INCLUSIVE SEARCH FOR NEW NONRESONANT PHENOMENA IN MULTILEPTON FINAL STATES WITH THE CMS DETECTOR AT THE LHC

By MAXIMILIAN HEINDL

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Physics and Astronomy

Written under the direction of

Sunil Somalwar

And approved by

New Brunswick, New Jersey

October 2021

ABSTRACT OF THE DISSERTATION

Inclusive search for new nonresonant phenomena in multilepton final states with the CMS detector at the LHC

By MAXIMILIAN HEINDL

Dissertation Director:

Sunil Somalwar

Many theories of physics beyond the Standard Model, such as the type-III seesaw mechanism, vector-like taus, or third generation scalar leptoquarks, can be probed at the Large Hadron Collider (LHC) by searching for signatures with multiple leptons. Therefore, an inclusive search for nonresonant signatures of phenomena beyond the Standard Model in events with three or more charged leptons (electrons, muons, and hadronic taus) is performed and presented in this dissertation. The data sample corresponds to a total integrated luminosity of $138\,\mathrm{fb}^{-1}$ recorded in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ by the CMS experiment at the LHC in 2016–2018. A signal model independent event categorization scheme is developed to probe multiple scenarios of the above-mentioned signal models. This scheme is based on the lepton charge and flavor multiplicities, invariant masses of lepton pairs, as well as on the multiplicity of b quark jets, and various kinematic variables. In the absence of any significant deviation from the background expectations, lower limits are set on the masses of type-III seesaw heavy fermions at 740 GeV for any branching fraction scenario to Standard Model leptons. Similarly, vector-like taus are excluded for masses below 920 GeV for the doublet model, and in the mass range of 120–180 GeV for the singlet model. In addition, scalar leptoquarks decaying exclusively to a top quark and a charged lepton of any flavor are excluded for masses below 1230 GeV.

Acknowledgements

This dissertation wouldn't have been possible without the support and guidance from many individuals and institutions. First and foremost, I would like to express my gratitude to my supervisor Prof. Sunil Somalwar. His advice and guidance were extremely valuable and helped me to stay calm even in harder times. In addition, I would like to thank the members of our research group at Rutgers University, especially Halil Saka, who shared the passion for soccer and Nutella crêpes, Olena Karacheban, Jay Vora, and Benjamin Rand for their helpful discussions, input, and cheerfulness. I would particularly like to thank Peter Thomassen, my predecessor in our research group, who introduced me to this wonderful team and thus, made this journey possible. Special thanks go to Prof. Scott Thomas for his consultancy on the theoretical side of particle physics.

Similarly, I want to thank our colleagues from the group of Prof. Sourabh Dube at IISER Pune for their close collaboration over the last years, specifically Angira Rastogi, Anshul Kapoor, Shubhanshu Chauhan, and Kunal Kothekar. Furthermore, I would like to thank my colleagues from the HCAL calibration team for their collaboration on the interesting detector work, especially Prof. John Paul Chou, Prof. Yurii Maravin, and Mary Hadley.

I am also deeply grateful for the support of my whole family, in particular my parents Conny and Stefan Heindl, and my devoted partner Alesia Pérez Prada. Without their belief in me, their encouragement, love and sacrifices, I would have not been able to finish this longstanding project.

This dissertation contains material from a CMS Collaboration publication [1] and an internal CMS Collaboration documentation [2], both of which I have co-authored.

In memory of Maria Römmelt and Herbert Heindl.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The idea of elementary components as the building blocks of all objects in the universe started with the atomism in ancient Greece [3]. Over the last 2500 years, this concept has steadily evolved into the Standard Model of particle physics (SM) [4, 5, 6], which has been proven to be one of the grand achievements in science. Since the 1950s, theoretical predictions, as well as experimental observations, worked hand in hand to develop the theory of the SM describing the fundamental interactions and properties of the elementary particles. The success of the SM can be evaluated by the exceptional precision of its theoretical predictions confirmed by experimental results probing the theory from the microscopic level of single elementary particles to the macroscopic scale of the universe [7].

Nonetheless, the Standard Model of particle physics has not been able to provide answers to all questions some discoveries have raised over the last decades. For instance, the SM predicts the neutrino to be massless, but the observation of neutrino oscillations indicates a massive neutrino. A possible solution to this inconsistency is the so-called seesaw mechanism [8], which extends the SM by adding new heavy particles acting as compensators for the very light neutrinos. Similarly, with the discovery of the Higgs boson [9, 10] the question arises if the possibility for a mass mechanism beyond the constraints of the SM Higgs sector can exist. Such a behavior can be achieved by the so-called vector-like fermions [11], which can be detected through their mixings with SM fermions. Furthermore, the unification of the electromagnetic and weak force to the electroweak interaction prompts the idea of a

further unification with the remaining strong force. Some of these Grand Unified Theories predict the existence of so-called leptoquarks, a new particle combining leptonic and baryonic characteristics [12].

To further probe the SM and the predictions of these theories beyond the Standard Model, the Large Hadron Collider (LHC) has been built as the latest experimental apparatus of its type [13]. This massive accelerator collides charged particles at very high energies and detectors around the point of collision reconstruct the scattered particles with very high accuracy. In this dissertation, we conducted a search for the type-III seesaw mechanism, vector-like leptons, and leptoquarks using proton-proton collision data collected by the CMS detector [14] at the LHC from 2016 to 2018. The focus lies on events with multilepton signatures, containing at least 3 charged leptons (electrons, muons, or hadronically decaying tau leptons). In addition, the results are shown in an inclusive table scheme allowing a simple reinterpretation of this result in the context of other signal models.

The dissertation starts with the principles of the Standard Model of particle physics in Chapter 2. This also includes an introduction into the type-III seesaw mechanism, vector-like leptons, and leptoquarks. Following this, Chapter 3 contains a description of the experimental apparatus, in particular the LHC accelerator and the CMS detector. The analysis itself is described in Chapters 4-8 with the results being shown in Chapter 9. The contents of Chapter 2 with the exception of Section 2.1, as well as the content of Chapters 4-9 substantially overlap with a CMS Collaboration publication [1] and an internal CMS Collaboration documentation [2], both of which I have co-authored.

Chapter 2

Theoretical overview

2.1 Standard Model

The Standard Model of particle physics is a relativistic quantum field theory that describes the fundamental interactions and properties of the elementary particles [4, 5, 6]. This includes the electromagnetic, weak and strong force, with the exception of the gravitational force. This section briefly summarizes the structure of the SM and if not explicitly stated otherwise is based on Refs. [4, 5, 6, 15, 16, 17].

Each particle in the SM is represented by a relativistic quantum field, and has either half-integer or integer spin values (in units of the Planck constant $\hbar$) assigned. Particles with half-integer spin are called fermions, particles with integer spin are referred to as bosons.

The fermionic particles are further divided into the group of leptons and the group of quarks, each of them containing six particles, subdivided into three so-called flavors. Leptons are split into the electron, muon and tau flavor:

$$\begin{pmatrix} \mathrm{e} \\ \nu_e \end{pmatrix}, \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}, \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} \tag{2.1}$$

Each flavor consists of an electrically charged (-1) lepton ℓ and an accompanying electrically neutral neutrino ν_ℓ. This analysis searches for events with three or more electrically charged leptons (electron, muon, or tau lepton).

The six quarks are grouped in the following flavors:

$$\begin{pmatrix} \text{up (u)} \\ \text{down (d)} \end{pmatrix}, \begin{pmatrix} \text{charm (c)} \\ \text{strange (s)} \end{pmatrix}, \begin{pmatrix} \text{top (t)} \\ \text{bottom (b)} \end{pmatrix} \tag{2.2}$$

All quarks in the upper row are called up-type, holding an electrical charge of $+2/3$, while the quarks in the lower row are called down-type, holding an electrical charge of $-1/3$. As each quark is subject to the strong interaction, it carries an additional color charge (red, green, or blue). In the SM, quarks form color neutral bound states through the process of hadronization with integer electrical charge, and are therefore not observed as individual particles. These particles are called baryons, if it consists of three quarks, or mesons, if two quarks form such a bound state.

Each fermion has its own antiparticle, which holds opposite charge, spin, lepton or baryon number, and color. Similar to the electrical charge, the lepton and baryon numbers are conserved quantum numbers in the SM.

Each of the three fundamental forces is represented by a gauge invariant field. The particles acting as the force carriers are the so-called gauge bosons. The electromagnetic force is invariant under the U(1) gauge group and mediated by the massless and electrically neutral photon (γ). The electromagnetic interaction happens between all charged particles. The electrically charged $W^\pm$ boson and the electrically neutral Z boson are both massive and carry the weak force, which is invariant under the SU(2) gauge group and acts on all flavored particles. The electromagnetic and the weak force have been unified into the combined theory of the electroweak interaction. The strong force is invariant under the SU(3) gauge group and acts on colored particles. It is mediated by the massless gluon (g), which has one of eight different color-anticolor combinations.

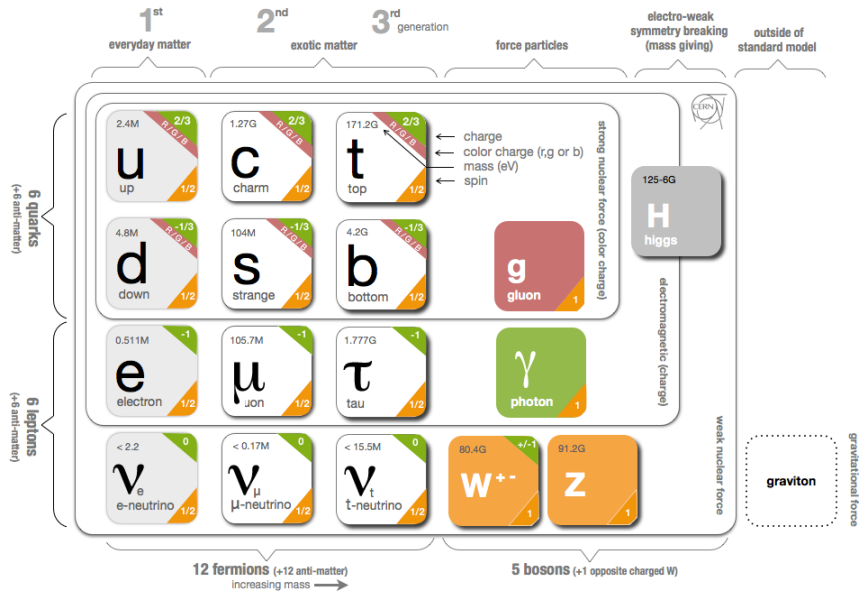

Figure 2.1: The elementary particles in the Standard Model [18].

The Higgs (H) boson has been the latest observed particle in the SM [9, 10]. Without the Higgs mechanism, all particles in the SM would be massless and degenerate, which contradicts the observations of massive fermions and gauge bosons. The interaction of the particles with the Higgs field, including the Higgs boson itself, generates the different masses between those particles. The photon and the gluon don't couple to the Higgs field and are therefore massless. A summary of all particles in the SM is shown in Figure 2.1.

Although the predictions of the SM on electromagnetic, weak and strong interactions have been confirmed with extraordinary precision, some observations raise open questions the SM can not address:

- Within the SM the neutrinos are assumed to be massless. However, the observation of neutrino oscillations between the different lepton flavors implies massive neutrinos [19].
- The Higgs boson was observed with a mass around the electroweak scale. However, the mass of the Higgs boson was expected to be in the order of 10^{15} GeV due to

corrections coming from top quark loops. The SM can not answer this so called Hierarchy Problem, i.e. why these values differ by several orders of magnitude [20].

- The theory of the SM does not account for the gravitational force, which is only described by General Relativity.
- The Standard Model can not present a suitable candidate for dark matter [21].

While the last two questions don't necessarily have to lie within the scope of the SM, the first two issues can create a precarious situation regarding the completeness of the Standard Model. Multiple theoretical models have been developed that can remedy these issues. Such extensions of the SM predict new particles and can therefore be probed with experiments like the LHC [13]. The neutrino mass problem can be explained by the so-called seesaw mechanism [8], while the Hierarchy Problem can be addressed by the existence of vector-like leptons [11] or leptoquarks [12]. Those three signal models are described in the following sections.

2.2 Type-III seesaw fermions

The seesaw mechanism introduces new heavy particles in order to account for the expected smallness of neutrino masses and flavor mixing [22, 23, 24, 25, 26, 27, 28, 29, 30]. Within the type-III seesaw model, the neutrino is considered a Majorana particle whose mass arises via the mediation of new massive fermions. These massive fermions are an SU(2) triplet of heavy Dirac charged leptons $\Sigma^{\pm}$, and a heavy Majorana neutral lepton Σ^{0}. At the LHC, these massive fermions may be pair-produced through electroweak interactions in both $\Sigma^{\pm}\Sigma^{\mp}$ and $\Sigma^{\pm}\Sigma^{0}$ pairs and decay to W, Z, H boson and SM lepton pairs via their mixings, yielding 27 distinct signal production and decay modes [8]. A complete decay chain example would be $\Sigma^{\pm}\Sigma^{0} \rightarrow \mathrm{W}^{\pm}\nu\mathrm{W}^{\pm}\ell^{\mp} \rightarrow \ell^{\pm}\nu\nu\ell^{\pm}\nu\ell^{\mp}$.

In this analysis, the $\Sigma^{\pm,0}$ are taken to be degenerate in mass and their decays to be prompt, rendering the interpretations independent of the precise mixing angle values. Electroweak

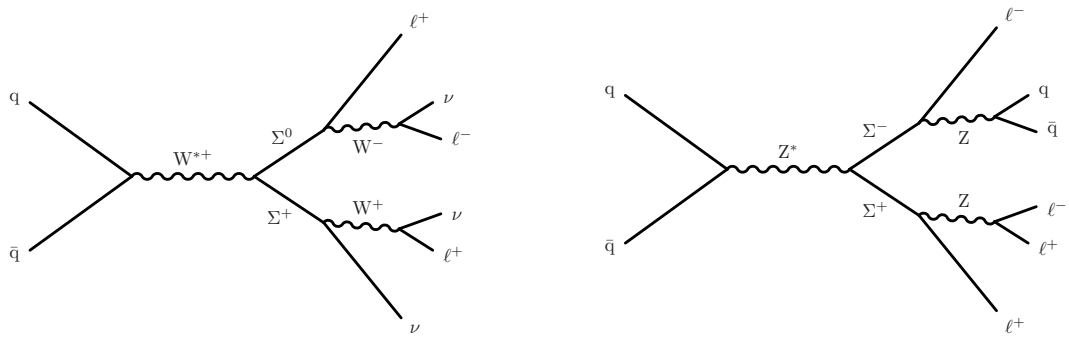

Figure 2.2: Example diagrams illustrating production and decay of seesaw type-III heavy fermion pairs at the LHC.

and low energy precision measurements enforce an upper bound on the mixing angles $V_\ell < 10^{-4}$ across all lepton flavors, which allows for prompt decays of heavy fermions at mass ranges accessible to current collider experiments [31, 32, 33, 34]. The free parameters of this model are the Σ mass, and the Σ decay branching fractions to SM lepton flavors $\mathcal{B}_\mathrm{e}$, $\mathcal{B}_\mu$, and $\mathcal{B}_\tau$, where $\mathcal{B}_\mathrm{e} + \mathcal{B}_\mu + \mathcal{B}_\tau = 1$.

Two diagrams exemplifying the production and decay of $\Sigma\Sigma$ pairs that result in multilepton final states are shown in Figure 2.2. The most stringent limits on the type-III seesaw model comes from a search conducted by the CMS Collaboration [35]. The search excluded type-III seesaw fermions with masses below 880 GeV in the lepton flavor-democratic scenario, where $\mathcal{B}_\mathrm{e} = \mathcal{B}_\tau = \mathcal{B}_\tau$.

2.3 Vector-like taus

Vector-like fermions are new particles whose left- and right-handed components transform similarly under the SM gauge symmetries, and hence their masses can be decoupled from the SM Higgs sector [36]. In particular, vector-like leptons can be produced via SU(2) gauge interactions at the LHC and that can yield signatures with energetic multileptons. The vector-like leptons may account for the mass hierarchy between the different generations of particles in the SM via their mixings with the SM leptons, and arise in a wide variety of models ranging from supersymmetry to grand unification [37, 38, 39, 40, 41, 42].

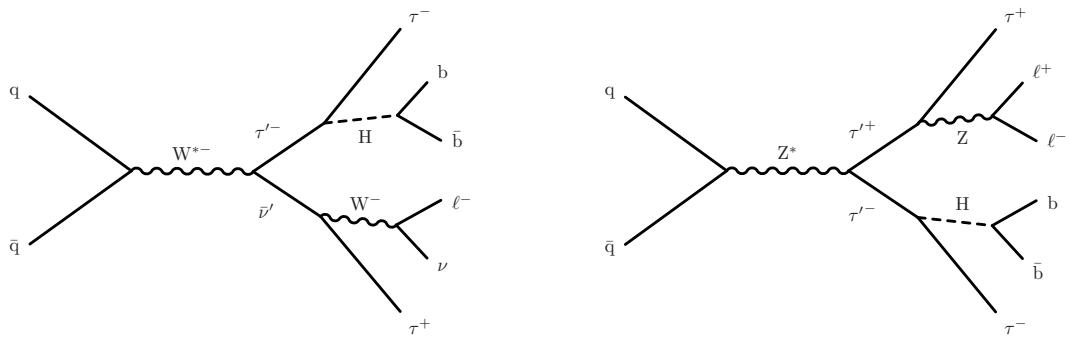

Figure 2.3: Example diagrams illustrating production and decay of doublet (left) and singlet (right) vector-like tau pairs at the LHC.

In this analysis, we consider SU(2) singlet and doublet vector-like leptons extensions to the SM with preferential coupling to the SM tau lepton, where the new vector-like particles are assumed to be mass-degenerate at tree level, and heavier than the SM gauge bosons [11, 43]. The singlet model introduces only a vector-like charged tau (anti-tau) lepton τ'^{-} (τ'^{+}), which can be pair produced via the $\mathrm{pp} \to \mathrm{Z}/\gamma^{\star} \to \tau'^{+}\tau'^{-}$ process at the LHC, and decays into $\mathrm{Z}\tau$, $\mathrm{H}\tau$, and $\mathrm{W}\nu$ pairs. In contrast, the doublet vector-like lepton model also has an accompanying vector-like tau neutrino (anti-neutrino) ν' ($\overline{\nu}'$), and the production modes also include the additional $\mathrm{pp} \to \mathrm{Z} \to \nu'\overline{\nu}'$ and $\mathrm{pp} \to \mathrm{W}^{+}(\mathrm{W}^{-}) \to \tau'^{+}\nu'(\tau'^{-}\overline{\nu}')$ processes. In the doublet vector-like lepton model, the neutral vector-like leptons always decay to a $\mathrm{W}^{\pm}\tau^{\mp}$ pair and the charged vector-like leptons decay to $\mathrm{Z}\tau$ or $\mathrm{H}\tau$ pairs. The decays of vector-like leptons are mediated via their mixings with the SM leptons through Yukawa interactions, and depend on the masses of vector-like leptons as well as the masses of SM bosons they decay into. Electroweak precision data allows the mixing between vector-like leptons and SM leptons $V' < 10^{-2}$, permitting prompt decays of vector-like leptons for mass values in the neighborhood of the electroweak scale [44, 45]. Similar to the type-III seesaw model, we assume prompt decays of vector-like leptons and our analysis is insensitive to the precise values of the actual mixing angles.

Two diagrams exemplifying the production and decay of vector-like lepton pairs that result in multilepton final states are shown in Figure 2.3. The most stringent constraints on the vector-like tau doublet model are from a search conducted by the CMS Collaboration [46], which excludes such particles in the mass range of 120–790 GeV. There are, so far, no direct

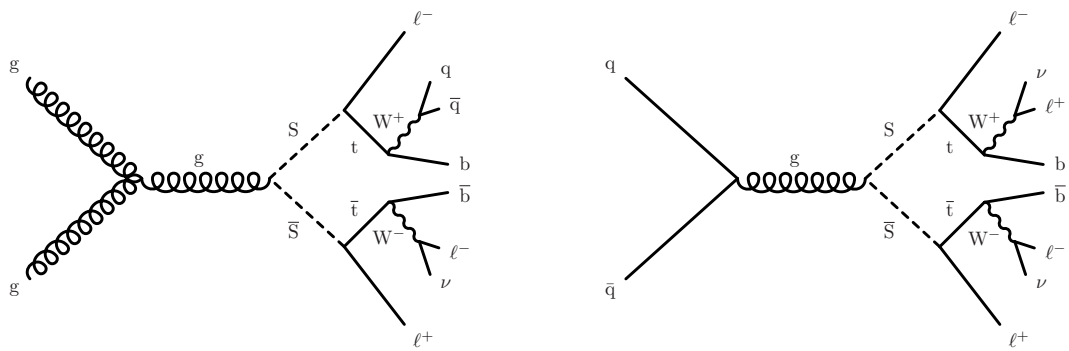

Figure 2.4: Example diagrams illustrating production and decay of leptoquark pairs at the LHC.

constraints on the vector-like tau singlet model from any of the LHC experiments. The L3 Collaboration at the LEP experiment placed a lower bound of $\sim 100\,\mathrm{GeV}$ on additional heavy leptons [47].

2.4 Leptoquarks

Leptoquarks are color-triplet scalar or vector bosons that carry both non-zero baryon and lepton quantum numbers and fractional electric charge [48]. Such particles commonly emerge in grand unification inspired theories [49, 50, 51] and certain supersymmetry models [52, 53]. Depending on the nature of the Yukawa coupling, such leptoquarks are expected to decay either to an up-type quark and a charged lepton or to a down-type quark and a neutrino, with branching fractions β and $1-\beta$, respectively.

In this analysis, we consider scalar leptoquarks (S) with $\beta = 1$, an electric charge of $-1/3$ and a non-zero Yukawa coupling to the top quark and to any of the charged leptons, $\lambda_{\mathrm{t}\ell}$, where $\ell = \mathrm{e}, \mu, \tau$ [12]. We assume that only one flavor of charged lepton coupling dominates at a time, leading to leptoquark branching fractions $\mathcal{B}_{\mathrm{e}} = 1$, $\mathcal{B}_{\mu} = 1$, and $\mathcal{B}_{\tau} = 1$. At the LHC, top-philic scalar leptoquarks are expected to be pair-produced via strong interactions independent of the unknown Yukawa coupling $\lambda_{\mathrm{t}\ell}$, and therefore, the production cross-section solely depends on the leptoquark mass. Furthermore, the leptoquark decays are assumed to be prompt with $\lambda_{\mathrm{t}\ell} < 10^{-1}$, in compliance with the bounds on such Yukawa couplings by leptonic Z decays [12, 54]. Similar to the magnitude of the mixing angles in the

type-III seesaw and vector-like lepton models, the analysis is independent of the absolute magnitude of the leptoquark Yukawa coupling beyond the promptness assumption.

Two diagrams exemplifying the production and decay of leptoquark pairs that result in multilepton final states are shown in Figure 2.4. The most stringent constraints on scalar leptoquarks with 100% branching fraction to a top quark and first, second, or third generation charged lepton are set by the ATLAS Collaboration, excluding such particles with masses below 1.48 TeV, 1.47 TeV and 1.43 TeV respectively [55, 56].

Chapter 3

Experimental apparatus

This chapter briefly outlines the experimental apparatus, in particular the Large Hadron Collider and the CMS detector, and if not explicitly stated otherwise is based on Refs. [13, 14, 57].

3.1 The Large Hadron Collider

The analysis presented in this dissertation is based on events from proton-proton (pp) collisions generated by the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN). It resides in a tunnel 100 m underground across the Swiss-French border and has a circumference of 26.7 km. The protons are produced from ionized hydrogen gas and accelerated through a complex of multiple linear and circular accelerators before reaching the LHC ring. At first, protons are accelerated to 50 MeV in the Linear Accelerator 2 (LINAC 2), which injects the beams into the Booster ring. After reaching an energy of 1.4 GeV the protons are transported to the Proton Synchrotron (PS), where they are accelerated up to 25 GeV. Finally, after reaching 450 GeV in the Super Proton Synchrotron (SPS) the beams get injected into the LHC. The whole complex is shown in Figure 3.1. Two proton beams split into 2808 bunches with 1.2×10^{11} protons each are then accelerated to a final energy of 6.5 TeV each. They collide every 25 ns at four interaction points with a total collision energy $\sqrt{s} = 13\,\text{TeV}$. At these interaction points

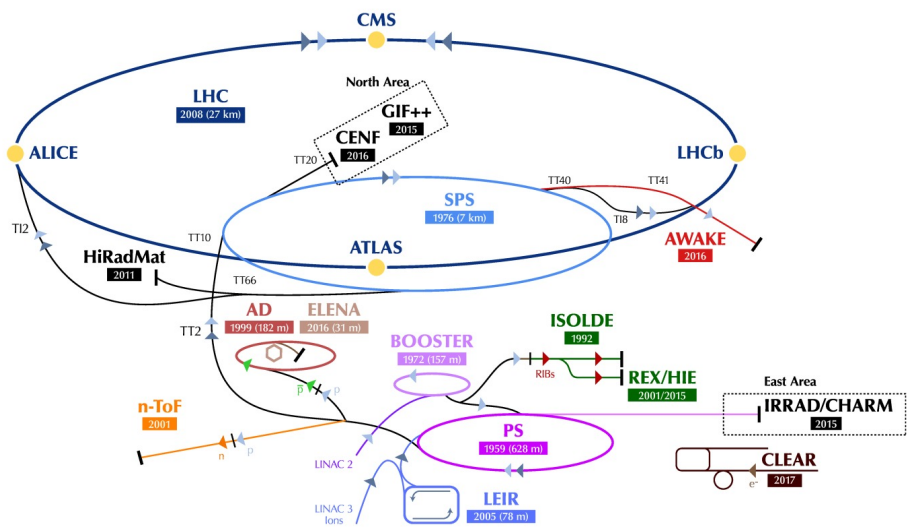

Figure 3.1: The LHC accelerator complex at CERN [61].

four experiments are located: A Large Ion Collider Experiment (ALICE) [58], A Toroidal LHC Apparatus (ATLAS) [59], the Compact Muon Solenoid (CMS) detector [14], and the Large Hadron Collider beauty (LHCb) detector [60].

The LHC consists of 1232 superconducting dipole and 392 superconducting quadrupole magnets cooled down to 1.9 K using superfluid helium and providing a magnetic field of 8.3 T to bend the proton beams around the ring structure. The instantaneous luminosity of the collider is given by

$$\mathcal{L} = \frac{N_b^2 n_b f \gamma}{4\pi\epsilon\beta^*} F \tag{3.1}$$

where N_b is the number of protons per bunch, n_b is the number of bunches per beam, f is the revolution frequency, γ is the relativistic gamma factor, ϵ is the normalized transverse beam emittance, β^* is the value of the beta function at the interaction point, and F is the geometrical reduction factor due to the beam-beam crossing angle at the interaction point. At the LHC the peak instantaneous luminosity is $\sim 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

3.2 The CMS detector

The Compact Muon Solenoid (CMS) experiment is one of the two general purpose detectors located at the LHC. The main part of the CMS detector is a superconducting solenoid, inside of which a silicon pixel and strip tracking system, a lead tungstate crystal electromagnetic calorimeter and a brass-scintillator hadron calorimeter are placed. Gas-ionization muon chambers and an iron return yoke are located outside the magnet.

The right-handed coordinate system of the detector is centered around the interaction point as its origin. The x-axis points radially inward towards the center of the LHC ring, the y-axis in the upward direction of the detector, and the z-axis along the counterclockwise proton beam direction. In cylindrical coordinates the azimuthal angle ϕ is measured in the transverse (x-y) plane, while the polar angle θ is measured in the y-z plane. As the polar angle θ is not a Lorentz-invariant quantity, the pseudorapidity η is defined as:

$$\eta = -\ln \tan(\theta/2) \tag{3.2}$$

An illustration of the CMS detector is shown in Figure 3.2.

3.2.1 Tracking system

The tracking system of the CMS detector provides a precise reconstruction of charged particle trajectories within $|\eta| < 2.5$ [64]. It consists of silicon-based p-n junctions assembled as separate pixel and strip detectors. When a charged particle passes through a layer of these junctions, it ionizes the silicon atoms, which results in the measurement of an electric current in this layer. The pixel detector is 97 cm in length and 30 cm in diameter, consisting of 3 barrel layers and 2 endcap disks. The pixel size is $100\,\mu\text{m} \times 150\,\mu\text{m}$. With its 66 million channels the pixel detector achieves a spatial resolution of $10\,\mu$m. The silicon strip detector is further away from the interaction point, is instrumented with 9.6 million channels, and is 5.5 m in length and 2.4 m in diameter. It is assembled in four inner barrel layers, six

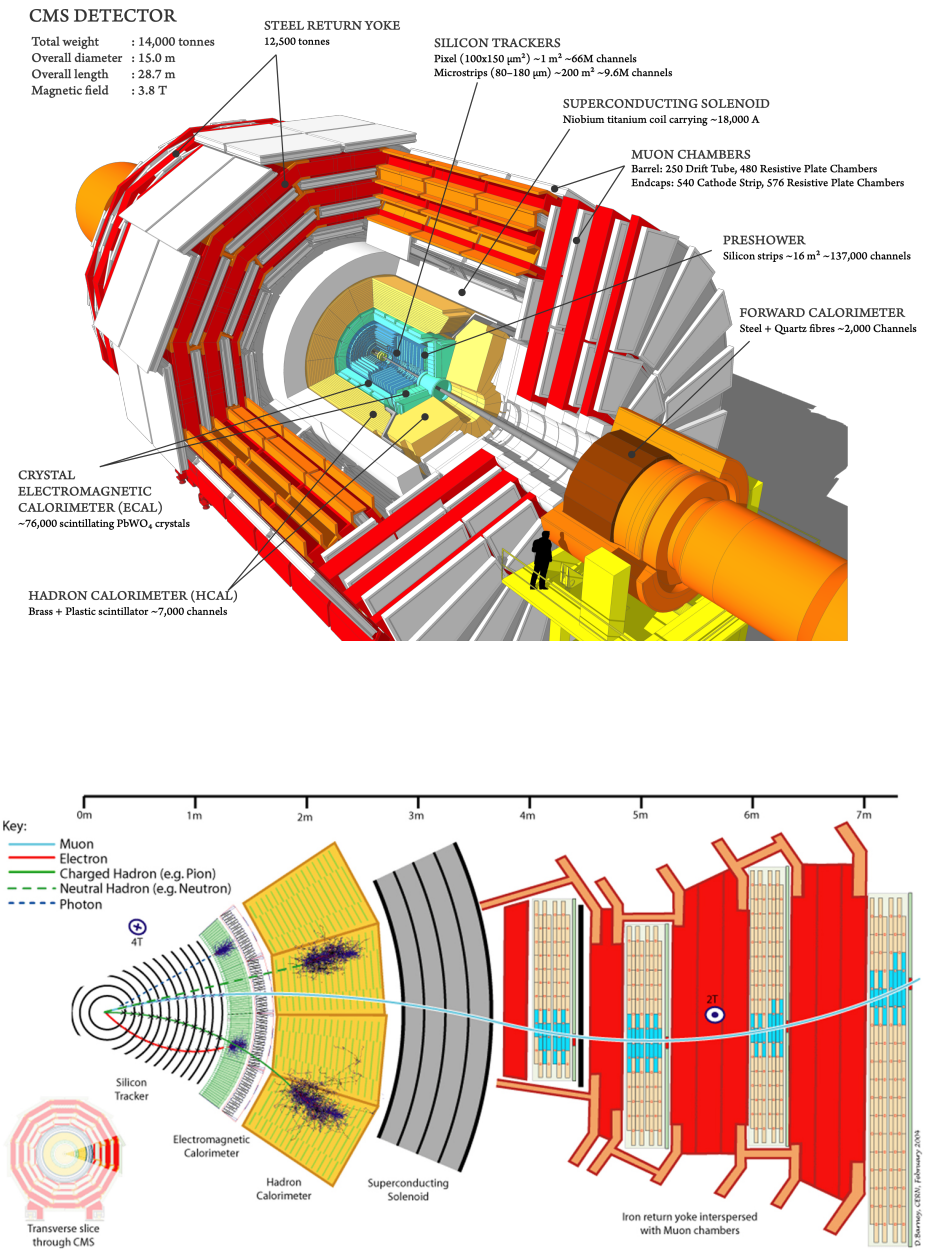

Figure 3.2: An illustration of the CMS detector with its subsystems (upper) [62] and a transverse slice through the CMS detector showing the detection of different particles (lower) [63].

outer barrel layers, three inner disks, and nine endcap disks. Each cell in the strip detector has a size of 10 cm × 80 μm. The strip detector achieves a spatial resolution of 20–70 μm, depending on the pseudorapidity of the reconstructed track.

3.2.2 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) aims to accurately detect electrons and photons by measuring the energy of charged particles [65]. It is composed of $\sim$ 76000 lead tungstate ($PbWO_4$) crystals, as this material combines a high density with optical transparency. Charged particles traveling through the crystals lose their energy in a showering process due to bremsstrahlung, ionization and e^+e^- pair production. The crystals get energetically excited and scintillate proportional to the amount of energy deposited, which is detected by photodiodes and -triodes to infer the energy of the incoming particle.

In the barrel ($|\eta| < 1.479$) section of the ECAL 61200 crystals are installed, each with front face dimensions of 22 mm × 22 mm and a length of 23 cm, which corresponds to $\sim$ 25 radiation lengths. The endcap ($1.479 < |\eta| < 3.0$) section is instrumented with $\sim$ 14600 crystals, each with front face dimensions of 28.6 mm × 28.6 mm and a length of 22 cm. In addition, the endcap unit consists of a 20 cm thick preshower detector in front of the crystal configuration, which is assembled out of 2 layers of lead and 2 layers of 2 mm wide silicon strips. This additional detector increases the granularity of the ECAL endcap regions helping to distinguish single photons from neutral pions decaying to two photons.

3.2.3 Hadron calorimeter

The hadron calorimeter (HCAL) is designed to measure the energy of hadron jets [66]. Unlike the ECAL, the HCAL is not a homogeneous, but a sampling calorimeter. It consists of a passive medium, which absorbs the energy of the incoming particle, and an active medium, which detects the particle shower produced during the energy absorption. The HCAL is divided into a barrel (HB), an endcap (HE), a forward (HF), and an outer (HO) detector.

The barrel (endcap) sections are instrumented with 15 (17) brass absorber plates and a steel absorber plate as the inner- and outermost layer. Between those absorber layers, a plastic scintillator with a thickness of 3.7 mm is assembled in a tile pattern. The light emitted by those scintillators is collected by wavelength-shifting fibers and channeled to hybrid photodiodes. The HB unit ($|\eta| < 1.3$) is divided into towers with a granularity of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. For the HE unit ($1.3 < |\eta| < 3.0$), the tower cell width varies ($0.87 < \Delta\phi < 0.175$, $0.087 < \Delta\eta < 0.35$) depending on the pseudorapidity.

The outer calorimeter is placed outside of the solenoid magnet and extends the depth of the HB system to improve the energy reconstruction accuracy. The HO system ($|\eta| < 1.3$) utilizes the material of the magnet coil as the absorber material and contains two layers of 10 mm thick scintillator tiles with an additional iron layer in between. Therefore, the combination of the HB, HE, and HO systems covers between 10.0 and 11.8 effective interaction lengths for each of the $\sim$ 7000 readout channels. To extend the calorimeter coverage to the forward region ($3.0 < |\eta| < 5.0$) the forward calorimeter is installed. It sits at a distance of 11.15 m to the interaction point on both sides and consist of quartz fibers with a length of 130 cm and 165 cm ($\sim$ 10 interaction lengths) inside of steel wedges. The Cherenkov light emitted by the shower particles traversing the quartz fibers is collected by $\sim$ 2000 photomultipler tubes, each covering a tower of size $\Delta\eta \times \Delta\phi = 0.175 \times 0.175$.

3.2.4 Superconducting solenoid

The central piece of the CMS detector is a 12.5 m long superconducting solenoid magnet with a diameter of 6 m [67]. It consist of niobium-titanium coils cooled down to 4.5 K using liquid helium. The solenoid produces a magnetic field of up to 3.8 T and is supplemented by an iron return yoke weighing 12500 tonnes to confine the magnetic flux to the whole volume of the detector. The magnet, together with the tracking and muon systems, allows for the high precision momentum and charge reconstruction of charged particles produced after the collisions by measuring the curvature of the particle trajectory.

3.2.5 Muon system

Muons can easily transverse the calorimeter system due to their lower loss of energy induced through bremsstrahlung compared to electrons. Hence, a separate muon identification system is located outside the solenoid magnet and calorimeters [68].

In the barrel ($|\eta| < 1.2$) section, 250 drift tube chambers (DT) are installed in 4 layers embedded in the return yoke. This leads to a spatial resolution of 77–123 μm in the transverse and 140–393 μm in the y-z plane. Each chamber covers an approximate area of $3.0\,\mathrm{m} \times 2.5\,\mathrm{m}$ and is filled with argon and carbon dioxide gas. Incoming muons ionize the gas, which results in an ion avalanche towards the cathode, allowing the reconstruction of the muon track.

In the endcap ($0.9 < |\eta| < 2.4$), 540 cathode strip chambers (CSC) are assembled in 6 layers, partially overlapping with the DT system. Similar to the DT chambers, each CSC is filled with argon and carbon dioxide gas. The spatial resolution of the CSC system is $\sim 80\,\mu$m in the transverse plane.

Each of the four DT layers as well as the three CSC layers with $|\eta| < 1.6$ are accompanied by more than 1000 resistive plate chambers (RPC). A RPC consists of two bakelite resistive plates separated by a 2 mm gap filled with freon gas. The plates are coated with a conductive graphite and an insulating foil, and each set of plates is equipped with metallic strips. The avalanche of electrons induced by the ionization of transversing muons is picked up by the strips to estimate the trajectory of the particle. Although the spatial resolution of the RPC system is relatively coarse (0.8–1.3 cm), the fast response time of < 3 ns is used for a precise timing determination of the muon tracks.

3.2.6 Trigger system

Around 1 billion pp collisions per second are performed at the LHC. However, only a fraction of those events can be stored due to the readout and computing time of the detector. Therefore, a multi-level triggering system is used to select potentially interesting events by

looking for predefined kinematic signatures [69]. The Level-1 (L1) trigger system performs quick event assessments based on the input from the ECAL and HCAL calorimeter and the muon system at each bunch crossing to reduce the rate to 100 kHz. The L1 trigger latency is $< 3\,\mu$s and the event data is kept in front-end buffers upon a positive L1 decision. All events passing the L1 trigger are forwarded to the high-level trigger (HLT) system which analyzes the full event information, including the tracking system. This step reduces the rate to 100 Hz and the events are transmitted to the permanent storage systems at CERN.

Chapter 4

Data and simulation samples

This chapter discusses the different datasets used in this analysis, as well as the simulated background and signal samples and their generator settings.

The data samples utilized in this search correspond to a total integrated luminosity of $138\,\mathrm{fb}^{-1}$ (36.3, 41.5, and $59.8\,\mathrm{fb}^{-1}$ in years 2016, 2017 and 2018, respectively), recorded in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ during the LHC from 2016–2018 (Run-2 period). A combination of isolated single-muon and single-electron triggers was used with the corresponding transverse momentum (p_{T}) thresholds of 24 and 27 GeV in 2016, 27 and 32 GeV in 2017, and 24 and 32 GeV in 2018. A complete list of the triggers and datasets as used in this analysis are given in Table A.1 in Appendix A.

Event samples from Monte Carlo (MC) simulations are used to estimate the rates of signal and relevant SM background processes. The WZ, $\mathrm{Z}\gamma$, $\mathrm{t\bar{t}Z}$, $\mathrm{t\bar{t}W}$, and triboson backgrounds are generated using MadGraph5_aMC@NLO (2.2.2 in 2016, 2.4.2 in 2017 and 2018 data analyses) [70] at next-to-leading order (NLO) precision. The ZZ background contribution from quark-antiquark annihilation is generated using POWHEG 2.0 [71, 72, 73] at NLO, whereas the contribution from gluon-gluon fusion is generated at leading order (LO) using MCFM 7.0.1 [74]. Backgrounds from Higgs boson production are generated at NLO using POWHEG and JHUGEN 7.0.11 [75, 76, 77, 78]. Simulated event samples for $\mathrm{Z}/\gamma^{\star}$ and $\mathrm{t\bar{t}}$

processes, generated at NLO with MadGraph5_aMC@NLO and POWHEG, respectively, are used for systematic uncertainty studies.

All signal samples are simulated at LO precision. The type-III seesaw and vector-like lepton samples are generated with MadGraph5_aMC@NLO 2.6.1, whereas the leptoquark samples are generated with PYTHIA 8.212 (8.230) in 2016 (2017 and 2018) [79]. The production cross-section for the type-III seesaw signal model is calculated at NLO plus next-to-leading logarithmic precision, assuming that the heavy leptons are SU(2) triplet fermions [80, 81]. Similarly, leptoquark and vector-like lepton cross-sections have been calculated at NLO accuracy [43, 82, 83]. The inclusive production cross-sections for all signal samples are listed in Table 4.1

All background and signal samples in 2016 are generated with the NNPDF3.0 NLO or LO parton distribution functions (PDF), with the order matching that in the matrix element calculations. In 2017 and 2018, the NNPDF3.1 next-to-next-to-leading order PDF [84, 85] are used. Parton showering, fragmentation, and hadronization for all samples are performed using PYTHIA 8.212 (8.230) in 2016 (2017 and 2018) .

Double counted partons generated with PYTHIA and MadGraph5_aMC@NLO are removed using the FxFx [86] or MLM [87] matching schemes. The response of the CMS detector is simulated using dedicated software based on the Geant4 toolkit [88], and the presence of multiple pp interactions in the same or adjacent bunch crossing (pileup) is incorporated by simulating additional interactions, that are both in-time and out-of-time with the hard collision according to the pileup in the data samples.

The simulated samples used in this analysis for all three years are listed in Tables A.2-A.4 in Appendix A. All samples are weighted to match the expected pileup profile of the corresponding dataset, and further corrections are applied to account for data-MC differences in lepton reconstruction, lepton identification, trigger efficiencies, and tagging efficiencies for jets of different flavor origins. Negative event weights are also applied as applicable.

Table 4.1: Total inclusive production cross-sections for type-III seesaw, vector-like tau and leptoquark signal models in pp collisions at 13 TeV.

Type-III seesaw:					
M_Σ (GeV)	$\sigma_{\Sigma\Sigma}$ (pb)	M_Σ (GeV)	$\sigma_{\Sigma\Sigma}$ (pb)	M_Σ (GeV)	$\sigma_{\Sigma\Sigma}$ (pb)
100	$3.43 \cdot 10^{1}$	550	$4.45 \cdot 10^{-2}$	1250	$4.65 \cdot 10^{-4}$
200	$2.71 \cdot 10^{0}$	700	$1.40 \cdot 10^{-2}$	1500	$1.19 \cdot 10^{-4}$
300	$5.77 \cdot 10^{-1}$	850	$5.00 \cdot 10^{-3}$		
400	$1.80 \cdot 10^{-1}$	1000	$1.96 \cdot 10^{-3}$		

Vector-like taus:					
$M_{\tau'}$ (GeV)	$\sigma_{\tau'\tau'}$ (pb)		$M_{\tau'}$ (GeV)	$\sigma_{\tau'\tau'}$ (pb)	
	$SU2$-singlet	$SU2$-doublet		$SU2$-singlet	$SU2$-doublet
100	$1.17 \cdot 10^{0}$	$1.69 \cdot 10^{1}$	550	$1.78 \cdot 10^{-3}$	$2.24 \cdot 10^{-2}$
125	$5.45 \cdot 10^{-1}$	-	600	$1.19 \cdot 10^{-3}$	$1.49 \cdot 10^{-2}$
150	$2.90 \cdot 10^{-1}$	$3.88 \cdot 10^{0}$	650	$8.05 \cdot 10^{-4}$	$1.01 \cdot 10^{-2}$
200	$1.05 \cdot 10^{-1}$	$1.36 \cdot 10^{0}$	700	$5.56 \cdot 10^{-4}$	$6.97 \cdot 10^{-3}$
250	$4.60 \cdot 10^{-2}$	$5.89 \cdot 10^{-1}$	750	$3.90 \cdot 10^{-4}$	$4.89 \cdot 10^{-3}$
300	$2.29 \cdot 10^{-2}$	$2.91 \cdot 10^{-1}$	800	$2.77 \cdot 10^{-4}$	$3.47 \cdot 10^{-3}$
350	$1.25 \cdot 10^{-2}$	$1.57 \cdot 10^{-1}$	850	$2.00 \cdot 10^{-4}$	$2.49 \cdot 10^{-3}$
400	$7.20 \cdot 10^{-3}$	$9.07 \cdot 10^{-2}$	900	$1.45 \cdot 10^{-4}$	$1.81 \cdot 10^{-3}$
450	$4.36 \cdot 10^{-3}$	$5.49 \cdot 10^{-2}$	950	$1.06 \cdot 10^{-4}$	$1.32 \cdot 10^{-3}$
500	$2.74 \cdot 10^{-3}$	$3.45 \cdot 10^{-2}$	1000	$7.80 \cdot 10^{-5}$	$9.71 \cdot 10^{-4}$

Leptoquarks:					
M_S (GeV)	σ_{SS} (pb)	M_S (GeV)	σ_{SS} (pb)	M_S (GeV)	σ_{SS} (pb)
200	$6.35 \cdot 10^{1}$	700	$6.45 \cdot 10^{-2}$	1400	$3.67 \cdot 10^{-4}$
300	$8.40 \cdot 10^{0}$	800	$2.68 \cdot 10^{-2}$	1700	$5.90 \cdot 10^{-5}$
400	$1.81 \cdot 10^{0}$	900	$1.19 \cdot 10^{-2}$	2000	$1.04 \cdot 10^{-5}$
500	$5.01 \cdot 10^{-1}$	1000	$5.51 \cdot 10^{-3}$		
600	$1.71 \cdot 10^{-1}$	1200	$1.35 \cdot 10^{-3}$		

Chapter 5

Object definitions

In this chapter the reconstruction and identification of different particles with the CMS detector is covered. We use the so-called particle-flow (PF) algorithm [89] to identify every particle as a reconstructed object in an event by using the information from the various parts of the CMS detector. We use electron, muon, tau, jet, and missing transverse momentum objects in this analysis as detailed below.

5.1 Electrons and muons

Muons are reconstructed by matching tracks in the inner tracker system with the tracks in the muon chambers. Muons used in this analysis have to satisfy a selection of $|\eta| < 2.4$ and $p_{\mathrm{T}} > 10\,\mathrm{GeV}$. To suppress muons originating from hadronic activity additional fit and matching quality criteria for the tracks in the tracking and muon system are required [90].

In this analysis, we use electrons with $|\eta| < 2.4$ and $p_{\mathrm{T}} > 10\,\mathrm{GeV}$. They are reconstructed by matching the path of a charged particle in the tracking system with energy clusters deposited in the ECAL. The energy of electron objects is determined from the momentum as specified by the tracker, the energy of the accompanying ECAL deposit, and the energy of all bremsstrahlung photons originating from the electron track. Electrons originating from photon conversions in detector material or from hadronic activity are suppressed by im-

Table 5.1: Electron and muon relative isolation working points.

	Loose	Tight
Electron I_{rel}	$< 100\%$	5–15% (inversely vs. p_{T})
Muon I_{rel}	$< 100\%$	$< 15\%$

posing shower shape, charge measurement and track quality requirements [91]. All selected electrons within a cone of $\Delta\mathrm{R} < 0.05$ around a selected muon are discarded to suppress contributions due to bremsstrahlung from muons, where $\Delta\mathrm{R}$ is the distance between a given pair of objects in the η-ϕ plane:

$$\Delta\mathrm{R} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{5.1}$$

Electrons and muons from hadronic activity originate mainly from semi-leptonic heavy flavor decays and are therefore accompanied by a hadronic jet. As a result, such leptons can be suppressed by imposing an isolation requirement. We use the so-called relative isolation I_{rel}, which is defined as

$$I_{rel} = I_{tot}/p_{\mathrm{T}} \tag{5.2}$$

where I_{tot} is the scalar sum of the transverse momentum of all charged hadrons, neutral hadrons, and photons within a cone of $\Delta\mathrm{R}$ around the lepton candidate. This quantity is corrected against contributions from pileup [90, 91]. The cone size for electrons is set to 0.3, for muons it is set to 0.4. We define two isolation selections, loose and tight, which are necessary for a data-driven background estimation method described in Section 7.1.1. The final object definition is always based on the tight isolation selection. The isolation working points for electrons and muons are given in Table 5.1.

In addition to these identification and isolation requirements, electrons and muons are required to pass certain selections on their displacement compared to the primary interaction point, as this also reduces the contribution from leptons originating from heavy flavor decays. Electrons must satisfy $|d_{\mathrm{z}}| < 0.1\,\mathrm{cm}$ and $|d_{\mathrm{xy}}| < 0.05\,\mathrm{cm}$ in the ECAL barrel ($\eta < 1.479$) and $|d_{\mathrm{z}}| < 0.2\,\mathrm{cm}$ and $|d_{\mathrm{xy}}| < 0.1\,\mathrm{cm}$ in the ECAL endcap ($\eta > 1.479$), where d_{z} and d_{xy} are the longitudinal and transverse impact parameters of electrons with

Table 5.2: Electron and muon displacement selections in 2016, 2017, and 2018.

	2016	2017	2018
$\mathrm{SIP}_{3\mathrm{D}}$	< 10.0	< 12.0	< 9.0
DeepCSV	< 0.6	< 0.4	< 0.3

respect to the primary interaction point. Similarly, muons must satisfy $|d_{\mathrm{z}}| < 0.1\,\mathrm{cm}$ and $|d_{\mathrm{xy}}| < 0.05\,\mathrm{cm}$. Additional lepton $\mathrm{SIP}_{3\mathrm{D}}$ and lepton DeepCSV requirements as summarized in Table 5.2 are placed on top of the object selection criteria given above. Lepton $\mathrm{SIP}_{3\mathrm{D}}$ is the 3-dimensional impact parameter significance of a given lepton with respect to the primary interaction point in the event, the so-called primary vertex (PV). Lepton DeepCSV is defined as the DeepCSV (b quark tagging) score [92] of the PF jet [89, 93, 94] nearest to the lepton with $\Delta\mathrm{R} < 0.4$. In the rare occurrence that no or multiple such jets are found per lepton, DeepCSV is set to be identically zero or the jet with the closest $\Delta\mathrm{R}$ is utilized. No identification or other selection criteria are applied to these jets other than a $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ selection.

5.2 Hadronic taus

Hadronically decaying tau lepton candidates (τ_{h}) are based on PF jet objects and reconstructed with the hadrons-plus-strips algorithm [95]. This algorithm combines 1 or 3 tracks with energy deposits in both calorimeters to identify the corresponding one-prong or three-prong tau decay modes. Taus are required to satisfy $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.3$.

To distinguish τ_{h} from jets originating from the hadronization of quarks or gluons, and from electrons or muons, the DeepTau [96] identification algorithm is used. All τ_{h} candidates within a cone of $\Delta\mathrm{R} < 0.5$ around a loose electron or muon candidate are rejected to further suppress the misidentification of light leptons as taus. Similar to electrons and muons, tau candidates must satisfy $|d_z| < 0.2\,\mathrm{cm}$ and a maximum lepton DeepCSV of 0.8 to suppress misidentified taus from heavy flavor decays. Taus also have a loose and tight selection, which is based on two working points of the anti-jet discriminator of the DeepTau identification.

5.3 Jets and missing transverse momentum

Jets used in this analysis are clustered using the AK4 algorithm [93, 94] and have to satisfy $p_{\mathrm{T}} > 30\,\mathrm{GeV}$, $|\eta| < 2.4$. To suppress the contribution of charged particles from pileup on reconstructed jets, a charged hadron subtraction technique is used, removing the energy of those particles that do not originate from the PV [89]. In addition, the jet momentum is also corrected for the impact of neutral particles from pileup [97]. Each jet has to fulfill additional quality criteria to reduce the impact of instrumental effects or reconstruction failures [98]. Furthermore, the jet candidate needs to be outside a cone of $\Delta\mathrm{R} < 0.4$ around a loose electron, muon or tau object. A subset of the final jet candidates originating from b quarks are identified using the DeepCSV b tagging algorithm [92].

The missing transverse momentum ($p_{\mathrm{T}}^{\mathrm{miss}}$) is calculated as the magnitude of the negative vectorial sum of the transverse momenta of all PF candidates. It is reconstructed using the Puppi algorithm with type-1 corrections [99], which includes the applied jet energy corrections given above.

Chapter 6

Event selection and categorization

In this chapter we describe the process of the event selection and categorization to maximize the sensitivity for signal models beyond the SM, including those probed specifically in this analysis. This goal is reached by creating a multidimensional classification based on various variables and quantities as depicted below.

This analysis targets events with 3 or more leptons recorded by the CMS detector during LHC Run-2. Due to trigger considerations, each event is required to have at least one muon with $p_{\mathrm{T}} > 26(29)$ GeV in 2016 and 2018 (2017) or at least one electron with $p_{\mathrm{T}} > 30(35)$ GeV in 2016 (2017 and 2018) that matches to a corresponding trigger object with $\Delta\mathrm{R} < 0.2$. In order to remove overlapping events in electron and muon primary datasets, events that trigger both electron and muon triggers are obtained from the muon primary dataset only. Throughout this analysis, we prioritize light leptons over taus, and consider the 4 leading p_{T} leptons in events with 5 or more lepton objects. This results in seven orthogonal channels (4L, 3L1T, 3L, 2L2T, 2L1T, 1L3T, 1L2T) as detailed in Table 6.1, where L represents the light lepton and T the hadronic tau lepton multiplicity. Furthermore, two additional 2 lepton channels are defined as 2L (exactly 2 light leptons) and 1L1T (exactly 1 light lepton and 1 tau), which are solely used for various background studies.

Table 6.1: Analysis channels, based on the electron, muon, and tau multiplicities per event.

	1 e/μ	2 e/μ	3 e/μ	≥ 4 e/μ
0 τ_h	-	(2L)	3L	4L
1 τ_h	(1L1T)	2L1T	3L1T	4L
2 τ_h	1L2T	2L2T	3L1T	4L
≥ 3 τ_h	1L3T	2L2T	3L1T	4L

To facilitate the event selections and categorizations in this analysis we define the following variables:

- M_{Z}: The SM Z boson mass, 91 GeV.
- $\Delta\mathrm{R}_{\mathrm{min}}$: Minimum ΔR between all leptons in an event.
- $\Delta\mathrm{R}_{\mathrm{min}}^{\tau_{\mathrm{h}}}$: Minimum ΔR between all hadronic taus in an event.
- N_{j}: Multiplicity of jets satisfying the respective selection criteria in an event.
- N_{j}^{15}: Multiplicity of all AK4 PF jets in the event with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$ and $|\eta| < 2.1$. This variable is only used in light lepton fake rates.
- N_{b}: Multiplicity of b-jets satisfying the respective selection criteria in an event.
- N_{trk}: Multiplicity of tracks with $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$ originating from the PV in the event.
- Scalar sums: We define L_{T} (H_{T}) as the scalar p_{T} sum of all leptons (jets) satisfying the selection requirements. Additionally, the scalar sum of L_{T}, H_{T}, and $p_{\mathrm{T}}^{\mathrm{miss}}$ is defined as S_{T}. $L_{\mathrm{T}} + p_{\mathrm{T}}^{\mathrm{miss}}$ is also used as a quantity of interest. For the signal models considered in this analysis, high signal mass hypotheses give rise to events with high L_{T}, H_{T}, $p_{\mathrm{T}}^{\mathrm{miss}}$ and S_{T}.
- Charge and flavor combinations: We count the number n of distinct opposite-sign (electric charge) same-flavor lepton pairs in an event to define OSSFn as an event quantity. Specific lepton-pairs are labeled as OS (opposite-sign), OSSF (opposite-sign, same-flavor), and OSOF (opposite-sign, other-flavor). Here, other-flavor implies different lepton flavors.

- Invariant and transverse masses: We define M_ℓ as the invariant mass of all leptons in the event, and M_{min} as the minimum invariant mass of all dilepton pairs in the event, irrespective of charge or flavor. The transverse mass for a single lepton i is defined as

$$M_{\text{T}}^{i} = \sqrt{2p_{\text{T}}^{\text{miss}} p_{\text{T}}^{i}[1 - \cos(\vec{p}_{\text{T}}^{\,\text{miss}}, \vec{p}_{\text{T}}^{\,i})]} \tag{6.1}$$

  where p_{T}^{i} is the p_{T} of lepton i.

  We define the M_{OSSF} variable in a given event as the OSSF dielectron or dimuon mass closest to the Z boson mass at 91 GeV, subject to some additional constraints, and label events with M_{OSSF} within 15 GeV of the Z boson mass (76–106 GeV mass window) as OnZ. Events without an OSSF dielectron or dimuon mass within 15 GeV of the Z boson mass are labeled as OffZ. In 3L OSSF1 events, the M_{T} variable is defined as M_{T}^{i} where lepton i is not part of the M_{OSSF} pair. However, in events with three electrons or muons, the M_{OSSF} and M_{T} variables are chosen simultaneously so that the event is OnZ and M_{T} is in the range of 50–150 GeV, if it is kinematically allowed. Similarly, in 4L OSSF2 events with four electrons or muons, the M_{OSSF} is chosen to give the maximum number of non-overlapping OSSF masses within the Z boson mass window.

In order to suppress contributions due to low-mass resonances from mesons and low-ΔR final state radiation, events with $M_{\text{min}} < 12$ GeV, $\Delta\text{R}_{\text{min}} < 0.2$ or $\Delta\text{R}_{\text{min}}^{\tau_{\text{h}}} < 0.5$ are vetoed. These baseline requirements are used throughout this analysis. All regions, which are used in the estimation of major SM backgrounds as described in Chapter 7, are called control regions and are removed in the following categorization schemes. All regions part of the categorization schemes are called signal regions. This allows the complete utilization of all multilepton events collected in Run-2.

Since the analysis aims to be as model independent as possible, each of these 3 or 4 lepton channels are further split into various lepton charge, flavor, mass, and kinematic regions depending on the dominant SM background processes. The primary event classification is performed based on the number of distinct OSSF dilepton pairs in the event considering

all lepton flavors. The allowed values for 4 lepton events are OSSF0, OSSF1, or OSSF2, whereas OSSF2 is disallowed for 3 lepton events.

Subsequently, each OSSF1 and OSSF2 category is further split based on the number of distinct OSSF dielectron or dimuon pairs whose mass is consistent with that of the Z boson in the predetermined window (76–106 GeV), yielding the OnZ category. It is also possible to have two distinct OnZ pairs in a 4L event, labelled as double-OnZ. If the event has no OnZ candidates, it is classified as an OffZ event. If the event has more than one flavor of OSSF pair, dielecton and dimuon pairs are prioritized over the ditau ones. If the OSSF pair is a dielectron or a dimuon pair, BelowZ (< 76 GeV), AboveZ (> 106 GeV) and MixedZ categories are defined, depending on the masses of all light lepton OSSF candidates with respect to the Z window. An event is classified as MixedZ if there are light lepton OSSF pairs in both BelowZ and AboveZ regions. If the OSSF pair is a ditau pair, no resonance is sought after due to the invisible component of the tau decays, and such pairs are only categorized as BelowZ or AboveZ with respect to M_{Z}.

In OSSF0 events, the mass of the OSOF dilepton pair with the largest mass is chosen for classification purposes as BelowZ or AboveZ. Similar to OSSF ditau pairs, OSOF dilepton pairs are also only categorized as BelowZ or AboveZ. OSSF0 events with no OSOF pairs are classified as same-sign (SS) events.

The 3L and 2L1T channels are further split into two, based on the values of either the M_{T}, or the minimum light lepton or tau p_{T} variables. In the 3L OnZ channel, an $M_{\mathrm{T}} > 150$ GeV criterion is used for this binary low/high classification, whereas a minimum light lepton $p_{\mathrm{T}} > 25$ GeV criterion is used for the rest of the 3L channel and a tau $p_{\mathrm{T}} > 50$ GeV criterion is used in the 2L1T channel.

This categorization scheme, detailed below in Table 6.2 and labelled as the fundamental scheme, yields 43 orthogonal selections labelled A1-G1. This scheme allows the complete utilization of multilepton events collected in Run-2, such that any event that does not populate a control region is a part of the so-called signal regions. In order to provide sensitivity to a large class of signal models beyond the SM, S_{T} and $L_{\mathrm{T}}+p_{\mathrm{T}}^{\mathrm{miss}}$ variables have

Table 6.2: Fundamental scheme of event categorization as a function of lepton charge and mass variables. Disallowed categories are marked with a hyphen, and categories marked with an asterix are inclusive of the unmarked ones in a given OSSFn channel. The 1L3T OSSF0 and OSSF1 events are combined into a single category.

		OSSF0			OSSF1				OSSF2		
		BelowZ	AboveZ	SS	OnZ	BelowZ	AboveZ	MixedZ	Single-OnZ	Double-OnZ	OffZ
3L	Low $p_{\mathrm{T}}/M_{\mathrm{T}}$	A1*		A2	A3	A4	A5	A6	–	–	–
	High $p_{\mathrm{T}}/M_{\mathrm{T}}$	A7*		A8	A9	A10	A11	A12	–	–	–
2L1T	Low p_{T}	B1	B2	B3	B4	B5	B6	–	–	–	–
	High p_{T}	B7	B8	B9	B10	B11	B12	–	–	–	–
1L2T		C1	C2	C3	–	C4	C5	–	–	–	–
4L		D1*			D2	D3*			D4	D5	D6
3L1T		E1*			E2	E3*			–	–	–
2L2T		F1*			F2*			–	F3	–	F4
1L3T		G1*			–	G1*		–	–	–	–

been chosen as the final discriminating variables in each of the 43 categories, producing the fundamental S_{T} table, and the fundamental $L_{\mathrm{T}} + p_{\mathrm{T}}^{\mathrm{miss}}$ table, respectively. The individual S_{T} and $L_{\mathrm{T}} + p_{\mathrm{T}}^{\mathrm{miss}}$ spectra in each of these 43 categories are evaluated in 200 GeV wide bins in order to ensure smooth and mostly monotonic falling expected background behavior. The first and last bins of the $L_{\mathrm{T}} + p_{\mathrm{T}}^{\mathrm{miss}}$ or S_{T} distribution are chosen with the requirement that the per-bin expected background yield is more than 1.0 to ensure robustness in statistical interpretations.

As illustrated in Chapter 9, the $L_{\mathrm{T}} + p_{\mathrm{T}}^{\mathrm{miss}}$ variable is expected to be more appropriate for signal models like type-III seesaw or vector-like leptons with low to moderate hadronic activity, whereas the S_{T} variable is expected to be more performant for the leptoquark model with high hadronic activity.

A second categorization scheme, the so-called advanced scheme, is also defined building on the fundamental S_{T} scheme. Each of the 43 fundamental scheme categories is first split, background statistics permitting, in up to three b tag multiplicities regions. Furthermore, each category in a given b tag multiplicity region is split, background statistics permitting, in up to four bins, using binary low or high $p_{\mathrm{T}}^{\mathrm{miss}}$ and H_{T} selection criteria. This results in a total of 204 orthogonal categories. The S_{T} variable binned in 200 GeV increments is used as the final discriminating variable, defining the advanced table, which is described in Tables 6.3-6.6.

Table 6.3: The binning of the $L_\mathrm{T} + p_\mathrm{T}^\mathrm{miss}$ or S_T variable for the fundamental and advanced table schemes in the 3L channel based on the categorization described in Table 6.2. The ranges, as well as the $p_\mathrm{T}^\mathrm{miss}$ and H_T requirements, are given in GeV. The first bins in the $L_\mathrm{T} + p_\mathrm{T}^\mathrm{miss}$ or S_T range contain the underflow, the last bins contain the overflow.

	Fundamental Tables				Advanced Table								
	$L_\mathrm{T} + p_\mathrm{T}^\mathrm{miss}$		S_T		0 b tag				1 b tag			≥ 2 b tag	
Cat.	Range	Bins	Range	Bins	$p_\mathrm{T}^\mathrm{miss}$	H_T	S_T Range	Bins	$p_\mathrm{T}^\mathrm{miss}$	S_T Range	Bins	S_T Range	Bins
A1	$[0, 800]$	1–4	$[50, 1650]$	1–8	< 125	< 150	$[100, 700]$	1–3	< 125	$[100, 1100]$	1–5	$[50, 1250]$	1–6
					< 125	> 150	$[200, 1000]$	4–7					
					> 125	< 150	$[250, 650]$	8–9	> 125	$[150, 1350]$	6–11		
					> 125	> 150	$[350, 1150]$	10–13					
A2	$[50, 450]$	5–6	$[150, 750]$	9–11	—	—	$[150, 750]$	14–16	—	—	—	—	—
A3	$[50, 1650]$	7–14	$[150, 2750]$	12–24	> 125	< 150	$[50, 1450]$	17–23	> 125	$[300, 2100]$	12–20	$[350, 1750]$	7–13
					> 125	> 150	$[250, 2650]$	24–35					
A4	$[100, 900]$	15–18	$[50, 1850]$	25–33	< 125	< 150	$[50, 650]$	36–38	< 125	$[0, 1200]$	21–26	$[100, 1300]$	14–19
					< 125	> 150	$[150, 1350]$	39–44					
					> 125	< 150	$[100, 700]$	45–47	> 125	$[100, 1500]$	27–33		
					> 125	> 150	$[300, 1500]$	48–53					
A5	$[150, 1150]$	19–23	$[0, 1800]$	34–42	< 125	< 150	$[0, 1000]$	54–58	< 125	$[100, 1100]$	34–38	$[100, 1300]$	20–25
					< 125	> 150	$[150, 1150]$	59–63					
					> 125	< 150	$[100, 900]$	64–67	> 125	$[200, 1200]$	39–43		
					> 125	> 150	$[300, 1300]$	68–72					
A6	$[50, 850]$	24–27	$[0, 1400]$	43–49	< 125	< 150	$[50, 650]$	73–75	< 125	$[150, 950]$	44–48	$[300, 1100]$	26–29
					< 125	> 150	$[200, 1000]$	76–79					
					> 125	< 150	$[200, 800]$	80–82	> 125	$[350, 1150]$	49–51		
					> 125	> 150	$[500, 1100]$	83–85					
A7	$[0, 1000]$	28–32	$[150, 1750]$	50–57	< 125	< 150	$[50, 650]$	86–88	< 125	$[150, 1150]$	52–56	$[150, 1350]$	30–35
					< 125	> 150	$[150, 950]$	89–92					
					> 125	< 150	$[150, 750]$	93–95	> 125	$[350, 1350]$	57–61		
					> 125	> 150	$[350, 1350]$	96–100					
A8	$[100, 500]$	33–34	$[50, 650]$	58–60	—	—	$[50, 650]$	101–103	—	—	—	—	—
A9	$[150, 1350]$	35–40	$[150, 2150]$	61–70	< 125	< 150	$[150, 2150]$	104–108	< 125	$[300, 1300]$	62–66	$[450, 1250]$	36–39
					< 125	> 150	$[400, 1800]$	109–115					
					> 125	< 150	$[200, 1000]$	116–119	> 125	$[350, 1350]$	67–71		
					> 125	> 150	$[500, 1700]$	120–125					
A10	$[100, 1100]$	41–45	$[0, 1800]$	71–79	< 125	< 150	$[0, 800]$	126–129	< 125	$[150, 1150]$	72–76	$[250, 1250]$	40–44
					< 125	> 150	$[200, 1400]$	130–135					
					> 125	< 150	$[150, 950]$	136–139	> 125	$[300, 1300]$	77–81		
					> 125	> 150	$[300, 1500]$	140–145					
A11	$[0, 1400]$	46–52	$[50, 2050]$	80–89	< 125	< 150	$[50, 1250]$	146–151	< 125	$[200, 1400]$	82–87	$[200, 1600]$	45–51
					< 125	> 150	$[200, 1600]$	152–158					
					> 125	< 150	$[200, 1200]$	159–163	> 125	$[300, 1500]$	88–93		
					> 125	> 150	$[400, 1800]$	164–170					
A12	$[100, 1100]$	53–57	$[150, 1750]$	90–97	< 125	< 150	$[100, 900]$	171–174	< 125	$[100, 1100]$	94–98	$[350, 1150]$	52–55
					< 125	> 150	$[250, 1450]$	175–180					
					> 125	< 150	$[300, 900]$	181–183	> 125	$[500, 1300]$	99–102		
					> 125	> 150	$[450, 1450]$	184–188					

Table 6.4: The binning of the $L_{\mathrm{T}} + p_{\mathrm{T}}^{\mathrm{miss}}$ or S_{T} variable for the fundamental and advanced table schemes in the 2L1T channel based on the categorization described in Table 6.2. The ranges, as well as the $p_{\mathrm{T}}^{\mathrm{miss}}$ and H_{T} requirements, are given in GeV. The first bins in the $L_{\mathrm{T}} + p_{\mathrm{T}}^{\mathrm{miss}}$ or S_{T} range contain the underflow, and the last bins contain the overflow.

	Fundamental Tables				Advanced Table								
	$L_{\mathrm{T}} + p_{\mathrm{T}}^{\mathrm{miss}}$		S_{T}		0 b tag				1 b tag			≥ 2 b tag	
Cat.	Range	Bins	Range	Bins	$p_{\mathrm{T}}^{\mathrm{miss}}$	H_{T}	S_{T} Range	Bins	$p_{\mathrm{T}}^{\mathrm{miss}}$	S_{T} Range	Bins	S_{T} Range	Bins
B1	$[100, 700]$	1–3	$[100, 1300]$	1–6	< 100	< 150	$[150, 550]$	1–2	< 100	$[0, 800]$	1–4	$[150, 950]$	1–4
					< 100	> 150	$[250, 850]$	3–5					
					> 100	< 150	$[0, 600]$	6–8	> 100	$[100, 1100]$	5–9		
					> 100	> 150	$[350, 1150]$	9–12					
B2	$[50, 850]$	4–7	$[0, 1600]$	7–14	< 100	< 150	$[50, 650]$	13–15	< 100	$[50, 1050]$	10–14	$[150, 1150]$	5–9
					< 100	> 150	$[300, 1100]$	16–19					
					> 100	< 150	$[200, 800]$	20–22	> 100	$[250, 1250]$	15–19		
					> 100	> 150	$[300, 1300]$	23–27					
B3	$[100, 500]$	8–9	$[150, 750]$	15–17	—	—	$[150, 750]$	28–30	—	—	—	—	—
B4	$[150, 950]$	10–13	$[0, 1800]$	18–26	> 100	< 150	$[50, 750]$	31–34	> 100	$[200, 1200]$	20–24	$[250, 1050]$	10–13
					> 100	> 150	$[400, 1600]$	35–40					
B5	$[100, 700]$	14–16	$[50, 1250]$	27–32	< 100	< 150	$[150, 550]$	41–42	< 100	$[50, 850]$	25–28	$[50, 1050]$	14–18
					< 100	> 150	$[100, 1100]$	43–47					
					> 100	< 150	$[0, 600]$	48–50	> 100	$[100, 900]$	29–32		
					> 100	> 150	$[300, 1100]$	51–54					
B6	$[0, 1000]$	17–21	$[150, 1550]$	33–39	< 100	< 150	$[50, 850]$	55–58	< 100	$[150, 1150]$	33–37	$[250, 1050]$	19–22
					< 100	> 150	$[200, 1200]$	59–63					
					> 100	< 150	$[200, 800]$	64–66	> 100	$[250, 1250]$	38–42		
					> 100	> 150	$[300, 1300]$	67–71					
B7	$[100, 700]$	22–24	$[50, 1250]$	40–45	< 100	< 150	$[150, 550]$	72–73	< 100	$[100, 700]$	43–45	$[400, 800]$	23–24
					< 100	> 150	$[150, 750]$	74–76					
					> 100	< 150	$[150, 550]$	77–78	> 100	$[250, 1050]$	46–49		
					> 100	> 150	$[350, 1150]$	79–82					
B8	$[0, 1000]$	25–29	$[150, 1750]$	46–53	< 100	< 150	$[150, 750]$	83–85	< 100	$[100, 1100]$	50–54	$[350, 1350]$	25–29
					< 100	> 150	$[250, 1050]$	86–89					
					> 100	< 150	$[150, 950]$	90–93	> 100	$[250, 1450]$	55–60		
					> 100	> 150	$[250, 1450]$	94–99					
B9	$[100, 500]$	30–31	$[100, 700]$	54–56	—	—	$[100, 700]$	100–102	—	—	—	—	—
B10	$[250, 1250]$	32–36	$[200, 2000]$	57–65	> 100	< 150	$[100, 1100]$	103–107	> 100	$[250, 1250]$	61–65	Incl.	30
					> 100	> 150	$[400, 2000]$	108–115					
B11	$[100, 900]$	37–40	$[50, 1450]$	66–72	< 100	< 150	$[100, 700]$	116–118	< 100	$[100, 900]$	66–69	$[250, 950]$	31–33
					< 100	> 150	$[250, 1050]$	119–122					
					> 100	< 150	$[150, 750]$	123–125	> 100	$[250, 1050]$	70–73		
					> 100	> 150	$[400, 1200]$	126–129					
B12	$[50, 1050]$	41–45	$[150, 1750]$	73–80	< 100	< 150	$[100, 900]$	130–133	< 100	$[200, 1000]$	74–77	$[400, 1000]$	34–36
					< 100	> 150	$[350, 1150]$	134–137					
					> 100	< 150	$[300, 900]$	138–140	> 100	$[350, 1350]$	78–82		
					> 100	> 150	$[600, 1200]$	141–143					

Table 6.5: The binning of the $L_\mathrm{T} + p_\mathrm{T}^\mathrm{miss}$ or S_T variable for the fundamental and advanced table schemes in the 1L2T channel based on the categorization described in Table 6.2. The ranges, as well as the $p_\mathrm{T}^\mathrm{miss}$ and H_T requirements, are given in GeV. The first bins in the $L_\mathrm{T} + p_\mathrm{T}^\mathrm{miss}$ or S_T range contain the underflow, and the last bins contain the overflow.

	Fundamental Tables				Advanced Table							
	$L_\mathrm{T} + p_\mathrm{T}^\mathrm{miss}$		S_T		0 b tag				1 b tag			≥ 2 b tag	
Cat.	Range	Bins	Range	Bins	$p_\mathrm{T}^\mathrm{miss}$	H_T	S_T Range	Bins	$p_\mathrm{T}^\mathrm{miss}$	S_T Range	Bins	S_T Range	Bins
C1	$[100, 500]$	46–47	$[100, 900]$	81–84	< 75	< 75	$[0, 400]$	1–2	< 75	$[100, 500]$	51–52	$[200, 600]$	78–79
					< 75	> 75	$[200, 600]$	3–4					
					> 75	< 75	Incl.	5	> 75	$[100, 700]$	53–55		
					> 75	> 75	$[150, 750]$	6–8					
C2	$[150, 750]$	48–50	$[100, 1100]$	85–89	< 75	< 75	$[150, 750]$	9–10	< 75	$[50, 650]$	56–58	$[100, 900]$	80–83
					< 75	> 75	$[100, 700]$	11–13					
					> 75	< 75	$[150, 750]$	14–16	> 75	$[200, 800]$	59–61		
					> 75	> 75	$[300, 900]$	17–19					
C3	$[50, 450]$	51–52	$[150, 550]$	90–91	—	—	$[150, 550]$	20–21	—	—	—	—	—
C4	$[50, 850]$	53–56	$[100, 1700]$	92–99	< 75	< 75	$[0, 600]$	22–24	< 75	$[0, 800]$	62–65	$[200, 1000]$	84–87
					< 75	> 75	$[150, 950]$	25–28					
					> 75	< 75	$[150, 750]$	29–31	> 75	$[150, 1150]$	66–70		
					> 75	> 75	$[150, 1350]$	32–37					
C5	$[50, 850]$	57–60	$[150, 1350]$	100–105	< 75	< 75	$[0, 600]$	38–40	< 75	$[150, 750]$	71–73	$[200, 800]$	88–90
					< 75	> 75	$[250, 850]$	41–43					
					> 75	< 75	$[50, 650]$	44–46	> 75	$[200, 1000]$	74–77		
					> 75	> 75	$[300, 1100]$	47–50					

Table 6.6: The binning of the $L_\mathrm{T} + p_\mathrm{T}^\mathrm{miss}$ or S_T variable for the fundamental and advanced table schemes in the 4L, 3L1T, 2L2T and 1L3T channels based on the categorization described in Table 6.2. The ranges, as well as the $p_\mathrm{T}^\mathrm{miss}$ and H_T requirements, are given in GeV. The first bins in the $L_\mathrm{T} + p_\mathrm{T}^\mathrm{miss}$ or S_T range contain the underflow, and the last bins contain the overflow. For the 3L1T and 2L2T channels, multiple categories are combined in the 1 or ≥ 2 b tag selections. These bins are marked with dagger characters.

	Fundamental Tables				Advanced Table							
	$L_\mathrm{T} + p_\mathrm{T}^\mathrm{miss}$		S_T		0 b tag				1 b tag			≥ 2 b tag	
Cat.	Range	Bins	Range	Bins	$p_\mathrm{T}^\mathrm{miss}$	H_T	S_T Range	Bins	$p_\mathrm{T}^\mathrm{miss}$	S_T Range	Bins	S_T Range	Bins
D1	Incl.	1	Incl.	1	—	—	Incl.	1	—	—	—	—	—
D2	$[150, 950]$	2–5	$[0, 1400]$	2–8	< 75	< 50	$[150, 550]$	2–3	< 75	$[200, 800]$	1–3	$[400, 1000]$	27–29
					< 75	> 50	$[200, 1000]$	4–7					
					> 75	< 50	$[100, 700]$	8–10	> 75	$[300, 1100]$	4–7		
					> 75	> 50	$[250, 1050]$	11–14					
D3	$[150, 750]$	6–8	$[150, 950]$	9–12	< 75	< 50	$[0, 400]$	15–16	—	$[250, 850]$	8–10	Incl.	30
					< 75	> 50	Incl.	17					
					> 75	—	Incl.	18					
D4	$[50, 1250]$	9–14	$[100, 1500]$	13–19	< 75	< 50	$[0, 1000]$	19–23	< 75	$[100, 900]$	11–14	$[250, 1050]$	31–34
					< 75	> 50	$[150, 1150]$	24-28					
					> 75	< 50	$[150, 750]$	29–31	> 75	$[250, 1050]$	15–18		
					> 75	> 50	$[400, 1200]$	32–35					
D5	$[100, 700]$	15–17	$[50, 1050]$	20–24	—	—	—	—	< 75	$[100, 900]$	19–22	Incl.	35
					—	—	—	—	> 75	Incl.	23		
D6	$[0, 800]$	18–21	$[100, 1100]$	25–29	< 75	< 50	$[0, 600]$	36–38	—	$[150, 750]$	24–26	Incl.	36
					< 75	> 50	$[150, 750]$	39–41					
					> 75	—	Incl.	42					
E1	$[100, 500]$	22–23	$[250, 650]$	30–31	—	—	Incl.	1	—	$[150, 950]$	21–24 †	$[250, 850]$	28–30 †
E2	$[100, 900]$	24–27	$[50, 1250]$	32–36	—	—	$[100, 1100]$	2–6	—		†		†
E3	$[150, 750]$	28–30	$[0, 1000]$	37–42	—	—	$[0, 800]$	7–10	—		†		†
F1	Incl.	31	$[50, 450]$	43–44	—	—	Incl.	11	—	$[200, 800]$	25–27 ‡	Incl.	31 ‡
F2	$[150, 550]$	32–33	$[100, 700]$	45–47	—	—	$[150, 550]$	12–13	—		‡		‡
F3	$[100, 700]$	34–36	$[150, 950]$	48–51	—	—	$[100, 900]$	14–17	—		‡		‡
F4	$[150, 550]$	37–38	$[200, 800]$	52–54	—	—	$[150, 550]$	18–19	—		‡		‡
G1	Incl.	39	Incl.	55	—	—	Incl.	20	—	—	—	—	—

Chapter 7

Background estimation

The estimation of the various multilepton background sources is summarized in this chapter. Multilepton events originate from a variety of SM processes with one or more bosons or top quarks, and can be broadly categorized as irreducible and reducible contributions.

A major class of background contributions arises from processes where all reconstructed leptons originate from decays of SM bosons. These contributions, dominated by processes such as WZ, ZZ, and $t\bar{t}Z$ production, constitute irreducible backgrounds, and such leptons are labelled as prompt leptons. Similarly, electrons and muons originating from the decays of prompt taus are considered to be prompt as well. Various other processes, such as $t\bar{t}W$, triboson, and top or vector boson associated Higgs production can also yield prompt multilepton signatures. These contributions are generally suppressed due to lower production cross-sections, and are the minor, rare irreducible backgrounds.

A subdominant contribution is due to leptons originating from initial- or final-state radiation photons that convert asymmetrically such that only one of the resultant leptons is reconstructed in the detector, or where an on-shell photon is misidentified as an electron. Such contributions are labelled as conversion leptons, and primarily originate from leptonic decays of Drell-Yan (DY) or $t\bar{t}$ processes in multilepton events. These conversion backgrounds are considered to be a part of the irreducible background contributions, and are estimated using inclusive or dedicated MC samples, such as $Z\gamma$.

In contrast, $t\bar{t}$, DY, and similar other processes that cannot yield 3 or more prompt or conversion leptons mainly contribute via leptons originating from semi-leptonic heavy flavor decays within jets or from other misidentified detector signatures. These are collectively labelled as misidentified (fake) leptons, and such contributions constitute the reducible backgrounds.

Prompt and conversion leptons are usually well isolated and have small displacements with respect to the primary vertex in the event, whereas misidentified leptons are typically poorly isolated or have large displacement values. While irreducible backgrounds are estimated using simulated samples that are normalized and validated in dedicated control regions in data for major processes, a data driven matrix method is used to estimate reducible background contributions that contain misidentified leptons.

In this analysis, various selections with 2, 3, and 4 lepton events are used as control regions to commission the MC, develop the matrix method, and normalize the leading irreducible background contributions, namely WZ, ZZ, $t\bar{t}Z$, and $Z\gamma$.

7.1 Misidentified lepton control regions

7.1.1 Matrix method

Misidentified lepton backgrounds (MisID) are estimated via a 3- or 4-dimensional implementation of a matrix method. The matrix method is a data-driven background estimation method based on the probabilities with which prompt and misidentified leptons pass a tight lepton selection given that they satisfy a loose lepton selection. These probabilities are called prompt (p) and fake (f) rates. The method relies on the assumption that these rates are universal and can be described as a function of the lepton and event dependent parameters. This assumption allows the measurement of these rates in background dominated control regions and their application to an orthogonal signal region. The algebra of a simplified 2-dimensional implementation of the matrix method can be found in Ref. [100]. The lepton definitions follow those in Chapter 5.

Table 7.1: Prompt and fake rate parameterizations for all lepton flavors.

	Primary Binning scheme	Variable	Correction Binning scheme	Variable
e prompt rate	$\lvert\eta\rvert : \{0, 1.5, 2.4\}$	p_{T}		
e fake rate	$\lvert\eta\rvert : \{0, 1.5, 2.4\}$	p_{T}	$p_{\mathrm{T}} : \{10, 15, 40\}$ GeV, $N_{\mathrm{j}} : \{0, \geq 1\}$	R_T (for DY) or N_{j}^{15} (for $\mathrm{t\bar{t}}$)
μ prompt rate	$\lvert\eta\rvert : \{0, 1.2, 2.4\}$	p_{T}		
μ fake rate	$\lvert\eta\rvert : \{0, 1.2, 2.4\}$	p_{T}	$p_{\mathrm{T}} : \{10, 15, 40\}$ GeV, $N_{\mathrm{j}} : \{0, \geq 1\}$	R_T (for DY) or N_{j}^{15} (for $\mathrm{t\bar{t}}$)
τ_h prompt rate	$\lvert\eta\rvert : \{0, 1.5, 2.3\}^{\dagger}$	p_{T}	$\lvert\eta\rvert : \{0, 1.5, 2.3\}^{\dagger}$	$\lvert\eta\rvert$
τ_h fake rate	$p_{\mathrm{T}} : \{20, 30, 50, 80, 150, \infty\}$ GeV †	ΔR_T	$\lvert\eta\rvert : \{0, 1.5, 2.3\}^{\dagger}$	$\lvert\eta\rvert$
			Inclusive	N_{trk}
	† in 1 and 3 prong separately			

Since this analysis relies on the use of isolated single lepton triggers, it is implicitly assumed that at least one triggering lepton in 3 or 4 lepton events is a prompt lepton. Although background contributions with multiple misidentified leptons are rare, the matrix method is able to predict such contributions with up to 2 (3) simultaneous misidentified leptons in 3 (4) lepton events. The primary ingredient of the matrix method is the determination of the prompt and fake rates for each lepton flavor, as detailed below.

7.1.2 Prompt rates

Prompt rates are measured using a tag-and-probe method [101] in various dilepton events. In data, prompt rates for electrons and muons are studied in a DY enriched set of opposite-sign ee and $\mu\mu$ events, respectively, with an OnZ OSSF pair. Similarly, prompt rates for taus are studied in a DY enriched set of opposite-sign eτ_h and $\mu\tau_h$ events. In order to increase the purity of prompt tau contributions from Z boson decays, 1L1T events are additionally required to have $M_{\mathrm{T}} < 40\,\mathrm{GeV}$, $\Delta\mathrm{R}_{\mathrm{min}} < 3.5$, $p_{\mathrm{T}}^{\mathrm{miss}} < 100\,\mathrm{GeV}$ and the mass of the OSOF dilepton pair in the range of 40–80 GeV. The M_{T} is computed with the light lepton and $p_{\mathrm{T}}^{\mathrm{miss}}$ vectors.

The leading p_{T} light lepton is chosen as the tag and is required to also satisfy the tight lepton selection, and the subleading p_{T} lepton is chosen as the probe. In MC samples, prompt rates have been measured for the DY and $\mathrm{t\bar{t}}$ processes, where reconstructed leptons kinematically matched to generator level prompt leptons ($\Delta\mathrm{R} < 0.2$) are used. In prompt rate measurements conducted in data, contributions due to misidentified probe leptons are estimated and subtracted using MC methods, which is a minor correction for electrons and

muons. Prompt rates play a subdominant role in the matrix method, and the corresponding impact of this uncertainty on the misidentified lepton background estimate is generally negligible.

Prompt rates are primarily parametrized as a function of the lepton p_{T} in the barrel and endcap regions separately. Electron and muon prompt rates are directly parametrized in p_{T} bins, whereas a continuous fit is used for tau prompt rates to minimize statistical fluctuations. Details of the parameterization is given in Table 7.1, and prompt rates for electrons, muons, and taus are given in Appendices B.1, C.1, and D.1.

7.1.3 Fake rates

DY and $\mathrm{t\bar{t}}$ processes are the most dominant SM contributions to the total misidentified lepton background in multilepton events. However, different light quark, heavy quark and gluon composition as well as different event kinematics of these two processes yield fake rates that may differ up to 50% from each other for the same lepton flavor. Therefore, dedicated data and MC measurements using a tag-and-probe method are performed in both processes, for all lepton flavors.

For electrons and muons, a DY enriched selection of data events with a misidentified lepton is created by having a 3L selection with an OnZ pair, $p_{\mathrm{T}}^{\mathrm{miss}} < 100\,\mathrm{GeV}$, $M_{\mathrm{T}} < 50\,\mathrm{GeV}$ and $N_{\mathrm{b}} = 0$. The leptons forming the OnZ pair are taken as the tag leptons, and the additional lepton is taken as the misidentified probe lepton, i.e. eeμ and $\mu\mu\mu$ events are used to measure muon fake rates, and eee and $\mu\mu$e events are used to measure the electron fake rates. In each bin where a fake rate measurement is performed, the low M_{T} ($M_{\mathrm{T}} < 50\,\mathrm{GeV}$) region is used to compute the fake rate, whereas the corresponding medium M_{T} ($50 < M_{\mathrm{T}} < 150\,\mathrm{GeV}$) region is used to measure the per-bin in-situ WZ normalization. The WZ process is the dominant source for prompt lepton contamination in fake rate measurements, and its contribution, alongside other contributions due to prompt leptons, are subtracted in each bin accordingly. For taus, a similar DY enriched set of data events with a misidentified tau probe object is obtained by requiring 2L1T OnZ events with $p_{\mathrm{T}}^{\mathrm{miss}} < 100\,\mathrm{GeV}$, i.e. ee$\tau$ and

$\mu\mu\tau$. Unlike the selection for light leptons, this region has negligible contamination from prompt tau processes, and therefore subtraction does not play an important role.

Similarly, a dileptonic $t\bar{t}$ MC sample is studied for the fake rates of all lepton flavors, using 3 lepton events where generator-level ΔR matching is used to distinguish between the prompt ($\Delta R < 0.2$) and misidentified ($\Delta R > 0.2$) leptons, and the misidentified lepton is taken to the probe object. Additionally, misidentified probe leptons matching to a generator level photon are also vetoed in order to remove contributions from photon conversions. In all DY and $t\bar{t}$ selections, the tag leptons need to also pass the tight lepton requirements, and at least one of them is required match to a trigger object ($\Delta R < 0.2$).

Lepton fake rates are measured in multiple orthogonal regions, utilizing lepton and event level kinematic variables including the lepton recoil variable. The transverse recoil vector for any given lepton, $\vec{p}_R$, is calculated as the 2-dimensional (xy-plane) vector sum of momenta of all other physics objects in the event. Then, the projection of the transverse recoil vector along the lepton transverse momentum axis is defined as:

$$r_T = -\vec{p}_R \cdot \vec{p}_T / p_{\mathrm{T}} \tag{7.1}$$

This ensure that r_T is positive when it is in opposite direction to the lepton. The physics objects include all other leptons and Puppi AK4 jets with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ cleaned against the loose leptons ($\Delta R > 0.4$). For electrons and muons, $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is also included in the calculation of recoil vector as it is found to improve the modeling of the M_{T} distribution.

The electron and muon fake rates are primarily extracted as a function of the lepton p_{T} in the barrel ($|\eta| < 1.5$ for electrons, $|\eta| < 1.2$ for muons) and endcap ($|\eta| > 1.5$ for electrons, $|\eta| > 1.2$ for muons) regions orthogonally, where a continuous fit is used to minimize statistical fluctuations. In order to account for certain event characteristics that affect the fake rates, secondary correction factors are applied to these fake rates. For the DY fake rates of light leptons, these correction factors are measured for each flavor as a function of the relative transverse recoil R_T in bins of low and high lepton p_{T} as well as low and high jet multiplicity (N_{j}) with respect to the average fake rate of given bin, where R_T

is defined as:

$$R_T = r_T / p_{\mathrm{T}} \tag{7.2}$$

For the $\mathrm{t\bar{t}}$ fake rates, the N_{j}^{15} variable is used for the correction factor parameterization instead. The final fake rates are obtained by the product of the initial fake rates and the corresponding correction factors. Correction factors are used instead of orthogonal multidimensional measurements in order to maintain sufficient event yields in each bin.

A similar scheme is also applied to taus. The tau fake rates are primarily measured as a function of the delta transverse recoil ΔR_T in orthogonal p_{T} regions for 1 and 3 prong taus using a continuous fit, where ΔR_T is defined as:

$$\Delta R_T = r_T - p_{\mathrm{T}} \tag{7.3}$$

Correction factors are applied to these fake rates as a function of tau $|\eta|$ in the barrel ($|\eta| < 1.5$) and endcap ($|\eta| > 1.5$) regions for 1 and 3 prong separately, and as a function of N_{trk} in the event inclusively.

The operational differences in the treatment of light lepton and tau fake rate parameterizations originate from differences in available statistics as well as the isolation characteristics among different lepton flavors. Details of the fake rate parameterizations are given in Table 7.1, and all fake rates and corrections are given in Appendices B.2, C.2, and D.2.

The final misidentified background estimation for all lepton flavors is based on a weighted average of the DY and $\mathrm{t\bar{t}}$ based fake and prompt rate measurements. The estimation is based on the expected DY-$\mathrm{t\bar{t}}$ composition in each signal region category described in Table 6.2 obtained from simulated samples.

7.1.4 Validation in data control regions

The measured prompt and fake rates are first commissioned by using the same event selections as used for the measurement of the fake rates, i.e. 3L, OnZ, $p_{\mathrm{T}}^{\mathrm{miss}} < 100\,\mathrm{GeV}$, $M_{\mathrm{T}} < 50\,\mathrm{GeV}$ and $N_{\mathrm{b}} = 0$ events for light leptons, and 2L1T, OnZ, $p_{\mathrm{T}}^{\mathrm{miss}} < 100\,\mathrm{GeV}$ events

for taus. The results of these tests are shown in Figures 7.1 and 7.2 for misidentified light leptons and taus, respectively, where we observe a very good modeling of observed data as a function of key kinematic variables.

7.1.5 Validation in $t\bar{t}$ MC sample

Since the misidentified lepton control regions in data are DY dominated, the performance of the matrix method is also tested in $t\bar{t}$ MC samples using the $t\bar{t}$ MC based rates. An inclusive selection of 3L and 2L1T $t\bar{t}$ MC events are used to study the performance of the matrix method with a misidentified light lepton or tau object, where at least one reconstructed light lepton or tau lepton is required not to be matched ($\Delta R > 0.2$) to a generator level prompt lepton. The MC prediction in the tight lepton selection is treated as the observation, whereas the events satisfying the loose lepton selection are used for the matrix method with $t\bar{t}$ MC based rates to yield a misidentified lepton background estimate. The validation test results as a function of key kinematic variables are highlighted in Figure 7.3 and further detailed in Appendix E. We observe fairly good agreement between the yields of the matrix method and the direct predictions of the MC samples. In order to account for potential systematic trends visible in the tail regions of distributions in some variables and years, a conservative minimum 20% uncertainty is enforced on $t\bar{t}$ MC based fake rates based on the these studies for both light leptons and taus. A more detailed description of the uncertainty treatment of MisID backgrounds is given in Chapter 8. As the $t\bar{t}$ MC samples used in these studies have at least 10 times more effective luminosity than the data collected during the LHC Run-2, these deviations are not a particular source of concern.

7.1.6 Validation in semi-tight control regions

The inversion of the lepton isolation and lepton displacement requirements yield a selection of $t\bar{t}$ enriched data events with a misidentified lepton, and therefore can be used to test the performance of the $t\bar{t}$ MC based prompt and fake rates in data. The requirement that a lepton should fail the tight selection requirements ensures that these selections are

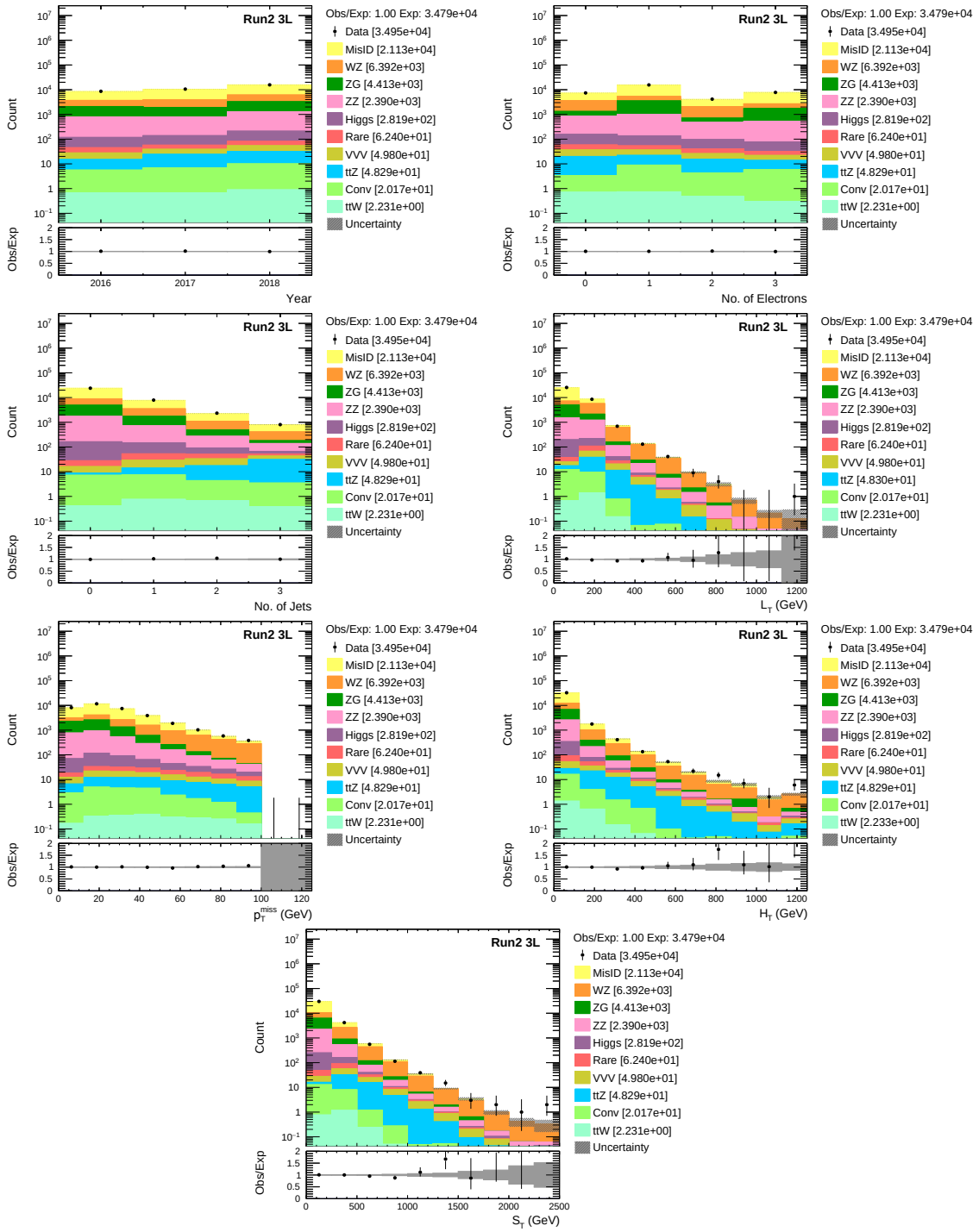

Figure 7.1: Misidentified light lepton control region for the Run-2 dataset. The uncertainties are statistical only.

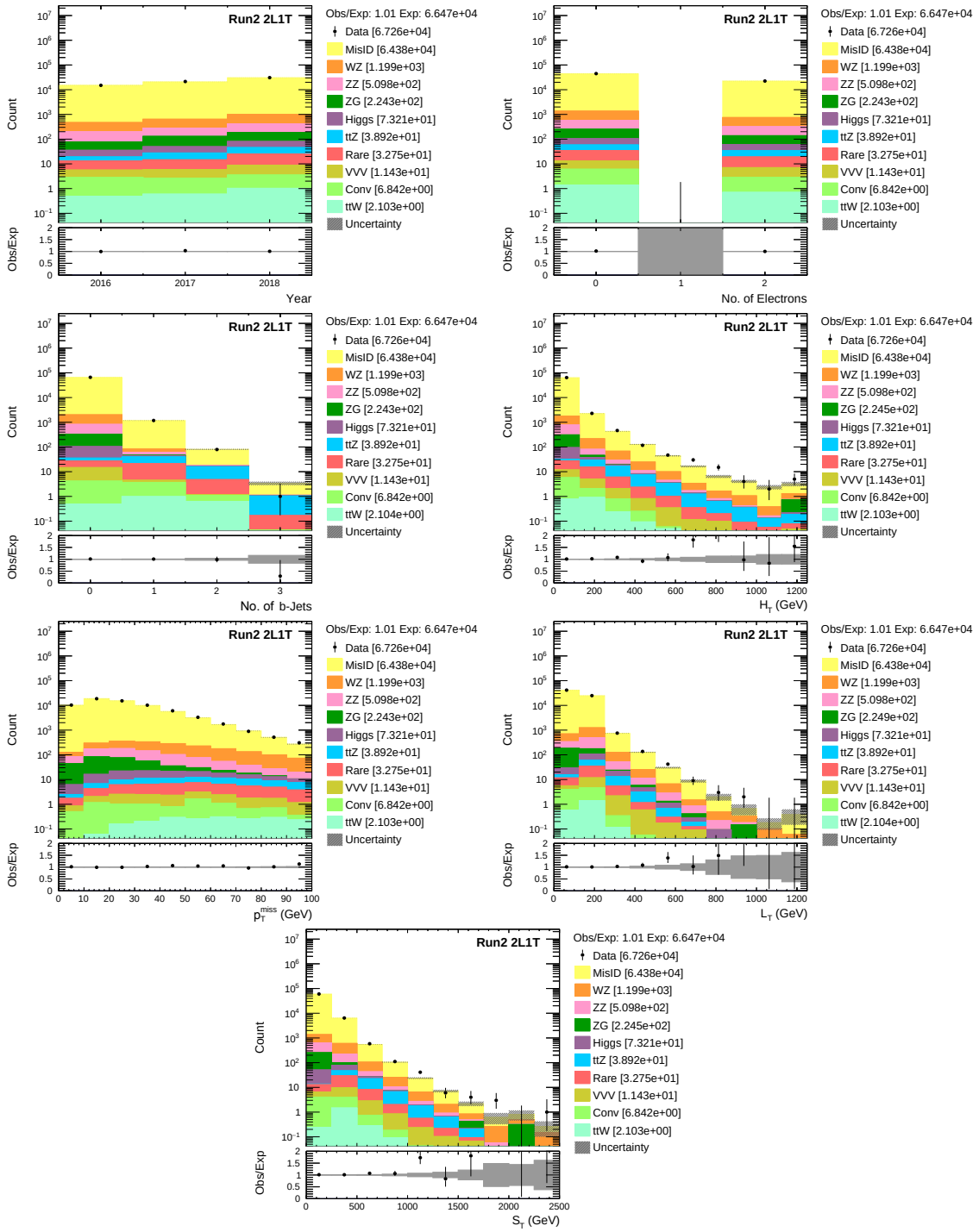

Figure 7.2: Misidentified tau control region for the Run-2 dataset. The uncertainties are statistical only.

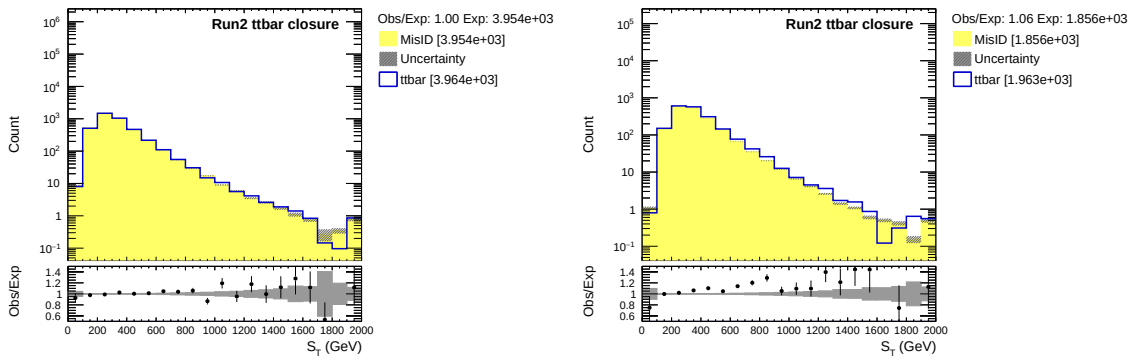

Figure 7.3: Misidentified light lepton $t\bar{t}$ MC validation tests in 3L events for the Run-2 dataset (left) and misidentified tau lepton $t\bar{t}$ MC validation test in 2L1T events for the Run-2 dataset (right). The uncertainties are statistical only.

orthogonal to the signal regions used in the analysis. These selections are defined for light leptons and taus as follows.

For leptons of any flavor that fail the tight lepton selection as described in Chapter 5, a semi-tight selection criteria is defined with slightly relaxed isolation requirements, whereas all other selection requirements are kept unchanged. For muons and electrons, the minimum relative isolation requirement is set to 37.5% and 12.5–37.5% respectively, while for taus, the working point of the anti-jet discriminator is relaxed. Furthermore, for light leptons, an alternate semi-tight selection is defined, where the displacement cuts used in the tight selections are inverted and only those with a lepton DeepCSV of more than 0.95 are vetoed. The inversion of displacement cuts is not applied to tau leptons, as they are expected to have non-zero displacements with respect to the PV due to non-zero tau lifetime. All semi-tight definitions are chosen as narrowly as possible not to diverge from the tight lepton selections, but to also obtain a sufficient yield of misidentified lepton background events. The loose lepton selections remain unchanged in all semi-tight selections. The semi-tight events where isolation cuts have been inverted are referred to as isolation semi-tight events, and those where displacement cuts have been inverted are referred to as displacement semi-tight events.

For the isolation and displacement semi-tight validation tests of light leptons, a selection of 3L events where one lepton has failed the tight selection but satisfies the semi-tight selection criteria is used. The leading light lepton is required to be tight and matched to a trigger

object, in order to ensure that the inversion of the isolation or displacement requirements do not clash with the trigger requirements. In the isolation semi-tight selection, events with no OSSF pair and at least 1 jet, or with an OSSF pair and at least 1 jet and $p_{\mathrm{T}}^{\mathrm{miss}} > 75\,\mathrm{GeV}$ are used to ensure a high purity $\mathrm{t\bar{t}}$ background with a misidentified light lepton. Similarly, in displacement semi-tight selection, events with no OSSF pair and $p_{\mathrm{T}}^{\mathrm{miss}} > 50\,\mathrm{GeV}$, or with an OSSF pair and $p_{\mathrm{T}}^{\mathrm{miss}} > 50\,\mathrm{GeV}$ and at least 1 jet are used.

For the isolation semi-tight validation tests of taus, a selection of 2L1T events is used where the tau lepton has failed the tight selection but satisfies the isolation semi-tight selection criteria. The leading light lepton is required to be tight and triggering as well. Events are required to have $p_{\mathrm{T}}^{\mathrm{miss}} > 50\,\mathrm{GeV}$, $N_{\mathrm{b}} \geq 1$, and either no OSSF pair or an OffZ OSSF pair, to ensure very high $\mathrm{t\bar{t}}$ purity with a misidentified tau.

Following the $\mathrm{t\bar{t}}$ MC based prompt and fake rate measurement strategies used for the tight lepton selections, dedicated $\mathrm{t\bar{t}}$ MC based fake and prompt rates are measured for each lepton flavor using the isolation and displaced semi-tight definitions separately, and are applied via the matrix method to estimate the semi-tight misidentified lepton background contributions. The results of these validation tests for the isolation semi-tight selections as well as the displacement semi-tight selection are given in Figures 7.4, 7.5, and 7.6, where good overall agreement is observed within statistical uncertainties in most bins. The deviations are within the 20% minimum relative uncertainty already derived from the $\mathrm{t\bar{t}}$ MC validation tests discussed in Section 7.1.5. Therefore, we don't derive an additional uncertainty on the $\mathrm{t\bar{t}}$ MC based fake rates.

7.2 ZZ control region

The ZZ process is the primary background component in the 4L channel, and is estimated using NLO POWHEG and LO MCFM samples for quark- and gluon-fusion production modes, respectively. A selection of events > 99% pure in $\mathrm{ZZ} \to 4\ell$ process is obtained by requiring 4L OSSF2 events to have two distinct M_{OSSF} values OnZ (i.e. double-OnZ), and $N_{\mathrm{b}} = 0$.

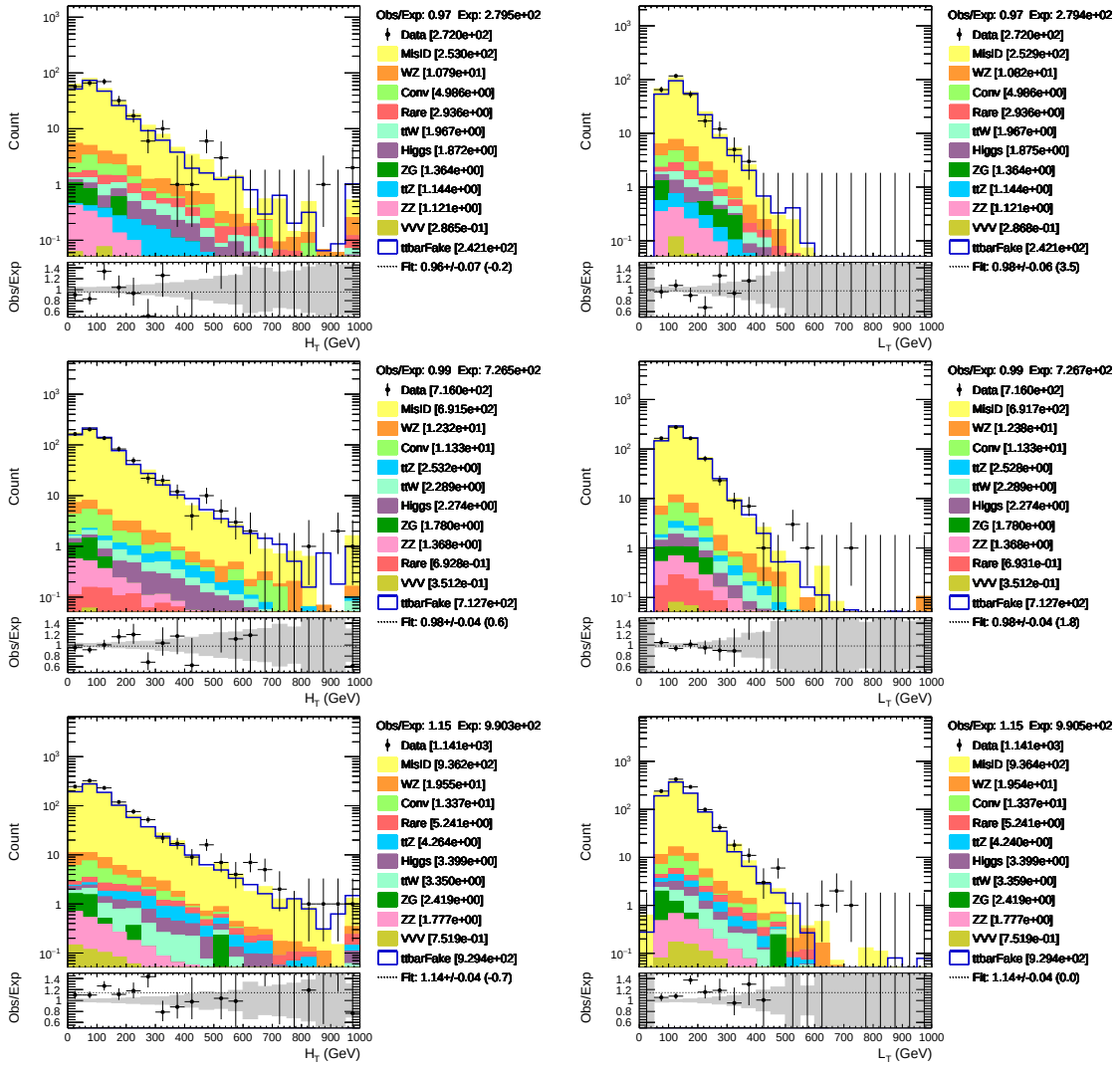

Figure 7.4: Isolation semi-tight validation region for light leptons in 3L events in 2016 (upper row), 2017 (middle row), and 2018 (lower row). The misidentified (fake) background estimation from the simulated $t\bar{t}$ sample is shown for comparison. The uncertainties are statistical only.

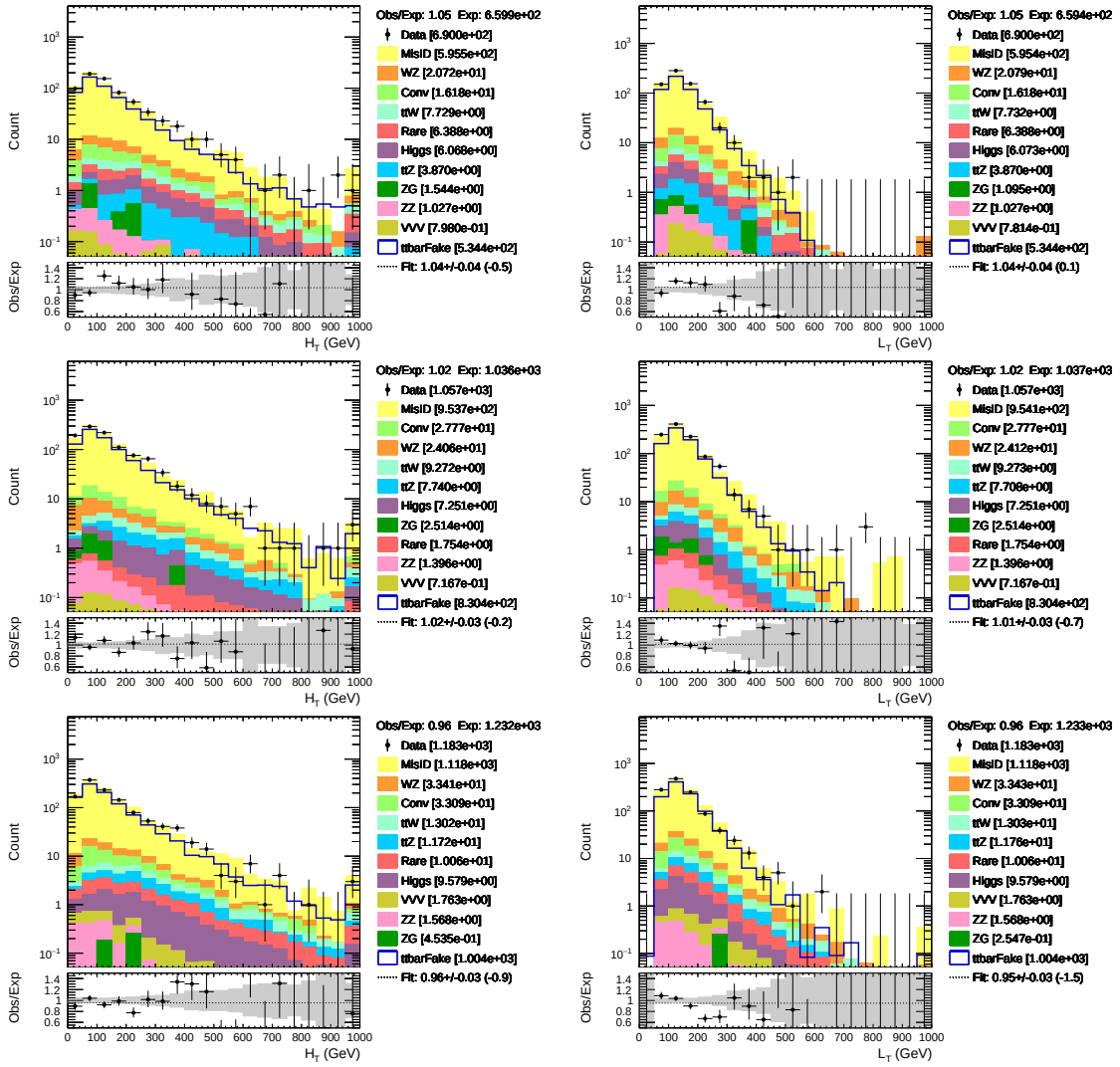

Figure 7.5: Displacement semi-tight validation region for light leptons in 3L events in 2016 (upper row), 2017 (middle row), and 2018 (lower row). The misidentified (fake) background estimation from the simulated $t\bar{t}$ sample is shown for comparison. The uncertainties are statistical only.

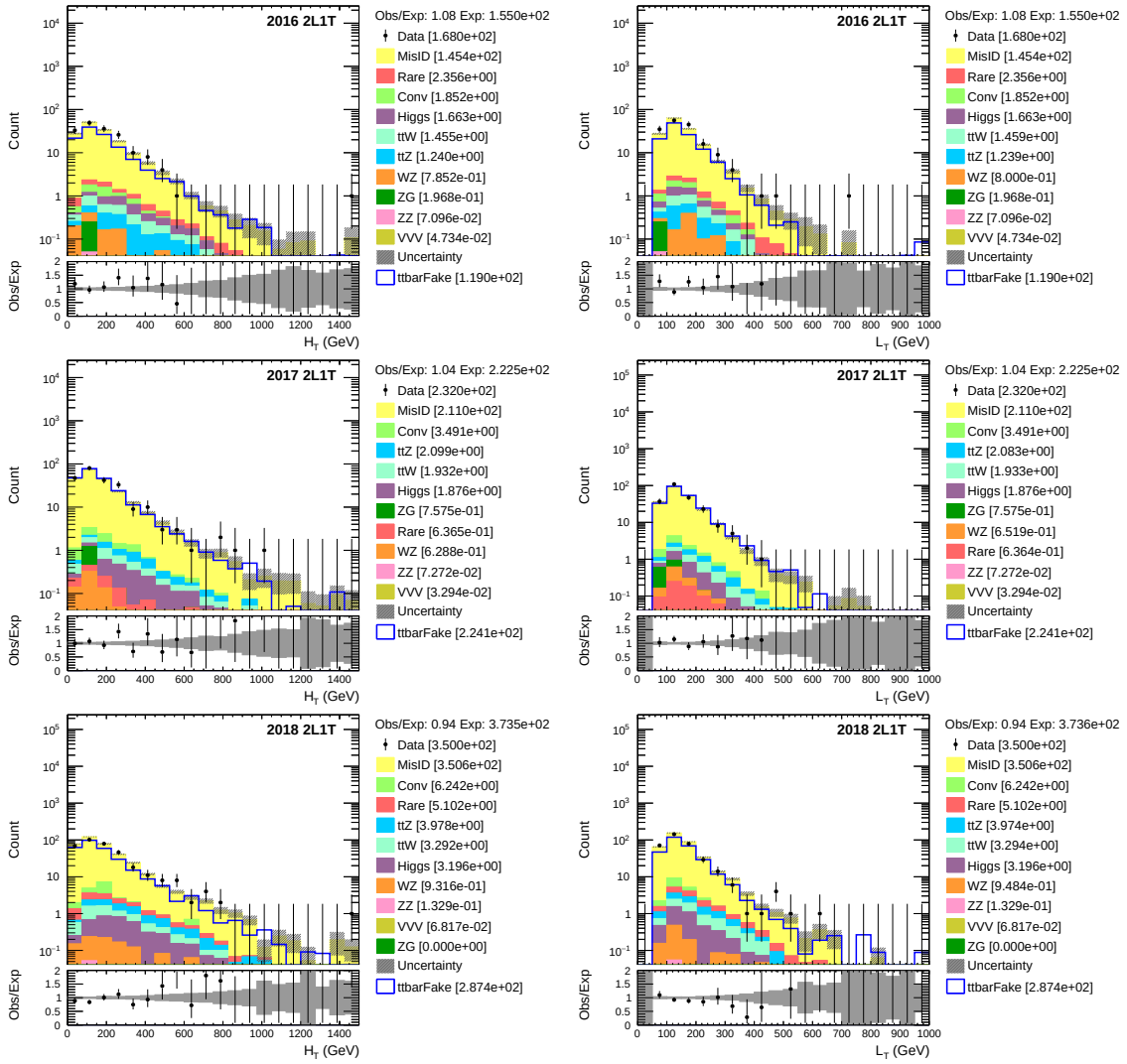

Figure 7.6: Isolation semi-tight validation region for taus in 2L1T events in 2016 (upper row), 2017 (middle row), and 2018 (lower row). The misidentified (fake) background estimation from the simulated $t\bar{t}$ sample is shown for comparison. The uncertainties are statistical only.

We observe normalization factors of 1.05 ± 0.05, 0.97 ± 0.04, and 1.00 ± 0.04 in 2016, 2017, and 2018, respectively. The quoted uncertainties include only the statistical component as the contamination from processes other than ZZ is negligible.

Data driven corrections are applied to the ZZ MC sample to improve the modeling of this background component in this analysis. Differences in the description of the jet multiplicity distribution are used to reweight the ZZ samples in 0, 1, 2 and ≥ 3 jet bins. This is a known deficiency of POWHEG samples that include only up to 1 jet at the generator matrix level. The size of the jet multiplicity corrections are also applied as systematic uncertainties. The ZZ samples are also reweighted as a function of the visible diboson p_{T} to match the MC distribution to that of the data. The visible diboson p_{T} is defined as the vectorial sum of the transverse momenta of all reconstructed leptons in the event. The correction is propagated from the reconstructed visible diboson p_{T} to the generator level visible diboson p_{T}, as illustrated in Figures 7.7 and 7.8, and then applied as a function of the generator level visible diboson p_{T} throughout the analysis. This ensures a proper treatment of the correction in analysis channels including tau leptons. An envelope is defined on the diboson p_{T} reweighting procedure to account for the per-bin statistical and systematic uncertainties, and the resulting uncertainty is propagated to the ZZ background accordingly.

The distributions of key kinematic and event variables in the ZZ control region for the Run-2 dataset are given in Figure 7.9. Good agreement is observed with respect to predictions across all variables.

7.3 WZ control region

The WZ $\rightarrow 3\ell\nu$ process is the primary irreducible background source for the 3 lepton channels. This background is estimated using NLO MADGRAPH5_AMC@NLO samples that are normalized to data using a selection of 3L events with M_{OSSF} OnZ, $p_{\mathrm{T}}^{\mathrm{miss}} < 125\,\mathrm{GeV}$, $50 < M_{\mathrm{T}} < 150\,\mathrm{GeV}$ and $N_{\mathrm{b}} = 0$. In this control region, the minimum lepton p_{T} threshold is raised to $20\,\mathrm{GeV}$ to suppress the misidentified lepton contributions, and the selection

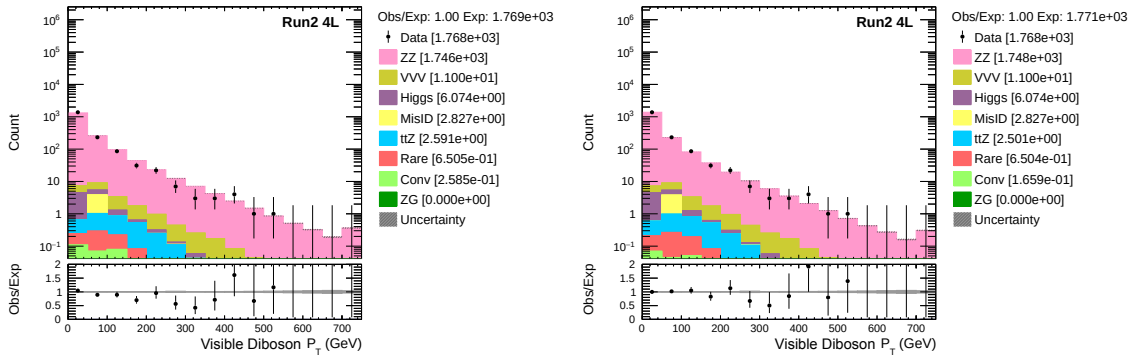

Figure 7.7: The visible diboson p_{T} distribution in ZZ control region for the Run-2 dataset before (left) and after (right) the diboson p_{T} correction. The uncertainties are statistical only.

yields a set of events $> 75\%$ pure in WZ. We observe normalization factors of 0.89 ± 0.03, 0.89 ± 0.05, 0.91 ± 0.03, in 2016, 2017, and 2018, respectively. The quoted uncertainties include the statistical component and the systematic component due to contamination from processes other than WZ.

Following the treatment of the ZZ MC samples, data driven corrections are applied to the WZ MC sample to improve the modeling of this background component in this analysis. Differences in the description of the jet multiplicity distribution are used to reweight the WZ samples in 0, 1, 2 and ≥ 3 jet bins. This is a known deficiency of MadGraph5_aMC@NLO samples that include only up to 2 jets at the generator matrix level. The size of the jet multiplicity corrections are also applied as systematic uncertainties. The WZ samples are also reweighted as a function of the visible diboson p_{T} to match the MC distribution to that of the data. The corrections are applied as a function of the generator level visible diboson p_{T} throughout the analysis, as illustrated in Figures 7.10 and 7.11.

The distributions of key kinematic and event variables in the WZ control region for the Run-2 dataset are given in Figure 7.12. Good agreement is observed with respect to predictions across all variables.

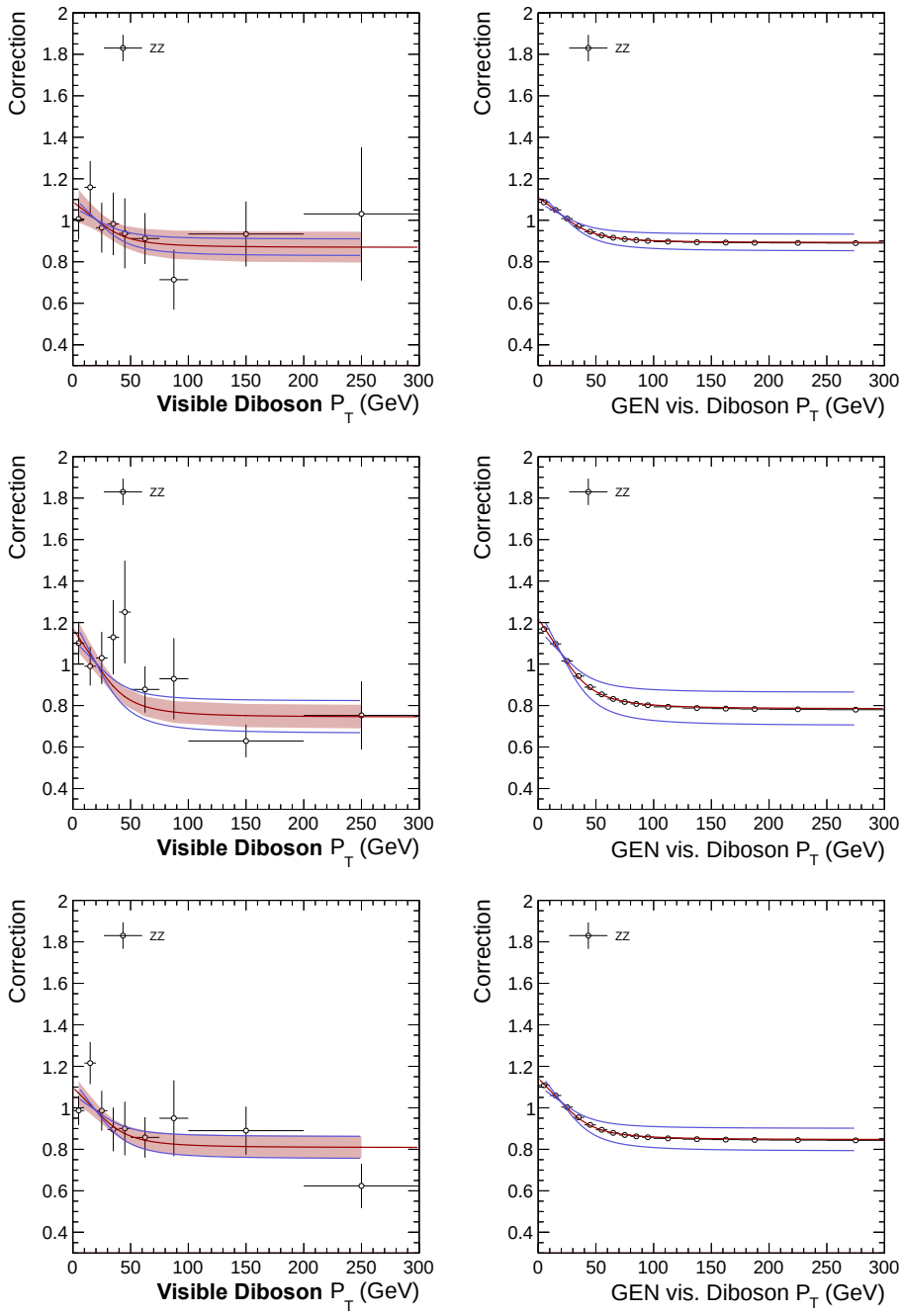

Figure 7.8: The visible diboson p_{T} correction measurement in ZZ control region in 2016 (upper left), in 2017 (middle left), in 2018 (lower left), and the generator level visible diboson p_{T} correction in 2016 (upper right), in 2017 (middle right), in 2018 (lower right). The red line represents an ad-hoc fit function, and the purple lines represent the systematic uncertainty envelope of the measurement. The uncertainties are statistical only.

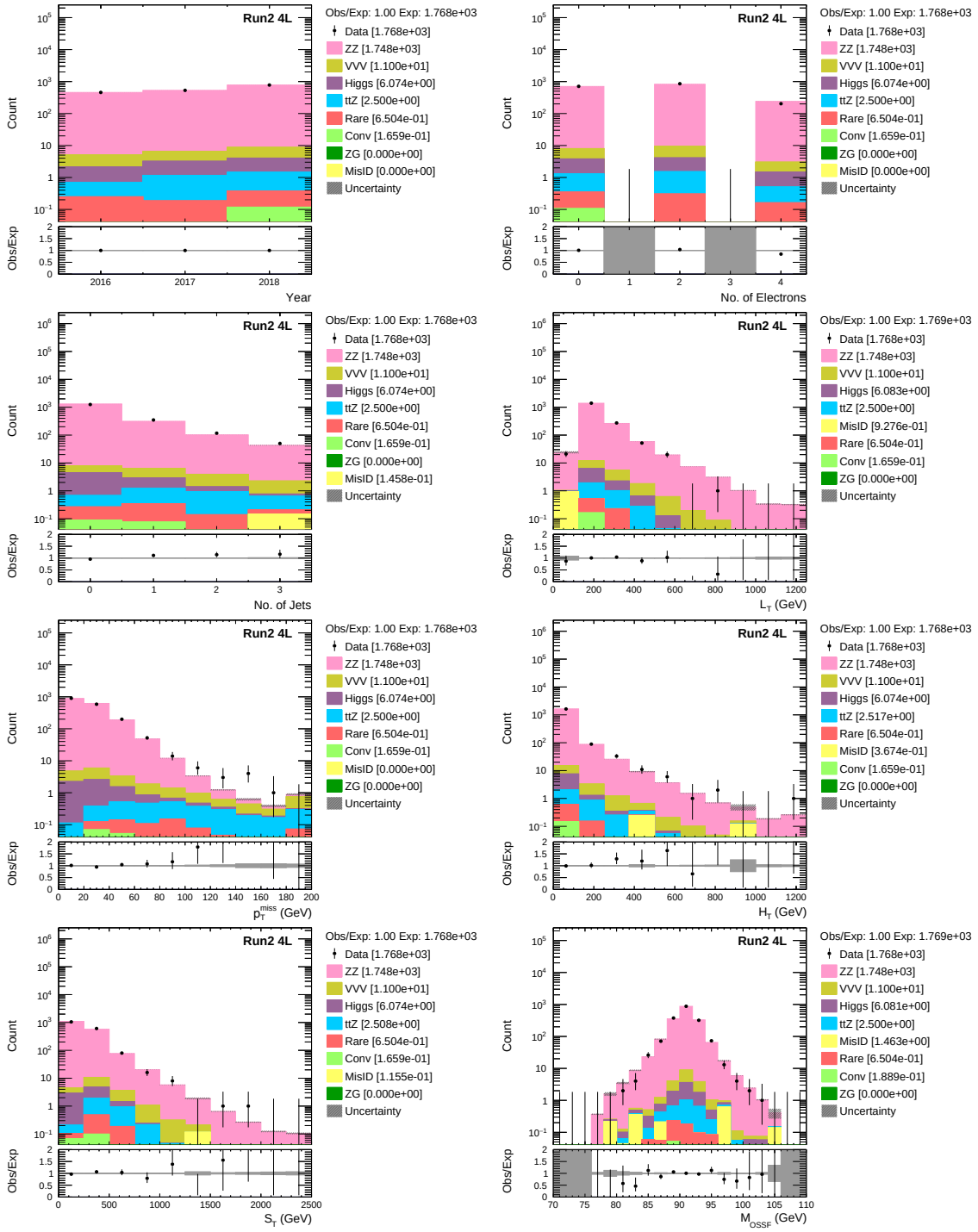

Figure 7.9: ZZ control region for the Run-2 dataset. The uncertainties are statistical only.

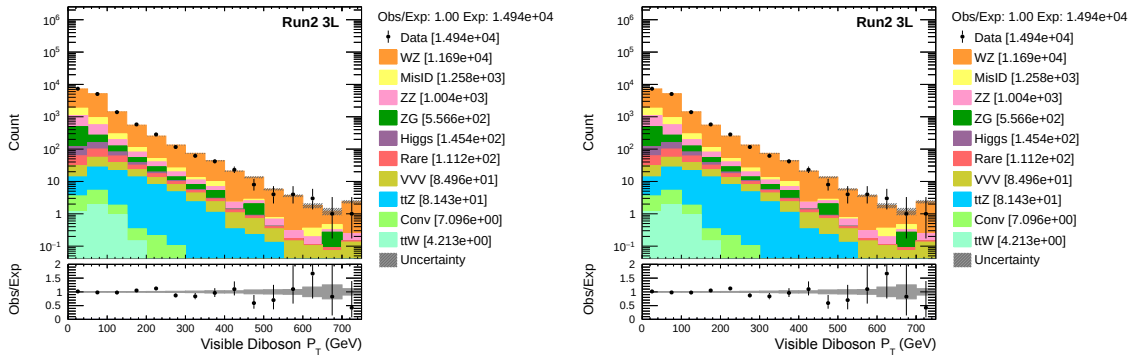

Figure 7.10: The visible diboson p_{T} distribution in WZ control region for the Run-2 dataset before (left) and after (right) the diboson p_{T} correction. The uncertainties are statistical only.

7.4 $\mathrm{t\bar{t}Z}$ control region

$\mathrm{t\bar{t}Z}$ production is a major irreducible SM background process for the 3L and 4L channels with $N_{\mathrm{b}} > 0$. This background is estimated using NLO MadGraph5_aMC@NLO samples and is normalized to data using a selection of 3L events with M_{OSSF} OnZ, $p_{\mathrm{T}}^{\mathrm{miss}} < 125\,\mathrm{GeV}$, $M_{\mathrm{T}} < 150\,\mathrm{GeV}$, $N_{\mathrm{b}} > 0$, $N_{\mathrm{j}} > 2$, and $S_{\mathrm{T}} > 350\,\mathrm{GeV}$. This $\mathrm{t\bar{t}Z}$ control region selection is orthogonal and complementary to the matrix method control region and the WZ control region selections. Similar to the selection in the WZ control region, the minimum lepton p_{T} threshold is raised to 20 GeV to suppress the misidentified lepton contributions, and the selection yields a set of events $\sim 50\%$ pure in $\mathrm{t\bar{t}Z}$. We observe normalization factors of 0.80 ± 0.22, 1.35 ± 0.22, and 1.28 ± 0.21, in 2016, 2017, and 2018, respectively. The quoted uncertainties include the statistical component and the systematic component due to contamination from processes other than $\mathrm{t\bar{t}Z}$.

The distributions of key kinematic and event variables in the $\mathrm{t\bar{t}Z}$ control region for the Run-2 dataset are given in Figure 7.13, illustrating good data agreement with respect to predictions.

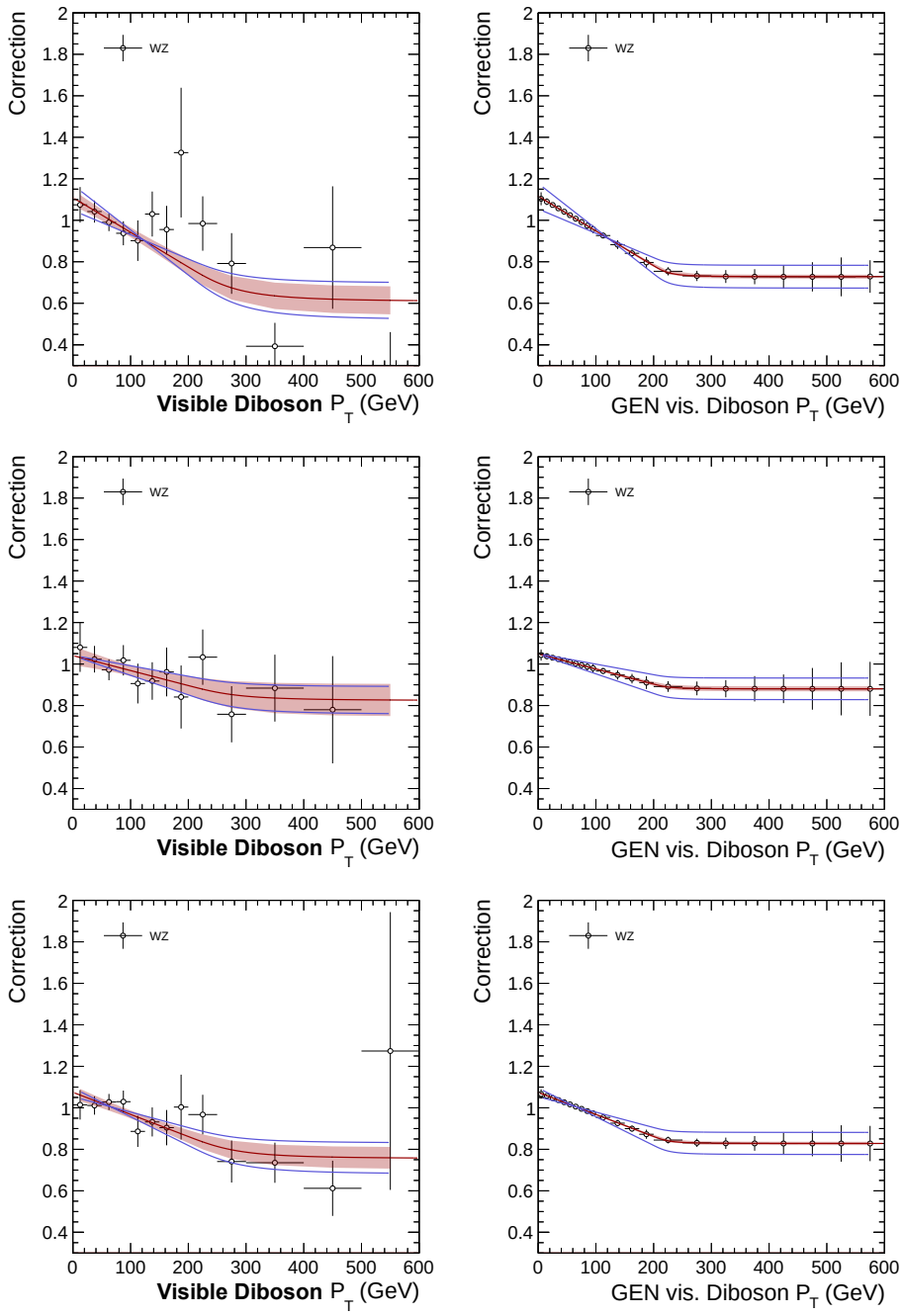

Figure 7.11: The visible diboson p_{T} correction measurement in WZ control region in 2016 (upper left), in 2017 (middle left), in 2018 (lower left), and the generator level visible diboson p_{T} correction in 2016 (upper right), in 2017 (middle right), in 2018 (lower right). The red line represents an ad-hoc fit function, and the purple lines represent the systematic uncertainty envelope of the measurement. The uncertainties are statistical only.

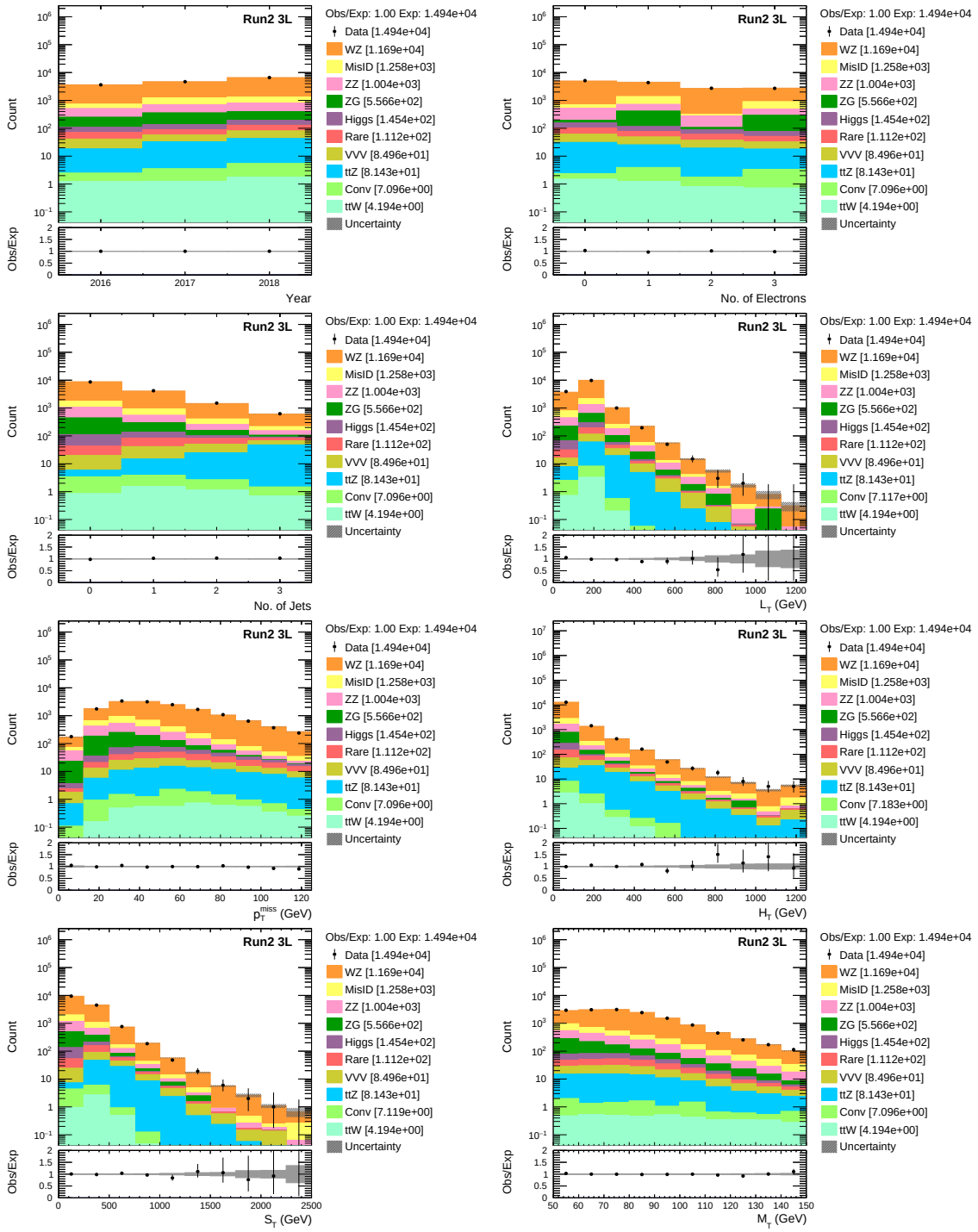

Figure 7.12: WZ control region for the Run-2 dataset. The uncertainties are statistical only.

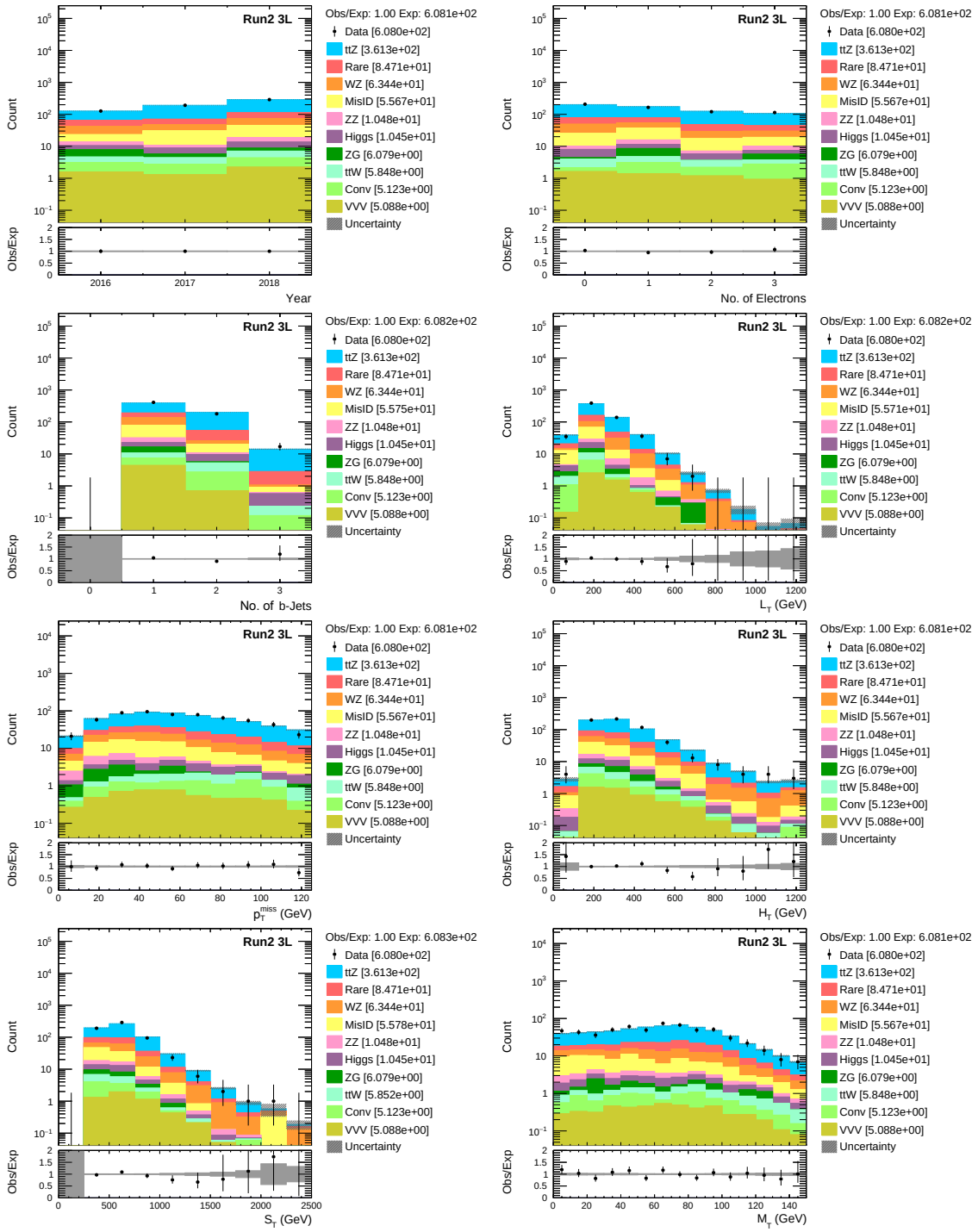

Figure 7.13: $t\bar{t}Z$ control region for the Run-2 dataset. The uncertainties are statistical only.

7.5 3L OnZ control region

Since we used various subsets of the 3L OnZ, $M_{\mathrm{T}} < 150\,\mathrm{GeV}$ events for the determination of light lepton fake rates as well as the normalization of WZ and $\mathrm{t\bar{t}Z}$ MC samples, we scrutinize this region inclusively in order to make sure all background components fit together as a function of key kinematic and event variables. These distributions are given in Figure 7.14.

7.6 Zγ control region

Background contributions due to internal and external photon conversions are estimated via simulated samples. The most dominant conversion background is final state radiation in dileptonic DY events, and the performance of this minor background is commissioned in a dedicated 3L control region. Events with 3 light leptons, $N_{\mathrm{b}} = 0$ and M_{OSSF} BelowZ are considered if $M_{3\ell}$ is within 15 GeV of the Z-boson mass, i.e. $\mathrm{Z} \to \ell\ell + \gamma^{(*)} \to \ell\ell\ell'\ell'$ where one of the ℓ's is too soft and does not satisfy our lepton selection criteria. We use a dedicated Zγ MC sample generated at NLO MadGraph5_aMC@NLO, and we require a generator level match $\Delta\mathrm{R} < 0.2$ between the reconstructed leptons and the prompt leptons and photons in MC in order to veto misidentified lepton contributions. We observe normalization factors of 0.81 ± 0.03, 0.87 ± 0.06, and 0.96 ± 0.05 in 2016, 2017, and 2018, respectively. The quoted uncertainties include the statistical component and the systematic component due to contamination from processes other than Zγ.

The distributions of key kinematic and event variables in the Zγ control region are given in Figure 7.15. Good agreement is observed with respect to predictions across all variables of interest, but since we observe up to 10% fluctuations in certain flavor bins, we assign a 10% overall normalization uncertainty to account for such lepton flavor variations.

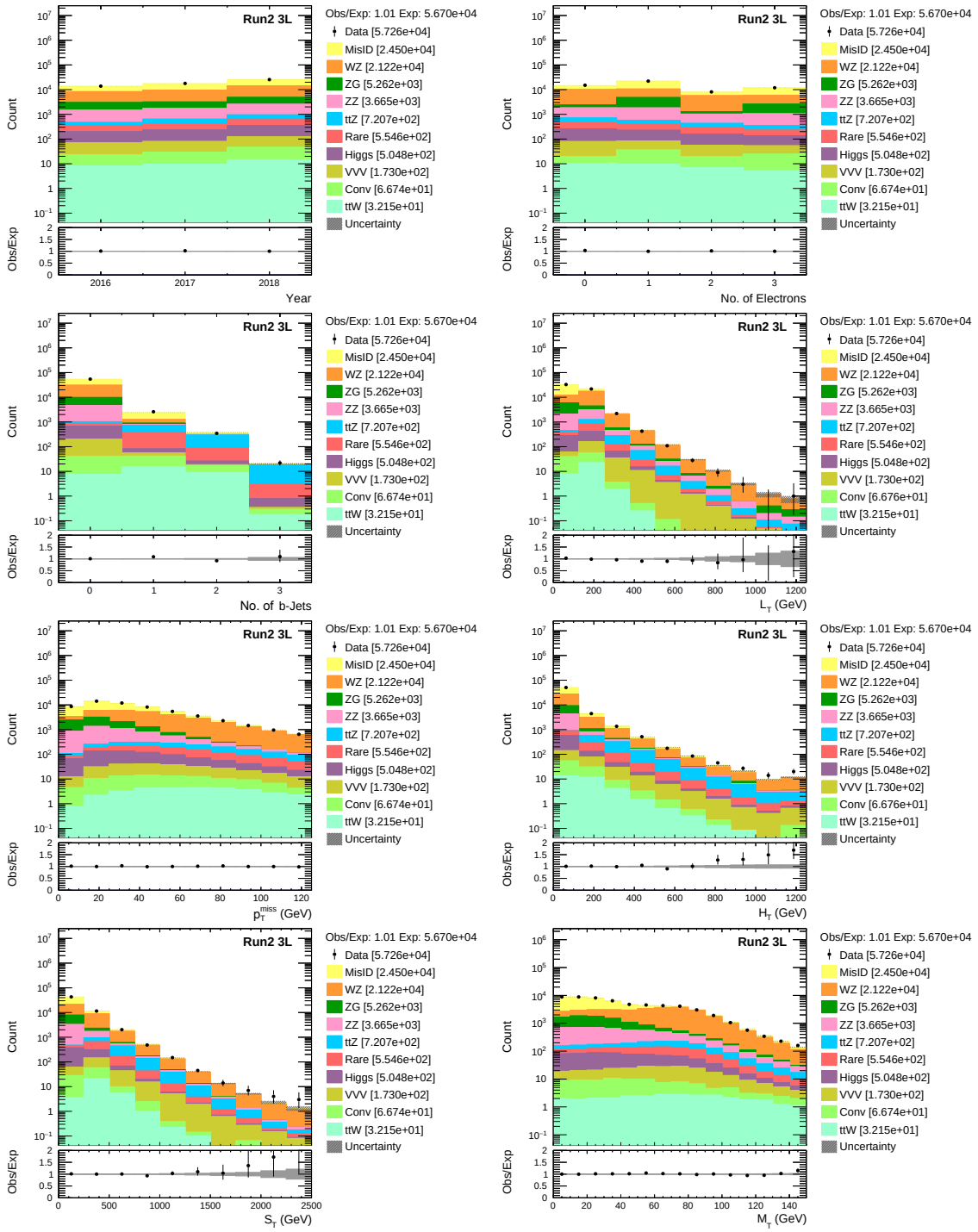

Figure 7.14: Combined 3L OnZ, $M_T < 150\,\text{GeV}$ control region (MisID + WZ + $t\bar{t}Z$) for the Run-2 dataset. The uncertainties are statistical only.

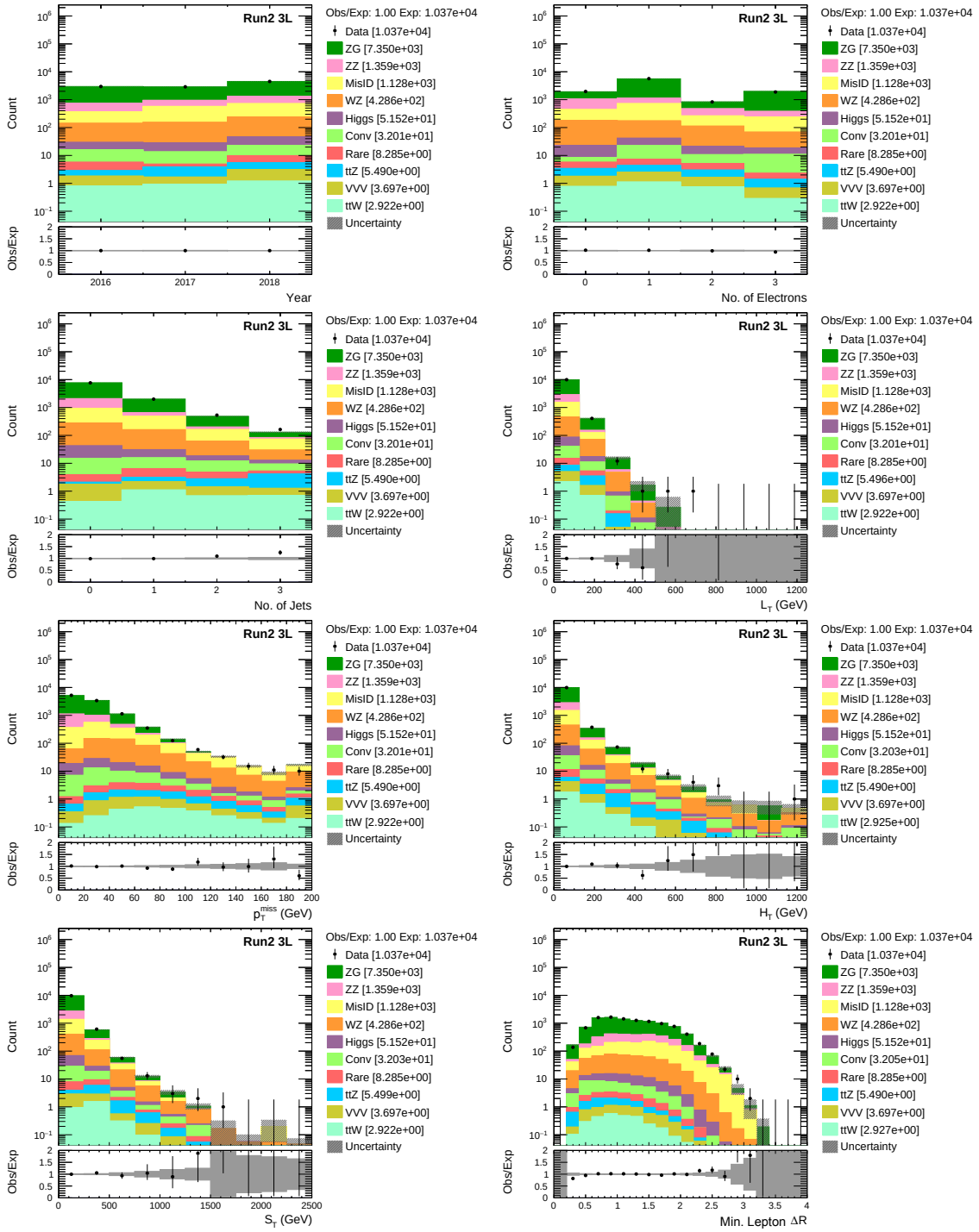

Figure 7.15: Zγ control region for the Run-2 dataset. The uncertainties are statistical only.

7.7 Dilepton control regions

Although the analysis is conducted in 3 or more leptons, dilepton events from leptonic DY or $t\bar{t}$ processes provide high statistics control regions where the object and event selections can be commissioned. DY events are also used for the measurement of electron-charge flip correction factors for simulated samples, as described below.

7.7.1 DY 2L OS control region

A set of events enriched in the $Z \to \mu\mu/ee$ processes are created in a 2L OnZ selection. The contribution due to misidentified (fake) leptons is estimated from MC samples, where at least one reconstructed lepton is not matched ($\Delta R > 0.2$) to a generator level prompt lepton. The NLO MadGraph5_aMC@NLO DY MC sample is normalized to the observed data, and the Z-p_T shape, as well as the jet multiplicity of the MC sample, are corrected to agree with that of the data shape using the events in the dimuon channel. We observe that this Z-p_T correction is also valid in the dielectron channel, and therefore conclude that this is a deficiency of the MC sample.

The distributions of key kinematic and event variables in the $Z \to \mu\mu/ee$ control region are given in Figures 7.16 and 7.17. Good agreement with respect to predictions is observed across all other variables of interest.

7.7.2 DY 2L SS control region

A set of events enriched in prompt electrons, for which the sign of the charge is mismeasured (flipped), is created by a same-sign dielectron OnZ selection. The events are required to also have $M_T < 70\,\text{GeV}$ computed with the leading electron and p_T^{miss}, and $\Delta R < 3.5$ between the dielectron pair to further increase the purity of the $Z \to ee$ process and suppress contributions due to other backgrounds including misidentified (fake) leptons. These events

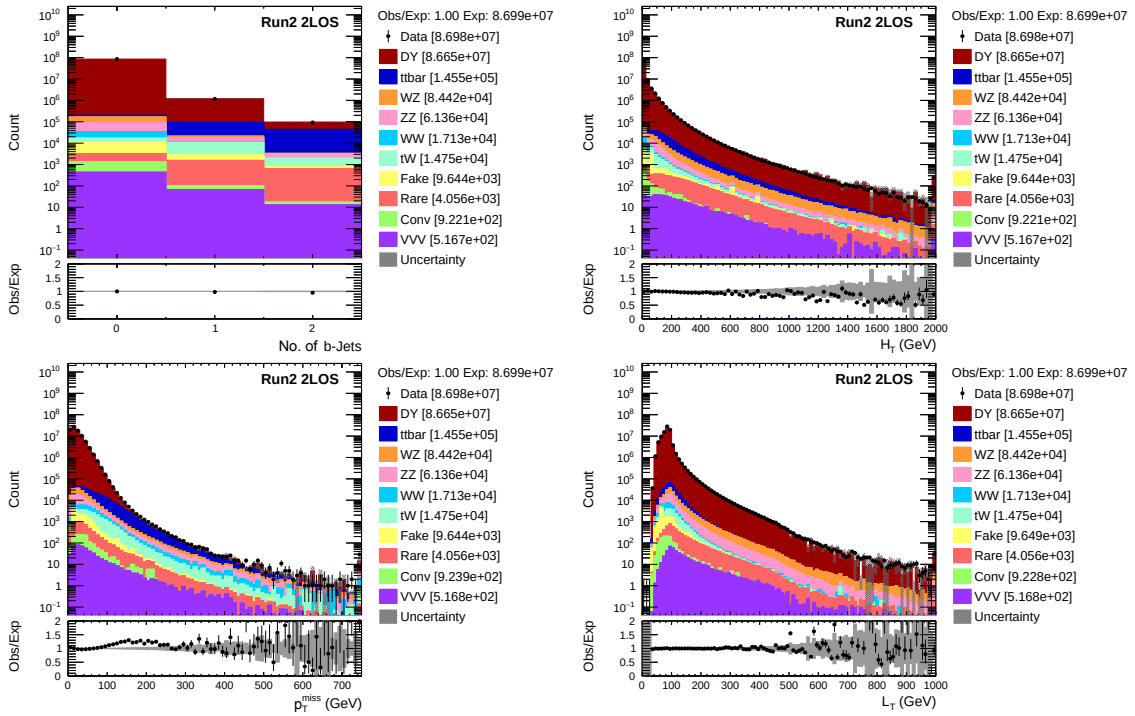

Figure 7.16: 2L Z $\rightarrow \mu\mu$ control region for the Run-2 dataset. The uncertainties are statistical only.

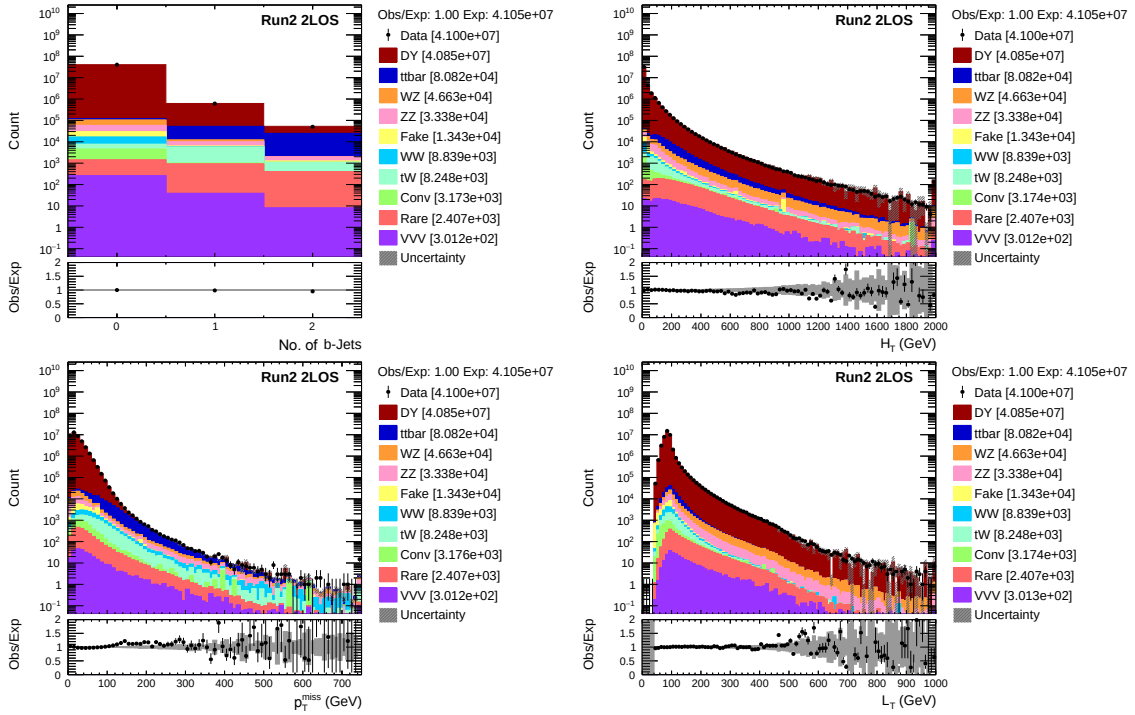

Figure 7.17: 2L Z $\rightarrow$ ee control region for the Run-2 dataset. The uncertainties are statistical only.

Table 7.2: Electron charge flip corrections in three orthogonal bins of $|\eta|$. The uncertainties include the statistical component and the systematic component due to subtraction of processes without a mismeasured charge.

	2016	2017	2018
IB	0.98 ± 0.17	1.23 ± 0.29	0.94 ± 0.26
OB	1.19 ± 0.05	1.33 ± 0.11	1.13 ± 0.08
E	0.93 ± 0.02	1.26 ± 0.04	1.30 ± 0.03

are utilized to derive a per-electron correction factor for selected electrons in MC whose charge is mismeasured.

Since electron charge flip probability depends on the material budget of the detector, this correction is performed in three orthogonal bins of $|\eta|$, namely the inner barrel $|\eta| < 1.0$ (IB), the outer barrel $1.0 < |\eta| < 1.479$ (OB), and the endcap $|\eta| > 1.479$ (E) region [91, 102]. The dielectron events are therefore split into pure IB-IB, OB-OB, and E-E categories, and the remaining mixed-η events are used to verify the measured correction factors. The measurements are repeated for each year separately in order to account for varying detector conditions during Run-2. The correction factors for all three $|\eta|$ bins are given in Table 7.2, the distributions of the yearly breakdown in each category are illustrated in Figure 7.18.

7.7.3 DY 1L1T control region

A set of events enriched in prompt taus originating from $\mathrm{Z} \to \tau\tau$ decays is created by an opposite-sign 1L1T selection, where events are required to have $M_{\mathrm{T}} < 40\,\mathrm{GeV}$ computed with the light lepton and $p_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\mathrm{R} < 3.5$ between the lepton pair, $p_{\mathrm{T}}^{\mathrm{miss}} < 100\,\mathrm{GeV}$, and a dilepton mass of 40–120 GeV. The triggering light lepton originates from a leptonic decay of one of the prompt taus coming from the Z boson. A 2D implementation of the matrix method is used to estimate the contribution due to misidentified leptons. The normalization and Z-p_{T} correction of the NLO MadGraph5_aMC@NLO DY MC are taken from the 2L OS control region and are observed to be valid in the 1L1T channel as well.

The distributions of key kinematic and event variables in the $\mathrm{Z} \to \tau\tau$ control region are given in Figure 7.19 illustrating overall good agreement between observations and the predictions.

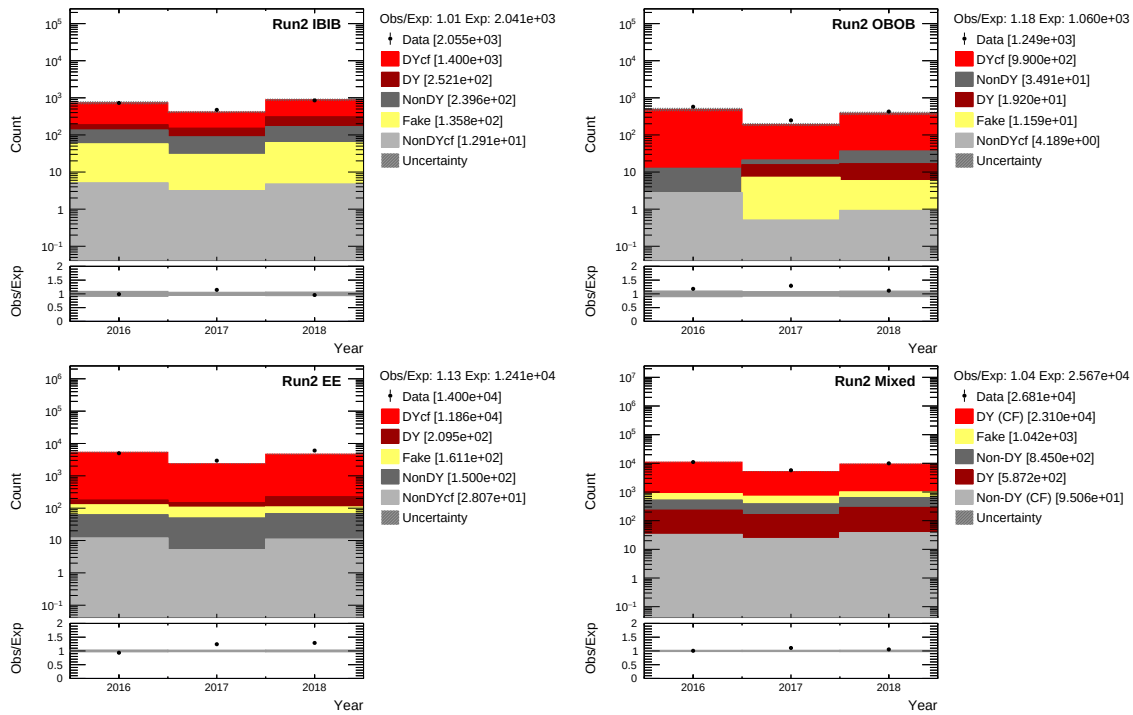

Figure 7.18: 2L SS control region for the Run-2 dataset in the IB-IB (upper left), OB-OB (upper right), and E-E (lower left) category before correction factors are applied, and in the mixed-η (lower right) category after correction factors are applied. Charge flip contributions from MC samples are indicated by the label "CF". The uncertainties are statistical only.

The prompt tau contributions originating from $\mathrm{Z} \to \tau\tau$ decays correspond to the mass peak approximately below 80 GeV, whereas those above 80 GeV are mostly originating from $\mathrm{Z} \to \ell\ell$ contributions where a light lepton is misidentified as a hadronic tau.

7.7.4 $\mathrm{t\bar{t}}$ 2L control region

A set of events enriched in $\mathrm{t\bar{t}}$ process is created in the 2L channel. We define a selection with an opposite-sign $\mathrm{e}\mu$ pair and $N_\mathrm{j} > 1$ as the main control region. In addition, a separate enriched region with OffZ opposite-sign same flavor pairs, $p_\mathrm{T}^\mathrm{miss} > 50\,\mathrm{GeV}$, $N_\mathrm{b} > 0$ and $N_\mathrm{j} > 2$ is defined. A 2D implementation of the matrix method is used to estimate the contribution due to misidentified leptons.

The NLO POWHEG $\mathrm{t\bar{t}}$ MC sample is normalized to the observed data in the main opposite-sign $\mathrm{e}\mu$ control region. Furthermore, the S_T shape of the MC sample is corrected to agree

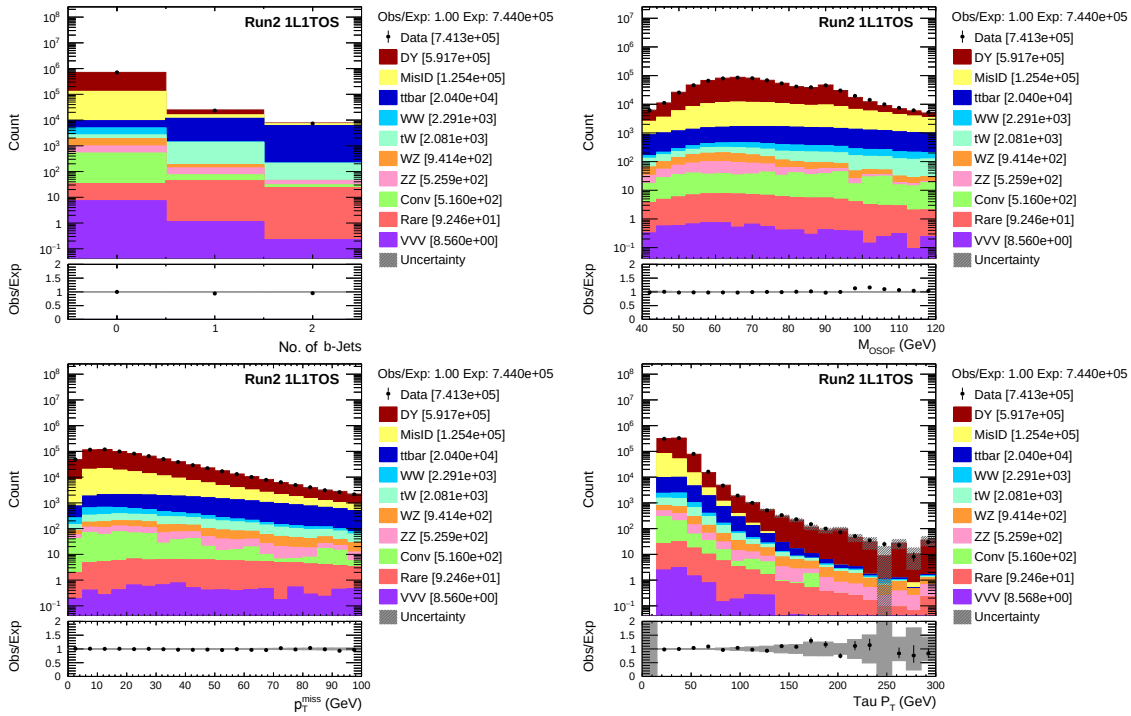

Figure 7.19: 1L1T Z $\rightarrow \tau\tau$ control region for the Run-2 dataset. The uncertainties are statistical only.

with that of the data in the opposite-sign $\mu\mu$ region and validated in the corresponding opposite-sign ee region. The corrected S_{T} shapes for the OS eμ selection, as well as the OS $\mu\mu$ and OS ee selections are shown in Figure 7.20. The distributions of key kinematic and event variables in the $\mathrm{t\bar{t}}$ control region are given in Figure 7.21 illustrating overall good agreement between observations and the predictions.

7.8 Preselection bins

Since we probe the entire multilepton landscape, a coarse-grained binning scheme is used as a preselection to assess the data quality and the performance of the background estimation for the various SM contributions. We use OSSFn, M_{OSSF}, and N_{b} as the main variables to define the preselection binning scheme for each channel. No selection criteria in potentially signal sensitive variables, such as $L_{\mathrm{T}} + p_{\mathrm{T}}^{\mathrm{miss}}$ or S_{T}, are imposed. Thus, this preselection of events is highly dominated by SM backgrounds in each bin and shows each control region in

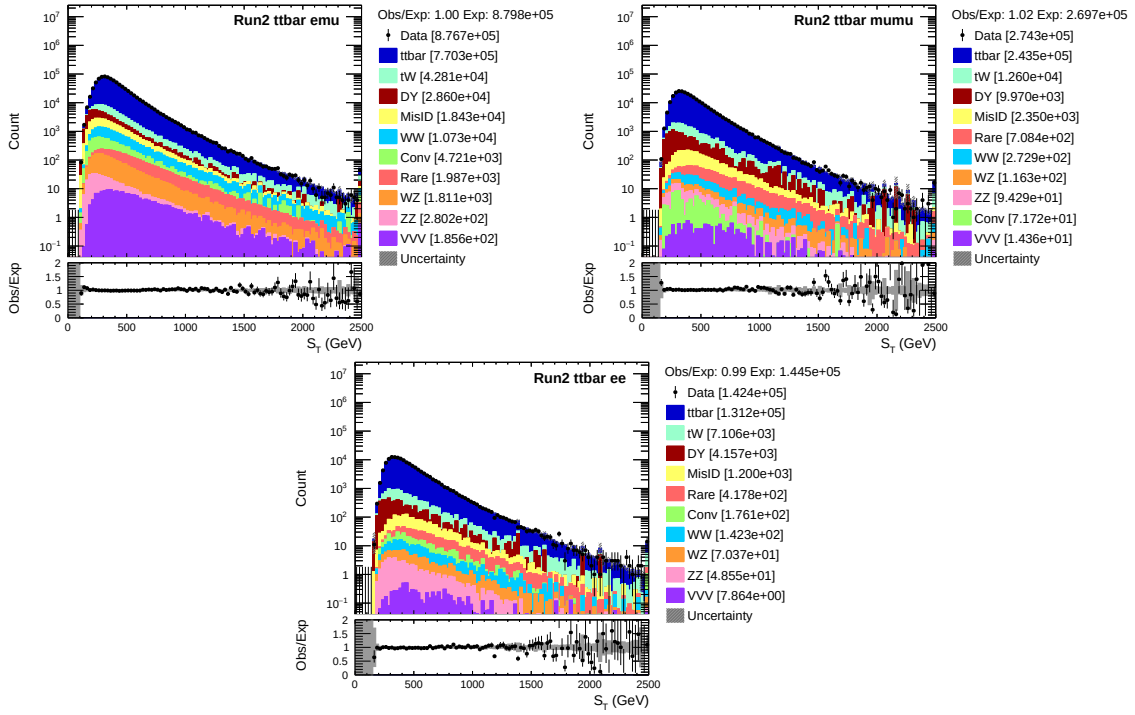

Figure 7.20: 2L $t\bar{t}$ S_T shapes in the OS eμ (top left), in the OS $\mu\mu$ (top right), and in the OS ee (bottom) control region for the Run-2 dataset. The uncertainties are statistical only.

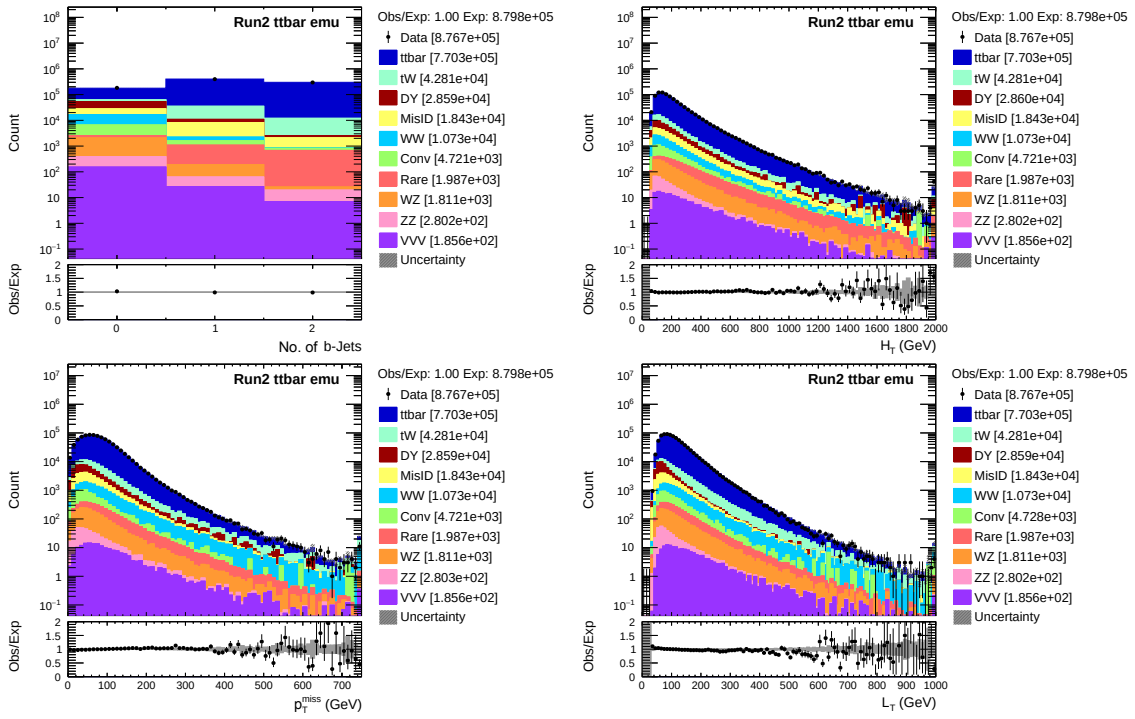

Figure 7.21: 2L $t\bar{t}$ OS eμ control region for the Run-2 dataset. The uncertainties are statistical only.

the 3L, 2L1T and 4L channels as a dedicated bin. The 1L3T, 2L2T and 3L1T channels are combined into one bin to ensure a sufficient SM background estimation avoiding potentially signal sensitive bins. The scheme of the preselection bins is given in Table 7.3.

We observe very good overall agreement between data and expected background yields across all preselection bins in all channels, as shown in Figure 7.22.

Table 7.3: Preselection binning scheme.

Bin	3L	2L1T	1L2T	4L	3L1T+2L2T +1L3T
1	OSSF1, OnZ, 0B, $M_{\mathrm{T}} < 50\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mathrm{miss}} < 125\,\mathrm{GeV}$	OSSF1, OnZ, $p_{\mathrm{T}}^{\mathrm{miss}} < 100\,\mathrm{GeV}$	OSSF1, 0B, $M_{\mathrm{T}}^{\mathrm{e},\mu} < 50\,\mathrm{GeV}$	OSSF2, Double-OnZ, 0B	Inclusive
2	OSSF1, OnZ, 0B, $50 < M_{\mathrm{T}} < 150\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mathrm{miss}} < 125\,\mathrm{GeV}$	OSSF1, OnZ, 0B, $p_{\mathrm{T}}^{\mathrm{miss}} > 100\,\mathrm{GeV}$	OSSF1, 0B, $M_{\mathrm{T}}^{\mathrm{e},\mu} > 50\,\mathrm{GeV}$	OSSF2, OnZ/OffZ, 0B	
3	OSSF1, OnZ, 1+B, $M_{\mathrm{T}} < 150\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mathrm{miss}} < 125\,\mathrm{GeV}$	OSSF1, BelowZ, 0B	OSSF0, 0B	OSSF1 or OSSF0, 0B	
4	OSSF1, BelowZ, 0B, $M_{3\ell}$ OnZ	OSSF1, AboveZ, 0B	1+B	1+B	
5	OSSF1, OnZ, 0B, $M_{\mathrm{T}} > 150\,\mathrm{GeV}$ or $p_{\mathrm{T}}^{\mathrm{miss}} > 125\,\mathrm{GeV}$	OSSF0, 0B, SS ee OnZ			
6	OSSF1, BelowZ, 0B	OSSF0, 0B			
7	OSSF1, MixedZ, 0B	1+B			
8	OSSF1, AboveZ, 0B				
9	OSSF0, 0B				
10	1+B				

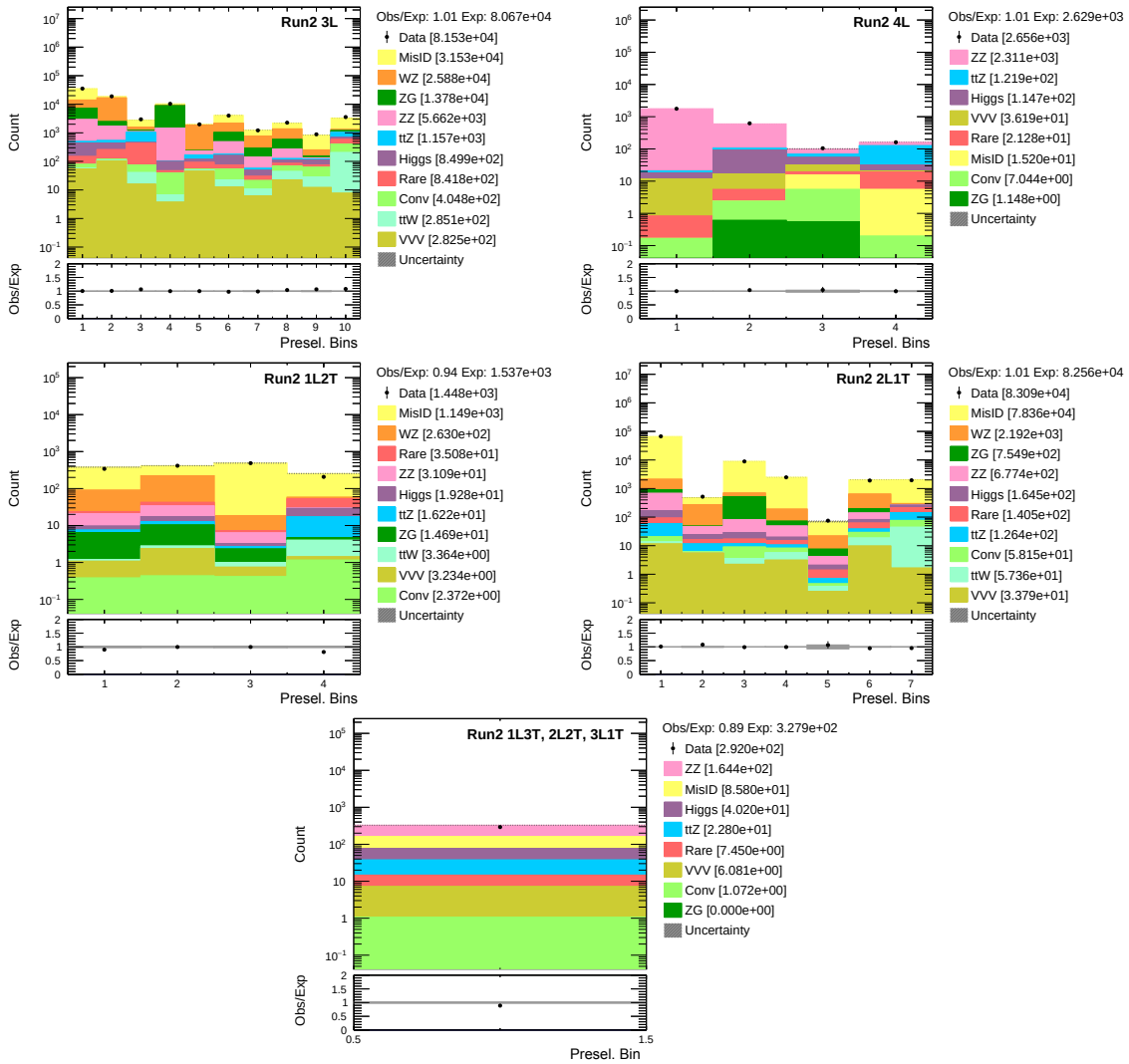

Figure 7.22: Preselection bins for the Run-2 dataset. The uncertainties are statistical only.

Chapter 8

Systematic uncertainties

The outline of all systematic uncertainties is given in this chapter. Their estimation and propagation to the analysis is described in the following paragraphs.

All background and signal estimates have statistical uncertainties due to the finite number of events in simulated samples or in data used for the matrix method. Although we utilize high-statistics MC samples, and therefore these statistical uncertainties are typically small, we propagate those uncertainties to the analysis.

The relative inclusive normalization uncertainties for WZ, Zγ, t$\bar{\text{t}}$Z, and ZZ are 3–5%, 10%, 15–25%, and 4–5% respectively, in all three years. All normalizations and uncertainties, including the correction of the jet multiplicity distribution and the diboson p_{T}, are measured in dedicated control regions as discussed in Sections 7.2-7.3. These uncertainties lead to variations of 5–30% for the jet multiplicity correction, and 5–15% for the diboson p_{T} correction, and are taken to be uncorrelated between the different channels.

For the subdominant, rare background processes such as t$\bar{\text{t}}$W, triboson, or associated Higgs production, a 50% systematic uncertainty is applied on the theoretical LO or NLO cross-sections to cover any higher order effects as well as renormalization and factorization scale uncertainties [103].

A number of systematic uncertainty sources are considered to account for differences in the modeling of pileup, lepton trigger, reconstructed energy scale and resolution, object identification, as well as lepton isolation between data and MC events. The pileup reweighting uncertainty is evaluated by varying the minimum bias cross-section [104] used in the correction procedure up and down by 5%, and is applied to all MC based backgrounds. The changes in the yields of WZ, ZZ, and $t\bar{t}Z$ backgrounds are observed to be $< 3\%$. The uncertainties on the electron, muon, and tau reconstruction, identification, and isolation efficiencies are applied, and are cumulatively in the range of 1–5% for electrons, $< 2\%$ for muons. and in the range of 5–15% for taus. The individual lepton trigger efficiencies in data and MC have systematic uncertainties of 1–4% per lepton leg, and the final uncertainty on the trigger efficiency weight per event is conservatively estimated by varying the data and MC efficiencies up and down in an anti-correlated way (i.e. vary data efficiency by $+2\%$ and MC efficiency by -2%). This generally results in a negligible variation due to the availability of multiple triggering leptons in a large part of the targeted phase space, and is observed to be at most 3% at very low values of L_{T}. Uncertainties on the b tag efficiencies applied to MC samples typically cause a $< 10\%$ variation in the yields throughout the signal regions.

The uncertainties on the jet, tau, muon, and electron energy scale are applied at the per-object level, where the corresponding object p_{T} is varied up and down. Similarly, the uncertainty on isolated (unclustered) energy deposits in the calorimeter is propagated to the analysis by varying their amount up and down. We also apply the energy resolution corrections for electrons and muons. All energy scale and resolution variations are then propagated to the kinematic quantities such as lepton p_{T}, L_{T}, $p_{\mathrm{T}}^{\mathrm{miss}}$, M_{T}, H_{T}, S_{T} and M_{OSSF}. To avoid statistical fluctuations driving these uncertainties, the energy scale variations are smoothened based on the running average between connected bins in the tables. Among these uncertainties, the lepton energy scales are the dominant sources of uncertainty, typically resulting in variations up to $\sim 5\%$, whereas all other energy scale and resolution uncertainties are smaller in comparison. These uncertainties apply to signal processes as well as to background contributions due to irreducible (prompt) processes, both of which are estimated using simulated samples.

Electron charge misidentification uncertainties are conservatively taken as the size of the largest correction as derived in the 2L same-sign control region, and hence a 30% relative uncertainty is assigned to such contributions. Additionally, for all rare background estimates as well as signal yields obtained directly from simulation using theoretical cross-sections, luminosity uncertainties are in the range of 1.2–2.5% in each year of data collection. Similarly, uncertainties due to choices of factorization scale, renormalization scale and PDF are also evaluated for signal and dominant irreducible background processes, yielding variations of up to 10% in the signal regions.

The misidentified lepton background contributions estimated via the data-driven matrix method have associated systematic uncertainties dominated by the uncertainties of the lepton fake rates. The relative uncertainties on the fake rates are typically in the 10–30% range. As we use an extrapolation of the fake rate measurements for light lepton $p_{\mathrm{T}} > 50$ GeV, and tau $p_{\mathrm{T}} > 80\,\mathrm{GeV}$, we double these uncertainties and assign a 60% relative uncertainty for all high p_{T} leptons. Therefore, lepton fake rates have typical relative uncertainties of 10% for low lepton p_{T} ($10 < p_{\mathrm{T}} < 20\,\mathrm{GeV}$ for light leptons, $10 < p_{\mathrm{T}} < 30\,\mathrm{GeV}$ for taus), 30% for medium lepton p_{T} ($20 < p_{\mathrm{T}} < 50\,\mathrm{GeV}$ for light leptons, $30 < p_{\mathrm{T}} < 80\,\mathrm{GeV}$ for taus), and 60% for high lepton p_{T} ($p_{\mathrm{T}} > 50\,\mathrm{GeV}$ for light leptons, $p_{\mathrm{T}} > 80\,\mathrm{GeV}$ for taus), yielding variations in the range of 20-50% on the misidentified lepton background estimates. Independent of the statistical uncertainties on the fake rate measurements, we also consider a process dependent uncertainty on the lepton fake rates, which is estimated by comparing the relative spread of the misidentified lepton background yields calculated only with DY and only with $\mathrm{t\bar{t}}$ based fake and prompt rates. This relative spread is typically found to be in the 5–25% range for all lepton flavors. In order to account for different compositions of misidentified lepton origins in multilepton events, the systematic uncertainties for the misidentified lepton backgrounds in different table categories are taken to be uncorrelated.

The uncertainty sources considered in this analysis including the respective typical magnitudes, relevant processes, and the typical resulting variations, are summarized in Table 8.1.

Table 8.1: Sources, magnitudes, effective variations, and correlation model of systematic uncertainties in signal regions. Uncertainty sources marked as "Yes" under the correlation model have their nuisance parameters correlated across the 3 years of data collection.

Uncertainty source	Magnitude	Type	Processes	Variation	Correlation
Statistical	1–100%	per event	All MC samples	1–100%	No
Luminosity	1.2–2.5%	per event	Conv./Rare/Signal	1.2–2.5%	Yes
Electron/Muon reco., ID and iso. efficiency	1–5%	per lepton	All MC samples	2–5%	No
Tau reco., ID and iso. efficiency	5–15%	per lepton	All MC samples	5–25%	No
Lepton displacement efficiency	1–2%	per lepton	All MC samples	3–5%	No
Trigger efficiency	1–4%	per lepton	All MC samples	$< 3\%$	No
b tag efficiency	1–10%	per jet	All MC samples	2–5%	No
Pileup	5%	per event	All MC samples	$< 3\%$	Yes
PDF, fact./renorm. scale	$< 20\%$	per event	WZ/ZZ/$t\bar{t}$Z/Signal	$< 10\%$	Yes
Jet energy scale	1–10%	per jet	All MC samples	$< 5\%$	No
Unclustered energy scale	1–25%	per event	All MC samples	$< 2\%$	No
Muon energy scale and resolution	2%	per lepton	All MC samples	$< 5\%$	No
Electron energy scale and resolution	$< 2\%$	per lepton	All MC samples	$< 5\%$	Yes
Tau energy scale	$< 10\%$	per lepton	All MC samples	$< 5\%$	No
Electron charge misidentification	30%	per lepton	All MC samples	$< 25\%$	No
WZ normalization	3–5%	per event	WZ	3–5%	No
ZZ normalization	4–5%	per event	ZZ	4–5%	No
$t\bar{t}$Z normalization	15–25%	per event	$t\bar{t}$Z	15–25%	No
Conversion normalization	10–50%	per event	Zγ/Conv.	10–50%	No
Rare normalization	50%	per event	Rare	50%	No
Prompt and fake rates	10–60%	per lepton	MisID	20–50%	No
DY–$t\bar{t}$ process dependency	5–25%	per lepton	MisID	5–25%	Yes
Diboson jet multiplicity modeling	$< 30\%$	per event	WZ/ZZ	5–30%	No
Diboson p_{T} modeling	$< 30\%$	per event	WZ/ZZ	5–15%	No

Chapter 9

Results

This chapter concludes the description of the analysis by summarizing the results of the analysis, i.e. showing the distributions of the different table schemes and calculating the exclusion limits for the probed signal models. In the following figures, the diboson background processes (WZ and ZZ) are denoted as "VV", whereas the $t\bar{t}Z$ and $t\bar{t}W$ contributions are combined into "ttV". Background processes involving lepton conversion, including the $Z\gamma$ process, are labeled as "Conv.". The other irreducible background processes consisting of triboson, Higgs, and other rare SM contributions, are collectively given as "Rare" backgrounds.

9.1 Signal region distributions

The $L_{\mathrm{T}}+p_{\mathrm{T}}^{\mathrm{miss}}$ distributions of the fundamental table scheme for all channels with the Run-2 dataset are shown in Figure 9.1, where each histogram bin corresponds to an orthogonal $L_{\mathrm{T}}+p_{\mathrm{T}}^{\mathrm{miss}}$ bin in regions defined by the fundamental table scheme in Chapter 6. Similarly, the S_{T} distributions of the fundamental table scheme and the advanced table scheme are shown in Figure 9.2, and Figures 9.3-9.6, respectively.

The observations are found to be consistent with the expectations from Standard Model processes. We perform a goodness-of-fit test for each table scheme using the saturated

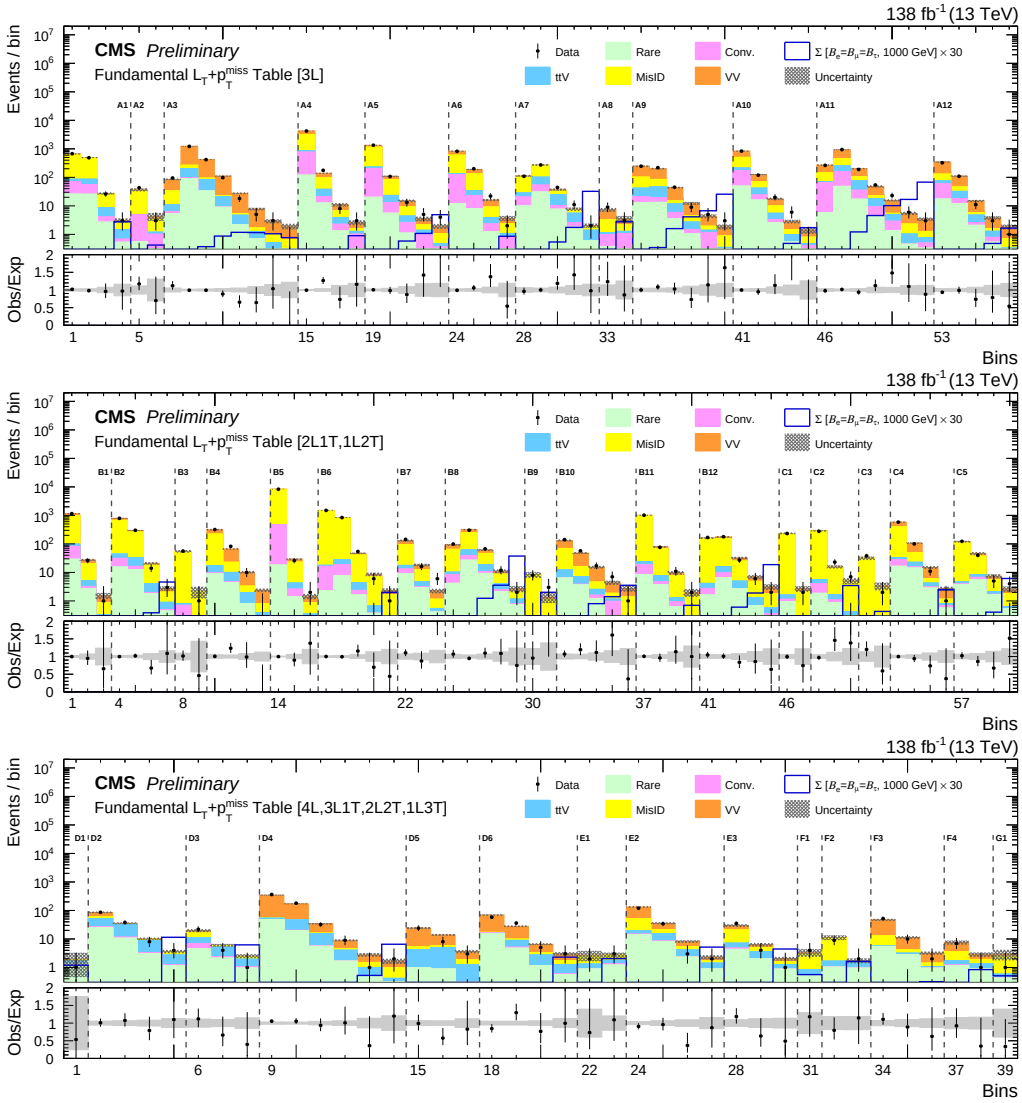

Figure 9.1: Combined signal region distributions of the fundamental $L_\mathrm{T}+p_\mathrm{T}^\mathrm{miss}$ table scheme for the Run-2 dataset.

model method [105] to quantify the agreement between the background estimation and the observed data, the so-called background-only hypothesis. For this hypothesis, we calculate a global p-value of 0.67, 0.53, and 0.11 for the fundamental $L_\mathrm{T}+p_\mathrm{T}^\mathrm{miss}$, the fundamental S_T, and the advanced table scheme, respectively.

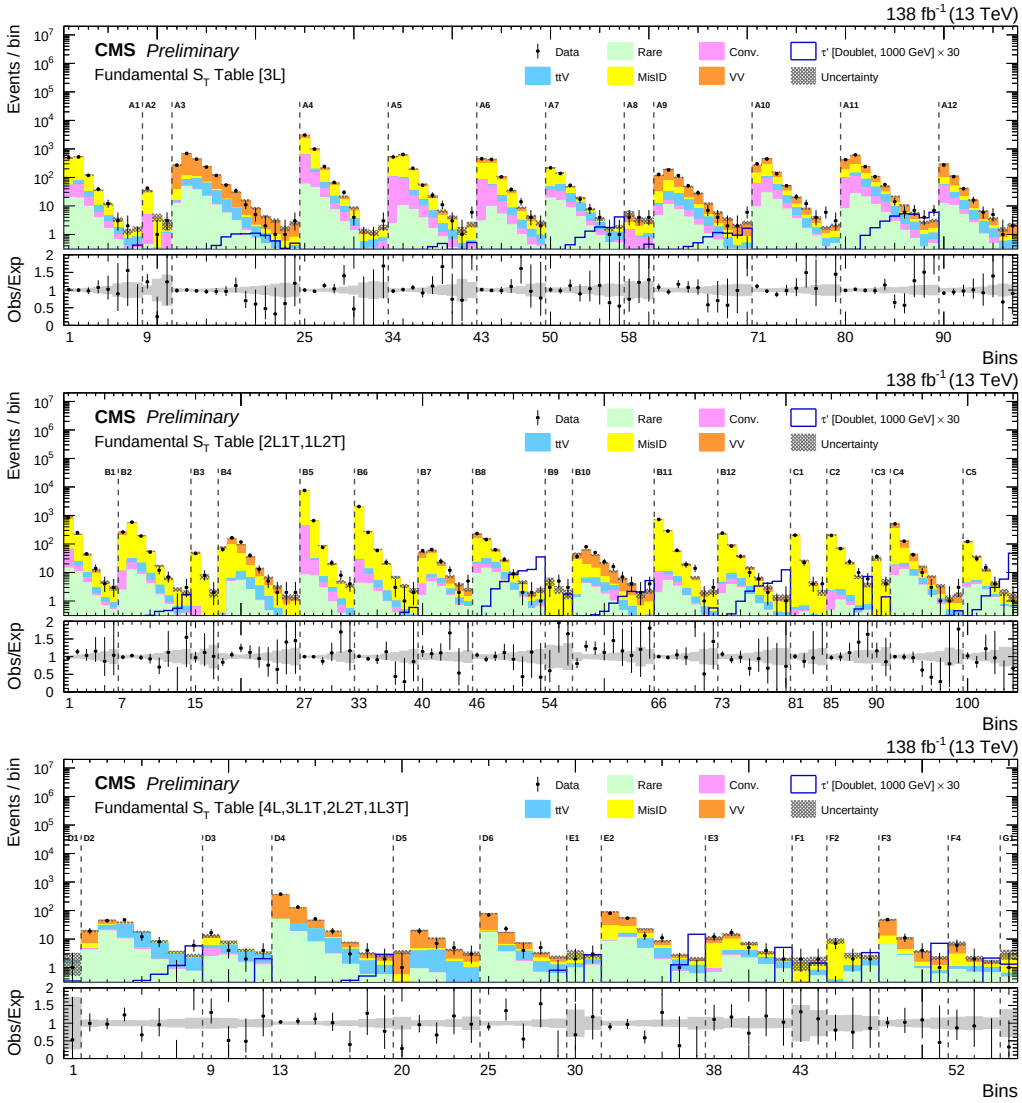

Figure 9.2: Combined signal region distributions of the fundamental S_{T} table scheme for the Run-2 dataset.

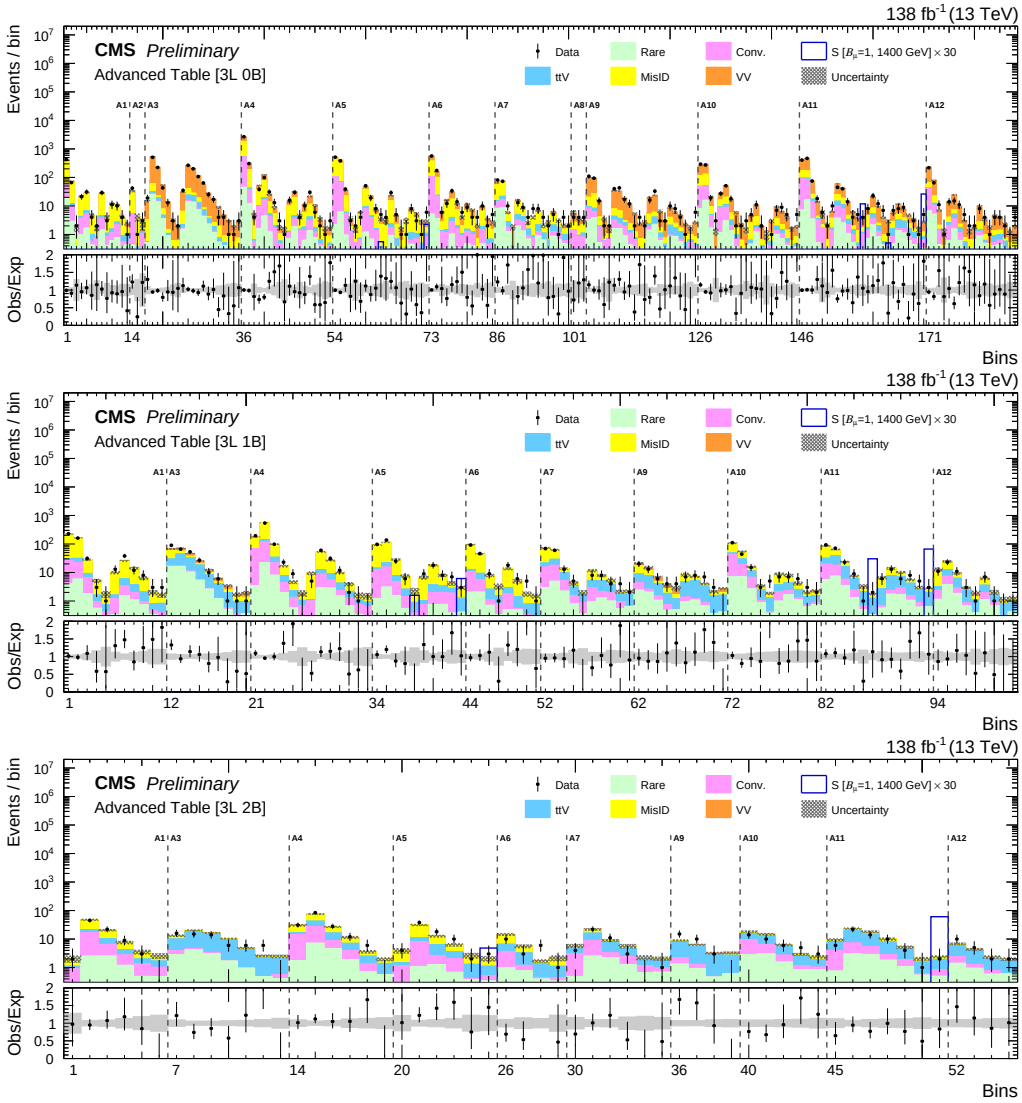

Figure 9.3: 3L signal region distributions of the advanced table scheme for the Run-2 dataset.

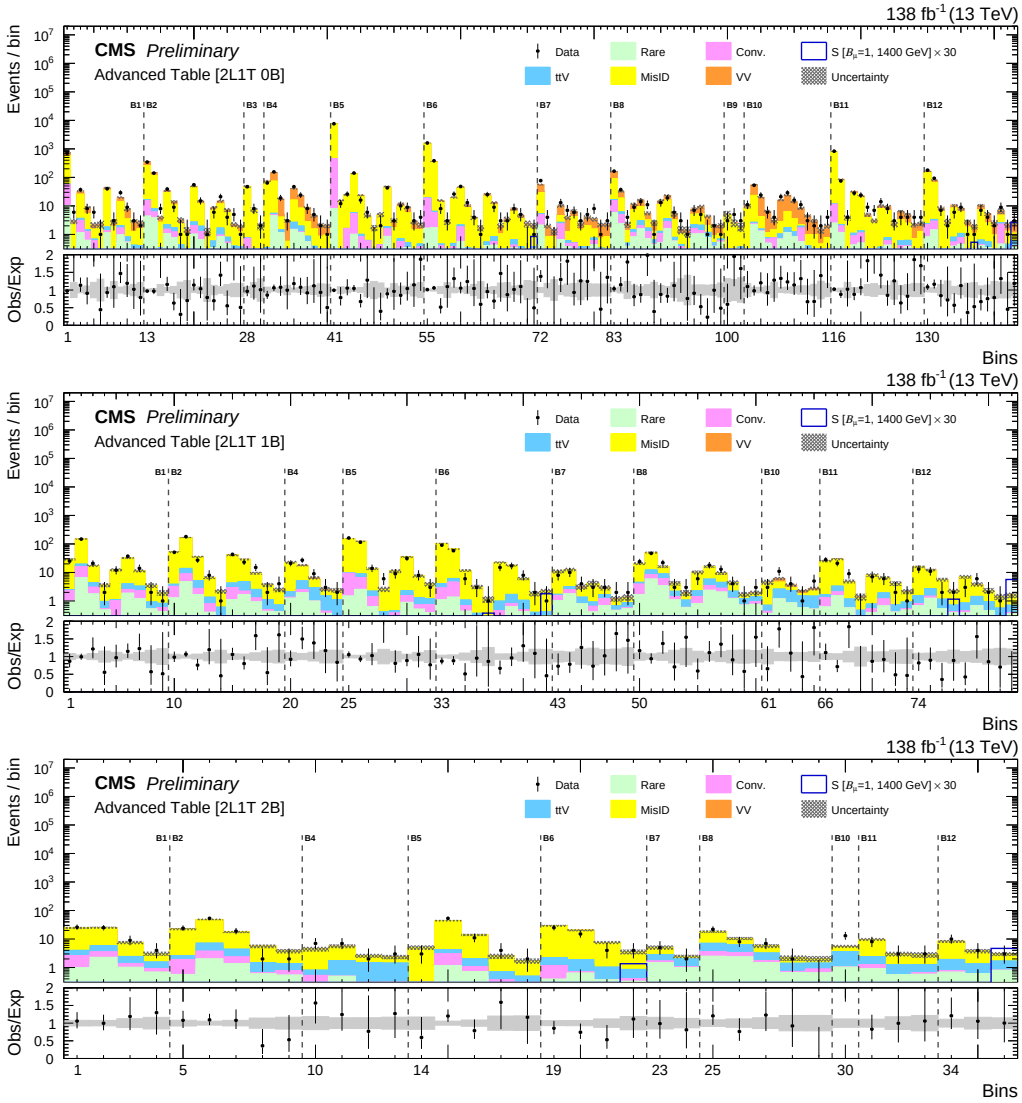

Figure 9.4: 2L1T signal region distributions of the advanced table scheme for the Run-2 dataset.

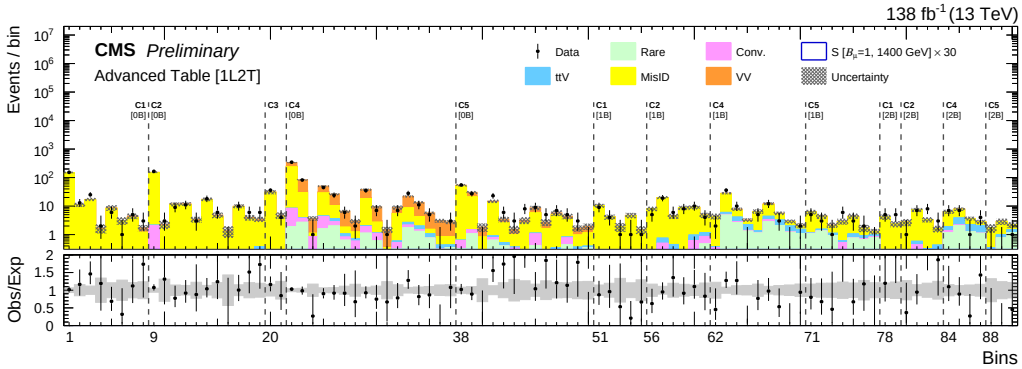

Figure 9.5: 1L2T signal region distribution of the advanced table scheme for the Run-2 dataset.

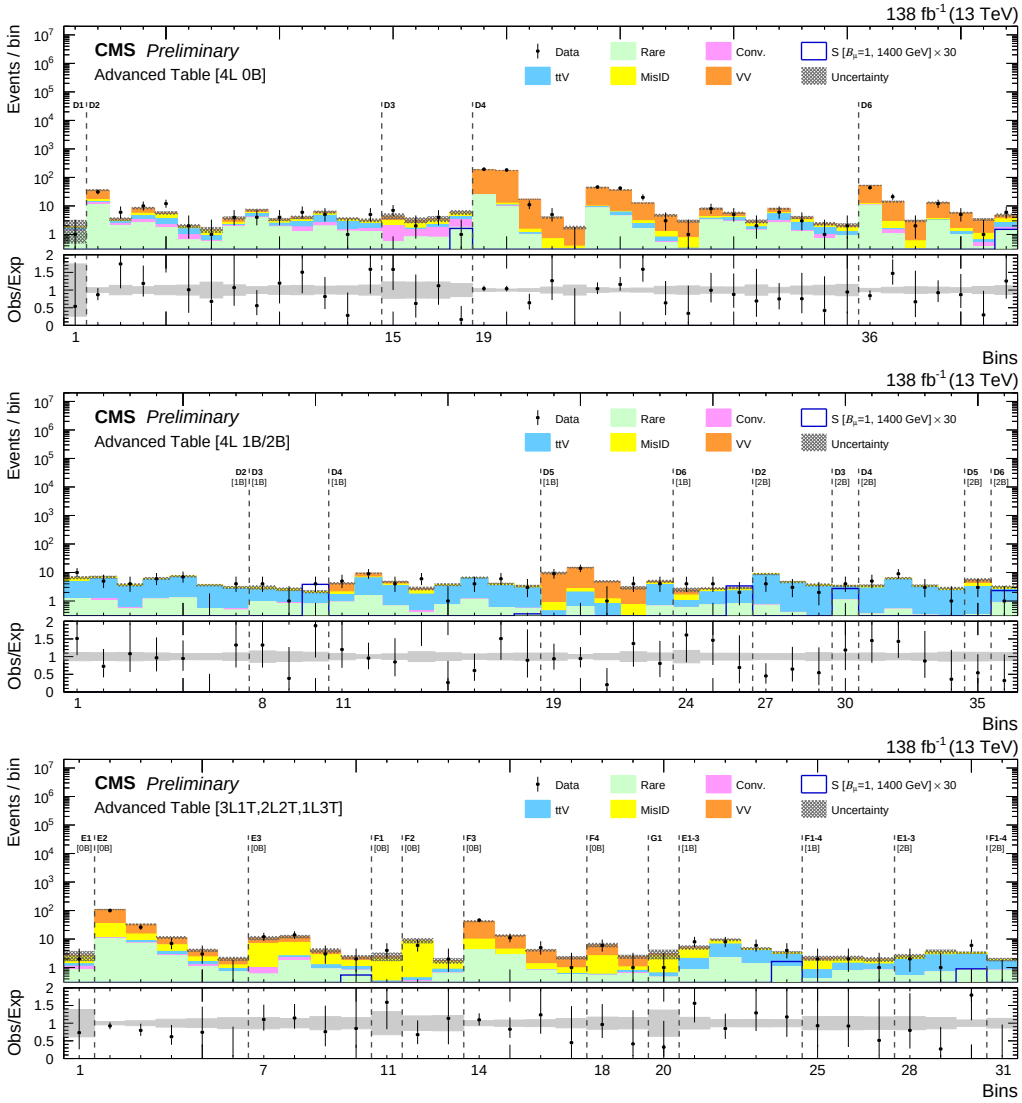

Figure 9.6: 4L, 3L1T, 2L2T and 1L3T signal region distributions of the advanced table scheme for the Run-2 dataset.

9.2 Exclusion limits

We calculate upper limits on the production cross section of the signal models probed in this analysis. These limits are evaluated separately for each of the three model independent table schemes. Each bin of the tables given in Figures 9.1-9.6 is treated as an independent counting experiment for each year and combined with the other bins in each table for all three years. We use a modified frequentist approach with the $\mathrm{CL_S}$ [106, 107, 108, 109] criterion, with a test statistic based on the binned profile likelihood in the asymptotic approximation. The upper limits are calculated at 95% confidence level (CL). The systematic uncertainties and their correlations as described in Section 8 are propagated to the likelihood as nuisance parameters with log-normal probability density functions. The statistical uncertainties are modeled with gamma functions.

The impact of the systematic uncertainties on the table schemes and their comparison is illustrated with the expected signal strength (r-value) ratio in Figures 9.7-9.9 on each signal model with the Run-2 dataset. We observe that at high masses the impact of systematic uncertainties is small. The high mass sensitive bins in the tables have low SM background population and therefore statistical uncertainties become the dominant nuisance. At low signal masses, signal events populate the SM background dominated phase space. Therefore, the impact of systematic uncertainties is the largest.

The fundamental $L_\mathrm{T}+p_\mathrm{T}^\mathrm{miss}$ table scheme is preferred for models with low hadronic activity, such as type-III seesaw and vector-like leptons. For models involving strong interactions, such as leptoquarks, the fundamental S_T scheme is preferred. For type-III seesaw heavy fermions, the fundamental scheme has a higher sensitivity for masses above 460 GeV assuming the lepton-flavor democratic or the $\mathcal{B}_\mathrm{e} = 1$ scenario, for masses above 430 GeV assuming the $\mathcal{B}_\mu = 1$ scenario, and for masses above 680 GeV assuming the $\mathcal{B}_\tau = 1$ scenario. Similarly, a higher sensitivity for vector-like taus in the fundamental scheme is reached for masses above 910 GeV in the doublet model and for masses above 630 GeV in the singlet model. For leptoquarks, the fundamental scheme is preferred for masses above 870 GeV

assuming the $\mathcal{B}_{\mathrm{e}} = 1$ scenario, for masses above 900 GeV assuming the $\mathcal{B}_{\mu} = 1$ scenario, and for masses above 720 GeV assuming the $\mathcal{B}_{\tau} = 1$ scenario. This transition is indicated by a vertical gray line in Figures 9.10-9.12. Limits from the fundamental table are shown above the gray line, the advanced table results are shown below the gray line.

Figure 9.10 shows the exclusion limits for the type-III seesaw heavy fermions for the lepton-flavor democratic scenario, as well as the $\mathcal{B}_{\mathrm{e}} = 1$, $\mathcal{B}_{\mu} = 1$ and $\mathcal{B}_{\tau} = 1$ scenarios. Type-III seesaw heavy fermions are observed (expected) to be excluded at 95% confidence level for masses below 940 (940) GeV assuming the lepton-flavor democratic scenario, for masses below 980 (980) GeV assuming the $\mathcal{B}_{\mathrm{e}} = 1$ scenario, for masses below 1060 (1050) GeV assuming the $\mathcal{B}_{\mu} = 1$ scenario, and for masses below 740 (770) GeV assuming the $\mathcal{B}_{\tau} = 1$ scenario.

Figure 9.11 shows the exclusion limits for the vector-like taus in the doublet model, as well as the singlet model. Vector-like taus are observed (expected) to be excluded at 95% confidence level for masses below 920 (920) GeV for the doublet model, while vector-like taus in the singlet model are observed to be excluded in the mass range of 120–180 GeV and expected to be excluded at a mass of 150 GeV.

Figure 9.12 shows the exclusion limits for scalar leptoquarks with coupling to a top quark and a charged lepton for the $\mathcal{B}_{\mathrm{e}} = 1$, $\mathcal{B}_{\mu} = 1$ and $\mathcal{B}_{\tau} = 1$ scenarios. Scalar leptoquarks with coupling to a top quark and a charged lepton are observed (expected) to be excluded at 95% confidence level for masses below 1230 (1340) GeV assuming the $\mathcal{B}_{\mathrm{e}} = 1$ scenario, for masses below 1300 (1430) GeV assuming the $\mathcal{B}_{\mu} = 1$ scenario, and for masses below 1230 (1200) GeV assuming the $\mathcal{B}_{\tau} = 1$ scenario.

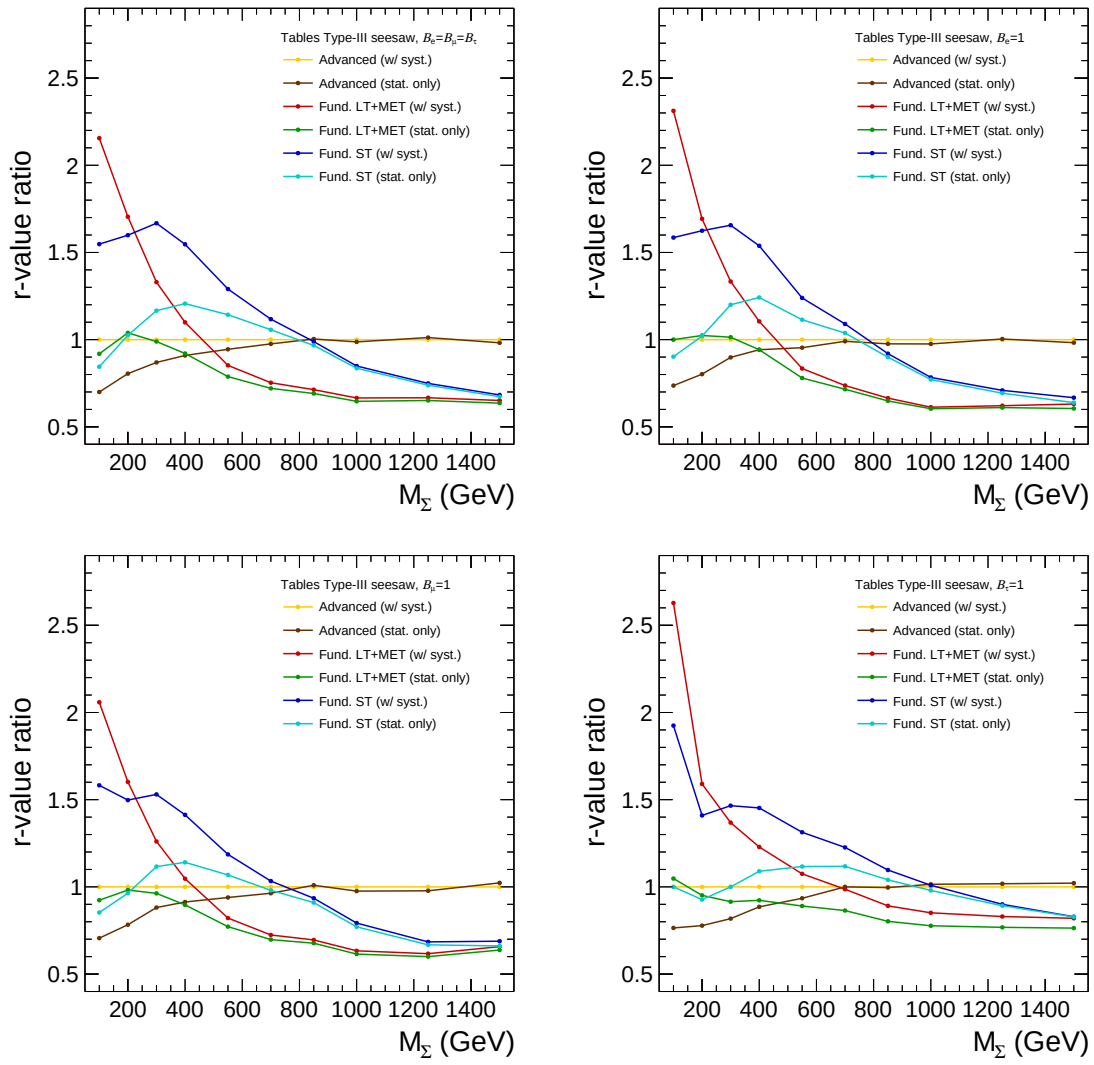

Figure 9.7: The impact of the systematic uncertainties and the comparison between the different table schemes on the expected upper cross-section limits for the the different lepton flavor branching ratio scenarios for type-III seesaw heavy fermions. The results are presented as an r-value ratio with respect to the advanced table scheme with systematic uncertainties.

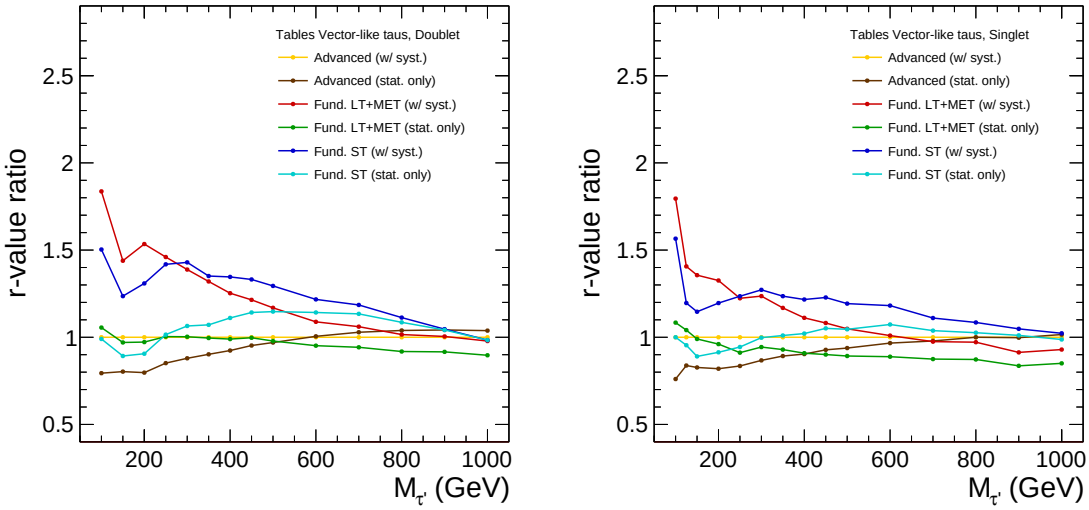

Figure 9.8: The impact of the systematic uncertainties and the comparison between the different table schemes on the expected upper cross-section limits for the doublet and singlet vector-like tau signal models. The results are presented as an r-value ratio with respect to the advanced table scheme with systematic uncertainties.

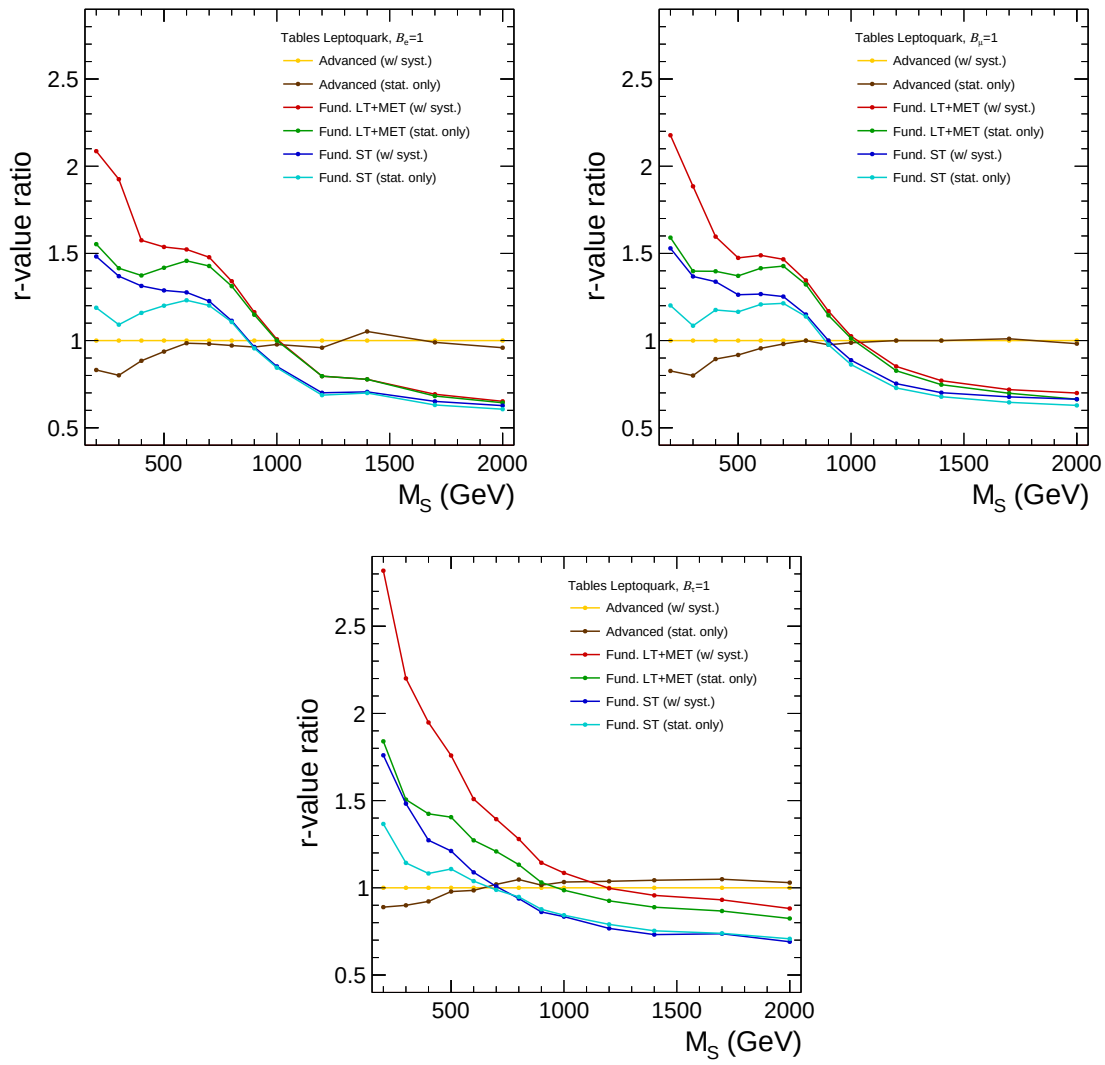

Figure 9.9: The impact of the systematic uncertainties and the comparison between the different table schemes on the expected upper cross-section limits for the the different lepton flavor branching ratio scenarios for scalar leptoquarks with coupling to a top quark and a charged lepton. The results are presented as an r-value ratio with respect to the advanced table scheme with systematic uncertainties.

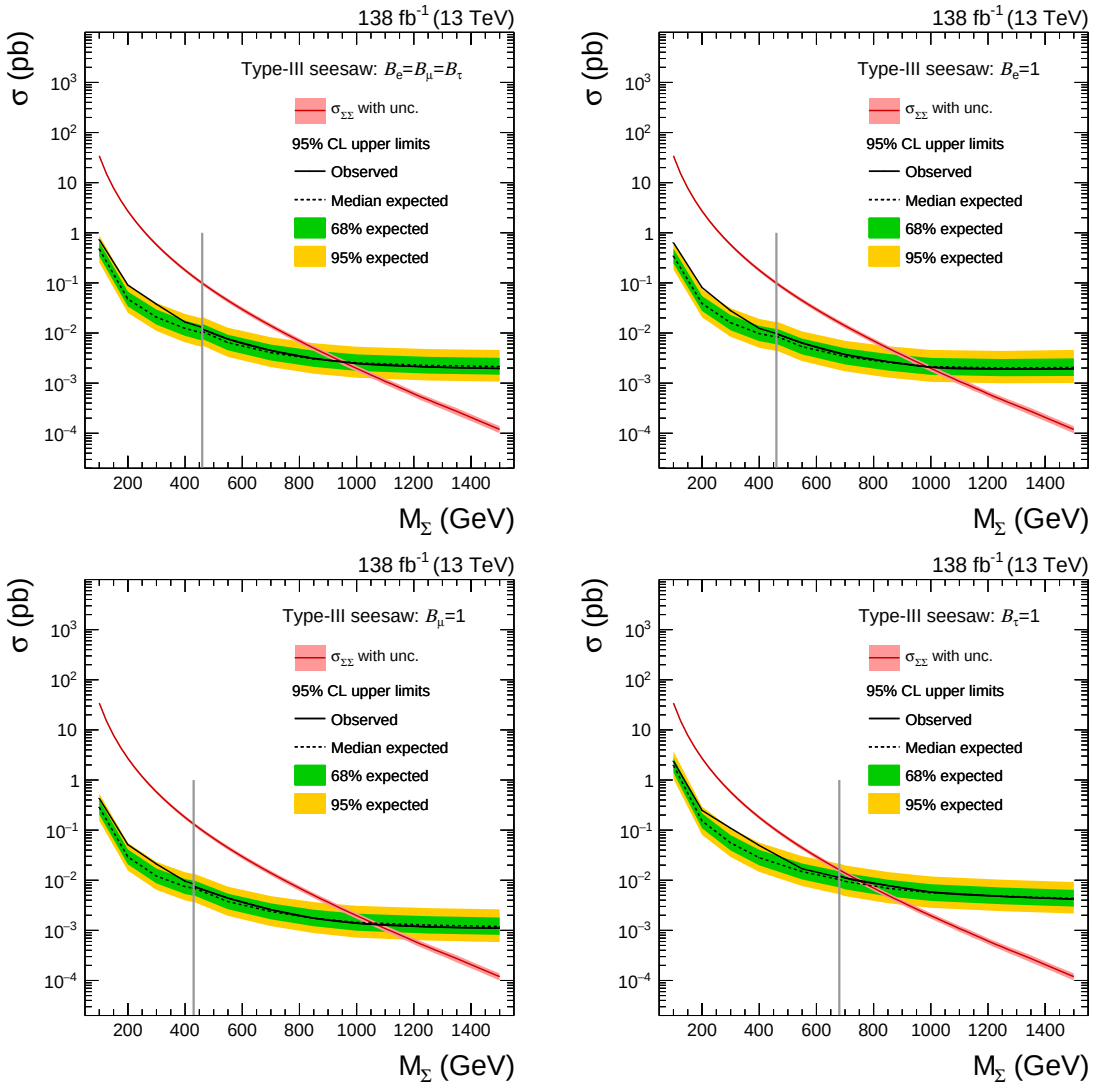

Figure 9.10: Exclusion limits at 95% CL for the type-III seesaw fermions in the flavor democratic (upper left), the $\mathcal{B}_e = 1$ (upper right), the $\mathcal{B}_\mu = 1$ (lower left) and the $\mathcal{B}_\tau = 1$ (lower right) scenario. Limits from the fundamental table are shown above the gray line, the advanced table results are shown below the gray line.

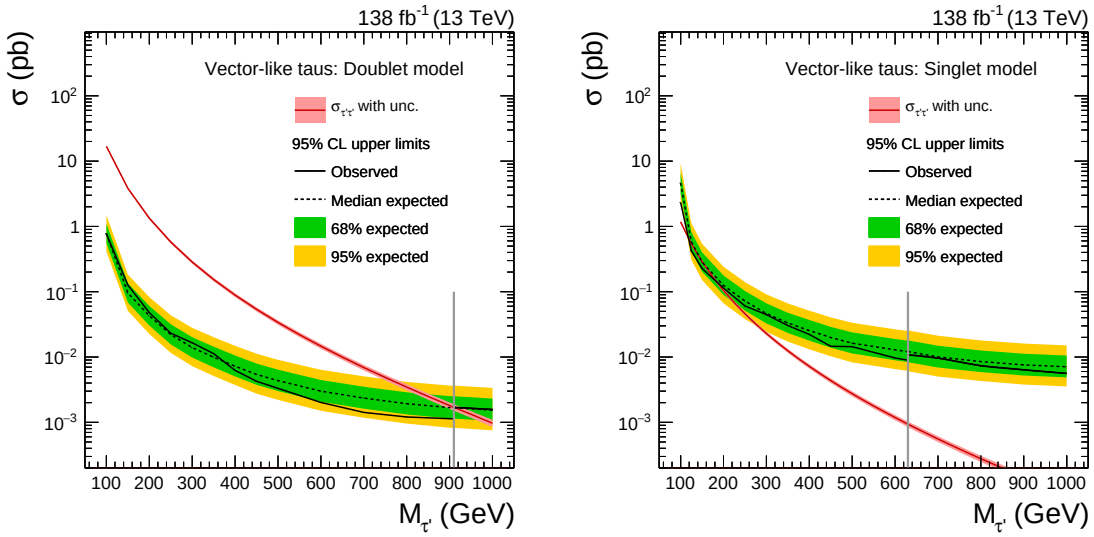

Figure 9.11: Exclusion limits at 95% CL for the vector-like taus in the doublet (left) and singlet (right) models. Limits from the fundamental table are shown above the gray line, the advanced table results are shown below the gray line.

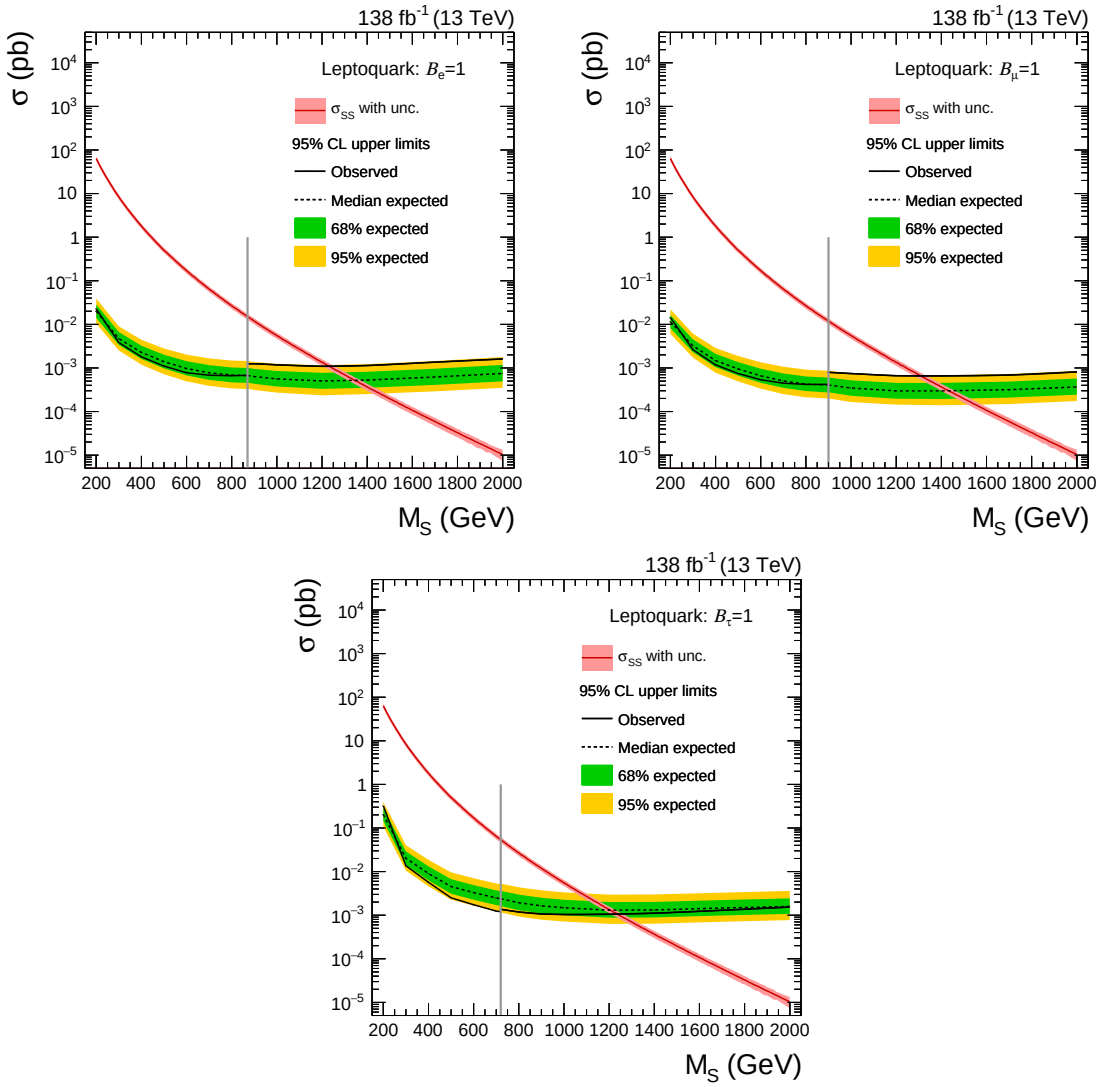

Figure 9.12: Exclusion limits at 95% CL for scalar leptoquarks with coupling to a top quark and a charged lepton in the $\mathcal{B}_e = 1$ (upper left), the $\mathcal{B}_\mu = 1$ (upper right) and the $\mathcal{B}_\tau = 1$ (lower) scenario. Limits from the fundamental table are shown above the gray line, the advanced table results are shown below the gray line.

Chapter 10

Conclusions

We performed a search for new physics in multilepton events in $138\,\mathrm{fb}^{-1}$ of pp collision data collected by the CMS detector during the LHC Run-2. We have interpreted the search results using three different model independent table schemes. The observations are found to be consistent with the expectations from Standard Model processes.

We calculated upper limits at 95% confidence level on the production cross section of the different signal models probed in this analysis. Type-III seesaw heavy fermions are observed (expected) to be excluded for masses below 740 (770) GeV for any branching fraction scenario to Standard Model leptons. The vector-like taus in the doublet model are observed (expected) to be excluded or masses below 920 (920) GeV. Vector-like taus in the singlet model are expected to be excluded at mass 150 GeV while they are observed to be excluded in the mass range 120–180 GeV. Scalar leptoquarks decaying exclusively to a top quark and a charged lepton of any flavor are observed (expected) to be excluded with masses below 1230 (1200) GeV.

Appendices

A List of data and MC samples

Table A.1: Single lepton datasets as well as corresponding run ranges, luminosities, and triggers in 2016, 2017, and 2018.

Dataset Name	Run Range	L (fb^{-1})
/SingleElectron/Run2016B-17Jul2018_ver2-v1/MINIAOD	272007-275376	5.8
/SingleElectron/Run2016C-17Jul2018-v1/MINIAOD	275657-276283	2.6
/SingleElectron/Run2016D-17Jul2018-v1/MINIAOD	276315-276811	4.2
/SingleElectron/Run2016E-17Jul2018-v1/MINIAOD	276831-277420	4.0
/SingleElectron/Run2016F-17Jul2018-v1/MINIAOD	277772-278808	3.1
/SingleElectron/Run2016G-17Jul2018-v1/MINIAOD	278820-280385	7.6
/SingleElectron/Run2016H-17Jul2018-v1/MINIAOD	280919-284044	8.7
	Trigger: HLT_Ele27_WPTight_Gsf_v	
/SingleMuon/Run2016B-17Jul2018_ver2-v1/MINIAOD	272007-275376	5.8
/SingleMuon/Run2016C-17Jul2018-v1/MINIAOD	275657-276283	2.6
/SingleMuon/Run2016D-17Jul2018-v1/MINIAOD	276315-276811	4.2
/SingleMuon/Run2016E-17Jul2018-v1/MINIAOD	276831-277420	4.0
/SingleMuon/Run2016F-17Jul2018-v1/MINIAOD	277772-278808	3.1
/SingleMuon/Run2016G-17Jul2018-v1/MINIAOD	278820-280385	7.6
/SingleMuon/Run2016H-17Jul2018-v1/MINIAOD	280919-284044	8.7
	Triggers: HLT_IsoMu24_v or HLT_IsoTkMu24_v	
/SingleElectron/Run2017B-31Mar2018-v1/MINIAOD	297046-299329	4.8
/SingleElectron/Run2017C-31Mar2018-v1/MINIAOD	299368-302029	9.6
/SingleElectron/Run2017D-31Mar2018-v1/MINIAOD	302030-303434	4.3
/SingleElectron/Run2017E-31Mar2018-v1/MINIAOD	303824-304797	9.3
/SingleElectron/Run2017F-31Mar2018-v1/MINIAOD	305040-306462	13.5
	Triggers: HLT_Ele27_WPTight_Gsf_L1DoubleEG_v or HLT_Ele32_WPTight_Gsf_v	
/SingleMuon/Run2017B-31Mar2018-v1/MINIAOD	297046-299329	4.8
/SingleMuon/Run2017C-31Mar2018-v1/MINIAOD	299368-302029	9.6
/SingleMuon/Run2017D-31Mar2018-v1/MINIAOD	302030-303434	4.3
/SingleMuon/Run2017E-31Mar2018-v1/MINIAOD	303824-304797	9.3
/SingleMuon/Run2017F-31Mar2018-v1/MINIAOD	305040-306462	13.5
	Triggers: HLT_IsoMu27_v	
/EGamma/Run2018A-17Sep2018-v2/MINIAOD	315252-316995	14.0
/EGamma/Run2018B-17Sep2018-v1/MINIAOD	317080-319310	7.1
/EGamma/Run2018C-17Sep2018-v1/MINIAOD	319337-320065	6.9
/EGamma/Run2018D-PromptReco-v2/MINIAOD	320673-325175	31.9
	Triggers: HLT_Ele32_WPTight_Gsf_v	
/SingleMuon/Run2018A-17Sep2018-v2/MINIAOD	315252-316995	14.0
/SingleMuon/Run2018B-17Sep2018-v1/MINIAOD	317080-319310	7.1
/SingleMuon/Run2018C-17Sep2018-v1/MINIAOD	319337-320065	6.9
/SingleMuon/Run2018D-PromptReco-v2/MINIAOD	320673-325175	31.9
	Triggers: HLT_IsoMu24_v	

Table A.2: The list of MC samples and corresponding cross-sections in 2016.

Process name	DAS name	Cross-section (pb)
DYJetsToLL_M-10to50_AMC	/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/†	18610.0
DYJetsToLL_M-50_AMC	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/†	5765.0
ZGToLLG_01J_lowMLL_lowPT_AMC	/ZGToLLG_01J_5f_lowMLL_lowGPt_TuneCP5_13TeV-amcatnloFXFX-pythia8/†	172.8
TTTo2L2Nu_POW	/TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8/†	88.29
TTToSemiLeptonic_POW	/TTToSemilepton_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8/†	365.34
WZTo3LNu_AMC	/WZTo3LNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/†	5.052
WZ2L2Q_AMC	/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8/†	6.331
ZZTo4L_POW	/ZZTo4L_13TeV_powheg_pythia8_ext1/†	1.325
ZZTo4L_POW	/ZZTo4L_13TeV_powheg_pythia8/†	1.325
ZZTo2L2Nu_POW	/ZZTo2L2Nu_13TeV_powheg_pythia8_ext1/†	0.5644
ZZTo2L2Q_AMC	/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8/†	3.688
GluGluToContinToZZTo4mu_MCFM	/GluGluToContinToZZTo4mu_13TeV_MCFM701_pythia8/†	0.002703
GluGluToContinToZZTo4e_MCFM	/GluGluToContinToZZTo4e_13TeV_MCFM701_pythia8/†	0.002703
GluGluToContinToZZTo4tau_MCFM	/GluGluToContinToZZTo4tau_13TeV_MCFM701_pythia8/†	0.002703
GluGluToContinToZZTo2e2tau_MCFM	/GluGluToContinToZZTo2e2tau_13TeV_MCFM701_pythia8/†	0.005423
GluGluToContinToZZTo2e2mu_MCFM	/GluGluToContinToZZTo2e2mu_13TeV_MCFM701_pythia8/†	0.005423
GluGluToContinToZZTo2mu2tau_MCFM	/GluGluToContinToZZTo2mu2tau_13TeV_MCFM701_pythia8/†	0.005423
WWTo2L2Nu_POW	/WWTo2L2Nu_13TeV-powheg/†	11.08
WGToLNuG_01J_AMC	/WGToLNuG_01J_5f_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/†	489.0
WJetsToLNu_AMC	/WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/†	61526.7
WLLJJ_WToLNu_EWK	/WLLJJ_WToLNu_EWK_TuneCUETP8M1_13TeV_madgraph-madspin-pythia8/†	0.0176
WpWpJJ_EWK-QCD_MG	/WpWpJJ_EWK-QCD_TuneCUETP8M1_13TeV-madgraph-pythia8/†	0.04932
WmWmJJ_POW	/WmWmJJ_13TeV-powheg-pythia8_TuneCUETP8M1/†	0.0079
ZZJJTo4L_EWK_MG	/ZZJJTo4L_EWK_13TeV-madgraph-pythia8/†	0.0004534
ZZTo4L_DPS_PYT	/ZZTo4L_DoubleScattering_13TeV-pythia8/†	0.00970
WWTo2L2Nu_DPS_PYT	/WWTo2L2Nu_DoubleScattering_13TeV-pythia8/†	0.1703
WWW_DiLeptonFilter_AMC	/WWW_4F_DiLeptonFilter_TuneCUETP8M1_13TeV-amcatnlo-pythia8/†	0.00728
WWZ_AMC	/WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/†	0.1651
WZZ_AMC	/WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/†	0.05565
ZZZ_AMC	/ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/†	0.01398
ST_tW_top_NoFullyHadronicDecays_POW	/ST_tW_top_5f_NoFullyHadronicDecays_13TeV_PSweights-powheg-pythia8/†	19.48
ST_tW_top_NoFullyHadronicDecays_POW	/ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1/†	19.48
ST_tW_antitop_NoFullyHadronicDecays_POW	/ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV_PSweights-powheg-pythia8/†	19.51
ST_tW_antitop_NoFullyHadronicDecays_POW	/ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1/†	19.51
ST_t-channel_top_POW	/ST_t-channel_top_4f_inclusiveDecays_13TeV_PSweights-powhegV2-madspin/†	119.7
ST_t-channel_top_POW	/ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1/†	119.7
ST_t-channel_antitop_POW	/ST_t-channel_antitop_4f_inclusiveDecays_13TeV_PSweights-powhegV2-madspin/†	71.74
ST_t-channel_antitop_POW	/ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1/†	71.74
ST_s-channel_leptonDecays_AMC	/ST_s-channel_4f_leptonDecays_13TeV_PSweights-amcatnlo-pythia8/†	3.74
ST_s-channel_leptonDecays_AMC	/ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1/†	3.74
ST_tWll_MG	/ST_tWll_5f_LO_13TeV-MadGraph-pythia8/†	0.01103
TTZToLL_M-1to10_AMC	/TTZToLL_M-1to10_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/†	0.05324
TTZToLLNuNu_M-10_AMC	/TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8/†	0.2728
TTWJetsToLNu_AMC	/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/†	0.2149
tZq_ll_MG	/tZq_ll_4f_PSweights_13TeV-amcatnlo-pythia8/†	0.0942
tZq_ll_MG	/tZq_ll_4f_13TeV-amcatnlo-pythia8/†	0.0942
THW_MG	/THW_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/†	0.1503
THQ_MG	/THQ_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/†	0.3189
TTHH_MG	/TTHH_TuneCUETP8M2T4_13TeV-madgraph-pythia8/†	0.0006655
TTZZ_MG	/TTZZ_TuneCUETP8M2T4_13TeV-madgraph-pythia8/†	0.001386
TTWH_MG	/TTWH_TuneCUETP8M2T4_13TeV-madgraph-pythia8/†	0.001141
TTZH_MG	/TTZH_TuneCUETP8M2T4_13TeV-madgraph-pythia8/†	0.001129
TTWZ_MG	/TTWZ_TuneCUETP8M2T4_13TeV-madgraph-pythia8/†	0.002448
TTWW_MG	/TTWW_TuneCUETP8M2T4_13TeV-madgraph-pythia8/†	0.006995
TTTJ_MG	/TTTJ_TuneCUETP8M2T4_13TeV-madgraph-pythia8/†	0.0003974
TTTW_MG	/TTTW_TuneCUETP8M2T4_13TeV-madgraph-pythia8/†	0.0007342
TTTT_MG	/TTTT_TuneCUETP8M2T4_PSweights_13TeV-amcatnlo-pythia8/†	0.008213
TTTT_MG	/TTTT_TuneCUETP8M2T4_13TeV-amcatnlo-pythia8/†	0.008213
GluGluHToZZTo4L_JHU	/GluGluHToZZTo4L_M125_13TeV_powheg2_JHUGenV6_pythia8/†	0.01297
VBF_HToZZTo4L_JHU	/VBF_HToZZTo4L_M125_13TeV_powheg2_JHUGenV709_pythia8/†	0.001010
ggZH_HToTauTau_ZToLL_POW	/ggZH_HToTauTau_ZToLL_M125_13TeV_powheg_pythia8/†	0.0007768
GluGluZH_HToWW_ZTo2L_POW	/GluGluZH_HToWW_ZTo2L_M125_13TeV_powheg_pythia8/†	0.002648
GluGluToZH_HToZZTo4L_JHU	/GluGluToZH_HToZZTo4L_M125_13TeV_JHUGenV723_pythia8/†	0.00003276
bbH_HToZZTo4L_JHU	/bbH_HToZZTo4L_M125_13TeV_JHUgenV702_pythia8/†	0.0001303
VHToNonbb_DiLeptonFilter_AMC	/VHToNonbb_M125_DiLeptonFilter_TuneCUETP8M1_13TeV_amcatnloFXFX_madspin_pythia8/†	0.09177
ttHToNonbb_POW	/ttHToNonbb_M125_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8/†	0.2118
	† RunIISummer16MiniAODv3-PUMoriond17_94X_mcRun2_asymptotic_v3*/MINIAODSIM	

Table A.3: The list of MC samples and corresponding cross-sections in 2017.

Process name	DAS name	Cross-section (pb)
DYJetsToLL_M-10to50_MG	/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8/†	18610.0
DYJetsToLL_M-50_AMC	/DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8/†	5765.0
ZGToLLG_01J_lowMLL_lowPT_AMC	/ZGToLLG_01J_5f_lowMLL_lowGPt_TuneCP5_13TeV-amcatnloFXFX-pythia8/†	172.8
TTTo2L2Nu_POW	/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/†	88.29
TTTo2L2Nu_POW	/TTTo2L2Nu_TuneCP5_PSweights_13TeV-powheg-pythia8/†	88.29
TTToSemiLeptonic_POW	/TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8/†	365.34
TTToSemiLeptonic_POW	/TTToSemiLeptonic_TuneCP5_PSweights_13TeV-powheg-pythia8/†	365.34
WZTo3LNu_AMC	/WZTo3LNu_TuneCP5_13TeV-amcatnloFXFX-pythia8/†	5.052
WZ2L2Q_AMC	/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8/†	6.331
ZZTo4L_POW	/ZZTo4L_13TeV_powheg_pythia8/†	1.325
ZZTo2L2Nu_POW	/ZZTo2L2Nu_13TeV_powheg_pythia8/†	0.5644
ZZTo2L2Q_AMC	/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8/†	3.688
GluGluToContinToZZTo4mu_MCFM	/GluGluToContinToZZTo4mu_13TeV_TuneCP5_MCFM701_pythia8/†	0.002703
GluGluToContinToZZTo4e_MCFM	/GluGluToContinToZZTo4e_13TeV_TuneCP5_MCFM701_pythia8/†	0.002703
GluGluToContinToZZTo4tau_MCFM	/GluGluToContinToZZTo4tau_13TeV_TuneCP5_MCFM701_pythia8/†	0.002703
GluGluToContinToZZTo2e2tau_MCFM	/GluGluToContinToZZTo2e2tau_13TeV_TuneCP5_MCFM701_pythia8/†	0.005423
GluGluToContinToZZTo2e2mu_MCFM	/GluGluToContinToZZTo2e2mu_13TeV_TuneCP5_MCFM701_pythia8/†	0.005423
GluGluToContinToZZTo2mu2tau_MCFM	/GluGluToContinToZZTo2mu2tau_13TeV_TuneCP5_MCFM701_pythia8/†	0.005423
WWTo2L2Nu_POW	/WWTo2L2Nu_NNPDF31_TuneCP5_13TeV-powheg-pythia8/†	11.08
WWTo2L2Nu_POW	/WWTo2L2Nu_NNPDF31_TuneCP5_PSweights_13TeV-powheg-pythia8/†	11.08
WGToLNuG_01J_AMC	/WGToLNuG_01J_5f_TuneCP5_13TeV-amcatnloFXFX-pythia8/†	489.0
WJetsToLNu_MG	/WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8/†	61526.7
WLLJJ_WToLNu_EWK	/WLLJJ_WToLNu_EWK_TuneCP5_13TeV_madgraph-madspin-pythia8/†	0.0176
WpWpJJ_EWK-QCD_MG	/WpWpJJ_EWK-QCD_TuneCP5_13TeV-madgraph-pythia8/†	0.04932
WmWmJJ_EWK_POW	/WmWmJJ_EWK_TuneCP5_13TeV-powheg-pythia8/†	0.00703
ZZJJTo4L_EWK_MG	/ZZJJTo4L_EWK_TuneCP5_13TeV-madgraph-pythia8/†	0.0004534
ZZTo4L_DPS_PYT	/ZZTo4L_TuneCP5_DoubleScattering_13TeV-pythia8/†	0.00970
WWTo2L2Nu_DPS_PYT	/WWTo2L2Nu_DoubleScattering_13TeV-pythia8/†	0.1703
WWZTo4L2Nu_AMC	/WWZTo4L2Nu_4f_TuneCP5_13TeV_amcatnlo_pythia8/†	0.0020670
WWZTo3L1Nu2Q_AMC	/WWZTo3L1Nu2Q_4f_TuneCP5_13TeV_amcatnlo_pythia8/†	0.0080390
WZZTo3L1Nu2Q_AMC	/WZZTo3L1Nu2Q_4f_TuneCP5_13TeV_amcatnlo_pythia8/†	0.0027190
WZZTo4LX_AMC	/WZZ_ZTo2L_WToAll_4f_TuneCP5_13TeV_amcatnlo_pythia8/†	0.0006299
ZZZTo4LX_AMC	/ZZZJetsTo4L2Nu_4f_TuneCP5_13TeV_amcatnloFXFX_pythia8/†	0.0001907
ST_tW_top_NoFullyHadronicDecays_POW	/ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8/†	19.48
ST_tW_top_NoFullyHadronicDecays_POW	/ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_PSweights_13TeV-powheg-pythia8/†	19.48
ST_tW_antitop_NoFullyHadronicDecays_POW	/ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8/†	19.51
ST_tW_antitop_NoFullyHadronicDecays_POW	/ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_PSweights_13TeV-powheg-pythia8/†	19.51
ST_t-channel_top_POW	/ST_t-channel_top_5f_TuneCP5_13TeV-powheg-pythia8/†	119.7
ST_t-channel_antitop_POW	/ST_t-channel_antitop_5f_TuneCP5_PSweights_13TeV-powheg-pythia8/†	71.74
ST_s-channel_leptonDecays_AMC	/ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8/†	3.74
ST_s-channel_leptonDecays_AMC	/ST_s-channel_4f_leptonDecays_TuneCP5_PSweights_13TeV-amcatnlo-pythia8/†	3.74
ST_tWll_MG	/ST_tWll_5f_LO_TuneCP5_PSweights_13TeV-madgraph-pythia8/†	0.01103
TTZToLL_M-1to10_AMC	/TTZToLL_M-1to10_TuneCP5_13TeV-amcatnlo-pythia8/†	0.05324
TTZToLLNuNu_M-10_AMC	/TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8/†	0.2728
TTZToLLNuNu_M-10_AMC	/TTZToLLNuNu_M-10_TuneCP5_PSweights_13TeV-amcatnlo-pythia8/†	0.2728
TTWJetsToLNu_AMC	/TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/†	0.2149
TTWJetsToLNu_AMC	/TTWJetsToLNu_TuneCP5_PSweights_13TeV-amcatnloFXFX-madspin-pythia8/†	0.2149
tZq_ll_MG	/tZq_ll_4f_ckm_NLO_TuneCP5_PSweights_13TeV-amcatnlo-pythia8/†	0.0942
THW_MG	/THW_5f_Hincl_13TeV_madgraph_pythia8/†	0.1503
THQ_MG	/THQ_4f_Hincl_13TeV_madgraph_pythia8/†	0.3189
TTHH_MG	/TTHH_TuneCP5_13TeV-madgraph-pythia8/†	0.0006655
TTZZ_MG	/TTZZ_TuneCP5_13TeV-madgraph-pythia8/†	0.001386
TTWH_MG	/TTWH_TuneCP5_13TeV-madgraph-pythia8/†	0.001141
TTZH_MG	/TTZH_TuneCP5_13TeV-madgraph-pythia8/†	0.001129
TTWZ_MG	/TTWZ_TuneCP5_13TeV-madgraph-pythia8/†	0.002448
TTWW_MG	/TTWW_TuneCP5_13TeV-madgraph-pythia8/†	0.006995
TTTJ_MG	/TTTJ_TuneCP5_13TeV-madgraph-pythia8/†	0.0003974
TTTW_MG	/TTTW_TuneCP5_13TeV-madgraph-pythia8/†	0.0007342
TTTT_MG	/TTTT_TuneCP5_PSweights_13TeV-amcatnlo-pythia8/†	0.008213
GluGluHToZZTo4L_JHU	/GluGluHToZZTo4L_M125_13TeV_powheg2_JHUGenV7011_pythia8/†	0.01297
VBF_HToZZTo4L_JHU	/VBF_HToZZTo4L_M125_13TeV_powheg2_JHUGenV7011_pythia8/†	0.001010
ggZH_HToTauTau_ZToLL_POW	/ggZH_HToTauTau_ZToLL_M125_13TeV_powheg_pythia8/†	0.0007768
GluGluZH_HToWW_ZTo2L_POW	/GluGluZH_HToWW_ZTo2L_M125_13TeV_powheg_pythia8_TuneCP5/†	0.002648
GluGluToZH_HToZZTo4L_JHU	/GluGluToZH_HToZZTo4L_M125_13TeV_JHUGenV723_pythia8/†	0.00003276
bbH_HToZZTo4L_JHU	/bbH_HToZZTo4L_M125_13TeV_JHUGenV7011_pythia8/†	0.0001303
VHToNonbb_DiLeptonFilter_AMC	/VHToNonbb_M125_DiLeptonFilter_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8/†	0.09177
ttHToNonbb_POW	/ttHToNonbb_M125_TuneCP5_13TeV-powheg-pythia8/†	0.2118

† RunIIFall17MiniAODv2-PU2017_12Apr2018_94X_mc2017_realistic_v14*/MINIAODSIM

Table A.4: The list of MC samples and corresponding cross-sections in 2018.

Process name	DAS name	Cross-section (pb)
DYJetsToLL_M-10to50_MG	/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8/†	18610.0
DYJetsToLL_M-50_AMC	/DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8/†	5765.0
ZGToLLG_01J_lowMLL_lowPT_AMC	/ZGToLLG_01J_5f_lowMLL_lowGPt_TuneCP5_13TeV-amcatnloFXFX-pythia8/†	172.8
TTTo2L2Nu_POW	/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/†	88.29
TTToSemiLeptonic_POW	/TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8/†	365.34
WZTo3LNu_AMC	/WZTo3LNu_TuneCP5_13TeV-amcatnloFXFX-pythia8/†	5.052
WZ2L2Q_AMC	/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8/†	6.331
ZZTo4L_POW	/ZZTo4L_TuneCP5_13TeV_powheg_pythia8/†	1.325
ZZTo4L_POW	/ZZTo4L_13TeV_powheg_pythia8_TuneCP5/†	1.325
ZZTo2L2Nu_POW	/ZZTo2L2Nu_TuneCP5_13TeV_powheg_pythia8/†	0.5644
ZZTo2L2Q_AMC	/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8/†	3.688
GluGluToContinToZZTo4mu_MCFM	/GluGluToContinToZZTo4mu_13TeV_TuneCP5_MCFM701_pythia8/†	0.002703
GluGluToContinToZZTo4e_MCFM	/GluGluToContinToZZTo4e_13TeV_TuneCP5_MCFM701_pythia8/†	0.002703
GluGluToContinToZZTo4tau_MCFM	/GluGluToContinToZZTo4tau_13TeV_TuneCP5_MCFM701_pythia8/†	0.002703
GluGluToContinToZZTo2e2tau_MCFM	/GluGluToContinToZZTo2e2tau_13TeV_TuneCP5_MCFM701_pythia8/†	0.005423
GluGluToContinToZZTo2e2mu_MCFM	/GluGluToContinToZZTo2e2mu_13TeV_TuneCP5_MCFM701_pythia8/†	0.005423
GluGluToContinToZZTo2mu2tau_MCFM	/GluGluToContinToZZTo2mu2tau_13TeV_TuneCP5_MCFM701_pythia8/†	0.005423
WWTo2L2Nu_POW	/WWTo2L2Nu_NNPDF31_TuneCP5_13TeV-powheg-pythia8/†	11.08
WGToLNuG_01J_AMC	/WGToLNuG_01J_5f_TuneCP5_13TeV-amcatnloFXFX-pythia8/†	489.0
WJetsToLNu_MG	/WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8/†	61526.7
WLLJJ_WToLNu_EWK	/WLLJJ_WToLNu_EWK_TuneCP5_13TeV_madgraph-madspin-pythia8/†	0.0176
WpWpJJ_EWK-QCD_MG	/WpWpJJ_EWK-QCD_TuneCP5_13TeV-madgraph-pythia8/†	0.04932
WmWmJJ_EWK_POW	/WmWmJJ_EWK_TuneCP5_13TeV-powheg-pythia8/†	0.00703
ZZJJTo4L_EWK_MG	/ZZJJTo4L_EWK_TuneCP5_13TeV-madgraph-pythia8/†	0.0004534
ZZTo4L_DPS_PYT	/ZZTo4L_TuneCP5_DoubleScattering_13TeV-pythia8/†	0.00970
WWTo2L2Nu_DPS_PYT	/WWTo2L2Nu_DoubleScattering_13TeV-pythia8/†	0.1703
WWZTo4L2Nu_AMC	/WWZTo4L2Nu_4f_TuneCP5_13TeV_amcatnlo_pythia8/†	0.0020670
WWZTo3L1Nu2Q_AMC	/WWZTo3L1Nu2Q_4f_TuneCP5_13TeV_amcatnlo_pythia8/†	0.0080390
WZZTo3L1Nu2Q_AMC	/WZZTo3L1Nu2Q_4f_TuneCP5_13TeV_amcatnlo_pythia8/†	0.0027190
WZZTo4LX_AMC	/WZZ_ZTo2L_WToAll_4f_TuneCP5_13TeV_amcatnlo_pythia8/†	0.0006299
ZZZTo4LX_AMC	/ZZZJetsTo4L2Nu_4f_TuneCP5_13TeV_amcatnloFXFX_pythia8/†	0.0001907
ST_tW_top_NoFullyHadronicDecays_POW	/ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8/†	19.48
ST_tW_antitop_NoFullyHadronicDecays_POW	/ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8/†	19.51
ST_t-channel_top_POW	/ST_t-channel_top_5f_TuneCP5_13TeV-powheg-pythia8/†	119.7
ST_t-channel_antitop_POW	/ST_t-channel_antitop_5f_TuneCP5_13TeV-powheg-pythia8/†	71.74
ST_s-channel_leptonDecays_MG	/ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-madgraph-pythia8/†	3.74
ST_tWll_MG	/ST_tWll_5f_LO_TuneCP5_PSweights_13TeV-madgraph-pythia8/†	0.01103
TTZToLL_M-1to10_AMC	/TTZToLL_M-1to10_TuneCP5_13TeV-amcatnlo-pythia8/†	0.05324
TTZToLLNuNu_M-10_AMC	/TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8/†	0.2728
TTWJetsToLNu_AMC	/TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/†	0.2149
tZq_ll_MG	/tZq_ll_4f_ckm_NLO_TuneCP5_13TeV-madgraph-pythia8/†	0.0942
THW_MG	/THW_5f_Hincl_13TeV_madgraph_pythia8/†	0.1503
THQ_MG	/THQ_4f_Hincl_13TeV_madgraph_pythia8/†	0.3189
TTHH_MG	/TTHH_TuneCP5_13TeV-madgraph-pythia8/†	0.0006655
TTZZ_MG	/TTZZ_TuneCP5_13TeV-madgraph-pythia8/†	0.001386
TTWH_MG	/TTWH_TuneCP5_13TeV-madgraph-pythia8/†	0.001141
TTZH_MG	/TTZH_TuneCP5_13TeV-madgraph-pythia8/†	0.001129
TTWZ_MG	/TTWZ_TuneCP5_13TeV-madgraph-pythia8/†	0.002448
TTWW_MG	/TTWW_TuneCP5_13TeV-madgraph-pythia8/†	0.006995
TTTJ_MG	/TTTJ_TuneCP5_13TeV-madgraph-pythia8/†	0.0003974
TTTW_MG	/TTTW_TuneCP5_13TeV-madgraph-pythia8/†	0.0007342
TTTT_MG	/TTTT_TuneCP5_13TeV-amcatnlo-pythia8/†	0.008213
GluGluHToZZTo4L_JHU	/GluGluHToZZTo4L_M125_13TeV_powheg2_JHUGenV7011_pythia8/†	0.01297
VBF_HToZZTo4L_JHU	/VBF_HToZZTo4L_M125_13TeV_powheg2_JHUGenV7011_pythia8/†	0.001010
ggZH_HToTauTau_ZToLL_POW	/ggZH_HToTauTau_ZToLL_M125_13TeV_powheg_pythia8/†	0.0007768
GluGluZH_HToWW_ZTo2L_POW	/GluGluZH_HToWW_ZTo2L_M125_13TeV_powheg_pythia8_TuneCP5_PSweights/†	0.002648
GluGluToZH_HToZZTo4L_JHU	/GluGluToZH_HToZZTo4L_M125_13TeV_JHUGenV723_pythia8/†	0.00003276
bbH_HToZZTo4L_JHU	/bbH_HToZZTo4L_M125_13TeV_JHUGenV7011_pythia8/†	0.0001303
VHToNonbb_DiLeptonFilter_AMC	/VHToNonbb_M125_DiLeptonFilter_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8/†	0.09177
ttHToNonbb_POW	/ttHToNonbb_M125_TuneCP5_13TeV-powheg-pythia8/†	0.2118

† RunIIAutumn18MiniAOD-102X_upgrade2018_realistic_v15*/MINIAODSIM

B Electron prompt and fake rates

B.1 Electron prompt rates

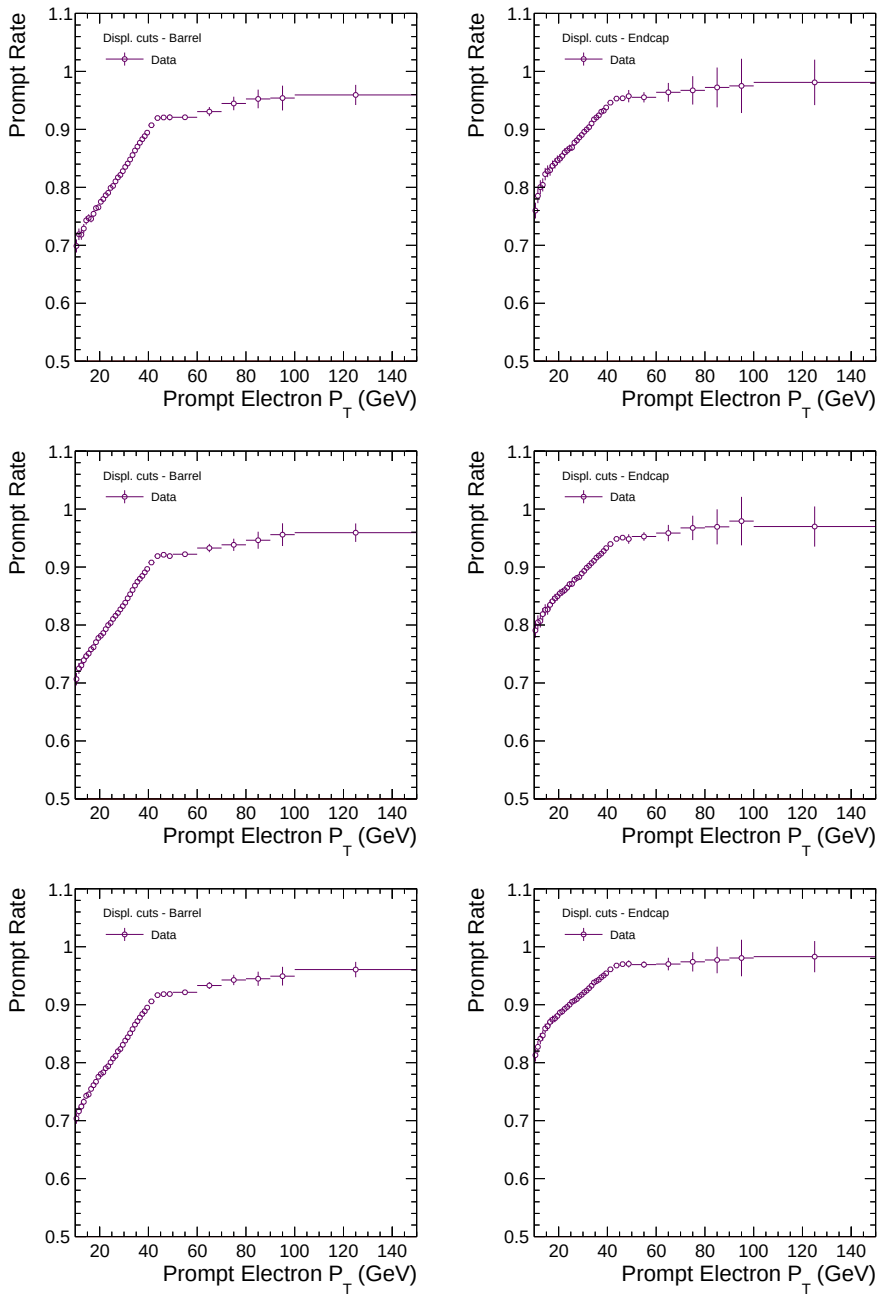

Figure B.1: Electron DY prompt rates in 2016 (upper), 2017 (middle) and 2018 (lower). The uncertainties are statistical only. Edge bins include over- and underflows.

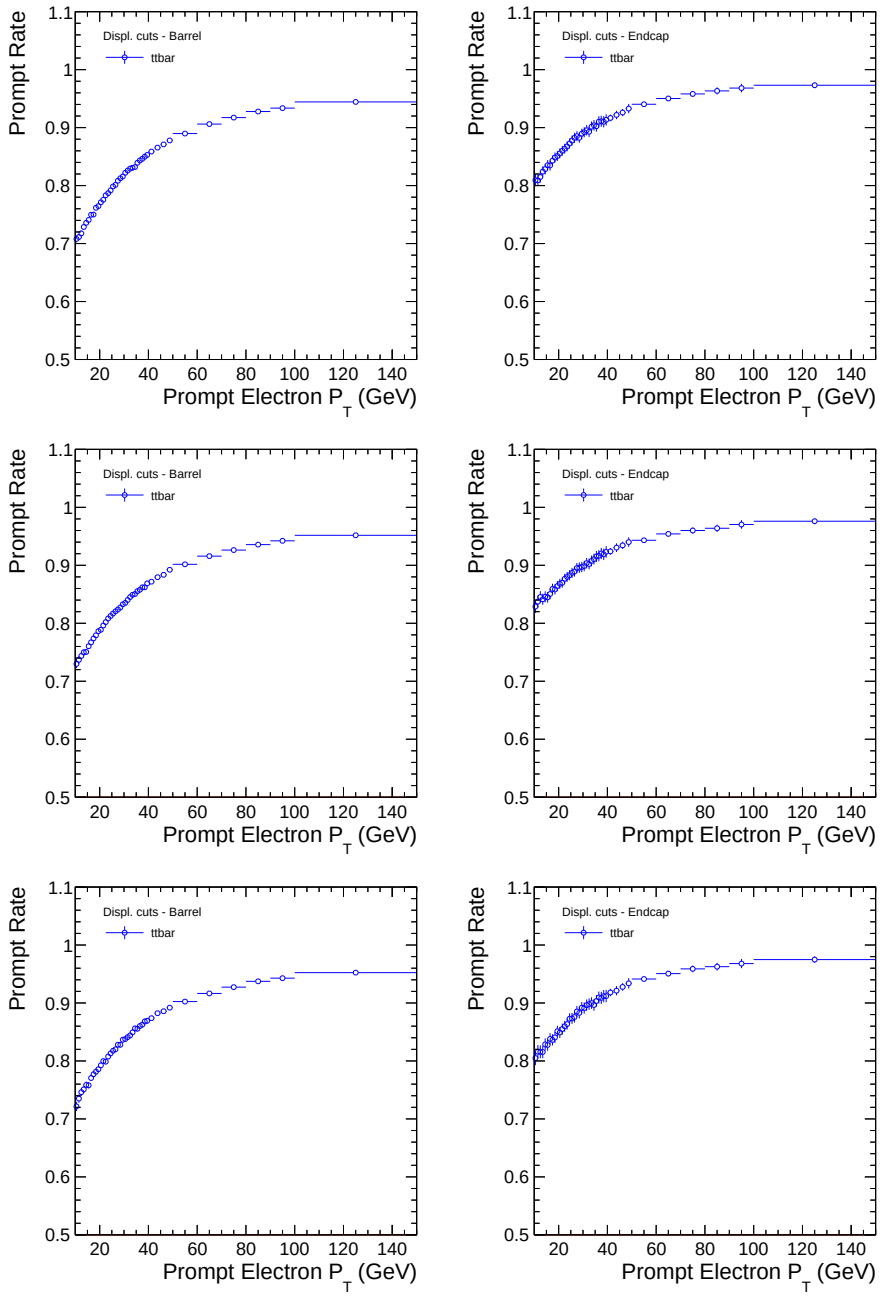

Figure B.2: Electron $t\bar{t}$ prompt rates in 2016, 2017 and 2018. The uncertainties are statistical only. Edge bins include over- and underflows.

B.2 Electron fake rates

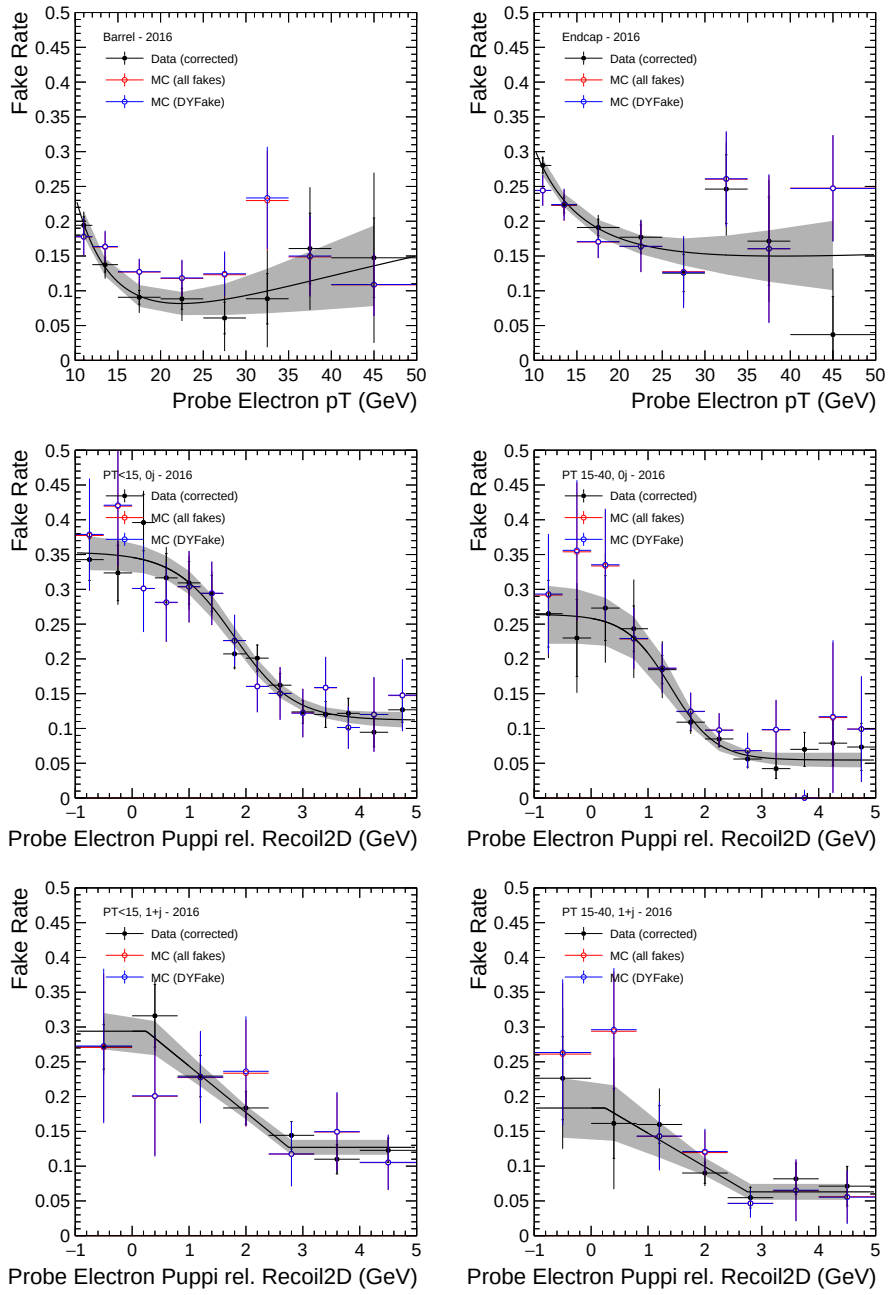

Figure B.3: Electron DY fake rates in 2016. The uncertainties are statistical only. Edge bins include over- and underflows.

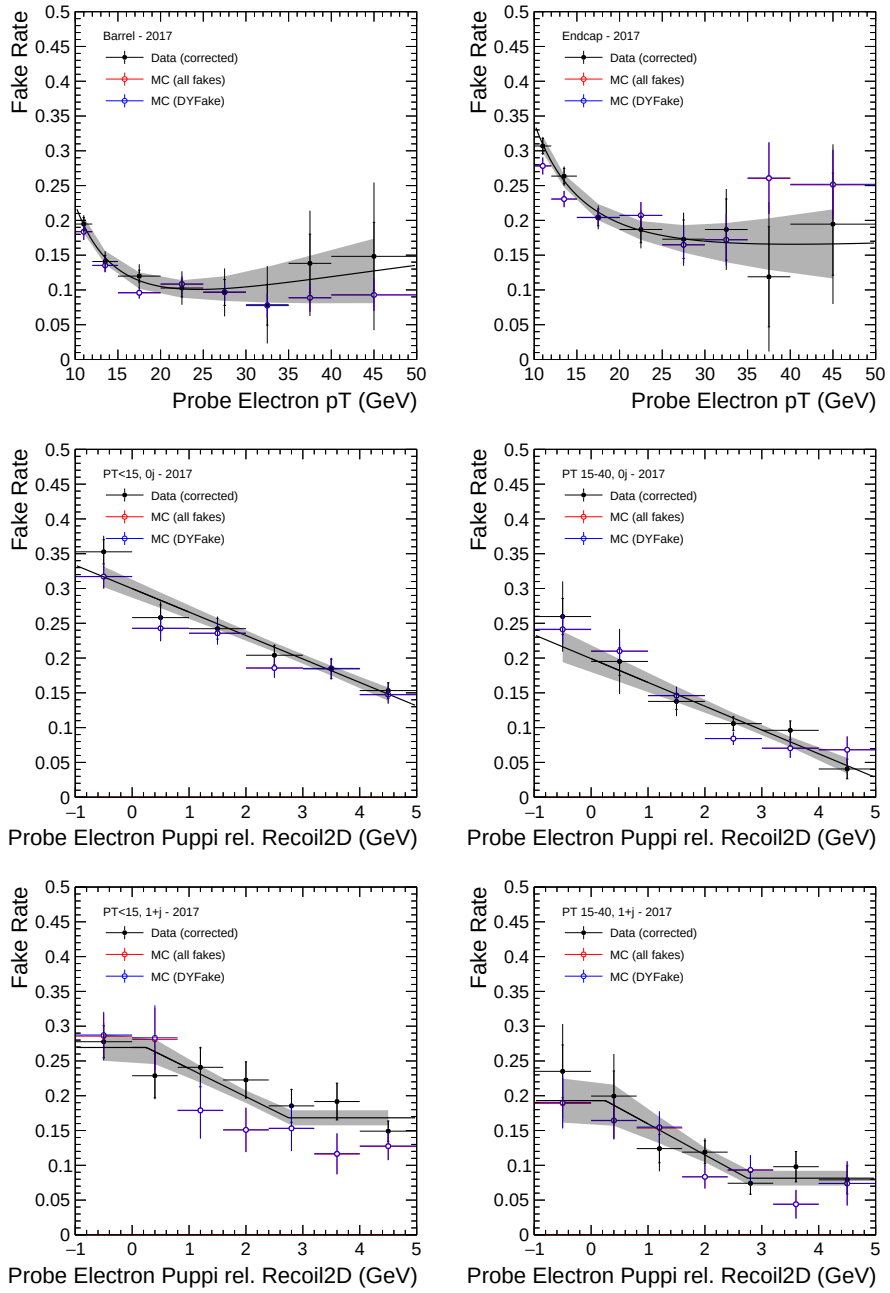

Figure B.4: Electron DY fake rates in 2017. The uncertainties are statistical only. Edge bins include over- and underflows.

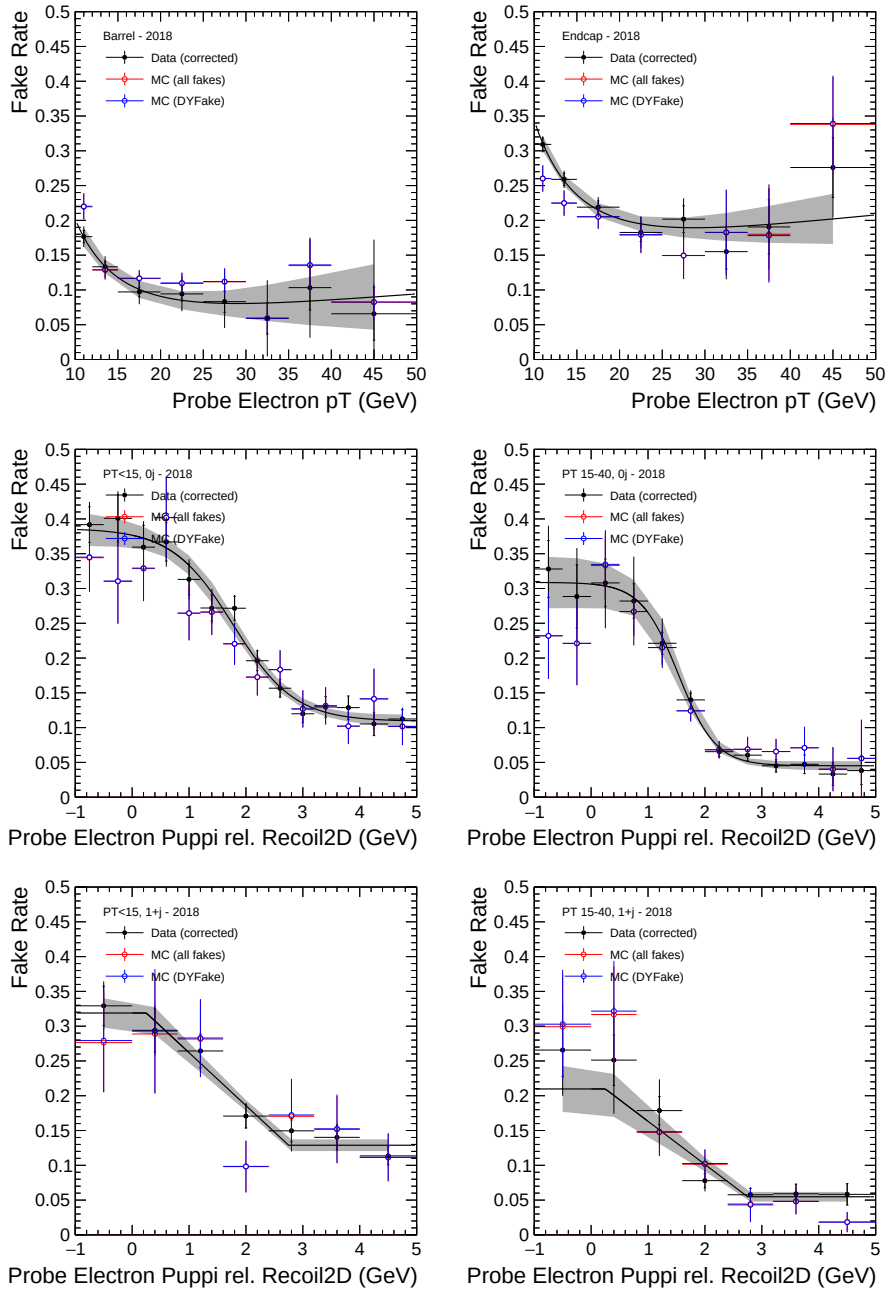

Figure B.5: Electron DY fake rates in 2018. The uncertainties are statistical only. Edge bins include over- and underflows.

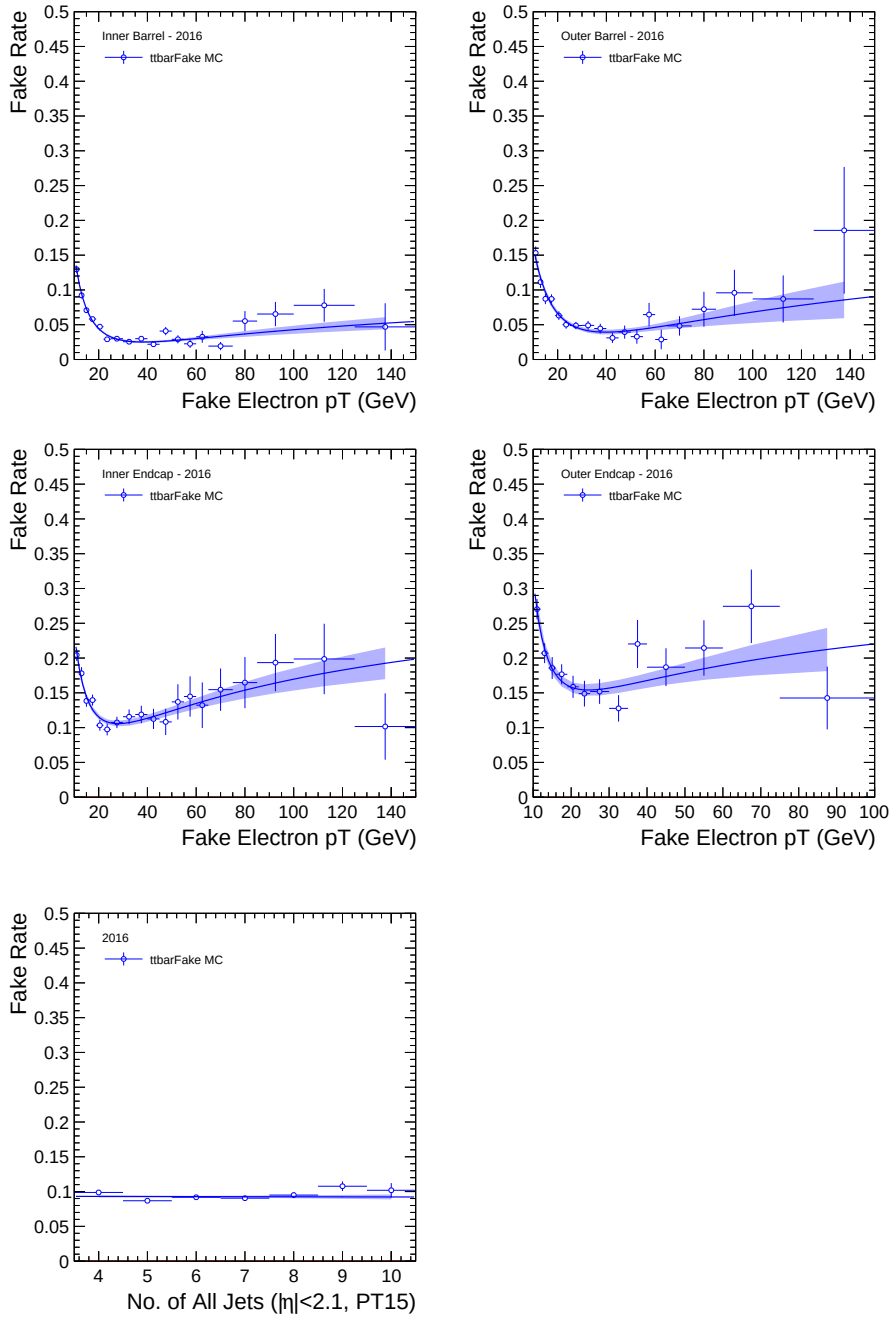

Figure B.6: Electron $t\bar{t}$ fake rates in 2016. The uncertainties are statistical only. Edge bins include over- and underflows.

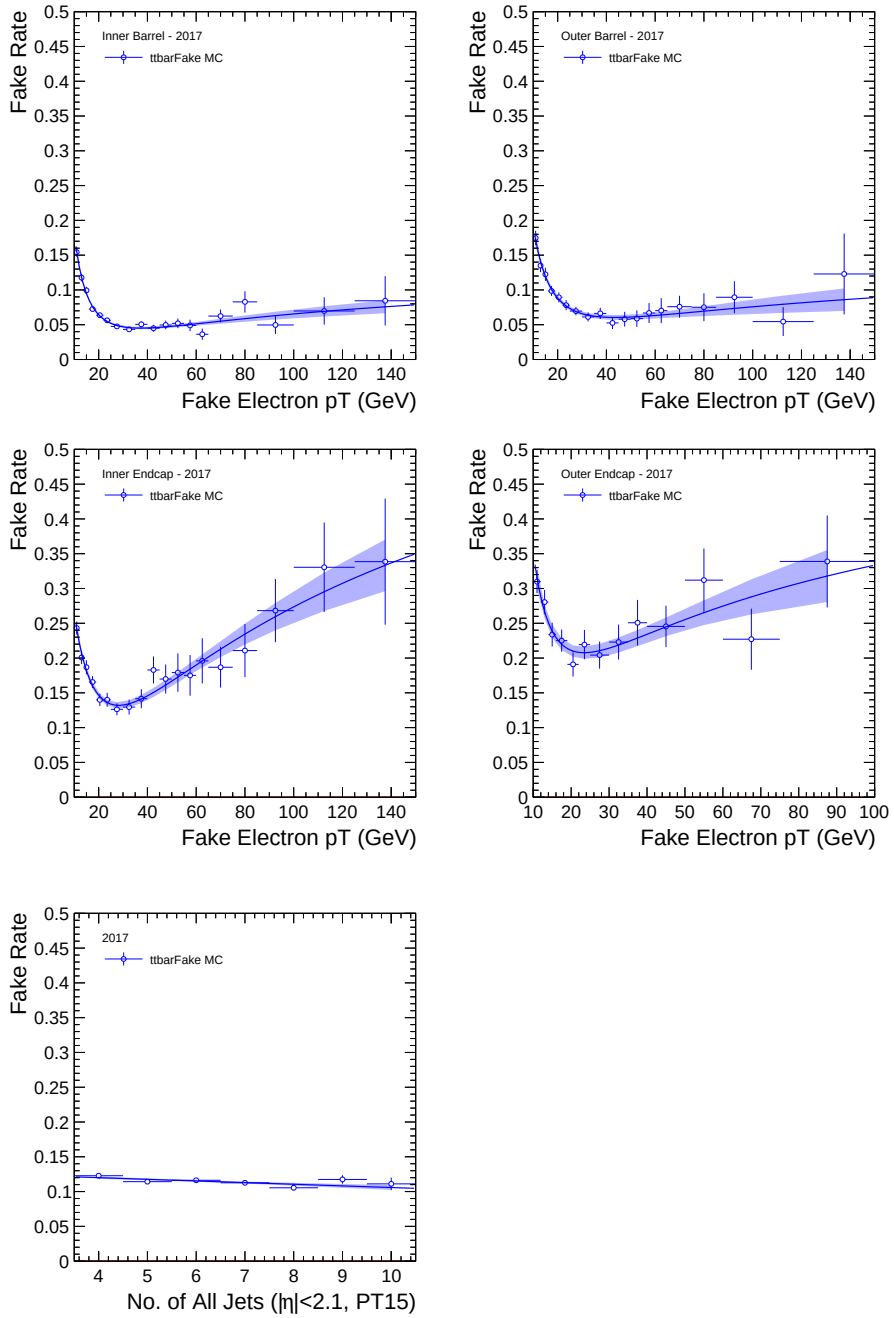

Figure B.7: Electron $t\bar{t}$ fake rates in 2017. The uncertainties are statistical only. Edge bins include over- and underflows.

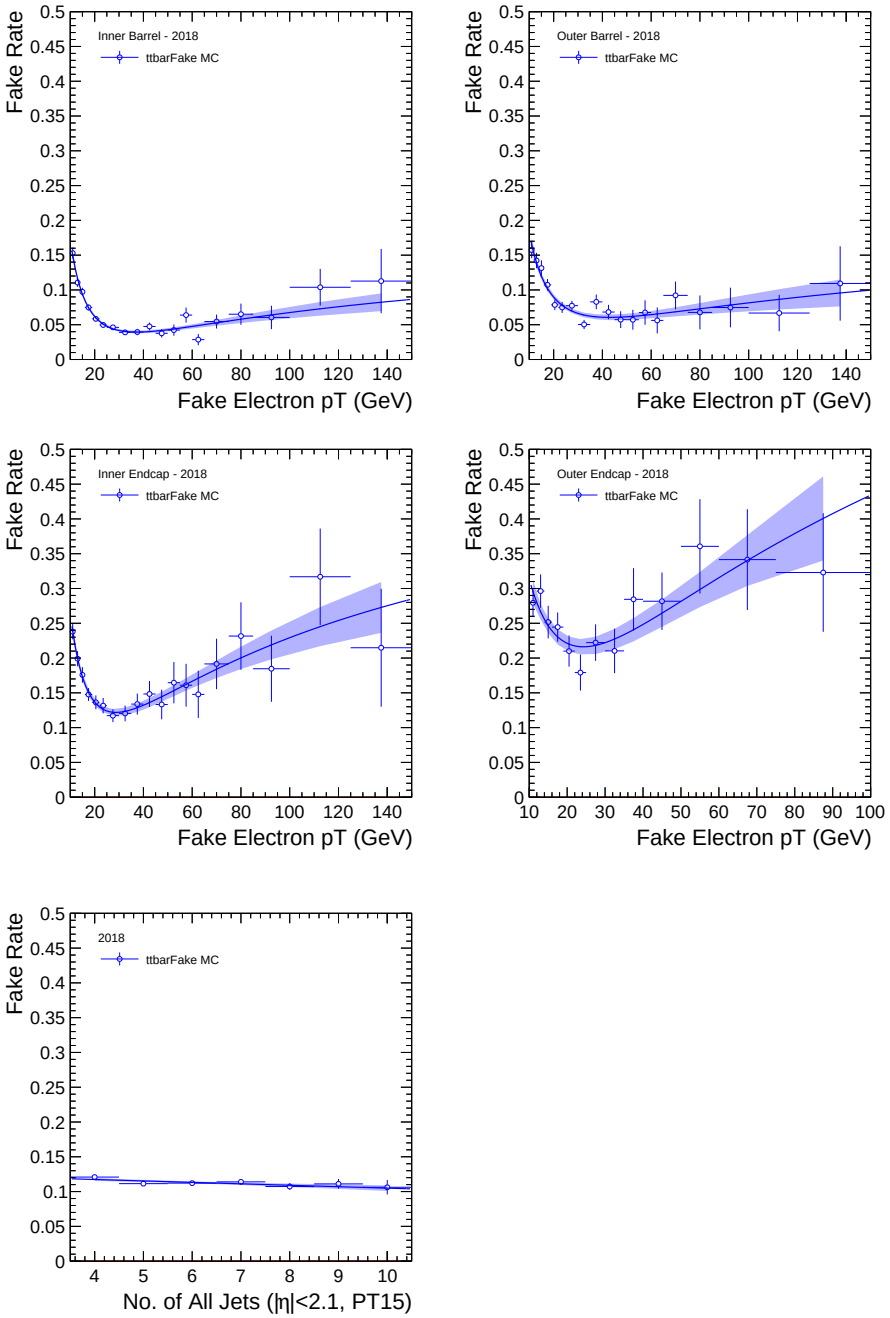

Figure B.8: Electron $t\bar{t}$ fake rates in 2018. The uncertainties are statistical only. Edge bins include over- and underflows.

C Muon prompt and fake rates

C.1 Muon prompt rates

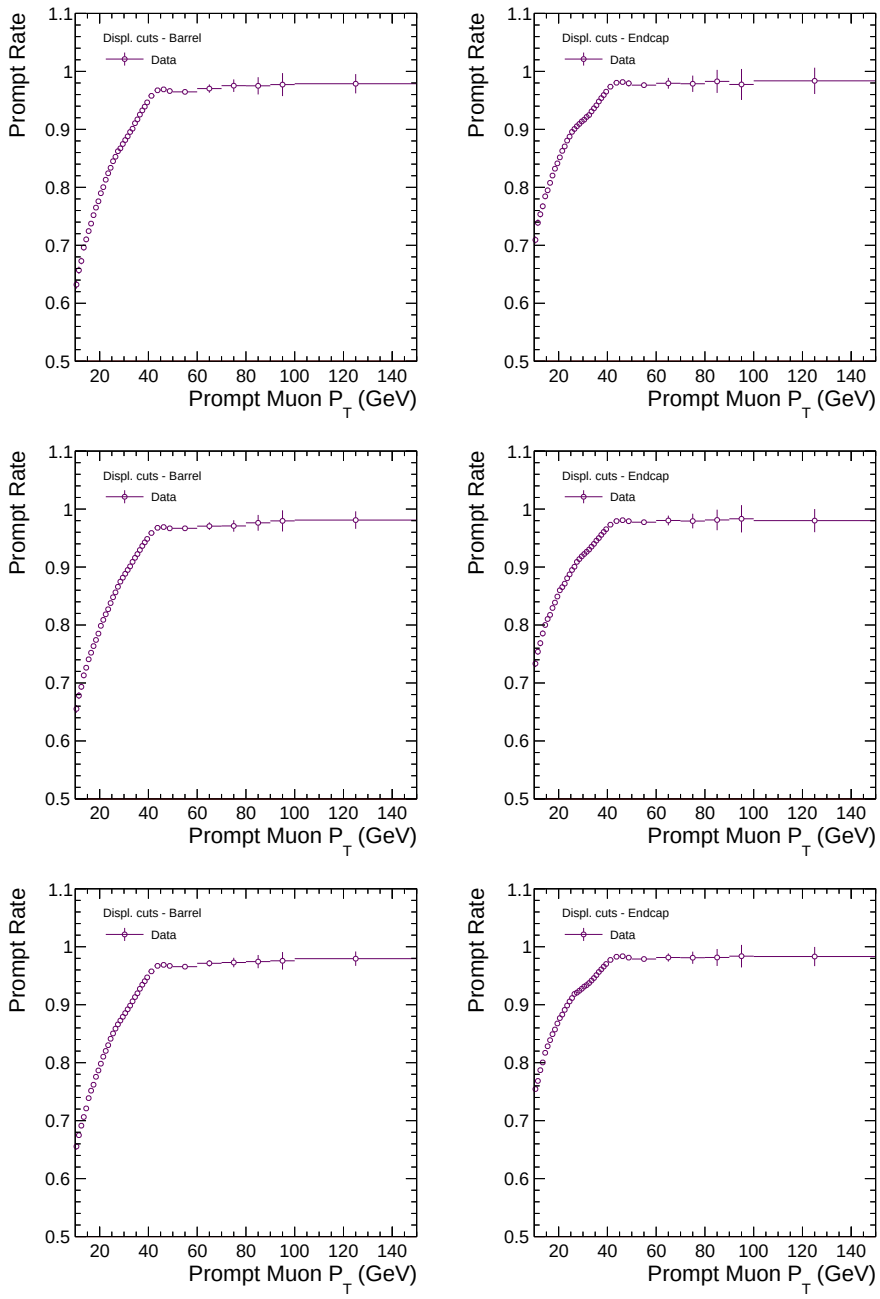

Figure C.9: Muon DY prompt rates in 2016 (upper), 2017 (middle) and 2018 (lower). The uncertainties are statistical only. Edge bins include over- and underflows.

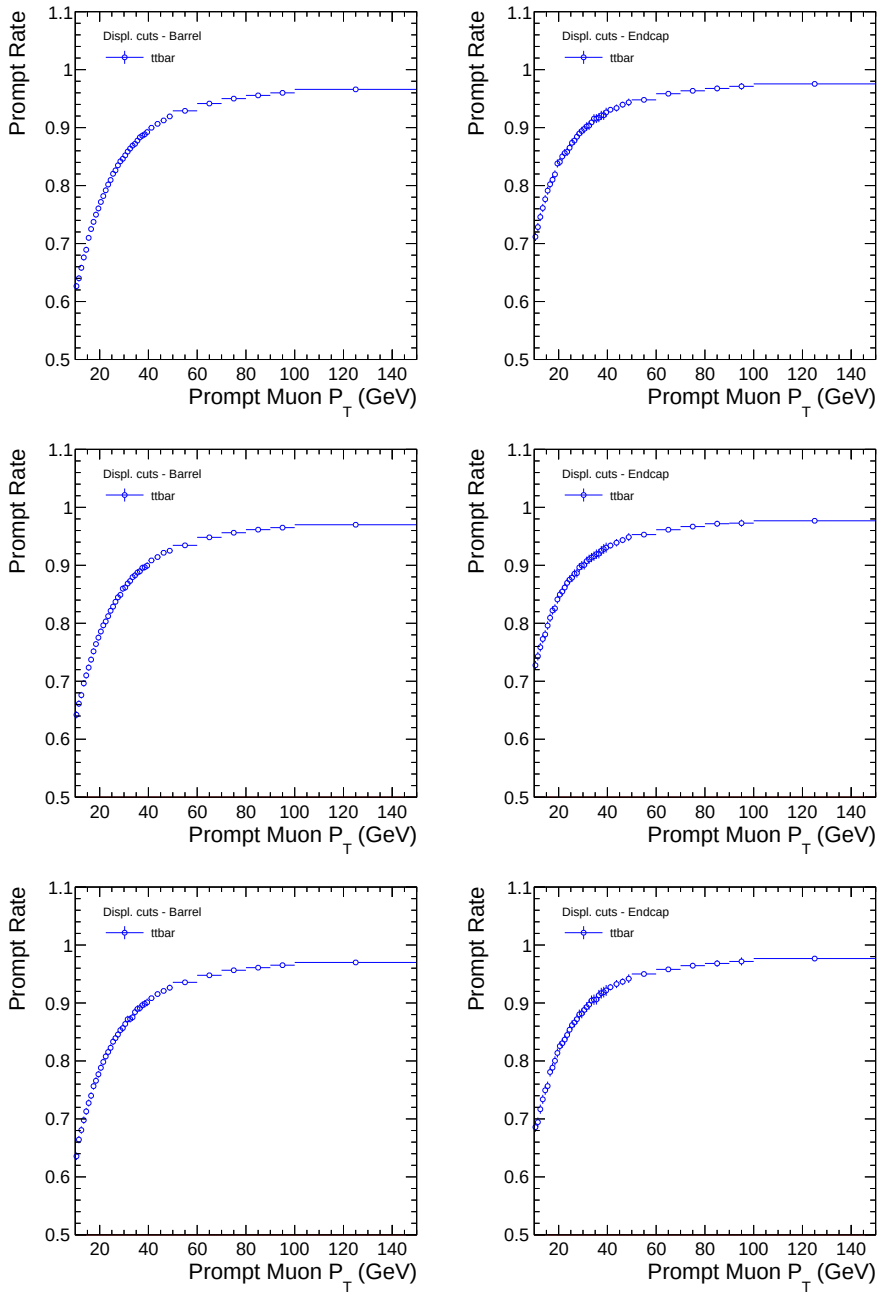

Figure C.10: Muon $t\bar{t}$ prompt rates in 2016 (upper), 2017 (middle) and 2018 (lower). The uncertainties are statistical only. Edge bins include over- and underflows.

C.2 Muon fake rates

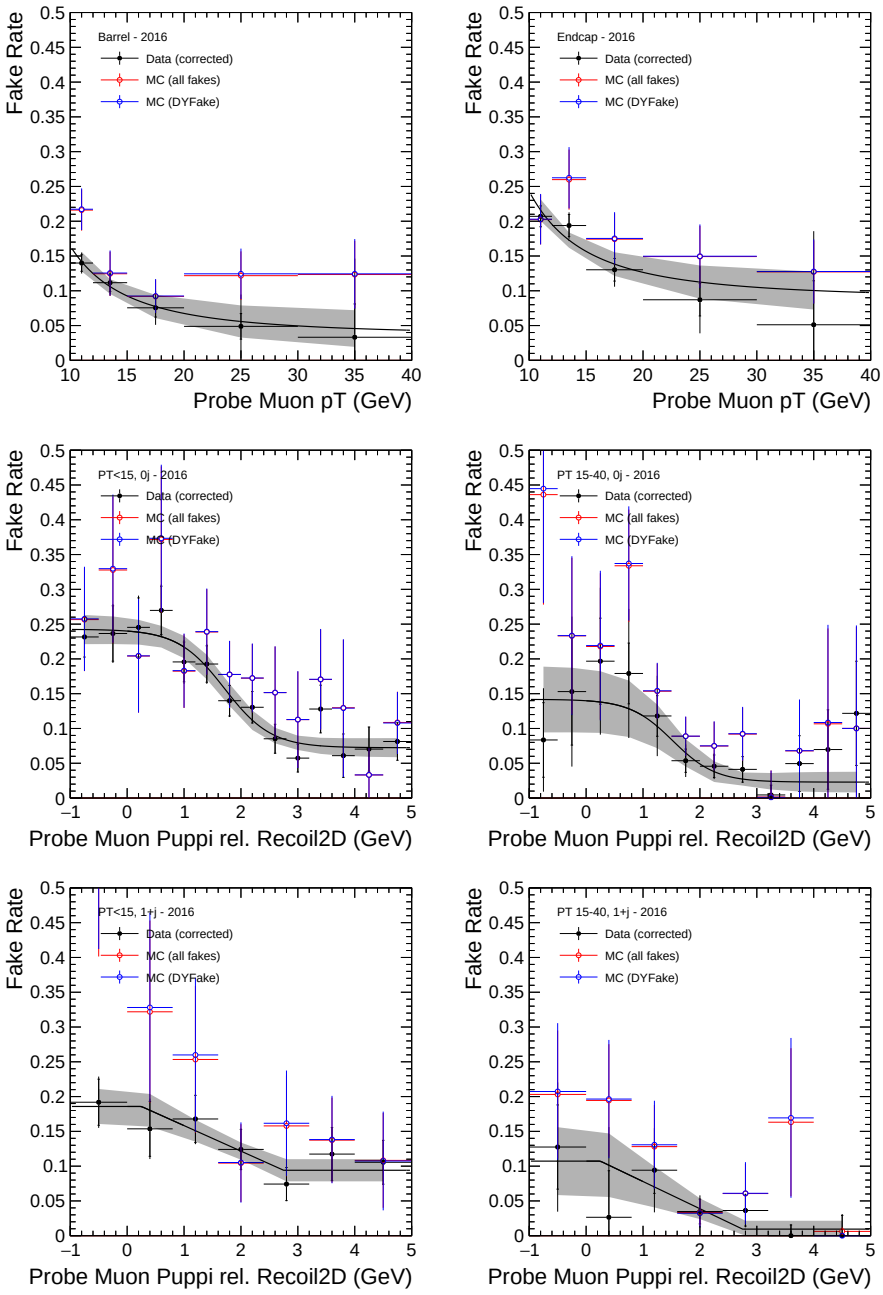

Figure C.11: Muon DY fake rates in 2016. The uncertainties are statistical only. Edge bins include over- and underflows.

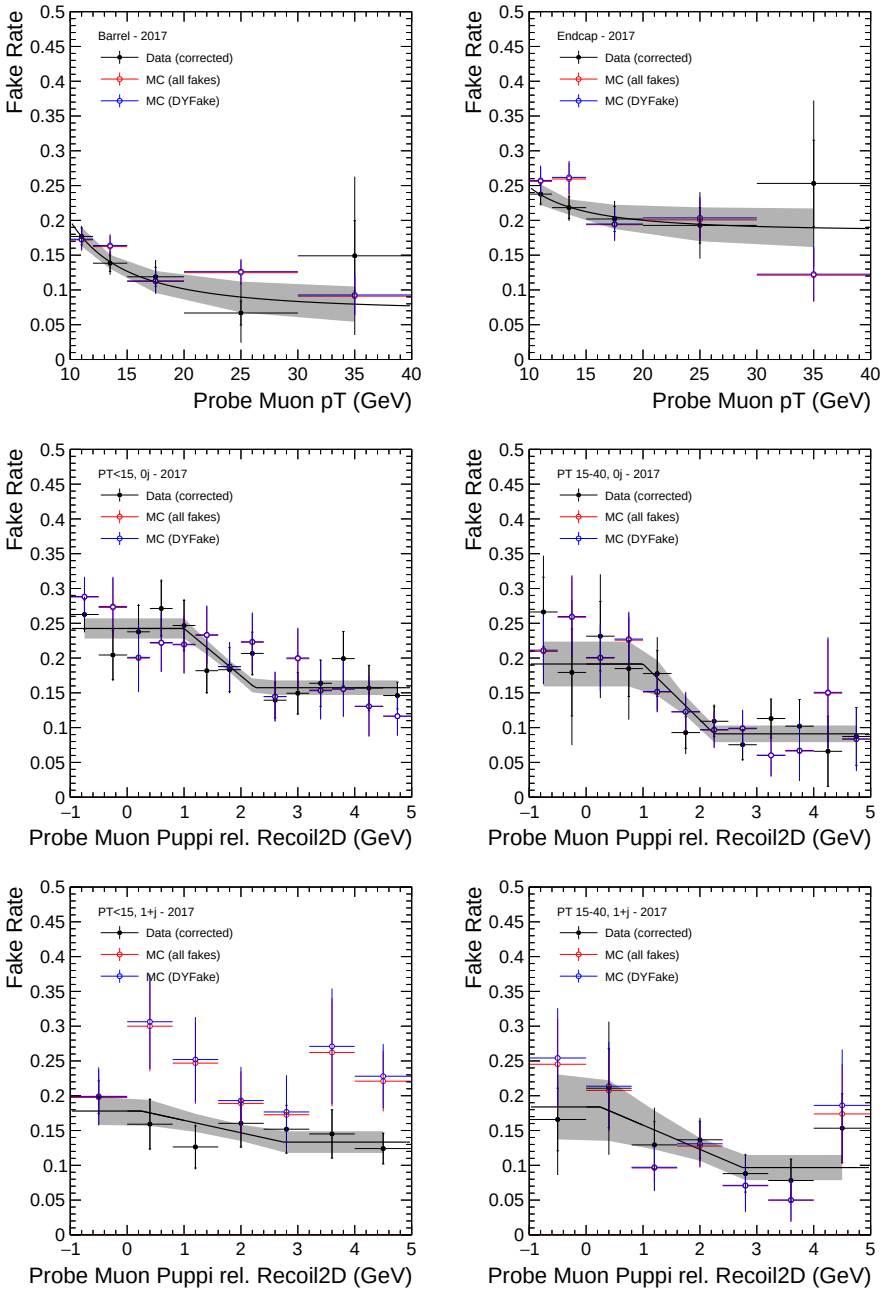

Figure C.12: Muon DY fake rates in 2017. The uncertainties are statistical only. Edge bins include over- and underflows.

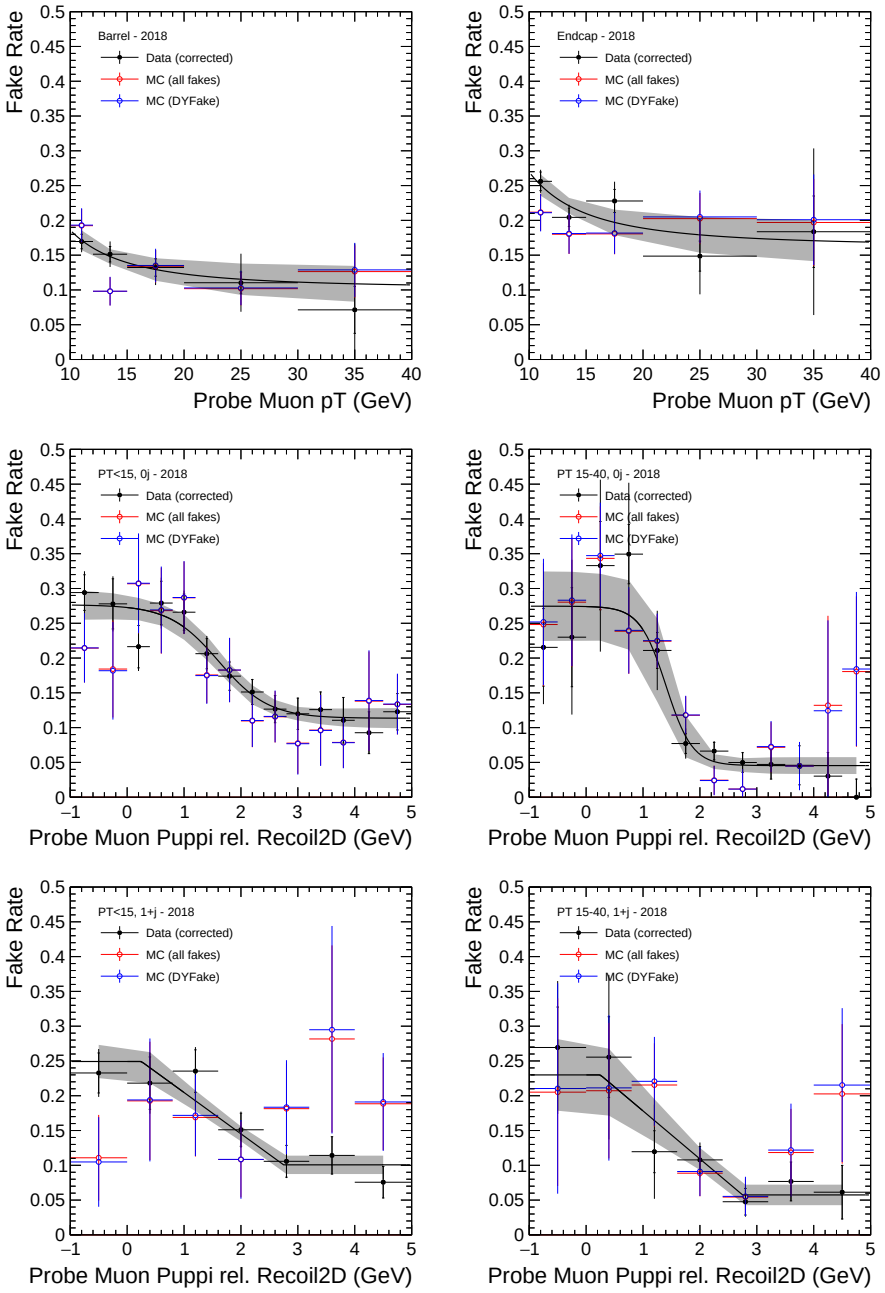

Figure C.13: Muon DY fake rates in 2018. The uncertainties are statistical only. Edge bins include over- and underflows.

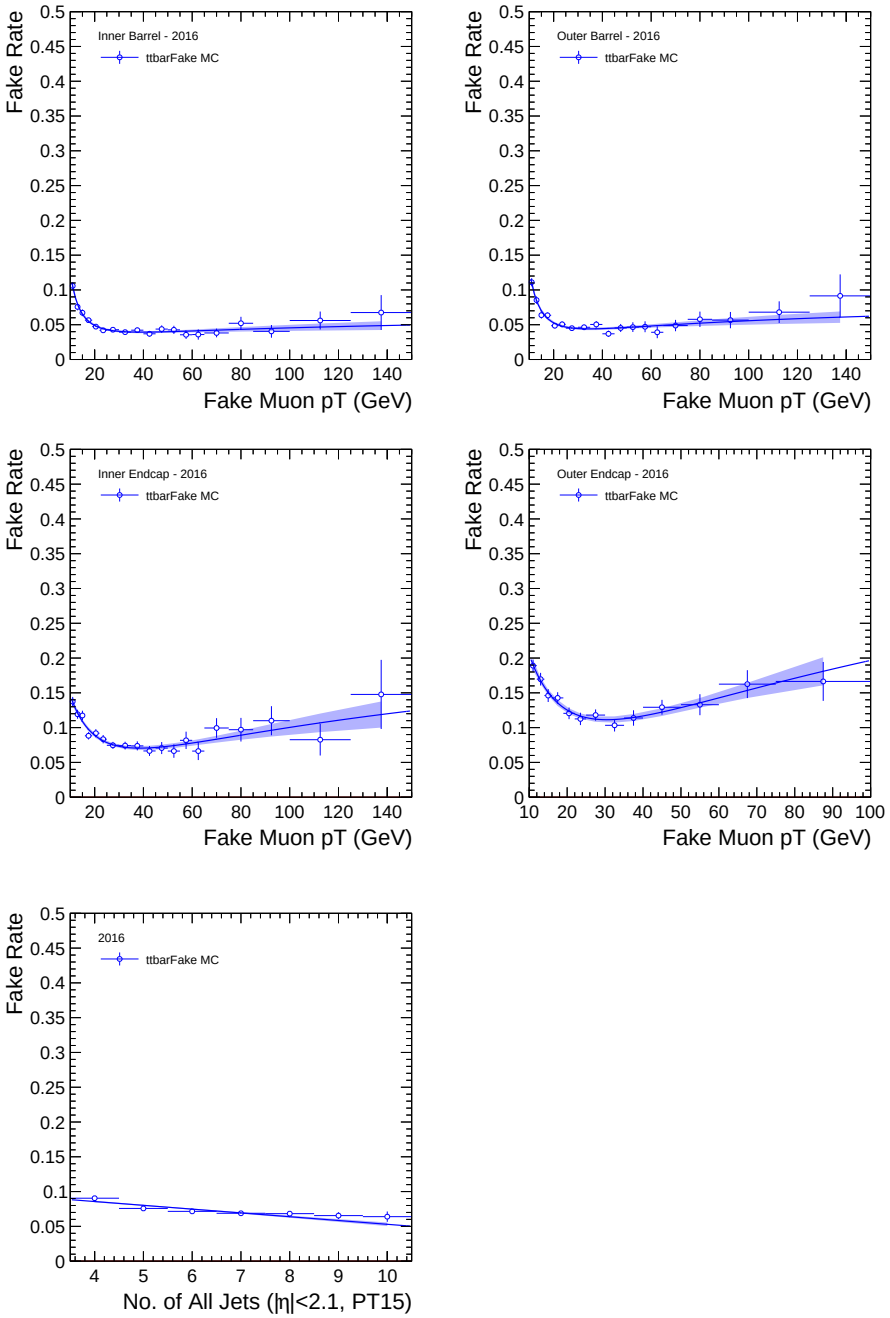

Figure C.14: Muon $t\bar{t}$ fake rates in 2016. The uncertainties are statistical only. Edge bins include over- and underflows.

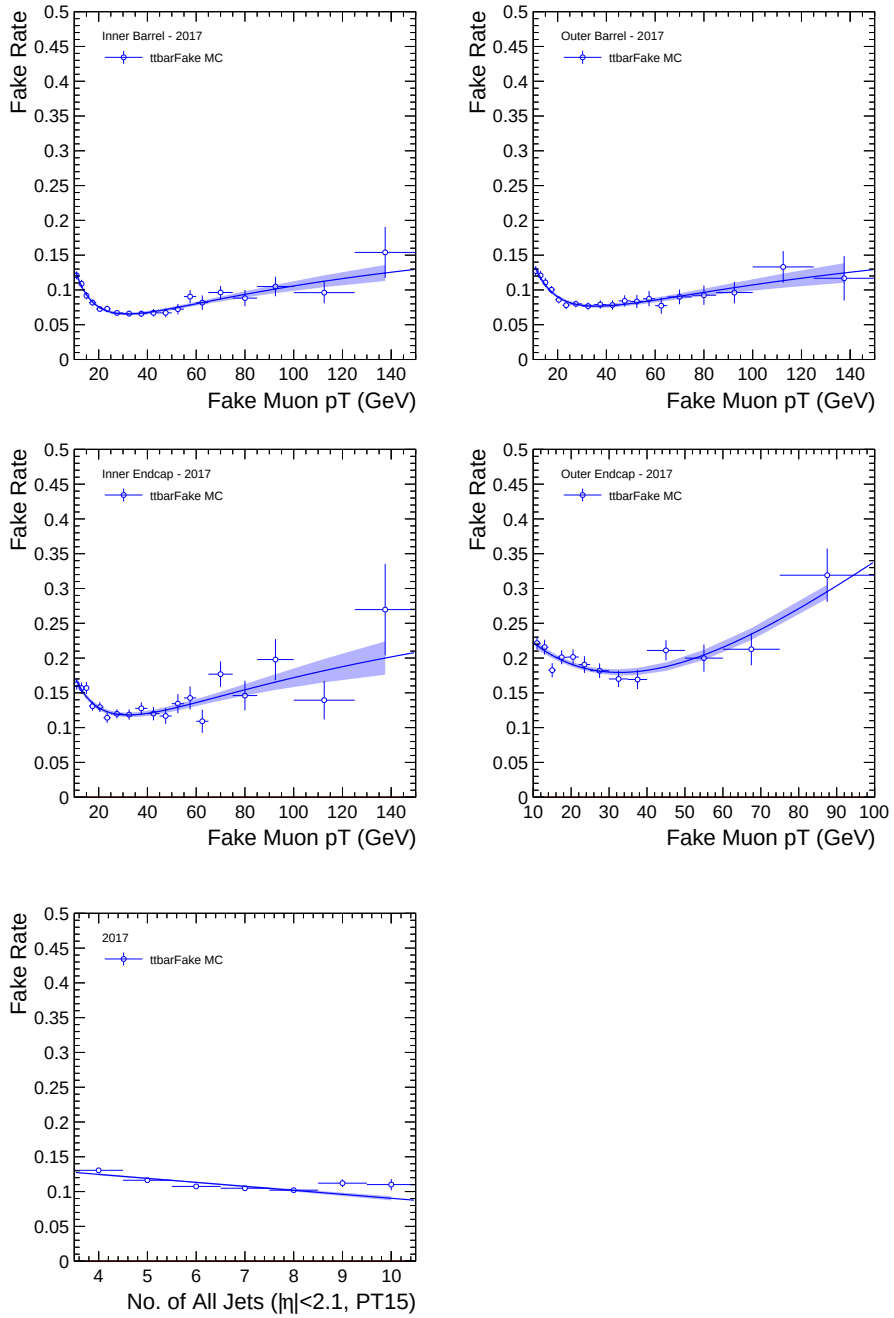

Figure C.15: Muon $t\bar{t}$ fake rates in 2017. The uncertainties are statistical only. Edge bins include over- and underflows.

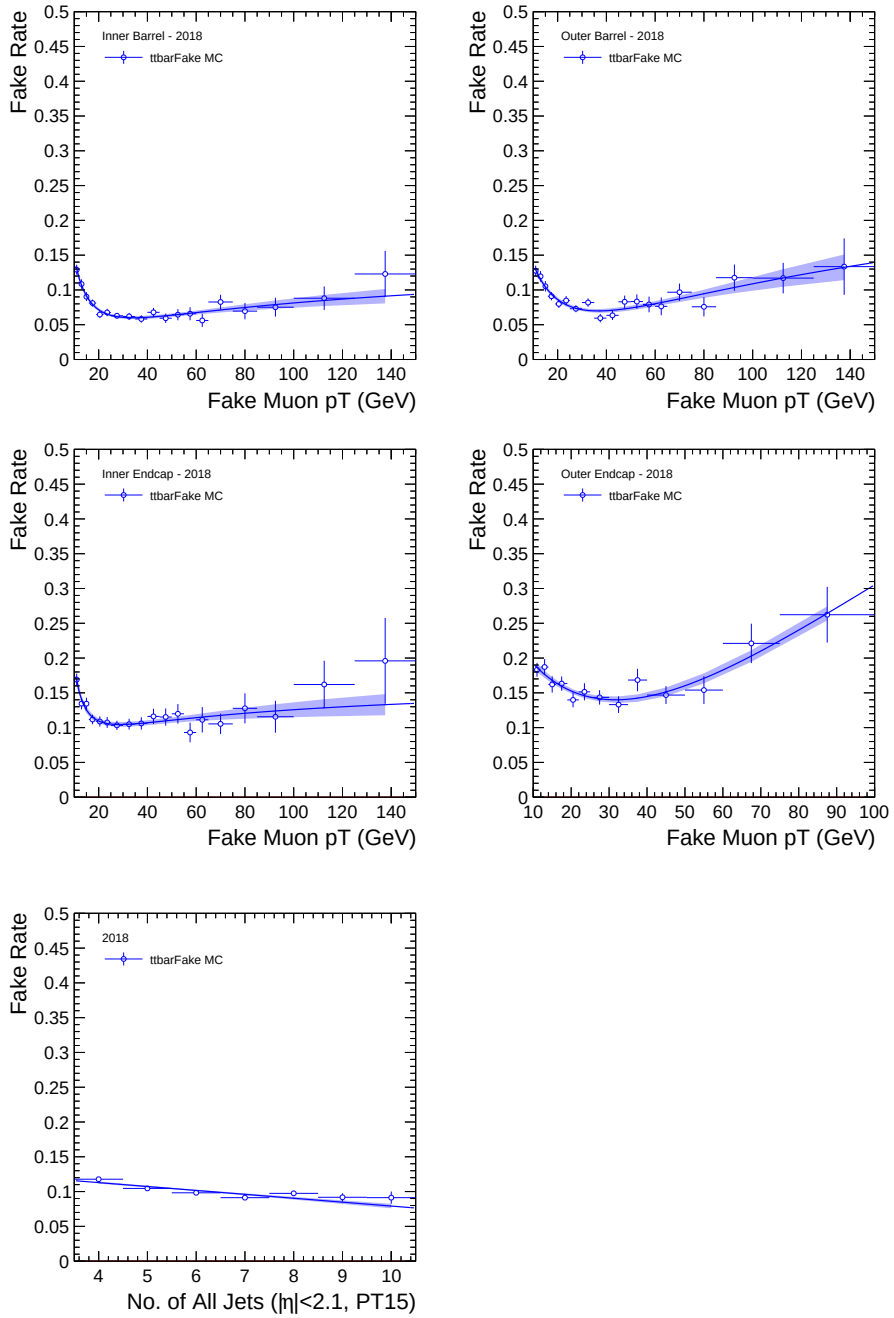

Figure C.16: Muon $t\bar{t}$ fake rates in 2018. The uncertainties are statistical only. Edge bins include over- and underflows.

D Tau prompt and fake rates

D.1 Tau prompt rates

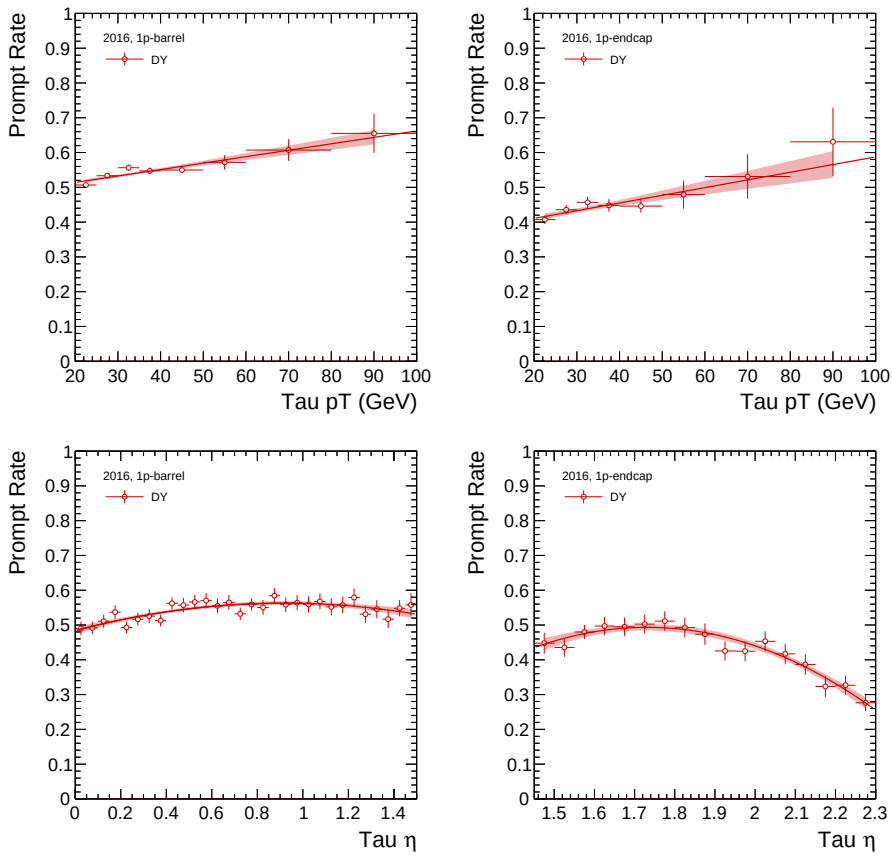

Figure D.17: Tau 1-prong DY prompt rates in 2016. The uncertainties are statistical only. Edge bins include over- and underflows.

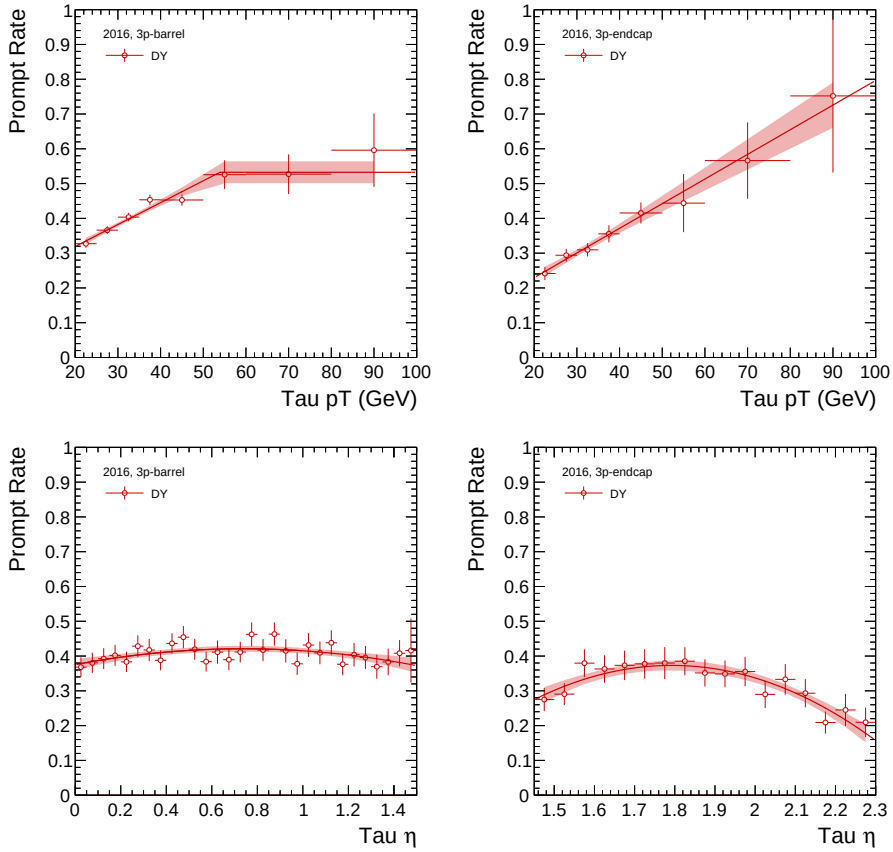

Figure D.18: Tau 3-prong DY prompt rates in 2016. The uncertainties are statistical only. Edge bins include over- and underflows.

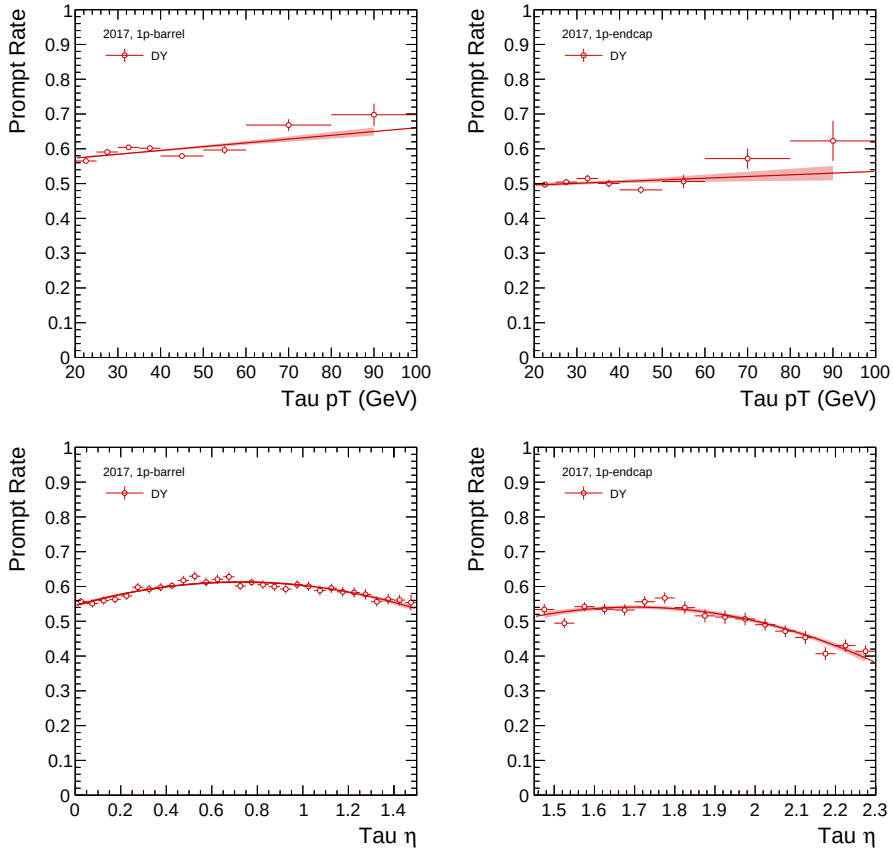

Figure D.19: Tau 1-prong DY prompt rates in 2017. The uncertainties are statistical only. Edge bins include over- and underflows.

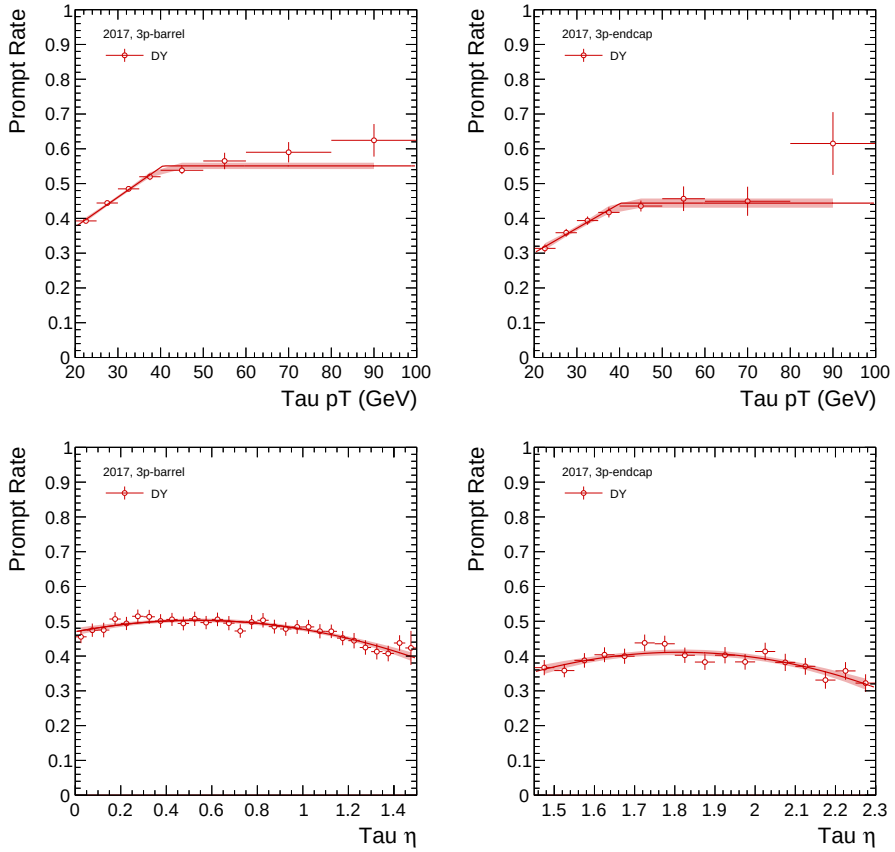

Figure D.20: Tau 3-prong DY prompt rates in 2017. The uncertainties are statistical only. Edge bins include over- and underflows.

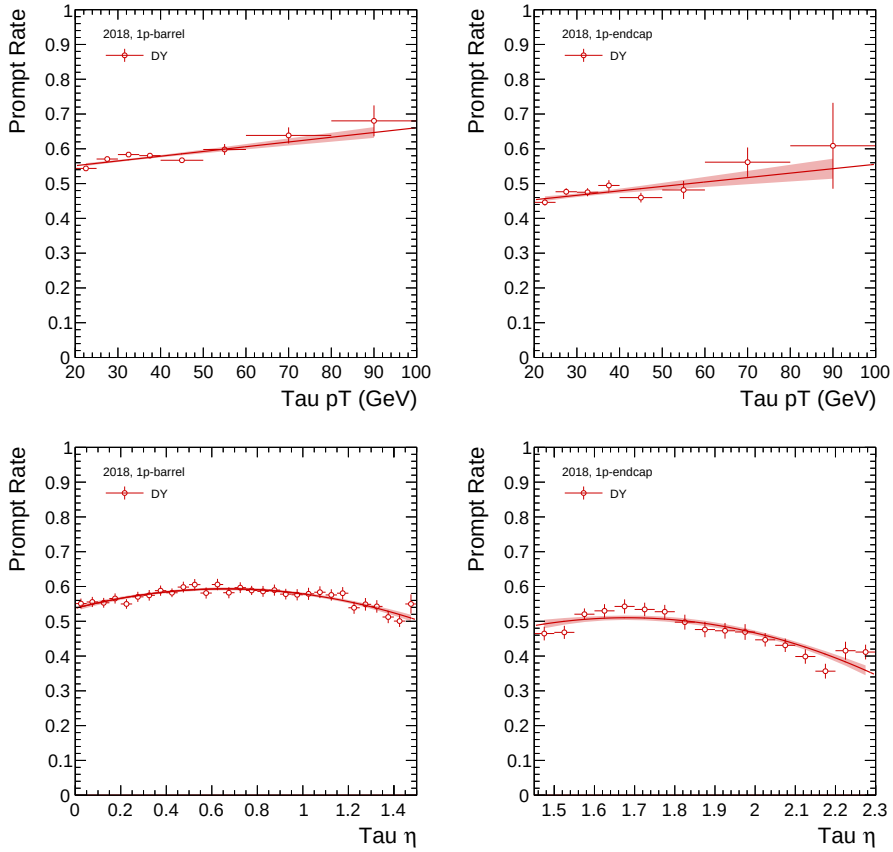

Figure D.21: Tau 1-prong DY prompt rates in 2018. The uncertainties are statistical only. Edge bins include over- and underflows.

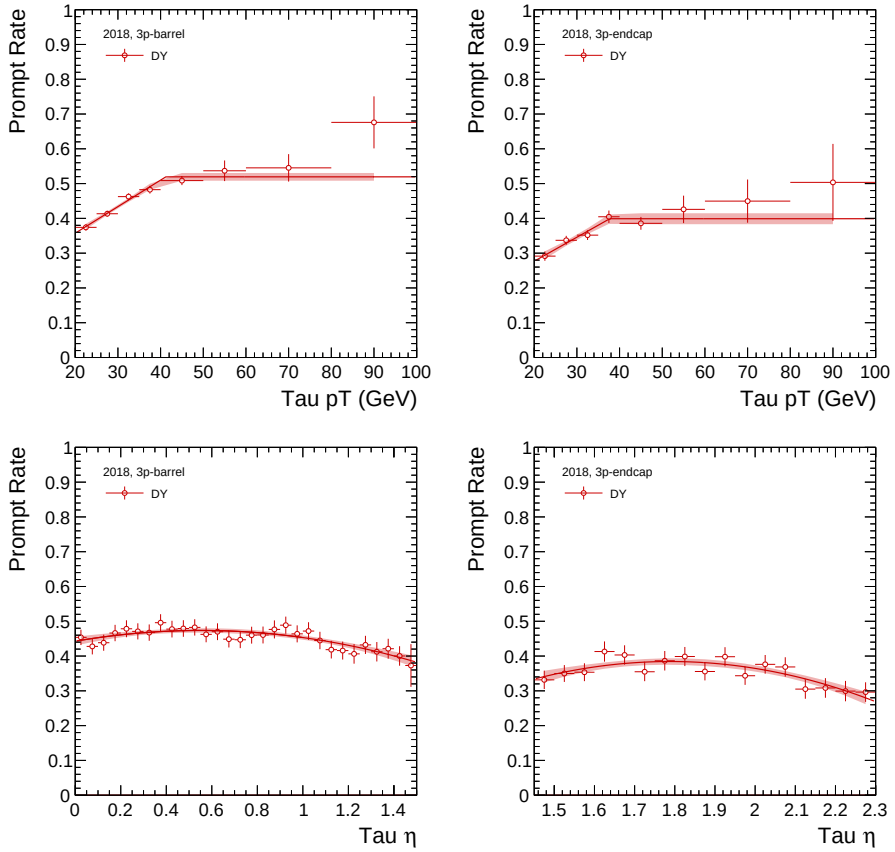

Figure D.22: Tau 3-prong DY prompt rates in 2018. The uncertainties are statistical only. Edge bins include over- and underflows.

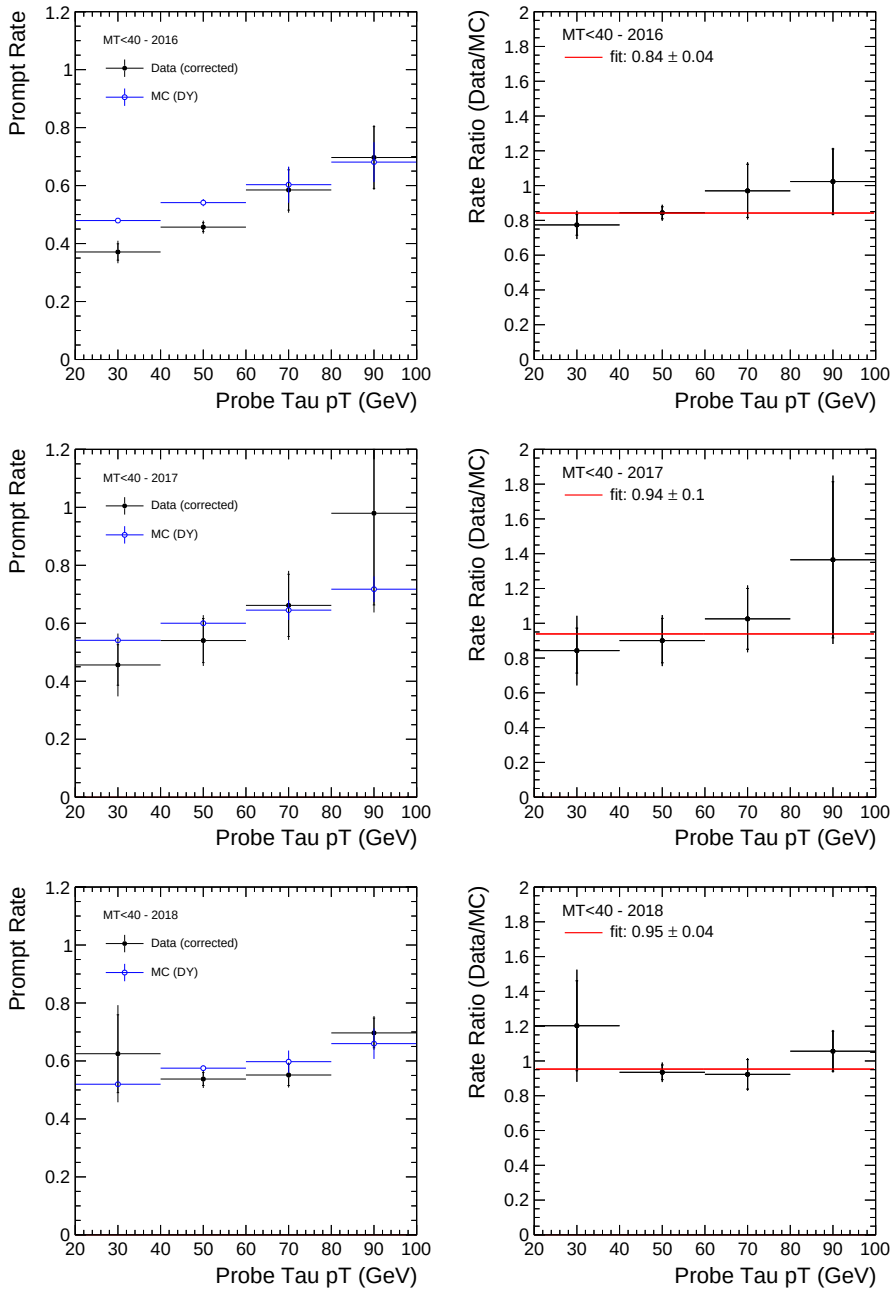

Figure D.23: Tau DY prompt rates in 2016, 2017, and 2018. The uncertainties are statistical only. Edge bins include over- and underflows.

D.2 Tau fake rates

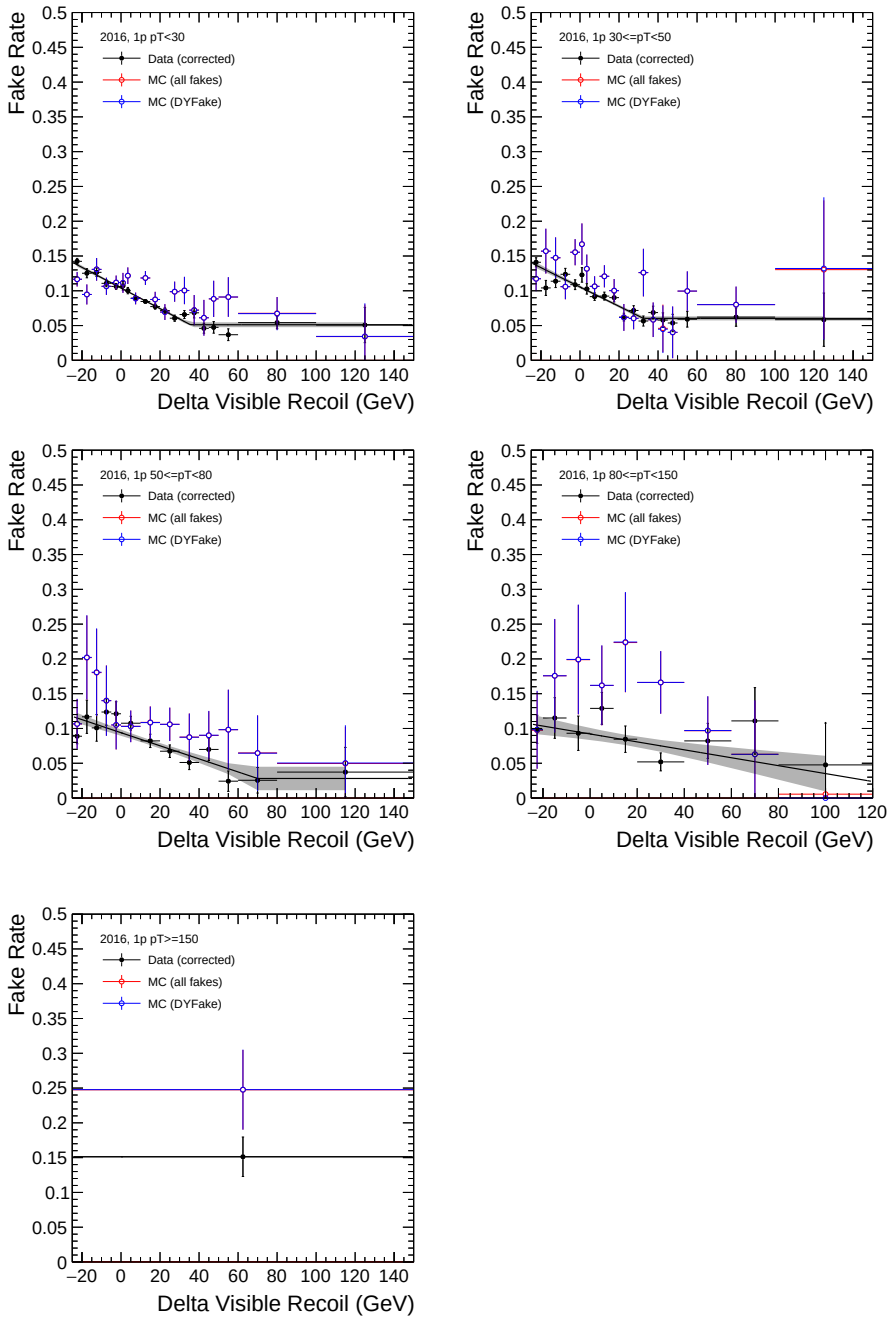

Figure D.24: Tau 1-prong DY fake rates in 2016. The uncertainties are statistical only. Edge bins include over- and underflows.

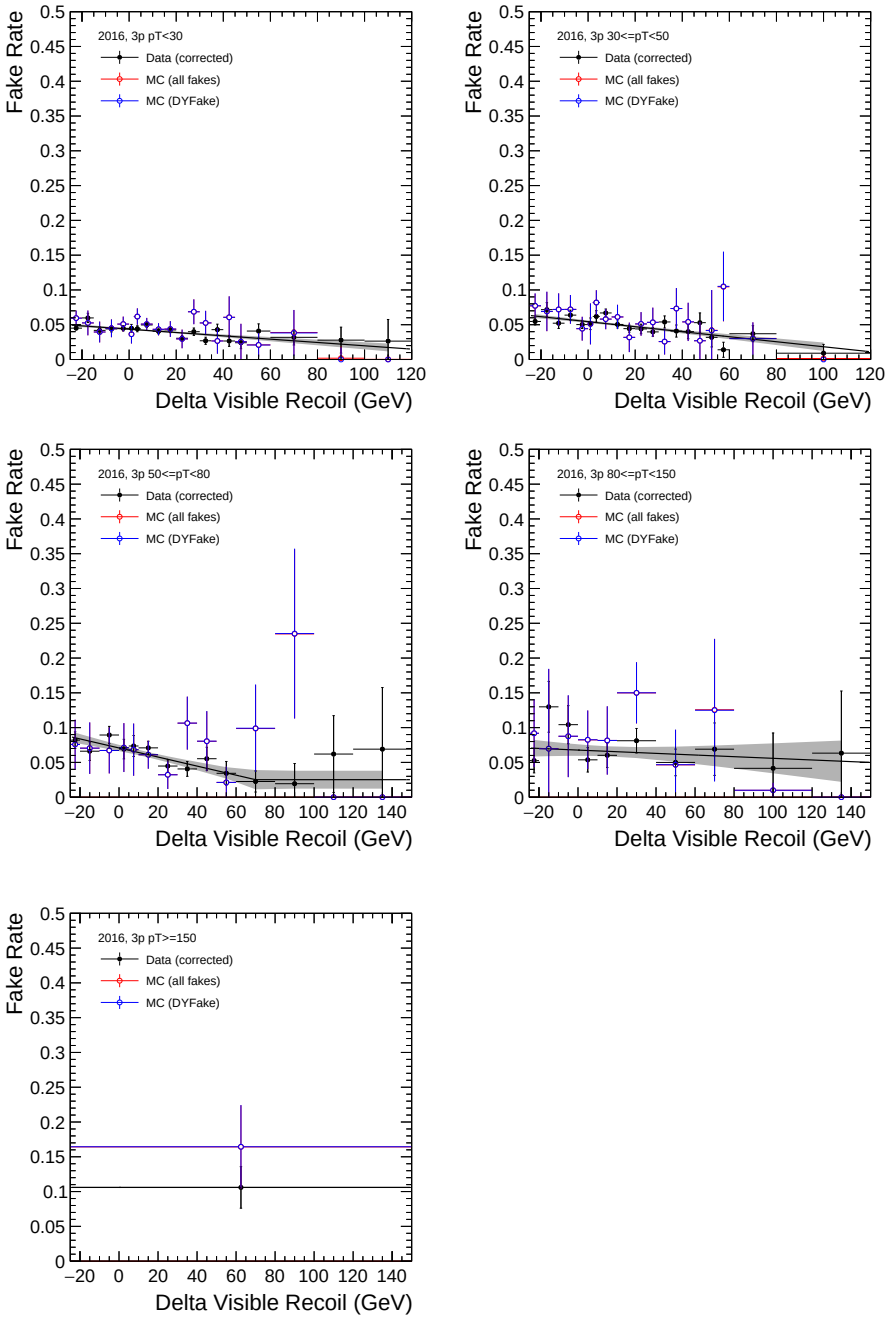

Figure D.25: Tau 3-prong DY fake rates in 2016. The uncertainties are statistical only. Edge bins include over- and underflows.

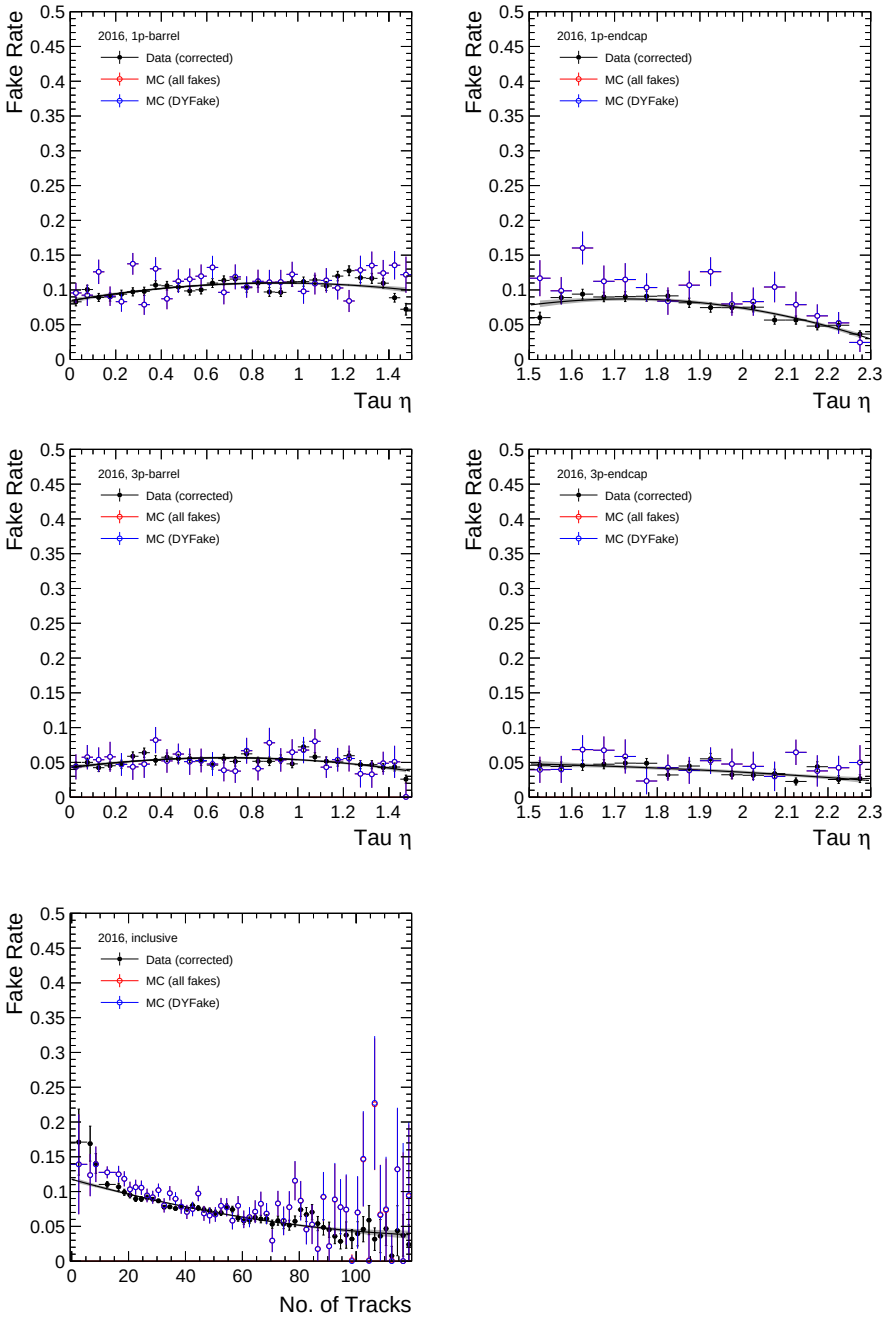

Figure D.26: Tau DY fake rate correction factors in 2016. The uncertainties are statistical only. Edge bins include over- and underflows.

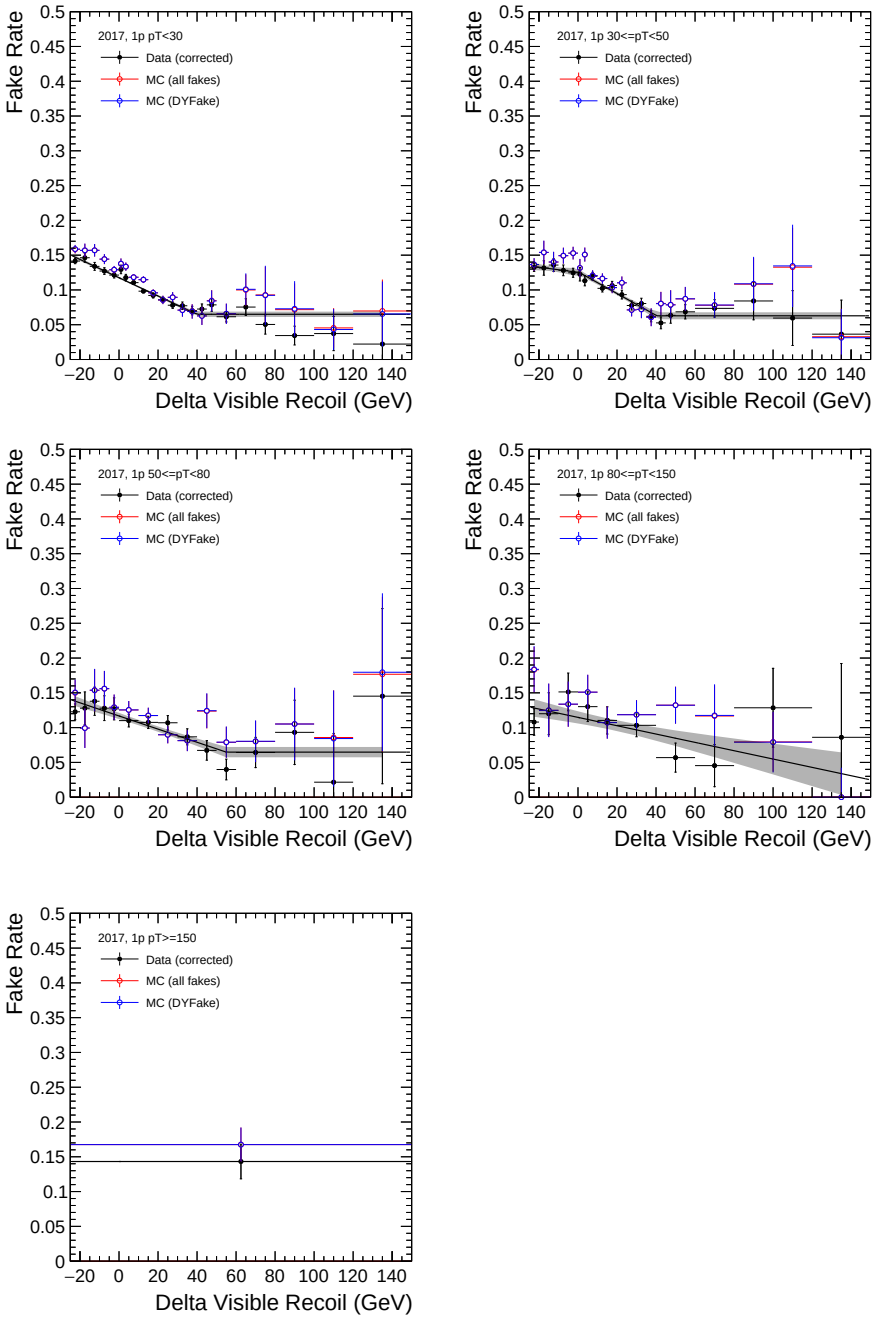

Figure D.27: Tau 1-prong DY fake rates in 2017. The uncertainties are statistical only. Edge bins include over- and underflows.

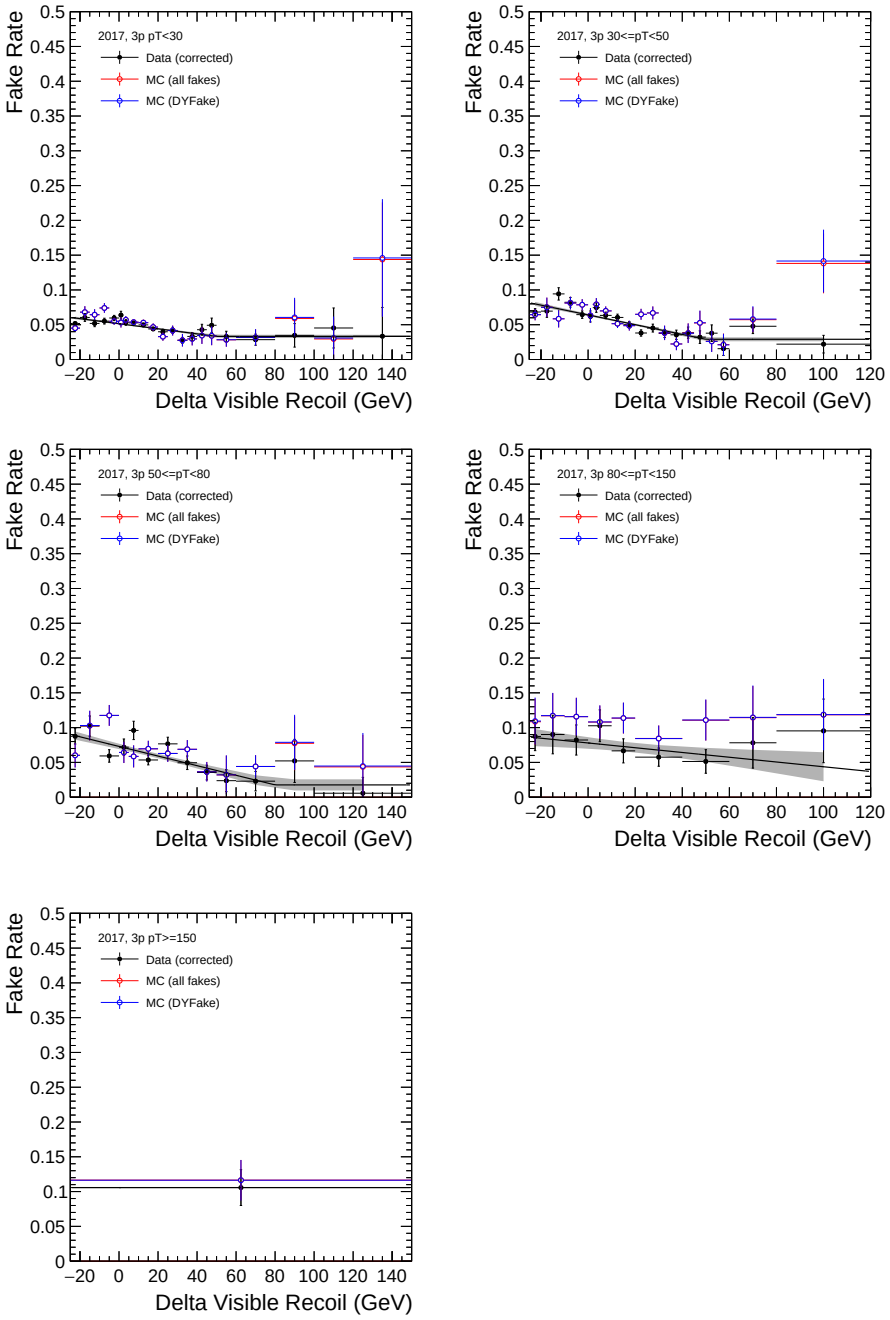

Figure D.28: Tau 3-prong DY fake rates in 2017. The uncertainties are statistical only. Edge bins include over- and underflows.

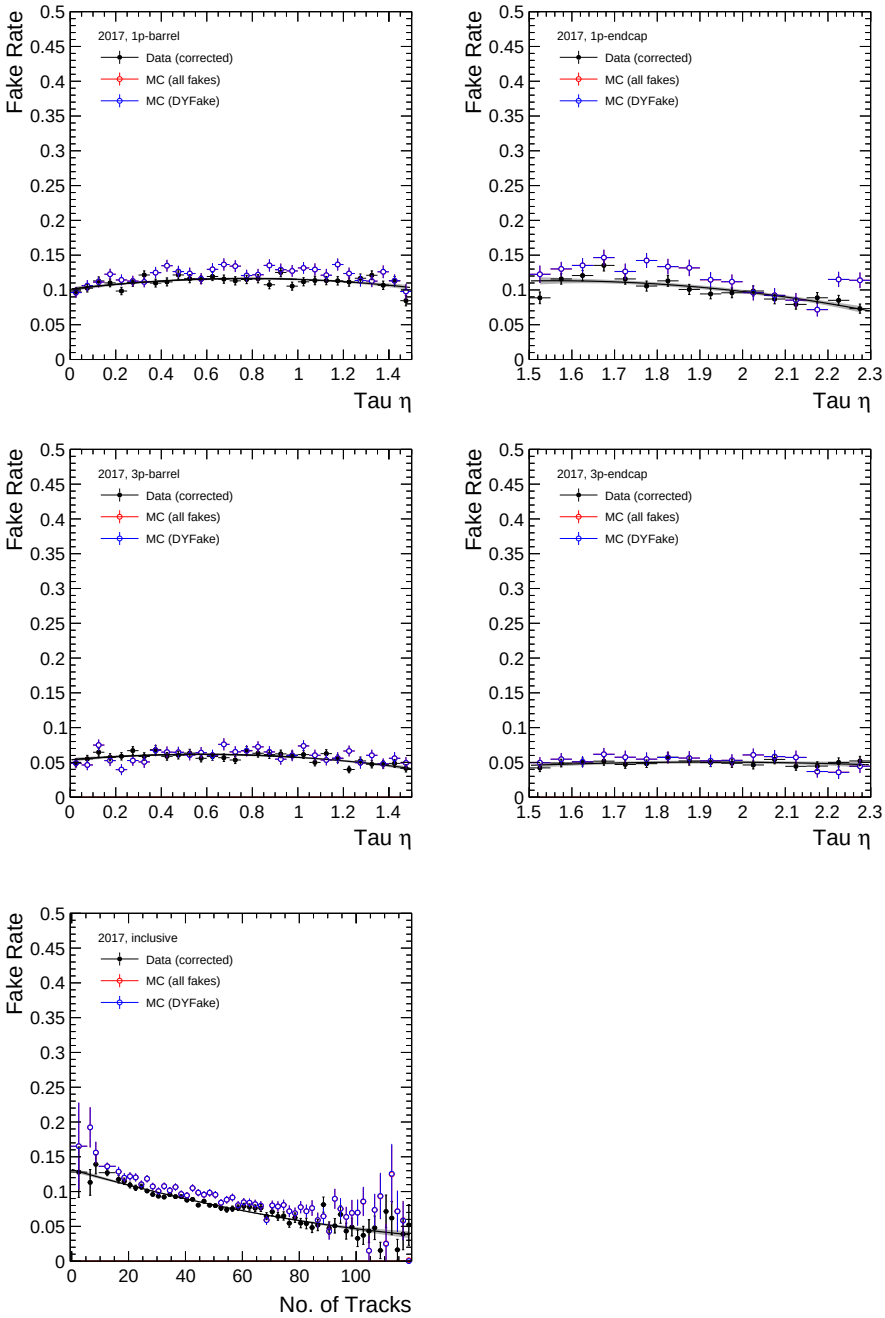

Figure D.29: Tau DY fake rate correction factors in 2017. The uncertainties are statistical only. Edge bins include over- and underflows.

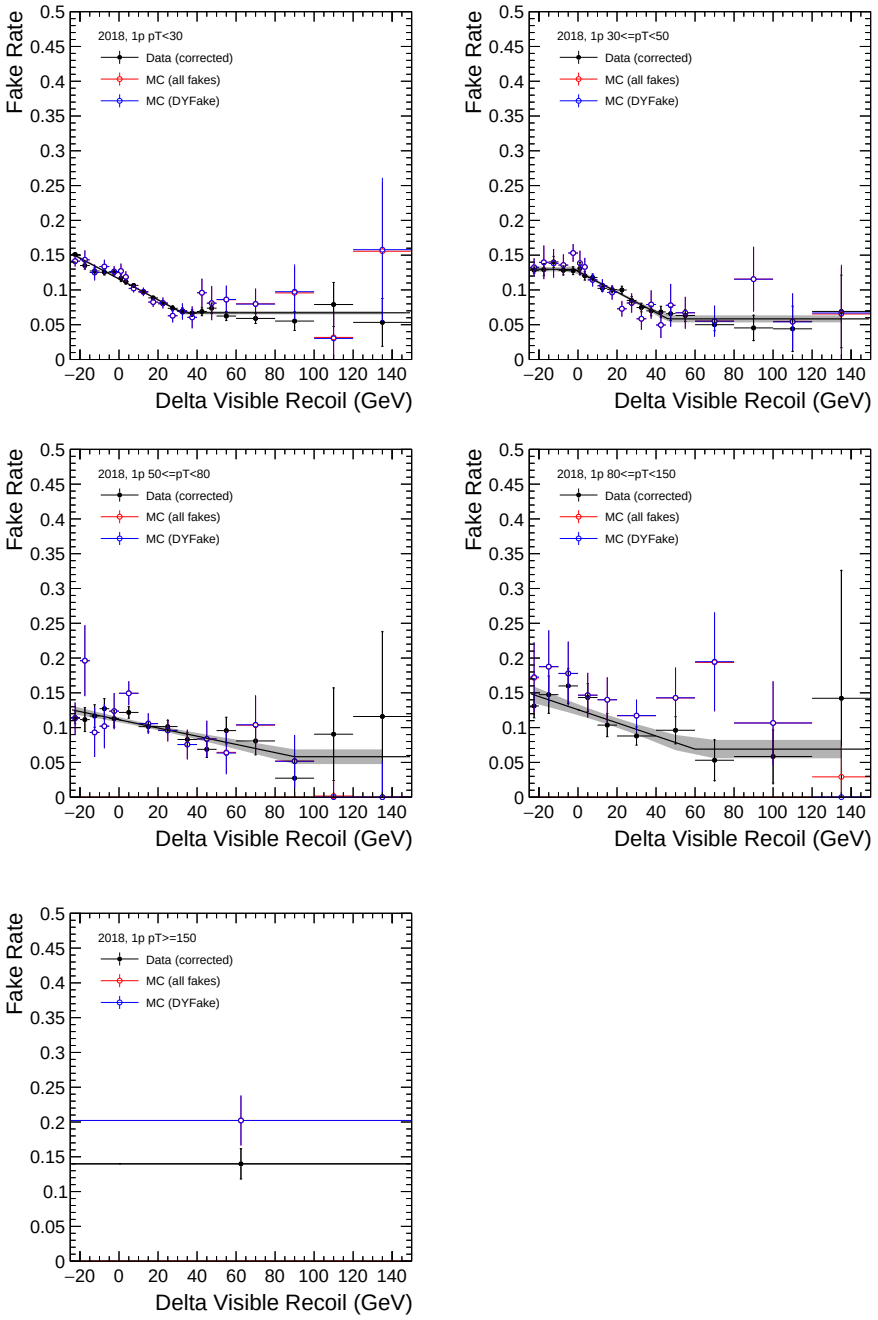

Figure D.30: Tau 1-prong DY fake rates in 2018. The uncertainties are statistical only. Edge bins include over- and underflows.

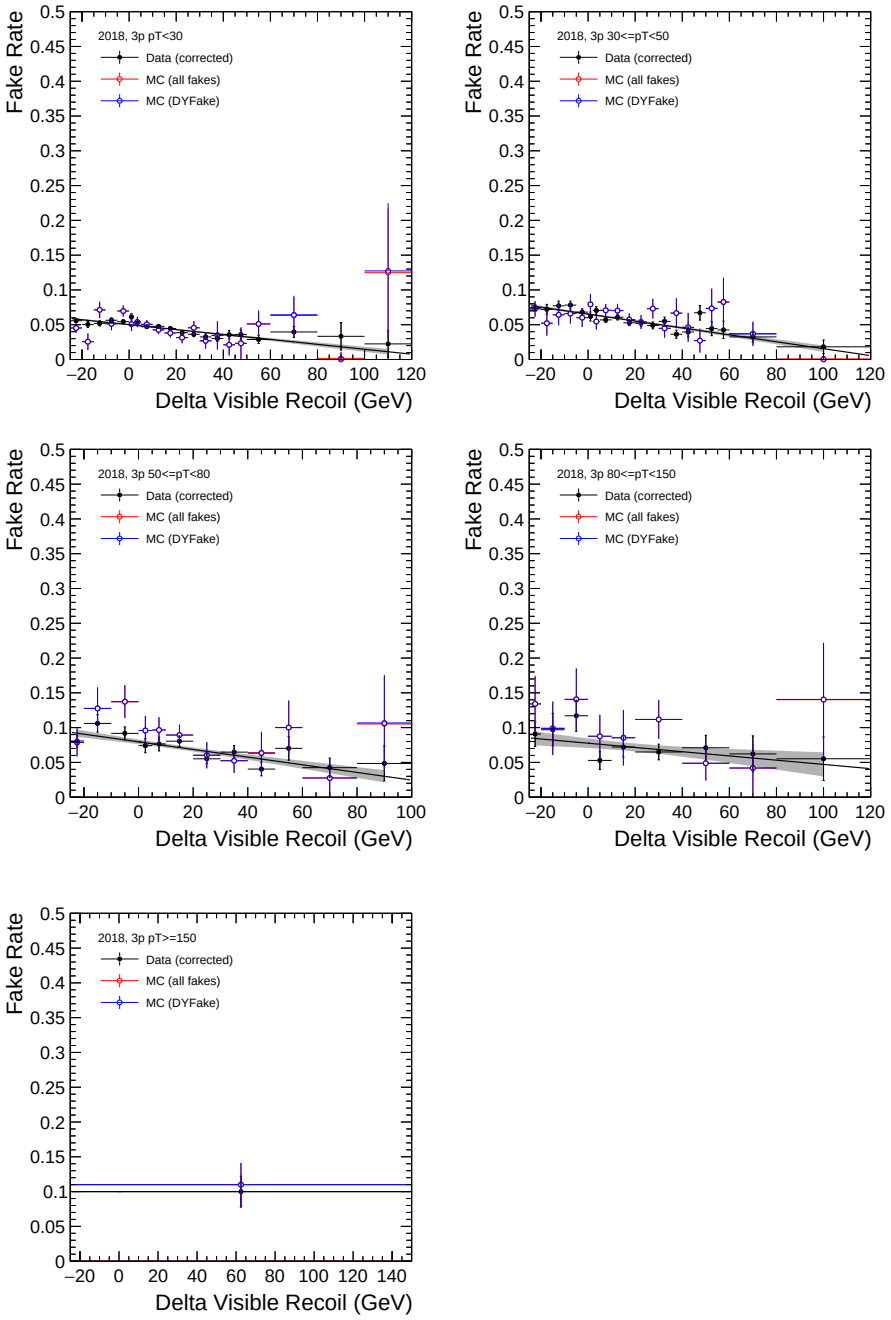

Figure D.31: Tau 3-prong DY fake rates in 2018. The uncertainties are statistical only. Edge bins include over- and underflows.

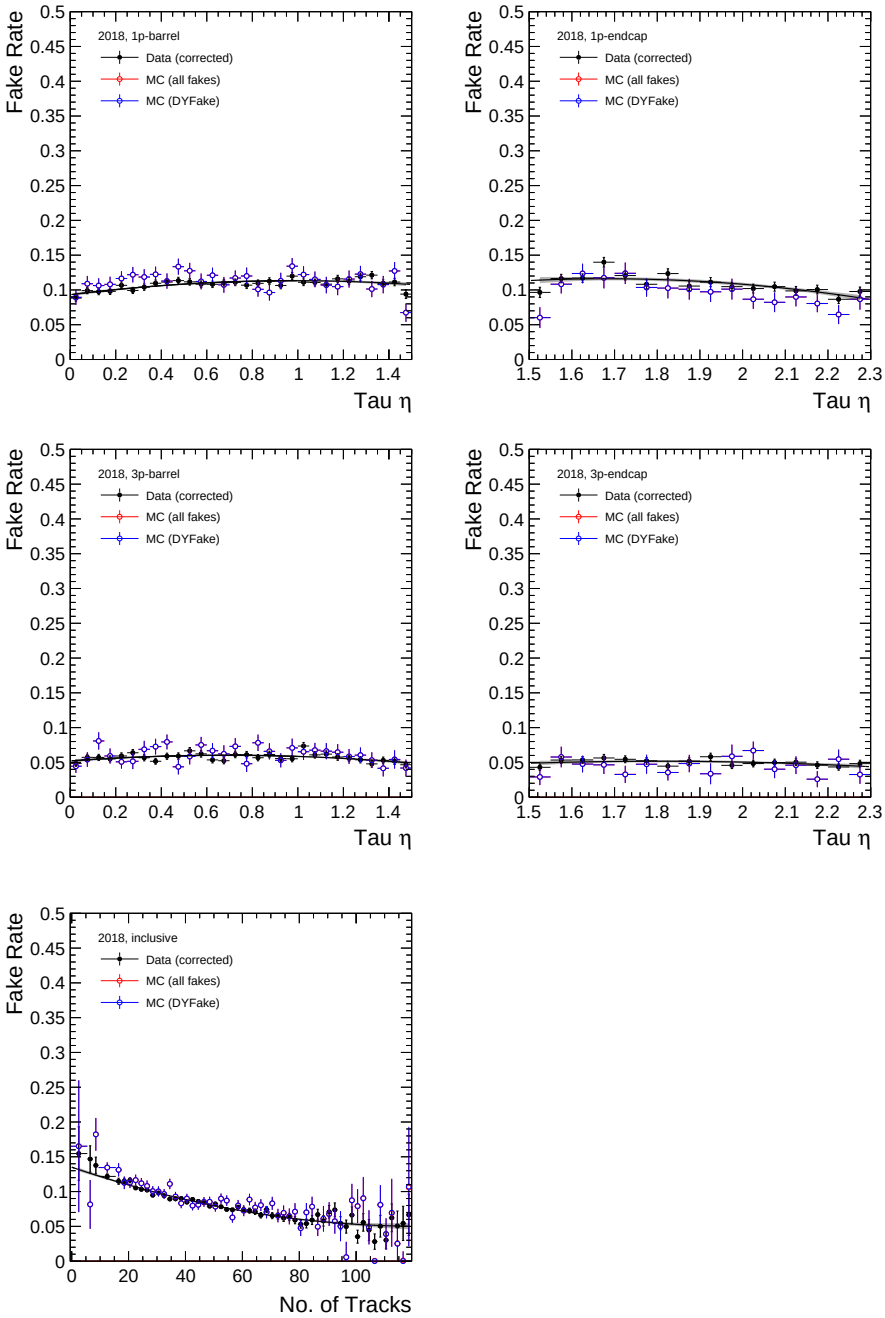

Figure D.32: Tau DY fake rate correction factors in 2018. The uncertainties are statistical only. Edge bins include over- and underflows.

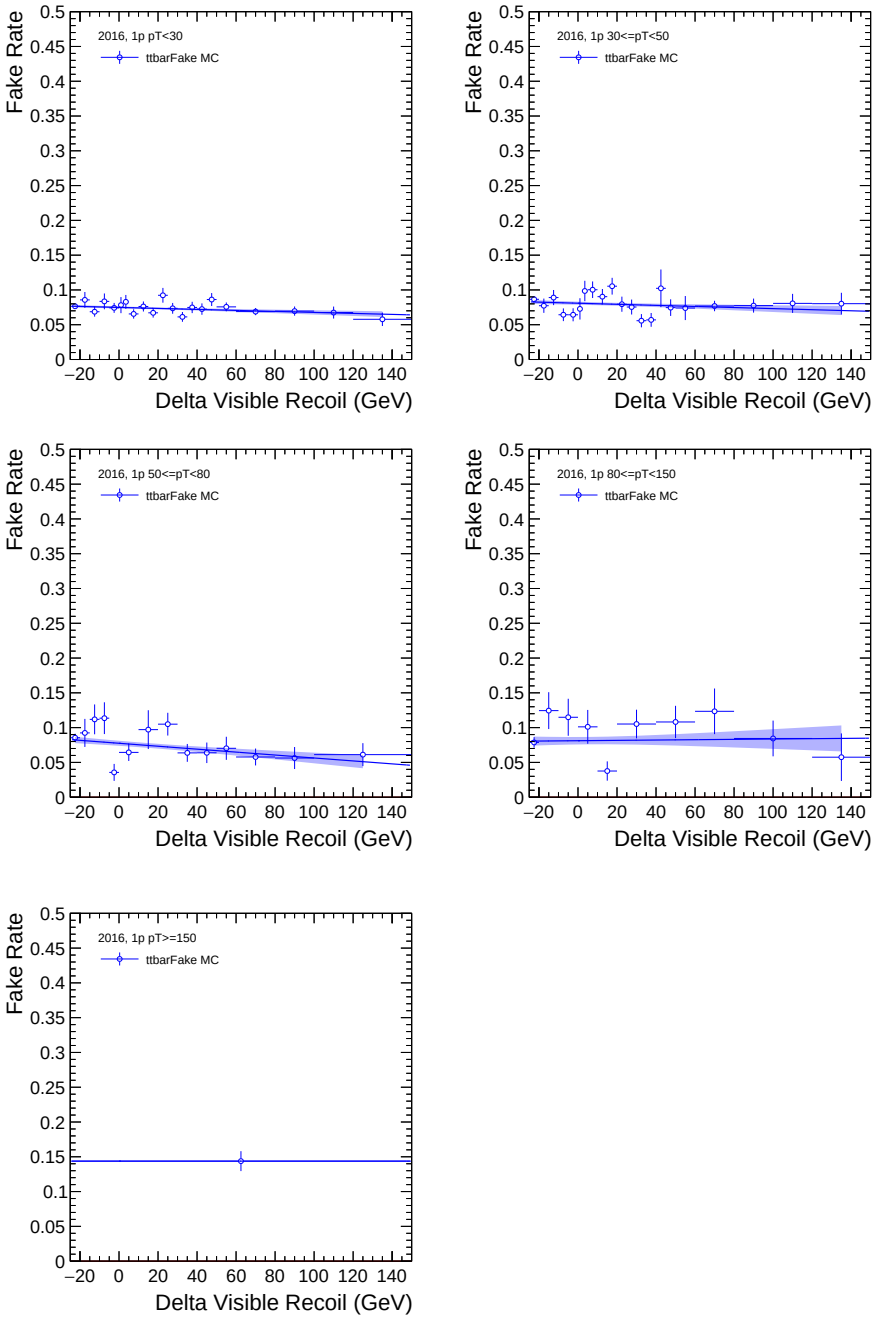

Figure D.33: Tau 1-prong $t\bar{t}$ fake rates in 2016. The uncertainties are statistical only. Edge bins include over- and underflows.

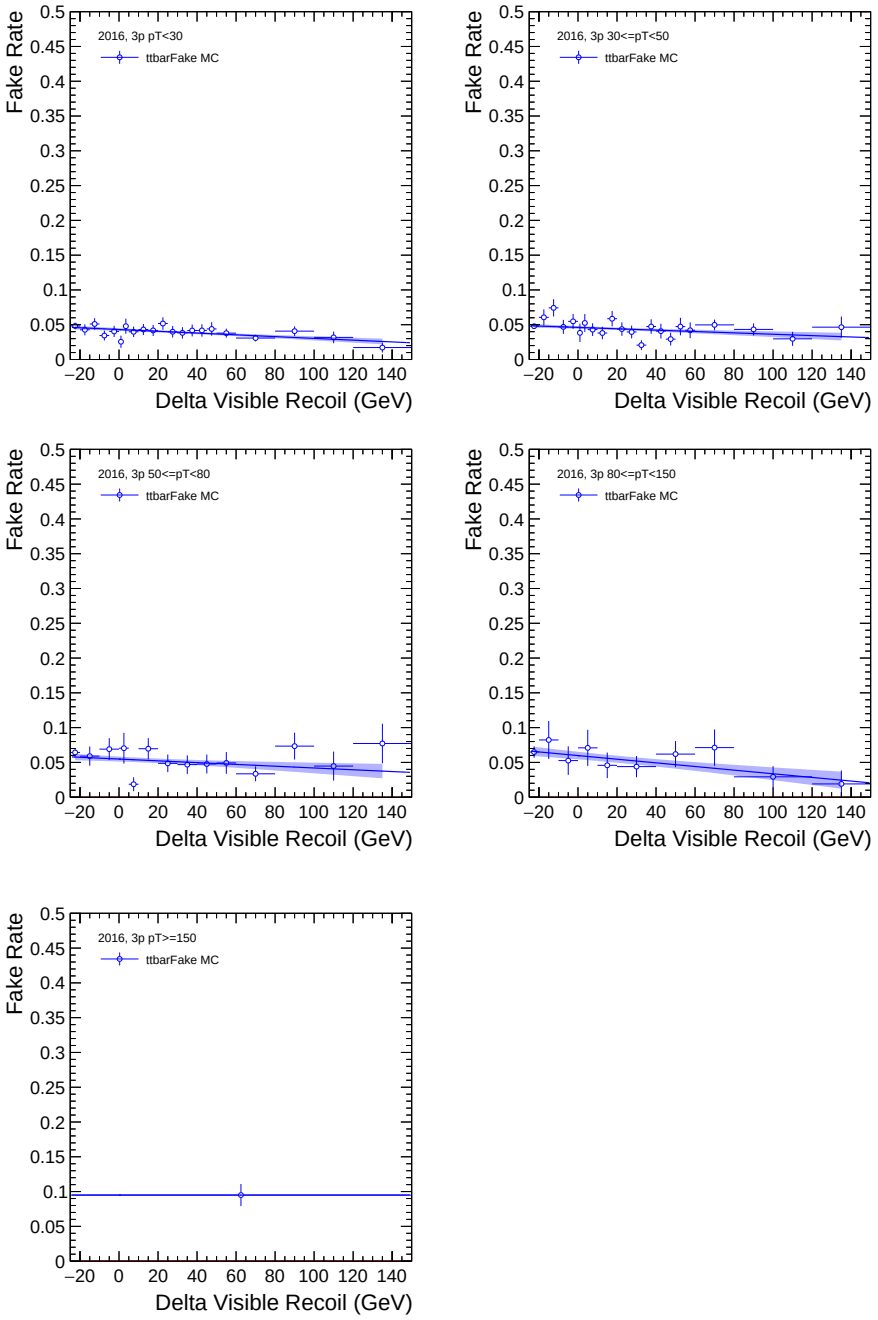

Figure D.34: Tau 3-prong $t\bar{t}$ fake rates in 2016. The uncertainties are statistical only. Edge bins include over- and underflows.

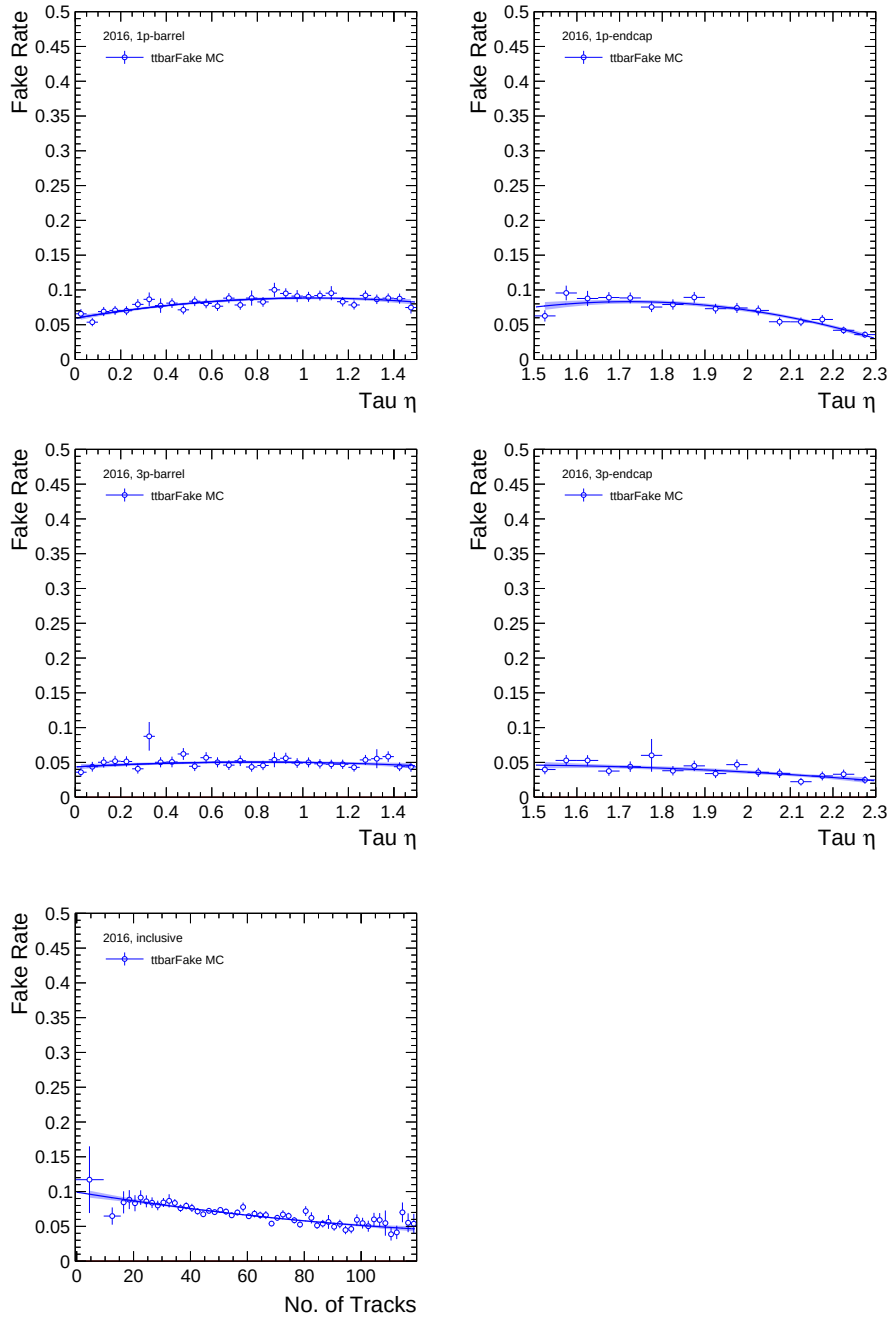

Figure D.35: Tau $t\bar{t}$ fake rate correction factors in 2016. The uncertainties are statistical only. Edge bins include over- and underflows.

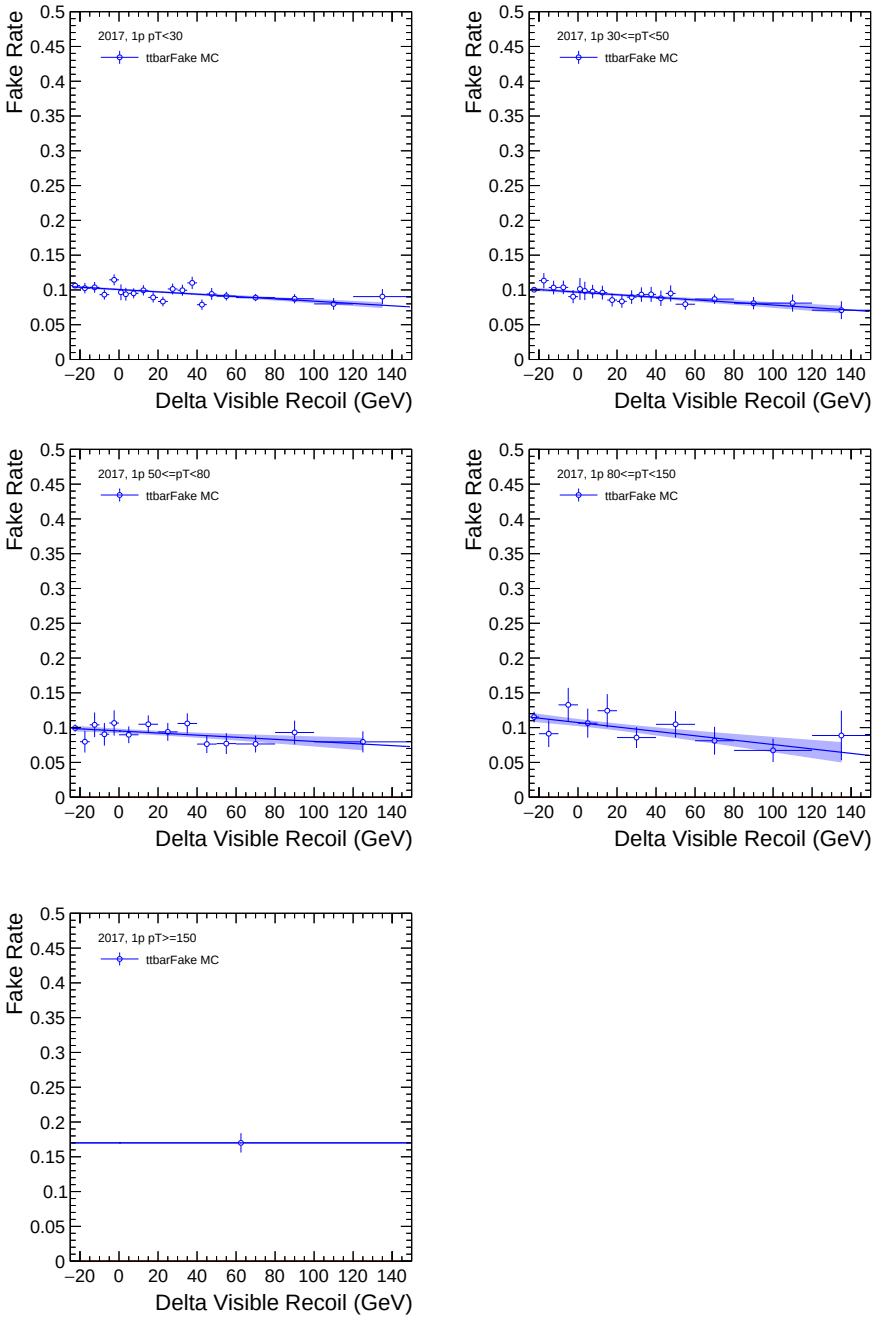

Figure D.36: Tau 1-prong $t\bar{t}$ fake rates in 2017. The uncertainties are statistical only. Edge bins include over- and underflows.

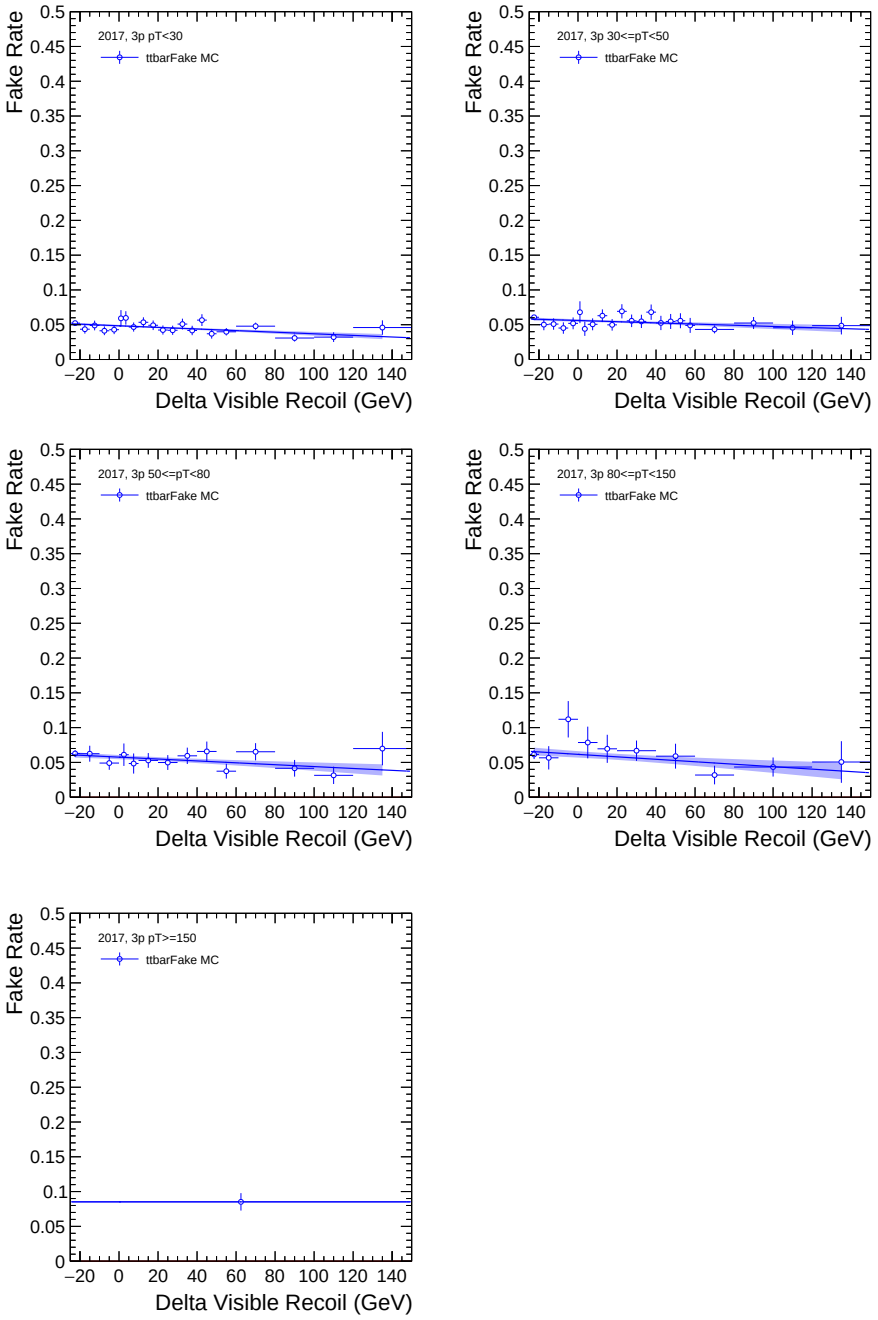

Figure D.37: Tau 3-prong $t\bar{t}$ fake rates in 2017. The uncertainties are statistical only. Edge bins include over- and underflows.

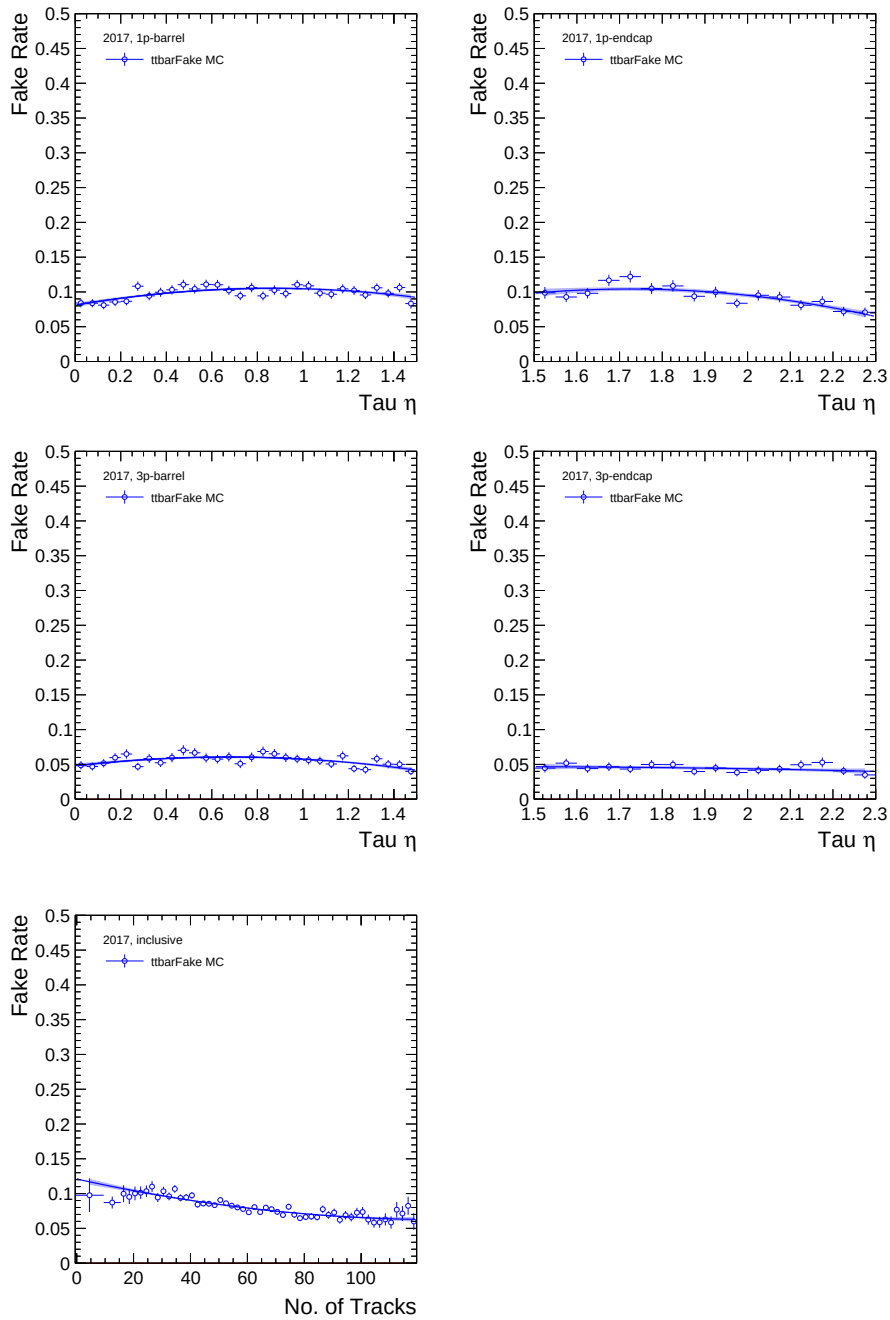

Figure D.38: Tau $t\bar{t}$ fake rate correction factors in 2017. The uncertainties are statistical only. Edge bins include over- and underflows.

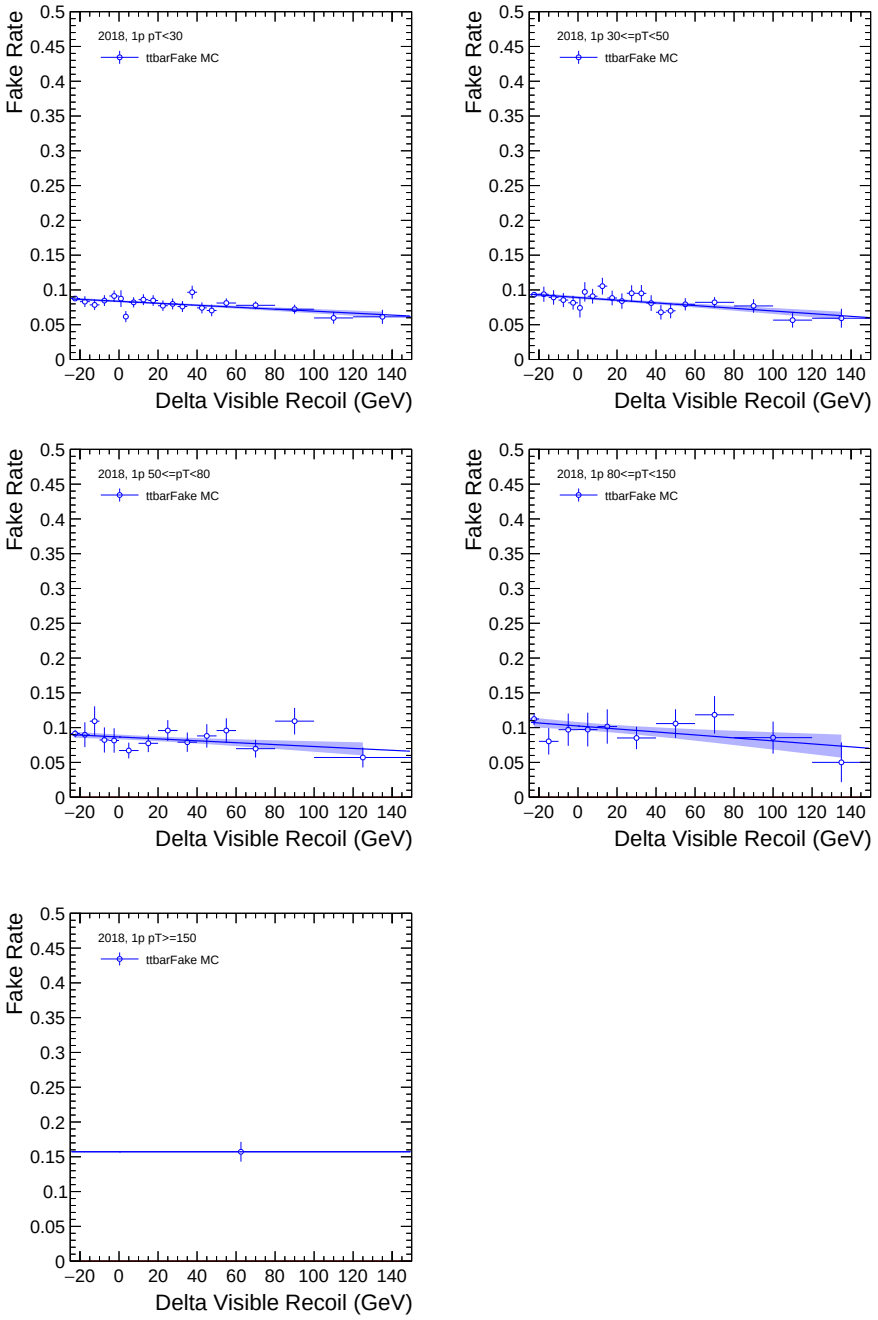

Figure D.39: Tau 1-prong $t\bar{t}$ fake rates in 2018. The uncertainties are statistical only. Edge bins include over- and underflows.

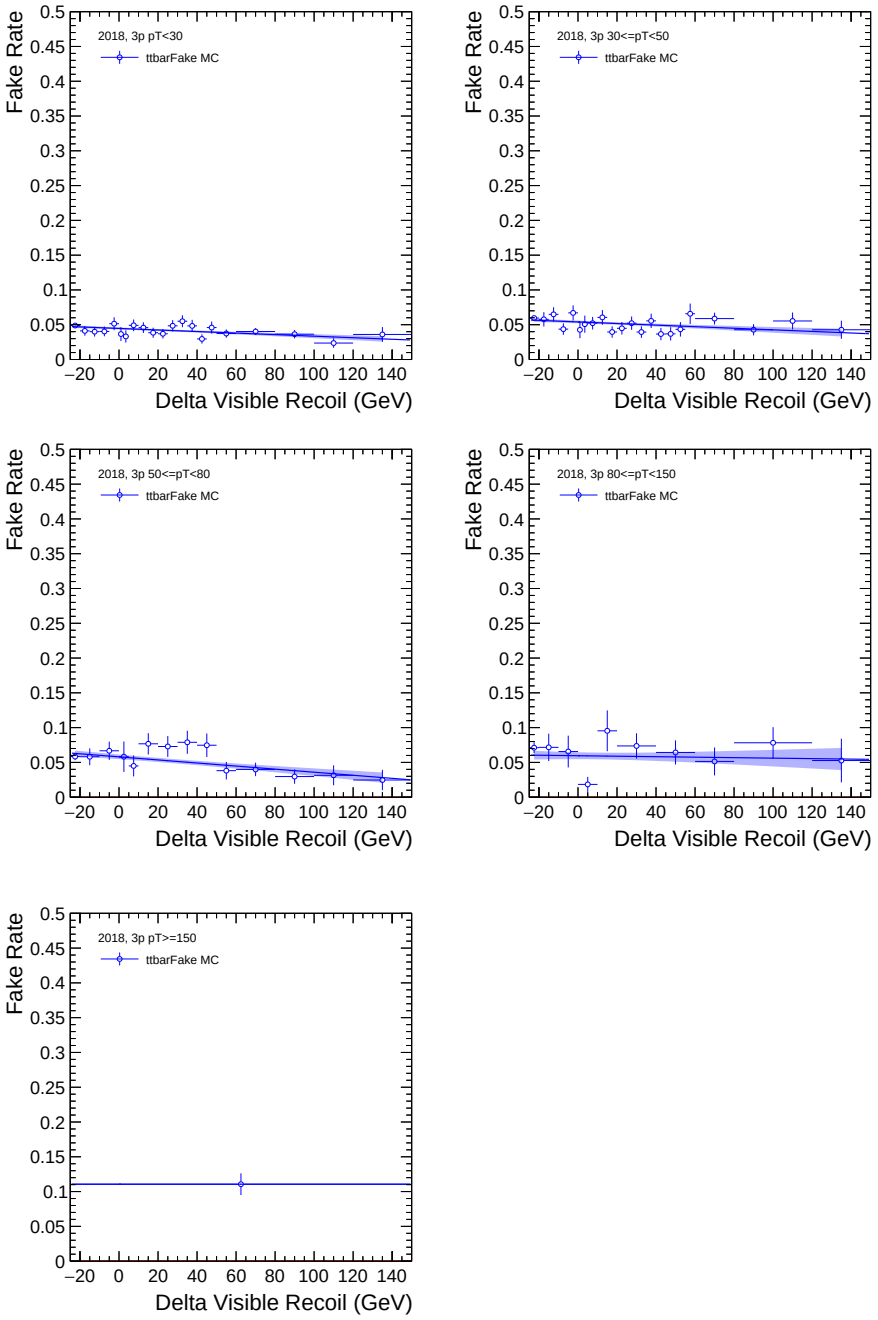

Figure D.40: Tau 3-prong $t\bar{t}$ fake rates in 2018. The uncertainties are statistical only. Edge bins include over- and underflows.

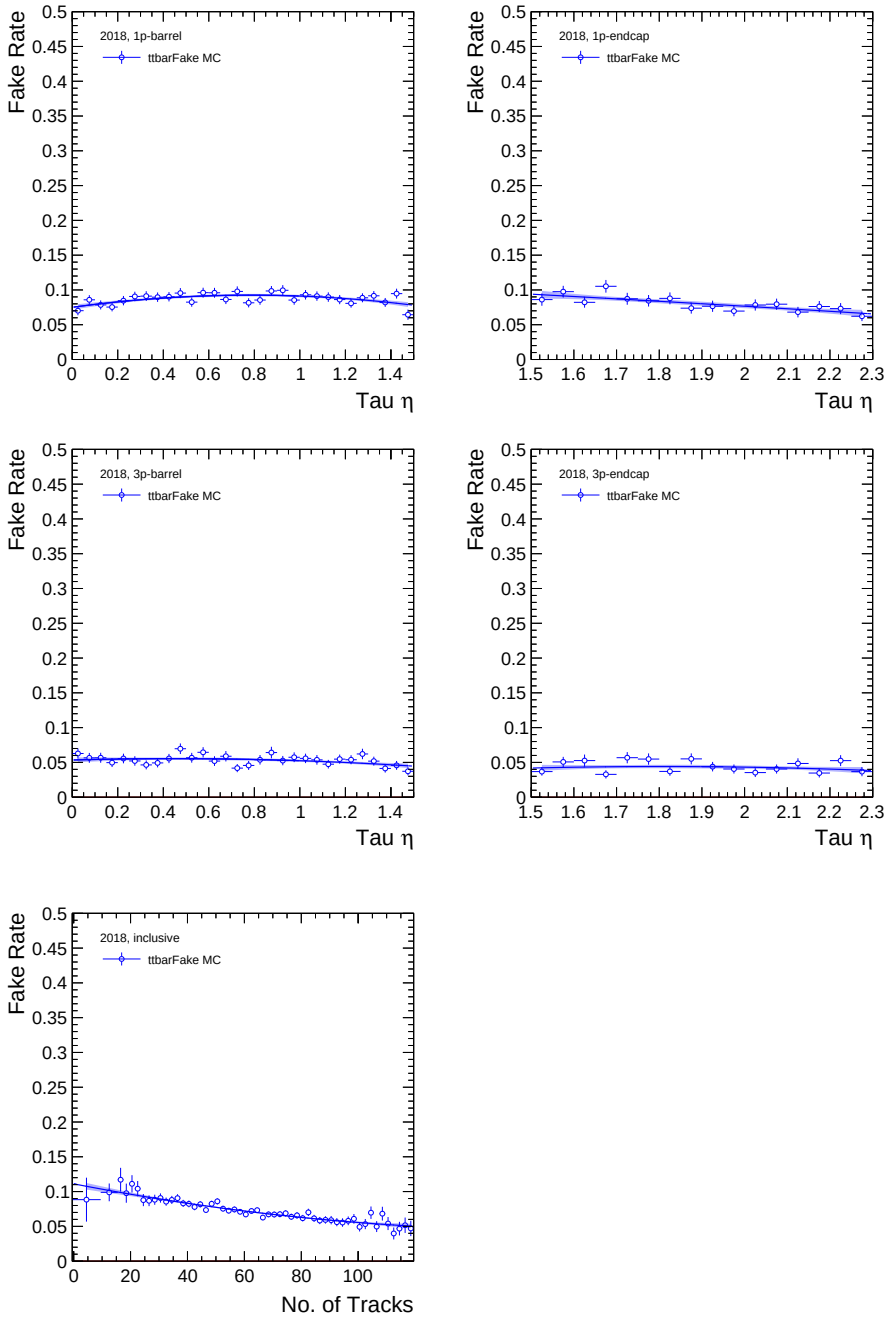

Figure D.41: Tau $t\bar{t}$ fake rate correction factors in 2018. The uncertainties are statistical only. Edge bins include over- and underflows.

E Misidentified lepton validation in $t\bar{t}$ MC

E.1 Misidentified light lepton validation

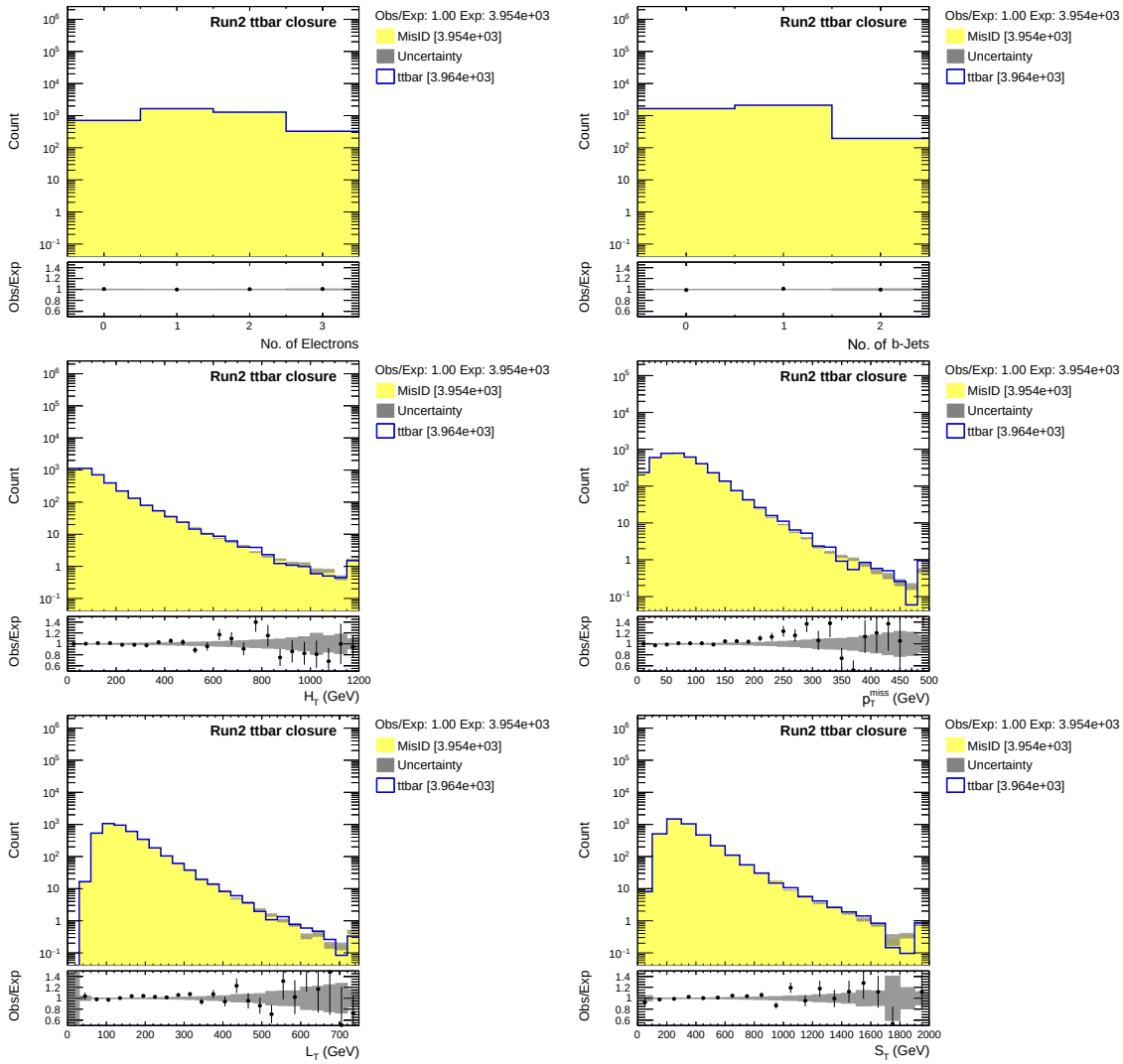

Figure E.42: Misidentified light lepton validation in $t\bar{t}$ MC in Run-2. The uncertainties are statistical only.

E.2 Misidentified tau validation

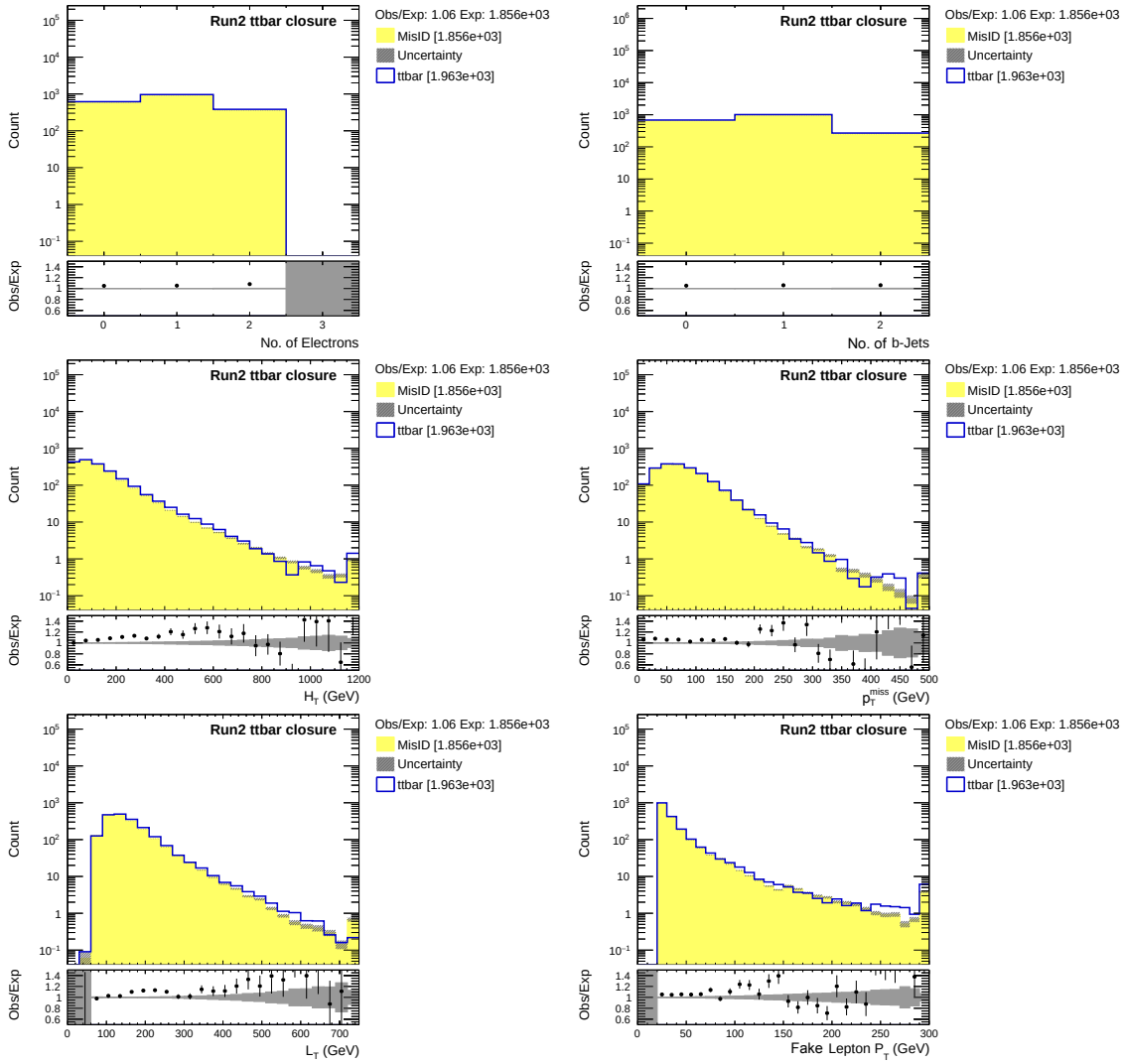

Figure E.43: Misidentified tau validation in $t\bar{t}$ MC in Run-2. The uncertainties are statistical only.

Bibliography

[1] CMS Collaboration. *Inclusive nonresonant multilepton probes of new phenomena at $\sqrt{s}$ =13 TeV*. CMS Physics Analysis Summary CMS-PAS-EXO-21-002. 2021.

[2] CMS Collaboration. *Inclusive nonresonant multilepton probes of new phenomena at $\sqrt{s}$ =13 TeV*. Unpublished internal CMS Collaboration documentation CMS-AN-20-052. 2021.

[3] J. C. Gregory. "A Short History of Atomism: from Democritus to Bohr". In: *Nature* 130 (1932), p. 947. DOI: 10.1038/130947b0.

[4] I. J. R. Aitchison and A. J. G. Hey. *Gauge Theories in Particle Physics: Volume I: From Relativistic Quantum Mechanics to QED, Third Edition*. Graduate Student Series in Physics. CRC Press, 2002. ISBN: 9780849387753.

[5] I. J. R. Aitchison and A. J. G. Hey. *Gauge Theories in Particle Physics, Volume II: QCD and the Electroweak Theory, Third Edition*. Graduate Student Series in Physics. CRC Press, 2003. ISBN: 9780849387760.

[6] M. E. Peskin and D. V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995. ISBN: 9780201503975.

[7] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.

[8] C. Biggio and F. Bonnet. "Implementation of the type-III Seesaw Model in FeynRules/MadGraph and Prospects for Discovery with Early LHC Data". In: *Eur. Phys. J. C* 72 (2012), p. 1899. DOI: 10.1140/epjc/s10052-012-1899-z.

[9] S. Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett. B* 716.1 (2012), p. 30. DOI: 10.1016/j.physletb.2012.08.021.

[10] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716.1 (2012), p. 1. DOI: 10.1016/j.physletb.2012.08.020.

[11] N. Kumar and S. P. Martin. "Vectorlike Leptons at the Large Hadron Collider". In: *Phys. Rev. D* 92.11 (2015), p. 115018. DOI: 10.1103/PhysRevD.92.115018.

[12] S. Davidson and P. Verdier. "Leptoquarks decaying to a top quark and a charged lepton at hadron colliders". In: *Phys. Rev. D* 83 (2011), p. 115016. DOI: 10.1103/PhysRevD.83.115016.

[13] L. Evans and P. Bryant. "LHC Machine". In: *JINST* 3.08 (2008), S08001. DOI: 10.1088/1748-0221/3/08/s08001.

[14] S. Chatrchyan et al. "The CMS experiment at the CERN LHC". In: *JINST* 3.08 (2008), S08004. DOI: 10.1088/1748-0221/3/08/s08004.

[15] C. Tully and G. Kane. "Elementary Particle Physics in a Nutshell." In: *American Journal of Physics* 80 (2012), p. 355. DOI: 10.1119/1.3683255.

[16] C. P. Burgess and G. D. Moore. *The standard model: A primer*. Cambridge University Press, 2006. ISBN: 9781107404267.

[17] M. D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014. ISBN: 9781107034730.

[18] A. Purcell. "Go on a particle quest at the first CERN webfest". In: *CERN Bulletin* 35-36 (2012).

[19] F. Capozzi et al. "Status of three-neutrino oscillation parameters, circa 2013". In: *Phys. Rev. D* 89 (9 2014), p. 093018. DOI: 10.1103/PhysRevD.89.093018.

[20] S. Martin. "A Supersymmetry primer". In: *Adv. Ser. Direct. High Energy Phys.* 21 (2010), p. 1. DOI: 10.1142/9789812839657_0001.

[21] A. Pomarol. *Beyond the Standard Model*. 2012. arXiv: 1202.1391 [hep-ph].

[22] P. Minkowski. "$\mu \to e\gamma$ at a Rate of One Out of 10^9 Muon Decays?" In: *Phys. Lett. B* 67 (1977), p. 421. DOI: 10.1016/0370-2693(77)90435-X.

[23] R. N. Mohapatra and G. Senjanovic. "Neutrino Mass and Spontaneous Parity Nonconservation". In: *Phys. Rev. Lett.* 44 (1980), p. 912. DOI: 10.1103/PhysRevLett.44.912.

[24] M. Magg and C. Wetterich. "Neutrino Mass Problem and Gauge Hierarchy". In: *Phys. Lett. B* 94 (1980), p. 61. DOI: 10.1016/0370-2693(80)90825-4.

[25] R. N. Mohapatra and G. Senjanovic. "Neutrino Masses and Mixings in Gauge Models with Spontaneous Parity Violation". In: *Phys. Rev. D* 23 (1981), p. 165. DOI: 10.1103/PhysRevD.23.165.

[26] J. Schechter and J. W. F. Valle. "Neutrino Masses in SU(2) x U(1) Theories". In: *Phys. Rev. D* 22 (1980), p. 2227. DOI: 10.1103/PhysRevD.22.2227.

[27] J. Schechter and J. W. F. Valle. "Neutrino Decay and Spontaneous Violation of Lepton Number". In: *Phys. Rev. D* 25 (1982), p. 774. DOI: 10.1103/PhysRevD.25.774.

[28] R. Foot et al. "Seesaw Neutrino Masses Induced by a Triplet of Leptons". In: *Z. Phys. C* 44 (1989), p. 441. DOI: 10.1007/BF01415558.

[29] R. N. Mohapatra. "Mechanism for Understanding Small Neutrino Mass in Superstring Theories". In: *Phys. Rev. Lett.* 56 (1986), p. 561. DOI: 10.1103/PhysRevLett.56.561.

[30] R. N. Mohapatra and J. W. F. Valle. "Neutrino Mass and Baryon Number Nonconservation in Superstring Models". In: *Phys. Rev. D* 34 (1986), p. 1642. DOI: 10.1103/PhysRevD.34.1642.

[31] A. Abada et al. "Low energy effects of neutrino masses". In: *JHEP* 12 (2007), p. 061. DOI: 10.1088/1126-6708/2007/12/061.

[32] A. Abada et al. "$\mu \to e\gamma$ and $\tau \to \ell\gamma$ decays in the fermion triplet seesaw model". In: *Phys. Rev. D* 78 (2008), p. 033007. DOI: 10.1103/PhysRevD.78.033007.

[33] F. del Aguila, J. de Blas, and M. Perez-Victoria. "Effects of new leptons in Electroweak Precision Data". In: *Phys. Rev. D* 78 (2008), p. 013010. DOI: 10.1103/PhysRevD.78.013010.

[34] D. Goswami and P. Poulose. "Direct searches of type-III seesaw triplet fermions at high energy e^+e^- collider". In: *Eur. Phys. J. C* 78 (2018), p. 42. DOI: 10.1140/epjc/s10052-017-5478-1.

[35] A. M. Sirunyan et al. "Search for physics beyond the standard model in multilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 03 (2020), p. 051. DOI: 10.1007/JHEP03(2020)051.

[36] F. del Aguila et al. "Vector Like Fermion and Standard Higgs Production at Hadron Colliders". In: *Nucl. Phys. B* 334 (1990), p. 1. DOI: 10.1016/0550-3213(90)90655-W.

[37] S. P. Martin. "Extra vector-like matter and the lightest Higgs scalar boson mass in low-energy supersymmetry". In: *Phys. Rev. D* 81 (2010), p. 035004. DOI: 10.1103/PhysRevD.81.035004.

[38] J. Halverson, N. Orlofsky, and A. Pierce. "Vectorlike Leptons as the Tip of the Dark Matter Iceberg". In: *Phys. Rev. D* 90.1 (2014), p. 015002. DOI: 10.1103/PhysRevD.90.015002.

[39] R. Dermisek and A. Raval. "Explanation of the Muon g-2 Anomaly with Vector-like Leptons and its Implications for Higgs Decays". In: *Phys. Rev. D* 88 (2013), p. 013017. DOI: 10.1103/PhysRevD.88.013017.

[40] M. Endo et al. "Higgs Mass and Muon Anomalous Magnetic Moment in Supersymmetric Models with Vector-Like Matters". In: *Phys. Rev. D* 84 (2011), p. 075017. DOI: 10.1103/PhysRevD.84.075017.

[41] G. Hiller et al. "Model Building from Asymptotic Safety with Higgs and Flavor Portals". In: *Phys. Rev. D* 102.9 (2020), p. 095023. DOI: 10.1103/PhysRevD.102.095023. arXiv: 2008.08606 [hep-ph].

[42] F. F. Freitas et al. *Phenomenology of vector-like leptons with Deep Learning at the Large Hadron Collider*. 2020. arXiv: 2010.01307 [hep-ph].

[43] P. N. Bhattiprolu and S. P. Martin. "Prospects for vectorlike leptons at future proton-proton colliders". In: *Phys. Rev. D* 100.1 (2019), p. 015033. DOI: 10.1103/PhysRevD.100.015033.

[44] R. Dermisek, A. Raval, and S. Shin. "Effects of vectorlike leptons on $h \to 4\ell$ and the connection to the muon g-2 anomaly". In: *Phys. Rev. D* 90.3 (2014), p. 034023. DOI: 10.1103/PhysRevD.90.034023.

[45] R. Dermisek et al. "Limits on Vectorlike Leptons from Searches for Anomalous Production of Multi-Lepton Events". In: *JHEP* 12 (2014), p. 013. DOI: 10.1007/JHEP12(2014)013.

[46] A. M. Sirunyan et al. "Search for vector-like leptons in multilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. D* 100.5 (2019), p. 052003. DOI: 10.1103/PhysRevD.100.052003.

[47] P. Achard et al. "Search for heavy neutral and charged leptons in e^+e^- annihilation at LEP". In: *Phys. Lett. B* 517 (2001), p. 75. DOI: 10.1016/S0370-2693(01)01005-X.

[48] W. Buchmuller, R. Ruckl, and D. Wyler. "Leptoquarks in Lepton - Quark Collisions". In: *Phys. Lett. B* 191 (1987). [Erratum: Phys. Lett. B 448, 320 (1999)], p. 442. DOI: 10.1016/0370-2693(87)90637-X.

[49] J. C. Pati and A. Salam. "Lepton Number as the Fourth Color". In: *Phys. Rev. D* 10 (1974). [Erratum: Phys. Rev. D 11, 703 (1975)], p. 275. DOI: 10.1103/PhysRevD.10.275.

[50] H. Georgi and S. L. Glashow. "Unity of All Elementary Particle Forces". In: *Phys. Rev. Lett.* 32 (1974), p. 438. DOI: 10.1103/PhysRevLett.32.438.

[51] H. Fritzsch and P. Minkowski. "Unified Interactions of Leptons and Hadrons". In: *Annals Phys.* 93 (1975), p. 193. DOI: 10.1016/0003-4916(75)90211-0.

[52] S. Weinberg. "Supersymmetry at Ordinary Energies. 1. Masses and Conservation Laws". In: *Phys. Rev. D* 26 (1982), p. 287. DOI: 10.1103/PhysRevD.26.287.

[53] R. Barbier et al. "R-parity violating supersymmetry". In: *Phys. Rept.* 420 (2005), p. 1. DOI: 10.1016/j.physrep.2005.08.006.

[54] J. K. Mizukoshi, O. J. P. Eboli, and M. C. Gonzalez-Garcia. "Bounds on scalar leptoquarks from Z physics". In: *Nucl. Phys. B* 443 (1995), p. 20. DOI: 10.1016/0550-3213(95)00162-L.

[55] G. Aad et al. "Search for pair production of scalar leptoquarks decaying into first- or second-generation leptons and top quarks in proton–proton collisions at $\sqrt{s}$= 13 TeV with the ATLAS detector". In: *Eur. Phys. J. C* 81 (2020). DOI: 10.1140/epjc/s10052-021-09009-8.

[56] G. Aad et al. "Search for pair production of third-generation scalar leptoquarks decaying into a top quark and a τ-lepton in pp collisions at $\sqrt{s}$= 13 TeV with the ATLAS detector". In: *JHEP* 2106 (2021). DOI: 10.1007/JHEP06(2021)179.

[57] O. S. Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-1.

[58] K. Aamodt et al. "The ALICE experiment at the CERN LHC". In: *JINST* 3.08 (2008), S08002. DOI: 10.1088/1748-0221/3/08/s08002.

[59] G. Aad et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3.08 (2008), S08003. DOI: 10.1088/1748-0221/3/08/s08003.

[60] A. Augusto Alves et al. "The LHCb Detector at the LHC". In: *JINST* 3.08 (2008), S08005. DOI: 10.1088/1748-0221/3/08/s08005.

[61] E. Mobs. *The CERN accelerator complex*. CERN Graphics CERN-GRAPHICS-2019-002. 2019.

[62] T. Sakuma. *Cutaway diagrams of CMS detector*. CERN Outreach CMS-OUTREACH-2019-001. 2019.

[63] L. R. F. Castillo. "The ATLAS and CMS detectors". In: *The Search and Discovery of the Higgs Boson*. 2053-2571. Morgan and Claypool Publishers, 2015. ISBN: 9781681740782. DOI: 10.1088/978-1-6817-4078-2ch4.

[64] CMS Collaboration. *The CMS tracker system project: Technical Design Report*. Technical design report CMS. Geneva: CERN, 1997.

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

[66] CMS Collaboration. *The CMS hadron calorimeter project: Technical Design Report*. Technical design report CMS. Geneva: CERN, 1997.

[67] CMS Collaboration. *The CMS magnet project: Technical Design Report*. Technical design report CMS. Geneva: CERN, 1997.

[68] CMS Collaboration. *The CMS muon project: Technical Design Report*. Technical design report CMS. Geneva: CERN, 1997.

[69] V. Khachatryan et al. "The CMS trigger system". In: *JINST* 12 (2017), P01020. DOI: 10.1088/1748-0221/12/01/P01020.

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

[71] P. Nason. "A New method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040.

[72] S. Frixione, P. Nason, and C. Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070.

[73] S. Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043.

[74] J. M. Campbell and R. K. Ellis. "MCFM for the Tevatron and the LHC". In: *Nucl. Phys. Proc. Suppl.* 205 (2010), p. 10. DOI: 10.1016/j.nuclphysbps.2010.08.011.

[75] Y. Gao et al. "Spin determination of single-produced resonances at hadron colliders". In: *Phys. Rev. D* 81 (2010), p. 075022. DOI: 10.1103/PhysRevD.81.075022.

[76] S. Bolognesi et al. "On the spin and parity of a single-produced resonance at the LHC". In: *Phys. Rev. D* 86 (2012), p. 095031. DOI: 10.1103/PhysRevD.86.095031.

[77] I. Anderson et al. "Constraining anomalous HVV interactions at proton and lepton colliders". In: *Phys. Rev. D* 89 (2014), p. 035007. DOI: 10.1103/PhysRevD.89.035007.

[78] A. V. Gritsan et al. "Constraining anomalous Higgs boson couplings to the heavy flavor fermions using matrix element techniques". In: *Phys. Rev. D* 94 (2016), p. 055023. DOI: 10.1103/PhysRevD.94.055023.

[79] T. Sjöstrand et al. "An Introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), p. 159. DOI: 10.1016/j.cpc.2015.01.024.

[80] B. Fuks et al. "Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV". In: *JHEP* 10 (2012), p. 081. DOI: 10.1007/JHEP10(2012)081.

[81] B. Fuks et al. "Precision predictions for electroweak superpartner production at hadron colliders with Resummino". In: *Eur. Phys. J. C* 73 (2013), p. 2480. DOI: 10.1140/epjc/s10052-013-2480-0.

[82] M. Kramer et al. "Pair production of scalar leptoquarks at the CERN LHC". In: *Phys. Rev. D* 71 (2005), p. 057503. DOI: 10.1103/PhysRevD.71.057503.

[83] J. Blumlein, E. Boos, and A. Kryukov. *Leptoquark pair production cross-sections at hadron colliders*. 1998. arXiv: hep-ph/9811271 [hep-ph].

[84] R. D. Ball et al. "Parton distributions for the LHC Run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040.

[85] R. D. Ball et al. "Parton distributions from high-precision collider data". In: *Eur. Phys. J. C* 77 (2017), p. 663. DOI: 10.1140/epjc/s10052-017-5199-5.

[86] R. Frederix and S. Frixione. "Merging meets matching in MC@NLO". In: *JHEP* 12 (2012), p. 061. DOI: 10.1007/JHEP12(2012)061.

[87] S. Hoeche et al. *Matching Parton Showers and Matrix Elements*. 2006. arXiv: 0602031 [hep-ph].

[88] S. Agostinelli et al. "GEANT4 - a simulation toolkit". In: *Nucl. Instrum. Meth.* 506 (2003), p. 250. DOI: 10.1016/S0168-9002(03)01368-8.

[89] A. M. Sirunyan et al. "Particle-flow reconstruction and global event description with the CMS detector". In: *JINST* 12 (2017), P10003. DOI: 10.1088/1748-0221/12/10/P10003.

[90] A. M. Sirunyan et al. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JINST* 13 (2018), P06015. DOI: 10.1088/1748-0221/13/06/P06015.

[91] V. Khachatryan et al. "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8TeV$". In: *JINST* 10.06 (2015), P06005. DOI: 10.1088/1748-0221/10/06/P06005.

[92] A. M. Sirunyan et al. "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV". In: *JINST* 13 (2018), P05011. DOI: 10.1088/1748-0221/13/05/P05011.

[93] M. Cacciari, G. P. Salam, and G. Soyez. "The anti-k_{T} jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063.

[94] M. Cacciari, G. P. Salam, and G. Soyez. "FastJet User Manual". In: *Eur. Phys. J. C* 72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2.

[95] A. M. Sirunyan et al. "Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ TeV". In: *JINST* 13.10 (2018), P10005. DOI: 10.1088/1748-0221/13/10/P10005.

[96] CMS Collaboration. *Performance of the DeepTau algorithm for the discrimination of taus against jets, electron, and muons*. CMS Detector Performance Summary CMS-DP-2019-033. 2019.

[97] A. M. Sirunyan et al. "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV". In: *JINST* 12 (2017), P02014. DOI: 10.1088/1748-0221/12/02/P02014.

[98] CMS Collaboration. *Jet algorithms performance in 13 TeV data*. CMS Physics Analysis Summary CMS-PAS-JME-16-003. 2017.

[99] D. Bertolini et al. "Pileup Per Particle Identification". In: *JHEP* 10 (2014), p. 059. DOI: 10.1007/JHEP10(2014)059.

[100] V. Khachatryan et al. "Search for Third-Generation Scalar Leptoquarks in the tτ Channel in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV". In: *JHEP* 07 (2015). [Erratum: JHEP 11, 056 (2016)], p. 042. DOI: 10.1007/JHEP11(2016)056.

[101] Serguei Chatrchyan et al. "Measurement of the Inclusive W and Z Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV". In: *JHEP* 10 (2011), p. 132. DOI: 10.1007/JHEP10(2011)132.

[102] J. Rembser. "CMS Electron and Photon Performance at 13 TeV". In: *J. Phys. Conf. Ser.* 1162.1 (2019), p. 012008. DOI: 10.1088/1742-6596/1162/1/012008.

[103] M. Cacciari et al. "The $t\bar{t}$ cross-section at 1.8 and 1.96 TeV: A study of the systematics due to parton densities and scale dependence". In: *JHEP* 04 (2004), p. 068. DOI: 10.1088/1126-6708/2004/04/068.

[104] R. Field. *Min-Bias and the Underlying Event at the LHC.* 2021. arXiv: 1202.0901 [hep-ph].

[105] S. Baker and R. D. Cousins. "Clarification of the use of CHI-square and likelihood functions in fits to histograms". In: *Nucl. Instrum. Meth.* 221 (1984), p. 437. DOI: 10.1016/0167-5087(84)90016-4.

[106] T. Junk. "Confidence level computation for combining searches with small statistics". In: *Nucl. Instrum. Meth.* 434 (1999), p. 435. DOI: 10.1016/S0168-9002(99)00498-2.

[107] A. L. Read. "Presentation of search results: The CL_s technique". In: *J. Phys. G* 28 (2002), p. 2693. DOI: 10.1088/0954-3899/28/10/313.

[108] G. Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. C* 71 (2011). [Erratum: Eur. Phys. J. C 73, 7 (2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0.

[109] ATLAS and CMS Collaborations. *Procedure for the LHC Higgs boson search combination in Summer 2011.* CMS Note CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11. 2011.