TECHNISCHE
UNIVERSITÄT
DRESDEN

Multivariate Analysis of Polarized Vector Boson Scattering of a W and Z Boson with the ATLAS Detector

Master-Arbeit
zur Erlangung des Hochschulgrades
Master of Science
im Master-Studiengang Physik

vorgelegt von

Anna Sophie Tegetmeier
geboren am 09.07.1997 in Hamburg

Institut für Kern- und Teilchenphysik
Fakultät Physik
Bereich Mathematik und Naturwissenschaften
Technische Universität Dresden
2022

Eingereicht am 07. November 2022

1. Gutachter: Dr. Frank Siegert
2. Gutachter: Prof. Dr. Arno Straessner

Summary

Abstract
English:
The scattering of two longitudinally polarized gauge bosons can be used to test details of the mechanism of electroweak symmetry breaking and to search for physics beyond the Standard Model of particle physics. This process can be studied in the electroweak production of a WZ pair in association with two jets with the final state $\ell\nu\ell\ell jj$. In this thesis this process was investigated for a center of mass energy $\sqrt{s} = 13$ TeV and an integrated luminosity of 139 fb^{-1}. A validation of the simulation of this processes with Madgraph that includes the polarization information of the bosons was performed. A multivariate analysis using Boosted Decision Trees was performed using events with longitudinally polarized W and Z boson (WLZL), longitudinally polarized W boson (WLZX), longitudinally polarized Z boson (WXZL) or at least one longitudinally polarized boson (min1L) as signal and the polarizations that are not part of the signal as well as other non-polarization backgrounds as background. Maximum Likelihood fits to Asimov data were performed using the BDT outputs. Different fit configurations were used including a stat-only fit and fits that include systematic uncertainties of the non-polarization backgrounds. For the WLZL signal all fits lead to significances of less than 1 standard deviation (σ). All fits to the mixed signal configurations WLZX and WXZL result in significances of less than $3\,\sigma$. For the min1L signal a significance of about $3\,\sigma$ could be obtained by the fit if the background polarization was assumed to be constrained by 25% systematic uncertainties. Fits using higher integrated luminosities indicate that evidence of the WLZL process might be seen at the High Luminosity LHC and evidence of WLZX and WXZL in Run 3 if the background polarizations can be sufficiently constrained.

Zusammenfassung
Deutsch
Die Streuung zweier longitudinal polarisierter Eichbosonen kann genutzt werden, um Details des Mechanismus der elektroschwachen Symmetriebrechung zu testen und um nach Physik jenseits des Standardmodells der Teilchenphysik zu suchen. Dieser Prozess kann bei der elektroschwachen Produktion eines WZ Paares zusammen mit zwei Jets mit dem Endzustand $\ell\nu\ell\ell jj$ untersucht werden. In dieser Arbeit wurde dieser Prozess für eine Schwerpunktsenergie $\sqrt{s} = 13$ TeV und eine integrierte Luminosität von 139 fb^{-1} untersucht.
Eine Validierung der Simulation der polarisierten Prozesse durch Madgraph wurde durchgeführt. Für eine multivariate Analyse wurden Boosted Decision Trees trainiert, wobei Ereignisse mit longitudinal polarisierten W- und Z-Bosonen (WLZL), longitudinal polarisierten W-Boson (WLZX), longitudinal polarisierten Z-Boson (WXZL) oder mindestens einem longitudinal polarisierten Boson (min1L) als Signal und die anderen Polarisationen sowie weitere Untergründe als Untergrund verwendet wurden. Maximum-Likelihood-Fits gegen Asimov-Daten wurden unter Verwendung der BDT-Werte durchgeführt. Verschiedene Fit Einstellungen wurden verwendet, darunter ein stat-only Fit und Fits, die systematische Unsicherheiten der nicht-Polarisations Untergründe beinhalten. Für das WLZL Signal führen alle Fits zu Signifikanzen von weniger als einer Standardabweichung (σ). Alle Fits für die gemischten Signalkonfigurationen WLZX und WXZL führen zu Signifikanzen von weniger als $3\,\sigma$. Für das min1L Signal kann eine Signifikanz von etwa $3\,\sigma$ erhalten werden, wenn angenommen wird, dass die Untergrundpolarisation mit 25% systematischer Unsicherheit eingeschränkt ist. Fits, die für höhere integrierte Luminositäten durchgeführt wurden, deuten darauf hin, dass Hinweise auf den WLZL-Prozess im High Luminosity LHC und Hinweise auf WLZX und WXZL im Run 3 gefunden werden könnten, falls die Untergrundpolarisationen ausreichend eingeschränkt werden können.

Contents

1 Introduction **1**

2 Theoretical Foundations **3**

2.1 The Standard Model of Particle Physics . . . 3

2.1.1 Electroweak Theory . . . 5

2.1.2 Electroweak Symmetry Breaking . . . 7

2.2 Vector Boson Scattering . . . 10

2.3 Polarization . . . 13

2.3.1 Polarization in VBS . . . 13

2.3.2 Dependence of Variables on the Polarizations . . . 14

3 Experiment **19**

3.1 The Large Hadron Collider . . . 19

3.2 The ATLAS Detector . . . 20

3.2.1 Coordinate System and Common Variables . . . 20

3.2.2 The Inner Detector . . . 21

3.2.3 The Calorimeters . . . 22

3.2.4 The Muon System . . . 23

4 Simulation of Events **25**

4.1 Basic Principal . . . 25

4.2 Simulation of Polarized Events . . . 27

5 Study of the Simulated Events **29**

5.1 Study on Truth Level . . . 29

5.1.1 Definition of Variables . . . 29

5.1.2 Phase Space Definition . . . 30

5.1.3 Effects of the Approximations . . . 31

5.1.4 Comparison to the Analytical Function . . . 32

5.1.5 Study of the Different Polarizations . . . 33

5.1.6 Comparison of the Parton Parton and Helicity Frame . . . 36

5.1.7 Validation of Samples with τ Decays . . . 41

5.1.8 Comparison of the Local and Official Samples . . . 43
5.2 Study at Reconstruction Level . . . 47
5.2.1 Object Selection . . . 47
5.2.2 Variables at Reconstruction Level . . . 48
5.2.3 Comparison of Distributions Between Truth and Reco Level . . . 50
5.2.4 Phase Space Definition at Reconstruction Level . . . 51
5.2.5 Inclusion of Backgrounds . . . 52

6 Extraction of the Longitudinal Polarization 59
6.1 Study with BDTs . . . 59
6.1.1 Introduction to Boosted Decision Trees . . . 60
6.1.2 Evaluation of the Performance . . . 61
6.1.3 Overtraining Conditions . . . 63
6.1.4 Variable and Metaparameter Optimization . . . 64
6.1.5 Input Variables and Preprocessing . . . 65
6.1.6 Results . . . 67
6.2 Studies of the Sensitivity to the Longitudinal Polarization . . . 77
6.2.1 Likelihood Function and Significance . . . 77
6.2.2 Fit Strategy . . . 80
6.2.3 Fit Results . . . 81
6.2.4 Fit to Pseudo Data . . . 84

7 Outlook for Run 3 and the High-Luminosity LHC 93
7.1 Fits at Higher Luminosities . . . 94
7.2 Cross Sections for $\sqrt{s} = 13.6$ TeV . . . 99

8 Summary and Outlook 101

A Mathematical Definitions 103

B Locally Produced Samples 104
B.1 Job Options . . . 104
B.1.1 Polarized Samples in the Helicity Frame . . . 104
B.1.2 Polarized Samples in the Parton Parton Frame . . . 108
B.1.3 Full Sample . . . 108
B.1.4 Unpolarized Sample . . . 108
B.2 Samples at Truth Level . . . 108

C Distributions of the Polarized Sample at Truth Level 111

D Comparison of Samples in the two Different Definitions of the Parton Parton Frame 118

E Distributions for Samples with τ Decays 121

F Comparison Plots for the Local and Official Samples 123

G Samples at Reconstruction Level 126

H Normalized Distributions of the Polarized Samples at Reco Level 129

I Background Distributions 133

J Ranking of the Input Variables 137

List of Figures 139

List of Tables 142

Bibliography 144

1 Introduction

Particle physics is the field of physics that deals with the fundamental building blocks of the universe and their interactions. The theory that describes our current knowledge about these elementary particles and their interactions is the Standard Model of particle physics (SM). Many predictions of the SM can only be probed at high energies. Such high energies can be achieved in particle accelerators like the Large Hadron Collider (LHC) at CERN. A great achievement was the observation of the Higgs boson by the ATLAS and CMS collaborations at the LHC in 2012. With that the particle content of the SM was completed. The measurements taken at the LHC provide the possibility to study the properties of the fundamental particles. So far no significant deviations from the SM have been found. However several phenomena cannot be explained by the SM. Examples are Dark Matter and the matter-antimatter asymmetry. Several theories have been formulated to describe physics beyond the SM. So far no additional particles, as predicted by many theories, have been observed. Using the Run 2 data from the LHC with a center of mass energy of $\sqrt{s} = 13$ TeV more precise measurements of the properties of the particles and their interactions can be done to probe the predictions of the SM. The large amount of data offers the possibility to measure processes that have not been observed so far. An example is the polarized vector boson scattering of the electroweak gauge bosons. Here the scattering of two longitudinally polarized bosons is of special interest. Without the Higgs boson this process leads to unphysical predictions. This connection makes this process sensitive to the details of the mechanism of electroweak symmetry breaking. The process of vector boson scattering is included when studying the electroweak production of VV pair production ($V = W, Z$) in association with two jets. The leptonic WZ channel is easier to reconstruct than channels with two W bosons due to only one neutrino in the final state, while it has a larger cross section than the ZZ channel. The electroweak $WZjj$ production was first observed by the ATLAS collaboration using a subset of the Run 2 data with an integrated luminosity of 36 fb^{-1} with a significance of 5.3 standard deviations [1]. The same process was measured using the full Run 2 data with an integrated luminosity of 137 fb^{-1} with a significance of 6.8 standard deviations [2].

In this thesis polarized vector boson scattering will be investigated in the electroweak production of polarized W and Z bosons in the final state of three leptons, one neutrino and two jets. First studies of this process have been done in [3] and [4]. In this thesis a multivariate analysis of this process is performed. In preparation for that samples with polarization information of

the W and Z bosons produced with the Monte-Carlo generator Madgraph are validated.
In chapter 2 an introduction to the SM of particle physics, specifically with the focus on vector boson scattering will be given. The experimental setup of the LHC and the ATLAS detector is described in chapter 3. The simulations of events are explained in chapter 4. The simulated events are studied at truth and reconstruction level in chapter 5. A validation of this samples is performed and variables that show discrimination power between the different polarizations of the gauge bosons are investigated. Additional backgrounds are included at reconstruction level. In chapter 6 Boosted Decision Trees are used to separate different signal processes which include events with longitudinal polarized vector bosons from the backgrounds. Maximum Likelihood fits are performed to extract the expected significance. An outlook for studying the process at higher luminosities during Run 3 and at the High Luminosity LHC is given in chapter 7. A summary and outlook for further studies is given in chapter 8.

2 Theoretical Foundations

In the following chapter a basic introduction into the theory behind the studied process is given. This follows closely the approach in [3] and [4].

2.1 The Standard Model of Particle Physics

In nature four fundamental interactions are known, namely the strong, weak and electromagnetic interaction and gravity. The first three are described in the theoretical framework called the Standard Model of Particle Physics (SM). Gravity is described by the theory of general relativity. While there are many attempts and ongoing efforts to find a theory that includes all four interactions, no widely accepted theory has been found so far. Since at scales relevant for particle physics the gravitational force is weak compared to the other forces, gravity does not play an important role when studying elementary particles and is typically ignored.
The SM is a relativistic quantum field theory that is based on local gauge symmetries. This means that the Lagrangian of the SM has to be invariant under certain transformations. A brief overview is given in the following.
The SM has a local

$$U(1)_Y \otimes SU(2)_L \otimes SU(3)_C$$

symmetry. Here the $SU(3)_C$ subgroup is responsible for strong interactions while the $U(1)_Y \otimes SU(2)_L$ subgroup represents the unification of the electromagnetic and weak interaction.
According to the Noether theorem [5] continuous symmetries result in conserved quantities. With this theorem the charges can be derived from the SM Lagrangian. For the strong interaction this is the colour charge $\vec{C}$, for the weak interaction the weak isospin I_W^3 and for electromagnetism the electric charge Q.
Particles interact with each other by the exchange of gauge bosons. To take part in an interaction the particles must hold the charge associated by the respective interaction. For the strong interaction the mediating particles are gluons g. The $W^{\pm}$ and the Z boson transmit the weak interaction and the photon γ the electromagnetic interaction.
Bosons are particles characterized by an integer spin. The gauge bosons, which are often also called vector bosons, all have a spin of 1. The only other known elementary particle that belongs to this group is the Higgs boson H. It is related to the mechanism that gives

particles their masses and has a spin of 0. The bosons that are part of the SM along with their corresponding interaction are summarized in Table 2.1.
The other group of particles are fermions which are characterized by a half-integer spin. All known elementary fermions have a spin of 1/2. Since the fermions are the building blocks of matter, they are also called matter particles. Fermions can further be grouped by their charges. Those who have no colour charge are called leptons and those with colour charge quarks. The group of leptons consists of charged leptons with an electric charge of 1 and neutrinos with asn electric charge of 0. There are three generations of known charged leptons, namely the electron e, the muon μ and the tau-lepton τ. They differ only in their mass, while a higher generation corresponds to a higher mass. Equivalently there are three generations of neutrinos, namely the electron neutrino ν_e, the muon neutrino ν_μ and the tau neutrino ν_τ.
For the quarks, those who have an electric charge of 2/3 are called up-type quarks and those with an electric charge of -1/3 down-type quarks. The quarks have one of the three colour charges red, green and blue (r, g, b). They are always grouped in a way such that in total they are colour neutral, which also leads to integer values in the electric charge. This is known as confinement. The three generations of the up-type quarks are the up u, the charm c and the top t quark. For the down-type quark these are the down d, the strange s and the bottom b quark.
A special property of the weak interaction is, that it violates parity and couples differently to different chiralities [6]. A field ψ can be separated into left and right chiral fields ψ_L and ψ_R by applying projection operators:

$$P_{L/R} = \frac{1}{2}\left(1 \mp \frac{\gamma^5}{2}\right). \tag{2.1}$$

Table 2.1: Overview of the bosons described in the SM and their associated interactions.

Boson	associated interaction	associated charge	spin
g	strong	$\vec{C}$	1
$W^\pm, Z$	weak	I_3^W	1
γ	electromagnetic	Q	1
H	-	-	0

Here γ^5 is the fifth Dirac matrix as explained in Appendix A. By construction the following relation applies:

$$\psi = \psi_L + \psi_R. \tag{2.2}$$

For the fermion fields the weak isospin is 1/2 for left-handed and 0 for right-handed fields. With that left-handed fields are represented as doublets and right-handed fields as singlets in the weak isospin. For the leptons of the first generation this can be noted as

$$L_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \tag{2.3}$$

$$\ell_R = e_R. \tag{2.4}$$

For the leptons of the other generations the notation is equivalent. Right-handed neutrinos do not participate in any interaction and no evidence of their existence has been found so far. Thus normally they are not included in the Lagrangian of the SM and are not mentioned above. For quarks a similar notation can be used, shown here again for the first generation:

$$Q_{L,\alpha} = \begin{pmatrix} u_\alpha \\ d_\alpha \end{pmatrix}_L \tag{2.5}$$

$$u_{R,\alpha} \tag{2.6}$$

$$d_{R,\alpha}. \tag{2.7}$$

$$\tag{2.8}$$

The α denotes the different colour charges.
An overview of the fermions together with their corresponding charges is given in Table 2.2. For each fermion also an anti-particle exist, for which all charges are reversed.

2.1.1 Electroweak Theory

The electroweak theory was developed in the 1960s by Glashow, Salam and Weinberg to unify the electromagnetic and weak interaction [7, 8, 9]. Using the notation introduced in the previous section to account for the parity violation of the weak interaction, the Lagrangian of the electroweak interaction can be written as

$$\begin{aligned} \mathcal{L} = \overline{L} i\gamma^\mu D_\mu L \quad &+ \overline{e}_R i\gamma^\mu D_\mu e_R \quad + \overline{u}_R i\gamma^\mu D_\mu u_R \\ + \overline{Q} i\gamma^\mu D_\mu Q \quad &+ \overline{d}_R i\gamma^\mu D_\mu d_R \quad + \overline{u}_R i\gamma^\mu D_\mu u_R \\ - \frac{1}{4} W^a_{\mu\nu} W^{a,\mu\nu} \quad &- \frac{1}{4} B_{\mu\nu} B^{\mu\nu}. \end{aligned} \tag{2.9}$$

Table 2.2: Overview of the fermions described by the SM and their corresponding charges. Anti-particles, which have the same masses but reversed charges are not included in the table. For the weak interaction only the charge for left-handed particles is shown. For right-handed ones the weak isospin is always 0.

	Generation			Charge		
	1.	2.	3.	Q	I_3^W	$\vec{C}$
fermions	ν_e	ν_μ	ν_τ	0	$+\frac{1}{2}$	$\vec{0}$
	e	μ	τ	-1	$-\frac{1}{2}$	$\vec{0}$
quarks	u	c	t	$+\frac{2}{3}$	$+\frac{1}{2}$	r, g, b
	d	s	b	$-\frac{1}{3}$	$-\frac{1}{2}$	r, g, b

For readability only the first generation of particles is included. The terms for particles of the other generations can be written equivalently. Here D_μ denotes the covariant derivative:

$$D_\mu = \partial_\mu + ig_W T^a W_\mu^a + ig_Y Y B_\mu. \tag{2.10}$$

It is introduced to ensure local gauge invariance and introduces the new gauge fields W_μ^a with a $= 1, 2, 3$ and B_μ with coupling constants g_W and g_Y. The T^a are the generators of the SU(2) symmetry with

$$T^a = \frac{1}{2}\sigma^a \tag{2.11}$$

and σ^a the Pauli matrices (see Appendix A). The weak hypercharge Y is the generator of the U(1) symmetry. It is related to the electric charge and the weak isospin through

$$Q = Y + T_3^W. \tag{2.12}$$

With the covariant derivative the Lagrangian is invariant under the simultaneous transformation of the fermion fields

$$\psi(x) \to \exp\left(\alpha(x)Y + i\beta(x)T^a\right)\psi \tag{2.13}$$

and the gauge fields

$$W_{\mu}^{a}(x) \rightarrow W_{\mu}^{a}(x) - \frac{1}{g_w}\partial_{\mu}\beta^{a}(x) - \epsilon_{abc}\beta^{b}(x)\beta^{c}(x) \tag{2.14}$$

$$B_{\mu}(x) \rightarrow B_{\mu}(x) - \frac{1}{g_Y}\partial_{\mu}\alpha^{a}(x). \tag{2.15}$$

In Equation 2.9 the first two terms describe the kinematics of the fermions and interactions between fermions and gauge bosons. The last term is included in the Lagrangian to describe the dynamics of these fields:

$$W_{\mu\nu}^{a} = \partial_{\mu}W_{\nu}^{a} - \partial_{\nu}W_{\mu}^{a} - g_W\epsilon_{abc}W_{\mu}^{b}W_{\nu}^{c} \tag{2.16}$$

$$B_{\mu\nu}^{a} = \partial_{\mu}B_{\nu}^{a} - \partial_{\nu}B_{\mu}^{a}. \tag{2.17}$$

It can be seen that Equation 2.16 includes terms that are quadratic in W_{μ}. In the Lagrangian this leads to terms that are of order three and four of the field which can be interpreted as self interactions among the gauge bosons.

2.1.2 Electroweak Symmetry Breaking

Mass terms in the Lagrangian can be written as

$$\mathcal{L}_{m_f} = m_f\psi\overline{\psi} \tag{2.18}$$

$$\mathcal{L}_{m_b} = \frac{1}{2}m_b^2 V_{\mu}V^{\mu} \tag{2.19}$$

for fermion and boson fields respectively. In Equation 2.9 no such quadratic terms and consequently no masses of the fermions and gauge bosons are included. This stands in contradiction to the observed masses of the fermions and the $W^{\pm}$ and Z bosons. Since the mass terms break local gauge invariance they cannot be included naively.
A solution of this problem is spontaneous symmetry breaking. For that a globally gauge invariant Lagrangian without mass terms is constructed. The gauge invariant is however spontaneously broken in its ground state which leads to mass terms of the fermions and bosons.
In the SM the electroweak symmetry breaking (EWSB) is described in the Brout-Englert-Higgs (BEG) mechanism [10, 11]. For this a $SU(2)_L$ doublet which contains two complex scalar fields is introduced:

$$\begin{pmatrix} \phi^{+} \\ \phi^{0} \end{pmatrix}. \tag{2.20}$$

The Lagrangian is than extended by kinematic terms of this fields and a new potential:

$$\mathcal{L}_{BEH} = (D_\mu\phi)^\dagger (D^\mu\phi) - V(\phi). \tag{2.21}$$

The potential is introduced as

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda\left(\phi^\dagger\phi\right)^2 \tag{2.22}$$

with free parameters $\mu^2 < 0$ and $\lambda > 0$.
Choosing one of the minima of the potential, the vacuum expectation value ϕ_0 of the field can be written as

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ \nu\end{pmatrix} \tag{2.23}$$

with $\nu = \sqrt{-\frac{\mu^2}{\lambda}}$.
This ground states breaks the $SU(2)_L \otimes U(1)_Y$ symmetry. It can be found that it is invariant under the usual symmetry transformation of quantum electrodynamics (QED) $U(1)_{em}$:

$$\phi_0(x) \to \exp\left(-i\alpha(x)\left(\frac{\sigma_3}{2} + Y\mathbb{1}\right)\right)\phi_0. \tag{2.24}$$

A first order expansion of ϕ_0 leads to

$$\phi(x) \approx \begin{pmatrix}\phi_1(x) + i\phi_2(x)\\ \frac{\nu+H(x)}{\sqrt{2}} + i\phi_3(x)\end{pmatrix} \tag{2.25}$$

Using gauge transformation this can be simplified to

$$\phi(x) \approx \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ \nu + H(x)\end{pmatrix}. \tag{2.26}$$

Here the linear term $H(x)$ introduces a new particle known as the Higgs boson.
The scalar fields ϕ_1, ϕ_2 and ϕ_3 can be identified as three massless Goldstone bosons as predicted by the Nambu-Goldstone theorem which states that every spontaneously broken continuous symmetry creates massless spin 0 Goldstone bosons. They can be interpreted as the longitudinal polarizations of the $W^\pm$ and Z bosons as further discussed in subsection 2.3.1.
Inserting ϕ_0 into the Lagrangian leads to

$$|D_\mu|^2 = \frac{1}{2}\left(\partial_\mu H\right)^2 + \frac{1}{8}\left(\nu + H\right)^2 |g_W W_\mu^3 - g_Y B_\mu|^2 + \frac{1}{8}g_W^2\left(\nu + H\right)^2 |W_\mu^1 - W_\mu^2|^2 \tag{2.27}$$

With that new mass eigenstates Z_μ, A and $W_\mu^\pm$ can be defined which diagonalize the mass terms:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^3\right) \tag{2.28}$$

$$Z_\mu = \frac{1}{\sqrt{g_W^2 + g_Y^2}}\left(g_W W_\mu^3 - g_Y B_\mu\right) \tag{2.29}$$

$$A = \frac{1}{\sqrt{g_W^2 + g_Y^2}}\left(g_W W_\mu^3 + g_Y B_\mu\right). \tag{2.30}$$

They correspond to the physical gauge bosons. With that the term in the Lagrangian becomes

$$|D_\mu|^2 = \frac{1}{2}\left(\partial_\mu H\right)^2 + \frac{1}{8}g_W^2\left(\nu + H\right)^2 W_\mu^- W^{\mu+} + \frac{1}{8}\left(g_W^2 + g_Y^2\right)\left(\nu + H\right)^2 Z_\mu Z^\mu. \tag{2.31}$$

From this the masses of the gauge fields can simply be read off:

$$m_W = \frac{1}{2}g_W \nu \tag{2.32}$$

$$m_Z = \frac{1}{2}\sqrt{g_W^2 + g_Y^2}\nu \tag{2.33}$$

$$m_A = 0. \tag{2.34}$$

The field A stands for the massless photon.
Besides quadratic terms of the gauge fields Equation 2.31 also contains terms of order HVV and $HHVV$ ($V = W^\pm, Z$). This terms correspond to interactions between the Higgs boson and the gauge fields as shown in Figure 2.1.

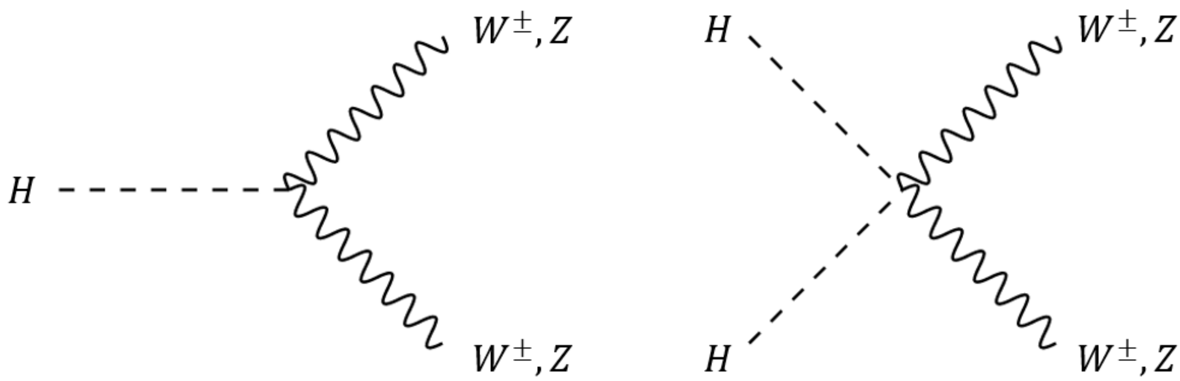


Figure 2.1: Feynman diagrams of the interactions of the Higgs boson with gauge bosons.

Inserting the vacuum expectation value in the potential one gets terms that are of order two, three and four of the Higgs boson. From the quadratic term the mass of the Higgs boson can be read off as

$$m_H = \sqrt{2\lambda\nu^2} \tag{2.35}$$

while the triple and quartic terms correspond to self interactions of the Higgs boson.
So far no mass terms for fermions have been introduced. Using the vacuum expectation value of the Higgs-field they can be inserted into the Lagrangian by hand. Looking only at the first generation of fermions, the first generation a mass term can be written as

$$\mathcal{L}_{Yukawa} = -\lambda_e \overline{L}\phi e_R - \lambda_d \overline{Q}\phi d_R - \lambda_u \overline{Q}\left(i\sigma_2\phi\right) u_R + \text{h.c.} \tag{2.36}$$

This leads to fermion masses

$$m_f = \frac{\lambda_f}{\sqrt{2}}\nu \tag{2.37}$$

where λ_f are the Yukawa couplings to the respective fermions.

2.2 Vector Boson Scattering

As mentioned in subsection 2.1.1, the SM predicts the scattering of the gauge bosons with each other. This is called vector boson scattering (VBS). Such processes have a very small cross section which makes it difficult to measure them. In this thesis the scattering of a $W^{\pm}$ with a Z boson is investigated. The Feynman diagrams contributing to this process are shown in Figure 2.2 Besides the Feynman diagram with a quartic vertex also two with triple vertices are

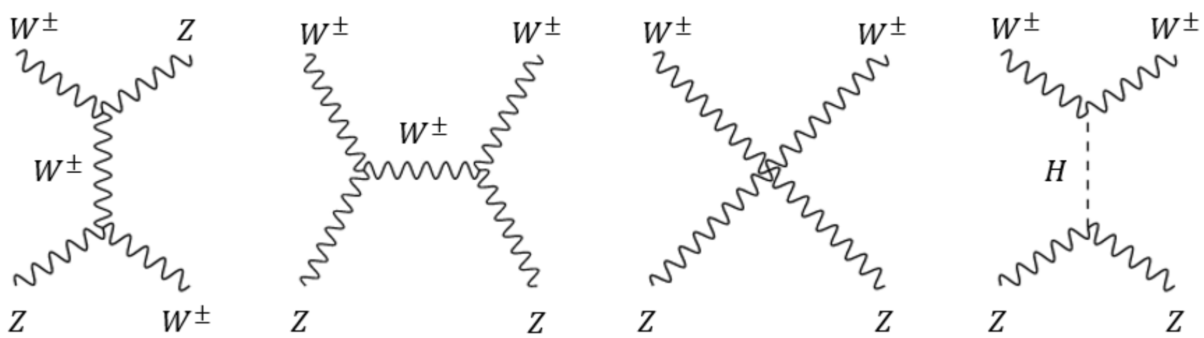


Figure 2.2: Feynmandiagrams contributing to the WZ VBS process.

part of the VBS process. In the last diagram the W and Z boson couple to the Higgs boson, as described in the previous section. This leads to a strong connection of VBS processes to EWSB which is discussed in more detail in subsection 2.3.1.
Since the W and Z boson are not stable the production and decay of them have to be taken

into account. At the LHC (see section 3.1) the W and Z bosons are emitted from quarks. The quarks get only weakly deflected which leads to a characteristic signature of the VBS process with a large rapidity difference and high invariant mass of the resulting jets.
In this thesis only the purely leptonic decay into electron, muons and neutrinos of the W and Z bosons is considered. While including the decay into tau leptons and the hadronic decay into quarks would increase the cross section, the chosen final state has the advantage of better background suppression and a better momentum resolution. With that the total final state is $\ell\nu_\ell\ell^\pm\ell^\mp jj$ with $\ell = e, \mu$. The first lepton and neutrino can stand either for a lepton and an anti-neutrino or an anti-lepton and a neutrino. The whole process is shown in Figure 2.3. The dashed circle stands for all the possible interactions of the W and Z boson as shown in Figure 2.2.

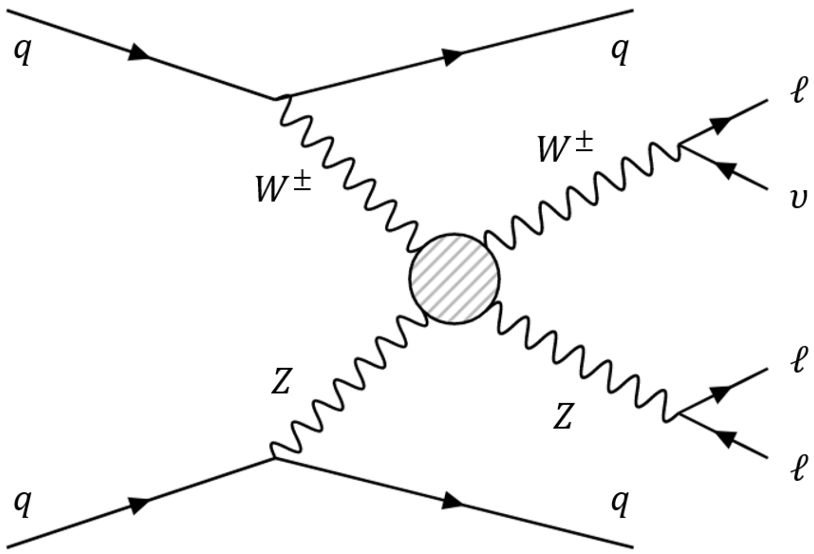


Figure 2.3: Feynmandiagrams of the WZ VBS process as it appears at the LHC.

It can be seen that the WZ VBS process is in leading order a purely electroweak process with a coupling order $\mathcal{O}(\alpha_{EW}^6)$, where α_{EW} denotes the electroweak coupling parameter. Non-VBS processes which have the same final state and the same coupling order are not gauge invariant separable from the WZ VBS process. Thus they have to be included in the signal definition. These processes are referred to as WZjj EW processes. In Figure 2.4 example Feynman diagrams of such processes are given. As shown with the diagram in the middle this also includes non-resonant diagrams where the leptons are not the direct decay products of the vector bosons.
A special process of the non-VBS processes with the same final state and the same coupling order is the so called tZj process. In this processes a b quark in the initial state radiates off a $W^\pm$ boson. The resulting top quark decays further into a b and a W boson so that this process always has a b quark in the initial and final state. An example diagram is shown

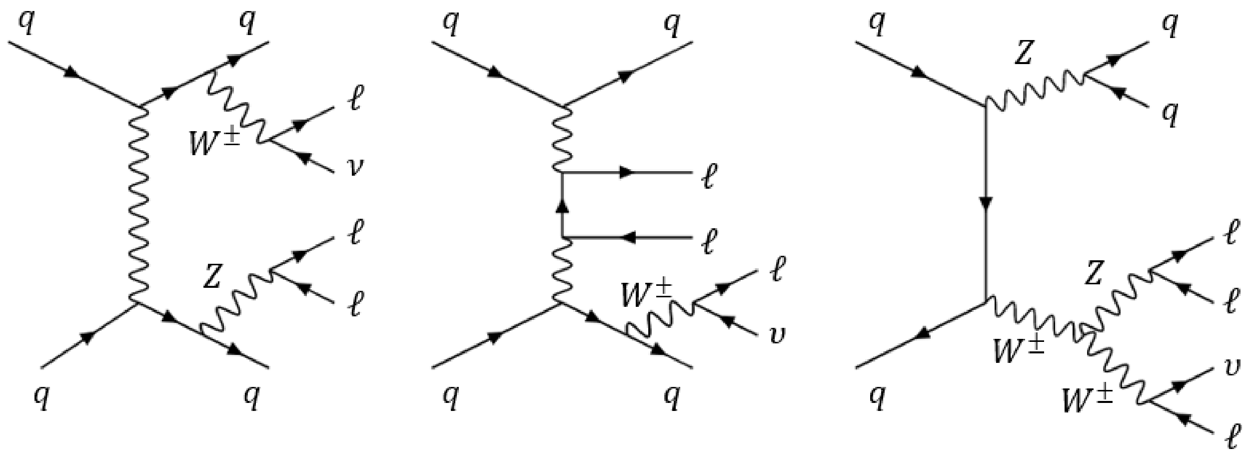


Figure 2.4: Example of non-VBS Feynman diagrams with same final state and same coupling order as the WZ VBS process.

in Figure 2.5. While it is not gauge invariantly separable from the signal process, b-tagging allows the identification processes with a b-quark in the final state. due to the W resonance tZj processes have a considerable contribution to processes with a $\ell\nu\ell\ell jj$ final state. Therefore processes with a b-quark in the initial or final state are often considered as an additional background. In this analysis however those processes are included in the signal process and not explicitly considered as a background contribution.

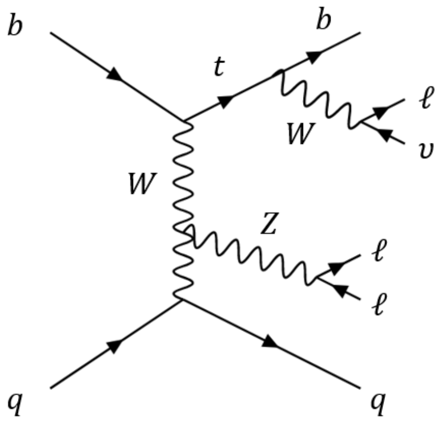


Figure 2.5: Example of a Feynman diagram of the tZj process.

The same final state as for the WZjj EW processes can also be produced in so called WZjj QCD processes. They have a coupling order $\mathcal{O}(\alpha_{EW}^4 \alpha_{QCD}^2)$. Since they are gauge invariantly separable from WZjj EW processes they are considered as a background process. In Figure 2.6 some examples of Feynman diagrams of such processes are shown.

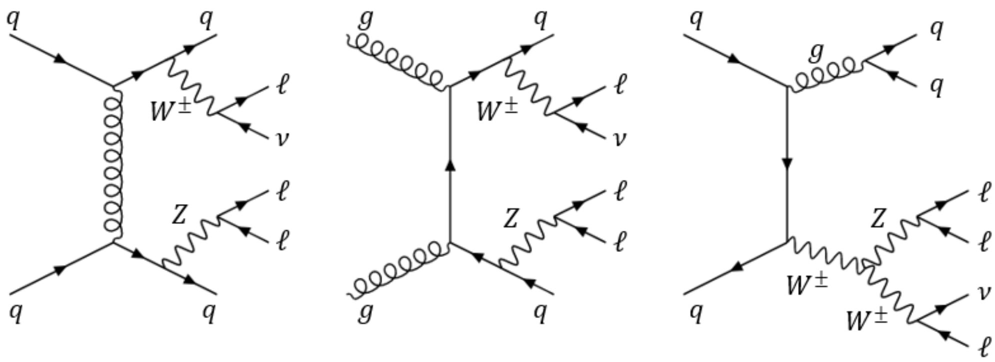


Figure 2.6: Example of WZjj QCD Feynman diagrams.

2.3 Polarization

Polarization is defined as the alignment of the spin of a particle with its momentum. It can be quantified by the helicity which is defined as the projection of the spin S of a particle onto its momentum $\vec{p}$:

$$h = S\,\frac{\vec{p}}{p}. \tag{2.38}$$

Depending on the spin of the particle different eigenvalues of the helicity are possible. A massless spin 1 particle can have helicities of $h = \pm 1$. This corresponds to a spin parallel or anti-parallel to the momentum. Together this is known as transverse polarization. Particles which are transversely polarized are denoted with an index T. For massive spin 1 particles like the $W^{\pm}$ and Z boson the helicity can additionally be 0. This means that the spin is perpendicular to the momentum. Particles with this polarization are called longitudinally polarized. They are denoted with an index L.

For massive particles the polarization is not Lorentz invariant and thus depends on the frame it is defined in. This can easily be seen when considering a boost into a frame in which the momentum changes direction. For massless particles which travel at the speed of light no such boost can be found and the helicity is Lorentz invariant.

2.3.1 Polarization in VBS

As mentioned in subsection 2.1.2 the longitudinal polarization is deeply linked to the mechanism of EWSB. The longitudinal polarization of the $W^{\pm}$ and Z bosons is introduced when this bosons get their masses. The massless Goldstone bosons that appear in the BEH mechanism are interpreted as the longitudinal polarization of those gauge bosons. Thus without the BEH mechanism they would not have a longitudinal polarization.

As shown in Figure 2.2 the VBS process includes a Feynman diagram with a Higgs boson contribution. In a SM without EWSB this channel does not appear. Looking at the scattering of longitudinally polarized bosons $W_L^\pm Z_L \to W_L^\pm Z_L$ the cross section of the VBS process without this diagram grows quadratically with the centre of mass energy. Consequently this process breaks unitarity which leads to unphysical predictions. By including the Higgs boson the increase of the cross section gets cancelled. This behaviour is shown with the blue lines in Figure 2.7 where the dashed one is the interaction without Higgs boson contribution and the full line the one with it.

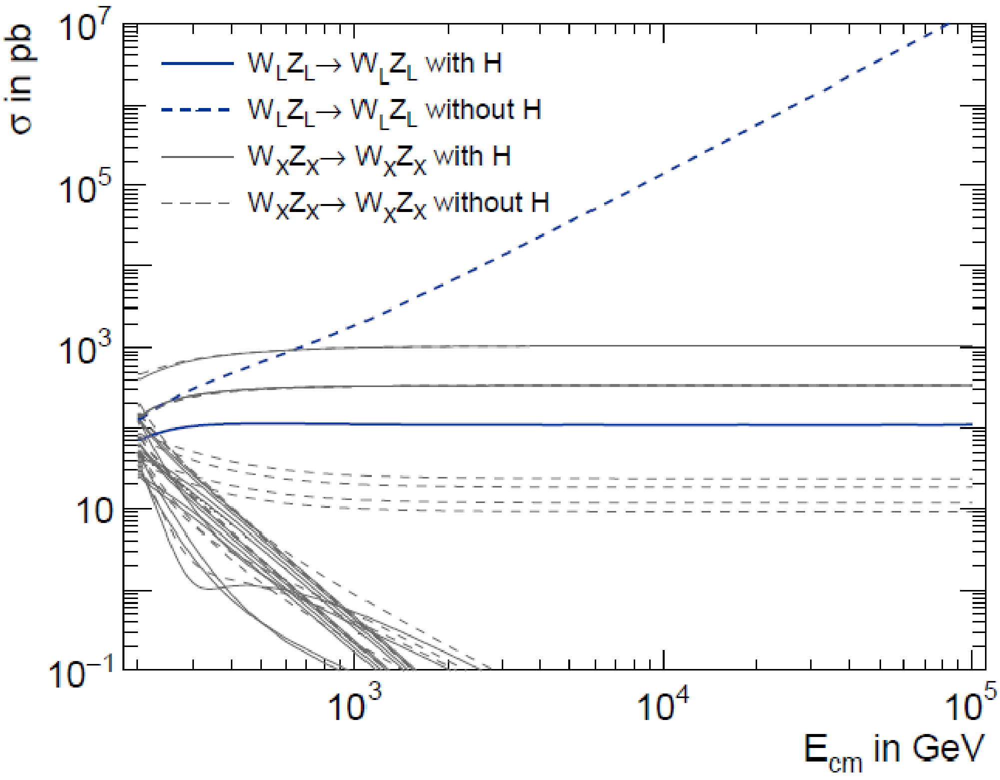


Figure 2.7: Dependence of the cross section of the scattering of a W and a Z boson on the centre of mass energy [3].

The scattering of bosons with other combinations of polarizations does not break unitarity independent of the Higgs boson contribution. In Figure 2.7 this is shown as the grey lines. This means that the study of the longitudinal polarization is of special interest since it provides the possibility to test details of the mechanism of EWSB.

2.3.2 Dependence of Variables on the Polarizations

To study the longitudinal polarization it is necessary to identify variables that are sensitive to the different polarizations.

The decay angle θ^* of the $W^\pm$ and the Z boson is a variable that is expected to show different behaviour for the different polarizations. It is defined as the angle between the momentum of a lepton in the rest frame of the boson and the momentum of the boson in the frame the polarization is defined in boosted into its rest frame. A schematic depiction is shown in Figure 2.8. For the $W^\pm$ boson the angle between the charged lepton and the $W^\pm$ boson is considered. For the Z boson the angle to the negatively charged lepton is used.

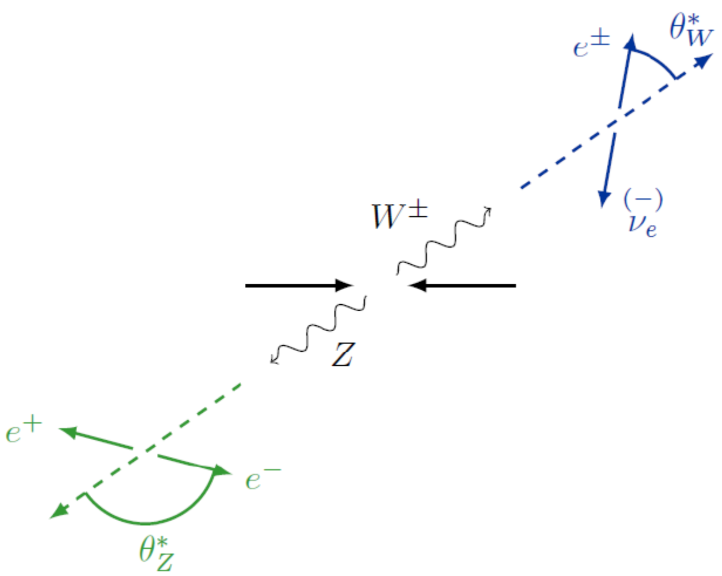


Figure 2.8: Visualization of the decay angle θ^*. The black lines are shown in the frame the polarization is defined in. The blue and green lines are depicted in the rest frame of the respective boson [3].

The dependence of this variable on the polarizations can be understood when considering the parity violating nature of the weak interaction. The W bosons couple only to left-handed particles and right-handed anti-particles. Looking as an example at the decay $W^+ \to e^+\nu_e$ this means that the positron has to be right-handed and the neutrino left-handed. If the W^+ is right-handed and considering spin conservation, the e^+ now has to fly in the same direction as the W^+ boson, so that the spin shows in the same direction as its momentum. For left-handed W^+ bosons on the other hand the e^+ flies in the opposite direction. This is visualized in Figure 2.9. For W^- bosons the opposite behaviour is expected. Longitudinal bosons are expected to have a decay angle of approximately $\theta^* \approx \frac{\pi}{2}$.
For the decay angle of the W and Z bosons analytical formulas can be found [12] as

$$\frac{1}{\sigma}\frac{d\sigma}{d\cos\theta^*_W} = \frac{3}{8}(1 \mp \cos\theta^*_W)^2 f_{-1} + \frac{3}{8}(1 \pm \cos\theta^*_W)^2 f_{+1} + \frac{3}{4}\sin^2\theta^*_W f_0 \quad \text{for } W^\pm. \tag{2.39}$$

and

$$\begin{aligned}\frac{1}{\sigma}\frac{d\sigma}{d\cos\theta^*_Z} &= \frac{3}{8}\left(1 + \cos^2\theta^*_Z - \frac{2(c_L^2 - c_R^2)}{(c_L^2 + c_R^2)}\right) f_{-1} \\ &+ \frac{3}{8}\left(1 + \cos^2\theta^*_Z + \frac{2(c_L^2 - c_R^2)}{(c_L^2 + c_R^2)}\right) f_{-1} + \frac{3}{4}\sin^2\theta^*_W f_0.\end{aligned} \tag{2.40}$$

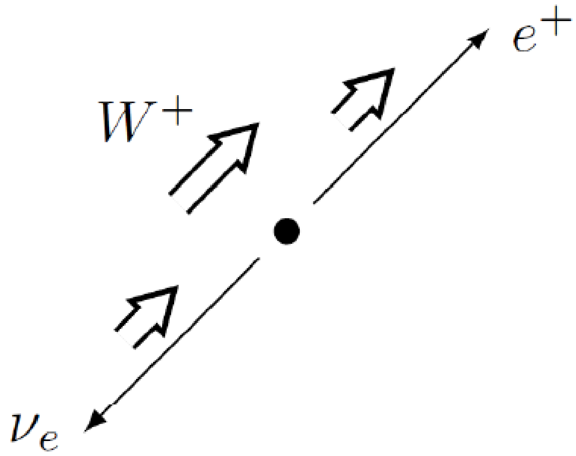


Figure 2.9: Visualization of spin and momenta in the decay of a "W^+ boson [3].

Here the f_i stand for the different helicity fractions and c_L and c_R for the left- and right-handed couplings of the fermions to the Z boson. The distributions of these formulas can be seen in Figure 2.10.

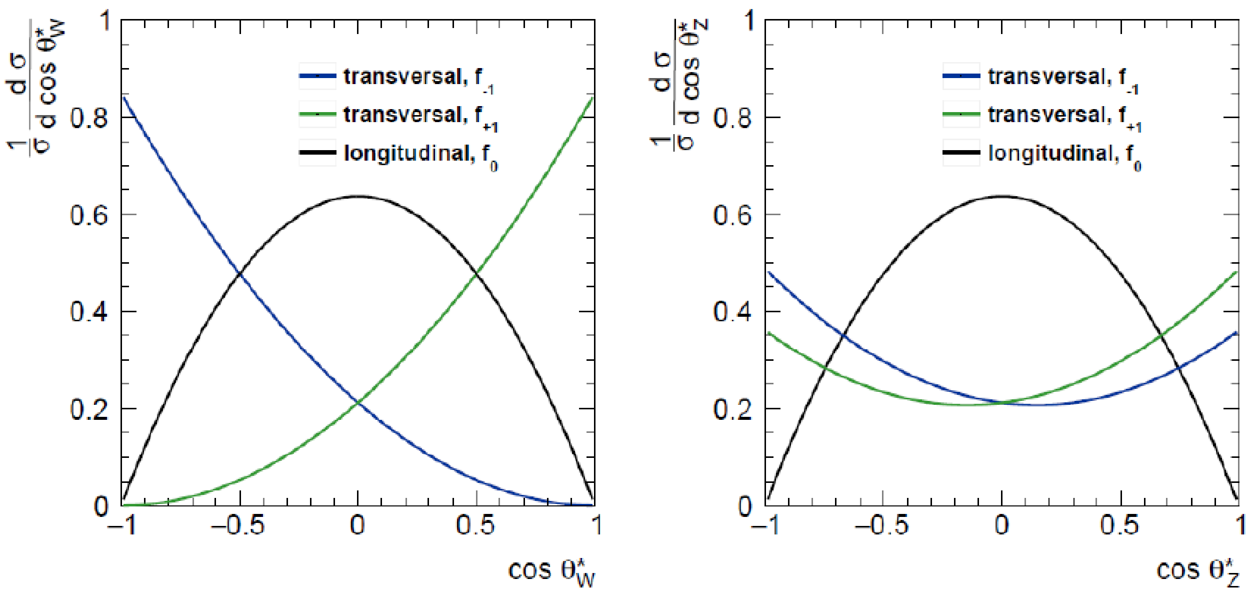


Figure 2.10: Distribution of the cross sections over the cosine of the decay angle θ^*. On the left side the distribution for a W^+ and on the right side for a Z boson is shown [3].

Similar to the leptons the W boson couples only to left-handed quarks and right-handed anti-quarks. Since the quarks are only slightly deflected their spin vector is not expected to change much. This means that the spin vector of the W boson should be orthogonal to the direction of movement of the quarks. If the W boson is emitted orthogonally it is expected to be transversely polarized and if emitted in forward or backward direction to be longitudinally polarized. This leads to a higher absolute value of the pseudorapidity $|\eta|$ and a lower absolute value of the transverse momentum $|p_T|$ of longitudinally polarized bosons. This is schematically shown in Figure 2.11. This connection only applies if the vector boson does not change helicity during the scattering. A study about the change of helicity during the scattering can be found in [3].

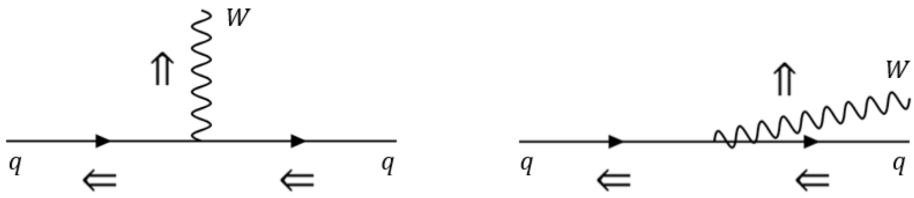


Figure 2.11: Schematic visualization of the spin and momenta in the scattering of a W boson from quarks.

3 Experiment

In this thesis only Monte-Carlo simulation and no actual data, taken from the Large Hadron Collider, are used. However in some parts of the thesis detector effects are taken into account in the simulations. In this chapter a brief overview over the LHC and the ATLAS detector is given.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [13] is the largest particle accelerator in the world. It has a 26.7 km long tunnel which is located between 45 to 170 m under the surface at the border of France and Switzerland. It is a circular proton-proton collider build to reach centre of mass collision energies up to $\sqrt{s} = 14$ TeV. However during Run 2 of the LHC only center of mass energies up to 13 TeV are reached. During Run 3 which started this year a centre of mass energy of 13.6 TeV will be achieved.
At the LHC the particles are collided at four interaction points with the four experiments ATLAS [14], CMS [15], ALICE [16] and LHCb [17]. While ATLAS and CMS are multi-purpose detectors, LHCb is specialised for B-physics and ALICE is designed for studies of strongly interacting matter and quark-gluon plasma with heavy iron collisions.
Protons are accelerated in several steps in pre-accelerators before they enter the main ring. In the LHC they do not form a continuous beam but are divided into bunches with up to 10^{11} protons per bunch. At the interaction points every 25 ns a crossing between two bunches with about 20 to 40 proton collisions takes place.
The event rate $\dot{N}$ at the LHC can be calculated [13] using

$$\dot{N}_{event} = L \cdot \sigma_{event}. \tag{3.1}$$

Here L denotes the machine luminosity and σ_{event} the cross section of the events that are studied. The luminosity depends on the beam parameters and can be written [13] as

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F \tag{3.2}$$

with N_b the number of particles per bunch, n_b the number of bunches per beam, f_{rev} the revolution frequency and γ_r the relativistic gamma factor. The the normalized transverse beam emittance ϵ_n and the amplitude modulation β^* are beam quality factors. F stands for the geometric luminosity reduction which accounts for the fact that the beams are crossed at an angle different from 0.

3.2 The ATLAS Detector

The simulation of detector effects used in this thesis are based on the setup of the ATLAS detector. The ATLAS detector is a multi-purpose detector with a hight of 25 meters and a lenght of 44 m. It consists of several sub-detectors which are specialized to measure different particles. A schematic view of the ATLAS detector can be seen in Figure 3.1. All information is taken from [14].

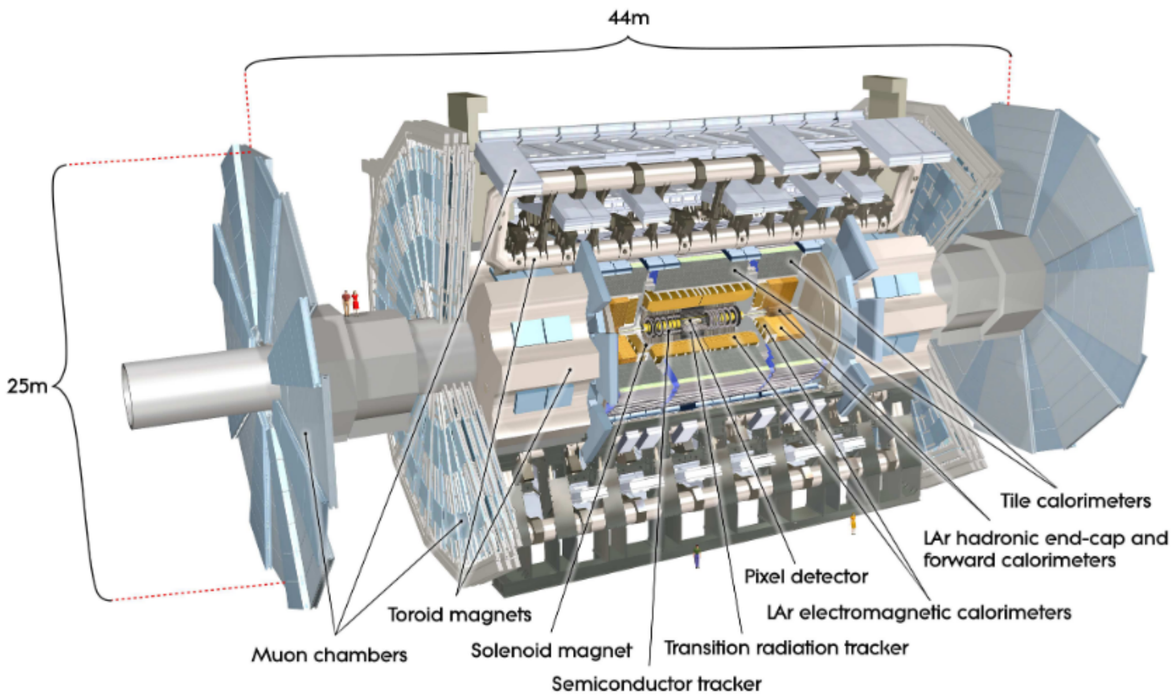


Figure 3.1: Cut away view of the ATLAS detector [14].

3.2.1 Coordinate System and Common Variables

For the ATLAS detector a right handed coordinate system is used. The x-axis points from the interaction point towards the center of the LHC, the y-axis points upward and the z-axis points along the beam axis. Instead of the Cartesian coordinates often ϕ, θ and R are used. Here the azimuthal angle ϕ is the angle between the projection onto the x-y plane and the x-axis and the polar angle θ is the angle to the z-axis. With R the smallest distance between the center of the coordinate system and the point is described.
In proton-proton collisions each of the partons in the proton carries a random fraction of the

total momentum of the proton. Thus the total momentum along the z-axis of the event is not known. This means that observables that are invariant under a boost along the z-axis are preferred.
Because of this a commonly used variable is the rapidity y:

$$y := \frac{1}{2} ln \left(\frac{E + p_z}{E - p_Z} \right). \tag{3.3}$$

Differences in the rapidity are invariant under such boosts. The rapidity is often replaced by the pseudorapidity η

$$\eta := -ln \left(\frac{\theta}{2} \right). \tag{3.4}$$

This variable has the advantage of an direct dependence on θ and is thus easier to calculate. It is only equal to the rapidity, if the mass of the particle can be neglected.
A variable that combines the two angles is the angular distance ΔR:

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}. \tag{3.5}$$

Another commonly used variable is the transverse momentum p_T:

$$p_T^2 = p_x^2 + p_y^2. \tag{3.6}$$

Since the proton is not expected to have a momentum in x and y direction the sum of the transverse momenta of all particles going out from the interaction must be zero. This can be used to define the missing transverse energy p_T^{miss} which is the difference of the sum of the transverse momenta to zero. In events with only one neutrino like WZ VBS process this can be used as an estimate of the transverse momentum of the neutrino which itself can not be measured by the ATLAS detector.

3.2.2 The Inner Detector

The Inner detector (ID) is used to precisely measure the tracks of charged particles. It consists of multiple tracking detectors which in combination are able to deal with the large track density in the detector while giving a high precision momentum measurement and vertex reconstruction. The ID is covered in a homogeneous magnetic field of 2 T which is generated by a central solenoid. The general layout of the ID is shown in Figure 3.2.
The Pixel and silicon microstrip (SCT) trackers located at small radii are precision tracking detectors . They cover a region $|\eta| < 2.5$. In the barrel region they are located on concentric cylinders around the beam axis while in the end cap region they are located on discs that are

perpendicular to the beam axis. The tracks normally cross three layers of the silicon pixel detectors. With them the highest highest granularity is achieved with a minimal pixel size of $50 \times 400\,\mu m^2$ in $R - \phi \times z$. After the pixel detector the tracks cross eight layers of the SCT detector.
These two detectors are surrounded by the Transition Radiation Tracker (TRT). This detector has a lower precision per point compared to the silicon detector but therefore is able to perform a larger number of measurements and has a longer measured track length. The TRT consists of 4 mm diameter straw tubes and is able to follow tracks in a region up to $|\eta| < 2.5$.

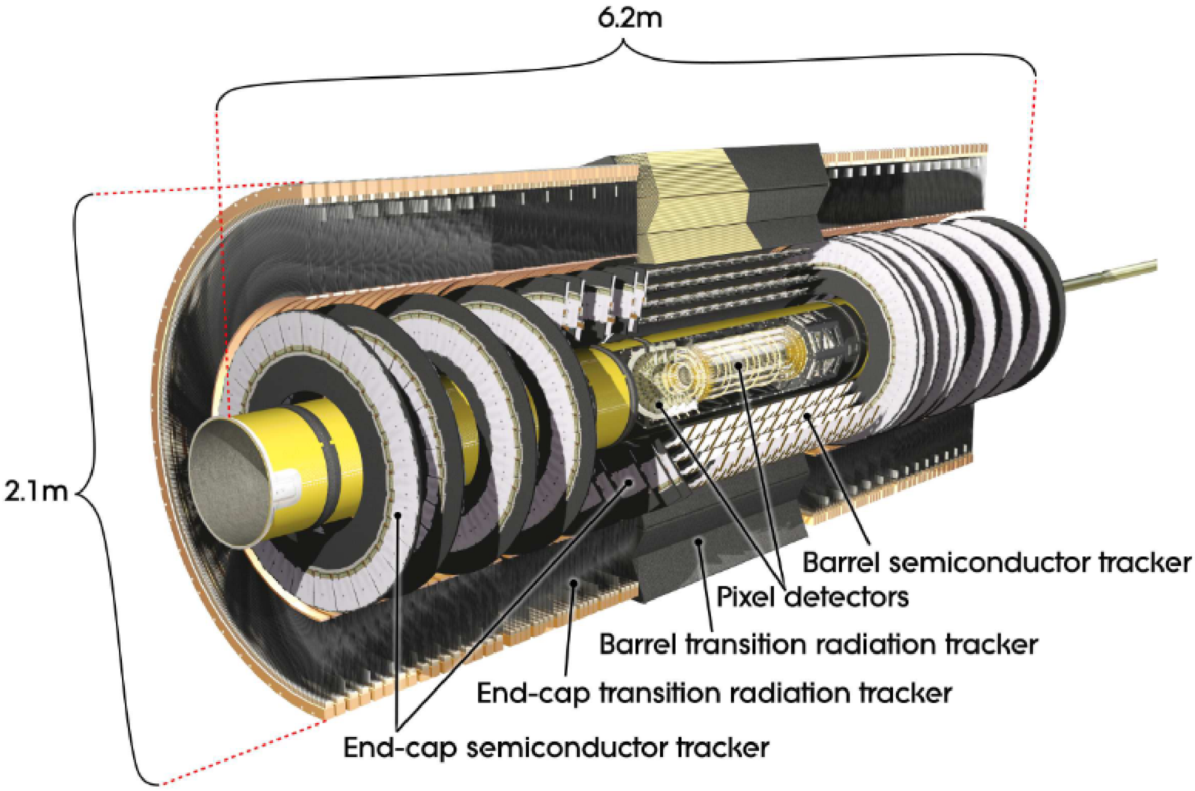


Figure 3.2: Cut away view of the ATLAS inner detector [14].

3.2.3 The Calorimeters

The calorimeters of the ATLAS detector are designed to measure the energy of the incoming particles. They cover a total range up to $|\eta| < 4.9$ and consists of different calorimeters specialised for different particles.
The lead-liquid argon (LAr) electromagnetic (EM) calorimeter is used to measure the energy of electrons and photons. It consists of a barrel part which covers a range $|\eta| < 1.475$ and two end-cap components with a coverage of $1.375 < |\eta| < 3.2$.
The EM calorimeter is surrounded by the hadronic calorimeters which are used for the energy measurements of hadrons and jets and are needed for the measurement of the missing transverse

energy. The Tile calorimeter covers a region $|\eta| < 1.0$ and has additionally two extended barrels at a range $0.8 < |\eta| < 1.7$. It uses steel as an absorber and scintillating tiles as the active material. The second calorimeter is the LAr hadronic end-cap calorimeter which covers a region $1.5 < |\eta| < 3.2$.
The third calorimeter is the LAr forward calorimeter with a coverage of $3.1 < |\eta| < 4.9$. It is build up from three modules per end-cap, where the first one is made of copper and deals with electromagnetic measurements while the other are made of tungsten and are used for hadronic reactions.
The layout of the calorimeters is shown in Figure 3.3.

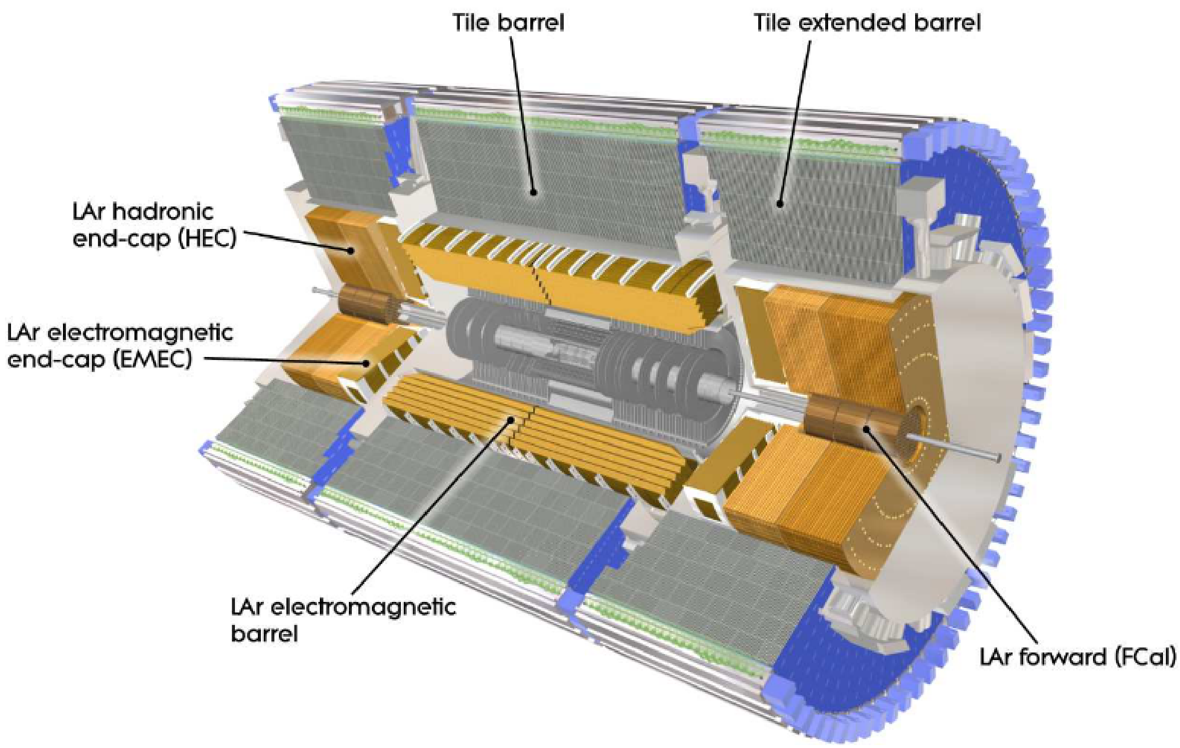


Figure 3.3: Cut away view of the ATLAS calorimeter system [14].

3.2.4 The Muon System

The last particles that have to be precisely measured by the ATLAS detector are muons. Similar to the inner detector the measurement of the muons is based on the deflection of their track in a magnetic field. The muon system consists of high-precision tracking chambers with a range $|\eta| < 2.7$ and a separate trigger with a range $|\eta| < 2.4$. For the most parts the tracking is done by Monitored Drift Tubes, but for the range $2 < |\eta| < 2.7$ Cathode Strip chambers are used. The tracking chamber uses Resistive Plate Chambers in the barrel region and Thin Gap Chambers in the end-cap regions. The muon system is depicted in Figure 3.4.

In general particles other than the muons get stopped in the calorimeters and do not reach the muon system. This leads to a very good identification of the muons.

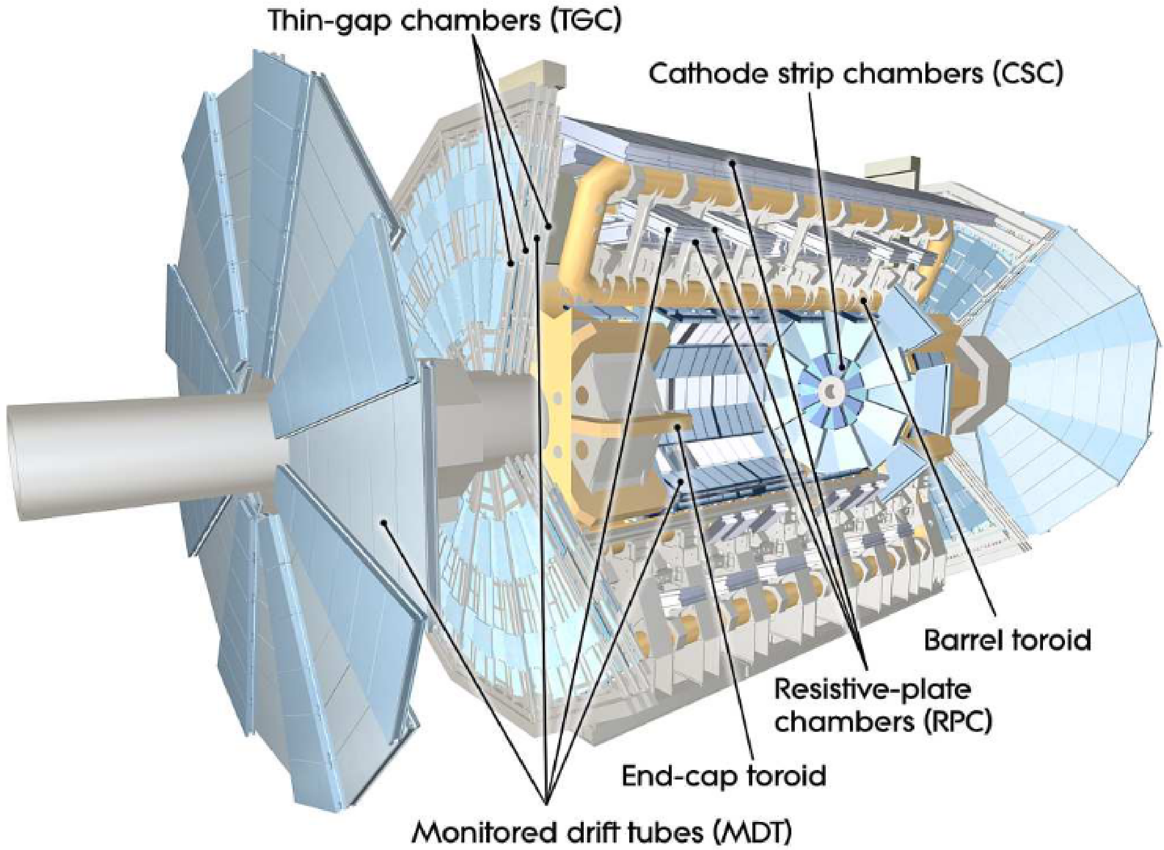


Figure 3.4: Cut away view of the ATLAS muon system [14].

4 Simulation of Events

For the analysis of the data taken at the LHC accurate simulations of the processes that take place during particle collisions are needed. Such simulations can be provided by Monte-Carlo generators. The following chapter gives a short description of the steps that are done to simulate events based on a given theory.

4.1 Basic Principal

The simulation of events is split into several phases that take place at different energy regimes and for that different approximations are used.
The first step of the simulation is the calculation of the hard process based on the matrix elements of the given process. For this usually fixed order perturbation theory is used. The squared matrix element is then integrated numerically over a given phase space. In order to deal with the high dimensionality and the complexity of the phase space for the integration Monte-Carlo Methods are used. For accelerators like the LHC, in which protons are collided, each of the partons inside the protons hold a fraction of the overall momentum of the proton. To account for this parton density functions (PDFs) are introduced which parametrize the possibility of a parton to have a specific fraction of the momentum of the proton. The PDFs are obtained from fits to data.
Parton showers are used to describe further QCD radiation from partons to account for higher order effects. For the coloured partons the transition into colourless hadrons need to be simulated. This hadronization happens at scales at which perturbation theory breaks down so that for this step phenomenological models are applied. Since many of the so simulated hadrons are unstable they have to be further decayed into hadrons that have large enough life times to be observed by the detector. Additionally at each step of the simulation QED bremsstrahlung has to be simulated.
Another effect that has to be included in the simulation are so called underlying events. With that the possible interactions of additional partons in the colliding protons, that not initiate the hard process, are described.
An overview of the different steps is given in Figure 4.1 for the example of a $t\bar{t}H$ process. Events that have passed these steps of the simulation are called truth level events. For the comparison to data also detector effects like the detector geometry and the response of the

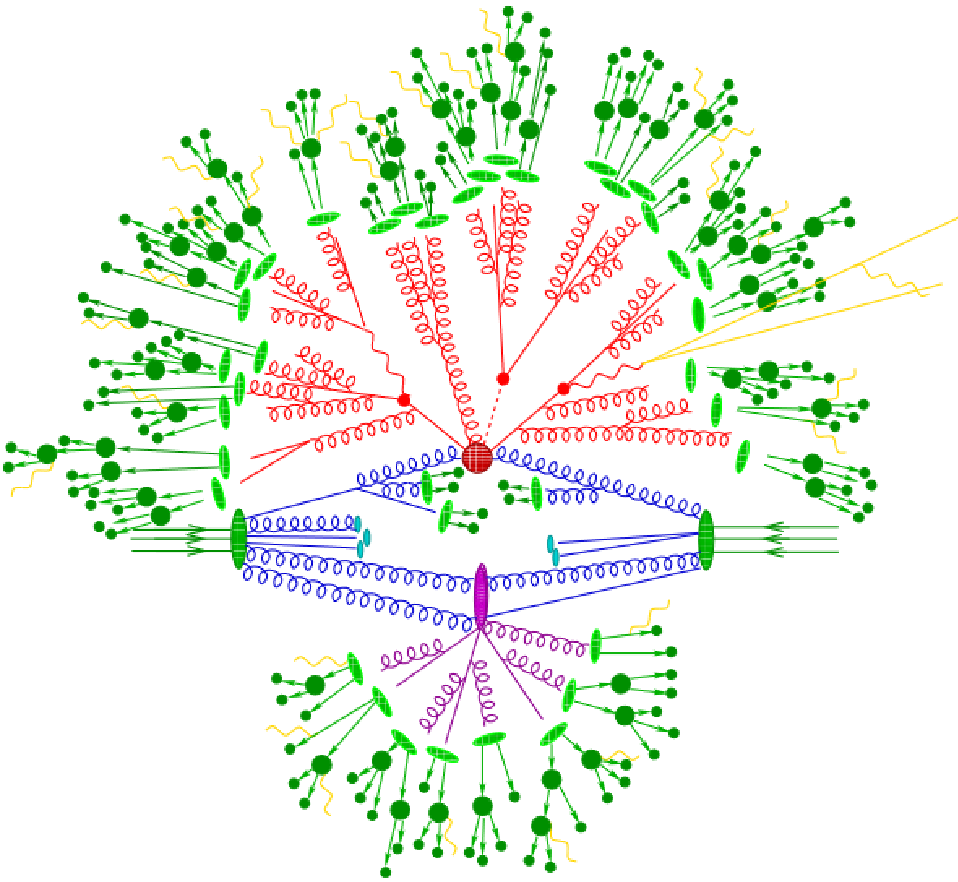


Figure 4.1: Pictorial representation of a $t\bar{t}H$ event as produced by an event generator. The hard interaction (big red blob) is followed by the decay of both top quarks and the Higgs boson (small red blobs). Additional hard QCD radiation is produced (red) and a secondary interaction takes place (purple blob) before the final-state partons hadronize (light green blobs) and hadrons decay (dark green blobs). Photon radiation occurs at any stage (yellow). [18]

detector material have to be simulated. This means that not longer the actual particles but their energy deposition in the detector are described for the events. In this step also energy depositions of particles that not originate from the main interaction are taken into account. As for real data, the detector output is then further processed in order to reconstruct the objects that entered the detector. After this step events are referred to be on reconstruction level or short reco level.

For the simulation of events at the LHC different Monte-Carlo event generators are available. A list of the officially produced samples that are used in this thesis together with their Monte-Carlo generators is given in Appendix G. This include the generators MadGraph5_aMC@NLO (Madgraph) [19], Pythia8 [20, 21], Herwig7 [22, 23] and Sherpa [18, 24].

4.2 Simulation of Polarized Events

For the analysis of the polarized WZ VBS channel samples with known helicity of the W and Z bosons are needed. A Monte-Carlo generator who offers the possibility to simulate samples with polarization information of the vector bosons is Madgraph [25]. The simulation of the polarized samples used in this analysis was done at leading order using Madgraph (version 2.9.5) interfaced with Phythia8 for the simulation of the parton shower, hadronization and underlying event. For the parton shower the local recoil dipole shower is used by setting `dipoleRecoil=on`. This was found to give a better description of the deep-inelastic scattering [26] and to be advantageous for the theoretical description of the VBS WZjj process [27].
As mentioned in section 2.3 the polarization depends on the frame in which it is defined in. Madgraph offers the possibility to specify the frame via the parameter `me_frame`. With this parameter the particles in whose rest frame the polarization is defined in are stated. For the polarization samples in this analysis the particles are accessed via the following numbers:

$$\underbrace{p}_{1}\underbrace{p}_{2} \rightarrow \underbrace{\ell}_{3}\underbrace{\nu}_{4}\underbrace{\ell}_{5}\underbrace{\ell}_{6}\underbrace{j}_{7}\underbrace{j}_{8}. \tag{4.1}$$

Here the p stands for the initial partons. For the majority of the samples used in this thesis the polarization was defined in the so called helicity frame which is the rest frame of the W and Z boson by setting `me_frame = 3,4,5,6`. A short study was done with samples in which the polarization was defined in the parton parton frame which corresponds to `me_frame = 1,2`.
For the simulation of the polarization different approximations have to be made. For the simulation of the polarization of the W and Z boson the generation process has to be split into the production and decay of the bosons. For that the decay chain syntax is used in this analysis which is further explained in [28, 29]. This means that only resonant diagrams can be included. While spin correlations and off shell effects are exactly taken into account, the virtuality of the decaying particle m_X^* is forced to be in the range

$$|m_V^* - m_V| \leq \texttt{bwcutoff}\Gamma_V. \tag{4.2}$$

Here m_V denotes the pole mass and Γ_V the width of the particle that gets decayed. The value of `bwcutoff` can be specified by the user. In this study the default value `bwcutoff = 15` is used. This imposes a loose on-shell condition.
The amplitude for the production and decay of a weak vector boson sums over the amplitudes with a the helicities λ for the intermediate vector boson $\mathcal{M}_\lambda^\mathcal{F}$ with $\mathcal{M}_\lambda^\mathcal{F} \sim \mathcal{M}_\mu^{\mathrm{prod}}\epsilon_\lambda^{\mu *}\epsilon_\lambda^{\mu}\mathcal{M}_\mu^{\mathrm{decay}}$:

$$\mathcal{M} \sim \sum_{\lambda=1}^{3} \mathcal{M}_\lambda^\mathcal{F}. \tag{4.3}$$

Here $\mathcal{M}_\mu^{\text{prod}}$ and $\mathcal{M}_\mu^{\text{decay}}$ are the matrix elements of the production and decay and the ϵ_λ^μ the polarization vectors. The cross section is proportional to the squared amplitude. Besides a sum over the squared amplitudes of the individual polarizations, this leads to interference terms between the polarizations:

$$|\mathcal{M}|^2 = \sum_\lambda |\mathcal{M}_\lambda^\mathcal{F}|^2 + \sum_{\lambda \neq \lambda'} \mathcal{M}_\lambda^{\mathcal{F}*} \mathcal{M}_{\lambda'}^\mathcal{F}. \tag{4.4}$$

These cancel only when the squared amplitude is integrated over the full azimuth angle ϕ of the decay products. In practice this is generally not possible. This leads to the exclusion of interferences between the polarizations as another approximation that is applied for the simulation of the polarizations.

Besides the polarized samples (pol) two other samples without polarization information were produced with Madgraph in order to see the effects of the approximations. For the full sample only the initial and final state are specified so that none of the approximations above are applied. For the unpolarized sample (unpol) the production and decay of the W and Z boson are simulated via the decay chain syntax but without polarization information so that interferences are included in this sample:

full: $pp \to \ell\nu\ell\ell jj$

unpol: $pp \to WZjj \to \ell\nu\ell\ell jj$

pol: $pp \to W_X Z_X jj \to \ell\nu\ell\ell jj \qquad (X = L, T)$

For all the samples generator level cuts are applied. They are summarized in Table 4.1.

Table 4.1: Generator level cuts used for the simulation of the polarized, unpolarized and full sample.

variable	cut
$p_T(j)$ [GeV]	> 15.0
$p_T(b)$ [GeV]	> 15.0
$p_T(\ell)$ [GeV]	> 4.0
$p_T(\text{heavy})$ [GeV]	> 2.0
$\lvert\eta(j)\rvert$	< 5.5
$\lvert\eta(\ell)\rvert$	< 5.0
$\Delta R(jj)$	> 0.4
$\Delta R(\ell\ell)$	> 0.2
$\Delta R(j\ell)$	> 0.2
$\Delta R(\gamma\ell)$	> 0.1
$m_{\ell\ell}$ [GeV]	> 4.0

5 Study of the Simulated Events

5.1 Study on Truth Level

As a first step for this thesis a validation of the polarized samples produced by Madgraph was performed. For that polarized samples as well as unpolarized samples and full samples as described in section 4.2 were produced locally at truth level for a centre of mass energy of $\sqrt{s} = 13$ TeV. A summary of all the locally produced samples and more details about their production can be found in section B.1. If not stated otherwise in this chapter only the samples in Table B.1a are used. For these samples only electrons and muons in the final state are allowed and the polarization is defined in the helicity frame. The polarized samples are further grouped according to the four different polarization combinations of the vector bosons denoted as WLZL, WLZT, WTZL and WTZT, where L stands for the longitudinal and T for the transverse polarization of the respective boson. A X behind one of the bosons denotes an arbitrary polarization of this boson. The two samples where the bosons have a different polarization are also referred to as mixed samples.
The samples were further processed and analysed using the Dresden frameworks ELCore and CAF. The analysis was done at a luminosity of 139 fb^{-1} which corresponds to the integrated luminosity achieved during Run 2 at the LHC.

5.1.1 Definition of Variables

For the variables of the jets, always the two jets with the highest momentum are used. They are referred to as the tagging jets. For the variables of the leptons, the W and Z leptons are chosen as follows: All same flavour opposite charge truth lepton pairs are taken as potential Z-pairs candidates. For every Z-pair candidate the lepton with the highest p_T and a flavour that fits to the truth neutrino is taken as a W lepton candidate from the remaining leptons. The product of the Breit-Wigner probabilities of the Z pair candidate and the associated pair of W-lepton and neutrino is calculated using the PDG [30] values of the W and Z mass and width. The Z leptons and the W lepton are chosen from the pairs with the highest value of this product. For the Z leptons the one with the higher p_T is referred to as the leading Z lepton $\ell_{Z,1}$ and the one with the lower p_T as the subleading Z lepton $\ell_{Z,2}$.
In the following some of the variables used in this analysis are defined:

Transverse mass of the W boson $m_T(W)$:

$$m_T(W) = \sqrt{2 \cdot p_T(\nu) \cdot p_T(\ell_W)\,(1 - \cos\Delta\varphi(\ell_W, \nu))} \tag{5.1}$$

Transverse mass of the diboson system $m_T(WZ)$:

$$m_T(WZ) = \sqrt{2 \cdot p_T(\nu) \cdot \sum_\ell \vec{p}_T \left(1 - \cos\Delta\varphi\left(\sum_\ell \vec{p}_T, \nu\right)\right)}.$$

Transverse momentum balance $p_T^{\text{bal}}(t_1, ..., t_N)$ of particles t_i :

$$p_T^{\text{bal}}(t_1, ..., t_N) = \frac{\sqrt{\sum_t (p_{x,t}) + \sum_t (p_{y,t})}}{\sum_t p_{T,t}}. \tag{5.2}$$

Centrality centrality$(x_1, ..., x_N)$ of particles x_i, with rapidities y_{x_i} of the particles and y_{j1} and y_{j2} of the tagging jets:

$$\text{centrality}(x_1, ..., x_N) = \max_{x_i} \left(\frac{|y_x - 0.5 \cdot (y_{j1} + y_{j2})|}{y_{j1} - y_{j2}}\right) \tag{5.3}$$

Scalar sum SW and SZ of the bosons, where MET denotes the missing transverse energy:

$$SW = p_T(l_W) + p_T(MET) \tag{5.4}$$

$$SZ = p_T(l_{Z,1}) + p_T(l_{Z,2}). \tag{5.5}$$

If not stated otherwise the truth neutrino momentum is used for the reconstruction of variables of the W boson.

5.1.2 Phase Space Definition

The phase space used for the truth analysis is based on the fiducial WZjj-EW phase space used in the previous WZ VBS study [31]. The phase space definition is shown in Table 5.1. In this analysis this phase space is referred to as the truth phase space. At least three leptons, one neutrino and two jets are required. Some basic kinematic cuts are applied for the leptons and jets. The cuts on the transverse mass of the W boson $m_T(W)$ and the invariant mass of the Z boson m_Z are applied to increase the contribution of resonant diagrams. To enhance the WZ VBS contribution a cut on the invariant mass of the tagging jets is applied.

Table 5.1: Phase space definition for the truth level analysis. Cuts above the line are applied in ELCore and cuts below the line in CAF.

Variable	Cut
$p_T(j)$ [GeV]	> 25
$\|\eta(j)\|$	< 4.5
$\Delta R(j, \ell)$	> 0.3
m_Z [GeV]	$\|m_Z - m_Z^{PDG}\| < 25$
$p_T(\ell_Z)$, $p_T(\ell_W)$ [GeV]	> 15, > 20
$\|\eta(\ell)\|$	< 2.5
$\Delta R(\ell_{Z,1}, \ell_{Z,2})$, $\Delta R(\ell_Z, \ell_W)$	> 0.2, > 0.3
Jet multiplicity	> 1.5
p_T(tagging jet)[GeV]	> 40
m_T^W [GeV]	> 30
m_Z [GeV]	$\|m_Z - m_Z^{PDG}\| < 10$
m_{jj} [GeV]	> 500

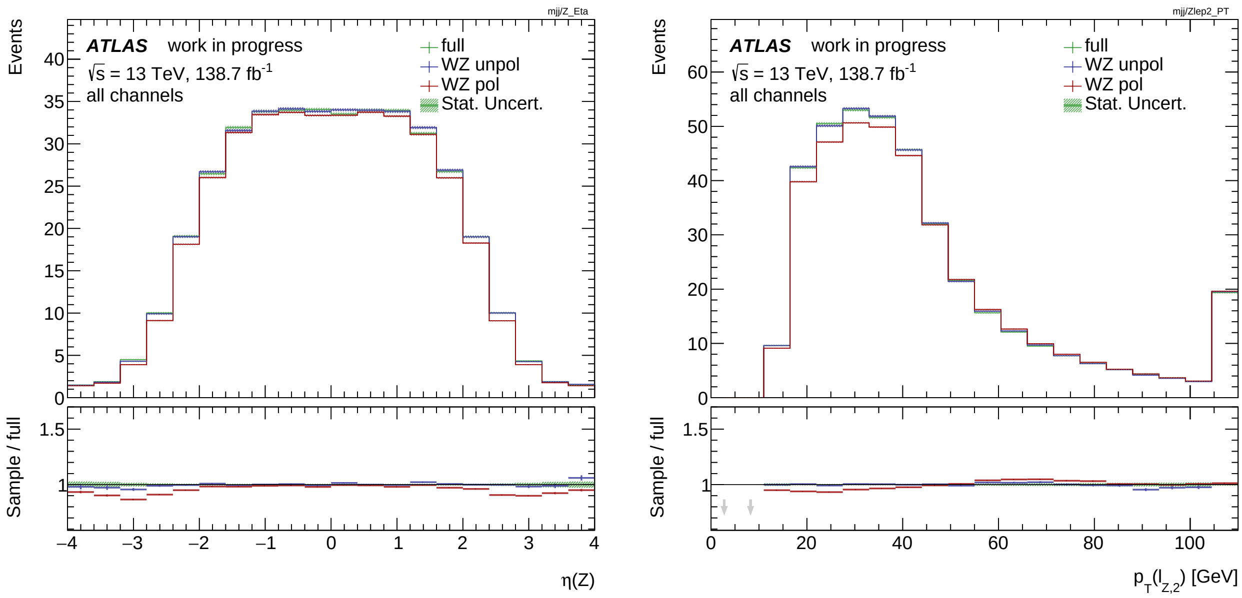
5.1.3 Effects of the Approximations

In order to investigate the effects of the approximations made for the simulation of the polarized sample as explained in section 4.2, the sum of the polarized samples was compared to the unpolarized and the full sample. The comparison of the full to the unpolarized sample shows the effects of the non-resonant diagrams and the off-shell effects while with the comparison of the polarized to the unpolarized sample the interference effects can be investigated.
In Table 5.2 the yields and ratios of the three samples are shown. A good agreement between the full and the unpolarized samples is found. The overall effects of non-resonant diagrams are very small with approximately 0.15 %. The overall size of the interference effects is with about 2.5 % larger but still in a magnitude that allow a reasonable study of the polarized samples.
For a more detailed study of the effects of the approximations the distributions of the three samples were compared for different variables. As an example the distributions for $\eta(Z)$ and $p_T(\ell_{Z,2})$ are shown in Figure 5.1. Only very small differences in the distributions of the full and

Table 5.2: Yields and ratios of the full, the unpol and the pol sample in the truth phase space.

	all
full	393.4 $\pm$ 0.6
WZ unpol	394.0 $\pm$ 0.6
WZ pol	384.04 $\pm$ 0.34
WZ unpol / full	1.0015 $\pm$ 0.0022
WZ pol / WZ unpol	0.9747 $\pm$ 0.0016
WZ pol / full	0.9762 $\pm$ 0.0018

the unpolarized sample are visible. For the two shown variables the shape of the differences is in some areas similar to the difference between the full and the polarized samples meaning that the non-resonant diagrams and off-shell effects may have an influence on the shape of these distributions. However the overall size of these effects is very small and the other variables that were looked at show a very good agreement for the shapes of the distributions of the full and the unpolarized samples. Therefore, in general, the differences between the full and the polarized sample can be attributed to the interference effects. They are studied in more detail in subsection 5.1.5.



(a) (b)

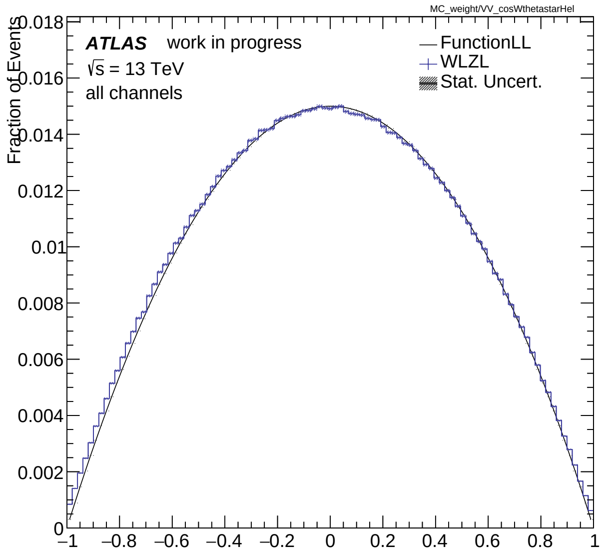
Figure 5.1: Distributions of the full, unpolarized and polarized sample for (a) the pseudo-rapidity of the Z boson $\eta(Z)$ and (b) the transverse monemtum of the subleading Z boson $p_T(l_{Z,2})$.

5.1.4 Comparison to the Analytical Function

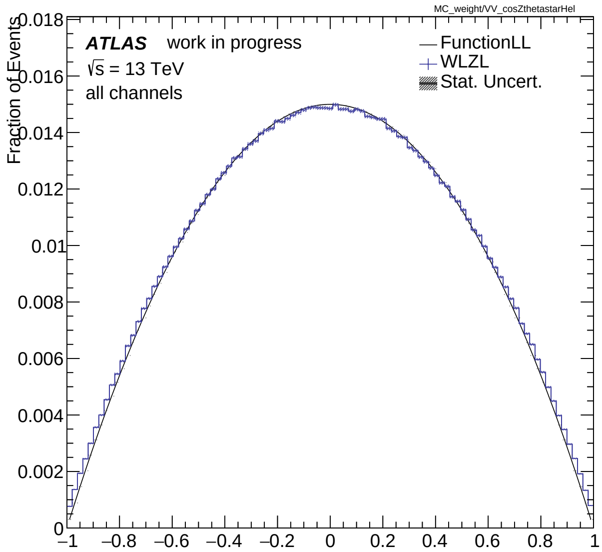
As explained in subsection 2.3.2 an analytical function is available for the decay angle of the W and Z boson . This can be used to further validate the polarized sample by comparing the decay angle distribution to this function. The analytical formula holds only true in a phase space without cuts. To account for that the decay angle distribution was plotted without applying the cuts mentioned in Table 5.1. Since for the generation of the samples generator cuts were already applied (see Table 4.1) some basic cuts are still included when plotting the the decay angle distribution.

For the closure test only the samples in which both bosons are longitudinally polarized were used . Thus only the last terms of Equation 2.39 are plotted. The parameter of the function was adjusted by hand to fit to the distribution.

The result are shown in Figure 5.2. It can be seen that an overall good closure is found. However some small deviations can be seen especially for larger absolute values of the cosine of the decay angle. There the values of the distribution are a bit larger than expected by the function and do not fall completely to 0 for $\cos\theta^*(W/Z) = \pm 1$.



(a) $\cos\theta^*(W)$



(b) $\cos\theta^*(Z)$

Figure 5.2: Normalized distribution of the cosine of the decay angle $\cos\theta^*$ of (a) the W boson and (b) the Z boson for the longitudinal longitudinal sample and the analytical function.

A reason for the deviations could be the cuts applied at generator level or additional emissions during the simulation of the parton shower. The latter could be further investigated by using matrix element particles, which are the particles that originate directly from the hard process.

5.1.5 Study of the Different Polarizations

For polarization studies it is important to have variables that are well modelled and show sensitivity to the different polarizations. To investigate this a study of the distributions of various variables was done. The distributions for the four polarization combinations were compared to each other and to the full sample to see if the variables that are considered for the analysis show unexpected behaviour and if shape differences between the different polarization combinations are visible. Furthermore the influence of the interference effects are studied in more detail.

In Table 5.3 the yields of the four polarized samples in the truth phase space are shown. Furthermore the resulting fractions of the different polarization combinations are given. It can be seen that the amount of longitudinally longitudinally (LL) polarized events is very small with only about 8%. The largest fraction has the sample with two transversely polarized bosons with about 53%. Even if the polarization of just of one of the bosons is considered the longitudinal polarization accounts for the W boson only for about 25% and for the Z boson only for about 29% of the events. This small fraction makes it difficult to separate the longitudinal polarization from the others.

In Figure 5.3 to Figure 5.7 the histograms of some of the investigated variables are shown. For each variable two plots are given. In the one at the top the stack of the four polarization samples is visible. This enables one to see the overall impact of the different polarizations and to compare the stack to the full sample. In the subplot the ratio of the stack to the full sample is given as the blue line which shows the interference effects. Additionaly the ratio of the sum of WTZT, WTZL and WLZT to the full sample is shown as the green line. Here the difference between the blue and the green line is the ratio of the WLZL sample to the full sample. This should ideally be large in comparison to the interference effects. In order to compare the shapes of the distributions of the different polarizations the plots at the bottom show the normalized distributions of the different polarizations as well as of the sum of the polarizations. Furthermore the shape of the interference effects are shown as the difference between the full and the sum of the polarized samples.

Since it is the goal to extract the longitudinal longitudinal polarization ideally the interference effects should be small in comparison to the WLZL sample and should have a different shape compared to this sample. Otherwise when doing a fit it would not be clear whether the fit extracts the actual polarization or if it sees the interference effects. Similar when using such a variable for BDTs, the BDT would than classify interference effects similar to events with longitudinal longitudinal polarization.

In Figure 5.3 the variables $p_T^{\mathrm{bal}}(jj)$ the transverse momentum of the diboson system $p_T(WZ)$ and the absolute value of the azimuthal angle difference of the two jets $|\Delta\phi(jj)|$ are shown as examples of variables that show some sensitivity to the longitudinal longitudinal polarization. It can be seen that for $p_T^{\mathrm{bal}}(jj)$ and $p_T(WZ)$ WLZL tends to be more at lower values than the other polarizations. However most LL events still lay in the region where also most events of the other polarizations appear. For these variables the interference effects have a similar shape as the longitudinal longitudinal polarization. However it can also be seen that the interference effects are small in general, with less than 10 % in all bins. Therefore these variables are still considered valuable for further studies in this thesis. For $|\Delta\phi(jj)|$ the longitudinally longitudinally polarized events appear mainly at large values of this variable. Only very small interference effects are visible. While the amount of the other polarizations at smaller values is larger than for the LL polarization, the maximum of their distributions lays at the same value as for the LL distribution.

In Figure 5.4 distributions of the decay angle of the W boson and the transverse momentum $p_T(W)$ and pseudorapidity $\eta(W)$ of the W are shown. These variables are expected to be sensitive to the polarization of the W boson as explained in subsection 2.3.2. It can be seen

Table 5.3: Yields of the full sample, the sum of all polarization samples and the individual samples of the four polarization combinations as well as the rations of the different polarizations to the sum of the polarizations in the truth phase space.

full	393.4 ± 0.6
sumPol	384.04 ± 0.34
WTZT	203.85 ± 0.30
WTZL	83.84 ± 0.12
WLZT	66.94 ± 0.10
WLZL	29.41 ± 0.04
WTZT / sumPol	0.5308 ± 0.0012
WTZL / sumPol	0.2183 ± 0.0008
WLZT / sumPol	0.1743 ± 0.0007
WLZL / sumPol	0.0766 ± 0.0004
WTZX / sumPol	0.7491 ± 0.0012
WLZX / sumPol	0.2509 ± 0.0007
WXZT / sumPol	0.7051 ± 0.0012
WXZL / sumPol	0.2949 ± 0.0008

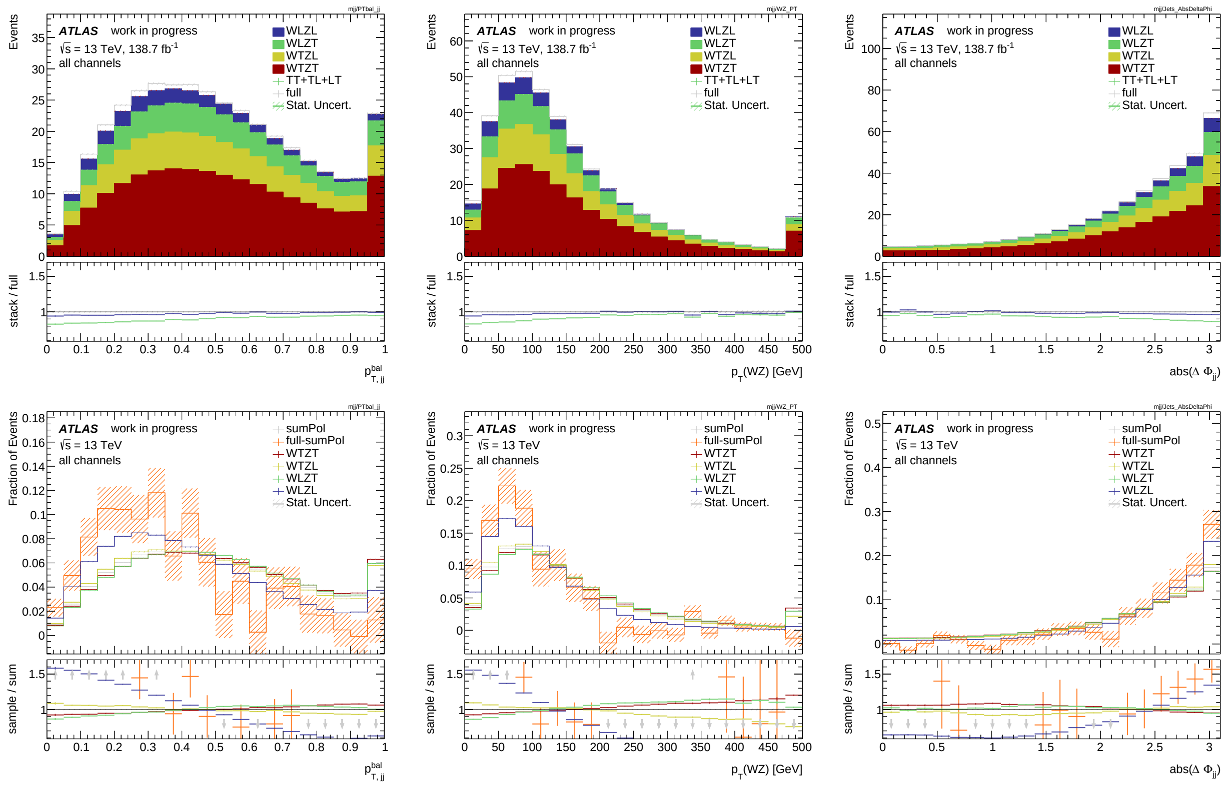
that this is the case for all three variables. Especially for the decay angle distribution shape differences between longitudinally and transversely polarized W bosons are clearly visible. While interference effects are visible in particular for $\cos\theta^*$ close to 1, this is a region with only a small LL contribution and the shape of the interference effects is different from the shape of any of the polarization sample. As expected for $p_T(W)$ the distributions where the W boson is longitudinally polarized are shifted to smaller values. For $\eta(W)$ the WLZL sample tends, as predicted, to larger and the WTZT sample to smaller absolute values. For mixed samples however no clear distinction between samples with longitudinally and samples with transversely polarized W boson can be found. For the transverse momentum $p_T(\ell_W)$ and pseudorapidity $\eta(\ell_W)$ of the W lepton a similar behaviour to that of the W boson can be found as shown in Figure 5.5. However, the distinction between the polarizations is less pronounced here. The missing transverse energy $p_T(MET)$ also shows a similar behaviour to the transverse momentum of the W boson. For all these variables only small interference effects are visible.

In Figure 5.6 three variables that show some sensitivity to the polarization of the Z boson can be seen. As expected for the decay angle distribution of the Z boson strong shape differences between samples with longitudinally and transversely polarized Z boson are visible. Almost no interference effects appear. Similar to the W boson events with longitudinal polarization of the Z boson tend to be at smaller values of the transverse momentum $p_T(Z)$ of the Z boson. While the interference effects are relatively small with less than 10% in all bins, their maximum in the area with a lot longitudinally polarized Z bosons. For the transverse momentum $p_T(\ell_{Z,2})$

of the subleading Z lepton events with longitudinally polarized Z lepton tend to be at larger values. For this variable the interference effects can get as large or larger than the longitudinal longitudinal polarization in regions with a considerable amount of this polarization. The shape is however more similar to the shapes of the distributions of transversely polarized Z bosons. For this variable also negative interference effects are visible in the tail of the distribution.
Variables for which larger interference effects are found are for example the transverse mass of the diboson system $m_T(WZ)$, the transverse momentum of the leading Z lepton $p_T(\ell_{Z,1})$ and $|\Delta\phi(\ell_{Z,1}, \ell_{Z,2})|$ of the leading and subleading Z boson. The distributions for these variables are shown in Figure 5.7. It can be seen that for $m_T(WZ)$ large interference effects appear in the area of the maximum of this distribution and are in the same magnitude or larger than the contribution of the WLZL sample. In the left tail negative interference effects of up to 40% are visible. Thus despite this variable showing shape differences for the different polarizations this variable was not used in the further analysis. A similar behaviour can be seen for $p_T(\ell_{Z,1})$ which was also excluded from further studies of the polarization in this thesis. In Figure 5.7c large differences between the full and the stacked polarized samples can be seen for $|\Delta\phi(\ell_{Z,1}, \ell Z, 2)|$ which makes this variable not suitable for the further analysis.
A first rough condition for which variables to use for the further analysis was made by requiring that the interference effects are less than 10% for all bins. With that the three variables mentioned above are excluded. Other variables that do not fulfil the conditions and are thus not used for further polarization studies are the pseudorapidity of the Z boson $\eta(Z)$, the scalar sum of the W and Z boson SW and SZ and the absolute value of the angular distance of the leading and subleading Z lepton $|\Delta R(\ell_{Z,1}, \ell_{Z,2})|$. Histograms of these variables can be found in Appendix C. This approach looks only bin for bin at the interference effects. The shape and comparison to the amount of the longitudinal polarization is not considered. Thus variables like the transverse momentum of the subleading Z boson or $|\Delta\phi(\ell_{Z1}, \ell_W)|$ (see Appendix C) are kept despite having interference effects as large as the LL polarized sample in some regions. For a more advanced rejection those effects could be taken into account.
The distributions of other variables that were used during this thesis are given in Appendix C. While some shape differences between the polarizations are visible it can already be seen that no variable is able to differentiate the longitudinal longitudinal polarization on its own. Thus for the further analysis multivariate techniques are used to enhance the separation of the longitudinal polarization.

5.1.6 Comparison of the Parton Parton and Helicity Frame

The helicity frame in which the polarization study was done so far seems as the rest frame of the W and Z boson intuitive for the definition of the polarization. Nevertheless it can be interesting to study the polarization in another frame. For that a brief study was done using samples in which the polarization is defined in the parton parton rest frame. For that



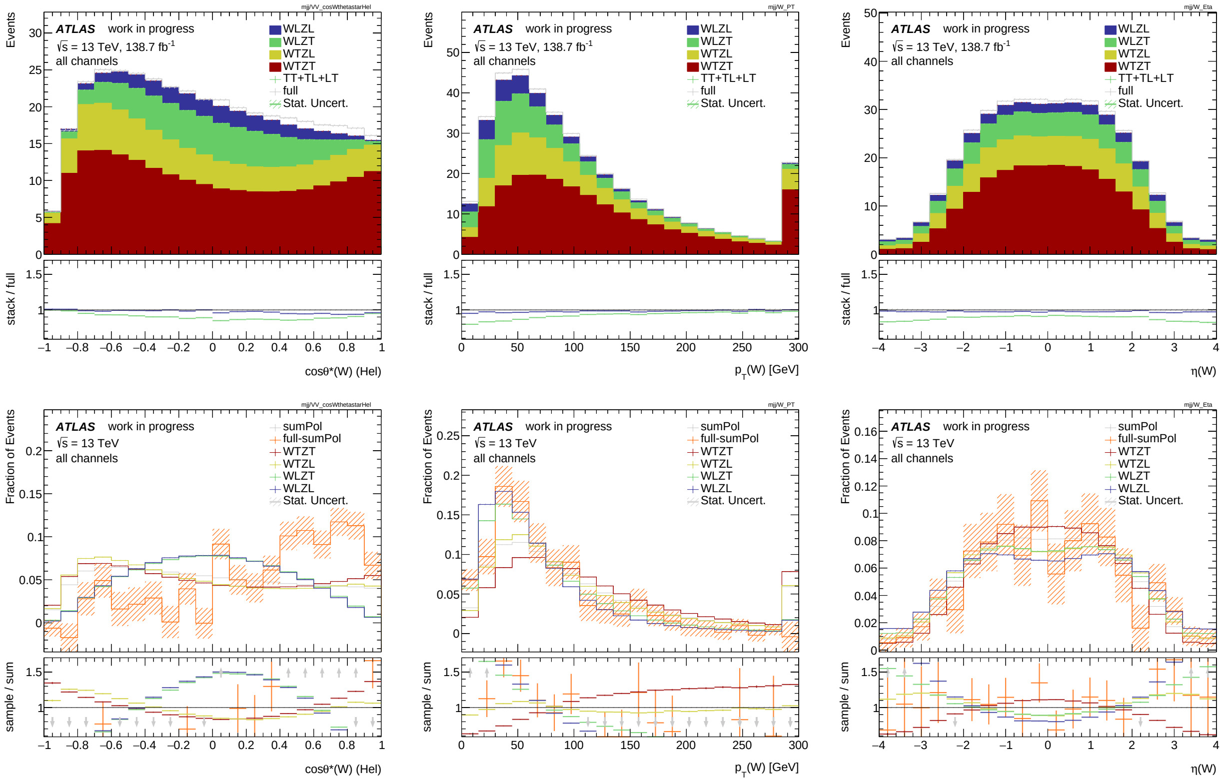
(a) (b) (c)

Figure 5.3: Distributions of (a) $p_{T,jj}^{\text{bal}}$ of the two jets (b) the transverse momentum of the diboson system $p_T(WZ)$ and (c) the absolute value of the difference of the azimuthal angles $|\Delta\phi(jj)|$ of the two jets. In the plots on the top the full sample and in the stack the four polarization combinations are shown. In the subplot the ratio of the stack to the full sample is shown as the blue line and the the ratio of the sum of WTZT, WTZL and WLZT to the full sample as the green line. In the plots at the bottom the normalized distributions of the four polarized samples and the sum of the polarized samples is plotted. Additionally the normalized distribution of the difference of the full sample to the sum of the polarized samples is shown. In the subplot the ratios of the individual samples to the sum of the polarization is given.

the samples with `me_frame = 1,2` were used [1]. They are summarized in Table B.2. While the parton parton frame can be easily defined in Madgraph it has the disadvantage that it is difficult to reconstruct experimentally since the kinematics of the partons are unknown.

In Table 5.4 the yields and fractions of the four polarization combinations are compared between the helicity and the parton parton frame. It can be seen that the fraction of LL polarized events is with only about 5% in the parton parton frame smaller than in the helicity frame with about 7.6%. The fractions of the samples in which only one of the bosons is longitudinally polarized are also smaller in the parton parton frame than in the helicity frame.

[1] Madgraph offers another possibility to define the parton parton frame by setting `me_frame = 3,4,5,6,7,8`. Samples defined with these two definitions should be identical within statistics. However differences between these samples were found. They are further discussed in Appendix D. The differences might be caused by the specific setup used to generate the samples.

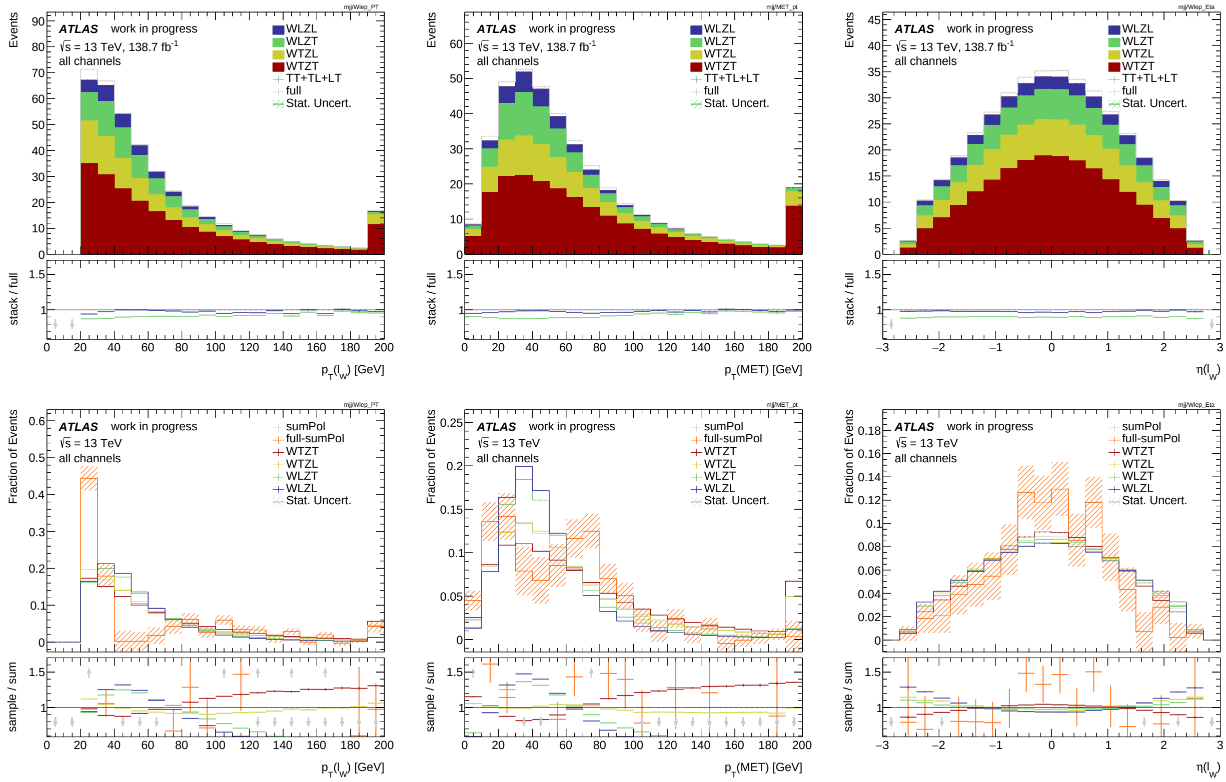


(a) (b) (c)

Figure 5.4: Distributions of (a) the cosine of the decay angle of the W boson $\cos\theta^*(W)$, (b) the transverse momentum of the W boson $p_T(W)$ and (c) the pseudorapidity of the W boson $\eta(W)$. In the plots on the top the full sample and in the stack the four polarization combinations are shown. In the subplot the ratio of the stack to the full sample is shown as the blue line and the the ratio of the sum of WTZT, WTZL and WLZT to the full sample as the green line. In the plots at the bottom the normalized distributions of the four polarized samples and the sum of the polarized samples is plotted. Additionally the normalized distribution of the difference of the full sample to the sum of the polarized samples is shown. In the subplot the ratios of the individual samples to the sum of the polarization is given.

The transversely transversely polarized events account for about 59% of the total events in the parton parton frame compared to 53% in the helicity frame.

Since only a small number of events was generated for the samples with the polarization defined in the parton parton frame, this samples show large statistical fluctuations. This make only a rough comparison between the two frames possible. However while the distributions of many variables look similar for the two reference frames, for some variables differences can be seen. One of them is the cosine of the angle of the W boson to the beam-axis in the WZ rest frame $\cos\theta(W)$. In the helicity frame events with longitudinally polarized Z boson appear mainly for high values of this variable meaning the W and Z boson are mainly scattered forward or backwards. For $\cos\theta \approx 0$ almost no events with longitudinally polarized Z boson appear. While this tendency is still visible for this angle for samples with polarization defined in the

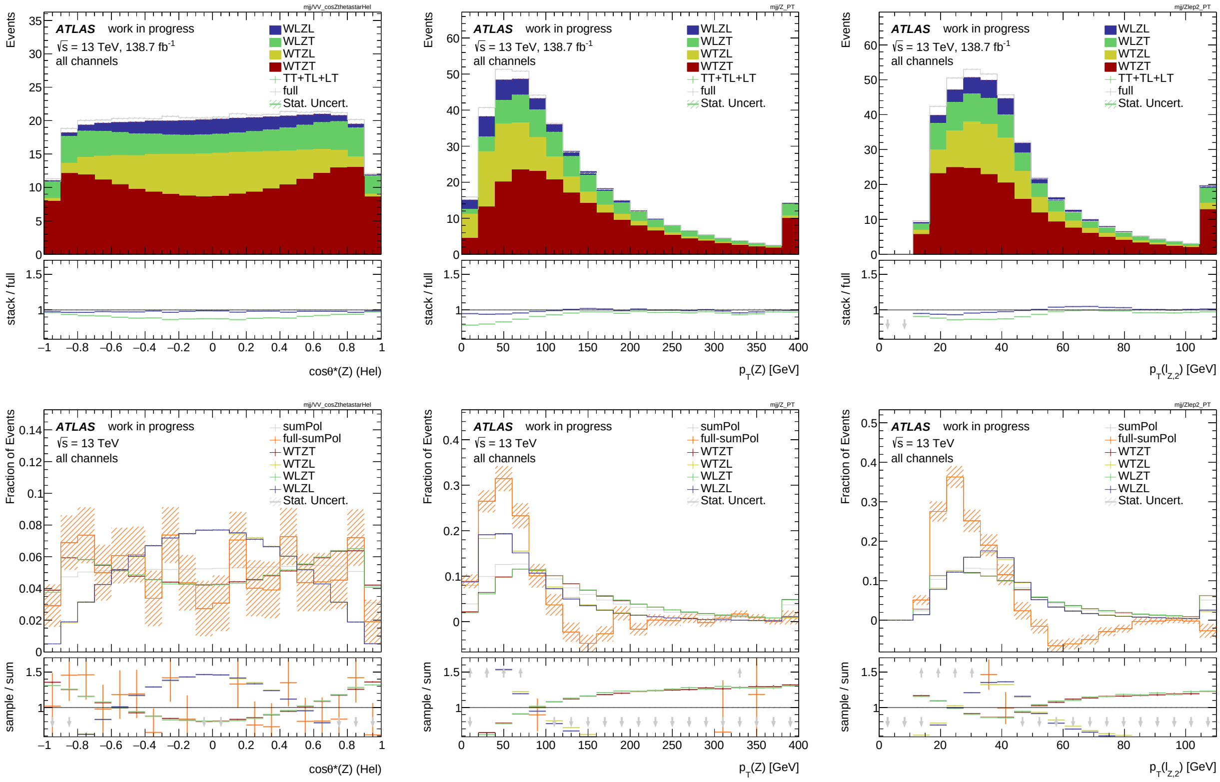


(a) (b) (c)

Figure 5.5: Distributions of (a) the transverse momentum of the W lepton $p_T(\ell_W)$, (b) the missing transverse energy $p_T(MET)$ and (c) the pseudorapidity of the W lepton $\eta(\ell_W)$. In the plots on the top the full sample and in the stack the four polarization combinations are shown. In the subplot the ratio of the stack to the full sample is shown as the blue line and the the ratio of the sum of WTZT, WTZL and WLZT to the full sample as the green line. In the plots at the bottom the normalized distributions of the four polarized samples and the sum of the polarized samples is plotted. Additionally the normalized distribution of the difference of the full sample to the sum of the polarized samples is shown. In the subplot the ratios of the individual samples to the sum of the polarization is given.

parton parton frame, a larger amount of events with longitudinally polarized Z bosons can be found for $\cos\theta \approx 0$. This is shown in Figure 5.8a and Figure 5.8b.

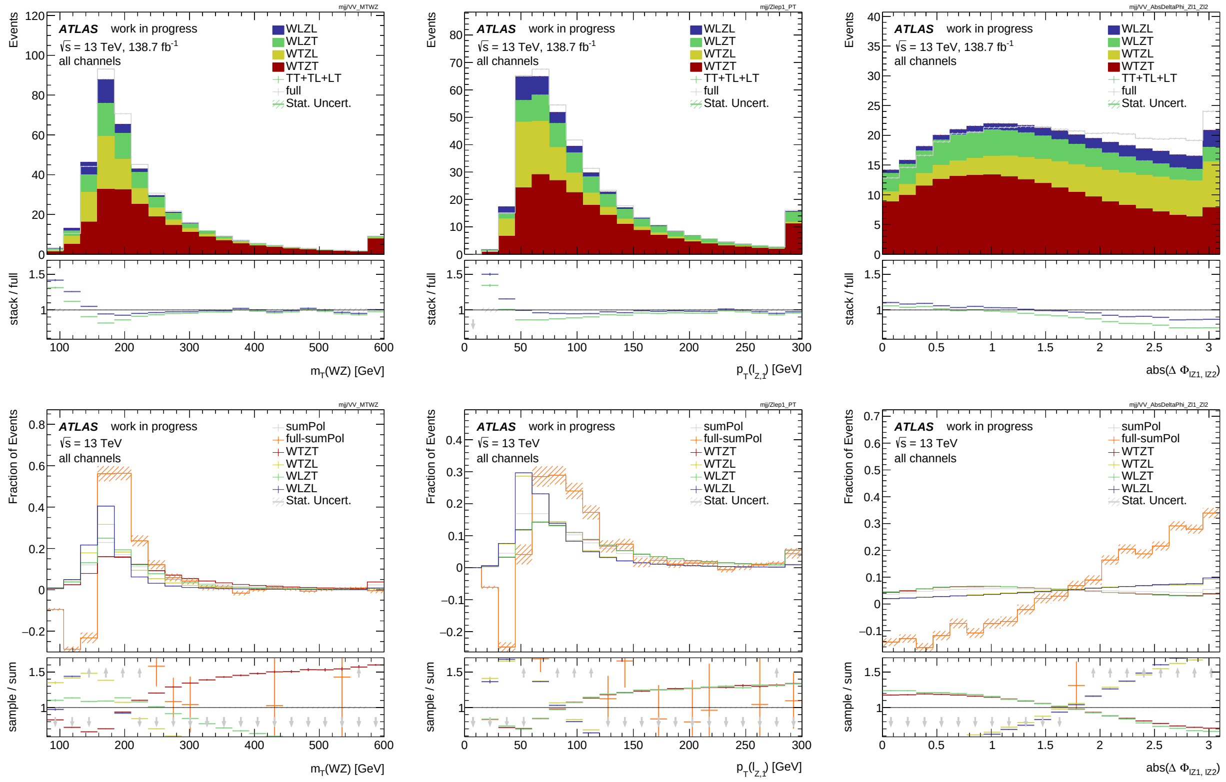
The tarnsversal momentum of the diboson system $p_T(WZ)$ and the variable $p_T^{bal}(jj)$ are also variable for that differences between the frames are visible. Their normalized distributions are shown in Figure 5.8c to Figure 5.8f. For the helicity frame samples the largest shape differences is found between the distributions with LL polarization and the distributions of the other polarizations. The shape differences between the other polarizations are small in comparison. For the parton parton frame samples the difference between the normalized distributions of WLZL and the mixed sample seems less strong. However this would need to be confirmed with samples with less statistical uncertainties. Another behaviour can be found for $p_T^{bal}(\ell\nu\ell\ell)$. While for this variable in the normalized distributions of the mixed samples with polarization defined in the helicity frame look different than for the samples where both



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

Figure 5.6: Distributions of (a) the cosine of the decay angle of the Z boson $\cos\theta^*(Z)$, (b) the transverse momentum of the Z boson $p_T(Z)$ and (c) the the transverse momentum of the subleading Z lepton $p_T(\ell_{Z,2})$. In the plots on the top the full sample and in the stack the four polarization combinations are shown. In the subplot the ratio of the stack to the full sample is shown as the blue line and the the ratio of the sum of WTZT, WTZL and WLZT to the full sample as the green line. In the plots at the bottom the normalized distributions of the four polarized samples and the sum of the polarized samples is plotted. Additionally the normalized distribution of the difference of the full sample to the sum of the polarized samples is shown. In the subplot the ratios of the individual samples to the sum of the polarization is given.

bosons are either longitudinally or transversely polarized, for the parton parton frame samples only the LL polarization shown a large shape difference compared to the other polarization combinations. This can be seen in Figure 5.9a and Figure 5.9b.
The distribution of the centrality of the Z boson can be seen in Figure 5.9c and Figure 5.9d. For the helicity frame samples only small shape differences between the different polarizations are visible. For the parton parton frame samples some sensitivity to the polarization of the Z boson is visible. Events with longitudinally polarized Z boson have a tendency to be at larger values.
In Figure 5.9e and Figure 5.9f the transverse momentum of the subleading Z lepton is shown as an example for a variable for that no large differences between the frames can be seen.
Since already some differences are found a further study of the samples with polarization



(a) (b) (c)

Figure 5.7: Distributions of (a) the transverse mass of the diboson system $m_T(WZ)$, (b) the transverse momentum of the leading Z lepton $p_T(\ell_{Z,1})$ and (c) the absolute values of the difference of the azimuthal angle of the leading and subleading Z leptons $|\Delta\phi(\ell_{Z,1}, \ell_{Z,2})$. In the plots on the top the full sample and in the stack the four polarization combinations are shown. In the subplot the ratio of the stack to the full sample is shown as the blue line and the the ratio of the sum of WTZT, WTZL and WLZT to the full sample as the green line. In the plots at the bottom the normalized distributions of the four polarized samples and the sum of the polarized samples is plotted. Additionally the normalized distribution of the difference of the full sample to the sum of the polarized samples is shown. In the subplot the ratios of the individual samples to the sum of the polarization is given.

defined in the parton parton frame might be interesting. For a more detailed study more events would need to be generated. For further studies in this thesis only the samples where the polarization is defined in the helicity frame are used.

5.1.7 Validation of Samples with τ Decays

The samples used so far have been produced with only electrons and muons in the final state. Processes where the W or the Z decays in at least one τ lepton have not been considered yet. Since the τ has a very short lifetime of only about 2.9 10^{-13} s [30] it can not be measured directly by the ATLAS detector. Because of its high mass the τ can decay hadronically, mostly into pions, or leptonically into electrons or muons. The fully leptonic decay accounts for only about 35.21% [30] of the decays. Thus a large number of events with a τ decays do not fulfill

Table 5.4: Yields and fractions of the polarized samples in the helicity and the parton parton frame.

	helicity frame	parton parton frame
sumPol	384.04 $\pm$ 0.34	382 $\pm$ 4
WTZT	203.85 $\pm$ 0.30	225 $\pm$ 3
WTZL	83.84 $\pm$ 0.12	78 $\pm$ 1
WLZT	66.94 $\pm$ 0.10	59.5 $\pm$ 0.8
WLZL	29.41 $\pm$ 0.04	20.32 $\pm$ 0.30
TT / sum	0.5308 $\pm$ 0.0012	0.588 $\pm$ 0.013
TL / sum	0.2183 $\pm$ 0.0008	0.203 $\pm$ 0.008
LT / sum	0.1743 $\pm$ 0.0007	0.156 $\pm$ 0.007
LL / sum	0.0766 $\pm$ 0.0004	0.053 $\pm$ 0.004

the required number of leptons or do not pass other phase space requirements. Therefore their contribution on the VBS events is expected to be very low. Nevertheless it is important to include these processes for completeness.
Samples with at least one τ in the final state of the hard process were produced for validation for the polarization and full samples. More information about their production can be found in section B.1. They are summarized in Table B.1b.
The yields of these samples in the truth phase space are given in Table 5.5. As expected the number of events of the τ samples is very low. In the truth phase space the distributions are dominated by statistical fluctuations. To be able to see actual shapes of the distributions the samples were looked at in a loose phase space. For that from the cuts listed in Table 5.1 only those up to the cut on the transverse momentum of the jets are applied. No unexpected behaviour was found for the used variables. The distributions of some basic variables can be seen in Appendix E.

Table 5.5: Event yields of the full and the four polarized samples. In the first column the yields for samples with only electron and muons in the final state are shown. The second column shows the yields for samples with at least one τ in the final state of the hard process.

	only e, mu	at least 1 tau
full	393.4 $\pm$ 0.6	0.168 $\pm$ 0.018
sum	384.04 $\pm$ 0.34	0.171 $\pm$ 0.010
WLZL	29.41 $\pm$ 0.04	0.0144 $\pm$ 0.0012
WLZT	66.94 $\pm$ 0.10	0.0269 $\pm$ 0.0027
WTZL	83.84 $\pm$ 0.12	0.040 $\pm$ 0.004
WTZT	203.85 $\pm$ 0.30	0.090 $\pm$ 0.009

5.1.8 Comparison of the Local and Official Samples

Since the study with the local samples showed good results, a official production of the polarization samples was requested [32]. To validate the official samples, a comparison between the local and official samples was done at truth level. For the full sample an officially produced sample was already available. This sample was produced with MadGraph version 2.6.2. To see if this sample should be produced with a newer version a comparison to the with Madgraph version 2.9.5 locally produced sample was done. The official samples are summarized in Table G.1.

The event yields for the full sample and the polarization samples with only electrons and muons in the final state are shown for the local and official production in Table 5.6. A good agreement within 1% can be seen. The comparison of the distributions shows good agreement within the uncertainties. In Figure 5.10 the distributions of the invariant mass and the transverse momentum of the W lepton are shown as an example. While the event yields of the official full sample are a bit lower for small m_{jj} values, in general a good agreement between all the official and local samples was found. The distributions of some additional variables can be seen in Appendix F. Since no large differences between the full sample produced with Madgraph version 2.6.2 and 2.9.5 was found, it was decided against requesting a new official sample.

Table 5.6: Comparison of the event yields of the locally and officially produced samples.

	official	local	local/official
full	390.8 ± 0.8	393.4 ± 0.6	1.0067 ± 0.0027
WTZT	205.15 ± 0.35	203.85 ± 0.30	0.9937 ± 0.0022
WTZL	84.52 ± 0.22	83.84 ± 0.12	0.9920 ± 0.0029
WLZT	67.48 ± 0.19	66.94 ± 0.10	0.9919 ± 0.0032
WLZL	29.30 ± 0.13	29.41 ± 0.04	1.004 ± 0.005

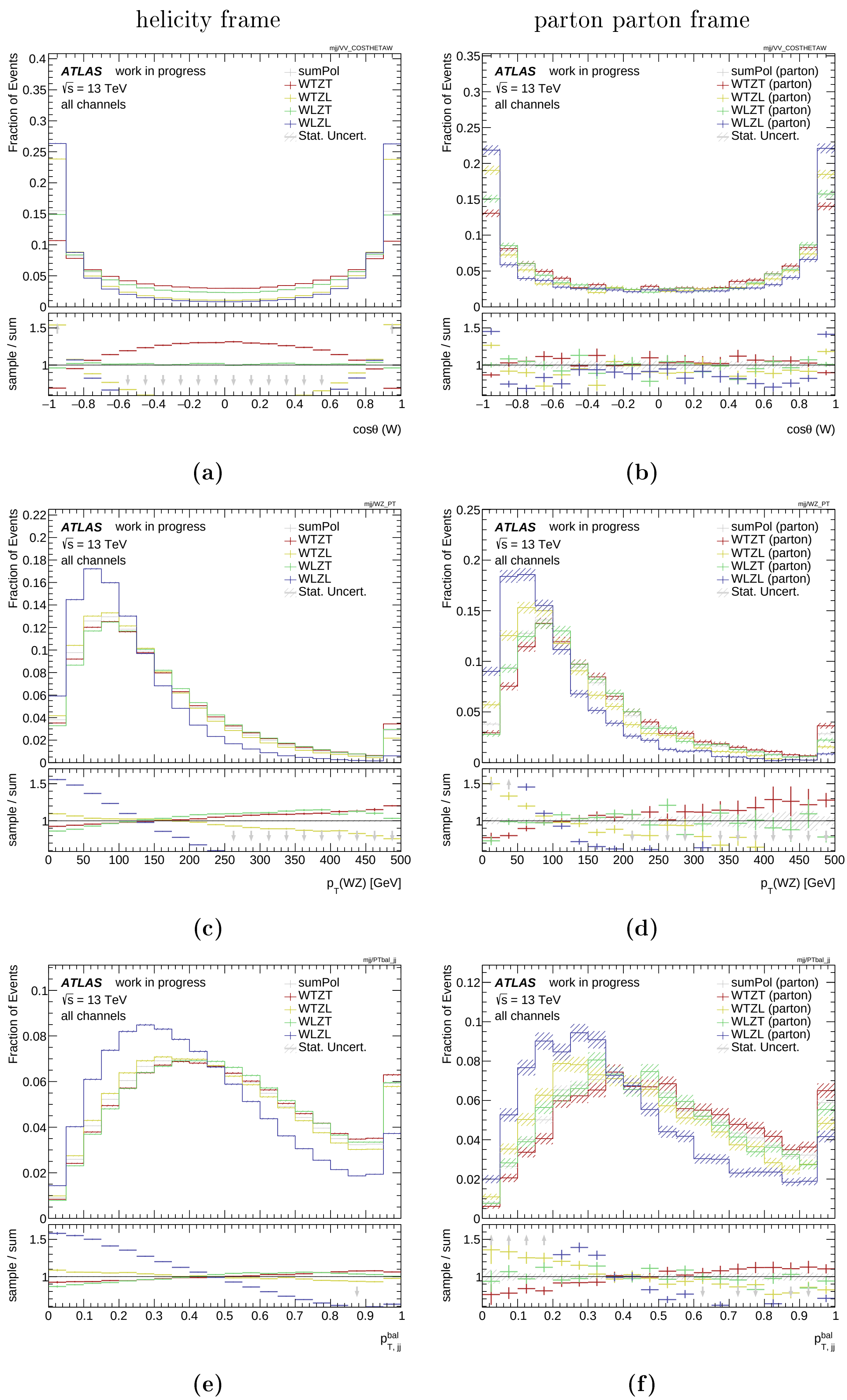
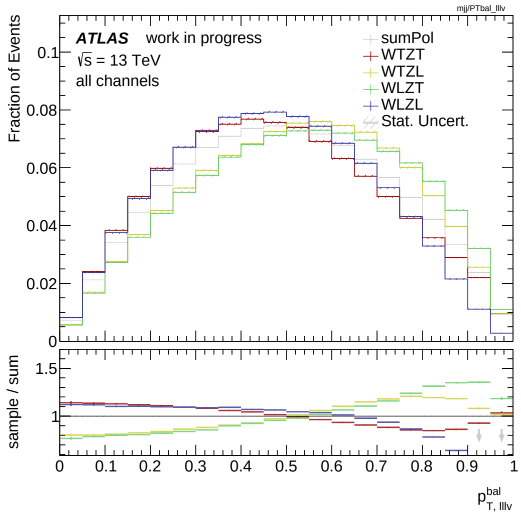
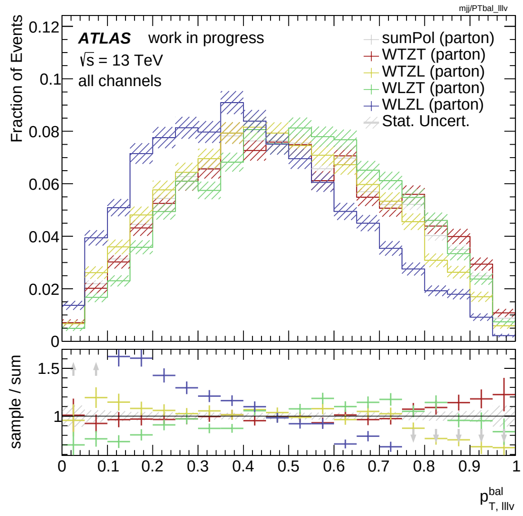


Figure 5.8: Normalized distributions of the different polarization combinations on the left in the helicity frame and on the right in the parton parton frame. From the top to the bottom the cosine of the angle of the W boson to the beam axis in the WZ rest frame $\cos\theta(\mathrm{W})$, the transverse momentum of the diboson system $p_T(WZ)$ and $p_T^{bal}(jj)$ of the two jets are shown.

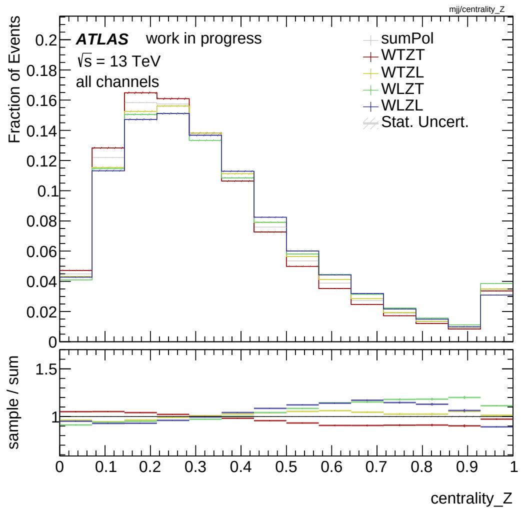
helicity frame | parton parton frame



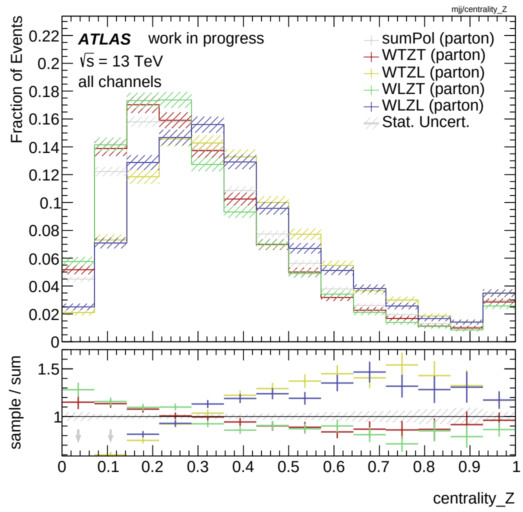
(a)



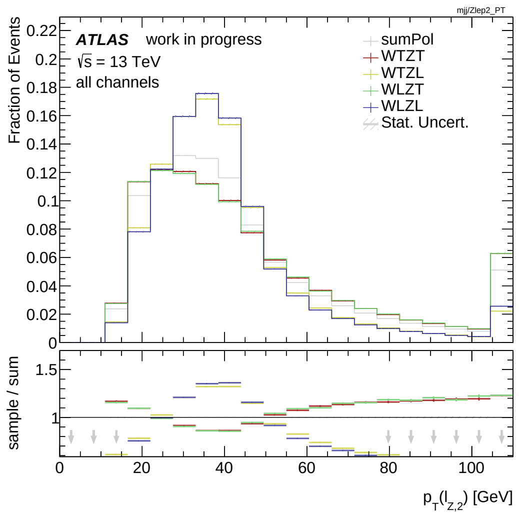
(b)



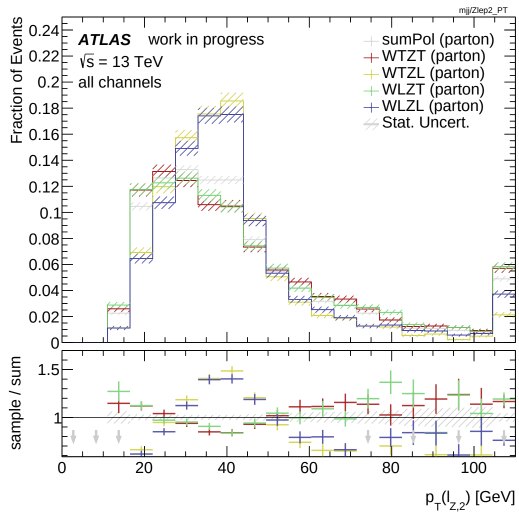
(c)



(d)

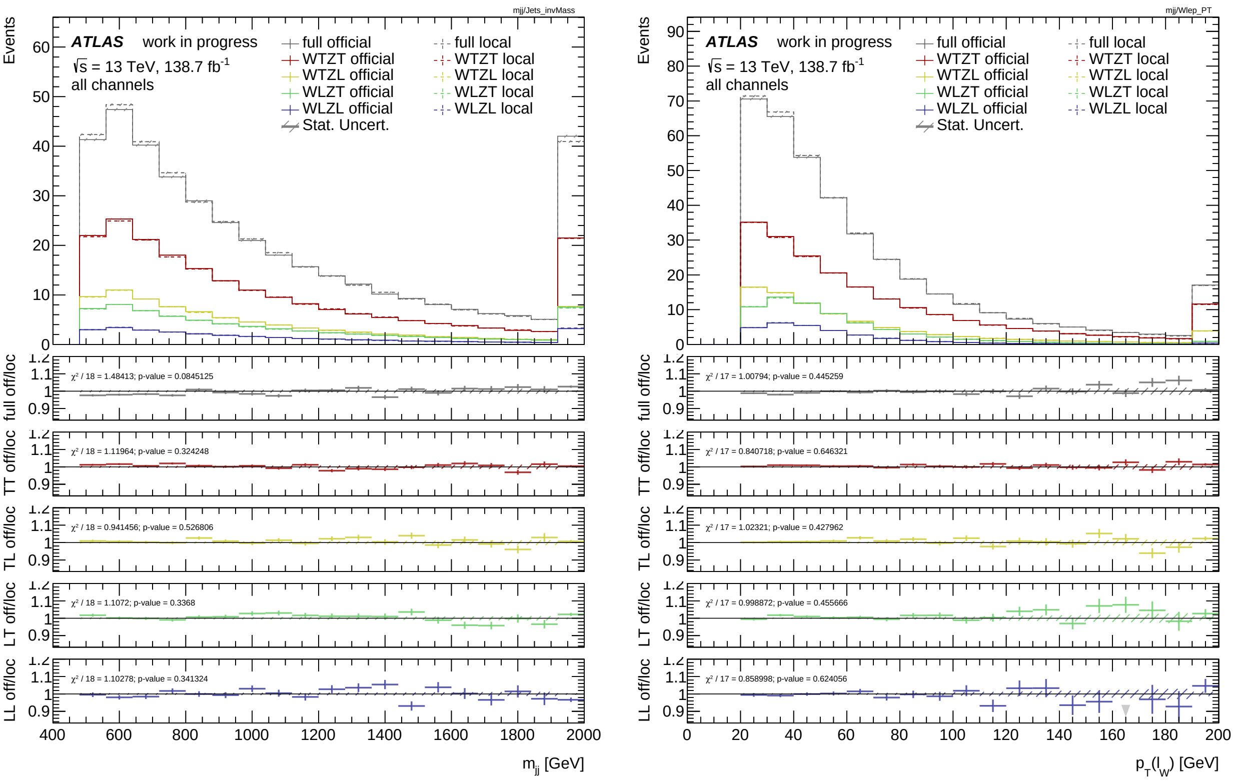


(e)



(f)

Figure 5.9: Normalized distributions of the different polarization combinations in the helicity frame (left) and in the parton parton frame (right). From the top to the bottom $p_T^{bal}(\ell\nu\ell\ell)$ of the three leptons and the neutrino, the centrality of the Z boson and the transverse momentum of the subleading Z lepton are shown.



(a) (b)

Figure 5.10: Comparison of the normalized distributions of the four polarization combinations between the locally (dashed lines) and officially (full lines) produced samples. In (a) the invariant mass of the two tagging jets m_{jj} and in (b) the transverse momentum of the W lepton are shown. In the subplots the ratio of the officially produced sample to the locally produced sample is shown for the full sample and every polarization combination.

5.2 Study at Reconstruction Level

So far the analysis was done at truth level. For a more realistic approach that includes the detector effects the samples are used at reconstruction level (reco level) for the further analysis. For the following analysis the polarization samples with only electrons and muons are always combined with the sample with at least one tau lepton in the final state.

5.2.1 Object Selection

At reconstruction level the physics objects have to be reconstructed from the simulated detector outputs. For the selection of the physics objects like electron, muons and jets, certain object selection criteria are applied on the kinematic properties of the objects. For the object selection the ELCore framework with Analysis Release 21.2.161 was used. The object selection is based one the one in the previous WZ VBS study in [31].

Leptons

For electrons and muons three levels of selections are used. They are referred to as baseline selection, Z selection and W selection. Each level is a subset of the previous one with more strict criteria applied. The electron and muon object selections are summarized in Table 5.7 and Table 5.8 respectively. There the d_0 and the z_0 describe the transverse and longitudinal impact parameters. An electron does not pass the electron object quality criteria if their cluster is effected by the presence of dead FEB or dead high voltage regions of the detector. For the electrons, object that show ambiguities between electrons and photons are vetoed in the W selection. This is done by requiring that the author flag of the electrons is 1 and using only object for that `DFCommonAddAmbiguity` ≤ 0 is fulfilled.

Jets

Jets are taken from the `AntiKt4EMPFlowJets_BTagging201903` collection. They are required to fulfil $p_T > 25$ GeV and $|\eta| < 4.5$. To remove pil-up jets, the pile-up jet vertex tagger `JVT` is used at the `Tight` working point for central jets and `fJVT` at the `Loose` working point for forward jets. For both a maximal p_T of 60 GeV is used.

Missing transverse energy

The missing transverse energy is calculating with the `METMarker` tool using baseline electrons, baseline muons and jets from the `AntiKt4EMPFlowJets` collection. Additional momentum energy deposits in the detector which are not associated with a reconstructed object are included via the soft term.

Overlap removal

Overlap removals have to be applied to avoid duplications when the detector outputs can be reconstructed as different objects. The overlap removals for different objects are applied in the following order:

- If a baseline electron and a baseline muon share a track the muon gets removed if it is calorimeter tagged. Otherwise the electron gets removed. If two baseline electrons share a track the one with the smaller transverse momentum gets removed
- Z selected electrons found within a range $0.2 < \Delta R < 0.4$ to a jet are removed. Z selected muons are removed if they are within a range $0.2 < \Delta R < 0.4$ to a jet that has less than three tracks.
- Jets with a distance $\Delta R < 0.2$ to a Z selection electron are removed. Jets with less than three tracks that are found within $\Delta R < 0.2$ to Z selected muons are removed.

5.2.2 Variables at Reconstruction Level

The leptons that are associated to the W and Z boson and that are thus used to reconstruct their variables are chosen as follows. All pairs of Z selected leptons with same flavour and opposite charge are taken as candidates for the Z leptons. If no such pair exist also pairs of leptons with same charge are used. Without such a pair the event is rejected. As the leading and subleading Z lepton $\ell_{Z,1}$ and $\ell_{Z,2}$ the pair of leptons with an invariant mass closest to the Z boson mass are chosen. The leading Z lepton is the one with the higher transverse momentum. Of the remaining leptons the one with the highest p_T that is W selected is chosen as the W lepton ℓ_W. If no such lepton exist the event gets rejected. For variables that depend only on the transverse momentum of the neutrino the missing transverse energy is used as an estimate for the neutrino. This applies to the transverse momentum of the W boson, the transverse momenta balance and the transverse mass of the W boson and the diboson system. For variables that require the full reconstruction of the neutrino momentum the z component is reconstructed by constraining the W mass to 80.385 GeV. For the decay of the W boson into a lepton and a neutrino, this can be solved for the z component of the neutrino. For this the lepton mass is neglected and the missing transverse energy is assumed to be the transverse momentum of the neutrino. This leads to two possible solutions of $p_z(\nu)$, from which the smaller one is chosen. With that the neutrino and subsequently the W boson can be reconstructed. This is used for the angular variables of the W boson, the invariant mass of the WZ system $m(WZ)$ and the decay angle distributions.

Table 5.7: Electron object selection used in this analysis.

Electron object selection			
Selection	Baseline Selection	Z Selection	W Selection
$p_T > 5$ GeV	✓	✓	✓
Electron object quality	✓	✓	✓
$\|\eta^{\text{cluster}}\| < 2.47$, $\|\eta\| < 2.5$	✓	✓	✓
`Loose_BL` identification	✓	✓	✓
$\|d_0/\sigma(d_0)\| < 3$	✓	✓	✓
$\|\Delta z_0 \sin\theta\| < 0.5$mm	✓	✓	✓
`Loose_VarRad` isolation	✓	✓	✓
e-to-μ and e-to-e overlap removal	✓	✓	✓
$p_T > 15$ GeV		✓	✓
Exclude $1.37 < \|\eta^{\text{cluster}}\| < 1.52$		✓	✓
`Medium` identification		✓	✓
`HighPtCaloOnly` isolation		✓	✓
e-to-jets overlap removal		✓	✓
$p_T > 20$ GeV			✓
`Tight` identification			✓
`Tight_VarRad` isolation			✓
`Unambiguous` author			✓
`DFCommonAddAmbiguity` ≤ 0			✓

Table 5.8: Muon object selection used in this analysis.

Muon object selection			
Selection	Baseline Selection	Z Selection	W Selection
$p_T > 5$ GeV	✓	✓	✓
$\|\eta\| < 2.7$	✓	✓	✓
`Loose` quality	✓	✓	✓
$\|d_0/\sigma(d_0)\| < 3$ (for $\|\eta\| < 2.5$ only)	✓	✓	✓
$\|\Delta z_0 \sin\theta\| < 0.5$mm (for $\|\eta\| < 2.5$ only)	✓	✓	✓
`PflowLoose_FixedRad` isolation	✓	✓	✓
μ-to-e overlap removal	✓	✓	✓
$p_T > 15$ GeV		✓	✓
$\|\eta\| < 2.5$		✓	✓
`Medium` quality		✓	✓
μ-to-jet overlap removal		✓	✓
$p_T > 20$ GeV			✓
`Tight` quality			✓
`PflowTight_FixedRad` isolation			✓

For the tagging jet variables, the ones of the two jets with the highest transverse momentum are chosen.

5.2.3 Comparison of Distributions Between Truth and Reco Level

To see the effects of the detector and of the object reconstruction, a comparison of the official sample between truth and reconstruction level was done. For this the truth phase space was used. Since not all objects are detected or reconstructed correctly by the detector a lower event yield is expected for the the samples at reconstruction level. This can be seen in Table 5.9 which compares the event yields at truth and reco level. When looking at the ratio between the full sample and the sum of the polarization samples it can be seen that this is larger at reco level. While at truth level the disagreement is about 1.1% at reco level it is at about 4.7%. Since the events in the polarization samples should be reconstructed equivalently to events were the polarization information is not known, this difference is not expected. To check if different generator level cuts could be responsible for this differences, they were compared between the full and the polarization samples. However the same generator level cuts were used. To see if the shape of the differences changes between truth and reco level, the distributions of some variables were investigated. As an example the pseudorapidity of the first tagging jet and the transverse momenta of the leading and subleading Z lepton are shown in Figure 5.11. The ratio of the full and the sum of the polarization samples is shown as the blue line in the subplot. No large differences are visible. Since the difference between the full sample and the sum of the polarization samples is not too large and have a similar shape to the one at truth level it was decided to continue with this.
To see the effects of the detector on the variables, the normalized distributions of the polarization samples were investigated at reco level. Since the kinematics of the neutrino can only be approximated using the missing transverse energy, distributions of variables of the W boson are expected to show differences at reco level. This can clearly be seen for the normalised distributions of the pseudorapidity and the decay angle of the W boson, as shown in Fig-

Table 5.9: Event yields of the official sample at truth and reco level in the truth phase space.

	truth	reco	reco/truth
full	390.8 $\pm$ 0.8	186.1 $\pm$ 0.6	0.48 $\pm$ 0.60
sumPol	386.5 $\pm$ 0.5	177.4 $\pm$ 0.4	0.46 $\pm$ 0.40
WTZT	205.15 $\pm$ 0.35	94.91 $\pm$ 0.28	0.46 $\pm$ 0.28
WTZL	84.52 $\pm$ 0.22	38.40 $\pm$ 0.17	0.45 $\pm$ 0.17
WLZT	67.48 $\pm$ 0.19	30.30 $\pm$ 0.15	0.45 $\pm$ 0.15
WLZL	29.30 $\pm$ 0.13	13.74 $\pm$ 0.10	0.47 $\pm$ 0.10
full / sum	0.9889 $\pm$ 0.0024	0.953 $\pm$ 0.004	

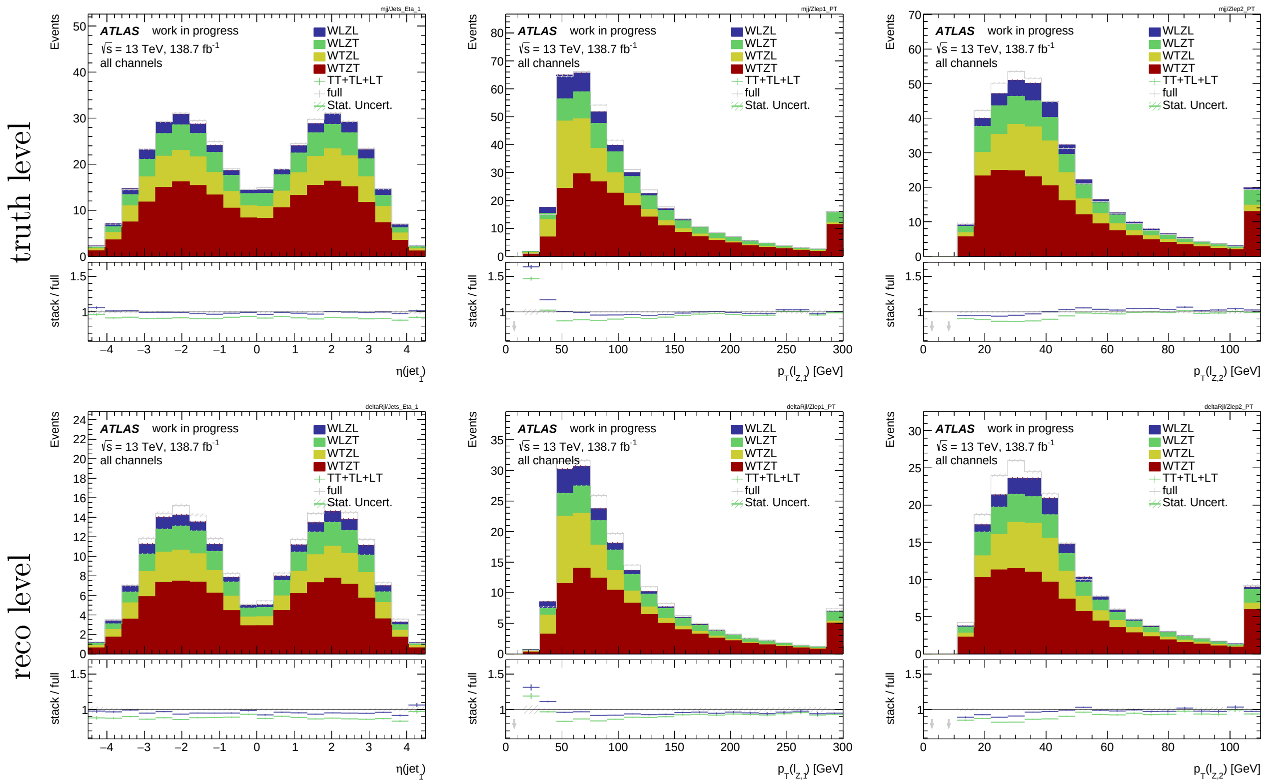
ure 5.12. The distributions of all the polarization combinations for the pseudorapidity of the W boson have a peak at zero at reco level. At truth level, all except the transverse transverse polarization showed a double maximum around zero. Shape differences between the different polarization combinations are still visible at reco level, where the distribution of the LL and the mixed polarizations is wider than the one of the transverse transverse polarization. For the decay angle distribution the difference between the distributions of events with longitudinal and transverse polarization of the W boson is less pronounced. Thus this variable show less differentiation power at reco level. Another variable that shows differences between truth and reco level is $\cos\theta(W)$. While still small, the fraction of events with longitudinally polarized Z boson around $\cos\theta(W) \approx 0$ is larger at reco level than at truth level. This can be seen in Figure 5.12c.

In Figure 5.13a the transverse momentum of the W boson is shown. The distribution of this variable is broader and covers larger values of this variable at reco level. For this variable the neutrino truth transverse momentum is replaced by the missing transverse momentum. However for the transverse mass of the WZ system $m_T(WZ)$, the transverse momentum of the W boson $p_T(W)$ and the transverse momentum balance $p_T^{bal}(\ell\nu\ell\ell)$, for that this replacement is also done, no large shape differences between truth and reco level are visible.

As expected the distributions generally tend to be a bit broader at reco level due to the reconstruction. However for most variables, either the shapes are still very similar at truth and reco level or changes in the distributions are not effecting the differences between the polarizations much. As an example in Figure 5.13b the pseudorapidity of the subleading Z boson is shown. While differences between truth and reco level are visible, the shape differences between the different polarization combinations are still similar. In Figure 5.13c the decay angle distribution of the Z boson is shown as an example of a variable with similar shapes at truth and reco level. The reco level distributions of other variables, that are used in this thesis, can be found in Appendix H.

5.2.4 Phase Space Definition at Reconstruction Level

For the further study some cuts are added additionally to the ones in the truth phase space. This was done to match the fiducial WZjj-EW phase space used in the previous WZ VBS study [31]. One of the new cuts require the tagging jets to be in opposite hemispheres to further increase the sensitivity to VBS processes. Additionally a veto on b-jets is included. This helps reducing the contribution of the tZj process that was described in section 2.2 and other backgrounds that will be described in the next section. The definition of the phase space is summarized in Table 5.10. This phase space is referred to as reco phase space during this thesis. It is used in all further studies done in this thesis.



(a) (b) (c)

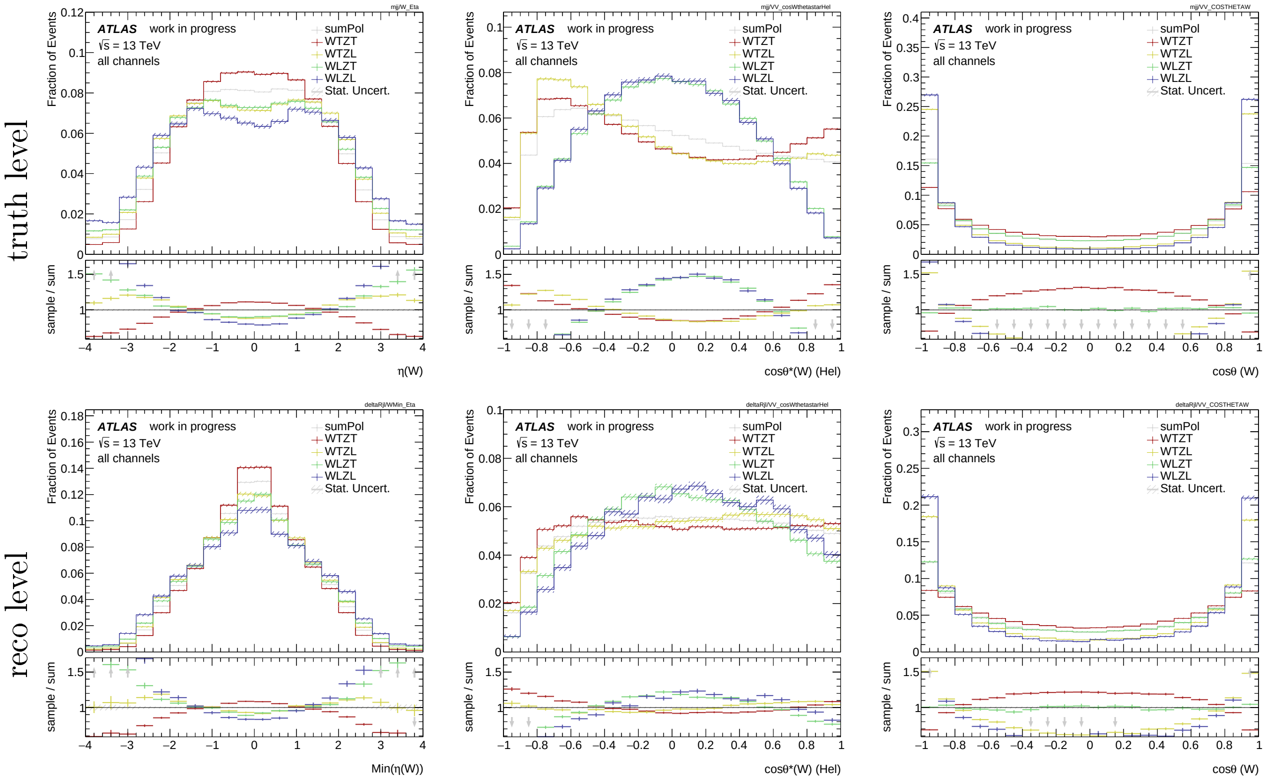
Figure 5.11: Distributions of a) the pseudorapidity of the first tagging jet $\eta(jet_1)$, (b) the transverse momentum of the leading Z lepton $p_T(\ell_{Z,1})$ and (c) the transverse momentum of the subleading Z lepton $p_T(\ell_{Z,2})$ are shown. The full sample is shown as the gray line and the different polarization combinations are shown in the stack. In the subplot the ratio of the stack to the full sample is shown as the blue line and the the ratio of the sum of WTZT, WTZL and WLZT to the full sample as the green line.

5.2.5 Inclusion of Backgrounds

For a proper study of the polarized WZjj EW events backgrounds have to be included. The backgrounds can be divided into two categories. Backgrounds with at least three reconstructed leptons that originate from the hard process are called prompt or irreducible backgrounds. If at least one of the reconstructed leptons is misidentified or originates not from the hard process the background is called non-prompt or reducible.

The largest background in this study is the WZ QCD background which is described in section 2.2. Another prompt background is the ZZ process, where one of the leptons in the final state is not reconstructed or identified in the detector. Other prompt backgrounds are the triboson processes VVV ($V = W, Z$) and $t\bar{t}V$ processes where top quarks are produced in association to a W or Z boson.

The largest non-prompt backgrounds that is included in the analysis is the pair production of top quarks $t\bar{t}$. The top quarks normally decays into a b quark and a W boson. For a leptonically decaying W boson the third lepton can be misidentified from a b quark or caused

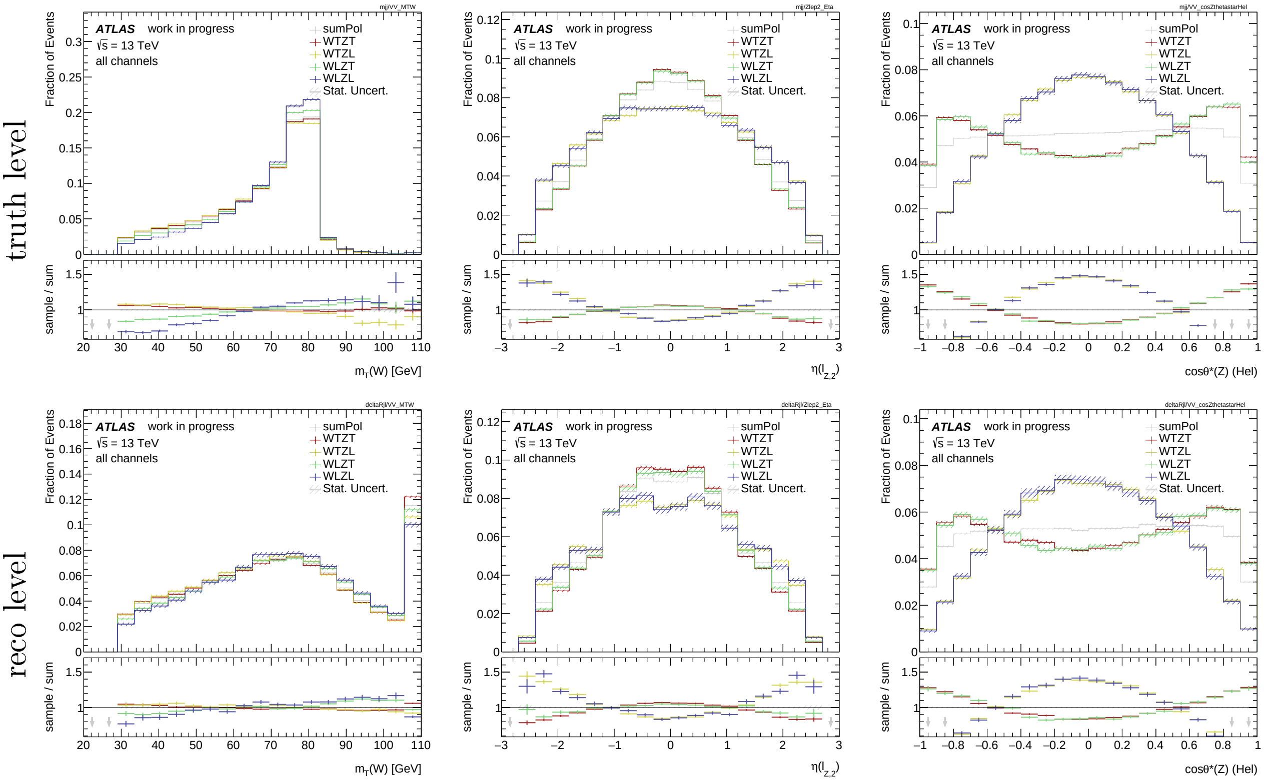


(a) (b) (c)

Figure 5.12: Normalized distributions of (a) the pseudorapidity of the W boson $\eta(W)$, (b) the cosine of the decay angle of the W boson $\cos\theta^{*}(W)$ and (c) the cosine of the angle of the W boson to the beam axis in the WZ rest frame at truth (top) and reco (bottom) level for the different polarization combinations.

by other radiations. Another non-prompt background is called single top. This includes s- and t-channel production or production in association of a W boson of the top quark. As for the $t\bar{t}$ background leptons can originate from the decay of the W boson, while misidentified jets can lead to the additional leptons. $Z + \gamma$ processes are another non-prompt background, where the photon causes the misidentification of the third lepton. The last non-prompt background that is included in this analysis are processes, where a Z boson is produced in association with a jet. The jet can be misidentified as the third lepton.

In this analysis these backgrounds will be referred to as the non-polarization backgrounds. The Monte-Carlo samples used for these backgrounds are summarized in Table G.2. As before when speaking of the polarizations or the full sample this always refers to WZjj EW processes. The polarizations that are taken as signal are referred to as signal polarization and those that are taken as background background polarization. In Table 5.11 the event yields of all the non-polarization background samples as well as of the different polarization combinations and the full sample are shown. The background is clearly dominated by the WZ QCD process with over 400 expected events. The second largest non-polarization background is the ZZ



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

Figure 5.13: Normalized distributions of (a) the transverse mass of the W boson $m_T(W)$, (b) the pseudorapidity of the subleading Z lepton $\eta(\ell_{Z,2})$ and (c) the cosine of the decay angle of the Z boson $\cos\theta^*(Z)$ at truth (top) and reco (bottom) level for the different polarization combinations.

process with about 30 expected events. For the polarizations it can be seen that the number of predicted events for the LL polarization is with about eight events very low. For the mixed polarization combinations about 20 events and for the transverse transverse one about 50 events are expected. In Table 5.12 the fraction of events with a τ decay of the polarization combinations is shown. They contribute to about 4% of the events. As a comparison in Table 5.11 also the event yields from the previous WZ VBS study are shown [31]. While the event yields of the WZ QCD sample is a bit higher and the event yields of most of the other backgrounds, especially the ttV sample, are a bit lower in this analysis, a reasonably good agreement can be seen.

The background distributions for some example variables are shown in Figure 5.14. Especially for the three shown jet variables, m_{jj}, $|\Delta R(jets)|$ and n_{jets} it can be seen that the combined non-polarization backgrounds have a different shape than all the different polarization combinations. Thus these variables are useful to separate the WZjj EW process from the non-polarization backgrounds. In Figure 5.14d and Figure 5.14e the transverse momenta of the W and the Z boson are shown. For these the shape of the non-polarization background distributions is close to the one of events with a transverse polarization of the respective bo-

Table 5.10: Definition of the reco phase space.

Variable	Cut
$p_T(\ell_Z)$, $p_T(\ell_W)$ [GeV]	> 15, > 20
$\|\eta(\ell)\|$	< 2.5
$\Delta R(\ell_{Z,1}, \ell_{Z,2})$, $\Delta R(\ell_Z, \ell_W)$	> 0.2, > 0.3
Jet multiplicity	> 1.5
p_T(tagging jet)[GeV]	> 40
m_T^W [GeV]	> 30
m_Z [GeV]	$\|m_Z - m_Z^{PDG}\| < 10$
m_{jj} [GeV]	> 500
$\eta(\text{jet1}) * \eta(\text{jet2})$	< 0
$\Delta R(j, \ell)$	> 0.3
N_{bjet}	0

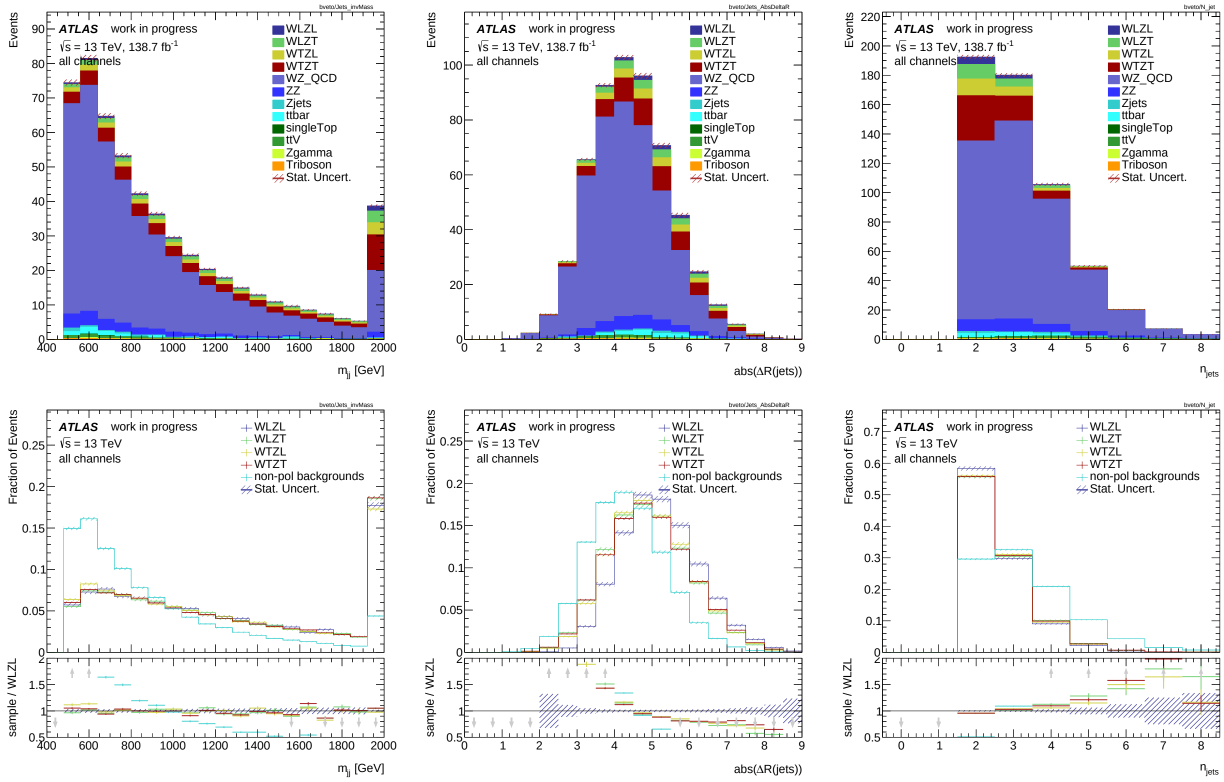
son. Thus when looking only at the polarization of the W or Z boson, the amount but not the shape of the overall background changes much. Therefore they still usable to differentiate these signal processes. For the $\cos\theta(W)$ distribution in Figure 5.14f the non-polarization backgrounds have a similar shape as the transverse transverse polarization, while the other polarization combinations have a different shape. Other background distributions for variables that are used in the further analysis can be found in Appendix I.

Table 5.11: Event yields of the non-polarization backgrounds and the different polarization combinations in the reco phase space. On the left the events yields in this analysis and on the right the event yield from the previous WZ VBS analysis [31] are shown. The WZjj EW full event yield combines the WZjj-EW and the tZj event yield and the ZZ event yield the ZZjj-QCD and ZZjj-EW event yields from [31].

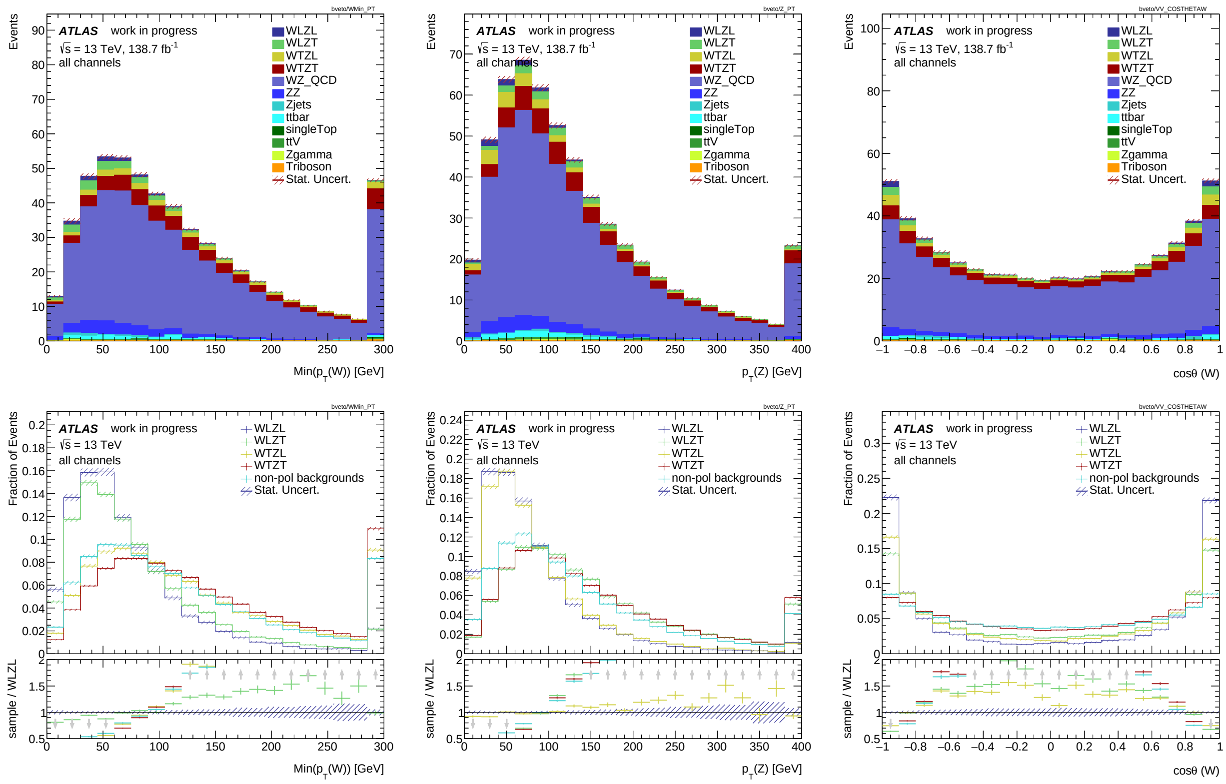
	event yields	WZ VBS note
WZ QCD	410 ± 2	368.71 ± 17.87
ZZ	27.22 ± 0.31	31.5 ± 2.44
Z+jets	3.5 ± 0.9	10.57 ± 4.23
ttbar	8 ± 1	12.91 ± 5.17
Single top	0.43 ± 0.25	
ttV	5.29 ± 0.18	22.37 ± 0.78
Z+gamma	1.0 ± 0.4	0.98 ± 0.39
Triboson	1.52 ± 0.05	2.16 ± 0.55
WLZL	8.55 ± 0.08	
WLZT	18.03 ± 0.11	
WTZL	20.27 ± 0.12	
WTZT	55.22 ± 0.21	
WZjj EW sumPol	102.06 ± 0.28	
WZjj EW full	110.7 ± 0.5	131.39 ± 3.36

Table 5.12: Fraction of events with at least one τ in the final state of the hard process of the respective polarization combination.

sum tau / all	0.040 ± 0.007
LL tau / all	0.043 ± 0.018
LT tau / all	0.036 ± 0.016
TL tau / all	0.041 ± 0.016
TT tau / all	0.040 ± 0.010



(a) (b) (c)



(d) (e) (f)

Figure 5.14: Distributions of (a) m_{jj} (b) $|\Delta R(jj)|$, (c) n_{jets}, (d) $p_T(W)$, (e) $p_T(Z)$ and (e) $\cos\theta(W)$. In the plots on the top the full sample and in the stack the four polarization combinations are shown at reco level. In the plots at the bottom the normalized distributions of the four polarized samples and the non-polarization background is shown. In the subplot the ratios of all the samples to the WLZL sample is given.

6 Extraction of the Longitudinal Polarization

The expected significance to observe the LL polarization can be extracted by performing a template fit based on a sensitive variable. As shown in Table 5.11 the contribution of events with two longitudinally polarized bosons is very low compared to the amount of the other polarizations and the non-polarization backgrounds. This makes it very difficult to extract the LL polarization especially because none of the considered variables shows very large shape differences. A way to improve the sensitivity is to use multivariate techniques such as boosted decision trees (BDTs) or neural networks to combine the information of multiple variables making it possible to use correlations between them. For this thesis BDTs were trained in order to separate WZjj EW events with longitudinally polarized bosons from the polarization and non-polarization backgrounds. The BDT outputs were used to perform a binned maximum likelihood fit to extract the expected significance for the respective polarization combinations.

6.1 Study with BDTs

For this thesis in total four BDTs with different signal configurations were trained. The tool that was used for the training utilizes the TMVA framework. The first BDT uses the longitudinal longitudinal polarization as signal and all the other polarizations and the non-polarization backgrounds as background (WLZL vs. rest, WLZL BDT). In Table 6.1 the number of signal and background events in the reco WZ VBS phase space is summarized. As already seen for the LL polarization only about 8 events are expected. This compares to about 550 background events.

To get higher statistics it can also be tried to extract processes where only one of the bosons has to be longitudinally polarized. For that two additional BDTs were trained. One of them requires for the signal process that the W boson is longitudinally polarized while the polarization of the Z boson is arbitrary (WLZX vs. rest, WLZX BDT). For the other BDT this condition is reversed (WXZL vs. rest, WXZL BDT). A fourth BDT was trained to separate WZ VBS events where at least one of the bosons is longitudinally polarized from all other backgrounds (min1L vs. rest, min1L BDT). This gives the largest statistics but is expected to show less shape differences between the signal and background polarizations. The number of expected signal and background events for each of these BDTs are also summarized in Table 6.1.

6.1.1 Introduction to Boosted Decision Trees

In this section the basic principles of Boosted decision trees will be discussed. More information can be found in [33]. Boosted decision trees are build up from multiple decision trees. In a decision tree an event is classified as signal or background based on a sequence of binary cuts. Each cut is based on the outcome of the previous cuts giving it a tree like structure. An example of such a tree can be seen in Figure 6.1.

The sequence of cuts is defined during the training of the decision tree. For the training labelled events are used that allow to check the classification of an event by the tree. The starting point of the training is a single node that contains all signal and background events. The gini-index $p(1-p)$ is calculated, where p is the signal purity. The maximal value of the gini-index is reached for $p = 0.5$, indicating a well mixed node. An increase of the signal or background fraction both leads to a decrease of the gini-index, with pure signal and background nodes giving a value of 0. For the node a grid search is done where for each variable a fixed number (nCuts) of equidistant cuts is tested. Each of these cuts splits the initial node into two subnodes. For all variables and all cut values the weighted sum of the gini-indices of the two subnodes is calculated and the variable and cut value which decreases the gini-index the most is chosen. This procedure is repeated for all of the so created subnodes until either the maximal depth of the tree (MaxDepth) is reached or the number of events fall below the minimal number of events per node (MinNodeSize). A final node where the predicted fraction of signal events is larger than 0.5 is classified as signal and the others as background. Events are classified according to the final node they fall into.

To improve the performance, multiple decision trees can be combined. For that a fixed number (NTrees) of decision trees is trained. In this thesis a technique called gradient boosting is used to combine these trees. With that the weights w_i of events i that are misclassified in previous

Table 6.1: Event yields of the signal and background of the respective BDT.

BDT	Signal	Background
WLZL vs. rest	8.55 ± 0.08	551 ± 2
WLZX vs. rest	26.57 ± 0.14	533 ± 2
WXZL vs. rest	28.81 ± 0.14	530 ± 2
min1L vs. rest	46.84 ± 0.18	512 ± 2

trees are increased by

$$w_i \rightarrow w_i \frac{1-f}{f} \tag{6.1}$$

where f is the misclassification rate of the previous trees. The weights of the entire sample are then renormalized to ensure that the sum of the weight stays constant. For each iteration of the boosting a tree weight γ is assigned to the newest tree. The tree weight is obtained by minimizing the sum over the loss functions of all events. The loss function for a single event is defined as

$$L(BDTscore_m(\gamma)) = \ln(1 + \exp(-2BDTscore_m(\gamma) \cdot y). \tag{6.2}$$

Here $BDTscore_m$ is the output of the boosted decision trees after iteration m and y the true label of the event.
To become less susceptible to statistical fluctuations the parameter Shrinkage can be introduced. It is used to scale the tree weight γ by a constant factor. For small values it reduces the the learning rate of the BDT. Another option to reduce the effects of statistical fluctuations that was used during this thesis is bagging. For that in each iteration step only a fraction (BaggedSampleFraction) of the training events is used for the training.
The result of all of the trees can be combined into a final output score. This so called BDT score has a range from -1 to 1, where lower values indicate that an event is background and higher that it is signal.

6.1.2 Evaluation of the Performance

The comparison of the BDT score distributions between the signal and the background samples or between two subsets of the input data can help to evaluate the performance of a BDT. In the first case a large disagreement indicates a good performance while in the second case a good agreement is desired. To quantify this several figures of merit (FoM) can be used. The ones that appear during this thesis will be presented below. They are inspired by the work done in [34]. The normalized BDT score distributions are denoted as n_1 and n_2.

Kolmogorov-Smirnov

For the Kolmogorov-Smirnov test the absolute difference of the values $n_{a,i}$ of the normalized distributions n_a in each bin i is calculated. The output is the largest of these differences:

$$Kol(n_1, n_2) = \max_i |n_{1,i} - n_{2,i}|. \tag{6.3}$$

A small value indicates a good agreement of the two distributions.

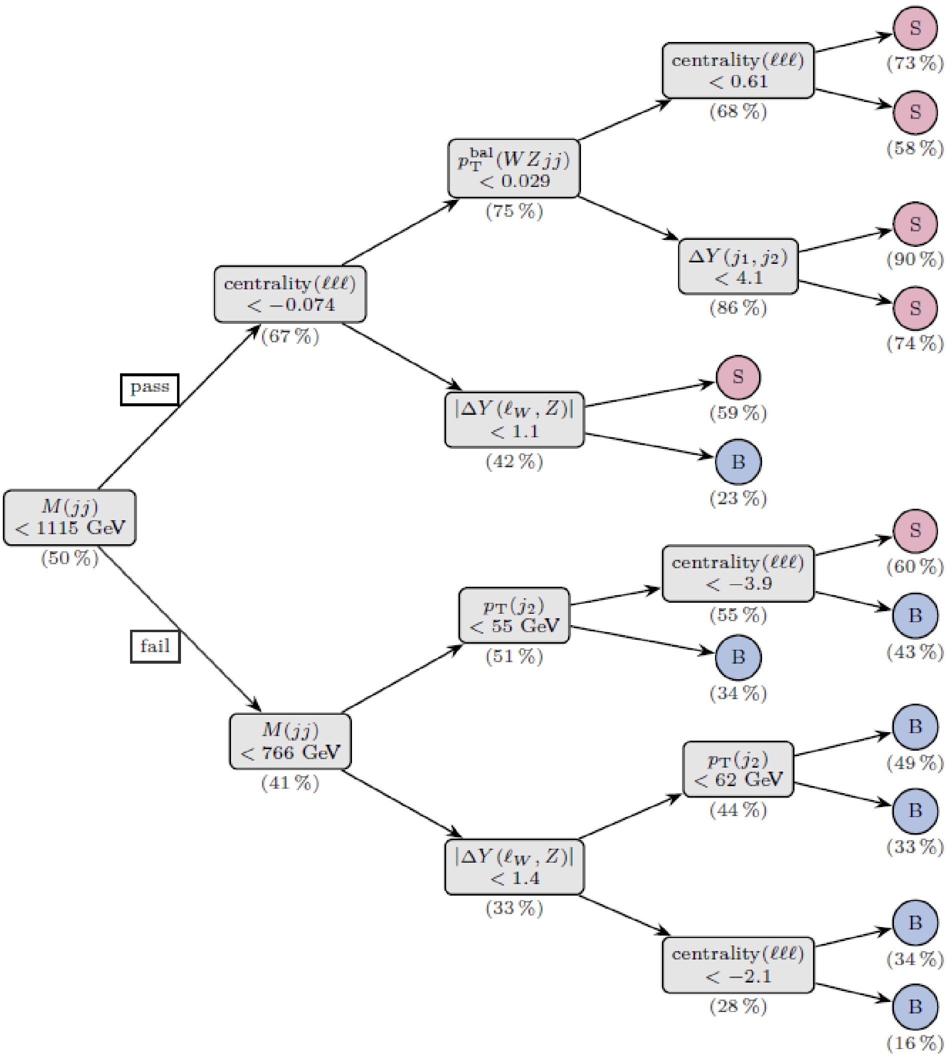


Figure 6.1: Example of a single decision tree. The cut used in each node is given in the gray boxes. The percentages below the nodes denote the signal purity in this node. [34]

Reduced χ^2

The χ^2 value is calculated as follows:

$$\chi^2 = \sum_{i=1}^{N_{bins}} \frac{(n_{1,i} - n_{2,i})^2}{\Delta n_{1,i}^2 + \Delta n_{2,i}^2}. \tag{6.4}$$

Here the $\Delta n_{a,i}$ stands for the absolute uncertainty in bin i for distribution n_a. To account for the summation over the bins the χ^2 value is divided by the degrees of freedom. A result close to 1 indicate good agreement and larger values disagreement.

Non-Overlap

For this figure of merit in each bin i the smaller value of the two normalized distributions is chosen. Then the difference of the sum of these minima to 1 is calculated:

$$f_{N-O}(n_1, n_2) = 1 - \sum_{i=1}^{N_{bins}} \min(n_{1,i}, n_{2,i}) \tag{6.5}$$

For distributions in good agreement the values of the normalized distributions are close together resulting in a value of f_{N-O} close to 0. Larger values indicate disagreement between the distributions.

Separation

This metric is defined as

$$f_{Sep}(n_1, n_2) = \sum_{i=1}^{N_{bins}} \frac{(n_{1,i} - n_{2,i})^2}{n_{1,i} + n_{2,i}}. \tag{6.6}$$

Bins that are empty are skipped. Large values indicate large disagreement.

Approximated ROC-integral

For the ROC-integral the background rejection is plotted over the signal acceptance and the curve is then integrated. A large area under the ROC curve indicates a good performance.

Estimated Significance

The significance (see subsection 6.2.1) can be estimated from the signal and background distributions by

$$Z_{SoB,sys} = \frac{S}{\sqrt{B + (\Delta B)^2}}. \tag{6.7}$$

Here S and B are the number of signal and background events in all the considered bins. With ΔB the systematic uncertainty of the background is taken into account. The significance is calculated for left- and right-sided cuts at each bin threshold and the maximum of these values is chosen as the output.

6.1.3 Overtraining Conditions

When training a BDT it has to be taken care that the BDT does not get not overtrained. If the BDT is overtrained it does no longer separate just the underlying distributions of the signal and background processes but gets sensitive to statistical fluctuations of the dataset it is trained

on. The discrimination between signal and background can then no longer be generalised to another dataset. To avoid this the dataset is split into a training dataset, which is used to train the BDT, and a test dataset, which is used to verify the performance. For BDTs with small overtraining effects the BDT score distributions for the test and the training dataset are in good agreement. Overtraining conditions based on the FoM given in subsection 6.1.2 are introduced. To be considered for further use the BDT has to fulfil these conditions.
For this analysis the simulated data was split equally into the test and training dataset using the following condition:

$$\begin{aligned} &\text{training sample} &&: \text{Event Number } \% \; 2 \neq 0 \\ &\text{test sample} &&: \text{Event Number } \% \; 2 = 0. \end{aligned}$$

The overtraining conditions consist of a Kolmogorov-Smirnov test of the BDT score distribution between the test and the training dataset for both signal Kol(S) and background Kol(B). This is used to ensure shape consistency between the samples. Another condition compares the estimated significances from the BDT score distributions of the test (te) and training (tr) datasets. A BDT with small overtraining should perform similar for both of these subsets and thus give a similar significance. The used overtraining conditions are shown in the following:

$$\begin{aligned} &\text{Kol(S)} &&< 0.0175 \\ &\text{Kol(B)} &&< 0.02 \\ &\frac{|Z_{SoB,sys,tr} - Z_{SoB,sys,te}|}{Z_{SoB,sys,tr}} &&< 0.05. \end{aligned}$$

Different values were tried for the different tests with the one above giving a reasonable overtraining rejection while allowing a good separation between the signal and background sample. They could be further optimized to get better overtraining results.

6.1.4 Variable and Metaparameter Optimization

In order to get a BDT that optimally separates the signal from the background processes, several parameters of the BDT and the choice of input variables can be optimized. This can be done by training BDTs with different combinations of variables and metaparameters and choosing the one that performs the best. As a metric to decide which BDT to choose the estimated significance on the test sample is used:

$$Z_{SoB,sys,te} = \frac{S}{\sqrt{B + (0.1B)^2}}.$$

A 10% uncertainty for the background was used. This corresponds to the systematic uncertainty of the largest background, the WZjj QCD background (see subsection 6.2.2).
The variable and metaparameter optimization was done in three steps which will be described in the following. In each step only those combinations which fulfil the overtraining conditions are considered.
In the first step a variable optimization is performed. The initial metaparameters used for this step are fixed and can be seen in Table 6.2. As the starting point for this step a BDT with a single variable is trained. Then the estimated significance of the test sample is calculated for a BDT after adding a single additional variable. This is done for each of the considered variables. The variable that increases the significance on the test sample the most is added to the initial variable and the procedure is repeated with the remaining variables. If the BDT includes more than one variable, for each of the already included variables it is also checked if dropping it gives the highest significance. This procedure is repeated until neither adding nor dropping a variable increases the significance any more.
In the second step a metaparameter optimization is performed. For that the variables from the first step are used. A grid search of all the metaparameters listed in Table 6.3 is performed. The combination which optimizes the metric the most is chosen for the final BDTs.
In the third step another variable optimization is performed with the metaparameters obtained in the second step and the variables from the first step as the initial variables. The same procedure as in the first variable optimization is used with the only difference that now it is also tried to add pairs of variables.
The final BDT that is used for the further analysis uses the metaparameters from step two and the resulting variables from step three.

6.1.5 Input Variables and Preprocessing

Variables that are considered in the optimization as input for the BDT are shown in Table 6.4 for the WLZL vs. rest case. They are ranked by decreasing non-overlap between the signal

Table 6.2: Metaparameters used at the start of the optimization. Some of them get optimized during the metaparameter optimization.

parameter	value
Boost Type	Grad
Ntrees	200
MaxDepth	4
MinNodeSize	2.5%
nCuts	20
Shrinkage	0.06
BaggedSampleFraction	0.6
NegWeightTreatment	Pray

Table 6.3: Range of values of the metaparameters that are tried during the metaparameter optimization.

Metaparameter	Considered values
Ntrees	200, 300,..., 1500
MaxDepth	3, 4, 5
MinNodeSize	1%, 1.5%, ..., 4%
Shrinkage	0.0025, 0.005, 0.0075, 0.01, 0.015, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1

and the background distributions. None of the variables that were found to have interference effects larger than 10% in any bin in the truth level study in subsection 5.1.5 were considered for the BDT training. For the other BDTs the same variables as for the WLZL BDT were used as candidates for the optimization. For them additionally the missing transverse energy was considered as a possible input variable. This variable might also be interesting as input for the WLZL BDT. The five highest ranking variables for these three BDTs are shown in Appendix J. As the starting variable for all BDTs except for the WLZX BDT the highest ranking variable was chosen. These are the cosine of the angle of the W boson to the z-axis $\cos\theta(W)$ for WLZL BDT, the transverse momentum of the Z boson $p_T(Z)$ for the WXZL BDT and the invariant mass of the jets m_{jj} for the min1L BDT. A slightly different preprocessing, that was previously used, resulted in the transverse momentum of the W boson $p_T(W)$ as the highest ranking variable for the WLZX BDT. This variable was used as the starting variable for this BDT. For the current preprocessing m_{jj} is classified as the highest ranking variable. Since the outputs of the figures of merits are very similar for both variables, this should not lead to worse training results.

For the BDT training the distributions are divided into 20 equidistant bins whose edges correspond to the cuts that are tried during the training. The outer limits are determined by the minimal and maximal value of the distribution. This can lead to a non optimal binning where only a few bins cover the area which contains the relevant information and thus reduce the discrimination power between signal and background. The sensitivity to the signal can be improved by preprocessing the variables by applying certain functions or by setting a minimal and maximal value by hand. The preprocessing used for this analysis is inspired by the work done in [34].

Variables whose distributions show a long tail are examples of variables for which preprocessing can improve the performance of the BDT. These are in particular the energy and transverse momenta of the different particles. The tail normally does not contain a lot of relevant information for the separation of the signal and background processes but is covered by most of the bins. To increase the number of bins and thus the number of possible cut values in the relevant region a logarithm can be applied on such observables.

Variables like the pseudorapidity or the cosine of the angle of the W boson to the z-axis are

symmetrical. Therefore the cut step size can be decreased when using the absolute value of these variables. While the relevant shape differences between signal and background for this variable are maintained, this can lead to the loss of informations about correlations with other variables. It has however been found that applying the absolute value does not lead to worse results.
Outlier events may lead to many bins which contain no or only very few events. These effects can be avoided by fixing the minimal and maximal value of the variables. All events with values outside this range are than treated as if they had the corresponding minimal or maximal value.
The preprocessing that was applied in this thesis can be seen in Table 6.4 for the individual variables. The minimum and maximum values were chosen relatively loose and still include most of the events. For further studies it could be beneficial to choose more strict values to improve the performance of the BDT. Since the range and shape of the distributions for the signal and background processes of the four different BDT do not vary a lot, the same preprocessing was applied for all four BDTs. This was checked by looking at the normalized distributions for all four BDTs and no need to adjust them individually was found.

6.1.6 Results

In Table 6.5 the metaparameters that result from the metaparameter optimization and thus are used for the final BDT are summarized for the four BDTs. It can be seen that for the WLZL BDT the optimization results in the initial parameters. In the optimization it can be seen that while especially BDTs with larger NTrees and smaller MinNodeSize increase the estimated significance all such BDTs were overtrained. Looking at the Kolmogorov-Smirnov test for the signal in Figure 6.2a it is visible that for the WLZL BDT before the metaparameter optimization the values are already close to the overtrainig limits. Adding more complexity can then easily lead to overtraining. For the other BDTs shown in the other plots in Figure 6.2 the values of the Kolmogorov-Smirnov test before the metaparameter optimization are much smaller than the overtraining limit and a strong increase after the metaparameter optimization is visible.
The input variables in the order in which they are included during the optimization are given in Table 6.6 for all four BDTs. Variables that are added as pairs during the second variable optimization are indicated by being grouped together by the dashed lines. While this has not to be the case, this can indicate correlations between them which might be interesting to investigate further. For the WLZL BDT the optimization results in 26 variables. The WLZX and WXZL BDTs use 25 and 24 variables respectively as input. The optimization of the min1L BDT results in 31 input variables. It can be seen that for all BDTs except the WLZL BDT the rapidity as well of the pseudorapidity of both of the jets are included. Since for the same jet these variables have very similar distributions, this can be an indication for overtrainig.

Table 6.4: List of input variables for the WLZL BDT that are considered during the optimization. They are ranked by decreasing non-overlap between signal and background. On the right side the preprocessing that is applied to these variables is shown.

Observable	Evaluation metric					Preprocessing		
	f_{N-O}	f_{Sep}	$f_{red\chi^2}$	$f_{Kolmogorov}$	f_{ROC}	function	min	max
$\cos\theta(W)$	0.30	0.113	266.29	0.30	0.66	abs		
$\|\Delta R(jets)\|$	0.28	0.124	508.40	0.28	0.65			
$p_T(W)$	0.27	0.103	317.23	0.27	0.63	log	0.2	3.2
$p_T(Z)$	0.26	0.092	274.53	0.26	0.62	log	0.2	3.1
$p_T(WZ)$	0.26	0.097	427.44	0.26	0.61	log	0.3	3.2
n_{jets}	0.24	0.092	428.04	0.24	0.50			
$p_T^{bal}(jj)$	0.24	0.078	205.32	0.24	0.63			
m_{jj}	0.23	0.074	170.32	0.23	0.62	log		3.8
$Y(jet_1)$	0.22	0.066	173.73	0.22	0.62	abs		
$\eta(jet_1)$	0.22	0.066	173.47	0.22	0.62	abs		
$\|\Delta Y(jets)\|$	0.22	0.079	282.69	0.22	0.61			
$p_T^{bal}(lllvjj)$	0.21	0.056	127.05	0.21	0.58	log	-3.8	-0.2
$E(l_{Z,1})$	0.21	0.067	209.36	0.21	0.60	log	1.4	3.4
$\|\Delta\phi(jets)\|$	0.21	0.061	167.01	0.21	0.59			
$centrality(lll)$	0.21	0.066	240.95	0.21	0.56	log	-1.5	
$centrality(Z)$	0.21	0.062	208.18	0.21	0.56	log	-2.1	
$p_T(l_W)$	0.20	0.065	250.61	0.20	0.57	log		3.0
$E(W)$	0.20	0.060	196.06	0.20	0.59	log		3.4
$E(Z)$	0.18	0.053	162.81	0.18	0.58	log		3.4
$\cos\theta^*(Z)$	0.17	0.045	122.46	0.09	0.48			
$\|\Delta R(l_{Z,2}, l_W)\|$	0.16	0.053	135.98	0.16	0.59			
$p_T(l_{Z,2})$	0.16	0.033	89.08	0.11	0.48	log		2.5
$p_T(jet_1)$	0.15	0.041	178.24	0.14	0.53	log		3.2
$E(jet_1)$	0.15	0.030	64.40	0.15	0.56	log		
$E(l_W)$	0.14	0.035	132.66	0.14	0.55	log	1.4	3.3
$p_T^{bal}(lllv)$	0.14	0.035	128.28	0.14	0.56			
$\cos\theta^*(W)$	0.12	0.026	65.89	0.11	0.53			
$E(jet_2)$	0.11	0.017	36.93	0.11	0.53	log		
$Y(jet_2)$	0.11	0.019	52.57	0.11	0.54	abs		
$\eta(jet_2)$	0.11	0.019	52.24	0.11	0.54	abs		5.0
$m_T(W)$	0.10	0.023	90.73	0.07	0.46	log		2.7
$m(WZ)$	0.10	0.018	47.36	0.10	0.52	log		3.4
$\|\Delta R(l_{Z,1}, l_W)\|$	0.08	0.014	33.87	0.08	0.52			
$p_T(jet_2)$	0.08	0.009	29.60	0.04	0.46	log		3.1
$E(l_{Z,2})$	0.08	0.013	37.20	0.04	0.47	log		3.0
$\eta(W)$	0.08	0.012	25.39	0.08	0.50	abs		
$\|\Delta\phi(l_{Z,1}, l_W)\|$	0.06	0.005	10.86	0.06	0.51			
$\eta(l_{Z,1})$	0.04	0.003	5.68	0.04	0.50	abs		
$Y(Z)$	0.04	0.002	3.85	0.04	0.49	abs		
$\eta(l_{Z,2})$	0.03	0.001	2.36	0.02	0.49	abs		
$\|\Delta\phi(l_{Z,2}, l_W)$	0.02	0.001	1.76	0.02	0.48			
$\eta(l_W)$	0.02	0.001	1.40	0.01	0.48	abs		
$Y(W)$	0.01	0.000	1.58	0.01	0.47	abs		

Increases of the performance of the training sample when including one of these variables while the other is already part of the BDT should only be caused by statistical fluctuations. To avoid such problems the choice of variables that are considered as input during the optimization or the overtraining conditions could be further optimized.
In Table 6.7 the results of the Kolmogorov-Smirnov test among other figures of merit that can be used to quantify the overtraining and the performance of the BDT are summarized for the four final BDTs. The BDT score distributions for the final BDTs are shown in Figure 6.3 to Figure 6.6. The plots on the upper left side compare the normalized BDT score distributions between the test and training sample for the signal and background. Thus they can be used as a visual check for overtraining.
The BDT score distribution of the WLZL BDT is shown in Figure 6.3. For the background a very good agreement between test and training samples can be seen. Fir the signal small differences between test and training dataset can be seen, which might indications slight overtraining. While for large BDT score values the disagreement between test and training sample is generally small, the results from the training sample lay constantly above the ones of the training sample with the exception of the last bin. For small BDT scores this is reversed. Especially for the bins with the lowest BDT score the disagreement between test and training sample for the signal get large. Since the agreement between training and test dataset is still good in the majority of the bins and the bins with high BDT score are most relevant when extracting the significance this is seen as acceptable for this analysis.
When looking at the distributions of the individual backgrounds for the WLZL BDT in Figure 6.3b it can be seen that the non-polarization backgrounds peak as expected at low BDT score values and the LL polarization at high BDT score values. The polarization backgrounds on the other hand are distributed rather equally over the whole BDT score range. This is shown in more detail in Figure 6.3c and Figure 6.3d for only the four polarization combinations. While the distribution of the transverse transverse polarization is shifted a little bit towards small values, no sharp peak in this region can be seen. The distributions of the two mixed samples are both very flat. This means that the BDT is not able to differentiate the polarization backgrounds very well from the LL polarization.
The BDT score distributions for the WLZX BDT are shown in Figure 6.4. The agreement between training and test dataset of the background as shown in Figure 6.4a are very good.

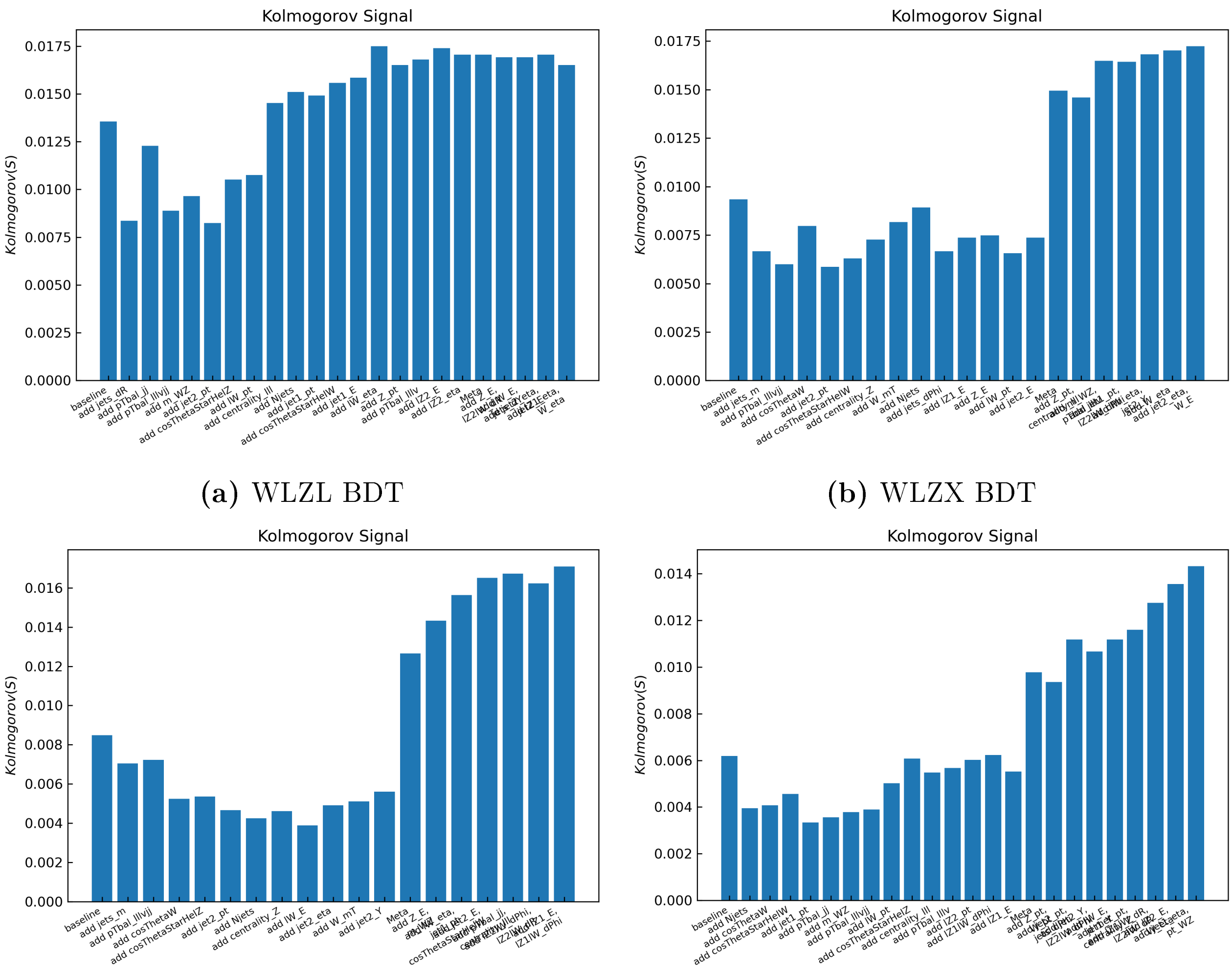
Table 6.5: Metaparameters used in the final BDTs as a result of the metaparameter optimization.

Metaparameter	WLZL vs. rest	WLZX vs. rest	WXZL vs. rest	min1L vs. rest
NTrees	200	900	1300	900
MaxDepth	4	5	5	4
MinNodeSize	2.5%	1.5%	1.5%	2.5%
Shrinkage	0.06	0.015	0.015	0.03

Table 6.6: Variables used in the final BDTs as a result of the optimization. The variables are given in the order in which they are included during the optimization. Variables that are added in the first variable optimization are given above and those that are added in the second variable optimization below the horizontal line. Variables that are added as pairs are indicated by the dashed lines.

	WLZL vs. Bkg	WLZX vs. Bkg	WXZL vs. Bkg	min1L vs. Bkg
1	$\cos\theta(W)$	$p_T(W)$	$p_T(Z)$	m_{jj}
2	$\Delta R(jets)$	m_{jj}	m_{jj}	n_{jets}
3	$p_T^{bal}(jj)$	$p_T^{bal}(lllvjj)$	$p_T^{bal}(lllvjj)$	$\cos\theta(W)$
4	$p_T^{bal}(lllvjj)$	$\cos\theta(W)$	$\cos\theta(W)$	$\cos\theta^*(W)$
5	$m(WZ)$	$p_T(jet_2)$	$\cos\theta^*(Z)$	$p_T(jet_1)$
6	$p_T(jet_2)$	$\cos\theta^*(W)$	$p_T(jet_2)$	$p_T^{bal}(jj)$
7	$\cos\theta^*(Z)$	$centrality(Z)$	n_{jets}	$m(WZ)$
8	$p_T(l_W)$	$m_T(W)$	$centrality(Z)$	$p_T^{bal}(lllvjj)$
9	$centrality(lll)$	n_{jets}	$E(l_W)$	$p_T(l_W)$
10	n_{jets}	$\Delta\phi(jets)$	$\eta(jet_2)$	$\cos\theta^*(Z)$
11	$p_T(jet_1)$	$E(l_{Z,1})$	$m_T(W)$	$centrality(lll)$
12	$\cos\theta^*(W)$	$E(Z)$	$Y(jet_2)$	$p_T^{bal}(lllv)$
13	$E(jet_1)$	$p_T(l_W)$	$E(Z)$	$p_T(l_{Z,2})$
14	$\eta(l_W)$	$E(jet_2)$	$m(WZ)$	$\Delta\phi(l_{Z,1}, l_W)$
15	$p_T(Z)$	$p_T(Z)$	$\eta(l_W)$	$E(l_{Z,1})$
16	$p_T^{bal}(lllv)$	$centrality(lll)$	$p_T(jet_1)$	$p_T(Z)$
17	$E(l_{Z,2}$	$m(WZ)$	$E(jet_2)$	$p_T(W)$
18	$\eta(l_{Z,2}$	$p_T^{bal}(lllv)$	$\cos\theta^*(W)$	$p_T(jet_2)$
19	$E(Z)$	$p_T(jet_1)$	$p_T^{bal}(jj)$	$\Delta\phi(jets)$
20	$\Delta R(l_{Z,2}, l_W)$	$\Delta\phi(l_{Z,2}, l_W)$	$centrality(lll)$	$Y(jet_2)$
21	$E(W)$	$\eta(l_W)$	$\Delta\phi(l_{Z,2}, l_W)$	$\Delta\phi(l_{Z,2}, l_W)$
22	$\Delta Y(jets)$	$Y(jet_2)$	$\Delta R(l_{Z,2}, l_W)$	$E(l_W)$
23	$\eta(jet_2)$	$\eta(W)$	$E(l_{Z,1})$	$Y(jet_1)$
24	$E(jet_2)$	$\eta(jet_2)$	$\Delta\phi(l_{Z,1}, l_W)$	$p_T(MET)$
25	$\eta(l_{Z,1})$	$E(W)$		$centrality(Z)$
26	$\eta(W)$			$\Delta R(l_{Z,1}, l_W)$
27				$\Delta R(l_{Z,2}, l_W)$
28				$E(l_{Z,2}$
29				$\eta(W)$
30				$\eta(jet_1)$
31				$p_T(WZ)$

For the signal a slight slope in the ratio plot is visible indicating mild overtraining. Since the deviations are not very large this is again considered as acceptable for this thesis. For the individual background a similar behaviour as for the WLZL BDT was found. While the non-polarization backgrounds have their maximum at small BDT score values, the polarization backgrounds are not separated well from the WLZX signal.
For both the WLZX BDT and the min1L BDT no significant signs of overtraining can be seen in Figure 6.5 and Figure 6.6a respectively. Large differences between test and training dataset appear only in bins with a low fraction of the signal. As for the previous BDTs no good separation of the polarization backgrounds to the signal can be seen. Especially for the min1L BDT a tendency of the background polarization towards higher BDT score values is visible.



(a) WLZL BDT

(b) WLZX BDT

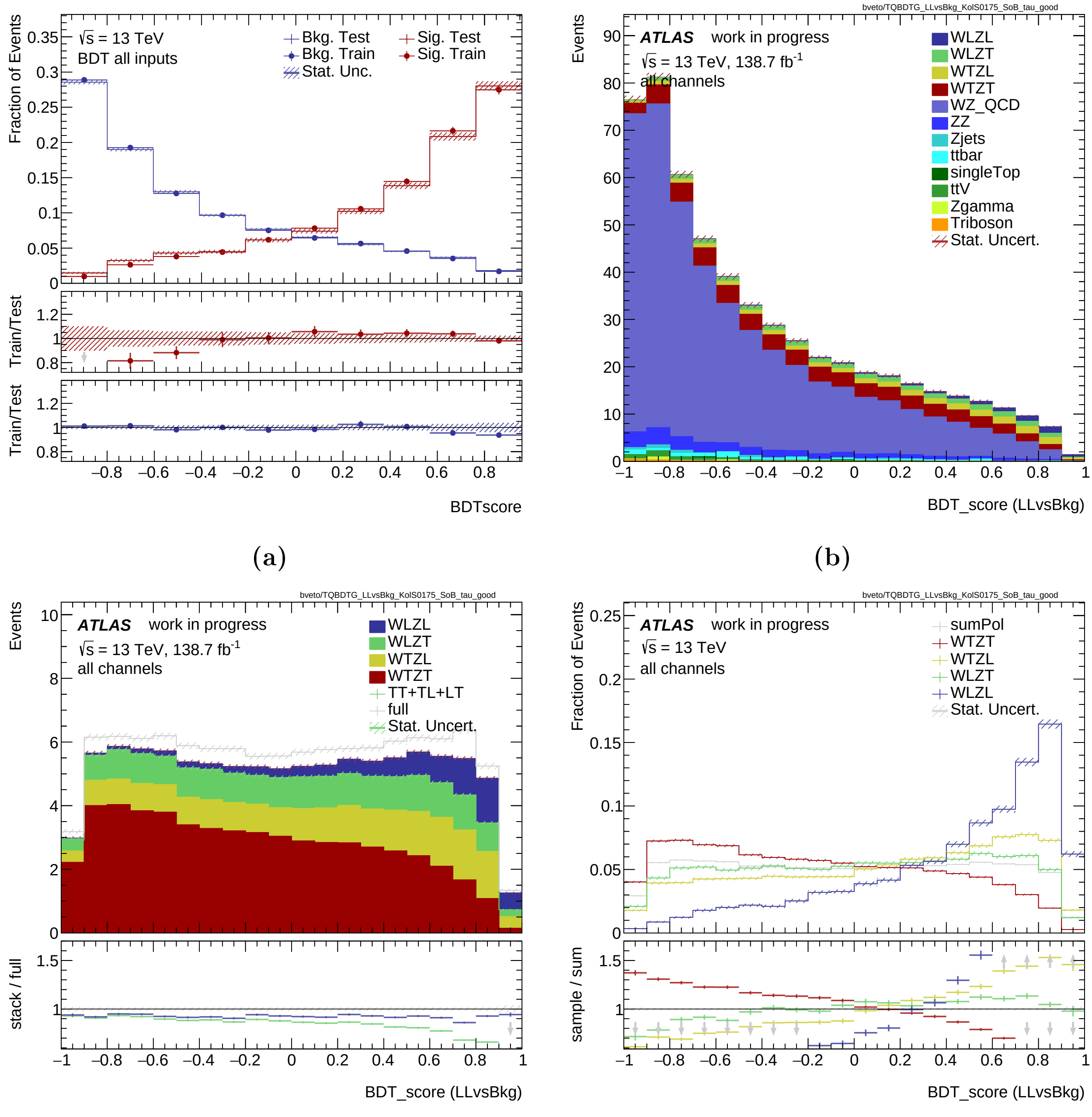
(c) WXZL BDT

(d) min1L BDT

Figure 6.2: Development of output of the Kolmogorow-Smirnow test of the signal during the optimization for (a) the WLZL BDT, (b) the WLZX BDT, (c) the WXZL BDT and (d) the min1L BDT.

Table 6.7: Values of some FoM for the final BDTs. A S (B) indicates that the signal (background) distributions of the training and test dataset are compared. The index te (tr) denotes that the FoM is calculated for the signal and background distributions of the test (training) datasets.

FoM	WLZL	WLZX	WXZL	min1L
Kol(S)	0.0165149	0.0172304	0.0170952	0.0143076
Kol(B)	0.00902377	0.0115445	0.0152551	0.0124431
red. χ^2(S)	1.41599	2.14282	2.58254	1.9991
red. χ^2(B)	1.24878	1.07346	1.67183	1.00327
$Z_{SoB,sys,tr}$	0.729175	1.90604	2.32802	3.40539
$Z_{SoB,sys,te}$	0.760617	1.89593	2.21905	3.44171
$f_{ROC,te}$	0.862914	0.82343	0.838741	0.816185
$f_{N-O,te}$	0.580175	0.502516	0.53291	0.492475
$f_{Sep,te}$	0.415833	0.331429	0.362226	0.317428



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

Figure 6.3: BDT distributions for the WLZL BDT. In (a) the normalized distributions of the signal (red) and background (blue) are shown for the test and training sample. In the subplots the ratios of the training and test dataset for the signal (top) and background (bottom) are given. In (b) the stacked event yields for all signal and background processes are shown. In (c) the distributions of the full sample and the stacked distributions of all four polarization combinations are shown. In the subplot the ratios of the sum of the four polarization combinations to the full sample (blue) and of the sum of WLTZ, WTZL and WTZT to the full sample (green) are shown. In (d) the normalized distributions of the four polarization combinations and the sum of the polarizations are given. The subplot shows the ratios of each of these samples to the sum of the polarizations.

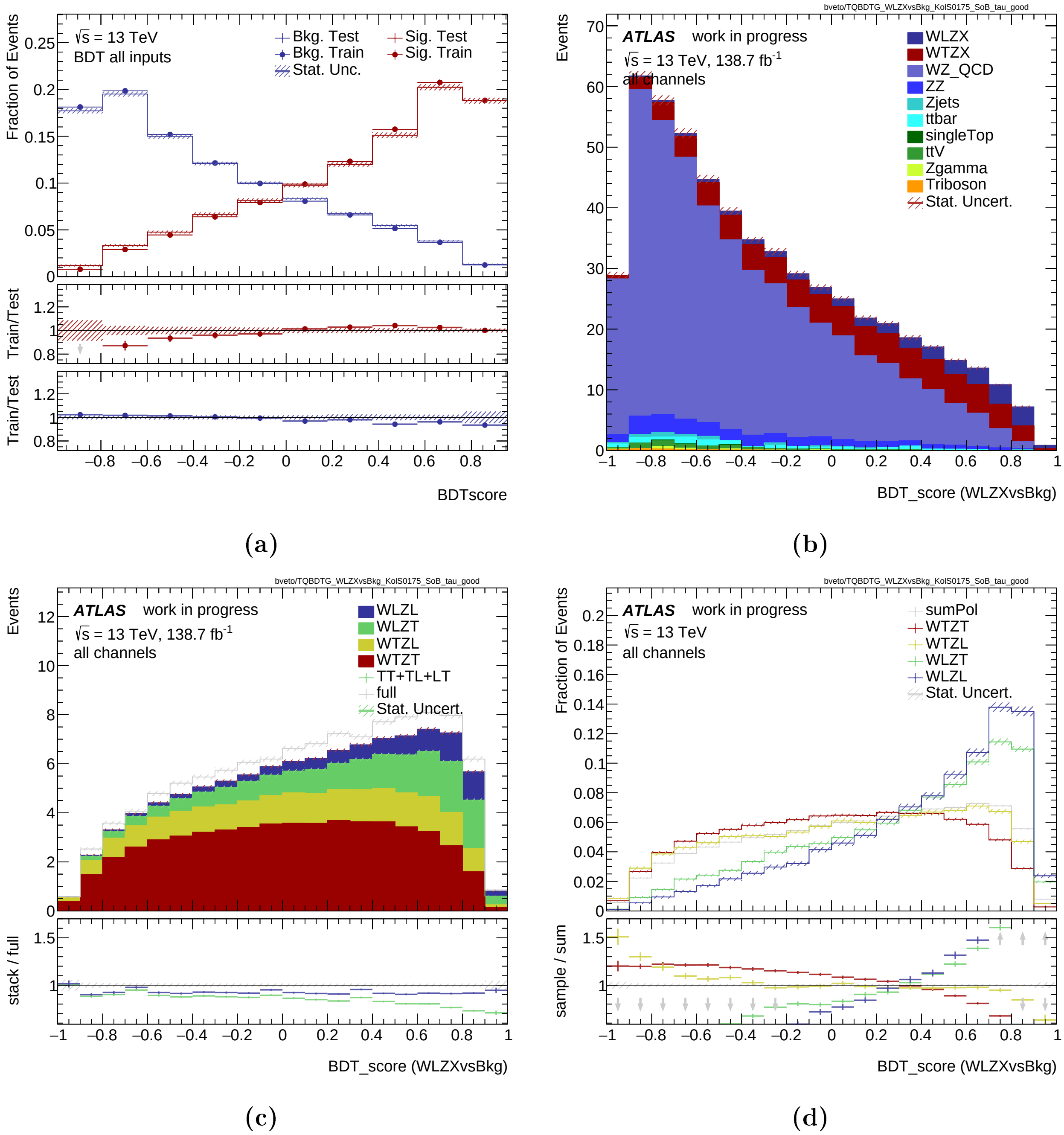


Figure 6.4: BDT distributions for the WLZX BDT. In (a) the normalized distributions of the signal (red) and background (blue) are shown for the test and training sample. In the subplots the ratios of the training and test dataset for the signal (top) and background (bottom) are given. In (b) the stacked event yields for all signal and background processes are shown. In (c) the distributions of the full sample and the stacked distributions of all four polarization combinations are shown. In the subplot the ratios of the sum of the four polarization combinations to the full sample (blue) and of the sum of WLTZ, WTZL and WTZT to the full sample (green) are shown. In (d) the normalized distributions of the four polarization combinations and the sum of the polarizations are given. The subplot shows the ratios of each of these samples to the sum of the polarizations.

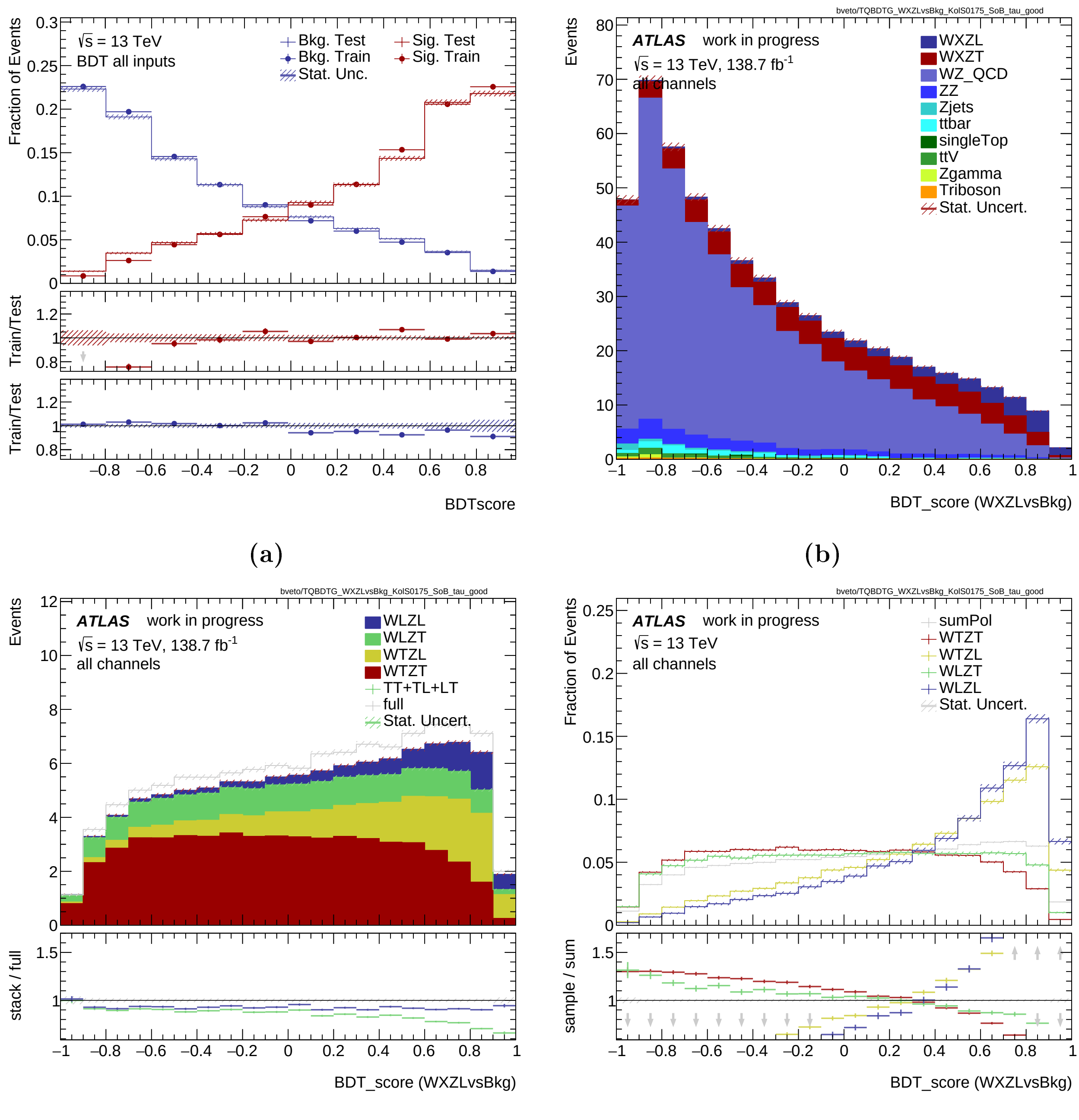


Figure 6.5: BDT distributions for the WXZL BDT. In (a) the normalized distributions of the signal (red) and background (blue) are shown for the test and training sample. In the subplots the ratios of the training and test dataset for the signal (top) and background (bottom) are given. In (b) the stacked event yields for all signal and background processes are shown. In (c) the distributions of the full sample and the stacked distributions of all four polarization combinations are shown. In the subplot the ratios of the sum of the four polarization combinations to the full sample (blue) and of the sum of WLTZ, WTZL and WTZT to the full sample (green) are shown. In (d) the normalized distributions of the four polarization combinations and the sum of the polarizations are given. The subplot shows the ratios of each of these samples to the sum of the polarizations.

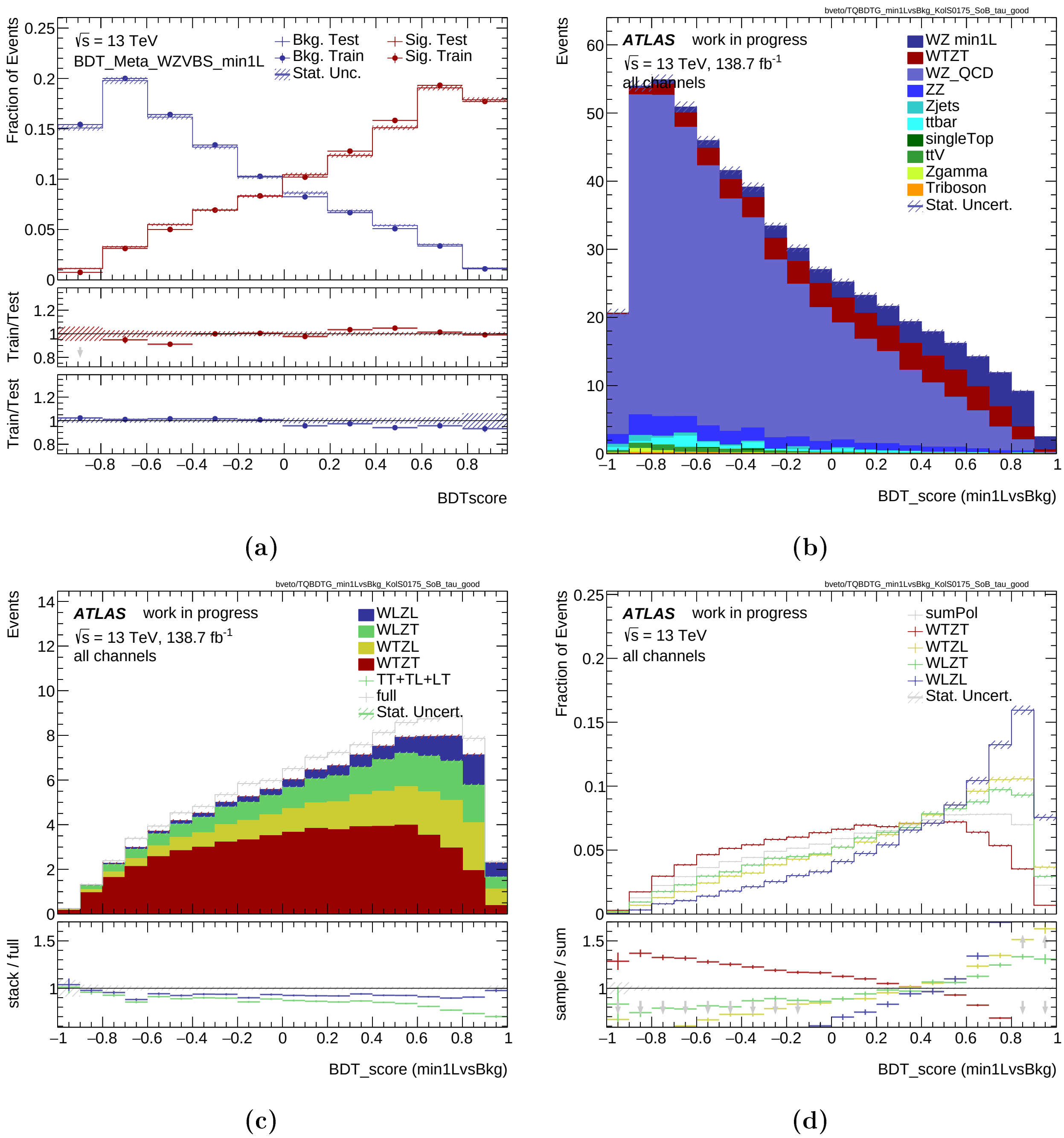


Figure 6.6: BDT distributions for the min1L BDT. In (a) the normalized distributions of the signal (red) and background (blue) are shown for the test and training sample. In the subplots the ratios of the training and test dataset for the signal (top) and background (bottom) are given. In (b) the stacked event yields for all signal and background processes are shown. In (c) the distributions of the full sample and the stacked distributions of all four polarization combinations are shown. In the subplot the ratios of the sum of the four polarization combinations to the full sample (blue) and of the sum of WLTZ, WTZL and WTZT to the full sample (green) are shown. In (d) the normalized distributions of the four polarization combinations and the sum of the polarizations are given. The subplot shows the ratios of each of these samples to the sum of the polarizations.

6.2 Studies of the Sensitivity to the Longitudinal Polarization

In order to see how sensitive we can be to the longitudinal polarization binned maximum likelihood fits using the BDT score distributions were performed to Asimov data. The framework that was used to do the fit is based on libraries by RooFit [35] and RooStat [36]. In the following a brief introduction to the statistical methods is given based on [34] and [37].

6.2.1 Likelihood Function and Significance

The likelihood function $L(\theta_1, ..., \theta_m)$ of m parameters θ_j for n observables x_i is defined as

$$L(\theta_1, ..., \theta_m) = c \cdot \prod_{i=1}^{n} f(x_i|\theta_1, ..., \theta_m). \tag{6.8}$$

Here $f(x_i|\theta_1, ..., \theta_m)$ is the probability density function (pdf) for the measurement of the observable x_i given a theory with the parameters $\theta_1, ..., \theta_m$. The c is an arbitrary constant.
In this study a binned likelihood is used, where the observables are the measured event yields n_i in the individual bins. The parameter of interest is the signal strength μ_S, which scales the amount of signal events. With that the expected number of events ν_i in bin i is:

$$\nu_i = \mu_S \, s_i + b_i. \tag{6.9}$$

Here the s_i and b_i are the signal and event yields in bin i for the signal and the background respectively. A signal strength of 0 corresponds to the background-only hypothesis, where no signal events are expected. When assuming a Poissonian distribution the likelihood function looks like

$$L(\mu_s) = \prod_{i=1}^{n} \text{Pois}(n_i|\mu_S s_i + b_i) = \prod_{i=1}^{n} \frac{(\mu_S s_i + b_i)^{n_i}}{n_i!} \cdot \exp(-(\mu_S s_i + b_i)). \tag{6.10}$$

In this likelihood only the statistical uncertainties of the number of expected events are included. This is the likelihood function that is used for the stat-only fit.
Other parameters, called nuisance parameters (NP), like systematic uncertainties on the backgrounds, can also be included in the likelihood function. If external measurements or other constraints are available, they can introduce an constrained term, based on an assumed pdf $f(\theta^{\text{ext}}|\theta)$ of the parameter θ and the external value θ^{ext}, to the likelihood function. Depending on the constrained term, the nuisance parameters can be categorized as shown in the following:

Unconstrained NP μ_B

If no explicit constrained term for the nuisance parameter is included, the nuisance parameter is denoted as μ_B. In the likelihood function they are treated similar to μ_S, but with a nominal value of 1. They can be used, if no further information about this parameter are available.

Poissonian NP γ

For this NP the constraint term follows a Poissonian distribution. In this study these nuisance parameters are included to account for statistical uncertainties caused by the limited amount of simulated events. For each bin one γ parameter is included, combining the statistical uncertainties of the simulations from all used samples. The γ parameters have a nominal value of 1. To achieve this value for each γ an additional scaling factor τ_i is included.

Gaussian NP α

The constraint term for this NP follows a Gaussian distribution. In this study they are used to include systematic uncertainties on the normalizations of the backgrounds into the likelihood function. The nominal value of these parameters is 0.

With that the expected number of events ν_i in bin i is

$$\nu_i = \mu_S\, s_i(\alpha) + b_i(\alpha, \mu_B). \tag{6.11}$$

With m_i predicted events from the simulation in bin i, the full likelihood can be written as follows:

$$L(\mu_S, \mu_B, \gamma, \alpha) = \prod_{i=1}^{n} Pois(n_i|\gamma_i, \nu_i) Pois(m_i|\gamma_i, \tau_i) \cdot \prod_{\alpha_p \in \alpha} G(0|\alpha_p, 1). \tag{6.12}$$

The likelihood function can be used to build a test statistic. A commonly used test statistic is the likelihood ratio:

$$\lambda(\mu_S) = \frac{L(\mu_S, \hat{\hat{\mu}}_B, \hat{\hat{\gamma}}, \hat{\hat{\alpha}})}{L(\hat{\mu}_S, \hat{\mu}_B, \hat{\gamma}, \hat{\alpha})} \tag{6.13}$$

For $L(\hat{\mu}_S, \hat{\mu}_B, \hat{\gamma}, \hat{\alpha})$, the parameters with the hat denote the maximum likelihood estimators of the corresponding parameter. They are obtained by maximizing the likelihood function. The conditional maximum likelihood parameters in $L(\mu_S, \hat{\hat{\mu}}_B, \hat{\hat{\gamma}}, \hat{\hat{\alpha}})$, denoted by a double hat, are

chosen to maximize the likelihood function for a given μ_S. With that the value of the test statistic is constrained between 0 ans 1.
The actual used test statistic, assuming the background-only hypothesis $\mu_S = 0$, is q_0:

$$q_0 = \begin{cases} -2\ln\lambda(0) & ; \hat{\mu}_S \geq 0 \\ 0 & ; \hat{\mu}_S \leq 0. \end{cases} \tag{6.14}$$

Here the signal strength estimator $\hat{\mu}_S$ is forced to be positive.
The test statistic is used to calculate the p-value. Assuming the background-only hypothesis it is defined as the probability to obtain a result equal or more extreme than what is actually observed, under the assumption of this hypothesis. The p-value can be obtained by integrating the pdf of the test statistic from the observed value to infinity:

$$p_{\mu_S} = \int_{q_{0,obs}}^{\infty} f(q_0|0) dq_0. \tag{6.15}$$

The pdf of the test statistic is normally not known. For only one parameter of interest and a large sample size the approximation [38, 37]

$$f(q_0|0) = \frac{1}{2}\delta(q_0) + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q_0}}\exp\left(-\frac{q_0}{2}\right) \tag{6.16}$$

can be used.
Instead of the p-value often the equivalent significance Z is used. They are connected by the relation

$$p_{\mu_S} = \int_{Z_{\mu_S}}^{\infty} G(x; 0, 1) dx. \tag{6.17}$$

Thus the significance is the distance in units of the standard deviation the observed value would have from the mean, so that the upper-tail probability is equal to p. In particle physics normally for a significance of $Z = 3\ \sigma$ or more one speaks of evidence and with a significance of $Z = 5\ \sigma$ or more of observation.
In this analysis no actual data were used. Instead a fit to Asimov-data was done. For Asimov-data the expected values of all the parameters are used. In this case the Asimov-data are the sum of all the polarization combinations and the non-polarization backgrounds. With that the expected significance can be obtained, which characterizes the sensitivity to the process of interest.

6.2.2 Fit Strategy

For each of the BDTs in section 6.1 several fits using the respective BDT score distributions were performed to Asimov-data. The same splitting into signal and background as for the corresponding BDT is used. Depending on the signal process they are called the WLZL fit, the WLZX fit, the WXZL fit and the min1L fit. For the fit a binning optimization is performed for each BDT distribution individually. The starting point of the binning optimization is the histogram with 20 bins. Starting from the last bin, neighbouring bins are merged if any of the following conditions are fulfilled:

- less than 10 events per bin
- decrease of the significance after merging the bins is less than 0.0015.

The resulting BDT score distributions are shown in Figure 6.7.
For each BDT score distribution four fits were performed using different configurations. For the first three fits, systematic uncertainties on the normalization on the non-polarization backgrounds are included. They are taken from the WZ VBS note [31] and are summarized in Table 6.8. Furthermore γ parameters as explained in the previous section are included. The uncertainties on the WZ QCD, ZZ and ttV backgrounds are the uncertainties of the cross sections of the respective process. The uncertainties of the ttbar, Z+jets and Z+gamma backgrounds result from comparing different methods used to estimate fake leptons. All polarizations that are not part of the signal are combined to a single background. For this background three different approaches are used. The first approach is to include this background as an unconstrained NP. In the second approach it is constrained by assuming a systematic uncertainty of 30% on this background. For the third fit the background polarization is fixed, i.e. it is not considered via a nuisance parameter. This allows to get a limit on the significance. During this thesis these three fits are referred to according to the treatment of the background polarization as follows:

(i) unconstrained

(ii) 30% syst. uncert.

(iii) fixed.

To get an absolute limit of the significance that can be obtained through the fit, a stat-only fit is performed. This means that only the statistical uncertainties that result from the amount of expected events are included. No nuisance parameters are used. This fit is referred to as

(iv) stat-only.

Table 6.8: Systematic uncertainties of the non polarization backgrounds. With exception of the single top uncertainty the uncertainties are taken from [31]. The single top uncertainty is an estimated value.

	syst. unc.
WZ QCD	9.6%
ZZ	14.4%
ttbar	40%
ttV	9.7%
Z+jets	40%
Triboson	25%
Z+gamma	40%
Single top	40%

6.2.3 Fit Results

In the following the results of the fits are shown for all the different signal configurations.

WLZL vs. rest

The results of the WLZL fit are shown below:

- unconstrained : $Z = 0.49\,\sigma$
- 30% syst. uncert. : $Z = 0.67\,\sigma$
- fixed : $Z = 0.84\,\sigma$
- stat-only : $Z = 0.88\,\sigma$

As expected the significance that results from the unconstrained is the smallest and grows the more the backgrounds get constrained. The significance for all fits is very low. Even the stat-only fit results only in a significance smaller than $1\,\sigma$. Thus it can be concluded that there is not enough statistical power to see the longitudinally longitudinally polarized WZ VBS process using Run 2 data.
In Figure 6.8 to Figure 6.10 the correlation plots between the different parameters and the pull plots for the first three fits can be seen. For the fit, where the background polarization is unconstrained (Figure 6.8), a strong anticorrelation between the normalization of the WZ QCD background and the background polarization is visible. The normalization of the WZ QCD background and the signal polarization is strongly correlated. The correlation between the signal and background polarization can be seen in the bottom right corner. They are very strong anticorrelated. This means that simultaneously increasing the background polarization while decreasing the signal polarization and the WZ QCD background or vice versa leads to similar results. This decreases the sensitivity to the signal polarization since a good agreement

with the Asimov-data can also be achieved by increasing just the background polarization. The strong anticorrelation between the signal and background polarization can be understood when looking back at the BDT score plots of the different polarizations in Figure 6.3c and Figure 6.3d. While the shapes of the signal and background polarization are not completely similar, they are also not separated very well.
For the fit, with 30% systematic uncertainty for the background polarization (Figure 6.9) correlations and anticorrelations between the same processes can be found. They are however less strong. No strong (anti-)correlations can be seen for the fit, where the background polarization is fixed (Figure 6.10).
The pull plots look similar for all fits. As for Asimov data expected they still are at their nominal values of 0. The green band in the pull plots show the pre-fit uncertainties. It can be seen that the Gaussian parameter of the WZ QCD background is overconstrained by the fit, i.e. its post-fit uncertainties are smaller than the pre-fit ones. This means that this background is more constrained by the Asimov-data than by the systematic uncertainties. It could be that for the measurement of this background in the other analysis further uncertainties are included, which would lead to larger uncertainties on this background.

WLZX vs. rest

The results of the WLZX fit are shown below:

- unconstrained : $Z = 0.97\,\sigma$
- 30% syst. uncert. : $Z = 1.67\,\sigma$
- fixed : $Z = 2.17\,\sigma$
- stat-only : $Z = 2.32\,\sigma$.

While larger than for the WLZL fit the significances are still small. In case the background polarizations are unconstrained a significance of less than $1\,\sigma$ is expected. Even in the limit of the stat-only fit a significance of less than $3\,\sigma$ is expected. Thus, as for the WLZL process, not enough events to see evidence of this process when using Run 2 data is available.
The correlation plots and pull plots for this process look similar to the WLZL fit. Again a strong anticorrelation between the signal and background polarization is visible which can be explained by the similar shapes of the BDT score distributions.

WXZL vs. rest

The results of the WXZL fit are shown below:

- unconstrained : $Z = 1.34\,\sigma$
- 30% syst. uncert. : $Z = 2.04\,\sigma$
- fixed : $Z = 2.48\,\sigma$
- stat-only : $Z = 2.64\,\sigma$.

The significances are a bit higher than for the WLZX vs. rest fit, indicating that the polarizations of the Z boson can be better separated than the ones of the W boson. However even the stat-only fit results in a significance less than $3\,\sigma$. Thus not enough statistical power to be sensitive to this process is available using Run 2 data.
The correlation and pull plots are similar to the ones of the already shown processes. Again a strong anticorrelation between the signal and background polarization is visible which can be explained by the similar shapes of the BDT score distributions

Min1L vs. rest

The results of the min1L fit are shown below:

- unconstrained : $Z = 1.57\,\sigma$
- 30% syst. uncert. : $Z = 2.89\,\sigma$
- fixed : $Z = 3.64\,\sigma$
- stat-only : $Z = 3.94\,\sigma$

- 25% syst. uncert. : $Z = 3.06\,\sigma$

This fit leads to the highest significances of all the signal configurations. This means that the higher statistic of this signal configuration is more important than the larger shape differences between the signal and background polarization for the other signal configurations. In case that the background polarization is unconstrained or constrained by 30% systematic uncertainty the fit results in a significance less than $3\,\sigma$. In case that the background polarization is fixed, the significance exceeds $3\,\sigma$. To see how much the background polarization needs to be constrained to get to this limit, the statistical uncertainty of the background polarization was varied. With a systematic uncertainty of 25% a significance of about $3\,\sigma$ was reached. Thus evidence of WZ VBS processes with at least one longitudinally polarized boson might be seen when using Run 2 data, if it is possible to restrict the background polarization enough.
The correlation and pull plots are similar to the ones of the already shown processes. Again a

strong anticorrelation between the signal and background polarization is visible which can be explained by the similar shapes of the BDT score distributions

6.2.4 Fit to Pseudo Data

Additionally to the fits to Asimov data, fits to different pseudo data were performed. For these fits instead of the sum of the non-polarization backgrounds and the different polarization combinations, the sum of the non-polarization backgrounds and the full sample was used. Thus in this pseudo data the interference effects between the polarizations are included. They are therefore closer to what real data are expected to look like. The results of the fits are shown below:

WLZL vs. rest

- unconstrained : $Z = 0.60\,\sigma$
- 30% syst. uncert. : $Z = 0.94\,\sigma$
- fixed : $Z = 1.24\,\sigma$
- stat-only : $Z = 1.33\,\sigma$.

WLZX vs. rest

- unconstrained : $Z = 1.06\,\sigma$
- 30% syst. uncert. : $Z = 1.94\,\sigma$
- fixed : $Z = 2.61\,\sigma$
- stat-only : $Z = 2.82\,\sigma$.

WLZL vs. rest

- unconstrained : $Z = 1.47\,\sigma$
- 30% syst. uncert. : $Z = 2.34\,\sigma$
- fixed : $Z = 2.91\,\sigma$
- stat-only : $Z = 3.14\,\sigma$.

WLZX vs. rest

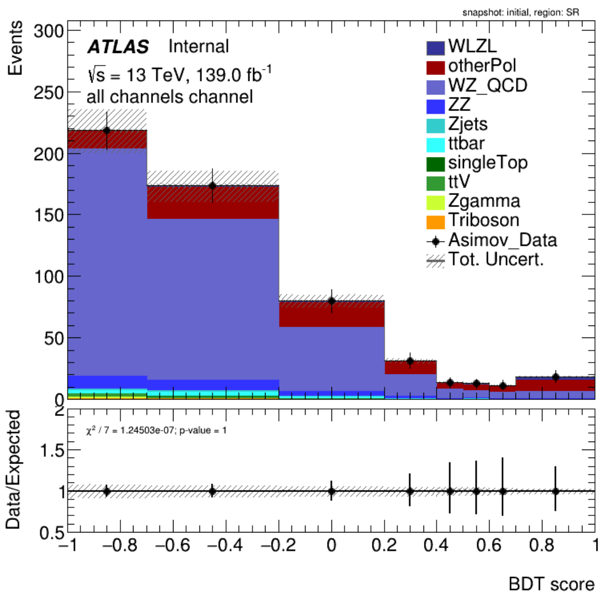
- unconstrained : $Z = 1.71\,\sigma$
- 30% syst. uncert. : $Z = 3.23\,\sigma$
- fixed : $Z = 4.13\,\sigma$
- stat-only : $Z = 4.51\,\sigma$.

The significances of all fits are higher than the one to the Asimov data. This can be understood when looking at the pre-fit plots shown in Figure 6.11 and Figure 6.12. In bins at small BDT score values a good agreement between the pseudo data and the stack of the Monte-Carlo samples is seen. For larger BDT score values the difference gets larger. The largest disagreement is found in the bin with the highest BDT score, which contains the largest amount of signal events. The shape of the interference effects is similar to the shape of the signal distribution. This means that increasing the signal sample leads to a better agreement of the pseudo data and the theoretical prediction by the Monte-Carlo samples and thus a larger fraction of the signal process is preferred by the fit. This is visible in the pre-fit plots shown on the right side of Figure 6.11 and Figure 6.12.

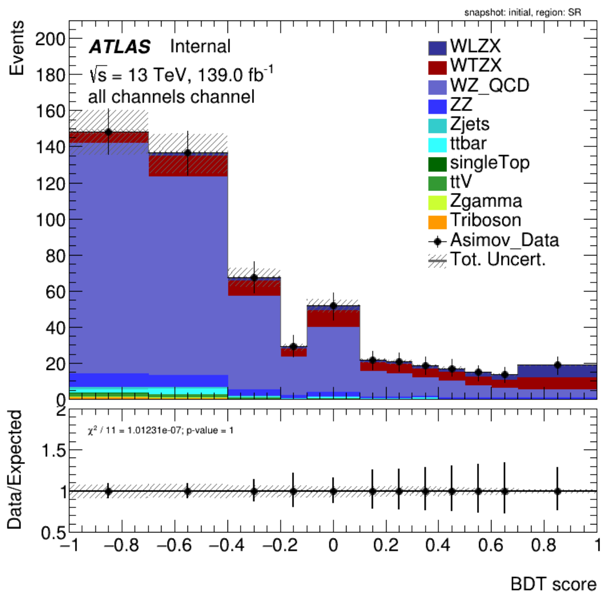
When looking at the BDT score distributions before the binning optimization in Figure 6.3c,

Figure 6.3c, Figure 6.3c and Figure 6.3c no large interference effects, visible as the blue line in the subplot, are seen. Only a very slight tendency to be larger at larger BDT scores is visible. This effect gets increased by the binning optimization since after the optimized binning especially the last bin at the highest BDT score is broader.

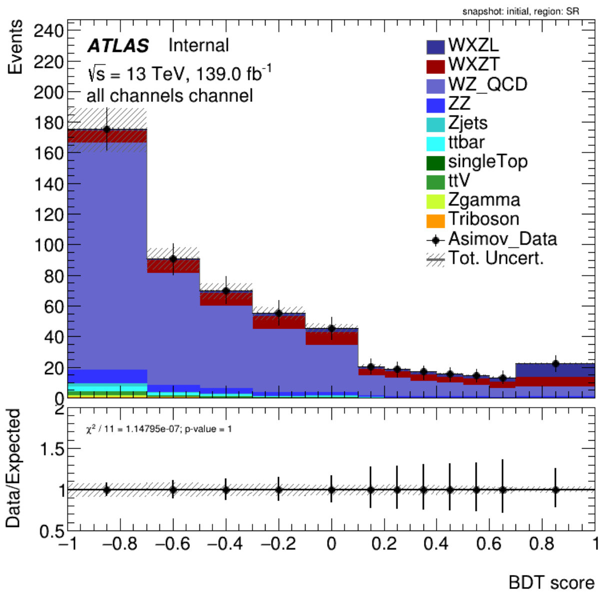
Since the interference effect have a noticeable effect on the expected significance it has to be decided how they are treated in the fit. In the ATLAS community several approaches on how to include the interference effects are discussed. One approach is to consider them as signal, since the interference is a consequence of having the different polarizations. In other approaches they are included in the background as additional uncertainties of the background template or as an additional template. They can also be included as a template to the fit and scale it proportional to the signal.



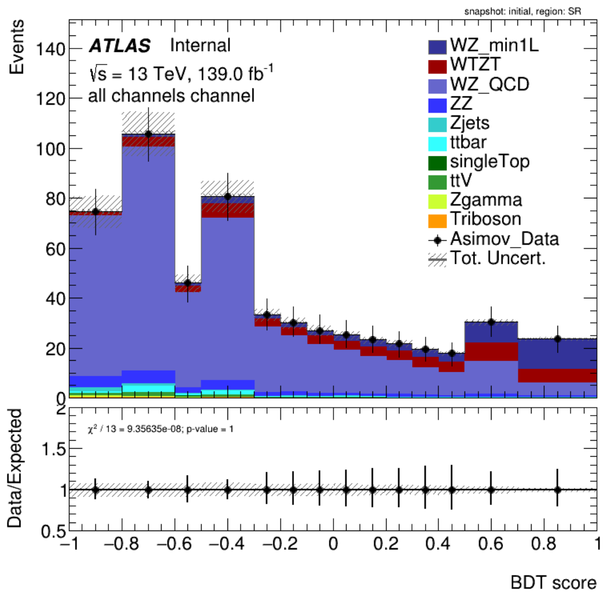
(a) WLZL BDT



(b) WLZX BDT

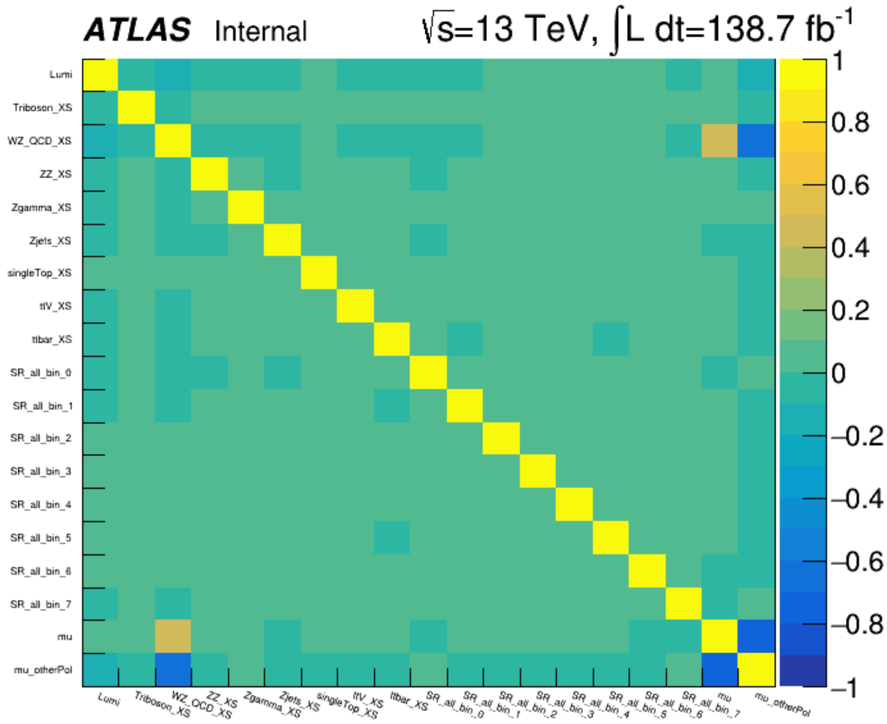


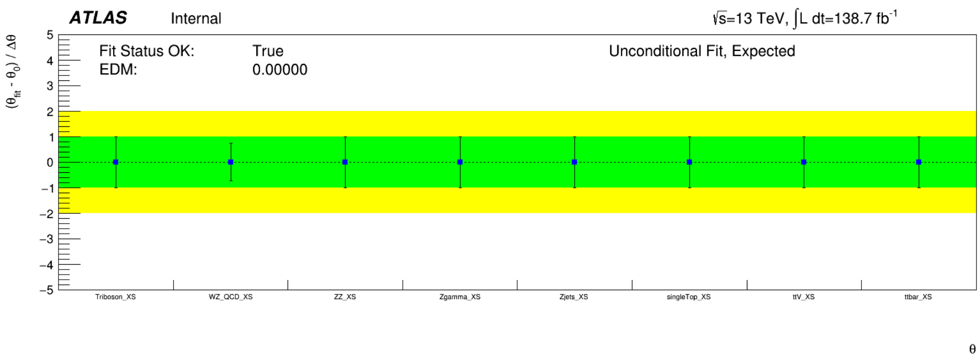
(c) WXZL BDT



(d) min1L BDT

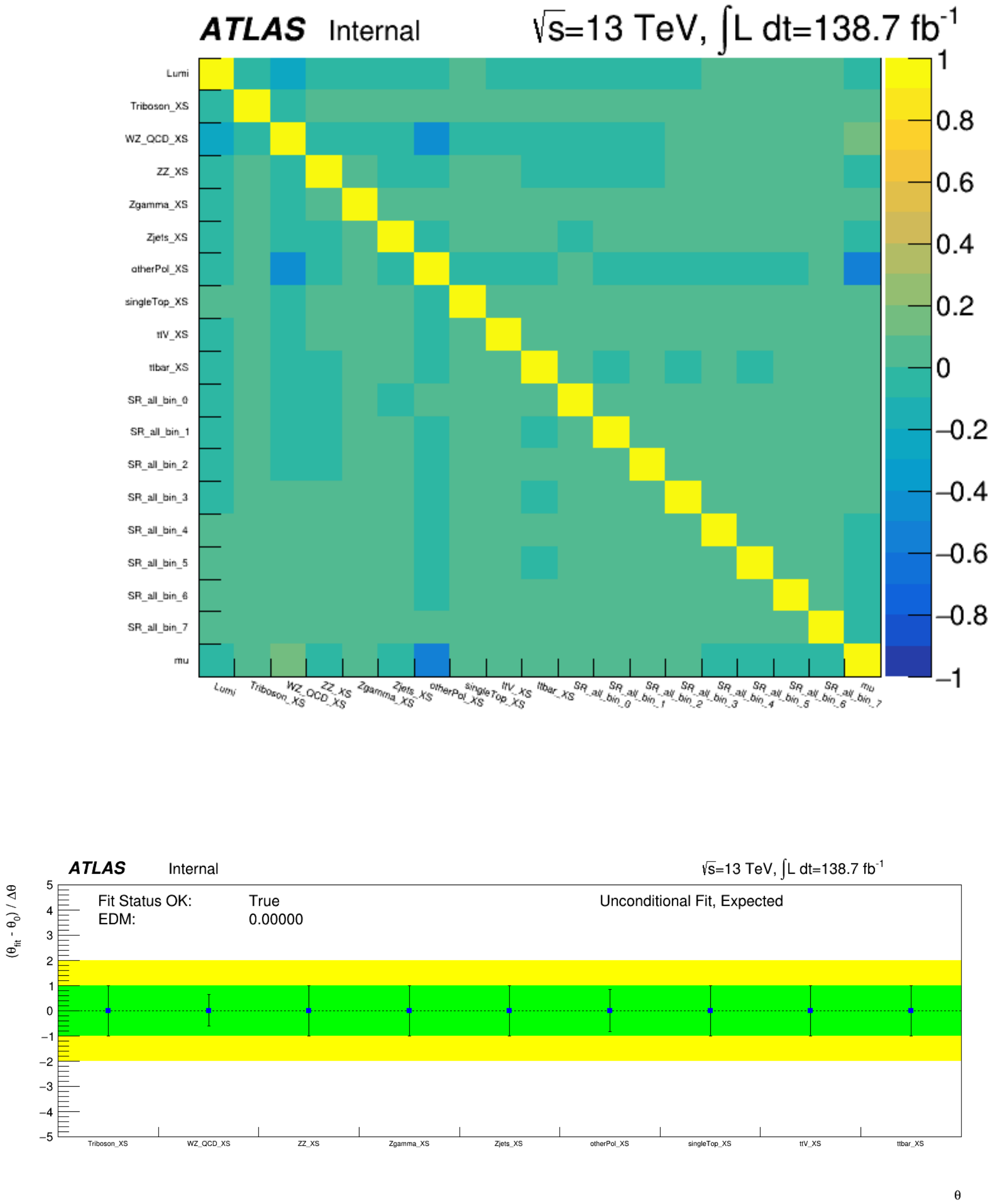
Figure 6.7: BDT score distributions for (a) the WLZL BDT, (b) the WLZX BDT, (c) the WXZL BDT and (d) the min1L BDT after the binning optimization. The Asimov-data are the sum of the non-polarization backgrounds and the sum of the polarizations.





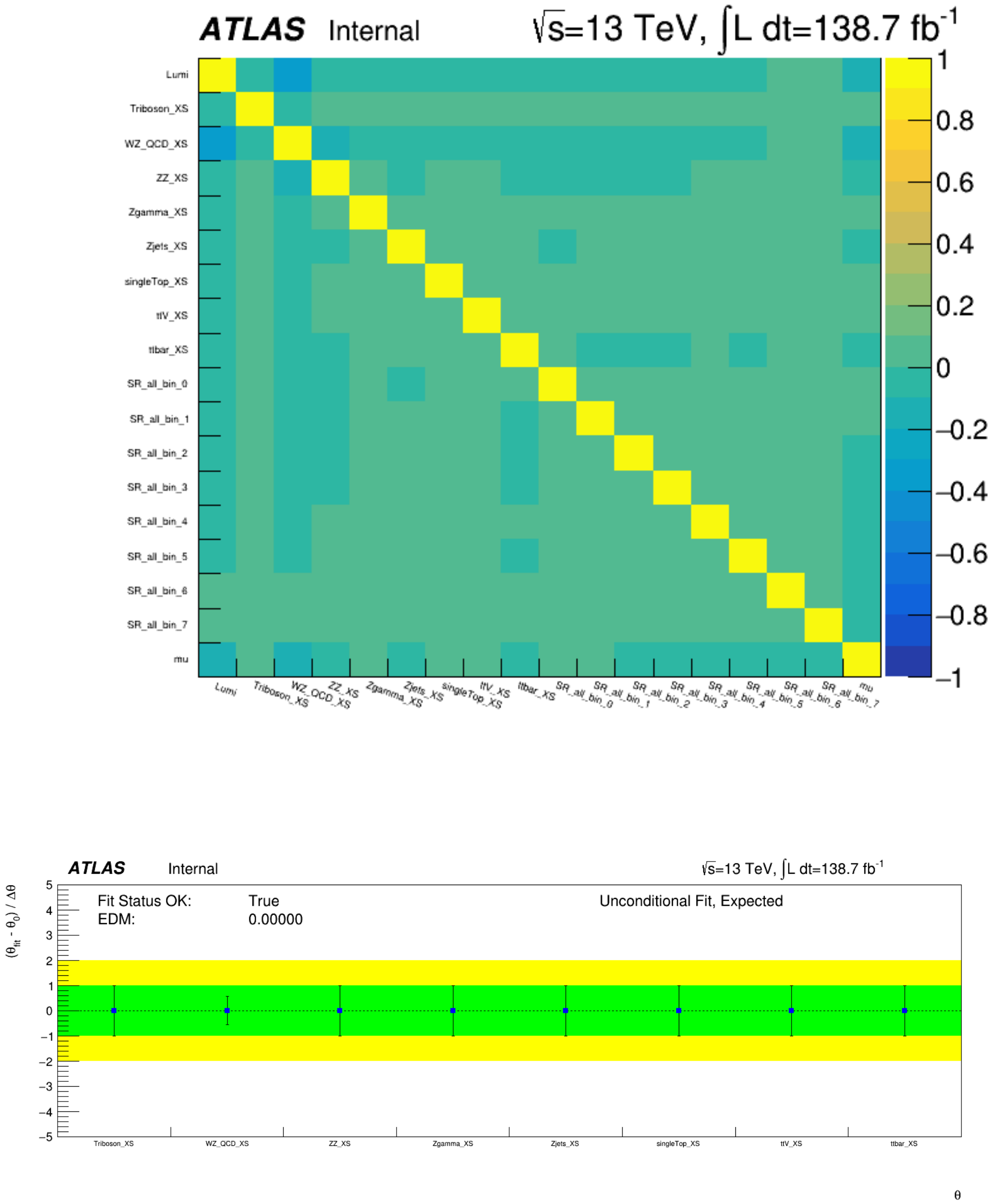
(a)

Figure 6.8: Correlation and pull plot for the WLZL fit, where the polarization background is unconstrained.



(a)

Figure 6.9: Correlation and pull plot for the WLZL fit, where the polarization background has a 30% systematic uncertainty.



(a)

Figure 6.10: Correlation and pull plot for the WLZL fit, where the polarization background is fixed

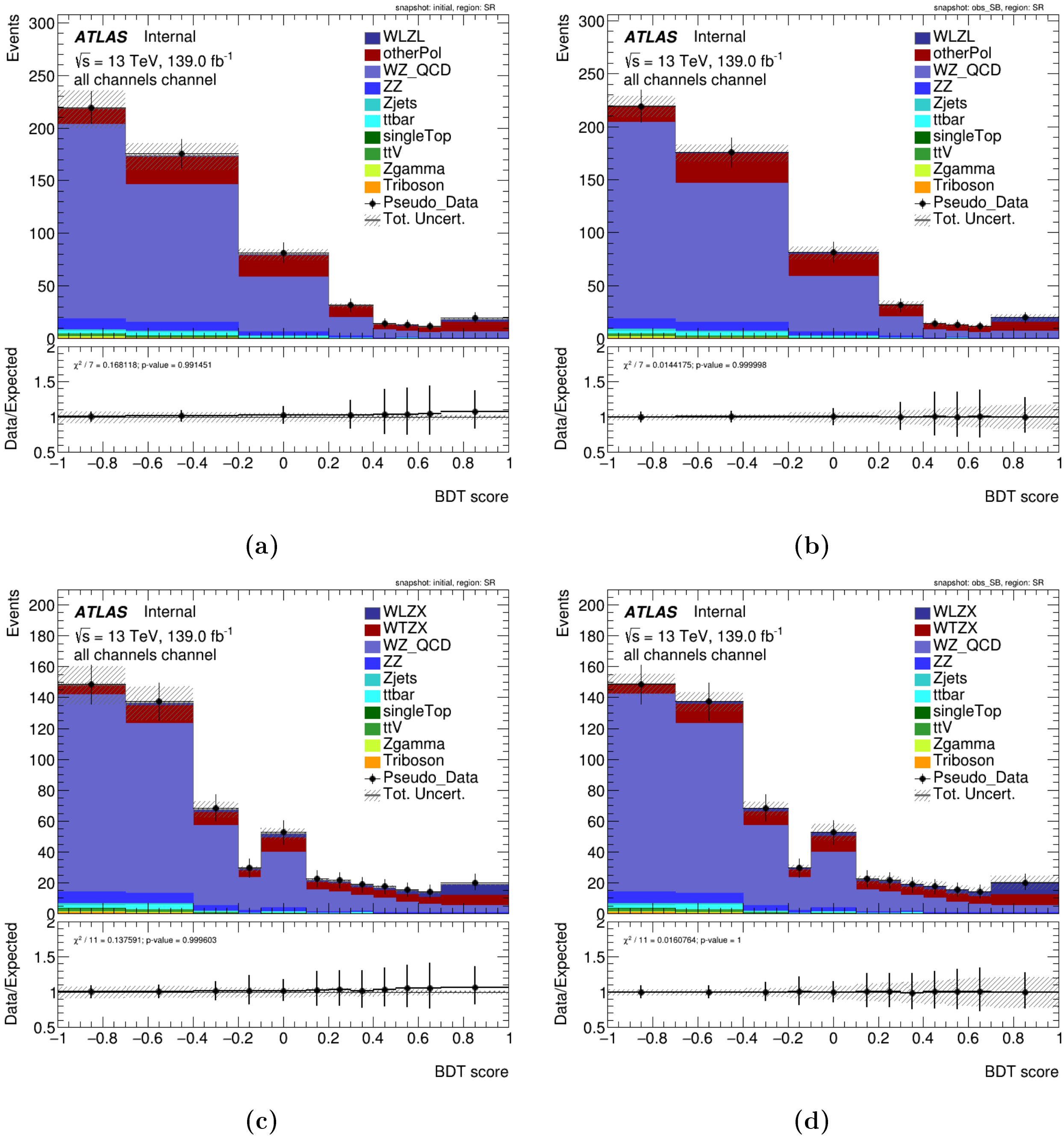
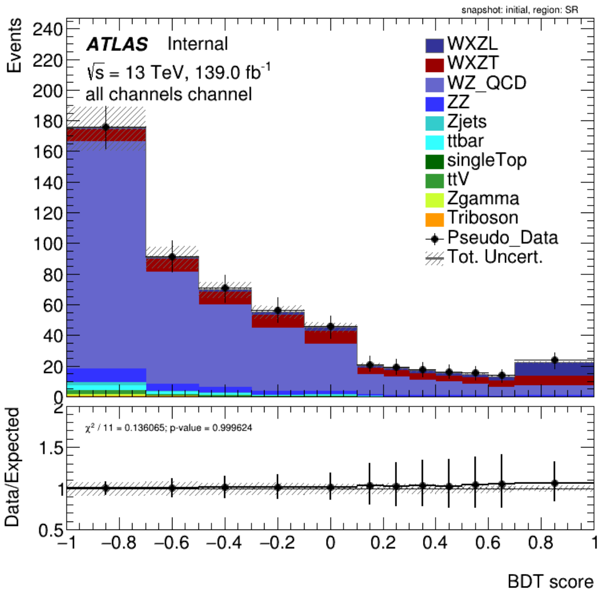
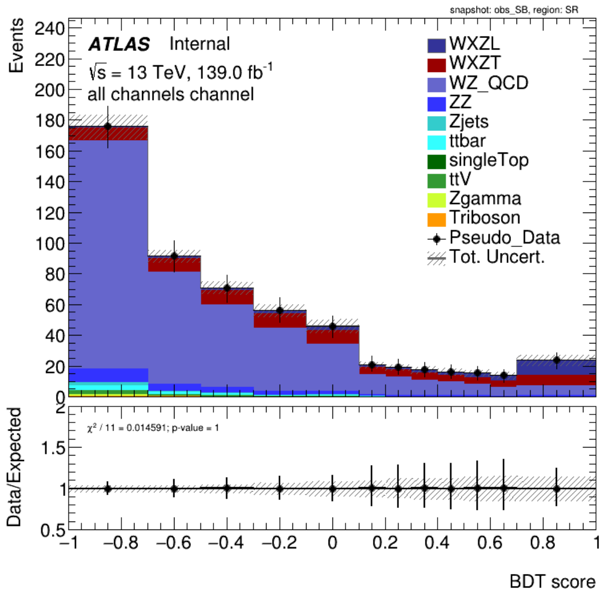


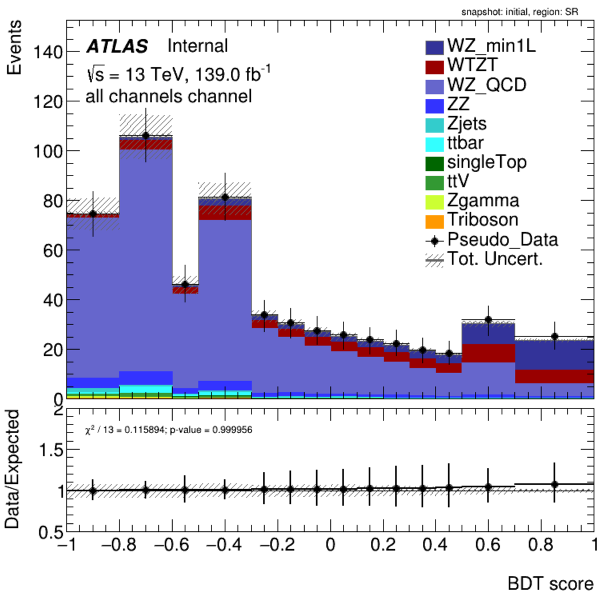
Figure 6.11: BDT score distributions after the binning optimization on the top for the WLZL BDT and on the bottom for the WLZX BDT. The pseudo-data are the sum of the non-polarization backgrounds and the full sample. On the right side the pre-fit and on the left side the post-fit plots are shown for the fit where the background polarizations are fixed.



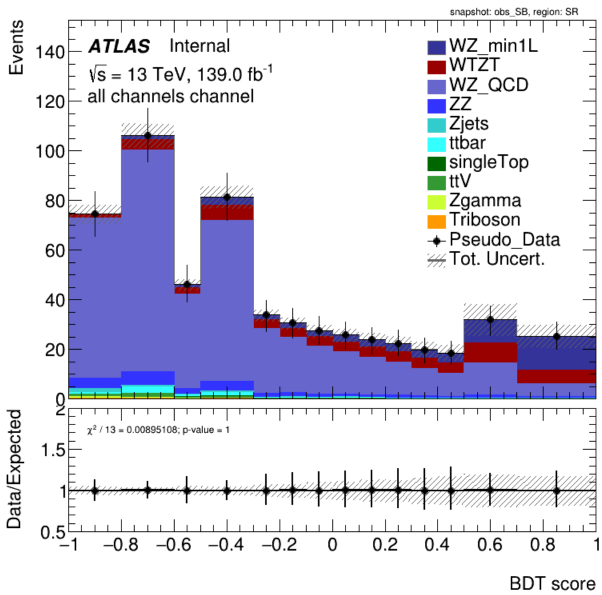
(a)



(b)



(c)



(d)

Figure 6.12: BDT score distributions after the binning optimization on the top for the WXZL BDT and on the bottom for the min1L BDT. The pseudo-data are the sum of the non-polarization backgrounds and the full sample. On the left side the pre-fit and on the right side the post-fit plots are shown for the fit where the background polarizations are fixed.

7 Outlook for Run 3 and the High-Luminosity LHC

In the previous chapter it has been seen, that with the integrated luminosity of 139 fb^{-1} that is achieved during Run2 not enough events to be sensitive to the LL polarization of the bosons are available. With Run 3 which started this year and the high-luminosity LHC (HL LHC) that is planed for further years the integrated luminosity will increase. A plan for further Runs at the LHC is shown in Figure 7.1. Collisions will take place at a higher center of mass energy of 13.6 TeV to 14 TeV. The initial goal for Run 3 was to reach an integrated luminosity of 300 fb^{-1} [39]. Newer predictions estimate an integrated luminosity up to 350 fb^{-1} [39]. For the HL LHC an integrated luminosity of 3000 fb^{-1} is targeted [39]. The ultimate performance is expected to be about 4000 fb^{-1} [39].

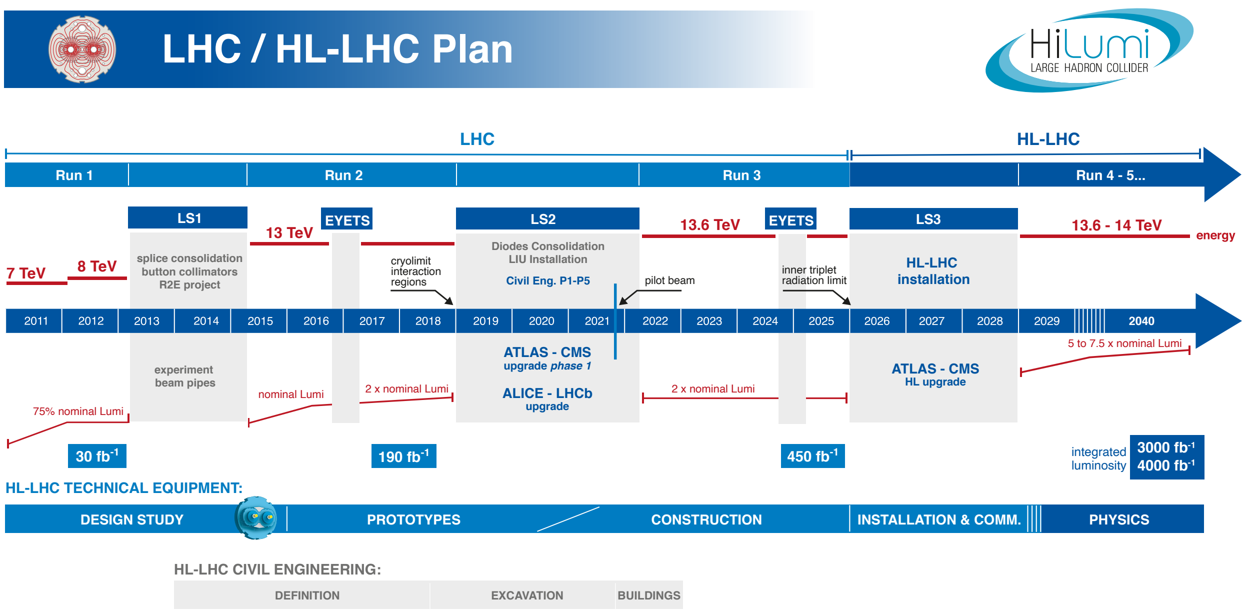


Figure 7.1: Plan for the LHC/HL LHC [40].

7.1 Fits at Higher Luminosities

To get an idea how much an increased luminosity improves the sensitivity to the longitudinal polarization, additional fits at integrated luminosities of 300 fb^{-1}, 3000 fb^{-1} and 4000 fb^{-1} were performed. This was done by increasing the luminosities of the mc16e samples so that the sum of the mc16a, mc16d and mc16e sample reaches the respective value. Other differences for Run 3 and the HL LHC compared to Run 2 like the increased center of mass energy of $\sqrt{s} = 13.6$ TeV and changes of the detectors are not taken into account for the fits. All fits are done equivalently to the previous chapter, using the BDT scores, Asimov-data and the same configurations. The binning optimization was performed anew for each luminosity.
The results of the fits are shown in Table 7.1 to Table 7.4. A visualisation of the results is shown in Figure 7.2.

WLZL vs. rest

For an integrated luminosity of 300 fb^{-1} the significances of all WLZL fits (see Table 7.1) are below $3\,\sigma$. The stat-only fit results in a significance of only $1.34\,\sigma$. Thus even when an integrated significance up to 350 fb^{-1} is reached, it is not to be expected to get sensitive to the longitudinal longitudinal polarization with Run 3 data. Fits where the background polarization is unconstrained or constrained by having 30% systematic uncertainties result for all luminosities in a significance less than $3\,\sigma$. In case that the background polarization is fixed, an expected significance of $3.61\,\sigma$ is reached for a luminosity of 3000 fb^{-1}. Thus it might be possible to see evidence of this process at the HL LHC.
The results of a previous study of the sensitivity to the longitudinal longitudinal polarization are shown in Figure 7.3. The details about this study can be found in [41]. No multivariate techniques were used to separate the signal polarization but a fit was done using the unrolled 2D $(\Delta R(jj), m_{jj})$ distribution. For an integrated luminosity of 3000 fb^{-1} a significance of about $2.5\,\sigma$ for the stat-only fit and about $1.7\,\sigma$ for the fit using systematic uncertainties are expected. In comparison the corresponding fit done in this study resulted in significances of $4.4\,\sigma$ of the stat-only fit and $2.1\,\sigma$ to $3.6\,\sigma$ for the fits that include systematic uncertainties. Thus despite other differences between the two fits, it can be seen that using multivariate techniques clearly improve the sensitivity to the longitudinal longitudinal polarization.

WLZX vs. rest

For an integrated luminosity of 300 fb^{-1} the WLZX fit (see Table 7.2) in the cases unconstrained and 30% syst. uncert. leads to a significance less than $3\,\sigma$. For the fixed fit a significance of $3.12\,\sigma$ is reached. Especially since the luminosity of Run 3 might be a bit higher than this, evidence of this process might be seen, if the background polarization can be con-

Table 7.1: Significances that result from the different WLZL fits for different integrated luminosities.

luminosity [fb^{-1}]	Significance [σ]			
	unconstrained	30 % sys	fixed	stat-only
139	0.49	0.67	0.84	0.88
300	0.75	0.94	1.27	1.34
3000	2.17	2.30	3.61	4.39
4000	2.42	2.54	4.00	5.07

strained enough. For an integrated luminosity of 3000 fb^{-1} all fits result in significances larger than $3\,\sigma$. For the fit where the background polarization is fixed, significances above $5\,\sigma$ are reached. Thus observation might be possible.

Table 7.2: Significances that result from the different WLZX fits for different integrated luminosities.

luminosity [fb^{-1}]	Significance [σ]			
	unconstrained	30 % sys	fixed	stat-only
139	0.97	1.67	2.17	2.32
300	1.37	2.08	3.12	3.40
3000	3.43	3.89	8.10	10.71
4000	3.74	4.17	8.85	12.36

WXZL vs. rest

Similar to the WLZX fit the WXZL fit (seeTable 7.3) results for an integrated luminosity of 300 fb^{-1} in the fixed and 30% syst. uncert. case in significances less than $3\,\sigma$. In the fixed case a significance of more than $3\,\sigma$ is reached. Thus evidence of this process might be seen using Run 3 data, if the background polarization can be sufficiently constrained. For an integrated luminosity of 3000 fb^{-1} all fits give an significance of at least $6\,\sigma$. Thus observation of this process might be possible at the HL LHC.

Table 7.3: Significances that result from the different WXZL fits for different integrated luminosities.

luminosity [fb^{-1}]	Significance [σ]			
	unconstrained	30 % sys	fixed	stat-only
139	1.34	2.04	2.48	2.64
300	2.07	2.78	3.73	4.01
3000	6.00	6.42	10.51	12.96
4000	6.68	7.07	11.67	14.96

Min1L vs. rest

For an integrated luminosity of 300 fb^{-1} the min1L fit (see Table 7.4) in the case 30% syst. uncert. leads to a significance of $3.74\,\sigma$. In the fixed case a significance higher than $5\,\sigma$ is reached. Thus evidence and maybe observation of this process can potentially be seen with the Run 3 data. With significances larger than $6\,\sigma$ for an integrated luminosity of 3000 fb^{-1} for all fits, observation at the HL LHC is likely possible.

Table 7.4: Significances that result from the different min1L fits for different integrated luminosities.

luminosity [fb^{-1}]	Significance [σ]			
	unconstrained	30 % sys	fixed	stat-only
139	1.57	2.89	3.64	3.94
300	2.46	3.74	5.35	5.93
3000	6.82	7.55	14.31	18.96
4000	7.50	8.18	15.81	21.89

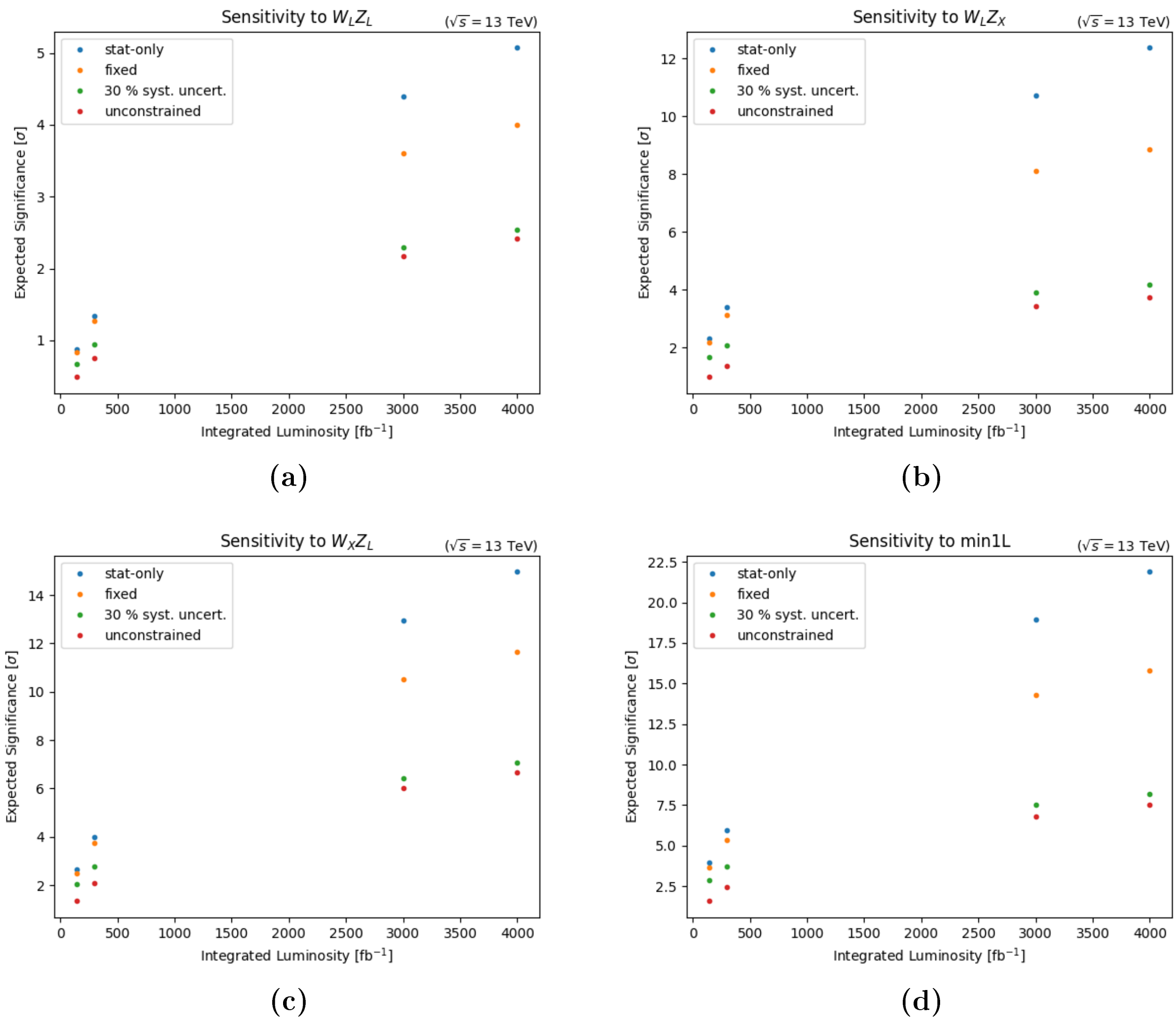


Figure 7.2: Significances that result from the fit with (a) WLZL ,(b) WLZX, (c) WXZL and (d) at least one longitudinally polarized boson as signal for different integrated luminosities. For the stat-only fit only statistical uncertainties are used. For the other fits systematic uncertainties on the non-polarization backgrounds are applied. The background polarization is unconstrained (orange), constrained by having 30% systematic uncertainty (green) or are fixed (red).

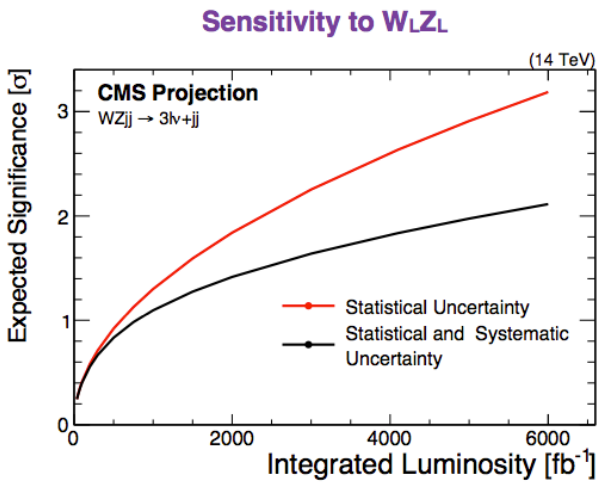


Figure 7.3: Expected significances from the study done in [41].

7.2 Cross Sections for $\sqrt{s} = 13.6$ TeV

To see the effect of the increased centre of mass energy for Run 3 and the HL LHC, samples with electrons and muons in the final state of the hard process were produced at the the new value $\sqrt{s} = 13.6$ TeV. The cross sections that are obtained from this sample are compared to the ones from the samples produced at $\sqrt{s} = 13$ TeV. This can be seen in Table 7.5. An increase of the cross section for the new samples of about 9%-10% can be observed. The increase of the cross section of the longitudinal longitudinal polarization seems slightly larger than the one of the transverse transverse polarization. Since the increase is still very similar for all samples, the fit should behave comparably for the different polarization combinations at the higher center of mass energy. However it also has to be investigated how the non-polarization backgrounds are effected by the higher center of mass energy.

Table 7.5: Cross sections of the polarization samples with electrons and muons in the final state of the hard process produced at $\sqrt{s} = 13$ TeV and $\sqrt{s} = 13.6$ TeV. In the right column the ratio of the cross sections of these samples are shown.

sample	σ [pb] for $\sqrt{s} = 13$ TeV	σ [pb] for $\sqrt{s} = 13.6$ TeV	13 TeV / 13.6 TeV
WLZL	0.001442248	0.0015772	0.91
WLZL OFM	0.000274386	0.0002989	0.92
WLZL OFP	0.00044687	0.0004855	0.92
WLZL SFM	0.000274267	0.0003019	0.91
WLZL SFP	0.000446725	0.0004909	0.91
WLZT	0.003660874	0.0040407	0.91
WLZT OFM	0.00068858	0.0007632	0.90
WLZT OFP	0.00114179	0.001258	0.91
WLZT SFM	0.000688444	0.0007635	0.90
WLZT SFP	0.00114206	0.001256	0.91
WTZL	0.004486316	0.0049505	0.91
WTZL OFM	0.000814423	0.000904	0.90
WTZL OFP	0.00142846	0.001574	0.91
WTZL SFM	0.000814353	0.0008995	0.91
WTZL SFP	0.00142908	0.001573	0.91
WTZT	0.01141561	0.012624	0.90
WTZT OFM	0.0020787	0.002307	0.90
WTZT OFP	0.00363	0.004002	0.91
WTZT SFM	0.00207874	0.002305	0.90
WTZT SFP	0.00362817	0.00401	0.90
sumPol	0.021005048	0.0231924	0.91

8 Summary and Outlook

In this thesis polarization in the vector boson scattering of a W and a Z boson at the LHC with the ATLAS detector was studied. Since without the Higgs boson the scattering of longitudinally longitudinally polarized bosons leads to unphysical predictions, this process can be used to probe details of the mechanism of EWSB. In this thesis samples with polarization information of the vector bosons were produced with Madgraph, corresponding to the full dataset of Run 2 with a center of mass energy $\sqrt{s} = 13$ TeV and an integrated luminosity of 139 fb^{-1}. The polarization was defined in the rest frame of the W and Z boson.
To validate these samples and to identify variables that show sensitivity to the longitudinal polarization a study of these samples was done at truth level. Effects of the different approximations were investigated. The overall effect of the approximations were found to be reasonably small. A comparison between the decay angle distributions of the W and Z boson to the analytical function was performed for the case of two longitudinally polarized bosons. While already a good closure was found, matrix element particles could be used for further studies to understand small deviations. No unexpected behaviour was found in the distributions of the investigated variables. However interference effects of more than 10% are observed in the distributions of some variables. To avoid difficulties in the interpretation of the interference effects those variables were excluded from the further analysis. In a qualitative investigation of the variables the transverse momentum balance of the two jets $p_T^{bal}(jj)$, the transverse momentum of the diboson system $p_T(WZ)$ and the absolute value of the azimuthal angle difference of the two jets $|\Delta\phi(jj)|$ were found to be most sensitive to the longitudinal longitudinal polarization. For the polarization of the single bosons the cosine of the decay angle as well as the transverse momenta and pseudorapidities of the bosons and the corresponding leptons showed some sensitivity to the longitudinal polarization.
A short study was done, with samples in which the polarization is defined in the parton-parton rest frame. A smaller fraction of the longitudinal longitudinal polarization was found. Since the sensitivity of some variables to the different polarizations were increased or decreased, a further study could be interesting.
An official sample for the polarization defined in the rest frame of the W and Z boson was requested and the study was continued at reconstruction level. Additional backgrounds to the WZ EW process were included. With that four BDTs were trained with the four signal configurations WLZL, WLZX, WXZL and at least one longitudinally polarized bosons. The

BDT outputs were used to perform maximum likelihood fits to Asimov data to extract the expected significance for the different signal processes. Besides a stat-only fit, fits that include systematic uncertainties on the normalizations of the non-polarization backgrounds and different constraints on the background polarization were performed. For the signal configurations with two longitudinally polarized bosons and the one with the mixed signal configurations all fits lead to significances of less than $3\,\sigma$. The highest significances obtained with the stat-only fit are $0.88\,\sigma$, $2.32\,\sigma$ and $2.64\,\sigma$ for WLZL, WLZX and WXZL as signal respectively. Thus especially for the longitudinal longitudinal polarization it is unlikely to be observe this process with Run 2 data. For the signal process with at least one longitudinally polarized boson a fit, where the background polarization is constrained by 25% systematic uncertainty, results in an expected significance of about $3\,\sigma$. Thus it might be possible to see evidence of this process when using Run 2 data.

For an outlook for Run 3 and the High Luminosity LHC a study with increased luminosities was done. For the WLZL signal significances over $3\,\sigma$ could be obtained for a fit where the background polarizations are fixed, for an integrated luminosity of 3000 fb^{-1}. However even for a significance of 4000 fb^{-1} none of the fits with less constraints of the background polarization results in a significance above 3σ. For the mixed polarization a fit where the background polarization are fixed resulted in significances above $3\,\sigma$ for an integrated luminosity of 300 fb^{-1}. Thus evidence of these processes might be seen when using Run 3 data when the background polarizations can be sufficiently constrained. The increase of the center of mass energy from 13 TeV to 13.6 TeV showed the same increase of the cross section of about 9% for all polarization combinations.

For further studies Deep Neural Networks could be used to further increase the separation of the the signal polarization from the backgrounds. An optimization of the input variables, the overtraining conditions and the range of the input variables of the BDTs might increase the performance of the BDTs. The variables $\cos\theta^*(W)$ and $\cos\theta^*(Z)$ can be used to create categories in which the different polarization combinations are enlarged. Simultaneous fits in such categories might help to reduce the correlations between the signal and background polarizations and improve the sensitivity. While due to the small number of events this might not be possible for the longitudinal longitudinal polarization, it could be tried to perform such fits for the other signal configurations or at higher integrated luminosities.

For further studies of the polarizations also polarization samples produced by other Monte-Carlo generator like Sherpa or PHANTOM could be used. For the polarization samples produced with Madgraph QCD scale uncertainties could be derived by varying the renormalization and factorization scales as done in [31]. Furthermore the differences observed when defining the polarizations in the two different definitions of the parton parton frame could be further investigated.

A Mathematical Definitions

Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{A.1}$$

γ-matrices (Dirac representation)

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \tag{A.2}$$

$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} \qquad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{A.3}$$

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{A.4}$$

B Locally Produced Samples

B.1 Job Options

Job options for the locally produced samples. The samples are split according to the flavour of the final state leptons and the charge of the W boson denoted as OFM, OFP, SFM, SFP. Here OF denotes samples in which the lepton associated to the W boson has a different flavour than the ones from the Z boson and SF samples in which all leptons have the same flavour. The M and P stand for minus and plus and denote the charge of the W boson and the associated lepton. Furthermore the samples are split in those were only electrons and muons are allowed in the final state and those were at least one of the leptons in the final state has to be τ lepton. The complete job option is only given for the OFM sample where both bosons are longitudinally polarized with e and μ in the final state (WLZL OFM) and in which the polarization is defined in the helicity frame (`me_frame = 3,4,5,6`). For the other samples only the relevant differences to this job option are stated.

B.1.1 Polarized Samples in the Helicity Frame

Samples with only e and μ in the final state

WLZL OFM (e, μ):

```
from MadGraphControl.MadGraphUtils import *

evgenConfig.generators = ["MadGraph", "Pythia8", "EvtGen"]

evgenConfig.description = 'MadGraph_W0Z0_lvlljj_OFM_EW6_masslessLeptons'
evgenConfig.keywords = ['SM', 'diboson', 'WZ', 'electroweak', '3lepton', 'VBS']
evgenConfig.contact = ['anna.sophie.tegetmeier@cern.ch']

runName='run_01'
process_total=None

evgenConfig.inputfilecheck = runName
```

```
#GridPack
do_gridpack_only = not is_gen_from_gridpack()

gridpack_mode=True
gridpack_dir='madevent/'
gridpack_compile=False
nevents = 5000 if do_gridpack_only else 12000

MADGRAPH_RUN_NAME = runName
MADGRAPH_GRIDPACK_LOCATION = gridpack_dir
MADGRAPH_CATCH_ERRORS=False


beamEnergy=-999
if hasattr(runArgs,'ecmEnergy'):
    beamEnergy = runArgs.ecmEnergy / 2.
else:
    raise RuntimeError("No center of mass energy found.")

#Fetch default LO run_card.dat and set parameters
extras = {'auto_ptj_mjj': False, 'ptl': 4.0, 'ptj': 15.0, 'drll': 0.2, 'ptb': 15.0,
    'dral': 0.1, 'mmll': 4.0, 'asrwgtflavor': 5, 'drjl': 0.2, 'systematics_program':
    'systematics', 'drjj': 0.4, 'systematics_arguments': "['--mur=0.5,1,2',
    '--muf=0.5,1,2',
    '--pdf=errorset,NNPDF30_nlo_as_0119@0,NNPDF30_nlo_as_0117@0,CT14nlo,MMHT2014nlo68clas118,
    PDF4LHC15_nlo_30_pdfas']", 'maxjetflavor': 5, 'me_frame': '3,4,5,6', 'pdlabel':
        "'lhapdf'", 'event_norm': 'sum', 'etal': 5.0, 'ptheavy': 2.0, 'use_syst':
        True, 'cut_decays': True, 'lhaid': 260000, 'etaj': 5.5}

extras["gridpack"] = '.true.' if do_gridpack_only else ".false."
extras["nevents"] = int(nevents)

process = """
    import model sm-no_masses

    define p = g u c d s u~ c~ d~ s~ b b~
    define j = g u c d s u~ c~ d~ s~ b b~

    generate p p > w-{0} z{0} j j QCD=0, w- > e- ve~, z > mu+ mu- @0
    add process p p > w-{0} z{0} j j QCD=0, w- > mu- vm~, z > e+ e- @0
```

```
    output -f

"""

process_dir = new_process(process)
modify_run_card(process_dir=process_dir, run_card_backup='run_card_backup.dat',
                settings=extras)
print_cards()
generate(required_accuracy=0.001,process_dir=process_dir,grid_pack=gridpack_mode,
      gridpack_compile=gridpack_compile, runArgs=runArgs)

output_suffix = gridpack_dir if do_gridpack_only else runName
runArgs.outputTXTFile = output_suffix+'._00001.events.tar.gz'
arrange_output(process_dir=process_dir,lhe_version=3,saveProcDir=True,runArgs=runArgs)

if not do_gridpack_only:

    runArgs.inputGeneratorFile=output_suffix+'._00001.events.tar.gz'

    include("Pythia8_i/Pythia8_A14_NNPDF23LO_EvtGen_Common.py")
    genSeq.Pythia8.Commands += ["SpaceShower:dipoleRecoil=on"]
    include("Pythia8_i/Pythia8_MadGraph.py")
    include("Pythia8_i/Pythia8_ShowerWeights.py")
```

WLZL OFP, SFM, SFP (e, μ):

```
  # WLZL OFP
   generate p p > w+{0} z{0} j j QCD=0, w+ > e+ ve, z > mu+ mu- @0
   add process p p > w+{0} z{0} j j QCD=0, w+ > mu+ vm, z > e+ e- @0

   # WLZL SFM
   generate p p > w-{0} z{0} j j QCD=0, w- > e- ve~, z > e+ e- @0
   add process p p > w-{0} z{0} j j QCD=0, w- > mu- vm~, z > mu+ mu- @0

   # WLZL SFP
   generate p p > w+{0} z{0} j j QCD=0, w+ > e+ ve, z > e+ e- @0
   add process p p > w+{0} z{0} j j QCD=0, w+ > mu+ vm, z > mu+ mu- @0
```

WLZT, WTZL, WTZT (e, μ):
Only the polarization of the bosons is changed shown for the OFM example. The flavour and charge combinations are done equivalently to the WLZL samples.

```
# WLZT OFM
 generate p p > w-{0} z{T} j j QCD=0, w- > e- ve~, z > mu+ mu- @0
 add process p p > w-{0} z{T} j j QCD=0, w- > mu- vm~, z > e+ e- @0

 #WTZL OFM
 generate p p > w-{T} z{0} j j QCD=0, w- > e- ve~, z > mu+ mu- @0
 add process p p > w-{T} z{0} j j QCD=0, w- > mu- vm~, z > e+ e- @0

 #WTZT OFM
 generate p p > w-{T} z{T} j j QCD=0, w- > e- ve~, z > mu+ mu- @0
 add process p p > w-{T} z{T} j j QCD=0, w- > mu- vm~, z > e+ e- @0
```

Samples with at least one τ in the final state

Samples with all the polarization, flavour and charge combinations as for the samples with only electrons and muons in the final state are produced. The only difference is that at least one of the leptons in the final state has to be a τ lepton. The resulting flavour combinations are shown for the WLZL samples:

```
# WLZL OFM
 generate p p > w-{0} z{0} j j QCD=0, w- > e- ve~, z > ta+ ta- @0
 add process p p > w-{0} z{0} j j QCD=0, w- > mu- vm~, z > ta+ ta- @0
 add process p p > w-{0} z{0} j j QCD=0, w- > ta- vt~, z > e+ e- @0
 add process p p > w-{0} z{0} j j QCD=0, w- > ta- vt~, z > mu+ mu- @0

 # WLZL OFP
 generate p p > w+{0} z{0} j j QCD=0, w+ > e+ ve, z > ta+ ta- @0
 add process p p > w+{0} z{0} j j QCD=0, w+ > mu+ vm, z > ta+ ta- @0
 add process p p > w+{0} z{0} j j QCD=0, w+ > ta+ vt, z > e+ e- @0
 add process p p > w+{0} z{0} j j QCD=0, w+ > ta+ vt, z > mu+ mu- @0

 # WLZL SFM
 generate p p > w-{0} z{0} j j QCD=0, w- > ta- vt~, z > ta+ ta- @0

 # WLZL SFP
 generate p p > w+{0} z{0} j j QCD=0, w+ > ta+ vt, z > ta+ ta- @0
```

B.1.2 Polarized Samples in the Parton Parton Frame

Samples with all the helicity, flavour and charge combinations are produced analogous to the ones in the helicity frame. Only samples with e and μ in the final state are generated. The only difference is the definition of the `me_frame`. Two possibilities for the definition of the parton parton frame are used. First `me_frame = 1,2` was set. This corresponds to the centre of mass frame of the two partons. As an alternative `me_frame = 3,4,5,6,7,8` was set. This corresponds to the centre of mass frame of all the final state particles. For both frame definitions all the samples are produced twice with different random seeds for the generation of the gridpack. For the first samples the default random seed was used (`iseed=0`) and for the second samples the random seed is set to `iseed=4321`.

B.1.3 Full Sample

Samples with only e and μ in the final state

For the full samples only the process definition is changed, shown here for the OFM sample. The flavour and charge combinations are similar to the ones shown for the polarized samples.

```
generate p p > mu+ mu- e- ve~ j j QCD=0 @0
add process p p > e+ e- mu- vm~ j j QCD=0 @0
```

Samples with at least one τ in the final state

Samples with all the polarization, flavour and charge combinations as for the samples with only electrons and muons in the final state are produced. The only difference is that at least one of the leptons in the final state has to be a τ lepton. The flavour and charge combinations are similar to the ones shown for the polarized samples.

B.1.4 Unpolarized Sample

For the unpolarized samples only the process definition is changed, shown here for the OFM sample. The flavour and charge combinations are similar to the ones shown for the polarized samples. Only samples with e and μ in the final state are produced.

```
generate p p > w- z j j QCD=0, w- > e- ve~, z > mu+ mu- @0
add process p p > w- z j j QCD=0, w- > mu- vm~, z > e+ e- @0
```

B.2 Samples at Truth Level

Table B.1: List of locally produced samples summarized under the names by which they are used in this analysis. The polarization for all the pol samples in this table are defined in the helicity frame by setting `me_frame = 3,4,5,6`. The leptons in the final state of the hard process are (a) either electrons or muons or (b) at least one of them is a τ . In the column flavour,charge samples in which the lepton from the W-decay has to be a different flavour than the ones from the Z decay are labelled with OF. If all leptons have the same flavour the samples are labelled with SF. The M and P stand for minus and plus and denote the charge of the W boson.

(a)

name	helicity	flavour,charge	N_{events}	σ in pb
full	-	OFM	1000000	0.005276
	-	OFP	1000000	0.008859
	-	SFM	1000000	0.005268
	-	SFP	1000000	0.008840
unpol	-	OFM	1000000	0.003863
	-	OFP	1000000	0.006661
	-	SFM	1000000	0.003861
	-	SFP	1000000	0.006658
pol	$W_L Z_L$	OFM	1000000	0.0002744
		OFP	994160	0.0004469
		SFM	1000000	0.0002743
		SFP	1000000	0.0004467
	$W_L Z_T$	OFM	1000000	0.0006886
		OFP	1000000	0.001142
		SFM	1000000	0.00068844
		SFP	1000000	0.001142
	$W_T Z_L$	OFM	1000000	0.0008144
		OFP	1000000	0.001428
		SFM	1000000	0.000814
		SFP	1000000	0.001429
	$W_T Z_T$	OFM	1000000	0.002079
		OFP	1000000	0.00363
		SFM	1000000	0.002079
		SFP	1000000	0.003628

(b)

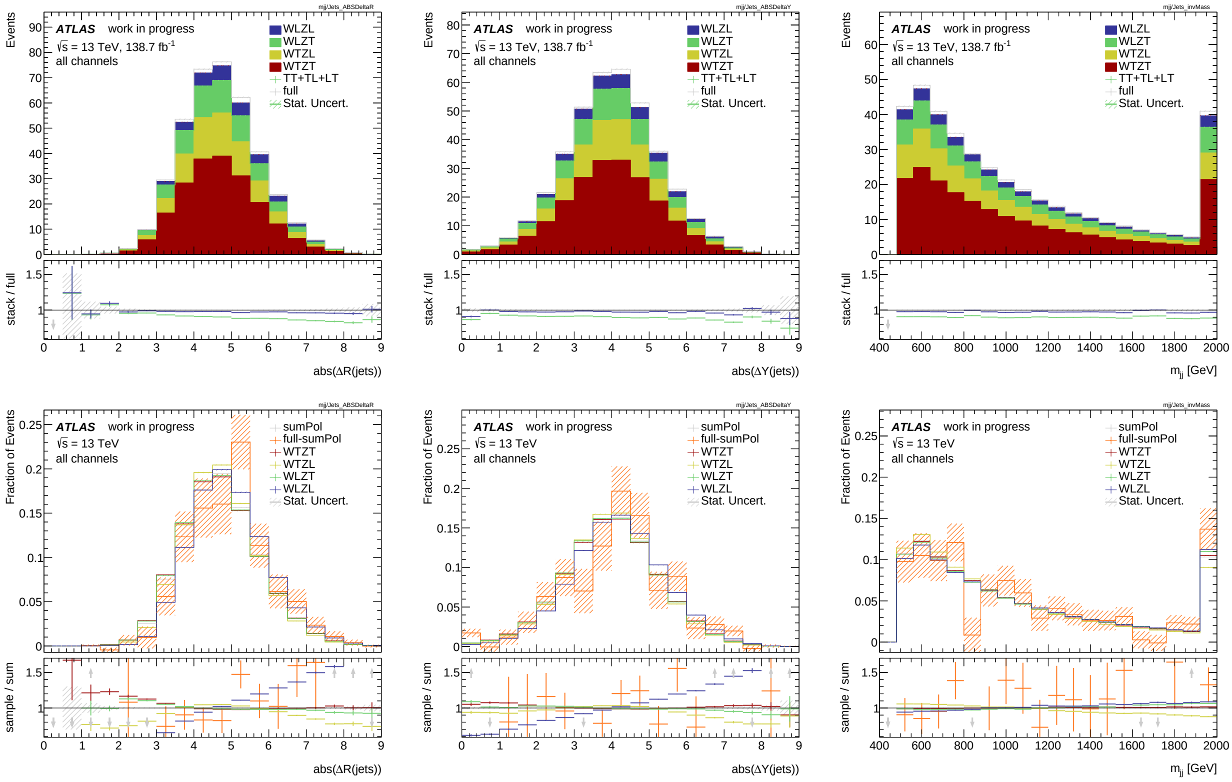
name	helicity	flavour	N_{events}	σ in pb
full	-	OFM	1000000	0.01055
	-	OFP	1000000	0.01772
	-	SFM	980000	0.002634
	-	SFP	1000000	0.004420
pol	$W_L Z_L$	OFM	1000000	0.0005487
		OFP	991601	0.0008938
		SFM	1000000	0.0001372
		SFP	1000000	0.0002233
	$W_L Z_T$	OFM	1000000	0.001377
		OFP	1000000	0.002283
		SFM	1000000	0.000344
		SFP	1000000	0.000571
	$W_T Z_L$	OFM	1000000	0.001629
		OFP	1000000	0.002857
		SFM	1000000	0.000407
		SFP	1000000	0.000715
	$W_T Z_T$	OFM	1000000	0.004158
		OFP	1000000	0.007260
		SFM	1000000	0.001039
		SFP	1000000	0.001814

Table B.2: List of locally produced samples summarized under the names by which they are used in this analysis. The polarization for all samples in this table are defined in the parton parton frame by setting `me_frame = 1,2` (`me12`) or `me_frame = 3,4,5,6,7,8` (`me345678`). The leptons in the final state of the hard process are either electrons or muons. In the column flavour,charge samples in which the lepton from the W-decay has to be a different flavour than the ones from the Z decay are labelled with OF. If all leptons have the same flavour the samples are labelled with SF. The M and P stand for minus and plus and denote the charge of the W boson.

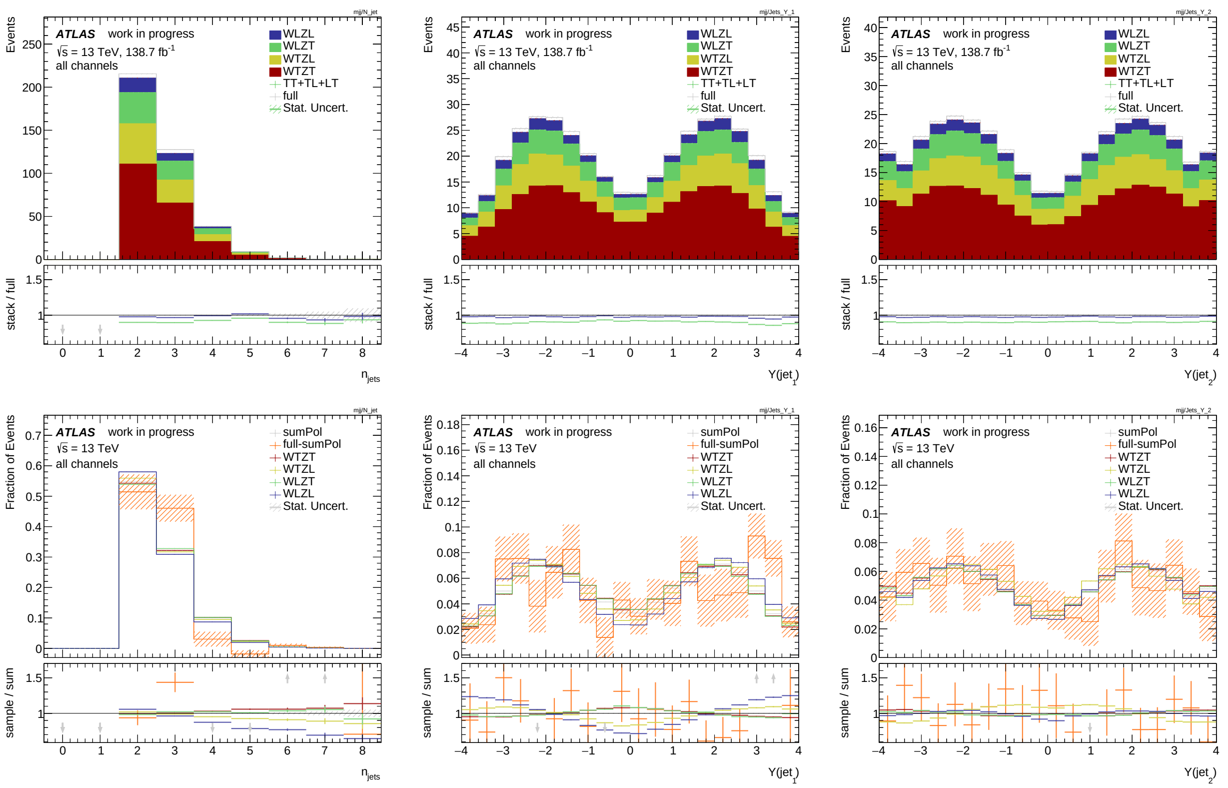
name	helicity	flavour,charge	N_{events}	σ in pb	
				`me12`	`me345678`
pol	$W_L Z_L$	OFM	10000	0.0002245	0.0002096
		OFP	10000	0.0003599	0.0003383
		SFM	10000	0.0002238	0.0002109
		SFP	10000	0.000359	0.000337
	$W_L Z_T$	OFM	10000	0.0005766	0.0005409
		OFP	10000	0.0009532	0.0008979
		SFM	10000	0.0005755	0.0005399
		SFP	10000	0.0009529	0.0008919
	$W_T Z_L$	OFM	10000	0.0007996	0.0007519
		OFP	10000	0.001399	0.001313
		SFM	10000	0.0007994	0.0007487
		SFP	10000	0.001401	0.00131
	$W_T Z_T$	OFM	10000	0.002251	0.00211
		OFP	10000	0.003934	0.003696
		SFM	10000	0.002251	0.002113
		SFP	10000	0.003929	0.00368
pol, diff. seed gridpack	$W_L Z_L$	OFM	10000	0.0002238	0.0002107
		OFP	10000	0.0003601	0.0003364
		SFM	10000	0.0002246	0.0002104
		SFP	10000	0.0003588	0.0003373
	$W_L Z_T$	OFM	10000	0.0005784	0.0005423
		OFP	10000	0.000955	0.0008975
		SFM	10000	0.0005797	0.0005427
		SFP	10000	0.000951	0.0008972
	$W_T Z_L$	OFM	10000	0.0008019	0.000751
		OFP	10000	0.001402	0.001309
		SFM	10000	0.0008055	0.00075
		SFP	10000	0.001398	0.001319
	$W_T Z_T$	OFM	10000	0.002256	0.002114
		OFP	10000	0.003935	0.003678
		SFM	10000	0.002245	0.002108
		SFP	10000	0.003937	0.003683

C Distributions of the Polarized Sample at Truth Level

In the following the distributions of the different polarizations for several variables are shown. In the plots on the top the full sample and in the stack the four polarization combinations are shown. In the subplot the ratio of the stack to the full sample is shown as the blue line and the the ratio of the sum of WTZT, WTZL and WLZT to the full sample as the green line. In the plots at the bottom the normalized distributions of the four polarized samples and the sum of the polarized samples is plotted. Additionally the normalized distribution of the difference of the full sample to the sum of the polarized samples is shown. In the subplot the ratios of the individual samples to the sum of the polarization is given.

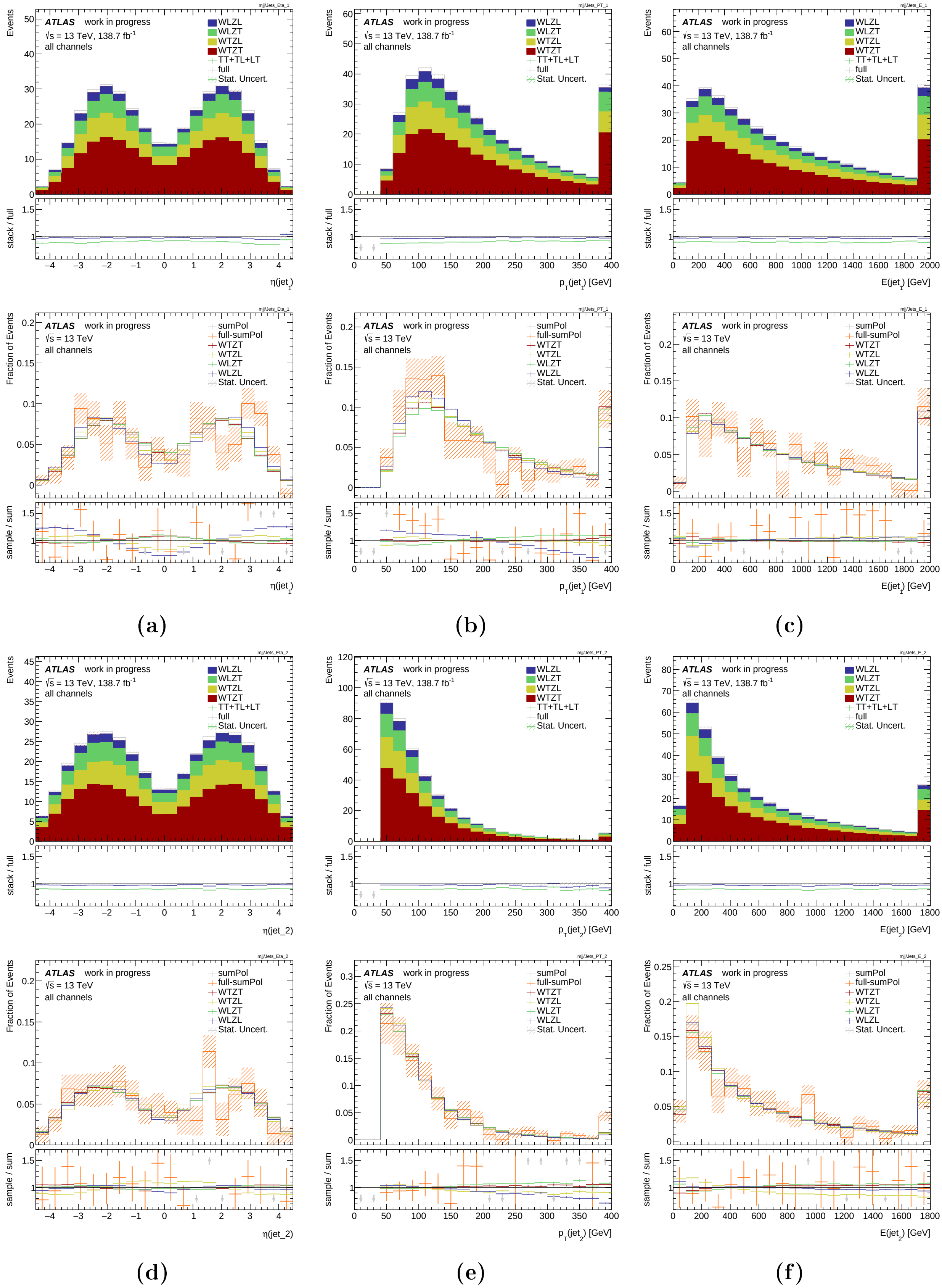


(a) (b) (c)



(d) (e) (f)

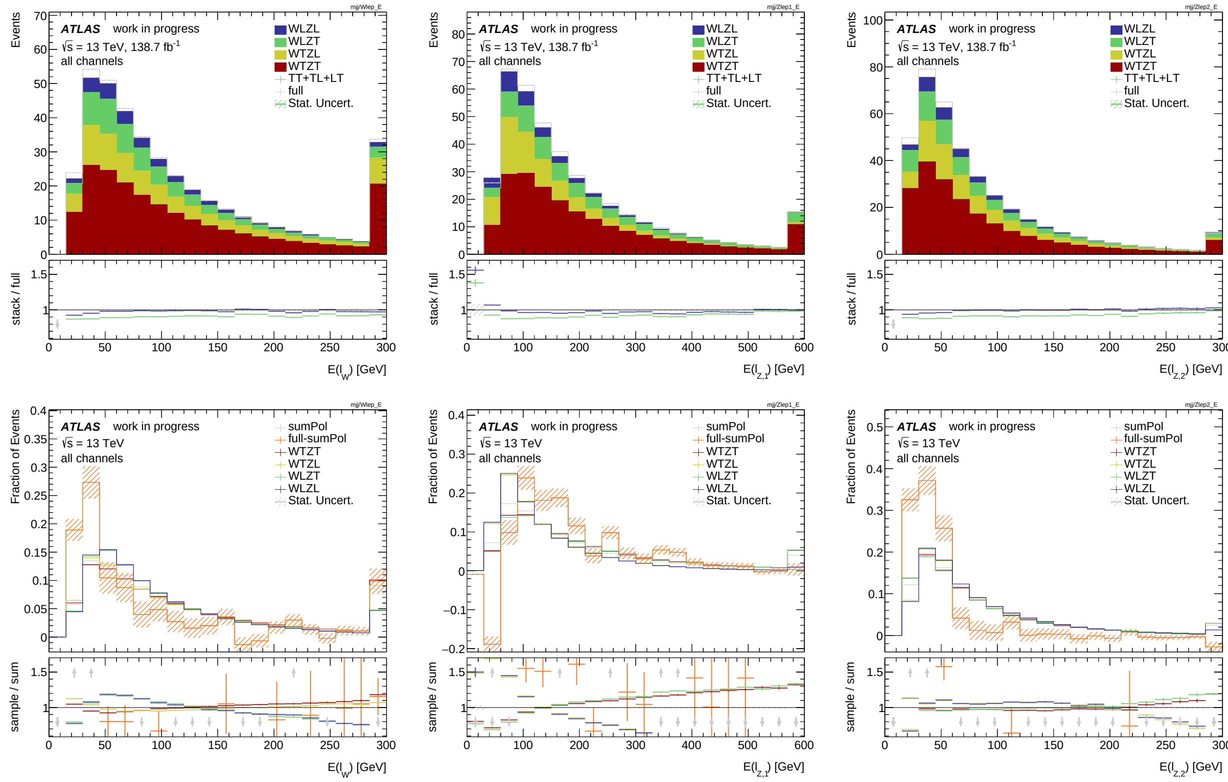
Figure C.1: Distributions of (a) the absolute value of the angular distance of the two jets $|\Delta R(jj)|$, (b) the absolute value of the rapidity difference of the two jets $|\Delta Y(jj)|$, (c) the invariant mass of the two jets m_{jj}, (d) the number of jets n_{jets}, (e) the rapidity of the first jet $Y(jet_1)$ and (f) the rapidity of the second jet $Y(jet_1)$.



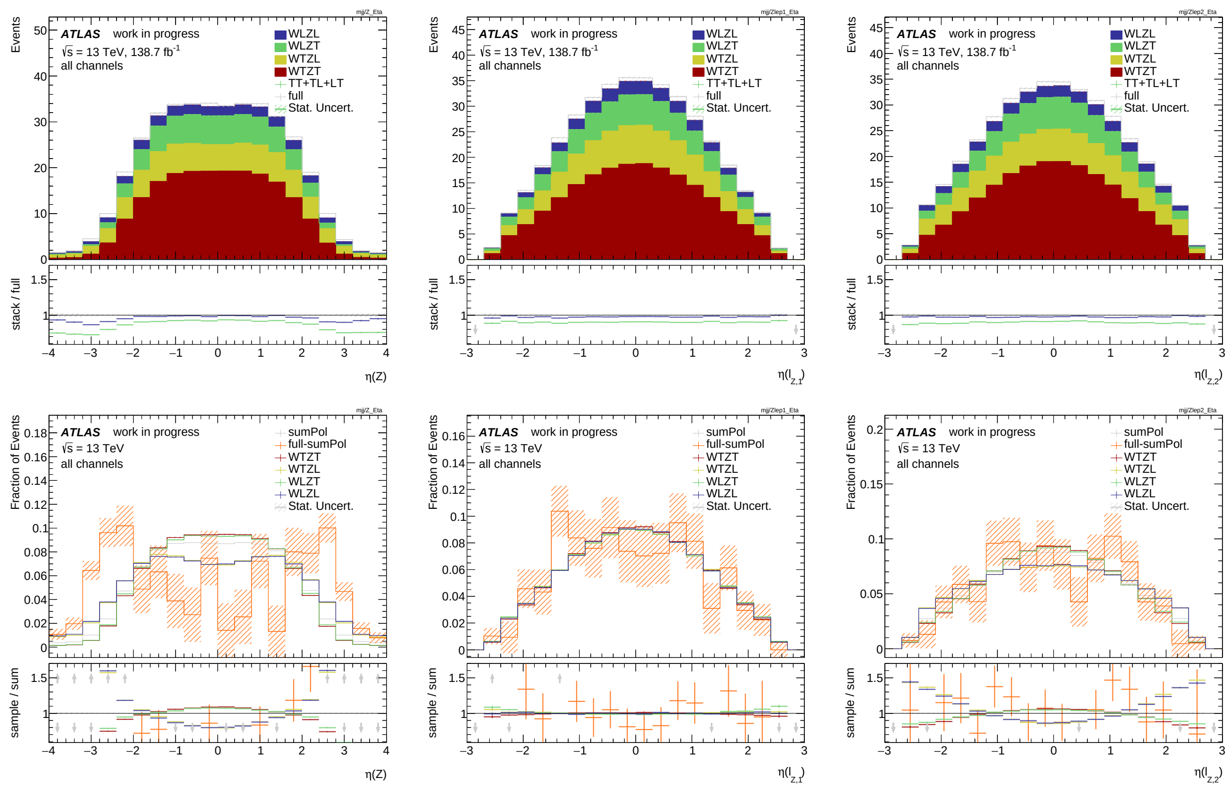
(a) (b) (c)

(d) (e) (f)

Figure C.2: Distributions of (a) the rapidity of the first jet $\eta(jet_1)$, (b) the transverse momentum of the first jet $p_T(jet_1)$, (c) the energy of the first jet $E(jet_1)$, (d) the rapidity of the second jet $\eta(jet_2)$, (e) the transverse momentum of the second jet $p_T(jet_2)$ and (f) the energy of the second jet $E(jet_2)$.



(a) (b) (c)



(d) (e) (f)

Figure C.3: Distributions of (a) the energy of the W lepton $E(\ell_W)$, (b) the energy of the leading Z lepton $E(\ell_{Z,1})$, (c) the energy of the subleading Z lepton $E(\ell_{Z,2})$, (d) the rapidity of the Z boson $\eta(Z)$, (e) the rapidity of the leading Z lepton $\eta(\ell_{Z,1})$ and (f) the rapidity of the Z boson $\eta(Z)$, (e) the rapidity of the subleading Z lepton $\eta(\ell_{Z,2})$.

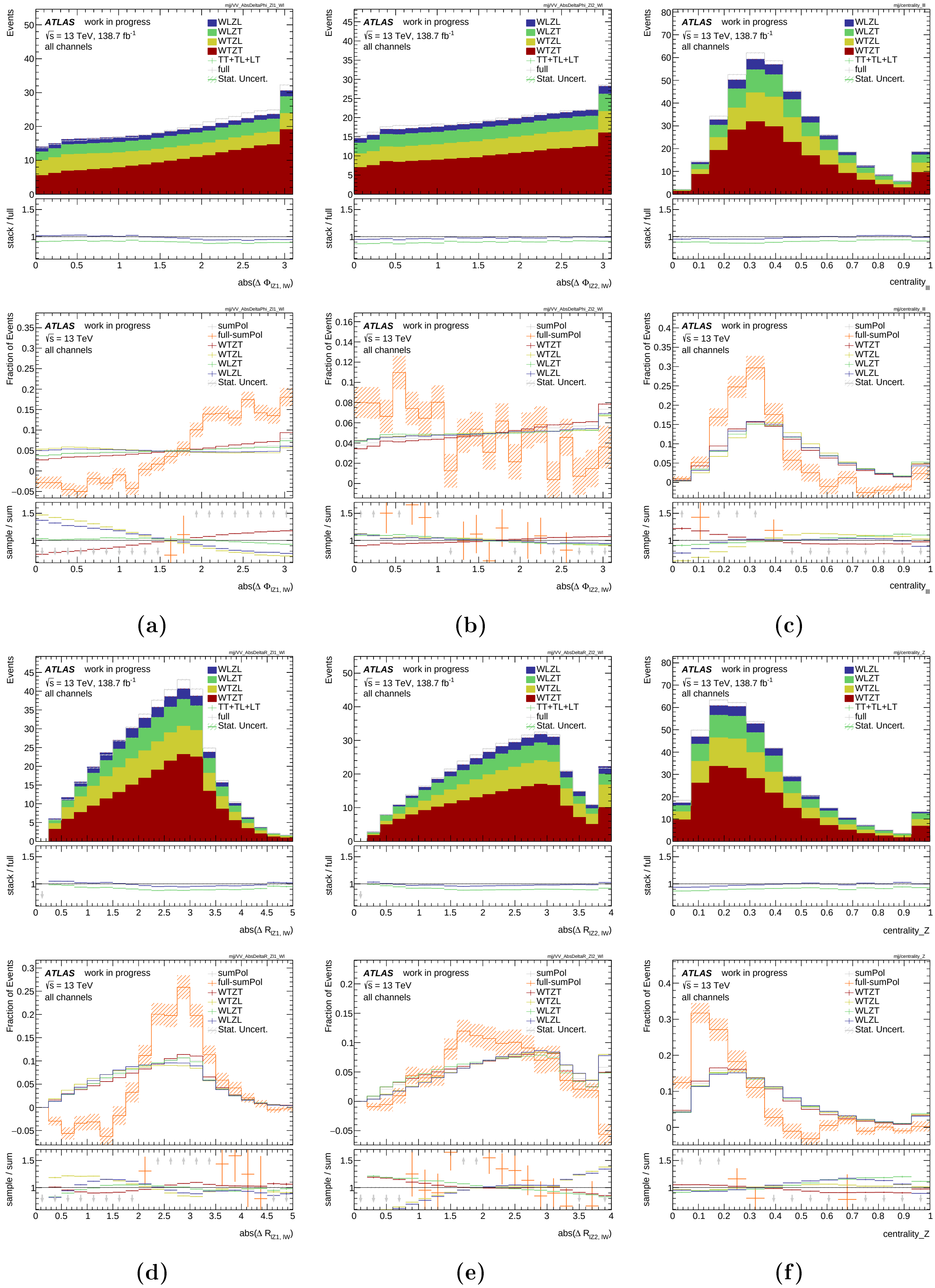


Figure C.4: Distributions of (a) the absolute value of the difference of the azimuthal angles of the leading Z lepton and the W lepton $|\Delta\phi_{lZ1,lW}|$, (b) the absolute value of the difference of the azimuthal angles of the subleading Z lepton and the W lepton $|\Delta\phi_{lZ2,lW}|$, (c) the centrality of the three leptons centrality$_{lll}$, (d) the absolute value of the angular distance of the leading Z lepton and the W lepton $|\Delta R_{lZ1,lW}|$ (e) the absolute value of the angular distance of the subleading Z lepton and the W lepton $|\Delta R_{lZ2,lW}|$ and (f) the centrality of the Z boson centrality(Z).

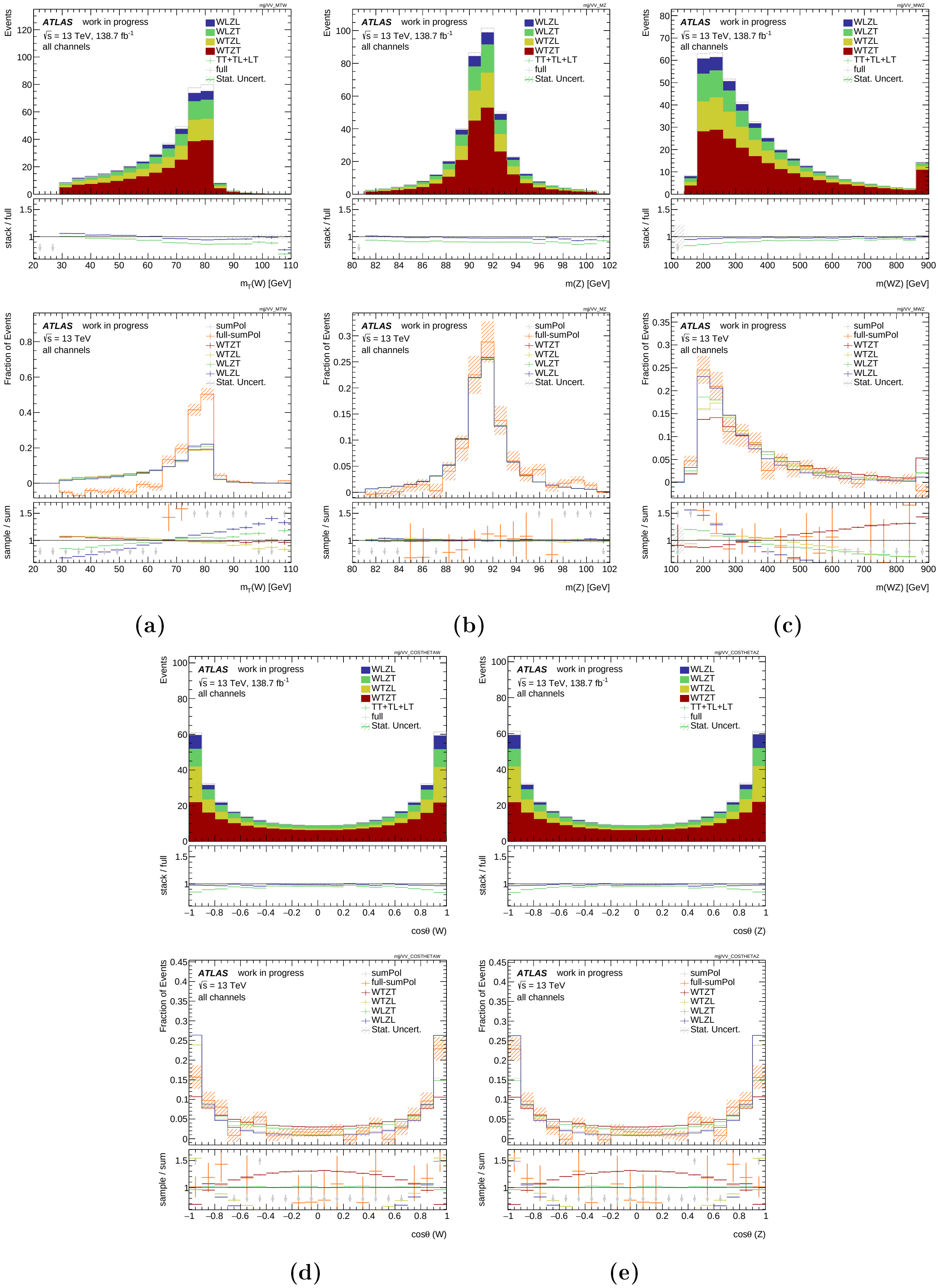
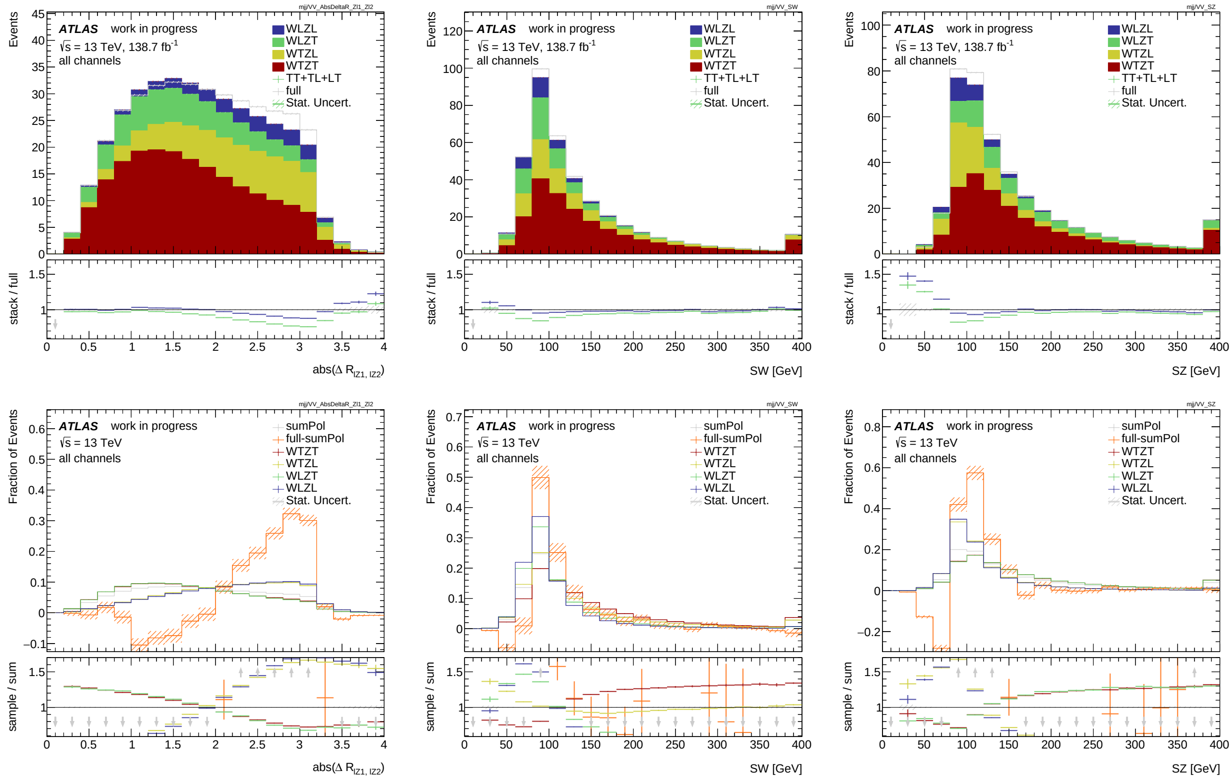
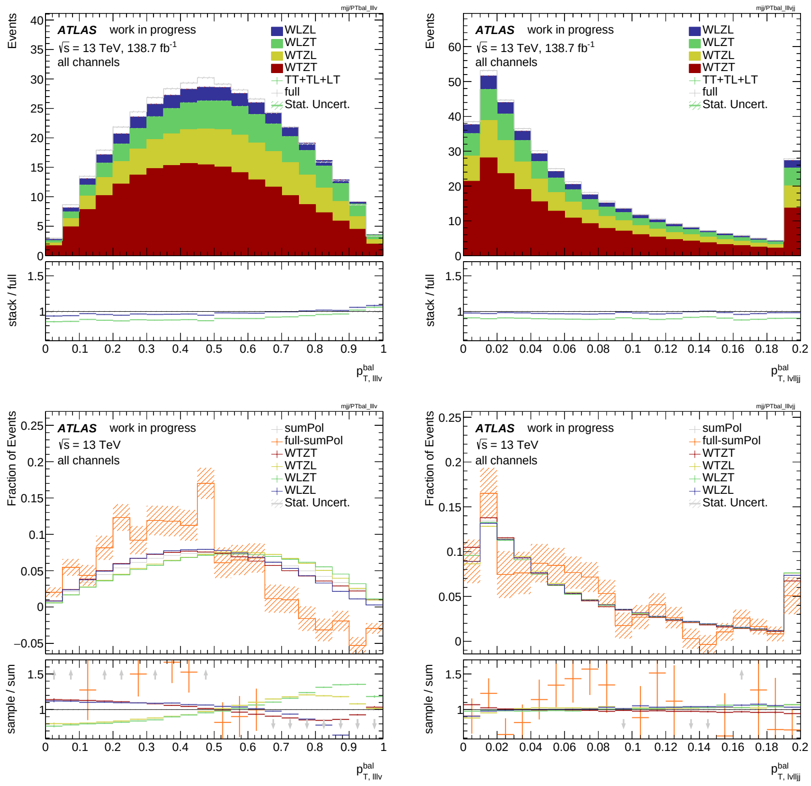


Figure C.5: Distributions of (a) the transverse mass of the W boson $m_T(W)$, (b) the invariant mass of the Z boson $m(Z)$, (c) the invariant mass of the diboson system $m(WZ)$, (d) the cosine of the angle of the W boson to the beam axis in the WZ rest frame and (f) the cosine of the angle of the Z boson to the beam axis in the WZ rest frame.



(a) (b) (c)



(d) (e)

Figure C.6: Distributions of (a) the absolute value of the angular distance of the leading and subleading Z lepton $|\Delta R_{lZ1,lZ2}|$, (b) the scalar sum of the W boson SW, (c) the scalar sum of the Z boson SZ, (d) $p^{\text{bal}}_{T,l\nu ll}$ and (f) $p^{\text{bal}}_{T,l\nu lljj}$.

D Comparison of Samples in the two Different Definitions of the Parton Parton Frame

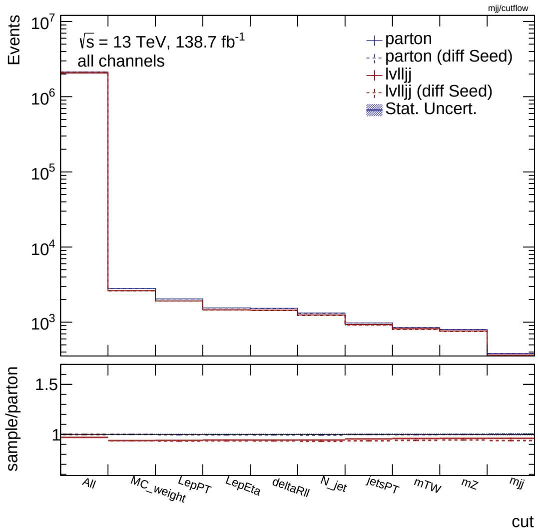
In this section samples with polarization defined in the parton parton frame for the two definitions `me_frame = 1,2` and `me_frame = 3,4,5,6,7,8` in Madgraph are compared. For both definitions two sets of samples with different seeds in the gridpack were produced. More infrmation about the samples is given in section B.1. The cutflow table for these four samples is shown in Table D.1. The cutlow is shown visually in Figure D.1a. It can be seen that the event yields of samples with `3,4,5,6,7,8` are constantly below the ones with `1,2`. The largest discrepancy is found in the total phase space with about 6%. Applying cuts reduces the discrepancy to about 4% in the truth VBS phase space. Especially the cut on the transverse momenta of the jets reduces the differences. To see if this is caused by shape differences between the samples the distributions for the transverse momenta of the first and second jets before the cut on these two variables are shown in Figure D.1b and Figure D.1c. Additionally the transverse mass of the W boson before the cut on this variable is shown in Figure D.1d. No large shape differences are visible.

Madgraph experts were informed about this [42]. As suggested the samples with different seeds are used to get an estimate of the statistical uncertainties. Especially in a phase space without many cuts the difference between samples with different seeds are clearly smaller than the difference between the samples with the different definitions of the frames. Thus it is unlikely that the discrepancy is caused by statistical fluctuations.

For the production of the samples `sde_strategy=1` and `dynamical_scale_choice = -1` was used. In [42] it is suggested that this setup can lead to small numerical errors in the determination of the scale, which may lead to the observed difference. To investigate this for further studies samples could be produced with another setup.

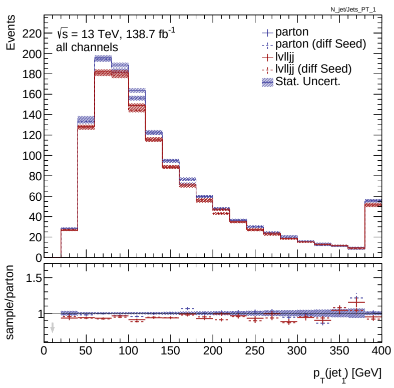
Table D.1: Cutflow table of samples with polarization defined in the parton parton frame by setting `me_frame = 1,2` (parton) or `me_frame = 3,4,5,6,7,8` (lvlljj). Samples that are produced with a different random seed in the gridpack are labelled with (a) and (b). The ratios between the samples with the same polarization definition but different seeds are shown in the columns parton (b)/(a) and lvlljj (b)/(a). The ratios between the samples (a) with the different polarization definitions is shown in the column lvlljj/parton (a).

	parton (a)	parton (b)	lvlljj (a)	lvlljj (b)	parton (b)/(a)	lvlljj (b)/(a)	lvlljj/parton (a)
All	2132000 ± 5000	2130000 ± 5000	2066000 ± 5000	2127000 ± 5000	0.999 ± 0.004	1.030 ± 0.004	0.969 ± 0.004
MC_weight	2799 ± 10	2796 ± 10	2624 ± 9	2618 ± 9	0.999 ± 0.005	0.997 ± 0.005	0.938 ± 0.005
LepPT	2046 ± 8	2036 ± 8	1921 ± 8	1907 ± 7	0.995 ± 0.006	0.993 ± 0.006	0.939 ± 0.005
LepEta	1551 ± 7	1542 ± 7	1463 ± 7	1449 ± 7	0.994 ± 0.006	0.991 ± 0.006	0.943 ± 0.006
deltaRll	1534 ± 7	1524 ± 7	1446 ± 7	1433 ± 7	0.994 ± 0.006	0.991 ± 0.006	0.942 ± 0.006
N_jet	1325 ± 7	1313 ± 6	1249 ± 6	1232 ± 6	0.991 ± 0.007	0.986 ± 0.007	0.943 ± 0.007
jetsPT	977 ± 6	976 ± 6	932 ± 5	914 ± 5	0.998 ± 0.008	0.980 ± 0.008	0.954 ± 0.008
mTW	855 ± 5	852 ± 5	819 ± 5	803 ± 5	0.997 ± 0.009	0.980 ± 0.009	0.958 ± 0.008
mZ	799 ± 5	798 ± 5	767 ± 5	754 ± 5	0.999 ± 0.009	0.983 ± 0.009	0.959 ± 0.009
mjj	382 ± 4	383 ± 4	367 ± 3	358 ± 3	1.002 ± 0.013	0.977 ± 0.013	0.960 ± 0.013



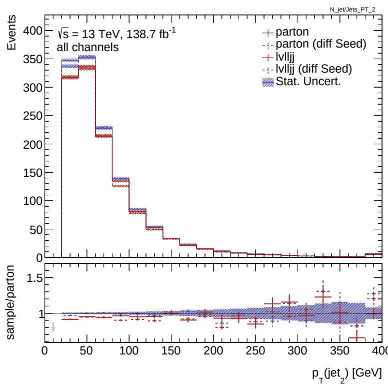
(a)

before jetsPT Cut (pt $>$ 40 GeV)



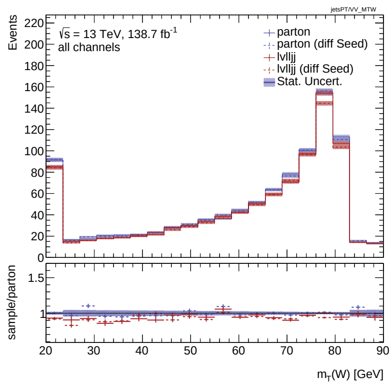
(b)

before jetsPT Cut (pt $>$ 40 GeV)



(c)

before mTW Cut (mTW $>$ 30 GeV)

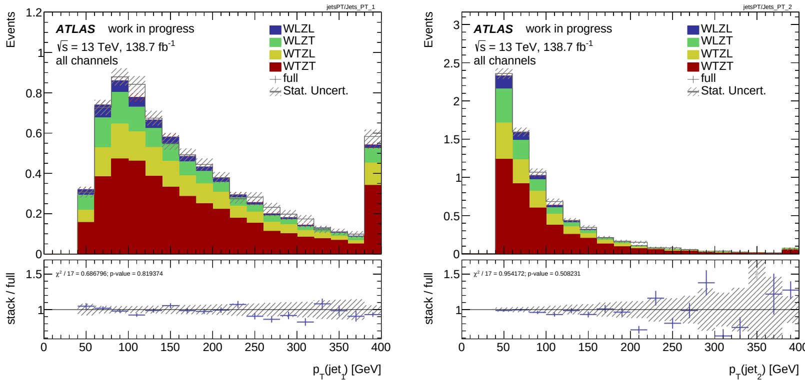


(d)

Figure D.1: Distributions for samples where the polarization is defined in the parton parton frame by setting `me_frame = 1,2` (parton, blue) or `me_freme = 3,4,5,6,7,8` (lvlljj, red). Samples are produced with different random seed in the gridpack (continuous and scattered line). In the subplot the ratio of each of the samples to parton sample is shown. In (a) the event yields for every step of the cutflow are shown. In (b) and (c) the distributions for the transverse moment of the first and second jet before the cut on these variables and in (d) the distribution for the transverse mass of the W boson before the cut on this variable are shown.

E Distributions for Samples with τ Decays

In the following the distributions of the four different polarization combinations for samples, where at least one of the leptons in the final state of the hard process has to be τ lepton. A more loose phase space than the truth VBS phase space, where only the cuts up to the cut on the transverse momenta of the jets are included, is used.



(a) (b) (c)

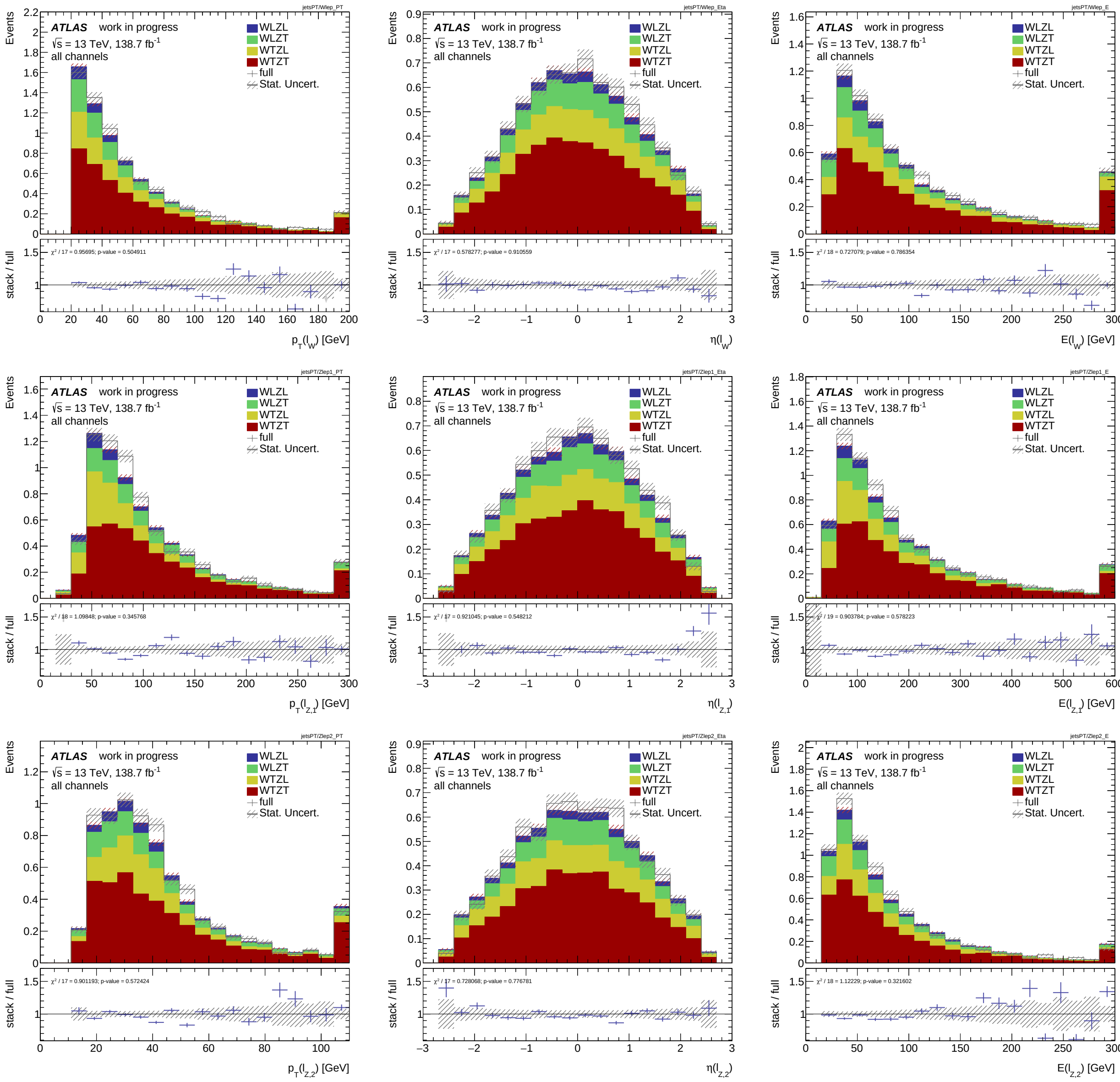
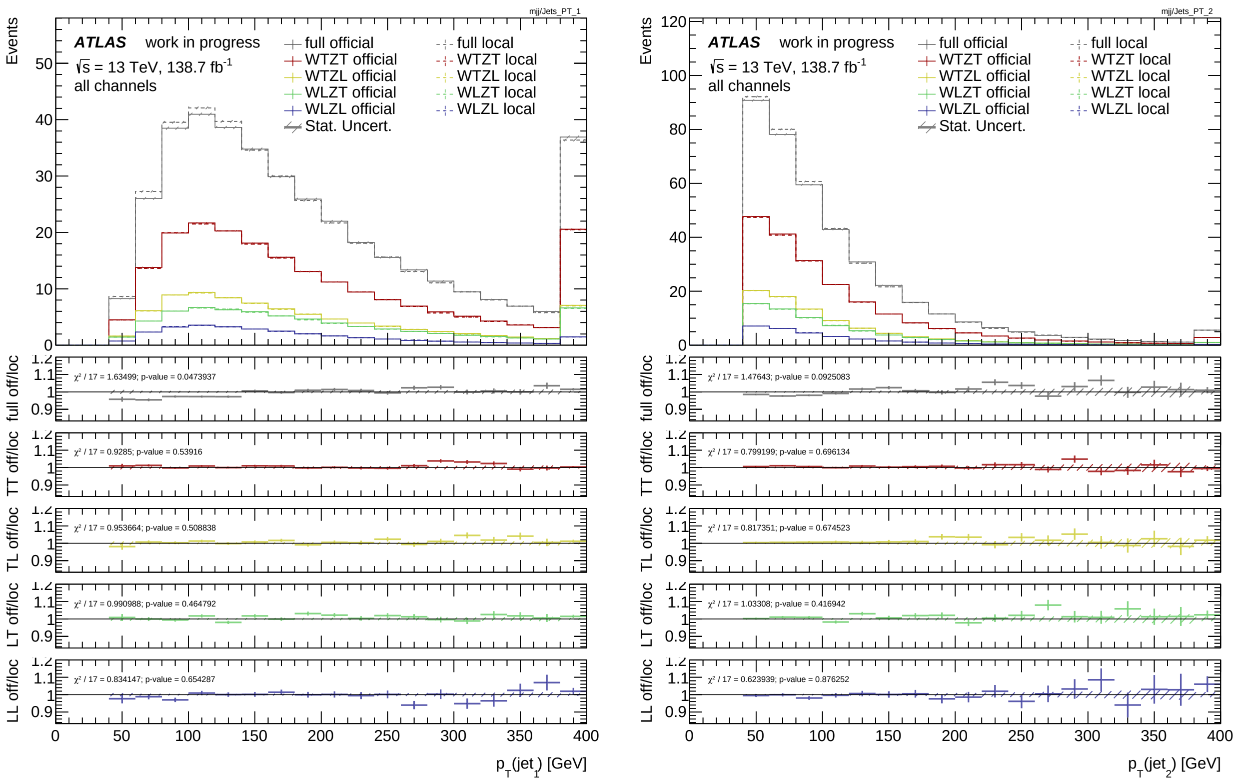


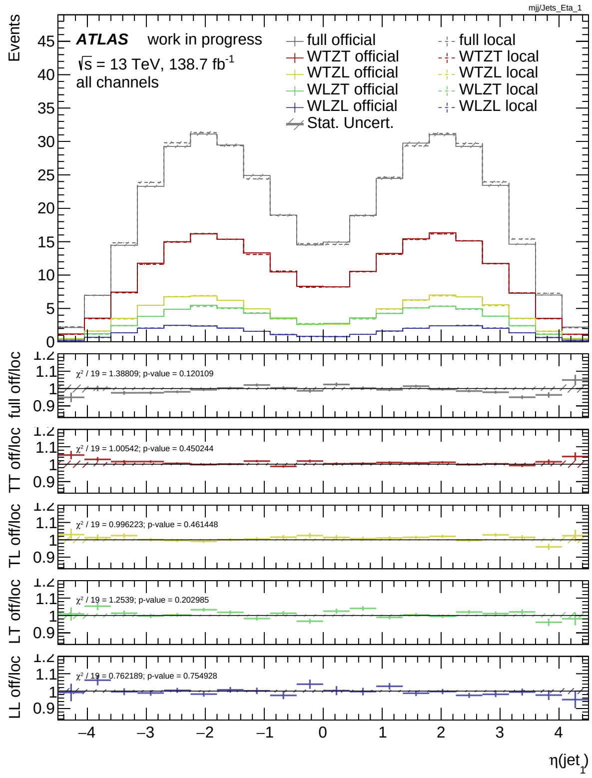
Figure E.1: Distributions of the four polarization combinations for samples with at least one τ in the final state of the hard process for basic variables of the jets and the W and Z boson.

F Comparison Plots for the Local and Official Samples

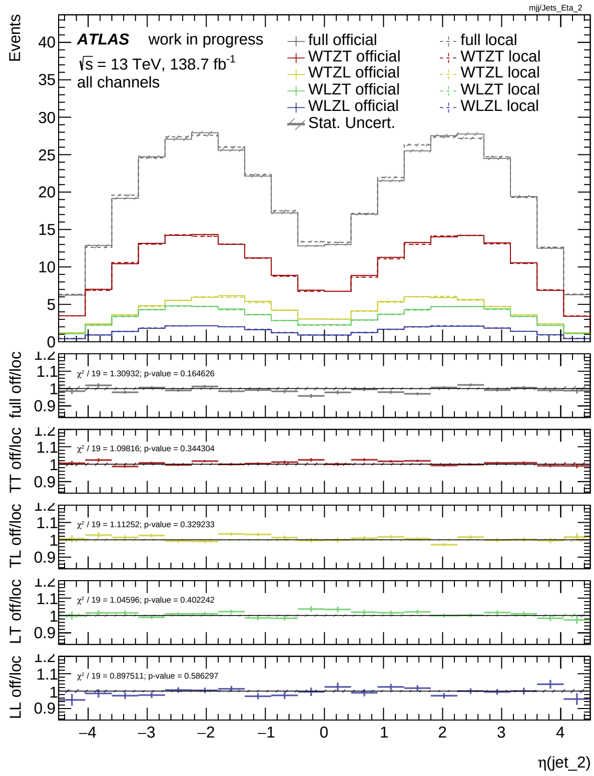


(a) **(b)**

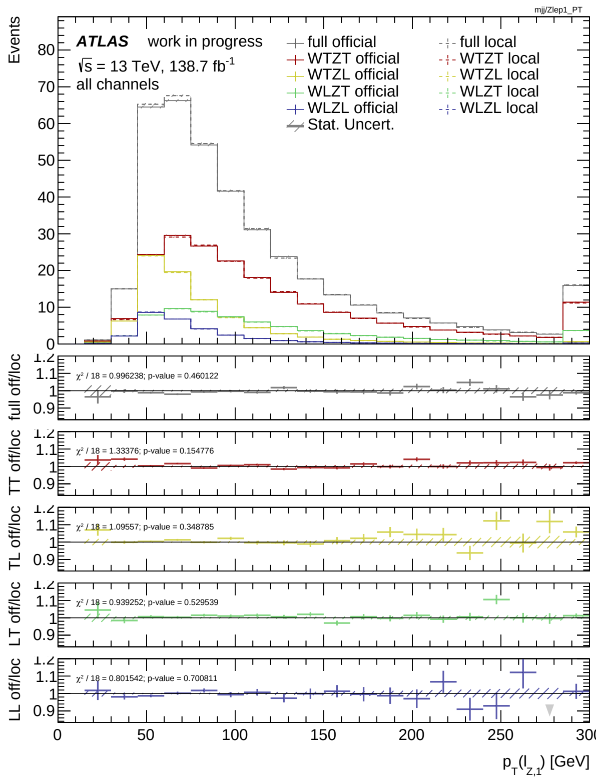
Figure F.1: Comparison of the normalized distributions of the four polarization combinations between the locally (dashed lines) and officially (full lines) produced samples. In the subplots the ratio of the officially produced sample to the locally produced sample is shown for the full sample and every polarization combination.



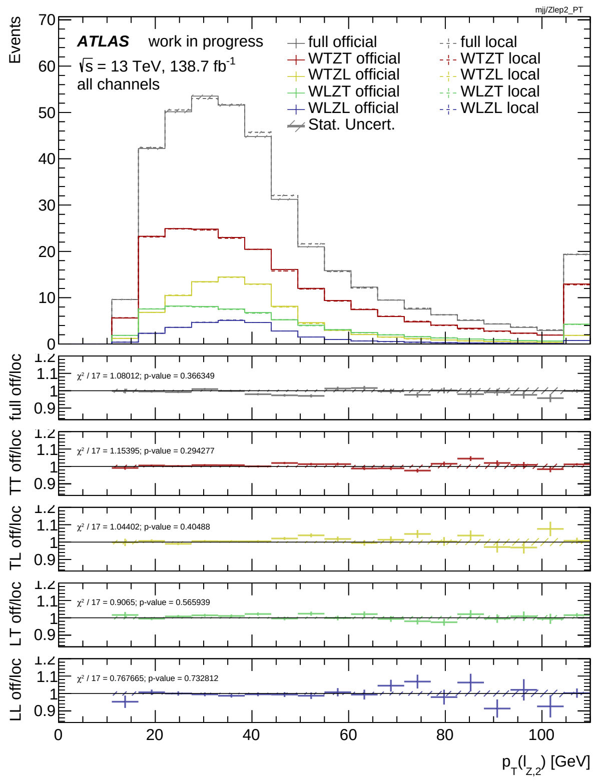
(a)



(b)

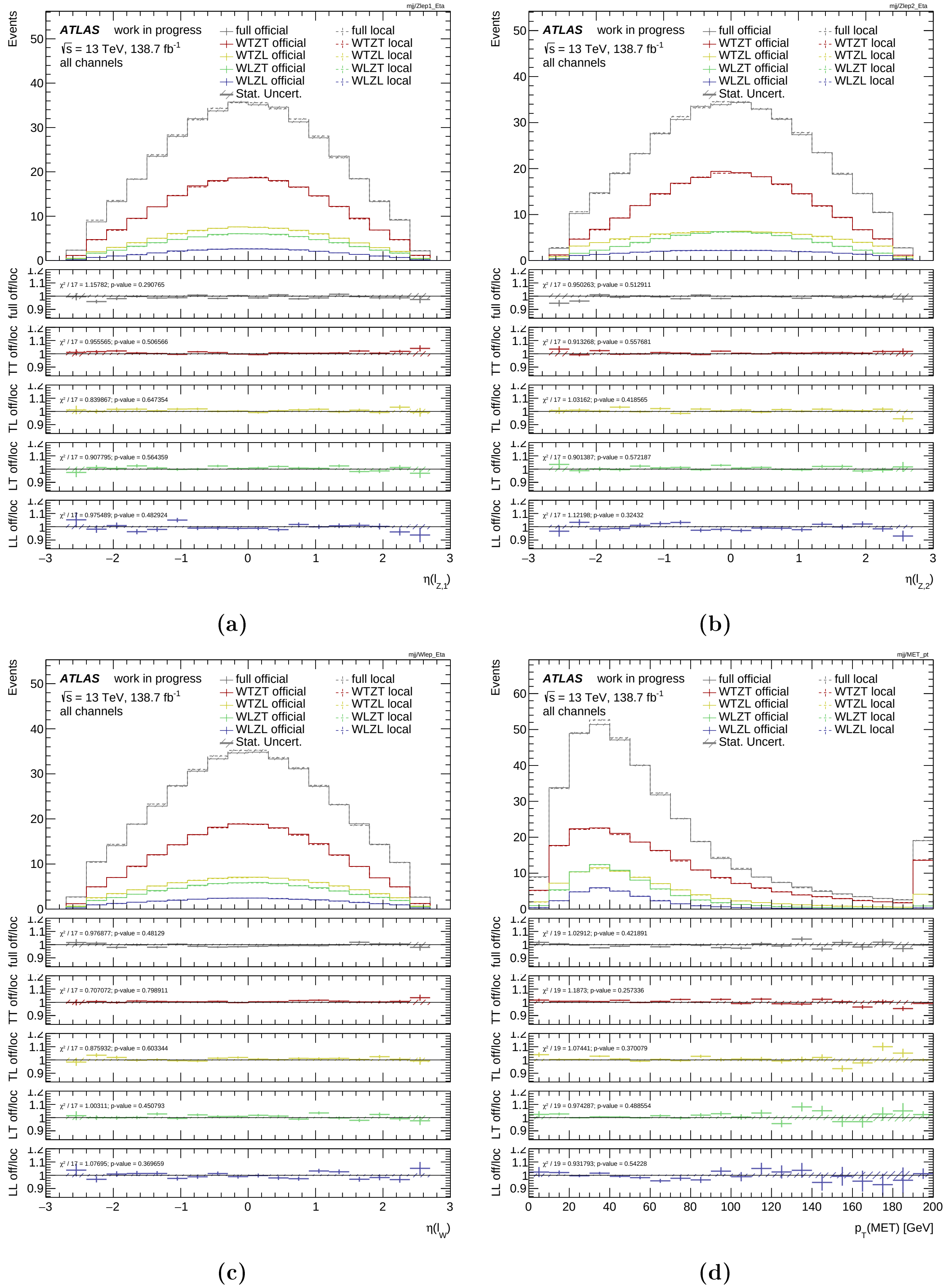


(c)



(d)

Figure F.2: Comparison of the normalized distributions of the four polarization combinations between the locally (dashed lines) and officially (full lines) produced samples. In the subplots the ratio of the officially produced sample to the locally produced sample is shown for the full sample and every polarization combination.



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

Figure F.3: Comparison of the normalized distributions of the four polarization combinations between the locally (dashed lines) and officially (full lines) produced samples. In the subplots the ratio of the officially produced sample to the locally produced sample is shown for the full sample and every polarization combination.

G Samples at Reconstruction Level

Table G.1: List of the full and the polarized signal samples used in this analysis at reco level summarized under the name by which they are referred to in this analysis. For all polarized samples the polarization is defined in the helicity frame. In the process definition the OF (opposite flavour) denotes that the lepton coming from the W boson has the not the same flavour as the two leptons from the Z decay while for SF all leptons have the same flavour. The letter after the SF/OF denotes the charge of the W boson where M stand for minus and P for plus.

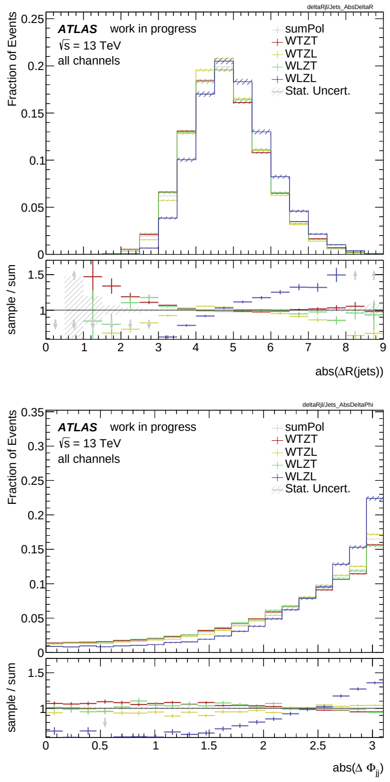
name	DSID	process	Generator	N_{events}	σ in pb	Filter eff.	k-factor
full	364739	$pp \to \ell\nu\ell\ell jj$ EW6 OFM	MadGraph + Pythia8	1.4M	0.015	1.0	1.0
	364741	$pp \to \ell\nu\ell\ell jj$ EW6 OFP		2M	0.026	1.0	1.0
	364742	$pp \to \ell\nu\ell\ell jj$ EW6 SFM		0.7M	0.0077	1.0	1.0
	364743	$pp \to \ell\nu\ell\ell jj$ EW6 SFP		0.9M	0.013	1.0	1.0
WLZL	511986	$W_LZ_Ljj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 OFM	MadGraph + Pythia8	70000	0.00027	1.0	1.0
	511987	$W_LZ_Ljj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 OFP		120000	0.00045	1.0	1.0
	511988	$W_LZ_Ljj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 SFM		70000	0.00027	1.0	1.0
	511989	$W_LZ_Ljj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 SFP		129000	0.00045	1.0	1.0
	514532	$W_LZ_Ljj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 OFM	MadGraph + Pythia8	30000	0.00055	1.0	1.0
	514533	$W_LZ_Ljj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 OFP		49000	0.00089	1.0	1.0
	514534	$W_LZ_Ljj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 SFM		30000	0.00014	1.0	1.0
	514535	$W_LZ_Ljj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 SFP		40000	0.00022	1.0	1.0
WLZT	511990	$W_LZ_Tjj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 OFM	MadGraph + Pythia8	169000	0,00069	1.0	1.0
	511991	$W_LZ_Tjj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 OFP		288000	0.0011	1.0	1.0
	511992	$W_LZ_Tjj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 SFM		169000	0.00069	1.0	1.0
	511993	$W_LZ_Tjj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 SFP		285000	0.0011	1.0	1.0
	514536	$W_LZ_Tjj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 OFM	MadGraph + Pythia8	50000	0.0014	1.0	1.0
	514537	$W_LZ_Tjj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 OFP		70000	0.0023	1.0	1.0
	514538	$W_LZ_Tjj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 SFM		50000	0.00034	1.0	1.0
	514539	$W_LZ_Tjj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 SFP		70000	0.00057	1.0	1.0
WTZL	511994	$W_TZ_Ljj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 OFM	MadGraph + Pythia8	207000	0.00081	1.0	1.0
	511995	$W_TZ_Ljj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 OFP		360000	0.0014	1.0	1.0
	511996	$W_TZ_Ljj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 SFM		210000	0.00081	1.0	1.0
	511997	$W_TZ_Ljj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 SFP		357000	0.0014	1.0	1.0
	514540	$W_TZ_Ljj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 OFM	MadGraph + Pythia8	49000	0.0016	1.0	1.0
	514541	$W_TZ_Ljj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 OFP		80000	0.0029	1.0	1.0
	514542	$W_TZ_Ljj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 SFM		50000	0.00041	1.0	1.0
	514543	$W_TZ_Ljj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 SFP		90000	0.00071	1.0	1.0
WTZT	511998	$W_TZ_Tjj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 OFM	MadGraph + Pythia8	470000	0.0021	1.0	1.0
	511999	$W_TZ_Tjj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 OFP		857000	0.0036	1.0	1.0
	512000	$W_TZ_Tjj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 SFM		467000	0.0021	1.0	1.0
	512001	$W_TZ_Tjj \to \ell\nu\ell\ell jj$ ($\ell = e, \mu$) EW6 SFP		847000	0.0036	1.0	1.0
	514544	$W_TZ_Tjj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 OFM	MadGraph + Pythia8	110000	0.0042	1.0	1.0
	514545	$W_TZ_Tjj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 OFP		190000	0.0073	1.0	1.0
	514546	$W_TZ_Tjj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 SFM		110000	0.0010	1.0	1.0
	514547	$W_TZ_Tjj \to \ell\nu\ell\ell jj$ (min 1 τ) EW6 SFP		190000	0.0018	1.0	1.0

Table G.2: List of the the background samples used in this analysis summarized under the name by which they are referred to in this analysis.

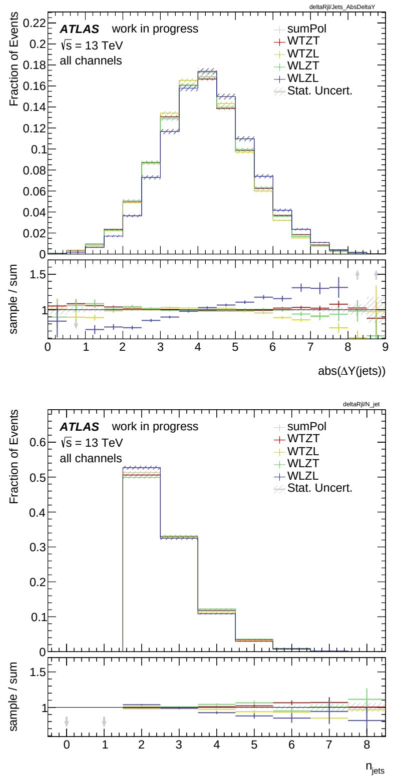
name	DSID	process	Generator	N_{events}	σ in pb	Filter eff.	k-factor
WZ QCD	364253	$WZ \to \ell\nu\ell\ell$ QCD	Sherpa 2.2.2	74M	4.57	1.0	1.0
ZZ	345705	$gg \to \ell\ell\ell\ell$	Sherpa 2.2.2	0.7M	0.0099	1.0	1.0
	345706	$gg \to \ell\ell\ell\ell$	Sherpa 2.2.2	30M	0.010	1.0	1.0
	364250	$q\bar{q} \to ZZ \to \ell\ell\ell\ell$	MadGraph + Pythia8	80M	1.252	1.0	1.0
	364254	$q\bar{q} \to ZZ \to \ell\ell\nu\nu$	MadGraph + Pythia8	70M	12.5	1.0	1.0
	364283	$ZZ \to \ell\ell\ell\ell$ EW6	MadGraph + Pythia8	0.98M	0.011	1.0	1.0
Z+jets	361106	$Z \to ee$	Powheg + Pythia8	1209702350	1901.1	1.0	1.026
	361107	$Z \to \mu\mu$		1193549900	1901.1	1.0	1.026
	361108	$Z \to \tau\tau$		133459000	1901.1	1.0	1.026
ttbar	410470	$t\bar{t}$	Powheg + Pythia8	467952000	729.7	0.54	1.14
single top	410658	t (t-channel)	Powheg + Pythia8	97557400	37.0	1.0	1.19
	410659	$\bar{t}$ (t-channel)		97121350	22.17	1.0	1.18
	410648	Wt (2ℓ)		3909400	3.997	1.0	0.94
	410649	$W\bar{t}$ (2ℓ)		3903900	3.99	1.0	0.95
	410646	Wt		39035000	37.93	1.0	0.94
	410647	$W\bar{t}$		39006000	37.91	1.0	0.95
	410644	t (s-channel)		7803000	2.03	1.0	1.02
	410645	$\bar{t}$ (s-channel)		7817000	1.27	1.0	1.02
ttV	410218	$t\bar{t}ee$	MadGraph + Pythia8	491500	0.037	1.0	1.12
	410219	$t\bar{t}\mu\mu$		4922000	0.037	1.0	1.12
	410155	$t\bar{t}W$		27036000	0.548	1.0	1.1
Z+gamma	366140	$Z\gamma \to ee\gamma$ ($7 < p_T^\gamma < 15$)	Sherpa 2.2.4	97104	46.29	1.0	1.0
	366141	$Z\gamma \to ee\gamma$ ($15 < p_T^\gamma < 35$)		465414	29.28	1.0	1.0
	366142	$Z\gamma \to ee\gamma$ ($35 < p_T^\gamma < 70$)		55151	5.16	1.0	1.0
	366143	$Z\gamma \to ee\gamma$ ($70 < p_T^\gamma < 140$)		48792	0.40	1.0	1.0
	366144	$Z\gamma \to ee\gamma$ ($140 < p_T^\gamma$)		68773	0.053	1.0	1.0
	366145	$Z\gamma \to \mu\mu\gamma$ ($7 < p_T^\gamma < 15$)		135223	46.28	1.0	1.0
	366146	$Z\gamma \to \mu\mu\gamma$ ($15 < p_T^\gamma < 35$)		643346	29.28	1.0	1.0
	366147	$Z\gamma \to \mu\mu\gamma$ ($35 < p_T^\gamma < 70$)		52290	5.16	1.0	1.0
	366148	$Z\gamma \to \mu\mu\gamma$ ($70 < p_T^\gamma < 140$)		68585	0.40	1.0	1.0
	366149	$Z\gamma \to \mu\mu\gamma$ ($140 < p_T^\gamma$)		79303	0.053	1.0	1.0
	366150	$Z\gamma \to \tau\tau\gamma$ ($7 < p_T^\gamma < 15$)		8376	46.25	1.0	1.0
	366151	$Z\gamma \to \tau\tau\gamma$ ($15 < p_T^\gamma < 35$)		24609	29.28	1.0	1.0
	366152	$Z\gamma \to \tau\tau\gamma$ ($35 < p_T^\gamma < 70$)		5624	5.15	1.0	1.0
	366153	$Z\gamma \to \tau\tau\gamma$ ($70 < p_T^\gamma < 140$)		5431	0.40	1.0	1.0
	366154	$Z\gamma \to \tau\tau\gamma$ ($140 < p_T^\gamma$)		9160	0.053	1.0	1.0
VVV	364242	$WWW \to 3\ell3\nu$	Sherpa 2.2.2	220000	0.0072	1.0	1.0
	364243	$WWZ \to 4\ell2\nu$		220000	0.0018	1.0	1.0
	364244	$WWZ \to 2\ell4\nu$		220000	0.0035	1.0	1.0
	364245	$WZZ \to 5\ell1\nu$		190000	0.00019	1.0	1.0
	364246	$WZZ \to 3\ell3\nu$		189000	0.0017	0.45	1.0
	364247	$ZZZ \to 6\ell0\nu$		139000	0.000014	1.0	1.0
	364248	$ZZZ \to 4\ell2\nu$		139000	0.00038	0.22	1.0
	364249	$ZZZ \to 2\ell4\nu$		139000	0.00038	0.44	1.0

H Normalized Distributions of the Polarized Samples at Reco Level

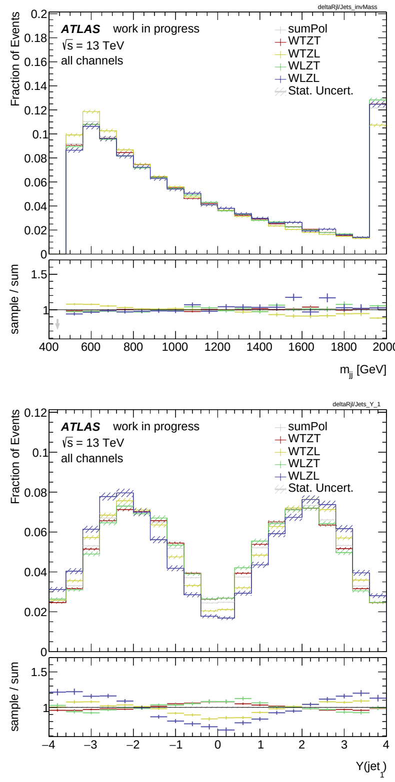
In the following the normalized distributions of the different polarization combinations and the sum of the polarizations at reco level are shown for variables that are used in this analysis. In the subplot the ratios of the individual samples to the sum of the polarizations is shown. The distributions are shown in the truth phase space.


ATLAS work in progress
√s = 13 TeV
all channels
sumPol
WTZT
WTZL
WLZT
WLZL
Stat. Uncert.
Fraction of Events
sample / sum
abs(ΔR(jets))
abs(Δ Φjj)

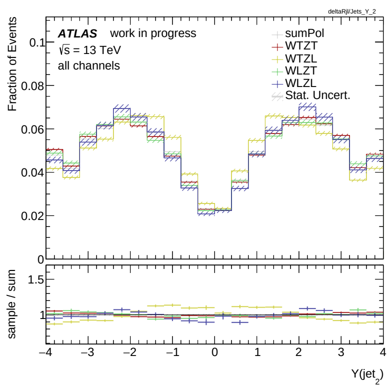
(a)


ATLAS work in progress
√s = 13 TeV
all channels
sumPol
WTZT
WTZL
WLZT
WLZL
Stat. Uncert.
Fraction of Events
sample / sum
abs(ΔY(jets))
njets

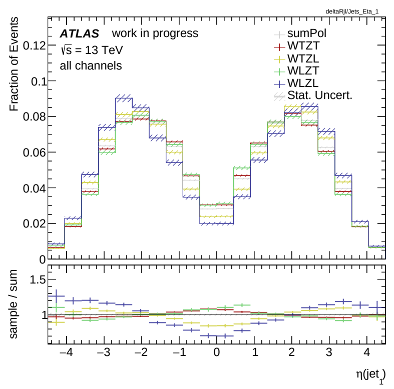
(b)


ATLAS work in progress
√s = 13 TeV
all channels
sumPol
WTZT
WTZL
WLZT
WLZL
Stat. Uncert.
Fraction of Events
sample / sum
mjj [GeV]
Y(jet1)

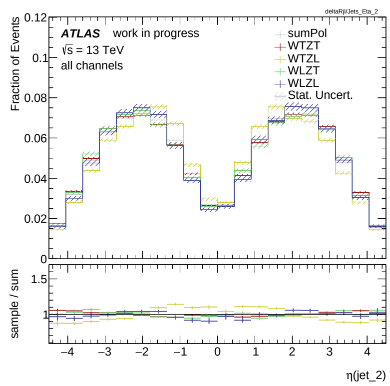
(c)


ATLAS work in progress
√s = 13 TeV
all channels
sumPol
WTZT
WTZL
WLZT
WLZL
Stat. Uncert.
Fraction of Events
sample / sum
Y(jet2)

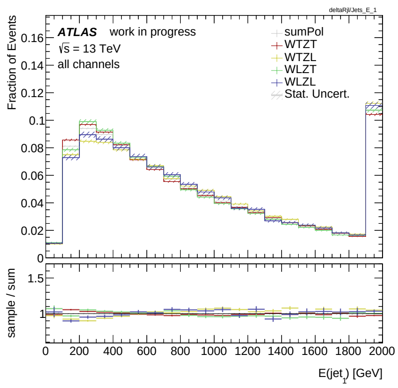
(d)


ATLAS work in progress
√s = 13 TeV
all channels
sumPol
WTZT
WTZL
WLZT
WLZL
Stat. Uncert.
Fraction of Events
sample / sum
η(jet1)

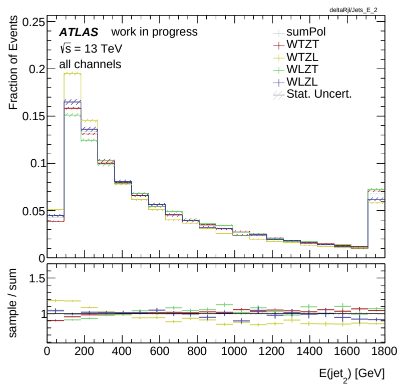
(e)


ATLAS work in progress
√s = 13 TeV
all channels
sumPol
WTZT
WTZL
WLZT
WLZL
Stat. Uncert.
Fraction of Events
sample / sum
η(jet_2)

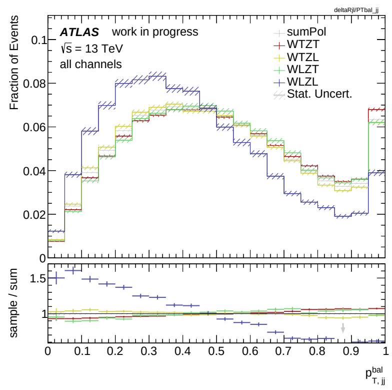
(f)


ATLAS work in progress
√s = 13 TeV
all channels
sumPol
WTZT
WTZL
WLZT
WLZL
Stat. Uncert.
Fraction of Events
sample / sum
E(jet1) [GeV]

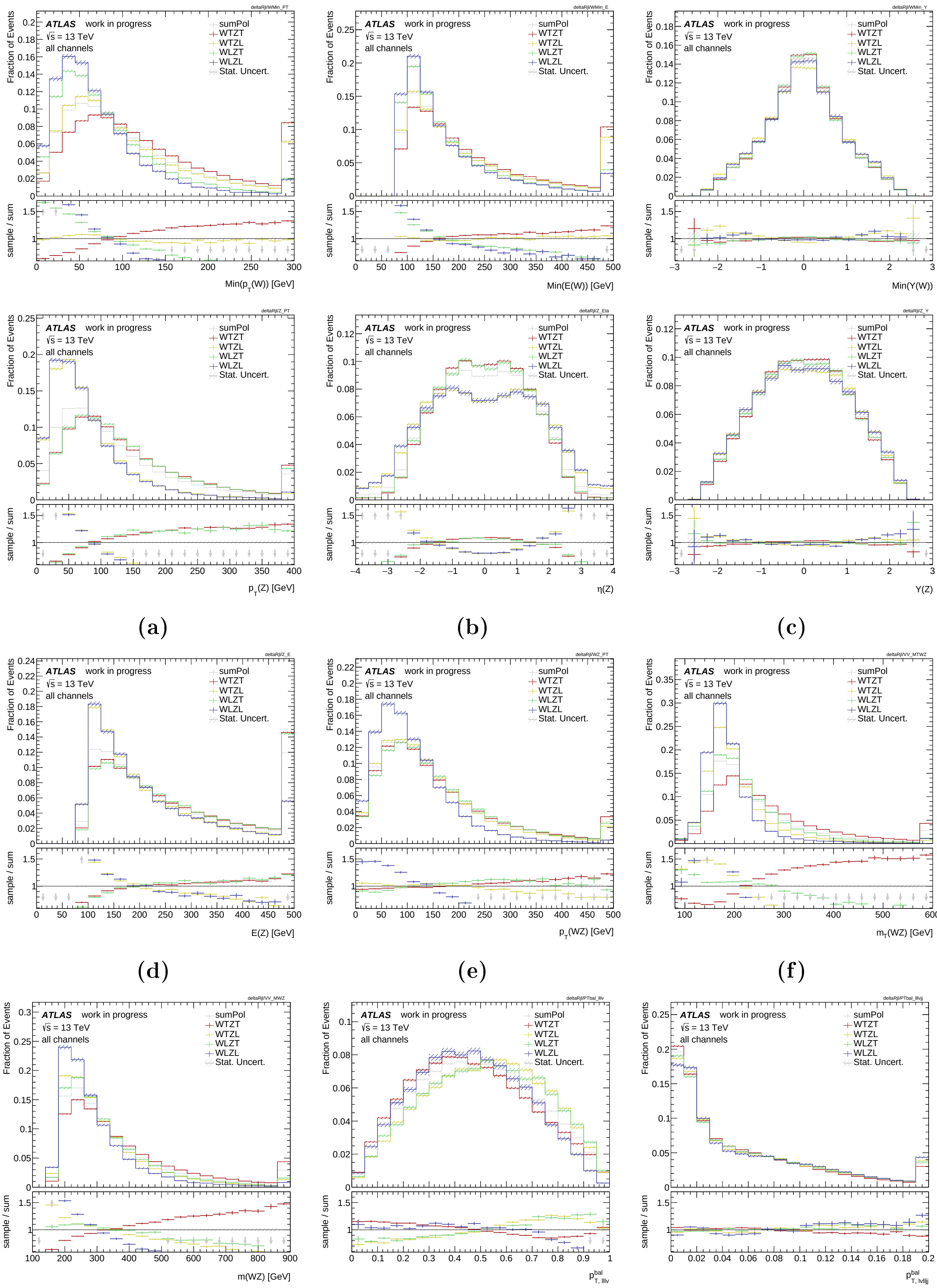
(g)


ATLAS work in progress
√s = 13 TeV
all channels
sumPol
WTZT
WTZL
WLZT
WLZL
Stat. Uncert.
Fraction of Events
sample / sum
E(jet2) [GeV]

(h)


ATLAS work in progress
√s = 13 TeV
all channels
sumPol
WTZT
WTZL
WLZT
WLZL
Stat. Uncert.
Fraction of Events
sample / sum
pbalT, jj

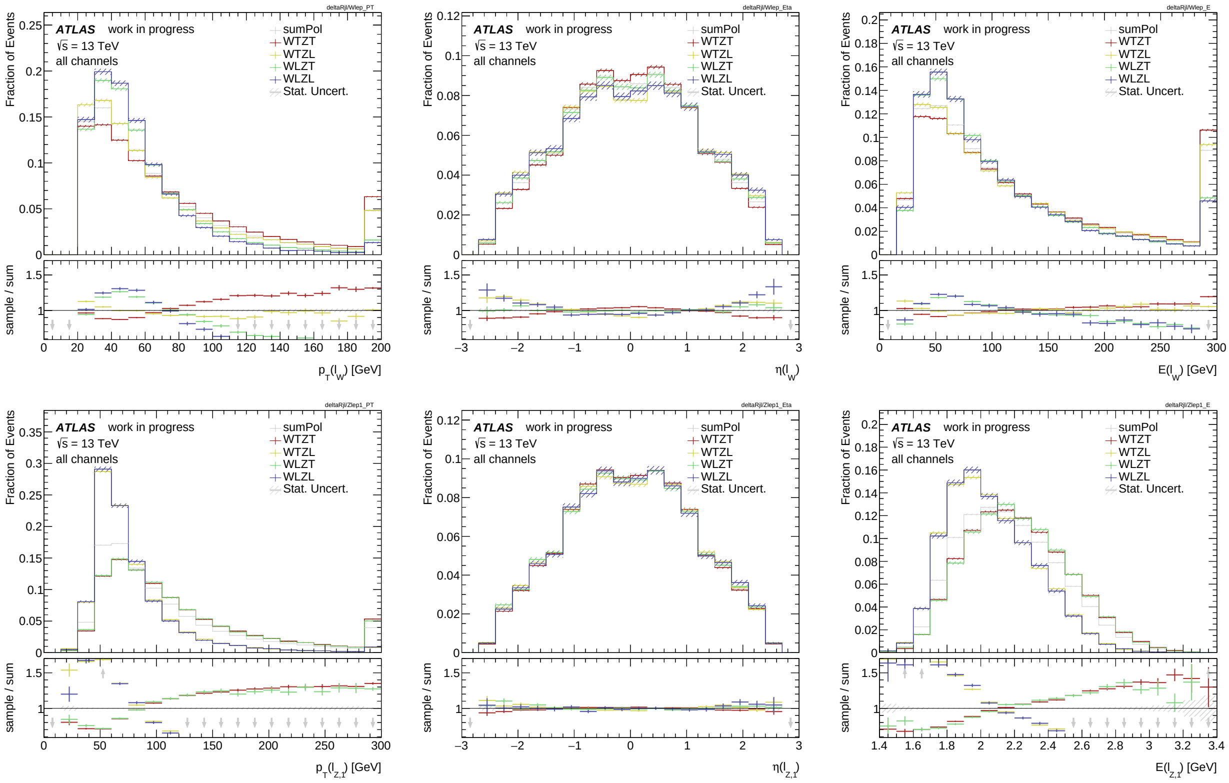
(i)


ATLAS work in progress
√s = 13 TeV
all channels
sumPol
WTZT
WTZL
WLZT
WLZL
Stat. Uncert.
Fraction of Events
sample / sum
Min(p_T(W)) [GeV]
Min(E(W)) [GeV]
Min(Y(W))
p_T(Z) [GeV]
η(Z)
Y(Z)
E(Z) [GeV]
p_T(WZ) [GeV]
m_T(WZ) [GeV]
m(WZ) [GeV]
p_T, lllν^bal
p_T, lνlljj^bal

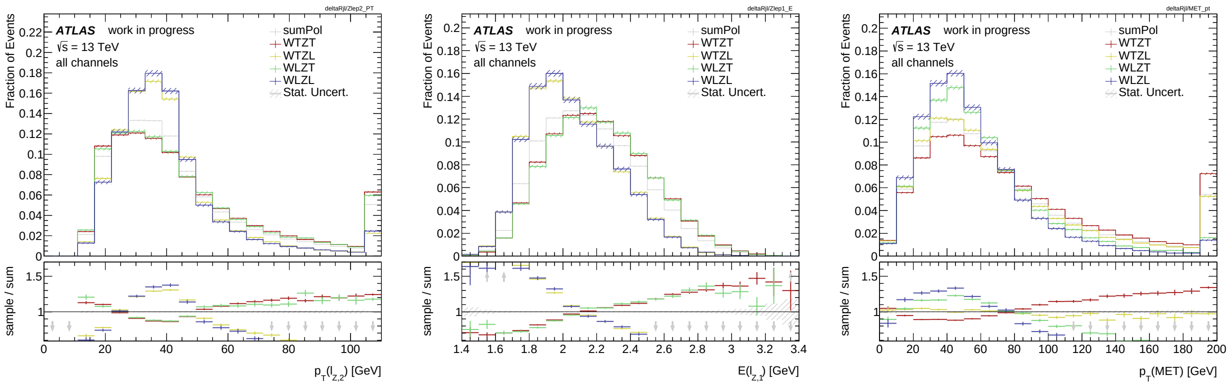
(a) (b) (c)

(d) (e) (f)

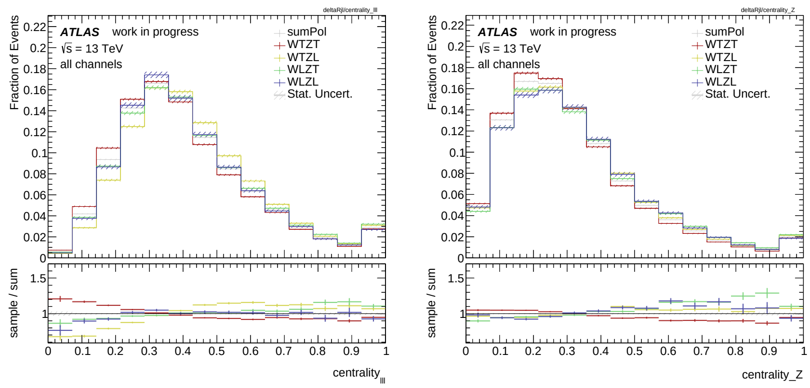
(g) (h) (i)


ATLAS work in progress
√s = 13 TeV
all channels
sumPol
WTZT
WTZL
WLZT
WLZL
Stat. Uncert.
Fraction of Events
sample / sum
p_T(l_W) [GeV]
η(l_W)
E(l_W) [GeV]
p_T(l_Z,1) [GeV]
η(l_Z,1)
E(l_Z,1) [GeV]

(a) (b) (c)


ATLAS work in progress
√s = 13 TeV
all channels
sumPol
WTZT
WTZL
WLZT
WLZL
Stat. Uncert.
Fraction of Events
sample / sum
p_T(l_Z,2) [GeV]
E(l_Z,1) [GeV]
p_T(MET) [GeV]

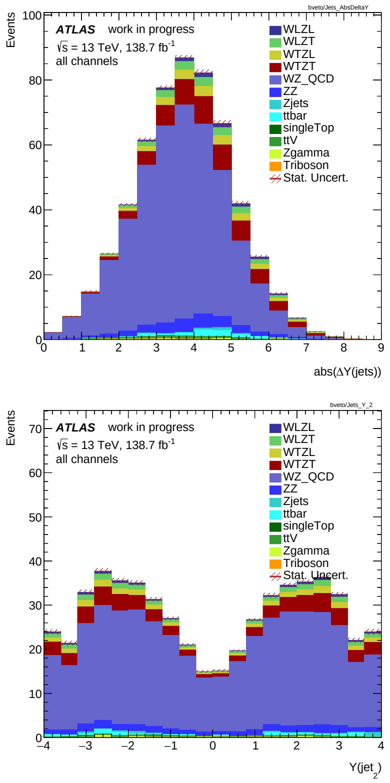
(d) (e) (f)


ATLAS work in progress
√s = 13 TeV
all channels
sumPol
WTZT
WTZL
WLZT
WLZL
Stat. Uncert.
Fraction of Events
sample / sum
centrality_lll
centrality_Z

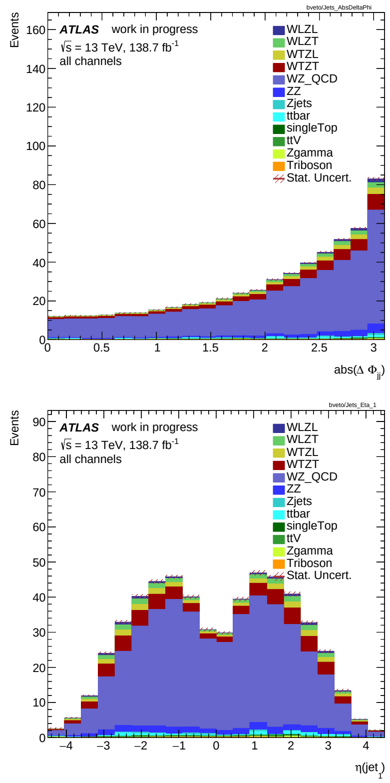
(g) (h)

I Background Distributions

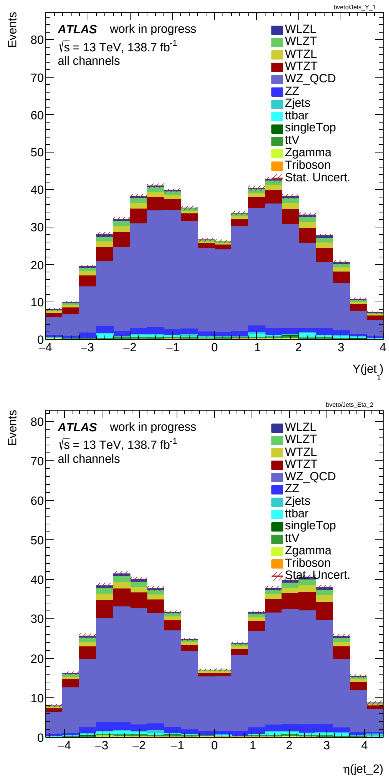
In the following the stacked distributions of the different polarization combinations and the non-polarization backgrounds are shown for variables used in this analysis. The distributions are shown in the reco phase space.


ATLAS work in progress
$\sqrt{s}$ = 13 TeV, 138.7 fb^{-1}
all channels
WLZL
WLZT
WTZL
WTZT
WZ_QCD
ZZ
Zjets
ttbar
singleTop
ttV
Zgamma
Triboson
Stat. Uncert.
Events
abs(ΔY(jets))
Y(jet_2)

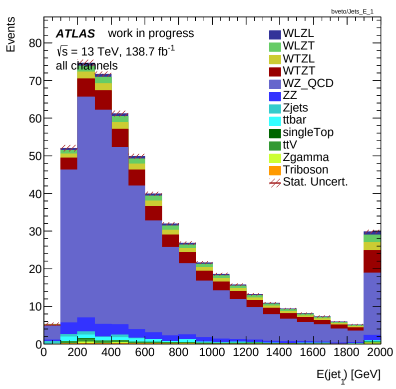
(a)


ATLAS work in progress
$\sqrt{s}$ = 13 TeV, 138.7 fb^{-1}
all channels
WLZL
WLZT
WTZL
WTZT
WZ_QCD
ZZ
Zjets
ttbar
singleTop
ttV
Zgamma
Triboson
Stat. Uncert.
Events
abs(Δ Φ_jj)
η(jet_1)

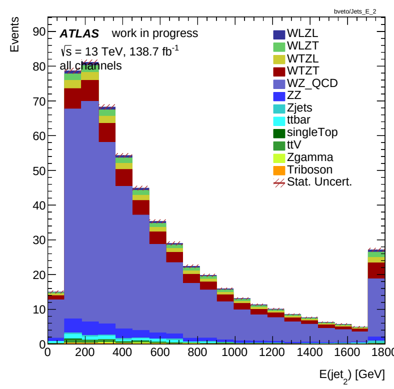
(b)


ATLAS work in progress
$\sqrt{s}$ = 13 TeV, 138.7 fb^{-1}
all channels
WLZL
WLZT
WTZL
WTZT
WZ_QCD
ZZ
Zjets
ttbar
singleTop
ttV
Zgamma
Triboson
Stat. Uncert.
Events
Y(jet_1)
η(jet_2)

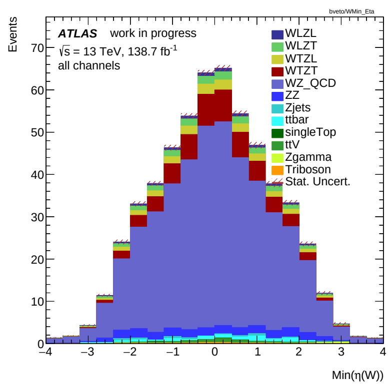
(c)


ATLAS work in progress
$\sqrt{s}$ = 13 TeV, 138.7 fb^{-1}
all channels
WLZL
WLZT
WTZL
WTZT
WZ_QCD
ZZ
Zjets
ttbar
singleTop
ttV
Zgamma
Triboson
Stat. Uncert.
Events
E(jet_1) [GeV]

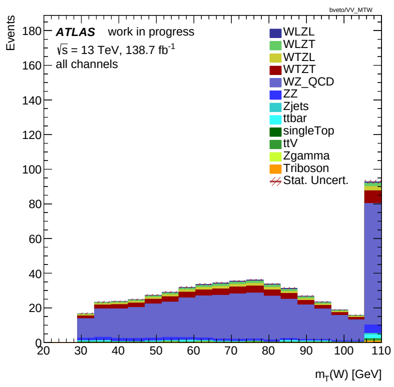
(d)


ATLAS work in progress
$\sqrt{s}$ = 13 TeV, 138.7 fb^{-1}
all channels
WLZL
WLZT
WTZL
WTZT
WZ_QCD
ZZ
Zjets
ttbar
singleTop
ttV
Zgamma
Triboson
Stat. Uncert.
Events
E(jet_2) [GeV]

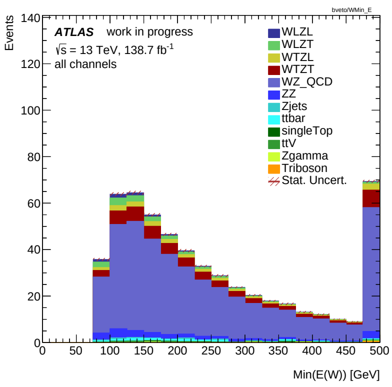
(e)


ATLAS work in progress
$\sqrt{s}$ = 13 TeV, 138.7 fb^{-1}
all channels
WLZL
WLZT
WTZL
WTZT
WZ_QCD
ZZ
Zjets
ttbar
singleTop
ttV
Zgamma
Triboson
Stat. Uncert.
Events
Min(η(W))

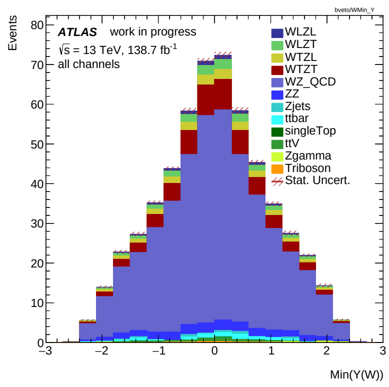
(f)


ATLAS work in progress
$\sqrt{s}$ = 13 TeV, 138.7 fb^{-1}
all channels
WLZL
WLZT
WTZL
WTZT
WZ_QCD
ZZ
Zjets
ttbar
singleTop
ttV
Zgamma
Triboson
Stat. Uncert.
Events
m_T(W) [GeV]

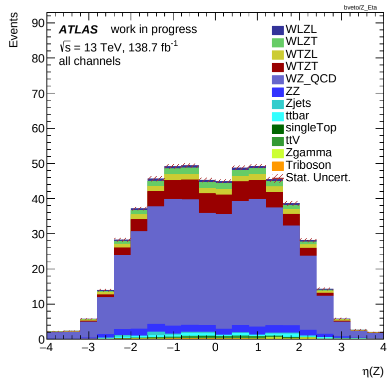
(g)

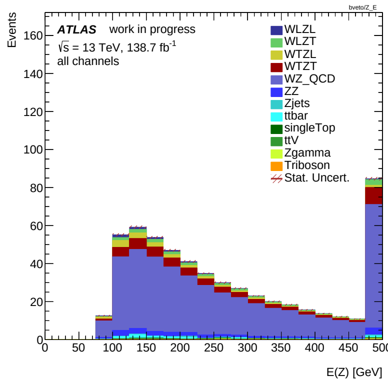

ATLAS work in progress
$\sqrt{s}$ = 13 TeV, 138.7 fb^{-1}
all channels
WLZL
WLZT
WTZL
WTZT
WZ_QCD
ZZ
Zjets
ttbar
singleTop
ttV
Zgamma
Triboson
Stat. Uncert.
Events
Min(E(W)) [GeV]

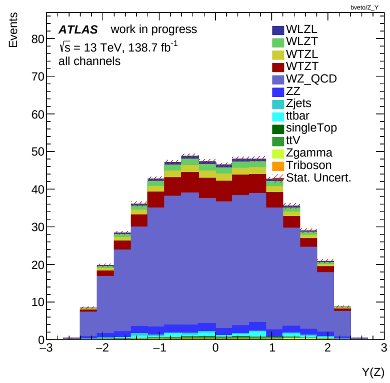
(h)

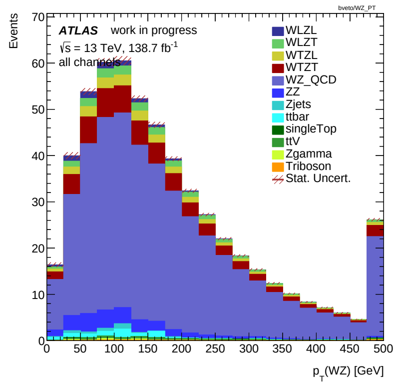

ATLAS work in progress
$\sqrt{s}$ = 13 TeV, 138.7 fb^{-1}
all channels
WLZL
WLZT
WTZL
WTZT
WZ_QCD
ZZ
Zjets
ttbar
singleTop
ttV
Zgamma
Triboson
Stat. Uncert.
Events
Min(Y(W))

(i)

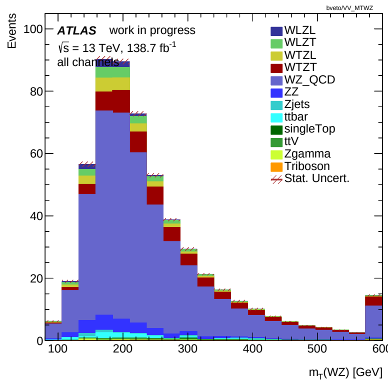

ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
η(Z)
Events


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
E(Z) [GeV]
Events

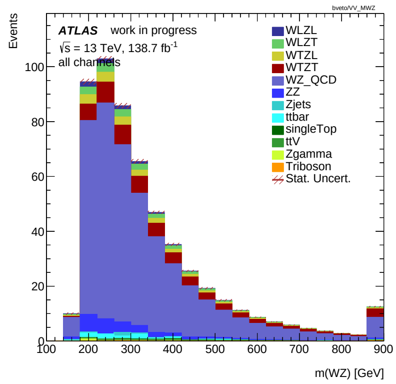

ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
Y(Z)
Events


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
pT(WZ) [GeV]
Events

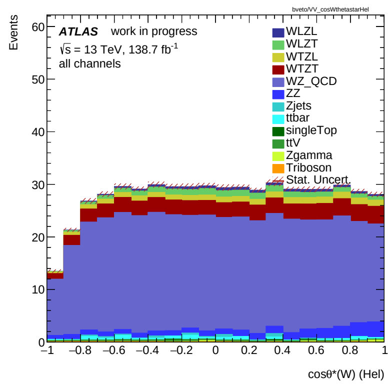
(a)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
mT(WZ) [GeV]
Events

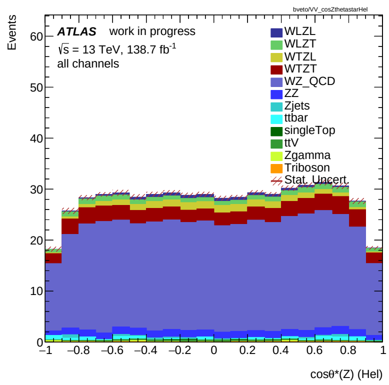
(b)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
m(WZ) [GeV]
Events

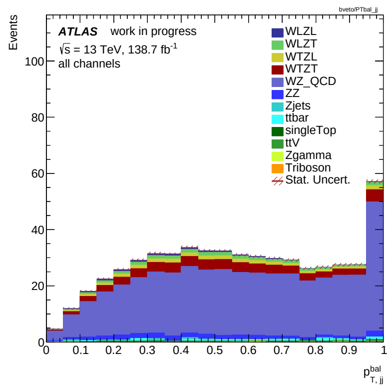
(c)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
cosθ*(W) (Hel)
Events

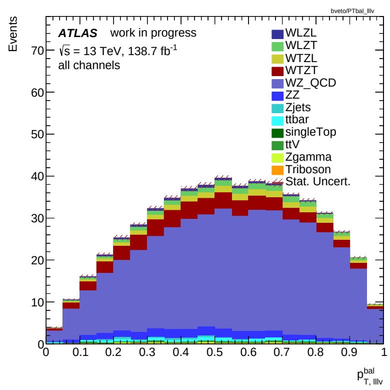
(d)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
cosθ*(Z) (Hel)
Events

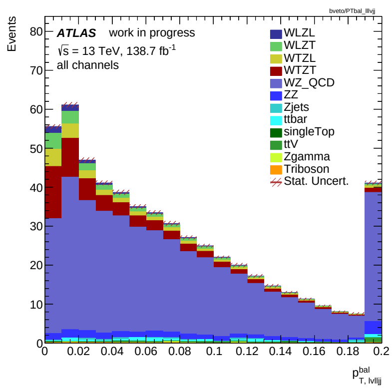
(e)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
pbal T, jj
Events

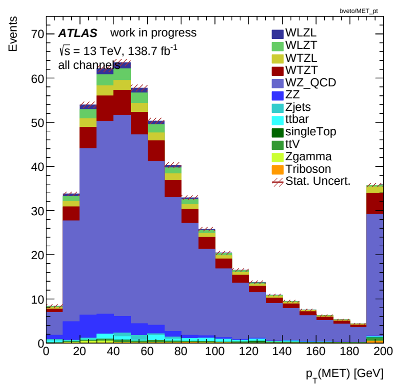
(f)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
pbal T, lllv
Events

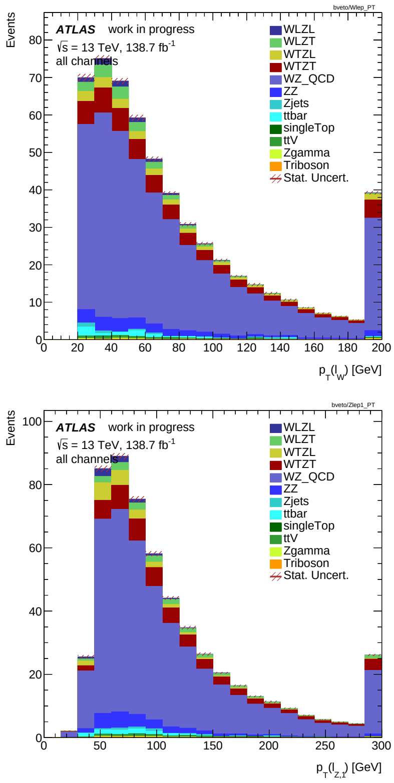
(g)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
pbal T, lvllj
Events

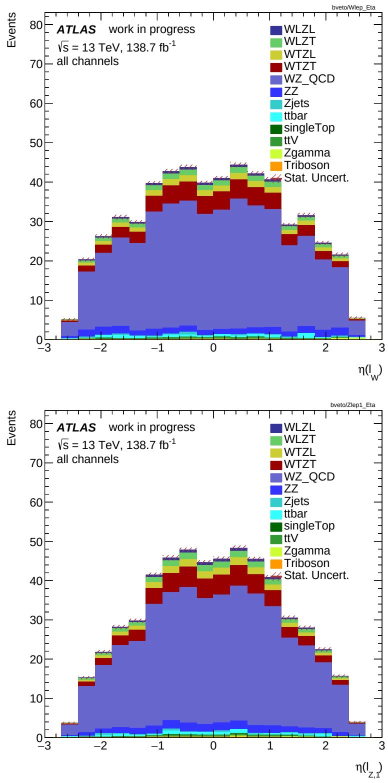
(h)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
pT(MET) [GeV]
Events

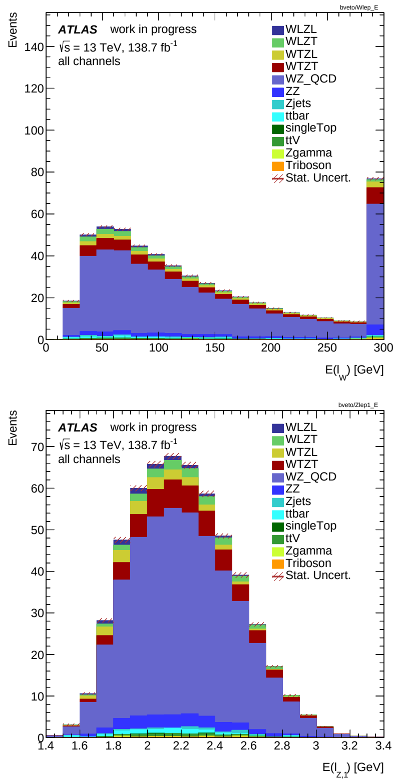
(i)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
WLZL
WLZT
WTZL
WTZT
WZ_QCD
ZZ
Zjets
ttbar
singleTop
ttV
Zgamma
Triboson
Stat. Uncert.
Events
p_T(l_W) [GeV]
p_T(l_Z,1) [GeV]

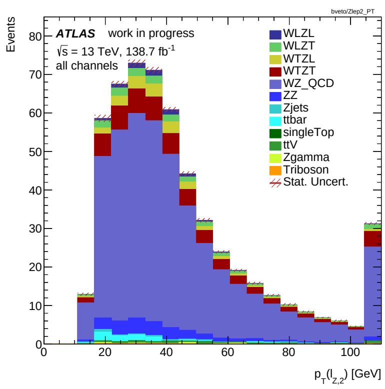
(a)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
Events
η(l_W)
η(l_Z,1)

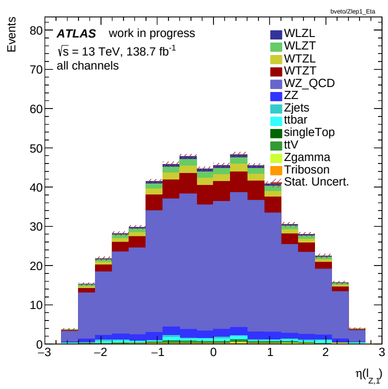
(b)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
Events
E(l_W) [GeV]
E(l_Z,1) [GeV]

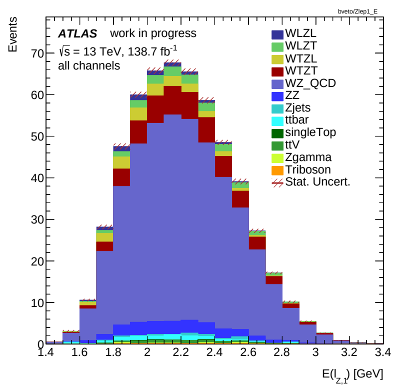
(c)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
Events
p_T(l_Z,2) [GeV]

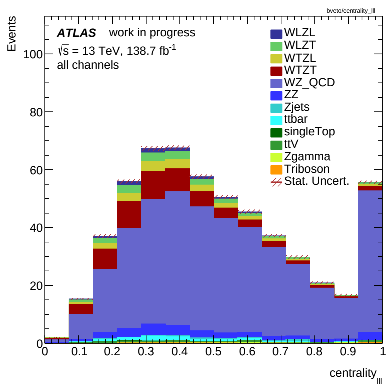
(d)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
Events
η(l_Z,1)

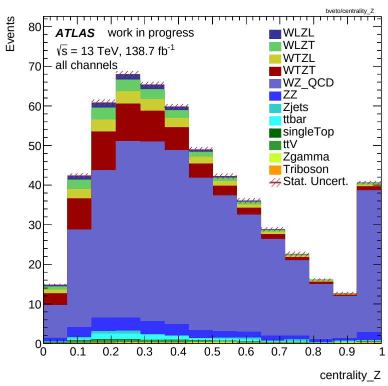
(e)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
Events
E(l_Z,1) [GeV]

(f)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
Events
centrality_lll

(g)


ATLAS work in progress
√s = 13 TeV, 138.7 fb-1
all channels
Events
centrality_Z

(h)

J Ranking of the Input Variables

Table J.1: The five highest ranking variables for the WLZX BDT that are considered during the optimization. They are ranked by decreasing non-overlap between signal and background. The same preprocessing as in Table 6.4 is applied.

ranking	Observable	f_{N-O}	f_{Sep}	$f_{red\chi^2}$	$f_{Kolmogorov}$	f_{ROC}
1	m_{jj}	0.24	0.082	387.98	0.24	0.63
2	$p_T(W)$	0.24	0.083	478.83	0.24	0.61
3	n_{jets}	0.23	0.084	604.16	0.23	0.50
4	$p_T^{bal}(lllvjj)$	0.22	0.065	324.70	0.22	0.59
5	$\Delta R(jets)$	0.21	0.069	395.34	0.21	0.60

Table J.2: The five highest ranking variables for the WLZX BDT that are considered during the optimization. They are ranked by decreasing non-overlap between signal and background. The same preprocessing as in Table 6.4 is applied.

ranking	Observable	f_{N-O}	f_{Sep}	$f_{red\chi^2}$	$f_{Kolmogorov}$	f_{ROC}
1	$p_T(Z)$	0.25	0.085	519.16	0.25	0.61
2	n_{jets}	0.23	0.088	656.30	0.23	0.50
3	m_{jj}	0.23	0.073	353.98	0.23	0.62
4	$\cos\theta(W)$	0.23	0.068	341.61	0.23	0.62
5	$p_T^{bal}(lllvjj)$	0.22	0.064	326.76	0.22	0.59

Table J.3: The five highest ranking variables for the min1L BDT that are considered during the optimization. They are ranked by decreasing non-overlap between signal and background. The same preprocessing as in Table 6.4 is applied.

ranking	Observable	f_{N-O}	f_{Sep}	$f_{red\chi^2}$	$f_{Kolmogorov}$	f_{ROC}
1	m_{jj}	0.25	0.082	527.26	0.25	0.63
2	n_{jets}	0.24	0.089	753.85	0.24	0.51
3	$p_T^{bal}(lllvjj)$	0.23	0.071	472.90	0.23	0.60
4	$\Delta R(jets)$	0.21	0.068	496.05	0.21	0.60
5	$\cos\theta(W)$	0.21	0.055	347.00	0.21	0.60

List of Figures

2.1 Feynman diagrams of the interactions of the Higgs boson with gauge bosons. . . . 9
2.2 Feynmandiagrams contributing to the WZ VBS process. 10
2.3 Feynmandiagrams of the WZ VBS process as it appears at the LHC. 11
2.4 Example of non-VBS Feynman diagrams with same final state and same coupling order as the WZ VBS process. 12
2.5 Example of a Feynman diagram of the tZj process. 12
2.6 Example of WZjj QCD Feynman diagrams. 13
2.7 Dependence of the cross section of the scattering of a W and a Z boson on the centre of mass energy [3]. 14
2.8 Visualization of the decay angle θ^* 15
2.9 Visualization of spin and momenta in the decay of a "W^+ boson [3]. 16
2.10 Distribution of the cross sections over the cosine of the decay angle θ^* 16
2.11 Schematic visualization of the spin and momenta in the scattering of a W boson from quarks. 17

3.1 Cut away view of the ATLAS detector [14]. 20
3.2 Cut away view of the ATLAS inner detector [14]. 22
3.3 Cut away view of the ATLAS calorimeter system [14]. 23
3.4 Cut away view of the ATLAS muon system [14]. 24

4.1 Pictorial representation of a $t\bar{t}H$ event as produced by an event generator. 26

5.1 Distributions of the full, unpolarized and polarized sample 32
5.2 Normalized distribution of the cosine of the decay angle $\cos\theta^*$ of the W boson and the Z boson for the longitudinal longitudinal sample and the analytical function. . 33
5.3 Distributions of the four polarization combinations and the full sample in the truth phase space 37
5.4 Distributions of the four polarization combinations and the full sample in the truth phase space 38
5.5 Distributions of the four polarization combinations and the full sample in the truth phase space 39

5.6 Distributions of the four polarization combinations and the full sample in the truth phase space . . . 40
5.7 Distributions of the four polarization combinations and the full sample in the truth phase space . . . 41
5.8 Normalized distributions of the different polarization combinations in the helicity frame and in the parton parton frame . . . 44
5.9 Normalized distributions of the different polarization combinations in the helicity frame and in the parton parton frame . . . 45
5.10 Comparison of the normalized distributions of the four polarization combinations between the locally and officially produced samples . . . 46
5.11 Distributions of the different polarization combinations and the full sample at truth and reco level in the truth phase space. . . . 52
5.12 Normalized distributions of the different polarization combinations and the sum of the polarizations at truth and reco level in the truth phase space. . . . 53
5.13 Normalized distributions of the different polarization combinations and the sum of the polarizations at truth and reco level in the truth phase space. . . . 54
5.14 Distributions of the non-polarization backgrounds and the different polarization combinations. . . . 57

6.1 Example of a single decision tree. . . . 62
6.2 Development of output of the Kolmogorow-Smirnow test of the signal during the optimization. . . . 71
6.3 BDT distributions for the WLZL BDT. . . . 73
6.4 BDT distributions for the WLZX BDT. . . . 74
6.5 BDT distributions for the WXZL BDT. . . . 75
6.6 BDT distributions for the min1L BDT. . . . 76
6.7 BDT score distributions and Asimov-data after the binning optimization. . . . 86
6.8 Correlation and pull plot for the WLZL fit, where the polarization background is unconstrained. . . . 87
6.9 Correlation and pull plot for the WLZL fit, where the polarization background has a 30% systematic uncertainty. . . . 88
6.10 Correlation and pull plot for the WLZL fit, where the polarization background is fixed . . . 89
6.11 Pre- and post-fit BDT score distributions of the WLZL BDT and the WLZX BDT after the binning optimization for the fit to pseudo data. . . . 90
6.12 Pre- and post-fit BDT score distributions of the WXZL BDT and the min1L BDT after the binning optimization for the fit to pseudo data. . . . 91

7.1 Plan for the LHC/HL LHC . . . 93

7.2 Significances that result from the fit with the different signal configurations for different integrated luminosities. 97
7.3 Expected significances from the study done in [41]. 98

C.1 Distributions of the different polarization combinations at truth level 112
C.2 Distributions of the different polarization combinations at truth level 113
C.3 Distributions of the different polarization combinations at truth level 114
C.4 Distributions of the different polarization combinations at truth level 115
C.5 Distributions of the different polarization combinations at truth level 116
C.6 Distributions of the different polarization combinations at truth level 117

D.1 Distributions for samples where the polarization is defined in the parton parton frame by setting `me_frame = 1,2` or `me_freme = 3,4,5,6,7,8` 120

E.1 Distributions of the four polarization combinations for samples with at least one τ in the final state of the hard process. 122

F.1 Comparison of the normalized distributions of the four polarization combinations between the locally and officially produced samples 123
F.2 Comparison of the normalized distributions of the four polarization combinations between the locally and officially produced samples 124
F.3 Comparison of the normalized distributions of the four polarization combinations between the locally and officially produced samples 125

List of Tables

2.1 Overview of the bosons described in the SM and their associated interactions. . . 4
2.2 Overview of the fermions described by the SM and their corresponding charges. . 6

4.1 Generator level cuts used for the simulation of the polarized, unpolarized and full sample. 28

5.1 Phase space definition for the truth level analysis. 31
5.2 Yields and ratios of the full, the unpol and the pol sample in the truth phase space. 31
5.3 Yields of the full sample, the sum of all polarization samples and the individual samples of the four polarization combinations in the truth phase space 35
5.4 Yields and fractions of the polarized samples in the helicity and the parton parton frame. 42
5.5 Event yields of the full and the four polarized samples for samples with only electrons and muons and samples with at least one τ in the final state of the hard process. 42
5.6 Comparison of the event yields of the locally and officially produced samples. . . 43
5.7 Electron object selection used in this analysis. 49
5.8 Muon object selection used in this analysis. 49
5.9 Event yields of the official sample at truth and reco level in the truth phase space. 50
5.10 Definition of the reco phase space. 55
5.11 Event yields of the non-polarization backgrounds and the different polarization combinations in the reco phase space. 56
5.12 Fraction of events with at least one τ in the final state of the hard process of the respective polarization combination. 56

6.1 Event yields of the signal and background of the respective BDT. 60
6.2 Metaparameters used at the start of the optimization. 65
6.3 Range of values of the metaparameters that are tried during the metaparameter optimization. 66
6.4 List of input variables for the WLZL BDT that are considered during the optimization. 68
6.5 Metaparameters used in the final BDTs as a result of the metaparameter optimization. 69
6.6 Variables used in the final BDTs as a result of the optimization. 70

6.7 Values of some FoM for the final BDTs. 72
6.8 Systematic uncertainties of the non polarization backgrounds. 81

7.1 Significances that result from the different WLZL fits for different integrated luminosities. 95
7.2 Significances that result from the different WLZX fits for different integrated luminosities. 95
7.3 Significances that result from the different WXZL fits for different integrated luminosities. 95
7.4 Significances that result from the different min1L fits for different integrated luminosities. 96
7.5 Cross sections of the polarization samples produced at $\sqrt{s} = 13$ TeV and $\sqrt{s} = 13.6$ TeV. 100

B.1 List of locally produced samples in the helicity frame 109
B.2 List of locally produced samples in the parton parton frame 110

D.1 Cutflow table of samples with polarization defined in the parton parton frame by setting `me_frame = 1,2` or `me_frame = 3,4,5,6,7,8` 119

G.1 List of the full and the polarized signal samples used in this analysis at reconstruction level. 127
G.2 List of the the background samples used in this analysis. 128

J.1 The five highest ranking variables for the WLZX BDT that are considered during the optimization. 137
J.2 The five highest ranking variables for the WLZX BDT that are considered during the optimization. 137
J.3 The five highest ranking variables for the min1L BDT that are considered during the optimization. 137

Bibliography

[1] Morad Aaboud et al. 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*, 793:469–492, 2019.

[2] Albert M Sirunyan et al. Measurements of production cross sections of WZ and same-sign WW boson pairs in association with two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 809:135710, 2020.

[3] Carsten Bittrich. Study of Polarization Fractions in the Scattering of Massive Gauge Bosons $W^{\pm}Z \to W^{\pm}Z$ with the ATLAS Detector at the Large Hadron Collider, Apr 2015. Presented 2015.

[4] Jan-Eric Nitschke. Studies on polarization in the scattering of two vector bosons with the ATLAS experiment at the LHC, Nov 2019. Presented 01 Nov 2019.

[5] E. Noether. Invariante variationsprobleme. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1918:235–257, 1918.

[6] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, Feb 1957.

[7] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579–588, 1961.

[8] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc. C*, 680519:367–377, 1968.

[9] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

[10] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[11] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[12] Vryonidou E. Stirling, W.J. Electroweak gauge boson polarisation at the lhc. *J. High Energ. Phys.*, 2012.

[13] Lyndon Evans and Philip Bryant. LHC machine. *Journal of Instrumentation*, 3(08):S08001–S08001, aug 2008.

[14] The ATLAS Collaboration. The ATLAS experiment at the CERN large hadron collider. *Journal of Instrumentation*, 3(08):S08003–S08003, aug 2008.

[15] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004–S08004, aug 2008.

[16] The ALICE Collaboration. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002–S08002, aug 2008.

[17] The LHCb Collaboration. The LHCb detector at the LHC. *Journal of Instrumentation*, 3(08):S08005–S08005, aug 2008.

[18] T Gleisberg, S Höche, F Krauss, M Schönherr, S Schumann, F Siegert, and J Winter. Event generation with SHERPA 1.1. *Journal of High Energy Physics*, 2009(02):007–007, feb 2009.

[19] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), jul 2014.

[20] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to pythia 8.2. *Computer Physics Communications*, 191:159–177, 2015.

[21] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *Journal of High Energy Physics*, 2006(05):026–026, may 2006.

[22] Manuel Bähr, Stefan Gieseke, Martyn A. Gigg, David Grellscheid, Keith Hamilton, Oluseyi Latunde-Dada, Simon Plätzer, Peter Richardson, Michael H. Seymour, Alexander Sherstnev, and Bryan R. Webber. Herwig++ physics and manual. *The European Physical Journal C*, 58(4):639–707, nov 2008.

[23] Johannes Bellm, Stefan Gieseke, David Grellscheid, Simon Plätzer, Michael Rauch, Christian Reuschle, Peter Richardson, Peter Schichtel, Michael H. Seymour, Andrzej Siódmok, Alexandra Wilcock, Nadine Fischer, Marco A. Harrendorf, Graeme Nail, Andreas Papaefstathiou, and Daniel Rauch. Herwig 7.0/herwig++ 3.0 release note. *The European Physical Journal C*, 76(4), apr 2016.

[24] Enrico Bothmann, Gurpreet Singh Chahal, Stefan Höche, Johannes Krause, Frank Krauss, Silvan Kuttimalai, Sebastian Liebschner, Davide Napoletano, Marek Schönherr, Holger Schulz, Steffen Schumann, and Frank Siegert. Event generation with sherpa 2.2. *SciPost Physics*, 7(3), sep 2019.

[25] Diogo Buarque Franzosi, Olivier Mattelaer, Richard Ruiz, and Sujay Shil. Automated predictions from polarized matrix elements. *Journal of High Energy Physics*, 2020(4), apr 2020.

[26] Baptiste Cabouat and Torbjörn Sjöstrand. Some dipole shower studies. *The European Physical Journal C*, 78(3), mar 2018.

[27] Barbara Jäger, Alexander Karlberg, and Johannes Scheller. Parton-shower effects in electroweak WZjj production at the next-to-leading order of QCD. *The European Physical Journal C*, 79(3), mar 2019.

[28] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5: going beyond. *Journal of High Energy Physics*, 2011(6), jun 2011.

[29] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), jul 2014.

[30] R. L. Workman and Others. Review of Particle Physics. *PTEP*, 2022:083C01, 2022.

[31] Bittrich et al. Study 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. Technical report, CERN, Geneva, 2021.

[32] Jira ticket for the polarization samples. `https://its.cern.ch/jira/browse/ATLMCPROD-9836`.

[33] A. et al. Hoecker. Tmva - toolkit for multivariate data analysis, 2007.

[34] Carsten Bittrich. Evidence for Scattering of Electroweak Gauge Bosons in the $W^{\pm}Z$ Channel with the ATLAS Detector at the Large Hadron Collider, 2020. Presented 14 Jul 2020.

[35] Wouter Verkerke and David Kirkby. The roofit toolkit for data modeling, 2003.

[36] Lorenzo Moneta, Kevin Belasco, Kyle Cranmer, Sven Kreiss, Alfio Lazzaro, Danilo Piparo, Gregory Schott, Wouter Verkerke, and Matthias Wolf. The roostats project, 2010.

[37] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2), feb 2011.

[38] Abraham Wald. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54(3):426–482, 1943.

[39] Aberle et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report.* CERN Yellow Reports: Monographs. CERN, Geneva, 2020.

[40] Lhc/ hl-lhc plan (last update february 2022). `https://hilumilhc.web.cern.ch/content/hl-lhc-project`. (visited on 17.10.2022).

[41] Andrea Dainese, Michelangelo Mangano, Andreas B Meyer, Aleandro Nisati, Gavin Salam, and Mika Anton Vesterinen. Report on the Physics at the HL-LHC, and Perspectives for the HE-LHC. Technical report, Geneva, Switzerland, 2019.

[42] https://answers.launchpad.net/mg5amcnlo/+question/701071.

Danksagung

Ich möchte mich hier bei all den Menschen bedanken, die mir während meines Studiums und meiner Masterarbeit begleitet und mir geholfen haben.
Zunächst möchte ich bei Frank Siegert für die Möglichkeit in seiner Arbeitsgruppe an diesem interessanten Thema zu arbeiten bedanken. Ein großer Dank gilt auch an Jan Erik Nitschke, der mich während meiner Masterarbeit betreut hat und immer Zeit für Rat und Fragen gehabt hat. I also want to thank Joany Manjarres for all the guidance and help she gave me during this thesis. Many thanks also the whole working group for the great atmosphere that made working there a very nice experience. Für das Korrekturlesen meiner Arbeit möchte ich mich bei Jan-Eric Nitschke, Max Stange, Tim Hermann, Erik Bachmann und Felix Bela bedanken.
Vielen Dank auch an meine Eltern, die mich immer unterstützt und motiviert haben.
Zuletzt möchte ich mich auch bei all meinen Freunden für all die schönen Erinnerungen an gemeinsame Aktivitäten während des Studiums danken.

Erklärung

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

Anna Sophie Tegetmeier
Dresden, November 2022