Bachelorarbeit

zur Erlangung des Hochschulgrades
Bachelor of Science
im Bachelor-Studiengang Physik

Eta-Correction for Data-Driven Estimate of Non-Prompt Background for Same Charged WW Scattering in the ATLAS Experiment

Christoph Peter Klimanek

geboren 12.10.1998 Datum in Dresden

Technische Universität Dresden
Institut für Kern- und Teilchenphysik
Fachrichtung Physik
Fakultät Mathematik und Naturwissenschaften

18.09.2020

Eingereicht am 15. September 2020

1. Gutachter: Prof. Dr. Michael Kobel
2. Gutachter: Dr. Frank Siegert
Betreuer: Max Stange und Dr. Carsten Bittrich

Zusammenfassung

Vektor-Bosonen-Streung (VBS) bietet die Möglichkeit die elektroschwache Symmetriebrechung zu untersuchen und die Vorhersagen des Standardmodells weiter zu überprüfen um mögliche Hinweise auf neue Physik jenseits des Standardmodells zu finden. Die $W^{\pm}W^{\pm}jj$-EWK Produktion ist ein vielversprechender Prozess um die VBS eingehender am LHC zu untersuchen. Jedoch ist dieser Signalprozess immer noch sehr anfällig für verschiedenste Untergründe. Einer dieser Untergründe resultiert aus Objekten, die fälschlicherweise als Leptonen des Signalprozesses gemessen werden, sogenannte 'non-prompte' oder auch 'fake' Leptonen.

In dieser Arbeit wird dieser fake Leptonen Untergrund in $W^{\pm}W^{\pm}jj$-EWK Produktion weiter untersucht und die Abschätzung von diesem Untergrund verbessert. Hierfür wird die 'fake factor'-Methode für zwei solche fake Leptonen produzierenden Prozesse $t\bar{t}$ und $W + jets$ verwendet. Dabei werden Daten aus Monte-Carlo Simulation mit zusätzlichen truth-level information verwendet, welche in Pythia und MadGraph generiert wurden. Die Prozesse wurden für Proton-Proton-Kollisionen im ATLAS Experiment mit einer Schwerpunktsenergie von $\sqrt{s} = 13$ TeV simuliert. Die integrierte Luminosität beträgt $138.7\,\text{fb}^{-1}$, was mit der Datenmenge des vollständigen run 2 vom LHC übereinstimmt. Es wird die Abhängigkeit des 'fake factors' von der Region des Detektors in welcher das fake Lepton rekonstruiert wurde, untersucht. Zur Beschreibung der Teilchenposition im Hinblick auf die Geometrie des Detektors wird die Pseudorapidität η gewählt. Es wird eine Kontrollregion erstellt, in welcher die 'fake factoren' in Abhängigkeit von p_{T} bzw. η berechnet werden. Darüberhinaus wird eine η abhängige Korrektur für den p_{T} abhängigen 'fake factor' für den $t\bar{t}$-Prozess erstellt. Die gewonnenen Ergebnisse werden durch einen Monte-Carlo Closure Test überprüft und die Qualität der Abschätzung des Untergrund zu beurteilt. Im Rahmen dieser Arbeit wurde festgestell, dass die zusätzliche Beachtung der Abhängigkeit des 'fake factor' von der Pseudorapidität η die Abschätzung des Untergrundes, für die untersuchten Monte-Carlo Simulation, signifikant verbessert. Diese Ergebnisse müssen jedoch noch für den Fall überprüft werden, dass ausschließlich reco-level Information für die Bestimmung der 'fake' Leptonen verwendet werden.

Abstract

Vector boson scattering (VBS) provides a unique way to study the electroweak symmetry breaking mechanism and further probe the predictions of the Standard Model to possibly discover any hint of new physics beyond the Standard Model. A very promising candidate to study such VBS processes at the LHC is the $W^{\pm}W^{\pm}jj$-EWK production. Although, it is still significantly susceptible to various background contributions. One of this background contribution originates from objects which are misidentified as signal leptons, denoted as non-prompt or fake leptons.

Thus this thesis further studies the fake lepton contribution in the $W^{\pm}W^{\pm}jj$-EWK production and aims to improve the fake lepton background estimation. Therefore a data driven fake factor study is conducted for two non-prompt background processes $t\bar{t}$ and $W_{+}jets$ using Monte Carlo simulations with truth-level information obtained from Pythia and MadGraph. The processes were simulated for pp-collision in the ATLAS experiment at a center of mass energy of $\sqrt{s} = 13$ TeV and an integrated luminosity of $138.7\,\text{fb}^{-1}$ in resemblance to the LHC full run 2 data. The dependency of the fake factor from the detector region where the object was reconstructed, represented by the pseudorapidity η, is studied. A control region is used to extract the fake factors binned in p_{T} respectively η. Furthermore an additional η-correction to the fake factor, that is calculated in bins of the transverse momentum, is derived for the $t\bar{t}$ process. A Monte Carlo Closure Test on the Control Region is employed to validate the quality of the fake lepton background estimation. It is found, that the inclusion of the η-dependency improves the fake lepton background estimation for the investigated Monte Carlo samples. This results remain to be validated using Monte Carlo samples with a lepton selection solely on reco-level.

Contents

Chapter 1

Introduction

The field of particle physics aims to understand the most fundamental particles making up our universe and their interactions. During the recent century a large number of new particles were discovered and a lot of theories proposed. Leading to the development of one comprehensive theory known as the Standard Model (SM) of Particle Physics. It has been extensively tested by various experiments and proven to be highly accurate - predicting new particles like the $W^{\pm}$ and Z-boson and the Higgs. The current formulation of the Standard Model is then assumed to be fully self-consistent by confirming the observation of a light SM Higgs at the Large Hadron Collider (LHC) in 2012. However there are still phenomena which can not be explained by today's Standard Model, suggesting the search for new physics beyond the Standard Model.

A candidate, that could possibly provide further insight, is the process of vector-boson scattering (VBS). In a VBS event two electroweak gauge bosons couple onto the same interaction vertex and scatter from each other $VV \rightarrow VV$ ($V = W, Z, \gamma^{1}$) . This is possible because of the underlying gauge symmetry of the electroweak (EWK) sector. According to the SM prediction this VBS processes are unitarized due to contributions of the Higgs boson at energies of 1 TeV and higher. Thus it provides a way to further examine the mechanism of electroweak symmetry breaking and search for new physics, by probing whether the scattering amplitude is fully or only partially unitarized by the light SM Higgs.

Amongst the VBS processes, the scattering of two same sign W bosons (WWss) is the process with the highest cross section, which is still relatively small in comparison to other processes observable at the LHC. Since the $W^{\pm}W^{\pm}$ scattering process is observed in a proton proton collision its topology consists of two quarks radiating a $W^{\pm}$ each, which then scatter and decay afterwards. The specific event topology and the consideration of only leptonic decays of the W for the analysis, significantly reduces SM backgrounds when compared to other VBS processes. Thus making the $W^{\pm}W^{\pm}jj$-EWK production one of the best channels for further VBS studies [1].

However there are multiple processes that can produce the same final state as the $W^{\pm}W^{\pm}jj$-EWK process, making it still susceptible to various background contributions. A proper background estimation is therefore very important to increase the signal to background ratio. Apart from prompt background producing events like $WZjj$ a significant contribution originates from non-prompt events. The non-prompt background arises from objects being misidentified as signal leptons. A data driven method is employed to better estimate this background, because Monte Carlo simulations are not expected to model this type of background very accurately. The aim of this thesis is to further improve the fake lepton background estimation for WWss scattering by studying its dependency of the detector region. Thus the fake factor method is applied to Monte Carlo samples with an additional truth level assignment for the $t\bar{t}$ and $W + jets$ background contributing processes .Furthermore an η-correction to the p_{T} binned fake factor is derived, as proposed and shown in Ref. [2, 3].

[1]Note that the process $\gamma\gamma \rightarrow \gamma\gamma$ is not possible, because of the U(1) gauge symmetry of QED. Hence there is no self-interaction between photons.

This thesis is structured as follows. Chapter 2 provides a brief overview about the Standard Model, especially the electroweak sector and explains the signal and background contribution in more detail. The basic geometry and composition of the ATLAS detector is presented in chapter 3. Monte Carlo simulations are briefly explained in chapter 4 as well as one of the main topics of this thesis the fake factor method. In Chapter 5 separate fake factor studies are performed for the faking leptons transverse momentum p_{T} and the detector region, denoted by the pseudorapidity η, providing a first impression on the fake lepton background. The respective η-correction is then derived and the results are evaluated by a Monte Carlo Closure Test in Chapter 6 and . The final chapter 7 concludes this thesis and provides an outlook on possible future analysis.

Chapter 2

Theory

2.1 Standard Model of Particle Physics

The Standard Model is one of the most successful, experimentally tested and accurate theories in physics up to this date. There are four known fundamental interactions between particles - the gravitational, electromagnetic, strong and weak interaction. The latter three are unified in one theory describing the dynamics and kinematics of all elementary constituents of matter.

Therefore the Standard Model is formulated in the framework of Quantum-Field-Theory (QFT), which is based on the principles of quantum mechanics and special relativity. It can be observed that the fundamental interactions satisfy certain symmetries, i.e. U(1) - electromagnetic interaction. Following the approach of Yang and Mills the mathematical expression is then obtained from the algebra of the particular symmetry group [4]. The transformations which conserve the local gauge symmetries of these interactions are denoted as local gauge transformations.

According to the Noether-Theorem, a conservation law exists to every continuous symmetry and vice versa. Therefore any of the fundamental interactions is associated with the conservation of certain quantum numbers, which are called charges (i.e: electric charge - electromagnetic interaction; color charge - strong interaction). Any of the particles that are able to partake in the interaction are then required to have a respective charge, that is not zero.

In Gauge Theories the most compact way to mathematically express them is by a Lagrangian density $\mathcal{L}$. The Standard Model Lagrangian density is explicitly lorentz invariant, which ensures that fundamental laws like invariance of the speed of light, momentum and energy conservation are valid in every reference frame. Furthermore it reflects on the gauge symmetry of the fundamental interaction. This symmetry is conserved by introducing new fields to the theory, called gauge fields, which compensate the effect of local gauge transformations and thus make the Lagrangian gauge invariant.[2, 3, 5]

Table 2.1: Overview on elementary particles of the Standard Model. Included fermions are assumed to be left-chiral. Also each fermion has an associated anti-particles with the same rest mass, but opposing charge, and are not listed here. Data taken from [5]

Particles	Quarks		Leptons		Gauge bosons			
	u,c,t	d,s,b	e,μ,τ	ν_e,ν_μ,ν_τ	$W^\pm$	Z	γ	g
electric charge Q [e]	$+\frac{2}{3}$	$-\frac{1}{3}$	0	0	±1	0	0	0
weak Isospin I_w^3	$+\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{2}$	$\pm$ 1	0	0	0
spin S	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	1	1	1

2.2 Standard Model Particles

The classical concept of particles is no longer valid in the scope of QFT. A more accurate depiction would be, that every particle is an excitation of the corresponding particles field and their interactions are then also mediated by particles, denoted as gauge bosons, which arise from their respective gauge-fields. One can distinguish between two fundamental types of particles - fermions with half-integer spin and gauge-bosons with integer spin. The former ones are the elementary constituents of matter, i.e. quarks make up protons and neutrons. While the latter ones are the exchange particles which mediate the interactions between the fermions, i.e. Quantum electrodynamics (QED) - photon γ .

As aforementioned all particles, which make up the Standard Model, can be characterized by their charges. The fermions can be divided in several subgroups - quarks, leptons. All quarks carry a color charge and are either denoted as "up-" with Q $= +1/3$ or "down"-type with Q $= -2/3$. Thereby those are the only fermions which take part in all three interactions. On the other hand leptons can be split into charged leptons, like electrons, and neutrinos as neutral leptons. Both do not posses any color charge, but an electric charge of Q $= \pm 1$ or 0 and therefore only participate in the electromagnetic and/or weak interaction (2.1).

Also there is a corresponding anti-particle to every particle in the Standard Model. Anti-particles possess the same rest mass as their particle counterparts but are charge conjugated. Furthermore it has been found that there exist three generations of fermions. The first generation consists of the stable particles making up our matter, while the other generations consist of identical but heavier, unstable copies of their first generation counterparts. [5]

2.3 Electroweak Theory

Electroweak Theory was developed by Glashow, Salam and Weinberg (1960s) and unifies the electromagnetic and weak interaction. It is described by a $SU(2)_L \times U(1)_Y$ local gauge theory with the corresponding conserved charges being the weak hypercharge $Y = Q - I_w^3$, electric charge Q and third component of the weak isospin I_w^3. The electroweak gauge bosons are then the charged $W^\pm$ boson, neutral Z boson and the photon γ.
According to gauge theory any gauge boson is assumed to be massless, because an explicit mass term would violate the local gauge invariance of the Lagrangian density or rather break the local gauge symmetry. Though experimental results prove that $W^\pm$ and Z boson are in fact massive particles.

This controversy is solved by introducing the concept of electroweak symmetry breaking (EWSB). The local gauge symmetry of the electroweak interaction is spontaneously broken by introducing a new symmetric field where the vacuum energy state is not invariant under symmetry operations - the Higgs field. Any Standard Model particle would then gain its mass by interacting with the Higgs field. Consequently the $W^\pm$ and Z boson gain its mass from the Higgs mechanism.

Another important observation is the phenomenon of triple and quartic gauge couplings. Since the $SU(2)_L \times U(1)_Y$ is non-abelian, its generators do not commute amongst each other. As a consequence the gauge fields in the Lagrangian do not commute as well and therefore higher orders terms of the gauge coupling constant α arise, where three and four gauge bosons couple to the same vertex. The strength of these gauge couplings are fully predicted by the theory and can be measured via the corresponding cross section. Any deviation from the Standard Model prediction would then manifest in the measured cross section and would hint towards new physics beyond the Standard Model. [2, 3, 5, 6]

2.4 Vector Boson Scattering

Vector Boson Scattering (VBS) is one of the electroweak processes measured at the LHC. In VBS, two vector bosons ($V = W,\ Z,\ \gamma$) interact with each other and then decay afterwards, thus it is called scattering $VV \rightarrow VV$. There are multiple channels possible, which contribute to this scattering process. According to the Standard Model predictions, VBS is either observed via quartic gauge coupling or the exchange of a γ, W, Z gauge boson or a Higgs boson, as depicted in Figure 2.1.

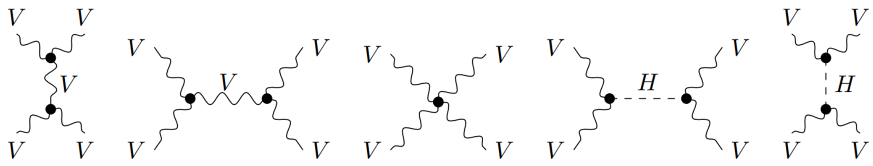


Figure 2.1: Feynman diagrams for all possible VBS channels (V=W, Z, γ). Only specific initial and final states, that do not violate fundamental conservation laws, like electric charge conservation, are allowed in the SM. Figure from [3]

Since the LHC is a proton-proton collider the initial vector bosons are emitted from two quarks in the proton-proton (pp) collision. Therefore, the VBS process is always accompanied by two jets, the tagging jets, which are measured closely to the beam axis but in opposite directions. Any of the VBS channels then requires six weak interaction vertices ($\mathcal{O}(\alpha_w^6)$), which leads to a very small cross section compared to other processes at the LHC. Because of that, the study of VBS requires large samples of data.

Vector boson scattering is an important process to study the electroweak symmetry breaking further and thus probe the SM electroweak theory. The contribution of a light SM Higgs Boson regularizes the amplitude of this process at high energies. In the absence of such, the amplitude would diverge, because the divergences arising from longitudinally polarized vector bosons at high energies are not contained. This would violate the unitarity, consequently resulting in a VV production higher than the pp collision rate. Any deviations from the precise SM predictions would then hint towards new BSM physics like modifications of the Higgs sector or additional resonances occurring in VV scattering.[1, 7, 8]

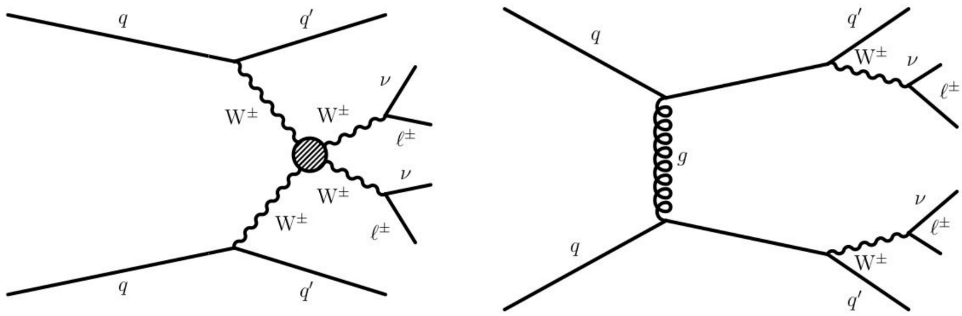
2.5 WWss Scattering Signal: $W^{\pm}W^{\pm} \rightarrow \ell^{\pm}\ell^{\pm}\nu\nu\ jj$

There are several VBS processes that are being studied at the LHC. The same charge WW and WZ, ZZ as well as $Z\gamma$ were already the subject of recent publications, reaching the observation level for the first time [2],[3]. This thesis focuses solely on the $W^{\pm}W^{\pm}$ scattering process. Among the possible VBS processes, the $W^{\pm}W^{\pm} \rightarrow W^{\pm}W^{\pm}$ has the largest electroweak (see Fig. 2.2a) over strong production (see Fig. 2.2b) cross-section ratio compared to other VBS processes [7].

In this process, two same sign W bosons are emitted by the two colliding quarks and then interact with each other (Section 2.4). The experimental signature of this process consists of the two tagging jets, as well as the decay products of the two W bosons, since they have a very short lifetime ($t \approx 10^{-25}s$). In leptonic decays, two lepton-neutrino pairs $W \rightarrow \ell\nu$ with the same electrical charge are produced. Although neutrinos are produced, they can not be measured in the detector. They are however identified during event reconstruction as missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$, because of energy and momentum conservation in the transverse plane (see Section 3.2).

Hadronic decays of the W$^{\pm}$ are excluded in this thesis, despite their approximately three times larger branching fractions. The reason being, that the event reconstruction is very complicated, due to the additional hadronization and pile up occuring in pp-collisions of the LHC. Furthermore, decays containing the short lived τ lepton are neglected, because it can also decay hadronically.

Any particular final state, of a particle interaction is described by an infinite number of Feynman diagrams, which contribute to the total probability amplitude. As those diagrams basically describe perturbations in terms of increasing order, it is reasonable to only consider the leading order diagrams. The leading order for any VBS process is then $\mathcal{O}(\alpha_w^6)$, as shown in Section 2.4 . Hence the final signal for W$^{\pm}W^{\pm} \rightarrow l^{\pm}l^{\pm}\nu\nu jj$ is described by the following Feynman diagram depicted in Figure 2.2. A specific signal region is then chosen for the measurement of this process according to the VBS signal topology (see Table B.1 in Appendix B).



(a) Representative Feynman diagram for $W^{\pm}W^{\pm}$ electroweak production process.

(b) Feynman Diagram for one possible $W^{\pm}W^{\pm}$ strong production process with QCD vertex from quark scattering mediated by a gluon (g).

Figure 2.2: Representative Feynman diagrams for production of same charged W boson pair in a pp collision at the LHC, decaying to two lepton neutrino pairs ($\ell\nu$) and accompanied by two quarks (q) detected as forward jets. The Dashed circle in (a) stands for every permitted VBS boson scattering diagrams shown in 2.1. New physics in the electroweak sector could modify this interaction. Figures from [9]

However, there are non-VBS processes, that can not be distinguished from the desired signal. Since one can not gauge invariantly separate those VBS and non-VBS diagrams, they have to be included in the definition of the signal as well and can not be treated as background (2.3).[1, 3, 7, 8]

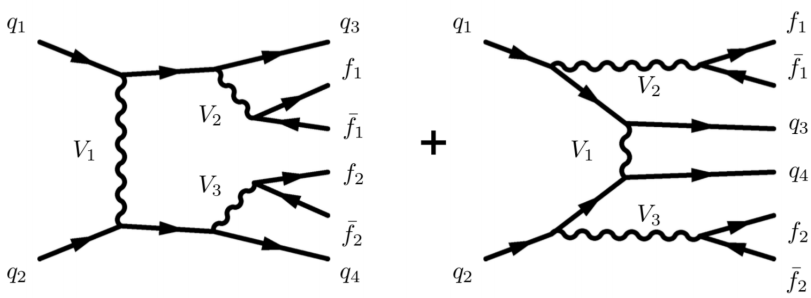


Figure 2.3: Representative Feynman diagrams for $VVjj$-electroweak production without VBS topology. The lines are labeled by quarks (q), vector bosons ($V = W, Z, \gamma$) and fermions (f). Figure from [1]

.

2.6 WWss Scattering Background

The term background in particle physics data analysis denotes any physical process, that contributes to your observed data, but is not part of your signal and can be distinguished in a gauge invariant way. Thus they yield a similar final state of particles and need to be treated thoroughly to identify the desired signal in the data.
In this thesis, background processes are either specified as prompt or non-prompt, which depends upon some criteria imposed on the final state particles. The prompt background will be elaborated briefly, whilst the Non-Prompt background is of main interest. For the process studied in this work such particle classification is only conducted for the final state leptons.

If the two leptons are found to be originating from a hard collision, they are called prompt. This process is then classified as prompt background or irreducible background. For WWss, the irreducible background mostly consists of $WZjj$-QCD and $W^{\pm}W^{\pm}jj$-QCD, which are both strong production processes in the order of $(\mathcal{O}(\alpha_s^2\alpha_w^4))$. Other processes contributing, are $WZjj$-EW and triboson processes, for example WWZ or WWW with a boson decaying hadronically. The $WZjj$ background in particular have to be considered, because it can be mistakenly measured as a signal event. This can happen when the opposite sign lepton originating from the Z decay is lost during reconstruction or fails the specific lepton identification requirements.[1, 3]

Contrary to the irreducible background that do contain two prompt leptons, there are backgrounds, called non-prompt, where an object is either falsely identified as a lepton or is not a lepton coming from a hard collision. Typical examples are, a jet faking a lepton and leptons, that originate from a secondary vertex or e/γ conversions. This kind of background often can be suppressed by imposing certain event selection criteria, thus also calling them reducible backgrounds. [3] Apart from the irreducible background the non-prompt lepton background has the second biggest background contribution (Ref. [1, 10, 11]) to the $W^{\pm}W^{\pm}jj$-EW signal. It is mainly produced by processes which have a very large cross section compared to the signal ($\approx 10^5$ times), like QCD-multijets, single t, $W + jets$ and $t\bar{t}$ production. The latter two are studied further in this work and also depicted in the following Feynman diagram (see Fig.2.6).
During the $W + jet$ process a W and a quark are produced. The W decays leptonic, while the quark is measured as a jet. It is possible that this respective jet is misidentified by the detector as a lepton. Thus the jet is then detected as second lepton, leading to the incorrect measurement of a WWss scattering event. In the $t\bar{t}$ process actually two opposite charged W are produced, leading could decay semileptonic or leptonic. In case of the semileptonic $t\bar{t}$ process one W decays leptonic, while the other one decays hadronically, resulting in a jet that would fake the second lepton. On the other hand, in a fully leptonic process two opposite charged leptons are produced. Thus one of the leptons needs to misidentified as an opposite charge lepton. Such events, denoted as chargeflips, are excluded from the contribution to the non-prompt background in this thesis.

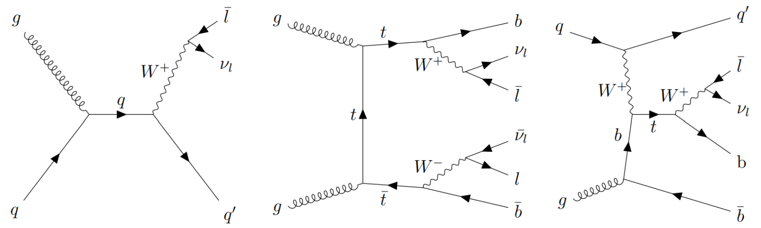


Figure 2.4: Feynman diagramms for the $W + jets$ (left), $t\bar{t}$ (middle) and single t (right) processes, that contribute to non-prompt background in WWss scattering. Curled lines are gluons (g), waved lines W bosons and straight lines the respective fermions. Figure from [10]

Chapter 3

Experiment

3.1 Large Hadron Collider

The Large Hadron collider (LHC) is the worlds largest particle accelerator and circular collider. It was build by the European Organization for Nuclear Research (CERN) to solve some of the remaining questions in the SM like the mechanism of EWSB and explore the field of high enery physics further. The collider is located near Geneva, Switzerland, in the same tunnel as the previous Large Electron-Positron Collider (LEP), 100 m below the surface.

Proton bunches, with approximately 10^{11} protons, are accelerated inside of two opposite circular beam pipes with a circumference of about 26.7 km. At the locations of the experiments the bunches "cross over" and proton-proton collisions with a current center of mass energy of $\sqrt{s} = 13$ TeV are measured. There are two multi purpose detectors which work independently - A Toroidal LHC ApparatuS (ATLAS) and Compact Muon Solenoid (CMS). They are designed to take high luminosity data and are now being upgraded and gradually prepared for the High-Luminosity upgrade, before the LHC run 3 starts.

In particle physics the Luminosity L is an important quantity to characterize the performance of a particle accelerator, since it describes the detected number of events per time and the respective cross section. If the luminosity is integrated with respect to time one obtains the integrated Luminosity, which then simply is the ratio between the observed number of events and the cross section. Therefore the integrated Luminosity basically display how much data is available. At the end of Run 2, both ATLAS and CMS had measured an integrated luminosity of around $140\,\mathrm{pb}^{-1}$ (Ref. [12]).

3.2 ATLAS Detector

ATLAS is a multi purpose detector, which was designed for precision measurement of Standard Model predictions including the search for new physics.

To properly describe a collision event, observed at the ATLAS detector, a specific coordinate system has to be introduced first. The collision point is chosen as the origin. The x-axis is pointing from the collision point towards the center of the LHC and the y-axis points upwards. Therefore the x-y plane is transverse to the beam axis. The z-axis then is parallel to the beam axis, while its direction is determined by forming a Cartesian right-handed coordinate system.

It is convenient to introduce additional variables to account for the cylindrical/spherical geometry as well. The azimuthal angle Φ is measured from in the x-y plane starting from the x-axis, whilst the polar angle Θ is the angle with respect to the z-axis. A depiction can be found in Figure 3.1.

Since the polar angle Θ is not invariant under Lorentz transformations it is rarely used. However the pseudorapidity η as a new angular variable is much more convenient, because it

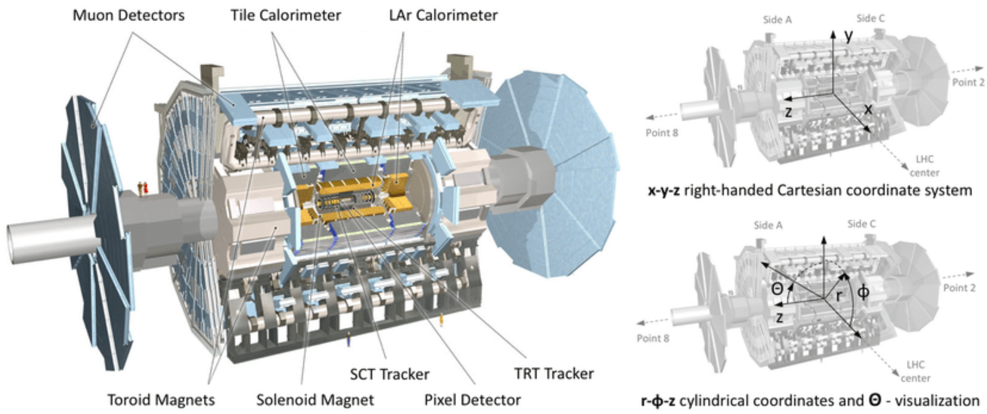


Figure 3.1: Overview over the ATLAS detector with the most important detector components labeled. On the right side, a depiction of the two possible coordinate systems, with respect to the detector geometry, is given. Figure from [14]

can be used to define some quantities like angular separation, that are then invariant under Lorentz boosts along the beam axis.

$$\eta = -\ln \left[\tan(\frac{\Theta}{2}) \right] \tag{3.1}$$

In a collision event, hundreds of particles are produced and then measured and identified by the detector. Therefore the detector actually consists of multiple subdetectors, which are arranged in a layout similar to an onion. The most inner detector the Tracker surrounds the collision point, with a coverage of $|\eta| < 2.5$. It is designed to track trajectories and momentum of charged particles. Particle tracks are then reconstructed up to their interaction vertex using pattern recognition and data from the tracking detector with high granularity - pixel detector, silicon microstrip detector (both semiconductor based) and transition radiation tracker (gaseous detector). Since the detectors are immersed in a strong magnetic field, the sign of the electric charge and the particle momentum can be determined from the bending radius and direction of the particle inside the magnetic field.

Next are the electromagnetic (ECAL) and hadronic calorimeter (HCAL), that measure the energy of the particles and cover an area of $|\eta| < 3.2$ and $|\eta| < 4.9$ respectively. The ATLAS ECAL is a lead (absorber) and liquid Argon (active material) based sampling calorimeter in the barrel and end-cap region of the detector, enclosing the Tracker. Electromagnetic particles like photons and electrons interact with the material creating a particle shower and depositing all of their energy in the calorimeter.
Hadrons and other particles not stopped in the ECAL deposit their energy in the surrounding HCAL, except for muons. It is an additional sampling calorimeter consisting of steel (absorber) and scintillator tiles (active material) which covers the barrel area up to very forward regions inside the end caps.

The outermost detector layer is the Muon Spectrometer, which is build to measure the momenta and trajectories of incoming muons in a region of $|\eta| < 2.7$. A muon track is measured by the detector using the ionization of particles in gaseous chambers and the momentum is determined due to a magnetic field bending the respective trajectory.

Neutrinos are the only particles which can not be measured by the ATLAS detector, because they interact exclusively through weak interaction and have a very small cross section. However they are still perceived through missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$, because the total momentum in the transverse plane should be zero according to energy-momentum conservation [2, 13].

Chapter 4

Data Analysis

4.1 Monte Carlo Simulation

In the following Section only the basic concept of Monte Carlo (MC) simulations will be explained. A deeper discussion with much more details is presented in Ref. [15].

Accurate and reliable simulations with associated uncertainties are an essential component of analysis in experimental particle physics. Monte Carlo simulation utilizes random numbers for repeated sampling to obtain the statistical properties of a certain process with respect to the probability distribution predicted by the considered theory.
Using factorization the complexity of calculating a particle collision event is reduced by separating the processes into multiple regimes like low and high energy. The full simulation, from particle to detector level, then consists of multiple steps.

First the processes of the collision event are simulated by a specialized MC Generator. Therefore the matrix elements of the corresponding Feynman diagrams are calculated according to the Feynman rules. The cross section is now obtained from integrating the matrix element over the phase space using the Monte Carlo method. Thus the integration is done numerically by summing over a sufficiently large sample set of random points in the phase space, instead of every single point. In the MC method this sample set is generated in a way, that the density distribution of the points is representative of the density distribution implied by the theory of this process. For instance, the initial four momentum of a proton-proton collision in the LHC is estimated by applying the parton density function (PDF), but for a full modeling a multitude of other particle properties have to be considered as well.
The generated particles contain all information related to the physics of the collision event without any additional effects i.e. detector physics. Hence they are also referred to as particles on truth-level.

Next the particle interaction with the detector is simulated, because a detector can not measure particles on truth level. In the final step the same reconstruction algorithms, also used for actual data, are applied to obtain simulated events on reconstruction (reco) level, which can then be compared to experimental data [3, 16].

MC simulations are used for example to estimate irreducible backgrounds or lepton identification efficiencies, where detector deficiencies are accounted for by deriving a respective scale factor. For the WWss scattering, background estimation of processes like triboson processes, $W^{\pm}W^{\pm}jj$-QCD are performed by using MC simulations, whereas for $WZjj$ only the shape is taken from MC and normalization is derived from data taken from the corresponding control region [11].

The utilization of MC simulations relies on the assumption that the modelling and obtained estimations are accurate and reliable. However background processes that are not very well modelled by MC simulations, would result in non-reliable a background prediction, if estimated from MC data. Background estimations, that do not rely on MC simulation, are explained further in the next Section.

4.2 Datadriven Background Estimation using Fake Factor Method

As explained in Section 2.4 the non-prompt lepton background, in the following denoted as fake lepton background, originates from particles being misidentified as leptons. The misidentified leptons arise either from decays of hadrons with b or c-quark content (heavy flavour) or from misidentified light flavour jets. These misidentified leptons are produced in the jet fragmentation process and thus will be spatially correlated with jets. The respective jet fragmentation is extremely difficult to model, and since the examined events, that pass the selection criteria, are rather rare, they are not expected to be well simulated by MC. Such fake backgrounds are then estimated using data driven methods, where control regions in data are used to model the background, and transfer respective factors to extrapolate from the control regions into the signal region. In this thesis an estimation of the fake lepton background for WWss is obtained by the Fake Factor method (Ref. [2, 17]).

Since measured data only have particle information on reco-level it is not possible to tell whether the selected object is fake or indeed a lepton. Thus a similar but adjusted differentiation has to be introduced for the respective reco level objects. For this thesis the two object level selections are defined as (see Tables B.2 and B.2):

- Analysis (Ana) leptons ℓ_A: objects passing the Analysis selection and thus would contribute to the number of actual signal leptons. However they are not inevitably real leptons, but could also be hadronic jets or leptons originating from secondary vertices or photon conversions.
- Non-Analysis ("Nonana") leptons ℓ_N: objects failing at the Analysis selection, but still pass the looser Non-Analysis selection. Thus they are assumed to be more likely non-prompt leptons.

An additional selection is performed on truth-level utilizing the object type information obtained from the IFFTruthClassifier[1]. It is important to note that a truth-level classification is only provided by MC-simulations, where as this is not possible for actual data.

- Prompt lepton: An electron or muon originating from the decay of a Wboson.
- Fake lepton: Objects, than are misidentified as lepton during the signal selection.

The selected reco-level objects can be related to the truth-level-objects by introducing a probability of passing or failing the respective selection. Since the final state of WWss consists of two leptons, the events contributing to the fake lepton background have two leptons as well. Thus such a probability needs to be assigned for each lepton (index i∈{1,2}).

- e_i: probability that a Prompt lepton gets Ana selected
- $\bar{e}_i = 1 - e_i$: probability that a Prompt lepton fails the Ana selection and is selected as Nonana
- f_i: probability that a Fake lepton gets Ana selected
- $\bar{f}_i = 1 - f_i$: probability that a Fake lepton fails the Ana selection and is selected as Nonana

Using matrix notation the number of events $N_{m,n}$, with m,n ∈ {A,N}, passing a certain combination of selection criteria on reco-level can be expressed in terms of the number of events $N_{k,l}$, with k,l ∈ {P,F}, on truth-level, that consist of a certain lepton combination. For

[1]A tool (algorithm) for classifying electrons and muons based on based on their type and origin.

instance, the number of events corresponding to the observed lepton signal is then denoted by N_{AA}, because both are required to pass the Analysis selection.

$$\begin{pmatrix} N_{AA} \\ N_{AN} \\ N_{NA} \\ N_{NN} \end{pmatrix} = \begin{pmatrix} e_1 e_2 & e_1 f_2 & f_1 e_2 & f_1 f_2 \\ e_1 \bar{e_2} & e_1 \bar{f_2} & f_1 \bar{e_2} & f_1 \bar{f_2} \\ \bar{e_1} e_2 & \bar{e_1} f_2 & \bar{f_1} e_2 & \bar{f_1} f_2 \\ \bar{e_1}\bar{e_2} & \bar{e_1}\bar{f_2} & \bar{f_1}\bar{e_2} & \bar{f_1}\bar{f_2} \end{pmatrix} \cdot \begin{pmatrix} N_{PP} \\ N_{PF} \\ N_{FP} \\ N_{FF} \end{pmatrix} \quad . \tag{4.1}$$

This equation can be rearranged by multiplying each row with a suitable term:

$$\begin{pmatrix} N_{AA} \\ -N_{AN} \cdot \frac{f_2}{\bar{f_2}} \\ -N_{NA} \cdot \frac{f_1}{\bar{f_1}} \\ N_{NN} \cdot \frac{f_1}{\bar{f_1}} \frac{f_2}{\bar{f_2}} \end{pmatrix} = \begin{pmatrix} e_1 e_2 & e_1 f_2 & f_1 e_2 & f_1 f_2 \\ -e_1 \bar{e_2} \cdot \frac{f_2}{\bar{f_2}} & -e_1 \bar{f_2} \cdot \frac{f_2}{\bar{f_2}} & -f_1 \bar{e_2} \cdot \frac{f_2}{\bar{f_2}} & -f_1 \bar{f_2} \cdot \frac{f_2}{\bar{f_2}} \\ -\bar{e_1} e_2 \cdot \frac{f_1}{\bar{f_1}} & -\bar{e_1} f_2 \cdot \frac{f_1}{\bar{f_1}} & -\bar{f_1} e_2 \cdot \frac{f_1}{\bar{f_1}} & -\bar{f_1} f_2 \cdot \frac{f_1}{\bar{f_1}} \\ \bar{e_1}\bar{e_2} \cdot \frac{f_1}{\bar{f_1}} \frac{f_2}{\bar{f_2}} & \bar{e_1}\bar{f_2} \cdot \frac{f_1}{\bar{f_1}} \frac{f_2}{\bar{f_2}} & \bar{f_1}\bar{e_2} \cdot \frac{f_1}{\bar{f_1}} \frac{f_2}{\bar{f_2}} & \bar{f_1}\bar{f_2} \cdot \frac{f_1}{\bar{f_1}} \frac{f_2}{\bar{f_2}} \end{pmatrix} \cdot \begin{pmatrix} N_{PP} \\ N_{PF} \\ N_{FP} \\ N_{FF} \end{pmatrix} \quad . \tag{4.2}$$

Adding the rows (1), (2), (3) and (4), Equation 4.3 is obtained.

$$\begin{aligned} N_{AA} - N_{AN} \cdot \frac{f_2}{\bar{f_2}} - N_{NA} \cdot \frac{f_1}{\bar{f_1}} - N_{NN} \cdot \frac{f_1}{\bar{f_1}} \frac{f_2}{\bar{f_2}} &= \left(e_1 e_2 - e_1 \bar{e_2} \cdot \frac{f_2}{\bar{f_2}} - \bar{e_1} e_2 \cdot \frac{f_1}{\bar{f_1}} \right) \cdot N_{PP} \\ &+ \bar{e_1}\bar{e_2} \cdot \frac{f_1}{\bar{f_1}} \frac{f_2}{\bar{f_2}} \cdot N_{PP} \quad . \end{aligned} \tag{4.3}$$

The number of fake leptons can be calculated by rearranging the terms of Equation 4.3.

$$\begin{aligned} N_{fake} = N_{AA} - e_1 e_2 N_{PP} &= (N_{AN} - e_1 \bar{e_2} N_{PP}) \cdot \frac{f_2}{\bar{f_2}} + (N_{NA} - \bar{e_1} e_2 N_{PP}) \cdot \frac{f_1}{\bar{f_1}} \\ &- (N_{NN} - \bar{e_1}\bar{e_2} N_{PP}) \cdot \frac{f_1}{\bar{f_1}} \frac{f_2}{\bar{f_2}} \quad . \end{aligned} \tag{4.4}$$

By identifying the terms consisting of $N_{m,n} - a \cdot N_{PP}$, where a is the product of e and $\bar{e}$, as the respective reducible background $N^{red}_{m,n}$ contribution to the fake lepton events, Equation 4.4 can be simplified further. For instance the term $e_1 \bar{e_2} N_{PP}$ resembles the number of prompt lepton events, where both are Ana selected. Thus this would just be the prompt background contribution for this specific combination of Ana/Nonana selections, which can be calculated using MC simulation. The remaining number of N_{AA} events after subtracting the number of prompt events indeed corresponds to the fake background, as defined in Section 2.6.

Note that the contribution of double fake events, which corresponds to the product of F_1 and F_2 is neglected for this thesis (see Eq. 4.5). The reason being that, the contribution of double fake events to the total fake lepton background is expected to be sufficiently small.

Besides that, a Fake Factor F_i is defined for every lepton, which represents the ratio of a fake lepton being Ana selected over it being Nonana selected. So far no specific enumeration is chosen for the first and second lepton. Thus in general the Fake Factors would depend upon that.
It is also important to acknowledge, that the fake factor is also lepton flavour dependent, because muons mainly originate from in-flight decays of pions and kaons, or heavy-flavor quark (b or c) decays, while electron fakes are mainly a jet is misidentified as an electron or originate from photon conversions Therefore separate electron and muon fake factors have to be calculated.

$$N_{fake} \approx N^{red}_{AN} \cdot F_2 + N^{red}_{NA} \cdot F_1 + \mathcal{O}(F^2) \quad \text{with} \quad F_i = \frac{f_i}{\bar{f_i}} \quad . \tag{4.5}$$

A convenient way to enumerate the leptons is to order them according to their p_T value with the first lepton being the leading and the second one being the subleading. Since the object selection criteria are the same for both, the leading and subleading, the respective fake factors are the same $F_1 = F_2$.

Calculating the fake factors using data from the signal region and then applying it there would be circular. To properly estimate them, a control Region is defined, which is enriched in fake lepton producing events and has as little contamination as possible from real leptons. Furthermore this control region needs to be orthogonal to the signal region, while the composition of fake events has to be similar in both. If these conditions are satisfied the fake factor can be assumed to be similar $F_{CR} \approx F_{SR}$ and the calculated fake factor F_{CR} can be applied in the signal Region to estimate the fake background contribution.

The Fake Factor in the Control Region can then be calculated according to Equation 4.5. Thus the probabilities are just the number of fake events passing or failing the Analysis selection per total number of events in the Control Region N_{ctrl}. More details on the definition of the Control Region and the Ana/Nonana-selection are given in Section 5.1.

$$F = \frac{f}{\bar{f}} = \frac{N_{AA} - N_{AA}^{Prompt,MC}}{N_{ctrl}} \cdot \left(\frac{N_{AN} - N_{AN}^{Prompt,MC}}{N_{ctrl}} \right)^{-1} = \frac{N_{AA} - N_{AA}^{Prompt,MC}}{N_{AN} - N_{AN}^{Prompt,MC}} . \quad (4.6)$$

The probability of having misidentified leptons, depends on the transverse momentum of the faking object, the detector response, the occupancy of the detector among others. To capture such dependencies it is possible to calculate fake factors of Equation 4.6 in bins of p_{T} or η.

The number of events in the Control Region producing prompt leptons $N_{AN}^{Prompt,MC}$ are estimated using Monte Carlo simulation. This contribution needs to be subtracted from the respective total number of events, because such prompt events do not represent actual fake lepton events.

Since this study is conducted using truth-level information obtained from the examined MC samples, Equation 4.6 actually only depends on the number of events, where the fake lepton is Ana N_A or Nonana N_N selected. The lepton then are assigned according to the provided truth level information, because on truth level one lepton is always known to be fake while the other has to be prompt then. However it is important to point out, that this is no longer the case, when using data with no truth level information. Thus a proper assignment like Tag and Probe leptons has to be chosen, but will not be discussed any further because it would be beyond the scope of this study [18].

Chapter 5

Fake Factor Extraction for $W + jets$ and $t\bar{t}$ Process

In the following chapter the dependencies of the fake factor from the fake leptons transverse momentum and the detector region, where it is reconstructed, are investigated for the $t\bar{t}$ and W+*jets* process. Both contribute significantly to the fake lepton background in WWss scattering [11]. During both processes a W and jet can be produced, which are then misidentified by the detector as a secondary lepton. The fake factors are calculated in a control region using Monte Carlo samples, that are generated in MadGraph and Pythia. A full sample list is included in Appendix A. These samples are simulated for pp collision at the ATLAS detector with a $\sqrt{13\text{ TeV}}$ and an integrated luminosity of 138.7fb^{-1}. This thesis uses a lepton selection for the fake factor calculation, which is based on additional truth level information. The evaluation and calculation was done on the CommonAnalysisFramework (CAF).

5.1 Definition of Control Region and Object selection

The fake factor method is based on the proper estimation of the probabilities of a fake lepton passing or failing a certain object level selection. Thus, the fake factor has to be calculated in a control region enriched with fake lepton events from similar origin as the signal region, while both regions remain disjoint. Therefore, a two lepton control region is chosen, because it can be adjusted to be kinematically closer to the WWss signal region B.1. Only a brief overview on the definition of the WWss control region and the Ana/Nonana-selection is provided in this section. A more detailed and in depth discussion on this topic is presented in Ref. [18].

The cuts that are applied to define a suitable control region are shown in Table 5.1. For the fake lepton background estimation, this is supposed to consist only of events where exactly two Ana and/or Nonana leptons are observed with at least one of them being Ana selected. Furthermore, events, where additional leptons pass the Veto lepton selection (see Tables B.3,B.2) are rejected, to lower the prompt contribution from WZ events. Additionally, the two leptons are required to have the same charge to decrease the prompt contribution from $Z + jets$ events.

Events which fail at any cut, that defines the desired signal region, are then included in the control region. Thus making the Control Region fully disjoint to this region, while retaining a high contribution from non-prompt lepton events.

However, the remaining control region is dominated by chargeflip leptons. Leptons denoted as chargeflip are either prompt leptons with a misidentified charge during reconstruction, or originate from an oppositely charged prompt lepton. Such chargeflip events, "mostly" electrons, are explicitly distinguished from the fake lepton background are not to be included in the desired control region. To decrease their contribution and adjust the fake lepton composition to be more similar to the signal region, an ECIDS [1] cut is applied. Contribution of charge flips in the ee-channel, mostly originating from Z decays, are reduced further by applying a m_Z window cut for the invariant dilepton mass m_{ll} and a missing transverse momentum cut. Additionally, events in the $\mu\mu$-channel with no jets are rejected to improve the agreement between data and Monte Carlo simulations.

Since the control region is enriched in events producing Ana and Nonana leptons, an exact definition of these two object selections is necessary. The object-level information are obtained during reconstruction of the collision event using data from the detector.

Both Ana and Nonana leptons need to pass some basic criteria on their kinematic properties and associated tracks. For instance, muons and electrons need to pass a transverse momentum threshold of $p_{\mathrm{T}} = 27$ GeV.
In addition, an Analysis electron needs to be isolated and electron candidates with ambiguous AUTHOR, i.e. electron candidates also identified by photon reconstruction algorithms, are vetoed. A Non-Analysis electron is required to fail the isolation and/or the identification criteria.
Similarly, Analysis muons also need to pass the isolation criteria, while Non-Analysis muons have a looser isolation and/or their track is effectively to distant to the primary vertex, denoted by the transverse impact parameter d_0.
The total list of selection criterias involved in the Analysis and Non-Analysis selection, are given in Appendix B.1.

Table 5.1: Event selection cuts used to define the fake factor control region (CR) for WWss scattering.(Ref.[18])

Exactly two Ana/Non-Ana leptons with identical electrical charge
One Analysis lepton
No additional Veto leptons
Electrons have to pass ECIDS
Failing at least one cut of the SR
In ee channel: $(m_{ll} < 79GeV)\|(98GeV < m_{ll})$ and $E_T^{miss} > 19GeV$
In $\mu\mu$ channel: $n_{jet} > 0$

5.2 Monte Carlo Closure Test

A Monte Carlo Closure Test can be performed in the control region to check whether the technical framework is working properly and to validate the extraction of the fake factor. During this test the fake factor is applied as an event weight to the number of events N_{AN} consisting of an Ana and Nonana selected leptons. Thus an estimation of the fake leptons which pass the Ana selection is obtained. By doing this in the Control Region the denominator of the Fake Factor cancels out with the total number of selected Ana and Nonana events of the Monte Carlo prediction N_{AN}^{fake}, since those are the same, and only the numerator of the Fake Factor N_{AA}^{fake} remains (Eq.5.1).

$$N^{fake,data-driven} = N_{AN}^{fake,MC} \cdot F = N_{AN}^{fake} \cdot \frac{N_{AA}^{fake}}{N_{AN}^{fake}} = N_{AA}^{fake} = N_{AA}^{fake} \ . \quad (5.1)$$

Hence the total number of two Ana selected leptons from the Monte Carlo simulation N_{AA}^{fake} and Fake Factor estimation should agree perfectly regardless of their uncertainties. This

[1]Electron Charge IDentification Selector, a tool/(neural network) trained using a boosted decision tree (BDT) to accurately and efficiently identify leptons

result is then denoted as an ideal closure. However the agreement of the integrated number of events is not the only aspect, that indicates a successful plot. The obtained distributions of N_{AA} should also be in good agreement with the MC prediction. Although small deviations for the respective lepton flavour channels are acceptable, since for example the binning of the distribution and applied fake factor do not match.

Therefore, a successful closure, when performing the Monte Carlo Closure Test in the control region, is indicated by a close agreement for the integrated number of events as well as a good agreement between the predicted MC distribution and data driven estimation of the fake lepton background.

5.3 p_{T} dependent Fake Factor

To gain a first impression on the fake lepton background contribution from $t\bar{t}$ and $W{+}jets$ to the WWss process, the fake factor $F(p_{\mathrm{T}})$ is obtained. It is extracted from the Control Region and binned in the fake lepton transverse momentum p_{T}. The calculation of the fake factor presented in this section is also applicable for further extractions of the fake factor in the remaining part of this thesis. A Monte Carlo Closure Test is then performed in the Control Region to validate the fake factor Method and its technical implementation.

As explained in Section 4.2 the fake factor is determined as the ratio of events with the fake being selected as Ana, to it being selected as a Nonana lepton. Since the employed MC samples contain only truth-level information one of the lepton is always assigned to be fake, while the other is prompt. Therefore the prompt contribution in Equation 4.6 is zero. This is of course not longer the case, when using a looser assignment like in actual data. Also, remember that in this study the contribution of double fake events is neglected.

Therefore, the distribution binned in the Probe lepton transverse momentum p_{T} is obtained for two different conditions. In the first distribution, both leptons are selected as Ana (AA). While in the second distribution, the assigned fake lepton is Nonana selected and the prompt is selected as Ana (NA). These are then binned in intervals of the transverse momentum of the fake lepton. Later, a broad binning is chosen for the fake factor to lower its statistical uncertainties.

By dividing the two histograms (Ana/Nonana) the fake factor is calculated. It is reasonable to assume that the Fake Factor is also lepton flavor dependent (see Section 4.2), which leads to separate fake factors for electrons and muons, as can be observed from Figures 5.1 and 5.2. The fake factors of both processes show similar behaviours for electrons and muons each, where $F(p_{\mathrm{T}})$ increases with increasing transverse momentum of the fake lepton. Also it can be observed, that in general $F(p_{\mathrm{T}})$ is larger for electron than for muons, particularly in the $t\bar{t}$ process. On the other hand the statistical uncertainties on $F(p_{\mathrm{T}})$ are bigger for $W{+}jets$ compared to $t\bar{t}$, especially for high p_{T} muons with uncertainties of around 30 %.

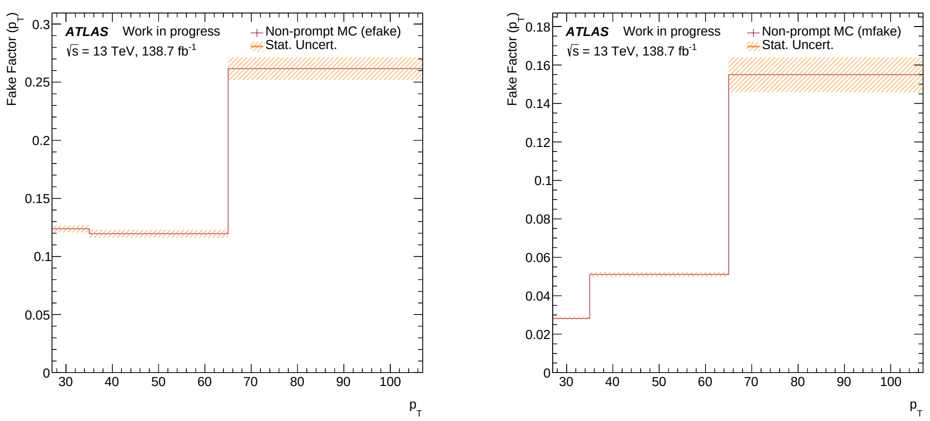
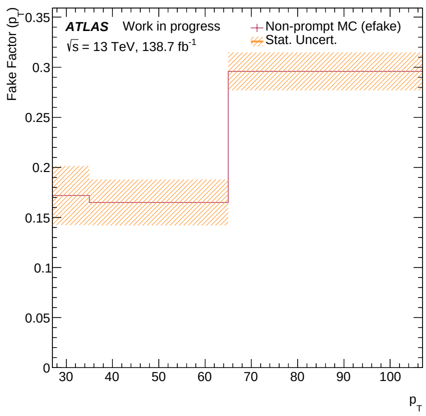


Figure 5.1: The calculated electron (left) and muon (right) fake factors $F(p_{\mathrm{T}})$ (red line) with the respective statistical uncertainties (orange) for the $t\bar{t}$ process.



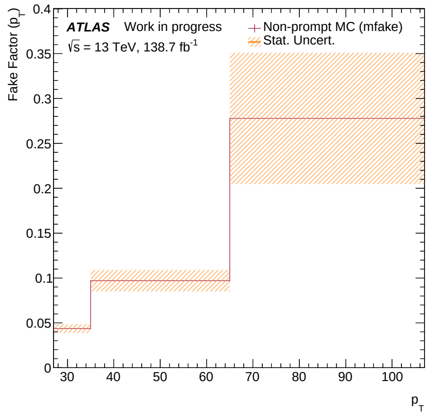


Figure 5.2: The calculated electron (left) and muon (right) fake factors $F(p_{\mathrm{T}})$ (red line) with the respective statistical uncertainties (orange) for the W+$jets$ process.

Table 5.2: Overview of number of events from integrated Monte Carlo Closure Test for the different lepton flavour channels of W+$jets$, only statistical uncertainties are observed. Figures are not rounded, to emphasize on the expected accurate agreement between Monte Carlo and data-driven estimation. The P and F denotes Prompt and Fake, while A and N denotes Ana and Nonana selected leptons.

	ee	$e\mu$	μe	$\mu\mu$	all
MC PF	674.1 ± 80	472.5 ± 50	557.1 ± 80	88.4 ± 9	1792.0 ± 120
MC FP	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Data-driven AN	633.7 ± 29	449.8 ± 20	631.0 ± 34	77.7 ± 3	1792.0 ± 50
Data-driven NA	0.0 ± 0.0	0.0 ± 2	0.0 ± 3	0.0 ± 0.0	0.0 ± 4
Total MC	674.1 ± 80	472.5 ± 50	557.1 ± 80	88.4 ± 9	1792.0 ± 120
Datadriven	633.7 ± 29	449.8 ± 20	631.0 ± 34	77.7 ± 3	1792.0 ± 50

Table 5.3: Overview of number of events from integrated Monte Carlo Closure Test for the different lepton flavour channels of $t\bar{t}$, only statistical uncertainties are observed. Figures are not rounded, to emphasize on the expected accurate agreement between Monte Carlo and data-driven estimation. The P and F denotes Prompt and Fake, while A and N denotes Ana and Nonana selected leptons.

	ee	$e\mu$	μe	$\mu\mu$	all
MC PF	260.3 ± 6	324.2 ± 7	319.8 ± 7	332.0 ± 7	1236.3 ± 14
MC FP	0.0 ± 0	0.0 ± 0	0.0 ± 0	0.0 ± 0	0.0 ± 0
Data-driven AN	271.8 ± 3	338.3 ± 2	313.1 ± 2	313.1 ± 2	1236.3 ± 4
Datad-riven NA	0.0 ± 0	0.0 ± 0.5	0.0 ± 0.33	0.0 ± 0	0.0 ± 0.6
Total MC	260.3 ± 6	324.2 ± 7	319.8 ± 7	332.0 ± 7	1236 ± 14
Datadriven	271.8 ± 3	338.3 ± 2	313.1 ± 2	313.1 ± 2	1236 ± 5

Extracting the Fake Factor in the Control Region and applying it there results in an ideal closure, for both considered processes $t\bar{t}$ and W+$jets$. Results are shown in Tables 5.2 and 5.3. The obtained total number of events from the data-driven background estimation of the Fake Factor agrees perfectly with the MC simulation (Prompt+Fake, both Ana) for the all channel, independent of the uncertainties. This is expected, because a perfect truth level assignment was used, proving that the technical implementation of the Fake Factor method in the framework is working properly.

The agreement between the η-distributions of the Monte Carlo prediction and fake lepton background estimation using only $F(p_T)$ seems already reasonably good for $t\bar{t}$, as can be seen in the previous figures. The leading and subleading lepton eta distributions agree with the simulation within their respective uncertainties. This can be probably explained by the strong rejection of chargeflip. Since chargeflip events, that are highly η-dependent, are explicitly excluded in this study and their contribution is being subtracted from the observed events, they should not influence the η-distribution anymore, possibly contributing to the observed results. However there seems to be some systematic trends which are not properly estimated by the $F(p_T)$.

In the channels with the fake electron being the leading electron (ee and $e\mu$) the events in the barrel area of the ECAL ($|\eta| < 1.4$) seem to get systematically overestimated (see Fig. 5.3 and Appendix C.2). While the events in the end cap area ($|\eta| > 1.5$) get slightly underestimated. Furthermore, the channels with a muon as the the leading lepton (μe and $\mu\mu$) seem to get underestimated as well.

The respective η-distributions for W+*jets* suffer from large statistical uncertainties even though a much coarser binning is imposed in comparison to $t\bar{t}$. Despite that the estimation and MC prediction mostly agree, except for the central part of the barrel, where the data-driven results overestimate the MC data significantly.

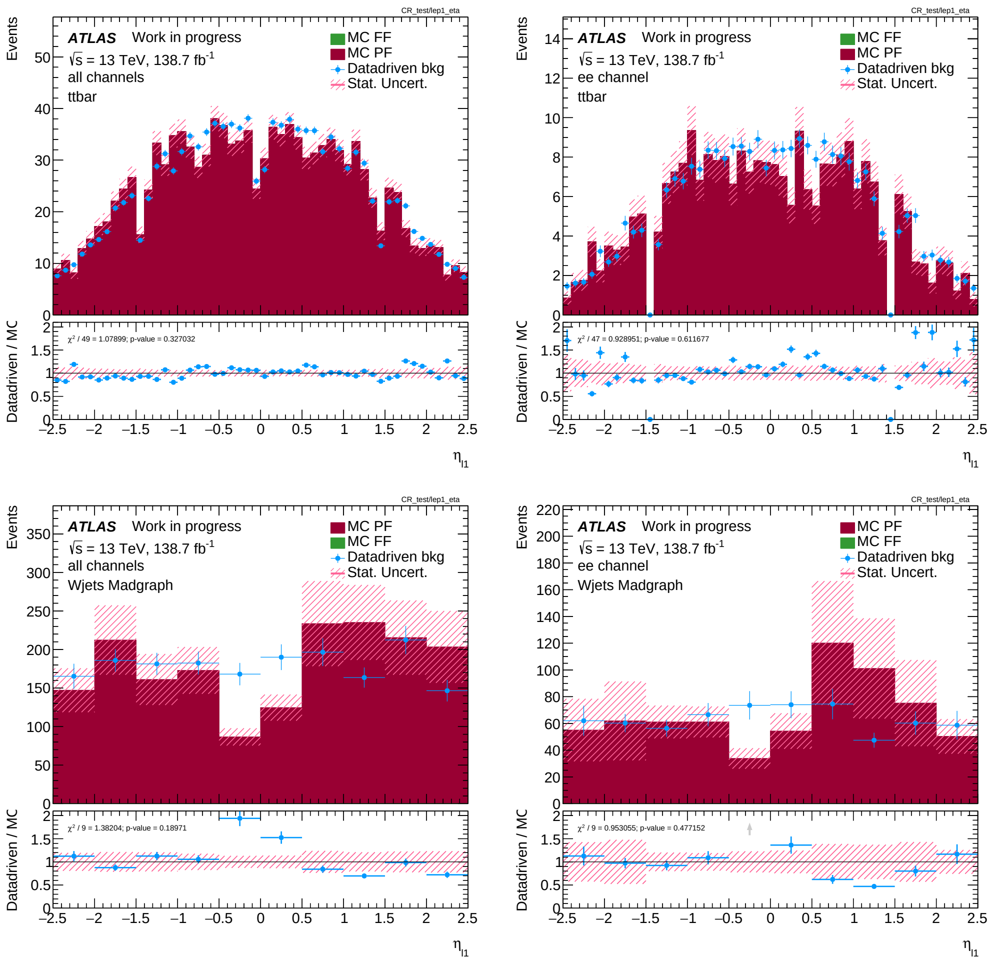


Figure 5.3: Event distributions for pseudorapidity of leading leptons, denoted by η_{l1} (upper row: $t\bar{t}$, lower row: W+*jets*. The channel description always refers to the leading and then subleading lepton, except for the sum of all channels. "Prompt+Fake" events predicted by Monte Carlo (red) are compared with the data-driven estimation (blue) from the fake factor, binned in the fake leptons p_T, including a ratio plot. The respective statistical uncertainties are show as well.

5.4 η dependent Fake Factor

It is reasonable to assume that the fake lepton background is not only determined by the detector physics involved in measuring the particle energies like p_T. The fake lepton contribution could also depend on the region where the object is reconstructed, which is also shown by other studies [2, 3]. This is to be expected because the ATLAS detector is not homogeneous and consists of multiple subdetectors with different coverages. A convenient variable for describing the detector geometry is η as mentioned in Section 3.2 already. Therefore the fake factor $F(\eta)$ is now binned in the fake leptons pseudorapidity instead of the transverse momentum and calculated to check for any dependencies on the detector region.

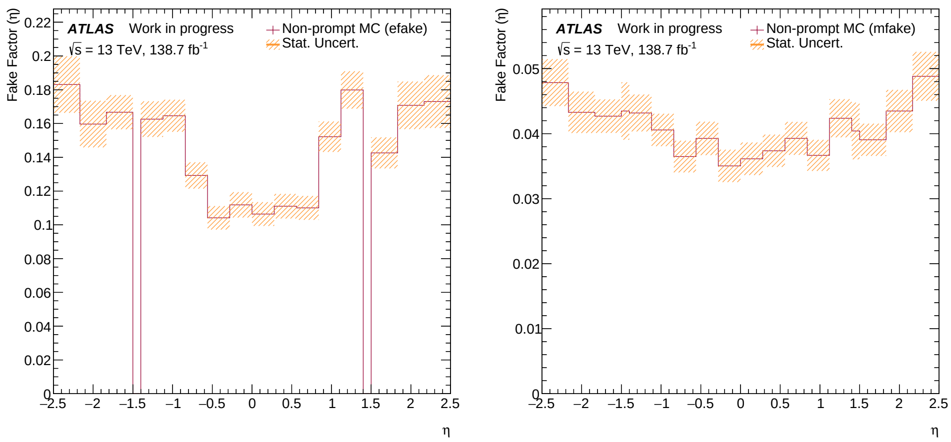


Figure 5.4: The calculated electron (left) and muon (right) fake factors $F(p_T)$ (red line) with the respective statistical uncertainties (orange) for the $t\bar{t}$ process.

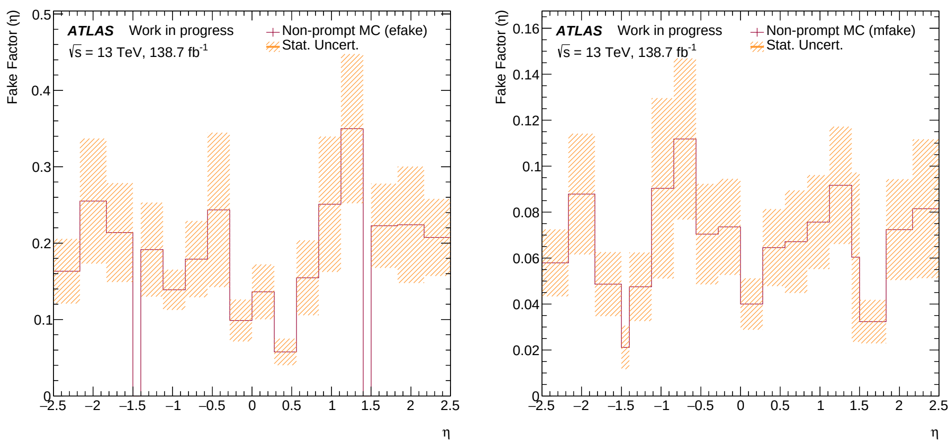


Figure 5.5: The calculated electron (left) and muon (right) fake factors $F(p_T)$ (red line) with the respective statistical uncertainties (orange) for the W+$jets$ process.

It can be observed from Figures 5.4, 5.5 that for$t\bar{t}$ and W+$jets$, F(η) is not a "flat distribution". Thus it implies that the number of fake leptons passing the Ana selection is somehow correlated to the pseudorapidity. Consequently the fake lepton background is then also expected to be η-dependent.

Although, it is important to note that the fake factors for the W+$jets$ process are dominated by statistical Monte Carlo uncertainties of up to 40 %. Thus no specific symmetry or behaviour is observable for this process regardless of the fake lepton flavour.

Contrary to that, both fake factors from the $t\bar{t}$ seem to be rather symmetric around $\eta = 0$. This will be examined further in Section **??**. Further, it can be observed that the that the relative fake event contribution seems to be the lowest in the central part of the barrel area ($\eta \approx 0$) of the ECAL and increases throughout the end-cap ($|\eta| < 2.5$). Also the transition area of the ECAL from the barrel to the end-caps ($1.37 < |\eta| < 1.52$) is easily spotted. There $F(\eta)$ suddenly drops down to zero. Though it does not imply that no fake lepton contribtuion is expected in this area of the detector, but rather events measured in this part of the detector are rejected due to the Ana and Nonana selection. This can be observed in the respective lepton η-distributions, included in the Appendix B.3.

A similar Fake Factor behaviour is observed for the muon fake contribution as well. However the transition area of the muon detector system is substantially covered, enabling a continuous measurement in η. The detector resolution only decreases perpendicular to the beam axis ($\eta \approx 0$) as a result of the detector structure (see Appendix B.3 as well).

Using the extracted η-FF, a Monte Carlo Closure Test on the control Region was performed. As expected, the results show an ideal closure in the total number of events, which also validates the obtained $F(\eta)$. More details on the obtained results as well as the respective η-distributions of the fake lepton background estimation are included in the Appendix C.1 and C.2.

Chapter 6

Fake Factor η Correction for $t\bar{t}$

In chapter 5 it was shown that the Fake Factor, for the fake lepton background of WWss, depends on the fake lepton transverse momentum p_{T}, but also on the region of the detector where it was reconstructed/measured. Implying that both quantities need to be considered to correctly estimate the fake lepton background. A 2D-binning of the Fake Factor meaning that it is calculated in bins of p_{T} and η would be ideal. This would preserve any correlations between these two variables and probably improve the background estimation significantly. However, this is not possible for this study due to limited statistics. Such a 2D-binned Fake Factor would simply be dominated by the statistical uncertainties.Therefore, a different approach is chosen for this thesis. As done in other studies, like for WZ [2, 3] a correction, with respect to η, of the $F(\eta)$ is derived and examined.

6.1 Combination of p_{T} and η based Fake Factors

Since a 2D-binning of the Fake Factor is not possible, the p_{T} and η binned fake factors need to be determined separately and applied consecutively during the Monte Carlo Closure Test. However applying them together as an event weight, without any adjustments, would result in an unsuccessful Closure Test, because both are already properly estimating the number of fake events on their own.

Since this study aims to further improve the fake background estimation of the p_{T} binned fake factor, only the fake factor binned in η needs to be adjusted while the $F(p_{\mathrm{T}})$ remains unchanged. Therefore applying the η fake factor or rather the desired η-correction should not change the total number of estimated fake events but eventually improve the shape of the eta distribution. Thus the η fake factor needs to be normalized accordingly. The specific details involved in normalizing it are explained in Section 6.2 .

To gain a first impression on the effect of considering both p_{T} and η, when estimating the fake lepton background, a successful Monte Carlo Closure Test is performed (see Table 6.1). As can be seen from the Figure 6.1 the fake background estimation (figure 5.3) for the lepton η-distribution improves, if both Fake Factors are applied.

For the ee- and $e\mu$-channel both, the systematic overestimation inside the barrel area as well as the slight underestimation in the end-caps, are being corrected.
Whilst the μe- and $\mu\mu$-channel η-distributions also get shaped slightly better. However they still seem to get systematically underestimated, but the deviations are still within the respective statistical uncertainties.

However the distributions obtained from $W{+}jets$-sample are still dominated by statistical uncertainties even when using a very broad binning. Furthermore, only minor improvements can be observed and the overestimation of the central barrel area is still present. This suggests that the observed results are significantly limited by the available statistics and quality of the MC-simulation. Any further employment of analytical methods like a fit is expected to have very large uncertainties and would not be reasonable. Thus this sample will be excluded for

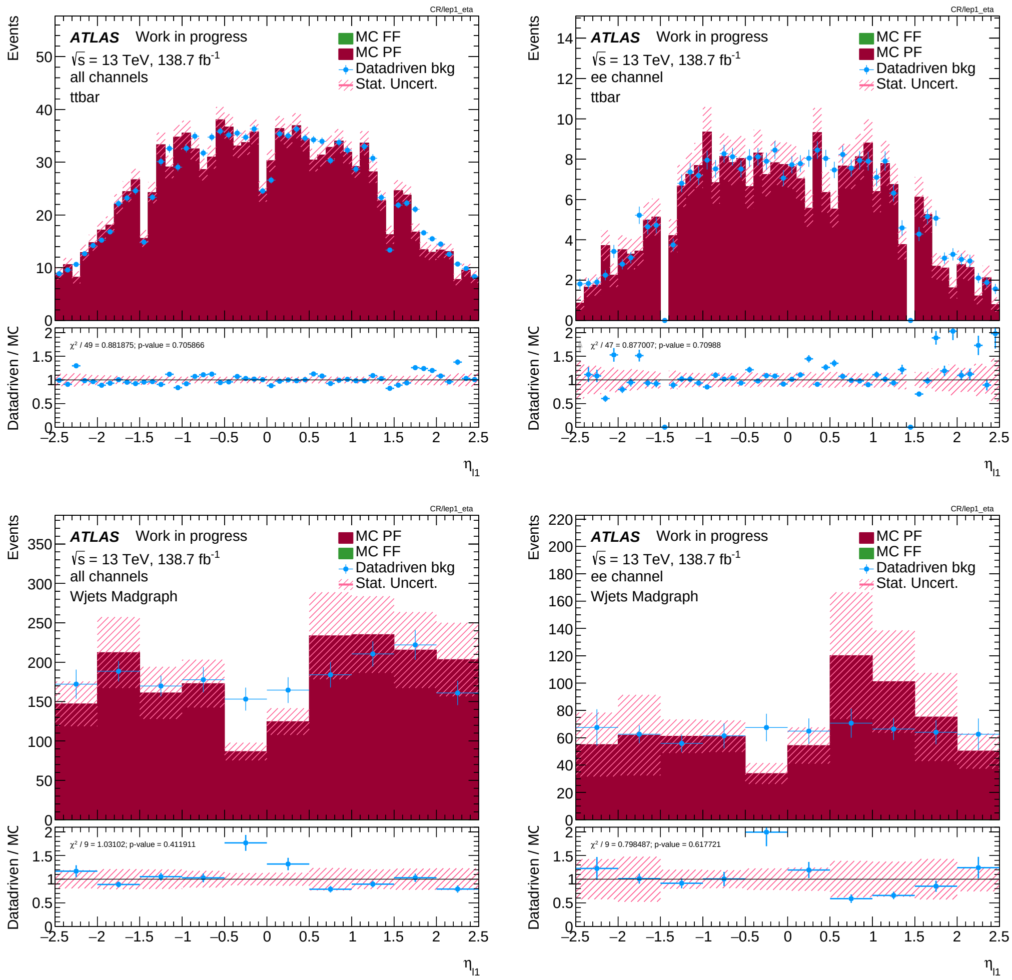


Figure 6.1: Event distributions for pseudorapidity of leading leptons, denoted by η_{l1} for $t\bar{t}$ (upper row) or W+*jets* (lower row). The channel description always refers to the leading and then subleading lepton, except for the sum of all channels. ”Prompt+Fake” events predicted by Monte Carlo (red) are compared with the data-driven estimation (blue) from both fake factor applied together, binned in the fake leptons p_{T} and η respectively, including a ratio plot (note: the applied η fake factor is rescaled according to 6.2). The respective statistical uncertainties are show as well.

the remaining part of this thesis. Therefore an η-correction to $F(\eta)$ is derived only for the $t\bar{t}$-sample.

Unfortunately the χ^2-test is not sensible to improvements of systematic shifts. It is only evaluating the agreement between MC and Data in each bin on its own. Thus it does not depend on the order of the tested bins or is not sensible to bin shuffling either. However applying $F(\eta)$ improves the modelling across multiples adjacent bins as for example seen in the ECAL barrel especially. Therefore another test, which is also sensitive to changes on the adjacent bins is needed to properly quantify and evaluate the possible improvements of the η-correction. An additional test will be performed for a proper comparison of the obtained results in the final section of this chapter.

6.2 Normalization of η Correction

As previously explained the η fake factor respectively the η-correction need to be normalized properly to only improve the shape of the background estimation but not change the total number of events. Since the total number of doubly Ana selected events is already properly estimated by the p_T fake factor. Thus the applied correction should only account for the η dependency.

The η Fake Factor distribution is normalized by the mean η fake factor of the Control Region [2]. This mean η fake factor is calculated by binning the lepton η-distribution in just one inclusive bin from $\eta = -2.5$ to $\eta = 2.5$.

The normalization is validated using the Monte Carlo Closure Test on the Control Region. A successful Closure is observed (Table 6.1) and only the modelling of the fake background is improved by applying the normalized $F(\eta)$ distribution, as could be observed in Section 6.1. It is important to note that no ideal closure is expected and deviations within the uncertainties are acceptable.

Table 6.1: Overview of number of events from integrated Monte Carlo Closure Test for the different lepton flavour channels of $t\bar{t}$, only statistical uncertainties are observed. During the data-driven estimation both the p_{T} and the normalized η based fake factor are combined. The P and F denotes Prompt and Fake, while A and N denotes Ana and Nonana selected leptons.

	ee	$e\mu$	μe	$\mu\mu$	all
MC PF	260.3 ± 6	324.2 ± 7	319.8 ± 7	332.0 ± 7	1236.3 ± 14
MC FP	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Datadriven AN	273.2 ± 3	340.5 ± 2	311.6 ± 2	313.2 ± 2	1238.4 ± 5
Datadriven NA	0.0 ± 0.0	0.0 ± 0.4	0.0 ± 0.34	0.0 ± 0.0	0.0 ± 0.6
total MC	260.3 ± 6	324.2 ± 7	319.8 ± 7	332.0 ± 7	1236.3 ± 14
Datadriven	273.2 ± 3	340.5 ± 2	311.6 ± 2	313.2 ± 2	1238.4 ± 5

6.3 Symmetrization of η Correction

According to Figures 5.4 the $F(\eta)$ for $t\bar{t}$ seems to be symmetric around $\eta = 0$. The same can be observed for the corresponding lepton eta distributions [Ana and Nonana: reference to Appendix]. It is reasonable to make such an assumption since the detector itself is symmetric around $\eta = 0$ and therefore the fake lepton events are expected to depend on $|\eta|$ only. To verify this assumption the asymmetry between the negative and positive half of the fake lepton η-distribution is calculated. The results are shown in Figure 6.2.

The asymmetries for fake electrons and muons in the central barrel area mostly fluctuate between $\pm 5\,\%$. However with increasing $|\eta|$ the asymmetry increases as far as $10\,\%$ for the most forward detector regions, especially for the electron distributions. Although a possible explanation could be that the overall statistics start to rapidly decrease in these areas of the detector, leading to a higher fluctuation of the corresponding asymmetry factor, because of a smaller number of events per bin. To determine whether this asymmetry arises from low statistics or physical effects a sample with higher statistics would be needed. However the observed asymmetry is accepted to be sufficiently small in comparison to the normalized η fake factor, including the respective uncertainties.

In this thesis $F(\eta)$ is then symmetrized and binned in $|\eta|$ instead (Ref. [2, 3]). Effectively doubling the amount of statistics per bin, which should improve the background estimation and lower the statistical uncertainties on the η fake factor. Although this could introduce further small deviations in the Closure Test, because of the differences arising from the symmetrization.

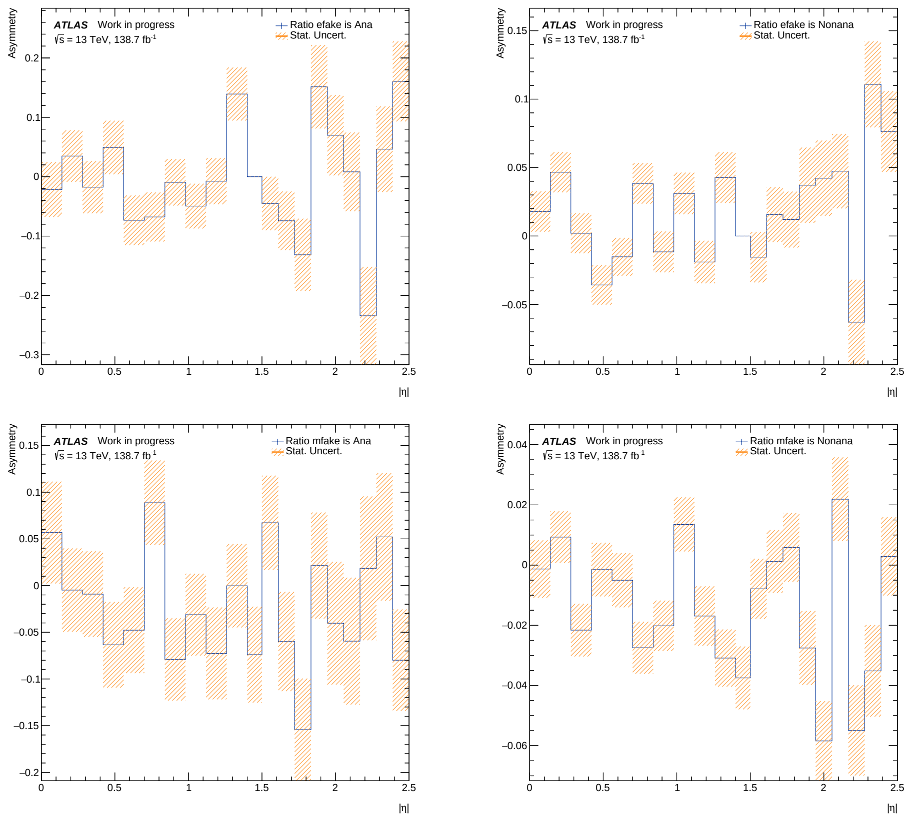


Figure 6.2: The asymmetry between the negative and positive half of the fake lepton η-distribution (blue) for the $t\bar{t}$ process with its respective statistical uncertainties (orange). Left column shows the asymmetry for the fake lepton being Ana selected, while the right column shows it for the lepton being Nonana selected. The respective fake lepton flavour is given in the legend (upper row: electrons, lower row: muons).

6.4 Fitted η Correction Function

The normalized $F(\eta)$ is now fitted as a function of $|\eta|$. This is done to extrapolate the per bin resolution of the η-correction. Expected to further improve the estimation of the η-dependency of fake lepton background.

The obtained fit, like the η fake factor, depends on the lepton flavour then. For the electrons a parabolic fit is applied for the barrel area ($|\eta| < 1.4$). A linear fit is chosen for the end-caps ($|\eta| > 1.5$) instead. Thus the derived correction for electrons depends on a total of 4 free parameters, because of the symmetrization around $\eta = 0$.
For the respective fake factor of the muons a parabolic fit is applied inclusive over the total detector area ($|\eta| < 2.5$). It then only depends of 2 free parameters.

$$C_e(\eta) = \begin{cases} a + b \cdot \eta^2 & , |\eta| < 1.4 \\ a + b \cdot \eta & , |\eta| > 1.5 \end{cases} \quad \text{and} \quad C_\mu(\eta) = a + b \cdot \eta^2 \ , |\eta| < 2.5 \tag{6.1}$$

This dependence on the lepton flavour could possibly be explained by the different detector geometries. The coverage of the ECAL is not continuous, whereas the muon detector, also consisting of multiple subdetectors, provides a more continuous coverage especially in the transition area. It is therefore not unlikely that different fake lepton behaviours are observed for different regions of the detector. Although this would need further investigation using a sample with higher statistics.

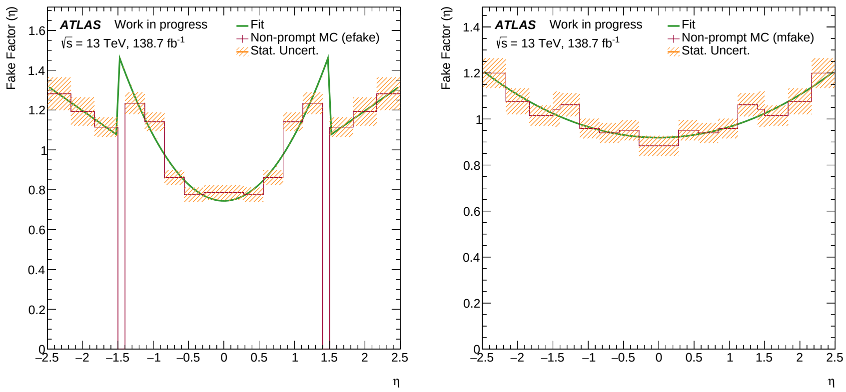


Figure 6.3: Fit (green) of the normalized η fake factor distribution (red) for $t\bar{t}$ process. The fake factor for the fake electron (muon) is shown on the left (right). Displayed statistical uncertainties (orange) belong to the binned fake factor distribution only, not the fit.

Figures 6.3 and Table **??**, which is included in the Appendix B.2, show the obtained fit results. To validate their quality the corresponding χ^2/NDF are provided as well. Considering the limited amount of statistics the fit results are acceptable. It is important to note, that the linear fit of the η-correction for electrons is only using three bins for which reason the χ^2/NDF value is probably underestimated. This is also reflected by the large uncertainties of the fit.

Since one of the limiting aspects of this Fake Factor study is the amount of statistics a thorough handling of the uncertainties is required. The correlation between the respective fit parameters are about ±0.4 and -0.7 (see Table B.4 in Appendix). Thus the statistical uncertainties are calculated using error propagation with correlated parameters (see eq. 6.2), where σ_a, σ_b denote the standard deviations of the respective parameters and ρ_{ab} their correlation.

$$\text{electrons:} \quad \sigma_{C_\eta} = \begin{cases} \sqrt{\sigma_a^2 + \sigma_b^2 \cdot \eta^4 + 2 \cdot \rho_{ab}\sigma_a\sigma_b \cdot \eta^2} & , |\eta| < 1.4 \\ \sqrt{\sigma_a^2 + \sigma_b^2 \cdot \eta^2 + 2 \cdot \rho_{ab}\sigma_a\sigma_b \cdot \eta} & , |\eta| > 1.5 \end{cases} \tag{6.2}$$

$$\text{muons:} \quad \sigma_{C_\eta} = \sqrt{\sigma_a^2 + \sigma_b^2 \cdot \eta^4 + 2 \cdot \rho_{ab}\sigma_a\sigma_b \cdot \eta^2} \;, |\eta| < 2.5 \tag{6.3}$$

6.5 Verification of η Correction

The obtained η-correction is applied in the Monte Carlo Closure Test as an additional event weight next to the p_T fake factor. Because of the technical framework used for this study the fit-function can not be applied itself. Therefore it needs to be converted into a histogram (see Fig.B.1 in Appendix). For this thesis a histogram with 100 bins over the range $|\eta| < 2.5$ is chosen. The resulting bin size is sufficiently small, because the η-distribution only consists of 50 bins only.

As explained in Section 6.1 the χ^2-test is not very suitable to validate the improvement of systematic shifts in the fake lepton background estimation, that were observed during this study. Additionally a Kolmogorov-Smirnov-test is now performed. This test contrary to the χ^2-test is sensitive to the change between adjacent bins. In this test the cumulative-distribution-function (CDF) of the MC estimation and the datadriven background are compared. The agreement of the two distributions can then be estimated from the maximum distance between the two CDFs across the range of all bins.

In Figure 6.4 an η-distribution, where the η-correction is applied, is compared to that, where only the $F(p_\mathrm{T})$ is used for the fake background estimation. Additionally the ratio between MC and the datadriven background estimation is provided as well as a comparison between the respective CDF from the Kolmogorov-Smirnov test. Only some exemplary distribution are included in this section, while the remaining are shown in Appendix C.2.6. The obtained results prove that the η-correction improves the modelling of the fake lepton background estimation. This is clearly observable from the comparison of the respetive values for the Kolmogorov-Smirnov statistic. The maximum distance between the CDF of MC simulation and datadriven background decreases across all channels, when the derived correction is applied. Results of the Kolmogorv-Smirnov test are summarized in Table 6.3.

Furthermore similar improvements to the modelling of the η-distribution as in Section 6.1 can be observed. For the ee-channel the agreement with the prediction is much better, because the overestimated number of events in the barrel area is getting decreased. This improvement applies to other distributions with electron contributions regardless of it being leading or subleading. Although in the μe-channel, where the electron is the subleading lepton the number of events seems slightly underestimated for this region when compared to the prediction, but still agrees within the respective uncertainties. The agreement between the predicted μ distribution and datadriven background looks good for the mixed flavour channels too. In general the η-correction only partially improves the modelling for the end cap regions. This is rather expected, because the available statistics for this part of the detector are very limited. A sample with higher statistics would then be very beneficial for improving the quality of the fits, respectively the correction, especially with regard to the Fake Factor for electrons.

However in the $\mu\mu$-channel the distribution is being underestimated from the datadriven estimation after all. Since this is already present in the distributions with only $F(p_{\mathrm{T}})$ and gets not corrected by the η correction, it remains unclear why this is observed. Another observation, which is not properly understood, is the significant increase of events of subleading leptons in the mixed flavour $e\mu$-channel for $\eta < -2.4$, which corresponds to the parts of the detector closest to the beam axis. It is unclear whether this excess in events is due to the small amount of statistics in the end-cap region or a insufficient prediction of the MC sample.

In conclusion the η-correction improves the estimation of the fake lepton background in the $t\bar{t}$-sample considerably. It is important to note that the obtained results are using a perfect truth-level assignment and need to be validated on a reco-level assignment using a Closure Test. Thus the obtained results would be more relatable to actual data. Furthermore an datadriven estimation using actual data remains to be performed as well.

Table 6.2: Overview of number of events from integrated Monte Carlo Closure Test for the p_{T} fake factor with η correction in $t\bar{t}$ (note: only stat. uncertainties observed).

	ee	$e\mu$	μe	$\mu\mu$	all
MC PF	260.3 ± 6	324.2 ± 7	319.8 ± 7	332.0 ± 7	1236.3 ± 14
MC FP	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Datadriven AN	272.4 ± 3	339.6 ± 2	311.0 ± 2	312.9 ± 2	1235.9 ± 5
Datadriven NA	0.0 ± 0.0	0.0 ± 0.5	0.0 ± 0.34	0.0 ± 0.0	0.0 ± 0.6
total MC	260.3 ± 6	324.2 ± 7	319.8 ± 7	332.0 ± 7	1236.3 ± 14
Datadriven	272.4 ± 3	339.6 ± 2	311.0 ± 2	312.9 ± 2	1235.9 ± 5

Table 6.3: Comparison of Kolmogorv-Smirnov test results for estimation of η distribution of fake lepton background in $t\bar{t}$ with only the p_{T} binned fake factor against the estimation with an η corrected fake factor.

		Kolmogorov-Smirnov statistic				
		ee	$e\mu$	μe	$\mu\mu$	all
lead lep ℓ_1	only $F(p_T)$	0.0324	0.0263	0.0229	0.0288	0.0256
	η corrected $F(p_T)$	0.0244	0.0115	0.0244	0.0255	0.0163
sublead lep ℓ_2	only $F(p_T)$	0.0433	0.0171	0.0304	0.0360	0.0268
	η corrected $F(p_T)$	0.0249	0.0119	0.0152	0.0243	0.0115

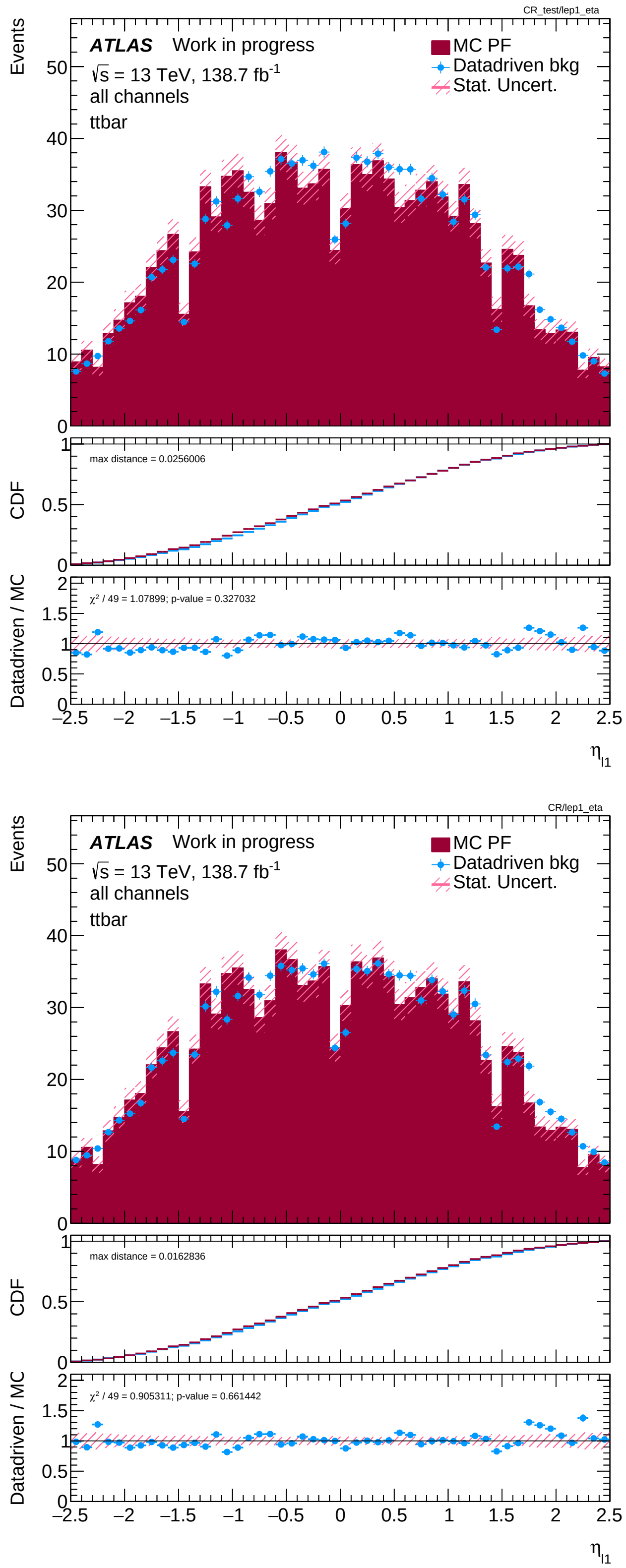


Figure 6.4: Exemplary comparison of estimated leading lepton η-distribution (upper row: with p_{T} fake factor only, lower row: with η corrected p_{T} fake factor) for all channels in $t\bar{t}$ (more in Appendix C.2.6) . Furthermore, CDF comparison for MC prediction and data-driven estimate, and respective ratio plots are included as auxiliary plots.

Chapter 7

Conclusion and Outlook

This thesis aims to improve the fake lepton background estimation for the $W^{\pm}W^{\pm}jj$-EWK scattering process. The fake lepton background is investigated for the $t\bar{t}$ and $W{+}jets$ process. An extensive fake factor study is conducted in a suitable control region to investigate the dependency of the fake factor from the pseudorapidity of the fake lepton, apart from the fake leptons transverse momentum p_{T}. This study is conducted using truth level information from MC simulations. The MC samples are generated in MADGRAPH and PYTHIA and simulate proton-proton collisions at $\sqrt{s} = 13\,TeV$ in the ATLAS detector with an integrated luminosity of $138.7\,\mathrm{fb}^{-1}$. Multiple Monte-Carlo Closure Test in the control region are performed to validate the obtained results and ensure the proper implementation of the fake factor method.

It is found that the fake factors of the $t\bar{t}$ and $W{+}jets$ process not only depend of the fake leptons transverse momentum, but its pseudorapidity as well. The electron fake factor for $t\bar{t}$ shows a parabolic dependency of η for the central area of the electromagnetic calorimeter inside the barrel ($|\eta| < 1.5$). A similar dependency is observed for muons for the total detector area, that was investigated ($|\eta| < 2.5$). However, the obtained results for $W{+}jets$ display large statistical uncertainties from the MC simulations, making it very difficult to determine any particular behaviour of the η-dependency. This could be investigated further in future studies on a sample with higher statistics or an improved MC generator. For that reason the $W{+}jets$ process is excluded for the remaining part of this thesis and only the $t\bar{t}$ process is investigated further.

Eventually an additional η-correction is derived for the p_{T} binned fake factor, because the current sample does not allow a fake factor binned in p_{T} and η simultaneously due to the limited amount of statistics. In the barrel region, a parabolic fit is applied to the properly rescaled η binned fake factor for both lepton flavours, to account for the respective η dependency. Furthermore a linear fit is used for the electron fake factor in the end-cap area of the detector ($|\eta| > 1.5$). For muons, the parabolic fit is extended up to the end-caps (see Fig.6.3). A good quality of the fits is observed, especially for the parabolic ones, and the corresponding statistical uncertainties are calculated using correlated error propagation. The obtained correction is stronger for electrons than for muons. However the specific cause of this behaviour remains unknown and could be examined in the future.

Applying the η-corrected fake factor is found to successfully improve the shape of the data driven estimation of the η-distributions by correcting some observed systematic shifts in the barrel area, especially for electrons, while the systematic underestimation of the distributions for muons is not corrected. A Kolmogorov-Smirnov test is performed to validate the observed improvements (see Fig. 6.4, C.2.6). Although, it is important to note that the agreement between the predicted Monte-Carlo distribution and the fake background estimation from the solely p_{T} binned fake factor is already reasonably good. Due to the limited amount of statistics the fake factor was not binned p_{T} and η simultaneously. However such a 2D-binning would maintain any correlation between these two quantities, which could provide further insight on the fake lepton background. It is important to note that this will not be feasible with the full run 2 data and probably requires a significantly larger sample size.

Since this thesis is conducted with a lepton selection in the control region based on truth-level information, its results remain to be validated for a selection on reco-level, using a Monte-Carlo closure test in the control region. This would require a proper assignment of the leptons (Tag and Probe) in the calculation of the fake factor. The methodology presented in this thesis is not limited to this particular control region or process, but could be also applied in separate fake factor studies, where an η-correction is expected to be necessary. Thus, it would be very interesting to conduct such a study on chargeflip events. These events are currently excluded from the non-prompt background definition of this thesis, but are expected to be very dependent on p_{T} as well as η, making its background estimation a very likely candidate for an η-correction.

Appendix A

List of all samples

Table A.1: List of Monte Carlo samples for $W + jets$ process

Dataset	nEvents	kFactor	xsec	genFilterEff
mc16_13TeV.363600.MGPy8EG_N30NLO_Wenu_Ht0_70_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	31237000	1.12	16719.0	0.83756
mc16_13TeV.363600.MGPy8EG_N30NLO_Wenu_Ht0_70_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	41282000	1.12	16719.0	0.83756
mc16_13TeV.363600.MGPy8EG_N30NLO_Wenu_Ht0_70_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	24990000	1.12	16719.0	0.83756
mc16_13TeV.363601.MGPy8EG_N30NLO_Wenu_Ht0_70_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	20215800	1.12	16720.0	0.13893
mc16_13TeV.363601.MGPy8EG_N30NLO_Wenu_Ht0_70_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	26754000	1.12	16720.0	0.13893
mc16_13TeV.363601.MGPy8EG_N30NLO_Wenu_Ht0_70_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	16159800	1.12	16720.0	0.13893
mc16_13TeV.363602.MGPy8EG_N30NLO_Wenu_Ht0_70_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	11034100	1.12	16719.0	0.024676
mc16_13TeV.363602.MGPy8EG_N30NLO_Wenu_Ht0_70_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	14888000	1.12	16719.0	0.024676
mc16_13TeV.363602.MGPy8EG_N30NLO_Wenu_Ht0_70_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	8998750	1.12	16719.0	0.024676
mc16_13TeV.363603.MGPy8EG_N30NLO_Wenu_Ht70_140_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	6201000	1.12	755.32	0.71219
mc16_13TeV.363603.MGPy8EG_N30NLO_Wenu_Ht70_140_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	8200000	1.12	755.32	0.71219
mc16_13TeV.363603.MGPy8EG_N30NLO_Wenu_Ht70_140_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	4979500	1.12	755.32	0.71219
mc16_13TeV.363604.MGPy8EG_N30NLO_Wenu_Ht70_140_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	5622700	1.12	755.77	0.24182
mc16_13TeV.363604.MGPy8EG_N30NLO_Wenu_Ht70_140_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	7447000	1.12	755.77	0.24182
mc16_13TeV.363604.MGPy8EG_N30NLO_Wenu_Ht70_140_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	4487600	1.12	755.77	0.24182
mc16_13TeV.363605.MGPy8EG_N30NLO_Wenu_Ht70_140_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	6153150	1.12	755.65	0.050389
mc16_13TeV.363605.MGPy8EG_N30NLO_Wenu_Ht70_140_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	8197000	1.12	755.65	0.050389
mc16_13TeV.363605.MGPy8EG_N30NLO_Wenu_Ht70_140_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	4951000	1.12	755.65	0.050389
mc16_13TeV.363606.MGPy8EG_N30NLO_Wenu_Ht140_280_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	6791000	1.12	318.89	0.6660
mc16_13TeV.363606.MGPy8EG_N30NLO_Wenu_Ht140_280_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	8999000	1.12	318.89	0.6660
mc16_13TeV.363606.MGPy8EG_N30NLO_Wenu_Ht140_280_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	5467500	1.12	318.89	0.6660
mc16_13TeV.363607.MGPy8EG_N30NLO_Wenu_Ht140_280_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	4999800	1.12	319.96	0.268
mc16_13TeV.363607.MGPy8EG_N30NLO_Wenu_Ht140_280_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	6572000	1.12	319.96	0.268
mc16_13TeV.363607.MGPy8EG_N30NLO_Wenu_Ht140_280_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	3997450	1.12	319.96	0.268
mc16_13TeV.363608.MGPy8EG_N30NLO_Wenu_Ht140_280_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	5042000	1.12	319.48	0.069443
mc16_13TeV.363608.MGPy8EG_N30NLO_Wenu_Ht140_280_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	6653000	1.12	319.48	0.069443
mc16_13TeV.363608.MGPy8EG_N30NLO_Wenu_Ht140_280_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	4041000	1.12	319.48	0.069443
mc16_13TeV.363609.MGPy8EG_N30NLO_Wenu_Ht280_500_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	2736500	1.12	73.558	0.61954
mc16_13TeV.363609.MGPy8EG_N30NLO_Wenu_Ht280_500_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	3475000	1.12	73.558	0.61954
mc16_13TeV.363609.MGPy8EG_N30NLO_Wenu_Ht280_500_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	2193000	1.12	73.558	0.61954
mc16_13TeV.363610.MGPy8EG_N30NLO_Wenu_Ht280_500_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	2487600	1.12	73.534	0.28662
mc16_13TeV.363610.MGPy8EG_N30NLO_Wenu_Ht280_500_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	3035000	1.12	73.534	0.28662
mc16_13TeV.363610.MGPy8EG_N30NLO_Wenu_Ht280_500_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	1992600	1.12	73.534	0.28662
mc16_13TeV.363611.MGPy8EG_N30NLO_Wenu_Ht280_500_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	2497800	1.12	73.56	0.097328
mc16_13TeV.363611.MGPy8EG_N30NLO_Wenu_Ht280_500_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	3249000	1.12	73.56	0.097328
mc16_13TeV.363611.MGPy8EG_N30NLO_Wenu_Ht280_500_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	1999050	1.12	73.56	0.097328
mc16_13TeV.363612.MGPy8EG_N30NLO_Wenu_Ht500_700_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	1468000	1.12	11.517	0.58696
mc16_13TeV.363612.MGPy8EG_N30NLO_Wenu_Ht500_700_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	1716000	1.12	11.517	0.58696
mc16_13TeV.363612.MGPy8EG_N30NLO_Wenu_Ht500_700_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	1184500	1.12	11.517	0.58696
mc16_13TeV.363613.MGPy8EG_N30NLO_Wenu_Ht500_700_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	494200	1.12	11.518	0.29906
mc16_13TeV.363613.MGPy8EG_N30NLO_Wenu_Ht500_700_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	645000	1.12	11.518	0.29906
mc16_13TeV.363613.MGPy8EG_N30NLO_Wenu_Ht500_700_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	396000	1.12	11.518	0.29906
mc16_13TeV.363614.MGPy8EG_N30NLO_Wenu_Ht500_700_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	364300	1.12	11.513	0.1165
mc16_13TeV.363614.MGPy8EG_N30NLO_Wenu_Ht500_700_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	489000	1.12	11.513	0.1165
mc16_13TeV.363614.MGPy8EG_N30NLO_Wenu_Ht500_700_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	295500	1.12	11.513	0.1165
mc16_13TeV.363615.MGPy8EG_N30NLO_Wenu_Ht700_1000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	477500	1.12	4.0173	0.56731
mc16_13TeV.363615.MGPy8EG_N30NLO_Wenu_Ht700_1000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	589000	1.12	4.0173	0.56731
mc16_13TeV.363615.MGPy8EG_N30NLO_Wenu_Ht700_1000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	386000	1.12	4.0173	0.56731
mc16_13TeV.363616.MGPy8EG_N30NLO_Wenu_Ht700_1000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	246600	1.12	4.0201	0.30526
mc16_13TeV.363616.MGPy8EG_N30NLO_Wenu_Ht700_1000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	331000	1.12	4.0201	0.30526
mc16_13TeV.363616.MGPy8EG_N30NLO_Wenu_Ht700_1000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	196600	1.12	4.0201	0.30526
mc16_13TeV.363617.MGPy8EG_N30NLO_Wenu_Ht700_1000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	127700	1.12	4.0189	0.12996
mc16_13TeV.363617.MGPy8EG_N30NLO_Wenu_Ht700_1000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	161000	1.12	4.0189	0.12996
mc16_13TeV.363617.MGPy8EG_N30NLO_Wenu_Ht700_1000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	98000	1.12	4.0189	0.12996
mc16_13TeV.363618.MGPy8EG_N30NLO_Wenu_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	94000	1.12	1.3238	0.54776
mc16_13TeV.363618.MGPy8EG_N30NLO_Wenu_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	124000	1.12	1.3238	0.54776
mc16_13TeV.363618.MGPy8EG_N30NLO_Wenu_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	84500	1.12	1.3238	0.54776
mc16_13TeV.363619.MGPy8EG_N30NLO_Wenu_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	62000	1.12	1.3249	0.31034
mc16_13TeV.363619.MGPy8EG_N30NLO_Wenu_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	87000	1.12	1.3249	0.31034
mc16_13TeV.363619.MGPy8EG_N30NLO_Wenu_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	53400	1.12	1.3249	0.31034
mc16_13TeV.363620.MGPy8EG_N30NLO_Wenu_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	60000	1.12	1.326	0.144
mc16_13TeV.363620.MGPy8EG_N30NLO_Wenu_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	87000	1.12	1.326	0.144
mc16_13TeV.363620.MGPy8EG_N30NLO_Wenu_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	53000	1.12	1.326	0.144
mc16_13TeV.363621.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r10201_p397	6600	1.12	0.042285	0.52634
mc16_13TeV.363621.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r10724_p397	14400	1.12	0.042285	0.52634
mc16_13TeV.363621.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r9364_p397	7000	1.12	0.042285	0.52634
mc16_13TeV.363622.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r10201_p397	8100	1.12	0.042199	0.31693
mc16_13TeV.363622.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r10724_p397	14200	1.12	0.042199	0.31693
mc16_13TeV.363622.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r9364_p397	7300	1.12	0.042199	0.31693
mc16_13TeV.363623.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e4835_s3126_r10201_p397	4600	1.12	0.045098	0.20867
mc16_13TeV.363623.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e4835_s3126_r10724_p397	4850	1.12	0.045098	0.20867
mc16_13TeV.363623.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e4835_s3126_r9364_p397	9000	1.12	0.045098	0.20867

Table A.2: List of Monte Carlo samples for $W + jets$ process

Dataset	nEvents	kFactor	xsec	genFilterEff
mc16_13TeV.363618.MGPy8EG_N30NLO_Wenu_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	94000	1.12	1.3238	0.54776
mc16_13TeV.363618.MGPy8EG_N30NLO_Wenu_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	124000	1.12	1.3238	0.54776
mc16_13TeV.363618.MGPy8EG_N30NLO_Wenu_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	84500	1.12	1.3238	0.54776
mc16_13TeV.363619.MGPy8EG_N30NLO_Wenu_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	62000	1.12	1.3249	0.31034
mc16_13TeV.363619.MGPy8EG_N30NLO_Wenu_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	87000	1.12	1.3249	0.31034
mc16_13TeV.363619.MGPy8EG_N30NLO_Wenu_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	53400	1.12	1.3249	0.31034
mc16_13TeV.363620.MGPy8EG_N30NLO_Wenu_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	60000	1.12	1.326	0.144
mc16_13TeV.363620.MGPy8EG_N30NLO_Wenu_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	87000	1.12	1.326	0.144
mc16_13TeV.363620.MGPy8EG_N30NLO_Wenu_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	53000	1.12	1.326	0.144
mc16_13TeV.363621.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r10201_p397	6600	1.12	0.042285	0.52634
mc16_13TeV.363621.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r10724_p397	14400	1.12	0.042285	0.52634
mc16_13TeV.363621.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r9364_p397	7000	1.12	0.042285	0.52634
mc16_13TeV.363622.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r10201_p397	8100	1.12	0.042199	0.31693
mc16_13TeV.363622.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r10724_p397	14200	1.12	0.042199	0.31693
mc16_13TeV.363622.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r9364_p397	7300	1.12	0.042199	0.31693
mc16_13TeV.363623.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e4835_s3126_r10201_p397	4600	1.12	0.045098	0.20867
mc16_13TeV.363623.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e4835_s3126_r10724_p397	4850	1.12	0.045098	0.20867
mc16_13TeV.363623.MGPy8EG_N30NLO_Wenu_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e4835_s3126_r9364_p397	9000	1.12	0.045098	0.20867
mc16_13TeV.363624.MGPy8EG_N30NLO_Wmunu_Ht0_70_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	31241000	1.12	16720.0	0.83759
mc16_13TeV.363624.MGPy8EG_N30NLO_Wmunu_Ht0_70_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	41422000	1.12	16720.0	0.83759
mc16_13TeV.363624.MGPy8EG_N30NLO_Wmunu_Ht0_70_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	24993000	1.12	16720.0	0.83759
mc16_13TeV.363625.MGPy8EG_N30NLO_Wmunu_Ht0_70_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	12492000	1.12	16720.0	0.1389
mc16_13TeV.363625.MGPy8EG_N30NLO_Wmunu_Ht0_70_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	16555000	1.12	16720.0	0.1389
mc16_13TeV.363625.MGPy8EG_N30NLO_Wmunu_Ht0_70_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	9993600	1.12	16720.0	0.1389
mc16_13TeV.363626.MGPy8EG_N30NLO_Wmunu_Ht0_70_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	11216950	1.12	16718.0	0.024676
mc16_13TeV.363626.MGPy8EG_N30NLO_Wmunu_Ht0_70_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	14887000	1.12	16718.0	0.024676
mc16_13TeV.363626.MGPy8EG_N30NLO_Wmunu_Ht0_70_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	8992500	1.12	16718.0	0.024676
mc16_13TeV.363627.MGPy8EG_N30NLO_Wmunu_Ht70_140_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	6818500	1.12	755.39	0.71228
mc16_13TeV.363627.MGPy8EG_N30NLO_Wmunu_Ht70_140_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	9039000	1.12	755.39	0.71228
mc16_13TeV.363627.MGPy8EG_N30NLO_Wmunu_Ht70_140_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	5476000	1.12	755.39	0.71228
mc16_13TeV.363628.MGPy8EG_N30NLO_Wmunu_Ht70_140_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	5585000	1.12	755.79	0.24172
mc16_13TeV.363628.MGPy8EG_N30NLO_Wmunu_Ht70_140_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	7419000	1.12	755.79	0.24172
mc16_13TeV.363628.MGPy8EG_N30NLO_Wmunu_Ht70_140_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	4497300	1.12	755.79	0.24172
mc16_13TeV.363629.MGPy8EG_N30NLO_Wmunu_Ht70_140_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	5309000	1.12	755.73	0.050382
mc16_13TeV.363629.MGPy8EG_N30NLO_Wmunu_Ht70_140_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	7101000	1.12	755.73	0.050382
mc16_13TeV.363629.MGPy8EG_N30NLO_Wmunu_Ht70_140_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	4271000	1.12	755.73	0.050382
mc16_13TeV.363630.MGPy8EG_N30NLO_Wmunu_Ht140_280_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	6785000	1.12	318.8	0.66598
mc16_13TeV.363630.MGPy8EG_N30NLO_Wmunu_Ht140_280_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	8958000	1.12	318.8	0.66598
mc16_13TeV.363630.MGPy8EG_N30NLO_Wmunu_Ht140_280_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	5471000	1.12	318.8	0.66598
mc16_13TeV.363631.MGPy8EG_N30NLO_Wmunu_Ht140_280_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	4985650	1.12	320.02	0.26821
mc16_13TeV.363631.MGPy8EG_N30NLO_Wmunu_Ht140_280_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	6571000	1.12	320.02	0.26821
mc16_13TeV.363631.MGPy8EG_N30NLO_Wmunu_Ht140_280_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	3797750	1.12	320.02	0.26821
mc16_13TeV.363632.MGPy8EG_N30NLO_Wmunu_Ht140_280_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	4989000	1.12	319.49	0.073364
mc16_13TeV.363632.MGPy8EG_N30NLO_Wmunu_Ht140_280_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	6571000	1.12	319.49	0.073364
mc16_13TeV.363632.MGPy8EG_N30NLO_Wmunu_Ht140_280_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	3996340	1.12	319.49	0.073364
mc16_13TeV.363633.MGPy8EG_N30NLO_Wmunu_Ht280_500_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	2488000	1.12	73.567	0.61912
mc16_13TeV.363633.MGPy8EG_N30NLO_Wmunu_Ht280_500_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	3286000	1.12	73.567	0.61912
mc16_13TeV.363633.MGPy8EG_N30NLO_Wmunu_Ht280_500_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	1993000	1.12	73.567	0.61912
mc16_13TeV.363634.MGPy8EG_N30NLO_Wmunu_Ht280_500_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	2862400	1.12	73.564	0.2864
mc16_13TeV.363634.MGPy8EG_N30NLO_Wmunu_Ht280_500_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	3753000	1.12	73.564	0.2864
mc16_13TeV.363634.MGPy8EG_N30NLO_Wmunu_Ht280_500_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	2293400	1.12	73.564	0.2864
mc16_13TeV.363635.MGPy8EG_N30NLO_Wmunu_Ht280_500_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	2702000	1.12	73.585	0.09728
mc16_13TeV.363635.MGPy8EG_N30NLO_Wmunu_Ht280_500_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	3554000	1.12	73.585	0.09728
mc16_13TeV.363635.MGPy8EG_N30NLO_Wmunu_Ht280_500_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	2198800	1.12	73.585	0.09728
mc16_13TeV.363636.MGPy8EG_N30NLO_Wmunu_Ht500_700_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	1469000	1.12	11.521	0.5871
mc16_13TeV.363636.MGPy8EG_N30NLO_Wmunu_Ht500_700_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	1934000	1.12	11.521	0.5871
mc16_13TeV.363636.MGPy8EG_N30NLO_Wmunu_Ht500_700_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	1185000	1.12	11.521	0.5871
mc16_13TeV.363637.MGPy8EG_N30NLO_Wmunu_Ht500_700_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	375000	1.12	11.519	0.29891
mc16_13TeV.363637.MGPy8EG_N30NLO_Wmunu_Ht500_700_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	492000	1.12	11.519	0.29891
mc16_13TeV.363637.MGPy8EG_N30NLO_Wmunu_Ht500_700_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	297000	1.12	11.519	0.29891
mc16_13TeV.363638.MGPy8EG_N30NLO_Wmunu_Ht500_700_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	724000	1.12	11.516	0.11659
mc16_13TeV.363638.MGPy8EG_N30NLO_Wmunu_Ht500_700_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	889000	1.12	11.516	0.11659
mc16_13TeV.363638.MGPy8EG_N30NLO_Wmunu_Ht500_700_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	590000	1.12	11.516	0.11659
mc16_13TeV.363639.MGPy8EG_N30NLO_Wmunu_Ht700_1000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	475000	1.12	4.0222	0.56688
mc16_13TeV.363639.MGPy8EG_N30NLO_Wmunu_Ht700_1000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	631000	1.12	4.0222	0.56688
mc16_13TeV.363639.MGPy8EG_N30NLO_Wmunu_Ht700_1000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	387500	1.12	4.0222	0.56688
mc16_13TeV.363640.MGPy8EG_N30NLO_Wmunu_Ht700_1000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	126000	1.12	4.0238	0.30558
mc16_13TeV.363640.MGPy8EG_N30NLO_Wmunu_Ht700_1000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	135000	1.12	4.0238	0.30558
mc16_13TeV.363640.MGPy8EG_N30NLO_Wmunu_Ht700_1000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	98200	1.12	4.0238	0.30558
mc16_13TeV.363641.MGPy8EG_N30NLO_Wmunu_Ht700_1000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	127000	1.12	4.0164	0.13021
mc16_13TeV.363641.MGPy8EG_N30NLO_Wmunu_Ht700_1000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	138000	1.12	4.0164	0.13021
mc16_13TeV.363641.MGPy8EG_N30NLO_Wmunu_Ht700_1000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	97700	1.12	4.0164	0.13021
mc16_13TeV.363642.MGPy8EG_N30NLO_Wmunu_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	60000	1.12	1.3305	0.54761
mc16_13TeV.363642.MGPy8EG_N30NLO_Wmunu_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	85000	1.12	1.3305	0.54761
mc16_13TeV.363642.MGPy8EG_N30NLO_Wmunu_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	50500	1.12	1.3305	0.54761
mc16_13TeV.363643.MGPy8EG_N30NLO_Wmunu_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	61000	1.12	1.3239	0.31071
mc16_13TeV.363643.MGPy8EG_N30NLO_Wmunu_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	86000	1.12	1.3239	0.31071
mc16_13TeV.363643.MGPy8EG_N30NLO_Wmunu_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	53400	1.12	1.3239	0.31071
mc16_13TeV.363644.MGPy8EG_N30NLO_Wmunu_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	103000	1.12	1.3264	0.14484
mc16_13TeV.363644.MGPy8EG_N30NLO_Wmunu_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	134000	1.12	1.3264	0.14484
mc16_13TeV.363644.MGPy8EG_N30NLO_Wmunu_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	88600	1.12	1.3264	0.14484
mc16_13TeV.363645.MGPy8EG_N30NLO_Wmunu_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r10201_p397	6400	1.12	0.042434	0.5250
mc16_13TeV.363645.MGPy8EG_N30NLO_Wmunu_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r10724_p397	14600	1.12	0.042434	0.5250
mc16_13TeV.363645.MGPy8EG_N30NLO_Wmunu_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r9364_p397	6800	1.12	0.042434	0.5250
mc16_13TeV.363646.MGPy8EG_N30NLO_Wmunu_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r10201_p397	8100	1.12	0.042132	0.31907
mc16_13TeV.363646.MGPy8EG_N30NLO_Wmunu_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r10724_p397	14400	1.12	0.042132	0.31907
mc16_13TeV.363646.MGPy8EG_N30NLO_Wmunu_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r9364_p397	7100	1.12	0.042132	0.31907
mc16_13TeV.363647.MGPy8EG_N30NLO_Wmunu_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e5136_s3126_r10201_p397	7400	1.12	0.042097	0.16318
mc16_13TeV.363647.MGPy8EG_N30NLO_Wmunu_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e5136_s3126_r10724_p397	15550	1.12	0.042097	0.16318
mc16_13TeV.363647.MGPy8EG_N30NLO_Wmunu_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e5136_s3126_r9364_p397	7700	1.12	0.042097	0.16318
mc16_13TeV.363648.MGPy8EG_N30NLO_Wtaunu_Ht0_70_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	31221000	1.12	16702.0	0.83756
mc16_13TeV.363648.MGPy8EG_N30NLO_Wtaunu_Ht0_70_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	37623000	1.12	16702.0	0.83756
mc16_13TeV.363648.MGPy8EG_N30NLO_Wtaunu_Ht0_70_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	24991000	1.12	16702.0	0.83756
mc16_13TeV.363649.MGPy8EG_N30NLO_Wtaunu_Ht0_70_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	12416000	1.12	16703.0	0.13889
mc16_13TeV.363649.MGPy8EG_N30NLO_Wtaunu_Ht0_70_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	15249000	1.12	16703.0	0.13889
mc16_13TeV.363649.MGPy8EG_N30NLO_Wtaunu_Ht0_70_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	9997600	1.12	16703.0	0.13889
mc16_13TeV.363650.MGPy8EG_N30NLO_Wtaunu_Ht0_70_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	10820600	1.12	16702.0	0.024679
mc16_13TeV.363650.MGPy8EG_N30NLO_Wtaunu_Ht0_70_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	12906000	1.12	16702.0	0.024679
mc16_13TeV.363650.MGPy8EG_N30NLO_Wtaunu_Ht0_70_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	8726550	1.12	16702.0	0.024679

Table A.3: List of Monte Carlo samples for $W + jets$ process

Dataset	nEvents	kFactor	xsec	genFilterEff
mc16_13TeV.363651.MGPy8EG_N30NLO_Wtaunu_Ht70_140_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	5253000	1.12	755.08	0.71247
mc16_13TeV.363651.MGPy8EG_N30NLO_Wtaunu_Ht70_140_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	6131000	1.12	755.08	0.71247
mc16_13TeV.363651.MGPy8EG_N30NLO_Wtaunu_Ht70_140_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	4591500	1.12	755.08	0.71247
mc16_13TeV.363652.MGPy8EG_N30NLO_Wtaunu_Ht70_140_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	5366000	1.12	755.24	0.24122
mc16_13TeV.363652.MGPy8EG_N30NLO_Wtaunu_Ht70_140_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	5960000	1.12	755.24	0.24122
mc16_13TeV.363652.MGPy8EG_N30NLO_Wtaunu_Ht70_140_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	4392700	1.12	755.24	0.24122
mc16_13TeV.363653.MGPy8EG_N30NLO_Wtaunu_Ht70_140_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	3894000	1.12	755.14	0.05052
mc16_13TeV.363653.MGPy8EG_N30NLO_Wtaunu_Ht70_140_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	5958000	1.12	755.14	0.05052
mc16_13TeV.363653.MGPy8EG_N30NLO_Wtaunu_Ht70_140_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	4332000	1.12	755.14	0.05052
mc16_13TeV.363654.MGPy8EG_N30NLO_Wtaunu_Ht140_280_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	4129000	1.12	315.27	0.667
mc16_13TeV.363654.MGPy8EG_N30NLO_Wtaunu_Ht140_280_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	4106000	1.12	315.27	0.667
mc16_13TeV.363654.MGPy8EG_N30NLO_Wtaunu_Ht140_280_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	4040000	1.12	315.27	0.667
mc16_13TeV.363655.MGPy8EG_N30NLO_Wtaunu_Ht140_280_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	4705650	1.12	315.97	0.26829
mc16_13TeV.363655.MGPy8EG_N30NLO_Wtaunu_Ht140_280_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	5810600	1.12	315.97	0.26829
mc16_13TeV.363655.MGPy8EG_N30NLO_Wtaunu_Ht140_280_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	3841900	1.12	315.97	0.26829
mc16_13TeV.363656.MGPy8EG_N30NLO_Wtaunu_Ht140_280_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	3432000	1.12	315.63	0.073217
mc16_13TeV.363656.MGPy8EG_N30NLO_Wtaunu_Ht140_280_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	3672000	1.12	315.63	0.073217
mc16_13TeV.363656.MGPy8EG_N30NLO_Wtaunu_Ht140_280_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	2876000	1.12	315.63	0.073217
mc16_13TeV.363657.MGPy8EG_N30NLO_Wtaunu_Ht280_500_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	1278500	1.12	72.217	0.62227
mc16_13TeV.363657.MGPy8EG_N30NLO_Wtaunu_Ht280_500_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	1165000	1.12	72.217	0.62227
mc16_13TeV.363657.MGPy8EG_N30NLO_Wtaunu_Ht280_500_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	1583500	1.12	72.217	0.62227
mc16_13TeV.363658.MGPy8EG_N30NLO_Wtaunu_Ht280_500_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	1063000	1.12	72.227	0.28577
mc16_13TeV.363658.MGPy8EG_N30NLO_Wtaunu_Ht280_500_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	1161000	1.12	72.227	0.28577
mc16_13TeV.363658.MGPy8EG_N30NLO_Wtaunu_Ht280_500_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	1233200	1.12	72.227	0.28577
mc16_13TeV.363659.MGPy8EG_N30NLO_Wtaunu_Ht280_500_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	1596550	1.12	72.209	0.09744
mc16_13TeV.363659.MGPy8EG_N30NLO_Wtaunu_Ht280_500_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	1685000	1.12	72.209	0.09744
mc16_13TeV.363659.MGPy8EG_N30NLO_Wtaunu_Ht280_500_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	1587950	1.12	72.209	0.09744
mc16_13TeV.363660.MGPy8EG_N30NLO_Wtaunu_Ht500_700_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	404000	1.12	11.445	0.58928
mc16_13TeV.363660.MGPy8EG_N30NLO_Wtaunu_Ht500_700_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	499000	1.12	11.445	0.58928
mc16_13TeV.363660.MGPy8EG_N30NLO_Wtaunu_Ht500_700_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	594500	1.12	11.445	0.58928
mc16_13TeV.363661.MGPy8EG_N30NLO_Wtaunu_Ht500_700_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	243000	1.12	11.451	0.29837
mc16_13TeV.363661.MGPy8EG_N30NLO_Wtaunu_Ht500_700_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	287000	1.12	11.451	0.29837
mc16_13TeV.363661.MGPy8EG_N30NLO_Wtaunu_Ht500_700_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	283000	1.12	11.451	0.29837
mc16_13TeV.363662.MGPy8EG_N30NLO_Wtaunu_Ht500_700_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	142000	1.12	11.457	0.11742
mc16_13TeV.363662.MGPy8EG_N30NLO_Wtaunu_Ht500_700_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	190000	1.12	11.457	0.11742
mc16_13TeV.363662.MGPy8EG_N30NLO_Wtaunu_Ht500_700_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	198300	1.12	11.457	0.11742
mc16_13TeV.363663.MGPy8EG_N30NLO_Wtaunu_Ht700_1000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	107000	1.12	4.0205	0.56858
mc16_13TeV.363663.MGPy8EG_N30NLO_Wtaunu_Ht700_1000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	145000	1.12	4.0205	0.56858
mc16_13TeV.363663.MGPy8EG_N30NLO_Wtaunu_Ht700_1000_CVetoBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	159500	1.12	4.0205	0.56858
mc16_13TeV.363664.MGPy8EG_N30NLO_Wtaunu_Ht700_1000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	72000	1.12	4.0369	0.30442
mc16_13TeV.363664.MGPy8EG_N30NLO_Wtaunu_Ht700_1000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	88000	1.12	4.0369	0.30442
mc16_13TeV.363664.MGPy8EG_N30NLO_Wtaunu_Ht700_1000_CFilterBVeto.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	85400	1.12	4.0369	0.30442
mc16_13TeV.363665.MGPy8EG_N30NLO_Wtaunu_Ht700_1000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10201_p397	62000	1.12	4.0252	0.1309
mc16_13TeV.363665.MGPy8EG_N30NLO_Wtaunu_Ht700_1000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r10724_p397	57000	1.12	4.0252	0.1309
mc16_13TeV.363665.MGPy8EG_N30NLO_Wtaunu_Ht700_1000_BFilter.deriv.DAOD_STDM3.e4944_s3126_r9364_p397	83000	1.12	4.0252	0.1309
mc16_13TeV.363666.MGPy8EG_N30NLO_Wtaunu_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r10201_p397	51000	1.12	1.3201	0.54623
mc16_13TeV.363666.MGPy8EG_N30NLO_Wtaunu_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r10724_p397	77000	1.12	1.3201	0.54623
mc16_13TeV.363666.MGPy8EG_N30NLO_Wtaunu_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r9364_p397	47000	1.12	1.3201	0.54623
mc16_13TeV.363667.MGPy8EG_N30NLO_Wtaunu_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r10201_p397	52000	1.12	1.322	0.31142
mc16_13TeV.363667.MGPy8EG_N30NLO_Wtaunu_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r10724_p397	81000	1.12	1.322	0.31142
mc16_13TeV.363667.MGPy8EG_N30NLO_Wtaunu_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r9364_p397	47800	1.12	1.322	0.31142
mc16_13TeV.363668.MGPy8EG_N30NLO_Wtaunu_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e5136_s3126_r10201_p397	54000	1.12	1.3227	0.14439
mc16_13TeV.363668.MGPy8EG_N30NLO_Wtaunu_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e5136_s3126_r10724_p397	69000	1.12	1.3227	0.14439
mc16_13TeV.363668.MGPy8EG_N30NLO_Wtaunu_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e5136_s3126_r9364_p397	47000	1.12	1.3227	0.14439
mc16_13TeV.363669.MGPy8EG_N30NLO_Wtaunu_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r10201_p397	5800	1.12	0.04293	0.5286
mc16_13TeV.363669.MGPy8EG_N30NLO_Wtaunu_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r10724_p397	10000	1.12	0.04293	0.5286
mc16_13TeV.363669.MGPy8EG_N30NLO_Wtaunu_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e5136_s3126_r9364_p397	3800	1.12	0.04293	0.5286
mc16_13TeV.363670.MGPy8EG_N30NLO_Wtaunu_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r10201_p397	5800	1.12	0.042686	0.31633
mc16_13TeV.363670.MGPy8EG_N30NLO_Wtaunu_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r10724_p397	12500	1.12	0.042686	0.31633
mc16_13TeV.363670.MGPy8EG_N30NLO_Wtaunu_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e5136_s3126_r9364_p397	5500	1.12	0.042686	0.31633
mc16_13TeV.363671.MGPy8EG_N30NLO_Wtaunu_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e5136_s3126_r10201_p397	6300	1.12	0.042754	0.16623
mc16_13TeV.363671.MGPy8EG_N30NLO_Wtaunu_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e5136_s3126_r10724_p397	11350	1.12	0.042754	0.16623
mc16_13TeV.363671.MGPy8EG_N30NLO_Wtaunu_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e5136_s3126_r9364_p397	6050	1.12	0.042754	0.16623

Table A.4: List of Monte Carlo samples for $t\bar{t}$ process

Dataset	nEvents	kFactor	xsec	genFilterEff
mc16_13TeV.410470.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad.deriv.DAOD_STDM3.e6337_s3126_r10201_p397	149327000	1.13975636159	729.77	0.54382
mc16_13TeV.410470.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad.deriv.DAOD_STDM3.e6337_s3126_r10724_p397	150740000	1.13975636159	729.77	0.54382
mc16_13TeV.410470.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad.deriv.DAOD_STDM3.e6337_s3126_r9364_p397	119432000	1.13975636159	729.77	0.54382

Appendix B

Object and Event Selection

B.1 Selection criteria

Table B.1: Event selection cuts used to define the Signal region (CR) for WWss scattering (Ref.[18])

Exactly two Analysis leptons with identical electrical charge No Veto lepton
Electrons have to pass ECIDS
ee-channel: $\|\eta\| < 1.37$, $\|m_{ee} - m_Z\| > 15GeV$
$m_{ll'} > 20GeV$
$E_T^{miss} > 30GeV$
At least two jets Leading jet: $p_T > 65GeV$ Subleading jet: $p_T > 35GeV$
b-jet veto
$\|\Delta y_{jj}\| > 2$
$m_{jj} > 500GeV$

Table B.2: Different object level selections for muons in the Control-Region.(Ref.[18])

	Veto	Non-Analysis	Analysis
ID	Loose	Medium	Medium
p_T	$p_T > 6GeV$	$p_T > 27GeV$	$p_T > 27GeV$
η	$\|\eta\| < 2.7$	$\|\eta\| < 2.5$	$\|\eta\| < 2.5$
$\|z_0 \cdot \sin\theta\|$	$\|z_0 \cdot \sin\theta\| < 0.5mm$	$\|z_0 \cdot \sin\theta\| < 0.5mm$	$\|z_0 \cdot \sin\theta\| < 0.5mm$
$\frac{d_0}{\sigma_{d_0}}$	$\frac{d_0}{\sigma_{d_0}} < 3$	$\frac{d_0}{\sigma_{d_0}} < 10$	$\frac{d_0}{\sigma_{d_0}} < 3$
Isolation	FCTight	-	FCTight
Additional requirements	Not the leading or subleading Ana/Non-Ana lepton	Fail Gradient Isolation or $\frac{d_0}{\sigma_{d_0}} < 3$	-

Table B.3: Different object level selections for electrons in the Control-Region.(Ref.[18])

	Veto	Non-Analysis	Analysis
ID	$p_T < 25GeV$: Medium $p_T > 25GeV$: Tight	MediumLH	TightLH
p_T	$p_T > 6GeV$	$p_T > 27GeV$	$p_T > 27GeV$
η	$\|\eta\| < 2.47$	$\|\eta\| < 2.47$, excluding $1.37 < \|\eta\| < 1.52$	$\|\eta\| < 2.47$, excluding $1.37 < \|\eta\| < 1.52$
$\|z_0 \cdot \sin\theta\|$	$\|z_0 \cdot \sin\theta\| < 0.5mm$	$\|z_0 \cdot \sin\theta\| < 0.5mm$	$\|z_0 \cdot \sin\theta\| < 0.5mm$
$\frac{d_0}{\sigma_{d_0}}$	$\frac{d_0}{\sigma_{d_0}} < 5$	$\frac{d_0}{\sigma_{d_0}} < 5$	$\frac{d_0}{\sigma_{d_0}} < 5$
Isolation	Gradient	-	FCTight
Author	-	-	1
Additional requirements	Not the leading or subleading Ana/Non-Ana lepton	Fail Gradient Isolation or TightLH Identification	-

B.2 Additional Fit Results for η Correction

Table B.4: Fit results η-correction for $t\bar{t}$from section 6.4, including the covariance and correlation matrix (note: $\mp$ denotes the sign for the respective linear fit, meaning upper sign [lower sign] for $\eta < -1.5$ [$\eta > 1.5$])

ℓ	Fit	Parameter	$\frac{\chi^2}{NDF}$	Covariance: $\begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ba} & \sigma_b^2 \end{pmatrix}$	Correlation: $\begin{pmatrix} \rho_{aa} & \rho_{ab} \\ \rho_{ba} & \rho_{bb} \end{pmatrix}$
e	Linear	a = 1.20 ± 0.04 b = -0.25 ± 0.14	0.004	$\begin{pmatrix} 0.001583 & \mp 0.00225 \\ \mp 0.00225 & 0.01926 \end{pmatrix}$	$\begin{pmatrix} 1.0 & \mp 0.4076 \\ \mp 0.4076 & 1.0 \end{pmatrix}$
e	Parabolic	a = 0.744 ± 0.018 b = 0.329 ± 0.026	1.524	$\begin{pmatrix} 0.00033 & -0.00033 \\ -0.00033 & 0.0007 \end{pmatrix}$	$\begin{pmatrix} 1.0 & -0.7003 \\ -0.7003 & 1.0 \end{pmatrix}$
μ	Parabolic	a = 0.919 ± 0.016 b = 0.047 ± 0.007	0.497	$\begin{pmatrix} 0.00027 & -0.00008 \\ -0.00008 & 0.00005 \end{pmatrix}$	$\begin{pmatrix} 1.0 & -0.704 \\ -0.704 & 1.0 \end{pmatrix}$

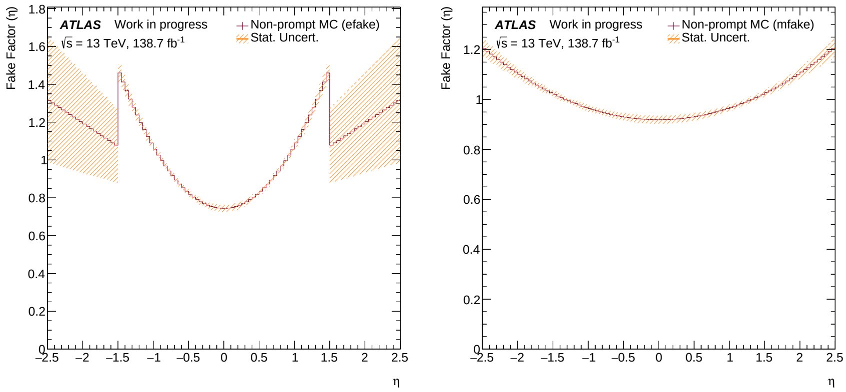


Figure B.1: Histograms (100 bins) of fitted η correction for the $t\bar{t}$ process, with the calculated statistical uncertainties (CL $= 1\,\sigma$). This histogram is used for the fake lepton background estimation during the Monte Carlo Closure Test. The correction for electrons (muons) is displayed on the left (right)

B.3 Fake Lepton Distributions

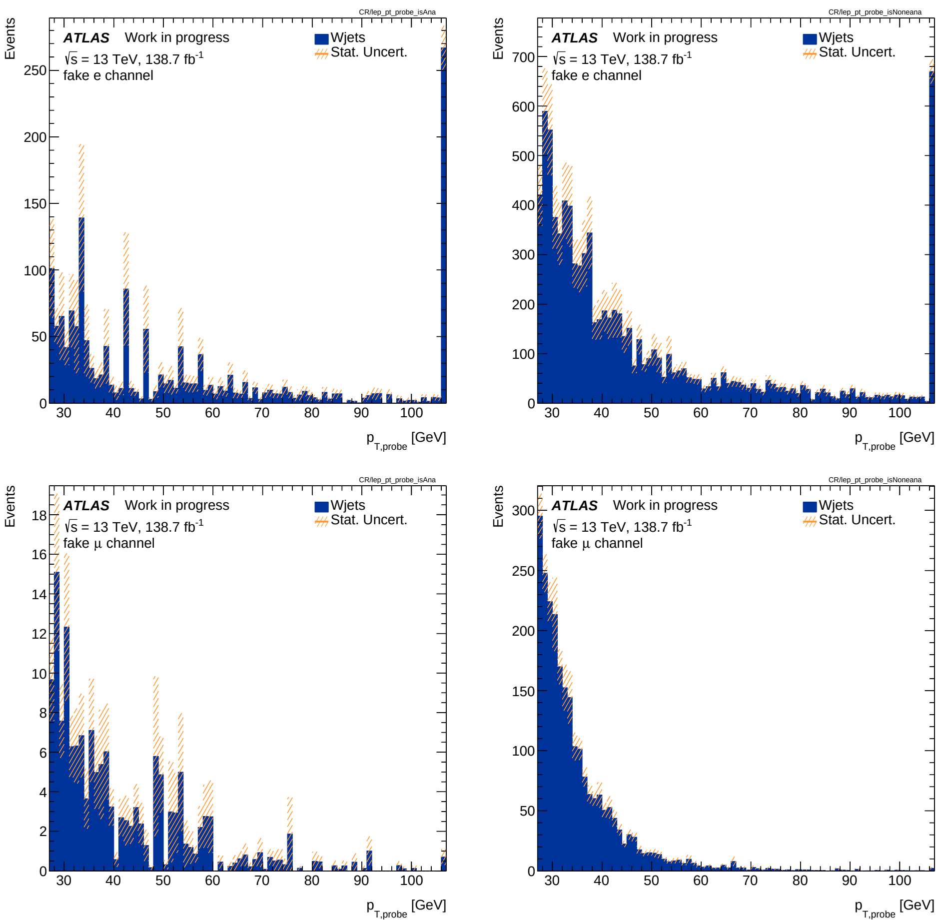


Figure B.2: p_{T} distributions of the fake lepton being Ana (Nonana) selected for the $W + jets$ process are shown on the left (right). The upper (lower) row corresponds to the fake lepton being an electron (muon). The displayed statistical uncertainties are obtained from Monte Carlo simulation.

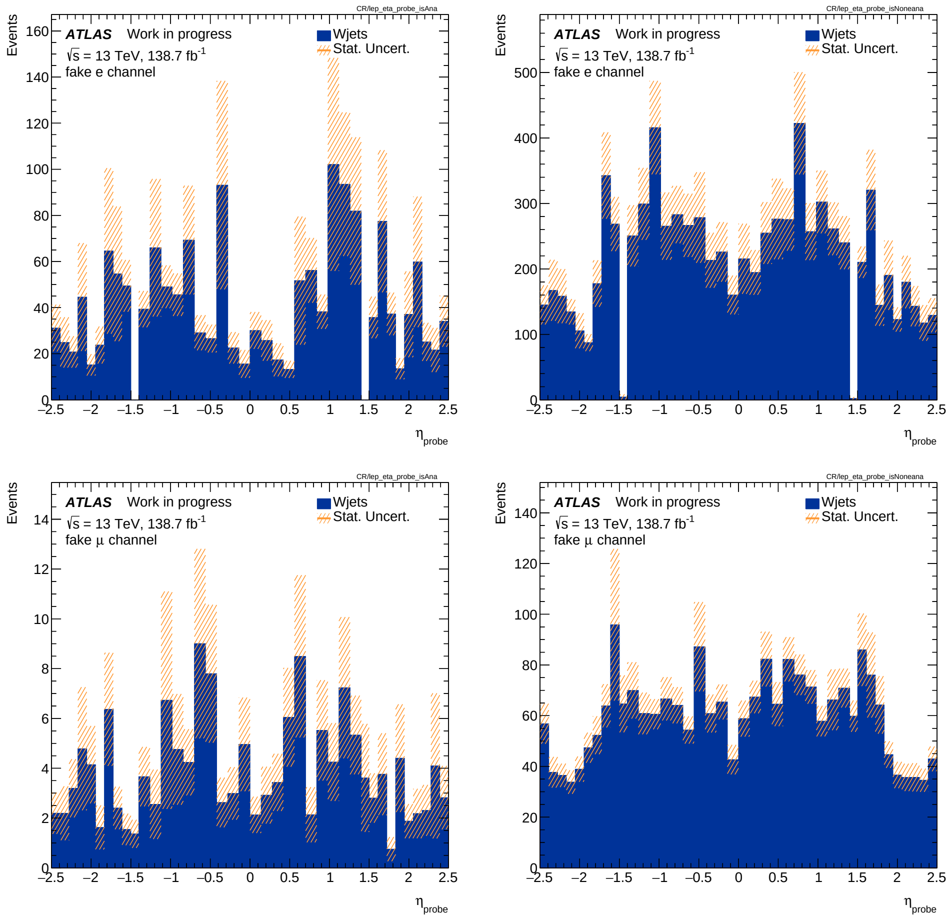


Figure B.3: η distributions of the fake lepton being Ana (Nonana) selected for the $W + jets$ process are shown on the left (right). The upper (lower) row corresponds to the fake lepton being an electron (muon). The displayed statistical uncertainties are obtained from Monte Carlo simulation.

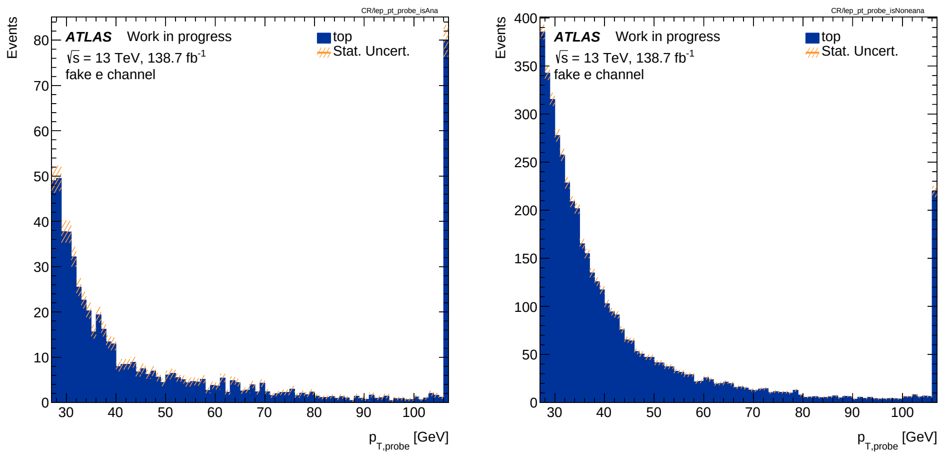


Figure B.4: p_{T} distributions of the fake electron being Ana (Nonana) selected for the $W + jets$ process are shown on the left (right). The displayed statistical uncertainties are obtained from Monte Carlo simulation.

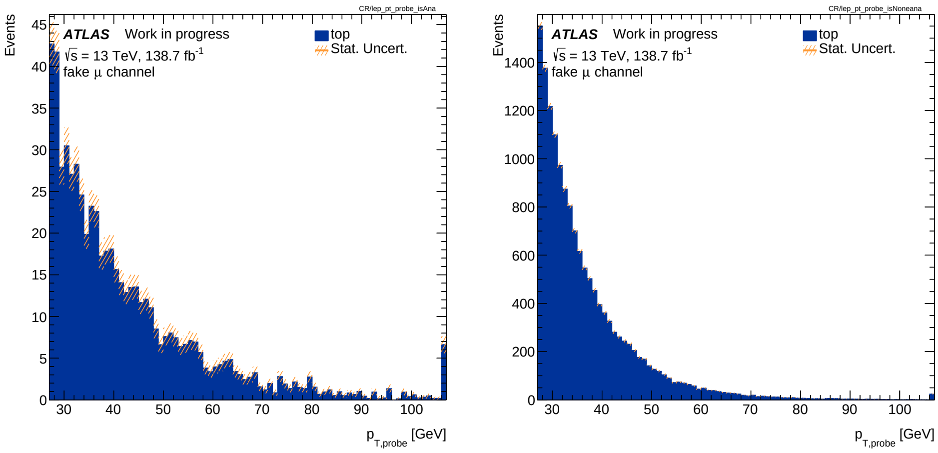


Figure B.5: fake lepton p_T distributions for $t\bar{t}$

Figure B.6: p_{T} distributions of the fake muon being Ana (Nonana) selected for the $t\bar{t}$ process are shown on the left (right). The displayed statistical uncertainties are obtained from Monte Carlo simulation.

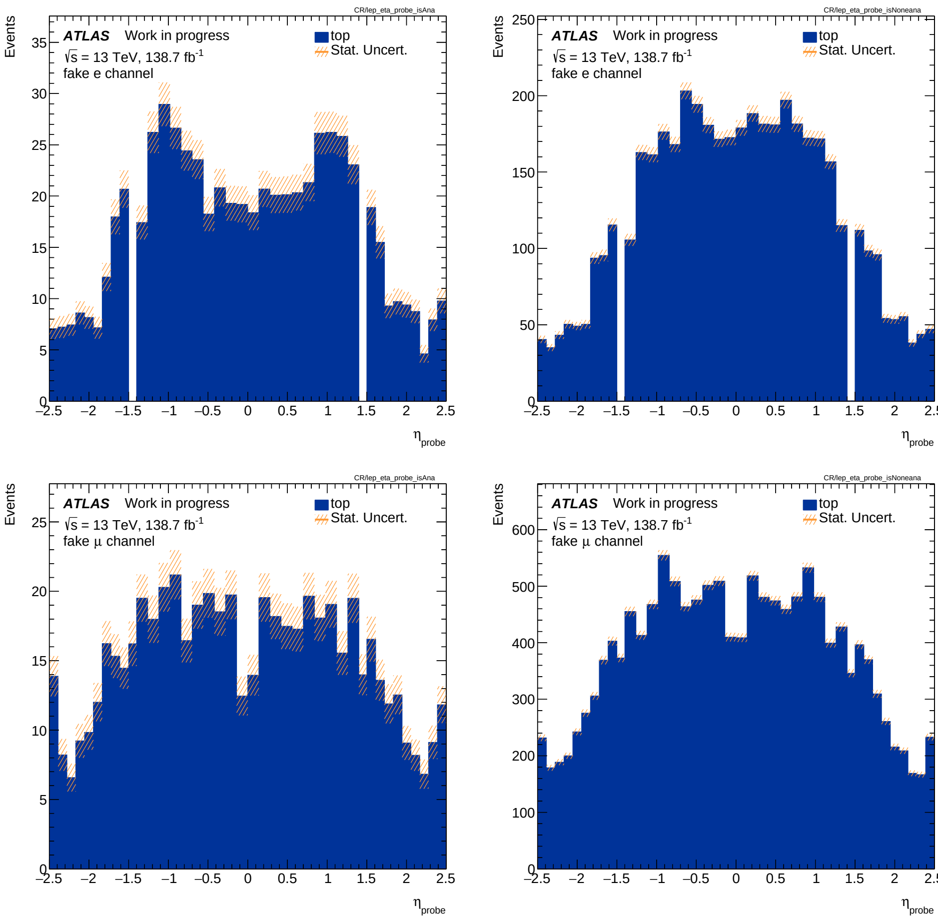


Figure B.7: η distributions of the fake lepton being Ana (Nonana) selected for the $t\bar{t}$ process are shown on the left (right). The upper (lower) row corresponds to the fake lepton being an electron (muon). The displayed statistical uncertainties are obtained from Monte Carlo simulation.

Appendix C

Monte-Carlo Closure Test

C.1 Additional Result Tables

Table C.1: Overview of number of events from integrated Monte-Carlo Closure Test with the fake factor binned in η for different lepton flavour channels in W+*jets*, only statistical uncertainties are observed. The P and F denotes Prompt and Fake, while A and N denotes Ana and Nonana selected leptons.

	ee	$e\mu$	μe	$\mu\mu$	all
MC PF	674.1 ± 80	472.5 ± 50	557.1 ± 80	88.4 ± 9	1792.0 ± 120
MC FP	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Datadriven AN	631.7 ± 33	419.2 ± 22	662.8 ± 40	78.3 ± 3	1792.0 ± 60
Datadriven NA	0.0 ± 0.0	0.0 ± 2	0.0 ± 3	0.0 ± 0.0	0.0 ± 4
total MC	674.1 ± 80	472.5 ± 50	557.1 ± 80	88.4 ± 9	1792.0 ± 120
Datadriven	631.7 ± 33	419.2 ± 22	662.8 ± 40	78.3 ± 3	1792.0 ± 60

Table C.2: Overview of number of events from integrated Monte-Carlo Closure Test with the fake factor binned in η for different lepton flavour channels in $t\bar{t}$, only statistical uncertainties are observed. The P and F denotes Prompt and Fake, while A and N denotes Ana and Nonana selected leptons.

	ee	$e\mu$	μe	$\mu\mu$	all
MC PF	260.3 ± 6	324.2 ± 7	319.8 ± 7	332.0 ± 7	1236.3 ± 14
MC FP	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Datadriven AN	270.1 ± 2	334.0 ± 2	316.0 ± 2	316.3 ± 1	1236.3 ± 4
Datadriven NA	0.0 ± 0.0	0.0 ± 0.33	0.0 ± 0.32	0.0 ± 0.0	0.0 ± 0.5
total MC	260.3 ± 6	324.2 ± 7	319.8 ± 7	332.0 ± 7	1236.3 ± 14
Datadriven	270.1 ± 2	334.0 ± 2	316.0 ± 2	316.3 ± 1	1236.3 ± 4

Table C.3: Overview of number of events from integrated Monte-Carlo Closure Test, with p_{T} fake factor and normalized η fake factor applied together, for different lepton flavour channels in W+*jets*, only statistical uncertainties are observed. The P and F denotes Prompt and Fake, while A and N denotes Ana and Nonana selected leptons.

	ee	$e\mu$	μe	$\mu\mu$	all
MC PF	674.1 ± 80	472.5 ± 50	557.1 ± 80	88.4 ± 9	1792.0 ± 120
MC FP	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Datadriven AN	643.5 ± 30	463.9 ± 21	617.6 ± 34	78.4 ± 2	1803.4 ± 50
Datadriven NA	0.0 ± 0.0	0.0 ± 2	0.0 ± 3	0.0 ± 0.0	0.0 ± 4
total MC	674.1 ± 80	472.5 ± 50	557.1 ± 80	88.4 ± 9	1792.0 ± 120
Datadriven	643.5 ± 30	463.9 ± 21	617.6 ± 34	78.4 ± 2	1803.4 ± 50

C.2 Additional η-Distributions from Closure Test

C.2.1 Closure Test with p_{T} binned Fake Factor in $Wjets$

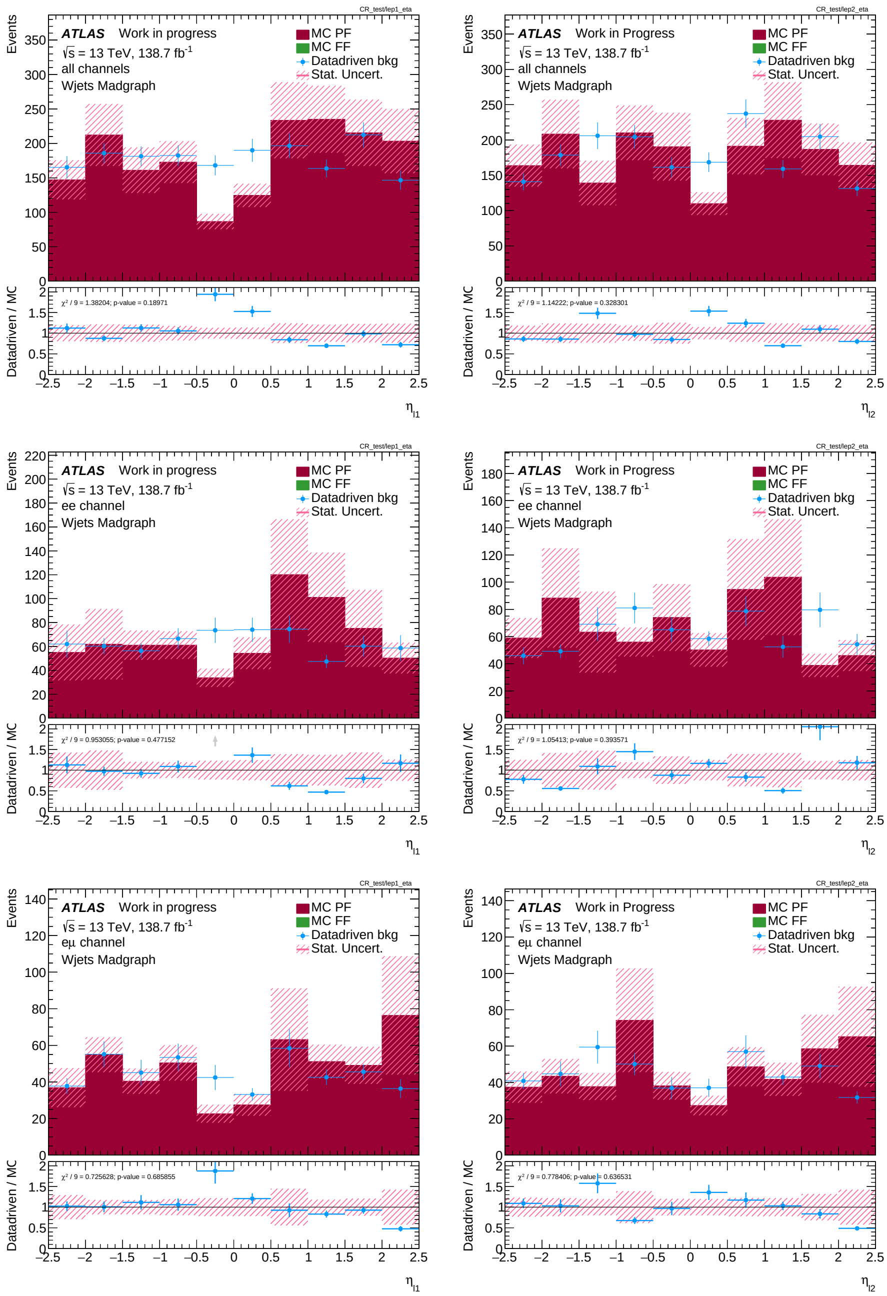


Figure C.1: Event distributions for pseudorapidity of leading (subleading) leptons on the left (right) in W+$jets$. The channel description always refers to the leading and then subleading lepton flavour, except for the sum of all channels. "Prompt+Fake" events predicted by Monte Carlo (red) are compared with the data-driven estimation (blue) from the fake factor, binned in the fake leptons p_{T}, including a ratio plot. The respective statistical uncertainties are shown as well.

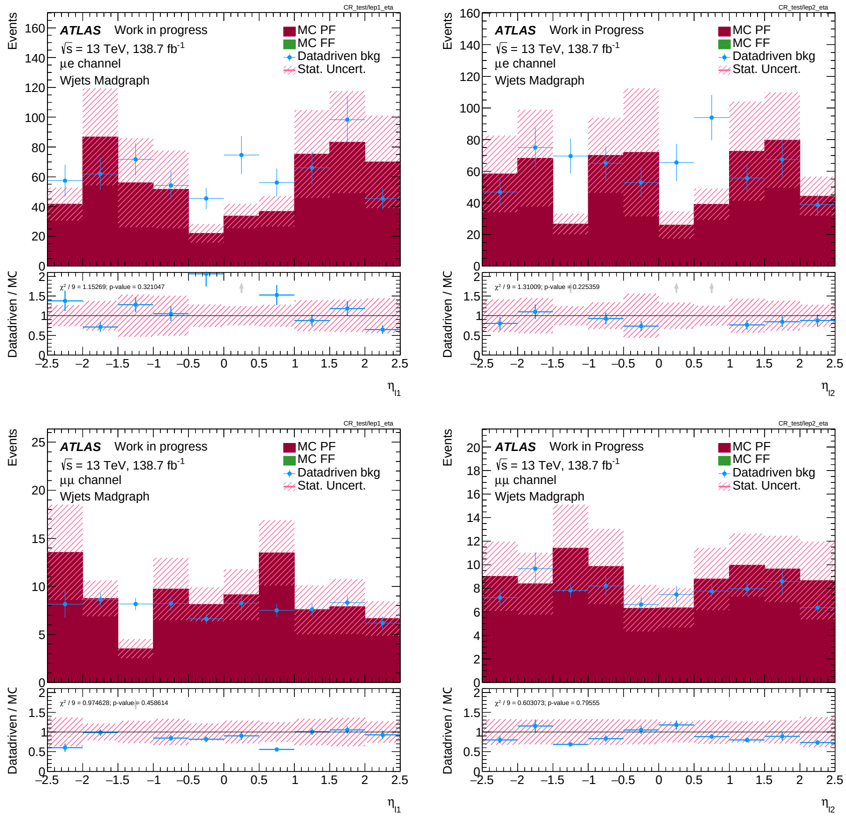


Figure C.2: Event distributions for pseudorapidity of leading (subleading) leptons on the left (right) in *W+jets*. The channel description always refers to the leading and then subleading lepton flavour, except for the sum of all channels. "Prompt+Fake" events predicted by Monte Carlo (red) are compared with the data-driven estimation (blue) from the fake factor, binned in the fake leptons p_{T}, including a ratio plot. The respective statistical uncertainties are shown as well.

C.2.2 Closure Test with η binned Fake Factor in $W jets$

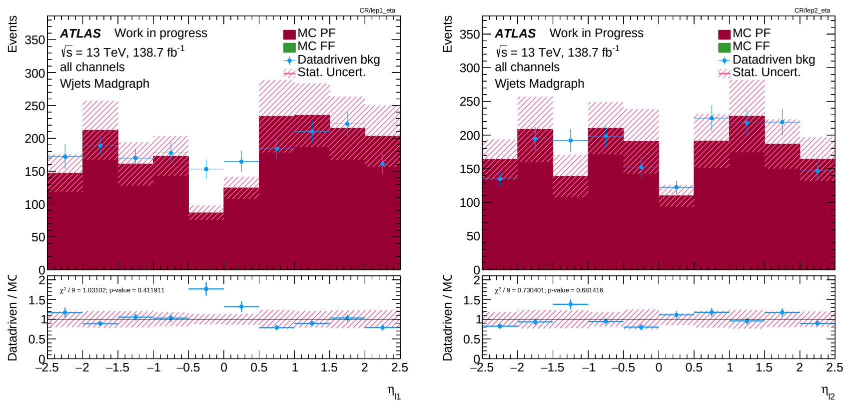


Figure C.3: Event distributions for pseudorapidity of leading (subleading) leptons on the left (right) in *W+jets*. The channel description always refers to the leading and then subleading lepton flavour, except for the sum of all channels. "Prompt+Fake" events predicted by Monte Carlo (red) are compared with the data-driven estimation (blue) from the fake factor, binned in the fake leptons η, including a ratio plot. The respective statistical uncertainties are shown as well.

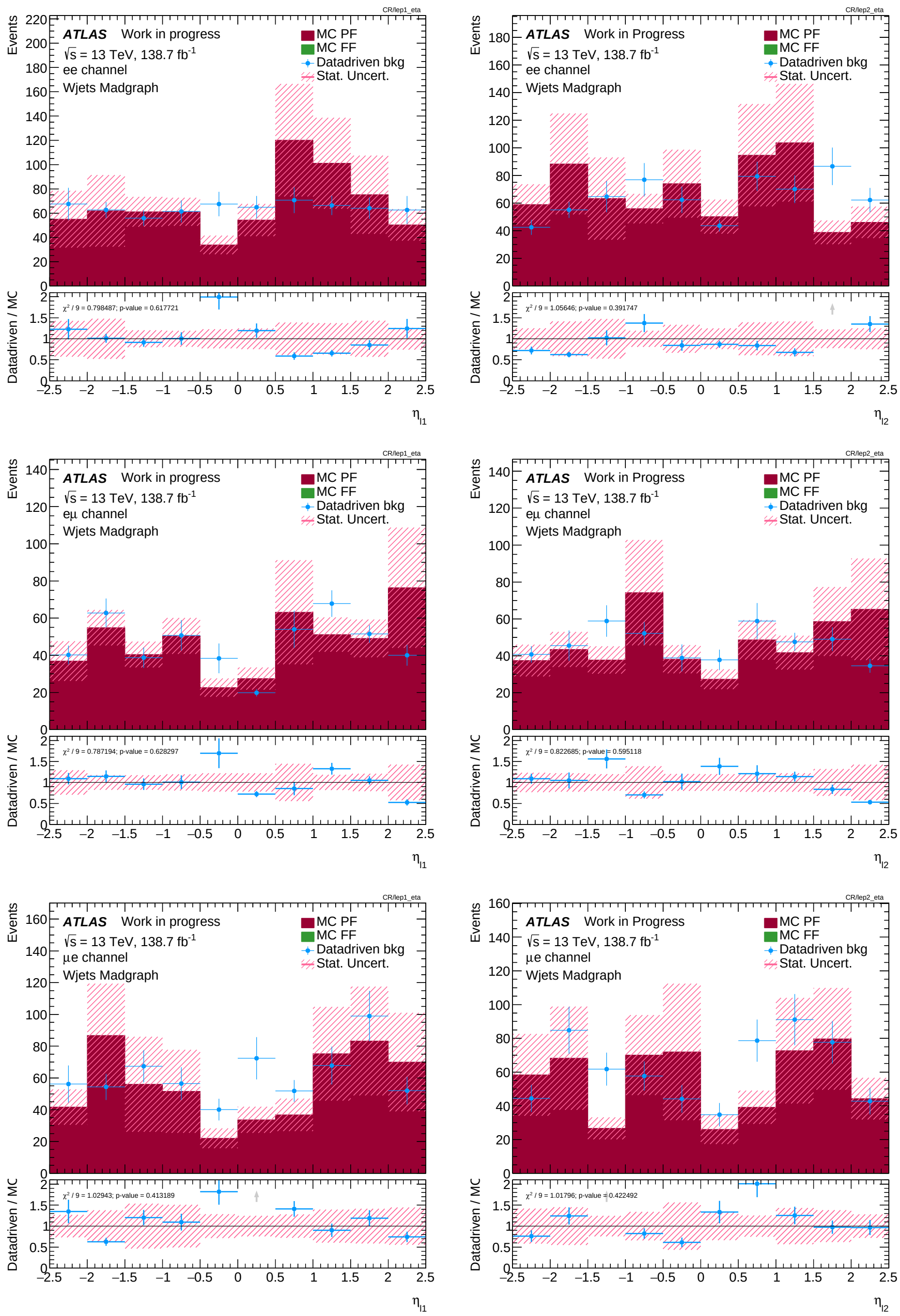


Figure C.4: Event distributions for pseudorapidity of leading (subleading) leptons on the left (right) in W+$jets$. The channel description always refers to the leading and then subleading lepton flavour, except for the sum of all channels. "Prompt+Fake" events predicted by Monte Carlo (red) are compared with the data-driven estimation (blue) from the fake factor, binned in the fake leptons η, including a ratio plot. The respective statistical uncertainties are shown as well.

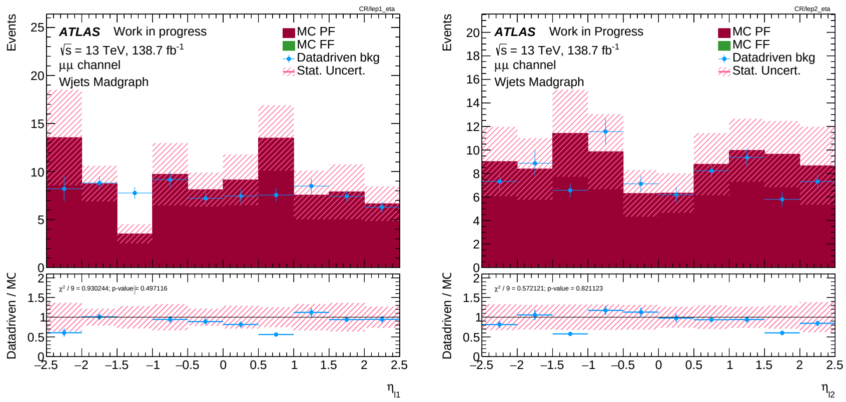


Figure C.5: Event distributions for pseudorapidity of leading (subleading) leptons on the left (right) in *W+jets*. The channel description always refers to the leading and then subleading lepton flavour, except for the sum of all channels. "Prompt+Fake" events predicted by Monte Carlo (red) are compared with the data-driven estimation (blue) from the fake factor, binned in the fake leptons η, including a ratio plot. The respective statistical uncertainties are shown as well.

C.2.3 Closure Test with combined p_T and η Fake Factor in $W jets$

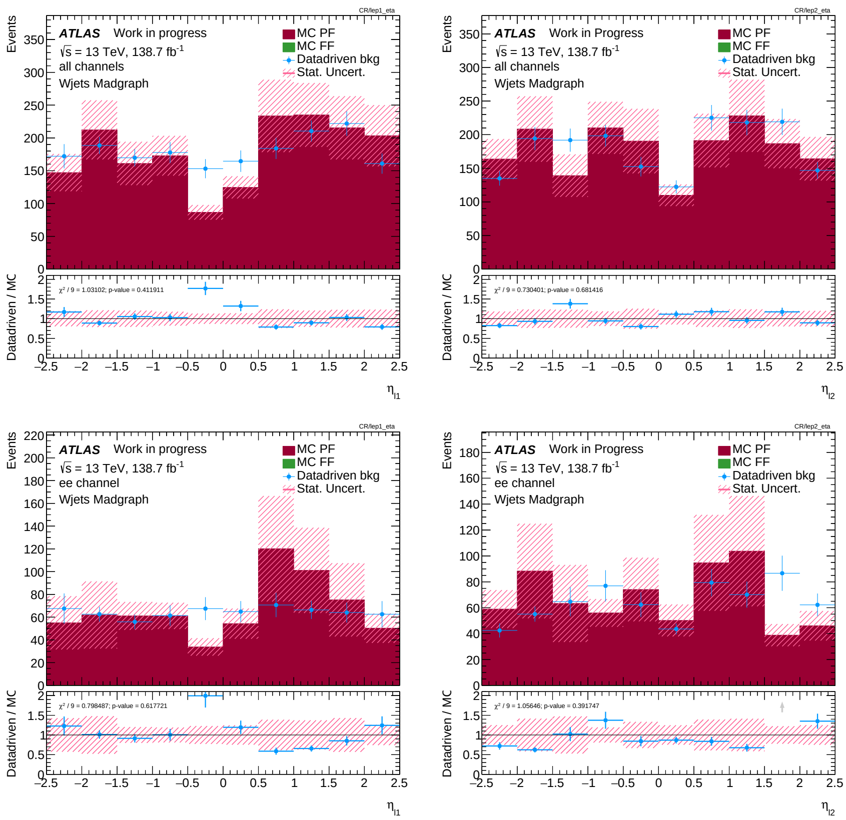


Figure C.6: Event distributions for pseudorapidity of leading (subleading) leptons on the left (right) in *W+jets*. The channel description always refers to the leading and then subleading lepton flavour, except for the sum of all channels. "Prompt+Fake" events predicted by Monte Carlo (red) are compared with the data-driven estimation (blue) from the combined fake factors, binned in the fake leptons p_T or η, including a ratio plot. The respective statistical uncertainties are shown as well.

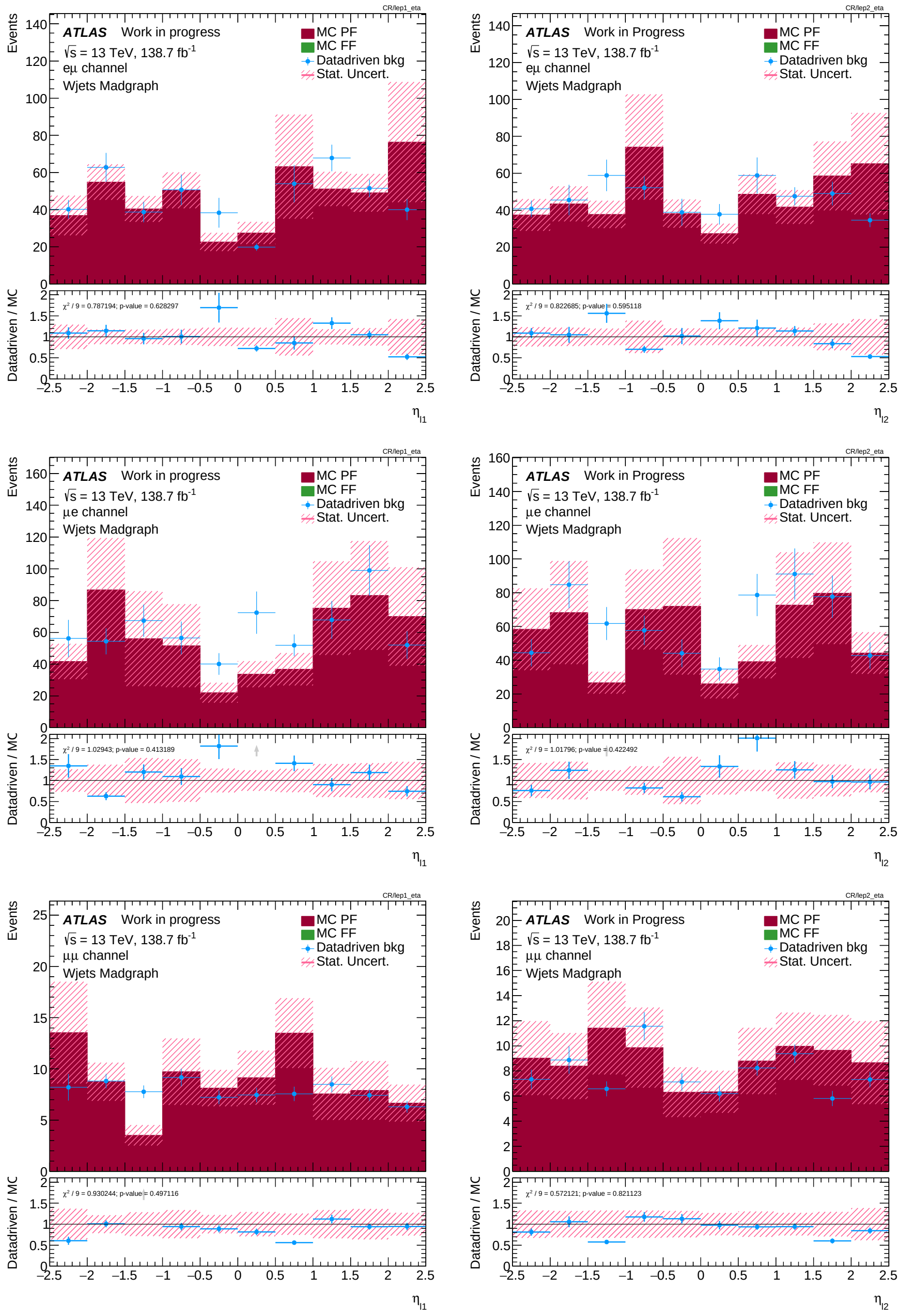


Figure C.7: Event distributions for pseudorapidity of leading (subleading) leptons on the left (right) in W+$jets$. The channel description always refers to the leading and then subleading lepton flavour, except for the sum of all channels. "Prompt+Fake" events predicted by Monte Carlo (red) are compared with the data-driven estimation (blue) from the combined fake factors, binned in the fake leptons p_T or η, including a ratio plot. The respective statistical uncertainties are shown as well.

C.2.4 Closure Test with η binned Fake Factor in $t\bar{t}$

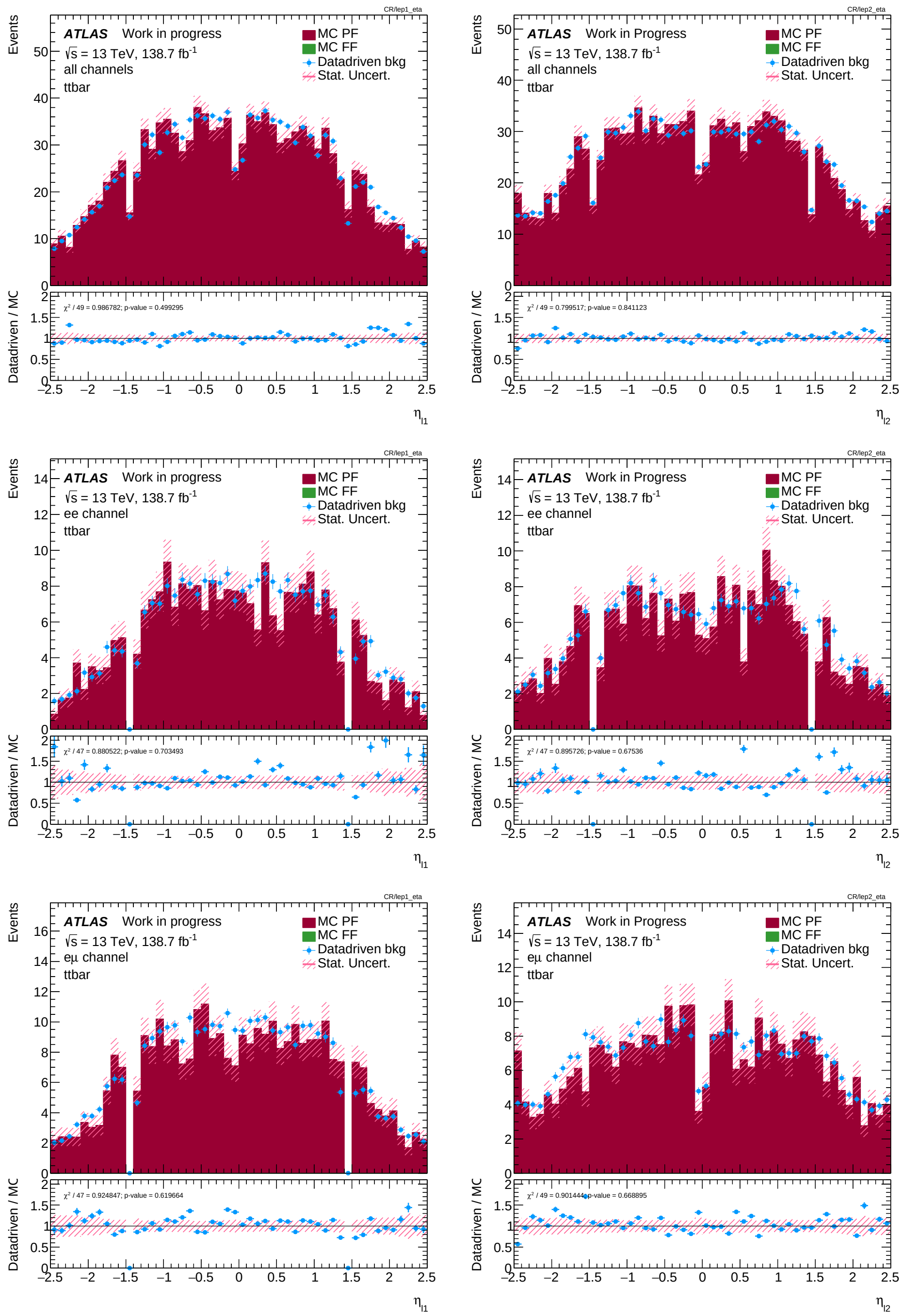


Figure C.8: Event distributions for pseudorapidity of leading (subleading) leptons on the left (right) in $t\bar{t}$. The channel description always refers to the leading and then subleading lepton flavour, except for the sum of all channels. "Prompt+Fake" events predicted by Monte Carlo (red) are compared with the data-driven estimation (blue) from the fake factor, binned in the fake leptons η, including a ratio plot. The respective statistical uncertainties are shown as well.

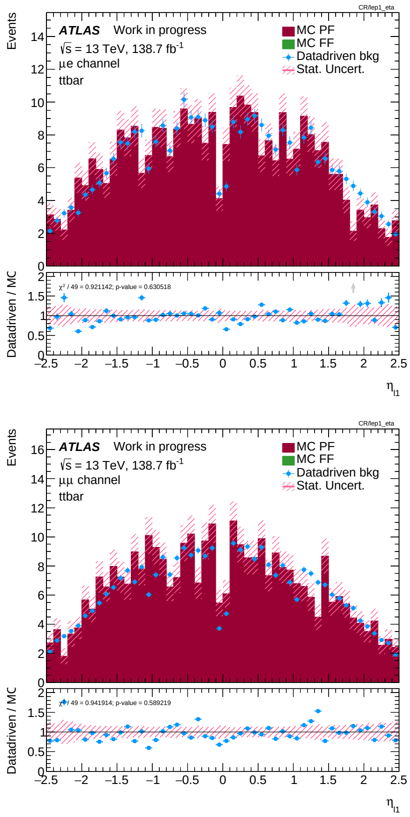


Figure C.9: Event distributions for pseudorapidity of leading (subleading) leptons on the left (right) in $t\bar{t}$. The channel description always refers to the leading and then subleading lepton flavour, except for the sum of all channels. "Prompt+Fake" events predicted by Monte Carlo (red) are compared with the data-driven estimation (blue) from the fake factor, binned in the fake leptons η, including a ratio plot. The respective statistical uncertainties are shown as well.

C.2.5 Closure Test with combined p_{T} and η Fake Factor in $t\bar{t}$

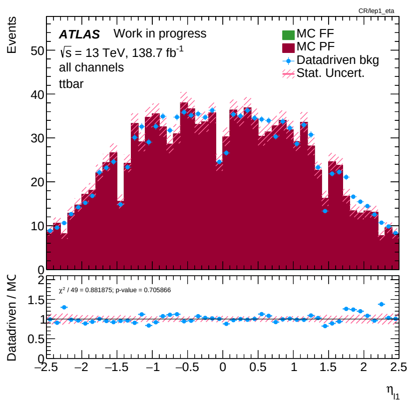


Figure C.10: Event distributions for pseudorapidity of leading (subleading) leptons on the left (right) in $t\bar{t}$. The channel description always refers to the leading and then subleading lepton flavour, except for the sum of all channels. "Prompt+Fake" events predicted by Monte Carlo (red) are compared with the data-driven estimation (blue) from the combined fake factors, binned in the fake leptons p_{T} or η, including a ratio plot. The respective statistical uncertainties are shown as well.

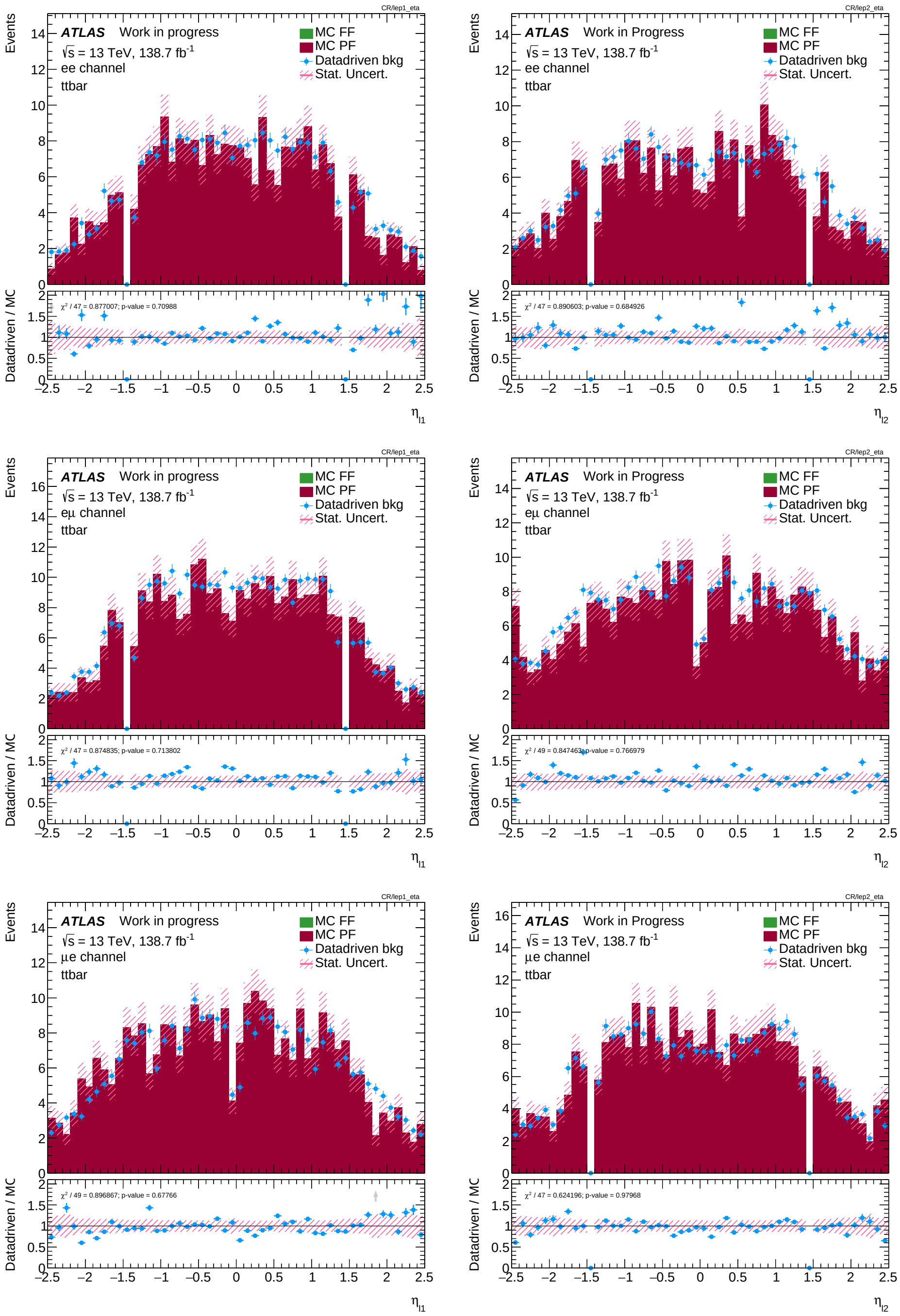


Figure C.11: Event distributions for pseudorapidity of leading (subleading) leptons on the left (right) in $t\bar{t}$. The channel description always refers to the leading and then subleading lepton flavour, except for the sum of all channels. "Prompt+Fake" events predicted by Monte Carlo (red) are compared with the data-driven estimation (blue) from the combined fake factors, binned in the fake leptons p_T or η, including a ratio plot. The respective statistical uncertainties are shown as well.

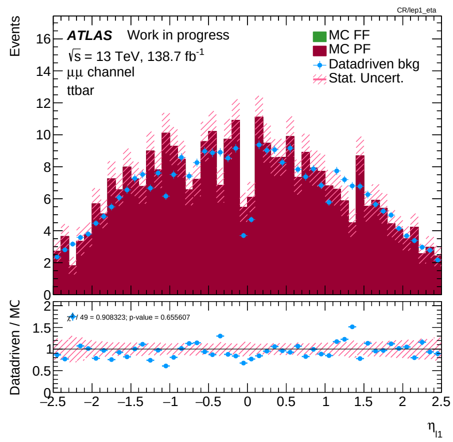


Figure C.12: Event distributions for pseudorapidity of leading (subleading) leptons on the left (right) in $t\bar{t}$. The channel description always refers to the leading and then subleading lepton flavour, except for the sum of all channels. "Prompt+Fake" events predicted by Monte Carlo (red) are compared with the data-driven estimation (blue) from the combined fake factors, binned in the fake leptons p_T or η, including a ratio plot. The respective statistical uncertainties are shown as well.

C.2.6 Verification of η Correction

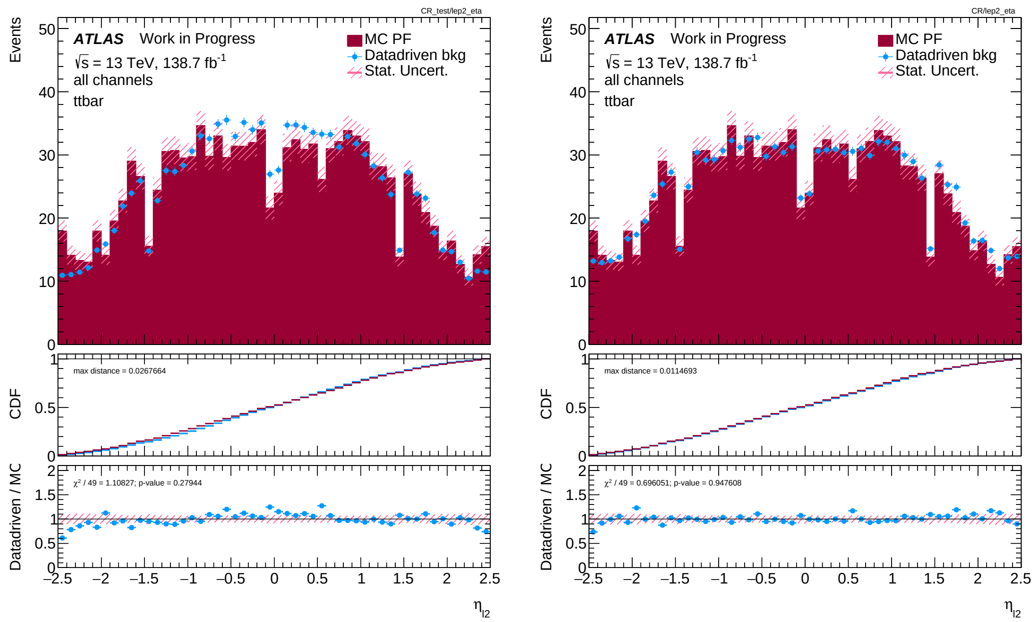


Figure C.13: Exemplary comparison of η-distribution (left column: with p_T fake factor only, right: with η corrected p_T fake factor) in $t\bar{t}$. The x-axis label $\eta_{\ell\,i}$ denotes the respective leading or subleading lepton distribution. The Channel description then always refers to the leading and then subleading lepton flavour, except for the sum of all channels. Furthermore, CDF comparison for MC prediction and data-driven estimate, and respective ratio plots are included as auxiliary plots.

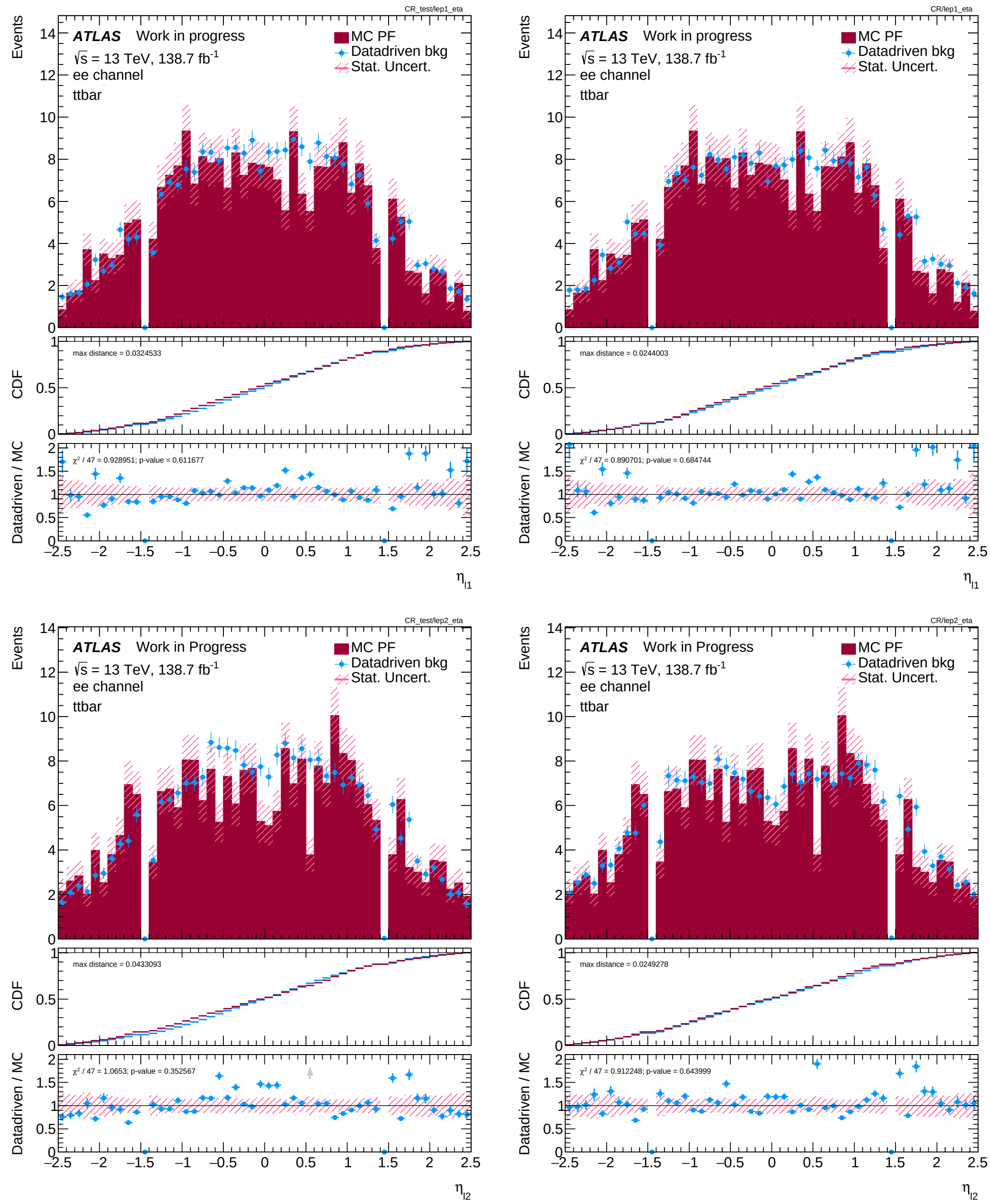


Figure C.14: Exemplary comparison of η-distribution (left column: with p_T fake factor only, right: with η corrected p_T fake factor) in $t\bar{t}$. The x-axis label $\eta_{\ell\, i}$ denotes the respective leading or subleading lepton distribution. The Channel description then always refers to the leading and then subleading lepton flavour, except for the sum of all channels. Furthermore, CDF comparison for MC prediction and data-driven estimate, and respective ratio plots are included as auxiliary plots.

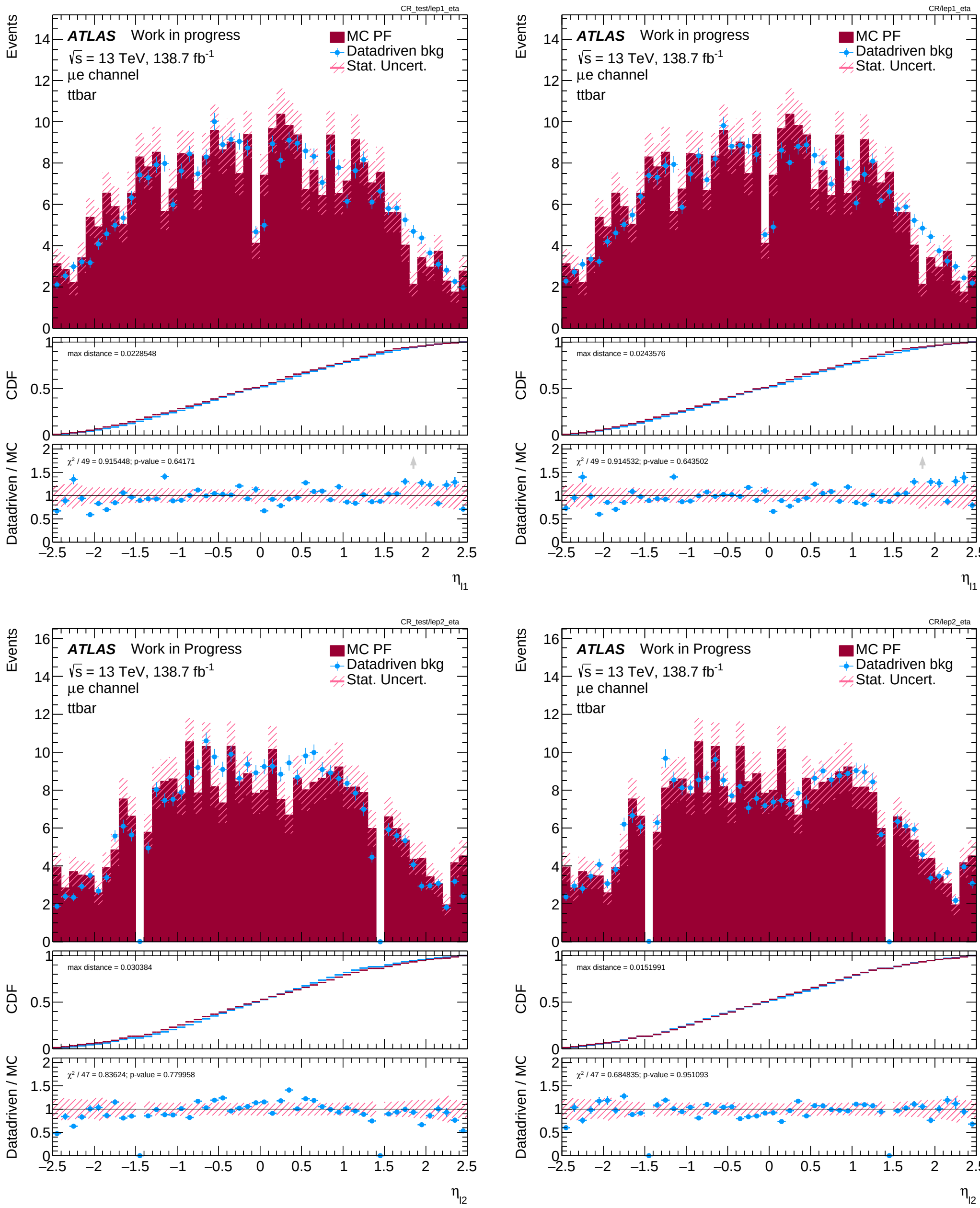


Figure C.15: Exemplary comparison of η-distribution (left column: with p_T fake factor only, right: with η corrected p_T fake factor) in $t\bar{t}$. The x-axis label $\eta_{\ell\, i}$ denotes the respective leading or subleading lepton distribution. The Channel description then always refers to the leading and then subleading lepton flavour, except for the sum of all channels. Furthermore, CDF comparison for MC prediction and data-driven estimate, and respective ratio plots are included as auxiliary plots.

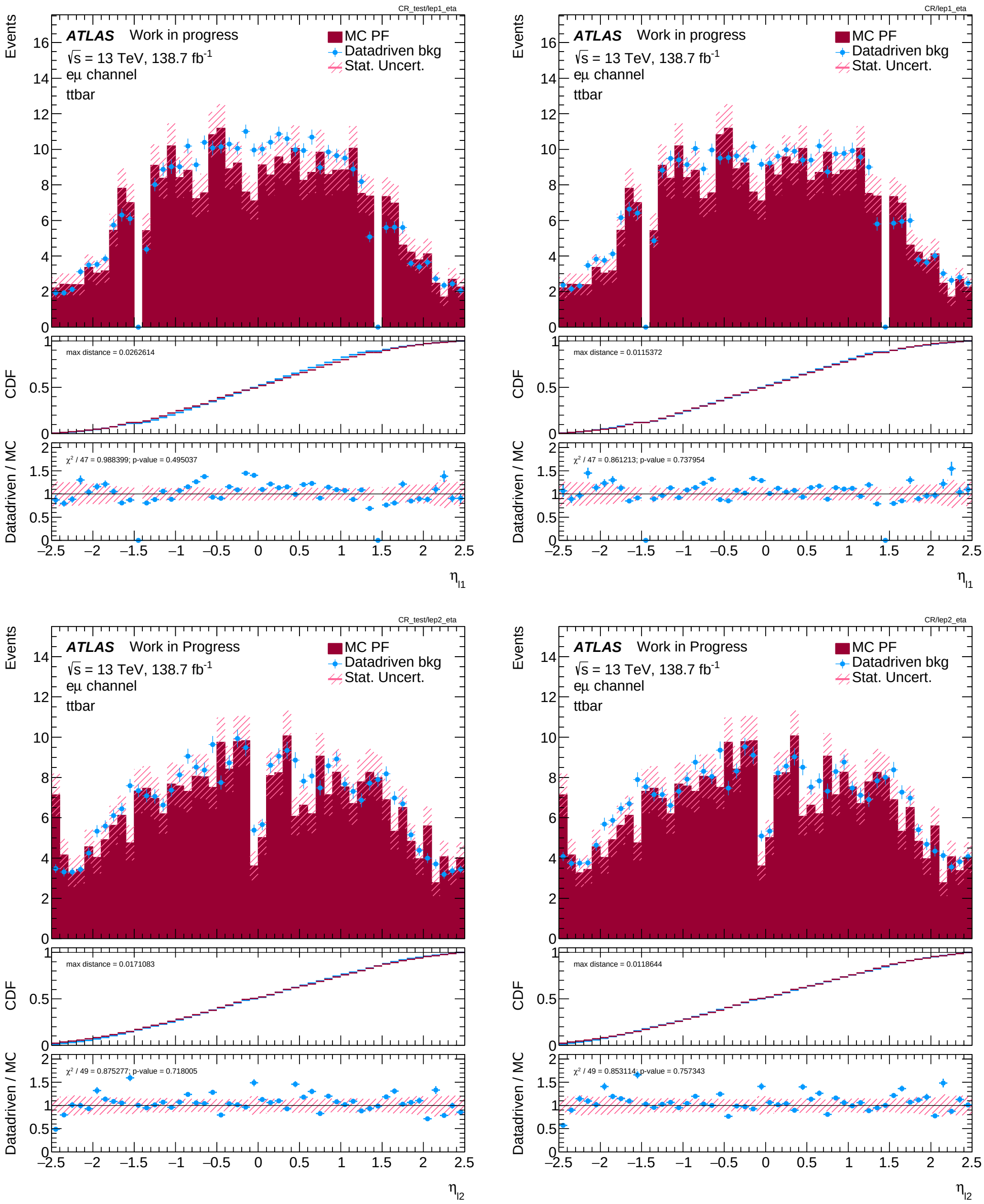


Figure C.16: Exemplary comparison of η-distribution (left column: with p_{T} fake factor only, right: with η corrected p_{T} fake factor) in $t\bar{t}$. The x-axis label $\eta_{\ell\, i}$ denotes the respective leading or subleading lepton distribution. The Channel description then always refers to the leading and then subleading lepton flavour, except for the sum of all channels. Furthermore, CDF comparison for MC prediction and data-driven estimate, and respective ratio plots are included as auxiliary plots.

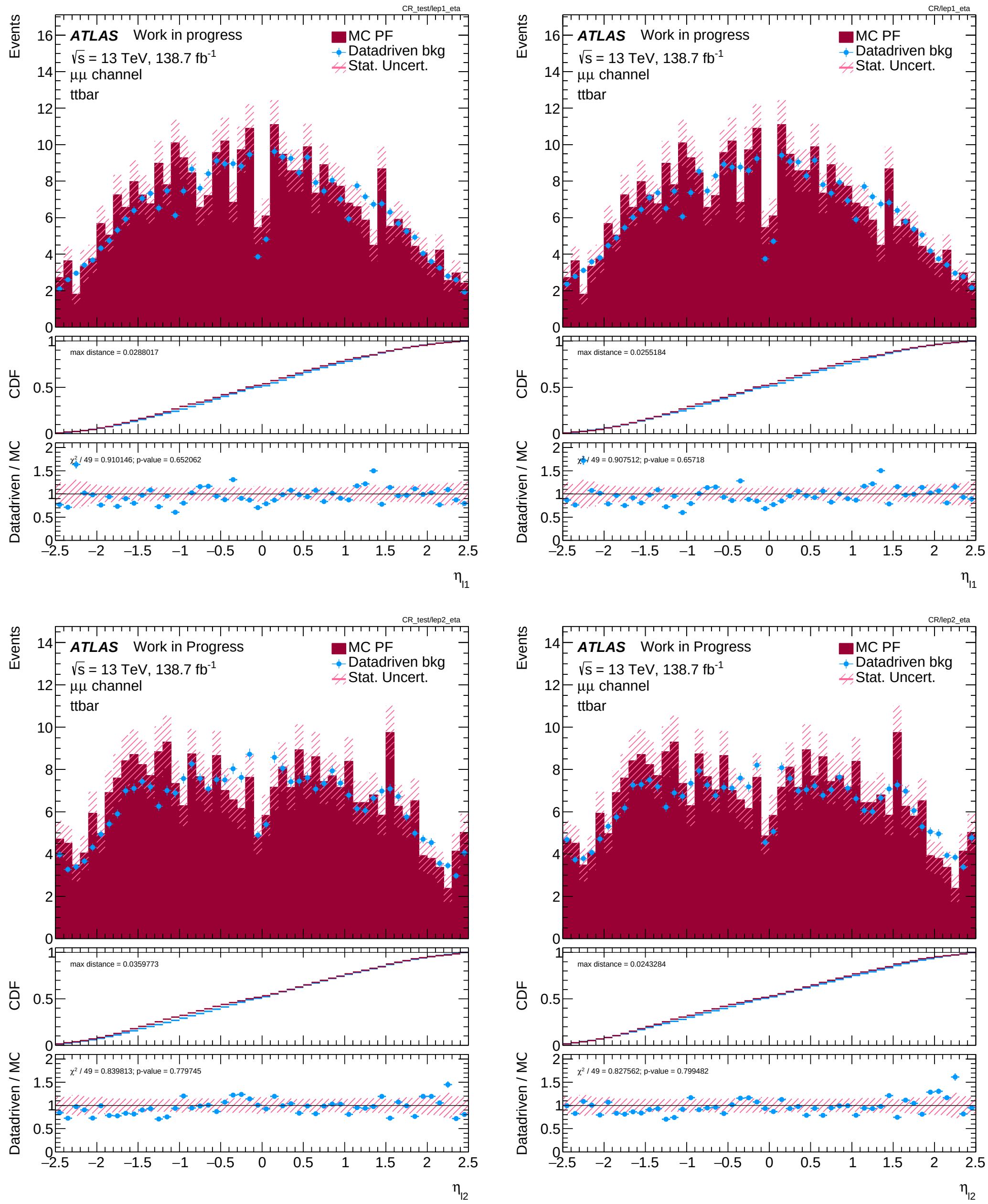


Figure C.17: Exemplary comparison of η-distribution (left column: with p_{T} fake factor only, right: with η corrected p_{T} fake factor) in $t\bar{t}$. The x-axis label $\eta_{\ell\, i}$ denotes the respective leading or subleading lepton distribution. The Channel description then always refers to the leading and then subleading lepton flavour, except for the sum of all channels. Furthermore, CDF comparison for MC prediction and data-driven estimate, and respective ratio plots are included as auxiliary plots.

Bibliography

[1] M. Aaboud et al., *Measurement of $W^\pm W^\pm$ vector-boson scattering and limits on anomalous quartic gauge couplings with the ATLAS detector. Measurement of $W^\pm W^\pm$ vector-boson scattering and limits on anomalous quartic gauge couplings with the ATLAS detector*, Phys. Rev. D **96** (2016) 012007. 34 p, URL: `https://cds.cern.ch/record/2230988` (cit. on pp. 1, 5–7).

[2] Angela Maria Burger, "Measurement of Vector Boson polarisation in WZ production at the LHC with the ATLAS detector. Mesure de la polarisation de bosons vecteurs lors de leur production par paire WZ au LHC avec le détecteur ATLAS", 2018, URL: `https://cds.cern.ch/record/2645514` (cit. on pp. 1, 3–5, 10, 12, 20, 23, 25).

[3] Carsten Bittrich, "Evidence for Scattering of Electroweak Gauge Bosons in the $W^\pm Z$ Channel with the ATLAS Detector at the Large Hadron Collider", 2020, URL: `https://cds.cern.ch/record/2719126` (cit. on pp. 1, 3–7, 11, 20, 23, 25).

[4] C. N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Phys. Rev. **96** (1 1954) 191, URL: `https://link.aps.org/doi/10.1103/PhysRev.96.191` (cit. on p. 3).

[5] Mark Thomson, *Modern Particle Physics*, Cambridge University Press, 2013 (cit. on pp. 3, 4).

[6] Carsten Bittrich, "Study of Polarization Fractions in the Scattering of Massive Gauge Bosons $W^\pm Z \to W^\pm Z$ with the ATLAS Detector at the Large Hadron Collider", 2015, URL: `https://cds.cern.ch/record/2014124` (cit. on p. 4).

[7] M. et al. Aaboud, *Observation of Electroweak Production of a Same-Sign W Boson Pair in Association with Two Jets in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*, Phys. Rev. Lett. **123** (16 2019) 161801, URL: `https://link.aps.org/doi/10.1103/PhysRevLett.123.161801` (cit. on pp. 5, 6).

[8] A. M. et al. Sirunyan, *Observation of Electroweak Production of Same-Sign W Boson Pairs in the Two Jet and Two Same-Sign Lepton Final State in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV*, Phys. Rev. Lett. **120** (8 2018) 081801, URL: `https://link.aps.org/doi/10.1103/PhysRevLett.120.081801` (cit. on pp. 5, 6).

[9] A. Sirunyan et al., *Measurements of production cross sections of WZ and same-sign WW boson pairs in association with two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **809** (2020) 135710. 42 p, URL: `https://cds.cern.ch/record/2716981` (cit. on p. 6).

[10] Moritz Poguntke, "Study of the non-prompt lepton background in the same charge WW analysis performed in ATLAS at 13 TeV", 2020, URL: `http://iktp.tu-dresden.de/IKTP/pub/20/20200127_Poguntke_Bachelorthesis.pdf` (cit. on p. 7).

[11] Franziska Iltzsche, "Observation of Electroweak Same-Charge $W^\pm W^\pm jj$ Production at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC", 2019, URL: `https://cds.cern.ch/record/2697175` (cit. on pp. 7, 11, 15).

[12] J. T. Boyd, *LHC Run-2 and Future Prospects*, 2020, arXiv: `2001.04370 [hep-ex]` (cit. on p. 9).

[13] G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003. 437 p, Also published by CERN Geneva in 2010, URL: `https://cds.cern.ch/record/1129811` (cit. on p. 10).

[14] Fabian Kuger, *Signal Formation Processes in Micromegas Detectors and Quality Control for large size Detector Construction for the ATLAS New Small Wheel*, 2017, arXiv: `1708.01624 [physics.ins-det]` (cit. on p. 10).

[15] Andy Buckley et al., *General-purpose event generators for LHC physics*, 2011, arXiv: `1101.2599 [hep-ph]` (cit. on p. 11).

[16] Carsten Bittrich, "Study of the Effects of Anomalous Quartic Gauge Couplings of the Scattering of Two Gauge Bosons V V V V at the Large Hadron Collider", 2012, URL: `http://iktp.tu-dresden.de/IKTP/pub/12/Bachelorarbeit-Carsten-Bittrich.pdf` (cit. on p. 11).

[17] John Alison, "The Road to Discovery: Detector Alignment, Electron Identification, Particle Misidentification, WW Physics, and the Discovery of the Higgs Boson", 2015, URL: `https://cds.cern.ch/record/1536507` (cit. on p. 12).

[18] Max V. Stange, "Data-driven estimation of the non-prompt background in same charged WW scattering within the ATLAS experiment using the fake factor method", To be published in 2020, MA thesis: Technische Universität Dresden (cit. on pp. 14–16, 37, 38).

Erklärung

Hiermit erkläre ich, dass ich diese Arbeit im Rahmen der Betreuung am `Institut für Kern- und Teilchephysik` ohne unzulässige Hilfe Dritter verfasst und alle Quellen als solche gekennzeichnet habe.

Christoph Klimanek,

Dresden, 18.09.2020