MASTER THESIS

FOR OBTAINING THE DEGREE OF MASTER OF SCIENCE (M.SC.)

Estimation of the electron charge misidentification background for the same-sign WW measurement with the ATLAS detector at $\sqrt{s} = 13$ TeV

José Antonio Fernández Pretel

Supervised by

Prof. Dr. Beate Heinemann

Dr. Oleg Kuprash

FAKULTÄT FÜR MATHEMATIK UND PHYSIK

ALBERT-LUDWIGS UNIVERSITY OF FREIBURG

Abstract

Electron charge misidentification was the largest source of background in the ee decay channel, and the third-largest background source in the all-inclusive leptonic $\ell\ell$ channel in the ATLAS measurement of electroweak production of same-sign W boson pairs in proton-proton collisions, made using the partial Run 2 dataset collected during 2015-2016. In this thesis, the sources of the charge misidentification are studied. The contribution of this background to the measurement of the electroweak production of two same-sign W bosons is then estimated using a data-driven method, based on the full Run 2 dataset, corresponding to an integrated luminosity of 139 fb^{-1} and a center-of-mass energy of $\sqrt{s} = 13$ TeV. The *Electron Charge ID Selector*, a tool based on a boosted decision tree developed by the *Egamma Combined Performance group* of the ATLAS Collaboration, is applied to reduce this background. The background contribution in a "Z-peak" control region in the ee channel, with the dilepton invariant mass m_{ee} satisfying $|m_{ee} - m_Z| < 15$ GeV, is reduced by a factor of 11. The background yield in the signal region (SR) in the $\ell\ell$ channel is reduced by a factor of 4 and corresponds to 10.3 $\pm$ 0.4 (stat) $\pm$ 6.7 (sys) events using the full Run 2 dataset. To increase the signal yields and to have a unified selection with other decay channels, a possibility for extension of the electron pseudorapidity range in the ee channel to the almost full acceptance of the inner tracking detector ($|\eta| < 2.47$) is checked. In this case, a background yield of 26.2 $\pm$ 0.8 (stat) $\pm$ 12.2 (sys) events is obtained, with a factor of 1.6 signal increase in the ee channel using the full Run 2 dataset.

Contents

Chapter 1

Introduction

Particle physics pursues a description of the fundamental constituents of matter and their interactions. The Standard Model of particle physics (SM), the most successful theory in terms of theoretical predictions known so far, achieves this goal by means of a unified description of quantum-mechanical principles (crucial at subatomic scales) and special relativity in a *quantum field theory* whose conservation laws are derived based on symmetry principles. The spontaneous breaking of the symmetry describing electroweak interactions causes three of the four mediators of the this force, namely the $W^{\pm}$ and Z bosons, to acquire mass by means of interactions with the *Higgs field*. The photon γ, as the fourth mediator, remains massless. Vector Boson Scattering (VBS) plays a crucial role in the Standard Model of particle physics for the understanding of the electroweak symmetry breaking mechanism, as well as the influence of the Higgs boson on the regularization of divergences in the theoretical calculations. Two gauge bosons in the final state accompanied by two jets can be produced at the LHC through strong and weak interactions, the latter being the mode sensitive to VBS. The production of two W bosons with identical electric charge and two jets is characterized by the largest ratio of the electroweak to strong modes, compared to other diboson final states. It is therefore an ideal candidate for testing the Standard Model predictions in the electroweak sector, by exploiting the quartic couplings involved in the theory and the linear polarization of the W bosons.

The electroweak production of same-sign W boson pairs is experimentally accessible at the Large Hadron Collider using leptonic decays, concretely electrons and muons. Since two leptons with the same charge are required in the final state, while opposite-sign final state leptons are much more probable to be produced in terms of production cross sections (e.g. through $Z+$ jets or $t\bar{t}$ events), inefficiencies in the electron reconstruction make the unlikely electron charge misidentification to become an important background contribution for the measurement. For the measurement of electroweak production of same-sign W boson pairs, the contribution due to electron charge misidentification is dominant in the ee channel and the third most important contribution in the all-inclusive $\ell\ell$ channel. In this thesis, the sources of the charge misidentification in the ATLAS detector are characterized. The contribution of this background is then estimated using a data-driven method for the measurement of the electroweak production of same-sign WW with two jets using proton-proton collisions recorded at the LHC by the ATLAS experiment for the full Run 2 dataset (139 fb^{-1} at $\sqrt{s}=13$ TeV). A Boosted Decision Tree (*Electron Charge ID Selector*, or ECIDS) developed by the *Egamma Combined Performance* *group* of the ATLAS Collaboration is applied to reduce this background by removing charge-misidentified electrons. By studying the signal significance in the signal region, the reduction of this background allows to explore a further relaxation of the selection criteria in the ee channel by considering the full angular acceptance of the inner detector, increasing the number of expected signal events.

The SM and its theoretical framework are outlined in Chapter 2, as well as the phenomenology of proton-proton collisions in hadron-collider experiments. The setup of the ATLAS detector at the LHC is presented in Chapter 3 along with the reconstruction algorithms of the different particles. The strategy to perform measurements on VBS and especially on electroweak production of same-sign W boson pairs is described in Chapter 4, introducing the analyzed dataset and the event selection in Chapter 5. Chapter 6 starts with an overview on the estimation of the charge misidentification background in previous analyses of this process carried out by the ATLAS Collaboration, including a quantitative study of the different sources of charge misidentification. The estimated background contribution due to charge misidentification for the measurement of the electroweak production of same-sign W boson pairs using almost the full Run 2 dataset at $\sqrt{s} = 13$ TeV is detailed in Chapter 7, along with its reduction using the ECIDS tool and further explorations of the signal region relaxation in the ee channel. The conclusions and outlook of this work are presented in Chapter 8.

Chapter 2

Theoretical Foundations

The theoretical framework provided by the Standard Model (SM) is used for the experimental measurements performed in this work. In this chapter, the fundamental concepts of the SM of particle physics are briefly outlined in Section 2.1. Electroweak gauge boson scattering, as a relevant piece for this work, is introduced in Section 2.2 in the context of the SM. The phenomenology of proton–proton collisions (pp) at high energies is described in Section 2.3, along with the approach for computational simulations. The process of interest in this work, vector boson scattering of W pairs of same-sign, its production in pp collisions at the LHC and topology will be addressed in Chapter 4.

2.1 The Standard Model of particle physics

Physics explains forces in nature based on four fundamental sources, known as electromagnetic, strong, weak and gravitational interactions. The fundamental particles of nature and their interactions are described with high precision by the SM, a quantum field theory (QFT). Section 2.1.1 provides an overview of the theoretical description of the SM. A deeper dive into its mathematical formalism is done in Section 2.1.2 in the context of *Quantum Electrodynamics*, the quantum field theory that describes electromagnetic interactions in the SM. Strong interactions are also described as QFT in the SM, whose overview is presented in 2.1.3. The Electroweak sector, the most relevant component of the SM in this work, is introduced in Section 2.1.4, whose predictions need to be suited by introducing a symmetry breaking mechanism that will be discussed in Section 2.1.5. The terms arising from this broken symmetry complete the SM Lagrangian.

2.1.1 Introduction

The SM describes the laws of nature at its most fundamental scale, where components of matter are expected to be indivisible. For this purpose, elementary particles and their interactions need to be addressed in detail, respecting the conservation laws of physics. The *Noether theorem* [2] plays a crucial role when connecting these two objectives by proving mathematically a direct relation between continuous gauge symmetries in a theory with physical conservation laws. As a *gauge theory*, electroweak (unified description of weak

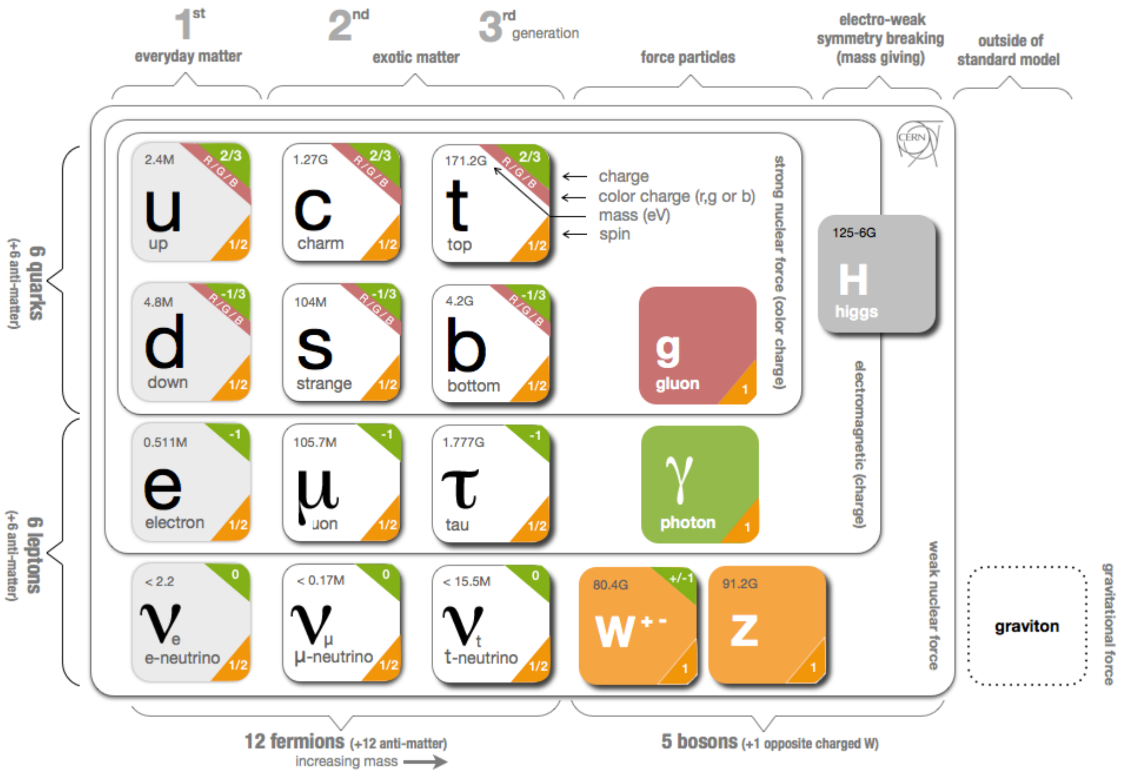


Figure 2.1: Elementary particles in the SM (Source: [1]).

and electromagnetic interactions) and strong interactions emerge in the SM by postulating local symmetry invariance under the non-abelian group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. Gravitational interactions can not be described yet by the SM. Every symmetry has a corresponding charge, which allows charged quantum fields (associated to particles) to interact under the corresponding symmetry. The concept of gauge invariance will be introduced in more detail in Section 2.1.2 in the context of *Quantum Electrodynamics*, the quantum field theory introducing the description of electrodynamics in the SM. Quantum fields are classified according to their *spin*, taking half-integer numbers[1] (1/2, 3/2, ...) for *fermionic* fields and integer values (0, 1, ...) for *bosonic* fields. *Fermions*, i.e. fundamental particles associated to a fermionic field, come in three *generations* and can be divided into six *quarks* and six *leptons* (and their corresponding antiparticles) whether they are charged under the $SU(3)_C$ symmetry or not, respectively. All generations share the same quantum numbers but differ in the particle masses. The classification of the fundamental particles in the SM is exposed in Figure 2.1 along with the mediators of the aforementioned interactions described in the SM.

The strong interaction, governed by the gauge group $SU(3)_C$, is described in the SM by mediation of eight *gluons*. Only quarks and gluons, particles carrying *color charge* (C), interact through the strong interaction. The color charge can take values of *red*, *green* and *blue*, and their corresponding anti-values. Only particles with vanishing net color have been experimentally observed so far. Following this evidence, two different kinds of par-

[1] Natural units $\hbar = c = 1$ are used in this thesis.

ticles can be produced, *mesons* (composed by two colored-opposed quarks) and *baryons* (superposition of all three colors or all three anti-colors). This behavior is know as *color confinement* and it is still to be proven analytically. Its behavior is commonly believed due to the fact that no free, colored particles have been experimentally observed. The coupling constant of strong interactions shows a peculiarity denoted by *asymptotic freedom*, that makes quarks and gluons to weakly[2] interact at higher energies, allowing to precisely compute scattering amplitudes by using perturbation theory at this energy scale. The QFT describing strong interactions within the SM is known as *Quantum Chromodynamics* (QCD) and its theoretical framework will be revisited in Section 2.1.3.

Since the ordering in first, second and third generation increase with the mass, second and third generations contain unstable particles that decay into lighter ones. Therefore, stable matter in the universe is composed only by elementary particles contained in the first generation. These decays are described in the SM by the electroweak interaction, whose formalism is outlined in Section 2.1.4. Electroweak interactions arise from the gauge group $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$, whose associated charges are the *weak hypercharge* (Y) and the *weak isospin* T. The third component of T (denoted by T_3) added to Y provides the *electric charge* $Q = T_3 + Y$. The *photon* γ, a massless gauge boson that conveys the interaction between electrically-charged particles, and three massive gauge bosons $W^\pm$ and Z, mediate the electroweak interactions among particles. Under the $\mathrm{SU}(2)_L$ group, quantum fields forming doublets are referred to as *left-handed* (LH) fermion fields, while singlet fields are know as *right-handed* (RH) fermion fields. This structure introduces *maximal parity violation*[3], as proven experimentally in β^- decays [3]. Weak interactions also introduce *CP violation* in the quark sector, regulating mass and interaction eigenstates of up-down quarks by means of the *CKM matrix* [4, 5].

Since the electroweak gauge symmetry does not allow gauge bosons to be massive, it is necessary to break this symmetry to generate massive *electroweak gauge bosons* W and Z in the SM according to experimental evidence [6, 7, 8, 9]. This is done by introducing a new complex scalar doublet in the theory with four new degrees of freedom [10, 11, 12]. Only one of them has physical meaning, the *Higgs field*, with the rest being removed by gauge transformations. A candidate for the SM Higgs boson was discovered at the LHC simultaneously by ATLAS [13] and CMS [14]. Its mass (about 125 GeV) and properties are consistent with the SM predictions [15]. The Electroweak Symmetry Breaking (EWSB) formalism will be visited in Section 2.1.5.

2.1.2 Fields and Gauge interactions in Quantum Electrodynamics

Quantum Electrodynamics (QED) is the QFT that describes electromagnetic interactions among fundamental particles in the SM model. It predicts quantum-mechanical phenomena at relativistic particle energies with astonishing accuracy. The illustration of its mathematical framework, introduced in this section, is also shared by the rest of quantum field theories in the SM.

Particles are described in the SM as *quantum fields*, with their propagation and interactions described in the *Lagrangian formalism*. This formalism is described in detail, for

[2]Not meaning that they interact by means of weak interactions.

[3]Discrete symmetry transformation that drives spatial coordinates to their opposite, $P(\vec{x}) = -\vec{x}$

instance, in Refs. [16, 17] on which this overview is based. Equations of motion for a freely propagating fermion spinor field ψ can be obtained from the *Euler-Lagrange equations* applied to the Lagrangian density[4]

$$\mathscr{L}_{\text{Dirac}} = \bar{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi\,, \tag{2.1}$$

where $\gamma^{\mu}\partial_{\mu}$ is the contraction of the *Dirac matrices* γ^{μ} and m is the mass of the fermion. This Lagrangian is invariant under *gauge transformations* of the fermion field[5], i.e., $\psi \to e^{i\alpha}\psi$, where α is a real phase factor. This Lagrangian is also known as *free Dirac Lagrangian*, since the corresponding equations of motion for the Dirac field ψ in $\mathscr{L}_{\text{Dirac}}$ correspond to the Dirac equation for ψ and $\bar{\psi}$. In order to address the conservation laws and previous knowledge from electrodynamics, the symmetry principles of the SM impose the invariance under *local gauge transformations* under the $\mathrm{U}(1)$ group, which means that the phase factor α is a function of space-time coordinates x,

$$\psi \to e^{i\alpha(x)}\psi\,. \tag{2.2}$$

However, the Lagrangian (2.1) is not invariant under Eq. (2.2). Introducing a new degree of freedom in the Dirac equation A^{μ} addresses the required local symmetry invariance, provided the transformation of the gauge boson field and the the so-called *covariant derivative* D_{μ}

$$\begin{aligned} A_{\mu} \to A'_{\mu} &= A_{\mu} + \partial_{\mu}\alpha(x), \\ \partial_{\mu} \to D_{\mu} &= \partial_{\mu} + iQA_{\mu}\,, \end{aligned} \tag{2.3}$$

where A_{μ} is again a massless free gauge boson field. The invariance of $\mathscr{L}_{\text{Dirac}}$ is now accomplished by applying the transformations (2.2) and (2.3) which introduce a new *interaction term* $q\bar{\psi}\gamma^{\mu}A_{\mu}\psi$ into the Lagrangian. The gauge-invariant Lagrangian for a spin-half fermion interacting through a gauge massless spin-one boson field A_{μ} is obtained by considering the free Dirac and interaction terms as well as the *kinematic term* for the gauge field $-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$,

$$\mathscr{L}_{\text{QED}} = \mathscr{L}_{\text{Dirac}} + \mathscr{L}_{\text{int}} + \mathscr{L}_{\text{kin}} = \bar{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi + Q\bar{\psi}\gamma^{\mu}A_{\mu}\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}\,. \tag{2.4}$$

$\mathscr{L}_{\text{QED}}$ constitutes the Lagrangian of QED. The corresponding equations of motion are the *Maxwell equations* known from classical electrodynamics. A_{μ} is then the electromagnetic vector potential, $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$ is the electromagnetic field tensor, and Q is the fermion charge. If N kinds of electrically charged fermions need to be considered in the same Lagrangian, only the sum $\sum_i^N \mathscr{L}_{\text{QED},i}$ for the different fermion fields ψ_i of electric charge Q_i is necessary. The introduced Lagrangian can account for nearly all observed phenomena from macroscopic scales down to 10^{-15} m.

Following the same prescription, strong interactions can be described by imposing local gauge invariance under the group $\mathrm{SU}(3)$ as discussed in the following. The arising theoretical framework is known as *Quantum Chromodynamics*. A very similar approach exploiting the group $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ provides a unified description of electromagnetic and weak interactions. The mentioned approach will be covered in Section 2.1.4.

[4] In the following, the Lagrangian density will be referred to as "*Lagrangian*".
[5] Global phase transformation in $\mathrm{U}(1)$ group

2.1.3 Quantum Chromodynamics

Strong interactions are delineated in the SM by a QFT known as *Quantum Chromodynamics* (QCD), whose derivation is based on the requirement of local gauge invariance under $\mathrm{SU}(3)_C$ transformation of quark color fields q. The Lagrangian of free color fields, which are also fermion fields, can be written analogously to Eq. (2.1) as

$$\mathscr{L}_{\text{Dirac},C} = \sum_j^3 \bar{q}_j \left(i\gamma^\mu \partial_\mu - m\right) q_j \, , \tag{2.5}$$

where j runs over the three color charges *red*, *blue* and *green*. Local gauge invariance is required in $\mathrm{SU}(3)_C$ by means of the transformation

$$q_j \to e^{i\alpha(x)\cdot \mathbf{T}} \, q_j \, , \tag{2.6}$$

where $\mathbf{T} = \{T_a\}$, along the index $a = 1, ..., 8$, is the set of eight generators of the $\mathrm{SU}(3)$ symmetry group. Since the generators $\mathbf{T}$ do not commute[6], the group is *non-Abelian*. Local gauge invariance is addressed analogously to Section 2.1.2 by introducing vector fields with transformation properties

$$\partial_\mu \to \partial_\mu + i g_s \frac{\lambda}{2} G_\mu^a \, , \tag{2.7}$$

$$G_\mu^a \to G_\mu^{',a} = G_\mu^a - \frac{1}{g_s}\partial_\mu \alpha_a(x) - f_{abc}\alpha_b(x) G_\mu^c \, , \tag{2.8}$$

coupling the eight *gluon fields* G_μ^a with a *coupling strength* g_s. Following the same approach, the imposition of local gauge invariance under $\mathrm{SU}(3)_L$ leads to the Lagrangian of QCD,

$$\mathscr{L}_{\text{QCD}} = \sum_j^3 \bar{q}_j \left(i\gamma^\mu \partial_\mu - m\right) q_j - g_s \left(\bar{q}_j \gamma^\mu \frac{\lambda_a}{2} q_j\right) G_\mu^a - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu} \, , \tag{2.9}$$

with $G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f_{abc} G_\mu^b G_\nu^c$ being the strength tensor. As a peculiarity of $\mathrm{SU}(3)$ and its non-Abelian structure, additional interaction terms arise in QCD from this strength tensor, allowing *self-interactions* between gluon fields.

2.1.4 The Electroweak sector of the Standard Model

In the SM, electromagnetic and weak interactions are described in an unified framework named the *electroweak theory* that requires local gauge invariance under transformations of the $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ group. In this group, fermion fields are represented by left-handed (LH) doublets χ_L and right-handed (RH) singlets ψ_R. For first-generation leptons, they are expressed as[7]

$$\chi_L(x) = \begin{pmatrix} \nu_e(x) \\ e^-(x) \end{pmatrix} \quad , \quad \psi_R(x) = e_R(x) \, . \tag{2.10}$$

[6]Commutation relations are $[T_a, T_b] = i f_{abc} T_c$, where f_{abc} are the *structure constants* of the group SU(3).

[7]For simplicity in notation, explicit dependence of fermion fields on space-time coordinates x will be omitted in the following.

Expressions for the first-generation quarks as well as second and third generations are analogous. As done in Sections 2.1.2 and 2.1.3, local gauge invariance under symmetry groups $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ are imposed by

$$\chi_L \to e^{i\beta(x)Y + i\alpha_a(x)\tau_a}\chi_L\,, \tag{2.11}$$

$$\psi_R \to e^{i\beta(x)Y}\psi_R\,, \tag{2.12}$$

the three *Pauli matrices* τ_a are the generators of $\mathrm{SU}(2)_L$ while the *weak hypercharge* Y appears as generator of the group $\mathrm{U}(1)$. $\beta(x)$ and $\alpha(x)$ are the phase transformations. The value of hypercharge for a LH fermion field takes $Y = -1$ and $Y = -2$ for RH fermion fields. By transforming the derivative ∂_μ to the covariant derivative, the Lagrangian describing interactions between massless gauge boson fields and massless fermion fields is obtained,

$$\begin{aligned}\mathscr{L}_{\text{EW}} = \mathscr{L}_{\text{EW},int} + \mathscr{L}_{\text{EW},kin} = \\ = \bar{\chi}_L\gamma^\mu\left[i\partial_\mu - g_w\frac{\tau_a}{2}W_\mu^a + \frac{g_w'}{2}B_\mu\right]\chi_L + \bar{\psi}_R\gamma^\mu\left[i\partial_\mu + g_w'B_\mu\right]\psi_R - \\ - \frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}\,,\end{aligned} \tag{2.13}$$

where the kinematic term $\mathscr{L}_{\text{EW},kin}$ for the gauge field strength tensors $B_{\mu\nu}$ and $W_{\mu\nu}$ was included. Interactions between massless gauge boson fields and massless fermion fields described in this theory are in contradiction with experimental evidence [6, 7, 8, 9, 18]. In order to address these, the *physical fields* $W_\mu^\pm$, A_μ and Z_μ are expressed as linear combinations of the original gauge fields B_μ, W_μ^a ($a = 1, 2, 3$),

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left[W_\mu^1 \mp iW_\mu^2\right] \quad , \quad \begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix}\begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}, \tag{2.14}$$

where the *Weinberg angle* θ_W needs to be

$$\cos\theta_W = \frac{g_w}{\sqrt{g_w^2 + g_w'^2}} \quad , \quad \sin\theta_W = \frac{g_w'}{\sqrt{g_w^2 + g_w'^2}}\,, \tag{2.15}$$

to obtain the QED current A_μ. The expansion of $\mathscr{L}_{\text{EW},kin}$ for electroweak gauge fields (2.13) introduces, in the same way as seen in Section 2.1.3 for QCD, *self-interaction terms* among the electroweak bosons. Fermion interactions with electroweak gauge bosons are introduced by $\mathscr{L}_{\text{EW},int}$.

Local gauge invariance under the symmetry group $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ forbids these gauge boson fields and the fermion fields to be massive[8]. Adding mass terms for these fields destroys *renormalization* and *unitarity*, crucial for predictability potential of the theory. In such a scenario, divergent cross sections predictions can arise for certain processes[9]. To address the experimental observations tackling the aforementioned inconveniences, additional terms need to be added to $\mathscr{L}_{\text{EW}}$. This necessity will involve the appearance of the *Higgs field* into the theory, that will allow the electroweak gauge bosons

[8] That is, terms of the kind $m_\psi\psi\bar{\psi}$, $m_W W_\mu W^\mu$, $m_B B_\mu B^\mu$ break the required local invariance when they are *ad-hoc* introduced into the Lagrangian.

[9] As an example, predicted cross section of vector boson scattering $W^+W^- \to W^+W^-$ would be directly proportional to the center-of-mass energy of the system, $\sigma \sim \frac{s}{m_W}$, where s is the center-of-mass energy and m_W is the mass of the W boson. According to this prediction, the cross section σ grows faster than the total cross section of the pp interaction.

to acquire the mass observed experimentally with high accuracy. In Section 2.1.5 these calculations are outlined.

2.1.5 Spontaneous Symmetry Breaking, the Higgs boson and the SM Lagrangian

In order to address the inconsistencies of the SM predictions with the measurements of electroweak interactions, further developments on the calculations provided in Section 2.1.4 are needed. Mass terms are introduced in $\mathscr{L}_{\text{EW}}$ by incorporating a new complex scalar boson field ϕ with non-vanishing expectation value in the ground state, that interacts with the gauge boson fields by means of the covariant derivative producing a *spontaneous symmetry breaking* of $\text{SU}(2)_L \otimes \text{U}(1)_Y$,

$$\mathscr{L}_{\text{EW}}^{\phi} = (D_\mu\phi)^\dagger D^\mu\phi - V(\phi) = (D_\mu\phi)^\dagger D^\mu\phi - \mu^2\phi^\dagger\phi + \lambda\phi^\dagger\phi\,, \tag{2.16}$$

with D_μ being the covariant derivative for $\text{SU}(2)_L \otimes \text{U}(1)_Y$. The potential $V(\phi)$ is the most general way to keep both renormalizability of the theory and gauge invariance with $\mu^2 < 0$ and $\lambda > 0$ having a local maximum and infinite-degenerated minima. This new complex scalar field has 4 degrees of freedom. The symmetry $\text{U}(1)_Q$ needs to be unbroken, so that non-zero hypercharge and isospin are required for ϕ. Choosing these features for the potential ϕ,

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad , \quad |\phi_0| = \sqrt{\frac{-\mu}{2\lambda}} = \frac{v}{\sqrt{2}} \tag{2.17}$$

where $|\phi_0|$ is the minimum value of ϕ, also known as its *vacuum expectation value*. The scalar field $h(x)$ is the so-called *Higgs field*. The interactions of this field with the electroweak gauge bosons, arising from Eq. (2.16), provide the corresponding mass terms in the Lagrangian. Values for electroweak gauge bosons masses are related to the vacuum expectation value of the Higgs field,

$$m_W = \frac{g\cdot v}{2} \quad , \quad m_H = \sqrt{2\mu^2} \quad , \quad m_Z = \frac{m_W}{\cos\theta_W} = \frac{v^2\sqrt{g^2+g'^2}}{2} \quad , \quad m_A = 0\,. \tag{2.18}$$

Along the same calculation, self-interactions of the Higgs field arise from the theory. The aforementioned couplings also resolve unitarity and renormalizability of the electroweak theory mentioned in Section 2.1.4.

Even though the Brout-Englert-Higgs mechanism incorporated the mass terms for the electroweak gauge bosons into the Lagrangian (2.16), the mass terms for fermions are still missing. The same Higgs doublet (2.17) is sufficient to explain also the fermion mass terms by means of the spontaneous symmetry breaking of the same group. Gauge invariant terms $\mathscr{L}_{\text{Yukawa}}$ under the $\text{SU}(2)_L \otimes \text{U}(1)_Y$ symmetry can be incorporated. As an example, the corresponding term for lepton fields χ_L and ψ_R is shown,

$$\mathscr{L}_{\text{Yukawa}} = -G_\ell\left[\bar{\chi}_L\phi\psi_R + \bar{\psi}_R\phi^*\chi_L\right]\,, \tag{2.19}$$

with arbitrary parameters G_ℓ, being $\ell = e, \mu, \tau$. Using the same parametrization (2.17) for the Higgs field, the mass terms for the different leptons and the interaction terms of the different leptons with the Higgs field h are

$$\mathscr{L}_{\text{Yukawa}} = \mathscr{L}_{\text{mass}} + \mathscr{L}_{h\ell\ell,int} = \frac{G_\ell v}{\sqrt{2}} \left(\bar{\ell}_L \ell_R + \bar{\ell}_R \ell_L\right) - \frac{G_\ell}{\sqrt{2}} \left(\bar{\ell}_L \ell_R + \bar{\ell}_R \ell_L\right) h \, . \tag{2.20}$$

Adding the Lagrangian terms from the electroweak sector (2.13) (describing kinematics for the vector gauge bosons, their self-interactions, lepton and quark kinematic energies and interactions with vector gauge bosons), the couplings of vector gauge boson fields to the Higgs (2.16) and the Lagrangian term describing lepton and quark masses and couplings to the Higgs field (2.19) for all fermions (leptons and quarks) provides the full Lagrangian for the Standard Model. The experimental evidence is predicted within its theoretical framework with high accuracy. All the precision measurements carried out in particle collider experiments so far, as shown in Figure 2.2 for the measurements in the ATLAS experiment[10], are in agreement with the SM. Nevertheless, multiple experimental observations in the recent years can not be explained in the context of its mathematical framework. The SM is, therefore, an incomplete theory for a complete understanding of the fundamental constituents of nature. For instance, attempts to incorporate gravitation as a quantum field theory within the SM lead to *non-renormalizable theories*. There is strong evidence of *neutrino oscillations* [20] that leads to the necessity of these to be massive, in contradiction with the SM description of these particles. Astronomical observations indicate presence of *dark matter* [21, 22, 23]. The *asymmetry matter-antimatter* has been also observed in cosmology [24]. Furthermore, the impossibility to find CP violation in strong interactions (known as *strong CP problem* [25]) or the difference between the scale of EW symmetry breaking scale ($v \approx 246$ GeV) with respect to the Planck scale (the so called *hierarchy problem* [26]) are not understood in the context of the SM.

2.2 Electroweak gauge vector boson scattering

The measurement of the scattering of massive gauge bosons, also called *Vector Boson Scattering* (VBS), turns out to be an arduous task due to the very small cross sections their predictions provide. Cross sections for electroweak production of gauge vector bosons take place within the lowest orders of magnitude along the experimental observations carried out at the LHC, as seen in Figure 2.2. Furthermore, the behavior of the Higgs boson as crucial piece for a renormalizable electroweak theory (discussed in Section 2.1.4) provides a window to test the EWSB mechanism in the SM with high precision. The scattering of massive gauge bosons is closely related to the EWSB mechanism as the longitudinal polarization modes correspond to the *Goldstone* modes of the broken symmetry [27]. Therefore, VBS is a very sensitive process to test the predictions of the SM, but also to set limits on models beyond this theory. However, transverse polarized gauge bosons are not related to EWSB mechanism [28]. VBS from transverse polarized gauge bosons is therefore not sensitive to the details of the EWSB mechanism since the unitarity problem does not arise by the scattering of vector gauge bosons in this polarization.

Deviations from SM predictions at the LHC energy regime can shed light to a path towards new indirect measurements of physics beyond the SM. Establishing limits on

[10] To date, cross-section measurements at the LHC and other collider experiments are in agreement with the SM predictions [15].

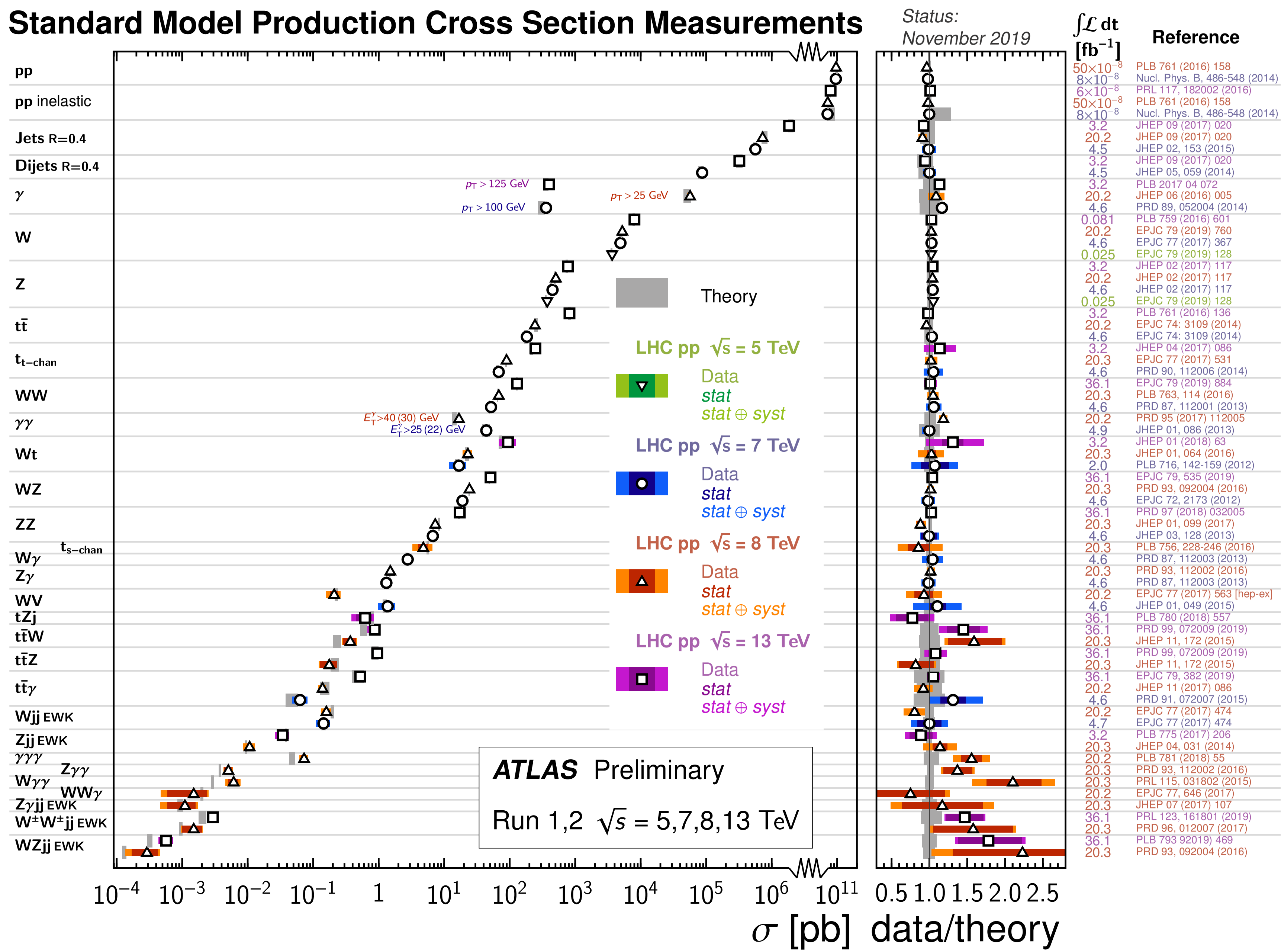


Figure 2.2: Overview of cross section measurements for a wide class of processes and observables, as obtained by the ATLAS [19] experiment at the LHC, for centre-of-mass energies of 7, 8 and 13 TeV. Theoretical predictions and their uncertainties are also shown on the right-hand side.

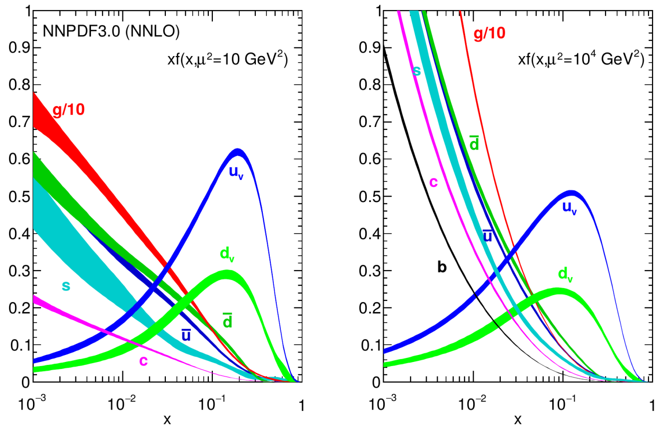


Figure 2.3: NNPDF3.0 unpolarized parton distribution functions of gluons, valence and sea quarks composing a proton at NNLO along different energy scales. (Source: [15]).

anomalous Quartic Gauge Couplings (aQGC) structure can narrow down effects of physics beyond the SM in EWSB[11]. This can be done by parameterizing deviations from the SM within an *Effective Field Theory* (EFT)[12].

2.3 Phenomenology of proton-proton collisions

The Lagrangian formalism studied so far turns out to be insufficient to compute predictions of specific observables to be directly tested in particle colliders. Indeed, protons are not elementary particles but hadrons, i.e. particles arising from bound states of quarks and gluons. This means that, in a hadron collision, processes of interest need to take into account the proton constituents and all the possible initial states for theoretical considerations. These components are also known as *partons*. They are considered to carry a certain fraction of the proton energy when a collision of partons is studied. To account for the fractions of proton momentum among its constituents, *Parton Distribution Fractions* $f_i(x_1, Q^2)$ (PDF) are defined. A PDF represents the momentum distribution of a parton q_i inside a proton at a given energy scale Q^2, with x_i being the momentum fraction of the parton q_i with respect to the proton momentum. The PDFs are determined experimentally using processes such as *Deep Inelastic Scattering* (DIS)[13]. At the moment, these functions can not be obtained from theoretical considerations and experimental results are indispensable[14]. Essential measurements on the knowledge of the proton substructure by DIS were carried out at the HERA ep collider by the H1 [30] and ZEUS[31] experiments. Results on the proton PDFs are shown in Figure 2.3 for two different energy scales Q^2. The valence quarks of the proton (whose composition is uud) are visibly enhanced. Due to the factorization theorem, it is possible to measure the PDFs at a certain energy scale Q^2 and extrapolate results to different energy scales Q_0^2 by means of the

[11] For details on the formalism and references, see Section 54 of Ref. [15].
[12] For an extensive review and basic concepts relevant to the construction of such EFTs, see Ref. [29].
[13] For a detailed coverage of DIS and DGLAP evolution equations, see for example Chapter 18 in Ref. [15] or [17].
[14] See Ref. [15], Section 18 for details.

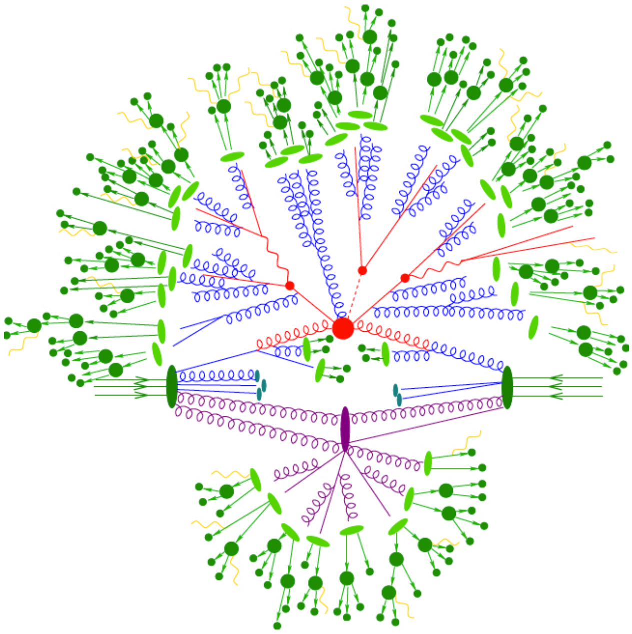


Figure 2.4: Pictorial representation of a *pp* collision produced with *Sherpa* 1.1 [35]. First, the initial state of the process (taking in consideration the substructure of both protons in the collision) is simulated. Initial state radiation of gluons and photons at this stage is taken into account. The hard processes comes next. The signal process of interest (corresponding to the production of two jets in this figure) are depicted in red. The underlying event is simulated as well (purple lines). Gluons experience *fragmentation* (light green) and hadronize into leptons. Hadrons(dark green arrows) decay into lighter particles. The hadron decay vertices are shown as dark green circles. Additionally, photons are radiated off the leptons (yellow curly lines). The beam remnants (blue) are also taken into account in MC simulations. (Source: [35]).

DGLAP evolution equations [32, 33, 34]. Following this approach, the cross section for a process X in a pp collision, $p_1p_2 \to X$, can be written as sum of partial partonic cross sections $\sigma_{q_1q_2} \to X$,

$$\sigma_{p_1p_2\to X} = \int_0^1 dx_1 dx_2 \sum_{q_1,q_2} f_{q_1}(x_1, Q^2) f_{q_2}(x_2, Q^2)\, d\sigma_{q_1q_2\to X}(x_1, x_2, s, Q^2)\,, \tag{2.21}$$

with $d\sigma_{q_1q_2\to X}(x_1, x_2, s, Q^2)$ being the *partonic cross section* where two partons q_1 and q_2 in the initial state produce the final state X.

The calculation of partonic cross sections is not sufficient to reproduce the real conditions observed in hadron collider physics. For instance, parton showers[15], hadronization and underlying events are produced in such a collisions. These phenomena are not considered in the computation exposed in Eq. (2.21). At the moment, soft hadronic phenomena such as hadronization and the underlying event production can not be calculated from first principles, relying on QCD-inspired models[16]. The consideration of these phenomena in simulations is crucial in experimental hadron collider physics to know the

[15] Perturbative evolution gives rise to a shower of quarks and gluons (partons). Multi-parton final states from leading and higher order calculations of matrix elements are linked to these parton showers using factorization prescriptions. For an overview and comparisons between different models, see [36, 37].

[16] A dedicated review on these models throughout different Monte Carlo Event Generators can be found in Ref. [15].

detector performance and the theoretical predictions. *Monte Carlo methods* (MC) suit excellently the quantum-mechanical behavior of such a physical processes. In Figure 2.4 a pictorial graph of a pp collision produced with *Sherpa* 1.1 [35] is shown. Simulations in MC can be decomposed in two parts. The first corresponds to the *matrix element generation*, where mathematical computations for the process of interest are processed in different orders of perturbation theory. The second part deals with the parton shower and hadronization modelling. Several MC generators are used in this work to estimate background contributions and expected signal yields as predicted by the SM. MC generators such as *Pythia* [38, 39] or *Herwig* [40] proceed with both steps at the same time, using mostly leading-order matrix elements, which might not be sufficient for detailed studies. Nevertheless, higher orders in perturbation theory can be computed very precisely by interfacing the parton shower and hadronization with the matrix-element generators, such as *MadGraph* [41].

Chapter 3

The ATLAS detector at the LHC

The ATLAS Collaboration conducts fundamental research in particle physics through experimental measurements in pp collisions[1] at the Large Hadron Collider, located at the European Organization for Nuclear Research (CERN) in Geneva, Switzerland. More than 3000 members from 183 institutions, representing 38 countries all around the world, constitute this collaboration [42]. In this chapter, the Large Hadron Collider accelerator facilities and specifications are introduced in Section 3.1. In its context, the ATLAS detector is described in more detail in Section 3.2. To identify and reconstruct different kinds of particles, multiple algorithms and selection criteria are applied on the signal recorded in the different components of the ATLAS detector. The methodology followed in the ATLAS experiment from 2015 until 2018 is presented in Section 3.3.

3.1 The Large Hadron Collider at CERN

At CERN, the European Organization for Nuclear Research founded in 1954 in Geneva (Switzerland), a particle collider with unprecedented values of center-of-mass energies and rate of particle collisions is installed, the Large Hadron Collider (LHC). Hadrons (either protons or heavy ions) are accelerated approximately a hundred meters underground in a circular collider 26,7 km long. The LHC is composed of eight arcs and eight straight sections installed in the tunnel that previously hosted the LEP accelerator [44]. In collider experiments, the event production rate of certain process with production cross section σ is inferred from

$$\dot{N} = \frac{dN}{dt} = L \cdot \sigma \,, \tag{3.1}$$

where L corresponds to the *instantaneous luminosity*, representing the flux of colliding particles. The instantaneous luminosity L is defined by the experimental setup and it has a dimension of number of particles per area per time. To explore rare events, i.e., processes with small cross sections, the instantaneous luminosity needs to be increased to achieve a considerable statistical set to claim an evidence, a discovery or to carry out a precision measurement. The total number of interactions is then given by the *integrated*

[1]Eventually, heavy ions are accelerated at the LHC for specific studies and their products are measured with the ATLAS experiment. Nevertheless, only pp collisions are relevant for this thesis.

	LHC Runs				
	Run 1		Run 2		
	2011	2012	2015-2016	2017	2018
$\sqrt{s}$ [TeV]	7.0	8.0	13.0		
Integrated luminosity [fb^{-1}]	5.6	20.3	36.2	44.3	58.5

Table 3.1: Center of mass energy and integrated luminosities along different LHC Runs [43].

luminosity $\mathcal{L}$

$$N = \sigma \cdot \int L(t)dt \equiv \mathcal{L} \cdot \sigma \,, \tag{3.2}$$

whose value is measured in units of inverse barn. At the LHC, units of inverse femtobarn are commonly used.

Designed to produce particle collisions up to center-of-mass energies of $\sqrt{s} = 14$ TeV and $L = 10^{34}$ m^{-2}s^{-1}, limitations in technology have made the LHC to steadily increase its center-of-mass energy and instantaneous luminosity along time [43]. Periods between upgrades of both accelerator and detectors are known as *Runs*. The different LHC Runs at the LHC are classified in table 3.1. Values of center-of-mass energies and integrated luminosities are also given. The evolution of (cumulative) delivered integrated luminosity by the LHC along years for Runs 1 and 2 are also shown in Figure 3.1. Excellent performance of the LHC during the Run 2 has provided an important increase of instantaneous luminosity from 2016 until 2018 (blue curves).

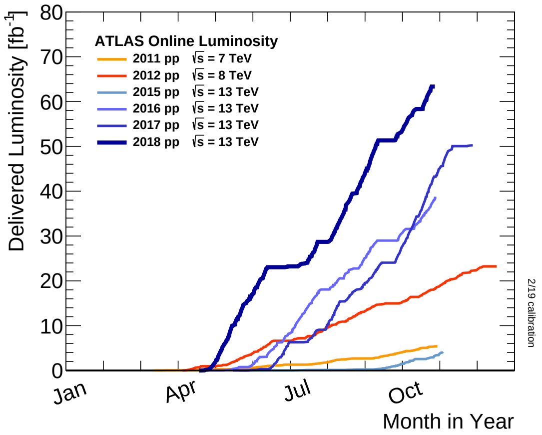


Figure 3.1: Cumulative luminosity versus day delivered to ATLAS during stable beams in *pp* collisions. Source [45].

Opposite magnetic dipole fields are needed to bend two counter-rotating proton beams on a circular path. A total of 1232 superconducting niobium-titanium (NbTi) dipole magnets, that produce a magnetic field of 8.3 T below critical temperature with superfluid helium at 1.9 K, are used to drive the particle beams. Bunches composed of $1.1 \cdot 10^{11}$ protons are obtained by ionizing hydrogen using electric fields. The extracted protons are then injected into the pre-acceleration chain. The pre-acceleration facilities of the LHC are presented in Figure 3.2, and it is constituted by the linear accelerator LINAC2, that accelerates the protons up to 50 MeV; the *Booster* and the *Proton-Synchrotron* (PS), speeding up the particles from 1.4 GeV up to 15 GeV; and finally the *Super Proton Synchrotron* (SPS) where protons are accelerated up to 450 GeV [46]. Protons at 450 GeV are then injected into the LHC and accelerated with a separation of 25 ns (approximately 7.5 m for maximum proton speed) by 16 radio-frequency cavities. The nominal filling scheme of the LHC allows the accelerator to introduce 2808 bunches per injection period per beam, colliding with a specific *crossing angle* to prevent parasitic collisions on both sides of the desired interaction point. The

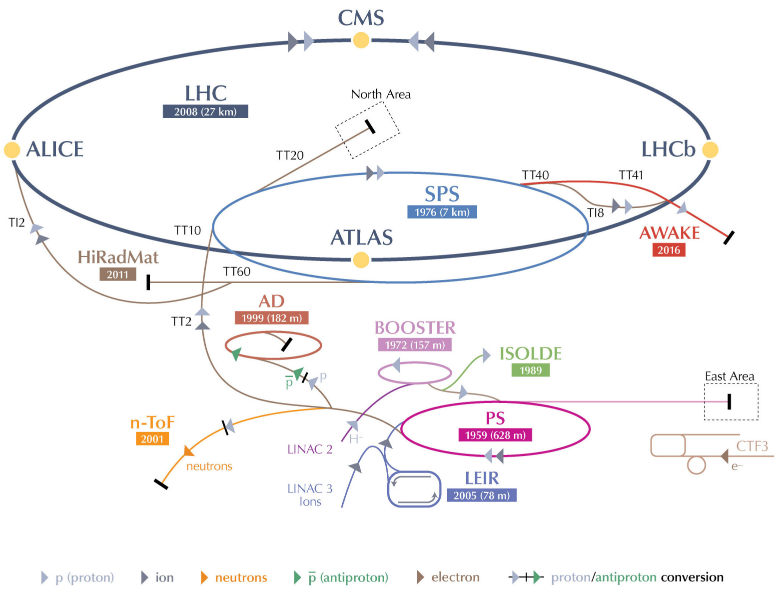


Figure 3.2: Particle acceleration facilities at CERN. Adapted from [48].

resulting instantaneous luminosity is given by

$$L = \frac{f\, n_b\, N_1\, N_2}{2\pi\sqrt{(\sigma_{x,1}^2 + \sigma_{x,2}^2)}\sqrt{(\sigma_{y,1}^2 + \sigma_{y,2}^2)}} \times S\,, \tag{3.3}$$

where f is the revolution frequency, n_b is the number of bunches, N_i is the number of protons in bunch $i = 1, 2$, and $\sigma_{x,i}$ and $\sigma_{y,i}$ denote the beam width of bunch i along x and y axes in the transverse plane, respectively. As a good approximation, beam widths for both bunches can be assumed to be equal. The factor S depends on the operating conditions (such as the crossing angle) and it usually takes values around 0.8 [47].

The collisions take place in four different *interaction points*, with their products recorded by eight different experiments. The four largest experiments that are set along the interaction points are ATLAS and CMS, general-purpose detectors of complementary design; the ALICE detector, that focuses on heavy-ion collisions and studies of confinement in QCD and properties of the quark-gluon plasma; and LHCb, which is designed to carry out optimized measurement on B-hadrons. TOTEM (that measures elastic scattering phenomena and total cross sections) is installed at the same interaction point as CMS. LHCf (that studies the origin of ultra-high-energy cosmic rays) and FASER (for searches of new light and weakly coupled elementary particles) are installed along with ATLAS; and MoEDAL (for direct searches of magnetic monopoles) sits with LHCb. This work is based on data recorded using the ATLAS experiment in pp collisions for the Run 2 period. The different components of the ATLAS detector are described in the following.

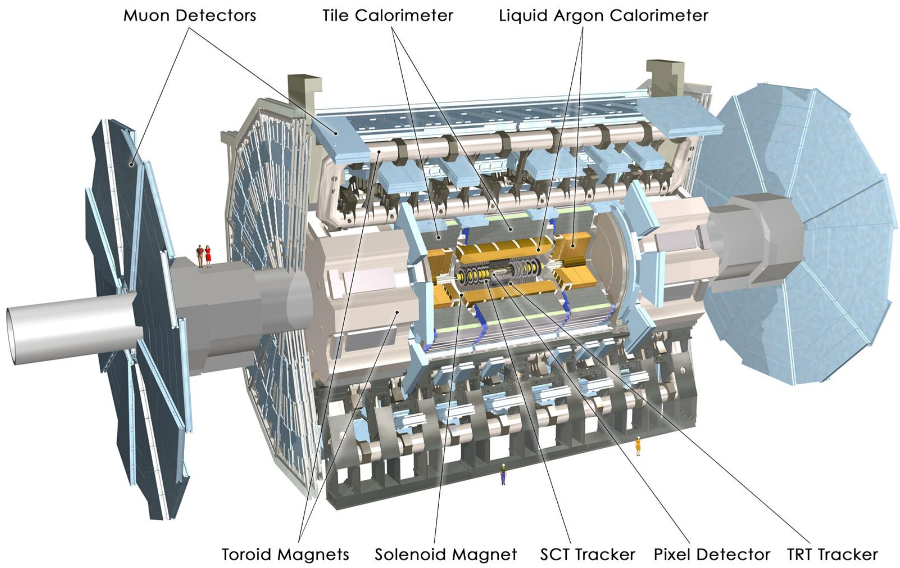


Figure 3.3: Schematic view of the ATLAS detector and its components. Taken from [49].

3.2 The ATLAS Detector

The ATLAS experiment[2] [50] is a particle detector at the LHC designed for precision measurements and searches for new phenomena in high-luminosity environments. The ATLAS detector is composed of different cylindrical layers that complement each other. The overall layout of detector is presented in Figure 3.3. According to this shape, a suitable coordinate system is presented in Section 3.2.1. The layers can be divided into four main groups according to their roles. The *inner detector* measures tracks of charged particles with very high precision and it is covered in Section 3.2.2. Measurements on the particle energies for photons and electrons, and separately for most of the hadrons, are carried out by the electromagnetic and hadronic *calorimeters*, respectively, described in Section 3.2.3. Finally, the *muon spectrometer* is covered in Section 3.2.4, used for precise measurements of muon four-momenta. Technical details of the ATLAS detector can be found in [50].

3.2.1 Coordinate System

At the ATLAS detector, the x-axis of the right-handed coordinate system points towards the center of the LHC, the y-axis is defined upwards and the z-axis is parallel to the beam pipe. For convenience, spherical coordinates are normally used. The origin of the coordinate system is determined *per event* at the nominal interaction point (whose definition and reconstruction are described in Section 3.3.1). The angular coordinates are the *polar angle* θ, defined as the angle from the beam axis and the *azimuthal angle* ϕ, which is measured around the beam axis in the transverse plane (corresponding to the $x-y$

[2] A Toroidal LHC ApparatuS.

plane). The *rapidity* y is defined with respect to the beam axis as

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) . \tag{3.4}$$

Contrarily to other space coordinates, the difference of rapidity between two objects is invariant under Lorentz boosts along the z-axis. Therefore, the rapidity is more convenient to describe kinematics in hadron collider experiments than the polar angle θ. Nevertheless, the *pseudorapidity* η is used instead since it is identical to the rapidity for massless particles (and they coincide in the high energy limit), directly related to the polar angle θ. It is defined as

$$\eta = - \ln \tan \left(\frac{\theta}{2} \right) , \tag{3.5}$$

taking values from 0 (transverse plane, $\theta = 90°$) to ∞ or $-\infty$ (when $\theta \to 0°$ or $\theta \to 180°$, respectively) along the beam axis. Based on these definitions, transverse variables such as the *transverse momentum* p_T are well defined in the $x - y$ plane in terms of η and ϕ as[3]

$$p \equiv |\vec{p}| = p_T \cdot \cosh \eta . \tag{3.6}$$

The angular separation in the pseudorapidity-azimuthal angle space between two objects can also be defined as

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} . \tag{3.7}$$

3.2.2 Inner Detector

The ATLAS Inner Detector (ID) [52, 53] provides a precise pattern recognition, high momentum resolution, and accurate measurements of tracks for electrically-charged particles above 0.5 GeV within a pseudorapidity of $|\eta| < 2.5$ and full coverage in ϕ. This allows a confident reconstruction of primary and secondary vertices. It is immersed in a 2 T magnetic field generated by a *thin superconducting solenoid* that surrounds the ID. Momenta of charged particles is then measured with high precision through the trajectory bending. The ID is composed of three complementary subdetectors that work independently. In Figure 3.4, the different subcomponents of the ID are presented. In the *barrel*, the ID subsystems are installed in concentric layers around the beam pipe in the central region. In contrast, the different layers are installed as transverse discs along the beam pipe in the *forward region*. Regions at both halves of the detector cylinder containing these transverse disks are referred to as *end-caps*.

Three barrel layers of *silicon pixel* allow for high resolution at pattern recognition from 33 up to 155 mm of the detector radii. Each layer provides a spatial resolution of 10 μm along the azimuthal direction and 115 μm along the axial (radial) direction of the barrel layers (disc layers). The high granularity of the pixel detector and its proximity to the interaction point is crucial for the reconstruction of primary and secondary vertices with high resolution. The ID was expanded in 2014 by inserting a new, innermost barrel layer called the *Insertable B-Layer* (IBL) [54] at 33.5 mm from the beam pipe. The IBL improves the track reconstruction performance by more than 40% [55], providing a better precision at vertex reconstruction and significant rejection of photon conversion.

[3] Properties $\cot\theta = \sinh\eta$ and $\cosh\eta = \sin^{-1}\theta$ are useful to derive these relations.

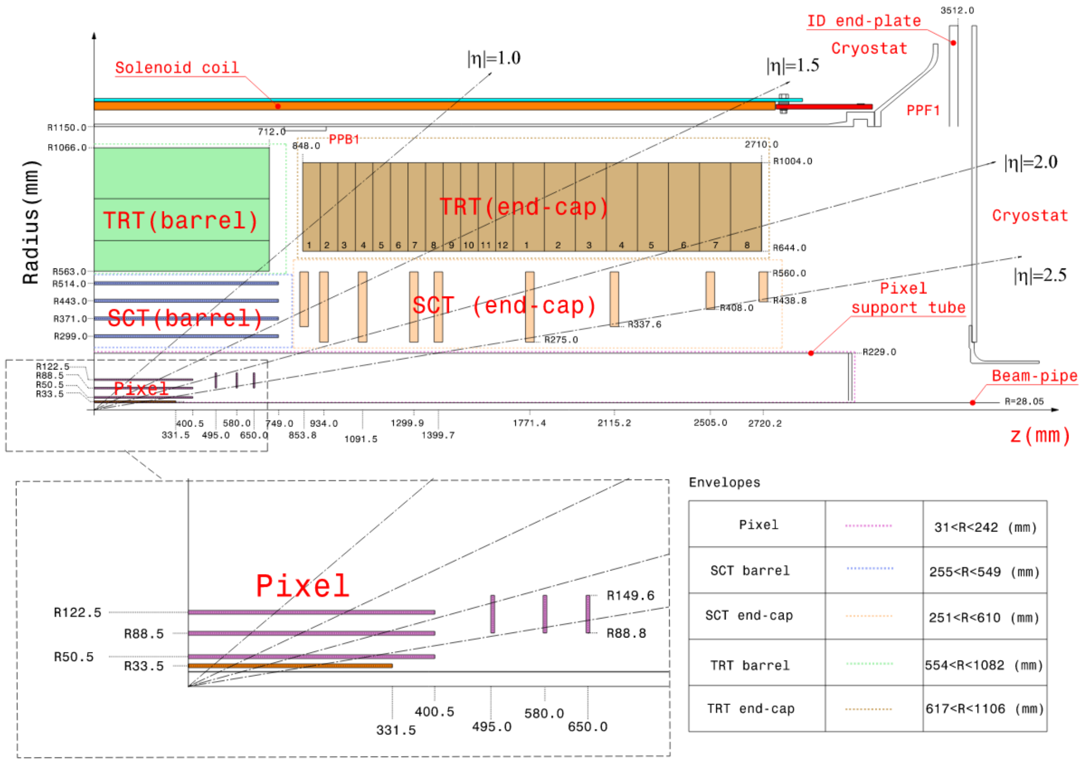


Figure 3.4: The $r - z$ cross section view of the layout of the ATLAS Inner Detector for Run 2. The top panel shows the entire Inner Detector, whereas the bottom panel shows a magnified view of the Pixel detector region, where the additional layer for the pixel detector introduced for Run 2 is visible. Taken from [51].

Four and two more layers of stereo pairs of *silicon microstrip* (SCT) span the detector radius from 299 up to 560 mm in the barrel and end-cap regions, respectively. Each module is composed of two layers of silicon micro-strip detector sensors glued back to back with a relative stereo angle of 40 mrad [56] to improve the measurement of the z coordinate and the radial coordinate for tracks of charged particles in the barrel and disc layers, respectively. The SCT layers provide around eight hits per track with an intrinsic resolution of 17 μm$\times$ 580 μm per layer.

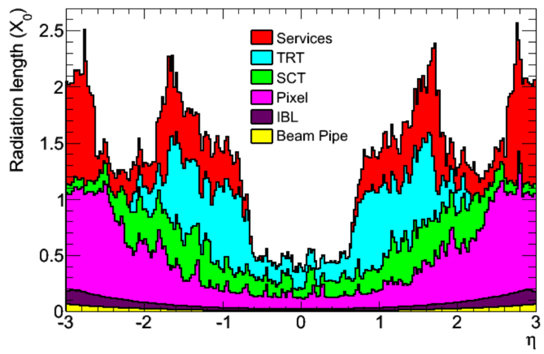


Figure 3.5: Material budget of the ATLAS inner detector shown in units of radiation lengths X_0 as a function of η for the different components of the ID. External IBL supports and services outside the active tracking volume are not included. Taken from [54].

Finally, several layers of gaseous straw tubes are compressed in the *Transition Radiation Tracker* (TRT) at the outer-most part, spanning the radial region from 563 up to 1066 mm. The TRT subcomponent of the ID provides particle identification and improvements on the momentum resolution for $|\eta| < 2.0$ and $p_T > 0.5$ GeV, with a substantial discriminating power between electrons and charged hadrons.

The amount of material in the ID detectors and its influence on the particles passing through it has implications in the particle reconstruction, especially for electrons and photons. Electrons can emit photons through *bremsstrahlung* and, similarly, photons can interact with the material producing e^+e^- pairs (also referred to as *photon conversions*). The probability of reconstruction for these particles is then affected by these phenomena. A detailed knowledge of the material budget is therefore crucial to account for this issue and carry out the pertinent corrections. The amount of material of the ID in the ATLAS experiment is shown in Figure 3.5. The IBL introduces additional material into the ID corresponding to 1.5% of a radiation length [51]. While the distributions are independent on ϕ, there is however a strong dependence on $|\eta|$. The increase of the material budget along η has implications in the reconstruction of the particle energies.

3.2.3 Calorimeter System

After passing through the ID, particles produced in *pp* collisions reach the calorimetry system [50], designed to efficiently perform at measuring energy and position of different kinds of particles. The *electromagnetic calorimeter* (EMCAL) is divided in two main components, the *presampler* (PS) and a high granularity liquid-argon (LAr) detector also divided in *front*, *middle* and *back* layers with granularities of 0.0031×0.1, 0.025×0.0245 and 0.05×0.025 in the $\eta \times \phi$, respectively. The finest granularity, offered by the front layer up to $|\eta| < 2.5$, allows for particle discrimination against multiple photon showers such as the ones produced in π^0 decays. Energy of electrons and photons is measured by fully absorbing them after the production of *particle showers*, whose shape allows to identify different kinds of particles interacting with the EMCAL. The energy resolution of the EMCAL is $10\%/\sqrt{\text{GeV}} \oplus 0.17\%$.

Hadronic objects (such as jets) have more penetration power and they are absorbed in the *scintillator-tile hadronic calorimeter* (HCAL) that surrounds the EMCAL. The HCAL is divided in three longitudinal segments with a coarser segmentation than the EMCAL (0.1×0.1 for the barrel region in the $\eta \times \phi$ space, getting coarser at higher η). A worse energy resolution of $50\%/\sqrt{\text{GeV}} \oplus 3\%$ is available for the HCAL in the barrel and end-cap layers.

3.2.4 Muon Spectrometer

Muons are the only charged particles that traverse the calorimeter system. Therefore, the recorded signal of charged particles outside the calorimeter system is a unique signature attributed to muons. The ATLAS Muon Spectrometer (MS) is made of 5500 individual muon chambers. The large size of the ATLAS experiment is primarily caused by an air-core *toroid magnet system*, providing a magnetic field to bend the muon tracks in the $r - z$ plane (instead of the $r - \phi$ plane as done in the ID) to obtain a high momentum resolution. The magnet system provides a magnetic field of 0.5 T in the central region and 1 T in the end-caps. Fast trigger of the detector relies on the Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGC) installed in the barrel and end-cap components of the muon spectrometer respectively up to $|\eta| < 2.4$. Details on the different subcomponents of the MS and their technologies can be found in Ref. [50].

3.2.5 Trigger and Data Acquisition Systems

The *ATLAS trigger system* is designed to select the pp collisions of interest for readout at an assumable rate that can be dealt by the storage capacities. The trigger system of the ATLAS detector in Run 2 is divided in two levels. A hardware-based first level trigger (Level-1) uses custom electronics to determine Regions of Interest (RoIs) in the detector, taking information from the coarse-granularity calorimeter and the muon detector as inputs. At this trigger level, the event rate is reduced from the LHC bunch crossing rate (of approximately 30 MHz) to 100 kHz with a decision time of 2.5 μs. The RoIs defined at Level-1 are sent to a software-based high level trigger (HLT) where detector information using the full granularity in either the RoI or the whole event is exploited by sophisticated selection algorithms. This second trigger level decrease the rate from the Level-1 output rate to approximately 1kHz within 200 ms of processing time. Further details and references can be found in Ref. [57].

3.3 Object Reconstruction and Particle Identification

The ATLAS detector uses the recorded signature of the event to reconstruct physical objects that are relevant in physics. To build these objects, different algorithms and selection criteria are applied to identify and reconstruct different kinds of particles, their interaction vertices and tracks with reliable identification and isolation performance with respect to background processes. In this section, the methodology followed in the ATLAS experiment in Run 2 is presented

3.3.1 Track and Vertex reconstruction

Particle tracks can be reconstructed for charged particles along the full ϕ range and $|\eta| < 2.5$ using the hits recorded within the ID. The MS also comes into play for the case of muons. These tracks are fitted to return a helix trajectory defined by a set of five parameters, denoted by d_0, z_0, ϕ_0, $\cot\theta$ and Q/p_T. The transversal and longitudinal *impact parameters*, d_0 and z_0, are defined as the distance of closest approach to the interaction vertex and its z coordinate, respectively. The parameter ϕ_0 is the azimuth at the point of closest approach, and $\cot\theta$ gives the inverse slope of the track in the $r - z$ plane. The *inverse transverse momentum* Q/p_T, with Q the charge of the particle, can be deduced from the measured radius of curvature. Their corresponding uncertainties are denoted by $\sigma(d_0)$, $\sigma(z_0)$, and so on. The ratios of the impact parameters over their uncertainties ($d_0/\sigma(d_0)$ or $z_0/\sigma(z_0)$, for transverse and longitudinal impact parameters, for instance) are referred to as *significances* and they are commonly used in analysis. Once these parameters are known, the analytical path of the track trajectory is completely defined. Details on the different approaches for particle track reconstruction can be found in Ref. [58].

Many primary vertices can be produced in a single bunch crossing, mainly due to superposition of multiple inelastic pp interactions. These multiple interaction points are usually referred to as *pile-up interactions*. The number of inelastic pp interactions per

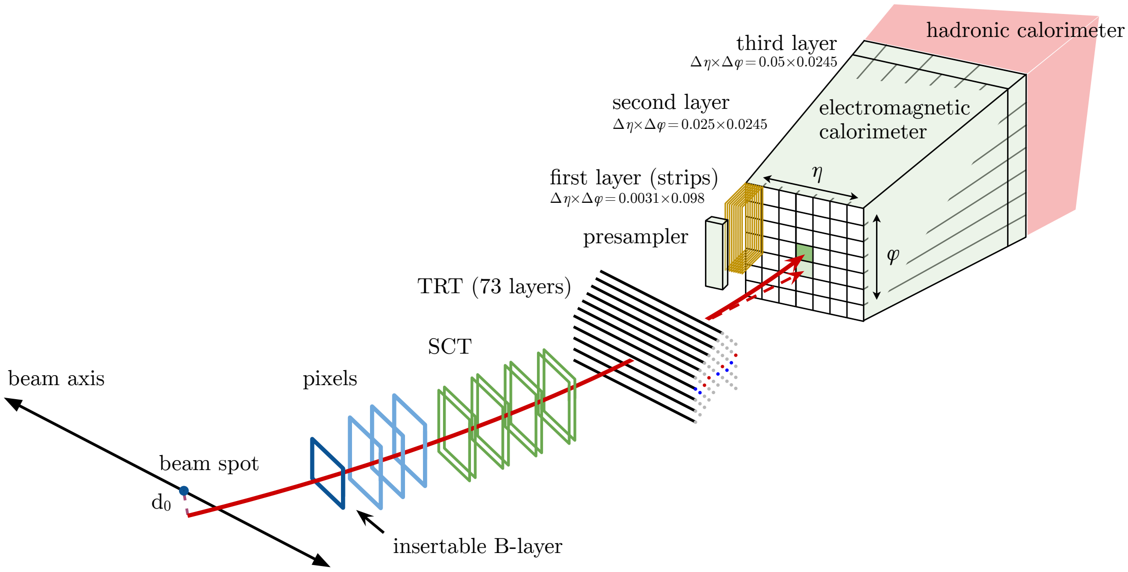


Figure 3.6: Illustration of the electron reconstruction and identification through its path along the different layers within the ATLAS inner detector. The hypothetical path of an electron (red) traverses the tracking system (pixel detectors, SCT and lastly the TRT) and then enters the electromagnetic calorimeter. Interaction of this electron with the detector material along its path through the tracking system can produce photons by means of bremsstrahlung. The dashed red line shows the path for one of these photons. Both particles would be finally absorbed by the electromagnetic calorimeter, with their energy being measured in different towers of the EMCAL that compose the cluster. Taken from [62].

bunch crossing under constant beam conditions is denoted by μ and is directly related to the instantaneous luminosity.

Primary vertex reconstruction is essential when inferring the full kinematic properties of the event. It is identified by matching reconstructed tracks pointing to an original common vertex. Details on the vertex reconstruction and pile-up determination can be found in Ref. [59]. A resolution of the longitudinal vertex position about 30 μm is achieved for events with high multiplicity of reconstructed tracks, with a resolution of the transverse position better than 20 μm [60, 61]. Estimations of the tracking efficiency with η and its dependency on η and the number of interactions per bunch crossing μ can be found in Ref. [61]. The used algorithms are shown to perform well in very dense environments. The track reconstruction efficiency decreases with greater values of $|\eta|$, being close to 90% for low μ and decreasing to 65-70% for denser pile-up environments.

3.3.2 Electrons

In the central region ($|\eta| < 2.5$), electron reconstruction is based on three fundamental components that characterize the electron signature: clusters of energy deposits recorded in the EMCAL, reconstructed tracks in the inner detector and the matching between them. Figure 3.6 illustrates the different elements of the ATLAS detector coming into play on electron reconstruction. The reconstruction starts from the energy deposits (clusters) in the electromagnetic calorimeter. Photons also follow a similar procedure to the one described here, taking into account that their presence in the inner detector can be only inferred from conversion vertices. Electron (and photon) energies are obtained using *dynamically-sized clustering algorithms* [63]. The overall cluster forms what is named *supercluster*. Details on this algorithm can be found in Ref. [63]. With respect to the older sliding-window approach [64], improvements in energy resolution of 30-40% are possible using superclusters. The resolution of measurements on the dilepton invariant mass are also improved by 5-10%, which are maintained for detector environments with higher pile-up [65].

An electron is reconstructed if at least one track is matched to the supercluster. In the case where several tracks are matched to the same cluster, tracks with hits in the pixel detector are preferred, and then tracks with hits in the SCT. Within each category, tracks with a better ΔR (see eq. (3.7)) match to the cluster in the second layer of the EMCAL are preferred. Following these preferences, a best-matching track is chosen to retrieve the electron kinematics. The four-momentum of central electrons can be computed using information from both the built supercluster and the highest-ranked track matched to it. The electron energy is given by the energy of the cluster. The ϕ and η coordinates of the electron candidate are taken from the corresponding parameters of the best-matching track at the vertex. The electric charge of the electron is also determined from the curvature of the best-matched track. The electron charge can be misidentified either by assignment of the incorrect track to the electron candidate or by a poor measurement of the track curvature in approximately 0.3% of the electrons. The sources of charge misidentification in the ATLAS detector are relevant for this thesis and they are studied quantitatively in Chapter 6. The procedure for energy calibration closely follows [66], updated with the aforementioned algorithm using superclusters for the energy reconstruction. The new energy measurement through the dynamical clustering algorithm provides variations of the energy scale with the average collisions per bunch crossing μ lower than 0.1%. Efficiencies in electron reconstruction within the energy regime to be used in this thesis are within 96.5-99.5% ($p_T > 20$ GeV) depending on the pseudorapidity. Deviations of the order of 5% can also be found for the increasing number of inelastic interactions per bunch crossing [63].

Reconstruction algorithms do not discriminate against background processes such as hadronic jets that can mimic electron signatures. In order to remove these contributions *electron identification selections* are defined. The identification of prompt electrons relies on a likelihood discriminant constructed from a set quantities related to the primary electron track, lateral and longitudinal development of the electromagnetic shower and the spatial compatibility track-cluster. Detailed descriptions of these variables and their definitions can be found in [62]. Three reference working points are defined, with increasing background rejection efficiencies, called *loose*, *medium* and *tight*. A list of variables involved in each of them along with their performances can be found in [67]. For the *Tight* working point (denoted by `TightLH`, used in this work), electron identification efficiencies measured in the data taking period 2015-2016 oscillate between 85-95% increasing with greater p_T and lower $|\eta|$ [62].

Furthermore, *isolation criteria* are needed when balancing between a highly-efficient identification of prompt electrons and good background rejection from heavy-flavour decays or light hadrons misidentified as electrons. Different working points are also defined for isolation [63]. The *Gradient isolation* working point will be used throughout this thesis, designed to give an η-uniform isolation efficiency of 90% at p_T = 25 GeV and

99% p_T = 60 GeV.

To reject the electrons with wrongly reconstructed charge, the *Electron Charge ID Selector* (ECIDS) has been developed based on a *Boosted Decision Tree* (BDT) discriminator, taking electrons with correctly reconstructed charges as signal events and charge-misidentified electrons as background events. To avoid the mismodelling of charge misidentification sources in Monte Carlo, the algorithm is trained using the 2015-2017 dataset (80 fb^{-1}) in $z \to ee$ decays. Input quantities to the BDT are the electron p_T and η, the charged transverse parameter (product $Q \times d_0$), the average charge of all tracks matched to the electron weighted by their number of hits in the SCT detector ($\bar{q}_{\mathrm{SCT}}$), the ratio of energy measurements taken by the calorimetry and tracking systems (E/p) and the $\Delta\phi$ between the cluster position in the second layer of the EMCAL and the momentum-rescaled track, extrapolated from the perigee, times the charge Q ($\Delta\phi_{\mathrm{res}}$). Approximately 90% of electrons with `TightLH` identification and `Tight` isolation requirements with incorrect electric charge are removed. Further details are given in Refs. [63, 68]. The ECIDS tool is applied in this thesis to reduce the contributions due to electron charge misidentification.

3.3.3 Muons

When traversing the ID and calorimetry system, muons lose a small fraction of their energy. Their reconstruction is based on the tracks fitted by the ID and MS together, along with energy depositions in the calorimetry system. Since muons are not absorbed by the detector components, their four-momenta are inferred from the reconstructed track. Details on muon reconstruction and performance measurements at $\sqrt{s} = 13$ TeV can be found in Ref. [69]. Muon reconstruction efficiency close to 99% within $|\eta| < 2.5$ and $5 < p_T < 100$ GeV. The isolation efficiency varies between 93 and 100%. Momentum resolution is measured to be 2.9% in $Z \to \mu\mu$ decays [69].

3.3.4 Jets

Color-charged particles produced in *pp* collisions hadronize before reaching the detector. This phenomenon is realized in the detector by the production of *jets*, whose signal is composed of groups of energy depositions and assigned reconstructed tracks. Jets are reconstructed by using the anti-k_T algorithm [70], which idealizes the jet as a cone spanned from the production vertex up to the calorimetry system. Selection efficiencies are greater than 95% [71] for p_T >20 GeV, increasing up to 99.5% for p_T >100 GeV.

The spatial resolution of the ID can resolve distances of several millimeters, which is sufficient for the flight length of hadrons containing b-quarks[4]. Based on the transverse and longitudinal impact parameters of the involved tracks and the reconstructed secondary vertices, the identification of *b-jets* (i.e. jets produced by b-hadrons) can be accomplished. *Artificial Neural Networks* exploit the characteristic topology of these decays to provide an efficient discrimination of b-jets from the rest of jet flavors over a large kinematic range. Details on the b-jet identification algorithms used in ATLAS during Run 2 and measurements of performance and efficiency are given in Refs. [72, 73]. Working

[4]Their typical lifetime is of the order of picoseconds [15].

points of 85%, 80%, 70%, or 60% are given according to the efficiency of b-jet identification.

3.3.5 Missing transverse momentum

Neutrinos and other hypothetical purely weakly interacting particles do not produce any signal on the different detector layers. Consequently, these particles introduce imbalances in the sum of total momenta of the measurable particles. An inference of their momentum can be done in the transverse plane using the *missing transverse momentum* E_T^{miss}. The components of E_T^{miss} and its azimuthal angle ϕ^{miss} can be obtained from

$$\left.\begin{array}{l} E_x^{\text{miss}} = -\sum_i E_i \sin\theta_i \cos\phi_i \\ E_y^{\text{miss}} = -\sum_i E_i \sin\theta_i \cos\phi_i \end{array}\right\} \begin{array}{r} E_T^{\text{miss}} = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2}\,, \\ \phi^{\text{miss}} = \arctan\left(\frac{E_y^{\text{miss}}}{E_x^{\text{miss}}}\right)\,. \end{array} \tag{3.8}$$

The resolution and performance of E_T^{miss} are highly dependent on the reconstruction efficiencies of every object i entering into the sum given by Eq. (3.8). Fully calibrated electrons, muons, photons, hadronically decaying τ-leptons, and jets reconstructed from calorimeter energy deposits and charged-particle tracks are used for the computation of Eq. (3.8), as well as soft-hadronic activity measured by reconstructed charge-particles tracks that were not associated with hard objects. The influence of pile-up is notorious [74]. Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV and further details in its reconstruction can be found in Ref. [75].

Chapter 4

Electroweak same-sign W pair production with two jets and its topology

Vector boson pair production is a proccess sensitive to electroweak vector boson scattering that can be studied by experiments at the LHC. Longitudinally polarized vector bosons in the final state are particularly interesting due to the relation to the electroweak symmetry breaking mechanism. As introduced in Section 2.2, the low production rate of these processes leads to challenging analyses for their measurements. In this chapter, the strategy to perform measurements on VBS at the LHC with special emphasis on the production of same-sign W boson pairs and two jets in the final state is presented. In Section 4.1, same-sign WW pair production is highlighted among different possible final states in VBS. In Section 4.2, its signature and arising backgrounds (Sec. 4.2.1) are discussed.

4.1 Electroweak same-sign W pair production with two jets for Vector Boson Scattering measurements

When two protons collide, interactions of vector bosons can take place when they are radiated out of the initial-state partons that take place in the hard process. The vector bosons are then deflected from the beam direction and their decays enter the detector volume. The two incoming partons involved are measurable in the final state as jets and they normally follow the direction of the collision axis. These two jets are denoted as *tagging jets* and carry the highest momenta in the event. A simplified presentation of the most relevant Feynman diagrams taking place in diboson production with two jets $VVjj$ is presented in Figure 4.1. Only leading-order (LO) diagrams are taken into account. Arising diagram contributions to the scattering amplitude of $VVjj$ proportional to $\alpha_{\text{EW}}^4\alpha_s^2$ are referred to as *QCD production* while terms proportional to the electroweak coupling constant as α_{EW}^6 are denoted by *electroweak production* (abbreviated as EWK). EWK production includes the contributions from triple and quartic gauge boson self-interactions diagrams, along with the mediation of Higgs boson. These diagrams contain the processes involving VBS (first row). Gauge bosons can also be radiated from one of

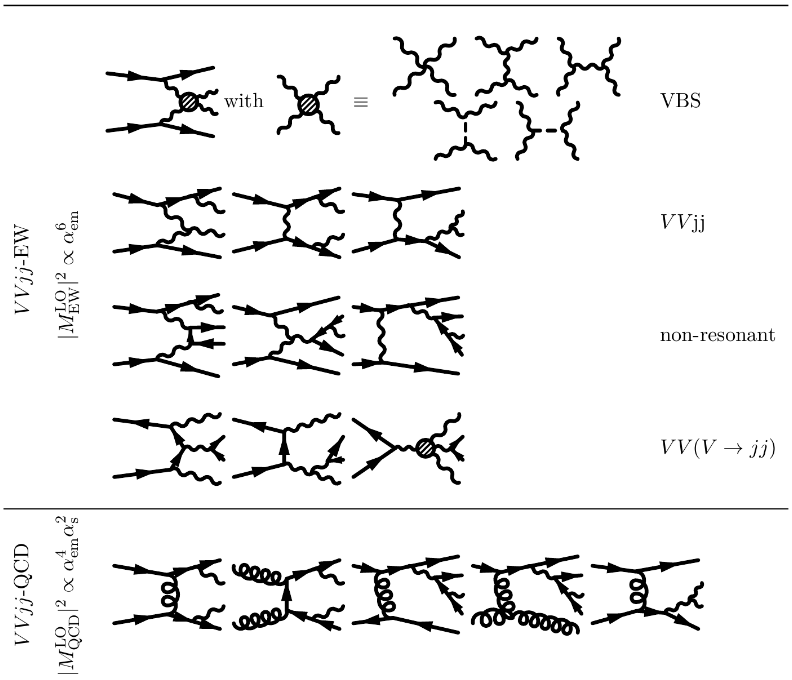


Figure 4.1: Feynman diagrams of diboson production in association with two jets at tree level. In the figure, α_{em} refers to the electroweak coupling (α_{EW} in the text). Although this is supposed to be a complete list, some diagrams are understood as representations of a whole set of diagrams with e.g. one final state gauge boson moved to another leg. If not required to construct a well-defined Feynman diagram, the possible decay products of the final state gauge bosons are not shown for simplicity. Source: [76].

the tagging quarks producing electroweak $VVjj$ in the final state (second row). Additionally, EWK $VVjj$ can be produced by non-resonant diagrams (third row) and triple production of gauge vector boson when one of them decays hadronically (fourth row). Differently to the EWK production, gluon-quark and gluon-gluon interactions arise as QCD sources of $VVjj$ (fifth row). As gluons are abundant constituents of the proton (see PDFs in Figure 2.3), QCD production is the dominant source for $VVjj$ in the final state. Even though QCD production can take place through triple gauge boson self-interactions (diagram on the right-hand side of the fifth row), it does not have any contribution from VBS. Hence, the QCD production of $VVjj$ is not sensitive to the electroweak symmetry breaking mechanism.

In quantum field theory, the total production cross section of $VVjj$ is proportional to the squared absolute module of the sum of all contributing scattering amplitudes,

$$\begin{aligned} \sigma &\propto |\mathscr{M}_{\mathrm{EWK}} + \mathscr{M}_{\mathrm{QCD}}|^2 \\ &\propto |\mathscr{M}_{\mathrm{EWK}}|^2 + |\mathscr{M}_{\mathrm{QCD}}|^2 + 2\,\mathrm{Re}\left(\mathscr{M}_{\mathrm{EWK}}\bar{\mathscr{M}}_{\mathrm{QCD}}\right) \\ &\equiv \sigma_{\mathrm{EWK}} + \sigma_{\mathrm{QCD}} + \sigma_{\mathrm{int}}\,, \end{aligned} \tag{4.1}$$

i.e. the total cross section can be seen as the sum of the pure electroweak production cross section σ_{EWK} (proportional to $|\mathscr{M}_{\mathrm{EWK}}|^2$), the pure-QCD production cross section σ_{QCD} (proportional to $|\mathscr{M}_{\mathrm{QCD}}|^2$) and an *interference term* σ_{int} proportional to the coupling $\alpha_s\alpha_{\mathrm{EW}}^5$, whose net contribution can be positive or negative. Studies on the contribution of σ_{int} claim that it is of the order of 3% of the total cross section and can be neglected as a first approximation [76]. However, its contribution might be crucial for precision measurements [77, 78, 79].

The electroweak component in Eq. (4.1) paves the way to test the electroweak symmetry breaking mechanism presented in Chapter 2. The isolation of this component with respect to the QCD production is crucial to proceed with the desired precision tests in this direction. Different processes involving VBS can be considered for this purpose according to their separation capabilities between QCD and EWK components. Since vector bosons are unstable particles of very short lifetime[1], they are measured in their *hadronic* or *leptonic* decay modes. In experimental particle physics, these modes are also referred to as *channels*. Even though the hadronic decay modes have a larger branching ratio, the separation of QCD and EWK components is unfeasible in this scenario due to the multiple-jet production rate at the LHC. Hadronic channels are therefore discarded. Contrarily, the leptonic channel ensures a good performance for these measurements. Experiments such as ATLAS provide lepton reconstruction and identification efficiencies for electrons and muons, as presented in Section 3.3. These efficiencies are considerably lower for τ leptons[2]. Therefore, only leptonic decay channels into electrons and muons are considered for VBS measurements.

The VBS process involving two same-sign W bosons has the largest signal-to-background ratio of all the VBS processes at the LHC, as found by the ATLAS [80, 81, 82] and CMS [83, 84] Collaborations at the center-of-mass energy of $\sqrt{s} = 8$ and $\sqrt{s} = 13$ TeV. Measurements on same-sign WW scattering (and VBS processes in general) are currently limited by statistics, an aspect that will improve in the near future for the results using the full Run 2 dataset, as well as analyses in future LHC Runs. Furthermore, the $W^{\pm}W^{\pm}$ scattering is the simplest VBS process for theoretical computations since the double-charge structure of the leptonic final state limits the number of partonic processes and the total number of Feynman diagrams to consider, while keeping all the interesting features of VBS. Hence, within the current experimental limitations at the LHC, electroweak production of same-sign W boson pairs turns out to be the best candidate to test the electroweak symmetry breaking mechanism and constraint the Higgs couplings with precision measurements within the SM. A more detailed discussion on the possible final states from $VVjj$, their sensitivity to VBS and experimental limitations can be found in [77].

4.2 Experimental signature of same-sign W pair production with two jets

The electroweak production of same-sign W boson pairs with two jets ($W^{\pm}W^{\pm}jj$) is characterized by the presence of two forward jets originated by the outgoing partons in the pp collision, two charged leptons (either electrons or muons) of identical electric charge and missing transverse momentum provided by the neutrinos arising from the W boson decays. Aiming an increase in sensitivity for the electroweak production of $W^{\pm}W^{\pm}jj$ and

[1]Half-life for electroweak vector bosons is of the order of 10^{-25} seconds so that they do not even leave the primary vertex (in terms of spatial resolution in the detector, see Section 3.3.1) before they decay into lighter particles.

[2] They decay rapidly into lighter leptons and neutrinos, introducing an important contribution in the missing transverse energy to be measured.

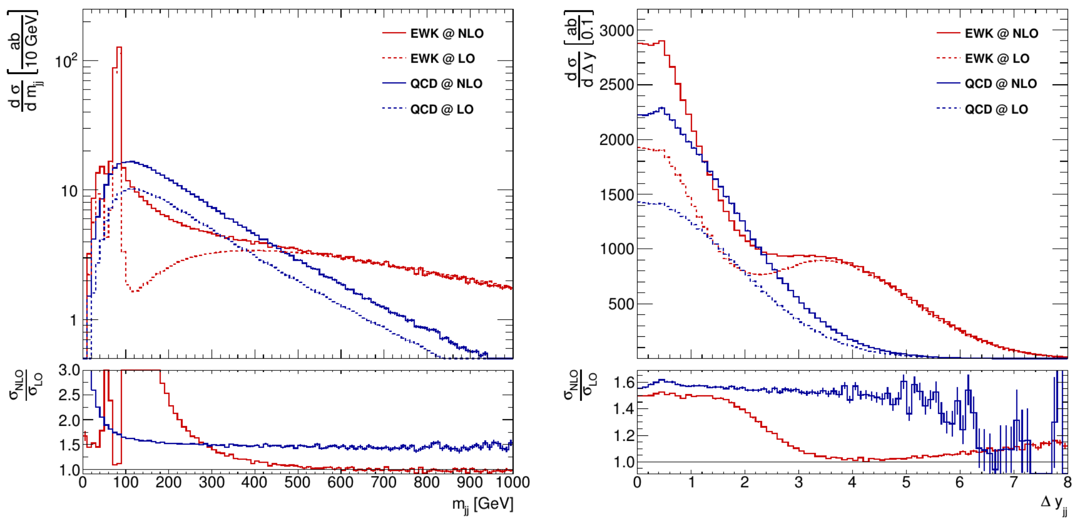


Figure 4.2: Differential cross section distributions are shown for the electroweak (red) and QCD (blue) production of the $pp \to e^{\pm}\mu^{\pm}\nu_e\nu_m jj$ process at $\sqrt{s} = 8$ TeV at LO and NLO QCD accuracy are presented. LO contributions are shown as dashed lines while contributions from NLO are represented by solid lines. The invariant mass of the two tagging jets m_{jj} (left) and the absolute value of the difference in rapidity between the two tagging jets $|\Delta y_{jj}|$ (right) are shown. At the bottom of each plot, the ratio between NLO and LO is shown for both QCD and EWK components, showing the relevance of considering such a contribution from the perturbative series along kinematic regions. Source: [77].

VBS processes (motivated in Section 4.1), the suppression of QCD production is studied. Differential cross sections for electroweak and QCD production of $pp \to e^{\pm}\mu^{\pm}\nu_e\nu_m jj$ are presented in Figure 4.2. The most remarkable differences for QCD and EWK production cross sections can be found on the dijet kinematics in $W^{\pm}W^{\pm}jj$. A clear separation on the absolute module of the difference in rapidity of the two tagging jets Δy_{jj} is observed. Jets from EWK production of $W^{\pm}W^{\pm}jj$ exhibit a maximum around $|\Delta y_{jj}| = 4$ at LO. QCD-produced tagging jets are more prone to be fully transversal to the collision axis (maximum at $|\Delta y_{jj}| = 0$). Additionally, the production cross section of $W^{\pm}W^{\pm}jj$ as a function of the invariant mass of both tagging jets m_{jj} shows a maximum at low values for the QCD component. This distribution falls rapidly with greater values of m_{jj}. This dependence is shifted to greater values in EWK production. Therefore, measurements of $W^{\pm}W^{\pm}jj$ are more sensitive to the EWK production when considering regions of high invariant mass of the two tagging jets as well as jet rapidities with $|\Delta y_{jj}| > 2$. The observed differences can be exploited to find an optimal signal region for measurements on EWK production of $W^{\pm}W^{\pm}jj$.

4.2.1 Background processes

The low production rate of electroweak same-sign W boson pairs along with two jets leads to other SM processes to become prominent background contributors for its measurement. Especially, low inefficiencies in lepton reconstruction and identification become important when reconstructing same-sign leptons in the final state from other SM processes. Different categories can be defined for the classification of the background contributions to the measurement of electroweak production of $pp \to W^{\pm}W^{\pm}jj \to$

$\ell^{\pm}\nu\ell^{\pm}\nu jj$. These categories are

- **QCD production of same-sign W boson pairs** with two jets in the final state (abbreviated as ssWW-QCD) lacks interest for accessing the electroweak symmetry breaking mechanism of the SM. Then, its contribution is removed for a better isolation of the EWK component of the $W^{\pm}W^{\pm}jj$ production.

- **Prompt lepton background contributions** arise from other SM processes that mimic the final state of interest. Processes such as $WZjj$, $ZZjj$ or $t\bar{t} + W^{\pm}/Z$ compose this category when some of the final-state leptons are lost outside the event selection or due to misreconstruction. The contribution of WZ is normally considered in its own category since it constitutes the main background source for this measurement.

- **Non-prompt lepton production** is a crucial background source at the production rate scale of the EWK production of $W^{\pm}W^{\pm}jj$. The presence of charged leptons produced by secondary interactions (e.g. hadron decays into leptons) can be reconstructed as fake prompt electrons and muons. Furthermore, imperfections from lepton identification algorithms can allow peculiar detector signals to be reconstructed as lepton candidates (even when there might not be a lepton interacting with the detector). Due to the large cross section for multijet production or $W^{\pm}+$ jets, non-prompt lepton production constitutes a prominent background source even though the misidentification rates are very low.

- **Prompt photon conversion** into $e^{+}e^{-}$ pairs, where only one of the electrons passes reconstruction, identification and selection criteria, can mimic the final state of $W^{\pm}W^{\pm}jj \rightarrow \ell^{\pm}\nu\ell^{\pm}\nu jj$ in processes such as $W\gamma$ or $Z\gamma$.

- **Electron charge misidentification** (also referred to as *charge flip* or *charge misID* for brevity) becomes an important background source in decay channels involving electrons. In SM processes decaying into two opposite-sign electrons (whose production rate is orders of magnitude greater than the production rate of $W^{\pm}W^{\pm}jj$, e.g. $Z +$ jets, $t\bar{t}$ or $W^{\pm}W^{\mp} +$ jets), background photons radiated off one of the prompt electrons can convert into an $e^{+}e^{-}$ pair. The track of the electron in this pair with opposite charge respect to the original prompt electron can be wrongly tagged as the prompt electron track, providing the wrong electric charge from its curvature. These will then pass the selection criteria applied for the $W^{\pm}W^{\pm}jj$ measurement. Additionally, inefficiencies of the track reconstruction algorithm can lead to misidentification of the track curvature orientation that leads to a misidentified electric charge. Charge misidentification has been proved to be negligible for muons in previous ATLAS analyses [85].

In this thesis, the contributions due to the electron charge misidentification are estimated for the measurement of electroweak production of $pp \rightarrow W^{\pm}W^{\pm}jj \rightarrow \ell^{\pm}\nu\ell^{\pm}\nu jj$ using the full Run 2 dataset recorded with the ATLAS detector at $\sqrt{s} = 13$ TeV. In Chapter 6, an overview of the relevance of this background contribution along the analyses carried out by the ATLAS Collaboration is given (Section 6.1) and studies on charge misidentification sources (Section 6.2) along with an analysis in detector simulation (Section 6.3) are presented.

Chapter 5

Dataset, simulated events and event selection

In this chapter, the data and MC sets used for the measurements presented in this work are described. The used samples are briefly described in 5.1. From these samples, electron, muon and other physical objects are defined for physics measurements by applying several selection criteria. Selections applied for this work, as well as the definition of the signal region (SR) for VBS measurements are described in Section 5.2

5.1 Dataset and simulated events

In this work, the full Run 2 dataset recorded with the ATLAS detector in *pp* collisions is used. The data taking periods are referred to as 2015-2016 (36 fb^{-1}), 2017 (44 fb^{-1}) and 2018 (58 fb^{-1}) datasets according to the classification given in Table 3.1. To avoid biases during the measurement, the blinding strategy of the analysis only allows looking at data events from 2015-2016 in the SR (defined in Section 5.2), as measurements were already published by the ATLAS Collaboration for this dataset [82].

MC simulations with Madgraph [41] in $Z \rightarrow ee +$ jets are used for the study of the charge misidentification sources in Chapter 6. The same samples are used for the computation of the charge misidentification probabilities, energy correction factors and closure tests in Chapter 7. Generated events are processed through a detector simulation [86] based on *GEANT4* [87]. The events are reconstructed identically as the data, using the same reconstruction algorithms. The simulation includes pile-up effects as described in Ref [88]. Special *MC campaigns* are designed to reproduce the detector environments during these periods of data taking. These are called *mc16a*, *mc16d* and *mc16e* for 2015-2016, 2017 and 2018 data-taking periods, respectively. Signal and background predictions are also produced in MC simulation.

The predictions of the prompt background contributions are produced in MC simulations. This category is divided into WZ (containing QCD and EWK production), ZZ and *top* (including single-top, $t\bar{t}$ and $t\bar{t} + X$, $X = W, Z$ contributions). Contributions from photon conversions are also computed in MC, including samples in $W\gamma$ and $Z\gamma$.

This category is labeled as $V\gamma$. Additionally, signal samples (i.e. EWK production of $W^{\pm}W^{\pm}jj$, referred to as ssWW EWK) and QCD production of $W^{\pm}W^{\pm}jj$ (ssWW QCD) are produced in MC. Background yields arising from electron charge misidentification are estimated using a data-driven method that is the main focus of this thesis. Chapter 7 describes the methodology in detail. Lastly, non-prompt background is simulated using data-driven techniques. The contributions from this background source, as well as from other prompt background sources (e.g. triboson), are neglected in the control regions used along this thesis as they are small in these and do not affect the conclusions. The list of samples used in this thesis is given in Appendix A.

The *Physics-xAOD* (PxAOD) event data model is used for data and MC samples, analyzed using the ATLAS Common Analysis Framework (CAF). Two ATLAS software releases are used, 20.7 (used during the 2015-2016 analysis [82]) and release 21. These will be referred to as Release 20 and Release 21 along the next chapters. Only differences in the electron energy reconstruction (with the latest using the topological cell clustering-based approach described in Section 3.3.2) were relevant between software releases for this work.

5.2 Selection criteria

5.2.1 Event preselection

Events are selected online by single-electron or single-muon triggers [89]. These are listed in Appendix A.1 for different data taking periods. The trigger efficiency in simulated events is corrected to the efficiency in data based on efficiencies scale factors provided by the Egamma and muon CP groups. To ensure that one of the signal leptons fired the trigger, each event is required to have at least one electron or muon matched to a trigger object, with a p_T above the p_T threshold of that object.

As many collisions per bunch crossing can occur, the vertex with greatest sum of squared transverse momenta from the associated tracks is selected as the primary vertex, which is used to compute all its associated quantities such as track impact parameters. Only events where the primary vertex has at least three associated tracks are considered. Contributions from additional pp collisions are simulated in MC by overlaying minimum bias events from Pythia8. Differences in the pileup conditions in data and simulation are corrected by reweighting the distribution of the average number of interactions per bunch crossing $\langle\mu\rangle$. The distribution in data is shifted by a factor of 1/1.09 in the procedure.

5.2.2 Object selection

In this section, the selection criteria for all the different objects used in the measurement of same-sign W boson pair production along with two jets is presented.

For leptons, two sets of selection criteria are defined. First, loose *preselection critera* are imposed, which have been optimized to reject multilepton processes. Ultimately, tighter *selection criteria* are applied, aiming at a selection of the signal leptons with high purity. Muons are required to fulfil a set of quality requirements recommended by the Muon CP group. Kinematic and geometrical cuts are applied to reject low-p_T muons coming from additional pile-up collisions and within the detector acceptance, imposing `Loose` identification with no isolation requirement, with $p_T > 3$ GeV and $|\eta| < 2.7$ for the muon preselection. Additionally, the transverse and longitudinal impact parameters must fulfill $|z_0 \sin\theta| < 0.5$ mm and $|d_0/\sigma_{d_0}| < 10$. The signal muons are further selected with `Medium` identification and `FixedCutTight` isolation criteria to reduce the contributions from background sources. The selected muon must have $p_T > 27$ GeV, $|\eta| < 2.5$, and impact parameters of $|z_0 \sin\theta| < 0.5$ mm and $|d_0/\sigma_{d_0}| < 3$ to ensure its origin at the primary reconstructed vertex. The electron energy is reconstructed in the EMCAL. The electron preselection requires `Loose` identification and no isolation requirement, with $p_T > 4.5$ GeV and $|\eta| < 2.47$. Additionally, the electron track must have $|z_0 \sin\theta| < 0.5$ mm and $|d_0/\sigma_{d_0}| < 5$. Then, signal electrons are selected with `Tight` identification and `Gradient` isolation requirements, having $p_T > 27$ GeV, $|\eta| < 2.5$ and impact parameters of $|z_0 \sin\theta| < 0.5$ mm and $|d_0/\sigma_{d_0}| < 3$. The crack region $1.37 < |\eta| < 1.52$ is excluded for the electron selection.

Jets are reconstructed from the topological calorimeter clusters using the anti-k_T algorithm [70] with a resolution parameter of $R = 0.4$ and preselected by requiring $p_T >$ 20 GeV and $|\eta| < 4.5$. The selected, calibrated jets are required to have $p_T > 30$ GeV in the forward region of the detector ($2.4 < |\eta| < 4.5$) and $p_T > 25$ GeV in the central region ($|\eta| < 2.4$). For jets passing this selection, the `MV2c10` tagger is used to identify b-jets with the 85% efficiency working point.

The missing transverse momentum E_T^{miss} is determined by the energy collected by the electromagnetic and hadronic calorimeters and by the muons measured by the muon spectrometer and inner detector.

The preselected objects are subject to an *overlap removal* procedure that attempts to remove possible duplications due to the same particle being reconstructed by two or more reconstruction algorithms. The standard approach used in ATLAS can be found in Ref. [90].

5.2.3 Event selection and the signal region

For the event selection, two isolated, well reconstructed same-sign leptons (either electron or muon) are required, having $p_T > 27$ GeV. In order to suppress contributions from WZ, events with only two leptons are kept. B-jets are vetoed to suppress contributions from $t\bar{t}$. Additionally, the invariant mass of the two leptons is required to be $m_{\ell\ell} > 20$ GeV, with an additional Z window veto in the ee channel ($|m_{ee} - m_Z| >$ 15 GeV) to mitigate contributions due to charge misidentification. To suppress the charge misidentification background, the electron pseudorapidity in the ee channel is restricted to $|\eta| < 1.37$. The missing transverse energy is also required to fulfill $E_T^{\text{miss}} > 30$ GeV.

The SR pursues the selection of electroweak VBS processes motivated by the discussion in Chapter 4. For this purpose, two jets of $p_{T,j0} > 65$ GeV and $p_{T,j1} > 35$ GeV are required. The pseudorapidity of both must satisfy $|\eta| < 4.5$, with a difference of rapidity $|\Delta y_{jj}| > 2$. Finally, the invariant mass of the two jets must be $m_{jj} > 500$ GeV.

Chapter 6

The charge misidentification background and its sources

Charge misidentification becomes an important background contribution to the measurement of processes with two same-sign leptons in the final states involving electrons. As the main topic of this thesis, the relevance of the charge misidentification background in the context of same-sign WW production measurements in hadron colliders will be reviewed in 6.1. The sources of charge misidentification are classified in Section 6.2. In Section 6.3, kinematic and track properties of the different identified electron types are presented. The estimation of the background contribution to the measurement of $W^{\pm}W^{\pm}jj$ using the full Run 2 dataset (139 fb^{-1}) is computed for the ATLAS Collaboration. Methodology and results are discussed in Chapter 7, as well as approaches for its reduction.

6.1 Relevance of the charge misidentification background for the same-sign WW analysis

The measurement of electroweak production of $W^{\pm}W^{\pm}jj$ deals with contributions from electron charge misidentification as one of the most important background sources for the leptonic channels involving electrons. Concretely, it is the third largest contribution, after WZ and non-prompt sources. Proportions are presented in Figure 6.1 for the background and signal yields in the signal region considering the sum of all measured decay channels (referred to as *all-inclusive* $\ell\ell$ channel). It is worth to notice the uncertainties that come along these yields. In the values provided in Figure 6.1, systematic and statistical uncertainties are summed in quadrature. Uncertainties of 28%, 52% and 26% with respect to the nominal value are given for contributions of WZ, non-prompt and e/γ conversions, respectively. The uncertainty of the estimated non-prompt background yields is approximately 50% in the $\mu\mu$ final states and varies between 40% and 90% for the ee and $e\mu$ decay channels. Theoretical uncertainties dominate the WZ contribution.

The focus of this thesis is the electron charge misID. For the analysis using the dataset of 36 fb^{-1}, the main sources of uncertainty for this background contribution come from

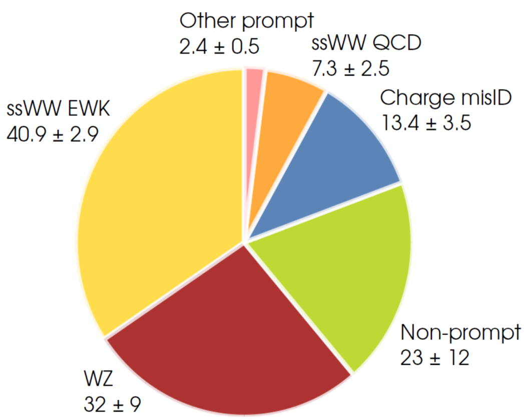


Figure 6.1: Number of events in the signal region for the electroweak production of same-sign W boson pairs (signal, abbreviated as *ssWW EWK*) and different background yields contributing to the measurement. The yields from the four channels taking place in the measurement (ee, $e\mu$, μe and $\mu\mu$) are summed. Both prompt photon conversions and background sources due to electron charge misidentification are considered in the blue yield. Source: [82].

	e^+e^+	e^-e^-	$e^+\mu^+$	$e^-\mu^-$	μ^+e^+	μ^-e^-
Yield ($\pm$ stat)	1.33 ± 0.08	1.17 ± 0.08	2.36 ± 0.28	1.85 ± 0.21	1.11 ± 0.14	1.10 ± 0.15
SF up	+0.13	+0.10	+0.08	+0.06	+0.03	+0.03
SF down	-0.10	-0.10	+0.03	-0.04	-0.02	-0.02
Energy corr	$\pm$0.05	$\pm$0.03	$\pm$0.20	$\pm$0.18	$\pm$0.01	$\pm$0.01

Table 6.1: Estimated event yields for the electron charge misID background contribution in the signal region are presented for the last analysis in ATLAS using the 2015-2016 dataset [82]. In this table, the first lepton in the pair indicates the lepton with the highest measured p_T in the event. Central values and systematic uncertainties are listed in the table for the ee, $e\mu$ and μe channels, for positive and negative final states separately. Systematic uncertainties due to the used methodology arise from variations on the applied scale factors (SF).

statistical uncertainties in the measurement of the charge misID rates (whose computation is detailed in Chapter 7). Other possible biases are in the estimation method, being less than 10% for the forward region $|\eta| > 2$ and up to 20% in the central region [82]. Two main analyses in ATLAS, using Run 1 [80] and the partial 2015-2016 dataset of Run 2 [91], are available to track the improvements on the charge misID background determination. Detailed values along decay channels are provided in Table 6.1 for background yields, statistical and systematic uncertainties in the signal region. Even though the selection criteria for the signal region was optimized for the later (so that the compared signal regions might differ), important improvements can be noticed for the charge misID background estimation. Aside statistical uncertainties that are reduced due to the use of larger datasets, systematic uncertainties were importantly reduced, going from total uncertainties of 19% (ee channel) and 37% ($e\mu$ channel) in Run 1 [80] to 7% (ee) and 9.6% ($e\mu$) of the nominal value in the 2015-2016 dataset [91], respectively. Furthermore, the decomposition of sources for systematic uncertainties on the latest available analysis is found for individual channels in lepton flavor and electric charges in the final state [91]. For the ee channel, the main source of systematic uncertainty is found in the used methodology through the use of scale factors (SF). The restrictions $|\eta| < 1.37$ and a Z mass window

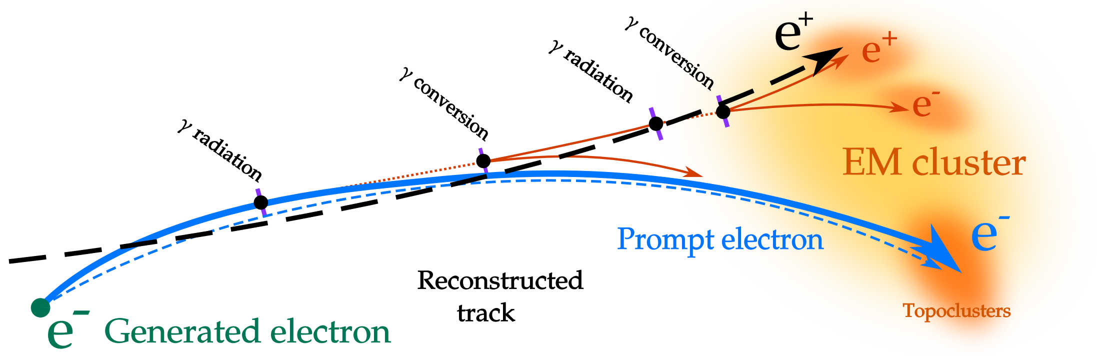


Figure 6.2: Sketch on the electron charge misidentification due to bremsstrahlung radiation when interacting with the detector material. The electron produced in the process of interest (green) is referred to as *generated electron*. Once it enters the detector, the genrated electron becomes the *prompt electron* (blue) which interacts with the detector material (linac rectangles) along its path. At the interaction points (black dots) photons can be produced through bremsstrahlung (labeled as γ radiation), photons that can subsequently decay into e^+e^- pairs (γ conversions). The overall shower of electrons and photons is then absorbed by the electromagnetic calorimeter, resulting in a spreading cluster. Among the various assigned tracks to the electron candidate (represented by dashed lines) for the subsequently produced electrons and photons, the one from a electron with opposite charge can be selected as the best-matching track, retrieving the wrong electric charge for the electron candidate (black e^+).

$|m_Z - m_{\ell\ell}| < 15$ GeV are applied to suppress the contributions of this background in the ee channel. The main source in the $e\mu$ channels comes from the applied method for the energy correction (discussed in detail in Chapter 7), contrarily to the μe channel. The reason of this difference was not found in the reference. Since the first lepton of the pair labeling the channel refers to the lepton of highest reconstructed p_T, the different background yields observed between the $e\mu$ and μe channels shows an increase of the electron charge misidentification phenomenon at higher energies. Even though the full pseudorapidity range is considered for $e\mu$ and μe channels, statistical uncertainties are dominant for the μe channel, where the expected background yields are lower.

6.2 Sources of charge misidentification in the ATLAS detector

The ATLAS detector provides excellent electron reconstruction efficiencies throughout a wide range in transverse momentum and pseudorapidity. Nevertheless, charge misreconstruction becomes an important background contribution when considering decays into final states with two same-sign leptons in measurements with constrained cross sections. For these cases, the expected number of signal events can become comparable to the unlikely effect of charge misidentification.

The electric charge of the electron is inferred from the track curvature reconstructed in the inner detector. The inner detector material (whose distribution is shown in Figure 3.5) lessens the signature quality left by the electrons in terms of irregular track-to-cluster assignments or poor reconstruction of the track curvature. Previous studies on charge misidentification in the ATLAS detector confirm this hypothesis [92], identifying *bremsstrahlung* and *straight-like tracks* produced at high-p_T as main sources of electron charge misID.

In Figure 6.2, an illustration on the misidentification of the electron charge due to bremsstrahlung is presented. An electron is created in the process of interest, e.g, by a MC generator. This electron, referred to as *generated electron*, is shown in green on the left-hand side of Figure 6.2. The generated electron enters the detector (or seeds the detector simulation in MC), becoming the *prompt electron* along its path. Its trajectory is depicted by a solid blue arrow. While the prompt electron travels through the detector material (shown as small lilac rectangles), interactions take place. Points of interaction are shown as black dots. These interactions can produce photons (γ *radiation*) through bremsstrahlung that can later decay into e^+e^- pairs (γ *conversion*). In the end, the most energetic component of the shower produced by these subsequent production of electrons and photons is absorbed by the electromagnetic calorimeter as a spreading cluster (blurry orange EM cluster). Lastly, multiple tracks (dashed arrows) are reconstructed by matching the EM topoclusters to the hits produced in the ID by the prompt and the subsequently produced electrons. Among all the tracks assigned to the electron candidate, the reconstructed electron charge (black e^+) is then inferred from the curvature of the best-matching track (that can be taken from an electron with opposite charge respect to the prompt electron) that fulfills the preferences described in Section 3.3.1, e.g. when the electron with wrong charge has the track of highest transverse momentum.

Electrons can also radiate by *Final State Radiation* (FSR), mimicking the signature presented in Figure 6.2. A Feynman diagram for this process is shown in Figure 6.3. The electron in the final state emits a photon collinear to its trajectory, acquiring an important amount of the electron four-momentum. Then, when traversing the detector material, the FSR photon produces an electron-positron pair that is reconstructed as electron candidate. Contrarily to bremsstrahlung (produced by interactions with the detector material), FSR radiation takes place at generation level and no detector effects are involved a priori. Consequently, FSR is not considered as a source of charge misID.

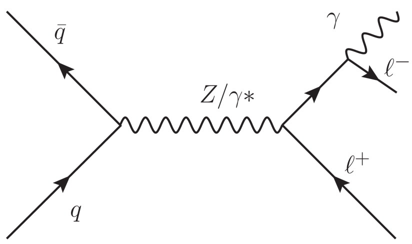


Figure 6.3: Feynman diagram for FSR radiation in $Z/\gamma*$ leptonic decay into $\ell^+\ell^-$.

Charge misidentification can also arise from a mismeasurement of the electron track curvature. For a charged particle traveling with velocity $\vec{v}$ within a homogeneous magnetic field $\vec{B}$ such that $\vec{v} \perp \vec{B}$, the trajectory is bent with a certain radius of curvature proportional p_T/Q, where Q is the charge of the particle, in a plane perpendicular to $\vec{B}$. This proportionality depends on the intensity of the magnetic field, $|\vec{B}|$ [93]. For a constant B, the radius of curvature is only dependent on p_T, with the trajectory being more bent for decreasing p_T. Charged particles with very high transverse momentum can, therefore, pass through the inner detector with very low curvature. Furthermore, the fit quality of the track is reduced within regions of high $|\eta|$. For instance, the end-cap layers of the TRT detector do not record signature for particles with $|\eta| \gtrsim 2.0$ (see Figure 3.4). Particle trajectories within the inner detector are difficult to be bent along the forward region since these particles travel parallel to the direction of the magnetic field. In Figure 6.4, a illustration of this source for charge misID is provided. A high-p_T prompt electron is poorly bent by the magnetic field (blue). The signature left by the prompt electron (blue dot) in the EMCAL (orange spot) is correctly assigned to the correspondent generated particle (green). Due to a reduced bending of the reconstructed track and a poorer quality of its fit (shown as a dashed line), the orientation of the track bending can be misleading, providing a wrong reconstruction of the electric charge.

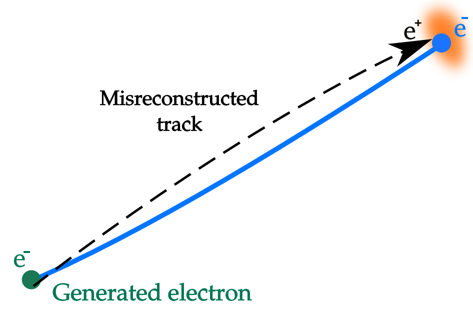


Figure 6.4: Sketch for high p_T track misreconstruction.

To verify the aforementioned sources, charge misidentification is studied in the ATLAS detector using MC and detector simulation in Section 6.3.

6.3 Investigations of charge misidentification sources in detector simulations

In this section, the sources of charge misidentification described in Section 6.2 are analyzed using MC processed through detector simulation. For these, *truth information* is available from the simulation to define the sources of charge misidentification. In Section 6.3.1 the accessed truth information is defined, with Section 6.3.2 showing the arising kinematic differences between correctly and wrongly reconstructed electron charges. In Section 6.3.3 the different observed sources of charge misID are studied at truth level using both their kinematic and tracking features.

6.3.1 Methodology and truth information assignment

In detector simulation, the truth information on the reconstructed and generated electron is retrieved. It helps to determine whether the reconstruction of the charge was successful. The electron is identified at generator level. In the final state, only the truth information on the simulated electron closest to the reconstructed best-matching track of the electron candidate is available. The electrons of interest are divided into *prompt* and *background* electrons at truth level. The process of origin is also identified, e.g. electrons from Z boson decays, photon conversions or FSR photons. The transverse momentum of a particle at truth level p_T^{Truth} is also interesting to study the resolution in energy reconstruction. The true charge of the reconstructed electrons is needed to identify charge misidentification in the detector simulation. This true charge is inferred from the electron at generation level, which is different from the reconstructed charge obtained from the track reconstruction in case charge misID takes place.

Charge misidentification is fundamentally caused by imperfections in the electron reconstruction algorithms. For MC, these come into play in detector simulation. Consequently, no considerable differences among either MC generators or physical processes are expected[1].

For the studies on charge misidentification sources presented in this section, events in $Z \to ee$ MC pass single-lepton trigger and matching. Electrons are selected with `TightLH` identification and `Gradient` isolation. Additionally, the overall acceptance of the inner detector, $|\eta| < 2.5$, is considered (excluding the crack region in the EMCAL, $1.37 < |\eta| < 1.52$). The electron energy is reconstructed using the signature in the electromagnetic calorimeter. Additionally, electrons must satisfy $p_T > 20$ GeV and the invariant mass of the two electrons is required to be $m_{\ell\ell} > 20$ GeV to remove contributions from soft-QCD processes.

Electrons passing these criteria are divided into four different categories according to the description given in Section 6.2. Firstly, prompt electrons are characterized by selecting Z boson decays as origin of the generated and reconstructed levels of the electron. Secondly, electrons arising from photon conversions are defined as a prompt electron produced in a Z boson decay at generator level, whose reconstructed electron in the detector simulation is found to be a background electron arising from a photon conversion. Additionally, each of these two categories is splitted into cases where charges at generator and reconstructed level agree or disagree, i.e. they involve correct or wrong electron charge identification, respectively. Technical details for the definition of these categories are given in Appendix B.3.

The described methodology is applied to compare kinematic distributions for charge misidentified and correctly reconstructed electrons. Different charge misID sources are further investigated in Section 6.3.3.

6.3.2 Kinematic properties of charge-misreconstructed electrons

The electron signature for correctly and wrongly reconstructed charges is studied in this section. To separate both contributions, the agreement between generated and reconstructed electric charges is tested for electrons passing the selection criteria. In Figure 6.5, the reconstructed electron p_T, the pseudorapidity $|\eta|$, the invariant mass of both electrons, the significance of the transverse impact parameter d_0/σ_{d_0}, the ratio between energy reconstructed in the EMCAL over energy measurement from the track curvature E/p and the number of tracks assigned to the electron candidate (*nTracksPerParticle*) are presented for the full Run 2.

Correctly reconstructed electrons (solid lines) show a strong tendency to be produced with transverse momenta measured in the EMCAL around 45 GeV, according to the kinematic constraints of the $Z \to ee$ decay. The charge is correctly reconstructed in 99.6% of the cases. Contrarily, charge misID constitute a rare phenomenon (0.3% of the electrons) and tend to be misidentified at greater values of p_T (dashed lines). Important differences are observed on the $|\eta|$ distribution for both electron types. Correctly reconstructed charges are mostly found in central regions of the detector. Electron charge misidenti-

[1] FSR radiation, whose final state mimics the signature produced by bremsstrahlung radiation, do depend on the used generator.

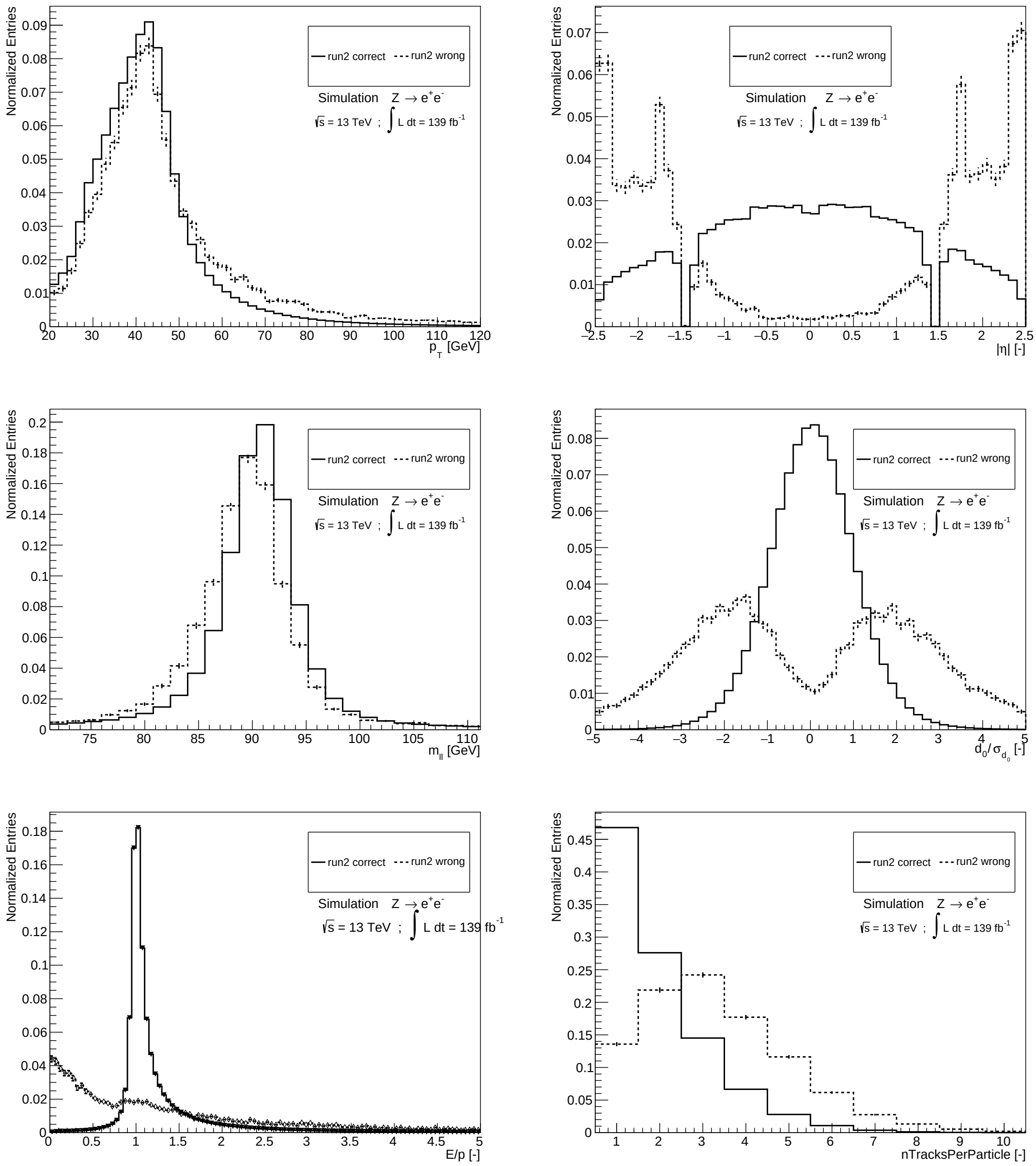


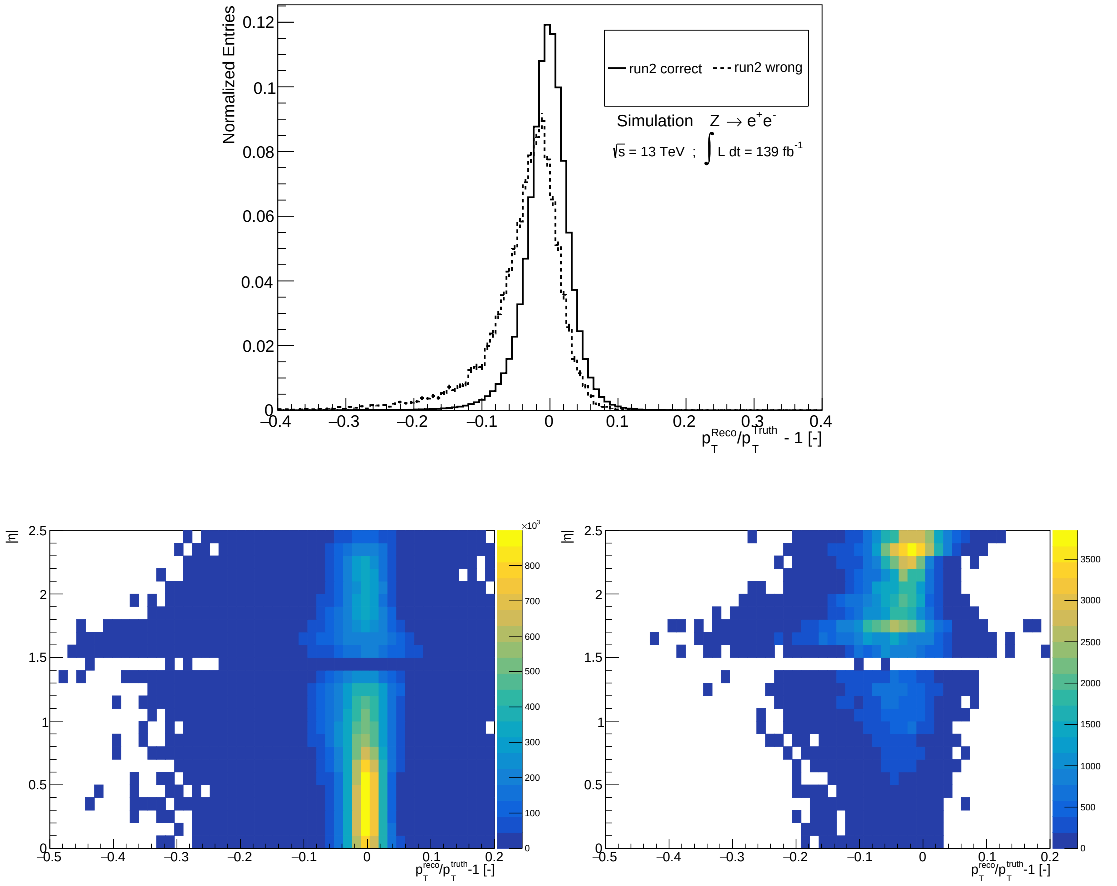
Figure 6.5: Kinematic and track distributions for electrons with correct (solid line) and wrong charge reconstruction (dashed line) in simulated $Z \to ee$ decays using the full Run 2 sample (139 fb^{-1}). On top, the reconstructed electron p_T (left) and the pseudorapidity $|\eta|$ (right). The central row contains the invariant mass reconstructed with both electrons $m_{\ell\ell}$ (left) and the significance of the transverse impact parameter d_0/σ_{d_0} (right). Below, the ratio between energy reconstructed in the EMCAL over energy measurement from the track curvature E/p (left) and the number of tracks assigned to the electron candidate (*nTracksPerParticle*, right). Only statistical uncertainties are shown. At least one charge -misidentified electron is considered for the $m_{\ell\ell}$ distribution of wrongly reconstructed electrons.

fication takes place at larger values of $|\eta|$, concretely for $|\eta| > 1.52$ and prominently in the region $2.3 < |\eta| < 2.5$. The dependency of charge misID with η is observed to be in agreement with the material distribution of the inner detector shown in Figure 3.5. An important fraction of charge-misID electrons is found in the most-forward regions of the inner detector ($\eta \gtrsim 2.3$). Electron charge misidentification is enhanced in this region where tracks difficultly bend when getting parallel to the direction of the magnetic field within the inner detector. The reconstruction of the Z boson invariant mass shows excellent performance when two correctly reconstructed electrons are involved[2]. When at least one charge-misidentified electron comes into the computation of $m_{\ell\ell}$ a worsened reconstruction of the Z boson mass is observed, with the distribution being shifted to lower values of energy. Energy leakage due to interactions with the detector material cause this smearing of the invariant mass reconstruction. Electrons with wrongly reconstructed charges show a track misalignment when observing the d_0 significance. The best-matched reconstructed track to these electron candidates do not cross the primary vertex, being the transverse impact parameter larger (in absolute module) than its corresponding uncertainty by a factor of 2 in most of the cases. This indicates that the reconstruction of the best matching track was induced by an electron that was not produced at the primary vertex. For electrons with correctly reconstructed charges, the extrapolation of the reconstructed track tends to be close the interaction vertex, providing a small value of the transverse parameter d_0 within its uncertainty σ_{d_0}. The ratio E/p turns out to be a very promising variable to distinguish between correctly and wrongly reconstructed electron charges. While the energy measurements from the calorimetry and tracking systems are in excellent agreement for correctly reconstructed electrons (by showing a prominent peak at $E/p = 1$), this agreement is completely lost when the charge is misidentified, tending to values of E/p close to zero. Since the energy reconstruction algorithm in the EMCAL performs well for correctly and wrongly reconstructed electrons as seen from the p_T and $m_{\ell\ell}$ distributions, the momentum inferred from the track is therefore overestimated when compared to the measurement from the electromagnetic calorimeter in most of the cases. While only one track is assigned to correctly reconstructed prompt electrons, additional tracks are normally assigned when the electric charge misidentified, showing the influence of subsequent production of electrons and photons on the charge misID.

Interactions with the detector material influence strongly the *energy resolution* (computed as $p_T^{\text{Reco}}/p_T^{\text{Truth}}-1$), as seen in Figure 6.6. While correctly reconstructed electrons show a sharp Gaussian-like distribution centered at zero (i.e. both good precision and accuracy at the same time), the energy resolution is smeared to negative values for wrongly reconstructed electrons in an asymmetric way. The tails do not correspond to Gaussian distributions after this smear. The detector reconstruction of transverse momentum using the electromagnetic calorimeter p_T^{Reco} underestimates the true transverse momentum of the electron, as it loses part of its energy along subsequent interactions with the detector material. The strong dependence of the energy resolution η clarify this fact. Especially, the energy resolution notably worse with increasing $|\eta|$, as the detector material the particle traverses becomes denser.

In conclusion, electrons with a wrong charge identification are normally found in regions of the inner detector with high pseudorapidity ($|\eta| > 1.52$) and importantly

[2]Improvements after using the dynamical, topological cell clustering-based approach for the energy reconstruction are compared with respect to the slicing-window approach for different kinematic distributions. The results are presented in Appendix B.1.



(a) Correctly reconstructed electron charges.

(b) Wrongly reconstructed electron charges.

Figure 6.6: On top, the energy resolution for correctly (solid) and wrongly reconstructed electrons (dashed) is shown in simulated $Z \to ee$ decays using the full Run 2 sample (139 fb^{-1}). Only statistical uncertainties are shown. Below, the dependence of the energy resolution for correctly (a) and wrongly (b) reconstructed electrons with $|\eta|$ is studied in differential, two-dimensional distributions with the 36 fb^{-1} sample. The slicing-window approach is used for the energy reconstruction for the results in the 36 fb^{-1} sample in the release 20.7 of the ATLAS software.

for the most-forward regions where the trajectory is poorly bent by the magnetic field ($2.3 < |\eta| < 2.5$). The misidentification tends to occur at slightly greater transverse momentum when compared to satisfactory charge reconstruction. When at least one misreconstructed electron charge takes place in the computation of the invariant mass of leading and subleading lepton in $Z \to ee$ events, the Z boson mass is underestimated due to energy leakage after the electron interacts with the detector material. The best-matching track for electrons with a misidentified charge shows a poor extrapolation in the vicinity of the primary vertex, as can be seen from its d_0 significance. This shows that charge misidentification is related to inefficiencies in the reconstruction of the electron candidate track. In Section 6.3.3, the sources for this component are identified and studied independently.

Charge identification	Electron category	Portion in sample [%]	
Correct charge	Prompt	99.0	99.6
	γ conversion	0.6	
Wrong charge	Charge misID γ conversion	0.2	0.3
	Charge misID prompt	0.11	

Table 6.2: Contributions to charge misidentificatio nfrom different electron categories characterized at truth level in $Z \to ee$ decays using the 36 fb^{-1} dataset.

6.3.3 Kinematic properties of charge-misidentified electrons for different sources

In the previous section, the main differences between charge-misidentified and correctly reconstructed electrons were outlined, along with a description of their sources. Such a sources are characterized at truth level and their kinematic and track properties are studied separately[3].

The electron categories to be considered are selected as described in Section 6.3.1. Table 6.2 summarizes the identified electron types. Prompt electrons with correctly identified charges are found in 99.0% of the cases. The electric charge for this electron category is misidentified in 0.1% of the studied electrons. 0.8% of the electrons are found to be reconstructed after bremsstrahlung, where the generated electron charge is misidentified 25% of the cases (constituting 0.2% of the electrons in the sample). FSR electrons compose 0.04% of the total electron sample. The main source of charge misidentification is found to arise from bremsstrahlung (64.5% of charge misidentified electrons), followed by misidentified prompt charges (35.5%).

Figure 6.7 shows the correlation between different electron categories per event observed for the first and second lepton with the highest p_T of the event (denoted by leading lepton ℓ_0 and subleading lepton ℓ_1, respectively). For 97.8% of the events in the studied sample, two prompt electrons from the Z decay are correctly reconstructed. In 1.3% of the events, two correctly identified electron charges are involved, one of them from a prompt electron and the other arising from a photon conversion. Events involving charge misidentification come in the third place. 0.4% of the Z decays are reconstructed with same-sign electron charges in the final state, when one of the electrons is produced after photon conversions and the other being a correctly reconstructed prompt electron. Lastly, 0.2% of the events include two prompt electrons, with one of their charges being wrongly reconstructed. Events including FSR electrons correspond to 0.08% of the cases. The two electron charges from the Z decay can also provide opposite-sign leptons in the final state when both charges are misidentified. This situation is found in 0.0015% of the cases. The electric charge tends to be correctly reconstructed for the leading and subleading leptons of the event identically when photon conversions are involved. Correlations between FSR and prompt electrons show the same balance. However, the charge misidentified prompt electron turns out to be the leading lepton in approximately 60% of the events. On the studied $Z \to ee$ decays, same-sign electrons in the final state are observed in 0.6% of the events.

[3]Results presented in this section use an older ATLAS software release were the dynamic, topological clustering presented in Section 3.3.2. Comparisons between the two software releases are presented in Appendix B.1. Although the energy reconstruction is found to improve, the conclusions on this Section are not affected by this new algorithm.

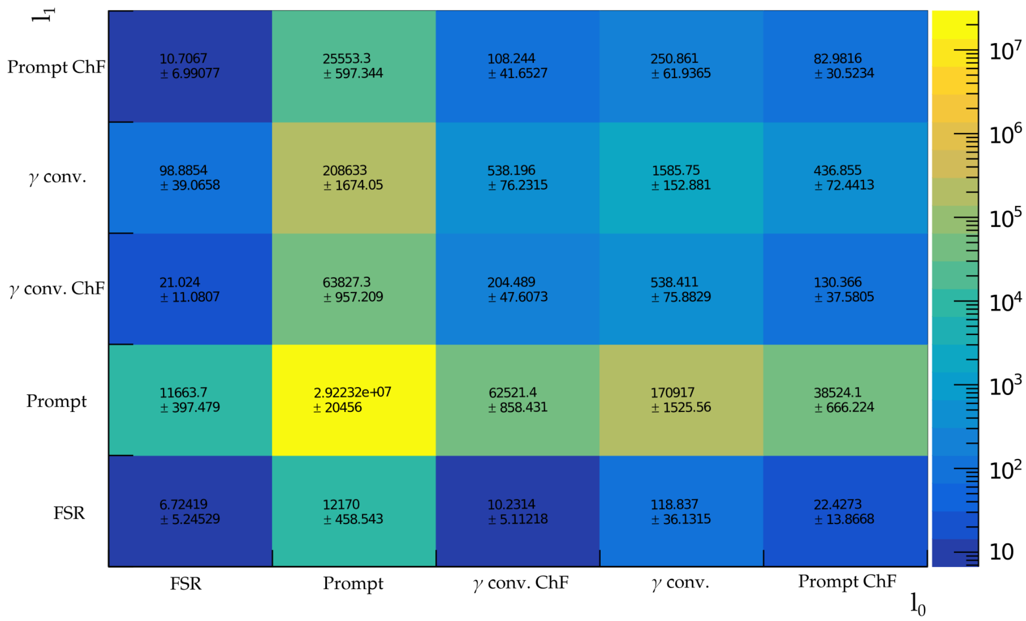


Figure 6.7: Correlation between electron types for leading (horizontal axis) and subleading electrons (vertical axis) involved per event in truth studies of charge misidentification in $Z \to ee$ decays using the 36 fb^{-1} dataset. These results are produced using the release 20.7 of the ATLAS software.

The kinematic properties of the different electron categories can be seen in Figure 6.8 for the different electron types, where the distributions for transverse momentum and pseudorapidity are shown. Correctly identified prompt electrons (blue) follow identical distributions as in Figure 6.5 for correctly reconstructed electrons, as they constitute 99% of the studied electrons. A prominent tendency to be produced with transverse momentum around 45 GeV is also observed for correctly reconstructed electrons arising from photon conversion, with a smear to regions of greater p_T. Electrons arising from photon conversions with correctly identified charges are found within the region $1.52 < |\eta| < 2.2$, following the material distribution of the inner detector. This tendency is further shifted to higher values of η for charge-misidentified electrons produced after photon conversions. Nevertheless, these tend to produce greater values of p_T in the EMCAL when compared to correctly reconstructed electrons arising from bremsstrahlung even though they are also reconstructed after experiencing the same phenomenon. A more prominent tendency to higher values of p_T is found for wrongly reconstructed prompt electrons. These are mostly found within very forward regions of the inner detector ($2.3 < |\eta| < 2.5$), where tracks for electrons with high-p_T are barely bent and no TRT hits can be found.

The effects of interactions between the electrons and the detector material can be also observed when reconstructing the invariant mass of leading and subleading lepton in Figure 6.9. As previously observed in Figure 6.5, an excellent reconstruction of the Z mass can be achieved when two prompt electrons are reconstructed in $Z \to ee$ decays. This reconstruction is slightly worsened when the charge from one of the prompt electrons is wrongly identified. Even though the prompt electron was completely identified, the observed tendency for these to be produced at $|\eta| > 2.3$ involves that they interacted

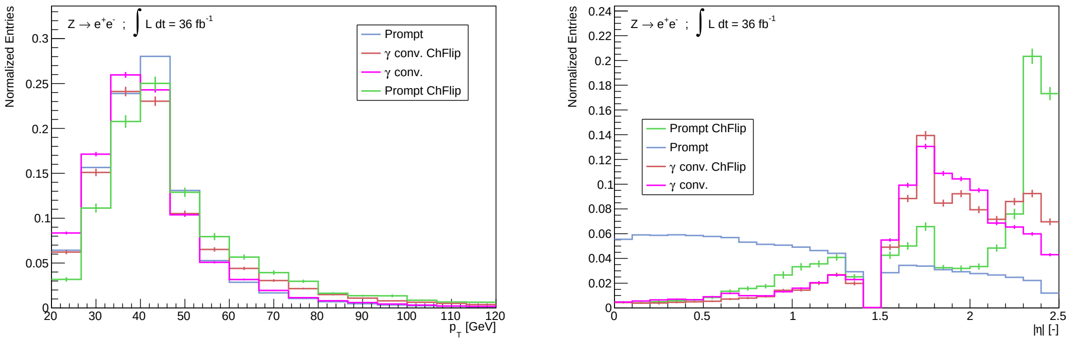


Figure 6.8: Transverse momentum p_T reconstructed by the electromagnetic calorimeter (left) and pseudorapidity $|\eta|$ (right) distributions for the different electron types identified at truth level for studies of charge misidentification sources in simulated $Z \to ee$ decays using the 36 fb^{-1} dataset. These results are produced using the release 20.7 of the ATLAS software.

with an imporant amount of inner detector material, lessening the energy measurement. However, the smear of $m_{\ell\ell}$ to lower energies observed in Figure 6.5 is reproduced well when at least one reconstructed electron arising from photon conversions is involved, either their charge is correctly (pink curve) or wrongly identified (red distribution). Additionally, charge misidentification from electrons arising after photon conversions constitutes, as presented in Table 6.2, about 75% of wrongly reconstructed charges, while the proportion of these electron type with correct charge reconstruction compared to correctly reconstructed prompt electrons is negligible. The smearing of the $m_{\ell\ell}$ distribution is therefore dominated by electrons induced by bremsstrahlung and, consequently, by energy leakage after interactions with the inner detector material.

The information on the best-matched track for the electron candidates is also accessed. In Figure 6.10 the d_0 significance, the number of assigned tracks to the electron candidate, the ratio E/p, the transverse momentum provided by the track p_T^{Track}, the distribution of charged inverse momentum Q/p inferred from the track, the total number of hits in the TRT component of the inner detector, the total number of hits of the track and the $\chi^2/d.o.f$ of the track fit are shown. Prompt electrons (with either correctly and wrongly reconstructed charges) provide tracks whose extrapolation crosses the interaction vertex within uncertainties, as seen from their distributions of d_0 significance. In contrast, electrons involving charge misID (either prompt or after photon conversions) present a mismatch when extrapolating the track to the primary vertex. The tendency of this ratio for charge-misidentified prompt electrons is shifted to lower values with a more irregular shape. But just one track is assigned to correctly reconstructed prompt electrons in approximately 75% of the cases. Additional tracks are likely to be assigned when the charge of prompt electrons is misidentified, reducing the proportion of electrons with only one track to about 35%. At the same time, electron candidates arising from photon conversions are prone to have three tracks assigned in approximately 30% of the cases. Around 85% of the charge-misID electrons present more than one track, while 60% have three or more. This is caused by multiple tracks assigments related to the prompt and subsequent photon conversions (see the sketch in Figure 6.2). As shown in Table 6.2, about 65% of the cases the track with opposite sign results as the best-matching track for the electron candidate. Prompt electrons with correctly identified charge show a similar distribution

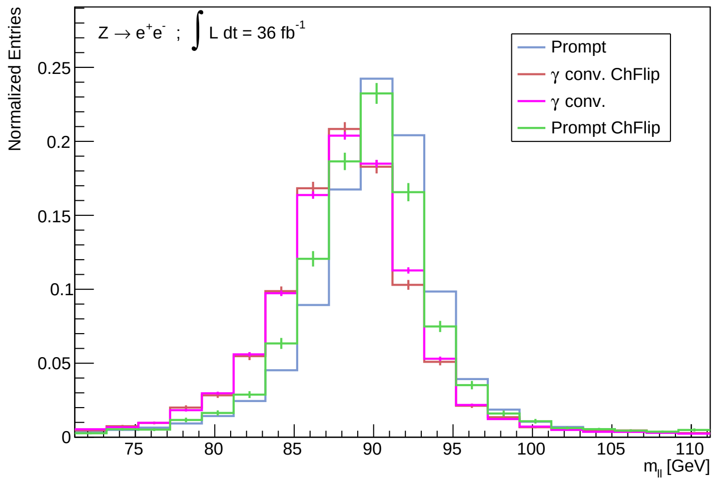


Figure 6.9: Influence from different electron types identified in truth studies for charge misidentification on the invariant mass $m_{\ell\ell}$ of both electrons in Z boson decays using the 36 fb^{-1} sample. At least one electron from photon conversions or charge-misidentified prompt electron is considered for the $m_{\ell\ell}$ distribution for these electron categories. For correctly reconstructed prompt electrons, two electrons of this category are required for the reconstruction of $m_{\ell\ell}$. The slicing-window approach is used for the energy reconstruction for the results in the 36 fb^{-1} sample in the release 20.7 of the ATLAS software.

for p_T^{Track} compared to the transverse momentum reconstructed in the EMCAL (presented in Figure 6.8). This is expected since both measurements agree as previously seen in Figure 6.5 for the ratio E/p. The prominent maximum around 45 GeV is lost for electrons when photon conversions come into play, whose tracks are prone to provide lower values of transverse momentum. Indeed, the radius of curvature is steadily reduced along the path through the inner detector material due to the reduction of the prompt lepton energy. Prompt electrons with wrongly reconstructed charges provide, in comparison to the measurement from the EMCAL, a completely overestimated value p_T^{Track}. The ratio E/p shows more clearly this disagreement between energy measurements from the calorimetry and tracking systems. The agreement is fantastic for prompt electrons with correctly identified charges, visible by a sharp peak at $E/p = 1$. Albeit, this agreement is lost when their charge is wrongly reconstructed, tending to provide ratios E/p around zero. This tendency was already observed in Figure 6.5, and it can be now understood better when comparing the distributions of p_T^{Track} and p_T for the different electron types. Uncorrelated measurements of the electron energy from the track and EMCAL are observed. For charge-misID electrons induced by photon conversions, a tendency to low values of E/p is found, with correctly charge-identified electrons after photon conversions showing a slight tendency to greater values. Mismeasurements of the track momentum can be better understood through the distributions of Q/p^{Track}, inversely proportional to the measured radius of curvature. Correctly charge-identified prompt electrons show preferred Q/p^{Track} values around 0.025 GeV^{-1}, in agreement with the kinematic constraints of $Z \to ee$ decays. Since p^{Track} is overestimated for other electron types, their distributions are shifted towards values around zero. For electrons arising from photon conversions, the already observed underestimated radius of curvature (from the p_T^{Track} distributions when compared to p_T) is in agreement with the lower values of $|Q/p^{\text{Track}}|$. The number of

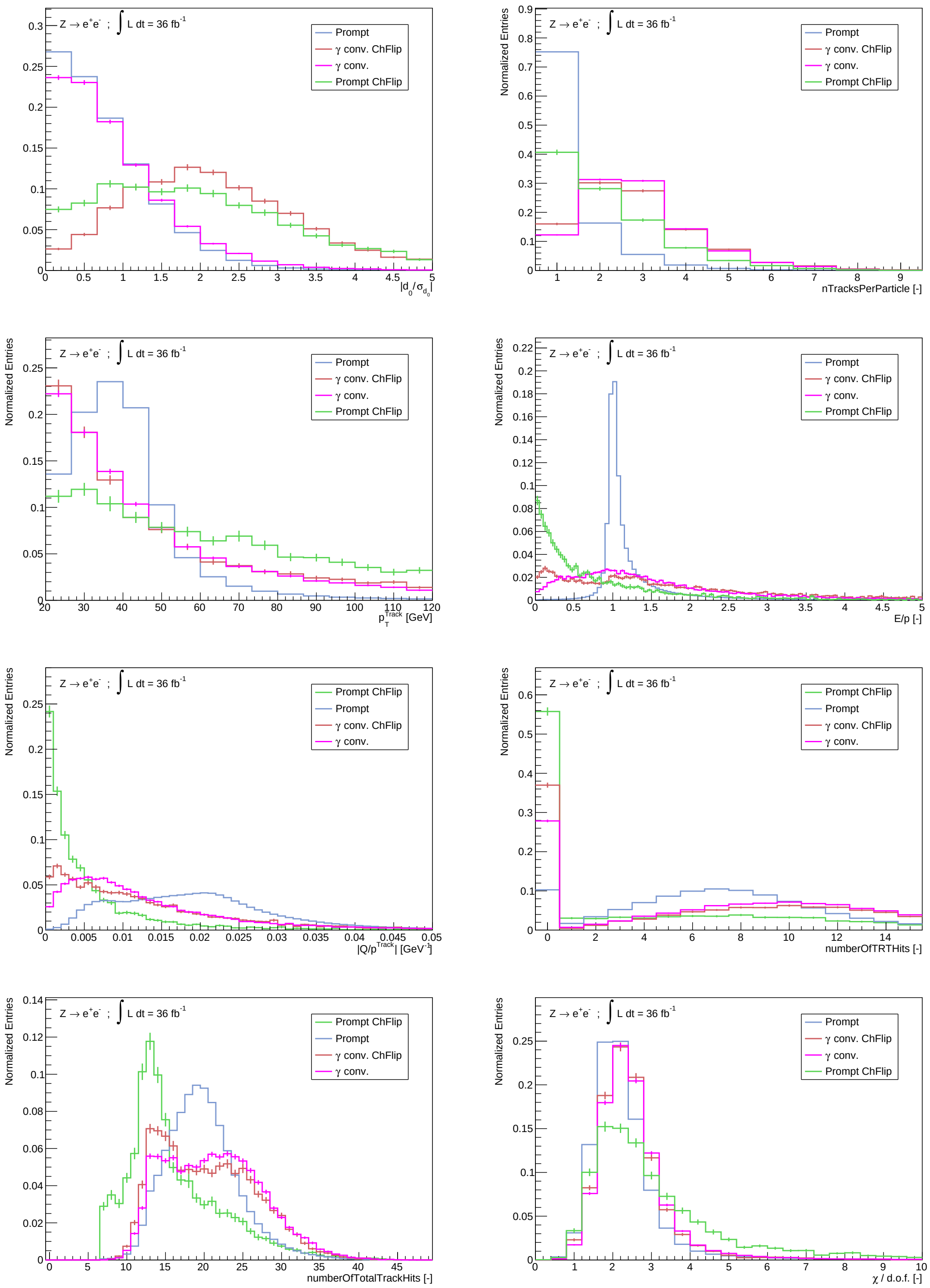


Figure 6.10: Tracking properties for the different identified electron types in truth studies of charge misidentification sources in simulated $Z \to ee$ decays for 36 fb^{-1}. The d_0 significance and the number of assigned tracks to the electron candidate (denoted by *nTracksPerParticle*) are presented on the upper row. Below, the ratio between the electron energy reconstructed in the electromagnetic calorimeter over the track momentum E/p and the transverse momentum provided by the track (p_T^{Track}) are shown. The third row includes the distribution of charged inverse momentum $|Q/p^{\mathrm{Track}}|$, inferred from the track and the total number of TRT hits. The total number of track hits (calculated as the sum of hits in the IBL, the pixel detector, the SCT and the TRT layers hits) and the $\chi^2/d.o.f$ of the track fit (where the number of degrees of freedom is then computed as the total number of track hits) are presented at the bottom row. Only statistical uncertainties are shown. These results are produced using the release 20.7 of the ATLAS software.

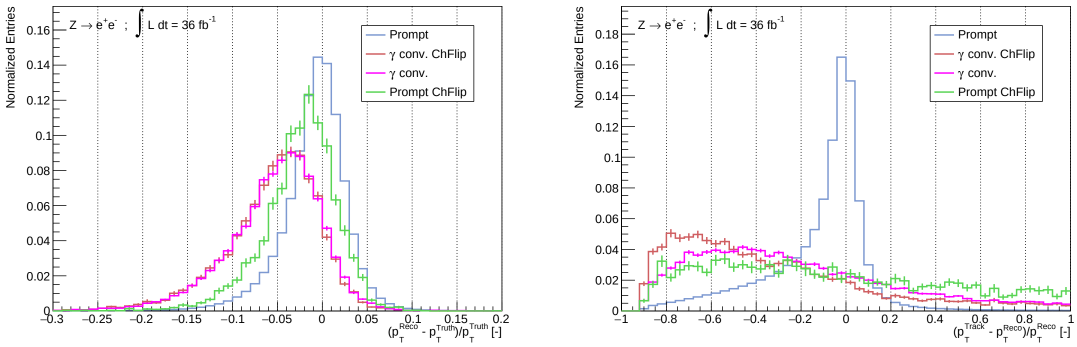


Figure 6.11: Energy resolution $(p_T^{\text{Reco}} - p_T^{\text{Truth}})/p_T^{\text{Truth}}$ and relative difference between the reconstructed transverse momentum measured in the electromagnetic calorimeter p_T^{Reco} and the transverse momentum measured with the reconstructed track in the inner detector p_T^{Track} (computed as $(p_T^{\text{Track}} - p_T^{\text{Reco}})/p_T^{\text{Reco}}$) for different electron types at truth level using $Z \to ee$ decays in 36 fb^{-1}. Only statistical uncertainties are shown. These results are produced using the release 20.7 of the ATLAS software.

TRT hits is considerably reduced for the case of charge misidentified prompt electrons, showing no hits in this detector in 70% of the cases due to their tendency to be produced at very high $|\eta|$. Only 10% of prompt electrons with correctly reconstructed charge do not produce hits in the TRT layers of the ID, as they are mostly observed in the central regions of the detector. This feature is also observed in the total amount of hits, computed as the sum of pixel (that includes the IBL), SCT and TRT hits. In the pixel and SCT layers, 4 and 8 hists are normally found for all the studied electron categories. Both TRT and the total amount of hits are not notably altered for correctly reconstructed prompt electrons (with 19 hits in most of the best-matched tracks, 7 of them from the TRT) and reconstructed electrons arising from photon conversions. The $\chi^2/d.o.f.$ of the track, providing information on the quality of the fit that builds the track, is considerably worsened for charge-misidentified prompt electrons. The mentioned reduction of the degrees of freedom influences this lessen. In comparison to correctly reconstructed prompt electrons, photon conversion electrons show a slightly worse track fit.

The *energy resolution*, computed as the relative difference between the transverse momentum reconstructed in the EMCAL (p_T^{Reco}) and p_T^{Truth}, provides important information on the energy reconstruction performance of the detector. In Figure 6.11, the energy resolution and relative difference between p_T^{Reco} and transverse momentum inferred from the reconstructed track in the inner detector p_T^{Track} are presented. Correctly reconstructed prompt electrons show an excellent energy resolution in both precision and accuracy. The correlation between p_T^{Reco} and p_T^{Track} is worsened in comparison, with a clear tendency for the measurement p_T^{Track} to underestimate the particle transverse momentum. These electrons compose the tail $E/p > 1$ in Figure 6.8. Electrons arising from photon conversions provide a worse energy resolution, with the center of the distribution being shifted to negative values of $(p_T^{\text{Reco}} - p_T^{\text{Truth}})/p_T^{\text{Truth}}$ and with its width being increased, i.e. both accuracy and precision are worsened, respectively. For these electrons, p_T^{Reco} is on average lower than p_T^{Truth} due to energy leakage after interactions with the detector material. In most of the electrons arising from photon conversions, p_T^{Track} is underestimated when compared to the reconstructed p_T^{Reco}. The agreement between p_T^{Track} and p_T^{Reco} per electron candidate is lost for these electron types, as notable from the flattened distribution. For charge misidentified prompt electrons, the energy resolution improves with respect to photon conversion electrons, but its precision is worsened when compared to correctly reconstructed prompt electrons. The distribution $(p_T^{\text{Track}} - p_T^{\text{Reco}})/p_T^{\text{Reco}}$ is flat, showing complete uncorrelated values of both measurements of transverse momentum taken from calorimetry and tracking systems.

In conclusion, charge misidentification takes place in the ATLAS detector for a reduced amount of the reconstructed electron candidates (0.3%), arising from two main sources. Around 64.5% of the electrons with wrongly reconstructed charge arise after bremsstrahlung radiation when these interact with the detector material. Since the material budget to be traversed by the electron increases in more forward regions of the inner detector, the electron candidates induced by photon conversions are found to be produced within this region. Energy leakage after these interactions also worsen the energy resolution for the measurement of the particle energy in the electromagnetic calorimeter with increasing $|\eta|$, smearing the reconstruction of the invariant mass to lower values of energy. For electrons produced after photon conversions two or three tracks are assigned on average, with the best-matching track providing a poor reconstruction of the particle momentum and inconsistent extrapolation to the primary vertex. Bremsstrahlung radiation does not always involves charge misidentification. 75% of these electron candidates arising from bremsstrahlung were found to provide the same charge as the generated electron. Electrons induced by Final State Radiation (FSR) radiation provide a similar signature as electrons arising from photon conversions. This phenomenon is found in approximately 0.04% of the studied events. Since this process is considered at generator level, no detector effects influence the reconstruction of the track. Therefore, FSR-induced electrons are not considered as a source of charge misidentification. The second source of charge misidentification is found to arise from straight-like tracks when the tracking algorithm has difficulties to properly measure the track curvature of the prompt electron trajectory. One or two tracks are assigned to these particles approximately 40% and 28% of the cases. The extrapolation of the best-track candidate (that is used to retrieve the electron kinematics) is far from the primary vertex and shows a poorer $\chi^2/d.o.f.$ for the fit. The measurement of the electron energy from this track is underestimated and it is uncorrelated to the energy reconstructed in the electromagnetic calorimeter. Prompt electrons with wrongly reconstructed charges are found within the very forward region of the inner detector ($2.3 < |\eta| < 2.5$) where the track is poorly bent by the magnetic field. Contrarily to the charge misidentification after photon conversions (where the wrongly reconstructed electron is found in leading and subleading leptons with similar probability), this source is found in the leading lepton in approximately 60% of the events where this source of charge misID is found. The properties of charge misidentified electrons observed in this chapter are considered for the estimation of the charge misidentification background presented in Chapter 7. The estimated background yield is used for the measurement of same-sign W boson pairs with two jets in the final state using the full Run 2 dataset recorded with the ATLAS experiment.

Chapter 7

Electron charge misidentification background estimation for the ssWW analysis

Electron charge misidentification becomes an important background when measuring decays of same-sign W boson pairs into electrons especially when physical processes producing opposite-sign electron pairs in the final state (whose production ratio is orders of magnitude larger than the physical process of interest) pass the selection criteria after one of their charges is misidentified due to detector inefficiencies. The sources for charge misidentification were studied in Chapter 6. The main purpose of this thesis is to estimate the influence of this phenomenon for the measurement of same-sign W boson pair production with the ATLAS experiment as a background contribution, using the full Run 2 dataset (139 fb^{-1}) recorded in pp collisions at the center-of-mass energy of $\sqrt{s}$ = 13 TeV. The methodology is presented in Section 7.1 with new modifications. Results of the validation of the method and the estimated background yield in the signal region are presented in Section 7.2. The background is reduced by applying the ECIDS tool, introduced in Section 3.3.2. Its implementation and resulting background yields are described in Section 7.3.

7.1 Methodology

To estimate the charge misidentification background yields, a data-driven method is used. For this purpose, opposite-sign data events are weighted for the estimation of event yields with same-sign leptons in the final state. The weights take into account the charge misidentification probabilities as a function of p_T and $|\eta|$, which are determined separately. Their computation is presented in Section 7.1.1. The weights are then computed as the probability for a same-sign dilepton event to arise from opposite-sign leptons due to charge misidentification, according to the kinematics of leading and subleading leptons, as presented in Section 7.1.2. As the mismeasurement of the electron energies for wrongly identified electrons is not involved in the considered opposite-sign data events, further energy corrections need to be carried out to mimic the energy leakage. The applied energy corrections are described in Section 7.1.3.

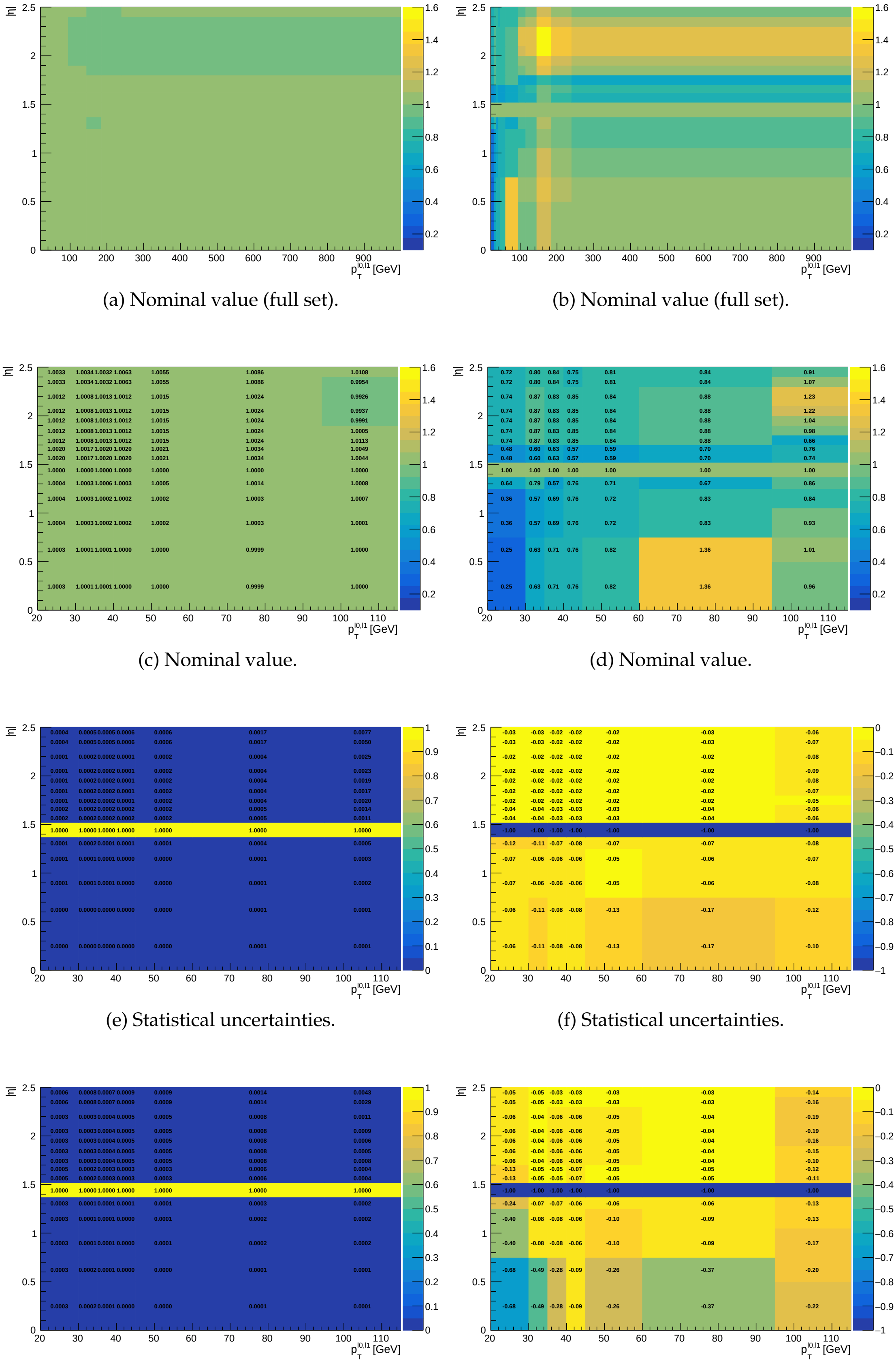
7.1.1 Measurement of charge misidentification probabilities

To compute the charge misidentification probabilities using truth information, the $Z \to e^+e^-$ MC samples after detector simulation are used. The reconstructed Z boson decays with same-sign electrons in the final state will involve charge misidentification, a signature that can be exploited for the computation of the charge misidentification probabilities in data and MC through a *side-band substraction* at the Z mass resonance [68]. Since effects of detector interactions on charge misidentification are not described precisely enough in the detector simulation, deviations from the charge misID probability distributions between data ($\varepsilon_{\text{data}}$) and MC ($\varepsilon_{\text{MC}}$) are observed. To correct for such a deviations, *scale factors* (SF) are applied. The SF are calculated for wrongly and correctly reconstructed electrons as

$$\text{SF}_{\text{wrong}} = \frac{\varepsilon_{\text{data}}}{\varepsilon_{\text{MC}}}, \qquad \text{SF}_{\text{correct}} = \frac{1-\varepsilon_{\text{data}}}{1-\varepsilon_{\text{MC}}}. \tag{7.1}$$

The distributions of $\text{SF}_{\text{correct}}$ and SF_{wrong} are calculated and delivered to the ATLAS collaboration by the *Egamma Combined Performance (CP) group* for different electron identification and isolation working points along with statistical and systematic uncertainties. Details on the computation of the used SF and their sources of uncertainties can be found in Refs. [63, 68]. The provided distributions for `TightLH` identification and `Gradient` isolation working points, statistical and systematic uncertainties are shown in Figure 7.1. Scale factors are mostly close to unity for correctly reconstructed electrons along the entire described region in p_T and $|\eta|$, since charge misidentification probabilities are small (of the order of 0.3% according to results in Chapter 6). Low statistical and systematic uncertainties (up to about 0.1% for bins with largest systematic uncertainties) come along with distributions of $\text{SF}_{\text{correct}}$. The crack region in the electromagnetic calorimeter is not taken into account, with the SF of $1.0 \pm 1.0\,(\text{stat}) \pm 1.0\,(\text{sys})$ within $1.37 < |\eta| < 1.52$. The distributions of SF_{wrong} show more interesting values. Especially in the most populated region of transverse momentum (about $20 < p_T < 60$ GeV for Z boson decays, as observed in Section 6.3.2), scale factors for wrongly reconstructed electrons show values lower than one. This means that, since the charge misidentification rates computed in such a cases are overestimated in MC as provided by Eq (7.1), a data-driven method applying ϵ_{MC} to data would also overestimate the outcoming background yields. Different tendencies are found for $p_T > 95$ GeV, where scale factors take values greater than one. This region is particularly affected by the larger statistical and systematic uncertainties, reaching the order of 20%. According to the kinematics and correct or wrong identification of the charges of leading and subleading leptons (by considering the truth information), values of $\text{SF}_{\ell 0}$ and $\text{SF}_{\ell 1}$ are retrieved. These are then applied to MC samples by event weights $w_{\text{SF}} = \text{SF}_{\ell 0} \times \text{SF}_{\ell 1}$.

To calculate the charge misidentification probabilities, electrons are selected similarly as in Section 6.2. *Madgraph* $Z \to ee$ MC samples passing single-lepton trigger and matching are used. Electron identification and isolation are required to pass `TightLH` and `Gradient` criteria, respectively. Excluding the crack region of the electromagnetic calorimeter ($1.37 < |\eta| < 1.52$), the almost complete acceptance of the inner detector, $|\eta| < 2.5$, is considered. The electron energy is required to be reconstructed using the



(a) Nominal value (full set).

(b) Nominal value (full set).

(c) Nominal value.

(d) Nominal value.

(e) Statistical uncertainties.

(f) Statistical uncertainties.

(g) Systematic uncertainties.

(h) Systematic uncertainties.

Figure 7.1: Scale factors for correctly (left) and wrongly (right) reconstructed electrons using `TightLH` identification and `Gradient` isolation. The full set of SF (upper row), values within $20 < p_T < 115$ GeV (second row) statistical uncertainties (third row) and systematic uncertainties (bottom row) are also shown for the full range in $|\eta|$ and $20 < p_T < 115$ GeV. [68].

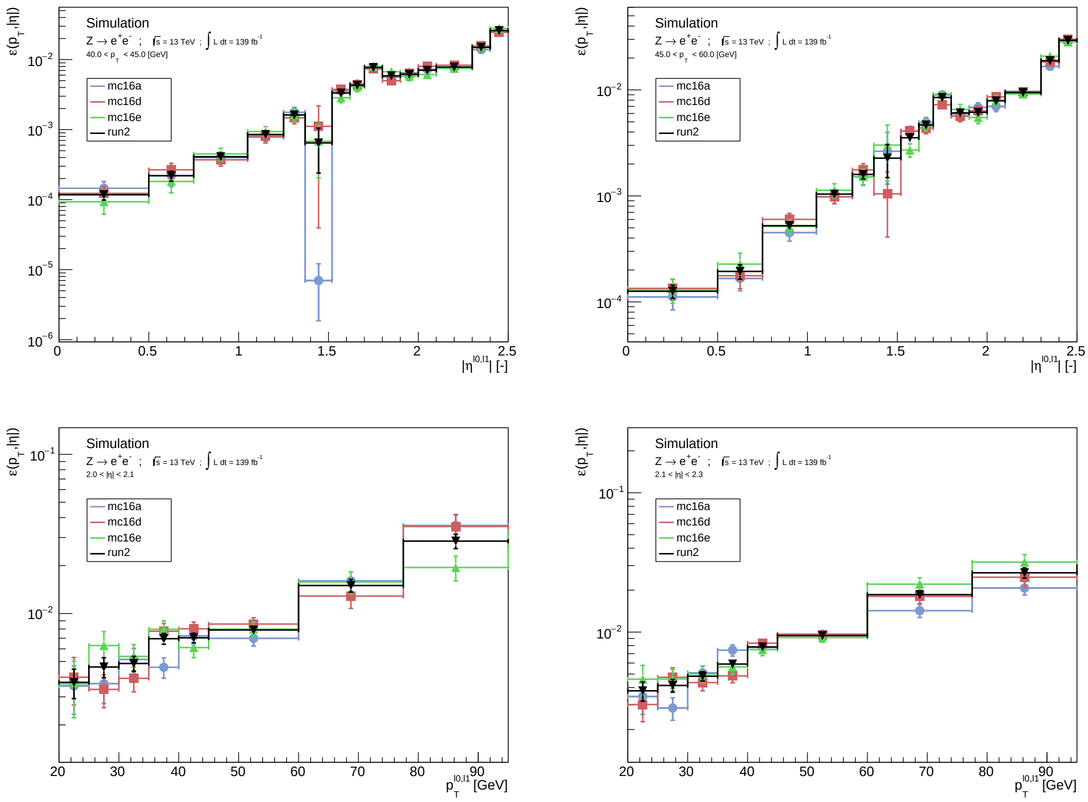


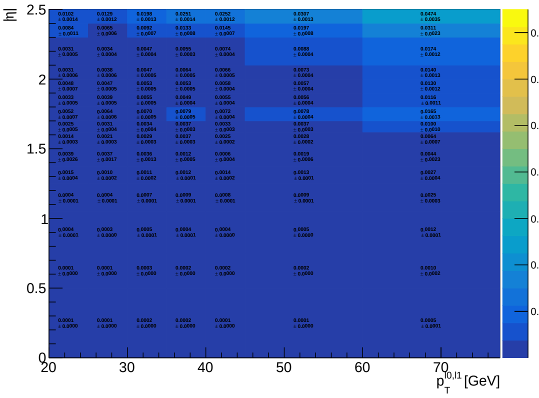
Figure 7.2: Comparison of charge misidentification probabilities for mc16a (blue), mc16d (red) and mc16e (green) MC campaigns along with distributions for the full run-2 set (black) in the most populated bins in p_T (top) and $|\eta|$ (bottom). The distributions were computed by exploiting the truth information in $Z \rightarrow ee$ Madgraph MC samples. Reconstructed electrons originated from FSR were removed for the computation and the scale factors shown in Figure 7.1 were applied. Only statistical uncertainties are shown.

signature in the electromagnetic calorimeter. The invariant mass of the two electrons must satisfy $m_{\ell\ell} > 20$ GeV to remove contributions from soft-QCD processes. Only electrons with transverse momentum $p_T > 20$ GeV are considered since the scale factors are only supported within this range. To avoid the influence in charge misidentification from FSR-induced electrons and other background processes, their contributions are removed at truth level. Electron candidates fulfilling this selection criteria are used to calculate the charge misidentification probabilities differentially in $|\eta|$[1] and p_T as

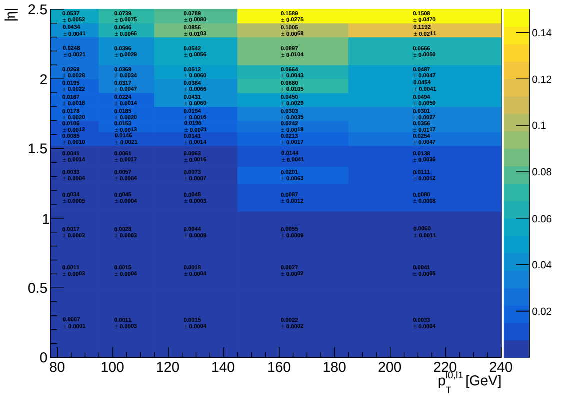
$$\varepsilon_{\text{misID},ij} = \frac{N_{\text{misID},ij}}{N_{\text{all},ij}}\,, \tag{7.2}$$

i.e. the rate between the number of electrons with wrongly identified charge N_{misID} (characterized at truth level within a certain bin ij of the $p_T - |\eta|$ distribution) over the total number of electrons N_{all} (wrongly plus correctly reconstructed charges, within the same bin ij).

[1] As presented in Figure 6.5, the η distribution for charge-misidentified electrons is symmetric along this variable, with these electrons found within forward regions of the inner detector (high $|\eta|$). Therefore, using the dependence on $|\eta|$ reduces statistical uncertainties for the computation of charge misidentification probabilities.



(a) $20 < p_T < 77.5$ GeV.



(b) $77.5 < p_T < 240$ GeV.

Figure 7.3: Charge misidentification probabilities with statistical uncertainties computed with electron candidates for bins within $20 < p_T < 77.5$ GeV (left) and 77.5 GeV $< p_T < 240$ GeV (right). The scale factors shown in Figure 7.1 are applied. Electrons for `TightLH` identification and `Gradient` isolation are chosen.

The charge misidentification probabilities are computed separately for different MC campaigns and the full Run 2 dataset. One-dimensional slices of the two-dimensional distribution are shown in 7.2 for the regions with largest electron population. The full set of distributions can be found in Section C.5. The charge misidentification probabilities agree within statistical uncertainties across MC campaigns. Therefore, only distributions for full Run 2 are considered for the background estimation to reduce statistical fluctuations. The obtained two-dimensional distribution is presented in Figure 7.3. The charge misidentification probabilities increase with increasing reconstructed electron p_T and $|\eta|$. The binning chosen for the charge misidentification probabilities is based on the binning of the scale factors. The binning in p_T is divided in 13 bins as $\{20.0, 25.0, 30.0, 35.0, 40.0, 45.0, 60.0, 77.5, 95.0, 115.0, 145.0, 185.0, 240.0, 999.0\}$ while 15 bins $\{0.0, 0.5, 0.75, 1.05, 1.25, 1.37, 1.52, 1.62, 1.7, 1.8, 1.9, 2.0, 2.1, 2.3, 2.4, 2.5\}$ are set for the binning in $|\eta|$. Probabilities take values from 0.003% up to approximately 15%. Along the $|\eta|$ axis, these probabilities can increase by two orders of magnitude within a certain region of p_T. When a $|\eta|$ region is chosen, the increase along p_T is found to be weaker. These dependences are in agreement with the distributions presented in Section 6.3 for the kinematics of correctly and wrongly charge-identified electrons. Integrating the distributions $N_{\text{misID},ij}$ and $N_{\text{all},ij}$, the integrated charge misidentification probability along p_T and η is 0.3%, in agreement with the results in Table 6.2.

The charge misidentification probabilities play a crucial role for the background estimation, since they include the knowledge on the detector material influencing such a phenomenon in terms of the electron kinematics. They come into play through event weights to be applied on a dataset containing two charged leptons of opposite sign in the final state. In this way, the phenomenon of charge misidentification producing same-sign leptons in the final state is accounted for the estimation of the background yield.

7.1.2 Weights for the charge misidentification background estimation

For the data-driven background estimation, events with opposite-sign leptons in the final state (OS events) are used to compute the background yields for same-sign leptons in the final state (SS events) involving wrong charge reconstruction. Since charge-misidentified electrons constitute a small portion in the dataset (approximately 0.3% according to integrated distributions in Figure 7.3), the choice of weighting OS events reduces remarkably the corresponding statistical uncertainties for the background yield. In this way, given the two-dimensional distribution of charge misidentification probabilities $\varepsilon_{\text{misID}}$, the weights to apply on events with opposite-sign leptons can be computed as

$$w = \frac{\varepsilon_{\ell 0}(1-\varepsilon_{\ell 1}) + \varepsilon_{\ell 1}(1-\varepsilon_{\ell 0})}{(1-\varepsilon_{\ell 0})(1-\varepsilon_{\ell 1}) + \varepsilon_{\ell 0}\varepsilon_{\ell 1}} . \tag{7.3}$$

The numerator indicates the probability for a SS event to arise from an OS event. Considering kinematic variables of leading (ℓ_0) and subleading (ℓ_1) electrons, the numerator $\varepsilon_{\ell 0}(1-\varepsilon_{\ell 1}) + \varepsilon_{\ell 1}(1-\varepsilon_{\ell 0})$ contains the probability for leading (subleading) lepton to flip the charge while the charge of the subleading (leading) lepton is correctly reconstructed in the first (second) term. In the denominator, the probability for an OS event to remain as opposite-sign is expressed by the probability of reconstructing both charges satisfactorily ($(1-\varepsilon_{\ell 0})(1-\varepsilon_{\ell 1})$) plus the probability of misidentifying both at the same time ($\varepsilon_{\ell 0}\varepsilon_{\ell 1}$). Since misidentification of muon electric charges has been proven to be small in the ATLAS experiment [94], it is neglected in this work by considering the muon charge misidentification probabilities to be $\varepsilon_{\text{misID},\mu} = 0$.

Events characterized by two opposite-sign leptons are composed mainly by two prompt electrons whose energy is reconstructed in the electromagnetic calorimeter with very high efficiency (see Section 3.3.2). Therefore, when these events are weighted to retrieve background yields for SS events, energy leakage due to material interactions has not been experienced by these electrons. To faithfully represent the worsened kinematic reconstruction of electrons after charge misidentification, energy corrections need to be applied to take this leakage into account. The corrections to be applied are described in the following.

7.1.3 Electron energy correction

As seen in Section 6.3.3, the energy resolution worsens notably for charge misidentified electrons. The reconstructed energy in such a cases is considerably lower than the truth energy of the electron. The missing reconstructed energy is lost due to bremsstrahlung radiation after interactions with the detector material when inefficiencies for the energy reconstruction algorithm cannot account these subsequent photon conversions when building the electromagnetic cluster. The energy correction is then closely related to $|\eta|$ and p_T since the charge misidentification rates strongly depends on them by influence of the detector material. To account for this effect after the event weighting introduced in the previous section, the reconstructed electron p_T is corrected by the following expression

$$p_T^{\text{corrected}} = \frac{p_T^{\text{old}}}{\alpha} + dE , \tag{7.4}$$

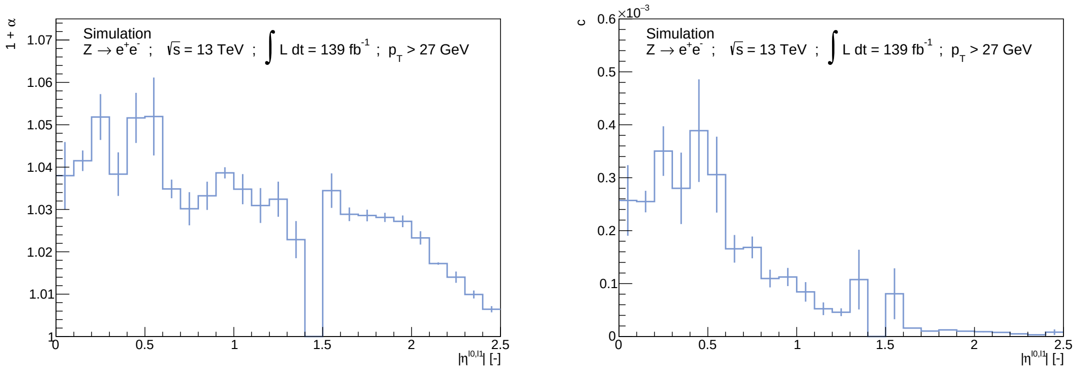


Figure 7.4: Energy correction factors α and c for charge misidentification background estimation computed using Madgraph $Z \to ee$ MC. Only electrons with $p_T > 27$ GeV are considered for the computation of these factors.

where α is the *residual energy scale correction factor*, dE is a *smearing term* and p_T^{old} is the energy measurement in the electromagnetic calorimeter for the electron. The energy correction factor α is computed by considering the averaged energy resolution in different bins i along $|\eta|$ of correctly ($R_{p_T}^{\text{right}}$) and wrongly ($R_{p_T}^{\text{wrong}}$) reconstructed electrons as

$$\alpha_i = \frac{1 + R_{p_T,i}^{\text{right}}}{1 + R_{p_T,i}^{\text{wrong}}} \,. \tag{7.5}$$

The smearing term dE is obtained as a random value from a Gaussian probability distribution centered at zero of the width $\frac{p_T^{\text{old}}}{\alpha}\sqrt{c}$, with c being determined differentially for the $|\eta|$-bin i as

$$c_i = \frac{\Delta^2_{\text{wrong},i}}{(1 + R_{p_T,i}^{\text{wrong}})^2} - \frac{\Delta^2_{\text{right},i}}{(1 + R_{p_T,i}^{\text{right}})^2} \,. \tag{7.6}$$

where Δ_{right} and Δ_{wrong} denote the standard error[2] of the mean energy resolution for correctly and wrongly reconstructed electrons, respectively.

Madgraph $Z \to ee$ MC samples are used for the calculation of the energy correction factors α and c. The results are presented in Figure 7.4. According to the SR selection, only the electrons satisfying $p_T > 27$ GeV are taken into account for the computation of the factors α and c. The energy scale correction factor α decreases for greater values of $|\eta|$, which is expected due to the better energy resolution observed in correctly reconstructed electrons especially at low $|\eta|$ (see Figure 6.6). The parameter c shows the same tendency and takes positive values smaller than 0.0008. This means that the averaged energy correction tend to be similarly worsened within similar uncertainties at high $|\eta|$, while wrong reconstructed electrons are worse reconstructed at low $|\eta|$ when compared to electrons with correctly reconstructed charges.

The energy correction is applied onto the transverse momentum of one randomly selected electron in the event through Eq (7.4). This selection takes into account the lepton charge misidentification probability of the lepton candidates reconstructed in the event.

[2] i.e. the standard deviation divided by the total number of entries.

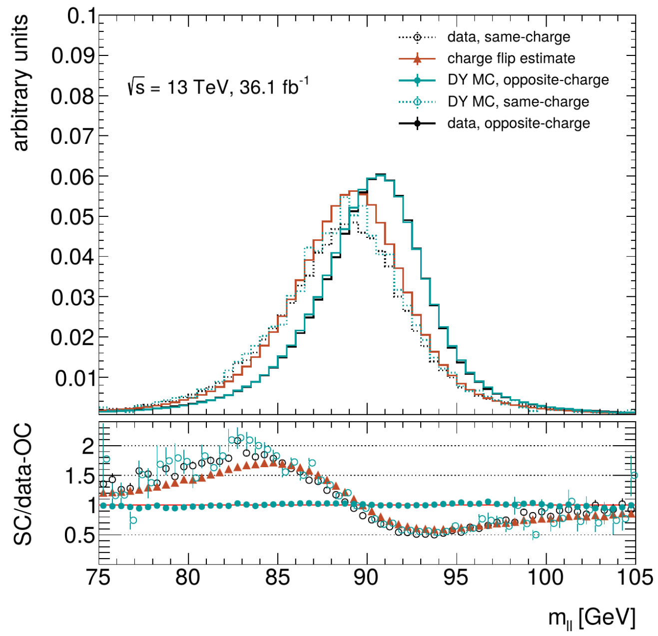


Figure 7.5: Normalized distributions for the reconstructed invariant mass of the Z boson by using data and MC with OS event (solid lines) and SS event (dashed lines) selections. The ratio with respect to OS data is also shown at the bottom. In orange, the charge misidentification background yield obtained by weighting OS data events with Eq 7.3 and applying the energy correction 7.4 on the transverse momentum of the electrons is presented. From the ratio plot, the shift of the $m_{\ell\ell}$ distribution is observed to reproduce consistently the SS event selection by correcting individual electrons, achieving also reduced statistical uncertainties. Source [91].

The sum of charge misidentification probabilities $\delta = \varepsilon_0 + \varepsilon_1 + ... + \varepsilon_N$ is computed throughout the ordered list in decreasing p_T for the N leptons in the event. A random number r is retrieved within 0 and δ. The electron i to be corrected is the one whose probability ε_i exceeds the value r along the cumulative sum δ. Since the ordering of the electrons is conserved along the cumulative sum, it is more likely for the leading lepton to be corrected if $\varepsilon_0 > \varepsilon_1$ and vice versa.

The effect on the reconstructed invariant mass of the leading and subleading leptons after correcting the electron p_T can be observed in Figure 7.5. The reconstructed $m_{\ell\ell}$ with OS events in both MC and data is consistent with the Z mass. However, when two leptons of the same charge are selected, the distribution is shifted to lower energies. This effect was already observed during the studies on the charge misidentification sources at truth level in Section 6.3.2. The ratio with respect to OS data is also shown at the bottom. In orange, the charge misidentification background yield obtained by weighting OS data events is presented and it reproduces consistently the effects due to energy leakage observed in Figure 6.5 for wrongly reconstructed electrons. This is observed in the ratio plot when compared to the SS event selection. Additionally, the weighting of the OS event selection reduces statistical uncertainties notably, as described in this section. In Section 7.2, the method is validated by means of a closure test and computations of the data-driven estimation in a dedicated Control Region (CR) at the Z boson resonance, along with background yields in the Signal Region.

7.2 Results

The methodology introduced in Section 7.1 is applied to estimate the background contribution due to charge misidentification to the measurement of electroweak same-sign W boson pair production with the ATLAS experiment. First of all, the methodology is validated in Section 7.2.1 by means of a closure test and a data-driven estimation in a Z window CR. Then, the obtained contributions of this background source in the SR are presented in Section 7.2.3. Sources of systematic uncertainties affecting the charge misidentification background yields in the SR are studied in Section 7.2.2.

7.2.1 Validations

The method described in Section 7.1 is validated in three steps. Firstly, a closure test is carried out by using Madgraph $Z \rightarrow ee$ MC samples (as for the computation of the charge misidentification probabilities and energy correction parameters) in a Z window CR. For this CR, the overall acceptance of the inner detector ($|\eta| < 2.5$) is considered, excluding the crack region in the EMCAL, $1.37 < |\eta| < 1.52$. The electron energy is required to be reconstructed using the signature in the electromagnetic calorimeter. Additionally, electrons with low transverse momentum $p_T < 27$ GeV are discarded and the invariant mass of the two electrons is required to be $m_{\ell\ell} > 20$ GeV to remove contributions from soft-QCD processes. To avoid the influence in charge misidentification from FSR-induced electrons and other background processes, a selection at truth level is applied on MC samples. A *third lepton veto* is also applied, i.e. only events with two leptons are considered. For the ee channel, the invariant mass of the two electrons passing the third lepton veto is required to be within 15 GeV of difference with respect to the Z boson mass ($|m_Z - m_{ee}| < 15$ GeV). Results in the Z window CR are shown in Figure 7.6 when comparing the SS event selection (black dots) to the estimation using the OS event selection (shown in blue) on which the method exposed in Section 7.1 is applied. Consistent agreement is found within 30% between SS and OS selections. An energy overcorrection is found on the p_T distribution of the leading lepton, which is not present for the subleading lepton. The ratio $p_T^{\ell 1}/p_T^{\ell 0}$ is found to agree between OS and SS event selections, presenting small underestimation from the OS selection for low ratios, when the the transverse momentum of the subleading lepton is lower than half of $p_T^{\ell 0}$. Additionally, the energy correction drives this ratio to values greater than one, propelling the leading lepton to have lower energy than the subleading lepton. Along with the mentioned underestimation for low ratios, a tendency to overestimate charge misidentification contributions for ratios close to one might indicate that the energy of the leading lepton is corrected more often than the subleading lepton, according to the expectations from the SS event selection. The arising $m_{\ell\ell}$ distribution obtained from the OS event selection shows deviations up to 40% for $m_{ee} < m_Z$ with respect to the expectation from SS events, where underestimation in the OS selection is observed. This might be related to the differences between distributions of corrected p_T of leading and subleading leptons. The η dependencies of charge misidentified electrons in pseudorapidity (that reproduce the results presented in Section 6.3) for leading and subleading lepton agree satisfactorily within 20%.

Further dilepton variables reconstructed with leading and subleading lepton are shown

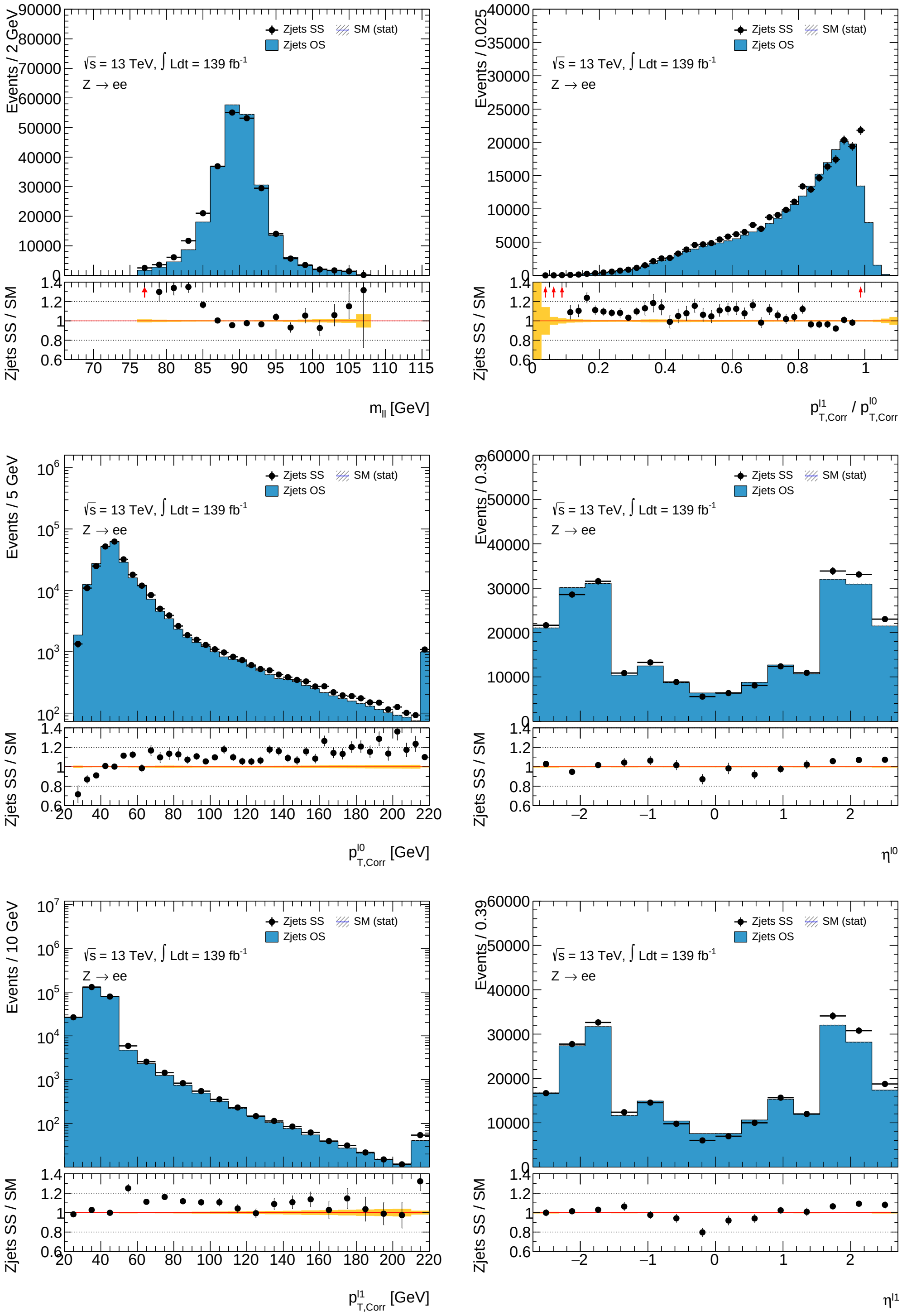


Figure 7.6: Kinematic variables for the closure test of the methodology presented in Section 7.1. Madgraph $Z \to ee$ MC samples are selected for full Run 2 in the Z window CR to have opposite-sign electrons in the final state (Zjets OS, shown in blue) on which the method is applied. The resulting yields are then compared to a event selection requiring same-sign leptons in the final state (Zjets SS, shown as black dots). On top, the invariant mass of constructed with leading and subleading lepton (left), the reconstructed transverse momentum of the leading lepton measured after energy correction $p^{\ell 0}_{T,\mathrm{Corr}}$ (center) and the pseudorapidity distribution of the leading lepton $\eta^{\ell 0}$ (right) are shown. On the bottom, the ratio $p^{\ell 1}_T / p^{\ell 0}_T$ of corrected transverse momentum of subleading over leading lepton (left), the reconstructed transverse momentum of the subleading lepton measured after energy correction $p^{\ell 1}_{T,\mathrm{Corr}}$ (center) and the pseudorapidity distribution for the subleading lepton $\eta^{\ell 1}$ (right). The ratio SS/OS is computed for the ration plot.

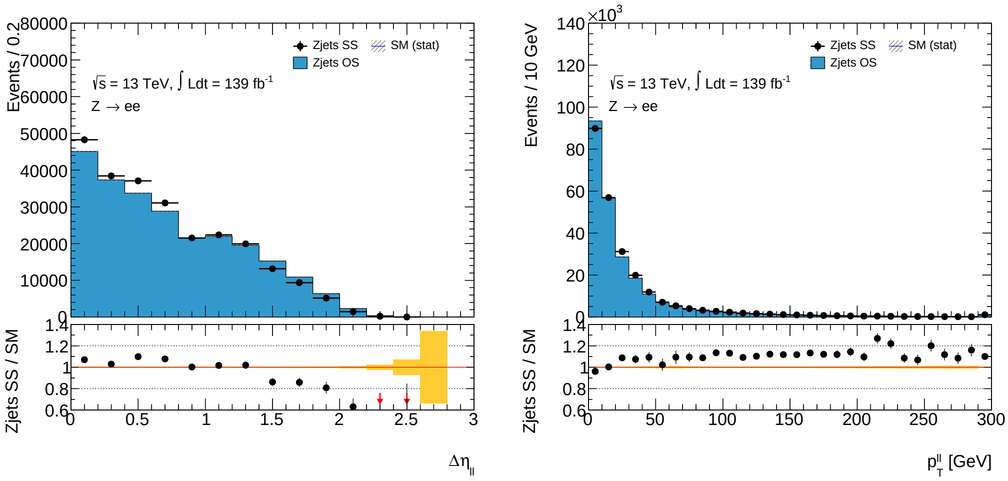


Figure 7.7: Dilepton variables for closure test of methodology presented in Section 7.1. Madgraph $Z \to ee$ MC samples are selected for full Run 2 in the Z window control region to have opposite-sign electrons in the final state (Zjets OS, shown in blue) on which the method is applied. The resulting yields are then compared to a event selection requiring same-sign leptons in the final state (Zjets SS, shown as black dots). The difference of pseudorapidity $\Delta\eta_{\ell\ell}$ (left) and the transverse component of three momentum $p_T^{\ell\ell}$ for the summed Lorentz vector of leading and subleading lepton (right) are presented. The ratio SS/OS is computed for the ration plot.

ee	mc16a [raw]	mc16d [raw]	mc16e [raw]	run2 [raw]
ZjetsSS	67000 ± 1000 [24576]	76890 ± 1080 [30903]	104400 ± 1600 [32815]	248100 ± 2200 [88294]
ZjetsOS	65960 ± 130 [2.50·10^6]	76090 ± 140 [2.93·10^6]	101240 ± 210 [3.08·10^6]	243300 ± 300 [8.52·10^6]
%	1.2 ± 1.5	1.0 ± 2.1	3.0 ± 1.5	1.9 ± 0.9

Table 7.1: Integrated number of events for OS and SS event selections in the Z window CR along different MC campaigns mc16a (corresponding to 2015-2016 dataset), mc16d (corresponding to 2017 dataset), mc16e (2018 dataset), and full Run 2 (2015-2018 dataset) for the closure test using $Z \to ee$ Madgraph MC samples. The raw number of events is included in squared brackets. The deviation between both selections is shown in percentage on the third row, whose uncertainty is propagated from the statistical uncertainties of the yields.

in Figure 7.7. The transverse component of three momentum $p_T^{\ell\ell}$ for the summed Lorentz vector of leading and subleading leptons and the difference of pseudorapidity $\Delta\eta_{\ell\ell}$ show agreement between OS and SS events within 20%. The method tends to overestimate the background yield for dilepton decays with $\Delta\eta_{\ell\ell} > 2$ up to 40%, i.e. when the two electrons in the Z boson decay are back-to-back along the beam axis. Besides, the method is observed to be consistent outside the Z window, as observed in Figure 7.8.

From the closure test using $Z \to ee$ MC samples, the method is shown to be reliable along the overall range of p_T and η when matching the weighted and energy-corrected OS events to SS events for the full Run 2 MC set. The integral values for both selections along different MC campaigns and full Run 2 set are presented in Table 7.1. For full Run 2, 243300 ± 300 events were observed after applying the method on the OS event selection, while 248100 ± 2200 are found for the SS selection. The integral values match between selections within statistical uncertainties in mc16a and mc16d. Underestimation of 3.0 ± 1.5% is found in mc16e, that drives the closure test for the OS event selection in $Z \to ee$ to underestimate the SS selection by 1.9 ± 0.9 %

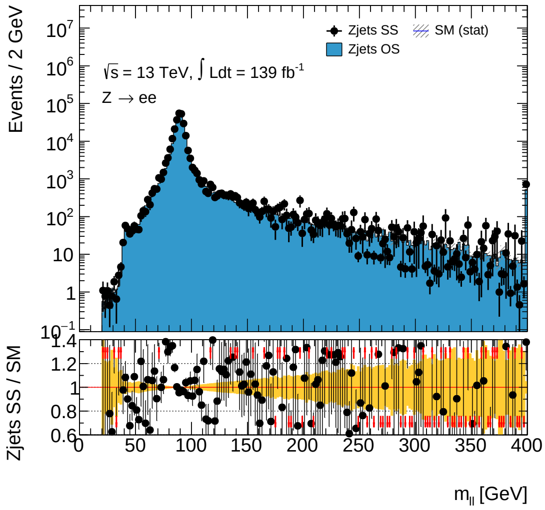


Figure 7.8: Invariant mass of constructed with leading and subleading lepton for the closure test of the methodology presented in Section 7.1. Madgraph $Z \to ee$ MC samples are selected for full Run 2 in the Z window CR to have opposite-sign electrons in the final state (Zjets OS, shown in blue) on which the method is applied. Results are compared to the SS event selection (Zjets SS, shown as black dots). In this plot, the criterion $|m_Z - m_{ee}| < 15$ GeV is removed. The ratio SS/OS is computed for the ration plot.

Secondly, $t\bar{t}$ MC samples were used in ee, $e\mu$, and μe channels for the full Run-2 set in the Z window control region, with the criterion $|m_Z - m_{ee}| < 15$ GeV dropped. The same charge misidentification probabilities and energy correction factors computed in $Z \to ee$ are used. Distributions for SS event selection and weighted and energy-corrected OS selection also agree within 20%. Compared to the $Z \to ee$ closure test presented in Figure 7.6, similar tendencies on the kinematic distributions of leading and subleading leptons are observed. However, the methodology shows an overall underestimation of $7.2 \pm 0.4\%$ from the OS event selection, probably due to a different behavior of the FSR-induced electrons among MC generators. The integral values for different MC campaigns are shown in Appendix C. Interesting differences are found when comparing the kinematic distributions of $t\bar{t}$ events decaying into $e\mu$ and μe channels. The most relevant distributions are presented in Figure 7.9, corresponding to the p_T of the leading lepton ℓ_0 (top) and subleading lepton ℓ_1 (center-top), the ratio $p_T^{\ell_1}/p_T^{\ell_0}$ between them (center-bottom) and the distribution of the invariant mass of the two leptons $m_{\ell\ell}$ (bottom). Since charge misidentification probabilities are zero for muons, the energy corrections should only be applied on the electron (found in ℓ_0 and ℓ_1 in the $e\mu$ and μe channels, respectively). Event weights given by Eq. (7.3) only consider the kinematics of the electron in the final state. The $e\mu$ channel shows a better estimation than the μe decays, where the muon is tagged as the the leading lepton. This can be seen from the transverse momentum of the muon in both channels, where no corrections are applied. The ratio SS/OS events for the $p_T^{\ell_1}/p_T^{\ell_0}$ distribution shows also a flat tendency in the $e\mu$ channel, while this ratio increases with greater $p_T^{\ell_1}/p_T^{\ell_0}$ for the μe channel. Similar dependencies are found for $m_{\ell\ell}$, being flatter for $e\mu$ decays. Asymmetries on the electron selection to apply the energy correction strategy might cause these differences. The methodology applied on the OS event selection underestimates the expected total number of events from the SS selection in the CR by $7.19 \pm 0.4\%$ and $7.6 \pm 0.6\%$ in the $e\mu$ and μe channels, respectively.

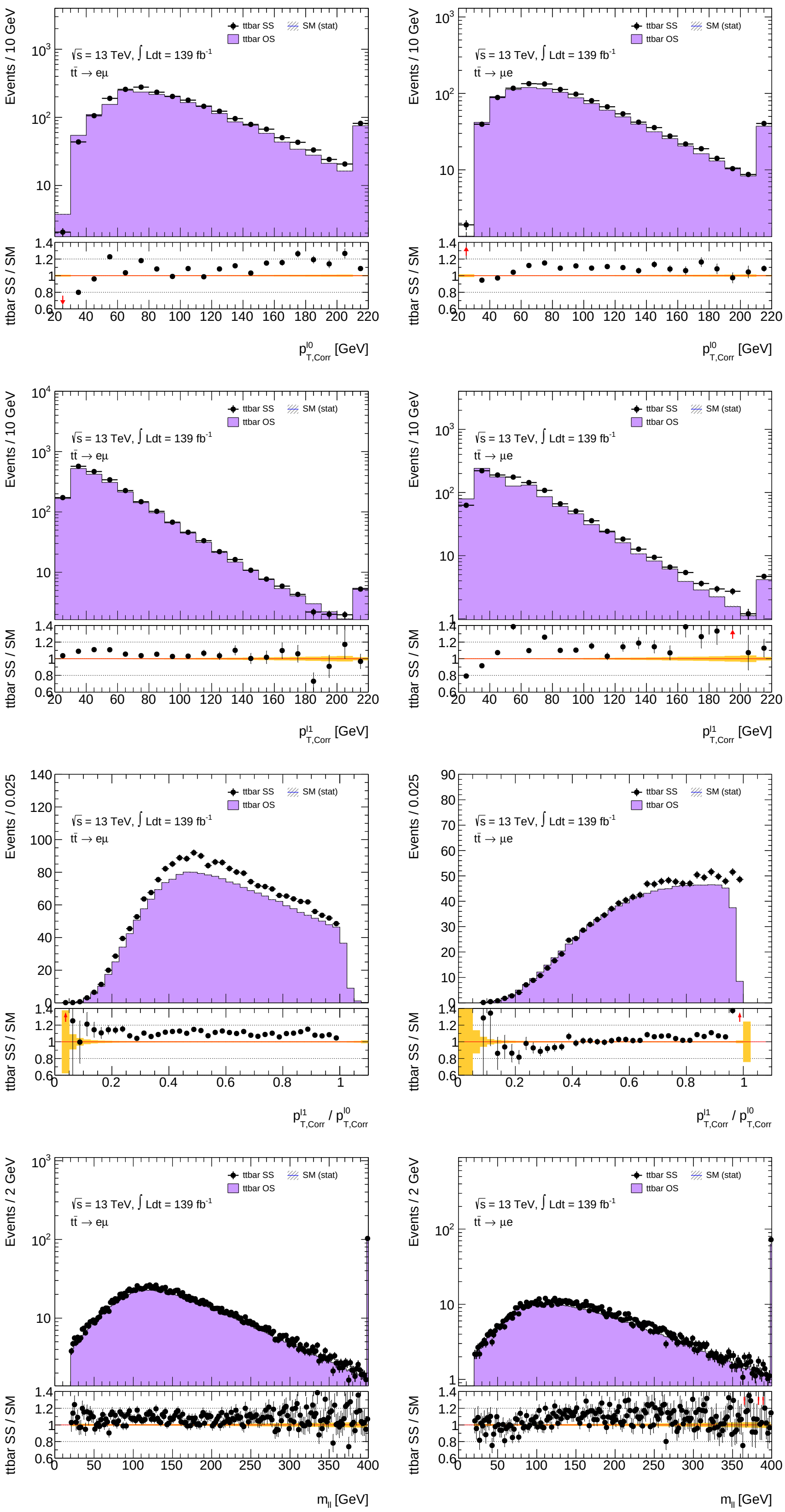


Figure 7.9: Kinematic distributions for SS and weighted and energy-corrected OS event selections in Powheg $t\bar{t}$ MC samples decaying into the $e\mu$ channel (left) and μe channel (right) for full Run 2 in the Z window CR. The p_T of the leading lepton (top), p_T of the subleading lepton (center), ratio $p_T^{\ell_1}/p_T^{\ell_0}$ of energy-corrected transverse momentum of subleading over leading leptons (top-right) and invariant mass of the two leptons $m_{\ell\ell}$ (bottom-right) are presented. The ratio SS/OS is computed for the ration plot.

ee channel	2015-2016	2017	2018	Run 2
Data	**51009**	**66286**	**86416**	**203711**
ssWW EWK	7	8	11	26
ssWW QCD	3	3	5	11
top	7	8	10	25
WZ	101	121	154	376
ZZ	10	12	16	38
Vγ	949	1208	543	2700
FSR	2155	3280	4233	9668
Charge flip	49534	58321	76135	183991
Total SM	**52766**	**62961**	**81107**	**196835**
(bkg-data)/data [%]	3.4	-5.0	-6.1	-3.4

Table 7.2: Number of events obtained in the Z window CR from the data-driven estimation in the ee channel for different datasets recorded by the ATLAS experiment in years 2015-2016, 2017 and 2018, along with the full Run 2 dataset. The category "*top*" includes contributions from single-top production, $t\bar{t}$ and ttX. "*WZ*" includes QCD and EWK production along with $WqqZ\ell\ell$, and $W\gamma$ and $Z\gamma$ compose the category "$V\gamma$". The FSR contribution is estimated using $Z \to ee$ MC with a truth selection.

Lastly, the data-driven estimation is carried out in the Z window CR after the validation with the closure test. The methodology introduced in Section 7.1 is again verified using data events selected with opposite-sign electrons in the final state. The results are shown in Figure 7.10 along with the rest of prompt background contributions. Since the fake background was not fully implemented at the time this thesis was written, their contribution was not considered. Non-prompt background contributions are expected to be negligible in this CR for the ee channel. However, they become important in $e\mu$ and μe channels. The validation through the data-driven estimation is therefore studied in ee decays, where the charge misidentification is dominant. Since FSR radiation cases were removed for the computation of charge misidentification probabilities by applying a truth requirement, their contribution is added using Madgraph $Z \to ee$ MC using truth information. Aside the contributions from charge misidentification and FSR radiation, the contributions from prompt backgrounds sum up to 1-2%, varying among campaigns. The yields from different contributions are provided in Table 7.2. Without accounting for non-prompt contributions, the total expected background from Standard Model sources (denoted by "*Total SM*") differs from the observed data events within 6% along data taking periods. The category "*top*" includes contributions from single-top production, $t\bar{t}$ and ttX after passing the truth selection. This avoids possible double counting of events in the total background estimation due to charge misidentification. However, the $t\bar{t}$ component is then populated by FSR-induced electrons, that might be in conflict with the FSR component. The influence of $t\bar{t}$ is considered to be negligible in any case. "*WZ*" includes QCD and EWK production along with $WqqZ\ell\ell$ after passing truth selection, and $W\gamma$ and $Z\gamma$ compose the category "$V\gamma$". The truth selection is disabled for this category. Kinematic properties in the Z window CR are presented in Figure 7.10. The background is correctly estimated in the CR within deviations of 25% with respect to data. The kinematic distributions improve with the separate treatment of FSR-induced electrons, as can be seen from the distributions shown in Section C.2. Dependencies for leading and subleading lepton are correctly estimated. Energy overcorrection is again found on the leading lepton p_T. The invariant mass of leading and subleading leptons $m_{\ell\ell}$ shows good agreement between data and background estimation, being underestimated for $m_{ee} < m_Z$ up to 30%. The pseudorapidity of leading and subleading lepton are well reproduced with respect to data within deviations of 20%. As also observed

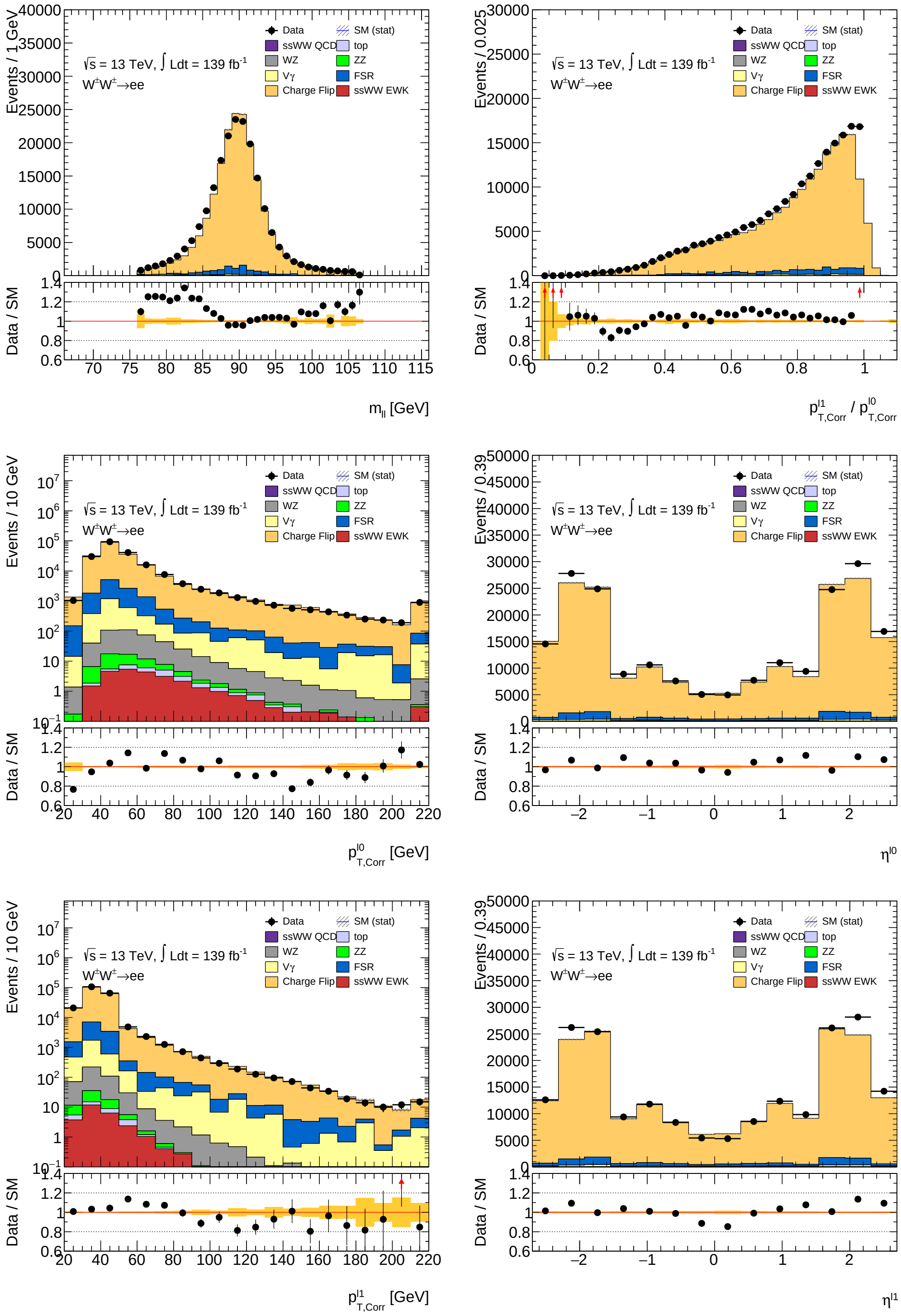


Figure 7.10: Expected yields from Standard Model sources along with data observation in the Z window CR in the ee channel for the full Run 2 dataset. A data-driven estimation is carried out for contributions due to charge misidentification. The category "top" includes contributions from single-top production, $t\bar{t}$ and ttX. "WZ" includes QCD and EWK production along with $WqqZ\ell\ell$, and $W\gamma$ and $Z\gamma$ compose the category "$V\gamma$". The FSR contribution is estimated using $Z \to ee$ MC by a truth requirement. The ratio Data/Total-SM is computed for the ration plot.

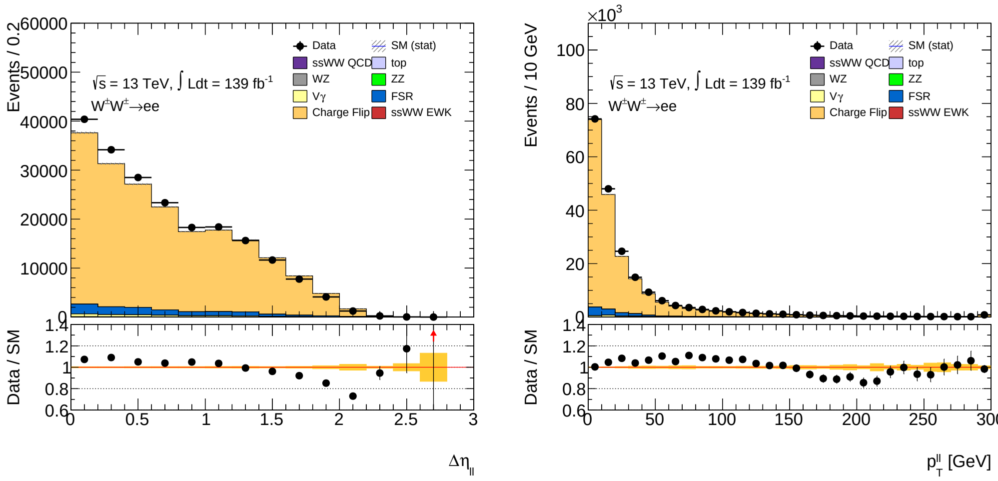


Figure 7.11: Dilepton variables for the data-driven background estimation due to charge misidentification using the methodology presented in Section 7.1. Data events with opposite-sign leptons in th final state are selected in the Z window control region (Zjets OS, shown in yellow) which the method is applied onto. The resulting yields are then compared to data events selected by requiring same-sign leptons in the final state (shown as black dots). The difference of pseudorapidity $\Delta\eta_{\ell\ell}$ (left), the transverse component of three momentum $p_T^{\ell\ell}$ for the summed Lorentz vector of leading and subleading lepton (center) and angular aperture between the two leptons $\Delta R_{\ell\ell}$ (right) are presented. The ratio Data/Total-SM is computed for the ration plot.

in the closure test using $Z \to ee$, the data-driven method tends to overpopulate central regions of the detector, providing an underestimation for $|\eta| > 2$. Additionally, the estimation of the ratio $p_T^{\ell 1}/p_T^{\ell 0}$ is consistent with the data. Dilepton distributions are also reproduced within the same uncertainties, as seen in Figure 7.11 for the $|\Delta\eta_{\ell\ell}|$ and $p_T^{\ell\ell}$ and they improve when compared to the results from the closure test using MC $Z \to ee$. The data is also well reproduced by the background estimation when the Z window criterion ($|m_Z - m_{\ell\ell}| > 15$ GeV for the ee channel) is disabled, as can be seen in Figure 7.12 for the reconstructed invariant mass of leading and subleading leptons, $m_{\ell\ell}$.

The data-driven estimation of the charge misidentification background contribution is found to be valid in the ee channel for the Z window CR, where this background contribution is dominant (as it constitutes more than 90% of the total background yield). The method is consistent to data along all the kinematic distributions within deviations of 25%, showing important improvements with respect to the previous analysis [91]. For the full Run 2 dataset, 196835 events were observed for the sum of background yields, where 183991 are produced by charge misidentification of one of the electrons. 203711 events are observed in data. Without the consideration of non-prompt background contributions, the integral values agree within 3.5%. Once the methodology has been proven to be reliable, the background yields in the SR can be computed. In Section 7.2.3, the background yields in the SR for the charge misidentification background are presented.

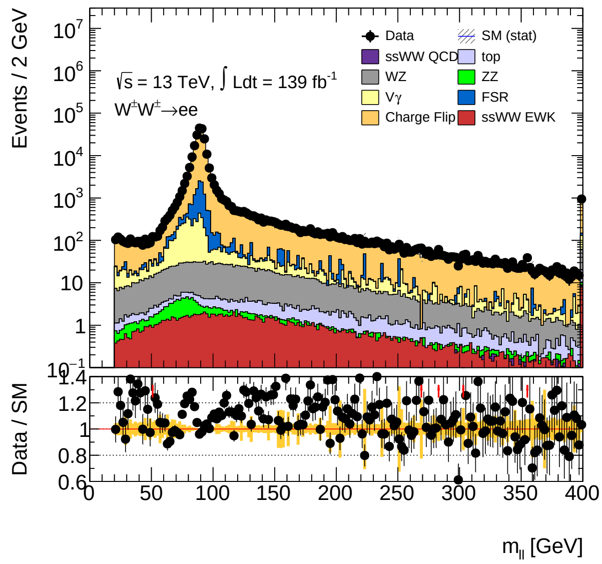


Figure 7.12: $m_{\ell\ell}$ distribution from sum of Standard Model sources along with data observation for the Z window CR in the ee channel for the full Run 2 dataset. The condition $|m_{\ell\ell} - m_Z| < 15$ GeV is removed to explore large values of $m_{\ell\ell}$. A data-driven estimation is carried out for contributions due to charge misidentification. The category *"top"* includes contributions from single-top production, $t\bar{t}$ and ttX. *"WZ"* includes QCD and EWK production along with $WqqZ\ell\ell$, and $W\gamma$ and $Z\gamma$ compose the category *"$V\gamma$"*. The FSR contribution is estimated using $Z \rightarrow ee$ MC by a truth requirement. The ratio Data/Total-SM is computed for the ration plot.

7.2.2 Sources of systematic uncertainties

For the estimated background yields due to charge misidentification, the sources of systematic uncertainties can be divided in two main categories. On the one hand, different approaches can be followed to compute the charge misidentification probabilities. The deviations due to the chosen methodology are considered from the systematic variations of the scale factors presented in Section 7.1.1. On the other hand, the smearing factor dE applied on the corrected electron p_T (introduced in Eq. (7.4)) might seem arbitrary. Therefore, systematic variations are also applied on this parameter.

The systematic variations of the scale factors are provided by the *Egamma Combined Performance (CP) group* [68] for different electron identification and isolation working points. No scale factors including the 2018 dataset are available at the time this thesis is written. Therefore, their distributions obtained from the 2015-2017 dataset are used to compute the charge misidentification background for the full Run 2 dataset. For the 2015-2017 dataset, the computation of the SF and its systematic uncertainties are investigated in more detail when compared to the previous analysis. These were presented in Figure 7.1 for `TightLH` electron identification and `Gradient` electron isolation, along with statistical and systematic uncertainties. The distributions used in the previous analysis can be found in Section C.1. Different systematic variations are applied when computing the charge misidentification probabilities in data and MC for the calculation of the scale factors:

- For both methods, a sideband substraction in $Z \to ee$ decays is carried out at the Z resonance to obtain the electron charge misidentification probabilities from OS and SS event selections. The width of the sideband is modified for systematic variations.
- For the computation of the charge misidentification probabilities in MC, differences for the removal of FSR-induced electrons are also considered for systematics.
- Non-closure of MC charge misidentification probabilities when compared to rates derived on electron selection at truth level is also taken into account.
- To reduce statistical fluctuations at high transverse momentum, a 1D×1D parameterization is exploited for the calculation of the scale factors at $p_T > 95$ GeV by assuming that the dependencies along p_T and η are decorrelated. Differences between the 1D×1D and the 2D parameterizations are also included as systematic variations.

Details on the treatment and the size of the different systematic sources, as well as the computation of the total systematic variation can be found in Ref. [68]. Additionally, variations on the energy correction smearing factor dE (introduced in Eq. (7.4)) are computed by assuming a conservative uncertainty of 100%. Both categories of systematic variations are considered separately, i.e., variations on the energy correction smearing are computed using the nominal distributions of the scale factors, while systematic variations from scale factors are considered applying the energy correction presented in Eq. (7.4). The total systematic uncertainty is given by the mean value of maximum upper and lower deviations for symmetrization. The total uncertainty of the background yields sums statistical and total systematic uncertainty in quadrature.

For the systematic variations due to SF uncertainties, charge misidentification probabilities and energy correction parameters are computed for nominal, up and down variations. The distributions for the variations on the charge misidentification probabilities are shown in Appendix C.6. Consequently, background yields for nominal, up and down variations are calculated using the corresponding distributions of probabilities and energy correction factors. Since the down variations of the scale factors can provide negative values, negative charge misidentification probabilities are found when weighting wrongly reconstructed electrons with negative scale factors as described in Section 7.1. In such a cases, the obtained probability is considered to be zero. This effect is realized at low p_T and low pseudorapidity, where the systematic uncertainties for wrongly reconstructed electrons are large (see Figure 7.1).

7.2.3 Charge misidentification background estimation in the signal region

In this section, the background yields due to charge misidentification are propagated to the SR and its systematic variations are computed for the 2015-2016 and the full Run 2 datasets. The first allows for comparisons with respect to the previous analysis. The last will be relevant for the current measurements on same-sign W boson pair production using the 139 fb^{-1} recorded dataset from 2015 until 2018.

The results for nominal background yields and their systematic and statistical uncertainties are presented in Table 7.3. The systematic variations from the SF and energy

36.1 fb^{-1}	ee	$e\mu$	μe	$\ell\ell$
Charge flip (Nom $\pm$ stat)	2.90 $\pm$ 0.15	6.6 $\pm$ 0.5	2.5 $\pm$ 0.2	12.0 $\pm$ 0.6
SF up	0.97	1.46	0.39	2.82
SF down	-0.95	-1.44	-0.37	-2.76
$dE \times 0$	-0.01	0.0	-0.001	-0.01
$dE \times 2$	0.03	0.0	0.0	0.03
total sys. variation	0.96	1.4	0.38	2.8
stat $\oplus$ sys	0.97	1.5	0.43	2.9

139 fb^{-1}	ee	$e\mu$	μe	$\ell\ell$
Charge flip (Nom $\pm$ stat)	11.8 $\pm$ 0.3	22.1 $\pm$ 0.9	9.4 $\pm$ 0.4	43.3 $\pm$ 1.1
SF up	3.93	4.7	1.74	10.3
SF down	-3.88	-4.6	-1.66	-10.2
$dE \times 0$	-0.06	0.003	-0.004	-0.07
$dE \times 2$	0.02	0.003	-0.0002	0.02
Total sys. variation	$\pm$ 3.9	$\pm$ 4.6	$\pm$ 1.7	$\pm$ 10.2
stat $\oplus$ sys	$\pm$ 3.9	$\pm$ 4.7	$\pm$ 1.8	$\pm$ 10.3

Table 7.3: Background yields in the electron channels (ee, $e\mu$ and μe) for the charge misidentification background contribution in the SR for the 2015-2016 dataset (top) and full Run 2 (bottom). Statistical uncertainties and systematic variations are shown.

correction smearing are presented separately. In Section 6.2, comparisons between correctly and wrongly reconstructed electrons regarding their electric charge were done. Correctly reconstructed electrons were found to be produced in central regions of the inner detector, while charge-misidentified electrons are mostly produced within $|\eta| > 1.52$, where the probability of interaction with the detector material is larger. In the Signal Region, the yields for the ee channel are constrained by $|\eta_{ee}| < 1.37$ and a Z window veto ($|m_Z - m_{ee}| > 15$ GeV), which reduces the charge misidentification background in this channel drastically. Additionally, the $e\mu$ channel present more events for this background contribution than the μe channel, since the charge misidentification is greater at high p_T, when the electron is tagged as the leading lepton. The $\mu\mu$ channel is not presented since the charge misidentification phenomenon is neglected for muons. The background yields for mc16a can be compared to the systematic variations from the previous analysis (presented in Table 6.1). Systematic variations arising from the SF are found to be the dominant source of uncertainty for the charge misidentification background yields and they increase considerably with respect to the previous analysis. In the ee channel, variations of 9% were found in the previous analysis, increasing up to 33% for the current analysis. In the $e\mu$ and μe channels the increase goes from 3.3% and 2.7% up to 21% and 15% respectively. When the $\ell\ell$ channel is considered, it grows from 4.8% of the nominal value up to 23%. In this analysis, this source of uncertainty surpass the statistical uncertainties and the variations from energy correction smearing in one and three orders of magnitude respectively. The asymmetry between the energy correction deviations found between $e\mu$ and μe channels in the previous analysis (whose variation on the energy correction smearing is computed as $dE \times 0$) is not found for the current analysis, as well as its dominance in the $e\mu$ channel. The proportions are conserved when considering the full 139 fb^{-1} dataset, since variations from scale factors are dominant. For this dataset, statistical uncertainties correspond to 7% and 9% of the total uncertainty of the background yield for ee and $\ell\ell$ channels.

An important amount of events arising in the SR due to charge misidentification are

Signal region ($\|\eta_{ee}\| < 2.5$)	ee	$e\mu$	μe	$\ell\ell$
Charge flip (Nom $\pm$ stat)	88.3 ± 2.1	22.1 ± 0.9	9.4 ± 0.4	119.9 ± 2.3
SF up	19.23	4.68	1.74	25.6
SF down	-19.02	-4.62	-1.66	-25.3
$dE \times 0$	0.011	0.003	-0.004	0.010
$dE \times 2$	0.13	0.003	-0.0002	0.13
Total sys. variation	$\pm$ 19	$\pm$ 5	$\pm$ 1.7	$\pm$ 25
stat $\oplus$ sys	$\pm$ 19	$\pm$ 5	$\pm$ 1.8	$\pm$ 26

Table 7.4: Background yields in the electron channels (ee, $e\mu$ and μe) for the charge misidentification background contribution in the SR with relaxed cut $|\eta_{ee}| < 2.5$ in the ee channel. Values for the full Run 2 dataset are presented with statistical uncertainties and different systematic variations.

removed by the requirements on pseudorapidity and the Z window veto. The measurement in the ee channel is consequently constrained with an additional reduction of the observed signal events. This treatment of the SR for the ee channel brings complications when defining different detector acceptances among channels. In order to increase the observed signal and avoid these complications, relaxing the constraints on pseudorapidity up to $|\eta| < 2.5$ is desired (as done for the channels involving muons). The relaxation of the η_{ee} cut in the SR is studied in Table 7.4, where the background yields and corresponding uncertainties are provided. The background yields due to charge misidentification grow by a factor of 7.5 after increasing the acceptance in pseudorapidity for the ee channel. The total uncertainty is 22% of the nominal value. A similar proportion is found for the $\ell\ell$ channel. Since no modifications were done for final states involving muons, the background yields are identical for the $e\mu$ and μe channels. Since SF still dominate the total uncertainty, its reduced proportion is totally related to the increase of the background yield.

The relaxation of the pseudorapidity criterion can be also studied through the signal significance. Since the uncertainty of the background yields play a crucial role for this analysis, the inclusion of systematic uncertainties for the evaluation of the statistical significance is desired. For counting experiments following Poisson statistics, the following expression is used for the computation of the *statistical significance* Z [95]

$$Z = \frac{2\,B\,(\sqrt{S+B} - \sqrt{B})}{B + (\Delta B)^2}, \tag{7.7}$$

where B is the background estimate with Gaussian uncertainty ΔB and S is the observed signal. Since non-prompt background contributions were not included for this analysis, the total background will be estimated as the difference between data and signal events. The uncertainty of this background is assumed to arise due to charge misidentification only.

Since the data events are only known for the 2015-2016 recorded dataset due to the blinding strategy, the statistical significance can only be studied in this dataset using 36.1 fb^{-1}. The values before and after the η relaxation are presented in Table 7.5. After the increase in detector acceptance, 54% of the observed data events are expected to arise from charge misidentification sources. Only 10% are expected to originate from electroweak production of same-sign W boson pairs. As expected, values for $e\mu$ and μe remain unchanged. The significance Z is dramatically reduced by almost 1σ for this background estimation, despite the lower proportion of the background uncertainty. This

mc16a ($|\eta_{ee}| < 1.37$)	ee	$e\mu$	μe	$\ell\ell$
Data	15	44	41	144
Signal	5.7 ± 0.3	12.7 ± 0.4	12.6 ± 0.5	45.5 ± 0.8
Charge flip ($\pm$ stat $\oplus$ sys)	2.9 ± 0.9	6.6 ± 1.6	2.5 ± 0.5	12 ± 3
Z	1.5	1.9	2.1	3.8

mc16a ($|\eta_{ee}| < 2.5$)	ee	$e\mu$	μe	$\ell\ell$
Data	50	44	41	179
Signal	9.3 ± 0.4	12.7 ± 0.4	12.6 ± 0.5	49.2 ± 0.9
Charge flip ($\pm$ stat $\oplus$ sys)	22 ± 5	6.6 ± 1.6	2.5 ± 0.5	31 ± 6
Z	0.8	1.9	2.1	2.9

Table 7.5: Signal significances Z before (top) and after (bottom) relaxation of $|\eta|$ cut in the Signal Region, considering total uncertainties of the charge misidentification background estimation.

reduction might avoid the claim of a possible discovery or observation of the desired measurement. The relaxation in pseudorapidity for this background estimation is, therefore, unfeasible in a precision measurement.

Nevertheless, the detector acceptance could be increased if an appropriate strategy for the reduction of charge misidentified electrons is found. In Section 7.3, the Electron Charge ID Selector is applied to discard events with charge misidentified electrons. The method using this tool is validated as done for the presented background estimation through a closure test using $Z \to ee$ decays (Section 7.3.1), studying with $t\bar{t}$ decay channels involving muons. After validating the data-driven background estimation in the Z window CR, the feasibility of the cut relaxation is studied once again after the background reduction in Section 7.3.2.

7.3 Charge misidentification background estimation using Electron Charge ID Selector (ECIDS)

To reduce the background yields due to charge misidentification, the Electron Charge ID Selector (also referred to as *ECIDS*) is used. This Boosted Decision Tree (BDT) is developed and trained by the *Egamma CP group*. Using `TightLH` electron isolation, the tool provides a signal efficiency of 98% [68] for the removal of charge-misidentified electrons.

The methodology follows closely the procedure described in Section 7.1. The ECIDS tool requirement is added after electron identification. Only events where the leading and subleading leptons have correctly-identified charges, according to the output of the ECIDS tool, are considered. This requirement is applied on electrons in every decay channel. Charge misidentification probabilities and energy correction parameters need to be recomputed for the new set of wrongly reconstructed electrons, after applying a new set of SF for its dedicated electron charge identification.

A comparison between SF with and without the use of ECIDS are provided in Appendix C.4. Two main differences are observed on the coarser binning of the distribution and the drastic increase of systematic uncertainties. In order to optimize statistical fluctuations due to the reduced amount of wrongly charge-identified electrons, SF with four

ee	mc16a [raw]	mc16d [raw]	mc16e [raw]	run2 [raw]
ZjetsSS	5700 $\pm$ 300 [3965]	6300 $\pm$ 300 [5112]	8100 $\pm$ 400 [5599]	20100 $\pm$ 600 [14676]
ZjetsOS	5408 $\pm$ 9 [2.44$\times10^6$]	6190 $\pm$ 10 [2.85$\times10^6$]	8249 $\pm$ 15 [2.99$\times10^6$]	19850 $\pm$ 20 [8.28$\times10^6$]
%	4 $\pm$ 5	2 $\pm$ 4	-2 $\pm$ 5	1 $\pm$ 3

Table 7.6: Integrated number of events for OS and SS event selections in the Z window CR along different MC campaigns mc16a (corresponding to 2015-2016 dataset), mc16d (corresponding to 2017 dataset), mc16e (2018 dataset), and full Run 2 (2015-2018 dataset) for the closure test using $Z \to ee$ Madgraph MC samples using removal of charge-misidentified electrons with the ECIDS selection tool. The raw number of events is included in squared brackets. The deviation between both selections is shown in percentage on the third row, whose uncertainty is propagated from the statistical uncertainties of the yields.

bins in p_T {20.0, 60.0, 90.0, 130.0, 999.9} and seven bins in $|\eta|$ {0.0, 0.75, 1.25, 1.37, 1.52, 1.7, 2.3, 2.5} are provided. However, systematic uncertainties increase to more than 100% of the nominal value for some of these bins, especially at low $|\eta|$ and transverse momentum. This is mostly caused by an asymmetric removal of electrons in terms of η and charge sign, that are accounted as a new source of systematic variations [68].

The charge misidentification probabilities are computed using the same binning as the distributions when ECIDS tool was not implemented. The binning in p_T is divided in 13 bins as {20.0, 25.0, 30.0, 35.0, 40.0, 45.0, 60.0, 77.5, 95.0, 115.0, 145.0, 185.0, 240.0, 999.0} while 15 bins {0.0, 0.5, 0.75, 1.05, 1.25, 1.37, 1.52, 1.62, 1.7, 1.8, 1.9, 2.0, 2.1, 2.3, 2.4, 2.5} are set for the binning in $|\eta|$. After applying the ECIDS selection, the probability of charge misreconstruction observed at truth level in detector simulation is reduced considerably. This reduction is especially relevant at low $|\eta|$ and p_T (coinciding with the most populated region due to $Z \to ee$ kinematics), where reductions of a factor of 10 are achieved. This reduction is weaker for bins at higher p_T. An integrated probability for charge misidentification of 0.025% is achieved with the use of this BDT with the selection applied for the computation of the charge misidentification rates. Studies of the charge misidentification probabilities for different MC campaigns show no systematic deviations, exploiting the full Run 2 dataset for the computation of these distributions. The comparisons can be found in Appendix C.7. Since the binning of these distributions was not optimized for the use of the ECIDS tool, large statistical uncertainties are observed on the charge misidentification probabilities, especially at low $|\eta|$ and p_T where the population of charge misidentified electrons is low.

The charge misidentification rates using the full Run 2 dataset are found to perform well for the background estimation in the Z window CR. As done in Section 7.2.1, the method is validated by means of a closure test using MC $Z \to ee$ decays. Tests on the $e\mu$ and μe channels are also carried out using $t\bar{t}$ leptonic decays. Additionally, the data-driven method is validated in the same CR. The corresponding background yields in the SR are given in Section 7.3.2.

7.3.1 Validations

Using Madgraph MC samples for $Z \to ee$ decays, a closure test of the methodology presented in Section 7.1 with the ECIDS tool selection is performed. The CR is selected as described in Section 7.2.1. Kinematic distributions for leading (ℓ_0) and subleading electrons (ℓ_1), the reconstructed invariant mass $m_{\ell\ell}$ and the ratio $p_T^{\ell_1}/p_T^{\ell_0}$ of energy-corrected

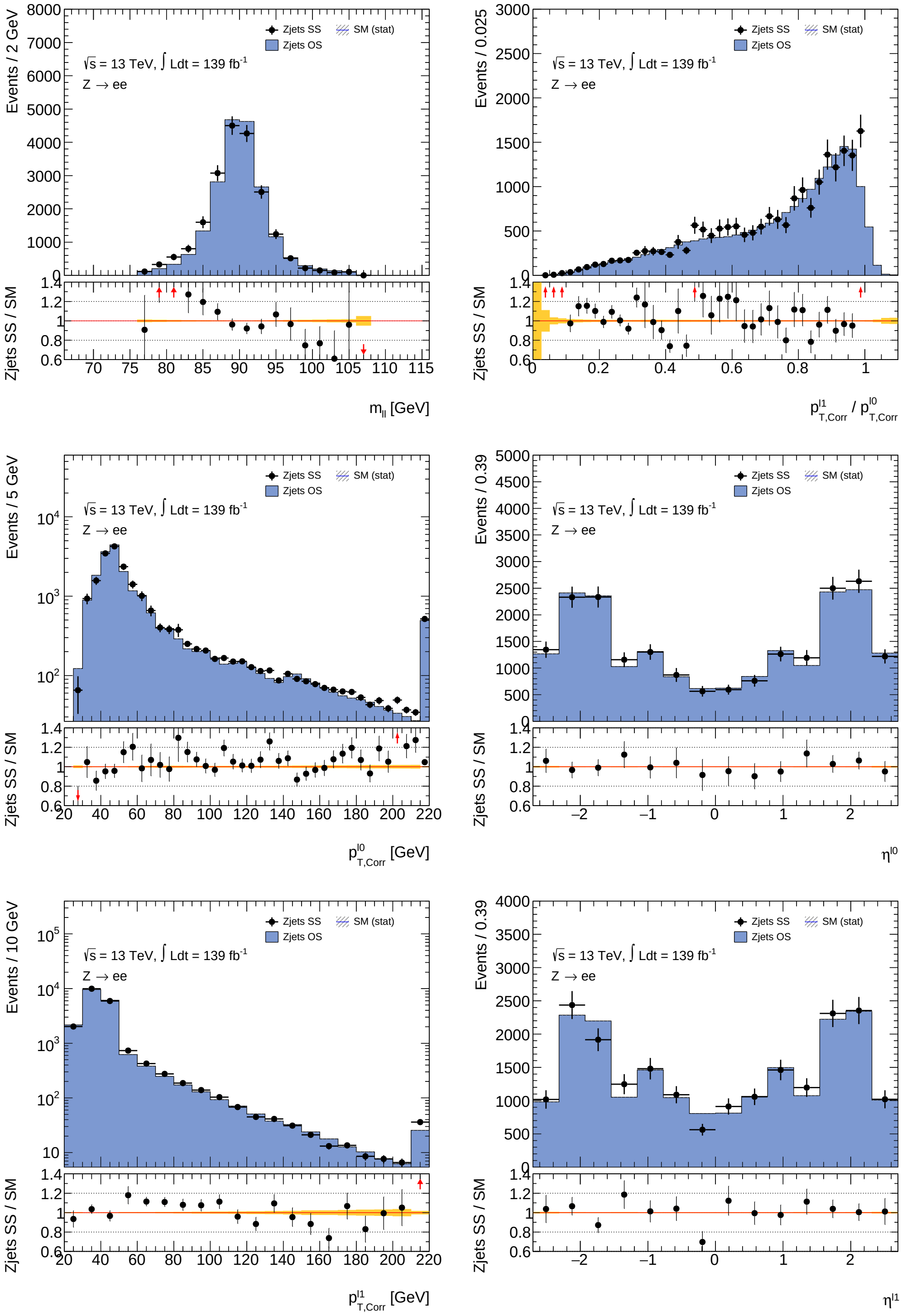


Figure 7.13: Kinematic variables for the closure test of the methodology presented in Section 7.1. Madgraph $Z \to ee$ MC samples are selected for full Run 2 in the Z window CR to have opposite-sign electrons in the final state (Zjets OS, shown in blue) on which the method is applied. The resulting yields are then compared to a event selection requiring same-sign leptons in the final state (Zjets SS, shown as black dots). On top, the invariant mass of the leading and subleading lepton (left), the reconstructed transverse momentum of the leading lepton measured after energy correction $p_{T,\mathrm{Corr}}^{\ell 0}$ (center) and the pseudorapidity distribution of the leading lepton $\eta^{\ell 0}$ (right) are shown. On the bottom, the ratio $p_T^{\ell 1}/p_T^{\ell 0}$ of corrected transverse momentum of subleading over leading lepton (left), the reconstructed transverse momentum of the subleading lepton measured after energy correction $p_{T,\mathrm{Corr}}^{\ell 1}$ (center) and the pseudorapidity distribution for the subleading lepton $\eta^{\ell 1}$ (right). Only statistical uncertainties are shown. The ratio SS/OS is computed for the ration plot.

ee channel	mc16a	mc16d	mc16e	run2
Data	**5389**	**6115**	**7952**	**19456**
ssWW EWK	7	8	11	26
ssWW QCD	3	3	4	11
top	3	3	4	10
WZ	96	114	147	357
ZZ	10	11	15	35
Vγ	164	186	136	486
FSR	609	797	1008	2413
Charge Flip	4538	5297	6922	16757
Total SM	**5429**	**6419**	**8247**	**20095**
(bkg-data)/data [%]	0.7	4.5	3.7	3.3

Table 7.7: Events obtained in the Z window CR for data and Standard Model expectations in the ee channel for different datasets recorded by the ATLAS experiment in years 2015-2016, 2017 and 2018, along with the full Run 2 dataset. The contributions due to charge misidentification are computed by the data-driven method presented in Section 7.1. The category *"top"* includes contributions from single-top production, $t\bar{t}$ and ttX after passing the selection at truth level. *"WZ"* includes QCD and EWK production along with $WqqZ\ell\ell$, and $W\gamma$ and $Z\gamma$ compose the category *"$V\gamma$"*. The FSR contribution is estimated using $Z \to ee$ MC by a truth requirement.

transverse momentum of subleading over leading lepton are presented in Figure 7.13. The use of the ECIDS selection of correctly reconstructed electrons makes the SS event selection to drop by a factor of 12 with respect to the closure test in Section 7.2.1 where the ECIDS tool was not applied. Consistently, the reduction in the probability of charge misidentification allows a correct estimation of the charge misidentification distributions using OS event selections. ECIDS also removes events from the OS selection, as can be seen from the raw amount of events when compared to 7.1. The kinematic distributions are in agreement with the closure test with no ECIDS tool applied within 20% between the two event selections. Integrated number of events agree along MC campaigns within statistical uncertainties. Their values are presented in Table 7.6.

The methodology can be tested for the estimation of charge misidentification in the $e\mu$ and μe channels by exploiting $t\bar{t}$ decays. The kinematic distributions from Powheg $t\bar{t}$ MC samples decaying into leptonic $e\mu$ and μe channels are compared in Figure 7.14. Once again, asymmetries are found between $e\mu$ and μe channels, especially for the transverse momentum of the muon (leading lepton in the μe channel and subleading lepton in the $e\mu$ channel). The ratio between SS and OS events selections seems to be dependent on the muon p_T, decreasing for greater values of this variable. This tendency is lost when the muon is tagged as the subleading lepton. These results were not observed when the ECIDS selection was disabled during the tests using $t\bar{t}$ MC, as presented in Figure 7.9. The reduced statistics for events when the charge misidentified electron is tagged as the subleading lepton does not allow to make strong conclusions on the differences observed between the energy-corrected ratio $p_T^{\ell_1}/p_T^{\ell_0}$. The flat tendency of the ratio present without the use of the ECIDS tool in the $e\mu$ channel is not present when the tool is applied. The $e\mu$ channel shows four times more events than the μe, since charge misidentification is more likely to occur at high electron transverse momenta. The ECIDS tool removes a factor of events of 7.8, 4.5 and 8.6 of the total yields between SS and OS event selections, that differ by $16.9 \pm 2.5\%$, $5.4 \pm 0.9\%$ and $15.2 \pm 1.5\%$ for the full Run 2 dataset in the ee, $e\mu$ and μe decay channels, respectively. The integral values for both selections along MC campaigns are provided in Appendix C.9.

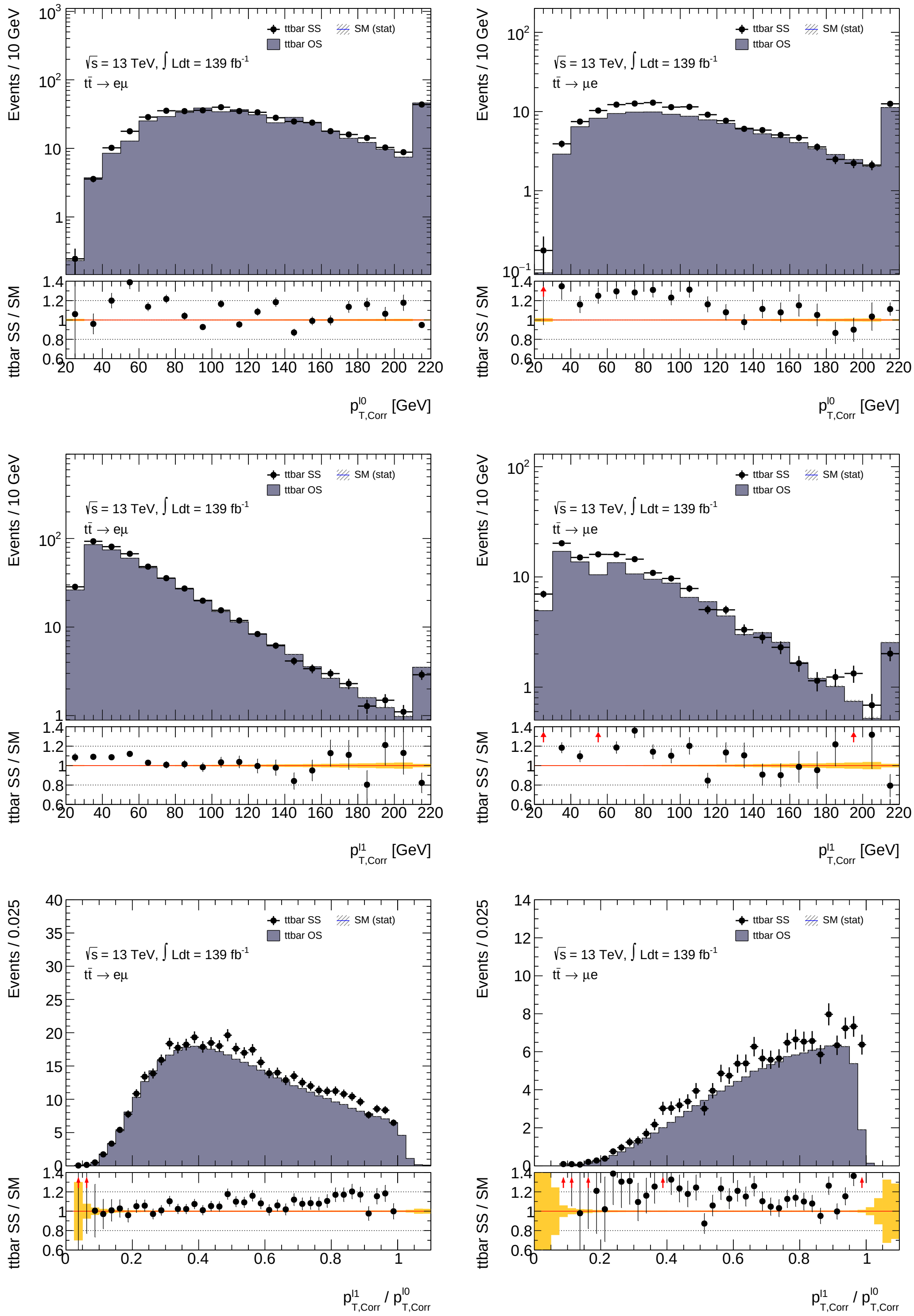


Figure 7.14: Kinematic distributions for SS and weighted and energy-corrected OS event selections in Powheg $t\bar{t}$ MC samples decaying into the $e\mu$ channel (left) and μe channel (right) for full Run 2 in the Z window CR. The p_T of the leading lepton (top), p_T of the subleading lepton (center), ratio $p_T^{\ell_1}/p_T^{\ell_0}$ of energy-corrected transverse momentum of subleading over leading leptons (top-right) and invariant mass of the two leptons $m_{\ell\ell}$ (bottom-right) are presented. The ratio SS/OS is computed for the ration plot.

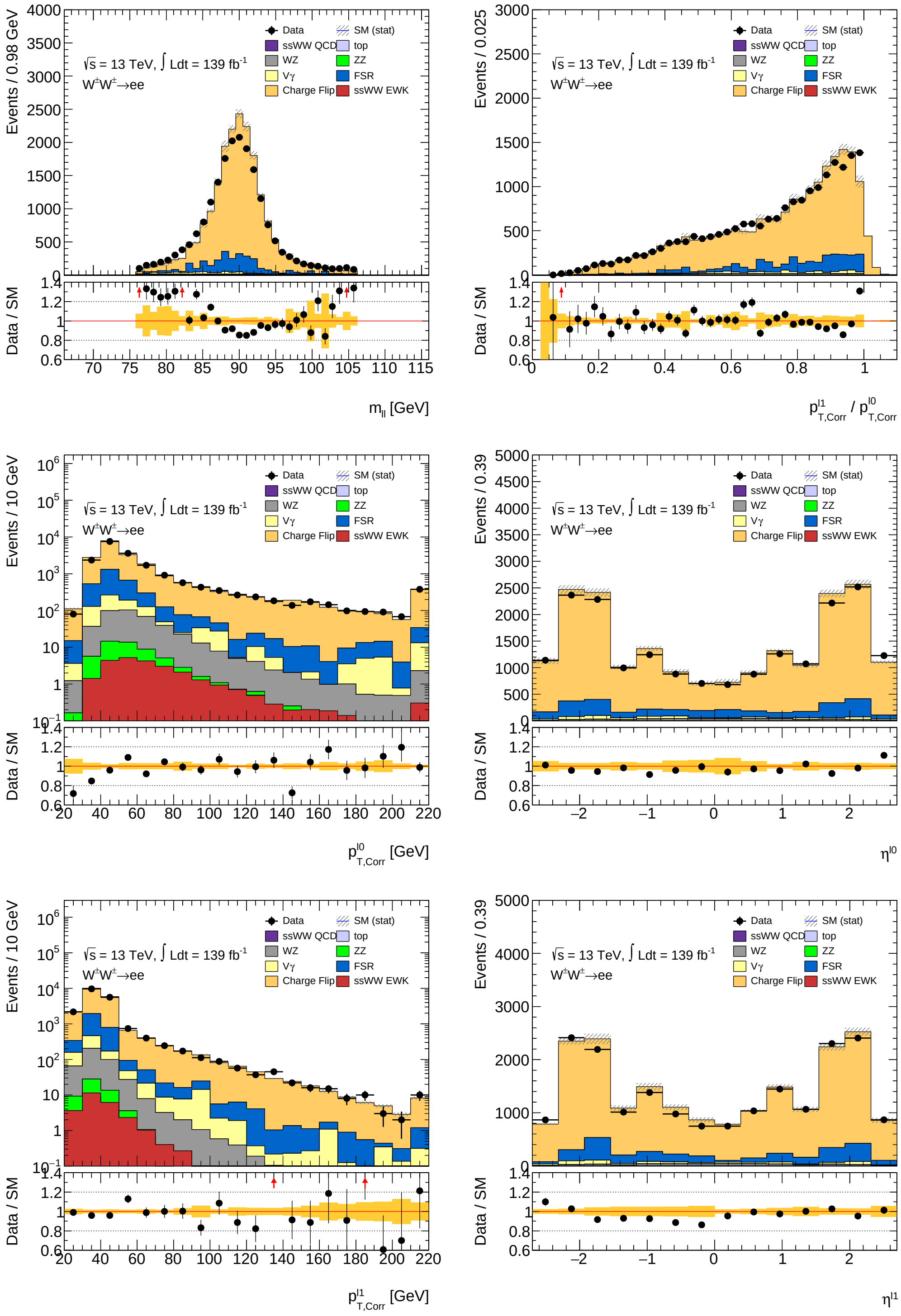


Figure 7.15: Expected yields from Standard Model sources along with data observation in the Z window CR for electrons in the ee channel satisfying the ECIDS selection for the full Run 2 dataset. A data-driven estimation is carried out for contributions due to charge misidentification. The category "*top*" includes contributions from single-top production, $t\bar{t}$ and ttX. "*WZ*" includes QCD and EWK production along with $WqqZ\ell\ell$, and $W\gamma$ and $Z\gamma$ compose the category "$V\gamma$". The FSR contribution is estimated using $Z \to ee$ MC by a truth requirement. The ratio Data/Total-SM is computed for the ration plot.

Lastly, the methodology is validated for the data-driven estimation using the Z window CR. The estimation is presented in Figure 7.15 for different kinematic distributions on leading (ℓ_0) and subleading leptons (ℓ_1), the reconstructed invariant mass and the ratio of energy-corrected transverse momenta $p_T^{\ell_1}/p_T^{\ell_0}$. Compared to the results presented in Section 7.2.1, slight improvements between data and Standard Model predictions are found after the ECIDS selection is required, especially in the $|\eta|$ and p_T distributions for leading and subleading leptons. Statistical uncertainties are notably larger and they arose for the constrained contribution for FSR-induced electrons. Kinematic distributions for data and SM expectations agree within 20% for $|m_{\ell\ell} - m_Z| < 10$ GeV. Nevertheless, deviations up to 30% can be observed at the edges of the Z window. Table 7.7 contains the observed data events and the Standard Model expectations for signal and background along different periods of data taking and the full Run 2 dataset. For the full Run 2 dataset, the background yields due to charge misidentification are 11 times lower after passing the ECIDS selection. The rest of Standard Model contributions also pass this selection. Events from top-quark sources (including ttX, single-top and $t\bar{t}$ events) are reduced by a factor of 2, Vγ by a factor of 5 and FSR events by a factor of 4. This reduction is expected for signatures from these background sources that mimic the bremsstrahlung radiation originating electron charge misidentification. Neither WZ nor same-sign WW (for both QCD and EWK production) suffer considerable reductions after the ECIDS tool is applied. Non-prompt background contributions are again not considered. These are expected to be negligible in the ee channel within the Z window CR. The observed amount of data events matches the Standard Model expectations within 5% along all data taking periods without considering this background source. Dilepton distributions show no remarkable differences after the ECIDS selection respect to distributions shown in Section 7.2.1.

The methodology presented in Section 7.1 for the data-driven estimation of background contributions due to electron charge misidentification is found to be valid after the ECIDS selection is required. This tool removes these contributions by a factor of 11 in the Z window CR. No notable reductions of the signal are observed. In the ee channel (where contributions due to charge misidentification constitute more than 80% of the total SM contributions), the kinematic distributions agree within 20% in most of the CR for the full Run 2 dataset, with larger deviations up to 30% being observed at large values of $|m_{\ell\ell} - m_Z|$. The confirmation of the method allows to propagate the background estimation originated by charge misidentified electrons to the Signal Region. The results will be exposed in Section 7.3.2.

7.3.2 Background estimation in the signal region using ECIDS

For the full Run 2 dataset, the data-driven estimation validated in Section 7.3.1 is propagated to the SR after applying the ECIDS selection. The arising contributions are presented in Table 7.8 along with systematic variations. New charge misidentification probabilities are calculated for the systematic variations of the SF. The full set of distributions is shown in Appendix C.8. When compared to the results presented in Table 7.3 (where the ECIDS selection is disabled), the overall background yields due to charge misidentification in the all-inclusive channel ($\ell\ell$) is reduced by a factor of 4 in the full Run 2 dataset. This reduction is weaker (factor of 3) in the ee channel. The total systematic uncertainties increase from 23% up to 62% and from 33% up to 93% of the nominal value for the

139 fb^{-1}	ee	$e\mu$	μe	$\ell\ell$
Charge flip (Nom $\pm$ stat)	3.63 $\pm$ 0.11	5.5 $\pm$ 0.3	1.25 $\pm$ 0.08	10.3 $\pm$ 0.4
SF up	+4.2	+2.8	+0.9	+7.8
SF down	-2.5	-2.3	-0.7	-5.5
$dE \times 0$	-0.009	0.0	-0.0002	-0.009
$dE \times 2$	+0.013	+0.001	-0.0008	+0.012
total sys. variation	$\pm$ 3.4	$\pm$ 2.5	$\pm$ 0.8	$\pm$ 6.7
stat $\oplus$ sys	$\pm$ 3.4	$\pm$ 2.6	$\pm$ 0.8	$\pm$ 6.7

139 fb^{-1} ($\lvert\eta_{ee}\rvert < 2.5$)	ee	$e\mu$	μe	$\ell\ell$
Charge flip (Nom $\pm$ stat)	19.5 $\pm$ 0.7	5.5 $\pm$ 0.3	1.25 $\pm$ 0.08	26.2 $\pm$ 0.8
SF up	+10.0	+2.8	+0.8	+13.6
SF down	-7.9	-2.3	-0.7	-10.9
$dE \times 0$	-0.024	0.0	-0.0002	-0.025
$dE \times 2$	-0.010	+0.0006	-0.0008	-0.011
total sys. variation	$\pm$ 8.9	$\pm$ 2.5	$\pm$ 0.8	$\pm$ 12.2
stat $\oplus$ sys	$\pm$ 9.0	$\pm$ 2.6	$\pm$ 0.8	$\pm$ 12.3

Table 7.8: Background yields due to charge misidentification and systematic uncertainties in the SR using ECIDS in the full Run 2 dataset.

$\ell\ell$ and ee channels, respectively. Systematic variations are more strongly dominated by the SF uncertainty than in the analysis without ECIDS selection. Statistical uncertainties constitute about 3% and 6% of the total yield uncertainty for the full Run 2 dataset in the ee and $\ell\ell$ channels, respectively. Systematic variations on the energy correction smearing factor dE lie within statistical fluctuations and they do not influence the estimated background yield in the Signal Region. Variations on dE are practically inexistent for $e\mu$ and μe channels, where only the energy of one lepton is varied. The event kinematics is not strongly notably deviated for these cases and the same selection criteria are passed as for the nominal background yield. The relaxed $|\eta|$ cut in the SR is considered at the bottom table. The estimated background yields arising from charge misidentification grow by a factor of 2.4 in the $\ell\ell$ channel, being the yield 5.1 times larger in the ee channel. As expected, $e\mu$ and μe remain unchanged after this relaxation. The increase in statistics does not provide relevant systematic variations from energy smearing. Nevertheless, this increase reduces the proportion on the deviations from SF uncertainties up to 45% for both ee and $\ell\ell$ channels. After the relaxation of $|\eta_{ee}|$, 74% of the yields due to charge misidentification in the $\ell\ell$ channel are produced from decays into two electrons, the same proportion found where the ECIDS tool was disabled.

The statistical significances can be computed for the 2015-2016 dataset, as given by Eq. (7.7). The results using the ECIDS selection are presented in Table 7.9. Signal events, as observed in the Z window CR, are not strongly affected by the ECIDS tool selection. A similar reduction of events is observed in data and the estimated yields due to charge misidentification for ee and $\ell\ell$ channels. The ECIDS selection of correctly charge-identified electrons increases the statistical significance of the signal, for the $\ell\ell$ channel, by 0.3σ in the Signal Region. Only 0.1σ is increased in the ee channel. This growth of Z is twice larger for $e\mu$ and μe channels.

The suppression of the charge misidentification by ECIDS stops the growth of its background yields after relaxing the $|\eta|$ cut. While the number of events arising from charge misidentification in the SR increased by a factor of 7.6 in the ee channel when the

mc16a ($|\eta_{ee}| < 1.37$)	ee	$e\mu$	μe	$\ell\ell$
Data	14	40	38	136
Signal	5.7 ± 0.3	12.5 ± 0.5	12.4 ± 0.5	45.1 ± 0.8
Charge flip ($\pm$ stat $\oplus$ sys)	0.9 ± 0.8	1.7 ± 0.8	0.32 ± 0.19	2.9 ± 1.7
Z	1.6	2.1	2.2	4.1

mc16a ($|\eta_{ee}| < 2.5$)	ee	$e\mu$	μe	$\ell\ell$
Data	36	40	38	158
Signal	9.0 ± 0.4	12.5 ± 0.5	12.4 ± 0.5	48.5 ± 0.9
Charge flip ($\pm$ stat $\oplus$ sys)	5 ± 3	1.7 ± 0.8	0.32 ± 0.19	7 ± 3
Z	1.3	2.1	2.2	3.8

Table 7.9: Signal significances using 36 fb^{-1} for relaxation of the $|\eta_{ee}|$ cut for the ee channel in the SR after applying the ECIDS selection..

ECIDS selection was not applied, a factor 5.5 more events are observed after the ECIDS selection. The same signal events are observed with and without ECIDS. In the all-inclusive channel, 4.4 more events were observed after the relaxation when ECIDS was not applied. Indeed, the estimated contribution due to charge misidentification after the relaxation is lower than the contribution of this background in the SR when the tool is disabled. This fact is also observed by the similar values of Z. Since the yields in the μe and $e\mu$ channels are not modified, the differences can also be observed in the weaker reduction of Z with respect to the no-ECIDS configuration. For the use of ECIDS, only a decrease of 0.3σ is observed for the $\ell\ell$ channel, even the uncertainties of the charge misidentification contribution were proportionally larger (up to 90%) from the systematic variations of the ECIDS scale factors.

The reduction of charge misidentification background by ECIDS makes the relaxation of the $|\eta|$ cut affordable in terms of reduction of signal significance to pave the way for a improved unfolding of the final measurement. Further relaxation of the signal region could be explored. A lower transverse momentum threshold of $p_T > 20$ GeV could be assessed for a reduced contribution of background yields due to charge misidentification. However, non-prompt contributions rise remarkably in this region [91]. Consequently, the relaxation in p_T is not considered in this work.

7.4 Outlook

The estimation of the background contributions due to charge misidentification presented in this work can be improved from different perspectives.

In Section 7.2.1, the test on the used methodology provided different results for $Z \to ee$ and $t\bar{t}$ MC. The observed underestimation for $t\bar{t}$ when using the charge misidentification probabilities computed in $Z \to ee$ might hint to possible biases in the methodology, introducing a dependency on the used physical process. These discrepancies are introduced from the consideration of the reconstructed p_T for the computation of the charge misidentification probabilities, according to studies carried out by the Egamma CP group [96]. The energy responses coming into numerator and denominator of Eq.7.2 differ for wrongly and correctly reconstructed electrons, introducing possible biases into the computation of these distributions. Instead, the truth transverse momentum can be exploited. The probabilities calculated in $Z \to ee$ and $t\bar{t}$ agree after this consideration. Since the p_T^{Truth} is not defined in data (where the method is applied) a different approach is needed for the energy correction. The energy responses are studied in MC for both correctly and reconstructed electrons in different η bins. The value of p_T^{Truth} is then inferred from Crystal-Ball fits on these distributions from the value of the reconstructed p_T. Furthermore, the corrected p_T to be applied on electrons from the OS event selection in the data-driven estimation exploit these fits by applying a convolution from the energy response of correctly reconstructed electrons to reproduce a energy response for wrongly identified electron charges. $t\bar{t}$ tests are claimed to agree using $Z \to ee$ charge misidentification probabilities after these considerations [96].

The reduced statistics after the ECIDS selection increases the statistical fluctuations in the distributions of charge misidentification probabilities. These fluctuations are especially strong at low p_T and $|\eta|$. Results can be found in Appendix C.7. To avoid these fluctuations to be driven into the background estimation, the binning could be optimized accordingly. In this thesis, the same binning as the distributions when no ECIDS selection is applied were used for the full Run 2 dataset, showing a good reliability in the method validation. Nevertheless, further improvements might be found after this optimization.

The scale factors used for the ECIDS selection involve large systematic uncertainties. This increase with respect to the scale factors presented in Section 7.1.1 appears due to dependencies on the charge-misidentified electron rejection with the charge sign along with asymmetries in the η distribution [68]. These inefficiencies are stronger at high η, where larger yields due to charge misidentification are expected. Deviations between charge misID rates for electrons and positrons are accounted as a new source of systematic uncertainty for the scale factors using the ECIDS selection, added on top of the sources described in Section 7.2.2. Understanding these asymmetries could improve considerably the statistical significance of the measurement. Increases of about 0.2σ are expected from the computation of Eq 7.7 considering an uncertainty of the nominal value of 30% closer to the uncertainties when the ECIDS selection was disabled, especially in the ee channel where the background is dominant.

Further strategies for the reduction of charge misidentification contributions can also be explored by either applying additional cuts, the development of more sophisticated algorithms or both at the same time. Selection optimization of variables such as the ratio E/p seems promising according to the results presented in Chapter 6. Even though the dynamical, topological cell clustering-based approach for the energy reconstruction was not explored in this thesis, the geometry and signature of the different topocluster for a same electron candidate could shed some light into new variables for rejection of charge misidentified electrons. The resulting reduction in statistics seems to be affordable for analyses using the full Run 2 dataset, as shown in this thesis for the computation of charge misidentification probabilities after the ECIDS selection.

Chapter 8

Conclusions

The full Run 2 dataset, corresponding to an integrated luminosity of 139 fb^{-1} of proton–proton collisions recorded by the ATLAS experiment at a center-of-mass energy of $\sqrt{s} = 13$ TeV allows carrying out precision measurements of vector boson scattering, in particular same-sign W boson pairs. Inefficiencies in the electron reconstruction algorithms might cause events with opposite-sign leptons satisfy the selectio criteria after experiencing charge misidentification. The effect is important for electrons, while it was proved to be negligible for muons. In this thesis, the contribution of charge misidentification background is estimated for the measurement of same-sign W boson pairs.

In order to understand the background contribution, the sources of charge misidentification are studied using $Z \to ee$ MC events. The different electron categories are identified exploiting the information at generation, detector simulation and reconstruction levels. Charge misidentification is found to be a rare process, arising in 0.3% of the considered electron candidates. Two main sources of charge misID are found, coinciding with previous studies carried out by the ATLAS Collaboration [92]. The dominant source is induced by bremsstrahlung radiation with subsequent photon conversions, arising in 64.5% of the charge-misidentified electrons. This is caused by interactions with the material budget of the inner tracker of the ATLAS detector. Therefore, electrons induced by this source are found within $|\eta| > 1.52$, agreeing with the distribution of the inner detector material. Due to the energy leakage produced by the subsequent radiation of photons, the energy resolution for these electrons ($p_T^{\text{Reco}}/p_T^{\text{Truth}} - 1$) is smeared to negative values. Due to this energy loss, the reconstruction of the invariant mass of the two electrons is also shifted to lower values when at least one electron induced by bremsstrahlung comes into the computation. n this charge flip category, more than one track is assigned to electron candidates in approximately 90% of cases, hinting that the best-matching track for the electron candidate was wrongly selected. This selected track does not match the primary vertex, as observed on the d_0 significance. The second source is found in 35.4% of the charge-misidentified electrons and it is produced by straight-like tracks of high-p_T electrons, when the tracking system has difficulties to resolve the track curvature. The orientation of the track bending can be then mismeasured, providing the wrong electric charge for the electron candidate. One track is assigned to 40% of these electron candidates, with its extrapolation also disagreeing with the position of the primary vertex. The electron energy retrieved by this track is underestimated when compared to the measurement in calorimetry. Since these electrons are mostly found in the most-forward region of the inner detector ($2.3 < |\eta| < 2.5$), no TRT hits are found in 56% of the cases, showing a poorer fit quality. Additionally, particle tracks in this region of the inner detector are bent to a smaller extent than in the central region since the electron travels parallel to the orientation of the magnetic field. This source is preferentially found on the leading lepton of the event in 60% of the cases where this source is found. FSR-induced electrons (found in 0.04% of the studied electron candidates) are not considered as a source of charge misidentification, since it is originated at generator level and no detector effects are involved in its production.

For the data-driven estimation of the charge misidentification background, $Z \rightarrow ee$ MC events are used to determine the probability of charge misidentification as a function of p_T and $|\eta|$. For this computation, FSR-induced electrons are removed at truth level. To correct for the discrepancies between these distributions obtained from data and MC, scale factors are applied. Statistical uncertainties in different parameterizations of these scale factors, especially at high p_T, dominate the systematic uncertainties reaching 15% of the nominal value. The charge misidentification probabilities are then used to estimate the background contributions by weighting an opposite-sign event selection to estimate the charge misID background yields in same-sign final states. In this way, statistical uncertainties are notoriously reduced for the estimated background yields. To consistently represent the worsened kinematic properties of charge-misidentified electrons, energy corrections are applied to take the observed energy leakage into account. An energy scale correction and smearing factors are applied taking into account the energy resolution of correctly and wrongly reconstructed electrons into account. The energy correction is applied on a randomly chosen electron in the event, with their charge misidentification probabilities are accounted for the decision. The methodology is validated in a Z-peak control region with a closure test using $Z \rightarrow ee$ MC events, with the differential distributions of opposite-sign and same-sign selections agreeing within 20%. The integrated number of events are found to be consistent within 3%. A data-driven estimation is also carried out in the ee channel for the Total-SM expected yields and data, finding a similar agreement. The background is then propagated to the signal region, obtaining 43.3 ± 1.1 (stat) $\pm$ 10.2 (sys) events in the $\ell\ell$ channel. Systematic uncertainties are dominated by the statistical fluctuations in the computation of the scale factors.

The background is then reduced by applying the *Electron Charge ID Selector* (ECIDS), a tool developed by the *Egamma combined performance group* based on a boosted decision tree discriminator between electrons with correctly (signal) and wrongly (background) reconstructed charges. After the selection of electrons through this algorithm, the methodology is reproduced. A similar agreement within 20% is found in the closure test using $Z \rightarrow ee$ MC. Considering the integrated number of events, the ECIDS selector reduces the charge misID background contribution by a factor of 11 in the control region, which causes an important increase of the statistical uncertainties. A similar agreement within 20% between data and total-SM expectations is found for the differential distributions in the Z-peak control region using the data-driven method. The contributions in the signal region due to charge-misidentification are reduced by a factor of 4 when using the ECIDS, with a contribution of 10.3 ± 0.4 (stat) $\pm$ 6.7 (sys) events in the $\ell\ell$ channel. A reduction of 1% is found for the signal events.

The reduction of the background contribution using ECIDS offers a possibility for extension of the electron pseudorapidity range in the ee channel to the almost full acceptance of the inner tracking detector ($|\eta| < 2.47$). This relaxation of the selection criteria also allows having a unified selection with other decay channels and to increase the signal yields. A signal significance is used to quantify the feasibility of this relaxation, considering also the uncertainties of the charge misidentification background estimation. In this case, a background yield of 26.2$\pm$0.8 (stat)$\pm$12.2 (sys) events is obtained in the $\ell\ell$ channel, with a factor of 1.6 of signal increase in the ee channel using the full Run 2 dataset.

Appendix A

Technical details on the dataset and event selection

A.1 List of triggers

Data period	Electron trigger
2015	HLT_e24_lhmedium_L1EM20VH, HLT_e60_lhmedium, HLT_e120_lhloose
2016	HLT_e26_lhtight_nod0_ivarloose, HLT_e60_lhmedium_nod0, HLT_e140_lhloose_nod0
2017	HLT_e26_lhtight_nod0_ivarloose, HLT_e60_lhmedium_nod0, HLT_e140_lhloose_nod0
2018	HLT_e26_lhtight_nod0_ivarloose, HLT_e60_lhmedium_nod0, HLT_e140_lhloose_nod0

Data period	Muon trigger
2015	HLT_mu20_iloose_L1MU15, HLT_mu50
2016	HLT_mu26_ivarmedium, HLT_mu50
2017	HLT_mu26_ivarmedium, HLT_mu50
2018	HLT_mu26_ivarmedium, HLT_mu50

Table A.1: Single-lepton triggers applied for electrons and muons along different data-taking periods.

A.2 List of samples

For the ATLAS datasets, a convention is followed to identify the samples

$$\underbrace{\texttt{data17\_13TeV}}_{\text{Project name}}\underbrace{\texttt{.00339849}}_{\text{Run number}}\underbrace{\texttt{.physics\_Main}}_{\text{Stream type}}\underbrace{\texttt{.deriv}}_{\text{Production step}}\underbrace{\texttt{.DAOD\_TOPQ4.}}_{\text{Data type}}\underbrace{\texttt{.889\_m1902\_p3402}}_{\text{Production tags}}$$

The naming of the MC samples follows the following convention

$$\underbrace{\texttt{mc16\_13TeV}}_{\text{Project name}}\texttt{.}\underbrace{\texttt{361510}}_{\text{Dataset identifier (MCID)}}\texttt{.}\underbrace{\texttt{MadGraphPythia8EvtGen}}_{\text{Event generator(s)}}\texttt{\_}\underbrace{\texttt{A14}}_{\text{tune}}\underbrace{\texttt{NNPDF23LO}}_{\text{PDF set}}\texttt{\_}\underbrace{\texttt{Ztautau\_Np0}}_{\text{Physics process}}$$

$$\underbrace{\texttt{.deriv.DAOD}}_{\text{Data format}}\texttt{\_}\underbrace{\texttt{STDM3}}_{\text{Derivation flavor}}\texttt{.}\underbrace{\texttt{e3898\_s3126\_r9364\_r9315\_p3975}}_{\text{Production tags}}$$

Data

data15_13TeV.*.physics_Main.deriv.DAOD_STDM3*p3896
data16_13TeV.*.physics_Main.deriv.DAOD_STDM3*p3896
data17_13TeV.*.physics_Main.deriv.DAOD_STDM3*p3896
data18_13TeV.*.physics_Main.deriv.DAOD_STDM3*p3896

mc16a

mc16_13TeV.363147.MGPy8EG_N30NLO_Zee_Ht0_70_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363148.MGPy8EG_N30NLO_Zee_Ht0_70_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363149.MGPy8EG_N30NLO_Zee_Ht0_70_BFilter.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363150.MGPy8EG_N30NLO_Zee_Ht70_140_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315
mc16_13TeV.363151.MGPy8EG_N30NLO_Zee_Ht70_140_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363152.MGPy8EG_N30NLO_Zee_Ht70_140_BFilter.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363153.MGPy8EG_N30NLO_Zee_Ht140_280_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363154.MGPy8EG_N30NLO_Zee_Ht140_280_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363155.MGPy8EG_N30NLO_Zee_Ht140_280_BFilter.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363156.MGPy8EG_N30NLO_Zee_Ht280_500_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363157.MGPy8EG_N30NLO_Zee_Ht280_500_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363158.MGPy8EG_N30NLO_Zee_Ht280_500_BFilter.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363159.MGPy8EG_N30NLO_Zee_Ht500_700_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363160.MGPy8EG_N30NLO_Zee_Ht500_700_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363161.MGPy8EG_N30NLO_Zee_Ht500_700_BFilter.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363162.MGPy8EG_N30NLO_Zee_Ht700_1000_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363163.MGPy8EG_N30NLO_Zee_Ht700_1000_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363164.MGPy8EG_N30NLO_Zee_Ht700_1000_BFilter.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363165.MGPy8EG_N30NLO_Zee_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363166.MGPy8EG_N30NLO_Zee_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363167.MGPy8EG_N30NLO_Zee_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363168.MGPy8EG_N30NLO_Zee_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363169.MGPy8EG_N30NLO_Zee_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363170.MGPy8EG_N30NLO_Zee_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e4866_s3126_r9364_r9315_p3975
mc16_13TeV.363356.Sherpa_221_NNPDF30NNLO_ZqqZll.deriv.DAOD_STDM3.e5525_s3126_r9364_r9315_p3975
mc16_13TeV.364250.Sherpa_222_NNPDF30NNLO_llll.deriv.DAOD_STDM3.e5894_s3126_r9364_r9315_p3975
mc16_13TeV.364253.Sherpa_222_NNPDF30NNLO_lllv.deriv.DAOD_STDM3.e5916_s3126_r9364_r9315_p3975
mc16_13TeV.364254.Sherpa_222_NNPDF30NNLO_llvv.deriv.DAOD_STDM3.e5916_s3126_r9364_r9315_p3975
mc16_13TeV.364283.Sherpa_222_NNPDF30NNLO_lllljj_EW6.deriv.DAOD_STDM3.e6055_e5984_s3126_r9364_r9315_p3975
mc16_13TeV.364284.Sherpa_222_NNPDF30NNLO_lllvjj_EW6.deriv.DAOD_STDM3.e6055_e5984_s3126_r9364_r9315_p3975
mc16_13TeV.364286.Sherpa_222_NNPDF30NNLO_llvvjj_ss_EW4.deriv.DAOD_STDM3.e6055_e5984_s3126_r9364_r9315_p3975
mc16_13TeV.364287.Sherpa_222_NNPDF30NNLO_llvvjj_ss_EW6.deriv.DAOD_STDM3.e6055_e5984_s3126_r9364_r9315_p3975
mc16_13TeV.364500.Sherpa_222_NNPDF30NNLO_eegamma_pty_7_15.deriv.DAOD_STDM3.e5928_e5984_s3126_r9364_r9315_p3970
mc16_13TeV.364501.Sherpa_222_NNPDF30NNLO_eegamma_pty_15_35.deriv.DAOD_STDM3.e5928_e5984_s3126_r9364_r9315_p3970
mc16_13TeV.364502.Sherpa_222_NNPDF30NNLO_eegamma_pty_35_70.deriv.DAOD_STDM3.e5928_e5984_s3126_r9364_r9315_p3970
mc16_13TeV.364503.Sherpa_222_NNPDF30NNLO_eegamma_pty_70_140.deriv.DAOD_STDM3.e5928_e5984_s3126_r9364_r9315_p3970
mc16_13TeV.364504.Sherpa_222_NNPDF30NNLO_eegamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_e5984_s3126_r9364_r9315_p3970
mc16_13TeV.364505.Sherpa_222_NNPDF30NNLO_mumugamma_pty_7_15.deriv.DAOD_STDM3.e5928_e5984_s3126_r9364_r9315_p3970
mc16_13TeV.364506.Sherpa_222_NNPDF30NNLO_mumugamma_pty_15_35.deriv.DAOD_STDM3.e5988_e5984_s3126_r9364_r9315_p3970
mc16_13TeV.364507.Sherpa_222_NNPDF30NNLO_mumugamma_pty_35_70.deriv.DAOD_STDM3.e5928_e5984_s3126_r9364_r9315_p3970
mc16_13TeV.364508.Sherpa_222_NNPDF30NNLO_mumugamma_pty_70_140.deriv.DAOD_STDM3.e5928_e5984_s3126_r9364_r9315_p3970
mc16_13TeV.364509.Sherpa_222_NNPDF30NNLO_mumugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_e5984_s3126_r9364_r9315_p3970
mc16_13TeV.364510.Sherpa_222_NNPDF30NNLO_tautaugamma_pty_7_15.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3970
mc16_13TeV.364511.Sherpa_222_NNPDF30NNLO_tautaugamma_pty_15_35.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3970
mc16_13TeV.364512.Sherpa_222_NNPDF30NNLO_tautaugamma_pty_35_70.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3970
mc16_13TeV.364513.Sherpa_222_NNPDF30NNLO_tautaugamma_pty_70_140.deriv.DAOD_STDM3.e5982_s3126_r9364_r9315_p3970
mc16_13TeV.364514.Sherpa_222_NNPDF30NNLO_tautaugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3970
mc16_13TeV.364521.Sherpa_222_NNPDF30NNLO_enugamma_pty_7_15.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364522.Sherpa_222_NNPDF30NNLO_enugamma_pty_15_35.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364523.Sherpa_222_NNPDF30NNLO_enugamma_pty_35_70.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364524.Sherpa_222_NNPDF30NNLO_enugamma_pty_70_140.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364525.Sherpa_222_NNPDF30NNLO_enugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364526.Sherpa_222_NNPDF30NNLO_munugamma_pty_7_15.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364527.Sherpa_222_NNPDF30NNLO_munugamma_pty_15_35.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364528.Sherpa_222_NNPDF30NNLO_munugamma_pty_35_70.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364529.Sherpa_222_NNPDF30NNLO_munugamma_pty_70_140.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364530.Sherpa_222_NNPDF30NNLO_munugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364531.Sherpa_222_NNPDF30NNLO_taunugamma_pty_7_15.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364532.Sherpa_222_NNPDF30NNLO_taunugamma_pty_15_35.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364533.Sherpa_222_NNPDF30NNLO_taunugamma_pty_35_70.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364534.Sherpa_222_NNPDF30NNLO_taunugamma_pty_70_140.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364535.Sherpa_222_NNPDF30NNLO_taunugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_s3126_r9364_r9315_p3975
mc16_13TeV.364739.MGPy8EG_NNPDF30NLO_A14NNPDF23LO_lvlljjEW6_OFMinus.deriv.DAOD_STDM3.e7421_e5984_s3126_r9364_r9315_p3975
mc16_13TeV.364740.MGPy8EG_NNPDF30NLO_A14NNPDF23LO_lvlljjEW6_OFPlus.deriv.DAOD_STDM3.e7421_e5984_s3126_r9364_r9315_p3975
mc16_13TeV.364741.MGPy8EG_NNPDF30NLO_A14NNPDF23LO_lvlljjEW6_SFMinus.deriv.DAOD_STDM3.e7421_e5984_s3126_r9364_r9315_p3975
mc16_13TeV.364742.MGPy8EG_NNPDF30NLO_A14NNPDF23LO_lvlljjEW6_SFPlus.deriv.DAOD_STDM3.e7421_e5984_s3126_r9364_r9315_p3975
mc16_13TeV.410155.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttW.deriv.DAOD_STDM3.e5070_s3126_r9364_r9315_p3975
mc16_13TeV.410218.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttee.deriv.DAOD_STDM3.e5070_s3126_r9364_r9315_p3975
mc16_13TeV.410219.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttmumu.deriv.DAOD_STDM3.e5070_s3126_r9364_r9315_p3975
mc16_13TeV.410220.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_tttautau.deriv.DAOD_STDM3.e5070_s3126_r9364_r9315_p3975
mc16_13TeV.410470.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad.deriv.DAOD_STDM3.e6337_e5984_s3126_r9364_r9315_p3975
mc16_13TeV.410472.PhPy8EG_A14_ttbar_hdamp258p75_dil.deriv.DAOD_STDM3.e6348_e5984_s3126_r9364_r9315_p3975
mc16_13TeV.410560.MadGraphPythia8EvtGen_A14_tZ_4fl_tchan_noAllHad.deriv.DAOD_STDM3.e5803_s3126_r9364_r9315_p3975
mc16_13TeV.410644.PowhegPythia8EvtGen_A14_singletop_schan_lept_top.deriv.DAOD_STDM3.e6527_s3126_r9364_p3975
mc16_13TeV.410645.PowhegPythia8EvtGen_A14_singletop_schan_lept_antitop.deriv.DAOD_STDM3.e6527_s3126_r9364_p3975
mc16_13TeV.410648.PowhegPythia8EvtGen_A14_Wt_DR_dilepton_top.deriv.DAOD_STDM3.e6615_e5984_s3126_r9364_r9315_p3975
mc16_13TeV.410649.PowhegPythia8EvtGen_A14_Wt_DR_dilepton_antitop.deriv.DAOD_STDM3.e6615_e5984_s3126_r9364_r9315_p3975
mc16_13TeV.410658.PhPy8EG_A14_tchan_BW50_lept_top.deriv.DAOD_STDM3.e6671_e5984_s3126_r9364_r9315_p3975

mc16_13TeV.410659.PhPy8EG_A14_tchan_BW50_lept_antitop.deriv.DAOD_STDM3.e6671_e5984_s3126_r9364_r9315_p3975

mc16d

mc16_13TeV.363147.MGPy8EG_N30NLO_Zee_Ht0_70_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363148.MGPy8EG_N30NLO_Zee_Ht0_70_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363149.MGPy8EG_N30NLO_Zee_Ht0_70_BFilter.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363150.MGPy8EG_N30NLO_Zee_Ht70_140_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363151.MGPy8EG_N30NLO_Zee_Ht70_140_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363152.MGPy8EG_N30NLO_Zee_Ht70_140_BFilter.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363153.MGPy8EG_N30NLO_Zee_Ht140_280_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363154.MGPy8EG_N30NLO_Zee_Ht140_280_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363155.MGPy8EG_N30NLO_Zee_Ht140_280_BFilter.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363156.MGPy8EG_N30NLO_Zee_Ht280_500_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363157.MGPy8EG_N30NLO_Zee_Ht280_500_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363158.MGPy8EG_N30NLO_Zee_Ht280_500_BFilter.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363159.MGPy8EG_N30NLO_Zee_Ht500_700_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363160.MGPy8EG_N30NLO_Zee_Ht500_700_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363161.MGPy8EG_N30NLO_Zee_Ht500_700_BFilter.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363162.MGPy8EG_N30NLO_Zee_Ht700_1000_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363163.MGPy8EG_N30NLO_Zee_Ht700_1000_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363164.MGPy8EG_N30NLO_Zee_Ht700_1000_BFilter.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363165.MGPy8EG_N30NLO_Zee_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363166.MGPy8EG_N30NLO_Zee_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363167.MGPy8EG_N30NLO_Zee_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363168.MGPy8EG_N30NLO_Zee_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363169.MGPy8EG_N30NLO_Zee_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363170.MGPy8EG_N30NLO_Zee_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e4866_s3126_r10201_r10210_p3975
mc16_13TeV.363356.Sherpa_221_NNPDF30NNLO_ZqqZll.deriv.DAOD_STDM3.e5525_s3126_r10201_r10210_p3975
mc16_13TeV.364250.Sherpa_222_NNPDF30NNLO_llll.deriv.DAOD_STDM3.e5894_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364253.Sherpa_222_NNPDF30NNLO_lllv.deriv.DAOD_STDM3.e5916_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364254.Sherpa_222_NNPDF30NNLO_llvv.deriv.DAOD_STDM3.e5916_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364283.Sherpa_222_NNPDF30NNLO_lllljj_EW6.deriv.DAOD_STDM3.e6055_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364284.Sherpa_222_NNPDF30NNLO_lllvjj_EW6.deriv.DAOD_STDM3.e6055_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364286.Sherpa_222_NNPDF30NNLO_llvvjj_ss_EW4.deriv.DAOD_STDM3.e6055_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364287.Sherpa_222_NNPDF30NNLO_llvvjj_ss_EW6.deriv.DAOD_STDM3.e6055_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364500.Sherpa_222_NNPDF30NNLO_eegamma_pty_7_15.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364501.Sherpa_222_NNPDF30NNLO_eegamma_pty_15_35.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364502.Sherpa_222_NNPDF30NNLO_eegamma_pty_35_70.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364503.Sherpa_222_NNPDF30NNLO_eegamma_pty_70_140.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364504.Sherpa_222_NNPDF30NNLO_eegamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364505.Sherpa_222_NNPDF30NNLO_mumugamma_pty_7_15.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364506.Sherpa_222_NNPDF30NNLO_mumugamma_pty_15_35.deriv.DAOD_STDM3.e5928_s3126_r10201_p3970
mc16_13TeV.364507.Sherpa_222_NNPDF30NNLO_mumugamma_pty_35_70.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364508.Sherpa_222_NNPDF30NNLO_mumugamma_pty_70_140.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364509.Sherpa_222_NNPDF30NNLO_mumugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364510.Sherpa_222_NNPDF30NNLO_tautaugamma_pty_7_15.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364511.Sherpa_222_NNPDF30NNLO_tautaugamma_pty_15_35.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364512.Sherpa_222_NNPDF30NNLO_tautaugamma_pty_35_70.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364513.Sherpa_222_NNPDF30NNLO_tautaugamma_pty_70_140.deriv.DAOD_STDM3.e5982_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364514.Sherpa_222_NNPDF30NNLO_tautaugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364521.Sherpa_222_NNPDF30NNLO_enugamma_pty_7_15.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364522.Sherpa_222_NNPDF30NNLO_enugamma_pty_15_35.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364523.Sherpa_222_NNPDF30NNLO_enugamma_pty_35_70.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364524.Sherpa_222_NNPDF30NNLO_enugamma_pty_70_140.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364525.Sherpa_222_NNPDF30NNLO_enugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364526.Sherpa_222_NNPDF30NNLO_munugamma_pty_7_15.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364527.Sherpa_222_NNPDF30NNLO_munugamma_pty_15_35.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364528.Sherpa_222_NNPDF30NNLO_munugamma_pty_35_70.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364529.Sherpa_222_NNPDF30NNLO_munugamma_pty_70_140.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364530.Sherpa_222_NNPDF30NNLO_munugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364531.Sherpa_222_NNPDF30NNLO_taunugamma_pty_7_15.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364532.Sherpa_222_NNPDF30NNLO_taunugamma_pty_15_35.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364533.Sherpa_222_NNPDF30NNLO_taunugamma_pty_35_70.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364534.Sherpa_222_NNPDF30NNLO_taunugamma_pty_70_140.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364535.Sherpa_222_NNPDF30NNLO_taunugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364739.MGPy8EG_NNPDF30NLO_A14NNPDF23LO_lvlljjEW6_OFMinus.deriv.DAOD_STDM3.e7421_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364740.MGPy8EG_NNPDF30NLO_A14NNPDF23LO_lvlljjEW6_OFPlus.deriv.DAOD_STDM3.e7421_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364741.MGPy8EG_NNPDF30NLO_A14NNPDF23LO_lvlljjEW6_SFMinus.deriv.DAOD_STDM3.e7421_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.364742.MGPy8EG_NNPDF30NLO_A14NNPDF23LO_lvlljjEW6_SFPlus.deriv.DAOD_STDM3.e7421_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.410155.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttW.deriv.DAOD_STDM3.e5070_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.410218.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttee.deriv.DAOD_STDM3.e5070_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.410219.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttmumu.deriv.DAOD_STDM3.e5070_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.410220.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_tttautau.deriv.DAOD_STDM3.e5070_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.410470.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad.deriv.DAOD_STDM3.e6337_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.410472.PhPy8EG_A14_ttbar_hdamp258p75_dil.deriv.DAOD_STDM3.e6348_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.410560.MadGraphPythia8EvtGen_A14_tZ_4fl_tchan_noAllHad.deriv.DAOD_STDM3.e5803_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.410644.PowhegPythia8EvtGen_A14_singletop_schan_lept_top.deriv.DAOD_STDM3.e6527_s3126_r10201_p3975
mc16_13TeV.410645.PowhegPythia8EvtGen_A14_singletop_schan_lept_antitop.deriv.DAOD_STDM3.e6527_s3126_r10201_p3975
mc16_13TeV.410648.PowhegPythia8EvtGen_A14_Wt_DR_dilepton_top.deriv.DAOD_STDM3.e6615_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.410649.PowhegPythia8EvtGen_A14_Wt_DR_dilepton_antitop.deriv.DAOD_STDM3.e6615_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.410658.PhPy8EG_A14_tchan_BW50_lept_top.deriv.DAOD_STDM3.e6671_e5984_s3126_r10201_r10210_p3975
mc16_13TeV.410659.PhPy8EG_A14_tchan_BW50_lept_antitop.deriv.DAOD_STDM3.e6671_e5984_s3126_r10201_r10210_p3975

mc16e

mc16_13TeV.363147.MGPy8EG_N30NLO_Zee_Ht0_70_CVetoBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363148.MGPy8EG_N30NLO_Zee_Ht0_70_CFilterBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363149.MGPy8EG_N30NLO_Zee_Ht0_70_BFilter.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363150.MGPy8EG_N30NLO_Zee_Ht70_140_CVetoBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363151.MGPy8EG_N30NLO_Zee_Ht70_140_CFilterBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363152.MGPy8EG_N30NLO_Zee_Ht70_140_BFilter.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363153.MGPy8EG_N30NLO_Zee_Ht140_280_CVetoBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363154.MGPy8EG_N30NLO_Zee_Ht140_280_CFilterBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363155.MGPy8EG_N30NLO_Zee_Ht140_280_BFilter.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363156.MGPy8EG_N30NLO_Zee_Ht280_500_CVetoBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363157.MGPy8EG_N30NLO_Zee_Ht280_500_CFilterBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363158.MGPy8EG_N30NLO_Zee_Ht280_500_BFilter.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363159.MGPy8EG_N30NLO_Zee_Ht500_700_CVetoBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363160.MGPy8EG_N30NLO_Zee_Ht500_700_CFilterBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363161.MGPy8EG_N30NLO_Zee_Ht500_700_BFilter.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363162.MGPy8EG_N30NLO_Zee_Ht700_1000_CVetoBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363163.MGPy8EG_N30NLO_Zee_Ht700_1000_CFilterBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363164.MGPy8EG_N30NLO_Zee_Ht700_1000_BFilter.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363165.MGPy8EG_N30NLO_Zee_Ht1000_2000_CVetoBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363166.MGPy8EG_N30NLO_Zee_Ht1000_2000_CFilterBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363167.MGPy8EG_N30NLO_Zee_Ht1000_2000_BFilter.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363168.MGPy8EG_N30NLO_Zee_Ht2000_E_CMS_CVetoBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363169.MGPy8EG_N30NLO_Zee_Ht2000_E_CMS_CFilterBVeto.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363170.MGPy8EG_N30NLO_Zee_Ht2000_E_CMS_BFilter.deriv.DAOD_STDM3.e4866_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.363356.Sherpa_221_NNPDF30NNLO_ZqqZll.deriv.DAOD_STDM3.e5525_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364250.Sherpa_222_NNPDF30NNLO_llll.deriv.DAOD_STDM3.e5894_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364253.Sherpa_222_NNPDF30NNLO_lllv.deriv.DAOD_STDM3.e5916_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364254.Sherpa_222_NNPDF30NNLO_llvv.deriv.DAOD_STDM3.e5916_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364283.Sherpa_222_NNPDF30NNLO_lllljj_EW6.deriv.DAOD_STDM3.e6055_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364284.Sherpa_222_NNPDF30NNLO_lllvjj_EW6.deriv.DAOD_STDM3.e6055_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364286.Sherpa_222_NNPDF30NNLO_llvvjj_ss_EW4.deriv.DAOD_STDM3.e6055_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364287.Sherpa_222_NNPDF30NNLO_llvvjj_ss_EW6.deriv.DAOD_STDM3.e6055_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364501.Sherpa_222_NNPDF30NNLO_eegamma_pty_15_35.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364502.Sherpa_222_NNPDF30NNLO_eegamma_pty_35_70.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364503.Sherpa_222_NNPDF30NNLO_eegamma_pty_70_140.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364504.Sherpa_222_NNPDF30NNLO_eegamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364506.Sherpa_222_NNPDF30NNLO_mumugamma_pty_15_35.deriv.DAOD_STDM3.e5928_s3126_r10724_p3970
mc16_13TeV.364507.Sherpa_222_NNPDF30NNLO_mumugamma_pty_35_70.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364508.Sherpa_222_NNPDF30NNLO_mumugamma_pty_70_140.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364509.Sherpa_222_NNPDF30NNLO_mumugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364511.Sherpa_222_NNPDF30NNLO_tautaugamma_pty_15_35.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364512.Sherpa_222_NNPDF30NNLO_tautaugamma_pty_35_70.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364513.Sherpa_222_NNPDF30NNLO_tautaugamma_pty_70_140.deriv.DAOD_STDM3.e5982_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364514.Sherpa_222_NNPDF30NNLO_tautaugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364522.Sherpa_222_NNPDF30NNLO_enugamma_pty_15_35.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364523.Sherpa_222_NNPDF30NNLO_enugamma_pty_35_70.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364524.Sherpa_222_NNPDF30NNLO_enugamma_pty_70_140.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364525.Sherpa_222_NNPDF30NNLO_enugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364527.Sherpa_222_NNPDF30NNLO_munugamma_pty_15_35.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364528.Sherpa_222_NNPDF30NNLO_munugamma_pty_35_70.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364529.Sherpa_222_NNPDF30NNLO_munugamma_pty_70_140.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364530.Sherpa_222_NNPDF30NNLO_munugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364532.Sherpa_222_NNPDF30NNLO_taunugamma_pty_15_35.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364533.Sherpa_222_NNPDF30NNLO_taunugamma_pty_35_70.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364534.Sherpa_222_NNPDF30NNLO_taunugamma_pty_70_140.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364535.Sherpa_222_NNPDF30NNLO_taunugamma_pty_140_E_CMS.deriv.DAOD_STDM3.e5928_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364739.MGPy8EG_NNPDF30NLO_A14NNPDF23LO_lvlljjEW6_OFMinus.deriv.DAOD_STDM3.e7421_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364740.MGPy8EG_NNPDF30NLO_A14NNPDF23LO_lvlljjEW6_OFPlus.deriv.DAOD_STDM3.e7421_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364741.MGPy8EG_NNPDF30NLO_A14NNPDF23LO_lvlljjEW6_SFMinus.deriv.DAOD_STDM3.e7421_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.364742.MGPy8EG_NNPDF30NLO_A14NNPDF23LO_lvlljjEW6_SFPlus.deriv.DAOD_STDM3.e7421_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.410155.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttW.deriv.DAOD_STDM3.e5070_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.410218.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttee.deriv.DAOD_STDM3.e5070_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.410219.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttmumu.deriv.DAOD_STDM3.e5070_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.410220.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_tttautau.deriv.DAOD_STDM3.e5070_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.410470.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad.deriv.DAOD_STDM3.e6337_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.410472.PhPy8EG_A14_ttbar_hdamp258p75_dil.deriv.DAOD_STDM3.e6348_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.410560.MadGraphPythia8EvtGen_A14_tZ_4fl_tchan_noAllHad.deriv.DAOD_STDM3.e5803_e5984_s3126_r10724_r10726_p3975
mc16_13TeV.410644.PowhegPythia8EvtGen_A14_singletop_schan_lept_top.deriv.DAOD_STDM3.e6527_s3126_r10724_p3975
mc16_13TeV.410645.PowhegPythia8EvtGen_A14_singletop_schan_lept_antitop.deriv.DAOD_STDM3.e6527_s3126_r10724_p3975
mc16_13TeV.410648.PowhegPythia8EvtGen_A14_Wt_DR_dilepton_top.deriv.DAOD_STDM3.e6615_e5984_s3126_s3136_r10724_r10726_p3975
mc16_13TeV.410649.PowhegPythia8EvtGen_A14_Wt_DR_dilepton_antitop.deriv.DAOD_STDM3.e6615_e5984_s3126_s3136_r10724_r10726_p3975
mc16_13TeV.410658.PhPy8EG_A14_tchan_BW50_lept_top.deriv.DAOD_STDM3.e6671_e5984_s3126_s3136_r10724_r10726_p3975
mc16_13TeV.410659.PhPy8EG_A14_tchan_BW50_lept_antitop.deriv.DAOD_STDM3.e6671_e5984_s3126_s3136_r10724_r10726_p3975

Appendix B

Additional Material on studies of charge misidentification sources

B.1 Comparisons in electron kinematics between software releases and MC campaigns

In this appendix, some comparisons for electron kinematics measured between software releases in Release 20.7 (used for the measurements of electroweak same-sign WW production in ATLAS with 2015-2016 dataset [82]) and Release 21 (used during this thesis). The electron selection follows Section 6.3.1 for both releases.

Considerable improvements on the $m_{\ell\ell}$ reconstruction were observed after the ATLAS software release was updated from release 20.7 to 21.2, where the new dynamical, topological cell clustering-based algorithm [65] (described in Sec 3.3.2) is used for the electron energy reconstruction.

B.2 Charge misidentification between MC campaings

In Figure B.2, the reconstructed electron p_T in the EMCAL, the pseudorapidity $|\eta|$, the invariant mass reconstructed with leading and subleading leptons, the significance of the transverse impact parameter d_0/σ_{d_0}, the ratio between energy reconstructed in the EMCAL over energy measurement from the track curvature E/p and the number of reconstructed tracks assigned to the electron candidate (*nTracksPerParticle*) are presented for MC campaigns mc16a (shown in blue), mc16d (red) and mc16e (green) as well as for the full Run 2 set (presented in black). Deviations among Monte Carlo campaigns are within statistical uncertainties and no systematic variations are observed. A consistent description of charge misidentification can be obtained for the full Run 2 sample. The charge is correctly reconstructed in 99.68% of the cases for the full Run 2 sample (99.69%, 99.69% and 99.67% for mc16a, mc16d and mc16e respectively). On the other hand, electron charges constitute a rare phenomenon (0.32% of the electrons for full Run 2, 0.31%, 0.31% and 0.33% for mc16a, mc16d and mc16e respectively) and tend to be misidentified at greater values of p_T (dashed lines).

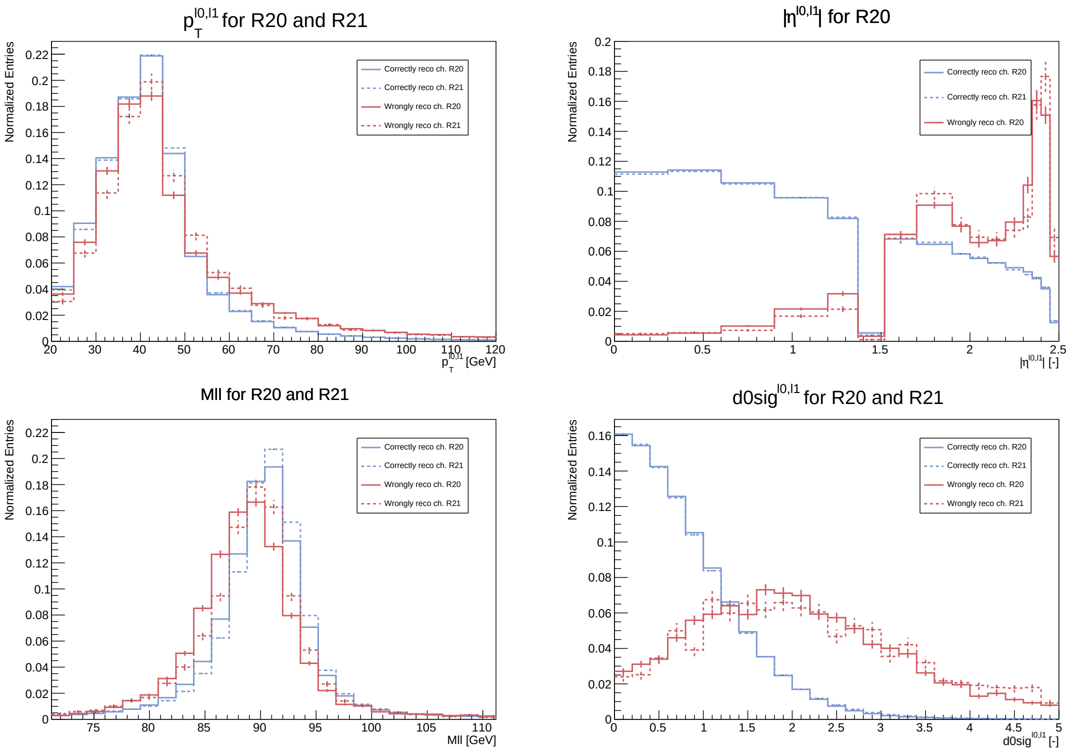


Figure B.1: Comparisons on kinematic and tracking distributions for electrons with correct and wrong charge reconstruction using the sliding-window (release 20.7 of the ATLAS software, shown in solid lines) and topological cell clustering-based (release 21.2 of the ATLAS software, shown in dashed lines) algorithms for the energy reconstruction for both correctly (blue) and wrongly (red) reconstructed electrons. The electron selection follows Section 6.3.1 in both releases. Only statistical uncertainties are shown.

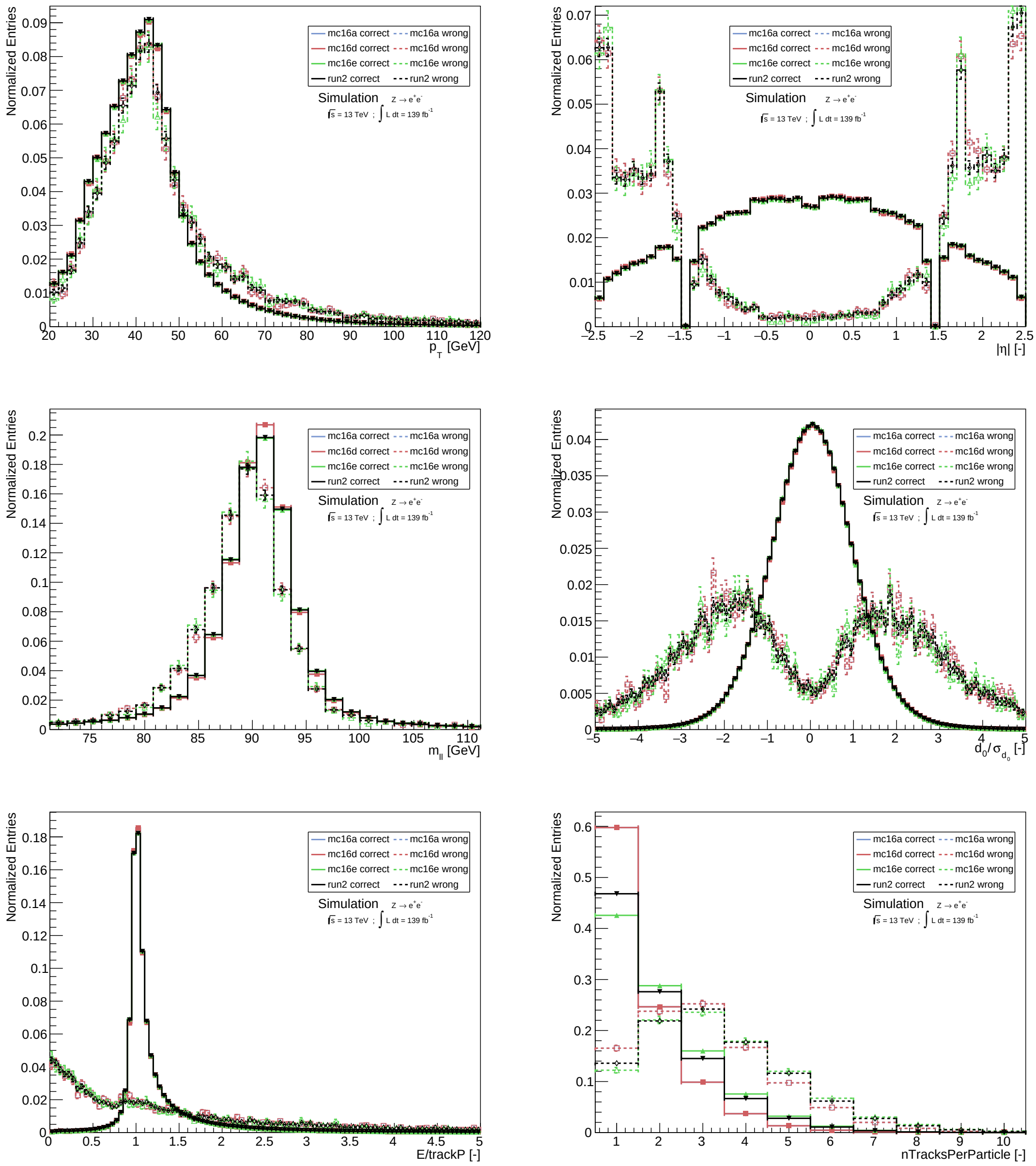


Figure B.2: Kinematic and track distributions for electrons with correct (solid line) and wrong charge reconstruction (dashed line) in simulated $Z \rightarrow ee$ decays using the full Run 2 sample (139 fb^{-1}). Distributions for mc16a (blue), mc16d (red) and mc16e (green) are also shown individually. On top, the reconstructed electron p_T (left) and the pseudorapidity $|\eta|$ (right). The central row contains the invariant mass reconstructed with both electrons $m_{\ell\ell}$ (left) and the significance of the transverse impact parameter d_0/σ_{d_0} (right). Below, the ratio between energy reconstructed in the EMCAL over energy measurement from the track curvature E/p (left) and the number of tracks assigned to the electron candidate (*nTracksPerParticle*, right). Only statistical uncertainties are shown. At least one charge-misidentified electron is considered for the $m_{\ell\ell}$ distribution of wrongly reconstructed electrons.

TruthType		TruthOrigin	
Value	Description	Value	Description
0	Unknown electron type	5	γ conversions
2	Isolated electron	13	Z decays
3	Non-isolated electron	40	FSR radiation
4	Background electron		
15	Non isolated Photon		

Table B.1: Description of relevant values of `TruthType` and `TruthOrigin` during studies on charge misidentification sources at truth level.

Notable deviations are found on the distribution of the number of tracks assigned to the electron candidate (*nTracksPerParticle*) among campaigns. Respect to mc16a and mc16d, the number of events with only one track assigned drops by 18%. These electrons candidates allocate additional tracks in mc16e. The denser detector environment (i.e. the increased number of pile-up interactions) can explain this effect due to the reduced efficiency in track reconstruction discussed in Section 3.3.1. Nevertheless, the track corresponding to the prompt electron signature is normally the best-matching track for the electron candidate, as the found kinematic distributions for mc16e agree with previous MC campaigns.

B.3 Technical details on the truth selection of electron categories

In detector simulation, truth information on the reconstructed and generated electron is retrieved. This information helps to determine whether the reconstructed charge was identified correctly. The electron can be identified by its PDG ID [15], with `PDGID == 11` and `PDGID == -11` identifying electrons and positrons, respectively. The characterization of electrons is then applied at generator level. In the final state, only the truth information on the generated electrons closest to the reconstructed electron candidates are considered. The ATLAS Athena tool `MCTruthClassifier` is used to retrieve information on the electron type and physical process of origin making use of two variables, `TruthType` and `TruthOrigin`, respectively. Within this tool, electrons of interest are divided into three main categories. *Isolated* (or *prompt*) electrons are signal particles, e.g. electrons from boson decay $Z \to ee$. `TruthType==2` is assigned to them. *Non-isolated* electrons, which are originated from decay of the charmed or b-flavored mesons and baryons (as exception decay of J/ψ included) and have `TruthType==3`. And lastly *background* electrons that correspond to all the rest including bb and cc mesons decays or photon conversions, with `TruthType==4`. Similar determination of the particle type can be determined for photons, muons and τ-leptons. Photons will be also relevant for our studies, with `TruthType==14` being assigned to isolated photons and `TruthType==15` to none isolated photons. Besides, the process of origin is identified by `TruthOrigin`, e.g. particles from Z boson decays (`TruthOrigin==13`), photon conversions (`TruthOrigin==5`) or FSR photons (`TruthOrigin==40`). Table B.1 summarizes the relevant values for `TruthType` and `TruthOrigin`.

The transverse momentum of a particle at truth level p_T^{Truth} (i.e. calculated using information from the detector simulation) is computed as a *dressed transverse momentum*, i.e. the sum of the particle momenta plus particle momenta from subsequent produced secondary particles (for instance, particles radiated through bremsstrahlung or photon conversions)[1]. This sum takes place from the beginning of the detector simulation up to the final state. Only truth information on the mother particle and best-matched electron to the reconstructed track in the final state are available at the PxAOD level. The true charge of the reconstructed electrons is needed to identify charge misidentification in the detector simulation. This true charge is inferred from the sign of the `PDGID` of the particle at generation level `firstEgMotherPdgID`, which is different from the reconstructed charge obtained from the track reconstruction in case charge misidentification occurs. Following the criterion in [92], charge misidentification for an electron traversing the ATLAS detector (also referred to as *charge flip*) is identified at truth level using the expression

$$\texttt{recoCharge} \times \texttt{firstEgMotherPdgID} > 0\,, \tag{B.1}$$

with `recoCharge` being the variable identifying the reconstructed charge from the track curvature and `firstEgMotherPdgID` = $\pm$`11` identifying the mother particle seeding the detector simulation as electron (`+11`) or positron (`-11`). Since the signs of `firstEgMotherPdgID` and `recoCharge` are opposite for electrons and positrons, the product between them for a wrongly reconstructed charge is positive. The definition of charge flip given in Eq (B.1) stands for particles with antiparticle partner, as expected. Neutral particles have positive PDG IDs [15]. Therefore, the charge misidentification for electrons produced from neutral particle decays at generator level will be assigned when a positron is reconstructed, with correct reconstruction being found for electrons. For these cases, the definition of charge misidentification is therefore not well defined. Their contributions to charge misidentification phenomena need to be treated separately.

Track variables are extracted from the `xAOD::TrackParticle` container within the PxAOD samples. Charge misidentification is fundamentally caused only by detector imperfections. Consequently, no considerable differences among either Monte Carlo generators or physical processes (Z boson decays, $t\bar{t}$, J/ψ, ...) are expected[2].

For the studies on charge misidentification sources presented in Section 6.3.3, The contribution from FSR-induced electrons and other background processes is removed by introducing the following selection [92] to avoid the influence in charge misidentification,

```
TruthType == 2 || (TruthOrigin == 5 && |FirstEgMotherPdgID|== 11).
```
(B.2)

The factor `TruthType == 2` keeps prompt electrons while the second factor allows to select any reconstructed particle originated after photon conversion (`TruthOrigin == 5`) that was originally seeded into the detector simulation as an electron (`|FirstEgMotherPdgID|== 11`). The described methodology is applied to produce kinematic distributions for charge misidentified and correctly reconstructed electrons. Comparisons on these distributions are presented in the Section 6.3.2.

Correctly identified electrons in the final state that have been produced from Z boson decays are referred to as *prompt* electrons. These can be characterized in the final

[1]The truth p_T defined here is a priory different from the *generated* p_T corresponding to the transverse momentum of the electron at generation level in Monte Carlo. Their values are assumed to be identical.

[2]Sources of charge misidentification such as FSR radiation do depend on the used generator. However, their influence in charge misidentification will be proven to be small.

Electron type	Charge flip (Eq (B.1))	`firstEgMotherTruthType`	`firstEgMotherTruthOrigin`	`TruthType`	`TruthOrigin`
Prompt	False	2	13	2	13
γ conv	False	2	13	4	5
Charge misID γ conv	True	2	13	4	5
Charge misID Prompt	True	2	13	2	13
FSR	—	—	—	—	40

Table B.2: Values of `FirstEgMotherTruthType` and `FirstEgMotherTruthOrigin`, `TruthType` and `TruthOrigin` used for the electron categories identified in truth studies with $Z \rightarrow ee$ Madgraph MC of charge misidentification. Values for `FirstEgMotherTruthType` and `FirstEgMotherTruthOrigin` for FSR-tagged electrons are selected with only the `TruthOrigin` requirement.

state at the truth level by requiring an isolated electron coming from a Z boson decay `TruthType==2` and `TruthOrigin==13`. The same values need to be applied on the mother particle initiating the detector simulation, i.e. `firstEgMotherTruthType==2` and `firstEgMotherTruthOrigin==13` (see Table B.1). The charge is required to be correctly reconstructed by applying Eq (B.1). Charge misidentification in prompt electrons is set in the same way as the prompt electrons but requiring charge flip after reconstruction using Eq (B.1). This source will be denoted as *charge misidentified prompts* (or *charge misID prompts*).

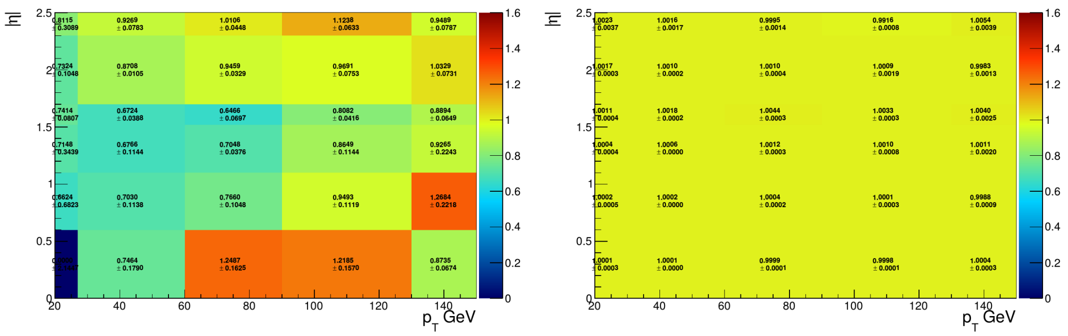
Reconstructed electrons from *photon conversions* (labeled also as *charge misID γ conversions*) will be described by selecting background electrons arising from this process, i.e. setting `TruthType==4` and `TruthOrigin==5` at truth level. In order to properly require charge flip, an isolated electron with Z boson decay as origin is demanded for the mother particle by applying `firstEgMotherTruthType==2` and `firstEgMotherTruthOrigin==13`. The charge can also be correctly reconstructed after photon conversions. By imposing charge flip using Eq (B.1), both cases will be referred to as charge misidentified photon conversions or photon conversions, respectively.

Definition of *FSR electrons* at truth level is characterized by the requirement `TruthType==40` and `TruthOrigin==40`, i.e. the reconstructed electron arises from a non-isolated photon through FSR radiation. Since the PDG ID does not provide information on the charge of the mother particle for these, both electrons involving charge flip and correct reconstruction according to Eq B.1 will be considered in this group.

Appendix C

Additional distributions on charge misidentification background estimation

C.1 Scale factors for charge misidentification probabilities used in the 2015-2016 analysis



(a) (b)

Figure C.1: Scale factors applied in the previous analysis using the 36.1 fb^{-1} dataset recorded in 2015-2016 for the computation of charge misidentification probabilities, as defined in (7.1). The distributions are presented for wrongly reconstructed electron charges (a) and correctly reconstructed charges (b) along with systematic uncertainties. These scale factors are only available up to 145 GeV. Taken from [91].

C.2 Comparisons of the charge misidentification data-driven background estimation with respect to the previous analysis.

Data and MC samples from the published analysis dataset processes to produce the background yield in the Z window control region and in the signal region. Using both the implementation and the source code from [91], the cutflows are compared for data and the charge misidentification background yield along the different cut stages. Identical cutflows were achieved on both setups. Right after the match, the randomization procedure for the energy correction is corrected. While the correction of p_T was applied on both leading (ℓ_0) and subleading (ℓ_1) leptons when the subleading lepton was chosen by the randomization procedure presented in Section 7.1 on the results from [91], only the chosen electron is corrected after these modifications. The agreement between data and the charge misID background estimation is found to improve for the p_T distributions after these considerations. Comparisons are presented in Figures C.2 and C.3.

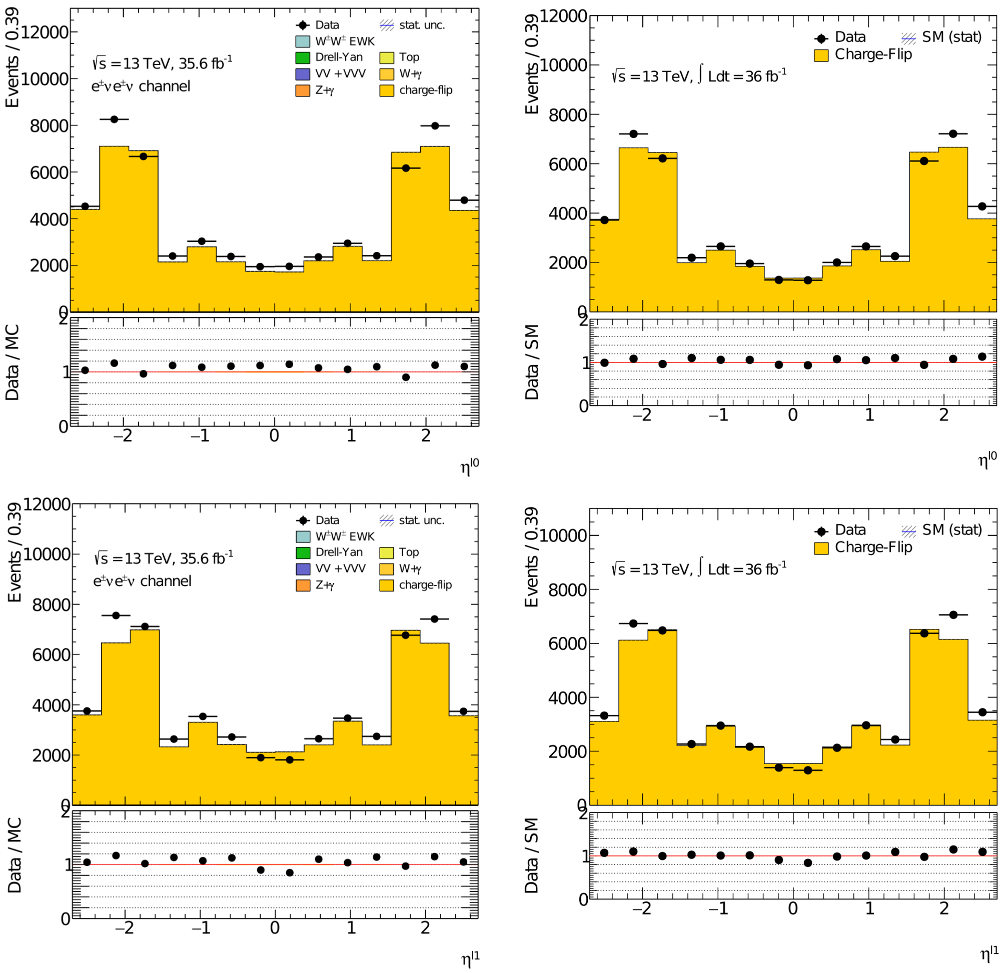


Figure C.2: Comparisons of the η distribution for leading (ℓ_0) and subleading (ℓ_1) leptons after modifications on energy correction (right) respect to the estimation in [91] (left). The same PxAOD samples in Release 20.7 from the previous analysis are used in both configurations. FSR electrons were not removed for these results, following the procedure in Ref. [91]. Therefore, they were also included for the distributions of charge misidentification probabilities and energy correction parameters.

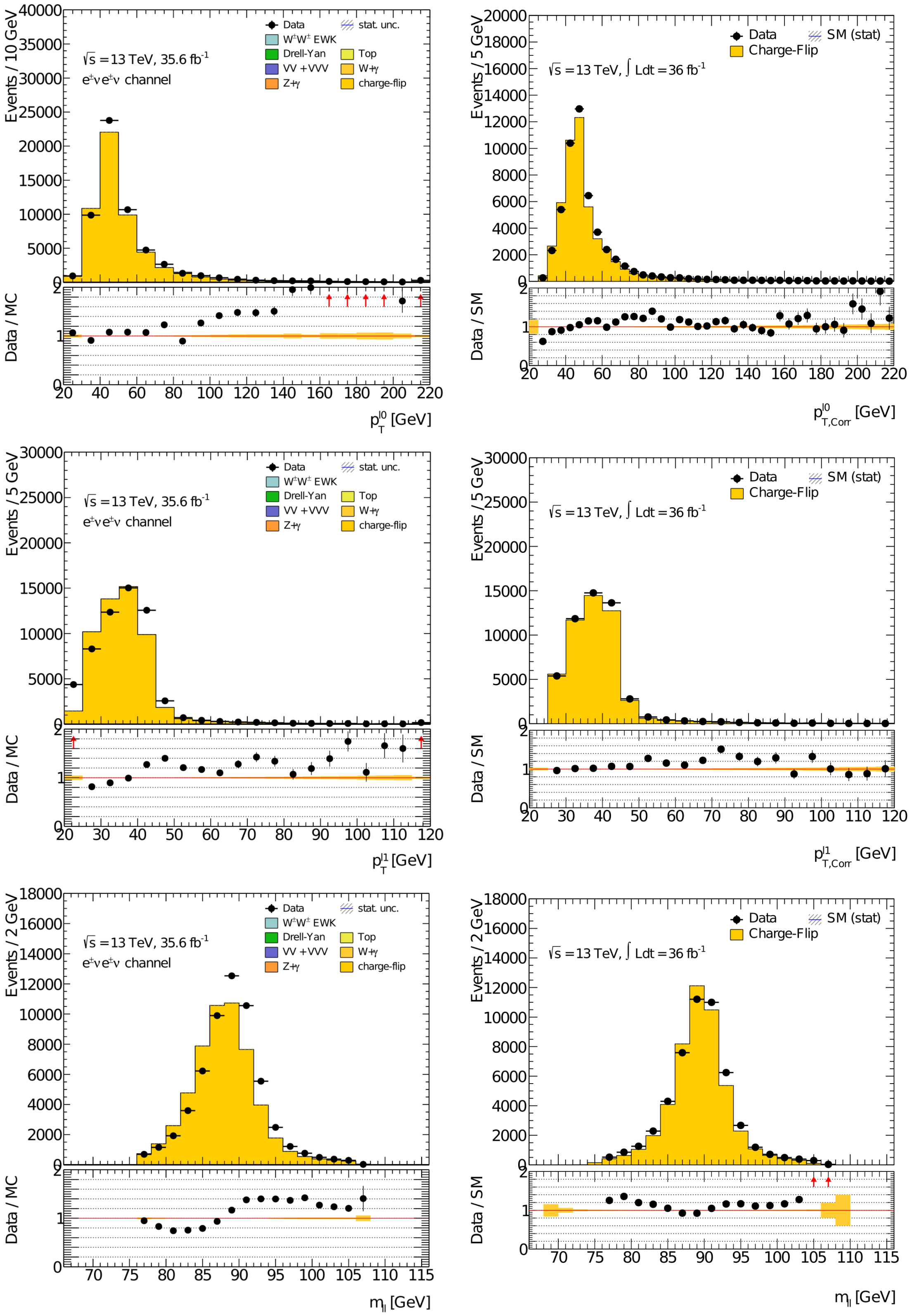


Figure C.3: Comparisons of the p_T for leading (ℓ_0) and subleading (ℓ_1) leptons and their invariant mass $m_{\ell\ell}$ distributions after modifications on energy correction (right) respect to the code from [91] (left). The same PxAOD samples in Release 20.7 from the previous analysis are used in both configurations. FSR electrons were not removed for these results, following the procedure in Ref. [91]. Therefore, they were also included for the distributions of charge misidentification probabilities and energy correction parameters.

C.3 $t\bar{t}$ tests in MC

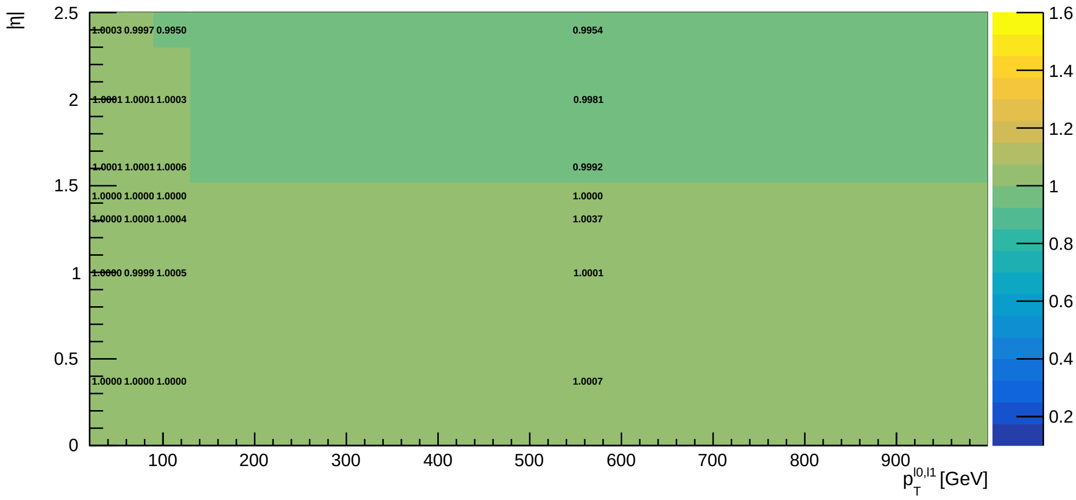
ee	mc16a [raw]	mc16d [raw]	mc16e [raw]	run2 [raw]
ttbarSS	93.2 ± 1.9 [2479]	116.7 ± 2.2 [3126]	155 ± 3 [4108]	365 ± 4 [9713]
ttbarOS	88.5 ± 0.3 [435834]	103.4 ± 0.3 [510077]	137.2 ± 0.3 [670680]	329.0 ± 0.5 [$1.61 \cdot 10^6$]
%	5.1 ± 2.0	9.1 ± 0.7	8.0 ± 0.6	7.2 ± 0.4

em	mc16a [raw]	mc16d [raw]	mc16e [raw]	run2 [raw]
ttbarSS	572 ± 5 [14963]	728 ± 5 [19496]	955 ± 6 [25234]	2250 ± 10 [59693]
ttbarOS	552.9 ± 0.8 [$2.67 \cdot 10^6$]	661.4 ± 0.9 [$3.23 \cdot 10^6$]	878.6 ± 1.0 [$4.24 \cdot 10^6$]	2092.8 ± 1.5 [$1.02 \cdot 10^7$]
%	3.4 ± 0.8	9.1 ± 0.7	8.0 ± 0.6	7.19 ± 0.4

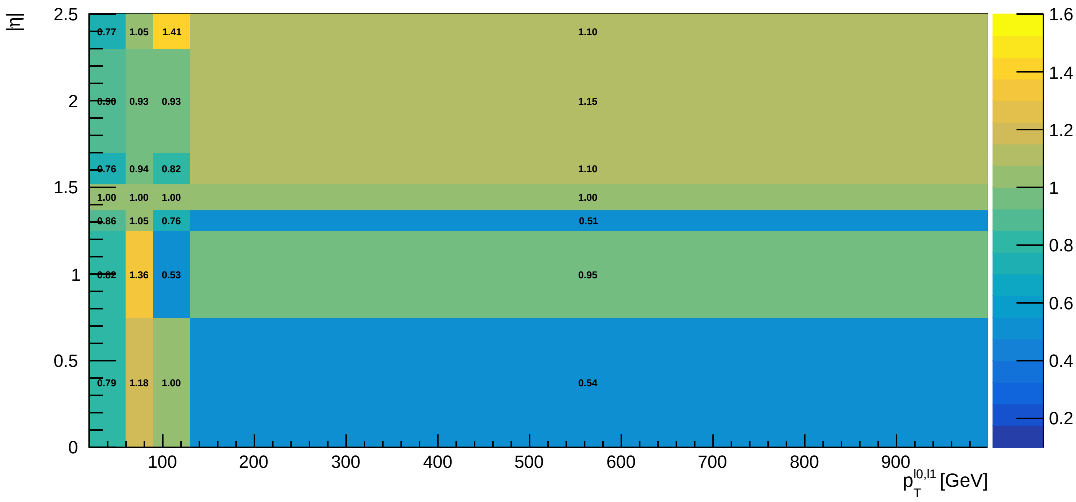
me	mc16a [raw]	mc16d [raw]	mc16e [raw]	run2 [raw]
ttbarSS	293 ± 3 [7755]	370 ± 4 [9884]	481 ± 4 [12805]	1144 ± 7 [30444]
ttbarOS	279.5 ± 0.4 [$2.40 \cdot 10^6$]	334.0 ± 0.5 [$2.86 \cdot 10^6$]	444.1 ± 0.5 [$3.75 \cdot 10^6$]	1057.5 ± 0.9 [$9.01 \cdot 10^6$]
%	4.8 ± 1.2	9.7 ± 1.0	7.7 ± 0.9	7.6 ± 0.6

Table C.1: Integrated number of events for OS and SS event selections in the Z window CR along different MC campaigns mc16a (corresponding to 2015-2016 dataset), mc16d (corresponding to 2017 dataset), mc16e (2018 dataset), and full run-2 (2015-2018 dataset) for the closure test using $t\bar{t}$ MC in the ee, $e\mu$ and μe channels. The requirement $|m_{ee} - m_Z| < 15$ GeV is not applied. The raw number of events is included in squared brackets. The deviation between both selections is shown in percentage on the third row, whose uncertainty is propagated from the statistical uncertainties of the yields.

C.4 Scale factors for the ECIDS tool

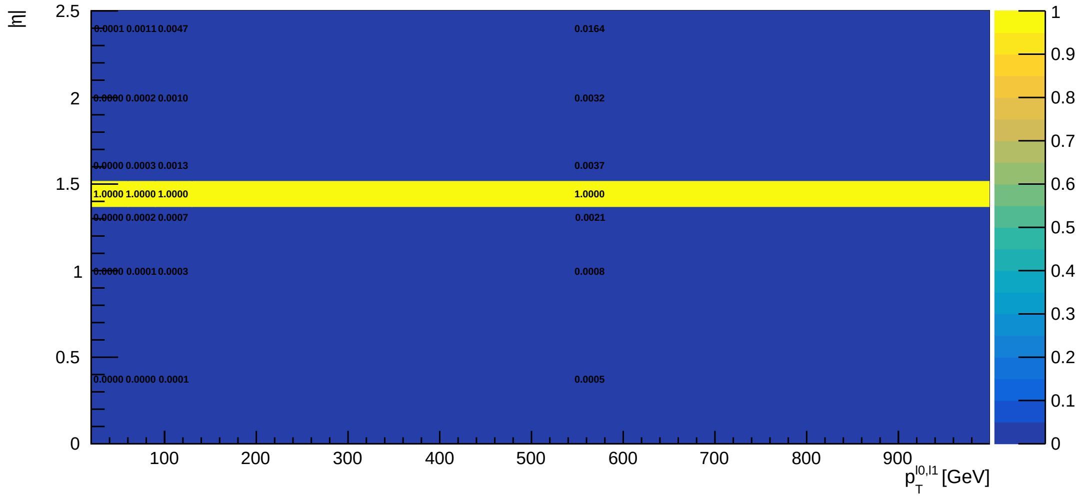


(a) Correctly reconstructed electrons.

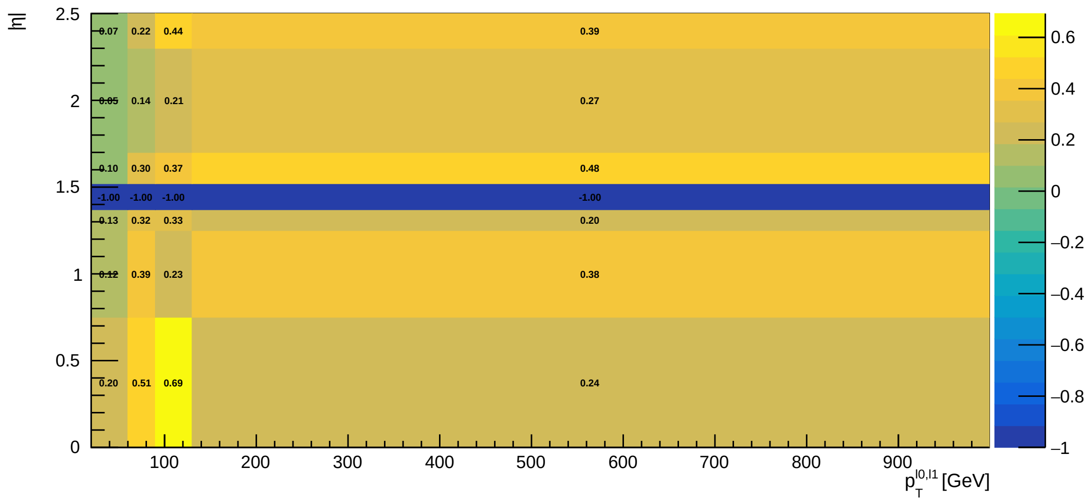


(b) Wrongly reconstructed electrons.

Figure C.4: Scale factors for `TightLH` and `Gradient` isolation electrons using the ECIDS for correctly (a) and wrongly (b) identified electron charges [68].

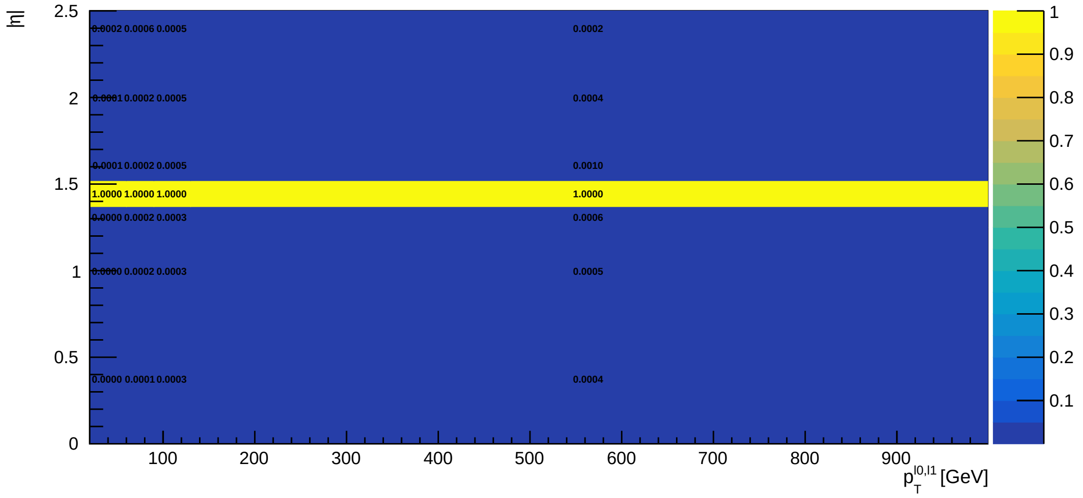


(a) Correctly reconstructed electrons.

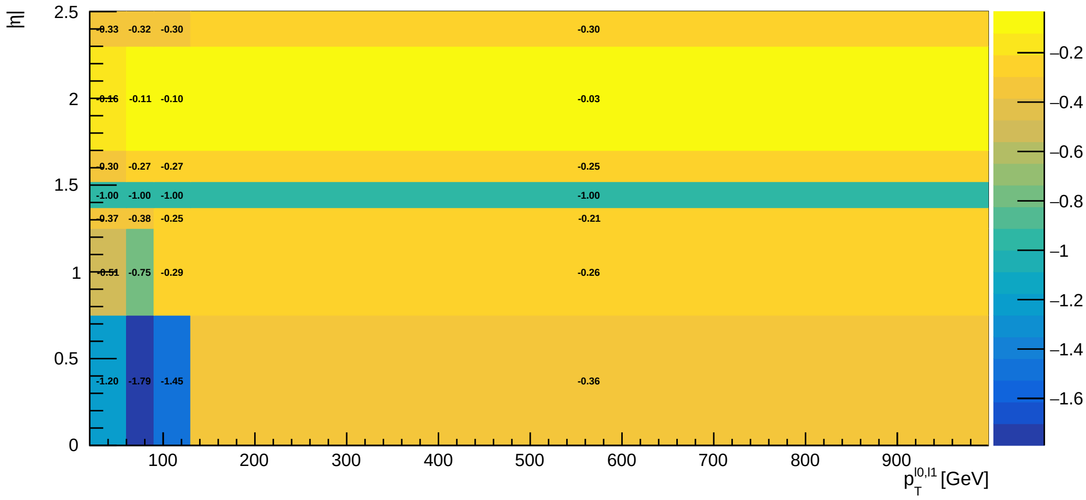


(b) Wrongly reconstructed electrons.

Figure C.5: Statistical uncertainties on the scale factors for `TightLH` and `Gradient` isolation electrons using the ECIDS for correctly (a) and wrongly (b) identified electron charges [68].

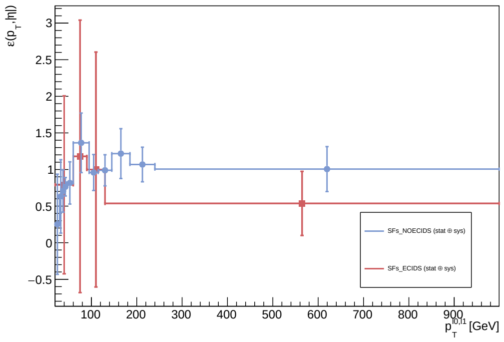


(a) Correctly reconstructed electrons.

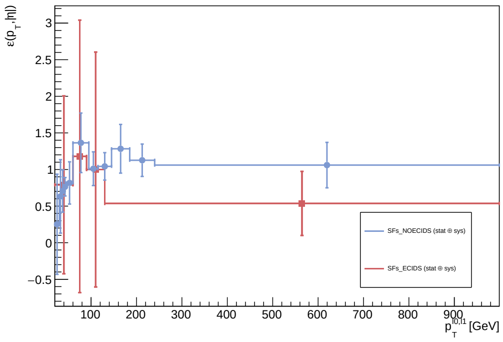


(b) Wrongly reconstructed electrons

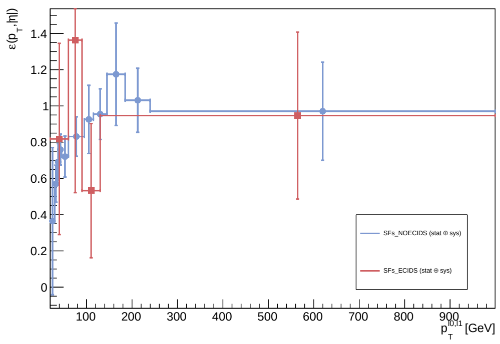
Figure C.6: Total systematic uncertainties on the scale factors using the ECIDS selection for `TightLH` and `Gradient` isolation electrons using the ECIDS for correctly (a) and wrongly (b) identified electron charges [68].



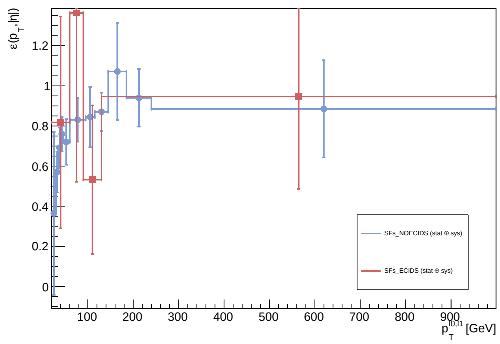
(a) $0 < |\eta| < 0.5$



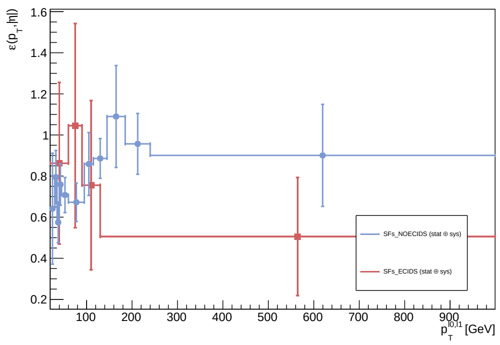
(b) $0.5 < |\eta| < 0.75$



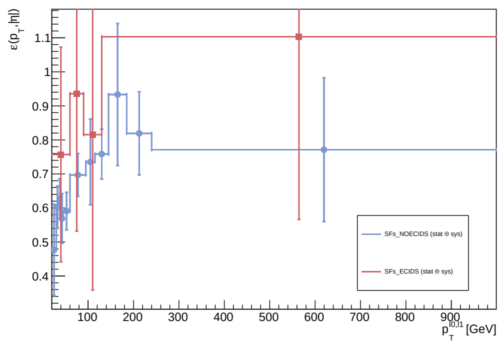
(c) $0.75 < |\eta| < 1.05$



(d) $1.05 < |\eta| < 1.25$

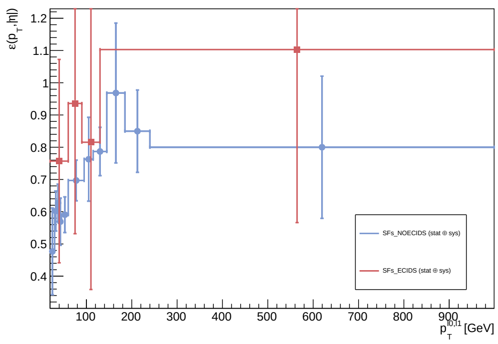


(e) $1.25 < |\eta| < 1.37$

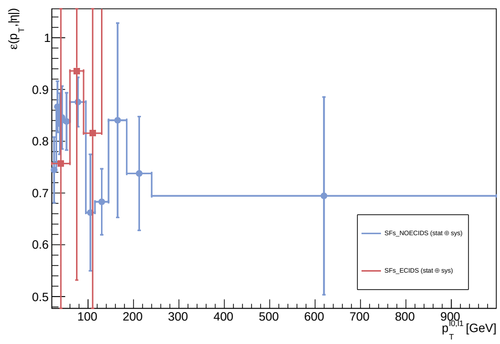


(f) $1.52 < |\eta| < 1.62$

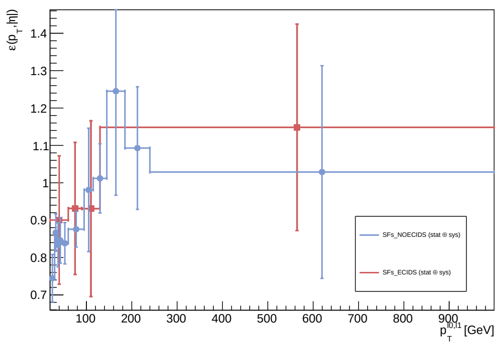
Figure C.7: Comparisons of the scale factors provided for `TightLH` electron identification and `Gradient` isolation with (red) and without (blue) the use of the ECIDS tool in different η bins for $|\eta| < 1.62$. The binning of the scale factors without ECIDS label the plots. The comparison is done respect to the closest bin in $|\eta|$ for the ECIDS scale factors. The sum in quadrature of statistical and systematic uncertainties is shown.



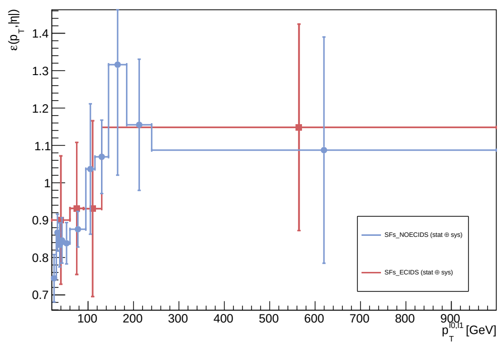
(a) $1.62 < |\eta| < 1.7$



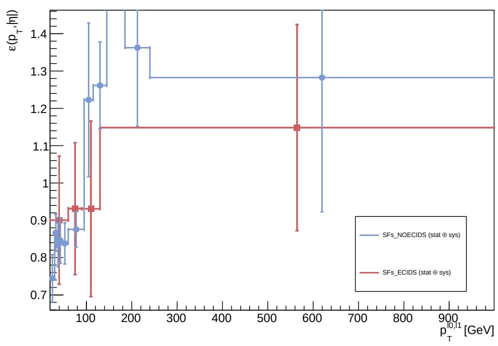
(b) $1.7 < |\eta| < 1.8$



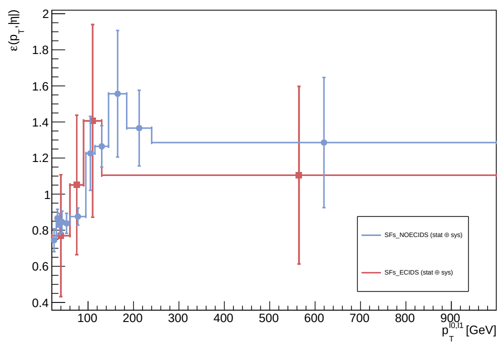
(c) $1.8 < |\eta| < 1.9$



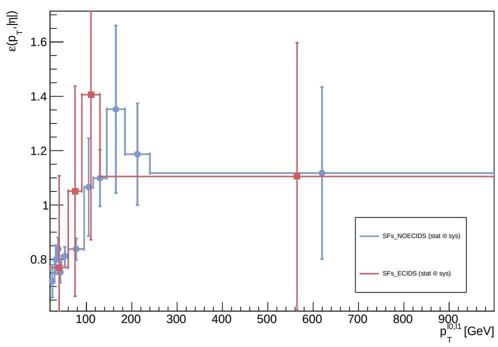
(d) $1.9 < |\eta| < 2.0$



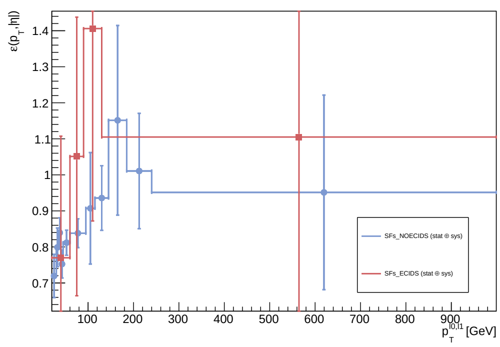
(e) $2.0 < |\eta| < 2.1$



(f) $2.1 < |\eta| < 2.3$



(g) $2.3 < |\eta| < 2.4$



(h) $2.4 < |\eta| < 2.5$

Figure C.8: Comparisons of the scale factors provided by the *Egamma CP group* for `TightLH` electron identification and `Gradient` isolation with (red) and without (blue) the use of the ECIDS tool in different η bins for $1.62 < |\eta| < 2.5$. The binning of the scale factor distributions without the use of the ECIDS tool label the plots. The comparison is done respect to the closest bin in $|\eta|$ for the ECIDS scale factors. The sum in quadrature of statistical and systematic uncertainties is shown.

C.5 Comparison of charge misidentification probabilities in different MC campaigns

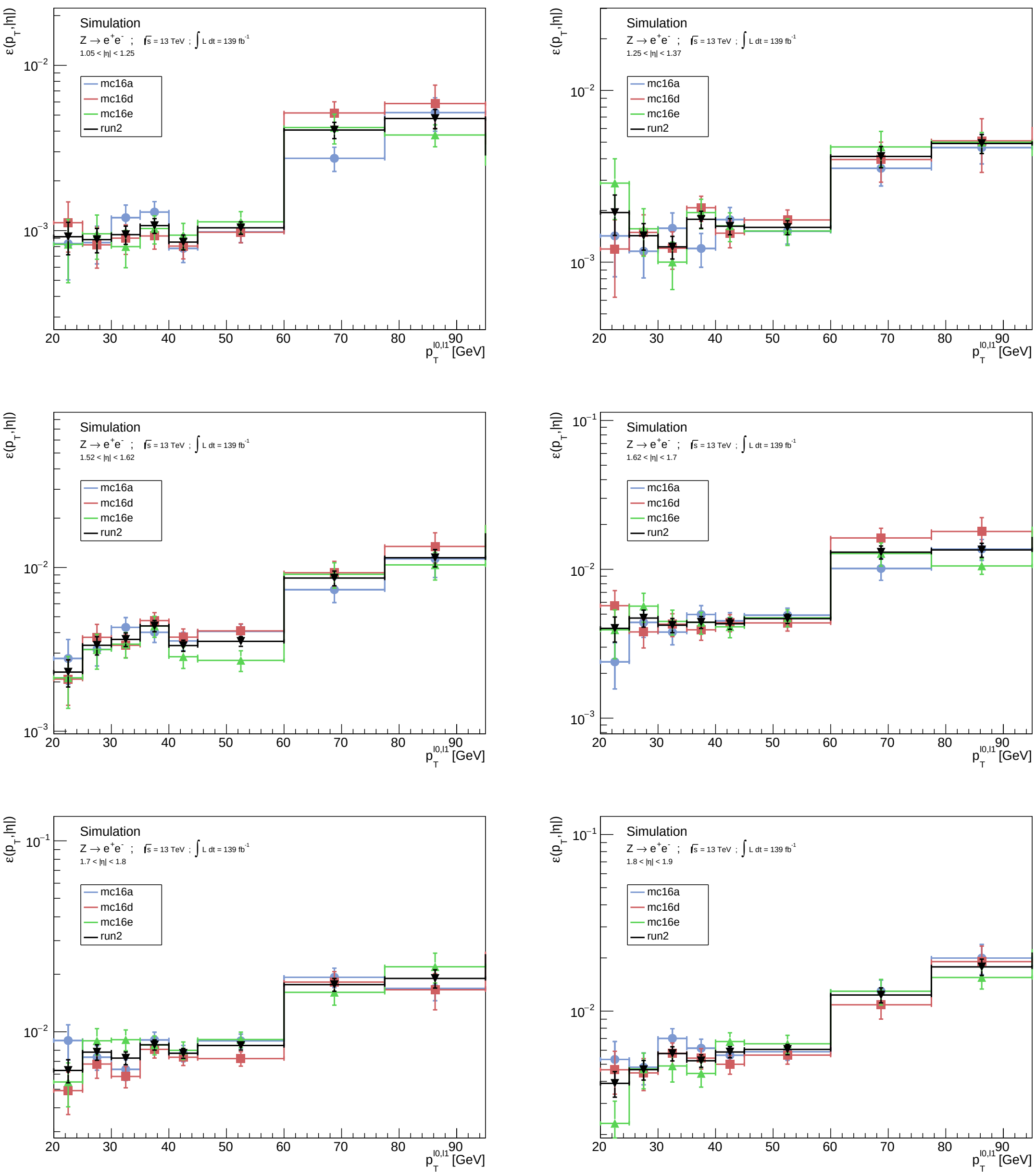


Figure C.9: Comparison of charge misidentification probabilities for mc16a (shown in blue), mc16d (shown in red) and mc16e (shown in green) Monte Carlo campaigns along with distributions for the full Run 2 set (black) for pseudorapidity bins within $0 < |\eta| < 1.9$. The distributions were computed by exploiting the truth information in $Z \to ee$ Madgraph Monte Carlo samples. Reconstructed electrons originated from FSR were removed for the computation and the scale factors shown in Figure 7.1 were applied. Only statistical uncertainties are shown. The region $1.37 < |\eta| < 1.52$ is not considered.

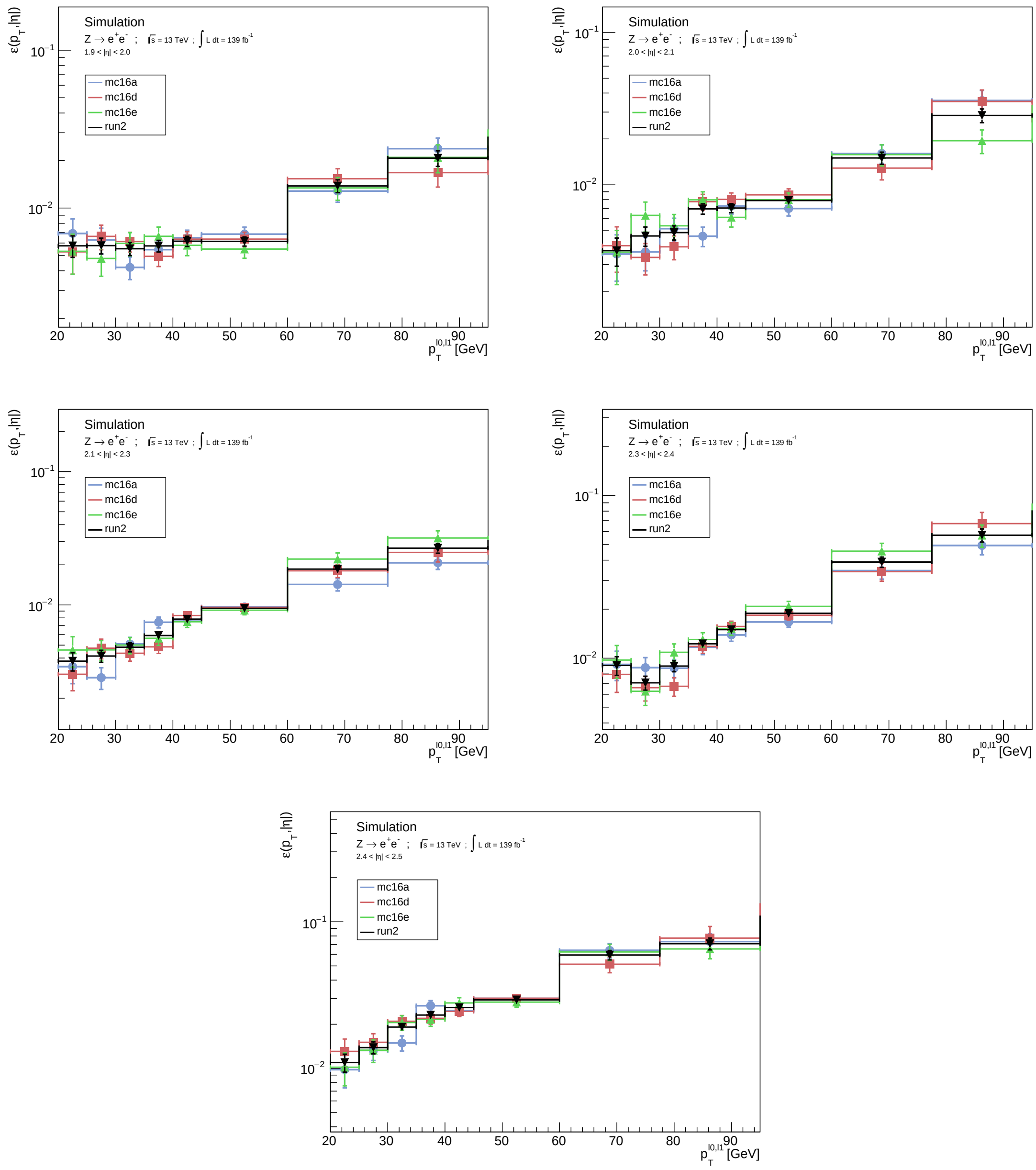


Figure C.10: Comparison of charge misidentification probabilities for mc16a (shown in blue), mc16d (shown in red) and mc16e (shown in green) Monte Carlo campaigns along with distributions for the full Run 2 set (black) for pseudorapidity bins within $1.9 < |\eta| < 2.5$. The distributions were computed by exploiting the truth information in $Z \to ee$ Madgraph Monte Carlo samples. Reconstructed electrons originated from FSR were removed for the computation and the scale factors shown in Figure 7.1 were applied. Only statistical uncertainties are shown.

C.6 Systematic variations on charge misidentification probabilities

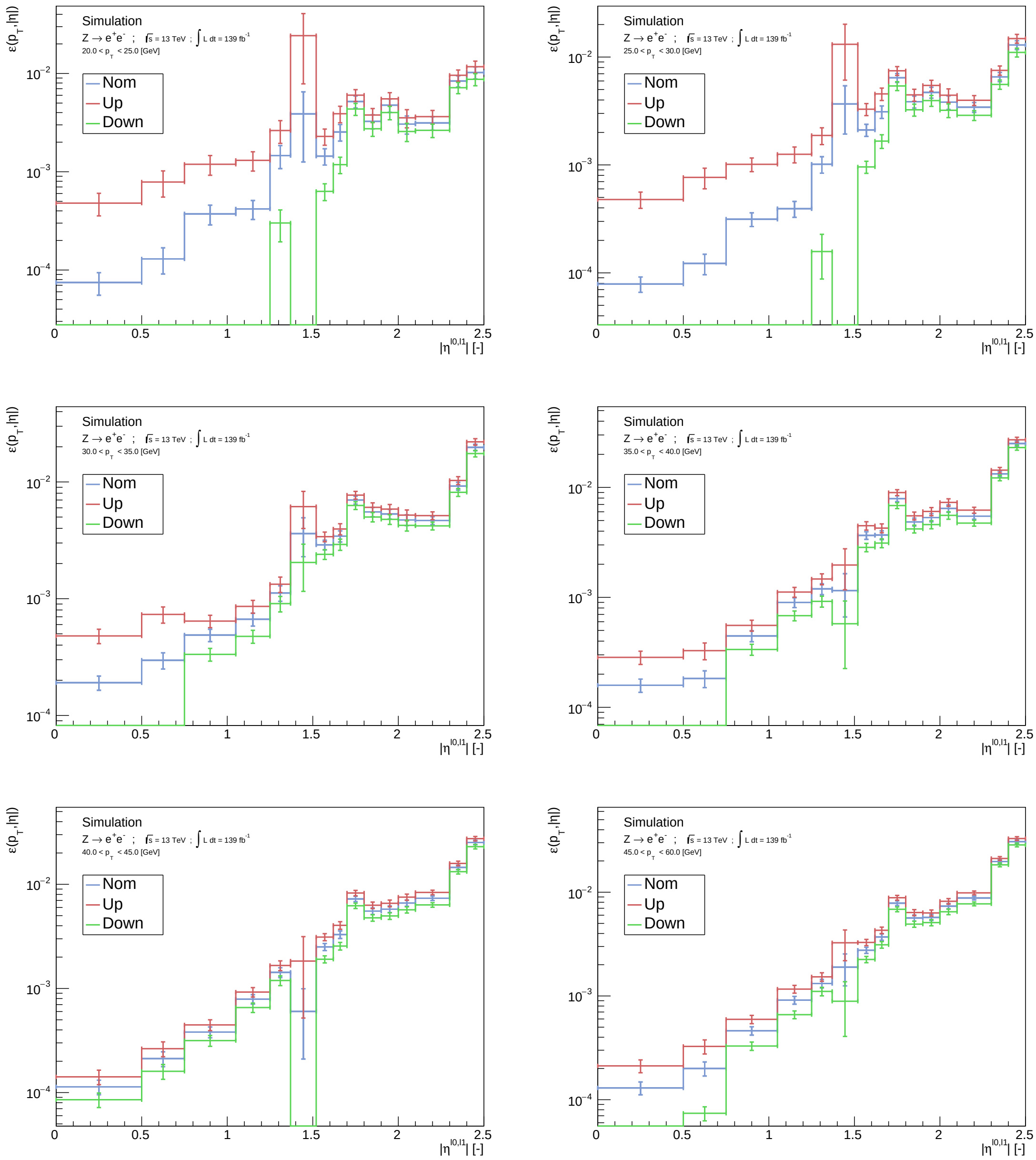


Figure C.11: Systematic variations on the scale factors for the computation of the charge misidentification probabilities using the full Run 2 sample in p_T bins within $20 < p_T < 60$ GeV. Reconstructed electrons originated from FSR were removed for the computation and the scale factors shown in Figure 7.1 were applied. Only statistical uncertainties are shown. The region $1.37 < |\eta| < 1.52$ is not considered.

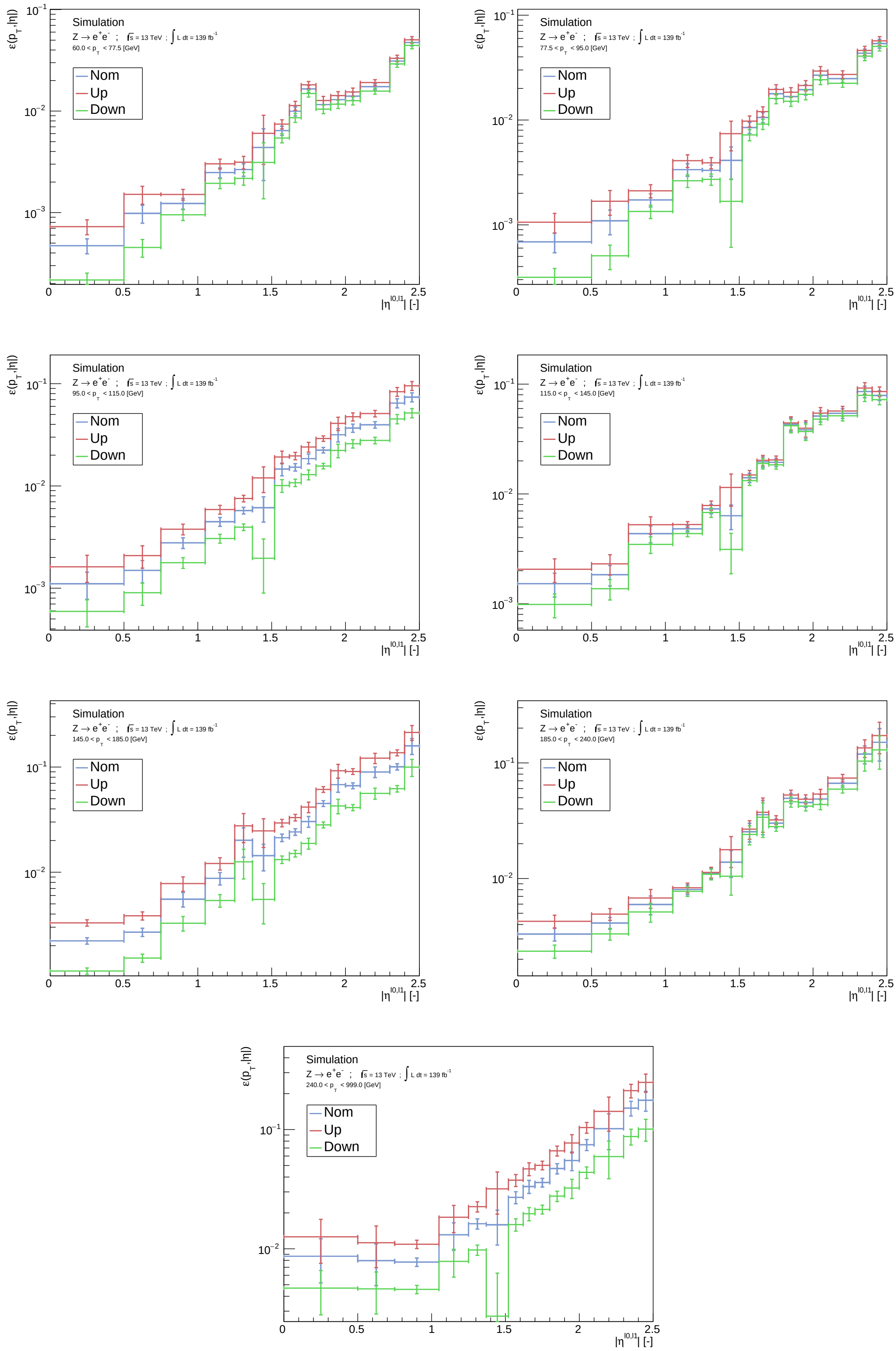


Figure C.12: Systematic variations on the scale factors on the charge misidentification probabilities for the full Run 2 sample in p_T bins within $p_T > 60$ GeV. Reconstructed electrons originated from FSR were removed for the computation and scale factors shown in Figure 7.1 were applied. Only statistical uncertainties are shown. The region $1.37 < |\eta| < 1.52$ is not considered.

C.7 Comparison of charge misidentification probabilities in different MC campaigns using ECIDS

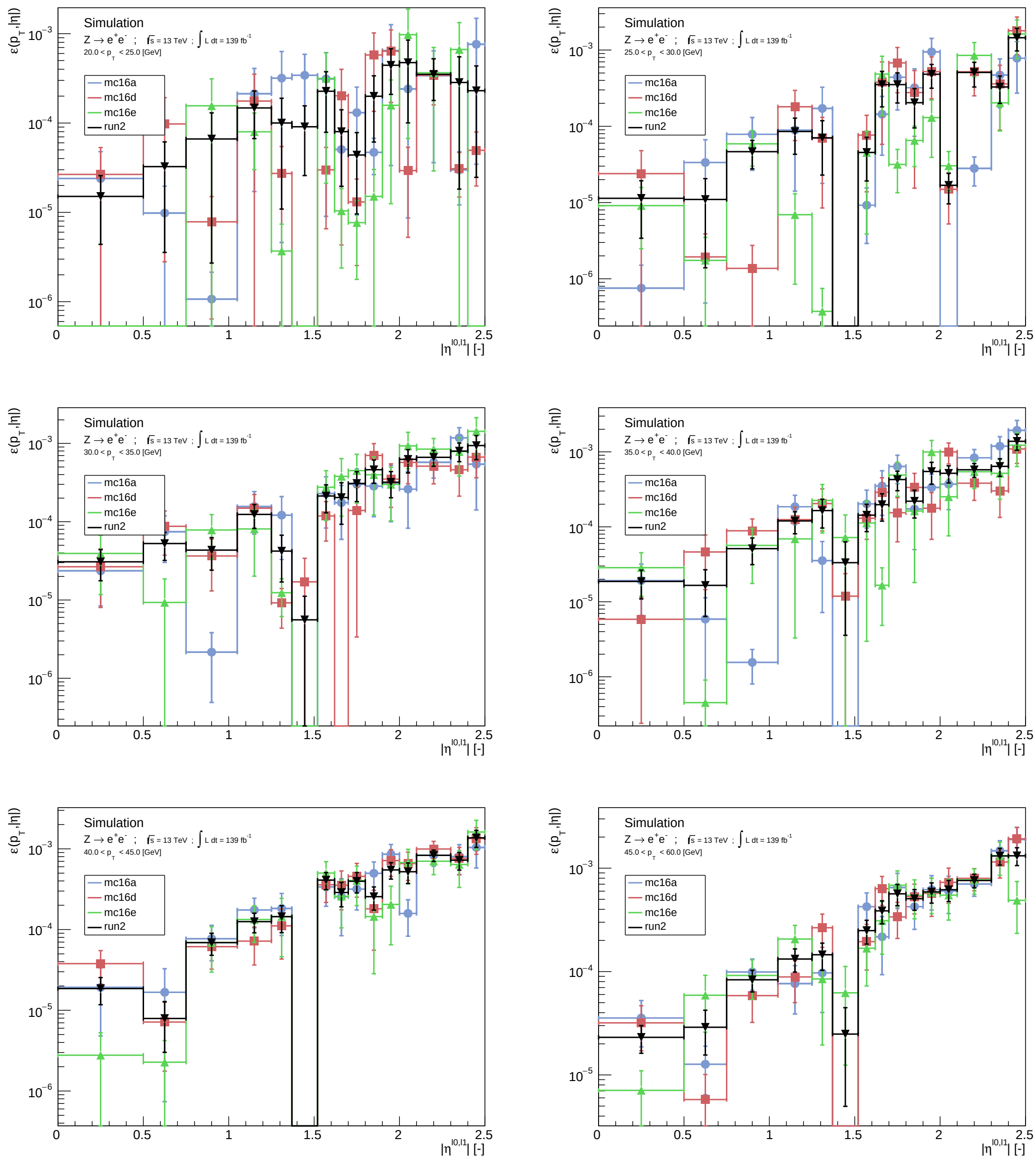


Figure C.13: Comparison of charge misidentification probabilities for mc16a (shown in blue), mc16d (shown in red) and mc16e (shown in green) Monte Carlo campaigns along with distributions for the full Run 2 set (black) for pseudorapidity bins within $20 < p_T < 60$ GeV after the ECIDS selection. The distributions were computed by exploiting the truth information in $Z \rightarrow ee$ Madgraph MC samples. Reconstructed electrons originated from FSR were removed for the computation and the scale factors shown in Figure C.4 were applied. Only statistical uncertainties are shown. The region $1.37 < |\eta| < 1.52$ is not considered in the analysis.

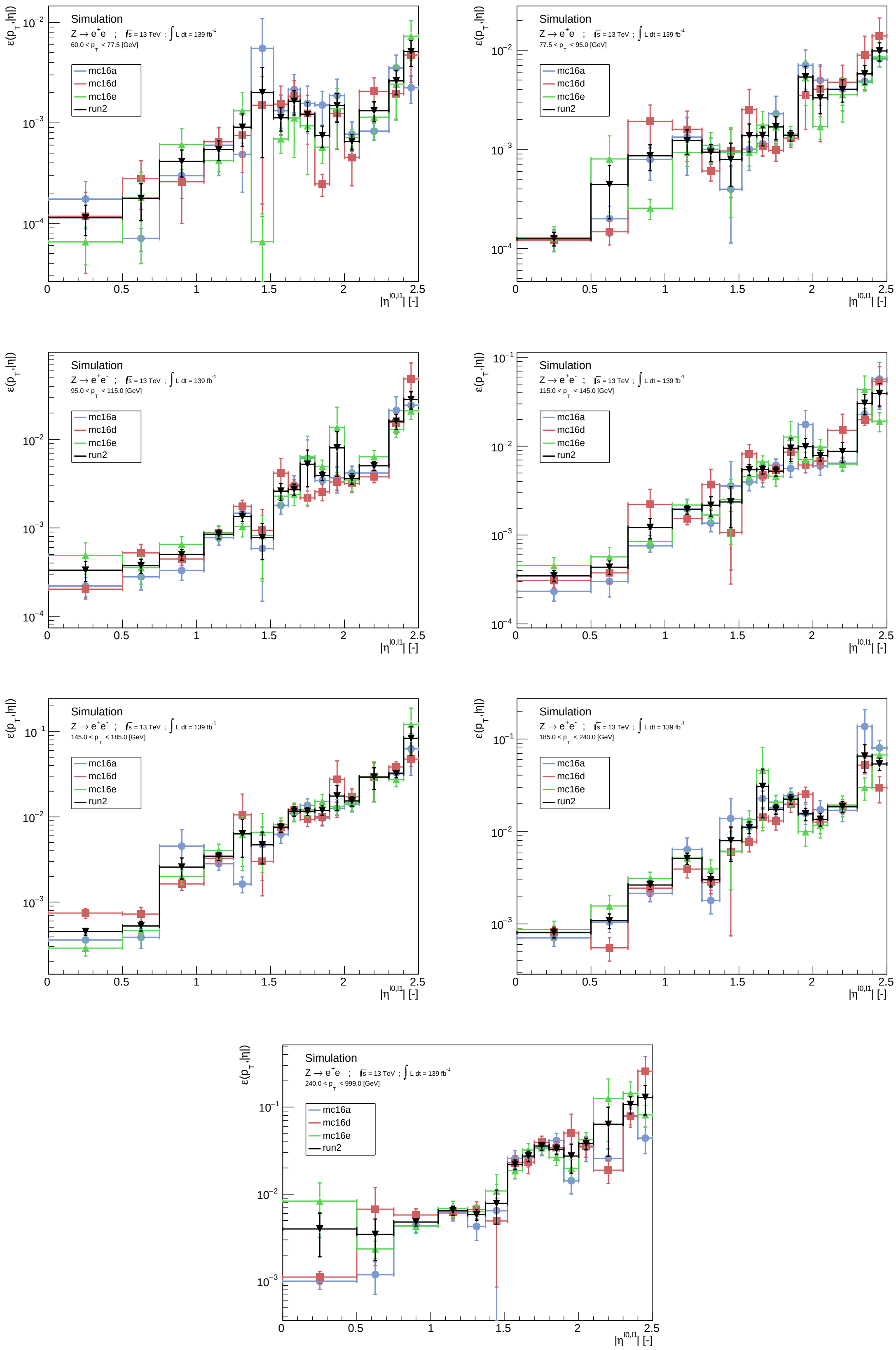


Figure C.14: Comparison of charge misidentification probabilities for mc16a (shown in blue), mc16d (shown in red) and mc16e (shown in green) Monte Carlo campaigns along with distributions for the full Run 2 set (black) for pseudorapidity bins within $p_T > 60$ GeV after the ECIDS selection. The distributions were computed by exploiting the truth information in $Z \rightarrow ee$ Madgraph MC samples. Reconstructed electrons originated from FSR were removed for the computation and the scale factors shown in Figure C.4 were applied. Only statistical uncertainties are shown. The region $1.37 < |\eta| < 1.52$ is not considered in the analysis.

C.8 Systematic variations on charge misidentification probabilities using ECIDS

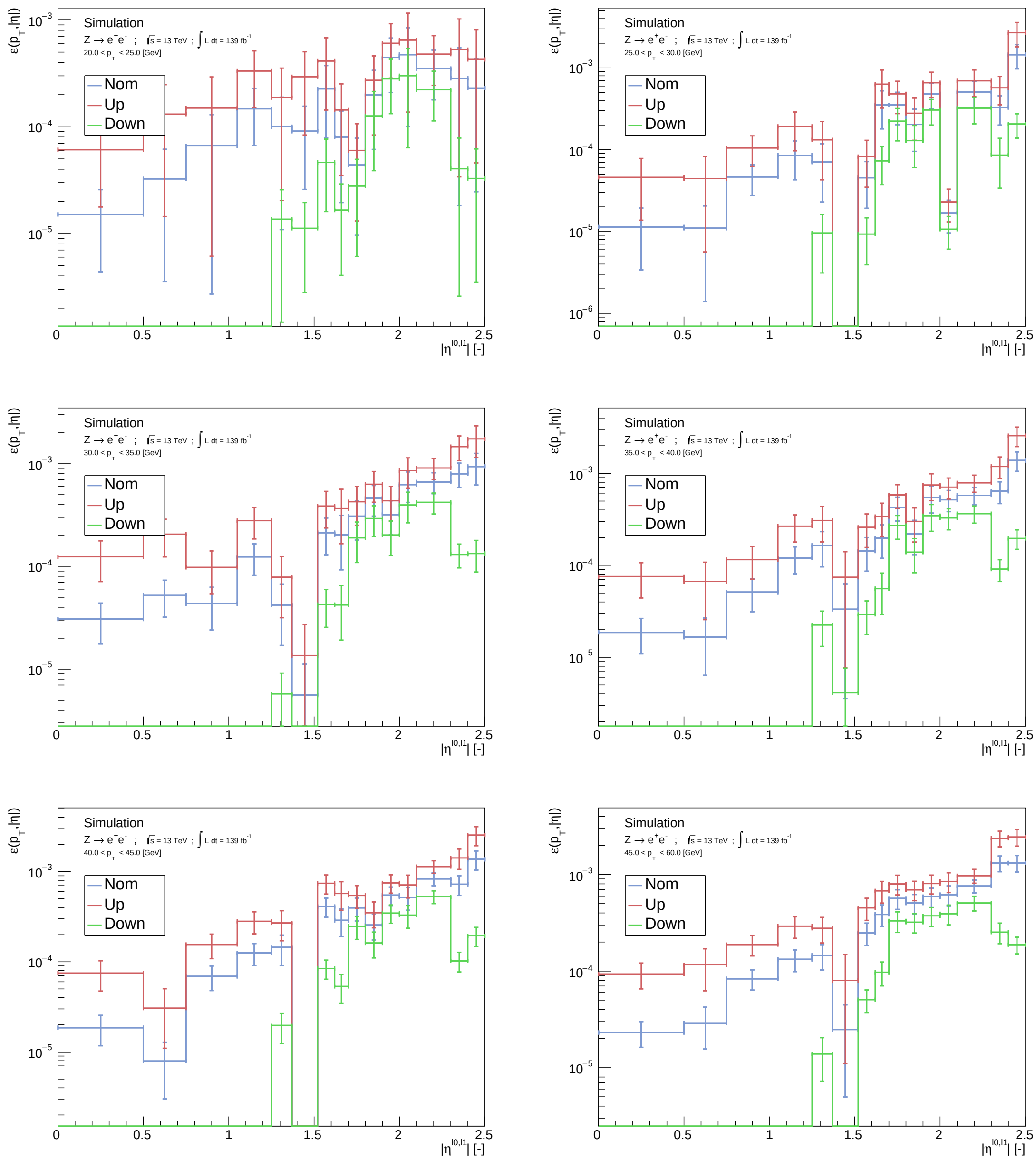


Figure C.15: Systematic variations on the scale factors for the computation of the charge misidentification probabilities using the full Run 2 sample in p_T bins within $20 < p_T < 60$ GeV. Reconstructed electrons originated from FSR were removed for the computation and the scale factors shown in Figure 7.1 were applied. Only statistical uncertainties are shown. The region $1.37 < |\eta| < 1.52$ is not considered.

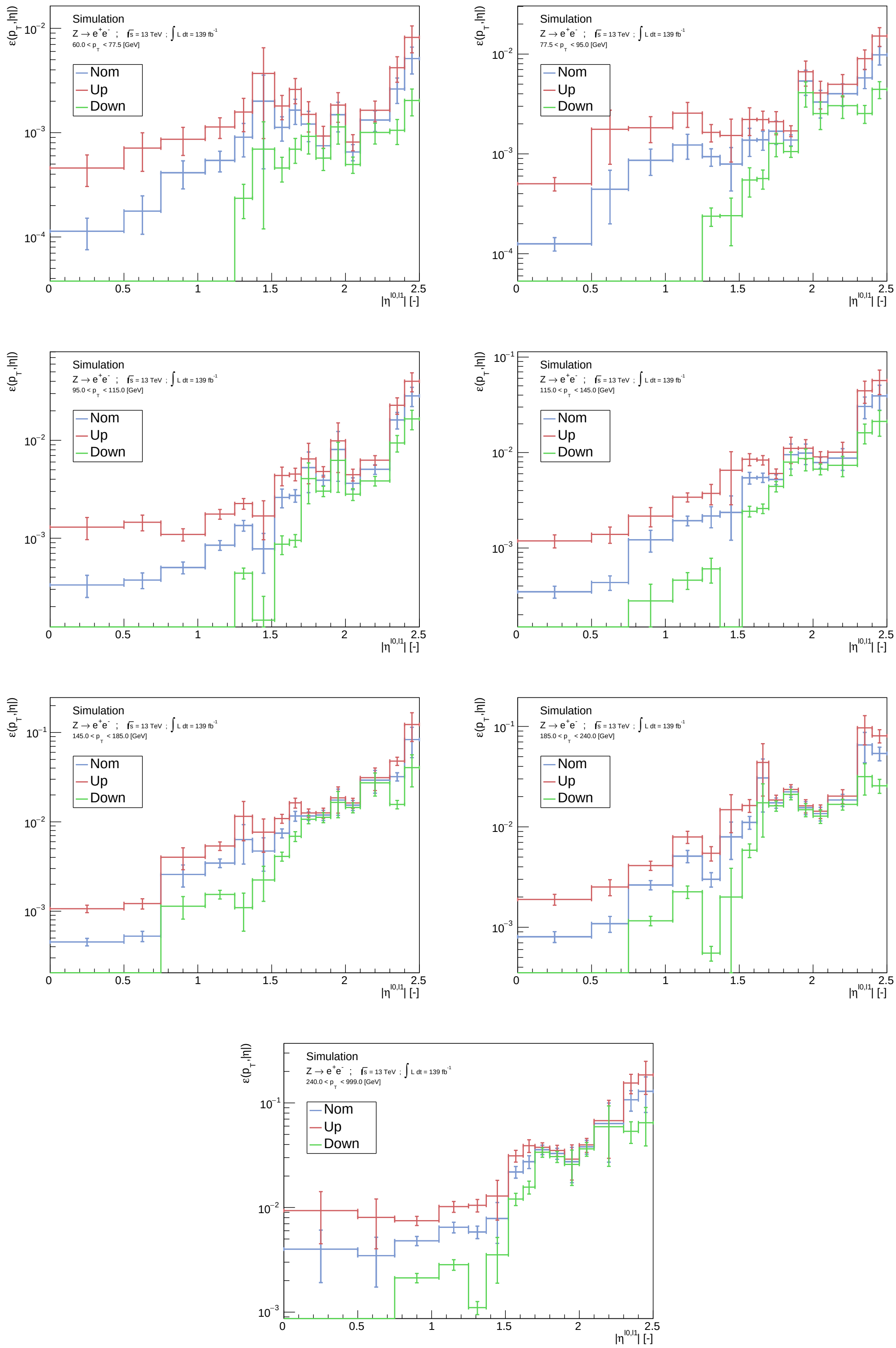


Figure C.16: Systematic variations on the scale factors on the charge misidentification probabilities for the full Run 2 sample in p_T bins within $p_T > 60$ GeV. Reconstructed electrons originated from FSR were removed for the computation and the scale factors shown in Figure 7.1 were applied. Only statistical uncertainties are shown. The region $1.37 < |\eta| < 1.52$ is not considered.

C.9 Additional distributions for the $t\bar{t}$ test in MC using ECIDS

ee	mc16a [raw]	mc16d [raw]	mc16e [raw]	run2 [raw]
ttbarSS	12.4 ± 0.7 [330]	15.9 ± 0.8 [423]	17.9 ± 0.9 [478]	46.2 ± 1.4 [1231]
ttbarOS	10.40 ± 0.03 [436321]	11.99 ± 0.04 [506800]	16.01 ± 0.04 [665734]	38.40 ± 0.07 [1.60×10^6]
%	16.1 ± 4.7	24.7 ± 3.8	10.5 ± 4.3	16.9 ± 2.5

em	mc16a [raw]	mc16d [raw]	mc16e [raw]	run2 [raw]
ttbarSS	115.9 ± 2.2 [3021]	149.7 ± 2.5 [4023]	197 ± 3 [5203]	463 ± 4 [12247]
ttbarOS	115.7 ± 0.2 [2.63×10^6]	137.9 ± 0.3 [3.18×10^6]	183.7 ± 0.2 [4.17×10^6]	437.4 ± 0.4 [9.99×10^6]
%	0.18 ± 1.88	7.8 ± 1.5	6.7 ± 1.4	5.4 ± 0.9

me	mc16a [raw]	mc16d [raw]	mc16e [raw]	run2 [raw]
ttbarSS	36.8 ± 1.2 [967]	47.2 ± 1.4 [1249]	59.5 ± 1.6 [1581]	143.6 ± 2.4 [3797]
ttbarOS	32.3 ± 0.1 [2.39×10^6]	38.4 ± 0.1 [2.85×10^6]	51.1 ± 0.1 [3.74×10^6]	121.8 ± 0.1 [8.98×10^6]
%	12 ± 3	18.5 ± 2.4	14.1 ± 2.3	15.2 ± 1.5

Table C.2: Integrated number of events after the ECIDS selection for OS and SS event selections in the Z window CR along different MC campaigns mc16a (corresponding to 2015-2016 dataset), mc16d (corresponding to 2017 dataset), mc16e (2018 dataset), and full run-2 (2015-2018 dataset) for the closure test using $t\bar{t}$ MC in the ee, $e\mu$ and μe channels. The requirement $|m_{ee} - m_Z| < 15$ GeV is not applied. The raw number of events is included in squared brackets. The deviation between both selections is shown in percentage on the third row, whose uncertainty is propagated from the statistical uncertainties of the yields.

List of Figures

List of Tables

References

[1] Andrew Purcell. Go on a particle quest at the first CERN webfest. Le premier webfest du CERN se lance à la conquête des particules (https://cds.cern.ch/record/1473657). (BUL-NA-2012-269. 35/2012):10, Aug 2012.

[2] E. Noether. Invariant variation problems. *Nachr. d. Koenig. Gesellsch.d. Wiss. zu Goettingen, Math-phys. Klasse*, page 235–257, 1918.

[3] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental Test of Parity Conservation in Beta Decay. *Phys. Rev.*, 105:1413–1414, 1957.

[4] Nicola Cabibbo. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10:531–533, 1963. [,648(1963)].

[5] Makoto Kobayashi and Toshihide Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[6] M. Banner et al. Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider. *Phys. Lett.*, 122B:476–485, 1983. [,7.45(1983)].

[7] G. Arnison et al. Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV. *Phys. Lett.*, 122B:103–116, 1983. [,611(1983)].

[8] G. Arnison et al. Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider. *Phys. Lett.*, 126B:398–410, 1983. [,7.55(1983)].

[9] P. Bagnaia et al. Evidence for $Z^0 \to e^+e^-$ at the CERN anti-pp Collider. *Phys. Lett.*, 129B:130–140, 1983. [,7.69(1983)].

[10] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, 1964. [,157(1964)].

[11] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964. [,160(1964)].

[12] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett.*, 13:585–587, 1964. [,162(1964)].

[13] Georges Aad et al. Evidence for the spin-0 nature of the Higgs boson using ATLAS data. *Phys. Lett.*, B726:120–144, 2013.

[14] Serguei Chatrchyan et al. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[15] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev.*, D98(3):030001, 2018.

[16] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.

[17] F. Halzen and Alan D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. 1984.

[18] F. J. Hasert et al. Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment. *Phys. Lett.*, 46B:138–140, 1973. [,5.15(1973)].

[19] Summary plots from the ATLAS Standard Model physics group (https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2019-024/).

[20] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, 1998.

[21] F. Zwicky. On the Masses of Nebulae and of Clusters of Nebulae. *Astrophys. J.*, 86:217–246, 1937.

[22] Vera C. Rubin and W. Kent Ford, Jr. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *Astrophys. J.*, 159:379–403, 1970.

[23] J. Silk et al. *Particle Dark Matter: Observations, Models and Searches*. Cambridge Univ. Press, Cambridge, 2010.

[24] Michael Dine and Alexander Kusenko. The Origin of the matter - antimatter asymmetry. *Rev. Mod. Phys.*, 76:1, 2003.

[25] R. D. Peccei. The Strong CP problem and axions. *Lect. Notes Phys.*, 741:3–17, 2008. [,3(2006)].

[26] Eldad Gildener. Gauge Symmetry Hierarchies. *Phys. Rev.*, D14:1667, 1976.

[27] G. J. Gounaris, R. Kogerler, and H. Neufeld. Relationship Between Longitudinally Polarized Vector Bosons and their Unphysical Scalar Partners. *Phys. Rev.*, D34:3257, 1986.

[28] Michael S. Chanowitz and Mary K. Gaillard. The TeV Physics of Strongly Interacting W's and Z's. *Nucl. Phys.*, B261:379–431, 1985.

[29] Ilaria Brivio and Michael Trott. The Standard Model as an Effective Field Theory. *Phys. Rept.*, 793:1–98, 2019.

[30] I. Abt et al. The H1 detector at HERA. *Nucl. Instrum. Meth.*, A386:310–347, 1997.

[31] ZEUS Status Report 1993 (https://www-zeus.desy.de/bluebook/bluebook.html).

[32] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov. Phys. JETP*, 46:641–653, 1977. [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[33] Vladimir Naumovich Gribov and L N Lipatov. Deep inelastic ep scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15(4):438–450, 1972.

[34] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys.*, B126:298–318, 1977.

[35] T. Gleisberg, Stefan. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter. Event generation with SHERPA 1.1. *JHEP*, 02:007, 2009.

[36] Stephen Mrenna and Peter Richardson. Matching matrix elements and parton showers with HERWIG and PYTHIA. *JHEP*, 05:040, 2004.

[37] Johan Alwall et al. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J.*, C53:473–500, 2008.

[38] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006.

[39] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015.

[40] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber. HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP*, 01:010, 2001.

[41] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5 : Going Beyond. *JHEP*, 06:128, 2011.

[42] The ATLAS Collaboration webpage (`https://atlas.cern/discover/collaboration`).

[43] Oliver Sim Brüning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. *LHC Design Report*, volume 1 of *CERN Yellow Reports: Monographs*. CERN, Geneva, 2004.

[44] LEP Design Report: Vol.2. The LEP Main Ring. 1984.

[45] The ATLAS Collaboration. Public results (`https://twiki.cern.ch/twiki/bin/view/AtlasPublic/WebHome#Detector`).

[46] Michael Benedikt, Paul Collier, V Mertens, John Poole, and Karlheinz Schindl. *LHC Design Report*, volume 3 of *CERN Yellow Reports: Monographs*. CERN, Geneva, 2004.

[47] Ewen H. Maclean. Accelerator Physics and the LHC. (HASCO Summer School 2019).

[48] Fabienne Marcastel. CERN's Accelerator Complex. La chaîne des accélérateurs du CERN (`https://cds.cern.ch/record/1621583`). Oct 2013. General Photo.

[49] Joao Pequenao. Computer generated image of the whole ATLAS detector (`https://cds.cern.ch/record/1095924`). Mar 2008.

[50] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[51] Studies of the ATLAS Inner Detector material using $\sqrt{s}$ =13 TeV pp collision data. Technical Report ATL-PHYS-PUB-2015-050, CERN, Geneva, Nov 2015.

[52] *ATLAS inner detector: Technical Design Report, 1*. Technical Design Report ATLAS. CERN, Geneva, 1997.

[53] ATLAS inner detector: Technical design report. Vol. 2. 1997.

[54] M Capeans, G Darbo, K Einsweiller, M Elsing, T Flick, M Garcia-Sciveres, C Gemme, H Pernegger, O Rohne, and R Vuillermet. ATLAS Insertable B-Layer Technical Design Report. Technical Report CERN-LHCC-2010-013. ATLAS-TDR-19, Sep 2010.

[55] Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV. Technical Report ATL-PHYS-PUB-2015-018, CERN, Geneva, Jul 2015.

[56] A Abdesselam and T Akimoto. The Barrel Modules of the ATLAS SemiConductor Tracker. Technical Report ATL-INDET-PUB-2006-005. ATL-COM-INDET-2006-009. CERN-ATL-COM-INDET-2006-009, CERN, Geneva, Jul 2006.

[57] Aranzazu Ruiz-Martinez and ATLAS Collaboration. The Run-2 ATLAS Trigger System. Technical Report ATL-DAQ-PROC-2016-003, CERN, Geneva, Feb 2016.

[58] T. Cornelissen, M. Elsing, I. Gavrilenko, W. Liebig, E. Moyse, and A. Salzburger. The new ATLAS track reconstruction (NEWT). *J. Phys. Conf. Ser.*, 119:032014, 2008.

[59] Morad Aaboud et al. Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC. *Eur. Phys. J.*, C77(5):332, 2017.

[60] Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV. Technical Report ATL-PHYS-PUB-2015-026, CERN, Geneva, Jul 2015.

[61] Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment. Technical Report ATLAS-CONF-2012-042, CERN, Geneva, Mar 2012.

[62] Morad Aaboud et al. Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s}$ = 13 TeV. *Eur. Phys. J.*, C79(8):639, 2019.

[63] Georges Aad et al. Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data. *JINST*, 14(12):P12006, 2019.

[64] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal. Calorimeter clustering algorithms: Description and performance. 2008.

[65] Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach. Technical Report ATL-PHYS-PUB-2017-022, CERN, Geneva, Dec 2017.

[66] Morad Aaboud et al. Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data. *JINST*, 14(03):P03017, 2019.

[67] Morad Aaboud et al. Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data. *Eur. Phys. J.*, C77(3):195, 2017.

[68] Christos Anastopoulos, Jean-Francois Arguin, Cyril Becot, Olympia Dartsi, Lucia Di Ciaccio, Alexandra Fell, Lucas Macrorie Flores, Dominique Godin, Sarah Heim, Joe Kroll, Yanwen Liu, Kristin Lohwasser, Kazuya Mochizuki, Emmanuel Monnier, Miha Muskinja, Andreas Petridis, Pavel Podberezko, Thomas Dennis Powell, Pascal Pralavorio, Joseph Reichert, Elias Michael Ruettinger, Abhishek Sharma, Philip Sommer, Grigore Tarna, Savannah Jennifer Thais, Evelyn Thomson, Artur Trofymov, and Hanlin Xu. Electron identification and efficiency measurements in 2017 data. Technical Report ATL-COM-PHYS-2018-1727, CERN, Geneva, Jan 2019.

[69] Georges Aad et al. Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV. *Eur. Phys. J.*, C76(5):292, 2016.

[70] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[71] Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector. Technical Report ATLAS-CONF-2015-029, CERN, Geneva, Jul 2015.

[72] Georges Aad et al. Performance of b-Jet Identification in the ATLAS Experiment. *JINST*, 11(04):P04008, 2016.

[73] Georges Aad et al. ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J.*, C79(11):970, 2019.

[74] Pile-up Suppression in Missing Transverse Momentum Reconstruction in the ATLAS Experiment in Proton-Proton Collisions at sqrts = 8 TeV. Technical Report ATLAS-CONF-2014-019, CERN, Geneva, May 2014.

[75] Morad Aaboud et al. Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV. *Eur. Phys. J.*, C78(11):903, 2018.

[76] Philipp Anger. *Probing Electroweak Gauge Boson Scattering with the ATLAS Detector at the Large Hadron Collider*. PhD thesis, Dresden, Tech. U., 2014-06-23.

[77] Christian Gumpert. Measurement of Electroweak Gauge Boson Scattering in the Channel $pp \to W^{\pm}W^{\pm}jj$ with the ATLAS Detector at the Large Hadron Collider, Dec 2014. Presented 27 Feb 2015.

[78] Francisco Campanario, Matthias Kerner, Le Duc Ninh, and Dieter Zeppenfeld. Next-to-leading order QCD corrections to W^+W^+ and W^-W^- production in association with two jets. *Phys. Rev.*, D89(5):054009, 2014.

[79] Benedikt Biedermann, Ansgar Denner, and Mathieu Pellen. Complete NLO corrections to W^+W^+ scattering and its irreducible background at the LHC. *JHEP*, 10:124, 2017.

[80] Georges Aad et al. Evidence for Electroweak Production of $W^{\pm}W^{\pm}jj$ in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector. *Phys. Rev. Lett.*, 113(14):141803, 2014.

[81] Morad Aaboud et al. Measurement of $W^{\pm}W^{\pm}$ vector-boson scattering and limits on anomalous quartic gauge couplings with the ATLAS detector. *Phys. Rev.*, D96(1):012007, 2017.

[82] Morad Aaboud et al. 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):161801, 2019.

[83] Vardan Khachatryan et al. Study of vector boson scattering and search for new physics in events with two same-sign leptons and two jets. *Phys. Rev. Lett.*, 114(5):051801, 2015.

[84] Albert M Sirunyan et al. 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):081801, 2018.

[85] B Heinemann, M Hurwitz, and L Skinnari. Search for Anomalous Production of Prompt Like-sign Muon Pairs and Constraints on Physics Beyond the Standard Model. Technical Report ATL-COM-PHYS-2011-1349, CERN, Geneva, Oct 2011.

[86] G. Aad et al. The ATLAS Simulation Infrastructure. *Eur. Phys. J.*, C70:823–874, 2010.

[87] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[88] Morad Aaboud et al. Measurement of the cross-section for producing a W boson in association with a single top quark in pp collisions at $\sqrt{s} = 13$ TeV with ATLAS. *JHEP*, 01:063, 2018.

[89] Georges Aad et al. Performance of electron and photon triggers in ATLAS during LHC Run 2. *Eur. Phys. J.*, C80(1):47, 2020.

[90] D Adams and et. al. Anastopoulos. Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014. Technical Report ATL-PHYS-INT-2014-018, CERN, Geneva, Jul 2014.

[91] Carsten Bittrich, William Kennedy Di Clemente, Emily Marie Duffield, Cong Geng, Giulia Gonella, Jun Guo, Beate Heinemann, Tim Herrmann, Franziska Iltzsche Speiser, Michael Kobel, Karsten Koeneke, Joseph Ira Kroll, Shu Li, Jianbei Liu, Yanlin Liu, Joany Andreina Manjarres Ramos, Claire Alexandra Lee, Monika Mittal, Chilufya Mwewa, Rustem Ospanov, Simone Pagan Griso, Karolos Jozef Potamianos, Marjorie Shapiro, Philip Sommer, Stefanie Todt, Yusheng Wu, Wenhao Xu, Sahal Yacoob, Haijun Yang, Liqing Zhang, Zhengguo Zhao, Bing Zhou, Junjie Zhu, and Marc-Andre Pleier. Support note for measurement of electroweak $W^{\pm}W^{\pm}jj$ production at $\sqrt{s} = 13$ TeV. Technical Report ATL-COM-PHYS-2018-252, CERN, Geneva, Mar 2018.

[92] Jean-Francois Arguin, Jerome Claude, Giulia Gonella, Borut Paul Kersevan, Karsten Koeneke, Kristin Lohwasser, Kazuya Mochizuki, Miha Muskinja, Caroline Piaulet, and Philip Sommer. Support Note for Electron ID: charge mis-identification. Technical Report ATL-COM-PHYS-2017-1014, CERN, Geneva, Jul 2017.

[93] Claus Grupen and Boris Schwartz. *Particle detectors*, volume 26. Cambridge Univ. Pr., Cambridge, UK, 2008.

[94] J Almond, L Bianchini, C Blocker, J Klinger, A Maslennikov, M Owen, K Skovpen, Y Tikhonov, and U Yang. Search for heavy neutrino, W_R and Z_R gauge bosons in events with two high-P_T leptons and jets with the ATLAS detector in pp collisions at $\sqrt{s}$ = 8 TeV. Technical Report ATL-COM-PHYS-2013-810, CERN, Geneva, Jun 2013. This is the supporting note for EXOT-2012-24, which is now published as JHEP07 (2015) 162.

[95] Robert D. Cousins, James T. Linnemann, and Jordan Tucker. Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process. *Nucl. Instrum. Meth.*, A595(2):480–501, 2008.

[96] Philip Sommer. Data-driven estimation of background from electron charge misidentification (https://indico.cern.ch/event/861548/contributions/3631706/attachments/1940227/3216722/present.pdf).

Acknowledgements

Writing a scientific thesis is a challenging experience that requires the guidance of experienced people. I wish to express my deepest gratitude to Prof. Dr. Beate Heinemann, who gave me the opportunity to write this thesis. I did not only learned experimental particle physics during our conversations, but I also grew personally and as a physicist.

I am indebted to Dr. Oleg Kuprash for his wholehearted support and patience during this year. Without his help, the goal of this project would not have been realized. I also want to thank Shalu Solomon for her unfailing support and enthusiastic discussions. I am happy to have worked with such a talented woman, whose contagious motivation is now part of me. I also wish to thank all the people in the same-sign WW analysis group, especially Karolos Potamianos and Cyril Becot whose assistance was crucial in the completion of this work.

I would like to pay my special regards to Giulia Gonella and Philip Sommer for their assistance and fruitful discussions on the technical details of the analysis and the charge misidentification phenomenon. Furthermore, my sincere gratitude to Andrea Knue and Thorwald Klapdor-Kleingrothaus. Your comments and remarks were of great help for the development of this work.

I also want to thank the experts and developers of the Common Analysis Framework, and especially Ralf Gugel, for his unconditioned help in the use of this software.

The studies presented in this thesis also take advantage of the work carried out by thousands of scientists that directly or indirectly contributed to the results achieved by the ATLAS Collaboration. This mention is well deserved.

I am grateful to the University of Freiburg for giving me the opportunity to study this extraordinary program and offering me all its facilities in full. This work was carried out using the Black Forest Grid. I want to thank their administrators and technical staff for keeping this facility in its best performance.

I wish to acknowledge the support and great love of my mother. You are my reference in effort and perseverance, and I would not be writing these words without your unwavering support. Special mention to Juan Fierro, whom I am lucky to have as a friend. I can not express how much I owe you. To all my friends and colleagues in the department, with special mention to Steffen Ludwig, Felicia Volle, Thea Engler, Kim Heidegger, Vlad Plesanovs and Zuzana Rurikova. Thank you all for such a lovely time in this group.