# Observation of Longitudinal-Longitudinal Interactions at High $p_T^Z$ and Studies of the Radiation Amplitude Zero Effect in $W^\pm Z$ Events with the ATLAS Detector at the Large Hadron Collider

**Dissertation**

zur Erlangung des akademischen Grades
Doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt der
Fakultät für Physik
der Technischen Universität Dresden


von
M.Sc. Jan-Eric Nitschke
geboren am 07.12.1995 in Dresden

Eingereicht am 1. Februar 2024

Verteidigt am 17. April 2024

| | |
|---|---|
| 1. Gutachter: | Prof. Dr. Michael Kobel |
| 2. Gutachter: | Dr. Frank Siegert |
| 3. Gutachter: | Prof. Dr. Joany Manjarrés |

## Zusammenfassung


Die Polarisation von Vektorbosonen bietet einzigartige Möglichkeiten zur Erforschung des Standardmodells (SM) der Teilchenphysik und Effekte neuer Physik. Zwei solcher Ansätze untersuchen exakte Aufhebungen bei hohen Energien in $W^{\pm}Z$-Ereignissen, bei denen beide Bosonen transversal polarisiert sind (TT), und die Energieabhängigkeit des Anteils der Ereignisse, bei denen beide Bosonen longitudinal polarisiert sind (00). Ersteres ist als "radiation amplitude zero" (RAZ) Effekt bekannt.

Um dies zu erreichen, werden Ereignisse mit leptonisch zerfallenden $W^{\pm}Z$-Bosonen aus den Run 2 Daten des ATLAS-Experiments am LHC ausgewählt. Diese Daten wurden bei einer Schwerpunktsenergie von $\sqrt{s} = 13\,\text{TeV}$ aufgezeichnet und entsprechen einer integrierten Luminosität von $140\,\text{fb}^{-1}$. Der Anteil der doppelt longitudinal polarisierten Ereignisse in einem Bereich mit $p_T^{WZ} < 70\,\text{GeV}$ und $100 < p_T^Z \leq 200\,\text{GeV}$ wir mit $18.9\ \pm^{3.4}_{3.3}\,(\text{stat.})\ \pm^{2.4}_{2.2}\,(\text{syst.})\,\%$ gemessen. Wenn der $p_T^Z$-Schnitt auf $p_T^Z > 200\,\text{GeV}$ geändert wird, beträgt der Wert $13\ \pm^{9}_{8}\,(\text{stat.})\ \pm^{2}_{2}\,(\text{syst.})\,\%$. Ersteres geht mit einer Signifikanz von $5.2\sigma$ gegenüber der Nullhypothese, dass kein 00-Zustand existiert, einher. Wird die Messung mit einem eingeschränkteren Modell durchgeführt, werden auch im Bereich $p_T^Z > 200\,\text{GeV}$ Hinweise auf den 00-Zustand mit einer Signifikanz von $3.2\sigma$ gefunden. Unter Verwendung von $100 < p_T^Z \leq 200\,\text{GeV}$ wird eine untere Schranke von $2.8\sigma$ für die Signifikanz der Zurückweisung der Hypothese gefunden, dass die Anteile der longitudinal polarisierten $W^{\pm}$- und $Z$-Bosonen unabhängig voneinander sind.

Der RAZ-Effekt wird durch die Entfaltung der Größen $\Delta Y(WZ)$ und $\Delta Y(\ell_W Z)$ für den TT-Polarisationszustand und die Summe der Polarisationen in Regionen untersucht, in denen $p_T^{WZ}$ kleiner als $10, 20, 40$, oder $70\,\text{GeV}$, oder uneingeschränkt ist. Hier wird der Effekt als eine Verringerung des differentiellen Wirkungsquerschnitts in der Nähe von $\Delta Y = 0$ realisiert. Um die Größe dieser Absenkungen zu quantifizieren wird eine Tiefenvariable ($D$) eingeführt. Für den TT-Zustand wird ihr Wert in der Region mit $p_T^{WZ} < 70\,\text{GeV}$ als $0,68 \pm 0,10$ und $0,29 \pm 0,06$ gemessen, wenn $\Delta Y(WZ)$ bzw. $\Delta Y(\ell_W Z)$ verwendet werden. Ersteres entspricht einer geschätzten Signifikanz von $D/\Delta D = 7.0$ und weist auf eine genaue Messung der Tiefe hin.

## Abstract

Polarization of vector bosons offers unique avenues for exploring the Standard Model (SM) of particle physics and effects going beyond. Two such approaches investigate delicate gauge cancellations at high energies in $W^{\pm}Z$ events where both bosons are transversally polarized (TT) and the energy dependence of the fraction of events where both bosons are longitudinally polarized (00). The former is known as the radiation amplitude zero (RAZ) effect.

To achieve this, events with leptonically decaying $W^{\pm}Z$ bosons are selected in the Run 2 data of the ATLAS experiment at the LHC. This data was recorded at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ and encompasses an integrated luminosity of $140\,\text{fb}^{-1}$. The fraction of doubly longitudinally polarized events in a region with $p_T^{WZ} < 70\,\text{GeV}$ and $100 < p_T^Z \leq 200\,\text{GeV}$ is measured to be $18.9 \pm^{3.4}_{3.3}\,(\text{stat.}) \pm^{2.4}_{2.2}\,(\text{syst.})\,\%$. When the $p_T^Z$ cut is changed to $p_T^Z > 200\,\text{GeV}$, the value is determined to be $13 \pm^{9}_{8}\,(\text{stat.}) \pm^{2}_{2}\,(\text{syst.})\,\%$. The former is accompanied by a significance of $5.2\sigma$ over the null hypothesis that no 00 state exists. If the measurement is performed with a more restricted model, evidence for the 00 state is also found in the $p_T^Z > 200\,\text{GeV}$ region with a significance of $3.2\sigma$. Using $100 < p_T^Z \leq 200\,\text{GeV}$, a lower bound of $2.8\sigma$ is determined for the significance of rejecting the hypothesis that the fractions of longitudinally polarized $W^{\pm}$ and $Z$ bosons are independent of each other.

The RAZ effect is probed through the unfolding of the $\Delta Y(WZ)$ and $\Delta Y(\ell_W Z)$ observables for the TT polarization state and the sum of polarizations in regions requiring $p_T^{WZ}$ to be smaller than $10, 20, 40$, or $70\,\text{GeV}$, or unconstrained. Here, the effect is realized as a reduction of the differential cross-section around $\Delta Y = 0$. A depth variable ($D$) is introduced to quantify the size of this dip and measured to be $0.68 \pm 0.10$ and $0.29 \pm 0.06$ for the TT state in the region with $p_T^{WZ} < 70\,\text{GeV}$ using $\Delta Y(WZ)$ and $\Delta Y(\ell_W Z)$, respectively. The former represents an estimated significance of $D/\Delta D = 7.0$ and indicates a precise measurement of the depth.

# Contents

# Chapter 1

# Introduction

The strive for knowledge and an understanding of the world we live in has long been a driving force for human endeavors. People as early as Thales of Miletus (around 600 BC) or Hypatia of Alexandria (around 400 AD) have been documented studying subjects like physics, biology, astronomy, and mathematics. In modern times, the study of the universe and all of its properties has become more formalized, and those who examine the fundamental building blocks of reality and their interactions are called physicists. Some of the most renowned ones in scientific history include Isaac Newton [1] for groundbreaking contributions to classical mechanics and gravity, Marie SkodowskaCurie [2] for pioneering research on radioactivity, and Stephen Hawking [3] for significant advancements in the understanding of black holes.

The current status of fundamental physics is such that two theories together describe all four known fundamental interactions in the universe. The first is Einstein's "General relativity" (GR), which describes gravity as a curvature of spacetime caused by the presence of matter and energy [4]. The second theory is the "Standard Model" of particle physics (SM), which describes the electromagnetic, weak, and strong nuclear forces. One major success of the SM was the prediction of the Higgs boson, which was discovered in 2012 by the ATLAS and CMS collaborations at the Large Hadron Collider (LHC) [5, 6].

Despite this success of the Standard Model, there are still many unresolved questions and mysteries in fundamental physics. For instance, there is currently no complete theory that unifies GR with the Standard Model [7–9], which is a major goal of modern theoretical physics. Additionally, the existence of dark matter [10] and dark energy [11], which make up a significant portion of the universe's mass and energy, remain unexplained. Other puzzles include the origin of neutrino masses [12], the observed asymmetry between matter and antimatter in the universe [13, 14], and the anomalous mass of the W boson [15, 16]. There is also the hierarchy problem [17], which refers to the large discrepancy between the energy scales at which gravity and the other fundamental forces operate. Finally, the question of quantum triviality [18], which concerns the consistency and stability of the Standard Model at very high energies, remains an open and active area of research. These open questions indicate that it is important to perform further measurements and probes of the Standard Model to gather more hints about where a more complete theory may lie.

One avenue for study is a rather direct probe of electroweak symmetry breaking (EWSB) via weak boson polarizations. As longitudinal polarization modes only exist for massive bosons, they are directly linked to the Higgs mechanism, which provides the $W^{\pm}$ and $Z$ bosons with their mass.

The holy grail of weak boson polarization measurements would be the direct $V_L V_L \to V_L V_L$ scattering. However, as there is currently no polarized weak boson accelerator and the short lifetime of the weak bosons makes direct detection impossible with current means, the next best approach is the extraction of polarization fractions in vector boson scattering processes at the LHC. As these processes themselves are also incredibly rare and have only first been observed in 2018/2019 by the ATLAS and CMS collaborations [19–21], a significant measurement of joint boson polarization fractions has so far been unsuccessful [22].

On the other side of the spectrum are the basic but vital steps of measuring polarization in single and multi-boson final states [23–25]. The most recent measurement by the ATLAS experiment has, for the first time, been able to measure all four combinations of transverse and longitudinal polarization in $W^{\pm}Z$ diboson final states [26].

The work presented in this thesis represents another step from these inclusive final states toward the ultimate goal of getting a better understanding of EWSB via vector boson polarization. For that, a specific phase space enhancing the fraction of events with two longitudinally polarized bosons is studied alongside regions designed to investigate the radiation amplitude zero effect [27, 28].

Towards that, Chapter 2 of this work lays the theoretical foundation of the SM and polarization. Chapter 3 describes the experimental setup of the LHC and the ATLAS detector used to obtain the data presented. The simulation techniques and results used to generate the theory predictions are presented in Chapter 4. The specific selections making up the unique phase space used in this work are described in Chapter 5. With all required ingredients in place, Chapter 6 shows a qualitative validation of the signal samples. Systematic uncertainties described in Chapter 7 and a multivariate classifier presented in Chapter 8 are used for the joint boson fraction measurement shown in Chapters 9 and 10. A description of the process of unfolding data distributions, along with the results of applying that procedure to variables sensitive to the radiation amplitude zero effect, makes up Chapter 11. Chapter 12 provides a conclusion of this work and an outlook on future studies.

# Chapter 2

# Theoretical Foundations

All science aims to create and validate theories and models that describe observed nature and can predict yet not seen aspects of it.

The theory that best does so in the realm of high-energy particle physics is the Standard Model of particle physics (SM). Formulated as a gauge quantum field theory, the SM describes the fundamental set of particles - leptons, gauge bosons, and the Higgs boson - as well as their interactions via the electromagnetic, weak, and strong forces.

This chapter discusses the theoretical foundations underpinning the work of this thesis, beginning with an overview of the SM, its mathematical description and particle content.

This thesis aims to measure $W^{\pm}$ and $Z$ bosons and their polarization. Towards that, Section 2.1.2 discusses the electroweak sector and electroweak symmetry breaking (EWSB) more in-depth. Finally, Section 2.2 explains the intricacies of polarization in these processes.

## 2.1 Standard Model of particle physics

The Standard Model was built up in many stages, starting with Pauli's [29] work on the interaction of particles with the electromagnetic field. This was expanded on by Yang and Mills [30] with the introduction of non-abelian gauge theories in search of a theory describing the strong interaction.

The term"Standard Model" was first used by Pais and Treiman [31] in 1975 to refer to the electroweak theory with four quarks. This was contributed to by Glashow [32] by combining the electromagnetic and weak interactions in 1961 as well as Weinberg [33] and Salam [34] which incorporated the Higgs mechanism [35–37]. The SM was completed to its modern form through the theory of quantum chromodynamics (QCD), which describes the strong interactions [38–40]. Tables 2.1 and 2.2 summarize the particle content of the SM, respectively showing the fermions and bosons currently known to exist. All of the listed particles with electric charge $\neq 0$ have a corresponding antiparticle with flipped charges but identical mass[1].

Many excellent introductions into the SM exist [45–49]. This section summarizes the key points relevant to this thesis based on these references, focusing on Ref. [45].

### 2.1.1 Gauge Theories

The Standard Model is defined by its Lagrangian. This is a formula containing the particle content as well as the interactions these particles take part in. From it, the equations of motion can be derived using the Euler-Lagrange equations. As a gauge theory, the Standard Model Lagrangian, and thus all observable quantities, are invariant under a specific set of local transformations of the kind:

$$\psi(\vec{x}, t) \rightarrow \psi'(\vec{x}, t) = e^{i\chi(\vec{x},t)}\psi, \tag{2.1}$$

[1] It is not yet known whether neutrinos are their own antiparticles or if antineutrinos are distinct [41, 42].

**Table 2.1:** Properties of left-handed Standard Model fermions. Each one has a spin of $\frac{1}{2}$ and an antiparticle with the same mass but all charges are reversed. Shown are the particles' masses, electric charge $Q$, and the third component of the weak isospin $T_3$ for the left-handed variant. The weak isospin of right-handed particles and left-handed antiparticles is 0 instead. The only exception is neutrinos, where right-handed particles and left-handed antiparticles have not been observed. While the observation of neutrino oscillation [43] requires at least two neutrino mass eigenstates to have non-zero mass, the exact values are not yet known. Instead, only upper limits have been measured. The values quoted here are determined as the square roots of $m^2_{\nu_\alpha} = \sum_i |U_{\alpha i}|^2 \, m^2_{\nu_i}$ with $\alpha \in e, \mu, \tau$ and $i \in 1, 2, 3$. The values of the masses are taken from Ref. [44].

| Generation | Fermion | | Mass in MeV | $Q$ | $T_3$ | colored |
|---|---|---|---|---|---|---|
| 1st | $e^-$ | electron | 0.511 | $-1$ | $-\frac{1}{2}$ | no |
| | $\nu_e$ | electron-neutrino | $< 1.1 \times 10^{-6}$ | 0 | $\frac{1}{2}$ | no |
| | $u$ | up-quark | $2.16^{+0.49}_{-0.26}$ | $\frac{2}{3}$ | $\frac{1}{2}$ | yes |
| | $d$ | down-quark | $4.67^{+0.48}_{-0.17}$ | $-\frac{1}{3}$ | $-\frac{1}{2}$ | yes |
| 2nd | $\mu^-$ | muon | 105.7 | $-1$ | $-\frac{1}{2}$ | no |
| | $\nu_\mu$ | muon-neutrino | $< 0.19$ | 0 | $\frac{1}{2}$ | no |
| | $c$ | charm-quark | $(1270 \pm 20)$ | $\frac{2}{3}$ | $\frac{1}{2}$ | yes |
| | $s$ | strange-quark | $93.4^{+8.6}_{-3.4}$ | $-\frac{1}{3}$ | $-\frac{1}{2}$ | yes |
| 3rd | $\tau^-$ | tau | $(1776.86 \pm 0.12)$ | $-1$ | $-\frac{1}{2}$ | no |
| | $\nu_\tau$ | tau-neutrino | $< 18.2$ | 0 | $\frac{1}{2}$ | no |
| | $t$ | top-quark | $(172.69 \pm 0.30) \times 10^3$ | $\frac{2}{3}$ | $\frac{1}{2}$ | yes |
| | $b$ | bottom-quark | $4180^{+30}_{-20}$ | $-\frac{1}{3}$ | $-\frac{1}{2}$ | yes |

where $\psi$ is the wave function of a matter field and $\chi(\vec{x}, t)$ is a space-time dependent scalar field. For the Lagrangian to satisfy this invariance, the existence of additional fields and particles is required. This leads to interaction terms mediated by the associated gauge bosons. According to Noether's theorem [50], there exists a corresponding conserved charge to each continuous symmetry. The internal symmetry obeyed by the SM is that of the non-abelian group

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y, \tag{2.2}$$

where the $SU(3)_C$ and $SU(2)_L \otimes U(1)_Y$ parts are related to the strong and electroweak interactions, respectively. The theory describing the former is called quantum chromodynamics (QCD). It is unknown and still an open research question if these are all the internal symmetries of nature or why those three are realized. However, their discovery, as well as the predictive power of a theory built just from requiring this symmetry to hold, are astonishing.

As a fully relativistic theory, the SM also satisfies the global Poincaré symmetry. The conserved quantities associated with this symmetry through Noether's theorem [50] are [51]:

- energy
- (linear) momentum
- angular momentum
- center-of-mass velocity.

### 2.1.2 Standard Model Lagrangian

Once the symmetries of the SM are set, it is possible to construct its Lagrangian. The SM Lagrangian consists of three components describing the electroweak theory, the electroweak

**Table 2.2:** Properties of all Standard Model bosons sorted by the interaction they are mediating. Shown are the particles' masses, electric charge $Q$, and the third component of the weak isospin $T_3$. The number listed for the gluon mass is a "Theoretical value. A mass as large as a few MeV may not be precluded." The values of the masses are taken from Ref. [44].

| Interaction | Boson | | Mass in GeV | Spin | $Q$ | $T_3$ | colored |
|---|---|---|---|---|---|---|---|
| electromagnetic | $\gamma$ | photon | $< 1 \times 10^{-27}$ | 1 | 0 | 0 | no |
| weak | $W^{\pm}$ | W bosons | $(80.377 \pm 0.012)$ | 1 | $\pm1$ | $\pm1$ | no |
| | $Z$ | Z boson | $(91.1876 \pm 0.0021)$ | 1 | 0 | 0 | no |
| strong | $g$ | gluons | 0 | 1 | 0 | 0 | yes |
| - | $H$ | Higgs | $(125.25 \pm 0.17)$ | 0 | 0 | $-\frac{1}{2}$ | no |

symmetry breaking required for fermion and boson masses, and quantum chromodynamics. It can be summarized as:

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{EWSB} + \mathcal{L}_{QCD}. \tag{2.3}$$

The following sections explore each of these parts in a bit more depth.

### Electroweak Theory

The first evidence for the now combined electroweak theory was seen at the laboratory of the European Organization for Nuclear Research (CERN). Neutral current interactions were first observed by the Gargamelle experiment in 1973, which was the experimental indication of the $Z$ boson [52]. The full discovery of the $W$ and the $Z$ boson had to wait until 1983, when it was achieved using the SPS Proton-Antiproton Collider [53–56].

It has been known since the Wu experiment [57] in 1957 that the weak interaction violates parity. As a consequence, particles with different chiralities are treated differently. To handle this correctly, the electroweak interaction acts differently on left-handed spinors $\psi_L$, which form $SU(2)_L$ doublets, compared to their right-handed counterparts $\psi_R$, which are $SU(2)_L$ singlets. These components can be built from the fermionic wave function $\psi$ via

$$\psi_{L/R} = P_{L/R}\psi = \frac{1 \mp \gamma^5}{2}\psi, \tag{2.4}$$

with $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. The field content of the first generation of leptons can thus be written as:

$$L_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L = \begin{pmatrix} P_L\psi_{\nu_e} \\ P_L\psi_e \end{pmatrix}, \tag{2.5}$$

$$l_R = e_R = P_R\psi_e. \tag{2.6}$$

As right-handed neutrinos would have no electric, weak, or strong charge, they do not interact with any other particle. Due to this, they have not been observed and are not part of the SM. The analogous relations for the first generation of quarks read:

$$Q_{L\alpha} = \begin{pmatrix} u_\alpha \\ d_\alpha \end{pmatrix}_L = \begin{pmatrix} P_L\psi_{u_\alpha} \\ P_L\psi_{d_\alpha} \end{pmatrix} \tag{2.7}$$

$$u_{R\alpha} = P_R\psi_{u_\alpha}, \tag{2.8}$$

$$d_{R\alpha} = P_R\psi_{d_\alpha}. \tag{2.9}$$

Using the Dirac adjoint spinor $\bar{\psi} = \psi^\dagger\gamma^0$, the fermionic electroweak sector thus takes the form:

$$\begin{aligned}\mathcal{L}_{ew} = &\sum_{j=1}^{3} i\overline{L}_L^j \not{D} L_L^j + i\overline{l}_R^j \not{D} l_R^j \\ &+ i\overline{Q}_L^j \not{D} Q_L^j + i\overline{u}_R^j \not{D} u_R^j + i\overline{d}_R^j \not{D} d_R^j,\end{aligned} \tag{2.10}$$

where

$$\begin{aligned}\not{D} &= \gamma^\mu D_\mu, \\ D^\mu &= \partial^\mu - ig_1\frac{Y}{2}B^\mu - ig_2\frac{\tau_i}{2}W_i^\mu\end{aligned} \tag{2.11}$$

is the covariant derivative, which is needed for the Lagrangian to remain invariant under $SU(2)_L \otimes U(1)_Y$. Here, $W_\mu^a$ (a = 1,2,3) and $B_\mu$ are the gauge fields of the $SU(2)_L$ and $U(1)_Y$ subgroups, respectively, $\tau_i$ are the Pauli matrices and $g_1$ and $g_2$ represent the coupling strengths. The values of the coupling strengths have to be determined by experiments but are identical for the different particles. The sum over the indices $j$ accounts for the three known generations of leptons and quarks.

Focusing just on the first generation again and specializing to the $U(1)$ terms for the leptons results in:

$$\mathcal{L}_{lepton}(U(1)) = \frac{g_1}{2}\left[Y_L(\bar{\nu}_L\gamma^\mu\nu_L + \bar{e}_L\gamma^\mu e_L) + Y_R\bar{e}_R\gamma^\mu e_R\right]B_\mu \tag{2.12}$$

As left- and right-handed projections can have separate hypercharges $Y$, $Y_R$ and $Y_L$ have been introduced.

By defining the physical, charged gauge boson fields

$$W_\mu^\pm = (-W_\mu^1 \pm iW_\mu^2)/\sqrt{2}, \tag{2.13}$$

and considering that terms for right-handed fermions vanish due to them being singlets, the leptonic terms stemming from the $SU(2)$ are as follows:

$$\mathcal{L}_{leptons}(SU(2)) = \frac{g_2}{2}\left[\bar{\nu}_L\gamma^\mu\nu_L W_\mu^3 - \sqrt{2}\bar{\nu}_L\gamma^\mu e_L W_\mu^+ - \sqrt{2}\bar{e}_L\gamma^\mu\nu_L W_\mu^- - \bar{e}_L\gamma^\mu e_L W_\mu^3\right] \tag{2.14}$$

Knowing from experiment [58] that neutrinos do not interact electromagnetically, the pure neutrino interactions are investigated. The terms concerned with that in the Lagrangian are proportional to

$$-\frac{g_1}{2}Y_L B_\mu - \frac{g_2}{2}W_\mu^3. \tag{2.15}$$

The physical fields $A_\mu$ and $Z_\mu$ are then defined as orthogonal combinations of $B_\mu$ and $W_\mu^3$, such that $A_\mu$ does not interact with neutrinos. Correctly normalized, they are written as:

$$A_\mu = \frac{g_2 B_\mu - g_1 Y_L W_\mu^3}{\sqrt{g_2^2 + g_1^2 Y_L^2}}, \tag{2.16}$$

$$Z_\mu = \frac{g_1 Y_L B_\mu + g_2 W_\mu^3}{\sqrt{g_2^2 + g_1^2 Y_L^2}}, \tag{2.17}$$

where $A_\mu$ is the photon field.

Adding these relations into the interaction terms containing only electrons and requiring $A_\mu$ to couple equally to the left- and right-handed electrons, as that is what has been experimentally observed [59, 60], one finds that

$$Y_R = 2Y_L. \tag{2.18}$$

As $Y_L$ only occurs in combination with $g_1$, $Y_L$ can be chosen to be $Y_L = -1$. Using that and by comparison to the expected electromagnetic interaction with a coupling between photons and electrons of $-e$, it can be seen that

$$e = \frac{g_1 g_2}{\sqrt{g_2^2 + g_1^2}}. \tag{2.19}$$

The off-diagonal elements combining $\nu$ and $e$ give rise to the observed charged current interactions mediated by $W^\pm$ bosons [53]. It also makes sense to define the electroweak mixing angle $\theta_W$ here:

$$\sin\theta_W = \frac{g_1}{\sqrt{g_2^2 + g_1^2}}, \tag{2.20}$$

$$\cos\theta_W = \frac{g_2}{\sqrt{g_2^2 + g_1^2}}. \tag{2.21}$$

**Particle Masses**

So far the Lagrangian has contained no mass terms $m\bar{\psi}\psi = m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R)$ for fermions or $\frac{1}{2}m_B^2 B^\mu B_\mu$ for bosons. This is because the naive addition of such terms violates the SM's gauge symmetries. The solution to this discrepancy to observation [15, 16, 44] is provided by the Higgs mechanism [35–37].

The main idea lies in the introduction of a complex scalar $SU(2)$ doublet with a non-zero vacuum expectation value:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \frac{\phi_1 + i\phi_2}{\sqrt{2}} \\ \frac{\phi_3 + i\phi_4}{\sqrt{2}} \end{pmatrix}, \tag{2.22}$$

with a Lagrangian

$$\mathcal{L}_\phi = (D_\mu\phi)^\dagger(D^\mu\phi) - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2. \tag{2.23}$$

For $\mu^2 < 0$ and $\lambda > 0$, the minima of the potential term satisfy

$$\phi^\dagger\phi = \frac{-\mu^2}{2\lambda} = \frac{v}{2}. \tag{2.24}$$

An arbitrary choice of a minimum is

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{2.25}$$

i.e. $\phi_3 = v$, $\phi_1 = \phi_2 = \phi_4 = 0$ in Equation 2.22. Expanding around this minimum yields

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{2.26}$$

This choice breaks the $SU(2)_L \otimes U(1)_Y$ symmetry to the usual QED one of $U(1)_Q$ and leaves only the Higgs field $h$ from the original four degrees of freedom. The Goldstone theorem states that for each of the three broken generators, there is a corresponding Goldstone boson. These are gauged away and "eaten" by the $W$ and $Z$ bosons to form their third spin degree of freedom. Expanding the dynamic terms with the covariant derivative in Equation 2.23 results in mass terms for the $W^\pm$ and $Z$ bosons, giving them a longitudinal polarization. These three extra degrees of freedom are those of the Goldstone bosons that were gauged away. The photon is left massless. The resulting masses are

$$M_W = \frac{g_2 v}{2} \tag{2.27}$$

$$M_Z = \frac{1}{2}v\sqrt{g_1^2 + g_2^2} = \frac{M_W}{\cos\theta_W}. \tag{2.28}$$

This expansion also introduces three and four-point interactions between the Higgs boson and the now massive weak gauge bosons. Substituting Equation 2.26 into the potential in

Equation 2.23 also results in a mass term for the Higgs boson $m_h = \sqrt{2\lambda}v \approx 125\,\mathrm{GeV}$ as well as three and four-point self-interactions. The vacuum expectation value is related to the Fermi constant via

$$v = 2\frac{M_W}{g_2} = \frac{1}{\sqrt{\sqrt{2}G_F}} \approx 246\,\mathrm{GeV}, \tag{2.29}$$

which has to be determined experimentally, as none of the potential parameters in Equation 2.23 are predicted by the SM.

Using the relation $Q = T_3 + \frac{Y}{2}$ between the electric charge $Q$ with the weak isospin $T_3$ and the hypercharge $Y$, and requiring the Higgs field $h$ to be electrically neutral, means setting the hypercharge $Y = 1$ to compensate the weak isospin $T_3 = -\frac{1}{2}$. A more pedagogical introduction in multiple steps can be found in Ref. [45], while an in-depth and general derivation is presented in Ref. [48]. That reference also contains more information about the Goldstone bosons in different gauges and the delicate cancellations that occur in each of them for the calculation of observables.

The last major missing thing now is the masses of fermions. While the Higgs mechanism does not automatically introduce mass terms for the fermions, it opens up a way to add them while keeping gauge invariance. For leptons and quarks, they are added as follows:

$$\mathcal{L}_Y = \sum_{j=1}^{3}\left( y_\ell^j \overline{L}_L^j \phi \ell_R^j + \sum_{k=1}^{3} Y_d^{jk} \overline{Q}_L^j \phi d_R^k + Y_u^{jk} \overline{Q}_L^j i\tau_2 \phi^* u_R^k + h.c. \right), \tag{2.30}$$

where the $y_l^j$ are the Yukawa couplings of the leptons and $Y^{jk}$ are those of the quarks. The additional sum and index for the quark terms are due to the mixing between the quarks' flavor states. This is described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix, which can be used to transform between mass $(d, s, b)$ and flavor eigenstates $(d', s', b')$:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.31}$$

The expansion of $\phi$ around its minimum yields masses $m_f = \frac{y_f v}{\sqrt{2}}$, which are free parameters of the theory, as well as couplings between the Higgs field and the fermions.

### Boson Interactions and Kinematics

Matching observational facts, kinetic and interaction terms of the bosons are introduced to the Lagrangian:

$$\mathcal{L}_{bosons} = -\frac{1}{4}B^{\mu\nu}B_{\mu\nu} - \frac{1}{4}W_i^{\mu\nu}W_{\mu\nu}^i - \frac{1}{4}G_a^{\mu\nu}G_{\mu\nu}^a, \tag{2.32}$$

with

$$B^{\mu\nu} = \partial^\mu B^\nu - \partial^\nu B^\mu, \tag{2.33}$$
$$W_i^{\mu\nu} = \partial^\mu W_i^\nu - \partial^\nu W_i^\mu + g_2 \epsilon_{ijk} W_j^\mu W_k^\nu, \tag{2.34}$$
$$G_a^{\mu\nu} = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu + g_3 f_{abc} G_b^\mu G_c^\nu. \tag{2.35}$$

Here, $\epsilon_{ijk}$ and $f_{abc}$ are the $SU(2)$ and $SU(3)$ structure constants.

The nature of the $W^{\mu\nu}$ and $G^{\mu\nu}$ field tensors give rise to three and four-point interactions of the $W^\pm$ and $Z$ bosons as well as the gluons. Those for the weak bosons are shown in Figure 2.1.

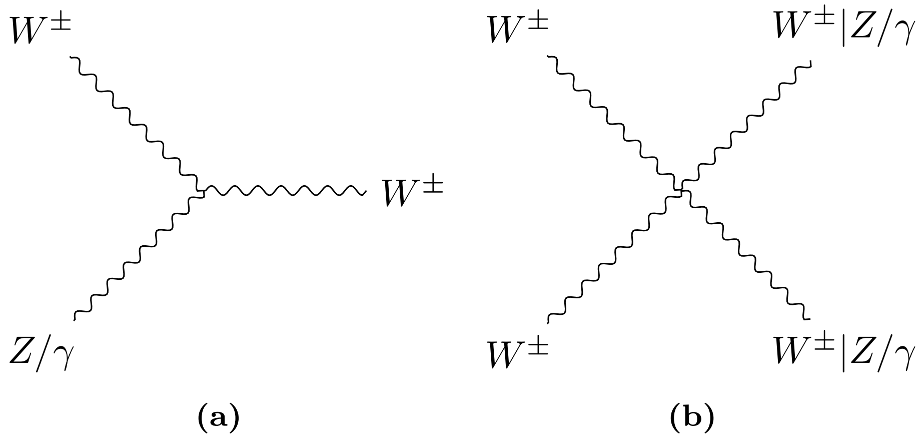



**Figure 2.1:** Feynman diagrams for three (a) and four (b) point interactions of electroweak gauge bosons. The "|" separator indicates that either the left or right option has to be chosen at all points in the diagram. For the "/", the sides can be picked independently from each other. The charges of the $W^\pm$ bosons have to be chosen such that the overall charge is conserved.

The full SM Lagrangian is finally

$$\begin{aligned}
\mathcal{L}_{SM} =& \sum_f i\bar{f}\not{D}f \\
&+ (D_\mu\phi)^\dagger(D^\mu\phi) - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \\
&- \frac{1}{4}B^{\mu\nu}B_{\mu\nu} - \frac{1}{4}W_i^{\mu\nu}W^i_{\mu\nu} - \frac{1}{4}G_a^{\mu\nu}G^a_{\mu\nu} \\
&+ \sum_{j=1}^{3}\left(y_\ell^j\overline{L}_L^j\phi\ell_R^j + \sum_{k=1}^{3}Y_d^{jk}\overline{Q}_L^j\phi d_R^k + Y_u^{jk}\overline{Q}_L^j i\tau_2\phi^* u_R^k + h.c.\right).
\end{aligned} \tag{2.36}$$

**Quantum Chromodynamics**

The full covariant derivative, including parts needed to remain invariant under $SU(3)_C$, is:

$$D^\mu = \partial^\mu - ig_1\frac{Y}{2}B^\mu - ig_2\frac{\tau_i}{2}W_i^\mu - ig_3\frac{\lambda_a}{2}G_a^\mu \tag{2.37}$$

where $\lambda^a$ are the Gell-Mann matrices and $G_a^\mu$ are the gluon fields and $g_3$ the strong coupling strength. As the leptons do not possess any color charge, the last, newly introduced, term in Equation 2.36, representing $3 \times 3$ matrices in color space, gives zero contribution to their dynamics. For the color-charged quarks, that term leads to a contribution to the SM Lagrangian of the form

$$\mathcal{L}_{QCD} \sim \frac{g_3}{2}\bar{q}_\alpha\gamma^\mu\lambda^a_{\alpha\beta}G^a_\mu q_\beta, \tag{2.38}$$

where $\alpha/\beta = 1, 2, 3$ are color indices. The gluon fields are all electrically neutral and interact with quarks similarly to the electromagnetic field. However, unlike photons, the interaction of a quark with a gluon can and does change the quark's color. This comes from the fact that not all of the generators are diagonal.

## 2.2 Polarization

This section gives an introduction to the polarization of vector bosons. First, Section 2.2.1 provides a quick overview of the foundational aspects needed to conceptually define polarizations for moving massive bosons. The equations for describing massive spin-1 particles in the SM are shown in Section 2.2.2. Section 2.2.3 summarizes the statistical considerations needed for polarization measurements in scattering experiments at particle colliders. Lastly, the caveats of polarization of intermediate particles are addressed in Section 2.2.4.

### 2.2.1 Spin and Polarization

In relativistic quantum mechanics, an operator analogous to the classical spin operators can be found in the Pauli-Lubanski operators. However, these only obey the desired relations for particles at rest [61–63]. Then, one can classify particles by the projection of their spin onto an arbitrary direction.

To generate particles with non-zero momentum, a Lorentz transformation can be applied. However, just stating the particle's transformed momentum is not enough for a clear classification of its spin. There are infinitely many frames that see the particle moving with the same momentum. As these frames can be related by rotations around the momentum vector, they can each report a different spin. Thus, the Lorentz transformation used has to be specified to achieve a proper classification.

The frame used in this thesis is the "modified helicity frame". Here, the spin quantization axis in the boson rest frame is defined by its momentum in the $W^{\pm}Z$ rest frame. As such, a particle that is observed in the state $|p; \lambda_{\text{mod}}\rangle$ in the laboratory frame will be found with a spin component $s_Z = \lambda_{\text{mod}}$ in its rest frame. This is achieved by a Lorentz boost from the laboratory frame into the $W^{\pm}Z$ rest frame and a subsequent rotation of the z-axis towards the direction of the particle's momentum, followed by a boost in that direction. As noted above, this specific transformation to the particle's rest frame needs to be employed to achieve the equality between $\lambda_{\text{mod}}$ and $s_Z$.

A visualization of the "modified helicity frame", alongside one for the "helicity frame", where the boson momentum, defining the spin quantization axis in the bosons rest frame, is defined in the laboratory system, is shown in Figure 2.2. It also demonstrates the definitions of the scattering angle $\cos\theta_V$ and the decay angles $\cos\theta^{\star}_{\ell V}$ or $\cos\theta_{\ell_V}$. The former is defined as the angle between the $W^{\pm}$ boson and the z-axis in the $WZ$ rest frame, and the latter are the angles between the direction of the boson momentum in the laboratory/$WZ$ rest frame and the lepton momentum in the boson rest frame. For the $W$ boson, the lepton is used for the decay angle, while the negatively charged lepton is used for the $Z$ boson.

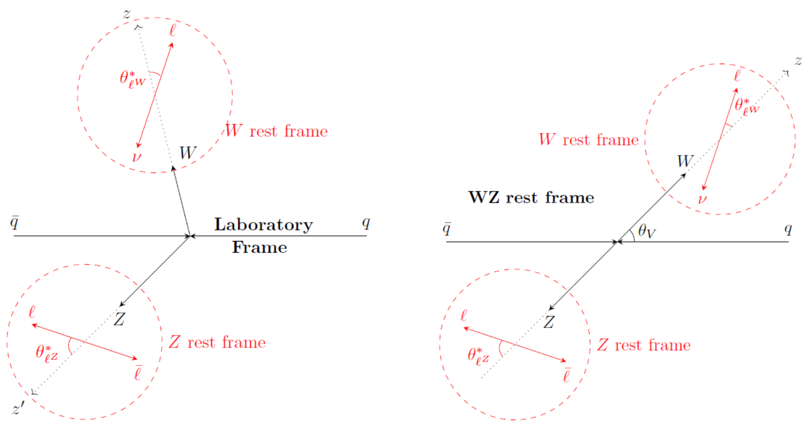



**Figure 2.2:** Visual representation of helicity and modified helicity frame. From Ref.[64]

### 2.2.2 Polarization of Massive Spin-1 Fields

To derive the basics of polarization of massive spin-1 particles, it is useful to start with the Proca equation

$$(\Box + m^2)B^{\mu} - \partial^{\mu}(\partial_{\nu}B^{\nu}) = 0. \tag{2.39}$$

This is the equation of motion for such particles as derived from the Lagrangian density, contained in Equation 2.36, via the EulerLagrange equations. Taking the divergence yields

$$\partial_\mu B^\mu = 0. \tag{2.40}$$

Employing a plane wave ansatz $B^\mu = \epsilon^\mu e^{-iq\cdot x}$ results in the condition

$$q_\mu \epsilon^\mu = 0. \tag{2.41}$$

For a spin-1 particle with mass, three spin degrees of freedom are characterized by three polarization vectors. For a particle traveling in $z$-direction these are

$$\epsilon^\mu_- = \frac{1}{\sqrt{2}}(0,-1,i,0),\ \ \epsilon^\mu_+ = \frac{1}{\sqrt{2}}(0,-1,-i,0),\ \ \epsilon^\mu_0 = (0,0,0,1). \tag{2.42}$$

With a Lorentz boost for a velocity corresponding to the momentum $p$ along the z-axis followed by a rotation, one can get the relations [65]

$$\epsilon^\mu_\pm(p) = \frac{e^{\pm i\phi}}{\sqrt{2}}\,(0,-\cos\theta\cos\phi \pm i\sin\phi, -\cos\theta\sin\phi \mp i\cos\phi, \sin\theta) \tag{2.43}$$

$$\epsilon^\mu_0(p) = \frac{p^0}{m}\left(\frac{|\vec{p}|}{p^0}, \cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta\right) \tag{2.44}$$

in the helicity basis.

Helicities + and − are called transversely polarized, right- and left-handed, respectively, while helicity 0 is called longitudinal polarization. In this work, the shorthand notation of 'T' for the transverse and '0' for longitudinal polarization will be used. The longitudinal polarization is also sometimes abbreviated with the letter 'L'. This should not lead to confusion with left-handed states as these are not studied directly in this thesis.

The polarization vectors obey the completeness relation

$$\sum_{\lambda=1}^{3} \epsilon^\mu_\lambda(p)\epsilon^{*\nu}_\lambda(p) = -g^{\mu\nu} + \frac{p^\mu p^\nu}{m^2}. \tag{2.45}$$

In the unitary gauge, the Feynman rules for internal and external lines of massive spin-1 particles can be found as



$$\frac{i(-g_{\mu\nu} + \frac{p_\mu p_\nu)}{m^2}}{p^2 - m^2 + i\epsilon}, \tag{2.46}$$

$$\epsilon_\lambda(p), \tag{2.47}$$

$$\epsilon^*_\lambda(p). \tag{2.48}$$

### 2.2.3 Spin Density Matrix

In collisions at the LHC, the polarizations of initial state particles are usually not known. Additionally, the polarization of any particular intermediate boson can not be determined directly. Due to this, a statistical description is required.

In cases like these, the use of a density matrix, or density operator, is advised. More information about spin density matrices can be found in Ref. [61, 66]. The spin density matrix, as any other density matrix, is positive semi-definite, Hermitian, and has a trace of one.

Provided that $\langle S, m|$ form a basis of the spin space, the spin density matrix can be written as

$$\hat{\rho} = \sum_{m=-S}^{m=+S} \omega_m |S,m\rangle\langle S,m|, \tag{2.49}$$

with $\omega_m$ between 0 and 1 being the probability of finding the system in the state $\langle S, m|$. The trace of one implies

$$\mathrm{Tr}(\rho) = \sum_{m=-S}^{m=+S} \omega_m = 1. \tag{2.50}$$

As discussed in Section 2.2.1, the classical spin formalism can be retained in relativistic quantum mechanics by moving to the helicity basis with the helicity operator

$$h = \frac{\mathbf{S} \cdot \vec{p}}{|\vec{p}|}. \tag{2.51}$$

In the helicity basis, the spin density matrix can be written according to Equation 2.49 with a basis of the polarization vectors from Equation 2.42. Then, in this basis, some components can be measured in an experiment. Here, the diagonal elements of the spin density matrix are called the polarization fractions, and the off-diagonal elements represent interference terms. Via Equation 2.50 it follows that

$$f_0 + f_L + f_R = f_0 + f_T = 1, \tag{2.52}$$

which can be trivially expanded to cases with two vector bosons.

### 2.2.4 Intermediate Polarized Bosons

As shown at the beginning of this chapter it is straightforward to define and access the polarization of external vector bosons as their polarization vectors show up directly in the matrix element. However, both the $W^\pm$ and the $Z$ boson are too short-lived to be measured directly as external particles with current accelerator and detector experiments. Thus, it is necessary to consider the case of intermediate polarized bosons such as those shown in the Feynman diagrams in Figure 2.3.

As the polarization vectors do not directly appear in the propagator as shown in Equation 2.47, they have to be reintroduced via the completeness relation

$$\sum_{\lambda=1}^{4} \epsilon^\mu_\lambda(p)\epsilon^{*\nu}_\lambda(p) = -g^{\mu\nu} + \frac{p^\mu p^\nu}{m^2}. \tag{2.53}$$

The relation here differs slightly from the one in Equation 2.45 by a fourth unphysical polarization, which vanishes for on-shell bosons or decays to massless leptons [67]. Applying this substitution and adding a term accounting for the finite width $\Gamma$ of the bosons leads to the following amplitude:

$$\mathcal{M} = \sum_\lambda \mathcal{M}^{\mathrm{prod}}_\mu \frac{i\epsilon^\mu_\lambda(p)\epsilon^{*\nu}_\lambda(p)}{p^2 - m^2 + i\Gamma m} \mathcal{M}^{\mathcal{D}}_\nu =: \sum_\lambda \mathcal{M}^{\mathcal{F}}_\lambda. \tag{2.54}$$

Here, $\mathcal{M}^{\mathcal{F}}_\lambda$ is the fully polarized amplitude for a single vector boson with polarization $\lambda$ and $\mathcal{M}^{\mathrm{prod}}_\mu$ and $\mathcal{M}^{\mathrm{D}}_\mu$ are the production and decay amplitudes. To calculate polarized cross-sections one subsequently uses just $\mathcal{M}^{\mathcal{F}}_\lambda$ instead of the full amplitude $\mathcal{M}$. One point to note here is that the sum of polarized cross-sections calculated this way will not (generally) be equal to the unpolarized cross-section calculated directly from $\mathcal{M}$. This is due to missing interference terms that arise when squaring the amplitude:

$$\underbrace{|\mathcal{M}|^2}_{\text{coherent sum}} = \underbrace{\sum_\lambda |\mathcal{M}^{\mathcal{F}}_\lambda|}_{\text{incoherent sum}} + \underbrace{\sum_{\lambda\neq\lambda'} \mathcal{M}^{\mathcal{F}}_\lambda \mathcal{M}^{*\mathcal{F}}_{\lambda'}}_{\text{interference}} \tag{2.55}$$

The interference terms only cancel when the amplitude can be integrated over the full range of $\phi$. This also means that no lepton selection criteria are applied. Under the same conditions, the angular distribution of $W^\pm/Z$ decays follows a simple analytical distribution.

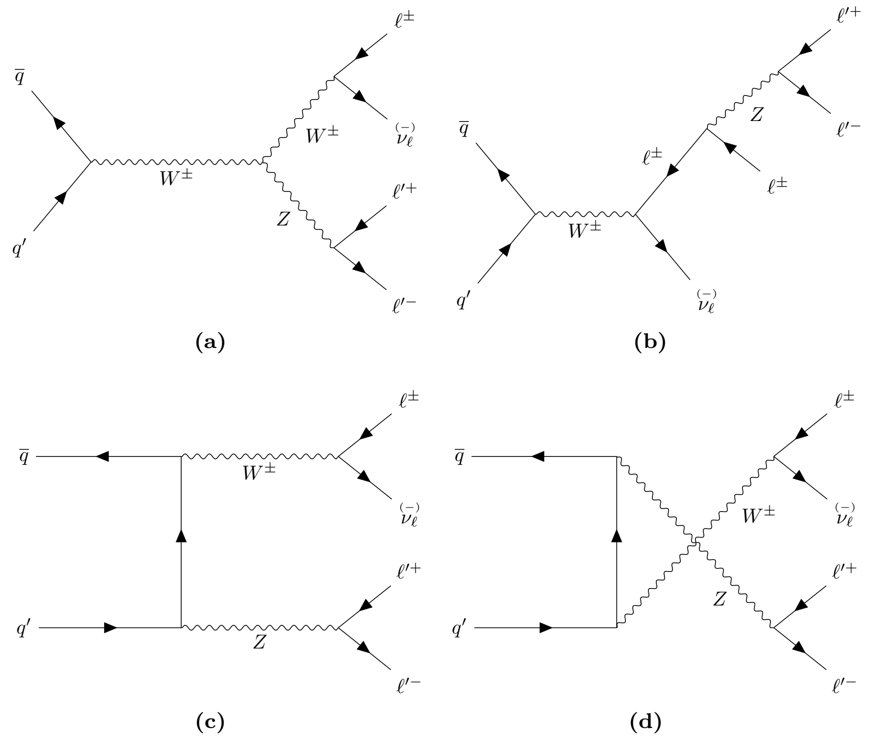



**Figure 2.3:** Leading order Feynman diagrams for the process $q\overline{q} \to \ell^{\pm}\overset{(-)}{\nu}_{\ell}\ell'^{+}\ell'^{-}$. Doubly resonant diagrams can be seen in (a), (c) and (d). Meanwhile, (b) shows an example of a non-resonant diagram where the matrix element can not be split cleanly into production and decay for each boson. Two more such diagrams exist with a $W^{\pm}$ or $Z$ boson being radiated from the neutrino instead. The fermion arrows related to the $W^{\pm}$ boson are set for the case of a $W^{+}$.

This distribution can be projected onto Legendre polynomials to extract the polarization fractions [67, 68].

In practice, this does not work, as realistic setups always include such lepton selection cuts. This necessitates the full simulation of this process via Monte Carlo generators. For more complex processes, like $W^{\pm}Z$ production, another difficulty arises in the form of non-resonant diagrams like the one in Figure 2.3b. The above procedure can not be applied to diagrams like these, as they can not be factorized into the production and decay of the bosons. However, simply ignoring such single- or non-resonant diagrams would break gauge invariance. Different Monte Carlo generators use different techniques to rectify this situation while introducing as minimal approximations or biases as possible. Some common approaches are:

- **Narrow-Width Approximation (NWA)**: This approximation replaces the denominator of the propagator with a delta function:

$$\frac{1}{p^2 - m^2 + i\Gamma m} \to \frac{\pi}{m\Gamma}\delta(p^2 - m^2) \tag{2.56}$$

  It is used by SHERPA [69, 70]

- **Double-Pole Approximation (DPA)**: In this approach, only the numerator of the propagator is adjusted while leaving the denominator unchanged. There, the momenta of the vector bosons are projected to on-shell values. The DPA is used in the PHANTOM [71] generator as well the RECOLA program [72, 73]. More information on this technique, as well as caveats on how to choose the projection, can be found in Ref. [67]. This technique is also often referred to as On-Shell Projection (OSP) in the literature.

- **Pure Diagram Selection**: This method vetoes all diagrams that are not fully resonant, introducing violations of gauge invariance. Optionally, a cut can be applied to select how on-shell the decay products have to be to limit the size of the gauge violation.

  This approach is used by MadGraph5_aMC@NLO [74, 75] with an on-shell condition $|m^\star - m| \leq \mathbf{bwcutoff} \cdot \Gamma$ and a default value of $\mathbf{bwcutoff} = 15$ [76, 77]. Here, $m^\star$ is the invariant mass of the boson produced in the event and $m$ the literature on-shell mass.

The accuracy of all three approaches is on the order $\mathcal{O}(\Gamma/M)$ [68, 74, 78].

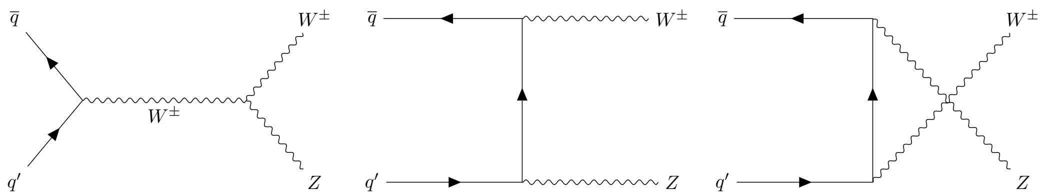
## 2.3 $W^\pm Z$ Production

Measuring polarization in the production of $W^\pm Z$ bosons is of interest for multiple reasons. The first is the direct link of the longitudinal polarization to the Higgs mechanism, as shown in Section 2.1.2. Another one is the prediction of amplitude zeros in the production of two transversally polarized bosons. While an experimental investigation of this effect is of interest in and of itself, this reduction in the transverse polarizations also causes a higher fraction of doubly longitudinal events, making it easier to study the first aspect.

The starting point to investigate this so-called radiation amplitude zero (RAZ) effect is the helicity amplitudes for the process

$$f_1(p_1)\bar{f}_2(p_2) \to W^\pm(p_W)Z(p_Z). \tag{2.57}$$

Based on the three leading order diagrams as shown in Figure 2.4 the helicity amplitudes can be derived as [27]




(a) (b) (c)

**Figure 2.4:** Leading order (a) s-, (b) t- and (c) u-channel Feynman diagrams for the process $f_1(p_1)\bar{f}_2(p_2) \to W^\pm(p_W)Z(p_Z)$.

$$\mathcal{M}\left(\lambda_{\mathrm{w}}, \lambda_{\mathrm{z}}\right) = F \epsilon_\mu^* \left(p_{\mathrm{w}}, \lambda_{\mathrm{w}}\right) \epsilon_\nu^* \left(p_{\mathrm{z}}, \lambda_{\mathrm{z}}\right) \bar{V}\left(p_2\right) T^{\mu\nu} \frac{1}{2}\left(1-\gamma_5\right) U\left(p_1\right), \tag{2.58}$$

with

$$\begin{aligned} T^{\mu\nu} &= g_-^{f_1} \gamma^\mu \frac{\left(\not{p}_1 - \not{p}_{\mathrm{z}}\right)}{u} \gamma^\nu \\ &+ g_-^{f_2} \gamma^\nu \frac{\left(\not{p}_1 - \not{p}_{\mathrm{w}}\right)}{t} \gamma^\mu \\ &+ \frac{Q_W \cot\theta_{\mathrm{w}}}{s - M_W^2}\left[\left(\not{p}_{\mathrm{z}} - \not{p}_{\mathrm{w}}\right) g^{\mu\nu} + 2p_{\mathrm{w}}^\nu \gamma^\mu - 2p_{\mathrm{z}}^\mu \gamma^\nu\right]. \end{aligned} \tag{2.59}$$

where $\lambda$ are the polarizations of the bosons with $\lambda = \pm 1, 0$, $\epsilon$ the bosons' polarization vectors, $V/U$ the (anti-)fermion spinors, $g_-^{f_i} = \frac{1}{\sin\theta_W \cos\theta_W}(T_3^f - Q_f \sin^2\theta_W)$ the couplings of the $Z$-boson to left-handed fermions and $s$, $t$ and $u$ the Mandelstam variables. The factor $F$ is

$$F = C \frac{e^2}{\sqrt{2}\sin\theta_W}, \tag{2.60}$$

where $C = \delta_{i_1 i_2} V_{f_1 f_2}$, with $i_1, i_2$ denoting the incoming quarks color indices and $V_{f_1 f_2}$ the quark mixing matrix element. Moving to the parton center-of-mass frame, which, at leading order, is the same as the $W^\pm Z$ center-of-mass frame, and writing these amplitudes out for each of the separate polarizations yield [27]

$$\begin{aligned}
\mathcal{M}(\pm,\mp) &= F\sin\theta\,(\lambda_\mathrm{w} - \cos\theta)\,X,\\
\mathcal{M}(\pm,\pm) &= F\sin\theta\left[\left(\lambda_\mathrm{w}\,(r_\mathrm{z} - r_\mathrm{w}) - \beta + \cos\theta + \beta\frac{\alpha - \beta\cos\theta}{1 - r_\mathrm{w}}\right)X + 2\beta Y\right],\\
\mathcal{M}(0,0) &= F\frac{\sin\theta}{\sqrt{2r_\mathrm{w}2r_\mathrm{z}}}\left[\left(-\beta\rho + \beta_\mathrm{w}\beta_\mathrm{z}\cos\theta + \beta\rho\frac{\alpha - \beta\cos\theta}{1 - r_\mathrm{w}}\right)X + 2\beta\rho Y\right],\\
\mathcal{M}(\pm,0) &= F\frac{(1 - \lambda_\mathrm{w}\cos\theta)}{\sqrt{2r_\mathrm{z}}}\left[\left(2\lambda_\mathrm{w}r_\mathrm{z} - \beta + \beta_\mathrm{z}\cos\theta + \beta\frac{\alpha - \beta\cos\theta}{1 - r_\mathrm{w}}\right)X + 2\beta Y\right],\\
\mathcal{M}(0,\pm) &= F\frac{(1 + \lambda_\mathrm{z}\cos\theta)}{\sqrt{2r_\mathrm{w}}}\left[\left(-2\lambda_\mathrm{z}r_\mathrm{w} - \beta + \beta_\mathrm{w}\cos\theta + \beta\frac{\alpha - \beta\cos\theta}{1 - r_\mathrm{w}}\right)X + 2\beta Y\right],
\end{aligned}\tag{2.61}$$

with $\theta$ being the scattering angle of the $W^\pm$-boson with respect to the $q_1$ direction as well as the following shorthands:

$$\begin{aligned}
X &= \frac{s}{2}\left(\frac{g_-^{f_1}}{u} + \frac{g_-^{f_2}}{t}\right),\\
Y &= g_-^{f_1}\frac{M_Z^2 s}{2u(s - M_W^2)},\\
\alpha &= 1 - r_W r_Z,\\
\beta &= (\alpha^2 - 4r_W r_Z)^{1/2},\\
r_V &= \frac{M_V^2}{s},\\
\beta_W &= 1 + r_W - r_Z,\\
\beta_Z &= 1 - r_W + r_Z,\\
\rho &= 1 + r_W + r_Z.
\end{aligned}$$

With the amplitude $\mathcal{M}(\pm,\mp)$ being independent of $Y$ it vanishes when $X = 0$, so when

$$\frac{g_-^{f_1}}{u} + \frac{g_-^{f_2}}{t} = 0 \quad \text{or} \quad \cos\theta_0 = \frac{\alpha}{\beta}\left(\frac{g_-^{f_1} + g_-^{f_2}}{g_-^{f_1} - g_-^{f_2}}\right). \tag{2.62}$$

Unlike the $W\gamma$ case, the position of the amplitude zero is dependent on the energy through $\alpha$ and $\beta$. However, this dependence is small, and in the limit of $s \gg M_V^2$, it lies at

$$\cos\theta_0 \simeq \begin{cases} \frac{1}{3}\tan^2\theta_W \simeq \phantom{-}0.1 & \text{for } d\bar{u} \to W^- Z,\\ -\frac{1}{3}\tan^2\theta_W \simeq -0.1 & \text{for } u\bar{d} \to W^+ Z.\end{cases} \tag{2.63}$$

At high energies, strong cancellations occur between the terms in square brackets, leaving only the $(\pm,\mp)$ and $(0,0)$ amplitudes non-zero. This can be seen in two steps.

First, for $s \to \infty$

$$\begin{aligned}
r_V &\to 0,\\
\beta &\to 1,\\
\alpha &\to 1,
\end{aligned}\tag{2.64}$$

thus, all the terms in the square brackets that include $X$ are proportional to $\frac{1}{s}$. As $Y \propto \frac{1}{s}$ this means that the full square bracket term scales as $\frac{1}{s}$. The factors before the square brackets grow as

$$1 \quad \text{for} \quad \mathcal{M}(\pm,\pm), \tag{2.65}$$

$$\sqrt{\frac{1}{\frac{1}{s}}} = \sqrt{s} \quad \text{for} \quad \mathcal{M}(\pm,0), \tag{2.66}$$

$$\sqrt{\frac{1}{\frac{1}{s}}} = \sqrt{s} \quad \text{for} \quad \mathcal{M}(0,\pm), \tag{2.67}$$

$$\sqrt{\frac{1}{\frac{1}{s^2}}} = s \quad \text{for} \quad \mathcal{M}(0,0). \tag{2.68}$$

Due to this, only the $(0,0)$ amplitude remains finite out of all those with a square bracket term as the $\frac{1}{s}$ and $s$ contributions cancel. Additionally, the $(\pm,\mp)$ contribution stays non-zero as it has no $s$ dependence at all. The amplitudes for the $(0,0)$ and $(\pm,\mp)$ states tend towards

$$\begin{aligned} \mathcal{M}(\pm,\mp) &\longrightarrow \frac{F}{\sin\theta}\left(\lambda_{\mathrm{w}} - \cos\theta\right)\left[\left(g_-^{f_1} - g_-^{f_2}\right)\cos -\theta\left(g_-^{f_1} + g_-^{f_2}\right)\right], \\ \mathcal{M}(0,0) &\longrightarrow \frac{F}{2}\sin\theta\frac{M_Z}{M_W}\left(g_-^{f_2} - g_-^{f_1}\right), \end{aligned} \tag{2.69}$$

in the limit $s \to \infty$. A derivation of these starting from Equations 2.61 is shown in Appendix B. When neglecting the weak hypercharge, this results in the squared amplitudes [79]

$$\begin{aligned} \left|\mathcal{M}_{WZ}^{\mathrm{TT}}\right|^2 &= \frac{g^4}{32}\left(\frac{s}{t} - \frac{s}{u}\right)^2 \sin^2\theta\left(1 + \cos^2\theta\right), \\ \left|\mathcal{M}_{WZ}^{00}\right|^2 &= \frac{g^4}{16}\sin^2\theta. \end{aligned} \tag{2.70}$$

It can be seen from this and $\left(\frac{s}{t} - \frac{s}{u}\right)^2 \propto \frac{\cos^2\theta}{\sin^4\theta}$ [27] that the TT states are dominant in the forward regions $|\cos\theta| \sim 1$ while the longitudinal polarizations dominate centrally with $\cos\theta \sim 0$. At full leading order, this means that in the $W^\pm Z$ center-of-mass frame, which is also the partonic center-of-mass frame, with $\theta = \frac{\pi}{2}$ the bosons are back to back with a rapidity[2] of $Y_W = Y_Z = 0$ and thus a rapidity difference of $\Delta Y_{WZ} = |Y_W - Y_Z| = 0$.

As such, the RAZ effect can be investigated by measuring the $\Delta Y(WZ)$ distribution and seeing a reduced cross-section around $\Delta Y(WZ) = 0$.

### 2.3.1 $W^\pm Z$ Production at Hadron Colliders

Measuring $W^\pm Z$ production at hadron colliders like the LHC provides two complications over the setup discussed so far in this section.

The first is the fact that either of the quarks initiating the process can come from either of the protons taking part in the collision. As such, it is not possible to determine the direction of the $q_1$ that defines the scattering angle in Equations 2.61. Due to this, the scattering angle is measured with respect to a fixed axis in practice, leading to a symmetric distribution that disturbs the exact zero specified in Equation 2.63. This is one of the reasons why it is useful to investigate the RAZ effect using the rapidity difference, where the reduction is expected at zero and thus not affected by the symmetrization.

The second complication originates from the unstable nature of the $W^\pm$ and $Z$ bosons. This necessitates the inclusion of the bosons' decays in defining the final state. The work in this thesis focuses solely on the fully leptonic final states due to its clean signature and good reconstructibility. This leads to the observed final state of $\ell^\pm \overset{(-)}{\nu}_\ell \ell'^+ \ell'^-$, with its corresponding Feynman diagrams already shown in Figure 2.3. Any number of additional jets is allowed to still be classified as signal, as these can always originate from initial state radiation.

[2] Rapidity is introduced in section 3.2 in Equation 3.1.

In the detector, other reactions can mimic the signal processes. These either come from processes at different coupling orders, ones with additional particles that are not reconstructed, or through the misidentification of one object as another. Some of the main backgrounds include the $ZZ$, $W^{\pm}Zjj$-$EW$, $t\bar{t}V$, $t\bar{t}$, $VVV$, $V\gamma$ and $Z+j$ processes, with $ZZ$, $t\bar{t}V$ and $t\bar{t}$ being the largest ones. Example Feynman diagrams for the first four of these processes are shown in Figure 2.5.

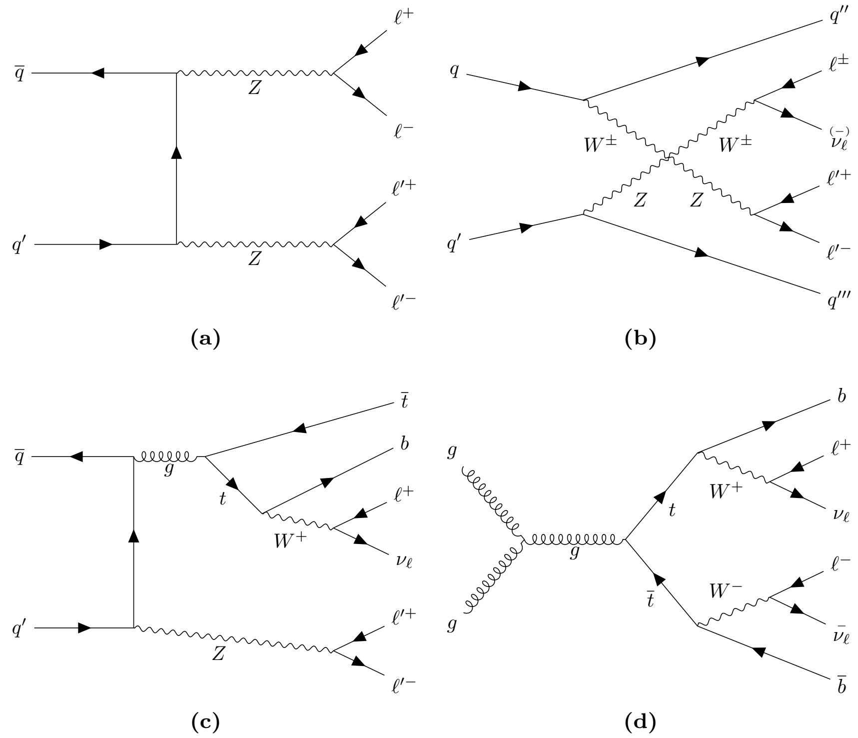



**Figure 2.5:** Example Feynman diagrams for background processes to $W^{\pm}Z$ production. Shown are the (a) $ZZ$, (b) $W^{\pm}Zjj$-$EW$, (c) $t\bar{t}Z$ and (d) $t\bar{t}$ processes.

The $ZZ$ process can mimic the $\ell^{\pm}\overset{(-)}{\nu}_{\ell}\ell'^{+}\ell'^{-}$ finale state when one of the four leptons is not reconstructed. For $t\bar{t}Z$, the produced particles can be the same as those from $W^{\pm}Z$ in case the $W$ from the top quark decays hadronically. If the $W$ decays leptonically, one of the additional leptons can fail to be reconstructed, which leaves three leptons in the final state.

The situation is slightly different for the $W^{\pm}Zjj$-$EW$ process. It naturally produces the expected leptonic finale state. The reason it is not considered as signal here is due to the unavailability of polarized samples at the start of the analysis. A gauge-invariant separation is possible due to the different coupling orders between $W^{\pm}Zjj$-$EW$ and the signal process.

An example of a reducible background is $t\bar{t}$, where the correct final state can not be reached through a particle not being detected. Instead, one of the jets has to be misidentified as a lepton.

# Chapter 3

# Experimental Setup

The data analyzed in this thesis was recorded by the ATLAS experiment [80] at the Large Hadron Collider (LHC) [81].

This chapter briefly summarizes the key aspects of both machines needed to understand their basic functionality as well as all related elements of this thesis. The information is taken from their respective technical reports [80–82] where more in-depth discussion and descriptions can be found.

## 3.1 The Large Hadron Collider

Situated beneath the Franco-Swiss border, the LHC serves as the preeminent particle accelerator, facilitating high-energy collisions crucial for probing the fundamental constituents of matter. The LHC was built in the 26.7 km long tunnel initially designed for the LEP experiment [83] at CERN. It is designed to accelerate protons and collide them at a center of mass energy of 14 TeV with an unprecedented luminosity of $1 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. It can also collide heavy ions at different energy ranges and luminosities [84].

The basic layout of the LHC follows the LEP tunnel geometry as shown in Figure 3.1. The LHC comprises eight arcs and eight straight sections, with each straight segment spanning approximately 528 meters. These straight sections can function as either experimental or utility insertions. The two general-purpose experiments are positioned on opposite sides of the ring. The ATLAS experiment is situated at Point 1, while the CMS experiment [85] occupies Point 5. Two additional experiments can be found at Point 2 and Point 8, respectively, housing the ALICE [86] and LHCb [87] experiments. A circular path for the beams is maintained by 1,232 dipole magnets. Additionally, 392 quadrupole magnets are employed to ensure beam focus [82].

### 3.1.1 Injector Chain

The magnet setup of the LHC does not allow for the acceleration of protons from rest up to their target energy of 6.5 TeV. Due to this, a series of pre-accelerators is used to bring the protons to an energy of 450 GeV before they are injected into the LHC. The accelerator complex present at CERN in 2022 is shown in Figure 3.2.

The LHC's injector chain is a series of particle accelerators that prepare and accelerate particles before they enter the main LHC ring. This chain consists of several accelerators, each progressively increasing the energy of the particles. A description of the chain alongside the main challenges to get it running is given in Ref. [89]. The main components, as they were used for proton-proton collisions during the time that the data for this thesis was taken, are quickly summarized below:

1. Linear Accelerator (Linac 2) [90, 91]: Protons are supplied by ionizing hydrogen gas and accelerated in a linear accelerator (Linac 2) to an energy of approximately 50 MeV.

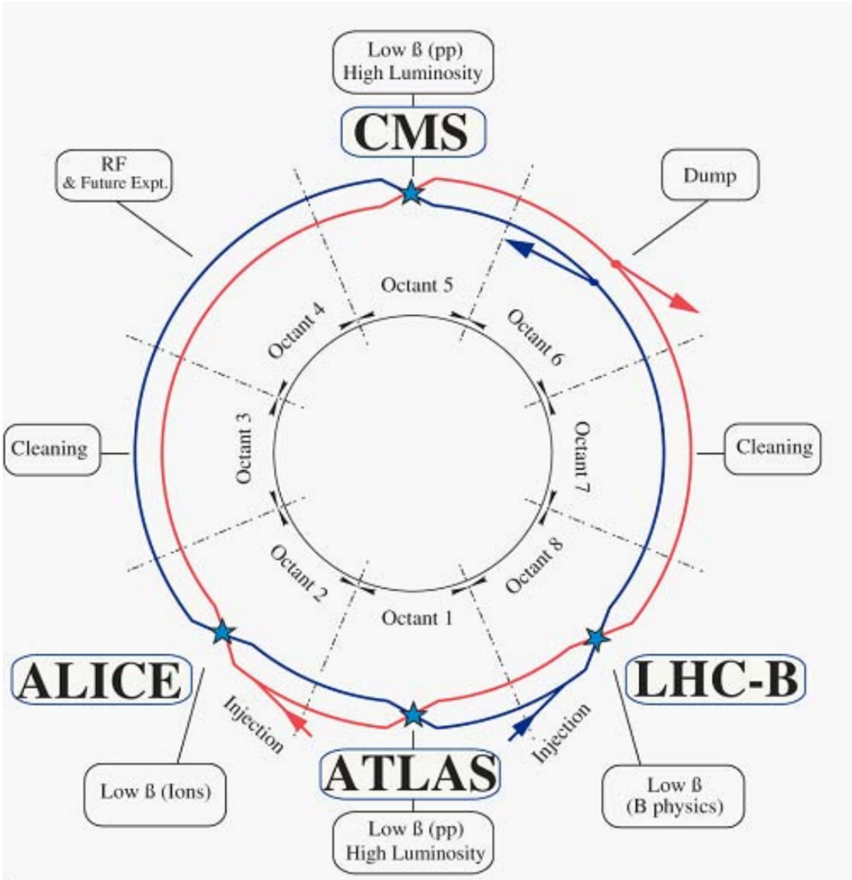



**Figure 3.1:** "Schematic layout of the LHC (Beam 1- clockwise, Beam 2 anticlockwise)." From Ref. [81].

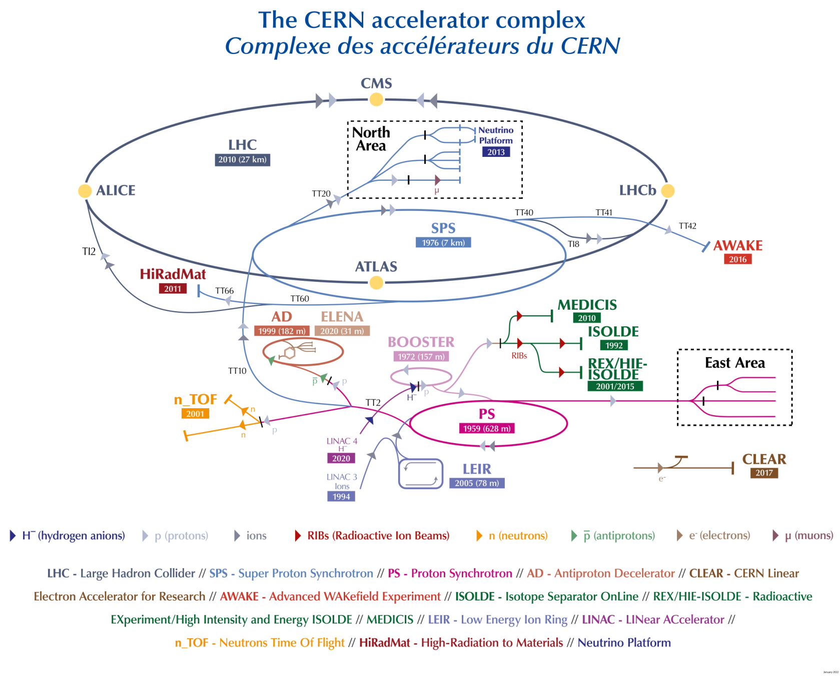



**Figure 3.2:** "The CERN accelerator complex layout in 2022." From Ref. [88].

2. Proton Synchrotron Booster (PSB) [92]: The proton beams from Linac 2 are further accelerated in the Proton Synchrotron Booster (PSB), increasing their energy to around 1.4 GeV.

3. Proton Synchrotron (PS) [93]: The PS takes the beams from the PSB and accelerates them to even higher energies, reaching approximately 25 GeV.

4. Super Proton Synchrotron (SPS) [94]: The Super Proton Synchrotron (SPS) serves as the final injector, taking the beams from the PS and accelerating them even closer to the speed of light, achieving energies of about 450 GeV. These beams are then injected into the main LHC ring.

The performance of the injector chain during Run 2 of the LHC, the period relevant to this thesis, is discussed in Ref. [95]. Starting in 2010, work was done to upgrade the LHC's injector chain. A description of this project can be found in Ref. [96].

### 3.1.2 Experiments at the LHC

The LHC and its injector chain host a variety of different experiments studying the Standard Model, supersymmetry, exotic isotopes, and cosmic rays.

At the time of writing, during Run 3 of the LHC, nine detector experiments hosted directly at the LHC are analyzing the particles produced in its collisions.

ATLAS (A Toroidal LHC Apparatus) [80]: One of the two general-purpose detectors at the LHC. It investigates a wide variety of topics. These range from probing all the aspects of the SM to physics beyond the Standard Model, like supersymmetry or dark matter. More information about the ATLAS experiment is given in Section 3.2.

CMS (Compact Muon Solenoid) [85]: The second general purpose detector at the LHC. It and the ATLAS detector serve as independent verification of each other's measurements and discoveries. Additionally, different choices were made when designing the two detectors, resulting in different strengths and weaknesses.

ALICE (A Large Ion Collider Experiment) [86]: Unlike the first two mentioned experiments, ALICE's main focus are not proton-proton collisions. Instead, it aims to study the strong interaction and, specifically, quark-gluon plasma through the results of heavy ion collisions.

LHCb (Large Hadron Collider beauty) [87]: An experiment with the main physics goal of studying CP violation in rare decays of charm and beauty or b hadrons. LHCb uses an unusual detector design. Instead of having a cylindrical detector around the whole collision point, its detector is concentrated in the forward direction.

TOTEM (TOTal Elastic and diffractive cross-section Measurement) [97, 98]: While technically an independent measurement, TOTEM is situated around the interaction point feeding the CMS experiment. Its main physics goals are a measurement of the total pp cross-section as well as studying elastic and diffractive scattering. With components positioned more than 200 m from either side of the collision point, it represents the LHC's 'longest' experiment.

LHCf (Large Hadron Collider forward) [99, 100]: Similar to TOTEM, this experiment sits around one of the general-purpose detectors. In this case, it lays 140 m from either side of the ATLAS collision point. It is capable of measuring neutral particles in the most extreme forward region up to zero degrees. This information can then be used to tune models utilized in the simulation of air showers induced by ultra-high-energy cosmic rays in the Earth's atmosphere.

MoEDAL-MAPP (Monopole and Exotics Detector at the LHC - MoEDAL Apparatus for Penetrating Particles) [101, 102]: Sharing the interaction point with LHCb, it is the first of a series of experiments dedicated to searches for new physics that the general purpose detectors are likely not sensitive to. While its main goal is the search for magnetic monopoles, it is generally looking for highly ionizing, feebly interacting, and very long-lived neutral particles.

FASER (ForwArd Search ExpeRiment) [103, 104]: Positioned 480 m away from the ATLAS detector it aims to measure light, extremely weakly-interacting particles. At low $p_T$, such particles would be produced in large numbers, even with low production cross-sections. As they are expected to have high energies and small cross-sections, they are also expected to travel macroscopic distances before decaying. Given the restricted size of the general purpose detectors together with their limited angular range, a detector placed $\approx 500$ m downstream of the collision point can fill this gap.

SND@LHC (Scattering and Neutrino Detector at the LHC) [105, 106]: Also located 480 m downstream from the ATLAS detector this experiment aims to study neutrinos. With its position slightly off-axis compared to FASER, it can measure neutrinos at larger angles.

### 3.1.3 Run 2 Performance

As the data for this thesis was taken during Run 2 of the LHC, its performance and configuration are discussed here briefly. More detailed information can be found in Ref. [107].

Run 2 of the LHC lasted from April 2015 until December 10th 2018, operating at a center-of-mass energy of $\sqrt{s} = 13$ TeV. During this time, it operated with a stable beam for a total of 5861 h and delivered an integrated luminosity of 160 fb$^{-1}$ to CMS and ATLAS. The integrated luminosity per year for the years 2011 to 2018 can be seen in Figure 3.3.

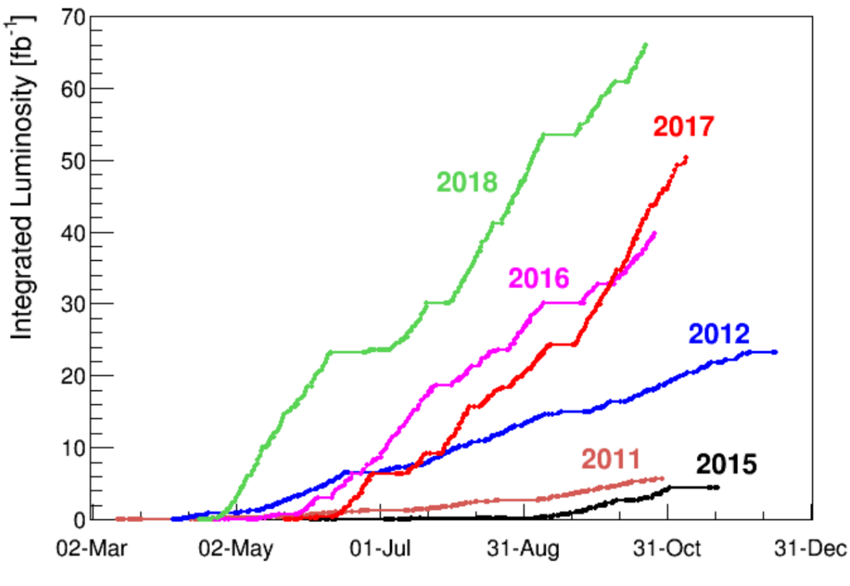



**Figure 3.3:** "Overview of the integrated LHC luminosity." From Ref. [107].

Set up for a nominal bunch spacing of 25 ns, operation was switched to that value in July 2015 after initially starting with a wider one of 50 ns. The number of bunches per ring fluctuated between 1868 and 2556, never quite reaching the design value of 2808. Meanwhile, the number of protons per bunch ranged from $1.1 \times 10^{11}$ to $1.25 \times 10^{11}$. The maximum was reached throughout parts of 2017, together with the minimal quoted value of bunches per ring. This leads to a higher than usual number of mean interactions per crossing $\langle\mu\rangle$ as shown in Figure 3.4.

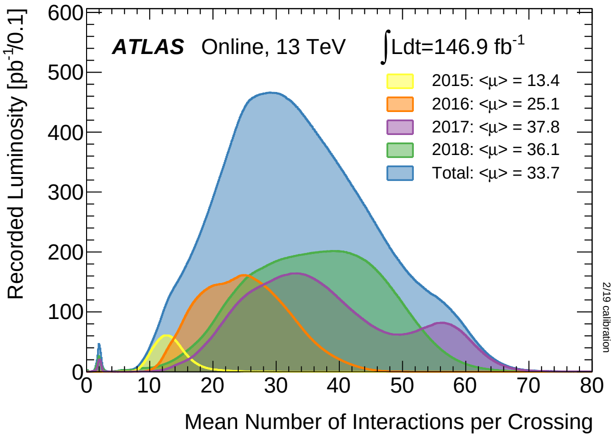



**Figure 3.4:** "**Number of Interactions per Crossing**": "Shown is the luminosity-weighted distribution of the mean number of interactions per crossing for the 2015 – 2018 pp collision data at 13 TeV center-of-mass energy. All data recorded by ATLAS during stable beams is shown, and the integrated luminosity and the mean mu value are given in the figure. The mean number of interactions per crossing corresponds to the mean of the Poisson distribution of the number of interactions per crossing calculated for each bunch. It is calculated from the instantaneous per bunch luminosity as $\mu = L_{\text{bunch}} \times \sigma_{\text{inel}}/f_r$ where $L_{\text{bunch}}$ is the per bunch instantaneous luminosity, $\sigma_{\text{inel}}$ is the inelastic cross-section which we take to be 80 mb for 13 TeV collisions, and $f_r$ is the LHC revolution frequency." From Ref. [108].

## 3.2 The ATLAS Experiment

ATLAS is one of the prominent components of the LHC and plays a pivotal role in capturing and analyzing the outcomes of these collisions with remarkable precision. It represents one of the largest collaborative efforts ever attempted in science, with over 5000 members and almost 3000 scientific authors across 182 Institutions and 43 countries [109].

This section provides an overview of key components of the ATLAS experiment: the physical ATLAS detector, object reconstruction procedures, and the data-taking and processing workflow.

### 3.2.1 The ATLAS Detector

As a general-purpose detector at the LHC, the ATLAS detector has to be able to measure and withstand very high interaction rates, radiation doses, particle multiplicities and energies. On top of that, the measurements have to be done with an incredibly high precision.

This section first describes the general coordinate system used for the ATLAS detector, followed by a very short introduction to its technical components. A detailed overview can be found in Ref. [80].

#### Coordinate system

The ATLAS detector uses a right-handed coordinate system with the interaction point of the collisions defined as the origin. The x-, y- and z-axis are defined as follows: The positive x-axis points from the origin to the center of the LHC ring, while the positive y-axis points upwards. Lastly, the z-axis points along the beam direction, with the positive side determined by the coordinate system being right-handed. Positions and directions are commonly described using the angles $\phi$ and $\theta$. Here, $\phi$ describes the azimuthal angle around the beam axis and $\theta$ the polar angle from the beam axis. The rapidity and pseudorapidity are then defined as

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \tag{3.1}$$

and

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

respectively. The transverse values (denoted by a subscript '$T$') $p_T$, $E_T$, $E_T^{miss}$ are defined in the x-y plane. Lastly, $\Delta R$ describes the distance in the pseudorapidity-azimuthal angle space as

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{3.3}$$

### Components

The basis for the design of all the other detector components lies in the nominal forward-backwards symmetry of the detector as well as its magnet configuration. It is comprised of a thin superconducting solenoid around the inner-detector cavity and three large superconducting toroids around the calorimeters. A cutaway view of the ATLAS detector with all of its components can be seen in Figure 3.5.

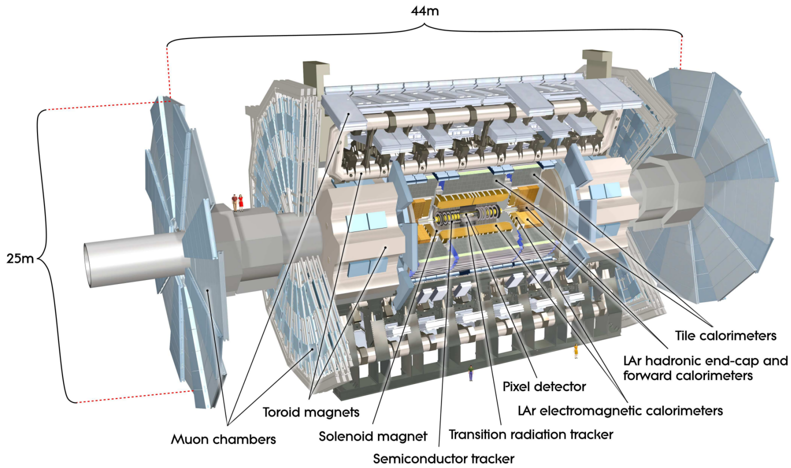



**Figure 3.5:** "Cut-away view of the ATLAS detector. The dimensions of the detector are 25 m in height and 44 m in length. The overall weight of the detector is approximately 7000 tonnes." From Ref. [80].

To be able to investigate as many physics processes as possible, the ATLAS detector needs to detect the resulting particles from each of them. A particle must travel a minimum distance of around 35 mm from the center to the border of the beam pipe in order to reach the detector. As such, only particles with a lifetime long enough to traverse this distance before decaying have any chance of being measured. An additional constraint for the possibility of efficient detection is the interaction via either the strong or electromagnetic force. Particles that only interact via the weak force or gravity have no realistic chance of being detected over the length of the ATLAS detector. As such, it is built with the target of detecting and separating as many of the remaining particles as possible. These include electrons, muons, photons, and light hadrons. For this purpose, ATLAS is built in layers that excel at detecting and measuring the properties of different classes of these particles. From the beam pipe going outwards, the main components are:

**Inner detector:** Submersed in a 2 T solenoidal field, it serves to perform pattern recognition as well as precise momentum and vertex measurements and identification of electrons. In the inner part, this is achieved using a combination of discrete high-resolution semiconductor pixel and stripe detectors. Straw-tube tracking detectors that can generate and detect transition radiation are used in the outer part. Covering the range $|\eta| < 2.5$, it can measure a particle's momentum and charge through the curvature of its path in the magnetic field.

**Electromagnetic calorimeter:** Consisting of lead-LAr detectors with accordion-shaped kapton electrodes and lead absorber plates, its main purpose is to measure the energy of light electromagnetically interacting particles. For this, it has high granularity to achieve good position and energy resolution. It is split into two regions. The first is the barrel part, which covers a range of $|\eta| < 1.475$. The two end-cap components cover the remaining range of $1.375 < |\eta| < 3.2$. The calorimeter consists of three sections in depth in the precision region $|\eta| < 2.5$ and two with coarser transverse granularity in the remaining region $2.5 < |\eta| < 3.2$.

**Hadronic calorimeter:** The hadronic calorimeter comprises multiple parts. The first, placed directly outside the EM calorimeter, is a scintillator-tile calorimeter covering $|\eta| < 1.7$ and using steel as the absorber. It is separated into a large barrel $|\eta| < 1.0$ and two smaller extended barrel cylinders $0.8 < |\eta| < 1.7$ on either side of the central barrel.

The Hadronic End-cap Calorimeter (HEC) and the Forward Calorimeter (FCal) again use liquid argon as their active material. They cover the ranges $1.5 < |\eta| < 3.2$ and $3.1 < |\eta| < 4.9$, respectively. They also differ in the absorber material they use. The HEC uses copper, while the FCal uses both copper and tungsten. Here, the copper component is optimized for electromagnetic measurements, while the tungsten ones aim to measure hadronic interactions.

**Muon system/Muon spectrometer:** The muon system is the outermost layer of the ATLAS detector and gives it its distinct overall appearance. It is composed of an air-core toroid system with a long barrel and two end-cap magnets. Its excellent muon momentum resolution is achieved using three layers of high-precision tracking chambers composed of Monitored Drift Tubes and Cathode Strip Chambers. The muon system covers the range $|\eta| < 2.7$.

A summary of the mentioned components, their resolution, and angular coverage are given in Table 3.1.

**Table 3.1:** "General performance goals of the ATLAS detector. Note that, for high-$p_T$ muons, the muon-spectrometer performance is independent of the inner-detector system. The units for $E$ and $p_T$ are in GeV." From Ref. [80].

| **Detector component** | **Required resolution** | **$\eta$ coverage** | |
|---|---|---|---|
| | | Measurement | Trigger |
| Tracking | $\sigma_{p_T}/p_T = 0.05\%\ p_T \oplus 1\%$ | $\pm 2.5$ | |
| EM calorimetry | $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$ | $\pm 3.2$ | $\pm 2.5$ |
| Hadronic calorimetry (jets) | | | |
| barrel and end-cap | $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$ | $\pm 3.2$ | $\pm 3.2$ |
| forward | $\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$ | $3.1 < \lvert\eta\rvert < 4.9$ | $3.1 < \lvert\eta\rvert < 4.9$ |
| Muon spectrometer | $\sigma_{p_T}/p_T$=10% at $p_T$ = 1 TeV | $\pm 2.7$ | $\pm 2.4$ |

### 3.2.2 Object Reconstruction

Objects used for analysis have to be reconstructed from the signals measured by the ATLAS detector. This chapter gives a summary of how the reconstruction works for the particles used in this thesis. More information about used working points and uncertainties are given in Chapters 5 and 7, respectively.

#### Electrons and Photons

This section summarizes the basic points of Ref. [110]. More information can be found there.

The first step in the reconstruction of electrons and photons is the creation of EM topo clusters. These are clusters formed from the electromagnetic and hadronic calorimeters that have a total energy above 400 MeV and a ratio of $\frac{E_{EM}}{E_{\text{tot}}} > 0.5$. Next, tracks are constructed in the inner detector, matched to clusters and refitted.

Electron candidates are then found as EM topo clusters with one or more matched tracks. Matching criteria and the procedure employed to handle multiple matches can be found in Ref. [111].

Photon candidates are either EM topo clusters that are matched to at least one conversion vertex or clusters matched to neither tracks nor conversion vertices. From these candidates, final clusters are built from all of those clusters that are matched to the candidates. Bremsstrahlung can cause one candidate to be reconstructed as two clusters. Due to this, multiple close-by clusters are merged.

Usually, not all of these reconstructed electrons and photons are used in an analysis. To select those particles that come from processes of interest, certain quality criteria are applied. The tighter the criteria, the higher the background rejection. However, that also decreases the efficiency for desired particles. The criteria used for this analysis are discussed later in Chapter 5.

#### Muons

Full information about muon reconstruction as well as identification criteria and efficiencies for identification, vertex association and isolation can be found in Ref. [112].

Reconstruction of muons is mainly based on information from the inner detector and the muon system. Additionally, calorimeter information can be used to determine track parameters and to account for cases with large energy loss.

In the muon system, first preliminary track candidates are built from segments using a loose constraint based on a parabolic trajectory. Here, a segment is created by individual hits in the system. More information from the muon system is added and a global $\chi^2$ fit of the trajectory is performed. From that, outlier hits can be removed, and missed hits can be added to the candidate. Furthermore, lower-quality tracks that share a lot of hits with higher-quality ones are removed. Using the complete detector information, five muon types are defined:

**Combined muons:** Represent muon system tracks matched to inner detector tracks.

**Inside-out combined muons:** Reconstructed from inner detector tracks that are extrapolated to the muon system and matched to at least three hits. The information from the inner detector is combined with the hits in the muon system and energy loss in the calorimeters.

**Muon-spectrometer extrapolated muons:** These are muons reconstructed from just the muon system if a track from there can not be matched to a track in the inner detector.

**Segment-tagged muons:** Inner detector tracks extrapolated to the muon system that tightly match a segment in the muon system are used to reconstruct segment-tagged muons.

**Calorimeter-tagged muons:** Lastly, these muons are reconstructed from inner detector tracks that can be matched to energy deposits in the calorimeters that are consistent with minimum-ionizing particles. Due to large background contamination, this algorithm requires tracks to have a $p_T$ of at least 5 GeV.

#### Jets

More information about the jet reconstruction algorithms used at ATLAS can be found in Ref. [113].

The main reconstruction method used is the anti-$k_T$ algorithm [114] with a radius parameter $R = 0.4$. The inputs for this algorithm come from Monte Carlo simulations or data. Here, the first step is calorimeter cells, which are clustered using the nearest neighbor algorithm. Current ATLAS analyses combine this with information from the inner detector to build **PFlow** [115] jets. One aspect is the subtraction of deposited energy from charged particles. This is replaced by momenta of inner detector tracks matched to those clusters. These **PFlow** jets exhibit better energy and angular resolution, reconstruction efficiency and pileup stability. Further pileup suppression is achieved using Jet vertex tagging (JVT) [116].

**Missing Transverse Momentum**

Details of the Missing transverse momentum (MET/$E_T^{\text{miss}}$) reconstruction in ATLAS can be found in Ref. [117]. This section summarizes the core information from that reference.

The $E_T^{\text{miss}}$ reconstructed in ATLAS is comprised of two main components. The first component arises from signals associated with *hard-events*, encompassing fully reconstructed and calibrated leptons, photons, and jets (referred to as hard objects). The second contribution to $E_T^{\text{miss}}$ originates from *soft-event* signals. These are comprised of charged-particle tracks linked to the hard-scatter vertex but not to any hard object.

The reconstruction algorithms for each of these contributions, as quickly summarized earlier in this chapter, are independent of each other. This can lead to double counting of contributions if they are reconstructed as multiple particles. To avoid this a *signal ambiguity resolution* is introduced to reject already counted terms.

### 3.2.3 Trigger and Data Taking

With collisions happening at a rate of roughly 40 MHz, the ATLAS trigger system has to provide an enormous rejection of minimum-bias processes. This is required, as physical limitations only allow for an event data recording rate of $\approx 1.2\,\text{kHz}$. Additionally, a high efficiency for the processes of interest has to be maintained to not hinder their exploration. To achieve this task, the ATLAS trigger system works in two stages.

First, the Level-1 (L1) trigger system uses a subset of the full detector information to reduce the event rate down to 100 kHz. It does so by using custom-made electronics and is mainly limited by the bandwidth of the readout. Looking primarily for high-$p_T$ muons, electrons/photons, jets, and $\tau$-lepton decays, as well as large $E_T^{\text{miss}}$, it makes its decisions and defines Regions-of-Interest (RoI's).

The second step, the High-Level-Trigger (HLT), combining the Level-2 and event filter steps from Run 1, uses the information from these RoI's and fast trigger algorithms to reject some events as early as possible. The final selection is performed using algorithms similar to those used for offline reconstruction and can access information from the full detector if necessary. Unlike the L1 trigger, the HLT is almost entirely based on commercially available computers and networking hardware. The average output rate of the HLT is 1.2 kHz, corresponding to a total reduction of a factor of $10^4$.

Figure 3.6 shows an overview of the trigger and data acquisition systems of the ATLAS detector during Run 2. More information on the full trigger system can be found in Refs. [80, 118].

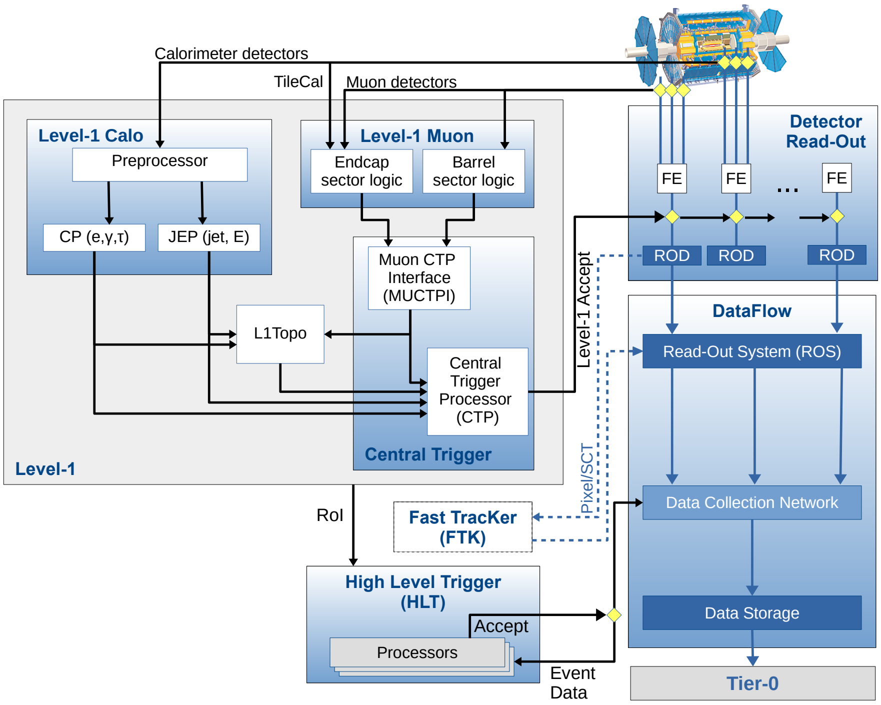



**Figure 3.6:** "The ATLAS TDAQ [Trigger and Data Acquisition] system in Run 2 showing the components relevant for triggering as well as the detector read-out and data flow." From Ref. [118].

# Chapter 4

# Simulation

The purpose of science is to build theories describing nature and verifying (or disproving) their correctness/usefulness by comparing these theories with reality. The complexity of the theories and high multiplicity of final states in particle physics, and especially proton–proton collisions, necessitate the use of powerful simulation software to generate predictions that can be compared to experimental data. In Section 4.1.1, this chapter offers a very brief overview of the general methodology of how these programs operate. A more general overview of that can be found in Ref. [119]. Sections 4.2 and 4.3 describe the data and the Monte Carlo samples used in this thesis, respectively.

## 4.1 MC Simulations

The simulation of proton-proton collisions happens in steps due to the extremely high complexities of the states that enter the detector as well as the vast energy scales. This section gives a brief overview of each of these steps in Section 4.1.1, followed by a description of how the combination of two of these steps works when going beyond leading-order predictions in Section 4.1.2.

### 4.1.1 Simulation Steps

Figure 4.1 shows a representation of the full simulation of a single event as generated by an MC event generator. The different colors represent the different steps of the simulation. This section will give a quick explanation of each of these steps.

#### Hard process

The simulation step that handles processes at the highest energy is the hard process represented by the central, dark red blob in Figure 4.1. It lays the ground for the main characteristics of kinematic distributions as well as the number of events that are expected to occur for a given process according to

$$N = \sigma \cdot \int L(t) dt. \tag{4.1}$$

Here, $L$ is the luminosity and $\sigma$ is the cross-section of the process of interest. The overall cross-section of a process depends on the particles involved and the considered kinematic regions. For any infinitesimal phase-space volume, the differential cross-section is proportional to the matrix element $\mathcal{M}$:

$$d\sigma \propto |M|^2 \, d\Phi. \tag{4.2}$$

For a given process, the matrix element can be determined from the Feynman rules for all the diagrams that connect the initial and final states. In general, there is an infinite number of diagrams connecting any two states with a physical transition. However, if the involved couplings are smaller than 1, diagrams that go beyond the smallest order in the couplings are

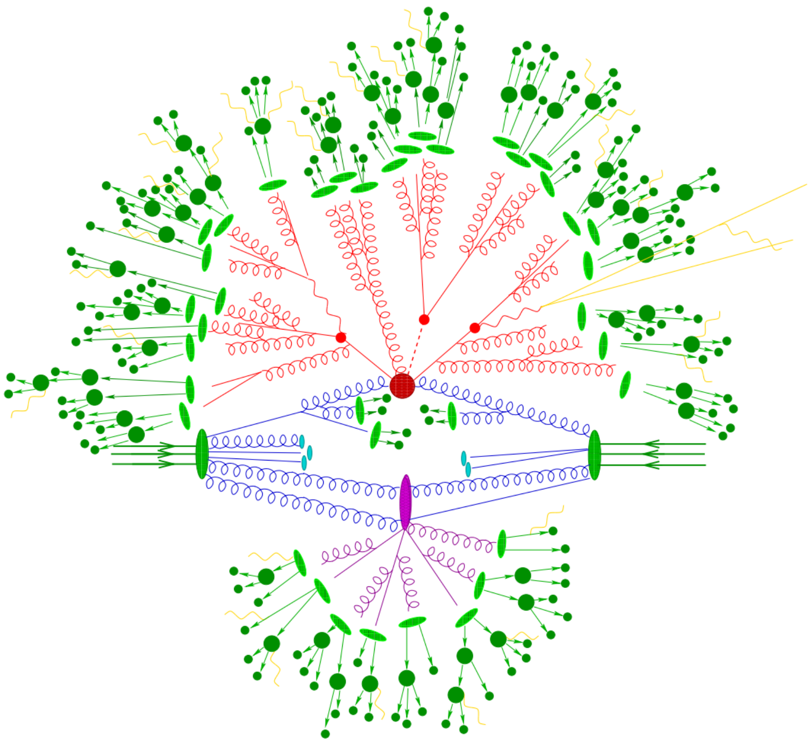


**Figure 4.1:** Representation of the simulation steps in an MC event generator for a $t\bar{t}h$ event. The most energetic part is the hard interaction (large, dark red circle). Afterward, the produced top quarks and the Higgs boson decay (small, bright red circles). At any of the points so far, additional QCD radiation can be produced (red lines). The resulting partons hadronize (bright green ovals), and some of the resulting hadrons subsequently decay (dark green circles). In any collision, secondary interactions can occur (purple oval) and progress through similar steps. Photons can be emitted during any of these steps (yellow lines). From Ref.[120].

suppressed. This lowest order for any given process is referred to as leading-order (LO), while those including an additional one are called next-to-leading order (NLO). Depending on the process, this additional order can come from any of the fundamental forces. If no specification is given, the additional coupling present is the strong one $\alpha_S$. Additional couplings of either the weak or electromagnetic force are usually grouped under NLO EW.

Using the factorization theorem [121], the cross-section can be calculated as [119]

$$\sigma = \sum_{a,b} \int_0^1 dx_a dx_b \int d\Phi_n \mathrm{PDF}_a(x_a, \mu_F)\mathrm{PDF}_b(x_b, \mu_F) \cdot \frac{1}{2x_a x_b s} \left|\mathcal{M}\right|^2 (\Phi_n; \mu_F, \mu_R). \quad (4.3)$$

Here, a and b are possible initial state partons. Depending on the process of interest, their flavor is not necessarily known and, as such, has to be summed over. Additionally, with the LHC being a proton-proton collider, the energy of each constituent quark can not be determined exactly. They are represented by the $x_i$ as the fraction of the proton's energy that the parton possesses. As such, they have to be integrated over according to parton distribution functions (PDF), which parameterize these fractions for the different partons and are usually obtained from external measurements. The factorization scale $\mu_F$ separates long and short-distance physics and describes the scale at which the transition from PDF to matrix element description happens. $\mu_R$ is the renormalization scale required for regularization. It represents the energy scale at which divergent loop integrals are regulated by adjusting coupling constants and masses. Both of these scales can be set by hand or determined dynamically by the processes and each event's kinematics. They are usually chosen to be nominally equal. The $s$ represents the Mandelstam variable encoding the overall energy of the process, and $\mathcal{M}$ is the matrix element for the transition from the initial to the final state for the process of interest.

Neglecting the PDFs, only the integral over the phase space $\Phi_n$ remains. This has a very high dimensionality of $3n - 4$, where $n$ represents the number of final state particles. This makes a classical analytical calculation extremely difficult. Due to that, the integration is usually done numerically using Monte Carlo methods [122, 123].

Complex procedures have been developed to achieve optimal and fast integration [124] and eventually event generation. Here, after the phase space integration, the phase space is sampled to produce events. Usually, the events are first produced with a weight proportional to the cross-section for easier integration and generation, and finally partially vetoed to unweight them. With that, each final event will have roughly the same weight, as is expected in data, while the kinematics still follow the expected distributions.

#### Parton shower

Usually, only a few, zero to two, QCD emissions are included directly in the hard matrix element. All others have to be produced via a Markov chain approach called parton shower [125]. Hereby, partons are emitted successively according to splitting functions until their energy is too low. Care has to be taken when handling recoils and the change of scales throughout the emissions. Additionally, splitting functions can diverge for low energies and small angles. A detailed overview of parton shower generators can be found in Ref. [126]. This phase is shown in Figure 4.1 in the form of bright red lines.

#### Hadronization

Color confinement dictates that the quarks and gluons created by the hard processes and the parton shower can not be observed directly. Instead, only hadrons or their decay products can be measured in the detector. As hadronization happens at low energies where perturbative calculations of QCD break down, phenomenological models fitted to data are used. The light green blobs in Figure 4.1 represent this stage of the event simulation.

Two main approaches exist to model this effect. These are the cluster [127, 128] and string [129, 130] models, both ideas from Ref. [131]. The string model is based on confinement and the linear potential between quarks. In this model, quarks and antiquarks are connected by strings representing the potential energy between them. At large enough separations, the potential energy becomes so large that it is energetically favorable to break the string and create a new quark-antiquark pair. Meanwhile, the cluster model is derived from the concept of preconfinement [132]. This comes from the observation that late in a parton shower at low energies, partons are grouped in colorless clusters. These can then decay into the observed final-state hadrons. A slightly more in-depth, but still pedagogical, explanation of this is given in Ref. [133].

The author of this thesis implemented an interface [134] in Sherpa, which uses the cluster model [135], for hadronization with Pythia8, using the string model [136, 137], to evaluate uncertainties arising from the different approaches.

#### Decays

A lot of the hadrons produced during hadronization, as well as $\tau$-leptons and potentially other BSM particles produced in the hard process, are unstable and decay before or in the detector. These decays also have to be simulated, shown in dark green in Figure 4.1, while taking great care to correctly model spin correlations. Some Monte Carlo generators like Sherpa [138], Pythia [136, 139] and Herwig [140] are able to handle these decays themselves. However, there are also dedicated packages available like TAUOLA [141] and EvtGen [142] that can be interfaced with.

To achieve the best possible results, current implementations generally use both calculations from matrix elements as well as results from phenomenological models based on observed data to model the decays of heavy, unstable particles.

#### QED radiation

The yellow lines in Figure 4.1 indicate QED radiation. This can occur at any stage of the event generation and must be simulated correctly. One approach to do that is by employing functionally the same algorithm used for the QCD shower. Another option is to follow the Yennie-Frautschi-Suura (YFS) formalism [143].

#### Underlying event

The last elements of Figure 4.1 that have not been discussed are those in purple. They represent the so-called underlying event. Here, remnants of the initial state hadrons can participate in further hard interactions. A general discussion of the underlying event and its components is given in Ref. [144].

As factorization theorems do not handle this scenario, their simulation is based on phenomenological models. The most significant contribution to the underlying event is multiple parton interactions. A model for simulating these was first given in Ref. [145].

### 4.1.2 Matching and Merging

Parton showers give approximations of the contributions from higher-order diagrams. Some processes, like the $W^{\pm}Z$ production discussed in this thesis, need to go beyond these approximations and require more accurate predictions.

There are two common approaches for calculating higher-order predictions and harmonizing them with parton shower setups to avoid double counting. A summary will be given below, while a more in-depth discussion of matching and merging can be found in Ref. [126].

#### Matching

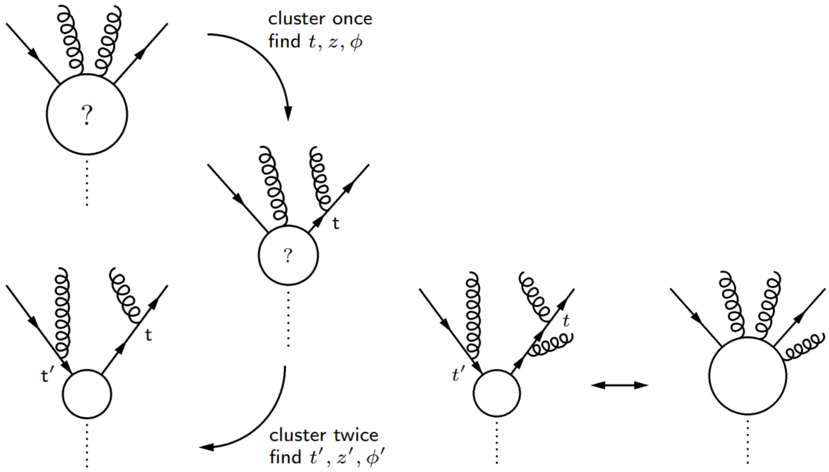
For matching, matrix elements of NLO or higher are calculated. Subsequently, the parton shower contributions from lower orders are computed and subtracted from the higher order calculation as these would be the components that would be double counted. The result can then be processed as normal by the parton shower. Two of the most widely used approaches for matching are the aMC@NLO [146] and Powheg [147, 148] methods. A comparison of the two methods can be found in Ref. [149].

#### Merging

In the second approach, additional (separate) tree-level calculations are performed with additional real emissions. All of these processes have to be combined with each other and the parton shower. This also suffers from issues of double counting as both the additional tree-level matrix elements and the parton shower include effects from real corrections.

The basic idea to avoid this double counting in merging is to veto parton shower emissions that are already included in the additional matrix elements. For that, it is necessary to find a mapping between the matrix element configurations and the parton shower branchings. This requires the parton shower calculation to be invertible. Then, any $n+1$, and in general $n+k$, configuration can be mapped back to the original $n$ particle configuration. A representation of this can be seen in Figure 4.2a.

Afterward, the shower can add further emissions at any point in the configuration. This is shown in Figure 4.2b, demonstrating a "truncated shower", where the evolution is truncated at a scale where gluons from the hard scattering calculation have to be reimplemented. When a parton shower emission gives the same multiplicity as one of the hard matrix elements and the emission is hard (constitutes its own jet), it has to be vetoed to avoid double counting. Instead, it should be simulated from the matrix element as shown in Figure 4.2b. This introduces the merging scale $Q_{cut}$, which is a parameter of the merging algorithm and usually of the order $\mathcal{O}(25\,\text{GeV})$.




**(a)** **(b)**

**Figure 4.2:** (a): "Sketch of a clustering sequence identifying the $e^+e^- \to q\bar{q}$ inclusive reaction in a $e^+e^- \to q\bar{q}gg$ final state." (b): "Sketch of a truncated parton-shower emission in the existing history and its correspondence to the higher-order tree-level matrix element." From Ref. [126].

Multiple schemes exist for merging such *multi-leg* setups with parton showers. Two common ones are the MLM [150] and the CKKW (Catani-Krauss-Kuhn-Webber) [151] schemes. Comparisons of these can be found in Refs. [150, 152].

For the CKKW algorithm, an updated version called CKKW-L [153, 154] exists. It is implemented in Pythia8 and Sherpa, with the former being used for the signal samples of this thesis. These are described in Section 4.3.1. The basic steps are the same, but the CKKW-L algorithm takes a different approach to calculating the Sudakov factors and vetoing in the shower. Specifically, CKKW-L builds a full cascade history with intermediate steps. Then, contrary to CKKW, where the smallest scale is always chosen for each clustering, all possible ordered shower histories are considered, and one is chosen according to a probability proportional to the product of the relevant branching probabilities. More detailed descriptions and further differences can be found in the original papers and the comparison references mentioned above.

## 4.2 Data Samples

The data used in this thesis was taken during Run 2 of the LHC, spanning from 2015 to 2018. Only proton-proton collisions at an energy of 13 TeV with a bunch-crossing of 25 ns are analyzed. Good run lists (GRL) remove events recorded during periods when the detector was not functioning correctly [155]. A representation of their impact on the luminosity can be found in Figure 4.3 and the selected GRLs are given in Appendix A.1. The data comprises a total integrated luminosity of 140.1 $\mathrm{fb}^{-1}$ [156, 157]. This, as well as the calibrated uncertainty of 0.83 % [156] was determined using the LUCID-2 detector [158].

## 4.3 MC Samples

This section summarizes the signal and background samples used in the polarization fraction and unfolding measurements. First, the polarized signal samples are discussed in Section 4.3.1. Next, unpolarized samples of the signal processes are introduced in Section 4.3.2. Finally, background samples are summarized in Section 4.3.3.

After their simulation, as described below, every generated event is passed through the

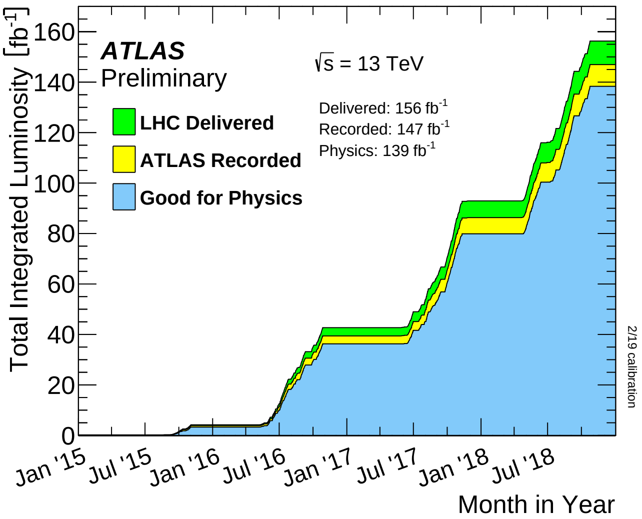



**Figure 4.3: "Total Integrated Luminosity and Data Quality in 2015-2018"**: "Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions at 13 TeV centre-of-mass energy in 2015-2018. The complete pp data sample in 2018 is shown. The delivered luminosity accounts for the luminosity delivered from the start of stable beams until the LHC requests ATLAS to put the detector in a safe standby mode to allow a beam dump or beam studies. The recorded luminosity reflects the DAQ inefficiency, as well as the inefficiency of the socalled "warm start": when the stable beam flag is raised, the tracking detectors undergo a ramp of the high-voltage and, for the pixel system, turning on the preamplifiers. The data quality assessment shown corresponds to the All Good efficiency shown in the 2015-2018 Full Dataset DQ tables here [Hyperlink to: `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/DataQualityResults`]. The All Good Data Quality criteria require all reconstructed physics objects to be of good data quality." From Ref. [108].

full ATLAS detector simulation based on GEANT4 [159, 160]. Afterward, reconstruction is performed with the same algorithms used for the data. Additional bunch crossings are added to the MC samples proportionally to the numbers occurring in the respective data periods. This is done to simulate pileup effects.

### 4.3.1 Signal Simulations for Polarized $W^{\pm}Z$ Production

The main polarized signal samples used in this thesis are produced with MadGraph 2.7.3 [74]. They were validated and requested by the author of Ref. [161] and make use of the feature introduced in version 2.7.0 to specify the polarization state of particles [75]. For massive spin-1 bosons, the available options are transverse and longitudinal polarization. It is currently not possible to separate left- and right-handed helicity states.

The samples are split by the bosons' decay channels. Specifically, one set of samples is generated where at least one boson decays into a $\tau$-lepton, while another set requires zero $\tau$-leptons. For the case without $\tau$-leptons, an example process definition for the case with two longitudinally polarized bosons is given by:

```
generate p p > wpm{0} z{0}, w+ > l+ vl, w- > l- vl~, z > l+ l-.
```
(4.4)

As described in Section 2.2, polarization is not a Lorentz invariant quantity. It is thus necessary to specify the frame that determines the boson polarization in the helicity basis. MadGraph's default value is the parton-parton center-of-mass frame. While that is identical to the desired $W^{\pm}Z$ rest frame at leading order, that is not the case anymore when including additional jets and the parton shower. Thus, the frame is specified using the `me_frame` setting.

For that, a collection of particles is specified. Their combined rest frame is then used as the defining frame. The particles are specified by the order they appear in the process definition, with intermediate particles substituted by their resolved decays. So for a definition

```
generate pp > w+{0} z{0} j, w+ > l+ vl, z > l+ l-,
```

the following matchings are given:

| | | | | | | |
|---|---|---|---|---|---|---|
| $q_1$ | $q_2$ | $l+$ | $vl$ | $l+$ | $l-$ | $j$ |
| 1 | 2 | 3 | 4 | 5 | 6 | 7. |

Thus, the final setting used is

```
'me_frame' : '3,4,5,6'
```

to select the $W^{\pm}Z$ center-of-mass frame. The notation is explained in short in Ref. [162, 163] and in more detail in Ref. [75].

Part of this thesis focuses on events with high transverse momenta up to $p_T^Z > 200\,\text{GeV}$. To produce enough events in that region to allow for a good statistical analysis while still keeping the generation time reasonable, the sample production is split. Separate sets of samples are produced for $p_T^Z < 150\,\text{GeV}$ and $p_T^Z \geq 150\,\text{GeV}$.

MadGraph is currently only able to simulate polarized processes at LO. Due to this, and knowing that the NLO corrections to $W^{\pm}Z$ production are sizeable [164], some of these corrections are recovered by including an additional jet in the matrix element. The showering and merging of the two matrix element calculations is done using Pythia8 8.244. The merging algorithm used is CKKW-L with a merging scale of 25 GeV.

Modeling of the PDF is done nominally via the `NNPDF3.0NLO` PDF set [165]. Its variations, as well as some other sets, are used when evaluating the uncertainty caused by this choice. This is discussed in more detail in Chapter 7. The sample requests can be found in Ref. [166], while the final job options for the case of two longitudinal bosons with $p_T^Z \geq 150\,\text{GeV}$ can be found in Ref. [167]. The other job options are stored under equivalent links with IDs 507020 through 507034. A summary and some additional information about the signal samples can be found in Appendix A.2.

### 4.3.2 Signal Simulations for Unpolarized $W^{\pm}Z$ Production

The lack of MC generators with the ability to generate polarized samples at NLO[1] necessitates the use of a leading order + merged setup. It is important to ensure that this, as well as the approximations made due to the decay chain syntax and the lack of polarization interference effects, do not introduce significant mismodeling. One step in doing so is a comparison against simulations that do not make these approximations and contain full NLO effects. This is discussed in Chapter 6. For this and related purposes, three pre-existing MC samples are used.

The first is generated with Powheg+Pythia8 with NLO QCD matrix elements matched to parton showers (PS). The CT10NLO PDF set is used for the hard-scattering process, while Pythia8 uses the AZNLO CTEQ6L1 tune [161].

Another sample used for this cross-check is generated at NLO in QCD with MadGraph5_aMC@NLO and Pythia8. Unlike the Powheg+Pythia8, this one also includes up to two jets directly from the matrix element. They are matched and merged to the parton shower using the FxFx merging scheme [74]. The merging scale is set at 25 GeV and the parton distribution function set is `NNPDF3.0`. The default dynamic renormalization and factorization scale set by MadGraph5_aMC@NLO (option "-1") is used [161].

Lastly, a sample generated with Sherpa 2.2.2 is used. This sample is also chosen partially for deriving concrete uncertainties due to NLO corrections in Chapter 7 and used to generate pseudo data for early fit validations. However, these are not shown in this thesis. The sample

[1] Recent developments of Sherpa make it now useable for this purpose [70]. However, these were not available in time to be used in this work.

is produced with up to 3 additional jets from the matrix element. The processes with zero and one partons are generated at NLO, while the case with two and three is generated at LO. A merging cut of 20 GeV is used alongside the `NNPDF3.0` PDF set [161]. Matching and merging are performed using the MC@NLO [149] and MEPS@NLO [168, 169] methods, respectively.

Additional information about these samples is also included in Appendix A.2.

### 4.3.3 Simulations for Background Processes

Besides the signal process described above, there are also other physical processes that lead to the same final state in the detector. Those that contain the final state with at least three leptons form the irreducible background and are estimated using MC samples directly. Such processes can be confused for the signal when some additional particles fall outside of the detector range or are not reconstructed due to the limited efficiency. The clearest example of this is $ZZ$ production, where one of the leptons can fail to be reconstructed, leading to the identical three-lepton final state expected for $W^{\pm}Z$ production. Other processes only mimic the signal process when another object, like a jet, is misreconstructed as a lepton. These backgrounds are called reducible and are estimated using a data-driven approach, as explained in Refs. [24, 26, 161, 170, 171]. However, these also make use of MC samples. An overview of samples used for the estimation of the backgrounds is taken from Ref. [161] and given below:

1. $ZZ$: Simulated with Sherpa 2.2.2 with NLO QCD accuracy for zero and one jet and LO for two and three jets.
2. $t\bar{t}Z$ and $t\bar{t}W$: Generated at NLO QCD with MadGraph5_aMC@NLO.
3. $VVV$: Sherpa 2.2.2 is used to generate these samples at NLO accuracy with no additional jets. One and two additional jets are also included directly from the matrix element. However, these are only simulated at LO.
4. $WZjj$-$EW$: Electroweak $W^{\pm}Z$ production is also considered as a background. It is simulated at LO accuracy using Madgraph+Pythia8 and includes all processes that produce an $\ell\nu\ell\ell jj$ final state with a coupling order of $\alpha_{\mathrm{EW}}^{6}$, regardless of the presence of actual $W^{\pm}$ and $Z$ propagators.
5. $t\bar{t}$: Simulated with Powheg+Pythia8 at NLO.
6. $Z\gamma$ and $W\gamma$: Version 2.2.2 of Sherpa is used to generate these samples at NLO accuracy.
7. $Z + j$: Nominally generated with Powheg+Pythia8 at NLO. Photos++ [172, 173] is used to simulate the final QED radiation.

Some more information about the utilized background samples can be found in Ref. [161] and in Appendix A.3.

# Chapter 5

# Selection

To measure the signal process, it is necessary to separate it from all other ones occurring in the vast number of LHC collisions. One step in doing so is requiring the collision result to contain the expected final state particles. For that, it is first necessary that the event contains objects that can be reconstructed as these particles. Depending on the exact particle kinematics, there can be a significant chance that objects are reconstructed incorrectly. To minimize the effect of this on the analysis, multiple criteria are required to be fulfilled for a reconstructed particle to be used in the analysis. These criteria are discussed in Section 5.1. Afterward, criteria are applied to the event as a whole, such as requiring the existence of a pair of leptons that could reasonably come from a $Z$ boson. These sets of criteria are summarized in Section 5.2. Lastly, the selection applied at the truth level is discussed in Section 5.4.

The implementation of the reconstructed level analysis was initially done by the ATLAS group at the University of Science and Technology of China [174, 175] with early cross-checks and validations performed by the author of this thesis [176]. This was later taken over by the author of Ref. [161], who also performed the truth level analysis using the RIVET toolkit [177]. As such, the content of this chapter follows Ref. [161].

The reconstructed level analysis is done using Analysis Release 21.2.188 [178]. Object calibrations and scale factors are applied following the recommendations of the combined performance groups [179]. The samples use the STDM5 derivations (`pTags` 4097/4095), where events with three leptons ($e$ or $\mu$) and tight slimming are preselected [180].

## 5.1 Object Selection

The main objects of interest in this analysis are electrons and muons. Their selection is described in Section 5.1.1. The analysis also makes use of jets and missing transverse momentum. The criteria for these objects to be used in the analysis are described in Sections 5.1.2 and 5.1.3, respectively. Lastly, Section 5.1.4 summarizes the overlap removal procedure applied to avoid using one underlying object as both a lepton and a jet.

### 5.1.1 Lepton Object Selection

Electrons and muons undergo a three-tiered process of object selection, categorized as baseline leptons, $Z$ leptons, and $W$ leptons. These selection criteria are implemented sequentially, with each level designed to be more stringent and encompass a narrower set of candidates than the preceding one. Consequently, every $W$ lepton also qualifies as a $Z$ lepton, and both $W$ and $Z$ leptons meet the criteria for baseline leptons. Baseline leptons are subject to the least restrictive selection criteria to efficiently screen out four-lepton events. Signal leptons must adhere to the more rigorous standards set by $W$ and $Z$ lepton selections to minimize the impact of fake-lepton backgrounds.

### Muon Object Selection

Before any of the following selections are applied, the momenta of all muons are calibrated. Baseline, $Z$, and $W$ selection criteria are summarized in Table 5.1.

The $p_\mathrm{T}$ progressing from $p_\mathrm{T} > 5\,\mathrm{GeV}$ for the baseline selection, through $p_\mathrm{T} > 15\,\mathrm{GeV}$ for $Z$ muons up to $p_\mathrm{T} > 20\,\mathrm{GeV}$ for $W$ ones clearly show the progression from loser to tighter cuts. Another one is the requirement of $|\eta| < 2.7$ for baseline muons. This, alongside the other criteria, is looser than the $|\eta| < 2.5$ applied to $Z$ and $W$ muons to increase the rejection of the $ZZ^{(*)} \to \ell\ell\ell'\ell'$ process. See Section 5.2.1 for how this is achieved exactly.

Even for baseline muons, those with $|\eta| < 2.5$ are required to pass the lepton-vertex association selection [181]. There are additional requirements on the significance of the track's transverse impact parameter relative to the beamline, as well as the longitudinal impact parameter, $z_0$, in the form of $|d_0/\sigma_{d_0}| < 3$ and $|z_0 \cdot \sin(\theta)| < 0.5\,\mathrm{mm}$, respectively. Here $z_0$ describes the difference between the value of $z$ of the point on the track at which $d_0$ is defined and the longitudinal position of the primary vertex.

The `MuonSelectionTool` [182] defines working points which classify muons according to a selection of reconstruction quality criteria. Different working points have different levels of efficiency and background rejections. Baseline muons are required to pass the `Loose` quality selection. Unlike those criteria used for the $Z$ and $W$ muons, this one includes calorimeter-tagged muons. See Section 3.2.2 for a description of the types of reconstructed muons. $Z$ muons use the *Medium* quality working point, while $W$ muons are required to pass the criteria defined by the `Tight` working point.

When it comes to isolation requirements, the `PflowLoose_FixedRad` one is used for baseline and $Z$ muons, while $W$ muons use the stricter `PflowTight_FixedRad` requirement.

**Table 5.1:** Three levels of muon object selection used in the analysis. Content from Ref. [161].

| **Muon object selection** | | | |
|---|---|---|---|
| Selection | Baseline selection | $Z$ selection | $W$ selection |
| $p_\mathrm{T} > 5$ GeV | ✓ | ✓ | ✓ |
| $\lvert\eta\rvert < 2.7$ | ✓ | ✓ | ✓ |
| `Loose` quality | ✓ | ✓ | ✓ |
| $\lvert d_0^{\mathrm{BL}}/\sigma(d_0^{\mathrm{BL}})\rvert < 3$ *(for* $\lvert\eta\rvert < 2.5$ *only)* | ✓ | ✓ | ✓ |
| $\lvert\Delta z_0^{\mathrm{BL}} \sin\theta\rvert < 0.5$ mm *(for* $\lvert\eta\rvert < 2.5$ *only)* | ✓ | ✓ | ✓ |
| `PflowLoose_FixedRad` isolation | ✓ | ✓ | ✓ |
| Overlap removal | ✓ | ✓ | ✓ |
| $p_\mathrm{T} > 15$ GeV | | ✓ | ✓ |
| `Medium` quality | | ✓ | ✓ |
| $p_T > 20$ GeV | | | ✓ |
| `Tight` quality | | | ✓ |
| `PflowTight_FixedRad` isolation | | | ✓ |

### Electron Object Selection

The calibration of the momentum for raw reconstructed electrons involves utilizing the `ElectronPhotonFourMomentumCorrection` package [183]. Every electron must satisfy object quality criteria, ensuring the exclusion of electrons with clusters influenced by the presence of dead FEB and high voltage regions of the detector. The selection process for candidate electrons involves three levels, as outlined in Table 5.2. These utilize the electron CP recommendations [184].

The transverse momentum criteria for electrons mirror those for muons. They progress through $p_\mathrm{T} > 5\,\mathrm{GeV}$,$p_\mathrm{T} > 15\,\mathrm{GeV}$ and $p_\mathrm{T} > 20\,\mathrm{GeV}$ for baseline, $Z$ and $W$ electrons.

Here, $p_T$ is calculated from the four-vector of an electron based on its calorimetric energy, while direction comes from the tracker.

Also, like muons, electrons are required to have $|\eta| < 2.5$. In this case, this refers to the value in the tracking volume. An additional criterium of $|\eta^{\text{cluster}}| < 2.47$ for the reconstruction of the cluster. While baseline electrons are allowed to be reconstructed in the region of $1.37 < |\eta^{\text{cluster}}| < 1.52$, this is not allowed for those that pass the $Z$ and $W$ selections. This exclusion is due to the calorimeter crack.

Each level of the electron selection imposes likelihood (LH) identification criteria on electron candidates [185, 186]. The criteria progress through `LooseLH+BLayer`, `MediumLH` and `TightLH` identification. Their respective efficiencies are in the ranges $84-96\%$, $72-93\%$ and $68-88\%$ for electrons with $10 < p_{\text{T}} < 80\,\text{GeV}$. Like muons, electrons are required to pass the lepton-vertex association selection [181]. They also have the same requirements on $z_0$ of $|z_0 \cdot \sin(\theta)| < 0.5\,\text{mm}$. However, the criterium on the significance of the track's transverse impact parameter is changed to $|d_0^{\text{BL}}/\sigma(d_0^{\text{BL}})| < 5$. Isolation requirements of `Loose_VarRad`, `HighPtCaloOnly` and `Tight_VarRad` are required for electrons to pass the baseline, $Z$ and $W$ selections, respectively.

Ambiguities between electrons and photons are removed using the `EGammaAmbiguityTool`. This is only done for the $W$ selection. Its role is to suppress the $Z\gamma$ background by requiring $W$ selected electrons not to pass photon reconstruction. The introduced inefficiency on true electrons from $W$ decays is roughly $\sim 5\,\%$. The dedicated `DFCommonAddAmbiguity` tagger [187] serves as another rejection for converted photons misidentified as electrons. The inefficiency here is $\sim 0.5\,\%$. This is significantly outweighed by its reduction of the $Z\gamma$ backgrounds by $35\,\%$ in the inclusive $W^{\pm}Z$ event selection, which is described in Section 5.2.1.

A correction to the MC due to the misidentification of the charge of the electrons is applied. Electron charge misidentification rates differ between Monte Carlo simulations and data. Scale factors correcting for this, as well as related uncertainties, are obtained using the official `EgammaChargeMisIdentificationTool` [188] tool. These scale factors are applied to all $Z$ and $W$ selected electrons.

**Table 5.2:** Three levels of electron object selection used in the analysis. Content from Ref. [161].

| **Electron object selection** | | | |
|---|---|---|---|
| Selection | Baseline selection | $Z$ selection | $W$ selection |
| $p_{\text{T}} > 5$ GeV | ✓ | ✓ | ✓ |
| Electron object quality | ✓ | ✓ | ✓ |
| $\|\eta^{\text{cluster}}\| < 2.47$, $\|\eta\| < 2.5$ | ✓ | ✓ | ✓ |
| `LooseLH+BLayer` identification | ✓ | ✓ | ✓ |
| $\|d_0^{\text{BL}}/\sigma(d_0^{\text{BL}})\| < 5$ | ✓ | ✓ | ✓ |
| $\|\Delta z_0^{\text{BL}} \sin\theta\| < 0.5$ mm | ✓ | ✓ | ✓ |
| `Loose_VarRad` isolation | ✓ | ✓ | ✓ |
| Overlap removal | ✓ | ✓ | ✓ |
| $p_T > 15$ GeV | | ✓ | ✓ |
| Exclude $1.37 < \|\eta^{\text{cluster}}\| < 1.52$ | | ✓ | ✓ |
| `MediumLH` identification | | ✓ | ✓ |
| `HighPtCaloOnly` isolation | | ✓ | ✓ |
| $p_T > 20$ GeV | | | ✓ |
| `TightLH` identification | | | ✓ |
| `Tight_VarRad` isolation | | | ✓ |
| `DFCommonAddAmbiguity` $\leq 0$ | | | ✓ |

### 5.1.2 Jet Object Selection

The jets used in this analysis are reconstructed from particle flow objects using the anti-$\mathrm{k_t}$ algorithm with a radius parameter of $\Delta R = 0.4$. Jets are required to have $p_\mathrm{T} > 30\,\mathrm{GeV}$ and $|\eta| < 4.5$ to be considered. Calibration and uncertainties are done according to the consolidated jet recommendations from ATLAS [189].

Jet vertex tagging is used to remove pile-up jets. This is done following the recommendations in Ref. [190] by using the pile-up jet vertex taggers `JVT` for central jets and `fJVT` for forward jets. Specifically, `JVT` uses the `Tight` working point while `fJVT` makes use of its `Loose` working point as stated in the recommendations detailed in Ref. [190]. Also, in accordance with the recommendations, the taggers are applied to all jets with $p_T < 60\,\mathrm{GeV}$. Data/MC calibrations from May 2020 are used to correct the MC efficiencies. These are defined in Ref. [190].

### 5.1.3 Missing Transverse Momentum

In the analysis presented in this thesis, $E_\mathrm{T}^\mathrm{miss}$ is reconstructed with the `METMakerTool` [191]. Calibrated electrons and muons passing the baseline selection and calibrated jets before any selection are used as inputs for this tool. Following the Jet/ETmiss group recommendations [192], the `PFlow` $E_\mathrm{T}^\mathrm{miss}$ reconstruction with the track-based measurement of the soft term is used in the analysis to circumvent pile-up induced degradation of calorimeter-based measurements of the soft term.

### 5.1.4 Object Overlap Removal

As most of the reconstruction algorithms are independent of each other, it is possible that one detector signal can be reconstructed as multiple different particles.

To resolve these potential ambiguities overlap removal procedures are performed to only keep one phyiscal particle per detector information. This is done among baseline level leptons and jets, but with isolation and $|d_0/\sigma_{d_0}|$ requirements released for leptons. The applied procedure follows the recommendations described in the Harmonization Document [193]. The implementation is performed similarly to the `AssociationUtils` tool [194] in its `Standard` overlap-removal configuration.

More overlap removal is performed for electrons sharing the same ID-track or those with clusters close to each other. In each case, the most energetic electron is kept. For the ID-track, this is classified using $p_T$, while the cluster $E_T$ is used in the other case.

Lastly, any calorimeter-tagged muon that shares an ID-track with an electron is removed.

## 5.2 Event selection

Two kinds of special phase spaces are used in the following analyses. They both build on an inclusive $W^\pm Z$ phase space, which has been used in other analyses studying that process [24, 26]. This baseline phase space, as well as those built on top of it, are summarized in this section.

### 5.2.1 Inclusive $W^\pm Z$ events selection

The selection of $W^\pm Z$ candidate events is summarized in Table 5.3.

One significant background for studies of $W^\pm Z$ events is those with two $Z$ bosons. They can mimic signal events if one lepton is not identified properly or falls outside of the detector range. This can especially be the case when tight criteria are applied to select the used leptons. To avoid this, a "$ZZ$ veto" is applied by counting leptons directly after overlap removal. This means leptons passing the baseline selection without considering the isolation and $d_0$ significance cuts.

Next, candidate events need to have exactly three leptons in accordance with the expectation of the final state. Additionally, each of these leptons needs to pass the $Z$ selection criteria. To ensure compatibility with the existence of a $Z$ boson, the event is required to have at least one pair of same-flavor opposite charge (SFOC) leptons. With three leptons total, there can be two such pairs. In that case, the pair with $m_{\ell\ell}$ closest to the $Z$ boson mass is chosen as the $Z$ boson candidate. The remaining lepton is then required to pass its respective $W$ lepton selection.

Increasing the transverse momentum threshold of the leading lepton to 25 GeV for the 2015 sample and to 27 GeV for 2016–2018 is done to improve the chances that the triggering lepton is above the trigger efficiency turn-on curve.

The last two steps in the selection are targeted towards ensuring that the reconstructed leptons originate from $W$ and $Z$ bosons. To do that, the SFOC lepton pair has to be close to the literature mass of the $Z$ boson. Specifically $|M_{\ell\ell} - M_Z| < 10\,\text{GeV}$. For the $W$ boson the requirement is that the transverse mass is above 30 GeV, where the transverse mass is defined by the $W$ candidate lepton and the missing transverse momentum as

$$m_{\mathrm{T}}^{W} = \sqrt{2 p_{\mathrm{T}}^{\ell} E_{\mathrm{T}}^{\mathrm{miss}} \left(1 - \cos \Delta\phi\right)}\,. \tag{5.1}$$

In this case, $\Delta\phi$ represents the angle between the constituents in the transverse plane.

Multiple of the variables used in polarization measurements require a reconstruction of the $W$-boson, specifically its center-of-mass frame. With the neutrino escaping the detector, this is not fully possible. To get an approximate result, the literature $W$-boson mass is used as a constraint to calculate the $z$ component of the neutrino momentum [24]. This results in a quadratic equation for the neutrino $p_z$. The smaller solution is chosen in case multiple real ones exist. If none exists, then the real component of the complex solutions is chosen.

**Table 5.3:** Overview of the inclusive event selection. Content from Ref. [161].

| Inclusive event selection | |
|---|---|
| Event cleaning | Reject LAr, Tile and SCT corrupted events and incomplete events |
| Primary vertex | Hard scattering vertex with at least two tracks |
| Trigger 2015 | `HLT_e24_lhmedium_L1EM20VH \|\| HLT_e60_lhmedium \|\| HLT_e120_lhloose`<br>`HLT_mu20_iloose_L1MU15 \|\| HLT_mu50` |
| Trigger 2016–2018 | `HLT_e26_lhtight_nod0_ivarloose \|\| HLT_e60_lhmedium_nod0 \|\| HLT_e140_lhloose_nod0`<br>`HLT_mu26_ivarmedium \|\| HLT_mu50` |
| $ZZ$ veto | Less than 4 baseline leptons |
| $N$ leptons | Exactly three leptons passing the $Z$ lepton selection |
| Leading lepton $p_{\mathrm{T}}$ | $p_{\mathrm{T}}^{\mathrm{lead}} > 25$ GeV (in 2015) or $p_{\mathrm{T}}^{\mathrm{lead}} > 27$ GeV (in 2016-2018) |
| $Z$ leptons | Two same flavor oppositely charged leptons passing $Z$-lepton selection |
| $Z$ invariant mass | $\lvert M_{\ell\ell} - M_Z \rvert < 10$ GeV |
| $W$ lepton | Remaining lepton passes $W$-lepton selection |
| $W$ transverse mass | $m_{\mathrm{T}}^{W} > 30$ GeV |
| $\Delta R$ | $\Delta R(\ell_Z^-, \ell_Z^+) > 0.2$, $\Delta R(\ell_Z, \ell_W) > 0.3$ |

### 5.2.2 Signal Region Definitions

The analyses performed for this thesis comprise two main measurements. The first is a measurement of the fraction of doubly longitudinal events at high energies, discussed in Chapter 9, and the second is a measurement to investigate the radiation amplitude zero effect. Each of these measurements has corresponding signal regions to enhance the signal of interest. These build on top of the inclusive $W^{\pm}Z$ selection described in the previous section in Table 5.3. They comprise of cuts on $p_T^{WZ}$ and $p_T^{Z}$. The $p_T^{WZ}$ cuts aim to reduce higher-order effects associated with jet activity. On the one hand, this allows for easier modeling of the processes with the existing MC generators, but it also serves to restore the RAZ effect, which is otherwise washed out through these higher-order corrections. The latter part also has the advantage of decreasing the size of the TT contribution and thus increasing the fraction of 00 events. Cuts on $p_T^{Z}$ drive the energy of the events to higher values. This results in a further increase of the 00 fraction and pushes toward the kinematic region where the Goldstone boson equivalence theorem holds [195, 196]. The values of the cuts are motivated by theory papers [197, 198]. The two sets of regions are summarized in Table 5.4.

**Table 5.4:** Overview of the selection criteria used to build the fiducial signal regions. Multiple values for a cut in a region category indicate that a region exists for each of the cut values. Content from Ref. [161].

| **Signal regions** | | |
|---|---|---|
| Selection | **Radiation Amplitude Zero** | **00-enriched** |
| Pass inclusive $WZ$ event selection | ✓ | ✓ |
| Transverse momentum of the $WZ$ system ($p_T^{WZ}$) | $< 10, 20, 40, 70\,\mathrm{GeV}$ | $< 70\,\mathrm{GeV}$ |
| Transverse momentum of the $Z$ boson ($p_T^Z$) | - | $(100, 200]\,, > 200\,\mathrm{GeV}$ |

## 5.3 Background Estimation

The inclusive selection discussed earlier mainly aims to reduce the impact of non-WZ processes and reaches values of around 20 % for their contribution to the total number of expected events [26]. While this is not their purpose, the signal region cuts reduce this further to the level of 10 % [161]. Despite the small amount of non-WZ events entering our signal regions, it is still necessary to estimate and include these background contributions in the statistical analysis. While most of the physical processes are treated individually for this purpose, backgrounds are grouped based on their origin into two categories for figures and tables. Specifically, the $VVV$, $ZZ$, $W^{\pm}Zjj$-*EW* and $t\bar{t}V$ processes are grouped into the *Prompt* category. All of these contain three prompt leptons and are estimated directly from MC samples as described in Section 4.3.3. The other backgrounds are combined under the *Non-prompt* label. These enter the signal regions through a misidentification of a detector signal as a lepton. As this is known to be badly modeled by MC generators, they are estimated using a data-driven approach as mentioned in Section 4.3.3.

## 5.4 Truth Selection

The object and event selections defined in the previous selections are deeply tied to the physical ATLAS detector. Values measured within them can not be interpreted without taking the physics of the detector itself into account. With detector simulation times on the order of minutes per event [199, 200], it is infeasible for most theoreticians to compare the predictions of their models to results measured purely at the reconstructed level. Due to this, it is helpful to provide measurements at the truth level where they can be compared more easily. To do this, it is necessary to define truth-level fiducial phase spaces close to the reconstructed level ones. Then, it is possible to transfer the results from the reconstructed phase space to the corresponding fiducial region.

In light of this, two sets of fiducial regions are defined. They are summarized in Table 5.5. The region without the $p_T^{WZ}$ and $p_T^Z$ criteria is the same as the one used for the inclusive diboson polarization measurement [26].

**Table 5.5:** Overview of the selection criteria used to build the analysis signal regions. Multiple values for a cut in a region category indicate that there exists a region for each of the cut values. Content from Ref. [161].

| Fiducial region selection | **RAZ** | **00-enriched** |
|---|---|---|
| Lepton $\lvert\eta\rvert$ | $< 2.5$ | |
| $p_{\mathrm{T}}$ of $\ell_Z$, $p_{\mathrm{T}}$ of $\ell_W$ [GeV] | $> 15$, $> 20$ | |
| $m_Z$ range [GeV] | $\lvert m_Z - m_Z^{\mathrm{PDG}}\rvert < 10$ | |
| $m_{\mathrm{T}}^W$ [GeV] | $> 30$ | |
| $\Delta R(\ell_Z^-, \ell_Z^+)$, $\Delta R(\ell_Z, \ell_W)$ | $> 0.2$, $> 0.3$ | |
| $p_T^{WZ}$ [GeV] | $< 10, 20, 40, 70$ | $< 70$ |
| $p_T^Z$ [GeV] | - | $(100, 200]\,, > 200\,\mathrm{GeV}$ |

The definitions of these regions rely on particle-level objects. Their definitions follow ATLAS recommendations [201]. They and the above event selections are chosen to be close to their reconstructed level counterparts.

Due to that, the leptons used for these phase spaces are prompt ones, meaning that they do not originate from hadron or $\tau$ decays. Specifically, they should be dressed leptons. These are reconstructed by clustering all photons in a cone of $\Delta R < 0.1$ around a bare lepton. Here, a bare lepton is one after QCD final state radiation. This mimics the way leptons, specifically electrons, are reconstructed in the ATLAS detector [202].

Next, the bosons are reconstructed from the leptons using the "resonant shape" algorithm [24]. It is based on the estimator

$$P = \left| \frac{1}{m^2_{(\ell^+,\ell^-)} - (m_Z^{\mathrm{PDG}})^2 + i\Gamma_Z^{\mathrm{PDG}} m_Z^{PDG}} \right|^2 \times \left| \frac{1}{m^2_{(\ell',\nu_{\ell'})} - (m_W^{\mathrm{PDG}})^2 + i\Gamma_W^{\mathrm{PDG}} m_W^{PDG}} \right|^2 . \tag{5.2}$$

Here, $m_Z^{\mathrm{PDG}}$, $m_W^{\mathrm{PDG}}$, $\Gamma_Z^{\mathrm{PDG}}$ and $\Gamma_W^{\mathrm{PDG}}$ are the world averages for the $W$ and $Z$ mass and width according to the Particle Data Group [44]. $P$ is calculated for every possible combination of $\ell - \ell$ and $\nu - \ell$ pairs that conform to the charge and flavor requirements of originating from the $Z$ or $W$ boson, respectively. The assignment chosen is the one with the highest value for $P$. Kinematic cuts are then applied on $p_T$, $\eta$, $m_Z$ and $m_T^W$, matching those in the reconstructed region. Lastly, separation cuts of $\Delta R(\ell_Z^-, \ell_Z^+) > 0.2$ and $\Delta R(\ell_Z, \ell_W) > 0.3$ are applied to avoid ambiguities between overlapping leptons that can occur at the reconstructed level.

# Chapter 6

# Validation

As already mentioned in Chapter 4, the current lack of MC generators that can simulate separate polarizations at full NLO[1] requires the use of an LO + merging setup. A qualitative validation of these samples is done in this chapter. They are performed in three steps. The first, in Section 6.1, is a comparison of the sum of the polarized samples with full NLO samples that only contain the coherent sum of all polarizations. Second, a comparison against NLO calculations done by theorists is performed. This is shown in Section 6.2. A comparison against real measured data is summarized in Section 6.3. Quantitative uncertainties are derived from these comparisons in Chapter 7.

Most of the figures in this section are produced by the author of Ref. [161] and largely taken directly from there.

## 6.1 Validation Against Unpolarized NLO Monte Carlo

In this section, the sum of the polarized signal samples is compared against the full NLO samples from POWHEG+PYTHIA8, MADGRAPH5_aMC@NLO, and SHERPA as described in Chapter 4. If the remaining missing NLO effects, as well as the approximations done in the polarized samples, are small, then the comparison should reveal no large deviations. The comparisons in this section are performed at the truth level in the region with $p_T^Z > 200\,\text{GeV}$ as described in Section 5.4.

Figure 6.1 shows shape-only comparisons between different NLO generators and the sum of polarizations for variables $p_T^{WZ}$, ${r_{21}}$[2], $p_T$ of the leptons[3] and $p_T^W$ after the $p_T^Z > 200\,\text{GeV}$ selection. Figure 6.2 shows more comparisons for variables lepton $\eta$, $\Delta Y(WZ)$, $\cos\theta_W$, and $\cos\theta_\ell(W)$ between NLO and LO simulation in the same 00-enhanced phase-space. More variables can be found in Appendix C

It can be seen that the sum of leading order samples agrees well within the band of NLO samples.

## 6.2 Validation Against NLO QCD Fixed Order Calculations

Another comparison is done against calculations performed by theorists [73, 203]. These are performed at fixed order without parton shower effects. They employ the double-pole approximation as described in Section 2.2.4. They can not be used to simulate individual events or perform GEANT4 simulation studies. However, the authors of Refs. [73, 203] have

[1] Recent developments in SHERPA now make it useable for this purpose [70]. However, these were not available in time to be used in this work.

[2] $r_{21}$ is defined as the ratio of the subleading to the leading boson momentum $r_{21} := \frac{p_T(V2)}{p_T(V1)}$

[3] The histogram is filled three times per event, once per lepton.

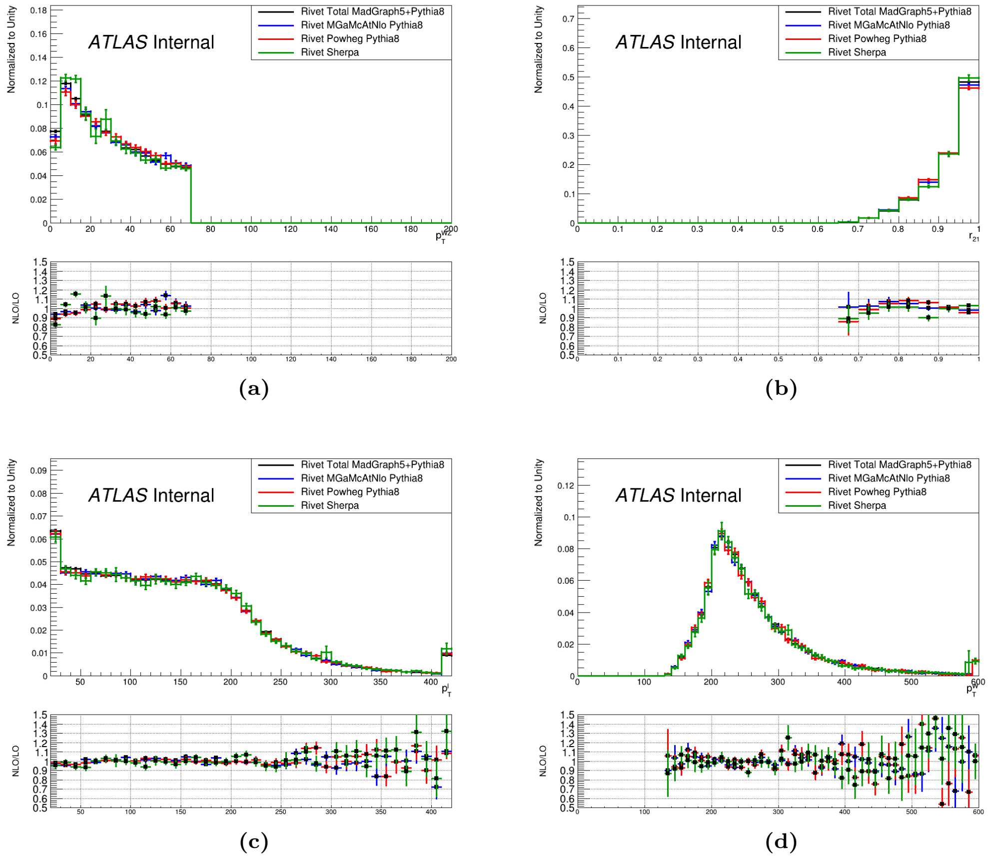



**Figure 6.1:** Comparison between NLO (Powheg+Pythia8, MadGraph5_aMC@NLO and Sherpa) and the sum of LO polarized samples for (a) $p_T^{WZ}$, (b) $r_{21}$, (c) $p_T$ for each of the three leptons and (d) $p_T^W$, in the $W_0Z_0$ enhanced phase space with $p_T^Z > 200$ GeV. From Ref. [161].

applied the same methods as described there to provide distributions for a phase space matching the 00 enhanced one with $p_T^Z > 200$ GeV as described in Section 5.4.

A comparison against them can not give a full (in)validation of the polarized Monte Carlo samples. This is due to the lack of parton shower simulation in the provided distribution. Additionally, only the decay channel $W \to e\nu_e; Z \to \mu\mu$ is included in the theory calculations, while all flavor combinations of electrons and muons are used in the polarized samples. Due to this, the comparison is done only at the shape level. Normalization differences are not investigated. Instead, the overall fractions as predicted by the NLO calculations and the polarized samples are shown in Table 6.1.

**Table 6.1:** "Predictions of diboson polarization fractions from the NLO QCD calculations and from the LO-merged MadGraph5_aMC@NLO samples for the signal region with $p_T^Z > 200$ GeV. It should be pointed out that the fractions from the NLO calculations are performed at the particle level, while the fractions for LO-merged MadGraph5_aMC@NLO samples are performed after parton showering." From Ref. [161].

| | $f_{00}$ | $f_{0T}$ | $f_{T0}$ | $f_{TT}$ |
|---|---|---|---|---|
| NLO calculations | 23.6% | 5.0% | 4.6% | 66.8% |
| LO-merged 0+1 jets samples | 23.8% | 5.8% | 5.2% | 65.2% |

To get at least some handle on the impact of the shower, the ratio $\frac{(NLO/LO_{merged})_{XX}}{(NLO/LO_{merged})_{TT}}$,

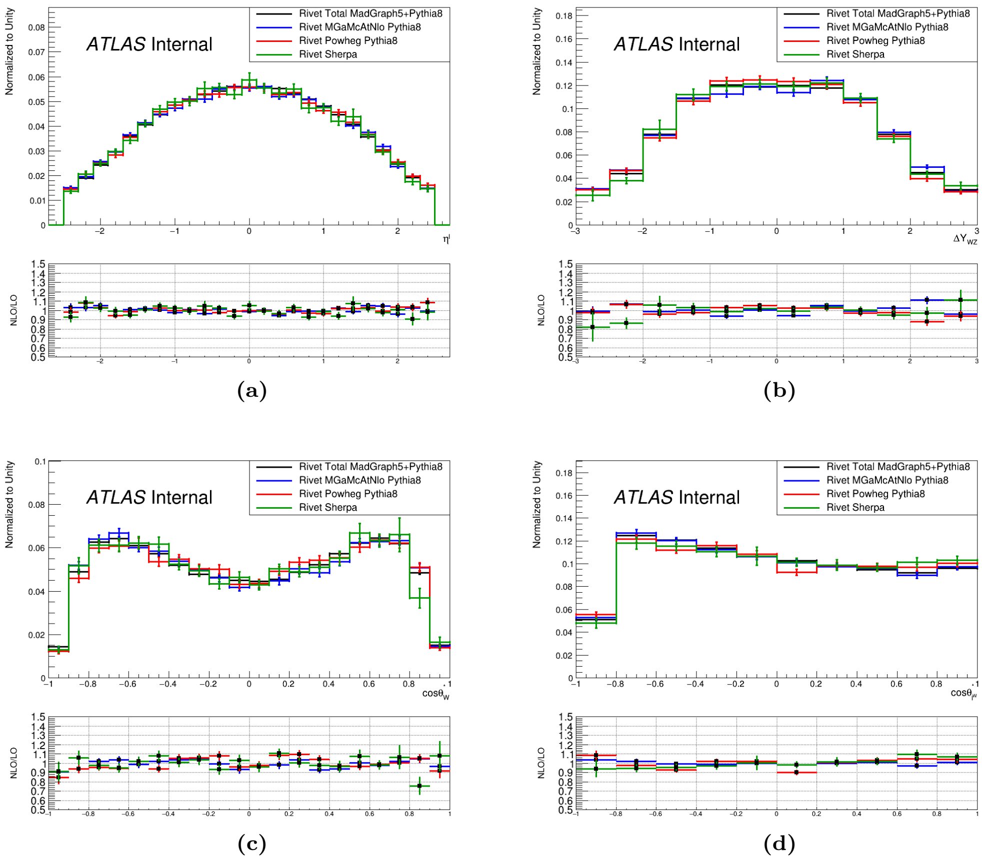


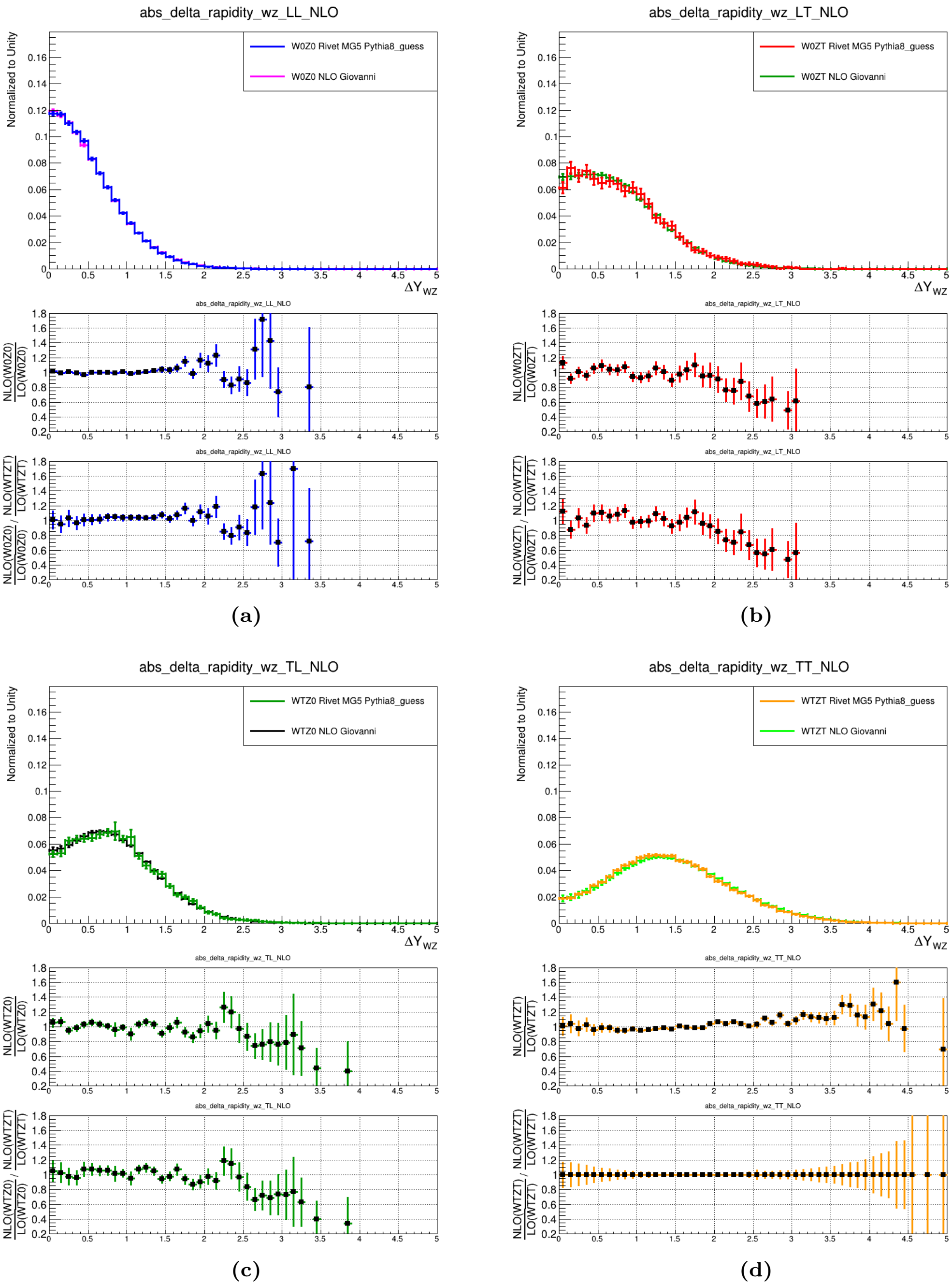

**Figure 6.2:** Comparison between NLO (Powheg+Pythia8, MadGraph5_aMC@NLO and Sherpa) and the sum of LO polarized samples for different kinematic variables: (a) shows the distribution of $\eta$ for all the three leptons, (b) $\Delta Y(WZ)$, (c) $\cos\theta_W$, and (d) $\cos\theta_\ell(W)$, in the $W_0Z_0$ enhanced phase space with $p_T^Z > 200\,\text{GeV}$. From Ref. [161].

where $XX \in (00, 0T, T0, TT)$, is used to get an estimate of its effect. If all polarizations share the same shape differences, as indicated by the ratios $(NLO/LO_{merged})_{XX}$, then it can reasonably be assumed that these are due to the shower. This is especially the case when considering the good agreement observed in the previous section.

Shape comparisons of $\Delta Y(WZ)$ and $\cos\theta_W$ between the NLO calculations and the polarized samples are shown in Figures 6.3 and 6.4, respectively. They include the comparisons for all four polarization states. They show very good agreement between the polarized samples and the theoretical calculations. Only the $\Delta Y(WZ)$ distribution of the 0T polarization state shows significant deviations at high values.

Additional variables are shown in Figures 6.5, 6.6, 6.7, 6.8 and 6.9. These only show the comparisons for the 00 and TT states for the $p_T^{WZ}$, $p_T^Z$, $m_T^{WZ}$, $W$ lepton $p_T^\ell(W)$, and $Z$ lepton $p_T^\ell(Z)$ variables, respectively. The mixed states are not shown here as their overall fraction in this region is small.

In these comparisons, it can be seen that some variables show shape differences from the theoretical calculations. This is most visible for $p_T^{WZ}$. However, the impact of the lack of a parton shower is expected to be most relevant for exactly this variable. Due to this, it makes sense to look at the bottom panel in the plots. While it is also not completely flat for the $p_T^{WZ}$ case, it still shows that the differences are very similar between the different polarization states, indicating that the deviations are mainly due to this modeling difference. The same effects can be seen for most of the other variables in the set.




(a) (b) (c) (d)

**Figure 6.3:** Shape comparison between NLO calculations by theorists and MC samples for (a) LL, (b) LT, (c) TL, (d) TT separately in the LL-enhanced phase space with $p_T^Z > 200\,\text{GeV}$ for $\Delta Y(WZ)$. The middle panel shows the ratio between the two predictions. The bottom panel shows the ratio of ratios: $\frac{(NLO/LO-merged)_{XX}}{(NLO/LO-merged)_{TT}}$ with $XX \in (LL, LT, TL, TT)$. From Ref. [161].

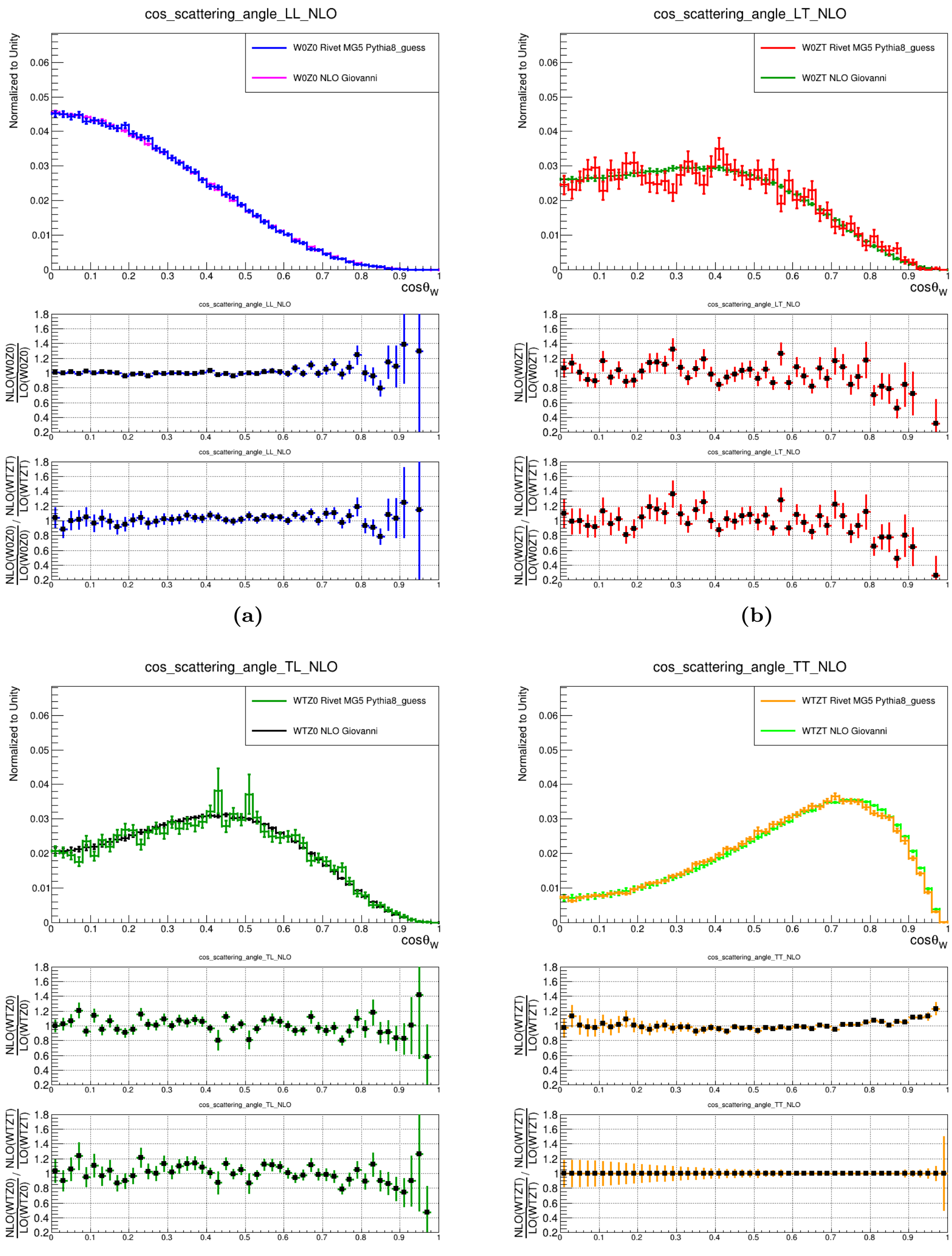
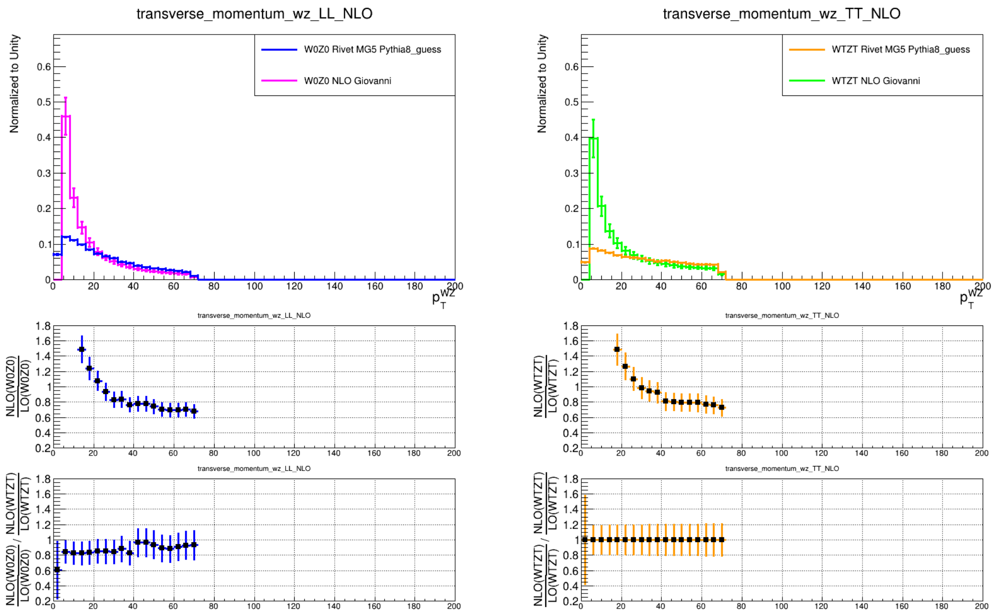
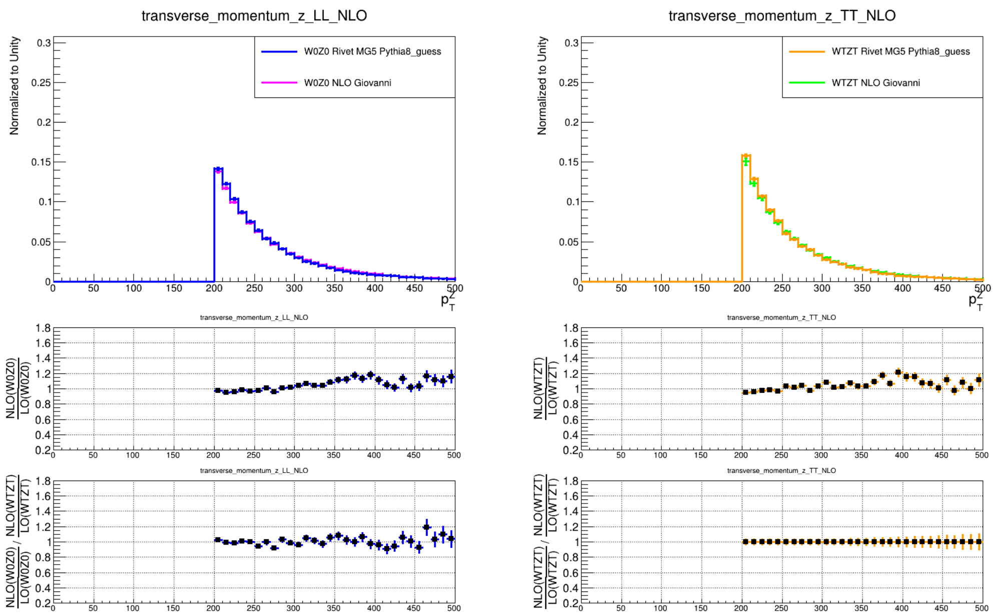



**Figure 6.4:** Shape comparison between NLO calculations by theorists and MC samples for (a) LL, (b) LT, (c) TL, (d) TT separately in the LL-enhanced phase space with $p_T^Z > 200\,\mathrm{GeV}$ for $\cos\theta_W$. The middle panel shows the ratio between the two predictions. The bottom panel shows the ratio of ratios: $\frac{(NLO/LO-merged)_{XX}}{(NLO/LO-merged)_{TT}}$ with $XX \in (LL, LT, TL, TT)$. From Ref. [161].



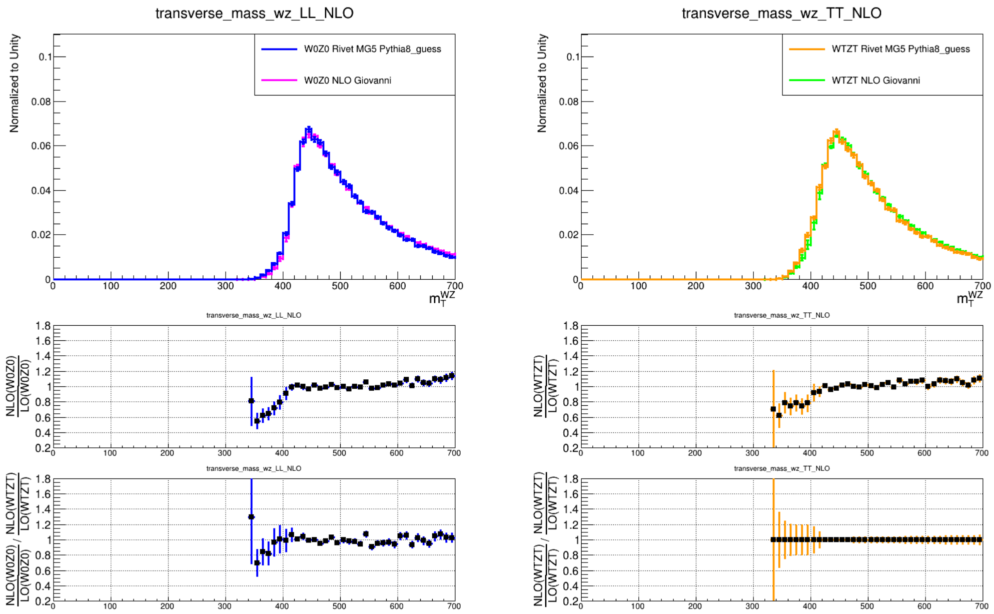

(a) (b)

**Figure 6.5:** Shape comparison between NLO calculations by theorists and MC samples for (a) LL and (b) TT separately in the LL-enhanced phase space with $p_T^Z > 200\,\text{GeV}$ for $p_T^{WZ}$. The middle panel shows the ratio between the two predictions. The bottom panel shows the ratio of ratios: $\frac{(NLO/LO-merged)_{XX}}{(NLO/LO-merged)_{TT}}$ with $XX \in (LL, TT)$. From Ref. [161].



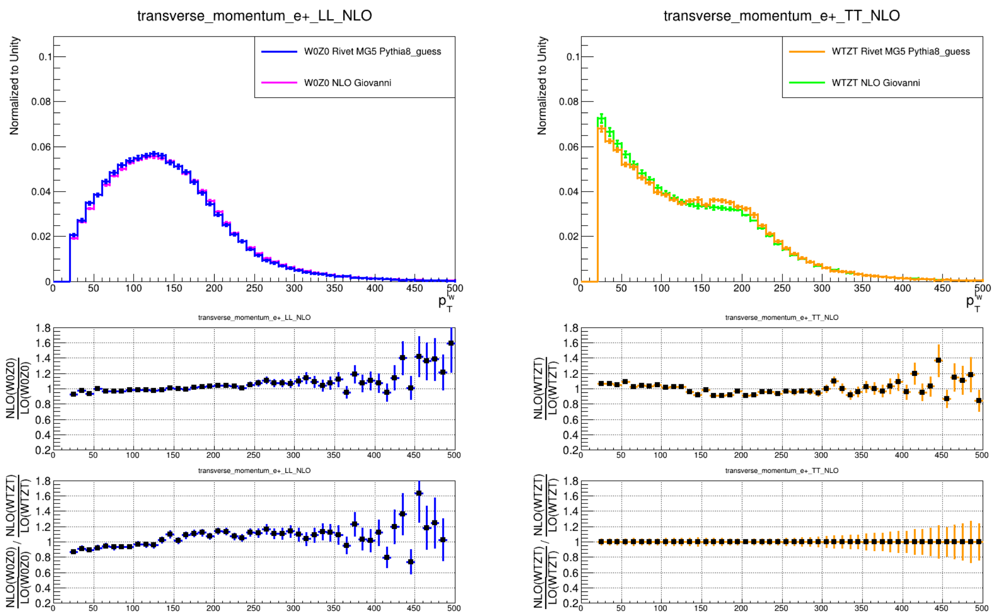

(a) (b)

**Figure 6.6:** Shape comparison between NLO calculations by theorists and MC samples for (a) LL and (b) TT separately in the LL-enhanced phase space with $p_T^Z > 200\,\text{GeV}$ for $p_T^Z$. The middle panel shows the ratio between the two predictions. The bottom panel shows the ratio of ratios: $\frac{(NLO/LO-merged)_{XX}}{(NLO/LO-merged)_{TT}}$ with $XX \in (LL, TT)$. From Ref. [161].



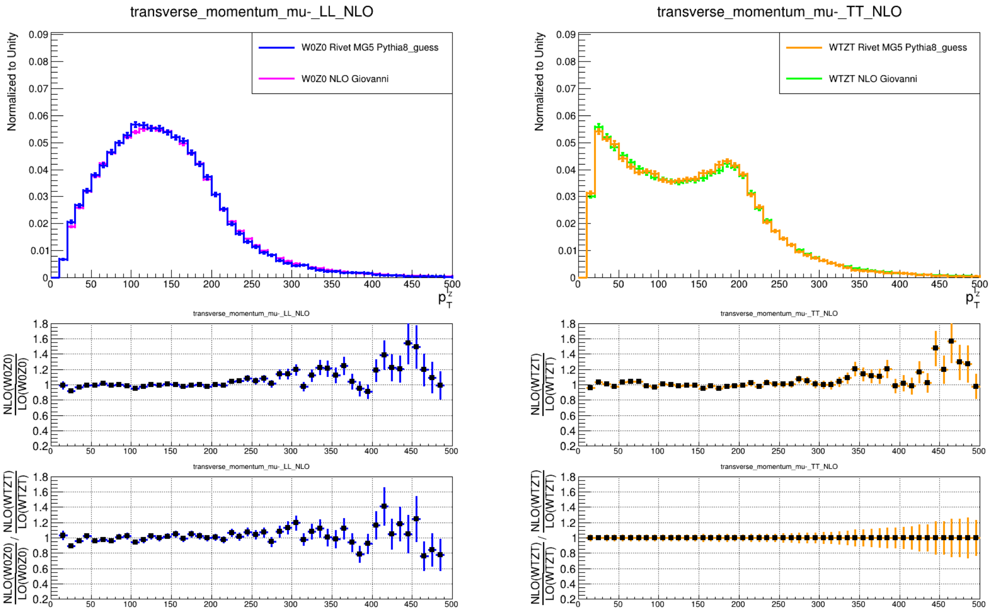

**(a)** **(b)**

**Figure 6.7:** Shape comparison between NLO calculations by theorists and MC samples for (a) LL and (b)TT separately in the LL-enhanced phase space with $p_T^Z > 200$ GeV for $m_{\mathrm{T}}^{WZ}$. The middle panel shows the ratio between the two predictions. The bottom panel shows the ratio of ratios: $\frac{(NLO/LO-merged)_{XX}}{(NLO/LO-merged)_{TT}}$ with $XX \in (LL, TT)$. From Ref. [161].




**(a)** **(b)**

**Figure 6.8:** Shape comparison between NLO calculations by theorists and MC samples for (a) LL and (b) TT separately in the LL-enhanced phase space with $p_T^Z > 200$ GeV for $p_T^{\ell}(W)$. The middle panel shows the ratio between the two predictions. The bottom panel shows the ratio of ratios: $\frac{(NLO/LO-merged)_{XX}}{(NLO/LO-merged)_{TT}}$ with $XX \in (LL, TT)$. From Ref. [161].




**(a)** **(b)**

**Figure 6.9:** Shape comparison between NLO calculations by theorists and MC samples for (a) 0 and (b) TT separately in the LL-enhanced phase space with $p_T^Z > 200\,\text{GeV}$ for $p_T^{\ell^-}(Z)$. The middle panel shows the ratio between the two predictions. The bottom panel shows the ratio of ratios: $\frac{(NLO/LO-merged)_{XX}}{(NLO/LO-merged)_{TT}}$ with $XX \in (LL, TT)$. From Ref. [161].

## 6.3 Validation Against Data

The last validation is a comparison against real data. Care has to be taken with such comparisons to not introduce any bias into the measurement. Due to this, data is usually blinded until an analysis has progressed enough that it is unlikely that biases will still be introduced. As such, comparisons were first performed in the inclusive region as it is very close to the already unblinded region used in Ref. [26].

Since LO calculations are used for different polarization states, a LO $\rightarrow$ NLO QCD scale factor of 1.489 is applied for the overall normalization. This is derived as the ratio of the SHERPA sample to the sum of polarized samples in the truth-inclusive region.

As a last step, reweighting factors correcting for data-MC differences in the $N_{\text{jets}}$ distribution are calculated. For this the simulated $WZ$ events are split into categories of $N_{\text{jets}} \in (0, 1, 2, 3+)$. Then, a factor is applied to each event in this bin in such a way that the final number of events in each of these categories matches between data and MC. The derived values are$(1.03, 0.84, 0.93, 1.42)$ for the categories, respectively. It has to be noted that these factors were derived with an NLO QCD k-factors of 1.525, which was derived the same way as above, except that the reconstructed level inclusive region was used.

Figures 6.10 to 6.13 show the comparisons for the inclusive region. In Figure 6.10 it can be seen that the jet multiplicity reweighting results in good agreement between data and MC. The other figures also show good modeling of most of the variables. The only visible differences can be seen for $\cos\theta_V$ and $\Delta Y(\ell_W Z)$.

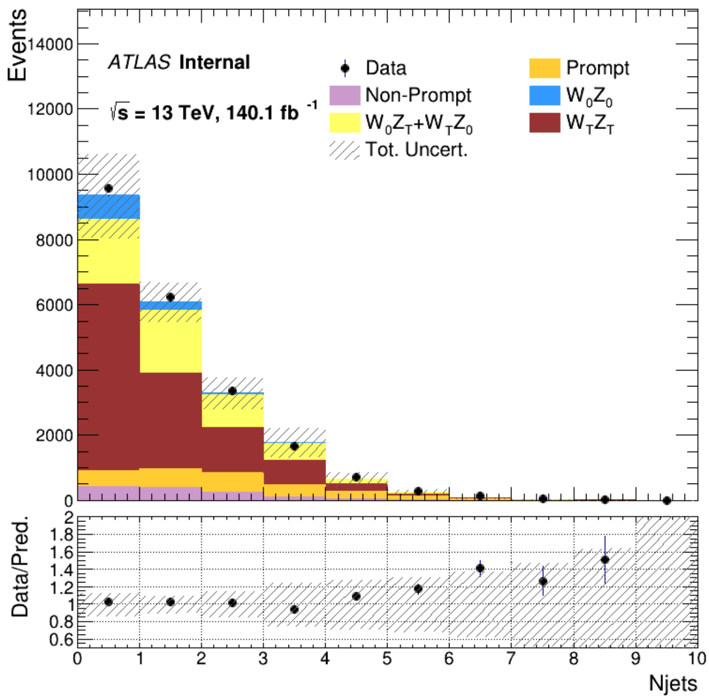



**Figure 6.10:** Comparison between data and prediction for $N_{\text{jets}}$. From the author of Ref. [161].

Later in the analysis, the data in the 00-enhanced regions was unblinded, and a comparison to data was performed there as well. In this case, the jet multiplicity weight derived from the inclusive region is still applied. However, the flat QCD k-factor is changed to 1.376 for the $100 < p_{\text{T}}^{Z} \leq 200\,\text{GeV}$ and 1.324 for the $p_{\text{T}}^{Z} > 200\,\text{GeV}$ region. Additionally, per sample k-factors due to EW effects are applied. The values for each sample and region are listed in Table 6.2 and originate from calculations provided by the authors of Ref. [204]. Figures 6.14 and 6.15 show the comparisons between data and MC for the $p_T^Z > 200\,\text{GeV}$ and $100 < p_T^Z \leq 200\,\text{GeV}$

regions, respectively. Good agreement is observed in the $100 < p_{\mathrm{T}}^{Z} \leq 200\,\mathrm{GeV}$ region. The bin-by-bin deviations in the $p_{\mathrm{T}}^{Z} > 200\,\mathrm{GeV}$ region are larger. However, they are still reasonably compatible within the statistical uncertainty. Due to this, no significant mismodelling is considered to be found.

**Table 6.2:** k-factors due to NLO EW effects. Content from the author of Ref. [161].

| **Polarization** | $100 < p_T^Z \leq 200$ **GeV** | $p_T^Z > 200$ **GeV** |
|---|---|---|
| 00 | 0.958 | 0.878 |
| 0T | 0.982 | 0.964 |
| T0 | 0.980 | 0.992 |
| TT | 0.950 | 0.889 |

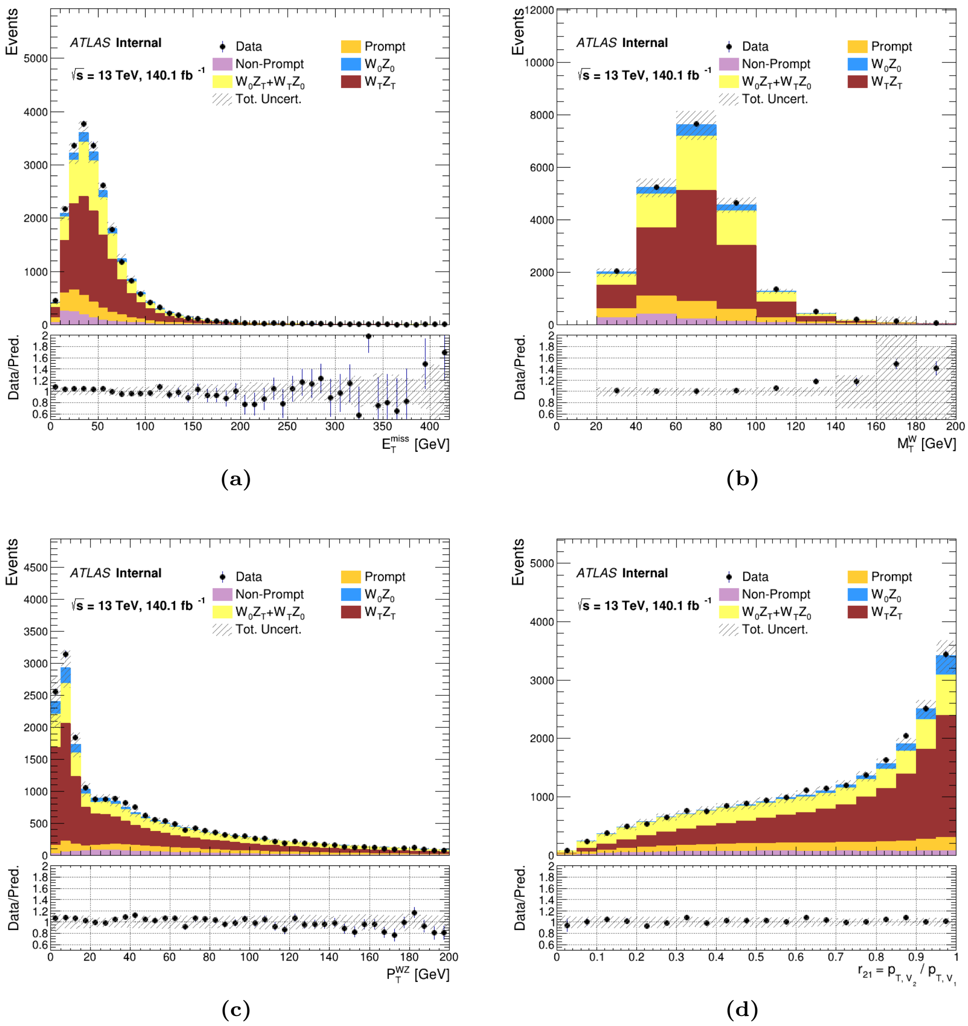



**Figure 6.11:** Comparison between data and prediction for (a) $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) $M_T^W$, (c) $p_T^{WZ}$, and (d) $r_{21}$ in the $WZ$ inclusive region. The non-prompt background is estimated using the matrix method. From the author of Ref. [161].

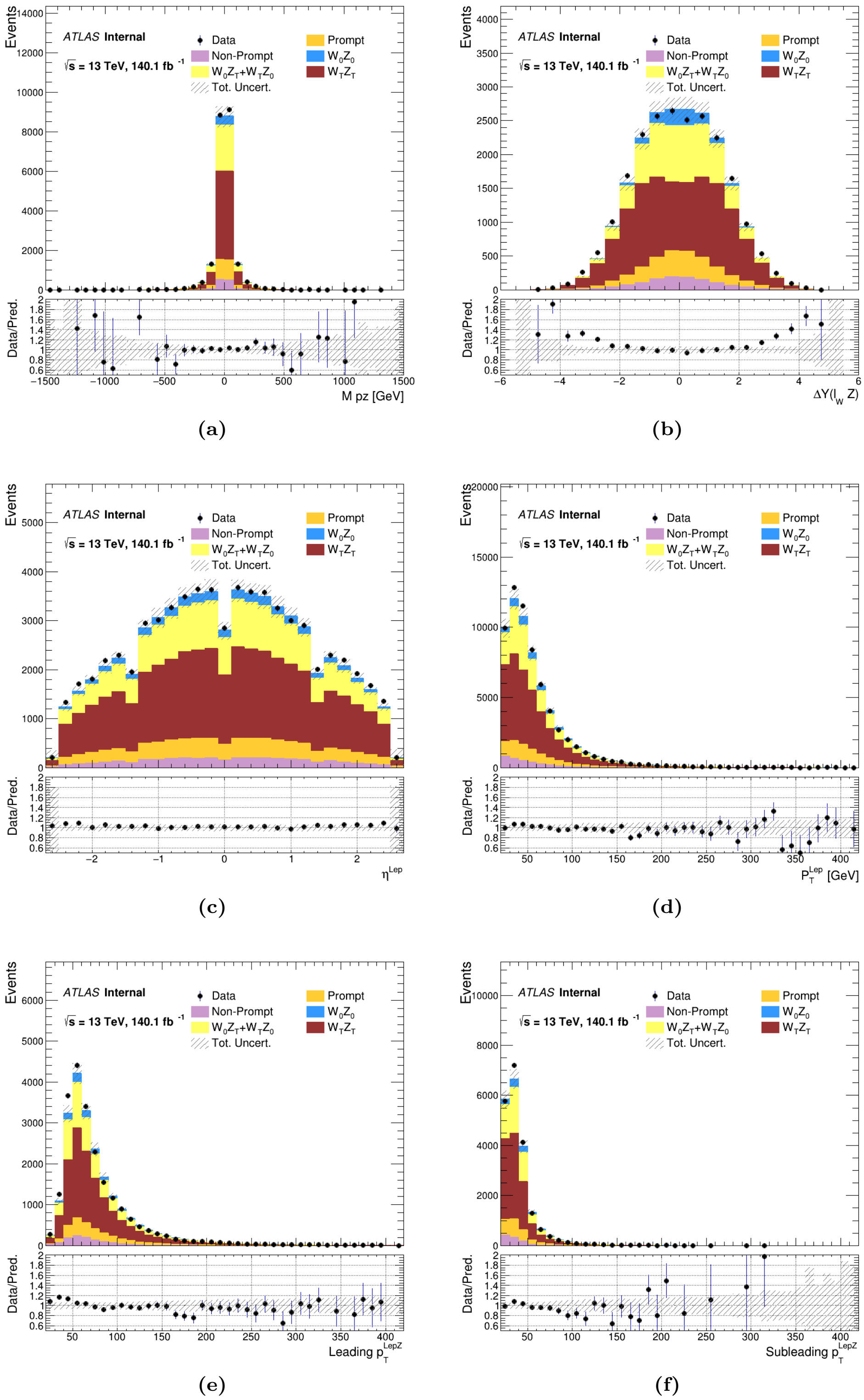
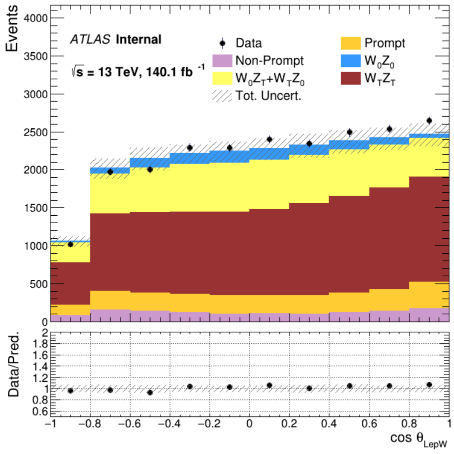
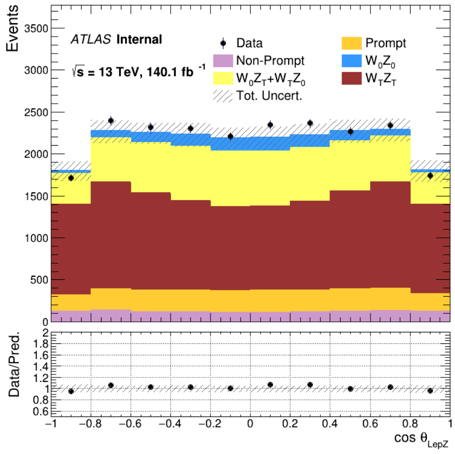


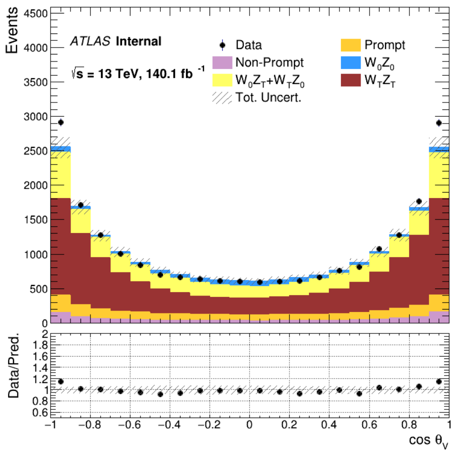

**Figure 6.12:** Comparison between data and prediction for (a) reconstructed neutrino $p_z$, (b) rapidity difference between the $\ell(W)$ and Z boson, (c) lepton $\eta$, (d) lepton $p_T$, (e) the leading $Z$ lepton $p_T$ and (f) the subleading $Z$ lepton $p_T$ in the $WZ$ inclusive region. The non-prompt background is estimated using the matrix method. From the author of Ref. [161].



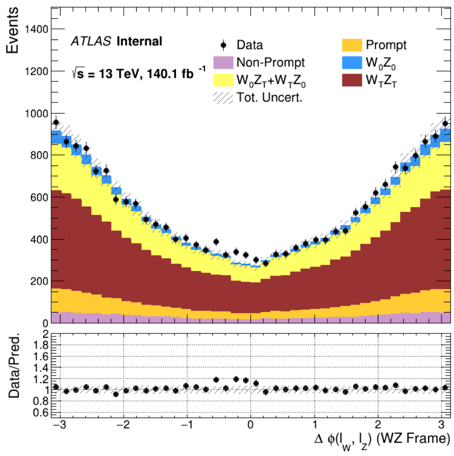

**(a)**




**(b)**

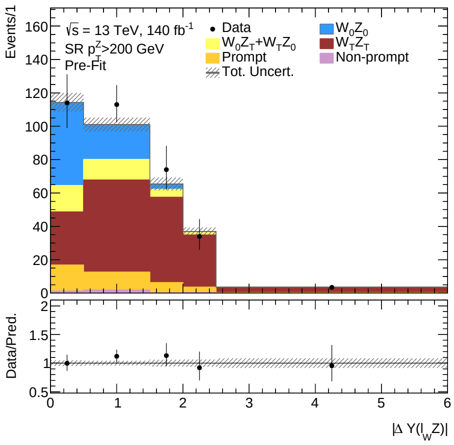



**(c)**

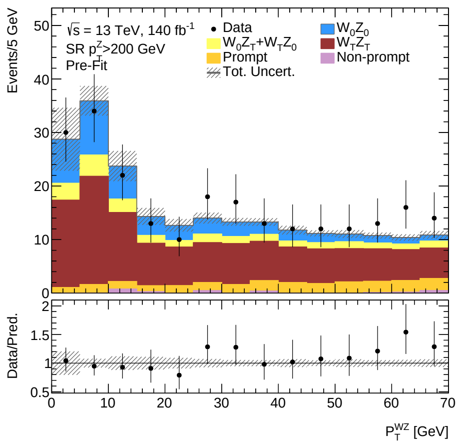



**(d)**

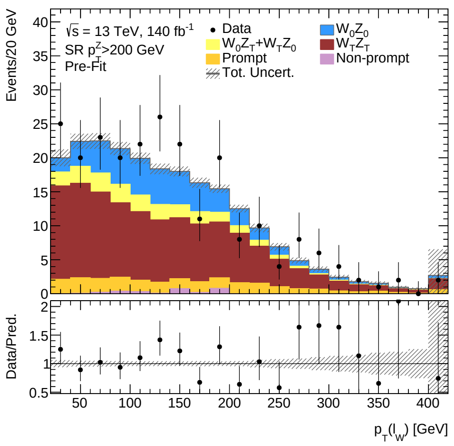
**Figure 6.13:** Comparison between data and prediction for the (a) $\cos\theta^*_\ell(W)$, (b) $\cos\theta^*_\ell(Z)$, (c) $\cos\theta_V$, and (d) $\Delta\phi$ between the $W$ and $Z$ leptons in WZ rest frame in the $WZ$ inclusive region. The non-prompt background is estimated using the matrix method. From the author of Ref. [161].



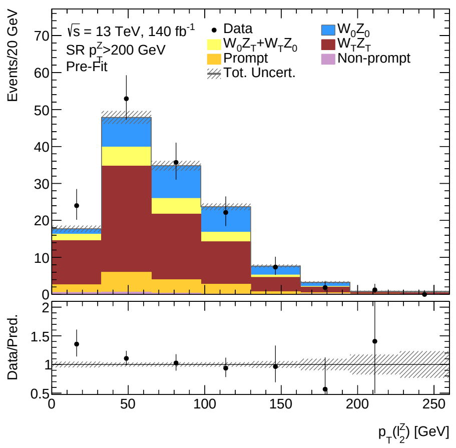

(a)



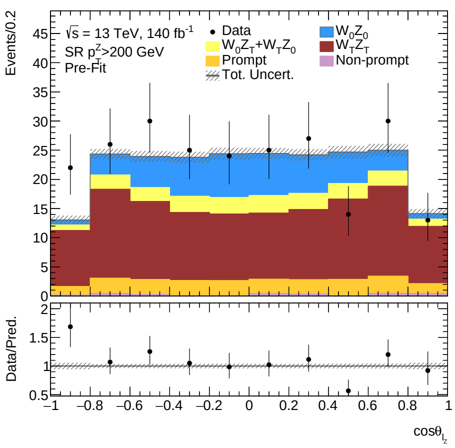

(b)



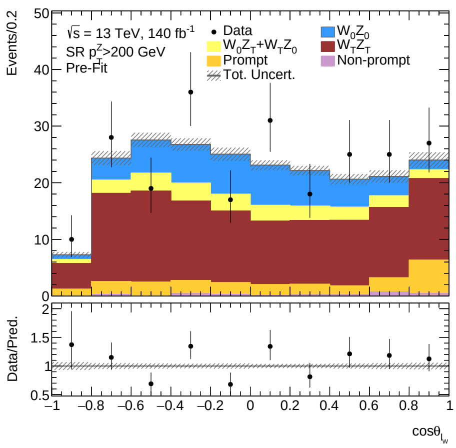

(c)




(d)




(e)




(f)

**Figure 6.14:** Comparison between data and prediction for (a) $|\Delta Y(\ell_W Z)|$, (b) $p_T^{WZ}$, (c) $p_T(\ell_W)$, (d) $p_T(\ell_2^Z)$, (e) $cos\theta_{\ell_Z}$ and (f) $cos\theta_{\ell_W}$ in the 00 enhanced region with $p_T^Z > 200$ GeV.

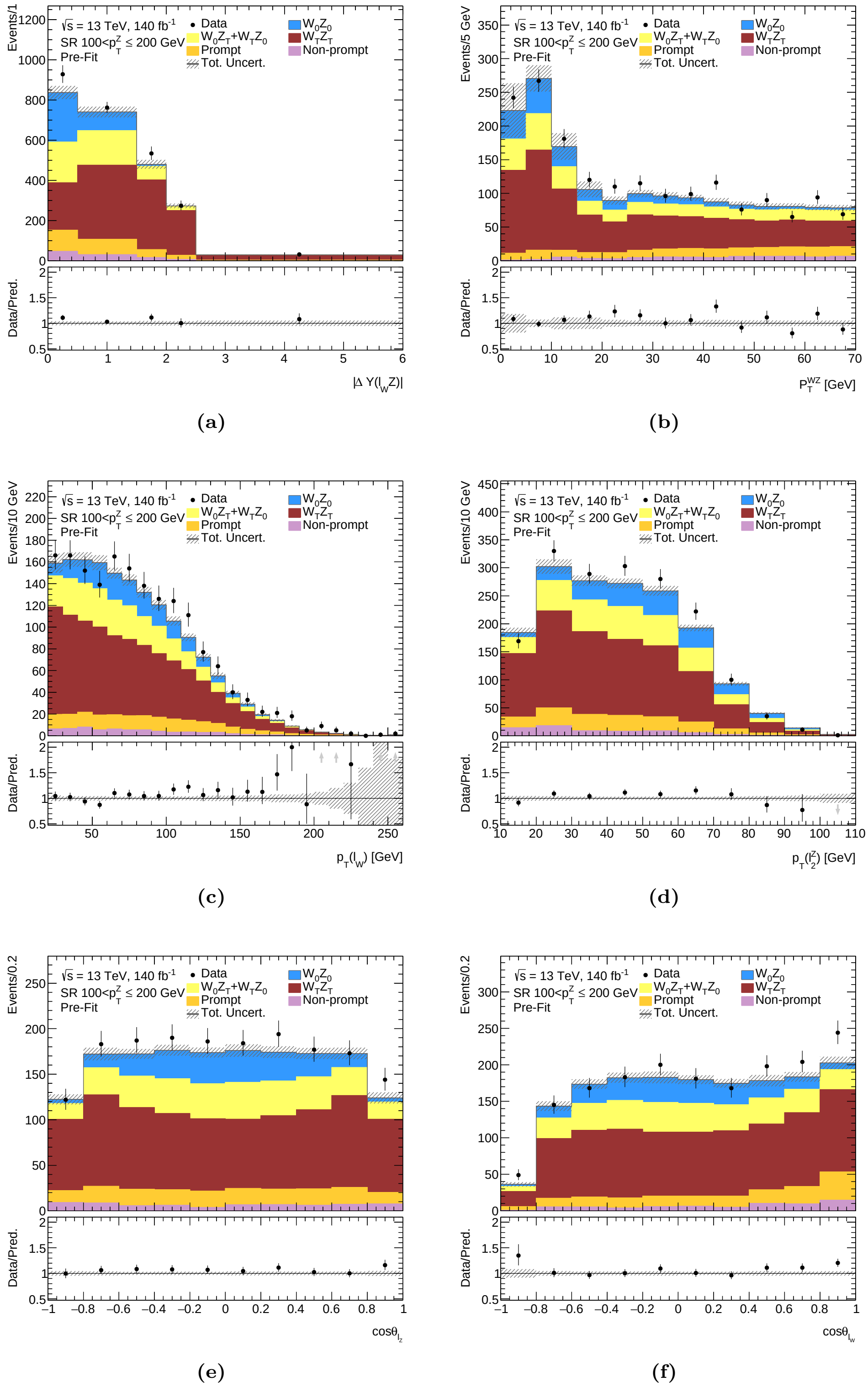



**Figure 6.15:** Comparison between data and prediction for (a) $|\Delta Y(\ell_W Z)|$, (b) $p_T^{WZ}$, (c) $p_T(\ell_W)$, (d) $p_T(\ell_2^Z)$, (e) $cos\theta_{\ell_Z}$ and (f) $cos\theta_{\ell_W}$ in the 00 enhanced region with $100 < p_T^Z \leq 200\,\text{GeV}$. The y-axis labels in these and similar figures of "Events/$X$" follow the ATLAS Style Guide [205–207]. The $X$ indicates the nominal bin size and for a bin of width "Size", the y-axis reports the bin's events multiplied by $X$/Size.

# Chapter 7

# Uncertainties

This chapter discusses all the sources of uncertainties affecting the measurements performed in this thesis. Section 7.1 covers all sources related to the experimental setup. Background-related uncertainties are summarized in Section 7.2. Lastly, uncertainties affecting the modeling of the signal predictions are described in Section 7.3.

## 7.1 Experimental Systematic Uncertainties

The first part of this section describes the uncertainty of the total luminosity measurement. The rest describes uncertainties related to the physics objects used in this thesis. As these are independent of any specific process or analysis, they are derived centrally by dedicated combined performance groups within ATLAS. They compile sets of variations covering all potential uncertainties. These typically include uncertainties in the object identification, reconstruction, isolation, and kinematics measurements. These variations are then applied to all MC samples and have their effects propagated to the final yields and distributions. Depending on the analysis, full or reduced sets of uncertainties can be used. The following sections summarize which sets and variations were used in this work.

### 7.1.1 Luminosity

The uncertainty in the combined 2015–2018 integrated luminosity is 0.83 % [156]. It is obtained using the LUCID-2 detector [158]. This uncertainty is applied to all processes that have their normalization estimated from MC. These are all the processes considered, except for the fake background, which uses a data-driven method. The uncertainty is applied in both the fraction measurement discussed in Chapter 9 as well as the RAZ measurement in Chapter 11.

### 7.1.2 Muons

Uncertainties for muons are derived by the Muon Combined Performance group. Information about muon reconstruction and how specific uncertainties are estimated can be found in Ref. [112]. Table 7.1 summarized the specific muon-related variations used in this thesis. These names later appear again in plots and descriptions related to the final measurements.

### 7.1.3 Electrons

A description of electron and photon reconstruction and associated uncertainties can be found in Refs. [110, 111]. As this analysis is very much statistically dominated and does not independently measure in different bins of $\eta$, the reduced '1NP_v1' model is used [161, 183]. It combines all physical effects in quadrature and considers them fully correlated across the whole $\eta$ range of the detector. The final list of electron and photon-related uncertainties is shown in Table 7.2.

**Table 7.1:** Summary of uncertainties related to muon measurements.

| Nuisance parameter name | Description |
|---|---|
| MUON_EFF_ISO_STAT | Muon statistical isolation efficiency uncertainty |
| MUON_EFF_ISO_SYS | Muon systematic isolation efficiency uncertainty |
| MUON_EFF_RECO_STAT | Muon reconstruction and identification efficiency statistical uncertainty for $p_T > 15\,\mathrm{GeV}$ |
| MUON_EFF_RECO_STAT_LOWPT | Muon reconstruction and identification efficiency statistical uncertainty for $p_T < 15\,\mathrm{GeV}$ |
| MUON_EFF_RECO_SYS | Muon reconstruction and identification efficiency systematic uncertainty for $p_T > 15\,\mathrm{GeV}$ |
| MUON_EFF_RECO_SYS_LOWPT | Muon reconstruction and identification efficiency systematic uncertainty for $p_T < 15\,\mathrm{GeV}$ |
| MUON_EFF_TTVA_STAT | Muon statistical track-to-vertex association efficiency uncertainty |
| MUON_EFF_TTVA_SYS | Muon systematic track-to-vertex association efficiency uncertainty |
| MUON_ID | Muon momentum resolution uncertainty: Inner detector |
| MUON_MS | Muon momentum resolution uncertainty: Muon system |
| MUON_SAGITTA_RHO | "Variations in the scale of the momentum (charge-dependent), based on combination of correction on combined (Z scale) and recombination of the corrections" From Ref. [208] |
| MUON_SAGITTA_RESBIAS | "Variations in the scale of the momentum (charge-dependent), based on the residual charge-dependent bias after correction" From Ref. [208] |
| MUON_SCALE | Muon momentum scale uncertainty |

**Table 7.2:** Summary of uncertainties related to electron measurements.

| NP name | Description |
|---|---|
| EG_RESOLUTION_ALL | Electron and photon energy resolution uncertainty |
| EG_SCALE_ALL | Electron and photon energy scale uncertainty |
| EG_SCALE_AF2 | Electron and photon Atlas Fast sim. energy scale uncertainty |
| EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR | Electron isolation efficiency uncertainty |
| EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR | Electron identification efficiency uncertainty |
| EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR | Electron reconstruction efficiency uncertainty |

### 7.1.4 Jets

Systematic variations on the jets are applied through the dedicated ATLAS "JetUncertaintiesTool". Separate uncertainties are provided related to jet energy scale and jet energy resolution [209]. For the first, a 'GlobalReduction' model is used. It combines the nearly 100 nuisance parameters into a set of just 20 and is intended for analyses not used within combinations. Similarly, a 'SimpleJER' model is used for the jet energy resolution uncertainties. This combines 34 original nuisance parameters down to 8. Additionally, smearing is only applied to MC events. These sets are appropriate as the analyses do not directly consider or apply cuts on jets. The list of resulting nuisance parameters is given in Table 7.3

### 7.1.5 Missing Transverse Momentum

As described in Chapters 3 and 5, missing transverse momentum is reconstructed using a hard and a soft term. Uncertainties on the hard term come directly from the uncertainties of the relevant object described in the previous sections. Additional uncertainties are then applied for the soft term [191, 210]. The resulting uncertainties are shown in Table 7.4.

More details on the reconstruction and evaluation of the uncertainties for $E_T^{\mathrm{miss}}$ are described in Ref. [117].

### 7.1.6 Flavor tagging

Flavor tagging is performed using the neural-network-based DL1r tagger. More information about flavor tagging and corresponding uncertainty determination can be found in Refs. [211, 212]. The total number of sources of uncertainties in the tagging is reduced from $\mathcal{O}(100)$ to $\mathcal{O}(10)$ through the use of eigenvector decomposition [213, 214].

The final set of flavor tagging uncertainties used in the analysis is given in Table 7.5.

**Table 7.3:** Summary of uncertainties related to jet measurements.

| NP name | Description |
|---|---|
| JET_EtaIntercalibration_NonClosure_2018data | Jet $\eta$ inter-calibration non-closure in 2018 data |
| JET_EtaIntercalibration_NonClosure_highE | Jet $\eta$ inter-calibration non-closure at high energyy |
| JET_EtaIntercalibration_NonClosure_negEta | Jet $\eta$ inter-calibration non-closure for negative $\eta$ |
| JET_EtaIntercalibration_NonClosure_posEta | Jet $\eta$ inter-calibration non-closure for positive $\eta$ |
| JET_EtaIntercalibration_TotalStat | Jet $\eta$ inter-calibration statistical uncertainty |
| JET_EtaIntercalibration_Modelling | Jet $\eta$ inter-calibration modelling uncertainty |
| JET_Grouped_NP_[1-3] | Jet energy scale uncertainty 1-3 |
| JET_Effective_NP_[1-7] | Jet energy scale uncertainty from the in situ analysis grouped |
| JET_Effective_NP_8restTerm | Remaining uncertainty from the from the in situ analysis |
| JET_BJES_Response | Jet energy scale uncertainty on b-jets |
| JET_fJvtEfficiency | Forward-Jet-Vertex-Tagging (fJVT) efficiency uncertainty |
| JET_JvtEfficiency | Jet-Vertex-Tagging (JVT) efficiency uncertainty |
| JET_Flavor_Composition | Jet flavour composition mismodelling uncertainty |
| JET_Flavor_Response | Jet flavour response mismodelling uncertainty |
| JET_JER_DataVsMC_MC16 | Jet energy resolution based on the difference between data and MC |
| JET_JER_EffectiveNP_[1-6] | Jet energy resolution (JER) uncertainties 1-6 |
| JET_JER_EffectiveNP_7restTerm | Remaining jet energy resolution (JER) uncertainty |
| JET_Pileup_OffsetMu | Pileup dependent uncertainty for the offsets of the mean number of interactions per bunch crossing |
| JET_Pileup_OffsetNPV | Pileup dependent uncertainty for the offsets of the number of primary vertices |
| JET_Pileup_PtTerm | Pileup dependent uncertainty of the $p_T$ term |
| JET_Pileup_RhoTopology | Pileup dependent uncertainty of the $\rho$ topology |
| JET_PunchThrough_MC16 | Uncertainty related to very high energy jets which escape the detector |
| JET_SingleParticle_HighPt | Jet energy scale uncertainty from behaviour of high $p_T$ jets |

**Table 7.4:** Summary of uncertainties related to met measurements.

| NP name | Description |
|---|---|
| MET_SoftTrk_ResoPara | MET track-based soft term resolution uncertainty along the $p_T$-Hard axis |
| MET_SoftTrk_ResoPerp | MET track-based soft term resolution uncertainty perpendicular to the $p_T$-Hard axis |
| MET_SoftTrk_Scale | MET track-based soft term offset along the $p_T$-Hard axis |

### 7.1.7 Pileup

Systematic uncertainties of the pileup reweighting are evaluated through the use of a dedicated “PileupReweightingTool” [215, 216], which is also used to perform the nominal pileup reweighting. They are realized by changing the value of a ‘DataScaleFactor’ used to improve agreement between data and MC. The corresponding uncertainty in the analysis is named `PRW_DATASF`.

## 7.2 Uncertainties on Background Contributions

Uncertainties on the non-prompt background are determined as part of the data-driven approach, giving the nominal estimation of the background. This is described in Ref. [161].

In the final analysis, they are represented by two variations:

`FakeBkg_Stat`: Uncertainty due to the limited statistics available for the background estimation

`FakeBkg_MetCut`: Uncertainty describing the impact of the choice of $E_T^{\text{miss}}$ cut on the estimation. Here the cut value is varied between the nominal 25/30 GeV and 50 GeV for the up and 20 GeV for the down variation.

**Table 7.5:** Summary of uncertainties related to flavor tagging measurements.

| NP name | Description |
|---|---|
| FT_EFF_Eigen_B_[0-8] | flavor-tagging efficiency uncertainties for b-jets 0-8 |
| FT_EFF_Eigen_C_[0-3] | flavor-tagging efficiency uncertainties for c-jets 0-3 |
| FT_EFF_Eigen_Light_[0-3] | flavor-tagging efficiency uncertainties for light jets 0-3 |
| FT_EFF_extrapolation | Uncertainty from extrapolating the tagging weights to high $p_T$ for b-jets |
| FT_EFF_extrapolation_from_charm | Uncertainty from extrapolating the mistagging to high $p_T$ for c-jets |

For the prompt backgrounds, flat normalization uncertainties are applied. Their overall size is estimated either from scale and pdf uncertainties or through differences between predicted and observed cross-sections in dedicated analyses. The chosen values for each process alongside sources motivating the values are given in Table 7.6.

**Table 7.6:** Background normalization uncertainties applied in the analysis.

| Sample | Unc. | Source |
|---|---|---|
| $ZZ$ | 15 % | [217–219] |
| $t\bar{t}+V$ | 13 % | [220, 221] |
| $WZjj$-EW | 20 % | [222] |
| $VVV$ | 20 % | [223] |

## 7.3 Signal Modeling Uncertainties

The last category considered is modeling uncertainties. These are process-specific uncertainties that arise due to an incomplete understanding of the processes of interest at the theoretical level. Some sources of such uncertainties are common to many ATLAS analyses. For these recommendations by the ATLAS Physics Modelling Group (PMG) exist [224]. A dedicated tool (PMGSystematicsTool) is provided for calculating these uncertainties [225]. These recommendations are provided and followed for the PDF 7.3.1, strong coupling 7.3.2, scale 7.3.3 and shower 7.3.4 uncertainties. Additional uncertainties are derived from the remaining NLO QCD and EW effects. For these, no central recommendations exist. The estimation is described in Sections 7.3.5 and 7.3.6, respectively.

The derivation and implementation of most of these uncertainties were performed by the author of Ref. [161], while the author of this thesis did the derivation of the nominal NLO EW uncertainties.

### 7.3.1 Parton Distribution Function (PDF)

As mentioned in Chapter 4, the PDF set used for the simulation of the signal samples is the NNPDF 3.0 PDF set. The central PDF used is the one with the number 260000. Following the PDF4LHC recommendations [226, 227] a 68 % confidence level for the PDF uncertainty has been obtained.

The PDF uncertainty is computed using 100 replicas of the NNPDF 3.0 PDF set. For each bin $i$, this is done according to

$$\Delta^{\mathrm{PDF}} N_i = \sqrt{\frac{1}{(n_{\mathrm{rep}}-1)} \sum_{k=1}^{n_{\mathrm{rep}}} \left( N_{i,\mathrm{NNPDF}}^{(k)} - \langle N_{i,\mathrm{NNPDF}} \rangle \right)^2} \,, \tag{7.1}$$

where $k \in \ 0, 1, 2.......n_{\mathrm{rep}}$ and $n_{\mathrm{rep}} = 100$, while $k = 0$ corresponds to the central PDF. Here, the mean of the PDFs is computed as

$$\langle N_{i,\mathrm{NNPDF}} \rangle = \frac{1}{n_{\mathrm{rep}}} \sum_{k=1}^{n_{\mathrm{rep}}} N_{i,\mathrm{NNPDF}}^{(k)} \,. \tag{7.2}$$

Other PDF sets such as PDF4LHC15_nlo_30_pdfas [226], CT14nlo [228], MMHT2014nlo68clas118 [229] were also investigated. However, they are not used further in this work as their inclusive cross-sections were found to be within 68% confidence interval determined from the NNPDF replicas [161]. The uncertainty is included in the fit under the name `theo_PDFUncertainty`.

### 7.3.2 Strong Coupling $\alpha_S$

Following the recommendation in Ref. [226], the PDF and $\alpha_\mathrm{s}$ uncertainties are calculated and combined, where variations of $\alpha_\mathrm{s} = 0.117$ and $\alpha_\mathrm{s} = 0.119$ are used along with the nominal value of $\alpha_\mathrm{s} = 0.118$. For the calculations, refer to Equation 7.3.

The uncertainty in the $i$-th bin due to the variation of the $\alpha_\mathrm{s}$, evaluated at the mass of the $Z$ boson, used is calculated as follows:

$$\Delta^{\alpha_\mathrm{s}} N_i = \left| \frac{N_i(\alpha_\mathrm{s} = 0.119) - N_i(\alpha_\mathrm{s} = 0.117)}{2} \right| . \tag{7.3}$$

These values come from the value and uncertainty of $\alpha_S(m_Z)$ according to the Particle Data Group [44]. In the fraction measurement, this uncertainty is included in the PDF uncertainty described in the previous section by adding it in quadrature to the result shown there. In the RAZ measurement, it is included as a standalone uncertainty.

### 7.3.3 Scale

As described in Chapter 4, the matrix element of a process does not only depend on the kinematics of the incoming and outgoing particles but also the (unphysical) renormalization $\mu_R$ and refactorization $\mu_F$ scales. These scales would not have an impact if calculations could be performed to all orders of perturbation theory. However, this is (currently) not possible. Due to this, a procedure is employed to estimate the impact of the choice of these scales on the measurement. This is done by varying each scale independently to values of 0.5, 1.0, and 2.0 of its nominal value. Out of all the nine possible combinations, the variations with $\{\mu_R, \mu_F\} \in (\{0.5, 2.0\}, \{2.0, 0.5\})$ are discarded to avoid large logarithms coming from the large scale differences [224].

The uncertainty on the expected number of events $N_i$ in bin $i$ is then calculated by

$$\Delta^{\mathrm{scale}} N_i = \max_j |N_i(\mu_{R,i}, \mu_{F,i}) - N_i(\mu_{R,0}, \mu_{F,0})| , \tag{7.4}$$

where $j$ and 0 represent the $j$-th scale variation and central scale, respectively. In the fraction measurement, this uncertainty is listed as `theo_ScaleUncertainty`.

### 7.3.4 Parton Shower

Technical issues with Herwig prevented the generation of samples that use an alternative generator for showering of the signal samples. Instead, uncertainties are derived by using eigen-variations of the A14 tune. These cover variations of ISR/FSR/MPI and color reconnection parameters in an independent way.

The recommendation is to estimate the effect of the five up and five down variations. Then, variations with small effects on the observables of interest should be neglected. Finally, the remaining ones are to be summed in quadrature.

An inspection of the individual variations showed that only one had a significant impact on the observables used in the fraction and RAZ measurements. As such, the maximum deviation between its up and down component towards the nominal is used as the full showering uncertainty. The name of this uncertainty in the fit is `theo_ShwrEigenTuneUnc`.

### 7.3.5 NLO QCD

While good agreement between the signal samples and NLO samples and calculations were found in Sections 6.1 and 6.2, some differences were still observed. To account for this, an uncertainty is added.

The NLO QCD corrections are applied based on the strategy used by the $ZZ$ polarization analysis [230, 231] as well as the inclusive $W^{\pm}Z$ polarization analysis [26]. A binwise k-factor is applied to each polarized sample individually by reweighting distributions of a variable having a significant impact on the results. For the RAZ regions, $\Delta Y(\ell_W Z)$ is used for reweighting while unfolding the $\Delta Y(\ell_W Z)$ distribution, and $\Delta Y(WZ)$ is used for reweighting while unfolding the $\Delta Y(WZ)$ distribution. For the 00 signal regions, the variables used for reweighting are $\Delta Y(\ell_W Z)$ and $\cos\theta_W$.

Two types of corrections are used:

1. NLO QCD (without interference): The basic QCD reweighting factor for each polarized sample is derived as

$$k_{\text{factor}} = \frac{MoCaNLO^{\text{Parton}}_{\text{Pol}}}{MGLO^{\text{Particle}}_{\text{Pol}}} \times \frac{\textit{Sherpa-NLO}^{\text{Particle}}_{\text{Incl}}}{MoCaNLO^{\text{Parton}}_{\text{Incl}}}, \tag{7.5}$$

where 'Pol' indicates the sample for which the uncertainty is derived and 'Incl' refers to samples without a split of polarization, so the coherent sum of all polarizations, including interference effects. The *MoCaNLO* input distributions are provided by the authors of Refs. [204, 232, 233], while *MGLO* refers to the signal samples described in Section 4.3.1 and *Sherpa-NLO* to the Sherpa 2.2.2 sample mentioned in Section 4.3.2. The first term reweights the respective polarized MadGraph sample to NLO. This also removes parton shower effects as these are not present in the *MoCaNLO* distributions. To account for this, the second term is used to add them back.

After the NLO QCD k-factor is applied, the interference effect

$$Interf^{\text{Reco}}_{\text{NLO}} = \left(1 - \frac{\sum_{\text{Pol}} MoCaNLO^{\text{Parton}}_{\text{Pol}}}{MoCaNLO^{\text{Parton}}_{\text{Incl}}}\right) \times \textit{Sherpa-NLO}^{\text{Reco}}_{\text{Incl}}, \tag{7.6}$$

is subtracted from it via

$$RewtFactor(\text{without interference}) = \frac{\textit{Sherpa-NLO}^{\text{Reco}}_{\text{Incl}} - Interf^{\text{Reco}}_{\text{NLO}}}{\sum_{\text{Pol}} \left(MGLO^{\text{Reco}}_{\text{Pol}} \times k_{\text{factor,Pol}}\right)}. \tag{7.7}$$

The same variables are used for calculating the NLO QCD and interference weights.

The total QCD correction is then the product of these two factors

$$kf_{\sigma_{\text{QCD}}} = k_{\text{factor}} \times RewtFactor(\text{without interference}), \tag{7.8}$$

$$d_{\sigma_{\text{NLO QCD}}} = d_{\sigma_{\text{MG-LO}}} \times kf_{\sigma_{\text{QCD}}}, \tag{7.9}$$

where the $d$ indicates distributions.

The uncertainty is calculated by taking the difference between this reweighted distribution and the nominal distribution, which has a constant NLO QCD kfactor applied.

$$d\Delta_{\text{QCD}} = d_{\sigma_{\text{NLO QCD}}} - Nominal \tag{7.10}$$

2. NLO QCD (with interference): Here, Equation 7.5 is still applied, but instead of Equation 7.7, Equation 7.11 is used:

$$RewtFactor(\text{with interference}) = \frac{\textit{Sherpa-NLO}^{\text{Reco}}_{\text{Incl}}}{\sum_{\text{Pol}} \left(MGLO^{\text{Reco}}_{\text{Pol}} \times k_{\text{factor,Pol}}\right)}. \tag{7.11}$$

$$kf_{\sigma_{\text{QCD(w.Inter)}}} = k_{\text{factor}} \times RewtFactor(\text{with interference}) \tag{7.12}$$

$$d_{\sigma_{\text{NLO QCD(w.Inter)}}} = d_{\sigma_{\text{MG-LO}}} \times kf_{\sigma_{\text{QCD(w.Inter)}}} \tag{7.13}$$

The uncertainty is applied to the nominal distribution as the difference between the NLO QCD correction with and without interference.

$$d\Delta_{\text{QCD(Inter Only)}} = d_{\sigma_{\text{NLO QCD(w.Inter)}}} - d_{\sigma_{\text{NLO QCD}}} \tag{7.14}$$

The uncertainties are referred to, respectively, as `theo_NLOQCDCorr_Unc_rewt` and `theo_NLOQCDCorrOnlyInter_Unc_rewt`. Each of them exists twice in the fit with the used variable appended to this name. The same is true for the NLO EW uncertainties discussed in the next section.

### 7.3.6 NLO EW

Next-to-Leading order corrections from the electroweak theory are expected to become sizeable at higher energies. Thus, it is of interest to incorporate these in the measurement. Just as with NLO QCD correction, there is currently no MC generator that can generate polarized samples with them included. Due to this, a similar approach is taken as in the previous section. It uses calculations at both NLO QCD and NLO EW. Such corrections have become available for $W^{\pm}Z$ production with both bosons in a definite polarization state [204, 232, 233]. These calculations have been provided by the authors of the above papers for all signal regions of the RAZ and 00 fraction measurements.

Figures 7.1 and 7.2 show comparisons of $\Delta Y(\ell_W Z)$ and $\Delta Y(WZ)$ distributions between the theoretical calculations for NLO QCD+EW and NLO QCD in the RAZ regions with $p_T^{WZ} < 70\,\mathrm{GeV}$. In this region, the corrections are relatively small and mostly flat in the shape of the variables of interest.

The situation is different in the regions used to extract the 00 polarization fraction. Examples for the region with $p_T^Z > 200\,\mathrm{GeV}$ are shown in Figures 7.3 and 7.4. Clear impacts on the shapes of $\Delta Y(\ell_W Z)$ and the average $p_T$ of the three leptons can be seen. Both variables show sensitivity to the polarization of the bosons and are thus of interest in the fraction measurement. To account for these effects, a reweighting approach similar to the one described in Section 7.3.5 is taken to estimate uncertainties from these values. Specifically, binwise corrections are calculated for two variables, $\Delta Y(\ell_W Z)$ and the average $p_T$ of the three leptons, via

$$kf_{\sigma_{\mathrm{QCD+EW/QCD}}} = d_{\sigma_{\mathrm{QCD+EW}}}/d_{\sigma_{\mathrm{QCD}}} \tag{7.15}$$

$$d_{\sigma_{\mathrm{NLO\ QCD+EW}}} = d_{\sigma_{\mathrm{MG\text{-}LO}}} \times kf_{\sigma_{\mathrm{QCD}}} \times kf_{\sigma_{\mathrm{QCD+EW/QCD}}} \tag{7.16}$$

$$d\Delta_{\mathrm{EW\ Only}} = d_{\sigma_{\mathrm{NLO\ QCD+EW}}} - d_{\sigma_{\mathrm{NLO\ QCD}}} \tag{7.17}$$

where $kf_{\sigma_{\mathrm{QCD}}}$ and $d_{\sigma_{\mathrm{NLO\ QCD}}}$ are the QCD k-factors and corrections from Equations 7.8 and 7.9, respectively. In the RAZ measurement, the corrections are calculated directly for the studied variables $\Delta Y(\ell_W Z)$ and $\Delta Y(WZ)$. This uncertainty can be identified in the fit as `theo_NLOOnlyEWCorr_Unc_rewt`.

#### Combining NLO QCD+EW Uncertainties

When using both NLO QCD and NLO EW corrections, there is always a question regarding how they should be combined. Generally, there are two approaches [230, 231]. An additive one

$$d_{\sigma_{\mathrm{NLO}_+}} = d_{\sigma_{\mathrm{LO}}} \times (1 + \delta_{\mathrm{QCD}} + \delta_{\mathrm{EW}}), \tag{7.18}$$

and a multiplicative one

$$d_{\sigma_{\mathrm{NLO}_\times}} = d_{\sigma_{\mathrm{LO}}} \times (1 + \delta_{\mathrm{QCD}}) \times (1 + \delta_{\mathrm{EW}}), \tag{7.19}$$

with

$$\delta_{\mathrm{QCD}} = \frac{d\Delta\sigma_{\mathrm{QCD}}}{d\sigma_{\mathrm{LO}}}, \delta_{\mathrm{EW}} = \frac{d\Delta\sigma_{\mathrm{EW}}}{d\sigma_{\mathrm{LO}}}. \tag{7.20}$$

The nominal approach used in Equation 7.16 is equivalent to the additive one as

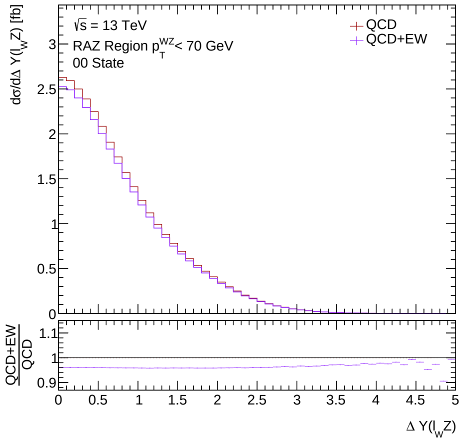
$$\begin{aligned} d_{\sigma_{\mathrm{NLO\ QCD+EW}}} &= d_{\sigma_{\mathrm{MG\text{-}LO}}} \times kf_{\sigma_{\mathrm{QCD}}} \times kf_{\sigma_{\mathrm{QCD+EW/QCD}}} && (7.21)\\ &= d_{\sigma_{\mathrm{LO}}} \times \frac{\cancel{d_{\sigma_{\mathrm{LO}}} + d\Delta_{\mathrm{QCD}}}}{d_{\sigma_{\mathrm{LO}}}} \times \frac{d_{\sigma_{\mathrm{LO}}} + d\Delta_{\mathrm{QCD}} + d\Delta_{\mathrm{EW}}}{\cancel{d_{\sigma_{\mathrm{LO}}} + d\Delta_{\mathrm{QCD}}}} && (7.22)\\ &= d_{\sigma_{\mathrm{LO}}} \times \frac{d_{\sigma_{\mathrm{LO}}} + d\Delta_{\mathrm{QCD}} + d\Delta_{\mathrm{EW}}}{d_{\sigma_{\mathrm{LO}}}} && (7.23)\\ &= d_{\sigma_{\mathrm{LO}}} \times (1 + \delta_{\mathrm{QCD}} + \delta_{\mathrm{EW}}). && (7.24) \end{aligned}$$

Here, MG-LO is the calculation from 0+1 jet MG samples, and LO is the actual theoretical leading order calculation without any jets.

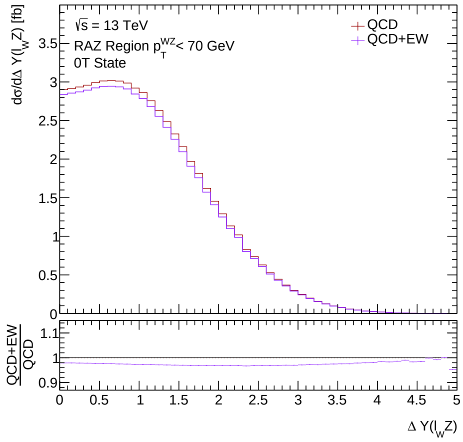
The multiplicative combination corrections are also calculated. Here, $d\Delta\sigma_{\mathrm{EW}}$ is also determined through distributions generated by a program provided by the authors of Ref. [204]. They are added as an uncertainty using the difference to the nominal additive approach

$$d\Delta_{\text{Multi Only}} = d_{\sigma_{\mathrm{NLO}_\times}} - d_{\sigma_{\mathrm{NLO}_+}}, \tag{7.25}$$

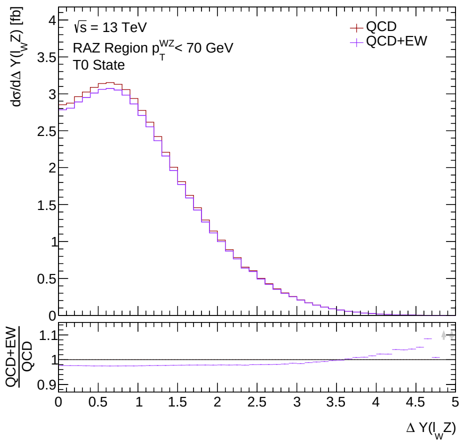
using the name `theo_NLOOnlyMultiCorr_Unc_rewt`.



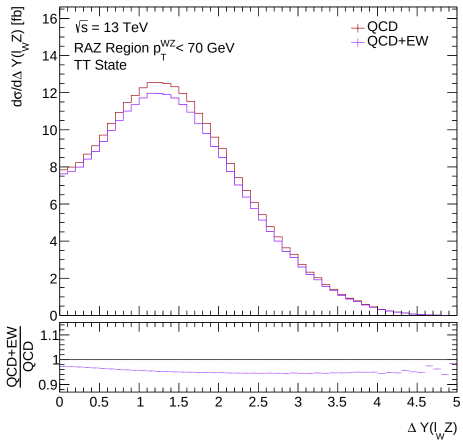

**(a)**



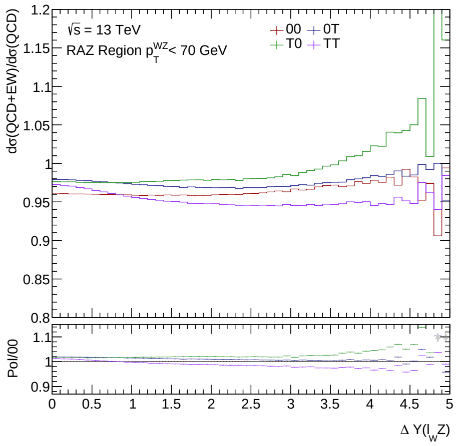

**(b)**




**(c)**




**(d)**




**(e)**

**Figure 7.1:** Comparison between NLO QCD + EW and NLO QCD of $\Delta Y(\ell_W Z)$ for the four polarizations in the RAZ region $p_T^{WZ} < 70\,\text{GeV}$. Shown are the comparisons for (a) 00, (b) 0T, (c) T0 and (d) TT. Additionally shown is (e) a comparison of NLO(QCD+EW)/NLO(QCD) for the four polarizations.

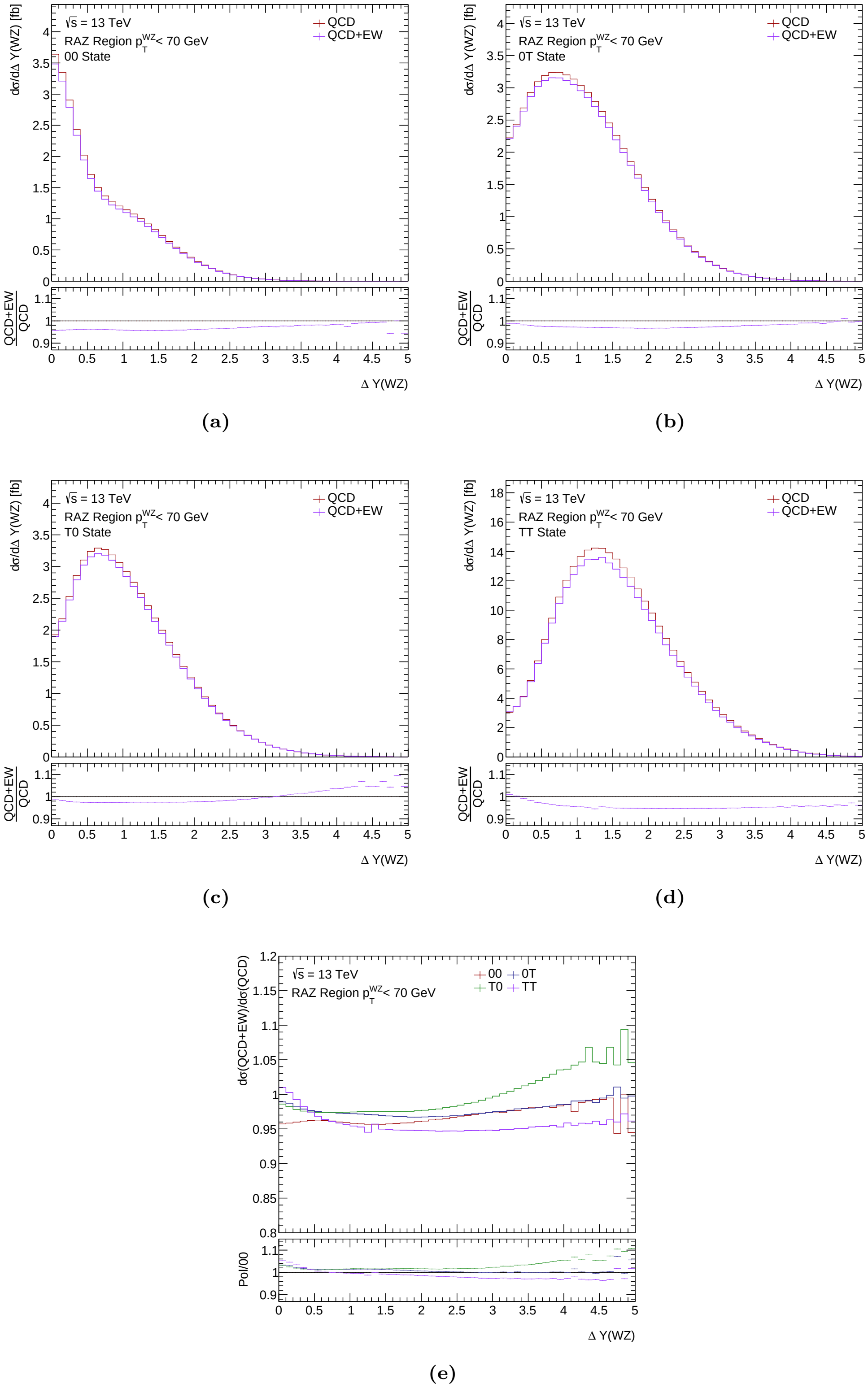



**Figure 7.2:** Comparison between NLO QCD + EW and NLO QCD of $\Delta Y(WZ)$ for the four polarizations in the RAZ region $p_T^{WZ} < 70\,\text{GeV}$. Shown are the comparisons for (a) 00, (b) 0T, (c) T0 and (d) TT. Additionally shown is (e) a comparison of NLO(QCD+EW)/NLO(QCD) for the four polarizations.

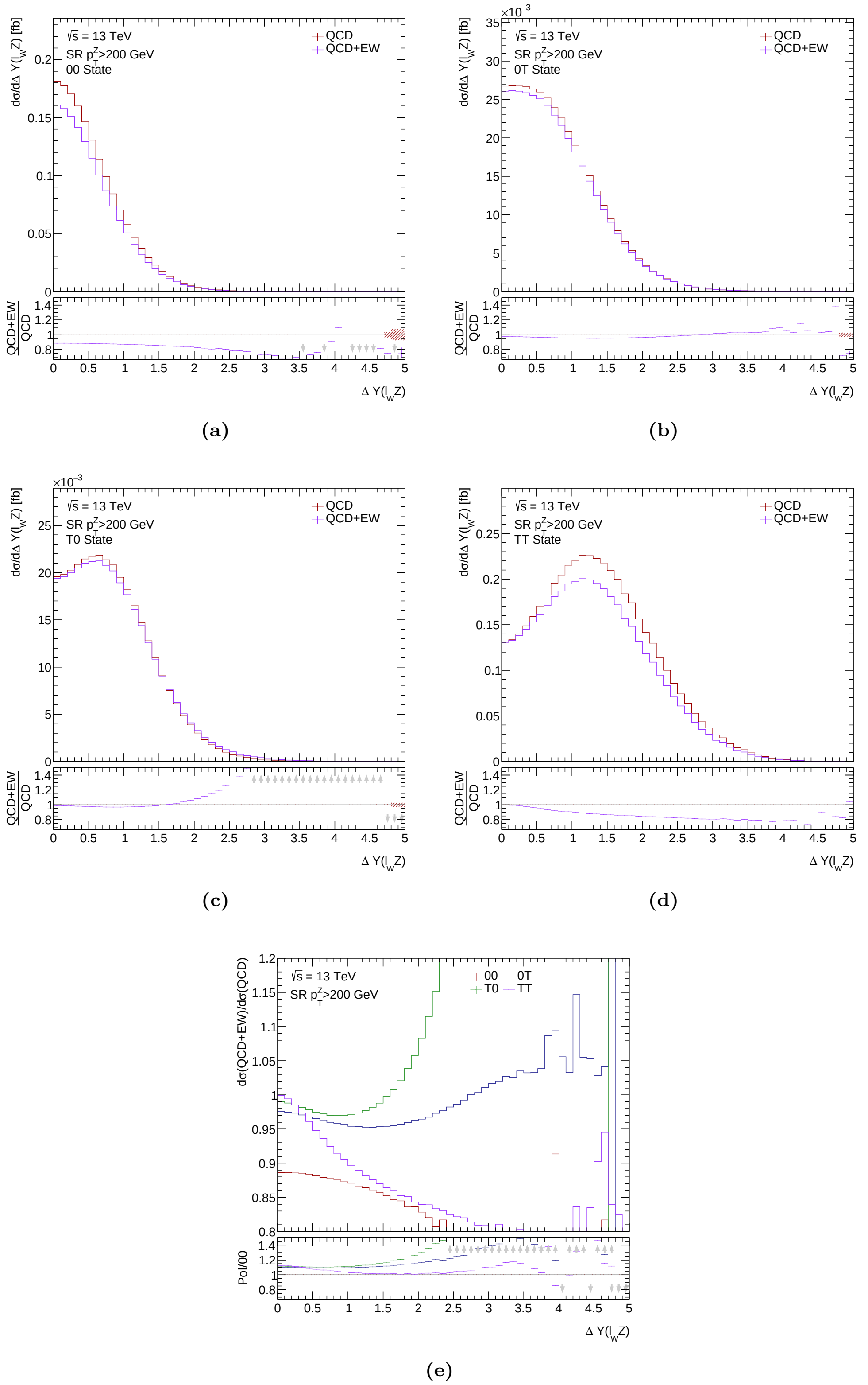
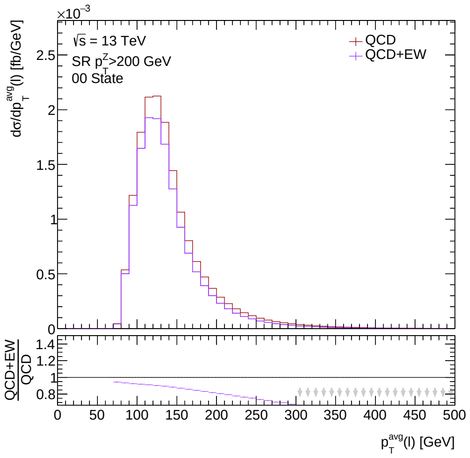
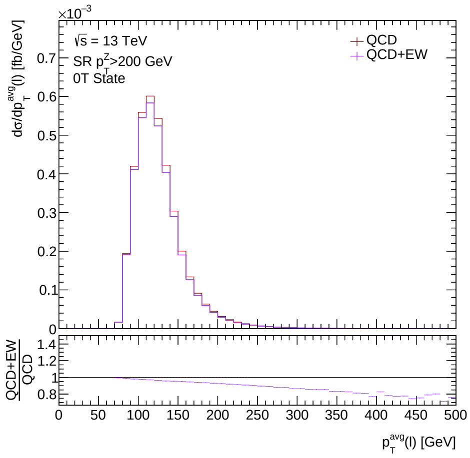


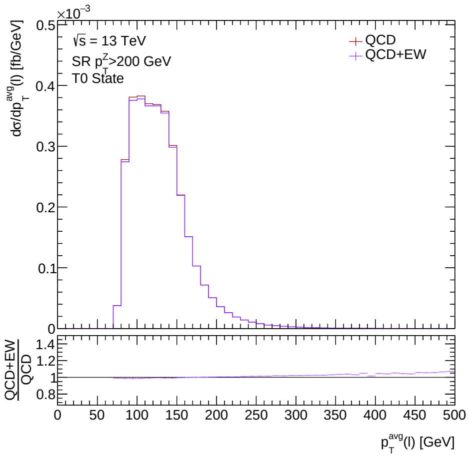

**Figure 7.3:** Comparison between NLO QCD + EW and NLO QCD of $\Delta Y(\ell_W Z)$ for the four polarizations in the 00 enhanced region with $p_T^Z > 200\,\text{GeV}$. Shown are the comparisons for (a) 00, (b) 0T, (c) T0 and (d) TT. Additionally shown is (e) a comparison of NLO(QCD+EW)/NLO(QCD) for the four polarizations.



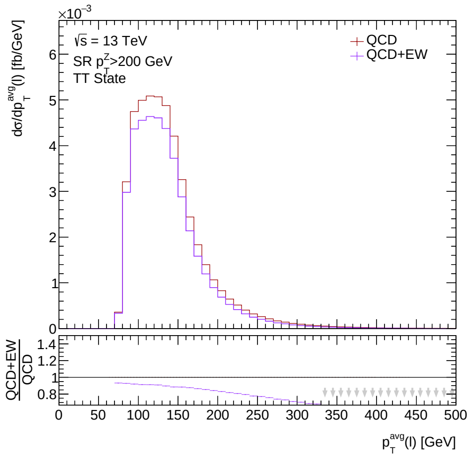

**(a)**



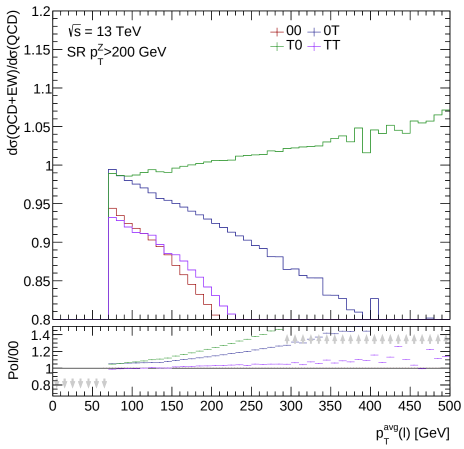

**(b)**




**(c)**




**(d)**




**(e)**

**Figure 7.4:** Comparison between NLO QCD + EW and NLO QCD of the average lepton $p_T$ for the four polarizations in the 00 enhanced region with $p_T^Z > 200\,\text{GeV}$. Shown are the comparisons for (a) 00, (b) 0T, (c) T0 and (d) TT. Additionally shown is (e) a comparison of NLO(QCD+EW)/NLO(QCD) for the four polarizations.

# Chapter 8

# Multivariate Analysis

To be able to measure smaller and smaller cross-sections, such as those of processes at the bottom right of Figure 8.1, it is necessary to utilize as much of the available information about the process of interest as possible. Doing so with the magnitude of available kinematic variables and their correlations has proven to be increasingly difficult for humans to do manually. As such, analysis strategies have slowly shifted from trying to optimize signal region cuts and fit variable selection by hand to utilizing multi-variate approaches instead. When following such a strategy, the manual tuning is reduced to a minimum and mainly done

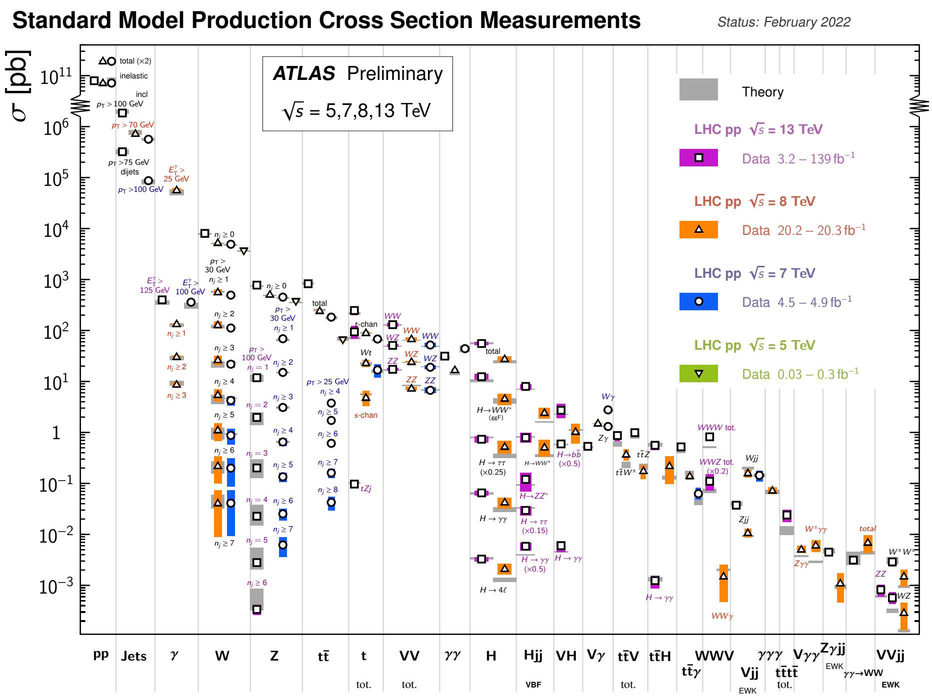



**Figure 8.1:** "Summary of several Standard Model total and fiducial production cross-section measurements .... The measurements are corrected for branching fractions, compared to the corresponding theoretical expectations. In some cases, the fiducial selection is different between measurements in the same final state for different center-of-mass energies $\sqrt{s}$, resulting in lower cross-section values at higher $\sqrt{s}$." From Ref. [234].

to select the specific physics target of interest as well as work around known mismodelings in detector or MC simulation.

While the general process of $W^{\pm}Z$ production is not at such a low cross-section, the focus on the 00 polarization, as well as the additional kinematic cuts, still necessitates the use of such an approach. The one used in this work makes use of boosted decision trees as implemented in ROOT's [235] *Toolkit for Multivariate Data Analysis* (TMVA). The following section gives a brief introduction to this topic. A more thorough description, as well as more technical implementation details, can be found in the TMVA user guide [236].

This chapter first covers the basics of boosted decision trees in Section 8.1. Afterward, the concrete setup and optimization procedure used for the classifier in the signal region with $p_T^Z > 200\,\text{GeV}$ is described in Sections 8.2 and 8.3, respectively. Finally, Section 8.4 presents the similarities and differences for the setup used in the region with $200\,\text{GeV} > p_T^Z > 100\,\text{GeV}$.

## 8.1 Boosted Decision Trees

One of the simplest multi-variate classifiers imaginable is a decision tree. Such a tree classifies events by taking individual yes/no decisions until a leaf node is reached. The classification of that event is then determined by which leaf node it arrived at. Each decision is only based on a singular variable. This effectively splits the parameter space into multiple hypercubes, each identified as signal- or background-like. Figure 8.2 shows an example of such a decision tree.

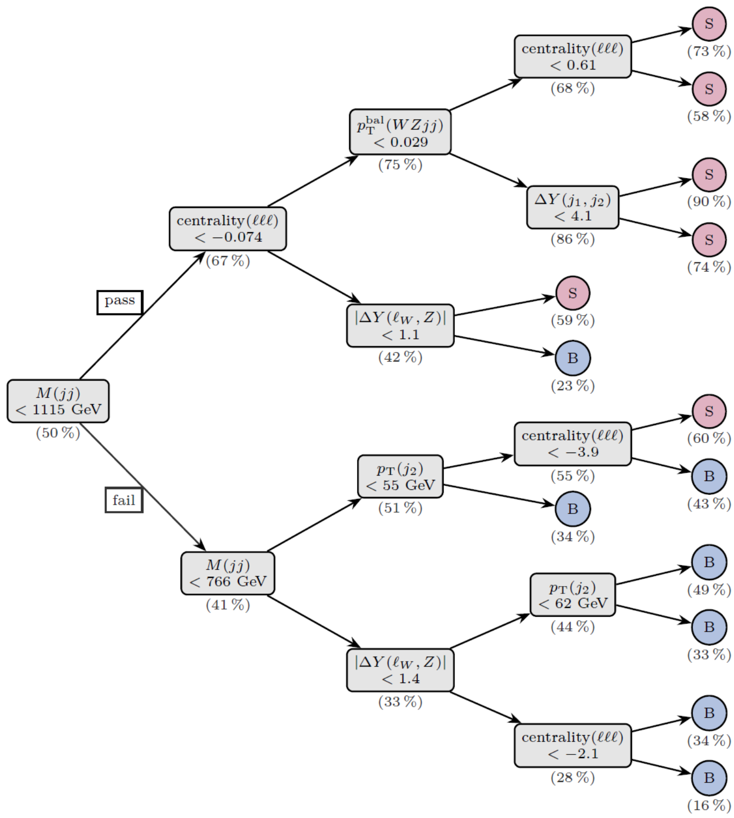



**Figure 8.2:** "Visualization of a single binary decision tree from the set of trees in the baseline BDT. Events are classified by evaluating the selection indicated in the current node (rounded rectangles) starting from the root node on the left. Events passing the indicated cut in each node follow the upper edge, those failing, follow the lower edge. Once a leaf node (circle) is reached the event is classified depending on whether the leaf node has a signal purity of more than 50% (signal leaf node), or as background otherwise. The signal purity in each node is indicated in grey below each node." From Ref. [237].

An individual decision tree is built iteratively. The starting point for "growing" such a tree is a root node containing the entire signal and background samples. These are supplied as vectors of selected observables for each event. This and subsequent nodes are then each split into two new ones. To determine how a given node should be split, a search over the available variables is performed. For each variable, the possible cut values are determined by the `nCuts` options. Based on this, `nCuts` equidistant values between the variable's minimum and maximum value in the training dataset are tried. A node is then split into two nodes based on the Gini Index $p \cdot (1-p)$, where $p = \frac{S}{S+B}$ is the signal purity. The variable and cut value that the splitting should be based on are chosen such that the increase in the Gini Index is maximized. This is assessed by comparing the weighted averages of the child nodes' indices to the Gini Index value of the parent node. The weighting is done according to their relative fractions of events. Nodes are split until either the depth of the tree is at the value of `MaxDepth` or the fraction of events in a node compared to the total training sample is smaller than `MinNodeSize`. Any node not split further is a leaf node and is classified as signal if the purity is larger than 0.5 and as background otherwise.

### 8.1.1 Boosting

To improve classification performance and reduce the influence of statistical fluctuations, a boosting approach is used. The main idea is to combine multiple base functions with low discriminative power, so-called "weak learners". Instead of using a single decision tree for the classification, a collection of multiple trees is trained and combined. The final boosted decision tree classifier has the form

$$F(x; P) = \sum_{m=0}^{M} \beta_m f(x; a_m); \quad P = \{(\beta_m, a_m) \mid m = 0, 1, 2, \ldots, M\} \tag{8.1}$$

with $x$ being the observable vectors for each event, $f(x; a_m)$ the individual trees with parameters $a_m$, and $\beta_m$ the real numbered weight of each tree.

The role of the boosting procedure is now to find the parameters $P$ such that the classifier has the smallest deviation to the true value $y$. This deviation is measured using a loss function $L(F, y)$, which defines the boosting procedure. As optimizing the whole loss function for the full boosted decision tree with all its weights and parameters of each tree is not feasible, an iterative approach is used. Here, a tree is added one by one with its parameters and weight optimized to minimize the total loss while leaving all the previous trees intact:

$$(\beta_m, a_m) = \arg\min_{\beta, a} \sum_{i=1}^{N} L(f_{m-1}(x_i) + \beta f(x_i; a), y_i), \tag{8.2}$$

where $f_{m-1}$ is the BDT up to the previous iteration, $f$ is the current tree to be added and $x_i$ is the observable vector for event $i$ out of $N$. This is done until the number of trees matches the specified `NTrees`.

The boosting algorithm employed in this work is TMVAs `Grad` implementation of gradient boosting and uses a loss function

$$L(F, y) = \ln(1 + e^{-2F(x)y}), \tag{8.3}$$

where $y$ is the vector of target labels. Unlike the earliest and most popular boosting method, AdaBoost [238], this loss function is more robust to outliers. However, it does not possess as easy a boosting algorithm. Instead, the minimization is done via steepest-descent. For this, the gradient of the loss function is calculated with respect to the predictions of the BDT of the previous iteration $f_{m-1}(x)$. The tree for the new iteration is then grown and adjusted to approximate this gradient. The addition of this new tree represents a step along the negative gradient. The value of $\beta_m$, controlled through the `Shrinkage` setting, then takes the role of a step length or learning rate. A value below one can increase the robustness of the BDT. To increase the stability of the classifier even further, the boosting is combined with a bagging-like resampling procedure. Here, every tree is trained only on a fraction of the whole sample. This option is enabled through the `UseBaggedBoost` switch. The value of `BaggedSampleFraction` controls the fraction of events that each tree sees.

In practice, the constituent trees forming a BDT constructed through gradient boosting are typically regression trees. These represent a generalization of decision trees. Unlike decision trees, where every leaf is labeled as signal or background, regression trees assign labels from a continuous range of numbers that are indicative of the node's signal purity.

More information about the used implementation can be found in Ref. [236], and a more in-depth discussion of boosting is shown in Ref. [239].

## 8.2 Training Setup

This section summarizes the overall setup for the BDT used in this work. The implementation uses an interface to TMVA and was developed for Ref. [237].

While multiple settings mentioned in the previous section are optimized with the procedure described in Section 8.3, some are fixed from the start. As mentioned in the previous section, one of these is the chosen boosting algorithm, which is set as `BoostType` = *Grad*. Bagging is enabled by setting `UseBaggedBoost` and the fraction of events that each tree sees is set to 60 % via `BaggedSampleFraction` = 0.6. As not many events with negative weights are expected, they are treated as normal using the default setting of `NegWeightTreatment` = *pray*.

All other settings not specified here or listed for optimization in Section 8.3 are kept at their default values.

### 8.2.1 Training Inputs

The BDT is trained to discriminate the 00 polarization (signal) state against all the other polarization states (background). The training does not include the other non-WZ-QCD background processes due to their low impact and MC statistics in this region. The signal and background sets are split into two statistically independent sub-samples according to the parity of each event's EventNumber. The classification method is trained on the first sample, the training sample, to create the BDT score distribution. It is then tested on the second sample to check the performance and potential overtraining of the classifier.

### 8.2.2 Training Variables

The training is performed using the events that pass the selection defined in Section 5.2.2. The procedure for building BDTs, as described in the previous section, has two major shortcomings that necessitate a preprocessing of the input variables to achieve the best performance possible.

The first is the fact that, at any point, only one-sided cuts are considered. Window cuts are only possible through two individual cuts at different layers but in the same branch of a tree. However, these are not guaranteed to be tried as it is unlikely that the first of the two cuts required would be the one with the highest separation power. As such, it is useful to use the absolute value for variables that are symmetric around zero but have a difference in their absolute values between signal and background.

The second concerns variables containing long but small tails in their distribution. For a fixed value of `nCuts`, the tried cut values for building a tree are taken from equidistant values between the maximum and minimum values of the variable. As such, long tails or a single outlier event at extremely high energy can cause very few cuts to be tested in regions of interest at lower energy. To circumvent this issue, two transformations are applied. The first is the use of logarithms for momentum-related distributions. This significantly reduces the impact of outlier events and long tails. Additionally, many distributions have hard minimum and maximum thresholds applied. This replaces events with low or high values with the respective threshold value.

The sensitivity of several observables to the different diboson polarization states was investigated. Table 8.1 shows the full list of observables considered for the training. It also includes the transformations applied to the variables in the two signal regions as well as values for the following figures of merit when using `nCuts` = 20. For all evaluation metrics $n_{1,i}$ ($h_{1,i}$) and $n_{2,i}$ ($h_{2,i}$) are the values of the $i$th bin of the normalized (cross-section scaled) distributions of the two input histograms. In this case, these are usually the signal and background histograms.

- Non-overlap ($f_{\text{N-O}}$):

$$f_{N-O}(n_1, n_2) = 1 - f_O = 1 - \sum_{i=1}^{N_{bins}} min(n_{1,i}, n_{2,i}) \tag{8.4}$$

The non-overlap represents the difference in the overlap of the two normalized distributions from one. If both distributions are identical, then their minimum is always their value, the sum of which is one. As such, two identical distributions have a Non-overlap of zero. Differences between the two distributions lead to larger values up to 1.

- Separation ($f_{\text{Sep}}$):

$$f_{Sep}(n_1, n_2) = \frac{1}{2} \sum_{i=1}^{N_{bins}} \frac{(n_{1,i} - n_{2,i})^2}{n_{1,i} + n_{2,i}} \tag{8.5}$$

The separation is a performance evaluation metric for normalized distributions, already implemented in TMVA [236] and originally from Ref. [240]. For the calculation, bins where both contributions are zero are skipped. Identical distributions have a separation of zero, while differences in the shapes lead to larger values up to one.

- Reduced $\chi^2$ ($f_{\text{red. }\chi^2}$):

$$f_{\text{red. }\chi^2}(n_1, n_2) = \frac{1}{N_{bins} - 1} \sum_{i=1}^{N_{bins}} \frac{(n_{1,i} - n_{2,i})^2}{\Delta n_{1,i}^2 + \Delta n_{2,i}^2} \tag{8.6}$$

The $\chi^2$-test is commonly used to evaluate the compatibility of two distributions. Identical distributions lead to a value of zero, while compatible distributions are indicated by values around one. For incompatible distributions, this value can grow arbitrarily large, depending on the uncertainties.

- Approximated ROC-integral ($f_{\text{ROC}}$):

$$\text{AUC} = \sum_{i=1}^{N_{bins}} \left(\min(Rej(B)_i, Rej(B)_{i-1}) \times (Eff(S)_{i-1} - Eff(S)_i)\right), \tag{8.7}$$

where

$$\begin{aligned} Rej(B)_0 &= 0 \\ Eff(S)_0 &= 1 \\ Rej(B)_i &= \frac{\sum_{j=1}^{i-1} n_{B,j}}{\sum_{j=1}^{N_{bins}} n_{B,j}} \\ Eff(S)_i &= \frac{\sum_{j=i}^{N_{bins}} n_{S,j}}{\sum_{j=1}^{N_{bins}} n_{S,j}} \end{aligned}$$

The *receiver operating characteristic* (ROC) curve describes the background rejection over the signal acceptance of a classifier. The integral under this curve (AUC) is a common metric to evaluate classifiers. Random classifiers have a value of 0.5, while a perfect classifier achieves an integral of 1. In this context of quantifying the compatibility of two distributions larger values, up to 1, indicate that the distributions are different. A value close to 0.5 indicates similar or identical distributions, while values below 0.5 are representative of an anti-classifier.

- Estimated total discovery significance ($f_{\text{SoBZ30}}$):

$$Z_{30} = \max_i \frac{S_i}{\sqrt{B_i + (0.3 \cdot B_i)^2)}}, \tag{8.8}$$

with

$$S_i = \sum_{j=i}^{N_{bins}} h_{S,j} \tag{8.9}$$

$$B_i = \sum_{j=i}^{N_{bins}} h_{B,j} \tag{8.10}$$

$$\tag{8.11}$$

Unlike all the other metrics used, this one is not only sensitive to shape differences but also to the overall normalization of the two histograms. Additionally, this figure of merit is a good approximation of the discovery significance of the signal given a constant background with a 30 % systematic uncertainty.

**Table 8.1:** Overview of the variables considered for the MVA training. Shown are the values of the figures of merit for each variable using 20 equidistant cuts after applying the listed transformations. The variables are ordered by descending $f_{\text{N-O}}$.

| Observable | Figures of merit | | | | | Transformations | | |
|---|---|---|---|---|---|---|---|---|
| | $f_{\text{N-O}}$ | $f_{\text{Sep}}$ | $f_{\text{red. }\chi^2}$ | $f_{\text{ROC}}$ | $f_{\text{SoBZ30}}$ | Function | min | max |
| $\Delta Y(\ell_W Z)$ | 0.44 | 0.255 | 1980.38 | 0.75 | 1.33 | abs | | 3.2 |
| $\cos\theta_V$ | 0.40 | 0.215 | 1661.14 | 0.74 | 1.25 | abs | | |
| $\Delta Y(WZ)$ | 0.40 | 0.213 | 1645.45 | 0.72 | 1.25 | abs | | 3.2 |
| $\Delta Y(\ell_W, \ell_Z^-)$ | 0.39 | 0.207 | 1609.59 | 0.72 | 1.23 | abs | | 3.2 |
| $cos\theta_{\ell_Z}$ | 0.21 | 0.066 | 499.30 | 0.61 | 1.02 | abs | | |
| $p_T(\ell_2^Z)$ | 0.20 | 0.060 | 451.46 | 0.59 | 1.01 | log | | 2.6 |
| $M_T^{WZ}$ | 0.18 | 0.046 | 339.85 | 0.59 | 0.98 | log | 2.2 | 3.0 |
| $M_{WZ}$ | 0.18 | 0.046 | 338.40 | 0.59 | 0.98 | log | | 3.0 |
| $cos\theta_{\ell_W}$ | 0.16 | 0.037 | 277.30 | 0.58 | 0.98 | abs | | |
| $p_T(\ell_W)$ | 0.13 | 0.027 | 189.35 | 0.52 | 0.97 | log | | 2.5 |
| $P_T^{WZ}$ | 0.12 | 0.021 | 149.01 | 0.55 | 0.96 | log | | 2.7 |
| $E_T^{Miss}$ | 0.11 | 0.021 | 167.88 | 0.47 | 0.97 | log | 1.4 | 2.6 |
| $r_{21}$ | 0.11 | 0.017 | 120.86 | 0.49 | 0.96 | | | |
| $p_T(\ell_1^Z)$ | 0.10 | 0.012 | 84.73 | 0.52 | 0.96 | log | | 2.6 |
| $\Delta\Phi(\ell_W, l_Z^-)$(lab) | 0.09 | 0.020 | 156.98 | 0.47 | 0.97 | abs | | |
| $\Delta\Phi(\ell_W, l_Z^-)(WZ)$ | 0.09 | 0.019 | 151.90 | 0.47 | 0.97 | abs | | |
| $\eta(\ell_W)$ | 0.07 | 0.010 | 75.40 | 0.52 | 0.96 | abs | | |
| $p_T^W$ | 0.06 | 0.010 | 66.51 | 0.50 | 0.96 | log | 0.3 | 2.7 |
| $p_T^Z$ | 0.05 | 0.004 | 25.51 | 0.46 | 0.96 | log | 0.3 | 3 |
| $M_T^W$ | 0.03 | 0.005 | 37.27 | 0.44 | 0.96 | log | 1.5 | 2.4 |
| $\eta(\ell_Z^1)$ | 0.03 | 0.002 | 18.69 | 0.47 | 0.96 | abs | | |
| $M_Z$ | 0.02 | 0.001 | 9.57 | 0.47 | 0.96 | | 81.0 | 101.2 |
| $\eta(\ell_Z^2)$ | 0.02 | 0.001 | 8.01 | 0.47 | 0.96 | abs | | |
| $\phi(\ell_Z^1)$ | 0.01 | 0.000 | 1.18 | 0.47 | 0.96 | abs | | |
| $\phi(\ell_W)$ | 0.01 | 0.000 | 0.92 | 0.47 | 0.96 | abs | | |
| $\phi(\ell_Z^2)$ | 0.01 | 0.000 | 0.61 | 0.47 | 0.96 | abs | | |

Figure 8.3 shows comparisons of some of the included variables after the application of the transformations. It can be seen that $\Delta Y(\ell_W Z)$ offers the strongest separation power between signal and background. $p_T(\ell_W)$ and $P_T^{WZ}$ also still show some discriminating power, while the $\phi(\ell_Z^2)$ distributions for signal and background are identical. Due to this, the raw $\phi$ and $\eta$ distributions of the leptons and bosons are not considered during the optimization of the final BDT. They are not sensitive to differences between the signal and background and would likely only prolong the optimization process.

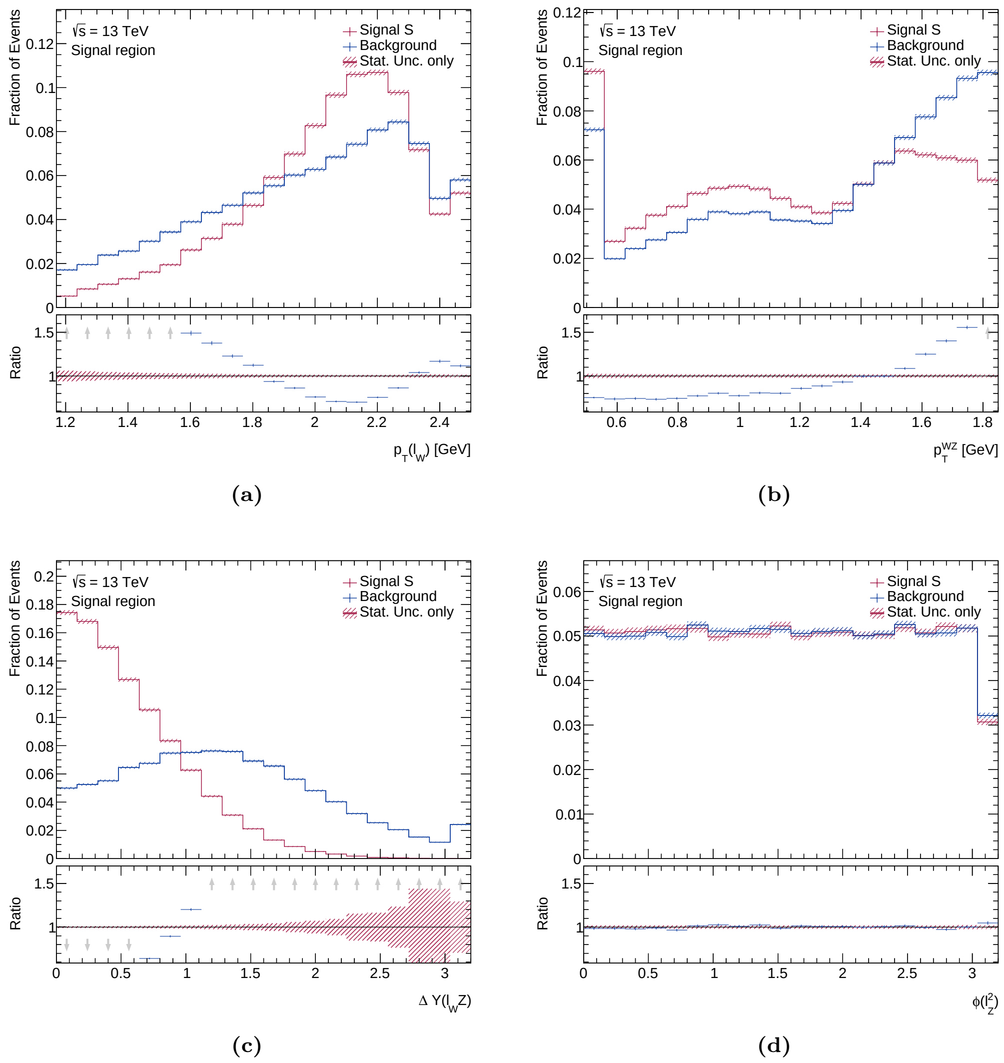



**Figure 8.3:** Comparison of normalized distributions between signal (00 polarization) and background (other polarizations) for (a) $p_T(\ell_W)$, (b) $P_T^{WZ}$, (c) $\Delta Y(\ell_W Z)$ and (d) $\phi(\ell_Z^2)$. The shown uncertainties only include the MC statistical uncertainties.

Two more variables are excluded from the optimization process for different reasons. The scattering angle $\cos\theta_V$ was excluded in an early step in the analysis due to mismodeling concerns seen in the inclusive region. This is visible in Figure 6.13c. Additionally, $p_T^Z$ and $p_T^W$ are not considered to not unnecessarily bias the classifier for the differential measurement. More information about this can be found in Appendix D.

## 8.3 Optimization

In order to get a BDT that performs optimally, an optimization is performed for picking both the variables to use and the meta-parameters of the BDT. In this context, the term 'meta-parameters' denotes the configurable settings mentioned in Section 8.1. The base procedure is taken from Ref. [237] with a complete reimplementation done first by the author of this thesis and later again by the author of Ref. [241].

Attempts to generalize the approach to also test and optimize deep neural networks encountered technical hurdles that could not be solved in the time frame of the underlying analysis. Instead, the author of Ref. [241] took on that task and developed the OPTIMA framework [242].

To maximize correct modeling and sensitivity while minimizing overtraining, the variables to use as input and the meta-parameters used were optimized. The optimization procedure starts with a variable optimization. Then, the meta-parameter optimization is done before a second variable optimization concludes the procedure. The meta-parameters of the initial BDT are displayed in Table 8.2.

### 8.3.1 First Variable Optimization

The variable optimization is done iteratively and uses the initial BDT's meta-parameters throughout. It starts with $\Delta Y(\ell_W Z)$, which has the best separation power between signal and background when considering a singular variable. From there, variables are added or dropped one by one, looking to maximize the approximated significance $f_{\mathrm{SoBZ30}}$ on the test set while not being overtrained. To quantify the overtraining, the red. $\chi^2$ between the BDT's output distributions on the training and test samples are compared for signal and background separately. If either value exceeds 2.0 or both are larger than 1.5, the configuration is considered overtrained. Both criteria roughly indicate the same strong overall level of disagreement between training and test distributions. Overtrained configurations are discarded, and the next best configuration is checked instead until one that is not overtrained is found. This overtraining check is not performed for the first two iterations as there is a high likelihood that all configurations are overtrained when using that few variables. If, in one iteration, the highest approximated significance is achieved by not adding any variables or by removing a variable, then the variable optimization is stopped.

**Table 8.2:** List of meta-parameters set at the beginning of the optimization.

| Parameter | Value |
|---|---|
| NTrees | 200 |
| MaxDepth | 4 |
| MinNodeSize | 2.5% |
| nCuts | 20 |
| Shrinkage | 0.06 |

### 8.3.2 Meta-parameter Optimization

For the meta-parameter optimization, a simple grid search over a multitude of values is performed while using the observables obtained from the first variable optimization. The values included in the search are:

- MaxDepth: 3, 4, 5,
- MinNodeSize: 1%, 2%, 2.5%, 4%,
- nCuts: 20, 50,
- Shrinkage: 0.01, 0.02, 0.04, 0.06,
- NTrees: 200, 400, 600, 800, 1000, 1300.

They are chosen based on the good results observed in Ref. [237] and mostly lay on both sides of the TMVA default values.

Afterward, the best configuration is chosen by selecting the one with the highest approximate significance while considering the same overtraining thresholds as before. The resulting optimized meta-parameter values are displayed in Table 8.3.

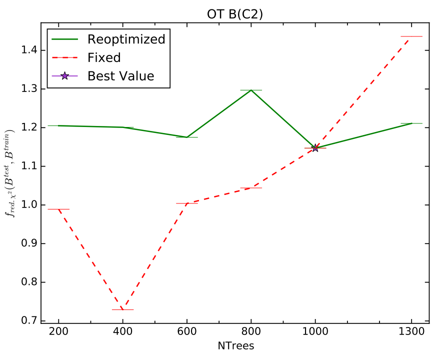
**Table 8.3:** List of optimized meta-parameters for the $p_T^Z > 200\,\text{GeV}$ region.

| Parameter | Value |
|---|---|
| NTrees | 1000 |
| MaxDepth | 4 |
| MinNodeSize | 1% |
| nCuts | 50 |
| Shrinkage | 0.02 |

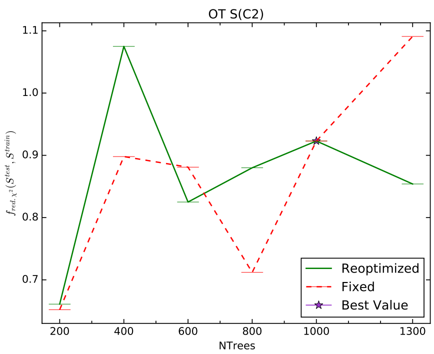
To gauge whether the chosen meta-parameter grid covers the range of optimal values, the impact of the individual meta-parameters is investigated.

The effect of changing the `NTrees` meta-parameter on the approx. significance and the $\chi^2$ overtraining values can be seen in Figure 8.4. Figure 8.4c shows that the test significance decreases by $\approx 0.1$ if `NTrees` is reduced from 1000 to 200, while the other parameters remain fixed. However, most of the sensitivity can be recovered if the other parameters are subsequently reoptimized by finding the best performing, already trained, grid point that matches the set value of `NTrees`. In that case, the drop in the significance is less than 0.02. The other two subfigures show a rough dependency of the overtraining metrics on the value of `NTrees`. Without reoptimization, the $\chi^2$ between the train and test sets increases for higher values of `NTrees` and decreases for lower ones. This matches the expectation that classifiers with more degrees of freedom are more prone to overtraining. Allowing the parameters to be reoptimized eliminates this effect. Instead, no strong dependence can be seen anymore. This is because the optimization aims to maximize the test significance. For highly flexible trees, it is generally lower than the training significance due to overtraining. With the reoptimization, the test significance is increased, decreasing the overall impact of the overtraining. Meanwhile, for lower values of `NTrees`, the BDT performs suboptimally with the set of remaining meta-parameter values. In that case, the best avenue to increase the test significance is to allow the trees more freedom. This makes the model more powerful, resulting in an increase in the approximate significance but also an increase in overtraining, as expected for more powerful models.

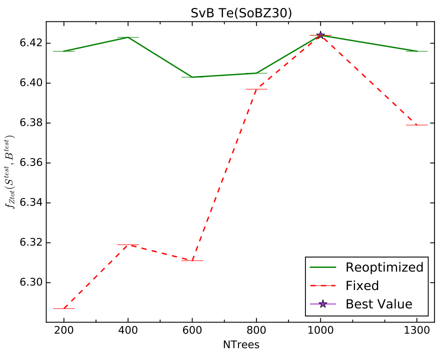
How different meta-parameters can work together to compensate or amplify these effects is demonstrated in Figures 8.5 and 8.6 for combinations of two and three meta-parameters, respectively. Figure 8.5 shows the overtraining and significance values for BDTs that have the `NTrees` and `Shrinkage` parameters fixed to specific values. Then, the other parameters are either kept constant at their value from the optimal configuration or reoptimized given the values of `NTrees` and `Shrinkage`. As above, the most extreme behavior can be seen when the other parameters are fixed. In that case, the overtraining values increase significantly when both the number of trees and the learning rate are set to large values, even more so than in the case where only `NTrees` is adjusted. The significance drops off when decreasing either the number of trees or the shrinkage. Curiously, it also decreases drastically when increasing both of these values. The reason for this is twofold. The first and biggest one is




**(a)**




**(b)**




**(c)**

**Figure 8.4:** Values of (a) $\chi^2$ between the BDT score distributions on the training and testing data for background and (b) signal events, and (c) $f_{\mathrm{SoBZ30}}$ on the test set, for BDTs with different values for the `NTrees` meta-parameter. For the red dashed line, the other parameters are kept constant, while for the green solid line, they are reoptimized. The purple star indicates the position of the overall optimal BDT.

that the significance of overtrained configurations is manually set to zero. This causes the low values on the top right corner of Figure 8.5e. The second is also due to the overtraining. This time, it causes the BDT to generalize poorly to unseen data, resulting in a reduction of the significance of the test set. These effects are all reduced when a reoptimization of the remaining parameters is performed. This is especially the case for the overtraining criteria. However, a smaller but analogous impact on the test significance remains. It stays roughly constant when setting either `NTrees` or `Shrinkage` to a high value, while the other is kept low. Setting both to lower values restricts the BDT and degrades the performance, while setting both to high values results in a slight increase in overtraining and worse generalization.

Equivalent results for an extension to three parameters are shown in Figure 8.6. Here, the `NTrees`, `Shrinkage`, and `MaxDepth` parameters are selected simultaneously, while all others are either fixed or reoptimized. The sizes of the balls, just like the colors, indicate the value of the overtraining metric or significance for each particular configuration. Larger sizes indicate larger values. The best-performing configuration is indicated by the star. Setting all three parameters to values that offer the BDT more free parameters leads to high overtraining metrics, even larger than in the case where only the first two parameters were modified. This can be seen particularly well in Figure 8.6c. Here, the middle layer represents identical values to Figure 8.5c. Increasing the maximum depth of the tree from four to five leads to $\chi^2$ values significantly larger than two. These are not shown anymore in the color grading but are shown through larger marker sizes. This consequently leads to a large number of overtrained configurations that have their significance set to zero. When the other parameters are reoptimized, this behavior is mitigated again. Each level in Figure 8.6f shows the expected result of optimal performance for the cases where only one parameter out of `NTrees` or `Shrinkage` has a larger value. However, the interplay for any of the other two parameter combinations is less pronounced.

Any other meta-parameter (combination) shows equivalent but weaker effects, overall indicating that `NTrees` and `Shrinkage` are the two most important meta-parameters.

### 8.3.3 Second Variable Optimization

The second variable optimization follows a similar process as the first, with one key difference. In this case, sets of two variables are considered to be added at each iteration step to take advantage of correlations. During this step, only the $\cos\theta_{\ell_W}$ variable was added (even if the algorithm can choose two variables if deemed beneficial).

The final list of variables used in the BDT training can be found in Table 8.4. They appear in the order in which they were added during the optimization.

**Table 8.4:** List of variables used in final BDT for the $p_T^Z > 200\,\text{GeV}$ region. Variables above the horizontal line were added during the first variable optimization step, while the one below was added in the second. Additionally, their rankings in the list of considered variables according to Table 8.1, ignoring variables eliminated from the optimization, and a physical description are shown.

| Training variable | Ranking | Definition |
|---|---|---|
| $\Delta Y(\ell_W Z)$ | 1 | Rapidity difference between the $W^\pm$ lepton and $Z$ boson |
| $p_T(WZ)$ | 10 | Transverse momentum of the $W^\pm Z$ system |
| $p_T(\ell^W)$ | 9 | $W$ lepton transverse momentum |
| $p_T(\ell_2^Z)$ | 5 | Subleading $Z$ lepton transverse momentum |
| $E_T^{\text{miss}}$ | 11 | Missing transverse momentum |
| $\cos\theta_{\ell_Z}$ | 4 | Cosine of the decay angle of the $Z$ lepton defined in the $WZ$ rest frame |
| $\cos\theta_{\ell_W}$ | 8 | Cosine of the decay angle of the $W$ lepton defined in the $WZ$ rest frame |

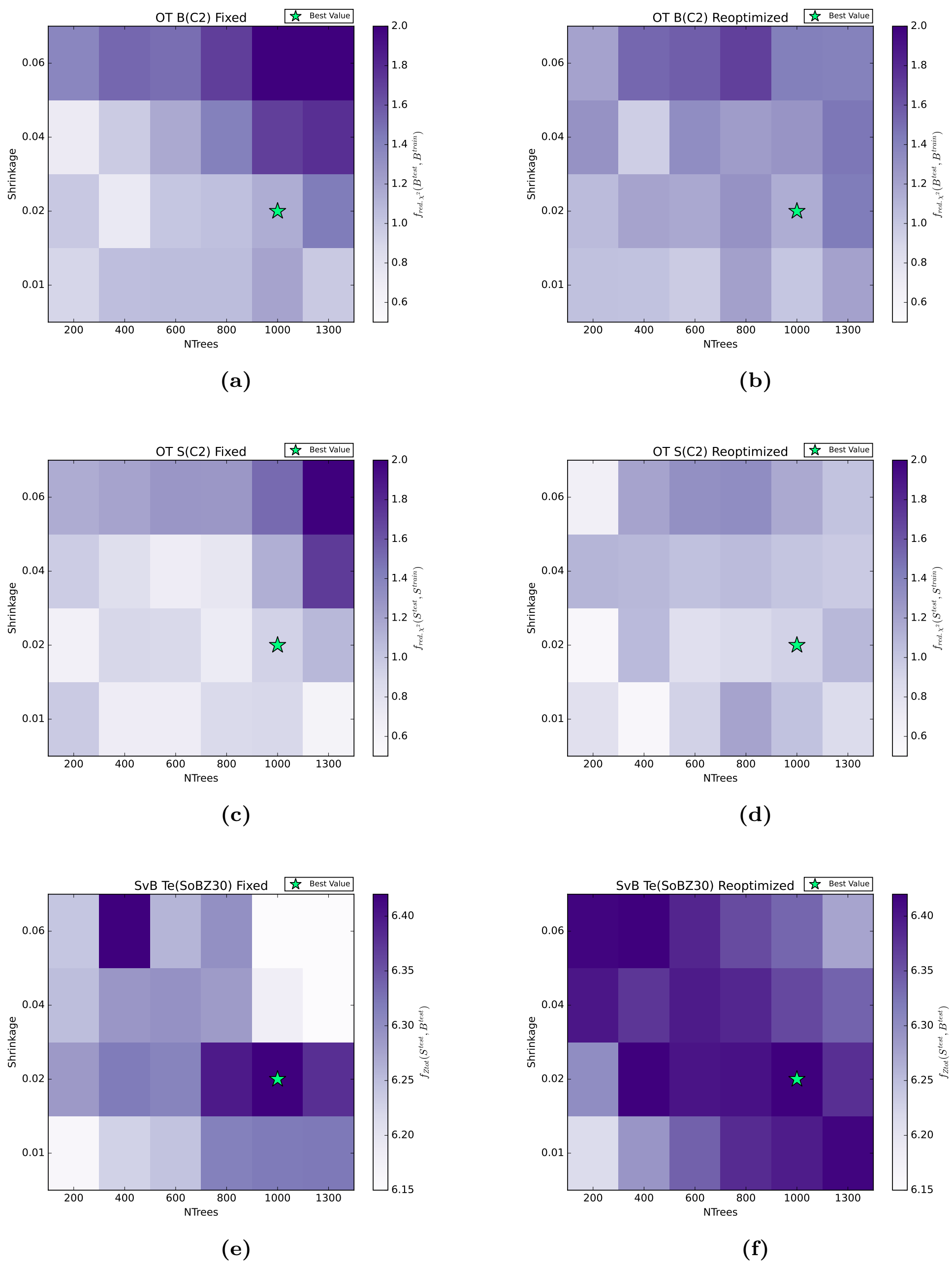



**Figure 8.5:** Values of (a) $\chi^2$ of the background and (c) signal, and (f) $f_{\mathrm{SoBZ30}}$ on the test set, for BDTs with different values for the `NTrees` and `Shrinkage` meta-parameters. There, all other meta-parameters are kept constant at their value of the best-performing BDT. Results where the other parameters are reoptimized are shown for (b) $\chi^2$ of the background and (d) signal, and (f) $f_{\mathrm{SoBZ30}}$ on the test set. The configuration of the overall best-performing BDT is marked with a star. Significances for overtrained configurations are set to zero.

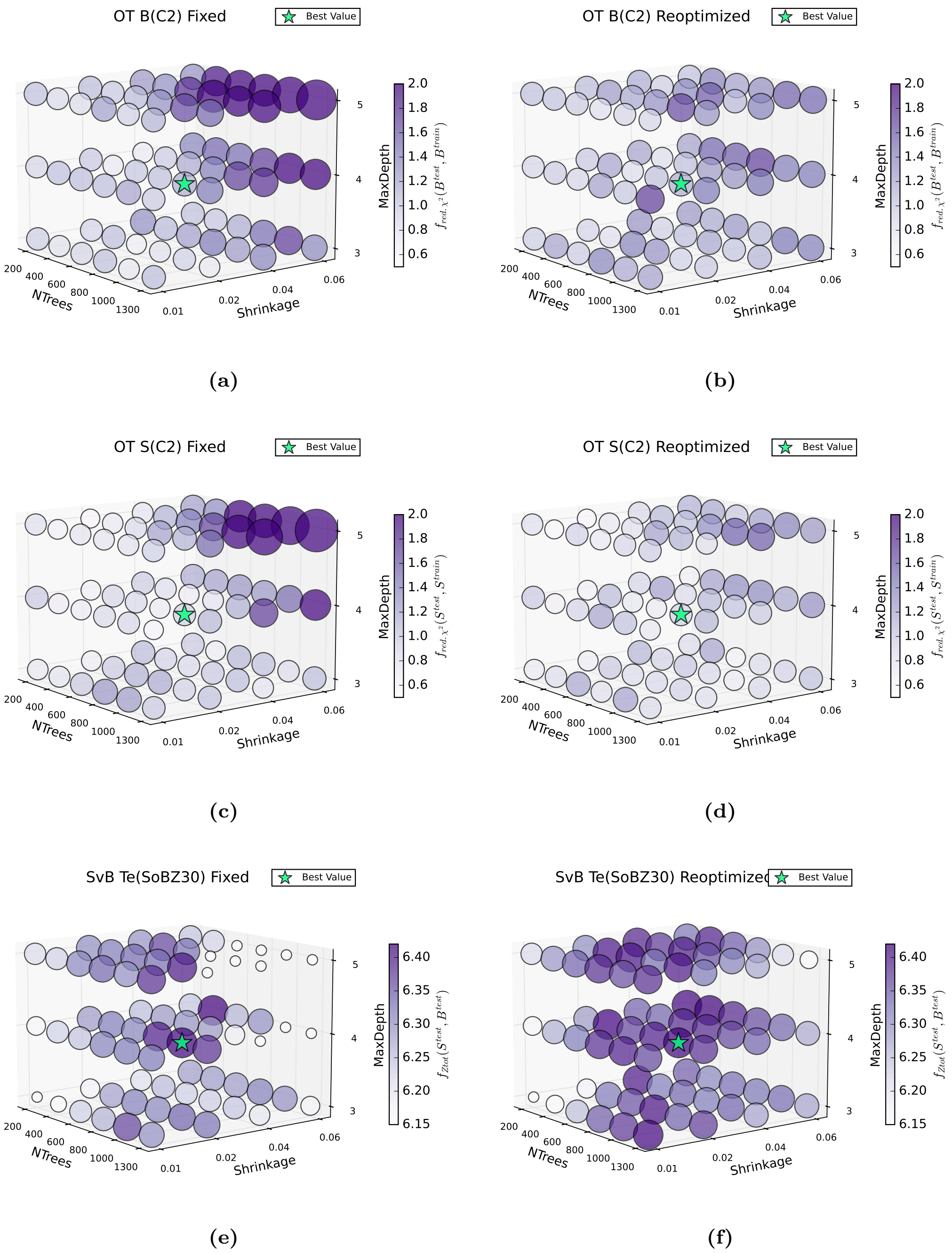
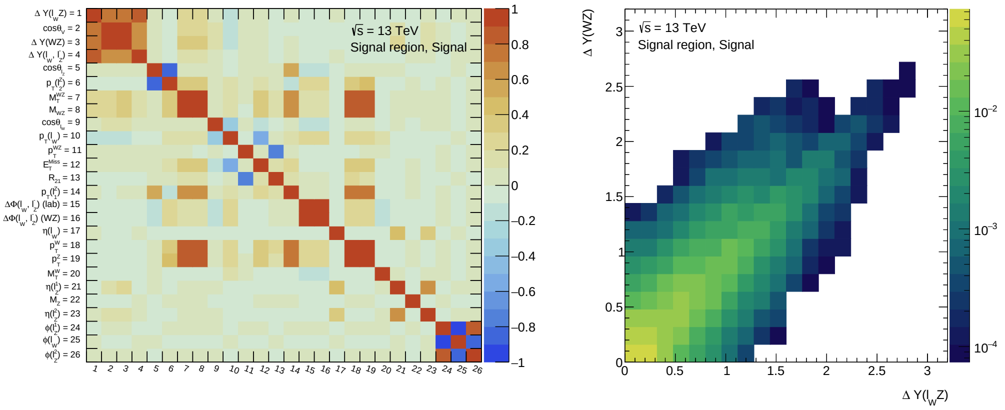



**Figure 8.6:** Values of (a) $\chi^2$ of the background and (c) signal, and (f) $f_{\mathrm{SoBZ30}}$ on the test set, for BDTs with different values for the `NTrees`, `Shrinkage` and `MadDepth` meta-parameters. There, all other meta-parameters are kept constant at their value of the best-performing BDT. Results where the other parameters are reoptimized are shown for (b) $\chi^2$ of the background and (d) signal, and (f) $f_{\mathrm{SoBZ30}}$ on the test set. The configuration of the overall best-performing BDT is marked with a star. Significances for overtrained configurations are set to zero.

### 8.3.4 Optimization Results

One point of interest here is that the chosen variables are not just the ones that are ranked highest in Table 8.1. Instead, the variables with ranks 2,3,6,7[1] are skipped. This can be understood by considering the linear correlations between variables as shown in Figure 8.7. It shows that variables ranked at places two and three have large correlations with the number 1 ranked variable, $\Delta Y(\ell_W Z)$, which is included in the training from the start. Due to this, there is only a small gain in sensitivity when also adding these variables. Instead, the procedure usually opts for variables with small linear correlations to the ones that have already been included.




(a) (b)

**Figure 8.7:** Shown are (a) the linear correlation coefficients between variables considered for the BDT training and (b) a normalized 2-D histogram of $\Delta Y(WZ)$ against $\Delta Y(\ell_W Z)$ for the region with $p_T^Z > 200\,\text{GeV}$.

Figure 8.8 shows the comparison of the final BDT's output score distribution on training and testing samples for signal and background events. A significant difference between the background and signal shape can be observed. The middle and bottom panels show the ratio between the BDT output on the training and testing samples for signal and background, respectively. The ratios agree within 10% and most of the time within the statistical uncertainties, indicating a good stability of the training method.

The progression of $f_{\text{SoBZ30}}$ on the test set and $\chi^2$ overtraining values through the optimization can be seen in Figure 8.9. From these, it is clear that the main sensitivity of the BDT comes from the first six variables, up to $\cos\theta_{\ell_Z}$. The optimization of the meta-parameters and addition of $\cos\theta_{\ell_W}$ only combine for an increase of 0.03. Figure 8.9a also shows the remarkable increase in the test significance of the BDT over the input variables. The best individual variable achieves $f_{\text{SoBZ30}}(Test) = 1.33$. Already the BDT that only contains this variable manages to reach $f_{\text{SoBZ30}}(Test) = 3.01$ while the final BDT increases this to $f_{\text{SoBZ30}}(Test) = 6.45$. Subfigure 8.9c also shows the reason why the overtraining criteria are only applied starting at the third iteration of the variable optimization. The baseline BDT, which contains only one variable, is extremely overtrained with a $\chi^2$ value between the training and test background histograms of approximately 22. The further iterations reduce this to acceptable levels. The signal overtraining value never reaches above 1.04 and ends at a value of 0.80 for the final BDT. The values for the background are somewhat larger, ending at a still acceptable value of 1.39. All of this is a first indication that a robust and sensitive classifier has been trained.

Further validation can be achieved by looking at the distributions of the final BDT. Figure 8.10 shows these for the individual polarizations. A clear separation can be seen

[1] Only counting those not explicitly excluded from training.

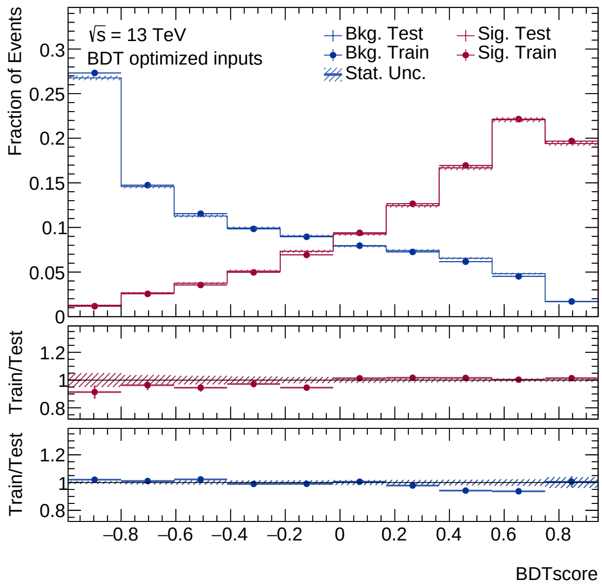


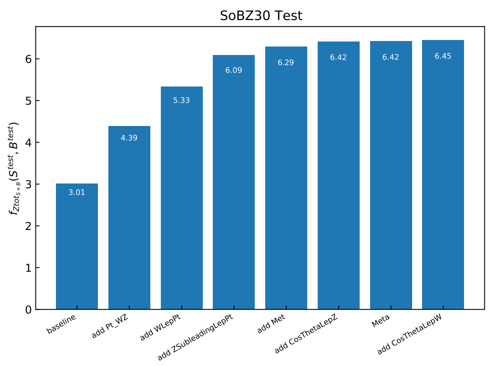

**Figure 8.8:** Comparison of the BDT score distribution of the signal (00 polarization, red) and the sum of the backgrounds (other polarizations, blue) for the training samples (dots) and the testing samples (bars). The middle and bottom panels show the ratio between training and testing for signal and background, respectively.

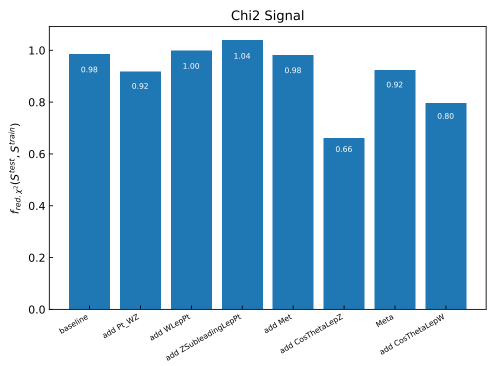
between the 00 and TT states, which make up the total signal and the majority of the background training input for the BDT. The shapes of the mixed 0T and T0 states lie somewhere between those two. They are effectively identical, leading to difficulties when trying to distinguish between them and, to a smaller degree, between them and the 00 polarization.

Lastly, the BDT score distributions obtained by using Sherpa and the sum of the polarized Madgraph samples are compared in Figure 8.11. These distributions agree within their uncertainties, giving the final indication that the training and optimization procedure has created a robust BDT and that the sum of polarization samples accurately model the full process, including NLO QCD effects.

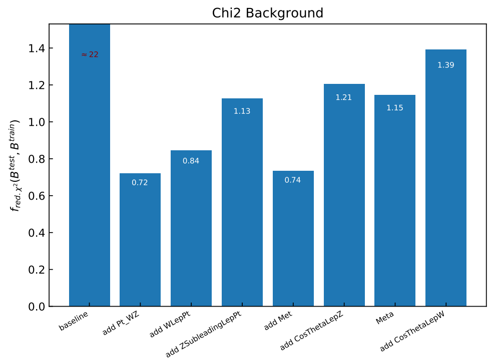
## 8.4 Setup for the $100 < p_T^Z \leq 200\,\mathrm{GeV}$ Region

The training input for the BDT in the $100 < p_T^Z < 200\,\mathrm{GeV}$ region is again derived as discussed in Section 5.2.2. However, here only a cut of 100 GeV is used without an upper bound. Additionally, the list of final variables is reused from the $p_T^Z > 200\,\mathrm{GeV}$ region. The only change is that the upper limit for $p_T(\ell_2^Z)$ is set to 2.8 instead of the 2.6 used for the higher $p_T$ region. Both the lack of an upper $p_T^Z$ cut as well as the different limit on $p_T(\ell_2^Z)$ originate from the practical evolution of the analysis and were deemed to not have large enough impacts to change later on. The meta-parameters are then taken from a separate optimization, and a distinct final BDT is trained for this region. The final meta-parameter settings are listed in Table 8.5.



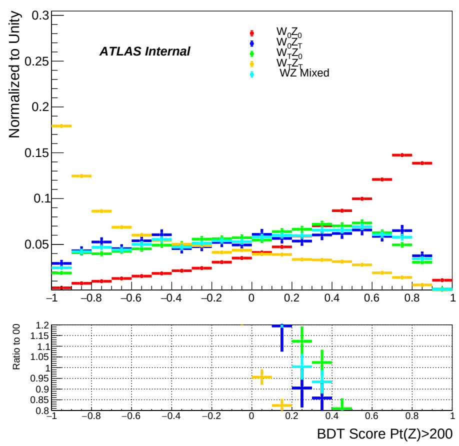



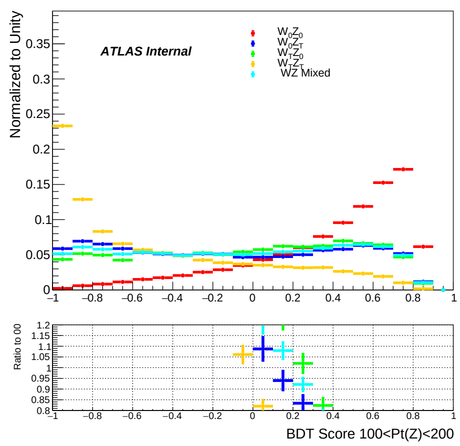

**(a)** **(b)**




**(c)**

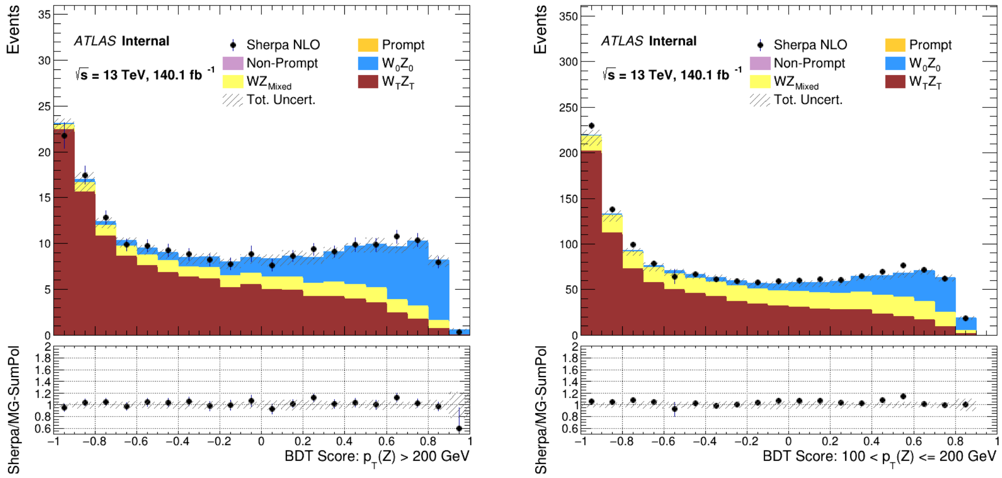
**Figure 8.9:** Progression of (a) $f_{\text{SoBZ30}}$ on the test set, (b) $\chi^2$ between the BDT score distributions on the training and testing data for signal and (c) background events through the optimization procedure.




**(a)** **(b)**

**Figure 8.10:** BDT score distributions for (a) $p_T^Z > 200\,\text{GeV}$ and (b) $100\,\text{GeV} < p_T^Z \leq 200\,\text{GeV}$ obtained for different 00, 0T, T0, and TT polarization states. There is a clear shape difference between the 00 state and the other polarization contributions. The shapes of the mixed polarization (0T and T0) are almost identical to each other. The "WZ Mixed" entry represents the sum of the 0T and T0 states. From the author of Ref. [161].




(a) (b)

**Figure 8.11:** Comparison between the inclusive NLO Sherpa sample (no separation of polarization states) and the LO-merged MadGraph sample (all polarization states summed together) for the BDT Score in the regions (a) $p_T^Z > 200\,\text{GeV}$ and (b) $100\,\text{GeV} < p_T^Z \leq 200\,\text{GeV}$. From the author of Ref. [161].

**Table 8.5:** List of optimized meta-parameters for the $100 < p_T^Z \leq 200\,\text{GeV}$ region.

| Parameter | Value |
|---|---|
| NTrees | 800 |
| MaxDepth | 5 |
| MinNodeSize | 1% |
| nCuts | 50 |
| Shrinkage | 0.01 |

# Chapter 9

# Fraction Measurement Setup

Currently, no vector boson polarization detector exists. As such, it is not straightforward to measure the polarization fraction in $W^{\pm}Z$ events. To do so regardless, a template fit approach is used. In that, MC predictions for the signal and background contributions are simulated. The differences in the predicted kinematic distributions are enhanced using the classifiers described in Chapter 8. Their contributions are then scaled such that the best match to observed data is found. From this, the different polarization fractions can be read off.

This chapter describes this approach for the measurement of the fraction of doubly longitudinally polarized WZ events at high values of $p_T(Z)$. Section 9.1 describes the general statistical approach used for this thesis. In Section 9.2, some more process/analysis-specific caveats are introduced. The set of different fit configurations is shown and explained in Section 9.3.

The introduction of the general statistical aspects at the beginning of this chapter uses the standard parametrization of the likelihood using scaling factors $\mu$ as they are commonly used in ATLAS analyses. Later in this chapter, the reparametrization to fractions is discussed. However, the general points are identical to those in the $\mu$ scenario. As that parametrization is slightly easier to understand and more common, it serves as the foundation of this chapter.

## 9.1 Statistical Analysis

The basis of the statistical analysis in this thesis is frequentist (also called classical) statistics where the probability of an event $A$ is the relative frequency with which that this event is measured given an infinite number of measurements [243]

$$P(A) = \lim_{n\to\infty} \frac{\text{number of occurrences of A in n measurements}}{n}. \tag{9.1}$$

Thus, under the assumption that each physical parameter has exactly one (unknown) value, the probability that a given value is realized in nature is either 0 % or 100 %. Without the addition of a degree of belief that any hypothesized theory is true, one can only determine the probability to measure some value under an assumed theory. It is not possible to state a probability that the theory is true given the measured data. Doing so requires the introduction of an aforementioned degree of belief to calculate

$$P(\text{theory}|\text{data}) \propto P(\text{data}|\text{theory}) \cdot P(\text{theory}), \tag{9.2}$$

as per Bayes' theorem [243]. The notation $P(A|B)$ describes the conditional probability of $A$ given that $B$ is true.

The same notation is applied to probability density functions (pdfs) as $f(n|\mu, \Theta)$, where $\mu$ and $\Theta$ describe theory parameters that are fixed. Then $n$ is randomly distributed according to $f$ given these parameter values. For a Poisson distribution, this means that

$$\text{Pois}(n|\nu) = \frac{\nu^n}{n!} e^{-\nu} \tag{9.3}$$

describes the probability to measure a number of observed events $n$ given an expected number of events $\nu$.

### 9.1.1 Likelihood

As it is not possible to prove (or disprove) a given theory, the likelihood $\mathcal{L} = P(data|theory)$ is used to quantify the compatibility of measured data with an assumed theory. Based on this, one then tries to rule out as many theories[1] as possible.

For the measurement of the fractions and the calculation of the significance of the hypothesis that the $f_{00}$ fraction is larger than zero a python wrapper [244] around ROOT's [235] HistFactory [245] is used. HistFactory itself is a wrapper around RooStats [246] and RooFit [247]. The statistical framework was originally developed for Refs. [20, 248] and later also applied in Ref. [237]. Significant further improvements were also made in Ref. [249]. This section follows the notation from Ref. [249] and only gives a basic introduction into binned likelihood fits as well as describes the points unique to this work. In all other points, the framework is used as described in that work. A quick overview of some more details of the fits and their validations is given in Section 9.1.4. In-depth looks, formulas, and derivations for these aspects can be found in Ref. [249].

The framework uses a binned maximum likelihood fit using the likelihood $\mathcal{L}(n|\mu,\Theta)$ as constructed by HistFactory's "MakeModelAndMeasurementFast":

$$\mathcal{L}(n|\mu,\Theta) = C_\gamma(\gamma)C_\lambda(\lambda)C_\alpha(\alpha)\prod_b \text{Pois}(n_b|\nu_b(\gamma_b,\lambda,\alpha,\phi)). \tag{9.4}$$

This likelihood is composed of a product of Poisson distributions of observed events $n_b$ for each bin $b$ of the signal region. The expected number of events in each bin $\nu_b$ depends on nuisance parameters that can be grouped into different categories.

- $\gamma$: Nuisance parameters that encode the uncertainty related to the finite number of simulated MC events
- $\alpha$: Set of nuisance parameters describing experimental and theoretical uncertainties
- $\lambda$: The nuisance parameter of relative luminosity scaling

Additionally, it depends on a set of unconstrained parameters $\phi$ that handle the normalization of signal samples and constitute parameters of interest. Each category of nuisance parameters has a corresponding constraint term stemming from external measurements.

#### Constraints

The MC statistical nuisance parameters $\gamma$ are constrained by

$$C_\gamma(\gamma) = \prod_b \text{Pois}(m_b|\gamma_b\tau_b). \tag{9.5}$$

This represents a Poisson distribution with $m_b$ being the weighted summed number of observed MC events in bin b and $\tau_b$ the weighted summed number of expected events in bin b before scaling with $\gamma$.

The nuisance parameter dealing with the luminosity $\lambda$ is constrained by a Gaussian distribution

$$C_\lambda(\lambda) = \text{Gaussian}(1|\lambda,\Delta\lambda_{\text{nom}}) \tag{9.6}$$

with an uncertainty obtained from an external measurement as stated in Section 7.1.1.

Experimental and theoretical uncertainties are also constrained using Gaussians normalized to have a mean value of zero and a standard deviation of one. These uncertainties, as described in Chapter 7, are implemented as three-point systematics with their nominal templates as well as up and down variations covering $\approx 68\,\%$ of the expected fluctuations.

$$C_\alpha(\alpha) = \prod_{\alpha_i\in\alpha} \text{Gaussian}(0|\alpha_i,1) \tag{9.7}$$

[1]In the sense that it is very unlikely for the measured data to have originated from this theory.

### 9.1.2 Maximum Likelihood Estimate

The best estimator [243] for the true values of the parameters based on the finite available data are those that form the theory that is most likely to produce the observed data. The values of the parameters $\mu$ and $\Theta$ that do that are the maximum likelihood estimates (MLE) $\hat{\mu}$ and $\hat{\Theta}$. They are determined by

$$\mathcal{L}(n|\hat{\mu}, \hat{\Theta}) = \max_{\mu,\Theta} \mathcal{L}(n|\mu, \Theta). \tag{9.8}$$

This minimization is done using ROOT's [235] Minuit2 package [250, 251]. The case where $\mu$ is free and not fixed to a particular value is called the "signal+background" (S+B) hypothesis.

### 9.1.3 Test Statistic, p-Value, and Significance

When trying to observe or find evidence for a particular physics process, one needs to be able to assess which of two hypotheses agrees better with data and by how much.

The first hypothesis is the background-only (B, B only or null) hypothesis $\mu = 0$, where the process of interest does not exist as predicted. The competing hypothesis is that where it does exist with an unknown size $\mu \geq 0$.

#### Test Statistic

The best test statistic to quantify this is the likelihood ratio [252]

$$t = -2 \ln \frac{\mathcal{L}(n|\mu = 0, \hat{\hat{\Theta}})}{\mathcal{L}(n|\hat{\mu}, \hat{\Theta})}, \tag{9.9}$$

where $\hat{\hat{\Theta}}$ describes the values of the set of parameters $\Theta$ that maximize the likelihood with $\mu$ being fixed to the given value.

The reason that $\mathcal{L}(n|\mu = 0, \hat{\hat{\Theta}})$ is not used as the test statistic itself is that it can lead to a premature rejection of the background-only hypothesis. This is because the rejection would not just be against $\mu = 0$ but against the whole model. Due to this, a rejection could occur due to mismodelling of polarization or non-polarization backgrounds or detector effects. Additionally, a rejection could more easily occur due to statistical fluctuations that are covered by neither the background-only nor the S+B hypothesis.

This test statistic is modified slightly to not reject the background-only hypothesis for a measured negative amount of signal events [253]:

$$q_0 = \begin{cases} -2 \ln \frac{\mathcal{L}(n|\mu=0,\hat{\hat{\Theta}})}{\mathcal{L}(n|\hat{\mu},\hat{\Theta})} & \hat{\mu} \geq 0, \\ 0 & \hat{\mu} < 0. \end{cases} \tag{9.10}$$

#### p-Value

To decide whether to reject the null hypothesis, the p-value

$$p_0 = \int_{q_{0,obs}}^{\infty} f(q_0|0) dq_0 \tag{9.11}$$

is used. It represents the probability of observing a test statistic value as extreme as, or more extreme than, the one calculated from the observed data, assuming the null hypothesis is true.

In particle physics, the p-value is usually translated to a significance $Z_0$

$$p_0 = \int_{Z_0}^{\infty} \text{Gaussian}(x|0, 1) dx \tag{9.12}$$

$$Z_0 = \Phi^{-1}(1 - p_0) \tag{9.13}$$

indicating the number of standard deviations from the mean of a Gaussian distribution. Here, $\Phi$ is the cumulative distribution function of a Gaussian with mean zero and standard deviation of one

$$\Phi(x) = \int_{-\infty}^{x} \text{Gaussian}(x|0,1)dx. \tag{9.14}$$

This transformation ensures that a Gaussian-distributed variable, when located $Z_0$ standard deviations above its mean, corresponds to an upper-tail probability equal to the original p-value.

#### Significance

The most accurate way to calculate this is by using toys thrown according to $\mathcal{L}(n|\mu = 0, \hat{\hat{\Theta}})$ and calculating the test-statistic for each of the toys. To get a reliable prediction from that, one needs at least a couple of toys that pass the measured value of $q_{0,obs}$. When targeting an observation, one desires to reach a significance of $5\sigma$ corresponding to a p-value of $2.87 \times 10^{-7}$. This means that the number of toys that have to be generated needs to be of the order $\mathcal{O}(10^8)$, which is not computationally feasible.

Instead, one can fall back to the asymptotic formulas, valid in the limit of a large number of observed events, from Wald and Wilks [253–255]. Those state that under a set of reasonable assumptions

$$-2\ln\frac{\mathcal{L}(n|\mu,\hat{\hat{\Theta}})}{\mathcal{L}(n|\hat{\mu},\hat{\Theta})} = \frac{(\mu-\hat{\mu})^2}{\sigma_\mu^2} + \mathcal{O}(1/\sqrt{N}), \tag{9.15}$$

where $\hat{\mu}$ follows a Gaussian distribution with mean $\mu' = 0$ and standard deviation $\sigma$. Special casing this for the discovery statistic $q_0$ with $\mu = 0$ and when neglecting the $\mathcal{O}(1/\sqrt{N})$ term due to $N$ being large one gets

$$q_0 = \begin{cases} \hat{\mu}^2/\sigma^2 & \hat{\mu} \geq 0, \\ 0 & \hat{\mu} < 0. \end{cases} \tag{9.16}$$

With this, the pdf of $q_0$ has the form

$$f(q_0|0) = \frac{1}{2}\delta(q_0) + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q_0}}e^{-q_0/2}, \tag{9.17}$$

with a corresponding cumulative distribution

$$F(q_0|0) = \Phi(\sqrt{q_0}). \tag{9.18}$$

Using Equation 9.11 the p-value of the $\mu = 0$ hypothesis can be calculated as

$$p_0 = 1 - F(q_0|0). \tag{9.19}$$

This results in a simple expression for the discovery significance

$$Z_0 = \Phi^{-1}(1-p_0) = \Phi^{-1}(1-(1-\Phi(\sqrt{q_0}) = \sqrt{q_0}. \tag{9.20}$$

### 9.1.4 Additional Details of the Fitting Setup

This section gives a short overview of some more details and other aspects of the fit. More in-depth looks at each of these points can be found in Ref. [249].

#### Expected Yields

The expected yields $\nu_b$ in Equation 9.4 depend on the variations of $\alpha$, modelling their uncertainties. For each of these systematic uncertainties, only three points are available: the nominal distribution with $\alpha_i = 0$ and the up and down variations with $\alpha_i = \pm 1$. In between and beyond these values, their impact has to be inter- or extrapolated. To do so, the impact of each nuisance parameter (NP) is split into a shape and a normalization aspect. This is done to avoid negative total yields for large variations.

This is ensured by the fact that the normalization impact is extrapolated exponentially while the shape effect is extrapolated linearly. Both aspects are interpolated using 6th-order polynomials, although with slightly different ones. The polynomial is chosen such that its value and first two derivatives are identical to the extrapolations at the boundary $|\alpha_i| = 1$ [245, 256, 257].

The combination of different NPs for one bin works such that first, the shape effects of all parameters are summed to the nominal [256]. Afterward, the normalizations are each multiplied to the result [257].

#### Monte Carlo Statistical Uncertainties

The statistical Monte Carlo uncertainties in Equation 9.4 are represented by the constraint term $C_\gamma(\gamma)$, which is expanded in Equation 9.5. It introduces $\tau_b$ and $\gamma_b$, which represent the nominal MC prediction, as well as its scaling. $\tau_b$ is constructed such that its relative uncertainty $\frac{\sqrt{\tau_b}}{\tau_b}$ matches the one of the nominal expected number of MC events in this bin [245, 258].

#### Likelihood Scans

A numerical calculation of the MLEs can run into problems when local maxima exist. In such a case, the result of the fit might not be the true MLEs.

To check for that, likelihood scans can be performed

$$-\Delta \ln \mathcal{L}(\mu) = -\ln \frac{\mathcal{L}(n|\mu, \hat{\hat{\Theta}})}{\mathcal{L}(n|\hat{\mu}, \hat{\Theta})} = -\ln \mathcal{L}(n|\mu, \hat{\hat{\Theta}}) + \ln \mathcal{L}(n|\hat{\mu}, \hat{\Theta}), \tag{9.21}$$

that calculate the conditional maximum likelihood for a range of specified $\mu$ values. By offsetting them with the unconditional MLE, the minimum is expected at exactly zero. Any value below indicates that the unconditional fit did not find the true global maximum likelihood.

The likelihood scans can additionally be used to read off the significance at $\mu = 0$ and determine the asymmetric standard deviation. This approach is used by Minos from Minuit2 [251] by determining values for $\mu$, such that

$$-\Delta \ln \mathcal{L}(\mu) = 0.5. \tag{9.22}$$

The same approach can be applied to other parameters besides $\mu$.

#### Uncertainties and Correlations

As described in the previous section, accurate asymmetric uncertainties for parameters can be obtained through Minos. Approximate symmetric uncertainties can also be obtained from the covariance matrix [243]. However, these are not used in this thesis. Instead, when symmetric uncertainties are shown for brevity, they are constructed by symmetrizing the uncertainties from Minos.

Nevertheless, this covariance matrix, as constructed by Hesse from Minos [251], is used for calculating the correlations between parameters in the fit as presented in this work.

#### Uncertainty Splitting and Breakdowns

It is often of interest to know not only the total uncertainty of a parameter but also the impact of individual sources. This information allows for better focus when aiming to reduce the total uncertainty. For example, if one finds that the dominant uncertainty is the data statistical one, then it might not make much sense to spend a lot of time reducing theoretical uncertainties further for this process. At least not yet. Theorists could better spend their time improving calculations for processes where theoretical uncertainties are dominant.

For this purpose, a set of conditional fits is performed where, subsequently, more and more parameters are fixed to their best-fit values. The procedure is briefly explained for the up uncertainty. The evaluation for the down uncertainties follows the same principle.

The procedure starts with the completely unconditional fit from which $\Delta\hat{\mu}^{up,total}$ is retrieved. Next, the fit is redone with a parameter, or set of parameters, $X$ set constant to their best-fit values from the unconditional fit. The up uncertainty due to the parameter(s) $X$ is subsequently calculated as the squared difference between $\Delta\hat{\mu}^{up,total}$ and the uncertainty from this more constrained fit $\Delta\hat{\hat{\mu}}^{up}(\hat{X})$ as

$$\Delta\hat{\mu}^{up,X} := \sqrt{(\Delta\hat{\mu}^{up,total})^2 - (\Delta\hat{\hat{\mu}}^{up}(\hat{X}))^2}. \tag{9.23}$$

The uncertainty for the next parameter(s) $Y$ is subsequently calculated as

$$\Delta\hat{\mu}^{up,Y} := \sqrt{(\Delta\hat{\hat{\mu}}^{up}(\hat{X}))^2 - (\Delta\hat{\hat{\mu}}^{up}(\hat{X},\hat{Y}))^2} \tag{9.24}$$

This is done for all parameter(-set) groupings until all parameters are set constant and the stat.-only uncertainty is determined from that. Then, by construction, the quadratic sum of all of these uncertainties equals the total uncertainty.

#### Pre- and Post-Fit Plots and Tables

An additional step in verifying the correct behavior of the fit and the applicability of the chosen model can be taken by inspecting the pre- and post-fit distributions themselves. For that, the content, as well as the uncertainties of each bin, have to be determined for the pre- and post-fit configurations. These can then be displayed in plots for easy visual inspection or tables for more detailed checks. This is done for each bin individually as well as the sum of all bins. Furthermore, the procedure is repeated for each sample by itself and the sum of all samples.

The bin contents for every combination of parameter values are extracted directly from the pdf. For each nuisance parameter, the up and down yield of a bin is determined by setting the parameter to their up/down values as determined by Equation 9.22. All other parameters are set to their best-fit value of the unconditional fit. The uncertainty corresponding to a specific parameter is then symmetrized as the average deviation from the nominal yield.

For the pre-fit yields, the uncertainties of each parameter are added in quadrature. For the post-fit yields, the correlations between the parameters as determined by Hesse [251] are taken into account.

Another complication is posed by statistical uncertainties. For the pre-fit case, this only represents the individual MC uncertainties for each sample. These are extracted directly from the input distributions. Post-fit, the statistical uncertainties come from the main parameter of interest.

#### Pull, Impact and Ranking Plots

It is often of interest to know how much a nuisance parameter's best-fit value and uncertainty differ from their pre-fit counterparts. Deviations in the best-fit value are called pulls, while changes in the uncertainties are referred to as the parameter being (over/under-)constrained.

Large changes indicate that the physics process under investigation offers a significant amount of new information compared to the external measurements that were done to obtain the parameter's value and uncertainties. If any of these appear, it needs to be checked if they can be understood from a physics point of view. Otherwise, they might indicate technical issues with the fit.

For this purpose each nuisance parameter $\theta_i \in \Theta$ is plotted with its post-fit value $\hat{\theta}_i$ as well as uncertainty $\Delta\theta_i$ against the pre-fit value and uncertainty $\theta_{i,0}$ and $\Delta\theta_{i,0}$.

Unconstrained parameters do not necessarily have pre-fit values or uncertainties determined by auxiliary measurements or physical arguments. Due to this, sensible but in the end arbitrary pre-fit values are chosen. This choice of these values does not impact any other results in this work, such as the best-fit values of the parameters of interest or their uncertainties. They are purely pedagogical. The nominal pre-fit values $\phi_{i,0}$ are set by fixing all normalization parameters $\mu$ to one in the formula of the parameter of interest as given in Section 9.2.1. The uncertainties are based on these but differ between parameter groups. Table 9.1 shows the chosen up and down uncertainties for each parameter group.

**Table 9.1:** Choice of pre-fit up- and down-uncertainties for unconstrained parameters. For each parameter(group) they are based on the chosen prefit value. The uncertainty for $N_{tot}$ is the one expected from a Poisson distribution, while the fractions use 100 % of their pre-fit value.

| Parameter | Down unc. | Up unc. |
|---|---|---|
| fractions $f$ | $-f_{pre}$ | $\min(1 - f_{pre}, f_{pre})$ |
| total scaling $N_{tot,pre}$ | $-\sqrt{N_{tot,pre}}$ | $\sqrt{N_{tot,pre}}$ |

An additional point of interest in inspecting a likelihood fit is understanding the impact of each nuisance parameter on the parameters of interest as well as getting an ordering of this with the highest impact. For this, each parameter is fixed to its pre- and post-fit up and down values. For each of these variations, the fit is redone and the relative difference of the best-fit value of the parameter of interest between this constrained fit and the unconstrained fit is calculated and plotted. The order of the parameters in the ranking plots is based on their absolute post-fit impact on the parameter of interest. The impact and ranking plots also contain the pull information on another axis.

#### Handling of Normalization of Theory Uncertainties

One of the final goals of this thesis is the measurement of polarization fractions. As shown in Section 9.2.1, the calculation of the fractions depends on both the measured yields across bins as well as the theory predictions for these fractions.

However, in these calculations, both the theory predictions as well as the implicitly measured scaling factors are affected by the theoretical uncertainties introduced in Section 7.3. To counteract this effect, all theory uncertainty templates are normalized to the nominal prediction. A thorough derivation of this can be found in Ref. [249].

## 9.2 Analysis Specific Aspects

The statistical analysis in this thesis has one significant point differentiating it from most other ATLAS analyses as well as another aspect that is different compared to the one discussed in Ref. [249]. The main one is the reparametrization to polarization fractions. This is discussed in Section 9.2.1. The smaller one is the use of smoothing in some fit configurations and is described in Section 9.2.2.

### 9.2.1 Reparametrizations

The main interest when investigating polarization is usually the individual polarization fractions and not the (pseudo) cross-sections of each polarization or the scaling factors $\mu$. Additionally, a global scaling factor that equally affects all polarization states is needed.

While the nominal fractions can be calculated from the post-fit $\mu$ values as the maximum likelihood estimator is invariant under parameter transformation [243], it proves more difficult to transfer the uncertainties correctly. While it is possible to use Gaussian error propagation using the correlation matrix from Hesse, this requires the uncertainties to be symmetrized. This can have particularly strong impacts when the $\mu$ parameters are small and their uncertainty is large.

For this reason, the likelihood is reparametrized before the fit to estimate the fractions and their uncertainties directly. This is done as shown in Refs. [259, 260] by expressing the $\mu$ parameters as functions of the new desired parameters of interest.

For a fit where each polarization has its own set of scaling factors, this would mean the transformation

$$(\mu, \mu_{TT}, \mu_{0T}, \mu_{T0}) \rightarrow (f_{00}, f_{TT}, f_{0T}, f_{T0}, \mu_{WZ}).$$

However, the new set of parameters has more parameters than the original and is thus over-specified. Knowing that the fractions have to sum to one, it is possible to drop one, for example, $f_{TT}$ of them and replace it with

$$f_{TT} = 1 - (f_{00} + f_{0T} + f_{T0}). \tag{9.25}$$

This gets rid of the redundancy and allows for a reparametrization using the following formulas

$$\mu = \frac{f_{00} \times N_{tot}}{N^{MC}_{Template,00}}, \tag{9.26}$$

$$\mu_{0T} = \frac{f_{0T} \times N_{tot}}{N^{MC}_{Template,0T}}, \tag{9.27}$$

$$\mu_{T0} = \frac{f_{T0} \times N_{tot}}{N^{MC}_{Template,T0}}, \tag{9.28}$$

$$\mu_{TT} = \frac{(1 - f_{00} - f_{0T} - f_{T0}) \times N_{tot}}{N^{MC}_{Template,TT}}. \tag{9.29}$$

where $N^{MC}_{Template,i}, i \in \{00, 0T, T0, TT\}$, is the integral over the respective polarizations Monte Carlo template. Here, $\mu_{WZ}$ was swapped out for $N_{tot}$, which serves the same role but makes for an easier parametrization. They have the relationship

$$\mu_{WZ} = \frac{N_{tot}}{\sum_i N^{MC}_{Template,i}}. \tag{9.30}$$

Their pre-fit values are set according to

$$N_{tot} = \sum_i (N^{MC}_{Template,i} \times \mu_i), \tag{9.31}$$

$$f_{00} = \frac{N^{MC}_{Template,00} \times \mu}{N_{tot}}, \tag{9.32}$$

$$f_{0T} = \frac{N^{MC}_{Template,0T} \times \mu_{0T}}{N_{tot}}, \tag{9.33}$$

$$f_{T0} = \frac{N^{MC}_{Template,T0} \times \mu_{T0}}{N_{tot}}, \tag{9.34}$$

$$f_{TT} = \frac{N^{MC}_{Template,TT} \times \mu_{TT}}{N_{tot}}, \tag{9.35}$$

$$\tag{9.36}$$

with all $\mu$'s set to their nominal pre-fit value of one.

Applying a reparametrization like this also has one major disadvantage. Using the regular parametrization with $\mu$'s, one can set their lower boundary to zero. With that, the implicit fractions are all constrained to be in the range $[0, 1]$ and sum to one.

After swapping to the fraction parametrization, the sum to one criteria is explicitly specified when removing the redundant fraction. However, it is now only possible to specify the validity range for those fractions that are explicit parameters in the likelihood. It is not possible anymore to guarantee it for the implicit, eliminated one. The upper bound is still kept as the explicit fractions can be given lower bounds of zero. This is not the case for the lower bound anymore. For example, all the explicit fractions could take a value of 0.8, leaving the implicit one with a value of

$$f_{TT} = 1 - 0.8 - 0.8 - 0.8 = -1.4. \tag{9.37}$$

This can potentially be rectified by adding additional constraints to the likelihood that penalize such parameter values. However, it was found in practice to make the fits less stable while still never being able to enforce those bounds completely. As such, all subsequently used main fit configurations always eliminate the fraction corresponding to the background polarization with the highest expected value and, thus, the lowest expected probability to become negative, even within uncertainties. An example fit demonstrating this issue is shown in Appendix E by instead eliminating a polarization state with a smaller expected value and larger uncertainties.

### 9.2.2 Smoothing

One common aspect shared between all fit configurations is the fact that a lot of the experimental uncertainties are affected by MC statistical fluctuations. This holds especially true for the Mixed polarizations. In fits to data, these fluctuations can cause spurious pulls or impacts of an uncertainty that is completely decoupled from its physical impacts and purely due to the MC statistical fluctuations. To avoid this, a smoothing procedure can applied to experimental uncertainties that can be affected by this. The smoothing is applied independently to the up and down variations of each uncertainty.

The basic steps are as follows:

1. Build a histogram representing the ratio of Variation/Nominal.
2. Smooth the histogram, in our case using ROOT's TH1.Smooth(1)
3. Multiply the ratio back onto the nominal histogram to retrieve a smoothed variation.

It is verified that the procedure only has a small impact on the templates by inspecting the distributions as well as performing Chi-squared and KolmogorovSmirnov tests to ensure that the variation histogram is only changed within its uncertainties. Additionally, the pull and ranking plots, as well as the final fit results, are compared with and without the application of smoothing to ensure that the impact on the results is small and within expectation.

However, smoothing also has downsides. The first one is that it makes it difficult to interpret the result in a pure frequentist way due to the additional ad-hoc information introduced into the fit procedure. As shown in Ref. [261], different smoothing procedures can lead to different results. That can make it challenging to motivate any particular choice, especially when they can arbitrarily reduce systematic uncertainties. Due to this, both approaches are investigated and compared.

## 9.3 Fitting Configurations

This section introduces the dimensions of the space of fit configurations that are used in this work, as well as the motivation to use each of them. First, the different tested parametrizations are introduced in Section 9.3.1, and the fitting regions are summarized in Section 9.3.2. Next, the configurations with different numbers of free parameters are discussed in Section 9.3.3. Section 9.3.4 summarizes the smoothing variations used. Finally, the levels at which the to-be-measured fractions are explored are described in Section 9.3.5.

Fits are then performed with each combination of the settings mentioned here, effectively performing a grid scan over all configurations to be able to assert the consistency of the observed effects of each setting. However, the results presented in Chapter 10 and the appendix only discuss one specific selection of the remaining settings for each dimension of the grid for brevity. All other combinations were checked to be consistent with those presented results.

### 9.3.1 Parametrization

Section 9.2.1 introduced how the likelihood can be reparametrized to polarization fractions. However, it also mentioned some downsides. To ensure that these are not realized in this analysis, additional fits are performed with the standard $\mu$ parametrization and compared against the fraction fits in terms of nominal results and significance.

This means that both of the parametrizations shown in Section 9.2.1 are used

$$(\mu, \mu_{TT}, \mu_{0T}, \mu_{T0}), \tag{9.38}$$

$$(f_{00}, f_{T0}, f_{0T}, N_{tot}). \tag{9.39}$$

For the cross-check, the $\mu$ parameters are calculated through the replacement formulas defining the parametrization and the nominal fit results. The calculation of their uncertainties also uses these formulas. They are then calculated using standard error propagation, taking into account the uncertainties of the actual fit parameters and their correlations

$$\Delta\mu(\phi) = \sqrt{\sum_{\phi_i, \phi_j \in \phi} \mu^{\mathrm{unc}_i}(\phi) C_{ij} \mu^{\mathrm{unc}_j}(\phi)}, \tag{9.40}$$

where $C$ is the correlation matrix as determined by Hesse and

$$\mu^{\mathrm{unc}_i}(\phi) = \frac{\left|\mu(\phi, \phi_i = \hat{\phi}_i + \Delta\hat{\phi}_i^{up}) - \mu_{\mathrm{nom}}(\phi)\right| + \left|\mu(\phi, \phi_i = \hat{\phi}_i - \Delta\hat{\phi}_i^{down}) - \mu_{\mathrm{nom}}(\phi)\right|}{2}. \tag{9.41}$$

#### Boson Polarization Independence

Joint $W^{\pm}Z$ polarizations have already been measured with the ATLAS experiment [26, 64]. This work separates itself from those by measuring the fractions for high values and as a function of $p_T(Z)$. It also has another related advantage. The previous work measured all four joint polarization states with a significance of more than $5\sigma$. However, those results are largely compatible with each boson's polarization being independent of the others. This means that similar results could be achieved with solely a single polarization measurement as already done in Ref. [24] and also repeated in the references above. As such, these measurements can not really be considered to measure the polarization fractions of interacting $W$ and $Z$ bosons. To quantify the small observed deviation, a transformation to a set of parameters better suited to investigate this is performed in Ref. [64]. There, the observed significance with respect to the independence hypothesis was measured to be $1.6\sigma$.

The same approach is followed in this thesis by swapping to the following set of parameters

$$(\mu, \mu_{TT}, \mu_{0T}, \mu_{T0}) \to (R_c, f_{W0}, f_{Z0}, N_{tot}), \tag{9.42}$$

with the transformations defined as

$$\mu = \frac{N_{tot} \times f_{W0} \times f_{Z0} \times R_c}{N^{MC}_{Template,00}}, \tag{9.43}$$

$$\mu_{0T} = \frac{N_{tot} \times f_{W0} \times (1 - (R_c \times f_{Z0}))}{N^{MC}_{Template,0T}}, \tag{9.44}$$

$$\mu_{T0} = \frac{N_{tot} \times f_{Z0} \times (1 - (R_c \times f_{W0}))}{N^{MC}_{Template,T0}}, \tag{9.45}$$

$$\mu_{TT} = \frac{N_{tot} \times (1 - f_{Z0}) \times (1 - f_{W0}) + (R_c - 1) \times f_{W0} \times f_{Z0}}{N^{MC}_{Template,TT}}, \tag{9.46}$$

where $R_c$, representing,

$$R_c = \frac{f_{00}}{f_{W0} \times f_{Z0}} \tag{9.47}$$

gives a measure of the (in)dependence between the longitudinal polarizations of the two bosons. The other relations of $f_{W0}$ and $f_{Z0}$ to the previous fraction parametrization is given by

$$f_{W0} = f_{0T} + f_{00}, \tag{9.48}$$

$$f_{Z0} = f_{T0} + f_{00}. \tag{9.49}$$

This parametrization respects the fact that the sum of fractions has to be one. However, it also suffers the same drawbacks as the direct fraction parametrization in that the implicitly defined fractions can become negative. This and some more properties of this parametrization are shown in Appendix F.

For the fits using this parametrization, the background-only hypothesis is characterized by a fixed value of $R_c = 1$. The test statistic in Equation 9.10 is also modified slightly, such that it is not set to zero for values of $R_c$ below one or even zero, as both are still indicative of correlations between the two bosons.

### 9.3.2 Fit Regions

Independent fits are performed in each of the two 00-enriched signal regions described in Section 5.2.2. Additionally Appendix G discusses fits performed in a region requiring just $p_T^Z > 100\,\text{GeV}$.

Uniquely to the fraction parametrization, a configuration is run where both regions are fit simultaneously with independent parameters for $N_{tot}$ and the respective fractions. This is discussed in Appendix H.

### 9.3.3 Number of Free Parameters

The parametrizations shown so far have always had one (effective) scaling parameter per polarization. However, as Chapter 8 has shown, it is currently not possible to differentiate between the $0T$ and $T0$ states using the available variables. This issue is mitigated in Ref. [26, 64] by employing different fit channels based on the two bosons' decay angles. This procedure did not prove fruitful for this analysis as the reduced number of data events made an effective splitting of the $p_T^Z > 200\,\text{GeV}$ region impossible while still maintaining enough bins to make use of the shape and events to be confident that the asymptotic formula would still hold. To be consistent between the regions, it was also not applied in the region with $100 < p_T^Z \leq 200\,\text{GeV}$.

Additionally, their smaller expected fractions compared to the inclusive phase space means they are not expected to be measured with any useful significance. Due to this, fit configurations are used where one combined scaling parameter is used for both of those states. Conversely they are combined and given the scaling factor $\mu_{Mixed}$, also called $\mu_{0T+T0}$. The reparametrization to fractions for this configuration works analogous to that with the full four parameters. Just like in that case, the $f_{TT}$ fraction is the one that gets eliminated using the sum criterion.

For the independence parametrization, the $f_{W0}$ and $f_{Z0}$ parameters are merged to $f_{WZ0}$ representing their average, with the individual ratios within that fixed to the expected SM ones.

Furthermore, even this fit configuration suffers from somewhat similar shapes between the 00 state of interest and the mixed states due to the limited data and MC statistics, and underlying kinematic distributions. Based on the resulting anticorrelations, a third configuration is introduced that combines the mixed and the TT state and gives them a combined scaling factor $\mu_{Rest}$. In the fraction parameterization, the corresponding fraction is eliminated, leaving the fit with the free parameters $f_{00}, N_{tot}$. The result is that the relative fractions of the TT and mixed states are the ones predicted by the SM. This limits the physical applicability of this parametrization to the comparison to BSM models or parameters whose relative couplings to the TT and mixed states match the ones from the SM. The most important class of such models are the ones that couple to the 00 state exclusively.

With the three-parameter configuration already having questionable interpretability for the independence parametrization, the two-parameter one is not used for that selection of parameters.

### 9.3.4 Smoothing

As the smoothing discussed in Section 9.2.2 has both advantages and disadvantages, configurations with and without smoothing are investigated. The option with smoothing applies it to all experimental systematic uncertainties as described in Section 7.1 that start with any of [EG_, EL_EFF_, MUON_, JET_, MET_].

### 9.3.5 Truth and Reconstructed Fractions

The expected and measured fractions, as described in Section 9.2.1, depend on the values used for $N^{MC}_{Template,i}$. If the integral over the respective polarization's fit input MC template is used, as described in that section, then the resulting fractions represent the ones at the reconstruction level in the signal region. This is due to the value effectively representing

$$N^{MC}_{Template,i} = C_i \cdot \sigma_i \cdot L, \tag{9.50}$$

where $C_i$ is the efficiency factor accounting for detector inefficiencies and migrations, $\sigma_i$ is the pseudo-cross-section of the given polarization, and $L$ is the nominal integrated luminosity.

If one desires to measure the fractions directly at the fiducial truth level, then one can do so by eliminating the factor $C_i$ and replace $N^{MC}_{Template,i}$ with $N^{fiducial}_{Template,i}$ as determined by the equation

$$N^{fiducial}_{Template,i} = \frac{N^{MC}_{Template,i}}{C_i}. \tag{9.51}$$

So by simply replacing the values of $N^{MC}_{Template,i}$ with their truth level counterparts, one can directly measure the polarization fractions at the fiducial level. Care has to be taken in this case so that the truth and reconstructed templates are created with a consistent setup. For example, the per-sample EWK k-factors applied at the reconstructed level also have to be applied at the truth level. Otherwise, the relationship between the values is not described by Equation 9.51, and the resulting fractions can be different.

The electroweak k-factors are thus also taken from Table 6.2. Additionally, the impact of the jet multiplicity reweighting is considered using the factors shown in Table 9.2.

**Table 9.2:** Overall reweighting factors on the individual polarizations in the signal regions due to the jet multiplicity reweighting.

| Signal Region | Polarization | Jet Multiplicity Factor |
|---|---|---|
| $100 < p_T^Z \leq 200\,\mathrm{GeV}$ | 00 | 0.984 |
| | 0T | 0.965 |
| | T0 | 0.969 |
| | TT | 0.969 |
| $p_T^Z > 200\,\mathrm{GeV}$ | 00 | 0.976 |
| | 0T | 0.969 |
| | T0 | 0.972 |
| | TT | 0.970 |

The final fiducial event numbers, also including the QCD k-factors, as used in the parametrization, are summarized in Table 9.3.

This same point applies to the independence parametrizations that are also mentioned in Section 9.3.1. However, that is not the case for the $\mu$ parameterizations, as these do not depend on the integral numbers. Consequently, the grid points corresponding to the truth configuration for that parametrization are not run.

**Table 9.3:** Final fiducial event numbers used in the reparametrization to fiducial fraction.

| Signal Region | Polarization | Fiducial event numbers |
|---|---|---|
| $100 < p_T^Z \leq 200\,\mathrm{GeV}$ | 00 | 380.740 |
| | 0T | 298.890 |
| | T0 | 273.742 |
| | TT | 1546.876 |
| $p_T^Z > 200\,\mathrm{GeV}$ | 00 | 77.127 |
| | 0T | 20.569 |
| | T0 | 19.129 |
| | TT | 213.110 |

# Chapter 10

# Fraction Measurement Results

The determination of the fiducial fractions using smoothed distributions with the two- and three-parameter configurations constitute the nominal result of the fraction measurement. The results for the three-parameter configuration are shown in Section 10.1, while those for the two-parameter configuration are presented in Section 10.2.

The following sections show and explain the impacts of the smoothing (10.3) and the use of fiducial instead of reconstructed fractions (10.4). The differences between results with two to four free parameters are investigated in section 10.5. Meanwhile, the equivalence of the fraction and $\mu$ parametrization is demonstrated in section 10.6. Lastly, Section 10.7 describes the results and additional issues present for the independence parametrization.

## 10.1 Fit With Three Unconstrained Parameters

This section investigated the fitting setup using three free parameters. First, the results of fitting the templates to Asimov data are shown, and some control plots are discussed. The Asimov data is constructed by setting all the parameters to their expectation values and using the resulting distribution as pseudo data with respective data statistical uncertainties.

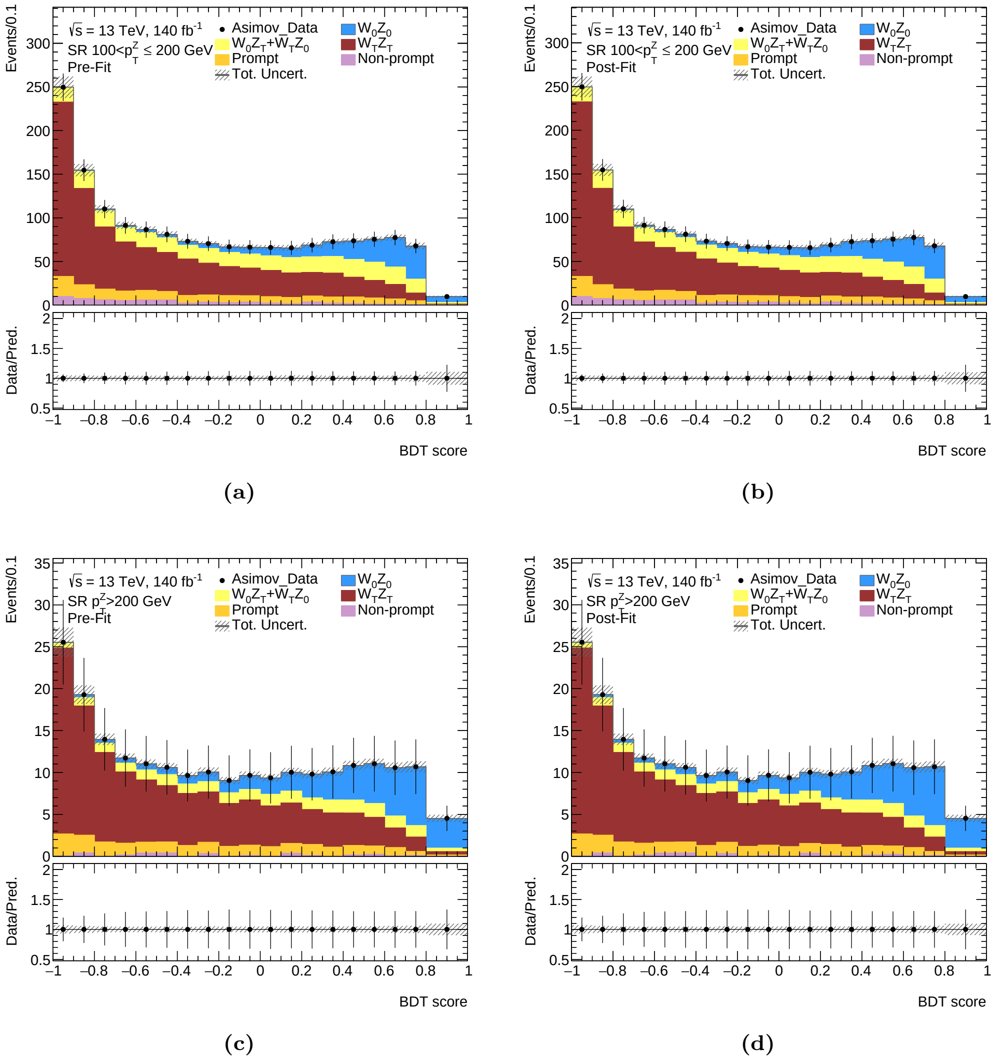
Afterward, the results when fitting to observed data are summarized, and similarities and differences to the Asimov fit are investigated.

### 10.1.1 Asimov Fit

As the analysis was constructed without looking at data, fits to Asimov data are used to gauge the sensitivity of the measurement as well as the stability of the fitting procedure. The respective pre- and post-fit distributions are shown in Figure 10.1. As the Asimov data exactly matches the expectation, only small differences in the uncertainties are expected due to correlations between the different nuisance parameters.

It can also be seen that the distributions of the prompt and non-prompt backgrounds are largely flat across the BDT score with a slight shift towards lower values. Out of those non-WZ backgrounds, the prompt ones dominate in both of the 00-enriched regions. While the relative contribution of the prompt background remains roughly equal across the two regions, this is not the case for the non-prompt one. This reduces from roughly 4 percent in the region with $100 < p_T^Z \leq 200\,\text{GeV}$ to $\approx 1\,\%$ in the one with $p_T^Z > 200\,\text{GeV}$. This can also be seen in Table 10.1, which shows the pre-fit yields.

The expected results with the three-parameter configuration are summarized in Table 10.2. It can be seen that the expected significance is higher in the lower $p_T$ region compared to the higher $p_T$ one. This is despite a lower expected $f_{00}$. The main reason here is the significantly reduced statistical uncertainty, which overcompensates the reduction in the expected fractions. The table also shows that the fraction values are statistically dominated in both regions. Only for $N_{tot}$ in the region with $100 < p_T^Z \leq 200\,\text{GeV}$ do the systematic uncertainties dominate. With both expected significances being below $5\sigma$, no observation of $f_{00}$ is expected in either




(a) (b) (c) (d)

**Figure 10.1:** Expected (a) pre- and (b) post-fit distributions in the 00-enriched $100 < p_T^Z \leq 200\,\text{GeV}$ region and (c) pre- and (d) post-fit distributions in the 00-enriched $p_T^Z > 200\,\text{GeV}$ region for the S+B hypothesis with the three sample strategy.

**Table 10.1:** Number of events for the expected signal and background, and observed data in the two signal regions for the joint polarization fraction measurement. The uncertainties shown include both statistical and systematic contributions.

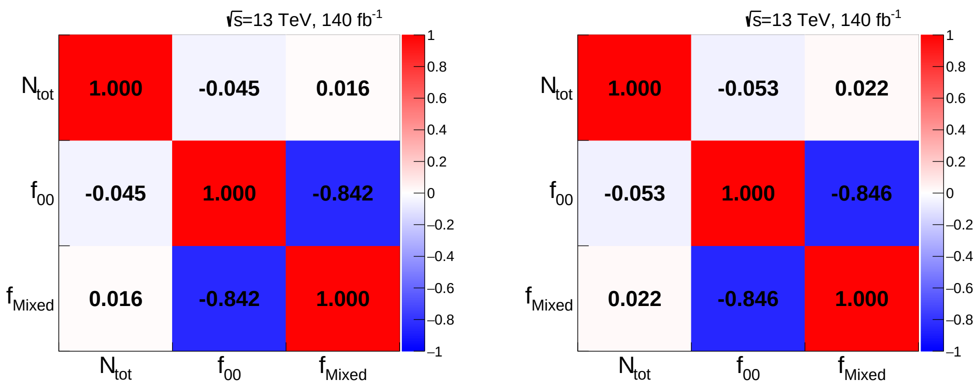
| **Process** | $100 < p_T^Z \leq 200\,\text{GeV}$ | $p_T^Z > 200\,\text{GeV}$ |
|---|---|---|
| $W_0Z_0$ | 222 ± 5 | 47.6 ± 1.5 |
| $W_0Z_T + W_TZ_0$ | 323 ± 12 | 23.7 ± 0.8 |
| $W_TZ_T$ | 856 ± 31 | 124 ± 4 |
| Prompt background | 169 ± 18 | 24.1 ± 2.7 |
| Non-prompt background | 68 ± 29 | 2.8 ± 1.1 |
| Total Expected | 1640 ± 60 | 222 ± 8 |
| Data | 1740 | 236 |

of the two regions. The 00-enriched $p_T^Z > 200\,\text{GeV}$ even has an expected significance below $3\sigma$, indicating that even the level of "evidence" might not be reached for this configuration. The table also shows that $f_{Mixed}$ can not become negative within its uncertainties, while the derived $\mu_{WZMixed}$ can. This is due to the symmetrization that occurs when propagating the uncertainties to the $\mu$'s.

**Table 10.2:** Asimov fit results using the smoothed nuisance parameter distributions in the configuration with three independent fit parameters for the fiducial fraction parametrization.

| 00-enriched region | $100 < p_T^Z \leq 200\,\text{GeV}$ | $p_T^Z > 200\,\text{GeV}$ |
|---|---|---|
| Significance | $4.285\,\sigma$ | $2.493\,\sigma$ |
| $f_{00}$ | $0.152\ \pm^{0.034}_{0.033}$ (stat) $\pm^{0.017}_{0.016}$ (sys) | $0.23\ \pm^{0.10}_{0.10}$ (stat) $\pm^{0.02}_{0.02}$ (sys) |
| $f_{Mixed}$ | $0.23\ \pm^{0.08}_{0.08}$ (stat) $\pm^{0.04}_{0.05}$ (sys) | $0.12\ \pm^{0.19}_{0.12}$ (stat) $\pm^{0.05}_{0.00}$ (sys) |
| $N_{tot}$ | $2500\ \pm^{70}_{70}$ (stat) $\pm^{100}_{100}$ (sys) | $330\ \pm^{26}_{25}$ (stat) $\pm^{14}_{13}$ (sys) |
| $f_{TT}$ | $0.62 \pm 0.05(\text{stat}) \pm 0.03(\text{sys})$ | $0.65 \pm 0.09(\text{stat}) \pm 0.01(\text{sys})$ |
| $\mu$ | $1.00 \pm 0.22(\text{stat}) \pm 0.12(\text{sys})$ | $1.0 \pm 0.4(\text{stat}) \pm 0.1(\text{sys})$ |
| $\mu_{WZTT}$ | $1.00 \pm 0.09(\text{stat}) \pm 0.06(\text{sys})$ | $1.00 \pm 0.16(\text{stat}) \pm 0.05(\text{sys})$ |
| $\mu_{WZMixed}$ | $1.00 \pm 0.34(\text{stat}) \pm 0.19(\text{sys})$ | $1.0 \pm 1.3(\text{stat}) \pm 0.2(\text{sys})$ |

Figures 10.2 and 10.3 show nuisance parameter correlations for the stat. only and full fits in both regions, respectively. The latter only includes nuisance parameters that have a correlation of at least 0.2 with any other NP. It can be seen that for both regions, there is an extremely large anticorrelation between the $f_{00}$ and $f_{Mixed}$ fraction parameters. This is due to both parameters being otherwise unconstrained in the fit and the two respective distributions being closest in shape, as was already shown in Figure 8.10. The expected correlations are $-0.828$ for the lower $p_T^Z$ region and $-0.838$ for the higher $p_T^Z$ one.

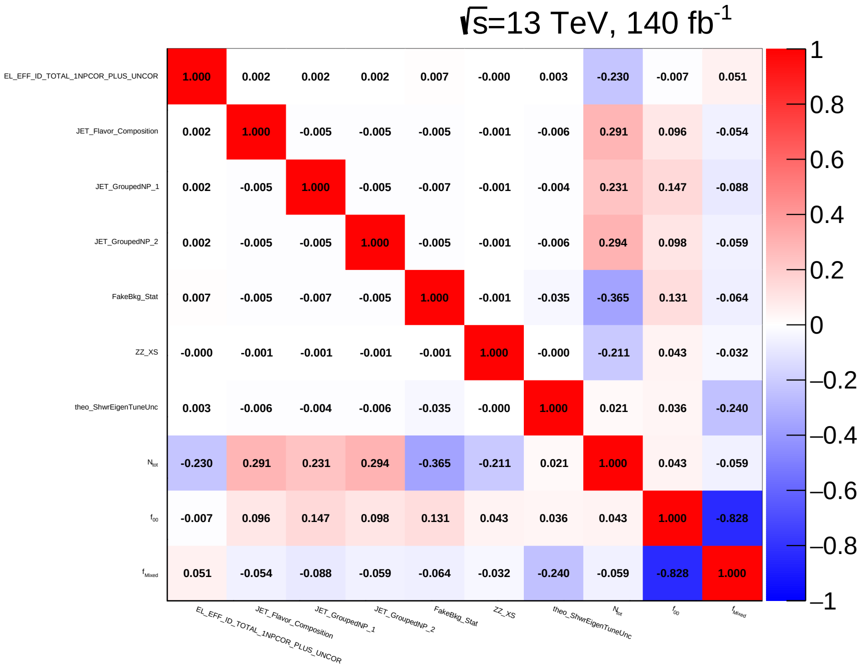



(a) (b)

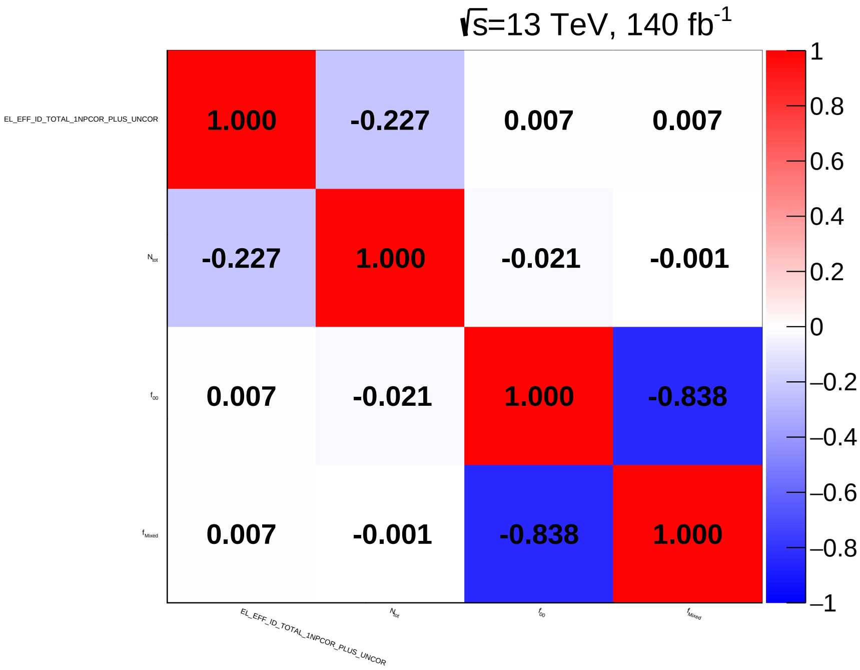
**Figure 10.2:** Expected post-fit correlations for the stat. only fit in the 00-enriched region with (a) $100 < p_T^Z \leq 200\,\text{GeV}$ and (b) $p_T^Z > 200\,\text{GeV}$ for the S+B hypothesis with the three sample strategy.

One interesting point about the correlations for the full fits is that there are significantly more parameters with at least one correlation above 20 % in the $100 < p_T^Z \leq 200\,\text{GeV}$ region. In the region with $p_T^Z > 200\,\text{GeV}$, there is only one parameter besides the unconstrained ones, while the lower $p_T^Z$ region includes seven. The largest correlations of these parameters, and the ones getting them included in the plot, are mostly with $N_{tot}$. As such, it makes sense that fewer parameters are present for the region where the uncertainty of $N_{tot}$ is statistically dominated.

Nuisance parameter rankings by their impact on $f_{00}$, $f_{Mixed}$ and $N_{tot}$ are shown in Figures 10.4 and 10.5. The top-ranked variables for the fractions concern NLO effects, shower, jet, and pileup uncertainties. The impact of the NLO uncertainties is one-sided as the uncertainties themselves are implemented as one-sided. This means the down variations exactly match the nominal prediction, thus having no impact on their value. The relevant uncertainties for $N_{tot}$ are similar, except that the uncertainties on the backgrounds, mainly the non-prompt one, also play a major to dominant role. Especially in the $100 < p_T^Z \leq 200\,\text{GeV}$ region.

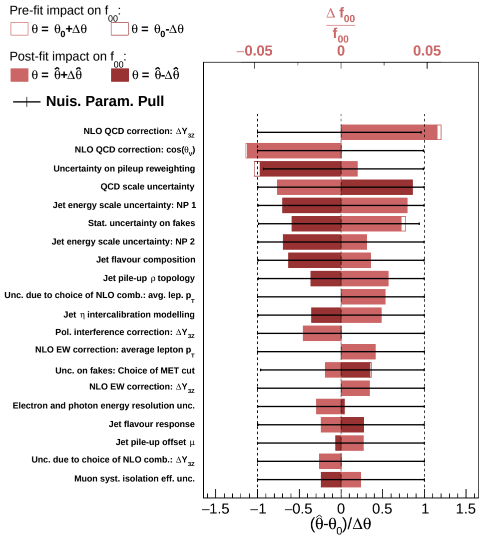



**(a)**

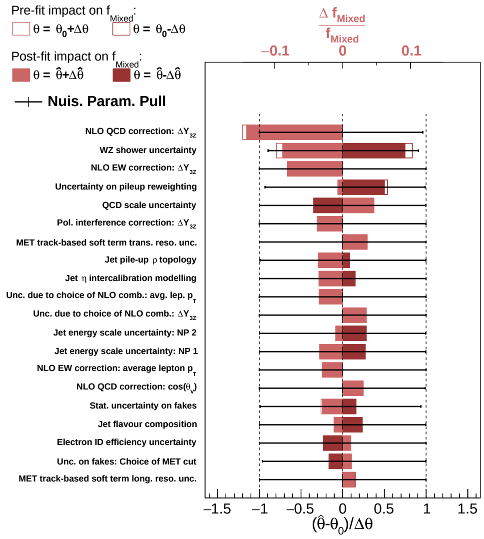



**(b)**

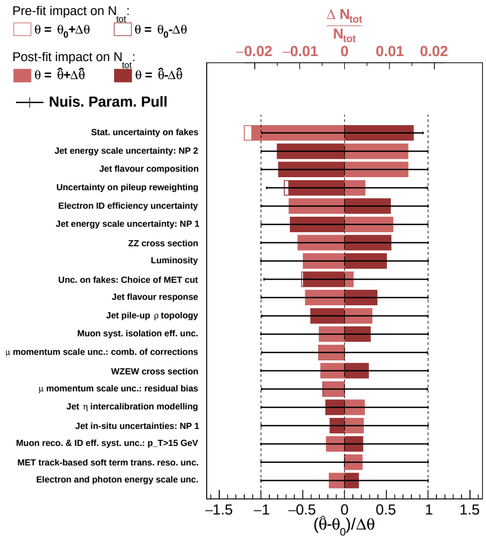
**Figure 10.3:** Expected post-fit correlations of nuisance parameters with a correlation of at least 0.2 with another NP for the full fit in the 00-enriched region with (a) $100 < p_T^Z \leq 200\,\text{GeV}$ and (b) $p_T^Z > 200\,\text{GeV}$ for the S+B hypothesis with the three sample strategy.




**(a)**

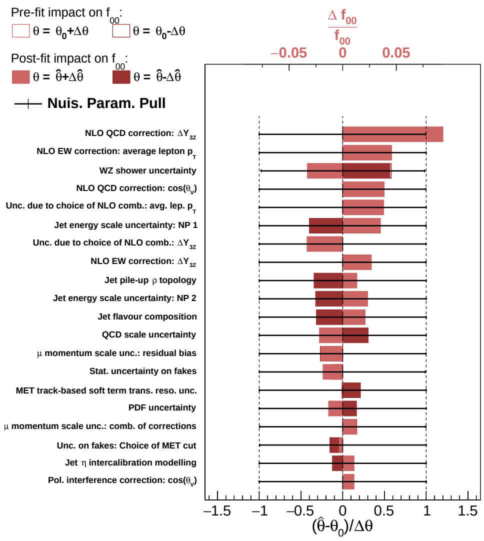



**(b)**

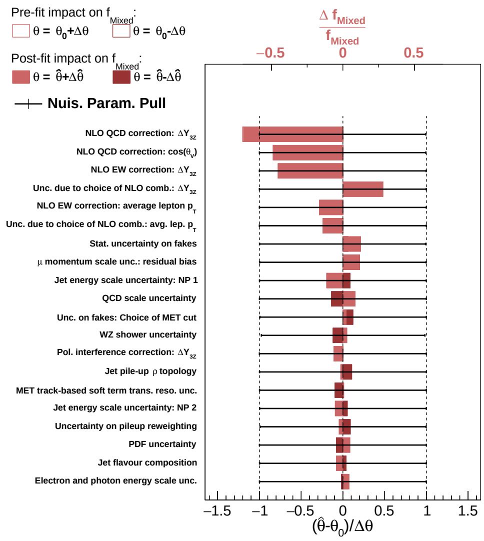



**(c)**

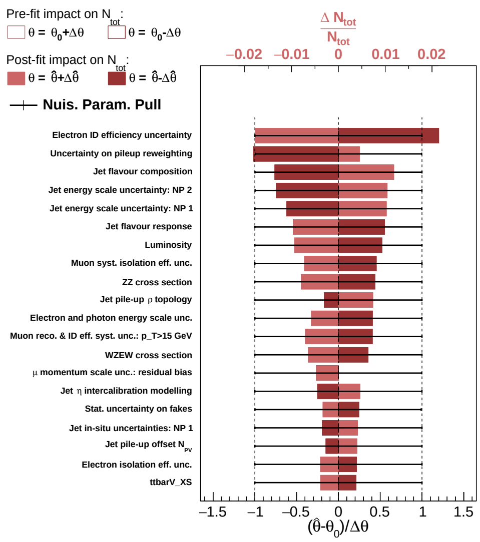
**Figure 10.4:** Ranking of the systematic nuisance parameters included in the fit in the 00-enriched $100 < p_T^Z \leq 200$ GeV region using the three sample strategy according to their impact on (a) $f_{00}$, (b) $f_{Mixed}$ and (c) $N_{tot}$ with Asimov data obtained with the expected background prediction. Only the top 20 parameters are shown. Nuisance parameters corresponding to MC statistical uncertainties are not considered here. The empty brown rectangles correspond to the pre-fit impact on the respective parameter and the filled brown ones to the post-fit impacts, both referring to the upper scale. The impact of each nuisance parameter, $\Delta f_{00}|\Delta f_{Mixed}|\Delta N_{tot}$, is computed by comparing the nominal best-fit value with the result of the fit when fixing the considered nuisance parameter to its best-fit value, $\hat{\theta}$, shifted by its pre-fit (post-fit) uncertainties $\pm\Delta\theta$ ($\pm\Delta\hat{\theta}$). The black points show the pulls of the nuisance parameters with respect to their nominal values, $\theta_0$. These pulls and their relative post-fit errors, $\Delta\hat{\theta}/\Delta\theta$, refer to the lower scale. For experimental uncertainties, which are broken down into several independent sources, the corresponding nuisance parameter (NP) index is reported.




(a)




(b)




(c)

**Figure 10.5:** Ranking of the systematic nuisance parameters included in the fit in the 00-enriched $p_T^Z > 200\,\text{GeV}$ region using the three sample strategy according to their impact on (a) $f_{00}$, (b) $f_{Mixed}$ and (c) $N_{tot}$ with Asimov data obtained with the expected background prediction. Only the top 20 parameters are shown. Nuisance parameters corresponding to MC statistical uncertainties are not considered here. The empty brown rectangles correspond to the pre-fit impact on the respective parameter and the filled brown ones to the post-fit impacts, both referring to the upper scale. The impact of each nuisance parameter, $\Delta f_{00}|\Delta f_{Mixed}|\Delta N_{tot}$, is computed by comparing the nominal best-fit value with the result of the fit when fixing the considered nuisance parameter to its best-fit value, $\hat{\theta}$, shifted by its pre-fit (post-fit) uncertainties $\pm\Delta\theta$ ($\pm\Delta\hat{\theta}$). The black points show the pulls of the nuisance parameters with respect to their nominal values, $\theta_0$. These pulls and their relative post-fit errors, $\Delta\hat{\theta}/\Delta\theta$, refer to the lower scale. For experimental uncertainties, which are broken down into several independent sources, the corresponding nuisance parameter (NP) index is reported.

### 10.1.2 Observed Fit

With the fits to Asimov data showing no concerning issues, the fits to observed data are performed.

The pre- and post-fit distributions for the two 00-enriched regions are shown in Figure 10.6, while the post-fit yields are summarized in Table 10.3. Generally, a good agreement can be observed between predictions and data, both pre- and post-fit and in both regions. Some larger deviations can be observed in the $p_T^Z > 200$ GeV region. However, these are still within the sizeable statistical uncertainties of the data.

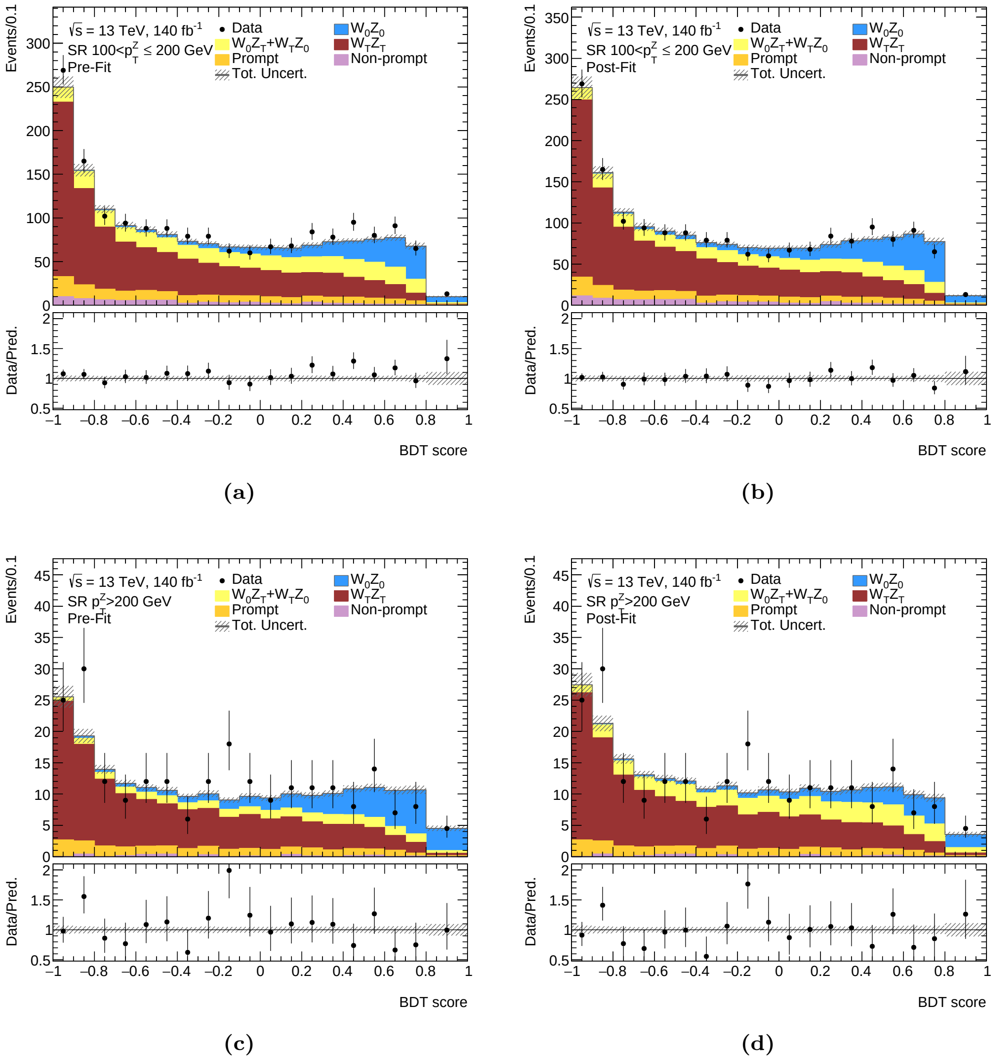


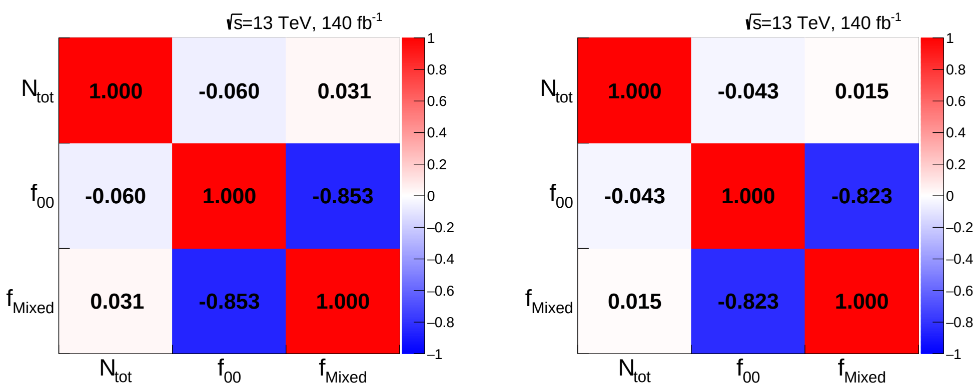

**Figure 10.6:** Observed (a) pre- and (b) post-fit distributions in the 00-enriched $100 < p_T^Z \leq 200$ GeV region and (c) pre- and (d) post-fit distributions in the 00-enriched $p_T^Z > 200$ GeV region for the S+B hypothesis with the three sample strategy.

Correlations between the parameters in the stat. only and full fits are shown in Figures 10.7 and 10.8. Similar to the case of the Asimov fit, the stat. only correlations are dominated by a very strong anti-correlation between the 00 and mixed fractions. This time the values of the correlations are $-0.835$ and $-0.818$ for the low and high $p_T^Z$ regions. Some small differences can be observed in the correlations of the full fits. Just like in the Asimov case, there are more nuisance parameters with correlations larger than 0.2 in the lower $p_T^Z$ region than in

**Table 10.3:** Observed post-fit yields in the configuration with three free parameters. Shown are the nominal values and statistical, theory, and experimental uncertainties.

| | $100 < p_T^Z \leq 200$ GeV | $p_T^Z > 200$ GeV |
|---|---|---|
| $W_0 Z_0$ | $290 \pm 50 \pm 0 \pm 10$ | $28 \pm 18 \pm 0 \pm 1$ |
| $W_0 Z_T + W_T Z_0$ | $280 \pm 110 \pm 0 \pm 10$ | $50 \pm 40 \pm 0 \pm 0$ |
| $W_T Z_T$ | $920 \pm 80 \pm 0 \pm 30$ | $132 \pm 26 \pm 1 \pm 4$ |
| Prompt background | $166 \pm 0 \pm 17 \pm 8$ | $24.2 \pm 0.0 \pm 2.1 \pm 1.6$ |
| Non-prompt background | $80 \pm 0 \pm 0 \pm 40$ | $2.8 \pm 0.0 \pm 0.0 \pm 1.0$ |
| Total Expected | $1740 \pm 40 \pm 20 \pm 60$ | $236 \pm 15 \pm 3 \pm 8$ |
| Data | $1740.0 \pm 0.0 \pm 0.0 \pm 0.0$ | $236.0 \pm 0.0 \pm 0.0 \pm 0.0$ |

the higher one. For the $100 < p_T^Z \leq 200$ GeV region, the majority of the included nuisance parameters are the same. The most significant addition is the one for the QCD uncertainty from the distribution of $\Delta Y(\ell_W Z)$, through its high correlation with $f_{00}$ and $f_{Mixed}$. This is included in both fit regions.



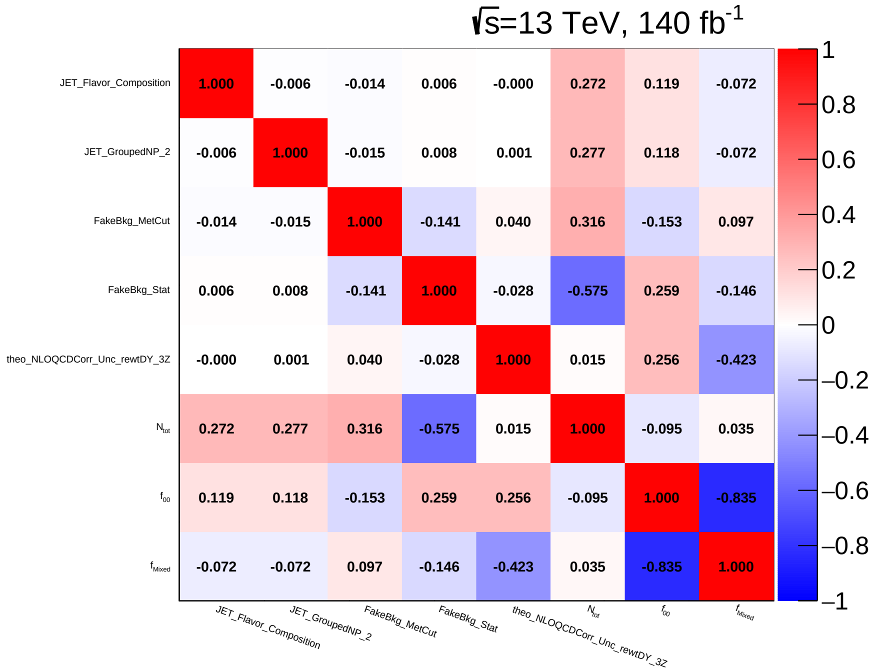

**(a)** **(b)**

**Figure 10.7:** Observed post-fit correlations for the stat. only fit in the 00-enriched region with (a) $100 < p_T^Z \leq 200$ GeV and (b) $p_T^Z > 200$ GeV for the S+B hypothesis with the three sample strategy.

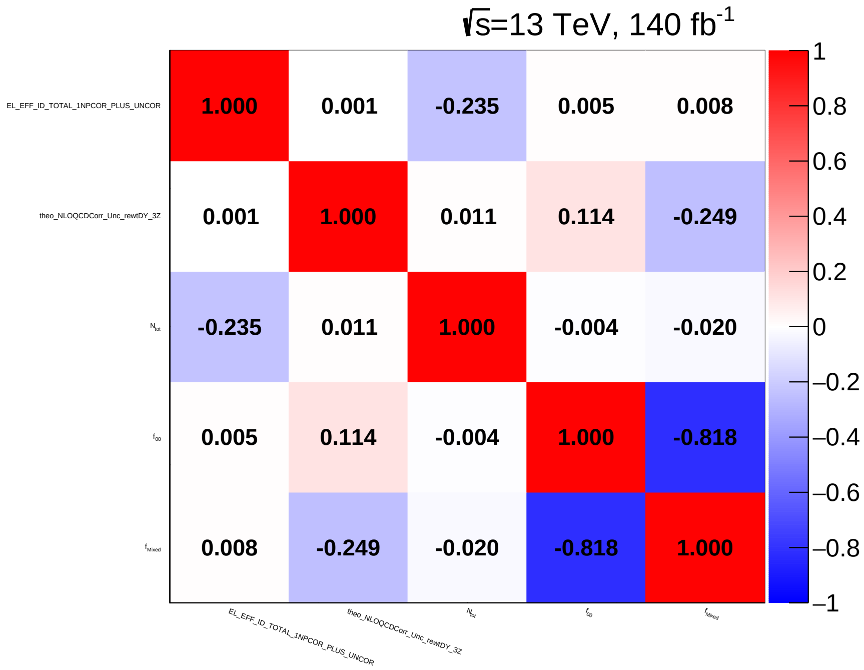
This inclusion can partially be understood by considering the nuisance parameters pull as shown in Figures 10.9 and 10.10. These figures show the 15 nuisance parameters with the strongest expected and observed constraints and the strongest observed pulls for the two 00-enriched regions. The strongest pulls are not shown for the Asimov fit, as they are all zero there. The main relevant point is the inclusion of this uncertainty within the top 3 highest-pulled ones. As all of the NLO uncertainties are implemented as one-sided uncertainties by setting their down variation equal to the nominal one, their correlations as their nominal value are reduced because down variations have no impact. A pull now allows both up and down variations to cause different fit configurations, increasing the observed correlations.

Another interesting point is the lack of the `PRW_DATASF` parameter in the correlations plots despite it being the highest pulled parameter, indicating a strong impact on the fit parameters. This is mostly because it has effects on both the $N_{tot}$ parameter as well as the fractions. This causes mild individual correlations, which lead to an increased pull of the parameter.

The post-fit values for the unconstrained fit parameters are shown alongside their uncer-

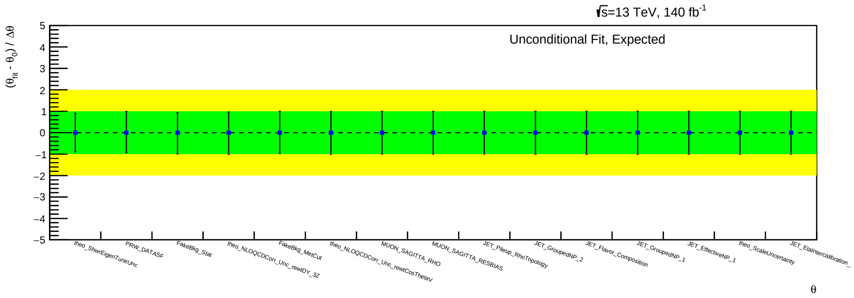



**(a)**

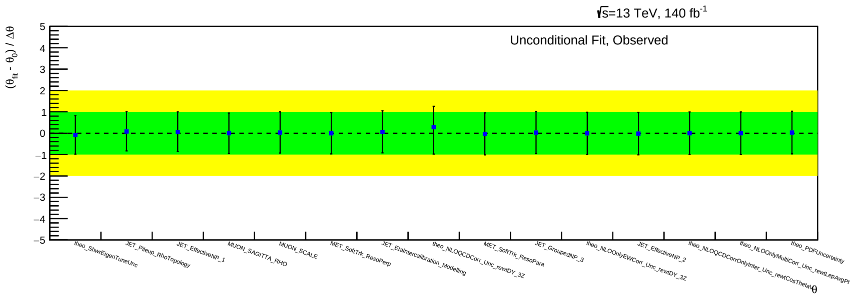



**(b)**

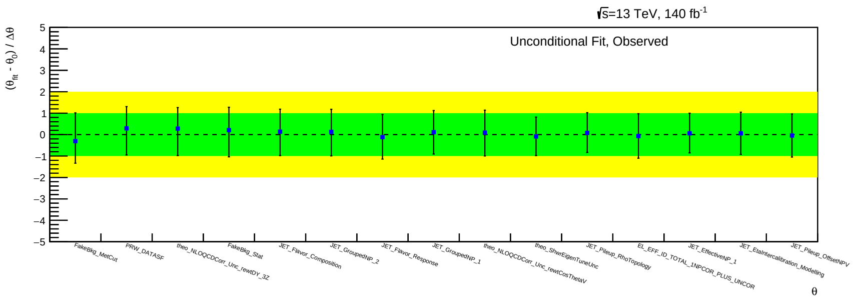
**Figure 10.8:** Observed post-fit correlations of nuisance parameters with a correlation of at least 0.2 with another NP for the full fit in the 00-enriched region with (a) $100 < p_T^Z \leq 200\,\text{GeV}$ and (b) $p_T^Z > 200\,\text{GeV}$ for the S+B hypothesis with the three sample strategy.




**(a)**

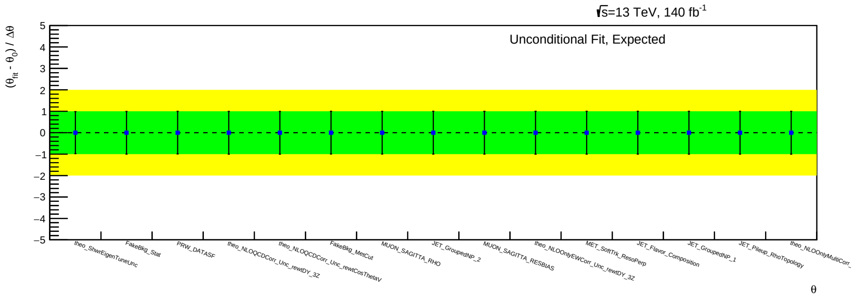



**(b)**

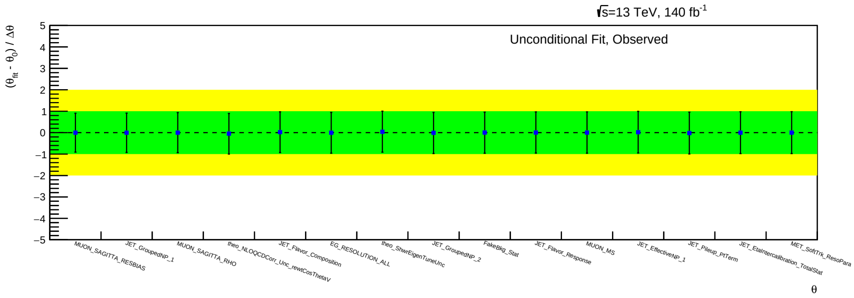



**(c)**

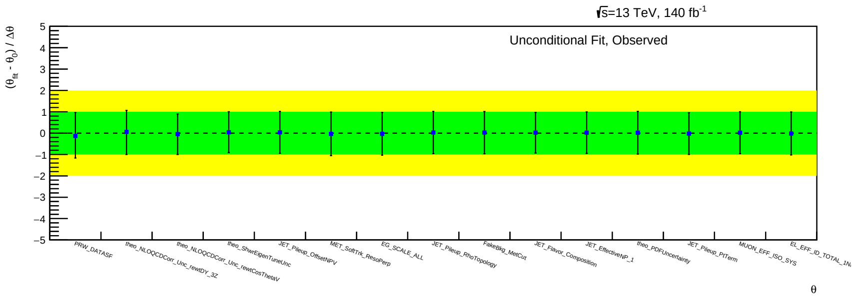
**Figure 10.9:** Post fit nuisance parameter pulls for the fit of the S+B hypothesis in the 00-enriched $100 < p_T^Z \leq 200\,\text{GeV}$ with the three sample strategy. Shown are (a) the most constrained parameters in the fit to Asimov data, (b) the most constrained and (c) the most pulled parameters in the fit to observed data.




**(a)**




**(b)**




**(c)**

**Figure 10.10:** Post fit nuisance parameter pulls for the fit of the S+B hypothesis in the 00-enriched $p_T^Z > 200\,\text{GeV}$ with the three sample strategy. Shown are (a) the most constrained parameters in the fit to Asimov data, (b) the most constrained and (c) the most pulled parameters in the fit to observed data.

tainties in Table 10.4. As already visible by comparing the pre- and post-fit distribution, it can be seen that the measured $f_{00}$ fraction in the $100 < p_T^Z \leq 200\,\text{GeV}$ region is larger than the expected one, while the opposite is the case for the region with $p_T^Z > 200\,\text{GeV}$. With only small changes in the uncertainties, this subsequently results in higher and smaller values for the significances in the two regions. The final measured values are $5.2\sigma$ in the lower $p_T$ region and $1.6\sigma$ in the higher $p_T$ one. With this, a non-zero fraction of doubly longitudinal events in a phase space with $p_T^{WZ} < 70\,\text{GeV}$ and $100 < p_T^Z \leq 200\,\text{GeV}$ has been observed. The same can not be said for the region with $p_T^Z > 200\,\text{GeV}$ where even the $3\sigma$ threshold required to claim “evidence” has not been reached.

The final value for $f_{00}$ in the region with $100 < p_T^Z \leq 200\,\text{GeV}$ with the three-parameter configuration is:

$$f_{00} = 0.189 \pm^{0.034}_{0.033}\,(\text{stat.}) \pm^{0.016}_{0.016}\,(\text{mod. syst.}) \pm^{0.018}_{0.015}\,(\text{exp. syst.}) \pm^{0.000}_{0.000}\,(\text{lumi.}) \tag{10.1}$$

in the region with $p_T^Z > 200\,\text{GeV}$ the value is:

$$f_{00} = 0.13 \pm^{0.09}_{0.08}\,(\text{stat.}) \pm^{0.02}_{0.02}\,(\text{mod. syst.}) \pm^{0.01}_{0.01}\,(\text{exp. syst.}) \pm^{0.00}_{0.00}\,(\text{lumi.}) \tag{10.2}$$

For both of these cases, the expected values are almost covered by the $1\sigma$ uncertainties. However, the deviations are in opposite directions. This results in a lower measured 00 fraction in the region with high $p_T^Z$ when the opposite was expected. With neither difference being statistically significant, further studies are needed to discern whether these are just fluctuations or could be a hint towards new physics.

**Table 10.4:** Observed fit results using the smoothed nuisance parameter distributions in the configuration with three independent fit parameters for the fiducial fraction parametrization.

| 00-enriched region | $100 < p_T^Z \leq 200\,\text{GeV}$ | $p_T^Z > 200\,\text{GeV}$ |
|---|---|---|
| Significance | $5.158\,\sigma$ | $1.595\,\sigma$ |
| $f_{00}$ | $0.189 \pm^{0.034}_{0.033}$ (stat) $\pm^{0.024}_{0.022}$ (sys) | $0.13 \pm^{0.09}_{0.08}$ (stat) $\pm^{0.02}_{0.02}$ (sys) |
| $f_{Mixed}$ | $0.18 \pm^{0.07}_{0.08}$ (stat) $\pm^{0.05}_{0.06}$ (sys) | $0.23 \pm^{0.17}_{0.18}$ (stat) $\pm^{0.06}_{0.10}$ (sys) |
| $N_{tot}$ | $2690 \pm^{80}_{70}$ (stat) $\pm^{130}_{120}$ (sys) | $355 \pm^{27}_{26}$ (stat) $\pm^{15}_{14}$ (sys) |
| $f_{TT}$ | $0.63 \pm 0.05$(stat) $\pm 0.04$(sys) | $0.64 \pm 0.12$(stat) $\pm 0.06$(sys) |
| $\mu$ | $1.33 \pm 0.24$(stat) $\pm 0.17$(sys) | $0.6 \pm 0.4$(stat) $\pm 0.1$(sys) |
| $\mu_{WZTT}$ | $1.09 \pm 0.09$(stat) $\pm 0.08$(sys) | $1.06 \pm 0.21$(stat) $\pm 0.11$(sys) |
| $\mu_{WZMixed}$ | $0.9 \pm 0.4$(stat) $\pm 0.3$(sys) | $2.1 \pm 1.6$(stat) $\pm 0.7$(sys) |

A breakdown of the uncertainties on $f_{00}$ in the two 00-enriched regions is shown in Table 10.5. The dominant uncertainty in both regions is the data statistical uncertainty. In the region with $p_T^Z > 200\,\text{GeV}$ it reaches a total of 64.5 %, while staying at a lower 18.0 % in the lower $p_T^Z$ region. Most uncertainties increase when going from the lower to the higher $p_T^Z$ region. One major exception is the uncertainty of the non-prompt background estimation, which decreases from 5.8 % to 0.8 % due to the reduced overall size of the related background. Another uncertainty with this effect, though smaller, is the one concerning interference effects between the different polarization states. The overall dominant systematic uncertainties in the region with $100 < p_T^Z \leq 200\,\text{GeV}$ come from the jet energy scale and resolution, the non-prompt background estimation, NLO QCD effects, and PDF, scale, and shower settings. In the $p_T^Z > 200\,\text{GeV}$ region, the jets, NLO QCD, and other theory uncertainties remain the dominant ones. Additionally, the MC statistical and other NLO effects become relevant. The impact of the uncertainty on the luminosity measurement only amounts to $0.1 - 0.2\,\%$. This is less than the 0.83 % from the actual measurement of the luminosity as mentioned in Section 4.2. This is one of the effects of the parametrization to fractions. The general impact of the luminosity uncertainty is absorbed by the $N_{tot}$ parameter, while the fractions are only impacted by the effective shape changes of the combined prompt and non-prompt background.

Likelihood scans of selected parameters are shown in Figures 10.11 and 10.12 for the $100 < p_T^Z \leq 200\,\text{GeV}$ and $p_T^Z > 200\,\text{GeV}$ regions, respectively. None of the scans show unexpected behavior, such as local minima or strong deviations from the expected parabolic shape. Nuisance parameter rankings by their impact on $f_{00}$, $f_{Mixed}$ and $N_{tot}$ are shown in Figures 10.13 and 10.14. The top-ranked variables match those expected from the systematic breakdown tables, with NLO, shower, and jet uncertainties being the most impactful on the fractions. The only deviation from this expectation is the pileup reweighting uncertainty in the region with $100 < p_T^Z \leq 200\,\text{GeV}$ which has an impact $\approx 5\,\%$ in the ranking plot but only $1.6\,\%$ in the breakdown table. This is due to two reasons. The first is the symmetrization of the uncertainties for the table. The second is due to the approach chosen to calculate the uncertainty breakdowns in the table as described in section 9.1.4. The uncertainties of contributions that are calculated early in the process and have similar impacts as ones calculated later are reduced. For $N_{tot}$, the most important nuisance parameters are related to jet and electron uncertainties, pileup, luminosity, and the non-prompt background. The parameters are also largely the same and in the same order as in the case of the fit to Asimov data. The total relative impact of the parameters on the fractions, however, changes significantly. This is especially the case for the $p_T^Z > 200\,\text{GeV}$ region and is mostly due to the change in the nominal value. Here, an increased value causes lower relative uncertainties, while the effect is the opposite for decreased values as seen for $f_{00}$. The NLO uncertainties are, again, mostly one-sided. They are not completely one-sided, as seen in the fits to Asimov data. This is because the pulls on these parameters cause their best-fit values to differ from the nominal value. Due to this, the down variation, still matching the nominal shape, is not equal to the best-fit one.

**Table 10.5:** Breakdown of uncertainties on $f_{00}$ in the fits to observed data with the three-parameter configuration.

| Source | Impact on $f_{00}$ [%] | |
|---|---|---|
| Experimental | $100 < p_T^Z \leq 200\,\text{GeV}$ | $p_T^Z > 200\,\text{GeV}$ |
| Luminosity | 0.1 | 0.2 |
| Electron calibration | 1.0 | 0.9 |
| Muon calibration | 1.1 | 1.3 |
| Jet energy scale and resolution | 5.9 | 9.0 |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 1.0 | 0.6 |
| Flavor-tagging inefficiency | 0.1 | 0.2 |
| Pileup modelling | 1.6 | 1.1 |
| Non-prompt background estimation | 5.8 | 0.8 |
| Modelling | | |
| Background, other | 1.4 | 1.6 |
| Model statistical | 2.5 | 5.6 |
| NLO QCD effects | 6.8 | 8.2 |
| NLO EW effects | 1.1 | 3.3 |
| Effect of additive vs multiplicative QCD+EW combination | 1.3 | 3.8 |
| Interference impact | 1.4 | 0.7 |
| PDF, Scales, and shower settings | 3.5 | 9.2 |
| Experimental and modelling | 12.1 | 17.7 |
| Data statistical | 18.0 | 64.5 |
| Total | 21.7 | 66.9 |

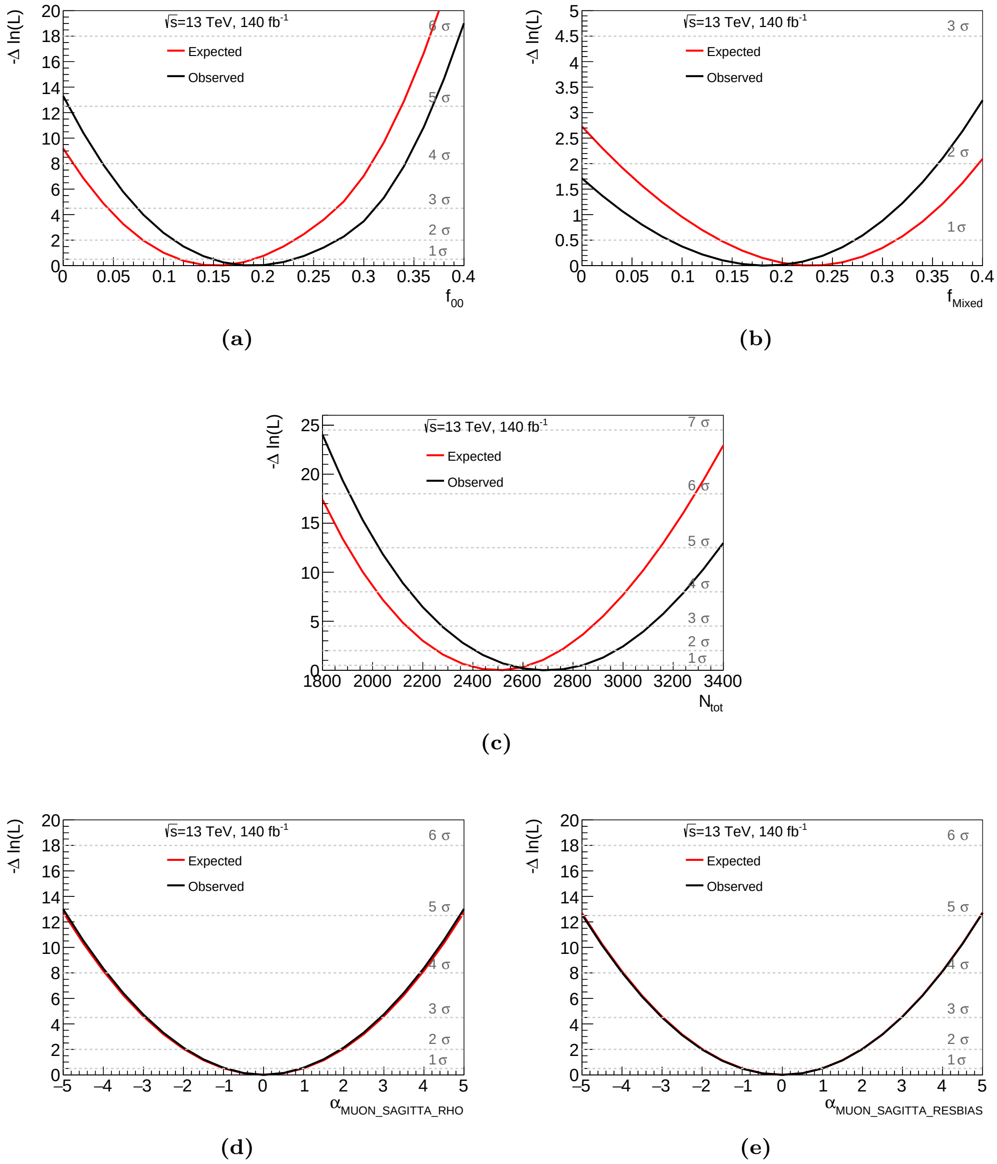
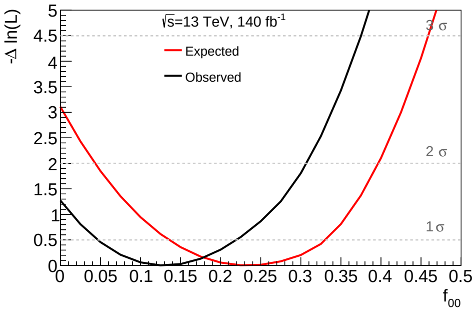
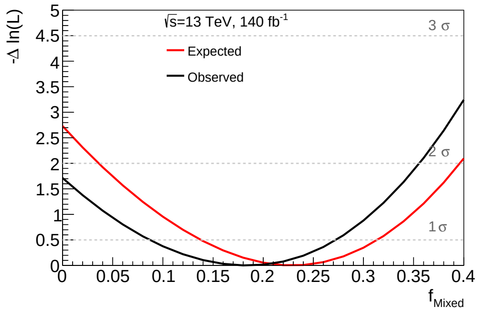


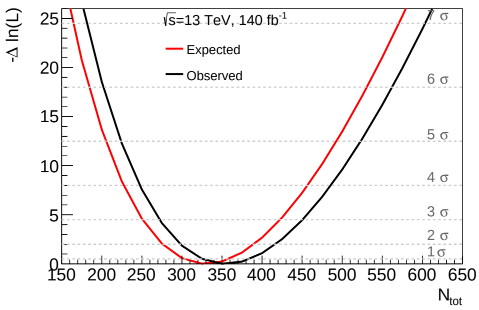

**Figure 10.11:** Likelihood scans for the full fits of the S+B hypothesis in the 00-enriched $100 < p_T^Z \leq 200\,\mathrm{GeV}$ region with the three sample strategy for the (a) $f_{00}$, (b) $f_{Mixed}$, (c) $N_{tot}$, (d) `MUON_SAGITTA_RHO` and (e) `MUON_SAGITTA_RESBIAS` parameters.



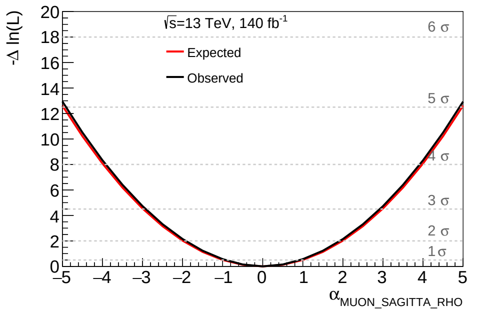

**(a)**



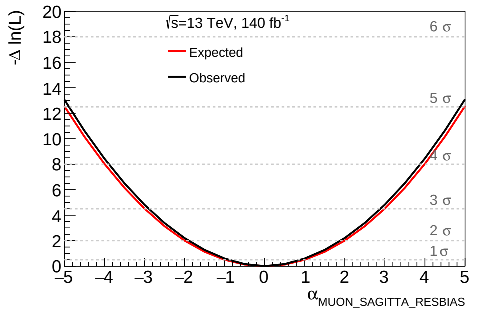

**(b)**




**(c)**




**(d)**




**(e)**

**Figure 10.12:** Likelihood scans for the full fits of the S+B hypothesis in the 00-enriched $p_T^Z > 200\,\text{GeV}$ region with the three sample strategy for the (a) $f_{00}$, (b) $f_{Mixed}$, (c) $N_{tot}$, (d) `MUON_SAGITTA_RHO` and (e) `MUON_SAGITTA_RESBIAS` parameters.

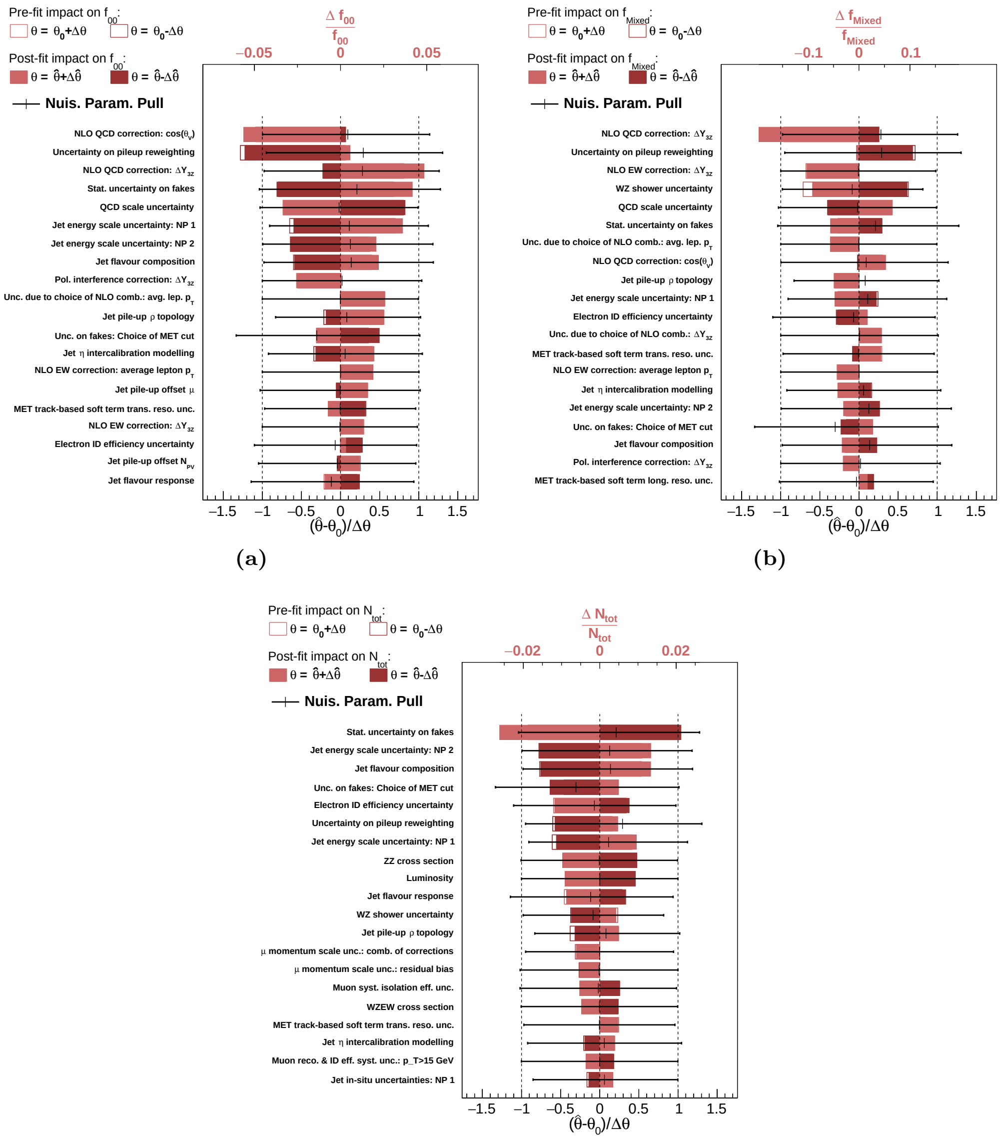
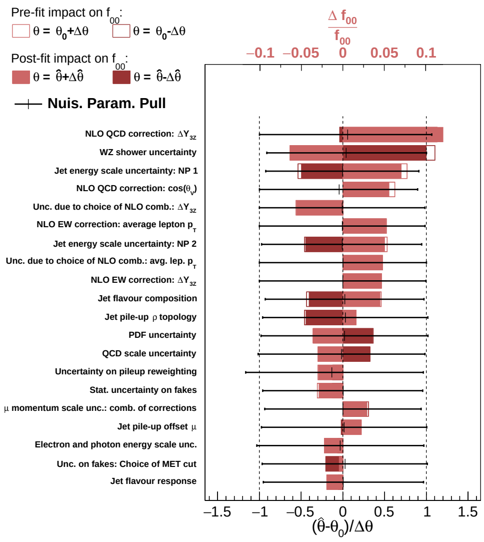
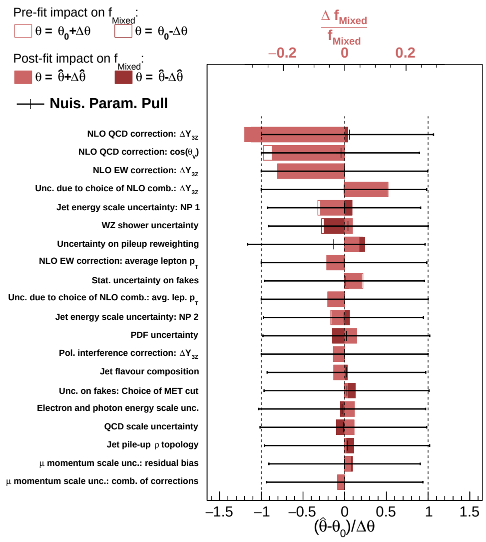


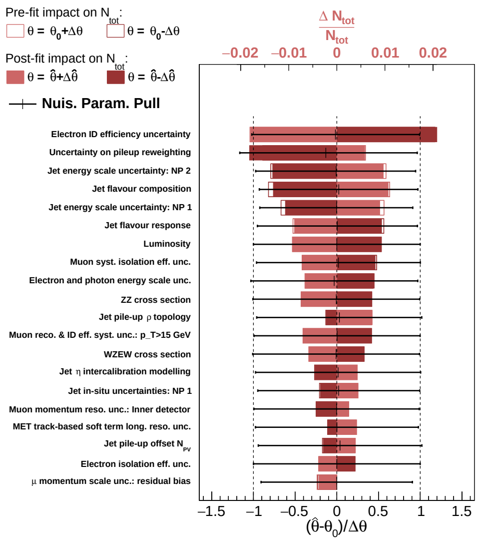

**Figure 10.13:** Ranking of the systematic nuisance parameters included in the fit in the 00-enriched $100 < p_T^Z \leq 200\,\text{GeV}$ region using the three sample strategy according to their impact on (a) $f_{00}$, (b) $f_{Mixed}$ and (c) $N_{tot}$ with data. Only the top 20 parameters are shown. Nuisance parameters corresponding to MC statistical uncertainties are not considered here. The empty brown rectangles correspond to the pre-fit impact on the respective parameter and the filled brown ones to the post-fit impacts, both referring to the upper scale. The impact of each nuisance parameter, $\Delta f_{00}|\Delta f_{Mixed}|\Delta N_{tot}$, is computed by comparing the nominal best-fit value with the result of the fit when fixing the considered nuisance parameter to its best-fit value, $\hat{\theta}$, shifted by its pre-fit (post-fit) uncertainties $\pm\Delta\theta$ $(\pm\Delta\hat{\theta})$. The black points show the pulls of the nuisance parameters with respect to their nominal values, $\theta_0$. These pulls and their relative post-fit errors, $\Delta\hat{\theta}/\Delta\theta$, refer to the lower scale. For experimental uncertainties, which are broken down into several independent sources, the corresponding nuisance parameter (NP) index is reported.




**(a)**




**(b)**




**(c)**

**Figure 10.14:** Ranking of the systematic nuisance parameters included in the fit in the 00-enriched $p_T^Z > 200\,\text{GeV}$ region using the three sample strategy according to their impact on (a) $f_{00}$, (b) $f_{Mixed}$ and (c) $N_{tot}$ with data. Only the top 20 parameters are shown. Nuisance parameters corresponding to MC statistical uncertainties are not considered here. The empty brown rectangles correspond to the pre-fit impact on the respective parameter and the filled brown ones to the post-fit impacts, both referring to the upper scale. The impact of each nuisance parameter, $\Delta f_{00}|\Delta f_{Mixed}|\Delta N_{tot}$, is computed by comparing the nominal best-fit value with the result of the fit when fixing the considered nuisance parameter to its best-fit value, $\hat{\theta}$, shifted by its pre-fit (post-fit) uncertainties $\pm\Delta\theta$ ($\pm\Delta\hat{\theta}$). The black points show the pulls of the nuisance parameters with respect to their nominal values, $\theta_0$. These pulls and their relative post-fit errors, $\Delta\hat{\theta}/\Delta\theta$, refer to the lower scale. For experimental uncertainties, which are broken down into several independent sources, the corresponding nuisance parameter (NP) index is reported.

## 10.2 Fit With Two Unconstrained Parameters

This section presents the results for the fits using the two-parameter configuration. They represent a reinterpretation of the same underlying data using a more constrained physics modeling. This allows for the setting of stronger constraints for such models. The structure is the same as the previous section's, in that first, the fits to Asimov data are shown, followed by those to observed data.

### 10.2.1 Asimov Fit

The pre- and post-fit distributions to the Asimov data using the two-parameter configuration are shown in Figure 10.15. The pre-fit yield numbers are omitted here as they are equivalent to the ones in Table 10.1. The only difference is that the TT and mixed contributions are summed together. As with the three-parameter configuration, the only differences between the pre- and post-fit plots are slight changes in the uncertainties due to constraints and correlations.

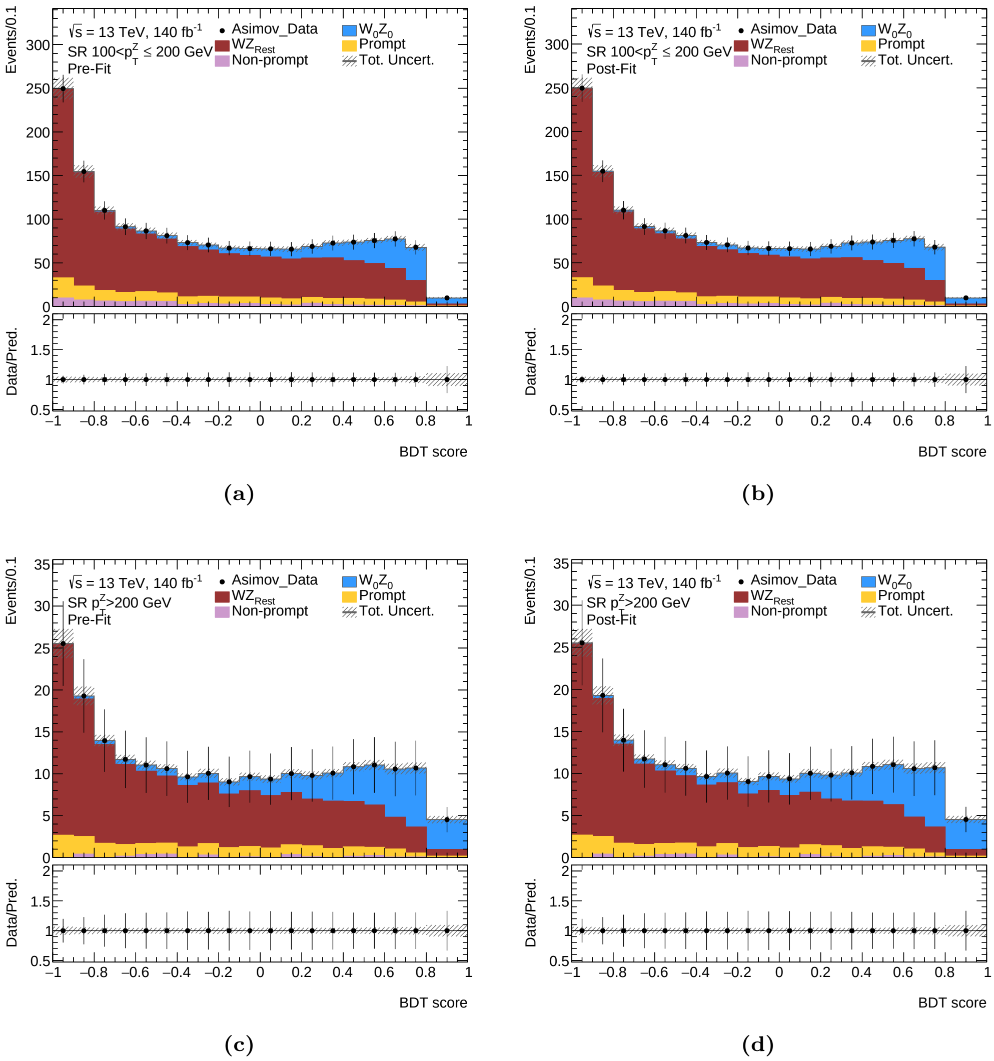


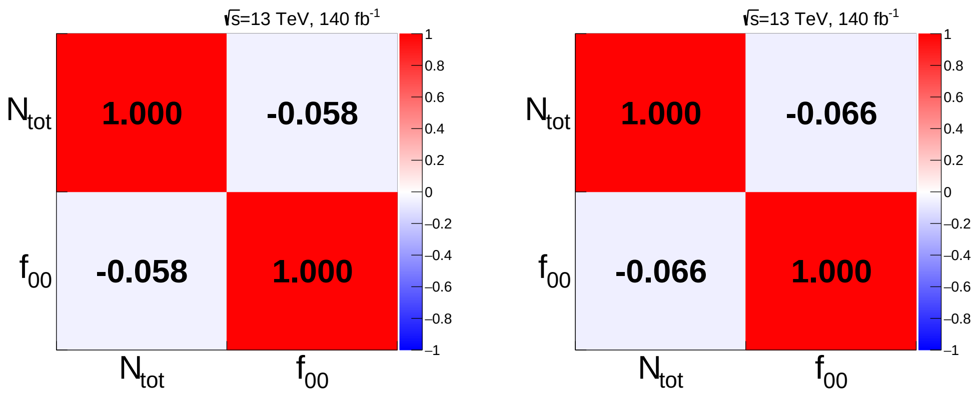

**Figure 10.15:** Expected (a) pre- and (b) post-fit distributions in the 00-enriched $100 < p_T^Z \leq 200$ GeV region and (c) pre- and (d) post-fit distributions in the 00-enriched $p_T^Z > 200$ GeV region for the S+B hypothesis with the two sample strategy.

The expected results of the fits using this configuration are shown in Table 10.6. The only significant difference to the results from the three-parameter case is the higher expected significances and the lower expected uncertainty on $f_{00}$. For example, the statistical and systematic uncertainties on $f_{00}$ in the region with $100 < p_T^Z \leq 200\,\mathrm{GeV}$ are reduced from $\approx 0.034$ and $\approx 0.017$ to 0.020 and 0.012. The nominal fit values are expectedly still the same, namely the pre-fit ones. Interestingly, the uncertainties of $N_{tot}$ are not reduced when switching from three to two unconstrained parameters. This is mostly expected, as $N_{tot}$ is explicitly decoupled from the fractions. All of this results in expected significances of $6.9\sigma$ for the region with $100 < p_T^Z \leq 200\,\mathrm{GeV}$ and 4.2 for the $p_T^Z > 200\,\mathrm{GeV}$ region, indicating potential “observation” and “evidence” of a non-zero $f_{00}$ value in these regions.

**Table 10.6:** Asimov fit results using the smoothed nuisance parameter distributions in the configuration with two independent fit parameters for the fiducial fraction parametrization.

| 00-enriched region | $100 < p_T^Z \leq 200\,\mathrm{GeV}$ | $p_T^Z > 200\,\mathrm{GeV}$ |
|---|---|---|
| Significance | $6.859\,\sigma$ | $4.225\,\sigma$ |
| $f_{00}$ | $0.152 \pm^{0.020}_{0.020}$ (stat) $\pm^{0.011}_{0.012}$ (sys) | $0.23 \pm^{0.06}_{0.05}$ (stat) $\pm^{0.02}_{0.02}$ (sys) |
| $N_{tot}$ | $2500 \pm^{70}_{70}$ (stat) $\pm^{100}_{100}$ (sys) | $330 \pm^{26}_{25}$ (stat) $\pm^{14}_{13}$ (sys) |
| $f_{Rest}$ | $0.848 \pm 0.020$(stat) $\pm 0.012$(sys) | $0.77 \pm 0.06$(stat) $\pm 0.02$(sys) |
| $\mu$ | $1.00 \pm 0.13$(stat) $\pm 0.09$(sys) | $1.00 \pm 0.25$(stat) $\pm 0.09$(sys) |
| $\mu_{WZRest}$ | $1.00 \pm 0.04$(stat) $\pm 0.04$(sys) | $1.00 \pm 0.10$(stat) $\pm 0.05$(sys) |

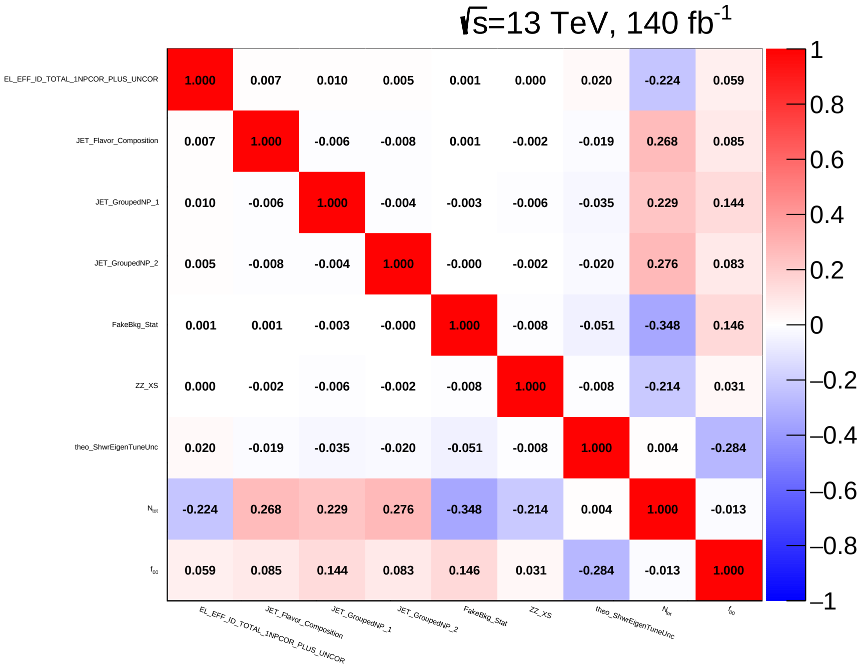
Post-fit parameter correlations for the stat. only and full fits are shown in Figures 10.16 and 10.17. The fits with no systematic uncertainties contain only two parameters, $f_{00}$ and $N_{tot}$.



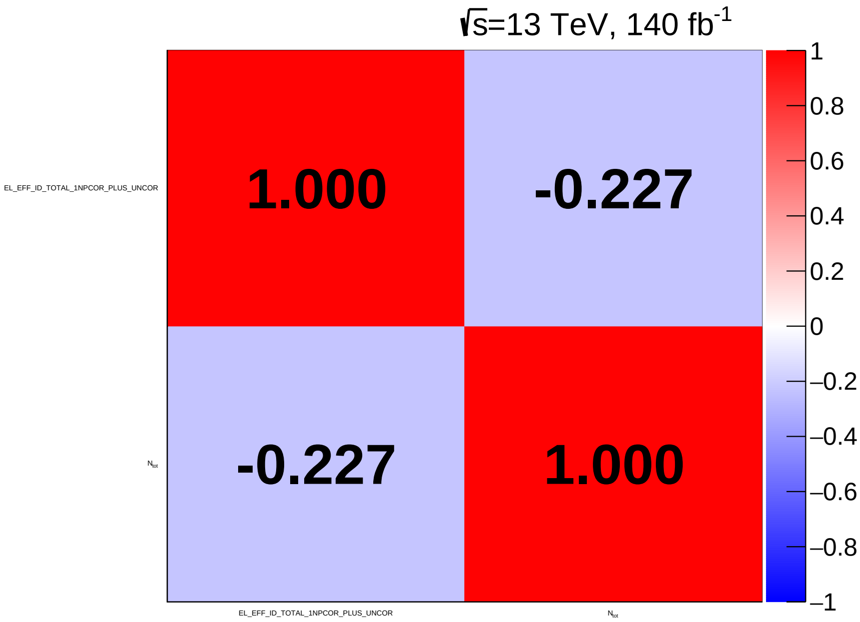

(a) (b)

**Figure 10.16:** Expected post-fit correlations for the stat. only fit in the 00-enriched region with (a) $100 < p_T^Z \leq 200\,\mathrm{GeV}$ and (b) $p_T^Z > 200\,\mathrm{GeV}$ for the S+B hypothesis with the two sample strategy.

Despite them being nominally decoupled, small correlations of $-0.058$ and $-0.066$ can be seen in the low and high $p_T^Z$ regions, respectively. These arise because smaller/higher values for $f_{00}$ might be best compensated by slightly increasing/reducing the overall number of events. This improves the agreement for high BDT score values at the expense of lower values. Whether this compensation should happen this way or the other way around is not immediately obvious and depends on the exact distributions as well as the size of the variations. For example, the background-only fits with $f_{00}$ set to zero both have a lower value




(a)




(b)

**Figure 10.17:** Expected post-fit correlations of nuisance parameters with a correlation of at least 0.2 with another NP for the full fit in the 00-enriched region with (a) $100 < p_T^Z \leq 200$ GeV and (b) $p_T^Z > 200$ GeV for the S+B hypothesis with the two sample strategy.

of $N_{tot}$, indicating a positive correlation between the two parameters.

The correlations of the full fit show the same general behavior as in the three-parameter case. There are fewer parameters with a correlation of above 0.2 in the high $p_T^Z$ region compared to the low $p_T^Z$ one, primarily due to larger statistical uncertainties. The nuisance parameters that do pass this threshold are also the same ones as in the other configuration and have it with $N_{tot}$. They also show smaller but non-zero correlations with the fraction parameter.

An overview of nuisance parameter impacts on $f_{00}$ and $N_{tot}$ is shown in Figure 10.18. The top-ranked parameters for $N_{tot}$ are almost completely identical to the ones from the three-parameter configuration. This is again mostly expected for the same reasons as the uncertainties are mostly identical between the two configurations. The situation is similar for the $f_{00}$ fraction in the high $p_T^Z$ region. Here, the top-ranked variables are again related to NLO EWK and QCD effects. The main difference is that they now cause reductions in the value of $f_{00}$, while increases are observed for the three-parameter configuration. These are, again, effects of how free mixed polarization states mediate between and compensate for the effects on the 00 and TT states. In the region with $100 < p_T^Z \leq 200\,\text{GeV}$ the most relevant uncertainties have changed a bit. Here, an uncertainty related to the interference between polarization states has become the dominant one. The shower uncertainty has also made its way into the twenty most important variables, up to the second spot.

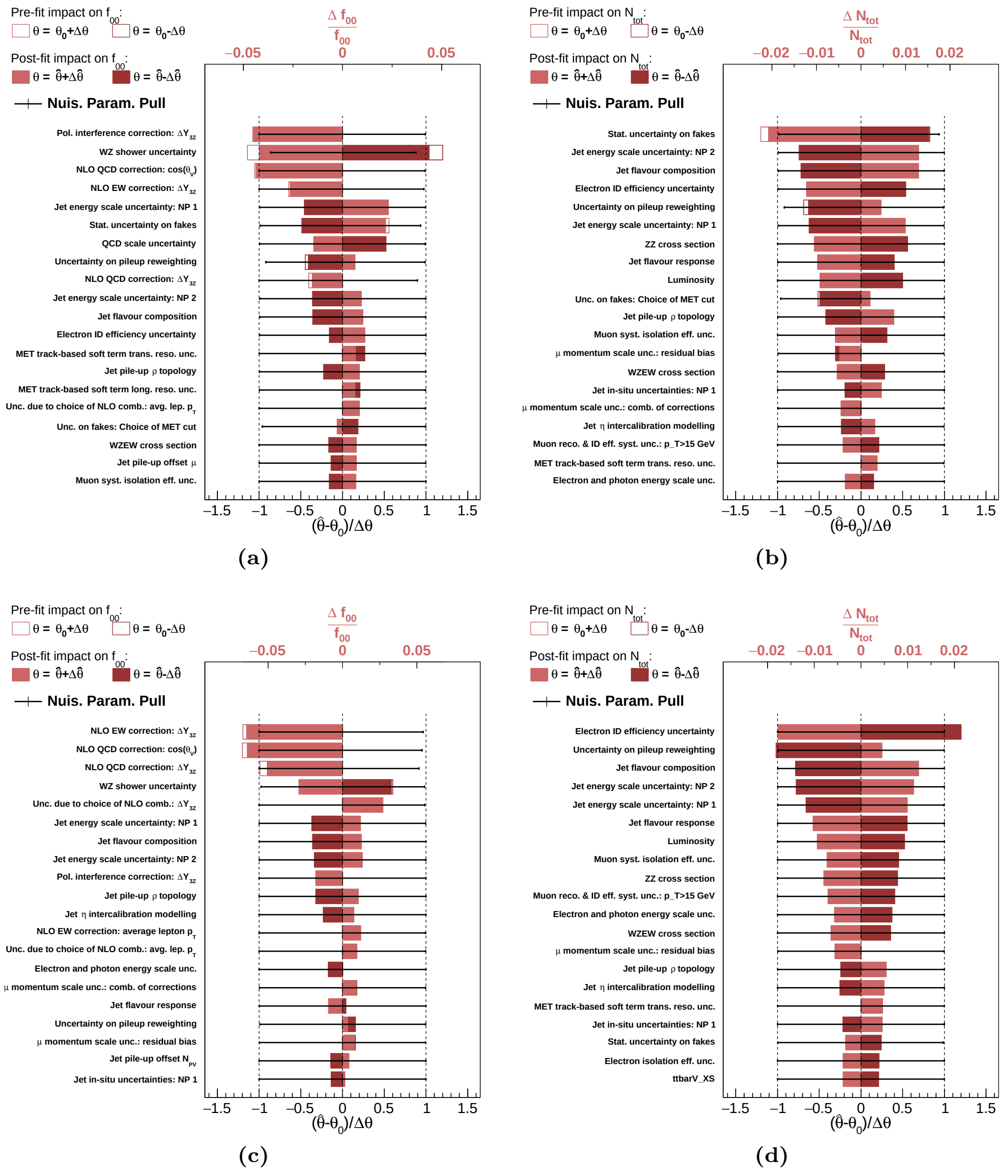



**Figure 10.18:** Ranking of the systematic nuisance parameters included in the fit in the 00-enriched $100 < p_T^Z \leq 200$ GeV region using the two sample strategy according to their impact on (a) $f_{00}$ and (b) $N_{tot}$ with Asimov data obtained with the expected background prediction and in the region with $p_T^Z > 200$ GeV for (c) $f_{00}$ and (d) $N_{tot}$. Only the top 20 parameters are shown. Nuisance parameters corresponding to MC statistical uncertainties are not considered here. The empty brown rectangles correspond to the pre-fit impact on the respective parameter and the filled brown ones to the post-fit impacts, both referring to the upper scale. The impact of each nuisance parameter, $\Delta f_{00}|\Delta f_{Mixed}|\Delta N_{tot}$, is computed by comparing the nominal best-fit value with the result of the fit when fixing the considered nuisance parameter to its best-fit value, $\hat{\theta}$, shifted by its pre-fit (post-fit) uncertainties $\pm\Delta\theta$ ($\pm\Delta\hat{\theta}$). The black points show the pulls of the nuisance parameters with respect to their nominal values, $\theta_0$. These pulls and their relative post-fit errors, $\Delta\hat{\theta}/\Delta\theta$, refer to the lower scale. For experimental uncertainties, which are broken down into several independent sources, the corresponding nuisance parameter (NP) index is reported.

### 10.2.2 Observed Fit

The pre- and post-fit distributions of the observed fit using the two-parameter configuration are shown in Figure 10.19. The pre-fit agreement between data and expectation is identical to the one in the three-parameter fits. The post-fit agreement is slightly worse due to the less flexible model. The post-fit event yields are shown in Table 10.7 and mirror the message from the plots.

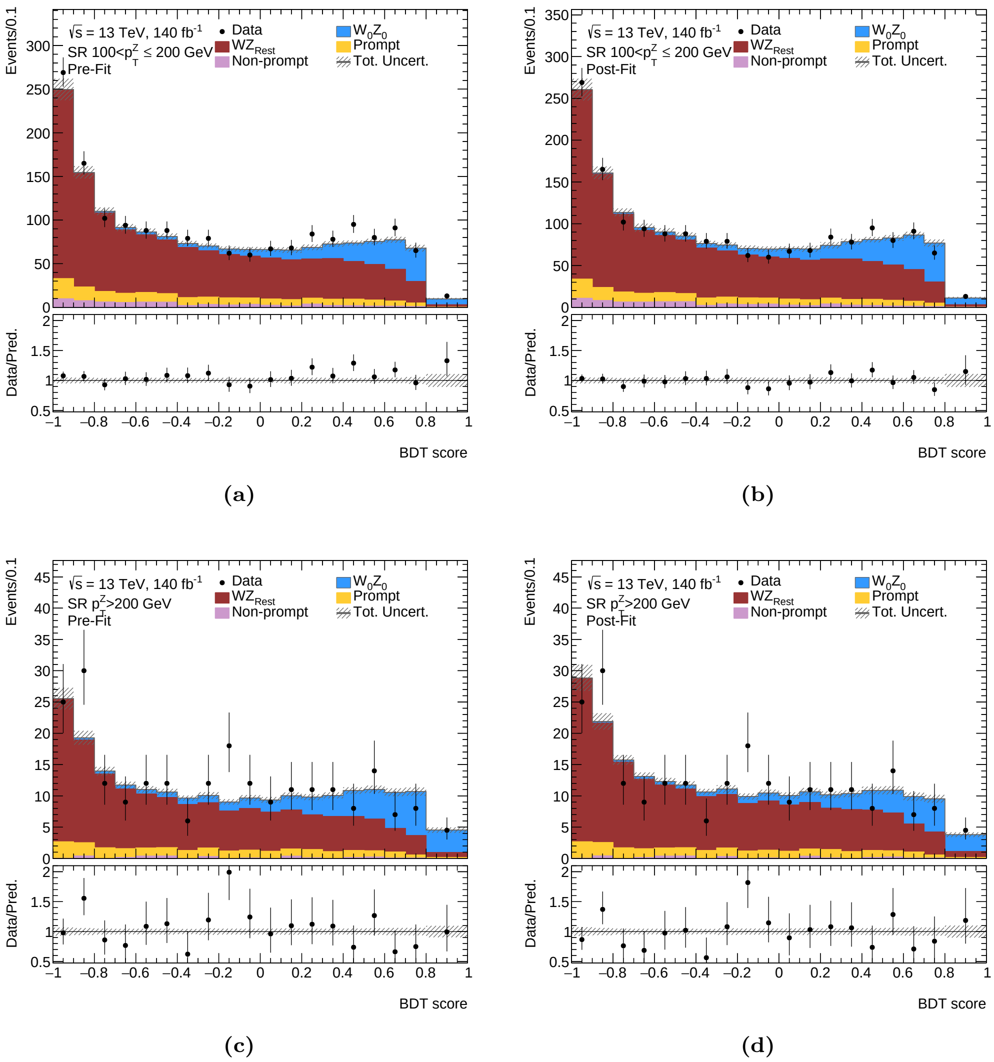



**Figure 10.19:** Observed (a) pre- and (b) post-fit distributions in the 00-enriched $100 < p_T^Z \leq 200\,\text{GeV}$ region and (c) pre- and (d) post-fit distributions in the 00-enriched $p_T^Z > 200\,\text{GeV}$ region for the S+B hypothesis with the two sample strategy.

The observed correlations in the stat. only and full fits are shown in Figures 10.20 and 10.21, respectively. The small correlation between $N_{tot}$ and $f_{00}$ is still present in the observed fit. The observed correlation coefficients are $-0.063$ and $-0.056$ for the low and high $p_T^Z$ regions.

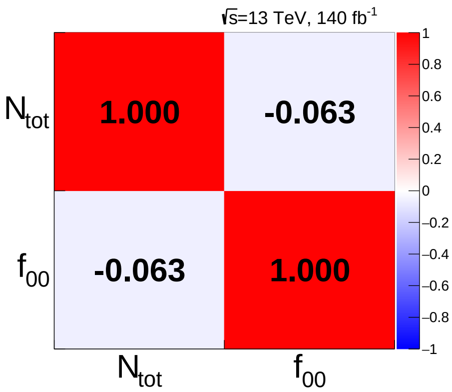
The correlations in the full fit show some more differences compared to the ones in the Asimov fit. In the $100 < p_T^Z \leq 200\,\text{GeV}$ the number of nuisance parameters with a correlation above 0.2 decreases from 7 to 6. The most noteworthy addition is again one of the NLO QCD uncertainties, this time from the reweighting of the $\cos\theta_V$ distribution. It has a large

**Table 10.7:** Observed post-fit yields in the configuration with two free parameters. Shown are the nominal values and statistical, theory, and experimental uncertainties.

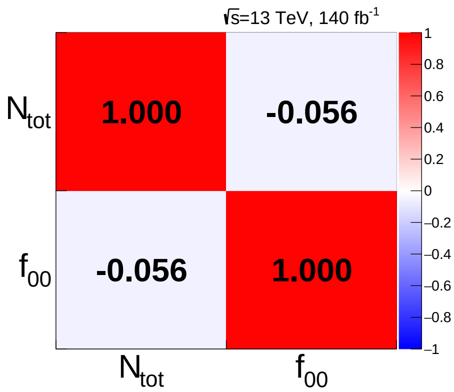
| | $100 < p_T^Z \leq 200$ GeV | $p_T^Z > 200$ GeV |
|---|---|---|
| $W_0Z_0$ | 271 $\pm 31$ $\pm$ 0 $\pm$ 7 | 36 $\pm 11$ $\pm$ 0 $\pm$ 1 |
| $WZ_{Rest}$ | 1230 $\pm 40$ $\pm 10$ $\pm 40$ | 173 $\pm 16$ $\pm$ 1 $\pm$ 6 |
| Prompt background | 168 $\pm$ 0 $\pm 17$ $\pm$ 8 | 24.1 $\pm$ 0.0 $\pm$ 2.1 $\pm$ 1.6 |
| Non-prompt background | 80 $\pm$ 0 $\pm$ 0 $\pm 40$ | 2.9 $\pm$ 0.0 $\pm$ 0.0 $\pm$ 1.0 |
| Total Expected | 1740 $\pm 40$ $\pm 20$ $\pm 60$ | 236 $\pm 15$ $\pm$ 3 $\pm$ 8 |
| Data | 1740.0 $\pm$ 0.0 $\pm$ 0.0 $\pm$ 0.0 | 236.0 $\pm$ 0.0 $\pm$ 0.0 $\pm$ 0.0 |

enough correlation with $f_{00}$ to make the cut. The region with $p_T^Z > 200$ GeV goes from two to four parameters that have at least one correlation coefficient above the threshold. Here, the previously missing $f_{00}$ is added alongside the NLO QCD uncertainty from $\Delta Y(\ell_W Z)$, which it is anti-correlated with. Both of these additions can again be understood by considering the pull plots in Figures 10.22c and 10.23c. Both of the mentioned NLO uncertainties are within the most pulled ones in their respective regions. This is due to them being able to improve the agreement with data, which is a difference with respect to the Asimov fit, where the best agreement is always found without any pulls.

Comparing the pulls and constraints to those observed for the three-parameter one, it can be seen that the ones for the two-parameter configuration are generally slightly larger. This is due to the reduced freedom of the fit to cover discrepancies by adjusting the unconstrained parameters. Additionally, even in this configuration, there are no unexpectedly large pulls or constraints, especially for experimental uncertainties determined by dedicated measurements. Instead, the largest ones are for theory uncertainties of the shower and NLO effects as well as the determination of the non-prompt background. However, even these are not constrained to a strong degree that would warrant further investigations.




(a)

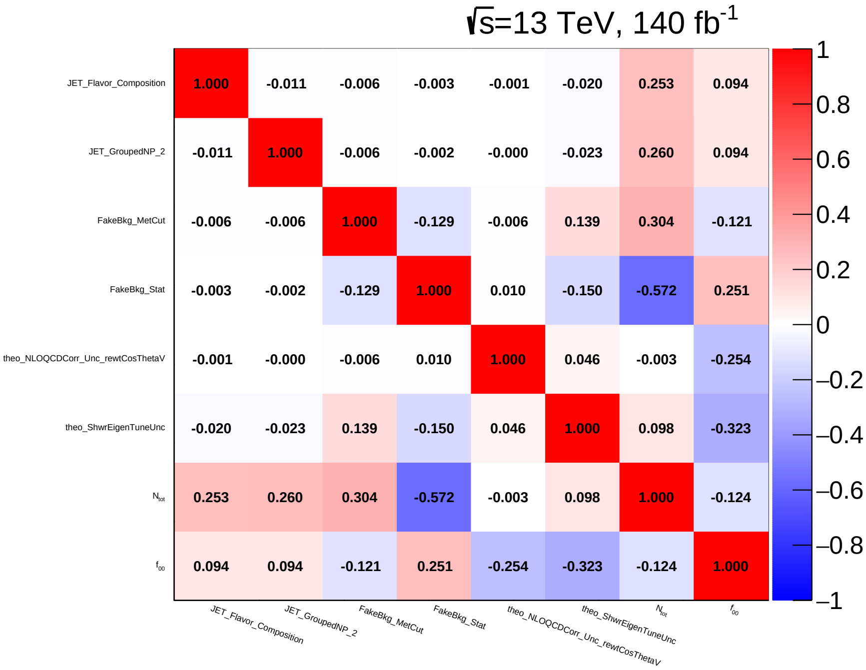



(b)

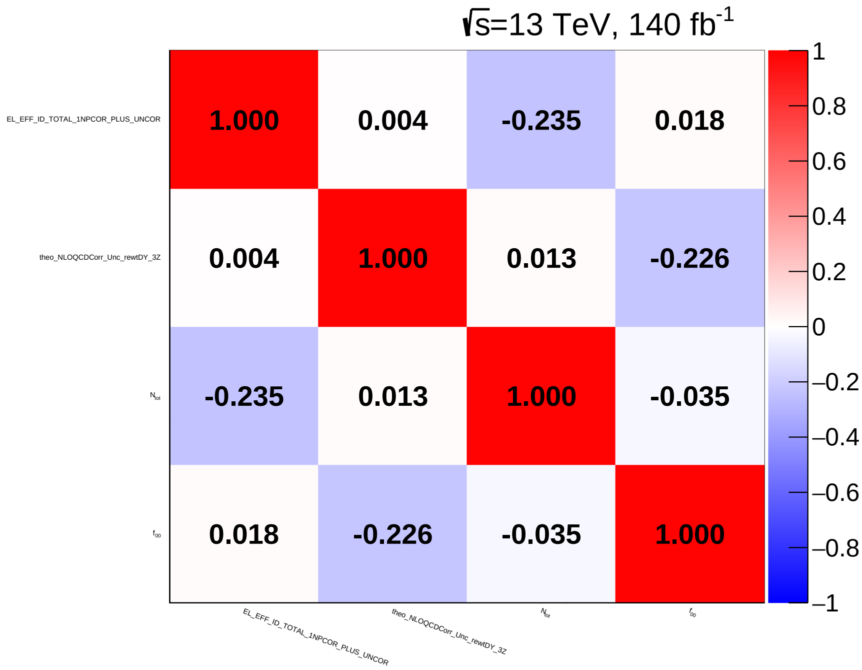
**Figure 10.20:** Observed post-fit correlations for the stat. only fit in the 00-enriched region with (a) $100 < p_T^Z \leq 200$ GeV and (b) $p_T^Z > 200$ GeV for the S+B hypothesis with the two sample strategy.

The post-fit values for the unconstrained fit parameters are shown alongside their uncertainties in Table 10.8. The general trend of more than expected 00 contributions in the lower $p_T^Z$ region and less in the region with $p_T^Z > 200$ GeV is the same as with the three-parameter configuration.

The observed result also follows the expected one in that the significances in both regions

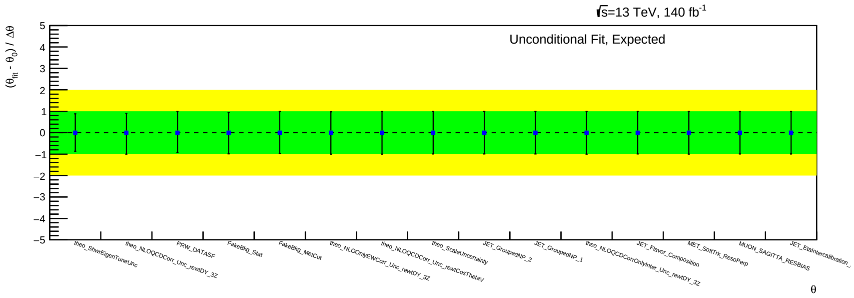



**(a)**

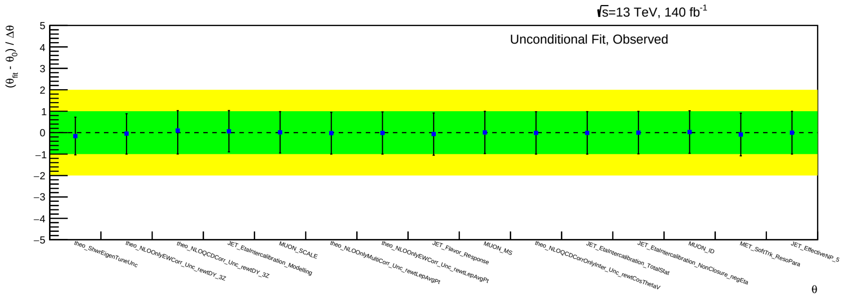



**(b)**

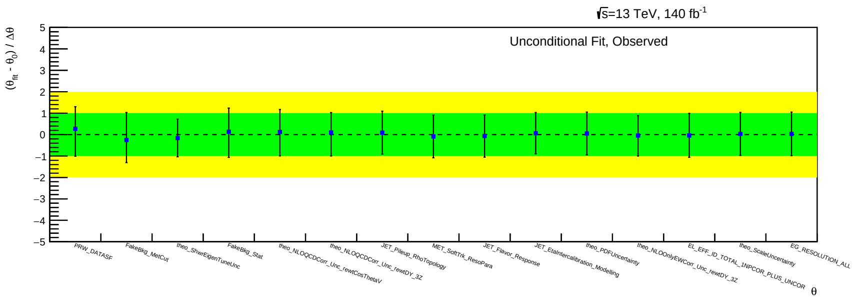
**Figure 10.21:** Observed post-fit correlations of nuisance parameters with a correlation of at least 0.2 with another NP for the full fit in the 00-enriched region with (a) $100 < p_T^Z \leq 200\,\text{GeV}$ and (b) $p_T^Z > 200\,\text{GeV}$ for the S+B hypothesis with the two sample strategy.




**(a)**




**(b)**




**(c)**

**Figure 10.22:** Post fit nuisance parameter pulls for the fit of the S+B hypothesis in the 00-enriched $100 < p_T^Z \leq 200\,\mathrm{GeV}$ with the two sample strategy. Shown are (a) the most constrained parameters in the fit to Asimov data, (b) the most constrained and (c) the most pulled parameters in the fit to observed data.

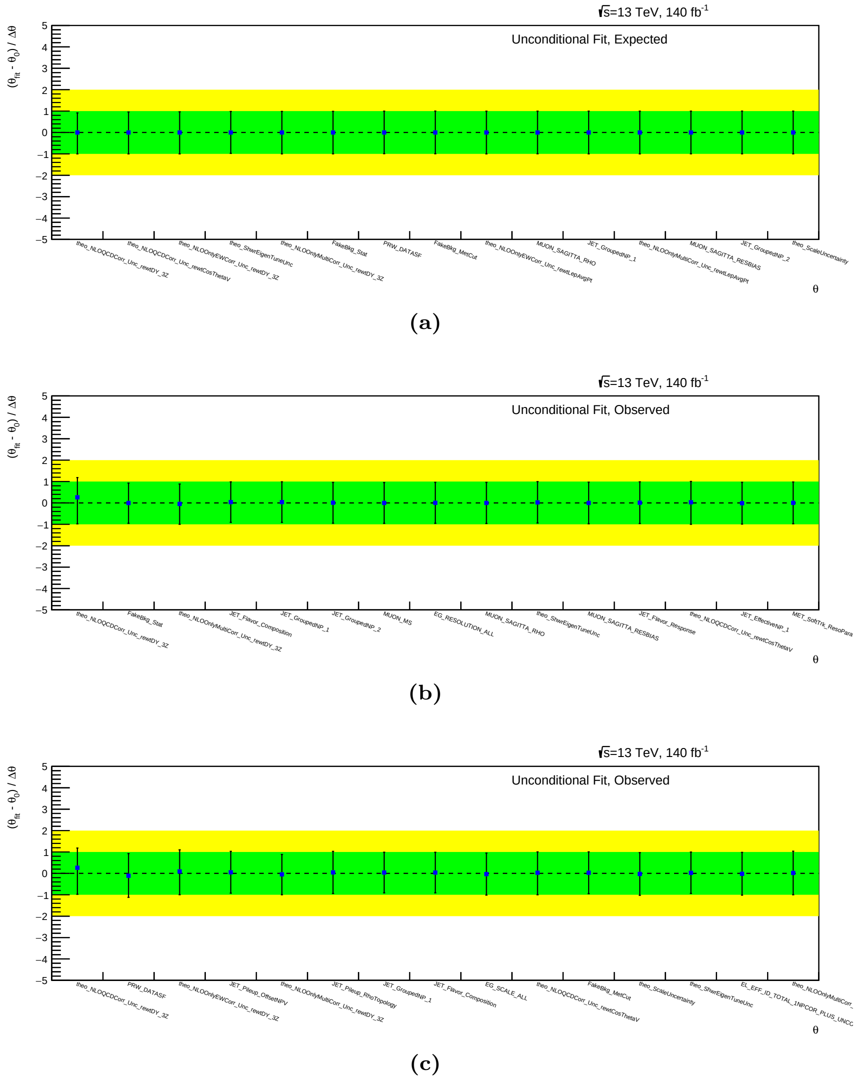



**Figure 10.23:** Post fit nuisance parameter pulls for the fit of the S+B hypothesis in the 00-enriched $p_T^Z > 200\,\text{GeV}$ with the two sample strategy. Shown are (a) the most constrained parameters in the fit to Asimov data, (b) the most constrained and (c) the most pulled parameters in the fit to observed data.

are larger than those in the three-parameter configuration. However, the nominal fit results of $f_{00}$ are not identical to those found for the three-parameter configuration. This is expected as this configuration can not reproduce the behavior of a mixed signal strength of $\approx 2$ while keeping $\mu_{WZTT} \approx 1$. Thus, different values of $f_{00}$ can not be compensated as freely. The final result of this is that the observed values for the doubly longitudinal fraction lie closer to the expected ones. With these results, it can also be seen that the values and uncertainties of $N_{tot}$ are not completely independent from the chosen configuration, as the value in the $100 < p_T^Z \leq 200$ GeV and the systematic up uncertainty are both lower in the two-parameter configuration than in the three-parameter one.

The final value for $f_{00}$ in the region with $100 < p_T^Z \leq 200$ GeV with the two-parameter configuration is:

$$f_{00} = 0.174 \pm^{0.020}_{0.020} \text{ (stat.)} \pm^{0.010}_{0.013} \text{ (mod. syst.)} \pm^{0.009}_{0.008} \text{ (exp. syst.)} \pm^{0.000}_{0.000} \text{ (lumi.)} \tag{10.3}$$

in the region with $p_T^Z > 200$ GeV the value is:

$$f_{00} = 0.16 \pm^{0.05}_{0.05} \text{ (stat.)} \pm^{0.02}_{0.02} \text{ (mod. syst.)} \pm^{0.01}_{0.01} \text{ (exp. syst.)} \pm^{0.00}_{0.00} \text{ (lumi.)} \tag{10.4}$$

As is the case with the three-parameter configuration, both values almost contain the expected values within $1\sigma$. The change in which regions' fraction is larger compared to the expectation also remains in this more constrained configuration. However, the difference is now smaller with the fraction in the $100 < p_T^Z \leq 200$ GeV region being just one percentage pointer larger than the one in the higher $p_T^Z$ region. This is in contrast to the almost six percentage points in the three-parameter configuration and the expectation that $f_{00}$ in the $p_T^Z > 200$ GeV region is approximately eight percentage points larger than the one in the lower $p_T^Z$ region.

**Table 10.8:** Observed fit results using the smoothed nuisance parameter distributions in the configuration with two independent fit parameters for the fiducial fraction parametrization.

| | $100 < p_T^Z \leq 200$ GeV | $p_T^Z > 200$ GeV |
|---|---|---|
| Significance | $7.726\,\sigma$ | $3.193\,\sigma$ |
| $f_{00}$ | $0.174 \pm^{0.020}_{0.020}$ (stat) $\pm^{0.013}_{0.015}$ (sys) | $0.16 \pm^{0.05}_{0.05}$ (stat) $\pm^{0.02}_{0.03}$ (sys) |
| $N_{tot}$ | $2680 \pm^{80}_{70}$ (stat) $\pm^{120}_{120}$ (sys) | $355 \pm^{27}_{26}$ (stat) $\pm^{15}_{14}$ (sys) |
| $f_{Rest}$ | $0.826 \pm 0.020$(stat) $\pm 0.014$(sys) | $0.84 \pm 0.05$(stat) $\pm 0.02$(sys) |
| $\mu$ | $1.23 \pm 0.14$(stat) $\pm 0.11$(sys) | $0.76 \pm 0.24$(stat) $\pm 0.11$(sys) |
| $\mu_{WZRest}$ | $1.05 \pm 0.04$(stat) $\pm 0.05$(sys) | $1.17 \pm 0.11$(stat) $\pm 0.06$(sys) |

A breakdown of the uncertainties on the fraction of 00 events is shown in Table 10.9. Overall, similar effects to those of the three-parameter configuration can be observed. This includes generally larger uncertainties in the $p_T^Z > 200$ GeV region, a dominant statistical uncertainty in both regions, with the effect being more pronounced in the higher $p_T^Z$ region, and a decrease in the uncertainties related to the non-prompt background and polarization state interference. A direct comparison with the values for the three-parameter configuration shows that total and most individual uncertainties are smaller in the configuration with two unconstrained parameters. The notable exceptions are the NLO QCD and EW uncertainties in the region with $p_T^Z > 200$ GeV, these increase from 8.2 % and 3.3 % in the three-parameter configuration to 9.1 % and 7.5 % the one with just two. This indicates that these effects generally become larger at high energies and have a sizeable impact on the $f_{00}$ and $f_{TT}$ parameters as the individual impact of the mixed states is reduced in the two-parameter configuration due to their direct coupling to the larger TT ones.

Likelihood scans of the free parameters and some nuisance parameters are shown in Figures 10.11 and 10.12 for the low and high $p_T^Z$ regions. They show the same general well-behaved shapes as already seen with the three-parameter configuration. Rankings of the most important nuisance parameters are given in Figure 10.26. The most important NPs are, again, those already seen for the Asimov fits. The agreement between the expected and observed rankings is even better than in the three-parameter case. The same holds for the overall relative impacts of each parameter. This is primarily due to the small difference in expected and observed values of $f_{00}$.

**Table 10.9:** Breakdown of uncertainties on $f_{00}$ in the fits to observed data with the two-parameter configuration.

| Source | Impact on $f_{00}$ [%] | |
|---|---|---|
| Experimental | $100 < p_T^Z \leq 200\,\mathrm{GeV}$ | $p_T^Z > 200\,\mathrm{GeV}$ |
| Luminosity | 0.1 | 0.1 |
| Electron calibration | 1.0 | 0.9 |
| Muon calibration | 0.6 | 0.6 |
| Jet energy scale and resolution | 3.1 | 4.8 |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 0.3 | 0.3 |
| Flavor-tagging inefficiency | 0.0 | 0.0 |
| Pileup modelling | 1.0 | 0.7 |
| Non-prompt background estimation | 3.6 | 0.6 |
| Modelling | | |
| Background, other | 0.9 | 0.9 |
| Model statistical | 1.6 | 2.2 |
| NLO QCD effects | 3.7 | 9.1 |
| NLO EW effects | 0.9 | 7.5 |
| Effect of additive vs multiplicative QCD+EW combination | 0.5 | 1.7 |
| Interference impact | 2.4 | 1.4 |
| PDF, Scales, and shower settings | 4.0 | 4.0 |
| Experimental and modelling | 8.1 | 13.9 |
| Data statistical | 11.4 | 31.4 |
| Total | 14.0 | 34.4 |

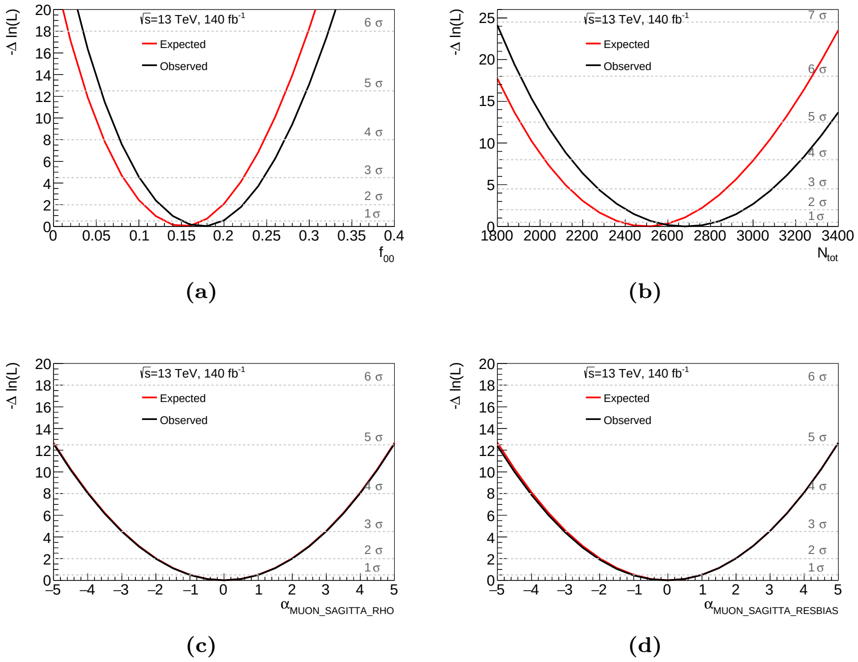



**Figure 10.24:** Likelihood scans for the full fits of the S+B hypothesis in the 00-enriched $100 < p_T^Z \leq 200$ GeV region with the two sample strategy for the (a) $f_{00}$, (b) $N_{tot}$, (c) `MUON_SAGITTA_RHO` and (d) `MUON_SAGITTA_RESBIAS` parameters.

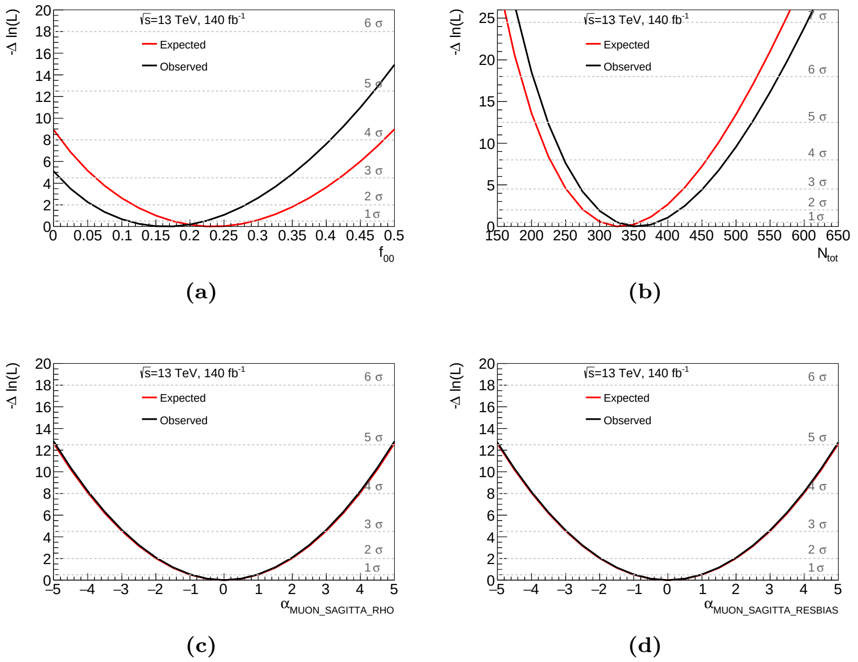
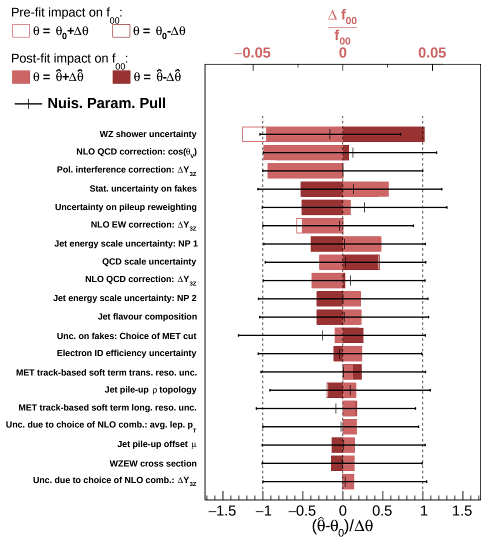
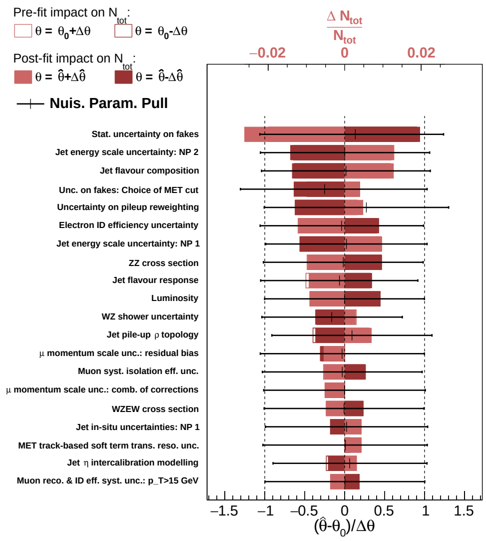


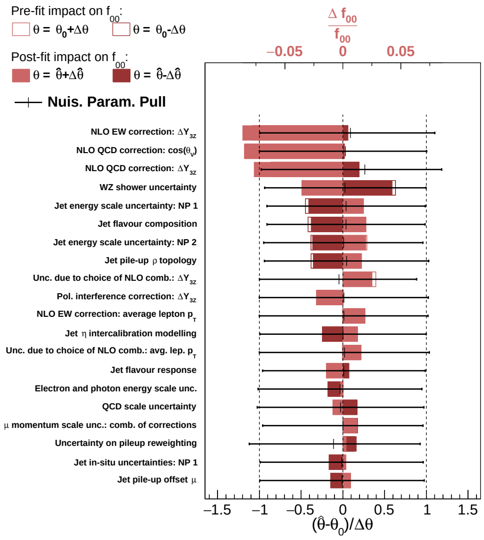

**Figure 10.25:** Likelihood scans for the full fits of the S+B hypothesis in the 00-enriched $p_T^Z > 200$ GeV region with the two sample strategy for the (a) $f_{00}$, (b) $N_{tot}$, (c) `MUON_SAGITTA_RHO` and (d) `MUON_SAGITTA_RESBIAS` parameters.



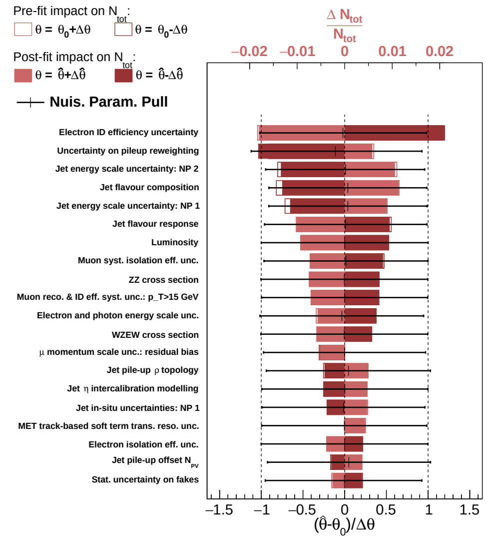

**(a)**




**(b)**




**(c)**

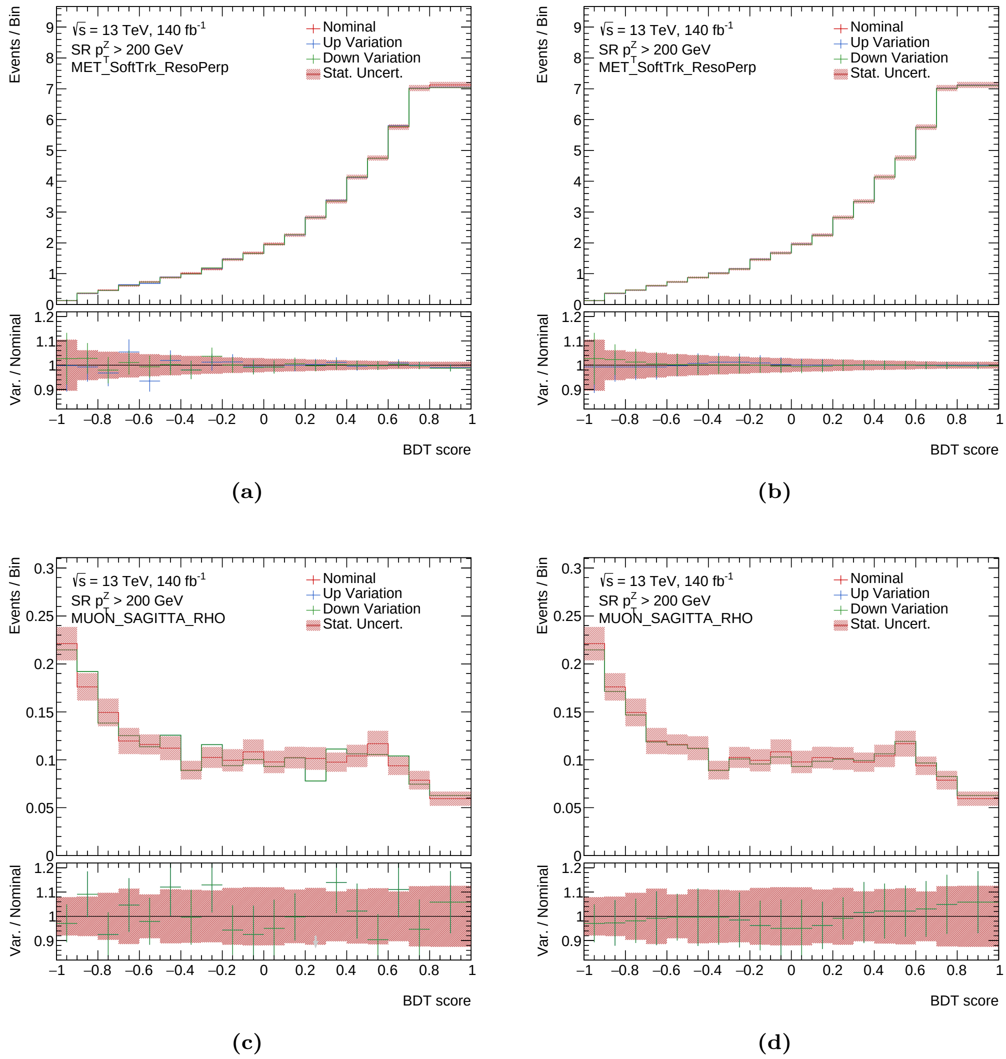



**(d)**

**Figure 10.26:** Ranking of the systematic nuisance parameters included in the fit in the 00-enriched $100 < p_T^Z \leq 200\,\text{GeV}$ region using the two sample strategy according to their impact on (a) $f_{00}$ and (b) $N_{tot}$ with data and in the region with $p_T^Z > 200\,\text{GeV}$ for (c) $f_{00}$ and (d) $N_{tot}$. Only the top 20 parameters are shown. Nuisance parameters corresponding to MC statistical uncertainties are not considered here. The empty brown rectangles correspond to the pre-fit impact on the respective parameter and the filled brown ones to the post-fit impacts, both referring to the upper scale. The impact of each nuisance parameter, $\Delta f_{00}|\Delta f_{Mixed}|\Delta N_{tot}$, is computed by comparing the nominal best-fit value with the result of the fit when fixing the considered nuisance parameter to its best-fit value, $\hat{\theta}$, shifted by its pre-fit (post-fit) uncertainties $\pm\Delta\theta$ ($\pm\Delta\hat{\theta}$). The black points show the pulls of the nuisance parameters with respect to their nominal values, $\theta_0$. These pulls and their relative post-fit errors, $\Delta\hat{\theta}/\Delta\theta$, refer to the lower scale. For experimental uncertainties, which are broken down into several independent sources, the corresponding nuisance parameter (NP) index is reported.

## 10.3 Smoothing Impact

The most direct way to assess the impact of the smoothing is a comparison of smoothed and unsmoothed templates. Two such comparisons can be found in Figure 10.27. It shows the smoothed and unsmoothed distributions for the `MET_SoftTrk_ResoPerp` nuisance parameter variations of the signal samples and the `MUON_SAGITTA_RHO` variations for the $VVV$ background in the signal region with $p_T^Z > 200\,\text{GeV}$. The latter represents the templates with the largest $\chi^2$ value between the smoothed and unsmoothed distributions for any sample, region, and parameter. The up and down variations for the `MUON_SAGITTA_RHO` parameter are nearly identical, leading to only the down variation being visible in the figure. These distributions show that the smoothing impact on the individual templates is small and within the MC statistical uncertainties. As such, it only reduces the impact of MC fluctuations without affecting the physical meaning of the individual nuisance parameters.




(a) (b) (c) (d)

**Figure 10.27:** Comparison of (a) unsmoothed and (b) smoothed distributions for variations of the `MET_SoftTrk_ResoPerp` nuisance parameter for the template of the $W_L Z_L$ signal in the region with $p_T^Z > 200\,\text{GeV}$. Also shown are the (c) unsmoothed and (d) smoothed distributions for variations of the `MUON_SAGITTA_RHO` nuisance parameter for the template of the $VVV$ background in the region with $p_T^Z > 200\,\text{GeV}$.

Furthermore, fits are performed in the two 00-enriched signal regions using the reconstructed fraction parametrization with and without smoothing applied. The resulting significances and post-fit parameters are shown in Tables 10.10 and 10.11 for the unsmoothed and smoothed fits, respectively. It can be seen that the smoothing only has a small impact on the significances, values, and uncertainties. The maximal significance decrease is found to be ≈ 0.09, while a maximum increase of ≈ 0.03 can be observed. The only fraction value that is changed within the significant digits is the $f_{Mixed}$ fraction in the observed fit in the $p_T^Z > 200\,\text{GeV}$ region. It is reduced by one percentage point from 0.24 to 0.23. The actual difference is lower still, with the more accurate values being 0.2369 and 0.2348, as also indicated by the sum of the rounded fractions being 1.01 originally and 1.00 after the smoothing, while the sum of unrounded values is 1.00 by construction.

**Table 10.10:** Fit results using the unsmoothed nuisance parameter distributions in the configuration with three independent fit parameters for the reconstructed fraction parametrization.

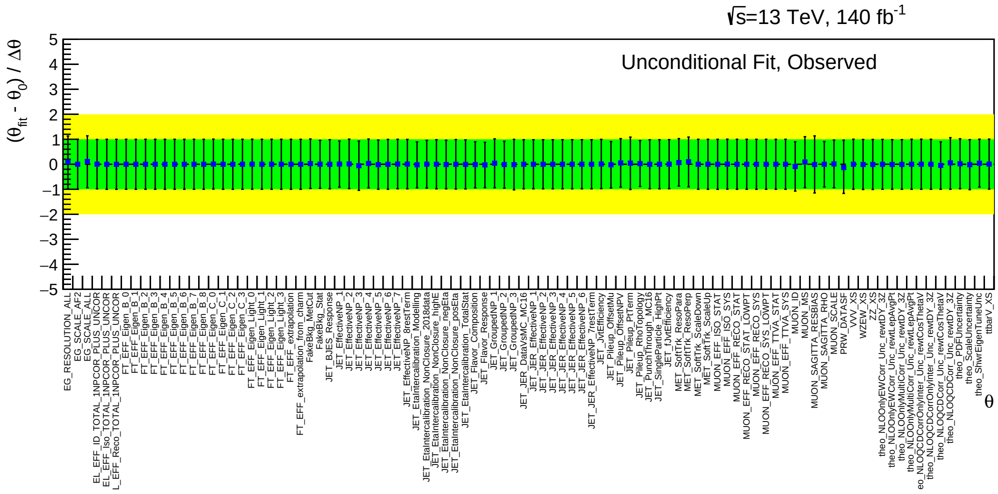
| | SR $100 < p_T^Z \leq 200\,\text{GeV}$ | | SR $p_T^Z > 200\,\text{GeV}$ | |
|---|---|---|---|---|
| | Exp. | Obs. | Exp. | Obs. |
| Significance | $4.374\,\sigma$ | $5.126\,\sigma$ | $2.494\,\sigma$ | $1.588\,\sigma$ |
| $f_{00}$ | $0.16\pm^{0.04}_{0.04}$ | $0.20\pm^{0.05}_{0.04}$ | $0.24\pm^{0.10}_{0.10}$ | $0.14\pm^{0.09}_{0.09}$ |
| $f_{Mixed}$ | $0.23\pm^{0.08}_{0.09}$ | $0.18\pm^{0.09}_{0.10}$ | $0.12\pm^{0.20}_{0.12}$ | $0.24\pm^{0.18}_{0.21}$ |
| $N_{tot}$ | $1400\pm^{70}_{70}$ | $1510\pm^{80}_{80}$ | $195\pm^{17}_{16}$ | $209\pm^{18}_{17}$ |
| $f_{TT}$ | $0.61 \pm 0.06$ | $0.62 \pm 0.06$ | $0.63 \pm 0.09$ | $0.63 \pm 0.13$ |
| $\mu$ | $1.00 \pm 0.25$ | $1.37 \pm 0.30$ | $1.0 \pm 0.4$ | $0.6 \pm 0.4$ |
| $\mu_{WZTT}$ | $1.00 \pm 0.11$ | $1.10 \pm 0.13$ | $1.00 \pm 0.17$ | $1.06 \pm 0.24$ |
| $\mu_{WZMixed}$ | $1.0 \pm 0.4$ | $0.8 \pm 0.5$ | $1.0 \pm 1.3$ | $2.1 \pm 1.7$ |

**Table 10.11:** Fit results using the smoothed nuisance parameter distributions in the configuration with three independent fit parameters for the reconstructed fraction parametrization.

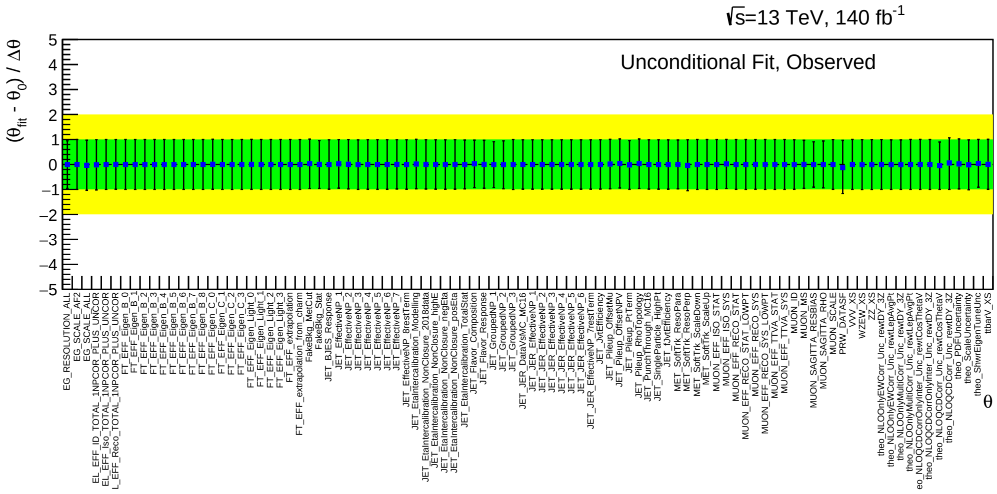
| | SR $100 < p_T^Z \leq 200\,\text{GeV}$ | | SR $p_T^Z > 200\,\text{GeV}$ | |
|---|---|---|---|---|
| | Exp. | Obs. | Exp. | Obs. |
| Significance | $4.284\,\sigma$ | $5.158\,\sigma$ | $2.493\,\sigma$ | $1.595\,\sigma$ |
| $f_{00}$ | $0.16\pm^{0.04}_{0.04}$ | $0.20\pm^{0.04}_{0.04}$ | $0.24\pm^{0.10}_{0.10}$ | $0.14\pm^{0.09}_{0.09}$ |
| $f_{Mixed}$ | $0.23\pm^{0.09}_{0.09}$ | $0.18\pm^{0.09}_{0.10}$ | $0.12\pm^{0.20}_{0.12}$ | $0.23\pm^{0.18}_{0.21}$ |
| $N_{tot}$ | $1400\pm^{70}_{70}$ | $1510\pm^{80}_{80}$ | $195\pm^{17}_{16}$ | $209\pm^{18}_{17}$ |
| $f_{TT}$ | $0.61 \pm 0.06$ | $0.62 \pm 0.06$ | $0.63 \pm 0.09$ | $0.63 \pm 0.13$ |
| $\mu$ | $1.00 \pm 0.25$ | $1.33 \pm 0.29$ | $1.0 \pm 0.4$ | $0.6 \pm 0.4$ |
| $\mu_{WZTT}$ | $1.00 \pm 0.11$ | $1.09 \pm 0.12$ | $1.00 \pm 0.17$ | $1.06 \pm 0.24$ |
| $\mu_{WZMixed}$ | $1.0 \pm 0.4$ | $0.9 \pm 0.4$ | $1.0 \pm 1.3$ | $2.1 \pm 1.7$ |

The beneficial impact of the smoothing procedure can be seen in Figures 10.28 and 10.29, which show the observed nuisance parameter pulls and ranking, respectively. The fit using unsmoothed distributions shows small but unexpected pulls and impacts for parameters displaying no significant shape differences in their variations and are not expected to have any impact on the measurement based on the physics effect they describe.

These are due to MC statistical fluctuations that align with those observed in the data. The smoothing removes these effects. This is especially visible for the `MET_SoftTrk_ResoPerp` and `MUON_SAGITTA_RHO` parameters which are some of the top-ranked parameters in Figure 10.29a despite their minimal to non-existent shape effects as previously shown for the former parameter in Figure 10.27. After the smoothing, the parameters have reduced in impact or even dropped out of the list of the twenty highest impact ones, as shown in Figure 10.29. The smoothing has the additional benefit of stabilizing the fit, leading to a faster convergence.



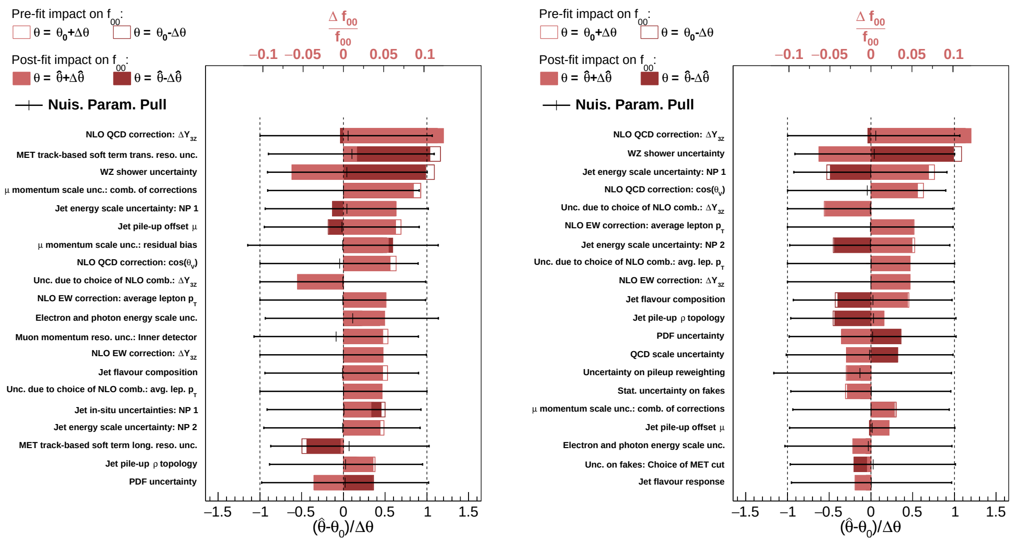

**(a)**




**(b)**

**Figure 10.28:** Observed nuisance parameter pulls in the fit using (a) the unsmoothed and (b) the smoothed nuisance parameter distributions in the configuration with three independent fit parameters for the reconstructed fraction parametrization in the region with $p_T^Z > 200\,\mathrm{GeV}$.

Due to this, the smoothed approach is used from here onwards and for the nominal results.




**(a)** **(b)**

**Figure 10.29:** Ranking of the systematic nuisance parameters included in the fit using (a) unsmoothed and (b) smoothed distributions in the region with $p_T^Z > 200\,\text{GeV}$ using the three-parameter strategy with the reconstructed fraction parametrization according to their impact on $f_{00}$ with data. Only the top 20 parameters are shown. Nuisance parameters corresponding to MC statistical uncertainties are not considered here. The empty brown rectangles correspond to the pre-fit impact on the respective parameter and the filled brown ones to the post-fit impacts, both referring to the upper scale. The impact of each nuisance parameter, $\Delta f_{00}$, is computed by comparing the nominal best-fit value with the result of the fit when fixing the considered nuisance parameter to its best-fit value, $\hat{\theta}$, shifted by its pre-fit (post-fit) uncertainties $\pm\Delta\theta$ ($\pm\Delta\hat{\theta}$). The black points show the pulls of the nuisance parameters with respect to their nominal values, $\theta_0$. These pulls and their relative post-fit errors, $\Delta\hat{\theta}/\Delta\theta$, refer to the lower scale. For experimental uncertainties, which are broken down into several independent sources, the corresponding nuisance parameter (NP) index is reported.

## 10.4 Differences Between Reconstructed and Fiducial Fractions

Next, the effect of switching from the reconstructed to the fiducial fraction parametrization, as described in Section 9.3.5, is investigated. As the physics scenarios of the background-only and S+B hypotheses are identical between both parametrizations, the expectation is that the significances are identical within the fit precision. Another expectation is that both parametrizations yield identical effective signal strengths $\mu$ as well as similar uncertainties. To verify that all of these expectations are met, the fits with smoothed distributions are compared in the scenario with three free parameters between the reconstructed and fiducial fraction parametrization.

The reconstructed results are already shown in the previous section in Table 10.11. The results using the fiducial parametrization are displayed in Table 10.12. It can be seen that the significances are identical except for the expected one in the region with $100 < p_T^Z \leq 200\,\mathrm{GeV}$, where the fiducial parametrization's is larger by 0.001. However, this is well within the fit and even machine accuracy. Good agreement can also be found for the calculated $\mu$'s and their uncertainties. The only deviation there is the observed uncertainty of $\mu_{WZTT}$ in the region with $p_T^Z > 200\,\mathrm{GeV}$ which differs by 0.01. This is still acceptable, especially considering that the actual difference is just 0.004, and such differences can arise because the actual uncertainties are symmetrized for the propagation.

More significant differences occur for the measured fractions. Here, deviations of up to 2 percentage points can be observed. Specifically, the $f_{00}$ fractions decrease while the $f_{TT}$ fractions increase. This indicates a better reconstruction efficiency for longitudinal events. The fiducial fractions represent a different physics quantity than the reconstructed ones. They are of more interest to the physics community outside of the ATLAS experiment as they do not depend on the details of the detector anymore.

As identical fit performance has been found and due to the higher relevance, the fiducial parametrization is chosen as the nominal one and used in any further studies in this chapter.

**Table 10.12:** Fit results using the smoothed nuisance parameter distributions in the configuration with three independent fit parameters for the fiducial fraction parametrization.

| | SR $100 < p_T^Z \leq 200\,\mathrm{GeV}$ | | SR $p_T^Z > 200\,\mathrm{GeV}$ | |
|---|---|---|---|---|
| | Exp. | Obs. | Exp. | Obs. |
| Significance | $4.285\,\sigma$ | $5.158\,\sigma$ | $2.493\,\sigma$ | $1.595\,\sigma$ |
| $f_{00}$ | $0.15\pm^{0.04}_{0.04}$ | $0.19\pm^{0.04}_{0.04}$ | $0.23\pm^{0.10}_{0.10}$ | $0.13\pm^{0.09}_{0.08}$ |
| $f_{Mixed}$ | $0.23\pm^{0.08}_{0.09}$ | $0.18\pm^{0.09}_{0.10}$ | $0.12\pm^{0.20}_{0.12}$ | $0.23\pm^{0.18}_{0.20}$ |
| $N_{tot}$ | $2500\pm^{120}_{120}$ | $2690\pm^{150}_{150}$ | $330\pm^{29}_{28}$ | $355\pm^{30}_{29}$ |
| $f_{TT}$ | $0.62 \pm 0.06$ | $0.63 \pm 0.06$ | $0.65 \pm 0.09$ | $0.64 \pm 0.13$ |
| $\mu$ | $1.00 \pm 0.25$ | $1.33 \pm 0.29$ | $1.0 \pm 0.4$ | $0.6 \pm 0.4$ |
| $\mu_{WZTT}$ | $1.00 \pm 0.11$ | $1.09 \pm 0.12$ | $1.00 \pm 0.17$ | $1.06 \pm 0.23$ |
| $\mu_{WZMixed}$ | $1.0 \pm 0.4$ | $0.9 \pm 0.4$ | $1.0 \pm 1.3$ | $2.1 \pm 1.7$ |

## 10.5 Impact of Number of Free Parameters

With each boson having three possible polarization states $(+, -, 0)$, there are a total of 9 possible states for the case with two bosons. The Monte Carlo generator used for the signal samples of this work is not able to generate the $+$ and $-$ states separately. Instead, only a transverse polarization, which includes both of these, can be defined. This now leaves four states for diboson events.

The most general physics models that could be investigated using such samples scale each of these four states separately. Other models might treat the 0T and T0 states the same. For such a model, only three free parameters are needed. Even more restrictive ones might only affect the 00 state or the TT, 0T, and T0 states in unison, such that their relative sizes to each other remain like the SM ones. This scenario leaves two free parameters.

While more free parameters allow the probing of wider physics scenarios, they are also more difficult to investigate experimentally. As such, all three cases are investigated here. The results for fits using three free parameters have already been shown in Table 10.12. Tables 10.13 and 10.14 show the ones using two and four free parameters, respectively.

**Table 10.13:** Fit results using the smoothed nuisance parameter distributions in the configuration with two independent fit parameters for the fiducial fraction parametrization.

| | SR $100 < p_T^Z \leq 200\,\mathrm{GeV}$ | | SR $p_T^Z > 200\,\mathrm{GeV}$ | |
|---|---|---|---|---|
| | Exp. | Obs. | Exp. | Obs. |
| Significance | $6.859\,\sigma$ | $7.726\,\sigma$ | $4.225\,\sigma$ | $3.193\,\sigma$ |
| $f_{00}$ | $0.152\pm^{0.023}_{0.023}$ | $0.174\pm^{0.024}_{0.025}$ | $0.23\pm^{0.06}_{0.06}$ | $0.16\pm^{0.06}_{0.06}$ |
| $N_{tot}$ | $2500\pm^{120}_{120}$ | $2680\pm^{140}_{150}$ | $330\pm^{29}_{28}$ | $355\pm^{30}_{29}$ |
| $f_{Rest}$ | $0.848 \pm 0.023$ | $0.826 \pm 0.024$ | $0.77 \pm 0.06$ | $0.84 \pm 0.06$ |
| $\mu$ | $1.00 \pm 0.16$ | $1.23 \pm 0.18$ | $1.00 \pm 0.27$ | $0.76 \pm 0.27$ |
| $\mu_{WZRest}$ | $1.00 \pm 0.06$ | $1.05 \pm 0.06$ | $1.00 \pm 0.11$ | $1.17 \pm 0.12$ |

**Table 10.14:** Fit results using the smoothed nuisance parameter distributions in the configuration with four independent fit parameters for the fiducial fraction parametrization.

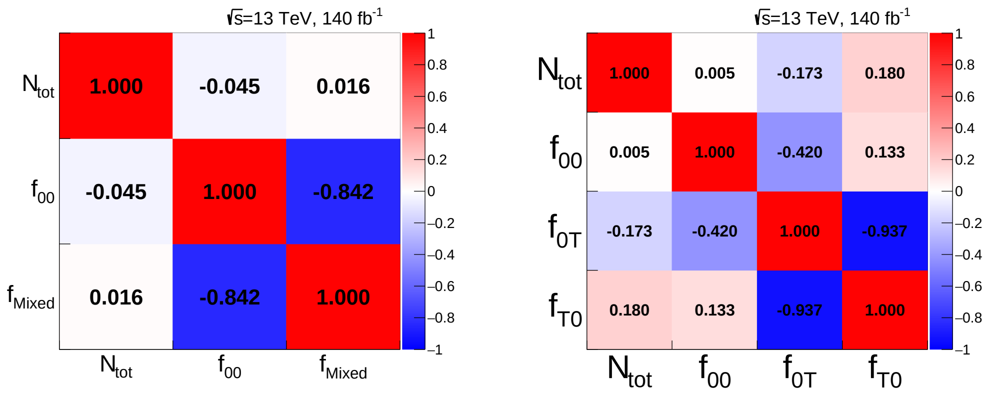
| | SR $100 < p_T^Z \leq 200\,\mathrm{GeV}$ | | SR $p_T^Z > 200\,\mathrm{GeV}$ | |
|---|---|---|---|---|
| | Exp. | Obs. | Exp. | Obs. |
| Significance | $3.875\,\sigma$ | $5.048\,\sigma$ | $2.279\,\sigma$ | $1.704\,\sigma$ |
| $f_{00}$ | $0.15\pm^{0.04}_{0.04}$ | $0.19\pm^{0.04}_{0.04}$ | $0.23\pm^{0.10}_{0.10}$ | $0.13\pm^{0.08}_{0.07}$ |
| $f_{0T}$ | $0.12\pm^{0.21}_{0.12}$ | $0.00\pm^{0.11}_{0.00}$ | $0.06\pm^{0.27}_{0.06}$ | $0.00\pm^{0.31}_{0.00}$ |
| $f_{T0}$ | $0.11\pm^{0.15}_{0.11}$ | $0.18\pm^{0.07}_{0.11}$ | $0.06\pm^{0.23}_{0.06}$ | $0.24\pm^{0.16}_{0.24}$ |
| $N_{tot}$ | $2500\pm^{120}_{120}$ | $2690\pm^{150}_{140}$ | $330\pm^{31}_{30}$ | $360\pm^{31}_{29}$ |
| $f_{TT}$ | $0.62 \pm 0.04$ | $0.63 \pm 0.09$ | $0.65 \pm 0.06$ | $0.63 \pm 0.21$ |
| $\mu$ | $1.00 \pm 0.27$ | $1.31 \pm 0.27$ | $1.0 \pm 0.5$ | $0.6 \pm 0.4$ |
| $\mu_{WZTT}$ | $1.00 \pm 0.08$ | $1.10 \pm 0.16$ | $1.00 \pm 0.13$ | $1.1 \pm 0.4$ |
| $\mu_{WZ0T}$ | $1.0 \pm 1.4$ | $0.0 \pm 0.5$ | $1.0 \pm 2.6$ | $0.0 \pm 2.7$ |
| $\mu_{WZT0}$ | $1.0 \pm 1.2$ | $1.8 \pm 0.9$ | $1.0 \pm 2.6$ | $4 \pm 4$ |

Larger significances and differences in the observed fractions can be seen in the two-parameter configuration compared to the three-parameter one. This is due to the mixed polarization states being closer in shape to the 00 ones, as was shown in Figure 8.10. Combining these with the TT state leaves less room to compensate for the missing contribution of the 00 state in the background-only fits.

Mostly, the opposite effect can be seen for the configuration where the T0 and 0T states are scaled separately. Here, a decrease of three of the four significances can be seen. The decrease in the fits to Asimov data is expected, as the additional freedom can improve the background-only fit, while the S+B one can not be improved from its already perfect state. Whether the significance of fits to observed data increases or decreases depends on the exact data and whether the additional freedom offers larger improvements to the B or S+B fits.

The more interesting phenomenon is the fact that the observed $f_{0T}$ fractions are 0 in both of the 00-enriched regions, while the sum of the mixed fractions is very close to the one measured in the configuration where they are combined: $f_{0T} + f_{T0} \approx f_{Mixed}$. This comes from the fact that their shapes are almost completely identical. With no other way of constraining the mixed states individually, the one that matches the data even just slightly better takes up all of the contribution for the other. This is exemplified by an expected stat. only correlation coefficient of $-0.937$ between the T0 and 0T states in the region with $100 < p_T^Z \leq 200\,\text{GeV}$. The observed ones can not be calculated by Hesse due to the best-fit value for $f_{0T}$ being exactly 0. This also causes the unusually large uncertainty on the implicit $f_{TT}$ fraction, as the correlation between the mixed states is not taken into account properly in the subtraction.

The expected correlations between the parameters in the stat. only fits for the three- and four-parameter configurations in the $100 < p_T^Z \leq 200\,\text{GeV}$ region are shown in Figure 10.30. It shows that a large correlation remains between $f_{00}$ and $f_{Mixed}$ in the three-parameter configuration. However, with an expected value of $-0.842$ for the stat. only fit, it is still smaller than the one observed in the four-parameter one. Additionally, the four-parameter configuration also shows significant (anti-)correlations between most of the other parameters except for $f_{00}$ and $N_{tot}$.




(a) (b)

**Figure 10.30:** Expected correlations in the stat. only fit in the region with $100 < p_T^Z \leq 200\,\text{GeV}$ using the fiducial fraction parametrization with (a) three independent parameters and (b) four independent parameters.

Due to this, the four-parameter configuration is not used as a nominal, and only the two- and three-parameter ones are kept for that. An additional issue with separating T0 and 0T is the strong dependence of the result and uncertainties on the choice of which fraction is treated as implicit. This is shown in Appendix E.

## 10.6 Equivalence of Parametrizations

The last step in verifying that the parametrization works correctly and does not impact the results outside of the fit's internal accuracy is a comparison between the fraction and $\mu$ parameterizations. For this, fits using the $\mu$ parametrization are performed using the smoothed distributions with two and three free parameters in both 00-enriched regions and compared against the equivalent fiducial fraction parametrization as already shown in Tables 10.12 and 10.13. These results are presented in Tables 10.15 and 10.16.

Perfect agreement can be observed for the significances between the two parametrizations. The best-fit $\mu$ values also agree with those calculated from the fraction fits to within the fit precision. The uncertainties determined directly by the fit using the signal strength parametrization and those propagated from the fraction one also agree within expectations.

All of this indicates that the reparametrization from signal strengths to fractions is successful and has no adverse impact on the measurement. Instead, it allows a direct and

**Table 10.15:** Fit results using the smoothed nuisance parameter distributions in the configuration with two independent fit parameters for the $\mu$ parametrization.

| | SR $100 < p_T^Z \leq 200$ GeV | | SR $p_T^Z > 200$ GeV | |
|---|---|---|---|---|
| | Exp. | Obs. | Exp. | Obs. |
| Significance | $6.859\,\sigma$ | $7.726\,\sigma$ | $4.225\,\sigma$ | $3.193\,\sigma$ |
| $\mu$ | $1.00\pm^{0.16}_{0.16}$ | $1.23\pm^{0.18}_{0.18}$ | $1.00\pm^{0.27}_{0.26}$ | $0.76\pm^{0.27}_{0.26}$ |
| $\mu_{WZRest}$ | $1.00\pm^{0.06}_{0.06}$ | $1.05\pm^{0.07}_{0.07}$ | $1.00\pm^{0.12}_{0.11}$ | $1.17\pm^{0.13}_{0.12}$ |

**Table 10.16:** Fit results using the smoothed nuisance parameter distributions in the configuration with three independent fit parameters for the $\mu$ parametrization.

| | SR $100 < p_T^Z \leq 200$ GeV | | SR $p_T^Z > 200$ GeV | |
|---|---|---|---|---|
| | Exp. | Obs. | Exp. | Obs. |
| Significance | $4.285\,\sigma$ | $5.158\,\sigma$ | $2.493\,\sigma$ | $1.595\,\sigma$ |
| $\mu$ | $1.00\pm^{0.26}_{0.25}$ | $1.33\pm^{0.30}_{0.29}$ | $1.0\pm^{0.4}_{0.4}$ | $0.6\pm^{0.4}_{0.4}$ |
| $\mu_{WZTT}$ | $1.00\pm^{0.12}_{0.11}$ | $1.09\pm^{0.13}_{0.12}$ | $1.00\pm^{0.23}_{0.21}$ | $1.06\pm^{0.26}_{0.22}$ |
| $\mu_{WZMixed}$ | $1.0\pm^{0.4}_{0.4}$ | $0.9\pm^{0.4}_{0.5}$ | $1.0\pm^{1.7}_{1.0}$ | $2.1\pm^{1.6}_{1.8}$ |

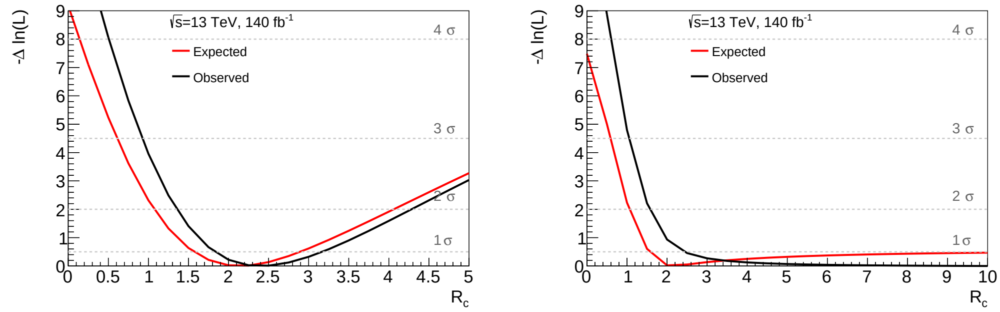
precise determination of the values and uncertainties of interest without relying on propagation afterward.

## 10.7 Results of Independence Parametrization

The reparametrization to the independence parameters, as introduced in Section 9.3.1, offers an interesting opportunity to find evidence for the existence of correlations between the two bosons' polarization states. This is especially the case, as the fully inclusive analysis in Ref. [26] only reached a value of $1.6\sigma$. However, there is one main problem with this measurement. The lack of differentiation between the mixed polarization states. This issue is also discussed in Appendix E. Here, it becomes visible in the results of the fit using the independence parametrization with four free parameters as shown in Table 10.17. The chosen parametrization can not ensure that both mixed fractions remain in the physically valid region of $[0, 1]$. As such, the resulting significances do not describe a real physics measurement and are thus not quoted for the fits to data.

**Table 10.17:** Fit results using the smoothed nuisance parameter distributions in the configuration with four independent fit parameters for the independence parametrization.

| | SR $100 < p_T^Z \leq 200$ GeV | | SR $p_T^Z > 200$ GeV | |
|---|---|---|---|---|
| | Exp. | Obs. | Exp. | Obs. |
| Significance | $2.114\,\sigma$ | - | $1.555\,\sigma$ | - |
| $R_c$ | $2\pm^{8}_{1}$ | $10\pm^{0}_{8}$ | $3\pm^{7}_{1}$ | $10\pm^{0}_{9}$ |
| $f_{W0}$ | $0.27\pm^{0.24}_{0.24}$ | $0.04\pm^{0.18}_{0.01}$ | $0.3\pm^{0.4}_{0.6}$ | $0.031\pm^{0.019}_{0.017}$ |
| $f_{Z0}$ | $0.26\pm^{0.21}_{0.21}$ | $0.49\pm^{0.04}_{0.16}$ | $0.3\pm^{0.3}_{0.6}$ | $0.47\pm^{0.11}_{0.34}$ |
| $N_{tot}$ | $2500\pm^{120}_{120}$ | $2700\pm^{140}_{140}$ | $330\pm^{35}_{33}$ | $364\pm^{31}_{30}$ |
| $\mu$ | $1.0 \pm 1.9$ | $1.4 \pm 2.7$ | $1.0 \pm 1.4$ | $0.7 \pm 0.5$ |
| $\mu_{WZTT}$ | $1.0 \pm 0.5$ | $1.2 \pm 0.5$ | $1.0 \pm 0.6$ | $1.10 \pm 0.29$ |
| $\mu_{WZ0T}$ | $1.0 \pm 2.5$ | $-1.4 \pm 2.6$ | $1 \pm 8$ | $-2.0 \pm 1.9$ |
| $\mu_{WZT0}$ | $1.0 \pm 2.5$ | $3 \pm 4$ | $1 \pm 7$ | $6.1 \pm 3.2$ |




(a) (b)

**Figure 10.31:** Likelihood scans of $R_c$ in the (a) three-parameter and (b) four-parameter configuration.

In the other parametrizations, this issue is circumvented by moving to the three-parameter configuration. The corresponding results are shown in Table 10.18. While the results are now physically valid, it is difficult to interpret them in the sense of rejecting the hypothesis of uncorrelated boson polarizations when this is explicitly a constraint on the fit through the merging of $f_{W0}$ and $f_{Z0}$.

**Table 10.18:** Fit results using the smoothed nuisance parameter distributions in the configuration with three independent fit parameters for the independence parametrization.

| | SR $100 < p_T^Z \leq 200\,\mathrm{GeV}$ | | SR $p_T^Z > 200\,\mathrm{GeV}$ | |
|---|---|---|---|---|
| | Exp. | Obs. | Exp. | Obs. |
| Significance | $2.153\,\sigma$ | $2.815\,\sigma$ | $1.564\,\sigma$ | $0.828\,\sigma$ |
| $R_c$ | $2.1\pm^{0.8}_{0.6}$ | $2.4\pm^{0.8}_{0.6}$ | $2.7\pm^{2.1}_{1.1}$ | $2.2\pm^{2.4}_{1.4}$ |
| $f_{WZ0}$ | $0.267\pm^{0.024}_{0.026}$ | $0.281\pm^{0.025}_{0.028}$ | $0.29\pm^{0.06}_{0.06}$ | $0.25\pm^{0.05}_{0.06}$ |
| $N_{tot}$ | $2500\pm^{120}_{120}$ | $2690\pm^{150}_{150}$ | $330\pm^{29}_{28}$ | $355\pm^{30}_{29}$ |
| $\mu$ | $1.00 \pm 0.26$ | $1.33 \pm 0.30$ | $1.0 \pm 0.5$ | $0.6 \pm 0.5$ |
| $\mu_{WZTT}$ | $1.00 \pm 0.08$ | $1.09 \pm 0.09$ | $1.00 \pm 0.21$ | $1.06 \pm 0.19$ |
| $\mu_{WZMixed}$ | $1.0 \pm 0.4$ | $0.9 \pm 0.5$ | $1.0 \pm 2.0$ | $2.1 \pm 2.0$ |

The likelihood scans for $R_c$, shown in Figure 10.31, depict the freedom the parameter gets towards higher values when the physicality condition for the mixed fractions is lifted.

## 10.7.1 Ensuring Physical Validity

There are three avenues to resolve this issue. Due to the low statistics and resulting poor sensitivity in the $p_T^Z > 200\,\mathrm{GeV}$ region, they are only investigated for the region with $100 < p_T^Z \leq 200\,\mathrm{GeV}$.

### Additional Constraints

One option is to set boundaries on all the explicit parameters, such that the unphysical regions are impossible. This is, however, very restrictive and also excludes a lot of physical regions. Another option is to add constraints to the likelihood, which penalize unphysical regions. However, these are found to lead to instabilities in the fit if the constraints are set strict. If they are set looser, the result can still lie in the unphysical region and only the extent is reduced. This approach was not pursued further due to this.

**Decay angle categories**

Another option is the use of decay angle channels, as briefly mentioned in Section 9.3.3 and implemented in Ref. [26, 64]. The approach was not used in the nominal measurement as the unphysical regions are easily mitigated in the fraction parametrization, no increase in the sensitivity on $f_{00}$ was found, and to be consistent between the two signal regions. As the $p_T^Z > 200\,\text{GeV}$ region is not considered for this measurement, and because better constraints on the mixed polarizations are pertinent for this configuration, the decay angle splitting is summarized and investigated. The splitting is performed based on the theoretical predictions that the decay angles of longitudinally polarized bosons are usually close to $|\cos\theta_\ell| = 0$, while transversally polarized ones tend towards $|\cos\theta_\ell| = 1$. This is referred to in Section 2.2.4 and visible in Figure 6.13. To make use of this, the events are split at values of $|\cos\theta_\ell| = 0.5$ to produce four independent channels. A visualization of these channels and the polarization states enhanced in each is shown in Figure 10.32. The respective pre-fit yields are summarized in Table 10.19. It can be seen that the combination of the individual channels agrees with the yields presented in Table 10.1. Additionally, the ratio of the yields for the mixed polarizations $\frac{W_0 Z_T}{W_T Z_0}$ reaches values of 2.6 and 0.5 in the $0T$ and $T0$ channels respectively. This is compared to the value 1.1 in the whole region. Due to this, an increase in one state, coupled with a decrease in the other, leads to different expected yields in these two channels, while the overall yield in the whole region stays the same.

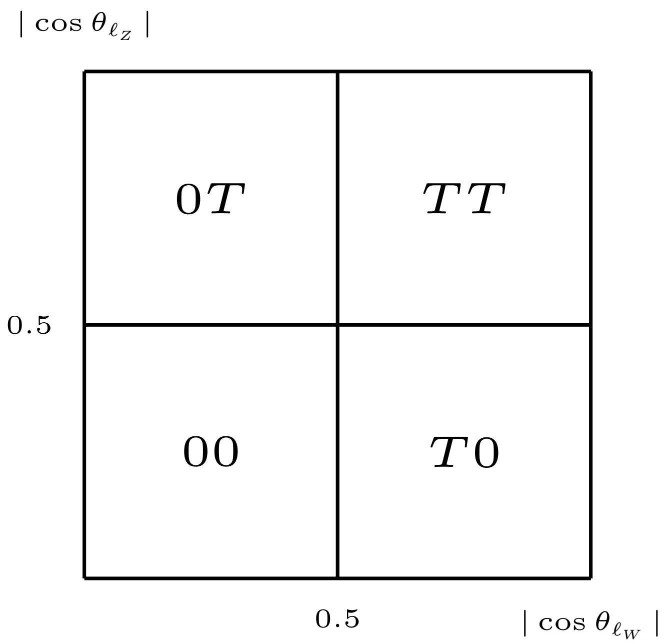



**Figure 10.32:** Visualization of the four channels in the decay angle splitting. Each channel is enhanced in the shown polarization combinations.

The fit results utilizing this splitting are shown in Table 10.20 and Figure 10.33. It can be seen that the use of the decay angle channels significantly improves the uncertainties of $f_{W0}$ and $f_{Z0}$ from more than 0.2 down to 0.06 in the expected fit. The observed fit results are also stabilized, with the best-fit value for $R_c$ now laying at $3.1\pm^{1.7}_{0.9}$ instead of $10\pm^{0}_{8}$. However, despite these improvements, the nominal fit results still reside within the unphysical region. As such, no observed significance is quoted for this setup either.

**Table 10.19:** Number of events for the expected signal and background as well as observed data in the $100 < p_T^Z \leq 200\,\text{GeV}$ region. They are shown for each of the four decay angle channels as well as the combination of them. The uncertainties shown include both statistical and systematic contributions.

| Channel | 00 | $0T$ | $T0$ | $TT$ | all |
|---|---|---|---|---|---|
| $W_0Z_0$ | $110.8 \pm 2.5$ | $42.8 \pm 1.4$ | $48.0 \pm 1.2$ | $20.0 \pm 0.7$ | $222 \pm 4$ |
| $W_0Z_T$ | $56.3 \pm 2.8$ | $58.4 \pm 2.4$ | $26.9 \pm 1.3$ | $28.6 \pm 1.5$ | $170 \pm 7$ |
| $W_TZ_0$ | $56.0 \pm 2.1$ | $22.5 \pm 1.1$ | $52.3 \pm 2.6$ | $21.5 \pm 1.3$ | $152 \pm 6$ |
| $W_TZ_T$ | $214 \pm 6$ | $237 \pm 8$ | $192 \pm 7$ | $213 \pm 8$ | $856 \pm 27$ |
| Prompt | $38 \pm 4$ | $34 \pm 4$ | $50 \pm 6$ | $46 \pm 6$ | $169 \pm 18$ |
| Non-prompt | $12 \pm 7$ | $17 \pm 8$ | $17 \pm 7$ | $22 \pm 7$ | $68 \pm 29$ |
| Total Expected | $487 \pm 14$ | $412 \pm 15$ | $387 \pm 13$ | $351 \pm 13$ | $1640 \pm 50$ |
| Data | 481 | 424 | 449 | 386 | 1740 |

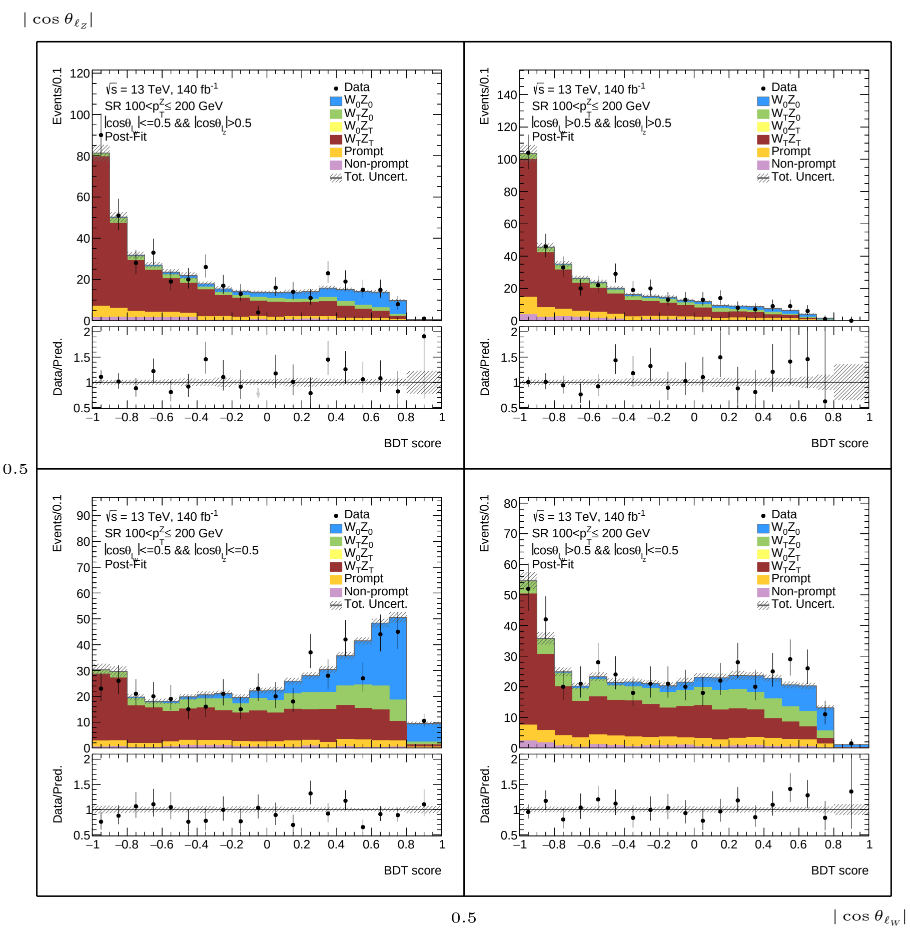



**Figure 10.33:** Observed post-fit distributions in the $100 < p_T^Z \leq 200\,\text{GeV}$ region for the four-parameter independence parametrization utilizing the decay angle splitting. Each distribution is shown in its respective channel.

**Table 10.20:** Fit results for the $100 < p_T^Z \leq 200$ GeV region using the smoothed nuisance parameter distributions in the configuration with four independent fit parameters for the independence parametrization using the decay angle splitting.

| | Exp. | Obs. |
|---|---|---|
| Significance | $2.223\,\sigma$ | - |
| $R_c$ | $2.1\pm^{0.7}_{0.6}$ | $3.1\pm^{1.7}_{0.9}$ |
| $f_{W0}$ | $0.27\pm^{0.06}_{0.06}$ | $0.16\pm^{0.06}_{0.06}$ |
| $f_{Z0}$ | $0.26\pm^{0.05}_{0.05}$ | $0.36\pm^{0.05}_{0.05}$ |
| $N_{tot}$ | $2500\pm^{110}_{110}$ | $2650\pm^{120}_{120}$ |
| $\mu$ | $1.00 \pm 0.24$ | $1.23 \pm 0.29$ |
| $\mu_{WZTT}$ | $1.00 \pm 0.07$ | $1.12 \pm 0.14$ |
| $\mu_{WZLT}$ | $1.00 \pm 0.32$ | $-0.2 \pm 0.7$ |
| $\mu_{WZTL}$ | $1.00 \pm 0.34$ | $1.8 \pm 0.4$ |

### Lower Bound

The third option does not allow for a full rectification of the issue but provides a way to set a lower limit on the significance. For this approach to work, it is necessary to realize that a smaller value of $\ln \mathcal{L}(n|\hat{\mu}, \hat{\Theta})$ can only lead to a lower value of the significance. This can be seen from combining Equations 9.9 and 9.20:

$$Z_0 = \sqrt{2\left(\ln \mathcal{L}(n|\hat{R_c}, \hat{\Theta}) - \ln \mathcal{L}(n|R_c = 1, \hat{\hat{\Theta}})\right)}. \tag{10.5}$$

As such, a lower bound for the significance can be found through the use of the full background-only fit and a restricted, physicality-conserving S+B fit. With such a setup, the value substituting for $\mathcal{L}(n|\hat{R_c}, \hat{\Theta})$ can only be underestimated, while this is never the case for $\mathcal{L}(n|R_c = 1, \hat{\hat{\Theta}})$. Additionally, it is unlikely for unphysical fit results to emerge in the background-only fit in the cases where a full fit would produce a value of $R_c$ larger than one. This is because $f_{W0} \times f_{Z0}$ reaches its maximum for $f_{W0} = f_{Z0}$ in the case where both are bounded below by zero and their sum is constrained, and unphysical results are impossible for configurations where $R_c = 1$ and $f_{W0} = f_{Z0}$. The restricted option used for the S+B fit is the three-parameter one described above, where the ratio of $f_{W0}$ to $f_{Z0}$ is fixed at its expected value. This eliminates unphysical regions where either of the mixed signal strengths becomes negative and is a special case of the four-parameter configuration. As such, the resulting likelihood value can only be smaller than the full one.

The best-fit values of the parameters for the S+B fits are already shown in Table 10.18. The results of the background-only fits are given in Table 10.21. It can be seen that the values for the longitudinal fractions for $W$ and $Z$ bosons are similar and that no unphysical signal strengths are reached.

The resulting likelihood values for the used S+B and background-only fits are summarized in Table 10.22, while the post-fit distributions are shown in Figure 10.34. The expected lower bound of the significance for rejecting the $R_c = 1$ hypothesis is found to be $2.11\sigma$. The final value measured in data is $2.82\sigma$, falling short slightly of the $3\sigma$ required to claim evidence. This makes it clear, however, that additional data taken in ongoing and future runs of the LHC will likely make it possible to achieve this threshold, especially when considering the improved opportunity to use the decay angle splitting for a better separation of the mixed polarization states.

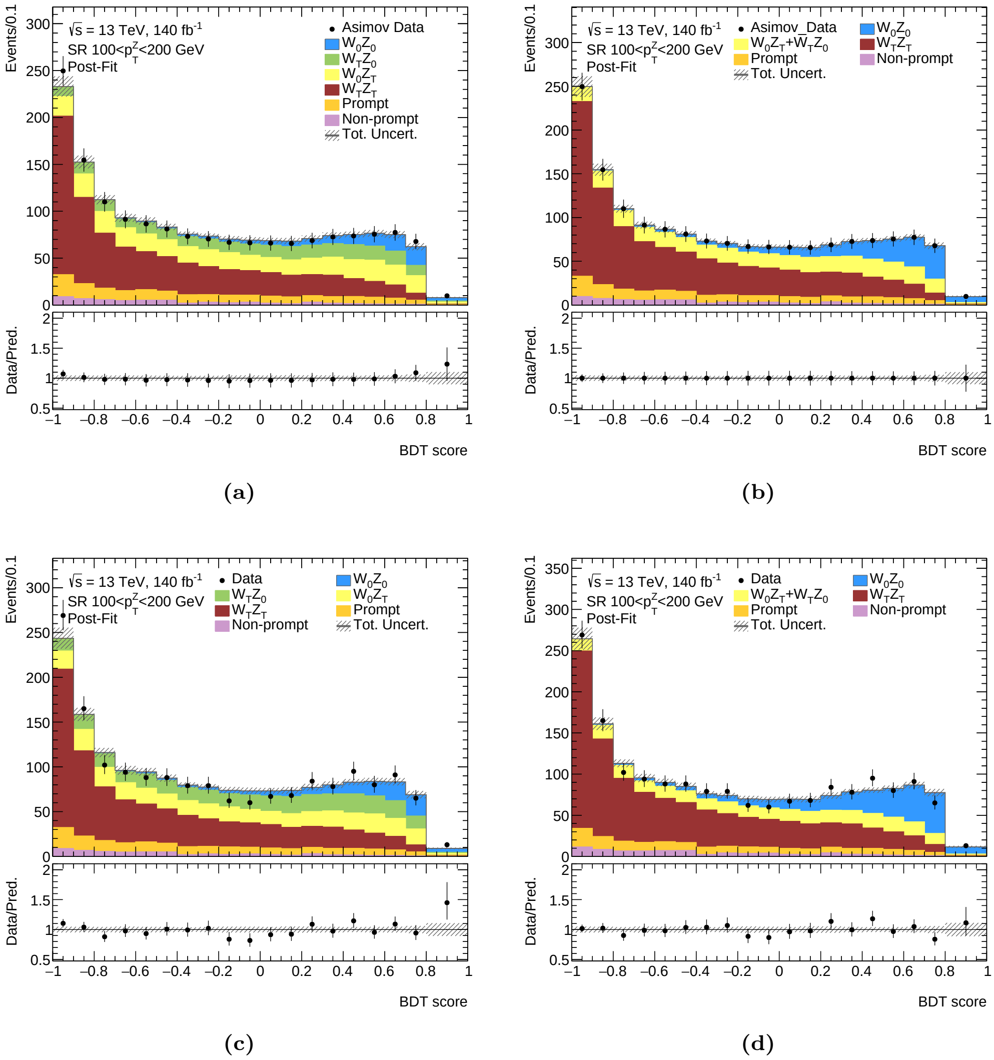



**Figure 10.34:** Expected (a) background-only and (b) S+B post-fit distributions and observed (c) background-only and (d) S+B post-fit distributions in the 00-enriched $100 < p_T^Z \leq 200\,\text{GeV}$ region for the independence parametrization. The background-only fits use four free parameters, while the S+B fits use three free parameters.

**Table 10.21:** Best fit values in the fits using the background-only hypothesis with smoothed distributions for the fiducial independence parametrization with four free parameters. Uncertainties omitted for brevity.

| | Expected | Observed |
|---|---|---|
| $R_c$ | 1.00 | 1.00 |
| $f_{W0}$ | 0.32 | 0.30 |
| $f_{Z0}$ | 0.24 | 0.29 |
| $N_{tot}$ | 2470 | 2650 |
| $\mu$ | 0.51 | 0.61 |
| $\mu_{WZTT}$ | 0.82 | 0.85 |
| $\mu_{WZ0T}$ | 2.03 | 1.91 |
| $\mu_{WZT0}$ | 1.47 | 1.95 |

**Table 10.22:** Log-likelihood values for the expected and observed S+B and background-only (B) fits, as well as the resulting lower bound of the significance for the independence parametrization. The S+B fits are performed with three free parameters, while the background-only ones use four free parameters.

| | S+B (3 Param.) | B (4 Param.) | $\sqrt{2\,(\ln \mathcal{L}(S+B) - \ln \mathcal{L}(B))}$ |
|---|---|---|---|
| Expected | 5813.57 | 5811.34 | 2.11 |
| Observed | 6281.04 | 6277.07 | 2.82 |

# Chapter 11

# Studies of the RAZ Effect

To investigate the radiation amplitude zero (RAZ) effect, distributions of rapidity differences are investigated as motivated in Chapter 2. If the radiation amplitude zero effect exists, fewer events with $\Delta Y(WZ) \approx 0$ are expected. This is specifically the case for the TT polarization state. Measuring this distribution with the ATLAS detector poses a significant challenge. The main contributing factor is the neutrino created during the leptonic decay of the $W$ boson. It escapes the detector without depositing any energy. To recover some of its kinematics, the missing transverse momentum is used as a proxy for $p_T^\nu$. The neutrino's momentum in the z-direction is approximated using the W boson mass constraint as described in Ref. [24]. However, both of these are just approximations and are not expected to reconstruct the neutrino momentum with full accuracy. This especially poses a challenge when very precise cancellations are expected. As such, the distribution measured with the reconstructed level observables is not expected to represent the underlying physics distribution accurately. This necessitates the use of unfolding to recover the fiducial level distributions from those measured with the ATLAS detector.

This chapter describes how this procedure works in Section 11.1. Section 11.2 describes the treatment of uncertainties in the unfolding, while the concrete inputs used for the work in this thesis are given in Section 11.3. Finally, Sections 11.4 and 11.5 present the results of the unfolding and describe how these distributions are used for studying the radiation amplitude zero effect.

## 11.1 Methodology for Unfolding

Unfolding is concerned with the relationship between the number of events observed in the detector and those at truth level. This can generally be described through the equation

$$n_i = \sum_j R_{ij} s_j + b_i. \tag{11.1}$$

Where $n_i$ is the total number of events observed in the $i$-th reconstructed bin, $s_j$ is the number of signal events in the $j$-th fiducial bin, $R_{ij}$ is the response matrix translating the signal between fiducial and reconstructed bins and $b_i$ is the number of background events in the $i$-th bin.

The response matrix combines three separate effects:

- Efficiency correction: Accounts for inefficiencies of the detector as well as migrations out of the signal region due to the detector resolution or imperfect information. It is defined as the ratio of events that pass the reconstructed and fiducial level selection over those that pass the fiducial one, regardless of whether or not they also pass the reconstructed one. It is split into bins at the fiducial level.

- Migration matrix: Describes migrations within the signal region between the fiducial and reconstructed level. It covers events that pass both the reconstructed and fiducial level signal regions. The content of bin $ij$ describes the probability that an event that is in truth bin $j$ ends up in reconstructed level bin $i$.

- Fiducial correction: Account for events migrating into the reconstructed signal region due to smearing effects. It is defined as the ratio of events passing the reconstructed and fiducial selections over those only required to pass the reconstructed selection. It is specified per reconstructed bin.

The effects' technical definitions are summarized in the following equations

$$\begin{aligned} f_i &= \frac{N^{R\cap T}_{reco,i}}{N^{R}_{reco,i}}, \\ \epsilon_j &= \frac{N^{R\cap T}_{truth,j}}{N^{T}_{truth,j}}, \\ M_{ij} &= \frac{N^{R\cap T}_{reco,i;truth,j}}{N^{R\cap T}_{truth,j}}, \end{aligned}$$

where $f_i, \epsilon_j, M_{ij}$ are the fiducial and efficiency corrections as well as the migration matrix. The $N$ represent number of events. The superscript indicates whether the events have to pass the truth selection, reconstructed selection, or both. The subscript specifies the bin and whether the reconstructed and truth value of the observable specifies the bin.

It is often useful to also define the quantities 'Stability' and 'Purity', which can be derived from the migration matrix. They are defined as [262]

$$S_j = \frac{N^{R\cap T}_{reco,j;truth,j}}{N^{R\cap T}_{truth,j}}, \tag{11.2}$$

$$P_i = \frac{N^{R\cap T}_{reco,i;truth,i}}{N^{R\cap T}_{reco,i}}. \tag{11.3}$$

The stability effectively describes the number of events that do not migrate out of a truth bin. Meanwhile, purity describes the number of events that did not migrate into a reconstruction-level bin.

All in all, this allows for a separation of the response matrix

$$R_{ij} = \frac{1}{f_i} \cdot M_{ij} \cdot \epsilon_j. \tag{11.4}$$

The whole process of folding and unfolding is visualized in Figure 11.1 by either going from the top left (truth selection) to the bottom right (reco selection) or doing the reverse. The goal of unfolding is now to obtain the truth distribution as the fiducial level, in the form of the $s_j$, from the measured distribution in the detector, the $n_i$. This can be achieved by inverting Equation 11.1

$$s_j = \sum_i R^{-1}_{ij} \left(n_i - b_i\right). \tag{11.5}$$

Multiple tools and approaches exist for performing these calculations. The one chosen in this work uses the VIPUnfolding tool [263], which is a wrapper around the RooUnfold toolkit [264]. The framework implements the determination of data statistical uncertainty using toys as specified in Ref. [262]. The author of this thesis made multiple significant changes to the framework in terms of bug fixes and feature enhancements. Merge requests for these into the main tool are pending or being prepared [265]. Additional improvements made by the author of Ref. [249] were also incorporated[1].

The unfolding itself is done through matrix inversion. Other approaches, such as Bin-By-Bin unfolding or regularized unfoldings like Iterative Bayesian Unfolding and Singular value

[1]These are summarized in the merge requests `https://gitlab.cern.ch/atlas-physics/sm/StandardModelTools_Unfolding/VIPUnfolding/-/merge_requests/8` and `https://gitlab.cern.ch/atlas-physics/sm/StandardModelTools_Unfolding/VIPUnfolding/-/merge_requests/11`

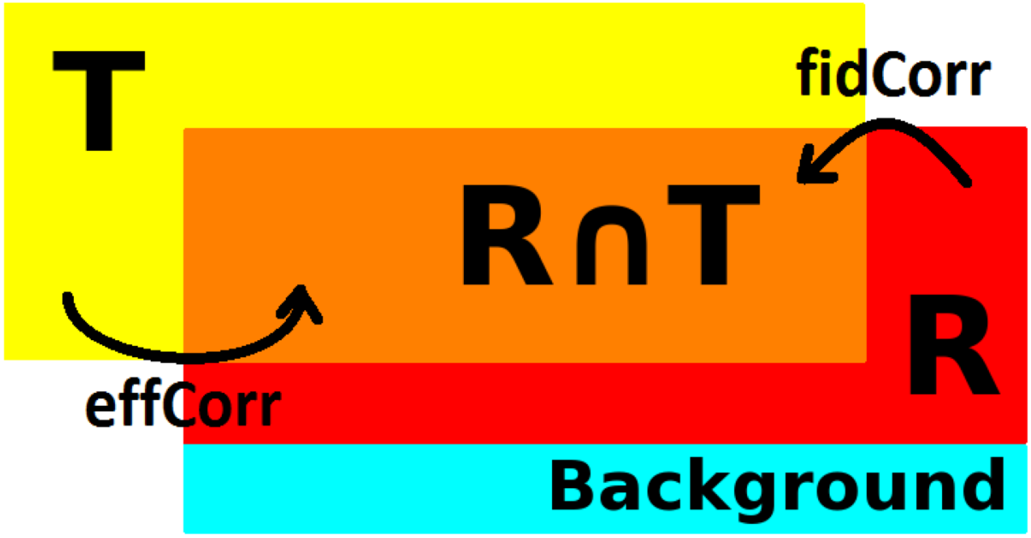



**Figure 11.1:** "Different selection levels of events. **R** contains all events passing the selection on reco-level (reco-selection), **T** contains all events passing the selection on truth-level (truth-selection). If an event passes both selection criteria, it is part of the joint phase space $\mathbf{R} \cap \mathbf{T}$. The background includes events which pass the selection on reco-level, but come from another process than the signal." From Ref. [249].

decomposition, exist and are discussed in more detail in Refs. [266, 267]. However, matrix inversion is the precise mathematical solution to the unfolding problem, can handle bin-to-bin migrations, and has no dependence on the prior fiducial input distribution.

As the unfolded distributions are intended for the study of the radiation amplitude zero effect, the overall normalization of the unfolded distribution is not relevant. As such, the unfolding performed normalized. This means that after unfolding, the distribution is normalized such that the sum over its bins is equal to 1. This is also the case for the treatment of uncertainties described in Section 11.2. There, the unfolded distributions corresponding to a systematic or statistical variation are normalized before any combination or calculation of bin-by-bin uncertainties is performed.

## 11.2 Treatment of Uncertainties

To assess the effects of uncertainties, the unfolding procedure is repeated with the variation effects applied coherently to all unfolding inputs. This includes efficiency and fiducial corrections and the migration matrix of the signal, as well as the background templates if any of these are affected. Any uncertainties that are not present for both signal and background use the nominal inputs for that component.

This produces a new unfolded distribution $N^k$ where $k$ describes the variation. The nominal distribution is referred to as $N^{nom}$. The related uncertainty distribution is then described by

$$\sigma^k = N^k - N^{nom}. \tag{11.6}$$

If an uncertainty is represented by individual up and down variations, then these are combined into a single uncertainty. Two different approaches for combining up and down uncertainties are used. For both combination options, the sign of the uncertainty distribution in a bin is determined by the up variation. The option used for most of the uncertainties is to average the up and down variations. In the case of the up-variation having a larger yield than the nominal, the resulting uncertainty in the $i$-th bin is

$$\sigma_i^k = \frac{\left|N_i^{k_{up}} - N_i^{nom}\right| + \left|N_i^{nom} - N_i^{k_{down}}\right|}{2}. \tag{11.7}$$

If the up variation has a smaller yield, then an additional factor of $-1$ is applied.

The other combination option, which is used for the flat background uncertainties, is to use the maximum variation value

$$\sigma_i^k = \max\left(\left|N_i^{k_{up}} - N_i^{nom}\right|, \left|N_i^{nom} - N_i^{k_{down}}\right|\right). \tag{11.8}$$

Again, a factor of $-1$ is applied if the up variation has a smaller yield than the nominal one.

The scale, shower, and PDF uncertainties follow similar approaches to match the ones described in Chapter 7. For the scale variations, the final uncertainty is taken as the envelope. Thus, each used variation is unfolded, and in each bin, the maximum deviation is chosen as the overall uncertainty with a positive sign. The shower uncertainties add the maxima of each variation's up and down components in quadrature. As only one variation has a significant impact, its respective maximum is used as the shower uncertainty. The PDF uncertainties are again combined using the standard deviation.

Statistical uncertainties from data, signal, and backgrounds are estimated separately using 2000 toys each. For data, its distribution is fluctuated around its nominal value using a Poisson distribution and then unfolded as usual. The uncertainty in each bin is calculated as

$$\sigma_i^{\text{stat}} = \sqrt{\frac{1}{2000-1}\sum_{j=1}^{2000}(N_i^{\text{stat},k} - N_i^{\text{stat,mean}})^2} \tag{11.9}$$

where

$$N_i^{\text{stat,mean}} = \frac{1}{2000}\sum_{j=1}^{2000} N_i^{\text{stat},k}. \tag{11.10}$$

The uncertainty on the background is estimated similarly, where the background is fluctuated using a Gaussian with the mean and uncertainties derived from the nominal background distribution. The fluctuated background is then subtracted from the data instead of the nominal one. The individual unfolded distributions are then combined to an uncertainty in the same way as for data. The same general approach is also employed for the estimation of the statistical uncertainty due to a finite signal sample. Here, the efficiency and fiducial corrections, as well as the migration matrix, are flucated around their nominal values using a Poisson distribution.

The combination of different uncertainty sources for the total uncertainty and category breakdowns is done through their quadrature sum

$$\sigma_{tot} = \sqrt{\sum_i \sigma_i^2}. \tag{11.11}$$

Covariance matrices are constructed for every uncertainty source via

$$\text{cov}_{ij}^k = (\sigma_i^k)(\sigma_j^k). \tag{11.12}$$

## 11.3 Unfolding Inputs

Individual unfoldings are performed for the $\Delta Y(WZ)$ and $\Delta Y(\ell_W Z)$ distributions in each of the RAZ regions described in Chapter 5, as well as the inclusive region without any cut on $p_T^{WZ}$. For each region, two sets of backgrounds and response matrices are used. One treats only the TT polarization as signal, as this is the one affected by the RAZ effect. The other considers all polarizations to be less model-dependent.

The first setup includes all non-WZ events as well as all non-TT WZ events as backgrounds. The fiducial and efficiency corrections, as well as the migration matrix, are then derived from the polarized MadGraph5_aMC@NLO samples with the TT polarization state. The individual event numbers are taken from the samples as produced directly by MadGraph and Pythia8 for the truth level and after the Geant4 simulation [159, 160] for the reconstructed level.

The resulting fiducial corrections are shown in Figure 11.2. It can be seen that there are only small differences between the correction for $\Delta Y(WZ)$ and $\Delta Y(\ell_W Z)$. This is partly because both are being considered in the same phase space, meaning that the overall corrections averaged over all bins are equal. Additionally, there are large correlations between the two variables. This was shown in the 00-enriched phase space in Figure 8.7. The correction factors are also largely independent of the concrete values for the rapidity differences, only showing slight reductions for very low and high absolute rapidity differences. The biggest changes are visible throughout the different regions. The tighter the cut on $p_T^{WZ}$, the lower the overall value of the correction factors. The factors for the bin $0.5 - 1.5$ change from roughly 0.9 in the inclusive region through $\approx 0.87, 0.85, 0.7$ down to 0.45 for the region with $p_T^{WZ} < 10\,\text{GeV}$. These arise from resolution effects of $p_T^{WZ}$ and thus changes in its value between the reconstructed and the truth level. This causes many events that pass the reconstructed level cut not to pass the fiducial level one.

A similar progression can be seen for the efficiency corrections shown in Figure 11.3. Here, the main difference is that the correction factors start at around 0.45 in the inclusive region and drop down to $\approx 0.3$ in the region with $p_T^{WZ} < 10\,\text{GeV}$. The drop is due to the same reason as the reduction in the values of the fiducial corrections. The overall lower value comes from the general lepton reconstruction and identification efficiencies around 80 % per lepton to pass the $Z$ or $W$ selection criteria described in Chapter 5, as well as from trigger efficiencies. The differences between $\Delta Y(\ell_W Z)$ and $\Delta Y(WZ)$ are again fairly minimal.

Significant differences, however, can be seen in the migration matrices and derived purities and stabilities as shown in Figures 11.4, 11.5 and 11.6. The migration matrices are almost perfectly diagonal for $\Delta Y(\ell_W Z)$, leading to purities and stabilities very close to one. This is not the case for the relationship between the truth and reconstructed values of $\Delta Y(\ell_W Z)$. Here, large migrations can be observed. The stability is highest for low values of the absolute rapidity differences with around 70 % and drops off for higher values, down to 40 %. For most bins, the majority of migrations happen towards lower absolute rapidity differences. The only exception is the bin around $|\Delta Y(\ell_W Z|$, where this is not possible. The migration matrices themselves, as well as the resulting stabilities, show almost no dependence on the value of the $p_T^{WZ}$ cut. This is somewhat different for the purity. Here, the main differences are a reduction of the purity of the bin $0.0 - 0.5$ from 40 % down to 20 %. The overall values are also in contrast to the fairly high stability observed in this bin. This effect is due to the nature of the RAZ effect, which reduces the number of events close to zero at the truth level. Due to this, even a moderate percentage of migration from the 0.5 to 1.5 bin down into the central one leads to a majority total contribution from migration for that bin. As the RAZ effect is expected to only exist at leading order and is thus stronger with tighter $p_T^{WZ}$ cuts, this reduction in purity for these tighter cuts is also expected.

The other configuration considers all polarized WZ events as signal. Thus, the contributions to the response matrix are derived from the sum of the polarized MadGraph samples. The background that gets subtracted from data at the beginning of the unfolding subsequently consists only of the prompt and non-prompt backgrounds that were already used in the fraction measurement. The resulting fiducial and efficiency corrections, as well as the migration matrices, purities and stabilities, are shown in Figures 11.7, 11.8, 11.9, 11.10 and 11.11.

The fiducial and efficiency corrections for this configuration are almost identical to those for unfolding only the TT state. This is expected as the TT state is the largest one, and the $p_T^{WZ}$ migrations and lepton selections should also be similar. The largest differences can be seen in the migration matrices and particularly the purities. Here, the purity of the leftmost bin starts at 50 % and never drops below 40 % as the contributions of the other polarization states in this bin serve to somewhat counteract the total impact of the migrations from the neighboring bin. Thus, the main differences in the final unfolded distribution are expected to mostly originate from the different background subtraction.

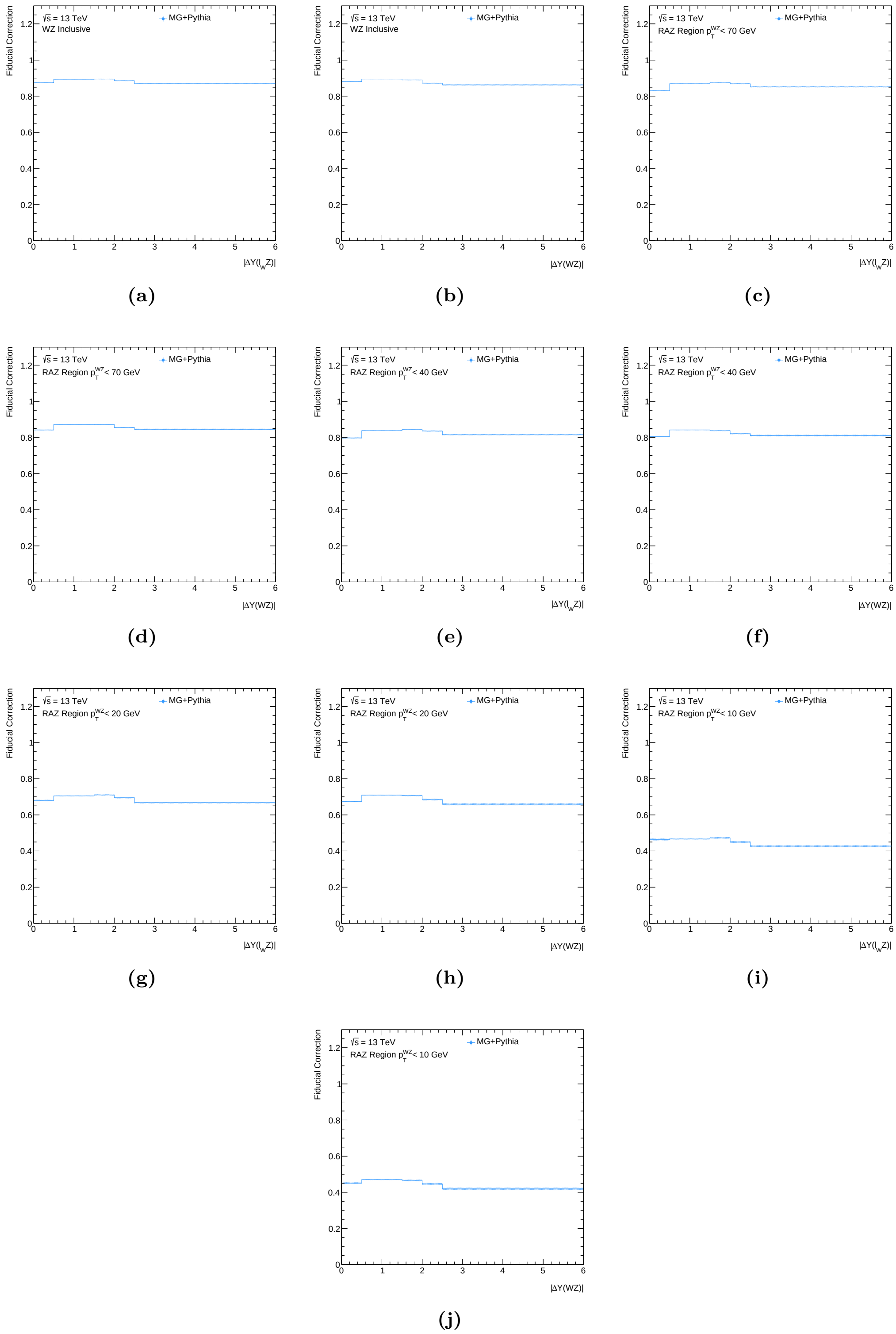
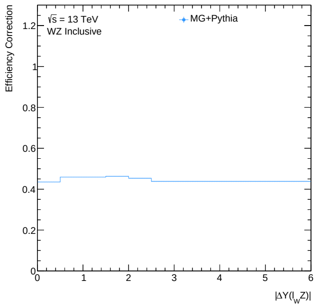
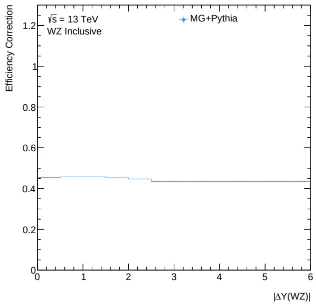


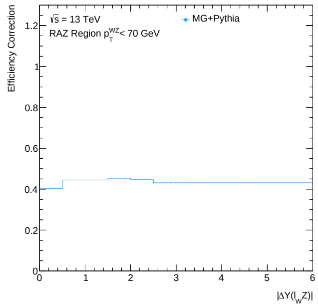

**Figure 11.2:** Fiducial correction inputs for the unfolding of the TT polarization state using $\Delta Y(\ell_W Z)$ in (a) the inclusive region without a cut on $p_T^{WZ}$ and the regions with (c) $p_T^{WZ} < 70\,\text{GeV}$, (e) $p_T^{WZ} < 40\,\text{GeV}$, (g) $p_T^{WZ} < 20\,\text{GeV}$, (i) $p_T^{WZ} < 10\,\text{GeV}$ and using $\Delta Y(WZ)$ in (b) the inclusive region, and the regions with (d) $p_T^{WZ} < 70\,\text{GeV}$, (f) $p_T^{WZ} < 40\,\text{GeV}$, (h) $p_T^{WZ} < 20\,\text{GeV}$, (j) $p_T^{WZ} < 10\,\text{GeV}$.



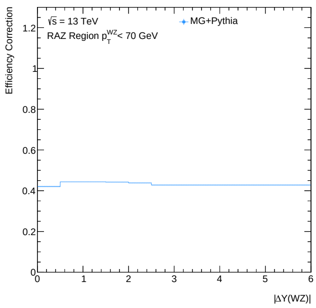

**(a)**



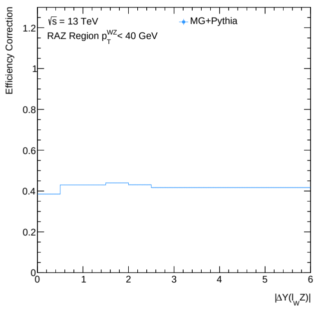

**(b)**



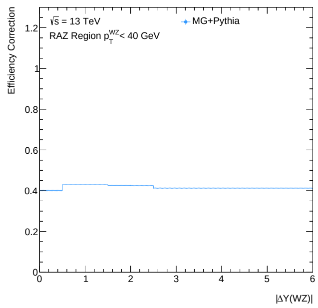

**(c)**



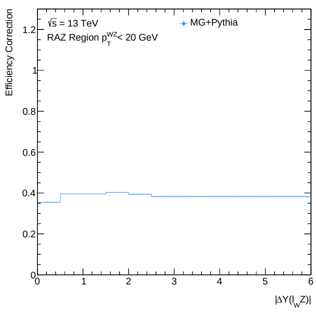

**(d)**



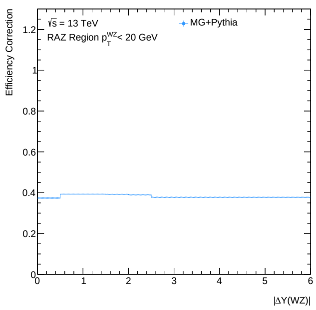

**(e)**



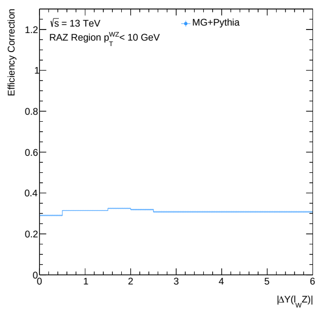

**(f)**



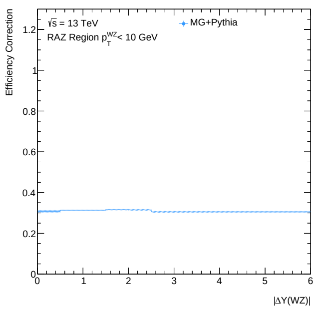

**(g)**




**(h)**




**(i)**




**(j)**

**Figure 11.3:** Efficiency correction inputs for the unfolding of the TT polarization state using $\Delta Y(\ell_W Z)$ in (a) the inclusive region without a cut on $p_T^{WZ}$ and the regions with (c) $p_T^{WZ} < 70\,\text{GeV}$, (e) $p_T^{WZ} < 40\,\text{GeV}$, (g) $p_T^{WZ} < 20\,\text{GeV}$, (i) $p_T^{WZ} < 10\,\text{GeV}$ and using $\Delta Y(WZ)$ in (b) the inclusive region, and the regions with (d) $p_T^{WZ} < 70\,\text{GeV}$, (f) $p_T^{WZ} < 40\,\text{GeV}$, (h) $p_T^{WZ} < 20\,\text{GeV}$, (j) $p_T^{WZ} < 10\,\text{GeV}$.

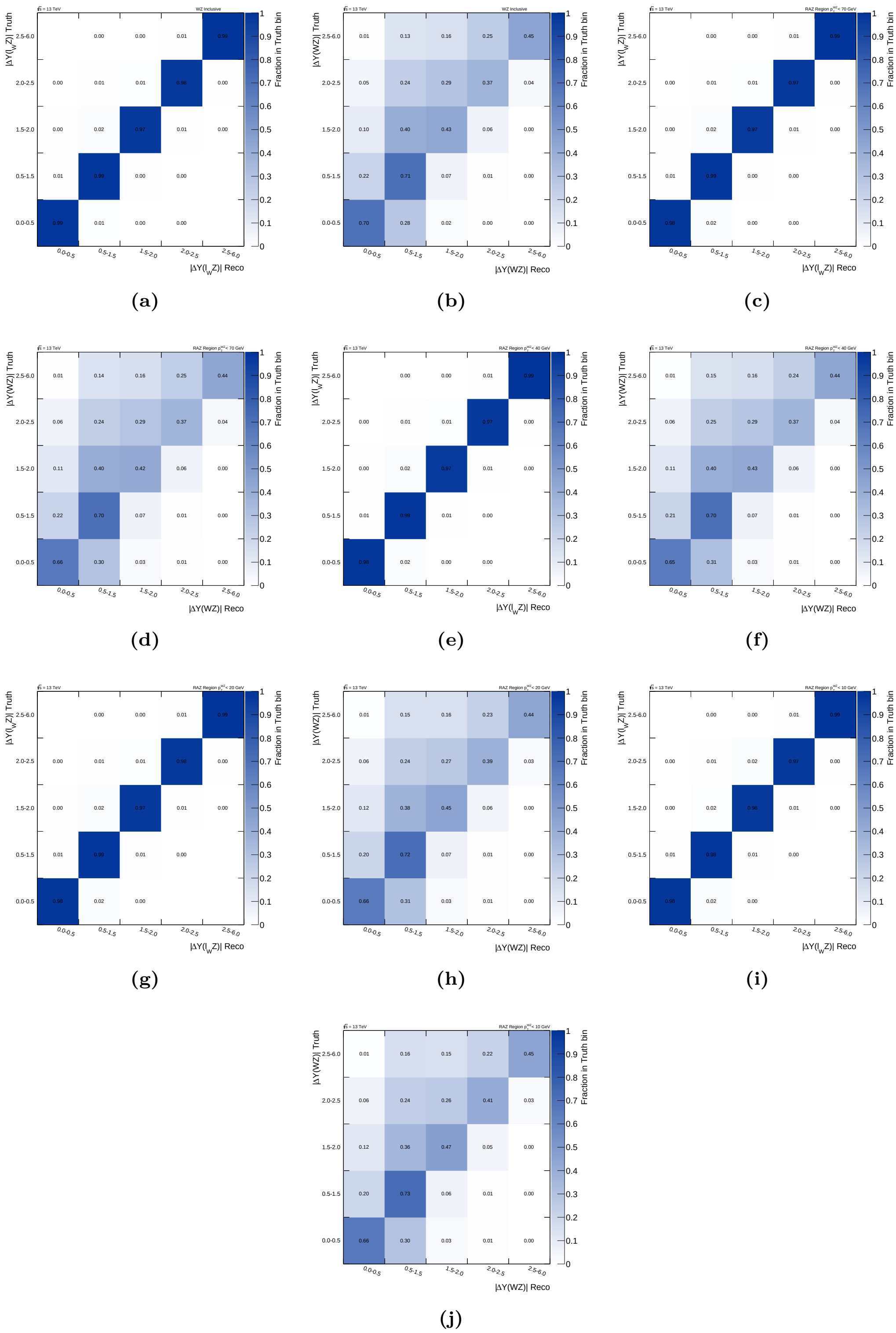



**Figure 11.4:** Migration matrix inputs for the unfolding of the TT polarization state using $\Delta Y(\ell_W Z)$ in (a) the inclusive region without a cut on $p_T^{WZ}$ and the regions with (c) $p_T^{WZ} < 70\,\mathrm{GeV}$, (e) $p_T^{WZ} < 40\,\mathrm{GeV}$, (g) $p_T^{WZ} < 20\,\mathrm{GeV}$, (i) $p_T^{WZ} < 10\,\mathrm{GeV}$ and using $\Delta Y(WZ)$ in (b) the inclusive region, and the regions with (d) $p_T^{WZ} < 70\,\mathrm{GeV}$, (f) $p_T^{WZ} < 40\,\mathrm{GeV}$, (h) $p_T^{WZ} < 20\,\mathrm{GeV}$, (j) $p_T^{WZ} < 10\,\mathrm{GeV}$.

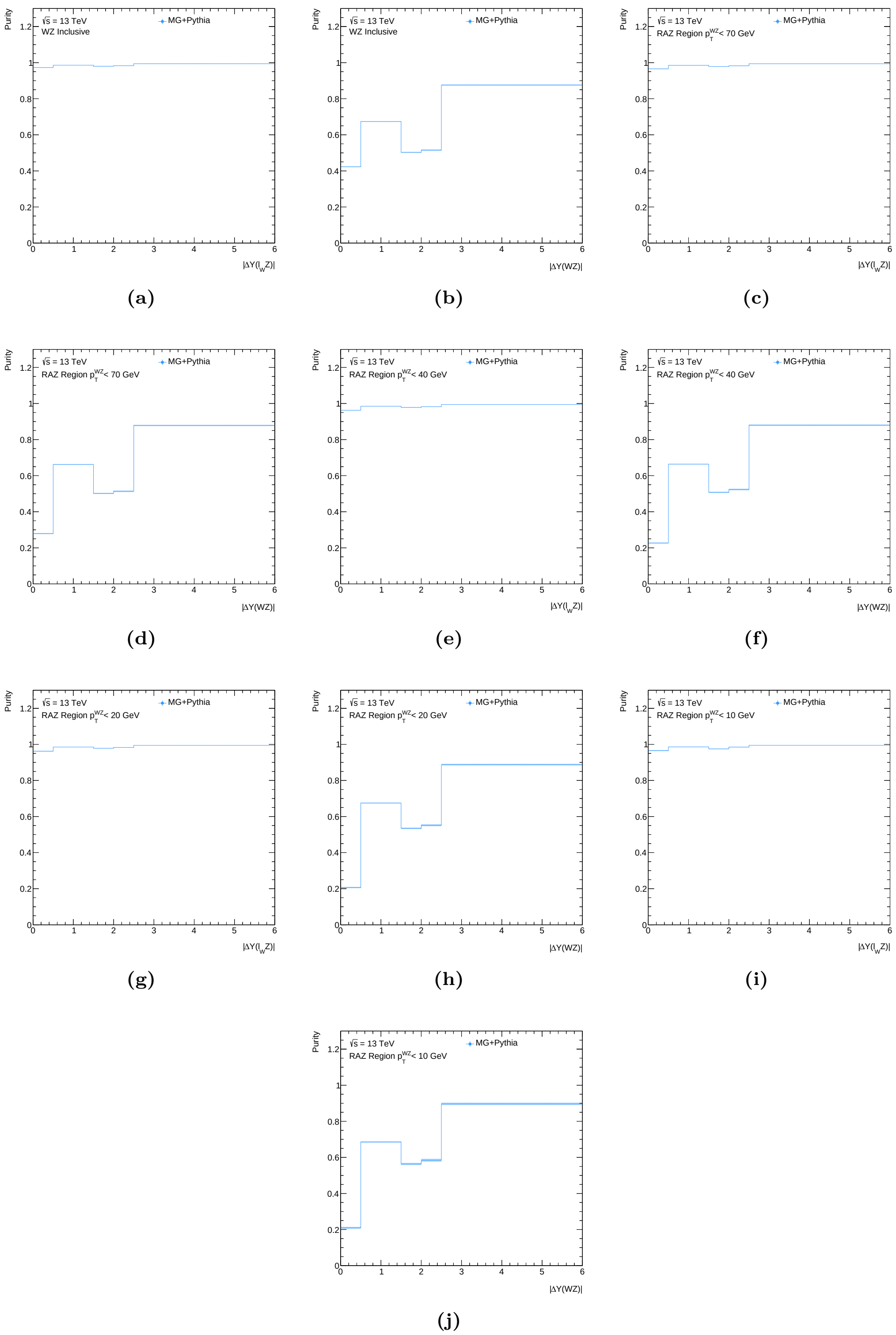



**Figure 11.5:** Purity for the unfolding of the TT polarization state using $\Delta Y(\ell_W Z)$ in (a) the inclusive region without a cut on $p_T^{WZ}$ and the regions with (c) $p_T^{WZ} < 70\,\text{GeV}$, (e) $p_T^{WZ} < 40\,\text{GeV}$, (g) $p_T^{WZ} < 20\,\text{GeV}$, (i) $p_T^{WZ} < 10\,\text{GeV}$ and using $\Delta Y(WZ)$ in (b) the inclusive region, and the regions with (d) $p_T^{WZ} < 70\,\text{GeV}$, (f) $p_T^{WZ} < 40\,\text{GeV}$, (h) $p_T^{WZ} < 20\,\text{GeV}$, (j) $p_T^{WZ} < 10\,\text{GeV}$.

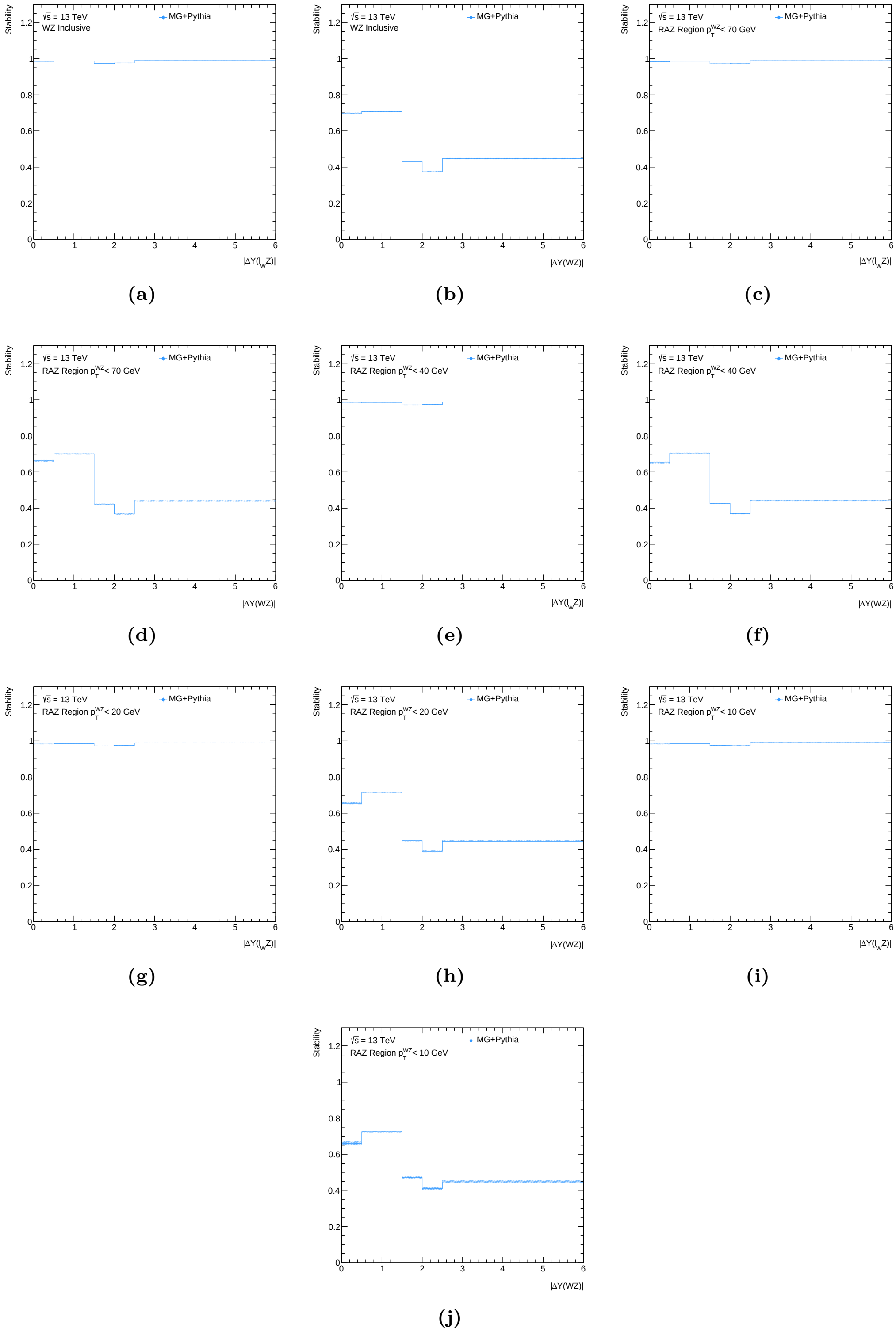



**Figure 11.6:** Stability for the unfolding of the TT polarization state using $\Delta Y(\ell_W Z)$ in (a) the inclusive region without a cut on $p_T^{WZ}$ and the regions with (c) $p_T^{WZ} < 70\,\text{GeV}$, (e) $p_T^{WZ} < 40\,\text{GeV}$, (g) $p_T^{WZ} < 20\,\text{GeV}$, (i) $p_T^{WZ} < 10\,\text{GeV}$ and using $\Delta Y(WZ)$ in (b) the inclusive region, and the regions with (d) $p_T^{WZ} < 70\,\text{GeV}$, (f) $p_T^{WZ} < 40\,\text{GeV}$, (h) $p_T^{WZ} < 20\,\text{GeV}$, (j) $p_T^{WZ} < 10\,\text{GeV}$.

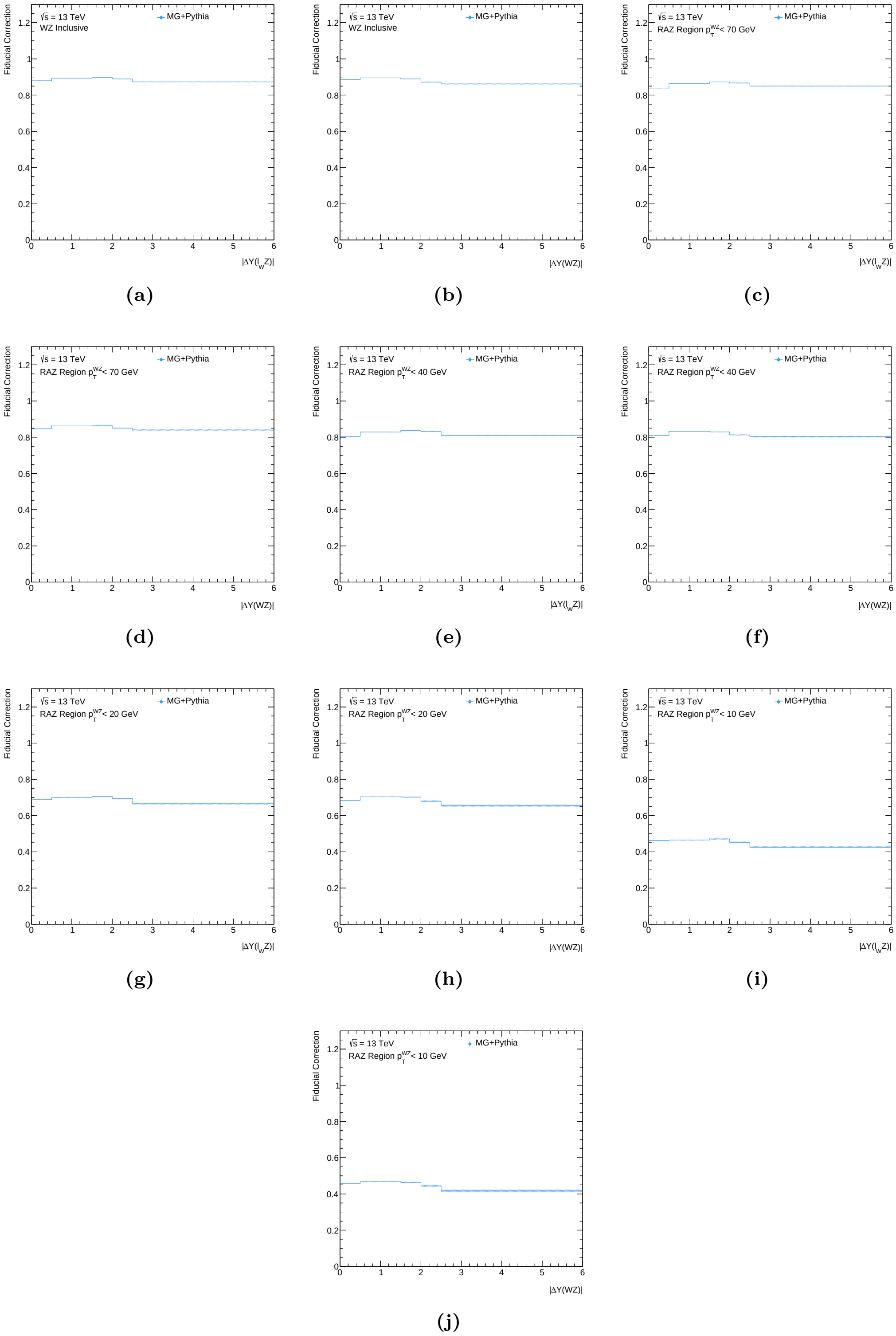



**Figure 11.7:** Fiducial correction inputs for the unfolding of the sum of polarization states using $\Delta Y(\ell_W Z)$ in (a) the inclusive region without a cut on $p_T^{WZ}$ and the regions with (c) $p_T^{WZ} < 70\,\text{GeV}$, (e) $p_T^{WZ} < 40\,\text{GeV}$, (g) $p_T^{WZ} < 20\,\text{GeV}$, (i) $p_T^{WZ} < 10\,\text{GeV}$ and using $\Delta Y(WZ)$ in (b) the inclusive region, and the regions with (d) $p_T^{WZ} < 70\,\text{GeV}$, (f) $p_T^{WZ} < 40\,\text{GeV}$, (h) $p_T^{WZ} < 20\,\text{GeV}$, (j) $p_T^{WZ} < 10\,\text{GeV}$.

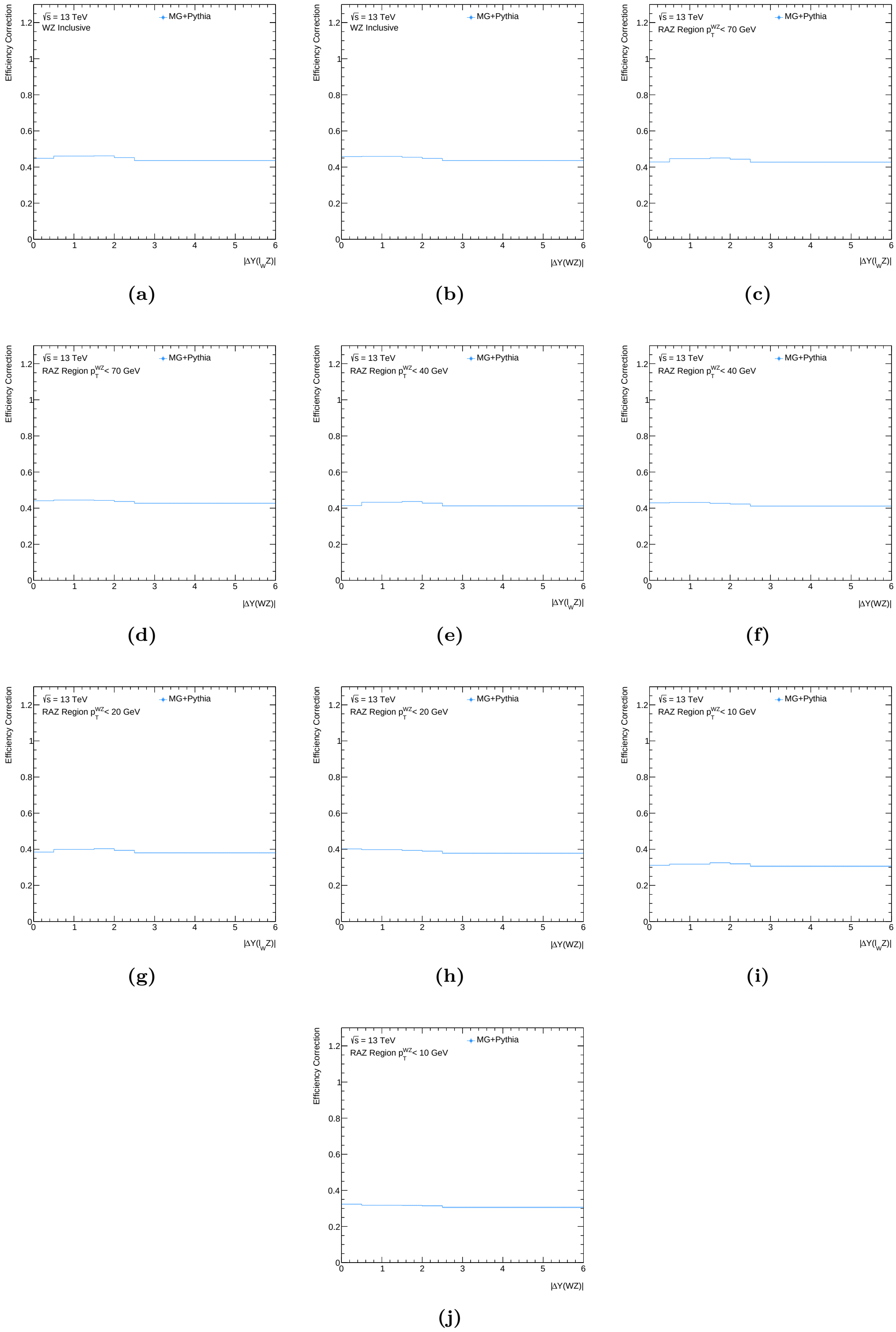



**Figure 11.8:** Efficiency correction inputs for the unfolding of the sum of polarization states using $\Delta Y(\ell_W Z)$ in (a) the inclusive region without a cut on $p_T^{WZ}$ and the regions with (c) $p_T^{WZ} < 70\,\text{GeV}$, (e) $p_T^{WZ} < 40\,\text{GeV}$, (g) $p_T^{WZ} < 20\,\text{GeV}$, (i) $p_T^{WZ} < 10\,\text{GeV}$ and using $\Delta Y(WZ)$ in (b) the inclusive region, and the regions with (d) $p_T^{WZ} < 70\,\text{GeV}$, (f) $p_T^{WZ} < 40\,\text{GeV}$, (h) $p_T^{WZ} < 20\,\text{GeV}$, (j) $p_T^{WZ} < 10\,\text{GeV}$.

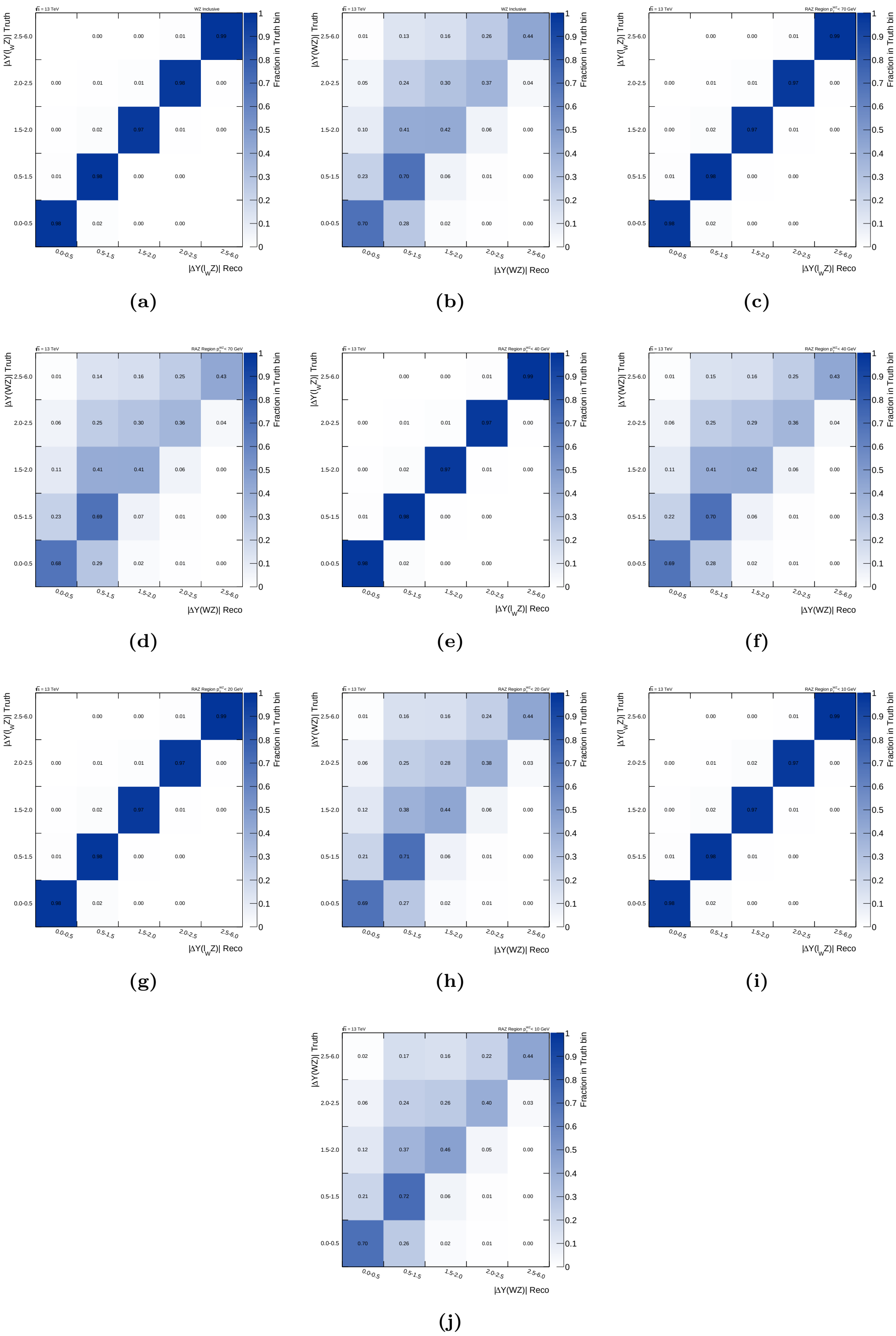



**Figure 11.9:** Migration matrix inputs for the unfolding of the sum of polarization states using $\Delta Y(\ell_W Z)$ in (a) the inclusive region without a cut on $p_T^{WZ}$ and the regions with (c) $p_T^{WZ} < 70\,\text{GeV}$, (e) $p_T^{WZ} < 40\,\text{GeV}$, (g) $p_T^{WZ} < 20\,\text{GeV}$, (i) $p_T^{WZ} < 10\,\text{GeV}$ and using $\Delta Y(WZ)$ in (b) the inclusive region, and the regions with (d) $p_T^{WZ} < 70\,\text{GeV}$, (f) $p_T^{WZ} < 40\,\text{GeV}$, (h) $p_T^{WZ} < 20\,\text{GeV}$, (j) $p_T^{WZ} < 10\,\text{GeV}$.

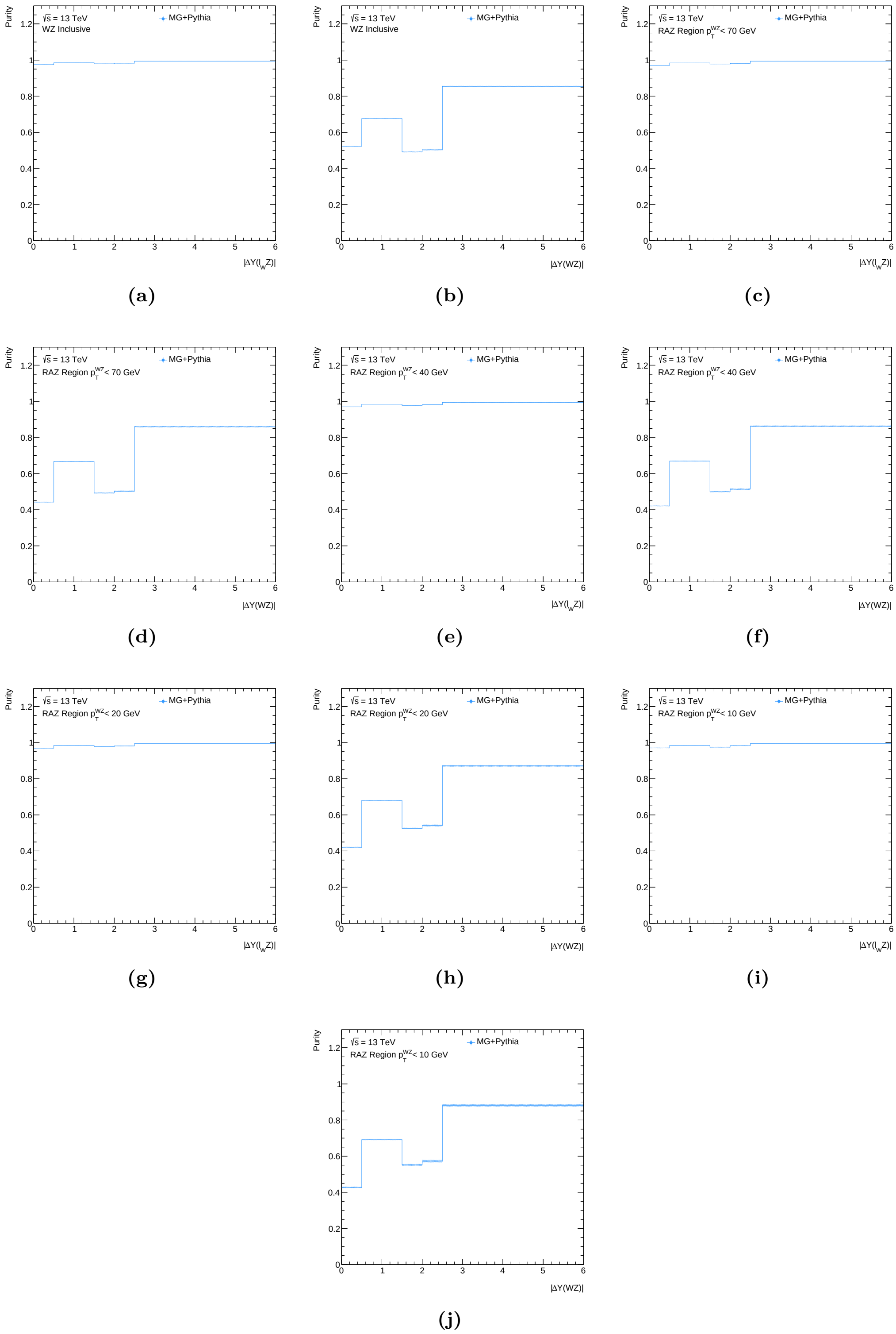



**Figure 11.10:** Purity for the unfolding of the sum of polarization states using $\Delta Y(\ell_W Z)$ in (a) the inclusive region without a cut on $p_T^{WZ}$ and the regions with (c) $p_T^{WZ} < 70\,\text{GeV}$, (e) $p_T^{WZ} < 40\,\text{GeV}$, (g) $p_T^{WZ} < 20\,\text{GeV}$, (i) $p_T^{WZ} < 10\,\text{GeV}$ and using $\Delta Y(WZ)$ in (b) the inclusive region, and the regions with (d) $p_T^{WZ} < 70\,\text{GeV}$, (f) $p_T^{WZ} < 40\,\text{GeV}$, (h) $p_T^{WZ} < 20\,\text{GeV}$, (j) $p_T^{WZ} < 10\,\text{GeV}$.

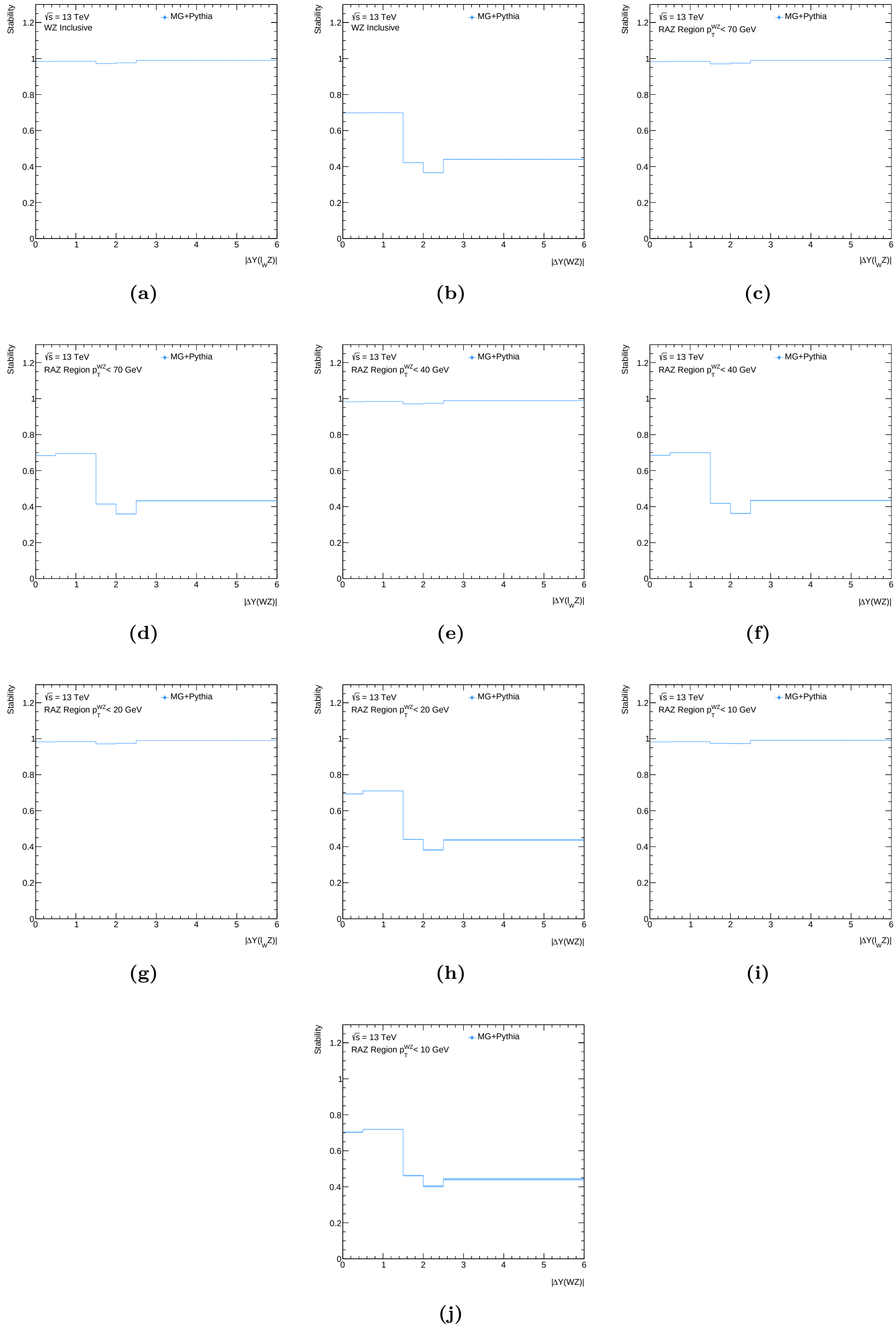



**Figure 11.11:** Stability for the unfolding of the sum of polarization states using $\Delta Y(\ell_W Z)$ in (a) the inclusive region without a cut on $p_T^{WZ}$ and the regions with (c) $p_T^{WZ} < 70\,\text{GeV}$, (e) $p_T^{WZ} < 40\,\text{GeV}$, (g) $p_T^{WZ} < 20\,\text{GeV}$, (i) $p_T^{WZ} < 10\,\text{GeV}$ and using $\Delta Y(WZ)$ in (b) the inclusive region, and the regions with (d) $p_T^{WZ} < 70\,\text{GeV}$, (f) $p_T^{WZ} < 40\,\text{GeV}$, (h) $p_T^{WZ} < 20\,\text{GeV}$, (j) $p_T^{WZ} < 10\,\text{GeV}$.

## 11.4 Unfolded Results

This section presents the results of the unfolding using $\Delta Y(\ell_W Z)$ and $\Delta Y(WZ)$ for both the TT polarization state as well as the sum of polarizations. Before any unfolding of data was done, a closure test was performed using simulated 00 events as for Asimov data, reconstructed and truth level signal, as well as the migration matrix. This was done to verify that the existing technical implementation works correctly. During that, a non-closure was found. The issue was investigated and discovered to be due to a multiple application of efficiency and fiducial corrections. The issue was fixed by the author[2], and full technical closure was seen afterward. The unfolding is performed separately in the (overlapping) RAZ regions as well as the inclusive region as defined in Chapter 5.

The following two Sections 11.4.1 and 11.4.2 show the results for the $\Delta Y(\ell_W Z)$ and $\Delta Y(WZ)$ unfolding, respectively.

### 11.4.1 Results for $\Delta Y(\ell_W Z)$

The full unfolded results using $\Delta Y(\ell_W Z)$ in all of the RAZ regions are summarized in Figure 11.12 for the TT polarization state and Figure 11.13 for the sum of polarizations. It can be seen that the relative cross-section in the bin around zero decreases from the inclusive region towards the $p_T^{WZ} < 10\,\text{GeV}$ region.

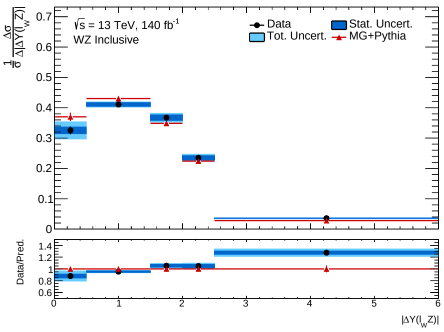
While the sum of polarization states has a visibly smaller dip at $\Delta Y(\ell_W Z) \approx 0$, its uncertainties are also overall lower. The main contributing factor is the reduced backgrounds that get subtracted from the data. One of the major uncertainty sources for the TT unfolding are normalization effects on the background. With most of the backgrounds having a peak structure around $\Delta Y(\ell_W Z) = 0$, any normalization change has a major impact on how low the dip is as zero.

The concrete unfolded bin entries and uncertainty breakdowns for the $p_T^{WZ} < 10\,\text{GeV}$ and inclusive regions are shown in Tables 11.1 and 11.2 for the unfolding of the TT polarization state and in Tables 11.3 and 11.4 for the sum of polarization states. It should be noted that the numbers can not be compared to the entries in the figures directly, as the values in the latter are divided by each bin's width. This is not the case for the tables. Identical tables for the other RAZ regions are shown in Appendix I.

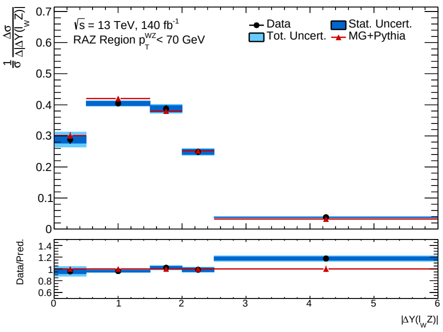
The depicted uncertainty breakdowns show that the dominant uncertainties in the unfolding of the TT polarization states are theory uncertainties. Mainly those related to NLO corrections as well as PDF, scale, and shower effects. The dominant experimental uncertainties are due to the Non-prompt background estimation in the inclusive region and the jet energy scale and resolution in the region with $p_T^{WZ} < 10\,\text{GeV}$. The data statistical uncertainty only plays a secondary role in the unfolding TT using $\Delta Y(\ell_W Z)$ except for the bin with $0.5 < \Delta Y(\ell_W Z) < 1.5$, where it is larger than the systematic uncertainties. The relative uncertainties are overall larger in the $p_T^{WZ} < 10\,\text{GeV}$ region. This holds for the statistical uncertainties due to fewer data and MC events in a region with more cuts, but also for the theory uncertainties as regions with jet vetoes generally pose more modeling issues and also depend more tightly on the jet energy measurements. In the bin with $|\Delta Y(\ell_W Z)| < 0.5$, the effect of larger relative uncertainties is exasperated due to the lower overall value.

The increase in the relative uncertainties for the region with the $p_T^{WZ}$ cut can also be observed in the unfolding of the sum of polarizations. One difference between the two setups is the impact of the data's statistical uncertainties. These are now the dominant ones in a plurality of the bins in both regions. This is due to a significant reduction in the NLO, PDF, scale, and showering uncertainties. The NLO uncertainties are derived directly at the truth level and exclusively for the WZ samples. As such, they have no impact on any of the remaining backgrounds and also not on fiducial and efficiency corrections or the migration matrix of the new signal. While the other theory uncertainties do have some impact on the response matrix, their dominant effect in the TT unfolding came from their impact on the normalization of the background polarizations. Nevertheless, they are still one of the largest

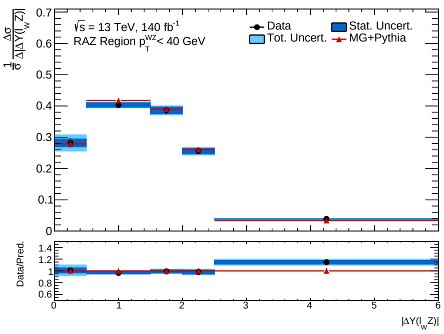
[2] The details are documented in `https://gitlab.cern.ch/atlas-physics/sm/StandardModelTools_Unfolding/VIPUnfolding/-/issues/6`



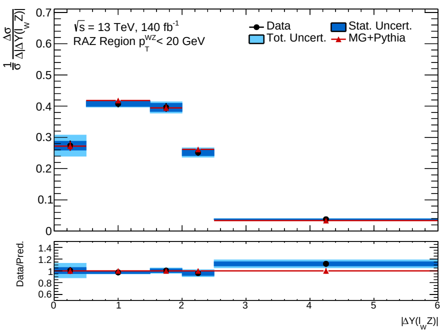

**(a)**



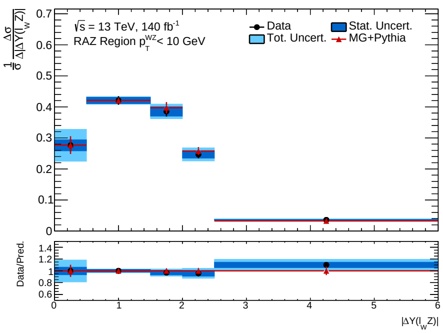

**(b)**




**(c)**

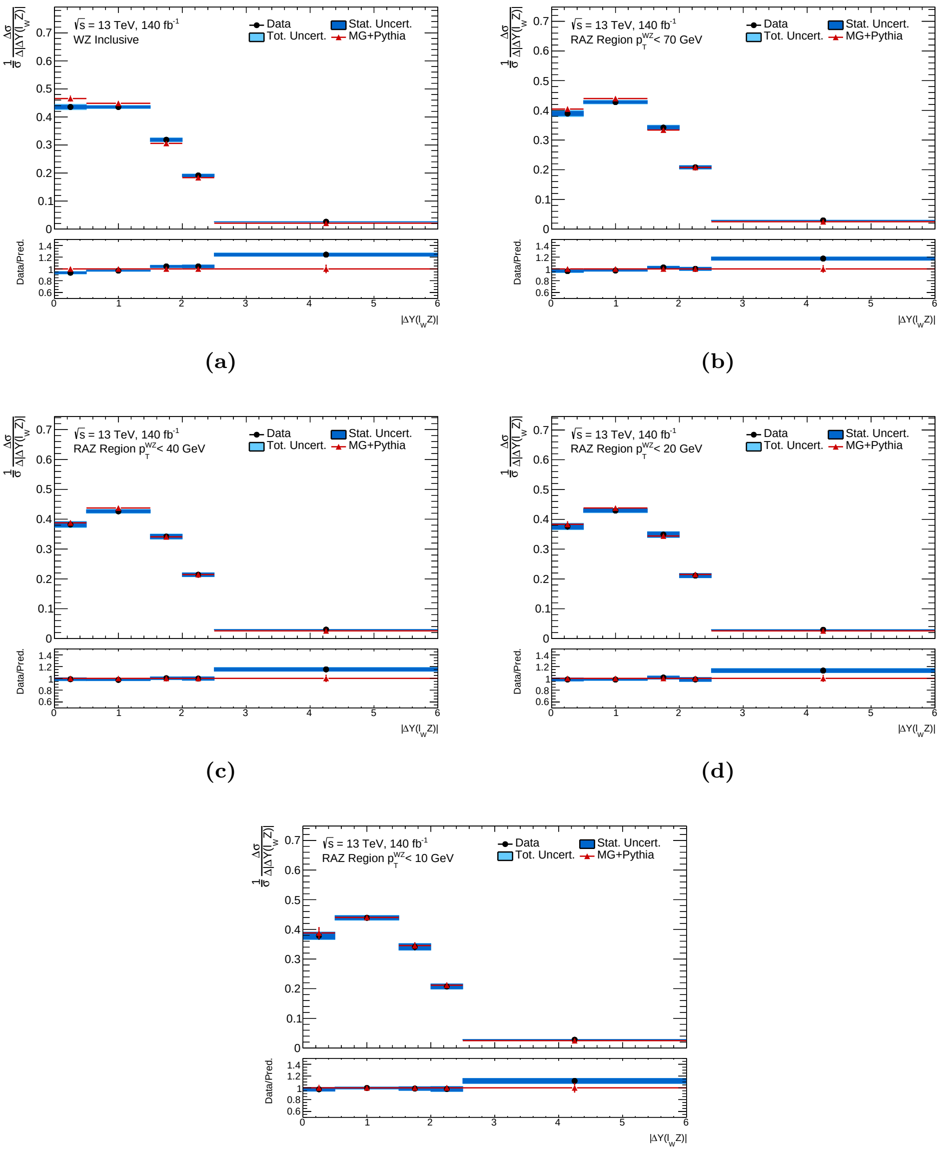



**(d)**




**(e)**

**Figure 11.12:** Unfolded distribution of the TT polarization state using $\Delta Y(\ell_W Z)$ in (a) the inclusive region without a cut on $p_T^{WZ}$ and the regions with (b) $p_T^{WZ} < 70\,\text{GeV}$, (c) $p_T^{WZ} < 40\,\text{GeV}$, (d) $p_T^{WZ} < 20\,\text{GeV}$, (e) $p_T^{WZ} < 10\,\text{GeV}$.




(a) (b) (c) (d) (e)

**Figure 11.13:** Unfolded distribution of the sum of polarization states using $\Delta Y(\ell_W Z)$ in (a) the inclusive region without a cut on $p_T^{WZ}$ and the regions with (b) $p_T^{WZ} < 70\,\text{GeV}$, (c) $p_T^{WZ} < 40\,\text{GeV}$, (d) $p_T^{WZ} < 20\,\text{GeV}$, (e) $p_T^{WZ} < 10\,\text{GeV}$.

uncertainties alongside Non-prompt background estimation and statistical data uncertainty. Another difference between the unfolding of the TT state and the sum of polarizations is that the data statistical uncertainties are overall smaller for the latter, due to the data fluctuations having a smaller relative size after the smaller background subtraction.

**Table 11.1:** Results and uncertainties for the unfolding of the TT polarization state using $\Delta Y(\ell_W Z)$ in the region with $p_T^{WZ} < 10\,\text{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.138 | 0.420 | 0.193 | 0.122 | 0.125 |
| Comb. Unc. | 18.6% | 2.95% | 6.14% | 8.58% | 9.58% |
| Data statistical | 6.54% | 2.45% | 3.98% | 4.89% | 4.66% |
| Experimental and modelling | 17.4% | 1.64% | 4.67% | 7.05% | 8.37% |
| Luminosity | 0.65% | 0.04% | 0.19% | 0.25% | 0.33% |
| Electron calibration | 0.85% | 0.10% | 0.20% | 0.38% | 0.61% |
| Muon calibration | 0.99% | 0.11% | 0.33% | 0.38% | 0.41% |
| Jet energy scale and resolution | 3.98% | 0.14% | 1.04% | 1.21% | 1.85% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 2.62% | 0.06% | 0.73% | 0.71% | 1.37% |
| Flavor-tagging inefficiency | 0.01% | 0.00% | 0.00% | 0.00% | 0.00% |
| Pileup modelling | 1.84% | 0.49% | 0.60% | 1.61% | 1.16% |
| Non-prompt background estimation | 1.32% | 0.13% | 0.46% | 0.17% | 0.13% |
| Background, other | 1.93% | 0.05% | 0.55% | 0.69% | 0.78% |
| Model statistical | 1.16% | 0.47% | 0.76% | 0.99% | 1.00% |
| NLO corrections | 11.4% | 0.77% | 2.86% | 4.30% | 5.07% |
| PDF, Scale and shower settings | 11.7% | 1.25% | 3.21% | 4.97% | 5.94% |

**Table 11.2:** Results and uncertainties for the unfolding of the TT polarization state using $\Delta Y(\ell_W Z)$ in the inclusive region.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.163 | 0.411 | 0.183 | 0.118 | 0.125 |
| Comb. Unc. | 8.81% | 2.18% | 4.00% | 5.15% | 6.49% |
| Data statistical | 3.49% | 1.60% | 2.59% | 3.16% | 2.96% |
| Experimental and modelling | 8.08% | 1.48% | 3.04% | 4.06% | 5.77% |
| Luminosity | 0.66% | 0.09% | 0.25% | 0.33% | 0.46% |
| Electron calibration | 0.53% | 0.09% | 0.18% | 0.25% | 0.51% |
| Muon calibration | 0.59% | 0.06% | 0.26% | 0.26% | 0.30% |
| Jet energy scale and resolution | 0.73% | 0.08% | 0.24% | 0.47% | 0.47% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 0.30% | 0.02% | 0.08% | 0.05% | 0.14% |
| Flavor-tagging inefficiency | 0.01% | 0.00% | 0.01% | 0.01% | 0.01% |
| Pileup modelling | 0.08% | 0.23% | 0.07% | 0.25% | 0.29% |
| Non-prompt background estimation | 3.64% | 0.20% | 1.40% | 1.33% | 2.09% |
| Background, other | 2.07% | 0.16% | 0.79% | 0.95% | 1.18% |
| Model statistical | 0.59% | 0.26% | 0.44% | 0.55% | 0.54% |
| NLO corrections | 4.34% | 0.53% | 1.17% | 2.26% | 3.23% |
| PDF, Scale and shower settings | 5.18% | 1.30% | 2.19% | 2.80% | 3.98% |

**Table 11.3:** Results and uncertainties for the unfolding of the sum of polarization states using $\Delta Y(\ell_W Z)$ in the region with $p_T^{WZ} < 10\,\text{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.189 | 0.439 | 0.171 | 0.104 | 0.097 |
| Comb. Unc. | 3.48% | 1.98% | 3.42% | 4.42% | 4.50% |
| Data statistical | 3.15% | 1.64% | 3.17% | 4.13% | 4.23% |
| Experimental and modelling | 1.49% | 1.11% | 1.29% | 1.58% | 1.54% |
| Luminosity | 0.03% | 0.00% | 0.01% | 0.02% | 0.02% |
| Electron calibration | 0.04% | 0.02% | 0.05% | 0.02% | 0.11% |
| Muon calibration | 0.10% | 0.09% | 0.18% | 0.08% | 0.30% |
| Jet energy scale and resolution | 0.23% | 0.15% | 0.10% | 0.25% | 0.19% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 0.20% | 0.07% | 0.21% | 0.20% | 0.25% |
| Flavor-tagging inefficiency | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| Pileup modelling | 0.14% | 0.22% | 0.01% | 0.68% | 0.02% |
| Non-prompt background estimation | 0.50% | 0.09% | 0.29% | 0.04% | 0.04% |
| Background, other | 0.69% | 0.02% | 0.31% | 0.40% | 0.45% |
| Model statistical | 0.55% | 0.31% | 0.60% | 0.84% | 0.93% |
| PDF, Scale and shower settings | 1.03% | 1.01% | 1.01% | 1.02% | 1.04% |

**Table 11.4:** Results and uncertainties for the unfolding of the sum of polarization states using $\Delta Y(\ell_W Z)$ in the inclusive region.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.217 | 0.434 | 0.159 | 0.096 | 0.091 |
| Comb. Unc. | 2.34% | 1.41% | 2.31% | 2.82% | 3.08% |
| Data statistical | 1.60% | 0.95% | 1.89% | 2.47% | 2.57% |
| Experimental and modelling | 1.71% | 1.04% | 1.33% | 1.37% | 1.70% |
| Luminosity | 0.04% | 0.00% | 0.02% | 0.03% | 0.04% |
| Electron calibration | 0.09% | 0.01% | 0.08% | 0.11% | 0.06% |
| Muon calibration | 0.04% | 0.05% | 0.06% | 0.07% | 0.21% |
| Jet energy scale and resolution | 0.18% | 0.05% | 0.13% | 0.13% | 0.14% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 0.09% | 0.02% | 0.03% | 0.02% | 0.06% |
| Flavor-tagging inefficiency | 0.00% | 0.00% | 0.00% | 0.01% | 0.00% |
| Pileup modelling | 0.30% | 0.09% | 0.10% | 0.01% | 0.09% |
| Non-prompt background estimation | 1.05% | 0.00% | 0.63% | 0.49% | 0.92% |
| Background, other | 0.61% | 0.04% | 0.36% | 0.44% | 0.55% |
| Model statistical | 0.26% | 0.15% | 0.30% | 0.41% | 0.44% |
| PDF, Scale and shower settings | 1.09% | 1.02% | 1.04% | 1.11% | 1.20% |

### 11.4.2 Results for $\Delta Y(WZ)$

The unfolded $\Delta Y(WZ)$ distributions are shown in Figures 11.14 and 11.15 analogously to the TT case. Similarly, the exact bin values and the systematic breakdowns for each bin are given in Tables 11.5 and 11.6 for the TT unfolding in the $p_T^{WZ} < 10\,\mathrm{GeV}$ and inclusive regions. The respective results for the unfolding of the sum of polarizations are depicted in Figures 11.7 and 11.8.

It can be seen that both configurations show an overall larger dip around zero than was the case with the unfolding of the $\Delta Y(\ell_W Z)$ distribution. This is expected, given that the RAZ predicts an exact zero only in the $\Delta Y(WZ)$ distribution, which gets smeared out by the decay of the $W$ boson in the $\Delta Y(\ell_W Z)$ distribution. The main motivation to also consider that distribution is the significantly easier reconstruction as shown in Section 11.3. An exact zero is, however, still not seen, even for the $\Delta Y(WZ)$ unfolding of the TT polarization state. This is also not unexpected due to the RAZ being washed out due to real higher order corrections, even with a tight cut on $p_T^{WZ}$. Additionally, the zero at $\Delta Y(WZ) = 0$ is also just approximate due to a neglection of the weak hypercharge. The exact zero is expected in $cos\theta_V$, which is, however, even more challenging to reconstruct and measure as well as smeared out due to it being impossible to determine which of the initial quarks originated from which proton.

The tables show that the unfolding of $\Delta Y(WZ)$ exhibits generally larger uncertainties in both configurations and regions. The uncertainties that mainly contribute to this increase are the data and modeling statistical ones. This is almost purely due to the significant migrations observed between the truth and reconstructed level for $\Delta Y(WZ)$. With that, small statistical fluctuations in the data or simulation can cause large, anticorrelated deviations in the final unfolded bins. Most of the other uncertainties have a similar impact per bin between the unfolding of $\Delta Y(WZ)$ and $\Delta Y(\ell_W Z)$. The significant increases in the bin with $|\Delta Y(WZ)| < 0.5$ are mostly due to the significantly smaller nominal value. The only other uncertainty category that shows a significant increase is the one for muon calibration. It shows an increase from around 0.3 % in the unfolding of $\Delta Y(\ell_W Z)$ up to $> 2\,\%$ in the unfolding of $\Delta Y(WZ)$ for the bins from 1.5 to 2.5. This change can be seen across the two shown regions and both unfolding configurations and is due to an up-fluctuation of the reconstructed signal in the $2.0 - 2.5$ bin. As this is compatible within the MC statistical uncertainties and small with respect to the dominant uncertainty, no further investigation was done.

**Table 11.5:** Results and uncertainties for the unfolding of the TT polarization state using $\Delta Y(WZ)$ in the region with $p_T^{WZ} < 10\,\mathrm{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.071 | 0.427 | 0.208 | 0.155 | 0.137 |
| Comb. Unc. | 45.2% | 6.09% | 11.5% | 12.7% | 10.7% |
| Data statistical | 24.7% | 5.76% | 9.96% | 10.3% | 7.25% |
| Experimental and modelling | 37.9% | 1.96% | 5.92% | 7.41% | 7.86% |
| Luminosity | 1.76% | 0.02% | 0.21% | 0.28% | 0.35% |
| Electron calibration | 2.71% | 0.14% | 0.51% | 0.56% | 0.82% |
| Muon calibration | 2.57% | 0.30% | 2.22% | 2.34% | 0.76% |
| Jet energy scale and resolution | 10.2% | 0.36% | 0.89% | 2.95% | 1.56% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 8.29% | 0.24% | 1.48% | 0.60% | 1.92% |
| Flavor-tagging inefficiency | 0.02% | 0.00% | 0.00% | 0.00% | 0.00% |
| Pileup modelling | 6.16% | 0.17% | 0.47% | 1.25% | 1.59% |
| Non-prompt background estimation | 3.81% | 0.34% | 0.45% | 0.20% | 0.06% |
| Background, other | 4.78% | 0.01% | 0.58% | 0.72% | 0.73% |
| Model statistical | 4.82% | 1.14% | 2.04% | 2.32% | 1.64% |
| NLO corrections | 14.9% | 0.81% | 2.81% | 2.77% | 2.78% |
| PDF, Scale and shower settings | 30.3% | 1.20% | 3.73% | 4.96% | 6.39% |

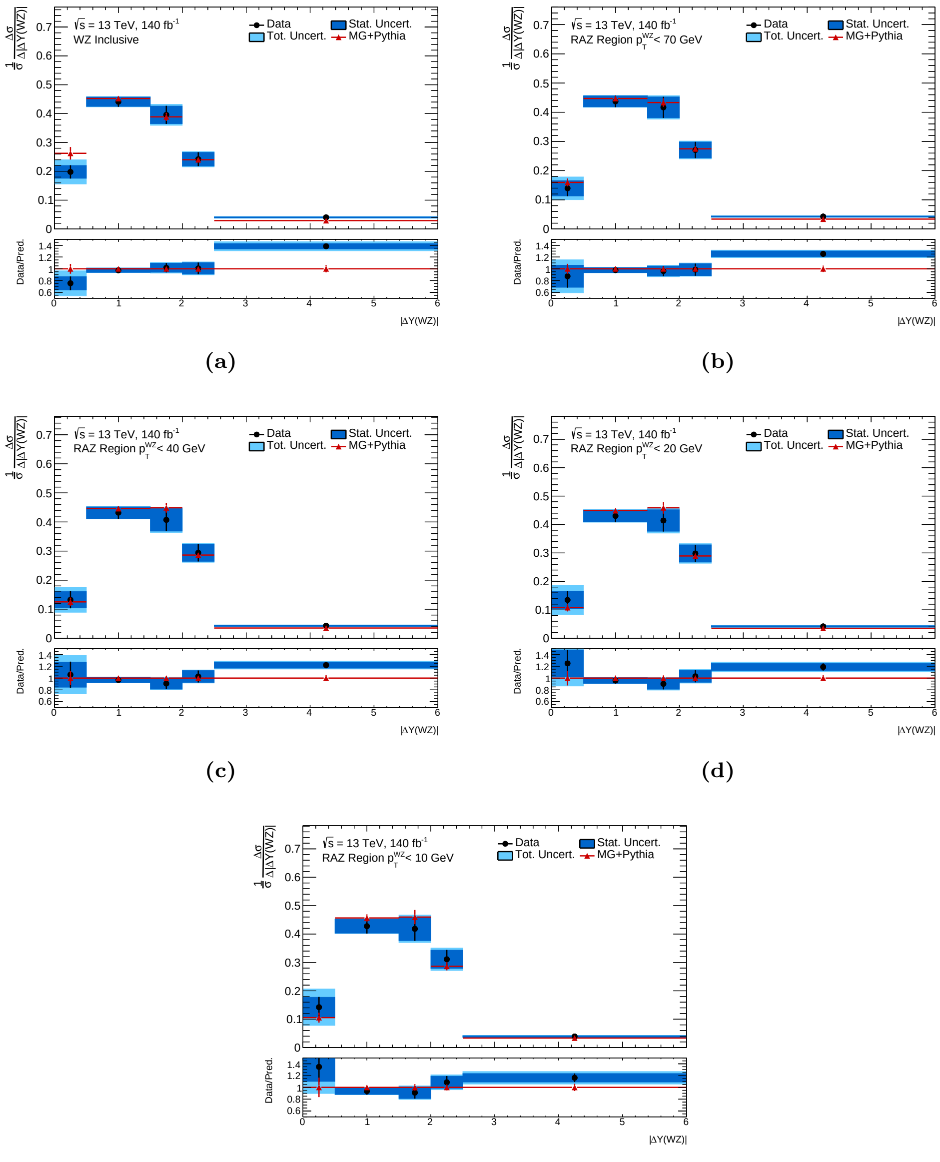
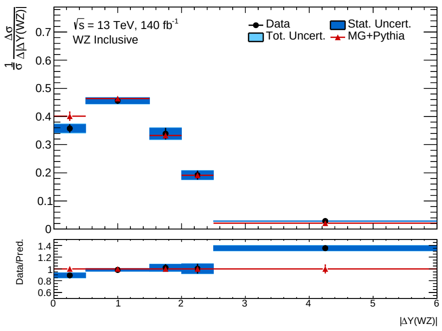
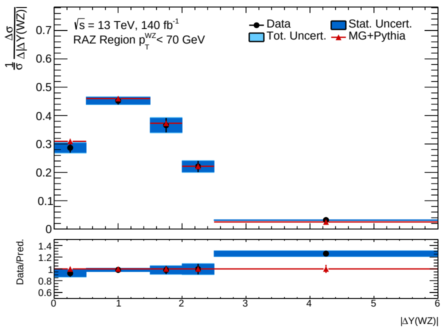


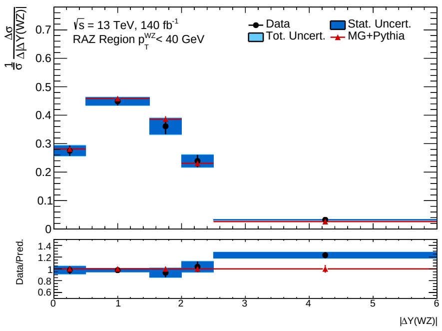

**Figure 11.14:** Unfolded distribution of the TT polarization state using $\Delta Y(WZ)$ in (a) the inclusive region without a cut on $p_T^{WZ}$ and the regions with (b) $p_T^{WZ} < 70\,\text{GeV}$, (c) $p_T^{WZ} < 40\,\text{GeV}$, (d) $p_T^{WZ} < 20\,\text{GeV}$, (e) $p_T^{WZ} < 10\,\text{GeV}$.



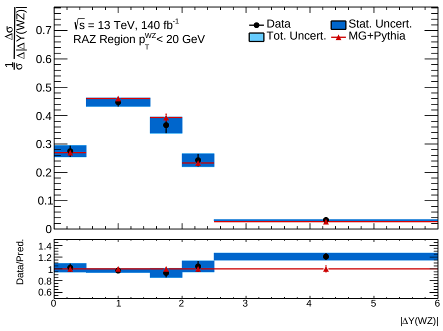

**(a)**



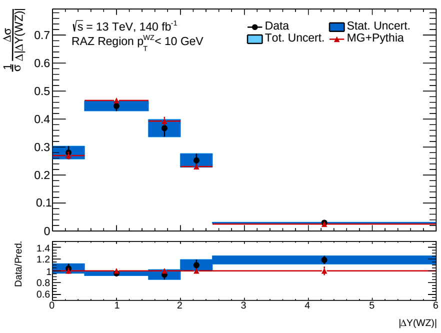

**(b)**




**(c)**




**(d)**




**(e)**

**Figure 11.15:** Unfolded distribution of the sum of polarization states using $\Delta Y(WZ)$ in (a) the inclusive region without a cut on $p_T^{WZ}$ and the regions with (b) $p_T^{WZ} < 70\,\text{GeV}$, (c) $p_T^{WZ} < 40\,\text{GeV}$, (d) $p_T^{WZ} < 20\,\text{GeV}$, (e) $p_T^{WZ} < 10\,\text{GeV}$.

**Table 11.6:** Results and uncertainties for the unfolding of the TT polarization state using $\Delta Y(WZ)$ in the inclusive region.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.099 | 0.441 | 0.198 | 0.120 | 0.141 |
| Comb. Unc. | 21.2% | 4.08% | 9.22% | 11.0% | 7.75% |
| Data statistical | 11.2% | 3.71% | 7.71% | 9.83% | 4.66% |
| Experimental and modelling | 17.9% | 1.70% | 5.05% | 5.00% | 6.19% |
| Luminosity | 1.54% | 0.04% | 0.30% | 0.33% | 0.50% |
| Electron calibration | 1.65% | 0.10% | 0.47% | 0.45% | 0.55% |
| Muon calibration | 1.80% | 0.37% | 2.14% | 2.04% | 0.66% |
| Jet energy scale and resolution | 2.17% | 0.22% | 0.61% | 0.88% | 0.61% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 1.55% | 0.33% | 0.67% | 0.94% | 0.31% |
| Flavor-tagging inefficiency | 0.02% | 0.01% | 0.02% | 0.03% | 0.00% |
| Pileup modelling | 0.25% | 0.00% | 0.26% | 0.34% | 0.47% |
| Non-prompt background estimation | 7.78% | 0.06% | 1.58% | 0.96% | 2.63% |
| Background, other | 4.45% | 0.07% | 0.88% | 0.90% | 1.20% |
| Model statistical | 2.13% | 0.73% | 1.62% | 2.07% | 0.99% |
| NLO corrections | 6.83% | 0.85% | 2.04% | 1.99% | 2.81% |
| PDF, Scale and shower settings | 13.2% | 1.14% | 3.10% | 2.94% | 4.40% |

**Table 11.7:** Results and uncertainties for the unfolding of the sum of polarization states using $\Delta Y(WZ)$ in the region with $p_T^{WZ} < 10\,\mathrm{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.140 | 0.447 | 0.183 | 0.126 | 0.103 |
| Comb. Unc. | 8.45% | 4.04% | 8.73% | 9.97% | 7.25% |
| Data statistical | 8.03% | 3.80% | 8.25% | 9.33% | 6.88% |
| Experimental and modelling | 2.63% | 1.37% | 2.85% | 3.51% | 2.28% |
| Luminosity | 0.06% | 0.00% | 0.01% | 0.02% | 0.02% |
| Electron calibration | 0.08% | 0.06% | 0.33% | 0.25% | 0.33% |
| Muon calibration | 0.23% | 0.19% | 1.77% | 2.20% | 0.53% |
| Jet energy scale and resolution | 0.42% | 0.19% | 0.61% | 1.32% | 0.68% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 0.81% | 0.14% | 0.63% | 0.35% | 0.56% |
| Flavor-tagging inefficiency | 0.01% | 0.00% | 0.00% | 0.00% | 0.00% |
| Pileup modelling | 0.03% | 0.05% | 0.20% | 0.24% | 0.33% |
| Non-prompt background estimation | 0.97% | 0.21% | 0.30% | 0.12% | 0.14% |
| Background, other | 1.08% | 0.01% | 0.33% | 0.42% | 0.40% |
| Model statistical | 1.50% | 0.71% | 1.67% | 2.00% | 1.50% |
| PDF, Scale and shower settings | 1.28% | 1.10% | 1.02% | 1.12% | 1.20% |

**Table 11.8:** Results and uncertainties for the unfolding of the sum of polarization states using $\Delta Y(WZ)$ in the inclusive region.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.178 | 0.456 | 0.169 | 0.096 | 0.100 |
| Comb. Unc. | 4.74% | 2.65% | 6.45% | 8.76% | 4.88% |
| Data statistical | 4.03% | 2.35% | 5.94% | 8.20% | 4.25% |
| Experimental and modelling | 2.51% | 1.22% | 2.50% | 3.07% | 2.39% |
| Luminosity | 0.07% | 0.00% | 0.03% | 0.03% | 0.04% |
| Electron calibration | 0.03% | 0.03% | 0.16% | 0.21% | 0.14% |
| Muon calibration | 0.20% | 0.24% | 1.48% | 1.74% | 0.35% |
| Jet energy scale and resolution | 0.59% | 0.16% | 0.46% | 0.47% | 0.35% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 0.48% | 0.22% | 0.55% | 0.76% | 0.21% |
| Flavor-tagging inefficiency | 0.01% | 0.00% | 0.01% | 0.02% | 0.01% |
| Pileup modelling | 0.56% | 0.01% | 0.34% | 0.53% | 0.00% |
| Non-prompt background estimation | 1.58% | 0.02% | 0.73% | 0.24% | 1.38% |
| Background, other | 0.93% | 0.05% | 0.40% | 0.41% | 0.57% |
| Model statistical | 0.73% | 0.45% | 1.22% | 1.73% | 0.91% |
| PDF, Scale and shower settings | 1.18% | 1.07% | 1.09% | 1.42% | 1.53% |

## 11.5 Unfolded Depth Calculation

The main goal of unfolding the rapidity differences is the investigation of the radiation amplitude zero effect. It can be characterized by an (approximate) zero amplitude of the TT polarization state close to $\Delta Y(WZ) = 0$ for pure leading order events. This is visible as a dip in the unfolded distributions.

To offer some quantification of this visual fact, a depth variable is defined as

$$\begin{aligned} D(\mathrm{Dip}, \mathrm{Peak}) &= 1 - \frac{2 \cdot \mathrm{Dip}}{\mathrm{Peak}}, \\ \mathrm{Dip} &= N(0.0 < |\Delta Y| < 0.5), \\ \mathrm{Peak} &= N(0.5 < |\Delta Y| < 1.5), \end{aligned} \tag{11.13}$$

where $N(|\Delta Y| < 0.5)$ and $N(0.5 < |\Delta Y| < 1.5)$ indicate the number of unfolded events with $|\Delta Y(\ell_W Z)\ |$ (or $|\Delta Y(WZ)\ |$) less than 0.5 or between 0.5 and 1.5. The factor of 2 in the denominator is to make sure the same rapidity difference range is used for both the denominator and numerator. Values above zero indicate a dip structure, while values below represent peaks.

Uncertainties on this quantity are derived using the bin-by-bin covariance matrix $V$ provided by VIPUnfolding via

$$\Delta D(\mathrm{Dip}, \mathrm{Peak}) = \sqrt{(\nabla D)^T \cdot V \cdot \nabla D}. \tag{11.14}$$

The resulting values are summarized in Figure 11.16 and Tables 11.9 and 11.10. Generally, good agreement between the theory predictions and observed data can be seen. The main exception is the inclusive region, where a sizably larger dip was measured than expected. The trend of a stronger dip with tighter $p_T^{WZ}$ cuts, and thus less jet activity, is visible for the Madgraph+Pythia8 predictions. While it can also be observed in the unfolded data, a clear tapering-off can be seen for the tightest cuts. For the depth calculated from $\Delta Y(WZ)$, there is even a small decline for the region with $p_T^{WZ} < 10\,\mathrm{GeV}$. However, they are all well within the uncertainties. Directly matching the expectations from the raw unfolded distributions, both a stronger dip, as well as larger uncertainties can be observed for the unfolding of $\Delta Y(WZ)$. This is the case for both the TT configuration and the unfolding of the sum of polarizations. A similar effect can be seen for the unfolding of TT when compared to the unfolding of all polarization states. The former has a larger overall dip but is also affected by the larger uncertainties seen in the previous section.

The combination of these effects results in the largest estimated significance, $D/\Delta D$, to be observed for the unfolding of $\Delta Y(WZ)$. Here, both the unfolding of just the TT state as well as the one for the sum of polarizations results in four regions with a value of $D/\Delta D$ above 5, while either the region with $p_T^{WZ}$ or the inclusive region end up at around 4. For the former, the main reason is the extremely large uncertainty of 0.157, while the latter is mostly hindered by the small value of the depth at around $D = 0.051$.

The overall smaller uncertainties in the unfolding of $\Delta Y(\ell_W Z)$ do not make up for the significantly less pronounced dips. Due to this, the unfolding of the TT polarization using $\Delta Y(\ell_W Z)$ does not reach $D/\Delta D = 5$ in any region and even falls below 3 in the most restrictive one. The unfolding of the sum of polarization states is mostly still competitive with the TT one, even for $\Delta Y(\ell_W Z)$, reaching a higher estimated significance in the $p_T^{WZ} < 10\,\mathrm{GeV}$ region. However, as the distribution shown in Figure 11.13a indicates, there is a flat structure around $\Delta Y(\ell_W Z) = 0$ in the inclusive region for the sum of polarizations instead of a dip. The Madgraph+Pythia8 prediction is even a peak structure. As such, the estimated significance for that particular configuration is very close to zero and the relative uncertainties are omitted.

As VIPUnfolding does not only store a total covariance matrix but also one per uncertainty, it is possible to apply Equation 11.14 to calculate a breakdown of the statistical and systematic uncertainty sources. These breakdowns for each RAZ region are shown with increasing tightness of the $p_T^{WZ}$ cut in Tables 11.11 to 11.15. The overall uncertainty progressions

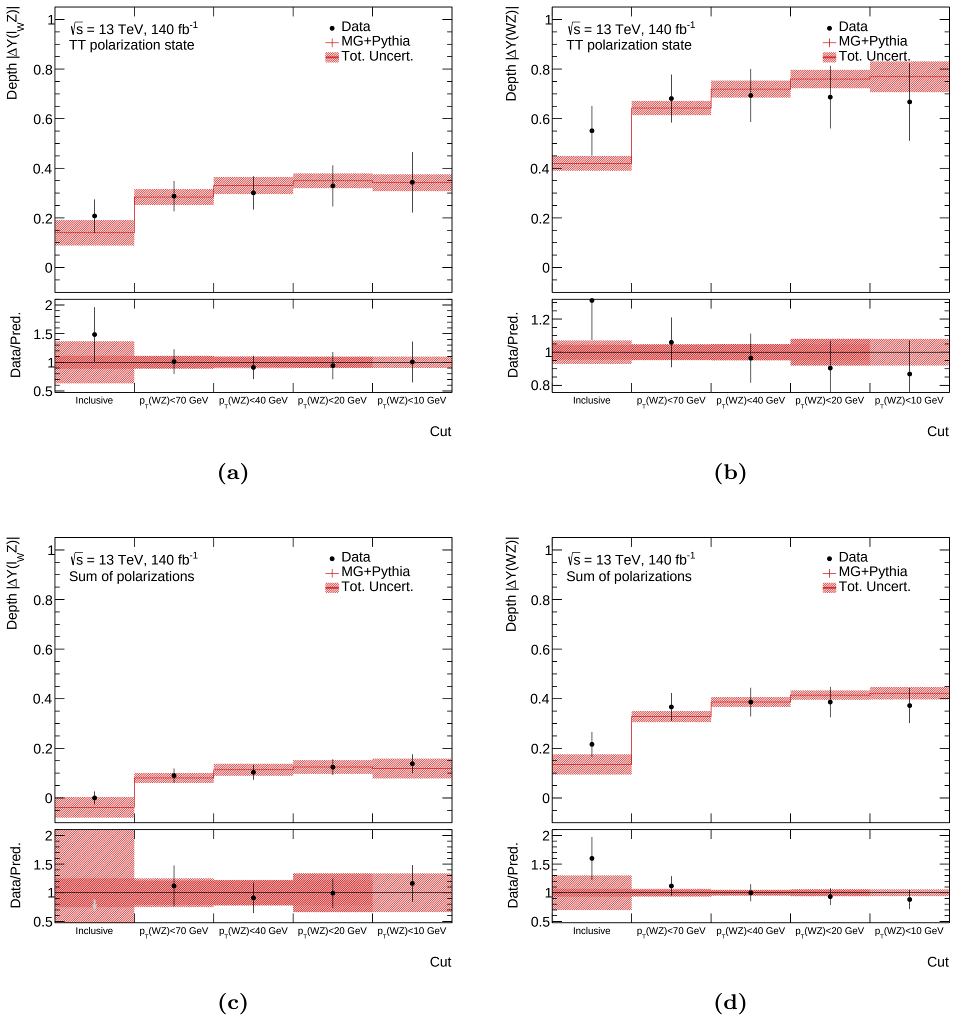



**Figure 11.16:** Progression of the depth values and their uncertainties with the tightness of the $p_T^{WZ}$ cut at the unfolded level for the TT polarization state using (a) $\Delta Y(\ell_W Z)$ and (b) $\Delta Y(WZ)$ and for the sum of polarization states using (c) $\Delta Y(\ell_W Z)$ and (d) $\Delta Y(WZ)$.

**Table 11.9:** Depth for $\Delta Y(WZ)$.

| | TT state | | | | Sum of polarizations | | | |
|---|---|---|---|---|---|---|---|---|
| Cut | Value | Uncertainty | | $D/\Delta D$ | Value | Uncertainty | | $D/\Delta D$ |
| Inclusive | 0.551 | 0.100 | 18.2 % | 5.50 | 0.216 | 0.051 | 23.4 % | 4.28 |
| $p_T(WZ) < 70$ GeV | 0.681 | 0.097 | 14.2 % | 7.03 | 0.367 | 0.056 | 15.3 % | 6.53 |
| $p_T(WZ) < 40$ GeV | 0.693 | 0.107 | 15.4 % | 6.48 | 0.386 | 0.058 | 15.0 % | 6.66 |
| $p_T(WZ) < 20$ GeV | 0.687 | 0.126 | 18.4 % | 5.44 | 0.387 | 0.061 | 15.8 % | 6.31 |
| $p_T(WZ) < 10$ GeV | 0.667 | 0.157 | 23.4 % | 4.26 | 0.372 | 0.071 | 19.0 % | 5.26 |

**Table 11.10:** Depth for $\Delta Y(\ell_W Z)$.

| | TT state | | | Sum of polarizations | | |
|---|---|---|---|---|---|---|
| Cut | Value | Uncertainty | $D/\Delta D$ | Value | Uncertainty | $D/\Delta D$ |
| Inclusive | 0.208 | 0.067 \| 32.2 % | 3.10 | 0.000 | 0.026 \| - | 0.00 |
| $p_T(WZ) < 70\,\text{GeV}$ | 0.287 | 0.061 \| 21.3 % | 4.69 | 0.090 | 0.029 \| 32.0 % | 3.13 |
| $p_T(WZ) < 40\,\text{GeV}$ | 0.301 | 0.067 \| 22.4 % | 4.46 | 0.103 | 0.030 \| 29.2 % | 3.42 |
| $p_T(WZ) < 20\,\text{GeV}$ | 0.329 | 0.083 \| 25.3 % | 3.96 | 0.124 | 0.032 \| 25.9 % | 3.86 |
| $p_T(WZ) < 10\,\text{GeV}$ | 0.344 | 0.122 \| 35.5 % | 2.81 | 0.138 | 0.038 \| 27.6 % | 3.63 |

between regions and the differences between the variables and configurations match those already discussed here. Additionally, the different effects of larger relative uncertainties per bin for small bins leading to smaller relative uncertainties for a larger dip are also mirrored there.

The most interesting observation is the increased effect of the statistical uncertainties compared to the ones seen for the raw values of the center two bins. This is best demonstrated for the case of the $\Delta Y(\ell_W Z)$ unfolding of the sum of polarizations in the region with $p_T^{WZ} < 10\,\text{GeV}$. Here, Table 11.3 indicates data statistical uncertainty of 3.15 % and 1.64 % in comparison to systematic uncertainties of 1.49 % and 1.11 % in the bins used for the depth calculation. Meanwhile, Table 11.15 shows an impact of 26.3 % on $D$ for the data statistical uncertainty and 8.3 % for systematic uncertainty. A similar increase can be observed for the MC statistical uncertainties. This increase in their impact is due to the bin-by-bin correlations. All of the other sizeable uncertainties are almost fully correlated between the two relevant bins, mainly because their largest impact comes from normalization effects on background subtraction. This always increases or decreases both of the bins at the same time. Meanwhile, the statistical uncertainties are evaluated independently for each bin. Due to the large migrations, this leads to sizeable anticorrelations between the individual bins, which increases the uncertainty of an observable that depends on their ratio.

Overall, the statistical uncertainties are the dominant ones for the depth when unfolding the sum of polarizations, regardless of the variable used or $p_T^{WZ}$ cut considered. For the unfolding of just the TT polarization state, the statistical uncertainties are smaller but still sizeable across all regions when unfolding $\Delta Y(\ell_W Z)$. The configuration where statistical and systematic uncertainties have the most similar impact is the unfolding of TT using $\Delta Y(WZ)$. Here, either is dominant depending on the exact region considered. Thus, a repetition of this measurement using the data from the LHC's Run3 or even the HL-LHC [268] promises great improvements in the sensitivity. Still, significant improvements in the estimation of the non-prompt background, NLO corrections, and shower modeling would be of great benefit, especially for the unfolding of the TT polarization state.

**Table 11.11:** Uncertainty on the depth for the TT polarization state and the sum of polarization states using $\Delta Y(\ell_W Z)$ and $\Delta Y(WZ)$ in the inclusive region. The uncertainties for $\Delta Y(\ell_W Z)$ from the sum of polarizations are omitted due to the value being close to zero.

| | Impact [%] | | | |
|---|---|---|---|---|
| Source | TT state | | Sum of polarizations | |
| Experimental | $\Delta Y(\ell_W Z)$ | $\Delta Y(WZ)$ | $\Delta Y(\ell_W Z)$ | $\Delta Y(WZ)$ |
| Luminosity | 2.2 | 1.2 | - | 0.3 |
| Electron calibration | 1.7 | 1.3 | - | 0.2 |
| Muon calibration | 2.2 | 1.4 | - | 1.2 |
| Jet energy scale and resolution | 2.6 | 1.8 | - | 2.6 |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 1.3 | 1.5 | - | 2.5 |
| Flavor-tagging inefficiency | 0.1 | 0.0 | - | 0.1 |
| Pileup modelling | 1.2 | 0.2 | - | 2.0 |
| Non-prompt background estimation | 13.1 | 6.3 | - | 5.8 |
| Modelling | | | | |
| Background, other | 7.3 | 3.6 | - | 3.4 |
| Model statistical | 2.9 | 2.2 | - | 4.0 |
| NLO corrections | 15.1 | 4.9 | - | 0.0 |
| PDF, Scale and shower settings | 16.1 | 10.3 | - | 0.9 |
| Unfolding uncertainty | 0.0 | 0.0 | - | 0.0 |
| Experimental and modelling | 27.3 | 14.1 | - | 9.0 |
| Data statistical | 17.1 | 11.5 | - | 21.6 |
| Total | 32.2 | 18.2 | - | 23.4 |

**Table 11.12:** Uncertainty on the depth for the TT polarization state and the sum of polarization states using $\Delta Y(\ell_W Z)$ and $\Delta Y(WZ)$ in the region with $p_T^{WZ} < 70\,\mathrm{GeV}$.

| | Impact [%] | | | |
|---|---|---|---|---|
| Source | TT state | | Sum of polarizations | |
| Experimental | $\Delta Y(\ell_W Z)$ | $\Delta Y(WZ)$ | $\Delta Y(\ell_W Z)$ | $\Delta Y(WZ)$ |
| Luminosity | 1.5 | 1.0 | 0.5 | 0.2 |
| Electron calibration | 0.9 | 0.8 | 1.7 | 0.5 |
| Muon calibration | 1.6 | 1.1 | 1.4 | 0.7 |
| Jet energy scale and resolution | 3.4 | 2.8 | 1.8 | 1.6 |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 1.3 | 1.7 | 2.2 | 2.2 |
| Flavor-tagging inefficiency | 0.0 | 0.0 | 0.1 | 0.0 |
| Pileup modelling | 0.0 | 0.5 | 3.4 | 0.4 |
| Non-prompt background estimation | 9.5 | 5.7 | 13.5 | 4.6 |
| Modelling | | | | |
| Background, other | 5.7 | 3.3 | 8.0 | 2.5 |
| Model statistical | 2.4 | 2.1 | 4.6 | 2.7 |
| NLO corrections | 9.2 | 1.4 | 0.0 | 0.0 |
| PDF, Scale and shower settings | 7.5 | 5.7 | 0.7 | 0.2 |
| Unfolding uncertainty | 0.0 | 0.0 | 0.0 | 0.0 |
| Experimental and modelling | 17.0 | 9.8 | 17.2 | 6.5 |
| Data statistical | 12.8 | 10.3 | 27.0 | 13.9 |
| Total | 21.3 | 14.2 | 32.0 | 15.3 |

**Table 11.13:** Uncertainty on the depth for the TT polarization state and the sum of polarization states using $\Delta Y(\ell_W Z)$ and $\Delta Y(WZ)$ in the region with $p_T^{WZ} < 40\,\text{GeV}$.

| | Impact [%] | | | |
|---|---|---|---|---|
| Source | TT state | | Sum of polarizations | |
| Experimental | $\Delta Y(\ell_W Z)$ | $\Delta Y(WZ)$ | $\Delta Y(\ell_W Z)$ | $\Delta Y(WZ)$ |
| Luminosity | 1.4 | 0.9 | 0.4 | 0.1 |
| Electron calibration | 1.1 | 1.0 | 1.2 | 0.6 |
| Muon calibration | 1.9 | 1.3 | 1.1 | 1.0 |
| Jet energy scale and resolution | 5.1 | 4.4 | 1.2 | 1.3 |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 1.0 | 2.1 | 1.7 | 2.6 |
| Flavor-tagging inefficiency | 0.0 | 0.0 | 0.1 | 0.0 |
| Pileup modelling | 0.3 | 1.2 | 3.3 | 0.1 |
| Non-prompt background estimation | 7.2 | 4.3 | 10.0 | 3.4 |
| Modelling | | | | |
| Background, other | 4.9 | 3.0 | 6.5 | 2.1 |
| Model statistical | 2.4 | 2.2 | 4.3 | 2.7 |
| NLO corrections | 11.3 | 2.8 | 0.0 | 0.0 |
| PDF, Scale and shower settings | 9.0 | 7.0 | 0.6 | 0.4 |
| Unfolding uncertainty | 0.0 | 0.0 | 0.0 | 0.0 |
| Experimental and modelling | 18.0 | 10.9 | 13.3 | 5.8 |
| Data statistical | 13.4 | 11.0 | 26.0 | 13.9 |
| Total | 22.4 | 15.4 | 29.2 | 15.0 |

**Table 11.14:** Uncertainty on the depth for the TT polarization state and the sum of polarization states using $\Delta Y(\ell_W Z)$ and $\Delta Y(WZ)$ in the region with $p_T^{WZ} < 20\,\text{GeV}$.

| | Impact [%] | | | |
|---|---|---|---|---|
| | TT state | | Sum of polarizations | |
| Source | $\Delta Y(\ell_W Z)$ | $\Delta Y(WZ)$ | $\Delta Y(\ell_W Z)$ | $\Delta Y(WZ)$ |
| Experimental | | | | |
| Luminosity | 1.2 | 0.9 | 0.3 | 0.1 |
| Electron calibration | 1.3 | 1.2 | 0.7 | 0.3 |
| Muon calibration | 1.9 | 1.3 | 1.1 | 1.3 |
| Jet energy scale and resolution | 8.1 | 6.2 | 2.0 | 1.1 |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 0.3 | 1.4 | 0.4 | 1.5 |
| Flavor-tagging inefficiency | 0.0 | 0.0 | 0.0 | 0.0 |
| Pileup modelling | 1.3 | 2.4 | 2.7 | 0.5 |
| Non-prompt background estimation | 4.2 | 2.6 | 5.7 | 2.1 |
| Modelling | | | | |
| Background, other | 4.0 | 2.5 | 4.9 | 1.8 |
| Model statistical | 2.3 | 2.3 | 4.1 | 2.8 |
| NLO corrections | 13.3 | 6.0 | 0.0 | 0.0 |
| PDF, Scale and shower settings | 13.1 | 9.4 | 0.7 | 0.5 |
| Unfolding uncertainty | 0.0 | 0.0 | 0.0 | 0.0 |
| Experimental and modelling | 21.5 | 13.9 | 9.3 | 4.6 |
| Data statistical | 13.3 | 12.1 | 24.2 | 15.2 |
| Total | 25.3 | 18.4 | 25.9 | 15.9 |

**Table 11.15:** Uncertainty on the depth for the TT polarization state and the sum of polarization states using $\Delta Y(\ell_W Z)$ and $\Delta Y(WZ)$ in the region with $p_T^{WZ} < 10\,\text{GeV}$.

| | Impact [%] | | | |
|---|---|---|---|---|
| | TT state | | Sum of polarizations | |
| Source | $\Delta Y(\ell_W Z)$ | $\Delta Y(WZ)$ | $\Delta Y(\ell_W Z)$ | $\Delta Y(WZ)$ |
| Experimental | | | | |
| Luminosity | 1.2 | 0.9 | 0.2 | 0.1 |
| Electron calibration | 1.4 | 1.3 | 0.4 | 0.1 |
| Muon calibration | 1.9 | 1.2 | 1.0 | 0.6 |
| Jet energy scale and resolution | 7.5 | 5.0 | 2.4 | 1.0 |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 4.9 | 4.1 | 1.7 | 1.4 |
| Flavor-tagging inefficiency | 0.0 | 0.0 | 0.0 | 0.0 |
| Pileup modelling | 2.6 | 3.0 | 2.3 | 0.0 |
| Non-prompt background estimation | 2.8 | 2.1 | 3.8 | 2.0 |
| Modelling | | | | |
| Background, other | 3.6 | 2.4 | 4.2 | 1.8 |
| Model statistical | 2.7 | 2.8 | 4.7 | 3.4 |
| NLO corrections | 21.5 | 7.1 | 0.0 | 0.0 |
| PDF, Scale and shower settings | 21.0 | 14.8 | 0.6 | 0.6 |
| Unfolding uncertainty | 0.0 | 0.0 | 0.0 | 0.0 |
| Experimental and modelling | 32.0 | 18.5 | 8.3 | 4.8 |
| Data statistical | 15.4 | 14.4 | 26.3 | 18.4 |
| Total | 35.5 | 23.5 | 27.6 | 19.0 |

# Chapter 12

# Conclusion

The study of weak boson polarization opens the opportunity to probe the SM, specifically electroweak symmetry breaking, in ways that were inaccessible before. Measurements of individual boson's polarizations in wide phase spaces have been possible for a while [23–25]. Simultaneous measurements of the polarization of two bosons have only recently started to become feasible [22, 26].

This work expands the knowledge of joint $W^{\pm}Z$ boson polarization using data collected by the ATLAS detector at the LHC. The data was collected at a center-of-mass energy of 13 TeV and constitutes an integrated luminosity of $140\,\mathrm{fb}^{-1}$. Two main aspects investigated in this work are the energy dependence of $f_{00}$ and the RAZ effect.

The former is achieved by using profile likelihood fits to independently measure the fraction of doubly longitudinal events for $p_T^Z$ in the bins $(100, 200]$ GeV and $(200, \infty]$ GeV, after also applying a $p_T^{WZ} < 70\,\mathrm{GeV}$ cut to reduce the impact of higher-order corrections. Boosted decision trees are used to enhance the separation of the different polarization states. The statistical analysis uses two separate approaches to provide more precise results for more restrictive models, given the sizeable statistical and systematic uncertainties.

Using three free fit parameters the $f_{00}$ fraction is measured to be:

$$p_T^Z \in (100, 200]\ \mathrm{GeV}\colon\ 18.9\ \pm^{3.4}_{3.3}\,(\mathrm{stat.})\ \pm^{1.6}_{1.6}\,(\mathrm{mod.\ syst.})\ \pm^{1.8}_{1.5}\,(\mathrm{exp.\ syst.})\,\%,$$

$$p_T^Z \in (200,\ \infty\,]\ \mathrm{GeV}\colon\ 13\ \ \pm^{9}_{8}\,(\mathrm{stat.})\ \pm^{2}_{2}\,(\mathrm{mod.\ syst.})\ \pm^{1}_{1}\,(\mathrm{exp.\ syst.})\,\%,$$

with the former providing observation of the 00 polarization state in the $100 < p_T^Z \leq 200\,\mathrm{GeV}$ region with a significance of $5.2\sigma$. A more restrictive model with just two free parameters results in different, but compatible, values:

$$p_T^Z \in (100, 200]\ \mathrm{GeV}\colon\ 17.4\ \pm^{2.0}_{2.0}\,(\mathrm{stat.})\ \pm^{1.0}_{1.3}\,(\mathrm{mod.\ syst.})\ \pm^{0.9}_{0.8}\,(\mathrm{exp.\ syst.})\,\%,$$

$$p_T^Z \in (200,\ \infty\,]\ \mathrm{GeV}\colon\ 16\ \ \pm^{5}_{5}\,(\mathrm{stat.})\ \pm^{2}_{2}\,(\mathrm{mod.\ syst.})\ \pm^{1}_{1}\,(\mathrm{exp.\ syst.})\,\%,$$

also providing observation of the 00 state for lower $p_T^Z$ and additionally giving evidence for it in the $p_T^Z > 200\,\mathrm{GeV}$ region with a significance of $3.2\sigma$. Both of these configurations are summarized in Figure 12.1. All of these results represent deviations from the Standard Model predictions with significances slightly above $1\sigma$. Additionally, in the region with $100 < p_T^Z \leq 200\,\mathrm{GeV}$, a lower bound of $2.8\sigma$ is found for the significance of rejecting the hypothesis that the fractions of longitudinally polarized $W^{\pm}$ and $Z$ bosons are independent of each other. To investigate whether these deviations from the SM and the value for the independence represent real differences and effects or just fluctuations, the uncertainties must be reduced. The statistical uncertainties are dominant in all of the presented results. As such, the data taken during Run 3 of the LHC or by the High Luminosity LHC could serve to answer these questions. Additional improvements can be made to the prediction and consideration of NLO corrections and the parton shower. The dominant experimental uncertainties come from the jet energy scale and resolution, as well as the prediction of non-prompt backgrounds. Further improvements can be made to the measurement through the use of artificial neural

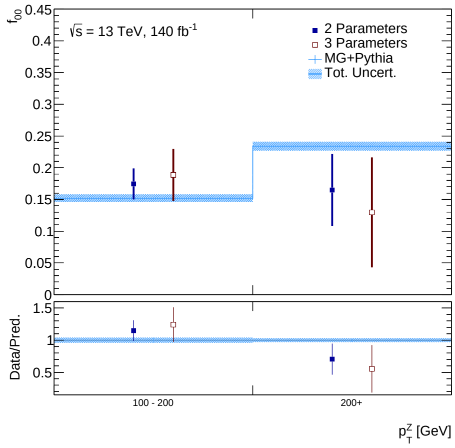



**Figure 12.1:** Summary of the three-parameter and two-parameter configuration $f_{00}$ fraction measurements compared to the Madgraph+Pythia8 prediction. The theoretical predictions include the effects of the EW k-factors described in Section 6.3, while uncertainties account for MC statistical uncertainties as well as theoretical uncertainties discussed in Section 7.3. Both the theoretical values and the uncertainties are provided by the author of Ref. [161].

networks for better polarization separation or the use of likelihood unfolding to take into account migrations between the two regions.

Meanwhile, the RAZ effect is studied using $W^{\pm}Z$ events in regions with $p_T^{WZ} < 10, 20, 40, 70\,\text{GeV}$ or where it is unconstrained. A depth variable is defined and calculated from unfolded $\Delta Y(\ell_W Z)$ and $\Delta Y(WZ)$ distributions for the TT polarization state and the sum of all polarizations. Values larger than zero indicate a dip, as expected from the RAZ effect. The results using $\Delta Y(WZ)$ for the TT polarization are shown in Figure 12.2. It demonstrates generally good agreement between expectations and data of existing dip behaviors, which increase with stronger cuts that reduce NLO effects. An estimated significance of $D/\Delta D = 7.0$ is obtained for the depth in the $p_T^{WZ} < 70\,\text{GeV}$ region, indicating an overall precise measurement of the dip. Similar overall behavior is observed for the other configurations. The largest uncertainties on the depth measurement are related to data statistics, NLO corrections, shower modeling, jet energy scale, and resolution, as well as the non-prompt background. As such, lowering the uncertainties in these areas would not only increase the precision of the fraction measurement but also help with an understanding of the RAZ effect for $W^{\pm}Z$ events. The unfolding step could also be improved through the use of likelihood-based unfolding methods, which can take correlations between uncertainties into account.

With first measurements of joint polarizations in multi-boson final states finally being available, the door is now open for the probing of $V_L V_L \to V_L V_L$ scattering as well as further studies of BSM effects utilizing the framework of effective field theories.

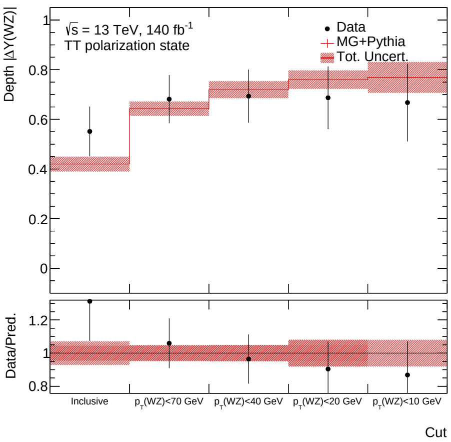



**Figure 12.2:** Progression of the depth values and their uncertainties with the tightness of the $p_T^{WZ}$ cut at the unfolded level for the TT polarization state using $\Delta Y(WZ)$.

# Appendix A

# Good Run Lists and MC Samples

This chapter contains information about the GRLs used to select useful luminosity blocks in Section A.1 as well as detailed information about the signal samples in Section A.2 and background samples in Section A.3.

## A.1 Good Run Lists

The GRLs employed in this thesis are [161]:

2015 data: `/cvmfs/atlas.cern.ch/repo/sw/database/GroupData/GoodRunsLists/data15_13TeV/20170619/data15_13TeV.periodAllYear_DetStatus-v89-pro21-02_Unknown_PHYS_StandardGRL_All_Good_25ns.xml`

2016 data: `/cvmfs/atlas.cern.ch/repo/sw/database/GroupData/GoodRunsLists/data16_13TeV/20180129/data16_13TeV.periodAllYear_DetStatus-v89-pro21-01_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml`

2017 data: `/cvmfs/atlas.cern.ch/repo/sw/database/GroupData/GoodRunsLists/data17_13TeV/20180619/data17_13TeV.periodAllYear_DetStatus-v99-pro22-01_Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim.xml`

2018 data: `/cvmfs/atlas.cern.ch/repo/sw/database/GroupData/GoodRunsLists/data18_13TeV/20190318/data18_13TeV.periodAllYear_DetStatus-v102-pro22-04_Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim.xml`

## A.2 Signal

Detailed information about the signal samples can be found in Table A.1.

## A.3 Background

Detailed information about the signal samples can be found in Table A.2.

**Table A.1:** Summary of signal MC samples used.

| DSID | Process | Generators | PDF | Events | Filter eff. | Cross-section [pb] |
|---|---|---|---|---|---|---|
| 507019 | $W_0Z_0 \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} \geq 150\,\mathrm{GeV}$) ($\ell = e,\mu$) final state | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~350K | 1.00 | 0.0024187 |
| 507020 | $W_0Z_0 \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} < 150\,\mathrm{GeV}$) ($\ell = e,\mu$) final state | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~550K | 1.00 | 0.020537 |
| 507021 | $W_0Z_T \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} \geq 150\,\mathrm{GeV}$) ($\ell = e,\mu$) final state | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~280K | 1.00 | 0.0085869 |
| 507022 | $W_0Z_T \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} < 150\,\mathrm{GeV}$) ($\ell = e,\mu$) final state | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~660K | 1.00 | 0.061592 |
| 507023 | $W_TZ_0 \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} \geq 150\,\mathrm{GeV}$) ($\ell = e,\mu$) final state | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~260K | 1.00 | 0.0025554 |
| 507024 | $W_TZ_0 \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} < 150\,\mathrm{GeV}$) ($\ell = e,\mu$) final state | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~660K | 1.00 | 0.065272 |
| 507025 | $W_TZ_T \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} \geq 150\,\mathrm{GeV}$) ($\ell = e,\mu$) final state | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~15M | 1.00 | 0.025954 |
| 507026 | $W_TZ_T \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} < 150\,\mathrm{GeV}$) ($\ell = e,\mu$) final state | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~1.66M | 1.00 | 0.30445 |
| 507027 | $W_0Z_0 \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} \geq 150\,\mathrm{GeV}$) ($\ell = e,\mu,\tau$) final state with $\tau$ decays considered | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~40K | 0.19282 | 0.003025 |
| 507028 | $W_0Z_0 \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} < 150\,\mathrm{GeV}$) ($\ell = e,\mu,\tau$) final state with $\tau$ decays considered | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~60K | 0.17178 | 0.02567 |
| 507029 | $W_0Z_T \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} \geq 150\,\mathrm{GeV}$) ($\ell = e,\mu,\tau$) final state with $\tau$ decays considered | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~70K | 0.20178 | 0.01072 |
| 507030 | $W_0Z_T \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} < 150\,\mathrm{GeV}$) ($\ell = e,\mu,\tau$) final state with $\tau$ decays considered | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~160K | 0.17366 | 0.076918 |
| 507031 | $W_TZ_0 \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} \geq 150\,\mathrm{GeV}$) ($\ell = e,\mu,\tau$) final state with $\tau$ decays considered | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~30K | 0.20127 | 0.003185 |
| 507032 | $W_TZ_0 \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} < 150\,\mathrm{GeV}$) ($\ell = e,\mu,\tau$) final state with $\tau$ decays considered | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~160K | 0.17607 | 0.081536 |
| 507033 | $W_TZ_T \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} \geq 150\,\mathrm{GeV}$) ($\ell = e,\mu,\tau$) final state with $\tau$ decays considered | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~350K | 0.19491 | 0.03242 |
| 507034 | $W_TZ_T \to \ell\nu\ell\ell$ ($p_{\mathrm{T}}^{Z} < 150\,\mathrm{GeV}$) ($\ell = e,\mu,\tau$) final state with $\tau$ decays considered | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | ~300K | 0.16602 | 0.38058 |
| 364253 | $WZ \to \ell\nu\ell\ell$ ($\ell = e,\mu,\tau$) | Sherpa 2.2.2 | NNLO NNPDF30 | ~74M | 1.00 | 4.57 |
| 361601 | $WZ \to \ell\nu\ell\ell$ ($\ell = e,\mu,\tau$) | Powheg+Pythia8 | NLO CT10 | ~30M | 1.00 | 4.51 |
| 361293 | $WZ \to \ell\nu\ell\ell$ ($\ell = e,\mu,\tau$) | MadGraph5_aMC@NLO + McAtNlo + Pythia 8 | NNLO NNPDF30 | ~3.9M | 0.34 | 1.70 |

**Table A.2:** Summary of background MC simulation.

| DSID | Process | Generators | PDF | Events | Filter eff. | Cross-section [pb] | k-factor |
|---|---|---|---|---|---|---|---|
| 364250 | $q\bar{q} \to ZZ \to \ell\ell\ell\ell$ | Sherpa 2.2.2 | NNLO NNPDF30 | 18758127 | 1.00 | 1.252 | 1.00 |
| 364283 | $ZZ \to \ell\ell\ell\ell$ EW6 | Sherpa 2.2.2 | NNLO NNPDF30 | 284282 | 1.00 | 0.011 | 1.00 |
| 345705 | $gg \to \ell\ell\ell\ell$ $(0 < m_{4\ell} < 130)$ | Sherpa 2.2.2 | NNLO NNPDF30 | 218297 | 1.00 | 0.0099 | 1.00 |
| 345706 | $gg \to \ell\ell\ell\ell$ $(m_{4\ell} > 130)$ | Sherpa 2.2.2 | NNLO NNPDF30 | 1559741 | 1.00 | 0.010 | 1.00 |
| 361106 | $Z \to ee$ | Powheg+Pythia8 | CT10 | 10046079 | 1.00 | 1901.1 | 1.00 |
| 361107 | $Z \to \mu\mu$ | Powheg+Pythia8 | CT10 | 13091504 | 1.00 | 1901.1 | 1.00 |
| 361108 | $Z \to \tau\tau$ | Powheg+Pythia8 | CT10 | 174191 | 1.00 | 1901.1 | 1.00 |
| 366140 | $Z\gamma \to ee\gamma$ $(7 < p_{\mathrm{T}}^{\gamma} < 15)$ | Sherpa 2.2.2 | NNLO NNPDF30 | 97104 | 1.00 | 46.293 | 1.00 |
| 366141 | $Z\gamma \to ee\gamma$ $(15 < p_{\mathrm{T}}^{\gamma} < 35)$ | Sherpa 2.2.2 | NNLO NNPDF30 | 465414 | 1.00 | 29.281 | 1.00 |
| 366142 | $Z\gamma \to ee\gamma$ $(35 < p_{\mathrm{T}}^{\gamma} < 70)$ | Sherpa 2.2.2 | NNLO NNPDF30 | 55151 | 1.00 | 5.1577 | 1.00 |
| 366143 | $Z\gamma \to ee\gamma$ $(70 < p_{\mathrm{T}}^{\gamma} < 140)$ | Sherpa 2.2.2 | NNLO NNPDF30 | 48792 | 1.00 | 0.40363 | 1.00 |
| 366144 | $Z\gamma \to ee\gamma$ $(p_{\mathrm{T}}^{\gamma} > 140)$ | Sherpa 2.2.2 | NNLO NNPDF30 | 68773 | 1.00 | 0.053018 | 1.00 |
| 366145 | $Z\gamma \to \mu\mu\gamma$ $(7 < p_{\mathrm{T}}^{\gamma} < 15)$ | Sherpa 2.2.2 | NNLO NNPDF30 | 135223 | 1.00 | 46.276 | 1.00 |
| 366146 | $Z\gamma \to \mu\mu\gamma$ $(15 < p_{\mathrm{T}}^{\gamma} < 35)$ | Sherpa 2.2.2 | NNLO NNPDF30 | 643346 | 1.00 | 29.28 | 1.00 |
| 366147 | $Z\gamma \to \mu\mu\gamma$ $(35 < p_{\mathrm{T}}^{\gamma} < 70)$ | Sherpa 2.2.2 | NNLO NNPDF30 | 78275 | 1.00 | 5.1566 | 1.00 |
| 366148 | $Z\gamma \to \mu\mu\gamma$ $(70 < p_{\mathrm{T}}^{\gamma} < 140)$ | Sherpa 2.2.2 | NNLO NNPDF30 | 68585 | 1.00 | 0.40322 | 1.00 |
| 366149 | $Z\gamma \to \mu\mu\gamma$ $(p_{\mathrm{T}}^{\gamma} > 140)$ | Sherpa 2.2.2 | NNLO NNPDF30 | 79303 | 1.00 | 0.05289 | 1.00 |
| 364242 | $WWW \to 3\ell 3\nu$ | Sherpa 2.2.2 | NNLO NNPDF30 | 59628 | 1.00 | 0.0072 | 1.00 |
| 364243 | $WWZ \to 4\ell 2\nu$ | Sherpa 2.2.2 | NNLO NNPDF30 | 121997 | 1.00 | 0.0018 | 1.00 |
| 364244 | $WWZ \to 2\ell 4\nu$ | Sherpa 2.2.2 | NNLO NNPDF30 | 5263 | 1.00 | 0.0035 | 1.00 |
| 364245 | $WZZ \to 5\ell 1\nu$ | Sherpa 2.2.2 | NNLO NNPDF30 | 135691 | 1.00 | 0.00019 | 1.00 |
| 364246 | $WZZ \to 3\ell 3\nu$ | Sherpa 2.2.2 | NNLO NNPDF30 | 63202 | 0.45 | 0.0017 | 1.00 |
| 364247 | $ZZZ \to 6\ell 0\nu$ | Sherpa 2.2.2 | NNLO NNPDF30 | 111922 | 1.00 | 0.000014 | 1.00 |
| 364248 | $ZZZ \to 4\ell 2\nu$ | Sherpa 2.2.2 | NNLO NNPDF30 | 73266 | 0.22 | 0.00038 | 1.00 |
| 364249 | $ZZZ \to 2\ell 4\nu$ | Sherpa 2.2.2 | NNLO NNPDF30 | 3190 | 0.44 | 0.00038 | 1.00 |
| 410470 | $t\bar{t}(\geqslant 1\ell)$ | Powheg+Pythia8 | | 24549057 | 0.54 | 729.7 | 1.00 |
| 410218 | $t\bar{t}ee$ | aMcAtNlo+Pythia 8 | | 2131415 | 1.00 | 0.037 | 1.00 |
| 410219 | $t\bar{t}\mu\mu$ | aMcAtNlo+Pythia 8 | | 2492926 | 1.00 | 0.037 | 1.00 |
| 410155 | $t\bar{t}W$ | aMcAtNlo+Pythia 8 | | 1951871 | 1.00 | 0.548 | 1.00 |
| 364739 | $W^{-}Zjj$-$EW$ OF | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | 1398000 | 1.00 | 0.015 | 1.00 |
| 364740 | $W^{+}Zjj$-$EW$ OF | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | 1996000 | 1.00 | 0.026 | 1.00 |
| 364741 | $W^{-}Zjj$-$EW$ SF | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | 699000 | 1.00 | 0.0077 | 1.00 |
| 364742 | $W^{+}Zjj$-$EW$ SF | MadGraph5_aMC@NLO + Pythia 8 | NLO NNPDF30 | 898000 | 1.00 | 0.013 | 1.00 |

# Appendix B

# RAZ Derivation

The basis for the derivation of the RAZ effect is the amplitudes for the individual polarization states [27]

$$
\begin{aligned}
\mathcal{M}(\pm,\mp) &= F\sin\theta\,(\lambda_{\mathrm{w}}-\cos\theta)\,X,\\
\mathcal{M}(0,0) &= F\frac{\sin\theta}{\sqrt{2r_{\mathrm{w}}2r_{\mathrm{z}}}}\left[\left(-\beta\rho+\beta_{\mathrm{w}}\beta_{\mathrm{z}}\cos\theta+\beta\rho\frac{\alpha-\beta\cos\theta}{1-r_{\mathrm{w}}}\right)X+2\beta\rho Y\right].
\end{aligned}
\tag{B.1}
$$

It was already shown in Chapter 2 that the other polarization states scale towards zero for $s \gg M_V^2$. The derivation of Equations 2.69 based on Equations B.1 is shown in the following:

$$
\begin{aligned}
\mathcal{M}(\pm,\mp) &= F\sin\theta\,(\lambda_{\mathrm{w}}-\cos\theta)\,X\\
&= F\sin\theta\,(\lambda_{\mathrm{w}}-\cos\theta)\,\frac{s}{2}\left(\frac{g_-^{f_1}}{u}+\frac{g_-^{f_2}}{t}\right)\\
&= F\sin\theta\,(\lambda_{\mathrm{w}}-\cos\theta)\,\frac{\cancel{s}}{\cancel{2}}\left(\frac{g_-^{f_1}}{-\frac{\cancel{s}}{\cancel{2}}(1+\cos\theta)}+\frac{g_-^{f_2}}{-\frac{\cancel{s}}{\cancel{2}}(1-\cos\theta)}\right)\\
&= F\sin\theta\,(\lambda_{\mathrm{w}}-\cos\theta)\left(\frac{-g_-^{f_1}}{1+\cos\theta}+\frac{-g_-^{f_2}}{1-\cos\theta}\right)\\
&= F\sin\theta\,(\lambda_{\mathrm{w}}-\cos\theta)\,\frac{-g_-^{f_1}(1-\cos\theta)-g_-^{f_2}(1+\cos\theta)}{\sin^2\theta}\\
&= \frac{F}{\sin\theta}\,(\lambda_{\mathrm{w}}-\cos\theta)\left[\left(g_-^{f_1}-g_-^{f_2}\right)\cos-\theta-\left(g_-^{f_1}+g_-^{f_2}\right)\right],
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{M}(0,0) &= F\frac{\sin\theta}{\sqrt{2r_\mathrm{w}2r_\mathrm{z}}}\left[\left(-\beta\rho+\beta_\mathrm{w}\beta_\mathrm{z}\cos\theta+\beta\rho\frac{\alpha-\beta\cos\theta}{1-r_\mathrm{w}}\right)X+2\beta\rho Y\right] \\
&= F\frac{s\cdot\sin\theta}{2M_W M_Z}\left[\frac{M_Z^2}{s}(-1+\cos\theta)X+2Y\right] \\
&= F\frac{s\cdot\sin\theta}{2M_W M_Z}\left[\frac{M_Z^2}{s}(-1+\cos\theta)\left[\left(\frac{-g_-^{f_1}}{1+\cos\theta}+\frac{-g_-^{f_2}}{1-\cos\theta}\right)\right]-2\frac{g_-^{f_1}M_Z^2}{s(1+\cos\theta)}\right] \\
&= \frac{F}{2}\sin\theta\frac{M_Z}{M_W}\left[(1-\cos\theta)\left[\left(\frac{g_-^{f_1}}{1+\cos\theta}+\frac{g_-^{f_2}}{1-\cos\theta}\right)\right]-2\frac{g_-^{f_1}}{1+\cos\theta}\right] \\
&= \frac{F}{2}\sin\theta\frac{M_Z}{M_W}\left[g_-^{f_2}+\frac{g_-^{f_1}(1-\cos\theta)}{1+\cos\theta}-2\frac{g_-^{f_1}}{1+\cos\theta}\right] \\
&= \frac{F}{2}\sin\theta\frac{M_Z}{M_W}\left[g_-^{f_2}+\frac{g_-^{f_1}-\cos\theta g_-^{f_1}-2g_-^{f_1}}{1+\cos\theta}\right] \\
&= \frac{F}{2}\sin\theta\frac{M_Z}{M_W}\left[g_-^{f_2}+\frac{-g_-^{f_1}(1+\cos\theta)}{1+\cos\theta}\right] \\
&= \frac{F}{2}\sin\theta\frac{M_Z}{M_W}\left(g_-^{f_2}-g_-^{f_1}\right),
\end{aligned}
$$

using Equations 2.64 and [27]

$$
\begin{aligned}
t &= -\frac{s}{2}(\alpha-\beta\cos\theta), \\
u &= -\frac{s}{2}(\alpha+\beta\cos\theta).
\end{aligned}
$$

# Appendix C

# More Comparisons to Unpolarized NLO MC Simulations

More comparison plots between the polarized samples and unpolarized NLO samples can be found in Figures C.1 and C.2.

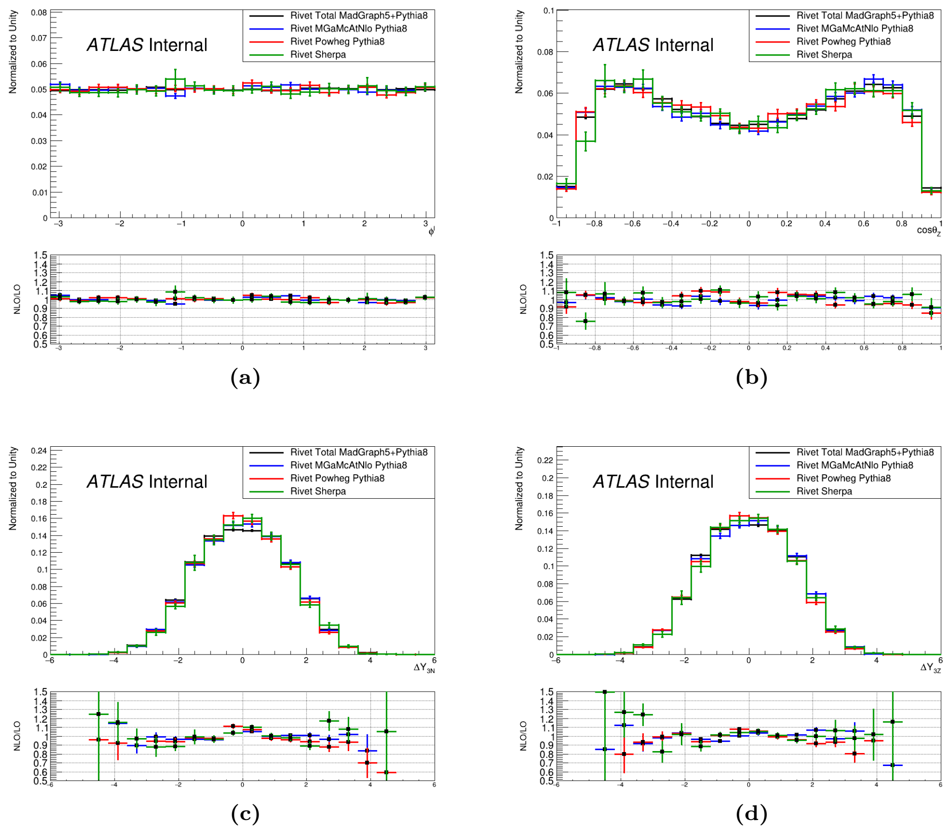



**Figure C.1:** Comparison between NLO (Powheg+Pythia8, MadGraph5_aMC@NLO and Sherpa) and the sum of LO polarized samples for (a) $\phi$ for all the three leptons, (b) $\cos\theta_Z$ ,(c) $\Delta Y(\ell_W \ell_Z^-)$ and (c) $\Delta Y(\ell_W Z)$, in the $W_L Z_L$ enhanced phase space with $p_T^Z > 200$ GeV, in the $W_0 Z_0$ enhanced phase space with $p_T^Z > 200$ GeV. From the author of Ref. [161].

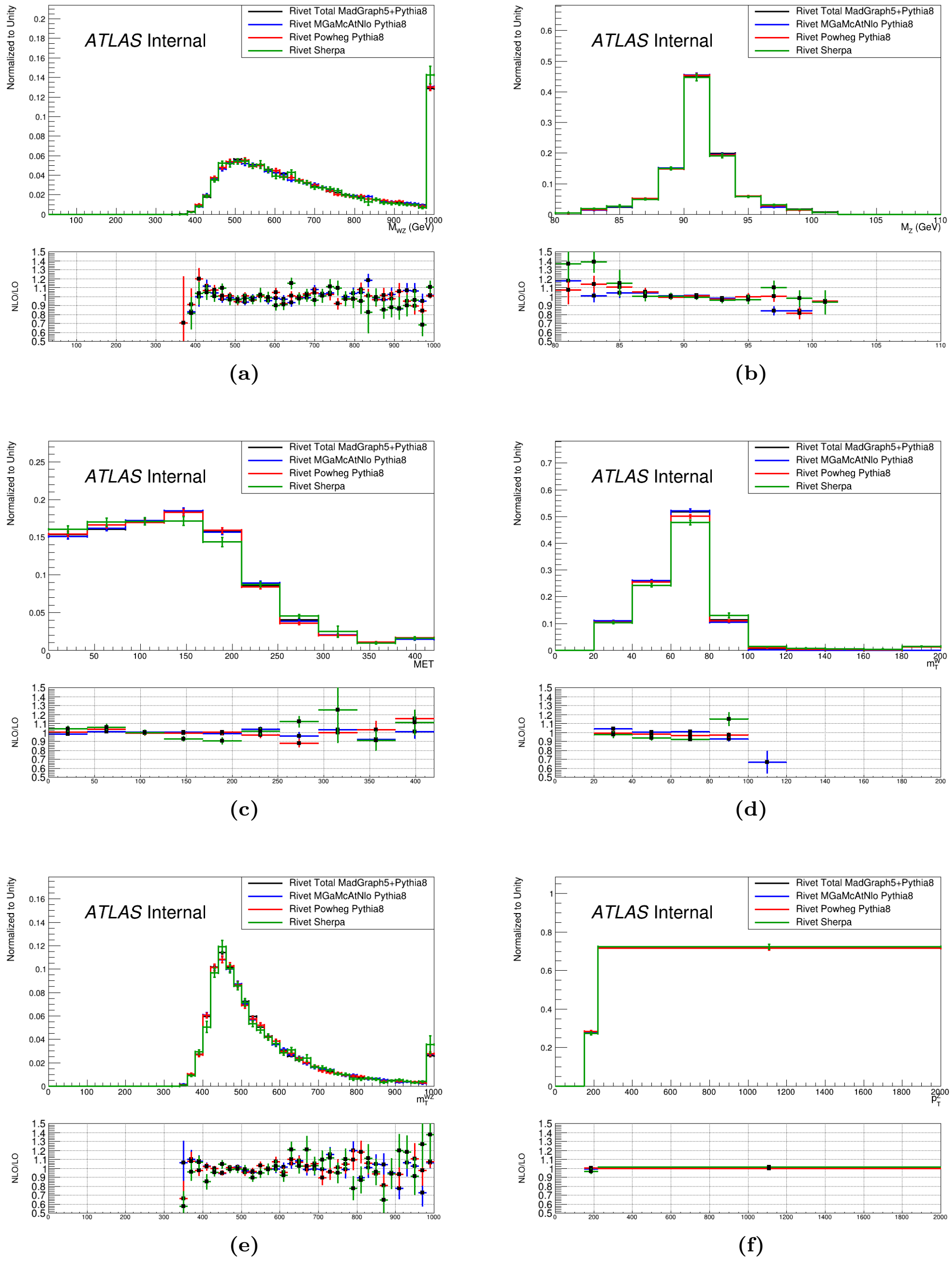



**Figure C.2:** Comparison between NLO (Powheg+Pythia8, MadGraph5_aMC@NLO and Sherpa) and the sum of LO polarized samples for (a) $M_{WZ}$, (b) $M_Z$, (c) $E_T^{\text{miss}}$, (d) $M_T^W$ (e) $M_T^{WZ}$ and (f) $p_T^Z$, in the $W_0Z_0$ enhanced phase space with $p_T^Z > 200\,\text{GeV}$. From the author of Ref. [161].

# Appendix D

# Fit with $p_T^Z$ Reweighted Signal

The sensitivity of the BDT template fit to physics beyond the SM has to be checked. Specifically for new physics that alters the $p_T^Z$-dependence of $f_{00}$ since this is exactly what is being measured in the 00-enhanced regions.

To do this, pseudo-data is created by stacking all the polarization and non-polarization backgrounds as well as a reweighted 00 signal. The signal region with $p_T^Z > 200$ GeV and $p_T^{WZ} < 70$ GeV is used. Except for the reweighted 00 signal, the pseudo-data is equivalent to Asimov data. Four different $p_T^Z$-dependent reweighting functions are investigated. They are summarized in Table D.1. The pseudo-data is then fitted to by the normal, not reweighted, templates. The fitting approach follows the one described in Chapter 9. This procedure is done for three different BDT variations. All three BDTs are generated according to the procedure described in Section 8.3. The training inputs are slightly different from those used in the final nominal analysis due to updates in the analysis release and luminosity since these studies were performed. As such, the exact BDTs differ from the nominal one. The conclusions should not be affected by this, however. The same holds for the applied k-factors during the fit. The differences are which variables are allowed for the optimization procedure. For the first BDT, all variables are allowed. The second BDT allows all variables except for $p_T^Z$ and $p_T^W$. Lastly, a BDT is trained with only angular variables. Table D.2 contains the final variables used for each BDT.

**Table D.1:** Reweighting functions used to investigate the influence of modifications to the $p_T^Z$ dependence of the 00 polarization component due to BSM physics on the fit results.

| Name | Function f($p_T^Z$) | Value at 200 GeV | Value at 400 GeV | Value at 1000 GeV |
|---|---|---|---|---|
| 0.05 | $1 + 0.05 \times (p_T^Z - 200)$ | 1 | 11 | 41 |
| 0.01 | $1 + 0.01 \times (p_T^Z - 200)$ | 1 | 3 | 9 |
| sq | $1 + (\frac{p_T^Z - 200}{200})^2$ | 1 | 2 | 17 |
| div | $\frac{200}{p_T^Z - 100}$ | 2 | 0.67 | 0.22 |

The results of stat-only fits to normal Asimov data using the three-sample configuration and the three different BDTs are summarized in Table D.3. All three fits reproduce the correct central values. However, the BDT using only angular variables has increased uncertainties and, thus, a decreased significance due to a poorer separation between the polarization signal and backgrounds.

Using the '0.05' reweighting function to produce pseudo-data produces fit results as shown in Table D.4. It can be seen that the BDT that includes $p_T^Z$ and $p_T^W$ produces results that are farthest (about 2 $\sigma$) from the truth values in the pseudo-data. Removing both $p_T^V$ from the optimization process improves the agreement by a minimal amount. The smallest deviation is found when using only angular variables. For the first two configurations, the truth value lies within two standard deviations from the fit result, while one standard deviation is enough for the third one. Figure D.1 shows the pre- and post-fit distributions for the fit of the full BDT

**Table D.2:** List of variables used in the three BDTs after optimization.

| Variables | | |
|---|---|---|
| Full BDT | No $p_T^V$ | Only Angular |
| $\Delta Y(\ell_W Z)$ | $\Delta Y(\ell_W Z)$ | $\Delta Y(\ell_W Z)$ |
| $P_T^{WZ}$ | $P_T^{WZ}$ | $cos\theta_{\ell_Z}$ |
| $p_T(\ell_W)$ | $p_T(\ell_W)$ | $cos\theta_{\ell_W}$ |
| $p_T(\ell_2^Z)$ | $p_T(\ell_2^Z)$ | $\Delta Y(WZ)$ |
| $cos\theta_{\ell_W}$ | $cos\theta_{\ell_W}$ | $\Delta\Phi(\ell_W, l_Z^-)(WZ)$ |
| $M_T^{WZ}$ | $M_T^{WZ}$ | $\Delta\Phi(\ell_W, l_Z^-)$(lab) |
| $r_{21}$ | $r_{21}$ | $\Delta Y(\ell_W, \ell_Z^-)$ |
| $M_{WZ}$ | $M_{WZ}$ | |
| $p_T^Z$ | | |
| $\Delta Y(WZ)$ | | |
| $\Delta Y(\ell_W, \ell_Z^-)$ | | |

**Table D.3:** Fitting results of the different BDTs to Asimov data.

| | Full | No $p_T(Z)$, $p_T(W)$ | Only Angular |
|---|---|---|---|
| $significance$ | $3.0\,\sigma$ | $3.0\,\sigma$ | $2.7\,\sigma$ |
| $f_{00}$ | $0.25\pm^{0.09}_{0.09}$ | $0.25\pm^{0.09}_{0.09}$ | $0.25\pm^{0.10}_{0.10}$ |
| $f_{Mixed}$ | $0.11\pm^{0.17}_{0.11}$ | $0.11\pm^{0.17}_{0.11}$ | $0.11\pm^{0.18}_{0.11}$ |
| $N_{tot}$ | $250\pm^{18}_{17}$ | $250\pm^{18}_{17}$ | $250\pm^{18}_{17}$ |
| $f_{TT}$ | 0.64 | 0.64 | 0.64 |
| $\mu$ | 1.0 | 1.0 | 1.0 |
| $\mu_{WZTT}$ | 1.0 | 1.0 | 1.0 |
| $\mu_{WZMixed}$ | 1.0 | 1.0 | 1.0 |

to the '0.05' pseudo-data.

The results for the other three reweighting functions can be found in Tables D.5, D.6 and D.7. For all of those cases, the full BDT produces results that have the largest disagreement to the truth value while removing $p_T^W$ and $p_T^Z$ gives a very slight improvement. Using only angular variables leads to a better agreement between fit results and truth values but also lower significances. For all of these configurations, the truth value lies within 1 $\sigma$ of the fit result.

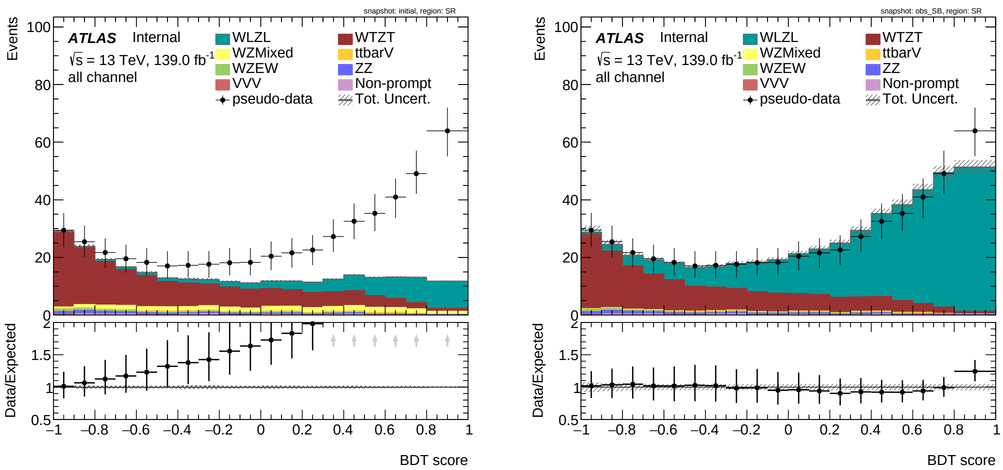
Using the two-sample configuration for the fit instead leads to better agreement between the fit and truth values at the expense of an overall worse agreement between pseudo-data and templates in the post-fit distribution. This can be seen in Table D.8 and Figure D.2.

To avoid having a direct dependence of the BDT on the variable of interest, $p_T^Z$, it was decided to omit $p_T^Z$ and $p_T^W$ from the final BDT optimization procedure. A small bias remains as a lot of variables are correlated with the transverse momenta of the bosons. However, no additional variables are removed due to this effect because:

- Most of the more realistic reweighting configurations still have the correct value within the $1\sigma$ uncertainty.
- Even the "Angular Only" BDT, which contains only variables with minimal individual correlation to $p_T^V$, shows a small bias.
- There is a sizeable drop in significance, which outweighs the slight reduction in the bias.

**Table D.4:** Fitting results of the different BDTs to pseudo-data produced with the '0.05' reweighting function.

| | Full | No $p_T^V$ | Only Angular | "Truth"-pseudo |
|---|---|---|---|---|
| $significance$ | $11.1\,\sigma$ | $11.1\,\sigma$ | $9.1\,\sigma$ | |
| $f_{00}$ | $0.68\pm^{0.03}_{0.04}$ | $0.68\pm^{0.03}_{0.04}$ | $0.66\pm^{0.03}_{0.07}$ | 0.617 |
| $f_{Mixed}$ | $0.0\pm^{0.05}_{0.0}$ | $0.0\pm^{0.05}_{0.0}$ | $0.0\pm^{0.10}_{0.00}$ | 0.057 |
| $N_{tot}$ | $490\pm^{25}_{25}$ | $490\pm^{26}_{25}$ | $490\pm^{26}_{25}$ | 490 |
| $f_{TT}$ | 0.32 | 0.33 | 0.34 | 0.33 |
| $\mu$ | 5.3 | 5.3 | 5.2 | 4.9 |
| $\mu_{WZTT}$ | 1.0 | 1.0 | 1.0 | 1.0 |
| $\mu_{WZMixed}$ | 0.0 | 0.0 | 0.0 | 1.0 |




(a)

(b)

**Figure D.1:** Shown are (a) pre-fit and (b) post-fit distributions for the fit of the full BDT to pseudo-data produced with the '0.05' reweighting function using the 3 sample fitting configuration.

**Table D.5:** Fitting results of the different BDTs to pseudo-data produced with the '0.01' reweighting function.

| | Full | No $p_T^V$ | Only Angular | "Truth"-pseudo |
|---|---|---|---|---|
| $significance$ | $5.2\,\sigma$ | $5.2\,\sigma$ | $4.3\,\sigma$ | |
| $f_{00}$ | $0.42\pm^{0.06}_{0.09}$ | $0.42\pm^{0.06}_{0.09}$ | $0.39\pm^{0.08}_{0.10}$ | 0.37 |
| $f_{Mixed}$ | $0.03\pm^{0.16}_{0.03}$ | $0.03\pm^{0.16}_{0.03}$ | $0.06\pm^{0.17}_{0.06}$ | 0.09 |
| $N_{tot}$ | $298\pm^{20}_{19}$ | $298\pm^{20}_{19}$ | $298\pm^{20}_{19}$ | 298 |
| $f_{TT}$ | 0.56 | 0.56 | 0.55 | 0.54 |
| $\mu$ | 2.0 | 2.0 | 1.9 | 1.8 |
| $\mu_{WZTT}$ | 1.0 | 1.0 | 1.0 | 1.0 |
| $\mu_{WZMixed}$ | 0.3 | 0.3 | 0.7 | 1.0 |

**Table D.6:** Fitting results of the different BDTs to pseudo-data produced with the 'sq' reweighting function.

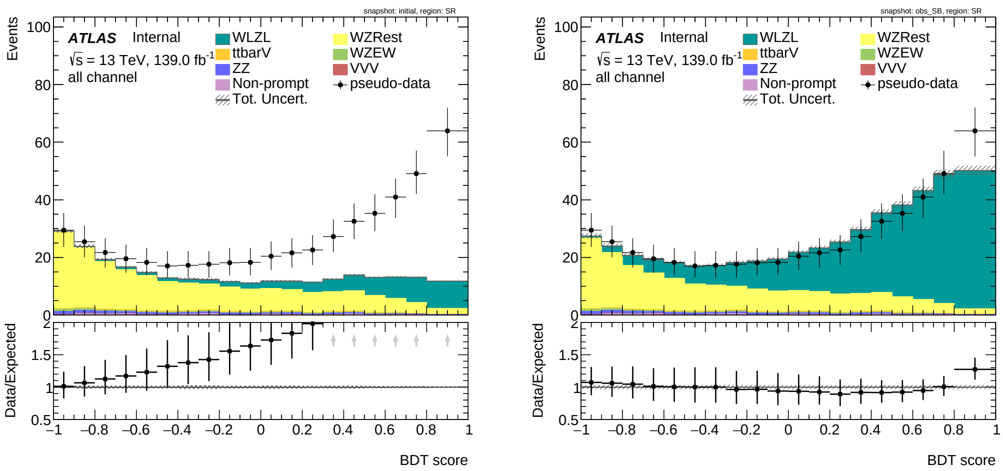
| | Full | No $p_T^V$ | Only Angular | "Truth"-pseudo |
|---|---|---|---|---|
| $significance$ | $4.3\,\sigma$ | $4.2\,\sigma$ | $3.5\,\sigma$ | |
| $f_{00}$ | $0.35\pm^{0.07}_{0.09}$ | $0.35\pm^{0.07}_{0.09}$ | $0.33\pm^{0.08}_{0.1}$ | 0.31 |
| $f_{Mixed}$ | $0.03\pm^{0.16}_{0.04}$ | $0.05\pm^{0.13}_{0.05}$ | $0.08\pm^{0.18}_{0.08}$ | 0.10 |
| $N_{tot}$ | $271\pm^{19}_{8}$ | $271\pm^{19}_{18}$ | $271\pm^{19}_{18}$ | 270 |
| $f_{TT}$ | 0.61 | 0.61 | 0.60 | 0.59 |
| $\mu$ | 1.5 | 1.5 | 1.4 | 1.3 |
| $\mu_{WZTT}$ | 1.0 | 1.0 | 1.0 | 1.0 |
| $\mu_{WZMixed}$ | 0.4 | 0.5 | 0.7 | 1.0 |

**Table D.7:** Fitting results of the different BDTs to pseudo-data produced with the 'div' reweighting function.

| | Full | No $p_T^V$ | Only Angular | "Truth"-pseudo |
|---|---|---|---|---|
| $significance$ | $3.3\,\sigma$ | $3.3\,\sigma$ | $3.1\,\sigma$ | |
| $f_{00}$ | $0.27\pm^{0.10}_{0.09}$ | $0.27\pm^{0.09}_{0.09}$ | $0.29\pm^{0.10}_{0.10}$ | 0.30 |
| $f_{Mixed}$ | $0.16\pm^{0.16}_{0.16}$ | $0.16\pm^{0.16}_{0.16}$ | $0.13\pm^{0.17}_{0.13}$ | 0.10 |
| $N_{tot}$ | $270\pm^{19}_{18}$ | $270\pm^{19}_{18}$ | $270\pm^{19}_{18}$ | 269 |
| $f_{TT}$ | 0.57 | 0.57 | 0.58 | 0.59 |
| $\mu$ | 1.2 | 1.2 | 1.2 | 1.3 |
| $\mu_{WZTT}$ | 1.0 | 1.0 | 1.0 | 1.0 |
| $\mu_{WZMixed}$ | 1.5 | 1.5 | 1.3 | 1.0 |

**Table D.8:** Fitting results of the different BDTs to pseudo-data produced with the '0.05' reweighting function for the 2 sample configuration.

| | Full | No $p_T^V$ | Only Angular | "Truth"-pseudo |
|---|---|---|---|---|
| $significance$ | $20.4\,\sigma$ | $20.4\,\sigma$ | $18.3\,\sigma$ | |
| $f_{00}$ | $0.65\pm^{0.04}_{0.04}$ | $0.65\pm^{0.04}_{0.04}$ | $0.63\pm^{0.04}_{0.04}$ | 0.62 |
| $N_{tot}$ | $490\pm^{10}_{25}$ | $490\pm^{10}_{25}$ | $490\pm^{11}_{5}$ | 490 |
| $f_{Rest}$ | 0.35 | 0.35 | 0.37 | 0.38 |
| $\mu$ | 5.1 | 5.1 | 4.9 | 4.9 |
| $\mu_{WZRest}$ | 0.9 | 0.9 | 1.0 | 1.0 |




**(a)** **(b)**

**Figure D.2:** Shown are (a) pre-fit and (b) post-fit distributions for the fit of the full BDT to pseudo-data produced with the ‘0.05’ reweighting function using the 2 sample fitting configuration.

# Appendix E

# Fit Eliminating a Smaller Fraction

As discussed in Section 10.5, the parametrizations using four free parameters suffer from large anti-correlations between the mixed polarization states. This causes a strong dependence of the fit results and uncertainties on the choice of which fraction is implicitly defined as 1 minus the sum of the other ones. The nominal choice is to use the fraction with the largest expected value and smallest expected uncertainties. This ensures that the fraction does not reach an unphysical value $< 0$, either as a best-fit value or within the uncertainties. Non-physical results are especially problematic in the S+B fit as they can inflate the measured significance.

To demonstrate the issue, fits are performed in the region with $p_T^Z > 200\,\text{GeV}$. These are shown in Table E.1 for the case where the T0 state is defined implicitly and in Table E.2 where the 0T state is chosen instead. These results can be compared to the ones shown in Table 10.14, where the TT state is eliminated.

**Table E.1:** Fit results using the smoothed nuisance parameter distributions in the configuration with four independent fit parameters for the fiducial fraction parametrization. Here $f_{T0}$ is defined implicitly via $f_{T0} = 1 - (f_{00} + f_{0T} + f_{TT})$.

| | Exp. | Obs. |
|---|---|---|
| Significance | $2.168\,\sigma$ | $1.704\,\sigma$ |
| $f_{00}$ | $0.23\pm^{0.10}_{0.11}$ | $0.13\pm^{0.08}_{0.07}$ |
| $f_{TT}$ | $0.65\pm^{0.15}_{0.13}$ | $0.63\pm^{0.13}_{0.11}$ |
| $f_{0T}$ | $0.1\pm^{0.6}_{0.1}$ | $0.00\pm^{0.31}_{0.00}$ |
| $N_{tot}$ | $330\pm^{31}_{33}$ | $360\pm^{31}_{29}$ |
| $f_{T0}$ | $0.06 \pm 0.31$ | $0.24 \pm 0.23$ |
| $\mu$ | $1.0 \pm 0.5$ | $0.6 \pm 0.4$ |
| $\mu_{WZLT}$ | $1 \pm 5$ | $0.0 \pm 2.7$ |
| $\mu_{WZTT}$ | $1.00 \pm 0.25$ | $1.07 \pm 0.22$ |
| $\mu_{WZTL}$ | $1 \pm 5$ | $4 \pm 4$ |

It can be seen that both configurations produce significant deviations compared to the nominal choice. While eliminating the T0 state does not cause many differences in the observed fit, it does so for the expected significance. This happens because the 0T state produces better results in the background-only fit to Asimov data and is set to a value of 0.92, leaving $f_{T0}$ with a large negative value. This leads to a reduction of the significance as the full unconstrained fit can not be improved through nonphysical parameters. The only change in the observed fit that does occur is that the uncertainty on $f_{TT}$ is smaller than the one calculated from propagation, as here the correlations between the mixed fractions are taken into account despite the nominal value of $f_{0T}$ being zero.

**Table E.2:** Fit results using the smoothed nuisance parameter distributions in the configuration with four independent fit parameters for the fiducial fraction parametrization. Here $f_{0T}$ is defined implicitly via $f_{0T} = 1 - (f_{00} + f_{T0} + f_{TT})$.

| | Exp. | Obs. |
|---|---|---|
| Significance | $2.279\,\sigma$ | $1.837\,\sigma$ |
| $f_{00}$ | $0.23\pm^{0.11}_{0.10}$ | $0.16\pm^{0.10}_{0.09}$ |
| $f_{TT}$ | $0.65\pm^{0.15}_{0.13}$ | $0.67\pm^{0.14}_{0.12}$ |
| $f_{T0}$ | $0.1\pm^{0.5}_{0.1}$ | $0.5\pm^{0.5}_{0.5}$ |
| $N_{tot}$ | $330\pm^{35}_{30}$ | $370\pm^{40}_{40}$ |
| $f_{0T}$ | $0.1 \pm 0.4$ | $-0.4 \pm 0.5$ |
| $\mu$ | $1.0 \pm 0.5$ | $0.8 \pm 0.5$ |
| $\mu_{WZTL}$ | $1 \pm 5$ | $10 \pm 10$ |
| $\mu_{WZTT}$ | $1.00 \pm 0.25$ | $1.17 \pm 0.26$ |
| $\mu_{WZLT}$ | $1 \pm 6$ | $-6 \pm 10$ |

The opposite can be observed when making $f_{0T}$ implicit. Here, the expected significance stays the same, as the best background-only fit sets $f_{T0} = 0$ whether the 0T or the TT states are set implicitly. Only the uncertainties on all of the fractions are larger. However, as already seen in Table 10.14, the shape of the distributions for the T0 state match better to observed data, thus increasing it as much as possible at the expense of the 0T state. With the latter now not having a hard lower bound of 0 anymore, the fit reaches unphysical regions for the S+B fit to observed data. This leads to a best-fit value of $f_{0T} = -0.4$ and an inflated significance.

Which of the two mixed states fits better to observed data can not easily be known before performing a fit. As such, it is best to leave both as explicit parameters in the fit to avoid unphysical regions.

# Appendix F

# Relations and Constraints in the Independence Parametrization

This section states and proves some more facts and relations regarding the independence parametrization of the fit introduced in Section 9.3.1.

Using Equations 9.47 and 9.48, it can be shown that

$$
\begin{aligned}
f_{TT} &= (1 - f_{Z0}) \times (1 - f_{W0}) + (R_c - 1) \times f_{W0} \times f_{Z0} \\
&= (f_{TT} + f_{0T})(f_{TT} + f_{T0}) + \left( \frac{f_{00}}{f_{W0} f_{Z0}} - 1 \right) f_{W0} f_{Z0} \\
&= (f_{TT} + f_{0T})(f_{TT} + f_{T0}) + \left( \frac{f_{00}}{f_{W0} f_{Z0}} - \frac{f_{W0} f_{Z0}}{f_{W0} f_{Z0}} \right) f_{W0} f_{Z0} \\
&= (f_{TT} + f_{0T})(f_{TT} + f_{T0}) + \left( \frac{f_{00} - f_{W0} f_{Z0}}{\cancel{f_{W0} f_{Z0}}} \right) \cancel{f_{W0} f_{Z0}} \\
&= (f_{TT} + f_{0T})(f_{TT} + f_{T0}) + (f_{00} - f_{W0} f_{Z0}) \\
&= (f_{TT} + f_{0T})(f_{TT} + f_{T0}) - (f_{00} + f_{0T})(f_{00} + f_{T0}) + f_{00} \\
&= f_{TT} f_{TT} + f_{TT} f_{0T} + f_{TT} f_{T0} + \cancel{f_{0T} f_{T0}} \\
&\quad - f_{00} f_{00} + f_{00} f_{0T} + f_{00} f_{T0} + \cancel{f_{0T} f_{T0}} \\
&\quad + f_{00} \\
&= f_{TT}(f_{TT} + f_{0T} + f_{T0}) - f_{00}(f_{00} + f_{0T} + f_{T0}) + f_{00} \\
&= f_{TT}(1 - f_{00}) - f_{00}(1 - f_{TT}) + f_{00} \\
&= f_{TT} - \cancel{f_{TT} f_{00}} - \cancel{f_{00}} + \cancel{f_{00} f_{TT}} + \cancel{f_{00}} \\
&= f_{TT}
\end{aligned}
$$

and

$$\begin{aligned}
f_{0T} &= f_{W0} \times (1 - (R_c \times f_{Z0})) \\
&= (f_{00} + f_{0T}) \times \left(1 - \frac{f_{00}}{f_{W0}\cancel{f_{Z0}}}\cancel{f_{Z0}}\right) \\
&= (f_{00} + f_{0T}) \times \left(1 - \frac{f_{00}}{f_{W0}}\right) \\
&= (f_{00} + f_{0T}) \times \left(1 - \frac{f_{00}}{f_{00} + f_{0T}}\right) \\
&= (f_{00} + f_{0T}) \times \left(\frac{\cancel{f_{00}} + f_{0T}}{f_{00} + f_{0T}} - \frac{\cancel{f_{00}}}{f_{00} + f_{0T}}\right) \\
&= \cancel{(f_{00} + f_{0T})} \times \frac{f_{0T}}{\cancel{f_{00} + f_{0T}}} \\
&= f_{0T}
\end{aligned}$$

Additionally, it can be shown that this parametrization respects that the sum of fractions has to be one

$$\begin{aligned}
f_{00} + f_{0T} + f_{T0} + f_{TT} = \\
&= R_c f_{W0} f_{Z0} \\
&\quad + f_{W0}(1 - R_c f_{Z0}) \\
&\quad + f_{Z0}(1 - R_c f_{W0}) \\
&\quad + (1 - f_{Z0})(1 - f_{W0}) + (R_c - 1) f_{W0} f_{Z0} \\
&= R_c f_{W0} f_{Z0} \\
&\quad + f_{W0} - R_c f_{W0} f_{Z0} \\
&\quad + f_{Z0} - R_c f_{W0} f_{Z0} \\
&\quad + 1 - f_{Z0} - f_{W0} + f_{W0} f_{Z0} + R_c f_{W0} f_{Z0} - f_{W0} f_{Z0} \\
&= 1
\end{aligned}$$

Like the fraction parametrization, this one suffers from the fact that the boundaries for all of the implicit fractions can not be defined. Only those for $f_{W0}$ and $f_{Z0}$ can be forced to be in the physical range. This means that all of the other fractions can, in principle, become larger than one, and all but $f_{00}$ can become smaller than zero. For $f_{00}$ this can be avoided by setting a lower bound of zero for $R_c$.

A set of example values where this happens is

$$\begin{aligned}
R_c &= 5, \\
f_{W0} &= 0.5, \\
f_{Z0} &= 0.5.
\end{aligned}$$

In that case, one gets

$$\begin{aligned}
f_{00} &= 1.25, \\
f_{0T} &= -0.75, \\
f_{T0} &= -0.75, \\
f_{TT} &= 1.25.
\end{aligned}$$

Consequently for

$$\begin{aligned}
R_c &= 0, \\
f_{W0} &= 1, \\
f_{Z0} &= 1,
\end{aligned}$$

the results are

$$
\begin{aligned}
f_{00} &= 0, \\
f_{0T} &= 1, \\
f_{T0} &= 1, \\
f_{TT} &= -1.
\end{aligned}
$$

# Appendix G

# Fit in the Inclusive $p_T^Z > 100\,\text{GeV}$ Region

At some point during the development of the analysis, it was planned to use an inclusive region with $p_T^Z > 100\,\text{GeV}$ alongside the region with $p_T^Z > 200\,\text{GeV}$. it was later discarded in favor of the exclusive $100 < p_T^Z \leq 200\,\text{GeV}$ due to an easier interpretability and combination of the results. The fit results using the fiducial fraction parametrization with two and three free parameters are listed in Tables G.1 and G.2, respectively, for completion sake. Following expectations, the significances are higher than those in the region with $100 < p_T^Z \leq 200\,\text{GeV}$ with higher $N_{tot}$ and 00 fractions between those of the two exclusive regions. Curiously, the deviations in the two regions almost exactly cancel the differences between expectations and data in those regions. Due to this, the observed and expected values agree almost perfectly.

**Table G.1:** Fit results using the smoothed nuisance parameter distributions in the configuration with two independent fit parameters for the fiducial fraction parametrization in the $p_T^Z > 100\,\text{GeV}$ region.

| | Exp. | Obs. |
|---|---|---|
| Significance | $7.692\,\sigma$ | $7.567\,\sigma$ |
| $f_{00}$ | $0.163\pm^{0.022}_{0.022}$ | $0.165\pm^{0.023}_{0.024}$ |
| $N_{tot}$ | $2830\pm^{130}_{130}$ | $3050\pm^{140}_{140}$ |
| $f_{Rest}$ | $0.837 \pm 0.022$ | $0.835 \pm 0.023$ |
| $\mu$ | $1.00 \pm 0.14$ | $1.09 \pm 0.16$ |
| $\mu_{WZRest}$ | $1.00 \pm 0.05$ | $1.08 \pm 0.06$ |

**Table G.2:** Fit results using the smoothed nuisance parameter distributions in the configuration with three independent fit parameters for the fiducial fraction parametrization in the $p_T^Z > 100\,\text{GeV}$ region.

| | Exp. | Obs. |
|---|---|---|
| Significance | $5.358\,\sigma$ | $5.313\,\sigma$ |
| $f_{00}$ | $0.162\pm^{0.034}_{0.033}$ | $0.16\pm^{0.04}_{0.04}$ |
| $f_{Mixed}$ | $0.22\pm^{0.08}_{0.09}$ | $0.22\pm^{0.08}_{0.09}$ |
| $N_{tot}$ | $2840\pm^{140}_{140}$ | $3060\pm^{140}_{140}$ |
| $f_{TT}$ | $0.62 \pm 0.06$ | $0.62 \pm 0.06$ |
| $\mu$ | $1.00 \pm 0.21$ | $1.08 \pm 0.24$ |
| $\mu_{WZTT}$ | $1.00 \pm 0.10$ | $1.07 \pm 0.12$ |
| $\mu_{WZMixed}$ | $1.0 \pm 0.4$ | $1.1 \pm 0.4$ |

# Appendix H

# Fit in Both Regions Simultaneously

Instead of fitting both of the $p_T^Z$ regions separately, it is possible to fit both of them simultaneously and exploit correlations between nuisance parameters to reduce uncertainties. Each region has separate free parameters for its overall scaling and fractions. However, all experimental and systematic nuisance parameters are treated as fully correlated across the regions. This effectively performs a likelihood unfolding under the assumption that there are no migrations between the regions. To cross-check how valid this assumption is, the migration matrices between the two nominal 00-enriched signal regions are investigated and shown in Figure H.1. The derived stabilities and purities are shown in Figures H.2 and H.3. A definition for both of these quantities is given in Chapter 11. The migrations are found to be small.

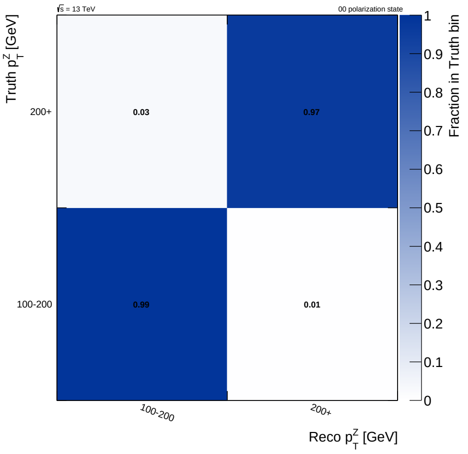
The setup for the simultaneous fits is as follows.

- Each region is treated completely independently with independent $N_{tot}$ and fraction parameters.
- Each region uses the same templates as for the individual fits.
- The nuisance parameters are fully correlated. This means that a nuisance parameter always has the same pulls/values in both regions.
- The background-only hypothesis is defined by setting the $f_{00}$ parameters in both regions to 0 simultaneously.

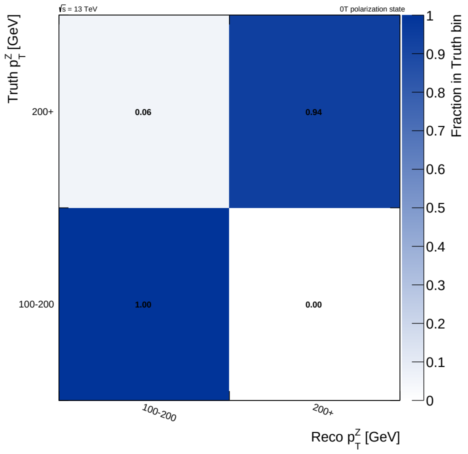
Using this setup, the quoted significance effectively describes the incompatibility of the data with the hypothesis that there are no doubly longitudinal contributions above $p_T^Z = 100\,\text{GeV}$.

The results of the fits using the two- and three-parameter configurations are shown in Tables H.2 and H.2, respectively. These results can be compared to those from the individual fits as shown in Tables 10.12 and 10.13 and those of the direct fit in the region with $p_T^Z > 100\,\text{GeV}$. The comparison to the individual fits shows increased significances. Additionally, small differences in the fit results to data can be observed for the values describing the region with 200 GeV. Here, $f_{00}$ and $f_{TT}$ are increased at the expense of $f_{Mixed}$. Meanwhile, the values and uncertainties in the $100 < p_T^Z \leq 200\,\text{GeV}$ remain the same within the significant digits. Only the propagated values for the $\mu$'s differ. Both of these effects can be understood through correlations of nuisance parameters.

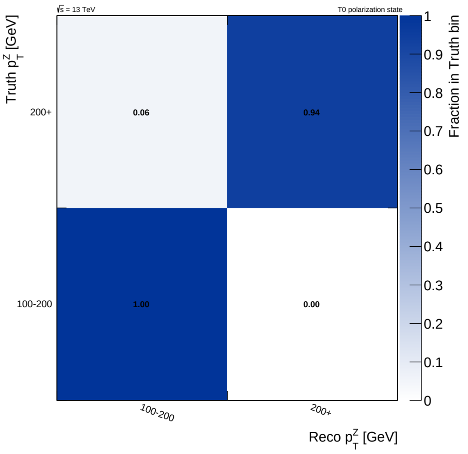
The best example is the `PRW_DATASF` nuisance parameter. Comparing Figures 10.9c and 10.10c, it can be seen that it is pulled up for the region with $100 < p_T^Z < 200\,\text{GeV}$ and pulled down for the $p_T^Z > 200\,\text{GeV}$ region. In the combined fit, this is not possible, leading to overall smaller pulls and, thus, different final results. This is demonstrated in Figure H.4a where this NP is now pulled up, but less than originally in the $100 < p_T^Z < 200\,\text{GeV}$ region.



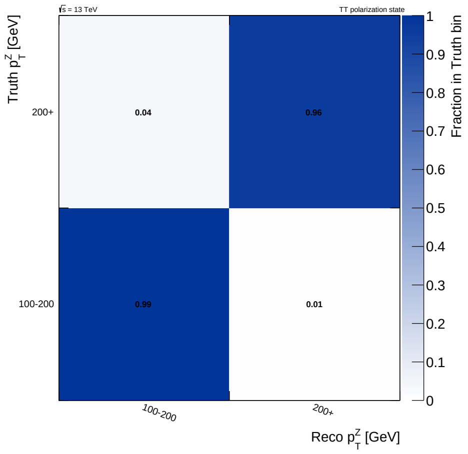

**(a)**




**(b)**




**(c)**




**(d)**

**Figure H.1:** Migration matrices for the (a) 00, (b) 0T, (c) T0, and (d) TT polarization states between the two 00-enriched regions.

Another interesting phenomenon can be seen in the observed correlations between the parameters for the individual regions, as shown in Figure H.4b. Namely, the fact that the separate $N_{tot}$ and fraction parameters for each region have correlations $\neq 0$. This is the most prominent for $N_{tot_{100}}$ and $N_{tot_{200}}$ with a correlation of 0.308. This is mostly due to nuisance parameters having similar impacts on the total event number in both regions.

These fits are not used as the nominal ones as a combined measurement should ideally take all, even small, migrations into account. Additionally, more studies and discussions would be needed to decide which nuisance parameters should be correlated across the regions and which should remain uncorrelated.

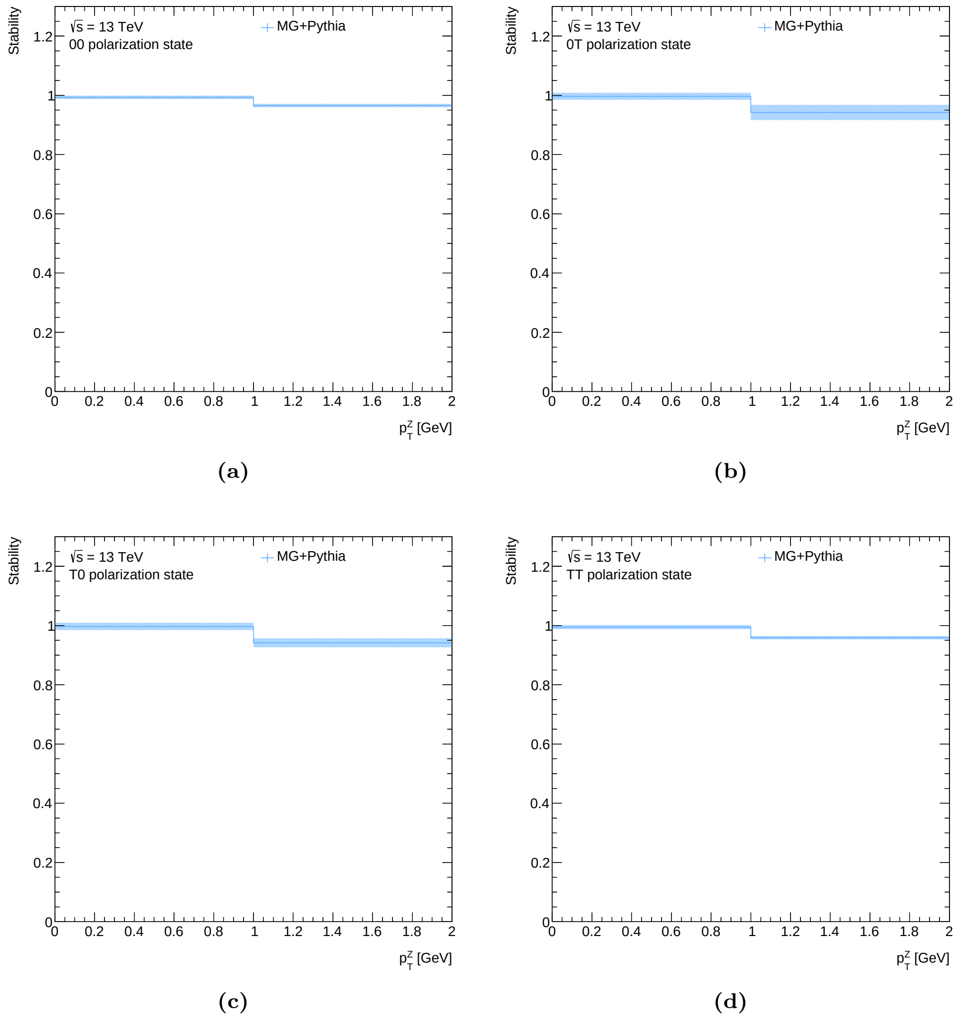
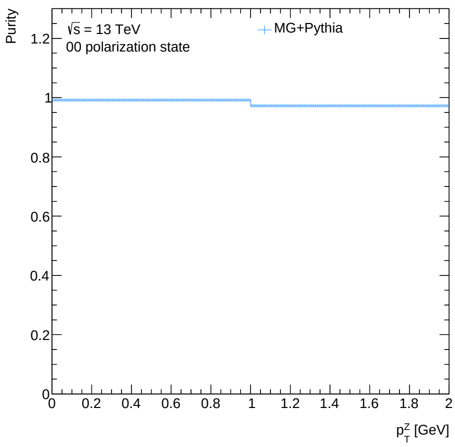


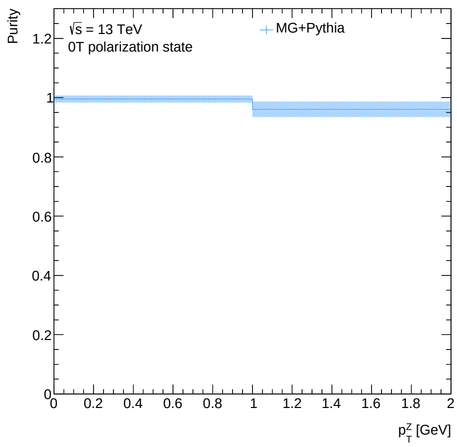

**Figure H.2:** Stabilities for the (a) 00, (b) 0T, (c) T0, and (d) TT polarization states between the two 00-enriched regions.

**Table H.1:** Fit results using the smoothed nuisance parameter distributions in the configuration with two independent fit parameters for the fiducial fraction parametrization in both 00-enriched regions simultaneously.

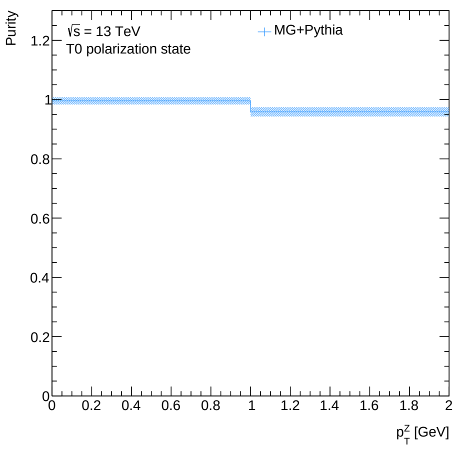
| | Exp. | Obs. |
|---|---|---|
| Significance | $7.495\,\sigma$ | $7.994\,\sigma$ |
| $N_{tot_{200}}$ | $330\pm^{29}_{28}$ | $356\pm^{31}_{30}$ |
| $f_{00_{200}}$ | $0.23\pm^{0.06}_{0.06}$ | $0.16\pm^{0.06}_{0.06}$ |
| $N_{tot_{100}}$ | $2500\pm^{120}_{120}$ | $2680\pm^{140}_{140}$ |
| $f_{00_{100}}$ | $0.152\pm^{0.023}_{0.023}$ | $0.172\pm^{0.024}_{0.024}$ |
| $f_{Rest_{200}}$ | $0.77 \pm 0.06$ | $0.84 \pm 0.06$ |
| $f_{Rest_{100}}$ | $0.848 \pm 0.023$ | $0.828 \pm 0.024$ |
| $\mu_{200}$ | $1.00 \pm 0.27$ | $0.76 \pm 0.27$ |
| $\mu_{WZRest_{200}}$ | $1.00 \pm 0.11$ | $1.18 \pm 0.13$ |
| $\mu_{100}$ | $1.00 \pm 0.16$ | $1.22 \pm 0.18$ |
| $\mu_{WZRest_{100}}$ | $1.00 \pm 0.06$ | $1.05 \pm 0.06$ |



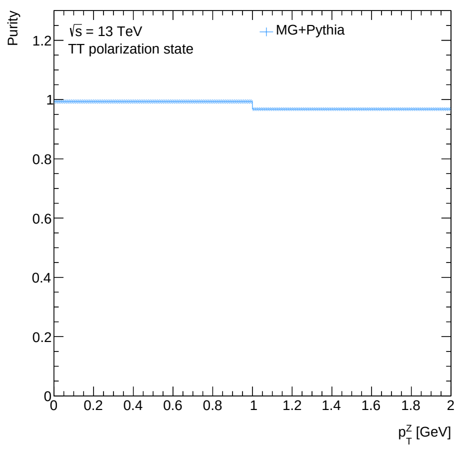

**(a)**




**(b)**




**(c)**

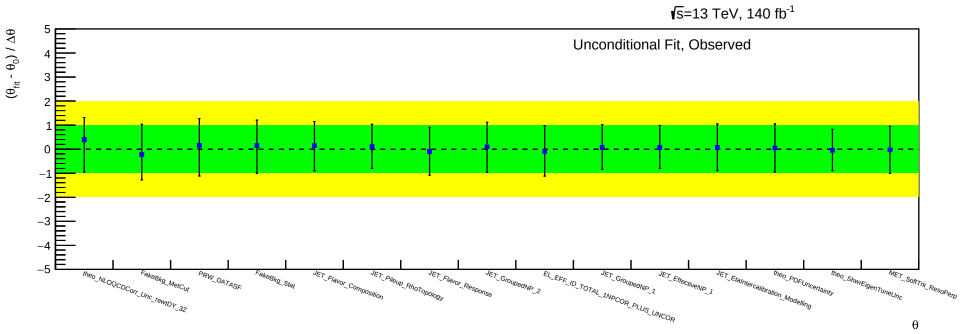



**(d)**

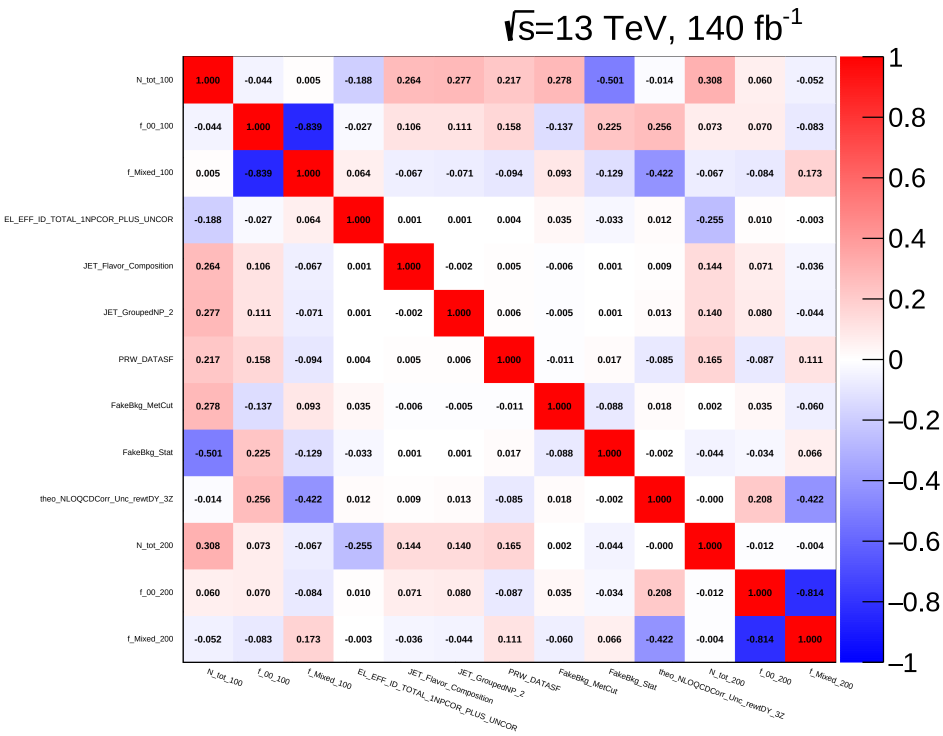
**Figure H.3:** Purities for the (a) 00, (b) 0T, (c) T0, and (d) TT polarization states between the two 00-enriched regions.

**Table H.2:** Fit results using the smoothed nuisance parameter distributions in the configuration with three independent fit parameters for the fiducial fraction parametrization in both 00-enriched regions simultaneously.

| | Exp. | Obs. |
|---|---|---|
| Significance | $4.823\,\sigma$ | $5.378\,\sigma$ |
| $f_{00_{200}}$ | $0.23\pm^{0.10}_{0.10}$ | $0.14\pm^{0.09}_{0.09}$ |
| $f_{Mixed_{200}}$ | $0.12\pm^{0.20}_{0.12}$ | $0.21\pm^{0.19}_{0.21}$ |
| $N_{tot_{200}}$ | $330\pm^{29}_{28}$ | $358\pm^{31}_{30}$ |
| $f_{00_{100}}$ | $0.15\pm^{0.04}_{0.04}$ | $0.19\pm^{0.04}_{0.04}$ |
| $f_{Mixed_{100}}$ | $0.23\pm^{0.08}_{0.09}$ | $0.18\pm^{0.09}_{0.10}$ |
| $N_{tot_{100}}$ | $2500\pm^{120}_{120}$ | $2690\pm^{140}_{140}$ |
| $f_{TT_{200}}$ | $0.65 \pm 0.09$ | $0.66 \pm 0.14$ |
| $f_{TT_{100}}$ | $0.62 \pm 0.06$ | $0.63 \pm 0.06$ |
| $\mu_{200}$ | $1.0 \pm 0.4$ | $0.6 \pm 0.4$ |
| $\mu_{WZTT_{200}}$ | $1.00 \pm 0.17$ | $1.10 \pm 0.25$ |
| $\mu_{WZMixed_{200}}$ | $1.0 \pm 1.3$ | $1.9 \pm 1.8$ |
| $\mu_{100}$ | $1.00 \pm 0.25$ | $1.34 \pm 0.29$ |
| $\mu_{WZTT_{100}}$ | $1.00 \pm 0.11$ | $1.10 \pm 0.12$ |
| $\mu_{WZMixed_{100}}$ | $1.0 \pm 0.4$ | $0.8 \pm 0.4$ |




(a)




(b)

**Figure H.4:** Observed post fit (a) nuisance parameter pulls for the most pulled parameters and (b) correlations of nuisance parameters with a correlation of at least 0.2 with another NP for the fit using the three-parameter configuration done in both 00-enriched regions simultaneously.

# Appendix I

# Breakdown of Uncertainties in Unfoldings

This appendix contains the bin values and uncertainty breakdowns for the remaining RAZ regions that were not shown in Chapter 11.

## I.1 Unfolding using $\Delta Y(\ell_W Z)$

The results and breakdowns for the $\Delta Y(\ell_W Z)$ unfolding are given in Tables I.1, I.2 and I.3 for the TT unfolding and Tables I.4, I.5 and I.6 for the sum of polarizations.

**Table I.1:** Results and uncertainties for the unfolding of the TT polarization state using $\Delta Y(\ell_W Z)$ in the region with $p_T^{WZ} < 20\,\text{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.137 | 0.408 | 0.198 | 0.126 | 0.132 |
| Comb. Unc. | 12.4% | 2.61% | 4.61% | 6.03% | 6.86% |
| Data statistical | 5.25% | 2.10% | 3.33% | 4.02% | 3.66% |
| Experimental and modelling | 11.2% | 1.55% | 3.19% | 4.49% | 5.80% |
| Luminosity | 0.65% | 0.06% | 0.19% | 0.25% | 0.33% |
| Electron calibration | 0.76% | 0.11% | 0.18% | 0.33% | 0.55% |
| Muon calibration | 0.93% | 0.11% | 0.30% | 0.38% | 0.32% |
| Jet energy scale and resolution | 4.41% | 0.44% | 1.41% | 1.22% | 2.64% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 0.19% | 0.05% | 0.14% | 0.09% | 0.06% |
| Flavor-tagging inefficiency | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| Pileup modelling | 1.05% | 0.41% | 0.37% | 1.05% | 0.81% |
| Non-prompt background estimation | 1.96% | 0.08% | 0.61% | 0.45% | 0.42% |
| Background, other | 2.03% | 0.10% | 0.60% | 0.71% | 0.82% |
| Model statistical | 0.92% | 0.36% | 0.57% | 0.74% | 0.74% |
| NLO corrections | 6.70% | 0.64% | 1.43% | 2.70% | 3.12% |
| PDF, Scale and shower settings | 7.13% | 1.19% | 2.17% | 2.94% | 3.79% |

## I.2 Unfolding using $\Delta Y(WZ)$

The results and breakdowns for the $\Delta Y(WZ)$ unfolding are given in Tables I.7, I.8 and I.9 for the TT unfolding and Tables I.10, I.11 and I.12 for the sum of polarizations.

**Table I.2:** Results and uncertainties for the unfolding of the TT polarization state using $\Delta Y(\ell_W Z)$ in the region with $p_T^{WZ} < 40\,\text{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.141 | 0.403 | 0.193 | 0.128 | 0.135 |
| Comb. Unc. | 9.38% | 2.40% | 3.88% | 4.95% | 5.28% |
| Data statistical | 4.58% | 1.91% | 2.91% | 3.58% | 3.37% |
| Experimental and modelling | 8.18% | 1.45% | 2.56% | 3.42% | 4.07% |
| Luminosity | 0.65% | 0.07% | 0.20% | 0.27% | 0.36% |
| Electron calibration | 0.57% | 0.11% | 0.10% | 0.29% | 0.51% |
| Muon calibration | 0.80% | 0.09% | 0.28% | 0.36% | 0.31% |
| Jet energy scale and resolution | 2.63% | 0.47% | 0.98% | 1.31% | 1.51% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 0.36% | 0.08% | 0.05% | 0.10% | 0.17% |
| Flavor-tagging inefficiency | 0.01% | 0.00% | 0.00% | 0.00% | 0.00% |
| Pileup modelling | 0.39% | 0.24% | 0.14% | 0.54% | 0.43% |
| Non-prompt background estimation | 3.13% | 0.03% | 1.11% | 0.80% | 0.99% |
| Background, other | 2.26% | 0.14% | 0.69% | 0.83% | 0.99% |
| Model statistical | 0.82% | 0.32% | 0.51% | 0.63% | 0.59% |
| NLO corrections | 4.64% | 0.55% | 0.77% | 1.75% | 1.98% |
| PDF, Scale and shower settings | 4.59% | 1.16% | 1.68% | 2.12% | 2.71% |

**Table I.3:** Results and uncertainties for the unfolding of the TT polarization state using $\Delta Y(\ell_W Z)$ in the region with $p_T^{WZ} < 70\,\text{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.144 | 0.404 | 0.194 | 0.123 | 0.134 |
| Comb. Unc. | 8.40% | 2.21% | 3.71% | 4.49% | 4.91% |
| Data statistical | 4.12% | 1.75% | 2.73% | 3.35% | 3.09% |
| Experimental and modelling | 7.31% | 1.34% | 2.50% | 2.98% | 3.81% |
| Luminosity | 0.67% | 0.08% | 0.22% | 0.28% | 0.38% |
| Electron calibration | 0.43% | 0.09% | 0.08% | 0.24% | 0.40% |
| Muon calibration | 0.67% | 0.05% | 0.25% | 0.26% | 0.26% |
| Jet energy scale and resolution | 1.61% | 0.26% | 0.69% | 0.81% | 0.75% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 0.50% | 0.01% | 0.20% | 0.05% | 0.18% |
| Flavor-tagging inefficiency | 0.01% | 0.00% | 0.00% | 0.01% | 0.01% |
| Pileup modelling | 0.22% | 0.22% | 0.21% | 0.35% | 0.29% |
| Non-prompt background estimation | 4.00% | 0.16% | 1.40% | 1.17% | 1.71% |
| Background, other | 2.46% | 0.16% | 0.79% | 0.94% | 1.12% |
| Model statistical | 0.77% | 0.30% | 0.48% | 0.59% | 0.57% |
| NLO corrections | 3.55% | 0.46% | 0.60% | 1.32% | 1.67% |
| PDF, Scale and shower settings | 3.76% | 1.14% | 1.54% | 1.88% | 2.48% |

**Table I.4:** Results and uncertainties for the unfolding of the sum of polarization states using $\Delta Y(\ell_W Z)$ in the region with $p_T^{WZ} < 20\,\text{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.188 | 0.428 | 0.174 | 0.105 | 0.103 |
| Comb. Unc. | 2.99% | 1.75% | 2.93% | 3.66% | 3.64% |
| Data statistical | 2.54% | 1.39% | 2.65% | 3.38% | 3.33% |
| Experimental and modelling | 1.56% | 1.07% | 1.25% | 1.39% | 1.46% |
| Luminosity | 0.04% | 0.00% | 0.02% | 0.02% | 0.02% |
| Electron calibration | 0.08% | 0.01% | 0.06% | 0.05% | 0.08% |
| Muon calibration | 0.09% | 0.10% | 0.16% | 0.10% | 0.27% |
| Jet energy scale and resolution | 0.27% | 0.07% | 0.20% | 0.34% | 0.52% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 0.05% | 0.02% | 0.07% | 0.07% | 0.05% |
| Flavor-tagging inefficiency | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| Pileup modelling | 0.22% | 0.16% | 0.06% | 0.39% | 0.00% |
| Non-prompt background estimation | 0.71% | 0.08% | 0.36% | 0.21% | 0.13% |
| Background, other | 0.72% | 0.03% | 0.34% | 0.41% | 0.47% |
| Model statistical | 0.43% | 0.23% | 0.45% | 0.61% | 0.65% |
| PDF, Scale and shower settings | 1.03% | 1.01% | 1.01% | 1.02% | 1.04% |

**Table I.5:** Results and uncertainties for the unfolding of the sum of polarization states using $\Delta Y(\ell_W Z)$ in the region with $p_T^{WZ} < 40\,\text{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.191 | 0.425 | 0.171 | 0.106 | 0.104 |
| Comb. Unc. | 2.85% | 1.63% | 2.66% | 3.27% | 3.36% |
| Data statistical | 2.21% | 1.24% | 2.28% | 2.97% | 3.03% |
| Experimental and modelling | 1.79% | 1.05% | 1.36% | 1.35% | 1.44% |
| Luminosity | 0.04% | 0.00% | 0.02% | 0.02% | 0.03% |
| Electron calibration | 0.12% | 0.01% | 0.12% | 0.04% | 0.07% |
| Muon calibration | 0.08% | 0.08% | 0.13% | 0.14% | 0.26% |
| Jet energy scale and resolution | 0.21% | 0.13% | 0.23% | 0.34% | 0.21% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 0.13% | 0.06% | 0.01% | 0.11% | 0.11% |
| Flavor-tagging inefficiency | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| Pileup modelling | 0.29% | 0.07% | 0.14% | 0.12% | 0.11% |
| Non-prompt background estimation | 1.10% | 0.04% | 0.65% | 0.35% | 0.43% |
| Background, other | 0.79% | 0.05% | 0.38% | 0.46% | 0.56% |
| Model statistical | 0.37% | 0.20% | 0.38% | 0.52% | 0.55% |
| PDF, Scale and shower settings | 1.02% | 1.01% | 1.01% | 1.02% | 1.04% |

**Table I.6:** Results and uncertainties for the unfolding of the sum of polarization states using $\Delta Y(\ell_W Z)$ in the region with $p_T^{WZ} < 70\,\text{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.195 | 0.427 | 0.171 | 0.104 | 0.103 |
| Comb. Unc. | 2.79% | 1.53% | 2.54% | 3.04% | 3.16% |
| Data statistical | 1.96% | 1.11% | 2.10% | 2.72% | 2.73% |
| Experimental and modelling | 1.98% | 1.04% | 1.42% | 1.37% | 1.60% |
| Luminosity | 0.05% | 0.00% | 0.02% | 0.03% | 0.03% |
| Electron calibration | 0.16% | 0.01% | 0.12% | 0.07% | 0.05% |
| Muon calibration | 0.12% | 0.04% | 0.07% | 0.08% | 0.20% |
| Jet energy scale and resolution | 0.23% | 0.09% | 0.26% | 0.27% | 0.11% |
| $E_{\text{T}}^{\text{miss}}$ scale and resolution | 0.20% | 0.01% | 0.13% | 0.01% | 0.12% |
| Flavor-tagging inefficiency | 0.00% | 0.00% | 0.00% | 0.01% | 0.00% |
| Pileup modelling | 0.26% | 0.07% | 0.04% | 0.02% | 0.13% |
| Non-prompt background estimation | 1.35% | 0.02% | 0.76% | 0.51% | 0.85% |
| Background, other | 0.84% | 0.05% | 0.42% | 0.50% | 0.60% |
| Model statistical | 0.34% | 0.18% | 0.35% | 0.46% | 0.49% |
| PDF, Scale and shower settings | 1.02% | 1.01% | 1.01% | 1.02% | 1.05% |

**Table I.7:** Results and uncertainties for the unfolding of the TT polarization state using $\Delta Y(WZ)$ in the region with $p_T^{WZ} < 20\,\text{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.067 | 0.429 | 0.207 | 0.148 | 0.145 |
| Comb. Unc. | 38.4% | 5.18% | 10.7% | 11.6% | 8.45% |
| Data statistical | 22.9% | 4.87% | 9.16% | 10.2% | 6.19% |
| Experimental and modelling | 30.8% | 1.78% | 5.57% | 5.50% | 5.75% |
| Luminosity | 1.93% | 0.02% | 0.22% | 0.27% | 0.36% |
| Electron calibration | 2.62% | 0.16% | 0.46% | 0.40% | 0.70% |
| Muon calibration | 2.66% | 0.67% | 2.93% | 2.55% | 0.72% |
| Jet energy scale and resolution | 13.9% | 0.38% | 1.93% | 2.49% | 2.22% |
| $E_{\text{T}}^{\text{miss}}$ scale and resolution | 2.72% | 0.44% | 0.87% | 0.87% | 0.49% |
| Flavor-tagging inefficiency | 0.03% | 0.00% | 0.00% | 0.00% | 0.01% |
| Pileup modelling | 5.20% | 0.12% | 0.62% | 0.21% | 1.38% |
| Non-prompt background estimation | 5.41% | 0.18% | 0.74% | 0.37% | 0.59% |
| Background, other | 5.43% | 0.01% | 0.66% | 0.71% | 0.79% |
| Model statistical | 4.31% | 0.97% | 1.92% | 2.17% | 1.32% |
| NLO corrections | 13.5% | 0.41% | 2.00% | 2.03% | 2.58% |
| PDF, Scale and shower settings | 21.0% | 1.09% | 2.92% | 2.64% | 3.92% |

**Table I.8:** Results and uncertainties for the unfolding of the TT polarization state using $\Delta Y(WZ)$ in the region with $p_T^{WZ} < 40\,\text{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.066 | 0.431 | 0.204 | 0.146 | 0.150 |
| Comb. Unc. | 32.5% | 5.02% | 10.7% | 11.1% | 6.82% |
| Data statistical | 21.3% | 4.66% | 9.51% | 10.0% | 5.31% |
| Experimental and modelling | 24.6% | 1.87% | 4.91% | 4.87% | 4.27% |
| Luminosity | 2.14% | 0.02% | 0.23% | 0.29% | 0.39% |
| Electron calibration | 2.36% | 0.21% | 0.49% | 0.45% | 0.60% |
| Muon calibration | 2.96% | 0.54% | 3.00% | 2.57% | 0.74% |
| Jet energy scale and resolution | 10.0% | 0.24% | 0.94% | 1.86% | 1.69% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 3.91% | 0.81% | 1.35% | 1.23% | 0.37% |
| Flavor-tagging inefficiency | 0.03% | 0.00% | 0.01% | 0.02% | 0.00% |
| Pileup modelling | 2.52% | 0.25% | 0.01% | 0.34% | 0.74% |
| Non-prompt background estimation | 9.54% | 0.27% | 0.91% | 0.97% | 1.24% |
| Background, other | 6.67% | 0.03% | 0.74% | 0.87% | 0.98% |
| Model statistical | 4.20% | 0.94% | 1.96% | 2.17% | 1.17% |
| NLO corrections | 6.73% | 0.48% | 1.42% | 1.21% | 1.27% |
| PDF, Scale and shower settings | 16.2% | 1.07% | 2.20% | 1.96% | 2.85% |

**Table I.9:** Results and uncertainties for the unfolding of the TT polarization state using $\Delta Y(WZ)$ in the region with $p_T^{WZ} < 70\,\text{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.070 | 0.436 | 0.208 | 0.135 | 0.148 |
| Comb. Unc. | 28.0% | 4.66% | 9.69% | 11.3% | 6.45% |
| Data statistical | 18.8% | 4.32% | 8.61% | 10.1% | 5.01% |
| Experimental and modelling | 20.7% | 1.74% | 4.44% | 5.02% | 4.06% |
| Luminosity | 2.09% | 0.02% | 0.25% | 0.30% | 0.42% |
| Electron calibration | 1.85% | 0.16% | 0.40% | 0.38% | 0.44% |
| Muon calibration | 2.47% | 0.48% | 2.88% | 2.90% | 0.71% |
| Jet energy scale and resolution | 5.91% | 0.24% | 0.47% | 1.91% | 0.64% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 3.20% | 0.52% | 0.97% | 1.36% | 0.26% |
| Flavor-tagging inefficiency | 0.05% | 0.00% | 0.01% | 0.02% | 0.01% |
| Pileup modelling | 0.94% | 0.21% | 0.31% | 0.36% | 0.57% |
| Non-prompt background estimation | 11.8% | 0.32% | 0.99% | 1.19% | 2.15% |
| Background, other | 6.98% | 0.03% | 0.87% | 0.95% | 1.12% |
| Model statistical | 3.75% | 0.89% | 1.83% | 2.20% | 1.11% |
| NLO corrections | 3.41% | 0.51% | 0.89% | 0.91% | 0.99% |
| PDF, Scale and shower settings | 12.5% | 1.10% | 1.99% | 1.68% | 2.57% |

**Table I.10:** Results and uncertainties for the unfolding of the sum of polarization states using $\Delta Y(WZ)$ in the region with $p_T^{WZ} < 20\,\text{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.137 | 0.448 | 0.182 | 0.120 | 0.110 |
| Comb. Unc. | 7.61% | 3.49% | 8.11% | 9.81% | 6.22% |
| Data statistical | 7.10% | 3.20% | 7.48% | 9.13% | 5.88% |
| Experimental and modelling | 2.72% | 1.40% | 3.14% | 3.59% | 2.03% |
| Luminosity | 0.06% | 0.00% | 0.02% | 0.02% | 0.02% |
| Electron calibration | 0.12% | 0.09% | 0.26% | 0.12% | 0.21% |
| Muon calibration | 0.43% | 0.45% | 2.27% | 2.36% | 0.45% |
| Jet energy scale and resolution | 0.76% | 0.17% | 0.49% | 0.90% | 0.51% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 0.63% | 0.30% | 0.77% | 0.70% | 0.38% |
| Flavor-tagging inefficiency | 0.01% | 0.00% | 0.00% | 0.00% | 0.00% |
| Pileup modelling | 0.18% | 0.11% | 0.10% | 0.79% | 0.46% |
| Non-prompt background estimation | 1.20% | 0.09% | 0.46% | 0.18% | 0.27% |
| Background, other | 1.16% | 0.01% | 0.37% | 0.40% | 0.43% |
| Model statistical | 1.30% | 0.61% | 1.52% | 1.90% | 1.26% |
| PDF, Scale and shower settings | 1.28% | 1.10% | 1.04% | 1.22% | 1.18% |

**Table I.11:** Results and uncertainties for the unfolding of the sum of polarization states using $\Delta Y(WZ)$ in the region with $p_T^{WZ} < 40\,\text{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.138 | 0.449 | 0.180 | 0.119 | 0.113 |
| Comb. Unc. | 7.17% | 3.37% | 8.35% | 9.62% | 5.39% |
| Data statistical | 6.36% | 3.04% | 7.66% | 8.87% | 5.03% |
| Experimental and modelling | 3.30% | 1.45% | 3.32% | 3.72% | 1.93% |
| Luminosity | 0.08% | 0.00% | 0.02% | 0.02% | 0.03% |
| Electron calibration | 0.22% | 0.12% | 0.27% | 0.20% | 0.16% |
| Muon calibration | 0.38% | 0.38% | 2.36% | 2.42% | 0.52% |
| Jet energy scale and resolution | 0.80% | 0.15% | 0.31% | 0.86% | 0.31% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 1.11% | 0.53% | 1.15% | 1.11% | 0.30% |
| Flavor-tagging inefficiency | 0.01% | 0.00% | 0.01% | 0.02% | 0.00% |
| Pileup modelling | 0.26% | 0.19% | 0.29% | 0.70% | 0.11% |
| Non-prompt background estimation | 1.97% | 0.14% | 0.47% | 0.45% | 0.63% |
| Background, other | 1.35% | 0.01% | 0.39% | 0.47% | 0.53% |
| Model statistical | 1.23% | 0.61% | 1.55% | 1.90% | 1.08% |
| PDF, Scale and shower settings | 1.21% | 1.09% | 1.04% | 1.17% | 1.18% |

**Table I.12:** Results and uncertainties for the unfolding of the sum of polarization states using $\Delta Y(WZ)$ in the region with $p_T^{WZ} < 70\,\text{GeV}$.

| Bin | 0.0 - 0.5 | 0.5 - 1.5 | 1.5 - 2.0 | 2.0 - 2.5 | 2.5 - 6.0 |
|---|---|---|---|---|---|
| Results | 0.143 | 0.452 | 0.182 | 0.110 | 0.111 |
| Comb. Unc. | 6.78% | 3.12% | 7.54% | 9.64% | 5.13% |
| Data statistical | 5.80% | 2.82% | 6.87% | 8.78% | 4.65% |
| Experimental and modelling | 3.51% | 1.35% | 3.10% | 3.99% | 2.16% |
| Luminosity | 0.09% | 0.00% | 0.03% | 0.03% | 0.03% |
| Electron calibration | 0.24% | 0.08% | 0.11% | 0.09% | 0.11% |
| Muon calibration | 0.25% | 0.33% | 2.22% | 2.60% | 0.45% |
| Jet energy scale and resolution | 0.78% | 0.16% | 0.52% | 1.28% | 0.23% |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ scale and resolution | 0.91% | 0.35% | 0.85% | 1.21% | 0.19% |
| Flavor-tagging inefficiency | 0.00% | 0.00% | 0.01% | 0.01% | 0.00% |
| Pileup modelling | 0.42% | 0.18% | 0.45% | 0.62% | 0.06% |
| Non-prompt background estimation | 2.44% | 0.20% | 0.42% | 0.48% | 1.23% |
| Background, other | 1.43% | 0.02% | 0.45% | 0.50% | 0.59% |
| Model statistical | 1.11% | 0.56% | 1.43% | 1.90% | 1.03% |
| PDF, Scale and shower settings | 1.13% | 1.07% | 1.03% | 1.22% | 1.19% |

# List of Figures

# List of Tables

# Bibliography

[1] Isaac Newton, Daniel Bernoulli, Colin MacLaurin, and Leonhard Euler, *Philosophiae naturalis principia mathematica*, vol. 1, excudit G. Brookman; impensis TT et J. Tegg, Londini, 1833 (cit. on p. 1).

[2] Marie Curie, *Radioactive substances : a translation from the French of the classical thesis presented to the Faculty of Sciences in Paris / by the distinguished Nobel Prize winner, Marie Curie*, English, Philosophical Library New York, 1961 94 p. : (cit. on p. 1).

[3] S. W. Hawking, *Black hole explosions*, Nature **248** (1974) 30 (cit. on p. 1).

[4] Albert Einstein, *The Field Equations of Gravitation*, Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys. ) **1915** (1915) 844 (cit. on p. 1).

[5] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: `1207.7214 [hep-ex]` (cit. on p. 1).

[6] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: `1207.7235 [hep-ex]` (cit. on p. 1).

[7] R. P. Woodard, *How Far Are We from the Quantum Theory of Gravity?*, Rept. Prog. Phys. **72** (2009) 126002, arXiv: `0907.4238 [gr-qc]` (cit. on p. 1).

[8] John F. Donoghue, *The effective field theory treatment of quantum gravity*, AIP Conf. Proc. **1483** (2012) 73, ed. by Waldyr Alves Rodrigues, Richard Kerner, Gentil Pires, and Carols Pinheiro, arXiv: `1209.3511 [gr-qc]` (cit. on p. 1).

[9] Steven Carlip, Dah-Wei Chiou, Wei-Tou Ni, and Richard Woodard, *Quantum Gravity: A Brief History of Ideas and Some Prospects*, Int. J. Mod. Phys. D **24** (2015) 1530028, arXiv: `1507.08194 [gr-qc]` (cit. on p. 1).

[10] Katherine Garrett and Gintaras Duda, *Dark Matter: A Primer*, Adv. Astron. **2011** (2011) 968283, arXiv: `1006.2483 [hep-ph]` (cit. on p. 1).

[11] Edmund J. Copeland, M. Sami, and Shinji Tsujikawa, *Dynamics of dark energy*, Int. J. Mod. Phys. D **15** (2006) 1753, arXiv: `hep-th/0603057` (cit. on p. 1).

[12] M. C. Gonzalez-Garcia and Yosef Nir, *Neutrino Masses and Mixing: Evidence and Implications*, Rev. Mod. Phys. **75** (2003) 345, arXiv: `hep-ph/0202058` (cit. on p. 1).

[13] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32 (cit. on p. 1).

[14] Werner Bernreuther, *CP violation and baryogenesis*, Lect. Notes Phys. **591** (2002) 237, arXiv: `hep-ph/0205279` (cit. on p. 1).

[15] CDF Collaboration, *High-precision measurement of the $W$ boson mass with the CDF II detector*, Science **376** (2022) 170 (cit. on pp. 1, 7).

[16] ATLAS Collaboration, *Improved W boson Mass Measurement using $\sqrt{s} = 7$ TeV Proton–Proton Collisions with the ATLAS Detector*, ATLAS-CONF-2023-004, 2023, URL: `https://cds.cern.ch/record/2853290` (cit. on pp. 1, 7).

[17] Nima Arkani-Hamed, Savas Dimopoulos, and G. R. Dvali, *The Hierarchy problem and new dimensions at a millimeter*, Phys. Lett. B **429** (1998) 263, arXiv: `hep-ph/9803315` (cit. on p. 1).

[18] David J. E. Callaway, *Triviality Pursuit: Can Elementary Scalar Particles Exist?*, Phys. Rept. **167** (1988) 241 (cit. on p. 1).

[19] CMS Collaboration, *Measurements of angular distance and momentum ratio distributions in three-jet and Z + two-jet final states in pp collisions*, Eur. Phys. J. C **81** (2021) 852, arXiv: `2102.08816 [hep-ex]` (cit. on p. 1).

[20] ATLAS Collaboration, *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** (2019) 161801, arXiv: `1906.03203 [hep-ex]` (cit. on pp. 1, 90).

[21] ATLAS Collaboration, *Observation of electroweak $W^{\pm}Z$ boson pair production in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **793** (2019) 469, arXiv: `1812.09740 [hep-ex]` (cit. on p. 1).

[22] CMS Collaboration, *Measurements of production cross sections of polarized same-sign W boson pairs in association with two jets in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **812** (2021) 136018, arXiv: `2009.09429 [hep-ex]` (cit. on pp. 1, 181).

[23] ATLAS and CMS Collaborations, *Combination of the W boson polarization measurements in top quark decays using ATLAS and CMS data at $\sqrt{s} = 8$ TeV*, JHEP **08** (2020) 051, arXiv: `2005.03799 [hep-ex]` (cit. on pp. 2, 181).

[24] ATLAS Collaboration, *Measurement of $W^{\pm}Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **79** (2019) 535, arXiv: `1902.05759 [hep-ex]` (cit. on pp. 2, 36, 40, 41, 43, 98, 149, 181).

[25] CMS Collaboration, *Measurement of the inclusive and differential WZ production cross sections, polarization angles, and triple gauge couplings in pp collisions at $\sqrt{s} = 13$ TeV*, JHEP **07** (2022) 032, arXiv: `2110.11231 [hep-ex]` (cit. on pp. 2, 181).

[26] ATLAS Collaboration, *Observation of gauge boson joint-polarisation states in $W^{\pm}Z$ production from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **843** (2023) 137895, arXiv: `2211.09435 [hep-ex]` (cit. on pp. 2, 36, 40, 42, 53, 63, 98, 99, 142, 144, 181).

[27] U. Baur, T. Han, and J. Ohnemus, *Amplitude zeros in $W^{\pm}Z$ production*, Phys. Rev. Lett. **72** (1994) 3941, arXiv: `hep-ph/9403248` (cit. on pp. 2, 14–16, 189, 190).

[28] Christopher Frye, Marat Freytsis, Jakub Scholtz, and Matthew J. Strassler, *Precision Diboson Observables for the LHC*, JHEP **03** (2016) 171, arXiv: `1510.08451 [hep-ph]` (cit. on p. 2).

[29] W. Pauli, *Relativistic Field Theories of Elementary Particles*, Rev. Mod. Phys. **13** (1941) 203 (cit. on p. 3).

[30] Chen-Ning Yang and Robert L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Phys. Rev. **96** (1954) 191, ed. by Jong-Ping Hsu and D. Fine (cit. on p. 3).

[31] A. Pais and S. B. Treiman, *How Many Charm Quantum Numbers Are There?*, Phys. Rev. Lett. **35** (1975) 1556 (cit. on p. 3).

[32] Sheldon L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961) 579, ISSN: 0029-5582, URL: `https://www.sciencedirect.com/science/article/pii/0029558261904692` (cit. on p. 3).

[33] Steven Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264 (cit. on p. 3).

[34] Abdus Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519** (1968) 367 (cit. on p. 3).

[35] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321, ed. by J. C. Taylor (cit. on pp. 3, 7).

[36] Peter W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508, ed. by J. C. Taylor (cit. on pp. 3, 7).

[37] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585, ed. by J. C. Taylor (cit. on pp. 3, 7).

[38] David J. Gross and Frank Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1343, ed. by J. C. Taylor (cit. on p. 3).

[39] H. David Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30** (1973) 1346, ed. by J. C. Taylor (cit. on p. 3).

[40] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Advantages of the color octet gluon picture*, Physics Letters B **47** (1973) 365, ISSN: 0370-2693, URL: `https://www.sciencedirect.com/science/article/pii/0370269373906254` (cit. on p. 3).

[41] Palash B. Pal, *Dirac, Majorana and Weyl fermions*, Am. J. Phys. **79** (2011) 485, arXiv: `1006.1718 [hep-ph]` (cit. on p. 3).

[42] Ettore Majorana, *Teoria simmetrica dellelettrone e del positrone*, Nuovo Cim. **14** (1937) 171 (cit. on p. 3).

[43] Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv: `hep-ex/9807003` (cit. on p. 4).

[44] Particle Data Group, *Review of Particle Physics*, PTEP **2022** (2022) 083C01 (cit. on pp. 4, 5, 7, 43, 63).

[45] Gordon Kane, *Modern Elementary Particle Physics: Explaining and Extending the Standard Model*, 2nd ed., Cambridge University Press, 2017 (cit. on pp. 3, 8).

[46] David Griffiths, *Introduction to elementary particles*, 2008, ISBN: 978-3-527-40601-2 (cit. on p. 3).

[47] Mark Thomson, *Modern Particle Physics*, Cambridge University Press, 2013 (cit. on p. 3).

[48] Michael E. Peskin and Daniel V. Schroeder, *An Introduction to quantum field theory*, Reading, USA: Addison-Wesley, 1995, ISBN: 978-0-201-50397-5 (cit. on pp. 3, 8).

[49] F. Halzen and Alan D. Martin, *QUARKS AND LEPTONS: AN INTRODUCTORY COURSE IN MODERN PARTICLE PHYSICS*, 1984, ISBN: 978-0-471-88741-6 (cit. on p. 3).

[50] E. Noether, *Invariante Variationsprobleme*, ger, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1918** (1918) 235, URL: `http://eudml.org/doc/59024` (cit. on p. 4).

[51] Sidney Coleman, *Notes from Sidney Coleman's Physics 253a*, 2013, arXiv: `1110.5013 [physics.ed-ph]` (cit. on p. 4).

[52] Gargamelle Neutrino Collaboration, *Observation of Neutrino Like Interactions without Muon or Electron in the Gargamelle Neutrino Experiment*, Nucl. Phys. B **73** (1974) 1 (cit. on p. 5).

[53] UA1 Collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at* $\sqrt{s} = 540$ *GeV*, Phys. Lett. B **122** (1983) 103 (cit. on pp. 5, 7).

[54] UA1 Collaboration, *Experimental observation of lepton pairs of invariant mass around 95 GeV/*$c^2$ *at the CERN SPS collider*, Phys. Lett. B **126** (1983) 398 (cit. on p. 5).

[55] UA2 Collaboration, *Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN* $\overline{p}p$ *collider*, Phys. Lett. B **122** (1983) 476 (cit. on p. 5).

[56] UA2 Collaboration, *Evidence for* $Z^0 \to e^+e^-$ *at the CERN* $\overline{p}p$ *collider*, Phys. Lett. B **129** (1983) 130 (cit. on p. 5).

[57] 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** (1957) 1413 (cit. on p. 5).

[58] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, *Detection of the Free Neutrino: a Confirmation*, Science **124** (1956) 103, eprint: `https://www.science.org/doi/pdf/10.1126/science.124.3212.103`, URL: `https://www.science.org/doi/abs/10.1126/science.124.3212.103` (cit. on p. 6).

[59] E. W. Koopmann, H. Van Krugten, B. Van Nooijen, and A. H. Wapstra, *Conservation of parity in electromagnetic interactions*, Nucl. Phys. B **19** (1970) 107 (cit. on p. 6).

[60] L. Wolfenstein, *TESTS FOR PARITY CONSERVATION IN THE ELECTROMAGNETIC INTERACTION OF HADRONS*, Nucl. Phys. B **72** (1974) 111 (cit. on p. 6).

[61] Elliot Leader, *Spin in Particle Physics*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge University Press, 2001 (cit. on pp. 10, 11).

[62] W. N. Polyzou, W. Glöckle, and H. Witaa, *Spin in Relativistic Quantum Theory*, Few-Body Systems **54** (2012) 1667, ISSN: 1432-5411, URL: `http://dx.doi.org/10.1007/s00601-012-0526-8` (cit. on p. 10).

[63] Dirk H. Rischke, *Symmetries in Quantum Mechanics and Particle Physics*, `https://itp.uni-frankfurt.de/~drischke/Script_Symmetries_Fudan2017.pdf`, Accessed: 2023-11-24 (cit. on p. 10).

[64] Luka Selem, *Measurement of gauge boson joint-polarisation states in WZ pair production at the LHC with the ATLAS detector. Mesure des états de co-polarisation des bosons de jauge en production de paire WZ au LHC avec le détecteur ATLAS*, Presented 27 Sep 2022, Savoie U., 2022, URL: `https://cds.cern.ch/record/2841405` (cit. on pp. 10, 98, 99, 144).

[65] H. Haber, *Polarization sum for massless spin-one particles, Part of Lecture Physics 218 (Advanced Quantum Field Theory) winter quarter 2016*, `https://scipp.ucsc.edu/%7Ehaber/ph218/polsum.pdf`, Accessed: 2023-11-01 (cit. on p. 11).

[66] C. Bourrely, E. Leader, and J. Soffer, *Polarization phenomena in hadronic reactions*, Physics Reports **59** (1980) 95, ISSN: 0370-1573, URL: `https://www.sciencedirect.com/science/article/pii/0370157380900174` (cit. on p. 11).

[67] Alessandro Ballestrero, Ezio Maina, and Giovanni Pelliccioli, *W boson polarization in vector boson scattering at the LHC*, JHEP **03** (2018) 170, arXiv: `1710.09339 [hep-ph]` (cit. on pp. 12, 13).

[68] Alessandro Ballestrero, Ezio Maina, and Giovanni Pelliccioli, *Polarized vector boson scattering in the fully leptonic WZ and ZZ channels at the LHC*, JHEP **09** (2019) 087, arXiv: `1907.04722 [hep-ph]` (cit. on pp. 13, 14).

[69] Diana Mareen Hoppe, *Implementation and phenomenology of polarized cross sections in the simulation of vector boson production with the SHERPA event generator*, Presented 12 Apr 2023, TUD Dresden University of Technology, 2023, URL: `https://cds.cern.ch/record/2872971` (cit. on p. 13).

[70] Mareen Hoppe, Marek Schönherr, and Frank Siegert, *Polarised cross sections for vector boson production with SHERPA*, 2023, arXiv: `2310.14803 [hep-ph]` (cit. on pp. 13, 35, 45).

[71] Alessandro Ballestrero, Aissa Belhouari, Giuseppe Bevilacqua, Vladimir Kashkan, and Ezio Maina, *PHANTOM: A Monte Carlo event generator for six parton final states at high energy colliders*, Comput. Phys. Commun. **180** (2009) 401, arXiv: `0801.3359 [hep-ph]` (cit. on p. 13).

[72] Stefano Actis et al., *RECOLA: REcursive Computation of One-Loop Amplitudes*, Comput. Phys. Commun. **214** (2017) 140, arXiv: `1605.01090 [hep-ph]` (cit. on p. 13).

[73] Ansgar Denner and Giovanni Pelliccioli, *NLO QCD predictions for doubly-polarized WZ production at the LHC*, Phys. Lett. B **814** (2021) 136107, arXiv: `2010.07149 [hep-ph]` (cit. on pp. 13, 45).

[74] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: `1405.0301 [hep-ph]` (cit. on pp. 14, 34, 35).

[75] Diogo Buarque Franzosi, Olivier Mattelaer, Richard Ruiz, and Sujay Shil, *Automated predictions from polarized matrix elements*, JHEP **04** (2020) 082, arXiv: `1912.01725 [hep-ph]` (cit. on pp. 14, 34, 35).

[76] *Different syntax for the same diagram leads to different cross-sections: Why?*, `https://answers.launchpad.net/mg5amcnlo/+faq/2778`, Accessed: 2023-11-01 (cit. on p. 14).

[77] *Meaning of various parameters in the run_card.dat*, `https://answers.launchpad.net/mg5amcnlo/+faq/2014`, Accessed: 2023-11-01 (cit. on p. 14).

[78] Mathieu Pellen, Rene Poncelet, and Andrei Popescu, *Polarised W+j production at the LHC: a study at NNLO QCD accuracy*, JHEP **02** (2022) 160, arXiv: `2109.14336 [hep-ph]` (cit. on p. 14).

[79] Da Liu and Lian-Tao Wang, *Prospects for precision measurement of diboson processes in the semileptonic decay channel in future LHC runs*, Phys. Rev. D **99** (2019) 055001, arXiv: `1804.08688 [hep-ph]` (cit. on p. 16).

[80] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on pp. 19, 21, 23–25, 27).

[81] Lyndon Evans and Philip Bryant, *LHC Machine*, JINST **3** (2008) S08001 (cit. on pp. 19, 20).

[82] Oliver Sim Brüning et al., *LHC Design Report*, CERN Yellow Reports: Monographs, Geneva: CERN, 2004, URL: `https://cds.cern.ch/record/782076` (cit. on p. 19).

[83] *LEP design report*, Report, Copies shelved as reports in LEP, PS and SPS libraries, Geneva: CERN, 1984, URL: `https://cds.cern.ch/record/102083` (cit. on p. 19).

[84] John M. Jowett and Michaela Schaumann, “Overview of Heavy Ions in LHC Run 2”, *9th LHC Operations Evian Workshop*, 2019 15 (cit. on p. 19).

[85] CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004 (cit. on pp. 19, 21).

[86] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002 (cit. on pp. 19, 21).

[87] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005 (cit. on pp. 19, 21).

[88] Fabienne Landua, *The CERN accelerator complex layout in 2022. Complexe des accélérateurs du CERN en janvier 2022*, (2022), General Photo, URL: `https://cds.cern.ch/record/2813716` (cit. on p. 20).

[89] Karlheinz Schindl, *The Injector Chain for the LHC; rev. version*, (1999), URL: `https://cds.cern.ch/record/384396` (cit. on p. 19).

[90] *Linear accelerator 2*, (2012), URL: `https://cds.cern.ch/record/1997427` (cit. on p. 19).

[91] *Linear accelerator 2*, `https://home.cern/science/accelerators/linear-accelerator-2`, Accessed: 2023-11-25 (cit. on p. 19).

[92] *The Proton Synchrotron Booster*, `https://home.cern/science/accelerators/proton-synchrotron-booster`, Accessed: 2023-11-25 (cit. on p. 21).

[93] *The Proton Synchrotron*, `https://home.cern/science/accelerators/proton-synchrotron`, Accessed: 2023-11-25 (cit. on p. 21).

[94] *The Super Proton Synchrotron*, `https://home.cern/science/accelerators/super-proton-synchrotron`, Accessed: 2023-11-25 (cit. on p. 21).

[95] H Bartosik et al., *Injectors Beam Performance Evolution during Run 2*, (2019) 51, URL: `https://cds.cern.ch/record/2750276` (cit. on p. 21).

[96] Hannes Bartosik and Giovanni Rumolo, *Performance of the LHC injector chain after the upgrade and potential development*, 2022, arXiv: `2203.09202 [physics.acc-ph]` (cit. on p. 21).

[97] TOTEM Collaboration, *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007 (cit. on p. 21).

[98] *TOTEM*, `https://home.cern/science/experiments/totem`, Accessed: 2023-11-25 (cit. on p. 21).

[99] LHCf Collaboration, *The LHCf experiment at the Large Hadron Collider: status and prospects*, PoS **ICRC2023** (2023) 444, URL: `https://cds.cern.ch/record/2868230` (cit. on p. 21).

[100] *LHCf*, `https://home.cern/science/experiments/lhcf`, Accessed: 2023-11-25 (cit. on p. 21).

[101] MoEDAL-MAPP Collaboration, "MoEDAL-MAPP, an LHC Dedicated Detector Search Facility", *Snowmass 2021*, 2022, arXiv: `2209.03988 [hep-ph]` (cit. on p. 21).

[102] *MoEDAL-MAPP*, `https://home.cern/science/experiments/moedal-mapp`, Accessed: 2023-11-25 (cit. on p. 21).

[103] FASER Collaboration, *FASER: ForwArd Search ExpeRiment at the LHC*, (2019), arXiv: `1901.04468 [hep-ex]` (cit. on p. 22).

[104] *FASER*, `https://home.cern/science/experiments/faser`, Accessed: 2023-11-25 (cit. on p. 22).

[105] SND@LHC Collaboration,
*SND@LHC: The Scattering and Neutrino Detector at the LHC*, (2022),
arXiv: 2210.02784 [hep-ex] (cit. on p. 22).

[106] *SND@LHC*, https://home.cern/science/experiments/sndlhc,
Accessed: 2023-11-25 (cit. on p. 22).

[107] Rende Steerenberg et al.,
*Operation and performance of the CERN Large Hadron Collider during proton Run 2*,
(2019) MOPMP031, URL: https://cds.cern.ch/record/2696126 (cit. on p. 22).

[108] ATLAS Collaboration, *Public ATLAS Luminosity Results for Run-2 of the LHC*,
https:
//twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2,
Accessed: 2023-11-26 (cit. on pp. 23, 34).

[109] *The Collaboration*, https://atlas.cern/Discover/Collaboration,
Accessed: 2023-11-25 (cit. on p. 23).

[110] ATLAS Collaboration,
*Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment*, (2023),
arXiv: 2308.13362 [hep-ex] (cit. on pp. 25, 59).

[111] ATLAS Collaboration,
*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. C **79** (2019) 639, arXiv: 1902.04655 [physics.ins-det]
(cit. on pp. 26, 59).

[112] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*,
Eur. Phys. J. C **81** (2021) 578, arXiv: 2012.00578 [hep-ex] (cit. on pp. 26, 59).

[113] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 689,
arXiv: 2007.02645 [hep-ex] (cit. on p. 26).

[114] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez,
*The anti-$k_t$ jet clustering algorithm*, JHEP **04** (2008) 063,
arXiv: 0802.1189 [hep-ph] (cit. on p. 26).

[115] ATLAS Collaboration,
*Jet reconstruction and performance using particle flow with the ATLAS Detector*,
Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 [hep-ex] (cit. on p. 27).

[116] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581,
arXiv: 1510.03823 [hep-ex] (cit. on p. 27).

[117] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV*,
Eur. Phys. J. C **78** (2018) 903, arXiv: 1802.08168 [hep-ex] (cit. on pp. 27, 60).

[118] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*,
JINST **15** (2020) P10004, arXiv: 2007.12539 [hep-ex] (cit. on pp. 27, 28).

[119] Andy Buckley et al., *General-purpose event generators for LHC physics*,
Phys. Rept. **504** (2011) 145, arXiv: 1101.2599 [hep-ph] (cit. on pp. 29, 30).

[120] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007,
arXiv: 0811.4622 [hep-ph] (cit. on p. 30).

[121] John C. Collins, Davison E. Soper, and George F. Sterman,
*Factorization of Hard Processes in QCD*,
Adv. Ser. Direct. High Energy Phys. **5** (1989) 1, arXiv: hep-ph/0409313
(cit. on p. 30).

[122] Stefan Weinzierl, *Introduction to Monte Carlo methods*, (2000), arXiv: hep-ph/0006269 (cit. on p. 31).

[123] Russel E. Caflisch, *Monte Carlo and quasi-Monte Carlo methods*, Acta Numerica **7** (1998) 1 (cit. on p. 31).

[124] Enrico Bothmann et al., *Efficient phase-space generation for hadron collider event simulation*, (2023), arXiv: 2302.10449 [hep-ph] (cit. on p. 31).

[125] Zoltan Nagy and Davison E. Soper, *What is a parton shower?*, Phys. Rev. D **98** (2018) 014034, arXiv: 1705.08093 [hep-ph] (cit. on p. 31).

[126] Stefan Höche, "Introduction to parton-shower event generators", *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, 2015 235, arXiv: 1411.4085 [hep-ph] (cit. on pp. 31–33).

[127] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, Nucl. Phys. B **238** (1984) 492 (cit. on p. 31).

[128] Jan-Christopher Winter, Frank Krauss, and Gerhard Soff, *A modified cluster-hadronization model*, Eur. Phys. J. C **36** (2004) 381, arXiv: hep-ph/0311085 (cit. on p. 31).

[129] Bo Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, *Parton fragmentation and string dynamics*, Phys. Rept. **97** (1983) 31 (cit. on p. 31).

[130] Bo Andersson, *The Lund Model*, vol. 7, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge University Press, 2023, ISBN: 978-1-00-940129-6, 978-1-00-940125-8, 978-1-00-940128-9, 978-0-521-01734-3, 978-0-521-42094-5, 978-0-511-88149-7 (cit. on p. 31).

[131] X. Artru and G. Mennessier, *String model and multiproduction*, Nucl. Phys. B **70** (1974) 93 (cit. on p. 31).

[132] D. Amati and G. Veneziano, *Preconfinement as a Property of Perturbative QCD*, Phys. Lett. B **83** (1979) 87 (cit. on p. 31).

[133] B. Webber, *Parton shower Monte Carlo event generators*, Scholarpedia **6** (2011) 10662, revision #128236 (cit. on p. 31).

[134] Jan-Eric Nitschke, *Sherpa <-> Pythia8 hadronisation interface*, ATL-COM-PHYS-2021-1033, 2021, URL: https://cds.cern.ch/record/2791755 (cit. on p. 31).

[135] Gurpreet Singh Chahal and Frank Krauss, *Cluster Hadronisation in Sherpa*, SciPost Phys. **13** (2022) 019, arXiv: 2203.11385 [hep-ph] (cit. on p. 31).

[136] Torbjörn Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph] (cit. on p. 31).

[137] *Fragmentation*, https://pythia.org/latest-manual/Fragmentation.html, Accessed: 2023-11-27 (cit. on p. 31).

[138] Enrico Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv: 1905.09127 [hep-ph] (cit. on p. 31).

[139] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Z. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv: hep-ph/0603175 (cit. on p. 31).

[140] Johannes Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, arXiv: 1512.01178 [hep-ph] (cit. on p. 31).

[141] S. Jadach, Z. Was, R. Decker, and Johann H. Kuhn, *The tau decay library TAUOLA: Version 2.4*, Comput. Phys. Commun. **76** (1993) 361 (cit. on p. 31).

[142] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152 (cit. on p. 31).

[143] D. R. Yennie, Steven C. Frautschi, and H. Suura, *The infrared divergence phenomena and high-energy processes*, Annals Phys. **13** (1961) 379 (cit. on p. 32).

[144] CDF Collaboration, *Studying the Underlying Event in Drell-Yan and High Transverse Momentum Jet Production at the Tevatron*, Phys. Rev. D **82** (2010) 034001, arXiv: `1003.3146 [hep-ex]` (cit. on p. 32).

[145] Torbjörn Sjöstrand and Maria van Zijl, *A Multiple Interaction Model for the Event Structure in Hadron Collisions*, Phys. Rev. D **36** (1987) 2019 (cit. on p. 32).

[146] Stefano Frixione and Bryan R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, arXiv: `hep-ph/0204244` (cit. on p. 32).

[147] Paolo Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: `hep-ph/0409146` (cit. on p. 32).

[148] Stefano Frixione, Paolo Nason, and Carlo Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: `0709.2092 [hep-ph]` (cit. on p. 32).

[149] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09** (2012) 049, arXiv: `1111.1220 [hep-ph]` (cit. on pp. 32, 36).

[150] Johan Alwall et al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, Eur. Phys. J. C **53** (2008) 473, arXiv: `0706.2569 [hep-ph]` (cit. on p. 33).

[151] S. Catani, F. Krauss, B. R. Webber, and R. Kuhn, *QCD Matrix Elements + Parton Showers*, JHEP **11** (2001) 063, arXiv: `hep-ph/0109231` (cit. on p. 33).

[152] Nils Lavesson and Leif Lonnblad, *Merging parton showers and matrix elements: Back to basics*, JHEP **04** (2008) 085, arXiv: `0712.2966 [hep-ph]` (cit. on p. 33).

[153] Leif Lönnblad, *Correcting the Colour-Dipole Cascade Model with Fixed Order Matrix Elements*, JHEP **05** (2002) 046, arXiv: `hep-ph/0112284` (cit. on p. 33).

[154] Leif Lönnblad and Stefan Prestel, *Matching tree-level matrix elements with interleaved showers*, JHEP **03** (2012) 019, arXiv: `1109.4829 [hep-ph]` (cit. on p. 33).

[155] *GoodRunListsForAnalysisRun2*, `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/GoodRunListsForAnalysisRun2`, Accessed: 2023-09-23 (cit. on p. 33).

[156] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **83** (2023) 982, arXiv: `2212.09379 [hep-ex]` (cit. on pp. 33, 59).

[157] *LuminosityForPhysics*, `https://twiki.cern.ch/twiki/bin/viewauth/Atlas/LuminosityForPhysics`, Accessed: 2023-09-23 (cit. on p. 33).

[158] G Avoni et al.,
*The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017, URL: https://cds.cern.ch/record/2633501 (cit. on pp. 33, 59).

[159] S. Agostinelli et al., *GEANT4 – a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on pp. 34, 152).

[160] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det] (cit. on pp. 34, 152).

[161] Prachi Atmasiddha, *Observations of the Radiation Amplitude Zero effect and of the Diboson Longitudinal-Longitudinal Interactions in High $p_T$ Region using WśZ Production with the ATLAS Detector*, Presented 09 Jun 2023, U. Michigan, Ann Arbor, 2023, URL: https://cds.cern.ch/record/2878509 (cit. on pp. 34–39, 41, 42, 45–56, 59, 61, 62, 86, 87, 182, 185, 191, 192).

[162] *How to generate polarized particles at LO*, https://answers.launchpad.net/mg5amcnlo/+faq/2243, Accessed: 2023-11-29 (cit. on p. 35).

[163] *Setting me_frame for polarized same when the production includes a decay chain*, https://answers.launchpad.net/mg5amcnlo/+question/690270, Accessed: 2023-11-29 (cit. on p. 35).

[164] S. Frixione, P. Nason, and G. Ridolfi, *Strong corrections to W Z production at hadron colliders*, Nucl. Phys. B **383** (1992) 3 (cit. on p. 35).

[165] NNPDF Collaboration, Richard D. Ball, et al., *Parton distributions for the LHC run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph] (cit. on p. 35).

[166] *WZ Polarized Samples Leptonic Final state After resolving CKKW-L merging issue with pythia*, https://its.cern.ch/jira/browse/ATLMCPROD-9286, Accessed: 2023-11-29 (cit. on p. 35).

[167] *WZ Polarized Samples Leptonic Final state*, https://gitlab.cern.ch/atlas-physics/pmg/mcjoboptions/-/tree/master/507xxx/507019, Accessed: 2023-11-29 (cit. on p. 35).

[168] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert, *QCD matrix elements + parton showers. The NLO case*, JHEP **04** (2013) 027, arXiv: 1207.5030 [hep-ph] (cit. on p. 36).

[169] Thomas Gehrmann, Stefan Hoche, Frank Krauss, Marek Schonherr, and Frank Siegert, *NLO QCD matrix elements + parton showers in $e^+e^-$ —> hadrons*, JHEP **01** (2013) 144, arXiv: 1207.5031 [hep-ph] (cit. on p. 36).

[170] ATLAS Collaboration, *Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC*, (2022), arXiv: 2211.16178 [hep-ex] (cit. on p. 36).

[171] ATLAS Collaboration,
*Measurements of $W^{\pm}Z$ production cross sections in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector and limits on anomalous gauge boson self-couplings*, Phys. Rev. D **93** (2016) 092004, arXiv: 1603.02151 [hep-ex] (cit. on p. 36).

[172] Piotr Golonka and Zbigniew Was, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, Eur. Phys. J. C **45** (2006) 97, arXiv: hep-ph/0506026 (cit. on p. 36).

[173] N. Davidson, T. Przedzinski, and Z. Was, *PHOTOS Interface in C++: Technical and physics documentation*, Comput. Phys. Commun. **199** (2016) 86, arXiv: 1011.0937 [hep-ph] (cit. on p. 36).

[174] *MainCode*, https://gitlab.cern.ch/yafu/MainCode/, Accessed: 2023-12-04 (cit. on p. 37).

[175] *WZPolarization*, https://gitlab.com/Higgs-125/WZPolarization, Accessed: 2023-12-10 (cit. on p. 37).

[176] *ELCore*, https://gitlab.cern.ch/atlas-germany-dresden-vbs-group/ELCore_setup, Accessed: 2023-12-04 (cit. on p. 37).

[177] Christian Bierlich et al., *Robust Independent Validation of Experiment and Theory: Rivet version 3*, SciPost Phys. **8** (2020) 026, arXiv: 1912.05451 [hep-ph] (cit. on p. 37).

[178] ATLAS Analysis Releases, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AnalysisRelease (cit. on p. 37).

[179] Combined Performance Tools for Physics Analysis, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PhysicsAnalysisWorkBookRel20CPRec (cit. on p. 37).

[180] Standard Model Analysis Software Group, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/StandardModelASG (cit. on p. 37).

[181] Tracking CP recomendations, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TrackingCPEOYE2015 (cit. on pp. 38, 39).

[182] Latest MCP recommendation for muon selection, https://twiki.cern.ch/twiki/bin/view/Atlas/MuonSelectionTool (cit. on p. 38).

[183] Electron four momentum correction, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronPhotonFourMomentumCorrection (cit. on pp. 38, 59).

[184] Latest Run 2 Electron Recommendations, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/LatestRecommendationsElectronIDRun2 (cit. on p. 38).

[185] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 [hep-ex] (cit. on p. 39).

[186] Latest egamma tune for electron identification, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EGammaIdentificationRun2 (cit. on p. 39).

[187] Additional $\gamma^*$ and material conversion veto, https://indico.cern.ch/event/862748/contributions/3635039/attachments/1946503/3235131/conversionsIFF2019.11.18.pdf (cit. on p. 39).

[188] Electron charge efficiency correction tool, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EgammaChargeMisIdentificationTool (cit. on p. 39).

[189] JetEtMiss Recommendations for Release 21 MC16 (2015-2018), https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetEtmissRecommendationsR21 (cit. on p. 40).

[190] Pileup jet recommendations, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PileupJetRecommendations (cit. on p. 40).

[191] Usage of Missing ET in analyses: rebuilding and systematics, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/METUtilities (cit. on pp. 40, 60).

[192] Etmiss Recommendations for Full Run2, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EtmissRecommendationsFullRun2 (cit. on p. 40).

[193] D Adams et al., *Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014*, ATL-COM-PHYS-2014-451, 2014, URL: https://cds.cern.ch/record/1700874 (cit. on p. 40).

[194] AssociationUtils - Overlap removal tools, https://gitlab.cern.ch/atlas/athena/-/tree/21.2/PhysicsAnalysis/AnalysisCommon/AssociationUtils (cit. on p. 40).

[195] Li Huang, Samuel D. Lane, Ian M. Lewis, and Zhen Liu, *Electroweak Restoration at the LHC and Beyond: The $Vh$ Channel*, Phys. Rev. D **103** (2021) 053007, arXiv: 2012.00774 [hep-ph] (cit. on p. 41).

[196] Michael S. Chanowitz and Mary K. Gaillard, *The TeV Physics of Strongly Interacting W's and Z's*, Nucl. Phys. B **261** (1985) 379 (cit. on p. 41).

[197] Roberto Franceschini, Giuliano Panico, Alex Pomarol, Francesco Riva, and Andrea Wulzer, *Electroweak Precision Tests in High-Energy Diboson Processes*, JHEP **02** (2018) 111, arXiv: 1712.01310 [hep-ph] (cit. on p. 41).

[198] Da Liu and Lian-Tao Wang, *Prospects for precision measurement of diboson processes in the semileptonic decay channel in future LHC runs*, Phys. Rev. D **99** (2019) 055001, arXiv: 1804.08688 [hep-ph] (cit. on p. 41).

[199] ATLAS Collaboration, *Optimizing the ATLAS Geant4 detector simulation software*, ATL-SOFT-PROC-2023-002, 2023, URL: https://cds.cern.ch/record/2851466 (cit. on p. 42).

[200] Zachary Marshall for the ATLAS Simulation Group, *Validation and performance studies for the ATLAS simulation*, J. Phys. Conf. Ser. **219** (2010) 032016, ed. by Jan Gruntorad and Milos Lokajicek (cit. on p. 42).

[201] ATLAS Collaboration, *Proposal for particle-level object and observable definitions for use in physics measurements at the LHC*, ATL-PHYS-PUB-2015-013, 2015, URL: https://cds.cern.ch/record/2022743 (cit. on p. 43).

[202] Elena Yatsenko et al., *Measurement of the $W^{\pm}Z$ boson pair production cross section in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*, ATL-COM-PHYS-2015-1502, 2015, URL: https://cds.cern.ch/record/2113195 (cit. on p. 43).

[203] Giovanni Pelliccioli and Giulia Zanderighi, *Polarised-boson pairs at the LHC with NLOPS accuracy*, (2023), arXiv: 2311.05220 [hep-ph] (cit. on p. 45).

[204] Thi Nhung Dao and Duc Ninh Le, *Enhancing the doubly-longitudinal polarization in WZ production at the LHC*, Commun. in Phys. **33** (2023) 223, arXiv: 2302.03324 [hep-ph] (cit. on pp. 53, 64–66).

[205] E F Eisenhandler and S Willocq, *ATLAS Style Guide*, ATL-GEN-PUB-2008-001, ATL-COM-GEN-2008-003, 2008, URL: https://cds.cern.ch/record/1110290 (cit. on p. 58).

[206] H Abramowicz et al., *Report of the ATLAS figures task force*, ATL-COM-GEN-2014-007, 2014, URL: https://cds.cern.ch/record/1707807 (cit. on p. 58).

[207] *Style guide for ATLAS Plots*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PubComPlotStyle#Axis_labels, Accessed: 2024-01-12 (cit. on p. 58).

[208] *TemplateForNewMCPPrescriptions*, https://twiki.cern.ch/twiki/bin/view/Main/TemplateForNewMCPPrescriptions#Supported_systematics_AN4, Accessed: 2024-01-15 (cit. on p. 60).

[209] Jet energy scale and jet energy resolution uncertainties for full run 2 analyses, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetUncertaintiesRel21Summer2018SmallR (cit. on p. 60).

[210] TSTSytematics, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TSTSytematics (cit. on p. 60).

[211] ATLAS Collaboration, *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, Eur. Phys. J. C **83** (2023) 681, arXiv: 2211.16345 [physics.data-an] (cit. on p. 60).

[212] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv: 1907.05120 [hep-ex] (cit. on p. 60).

[213] Ilaria Luise, *Two New Developments on the Statistical Treatment of Flavour Tagging Uncertainties in ATLAS*, (2022), URL: https://cds.cern.ch/record/2823374 (cit. on p. 60).

[214] ATLAS Collaboration, *Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV*, JHEP **08** (2018) 089, arXiv: 1805.01845 [hep-ex] (cit. on p. 60).

[215] Pileup reweighting, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ExtendedPileupReweighting (cit. on p. 61).

[216] William Buttinger, *Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data*, ATL-COM-SOFT-2015-119, 2015, URL: https://cds.cern.ch/record/2014726 (cit. on p. 61).

[217] ATLAS Collaboration, *$ZZ \to \ell^+\ell^-\ell'^+\ell'^-$ cross-section measurements and search for anomalous triple gauge couplings in 13 TeV pp collisions with the ATLAS detector*, Phys. Rev. D **97** (2018) 032005, arXiv: 1709.07703 [hep-ex] (cit. on p. 62).

[218] ATLAS Collaboration, *Measurement of ZZ production cross-sections in the four-lepton final state in pp collisions at $\sqrt{s} = 13.6$ TeV with the ATLAS experiment*, (2023), arXiv: 2311.09715 [hep-ex] (cit. on p. 62).

[219] *Standard Model Cross Sections for Diboson/Triboson Productions*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/SMDC14xsecs, Accessed: 2023-12-08 (cit. on p. 62).

[220] ATLAS Collaboration, *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) 072009, arXiv: 1901.03584 [hep-ex] (cit. on p. 62).

[221] *CrossSectionNLOttV*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/CrossSectionNLOttV, Accessed: 2023-12-08 (cit. on p. 62).

[222] ATLAS Collaboration, *Multi-boson simulation for 13 TeV ATLAS analyses*, ATL-PHYS-PUB-2016-002, 2016, URL: https://cds.cern.ch/record/2119986 (cit. on p. 62).

[223] ATLAS Collaboration, *Evidence for the production of three massive vector bosons with the ATLAS detector*, Phys. Lett. B **798** (2019) 134913, arXiv: 1903.10415 [hep-ex] (cit. on p. 62).

[224] PmgSystematicUncertaintyRecipess, `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PmgSystematicUncertaintyRecipes` (cit. on pp. 62, 63).

[225] PMGSystematicsTool, `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PMGSystematicsTool` (cit. on p. 62).

[226] Jon Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv: `1510.03865 [hep-ph]` (cit. on pp. 62, 63).

[227] Recommendation for using PDFs, `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PdfRecommendations` (cit. on p. 62).

[228] J. Gao et al., *CT10 next-to-next-to-leading order global analysis of QCD*, Phys. Rev. D **89** (2014) 033009, arXiv: `1302.6246 [hep-ph]` (cit. on p. 62).

[229] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C **75** (2015) 204, arXiv: `1412.3989 [hep-ph]` (cit. on p. 62).

[230] Bing Zhou et al., *Measurement of ZZ CP-violation and polarization in four-lepton dataset in 13 TeV proton-proton collisions with the ATLAS detector*, ATL-COM-PHYS-2022-075, 2022, URL: `https://cds.cern.ch/record/2801395` (cit. on pp. 63, 65).

[231] ATLAS Collaboration, *Evidence of pair production of longitudinally polarised vector bosons and study of CP properties in $ZZ \to 4\ell$ events with the ATLAS detector at $\sqrt{s} = 13$ TeV*, (2023), arXiv: `2310.04350 [hep-ex]` (cit. on pp. 63, 65).

[232] Julien Baglio, Le Duc Ninh, and Marcus M. Weber, *Massive gauge boson pair production at the LHC: a next-to-leading order story*, Phys. Rev. D **88** (2013) 113005, [Erratum: Phys.Rev.D 94, 099902 (2016)], arXiv: `1307.4331 [hep-ph]` (cit. on pp. 64, 65).

[233] Julien Baglio and Le Duc Ninh, *Fiducial polarization observables in hadronic $WZ$ production: A next-to-leading order QCD+EW study*, JHEP **04** (2019) 065, arXiv: `1810.11034 [hep-ph]` (cit. on pp. 64, 65).

[234] *Standard Model Summary Plots February 2022*, ATL-PHYS-PUB-2022-009, All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2022-009, 2022, URL: `https://cds.cern.ch/record/2804061` (cit. on p. 71).

[235] Rene Brun and Fons Rademakers, *ROOT – An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389** (1997) 81, ISSN: 0168-9002 (cit. on pp. 72, 90, 91).

[236] Andreas Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, 2007, arXiv: `physics/0703039 [physics.data-an]` (cit. on pp. 72, 74, 75).

[237] Carsten Bittrich, *Evidence for Scattering of Electroweak Gauge Bosons in the $W^{\pm}Z$ Channel with the ATLAS Detector at the Large Hadron Collider*, Presented 14 Jul 2020, Dresden, Tech. U., 2020, URL: `https://cds.cern.ch/record/2719126` (cit. on pp. 72, 74, 78, 79, 90).

[238] Yoav Freund and Robert E. Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, J. Comput. Syst. Sci. **55** (1997) 119 (cit. on p. 73).

[239] Trevor Hastie, Robert Tibshirani, and Jerome Friedman, "Boosting and Additive Trees", *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, New York, NY: Springer New York, 2009 337, ISBN: 978-0-387-84858-7, URL: `https://doi.org/10.1007/978-0-387-84858-7_10` (cit. on p. 74).

[240] BaBar Collaboration, *The BABAR physics book: Physics at an asymmetric B factory*, 1998 (cit. on p. 75).

[241] Erik Bachmann, *Polarization studies in same-sign W-boson pair production at the LHC with the ATLAS detector*, Presented 12 Apr 2023, TUD Dresden University of Technology, 2023, URL: https://cds.cern.ch/record/2872930 (cit. on p. 78).

[242] E. Bachmann, *OPTIMA*, https://gitlab.cern.ch/atlas-germany-dresden-vbs-group/optima, Accessed: 2023-10-16 (cit. on p. 78).

[243] G. Cowan, *Statistical data analysis*, Oxford University Press, USA, 1998 (cit. on pp. 89, 91, 93, 95).

[244] *StatisticTool*, https://gitlab.cern.ch/atlas-germany-dresden-vbs-group/StatisticTool/ and https://gitlab.cern.ch/atlas-physics/sm/ew/ssww-run2/StatTools; Accessed: 2023-11-15 (cit. on p. 90).

[245] ROOT Collaboration, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, CERN-OPEN-2012-016, 2012, URL: https://cds.cern.ch/record/1456844 (cit. on pp. 90, 93).

[246] Lorenzo Moneta et al., *The RooStats Project*, 2011, arXiv: 1009.1003 [physics.data-an] (cit. on p. 90).

[247] Wouter Verkerke and David Kirkby, *The RooFit toolkit for data modeling*, 2003, arXiv: physics/0306116 [physics.data-an] (cit. on p. 90).

[248] Franziska Iltzsche, *Observation of Electroweak Same-Charge $W^{\pm}W^{\pm}jj$ Production at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC*, Presented 27 Sep 2019, Dresden, Tech. U., 2019, URL: https://cds.cern.ch/record/2697175 (cit. on p. 90).

[249] Tim Herrmann, *Total and Differential Cross-Section Measurement and Interpretation of $W^{\pm}W^{\pm}$ Scattering in Effective Field Theories with the Run 2 Data of the ATLAS Experiment*, Presented 27 Oct 2023, Dresden, Tech. U., 2023 (cit. on pp. 90, 92, 95, 150, 151).

[250] M. Hatlo et al., *Developments of mathematical software libraries for the LHC experiments*, IEEE Trans. Nucl. Sci. **52** (2005) 2818 (cit. on p. 91).

[251] Fred JAMES and Matthias WINKLER, *Minuit 2*, https://root.cern/root/htmldoc/guides/minuit2/Minuit2.pdf, Accessed: 2023-11-14 (cit. on pp. 91, 93, 94).

[252] Jerzy Neyman and Egon Sharpe Pearson, *On the Problem of the Most Efficient Tests of Statistical Hypotheses*, Phil. Trans. Roy. Soc. Lond. A **231** (1933) 289 (cit. on p. 91).

[253] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv: 1007.1727 [physics.data-an] (cit. on pp. 91, 92), Erratum: Eur. Phys. J. C **73** (2013) 2501.

[254] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math.Stat. **9** (1938) 60 (cit. on p. 92).

[255] Abraham Wald, *Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large*, Transactions of the American Mathematical Society **54** (1943) 426, ISSN: 00029947, URL: http://www.jstor.org/stable/1990256 (visited on 11/15/2023) (cit. on p. 92).

[256] *PiecewiseInterpolation*, https://root.cern/doc/v608/HistoToWorkspaceFactoryFast_8cxx_source.html#l00514 and https://root.cern/doc/v608/PiecewiseInterpolation_8cxx_source.html#l00225; Accessed: 2023-11-15 (cit. on p. 93).

[257] *FlexibleInterpVar*, https://root.cern/doc/v608/HistoToWorkspaceFactoryFast_8cxx_source.html#l00728 and https://root.cern/doc/v608/FlexibleInterpVar_8cxx_source.html#l00453; Accessed: 2023-11-15 (cit. on p. 93).

[258] Ralf Wischnewski, *Errors in weighted histograms*, https://www-zeuthen.desy.de/~wischnew/amanda/discussion/wgterror/working.html, Accessed: 2023-11-15 (cit. on p. 93).

[259] Aaron Armbruster, *Advanced analysis / RooFit / HistFactory tutorial*, https://indico.cern.ch/event/634161/contributions/2678994/attachments/1524553/2383382/armbruster.pdf, Accessed: 2023-11-15 (cit. on p. 96).

[260] Aaron Armbruster, *parametrizeShape*, https://gitlab.cern.ch:8443/armbrust/RooFit-tutorial/-/blob/master/macros/parametrizeShape.C, Accessed: 2023-11-15 (cit. on p. 96).

[261] Spyros Argyropoulos, *Systematics pruning & smoothing*, https://indico.cern.ch/event/691683/contributions/2873279/attachments/1593521/2522846/PruningSmoothing.pdf, Accessed: 2023-11-15 (cit. on p. 97).

[262] Recommendations for unfolding, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/StandardModelUnfoldingNew (cit. on p. 150).

[263] Valerie Lang, Stefan Richter, and Matthias Schott, *VIPUnfolding*, https://gitlab.cern.ch/atlas-physics/sm/StandardModelTools_Unfolding/VIPUnfolding (cit. on p. 150).

[264] Tim Adye, *Unfolding algorithms and tests using RooUnfold*, (2011) 313, arXiv: 1105.1160 [physics.data-an] (cit. on p. 150).

[265] Jan-Eric Nitschke, *VIPUnfolding Fork*, https://gitlab.cern.ch/jnitschk/VIPUnfolding/-/tree/TimCodeWithFixes (cit. on p. 150).

[266] Tim Herrmann, *Study of Different Unfolding Methods of Kinematic Distributions of the $WZ \to WZ$ Scattering with Data and Simulations of the ATLAS Detector at the LHC*, Presented 2017, Dresden, Tech. U., 2017, URL: https://cds.cern.ch/record/2267776 (cit. on p. 151).

[267] Volker Blobel, *Unfolding Methods in Particle Physics*, (2011) 240, URL: https://cds.cern.ch/record/2203257 (cit. on p. 151).

[268] G. Apollinari, O. Brüning, T. Nakamoto, and Lucio Rossi, *High Luminosity Large Hadron Collider HL-LHC*, CERN Yellow Rep. (2015) 1, ed. by G Apollinari, I Béjar Alonso, O Brüning, M Lamont, and L Rossi, arXiv: 1705.08830 [physics.acc-ph] (cit. on p. 176).

# Danksagung

Mein aufrichtiger Dank gilt Michael Kobel und Frank Siegert, meinen engagierten Betreuern während meines Bachelor-, Master- und Promotionsstudiums. Ihre ständige Verfügbarkeit und unermüdliche Unterstützung waren für meinen akademischen Fortschritt von grundlegender Bedeutung.

Vielen Dank möchte ich auch gebenüber Joany Manjarrés ausdrücken. Dies gilt für die Unterstützung während ihrer Zeit hier in Dresden und danach von Toulouse aus. Ihre Hilfe bei physikalischen Fragen, aber auch bei der Vorbereitung von Vorträgen war für mich unentbehrlich. Das Navigieren durch die Prozesse einer ATLAS-Analyse von der Entwicklung (Abspaltung) über die Durchführung und schließlich die Finalisierung, mit Editorial Board, Standard Model Group und PAM Review und Approval wären ohne sie nicht möglich gewesen.

Besonderer Dank gilt Maren Bühring, Johann Voigt, Erik Bachmann, Mareen Hoppe, Max Stange und Tim Herrmann für das sorgfältige Korrekturlesen meiner Arbeit und ihre Bereitschaft, technische Fragen in den Bereichen Statistik, Machine learning, Unfolding and Polarisation zu beantworten. Ihre Beiträge haben die Qualität meiner Arbeit erheblich verbessert, und ihre verlässliche Unterstützung war von unschätzbarem Wert.

Ich bin dem gesamten Institut, insbesondere der Arbeitsgruppe, für die freundliche und kooperative Atmosphäre dankbar. Die Zusammenarbeit mit allen war eine angenehme Erfahrung, sowohl beruflich als auch persönlich.

Ein herzliches Dankeschön an meine Eltern und meine Schwester, die mich mein ganzes Leben lang unterstützt und die Grundlage für meine akademischen Bestrebungen geschaffen haben. Danken möchte ich auch Antonia, die mir während der Arbeit nicht nur emotional und mental zur Seite stand, sondern auch als unverzichtbare Quietscheente diente, wenn ich fachliche oder schreibtechnische Probleme zu lösen hatte.

Allen Genannten und den nicht explizit Genannten sei gesagt, dass ihre Beiträge eine entscheidende Rolle bei der Gestaltung meines akademischen Weges gespielt haben, wofür ich aufrichtig dankbar bin.

# Acknowledgments

My sincere thanks go to Michael Kobel and Frank Siegert, my dedicated supervisors during my bachelor's, master's, and doctoral studies. Their constant availability and tireless support were fundamental to my academic progress.

I would also like to express my gratitude to Joany Manjarrés. This applies to the support during her time here in Dresden and afterward from Toulouse. Her help with physics questions and the preparation of talks was indispensable for me. Navigating through the processes of an ATLAS analysis from development (spin-off) to implementation and ultimately finalization, with Editorial Board, Standard Model Group, and PAM Review and Approval would not have been possible without her.

Special thanks go to Maren Bühring, Johann Voigt, Erik Bachmann, Mareen Hoppe, Max Stange, and Tim Herrmann for their careful proofreading of my work and their willingness to answer technical questions in the areas of statistics, machine learning, unfolding, and polarization. Their contributions have greatly improved the quality of my work and their reliable support has been invaluable.

I am grateful to the whole institute, especially the working group, for the friendly and cooperative atmosphere. Working with everyone has been a pleasant experience, both professionally and personally.

A big thank you to my parents and sister who have supported me throughout my life and laid the foundation for my academic endeavors. I would also like to thank Antonia, who not only supported me emotionally and mentally during my work but also served as an indispensable rubber duck when I had to solve technical or writing problems.

To all those mentioned and those not explicitly named, their contributions have played a crucial role in shaping my academic path, for which I am sincerely grateful.

## Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt. Die vorliegende Dissertation wurde in der Zeit von August 2020 bis Februar 2024 im Institut für Kern- und Teilchenphysik unter der wissenschaftlichen Betreuung von Dr. Frank Siegert angefertigt. Es haben keine früheren erfolglosen Promotionsverfahren stattgefunden. Ich erkenne die Promotionsordnung des Bereichs Mathematik und Naturwissenschaften an der Technischen Universität Dresden von 23.02.2011 an.

Jan-Eric Nitschke
Dresden, 01.02.2024