likelihood distribution for $t\bar{t}\gamma$ in $e+$jets (L) and $\mu+$jets (R) channels with Asimov fit 2016 (top), 2017 (middle), and 2018 (bottom) data taking periods.

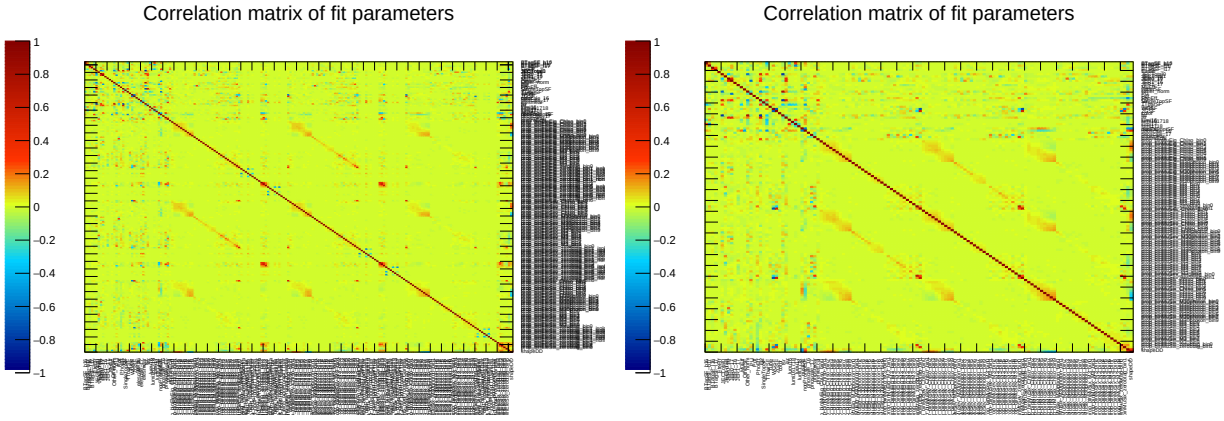


Figure 139: Double Negative log likelihood distribution for $t\bar{t}\gamma$ in $e+$jets (L) and $\mu+$jets (R) channels with Asimov fit for full RunII data.

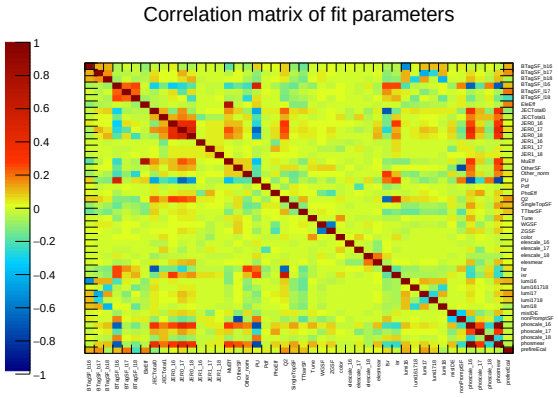


Figure 140: Double Negative log likelihood distribution for $t\bar{t}\gamma$ with Asimov fit for full RunII data.

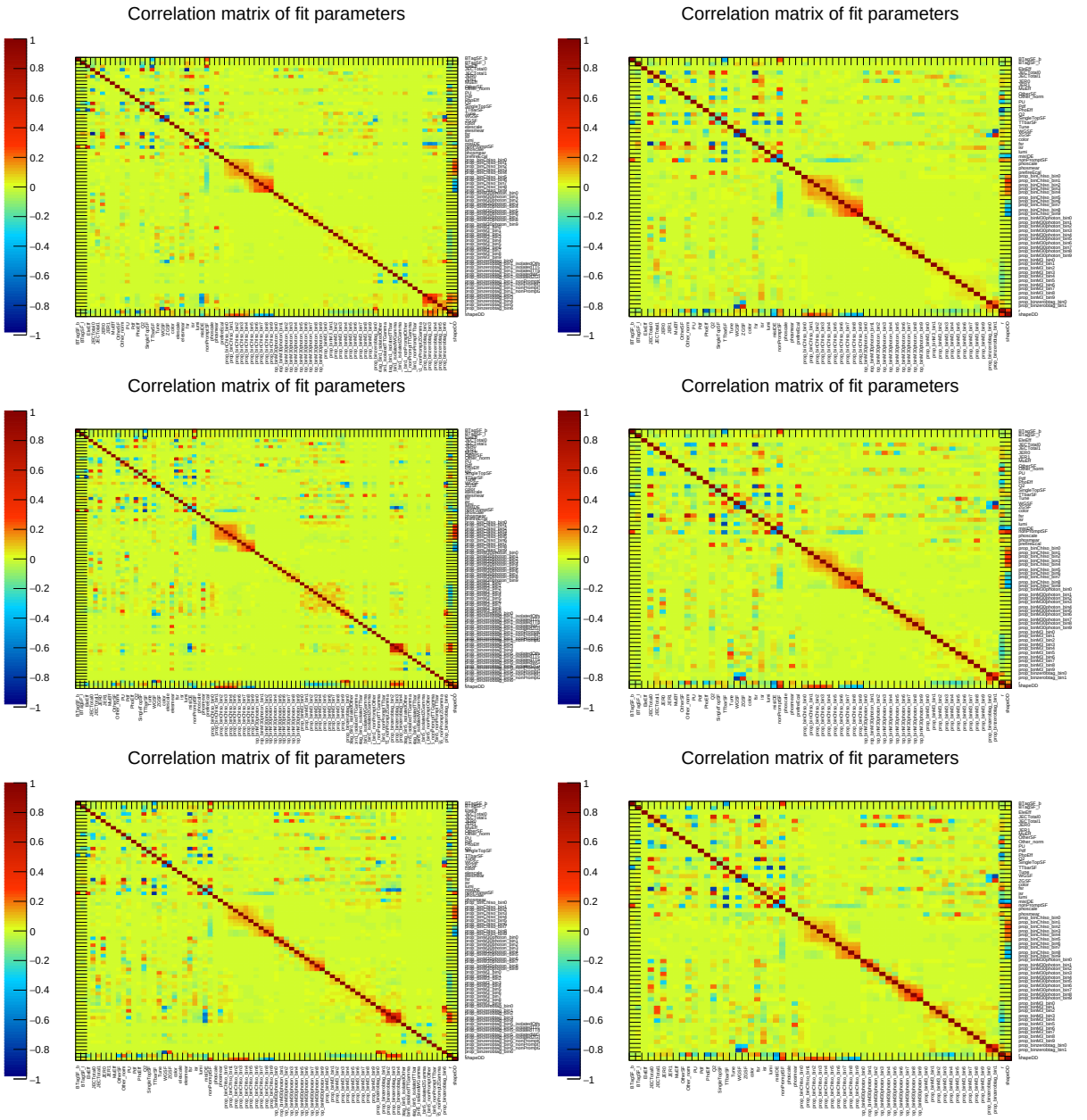


Figure 141: Double Negative log likelihood distribution for Ratio in $e+$jets (L) and $\mu+$jets (R) channels with Asimov fit 2016 (top), 2017 (middle), and 2018 (bottom) data taking periods.

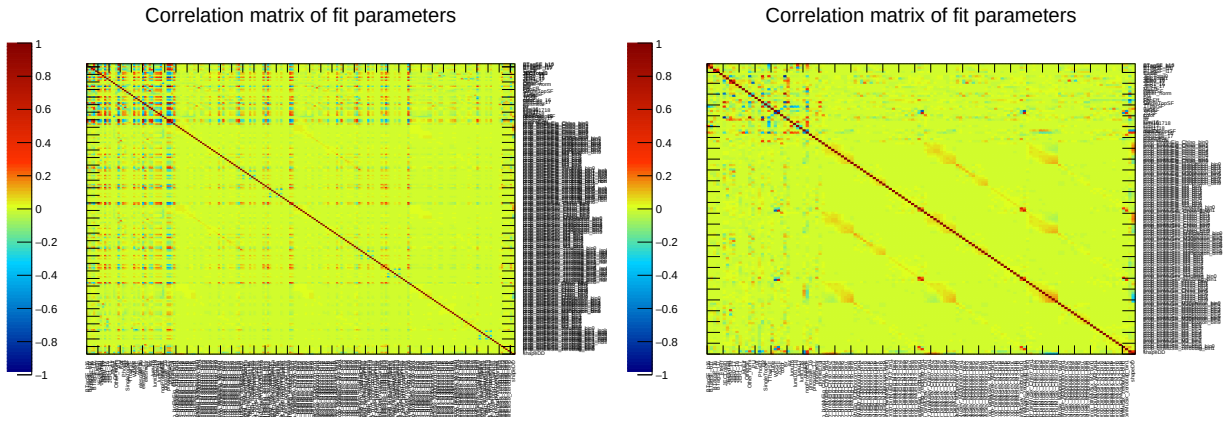


Figure 142: Double Negative log likelihood distribution for Ratio in $e+$jets (L) and $\mu+$jets (R) channels with Asimov fit for full RunII data.

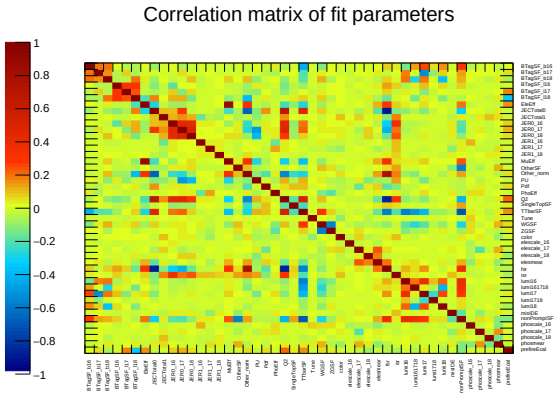


Figure 143: Double Negative log likelihood distribution for Ratio with Asimov fit for full RunII data.

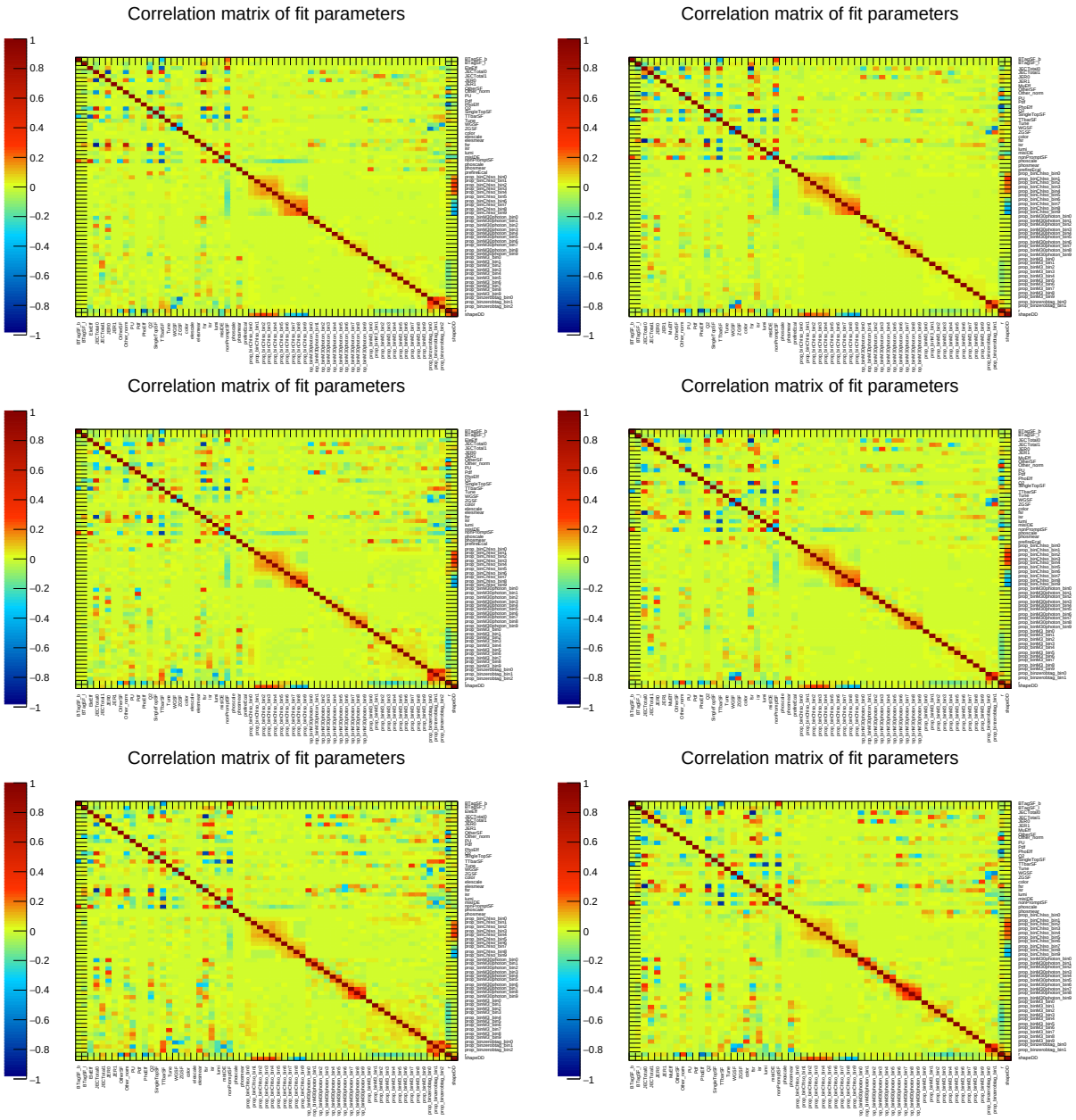


Figure 144: Double Negative log likelihood distribution for Ratio in $e+$jets (L) and $\mu+$jets (R) channels with data fit 2016 (top), 2017 (middle), and 2018 (bottom) data taking periods.

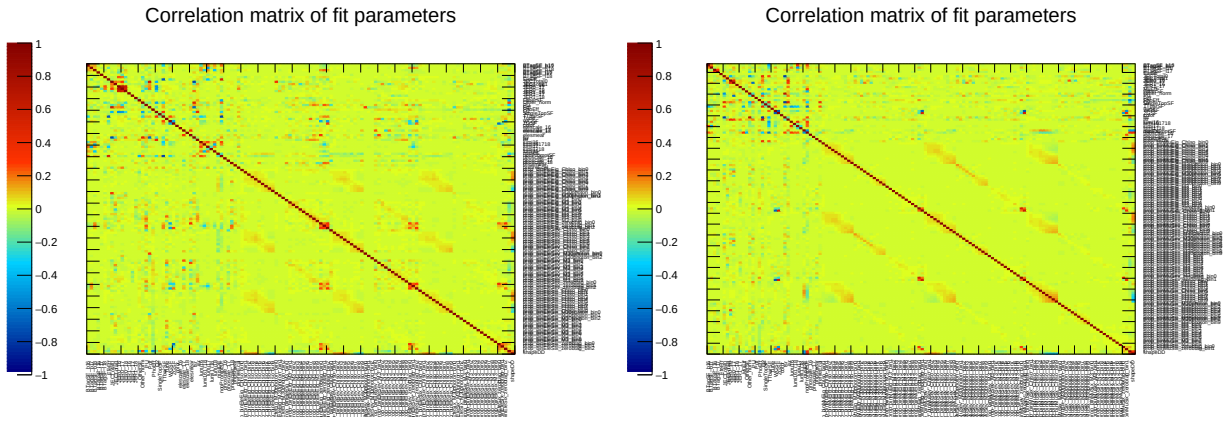


Figure 145: Double Negative log likelihood distribution for Ratio in $e+$jets (L) and $\mu+$jets (R) channels with data fit for full RunII data.

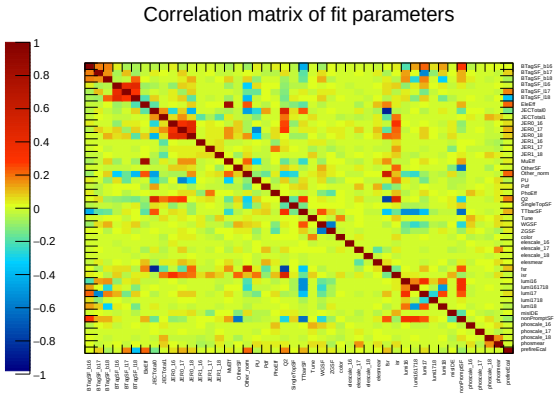


Figure 146: Double Negative log likelihood distribution for Ratio with data fit for full RunII data.

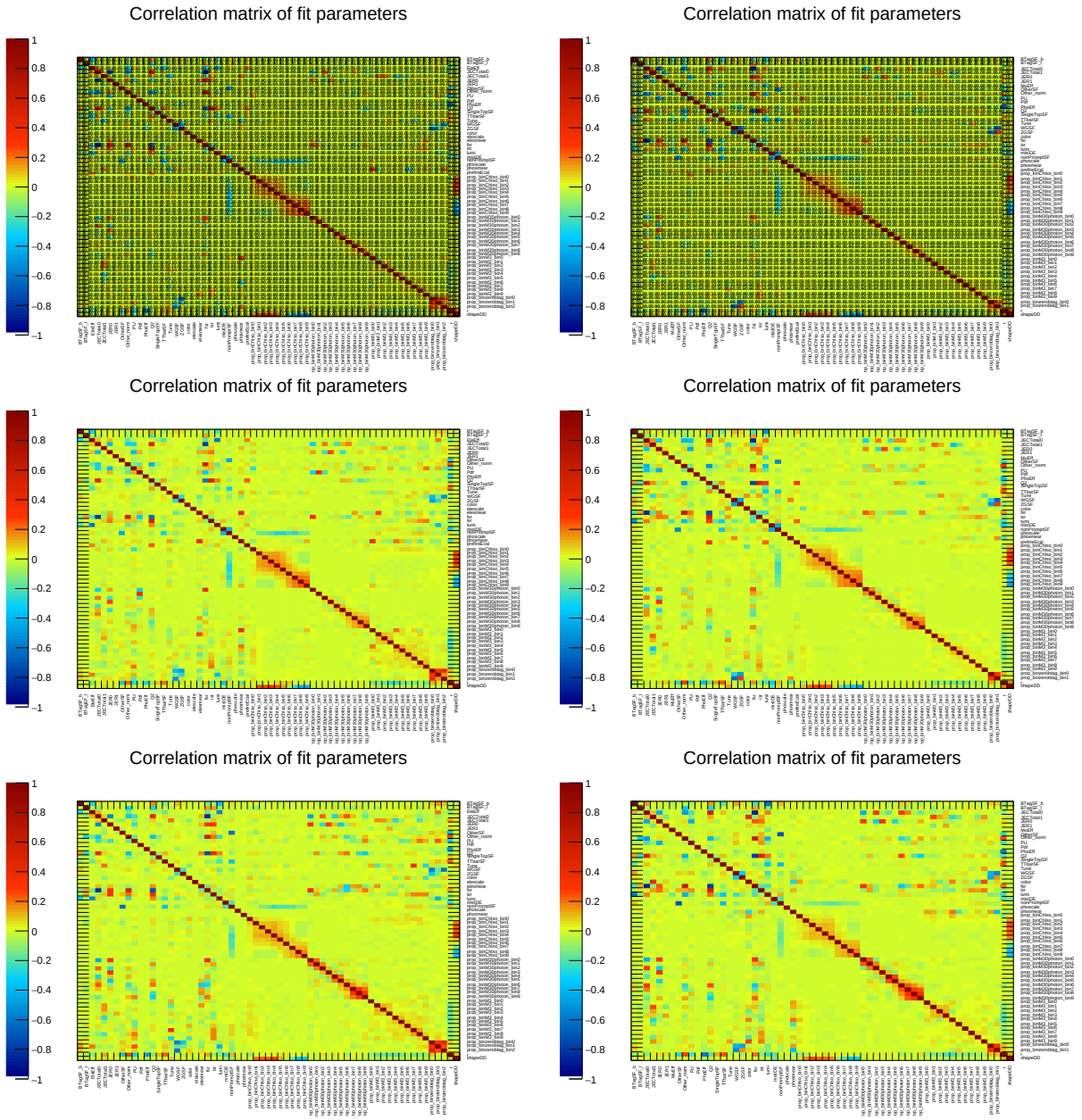


Figure 147: Double Negative log likelihood distribution for $t\bar{t}\gamma$ in $e+$jets (L) and $\mu+$jets (R) channels with data fit 2016 (top), 2017 (middle), and 2018 (bottom) data taking periods.

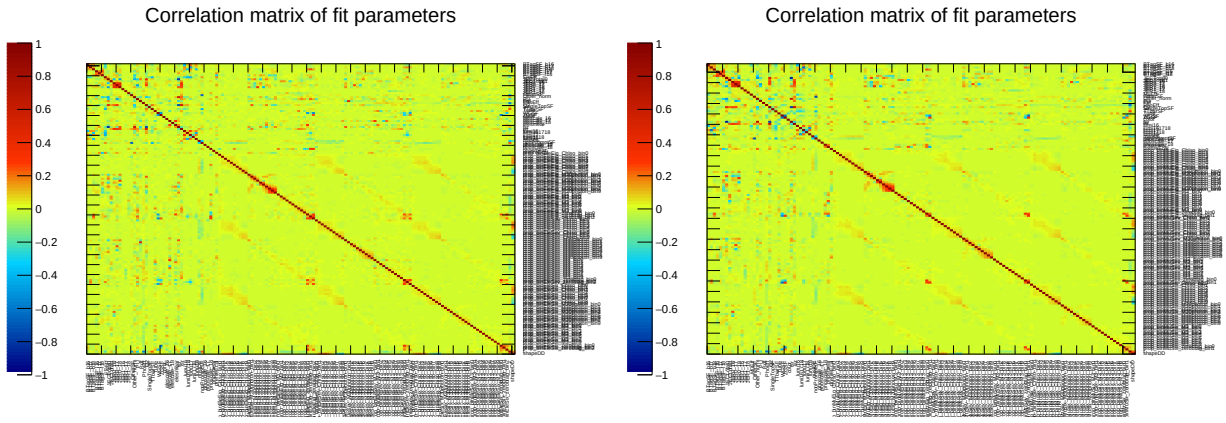


Figure 148: Double Negative log likelihood distribution for $\mathrm{t\bar{t}}\gamma$ in $e+$jets (L) and $\mu+$jets (R) channels with data fit for full RunII data.

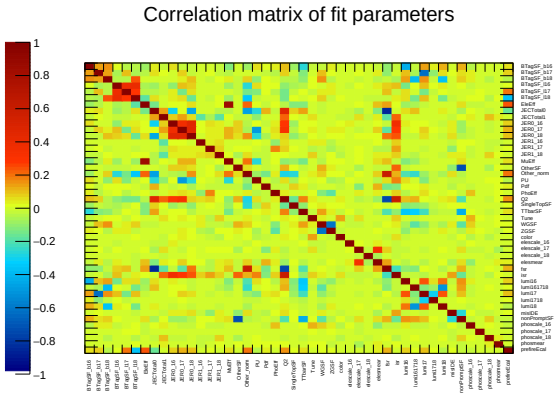


Figure 149: Double Negative log likelihood distribution for $\mathrm{t\bar{t}}\gamma$ with data fit for full RunII data.

APPENDIX H

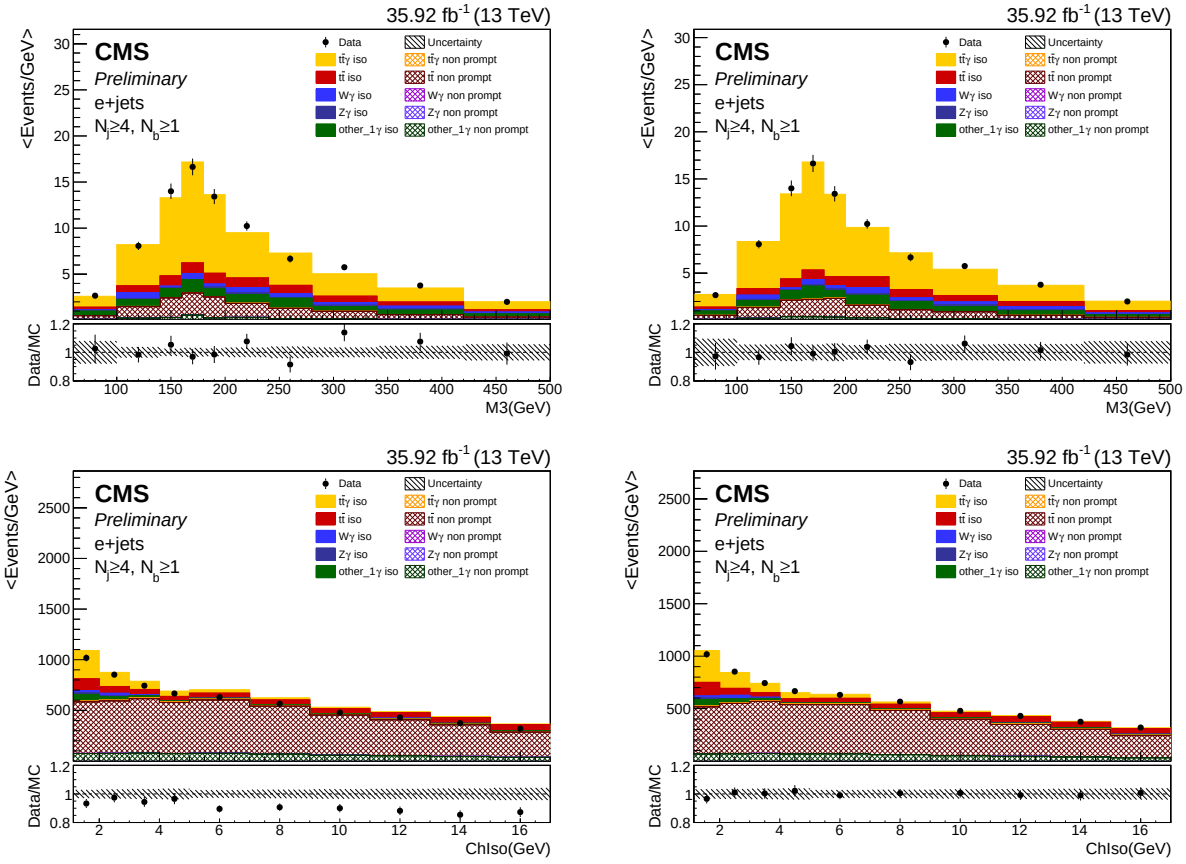


Figure 150: Pre-fit (left) and post-fit (right) plots for Data MC agreement for $\sigma_{t\bar{t}\gamma}$ fitting in the signal region for M_3 (top) and ChIso (bottom) for e channel in 2016 data.

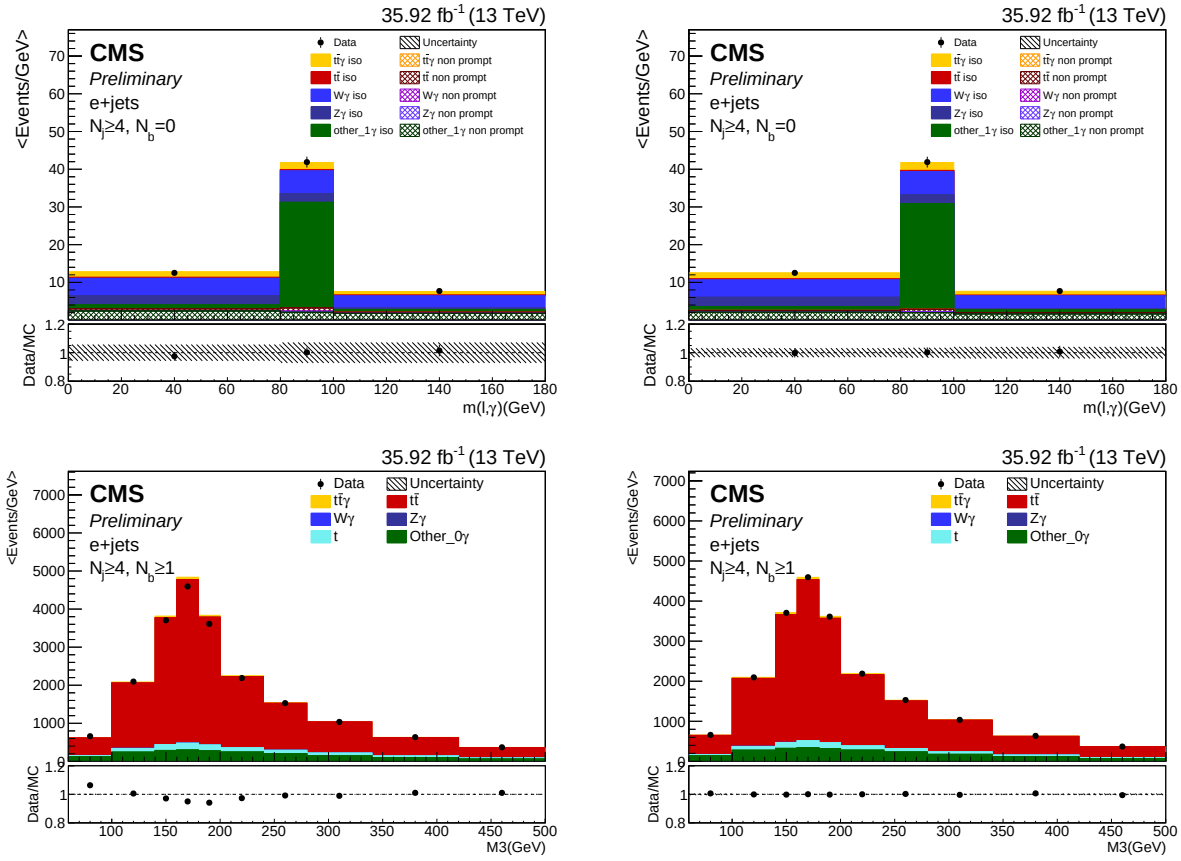


Figure 151: Pre-fit (left) and post-fit (right) plots for Data MC agreement for $\sigma_{t\bar{t}\gamma}$ fitting in zero btag control region (top) and zero photon control region (bottom) for e channel in 2016 data.

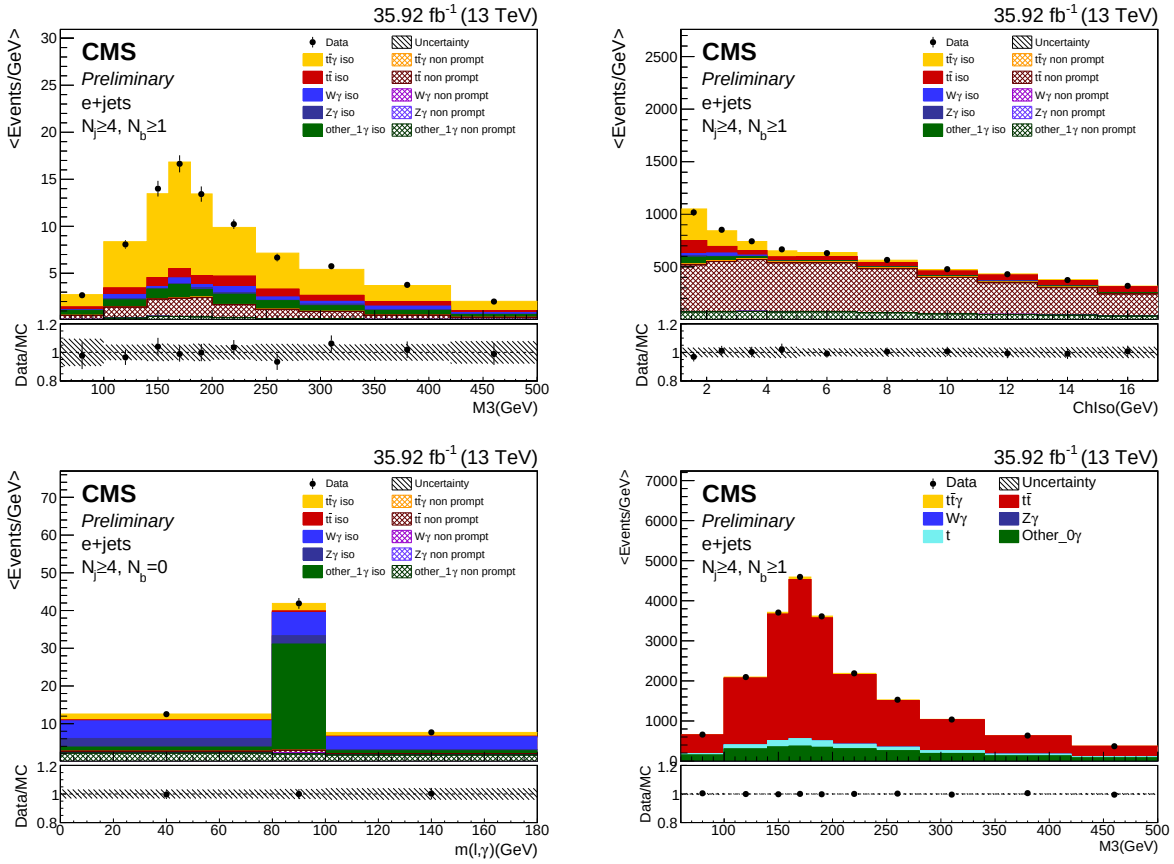


Figure 152: Post-fit plots for Data MC agreement in $\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$ fitting for M_3 and ChIso in the signal region (top left and top right, respectively), and zero btag control region and zero photon control region (bottom left and bottom right) for e channel in 2016 data.

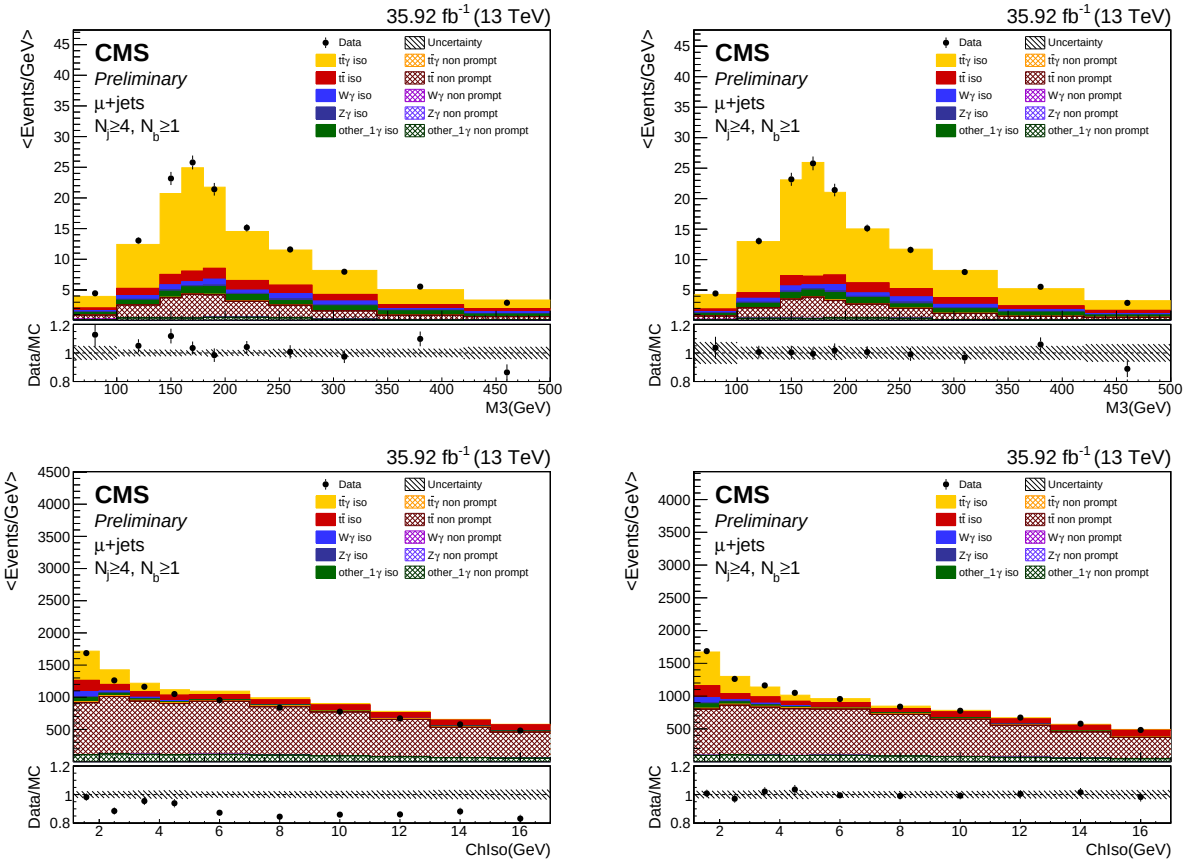


Figure 153: Pre-fit (left) and post-fit (right) plots for Data MC agreement for $\sigma_{t\bar{t}\gamma}$ fitting in the signal region for M_3 (top) and ChIso (bottom) for μ channel in 2016 data.

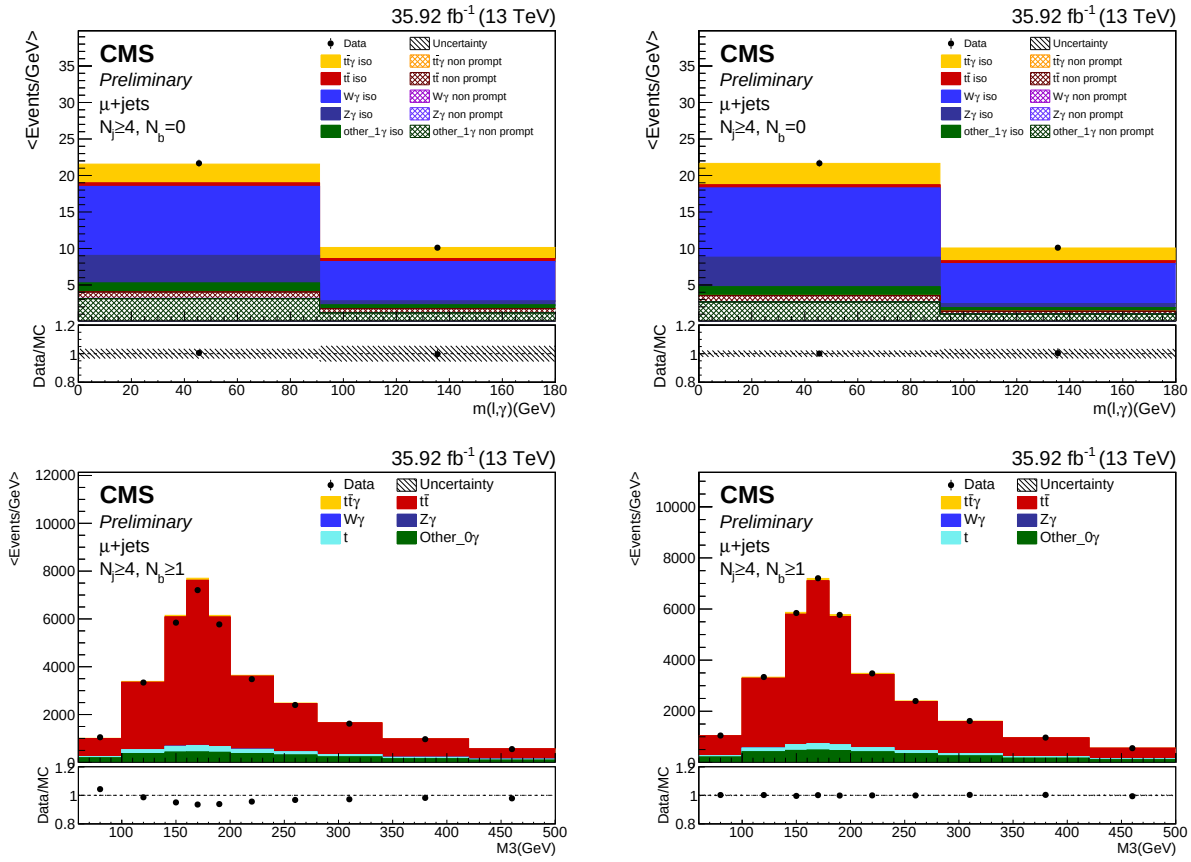


Figure 154: Pre-fit (left) and post-fit (right) plots for Data MC agreement for $\sigma_{t\bar{t}\gamma}$ fitting in zero btag control region (top) and zero photon control region (bottom) for μ channel in 2016 data.

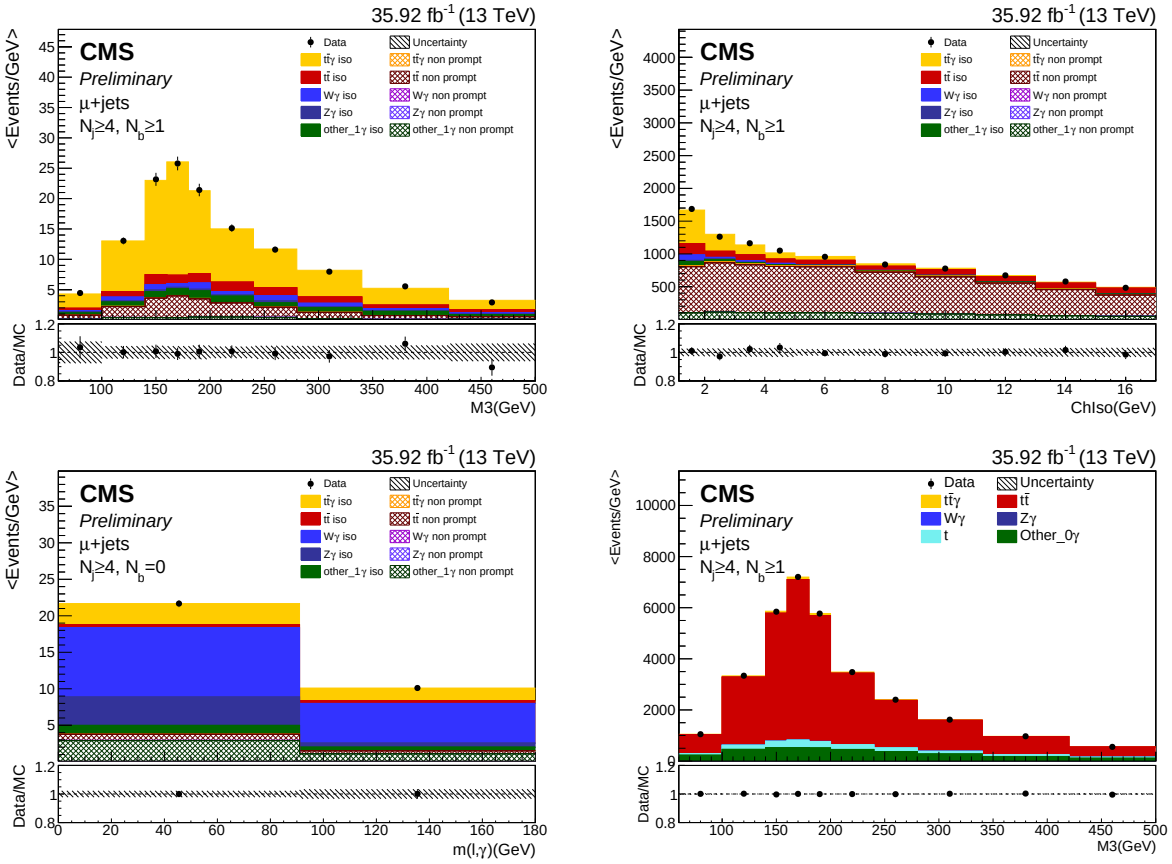


Figure 155: Post-fit plots for Data MC agreement in $\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$ fitting for M_3 and ChIso in the signal region (top left and top right, respectively), and zero btag control region and zero photon control region (bottom left and bottom right) for μ channel in 2016 data.

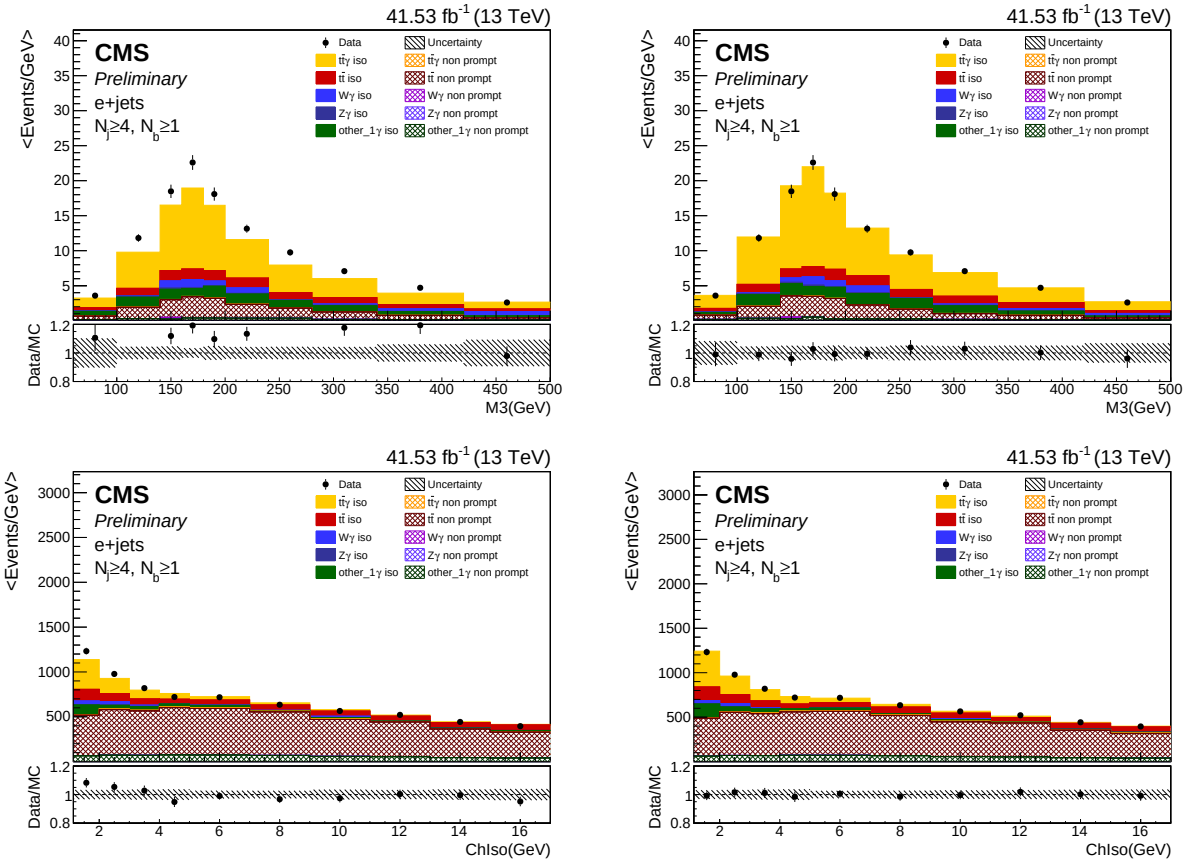


Figure 156: Pre-fit (left) and post-fit (right) plots for Data MC agreement for $\sigma_{t\bar{t}\gamma}$ fitting in the signal region for M_3 (top) and ChIso (bottom) for e channel in 2017 data.

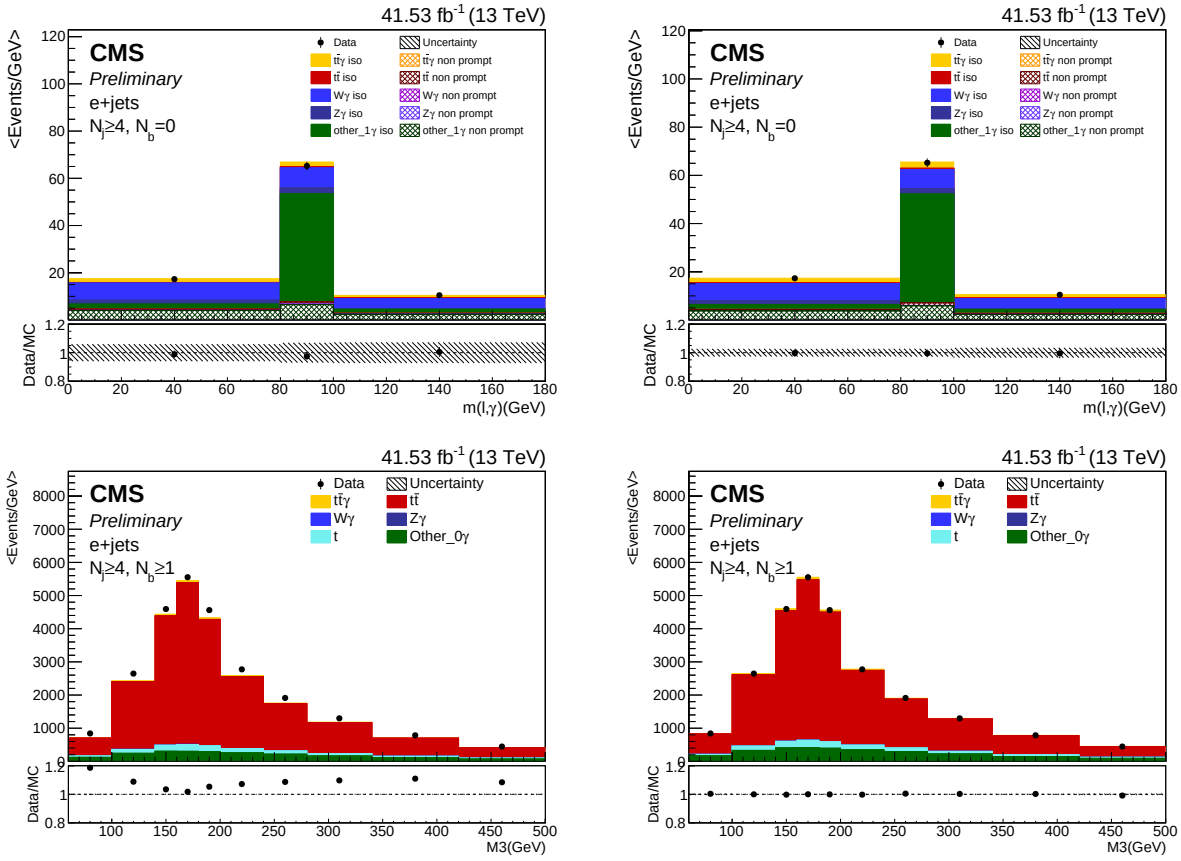


Figure 157: Pre-fit (left) and post-fit (right) plots for Data MC agreement for $\sigma_{t\bar{t}\gamma}$ fitting in zero btag control region (top) and zero photon control region (bottom) for e channel in 2017 data.

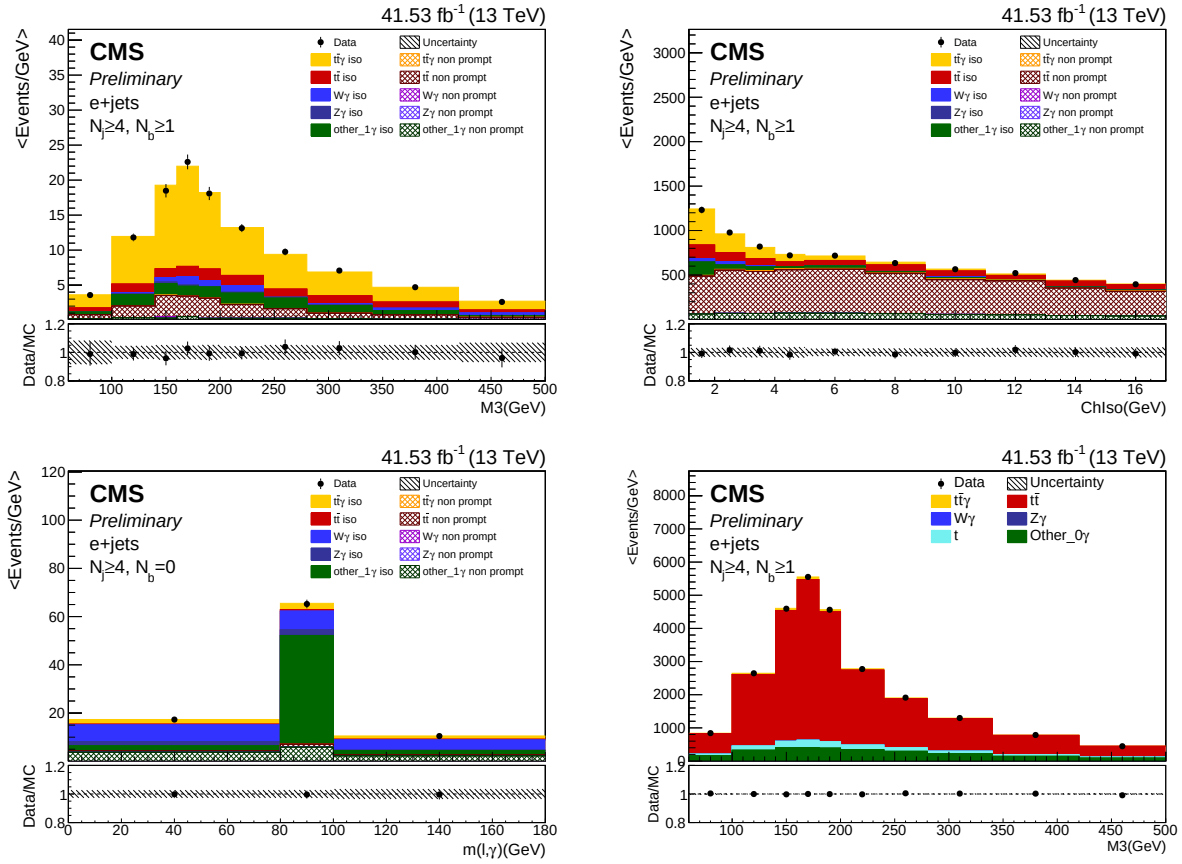


Figure 158: Post-fit plots for Data MC agreement in $\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$ fitting for M_3 and ChIso in the signal region (top left and top right, respectively), and zero btag control region and zero photon control region (bottom left and bottom right) for e channel in 2017 data.

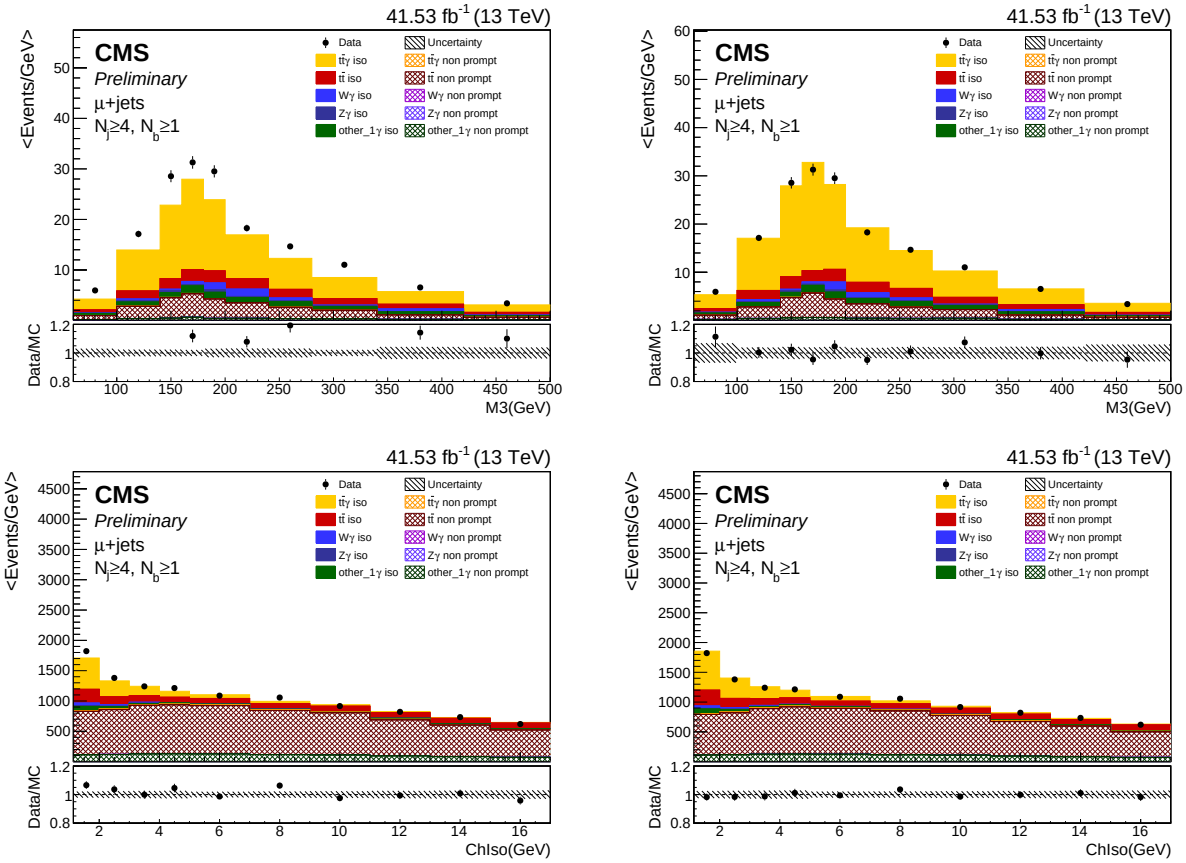


Figure 159: Pre-fit (left) and post-fit (right) plots for Data MC agreement for $\sigma_{t\bar{t}\gamma}$ fitting in the signal region for M_3 (top) and ChIso (bottom) for μ channel in 2017 data.

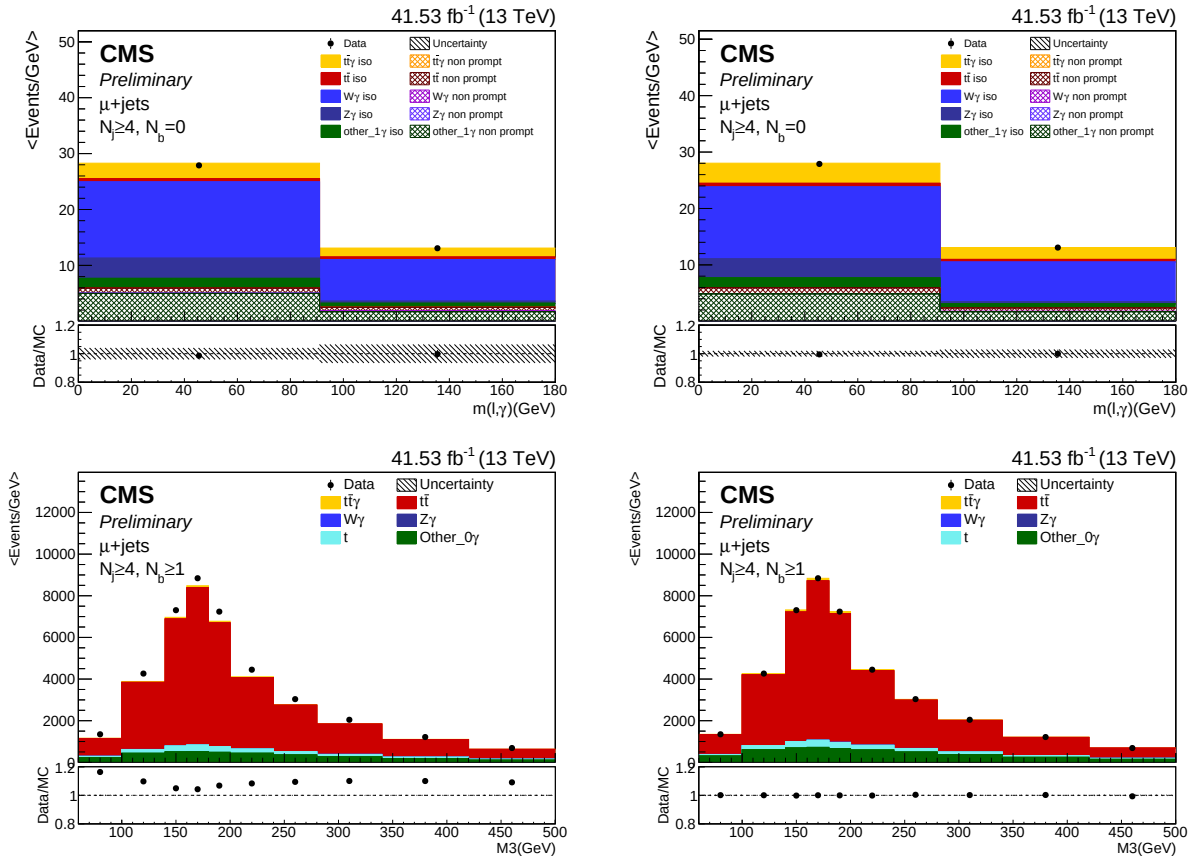


Figure 160: Pre-fit (left) and post-fit (right) plots for Data MC agreement for $\sigma_{t\bar{t}\gamma}$ fitting in zero btag control region (top) and zero photon control region (bottom) for μ channel in 2017 data.

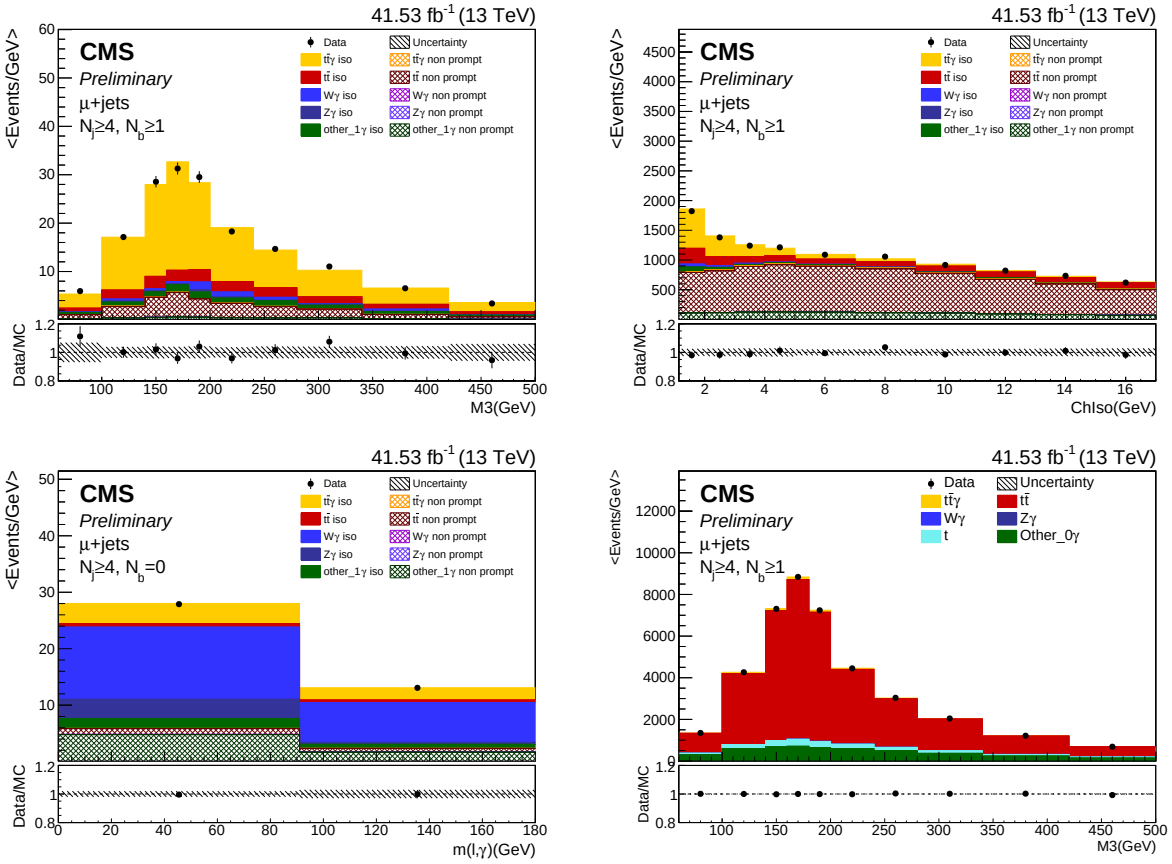


Figure 161: Post-fit plots for Data MC agreement in $\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$ fitting for M_3 and ChIso in the signal region (top left and top right, respectively), and zero btag control region and zero photon control region (bottom left and bottom right) for μ channel in 2017 data.

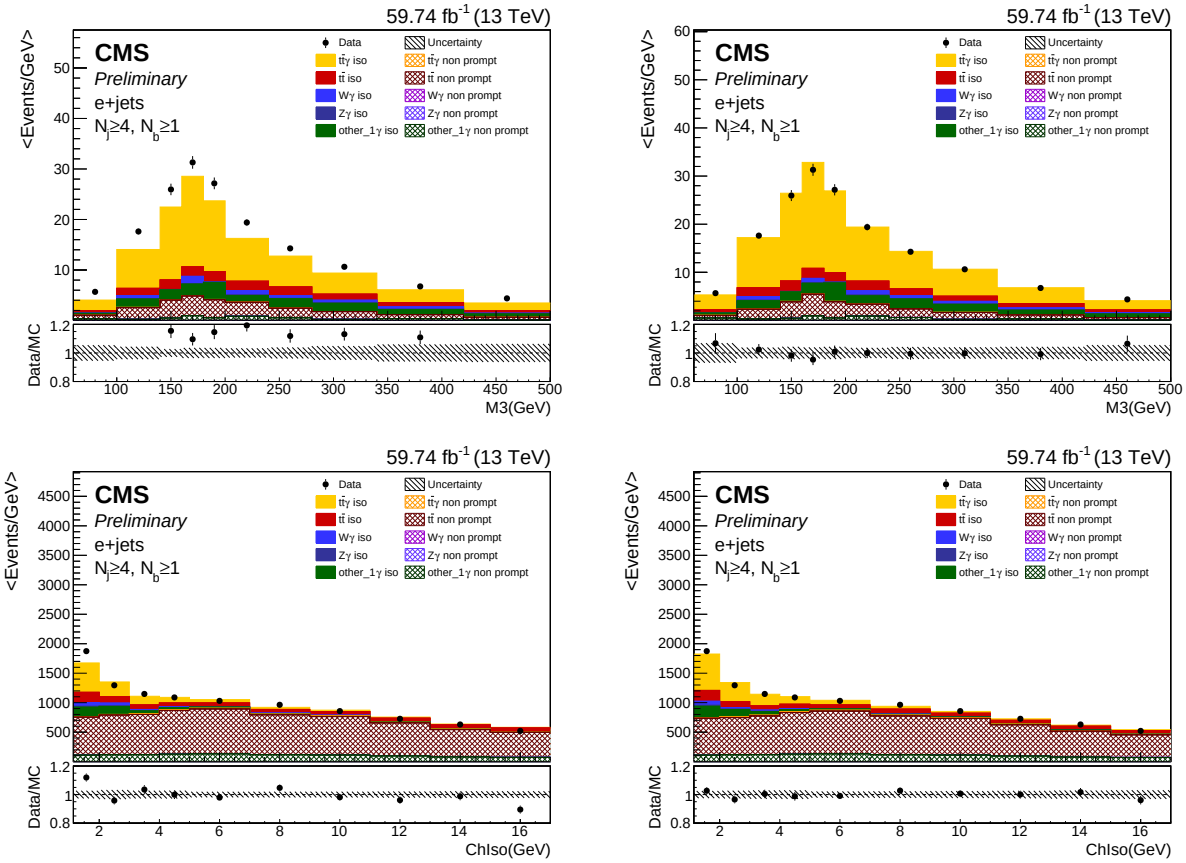


Figure 162: Pre-fit (left) and post-fit (right) plots for Data MC agreement for $\sigma_{t\bar{t}\gamma}$ fitting in the signal region for M_3 (top) and ChIso (bottom) for e channel in 2018 data.

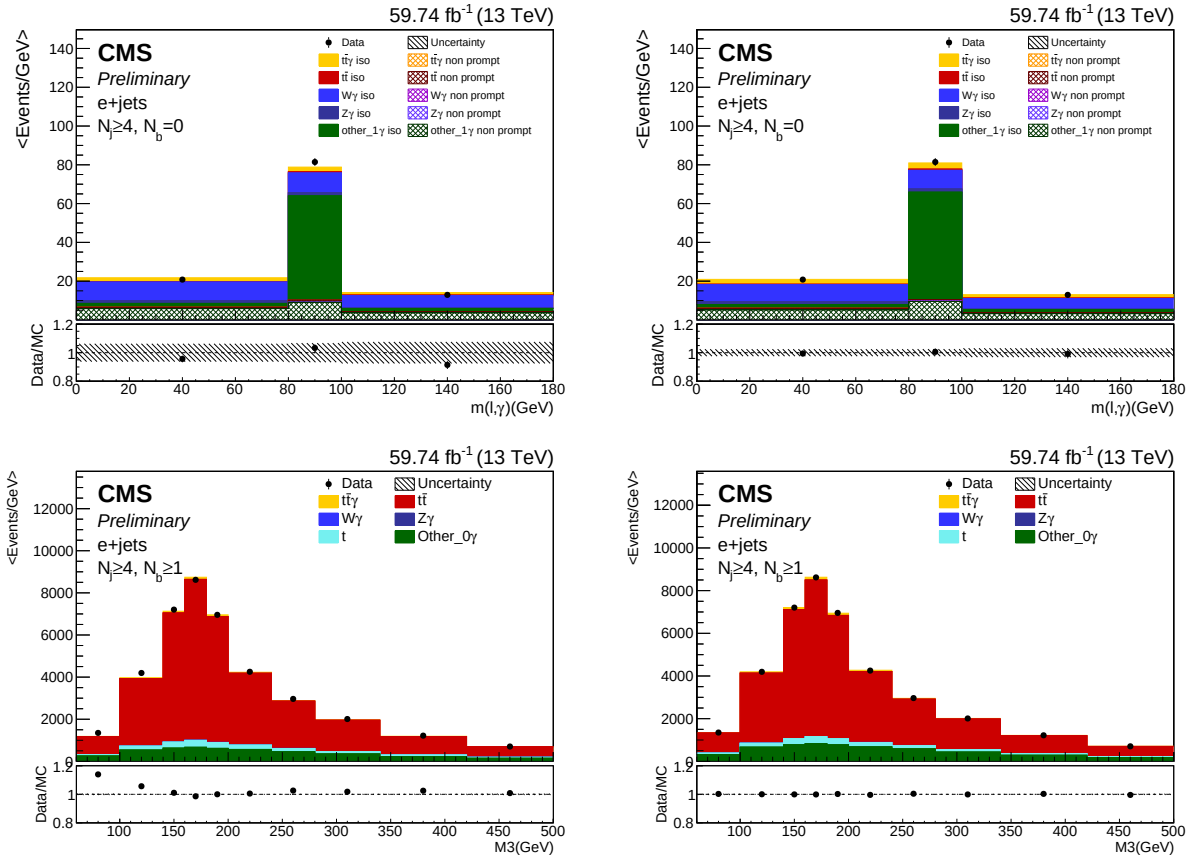


Figure 163: Pre-fit (left) and post-fit (right) plots for Data MC agreement for $\sigma_{t\bar{t}\gamma}$ fitting in zero btag control region (top) and zero photon control region (bottom) for e channel in 2018 data.

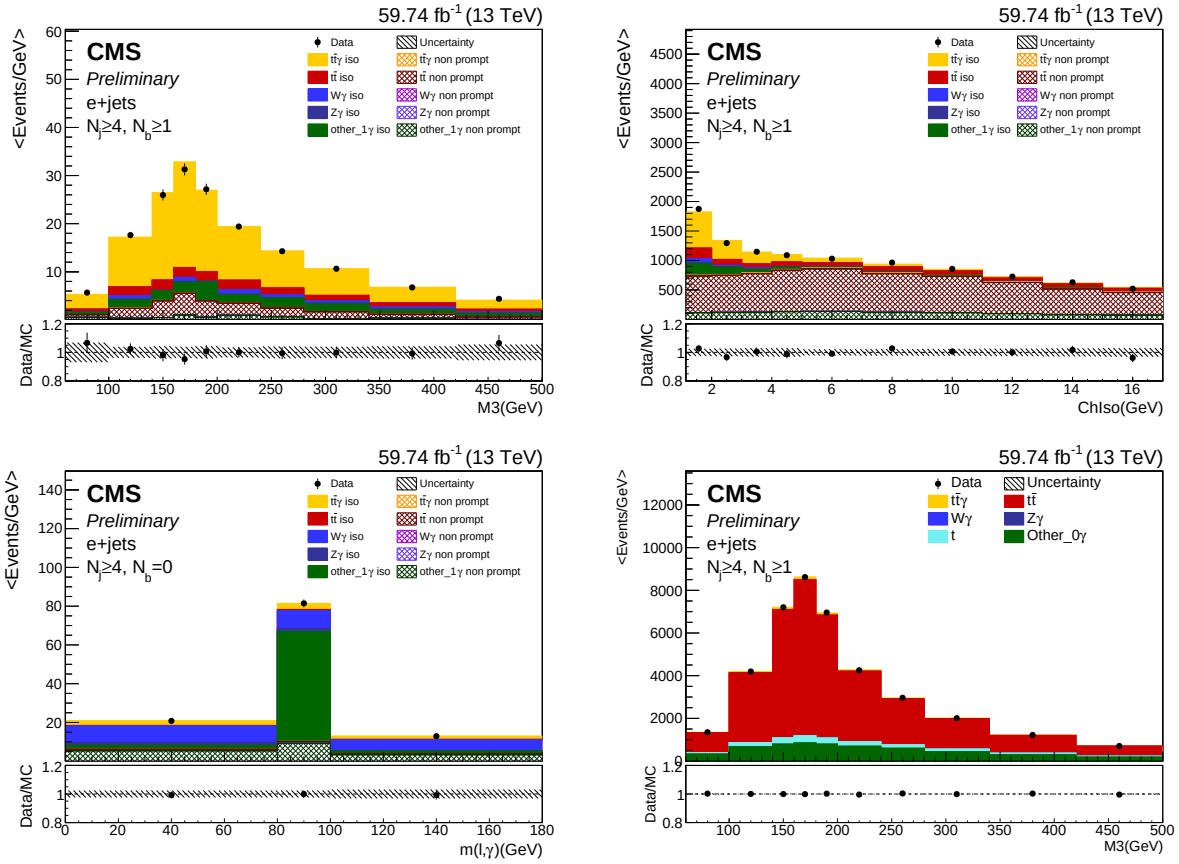


Figure 164: Post-fit plots for Data MC agreement in $\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$ fitting for M_3 and ChIso in the signal region (top left and top right, respectively), and zero btag control region and zero photon control region (bottom left and bottom right) for e channel in 2018 data.

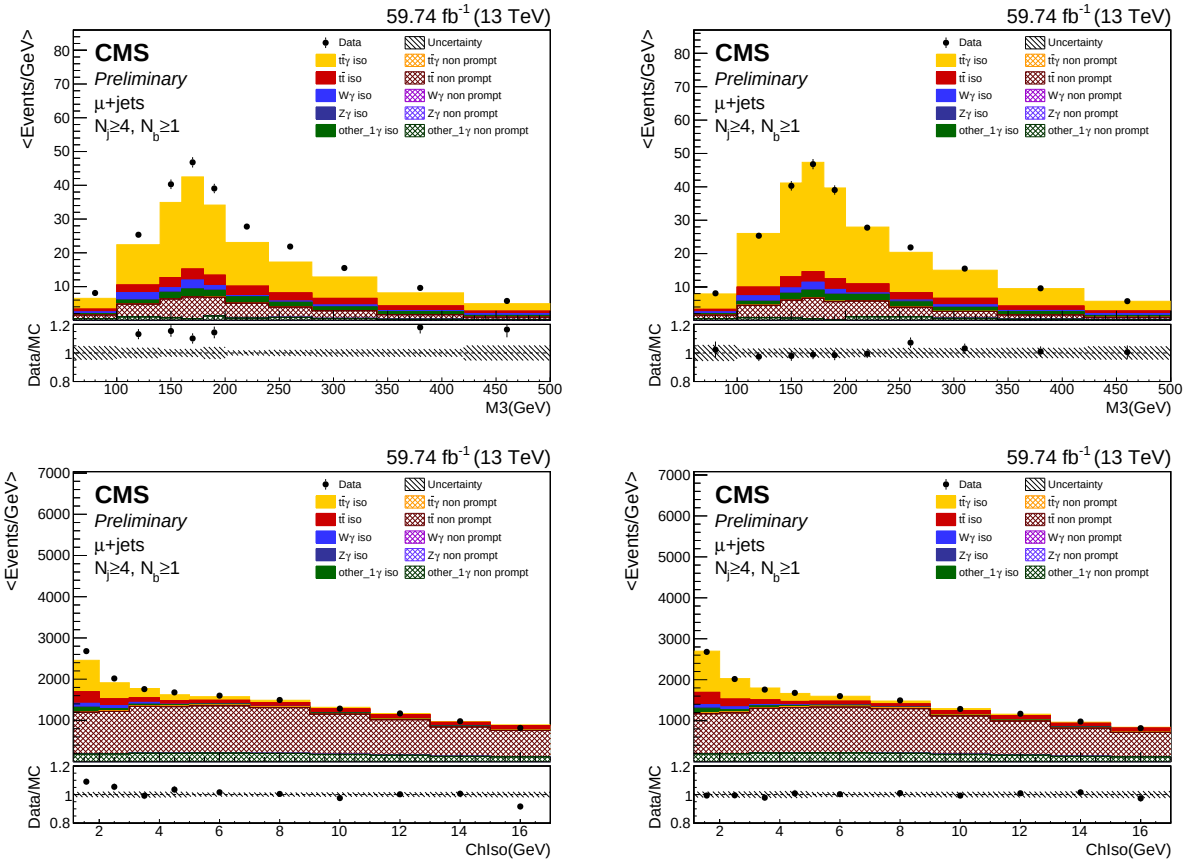


Figure 165: Pre-fit (left) and post-fit (right) plots for Data MC agreement for $\sigma_{t\bar{t}\gamma}$ fitting in the signal region for M_3 (top) and ChIso (bottom) for μ channel in 2018 data.

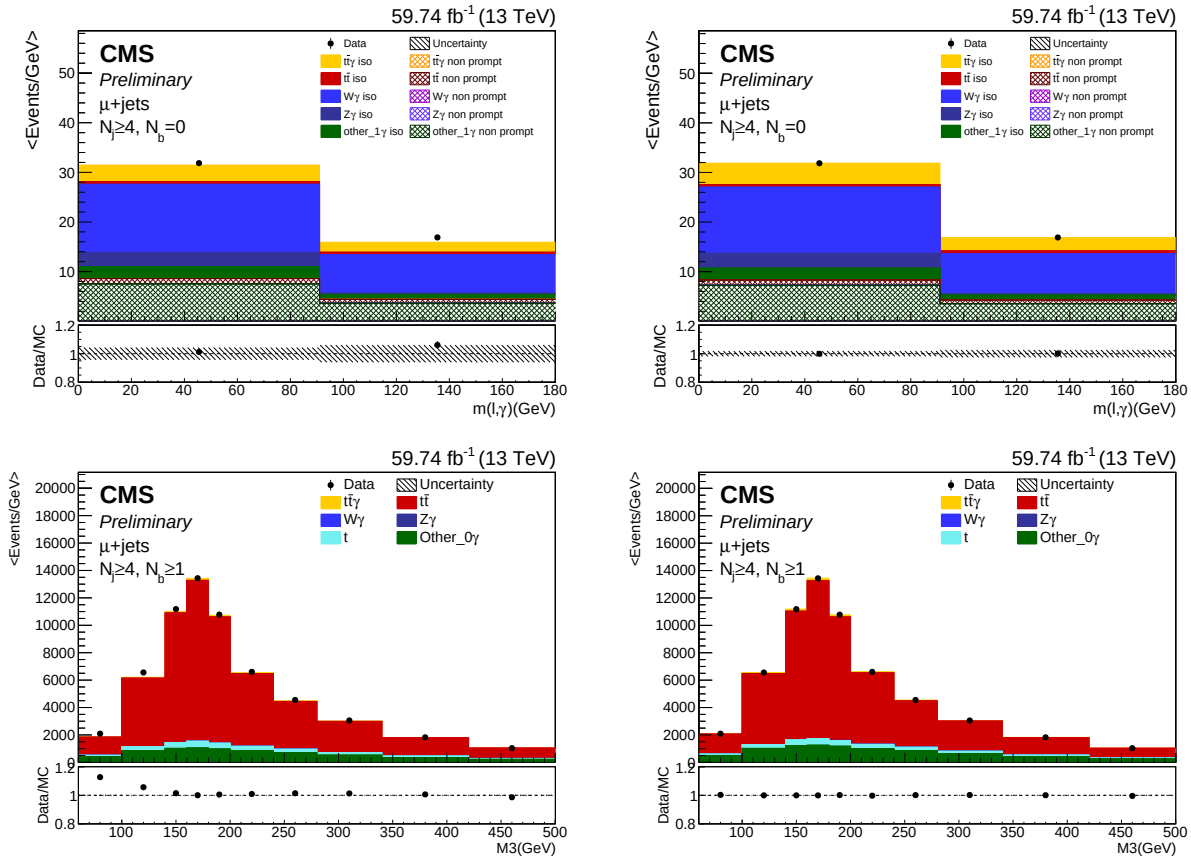


Figure 166: Pre-fit (left) and post-fit (right) plots for Data MC agreement for $\sigma_{t\bar{t}\gamma}$ fitting in zero btag control region (top) and zero photon control region (bottom) for μ channel in 2018 data.

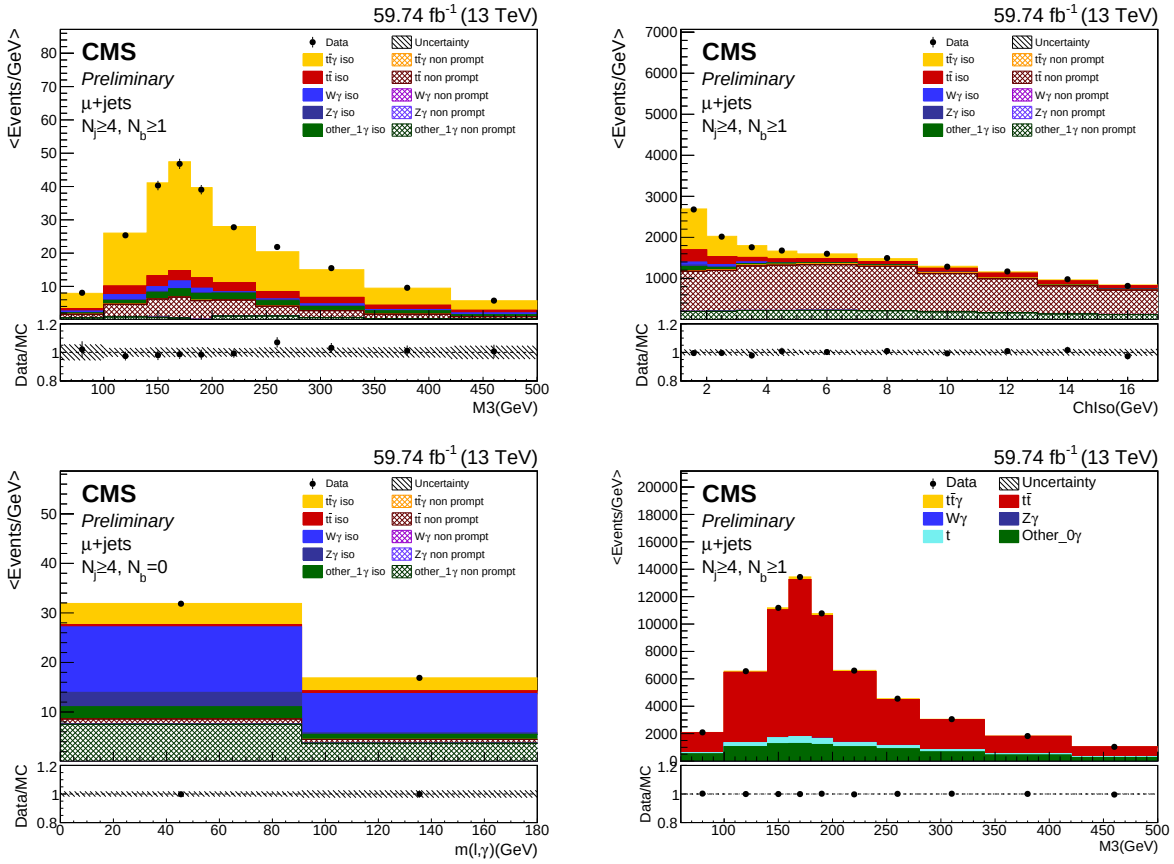


Figure 167: Post-fit plots for Data MC agreement in $\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$ fitting for M_3 and ChIso in the signal region (top left and top right, respectively), and zero btag control region and zero photon control region (bottom left and bottom right) for μ channel in 2018 data.

BIBLIOGRAPHY

[1] T. W. B. Kibble, "History of electroweak symmetry breaking", *J. Phys. Conf. Ser.* **626** (2015), no. 1, 012001, doi:10.1088/1742-6596/626/1/012001, arXiv:1502.06276.

[2] J. J. Aubert et al., "Experimental Observation of a Heavy Particle J", *Phys. Rev. Lett.* **33** (Dec, 1974) 1404–1406, doi:10.1103/PhysRevLett.33.1404.

[3] J. E. Augustin et al., "Discovery of a narrow resonance in e^+e^- annihilation", *Phys. Rev. Lett.* **33** (Dec, 1974) 1406–1408, doi:10.1103/PhysRevLett.33.1406.

[4] M. L. Perl et al., "Evidence for anomalous lepton production in $e^+ - e^-$ annihilation", *Phys. Rev. Lett.* **35** (Dec, 1975) 1489–1492, doi:10.1103/PhysRevLett.35.1489.

[5] S. W. Herb et al., "Observation of a dimuon resonance at 9.5 gev in 400-gev proton-nucleus collisions", *Phys. Rev. Lett.* **39** (Aug, 1977) 252–255, doi:10.1103/PhysRevLett.39.252.

[6] CDF Collaboration, "Observation of top quark production in $\bar{p}p$ collisions", *Phys. Rev. Lett.* **74** (1995) 2626–2631, doi:10.1103/PhysRevLett.74.2626, arXiv:hep-ex/9503002.

[7] D0 Collaboration, "Observation of the top quark", *Phys. Rev. Lett.* **74** (1995) 2632–2637, doi:10.1103/PhysRevLett.74.2632, arXiv:hep-ex/9503003.

[8] J. Aubert et al., "The ratio of the nucleon structure functions f2n for iron and deuterium", *Physics Letters B* **123** (1983), no. 3, 275–278, doi:https://doi.org/10.1016/0370-2693(83)90437-9.

[9] G. Arnison et al., "Experimental observation of lepton pairs of invariant mass around 95 gev/c2 at the cern sps collider", *Physics Letters B* **126** (1983), no. 5, 398–410, doi:https://doi.org/10.1016/0370-2693(83)90188-0.

[10] CMS Collaboration, M. Flechl, "Higgs boson discovery and recent results", in *22nd International Workshop on High Energy Physics and Quantum Field Theory*. 10, 2015. arXiv:1510.01924.

[11] J. F. Donoghue, E. Golowich, and B. R. Holstein, "Dynamics of the Standard Model", volume 35. Cambridge University Press, 2014.

[12] "The standard model", Jun, 2021. https://www.fnal.gov/pub/inquiring/matter/madeof/index.html.

[13] Super-Kamiokande Collaboration, "Evidence for oscillation of atmospheric neutrinos", *Phys. Rev. Lett.* **81** (Aug, 1998) 1562–1567, doi:10.1103/PhysRevLett.81.1562.

[14] M. Kobayashi and T. Maskawa, "CP-Violation in the Renormalizable Theory of Weak Interaction", *Progress of Theoretical Physics* **49** (02, 1973) 652–657, doi:10.1143/PTP.49.652.

[15] C. Campagnari and M. Franklin, "The Discovery of the top quark", *Rev. Mod. Phys.* **69** (1997) 137–212, doi:10.1103/RevModPhys.69.137, arXiv:hep-ex/9608003.

[16] A. Giammanco and G. Gagliardi, "Top quark at LHC", technical report, CERN, Geneva, May, 2005.

[17] G. Altarelli and G. Parisi, "Asymptotic freedom in parton language", *Nuclear Physics B* **126** (1977), no. 2, 298–318, doi:https://doi.org/10.1016/0550-3213(77)90384-4.

[18] A. Heinson, "Feynman Diagrams for Top Physics Talks and Notes", Oct, 2011. https://www-d0.fnal.gov/Run2Physics/top/top_public_web_pages/top_feynman_diagrams.html.

[19] "Top quark", Mar, 2021. https://en.wikipedia.org/w/index.php?title=Top_quark&oldid=1010072508.

[20] U. Baur, A. Juste, L. H. Orr, and D. Rainwater, "Probing electroweak top quark couplings at hadron colliders", *Phys. Rev. D* **71** (Mar, 2005) 054013, doi:10.1103/PhysRevD.71.054013.

[21] K. Melnikov, M. Schulze, and A. Scharf, "QCD corrections to top quark pair production in association with a photon at hadron colliders", *Phys. Rev. D* **83** (2011) 074013, doi:10.1103/PhysRevD.83.074013, arXiv:1102.1967.

[22] P. F. Duan et al., "QCD corrections to associated production of $t\bar{t}\gamma$ at hadron colliders", *Phys. Rev. D* **80** (Jul, 2009) 014022, doi:10.1103/PhysRevD.80.014022.

[23] "Physics results combined cmspublic twiki", Mar, 2021. https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined.

[24] CDF Collaboration, "Evidence for $t\bar{t}\gamma$ Production and Measurement of $\sigma_{t\bar{t}\gamma}/\sigma_{t\bar{t}}$", *Phys. Rev. D* **84** (2011) 031104, doi:10.1103/PhysRevD.84.031104, arXiv:1106.3970.

[25] 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* **91** (2015), no. 7, 072007, doi:10.1103/PhysRevD.91.072007, arXiv:1502.00586.

[26] 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", *JHEP* **11** (2017) 086, doi:10.1007/JHEP11(2017)086, arXiv:1706.03046.

[27] 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", *Eur. Phys. J. C* **79** (2019), no. 5, 382, doi:10.1140/epjc/s10052-019-6849-6, arXiv:1812.01697.

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

[29] U. Husemann, "Top-Quark Physics: Status and Prospects", *Prog. Part. Nucl. Phys.* **95** (2017) 48–97, doi:10.1016/j.ppnp.2017.03.002, arXiv:1704.01356.

[30] J. A. Aguilar-Saavedra, N. F. Castro, and A. Onofre, "Constraints on the Wtb vertex from early LHC data", *Phys. Rev. D* **83** (2011) 117301, doi:10.1103/PhysRevD.83.117301, arXiv:1105.0117.

[31] M. Aaboud et al., "Measurement of the $\mathrm{t\bar{t}}\gamma$ production cross section in proton-proton collisions at $\sqrt{s} = 8$ tev with the atlas detector", *Journal of High Energy Physics* **2017** (11, 2017) doi:10.1007/JHEP11(2017)086.

[32] CMS Collaboration, G. L. Bayatian and et. al, "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software". Technical Design Report CMS. CERN, Geneva, 2006.

[33] The CMS Collaboration, "CMS physics technical design report, volume II: Physics

performance", *Journal of Physics G: Nuclear and Particle Physics* **34** (April, 2007) 995–1579, doi:10.1088/0954-3899/34/6/s01.

[34] W. Barletta et al., "Working Group Report: Hadron Colliders", in *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*. 2013. arXiv:1310.0290.

[35] P. Brogi, "Study of inelastic processes in proton-proton collisions at the LHC with the TOTEM Experiment", Master's thesis, INFN, Pisa, 2011.

[36] Particle Data Group Collaboration, "Review of Particle Physics", *PTEP* **2020** (2020), no. 8, 083C01, doi:10.1093/ptep/ptaa104.

[37] "Luminosity Public Results", Apr, 2021. https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[38] CMS Collaboration, "The CMS experiment at the CERN LHC. The Compact Muon Solenoid experiment", *JINST* **3** (2008) S08004. 361 p, doi:10.1088/1748-0221/3/08/S08004.

[39] E. Focardi, "Status of the cms detector", *Physics Procedia* **37** (2012) 119–127, doi:https://doi.org/10.1016/j.phpro.2012.02.363.

[40] T. Lenzi, "Development and Study of Different Muon Track Reconstruction Algorithms for the Level-1 Trigger for the CMS Muon Upgrade with GEM Detectors", Master's thesis, U. Brussels (main), 2013.

[41] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018), no. 06, P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.

[42] G. Petrucciani, "The CMS experiment at the CERN LHC", pp. 15–58. Scuola Normale Superiore, Pisa, 2013. doi:10.1007/978-88-7642-482-3_2.

[43] CMS Collaboration, "The CMS tracker: addendum to the Technical Design Report", 2000. https://cds.cern.ch/record/490194.

[44] CMS Collaboration, Karimäki, V and Others, "The CMS tracker system project: Technical Design Report", 1997. https://cds.cern.ch/record/368412.

[45] A. Dominguez and D. e. a. Abbaneo, "CMS Technical Design Report for the Pixel Detector Upgrade", Sep, 2012. https://cds.cern.ch/record/1481838.

[46] L. Viliani, "CMS tracker performance and readiness for LHC Run II", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **824** (2016) 67–69, doi:https://doi.org/10.1016/j.nima.2015.09.046.

[47] "Silicon Strips Detector for CMS Experiment", Apr, 2021. https://cms.cern/detector/identifying-tracks/silicon-strips.

[48] The CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the cms tracker", *Journal of Instrumentation* **9** (Oct, 2014) P10009–P10009, doi:10.1088/1748-0221/9/10/p10009.

[49] CMS Collaboration, "The CMS electromagnetic calorimeter project: Technical Design Report". Technical design report. CERN, Geneva, 1997.

[50] CMS Collaboration, S. Chatrchyan et al., "Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV", 2013. doi:10.1088/1748-0221/8/09/P09009.

[51] CMS Collaboration, "Particle-flow reconstruction and global event description with

the CMS detector", *JINST* **12** (2017), no. 10, P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[52] G. Bayatian et al., "Design, performance and calibration of the CMS forward calorimeter wedges", *Eur. Phys. J. C* **53** (2008) 139–166, doi:10.1140/epjc/s10052-007-0459-4.

[53] CMS Collaboration, "Performance of the CMS Hadron Calorimeter with Cosmic Ray Muons and LHC Beam Data", *JINST* **5** (2010) T03012, doi:10.1088/1748-0221/5/03/T03012, arXiv:0911.4991.

[54] G. Cerminara, "The drift tube system of the cms experiment", *Nuclear Physics B - Proceedings Supplements* **172** (2007) 71–74, doi:https://doi.org/10.1016/j.nuclphysbps.2007.07.016.

[55] CMS Collaboration, "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **15** (2020), no. 10, P10017, doi:10.1088/1748-0221/15/10/P10017, arXiv:2006.10165.

[56] CMS Collaboration, A. Tapper and D. Acosta, "CMS technical design report for the level-1 trigger upgrade", 2013. https://inspirehep.net/literature/1261310.

[57] CMS Collaboration, "The CMS trigger system", *JINST* **12** (2017), no. 01, P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

[58] CMS Collaboration, "CMS the computing project, technical design report", 07, 2005. https://inspirehep.net/literature/692910.

[59] " Website explaining the CMS computing model", Apr, 2021. https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookComputingModel.

[60] C. Oleari, "The POWHEG-BOX", *Nucl. Phys. B Proc. Suppl.* **205-206** (2010) 36–41, doi:10.1016/j.nuclphysbps.2010.08.016, arXiv:1007.3893.

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

[62] J. Alwall et al., "MadGraph 5 : Going Beyond", *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.

[63] Sherpa Collaboration, "Event Generation with Sherpa 2.2", *SciPost Phys.* **7** (2019), no. 3, 034, doi:10.21468/SciPostPhys.7.3.034, arXiv:1905.09127.

[64] "Sherpa and open science grid: Predicting the emergence of jets", Aug, 2020. https://sciencenode.org/feature/sherpa-and-open-science-grid-predicting-emergence-jets.php.

[65] "Summary table 13 tev $\mathrm{t\bar{t}}$ cross section", May, 2021. https://twiki.cern.ch/twiki/bin/view/CMS/SummaryTable1G25ns.

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

[67] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473–500, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[68] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

[69] "References CMS cross section measurements", Jun, 2021. https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombinedRefs.

[70] S. Sumowidagdo, "A pedagogical derivation of the matrix element method in particle physics data analysis", *Journal of Physics: Conference Series* **988** (03, 2018) 012003, doi:10.1088/1742-6596/988/1/012003.

[71] Donato, Silvio, "Cms trigger performance", *EPJ Web Conf.* **182** (2018) 02037, doi:10.1051/epjconf/201818202037.

[72] Perrotta, Andrea, "Cms event reconstruction status in run 2", *EPJ Web Conf.* **214** (2019) 02015, doi:10.1051/epjconf/201921402015.

[73] "Physics Results TOP Event Displays". https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsTOPEventDisplays.

[74] R. Frühwirth, "Application of kalman filtering to track and vertex fitting", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262** (1987), no. 2, 444 – 450, doi:https://doi.org/10.1016/0168-9002(87)90887-4.

[75] CMS Collaboration, "CMS Tracking Performance Results from Early LHC Operation", *Eur. Phys. J. C* **70** (2010) 1165–1192, doi:10.1140/epjc/s10052-010-1491-3, arXiv:1007.1988.

[76] N. Akchurin et al., "First beam tests of prototype silicon modules for the cms high granularity endcap calorimeter", *Journal of Instrumentation* **13** (10, 2018) P10023–P10023, doi:10.1088/1748-0221/13/10/P10023.

[77] "Particle flow clustering cms twiki", Sep, 2020. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideParticleFlowClustering.

[78] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at sqrt(s) = 8 tev", arXiv:1502.02701.

[79] CMS Collaboration, "Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at sqrt(s) = 8 TeV", *JINST* **10** (2015), no. 08, P08010, doi:10.1088/1748-0221/10/08/P08010, arXiv:1502.02702.

[80] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[81] M. Cacciari, G. P. Salam, and G. Soyez, "Fastjet user manual", *Eur. Phys. J.* **C72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[82] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[83] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018), no. 05, P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[84] "B-tagging commissioning", Sep, 2020. https://hackmd.io/@cN0wHbxqQ52iMSByxQI6CQ/rkavdgSEG?type=view.

[85] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s}$ = 8 TeV", *JINST* **10** (2015), no. 02, P02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.

[86] "Dataformats cmspublic $<$ twiki", Jun, 2021. https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookDataFormats.

[87] A. Rizzi, G. Petrucciani, and M. Peruzzi, "A further reduction in cms event data for

analysis: the nanoaod format", *EPJ Web of Conferences* **214** (01, 2019) 06021, doi:10.1051/epjconf/201921406021.

[88] "Missing transverse energy optional filters for Run II". https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2#MiniAOD_805.

[89] CMS Collaboration, "The CMS trigger system", *JINST* **12** (2017), no. 01, P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

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

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

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

[93] CMS Collaboration, "Cut-based photon ID recommendations for Run II", 2016-2018. https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedPhotonIdentificationRun2#Cut_Based_Photon_ID_for_Run_2.

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

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

[96] 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", 2021. https://inspirehep.net/literature/1852697.

[97] A. Sirunyan et al., "Performance of the CMS level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Journal of Instrumentation* **15** (oct, 2020) P10017–P10017, doi:10.1088/1748-0221/15/10/p10017.

[98] CMS Collaboration, "Search for disappearing tracks in proton-proton collisions at $\sqrt{s}$ =13 tev", *Phys. Lett. B* **806** (2020) 135502, doi:10.1016/j.physletb.2020.135502, arXiv:2004.05153.

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

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

[101] "Egm scale and smear corrections for electrons and photons.", May, 2021. https://twiki.cern.ch/twiki/bin/view/CMS/EgammaMiniAODV2#Energy_Scale_and_Smearing.

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

[103] CMS Collaboration, "Muon reference efficiencies for 2017", 2017. https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2017.

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

[105] Javier Cuevas Maestro, Santiago Folgueras, Juan Rodrigo González Fernández, Enrique Palencia Cortezón, Sergio Sánchez Cruz, "Lepton efficiency uncertainties in TOP analyses", CMS Analysis Note CMS AN-2018/210, CERN LHC, 2018.

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

[107] "Documentation of the roostats -based statistics tools for higgs pag". http://cms-analysis.github.io/higgsanalysis-combinedlimit.

[108] M. M. A. Ghneimat, "Probing the top-quark coupling to the photon through the cross-section measurement of $t\bar{t}\gamma$ production in *pp* collisions with the ATLAS detector", Apr, 2018. Presented 17 May 2018. https://cds.cern.ch/record/2320139.

[109] CMS Collaboration, "TOP Systematic Uncertainties (Run2)", 2020. https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopSystematics.

[110] CMS Collaboration, "CMS luminosity measurements for the 2016 data-taking period", 2017. https://cds.cern.ch/record/2257069.

[111] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV", 2018. https://cds.cern.ch/record/2621960.

[112] CMS Collaboration, "CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV", 2019. https://cds.cern.ch/record/2676164.

[113] CMS Collaboration, "CMS B-Tagging and Vertexing", 2016-2018. https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsBTV.

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

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

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

[117] R. J. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples", *Comput. Phys. Commun.* **77** (1993) 219–228, doi:10.1016/0010-4655(93)90005-W.

[118] Otto Hindrichs, "Measurement of differential cross sections for the production of top quark pairs in lepton+jets events from pp collisions at sqrt(s) = 13 TeV", 2018.

[119] L. G. Xia, "Study of constraint and impact of a nuisance parameter in maximum likelihood method", *J. Phys. G* **46** (2019) 085004, doi:10.1088/1361-6471/ab02c0, arXiv:1805.03961.

ABSTRACT

MEASUREMENT OF THE CROSS SECTION OF TOP QUARK PAIRS PRODUCED IN ASSOCIATION WITH A PHOTON IN LEPTON + JETS EVENTS AT $\sqrt{s}$ = 13 TEV WITH FULL RUNII CMS DATA

by

NABIN POUDYAL

AUGUST 2021

Advisor: Dr. Robert Harr

Major: Physics

Degree: Doctor Of Philosophy

The inclusive production cross section of top quark pairs in association with a photon is measured in proton-proton collisions at the LHC with 13 TeV energy using the full RunII data collected by CMS in 2016, 2017, and 2018 with a total corresponding integrated luminosity of 137 fb^{-1}. The relative fraction of $\mathrm{t\bar{t}}\gamma$ events normalized to inclusive $\mathrm{t\bar{t}}$ production is measured. The cross section measurement provides important information about the electromagnetic coupling of the standard model top quark and is sensitive to physics beyond the standard model. The analysis is carried out in the in semileptonic decay channel with a well isolated high $\mathrm{P_T}$ lepton (electron and muon), at least four jets out of which at least one must be b-tagged, and an isolated photon. Photons may be emitted from initial state radiation, top quarks, and decay products of top quarks. The fiducial phase space with a photon of transverse momentum $\geq$ 20 GeV and $|\eta| < 1.4441$ is used. A simultaneous likelihood fit of control regions with the signal region is done to constraint the backgrounds and to extract the $\mathrm{t\bar{t}}\gamma$ cross section. The measurement of the ratio of $\mathrm{t\bar{t}}\gamma$ to $\mathrm{t\bar{t}}$ is $0.02055 \pm$

$0.00099\,(\text{syst.}) \pm 0.00099\,(\text{stat.})$ in e channel , $0.02156 \pm 0.00068\,(\text{syst.}) \pm 0.00068\,(\text{stat.})$ in μ channel , and $0.02203 \pm 0.00064\,(\text{syst.}) \pm 0.00064\,(\text{stat.})$ in $e + \mu$ channel , respectively. The measured inclusive cross section is $3.81 \pm 0.15\,(\text{syst.}) \pm 0.10\,(\text{stat.})$ pb in e channel , $3.87 \pm 0.11\,(\text{syst.}) \pm 0.07\,(\text{stat.})$ pb in μ channel , and $3.96 \pm 0.10\,(\text{syst.}) \pm 0.06\,(\text{stat.})$ pb in $e + \mu$ channel , respectively for full RunII data. The total inclusive cross section with a photon of transverse momentum $\geq$ 20 GeV is found to be $3.96 \pm 0.10\,(\text{syst.}) \pm 0.06\,(\text{stat.})$ pb with full RunII CMS data. The results are in agreement with the standard model next to leading order prediction.

AUTOBIOGRAPHICAL STATEMENT

Education

2021 Ph.D. in Physics **(Expected Graduation by June 2021), Wayne State U**, Detroit MI
Thesis Title : **"Measurement of the cross section of top quark pairs in association with a photon in lepton + jets final state events at LHC 13TeV energy"**

2014 Master of Science in Physics, **U of South Dakota**, Vermilion SD
Thesis Title : **"Design and Development of Gadolinium doped Cherenkov Neutron Detector"**

2011 Bachelor of Science in Physics, **Tribhuvan University**, Kathmandu Nepal
Major : Physics and Statistics

Research Experience

2021 **Research Assistant | Wayne State University, Detroit, MI**

- Data analysis on precision measurement of the cross section of top quark pairs in association with a photon in lepton + jets final state events at LHC 13TeV energy.
- Hardware work on Test Beam operation, analysis and characterization of prototypes of 2S Outer tracker module R&D for HL-LHC phase-II upgrade.
- C++ software development of Monicelli and Chewie Fermilab framework (Alignment and analysis software resp).
- Development of the Burn-in System which consisted of Cold box linked to Chiller, Heating strips, and dry air supply connected by solid state relays and controlled by electronic control system to test performance of 2S outer tracker module at temperature -33°C
- Generator level study of hadronic decay of W boson using the tau like jet algorithm in stand alone Pythia8 software, fastjet and ROOT6.

TestBeam | Monicelli & Chewie C++ framework development | Pythia8 | fastjet | ROOT | higgs couplings | higgsCombine

2014 **Research Assistant | University of South Dakota , Vermillion, SD**

- Design and Development of Gadolinium doped Cherenkov Neutron Detector for The Large Underground Xenon (LUX) experiment at Sandford Underground Laboratory(SURF)

Neutron Detector | Design and Development of instrument and software

Selected Publications

2020 Beam Test Performance of Prototype Silicon Detectors for the Outer Tracker for the Phase-2 Upgrade of CMS, JINST 15 (2020) P03014

2021 Test Beam Performance of CBC3-based Prototype Modules of the phase II CMS Outer Tracker before and after neutron irradiation [Under preparation]

2021 Measurement of the inclusive $t\bar{t}\gamma$ cross section in the single lepton channel with 137 fb^{-1} at $\sqrt{s}$ = 13 TeV using CMS full Run II data.

Conferences and Presentations

2020 **2020 APS April Meeting** , Remote Online, Virtual
Title : "Measurement of the cross section of top quark pairs in association with a photon in lepton + jets events at $\sqrt{s}$ = 13 TeV with CMS full RunII data"

2019 **2019 Fermilab User Meeting Poster presentation**, Fermilab, IL
Title : "FNAL testbeam analysis of irradiated CBC3-2S minimodule for Outer Tracker phase 2 upgrade for CMS CERN"

2018 **2018 USCMS Annual Meeting**, U of Minnesota, Minneapolis
Title : "FNAL testbeam analysis for CBC3-2S minimodule for OT phase 2 upgrade"

2017 **2017 USCMS Annual Meeting**, Wayne State University, Detroit

2017 **Wayne State University Graduate Research Day Poster Presentation**, Wayne State U, Detroit
Title : "Simulation of study of L1 trigger design for Hadronic decay of W boson for CMS"