

**Evidence for Scattering of
Electroweak Gauge Bosons in the
 $W^\pm Z$ Channel with the ATLAS Detector
at the Large Hadron Collider**

Dissertation

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

vorgelegt der
Fakultät für Physik
der Technischen Universität Dresden

von
M.Sc. Carsten Bittrich
geboren am 09.02.1989 in Görlitz

Eingereicht am 08. April 2020

Verteidigt am 14. Juli 2020

1. Gutachter: Prof. Dr. Michael Kobel
2. Gutachter: Prof. Dr. Sabine Lammers

Zusammenfassung

Das Standardmodell der Elementarteilchenphysik (SM) ist die fundamentale Theorie, um Elementarteilchen und deren wichtigste Wechselwirkungen, die elektromagnetische, die schwache und die starke Wechselwirkung, zu beschreiben. Die zugrundeliegende Symmetrie, die die schwache und die starke Wechselwirkung im SM beschreibt, ist komplexer als die der elektromagnetischen Wechselwirkung. Diese Symmetrie verlangt die Existenz von direkten Drei- und Vierteilchenwechselwirkungen zwischen den Botenteilchen der schwachen und starken Wechselwirkung, auch Eichbosonen genannt. Diese Selbstwechselwirkungen existieren nicht für das Botenteilchen der elektromagnetischen Wechselwirkung, dem Photon. Während die Dreiteilchenwechselwirkungen bereits an früheren Beschleunigerexperimenten studiert wurden, sind die direkten Vierteilchenwechselwirkungen eine direkte Vorhersage der Eichstruktur des SM, welche zu Beginn dieser Studie für die schwache Wechselwirkung noch nicht beobachtet wurde.

Ein Prozess, zu dem sowohl die Drei- als auch die Vierteilchenwechselwirkungen beitragen, ist die Streuung der Botenteilchen der elektroschwachen Wechselwirkung W^\pm, Z, γ auch Vektorbosonstreuung (VBS) genannt. Im SM gibt es neben den Selbstwechselwirkungen auch Beiträge durch den Austausch eines Higgs-bosons. Diese Beiträge sind sensitiv auf die Eigenschaften des Higgs-bosons, insbesondere die Kopplungen zu den Botenteilchen und auf die Details des Mechanismus durch den die W^\pm - und Z -Botenteilchen ihre Masse erhalten, der elektroschwachen Symmetriebrechung.

An Hadronenbeschleunigern wie dem Large Hadron Collider (LHC) kann VBS in Endzuständen mit den Zerfallsprodukten der zwei Botenteilchen zusammen mit zwei Jets studieren. Diese Jets haben eine sehr charakteristische Signatur, welche eine gute Unterdrückung von Untergründen erlaubt und dadurch Studien dieses komplexen Endzustandes trotz des geringen Wechselwirkungsquerschnitts ermöglicht.

Die ersten Hinweise auf die Streuung von Eichbosonen wurde basierend auf den Daten von Run 1 des LHC von der ATLAS-Kollaboration in dem $W^\pm W^\pm \rightarrow W^\pm W^\pm$ -Kanal im leptonischen Endzustand gefunden. Die CMS-Kollaboration hat die erste Beobachtung von VBS mit einer Signifikanz von 5,5 Standardabweichungen im gleichen Kanal in den Daten von 2015 und 2016 von Run 2 publiziert [1]. Diese Beobachtung wurde später von der ATLAS-Kollaboration in [2] mit Beiträgen des Autors, z.B. zur Modellierung des $W^\pm Z$ Untergrundprozesses und den zugehörigen Unsicherheiten bestätigt.

Der zweite Endzustand, für den VBS beobachtet wurde, ist die $W^\pm Z/\gamma \rightarrow W^\pm Z$ -Sreuung im leptonischen Endzustand. Diese Beobachtung wurde in [3] von der ATLAS-Kollaboration mit signifikanten Beiträgen des Autors erreicht. Der genutzte Datensatz wurde mit dem ATLAS-Detektor bei einer Schwerpunktsenergie von $\sqrt{s} = 13$ TeV in 2015 und 2016 in Run 2 des LHC aufgenommen und erreicht eine integrierte Luminosität von $36,1 \text{ fb}^{-1}$. In dieser Studie wurde der Datensatz mit dem gleichen allgemeinen Ansatz, aber mit Verbesserungen in verschiedenen Schlüsselaspekten erneut analysiert.

Ein umfassender Überblick über die verfügbaren Setups für Simulationen des Signalprozesses wird präsentiert. Zuvor gefundene Modellierungsprobleme der Partonschauersimulationen von SHERPA und früheren Versionen von PYTHIA 8 werden bestätigt. Die höchsten Ordnungen der Störungstheorie in der Matrixelementrechnung, die erreicht wurden, waren in führender Ordnung (LO) für den vollen VBS-Prozess und in zweiter Ordnung (NLO) bei Anwendung der VBF-Näherung. Für folgende Studien erscheint eine LO-Rechnung des vollen Prozesses inklusive eines zusätzlich abgestrahlten Partons mit Anpassung an die Partonschauersimulation am vielversprechendsten, bis volle NLO-Rechnungen für VBS in allen Kanälen verfügbar sind.

Besondere Bedeutung wird auf die Modellierung von Untergründen, vor allem der $W^\pm Z$ -Zweibosonen-Produktion mit zusätzlichen QCD-Emissionen und dem experimentellen Untergrund durch fehlidentifizierte Leptonen. Ein datengetriebener Ansatz wird angewendet und detailliert studiert, um verlässliche Vorhersagen des experimentellen Untergrundes zu erhalten. Die Vorhersage kann zum Beispiel durch zusätzliche Korrekturen signifikant verbessert werden. Diese werden durch dedizierte Tests der internen Konsistenz des Verfahrens mittels Simulationen gefunden.

Algorithmen des Maschinennlernens in Form von verstärkten Entscheidungsbäumen (BDT) werden trainiert und optimiert, um die Unterscheidung vom Signal und Untergrundprozessen zu verbessern.

Zusätzliche Hinweise für den Signalprozess mit einer Signifikanz von 3.44 Standardabweichungen werden mittels der Profile-Likelihood-Methode in einem gebinnten Maximum-Likelihood-Anpassung erreicht. Ein Wechselwirkungsquerschnitts von

$$\sigma_{\text{obs}} = 1,41 \pm_{0,40}^{0,46} (\text{stat}) \pm_{0,28}^{0,38} (\text{theo}) \pm_{0,13}^{0,13} (\text{sys}) \text{ fb}$$

wird gemessen, was in sehr guter Übereinstimmung mit SM-Vorhersagen in führender Ordnung Störungstheorie von $\sigma_{\text{pred}} = 1,33 \pm_{0,14}^{0,15} \text{ fb}$ ist.

Abstract

The Standard Model (SM) is the fundamental theory describing elementary particles and their main interactions at typical energy scales at collider experiments, the electromagnetic, the weak, and the strong interactions. The more complex underlying structure describing the weak and the strong interactions in the SM compared to the electromagnetic interaction necessitates direct three-point and four-point interactions among the mediators of the weak and strong interactions, called gauge bosons. Such self-interactions do not exist for the gauge boson of the electromagnetic interaction, the photon. While the three-point interaction was studied in detail in earlier collider experiments, the four-point interaction is a fundamental prediction of the SM, which was not observed for the weak interaction when starting this study.

One process, where both the three-point as well as the four-point interactions contribute is the scattering of electroweak gauge bosons W^\pm, Z, γ also referred to as vector boson scattering (VBS). In the SM, this scattering is mediated by gauge boson self-interactions, or via the exchange of a Higgs boson. The scattering contributions mediated by a Higgs boson are sensitive to the properties of the Higgs boson and the details of the mechanism in which the W^\pm and Z bosons acquire their masses, called electroweak symmetry breaking.

At hadron colliders such as the Large Hadron Collider (LHC), VBS is observable in a final state with the decay products of two gauge bosons in combination with two jets. These jets have a distinct signature allowing for good suppression of backgrounds and consequently for studies of the complex final state despite the low cross-sections.

The first evidence for a VBS process was presented based on the Run 1 dataset alone by the ATLAS collaboration in the $W^\pm W^\pm \rightarrow W^\pm W^\pm$ channel in the fully leptonic final state. The CMS collaboration published the first observation of VBS in the same channel using data from 2015 and 2016 of Run 2 [1], which was later confirmed by the ATLAS collaboration with contributions by the author, e.g. in the modelling of $W^\pm Z$ background processes and associated uncertainties[2].

The second boson channel for which VBS was observed was the $W^\pm Z/\gamma \rightarrow W^\pm Z$ boson channel in the fully leptonic final state. This observation was published by the ATLAS collaboration with significant contributions by the author in [3]. The studied dataset was collected with the ATLAS detector at a centre-of-mass energy $\sqrt{s} = 13$ TeV during 2015 and 2016 of Run 2 of the LHC and amounts to an integrated luminosity of 36.1 fb^{-1} . In this study, the dataset was re-analysed following the same overall approach but with improvements in several key aspects.

A comprehensive overview of available setups for reliable simulations of the signal process is presented. In a modelling study of the available setups, modelling issues in the parton shower simulation of SHERPA and earlier versions of PYTHIA 8 observed in earlier studies are confirmed. The best matrix-element accuracies in available setups are leading-order for the full VBS signal process and next-to-leading-order in the VBF approximation. For upcoming analyses, a leading-order calculation of the full process including an additional QCD emission merged with parton shower simulations is found to be most promising, before full next-to-leading order calculations become available for all boson channels in VBS.

Additional emphasis is set on the modelling of backgrounds, mainly $W^\pm Z$ diboson production in association with additional QCD emissions as well as the experimental background due to misidentified leptons. A data-driven approach is applied and studied in detail for a reliable estimate of the latter background. Significant improvements to the estimate, e.g. in the form of additional corrections, are found via dedicated tests of the self-consistency of the approach using simulations.

Machine-learning algorithms in the form of Boosted-Decision-Trees (BDT) are trained and optimized for improved separation of the background and signal processes. Evidence for the signal process is found with a significance of 3.44σ using the profile likelihood method in a binned maximum-likelihood fit. The fiducial cross-section is measured to be $\sigma_{\text{obs}} = 1.41 \pm_{0.40}^{0.46} (\text{stat}) \pm_{0.28}^{0.38} (\text{theo}) \pm_{0.13}^{0.13} (\text{sys}) \text{ fb}$, which is in good agreement with the leading-order SM prediction of $\sigma_{\text{pred}} = 1.33 \pm_{0.14}^{0.15} \text{ fb}$.

Contents

1	Introduction	1
2	Theoretical Framework	3
2.1	The Standard Model of Particle Physics	3
2.1.1	Overview	3
2.1.2	Primer to gauge theories	4
2.1.3	Building the Standard Model Lagrangian	7
2.1.4	Electroweak theory	10
2.1.5	Quantum chromo dynamics	11
2.1.6	Interactions of bosons	12
2.1.7	Introducing masses	13
2.1.8	Predictions and validations	17
2.1.9	Beyond the Standard Model	18
2.2	Scattering of Electroweak Gauge Bosons	20
2.2.1	Vector Boson Scattering at the Hadron Colliders.	22
2.2.2	Comparison of final states	24
3	Simulations and Modelling Studies	29
3.1	Methodology	29
3.1.1	Phases of event generation	29
3.1.2	Matching and merging	33
3.2	Overview of Monte-Carlo Generators	35
3.2.1	Sherpa	35
3.2.2	MadGraph5_aMC@NLO	35
3.2.3	Herwig 7	36
3.2.4	Pythia 8	36
3.2.5	Other event generators and tools	37
3.3	Problems and Caveats	37
3.3.1	Issues towards higher accuracy	39
3.4	Common Approximations	40
3.5	Process-specific Setups	43
3.5.1	$WZjj$ -EW6	43
3.5.2	$WZjj$ -EW4	46
3.5.3	$WZjj$ -EW5	47
3.6	Overview of Simulations	48
3.7	Modelling Studies	48
3.7.1	Phase space definition	48
3.7.2	Modelling of $WZjj$ -EW6	51
3.7.3	Modelling of $WZjj$ -EW4	56
3.7.4	Process comparison	57

4	Experiment	61
4.1	The Large Hadron Collider	61
4.2	The ATLAS Detector	65
4.2.1	General design	65
4.2.2	Coordinate system and common observables	66
4.2.3	Inner detector	68
4.2.4	Calorimeters	69
4.2.5	Muon system	71
4.2.6	Luminosity measurement	72
4.2.7	Trigger	72
4.3	Reconstruction of Analysis Objects	73
4.3.1	Electrons	74
4.3.2	Muons	75
4.3.3	Jets	77
4.3.4	Lepton isolation	80
4.3.5	Missing transverse energy	81
4.4	Processing Chain for ATLAS Data	82
4.4.1	ATLAS data model	82
4.4.2	Analysis workflow	83
5	Object and Event Selection	85
5.1	Preselection	85
5.2	Object Selection	86
5.2.1	Electrons	86
5.2.2	Muons	88
5.2.3	Jets	89
5.2.4	Overlap removal	90
5.2.5	Missing transverse momentum	90
5.3	Event Selection	90
6	Background Estimation	97
6.1	Prompt Backgrounds	98
6.2	Data-driven Non-prompt Estimation	101
6.2.1	The fake factor method	102
6.2.2	Fake factor determination	105
6.2.3	Closure tests	113
6.2.4	Results of matrix method	124
7	Multi-variate Event Classification	127
7.1	Basic Principles of Boosted Decision Trees	128
7.1.1	Evaluation of performance	131
7.2	Selection and Preprocessing of Input Variables	132
7.2.1	Modelling of input observables	134
7.2.2	Preprocessing	137
7.3	Optimization	138
7.3.1	Set of input observables	140
7.3.2	Meta variable optimization	142
7.3.3	Combination of optimizations	142
7.4	Evaluation at Truth-level	146
7.5	Conclusions	146

8	Uncertainties	151
8.1	Experimental Uncertainties	152
8.2	Modelling Uncertainties	153
8.2.1	PDF uncertainties	155
8.2.2	Strong coupling	156
8.2.3	Scale uncertainties	156
8.2.4	Parton shower and hadronisation	157
8.2.5	Merging	159
8.2.6	Missing contributions	160
8.2.7	Result for predicted fiducial cross-section	162
8.2.8	Caveats for uncertainties from two-point variations	162
9	Cross-section Measurement	163
9.1	Statistical Methodology	163
9.1.1	Likelihood construction	164
9.1.2	Profile likelihood ratio method	165
9.2	Fitting Setup	168
9.2.1	Technical implementation	168
9.2.2	Input processes	168
9.2.3	Input distributions	169
9.2.4	Input preparation	171
9.3	Expected Fit Results	172
9.3.1	Validation study using toy samples	176
9.4	Observed Fit Results	182
9.5	Interpretation	188
10	Conclusions & Outlook	191
A	Conventions and Notations	195
A.1	Natural Units	195
A.2	Einstein's Summation Convention	195
A.3	Relativistic and Dirac Notation	195
A.4	Group Theory	196
B	Auxiliary Information on Event Generation	197
B.1	Configuration Files and Notes on Private Productions	197
B.1.1	EW6 MG5Py8Fix	197
B.1.2	EW6 MG5H7	200
B.1.3	EW6 WhPy8	200
B.1.4	EW6 VBFNLOPy8	202
B.1.5	EW6 H7-VBFNLO LO	207
B.1.6	EW6 H7-VBFNLO NLO	209
B.1.7	EW5 MG5Py8	212
B.1.8	EW4 MG5Py8	214
B.2	Simulations at Reconstruction Level	215
C	Details of Typical Comparison Plots	221
D	Calibration of Forward Electrons	223
D.1	Event selection and simulation samples	223
D.2	Formalism and methodology	224
D.2.1	Definition of inter-calibration regions	224
D.2.2	Tag-and-probe approach	225
D.2.3	Fitting procedure	226
D.3	Systematic effects and results	228
D.3.1	Systematic uncertainties	228
D.3.2	Results	229

D.4	Validation	229
E	Data-driven Non-prompt Estimate	233
E.1	Object Selection	233
E.2	Monte-Carlo simulation based Closure Test	236
E.3	Further details on matrix-method results	236
F	BDT Input Modelling Validation	243
F.1	Comparison of Predictions for Signal Process for all Input Candidates	243
F.2	Truth-level comparison BDT inputs	250
F.3	Validation of Modelling of BDT inputs	253
F.3.1	WZ Control Region	253
F.3.2	Signal Region	262
G	Uncertainties	271
G.1	Modelling Uncertainties in Control Regions	271
G.1.1	PDF uncertainties	271
G.1.2	Strong coupling	273
G.1.3	Scale uncertainties	275
G.1.4	Parton Shower and Hadronization	277
H	Additional Fit Results	279
H.1	Fit to Asimov Dataset	279
H.1.1	Full list of values for nuisance parameters	279
H.1.2	Detailed post-fit event yields	284
H.2	Fit to Observed Dataset	287
H.2.1	Full list of values for nuisance parameters	287
H.2.2	Correlation between nuisance parameters	292
H.2.3	Full ranking of nuisance parameters	294
H.2.4	Detailed post-fit event yields	299

Chapter 1

Introduction

The goal of elementary particle physics is to understand the fundamental building blocks of matter and their interactions. Especially over the last century, there has been immense progress in identifying and studying these building blocks, called elementary particles. In the pursuit of this goal, a comprehensive theory was developed, describing all known elementary particles and their primary interactions at typical energies of particle colliders.

This theory, known as the Standard Model of particle physics (SM), was refined and extended over the years and several predictions of the SM were confirmed. The major predictions of the SM were the predictions of new particles, e.g. the W^\pm and Z bosons, the top quark, and the Higgs boson. All of these predictions have been confirmed, and the predicted particles have been observed experimentally.

The experiments grew in complexity and cost over the years. Many experiments of early particle physics can easily be recreated in lab courses and schools. Nowadays, the research in particle physics is driven by large collider experiments, run by up to thousands of technicians and physicists.

The combination of enormous efforts on the experimental and theoretical side, stimulating each other has led to enormous progress in the field over the last decades.

The SM without a Higgs boson predicts unphysical behaviour and is thus inconsistent at energies in the order of 1 TeV. The largest particle collider in the world, the Large Hadron Collider (LHC) was build to reach such energies. Thus, it was widely expected to find indications for a Higgs boson or new physics beyond the SM. In 2012, the ATLAS and CMS collaborations observed a Higgs-like boson.

In recent years the properties of the observed particle have been studied in more detail and are well compatible with predictions of the SM. With this Higgs boson, the SM is a consistent theory in the realm of collider experiments at energies far beyond the LHC and proposed future colliders.

Nevertheless, some fundamental questions cannot be answered within the SM as we know it. It does not incorporate gravitational effects and cannot explain some phenomena at cosmic scales like the amount of matter-antimatter asymmetry or Dark Matter, making the search of new physics beyond the Standard Model more relevant now.

Many physicists hoped to find indications for physics beyond the SM in the form of new particles or other deviations from the SM in the data collected at the LHC. However, in the currently analysed dataset, no evidence for such phenomena was found. Instead, measurements are in good agreement with the SM. In the absence of direct detections of new physics particles, predictions of the SM have to be studied more precisely turning SM measurements into indirect probes of new physics.

A fundamental prediction of the underlying symmetry structure of the SM is the existence of three-point and four-point self-interaction vertices among gauge bosons of the weak and the strong interaction. Such self-interaction does not exist for the electromagnetic interaction. The gauge boson of electromagnetism, the photon does not couple to itself. For the gluon, the gauge boson of the strong interaction, as well as the W^\pm and Z bosons, the gauge bosons of the weak interaction, the triple vertex has been studied at earlier collider experiments. Until

recently, no dedicated tests of the quartic coupling among gauge bosons were performed.

The most promising process with contributions of the quartic gauge coupling is the scattering of electroweak gauge bosons. In this process, two gauge bosons can interact via the triple and quartic gauge couplings. Studies of this process are, therefore, the first test of a fundamental prediction of the gauge structure of the SM. Since the Higgs boson can also mediate the scattering, studies of this scattering can also provide insight into the underlying theory of the Higgs boson and are sensitive to deviations from the SM predictions in this sector.

The scattering of longitudinally polarised massive gauge bosons is of particular interest. These polarisation states are intertwined with the mechanism of electroweak symmetry breaking, the BEH mechanism. Through this mechanism, the gauge bosons gain their mass, without which longitudinal polarisation states do not exist. In the SM, these states can be identified with the Goldstone bosons introduced in the BEH-mechanism. Primarily, at high energies, these states behave drastically different from the transversally polarised states.

The aforementioned unphysical predictions of the SM without the Higgs boson occur in this scattering process of purely longitudinally polarised bosons, contributing to the design decisions for the LHC. In this theory, the probability for this process increases with energy and predicts scattering probabilities larger than one, thus violating unitarity. In the SM this violation is cancelled by the contributions of the Higgs boson. If the properties of the Higgs boson observed in 2012 deviate from the SM predictions, the Higgs could only partially restore unitarity leading to drastic differences in the high-energy behaviour of the scattering. In a full theory, the unitarity then needs to be restored by additional more massive particles.

While the existence is commonly accepted and presumed, experimental observations of the scattering were not available when starting this work. The CMS observation reported the first-ever observation of scattering of electroweak gauge boson in the $W^\pm W^\pm \rightarrow W^\pm W^\pm$ channel [1]. The first observation in the $W^\pm Z/\gamma \rightarrow W^\pm Z$ channel performed in the fully leptonic final state was reported by the ATLAS collaboration in [3]. This publication was performed with significant contributions by the author and built the foundation for this work. For this work, the dataset was re-analysed with improvements in the signal and background modelling and changes to the object and event selection to improve consistency with other analyses of the vector boson scattering process.

This work is structured as follows. Chapter 2 provides an overview of the theoretical framework and describes the signal process of scattering of gauge bosons in more detail. An overview of the procedure and tests of the available setups for simulations with a focus on the studied final state are presented in Chapter 3. Chapter 4 introduces the LHC and the ATLAS detector as well as the reconstruction of physics objects from the recorded data. The signal region is defined via the object and event selection presented in Chapter 5. The modelling of background processes, in particular the data-driven estimation of non-prompt backgrounds, is improved and validated in Chapter 6. Multi-variate techniques are derived and optimised as described in Chapter 7 in order to improve the separation of the signal process from background processes. Using uncertainty estimates derived in Chapter 8, the signal process is studied using statistical techniques in Chapter 9. Chapter 10 concludes this study and provides an outlook to future analyses.

Chapter 2

Theoretical Framework

2.1 The Standard Model of Particle Physics

The goal of elementary particle physics is to describe the fundamental building blocks of nature and their interactions. The Standard Model of particle physics (SM) is the widely-accepted theory to achieve that. There are many excellent introductions to the Standard Model available [4–6]. It is not the goal of this work to try to add to this list. This section summarizes and follows the outline and notation of one of these existing introductions, considered excellent by the author [4]. For further details and calculations, the reader is highly recommended to inspect [4].

The notation is based on the usual relativistic notation, Einstein’s summation convention, and on natural units. All three conventions are used throughout this work and are described in more detail in Appendix A. A comprehensive overview of different notations commonly used to describe the SM Lagrangian and the resulting Feynman rules is given in [7].

2.1.1 Overview

Over the last few decades, scientists have found new particles which were considered at the time to be elementary, i.e. not to have a substructure. Technological advancements and continuous effort lead to a large number of discoveries of particles and physicists tried to find an underlying ordering system for a long time. In the sixties and seventies of the previous century, theorists caught up and formulated the theory that evolved to the SM as we know it today.

The SM consists of a description of interactions and elementary particles grouped according to their interactions described in terms of a single formula, called Lagrangian.

The SM is a theory of symmetries, that is the invariance under a specific set of transformations. These symmetries are used to derive underlying properties of interactions as well as a set of conserved quantities, called quantum numbers. Some of the conserved quantities are the charges, which dictate which interactions a particle participates in and how strong the interaction is. The charges are also used to identify and order particles. The currently known elementary particles can be arranged in groups of multiplets without obvious gaps for unknown particles.

There are four known fundamental interactions; gravity, the electromagnetic interaction, weak interaction, and the strong interaction.

Gravity is best described by the theory of general relativity, which is not part of the Standard Model. Formulations of gravity as relativistic quantum field theories and a possible unification with the other interactions are an active field of research, but no formulation is widely accepted. Since the effect of gravity is negligible when compared to the other interactions at typical energy scales, gravity is typically not considered when studying elementary particles in collider experiments.

Electromagnetism is a unified theory combining electricity and magnetism and was developed in the 19th century. The quantum field theory that describes electromagnetism is

call *quantum electro-dynamics (QED)*. Due to the readily observable effects also on macroscopic scales, electromagnetism was much easier to study compared to the weak or strong interactions.

The weak interaction was proposed only in 1933 to explain β -decay energy spectra. Electromagnetism and the weak interaction can be described in a unified theory called the electroweak theory.

The existence of a new force was proposed at the beginning of the 20th century to explain the formation and stability of nuclei. This new force is now known as the strong interaction. For nuclei to be stable, this additional contracting interaction has to counteract the electromagnetic repulsion caused by the positively-charged protons. The quantum field theory that describes the strong interaction is called *quantum chromo-dynamics (QCD)*. The strong interaction is now known to cause also the formation of hadrons from their fundamental building blocks. These hadrons are always neutral in the charge of the strong interaction, called colour charge. In a typical representation for quarks, there are three different colour charges (red, green, blue) and their anti-colours. Due to their substructure, hadrons can still interact via the strong interaction, e.g. to form nuclei, but only on very short distances in the order of ≈ 1 fm.

Gauge particles are the mediators of interactions. Charged particles emit such gauge particles, which travel through space-time and interact with other charged particles. In the SM, all of these gauge particles have a spin of 1 and are thus often called vector bosons. The known vector bosons are the photon for electromagnetic interaction, W^\pm and Z bosons for the weak interaction and the gluons G_a for the strong interaction. In total, there are eight gluons indicated by the index a with different colour combinations. Since all other properties are equivalent, they are usually not considered independently.

There is one additional boson which is considered to be elementary, called the Higgs boson. This boson has spin 0 and is not a gauge boson but arises from the mechanism that introduces masses (see Section 2.1.7).

In addition to these bosons, there are several spin- $\frac{1}{2}$ elementary particles, called fermions. These are the fundamental building blocks of matter and are identified by their charges.

All fermions in the SM participate in the weak interaction. Colour-neutral fermions, i.e. fermions not participating in strong interactions, are called *leptons*, while fermions with a non-zero colour-charge are called *quarks*. All quarks participate in the electromagnetic and weak interaction. Quarks can be grouped according to the absolute values of their electroweak charges $|Q|$ in up-type, with $|Q| = 2/3$ and down-type quarks with $|Q| = 1/3$. Leptons are as well further subdivided according to their electromagnetic charge. Leptons without an electromagnetic charge are called *neutrinos*.

In total, there are four differently charged states, not counting the different colour states for quarks; neutrino, charged lepton, up-type quark, and down-type quark. Each of these states occurs in three generations, differing only in their masses. An overview of the elementary particles in the Standard Model is given in Table 2.1. An antiparticle exists for each elementary particle with opposite charges and the same mass.

2.1.2 Primer to gauge theories

The SM is a gauge theory in the framework of relativistic quantum field theory. In relativistic quantum field theories, a theory is usually identified by its Lagrangian. The Euler-Lagrange equations provide a recipe on how to derive the equations of motion for a system with known Lagrangian. The Lagrangian has to be scalar in every relevant space, invariant under Lorentz transformations and local gauge transformations to yield consistent predictions[4].

These symmetries, especially for gauge transformations, are of great importance following Noether's theorem [10]. This theorem states that there is an associated conserved quantity for every continuous symmetry and vice versa. The theorem also describes how to derive the charge from a Lagrangian describing the system. By formulating symmetries that are observed in nature in terms of invariance under local gauge transformations, associated charges and properties of the interactions between particles with these charges can be derived.

Table 2.1: Overview of elementary particles in the Standard Model. Each fermion (spin = $\frac{1}{2}$) has an associated anti-particle, which have opposing charges and are not listed here. For each particle, the observed mass m and the electromagnetic charge Q are shown. The neutrino interaction eigenstates ν_e , ν_μ , and ν_τ mix to the mass-eigenstates ν_1 , ν_2 , and ν_3 . At least two mass eigenstates must have a non-zero mass to be consistent with the observation of neutrino oscillations. The listed upper limits for neutrino masses are derived from direct measurements [8]. The listed quarks are triplets of quarks, with one state for each colour charge. For details of the quark mass estimation procedures, see [9]. Masses of all particles other than neutrinos from [9].

	Name		Spin	Mass m	Electromagn. charge Q [e]
1 st Generation	ν_e	e -neutrino	1/2	< 1.1 eV	0
	e	electron	1/2	0.511 MeV	-1
	u	up quark	1/2	$2.16^{+0.49}_{-0.26}$ MeV	2/3
	d	down quark	1/2	$4.67^{+0.48}_{-0.17}$ MeV	-1/3
2 nd Generation	ν_μ	μ -neutrino	1/2	< 1.1 eV	0
	μ	muon	1/2	0.1057 GeV	-1
	c	charm quark	1/2	1.27 ± 0.02 GeV	2/3
	s	strange quark	1/2	93^{+11}_{-5} MeV	-1/3
3 rd Generation	ν_τ	τ -neutrino	1/2	< 1.1 eV	0
	τ	tauon	1/2	1.777 GeV	-1
	t	top quark	1/2	172.9 ± 0.4 GeV	2/3
	b	bottom quark	1/2	$4.18^{+0.03}_{-0.02}$ GeV	-1/3
Gauge bosons	γ, A	photon	1	0	0
	Z	Z boson	1	91.1876 ± 0.0021 GeV	0
	W^\pm	W^\pm boson	1	80.379 ± 0.012 GeV	± 1
	g	gluon	1	0	0
	H	Higgs boson	0	125.10 ± 0.14 GeV	0

In this framework, the theory should be invariant under certain gauge transformations, even if they are dependent on the point in space-time. A good example for such a transformation is the phase of a wave function $|\psi|^2$ (see [4]). One is allowed to pick a phase of choice independent of the choice at other places or other times. Such transformations are called local gauge transformations, and they can be written as

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

Here $\chi(\vec{x}, t)$ is a scalar field describing the phase changes in space-time. When testing the invariance of Schrödingers equation under such transformations one always needs to allow for the presence of an electromagnetic field via an vector potential \vec{A} and a scalar potential V and modify Schrödingers equation to be

$$\frac{1}{2m} \left(-i\nabla + e\vec{A} \right)^2 \psi = (i\partial/\partial t + eV) \psi, \quad (2.2)$$

where e is the electric charge of the electron. This adjusted equation is then invariant under the simultaneous transformations

$$\psi(\vec{x}, t) \rightarrow \psi'(\vec{x}, t) = e^{i\chi(\vec{x}, t)} \psi(\vec{x}, t) \quad (2.3)$$

$$\vec{A}(\vec{x}, t) \rightarrow \vec{A}'(\vec{x}, t) = \vec{A}(\vec{x}, t) + \frac{1}{e} \nabla \chi \quad (2.4)$$

$$V(\vec{x}, t) \rightarrow V'(\vec{x}, t) = V(\vec{x}, t) - \frac{1}{e} \frac{\partial}{\partial t} \chi(\vec{x}, t). \quad (2.5)$$

The notation can be simplified by defining the *covariant derivative* specific to a local gauge transformation

$$\vec{D} = -\nabla - ie\vec{A} \quad D^0 = \partial/\partial t - ieV. \quad (2.6)$$

Replacing the usual derivative with the covariant derivative of a certain gauge transformation in a candidate theory, makes the kinematic terms in the theory automatically gauge invariant under this specific gauge transformation.

We see that local gauge invariance dictates the presence of a vector field ($V; \vec{A}$) (or in relativistic notation A^μ) associated with a spin-1 particle. In this example, this particle can be interpreted as the photon. The presence of the photon can be derived like this by choosing the correct local gauge transformation. In order to describe the dynamics of the new gauge field A^μ a new term has to be added to the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \quad (2.7)$$

where the field strength tensor $F^{\mu\nu}$ is defined as

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \quad (2.8)$$

No underlying motivation for the particular set of local gauge transformations in the SM was found yet. Three groups of local gauge transformations for which the SM Lagrangian is invariant are known. These are discussed in detail below.

A procedure to find meaningful local gauge transformations is based on postulating new internal spaces. The idea is to group particles with similar properties concerning an observed interaction in multiplets. Such a multiplet could be postulated, e.g. for pairs of quarks, which have similar masses and behave similarly under the weak interaction. The two quarks types are thus grouped presumably in a spin-like doublet, where one represents the spin-up and the other the spin-down state. Due to the spin-like structure, these multiplets are typically called isospin multiplets. Postulating invariance under rotations in this internal isospin space leads to a candidate theory for the weak interaction.

The transformations describing such rotations generate a $SU(2)$ in group theory and can be written as

$$\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} \rightarrow \begin{pmatrix} q'_1 \\ q'_2 \end{pmatrix} = e^{i\vec{\epsilon}(\vec{x}, t) \cdot \vec{\tau}/2} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}, \quad (2.9)$$

where $\vec{\tau}$ represents the Pauli spin matrices (see Appendix A) and $\vec{\varepsilon}$ specifies the rotation in the isospin space. It is important to note that the order of rotations matters, i.e. the spin matrices do not commute:

$$[\tau_i, \tau_j] = 2i\epsilon_{ijk}\tau_k, \quad (2.10)$$

with the Levi-Civita symbol ϵ_{ijk} . Transformations where the order of application affects the result are called non-Abelian transformations. This has fundamental consequences for this work as discussed below (see Section 2.2).

Similar to the example of electrodynamics following Equation 2.1, a field W_i has to be introduced for each τ_i . It is again convenient to introduce a covariant derivative in an equivalent way.

The two examples above of a $U(1)$ and a $SU(2)$ symmetry have been confirmed, and the underlying theory in nature seems to be invariant under such transformations.¹

The third group of transformations that matches observations is a $SU(3)$ symmetry in an internal space for the strong interaction, called colour space. Similar to the $SU(2)$ case, such transformations can be expressed in terms of infinitesimal rotations, now using the eight 3×3 λ_i matrices rather than three 2×2 τ_i matrices (see Appendix A). As the interpretation as rotations implies, such transformations are non-Abelian similar to $SU(2)$. Once more, new fields have to be introduced to ensure gauge invariance under the appropriate transformations. For the $SU(3)$ with eight generators λ_i , eight fields G_i^μ are introduced.

A new covariant derivative D^μ can be introduced combining all three transformations

$$D^\mu = \partial^\mu - ig_1 \frac{Y}{2} B^\mu - ig_2 \frac{\tau_i}{2} W_i^\mu - ig_3 \frac{\lambda_a}{2} G_a^\mu, \quad (2.11)$$

where g_1 , g_2 , and g_3 are for the moment arbitrary real numbers representing the coupling strengths. The notation for the $U(1)$ transformations has been adjusted so the identification as electromagnetism isn't introduced ad-hoc. Here the hypercharge generator Y is introduced, which multiplies with a state-dependent scalar. It should be noted that each term is a singlet in the space of the other transformations. For example the last term is a 3×3 matrix in $SU(3)$ colour space, but a singlet in the $SU(2)$ isospin space.

When this covariant derivative is applied in a Lagrangian that describes the dynamics of fermion fields, the additional gauge bosons and interactions are automatically introduced to the Lagrangian in a gauge-invariant way. In other words, if one knows which particles to consider, almost all² of their interactions can be derived from that with few free parameters.

2.1.3 Building the Standard Model Lagrangian

The fundamental building blocks of matter are massive spin- $\frac{1}{2}$ fermions. The wave equation which describes the free relativistic kinematics of a fermionic field is the Dirac equation

$$(i\gamma^\mu \partial_\mu - m) \psi = 0, \quad (2.12)$$

where γ^μ are the γ matrices as defined in Appendix A³ and ψ the spinor representation of the fermionic field. This equation can be derived using the Euler-Lagrange equations from the Lagrangian

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi} (i\gamma^\mu \partial_\mu - m) \psi. \quad (2.13)$$

The chirality is introduced using the projection operators

$$P_L = \frac{1 - \gamma^5}{2}, \quad P_R = \frac{1 + \gamma^5}{2} \quad (2.14)$$

¹However, the symmetry must be broken spontaneously in the ground state to be consistent with massive elementary particles.

²We will see that we have to add other terms to the Lagrangian to explain masses. These will lead to additional interaction terms.

³The representation chosen here corresponds to [4] and slightly deviated from the usual chiral basis, e.g. in the sign of γ^5 .

where γ^5 is defined as

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3. \quad (2.15)$$

Each field ψ can easily be split in its left-handed and right-handed projections ψ_L and ψ_R

$$\psi = \frac{1 - \gamma^5}{2}\psi + \frac{1 + \gamma^5}{2}\psi = \psi_L + \psi_R. \quad (2.16)$$

In the chosen representation for γ^μ , it can be seen that

$$\gamma^5 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad (2.17)$$

such that the first (last) two components of ψ are identified as ψ_R (ψ_L).

The negative energy solutions from Equation (2.12) are interpreted as positive energy solutions of anti-particles. We can introduce the notation

$$\bar{\psi} = \psi^\dagger \gamma^0, \quad (2.18)$$

to represent the spinor of the anti-particle corresponding to a particle spinor ψ . The γ^0 switches the left- and right-handed contributions which is necessary since the left-handed anti-particle behaves like the right-handed particle and vice versa.

One expects the left- and right-handed contributions to interact in the same way. This expectation is equivalent to invariance under the transformation of parity, where all spatial coordinates change sign. No violations of this symmetry were observed in electromagnetism and the strong interaction, but the weak interaction was shown to violate this symmetry maximally. Left- and right-handed projections not only behave differently, but only left-handed fields participate in the interaction associated with $SU(2)$ transformations in isospin space.

So when constructing the Lagrangian for the SM, the fermionic fields need to be separated according to their chirality and only the left-handed fields are grouped in $SU(2)$ doublets. The right-handed fields remain singlets in weak isospin space and do not participate in this interaction.

So for the first generation of the leptonic sector, the participating fields can be written as

$$L_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L = \begin{pmatrix} P_L \psi_{\nu_e} \\ P_L \psi_e \end{pmatrix}, \quad (2.19)$$

$$e_R = P_R \psi_e. \quad (2.20)$$

Since right-handed neutrinos also do not participate in electromagnetic or strong interaction, and no experimental proof for their existence has been observed, they are usually not included in the Lagrangian. For quarks, the equivalent fields are introduced as

$$Q_{L\alpha} = \begin{pmatrix} u_\alpha \\ d_\alpha \end{pmatrix}_L = \begin{pmatrix} P_L \psi_{u_\alpha} \\ P_L \psi_{d_\alpha} \end{pmatrix}, \quad (2.21)$$

$$u_{R\alpha} = P_R \psi_{u_\alpha}, \quad (2.22)$$

$$d_{R\alpha} = P_R \psi_{d_\alpha}, \quad (2.23)$$

where the additional index α iterates over the three colour charges, i.e. the indices in the $SU(3)$ triplet.

The notation can be extended to the second and third generation of fermions. Except for masses, no experimental evidence for differences in the theory of each generation has been found so far. In order to indicate the generation an additional index $n = 1, 2, 3$ is added as superscript and summed over.

Table 2.2: Overview of gauge symmetries of full Standard Model Lagrangian and their properties. For each group, the associated gauge field, charge, and generators are listed. The next-to-last column indicates whether the group is abelian, and the last column indicates the possible values for the running index.

Group	Field	Charge	Generator		Index
$U(1)_Y$	B^μ	Hyper charge Y	Y	abelian	
$SU(2)_L$	W_i^μ	Weak isospin I_3	$1/2 \cdot \tau_i$	non-Abelian	i in $\{1, 2, 3\}$
$SU(3)_C$	G_a^μ	Colour	$1/2 \cdot \lambda_a$	non-Abelian	a in $\{1, 2, \dots, 8\}$

Table 2.3: Fermionic charge states in the Standard Model. States that only differ in their colours are grouped. For each group of states, the third component of the weak isospin I_3 , the hyper-charge Y , the electric charge Q , and the colour representations are listed. Since the three generations of fermions only differ by their masses, only the first generation of fermions is shown.

	Name		I_3	Y	Q	Colour representation
Leptons	left-handed	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$1/2$	-1	0	no
			$-1/2$	-1	-1	no
	right-handed	e_R	0	-2	-1	no
Quarks	left-handed	$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$1/2$	$1/3$	$2/3$	triple
			$-1/2$	$1/3$	$-1/3$	triple
	right-handed	u_R	0	$4/3$	$2/3$	triple
		d_R	0	$-2/3$	$-1/3$	triple

From these building blocks, we can derive the fermionic Lagrangian $\mathcal{L}_{\text{ferm}}$ by adding the necessary terms to derive the Dirac equation for a free massless⁴ fermion field

$$\mathcal{L}_{\text{Dirac, massless}} = \bar{\psi} i \gamma^\mu \partial_\mu \psi \quad (2.24)$$

for each of the introduced fields. After replacing ∂_μ with the covariant derivative as introduced in Equation (2.11), the fermionic Lagrangian $\mathcal{L}_{\text{ferm}}$ reads

$$\mathcal{L}_{\text{ferm}} = \sum_{n \in \{1, 2, 3\}} \sum_{\substack{f \in \{L, e_R, \\ Q_L, u_R, d_R\}}} \bar{f} i \gamma^\mu D_\mu f. \quad (2.25)$$

This fermionic Lagrangian $\mathcal{L}_{\text{ferm}}$ has a

$$U(1)_Y \otimes SU(2)_L \otimes SU(3)_C \quad (2.26)$$

symmetry and the covariant derivative introduces new gauge fields and terms which describe the interactions between fermions and gauge fields. An overview of the gauge symmetries is shown in Table 2.2 and of the charge properties of the different fields in Table 2.3. The following sections discuss these new terms independently for the electroweak and the strong interactions.

⁴Masses cause problems due to parity violation when introduced naively. This is discussed in more detail in Section 2.1.7

2.1.4 Electroweak theory

The $U(1)_Y$ symmetry can be derived following the example of electrodynamics and $SU(2)_L$ symmetry in isospin space is motivated via the weak interaction. When focussing on the first generation of leptons, the terms connected to the $U(1)_Y$ are

$$\mathcal{L}_{\text{ferm}}(U(1), 1. \text{ gen. leptons}) = \frac{g_1}{2} [Y_L (\bar{\nu}_L \gamma^\mu \nu_L + \bar{e}_L \gamma^\mu e_L) + Y_R \bar{e}_R \gamma^\mu e_R] B_\mu. \quad (2.27)$$

Y_L and Y_R are introduced to allow for independent hypercharges Y for left- and right-handed fields.

Since purely right-handed fields are singlets in $SU(2)_L$, the associated terms in the Lagrangian are zero. For the left-handed term, it is useful to define charge states for the fields W_μ^i in the following way

$$W_\mu^\pm = (-W_\mu^1 \pm iW_\mu^2)/\sqrt{2}. \quad (2.28)$$

Using these fields, the remaining leptonic interactions can be written as

$$\begin{aligned} \mathcal{L}_{\text{ferm}}(SU(2), 1. \text{ gen. leptons}) = \frac{g_2}{2} \bigg[& \bar{\nu}_L \gamma^\mu \nu_L W_\mu^3 - \sqrt{2} \bar{\nu}_L \gamma^\mu e_L W_\mu^+ \\ & - \sqrt{2} \bar{e}_L \gamma^\mu \nu_L W_\mu^- - \bar{e}_L \gamma^\mu e_L W_\mu^3 \bigg]. \end{aligned} \quad (2.29)$$

Equations (2.27) and (2.29) contain all interactions of leptons in the electroweak sector. This can be compared to knowledge from earlier experiments.

First, the neutrino needs to have an electromagnetic charge of zero. Terms in the leptonic interaction Lagrangian without electron fields are proportional to

$$-\frac{g_1}{2} Y_L B_\mu - \frac{g_2}{2} W_\mu^3. \quad (2.30)$$

The approach to achieve a vanishing electromagnetic charge of the neutrino is to rotate the orthogonal basis of the B_μ and W_μ^3 fields such that Equation (2.30) coincides with one of the rotated fields Z_μ . The other rotated field is by construction orthogonal and consequently does not interact with the neutrino. This other field will be identified as the photon field A_μ . Such rotation is achieved when using the definition

$$A_\mu = \frac{g_2 B_\mu - g_1 Y_L W_\mu^3}{\sqrt{g_2^2 + g_1^2 Y_L^2}}, \quad (2.31)$$

$$Z_\mu = \frac{g_1 Y_L B_\mu + g_2 W_\mu^3}{\sqrt{g_2^2 + g_1^2 Y_L^2}}, \quad (2.32)$$

where the rotation angle θ_w , called Weinberg-angle, can be expressed as

$$\sin \theta_w = \frac{g_1}{\sqrt{g_1^2 + g_2^2}}. \quad (2.33)$$

The leptonic interaction Lagrangian without contributions of the neutrino fields can be written in terms of the proposed definition of the fields A_μ and Z_μ as

$$\begin{aligned} & -A_\mu \left\{ \bar{e}_L \gamma^\mu e_L \left[\frac{g_1 g_2 Y_L}{\sqrt{g_2^2 + g_1^2 Y_L^2}} \right] + \bar{e}_R \gamma^\mu e_R \left[\frac{g_1 g_2 Y_R}{2\sqrt{g_2^2 + g_1^2 Y_L^2}} \right] \right\} \\ & -Z_\mu \left\{ \bar{e}_L \gamma^\mu e_L \left[\frac{g_1^2 Y_L^2 - g_2^2}{2\sqrt{g_2^2 + g_1^2 Y_L^2}} \right] + \bar{e}_R \gamma^\mu e_R \left[\frac{g_1^2 Y_R Y_L}{2\sqrt{g_2^2 + g_1^2 Y_L^2}} \right] \right\}. \end{aligned} \quad (2.34)$$

Since we want to identify the field A_μ as the photon, which couples equally to the left- and right-handed electrons we can fix $Y_R = 2Y_L$. Since Y_L and g_1 only occur in combination,

we can fix $Y_L = -1$ for convenience. When identifying the coupling of the electron and the photon as the electromagnetic coupling e , we can re-write

$$e = \frac{g_1 g_2}{\sqrt{g_1^2 + g_2^2}} \quad (2.35)$$

This approach is consistent with previous experiments and makes predictions about the interactions of the neutrino and electron fields with the new field Z_μ . Also, the rotation angle will affect other processes and thus provides a strong prediction and a principal constraint of the SM.

A generalized relation between the hypercharge Y (associated to $U(1)_Y$), the weak isospin T_3 (associated to $SU(2)_L$), and the electromagnetic charge Q known from electromagnetism is found to be

$$Q = T_3 + \frac{Y}{2}. \quad (2.36)$$

Where the weak isospin is the eigenvalue of the diagonal $SU(2)$ generator T_3 , i.e. the usual z -projection of the spin.

The remaining terms in the leptonic interaction Lagrangian containing neutrino and electron fields are

$$\frac{g_2}{\sqrt{2}} (\bar{\nu}_L \gamma^\mu e_L W_\mu^+ + \bar{e}_L \gamma^\mu \nu_L W_\mu^-) \quad (2.37)$$

and represent a transition between these states in isospin space mediated via the gauge fields W_μ^\pm . This matches the weak interactions that have been seen before the formulation of the theory.

The arguments can easily be translated to quarks by replacing the electron (neutrino) field with the down-quark (up-quark) field and adding the appropriate terms for the right-handed up-quark. From there, one can also easily extend the terms to the second and third generation with the appropriate substitutions.

This theory predicted the interaction of the electron and neutrino mediated by the Z_μ gauge field that was unobserved at the time. The fact that this particle had not been observed and that interactions via the W^\pm gauge bosons were estimated using the couplings to be larger than observed, lead to the assumption that the new gauge boson could be massive. A mass large enough would suppress the interactions and prohibit direct detection in earlier experiments due to limited available energies. The experimental observation of these gauge bosons in 1983 [11–14] was a great confirmation of the electroweak theory.

2.1.5 Quantum chromo dynamics

The remaining fermionic interactions are the strong interactions of the quarks. The theory that describes this is called *quantum chromodynamics (QCD)*. The Lagrangian for these interactions can be seen by focusing on the terms related to $SU(3)_C$ for a quark field q , that is

$$\mathcal{L} \sim \frac{g_3}{2} \bar{q}_\alpha \gamma^\mu \lambda_{\alpha\beta}^a G_\mu^a q_\beta. \quad (2.38)$$

Here, a separation in left- and right-handed fields is not necessary as both contributions interact equally. The indices α and β indicate an index in the colour-triplet. G_μ^a are the eight Gluon fields which couple to the colours as dictated by the eight $SU(3)_C$ generators in colour space. The Gell-Mann matrices (see Appendix A) are generators of the fundamental representation. The components of the $SU(3)$ generators dictate which gluon interacts with which colour states. For each combination of colour states, there is at least one gluon to mediate the interaction, implying that interactions between all colour state combinations exist. There is no direct interaction between different quark flavours, i.e. there is no vertex in QCD where for instance a u quark changes to an d or c quark.

The colour states are labeled red, blue, and green. These names were chosen in analogy to additive mixing in colour theory. The colour states can be combined in different ways to form a colour neutral state. The combination of all colours and their anti-colours is colour-neutral. Additionally, the combinations of red, blue, and green or anti-red, anti-blue, and anti-green are colour-neutral.

The structure of the $SU(3)_C$ group implies directly that the gluons have to be colour charged, which necessitates gluon self-interactions. These self-interactions are a fundamental difference to QED and cause significantly different properties for QCD.

An important difference to QED is the behaviour of the potential, e.g. between a particle and its anti-particle at large distances r . In electromagnetism, the potential converges with increasing distance to a constant value, usually set to zero. However, in QCD, the potential rises linearly with the distance. This behaviour can be motivated when considering the self-interactions of Gluons. This behaviour was experimentally confirmed in bound states of quark-antiquark pairs[15]. When trying to separate two quarks, one has to overcome a constant contracting force. Once the potential energy exceeds the production threshold, a new quark-antiquark pair is created, thus creating two separate pairs of bounded quarks. Typical distances at which these pair productions occur are about 1 fm. Due to the pair-creation, it is not possible to create or observe a single free quark. This effect is called *confinement*. Due to the confinement, only colour-neutral states are observable.

Color-neutral bound states of quarks are called *Hadrons*. The nominal constituent quarks, called valence quarks, determine the main properties of hadrons such as mass, lifetime, or radius. Hadrons are subdivided according to the number of valence quarks. Most of the known hadrons are either mesons with a quark and an anti-quark or baryons with three valence quarks.

Especially when accelerated to high energies, hadrons also contain additional quark-antiquark pairs created from vacuum fluctuations. The exact composition of a hadron can only be determined on a probabilistic basis. This is summarized for collider experiments in the *parton distribution function (PDF)*, which describes the probability that a gluon or quark of specific flavour participates in a given interaction with a given fraction of the hadron's momentum.

The probability for QCD radiation at low energies or small angles diverges. In order to avoid singularities in calculations, only observables that are not sensitive to this radiation of extra particles should be used. Observables like this can be build by clustering individual particles to *jets*, i.e. integrating over all such additional emissions. The clustering algorithm used here is discussed in more detail in Section 3.2.5.

2.1.6 Interactions of bosons

The most fundamental prediction of the SM as a gauge theory, namely the existence of gauge bosons, was experimentally confirmed by observation of the gauge bosons. In order to complete the SM Lagrangian, it is, therefore, necessary to include terms that describe the dynamics of the gauge fields themselves. This can be done analogously to electrodynamics when using generalized field strength tensors

$$B^{\mu\nu} = \partial^\mu B^\nu - \partial^\nu B^\mu \quad \text{for } U(1)_Y \quad (2.39)$$

$$W_i^{\mu\nu} = \partial^\mu W_i^\nu - \partial^\nu W_i^\mu + g_2 \epsilon_{ijk} W_j^\mu W_k^\nu \quad \text{for } SU(2)_L \quad (2.40)$$

$$G_a^{\mu\nu} = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu + g_3 f_{abc} G_b^\mu G_c^\nu \quad \text{for } SU(3)_C \quad (2.41)$$

with $i = 1, 2, 3$ and $a = 1, 2, \dots, 8$ and the structure constants of $SU(2)$ ϵ_{ijk} and of $SU(3)$ f_{abc} . The generalization is necessary to account for the non-Abelian nature of $SU(2)$ and $SU(3)$. Since transformations within these groups can be interpreted as higher-dimensional rotations, this property is equivalent to the fact that the order of rotations in multiple dimensions matters.

The bosonic Lagrangian can be written as

$$\mathcal{L}_{\text{bosons}} = -\frac{1}{4} B^{\mu\nu} B_{\mu\nu} - \frac{1}{4} W_i^{\mu\nu} W_{\mu\nu}^i - \frac{1}{4} G_a^{\mu\nu} G_{\mu\nu}^a. \quad (2.42)$$

It can be seen that the field strength tensors enter quadratically in this Lagrangian. For $B^{\mu\nu}$ this Lagrangian includes only terms quadratic in the field B^μ or its derivative $\partial^\mu B^\nu$, describing the dynamics of the B^μ field. $W_i^{\mu\nu}$ and $G_a^{\mu\nu}$, however, include a term that is

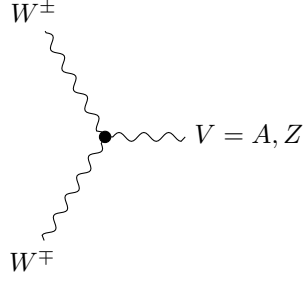


Figure 2.1: Feynman graph for the three-point interaction between the physical gauge fields in the electroweak theory.

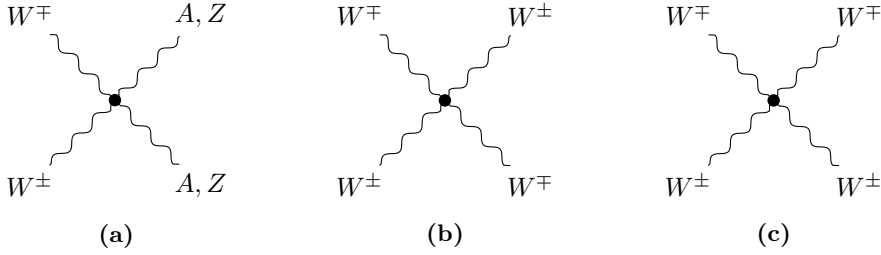


Figure 2.2: Feynman graphs for the four-point interaction between the physical gauge fields in electroweak theory.

already quadratic in the field, so the full Lagrangian also includes terms that are of order three and four in the fields W_i^μ and G_a^μ .

These terms can be identified as three- and four-point vertices of self-interactions among the gauge bosons. Since the individual gluon fields, G_a^μ have identical properties other than the colour charge, they are usually not treated individually, and thus the interactions are not discussed in more detail. These self-interactions and some implications are mentioned in Section 2.1.5.

For the fields W_j^μ , a more thorough discussion is of particular interest for this work. The terms describing the three-point vertex for this field always contain all three components in weak isospin space W_1^μ , W_2^μ , and W_3^μ . W_1^μ and W_2^μ mix to the physical W^\pm gauge bosons (see Equation (2.28)). The third component has contributions in the physical photon and the Z boson. So the resulting three-point interactions described by these terms are interactions between a pair of W^\pm bosons and either a photon or a Z boson as depicted in Figure 2.1. While charge conservation would allow a three-point interaction purely between Z bosons and photons, such an interaction is not part of the Lagrangian and not included in the SM.

The terms for the four-point interaction exist in all combinations where one of the fields W_i^μ does not contribute, and both other fields are included in quadrature. With the mixing to the physical bosons, the four-particle vertices include either four W^\pm bosons or two W^\pm bosons and two additional bosons, where each can be either a photon or a Z boson as shown in Figure 2.2. Similar to the three-point interactions, a vertex with four neutral electroweak gauge bosons does obey charge conservation but does not exist in the SM.

2.1.7 Introducing masses

One of the biggest fundamental problems during the development of the SM was that all particles in the Lagrangian have to be massless. No explicit mass terms are included as those would violate gauge invariance. A mass term for fermions, for instance, has the form

$$m\bar{\psi}\psi = m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R). \quad (2.43)$$

Such terms are not a singlet in weak isospin space and are not invariant under $SU(2)_L$ transformations. For bosons a mass term has the form

$$\frac{1}{2}m^2 B^\mu B_\mu, \quad (2.44)$$

which also adds uncanceled terms when transformed under $SU(2)_L \times U(1)_Y$.

Massless particles are, however, in apparent contradiction to observations. Fermions were known to be massive and also the predicted gauge bosons W^\pm and Z needed to be massive to explain the lack of earlier experimental evidence.

Spontaneous symmetry breaking: A way to overcome this problem is described by the Brout-Englert-Higgs (BEH) mechanism [16–20] (see [21] for a good review). This mechanism proposes a new field with a new potential assigned to it. The potential has its ground state for non-zero values of the fields. The overall potential and thus the overall Lagrangian is gauge invariant, but some symmetries are broken in the ground state. In that ground state, particles can obtain masses via interactions with the field. This concept is called *spontaneous symmetry breaking* and is also used in other fields of physics, e.g. condensed matter physics.

The BEH mechanism introduces two scalar complex fields ϕ^+ and ϕ^0 combined in a $SU(2)_L$ doublet ϕ

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \quad (2.45)$$

The SM Lagrangian is extended by new terms \mathcal{L}_{BEH}

$$\mathcal{L}_{\text{BEH}} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi), \quad (2.46)$$

where the first term describes the kinematic of the new fields and $V(\phi)$ is the new BEH potential introduced as

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \quad (2.47)$$

The full Lagrangian is invariant under the usual gauge transformations. In cases where $\mu^2 < 0$ and $\lambda > 0$, the potential has minima for

$$\phi^\dagger \phi = \frac{-\mu^2}{2\lambda} = \frac{v^2}{2}. \quad (2.48)$$

A choice for the field, for which the potential is minimized can be transformed to all other choices which minimize the potential using gauge transformations. One can freely pick a convenient value, called vacuum state ϕ_0 ,

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad (2.49)$$

and general expansions around this vacuum state can be transformed to

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \quad (2.50)$$

The number of degrees of freedom in the field is reduced from initially four, for two complex fields, to the one, for the expansion around the minimum. The effects of the remaining degrees of freedom can be removed using gauge transformations.

The choice of the minimum breaks some of the $U(1)_Y \otimes SU(2)_L$ symmetries. A symmetry group is found in the expansion around the minimum, which can be identified as the usual phase transformations for QED $U(1)_Q$. The Nambu-Goldstone theorem [22, 23] dictates the existence of a massless scalar boson for each of the broken symmetries, typically referred to as Goldstone bosons. The Goldstone bosons belong to the degrees of freedom lost when choosing a minimum.

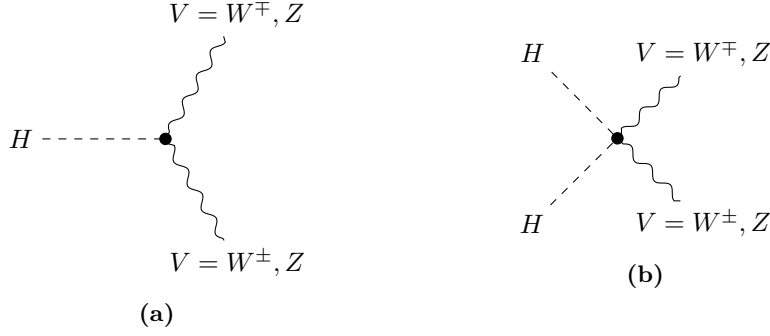


Figure 2.3: Feynman diagrams for the three-point (a) and four-point (b) interactions between the physical gauge fields W^\pm and Z and the Higgs field H .

Boson masses and interactions: When introducing the usual covariant derivative for the electroweak theory in \mathcal{L}_{BEH} and the expansion around the chosen minimum rather than the general Higgs field, mass terms for the gauge bosons can be derived. Requiring the Higgs field H to be electrically neutral and the choice of the $SU(2)_L$ doublet forces the hypercharge of H to be $Y_H = 1$. With this choice, one can derive the terms

$$(D_\mu \phi)^\dagger (D^\mu \phi) = \frac{1}{8} \left| \begin{pmatrix} g_1 B_\mu + g_2 W_\mu^3 & g_2 (W_\mu^1 - iW_\mu^2) \\ g_2 (W_\mu^1 + iW_\mu^2) & g_1 B_\mu - g_2 W_\mu^3 \end{pmatrix} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \right|^2 \quad (2.51)$$

$$= \frac{1}{4} (v + H)^2 g_2^2 W_\mu^+ W^{-\mu} + \frac{1}{8} (g_2^2 + g_1^2) (v + H)^2 Z_\mu Z^\mu. \quad (2.52)$$

A mass term for a boson in the Lagrangian has the form $cM_V^2 VV$, where $c = \frac{1}{2}$ for self-conjugate bosons and $c = 1$ else. For the W^\pm boson the mass term reads

$$M_W = g_2 \frac{v}{2} \quad (2.53)$$

For the Z boson and photon gauge fields, the gauge fields B^μ and W_3^μ have to be transformed using the same mixing as introduced earlier. It can easily be confirmed that the combination in the equation is the exact term used for the pure Z boson. That means the BEH mechanism provides a mass term for the Z boson of

$$M_Z = \frac{1}{2} v \sqrt{g_1^2 + g_2^2}, \quad (2.54)$$

while for the photon no mass term is introduced, thus

$$M_A = 0. \quad (2.55)$$

The pseudo-Goldstone bosons are identified as the longitudinal states of the massive gauge bosons W^\pm and Z . These states do not exist in the SM without the BEH mechanism and are introduced when assigning a mass to the bosons. This connection between the Goldstone bosons and thus the BEH mechanism and the longitudinal states of the electroweak gauge bosons is of particular interest for this work and is discussed further in Section 2.2.

In addition to the mass terms, terms with the structure HVV or $HHVV$ with $V = W^\pm, Z$ are derived. These terms describe three-point and four-point interactions between the Higgs field and gauge boson pairs W^+W^- or ZZ . The coupling strength is proportional to the mass of the gauge boson squared, and the Feynman diagrams for these interactions are shown in Figure 2.3.

Fermion masses: The BEH mechanism, as introduced before, covers the bosonic sector very well and introduces masses for gauge bosons consistent with experiment. For leptons,

additional terms have to be added to the overall Lagrangian. These terms describe interactions between the leptons and the BEH field

$$\mathcal{L}_{\text{Yukawa, leptons}} = y_\ell^i (\bar{L}^i \phi e_R^i + \phi^\dagger e_R^i L^i), \quad (2.56)$$

where $i = 1, 2, 3$ sums over the generations and with arbitrary couplings y_ℓ^i .

This full term is gauge invariant, and in order to understand the physical content of this Lagrangian, we can again replace the general BEH field ϕ with the expansion around the minimum. This yields mass terms for the leptons with

$$m_\ell = y_\ell v / \sqrt{2}. \quad (2.57)$$

The free parameter y_ℓ can now be chosen to add mass terms for each lepton manually. Additionally, an interaction of the lepton with the Higgs boson is included in this Lagrangian with a coupling strength of $\frac{m_\ell}{v}$. No terms for neutrino masses are included by definition.

For quarks, additional terms have to be added to produce mass terms for the state with positive weak isospin

$$\mathcal{L}_{\text{Yukawa, quarks}} = Y_d^{jk} \bar{Q}_L^j \phi d_R^k + Y_u^{jk} \bar{Q}_L^j i \tau_2 \phi^* u_R^k + \text{h.c.}, \quad (2.58)$$

with $j, k = 1, 2, 3$ summing over the generations and where the Yukawa matrices Y_d^{jk} and Y_u^{jk} are complex matrices which can be set for each quark to yield the correct mass terms. The Yukawa matrices are allowed to be non-diagonal, which describes a mixing between the flavour eigenstates of the quarks. The Yukawa matrices can be diagonalized, when the mass eigenstates d of the down-type quarks are transformed to their flavour eigenstate basis d' using the CKM matrix V_{CKM}

$$d'^i = \sum_j V_{\text{CKM}}^{ij} d_j. \quad (2.59)$$

The mixing angles are obtained experimentally, and the mixing is small in the strong sector. Similar mixing was also observed for neutrinos, where the mixing is, however, close the maximal.

This mechanism is less elegant compared to the BEH mechanism as it introduces one free parameter to the theory for each massive fermion. The theory does not predict relations between fermion masses, but essentially provides a way to make fermion masses work in the theoretical framework.

The BEH mechanism adds new free parameters as well, but motivates the origin of the mass of gauge bosons and provides a relation between the masses of the massive gauge bosons at tree level, i.e. without higher-order corrections, of

$$\frac{M_W}{M_Z} = \cos \theta_W. \quad (2.60)$$

A new prediction of the BEH mechanism is the existence of a massive scalar boson, the Higgs boson H . While the mass of the Higgs boson was a free parameter in the BEH mechanism as well, once the mass is known, most interactions can be worked out in this theory.

One of the two free parameters of the BEH potential can be expressed in terms of the vacuum expectation value v of the potential, which is proportional to the ratio of the W^\pm boson mass over the coupling g_2 . This ratio can be expressed in terms of the Fermi constant G_F , which is well known from muon lifetime measurements, so v can be fixed to

$$v = 2 \frac{M_W}{g_2} = \frac{1}{\sqrt{\sqrt{2} G_F}} \approx 246 \text{ GeV}. \quad (2.61)$$

The other free parameter λ of the BEH potential can be related to the mass of the Higgs boson M_H , which is measured to be [9]

$$M_H = 125.10 \pm 0.14 \text{ GeV}. \quad (2.62)$$

Now that we have identified the missing terms to obtain massive particles we can put together the full Standard Model Lagrangian

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{ferm}} + \mathcal{L}_{\text{boson}} + \mathcal{L}_{\text{BEH}} + \mathcal{L}_{\text{Yukawa}}. \quad (2.63)$$

with

$$\mathcal{L}_{\text{ferm}} = \sum_{f \in \{L^j, e_R^j, Q_L^j, u_R^j, d_R^j\}} \bar{f} i \gamma^\mu D_\mu f \quad (2.64)$$

$$\mathcal{L}_{\text{boson}} = -\frac{1}{4} B^{\mu\nu} B_{\mu\nu} - \frac{1}{4} W_i^{\mu\nu} W_{\mu\nu}^i - \frac{1}{4} G_a^{\mu\nu} G_{\mu\nu}^a \quad (2.65)$$

$$\mathcal{L}_{\text{BEH}} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi) \quad (2.66)$$

$$\mathcal{L}_{\text{Yukawa}} = y_\ell^j \bar{L}^j \phi e_R^j + Y_d^{jk} \bar{Q}_L^j \phi d_R^k + Y_u^{jk} \bar{Q}_L^j i \tau_2 \phi^* u_R^k + \text{h.c.} \quad (2.67)$$

where j, k in $\{1, 2, 3\}$ sums over the fermion generations, i in $\{1, 2, 3\}$ the $SU(2)_L$ gauge fields, a in $\{1, 2, \dots, 8\}$ the $SU(3)_C$ gauge fields.

From this Lagrangian \mathcal{L}_{SM} , the Feynman rules can be derived, which describe all interactions and their strengths. The full set of Feynman rules in the SM can be found in [7]. For the sake of simplicity, the sign conventions in the notation of [7] used in this work are

$$\eta = \eta' = \eta_e = -1 \quad (2.68)$$

$$\eta_Z = \eta_\theta = \eta_Y = 1. \quad (2.69)$$

2.1.8 Predictions and validations

The formulation of the SM with the BEH mechanism certainly ranks among the most substantial intellectual achievements of humankind and was awarded multiple Nobel prizes. Among the most important for the SM from those Nobel prizes, Gell-Mann was awarded in 1969 for the introduction of the concept of quarks, 1979 Glashow, Salam, and Weinberg were awarded for the introduction of the electroweak theory, and Higgs and Englert were awarded in 2013 for the BEH mechanism[24–26].

The success of a theory should be graded based on predictions of previously unknown phenomena. Some of such predictions of the SM are based on the ordering in $SU(2)_L$ doublets of quarks. When the electroweak theory was developed, only three quarks were known. The doublets required the existence of an additional quark to form a second doublet. This existence was confirmed with the observation of the J/Ψ meson in 1974[27, 28], which included a previously unknown quark flavour. This observation was a confirmation for the quark model and the Standard Model as a gauge theory. With similar arguments, an additional quark was postulated once the bottom quark as fifth quark was discovered. The top quark observation in 1995 was a further confirmation of the Standard Model.

The most fundamental predictions of the Standard Model as a gauge theory, however, were the existence of three gauge bosons, that were not known at the time. When the existence of the Gluon was experimentally confirmed in 1979 and of the W^\pm and Z bosons in 1983, the Standard Model became widely accepted. The main obstacle in the SM were the masses of the electroweak gauge bosons. The BEH mechanism and the Higgs boson as a consequence were predicted to overcome the fundamental obstacle of the Standard Model, the masses of elementary particles. The search for a Higgs boson was one of the most important goals for particle physics over the last decades. In 2012 the ATLAS and CMS experiments at the LHC announced the observation of a particle consistent with the SM Higgs boson after almost 50 years of search [29, 30]. Since then, the properties of the particle have been investigated in more detail, and no substantial deviations from the prediction for a SM-like Higgs boson have been found.

In addition to the prediction of new particles, relations between seemingly unrelated phenomena can be derived from the Standard Model. For instance, the mass ratio of the W^\pm and Z bosons as well as a forward-backwards asymmetry in $Z \rightarrow e^+e^-$ decays are related

to the Weinberg angle θ_W . These relations allow independent measurements of quantities, for instance, θ_W . These predictions are in very good agreement with experimental results.

Similar to the measurements of theory parameters, also process specific event rates are measured and compared to predictions. The event rate \dot{N} is expressed as

$$\dot{N} = L \cdot \sigma, \quad (2.70)$$

where the dependencies on different parameters are separated into two factors. There L is the instantaneous luminosity, and σ is the cross-section for the process of interest. The luminosity incorporates the experimental factors and is discussed in more detail in Section 4.1.

The cross-section includes the dependencies on the considered process, the types and energies of the initial particles, as well as selection criteria. From the Lagrangian, given a set of values for free SM parameters such as masses and couplings, the cross-section can be calculated.

First, all diagrams that contribute to a given process, i.e. that connect the initial state to a given final state and potentially fulfil additional criteria on the coupling structure have to be found. In theory, the number of diagrams is infinite because of additional internal particles or low-energetic emissions. However, since the couplings are in general smaller than 1, these additional internal particles or emissions are suppressed, and typically only the group of diagrams with the lowest number of couplings are dominant. Calculations, where only diagrams with the lowest number of couplings are considered, are referred to as leading order (LO). If all contributions with an additional coupling order are included, the calculation is considered next-to-leading order (NLO). In general, this nomenclature can be applied to additional QCD couplings (α_S) or other couplings, where in practice α_{EM} and α_W are often combined as α_{EW} . Unless stated otherwise, next-to-leading order (NLO) is only used for additional QCD couplings throughout this work.

From these diagrams, the matrix elements \mathcal{M} can be build using the Feynman rules. The differential cross-section for the volume in phase space $d\Phi$ are proportional to $|\mathcal{M}|^2$

$$d\sigma \propto |\mathcal{M}|^2 d\Phi. \quad (2.71)$$

When integrating over the full phase space defined by the selection criteria of interest, one needs to take into account all possible initial states. For this, the possible partons are summed over, and the possible initial momenta are integrated over.

A good overview of the agreement between experimental observations and theoretical predictions as obtained from simulations implementing this calculation is shown in Figure 2.4. The agreement is very good over a range of 14 orders of magnitude in cross-section for p - p collisions.

2.1.9 Beyond the Standard Model

While widely accepted and confirmed, the SM has internal issues. Among the most substantial issues is the hierarchy problem which arises from quantum corrections to the Higgs mass. Calculations based on the known particles show extensive corrections compared to the observed mass. While it is in principle possible to counteract these corrections by choosing the free parameter of the bare Higgs mass appropriately, this parameter has to be chosen to match the quantum corrections to many orders of magnitude. This need for finely tuned parameters is generally considered a conceptual issue, and it might be caused by additional particles counteracting the quantum corrections more naturally. This issue is one of the primary motivations to develop supersymmetry[32].

In addition to the internal theoretical issues, there are phenomena which the SM cannot explain or which do not agree to predictions within uncertainties. Among the most significant discrepancies are the magnetic moment of the μ and discrepancies in flavour physics, e.g. in the study of b decays[33].

Another apparent problem is the lack of neutrino masses. Since the observation of neutrino oscillations, it is known that at least two neutrinos must be massive. In the current formulation of the SM, neutrino masses are not included. The neutrino sector is an area of

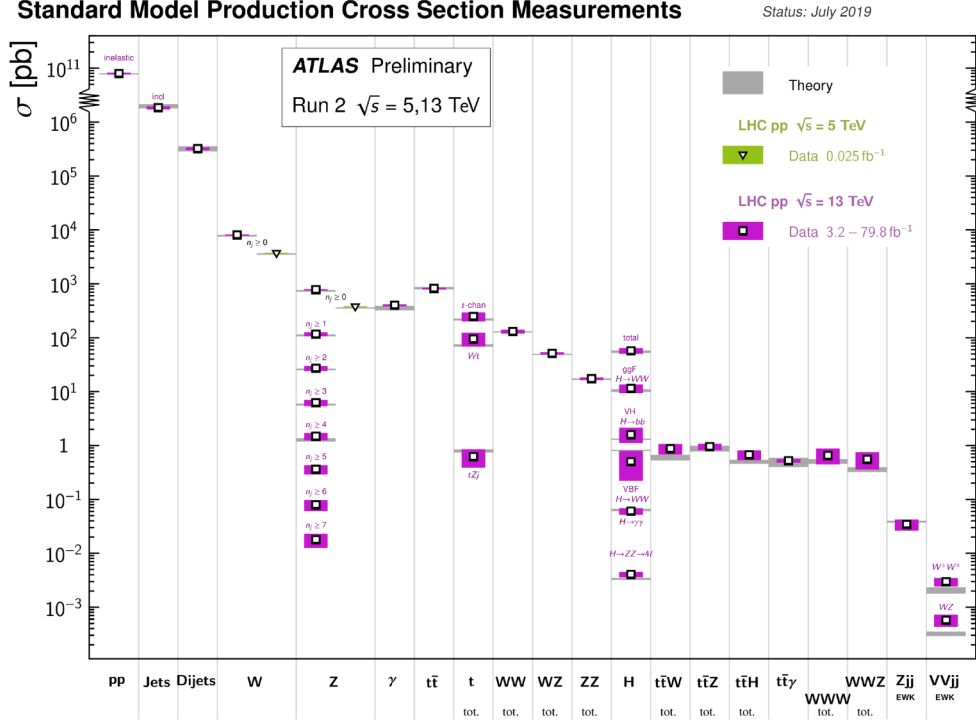


Figure 2.4: Summary of several Standard Model total and fiducial production cross-section measurements from Run 2, corrected for branching fractions, compared to the corresponding theoretical expectations. From: [31]

very active research. It is foreseen that many open questions will be answered over the next years, e.g. the mass hierarchy. Also, more fundamental questions are being studied. These include, e.g. whether the neutrino is a majorana particle.

Additionally, there are very fundamental phenomena on cosmic scales that cannot be explained within the SM and are addressed in extensions of the SM:

Matter-antimatter asymmetry While production of particles within the SM is mostly symmetric in the number of anti-matter and matter particles, a strong asymmetry is observed on macroscopic scales. The largest fraction of the visible universe consists of matter. The production of such asymmetry in the early universe is currently not well understood. The known sources of matter-antimatter asymmetry in the SM are not sufficient to explain the large observed asymmetry.

Dark matter The dependence of the velocity of stars from the distance to the centre of their galaxy is not consistent with the visible mass distribution. Large amounts of additional matter are necessary to explain the velocity with general relativity assuming standard gravity. This lead to the proposal of dark matter, a form of matter that does not participate in the electromagnetic interaction and thus is not visible with traditional astronomic instruments. There is also strong evidence for the existence of dark matter in other astronomical observations, e.g. the effect as gravitational lenses of matter or the effect on the cosmic microwave background.

There are many proposed expansions of the Standard Model to explain different subsets of these issues. Since this work studies a process predicted by the SM and an interpretation of potential deviations is beyond the scope of this work, no particular theory beyond the SM (BSM) is covered in detail. Popular expansions are, e.g. supersymmetry, extra dimensions, or string theory. The interested reader is referred to [34, 35] and reviews of [9].

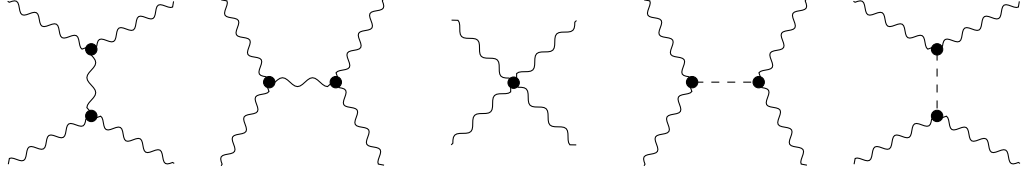


Figure 2.5: All leading-order Feynman diagrams that connect two initial and two final state massive gauge bosons in $VV \rightarrow VV$.

Predictions of these expansions, such as new particles, are actively being searched for in recorded data. No clear indications of new particles have been found yet, and instead, exclusion limits on model parameters were set. For many proposed expansion, typical exclusion limits are in the order of $\mathcal{O}(100 - 1000)$ GeV (see, e.g. [36]).

2.2 Scattering of Electroweak Gauge Bosons

The scattering of electroweak gauge bosons is a process where two gauge bosons interact with each other in the form $VV \rightarrow VV$. In this notation each V stands for any of the electroweak gauge bosons; $V = W^\pm, Z, \gamma$, while not all boson combinations are allowed. This process is also often called *vector boson scattering (VBS)*, which is used synonymously throughout this work.

The quartic gauge couplings, as shown in Figure 2.2 contribute to the scattering process, which opens the possibility of a direct experimental test of the existence of these couplings. However, a Lagrangian with the associated terms is only gauge invariant, when the terms that produce the triple gauge couplings shown in Figure 2.1 connected via a gauge boson propagator are included as well. Separating these diagrams is thus not possible in a gauge-invariant manner. For proper theory predictions, all of these diagrams have to be included in the calculation.

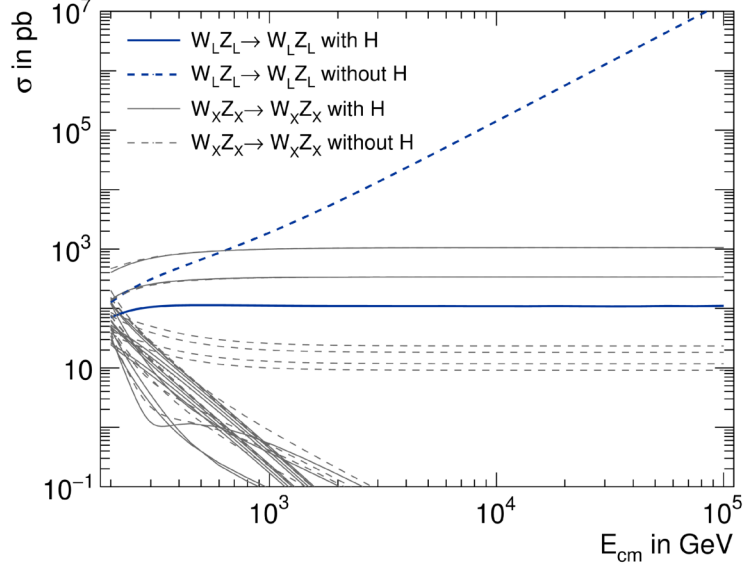
Via the couplings of massive gauge bosons to the Higgs boson shown in Figure 2.3, similar diagrams for $VV \rightarrow VV$ can be built for some boson combinations where the scattering is mediated by exchanging a Higgs boson. So at leading order, five Feynman diagrams are contributing to the scattering of electroweak gauge bosons. These are shown in Figure 2.5.

The existence of this scattering is a direct prediction of the gauge structure of the SM. The symmetry under non-Abelian transformations such as $SU(2)$ or $SU(3)$ directly implies the existence of direct couplings among the associated gauge bosons. The scattering mediated by such couplings is of particular interest in the electroweak $SU(2)_L$ case, due to the masses of the W^\pm and Z bosons.

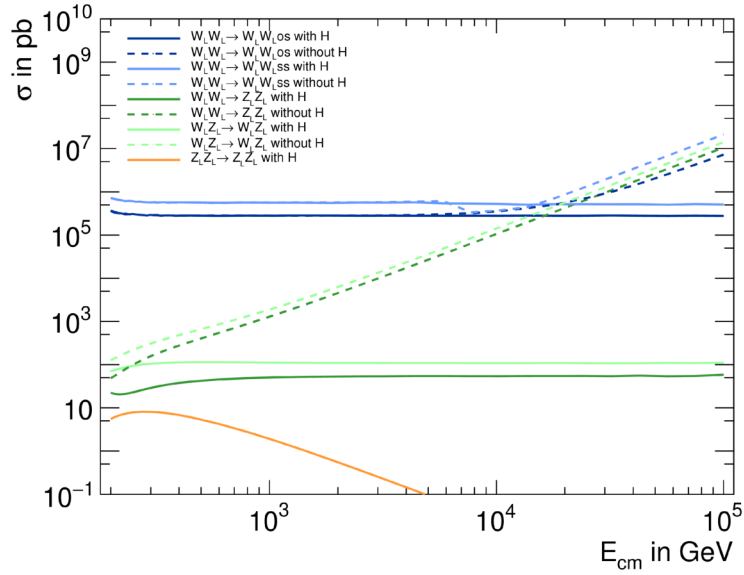
The polarization is defined as the projection of the spin of a particle onto its momentum. For massless spin-1 particles, such as photons and gluons, the polarization has only two allowed states, where the spin is parallel or anti-parallel to the momentum. These states are labelled as transversally polarized. Massive spin-1 particles have an additional allowed state, where the spin is perpendicular to the momentum. This state is labelled longitudinally polarized. As mentioned in Section 2.1.7, these states are connected to the Goldstone bosons.

The cross-sections for the scattering $V_L V_L \rightarrow V_L V_L$ of longitudinally polarized bosons without the Higgs-mediated contributions grow quadratically with the centre-of-mass energy $\propto E^2$. Such energy dependence breaks unitarity and is not physical. This unitarity violation can be interpreted as a scattering probability larger than 1 for sufficiently high energies. This prediction is a clear indication that the scattering of electroweak gauge bosons cannot be explained consistently in the SM without a Higgs boson. The scattering of gauge bosons with polarization states other than the purely longitudinal scattering does not break unitarity.

If one includes the SM Higgs boson as predicted by the BEH mechanism, i.e. with the couplings to the electroweak gauge bosons as shown in Figure 2.3, the additional contributions cancel the rise with increasing energy. This cancellation circumvents the violation of unitarity and makes the predictions physically meaningful. This behaviour is shown in Figure 2.6.



(a)



(b)

Figure 2.6: Dependency of the cross-section of polarization combinations of the scattering $VV \rightarrow VV$ on the centre-of-mass energy E_{cm} . Full lines show the cross-section including a SM-like Higgs boson and dashed lines for the case without a Higgs boson. All polarization combinations in the scattering $WZ \rightarrow WZ$ are shown in [2.6a](#). Blue lines show the scattering of purely longitudinal boson polarization while all other polarization combinations are shown in grey. All allowed boson combinations in the purely longitudinal scattering $V_L V_L \rightarrow V_L V_L$ are shown in [2.6b](#). All cross-sections are calculated using the Madgraph Monte Carlo event generator. [37](#)

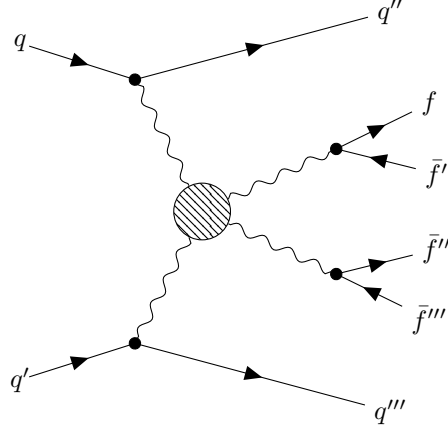


Figure 2.7: Schematic structure of Feynman diagram for the scattering of electroweak gauge boson scattering at the LHC. The dashed circle acts as a placeholder for any of the possible boson scattering diagrams shown in Figure 2.5.

The unitarity-violating prediction without the Higgs boson indicated that the SM was not complete without the Higgs boson at scales of about 1 TeV.

There needed to be some mechanism to restore unitarity in the scattering of longitudinal vector bosons. Depending on the coupling of the SM-like Higgs boson discovered in 2012 to the gauge bosons, the inclusion of this boson could only partially unitarize the scattering. For instance in models with multiple Higgs bosons, the inclusion of only one Higgs boson, i.e. the observed Higgs boson with a mass of ≈ 125 GeV, could only dampen the rise in cross-section with the scattering energy. In this scenario, only the inclusion of the second Higgs boson would completely cancel the rise and thus restore unitarity. In this prediction, substantial differences in the high-energy behaviour of the scattering of longitudinal bosons occur.

The existence of the scattering of electroweak gauge bosons is a fundamental prediction of the gauge structure of the SM. Studying this process also provides an opportunity to test the details of the mechanism of electroweak symmetry breaking.

2.2.1 Vector Boson Scattering at the Hadron Colliders.

Since the gauge bosons are not stable, the production of the initial bosons and the decay of the outgoing bosons have to be considered as part of the hard process. At hadron colliders such as the LHC (see Section 4.1), the initial bosons are radiated off by partons. The typical structure of Feynman diagrams for the full process from partons to the boson decay products is shown in Figure 2.7.

Since gluons cannot radiate the vector bosons, only quark-initiated diagrams occur at leading order. The initial quarks are lightly deflected from the beam axis since they only undergo an electroweak interaction. This implies a characteristic signature of two jets, called tagging jets, with a significant difference in rapidity and high invariant mass. In addition to these tagging jets, the final state contains the decay products of the vector bosons. This final state is usually denoted as $VVjj$, while explicitly also including diagrams with non-resonant bosons.

All couplings in the LO Feynman diagrams as shown in Figure 2.7 are in the electroweak interaction. The coupling structure of the squared matrix element for such diagrams is $|\mathcal{M}^2| \propto \alpha_{\text{EW}}^6$ at leading order, where α_{EW} denotes a generalization of α_{em} and α_{W} to a combined electroweak coupling only used for the determination of the order of a process. The same final state can also be produced with the coupling structure $|\mathcal{M}^2| \propto \alpha_{\text{EW}}^4 \alpha_{\text{S}}^2$, which has larger cross-sections due to the larger numerical value of α_{S} at typical hard scales.

Since only the $|\mathcal{M}^2| \propto \alpha_{\text{EW}}^6$ coupling structure has contributions of vector boson scattering, the definition of the signal process includes this requirement. The signal is the $VVjj$

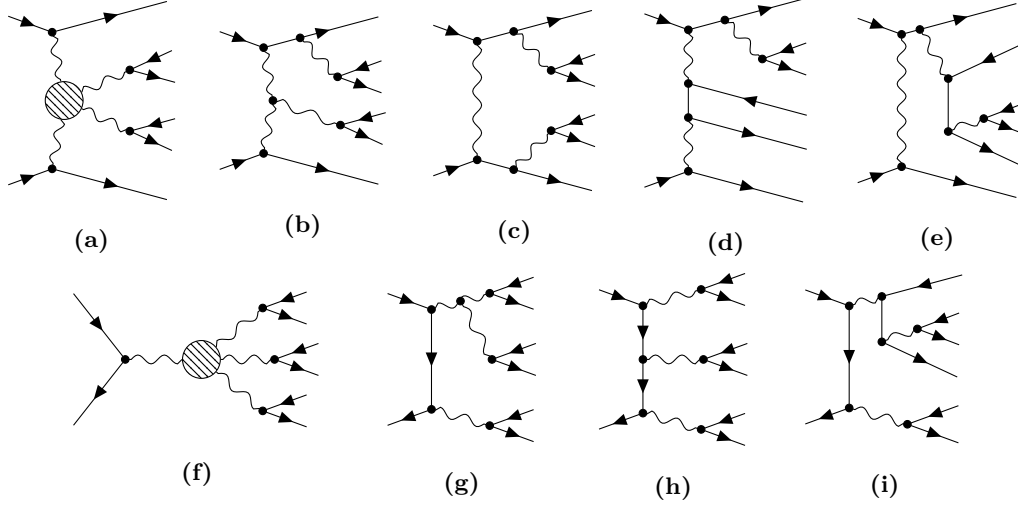


Figure 2.8: Feynman diagrams for the $VVjj$ -EW6 process at leading order perturbation theory. The dashed circle acts as a placeholder for any of the possible boson scattering diagrams shown in Figure 2.5. The top row shows diagrams in the t - or u -channel and the bottom row shows s -channel diagrams. In the s -channel diagrams, one of the outgoing fermion pairs has to represent quarks to result in the correct final state.

final state for a certain boson and decay product combination, where the diagrams with less than six electroweak interactions are vetoed and considered as background. This signal process is labelled as $VVjj$ -EW6. Depending on the study, sometimes also the interference contributions of order α_{EW}^5 are included in the signal process definition.

This signal definition includes many diagrams that do not contain vector boson scattering but have the same coupling structure. Examples for such diagrams are shown in Figure 2.8.

The different contributions are grouped by their topology as t - or u -channel, and s -channel, defined according to the quark lines. In the s -channel diagrams, the initial quarks are connected by a continuous fermion line. In the t - or u -channel contributions, the initial quark lines are connected to a final state quark each.

This has drastic effects on observed jet kinematic. In many of the s -channel diagrams, the final state quarks are the decay product of an electroweak gauge boson. These contributions lead to a distinct peak in the invariant mass of the two jets around 80 – 90 GeV and to increased contributions at lower invariant masses in general. This kinematic signature is a considerable contrast to the typical behaviour for typical vector boson scattering topologies with high invariant masses of the jets as mentioned above.

As indicated by the blob in Figure 2.8f some s -channel diagrams contain similar couplings as the pure vector boson scattering $VV \rightarrow VV$, realized in a $V \rightarrow VVV$ topology.

In nominal vector boson scattering diagrams, the final state fermions other than the tagging jets occur in pairs originating from a common electroweak gauge boson decay. Depending on the studied boson combination, the invariant mass of these pairs is used as a selection criterion to suppress backgrounds. Figures 2.8d, 2.8e, and 2.8i show *non-resonant diagrams*, where at least one of the fermion pairs does not originate from a common gauge boson decay.

The VBS signal process definition usually excludes terms of order $\mathcal{O}(\alpha_{\text{EW}}^4 \alpha_S^2)$. This background is historically often labelled QCD due to the occurrence of α_S compared to the VBS process, which is purely electroweak at LO. Throughout this work, the label $VVjj$ -EW4 is used instead. Diagrams for this process are shown in Figure 2.9.

Experimentally, one can not distinguish perfectly between a jet originating from the hard process, from the shower, or pile up. Thus, the two jets in the required final state can also be considered as radiative corrections to the diboson production of order $\mathcal{O}(\alpha_{\text{EW}}^4)$. The full

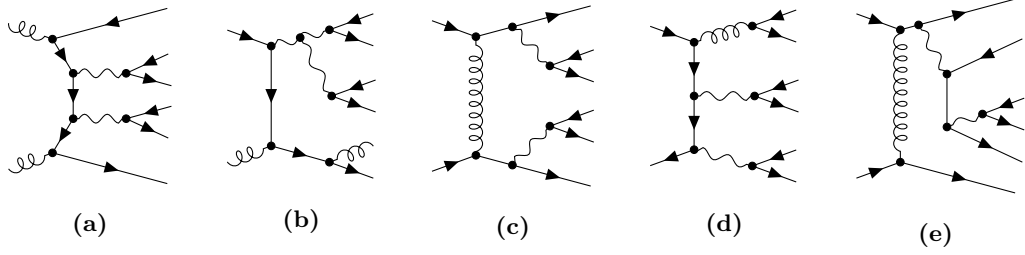


Figure 2.9: Feynman diagrams for the $VVjj$ -EW4 process at leading-order perturbation theory. Some diagrams, e.g. in Figures (a) and (b) can be interpreted as QCD radiative correction to diboson production VV motivating the inclusion of final states with lower jet multiplicities in the hard process calculation in simulations.

definition of this background consequently has to include final states with fewer jets in the final state. What usually is constant for this process is the coupling order in α_{EW} . Further motivating the label based on $\alpha_{EW} = 4$.

From Figure 2.9 it can be seen that in contrast to the $VVjj$ -EW6 process, there are gluon-induced diagrams. These diagrams are typically PDF-enhanced, i.e. particularly for low parton momenta, the probability of finding a gluon is much higher than to find a quark. Since α_S is also larger at typically energies than α_{EW} , the $VVjj$ -EW4 process is also less suppressed by the couplings leading to significantly larger cross-sections.

In order to study $VVjj$ -EW6, it is essential to find kinematic differences to $VVjj$ -EW4 to find selection criteria to suppress $VVjj$ -EW4 contamination. The different behaviour of the final state jets offers an excellent possibility to do that. In $VVjj$ -EW4 these are always colour connected, making additional QCD radiation along this connection more likely. Also the characteristic signature of high invariant masses and a significant rapidity separation for $VVjj$ -EW6 are not expected in $VVjj$ -EW4. A more detailed study of the kinematic differences between $VVjj$ -EW4 and $VVjj$ -EW6 is shown in Section 3.7.

When focussing on the $VVjj$ final state at leading order, also interference effects between the $VVjj$ -EW4 and $VVjj$ -EW6 processes have to be taken into account. Such interferences appear for all diagrams which have the same initial and final state and are allowed to be negative. These have the coupling order of $\mathcal{O}(\alpha_{EW}^5)$ and are suppressed if one of the processes is suppressed.

2.2.2 Comparison of final states

Vector boson scattering can occur with different boson combinations; $W^\pm W^\pm$, $W^\pm W^\mp$, $W^\pm Z$, $W^\pm \gamma$, ZZ , $Z\gamma$, $\gamma\gamma$. Due to the couplings to the Higgs boson and thus the connection to the mechanism of electroweak symmetry breaking, the scattering of massive gauge bosons is of particular theoretical interest. Experimentally, these channels have additional complexity due to the decay of the gauge bosons. While photons from the hard process can be identified directly or via electrons after conversion, the massive W^\pm and Z bosons are unstable.

The hadronic decays of gauge boson have high branching ratios. However, additional backgrounds contribute because of the resulting final state of, e.g. $\bar{\ell}\nu_\ell jjjj$ has also contributions with coupling order $\mathcal{O}(\alpha_{EW}^2 \alpha_S^4)$ at leading order. Also, the tagging and measurement of jets is far less accurate compared to leptons.

The decay modes containing charged leptons offer far cleaner signatures and are easier to identify. The major disadvantage of the leptonic final states is reduced branching ratios leading to lower signal cross-sections.

An overview of the currently available analyses of the ATLAS and CMS collaborations is shown in Table 2.4. Experimentally, the fully-leptonic decay modes offer the best sensitivity.

The $W^\pm W^\pm$ channel in the fully leptonic $\ell^\pm \ell^\pm \nu_\ell \nu_\ell jj$ final state has significantly smaller contamination due to the $VVjj$ -EW4 background process. Due to charge conservation, no

Table 2.4: Summary of available experimental measurements of electroweak gauge boson scattering at 13 TeV in different final states is shown. For each measurement, the collaboration, the observed and expected significances, and the observed and predicted fiducial cross-sections including total uncertainties in fb, the accuracy of the hard-process calculation, and the MC generator used for the nominal prediction are shown. The measurement in the $\ell^\pm \ell^\pm \nu_\ell \nu_\ell jj$ final state by the ATLAS collaboration uses an NLO QCD prediction for the signal process in the VBS-approximation (see Section 3.4) indicated by NLO*. References to the publications are indicated in the last column.

Final state	Significance [σ]		$\sigma(VVjj - \text{EW6})$ [fb]		Prediction	Ref.
	Obs.	Exp.	Obs.	Pred.		
$\ell^\pm \ell^\pm \nu_\ell \nu_\ell jj$						
- CMS	5.5	5.7	3.83 ± 0.75	4.25 ± 0.27	LO MG5_AMC	[1]
- ATLAS	6.5	6.5	$2.89^{+0.59}_{-0.56}$	$3.08^{+0.45}_{-0.46}$	NLO* POWHEGBOX	[2]
$\ell^\pm \ell^\mp \ell^\pm \nu_\ell jj$						
- CMS	2.2	2.5	$1.0^{+0.6}_{-0.5}$	$1.25^{+0.13}_{-0.11}$	LO MG5_AMC	[38]
- ATLAS	5.3	3.2	$0.57^{+0.26}_{-0.14}$	$0.321^{+0.028}_{-0.024}$	LO SHERPA	[3]
$\ell^\pm \ell^\mp \ell^\pm \ell^\mp jj$						
- CMS	2.7	1.6	$0.40^{+0.25}_{-0.18}$	$0.29^{+0.02}_{-0.03}$	LO MG5_AMC	[39]
- ATLAS	5.5	4.3	0.82 ± 0.21	0.61 ± 0.03	LO MG5_AMC	[40]
$\ell^\pm \ell^\mp \gamma jj$						
- CMS	3.9	5.2	3.2 ± 1.15	4.97 ± 0.29	LO MG5_AMC	[41]
- ATLAS	4.1	4.1	7.8 ± 2.0	7.8 ± 0.5	LO MG5_AMC	[42]
$V(\rightarrow jj)Vjj$						
- ATLAS	2.7	2.5	43^{+18}_{-17}	43.0 ± 2.4	LO MG5_AMC	[43]

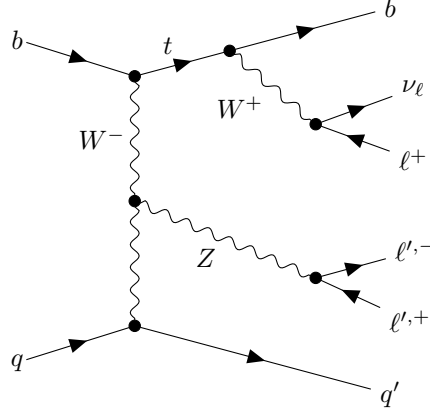


Figure 2.10: Example Feynman diagram for the tZj process at leading order perturbation theory. The upper quark line contains a top-quark resonance, motivating the definition of these contributions as a background process.

gluon-induced diagrams are allowed at LO. This process suffers from two different kinds of experimental backgrounds. The misidentification of other signals as prompt leptons is a prevalent background for analyses. Additionally, a misidentification of the charge of a lepton introduces another large source of background contamination. These experimental backgrounds make the analysis of this process more complicated, but overall this final state shows the largest expected significances.

The $W^\pm W^\mp$ channel suffers from higher background contaminations not only due to the $VVjj$ -EW4 process, but also originating from, e.g. $t\bar{t}$ and Z +jets production and no dedicated study of the VBS process in this channel is available. The $W^\pm Z$ channel in the leptonic $\ell^\pm \ell^\mp \ell^\pm \nu_\ell jj$ final state offers a good compromise between good reconstructability and a still reasonably high cross-section. The ZZ channel in the $\ell^\pm \ell^\mp \ell^\pm \ell^\mp jj$ final state has a very clean signature with small background contamination, but the cross-section is reduced due to the lower leptonic branching fraction.

The publication timeline of available observations confirms this assumption. The scattering of vector bosons was first observed in the $W^\pm W^\pm$ channel by the CMS collaboration in early 2018 using the dataset recorded in 2015 and 2016 [1]. Later in the same year, the ATLAS collaboration published observations in the $W^\pm W^\pm$ and $W^\pm Z$ channels using the equivalent dataset [2, 3]. In 2019, the ATLAS collaboration also observed VBS in the ZZ channel using the full Run 2 dataset [40].

The observation in the $W^\pm Z$ channel by ATLAS [3] was published in the context of this work. This work follows the analysis procedure that was developed for [3], but some aspects of the analysis are adjusted and improved.

An important difference between the $W^\pm Z$ channel and most other boson channels in VBS is the contribution of diagrams, where a single t quark is produced in association to a Z boson. When considering the decay chain $t \rightarrow bW^+ \rightarrow b\ell^+\nu_\ell$, the final state and coupling structure match those of the $WZjj$ -EW6 signal process. An example Feynman diagram is shown in Figure 2.10. This subprocess contributes significantly to the $\ell\ell\ell\nu jj$ final state due to the enhancement of the t resonance. The contributions are separated according to the quark flavours in the hard-process. The tZj process always has a b quark in the final and initial state so the VBS signal process $WZjj$ -EW6 is restricted to events without any b quark in either the initial or final state of the hard process calculation. An advantage of this definition is the clear distinction from the $WZjj$ -EW6 signal process without interference between these two processes. The tZj process in this definition still has contributions of VBS topologies, where the incoming Z boson is radiated off a b quark.

With these channel-specific restrictions, the definitions for the $WZjj$ -EW6 signal process

and its major backgrounds are complete. With these definitions, simulations can be created to study the kinematic differences and the expected distinct signature in detail.

Chapter 3

Simulations and Modelling Studies

A key ingredient for analyses of particle collisions is the availability of accurate and reliable simulations. Such simulations include the generation of events for a given process based on an assumed theory. The methods and procedures available to produce such simulations are introduced and discussed for the $\ell\ell\nu jj$ final state below. An excellent overview with more detailed discussion is presented in [44]. For a direct comparison to data, also the detector effects have to be taken into account as discussed in Chapter 4.

Information of events that have passed each step in the generation process is referred to as truth-level or particle-level. For a valid comparison to observed data, a simulation of detector effects and the subsequent reconstruction chain is necessary as well. After these further steps, the information is referred to as detector-level or reconstruction-level.

3.1 Methodology

A visualisation of the structure of a simulated hadron collision is shown in Figure 3.1. In order to reduce the complexity of the simulation, the full collision has to be split into multiple phases. The matrix element calculation in perturbation theory, in general, achieves the most accurate predictions. This approach is not valid at low momentum transfers due to the energy-dependent couplings. At low energies, due to the increasing couplings and infra-red logarithms at each order, higher-order terms are not suppressed sufficiently to justify their neglect. In order to model such non-perturbative effects, different models are available which have been tuned to previous measurements. This motivates a splitting of the full hadron collision along the energy scale of the individual interactions.

3.1.1 Phases of event generation

Hard Process: The part of the interaction at the largest scale comprises the hard process. The hard process is usually calculated in fixed-order perturbation theory from the matrix elements of the given scattering process. The full calculation is based on Equation (2.71) which can be written in more detail for proton-proton collisions as

$$\sigma = \sum_{a,b} \int_0^1 dx_a dx_b \int d\Phi_n \text{PDF}_a(x_a, \mu_F) \cdot \text{PDF}_b(x_b, \mu_F) \cdot \frac{1}{2x_a x_b s} |\mathcal{M}|^2(\Phi_n; \mu_F, \mu_R), \quad (3.1)$$

where a and b iterate over the allowed initial state partons, μ_F is the factorization scale, μ_R is the renormalization scale, s is the hadronic centre-of-mass energy squared, x_a and x_b are the momentum fractions of the partons a and b with the parton distribution function $\text{PDF}_a(x_a, \mu_F)$, $d\Phi_n$ is the differential phase space element for a n -particle final state and $\mathcal{M}(\Phi_n; \mu_F, \mu_R)$ is the matrix element for the transition from the initial to the final state.

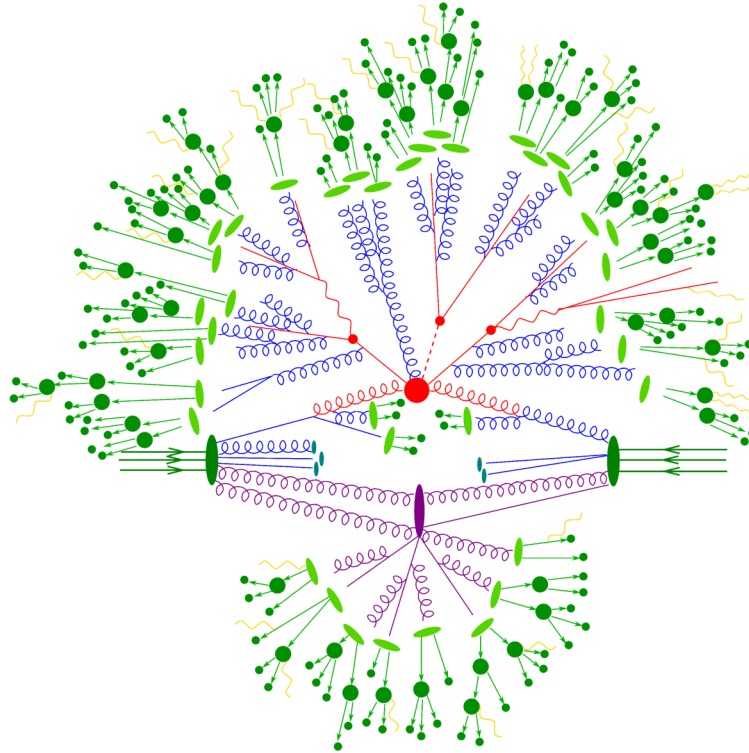


Figure 3.1: Schematic representation of a $t\bar{t}H$ event as produced by an event generator. The hard interaction (big red blob) is followed by the decay of both top quarks and the Higgs boson (small red blobs). Additional hard QCD radiation is produced (red), and a secondary interaction takes place (purple blob) before the final state partons hadronise (light green blobs) and hadrons decay (dark green blobs). Photon radiation occurs at any stage (yellow). From: [\[45\]](#).

This equation describes the factorization of the hard process from the resolution of the protons' substructure as defined in the factorization theorem. The low-energetic splittings in the proton can not be described accurately by perturbation theory. The effects of this are therefore modelled in PDFs. The factorization scale defines at which scale the transition from the hard process to PDFs occurs. While the evolution to higher scales is analytically possible in terms of the DGLAP equations, a fully analytical description of such effects is not known. Different PDF sets are available and are fitted to data from previous measurements.

In order to evaluate the hard process cross-section, the squared matrix element has to be determined. The matrix elements are either pre-defined in event generation software, put in using external libraries, or determined dynamically based on specified initial and final states given a set of restrictions, e.g. on propagators or the coupling order. While pre-defined matrix elements are usually very optimized and thus far more efficient, dynamically generated matrix elements offer great flexibility to specify the processes to be considered.

The matrix elements depend on the particle kinematics of the considered phase space point as well as on the factorization and renormalization scales μ_F and μ_R . These scales are unphysical, and a full calculation to all orders of perturbation theory would not depend on these scales.

The renormalization scale is introduced in regularization. Some loop-contributions in fixed order calculations need to be regularized to avoid divergencies which would be cancelled in higher-order contributions in a calculation to all orders. Such divergencies are avoided by adjusting physical parameters to parametrize the effect of loop corrections. The renormalization scale defines which corrections should be included in the effective coupling and which should be resolved. For instance, the strong coupling α_S is heavily affected by loop corrections, leading to a significant dependence of α_S on the renormalization scale μ_R .

These scales can be set to fixed values, but can also be determined dynamically based on the event kinematics. μ_R and μ_F are commonly chosen to be equal to each other and set to typical momentum transfers or resonance masses.

The squared matrix element is integrated numerically over the allowed region of phase space. Due to the dimensionality and the complexity of the phase space integrals, numerical integrations are usually performed using the Monte-Carlo method, which performs better than other integration methods specifically at high dimensions. The central procedure is to evaluate the integral not on a fixed grid, but at a randomly distributed set of points.

The matrix element will contain resonances and even divergencies which have to be treated with care in fixed-order calculations. Divergencies are typically excluded from the phase space using an appropriate set of selection criteria, e.g. on transverse momenta, invariant masses, or angular distances. Without further adjustments, resonances would cause severe inefficiencies.

An easy way to account for resonances is to assign a weight to each event based on the squared matrix element. However, such weights would vary by several orders of magnitude due to resonances. For experimental purposes, where each event has to pass a computationally expensive detector simulation as well, large variations of the event weight are problematic.

For this reason, an unweighting procedure is applied, and phase space points are vetoed with a probability depending on the squared matrix element, which causes inefficiencies during the generation of events. In general, this unweighting procedure is most efficient when the squared matrix element is either flat or can be transformed into a flat function. Phase space integrators apply such transformations automatically to reduce the inefficiencies and increase the performance of the phase space sampling.

In the first step, event generators optimize these transformations and improve the overall accuracy of the calculated total cross-section. Subsequently, the software generates events by a further sampling of the phase space.

Parton shower: Typically, no or only a few additional QCD emissions are included in the calculation of the hard scattering process due to the increased complexity of the phase space and the matrix elements. Further radiation from particles in the event is simulated

using parton showers using a Markov chain approach. Such radiation can stem from initial- and final state partons. This procedure evolves the parton ensemble from the scale of the hard process by iteratively adding parton emissions, down to a few GeV, where perturbation theory breaks down. This parton shower approach essentially approximates higher-order contributions to all orders.

Splitting functions for QCD radiations, i.e. probabilities for gluon emissions or quark-pair production are known analytically and diverge for low energetic radiations or radiations at small angles. These divergencies motivate the usage of infrared and collinear safe observables (see Section 2.1.5). The algorithm for clustering is defined in detail in Section 3.2.5.

Different implementations of parton showers exist, differing in the splitting functions, the kinematic scheme for the distribution of recoil, or the evolution variable. Most commonly used implementations are the k_{\perp} -ordered shower, the angular-ordered shower, and the dipole shower. The former two algorithms show differences in the treatment of soft emissions for two partons that are close to each other. In the k_{\perp} -ordered shower the splitting functions need to be adjusted to avoid double counting of such emissions. This double-counting is automatically avoided in angular-ordered showers.

In dipole showers, the particles are emitted from colour dipoles rather than single particles, which also avoids such double-counting. In each iteration of the shower algorithm, a colour-flow is assigned to the event, and the shower scale evolution is performed globally for the event. The modern versions of the most commonly used shower generators have implemented variants of the dipole shower.

The additional splittings during the parton shower simulation, especially for initial state radiation introduces a momentum recoil on the rest of the event. For the k_{\perp} -ordered and angular-ordered showers momentum conservation needs to be violated at intermediate steps to allow for the $1 \rightarrow 2$ splittings. Since dipole showers use $2 \rightarrow 3$ splittings, violations of momentum conservation can be avoided entirely. These recoil effects are propagated to the full event, and slight necessary momentum boosts and adjustments are applied to the full event to restore momentum conservation and the overall kinematic. More details are discussed in [46].

Hadronization: Once the event scale is evolved in the parton shower to $\mathcal{O}(1 \text{ GeV})$, perturbation theory breaks down. At this point, phenomenological models have to be applied to describe the formation of colourless hadrons from the unobservable coloured partons.

The models used in modern event generators are the string and the cluster model. These models describe the creation of quark-antiquark pairs during the separation of colour-connected partons according to their momenta.

The string model follows the reasoning mentioned in Section 2.1.5 and has many free parameters which are adjusted to match observed data.

In the cluster model, the formation is based on a property called pre-confinement. This property describes that colour-singlet clusters of partons can be formed at each value for the evolution scale with an invariant mass distribution that is independent of the scale of the hard process or the starting scale of the shower.

The cluster model includes more elements based on perturbative QCD and includes fewer free parameters. The overall agreement with observed data is slightly inferior compared to the string model.

Decay of unstable particles: Hadronization produces many unstable hadrons. The decay of these particles has to be included to achieve an observable final state of particles that are stable on a collider time scale. Special care must be taken in order to properly include the effects of spin correlations and excited unstable particles. The simulation of such decays is performed in some shower generators using external packages, e.g. EvtGen[47] for hadron or TAUOLA[48] for tau-lepton decays. Other shower generators have improved their internal decay libraries. Modern implementations employ matrix element calculations combined with models based on experimental data.

QED radiation: Two common approaches for the simulation of QED radiation are available. One is to employ the same shower algorithm as applied for QCD radiations while considering electromagnetically charged rather than coloured particles. A major issue compared to QCD radiation is the increased effect of terms, which are suppressed by a colour factor of $1/N_C^2$ in QCD. This approach has the advantage that QED and QCD radiations can be evaluated simultaneously.

An alternative approach is provided by the Yennie-Frautschi-Suura (YFS) formalism. Here radiation of multiple photons is simulated iteratively and finally corrected to match the event-wide analytical calculation for the given amount of photons. This approach allows the exact inclusion of higher-order corrections for the distributions of radiated photons.

Underlying event: The final ingredient for the simulation of hadron collisions is to include effects that are not directly linked to the hard interaction – the *underlying event*. The hard process is initiated by two individual partons of a pair of initial hadrons. The remaining particles of the same hadron pair have a non-zero probability of interacting with each other, and the effect of such interactions have been confirmed experimentally. Such additional interactions are called *multiple parton interactions (MPI)*. The high-energy effects of such interactions can be described using perturbative QCD dominated by dijet production in t -channel gluon exchange. In contrast, low-energy effects are usually modelled to match data. The particles produced by MPI also have effects on the main interaction, especially on the colour reconnections of the hard process and its radiations.

3.1.2 Matching and merging

Additional QCD radiation on top of the subprocess with lowest coupling order referred to as born process can be modelled either via the hard process or via the parton shower. Modelling as part of the hard process describes individual hard and wide-angle emissions better. On the other hand, the parton shower is the only way to model the multitude of emissions at low energies and low angles. The occurrence of large logarithms in these kinematic regions of phase space and the fastly increasing dimensionality of the phase space with each emission render an inclusion in the hard process calculation impossible.

When going beyond a LO calculation of the born process in the hard process, the combination with the parton shower is non-trivial, and many details need to be considered thoroughly. A naive implementation of an NLO calculation with a parton shower would introduce double-counting of emissions among other problems.

The double-counting can easily be seen in Z boson production followed by a leptonic decay $pp \rightarrow Z \rightarrow \ell^+ \ell^-$. In the leading order matrix-element calculation for this process, the transverse momentum of the dilepton system $p_T(\ell\ell)$ is fixed to 0 since no additional emissions are included. When applying a parton shower simulation to the leading-order calculation, the dilepton system recoils against all of the QCD radiations, mostly from initial state radiation, and thus gains a soft transverse momentum distribution. This simulation describes the experimental distribution reasonably well for $p_T(\ell\ell)$ below few GeV, but high transverse momenta are not described well.

The modelling can be improved by including the first emission in the hard process calculation, either by going to NLO or by adding $pp \rightarrow Zj \rightarrow \ell^+ \ell^- j$ as the second process. When adding the extra emission as an additional process, a minimum transverse momentum of the final state parton is required to avoid large logarithms. This expansion of the process definition significantly improves the agreement to measurements for high p_T , especially above the minimum transverse momentum of the jet. However, emissions in the transition region from the low- to high- p_T ranges are partially covered by both the hard-scattering and the parton shower simulations. Without further adjustments, such simulations overestimate the experimental data.

Depending on the setup of the matrix element calculation, different approaches are available for a proper combination of the hard process with the parton shower.

Subtraction approach—Matching. For the combination of a full NLO matrix element

calculation and the parton shower, *matching* algorithms are available that directly subtract effects to order α_S of the parton shower on the born process from the matrix element. In other words, the matrix element calculation is matched to the parton shower. The resulting events can then be processed with the respective parton shower.

Vetoing approach—Merging. For cases where multiple real emissions are included as individual processes at tree level, the combination is possible using *merging* algorithms. With such LO *multi-leg* setups it is possible to also include a second or third emission from the matrix element. The merging algorithm vetos parton shower emissions that overlap with the matrix element calculation.

Some setups also allow the combination of matching and merging to combine NLO multi-leg calculations with a parton shower properly. Both approaches are discussed in detail below.

Matching: The parton shower approximates the full calculation to all orders by including the leading logarithmic contributions in each order of α_S . Since these terms are also included in the full NLO calculation, they would be double-counted when combined naively. Two formalisms for NLO matching are available and commonly used, MC@NLO- and POWHEG-style matching. A good comparison is given in [49].

The basic principle of MC@NLO-style[50] matching algorithms is to calculate the leading logarithmic terms in that order explicitly and to subtract these effects from the full NLO matrix element calculation. These are the terms the parton shower adds when applied to the LO calculation. This subtraction avoids the double-counting but introduces the possibility for negative event weights.

The POWHEG [51, 52] formalism follows a slightly different approach to avoid negative event weights. Since the problematic double-counting only occurs for the first emission of the parton shower, this emission is always simulated in the ME generator. This is done in a parton-shower-like approach where the non-emission probabilities, included in so-called Sudakov form factors, are calculated from the real-emission matrix element rather than approximations of it. This removes the necessity to introduce subtraction terms and thus, negative weights. The hard process cross-section is then the same as for the full NLO calculation. This approach, however, moves part of the complexity to the parton shower simulation, which now has to be adjusted to match the emission pattern of the hard process generation. It has to be ensured that the already generated emission is the hardest in the shower simulation and that still the full kinematic region for shower emissions is covered. This can be done by using truncated and vetoed showers [49].

Merging: There are many different implementations of tree-level merging algorithms. The most commonly used implementations differ in details but are typically based on one of two basic schemes, the MLM-scheme[53] and the CKKW-scheme[54, 55]. A comparison study can be found in [53].

MLM merging is motivated based on phenomenological arguments. After the simulation of the parton shower, jet objects are clustered and matched based on the angular distance ΔR to the matrix element partons. An event is vetoed if a matrix element parton could not be matched to a jet or a jet is matched to multiple matrix element partons. This veto ensures that each matrix-element parton can be resolved as individual emission. If an event does not originate from the subprocess with the highest number of additional emissions in the hard-process calculation, the event is vetoed if an additional jet is clustered that is not matched to any of the matrix element partons. The performance of this scheme is dependent on the exact clustering and matching criteria used, and no formal accuracy can be derived.

An alternative scheme, based on an analytic formulation of emission probabilities, is available via the CKKW-scheme. Roughly speaking, a merging scale is introduced and used to divide the available phase space for emissions. If an event with additional emissions compared to the born process in the matrix-element calculation has an emission below that merging scale, the event is vetoed. For events from processes with a non-maximal number of considered emissions, parton shower emissions above the merging scale are vetoed. This veto

effectively removes the double-counting, and the parton shower (matrix element calculation) dominates for soft (hard) emissions. Typical merging scales are $\mathcal{O}(25 \text{ GeV})$ and additional corrections for non-emission probabilities and different scales for α_s are applied.

In the original CKKW-scheme, the Sudakov form factors to correct for the non-emission probabilities are derived analytically. An extension of this scheme, the CKKW-L scheme generates the Sudakov form factors dynamically from the shower by vetoing events rather than emissions, which is more flexible for new showers.

In recent years, implementations for NLO merging have become available. The procedures are mostly based on combinations of one of the mentioned merging schemes with either MC@NLO or POWHEG-style matching. The details of these implementations are beyond the scope of this work.

3.2 Overview of Monte-Carlo Generators

3.2.1 Sherpa

SHERPA is a general-purpose Monte Carlo event generator framework for the simulation of particle collisions in high-energy collider experiments [45, 56]. A focus in the framework design was the best possible merging of the matrix element calculation and the parton shower. The SHERPA generator framework achieves this in an automated fashion by covering the main aspects of the simulation internally without the need to rely on external interfaces. For the sake of flexibility interfaces to external libraries can be used for a collider setup with non-default PDFs or to produce events for BSM physics models.

The SHERPA framework includes two tree-level matrix-element generators AMEGIC[57] and COMIX[58], while for one-loop matrix elements SHERPA relies on external loop-matrix-element providers like BLACKHAT[59], OPENLOOPS[60], or RECOLA[61]. For the simulation of the parton shower, two algorithms are available. The CSSHOWER[62] is the default and is based on CataniSeymour dipole factorisation[63]. The DIRE[64] shower is an alternative shower and follows a hybrid approach between a dipole shower and standard collinear shower evolution. Hadronization is based on the cluster fragmentation model.

For matching of NLO matrix elements, MC@NLO-style matching is implemented. LO merging is available and implemented in the MEPS@LO[65] merging algorithm, which is based on the CKKW-L scheme and truncated showers. For merging of NLO processes the merging algorithm has been extended to MC@NLO-matched processes in the MEPS@NLO merging algorithm[66].

QED radiation is implemented with accuracy at leading order in α for some processes. A decay library for the simulation of τ and hadron decays is available. Many output formats like HepMC[67] and LHE[68] are available. Systematic variations of the renormalization and factorization scales, and the PDF-set can be calculated on-the-fly.

SHERPA supports per-process enhancement factors, which allow adjusting the proportions of produced events per subprocess. When a generation contains multiple subprocesses, e.g. different jet multiplicities, events for a subprocess with enhancement factor > 1 are produced more often, and the event weights are adjusted to correct the final distributions. Such enhancement factors allow a statistical enhancement of certain regions of phase space.

3.2.2 MadGraph5_aMC@NLO

MADGRAPH5_AMC@NLO [69] is a fully automated, general-purpose, and public event generator, which merges all the features of MADGRAPH5[70] and AMC@NLO[71, 72] in a unique framework. MADGRAPH5_AMC@NLO contains all ingredients to perform an automated calculation at NLO QCD. This includes the automatic cancellation of infrared singularities, as well as the evaluation of renormalized one-loop amplitudes and the matching to parton showers. Throughout this work also the short notation MG5_AMC is used.

The matching of the matrix element calculation is based on the MC@NLO approach. The simulation of the parton shower has to be performed in external tools, and the matching has

been validated for the commonly used parton shower simulations in HERWIG and PYTHIA (see below). Merging to parton showers is based on the MLM-scheme at tree-level. For NLO calculations the FxFx scheme[73], which combines the MLM merging scheme with MC@NLO-style matching, is available for merging. When PYTHIA8 is used for the parton shower simulation also CKKW-L style merging is available.

The automated approach of NLO calculations also holds for renormalizable extensions of the SM, if the necessary inputs are provided. For this, MADGRAPH5_AMC@NLO has full compatibility with FEYNRULES[74] the common FEYNRULES output format which can be converted into a MADGRAPH5_AMC@NLO model in an automated way.

Systematic variations of the renormalization and factorization scales, the PDF-set as well as model parameters, e.g. particle masses or BSM coupling parameters can be calculated on-the-fly.

MG5_AMC was used for simulations in this work in versions 5.2.6.3 and 5.2.6.5.

3.2.3 Herwig 7

HERWIG 7 [75, 76] is the most current release in the HERWIG event generator family, which combines and supersedes the original HERWIG, implemented in Fortran and the C++-based HERWIG++ release branch. HERWIG 7 is a general-purpose event generator with special emphasis on soft QCD, parton shower, and hadronization effects and can be used in standalone or in shower mode, where the parton shower can be simulated for events from a hard process generation of external tools.

For the simulation of the parton shower, both an angular ordered shower and a dipole shower are available. The hadronization is based on the cluster hadronization model. Hadron decays are simulated using matrix elements, including spin correlations and off-shell effects where possible [75].

In recent versions, the MATCHBOX [77] framework, for flexible NLO hard-process calculations combined with parton shower simulation is available when running in standalone mode. This framework relies on external libraries to provide tree or loop-level matrix elements because built-in matrix elements are only available for a few SM processes. Several external libraries like MG5_AMC (tree- and loop-level), OPENLOOPS (loop-level), or VBFNLO (tree- and loop-level, see below) are available and different combinations can be used for tree- and loop-level.

HERWIG 7 can simulate the hard process using these matrix elements and offers different approaches for combining the calculation and the parton shower simulation. For matching both the POWHEG and MC@NLO approaches are implemented. NLO merging of multi-leg processes was implemented in HERWIG 7 similar to UNLOPS[78], which is based on CKKW-L style merging combined with MC@NLO style matching and requires the usage of a dipole shower.

Systematic variations of the renormalization and factorization scales, the PDF-set as well as some shower parameters can be calculated on-the-fly for the shower simulation. Extending this approach to include variations in the simulation of the hard-process is being worked on, but not available at the time of writing.

HERWIG 7 was used for simulations in this work in version 7.1.3.

3.2.4 Pythia 8

PYTHIA 8 [79, 80] is a very commonly used event generator, focusing on the simulation of the parton shower and soft QCD effects. While PYTHIA 8 also can simulate the hard process, this feature is not used for this study. Instead, PYTHIA 8 is also able to simulate the parton shower and subsequent steps for events generated by external programs such as MG5_AMC, VBFNLO, or POWHEGBox.

The default shower in PYTHIA 8 is p_T -ordered, but a dipole shower is implemented as well. Hadronization is modelled using the Lund string fragmentation framework[81]. While basic modelling of particle decays is implemented, polarization and non-trivial matrix element

effects are neglected for many particles. PYTHIA 8 can be interfaced so that other libraries can be called for the decay of individual particle types.

3.2.5 Other event generators and tools

While the most important simulations for this work have been obtained using combinations of the event generators mentioned above, also several other generators are of interest:

VBFNLO A process-specific event generator for the hard process with focus of interactions of electroweak gauge bosons[82–84]. Matrix elements are optimized by hand and built-in, so only a limited set of processes is available. Standalone VBFNLO supports event generation at LO or the evaluation of the total cross-section at NLO. Using the HERWIG 7 MATCHBOX interface, the NLO matrix elements can also be used for event generation.

PowhegBox A process-specific event generator with an implementation of the POWHEG matching scheme[85]. Matrix elements for many processes have been implemented by hand with different orders of accuracy. Systematic variations for the renormalization and factorization scales as well as for the PDF set can be evaluated on-the-fly.

WHIZARD A general-purpose event generator for the simulation of a hard process[86, 87].

Another tool of special interest is FASTJET[88], implementing the *AntiKt* [89] algorithm for the clustering of jets. It recursively checks all distance parameters between pairs of input objects d_{ij} and between each input object and the beam d_{iB}

$$d_{ij} = \min \left(p_{T,i}^{2p}, p_{T,j}^{2p} \right) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = p_{T,i}^{2p}, \quad (3.2)$$

where p and R are free parameters of the algorithm and the other observables of the input objects follow the notation described in Section 4.2.2 below. If the smallest distance parameter in the event is d_{iB} object i is treated as jet candidate and removed from the list of input objects. If the minimal distance parameter is d_{ij} , objects i and j are merged. This procedure is repeated until no input object is left. Throughout this work the free parameters are chosen to be $p = -1$ and $R = 0.4$.

3.3 Problems and Caveats

Many problems in the modelling of VBS in different event generators have been found over the last years. One reason for this is the intrinsic complexity and high dimensionality of the phase space. Due to this, a full physics validation is beyond the scope of internal testing by the authors and problems are occasionally found in large comprehensive comparison studies. Sometimes also assumptions that are reasonable for other processes do not hold for VBS processes. Because of that, in some cases default settings of event generators have been shown to produce wrong results.

In diagrams with typical VBS topology, there is no colour flow at leading order between the two final state quark lines. When a third jet is included in the matrix element calculation, either via merging or by going to NLO, this third jet reflects the colour topology. Since this is not feasible for many final states, the third jet is commonly simulated as part of the parton shower. The non-trivial colour flow is problematic for some available shower generators since some parton shower generators assume a colour connection.

Such problematic behaviour has been observed for the PYTHIA 8 parton shower in [90]. Among other observables, the Zeppenfeld variable for the third-hardest jet $z(j_3)$ defined as

$$z(j_3) = \frac{1}{|\Delta Y_{jj}|} \left(Y(j_3) - \frac{Y(j_1) + Y(j_2)}{2} \right), \quad (3.3)$$

was studied and compared for different LO+PS simulations with a fixed order NLO calculation. In the NLO calculation, the third jet is mostly modelled from the matrix element

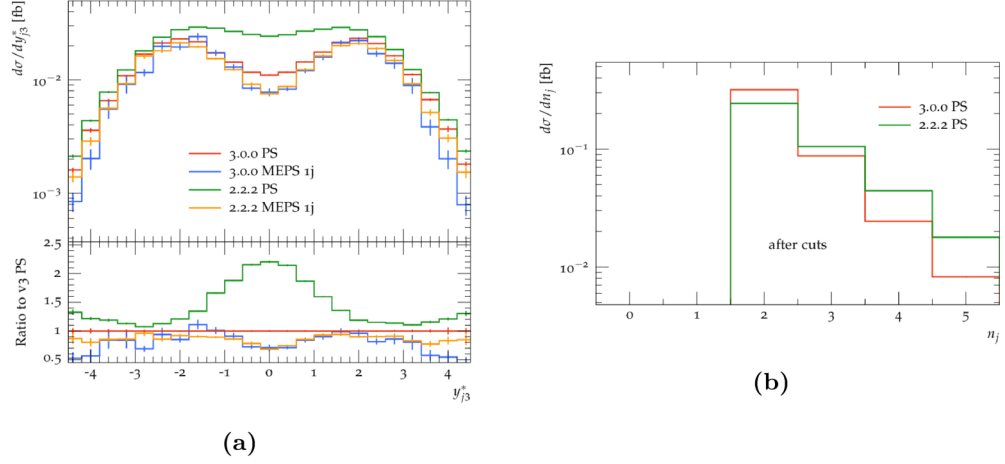


Figure 3.2: Differential cross-sections of the W^+W^+jj -EW6 process in the $e^+\mu^+$ channel as predicted by SHERPA in different settings. The phase-space is a typical phase-space defined for VBS-like processes with a minimum invariant mass of the tagging jets of 500 GeV and minimum absolute difference in rapidities between the tagging jets of 2. The details are listed in [94]. Figure (a) shows the rapidity of the third jet measured with respect to the center of the two tagging jets $Y^* = Y(j_3) - \bar{Y} = (Y(j_1) + Y(j_2))/2$ and (b) the exclusive number of jets with a minimum transverse momentum of 30 GeV. The green and red lines show the LO+PS predictions of the process and the yellow and blue lines represent a prediction where the first emission is also included in the hard process calculation using MEPS@LO merging to the parton shower. Sherpa in version 2.2.2 was used to determine the predictions represented by the yellow and green lines. The predictions in red and blue were determined using a pre-release version of SHERPA 3.0. From: [94].

calculation, while for the LO+PS simulations it is dominantly defined by the parton shower simulation. Large differences between PYTHIA 8 and HERWIG 7 parton showers were observed, where HERWIG 7 was much closer to the NLO fixed order calculation (see Fig. 19 in [90]). A similar comparison study was performed in [91], where different LO+PS setups were compared for the $\ell\ell\nu jj$ final state. Also, significant discrepancies between the parton-shower-based predictions of $z(j_3)$ consistent with [90] were found. Other comparisons to data taken by the CMS collaboration indicated that data favours the HERWIG 7 over the PYTHIA 8 prediction.

This mismodelling can be adjusted and mitigated by switching how the recoil of initial state radiation is distributed. By default, this is distributed to the full event, but better modelling is achieved here if only a single recoiler is picked based on the colour flow of the full event [92]. This can be done by setting `SpaceShower:dipoleRecoil = on` for versions > 8.230 . This setting is incompatible with the MC@NLO style matching. The proposed fix has been studied, e.g. in [93] and was shown to yield predictions more similar to the HERWIG 7 prediction for VBS in the $W^\pm W^\pm$ channel.

A similar issue was also found in SHERPA [94]. The default setting assumed a colour connection between the final state quarks and consequently overestimated the amount of QCD radiation between the tagging jets. This problematic behaviour can be seen in the distribution of the rapidity of the third jet $Y(j_3)$ shown in Figure 3.2. A fix for this mismodelling will be available in version 3.0 of SHERPA, which is not officially available yet and needs to be validated.

The issues in PYTHIA 8 and SHERPA mentioned above are of particular importance for this work. Two simulations for the signal process $WZjj$ -EW6 were prepared in the context of this work and used in [3]. The details of both simulations are discussed below in Section 3.6. They employ independent frameworks for each step of the simulation, but both were affected by the mismodelling of the third jet kinematic. In the light of these mismodellings, a more detailed modelling study for the $\ell\ell\nu jj$ final state is necessary and presented below in Section 3.7.

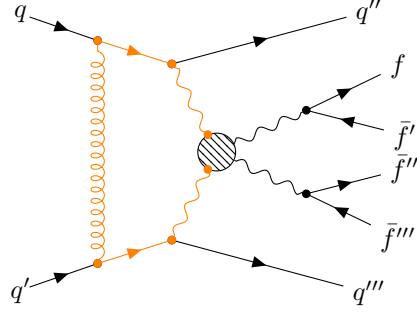


Figure 3.3: Schematic structure of a Feynman diagram for the scattering of electroweak gauge bosons at the LHC with a virtual QCD correction. The intermediate gluon connects the two quark lines and forms a 6-point loop together with the gauge bosons that initiate the boson scattering. The loop is highlighted in orange. The dashed circle acts as a placeholder for any of the possible boson scattering diagrams shown in Figure 2.5.

3.3.1 Issues towards higher accuracy

The observed issues in the parton shower modelling in combination with the expected difference between the kinematics of the third jet in the $VVjj$ -EW6 signal and the $VVjj$ -EW4 background indicate immense benefits from including the third jet in the hard process definition. This increase in accuracy is expected to yield more reliable predictions, which could then be studied in more detail and exploited for better suppression of the $VVjj$ -EW4 background.

The two available approaches for including the third jet in the matrix element are to either calculate the full process at NLO or to merge a multi-leg simulation of the two- and three-jet final states. However, different issues arise when trying to increase the accuracy of the hard-process calculation.

Dimensionality of phase space: Feynman diagrams with one additionally emitted jet have seven particles in the final state, leading to a high dimensionality of the phase space to be integrated. The numerical optimization of the integration is complicated and time-consuming and on the edge of what is technically feasible. While such simulations are available in the $W^\pm W^\pm$ channel, it is even more challenging for the $W^\pm Z$ channel. The asymmetry in the masses and decays of the W^\pm and Z bosons implies fewer symmetries in Feynman diagrams such that less of the matrix elements can be re-used during the simulation of the hard process. In practice, this results in increased time for the integration and optimization of the phase space. An increase by a factor of about 4 for the $W^\pm Z$ channel was found, when compared to the $W^\pm W^\pm$ channel. This increased computing time makes many available setups infeasible for the $W^\pm Z$ channel, that might work technically.

Complicated loop diagrams: When including the full virtual corrections necessary for the NLO calculation, Feynman diagrams would have to be added that show a very complex loop structure. Especially diagrams where a gluon is exchanged between the initial quark lines cause problems. An example is depicted in Figure 3.3. Such diagrams include 6-point loops where one has to integrate over all possible momenta. Often such complicated loops are not included for technical reasons in the simulation.

Since these loops would add a colour connection between the quark lines, a very different kinematic behaviour of consecutive radiations are expected. The typical rapidity gap between the two tagging jets expected at leading order might be reduced when including such loops in the full NLO calculation.

Soft photon divergency: In the typical LO calculation for VBS processes, one restricts the phase space by imposing a minimum on the transverse momentum of the two final state

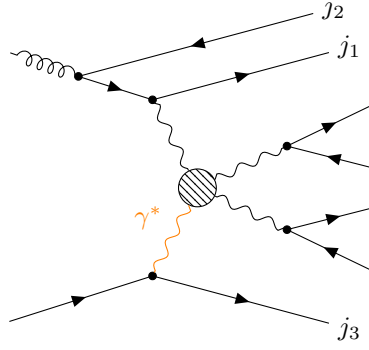


Figure 3.4: Schematic structure of a Feynman diagram for the scattering of electroweak gauge bosons at the LHC with a real emission as QCD correction. The emitted quark can be resolved as extra jet. If the momenta of the jets are such that j_3 has the lowest transverse momentum, no direct cut on $p_T(j_3)$ is applied. In these cases, a photon initiating the boson scattering, as highlighted in orange, is allowed to be soft. The dashed circle acts as a placeholder for any of the possible boson scattering diagrams shown in Figure 2.5.

partons. When integrating the phase space for real emissions as part of an NLO calculation, there is no explicit phase space cut on the emission, and the existing criteria are applied to the two hardest jets in the event.

For VBS topologies that can be initiated by a γ^* , this introduces a divergency originating from soft photons. If a photon is emitted from a quark line and the resulting quark is allowed to have negligible transverse momentum, the photon can be soft and thus introduces a divergence. An example diagram for this is shown in Figure 3.4. This divergency can cause numerical issues and instabilities in the phase space integration and should be avoided.

Possible solutions would be to factorize the process and include the photon in the PDF or to impose a low energy cut-off for the photon virtuality. Theorists typically follow the latter approach in calculations (see, e.g. [95]). Unfortunately, it is not clear how the available MC event generators handle this or whether they have a necessary special treatment in place. For SHERPA and HERWIG 7, no special treatment is applied as confirmed by the authors in private communication.

This problem only occurs when only a photon is emitted from a quark line. Such a photon emission is not possible for the first real emission in the $W^\pm W^\pm$ channel. All other channels of VBS are affected by this. The problem also does not occur when using a LO merged setup, where an explicit cut on the transverse momentum of the extra emission is typically applied.

3.4 Common Approximations

Resonant boson approximation: In order to reduce the complexity of the matrix element and to simplify the calculations, approximations are commonly applied. One group of approximations is applicable for all processes where unstable particles, for instance, electroweak gauge bosons in the case of VBS, are studied. The idea is to split the full process in the production and the decay of said bosons.

The strictest version of this is applied in the narrow-width-approximation, which is only valid for intermediate particles where the width is minimal compared to the mass, which is the case for the W^\pm and Z bosons. In this approximation, the full production process is considered with the bosons in the final state. Subsequently, the decay is simulated independently. In modern implementations, it is possible to include spin correlations and to re-introduce some off-shell effects to model the Breit-Wigner distribution of the particle's decay correctly[96].

The down-side is that other resonances that might be possible in the full process might be neglected. This is the case if the narrow-width-approximation is applied to a Z boson. Any Z boson that does not couple to neutrinos in a Feynman diagram can be replaced by

a virtual photon when adjusting the couplings accordingly. Also, contributions, where the intermediate particle is far off-shell, are usually not modelled correctly.

Another, less strict approximation is to consider only diagrams where a certain number of s -channel propagators of the intermediate particle exist. While this does not restrict the dimension of the phase space, this approximation limits the number of diagrams and consequently simplifies the calculation of the matrix element.

How substantial the effect of such approximations is for a given process significantly depends on the phase space requirements. If the decay products are experimentally accessible, it is possible to focus on the on-shell region by requiring an invariant mass close to the particle's mass. In the $\ell\ell\nu jj$ final state, such invariant mass cut can only directly be applied to the $\ell^+\ell^-$ pair originating from the Z boson decay. Also there, migrations from the significant γ^* contributions are possible if the photon decays to τ leptons which might decay leptonically. Also in the final state with three leptons of the same flavour, migrations can occur due to misassignment to the bosons. When introducing such approximations, one ideally needs to study the effect in the phase space of interest carefully.

One goal of this thesis is to find possible setups to produce reliable MC simulations for VBS processes that could be used as background also for further analyses within ATLAS, e.g. for analyses in the $W^\pm W^\pm$ channel. In such scenarios, a phase space restriction to the resonance mass is not possible. For this reason, such approximations are purposefully not considered from here on.

VBF approximation: Another group of commonly used approximations is specific to the VBS topology of two tagging jets. As mentioned in Section 2.2 the diagrams of the signal process $VVjj$ -EW6 can be grouped in s -, t -, or u -channel according to the quark lines in a gauge invariant way. With this grouping, the full matrix element can be written as

$$|\mathcal{M}_{\text{full}}|^2 \sim |s + t + u|^2, \quad (3.4)$$

where s , t , and u represent the sum of all diagrams of the respective group.

The s -channel diagrams do not include the $VV \rightarrow VV$ scattering and the characteristic signature of the tagging jets. The VBF approximation separates these s -channel diagrams from the signal process and neglects interferences between the different groups. In the full VBF approximation, the matrix element can be calculated as

$$|\mathcal{M}_{\text{VBF}}|^2 \sim |t|^2 + |u|^2. \quad (3.5)$$

This VBF approximation ensures that the final state jets are not colour connected for the signal process at LO.

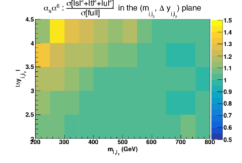
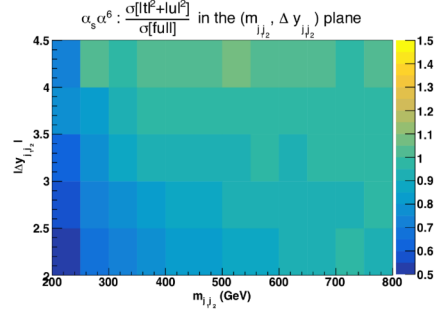
Sometimes the s -channel contributions are generated and added as a separate process, again neglecting the interferences between the different groups

$$|\mathcal{M}_{\text{VBF}+s}|^2 \sim |s|^2 + |t|^2 + |u|^2. \quad (3.6)$$

The effect of this approximation was studied in detail for the $W^\pm W^\pm$ channel at parton level in [90]. At LO this approximation is justified given typical selection criteria to find the tagging jets of a minimum invariant mass $M(jj) > 500$ GeV and a rapidity separation of $\Delta Y(jj) > 2$, with effects of about 1%. When adding a parton shower and detector simulation with pileup, additional jets can be found which can cause migrations of neglected contributions in the phase space of interest.

Such migrations can also be observed when going to NLO accuracy. Figure 3.5 shows the relative effect of the approximations at NLO differentially in the $M(jj)$ and $\Delta Y(jj)$. For the typical selection criteria mentioned above the effects of this approximation are still about 10 – 20% in some areas of phase space.

While applied commonly and by default in some generators, this approximation potentially yields problematic predictions. Applying sufficiently hard selection criteria to reduce the effect is not always feasible due to limited statistics in data.



(b)

(a)

Figure 3.5: Ratios of an approximated over the full matrix element calculation for double-differential distributions in the invariant tagging jet mass $M(jj)$ and the absolute rapidity difference $|\Delta Y(jj)|$ for the $WZjj$ -EW6 process at NLO QCD. The approximated matrix element calculation employ the VBF approximation (left) or the VBF approximation including the s -channel contributions as separate processes (right), respectively. From: [90].

3.5 Process-specific Setups

In order to find suitable setups for the simulation of events for VBS processes with wide applicability, combinations of the generators and approximations mentioned above are discussed. Setups are considered specific for the signal processes $WZjj$ -EW6, its irreducible background $WZjj$ -EW4 and the interference between the two $WZjj$ -EW5. The main motivation is to find setups that include a simulation of the third jet kinematic in the hard-process definition in order not to rely on parton shower simulations for this. In general, strict fixed-order calculations are not included, since the application of a parton shower is a crucial component for comparisons to experimental data.

During the technical tests of available setups, several generator specific caveats have been found. The configurations and some caveats are presented in Appendix B.1.

3.5.1 $WZjj$ -EW6

For the VBS signal process different levels of accuracy are discussed.

Leading order and parton shower ($2j@LO+PS$): The baseline for further comparisons is the matrix element calculation of the full process at LO combined with parton shower simulations. Many setups are available for such simulations, so only a selection is presented. Several other setups are available, which apply the VBF approximation. Among the latter are setups based on VBFNLO matrix elements, which includes for most boson channels the NLO implementations in POWHEGBOX. A study of different LO+PS predictions for the $\ell\ell\nu jj$ final state was presented in [91].

- SHERPA + SHERPA. SHERPA is capable of calculating the full VBS signal process at LO accuracy for all boson channels. As discussed above, this setup suffers from the mismodelling of the third-jet kinematic. This issue is expected to be fixed in SHERPA 3.0.

This setup was used for a large central simulation in the ATLAS collaboration, so it is available with immense sample size, including simulation of the detector. It was used as nominal sample in [3] and is used throughout this work.

- MG5_AMC+PYTHIA 8. The simulation of the hard process at LO without approximations is possible in MG5_AMC for all boson channels. The default shower modelling of PYTHIA 8 suffers from the mismodelling of the kinematic of the third jet, as discussed above.

This setup with default options was used for central production in the ATLAS collaboration and is available including simulation of the detector. It is used to estimate modelling uncertainties in [3].

In recent versions of PYTHIA 8, an option to mitigate the effect of the mismodelling was introduced. A simulation using this option was produced privately with limited statistics for this work, including the simulation of the detector. This setup is expected to reproduce the third jet kinematic better than the official samples. The short notation for this setup is adjusted to MG5PY8FIX to indicate the application of this fix.

Leading order merged and parton shower ($2, 3j@LO+PS$): The least computationally expensive way to include the third jet in the hard process simulation is to merge the $2j$ and $3j$ fixed order calculations with a parton shower. This setup avoids the complicated loops and depending on the implementation of the merging, the problem of soft photon divergencies.

- MG5_AMC + PYTHIA 8. This setup employs the CKKW-L merging scheme in the implementation in PYTHIA 8. A technical problem in the merging implementation in PYTHIA 8 versions available at the time was discovered in the context of this work. This problem affects processes with jets in the born process, such as VBS. In the available

version of PYTHIA 8, the different settings to define the hard process with jets were not successful.

A fix is available in recent versions of PYTHIA 8 and once validated, this setup is available in all boson channels.

- **SHERPA + SHERPA.** Merging is easily possible in SHERPA, but depending on the boson channel, the phase space integration is very CPU intensive. Stability issues were observed in earlier versions but could not be reproduced in dedicated tests. In the available versions, a problem in the merging algorithm specific to VBF topologies caused biases in the normalization in this setup [94].

For this work, the integration stability was tested with version 2.2.4 in the $\ell\ell\nu jj$ final state, and while being very CPU intensive, the integration was found to be stable. Due to the necessary amount of CPU time in combination with the known normalization issues, a full production has not been pursued.

In the upcoming version 3 of the SHERPA generator, the issues in the merging are expected to be fixed, and this merged setup is available in all boson channels.

- **MATCHBOX(MG5_AMC) + HERWIG 7.** In this setup, MG5_AMC is only used to provide matrix elements, and the hard-process simulation, including merging to the parton shower is done in the MATCHBOX framework in HERWIG 7.

In available versions of HERWIG 7, systematic variations affecting the hard process can not be evaluated on-the-fly. For this reason, this setup should only be used for validation purposes. In centrally installed versions of the HERWIG 7 generator, the interface to MG5_AMC did not work in time for this work, so this setup was not successfully tested during this work. A private installation was not pursued as it would not allow for subsequent central production of events, and the evaluation of theory uncertainties would be impractical without on-the-fly weights.

Once a version is available where the evaluation of systematic uncertainties is possible on-the-fly, the technical issues should be worked on to make this setup available.

All potential setups have issues in the currently available versions. However, all of these known issues should be fixed in upcoming versions. Since multiple setups are foreseen, this is a good baseline of accuracy for future comparisons in all boson channels. Unfortunately, none of the setups is available in time for this work.

NLO in VBF approximation and parton shower ($2j@$ NLO VBF + PS): A compromise in order to achieve the step to a full NLO calculation is to approximate the signal process using the VBF approximation. In this approximation, manually optimized matrix elements for the $t+u$ channel process in VBF approximation and the s -channel contributions are available separately for all boson channels in VBFNLO. The optimization drastically reduces the CPU time required to simulate the hard process. However, VBFNLO can not generate events at NLO, so its matrix elements need to be processed by another event generator.

- **MATCHBOX(VBFNLO) + HERWIG 7.** The MATCHBOX framework can integrate VBF"-NLO as an external matrix-element provider at tree- and loop-level. Event generation can thus be performed within HERWIG 7 at LO or NLO. While VBFNLO in principle includes the s channel contributions, they are not available within MATCHBOX. The authors have been informed about this and may include an option to include these contributions in future versions of HERWIG 7. Another problem is the lack of a possibility to evaluate systematic uncertainties efficiently, as mentioned above.

For validation purposes the $\ell\ell\nu jj$ final state in this setup was simulated both at LO + PS and NLO + PS in the VBF approximation. More details on the simulation can be found in Appendix B.1.

- POWHEGBOX + HERWIG 7 or POWHEGBOX + PYTHIA 8. The POWHEGBOX generator includes the VBS process in all boson channels in the leptonic decay modes at NLO in QCD. The details are process specific, but for most boson channels and most importantly for this work in the $W^\pm Z$ channel the implemented processes are based on the VBFNLO matrix elements and consequently also employ the VBF approximation. Similar as for the HERWIG 7 generator, the s channel contributions are not included. No further effort to implement this contribution was foreseen by the authors when asked about this.

POWHEGBOX allows interfacing to either HERWIG 7 or PYTHIA 8 for simulation of the parton shower. The combination to either of the two shower generators are available, and the other can be used for validation purposes.

The implementation of VBS in the $\ell\ell\nu jj$ final state at NLO became available after this modelling study was performed[97]. The possibility to evaluate theory uncertainties via on-the-fly weight variations is the only identified advantage over the existing setup based on MATCHBOX, so no simulation in this setup was performed.

NLO (only factorisable corrections) and parton shower ($2j@NLO$ (fact. only) + PS):

In order to avoid many-point loops mentioned in Section 3.3.1, MG5_AMC applies an approximation to exclude non-factorisable QCD corrections. That means for that gluon exchange between the different quark lines is excluded. Details on this can be found in [90]. In the full calculation, both the virtual and real emission diagrams typically include divergencies that cancel when adding both contributions. Neglecting unknown terms introduces the potential of uncanceled divergencies in the real-emission terms. In order to perform such simulations, an MG5_AMC internal pole-cancellation check has to be turned off. This setting, however, may introduce a bias in the simulation that can only be evaluated by comparing to a full NLO calculation.

- MG5_AMC + HERWIG 7. The MG5_AMC authors have performed the integration using MG5_AMC in [90] for the $W^\pm W^\pm$ channel for a particular lepton combination. However, the integration seems still unstable and could not be reproduced fully in the $W^\pm W^\pm$ channel for [93] but relied on a resonant boson approximation. In the context of this work, ultimately unsuccessful tests were made to apply the procedure in the $W^\pm Z$ channel.

In case the technical issues can be overcome, PYTHIA 8 is not available for the parton shower simulation, because the option to improve the modelling is not compatible with MC@NLO style matching applied in MG5_AMC. The parton shower should be simulated in HERWIG 7 to avoid the known mismodelling of the kinematics in the parton shower.

- MATCHBOX(MG5_AMC) + HERWIG 7. This setup is very similar to the setup above. Both setups use matrix elements from MG5_AMC and the parton shower from HERWIG 7. The difference is which framework handles the integration of the phase space and the event generation, which has important technical differences but should result in the same physics. Here, this is handled in the MATCHBOX framework in HERWIG 7.

The integration in MATCHBOX can be parallelized easier and in a more reliable way. Tests have been performed for this work where the event generation failed, indicating a not sufficiently optimized phase space. This problem might be caused by the divergencies mentioned in Section 3.3.1 since no special treatment for these issues is available. In order to overcome these issues, the approximation of the MG5_AMC matrix elements needs to be avoided, and a procedure for the soft-photon-induced divergency needs to be implemented. The later might be available in future versions of HERWIG 7.

This level of accuracy seems to be beyond technical feasibility. The phase space integration seems too unstable and additional theoretical problems can not be addressed.

Full NLO and parton shower ($2j@NLO + PS$): One of the potential issues was caused by the approximation applied by the MG5_AMC matrix elements. This issue could be addressed by using a different matrix-element provider at loop-level. OPENLOOPS 2 includes such matrix elements for some VBS processes including the $W^\pm Z$ channel in recent versions. Such a combination allows for following potential setups

- MATCHBOX(MG5_AMC + OPENLOOPS) + HERWIG 7. Here only the tree-level matrix elements are used from MG5_AMC and the loop-level matrix elements are provided by OPENLOOPS. The manually-optimized matrix elements from OPENLOOPS include all loop corrections but are more efficient compared to MG5_AMC. So the computation of the integral might be better in this setup compared to the case with loop-level matrix elements from MG5_AMC.

At the moment, the interface between OPENLOOPS and HERWIG 7 has not been fully validated, and further tests are needed. The central installation of OpenLoops and HERWIG 7 on CERNs resources unfortunately only provides OpenLoops versions that do not include the $W^\pm Z$ channel so the simulation could not be tested in the scope of this work.

- SHERPA(OPENLOOPS) + SHERPA. The OPENLOOPS matrix elements can also be used in SHERPA. For boson channels where the $2, 3j@LO + PS$ setup works with reasonable CPU resources, switching to this NLO setup using OPENLOOPS might be a good option for improvement. For the rather complicated $W^\pm Z$ channel this is assumed to be infeasible.

For stable integrations, the issue of soft photons needs to be addressed. Only once this is implemented, calculations at NLO without the VBF approximation are available in boson channels other than $W^\pm W^\pm$.

Recent advancements: While independent NLO QCD and NLO EW corrections have been presented before, a recent study presented fixed-order NLO (QCD + EW) corrections to the $WZjj$ -EW6 process [98] using two independent implementations. The phase space used in this study is modelled after [38]. While the results can not be directly compared due to the different phase space definitions and the lack of a parton shower simulation, it is instructive to consider the results.

When including NLO QCD corrections, the total cross-section is mildly reduced (-1.8%) which is well covered by the typical scale variations. Including only the NLO EW corrections reduces the cross-section by about 16% overall and about 30% in specific regions of phase space. The most considerable effects occur in high energy tails of distributions where Sudakov logarithms enhance NLO EW corrections. When combining NLO QCD and NLO EW corrections, the overall effect is about 18% . In most regions of phase space, the full NLO calculation is within the significant scale uncertainties of the LO prediction. However, especially where the NLO EW corrections are substantial, these seem not sufficient to cover the full uncertainties due to higher-order corrections.

Since especially the NLO EW corrections have a significant effect on the shape of distributions, a proper combination with a parton shower would be highly desirable in order to allow comparisons to experimental data. For the $W^\pm W^\pm$ channel, such combination recently became available in the POWHEGBOX framework using loop matrix elements from RECOLA [99]. An extension of this to other final states will significantly improve the accuracy of simulations and facilitate more reliable comparisons to data.

3.5.2 $WZjj$ -EW4

The requirements for the $WZjj$ -EW4 process are drastically different compared to the signal process. The final state with two jets can also be interpreted as extra emissions to the diboson process. These lower jet multiplicity final states also contribute to a detector level phase space with a requirement for two jets via additional radiations from the parton shower

or via pileup. Consequently, they need to be included in the hard process definition as well. Since the goal is to include the third jet in the hard process definition, only setups that allow merging of higher jet multiplicity final states are available. Also due to the standard colour-flow topology of the $VVjj$ -EW4 final state, all of the usual shower generators are available. Since multiple setups are available in principle, the focus is set on the $W^\pm Z$ channel here.

Sherpa: The nominal setup used in the ATLAS collaboration for such processes uses the SHERPA generator in version 2.2.2. In the $\ell\ell\nu jj$ final state, the achieved accuracy is the base process of diboson production $pp \rightarrow \ell\ell\nu$ with additional radiations of $0, 1j$ @NLO and $2, 3j$ @LO. This centrally produced dataset is the nominal simulation for $WZjj$ -EW4 used throughout this thesis.

In order to enhance the amount of generated events in phase spaces with the requirement of additional jets, enhancement factors (see Section 3.2.1) have been used for the higher jet multiplicities. The final state with one additional jet is enhanced by a factor of three and the higher jet multiplicities are enhanced by ten.

MG5_aMC + Pythia 8: An alternative setup is to use CKKW-L merging for the combination of the diboson production process with additional radiation with LO accuracy. The MG5_aMC implementation of this requires the usage of PYTHIA 8 for the parton shower implementation.

For this work, a dataset was produced that combines the final state $\ell\ell\nu$ with additional radiations of up to three additional partons in the matrix element calculation. Due to the lack of enhancement factors compared to Sherpa, the fraction of events that pass all selection criteria is considerably smaller. For this private production, a detector simulation of sufficient events was not feasible, so that this sample is only available at truth-level. A setup with this accuracy is equivalent to the nominal simulation for this process used in the experimental study by the CMS collaboration [38].

Matchbox(MG5_aMC) + Herwig: A more flexible merging implementation is available in the MATCHBOX framework, which allows similar setups as in SHERPA. For this work, several levels of accuracy were tested, but the tests were not successful due to technical issues. The errors were reported to the authors, so in future versions of HERWIG 7, it will be worth reconsidering this setup.

3.5.3 $WZjj$ -EW5

For studies of the interference of the signal and background process, a prediction of the interference contributions $WZjj$ -EW5 is necessary. The interference effects can be estimated based on an inclusive sample, where diagrams of both processes are included in the process definition. If simulations for the pure signal and background processes are subtracted, the resulting distribution corresponds to the interference contributions, which might be negative. Since the interference terms are typically small, compared to the sum of signal and background processes, this procedure results in large statistical uncertainties.

In order to circumvent this, MG5_aMC can restrict the process not on the level of matrix elements, but of squared matrix elements. Using such a restriction, the coupling order for interference terms in the squared matrix element can be specified directly. This procedure is much more efficient in estimating the interference term. Still, the calculation for this process is more complicated compared to either the signal or background process since all matrix elements of both processes have to be considered.

For $WZjj$ -EW5, the coupling order in the squared matrix element was set to $\alpha_{EW}^5 \alpha_s$. With this restriction, a dataset was simulated for this work at LO using MG5_aMC + PYTHIA 8. Details of the configuration can be found in Appendix B.1.

3.6 Overview of Simulations

In the previous section, different setups were compared, and many of them were used to produce simulations for validation purposes. An overview of the produced simulations for different processes with the $\ell\ell\nu jj$ final state is presented in Table 3.1. While simulations for the $WZjj$ -EW6 process were obtained with NLO accuracy when using the VBF approximation, for the full process, only samples with LO+PS accuracy were produced successfully.

Simulations for other background processes: For other backgrounds, the same simulations were used as in [3]. $ZZjj$ – EW4 and $ZZjj$ – EW6 processes were simulated with a similar setup for the respective processes in the $\ell\ell\nu jj$ final state using SHERPA version 2.2.2. For gluon-induced production of a ZZ diboson pair as well as triboson production VVV processes, SHERPA 2.1.1 was used with the CT10 PDF set. MG5_AMC with the NNPDF3.0 PDF set and interfaced to PYTHIA 8 for simulation of the parton shower, was used to simulate the production processes of a top-quark pair in association with gauge bosons $t\bar{t}V$ at NLO accuracy as well as the associated production of a single top quark and a Z boson at LO accuracy.

Additionally, for the study of non-prompt background contributions, POWHEGBOX interfaced with PYTHIA 8 is used with the CT10 PDF set in the matrix element calculations to simulate top-quark pair production $t\bar{t}$. Production of a massive gauge boson and a photon $V\gamma$ was simulated for up to one additional parton at NLO and for up to 3 additional partons at LO using SHERPA version 2.2.2 using the NNPDF3.0 PDF set. For the production of a single gauge boson in association with jets V + jets at NLO accuracy POWHEGBOX was used for the simulation with the CT10 PDF set and interfaced to PYTHIA 8 for simulation of the parton shower, overlap with $V\gamma$ was removed by vetoing events which contain a photon at truth level with a transverse momentum of at least 10 GeV. The full list of simulations used for the study at reconstruction level is shown in Appendix B.2.

For simulations that are compared to measured data, a simulation of the detector based on GEANT4[100] is performed. Multiple overlaid proton-proton collisions are simulated with the soft QCD processes of PYTHIA 8 in version 8.186 using the A2 set of tuned parameters[101] and the MSTW2008LO PDF set.

3.7 Modelling Studies

3.7.1 Phase space definition

In order to compare the different setups, a common set of phase spaces (PS) is defined. The fiducial phase space is modelled after the measurement in [3]. The exact definition is described below and summarized in Tables 3.2.

A necessary prerequisite for the definition of phase spaces is a clear definition of object candidates. Electron e , muon μ , tauon τ , and neutrino ν candidates are selected from the container of truth particles by requiring a compatible PDG ID. Potential particles originating from the detector simulation are vetoed using the barcode, a unique identifier of each particle in the event. Furthermore, the status is required to be equal to 1, vetoing instable or intermediate particles, e.g. before the shower. Non-prompt leptons, e.g. leptons from hadronic decays, are vetoed using a centrally provided algorithm called `MCTruthClassifier` based on the origins of truth particles.

In order to be less sensitive on the modelling of Bremsstrahlung, the four-momenta of electrons and muons are dressed. In this procedure, all photons in a cone of $\Delta R = 0.1$ around each candidate are collected, and their four-momenta are added to the candidate's four-momentum. This correction assumes the candidate emitted the collected photons.

Jet candidates are clustered using the AntiKt algorithm with a radius parameter $R = 0.4$ from truth particles with status 1, where leptonic decay products of W^\pm and Z bosons are

Table 3.1: Overview of available simulations for the different processes in the $\ell\ell\nu j$ final state at particle-level. A label as used in comparison plots is introduced for each sample in the first column. A short description of the applied generator setup is provided, where MB is used as an abbreviation for the MATCHBOX framework in HERWIG 7. Setups affected by the mismodelling of the kinematic of the third jet are marked with †. For each sample, the accuracy, the number of simulated events at particle-level (column “Truth”) and after full detector simulation and reconstruction (column “Reconstr.”), and potential comments are listed. Samples, where the detector simulation and reconstruction were performed privately, are marked with *.

Label	Generator Setup	Accuracy	Number of events		comment
			Truth	Reconstr.	
EW6 S2	SHERPA 2.2.2†	LO+PS	1471000	1471000	Central ATLAS sample (364284)
EW6 MG5Py8	MG5_AMC + PYTHIA 8 †	LO+PS	500000	500000	Central ATLAS sample (364499)
EW6 MG5Py8Fix	MG5_AMC + PYTHIA 8	LO+PS	1000000	159000*	
EW6 MG5H7	MG5_AMC + HERWIG 7	LO+PS	989712	159000*	
EW6 WhPy8	WHIZARD PYTHIA 8	LO+PS	100296	0	
EW6 VBFNLOPy8 tu	VBFNLO + PYTHIA 8	LO(VBF)+PS	174600	0	
EW6 VBFNLOPy8 stu	VBFNLO + PYTHIA 8	LO(VBF+s)+PS	420972	0	with s channel as separate process
EW6 H7-VBFNLO	MB(VBFNLO) + HERWIG 7	LO(VBF)+PS	1474600	0	
EW6 H7-VBFNLO	MB(VBFNLO) + HERWIG 7	NLO(VBF)+PS	84003	0	
tZj MG5H7	MG5_AMC + HERWIG 7	LO+PS	989712	159000*	
EW5 MG5Py8	MG5_AMC + PYTHIA 8	LO+PS	988329	0	
EW4 S2	SHERPA 2.2.2	(0,1j@NLO, 2,3j@LO)+PS	15645900	15645900	Central ATLAS sample (364253)
EW4 MG5Py8	MG5_AMC + PYTHIA 8	(0,1,2,3j@LO)+PS	985000	0	

Table 3.2: Overview of object selection criteria for the different leptonic object types at truth level. For the absolute value of the PDG ID, potentially multiple allowed values are shown. The last row indicates whether a dressing of the candidate’s four-momentum is applied.

Object selection criterion	electron e	muon μ	tauon τ	neutrino ν
PDG ID	11	13	15	12, 14, 16
Barcode	< 200000	< 200000	< 200000	< 200000
Status	1	1		1
MCTruthClassifierType	= IsoElectron	= IsoMuon	= IsoTau	= Neutrino
Dressing	✓	✓		

excluded. Jets are required to fulfil

$$p_T(j) > 25 \text{ GeV} \quad |Y(j)| < 4.5. \quad (3.7)$$

Furthermore, jets that are close to either an electron or muon candidate with $\Delta R < 0.3$ are vetoed. The two jets with highest p_T in the event are selected as tagging jets j_1 and j_2 ¹.

Events have to pass a set of preselections to ensure compatibility with the required final state $\ell\ell\nu jj$. For this, events are vetoed if any τ candidate is present and from here on only e and μ candidates are referred to as lepton candidates. Each event is required to have precisely three lepton candidates, and at least one neutrino and two jets.

The lepton candidates have to be assigned to the W^\pm and Z bosons in order to define the final phase spaces. Each combination of the three leading lepton candidates is tested in combination with the leading neutrino in the event. Only those assignments where the flavour and charge combination is compatible with the leptonic decays of one W^\pm and one Z boson are considered. If no compatible assignment is found, the event is vetoed. For each allowed assignment a score is calculated, measuring the compatibility of the invariant masses of the assigned decay products for each boson with the bosons’ nominal masses based on Breit-Wigner distributions. The assignment with the higher score is used. This assignment algorithm is commonly referred to as resonant shape algorithm.

The lepton assigned to the W^\pm boson is referred to as ℓ_W , and the leptons assigned to the Z boson, sorted by p_T are labelled $\ell_{Z,1}$ and $\ell_{Z,2}$, respectively. Events are classified according to the assignment in four channels; eee , μee , $\mu\mu e$, and $\mu\mu\mu$, where the flavour is indicated in the order $\ell_W\ell_{Z,1}\ell_{Z,2}$.

Based on these candidates and this assignment, three phase spaces are defined. These are strict subsets of each other. The definitions are described in the following and summarized in Table 3.3.

Loose phase space: This phase space is modelled to be close to the detector acceptance and to reduce γ^* contributions. While simulations of events typically require generator level restrictions, these were chosen to be looser than this phase space, including margins to allow for migrations due to, e.g. the parton shower simulation.

All assigned lepton candidates have to fulfil

$$p_T(\ell) > 15 \text{ GeV}, \quad |\eta(\ell)| < 2.5, \quad (3.8)$$

and for ℓ_W the transverse mass requirement is increased to $p_T(\ell_W) > 20 \text{ GeV}$. The invariant mass of the Z boson candidate $M(\ell_{Z,1}\ell_{Z,2})$ is required to fulfil $66 \text{ GeV} < M(\ell_{Z,1}\ell_{Z,2}) < 116 \text{ GeV}$. The leptons assigned to the Z boson have to be separated from each other by $\Delta R(\ell_{Z,1}, \ell_{Z,2}) > 0.2$ and from ℓ_W by $\Delta R(\ell_Z, \ell_W) > 0.3$. Both tagging jets are required to have a transverse momentum of $p_T(j) > 40 \text{ GeV}$.

¹This selection of tagging jets is chosen differently than in [3], where the second tagging jet is the leading jet from the set of jets in the opposite hemisphere of the detector. This assignment strategy is not invariant against boosts along the z axis and introduces a dependence on the laboratory frame. For this reason, here rather the commonly used definition of tagging jets is used.

Table 3.3: Overview of event selection criteria for the different phase spaces at truth level.

Event selection criterion	Loose PS	$WZjj$ PS	Fiducial PS
τ veto	✓	✓	✓
$p_T(\ell_Z) > 15$ GeV	✓	✓	✓
$p_T(\ell_W) > 20$ GeV	✓	✓	✓
$p_T(j) > 40$ GeV	✓	✓	✓
$ \eta(\ell) < 2.5$	✓	✓	✓
$ M(\ell\ell) - 91 \text{ GeV} < 25 \text{ GeV}$	✓	✓	✓
$\Delta R(\ell_{Z,1}, \ell_{Z,2}) > 0.2$	✓	✓	✓
$\Delta R(\ell_Z, \ell_W) > 0.3$	✓	✓	✓
$M_T(W) > 30 \text{ GeV}$		✓	✓
$ M(\ell\ell) - 91.1876 \text{ GeV} < 10 \text{ GeV}$		✓	✓
$M(jj) > 150 \text{ GeV}$		✓	✓
$\Delta Y(jj) > 2$			✓
$M(jj) > 500 \text{ GeV}$			✓

$WZjj$ phase space: Based on the Loose phase space, the $WZjj$ phase space adds requirements to select the typical topology of resonant bosons and to suppress contributions where the tagging jets originate from a hadronic gauge boson decay. The transverse mass of the W^\pm boson $M_T(W)$, defined as

$$M_T(W) = \sqrt{2 \cdot p_T(\nu) \cdot p_T(\ell_W)(1 - \cos \Delta\varphi(\ell_W, \nu))}. \quad (3.9)$$

is required to be larger than 30 GeV and the invariant mass of the Z boson $M(\ell_{Z,1} \ell_{Z,2})$ is required to be in a window of 10 GeV width around $m_Z = 91.1876$ GeV, for the nominal Z boson mass. The invariant mass of the tagging jets $M(jj)$ is required to be larger than 150 GeV.

Fiducial phase space: The $WZjj$ phase space is further restricted to increase the contribution of the signal process in the VBF topology. Events have to fulfil

$$M(jj) > 500 \text{ GeV}, \quad \Delta Y(jj) > 2. \quad (3.10)$$

3.7.2 Modelling of $WZjj$ -EW6

The integrated cross-sections in each of the three phase spaces are listed in Table 3.4 for all available simulations of the $WZjj$ -EW6 process.

Modelling at LO+PS: Several setups are available for LO+PS accuracy, allowing a detailed study of the full process. Comparing the predicted cross-sections in each phase space for the LO+PS predictions, there are considerable differences. As is discussed in Section 8.2 below, the theory uncertainties, especially from the renormalization scale, are vast at about 10 – 20%. Since different dynamic scales are used in the hard-process definitions, these contribute to the overall differences. For the MG5_AMC hard-process calculation, three different parton shower simulations are available and also substantial differences can be seen between them. While the default PYTHIA 8 showering model disagrees significantly from the HERWIG 7 shower simulation, the agreement is much better, when the alternative recoil scheme in PYTHIA 8 is used to avoid the mismodelling of the third jet kinematic.

Table 3.4: Overview of predicted cross-sections in fb for all available simulations of the $WZjj$ -EW6 process in all considered phase spaces at particle-level. The labels for the different simulations follow Table 3.1. The shown uncertainties contain only statistical uncertainties due to the limited number of generated events.

	cross-section [fb]		
	Loose PS	$WZjj$ PS	Fiducial PS
EW6 S2	3.001 ± 0.013	2.069 ± 0.011	1.152 ± 0.008
EW6 MG5Py8	3.170 ± 0.019	2.234 ± 0.016	1.366 ± 0.013
EW6 MG5Py8Fix	3.030 ± 0.014	2.098 ± 0.012	1.346 ± 0.009
EW6 MG5H7	2.965 ± 0.014	2.067 ± 0.012	1.331 ± 0.009
EW6 H7-VBFNLO LO	1.859 ± 0.006	1.464 ± 0.005	0.997 ± 0.004
EW6 H7-VBFNLO NLO	2.326 ± 0.018	1.789 ± 0.016	1.207 ± 0.012
EW6 VBFNLOPy8 tu	2.360 ± 0.012	1.846 ± 0.011	1.273 ± 0.009
EW6 VBFNLOPy8 stu	2.878 ± 0.013	1.994 ± 0.011	1.290 ± 0.009
EW6 WhizardPy8	3.122 ± 0.030	2.144 ± 0.025	1.373 ± 0.020

This behaviour is confirmed in the differential cross-sections in the Loose PS shown in Figure 3.6.² In observables with low sensitivity to the parton shower, e.g. the transverse mass of the diboson system $M_T(WZ)$, defined as

$$M_T(WZ) = \sqrt{2 \cdot p_T(\nu) \cdot \sum_{\ell} \vec{p}_T \left(1 - \cos \Delta\varphi \left(\sum_{\ell} \vec{p}_T, \nu \right) \right)}, \quad (3.11)$$

the overall agreement is good. The MG5_AMC and SHERPA setups deviate slightly probably caused by different scale choices. The WHIZARD setup, in general, agrees better to the MG5_AMC prediction for both the normalization and the shapes of differential predictions. In observables that are sensitive to the shower modelling like the invariant mass of the tagging jets $M(jj)$ large differences can be seen. The SHERPA and default PYTHIA 8 showers with known issues show large differences compared to the HERWIG 7 shower, especially for lower values of $M(jj) < 500$ GeV. With the adjusted recoil scheme in PYTHIA 8, the differences to HERWIG 7 are minor. In the tail of the $M(jj)$ distribution, all MG5_AMC based predictions agree, indicating that the hard process calculation dominates this region. There the SHERPA calculation predicts lower cross-sections, while WHIZARD agrees well with the MG5_AMC prediction. The different choices of dynamical scales could again cause the difference.

When going to more restrictive phase spaces, selection criteria are imposed on observables that are sensitive to the parton shower modelling. While these selections affect the integrated cross-sections, the agreement gets better as the rejected regions of phase space are more sensitive to the parton shower, see, e.g. Figure 3.6b.

Figure 3.7 shows differential cross-section predictions using several setups. Similar to the Loose PS, leptonic observables show excellent agreement. Observables that are sensitive to the parton shower modelling show considerable differences. The number of jets and the pseudo-rapidity of the third jet confirm consistent behaviour compared to earlier studies discussed in Section 3.3. Both SHERPA and the default PYTHIA 8 predict too much activity between the tagging jets, leading to higher predicted numbers of jets and additional third-jet activity for low pseudo-rapidities. The WHIZARD prediction agrees very well with the prediction by MG5_AMC, and both predictions employ the same parton shower simulation.

²There are many comparison plots in the style of the plot shown in Figure 3.6 with varying inlays throughout this work. A detailed description of an example with each type of inlay used throughout this work is given in Appendix C

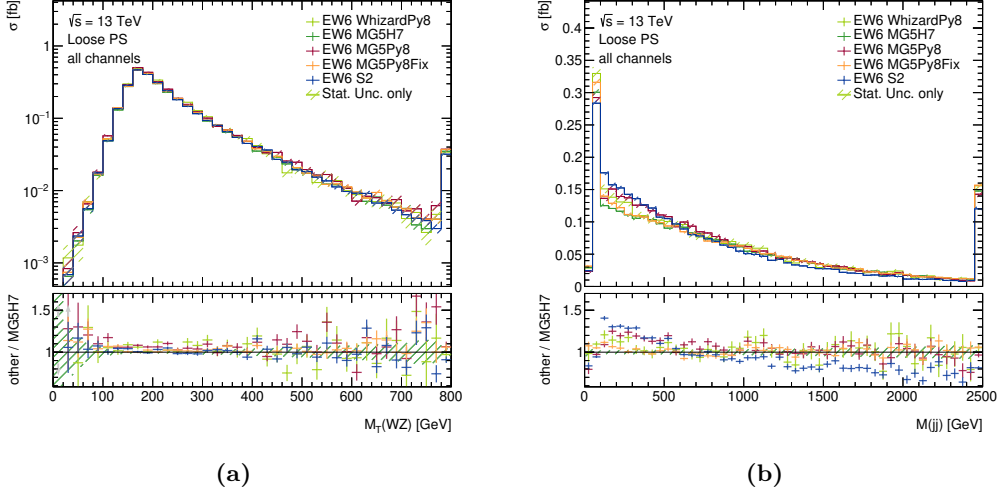


Figure 3.6: Comparison of differential cross-sections for different simulations of the $WZjj$ -EW6 process at LO+PS accuracy in the Loose phase space at particle level. Figure (a) shows the transverse mass of the $W^\pm Z$ system and Figure (b) the invariant mass of the tagging jets. Uncertainties only show statistical uncertainties due to the limited number of simulated events. The lower subplot shows the ratio of other predictions to the MG5H7 prediction.

Effect of VBF approximation: The effect of the VBF approximation as introduced in Section 3.4 at LO can be studied with the available simulations. A comparison of different predictions is shown in Figure 3.8. There the LO prediction for the full process is compared to the full VBF approximation as applied in VBFNLO and when using the VBFNLO matrix elements in the MATCHBOX framework of HERWIG 7.

While the predictions of VBFNLO and MATCHBOX should agree, significant differences, especially on the normalization, can be seen. An origin of such differences is the choice of renormalization and factorization scales. For the HERWIG 7 prediction, a scale based on the scalar sum of transverse masses of final state objects H_T is used. In VBFNLO the momentum transfer of the exchanged boson is used for the simulation with the VBF approximations and for the s -channel contributions the invariant mass of the full final state is used. Another source of differences might be the modelling of the parton shower.

Confirming the studies performed without the simulation of a parton shower in [90], the differences to the prediction of the full process are especially pronounced for low $|\Delta Y(jj)|$ and $M(jj)$. After imposing the requirements of $|\Delta Y(jj)| > 2$ and $M(jj) > 500$ GeV the LO predictions agree well.

In the VBFNLO prediction, also the s -channel contributions can be simulated as a separate process. The agreement to the simulation of the full process can be greatly enhanced to only a few percents by adding this contribution to the prediction for the VBF approximated process alone. This behaviour confirms the studies in [90]. Remaining differences can again be caused by different scale choices or differences in the modelling of the parton shower.

Additionally, the NLO prediction by MATCHBOX is included for comparisons. Sizable differences at the level of 20 – 30% are observed between the LO and NLO predictions. Unfortunately, the s -channel contributions are not available at NLO in this framework at the time of writing. This unavailability forbids a necessary study of the VBF approximation at NLO at the level of the hard process, i.e. without the simulation of the parton shower as done in [90].

An application of a k -factor to correct for additional contributions in the NLO calculation compared to the LO calculation is not applied here. Deriving such factor from the VBFNLO predictions using the VBF approximations is expected to overestimate the missing contributions when applied to a LO calculation of the full process.

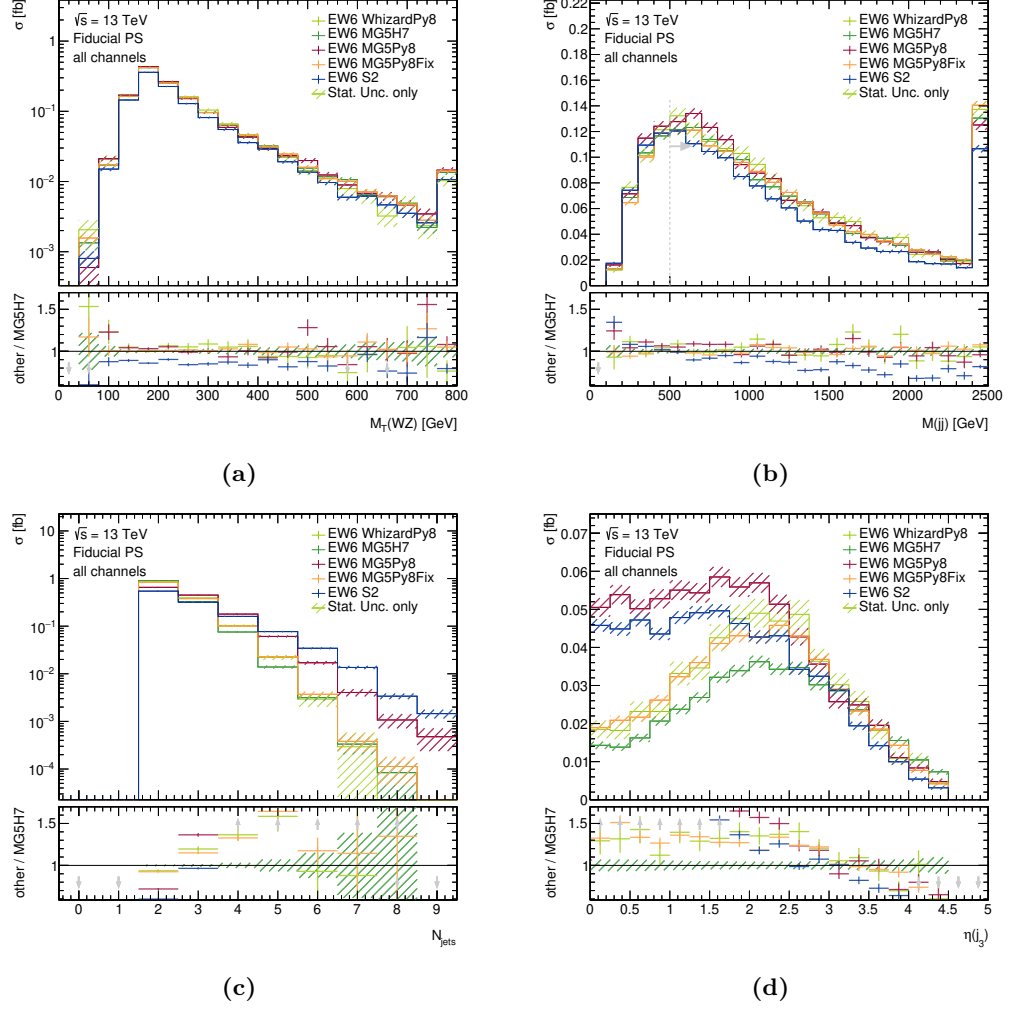


Figure 3.7: Comparison of differential cross-sections for different simulations of the $WZjj$ -EW6 process at LO+PS accuracy in the Fiducial phase space at particle level. Figure (a) shows the transverse mass of the $W^\pm Z$ system, (b) the invariant mass of the tagging jets, (c) the exclusive jet multiplicity, and (d) the absolute pseudo-rapidity of the third jet for events with at least three jets. Figure (b) also shows events, which pass all selection criteria in the Fiducial phase space other than the requirement on the minimum invariant tagging jet mass of $M(jj) > 500$ GeV, which is also indicated using a grey dotted line with the connected arrow. Uncertainties only show statistical uncertainties due to the limited number of simulated events. The lower subplot shows the ratio with respect to the MG5H7 prediction.

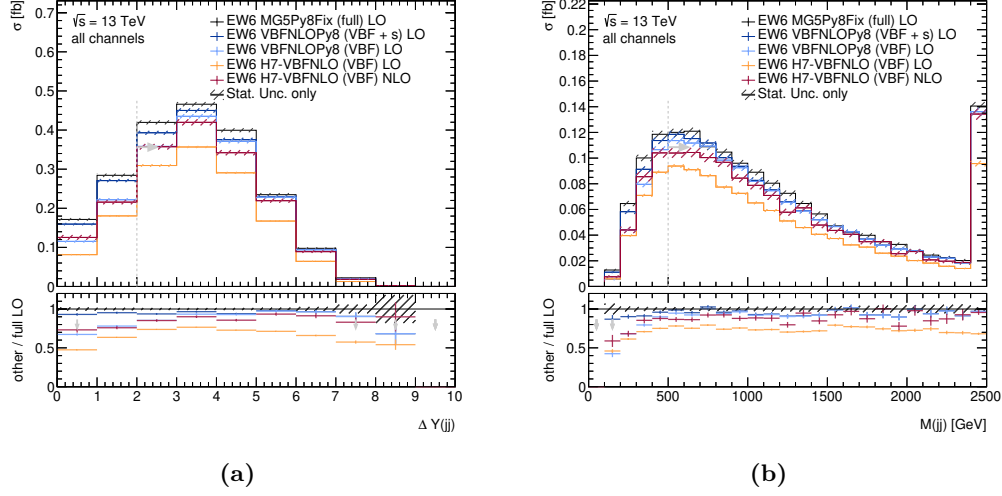


Figure 3.8: Comparison of the effect of different approximations of the full $WZjj$ -EW6 process on differential cross-sections from the $WZjj$ phase space towards the Fiducial phase space. The different approximations and accuracies of the hard-process calculations are indicated in the legend. Figure (a) shows the absolute rapidity separation of the tagging jets in the $WZjj$ phase space and (b) the invariant mass of the tagging jets after the requirement on the absolute rapidity separation of the tagging jets. The vertical grey dotted lines with a connected arrow indicate the imposed requirements in the respective observable towards the Fiducial phase space. Uncertainties only show statistical uncertainties due to the limited number of simulated events. The lower subplot shows the ratio with respect to the MG5H7-based prediction for the full process at LO.

Effect of scale choice: The MG5_AMC authors have implemented several dynamical scales[102] and in the setup used here, all of the dynamical scales are evaluated on-the-fly for each event so that a reweighting is possible. The default scale is the transverse mass of the $2 \rightarrow 2$ system resulting from a k_T clustering. Alternative scales are the total transverse energy of the event $\sum p_T$, the sum of the transverse masses H_T , the sum of the transverse masses divided by 2 $H_T/2$, or the partonic energy $\sqrt{\hat{s}}$. The former two are equivalent for massless final states, as is the case here, so $\sum p_T$ is not considered further. SHERPA uses a dynamical scale called STRICT_METS, where the process is clustered to a $2 \rightarrow 2$ process by inverting the parton shower.

Differential cross-sections for different dynamical scale choices for the MG5Py8Fix dataset are shown in Figure 3.9. For most differential cross-sections a change of the default scale only changes the normalization. For the observables shown in Figure 3.9, however, the shapes are affected as well. The default scale predicts more events at high tagging jet invariant masses $M(jj)$ and absolute rapidity differences $|\Delta Y(jj)|$ which corresponds to the typical VBS topology. Although the scale used in SHERPA is conceptually closest to the default scale in MG5_AMC, the SHERPA prediction is closest to the prediction using the H_T scale in MG5_AMC.

Modelling of tZj : As discussed in Section 2.2, the VBS signal process in the $\ell\ell\nu jj$ final state is further restricted compared to other boson channels due to the large contributions from the production of a single t -quark in association to an Z boson. This subprocess is usually referred to as tZj and contributes significantly due to the enhancement of the t resonance. In the signal process, events with a b -quark in either the initial or final state in the hard process are not included in the signal definition.

In the SHERPA simulations, this criterion was applied as an additional requirement in the hard process, but in the MG5_AMC hard process definition used here the b -quarks are allowed. The MG5_AMC samples can consequently be used to split the events according to the truth record in events that contain a b -quark in the initial or final state (tZj process)

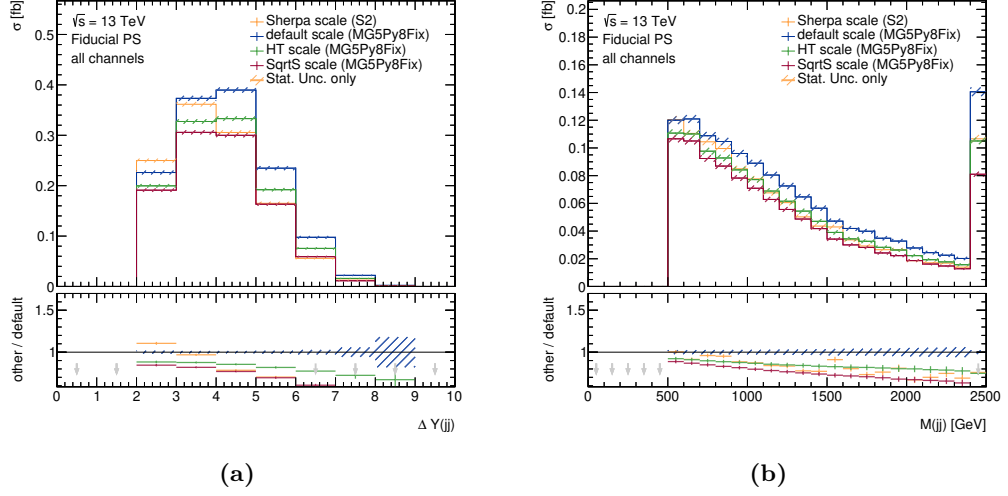


Figure 3.9: Comparison of differential cross-sections with different dynamical scale choices for the full $WZjj$ -EW6 process in the Fiducial phase space. The different scale choices are indicated in the legend. Figure (a) shows the absolute rapidity separation of the tagging jets and (b) the invariant mass of the tagging jets. Uncertainties only show statistical uncertainties due to the limited number of simulated events. The lower subplot shows the ratio with respect to the prediction using the default scale choice in MG5_AMC.

Table 3.5: Overview of predicted cross-sections in fb for simulations of the $WZjj$ -EW4 process in all considered phase spaces at particle-level. The labels for the different simulations follow Table 3.1. The last row also lists the ratio between the two predictions. The shown uncertainties contain only statistical uncertainties due to the limited number of generated events.

	cross-section [fb]		
	Loose PS	$WZjj$ PS	Fiducial PS
EW4 S2	52.01 \pm 0.10	29.65 \pm 0.07	5.97 \pm 0.04
EW4 MG5Py8	33.24 \pm 0.31	19.99 \pm 0.24	3.82 \pm 0.11
MG5Py8 / S2	0.639 \pm 0.008	0.674 \pm 0.010	0.641 \pm 0.023

and those events that do not ($WZjj$ -EW6 signal process). In the simulations shown earlier, only the events without b quark, i.e. of the signal process, were considered. The events with a b -quark are used at detector-level to model the tZj background.

A comparison of differential cross-sections for the two split subprocesses is shown in Figure 3.10. The predictions for the tZj process are less signal-like in typical VBS observables such as the absolute difference in rapidity $|\Delta Y(jj)|$ or the invariant mass of the tagging jets $M(jj)$. This finding confirms the choice to remove these contributions from the signal process definition.

3.7.3 Modelling of $WZjj$ -EW4

The predicted cross-sections for the simulations of the $WZjj$ -EW4 process are listed for the three phase spaces in Table 3.5. There are large differences between the two simulations which are mostly constant for the three phase spaces. A part of the difference is caused by the lack of virtual corrections to the 0 and 1 jet final state in the MG5_AMC calculation. Additional differences originate from the different choices of dynamical scales as discussed for the $WZjj$ -EW6 process. Especially due to the additional contributions, the SHERPA-based simulation is considered more reliable.

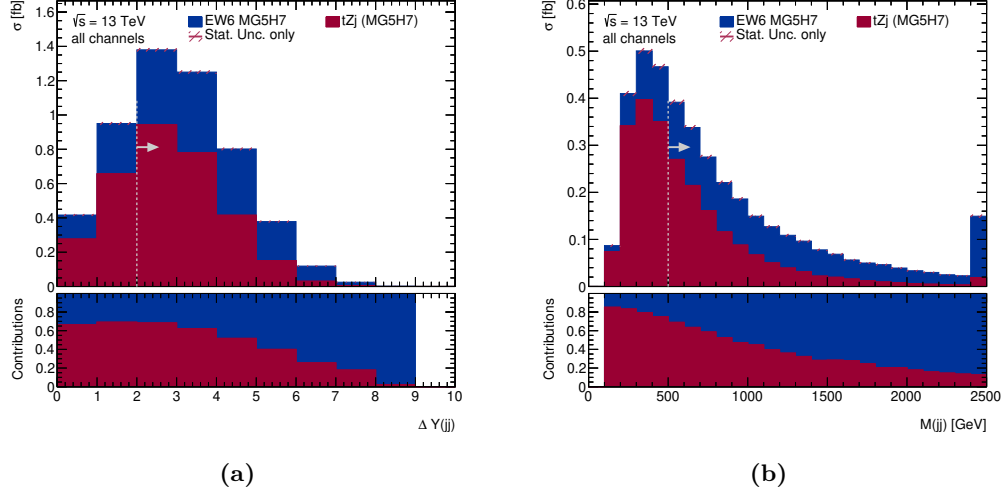


Figure 3.10: Comparison of the differential cross-sections for the tZj and the $WZjj$ -EW6 processes from the $WZjj$ phase space towards the Fiducial phase space. Figure (a) shows the absolute rapidity separation of the tagging jets in the $WZjj$ phase space and (b) the invariant mass of the tagging jets after the requirement on the absolute rapidity separation of the tagging jets. The vertical grey dotted lines with a connected arrow indicate the imposed requirements in the respective observable towards the Fiducial phase space. Uncertainties only show statistical uncertainties due to the limited number of simulated events. The lower subplot shows the relative composition of the sum of the processes.

Comparisons of predicted differential cross-sections are shown in Figure 3.11. For most observables, the predicted cross-sections agree other than the normalization difference. The transverse mass of the diboson system $M_T(WZ)$ and the invariant mass of the tri-lepton system $M(\ell\ell\ell)$ show a tendency of MG5_AMC to favour higher invariant masses. The pseudo-rapidity of the third jet $|Y(j_3)|$, as well as the invariant mass of the tagging jets $M(jj)$ and the relative lepton centrality $\zeta_{\text{rel}}(\ell\ell\ell)$, defined similarly to the zeppenfeld variable as

$$\zeta_{\text{rel}}(\ell\ell\ell) = \max_{\ell} \left| \frac{Y(\ell) - \frac{1}{2}(Y(j_1) + Y(j_2))}{Y(j_1) - Y(j_2)} \right|, \quad (3.12)$$

show very good agreement.

Only observables that are sensitive to the existence of jets in addition to the two required jets show significant differences. This is most obvious in the exclusive number of jets N_{jets} shown in Figure 3.11c. For events with precisely two jets, the predictions agree very well, but for higher jet multiplicities the disagreement is more severe.

Since both samples include the third jet in the matrix element calculation, better agreement, especially for $N_{\text{jets}} = 3$ was expected. The disagreement could be caused by a feature of the MEPS@NLO merging in SHERPA. When lower parton multiplicity subprocesses are available at NLO, as is the case here, a k -factor from LO to NLO is calculated and applied to subprocesses that are only available at LO. Such corrections are not applied in the CKKW-L merging in the MG5_AMC sample, which can explain part of the normalization difference.

Differences in the parton shower modelling most likely cause the mostly constant offset at even higher multiplicities.

3.7.4 Process comparison

The predicted cross-sections for the $WZjj$ -EW4 background process, the $WZjj$ -EW6 signal process, their $WZjj$ -EW5 interference, as well as the sum of all three processes in the three phase spaces are listed in Table 3.6. The $WZjj$ -EW4 process dominates in all phase spaces, due to its less suppressing coupling structure. The interference contribution was found to be positive and non-negligible compared to the $WZjj$ -EW6 signal process. In the Loose

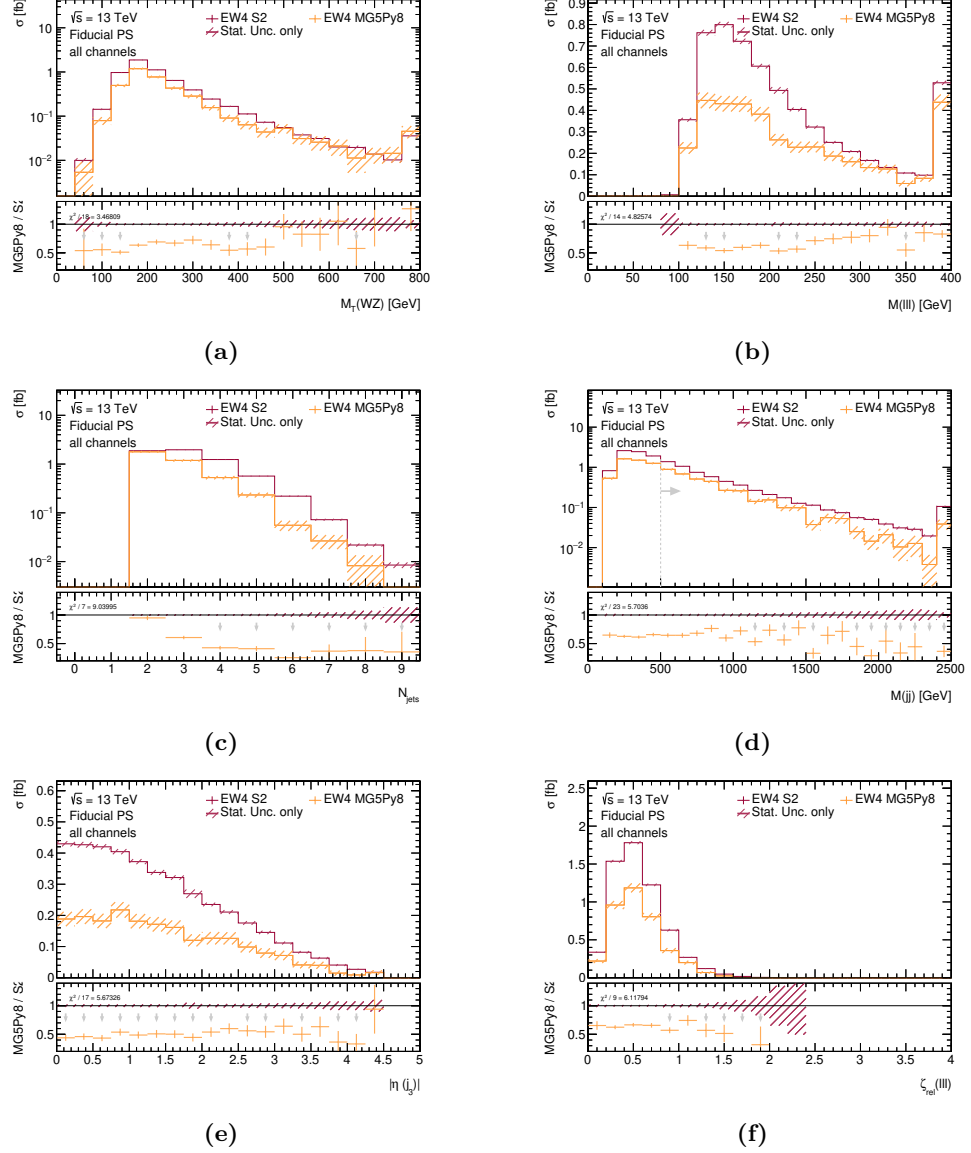


Figure 3.11: Comparison of differential cross-sections predictions for the $WZjj$ -EW4 process in the Fiducial phase space at particle level. Figure (a) shows the transverse mass of the $W^\pm Z$ system, (b) the invariant mass of the tri-lepton system, (c) the exclusive jet multiplicity, (d) the invariant mass of the tagging jets, (e), the absolute pseudo-rapidity of the third jet for events with at least three jets, and (f) the relative lepton centrality $\zeta_{\text{rel}}(\ell\ell\ell)$ as defined in the text. Figure (d) also shows events, which pass all selection criteria in the Fiducial phase space but the requirement on the minimum invariant tagging jet mass of $M(jj) > 500$ GeV, which is also indicated using a grey dotted line with the connected arrow. Uncertainties only show statistical uncertainties due to the limited number of simulated events. The lower subplot shows the ratio with respect to the SHERPA prediction.

Table 3.6: Overview of predicted cross-sections in fb for different processes with the $\ell\ell\nu jj$ final state and their sum in all considered phase spaces at particle-level. The labels for the different simulations follow Table 3.1 and indicate the simulation setup used for the prediction. The last three rows also indicate the relative contributions of each process to the sum in %. The shown uncertainties contain only statistical uncertainties due to the limited number of generated events.

	cross-section [fb]					
	Loose PS		$WZjj$ PS		Fiducial PS	
EW4 S2	52.01	± 0.10	29.65	± 0.07	5.97	± 0.04
EW5 MG5Py8	0.2961 ± 0.0029		0.2718 ± 0.0025		0.0529 ± 0.0011	
EW6 MG5Py8Fix	3.030	± 0.014	2.098	± 0.012	1.346	± 0.009
Sum	55.34	± 0.10	32.02	± 0.08	7.37	± 0.04
Rel Frac. EW4 [%]	93.99	± 0.26	92.60	± 0.33	81.0	± 0.7
Rel Frac. EW5 [%]	0.54	± 0.07	0.85	± 0.09	0.72	± 0.18
Rel Frac. EW6 [%]	5.48	± 0.12	6.55	± 0.16	18.3	± 0.4

PS, the interference amounts to about 10 % of the $WZjj$ -EW6 process. This contribution is reduced to about 4 % in the Fiducial PS. Also, the $WZjj$ -EW4 process is drastically reduced compared to the $WZjj$ -EW6 process in the Fiducial PS.

Especially the additional restriction from the $WZjj$ PS towards the Fiducial PS reduces the signal only by about 35 %, while the $WZjj$ -EW4 and $WZjj$ -EW5 processes are reduced by about 80 %. These selection criteria are $M(jj) > 500$ GeV and $|\Delta Y(jj)| > 2$, and are visualized in Figure 3.12. Both cuts reduce $WZjj$ -EW4 contributions significantly. Even after the $M(jj)$ requirement is already applied, the $WZjj$ -EW5 contribution at low $|\Delta Y(jj)|$ values is larger than the $WZjj$ -EW6 contribution. Since the contribution of $WZjj$ -EW5 is treated as uncertainty of the signal process $WZjj$ -EW6 in the final fit, the goal is to reduce this contribution as much as possible. Before the $|\Delta Y(jj)|$ requirement, the interference contribution still amounts to about 10 % of the $WZjj$ -EW6 process. Only when adding this requirement the relative contribution, and thus the systematic uncertainty on the signal process is reduced to about 4 % in the Fiducial PS. This reduction motivates the addition of this selection criterion compared to [3].

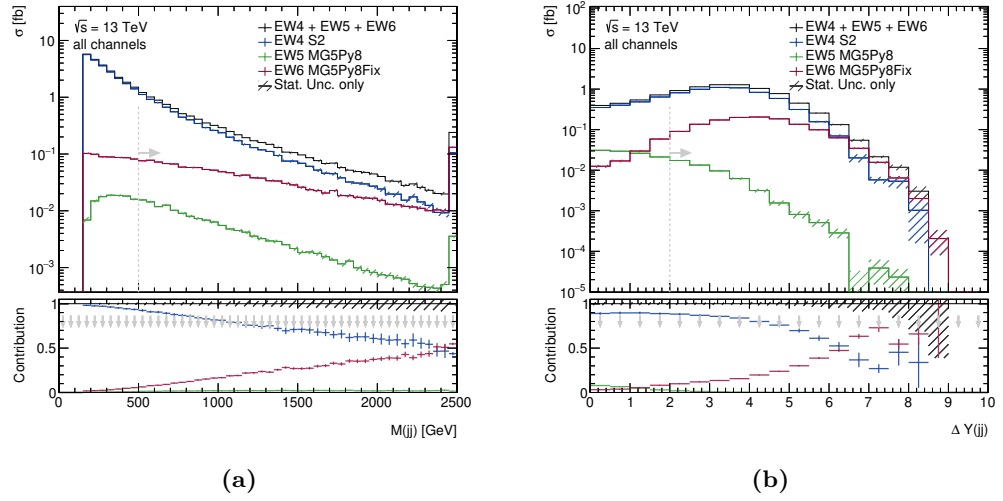


Figure 3.12: Comparison of the differential cross-sections for different processes with the $llvjj$ final state and their sum from the $WZjj$ phase space towards the Fiducial phase space. Figure (a) shows the invariant mass of the tagging jets in the $WZjj$ phase space and (b) the absolute rapidity separation of the tagging jets after the requirement on the invariant mass of the tagging jets. The vertical grey dotted lines with a connected arrow indicate the imposed requirements in the respective observable towards the Fiducial phase space. Uncertainties only show statistical uncertainties due to the limited number of simulated events. The lower subplot shows the relative contributions of each process to the sum of all processes.

Chapter 4

Experiment

The data studied for this work has been collected at the ATLAS detector in 2015 and 2016 during Run 2 of the Large Hadron Collider. While no studies dedicated to the experimental hardware were performed for this work, a basic knowledge of the experimental design and the reconstruction chain is instrumental for a good understanding of the final analysis of the data. The following chapter provides a broad overview. More detailed information can be found in [103, 104].

4.1 The Large Hadron Collider

The Large Hadron Collider (LHC)[103] is the largest particle accelerator in the world, was built from 1998 to 2008 by the European Organisation for Nuclear Research (CERN)¹, and is situated close to Geneva, Switzerland, crossing the border to France.

A primary design objective for the LHC was to be a discovery machine, trying to enable searches for new physics effects in a wide range of potential scenarios. These scenarios motivated the design goal to achieve unprecedented collision energies and luminosities. A matter-matter collider design was chosen to avoid limiting the available luminosity due to a bottleneck in production and preparation of an antimatter beam. The LHC can be operated in different beam modes, colliding either two proton beams, a proton and a beam of Pb ions, or two beams of Pb ions.

The LHC has a circumference of 27 km and consists of 8 straight sections connected by arcs. The straight sections can be used to host experiments or serve utilitarian purposes. A schematic overview of the sectors and their usage can be found in Figure 4.1. The main experiments hosted by the LHC are:

ATLAS (A Toroidal LHC ApparatuS) the largest detector designed for high luminosities and general-purpose physics searches [105]. ATLAS is hosted at Point 1.

CMS (Compact Muon Solenoid) an independently designed detector with similar purpose compared to ATLAS [106]. CMS is hosted at Point 5.

Alice (A Large Ion Collider Experiment) designed to study the properties of quark-gluon-plasma in heavy-ion collisions [107]. ALICE is hosted at Point 2.

LHC-b (Large Hadron Collider beauty) a detector specialized in B -physics to study CP -violation in interactions of bottom quarks [108]. LHC-b is hosted at Point 8.

As a cost-saving measure, the LHC was built in the pre-existing tunnel, which was initially constructed for the Large Electron-Positron (LEP) collider [109]. For the LEP collider, a tunnel layout of comparably long straight lines connected by short arcs was chosen to maximize the distance available for acceleration to counter substantial energy losses due to Bremsstrahlung. Due to the drastically reduced impact of Bremsstrahlung, while deflecting

¹derived from the French name Conseil européen pour la recherche nucléaire

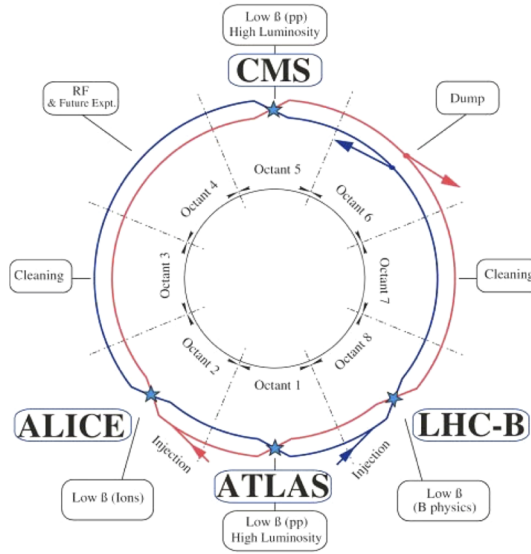


Figure 4.1: Schematic layout of the LHC. From: [103]

protons or Pb ions compared to electrons², a limiting factor for the LHC is the available magnetic field for deflecting the hadron beams given the predefined curvature of the arcs.

Another design challenge imposed by the LEP tunnel was the internal diameter of only a few meters within the tunnel. This spatial constraint prohibited less complicated designs using two pipes with an independent magnet and cooling system for the two beams traversing the tunnel in opposite directions. Instead, a more complicated twin-bore magnet design was necessary to force the anti-parallel beams on their trajectories along the tunnel. The beams have independent vacuum pipes for most of the path and are only crossed at the four interaction points, hosting the main experiments. The beam pipes share a cooling and magnet system, designed to provide anti-parallel magnetic fields in the two beam pipes. The use of superconducting NbTi cables, cooled to 2 K, allowed to reach field strengths of more than 8 T. With this field strength a maximum beam energy for proton beams of 7 TeV could be achieved.

The LHC is part of a large accelerator complex hosted at CERN, and the protons pass a multi-step pre-acceleration phase before being injected into the LHC. A schematic overview of the accelerator complex at CERN [110], including the injection chain is shown in Figure 4.2. More detailed information can be found in [110–112].

The preparation of the beams starts with a bottle of hydrogen. In the first step, some hydrogen is released to a chamber, where an electrical field splits the hydrogen atoms and separates electrons and protons. The protons are passed into a radio frequency quadrupole which groups the protons into bunches and accelerates them to about 750 keV. From there, the proton bunches are accelerated further by a linear accelerator, Linac2, to about 50 MeV. All subsequent increases in beam energy are performed in circular accelerators, specifically synchrotrons. The first is the Proton Synchrotron Booster (BOOSTER), which consists of a stack of four synchrotron rings. Each bunch is directed in one of the rings and accelerated to an energy of 1.4 GeV. The accelerated bunches, now travelling at 91 % of the speed of light, are filled with a well-defined spacing of about 330 ns into the next accelerator in the injection chain, the Proton Synchrotron (PS). In the nominal LHC filling scheme [114], one batch with four bunches and one with two batches are filled in the PS to utilize its capacity better. The PS splits each bunch in twelve new bunches with uniform spacing, resulting in a train of 72 bunches with a bunch spacing of 27.5 ns. After a further acceleration to 25 GeV this bunch train is passed to the Super Proton Synchrotron (SPS). The velocity of the bunches has now

²The radiative losses due to Bremsstrahlung are inversely proportional to the squared mass of the considered particle.

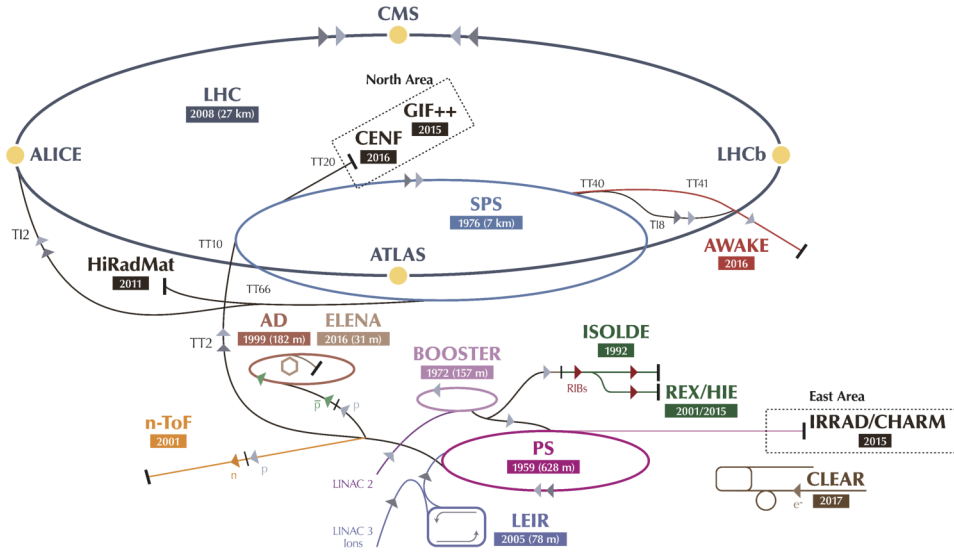


Figure 4.2: The CERN accelerator complex. Modified from [113]

increased to about 99.9 % of the speed of light, resulting in the final bunch spacing of 25 ns. The effect on the bunch spacing of any further increase in energy is not noticeable since the velocity is already approaching its theoretical threshold dictated by general relativity. The SPS is a 7 km-circumference accelerator and gathers up to four trains and accelerates them to 450 GeV. Finally, the trains are injected into the LHC. Several fills of the SPS are injected with pre-defined spacing into the LHC adding up to its nominal design number of bunches of 2808, i.e. 39 trains of 72 bunches each. [114] The spacing between the trains are governed by the time it takes to kick the bunches into or out-of their circular motion along with one of the accelerators during injection or for the beam dump, respectively.

Once the LHC has reached its maximum bunch capacity, the protons are accelerated to their final energy. The LHC was designed to reach 7 TeV beam energy, providing collisions at a centre-of-mass energy of $\sqrt{s} = 14$ TeV. Shortly after the LHC's initial operations, an incident occurred, causing severe damage to the machine and delaying the physics programme for years. Although the problem was found and measures were taken to prevent the problem from occurring again, it was decided to start the physics programme at a reduced beam energy of 3.5 TeV and 4 TeV during Run 1 and of 6.5 TeV during Run 2. Even at that reduced beam energy, the LHC was the highest-energy particle accelerator in the world from the beginning of its physics programme and has not been surpassed since.

A key factor for the performance of precision measurements and new physics searches, next to the available centre-of-mass energy \sqrt{s} , is the available number of events in data to reduce the effect of statistical uncertainties. As discussed in Section 2.1.8, the predicted rate of events per second \dot{N} can be calculated for a given \sqrt{s} as

$$\dot{N} = L \cdot \sigma, \quad (4.1)$$

where the instantaneous machine luminosity L summarizes the experiment-specific dependencies and σ is the cross-section and combines all other (e.g. process-specific) dependencies.

For a Gaussian beam distribution, the luminosity can be written as

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \varepsilon_n \beta^*} F \quad (4.2)$$

where N_b is 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 geometric luminosity reduction factor F is introduced to account for non-zero crossing angle of the beams at the interaction point. The normalized transverse beam emittance ε_n and the amplitude

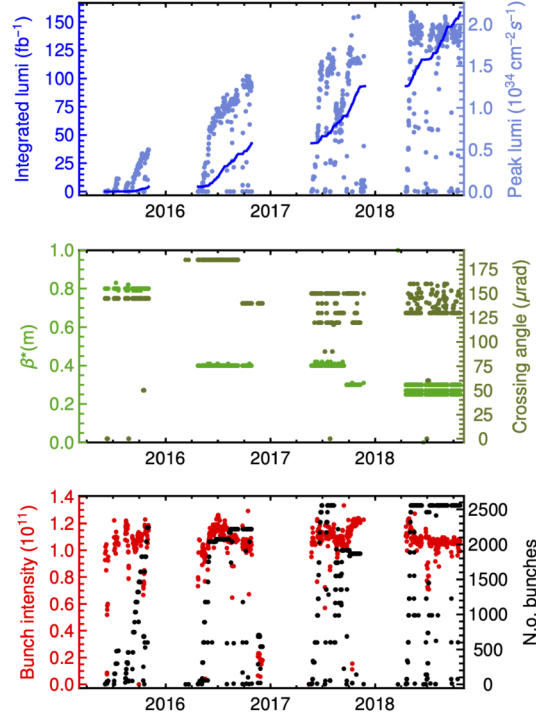


Figure 4.3: The integrated and peak luminosity (top), the β^* and the crossing angle (middle), and the bunch intensity and the number of bunches (bottom) over the years in LHC Run 2. The dots show the values at the start of stable beams for all Run 2 standard physics fills. Special runs have been excluded. From [116].

modulation at the collision point β^* are beam quality parameters, where ε_n is a measure for the oscillation of the particles around the central path along the detector. The amplitude modulation β^* quantifies how strong the beam is focused at the interaction point. The product $\varepsilon_n \cdot \beta^*$ is proportional to the beam area. [103, 115]

During Run 1 the LHC was operated with high safety margins, falling short compared to the design values for both the beam energy and the luminosity. The goal for Run 2 was to approach the design performance, motivating the increase in beam energy to 6.5 TeV. Other beam parameters were ramped up slowly during operation. Figure 4.3 shows the development of the LHC performance quantified in several parameters.

In 2015 the design luminosity of $1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ was not reached. While the number of bunches ramped up steadily to 2244, the design value of 2808 wasn't reached. Next to the number of bunches, the increased β^* of 0.8 m compared to the design value of 0.55 m resulted in a reduced peak luminosity of $5 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$. Over the following years the parameters were continuously optimized, so that the design luminosity was achieved during 2016, and was exceeded by the end of Run 2 in 2018 with a peak luminosity of $2.1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. [116–118]

Due to the good performance and availability, the LHC delivered more than 150 fb^{-1} of integrated luminosity during Run 2. Data taking stopped at the end of 2018 and is scheduled to continue with Run 3 in 2021 [119]. Until then, maintenance work and several upgrades are performed to the LHC and the detectors. In the following years, the LHC will be upgraded to the High-Luminosity LHC aimed to provide 5 – 7 times the nominal luminosity of the original LHC design values. [120]

4.2 The ATLAS Detector

The data studied in this work have been measured by the ATLAS detector. For a comprehensive review of the design of the ATLAS detector, see [105], which is briefly summarized in the following section.

4.2.1 General design

The ATLAS detector has been designed with a set of physics scenarios, from precise measurements of SM properties to possible observations of new physics phenomena, in mind. The most important benchmark for this was the search for a Higgs boson. However, also other scenarios for physics beyond the SM like the existence of new heavy gauge bosons W' or Z' , supersymmetric particles, or extra-dimensions have been considered.

These scenarios resulted in a list of performance goals:

- maximum available angular coverage to detect escaping particles
- good momentum resolution and reconstruction efficiency for charged-particles
- track reconstruction close to the interaction region to find secondary vertices for tagging of τ -leptons and b -jets
- good identification and momentum resolution for electrons, muons, and photons over a wide range of transverse energies
- hadronic calorimetry with full coverage for accurate jet and missing transverse energy measurements
- highly efficient triggering on low transverse-momentum objects with sufficient background rejection

The general conditions at LHC imposed additional demands on the design. In each bunch crossing, multiple p - p pair collisions coincide. During Run 2, typical numbers of interactions per bunch crossing were between 10 and 60. Most of these simultaneous interactions are not of interest, but the detector measures only the combined outcome of all of these interactions. So if the detector is read-out to measure a high-energy event of interest, one needs to try to minimize the effect of the additional overlayed detector signals, which is not connected to the primary interaction. This effect is called *in-time pileup*.

The bunch spacing of only 25 ns also implies that the decision whether an event should be saved and if so, the actual read-out process of the full detector, need to be very fast. There are necessary to minimize the effect of *out-of-time pileup*, i.e. measuring overlayed signals originating from other bunch crossings than the one of interest.

Also, due to a large amount of beam energy, the detector subsystems and their electronics need to be radiation-hard to withstand the stress imposed by the collisions.

The shape of the ATLAS detector is cylindrical and features an eight-fold azimuthal symmetry with respect to the beam pipe and can be seen in a cut-away view in Figure 4.4. With a length of 44 m and a diameter of 25 m the ATLAS detector is the largest of the experiments hosted at the LHC.

The subcomponents of ATLAS are arranged in multiple layers around the beam pipe and centred around the nominal interaction point. The inner-most component starting with a radius of only a few centimetres to about 1 m is the inner detector system designed for high-precision track reconstruction. The next layer going outwards is the electromagnetic calorimeter with excellent energy and position resolution. The hadronic calorimeter is designed to measure most kinds of particles escaping the electromagnetic calorimeter. Neutrinos and muons are the only particles in the SM that typically pass the hadronic calorimeter. While neutrinos cannot be measured directly in ATLAS, the outer-most part of the detector, the muon systems, is designed to identify and measure muons. Each of the subcomponents is discussed in subsequent sections.

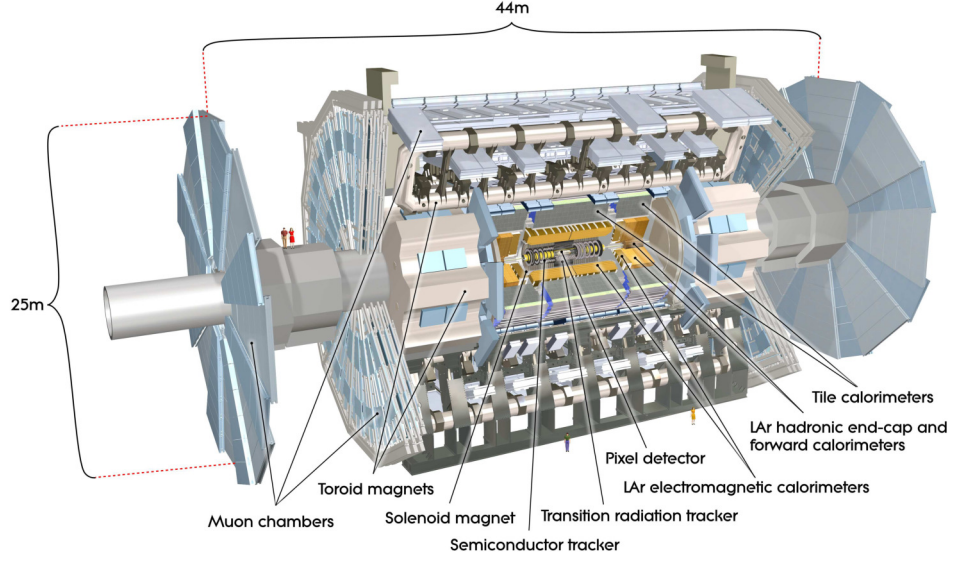


Figure 4.4: Cut-away view of the ATLAS detector. From: [105]

The ATLAS detector is immersed by a magnetic field generated by superconducting solenoidal magnets around the inner detector and large toroidal magnets in the muon chambers, allowing for precise momentum and charge measurements. The field in the inner detector was designed to be aligned with the beam axis and has a strength of 2 T. The strength of the toroidal magnetic field in the muon detectors is about 0.5 T (1 T) in the central (endcap) regions respectively.

4.2.2 Coordinate system and common observables

There are a well-defined coordinate system and a conventional notation for analyses in ATLAS. The coordinate system is a right-handed cartesian system, where the z -axis points along the beamline and the origin is set to the nominal interaction point. The x -axis is in the horizontal plane and the positive side points towards the centre of the LHC accelerator ring. The positive side of the vertical y -axis points upwards. Thus, the positive z -axis points in a counter-clockwise direction along the LHC ring.

The side of the ATLAS detector with positive z is labelled side A (for “anti-clockwise”) and with negative z side C (for “clockwise”).

The usual definitions for the azimuthal and polar angles θ and φ are adopted from the spherical coordinate system

$$\varphi = \arctan \frac{x}{z}, \quad (4.3)$$

$$\theta = \arccos \frac{z}{\sqrt{x^2 + y^2 + z^2}}. \quad (4.4)$$

The partons that take part in the hard interaction processes, only carry an unknown fraction of the proton’s momentum along the beam axis. The centre-of-mass system of the interacting partons and thus, the hard process is consequently unknown as well. For this reason, observables, that are invariant under boosts along the beam axis are preferred.

The azimuthal angle θ describes the angle between a certain direction and the beam axis and is not invariant under these transformations. The *rapidity* Y^3 is introduced as

$$Y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right). \quad (4.5)$$

³Often the lower-case letter y is used for the rapidity as well. This notation is not used throughout this work to avoid confusion with the cartesian coordinate y .

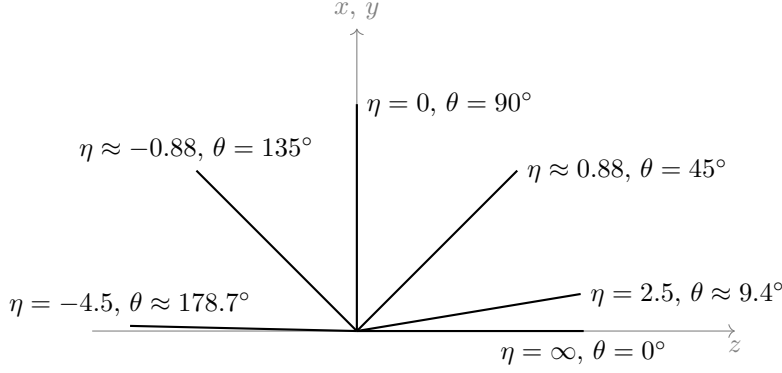


Figure 4.5: Relation of θ and η .

While also the rapidity itself is not invariant under boosts along the z -axis, differences in rapidity between two momenta are. Additionally, the *pseudo-rapidity* η is introduced as

$$\eta = -\ln \tan \frac{\theta}{2}. \quad (4.6)$$

This relation is visualized in Figure 4.5. For massless particles, the pseudo-rapidity η is equivalent to the rapidity Y . Due to the direct relation to θ and the advantage of invariance against boosts along the z -axis for the separation of massless particles, η is usually preferred over the azimuthal angle θ .

Based on the pseudo-rapidity η , an adjusted metric for angular distances between two directions is commonly introduced

$$\Delta R = \sqrt{\Delta\varphi^2 + \Delta\eta^2}, \quad (4.7)$$

where $\Delta\varphi$ is the smaller angle between the projections of the directions into the transverse plane. Throughout this work, this is adjusted for massive particles such as W^\pm or Z bosons or jets to use the rapidity Y rather than the pseudo-rapidity η

$$\Delta R = \sqrt{\Delta\varphi^2 + \Delta Y^2}. \quad (4.8)$$

unless stated otherwise. Electrons and muons are considered massless in this context. Since only the differences of rapidities enter in this distance, it is invariant under boosts along the z -axis.

Other commonly used observables are the *transverse momentum* p_T and the *missing transverse energy* E_T^{miss} . Both are measured in the x - y -plane. The missing transverse energy is a measure for the imbalance of the calorimeter signals. Momentum conservation dictates that the sum of all momenta of an interaction in that plane balance out. Imbalances can be caused by uncertainties in momentum measurements, signals wrongly assigned to the main event, or undetected particles, such as neutrinos or particles outside the acceptance of the detector.

A set of *impact parameters* is introduced to quantify the distance between a track and a point in space. The beam spot is usually used as a reference point. Unless stated otherwise, this will be the default point for impact parameters throughout this work. The beam spot is measured by averaging the main interaction point for many random events and is updated periodically after few minutes.[121] Depending on the beam conditions at LHC this beam spot might be shifted compared to the nominal interaction point by few mm in the transverse plane and few cm along the z -axis.[122] This procedure was chosen because the size of the beam spot can be measured more accurately compared to the position of primary vertices in the transverse plane.[123, 124]

The definition of the impact parameters is based on the *point of closest approach (pca)*. The pca is the point along the track which is closest to the beam spot when projected onto

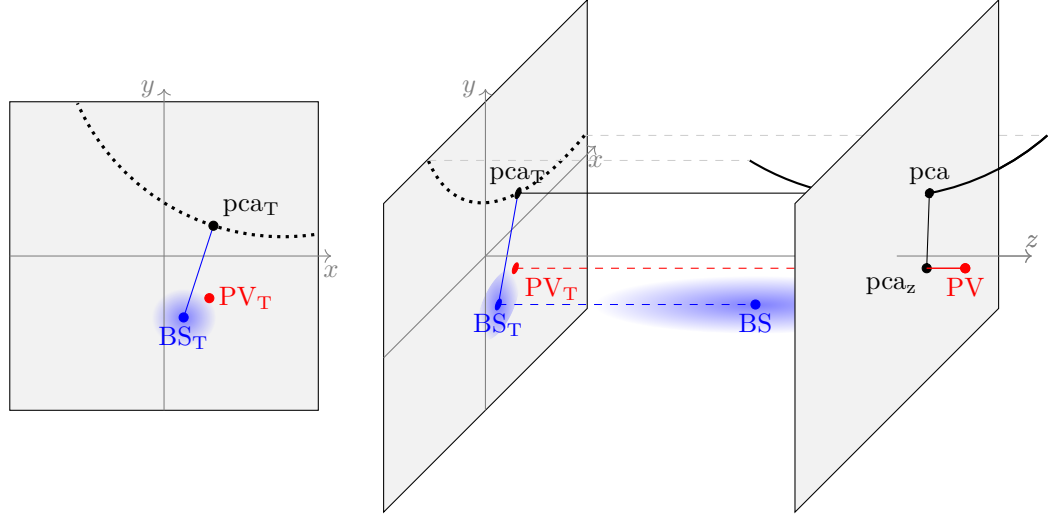


Figure 4.6: Visualisation of the definition of impact parameters. The left plot shows the projection onto the transverse plane, and the right plot shows a 3-dimensional view. The dotted line shows a projection of the track onto the transverse plane. The blue areas indicate the beam spot (BS) and its size and the red dots indicate the primary vertex (PV) and its projection onto the transverse plane (PV_T). The black dots are the point of closest approach (pca) and its projection (pca_T). The transverse impact parameter d_0 is the distance between BS_T and pca_T and is indicated by a blue line. The grey areas in the right plot indicate the transverse plane at $z = 0$ and at the pca. The longitudinal impact parameter z_0 is the distance along z between the pca and the PV, as indicated by a red line.

the transverse plane. The transverse impact parameter d_0 is the distance of this projection to the beam spot in the transverse plane. The longitudinal impact parameter z_0 is the distance along the z -axis of the pca to the primary vertex, i.e. not the beam spot. This is visualized in Figure 4.6.

4.2.3 Inner detector

The inner detector of ATLAS is designed for precise measurement of tracks of charged particles, necessary for accurate momentum and charge measurement. Similar to the overall ATLAS layout, the inner detector consists of a barrel part and additional disks to extend coverage in η . It consists of many layers, using different technologies to measure the position, where a charged particle did pass each layer. The information of all layers is combined to estimate the most likely set of tracks causing the hits.

The reconstructed tracks allow for an assignment of tracks to reconstructed vertices and thus to distinguish between tracks caused by pileup interactions from the hard process of interest. Additionally, the curvature of the tracks, caused by the magnetic field, is used to measure the charge and momentum of the particle. The general layout is shown in Figure 4.7.

For a precise measurement of primary vertices, the detectors need to be very close to the actual beamline to minimize the need for extrapolations. In order to increase precision compared to Run 1, an additional layer was added for Run 2, the insertable B-Layer (IBL)[126]. With this layer, the minimal distance to the centre of the beam pipe was reduced from 50.5 mm to less than 34 mm.

The innermost layers employ silicon pixel detectors, offering the best spatial resolution of $10\,\mu\text{m}$ in $R - \varphi$ and $115\,\mu\text{m}$ in z (R) for the barrel (in the disks). For the IBL the resolution in z could even be reduced to about $66\,\mu\text{m}$. [127]

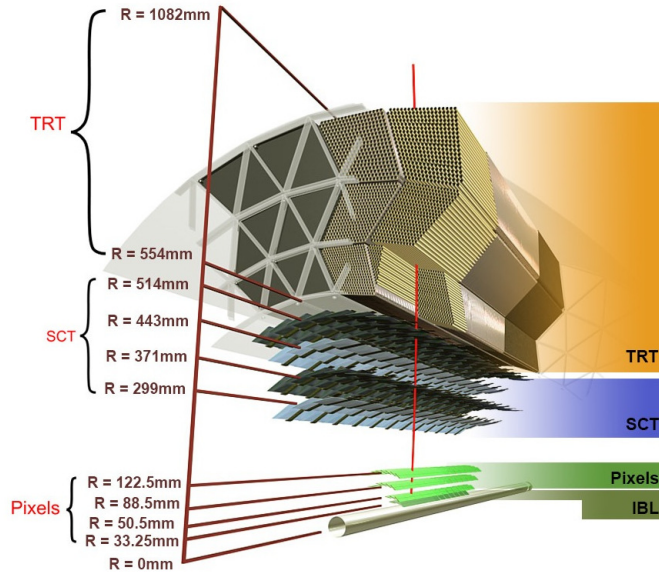


Figure 4.7: Sketch of the ATLAS inner detector showing all its components, including the new insertable B-layer (IBL). The distances to the interaction point are also shown. From: [125]

The subsequent layers, which employ silicone microstrips, form the *semiconductor tracker* (SCT). Each particle passes four pairs of two layers. Each pair allows for the precise measurement of a space point. These layers provide a resolution of $17 \mu\text{m}$ in $R - \varphi$ and $580 \mu\text{m}$ in z (R) for the barrel (in the disks). All of the pixel and microstrip detectors provide coverage up to $|\eta| < 2.5$.

The remaining layers of the inner detector, the *Transition Radiation Tracker* (TRT), covers only the region to $|\eta| < 2.0$ due to spatial constraints. These layers consist of straw tubes parallel to the beam axis and provide only $R - \varphi$ information with a resolution of $130 \mu\text{m}$.

This design of the inner detector benefits both from the high spatial precision close to the interaction points and the large number of hits and the considerable track length using the information measured by the TRT.

4.2.4 Calorimeters

There are two prerequisites for accurate energy measurement and suitable identification of particles using the calorimeter. Sufficient thickness avoids punch-through into the muon system and fine granularity help distinguishing shower shapes. For the measurement of $E_{\text{T}}^{\text{miss}}$, a maximum possible η coverage is necessary. The ATLAS calorimeter system (see Figure 4.8) fulfils both of these criteria very well. It can be subdivided into four components, the Electromagnetic Calorimeter (EM), hadronic tile calorimeters, the Hadronic Endcap Calorimeter (HEC), and the Forward Calorimeter (FCal).

The EM system has a cylindrical shape, with a barrel system covering $|\eta| < 1.475$ and independent endcap calorimeters (EMEC) covering $1.375 < |\eta| < 3.2$. This calorimeter system employs liquid argon (LAr) as the active medium and lead as absorber material. The material before and in the EM system amounts to more than 25 radiation lengths X_0 .

The hadronic tile calorimeter consists of a central barrel covering $|\eta| < 1.0$ and extensions on both sides, which enclose the endcaps of the other calorimeter systems. The coverage of η is thus extended to $|\eta| < 1.7$. The technology is based on lead as absorber medium and scintillators as a sampling medium.

The HEC extends the hadronic coverage to $|\eta| < 3.2$ and uses copper absorber and LAr as active medium. On each side, there are two independent wheels.

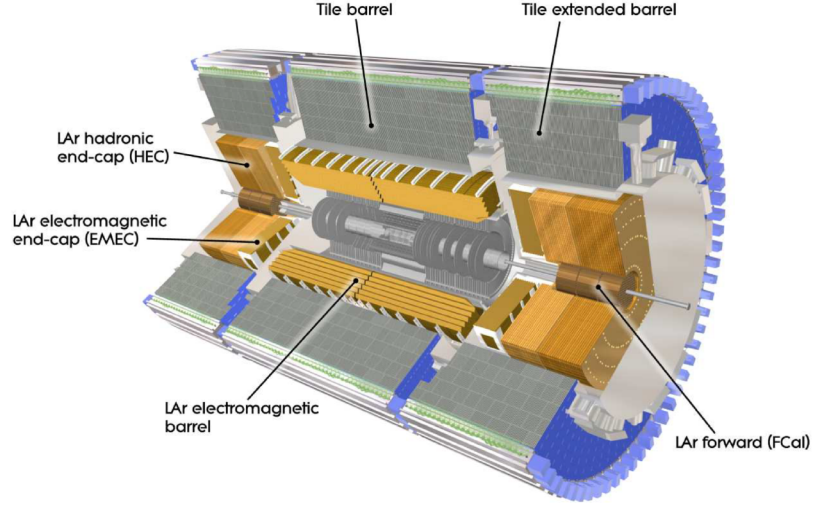


Figure 4.8: Cut-away view of the ATLAS calorimeter system. From: [105]

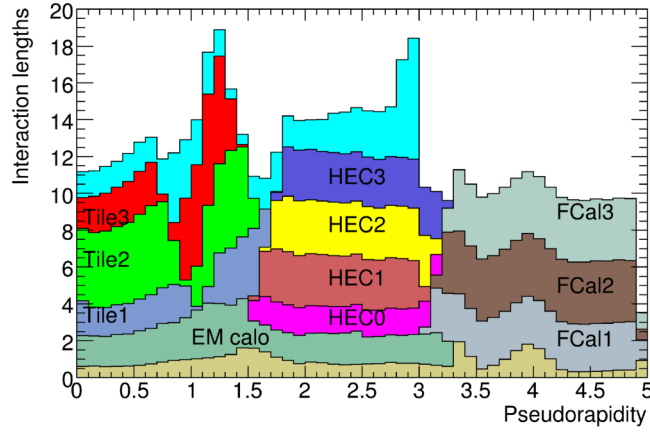


Figure 4.9: The cumulative amount of material, in units of interaction length, as a function of $|\eta|$, in front of the electromagnetic calorimeters, in the electromagnetic calorimeters themselves, in each hadronic compartment, and the total amount at the end of the active calorimetry. Also shown for completeness is the total amount of material in front of the first active layer of the muon spectrometer (up to $|\eta| < 3.0$). From: [105]

The LAr-based FCal consists of three modules on each side, extending the η coverage in the range of $3.1 < |\eta| < 4.9$. The module closest to the nominal interaction point uses copper as absorber material and functions as an electromagnetic calorimeter. The other two modules use tungsten and are hadronic calorimeters.

The cumulative amount of material, given in units of interaction length λ , as a function of $|\eta|$ is shown in Figure 4.9. The total thickness of about 10λ ensures a sufficiently low punch-through rate.

The granularity of the calorimeter is in general in the order of 0.1×0.1 in $\Delta\eta \times \Delta\varphi$. Only for $|\eta| < 2.5$, the electromagnetic calorimeter system provides substantially finer granularity, reaching 0.025×0.025 . This fine granularity allows for precise spatial measurements and identification of photons and electrons. In the FCal the granularity is much coarser for technical reasons. There the granularity is more coarse by factors between 20 and 50 in each dimension.

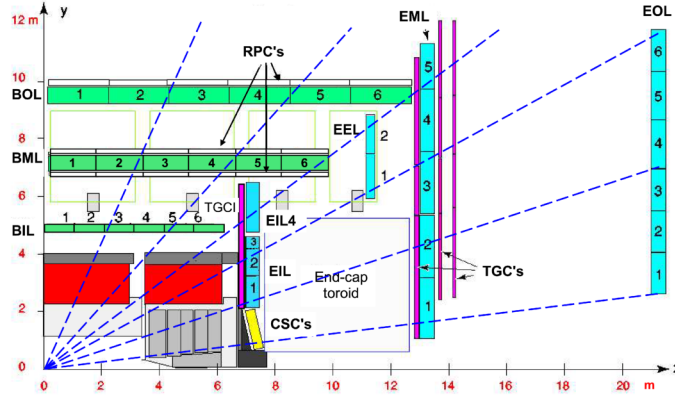


Figure 4.10: Cross-section of the muon system in a plane containing the beam axis. Dashed lines illustrate straight trajectories of hypothetical infinite-momentum muons. [105]

4.2.5 Muon system

The muon system provides accurate information for the identification and measurement of muons. Similar to the inner detector, the momentum measurement is based on the deflection of charged particles in the magnetic field, which permeates the detector. To allow for accurate measurement of the deflection of tracks, the alignment of the individual subcomponents and the strength of the magnetic field have to be measured precisely.

The muon system consists of four subcomponents; the *monitored drift tubes (MDT)*, the *cathode strip chambers (CSC)*, the *resistive place chambers (RPC)* and the *thin gap chambers (TGC)*. The MDT and CSC provide many hits with precise spatial accuracy, providing information for precise track reconstruction for pseudo-rapidities of $|\eta| < 2.7$. The RPC and TGC systems are focussed on precise time information and fast read-out to be used for fast and reliable triggering for $|\eta| < 2.4$. The overall layout of the muon system is shown in Figure 4.10.

The MDT chambers use small tubes with a diameter of ≈ 3 cm, filled with pressurized gas. An accuracy of $35 \mu\text{m}$ is achieved for the shortest distance between the muon passing the tube and the wire. This accuracy is possible using the timing and shape of the measured pulses caused by electrons produced along the track by ionization. No information on the coordinate along the tube is measured. The tubes are aligned along φ , such that the position of the track in the bending plane is measured.

The MDT chambers are used for the largest part of the detector for precision-tracking, covering $|\eta| < 2.7$ with a small crack around $\eta \approx 0$. Only at the inner-most layer in the endcap the MDT ends at $\eta \approx 2.0$ and the CSC provides necessary information for $2.0 < |\eta| < 2.7$. In this region the track density is the highest, exceeding the limits of safe operations for the MDT. The CSC provides measurements of both track coordinates via the induced charges on the grid of cathode strips. This technology allows higher counting rates and provides information on the timing with an accuracy of about 7 ns. The resolution in the bending plane is about $40 \mu\text{m}$ and about 5 mm in the φ direction.

The triggering system is build up from the RPC in the barrel ($|\eta| < 1.05$) and the TGC in the endcap ($1.05 < |\eta| < 2.4$). These systems provide accurate timing information for reliable identification of the bunch crossing a muon originates from as well as fast and coarse tracking information to be used in higher-level triggers. Additionally, the measurements of these subcomponents can be used to improve the track information when matched to the MDT measurements by providing a measurement of the track coordinate along φ with an accuracy of < 10 mm.

The RPC consists of three layers of chambers attached to an MDT chamber each. The RPCs employ two resistive plates with an electric field of about 4.9 kV/mm for each chamber. Electrons, resulting from ionization, form avalanches towards the anodes, which are

subsequently measured via capacitive coupling to metallic strips.

The TGCs are multi-wire proportional chambers with a potential of $\approx 3\text{ kV}$ and wire-to-wire distance of 1.8 mm , leading to short drift times. Due to the arrangement of the TGC in the detector, muons from the interaction point always have incidence angles $> 10^\circ$, thus at least partially avoiding the low-field region in the middle of two wires. The TGCs achieve a timing resolution of about 4 ns and spacial resolution of $2 - 7\text{ mm}$ in both coordinates.

4.2.6 Luminosity measurement

There are two dedicated detector systems built to measure the instantaneous luminosity delivered to ATLAS and the number of interactions per bunch crossing; *LUCID* (*luminosity measurement using Cerenkov integrating detector*) and *ALFA* (*absolute luminosity for ATLAS*).[105]

ALFA is situated in a distance of 240 m on both sides of the nominal interaction point and is as close as 1 mm to the beam. ALFA consists of scintillating fibre trackers in Roman pots and measures the absolute luminosity via elastic scattering at small angles. Such measurements are, however, only possible in special runs with high β^* .

LUCID only has a distance of 17 m from the nominal interaction point and is used for relative measurements of the instantaneous luminosity. LUCID detects inelastic $p\text{-}p$ scattering at high rapidities. During the long shut-down between Run 1 and Run 2, LUCID has been upgraded and renamed to LUCID-2. The upgrade extended the applicability to the more demanding conditions. In order to allow for more stable luminosity determination over long times, small amounts of radioactive sources are added to allow monitoring the gain stability of the photomultipliers.[128]

In contrast to the measurements in ALFA, the LUCID measurements have to be calibrated to translate this process to the instantaneous luminosity. For this calibration, so-called van-der-Meer scans are performed in special low-luminosity runs. This calibration is then transferred and applied to runs with the usual physics conditions.[129]

For the data-taking period considered for this work, an integrated delivered luminosity of 36.2 fb^{-1} was measured with an uncertainty of 2.1% . The main uncertainty was caused by the transfer of the calibration from special runs to nominal physics conditions.[129]

4.2.7 Trigger

The time between two bunch crossings of 25 ns is smaller than the time it takes to read-out the full detector ($\mathcal{O}(100\text{ ns})$). Thus it technically is not possible to read-out and to store each collision. Even when trying to store always the first collision after a read-out, the disk space that is needed to store the information of one event is about 1 MB , resulting in a hypothetical rate of about $\approx 10\text{ MHz} \cdot 1\text{ MB} = 10\text{ TB/s}$. In order to keep costs for storage and analysis of the data feasible, stricter and ideally also more intelligent filters have to be applied. The ATLAS trigger system implements such filters.

The trigger system was originally designed with three levels:

- the hardware-based L1 trigger, using only fast and coarse information and tagging *regions-of-interest* (*ROI*) in the event
- the software base L2 trigger that used more detailed information in these regions of interest to perform more accurate calculations
- the event filter, refining the decisions of the L2 layer and reducing the rate further.

The design also included the possibility to prescale a subset of triggers, such that only a fraction of the events passing the trigger are stored to reduce the rate while still enabling studies of these kinematic regions. An overall rate reduction to 200 Hz was the goal. This setup was used during Run 1, while the average final event streaming rate was allowed to be up to $\approx 700\text{ kHz}$ due to increased computing capabilities.[130]

In order to prepare the trigger system for the increased demands imposed for Run 2, several changes and upgrades have been performed.[131] Similar to the original design, the

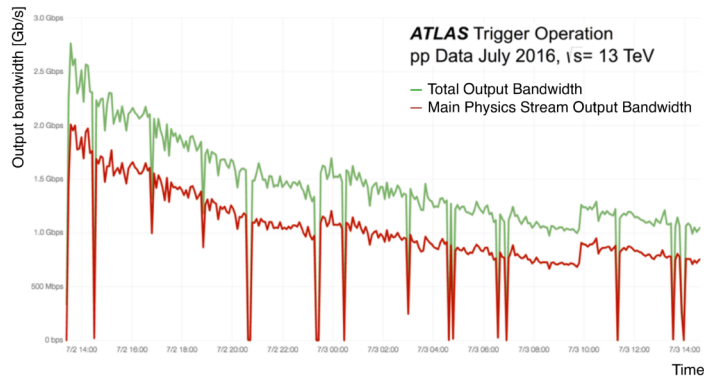


Figure 4.11: Total and central physics stream output bandwidth at the High-Level Trigger (HLT) as a function of time throughout a fill taken in July 2016 with a peak luminosity of $L = 1.02 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ and a peak pile-up of $\mu = 35$. From: [132]

first step is a hardware-based L1 trigger, using the information of subsets of the calorimeters (L1Calo) or the muon system (L1Muon), respectively. Additionally, topological triggers (L1Topo) based on FPGAs were added. This new part of the central trigger combines the information provided by the L1Calo and L1Muon systems and selects events based on geometrical or topological criteria, e.g. the missing transverse energy.

The *Central Trigger Processor (CPT)* forms the final L1 decision, taking into account dead-time criteria to ensure non-overlapping detector read-out windows and reliable performance. In case an event is accepted, the detector information is read-out and buffered in the *Read-Out System*.

The software-based *High-Level Trigger (HLT)* has access to the full detector data as well as coarsely reconstructed tracks, using a linear fit rather than a fit of a helix, and the ROI provided by the L1 triggers. This step is merged from the L2 trigger and the event filter from the original design into a shared computing farm for simplification and to improve the sharing resources.

The constraints on the current system are a maximum of 100 kHz on the L1 trigger rate due to the ATLAS read-out times and the goal to keep the average HLT physics output rate at around 1 kHz to ensure cost and time for storage and analysis of the recorded data stay on a functional level. An example of the development of the output bandwidth during an LHC physics fill is shown in Figure 4.11.

4.3 Reconstruction of Analysis Objects

The interpretation of the detector information in terms of physics objects as well as some corrections and classifications for these physics objects are very fundamental building blocks of all analyses. Within the ATLAS collaboration, the combined performance groups develop, maintain, optimise, and provide algorithms and tools along with recommendations to be used by physics analysis groups.

In general, the first step is to combine individual detector information to basic building blocks, i.e. to combine individual hits in the tracker to find track candidates and to combine calorimeter cells where a certain amount of energy was deposited to clusters.

From these tracks and clusters, candidates for physics objects such as electrons, muons, or jets are reconstructed. This reconstruction is in general very conservative, aiming for high efficiency in finding each object of the type of interest. In order to better distinguish correctly reconstructed objects from misidentified objects, e.g. jets that have been reconstructed as electrons, multiple working points of object identification are introduced for some object types. The tighter the working point, the higher the probability, that a reconstructed object was correctly identified.

Also, corrections are derived, e.g. for the energy or momentum of the object or to correct

for efficiency differences between simulated and recorded data. Such differences are typically caused by imperfect simulation of the detector, e.g. due to limited knowledge of the material distribution.

In the following, the most important recommendations are summarised for electrons, muons, jets, and the missing transverse momentum since these are the objects used in this work. A detailed example of how such recommendations can be derived is shown for the energy calibration of electrons outside of the tracking acceptance of the detector in Appendix D. Although not used in this work, this calibration was derived by the author as qualification project.

4.3.1 Electrons

Electrons have a distinct signature of charged particle tracks in the inner detector and clusters in the electromagnetic calorimeter. Typically one expects one track, but additional tracks can be present due to Bremsstrahlung. The energy deposits in the electromagnetic calorimeter are typically more narrow than clusters caused by jets.

The reconstruction of electrons bases on this signature and is applied in three steps. The procedure is described in detail in [133] and summarised in the following. Firstly, clusters are formed from the calorimeter information. Then tracks are formed and fitted to the hits in the inner detector. Finally, the tracks are matched in η - φ to the calorimeter clusters. Electron candidates are build from the combined information of both subdetectors.

In order to build clusters, the information of all calorimeter layers is first combined to calorimeter towers. These towers have a granularity of 0.025×0.025 in $\eta \times \varphi$ defined by the second layer, which has the least fine granularity. Towers that exceed a minimum threshold of $\sum E_T > 2.5$ GeV are used as a seed to build clusters using a sliding-window technique [134] with a window size 3×5 . For electrons with $p_T > 15$ GeV the reconstruction efficiency of the clusters is $> 99\%$.

Tracks are formed from the hit information in the inner detector. Track candidates are seeded from sets of three space points translated from hits in the silicon layers. Each of the candidates is then extrapolated using two different mass hypotheses, which differ in the amount of relative radiative losses due to Bremsstrahlung, and checked for consistency with the measured hits in the other layers and possibly dropped. A χ^2 fit is applied to improve the matching the remaining track candidates to the observed hits while resolving ambiguities of shared hits. Track candidates have to be associated with four hits in the silicon detector, and only those track candidates are kept that loosely match a calorimeter cluster in $\Delta\eta$ and $\Delta\varphi$ for the next step. Tracks passing these criteria are fit again using an optimised Gaussian-sum filter [135], which models energy losses of charged particles better. [136]

After this fit, all track candidates close to a calorimeter cluster in $\Delta\eta$ and $\Delta\varphi$ are assigned to that cluster to build an electron candidate. If multiple tracks are assigned, a primary track is chosen based on the angular distance to the cluster and the number of hits in the inner detector. The four-momentum of the electron candidate is defined by the calorimeter cluster's energy, which needs calibration in a subsequent step, and the direction of the primary track.

The signature defined above can also be caused by photons causing hits in the inner detector due to radiated electron-positron-pairs. Based on several criteria on the primary track and the ratio of energy deposits in the calorimeter and the momentum of the track, electron candidates are classified as an electron, photon, or ambiguous.

The total efficiency for electrons with $p_T > 15$ GeV is about 98% and depends on η and p_T . The resulting electron candidates are available in special containers in the ESD and xAOD files. Further algorithms for selection and corrections of electrons are available to be applied by the analysers.

Based on these reconstructed electron candidates, multiple operating points for identification of electrons are available to allow different compromises between accepting misidentified objects and identification efficiency of prompt electrons. Misidentified objects could be jets reconstructed as electrons, electrons from γ conversions, or non-prompt electrons originating from hadronic decays.

The identification is based on a likelihood combining several observables like ratios of individual layers including the first layer of the hadronic calorimeter to the tower energy, track impact parameters, lateral shower widths, and the track-cluster matching parameters $\Delta\eta$ and $\Delta\varphi$. Probability density functions (pdf) of each parameter for signal and background objects are estimated from simulations. A signal (background) likelihood L_S (L_B) is formed for each electron candidate by multiplying for each observable the value of the signal (background) pdf for the value of the observable. A discriminant d_L is defined as

$$d_L = \frac{L_S}{L_S + L_B} \quad (4.9)$$

and subsequently transformed with an inverse sigmoid function for easier handling.

Four operating points are defined via increasing thresholds for this discriminant in combination with additional simple cuts on other variables such as the number of hits in different subsystems of the inner detector; **VeryLoose**, **Loose**, **Medium**, and **Tight**. One additional operation point is introduced based on the **Loose** operating point by adding a requirement on a hit in the IBL; **LooseAndBLayer**. In contrast to Run 1, the working points are defined such that candidates passing a tighter working point also pass the looser working points.

A multi-step approach is performed to calibrate the energy of electrons based on the information of the associated calorimeter cluster as described in detail in [137]. In a first step, a layer-dependent correction is applied in data only to correct for known inhomogeneities in the response behaviour of the different layers. The energy of the electron candidate is estimated using multi-variate techniques from the energy deposits. These algorithms were trained on simulated data and are applied equivalently on data and simulation. Potential local non-uniformities in the calorimeter response can be caused by non-nominal high-voltage settings or geometric effects and are corrected for in data. The corrections are obtained by studying the ratio of the track momenta and the calorimeter cluster information in Z boson decay events.

The final step in the calibration procedure is an *in-situ* calibration, to correct for remaining differences between data and simulation. These differences are studied in $Z \rightarrow ee$ events with a well-identified electron-positron pair. The shapes of the invariant mass M_{ee} of the electron-positron pair in a window around the PDG mass of the Z boson is measured in data and compared to simulations. The peak originating from the decay of a Z boson is observed at different values, and the energy scale of electron candidates in data is corrected to match the peak position in simulations. Similarly, the width of the peak is smaller when using simulated data, and a smearing procedure is applied on simulated data to match the energy resolution in data. The in-situ calibration for electrons in the forward region of the detector was derived as part of this work. The procedure is described in more detail as an example of the general approach in Appendix D.

The overall calibration procedure is validated by studying invariant mass distributions in data and simulations for events selected for $J/\psi \rightarrow ee$ decays and for radiative Z decays $Z \rightarrow ee\gamma$.

4.3.2 Muons

The signature of muons in the ATLAS is very clear due to the muon system, with which typically only muons interact. The main challenge is to combine information from the inner tracker with the segments reconstructed from the muon system and to distinguish muons originating the hard process from muons from other origins, e.g. from hadron decays.

For the reconstruction of muons, track segments are first built from the information of the inner detector and the muon system separately. In the last step, these track segments are combined to build muon candidates as described in more detail in [138]. The reconstruction of tracks is equivalent to the reconstruction of tracks for electron candidates as described above and the same track candidates that pass the global χ^2 fit are used.

In order to build track candidates from the information of the muon system, first the hits in each of the three MDT layers are fitted using a straight line and combined with information of the trigger systems to build track segments. A Hough transformation[139]

is used to reduce computing time. These segments are used as seeds to build muon track candidates starting from segments in the middle layer as these are equipped with two trigger chambers and thus provide the best measurement orthogonal to the bending plane. The muon track candidates are accepted based on the number of hits and the fit quality. In general, at least two matching segments are required to build a track candidate, but in the transition from the barrel to the endcap, track candidates with only one high-quality segment are accepted. Finally, all track candidates are refined by performing a global χ^2 fit to the hits.

The ID and MS tracks, as well as information from the calorimeters, are combined to build muon candidates. Depending on which subsystem did provide information, muons are classified into four types:

Combined muons: track candidates from ID and MS are matched and a global refit of ID and MS hits is performed. This type provides the best quality of muon reconstruction.

Segment-tagged muons: track candidates from ID are accepted as muons, in case they are matched to at least one segment in the muon system when extrapolated. This type is primarily used to recover overall reconstruction efficiency for muons with low p_T and regions of reduced acceptance.

Calorimeter-tagged muons: track candidates from ID that can be matched to energy depositions in the calorimeter that are compatible with muons. This type has the lowest purity and is mostly used to cover the region of the central crack in the muon system at $|\eta| < 0.1$.

Extrapolated muons: track candidates from MS that are loosely matched to the interaction point. The energy loss in the calorimeter is estimated, resulting in a significantly reduced energy resolution. This type is used to extend the coverage in $|\eta|$ to 2.7, where no ID information is available.

A collection of muon candidates is built from all of these types while removing overlaps. This collection is stored in containers in the ESD and xAOD files.

The primary background contamination in the candidate muons is non-prompt muons originating from in-flight decays of, e.g. pions or kaons. These candidates typically show a poor fit quality and have incompatible momenta in the ID and MS measurements.

The observables used in muon identification are the normalized χ^2 of the combined track fit for combined muons, the absolute value of the difference in the transverse momentum measurements, the significance of the difference in the ratio of the charge over the momentum between the ID and MS tracks, and the number of hits and holes of the track in different subsystems.

Three operating points are offered with increasing purity; **Loose**, **Medium**, and **Tight**. Similar, as for electrons, the operating points are designed to be subsets of each other, i.e. all muon candidates passing the **Tight** (**Medium**) identification also pass the **Medium** (**Loose**) identification. The **Medium** operation point was designed to minimize systematic uncertainties and is thus the default recommendation, while the **Loose** (**Tight**) operation point increases the reconstruction efficiency (signal purity). An additional operating point **High- p_T** is introduced to specifically reduce the systematic uncertainties on the momentum measurement for muons with $p_T > 100$ GeV.

The identification efficiencies are about 91 % for the **Tight** working point, 96 % for the **Medium** and 98 % for the **Loose** working point.

The calibration of muons is estimated using combined muons only and is described in more detail in [138]. The momentum scale and resolution smearing of the ID and MS tracks are corrected and subsequently combined with a weighted average. The correction analytically depends on p_T and the correction factors are binned in η and φ .

The correction factors are estimated in $Z \rightarrow \mu\mu$ events that are selected with two well-reconstructed muons with opposite charges, passing the **Medium** identification. Template distributions in the invariant mass of the dimuon system, reconstructed in the detector subsystem of interest, are generated for different values of the correction factors. These are

fitted in data to extract the best-fit values for the correction factors in regions η and φ . For the MS, the momentum imbalance between the MS and the corrected ID measurements is also included in the fit.

The calibration is validated by studying the invariant mass of the dimuon system measured using the corrected, combined muon momenta in $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ events.

4.3.3 Jets

Reconstruction: The most commonly used collection of jets is reconstructed by running the same jet finder algorithm as introduced in Section 3.2.5, AntiKt with a radius parameter of $R = 0.4$ on special clusters of energy deposits in the calorimeter system. The so-called *topo-clusters* combine information of topologically connected calorimeter cells. By disregarding insignificant calorimeter cell signals that are not connected topologically to a cell with a significant energy deposit, a significant noise reduction is achieved. The procedure is summarized in the following, and more details can be found in [140, 141]

The motivation for using topo-clusters is to combine information on the deposited energy with three-dimensional direction reconstruction of particles based on the topological shape of a set of neighbouring cells with significant energy deposit.

The clustering process for the topo-clusters is based on the cell signal significance, i.e. the ratio of energy deposit in a cell over an estimate for the average noise, both measured on the electromagnetic energy scale.⁴ The cluster formation is seeded at each cell, where the cell signal significance exceeds a threshold of 4. Starting from these seeds, all neighbouring cells exceeding a cell signal significance of 2 are added to the proto-cluster recursively. The recursion stops once the energy deposits in each neighbouring cell of the proto-cluster are not considered significant enough; that is, each cell signal significance is smaller than 2.

When using this algorithm, cases can occur, where a proto-cluster has multiple local signal maxima. Depending on the actual distance and height of the maxima, the proto-cluster might be split. Each cell is either assigned to one of the proto-clusters or shared between proto-clusters according to a cell-specific weight.

The directional information of the topo-cluster η and φ with respect to the origin of the ATLAS coordinate system is derived as the signal-weighted barycentres using the absolute values of the signal cell significances. The energy assigned to the topo-cluster is the weighted sum of the energy deposits in the clusters. A four-momentum is build based on a hypothetical massless pseudo-particle with this energy and direction and assigned to the topo-cluster.

These four-momenta are used as nominal input for the jet clustering. Additionally, tracks reconstructed during the electron or muon reconstruction are assigned to jets in a process called ghost-association. These tracks are considered in the jet reconstruction with infinitesimal momentum. With this, the tracks do not affect the four-momentum of jets but can be assigned unambiguously.

Energy correction: The calibration of jets is split into multiple steps to match the energy scale of reconstructed jets to the scale of jets clustered from particle-level information and of harmonizing the energy measurement in data and simulation.

The first step is an origin correction, in which the origins of the four-momenta of reconstructed jets are corrected to point to the primary vertex. This correction improves the resolution in η .

In the subsequent steps, the effect of pileup should be minimized. Based on the jet-area, the estimated p_T density of pileup events is subtracted from the energy deposits to correct the jet energy. The p_T density is estimated by using the median density per jet for the full event, which is dominated by a large number of low-energy pileup jets. Further corrections for remaining pileup dependence are derived based on simulations. A linear dependence of the difference in p_T to the matched jet at truth level on the number of primary vertices and

⁴At the electromagnetic scale, the detector response is translated to energy such that the energy of electrons and photons are adequately reproduced. Necessary corrections for hadronically interacting particles are not included and need to apply after reconstruction.

the average number of interactions per bunch crossing $\langle\mu\rangle$ is observed. Linear correction factors are estimated in bins of the pseudorapidity η with a subsequent piece-wise linear fit in η to reduce the impact of statistical fluctuations. A residual logarithmical dependence of the correction factor on the transverse momentum p_T was found and assigned as pileup dependent systematic uncertainty.

The absolute value of the jet energy scale and the rapidity is calibrated using corrections derived from simulations of dijet events. The average energy response is measured from well isolated detector-level jets that are well matched to a particle-level jet within $\Delta R < 0.3$. The energy response is derived as a function of the rapidity of the detector-level jet and the energy of the particle-level jet and inverted numerically to obtain correction factors for the reconstructed jet energy. An additional bias in the jet rapidity is observed, especially pronounced in the transition ranges between different detector technologies or granularities. This bias is partially due to the inhomogeneous energy response. This bias is measured, and correction factors are derived equivalently to the jet energy response. The derived correction factors for η only correct the p_T and η of the jet momentum.

Additional dependencies of the energy response on other observables sensitive to the jet particle composition and the distribution of energy within the jet are observed and corrected to reduce systematic uncertainties. For each of these five observables, correction factors are derived depending on the transverse momentum of the particle-level jet and the rapidity of the detector-level jet. These corrections are defined such that the already-corrected average energy response is not affected. The observables for which such a correction is derived and applied are [142]

- the fraction of jet energy measured in the first layer of the hadronic Tile calorimeter for $|\eta| < 1.7$
- the fraction of jet energy measured in the third layer of the electromagnetic LAr calorimeter for $|\eta| < 3.5$
- the number of tracks assigned to the jet with a minimum transverse momentum of 1 GeV for $|\eta| < 2.5$
- the average p_T -weighted transverse distance in the η - φ plane between the jet axis and all assigned tracks with $p_T > 1$ GeV for $|\eta| < 2.5$
- the number of muon track segments ghost-associated with the jet for $|\eta| < 2.7$

In order to correct for remaining differences in the energy response between data and simulation due to imperfect simulations, an in-situ calibration is applied to data. The correction factors are derived using jets against well-measured reference objects. For example forward jets (i.e. jets with $2.5 < |\eta| < 4.5$) are balanced against central jets in dijet events, or high- p_T jets are balanced against multiple low- p_T jets in multi-jet events. The measured observable is the ratio of the transverse momentum of the jet of interest over the transverse momentum of the system of reference objects binned in p_T . This final correction factor is derived as the ratio of this ratio in data over the same ratio as measured in simulations.

The resolution of the jet energy measurement is corrected using two complementary methods as described briefly below and in more detail in [143].

The first method bases on the approximate scalar balance between the transverse momenta of the two leading jets in dijet events. The studied observable is the asymmetry of the scalar transverse momenta, i.e. the ratio of the scalar difference of the transverse momenta over the scalar sum. The relative resolution of the jet energy is proportional to the width of this asymmetry. In order to reduce the effect of soft radiations in dijet events, the two jets are required to be back-to-back in φ , and events with a third reconstructed jet are vetoed. The veto threshold on the transverse momentum of this jet is varied from 7 GeV to 20 GeV. The asymmetry is fitted linearly and extrapolated to lower thresholds. The nominal resolution is derived from the extrapolation to 0 GeV, and a systematic uncertainty is introduced to cover the effects of soft emissions. An additional correction on this asymmetry is estimated from simulations in truth dijet events and subtracted in quadrature.

A second, alternative method is based on the vector sum of the transverse momenta in dijet events and its contributions parallel and perpendicular to the bisection of the jets' momenta in the transverse plane. The effect of soft radiations is expected to be similar on average for both contributions. On the other hand, the effect of the detector resolution is not expected to be similar for the two projections. The projection along the bisection tends to be the sum of small contributions of the full transverse momenta. In contrast, the component perpendicular to the bisection is the difference between substantial contributions of the transverse momenta. So for determining this component, the nominal values are subtracted, while the uncertainties are added. Consequently, the detector resolution is expected to be larger transverse to the bisection. The resolution is derived as the squared difference of the perpendicular and parallel components along the bisection.

Pileup suppression: Since they are reconstructed from calorimeter information only, jets are very susceptible to in-time and out-of-time pileup. These pileup jets can originate from random fluctuations of pileup activity or other proton-proton collisions.

A first step in the suppression of pileup jets is the jet-area based p_T -subtraction during the calibration of jets. This correction reduces the p_T of many pileup jets below the default minimum p_T -threshold of 20 GeV and thus effectively removes these jets from the event. In order to further reduce the impact of pileup jets, many algorithms are available to select jets originating from the hard-scattering. These algorithms are described in more detail in [144] and are summarized in the following.

The algorithm used in most analyses of Run 1 data is based on the *jet vertex fraction* (JVF), which is defined for a jet j and a hard-scatter primary vertex PV^{HS} as

$$\text{JVF}(j, PV^{\text{HS}}) = \frac{\sum_m p_T(\text{track}_m, PV^{\text{HS}})}{\sum_l \sum_m p_T(\text{track}_m, PV^l)}, \quad (4.10)$$

where m iterates over all tracks assigned to the jet j and l over all primary vertices in the event. The term $p_T(\text{track}_i, PV^k)$ is defined as the transverse momentum of track_i if the track is matched to vertex PV^k and 0 otherwise.

JVF provides an estimate of the probability that a jet originates from the hard-scatter primary vertex. A problem with this definition is the intrinsic dependence on the number of primary vertices. Additional algorithms have been developed to correct for this, thus yielding results that are more robust against high pileup activity.

The *corrected jet vertex fraction* corrJVF is introduced using a modified assignment of tracks to vertices, where a priority is given to the primary vertex with the highest sum of track transverse momenta. Also, the definition in Equation (4.10) is adjusted by dividing the contributions from pileup vertices by the total number of pileup tracks times a constant, which is chosen to minimize the final dependency on the number of primary vertices.

An additional variable R_{p_T} is introduced as the ratio of transverse momenta. The numerator is the scalar sum of the transverse momenta of all tracks associated with a jet and matched to the hard-scatter primary vertex. This is divided by the fully calibrated transverse momentum of the jet. While R_{p_T} is close to 0 for pileup jets, it approximates the charged transverse momentum fraction of the jet for jets from the hard scattering process.

Both corrJVF and R_{p_T} are much less dependent on the number of primary vertices, and these variables are combined using a two-dimensional likelihood algorithm based on the k -nearest neighbours algorithm applied to simulated data in order to improve performance. The combination is called *jet vertex tagger* (JVT) and supersedes JVF as the default algorithm for suppression of pileup jets. A comparison of the performance of the individual taggers is shown in Fig. 4.12. It can be seen that JVT performs best and that the performance for JVT is mostly independent of the number of primary vertices.

All of the algorithms described above are based on the ghost associated tracks of a jet and are thus only applicable to jets with $|\eta| < 2.5$. Further algorithms have been developed to reject pileup jets also in the forward region with $2.5 < |\eta| < 4.5$. These algorithms are based on shower shape variables and comparisons of a forward jet's momentum with a simplified track-based calculation of the transverse momentum balance for each primary vertex other

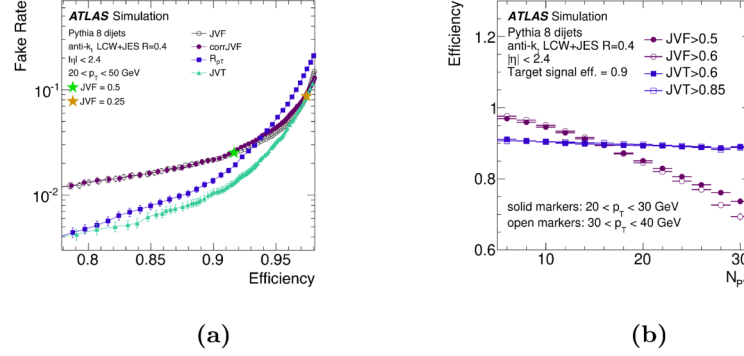


Figure 4.12: (a) Fake rate from pileup jets versus hard-scatter jet efficiency curves for JVF, corrJVF, R_{pT} , and JVT. The widely used JVF working points with cut values 0.25 and 0.5 are indicated with gold and green stars. (b) Primary vertex dependence of the hard-scatter jet efficiency for $20 < p_T < 30$ GeV (solid markers) and $30 < p_T < 40$ GeV (open markers) jets for fixed cuts of JVT (blue square) and JVF (violet circle) such that the inclusive efficiency is 90 %. From: [144]

than the primary vertex of the hard scattering. Since this is not applied, it is not discussed in detail here. More details can be found in [145].

Flavour tagging: Many analyses profit from additional information on the physics origin of jets. Such information can be provided using flavour tagging algorithms. These algorithms try to differentiate between jets originating from B -hadrons, C -hadrons, light hadrons, and gluons. For this work, only the tagging of jets originating from B -hadrons, the so-called b -tagging, is used and introduced shortly in the following. More details can be found in [146, 147]. The same algorithms can be applied with different cut-offs and optimizations to find c -quark jets. The tagging of gluons is based on other topological information, such as the number of associated tracks or the width of a jet. [148, 149]

The tagging of b -quark jets is based on the characteristic signature of a secondary vertex, which is displaced from the hard-scattering primary vertex due to the long lifetime of B -hadrons. Since precise tracking information is necessary to find these secondary vertices, b -tagging information is only available in the range of $|\eta| < 2.5$.

Three basic tagging algorithms are defined exploiting different parts of this signature. One tagger combines the impact parameters of the tracks associated with the jet using a likelihood-based approach. A second tagger is based on the explicit reconstruction of secondary vertices, where e.g. the significance of the distance from the hard-scatter primary vertex to each secondary vertex can be used as discriminating observable. The third basic tagger, called JetFitter [150] combines the information from multiple vertices and tries to reconstruct the decay chain of a B -hadron.

The default tagger in ATLAS is based on a *Boosted-Decision-Tree* (BDT) combining the output of the three taggers mentioned above. This approach was already used for analyses of data recorded during Run 1, but the current default for Run 2, called MV2, was much improved. The improvement was mainly achieved by better treatment of cases where one of the input taggers failed and optimization of the meta-parameters of the BDT.

4.3.4 Lepton isolation

For analyses using leptons, the *isolation* of a lepton can often be used to reduce the contamination due to leptons from hadron decays or other non-prompt leptons. There are different observables defined to quantify the isolation of a lepton candidate. This section summarizes the working points and the isolation variables the working points rely on in a general way. The lepton-flavour-specific definitions and further details can be found in [133, 138, 151].

The calorimeter-based E_T^{coneR} isolation variables are based on calorimeter activity in a cone around the actual signal of the lepton. The R in the superscript is a placeholder for a numeric value to indicate the actual cone size, e.g. E_T^{cone20} has a cone size of $\Delta R = 0.2$. The energies of all cells in this cone are added and corrected for the contributions of the lepton candidate, pileup, and underlying event contributions.

The track-based isolation variable p_T^{coneR} uses an equivalently defined cone. This variable is the sum of the momenta of all tracks within this cone around the lepton candidate. Only tracks that originate from the primary vertex and are not assigned to the lepton candidate are summed.

Variants with variable cone size labeled “varconeR” are introduced by using a p_T or E_T dependent cone, typically defined as the minimum of the ratio $10 \text{ GeV}/p_T$ or $10 \text{ GeV}/E_T$ and the baseline cone size indicated by the R, e.g. for $p_T^{\text{varcone20}}$:

$$\Delta R = \min(10 \text{ GeV}/p_T, 0.2) . \quad (4.11)$$

Two different working point types for the isolation are provided, typically based on relative isolations, where the previously introduced variables are divided by the candidates p_T or E_T respectively. Fixed cut working points apply strict upper thresholds on the relative isolation variables and efficiency targetted working points have p_T or E_T dependent upper thresholds designed to achieve a target efficiency. Details on the working points can be found at [152].

4.3.5 Missing transverse energy

An essential part of the reconstruction is the calculation of the imbalance of all measured transverse momenta in the event. Since the transverse momentum before the bunch crossing is negligible, the sum of all transverse momenta has to be negligible as well. The transverse momentum that is missing to balance all measured transverse momenta is called *missing transverse momentum* or to indicate, that the calculation is mostly based on the calorimeter information *missing transverse energy* (E_T^{miss}). This quantity gives a handle on neutrinos produced in the event, whose momentum that the ATLAS detector cannot measure. Another possible source of missing transverse momentum is energy-mismeasurements or false assignment to the considered primary vertex of analysis objects.

The calculation of E_T^{miss} is split into two components. First, the momenta of all analysis objects assigned to the hard scattering are summed after their corresponding calibrations to build the so-called hard-term. Additional momentum energy deposits of objects that failed the selection, but are still assigned to the hard scattering are combined to the soft-term. For both terms, the missing transverse momentum is calculated by inverting the direction of the momentum sum in the transverse plane. The total missing transverse momentum is then defined as

$$E_T^{\text{miss}} = E_T^{\text{miss, hard-term}} + E_T^{\text{miss, soft-term}} . \quad (4.12)$$

There are no fixed selection criteria for objects to enter the hard-term, but rather the recommendation to use the most inclusive level of selection for each type of object that affects the analysis. The types of objects that can enter the hard term are electrons, muons, photons, hadronically decaying taus and jets. If one of these types is not used in an analysis, it is not added to the hard-term directly but as hard-jets, since most signatures are also reconstructed as jets.

The fact that most signals are also reconstructed as jets implies, that signal would naively be double-counted. In order to avoid such double-counting, an overlap-removal procedure is applied on the level of topo-clusters for jets or the object itself for other object types.

In order to include low-energetic contributions to E_T^{miss} properly, the soft-term is included. Track information provided by the inner detector is used to distinguish contributions radiated off from the hard-scattering primary vertex from contributions originating from other primary vertices. Consequently, the soft-term can only be evaluated in the central region with $|\eta| < 2.5$. Based on the impact parameters concerning the hard-scatter vertex, each track with a minimum transverse momentum contributes to the soft term. In addition to the good

track-to-vertex association of the tracking information, the fast response reduces the amount of out-of-time pileup compared to calorimeter-based approaches to estimate the soft term.

It is again vital to remove double-counting, which can occur for tracks associated with objects that are already included in the hard-term. For electrons, photons and hadronically decaying taus, the association of tracks is based on radial distances (either $\Delta\eta$, $\Delta\varphi$, or ΔR). For muons, the matched ID-track of the muon, and for jets, ghost-associated tracks are removed from the soft term.

4.4 Processing Chain for ATLAS Data

The trigger output stream results in a data-rate of about 1 GB/s of events that are copied to long-term storage and directly passed through a processing chain described below. The stored version of the data can be used to run updated versions of the identification and reconstruction algorithms described above. Such reprocessing is applied to fix problems that might be observed during data taking and to make sure the consistent algorithms are run for the whole duration of data taking. This section presents an overview of the processing chain, highlighting parts with impact on the physics result and parts with direct contributions during the work on this thesis. More details on the data model and recommended workflows can be found in [153, 154].

4.4.1 ATLAS data model

The data of the trigger output stream are stored in a particular byte-stream format, called *RAW*. This format has an event size of about 1000 kB and is used as input for the processing chain for data recorded at ATLAS. The first step in the processing chain is to transfer the RAW data in a C++ object representation, which is easier to handle, while still keeping all information. Additionally, a first set of reconstruction algorithms is applied, to find, e.g. tracks and jets in the event. The format used for this is called *ESD* (*Event Summary Data*). Since the event size increases drastically to about 2700 kB this format is not stored but directly processed further. In the next step, different types of physics objects are reconstructed and calibrated from the input data. To reduce file size, information that is not used in standard analyses are not stored in the following format called *xAOD* (*extended Analysis Object Data*). The reduced event size is about 300 kB, which makes long-term storage of this format feasible while being easier accessible compared to the RAW format. In contrast to the software model used in Run 1, this format is not only readable in the ATLAS software framework, Athena. However, it is also compatible with the ROOT software framework[155], which is commonly used beyond ATLAS and high-energy physics.

The processing chain for simulated data is designed to re-use the same steps as for data as much as possible. This goal is achieved using a data format that can be used as an alternative to RAW when creating the ESD format.

Simulated data which was generated using a MC event generator (see Section 3.2) interfaced with Athena is stored in the EVNT format. The first processing step is to simulate how much energy is deposited in each active detector component for each event. This information is stored in a C++-object-based data format called *HITS*. The event size in this format is about 1000 kB. In the next step, this is overlaid with varying numbers of minimum bias events, either from data recorded at ATLAS or from simulations, to take the effect of pileup into account. After simulation of the effect of digitisation, the events are stored in a format called *RDO* (*Raw Data Object*), which contains the equivalent information as RAW, but is stored in C++ object representation. In a subsequent step, the effect of the triggers is simulated and stored in the same format. Once the trigger information is added, the same sets of reconstruction algorithms are applied to produce the ESD and all subsequent formats. Simulating the energy depositions to produce the HITS format is the most CPU-intense step⁵ in this chain. For this reason, the HITS format is the only format of the intermediary formats from EVNT to AOD that is usually stored.

⁵This step alone takes about 10 min per event and CPU

The main difference between simulated and recorded data in the AOD format is that for simulated data the information about the truth-level particles, i.e. particles before the detector simulation, are stored as well in order to study the effect of the detector better. Consequently, the event sizes are larger at about 400 kB for simulated data.

A release change affecting amongst other changes the reconstruction algorithms and data formats have been applied before the data taking of 2017. While it is possible to process the data from 2015 and 2016 with the new version of the software, the data taken in 2017 and 2018 is not available in the software release used in this work. For this reason, this thesis is limited to the 36.2 fb^{-1} of data recorded in 2015 and 2016.

4.4.2 Analysis workflow

The xAOD format is designed to cover all common physics analyses. Consequently, the resulting file sizes are still very sizable, and direct analysis of events in xAOD format would be very ineffective and are generally discouraged. Instead, a special framework was developed to produce derivations of the full xAOD files centrally, which are tailored for specific groups of processes. These derivations, called DxAODs use the same file format as xAODs, but typically only a subset of the available information is stored.

There are different filters which can also be combined:

skimming removing of events, e.g. to keep only events with a certain number of identified leptons,

thinning removing objects, e.g. to keep only tracks that are close to a lepton, photon, or jet,

slimming removing information about an object, e.g. to remove links from electrons to track objects that have been thinned,

augmentation calculating and adding higher-level observables, e.g. to store the number of tracks in a cone above a p_T -threshold around a jet, to allow thinning tracks around jets.

Additionally, a set of meta-data information such as the sum of event weights before applying the derivation filters, which is necessary to normalise the analysis results properly, are added.

Each physics and combined performance group developed and maintains a list of formats to cover the needs of that group. Each format should use a combination of these filters resulting in a size reduction to about 1 % or less of the original xAOD size.

A part of the work for this thesis was to maintain the derivation formats of the Standard Model physics group of ATLAS. During this time, a significant size reduction was achieved. Next to some general maintenance, the propagation of a specific set of metadata was improved as well. This metadata included the names of on-the-fly generated systematic variations of the event weight for simulated data events, which are used to estimate systematic uncertainties on the predictions. The access to these names drastically simplified the implementation of a proper combination of these weights.

The derivation format used in this work is called STDM5 (derivation format 5 of the Standard Model physics group). The derivation formats of the Standard Model group, in general, are designed to provide the minimal set of information necessary to apply all corrections by the combined performance groups for all reconstructed particle types and are thus widely usable. The STDM5 format applies a strong skimming to store only events with at least three identified leptons with $p_T > 6 \text{ GeV}$. The average size is reduced to 0.06 % for measured data in the STDM5 format. More information about the formats can be found at [156], a summary page developed by the author.

The recommended analysis workflow is to use one of the centrally produced derivation formats as input and to run a further reduction step, tailored to the specific analysis. Typically this step should include the application of the corrections recommended by the combined performance groups. The result of this step should then be in the order of $\mathcal{O}(GB)$, to allow for fast iteration in the production of new analysis results.

The recommended workflow was adopted for this work. Before this work, a framework was developed in the local institute for more straightforward implementation of physics analyses, reading derivations and writing simple ROOT-based files, called *MiniTrees*. For this thesis, this framework, called **ELCore**[157], was maintained, and an analysis was implemented based on it. With this analysis, simulated and recorded data were analysed using the reconstructed information, and for simulated data simultaneously using the truth-level information. This approach enabled fast and flexible analyses of the data in the subsequent step.

For the final analyses of the MiniTrees, a separate framework, called **CommonAnalysis-Framework** was adopted from [158] and maintained by the local working group. This framework allows to analyse the MiniTrees using small python scripts to produce histograms, write tables, study systematics, and quickly perform optimisations and cross-checks.

Chapter 5

Object and Event Selection

The typical cross-section for a VBS process is in the order of 1 fb as discussed in Chapter 3. The integrated luminosity recorded by the ATLAS detector during all data taking periods in 2015 and 2016 combined amounts to about 36 fb^{-1} . These numbers yield an estimated production of well below 100 events from a VBS process for this amount of data.

The criteria to select objects and events have to be considered carefully in order to maximize the efficiency to find the signal processes while reducing the backgrounds enough to increase sensitivity. The event and object selection criteria have been studied and optimized in the context of the official ATLAS measurement in the considered $\ell\ell\nu jj$ final state [3]. In this work, the criteria are mostly adopted from there with small deviations.

The event selection has been implemented from the ground up in the **ELCore** framework using centrally provided tools as part of the **AnalysisBase** release in version 2.4.43. All applied tools are applied in their default versions for this release. Non-default settings are described below.

5.1 Preselection

Before any complex selection criteria are applied, a basic preselection is applied that is not specific to the current analysis. This preselection ensures good detector operations to allow for meaningful comparisons between simulations and measured data.

Measured data is rejected if it was taken during a time interval of a few minutes which was flagged as problematic by independent data-quality validation. This selection is implemented in the **GoodRunsListSelectionTool**. The GoodRunLists applied are

- `data15_13TeV.periodAllYear_DetStatus-v79-repro20-02_DQDefects-00-02-02_PHYS_StandardGRL_All_Good_25ns.xml`
- `data16_13TeV.periodAllYear_DetStatus-v88-pro20-21_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml`,

where the first few characters indicate the corresponding year. The choice of these GoodRunLists reduces the total available luminosity to 36.1 fb^{-1} with a relative uncertainty of 2.1 %.

The simulated data are randomly assigned to a run during data taking according to the integrated luminosity of the individual runs. For each run, the distribution of the average number of interactions per bunch crossing is reweighted in simulations to match the observed distribution for the respective run. This is implemented in the **PileupReweightingTool**. The assignment to a run in simulations is also used in other situations where conditions have changed during data taking.

With this assignment, a year-dependent list of triggers is checked, and events which did not pass at least one of them are rejected. The considered triggers in measured data are:

2015 e: `HLT_e24_lhmedium_L1EM20VH, HLT_e60_lhmedium, HLT_e120_lhloose`

2015 μ : HLT_mu20_iloose_L1MU15, HLT_mu50

2016 e : HLT_e26_lhtight_nod0_ivarloose, HLT_e60_lhmedium_nod0,
HLT_e140_lhloose_nod0

2016 μ : HLT_mu26_ivarmedium, HLT_mu50

and for simulated data the trigger HLT_e24_lhmedium_L1EM20VH in 2015 is replaced by HLT_e24_lhmedium_L1EM18VH. These triggers correspond to the lowest un-prescaled single-lepton triggers of the corresponding periods. A per-event correction factor to harmonize trigger efficiencies in simulations and measured data is applied by combining centrally provided per-object trigger efficiencies extracted using the `AsgElectronEfficiencyCorrectionTool` and `MuonTriggerScaleFactors` tools.

Each event in measured data is checked for conditions known to cause problems, and affected events are rejected. Such problematic conditions include events with problems in the hadronic tile calorimeter, noise bursts in LAr-based calorimeters, or events after a restart of SCT or the timing, trigger and control (TTC) components.

Events are required to have at least one primary vertex reconstructed from at least two tracks and compatible with the pp interaction region. If several such vertices are present in the event, the one with the highest sum of p_T^2 of the associated tracks is selected.

5.2 Object Selection

For electrons and muons, multiple selections, i.e. sets of selection criteria, are employed. For jet candidates only one selection is used.

In the nominal event selection, there are three selections of leptons: **Baseline** leptons ultimately used to veto events with a fourth lepton, **Z-analysis** for leptons assigned to the Z boson, **W-analysis** for leptons assigned to the W^\pm boson. Special care was taken to ensure that the regions are inclusive, i.e. each lepton passing the **W-analysis** selection also passes the **Z-analysis** selection, and each lepton passing the **Z-analysis** selection also passes the **Baseline** selection.

5.2.1 Electrons

An overview of the applied selection criteria for electrons is presented in Table 5.1 and the individual criteria are discussed in detail in the following.

Corrections: Electron candidates from the central reconstruction are calibrated by smearing and scaling their energy using the `EgammaCalibrationAndSmearingTool` with the latest available recommendations.

A part of the energy of an electron candidate can leak from the central cells to the cone used for the calculation of isolation variables. The `IsolationLeakageCorrections` tool is applied on simulated and recorded data to correct for this effect. This tool also corrects for differences between simulation and recorded data.

Identification: Electron candidates in the **Baseline** selection are required to pass the `LooseAndBLayer` identification. In the **Z-analysis** selection the requirement is tightened to the **Medium** working point and the **W-analysis** selections requires candidates to pass the **Tight** working point.

Impact parameter: Using θ and the impact parameters d_0 and z_0 as defined in Section 4.2.2, the track assigned to an electron has to fulfil

$$|z_0 \cdot \sin(\theta)| < 0.5 \text{ mm}, \quad \left| \frac{d_0}{\sigma_{d_0}} \right| < 5, \quad (5.1)$$

where σ_{d_0} is the uncertainty of the d_0 measurement.

Table 5.1: Overview of different sets of object selection criteria for electrons. The overlap removal criteria correspond to steps 2 to 4 listed in Section 5.2.4.

Object selection criterium	<i>e</i> selection		
	Baseline	Z-analysis	W-analysis
Impact parameter $ z_0 \cdot \sin(\theta) < 0.5 \text{ mm}$	✓	✓	✓
d_0 significance $ d_0/\sigma_{d_0} < 5$	✓	✓	✓
Object quality	✓	✓	✓
Identification	LooseAndBLayer	Medium	Tight
Isolation	LooseTrackOnly	Gradient	Gradient
Transverse momentum p_T	$> 5 \text{ GeV}$	$> 15 \text{ GeV}$	$> 20 \text{ GeV}$
Track-based pseudo-rapidity $ \eta < 2.5$	✓	✓	✓
Calorimeter-based pseudo-rapidity $ \eta_{\text{calo}} < 2.47$	✓	✓	✓
$1.37 \leq \eta_{\text{calo}} \leq 1.52$		✓	✓
Ambiguous veto			✓
Overlap removal			
baseline $e-\mu$	✓	✓	✓
$e-e$		✓	✓
$e-j$		✓	✓

Object quality: Electron candidates are rejected if the calorimeter cluster is flagged as problematic due to issues in the electronics, e.g. problems in the high-voltage supply or if a cell in the cluster core is masked.

Ambiguity: The default list of reconstructed electrons, contains objects classified as electron and objects classified as ambiguous (see Section 4.3.1). Both classifications are allowed in the **Baseline** and **Z-analysis** selection, while in the **W-analysis** selection ambiguous objects are vetoed using the **EGammaAmbiguityTool**.

This definition differs from the object definition of [3], where the veto of ambiguous electrons was not part of the selection, but the requirement was added as event veto by hand. While this does not affect the nominal analysis, it has implications on the data-driven estimate of non-prompt backgrounds.

Pseudo-rapidity: The pseudo-rapidity η of an electron is usually defined by the assigned track. This track-based pseudo-rapidity is required to be $|\eta| < 2.5$ in all selections.

In order to veto so-called cracks, i.e. the transition region of different calorimeter subsystems, requirements on the pseudo-rapidity based on the second layer of the calorimeter η_{calo} are applied. **Baseline** electrons are required to have $|\eta_{\text{calo}}| < 2.47$ and in the **Z-analysis** and **W-analysis** selection also a veto of the crack is applied by vetoing candidates with $1.37 \leq |\eta_{\text{calo}}| \leq 1.52$.

Transverse momentum: The transverse momentum of the electron candidate p_T is required to be $p_T > 5 \text{ GeV}$ in the **Baseline** selection, $p_T > 15 \text{ GeV}$ in the **Z-analysis** selection, and $p_T > 20 \text{ GeV}$ in the **W-analysis** selection.

Isolation: Using the **IsolationSelectionTool**, electrons have to pass the **LooseTrackOnly** working point in the **Baseline** selection and the **Gradient** working point on both the **Z-analysis** and **W-analysis** selections. Both working points are efficiency targetted

Table 5.2: Overview of different sets of object selection criteria for muons. The overlap removal criteria correspond to steps 2 and 4 listed in Section 5.2.4.

Object selection criterium	μ selection		
	Baseline	Z-analysis	W-analysis
Impact parameter $ z_0 \cdot \sin(\theta) < 0.5 \text{ mm}$	✓	✓	✓
d_0 significance $ d_0/\sigma_{d_0} < 3$	✓	✓	✓
Identification	Loose	Medium	Tight
Isolation	FixedCutLoose	Gradient	Gradient
Transverse momentum p_T	$> 5 \text{ GeV}$	$> 15 \text{ GeV}$	$> 20 \text{ GeV}$
Absolute pseudo-rapidity $ \eta $	< 2.7	< 2.5	< 2.5
Overlap removal			
baseline $e\text{-}\mu$	✓	✓	✓
$\mu\text{-}j$		✓	✓

working points and differ in the target efficiency and whether the calorimeter based isolation variables are used on top of the track-based variables (see Section 4.3.4).

Efficiency corrections: Potential differences in the efficiency of selection criteria between simulated and measured data are corrected for by applying per-object correction factors. Such corrections are applied for electrons to correct for mismodelled efficiencies in the reconstruction, the identification, or the isolation requirements. The corrections are applied using the `AsgElectronEfficiencyCorrectionTool` tool.

5.2.2 Muons

The selection criteria a muon candidate has to pass in the different selections are described below and summarized in Table 5.2.

Energy corrections: The momentum of each reconstructed muon candidate is corrected using the `MuonCalibrationAndSmearingTool`. A charge-dependent correction on the muon momentum scale for the Sagitta bias is applied to data for 2016 only since this correction is not available for 2015 data.

Identification: Muon candidates in the **Baseline** selection have to pass the **Loose** working point as implemented in the `MuonSelectionTool`. The requirement is tightened to the **Medium** (**Tight**) working point for the **Z-analysis** (**W-analysis**) selection.

Impact parameter: Similar to electron candidates, the track assigned to a muon has to fulfil

$$|z_0 \cdot \sin(\theta)| < 0.5 \text{ mm}, \quad \left| \frac{d_0}{\sigma_{d_0}} \right| < 3, \quad (5.2)$$

where σ_{d_0} is the uncertainty of the d_0 measurement and d_0 , z_0 and θ are the usual impact parameters as defined in Section 4.2.2.

Pseudo-rapidity: The pseudo-rapidity η of each muon candidate is required to fulfil $|\eta| < 2.7$ to pass the **Baseline** selection. For the **Z-analysis** and **W-analysis** selections only muons with $|\eta| < 2.5$ are accepted.

Table 5.3: Overview of object selection criteria for jets. The overlap removal criteria correspond to step 1 listed in Section 5.2.4.

Object selection criterium	jet selection
Pileup reduction using JVT	Medium
Transverse momentum p_T	> 25 GeV
Absolute pseudo-rapidity $ \eta $	< 4.9
Overlap removal $j\text{-}\ell$	✓

Transverse momentum: The requirements for the transverse momentum p_T of muon objects are chosen analogously as for electron candidates. **Baseline** muons have to fulfil $p_T > 5$ GeV, **Z-analysis** selected muons have $p_T > 15$ GeV and the **W-analysis** selection requires $p_T > 20$ GeV for muons.

Isolation: Selection criteria on the isolation for muons are applied using the **Isolation-SelectionTool**. Candidates have to pass the **FixedCutLoose** working point in the **Baseline** selection as well as the **Z-analysis** selection and the **Gradient** working point in the **W-analysis** selections. The **FixedCutLoose** working point applies fixed cuts on the relative isolation variables, and the **Gradient** working point is efficiency targeted (see Section 4.3.4). For the kinematic region of muons in this selection, the **Gradient** working point was checked to be tighter than the **FixedCutLoose** working point.

Efficiency corrections: As for electron candidates, per-object correction factors are applied for muons to correct for differences in the reconstruction, the identification, as well as the isolation requirements between simulated and measured data. The corrections are applied using the **MuonEfficiencyScaleFactors** tool.

5.2.3 Jets

Jet candidates are build using the AntiKt algorithm with radius parameter $R = 0.4$ from topo-clusters of calorimeter cells as described in Section 4.3.3. Further selection requirements are detailed below and summarized in Table 5.3.

Energy corrections: Jets are calibrated using a combination of the **JetCalibrationTool**, the **JERTool** and the **JERSmearingTool** using the latest available recommendations.

Pileup suppression: The compatibility with the primary vertex is evaluated for each central ($|\eta| < 2.5$) jet using the **JetVertexTaggerTool** with the **Medium** working point. The efficiency for simulated data is corrected using a per-event scale factor to match the efficiency in measured data using the **JetJvtEfficiency** tool.

Flavour tagging: Jet candidates that pass the b -jet classification of the **BTaggingSelectionTool** with the **FixedCutBEff_70** working point are tagged as b -jets using the **MV2c10** tagger. Efficiency differences between simulated and measured data are corrected using an per-event b -tagging scale factor.

Pseudo-rapidity: In order to ensure compatibility with the detector acceptance, a requirement for the pseudo-rapidity η of each jet candidate of $|\eta| < 4.9$ is applied.

Transverse momentum: Jet candidates passing the full selection are required to have a minimum transverse momentum of $p_T > 25$ GeV.

5.2.4 Overlap removal

The selection criteria applied above can result in a duplication of objects, where the underlying object at truth level is reconstructed as two different object types. In order to avoid such duplications, the overlap of objects close to each other is removed. The metric to evaluate the distance between objects is the angular distance based on the rapidity ΔR as defined in Equation (4.8). Another essential criterion used for overlap removal is whether different objects share tracks.

The decision which object to use and which to reject is non-trivial and motivated based on physics arguments[159]. The different steps for the overlap removal are

1. Jet candidates are rejected that are close to a lepton candidate passing the **Z-analysis** selection. The minimum distance to an electron for a jet to be allowed is $\Delta R = 0.2$ and for muons it is $\Delta R = 0.4$. Jets with three or more tracks with $p_T > 500$ MeV each are not vetoed due to muons independent of their distance. This requirement on jets is applied before evaluating the jet-based per-event efficiency corrections.
2. overlap between electrons and muons is removed in case two candidates passing the **Baseline** selection share a track. In case the muon is calorimeter-tagged, the muon is rejected. Otherwise, the electron is rejected.
3. Electron candidates passing the **Z-analysis** selection are vetoed if they are within $\Delta R < 0.1$ with another electron candidate with higher p_T in the **Z-analysis** selection.
4. Electron and muon candidates passing the **Z-analysis** selection are removed if they are in a cone of $\Delta R < 0.4$ around a jet that passed the first step of overlap removal.

5.2.5 Missing transverse momentum

The missing transverse momentum in the event is calculated using the **METMaker** tool, which includes it's own implementation of overlap removal (see Section 4.3.5).

The hard term is compiled from electron and muon candidates passing the **Baseline** selection as well as all jets in the event after applying all energy corrections. The evaluation of the soft term is based on tracks.

5.3 Event Selection

Similar to the studies at truth level shown in Section 3.7, three different sets of event selection criteria are defined at detector-level. These regions of phase space are called $\ell\ell\nu jj$ region, $WZjj$ region, and signal region¹.

$\ell\ell\nu jj$ region: All events are required to have at least three leptons passing the **Z-analysis** selection. The leptons are assigned to the W^\pm and Z boson decay using the following algorithm:

1. Find all pairs of leptons with the same flavour and opposite charge (SFOC). Events without an SFOC lepton pair are rejected.
2. The SFOC lepton pair with an invariant mass closest to the nominal mass of the Z boson of $m_Z = 91.1876$ GeV is chosen and assigned to the Z boson and labelled $\ell_{Z,1}$ and $\ell_{Z,2}$ sorted by their transverse momentum p_T .
3. From the remaining leptons passing the **Z-analysis** selection, the one with highest transverse momentum p_T is chosen, assigned to the W^\pm boson, and labelled ℓ_W
4. The lepton assigned to the W^\pm boson ℓ_W is also required to pass the **W-analysis** selection. Otherwise, the event is rejected.

¹The single term “region” is used exclusively at detector level, while the term “phase space” is only used at truth level throughout this work.

This assignment splits the events in four channels for the different lepton flavour combinations; eee , $e\mu\mu$, μee , and $\mu\mu\mu$. The label for each channel is built from the flavour of each lepton in the order $\ell_W \ell_{Z,1} \ell_{Z,2}$.

At least one of the leptons assigned to a boson is required to have a transverse momentum of $p_T > 25$ GeV for runs in 2015 and $p_T > 27$ GeV in 2016 in order to match the trigger thresholds. For each trigger that accepted the event, the object reconstructed by the trigger algorithm is tested for compatibility with the three assigned leptons. The compatibility is evaluated based on the angular distance ΔR . Events for which no match can be found are discarded.

Only events with at least two jets are selected. The two jets with highest transverse momentum are selected as tagging jets j_1 and j_2 ². In order to suppress contamination from pileup jets, the two tagging jets are required to have a transverse momentum of $p_T(j) > 40$ GeV.

A second collection of jets with the same selection criteria as the default jets, other than a reduced requirement on the transverse momentum of $p_T > 20$ GeV is processed for each event. The overlap removal of jets due to close-by leptons and the pileup suppression is applied. These jets are tested for a mismeasurement due to hot calorimeter cells using the `JetCleaningTool`. All jets in this alternative collection are required to be classified as clean. Otherwise, the event is vetoed. The alternative collection is not used for anything else, and further mentions of jets always refer to the nominal jet collection.

These requirements define the $\ell\ell\ell\nu jj$ region. At this point, all single-object selection criteria are applied to the full final state. Further selection criteria of subsequent regions concern the full event, e.g. the number of b -tagged jets in the event, or combinations of objects, e.g. in the form of invariant masses or angular distances.

A comparison of measured data to simulated data with the corrections mentioned above is shown in Figure 5.1. The simulations are split according to the processes in the categories $WZjj$ -EW4, $WZjj$ -EW6, non-prompt, and prompt. The interference process $WZjj$ -EW5 is not included due to the lack of predictions, including detector effects. For both $WZjj$ -EW4 and $WZjj$ -EW6 the simulations from the SHERPA event generator are used due to the substantial number of simulated events available. The categorisation of background processes in prompt and non-prompt reflects on whether the hard process typically contains at least three leptons (prompt) or not (non-prompt). The background processes are discussed in more detail in Chapter 6.

The agreement is already very good considering the shown uncertainties for simulations only include the uncertainties due to the limited number of generated events. No theory or experimental uncertainties are shown. The most substantial discrepancies shown are the invariant mass of the Z boson candidate $M(\ell_{Z,1}\ell_{Z,2})$ in regions far off the nominal Z boson mass (see Figure 5.1a). Poorly modelled non-prompt backgrounds most likely cause this discrepancy. Such bad modelling is a typical issue in available simulations and motivates the switch to data-driven estimates for non-prompt backgrounds. Another potential reason for the discrepancy is additional contributing processes or regions of phase space not included in the simulations.

$WZjj$ region: The $WZjj$ region is based on the $\ell\ell\ell\nu jj$ region and imposes additional selection criteria designed to maximise the contributions of $W^\pm Z$ diboson production in the $\ell\ell\ell\nu jj$ final state. Events with at least four leptons passing the `Baseline` selection are vetoed to suppress contamination from processes like diboson ZZ production, where an additional lepton is expected. If any jet in the event is classified as a b -jet, the event is discarded in order to suppress contamination of $t\bar{t}$ and other t -quark based background processes.

Additional requirements are imposed for the $WZjj$ region on the transverse mass of the W^\pm boson $M_T(W) > 30$ GeV calculated using E_T^{miss} and ℓ_W in Equation (3.9). The invariant

²This selection of tagging jets is chosen differently compared to [3], where the second tagging jet is the leading jet from the set of jets in the opposite hemisphere of the detector. This assignment is not invariant against boosts along the z axis and introduces a dependence on the laboratory frame. For this reason, here the commonly used definition of tagging jets is used instead.

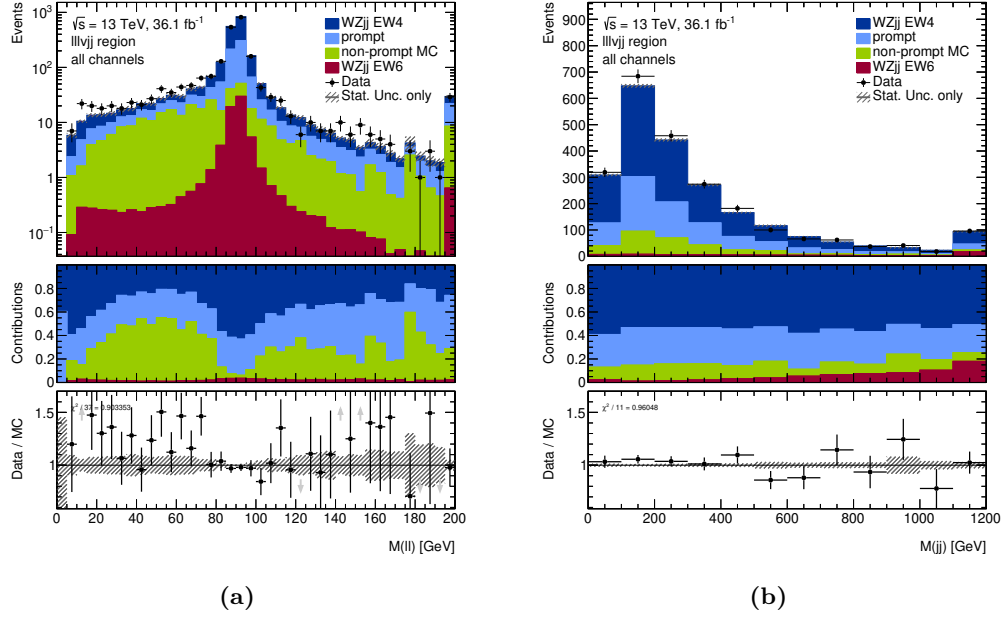


Figure 5.1: Comparison of predicted and observed differential event yields in the $lll\nu jj$ region. Figure (a) shows the event yields differentially in the invariant mass of the lepton pair assigned to the Z boson $M(\ell_{Z,1}\ell_{Z,2})$ and Figure (b) in the invariant mass of the two tagging jets. The middle inlay shows the relative composition of the total prediction and the lower inlay shows the ratio of the observed event yield with respect to the prediction. Uncertainties only show statistical uncertainties due to the limited number of simulated events.

mass of the lepton pair assigned to the Z boson $M(\ell_{Z,1}\ell_{Z,2})$ is required to be within a 10 GeV window around the nominal Z boson mass $m_Z = 91.1876$ GeV.

Each of the selection criteria from the $lll\nu jj$ region to the $WZjj$ region is visualized in Figure 5.2.

Signal region: Additional requirements are added to enhance contributions of the signal process $WZjj$ -EW6 compared to $WZjj$ -EW4 and $WZjj$ -EW5 to define the signal region. In order to achieve this, the invariant mass of the tagging jet system is required to be $M(jj) > 500$ GeV and the absolute difference in the rapidity of the tagging jets is restricted to $|\Delta Y(jj)| > 2$. Both criteria and their effects on the expected distributions are shown in Figure 5.3.

The restriction on $M(jj)$ increases the expected significance. The inlay in Figure 5.3a indicates even higher expected significances for increased selection criteria around $M(jj) > 800$ GeV. Since no systematic uncertainties are included in the evaluation of the expected significance, the final requirement was chosen to be 500 GeV to be consistent with earlier VBS analyses.

The further restriction on $|\Delta Y(jj)|$ only slightly increases the expected significance. When only considering this optimization, it would be reasonable not to apply the requirement. As discussed in Section 3.7.4 it is also reasonable to impose a requirement on $|\Delta Y(jj)|$ in order to suppress the $WZjj$ -EW5 interference contribution. This contribution is not included in the figures shown here. The value chosen here for the analysis of $|\Delta Y(jj)| > 2$ is a compromise between the suppression of $WZjj$ -EW4 and $WZjj$ -EW5 contributions and the available statistics in the signal region for control distributions.

An overview of the selection criteria applied to the different regions is presented in Table 5.4 and a comparison of event yields for each region is shown in Table 5.5. A discrepancy between the predicted and observed event yields is found for the signal region. This is not

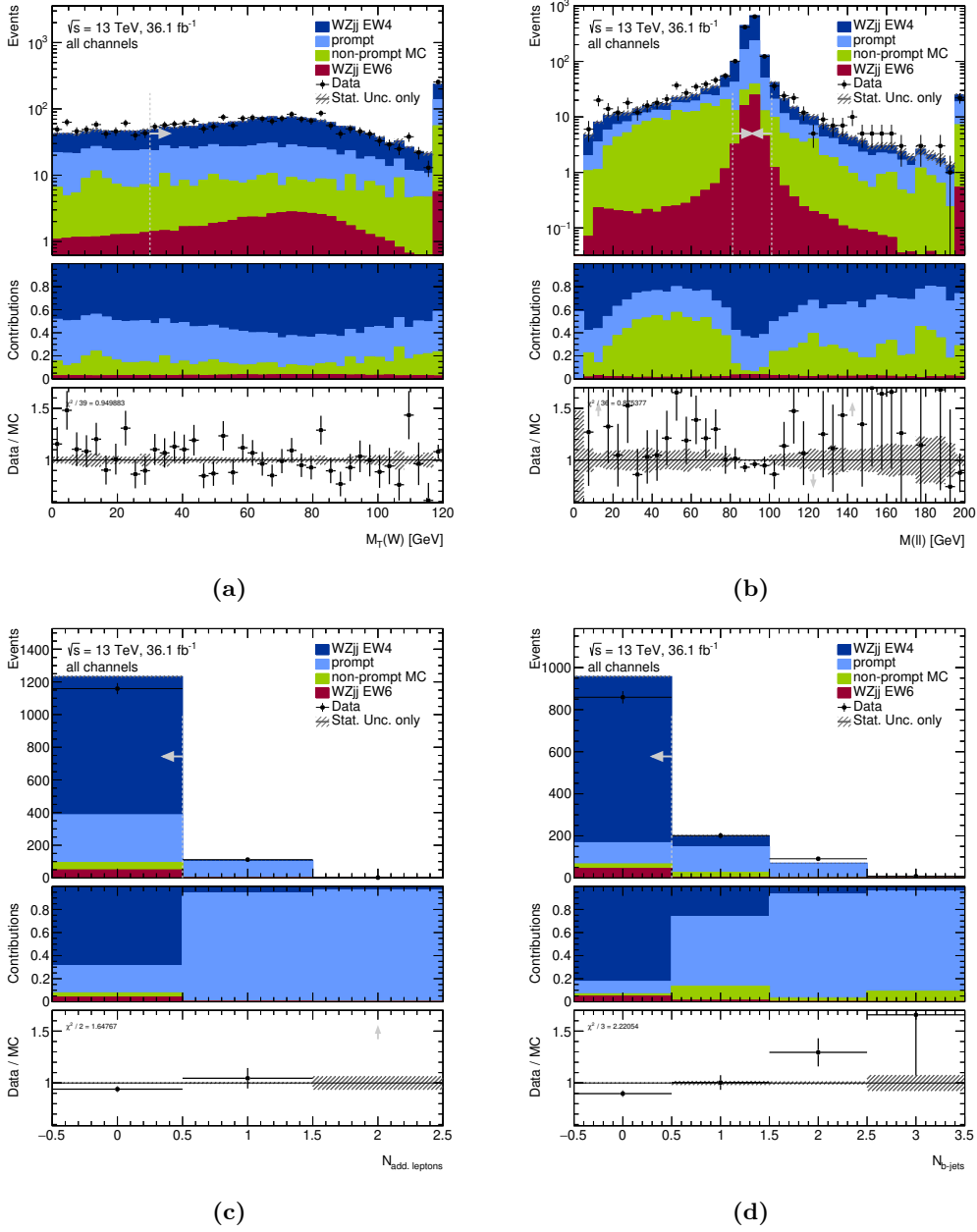


Figure 5.2: Comparison of predicted and observed differential event yields at different stages of the event selection from the $lllvjj$ region to the $WZjj$ region. The arrows indicate the applied selection criteria. In each figure, the selection criteria as indicated in the previous figures (alphabetical order of references) are applied on top of the $lllvjj$ region and the event yield is shown differentially in the observable of the applied selection criterion (grey dashed line with arrow). Figure (a) shows the selection criterion on the transverse mass of the W^\pm boson candidate, Figure (b) the invariant mass of the Z boson candidate, Figure (c) the number of additional leptons, and Figure (d) the number of jets tagged as b -jet. The middle inlay shows the relative composition of the total prediction and the lower inlay shows the ratio of the observed event yield with respect to the prediction. Uncertainties only show statistical uncertainties due to the limited number of simulated events.

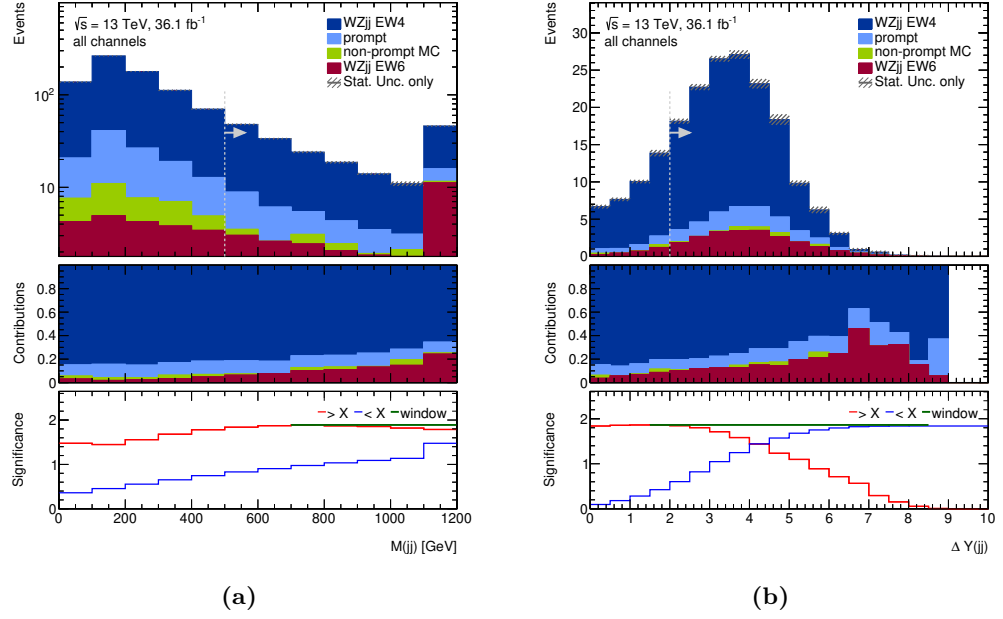


Figure 5.3: Predicted differential event yields at different stages of the event selection from the $WZjj$ region to the signal region. The arrows indicate the applied selection criteria. In each figure, a selection criterion is indicated (grey dashed line with arrow) and the event yield is shown differentially in the observable of the applied selection criterion. Figure (a) shows the selection criterion invariant mass of the tagging jets $M(jj)$ and Figure (b) the rapidity separation of the tagging jets after applying also the criterion on $M(jj)$. The middle inset shows the relative composition of the total prediction and the lower inset shows an estimate for the significance of different cut scenarios. The red (blue) curve accepts events larger (smaller) than the current value. The green line indicates the accepted ranges for the window cut with maximal significance. Uncertainties only show the statistical uncertainty due to the limited number of simulated events and the significance estimate additionally takes into account the expected statistical fluctuations of the data around the prediction.

Table 5.4: Overview of event selection criteria for the different regions at detector level. Different selection criteria are introduced in the text each region where a criterion is applied is marked with \checkmark .

Event selection criterion	$\ell\ell\nu jj$ region	$WZjj$ region	Signal region
Event cleaning	\checkmark	\checkmark	\checkmark
GoodRunList	\checkmark	\checkmark	\checkmark
Trigger	\checkmark	\checkmark	\checkmark
Primary Vertex	\checkmark	\checkmark	\checkmark
$\ell\ell\nu jj$ final state			
≥ 2 jets	\checkmark	\checkmark	\checkmark
≥ 3 Z-analysis leptons	\checkmark	\checkmark	\checkmark
One SFOC pair	\checkmark	\checkmark	\checkmark
ℓ_W is in W-analysis selection	\checkmark	\checkmark	\checkmark
Transverse momentum of leading lepton $p_T(\ell) > 25(27)$ GeV	\checkmark	\checkmark	\checkmark
Transverse momentum of subleading jet $p_T(j_2) > 40$ GeV	\checkmark	\checkmark	\checkmark
$M_T(W) > 30$ GeV		\checkmark	\checkmark
$ M(\ell\ell) - 91.1876 < 10$ GeV		\checkmark	\checkmark
four Baseline leptons veto		\checkmark	\checkmark
b -jet veto		\checkmark	\checkmark
$M(jj) > 500$ GeV			\checkmark
$\Delta Y(jj) > 2$			\checkmark

considered problematic, since the listed uncertainties only include uncertainties due to the limited amount of simulated events. Additionally, the nominal modelling of backgrounds is considered sub-optimal and is studied and improved in the following Chapter.

Table 5.5: Overview of observed and predicted event yields per group of processes in each region. The rows on the bottom also show the sum of all backgrounds, the sum of all simulation-based predictions, and the ratio of the observed yield over the predicted yield. The shown uncertainties include only uncertainties due to the limited amount of simulated events. Only the uncertainty on the ratio of the observed yield over the predicted yield also includes statistical uncertainties on the observed dataset.

	$\ell\ell\nu jj$ region		$WZjj$ region		Signal region	
Data	2340		859		129	
$WZjj$ EW4	1232	± 4	790	± 3	116	± 1
prompt	685	± 2	101.8	± 0.7	16.66	± 0.28
non-prompt MC	283	± 11	20	± 3	2.0	± 0.7
$WZjj$ EW6	69.35	± 0.30	45.39	± 0.24	22.17	± 0.17
total BKG	2201	± 12	912	± 4	135	± 1
total MC	2270	± 12	957	± 4	157	± 1
Data / MC	1.030	± 0.022	0.897	± 0.033	0.82	± 0.08

Chapter 6

Background Estimation

In Section 5.3, differential cross-section predictions are shown comparing simulations for the signal and background processes to the measured data. While the agreement was already mostly acceptable, no uncertainties were included yet. For the study of the signal model, it is crucial to have a good understanding of the background processes and their modelling uncertainties. The different sources of backgrounds are discussed in detail in this chapter, while detailed studies of the modelling uncertainties are presented in Chapter 8.

Backgrounds can be classified on whether all required final state objects were identified correctly or not. For the process studied in this work, such classification is done only for the final state leptons. If each of the three identified leptons stems from the hard-process, the background is called *prompt*. In case an identified lepton is misidentified or stems from another origin, e.g. a hadron decay, the background is labelled *non-prompt*.

Non-prompt backgrounds can usually be suppressed by tightening the object identification criteria on the signal leptons. For this reason, this background is often also called *reducible*. For a prompt background, such object identification criteria typically have similar effects as for the signal process. Hence, the contribution relative to the signal is typically constant for changes in the object identification criteria. This background is thus also called *irreducible* background. Imposing tighter selection criteria on the topology of the event can typically suppress such backgrounds.

Prompt background processes must feature the required leptons in the final state, but can also include additional leptons. The most important prompt background in this study is $WZjj$ -EW4, which was already discussed in detail in Section 2.2. It features the same final state as the $WZjj$ -EW6 signal process but has a different coupling order. The tZj background was already discussed there as well. It has the same coupling order as the signal process but includes a b -quark either in the initial or final state. Further prompt backgrounds are ZZ diboson production, where one of the four leptons is not identified in the detector. This background also includes contributions of order $\mathcal{O}(\alpha_{\text{EW}}^6)$. This background contribution is very similar to the signal process, and if one lepton is not identified in the detector, it mimicks the kinematic behaviour of the signal process. The production of three electroweak gauge bosons VVV with leptonic decays may also contribute with different boson channels as prompt background. The last group of prompt backgrounds considered in this study is associated production of a top-quark pair with an electroweak gauge boson, $t\bar{t}W$ and $t\bar{t}Z$. Examples for Feynman diagrams for the other prompt backgrounds are shown in Figure 6.1.

Non-prompt background processes are processes with fewer leptons in the final state but typically with much higher cross-section. A substantial non-prompt background process is pair production of top quarks $t\bar{t}$. In regions studied here typically both top-quarks decay leptonically, and one the b -quarks or other radiation causes the identification of the third lepton. The wrong identification can be caused either by producing a lepton in the decay chain or via misidentification.

Other top-based non-prompt backgrounds are merged and labelled single top. This process group includes s - and t -channel production of t -quarks as well as production in associa-

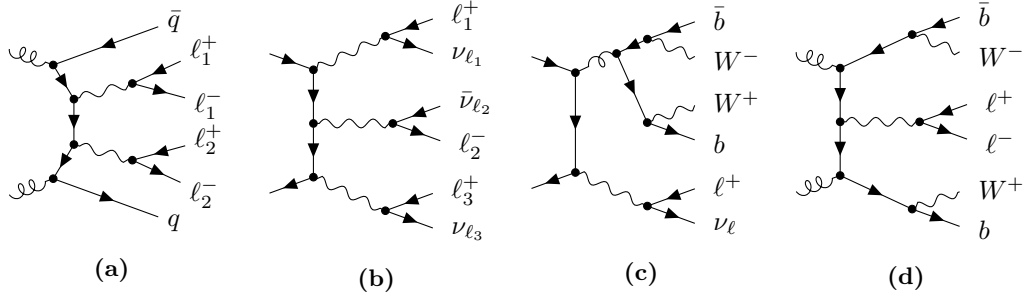


Figure 6.1: Example Feynman diagrams for the leading other prompt background processes next to the $WZjj$ -EW4 and tZj processes. Indices at leptons show flavour combinations, i.e. leptons with matching index are of the same flavour. Figure (a) shows an example for the ZZ process, Figure (b) for the VVV triboson process, and Figure (c) as well as Figure (d) for the $t\bar{t} + V$ process.

tion to a W^\pm boson. These processes can have one or two leptons in the hard process from leptonic decays of the W^\pm boson.

The third group of non-prompt processes is $Z\gamma$ diboson production, where the γ causes the identification of an electron, either directly by matching to a close-by track or by asymmetric conversion $\gamma^* \rightarrow e^+e^-$.

Other non-prompt processes can be the production of a single or two massive electroweak gauge bosons in association with jets in decay channels with only one or two final state leptons, e.g. Z +jets, $W^\pm W^\mp$ +jets, or $ZV(\rightarrow qq')$ +jets. For these processes, the third lepton typically originates from the jets, via misidentification or hadron decays.

The predicted numbers of events for each of these processes in the three regions are shown in Table 6.1 only including statistical uncertainties of the simulations. For the prompt background processes, the statistical uncertainties are, in general, much smaller compared to non-prompt processes. Since the simulations of the non-prompt backgrounds need to model the small probability to yield a non-prompt lepton identified and selected lepton on top of the other selection efficiencies. Only a very small fraction of the generated events are accepted for non-prompt backgrounds. For prompt backgrounds this overall efficiency is better.

For this reason, Monte-Carlo generated events will be used in this work for all prompt backgrounds, while the modelling of non-prompt backgrounds is based on measured data. Before the procedure for the estimation of non-prompt backgrounds is described in detail (see Section 6.2), the modelling of the prompt backgrounds is validated.

6.1 Prompt Backgrounds

In order to validate the modelling of the simulations for the most important prompt backgrounds, they are compared to measured data. For this, dedicated control regions are built to enhance the contributions of the prompt backgrounds. In order to avoid studying data in the signal region, the control regions (CR) are chosen to be orthogonal to the signal region, i.e. there is no overlap in events. This choice also reduces the contamination from signal events in the control regions, which would complicate the validation of background processes.

WZ control region: The most important background is the $WZjj$ -EW4 process. The $WZjj$ region was designed to include this process as well. This region is useful as a starting point for the definition of the control region for this process to avoid contributions of other backgrounds. The WZ control region is defined to contain all events in the $WZjj$ region, that do not pass criteria of the signal region. With this definition, orthogonality to the signal region is ensured, while the amount of statistics is maximised.

The predicted event yields are summarized in Table 6.2. The predicted purity of the $WZjj$ -EW4 process in this CR is 84% with a signal contamination of about 3%. The high

Table 6.1: Overview of observed and predicted event yields for an integrated luminosity of 36.1 fb^{-1} per processes in each region. The rows on the bottom also show the sum of all backgrounds, the sum of all simulation-based predictions, and the ratio of the observed yield over the predicted yield. The shown uncertainties include only uncertainties due to the limited amount of simulated events. Only the uncertainty on the ratio of the observed yield over the predicted yield also includes statistical uncertainties on the observed dataset.

	$\ell\ell\nu jj$ region		$WZjj$ region		Signal region	
Data	2340		859		129	
$WZjj$ -EW4	1232	± 4	790	± 3	116	± 1
$ZZjj$ -EW4	242	± 1	50.2	± 0.5	6.46	± 0.18
tZj	109.9	± 0.8	20.10	± 0.32	5.66	± 0.16
$t\bar{t}+V$	306	± 1	25.2	± 0.4	2.56	± 0.13
VVV	11.03	± 0.18	3.19	± 0.08	0.467	± 0.031
$t\bar{t}$	226	± 8	10	± 2	1.7	± 0.7
$V+\gamma$	38	± 7	5	± 1	0.29	± 0.11
$V+VV$	15	± 4	5	± 2	0	± 0
single top	5.4	± 1.0	0.38	± 0.22	0	± 0
$ZZjj$ -EW6	16.24	± 0.13	3.11	± 0.05	1.51	± 0.04
$WZjj$ -EW6	69.35	± 0.30	45.39	± 0.24	22.17	± 0.17
total BKG	2201	± 12	912	± 4	135	± 1
total MC	2270	± 12	957	± 4	157	± 1
Data / MC	1.030	± 0.022	0.897	± 0.033	0.82	± 0.08

Table 6.2: Overview of observed and predicted event yields for an integrated luminosity of 36.1 fb^{-1} per process in each control region. The rows on the bottom also show the sum of all backgrounds, the sum of all simulation-based predictions, and the ratio of the observed yield over the predicted yield. The shown uncertainties include only uncertainties due to the limited amount of simulated events. Only the uncertainty on the ratio of the observed yield over the predicted yield also includes statistical uncertainties on the observed dataset.

	WZ CR		b CR		ZZ CR	
Data	730		300		80	
$WZjj$ -EW4	674	± 3	57.2	± 0.9	2.53	± 0.14
$ZZjj$ -EW4	43.7	± 0.5	4.50	± 0.14	63.6	± 0.5
tZj	14.44	± 0.27	57.1	± 0.6	0.29	± 0.04
$t\bar{t}+V$	22.6	± 0.4	126.6	± 0.9	1.96	± 0.11
VVV	2.73	± 0.07	0.295	± 0.025	0.852	± 0.027
$ZZjj$ -EW6	1.599	± 0.034	0.203	± 0.019	4.97	± 0.07
non-prompt MC	18	± 3	27	± 4	0.38	± 0.27
$WZjj$ -EW6	23.22	± 0.17	2.51	± 0.07	0.138	± 0.012
total BKG	777	± 4	273	± 4	74.6	± 0.6
total MC	800	± 4	275	± 4	74.7	± 0.6
Data / MC	0.91	± 0.04	1.09	± 0.06	1.07	± 0.12

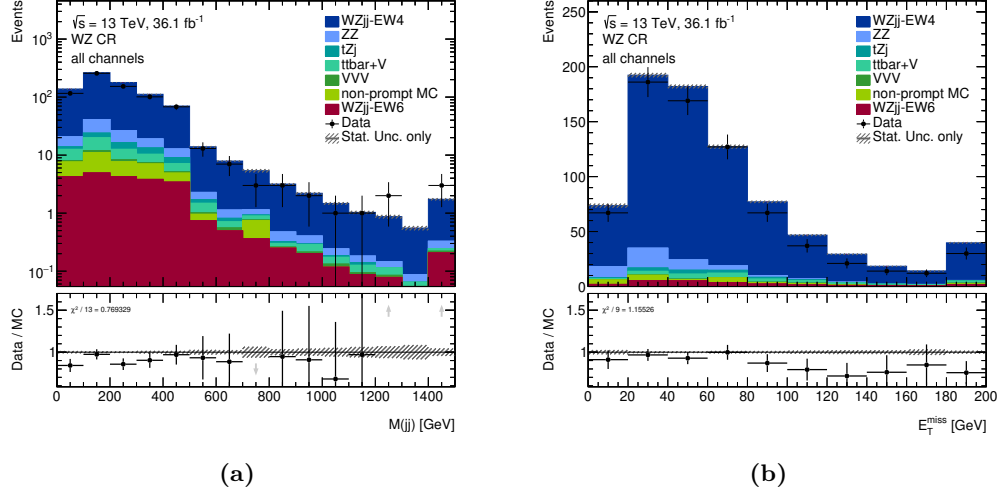


Figure 6.2: Differential event yields in the WZ control region. The nominal prediction is derived purely from simulations and shown uncertainties include only statistical uncertainties due to limited number of simulated events. Figure (a) shows the event yield differentially in the invariant mass of tagging jets $M(jj)$ and Figure (b) for the missing transverse energy. The tagging jet invariant mass $M(jj)$ has a larger step at 500 GeV introduced by the “or” combination of the $M(jj) < 500$ GeV and $|\Delta Y(jj)| < 2$ requirements in the definition of this control region.

purity allows for a proper evaluation of the modelling of the $WZjj$ -EW4 process. The ratio of observed events to the nominal prediction shows a discrepancy at the level of 10% of the total $WZjj$ -EW4 yield. As shown below in Chapter 8, this discrepancy is well within the scale uncertainties of the prediction.

Differential distributions of observed events compared to the nominal predictions from the MC generators in the WZ CR are shown in Figure 6.2. The distribution of the invariant mass of the tagging jets $M(jj)$ has a distinct step at 500 GeV. Following the definition of this control region, there is no additional requirement on absolute rapidity difference $|\Delta Y(jj)|$ for events with $M(jj) < 500$ GeV. While for events with a higher $M(jj)$, $|\Delta Y(jj)|$ has to be smaller than 2. No significant discrepancy in the shape of the $M(jj)$ distribution is observed.

The observed distribution of the missing transverse energy E_T^{miss} is softer, i.e. tends to smaller values, compared to the nominal prediction. The predicted number of events with $E_T^{\text{miss}} > 100$ GeV exceeds the observed number by about 20%, while for lower values the difference is less severe. Other backgrounds in this CR could cause this difference. For lower E_T^{miss} , the ZZ prompt backgrounds and non-prompt background are essential, for which especially the normalization might not be well modelled, too. The discrepancy might also be covered by experimental uncertainties, especially on the missing transverse momentum measurement itself, but also due to the modelling of pileup and other effects.

b control region: The contributions for the tZj and $t\bar{t}+V$ prompt background processes are significantly reduced when including the veto of b -tagged jets in the event. These processes are thus studied in a dedicated CR. This CR bases on the $WZjj$ region and is defined by inverting the veto of events with b -tagged jets.

The combined purity of this control region for the processes tZj and $t\bar{t}+V$ is about 67%. Other significant contributions are the $WZjj$ -EW4 process, which is studied in its dedicated control region, and non-prompt backgrounds which can also be modelled in a data-driven way. The predicted amount of signal contribution is about 1%.

Differential distributions in the number of b -tagged jets $N_{b\text{-jets}}$ and the missing transverse energy E_T^{miss} are shown in Figure 6.3. While the observed distributions are compatible with the predictions, no reliable conclusions can be drawn due to the limited amount of statistics

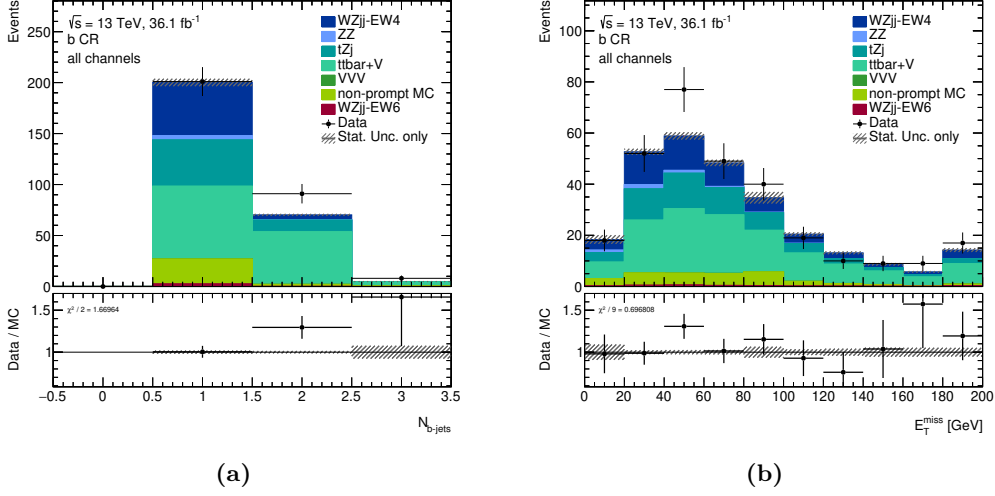


Figure 6.3: Differential event yields in the b control region. Nominal prediction is derived purely from simulations, and shown uncertainties include only statistical uncertainties due to the limited number of simulated events. Figure (a) shows the event yield differentially in the multiplicity of b -tagged jets and Figure (b) for the missing transverse energy.

in observed data and a large number of contributing processes in this control region.

ZZ control region: In order to study prompt backgrounds that include an additional fourth lepton, a control region is built based on the $WZjj$ region by inverting the veto on four **Baseline** leptons in the event. That is, this control region contains all events that have at least four **Baseline** leptons and pass all selection criteria for the usual $WZjj$ region. This control region is enriched in diboson ZZ production and is labelled ZZ CR.

The predicted and observed event yields for this control region are included in Table 6.2 as well. The purity in the ZZ process is 92% and the signal contamination is well below 1%. The observed and predicted event numbers differ by 7%, which is within the expected statistical fluctuations.

Distributions comparing the observed and expected differential event yield in the invariant mass of the tagging jets $M(jj)$ and the missing transverse momentum E_T^{miss} are shown in Figure 6.4. Within the statistical uncertainties, no significant discrepancy between the observation and the prediction is found.

6.2 Data-driven Non-prompt Estimation

The non-prompt background is expected to be much smaller compared to the backgrounds due to prompt processes, even when not including the $WZjj$ -EW4 process (see Table 5.5). Even though the overall contribution is small, the uncertainty is substantial. One driving factor is the limited amount of generated MC events. The regions of phase space considered here, are only populated by a minimal fraction of generated events of the full sample. This uncertainty is shown in e.g. Table 5.5. Additional sources of uncertainty are uncertainties in the modelling of the probability to cause an object that is misidentified as a signal lepton. This probability is typically not well described in MC simulations, and large additional uncertainties are necessary to cover this mismodelling. These uncertainties are on top of the typical theoretical uncertainties due to the PDF choice and to cover the lack of higher-order terms in the hard process calculation.

These significant uncertainties motivate a more sophisticated approach to model non-prompt backgrounds based on measured data in dedicated control regions. Many approaches are available for such data-driven (DD) estimates. An excellent overview is presented in

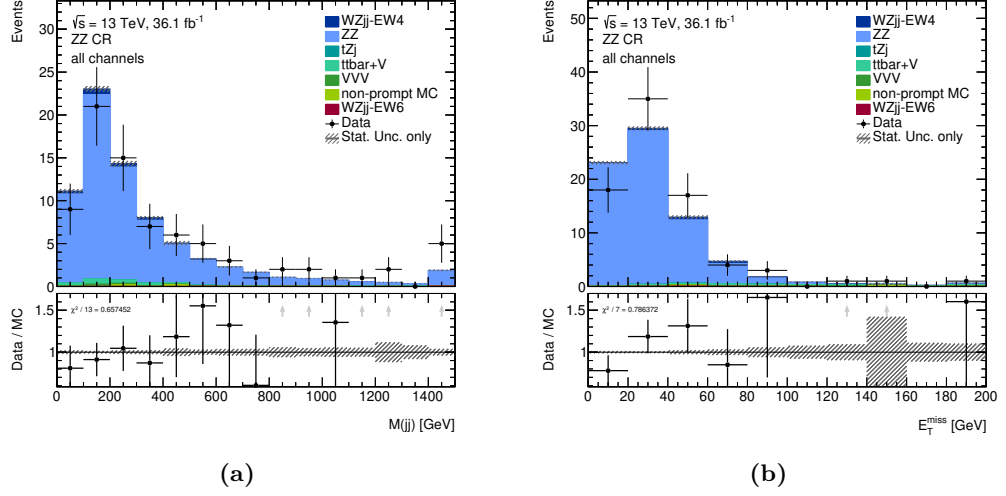


Figure 6.4: Differential event yields in the ZZ control region. Nominal prediction is derived purely from simulations and shown uncertainties include only statistical uncertainties due to limited number of simulated events. Figure (a) shows the event yield differentially in the invariant mass of the tagging jets and Figure (b) for the missing transverse energy.

[160]. The approach adopted in this work is often referred to as *Fake Factor* method or *Matrix-Method*.

The basic principle is to define a second, less-stringent selection of leptons. For this second selection, object identification criteria are lowered, while kinematic selection criteria are the same as in the nominal selection. First, the ratio of objects in the nominal to objects in this second selection is measured in a control region for each lepton flavour. This ratio is called fake factor and is applied to events, where one (or two) leptons are in the less-stringent selection rather than the nominal object selection to extrapolate to the signal region.

This approach is powerful since MC simulated predictions are only used to correct for the effect of processes with prompt objects in the nominal data-driven estimate. In addition to this, simulations are used to estimate uncertainties and to validate the procedure.

6.2.1 The fake factor method

The fake factor method necessitates the definition of an additional, less-stringent object selection. This selection has to be chosen with no overlap to the nominal object selection. Throughout this thesis, this set of object selection criteria is labeled *control* selection, as compared to the nominal *analysis* selection.¹ The relation between control and analysis object selections is visualized in Figure 6.5.

Depending on which of these detector-level selections contains a certain object, the object is classified as

- analysis lepton ℓ_A : lepton passing the analysis object selection
- control lepton ℓ_C : lepton passing the control object selection and not the analysis selection

An additional classification is performed on truth-level, based on the object type information provided by the `MCTruthClassifier`, mentioned in Section 3.7.1,

- prompt lepton ℓ_P : truth lepton tagged as `IsoElectron` or `IsoMuon`
- non-prompt lepton ℓ_N : truth lepton with other type classification.

¹These levels of object selection are often labeled “Tight” and “Loose”. These names are also often used for object identification and isolation working points, where Tight is a subset of Loose. For this reason, these names are not used here.

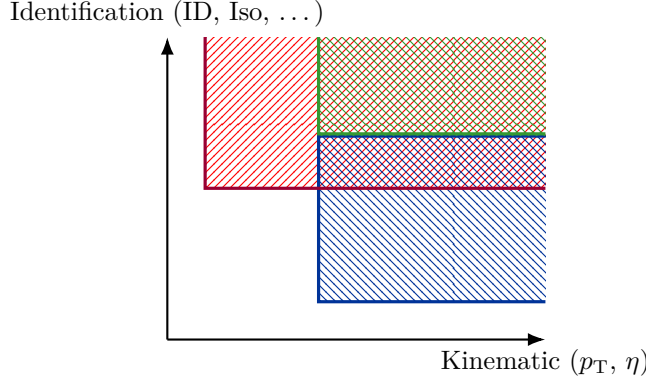


Figure 6.5: Schematic visualisation of different lepton selections. The x axis represents the strictness of all combined kinematic cuts (p_T and η), and the y axis represents a hypothetical combination of identification quality criteria (e.g. identification, isolation, overlap removal, ...) for a lepton. The green hatched (from top left to bottom right) area represents the analysis selection and the blue hatched (from top left to bottom right) area represents the control selection. The red hatched (from bottom left to top right) area represents the selection used for the veto of a fourth lepton.

Table 6.3: Overview of migration probabilities from the truth-level classification (P/N) to the detector-level classification (A/C) with the classes as described in the text.

		Detector-level	
		A	C
Truth-level	P	e	\bar{e}
	N	f	\bar{f}

These classifications are correlated, i.e. control (analysis) leptons are likely to be non-prompt (prompt). However, the probability f of a non-prompt lepton to be identified as analysis leptons is small but non-zero. Otherwise, such non-prompt background would not exist. The probability e for a prompt lepton to be identified as analysis lepton is assumed to be higher than f . The probabilities to be selected as control lepton rather than as analysis lepton are denoted using the bar notation, \bar{e} (\bar{f}) for prompt (non-prompt) leptons. The nomenclature of these migration probabilities is summarized in Table 6.3.

Based on these object classifications at detector-level (A/C) and truth-level (P/N), events with three leptons are classified by stringing together the labels of the individual leptons. The order of the labels follows the assignment of the leptons to the bosons. The first index is for the lepton assigned to the W^\pm boson, the second for the leading lepton assigned to the Z boson and the third for the subleading lepton assigned to the Z boson. In this notation, the number of events with three leptons in the analysis selection is N_{AAA} , while the number of events with three prompt leptons is written as N_{PPP} .

Using the migration probabilities, N_{AAA} can be decomposed according to the truth-level classification

$$N_{AAA} = e_1 e_2 e_3 N_{PPP} + \quad (6.1)$$

$$f_1 e_2 e_3 N_{NPP} + e_1 f_2 e_3 N_{PNP} + e_1 e_2 f_3 N_{PPN} + \quad (6.2)$$

$$f_1 f_2 e_3 N_{NNP} + f_1 e_2 f_3 N_{NPN} + e_1 f_2 f_3 N_{PNN} + \quad (6.3)$$

$$f_1 f_2 f_3 N_{NNN}, \quad (6.4)$$

where the migration matrices have an additional index i indicating which lepton they are associated to in the aforementioned order.

The right-hand side is split into multiple lines according to the number of non-prompt leptons. Since typically $f < e$, the overall event yields decrease with increasing number of

wrongly identified non-prompt leptons. Due to this decrease with $\frac{f}{e}$, the contribution from N_{NNN} is small and can be neglected. The first line, Equation (6.1), indicates the prompt contribution to the overall event number $N_{\text{AAA}}^{\text{prompt}}$. The remaining lines contribute to the total non-prompt background contribution $N_{\text{AAA}}^{\text{non-prompt}}$, which can be approximated when neglecting contributions from N_{NNN} .

$$N_{\text{AAA}}^{\text{non-prompt}} = N_{\text{AAA}} - e_1 e_2 e_3 N_{\text{PPP}} = N_{\text{AAA}} - N_{\text{AAA}}^{\text{prompt}}. \quad (6.5)$$

This procedure can be generalized to include *control lepton regions* where a control lepton was selected and assigned rather than an analysis lepton, using matrix notation for the possible combinations of control and analysis leptons

$$\begin{pmatrix} N_{\text{AAA}} \\ N_{\text{CAA}} \\ N_{\text{ACA}} \\ N_{\text{AAC}} \\ N_{\text{ACC}} \\ N_{\text{CAC}} \\ N_{\text{CCA}} \end{pmatrix} = \begin{pmatrix} e_1 e_2 e_3 & f_1 e_2 e_3 & e_1 f_2 e_3 & e_1 e_2 f_3 & e_1 f_2 f_3 & f_1 e_2 f_3 & f_1 f_2 e_3 \\ \bar{e}_1 e_2 e_3 & \bar{f}_1 e_2 e_3 & \bar{e}_1 f_2 e_3 & \bar{e}_1 e_2 f_3 & \bar{e}_1 f_2 f_3 & \bar{f}_1 e_2 f_3 & \bar{f}_1 f_2 e_3 \\ e_1 \bar{e}_2 e_3 & f_1 \bar{e}_2 e_3 & e_1 f_2 \bar{e}_3 & e_1 \bar{e}_2 f_3 & e_1 f_2 f_3 & f_1 \bar{e}_2 f_3 & f_1 f_2 \bar{e}_3 \\ e_1 e_2 \bar{e}_3 & f_1 e_2 \bar{e}_3 & e_1 f_2 \bar{e}_3 & e_1 e_2 f_3 & e_1 f_2 f_3 & f_1 e_2 f_3 & f_1 f_2 \bar{e}_3 \\ e_1 \bar{e}_2 \bar{e}_3 & f_1 \bar{e}_2 \bar{e}_3 & e_1 f_2 \bar{e}_3 & e_1 \bar{e}_2 f_3 & e_1 f_2 f_3 & f_1 \bar{e}_2 f_3 & f_1 f_2 \bar{e}_3 \\ \bar{e}_1 e_2 \bar{e}_3 & \bar{f}_1 e_2 \bar{e}_3 & \bar{e}_1 f_2 \bar{e}_3 & \bar{e}_1 e_2 f_3 & \bar{e}_1 f_2 f_3 & \bar{f}_1 e_2 f_3 & \bar{f}_1 f_2 \bar{e}_3 \\ \bar{e}_1 \bar{e}_2 e_3 & \bar{f}_1 \bar{e}_2 e_3 & \bar{e}_1 f_2 e_3 & \bar{e}_1 \bar{e}_2 f_3 & \bar{e}_1 f_2 f_3 & \bar{f}_1 \bar{e}_2 f_3 & \bar{f}_1 f_2 e_3 \end{pmatrix} \begin{pmatrix} N_{\text{PPP}} \\ N_{\text{NPP}} \\ N_{\text{PNP}} \\ N_{\text{PPN}} \\ N_{\text{PNN}} \\ N_{\text{NPN}} \\ N_{\text{NNP}} \end{pmatrix}. \quad (6.6)$$

In order to relate the possible combinations of control and analysis leptons, extrapolation is necessary. A convenient extrapolation is to define the fake factor F_i as

$$F_i = \frac{f_i}{e_i}, \quad (6.7)$$

which is equal to the ratio of non-prompt leptons passing the analysis selection over the control selection. Multiplying the appropriate fake factor to the number of events with a control lepton extrapolates to the number of events, where this lepton is in the analysis selection. By multiplying with the corresponding fake factors in each line and neglecting terms of order $\mathcal{O}(f^3)$, the migration matrix in Equation (6.6) can be rewritten as

$$\begin{pmatrix} e_1 e_2 e_3 & f_1 e_2 e_3 & e_1 f_2 e_3 & e_1 e_2 f_3 & e_1 f_2 f_3 & f_1 e_2 f_3 & f_1 f_2 e_3 \\ \bar{e}_1 e_2 e_3 F_1 & \bar{f}_1 e_2 e_3 & \bar{e}_1 f_2 e_3 F_1 & \bar{e}_1 e_2 f_3 F_1 & 0 & \bar{f}_1 e_2 f_3 & \bar{f}_1 f_2 e_3 \\ e_1 \bar{e}_2 e_3 F_2 & f_1 \bar{e}_2 e_3 F_2 & e_1 f_2 \bar{e}_3 & e_1 \bar{e}_2 f_3 F_2 & e_1 f_2 f_3 & 0 & f_1 f_2 \bar{e}_3 \\ e_1 e_2 \bar{e}_3 F_3 & f_1 e_2 \bar{e}_3 F_3 & e_1 f_2 \bar{e}_3 F_3 & e_1 e_2 f_3 & e_1 f_2 f_3 & f_1 e_2 f_3 & 0 \\ e_1 \bar{e}_2 \bar{e}_3 F_2 F_3 & 0 & e_1 f_2 \bar{e}_3 F_3 & e_1 \bar{e}_2 f_3 F_2 & e_1 f_2 f_3 & 0 & 0 \\ \bar{e}_1 e_2 \bar{e}_3 F_1 F_3 & \bar{f}_1 e_2 \bar{e}_3 F_3 & 0 & \bar{e}_1 e_2 f_3 F_1 & 0 & \bar{f}_1 e_2 f_3 & 0 \\ \bar{e}_1 \bar{e}_2 e_3 F_1 F_2 & \bar{f}_1 \bar{e}_2 e_3 F_2 & \bar{e}_1 f_2 e_3 F_1 & 0 & 0 & 0 & \bar{f}_1 f_2 e_3 \end{pmatrix}. \quad (6.8)$$

With this migration matrix, it can easily be verified, that adding lines 1, 5, 6, and 7 and subtracting lines 2, 3, and 4 yields 0 for every column but the first. When applying this procedure to the full equation and ordering in the fake factors, one obtains

$$N_{\text{AAA}} - e_1 e_2 e_3 N_{\text{PPP}} \approx (N_{\text{CAA}} - \bar{e}_1 e_2 e_3 N_{\text{PPP}}) F_1 - (N_{\text{ACC}} - e_1 \bar{e}_2 \bar{e}_3 N_{\text{PPP}}) F_2 F_3 + \quad (6.9)$$

$$(N_{\text{ACA}} - e_1 \bar{e}_2 e_3 N_{\text{PPP}}) F_2 - (N_{\text{CAC}} - \bar{e}_1 e_2 \bar{e}_3 N_{\text{PPP}}) F_1 F_3 + \quad (6.10)$$

$$(N_{\text{AAC}} - e_1 e_2 \bar{e}_3 N_{\text{PPP}}) F_3 - (N_{\text{CCA}} - \bar{e}_1 \bar{e}_2 e_3 N_{\text{PPP}}) F_1 F_2 \quad (6.11)$$

Each term with a common fake factor can be identified as the non-prompt contribution in a selection using equivalent arguments as in Equation (6.5)

$$\begin{aligned} N_{\text{AAA}}^{\text{non-prompt}} &\approx N_{\text{CAA}}^{\text{non-prompt}} F_1 + N_{\text{ACA}}^{\text{non-prompt}} F_2 + N_{\text{AAC}}^{\text{non-prompt}} F_3 \\ &\quad - N_{\text{ACC}}^{\text{non-prompt}} F_2 F_3 - N_{\text{CAC}}^{\text{non-prompt}} F_1 F_3 - N_{\text{CCA}}^{\text{non-prompt}} F_1 F_2. \end{aligned} \quad (6.12)$$

In order to obtain a data-driven estimate for $N_{\text{AAA}}^{\text{non-prompt}}$, the non-prompt contributions in each of the regions with a selected control lepton can be obtained by subtracting the prompt

contribution estimated using MC predictions from the observed number of events in data. Additionally, the fake factors F_i can be obtained from data in a region orthogonal to the signal region. For instance, the term associated with the CAA region can be written as

$$N_{\text{CAA}}^{\text{non-prompt}} F_1 \rightarrow \left(N_{\text{CAA}}^{\text{data}} - N_{\text{CAA}}^{\text{prompt MC}} \right) F_1^{\text{data}}. \quad (6.13)$$

For leptons of a different flavour, the underlying physics for the misidentification differ. For instance, a photon is more likely to be identified as an electron rather than as muon. This leads to differences in the fake factors for different lepton flavours. The fake factor is known to also depend on the transverse momentum p_T of the considered lepton. In this study, an additional complication arises from the different selection criteria for leptons assigned to the W^\pm boson and Z boson, respectively. The fake factor for a lepton candidate as used in this work depends on

- the lepton flavour
- the transverse momentum
- which boson the candidate is assigned to, i.e. whether the candidate is required to pass the **W-analysis** or **Z-analysis** selection.

The fake factor also depends on the origin of the non-prompt lepton, i.e. the underlying truth object that was misidentified. A photon has a different probability of being misidentified as electron compared to a light or even a heavy jet. Since this information is not accessible in data, it is not easily possible to correct for such effects.

6.2.2 Fake factor determination

When labelling the total number of identified non-prompt lepton candidates as N_N , the definition of the fake factor yields for large N_N

$$F = \frac{f}{\bar{f}} \approx \frac{N_{N \rightarrow A}}{N_N} \cdot \frac{N_N}{N_{N \rightarrow C}} = \frac{N_{N \rightarrow A}}{N_{N \rightarrow C}}, \quad (6.14)$$

where $N_{N \rightarrow A}$ ($N_{N \rightarrow C}$) is the number of non-prompt leptons passing the analysis (control) selection. Due to statistical fluctuations in the observed or predicted event numbers, the event yield ratio on the right-hand side is only an approximation for the true fake factor F .

While the fake factors F_i can, in principle, be determined using Equation (6.14) from simulated non-prompt events, this approach has similar disadvantages as a full non-prompt estimation from pure simulations. It is very beneficial to extract the fake factors from data in a control region orthogonal to the signal region.

The goal is to find a control region, where one lepton has a high probability of being non-prompt lepton, labelled probe lepton. In contrast, potential other leptons ideally have a high purity of prompt leptons. The fake factor is commonly estimated by evaluating the ratio of events in a control region where the probe lepton is selected as analysis lepton $N_{N \rightarrow A}^{\text{CR}}$ over those where the probe lepton is in the control selection $N_{N \rightarrow C}^{\text{CR}}$.

The fake factor method is based on a fundamental assumption. It is assumed that the fake factor as determined in the control region F^{CR} is the same as in the signal region F^{SR} , which is the fake factor that is applied to evaluate the final data-driven non-prompt prediction. This assumption needs to be tested using MC simulations.

This assumption is particularly problematic in combination with the different non-prompt origins. If the distribution of different non-prompt origins is significantly different in the control region and the signal region, it is not valid to assume that the fake factors are in agreement.

However, since adjusting the control object selection has different effects on the fake factors for different non-prompt origins, the definition can be tuned to decrease the effect of different non-prompt origin compositions. This approach has not been followed here but offers potential for improvement of the method.

With these general remarks, a control lepton selection and control region is chosen and summarized below.

Control lepton selection.

The procedure to define the control lepton selections is to construct an additional lepton selection by loosening the **Z-analysis** selection first. These new selections include the **Z-analysis** and **W-analysis** selections. The control lepton selection for leptons assigned to the Z boson is labelled **Z-control** and is defined as all leptons passing the looser selection, but not the **Z-analysis** selection. Similarly, the **W-control** selection is defined as all leptons passing the same, looser selection, but not the **W-analysis** selection.

For electrons, the looser selection is defined based on the **Z-analysis** selection. The identification working point is lowered to **LooseAndBLayer**. Additionally, the isolation requirement of passing the **Gradient** working point is replaced by a fixed minimum discussed below.

For muons, the identification working point is the same as in the **Z-analysis** selection. The isolation working point of **FixedCutLoose** is replaced as discussed below. Additionally, the restriction on the significance of the d_0 measurement $\left| \frac{d_0}{\sigma_{d_0}} \right|$ is not applied.

An isolation dependent bias on the lepton candidates transverse momentum p_T was observed during the work on this thesis. A similar bias was observed in [2], where a p_T was corrected to include the transverse momenta in a cone around the lepton track. This procedure is applied in this work as well using the replacement

$$p_T \rightarrow p_T + p_T^{\text{varcone30}}, \quad (6.15)$$

where $p_T^{\text{varcone30}}$ is the sum of transverse momenta of all tracks within a cone around the lepton's track. The cone size decreases with increasing transverse momentum (see Section 4.3.4). The correction is only applied for lepton candidates failing the isolation working point required in the respective **Z-analysis** selection. The correction only scales the momentum without adjusting its direction. A correction including the full four-momenta of the surrounding tracks is expected to perform better, but the necessary information is not available in the derivations used here.

For all subsequent steps, including the $p_T > 15$ GeV requirement, only the corrected p_T is used unless stated otherwise. In order to ensure, this p_T -correction does not introduce a bias in combination with the skimming applied on derivation level (see Section 4.4.2), a minimum of $p_T > 9$ GeV is required before applying the replacement. This minimum ensures the lepton candidate still passes the object selection criteria used in the skimming at derivation level.

In order to ensure the effect of the isolation-based p_T correction is not too large, a minimum isolation is required in the looser selection. For both lepton flavours a fixed maximum of 0.5 for the isolation variable $p_T^{\text{varcone30}}/p_T$ is required. The bias as well as the effect of the corrections are described in more detail below (see Section 6.2.3).

Another change compared to the **Z-analysis** selections applied for both lepton flavours is not to apply the overlap removal in the **Z-control** and **W-control** selections. When the overlap removal procedure is applied, many control leptons are rejected, which increases the fake factor and ultimately increases the statistical uncertainty. Not applying the overlap removal has important implications which are discussed below.

The control lepton selections used in this work are based on [3] but differ in significant ways, e.g. the isolation-based p_T correction, which was not applied before. The working points were used as in [3], but special care was taken to achieve consistency between the control regions and the signal region, especially for the overlap removal.

Also, the change to veto ambiguous electrons in the **W-analysis** selection rather than to apply this as event selection criterion covers this effect in the fake factor rather than to apply the criterion on misidentified but assigned leptons.

A full overview of the applied selection criteria for both muons and electrons is shown in Section E.1 in the appendix.

Effect of control leptons on event selection: The definition of control lepton regions, where control leptons are selected and assigned is not trivial. Some important aspects have to be considered in more detail in order to avoid biases.

Since the set of control leptons is not a subset of the **Baseline** leptons used for the veto of additional leptons (see Figure 6.5), events are allowed to have additional control leptons. However, if the assignment would be trivially extended to consider analysis and control leptons in the same way, the control lepton could be assigned, and an additional analysis lepton would be unassigned. In this case, an event that previously was in the nominal selection because the additional control lepton potentially does not veto the event would now be rejected due to the additional analysis lepton. In order to not affect the nominal selection, the assignment of leptons is extended in a more sophisticated manner.

The goal is to assign analysis leptons preferably. This goal is achieved by considering only analysis leptons for the assignment by default. Only in events with less than three analysis leptons, control leptons are recursively included ordered with decreasing transverse momentum until three leptons are available. So, for instance, in events with two analysis leptons, only the leading control lepton is considered. If less than three leptons that are either control or analysis selected are present in the event, the event is vetoed. No further adjustments to the assignment are needed.

Since no overlap removal is applied on control leptons, control leptons are often accompanied by close-by jets. These typically have the same underlying origin, e.g. if a non-prompt muon originates from a hadron decay, a control muon can be selected close by the original jet. The close-by jets biases the jet multiplicity and can as well affect the selection of the full event, in cases where they are selected as tagging jet. In order to avoid such double-counting, the effect of the overlap removal is replicated. In cases where a control lepton is selected and assigned, all jets within a cone of $\Delta R < 0.4$ are removed from the event.

The test, whether the object reconstructed by the trigger algorithm matches any of the selected leptons, is restricted to analysis leptons only. Consequently, the event-wide factor to correct for differences in the trigger efficiency between simulated and measured data is adjusted only to consider analysis leptons. Since none of the centrally provided correction factors is applicable for control leptons, no correction factors are applied.

In further aspects, the selection is adjusted by extending the considered leptons to the assigned leptons independent on whether they are in the control or analysis selection.

Control region for fake factor determination.

Based on the predicted yields from MC simulations (see Table 6.1), the process that contributes the most to the non-prompt background in each of the three main considered regions of phase space is $t\bar{t}$ production. However, due to the limited number of generated events, especially in the Signal region, no hard conclusions can be drawn from the integrated yields alone.

The predictions of MC simulations for the different non-prompt processes are compared differentially in the invariant mass of the tagging jets in the $\ell\ell\nu jj$ region and the $WZjj$ region in Figure 6.6. The signal region is a subset of the events with high $M(jj)$ in the $WZjj$ region. These numbers also do not allow for reliable conclusions on the contribution directly, due to the lack of other event selection criteria. However, by evaluating the trend of the relative composition of the individual non-prompt processes with increasing $M(jj)$, the plausibility can be checked. The contribution of $V + VV$ decreases with increasing $M(jj)$, while the single top contribution seems to follow a similar trend as $t\bar{t}$, but is overall very small on all phase spaces. The contribution of $V + \gamma$ decreases with increasing $M(jj)$ in the $WZjj$ region. The behaviour in the $\ell\ell\nu jj$ region is more complicated than the $WZjj$ region, but since the $WZjj$ region is closer to the signal region and there seems to be enough statistics for $t\bar{t}$ and $V + \gamma$, the $WZjj$ behaviour is considered more valuable. In all considered regions, the contribution of $V + \gamma$ to the total non-prompt prediction is between 10 – 20 %.

After having considered this plausibility check, the control region for the determination of the fake factor is designed to have a high contribution of $t\bar{t}$. The motivation is to focus on fully-leptonic $t\bar{t}$ decays as tag and search for an additional third lepton as the probe. In order to maximise the available number of events, the control region is defined with a minimal set of event selection criteria only.

The baseline is the typical event preselection of event cleaning, GoodRunList, Primary

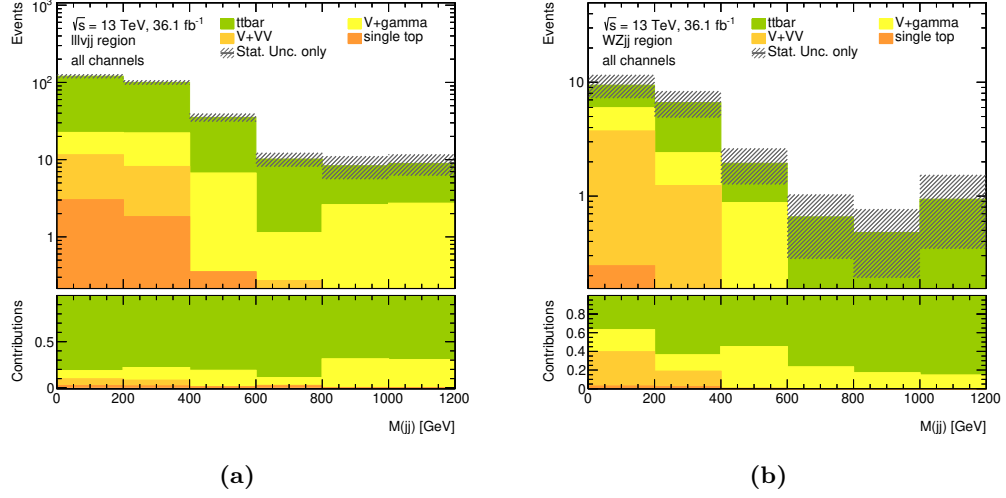


Figure 6.6: Differential event yield in the invariant mass of the tagging jets as predicted by simulations for the non-prompt background by process. Uncertainties only contain statistical uncertainties due to the limited number of simulated events. Figure (a) shows the differential event yield in the $llvjj$ region and Figure (b) for the $WZjj$ region.

vertex, and Trigger requirements using the same triggers as in the nominal selection. No additional requirements on the presence of jets are imposed. The selection of leptons considered in the assignment is applied in the same way as described above, i.e. analysis leptons are preferred, and control leptons are added until three leptons are available for the assignment. Events with less than three leptons, either passing the control or analysis selection are vetoed. If an event has more than three leptons passing the **Baseline** selection (see Section 5.2), the event is vetoed. This combination of criteria implies, only events with exactly three leptons in the assignment are considered. The leading lepton of those three is required to have a minimum transverse momentum of $p_T > 25$ GeV (> 27 GeV) for events measured in or assigned to 2015 (2016).

If a lepton pair of the same flavour and opposite charges is found among the considered leptons, the event is vetoed. In dileptonic $t\bar{t}$, the leptons typically have opposite charge, but the flavour is independent of each other. This requirement can be fulfilled by events where the three considered leptons have the same charge, or where two leptons have identical flavour and charge and the third lepton has different flavour and charge, e.g. $e^+e^+\mu^-$. Since one expects a pair of oppositely charged leptons in $t\bar{t}$ production, events with three leptons of the same charge are vetoed. The lepton of the flavour that only occurs once is selected as part of the tag lepton pair and required to pass the **Z-analysis** selection.

For the determination of the fake factor a probe lepton has to be chosen from the two leptons of identical flavour and charge. The fake factor is determined as the number of events where the probe lepton is an analysis lepton over the number of events where the probe lepton is a control lepton in ranges of the transverse momentum of the probe lepton. The other lepton is considered a tag lepton as well and is required to pass the **Z-analysis** selection. Contributions of events with misassigned leptons, i.e. events where the probe is actually a prompt lepton, have to be subtracted from data in order to estimate the data-driven non-prompt contribution.

One can think of several different algorithms to select a probe lepton from the two leptons with the identical flavour. Assignment algorithms differ in their probability to assign the leptons correctly. The higher the fraction of misassigned events, the higher the relative statistical uncertainties on the data-driven non-prompt contribution. Independent of the assignment algorithm, the fraction of non-prompt leptons assigned as the probe and the fraction of prompt leptons assigned as probe add up to 1.

An alternative treatment is to accept and evaluate both possible cases. In this case, each

Table 6.4: Overview of event selection criteria for the $t\bar{t}$ control region used for the determination of the fake factor.

Event selection criterion	$t\bar{t}$ control region
Event cleaning	✓
GoodRunList	✓
Trigger	✓
Primary Vertex	✓
Transverse momentum of leading lepton	✓
$p_T(\ell) > 25(27)$ GeV	✓
≥ 3 Z-analysis or Z-control leptons	✓
four Baseline leptons veto	✓
No SFOC pair	✓
Not three leptons of same charge	✓
tag: OFOC pair Z-analysis leptons	✓

assignment probability is 100%. This approach maximises the fraction of prompt leptons that have to be subtracted from data and is equivalent to an entirely random assignment. However, the benefit of this approach compared to the random assignment is an enhancement of the available statistics by a factor of 2, since no non-prompt candidates are rejected.

Other assignment algorithms cannot benefit from such an enhancement and need to compensate for the reduction of non-prompt leptons with significantly enhanced reduction of the prompt contribution due to a low misassignment probability. Several assignment algorithms were tested, and no algorithm was found that achieved a sufficient increase in the probability for a correct assignment.

An additional consideration is the introduction of potential bias due to the assignment. For instance, a hypothetical assignment based on the minimal angular distance to any jet in the event would select a biased subset of non-prompt events and is expected to bias the fake factor determination and consequently the data-driven estimate.

In this work, the approach to consider both possible assignments was followed. When determining the fake factor for the W-analysis selection, the potential probe lepton, that is not assigned is required to pass the W-analysis selection as well. This additional requirement is necessary to reduce the contamination of prompt leptons due to the misassignment.

The control region is labeled $t\bar{t}$ control region and based on a control region used in [161]. The selection requirements are summarized in Table 6.4. In this region, fake rates are determined for three different scenarios

ttbar Only events from simulations of the top-pair production ($t\bar{t}$) are used and events where the probe, i.e. the fake candidate lepton is classified by the `MCTruthClassifier` as `IsoElectron` or `IsoMuon`, i.e. prompt leptons are vetoed. No requirement for the classification of the tag leptons is imposed.

Non-prompt MC Events from simulations of all considered processes are used and events where the probe, i.e. the fake candidate lepton is classified by the `MCTruthClassifier` as `IsoElectron` or `IsoMuon` are vetoed. No requirement for the classification of the tag leptons is imposed.

Reduced data The fake factors are determined from the observed events in data reduced by the prompt contributions. The prompt contributions are evaluated from simulated events of all considered processes where the probe lepton is required to be classified by the `MCTruthClassifier` as `IsoElectron` or `IsoMuon`.

The resulting fake factors are compared for all three scenarios in Figure 6.7 and listed in detail for the scenario of reduced data in Table 6.5.

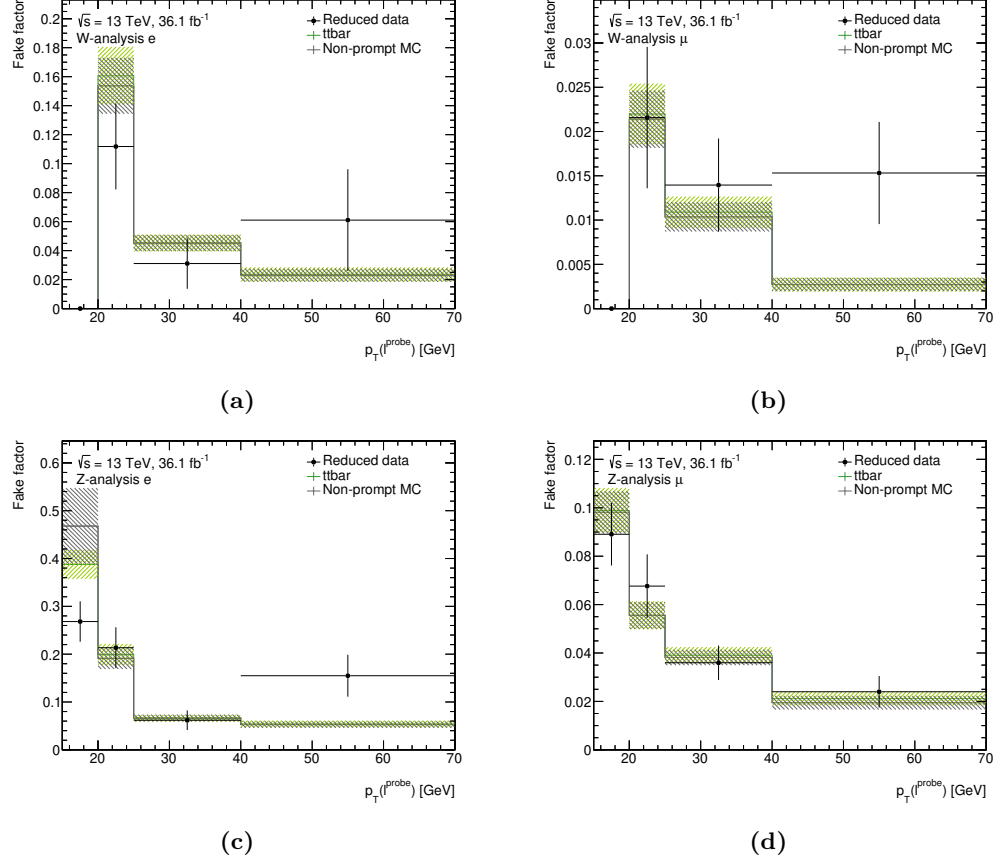


Figure 6.7: Comparison of fake factors for different transverse momentum ranges for the $t\bar{t}$ (green), the non-prompt MC (grey), and the reduced data scenarios (black data points). Figures in top row show fake factors for **W-analysis** selection and in bottom row for the **Z-analysis** selection. Fake factors for electrons (muons) are shown in the Figures on the left (right).

Table 6.5: Detailed overview of the fake factors derived for different ranges in the transverse momentum of the fake candidate p_T^{probe} in GeV in the reduced data scenario, i.e. from data. The header row shows the momentum range and other rows show the fake factors for each momentum range. The first column indicates for which object selection level and the second column for which lepton flavour the fake factors are derived.

p_T^{probe} [GeV]	[15 – 20]	[20 – 25]	[25 – 40]	[40 – ∞]
Z e	0.27 ± 0.04	0.21 ± 0.04	0.0618 ± 0.021	0.15 ± 0.04
Z μ	0.089 ± 0.013	0.068 ± 0.013	0.036 ± 0.007	0.024 ± 0.007
W e	0	0.11 ± 0.03	0.031 ± 0.017	0.06 ± 0.04
W μ	0	0.022 ± 0.008	0.014 ± 0.005	0.015 ± 0.006

Table 6.6: Overview of the predicted number of events for an integrated luminosity of 36.1 fb^{-1} in the $t\bar{t}$ control region, when one of the potential probe leptons is allowed to fail the **Z-analysis** selection. The columns show the event yields when the probe lepton is required to be an electron (second column), muon (third column), or whether both flavours are accepted (last column). Additionally, the predicted relative contributions of top processes, as well as the total predicted and observed yields are shown. The shown uncertainties for the simulations include only the statistical uncertainties due to the limited number of events in the simulation.

	probe e	probe μ	all
$WZjj$	15.3 ± 0.9	13.5 ± 0.8	29 ± 1
other prompt	22 ± 1	24 ± 2	46 ± 2
top	1465 ± 22	4090 ± 40	5560 ± 40
Z +jets	33 ± 19	120 ± 100	150 ± 100
other non-prompt	30 ± 1	12.8 ± 0.8	42 ± 1
both non-prompt	16 ± 2	18 ± 3	34 ± 4
frac. top [%]	93 ± 2	96 ± 3	95 ± 2
total MC	1581 ± 29	4280 ± 110	5860 ± 110
Data	1470	4530	6010

Validation of control region choice: Event yields are summarized in Table 6.6 where only one of the potential probe leptons is required to pass the analysis selection. The relative contributions of top backgrounds are very high and on a similar level as compared to the signal region. When only considering the shown uncertainties, the differences are significant. However, this is in a region where one of the leptons is allowed to be very loosely selected.

The respective yields, when requiring both potential probe leptons to pass the **Z-analysis** selection, are shown in Table 6.7. When requiring both leptons to pass the analysis selection, the agreement is excellent even when only considering the statistical uncertainties of the simulation. This indicates that experimental uncertainties on control leptons could cause the observed differences.

Another check of the primary assumption of equivalent fake factors in the signal and control regions is a comparison of the truth origin of identified non-prompt leptons. The truth origins according to the **MCTruthClassifier** tool are shown for each lepton flavour in Figure 6.8. Objects which could not be classified are labeled as **NonDefined**, and results of successful classifications are combined in four groups; **Photons**, **LightJet**, **BJet**, and **CJet**. In the $t\bar{t}$ control region the dominant among the successful contributions is **CJet** for both electrons and muons. The c -type quark decays by emitting a W^\pm boson, which can decay leptonically and thus produce a real electron or muon in equal fractions.

Interestingly, the fraction of leptons that could not be classified is very similar in the control region for electrons and muons. A plausible reason for the **MCTruthClassifier** to fail would be actual misidentifications, e.g. cases, where a jet is falsely identified as an electron, rather than cases, where an actual electron was produced by hadron decays and adequately identified in the detector. In such cases, no truth object of the same type could be matched to the reconstructed object. However, one should assume that a misidentification of muons is much less likely compared to electrons. This behaviour indicates additional issues that need further study.

In the signal region, the statistical uncertainties on the fractions are considerable. The large uncertainties in combination with the large fraction on the unsuccessful classification of the **MCTruthClassifier**, which cannot be understood easily make it hard to draw reliable conclusions from a comparison between the control and signal region.

Table 6.7: Overview of the predicted number of events for an integrated luminosity of 36.1 fb^{-1} in the $t\bar{t}$ control region, when requiring both potential probe leptons to pass the **Z-analysis** selection. The columns show the event yields when the probe lepton is required to be an electron (second column), muon (third column), or whether both flavours are accepted (last column). Additionally, the predicted relative contributions of top processes, as well as the total predicted and observed yields are shown. The shown uncertainties for the simulations include only the statistical uncertainties due to the limited number of events in the simulation.

	probe e	probe μ	all
$WZjj$	12.4 ± 0.8	9.8 ± 0.7	22 ± 1
other prompt	17.2 ± 0.7	17.0 ± 0.4	34.2 ± 0.9
top	171 ± 7	177 ± 8	348 ± 11
$Z+\text{jets}$	13 ± 13	0 ± 0	13 ± 13
other non-prompt	15.3 ± 0.8	0.74 ± 0.11	16.0 ± 0.8
both non-prompt	3 ± 1	1.95 ± 0.19	5 ± 1
frac. top [%]	74 ± 7	86 ± 5	79 ± 4
total MC	232 ± 15	206 ± 8	438 ± 17
Data	226	203	429

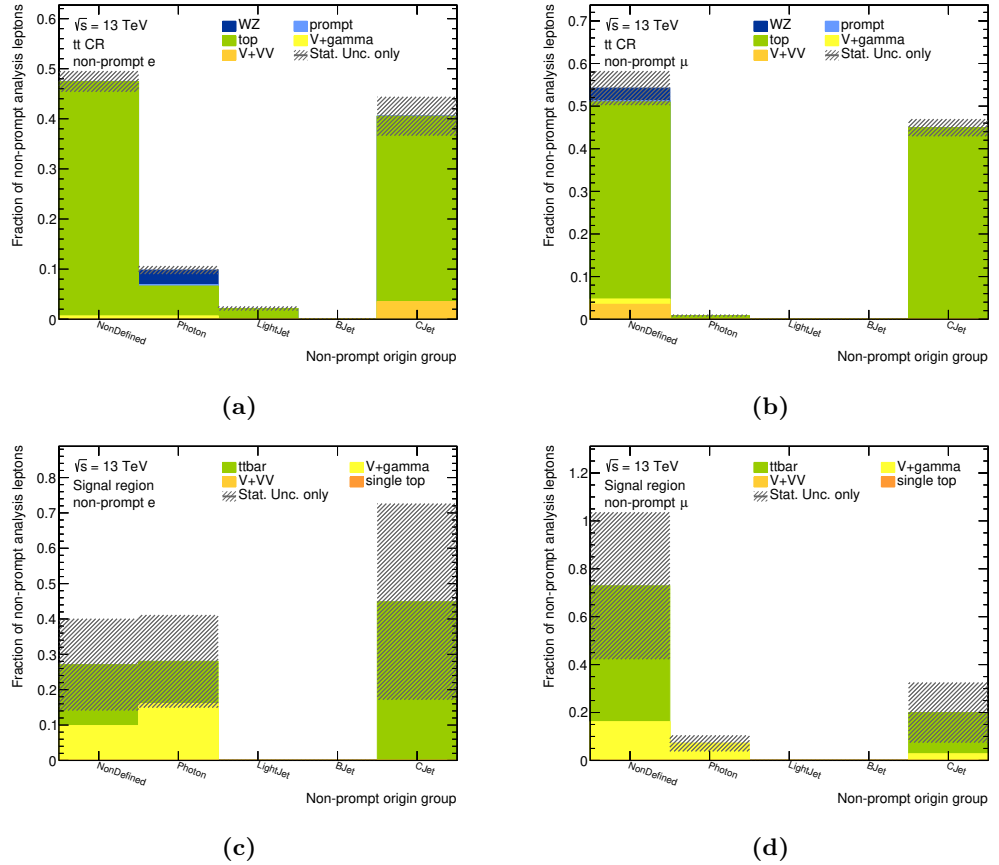


Figure 6.8: Fractions of truth origins in simulation-based non-prompt prediction for the $t\bar{t}$ CR and the signal region. The Figures in the top row show the $t\bar{t}$ CR and in the bottom row the signal region. Truth origin fractions for electrons (muons) are shown in Figures on the left (right) side. The different colours represent the different processes contributing to the non-prompt prediction with the colour assignment, as indicated in the legend.

6.2.3 Closure tests

The tests of underlying assumption proposed and performed in the previous section are reasonable indications on how to find a control region. However, as discussed there, the possible conclusions are limited due to limited statistics and residual effects not included in the test. There is a direct and comprehensive way to test the assumption in simulations. Such tests are possible in a simulation-based closure test of the procedure.

An ideal closure test is to apply the studied procedure on simulations in a way such that perfect agreement can be achieved. In closure tests that are set up like this, any deviation indicates a technical issue, problem in the implementation, or another issue that needs to be understood. Such a perfect closure test for the data-driven background estimation would be to extract a fake factor for a set of simulations in a control region and to apply this fake factor in the same set of simulations, in the same region while controlling several other selection criteria. In other words, the fake factor is applied to the same events where it was extracted.

In the full matrix method procedure as applied here, several complications are present like including the effect of events with two non-prompt leptons, the existence of two different selection levels, or the non-trivial assignment. It is possible to avoid each of these complications. This test was performed for the integrated event yield, and a perfect closure was achieved. However, this configuration in this test is deemed too far from the full matrix method procedure to draw meaningful conclusions from the concrete numbers, so they are not reported here. A more interesting comparison is to allow for such effects and to include them in the comparison. The final agreement can not be expected to be perfect anymore. Nevertheless, an agreement should be achieved within statistical uncertainties.

For this work, detailed closure tests were performed based on simulations of the top-quark production process. The $t\bar{t}$ control region is reasonably pure in this process, and for the sake of simplicity and efficiency, other processes were not included for most studies. Performing these closure tests motivated several changes compared to the previous study in [3], most notably adjustments in the overlap removal and p_T correction of control leptons (see Page 106 above). In order to justify these changes, the closure tests are described below.

Overlap between selected control leptons and jets: The starting point is to use the simulations of the $t\bar{t}$ process only and procedures more closely to [3], where the overlap between selected control leptons and jets is not removed, and the isolation-based p_T correction is not applied. Dedicated fake factors for this simulation are determined and applied following the matrix-method.

When comparing the event yields in the $t\bar{t}$ control region split by channels for the matrix-method prediction compared to the pure MC simulation, a good agreement is found. Differences are about 5 – 10 % and well within statistical uncertainties of the pure MC prediction or the fake factors, respectively of about 10 %. This consistency confirms that the technical setup works correctly. However, if the predicted event yields are compared in other regions of phase space, differences occur. The control region is adjusted step-by-step towards the signal region by inverting the veto of a lepton pair with the same flavour and opposite charge and adding additional requirements one at a time.

Following this procedure, a large disagreement occurs once the requirement of two jets is imposed. In order to understand the origin of this difference, the jet multiplicity is studied in detail. The exclusive jet multiplicity N_{jets} is compared for the pure MC and the matrix-method predictions in Figure 6.9a. Significant disagreement between the two predictions is found, where the matrix-method procedure predicts more jets compared to the pure MC prediction. Especially for events with low jet multiplicity, the distributions would agree much better, if the jet multiplicity for the matrix-method prediction would be shifted to lower values by one jet. In other words, the matrix-method procedure seems to predict one extra jet.

Since most events select exactly one control lepton, the lack of overlap removal between control leptons and jets is a reasonable assumption as origin of this disagreement. This assumption is cross-checked by evaluating the minimum distance between either of the two tagging jets and any of the three selected leptons. This check is performed in the $WZjj$

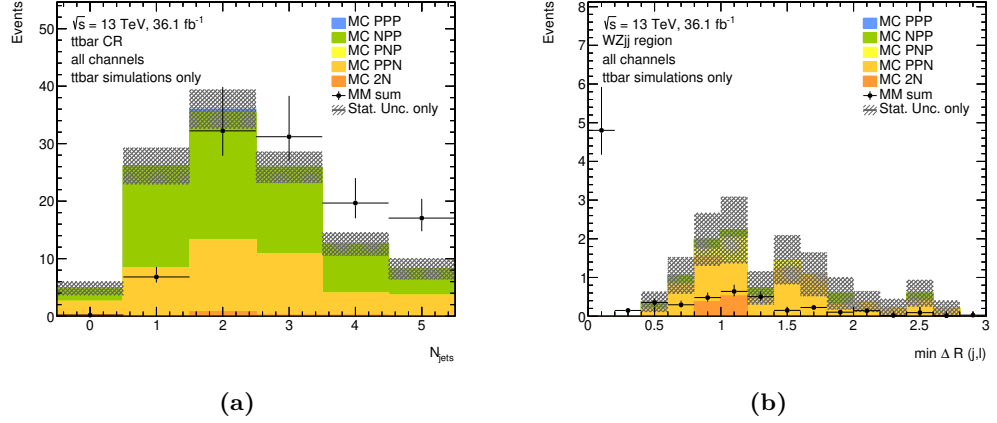


Figure 6.9: Comparison of predicted differential event yields for the $t\bar{t}$ process based on pure simulations to those of the plain MC-based matrix-method. The prediction based on simulations is divided according to the truth-level classification (P/N) of the three selected leptons. Figure (a) shows the event yield differentially in the $t\bar{t}$ control region and Figure (b) for the minimum angular distance between a tagging jet and any of the assigned leptons in the $WZjj$ region. Uncertainties include the statistical uncertainties due to the limited number of simulated events as well as systematic uncertainties on the fake factors.

region, where a requirement on the second jet is already imposed and the result is shown in Figure 6.9b. It can be seen, that for a majority of the events in the matrix-method prediction a tagging jet overlaps with a signal lepton within $\Delta R(j, \ell) < 0.4$. This indicates a double-counting by identifying the same underlying physical object as two different objects.

One procedure to avoid this problematic behaviour would be to apply the same overlap removal procedure as for the analysis lepton also to the control leptons. In this way, such overlap could not occur, but the number of control leptons would be significantly reduced, thus increasing the statistical uncertainties on the final data-driven prediction.

The solution applied in this work is to require a minimum angular distance of $\Delta R \geq 0.4$ for any jet to any of the three assigned leptons. Any jet that is closer to an assigned lepton is removed from the event. For pure MC simulations, nothing changes, since this overlap between analysis leptons and jets is already removed as part of the nominal selection. For the matrix-method prediction, however, one control lepton is selected in most events, and any jet that overlaps with this control lepton is removed, thus reducing the predicted jet multiplicity by 1.

The closure test is performed again with the additional requirement in the jet selection. This additional requirement does not change the fake factors, but the adjusted jet object selection changes the jet multiplicity distributions and the effect of event selection criteria. The respective distributions are shown in Figure 6.10.

As expected, the jet multiplicity distribution is in much better agreement, and no significant difference is found. The minimum distance between a tagging jet and an assigned lepton $\min \Delta R(j, \ell)$ is in much better agreement as well. Not only the peak around 0 for the matrix-method prediction was removed, but also the normalization agrees better than before.

The additional requirements to remove the overlap between jets and selected control leptons is well justified and necessary in order to avoid a bias due to the double-counting of objects. From here on, the jet selection criterion is applied for all samples.

Control lepton p_T -correction: The other main effect to be studied is the isolation-based p_T -correction applied to control leptons. Again, a closure test is performed for simulations of the $t\bar{t}$ process without applying this correction to study the effect.

The event yield comparison is unaffected by the addition of the overlap removal require-

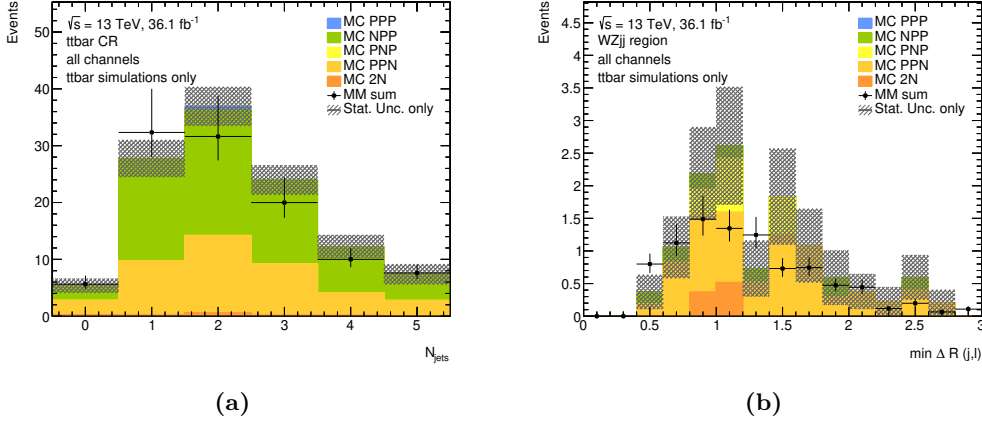


Figure 6.10: Comparison of predicted differential event yields for the $t\bar{t}$ process based on pure simulations to those of the MC-based matrix-method with the additional jet requirement mentioned in the text. The prediction based on simulations is divided according to the truth-level classification (P/N) of the three selected leptons. Figure (a) shows the event yield differentially in the $t\bar{t}$ control region and Figure (b) for the minimum angular distance between a tagging jet and any of the assigned leptons in the $WZjj$ region. Uncertainties include the statistical uncertainties due to the limited number of simulated events as well as systematic uncertainties on the fake factors.

ment between jets and control leptons, and the predicted yields are in agreement within their statistical uncertainties. By adjusting the region in which the predicted yields are compared towards the signal region, the effect of different requirements is tested. The first significant disagreement occurs, when requiring the invariant mass of the leptons assigned to the Z boson to be within a window of 10 GeV around the nominal mass of the Z boson. This behaviour indicates problems in the modelling of the invariant mass of this lepton pair.

The predicted distributions for the invariant mass of the lepton pair assigned to the Z boson $M(\ell_{Z,1}\ell_{Z,2})$ is shown in Figure 6.11 for the $t\bar{t}$ CR as well as the $\ell\ell\nu jj$ region. The predicted distributions disagree in both considered regions. In both regions, the matrix-method predicts lower invariant masses resulting in a shift of the peak to lower invariant masses. The regions are orthogonal due to the different requirements on the presence of an SFOC lepton pair in the event. While from the two identical leptons, the lepton with lower- p_T is used in combination with the single lepton of opposite charge in the $t\bar{t}$ CR, the $\ell\ell\nu jj$ region uses the default algorithm at detector-level, favouring the $M(\ell_{Z,1}\ell_{Z,2})$ to be closest to the nominal Z boson mass. Since the issue occurs in both regions, it can not be caused by the assignment.

A mismodelling of the momentum of any of the two leptons used for $M(\ell_{Z,1}\ell_{Z,2})$ can cause deviations of the invariant mass. Such mismodelling is partially corrected by using independent fake factors depending on the transverse momentum of the control lepton. The effect of this is visualized in Figure 6.12, where normalized predictions for the transverse momentum of the subleading Z lepton $p_T(\ell_{Z,2})$ and the lepton pair invariant mass $M(\ell_{Z,1}\ell_{Z,2})$ are compared before and after applying the fake factor. Before the application of the fake factor, both distributions show significant disagreement. When the fake factor is applied, the agreement in $p_T(\ell_{Z,2})$ is much improved, while the invariant mass distributions are mostly unchanged and still differ significantly. The remaining differences indicate that the underlying origin for the mismodelling is not fixed when only relying on the binning of the fake factor for the correction. Additional corrections for the kinematics are necessary.

Further tests are performed in the $\ell\ell\nu jj$ region due to the increased available statistic. According to the classification of the `MCTruthClassifier`, the subleading lepton assigned to the Z boson $\ell_{Z,2}$ is non-prompt for the majority of the $t\bar{t}$ contribution in this region. In order to understand the physics origin of the discrepancy, the effect is studied in different channels.

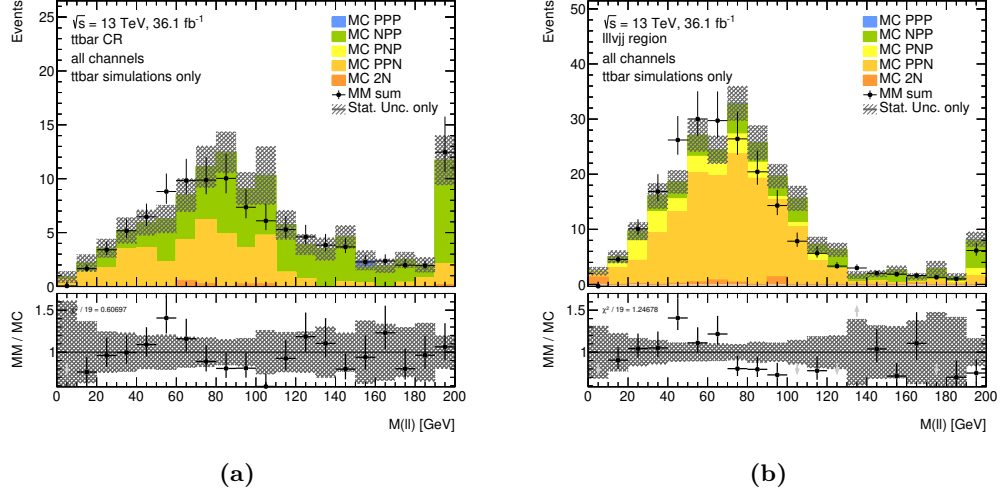


Figure 6.11: Comparison of predicted differential event yields in the invariant mass of the leptons assigned to the Z boson for the $t\bar{t}$ process based on pure simulations to those of the MC-based matrix-method. The prediction based on simulations is divided according to the truth-level classification (P/N) of the three selected leptons. Figure (a) shows the differential event yield in the $t\bar{t}$ control region and Figure (b) in the $WZjj$ region. Uncertainties include the statistical uncertainties due to the limited number of simulated events as well as systematic uncertainties on the fake factors.

The distribution of the invariant mass of the lepton pair $M(\ell_{Z,1}\ell_{Z,2})$ is shown for each channel in Figure 6.13. In the eee and μee channels, where the invariant mass is of an electron pair is evaluated, the difference is less significant, and no significant shift between the pure MC simulation and the matrix-method prediction is found. However, in the $e\mu\mu$ and $\mu\mu\mu$ channels, the distributions clearly differ. The origin of the discrepancy must be more pronounced for control muons compared to control electrons.

The mismodelling must be caused by a difference in the selection of **Z-control** muons compared to **Z-analysis** muons. The additional criteria which are only applied for **Z-analysis** muons and not for **Z-control** muons are

- tracking impact parameter d_0 significance below 5: $|d_0/\sigma_{d_0}| < 3$
- minimum isolation working point **FixedCutLoose**, which requires for muons[152]
 - $E_T^{\text{cone20}}/p_T < 0.3$
 - $p_T^{\text{varcone30}}/p_T < 0.15$,

where E_T^{cone20} and $p_T^{\text{varcone30}}$ are defined as introduced in Section 4.3.4.

By requiring the control muon also to pass one of these criteria at a time, the majority of the discrepancy could be associated with muons failing the $p_T^{\text{varcone30}}/p_T < 0.15$ requirement. Such discrepancy indicates a miscalibration of muons failing this isolation requirement.

This effect was also observed in [2], where the correction

$$p_T \rightarrow p_T + p_T^{\text{varcone30}} \quad (6.16)$$

was applied as described in detail in Section 6.2.2 above. Adjustments of the corrections where the contribution of $p_T^{\text{varcone30}}$ was scaled with a constant were tested but found to perform worse.

When re-running the closure test while applying the isolation-based p_T -correction, the predictions agree much better. Figure 6.14 compares the normalized predicted differential distributions for $p_T(\ell_{Z,2})$ and $M(\ell_{Z,1}\ell_{Z,2})$ before the application of the fake factors. Already before the application of the fake factor the predictions for the $M(\ell_{Z,1}\ell_{Z,2})$ distribution are in much better agreement to the case without the correction shown in Figure 6.12.

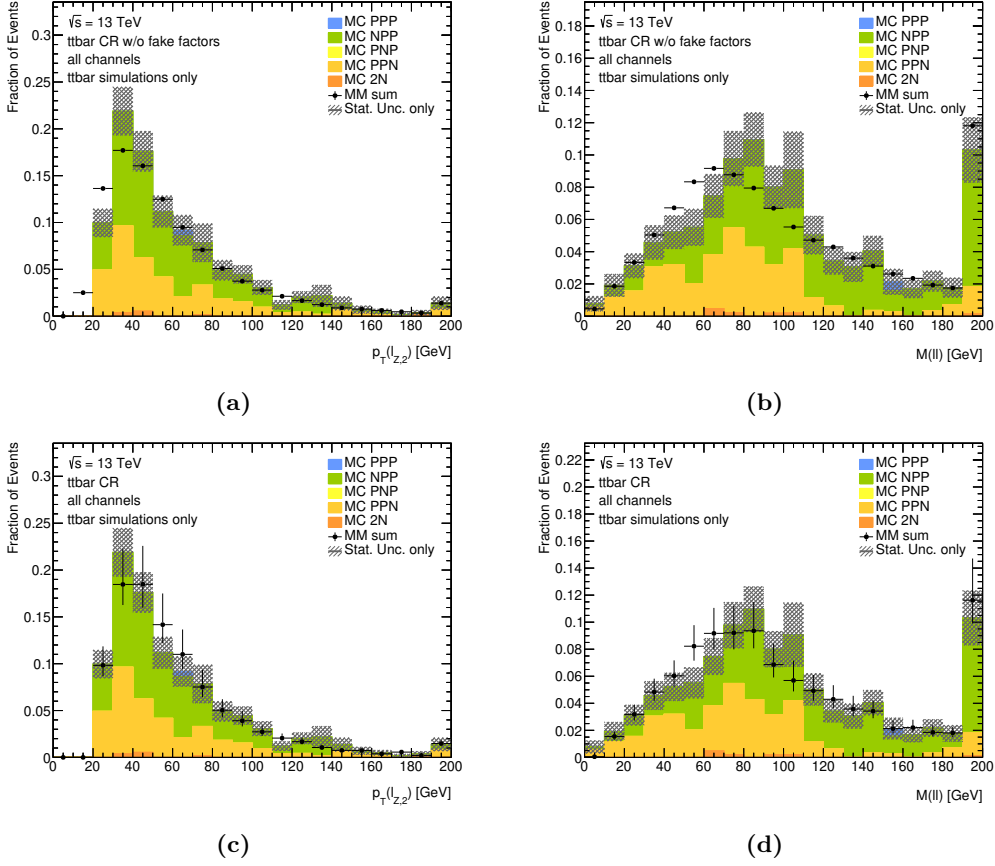


Figure 6.12: Comparison of predicted distributions in the invariant mass of the leptons assigned to the Z boson for the $t\bar{t}$ process in the $t\bar{t}$ control region based on pure simulations to those of the MC-based matrix-method as visualization of the kinematic effect of the fake factor. The prediction based on simulations is divided according to the truth-level classification (P/N) of the three selected leptons. Figures in the top row show predictions of the distributions before the application of the fake factors and Figures in the bottom row after the application. Figures on the left show the distribution in the transverse momentum of the subleading lepton assigned to the Z boson and Figures on the right for the invariant mass of the lepton pair assigned to the Z boson. Uncertainties include the statistical uncertainties due to the limited number of simulated events as well as systematic uncertainties on the fake factors.

The predicted distributions for the invariant mass of the lepton pair assigned to the Z boson $M(\ell_{Z,1}\ell_{Z,2})$ after the correction are shown in Figure 6.15 for the $t\bar{t}$ CR as well as the $\ell\ell\nu jj$ region. When comparing the corrected predictions to those shown in Figure 6.11 without the correction, it can be seen, that the correction fixes the bias towards smaller invariant masses in the matrix-method.

In general, there seems to be an issue with the calibration of non-isolated leptons. The calibration is typically derived for isolated leptons and directly applying such calibration to non-isolated leptons, as is necessary when considering control leptons, introduces a bias. It is considered necessary to check for similar mismodellings in future studies which employ data-driven estimates based on the fake factor method.

η correction: It has been observed in earlier studies, that the fake factor not only depends on the transverse momentum p_T but also on the pseudo-rapidity η of the non-prompt lepton candidate. For this reason, the fake factor is often derived in ranges of p_T and η simultaneously. Due to the limited available statistics, such 2-dimensional fake factors are not feasible

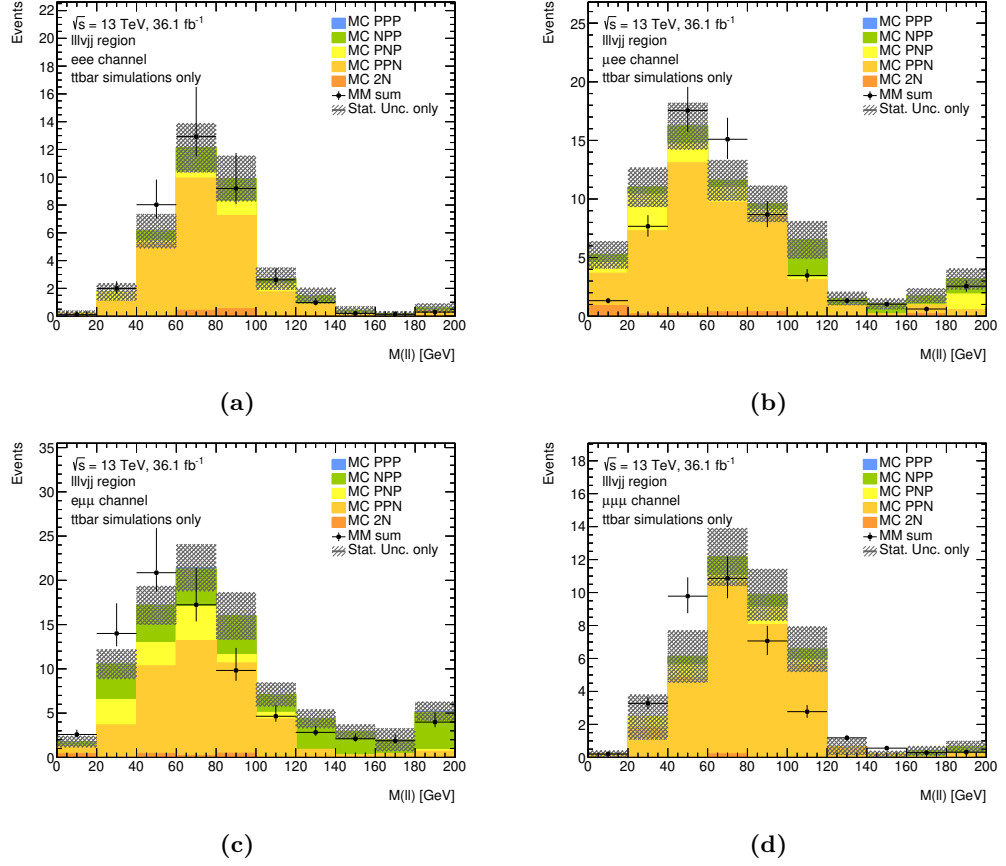


Figure 6.13: Comparison of predicted differential event yields in the invariant mass of the leptons assigned to the Z boson for the $t\bar{t}$ process in the $t\bar{t}$ control region based on pure simulations to those of the MC-based matrix-method for the different leptonic channels. The prediction based on simulations is divided according to the truth-level classification (P/N) of the three selected leptons. For each Figure, the considered channel is indicated in the top left corner. Uncertainties include the statistical uncertainties due to the limited number of simulated events as well as systematic uncertainties on the fake factors.

when using the $t\bar{t}$ control region to derive the fake factors. Instead, an η -based correction is derived, adjusting the p_T -dependent fake factor. This approach was applied in [3] and is followed here.

The discrepancy in the differential predictions in η in the $\ell\ell\ell\nu jj$ region are shown in Figure 6.16. The matrix-method predicts fewer events for forward leptons and more for central leptons.

In order to correct this bias, a $|\eta|$ -dependent correction is derived from pure MC simulations only. Due to the limited amount of events in data, especially in the more forward ($|\eta| > 1.5$) region, the statistical uncertainties were too large to derive a meaningful correction from data. With a larger dataset a data-driven extraction of this η correction should be targeted.

The η -correction factors C_η are determined similar to the fake factor, but in ranges of $|\eta|$ (ℓ^{probe})

$$C_\eta^i = \frac{f_{N \rightarrow A}^i}{f_{N \rightarrow C}^i}. \quad (6.17)$$

where $f_{N \rightarrow A}^i$ denotes the fraction of all events with an analysis lepton probe ($N_{N \rightarrow A}$), where $|\eta|$ (ℓ^{probe}) is in the currently considered range i and $f_{N \rightarrow C}^i$ is defined analogously for control leptons.

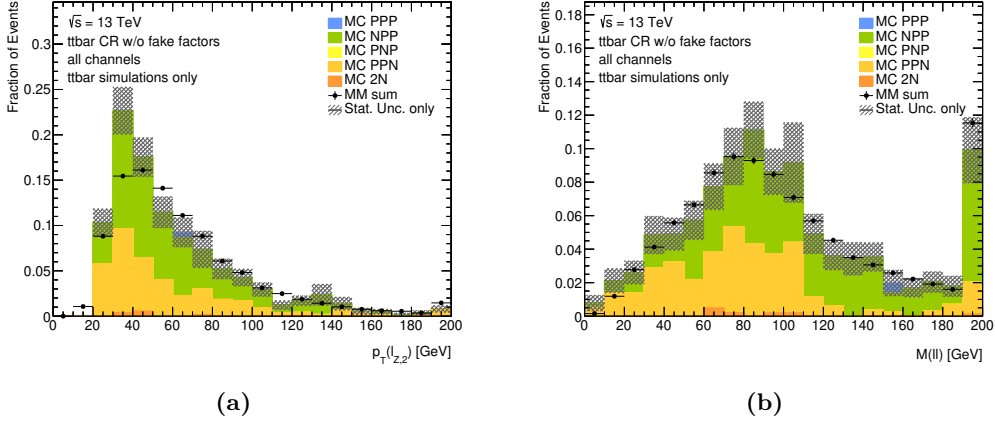


Figure 6.14: Comparison of predicted distributions for the $t\bar{t}$ process in the $t\bar{t}$ control region based on pure simulations to those of the MC-based matrix-method before the application of the fake factors with the isolation-based p_T correction. The prediction based on simulations is divided according to the truth-level classification (P/N) of the three selected leptons. Figure (a) shows the event yield differentially in the transverse momentum of the subleading lepton assigned to the Z boson and Figure (b) for the invariant mass of the lepton assigned to the Z boson. Uncertainties include the statistical uncertainties due to the limited number of simulated events as well as systematic uncertainties on the fake factors.

The main difference, is that the difference in the total event yield is already covered for in the fake factor, so C_η is only supposed to correct the η -dependency. In practice, the calculation is performed by scaling the $|\eta|(\ell^{\text{probe}})$ histograms to unit-area and by evaluating the ratio in each bin. The correction factors C_η are determined in the same region as the fake factors, before applying the fake factors for five equi-distant ranges from $|\eta|(\ell^{\text{probe}}) = 0$ to $|\eta|(\ell^{\text{probe}}) = 2.5$.

In order to avoid unphysical steps in the transition between $|\eta|(\ell^{\text{probe}})$ ranges the determined correction factors are fitted with a quadratic equation symmetric with respect $\eta = 0$

$$C_\eta = a + b \cdot \eta^2, \quad (6.18)$$

where a and b are free parameters. The best-fit values for a and b are determined for four different scenarios reflecting the two lepton flavours and two analysis selections using a χ^2 fit. The fit inputs as well as results are shown for all four scenarios in Figure 6.17.

The overall fit qualities are very good, and in each fit, a correlation between the two parameters of about -0.5 to -0.6 is obtained. The uncertainty in the correction factor is approximated using error propagation with correlated parameters

$$\sigma_{C_\eta} \approx \sqrt{\sigma_a^2 + \sigma_b^2 \cdot \eta^4 + 2 \cdot c_{ab} \sigma_a \sigma_b \eta^2}. \quad (6.19)$$

where σ_a (σ_b) is the standard deviation of parameter a (b) and c_{ab} is the correlation between the two parameters. The resulting systematic uncertainty is increased by a factor of 2 in order to cover potential differences in the η correction between simulated and measured data.

The corrected predictions for the pseudo-rapidity distributions are shown in Figure 6.18. The systematic uncertainties introduced by the fit of the $|\eta|$ -correction are not propagated here for technical reasons. The corrected predictions agree much better, especially differentially in $\eta(\ell_{Z,2})$.

The corrected prediction for $\eta(\ell_W)$ disagrees for large negative pseudo-rapidities. The differential prediction in $\eta(\ell_W)$ was asymmetric in η , and the matrix-method predicted fewer events for large negative pseudo-rapidities compared to the direct prediction of MC simulations. Neither of the individual distributions has a significant asymmetry, i.e. the differences

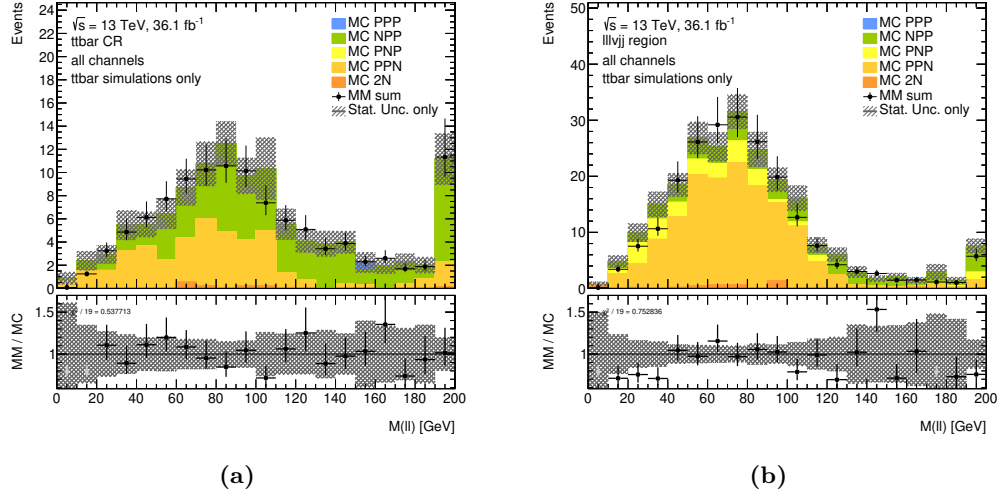


Figure 6.15: Comparison of predicted distributions for the $t\bar{t}$ process in the $t\bar{t}$ control region based on pure simulations to those of the MC-based matrix-method with the isolation-based p_T correction. The prediction based on simulations is divided according to the truth-level classification (P/N) of the three selected leptons. Figure (a) shows the event yield differentially in the transverse momentum of the subleading lepton assigned to the Z boson and Figure (b) for the invariant mass of the lepton assigned to the Z boson. Uncertainties include the statistical uncertainties due to the limited number of simulated events as well as systematic uncertainties on the fake factors.

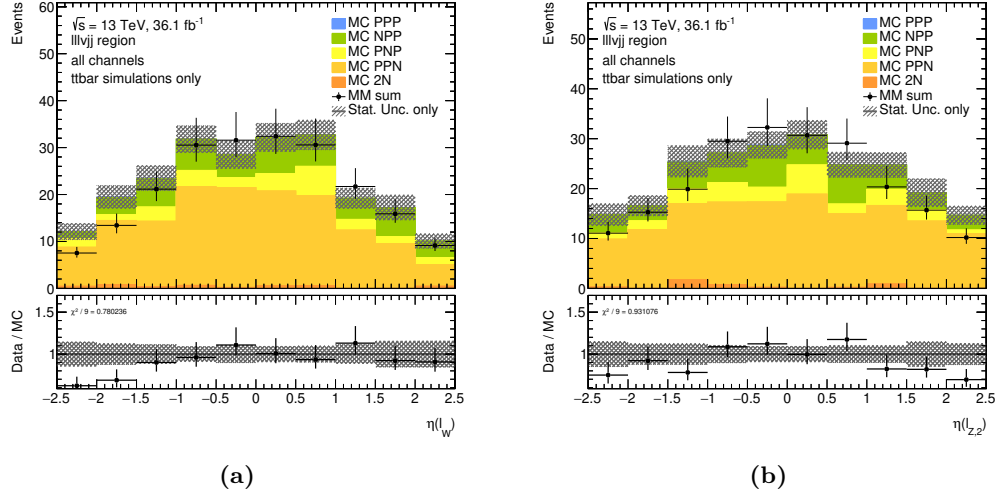


Figure 6.16: Comparison of predicted distributions for the $t\bar{t}$ process in the $lll\nu jj$ region based on pure simulations to those of the MC-based matrix-method with the isolation-based p_T correction. The prediction based on simulations is divided according to the truth-level classification (P/N) of the three selected leptons. Figure (a) shows the event yield differentially in the pseudo-rapidity of the lepton assigned to the W^\pm boson and Figure (b) for the pseudo-rapidity of the subleading lepton assigned to the Z boson. Uncertainties include the statistical uncertainties due to the limited number of simulated events as well as systematic uncertainties on the fake factors.

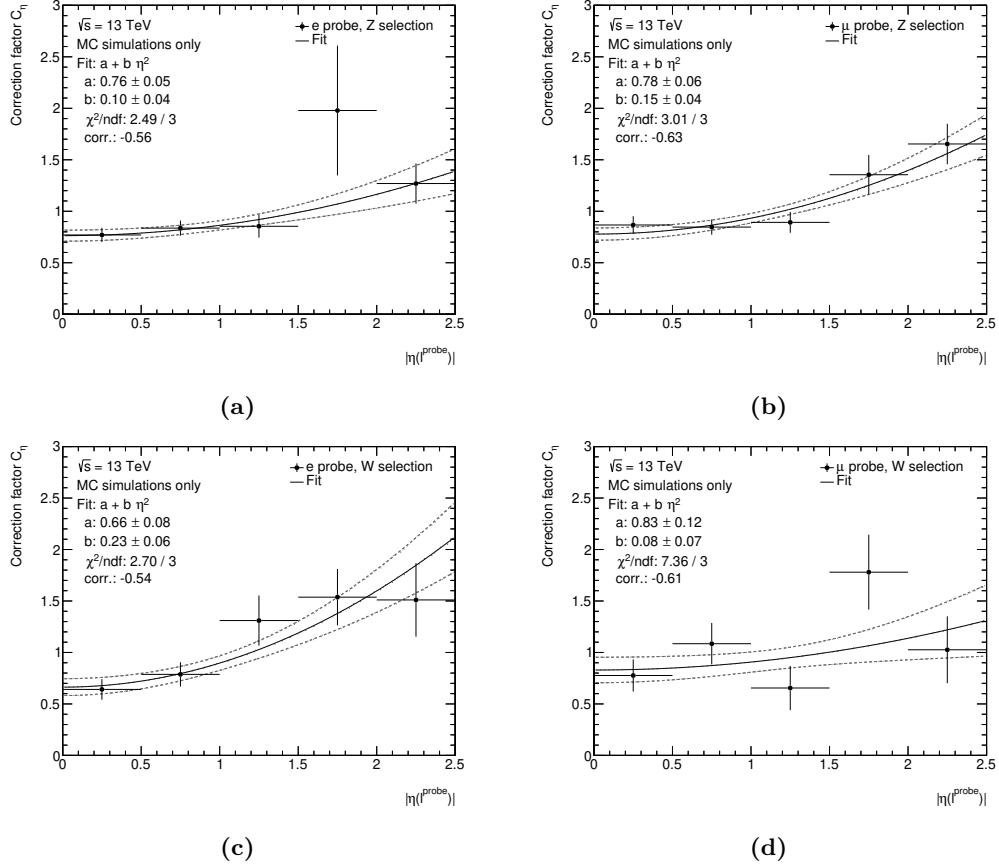


Figure 6.17: Correction factors for pseudo-rapidity distribution of the non-prompt lepton candidate. The factor is determined for different ranges in the absolute pseudo-rapidity and is shown using black data point. Additionally, a fitted continuous function used for the actual correction is shown (solid line) including an uncertainty band (dashed lines). The fitting function and the obtained parameters are shown in the top left corner of the plot. Figures in the top row show the correction factors for the **Z-analysis** selection and in the bottom row for the **W-analysis** selection. Figures on the left (right) side show the correction factors for electrons (muons).

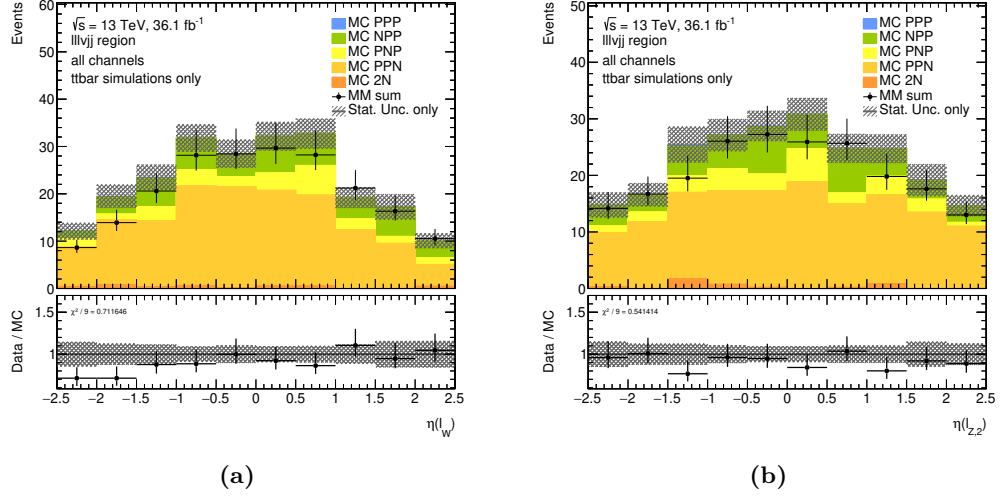


Figure 6.18: Comparison of predicted distributions for the $t\bar{t}$ process in the $\ell\ell\nu jj$ region based on pure simulations to those of the MC-based matrix-method with the isolation-based p_T -correction and the η correction. The prediction based on simulations is divided according to the truth-level classification (P/N) of the three selected leptons. Figure (a) shows the event yield differentially in the pseudo-rapidity of the lepton assigned to the W^\pm boson and Figure (b) for the pseudo-rapidity of the subleading lepton assigned to the Z boson. Uncertainties include the statistical uncertainties due to the limited number of simulated events as well as systematic uncertainties on the fake factors.

are well within statistical uncertainties, but the predictions are slightly asymmetric in opposing directions. Since the correction is defined symmetrically, the asymmetry is still present in the corrected predictions.

Due to the excellent agreement in the $\eta(\ell_{Z,2})$ distribution and the limited available statistics, the asymmetry is not investigated further.

Full MC closure test: The closure tests mentioned above were performed only using simulations for the $t\bar{t}$ production process for performance reason. A final closure test with all of the derived corrections is applied, including simulations for all considered processes. New simulation-based fake factors are derived and applied to the full set of MC simulations. The resulting prediction is compared to the direct prediction of the MC simulations.

The predicted event yields are shown per process in Table 6.8 for different regions of phase space. The numbers are also listed per lepton channel in Appendix E.2. The prediction for the $t\bar{t}$ control region agrees very well at a level of 3 % before applying the η -correction. These differences are already covered by the statistical uncertainties of the pure MC prediction at about 5 %. Once the η correction is applied, the differences increase to about 7 %. The systematic uncertainties on the fake factors due to the limited statistics in the control region cover the differences. In the $WZjj$ region, the agreement is greatly improved to about 3 %, which is small compared to the statistical uncertainties.

When comparing the predictions per process, it is evident that other processes are systematically smaller in the matrix-method prediction compared to the direct MC prediction. The effect is most significant for $V + \gamma$, where the MC statistic is sufficient to draw reliable conclusions. The contribution of this process is 50 % smaller in the matrix-method prediction.

This behaviour is also confirmed in comparisons of differential predictions. Figure 6.19 compares the distribution split by lepton flavour channel and the differential prediction in the missing transverse energy E_T^{miss} in the $\ell\ell\nu jj$ region and the $WZjj$ region. The direct prediction from MC simulations is split by process. It can be seen that only in regions of phase space where the $V + \gamma$ contribution is non-negligible, the predictions differ significantly. The $V + \gamma$ contributes mostly in channels where an electron is assigned to the W^\pm boson, i.e. the $e\mu\mu$ and eee channels, and for $E_T^{\text{miss}} < 40$ GeV. In other regions of phase space, the

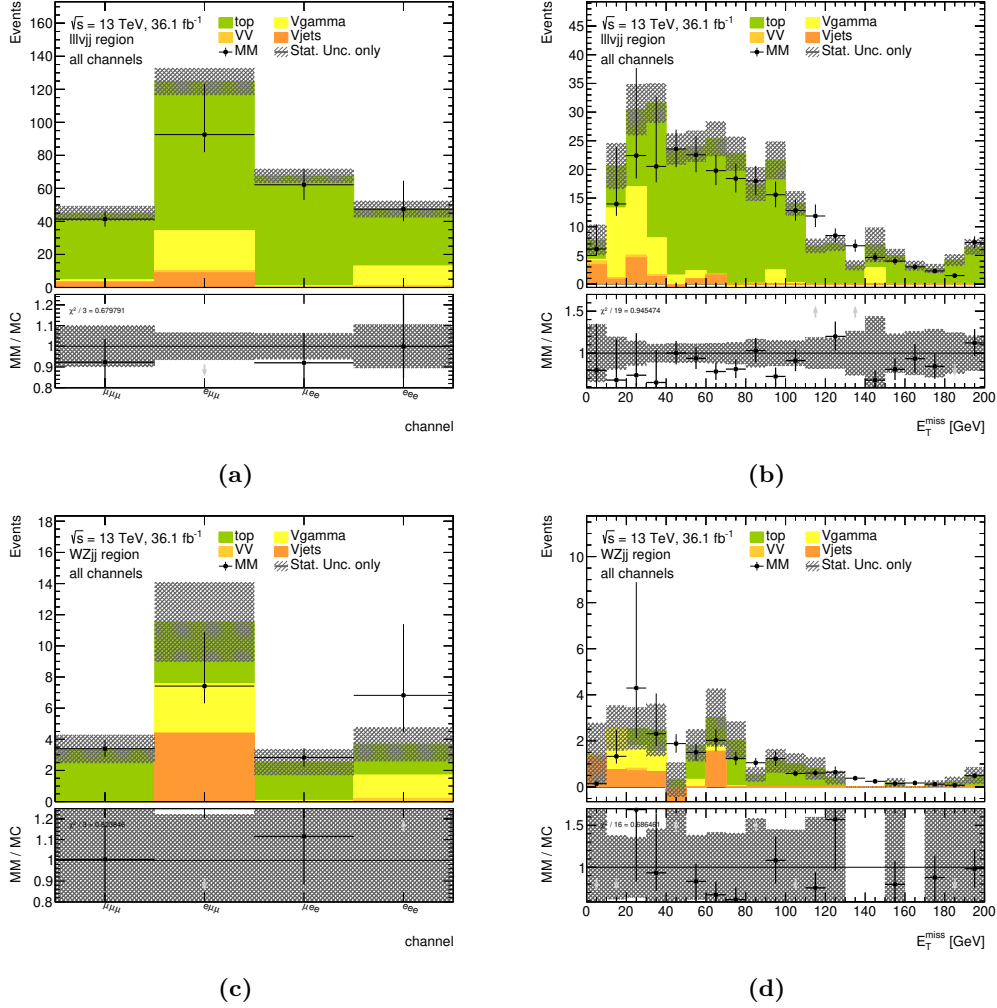


Figure 6.19: Comparison of predicted distributions for the $t\bar{t}$ process based on pure simulations (coloured bars) to those of the MC-based matrix-method including all derived corrections (data points). The prediction based on simulations is split by process. Figures in the top row show event yields in $\ell\ell\nu jj$ region and in the bottom row the $WZjj$ region. Figures on the left show the distribution over the lepton channels and on the right the missing transverse energy. Uncertainties include the statistical uncertainties due to the limited number of simulated events as well as systematic uncertainties on the fake factors.

Table 6.8: Overview of predicted number of events for an integrated luminosity of 36.1 fb^{-1} per process in each of the three regions for the simulation-based prediction (MC) as well as the matrix-method prediction (MM). Additional rows on the bottom show the total event yield for each of the prediction methods and the ratio of the matrix-method to the simulation-based prediction. Uncertainties are split in the statistical uncertainties due to the limited number of simulated events and systematic uncertainties of the matrix-method.

	ttbar CR			$\ell\ell\nu jj$ region			$WZjj$ region		
MC top	118	± 6	$^{+0}_{-0}$	232	± 8	$^{+0}_{-0}$	12	± 2	$^{+0}_{-0}$
MC VV	0.53 ± 0.18			$^{+0}_{-0}$	2.8 ± 0.7			$^{+0}_{-0}$	0.29 ± 0.12
MC $V\gamma$	0.4	± 0.4	$^{+0}_{-0}$	38	± 7	$^{+0}_{-0}$	5	± 1	$^{+0}_{-0}$
MC V +jets	0	± 0	$^{+0}_{-0}$	12	± 4	$^{+0}_{-0}$	4	± 2	$^{+0}_{-0}$
MM top	107	± 2	$^{+24}_{-14}$	210	± 0	$^{+40}_{-30}$	12	± 1	$^{+2}_{-2}$
MM VV	0.41 ± 0.05			$^{+0.15}_{-0.05}$	1.2 ± 0.1			$^{+0.6}_{-0.2}$	0.25 ± 0.04
MM $V\gamma$	0.02 ± 0.04			$^{+0.00}_{-0.00}$	15	± 1	$^{+10}_{-2}$	2	± 0
MM V +jets	4	± 2	$^{+2}_{-0}$	13	± 3	$^{+9}_{-2}$	5	± 2	$^{+4}_{-1}$
total MC	119	± 6	$^{+0}_{-0}$	285	± 11	$^{+0}_{-0}$	21	± 3	$^{+0}_{-0}$
total MM	111	± 3	$^{+25}_{-15}$	240	± 0	$^{+60}_{-30}$	20	± 2	$^{+8}_{-3}$
ratio MM / MC	0.93 ± 0.06			$^{+0.21}_{-0.21}$	0.86 ± 0.04			$^{+0.20}_{-0.20}$	1.0 ± 0.2
								$^{+0.4}_{-0.4}$	

predictions are in good agreement.

In order to conservatively cover remaining differences due to other contributing processes and the migration towards the signal region, an additional systematic uncertainty on the matrix-method prediction is introduced. In regions with significant $V + \gamma$ contributions, i.e. in the $e\mu\mu$ and eee channels for $E_T^{\text{miss}} < 40 \text{ GeV}$, this uncertainty is set to 100 %. For all other regions of phase space 10 % are added.

6.2.4 Results of matrix method

The final data-driven estimation for non-prompt backgrounds can be derived using fake factors, estimated from data in the $t\bar{t}$ control region. These fake factors are applied to events in control lepton regions for events in measured data and subtracting contributions where all three leptons are prompt, estimated from MC simulations. The corrections mentioned above and systematic uncertainties are applied.

The predicted event yields for the three regions of interest are listed in Table 6.9. More detailed listings split by lepton flavour channel are presented in Appendix E.3. The overall agreement of the predicted total event yields is very good.

Additional comparisons of differential yields in the $WZjj$ region are shown in Figure 6.20. In general, the predictions are in good agreement. The matrix-method prediction is more smooth, reflecting smaller statistical uncertainties. The total uncertainties of the matrix-method prediction include all systematics directly related to the matrix-method. These total uncertainties are on a similar order as the statistical uncertainties of the prediction from pure MC simulations. A large benefit for the matrix-method is that these uncertainties are mainly driven by systematics and are thus correlated between different bins. While for the pure MC prediction, the shown uncertainties are completely uncorrelated. Proper treatment of the MC predictions would require additional theoretical uncertainties on the PDF, the scale choice and similar to ensure conservative coverage of the predictions. These are usually additional systematic uncertainties between 10 % and 50 % for background processes, which makes the pure MC prediction much more unfavourable.

In summary, the data-driven background estimate using the matrix-method provides a good prediction for non-prompt backgrounds. A set of corrections has been derived motivated

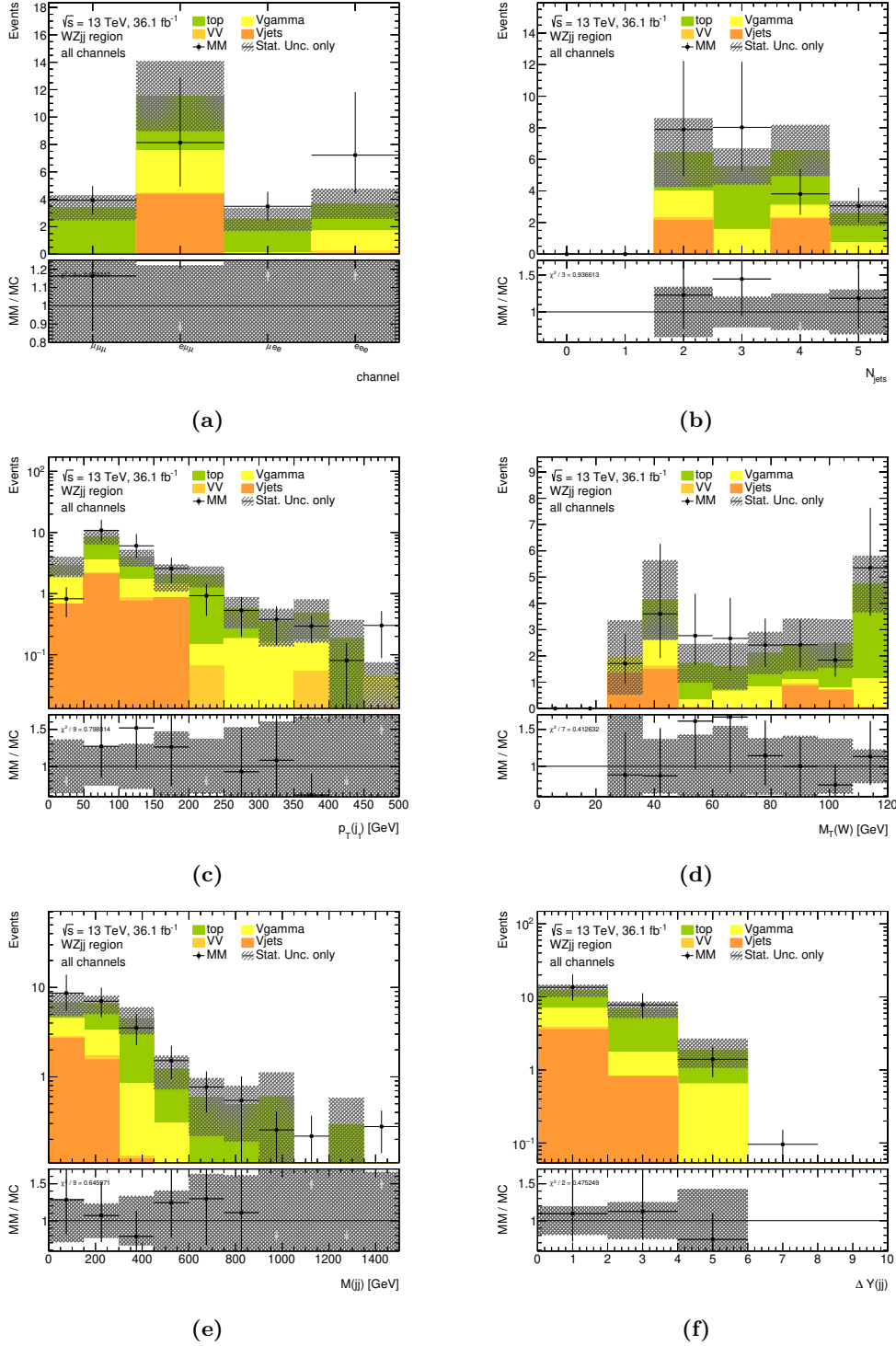


Figure 6.20: Comparison of predicted distributions for the $t\bar{t}$ process in the $WZjj$ region based on pure simulations (coloured bars) to those of the data-driven matrix-method including all derived corrections (data points). The prediction based on simulations is split by process. Figure (a) shows the predicted yield for each lepton channel, Figure (b) in the jet multiplicity, Figure (c) the transverse momentum of the leading jet, Figure (d) in the transverse mass of the W^\pm boson, Figure (e) the tagging jet invariant mass, and Figure (f) the rapidity separation of the tagging jets. Uncertainties include the statistical uncertainties due to the limited number of simulated events as well as systematic uncertainties on the fake factors.

Table 6.9: Overview of the predicted number of events for an integrated luminosity of 36.1 fb^{-1} per process in each of the three regions for the simulation-based prediction (MC) as well as the data-driven matrix-method prediction (MM). The simulation-based prediction is split according to the process. In contrast, the data-driven matrix-method prediction is split according to the detector-level classification (A/C), i.e. which selection stage each lepton candidate passed. Additional rows on the bottom show the total event yield for each of the prediction methods and the ratio of the matrix-method to the simulation-based prediction. Uncertainties are split into the statistical uncertainties due to the limited number of simulated events and systematic uncertainties of the matrix-method.

	$\ell\ell\nu jj$ region			$WZjj$ region			Signal region		
MC top	232	± 8	$^{+0}_{-0}$	12	± 2	$^{+0}_{-0}$	1.7	± 0.8	$^{+0}_{-0}$
MC $V + \gamma$	38	± 7	$^{+0}_{-0}$	5	± 1	$^{+0}_{-0}$	0.30	± 0.11	$^{+0}_{-0}$
MC VV	2.8	± 0.7	$^{+0}_{-0}$	0.29 ± 0.12	$^{+0}_{-0}$		0	± 0	$^{+0}_{-0}$
MC $V + \text{jets}$	12	± 4	$^{+0}_{-0}$	4	± 2	$^{+0}_{-0}$	0	± 0	$^{+0}_{-0}$
MM CAA	90	± 0	$^{+50}_{-40}$	10	± 2	$^{+7}_{-4}$	1.0	± 0.5	$^{+0.5}_{-0.5}$
MM ACA	44	± 1	$^{+13}_{-12}$	2.9 ± 0.3	$^{+0.9}_{-0.8}$		0.34	± 0.08	$^{+0.10}_{-0.09}$
MM AAC	139	± 3	$^{+34}_{-28}$	10	± 1	$^{+3}_{-2}$	1.22	± 0.19	$^{+0.34}_{-0.27}$
MM 2C	2	± 0	$^{+1}_{-1}$	-0.19 ± 0.05	$^{+0.10}_{-0.14}$		-0.040 ± 0.027	$^{+0.019}_{-0.035}$	
total MC	285	± 11	$^{+0}_{-0}$	21	± 3	$^{+0}_{-0}$	2.0	± 0.8	$^{+0}_{-0}$
total MM	270	± 10	$^{+90}_{-70}$	23	± 2	$^{+10}_{-7}$	2.5	± 0.6	$^{+0.9}_{-0.8}$
ratio MM / MC	0.94 ± 0.04	$^{+0.32}_{-0.32}$		1.1 ± 0.2	$^{+0.5}_{-0.5}$		1.2 ± 0.5	$^{+0.4}_{-0.4}$	

from extensive MC closure tests. A set of uncertainties has been derived in order to cover known mismodellings conservatively. The final prediction is compatible with direct MC predictions but is more accurate.

Chapter 7

Multi-variate Event Classification

As seen in the previous chapters and especially in Table 5.5, the contribution of backgrounds in the signal region is still very large. Especially the $WZjj$ -EW4 background is predicted to be much larger than the predicted $WZjj$ -EW6 process. The expected sensitivity to the signal process is rather small, due to the substantial background contributions and the associated uncertainties.

The experimental sensitivity can be increased by studying differential distributions in an observable that is sensitive to the signal process. Observables are sensitive if the predicted differential distributions differ significantly between the signal and background processes.

This effect can be enhanced by combining the information of multiple observables. An excellent way to combine information of multiple observables to increase the sensitivity to the signal process is to use machine learning. This work follows the approach of [3] and uses a *Boosted Decision Tree (BDT)* in order to combine the information of multiple observables into one output, typically referred to as **BDTscore**.

In [3], a BDT was trained to distinguish between the $WZjj$ -EW6 signal process and prompt background processes. Since the phase space in this work is defined differently, the BDT of $WZjj$ -EW6 cannot be directly applied. Instead, an equivalent BDT is trained using the same settings and input observables.

For the training of this new BDT and throughout this chapter, three datasets are used. The background process comprises of the $WZjj$ -EW4 prediction in combination with all other prompt backgrounds. These are the backgrounds that are considered in the training of the BDT. The remaining background, which is not included, is the non-prompt background, has a significant fraction of negative event weights. Such negative event weights are potentially problematic in the training process and therefore avoided. The consequence of this is that the BDT is not ideally trained to separate the non-prompt background from the signal. Since this is a subleading contribution to the overall background processes, this is accepted as a trade-off.

The signal process used for the training is the $WZjj$ -EW6 prediction derived using SHERPA as listed in Table 3.1. This sample has a known issue in the modelling of the shower kinematic as discussed in Section 3.3, but offers a large number of events. An alternative signal model derived using MG5_AMC and HERWIG 7 is studied. The modelling of the shower kinematic is expected to be better in this sample, which has been studied in Section 3.7.2. Since the detector simulation in this sample was performed privately, the available statistics are very limited and not sufficient for the training of multi-variate techniques.

The new BDT serves as a baseline, and the predicted event yield for the combined background and the signal process are shown differentially in the **BDTscore** of this baseline BDT in Figure 7.1.

In this chapter, the goal is to adopt the list of input observables of the BDT to avoid observables which are severely affected by modelling issues discussed in Section 3.3, and to

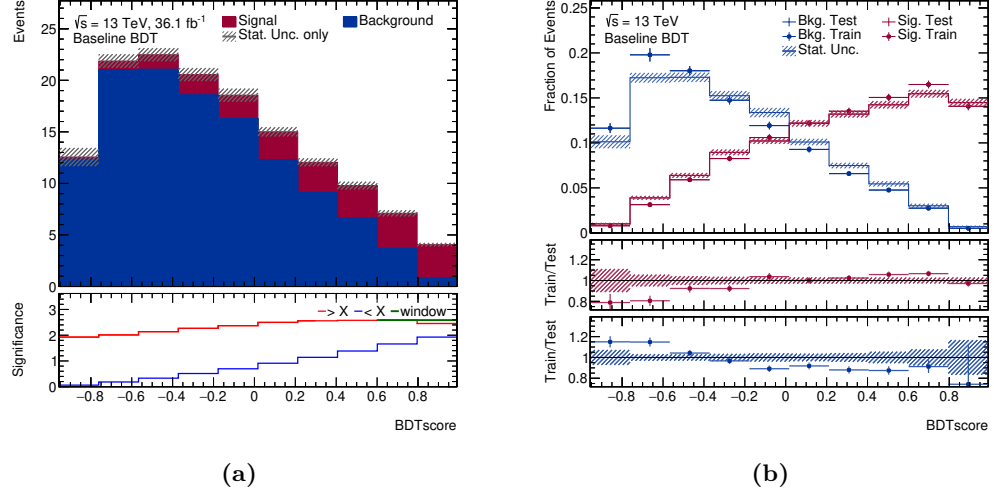


Figure 7.1: Comparison of distributions differentially in the **BDTscore** of the baseline BDT for signal (red) and background (blue) simulations in the signal region. The left plot shows the stacked differential event yields as estimated from the test datasets with an inlay indicating the estimated statistical significance for different cut scenarios. The right plot shows predicted distributions normalized to unit area for the training dataset (full circle) and the test dataset (hatched band) with two inlays showing the ratio of training over test fractions in each bin for the background and signal datasets respectively. The indicated uncertainties represent only the statistical uncertainties due to the limited amount of events in the simulations.

optimize the training of the BDT in order to re-increase performance.

7.1 Basic Principles of Boosted Decision Trees

Especially for the optimization of the performance, a basic understanding of BDTs is necessary. A brief introduction restricted to the basic principles and settings used in this work is presented below. For a more detailed description, including the details of the technical implementation in the **TMVA** framework are presented in [162].

The basic building blocks of BDTs are simple binary decision trees. Binary decision trees split the full region using a series of binary cuts using one input observable at a time into smaller subregions, called leaf nodes. Each of the leaf nodes is classified as either signal or background. The series of cuts has a tree-like structure, i.e. what criterion is checked next depends on the result of the previous decision. The resulting structure is visualized using an example decision tree of the baseline BDT in Figure 7.2.

Binary decision trees and their splitting criteria can be derived using labelled events, i.e. groups of events for the signal or background processes, respectively. The process of derivation of a decision tree is also referred to as growing or training and follows a simple procedure recursively.

At each node, the splitting criterion is searched based on the gini-index defined as $p \cdot (1-p)$ for a node with the signal purity p using the predicted background and signal processes. The gini-index has its maximum at $p = 0.5$ and decreases towards $p = 0$ and towards $p = 1$ to 0. The gini-index can be interpreted as a measure for how well mixed a node is. A node is split if its gini index is larger than the weighted combination of the gini-indices of the two subnodes created by a cut. The maximum decrease of the gini-index for all tested cut values in all observables is used for the split of the node.

The cut criterion with maximal separation is searched in a grid search, testing all input variables and testing **nCut** equidistant cut values over the full range of each observable. Since this is a simple grid search only, the computing time scales linearly with the number

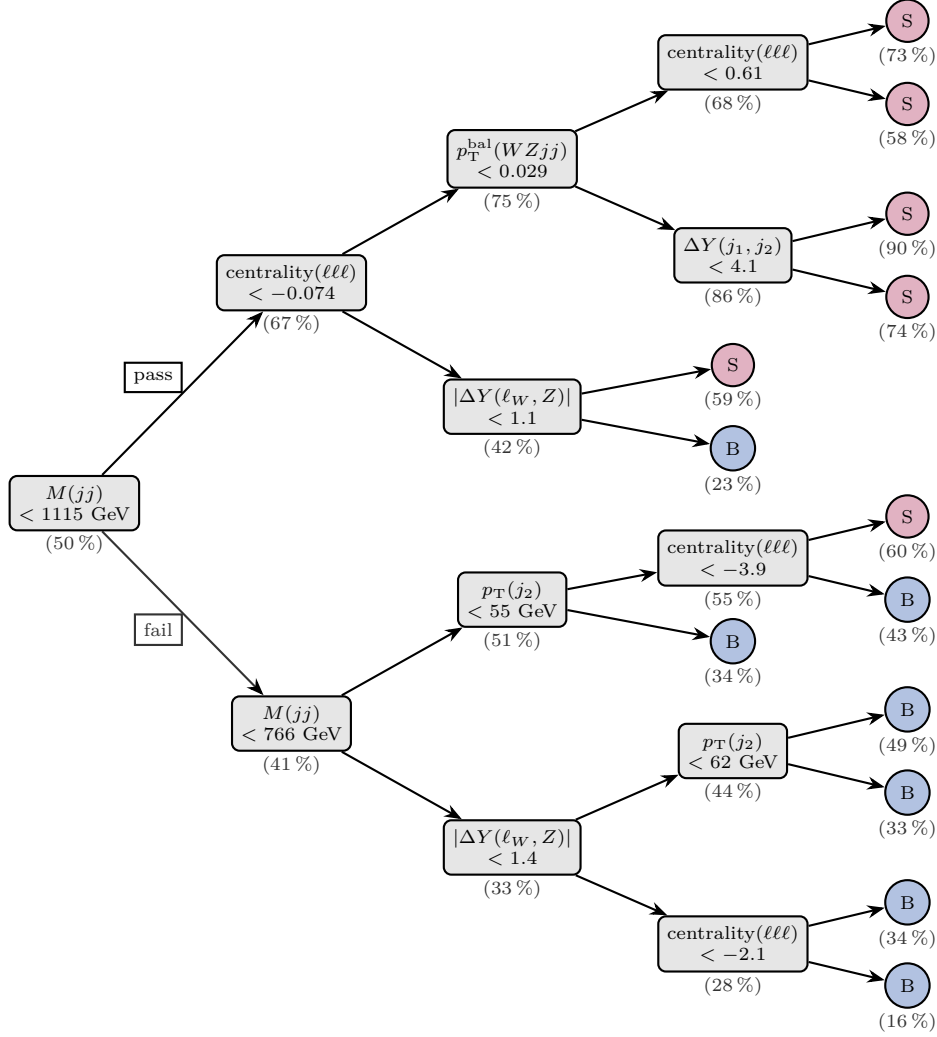


Figure 7.2: Visualization of a single binary decision tree from the set of trees in the baseline BDT. Events are classified by evaluating the selection indicated in the current node (rounded rectangles) starting from the root node on the left. Events passing the indicated cut in each node follow the upper edge, those failing follow the lower edge. Once a leaf node (circle) is reached the event is classified depending on whether the leaf node has a signal purity of more than 50 % (signal leaf node), or as background otherwise. The signal purity in each node is indicated in grey below each node.

of tested cuts per observable **nCut** and the number of input observables. The performance of the maximization, i.e. how well the actual best value is approached depends on the full range of the input parameter and the number of cuts **nCut**.

The determined cut separates the initial node in two new nodes, and for each of the new nodes, a series of stop conditions are tested to decide whether to split this node again or to accept this node as a leaf node. The stop conditions are

- maximum depth of the tree of **MaxDepth**, i.e. the total number of decisions in this chain of nodes
- minimum fraction of training events **MinNodeSize**

if any of these conditions fail, the node is not split further. Such nodes are accepted as leaf nodes and classified as a signal in case the predicted fraction of signal events is larger than 0.5 or as background otherwise.

The result of such a decision tree is a binary decision whether an event is classified as signal or background based on the classification of the leaf nodes the event falls in. A problem arises from the instability of such decision trees due to fluctuations in the set of training events. Such fluctuations have a substantial effect and can drastically affect the structure of the output tree. Systematic uncertainties caused by such migrations between the classified categories are quite large and can not be estimated easily.

A way to circumvent this issue is to train multiple decision trees, where the training set is reweighted or different subsets of events are used for the training in each decision tree. The parameter **NTrees** defines the number of requested trees. The results of all trees can then be combined to yield a final output score, typically ranging from -1 to 1.

The technique applied in this work for training and combining multiple trees is called *gradient boosting*. In this approach, the set of training events are iteratively reweighted, by increasing the weight w_i of a misclassified event i by

$$w_i \rightarrow w_i \cdot \frac{1-f}{f} \quad (7.1)$$

where f is the misclassification rate of the decision tree of the previous boosting iteration. After this reweighting, all event weights are decreased by a constant factor to correct the change in the sum of weights in each event sample.

In order to reduce sensitivity to statistical fluctuations in the training sample further, the reweighting in gradient boosting is combined with a *bagging* procedure. In this combination, each iteration of the training uses a different random subset of events of the training set. The fraction of events used in each iteration is given by the parameter **BaggedSampleFraction** and was chosen to be 0.6.

In each iteration m of the boosting, the newest tree is assigned a tree weight γ , while the weights of previous trees are not adjusted. A loss function is defined to quantify the deviation from the true label y (+1 for signal, -1 for background) in each event

$$L(\text{BDTScore}_m(\gamma)) = \ln(1 + \exp(-2 \cdot \text{BDTScore}_m(\gamma) \cdot y)) , \quad (7.2)$$

where $\text{BDTScore}_m(\gamma)$ is the output of the full boosted decision tree after iteration m and is obtained by appending the weighted decision tree in this iteration to the previous model. The tree weight γ is chosen to minimize the total loss function summed over all events and the boosted decision tree is adjusted to include the newest decision tree in the weighted average.

A new parameter called **Shrinkage** can be introduced scaling the parameter γ , that minimizes the loss function. By using $\text{Shrinkage} \cdot \gamma$ with small values for **Shrinkage** as tree weight, the effect of the newest iteration, i.e. a learning rate, can be reduced. This procedure increases the robustness and accuracy of the prediction but increases the computation time for the training.

The sensitivity of machine learning techniques to random fluctuations in the training data is referred to as *overtraining* and implies that the model can not be generalized to new data. Such overtraining artificially increases performance in the training data and may even

reduce performance in general and thus should be avoided. In general, overtraining occurs when there are too many free parameters in the model for a given number of training events. BDTs are typically prone to overtraining due to a large number of nodes and trees.

There is no clear criterion for overtraining, but an often-used procedure is to split the available training data into two non-overlapping subsets, where one is used for training (“training sample”), and the other (“test sample”) is exclusively used to evaluate the performance of the trained BDT on a statistically independent sample. If the performance in the test sample is equivalent to the training sample, the effect of overtraining is assumed to be small. This test is applied here, with a split such that the training and test samples have approximately the same numbers of events.

7.1.1 Evaluation of performance

An evaluation of the performance of multi-variate techniques is crucial for a reasonable optimization as well as for reliable tests to avoid overtraining. The easiest approach is to apply the BDT on two datasets and to compare the distributions of the **BDTscore**. A BDT with good performance should result in large disagreement between the distributions for signal and background samples, while the distributions in the test and training samples should agree.

It is helpful to use evaluation metrics to quantify the agreement of the two considered **BDTscore** distributions h_1 and h_2 . Several different evaluation metrics were considered in this work and are described in the following. Some evaluation metrics are applied on normalized distributions and are only based on the shapes of the distributions. In this cases the input distributions are labeled n_1 and n_2 .

Reduced χ^2 $f_{\text{red. } \chi^2}(n_1, n_2)$. A common test of compatibility of two distributions is the χ^2 -test. A χ^2 -value is calculated for two normalized distributions n_1 and n_2 as

$$\chi^2 = \sum_{i=1}^{N_{\text{bins}}} \frac{(n_{1,i} - n_{2,i})^2}{\Delta n_{1,i}^2 + \Delta n_{2,i}^2}, \quad (7.3)$$

where i iterates over all bins, $n_{k,i}$ is the predicted value in bin i for distribution n_k and $\Delta n_{k,i}$ is the absolute uncertainty in bin i for distribution n_k . This χ^2 value is divided by the number of degrees of freedom to counteract the sum of the bins. For distributions in good agreement, a reduced χ^2 of about 1 is expected, while larger values indicate disagreements.

Non-overlap $f_{\text{N-O}}(n_1, n_2)$. This evaluation metric is based on the overlap of distributions. The overlap is calculated by summing over the minima of the normalized distributions in each bin. For the sake of readability, the final value is determined by subtracting the fraction of overlapping integrals from 1

$$f_{\text{N-O}}(n_1, n_2) = 1 - f_{\text{O}}(n_1, n_2) = 1 - \sum_{i=1}^{N_{\text{bins}}} \min(n_{1,i}, n_{2,i}). \quad (7.4)$$

If distributions are in perfect agreement, the minimum in each bin corresponds to the value itself and since the distributions are normalized, the metric vanishes $f_{\text{N-O}} = 0$. Any deviation of n_1 from n_2 will increase the metric value $f_{\text{N-O}}$.

Separation $f_{\text{Sep}}(n_1, n_2)$ The separation is a performance evaluation metric for normalized distributions and is already implemented in the **TMVA** framework[162] and was originally proposed in [163]. It is defined as

$$f_{\text{Sep}}(n_1, n_2) = \sum_{i=1}^{N_{\text{bins}}} \frac{(n_{1,i} - n_{2,i})^2}{n_{1,i} + n_{2,i}}, \quad (7.5)$$

where bins that are empty in both distributions are skipped.

Approximated ROC-integral $f_{\text{ROC}}(\mathbf{n}_1, \mathbf{n}_2)$. A commonly used evaluation metric is the integral under the *receiver operating characteristic (ROC)* curve, commonly referred to as *area under the curve (AUC)*. For the ideal ROC curve, the background rejection is plotted differentially over the signal acceptance. The curve is determined by continuously varying a cut selection threshold. In the implementation used here, only binned distributions are available, introducing steps in the ROC curve. The discrete sampling of the ROC curve limits this metric only to approximate the full ROC integral.

Statistical significance $f_{\text{Zstat}}(\mathbf{h}_1, \mathbf{h}_2)$. A metric sensitive on the actual normalization of the compared processes is based on the significance, estimated only including statistical uncertainties as

$$Z = \frac{S}{\sqrt{S+B}}. \quad (7.6)$$

In this metric h_1 is assigned to be the signal process and h_2 is considered as background. Left- and right-sided cuts at each bin threshold are tested. For each tested scenario, S (B) is the sum over all accepted bins of h_1 (h_2). The maximum significance Z is chosen as output of the evaluation metric.

For evaluation of the classification performance, this metric is closest to the actual application of the **BDTscore** in a subsequent likelihood-based fit. For tests of overtraining this metric can not be considered directly since no direct conclusions on the agreement can be drawn. However, if f_{Zstat} is determined for signal compared to the background in the training dataset and subsequently in the test dataset, any differences between f_{Zstat} in the training and test dataset indicate effects of overtraining.

Total significance $f_{\text{Ztot}}(\mathbf{h}_1, \mathbf{h}_2)$. The total significance f_{Ztot} is determined equivalently to the statistical uncertainty, but including an estimated systematic uncertainty on the background yield of 30 % by adjusting the significance estimate to

$$Z = \frac{S}{\sqrt{S+B+(0.3 \cdot B)^2}}. \quad (7.7)$$

The same notation and caveats mentioned for f_{Zstat} apply.

7.2 Selection and Preprocessing of Input Variables

A new BDT is trained, with the set of input observables in the baseline BDT as a starting point. These observables can be grouped as follows

Transverse momenta The transverse momenta of the leading jet $p_{\text{T}}(j_1)$, the subleading jet $p_{\text{T}}(j_2)$, the W^\pm boson $p_{\text{T}}(W)$, and the Z boson $p_{\text{T}}(Z)$ are used. For the W^\pm and Z bosons the transverse momenta are evaluated from the assigned leptons or missing transverse energy in case of the W^\pm boson.

Rapidity The rapidities of the leading jet $Y(j_1)$ and the W^\pm boson were included directly, which relies on a reconstruction of the full momentum of the W^\pm boson. Assuming the $E_{\text{T}}^{\text{miss}}$ originates from a single neutrino of a W^\pm decay with the assigned lepton ℓ_W , the otherwise-missing longitudinal momentum of the neutrino can be estimated by constraining the invariant mass of ℓ_W and the hypothetical neutrino to $m_W = 80.379$ GeV following [164].

Tagging jet system The kinematic of the system of tagging jets is included via the absolute rapidity difference $\Delta Y(j_1, j_2)$, the absolute difference in the transverse plane $\Delta\phi(j_1, j_2)$, as well as the invariant mass M_{jj} .

It is noteworthy, that the rapidity of the subleading jet is not included directly, but only via the rapidity difference to the leading jet. This approach not only has the advantage of invariance against lorentz boosts along the beam axis (see Section 4.2.2), but also transforms the symmetrical distribution of $Y(j_2)$ to an observable that is easier

to use in a BDT, which only applies one-sided cuts at each step, while still including the information about the relation between the rapidities of the jets.

Leptonic system The rapidity difference in the leptonic system are summarized in the observable $|\Delta Y(\ell_W, Z)|$, which is the absolute rapidity difference between the lepton assigned to the W^\pm boson ℓ_W and the Z boson. Also, the transverse mass of the $W^\pm Z$ system $M_T(WZ)$ is used, which is defined equivalently to Equation (3.11), where $p_T(\nu)$ is replaced by E_T^{miss} .

Jet multiplicity The exclusive multiplicity of jets passing the jet object selection criteria listed in Section 5.2.3 N_{jets} is included. This observable is the only input observable that is not a floating-point number but an integer, which is usually considered problematic in BDTs.

Angular distance The observable $\Delta R(j_1, Z)$ measures the angular distance between the leading jet and the Z boson.

Lepton centrality The lepton centrality labelled $\text{centrality}(\ell\ell\ell)$ is defined based on the centrality of an object ℓ

$$\text{centrality}(\ell) = \min(Y(j_a) - Y(\ell), Y(\ell) - Y(j_b)), \quad (7.8)$$

where the rapidities of the tagging jets are ordered, so that $Y(j_a) > Y(j_b)$. The lepton centrality is the minimum of the object centralities of the three leptons ℓ_W , $\ell_{Z,1}$, and $\ell_{Z,2}$ and is positive, when all leptons are between the tagging jets in terms of the rapidity.

Transverse momentum balance The transverse momentum balance of a group of n particles p_i is defined as the absolute value of the vectorial sum over the scalar sum of the transverse momenta of the objects

$$p_T^{\text{bal}}(p_1, \dots, p_n) = \frac{|\sum_{i=1}^n \vec{p}_T(p_i)|}{\sum_{i=1}^n |\vec{p}_T(p_i)|}. \quad (7.9)$$

Based on this, p_T^{bal} of the hard system $p_T^{\text{bal}}(WZjj)$ is calculated from the four-momenta of the tagging jets and the W^\pm and Z bosons.

The set of observables used in the baseline BDT is first extended to build a set of available observables to be studied as input observables. Additional low-level observables like transverse momenta and rapidities are added for other particles as well as other high-level observables combining information of multiple objects.

Transverse momenta The list of transverse momenta to be considered is extended to include the leptons $p_T(\ell_W)$, $p_T(\ell_{Z,1})$, $p_T(\ell_{Z,2})$ as well as the missing transverse energy E_T^{miss} .

Rapidity The rapidities of the Z boson $Y(Z)$ and the leptons $Y(\ell_W)$, $Y(\ell_{Z,1})$, $Y(\ell_{Z,2})$ are tested as well. The rapidity of the second jet $Y(j_2)$ is included explicitly as well, although already indirectly included via the rapidity difference to the leading jet. This observable will allow testing the hypothesis that the rapidity difference performs better than the plain rapidity.

Tagging jet system For the sake of completeness, also the angular difference $\Delta R(j_1, j_2)$ between the tagging jets is included.

Leptonic system For the leptonic system, also the invariant masses of the Z boson $M(Z)$ as well as the tri-lepton system $M(\ell\ell\ell)$, and the transverse mass of the W^\pm boson $M_T(W)$ are included.

Lepton centrality The centrality as defined in Equation (7.8) is also included for the system of leptons assigned to the Z boson centrality($\ell\ell$).

Furthermore, an adjusted definition usually referred to as Zeppenfeld variable ζ ¹ is slightly adapted from the original [165] and defined for an object p with rapidity $Y(p)$ as

$$\zeta(p) = \left| \frac{Y(p) - 0.5 \cdot (Y(j_1) + Y(j_2))}{Y(j_1) - Y(j_2)} \right|, \quad (7.10)$$

with the rapidities of the tagging jets $Y(j_1)$ and $Y(j_2)$. This observable is 0 for objects exactly in the center of the two tagging jets and 0.5 for objects with the same rapidity as one of the tagging jets.

When considering $\zeta(p_1, \dots, p_n)$ of multiple objects, the maximum of the $\zeta(p_i)$ is used. The maximum of ζ of the two leptons assigned to the Z boson is labelled $\zeta(\ell\ell)$ and used as input as well as the maximum of ζ of all three leptons $\zeta(\ell\ell\ell)$.

Transverse momentum balance The p_T^{bal} as introduced before is used as well for other sets of particles; for the system of tagging jets $p_T^{\text{bal}}(jj)$, the three leptons in combination with $E_T^{\text{miss}} p_T^{\text{bal}}(\ell\ell v)$, and for the combination of the tagging jets, the three leptons, and $E_T^{\text{miss}} p_T^{\text{bal}}(\ell\ell v jj)$.

Angular distance Pair-wise angular distances are considered as well between the three signal leptons $\Delta R(\ell_W, \ell_{Z,1})$, $\Delta R(\ell_W, \ell_{Z,2})$, and $\Delta R(\ell_{Z,1}, \ell_{Z,2})$. Additionally, for each tagging jet, the minimum angular distance to any of the three signal leptons is included: $\min \Delta R(j_1, \ell)$ and $\min \Delta R(j_2, \ell)$.

7.2.1 Modelling of input observables

The full set of input observables is carefully validated using the alternative sample produced with MG5_AMC and HERWIG 7. Differences between the two signal predictions in an observable are accepted only in case the SHERPA signal sample is closer to the predicted distribution for the background processes. In this case, the BDT will not be trained ideally, but the MG5_AMC-based prediction is then expected to be classified as more signal-like.

All of the considered observables fulfil this requirement. The four shape-based evaluation metrics for the comparison of the signal predictions are listed for each observable in Table 7.1. The observables with the largest differences are the transverse momentum balance of the $WZjj$, and $\ell\ell v jj$ systems $p_T^{\text{bal}}(WZjj)$ and $p_T^{\text{bal}}(\ell\ell v jj)$, the absolute differences in the rapidity $\Delta Y(j_1, j_2)$ and the angular distance $\Delta R(j_1, j_2)$ between the tagging jets, the jet multiplicity N_{jets} , as well as the invariant mass of the tagging jets M_{jj} . Normalized differential distributions in those six observables are shown in Figure 7.3 for the two signal and the total background processes. For the other observables, the distributions of the SHERPA and MG5_AMC predictions agree within uncertainties and are shown in Appendix F.1 for completeness. In each distribution in almost every bin, the SHERPA-based signal prediction is between the background and the MG5_AMC-based signal prediction.

While the modelling of the jet multiplicity passed the general criterium on the modelling, the modelling is solely driven by the parton shower algorithm. Additionally, it is the only integer input for the BDT. In order to avoid issues, the jet multiplicity is removed from the full list of considered observables.

With this, the full list of considered observables is complete, and a new BDT is trained with the same settings to evaluate the effect of adding several new observables to the baseline set of observables. The predicted event yield, as well as normalized distributions for the combined background and the signal process, are shown differentially in the **BDTscore** for this new BDT in Figure 7.4. Quantitative comparisons of the different BDTs in each step of the full optimization process using the evaluation metrics are shown in Section 7.3.3 in Table 7.4.

¹Common definitions of the Zeppenfeld observable vary, e.g. without dividing by the rapidity separation or with an additional factor of 2.

Table 7.1: Overview of classification evaluation metrics comparing the SHERPA and MG5_AMC signal predictions for each considered input observable ordered by decreasing Non-Overlap $f_{\text{N-O}}$.

Observable	$f_{\text{N-O}}$	f_{Sep}	$f_{\text{red. } \chi^2}$	f_{ROC}
$p_{\text{T}}^{\text{bal}}(WZjj)$	0.12	0.026	21.40	0.42
$p_{\text{T}}^{\text{bal}}(\ell\ell vjj)$	0.12	0.026	20.37	0.42
$\Delta R(j_1, j_2)$	0.08	0.011	3.87	0.50
$\Delta Y(j_1, j_2)$	0.07	0.008	2.59	0.50
N_{jets}	0.06	0.007	5.14	0.38
M_{jj}	0.06	0.007	3.71	0.43
$Y(\ell_{Z,2})$	0.06	0.005	1.72	0.48
$\Delta R(j_1, Z)$	0.05	0.005	1.82	0.45
$p_{\text{T}}^{\text{bal}}(jj)$	0.05	0.004	1.21	0.49
$\zeta(\ell\ell)$	0.05	0.005	2.68	0.44
$\min \Delta R(j_1, \ell)$	0.05	0.004	1.45	0.46
$Y(j_2)$	0.05	0.003	1.23	0.48
$\Delta\phi(j_1, j_2)$	0.05	0.004	1.63	0.48
$p_{\text{T}}^{\text{bal}}(\ell\ell v)$	0.05	0.005	2.01	0.48
$Y(\ell_{Z,1})$	0.05	0.003	1.03	0.47
centrality($\ell\ell$)	0.05	0.004	2.45	0.47
$Y(Z)$	0.04	0.003	0.97	0.47
$Y(\ell_W)$	0.04	0.003	0.89	0.48
$Y(j_1)$	0.04	0.003	1.20	0.47
$\Delta R(\ell_W, \ell_{Z,2})$	0.04	0.003	1.49	0.47
centrality($\ell\ell$)	0.04	0.004	1.41	0.46
$ \Delta Y(\ell_W, Z) $	0.04	0.003	1.59	0.47
$\Delta R(\ell_{Z,1}, \ell_{Z,2})$	0.04	0.002	0.79	0.48
$Y(W)$	0.04	0.003	0.85	0.47
$M(Z)$	0.04	0.002	0.72	0.46
$\Delta R(\ell_W, \ell_{Z,1})$	0.04	0.004	1.27	0.47
$\min \Delta R(j_1, \ell)$	0.04	0.003	0.87	0.47
$p_{\text{T}}(\ell_W)$	0.03	0.003	2.99	0.35
$\zeta(\ell\ell)$	0.03	0.003	2.06	0.43
$p_{\text{T}}(j_1)$	0.03	0.004	2.54	0.39
$M_{\text{T}}(WZ)$	0.03	0.002	1.90	0.31
$p_{\text{T}}(j_2)$	0.03	0.003	1.80	0.32
$p_{\text{T}}(W)$	0.03	0.003	2.28	0.35
$p_{\text{T}}(\ell_{Z,2})$	0.02	0.002	1.01	0.37
$p_{\text{T}}(Z)$	0.02	0.001	0.99	0.38
$E_{\text{T}}^{\text{miss}}$	0.02	0.002	1.86	0.34
$p_{\text{T}}(\ell_{Z,1})$	0.02	0.002	1.70	0.35
$M(\ell\ell)$	0.01	0.002	2.15	0.35
$M_{\text{T}}(W)$	0.00	0.000	1.25	0.27

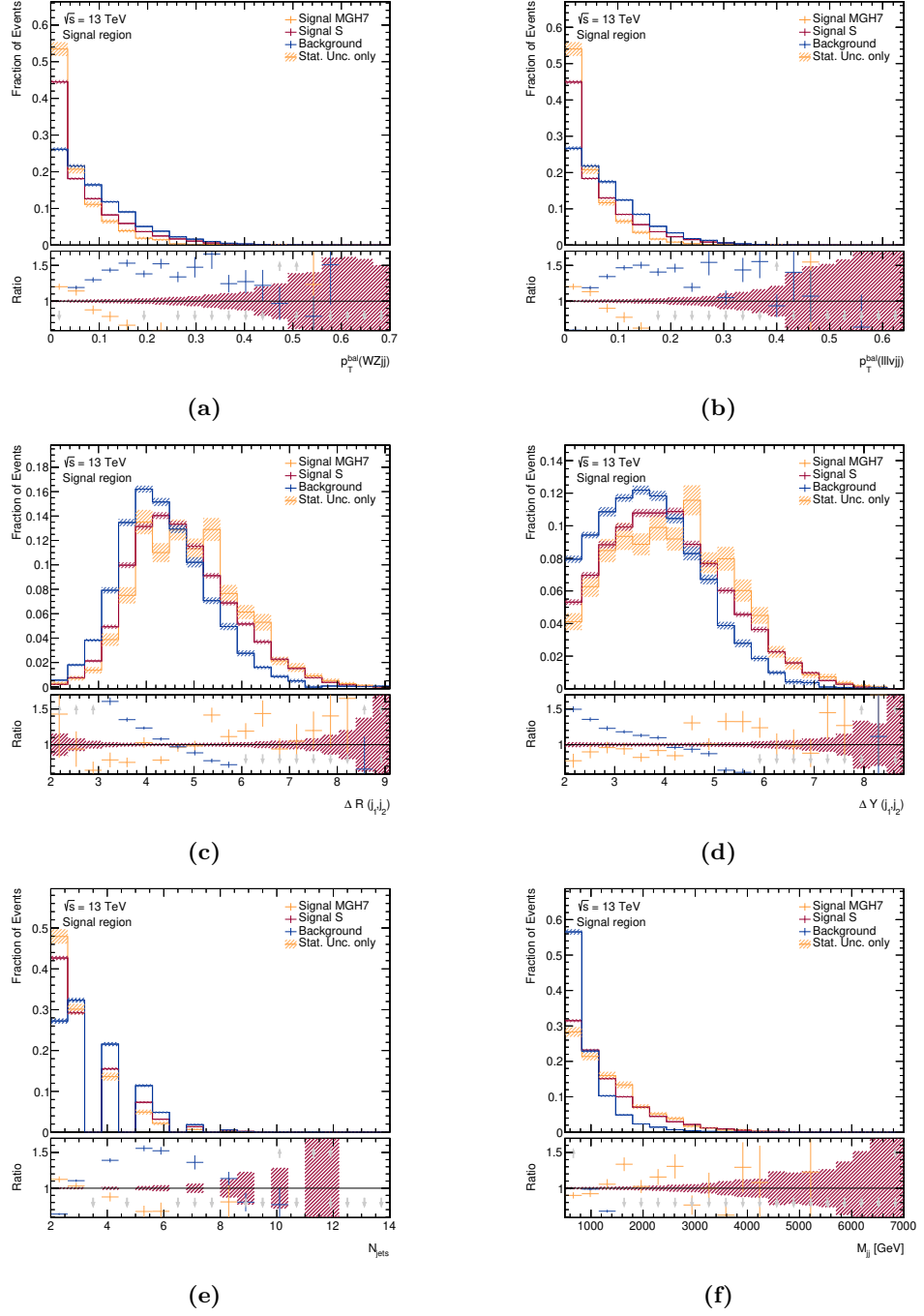


Figure 7.3: Comparison of normalized differential predictions for the SHERPA (red) and MG5_AMC-based (orange) signal prediction as well as the prediction for the combined prompt backgrounds (blue) in different input observables in the signal region. Inlays show the ratio of each of the other distributions with respect to the normalized SHERPA-based signal prediction. Figure (a) and Figure (b) show the transverse momentum balance of the $WZjj$ system ($p_T^{\text{bal}}(WZjj)$) and the $\ell\ell\nu jj$ system ($p_T^{\text{bal}}(\ell\ell\nu jj)$) respectively. Figure (c) and Figure (d) show angular distance ($\Delta R(j_1, j_2)$) and the distance in rapidity ($\Delta Y(j_1, j_2)$) between the tagging jets. Figure (e) shows the exclusive jet multiplicity N_{jets} and Figure (f) shows the invariant mass of the tagging jets M_{jj} . The indicated uncertainties represent only the statistical uncertainties due to the limited amount of events in the simulations.

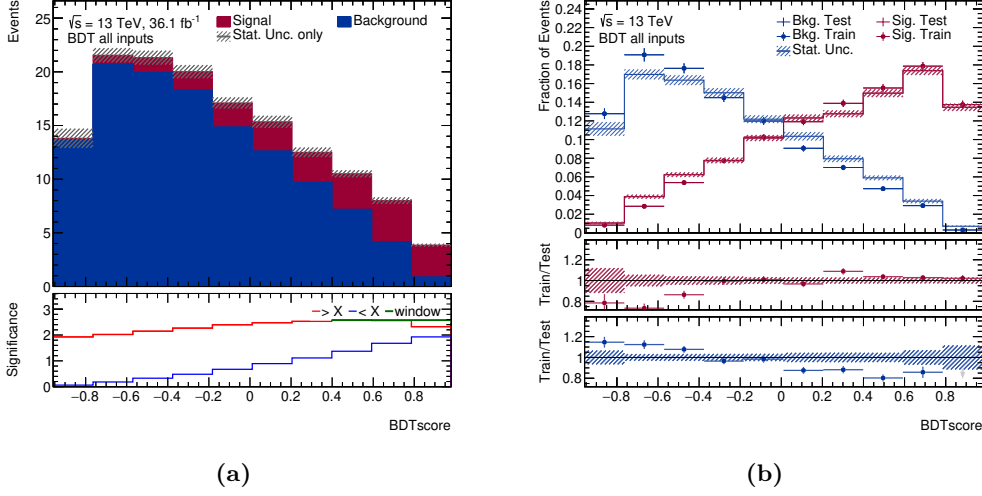


Figure 7.4: Comparison of distributions differentially in the **BDTscore** of the BDT trained using all considered input observables for signal (red) and background (blue) simulations in the signal region. The left plot shows the stacked differential event yields as estimated for different test datasets with an inset indicating the estimated statistical significance for different cut scenarios. The right plot shows predicted distributions normalized to unit area for the training dataset (full circle) and the test dataset (hatched band) with two insets showing the ratio of training over test fractions in each bin for the background and signal datasets respectively. The indicated uncertainties represent only the statistical uncertainties due to the limited amount of events in the simulations.

7.2.2 Preprocessing

It is useful to apply a preprocessing to the input observables in order to avoid technical and physical issues. The distributions in Figure 7.4 were designed to reproduce the available information to select the best possible cut during the training of the BDT in the TMVA implementation with the baseline settings. The range of each observable accurately represents the spread from the minimum to the maximum value and is separated in 20 bins. This definition means the tested cut values depend on the random values for the minimum and maximum. For instance in Figure 7.3f the occurrence of a random event in any dataset with slightly over 7 TeV invariant mass of the tagging jet dictates the step size between considered cut values and only very few of the tested cuts are actually in the area of interest around 1 TeV. For many of the observables, a large part of the tested distribution does not carry relevant information. Also, the small number of tested cuts in the actual region of interest indicates a non-optimal convergence and training. Different approaches are followed in order to avoid this technical issue.

A first step is to transform the shape of the input variables to reduce the effect of high-energy tails. It has been observed before that applying a logarithm on observables with such high-energy tails, increases the performance of training in machine learning. While the ordering of events is not changed, the effect of long tails with low overall contributions is reduced. This transformation is applied to all invariant masses, transverse masses, transverse momenta, and p_T^{bal} -based observables.

The effect of outlier events in tails can be avoided entirely by manually restricting the range of input observables. The range can be restricted by replacing values larger or lower than a threshold with that threshold. This approach is equivalent to typical over- or underflow bins in histograms. The events with larger values are still considered but not distinguished from each other in the affected observable anymore.

Another transformation that is applied is to use absolute value rather than allowing for negative rapidities. For individual particles, the rapidity is symmetric with respect to the $Y = 0$ axis. The main physics differences between the processes are how forward particular

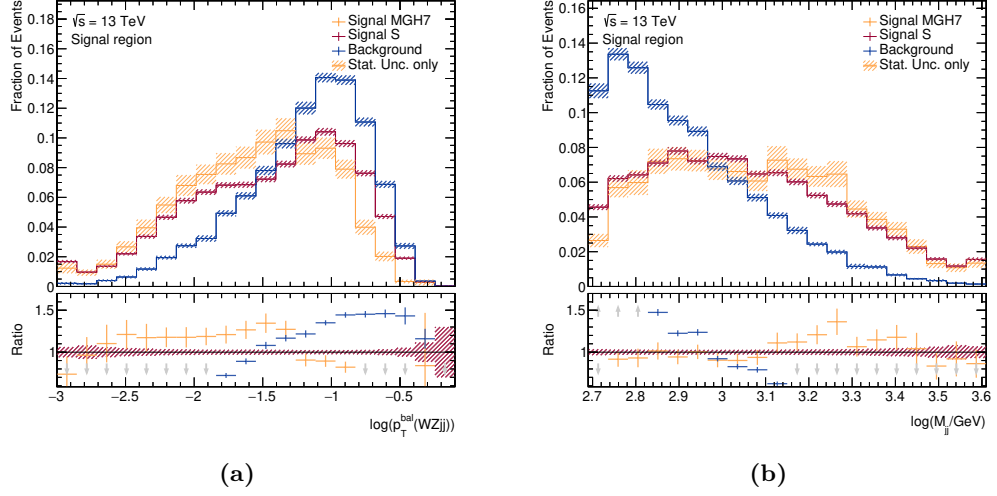


Figure 7.5: Comparison of normalized differential predictions for the SHERPA (red) and MG5_AMC-based (orange) signal prediction as well as the prediction for the combined prompt backgrounds (blue) in different input observables after preprocessing in the signal region. Inlays show the ratio of each of the other distributions with respect to the normalized SHERPA-based signal prediction. Figure (a) shows the transverse momentum balance of the $WZjj$ system $p_T^{\text{bal}}(WZjj)$ after preprocessing with a log-transformation and a restriction of the range. Figure (b) shows the invariant mass of the tagging jets M_{jj} after preprocessing with a log-transformation and a restriction of the range. The indicated uncertainties represent only the statistical uncertainties due to the limited amount of events in the simulations.

objects are. Since the BDT only applies simple left- or right-sided cuts, rather than, e.g. window cuts, the full potential of rapidities of individual objects cannot be evaluated by default since one of the two detector sides can be assigned to a single category in a single cut. For absolute values of rapidities, the forward and central parts can be differentiated appropriately.

Other observables compensate for the loss of information about rapidity differences between object pairs. For some object combinations, the rapidity differences are included directly, for other combinations the rapidity differences are available for the BDT indirectly via angular distances in ΔR or via centrality based observables.

A list of all observables with the applied transformations, including the evaluation metrics for the comparisons of the combined background and the SHERPA-based signal model is presented in Table 7.2. Examples for the effect on the distributions are shown for the transverse momentum balance of the $WZjj$ system as well as the invariant mass of the tagging jets $M(jj)$ in Figure 7.5.

Again a new BDT is trained with the same settings to evaluate the effect of the transformations. The predicted event yield as well as normalized distributions for the combined background and the signal process are shown differentially in the `BDTscore` for this new BDT in Figure 7.6.

7.3 Optimization

With the full list of considered input observables, a new BDT is constructed and optimized from the ground up. The goal is to minimize modelling uncertainties by using a number of input observables that is as low as reasonably possible while keeping or increasing the overall performance of the BDT.

The construction of the BDT is separated into three phases. In the first phase, the smallest reasonable set of input observables is determined with the meta-parameters of the

Table 7.2: Overview of classification evaluation metrics comparing the background and SHERPA-based signal predictions for each considered input observable as well as the applied transformations ordered by decreasing Non-Overlap $f_{\text{N-O}}$.

Observable	Evaluation metric				Transformations		
	$f_{\text{N-O}}$	f_{Sep}	$f_{\text{red. } \chi^2}$	f_{ROC}	Function	min	max
M_{jj}	0.27	0.102	142.21	0.65	log		3.6
centrality($\ell\ell$)	0.23	0.073	98.20	0.62		-2	3
$\zeta(\ell\ell)$	0.22	0.066	100.00	0.61			1.2
centrality($\ell\ell\ell$)	0.22	0.072	97.76	0.61		-2.2	2.8
$\zeta(\ell\ell\ell)$	0.22	0.064	97.92	0.61			1.2
$p_{\text{T}}^{\text{bal}}(WZjj)$	0.19	0.053	80.40	0.58	log	-3	
$p_{\text{T}}^{\text{bal}}(\ell\ell\nu jj)$	0.18	0.053	79.96	0.58	log	-3	
$\min \Delta R(j_1, \ell)$	0.15	0.038	39.55	0.55			6
$p_{\text{T}}(j_2)$	0.15	0.032	41.88	0.56	log		2.8
$Y(j_1)$	0.14	0.027	34.77	0.56	abs		
$ \Delta Y(\ell_W, Z) $	0.14	0.032	50.15	0.55			4
$\Delta R(j_1, j_2)$	0.14	0.034	42.33	0.56			8
$M(\ell\ell\ell)$	0.12	0.023	34.31	0.54	log		3
$p_{\text{T}}^{\text{bal}}(jj)$	0.12	0.019	26.06	0.53	log	-1.3	
$\Delta Y(j_1, j_2)$	0.12	0.025	28.71	0.55			7
$\Delta R(\ell_W, \ell_{Z,1})$	0.11	0.020	30.54	0.54			5
$\Delta\phi(j_1, j_2)$	0.10	0.015	19.75	0.53			
$\Delta R(\ell_W, \ell_{Z,2})$	0.09	0.017	23.25	0.53			5
$\Delta R(j_1, Z)$	0.08	0.008	12.11	0.48			7
$p_{\text{T}}(j_1)$	0.07	0.012	9.43	0.51	log		3
$\min \Delta R(j_1, \ell)$	0.06	0.008	11.11	0.48			5
$p_{\text{T}}^{\text{bal}}(\ell\ell\nu)$	0.06	0.006	8.81	0.49	log	-1.3	
$Y(\ell_W)$	0.06	0.005	6.29	0.51	abs		
$Y(\ell_{Z,1})$	0.05	0.005	5.87	0.51	abs		
$Y(\ell_{Z,2})$	0.05	0.004	5.24	0.50	abs		
$Y(Z)$	0.05	0.004	5.59	0.50	abs		
$M_{\text{T}}(WZ)$	0.05	0.007	9.93	0.49	log	1.9	3
$Y(j_2)$	0.05	0.003	5.30	0.47	abs		
$Y(W)$	0.04	0.002	3.34	0.49	abs		
$M_{\text{T}}(W)$	0.04	0.003	3.53	0.46	log		2.4
$p_{\text{T}}(\ell_W)$	0.03	0.001	2.05	0.48	log		2.7
$E_{\text{T}}^{\text{miss}}$	0.03	0.001	1.83	0.44	log		
$\Delta R(\ell_{Z,1}, \ell_{Z,2})$	0.02	0.001	1.41	0.47			4
$p_{\text{T}}(W)$	0.02	0.001	1.97	0.47	log	0.8	3
$M(Z)$	0.02	0.001	1.02	0.46	log		
$p_{\text{T}}(\ell_{Z,1})$	0.02	0.002	1.60	0.45	log		
$p_{\text{T}}(Z)$	0.02	0.001	0.98	0.46	log	1	3
$p_{\text{T}}(\ell_{Z,2})$	0.01	0.001	1.35	0.46	log		

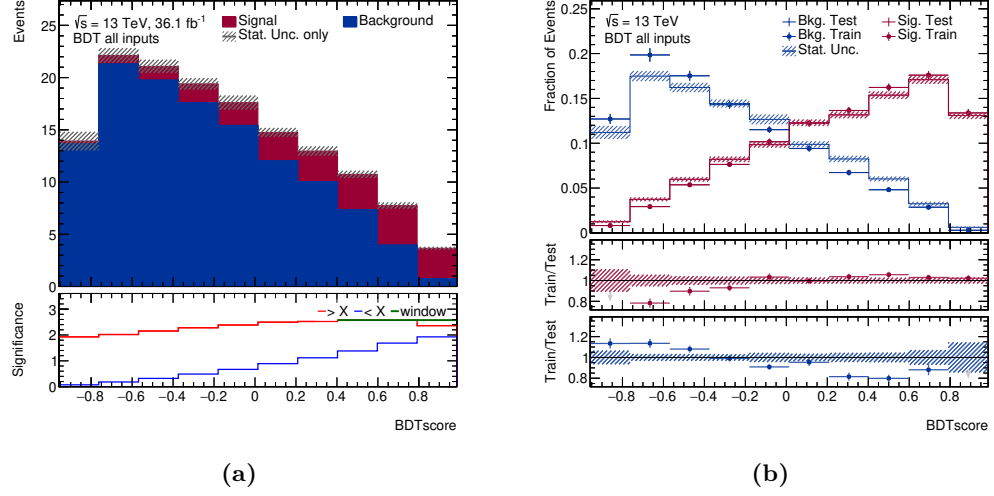


Figure 7.6: Comparison of distributions differentially in the **BDTscore** of the BDT trained using all considered input observables after pre-processing for signal (red) and background (blue) simulations in the signal region. The left plot shows the stacked differential event yields as estimated from the test datasets with an inlay indicating the estimated statistical significance for different cut scenarios. The right plot shows predicted distributions normalized to unit area for the training dataset (full circle) and the test dataset (hatched band) with two inlays showing the ratio of training over test fractions in each bin for the background and signal datasets respectively. The indicated uncertainties represent only the statistical uncertainties due to the limited amount of events in the simulations.

baseline BDT. In the second phase, the meta-parameters of the BDT are optimized with the newly determined set of input observables. The separation of phases is necessary in order to reduce the complexity of the optimization problem. The goal of the final phase is to combine the optimizations by advancing the search for input observables using the optimized set of meta-data.

7.3.1 Set of input observables

The search for the minimal set of input observables is performed iteratively, starting with a single observable. The invariant mass of the tagging jets M_{jj} is used as a starting point since this is the input observable with the strongest classification performance between signal and background. The meta-parameter **NTrees** is reduced to 100 compared to the baseline BDT in order to avoid overtraining due to the small number of input observables.

In each iteration step several adjustments to the setup are evaluated. The metric $f_{\text{red. } \chi^2}$ is applied to compare the **BDTscore** distributions for the test and training datasets individually for the signal ($f_{\text{red. } \chi^2}(S^{\text{test}}, S^{\text{train.}})$) and background processes ($f_{\text{red. } \chi^2}(B^{\text{test}}, B^{\text{train.}})$). If any of the two metrics is larger than 2, or if both are larger than 1.5, the respective BDT is considered overtrained and rejected. All remaining BDTs are ranked according to the evaluation of the signal vs background classification in the test dataset using the evaluation metric $f_{\text{Ztot}}(S^{\text{test}}, B^{\text{test}})$. The other considered evaluation metrics are also applied and reported for the performance evaluation on the test dataset.

The metric f_{Ztot} was chosen, because it takes the relative normalizations into account as opposed to shape-based metrics. The metrics based on significance estimates for cut-based scenarios have the additional advantage to focus on the parts of the distributions where the signal contributions are sizable. Due to the inclusion of an estimated systematic uncertainty, f_{Ztot} is also expected to yield results more consistent with the final fitting procedure discussed in Chapter 9.

First, the parameter **NTrees** is varied up and down by 100 in order to adapt to the increasing number of inputs. If one of these variations performs better than the previous

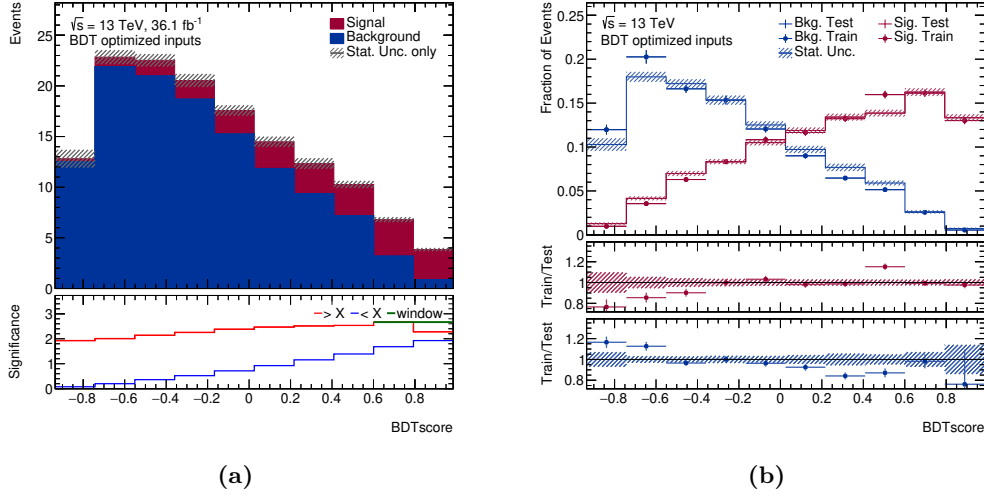


Figure 7.7: Comparison of distributions differentially in the **BDTscore** of the BDT with an optimized set of input observables and the meta-parameter configuration of the baseline BDT for signal (red) and background (blue) simulations in the signal region. The left plot shows the stacked differential event yields as estimated from the test datasets with an inlay indicating the estimated statistical significance for different cut scenarios. The right plot shows predicted distributions normalized to unit area for the training dataset (full circle) and the test dataset (hatched band) with two inlays showing the ratio of training over test fractions in each bin for the background and signal datasets respectively. The indicated uncertainties represent only the statistical uncertainties due to the limited amount of events in the simulations.

iteration and is not classified as overtrained, the meta-parameter is adjusted accordingly.

Afterwards, the set of input observables is varied by one observable in each possible combination. That means new BDTs are trained with reduced sets of inputs by removing any one of the input observables. With extended sets of inputs by adding any of the observables, that is considered but not included in the current set of input observables.

In case the best-ranking adjustment results from adding an observable, this observable is added to the set of inputs, and the next iteration step is started. If the BDT with the current settings or one with a reduced set of inputs performs best, the iterative procedure is stopped.

This procedure results in a set of 10 input observables: M_{jj} , $p_T^{\text{bal}}(\ell\ell\nu jj)$, $|\Delta Y(\ell_W, Z)|$, $\zeta(\ell\ell)$, $p_T(j_1)$, $M(\ell\ell)$, $p_T(\ell_{Z,2})$, $p_T(j_2)$, $\Delta R(j_1, j_2)$, $p_T(Z)$. The changes and the classification performances in each step are listed combined with each other step in the overall optimization chain in Table 7.4 below. The distributions of the BDT after this optimization phase are shown in Figure 7.7.

The performance using the $f_{Z_{\text{tot}}}$ metric is slightly increased compared to the baseline BDT. Also, for the other considered metrics, the performance is at least similar to the baseline BDT.

Interestingly, the BDT with this restricted set of input observables performs also better on the test sample than the BDT with the full set of pre-processed input observables. One possible reason is that a different metric is used during training of the BDT. Another potential reason is an overtraining of the BDT with all inputs. Having more input observables as potential inputs allow for more possibilities for statistical fluctuations, to which the training procedure might be sensitive. While the $f_{\text{red. } \chi^2}$ metrics are not considered problematic; a larger difference in performance between the test and training datasets is observed for the BDT with all input observables. This supports the motivation to remove unnecessary inputs from the BDT.

The final BDT shows mild indications of overtraining for the signal process consistently

around $f_{\text{red.}} \chi^2 \approx 1.5$ for the last four iterations. Since the level of disagreement is still small, the iterations are accepted nevertheless. Such disagreement is considered less problematic since the dataset used for the training and optimization of the BDT is not the sample used when comparing to data.

7.3.2 Meta variable optimization

A further improvement of the performance of the BDT can be achieved by optimizing the meta-parameters steering the structure and training procedure of the BDT. This optimization is performed using a multi-dimensional grid search of possible parameter values. For each combination of parameters, a new BDT is trained using the same set of input observables as determined in the first phase of the optimization.

The parameters included in the grid scan with the considered values are

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

Each parameter has been introduced in Section 7.1 and the values used in the baseline BDT are marked using an *.

This parameter grid results in 576 BDTs, which are subsequently filtered and ordered equivalently to the ordering in each iteration step in the first phase of the optimization. The highest-ranking BDT resulted from the parameter combination: **MaxDepth** = 5, **nCuts** = 20, **MinNodeSize** = 2 %, **Shrinkage** = 0.02, **NTrees** = 400. The performance of this BDT with optimized meta-parameters is listed in Table 7.4 below and the distributions of the **BDTscore** are shown in Figure 7.8.

When grouping all BDTs according to the value in an individual parameter, several aspects of the parameters can be studied.

Changes in the fractions of BDTs that are classified as overtrained in each such group agree with the expectations. The fractions increase, when increasing any of the parameters **MaxDepth** (from 45/192 to 125/192), **NTrees** (from 9/96 to 67/96), or **Shrinkage** (from 10/144 to 116/144). When grouping the BDTs according to **MinNodeSize**, the fractions decrease from 71/144 to 48/144 for increasing values. For the **nCuts** the fractions are consistent for both considered values.

The maximum values of the metric f_{Ztot} for all BDTs with a common value are shown in Figure 7.9 for the parameters with at least three considered values: **MinNodeSize**, **MaxDepth**, **NTrees**, and **Shrinkage**. In other words, these are the maximum performances when only fixing the shown parameter and optimizing all other parameters. These maxima are compared to the f_{Ztot} metrics when only changing the shown parameter as compared to the BDT with the best overall performance.

When varying one of the parameters **Shrinkage** and **NTrees** from the value in the best fit, it can be seen that the performance falls off rapidly when fixing all other parameters. However, when the other parameters are optimized again, the performance can almost wholly be restored. For **MaxDepth**, the optimized performances increase mildly when increasing the parameter value. Further tests were performed with **MaxDepth** = 6, but no further performance increase was found.

7.3.3 Combination of optimizations

With the optimized set of meta-parameters, the iteration for the inclusion of input observables is started again. In this final optimization phase, the set of tested configurations is increased to include also all possible configurations where two additional inputs are added. Since the

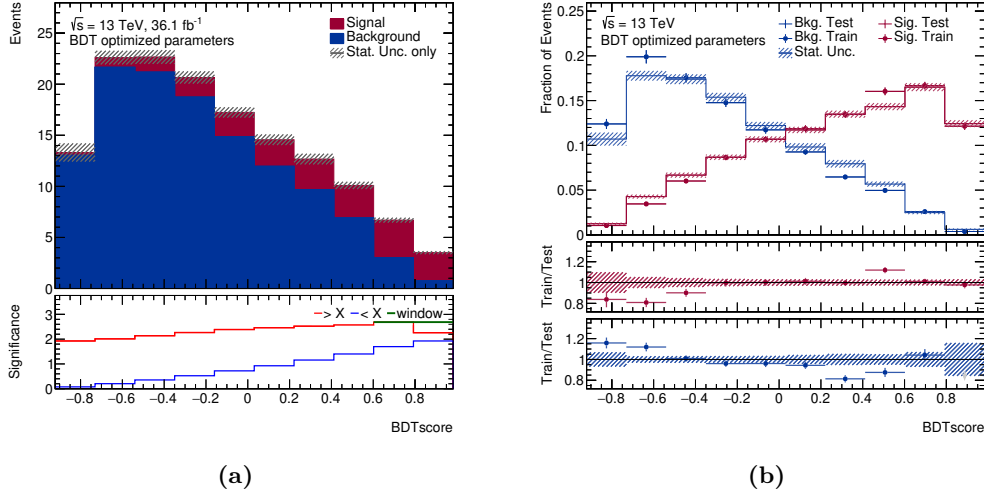


Figure 7.8: Comparison of distributions differentially in the **BDTscore** of the BDT with an optimized set of input observables and meta-parameters for signal (red) and background (blue) simulations in the signal region. The left plot shows the stacked differential event yields as estimated from the test datasets with an inset indicating the estimated statistical significance for different cut scenarios. The right plot shows predicted distributions normalized to unit area for the training dataset (full circle) and the test dataset (hatched band) with two insets showing the ratio of training over test fractions in each bin for the background and signal datasets respectively. The indicated uncertainties represent only the statistical uncertainties due to the limited amount of events in the simulations.

BDT is designed to employ also multi-variate information, combinations of input observables can be symbiotic. So the inclusion of a combination of two observables can be more beneficial to the overall performance than the inclusion of each of the individual observables would indicate.

After four additional iterations, no further performance increase without clear indications of overtraining were found. Six additional observables were added as input: $p_T(\ell_W)$, $Y(W)$, $M_T(W)$, centrality($\ell\ell$), $Y(\ell_W)$, and $\min \Delta R(j_1, \ell)$. The full set of input observables is listed in Table 7.3.

With this final set of optimizations, the set of input parameters and the configuration of meta-parameters is fixed. In the switch to the final setup, two other adjustments to the setup are applied. During the optimization, the signal dataset was split in test and training datasets of equal sizes. Since the SHERPA-based sample used for the training is not used further, the splitting is adjusted to increase the fraction of events in the training dataset to 90 % of the overall signal sample. This adjusted splitting reduces the probability of overtraining in the signal sample, which is expected to increase stability.

In order to avoid issues with overtraining of the background sample, a cross-training procedure is applied. Two independent BDTs are trained with inverted splitting in test and training samples so that each event is part of the test sample for precisely one of the BDTs. In the final analysis, each BDT is applied only to its test dataset and the resulting **BDTscore** distributions of each BDT are added. In this way, no BDT is applied to any event that was used during its training. The datasets are split according to whether each event number is even or odd. This definition allows for a reproducible, statistically independent, and unbiased splitting into two roughly equal fractions. A pair of cross-trained BDTs is trained using samples according to these splittings.

The list of input observables of the BDTs ranked according to the importance as provided by the TMVA package are shown in Table 7.3. Also, the change in the performance is evaluated when discarding each observable individually from the BDT using the evaluation metrics used before.

It can be seen that the performance evaluated using the $f_{Z_{\text{tot}}}$ metric is always reduced

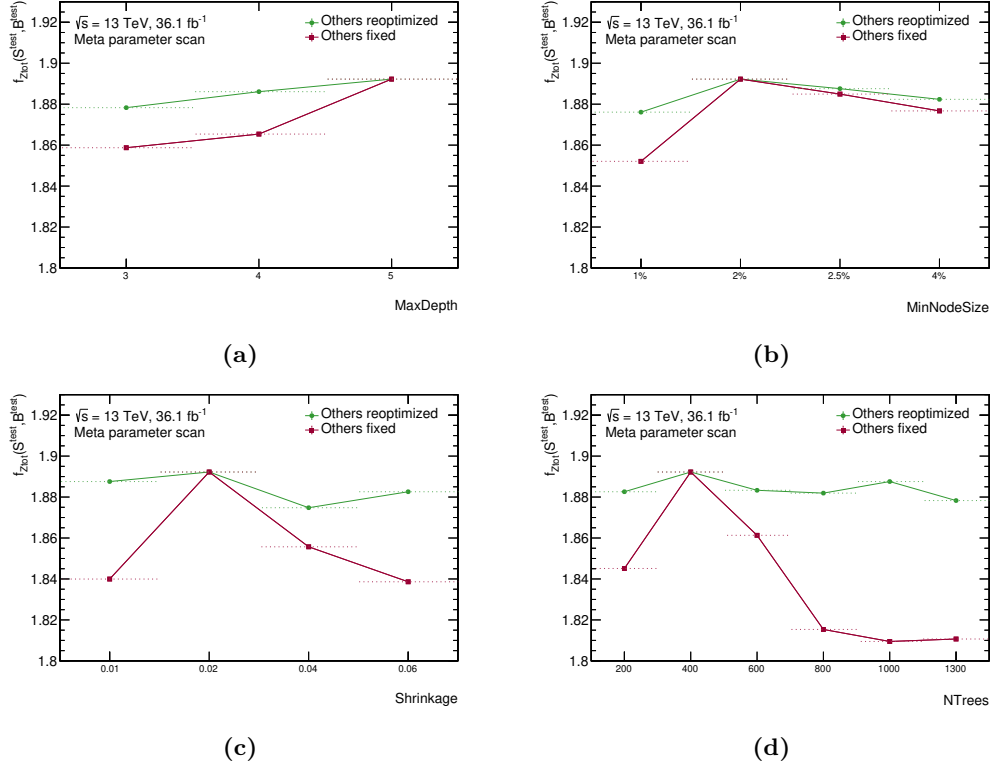


Figure 7.9: Signal vs background classification performances evaluated on the test dataset using the $f_{Z_{\text{tot}}}$ metric for scans of different meta-parameters of the BDT structure and training. The names of the considered parameters in each figure are indicated on the x axis. The maximum performances for any BDT in the grid scan for a given value of a parameter are shown in green. The red line indicates the performance when starting from the global best fit only the shown parameter is varied.

Table 7.3: List of input observables in the final BDTs listed with to their importance in each of the cross-trained BDTs as evaluated by the **TMVA** package. The input observables are ordered by descending importance in the BDT for even event numbers. Additionally, the performance change compared to the fully optimized final BDT is evaluated for the BDT for even event numbers when the respective observable is removed from the set of input observables quantified using the evaluation metrics comparing the signal and background test samples.

Observable	Importance [%]		Performance change				
	Even	Odd	Δf_{Ztot}	Δf_{ROC}	$\Delta f_{\text{N-O}}$	Δf_{Sep}	Δf_{Zstat}
M_{jj}	9.92	9.57	-0.034	< 0.001	-0.007	-0.002	-0.020
$\Delta R(j_1, j_2)$	8.32	5.60	-0.033	> -0.001	-0.006	-0.003	-0.017
$p_{\text{T}}^{\text{bal}}(\ell\ell\nu jj)$	7.53	8.01	-0.307	-0.020	-0.020	-0.035	-0.051
$ \Delta Y(\ell_W, Z) $	7.45	7.66	-0.032	-0.004	-0.005	-0.008	-0.013
$p_{\text{T}}(j_2)$	7.39	7.05	-0.019	-0.007	-0.020	-0.013	-0.050
$\min \Delta R(j_1, \ell)$	7.22	7.45	-0.042	-0.003	-0.006	-0.007	-0.020
centrality($\ell\ell\ell$)	6.53	7.27	-0.021	> -0.001	< 0.001	-0.001	< 0.001
$p_{\text{T}}(j_1)$	6.45	7.47	-0.045	-0.003	-0.009	-0.007	-0.022
$\zeta(\ell\ell\ell)$	6.25	6.81	-0.029	-0.001	-0.001	-0.001	-0.011
$M_{\text{T}}(W)$	5.49	5.34	-0.034	-0.001	-0.004	-0.003	-0.017
$M(\ell\ell\ell)$	5.23	5.02	-0.023	-0.001	-0.005	-0.001	-0.015
$Y(\ell_W)$	5.10	6.20	-0.020	-0.001	-0.002	-0.002	-0.006
$p_{\text{T}}(\ell_W)$	4.65	4.63	-0.017	> -0.001	-0.005	-0.002	-0.013
$p_{\text{T}}(Z)$	4.37	4.15	-0.015	-0.001	-0.008	-0.003	-0.017
$p_{\text{T}}(\ell_{Z,2})$	4.16	4.44	-0.028	-0.000	-0.005	-0.002	-0.014
$Y(W)$	3.96	4.33	-0.024	< 0.001	0.001	-0.001	-0.002

when dropping an input observable. Also, the other metrics indicate reduced performance for each observable, with few exceptions. All of these exceptions are compatible with an invariant performance. The importance ranking of the **TMVA** package does not show a clear correlation with any of the evaluation metrics. While the importance evaluates the impact of an observable in a given BDT, it has no handle on whether the training procedure can recover the performance loss, when removing a commonly used observable using other observables.

The predicted event yield, as well as normalized distributions for the combined background and the signal process, are shown differentially in the **BDTscore** for the final BDT in Figure 7.10.

7.4 Evaluation at Truth-level

In addition to the nominal evaluation of the BDT at detector-level, an evaluation at truth-level is helpful for further studies. Such evaluation is achieved by creating copies of the BDT configuration files and manually replacing the observables by the respective observables at truth-level. A comparison of all input distributions at truth level are shown in Section F.2 in the appendix.

The predicted cross-sections in the fiducial phase space differentially in the **BDTscore** of this truth-level BDT are studied in Figure 7.11

7.5 Conclusions

For comparison of the baseline BDT and the optimized BDT using the optimized set of input observables after pre-processing and the optimized configuration of meta-parameters, differential distributions for both BDTs are shown in Figure 7.10.

While the distributions of the different BDTs are quite similar at first glance, especially in extreme regions with $|\mathbf{BDTscore}| \approx 1$ there are apparent differences to be seen. A higher fraction of events in the range with the highest **BDTscore** is predicted for the baseline BDT as compared to the final BDT. It can also be seen that the baseline BDT shows more indications of overtraining for the signal as well as the background processes. A significant difference is also to be seen in the significance estimations for different cut scenarios (see Appendix C for a detailed description of the significance estimations). For the baseline BDT, the maximum significance of below 2.6 is achieved, while for the final BDT a maximum significance of more than 2.7 is achieved.

The classification performance of the final BDT, as well as each step during the optimization quantified using the evaluation metrics, are listed in Table 7.4. The two cross-trained final BDTs perform better than the original baseline BDT in each of the considered evaluation metrics.

The main increases in performance during the optimization were achieved while adding new input observables. In each iteration, the performance on the test sample could be increased until a plateau was reached after ten input observables. The effect of the meta-parameter optimization seems comparably small, but allowed a further inclusion of observables with the limited effect of overtraining, thus allowing for further increases due to the inclusion of additional input observables.

The applied transformation of the input observables also improved the overall performance and reduced overtraining. The overall effect can, however, only be determined by re-performing the optimization following the same procedure without applying the transformations. This test was not done due to the time- and computing-intensity of the optimization process.

The adjustments to the sample splitting after the actual optimization is beneficial as well. While the $f_{Z_{tot}}$ metric indicates a reduced performance, all other metrics are improved. Also, the effects of overtraining could be reduced. Any bias due to overtraining is avoided since no BDT is applied to an event used during its training process.

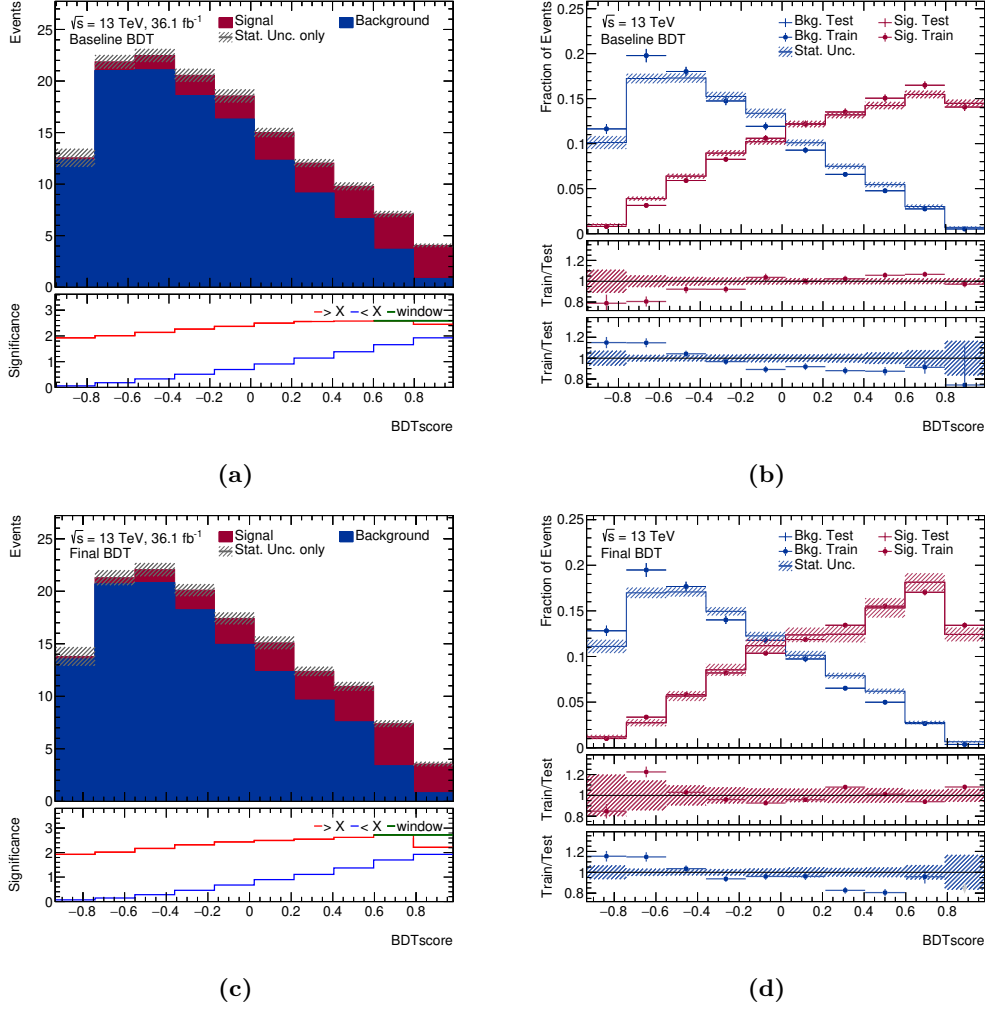


Figure 7.10: Comparison of distributions differentially in the BDTscore of the baseline BDT (top) compared to the optimized final BDT (bottom) for signal (red) and background (blue) simulations in the signal region. The left plot shows the stacked differential event yields as estimated from the test datasets with an inset indicating the estimated statistical significance for different cut scenarios. The right plot shows predicted distributions normalized to unit area for the training dataset (full circle) and the test dataset (hatched band) with two insets showing the ratio of training over test fractions in each bin for the background and signal datasets respectively. The indicated uncertainties represent only the statistical uncertainties due to the limited amount of events in the simulations.

Table 7.4: List of steps during the optimization of the BDT from the baseline BDT towards the final, cross-trained BDTs listed with different performance evaluation metrics. The overtraining is quantified by applying the $f_{\text{red.}} \chi^2$ metric to compare test and training samples for signal and background, respectively. The performance on the training datasets is evaluated using the f_{Ztot} metric comparing the signal and background predictions in the training sample. For the performance on the test samples, also the metrics f_{ROC} , $f_{\text{N-O}}$, f_{Sep} , and f_{Zstat} as introduced in Section 7.1.1 are used.

Optimization step	Overtraining		Training		Performance			
	$f_{\text{red.}} \chi^2(S)$	testing vs. training	f_{Ztot}	f_{Ztot}	f_{ROC}	$f_{\text{N-O}}$	f_{Sep}	f_{Zstat}
Baseline BDT	1.51	1.33	1.956	1.842	0.777	0.429	0.252	2.226
all considered inputs	1.47	1.71	2.004	1.820	0.778	0.426	0.253	2.218
apply transformations	1.25	1.61	2.017	1.828	0.779	0.430	0.254	2.229
adding M_{jj}	1.47	1.05	1.197	1.160	0.657	0.259	0.099	1.880
adding $p_{\text{T}}^{\text{bal}}(\ell\ell\nu jj)$	1.11	1.18	1.419	1.492	0.687	0.296	0.134	1.909
adding $ \Delta Y(\ell_W, Z) $	1.14	0.95	1.566	1.605	0.708	0.335	0.159	1.972
setting NTrees = 200	0.99	1.04	1.582	1.616	0.707	0.331	0.158	1.963
adding $\zeta(\ell\ell)$	1.08	1.41	1.824	1.747	0.748	0.382	0.208	2.102
adding $p_{\text{T}}(j_1)$	1.09	0.87	1.837	1.826	0.759	0.396	0.226	2.146
adding $M(\ell\ell)$	1.02	1.21	1.862	1.847	0.759	0.396	0.228	2.137
adding $p_{\text{T}}(\ell_{Z,2})$	1.43	1.02	1.865	1.856	0.760	0.401	0.228	2.149
adding $p_{\text{T}}(j_2)$	1.48	1.43	1.902	1.854	0.771	0.417	0.258	2.193
adding $\Delta R(j_1, j_2)$	1.59	1.01	1.921	1.880	0.773	0.423	0.247	2.215
adding $p_{\text{T}}(Z)$	1.55	1.17	1.939	1.883	0.773	0.424	0.246	2.221
Meta-parameter grid search	1.53	1.17	1.936	1.892	0.775	0.421	0.249	2.214
adding $p_{\text{T}}(\ell_W)$ and $Y(W)$	1.45	1.46	1.963	1.907	0.775	0.423	0.248	2.208
adding $M_{\text{T}}(W)$	1.74	1.32	1.965	1.913	0.776	0.422	0.251	2.220
adding centrality($\ell\ell$)	1.47	1.47	1.967	1.919	0.776	0.420	0.251	2.212
adding $Y(\ell_W)$ and $\min \Delta R(j_1, \ell)$	1.47	1.55	2.001	1.924	0.780	0.427	0.256	2.232
Final BDT Even	1.01	1.41	1.987	1.916	0.788	0.443	0.273	2.264
Final BDT Odd	0.97	1.47	2.067	1.877	0.792	0.451	0.278	2.285

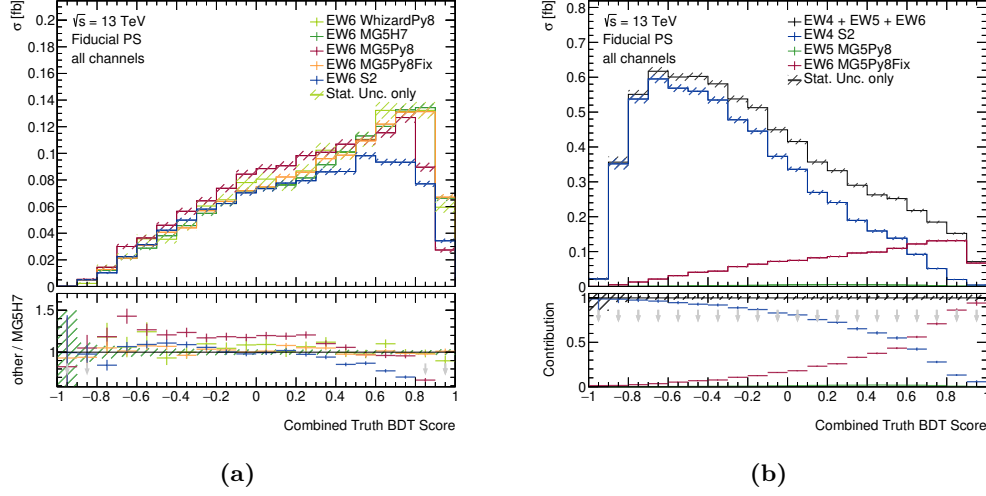


Figure 7.11: Comparison of distributions differentially in the truth-level **BDTscore** of the combined BDT in the fiducial phase space. The left plot shows the differential cross-sections for different simulations of the $WZjj$ -EW6 process with different colours, as indicated in the legend. The inset shows the ratio to the nominal simulation based on MG5_AMC and HERWIG 7. The nominal simulations for the $WZjj$ -EW4 background (blue), $WZjj$ -EW5 interference (green) and $WZjj$ -EW6 signal (red) processes are compared to the sum of the three processes (black). The inset shows the ratio to the sum of the processes, i.e. the overall contribution. The indicated uncertainties represent only the statistical uncertainties due to the limited amount of events in the simulations.

A downside of the optimized BDT is the slight increase in the number of input observables from 15 to 16. The modelling of each of the input observables needs to be validated to avoid biases between simulation and data. For this, systematic uncertainties on the predictions need to be evaluated. These uncertainties are discussed in detail in the following chapter, and a comparison of predicted distributions to data in the signal and WZ control regions for the **BDTscore**, as well as the input observables, is presented in Appendix F.3.

Chapter 8

Uncertainties

The final ingredient for quantitative comparisons of simulated and measured data is a realistic and reliable estimation of systematic uncertainties. Depending on the underlying origin of the uncertainty, systematic uncertainties are categorised, and different methods are employed to estimate uncertainties.

The main categorisation is the separation of experimental and modelling uncertainties. Experimental uncertainties arise from the simulation of the detector and the reconstruction and consequently do not depend on a specific process. Modelling uncertainties, also referred to as theoretical uncertainties, are specific to a dataset of simulated events and are added to ensure the simulation covers the underlying physics of the process of interest.

The basic principle of uncertainty estimation is to propagate changes to cover the effects of assumptions or approximations to the final analysis results. An example is a change in the predicted event yield in case one assumes a slightly different calibration of electrons since a different subset of events might be selected. These changes are evaluated for many different sources of uncertainties and are combined to yield the total systematic uncertainty.

For most uncertainties, there are up and down variations, which reflect the effects of an underlying physics origin. In the previous example of the calibration of electrons, the nominal calibration is measured with uncertainty and can be too high or too low. Two variations are necessary to cover the full effects of the underlying origin of the uncertainty. The variations are designed to cover the 1σ (68% CL) confidence interval of the underlying parameter.

Many uncertainties are estimated using variations, where the datasets are processed multiple times with changes, e.g. in the reconstruction algorithms. Some variations result in different subsets of selected events and differences in observables compared to the nominal dataset. For these variations new, independent MiniTrees are written (see Section 4.4.2).

Some systematics do not change properties of objects but only reweight the events. Such variations by definition can not result in different subsets of events, and for each event, the values of observables are unchanged. However, due to the changes in the weights of different events, the final predictions can vary in shape and normalisation. Technically, these variations are implemented by added a new entry storing the correction factor on the event weight in the existing MiniTrees.

The combination of systematic uncertainties is not trivial and ideally should only be performed specifically to the event yield or histogram of interest. In case systematic variations are correlated, these correlations have to be taken into account as well. In order to avoid this complexity, most variations are defined to be uncorrelated. This can be achieved by defining the variations in an uncorrelated way in the first place or by transforming the variations into an uncorrelated basis. For uncorrelated uncertainties $\Delta_1 N$ and $\Delta_2 N$, the changes compared to the nominal analysis result can be added in quadrature:

$$\Delta_{\text{tot}} N = \sqrt{(\Delta_1 N)^2 + (\Delta_2 N)^2}. \quad (8.1)$$

Unless stated otherwise, systematic uncertainties are assumed to be uncorrelated and added in quadrature.

8.1 Experimental Uncertainties

Since experimental uncertainties are not process-specific, procedures are derived centrally and applied to all simulations. Typically, the responsible combined performance groups derive and recommend at least one set of variations to cover potential uncertainties. These variations are applied to each simulated dataset individually, and the changes are propagated further.

For convenience, the individual uncertainties are grouped in tables as described in the following:

EG. This group includes systematic uncertainties due to the imperfect electron calibration. The underlying origins covered in this group are mismodelling of the electron energy resolution (`EG_RESOLUTION_ALL`) and the energy scale (`EG_SCALE_ALL`), with up and down variations. This is a reduced set of variations obtained by conservatively combining the full set of 18 variations for the electron energy resolution and 134 for the electron energy scale.

EL_EFF. This group covers systematic variations added to cover the effects of electron efficiency measurements. Two variations are included to cover the effects of imprecise measurement of the reconstruction (`EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR`), identification (`EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR`), isolation (`EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR`), and trigger (`EL_EFF_TriggerEff_TOTAL_1NPCOR_PLUS_UNCOR`) efficiencies and the trigger scale factor (`EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR`).

MUON. The group of systematic uncertainties in the muon measurements includes more variations even in the smallest available set.

Mismodellings in the muon calibration are included for the smearing of ID tracks (`MUON_ID`), smearing of MS tracks (`MUON_MS`), the momentum scale (`MUON_SCALE`), and a charge-dependent scale correction (`MUON_SAGITTA_RHO`) as well as the remaining charge-dependency after correction (`MUON_SAGITTA_RESBIAS`) with two variations each. Uncertainties in the isolation measurement are covered with two variations each for statistical effects (`MUON_ISO_STAT`) and systematic effects (`MUON_ISO_SYS`) in the measurement. Uncertainties in the measurement of the correction factors for the track-to-vertex-association are included as well with two variations each for the statistical (`MUON_TTVA_STAT`) and systematic (`MUON_TTVA_SYS`) components.

MUON_EFF. The uncertainties on the muon efficiencies include the per-object correction factors (`MUON_EFF_STAT/MUON_EFF_SYS`), as well as the contributions for muons with low transverse momentum (`MUON_EFF_STAT_LOWPT/MUON_EFF_SYS_LOWPT`) and the trigger efficiencies (`MUON_EFF_TrigStatUncertainty/MUON_EFF_TrigSystUncertainty`). For each of these effects, the statistical and systematic components of the uncertainties are separated and covered by two variations each.

MET. The uncertainties on the missing transverse momentum include uncertainties on the resolution (`MET_SoftTrk_Reso`) and scale (`MET_SoftTrk_Scale`) of the track-based soft-term contributions. Uncertainties on the hard-terms are covered by variations in the electron, muon, and jet uncertainties.

JET. This set of systematic variations covers the systematic mismodelling of effects of the jet energy scale as well as on the vertex tagging. The jet energy scale is determined using balance-based in-situ measurements (see Section 4.3.3) resulting in 75 fitted parameters with uncertainties. In order to reduce complexity these parameters have been combined to 16 effective parameters according to the central recommendations. These 16 parameters are grouped according to their source: 7 for statistical parameters (`JET_EffectiveNP_Statistical`), 4 for modelling parameters (`JET_EffectiveNP_Modelling`), 2 for detector-related parameters (`JET_EffectiveNP_Detector`), and 3 for mixed parameters (`JET_EffectiveNP_Mixed`). Additional parameters on the eta-intercalibration (`3x JET_EtaIntercalibration` for modelling,

non-closure, and statistical uncertainties), high-momentum jets (`JET_SingleParticle_HighPt`), as well as the pileup dependence (4x `JET_Pileup` for the offset in μ , the offset in the number of primary vertices, the p_T -term, and the ρ topology) are introduced. For each of these parameters, two variations are performed to include up- and down-variations.

Additional variations are added to include the effect of mismodelling of the flavour composition (`JET_Flavor_Composition`) and response (`JET_Flavor_Response`) of a sample, as well as the treatment of b -jets in general (`JET_BJES_Response`), and punch-through jets (`JET_PunchThrough_MC15`). For each of these effects two variations are included.

Mismodelling of the efficiency of the jet vertex tagging requirement is evaluated event-per-event with two additional variations (`JET_JvtEfficiency`).

JER. The uncertainty of the jet energy resolution is treated as a separate group in order to allow for individual discussions of this source, which has been observed before to be large for vector boson scattering topologies. One single variation (`JET_JER_SINGLE_NP`) is provided which needs to be symmetrised.

BTAG. For the mid-modelling of the uncertainties associated with the tagging of b -flavoured jets, a simplified model with a reduced set of five parameters in total is used. The five parameters combine the efficiency uncertainties per flavour group (bottom `FT_EFF_B_systematics`, charm `FT_EFF_C_systematics`, and light `FT_EFF_Light_systematics`), as well as two uncertainties on the extrapolation (`FT_EFF_extrapolation` and `FT_EFF_extrapolation_from_charm`). Two variations per parameter are included.

PileUp. This group covers the uncertainties on the process of reweighting the distribution of primary vertices from simulations to match data as described in Section 5.1. One parameter (`PRW_DATASF`) with two variations is included.

Lumi. The luminosity provides the global scaling of simulations to measured data and an uncertainty of 2.1% is included as a variation on the global normalisation for each simulation.

The relative effects of each of the mentioned groups of systematic uncertainties on each of the individual input processes considered in the analysis in the signal region are listed in Table 8.1. The leading uncertainties, in general, are those associated with jets. Depending on the considered process, the jet energy measurement, flavour tagging, or modelling of pileup activity are the leading experimental systematic. The jet energy measurement has an average effect of about 5% for most processes. The flavour tagging uncertainty is smaller for processes without contributions of top quarks, while for tZj and $t\bar{t} + V$ the flavour tagging uncertainty is the dominant uncertainty at about 20%. For many other processes, the uncertainty associated with the pileup reweighting is between 5 and 10% and thus of similar importance as the jet energy measurement.

8.2 Modelling Uncertainties

Modelling uncertainties depend on the considered process, and different approaches are followed to estimate the uncertainties. The modelling uncertainties for the data-driven non-prompt background estimation are introduced, discussed and validated in Section 6.2. For prompt background processes, only a conservative normalization uncertainty is added to cover mismodelling effects in order to reduce complexity. Only when determining the input uncertainties as used in the final fitting procedure, the modelling uncertainties of the $WZjj$ -EW4 and $WZjj$ -EW6 processes are discussed and evaluated in detail.

The normalization uncertainties are estimated on a process-by-process basis. For the main prompt background, a 10% normalization uncertainty is estimated based on the disagreement between data and predictions in the dedicated control region (see Section 6.1).

Table 8.1: Relative effect of experimental uncertainty groups on predicted yield in the signal region per process in percent. The labels for uncertainty groups follow the nomenclature used in Section 8.1. The total experimental uncertainty is given in the lowest line. Each column corresponds to a process treated as a separate process in the final fit. The group MC stat corresponds to the statistical uncertainty due to the limited amount of simulated events. For the data-driven non-prompt estimate, this includes the statistical uncertainty of observed data events in the control region where the fake factors are applied.

	$WZjj$ -EW4	$ZZjj$ -EW4	tZj	$t\bar{t} + V$	VVV	$ZZjj$ -EW6	non-prompt	$WZjj$ -EW6
BTAG	+4.79 -0.00	+3.72 -0.00	+17.30 -2.52	+22.82 -4.10	+4.56 -0.00	+4.55 -0.00	+0.00 -3.32	+2.96 -0.00
EG	+0.16 -0.23	+0.90 -0.00	+0.27 -0.16	+1.52 -0.00	+0.00 -0.84	+0.97 -0.00	+1.57 -0.87	+0.24 -0.67
EL_EFF	+1.38 -1.51	+2.35 -1.22	+1.81 -0.99	+2.32 -0.75	+1.28 -2.18	+2.14 -1.05	+2.91 -0.91	+1.65 -1.36
FAKE	+0.00 -0.00	+0.00 -0.00	+0.00 -0.00	+0.00 -0.00	+0.00 -0.00	+0.00 -0.00	+60.53 -35.31	+0.00 -0.00
JER	+0.42 -0.00	+6.45 -0.00	+1.36 -0.00	+0.00 -1.79	+0.00 -1.00	+0.25 -0.00	+0.00 -2.13	+0.00 -2.10
JET	+5.38 -5.50	+9.37 -6.78	+4.99 -3.53	+6.13 -2.95	+4.75 -7.05	+2.06 -2.17	+6.93 -4.11	+0.65 -3.89
Lumi	+2.10 -2.10	+2.10 -2.10	+2.10 -2.10	+2.10 -2.10	+2.10 -2.10	+2.10 -2.10	+1.54 -1.54	+2.10 -2.10
MET	+0.00 -0.11	+1.04 -1.10	+0.97 -0.20	+0.91 -0.00	+0.00 -0.78	+3.54 -1.21	+1.49 -0.00	+0.00 -0.52
MUON	+0.38 -0.92	+1.44 -0.06	+1.73 -0.08	+1.71 -1.51	+0.28 -2.57	+1.71 -0.00	+6.93 -2.55	+0.71 -0.46
MUON_EFF	+0.45 -0.63	+1.40 -0.00	+1.21 -0.19	+1.76 -0.00	+0.09 -1.20	+1.38 -0.06	+2.25 -0.00	+0.75 -0.44
PileUp	+6.05 -2.17	+5.47 -0.31	+8.91 -0.56	+9.75 -0.00	+12.43 -0.00	+2.30 -4.07	+5.98 -0.00	+4.95 -3.46
MC stat	+1.02 -1.02	+2.73 -2.73	+2.91 -2.91	+5.17 -5.17	+6.75 -6.75	+2.45 -2.45	+21.88 -21.88	+2.51 -2.51
Total Uncertainty	+9.81 -6.64	+14.02 -7.79	+20.66 -5.75	+26.44 -7.92	+15.80 -10.71	+7.97 -5.85	+65.54 -42.06	+6.95 -6.72

For the ZZ background, two different normalization uncertainties are applied for the $ZZjj$ -EW4 and $ZZjj$ -EW6 contributions. Due to the small contribution of the $ZZjj$ -EW6 process, no direct comparison to data is applied, but the uncertainty is estimated conservatively based on the dedicated analysis of this process in [40]. There the difference between the predicted and observed cross-sections is about 35 %, so here a more conservative normalization uncertainty of 40 % is applied. For the $ZZjj$ -EW4 the normalization uncertainty is estimated to be 10 % based on comparisons in the dedicated ZZ control region.

For each of the other considered prompt backgrounds, tZj , $t\bar{t} + V$, VVV , a modelling uncertainty of 30 % is estimated. Uncertainties of roughly this size are often observed for leading-order scale variations, motivating this choice.

These normalization uncertainties are mostly used for plotting purposes and qualitative statements. Comparisons of the predicted and observed event yields differentially in the input observables used in the BDT, using these coarse modelling uncertainties are shown in Section F.3 in the appendix for the signal region as well as the WZ control region. For the final cross-section measurement using a fitting approach, the uncertainties are estimated more conservatively or in greater detail.

For a detailed comparison of the modelling uncertainties, different sources of modelling uncertainties are considered. For some known sources, good commonly-used approaches exist to evaluate the related uncertainties, for others, no estimates are available.[166] Since the application of the particular modelling uncertainties is limited to the inputs for the fitting procedure, the methods are introduced here using the same bin ranges for the combined **BDTscore** as used in the fit in the signal region. The choice of this set of bin ranges is discussed in Section 9.2.3. The methods are also applied to derive uncertainties for input distributions in control regions where applicable. The effects of these uncertainties in the control regions as derived for the $WZjj$ -EW4 and $WZjj$ -EW6 processes are collected in Appendix G.1

8.2.1 PDF uncertainties

The parton distribution functions are a necessary input for simulations and are obtained from fits to earlier measurements. From these fits, variations of the fit parameters after decorrelation are provided, which can be propagated to estimate the effect of variations within the individual fits. These sets of variations are referred to as eigen-sets. Different PDF sets exist, employing different fitting approaches.

The recommended way to evaluate modelling uncertainties due to the PDF is to use the full set of variations of the **NNPDF30_nn1o** PDF set[167] as a baseline. For the **NNPDF30_nn1o** PDF set the variations do not represent variations of fit parameters, but different replicas of neuronal nets. The uncertainty on an event yield N is evaluated based on the variance of the predictions of the different replicas following [167]

$$\sigma_{\text{NNPDF}}^2 = \frac{1}{100} \sum_{i=1}^{100} (N_{\text{NNPDF}}^k - N_{\text{NNPDF}}^{\text{Nominal}})^2, \quad (8.2)$$

where N_{NNPDF}^k are the predictions of replica k and $N_{\text{NNPDF}}^{\text{Nominal}}$ is the prediction using the nominal **NNPDF30_nn1o** PDF set.

Additionally, the nominal predictions for two alternative PDF sets are compared to the nominal **NNPDF3.0**-based prediction. The maximum deviation from the nominal in each direction of the variations of the alternative PDF sets and the internal variations of the **NNPDF3.0** is used as total PDF uncertainty. This treatment ensures that the uncertainty covers both types of variations. The alternative PDF sets used in this work are **CT14nn1o** and **MMHT2014nn1o**.

In the available simulations, special care was taken to propagate the variations from the event generation to the final MiniTrees. With these variations, a full evaluation of the dependency on the PDF is possible rather than a more straightforward reweighting approach.

The resulting uncertainties due to the choice of PDF are visualized for the $WZjj$ -EW4 and $WZjj$ -EW6 processes in Figure 8.1. The uncertainties are approximately 2 % for both

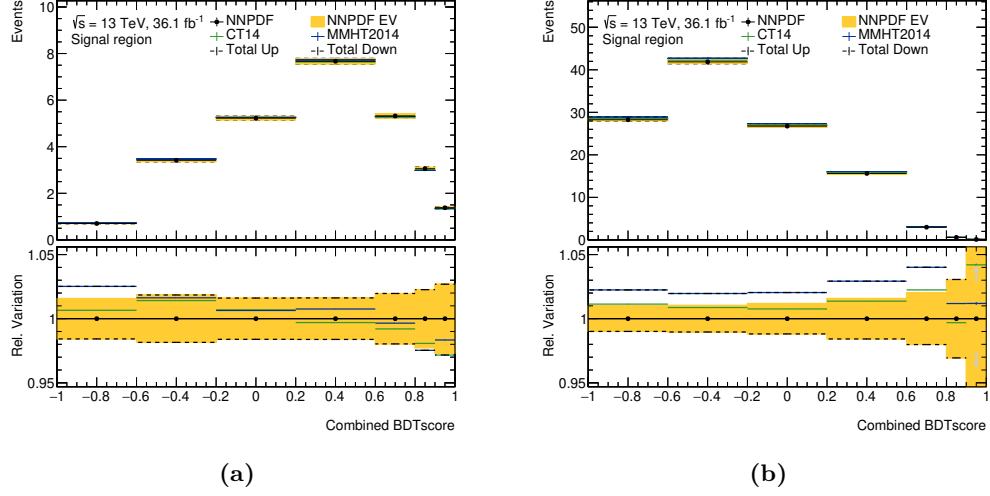


Figure 8.1: Uncertainties due to mismodelling of the PDF on the predicted event yields differentially in the **BDTscore** for the signal region. Figure (a) shows the PDF uncertainties for the $WZjj$ -EW6 signal process and Figure (b) for the $WZjj$ -EW4 background process. The nominal prediction (black) is shown with the uncertainty derived from the NNPDF3.0 eigen-set (yellow band). For comparison the predicted yields using the nominal MMHT2014 $_{\text{nnlo}}$ (blue) and CT14 (green) PDF sets are shown. The total PDF uncertainty derived from these comparisons is shown using black, dashed lines. The lower inlay shows the ratio to the nominal prediction in each bin.

processes except for very high values of the **BDTscore**, where the uncertainties increase. Especially for the $WZjj$ -EW4 background process, the uncertainty increases drastically to about 30 %.

For $WZjj$ -EW6 the eigen-set variations of the NNPDF3.0 PDF cover the variations to other PDF sets in most kinematic regions with few exceptions. In cases where the variation to a different PDF set is not covered by the symmetric uncertainty band estimated from the eigen-set, the uncertainty is increased in an asymmetric way to cover the alternative predictions as well. For the $WZjj$ -EW4 process, the MMHT2014 $_{\text{nnlo}}$ prediction is consistently larger compared to the nominal NNPDF3.0-based prediction by about 3 %, which is outside of the uncertainty band estimated using the NNPDF3.0 eigen-set.

8.2.2 Strong coupling

The dependencies of the parton distribution function on the strong coupling α_S are evaluated by including variations of the PDF sets using values for α_S evaluated at the Z boson mass which are varied by 0.001 around the nominal value of 0.118. The resulting uncertainties for the signal region are shown processes in Figure 8.2.

For the $WZjj$ -EW6 signal process the α_S uncertainties are negligible and well below 1 %. For $WZjj$ -EW4, the effect is similar to the effect of the overall PDF uncertainties. The relative uncertainties are about 2 % in most kinematic regions and up to about 17 % at very high values of the **BDTscore**. The drastic difference in effect between the two processes is expected since only purely electroweak diagrams are included in the hard process calculation for the $WZjj$ -EW6 process.

8.2.3 Scale uncertainties

As discussed in Section 3.1, the hard process calculation depends on the unphysical renormalization and factorization scales. Since full calculations to all orders of perturbation theory would not depend on these scales, the scales are commonly used to estimate the effect of neglecting higher-order contributions in the hard-process calculation.

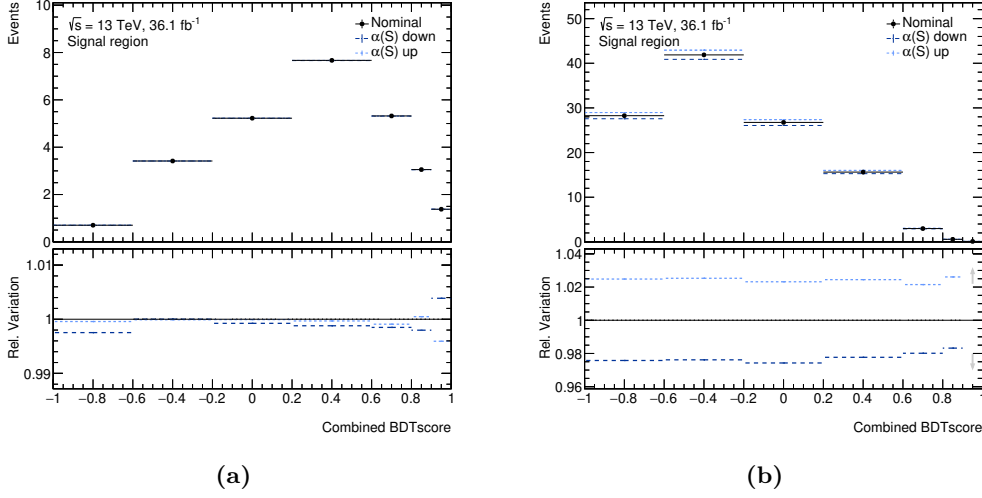


Figure 8.2: Uncertainties due to mismodelling of α_S on the predicted event yields differentially in the `BDTscore` for the signal region. Figure (a) shows the α_S uncertainties for the $WZjj$ -EW6 signal process and Figure (b) for the $WZjj$ -EW4 background process. The nominal prediction (black) is compared to the predicted yields using the variations of $\alpha_S = 0.117$ (dashed, dark blue) and $\alpha_S = 0.119$ (dotted, light blue). The lower inlay shows the ratio to the nominal prediction in each bin.

The common recommendation is to vary both scales independent of each other by a factor of 2 up and down. These variations result in eight combinations with variations from the nominal setting, from which the two combinations where the two scales are varied in opposing directions are neglected. The maximum deviation of any of the six remaining combinations from the nominal prediction is used as scale uncertainty.

The effect of these scale variations on the predicted differential distributions for the $WZjj$ -EW4 and $WZjj$ -EW6 processes are shown in Figure 8.3. It can be seen that the effect of the renormalisation scale on the $WZjj$ -EW6 process is negligible. This is plausible since the renormalisation scale mostly enters via the running of the α_S coupling. Since this coupling doesn't enter in the considered diagrams for this process, no significant effect is expected. The effect of the factorisation scale is between 5 and 10 % in the signal region.

For the $WZjj$ -EW4 process, the effect of the factorisation scale is only a few per cent and minor compared to the effect of the renormalisation scale. The renormalisation scale variations mostly vary the overall event yield by between 20 and 30 %, while the effect is similar for most values of the `BDTscore`. Only at high `BDTscore` values around 1, the relative effect is reduced compared to other values.

8.2.4 Parton shower and hadronisation

During the simulation of the parton shower and hadronisation effects, different models exist to model non-perturbative effects. While some generators include variations of internal parameters, the most common recommendation is to compare two different shower algorithms and implementations whenever possible.

For the $WZjj$ -EW6 process, both approaches are followed and included. The nominal simulation employs the HERWIG 7 generator for the simulation of the parton shower effects. In the available version, HERWIG 7 is able to include on-the-fly weights for shower-internal scale variations. A systematic uncertainty is derived by varying both scales simultaneously. In principle, a correlation with the scale variations of the hard-process calculations is possible and plausible. For simulations where merging was applied, such correlation is the recommended approach, while for simulations without merging, there is no clear recommendation and both approaches are considered acceptable.[166] Since the available simulation does not employ merging, for the sake of simplicity, the uncorrelated approach is followed.

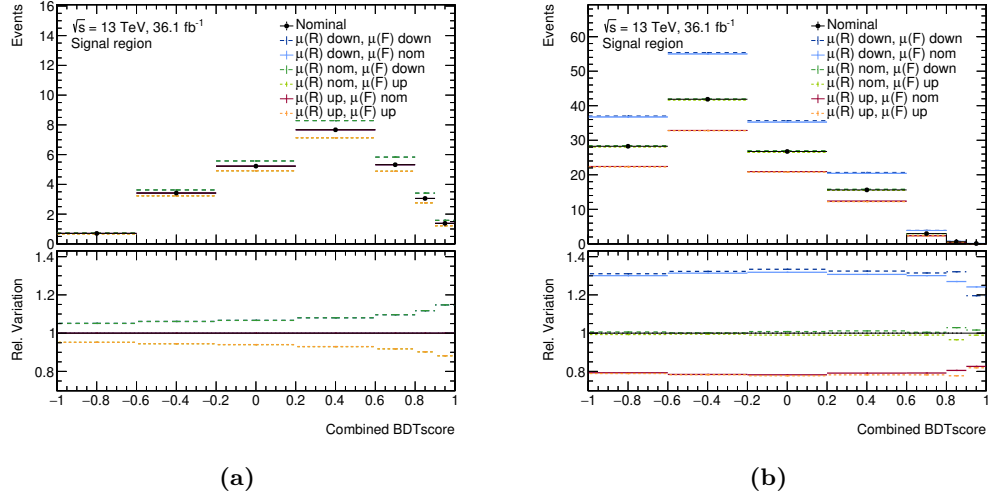


Figure 8.3: Uncertainties due to the choice of scales on the predicted event yields differentially in the BDTscore for the signal region. Figure a shows the scale uncertainties for the $WZjj\text{-EW6}$ signal process and Figure b for the $WZjj\text{-EW4}$ background process. The prediction using nominal scale choices (full line, black) is compared to the predicted yields using up and down variations of the renormalization (in legend as $\mu(R)$) and factorization scales (in legend as $\mu(F)$). The colours represent different choices of the renormalization scale, where variations with nominal $\mu(R)$ are green, and up (down) variations are blue (red). Variations with the nominal factorization scale are shown in full lines, and up (down) variations of $\mu(F)$ are indicated using a darker (lighter) line colour and dashed (dotted) lines. The lower inlay shows the ratio to the nominal prediction in each bin.

Also, an alternative simulation using the PYTHIA 8 generator for the simulation of parton shower effects was produced privately, including the simulation of detector effects. For each bin, a systematic uncertainty is derived as half of the difference between the nominal and alternative samples. The variation is symmetrised by including the same absolute variation in the opposite direction. In this definition of the variations, the difference between the up and down variations is equal to the difference between the two alternative simulations.

Both modelling uncertainties for the $WZjj\text{-EW6}$ process are shown in Figure 8.4. The variations of the renormalization and factorization scales during the shower simulations have a substantial impact on the prediction, both on the normalization and the shapes. For the comparison between the nominal and alternative shower model, no statistically significant difference was found given the limited number of events due to the private detector simulation. Nevertheless, an uncertainty is derived using these two models, yielding an effect of about 5% varying strongly with the BDTscore . When comparing the truth-level application of the BDT as shown in Figure 7.11 for the two considered samples, the resulting uncertainty conservatively covers the differences at truth level and seems to be mostly caused by statistical fluctuations between the two simulations at detector-level. Due to the good agreement at truth-level, the modelling uncertainty is not extended to include the statistical uncertainties of the two models explicitly.

For the $WZjj\text{-EW4}$ process no alternative simulation is available at detector-level. The only viable alternative to the nominal SHERPA-based simulation is the privately produced MG5_AMC sample using PYTHIA 8 for showering. A private detector simulation for this sample is not feasible due to the low overall efficiency compared to the $WZjj\text{-EW6}$ sample due to the final states with lower jet multiplicity.

Instead of a direct comparison after detector simulation, the truth-level prediction using the MG5_AMC + PYTHIA 8-based simulation for the BDTscore is folded using the information extracted from the SHERPA-based simulation. The truth-level BDTscore in the fiducial phase space is approximated using the procedure described in Section 7.5.

The folding is performed using the efficiency correction, the response matrix, and the

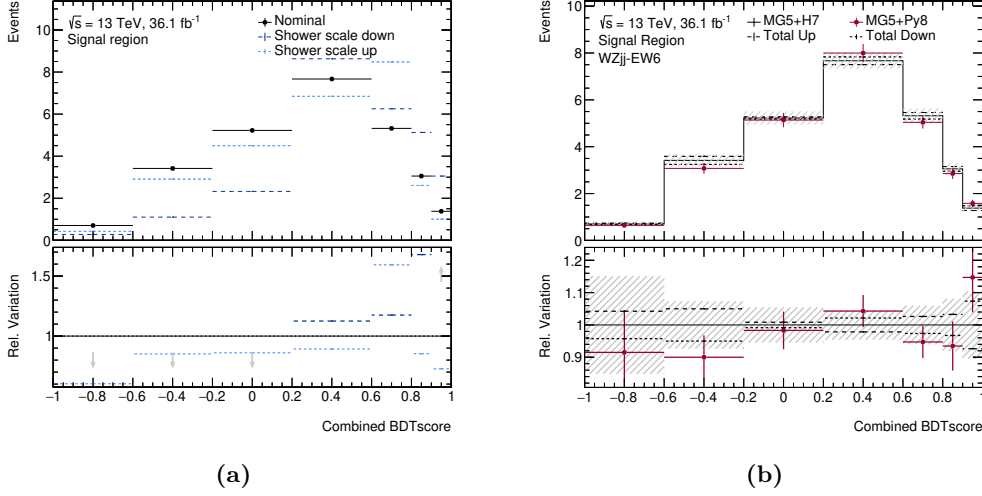


Figure 8.4: Uncertainties due to other mismodelling on the predicted event yields differentially in the BDTscore for the $WZjj$ -EW6 signal process in the signal region. Figure (a) shows the scale uncertainties in the parton shower simulation and Figure (b) additional shower mismodelling uncertainties estimated from a comparison to an alternative shower generator. The shower scale uncertainties are propagated to the fit as shown and the derived up and down variations for the shower uncertainty are indicated by black dashed and dotted lines. The hatched error band represents the statistical uncertainty on the nominal prediction due to the limited number of generated events. The lower inlay shows the ratio to the nominal prediction in each bin.

fiducial correction predicted by the SHERPA-based simulation. The efficiency correction corrects for contributions in the fiducial phase space, that do not pass the selection criteria for the signal region at detector level. The truth-level distribution constrained to events passing the criteria for the fiducial phase space and the signal region is folded using the response matrix, which represents the conditional probability for events to be in a certain bin of the detector level BDTscore , given the value of the truth-level BDTscore assigned to the event, is in the considered range. Following this definition, each row in the response matrix is normalized to 1. The response matrix as estimated from the SHERPA-based simulation is shown in Figure 8.5a. After that folding, the fiducial correction is used to correct the detector-level BDTscore distribution to include events that did not pass the selection criteria of the fiducial phase space.

The resulting distribution, when applied to the MG5_AMC +PYTHIA 8-based simulation, is normalized to the SHERPA-based prediction. The final uncertainty is derived by symmetrizing half of the difference between the folded prediction using MG5_AMC and the nominal SHERPA prediction. The final uncertainty is shown in Figure 8.5b.

The uncertainty shows a definite slope to higher event rates for higher BDTscore . For most bins, the uncertainty is around 20 %, while it reaches 100 % in the last bin. The direction of the slope is plausible when considering that the BDT was trained using the SHERPA-based simulation and kinematic differences as expected, especially for the jet kinematics can increase the BDTscore . The uncertainty is exceptionally large in the last bin partially due to its reduced range covering only a range of 0.1 in the BDTscore . High statistical uncertainties in that bin can artificially increase this uncertainty.

8.2.5 Merging

For simulations where merging procedures were applied, a typical approach to estimate uncertainties is to vary parameters for the merging algorithm, e.g. the merging scale in CKKW-(L) merging. For the $WZjj$ -EW6 simulations used in this work, no merging is applied. For the nominal simulation of the $WZjj$ -EW4 process, the MEPS@NLO merging of SHERPA was

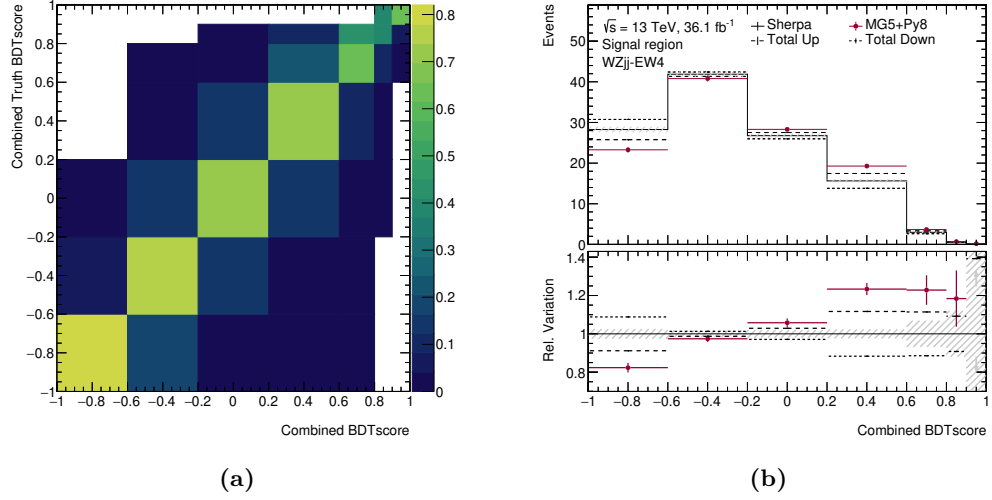


Figure 8.5: Comparisons for the modelling uncertainty on the $WZjj$ -EW4 background process. Figure (a) shows the response matrix for the $BDTscore$ distribution derived from the nominal prediction. Figure (b) shows a comparison of the folded MG5_AMC prediction (red) to the nominal SHERPA prediction (black) as well as the modelling uncertainty (black dashed or dotted lines) derived from the difference between the two predictions. The hatched error band represents the statistical uncertainty on the nominal prediction due to limited number of generated events. The lower inlay shows the ratio to the nominal prediction in each bin.

applied. Since the alternative sample simulated with MG5_AMC and PYTHIA 8 applies not only an alternative shower, but also a different merging algorithm, the modelling uncertainty derived from the difference between the two samples is assumed to cover potential mismodelling due to the merging parameters. No further uncertainty is derived.

8.2.6 Missing contributions

A very important contribution to potential mismodelling is due to missing contributions in the simulations, such as interferences between processes or diagrams of higher coupling order. For the $WZjj$ -EW6 signal process, there are three types of missing contributions, the lack of the modelling of the interference between the signal and the $WZjj$ -EW4 background process, missing contributions with higher-order in α_S , and missing contributions with higher-order in α_{EW} .

Interferences: The interference contributions, labeled as $WZjj$ -EW5, are discussed above in Section 3.7.4. The predicted yield was found to be about 5% of the $WZjj$ -EW6 process in the fiducial phase space. Since no prediction of the interference contribution is available, including detector-effects, the interference cannot be included as background in the final fit. Consequently, any contributions of the interference process may be interpreted as a part of the signal in the fit. This definition adjusts the interpretation of the final result to cover not only the $WZjj$ -EW6 process but also the $WZjj$ -EW5 interference.

In order to account for the shape effect on the predicted signal distribution, a similar procedure is followed as applied for the modelling uncertainties of the $WZjj$ -EW4 process. The truth-level distribution in the $BDTscore$ of the combined $WZjj$ -EW5 and $WZjj$ -EW6 contributions are folded using the $WZjj$ -EW6 response matrix, efficiency and fiducial corrections as predicted by the nominal MG5_AMC +HERWIG 7 simulation. From the shape differences between the combination of $WZjj$ -EW5 and $WZjj$ -EW6 to the pure $WZjj$ -EW6 prediction, an uncertainty is derived. The response matrix, as well as the final uncertainty, are shown in Figure 8.6. The uncertainty is found to be small at a level of few percents. The folded prediction for the $WZjj$ -EW5 and $WZjj$ -EW6 contributions tends to smaller values

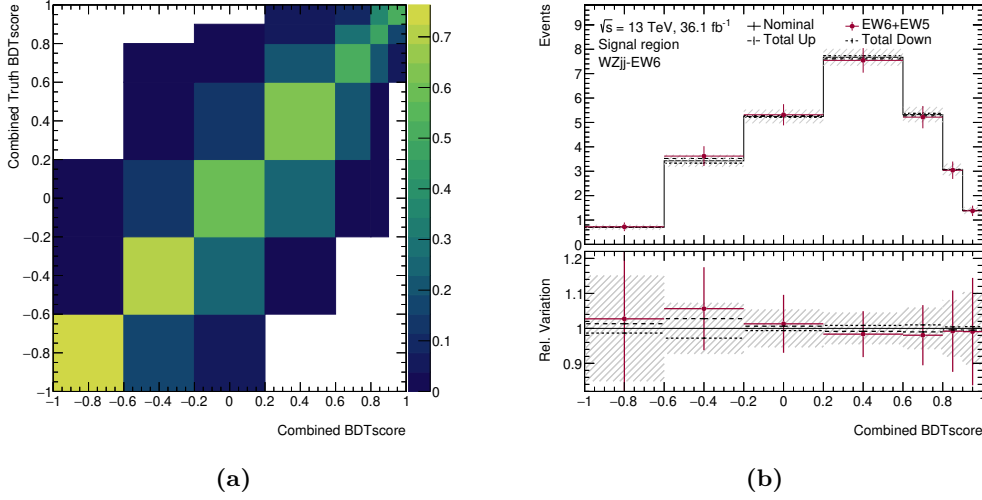


Figure 8.6: Comparisons for the modelling uncertainty due to the missing contributions of the $WZjj$ -EW5 interferences on the $WZjj$ -EW6 signal process. Figure (a) shows the response matrix for the **BDTscore** distribution derived from the nominal prediction. Figure (b) shows a comparison of the folded prediction when adding the interference contribution (red) to the nominal prediction (black) as well as the modelling uncertainty (black dashed or dotted lines) derived from the difference between the two predictions. The hatched error band represents the statistical uncertainty on the nominal prediction due to the limited number of generated events. The lower inlay shows the ratio to the nominal prediction in each bin.

of the **BDTscore**. This is plausible since the interference contribution is expected to be the largest in kinematic regions, where the $WZjj$ -EW4 and $WZjj$ -EW6 processes contribute, rather than in regions where one of the individual contributions dominates.

Since the folded distribution is scaled to the same integral as the pure $WZjj$ -EW6 prediction, it predicts smaller yields at large values of the **BDTscore**. Overall, the effect is rather small, especially compared to the systematic uncertainty introduced by the limited amount of simulated events of the nominal $WZjj$ -EW6 sample.

Higher-order terms: Other missing contributions due to missing diagrams of higher coupling orders are not feasible to estimate directly. Typically, the effect of such missing terms is estimated using scale variations (see Section 8.2.3) for terms with higher order in α_S . For the estimation of the effect of α_{EW} no such recipe exists.

For the $WZjj$ -EW4 process, the effect of terms with additional couplings of α_S is covered by the sizeable effect of the scale variations. Terms of higher order of α_{EW} do not exist as part of this process. Following the process definitions from Section 2.2.1 such terms are part of the $WZjj$ -EW5 or even $WZjj$ -EW6 processes by definition. Some of these terms, e.g. photon-induced diagrams, are not included in the current simulations as well. Once such terms are included in the simulations, clear definitions of the signal and background processes become more challenging, and it will be non-trivial to ensure the simulations are generated according to the definitions. For the strict process definition used here, the effect of higher-order terms on the $WZjj$ -EW4 process is assumed to be covered by the given uncertainties.

For the $WZjj$ -EW6 process in principle diagrams of higher-order in any of the two couplings α_S and α_{EW} ¹ exist. While dedicated numerical fixed-order calculations at NLO-EW and NLO-QCD have been performed[98], no corrections to the full process are available at particle-level. Additionally, the BDT can not be applied directly within the fixed-order calculations, so no prediction of the **BDTscore** can be determined. While the effect is expected

¹The signal process was defined as the set of diagrams with at least six couplings.

to be large, based on the fixed-order results, no reliable uncertainty can be estimated.

8.2.7 Result for predicted fiducial cross-section

The modelling uncertainties also apply for the predicted fiducial cross-section. For the nominal prediction based in MG5__AMC +HERWIG 7, a nominal fiducial cross-section in the fiducial phase space of 1.331 ± 0.009 fb was determined in Section 3.7.2, where the uncertainty is due to the limited amount of simulated events.

The modelling uncertainty of the shower estimated from the comparison to a MG5__AMC +PYTHIA 8 based simulation is not included. Additionally, the uncertainty to account for the effect of the missing interference contributions does not apply here since the predicted fiducial cross-section is presented for the $WZjj$ -EW6 process alone. The remaining modelling uncertainties are due to the choice of a PDF, the value of α_S , the scale choice, and the shower scale. Following the procedures mentioned above for the predicted cross-sections in the fiducial phase space, yields

$$1.33 \pm 0.01 \text{ (MC stat.)} \pm 0.02 \text{ (PDF)} \pm {}^0_{<0.01} (\alpha_S) \pm {}^{0.11}_{0.10} \text{ (scale)} \pm {}^{0.09}_{0.10} \text{ (shower)} \text{ fb} \quad (8.3)$$

$$1.33 \pm {}^{0.15}_{0.14} \text{ fb} . \quad (8.4)$$

8.2.8 Caveats for uncertainties from two-point variations

The way to define a systematic uncertainty from two alternative simulations is not trivial, and no clear recommendation was found. An alternative proposal for the construction is to treat the variations as uncorrelated across bins rather than completely correlated. This introduces one independent uncertainty per bin rather than one overall uncertainty. This approach was tried for this work, and the decorrelated uncertainties yielded less conservative results compared to the procedure assuming correlations across bins. Especially for fits which favour a varied distribution over the nominal distribution² for a background process, in the decorrelated approach potentially several uncertainties need to be varied rather than a single uncertainty in the correlated approach. While the decorrelated approach is more flexible in the form of the adjustments, in order to obtain a similar effect as in the correlated approach, multiple bins have to be adjusted simultaneously. These variations from the nominal predictions are treated as less probable in the likelihood approach, which leads to higher significances for the rejection of the background-only hypothesis when assuming the existence of a signal.

²That is, a fit that favours a non-zero nuisance parameter.

Chapter 9

Cross-section Measurement

With validated estimates of the signal and background predictions as well as their systematic uncertainties, all ingredients for the final cross-section measurement are in place. The ingredients are combined using a profile-likelihood-based fit from which an observation significance and the cross-section of the signal process are determined. Before the presentation of the results, the basic principles of the used statistical methods are outlined.

9.1 Statistical Methodology

The methods used in the statistical analysis of the data are briefly summarized, following the notation and partially based on [168], [169] and [170], each presenting excellent introductions to the topic.

The likelihood is introduced as the conditional probability for the measurement of a fixed set of data given a theory hypothesis $P(\text{data}|\text{theory})$. Since the measured set of data is regarded as fixed, the likelihood is often written as $L(\text{theory})$. Different theories or different parameter values of a theory can be compared using the likelihood of a given measurement. Based on the likelihood, hypotheses can be tested, and parameter values can be estimated by maximizing the likelihood of a measurement.

The theory used in this work comprises of the signal and background models combined including a parameter μ_S for the signal strength, which scales the amount of predicted signal compared to the nominal prediction. The predicted total number of events can be written as

$$\mu_S \cdot S + B. \quad (9.1)$$

With the probability density function (pdf)¹ $f(x_i; \mu_S)$ of a measurement of n observables x_i given the theory with signal strength μ_S , the Likelihood can be expressed as

$$L(\mu_S) = \prod_{i=1}^n f(x_i; \mu_S). \quad (9.2)$$

While the measurement can in principle be evaluated event-by-event for a physical observable such as an invariant mass or a **BDTscore**, often binned Likelihoods are used, where each observable represents the event yield in a range of the underlying physical observable. This only changes the choice of the observables and the structure of Equation (9.2) is unchanged.

The signal strength is used for the calculation of the final cross-section and is therefore referred to as parameter-of-interest. It is useful to consider additional parameters θ as well, called nuisance parameters (NP) of the theory for which the final value is not of interest in itself. Such parameters can be used to model the effect of systematic uncertainties. For example, the uncertainty in the measurement of the integrated luminosity affects the signal, as

¹Note the different capitalization for the abbreviation as compared to “PDF”, used throughout this work for parton distribution function.

well as the background predictions and deviations from the nominal value, can be quantified using a nuisance parameter.

The information of external measurements of the additional parameters θ , in general, can be included in the likelihood by including (potentially hypothetical) external measurements of the parameter. A new term, referred to as constraint term, is included in the likelihood-based on an assumed pdf $f(\theta^{\text{ext}}; \theta)$ of the externally measured value θ^{ext} and its true value θ

$$L(\mu_S, \theta) = \left(\prod_{i=1}^n f(x_i; \mu_S, \theta) \right) \cdot f(\theta^{\text{ext}}; \theta). \quad (9.3)$$

The pdfs $f(\theta^{\text{ext}}; \theta)$ are based on external measurements and include the associated uncertainties. While ideally, the full likelihood of the external measurement can be included, often it is approximated using a Gaussian distribution $G(\theta^{\text{ext}}; \theta, \sigma_\theta)$.

9.1.1 Likelihood construction

The framework used for the construction of the Likelihood is called **HistFactory** and is described in detail in [171]. This framework employs the **Roofit**[172] implementations of fundamental statistical tools and simplifies the construction of complex likelihoods for binned histograms in multiple channels and regions.

The likelihood for a measurement n_j is based on the conditional probabilities $P(n_j | \mu_S s_j + b_j)$ for the measured value n_j given the predicted value $\mu_S s_j + b_j$ in each bin j assuming a Poisson distribution

$$P(n_j | \mu_S s_j + b_j) = \text{Pois}(n_j | \mu_S s_j + b_j) = \frac{(\mu_S s_j + b_j)^{n_j}}{n_j!} e^{-(\mu_S s_j + b_j)}. \quad (9.4)$$

The baseline likelihood covering only statistical uncertainties is determined by multiplying the conditional probabilities over all bins j and regions r

$$L^{\text{stat. only}}(\mu_S) = \prod_{r \in \text{regions}} \prod_{j_r \in \text{bins in } r} \text{Pois}(n_{j_r} | \mu_S s_{j_r} + b_{j_r}) \quad (9.5)$$

Similar to Equation (9.3), this baseline likelihood can be extended to cover systematic uncertainties via a set of nuisance parameters θ_n as well as their constraint terms from external measurements. The predicted expectation values for the signal and background process s_{j_r} and b_{j_r} now depend on the nuisance parameters θ .

Each nuisance parameter can be categorized according to the associated constraint term:

Unconstrained NP μ_B . For some nuisance parameters, no explicit constraint term is included at all. This is used for the normalization uncertainty of the leading prompt backgrounds $WZjj$ -EW4, $ZZjj$ -EW4, $t\bar{t} + V$, and tZj . The technical implementation in the likelihood is equivalent to the signal strength μ_S , which is not considered a nuisance parameter, but a parameter-of-interest. The different interpretation is indicated here using index B for the nuisance parameter.

This approach implies that the predicted normalizations are not used for these processes, and only data in dedicated control regions included in the likelihood is used to derive the normalization. The nominal value of these NP is 1, and they are applied multiplicatively in each bin of the predictions.

Poissonian NP γ . The limited amount of simulated events is considered as systematic uncertainty in the final measurement, where the observed prediction is considered as an external counting measurement. The associated uncertainties are constrained using a Poissonian distribution to reflect the nature of the counting experiment.

The nominal values of these NP are 1. In order to achieve this nominal value, an additional scale factor τ_{j_r} is introduced for each γ_{j_r} parameter in the expectation value of the Poissonian constraint term.

In **HistFactory**, one γ parameter is added for each bin, i.e. all statistical uncertainties from the simulation are combined for all processes in order to reduce the complexity of the likelihood and to increase performance. This parameter τ_{j_r} depends on the considered bin j_r , but can safely be treated as a fixed parameter.

Gaussian NP α . All other, in particular, all experimental, uncertainties are included using a Gaussian constraint term. The up- and down-variations of the underlying parameter affect the predicted yields and distributions. These effects are propagated to the likelihood, as discussed in Chapter 8.

The α nuisance parameters follow a normal distribution, and in each bin, the effects on the predictions are estimated based on interpolations between the nominal and the variations. The nominal prediction is mapped to $\alpha = 0$, and the potentially asymmetric effects of the variations are mapped. $\alpha = +1$ results in the provided up variation and $\alpha = -1$ in the down variation of the underlying parameter.

In the technical implementation of Gaussian NPs α , the effect of each α is separated in normalization and shape effects with different schemes for interpolation or extrapolation in order to improve stability of the fit. For both effects, polynomial interpolation is used, while the extrapolation method differs for the normalization and the shape effect. For the normalization effect, exponential extrapolation is used, and linear extrapolation is used for the shape effect.

When using the short notation ν_{j_r} for the nominal expected number of event in bin j_r of region r , defined as

$$\nu_{j_r} = \mu_S \cdot s_{j_r}(\alpha) + b_{j_r}(\alpha, \mu_B), \quad (9.6)$$

the full likelihood can be written as

$$L(\mu_S, \mu_B, \gamma, \alpha) = \prod_{\substack{r \in \\ \text{regions}}} \prod_{\substack{j_r \in \\ \text{bins in } r}} \text{Pois}(n_{j_r} | \gamma_{j_r} \cdot \nu_{j_r}) \text{Pois}(m_{j_r} | \gamma_{j_r} \cdot \tau_{j_r}) \cdot \prod_{\alpha_p \in \alpha} G(0; \alpha_p, 1). \quad (9.7)$$

This requires that the effect of potential differences between external measurements of nuisance parameters α and their values used for the evaluation of the nominal predictions can be neglected.

The notations used in Equation (9.7) are defined as

μ_S : Scale factor for signal strength,

μ_B : Set of nuisance parameters without explicit constraint term,

γ : Set of nuisance parameters with Poissonian constraint term,

α : Set of nuisance parameters with Gaussian constraint term,

τ_{j_r} : Scale factor introduced to allow for nominal value of associated γ_{j_r} to be 1,

n_{j_r} : Measurement of event yield in bin j_r ,

m_{j_r} : Observed prediction from simulations for event yield in bin j_r as measurement of true underlying prediction of simulation.

9.1.2 Profile likelihood ratio method

The chosen observable used for the evaluation of statistical tests is referred to as test statistic. While it is possible to use the likelihood as test statistic directly, no direct information is contained from the numerical value of the likelihood given a hypothesis alone. For this reason, test statistics are often based on the profile likelihood ratio λ

$$\lambda(\mu_S) = \frac{L(\mu_S, \hat{\mu}_B, \hat{\gamma}, \hat{\alpha})}{L(\hat{\mu}_S, \hat{\mu}_B, \hat{\gamma}, \hat{\alpha})}, \quad (9.8)$$

where the likelihood in the denominator is evaluated at its maximum, i.e. the maximum-likelihood estimators (indicated by the single hat, e.g. $\hat{\alpha}$) of each parameter are used. The parameter of interest, here the signal strength parameter μ_S , in the numerator is a free parameter in the profile likelihood ratio $\lambda(\mu_S)$ and constrains the maximization of the associated likelihood. That means the remaining parameters, i.e. the nuisance parameters, are chosen to maximize the likelihood for a given μ_S . These constrained maximum likelihood estimators of the parameters depend on μ_S and are indicated by a double-hat (e.g. $\hat{\hat{\alpha}}$).

This profile likelihood ratio is constrained to values of $0 \leq \lambda(\mu_S) \leq 1$ and quantifies the compatibility to data of a theory hypothesis described by a given μ_S compared to the hypothesis with global maximum likelihood estimators.

The actual test statistic used throughout this work is q_0 for the discovery of a positive signal (see [169])

$$q_0 = \begin{cases} -2 \ln \lambda(0) & \hat{\mu}_S \geq 0, \\ 0 & \hat{\mu}_S < 0, \end{cases} \quad (9.9)$$

where the profile likelihood ratio for the background-only hypothesis $\lambda(0)$ is used. This test statistic quantifies the compatibility of the background-only hypothesis with data. The value of q_0 is required to be positive and larger values indicate larger incompatibility.

Since the signal and as well as the integrated interference with the main background are predicted to be positive, cases with a negative maximum-likelihood estimator for the signal strength $\hat{\mu}_S$ are treated separately. For a measurement with $\hat{\mu}_S < 0$, the test statistic q_0 is defined to imply perfect compatibility ($q_0 = 0$) with the background-only hypothesis. This adjustment ensures incompatibility with the background-only hypothesis due to an under fluctuation is not interpreted as evidence for the proposed signal.

Hypothesis tests: For a hypothesis test, the p -value of a measurement is used, defined as the probability for a random fluctuation predicted by the tested theory to be equally or less compatible with the tested hypothesis than the given measurement. Hypotheses that yield a p -value below a threshold are rejected.

For the sake of readability, the p -value is often transformed to a significance Z such that

$$p_{\mu_S} = \int_{Z_{\mu_S}}^{\infty} G(x; 0, 1) dx. \quad (9.10)$$

The significance Z can also be interpreted as the distance in units of the standard deviation σ a measurement of a Gaussian distributed observable is moved from its expectation value to yield the same one-sided p -value.

This transformation can also be applied using the inverse cumulative distribution function of the normal Gaussian distribution Φ^{-1} as

$$Z_{\mu_S} = \Phi^{-1}(1 - p_{\mu_S}). \quad (9.11)$$

For testing a hypothesis defined by a given signal strength μ_S , it is necessary to know the pdf $f(q_0|\mu_S)$ of the test statistic q_0 given a value for the signal strength μ_S . From this distribution, the p -value of a measurement given a hypothesis of μ_S can easily be determined via integration from the measured value of the test statistic $q_{0, \text{obs}}$ to infinity

$$p_{\mu_S} = \int_{q_{0, \text{obs}}}^{\infty} f(q_0|\mu_S) dq_0. \quad (9.12)$$

A reliable but potentially very computationally intensive way to obtain the distribution $f(q_0|\mu_S)$ is to sample the distribution using toy datasets. The full evaluation of the test statistic, including the maximum likelihood estimators in the constrained and the unconstrained case, has to be performed for each toy dataset. With a typical number of NP with about ten bins, the evaluation of the test statistic for one toy already takes a few minutes on a single CPU core.

In order to claim the discovery of a signal, the incompatibility of the background-only hypothesis with the measurement needs to have a significance of $Z_0 = 5\sigma$, which translates to a p -value of $2.87 \cdot 10^{-7}$. For a reliable estimate of the p -value, at least a few toys need to pass this p -value, leading to an estimated number of necessary toys of between 5 and 10 million. An estimated CPU time is in the order of $10^5 - 10^6$ CPUh, which makes such calculations inaccessible.

An alternative approach relies on the large sample limit and is based on [173], which includes a generalization of Wilks theorem [174]. For one parameter of interest, it is shown in [173] that

$$-2 \ln \lambda(\mu_S) = \frac{(\mu_S - \hat{\mu}_S)^2}{\sigma_{\mu_S}^2} + \mathcal{O}(1/\sqrt{N}), \quad (9.13)$$

where N is the sample size and σ_{μ_S} is the standard deviation of the maximum-likelihood estimator $\hat{\mu}_S$ around the true value of the signal strength parameter.

From these results an estimate of the pdf $f(q_0|0)$ of the test statistic q_0 in the background-only hypothesis is presented in [169] as

$$f(q_0|0) = \frac{1}{2} \delta(q_0) + \frac{1}{2} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{q_0}} \exp\left(-\frac{q_0}{2}\right). \quad (9.14)$$

Also in this special case, in the large sample limit the discovery significance is

$$Z_0 = \sqrt{q_0} = \begin{cases} \sqrt{-2 \ln \lambda(0)} & \hat{\mu}_S \geq 0, \\ 0 & \hat{\mu}_S < 0, \end{cases} \quad (9.15)$$

These two approaches for the estimation of $f(q_0|0)$ and thus the discovery significance have different regions of applicability and validity. Depending on the situation, the applicability has to be evaluated in each analysis.

Confidence intervals: Another important result presented in this work is an estimator as well as a confidence interval for the fiducial cross-section derived from the signal strength. The estimator is obtained from a maximization of the likelihood presented in Equation (9.7). The determination of the confidence interval, generally defined as a random interval which contains the actual value of a parameter with a given probability, called confidence level, is not as straight forward. Typically, the confidence interval is presented for a confidence level of 68% equivalent to an interval with one standard deviation around the mean for a Gaussian distribution.

The accurate approach is performed using a confidence belt construction. In this construction, a scan of different values for the true parameter μ_S is performed.² For each value of μ_S the distribution $f(\hat{\mu}_S; \mu_S)$ of the estimator $\hat{\mu}_S$ is determined, e.g. using a toy-based study. Each toy measurement is obtained by sampling the pdf of the measured parameter given the currently considered value of μ_S and the $\hat{\mu}_S$ is determined using the same methods as in the nominal fit to the observation. A confidence interval $[\hat{\mu}_S^{\min}(\mu_S), \hat{\mu}_S^{\max}(\mu_S)]$ is determined for $\hat{\mu}_S$ with the desired construction, e.g. a confidence level of 68% with equal probabilities to obtain a value outside of the interval towards higher and lower values. The confidence intervals vary with μ_S and the dependency on the true value μ_S define the confidence belt. The final confidence interval is defined as the set of values of μ_S for which the observed value of $\hat{\mu}_S$ is in the confidence interval $[\hat{\mu}_S^{\min}(\mu_S), \hat{\mu}_S^{\max}(\mu_S)]$.

The full approach via the confidence belt is potentially very CPU-intensive because a sufficiently large number of toys has to be determined for a large number of values in μ and for each toy the full test statistic has to be re-evaluated.

A commonly used approximation is available using the profile likelihood ratio $\lambda(\mu_S)$. If Wilks theorem is applicable, the value of the $\lambda(\mu_S)$ can directly be related to the significance Z_{μ_S}

$$Z_{\mu_S} = \sqrt{-2 \ln \lambda(\mu_S)} \quad (9.16)$$

²The approach is applicable for other parameters θ as well, but the sake of simplicity the parameter of interest μ_S is used as example here.

(see [168]). This implies, a confidence interval for a confidence level of 68 % can be constructed from all values of μ_S , where $\ln \lambda(\mu_S) > 0.5$. Such intervals are also referred to as likelihood intervals.

While this approach is only an approximation to the confidence belt approach, it can be shown to converge towards the classical confidence interval in the large sample limit. The computing time in this approach is significantly reduced since the problem is reduced to finding μ_S for which the likelihood has a specific value, which is efficiently implemented in the MINOS method in the MINUIT framework[175]. A good discussion of the applicability of this method even for non-parabolic profile likelihoods is presented in [176].

9.2 Fitting Setup

The final fitting follows the approach of [3] and is performed simultaneously in 4 regions; the signal region, the WZ control region, the b control region, and the ZZ control region. The choice of distributions for the fit differs from [3] as discussed in Section 9.2.3 below.

9.2.1 Technical implementation

The technical implementation of the fit is independent of [3]. As an alternative, a package that was developed for [177] and [2] was applied, due to the easier interfacing with the existing event selection and plotting frameworks. For the application in this work, the possibility for MC-based toy studies was implemented. In addition to this, only minor adjustments were necessary.

The fitting package is based on libraries provided by RooFit[172] and RooStats[178]. The required input information for the fit as well as the construction of the likelihood function is performed using HistFactory[171] and follows the description of Section 9.1.1. The fitting framework determines results by default using the asymptotic formulas of the profile likelihood approach. The applicability of this approach is discussed briefly in the previous section and is cross-checked using a MC-based toy study in Section 9.3.1.

9.2.2 Input processes

The grouping of individual processes as input to the fit defines how fine-grained the different background contributions can be adjusted during the fitting procedure. The full list of process groups as input for the fit is

- $WZjj$ -EW6. The signal process is modelled in the nominal fit using the LO simulations based on MG5_AMC for the hard process and HERWIG 7 for the simulation of the parton shower. The normalization compared to the predicted event yield is defined by the parameter-of-interest μ_S . Detailed modelling uncertainties are derived as discussed in Chapter 8.
- $WZjj$ -EW4. The leading prompt background process is modelled in the nominal fit using SHERPA-based simulations. By applying an unconstrained nuisance parameter $\mu_{WZjj\text{-}EW4}$ on the normalization of this background, the final normalization is purely determined from data mostly in the WZ control region during the main fit. Additionally, detailed modelling uncertainties are derived as discussed in Chapter 8.
- $ZZjj$ -EW4. The subleading prompt background process is modelled in the nominal fit using SHERPA-based simulations. By applying an unconstrained nuisance parameter $\mu_{ZZjj\text{-}EW4}$ on the normalization of this background, the final normalization is purely determined from data mostly in the ZZ control region during the main fit.
- tZj . The tZj background process is modelled in the nominal fit using MG5_AMC-based simulations interfaced with PYTHIA 8 parton shower simulation. This simulation is obtained by inverting the veto on b quarks in the initial or final state of the matrix element calculation in the MG5_AMC-based signal sample. By applying an

unconstrained nuisance parameter μ_{tZj} on the normalization of this background, the final normalization is purely determined from data mostly from the b control region during the main fit.

- $t\bar{t} + V$. The production of a top-quark pair in association with a vector boson is modelled in the nominal fit using SHERPA-based simulations. By applying an unconstrained nuisance parameter $\mu_{t\bar{t}+V}$ on the normalization of this background, the final normalization is purely determined from data mostly in the b control region during the main fit.
- Non-prompt. The non-prompt background is estimated using the matrix-method using the associated uncertainties as discussed in detail in Section 6.2.
- VVV . The production of three electroweak gauge bosons is included in several final states and is based on SHERPA predictions. A total normalization uncertainty of 30 % is applied.
- $ZZjj$ -EW6. The scattering of electroweak gauge bosons in the fully-leptonic ZZ final state is treated as a separate input to the fit. The overall impact is minimal, but this process features a similar kinematic behaviour of the tagging jets compared to the signal process. A normalization theory uncertainty of 40 % is applied.

9.2.3 Input distributions

In order to perform a sufficient number of fits to toy datasets for the validation of the results, the input distributions were chosen to reduce computing time. Wherever feasible, the number of bins of the input distributions is reduced.

In the WZ as well as the ZZ control region, the predicted event yields show a high purity of the $WZjj$ -EW4 and $ZZjj$ -EW4 process, respectively (see Table 6.2). Due to the high purity, these regions only have limited possibility to restrict other backgrounds. For this reason, only the integrated event yields are included in the fit rather than a binned distribution.

In this way, the regions are mainly used to derive data-driven normalizations of the $WZjj$ -EW4 and $ZZjj$ -EW4 processes in order to not rely on the simulation-based predictions. While the fit in a differential distribution could be used to constrain other nuisance parameters, e.g. associated with modelling uncertainties, the applicability of these constraints for the signal region is not trivial and might depend on other parameters, e.g. different kinematics of the underlying object.

In the b control region, multiple processes are predicted to contribute significantly to the total event yield; $WZjj$ -EW4, tZj , and $t\bar{t} + V$. The $WZjj$ -EW4 process is well-constrained from the WZ control region, but it is still useful to fit the event yield differentially in order to allow for a separation between tZj and $t\bar{t} + V$. As it can be seen in Figure 9.1, the multiplicity of jets provides a good handle on this. While the $t\bar{t} + V$ process contributes mostly for very high multiplicities of at least four jets, the tZj process contributes to similar amounts in each of the three bins.

Similar discrimination between the different background processes would also be possible using the multiplicity of b -tagged jets. Due to the limited set of available simulations, some uncertainties are estimated by folding an alternative simulation from truth to reconstruction-level (see Chapter 8). Such folding is very sensitive to experimental effects. In order to improve the reliability of this folding, the jet multiplicity is used rather than counting only b -tagged jets. This choice reduces the additional dependency of the experimental flavour tagging techniques.

In the signal region, the event yield is fitted differentially in the **BDTscore** to improve the separation of the signal process from backgrounds. A variable bin size is used to ensure signal sensitivity at high values for the **BDTscore** while reducing the number of overall bins.

In order to decide on the final binning, the predicted distributions for the **BDTscore** are compared for each process in Figure 9.2 with a uniform binning.

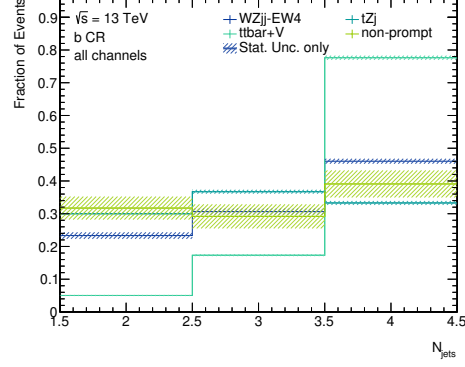


Figure 9.1: Relative predicted event fractions per range in exclusive jet multiplicity N_{jets} for processes with largest contributions in b control region. Colours represent different processes as indicated by the legend. Contributions with more than four jets are shown as part of the last bin. Uncertainties only show statistical uncertainties due to the limited number of simulated events.

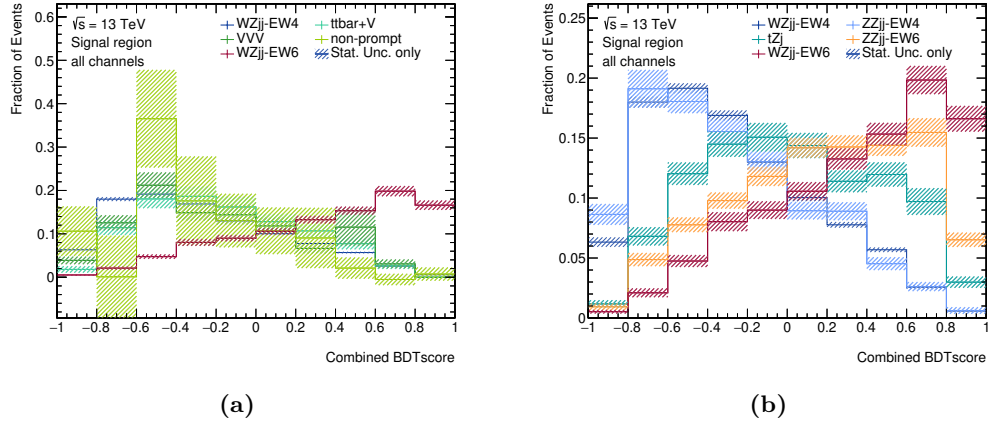


Figure 9.2: Relative predicted event fractions per range in BDTscore for different subsets of processes in signal region. Both plots show the $WZjj\text{-EW6}$ signal process (red) and the $WZjj\text{-EW4}$ leading background (dark blue) for comparison. Figure (a) additionally shows predictions for the $t\bar{t} + V$, VVV , and non-prompt backgrounds and Figure (b) for the $ZZjj\text{-EW4}$, $ZZjj\text{-EW6}$, and tZj backgrounds. Uncertainties only show statistical uncertainties due to the limited number of simulated events.

It can be seen that most backgrounds show a very similar distribution in the combined **BDTscore** with a maximum between -0.8 and -0.4 and a steady decrease for larger values. The only backgrounds with significantly different predictions are the tZj and $ZZjj$ -EW6 processes. As discussed in Section 2.2.1, the tZj process includes contributions of diagrams with the typical VBS signature, where a Z boson is radiated of a b -quark line. For these contributions, a similar **BDTscore** distribution as for the signal is expected. Due to the dominance of other contributions, e.g. with an s -channel top-quark resonance, the overall distribution is a mixture with a typical background distribution and is found to have a maximum around 0. For the $ZZjj$ -EW6 process, the distribution of the **BDTscore** is very similar to the signal prediction. For this process, the typical VBS signature is expected in the same way as for the signal process, just with differences in the lepton kinematics. Since both processes include contributions with a VBS signature, the classification of these processes is expected to be rather signal-like as confirmed by the distributions.

The small shape differences in the predicted distributions for most backgrounds imply that the fitting procedure will not be able to distinguish between the processes using a fine binning for background-dominated regions of the **BDTscore**. For those regions, a very coarse binning with an interval length of 0.4 is chosen to increase computing performance and stability of the fit.

For regions predicted to be signal-dominated, a finer binning is used to keep the information and to build regions with very high signal to background ratio. In this region, the predicted number of events, especially of the background-only prediction is minimal and even below 1. These bins contribute significantly to the observed significance. However, the applied statistical methods have to be carefully validated and cross-checked since some methods are derived in the large sample limit.

The binning used in the final fit features seven bins and is defined by thresholds at

$$[-1, -0.6, -0.2, 0.2, 0.6, 0.8, 0.9, 1]. \quad (9.17)$$

9.2.4 Input preparation

The inputs to the final fit need some minor adjustments in order to ensure the stability of the fit. For a proper estimation of uncertainties, the nominal predictions have to be non-zero and ideally positive in each bin for each considered process. However, especially in the region of large **BDTscore**, the nominal prediction of some processes is zero or even negative. This is indicated in the predicted event yields per process as listed in Table 9.1 below.

A special treatment to the affected histograms is applied. For bins with negative predicted event yield, the predicted event yield is instead conservatively estimated to be the absolute value of the nominal prediction. Such negative predictions only occur for the two bins with the highest **BDTscore** for the non-prompt background process.

For bins predicting zero events, i.e. regions not populated by simulated events, a reasonable scale for the simulations has to be estimated. This estimation is based on a per-event weight in each non-zero bin of the histogram for the given process. No information on individual events is available at this stage. Only the sum of weights of all events in a given bin is known. For this reason, homogeneous weights are assumed to perform this estimation. Under this assumption, a hypothetical number of simulated events N_{MC} can be estimated from the relative uncertainty ΔN_{rel} in a bin using the relation

$$\Delta N_{rel} = \frac{1}{\sqrt{N_{MC}}} . \quad (9.18)$$

If the predicted event yield in a bin is divided by N_{MC} , an estimated average weight can be extracted. These estimated average weights were found to be stable between bins for the affected histograms. In order to be conservative, this largest average weight for any bin of the considered histogram is used as a corrected prediction. Additionally, the relative uncertainty is set to 100% mimicking the prediction of a single simulated event with the estimated average weight. The only affected bins for this corrections are the last bin for the $t\bar{t} + V$, and the two last bins for the VVV processes in the signal region.

When applying this procedure to the fit input distributions, the stability and validity of the fits increased by avoiding unphysical predictions. However, the procedures have side-effects which need to be cancelled as well. First of all, the integrated event yield in a given region is increased. This increase is corrected by scaling the distribution to the same integral as before the corrections. Also, to ensure the effects of systematics are propagated correctly, the corrections above are directly applied only to the nominal predictions. The absolute effect of the corrections in each bin is stored and applied to each systematic correction. In this way, consistent corrections are applied to all variations, rather than determining a different correction for each variation.

Background-only control fit: Another effect that needs to be corrected in the fit inputs is a potential bias in the predicted normalization for certain backgrounds. The background normalization for the $WZjj$ -EW4, $ZZjj$ -EW4, tZj , and $t\bar{t}+V$ processes are not constrained to their simulated-based predictions in the fit. Instead, these normalizations are taken from data in dedicated control regions as part of the main fit. Since the expected results for the fit strongly depend on the assumed normalizations of these backgrounds, a potential bias due to a normalization mismodelling needs to be avoided.

This is done equivalently to [3] or [2], by applying a background-only fit to data in the control regions only. Since the signal contributions in these control regions are expected to be small, the normalizations of the backgrounds can be derived directly from data in this way.

The observed normalization factors for the background-only control region fit are

$$\mu_{WZjj\text{-EW4}} = 0.881 \qquad \mu_{ZZjj\text{-EW4}} = 1.076 \qquad (9.19)$$

$$\mu_{tZj} = 1.535 \qquad \mu_{t\bar{t}+V} = 0.961. \qquad (9.20)$$

These normalization factors are applied to the nominal predictions of the background in the main fit. It should be noted that these normalization factors do not affect the observed results of the fit to data in all regions. Since the normalizations of these backgrounds are not constrained any pure normalization effect can be cancelled, and the final scaling is solely based on the best-fit to data and is independent of the initial normalization. This independence has been validated in the used technical implementation of the fit as a cross-check.

Only the expected results are affected by construction. These are determined by fitting to a hypothetical data distribution equal to the nominal prediction, which depends on the initial normalization of these backgrounds.

The corrected input yields for the fits are listed differentially in the **BDTscore** in Table 9.2 for the signal region. Table 9.3 shows the corrected yields integrated for the WZ and ZZ CR and differentially in the jet multiplicity for the b CR.

Comparisons of the predicted and observed event yields differentially in the input observables used in the BDT, when applying these input corrections are shown in Section F.3 in the appendix for the signal region as well as the WZ control region.

9.3 Expected Fit Results

Before a fit to the observed dataset is performed, the expected results are determined by performing a fit to a so-called Asimov dataset. This Asimov dataset is defined as the nominal prediction of the signal+background hypothesis, given a signal strength of $\mu_S = 1$. Only the background normalizations are set to the values determined by the background-only control-region fit.

The results obtained from this fit to Asimov data are a good handle for a cross-analysis comparison because these are not affected by statistical fluctuations in data and provide an estimate for the sensitivity of an analysis.

A scan of the profile likelihood ratio over the signal strength parameter μ_S is shown for the fit, including all uncertainties, in Figure 9.3. The negative logarithmic profile likelihood ratio $-\ln \lambda(\mu_S)$ has its minimum at the best-fit value. As expected in the fit to Asimov data,

Table 9.1: Predicted event yields per range in the **BDTscore** in the signal region for each process considered in the fit before input corrections. The last column additionally lists the total predicted event yield in the signal region. The two bottom rows show the total expected event yield and the derived distribution of the Asimov dataset. Uncertainties show the total uncertainty determined by combining the effect of all nuisance parameters at their pre-fit values and uncertainties.

	$[-1, -0.6)$	$[-0.6, -0.2)$	$[-0.2, 0.2)$	$[0.2, 0.6)$	$[0.6, 0.8)$	$[0.8, 0.9)$	$[0.9, 1]$	total
$WZjj\text{-}EW4$	28 \pm 8	42 \pm 12	27 \pm 8	16 \pm 5	3.0 \pm 0.9	0.57 \pm 0.20	0.10 \pm 0.08	116 \pm 33
$ZZjj\text{-}EW4$	1.79 \pm 0.33	2.17 \pm 0.25	1.42 \pm 0.20	0.87 \pm 0.13	0.17 \pm 0.04	0.033 \pm 0.026	0.008 \pm 0.009	6.5 \pm 0.6
tZj	0.45 \pm 0.06	1.50 \pm 0.19	1.67 \pm 0.19	1.32 \pm 0.14	0.55 \pm 0.10	0.15 \pm 0.04	0.017 \pm 0.011	5.7 \pm 0.6
$ZZjj\text{-}EW6$	0.09 \pm 0.04	0.27 \pm 0.11	0.39 \pm 0.16	0.43 \pm 0.18	0.23 \pm 0.10	0.066 \pm 0.029	0.032 \pm 0.014	1.5 \pm 0.6
VVV	0.077 \pm 0.028	0.17 \pm 0.06	0.12 \pm 0.04	0.085 \pm 0.034	0.014 \pm 0.008	0.0000 \pm 0.00005	0.0 \pm 0.0	0.47 \pm 0.15
$t\bar{t} + V$	0.34 \pm 0.07	0.94 \pm 0.14	0.74 \pm 0.13	0.47 \pm 0.09	0.068 \pm 0.026	0.000 \pm 0.012	0.0 \pm 0.0	2.56 \pm 0.31
Non-prompt	0.5 \pm 0.7	2.6 \pm 1.1	1.1 \pm 0.7	0.5 \pm 0.5	0.00 \pm 0.06	0.05 \pm 0.08	0.0000 \pm 0.0017	4.8 \pm 2.3
$WZjj\text{-}EW6$	0.70 \pm 0.24	3.4 \pm 1.2	5.2 \pm 1.5	7.7 \pm 1.2	5.3 \pm 1.7	3.1 \pm 1.3	1.4 \pm 1.1	27 \pm 8
Total Expected	32 \pm 8	53 \pm 12	37 \pm 8	27 \pm 5	9.3 \pm 2.0	3.9 \pm 1.4	1.5 \pm 1.1	164 \pm 34
Asimov data	32 \pm 6	53 \pm 7	37 \pm 6	27 \pm 5	9.3 \pm 3.1	3.9 \pm 2.0	1.5 \pm 1.2	164 \pm 13

Table 9.2: Predicted event yields per range in the **BDTscore** in the signal region for each process considered in the fit after input corrections. The last column additionally lists the total predicted event yield in the signal region. The two bottom rows show the total expected event yield and the derived distribution of the Asimov dataset. Uncertainties show the total uncertainty determined by combining the effect of all nuisance parameters at their pre-fit values and uncertainties.

	$[-1, -0.6)$	$[-0.6, -0.2)$	$[-0.2, 0.2)$	$[0.2, 0.6)$	$[0.6, 0.8)$	$[0.8, 0.9)$	$[0.9, 1]$	total
$WZjj\text{-}EW4$	25 \pm 7	37 \pm 10	24 \pm 7	14 \pm 4	2.6 \pm 0.8	0.50 \pm 0.18	0.09 \pm 0.07	102 \pm 29
$ZZjj\text{-}EW4$	1.93 \pm 0.35	2.34 \pm 0.26	1.53 \pm 0.21	0.93 \pm 0.14	0.18 \pm 0.05	0.035 \pm 0.027	0.008 \pm 0.010	7.0 \pm 0.7
tZj	0.69 \pm 0.08	2.30 \pm 0.28	2.56 \pm 0.27	2.03 \pm 0.20	0.84 \pm 0.13	0.23 \pm 0.05	0.027 \pm 0.014	8.7 \pm 0.9
$ZZjj\text{-}EW6$	0.09 \pm 0.04	0.27 \pm 0.11	0.39 \pm 0.16	0.43 \pm 0.18	0.23 \pm 0.10	0.066 \pm 0.029	0.032 \pm 0.014	1.5 \pm 0.6
VVV	0.075 \pm 0.028	0.16 \pm 0.05	0.12 \pm 0.04	0.083 \pm 0.033	0.014 \pm 0.008	0.005 \pm 0.005	0.005 \pm 0.005	0.47 \pm 0.15
$t\bar{t} + V$	0.32 \pm 0.07	0.90 \pm 0.14	0.71 \pm 0.12	0.45 \pm 0.09	0.065 \pm 0.025	0.000 \pm 0.012	0.007 \pm 0.008	2.46 \pm 0.30
Non-prompt	0.5 \pm 0.6	2.6 \pm 1.1	1.1 \pm 0.7	0.5 \pm 0.5	0.02 \pm 0.08	0.05 \pm 0.08	0.011 \pm 0.005	4.8 \pm 2.3
$WZjj\text{-}EW6$	0.70 \pm 0.24	3.4 \pm 1.2	5.2 \pm 1.5	7.7 \pm 1.2	5.3 \pm 1.7	3.1 \pm 1.3	1.4 \pm 1.1	27 \pm 8
Total Expected	29 \pm 7	49 \pm 11	35 \pm 7	26 \pm 5	9.3 \pm 2.0	4.0 \pm 1.4	1.6 \pm 1.1	154 \pm 30
Asimov data	29 \pm 5	49 \pm 7	35 \pm 6	26 \pm 5	9.3 \pm 3.1	4.0 \pm 2.0	1.6 \pm 1.3	154 \pm 12

Table 9.3: Predicted event yields in the control regions for each process considered in the fit after input corrections. Table (a) lists the predictions for the b control region per range in the jet multiplicity, and the total predicted yield in the last column. Table (b) lists the total predictions for the ZZ control region, and Table (c) for the WZ control region. The two bottom rows show the total expected event yield and the derived distribution of the Asimov dataset. Uncertainties show the total uncertainty determined by combining the effect of all nuisance parameters at their pre-fit values and uncertainties.

(a)							
	2		3		≥ 4		total
$WZjj$ -EW4	11.7	± 3.3	15	± 5	23	± 8	50 ± 14
$ZZjj$ -EW4	1.03	± 0.13	1.47	± 0.12	2.35	± 0.31	4.8 ± 0.4
tZj	26.3	± 2.5	32.2	± 2.1	29	± 4	88 ± 6
$ZZjj$ -EW6	0.035	± 0.016	0.059	± 0.028	0.11	± 0.05	0.20 ± 0.08
VVV	0.039	± 0.015	0.091	± 0.035	0.16	± 0.06	0.29 ± 0.10
$t\bar{t} + V$	6.1	± 0.8	21.1	± 2.4	94	± 7	122 ± 7
Non-prompt	11	± 4	10.2	± 3.3	14	± 5	35 ± 11
$WZjj$ -EW6	0.52	± 0.18	0.40	± 0.25	0.4	± 0.4	1.4 ± 0.7
Total Expected	57	± 6	81	± 7	164	± 16	301 ± 24
Asimov data	57	± 8	81	± 9	164	± 13	301 ± 17

(b)		(c)	
	total		total
$WZjj$ -EW4	2.2 ± 0.7	$WZjj$ -EW4	590 ± 160
$ZZjj$ -EW4	68 ± 5	$ZZjj$ -EW4	47.0 ± 3.3
tZj	0.44 ± 0.07	tZj	22.2 ± 2.1
$ZZjj$ -EW6	5.0 ± 2.0	$ZZjj$ -EW6	1.6 ± 0.6
VVV	0.85 ± 0.26	VVV	2.7 ± 0.8
$t\bar{t} + V$	1.88 ± 0.25	$t\bar{t} + V$	21.8 ± 2.6
Non-prompt	1.2 ± 0.7	Non-prompt	41 ± 19
$WZjj$ -EW6	0.09 ± 0.05	$WZjj$ -EW6	19.9 ± 1.4
Total Expected	80 ± 6	Total Expected	750 ± 160
Asimov data	80 ± 9	Asimov data	750 ± 27

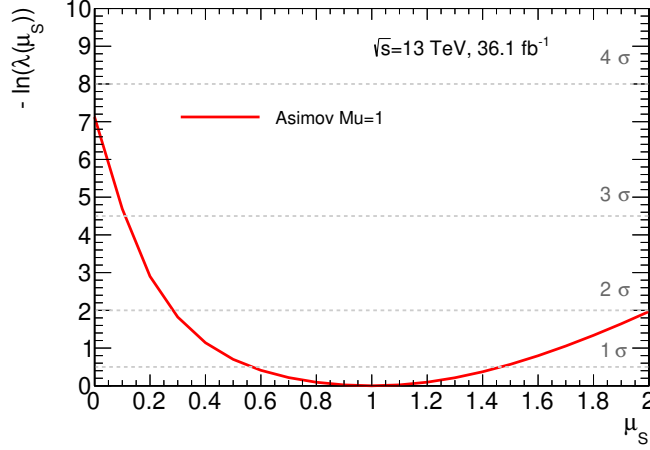


Figure 9.3: Profile likelihood scan in the signal strength parameter μ_S for the fit to Asimov data derived for $\mu_S = 1$. For each value of μ_S , the fit is reoptimized, and the negative logarithm of the profile likelihood ratio $-\ln \lambda(\mu_S)$ is shown. The value of the test statistic q_0 can be read off as the value of the red curve for $\mu_S = 0$. Thresholds for different background rejection significances are shown in grey dashed lines based on the asymptotic formulas.

the minimum is placed at $\mu_S = 1$, which was used to define the Asimov dataset. The test statistic q_0 is determined by doubling the shown value of the curve at $\mu_S = 0$ and is found to be $q_0 = 14.24$.

The fit to Asimov data yields an expected significance when only considering statistical uncertainties of $Z_{\text{exp, stat-only}} = 4.10\sigma$, which was found based on the asymptotic formulas. When adding theory modelling uncertainties, the significance is reduced to $Z_{\text{exp, stat+theo}} = 3.96\sigma$. When including the full set of uncertainties, the full expected significance of $Z_{\text{exp}} = 3.77\sigma$ is derived.

The estimate for the signal strength μ_S and the associated likelihood limits are

$$\mu_S = 1.00 \pm 0.36(\text{stat}) \pm 0.27(\text{theo}) \pm 0.14(\text{sys}) = 1.0 \pm 0.5_{0.4}^{0.5}. \quad (9.21)$$

Here the likelihood limits are split according to the associated sources of uncertainty. The largest contribution to the overall uncertainty is the statistical uncertainty. For a full comparison of the effect of individual uncertainty origins, the cross-section and the associated uncertainty measurements should be considered. A direct comparison here is problematic, since potential normalization effects of modelling uncertainties of the signal process $WZjj$ -EW6 are 100% anti-correlated to the signal strength. Since the normalization is taken from Asimov data in the fit, the uncertainty on the signal strength includes these anti-correlations to modelling uncertainties of the predicted cross-section. When combining the signal strength with the predicted cross-section, the anti-correlated contributions to the modelling uncertainties of both values cancel. Consequently, the measured cross-section should not be affected by these normalization effects, while the uncertainties of the signal strength are. The proper treatment is discussed for the observed result, and no cross-section for the expected result is reported.

The stability of the fit can be tested by checking the post-fit values for the nuisance parameters. Since the Asimov dataset matches perfectly to the nominal prediction, no significant pulls of the nuisance parameters are expected. The post-fit values and likelihood limits of each Gaussian nuisance parameter are shown in Figure 9.4.

As can be seen, all post-fit values of the Gaussian NP are still at their nominal values at 0.0. For the background normalization parameters of the main backgrounds which are not

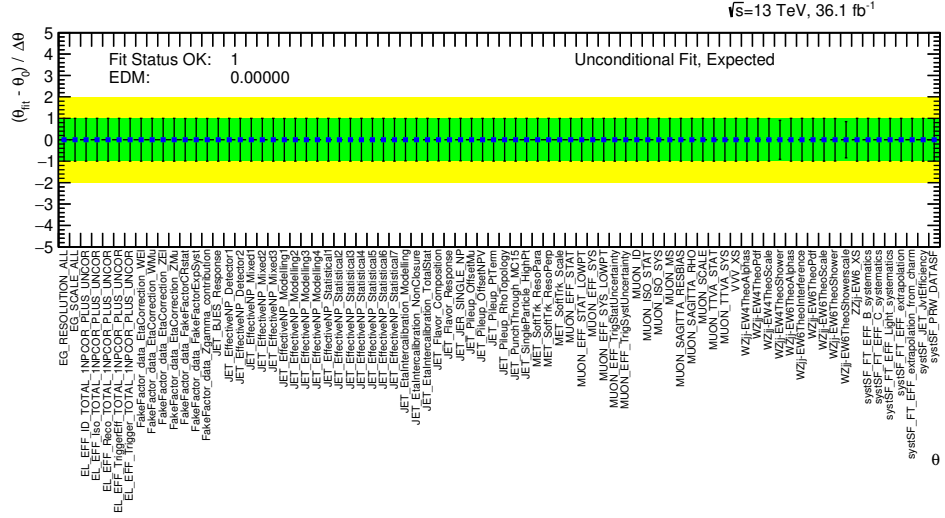


Figure 9.4: Comparison of pre and post-fit values of all Gaussian NP α for the fit with unconstrained signal strength to the Asimov dataset. The green band indicates the pre-fit uncertainties around the nominal value of 0. The black data points indicate the post-fit values and uncertainties.

shown in the figure, the post-fit values are

$$\mu_{WZjj\text{-}EW4} = 0.88 \pm_{0.22}^{0.26} \quad \mu_{ZZjj\text{-}EW4} = 1.08 \pm_{0.16}^{0.18} \quad (9.22)$$

$$\mu_{tZj} = 1.5 \pm_{0.5}^{0.6} \quad \mu_{t\bar{t}+V} = 0.96 \pm_{0.22}^{0.24} \quad (9.23)$$

While the uncertainties on the normalizations are large, the best-fit estimates are in very good numerical agreement at the per-mille level with their initial values determined in the background-only control-region fit.

The post-fit event yield distributions in all regions for the fit to Asimov data are shown in Figure 9.5. A detailed list of post-fit yields including their uncertainties for each input processes of the fit in each region are given in Section H.1.1 in the appendix.

9.3.1 Validation study using toy samples

The profile likelihood ratio shows a clear asymmetric behaviour, as seen in Figure 9.3. Such asymmetry is expected for the low event numbers, especially in the bins with highest `BDTscore`, where the statistical uncertainties can not yet be considered Gaussian, but have to be treated according to a Poisson statistic. This asymmetry can potentially indicate problems in the asymptotic formulas. For this reason, toy studies of the full fit were performed to validate the results of the asymptotic formulas.

Unfortunately, issues in closure tests were found for each tested way to generate toy datasets. The correct way to create toy datasets covering the effects of all types of nuisance parameters was not found, rendering the toy generation untrustworthy and the validation study inconclusive. The study is documented below for the sake of documentation.

The toy datasets were generated for the background-only hypothesis, i.e. the signal strength parameter μ_S was set to 0. The toy distributions are generated using tools of the `RooFit` framework based on the pdf provided by `HistFactory`. While it is technically possible to randomize the values of the nuisance parameters simultaneously, this was found to be too inefficient, and the generation of a toy dataset took several minutes already when adding only a few NPs. Since this method does not scale to a large number of NP in the full model, a different approach is followed.

Instead, a random value for each NP is determined prior to the actual generation of the toy dataset. The randomization of the NP values depends on the type of NP and follows

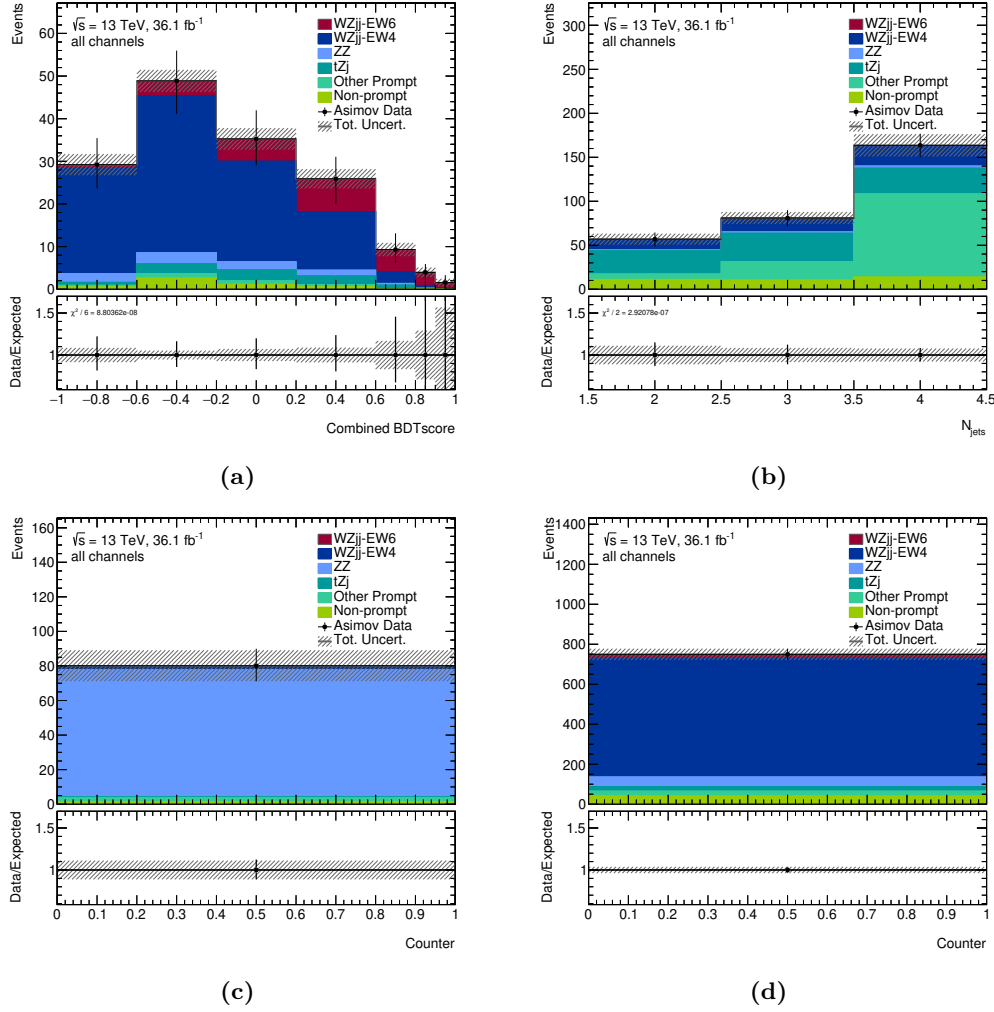


Figure 9.5: Post-fit event yield distributions for all regions for the fit to Asimov data. Figure (a) shows the differential distribution in the BDTscore for the signal region, and Figure (b) differentially in the jet multiplicity in the b control region. Figure (c) shows the total event yield for the ZZ control region and Figure (d) for the WZ control region. The effects of all nuisance parameters are propagated and the associated uncertainties are symmetrized in order to allow for a proper treatment of correlations between nuisance parameters.

Table 9.4: Detailed numbers for the definition of high statistics and low statistics simplified models. Choice of numbers is motivated in the text.

	Signal			Background		
	Bin 1	Bin 2	Bin 3	Bin 1	Bin 2	Bin 3
High statistics model						
Nominal	290	360	180	15300	7500	800
Statistical uncertainty	25	30	16	765	375	40
Gaussian NP 1	0	0	0	200	-50	0
Gaussian NP 2	0	0	0	100	20	-10
Gaussian NP 3	0	0	0	80	40	20
Low statistics model						
Nominal	2.9	3.6	1.6	1.53	0.75	0.08
Statistical uncertainty	0.25	0.3	0.16	0.01	0.01	0.001
Gaussian NP 1	0	0	0	0.56	0.12	0.05

pdfs motivated by the likelihood presented in Equation (9.7). For Gaussian NP α a random value is generated from a Gaussian pdf with mean of $\mu = 0$ and standard deviation of $\sigma = 1$. For Poissonian NP γ , a Poissonian pdf with an expectation value equal to the associated τ parameter from Equation (9.7) is used to generate a random value. This random value is divided by τ to determine the actual value of the γ parameter. For unconstrained NP μ_B no constraint term is present, so a uniform pdf has to be applied. In order to ensure that the mean value of the generated values is equal to the nominal value of the μ_B parameters μ_B^{nom} , the considered range for the NP is restricted to the interval from 0 to twice the nominal value $[0, 2 \cdot \mu_B^{\text{nom}}]$. By randomizing all nuisance parameters according to these prescriptions, all possible rather than only the purely statistical variations of the background-only model are tested. Special care was taken to ensure that each toy was generated using a unique seed to ensure statistical independence of the toy datasets.

The considered model is fit to each toy dataset in order to determine the value of the test statistic q_0 and only toys resulting in successful fits are considered. For some toy datasets, the fits may not converge properly, making the value of the test statistic unreliable. For cases resulting in a non-negligible amount of unsuccessful fits to toy datasets, the fraction is explicitly mentioned.

The generation of toy datasets is cross-checked for very simplified models with one region of 3 bins and only few NP. The definition of these models are summarized in Table 9.4. A low statistic model is defined, where the background-only prediction is roughly equal to the background prediction for the three bins with the highest **BDTscore** in the signal region of the main fit model. Due to this definition, effects of the Poissonian uncertainties for low numbers are expected to be similar as for the full fit. Poissonian NP of up to 10 %, a Gaussian NP of about 50 %, a luminosity uncertainty, and an unconstrained NP on the normalization are included.

A second simplified model referred to as the high statistics model is defined, where the background prediction is increased by a factor of 10000. Two additional Gaussian NP are included, and the effects of all NP are scaled to 1 – 5 % in the last bin. In this model, the lowest nominal predicted yield is 800. So for this model, the asymptotic formulas are assumed to be valid.

For both models, about 100000 toy distributions are generated, and the value of the test statistic q_0 in the corresponding model is determined for each toy dataset. The distributions of the test statistic q_0 are shown in Figure 9.6 and compared to the predictions of the asymptotic formulas. The predicted distribution for this test statistic following the asymptotic

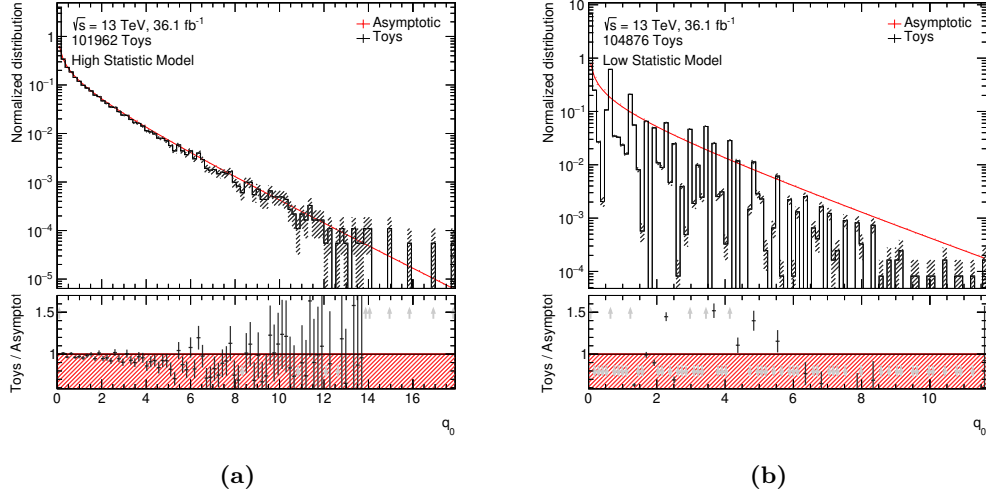


Figure 9.6: Probability distribution function of the test statistic q_0 derived from the asymptotic formulas (red) and from toy experiments using about 100k MC toy sample distributions (black). The exact number of toys is indicated in the top left corner of the plot. Figure (a) shows the pdf for the high statistics model and Figure (b) for the low statistics model as described in the text. The lower inlay shows the ratio of the toy-based distribution with respect to the distribution predicted by the asymptotic formulas.

formulas [169] is

$$f(q_0|0) = \frac{1}{2}\delta(q_0) + \frac{1}{2} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{q_0}} \exp\left(-\frac{q_0}{2}\right). \quad (9.24)$$

For the high statistic model, a very good agreement with the predictions of the asymptotic formulas is observed. For the low statistic model, the distribution determined using toys shows large fluctuations probably caused by the discrete nature of the toy datasets. Only integer numbers are generated in each event mimicking an actual measurement. Due to this, the number of possible distributions is greatly reduced, and many toy datasets are expected to agree by coincidence.

From these distributions, significances can be determined for each value of the test statistic q_0 . The asymptotic significance $Z_{\text{asymptotic}}$ is evaluated as $\sqrt{q_0}$ and the toy-based significance Z_{toys} can be calculated via the fraction of toy datasets resulting in an equal or larger value for q_0 . In Figure 9.7 the two significances are compared by evaluating the average Z_{toys} for all toys within a range of $Z_{\text{asymptotic}}$. The uncertainties of Z_{toys} include the standard deviation of toys within this range of $Z_{\text{asymptotic}}$ as well as an uncertainty due to the limited number of toy datasets estimated as the variance of the p -value

$$\Delta p = \sqrt{p \cdot (1 - p) / N_{\text{toys}}}. \quad (9.25)$$

If the asymptotic formulas are applicable $Z_{\text{asymptotic}}$ and Z_{toys} should agree within uncertainties. For the high statistic model, an unexpected constant offset of below 0.05 is found. The origin of this offset was not found in several variations of the model definition. Other than this offset, the two significances are in excellent agreement within uncertainties. For the low statistic model, the comparison shows erratic behaviour for low values of $Z_{\text{asymptotic}}$. This behaviour is assumed to be caused by the discreteness of the event numbers due to the low yields. Additionally, at higher values of $Z_{\text{asymptotic}}$, there is an offset of about 0.4σ . These severe discrepancies could be caused by a non-applicability of the asymptotic formulas or an issue in the generation of the toy datasets.

In this model, the fraction of toys resulting in unsuccessful fits was about 1%. Since the problematic behaviour already occurs at low significances any different behaviour of the 1% of toys with unsuccessful fits can not be the main origin of differences between the significance estimates.

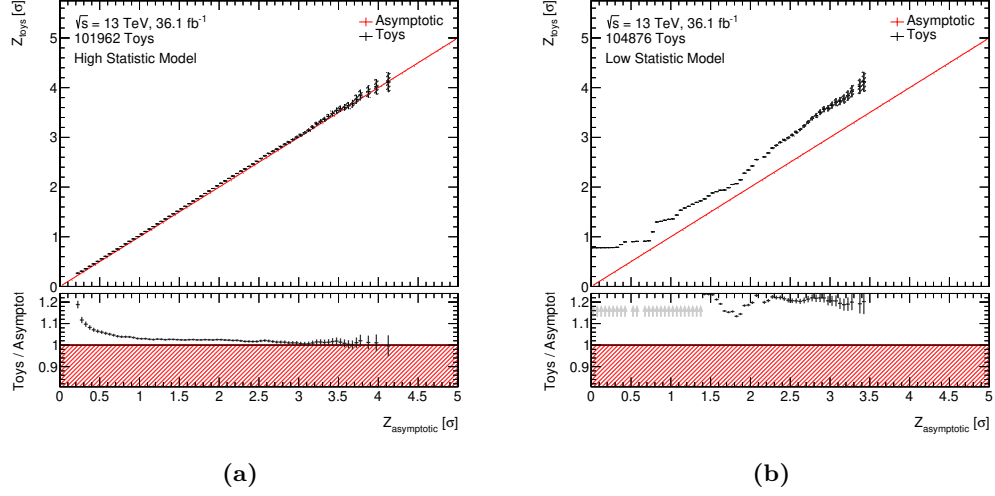


Figure 9.7: Comparison of the significance estimates based on the asymptotic formulas (x -axis) and the toy-experiment (y -axis). The averaged toy-based estimate and its uncertainty are shown as a black data point in each range of $Z_{\text{asymptotic}}$. The red line shows a scenario of perfect agreement between the two estimates. Figure (a) shows the significance comparison for the high statistics model and Figure (b) for the low statistics model as described in the text. The lower inlay shows the ratio of the toy-based estimate with respect to the estimate using the asymptotic formulas.

To further study the discrepancies, adjustments of the low statistic model are studied, where the considered set of NP is reduced. The most substantial part of the discrepancy could be traced to the handling of unconstrained NP. In a model without these NP, the discrepancies between $Z_{\text{asymptotic}}$ and Z_{toys} are greatly reduced. The corresponding comparisons for this further simplified model are shown in Figure 9.8.

It can be seen in Figure 9.8a that the distribution of the q_0 test statistic still has clear local minima and maxima, but the effect is far less pronounced as in the model including the unconstrained NP (see Figure 9.6b). The problematic behaviour in the comparison of the significances is now much clearer. In this model, a smooth dependency is observed, where Z_{toys} is significantly larger than $Z_{\text{asymptotic}}$ for low significances. This effect is assumed to be caused by the non-Gaussian distribution of the statistical uncertainties, based on the comparison to the high statistic model. The difference between the two significance estimates decreases with increasing $Z_{\text{asymptotic}}$ and is small compared to the statistical uncertainties of Z_{toys} for $Z_{\text{asymptotic}} > 3$.

While the problematic behaviour in the low statistic model is not understood, a further study of the underlying origins of these discrepancies is beyond the scope of this work and the procedure is applied to the full fit model. Toys were generated iteratively until the resulting statistical uncertainty on the expected Z_{toys} is considered acceptable. The distribution of the test statistic q_0 as well as the comparison of Z_{toys} and $Z_{\text{asymptotic}}$ for different values of the test statistic are shown in Figure 9.9, where also the value of the test statistic for the fit to Asimov data $q_0 = 14.24$ is indicated. For this q_0 , a toy-based significance of

$$Z_{\text{toys}} = 3.60 \pm_{0.07}^{0.09} \sigma \quad (9.26)$$

is determined. In this model, the fraction of toys resulting in unsuccessful fits is quite large at about 18 %. This large fraction could indicate problematic behaviour in the fit, e.g. caused by highly correlated NP. This issue limits the trust in the toy-based significance, which could be significantly altered by the large fraction of toy distributions which could not be evaluated properly.

Interestingly, the determined distribution of the test statistic and the comparison of significance estimates do not show strong indications of problematic behaviour observed for

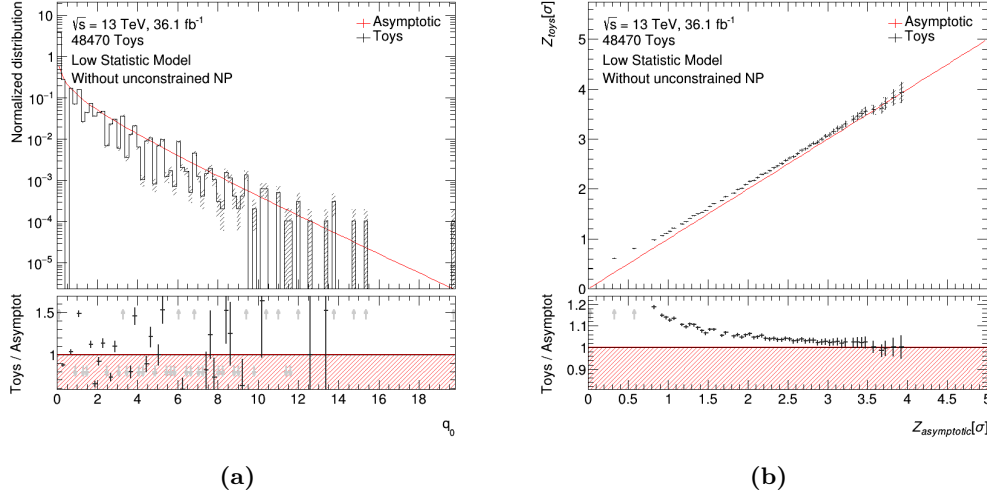


Figure 9.8: Comparison of toy-based results to results based on the asymptotic formulas for the low statistics model without including unconstrained NP. Figure (a) shows the pdf of the test statistic q_0 . Figure (b) a comparison of the significance estimates based on the asymptotic formulas (x -axis) and on the toy-experiment (y -axis). The averaged toy-based estimate and its uncertainty are shown as a black data point in each range of $Z_{\text{asymptotic}}$. The red line shows a scenario of perfect agreement between the two estimates. The lower inlay shows the ratio of the toy-based result to the result using the asymptotic formulas.

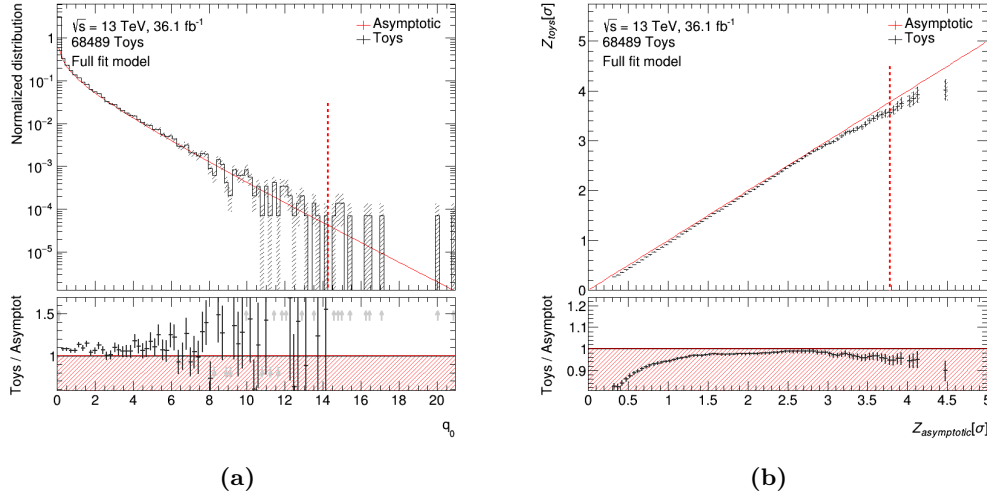


Figure 9.9: Comparison of toy-based results to results based on the asymptotic formulas for the full fit model. Figure (a) shows the pdf of the test statistic q_0 . Figure (b) a comparison of the significance estimates based on the asymptotic formulas (x -axis) and on the toy-experiment (y -axis). The averaged toy-based estimate and its uncertainty are shown as a black data point in each range of $Z_{\text{asymptotic}}$. The red line shows a scenario of perfect agreement between the two estimates. The lower inlay shows the ratio of the toy-based result to the result using the asymptotic formulas. The value of the test statistic for the fit to the Asimov dataset is indicated by a vertical red dashed line.

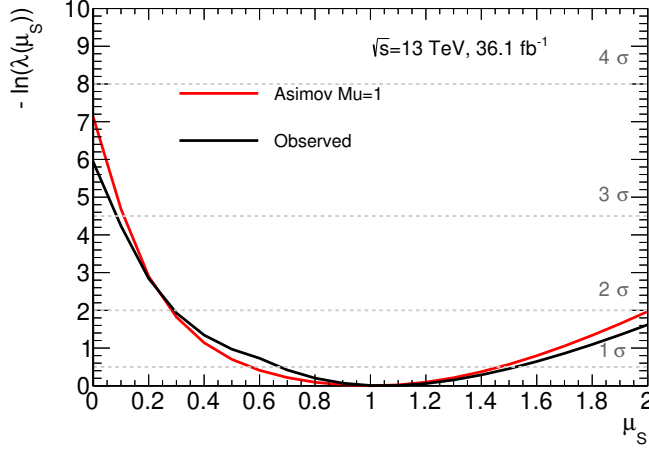


Figure 9.10: Profile likelihood scan in the signal strength parameter μ_S for the fit to the observed dataset (black) and the Asimov dataset with $\mu_S = 1$ (red). For each value of μ_S each fit is reoptimized and the negative logarithm of the profile likelihood ratio $-\ln \lambda(\mu_S)$ is shown. The value of the test statistic q_0 can be read off as the value of the red curve for $\mu_S = 0$. Thresholds for different background rejection significances are shown in grey dashed lines based on the asymptotic formulas.

the simplified models. The distribution of the test statistic does not show clear local minima and maxima. These are assumed to be smeared out by the presence of a larger number of additional bins compared to the simplified model.

The comparison of Z_{toys} and $Z_{\text{asymptotic}}$ shows a small offset of about 0.04σ as previously observed for the high statistic model, but here Z_{toys} is smaller than $Z_{\text{asymptotic}}$. The discrepancy of larger Z_{toys} for low $Z_{\text{asymptotic}}$ as observed in the low statistic model without unconstrained NP as well as the problematic behaviour of the low statistic model with unconstrained NP are not found here.

When considering the large fraction of toys with unsuccessful fits and the unexpected problematic behaviour in the low statistic model, the significance estimate based on the toys is only regarded as cross-check rather than replacing the significance based on the asymptotic formulas as the nominal result for the significance.

9.4 Observed Fit Results

Since the fit stability was confirmed in the fit to the Asimov dataset, the model is fitted to the observed dataset in a final step. A scan of the profile likelihood ratio for different values of the signal strength μ_S is shown in Figure 9.10.

It can be seen that the profile likelihood curve has a lower value at $\mu_S = 0$ and that the best-fit value is slightly shifted from its nominal value at 1. The derived value for the test statistic is $q_0 = 11.86$, which can be translated using the asymptotic formulas to a significance of

$$Z_{\text{obs}} = 3.44\sigma. \quad (9.27)$$

When only considering statistical uncertainties a significance of $Z_{\text{obs, stat-only}} = 3.70\sigma$ is found and when adding modelling uncertainties, the significance decreases to $Z_{\text{obs, stat+theo}} = 3.58\sigma$. The significance estimate based on the toy-study is $Z_{\text{obs, toys}} = 3.29 \pm_{0.04}^{0.05}\sigma$.

The best-fit value of the signal strength is found to be

$$\mu_S = 1.05 \pm_{0.30}^{0.35}(\text{stat}) \pm_{0.22}^{0.31}(\text{theo}) \pm_{0.09}^{0.10}(\text{sys}) = 1.1 \pm_{0.4}^{0.5}, \quad (9.28)$$

where the likelihood limits are split according to the associated sources of uncertainty. The leading source of uncertainty is of statistical nature, while also the modelling uncertainties contribute to a similar amount.

uncertainties, which are measured in dedicated analyses, the small pulls are an indication of compatibility with these dedicated measurements.

In the background-only fit also other NP show non-negligible pulls, e.g. for the efficiency of jet flavour tagging and energy measurement, as well as on the pileup reweighting. In the background-only fitting procedure, incompatibility with external measurements reflected as pulls in the associated NP is necessary to be compatible with the observed data in the signal region. The background-only hypothesis is clearly less compatible with data observed here and in external measurements than the full model, including the signal process. This incompatibility causes higher constraints in the background-only fit and consequently results in the observed rejection significance of the background-only hypothesis.

For further comparison of the effects of the different NP, a ranking is compiled based on the impact of the nuisance parameter on the signal strength. The impact for an individual NP θ is evaluated by reperforming the fit while fixing the respective NP to the up- and down-variations around its best-fit value $\hat{\theta}$ using its post-fit uncertainties. From these restricted fits, the signal strength estimators $\hat{\mu}_S^{\text{up, post}}$ and $\hat{\mu}_S^{\text{down, post}}$ are determined. The impact of the NP on μ_S is determined as the maximum distance to the best-fit value of the signal strength $\hat{\mu}_S$ and is used to order the NP.

In order to test the impact on the signal strength of a potential constraint of each NP, the fit is also reperformed while fixing the NP to the respective variations using the pre-fit uncertainties, yielding the restricted signal strength parameters $\hat{\mu}_S^{\text{up, post}}$ and $\hat{\mu}_S^{\text{down, post}}$. The impacts on the signal strength of the 20 highest-ranked Gaussian NPs α are compared in Figure 9.12. The full ranking, including all nuisance parameters, is listed in Appendix H.2.3. It is emphasised that for technical reasons, only Gaussian NPs are shown in Figure 9.12, the 12 highest ranked NPs are Gaussian NPs, so conclusions based on the figure are not affected significantly by this filtering.

The NP with the highest impact on the signal strength are two of the modelling uncertainties of the $WZjj$ -EW6 signal process. These are the modelling uncertainties with the highest impact on the predicted yield for high values of the **BDTscore** in the signal region. In the best-fit scenario, the post-fit signal yield, including effects of all NP and the signal strength, is determined from data. Any change introduced to predicted yields by varying the NP associated to the signal modelling is balanced in the unrestricted fit by adjusting the signal strength such that the overall post-fit yield is reduced³. This dependence introduces strong anti-correlations of these NP with the signal strength and consequently a substantial impact of variations of said NP on the signal strength.

Also, the third-highest-ranked NP is associated to a modelling uncertainty, the shower uncertainty of the $WZjj$ -EW4 background. This shower uncertainty is the uncertainty with the largest increase of the predicted yields with high values of the **BDTscore** in the signal region for this process. An increase of this predicted background yield reduces the signal yield necessary to fit the observed data distribution.

The ranking with the highest impact on the signal strength for three modelling uncertainties is consistent with the separation of systematic uncertainties on the signal strength listed in Equation (9.28). Modelling uncertainties are the dominant systematic uncertainties in general, and the ranking provides a more detailed insight into the composition of the combined modelling uncertainty. The experimental uncertainty on the measured signal strength is smaller than the modelling uncertainty, reflected in the smaller impact of NPs associated with experimental uncertainties.

Post-fit yields: The post-fit event yield distributions in all regions for the fit to data are shown in Figure 9.13. The effects of all nuisance parameters are propagated, and the associated uncertainties are symmetrised. This procedure allows for proper treatment of correlations between nuisance parameters. A detailed list of post-fit yields including their uncertainties for each input processes of the fit in each region are given in Section H.2.4 in the appendix.

³A full cancellation is not achieved in all cases due to the shape effects of NPs

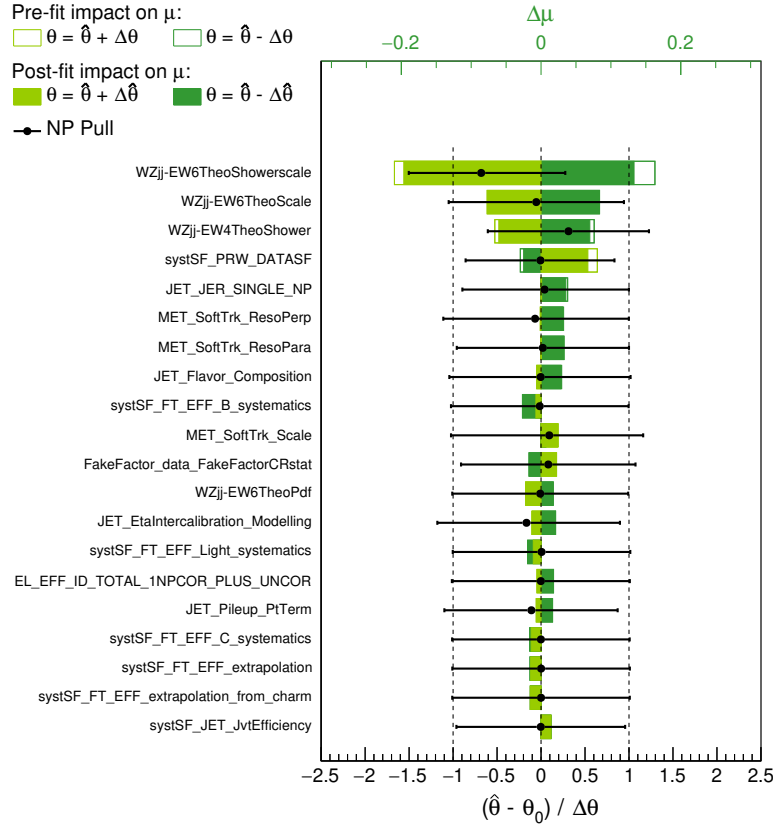


Figure 9.12: Impacts on the signal strength μ_S (upper x axis) and comparison of pre and post-fit values (lower x axis) for the 20 highest ranked Gaussian nuisance parameters α for the full fit to the observed dataset. The pre-fit values are 0 ± 1 , as indicated by the black dashed lines. The post-fit values are indicated using black data points. The impacts for up (down) post-fit 1σ variations of each NP on μ_S are indicated using full dark (light) green bars. For impacts based on the pre-fit 1σ variations only outlines of the bars are shown.

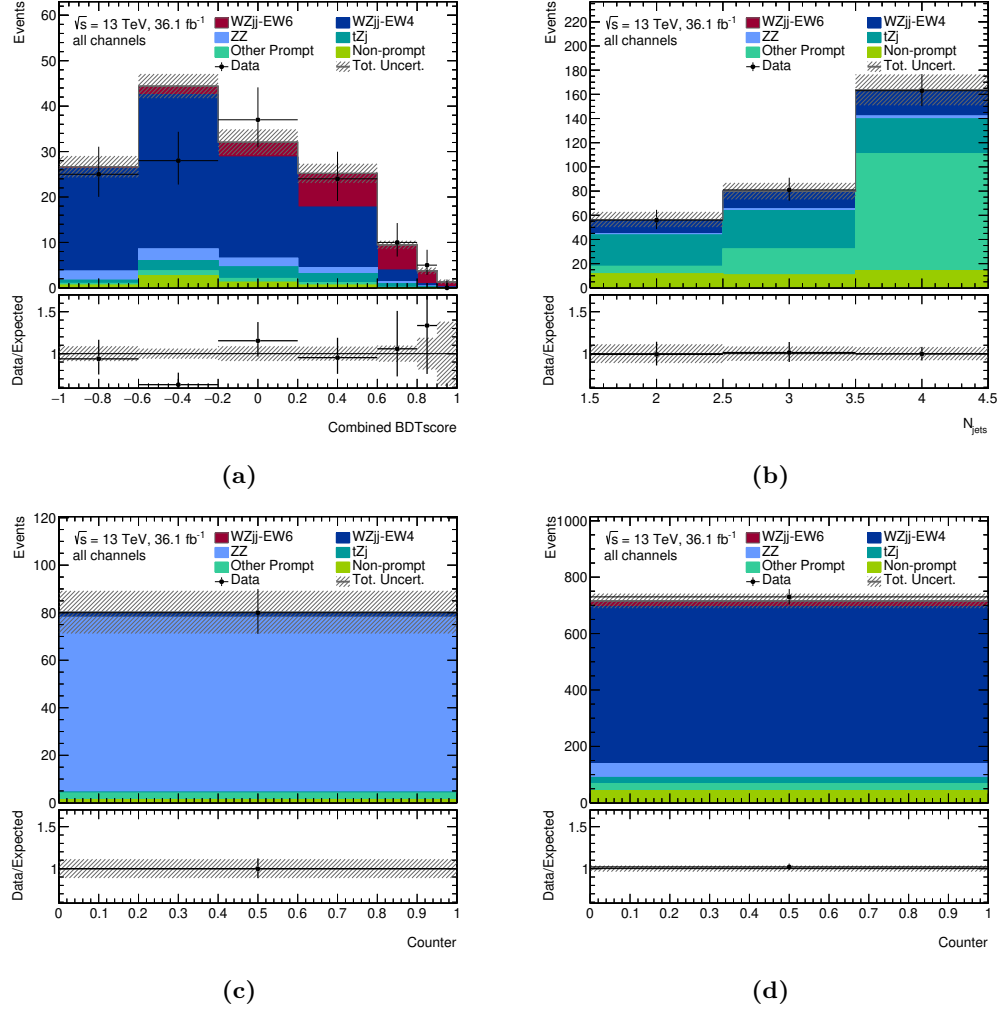


Figure 9.13: Post-fit event yield distributions for all regions for the fit to the observed dataset. Figure (a) shows the differential distribution in the **BDTscore** for the signal region, and Figure (b) differentially in the jet multiplicity in the b control region. Figure (c) shows the total event yield for the ZZ control region and Figure (d) for the WZ control region. The effects of all nuisance parameters are propagated, and the associated uncertainties are symmetrized to allow for a treatment of correlations between nuisance parameters.

Overall a good agreement of the predicted post-fit event yields and the observed distribution in data is found. In the control regions, excellent agreement is found, caused by the large flexibility of the modelling of normalizations of the main backgrounds. In the signal region, a very good agreement is found for most bins of the `BDTscore`. Only the bin from $[-0.6, -0.2)$ shows a local deviation of about two standard deviations. When considering the excellent agreement in the other ranges, it is plausible to assume that statistical fluctuations cause this deviation, which is, thus, not considered problematic here.

Cross-section extraction: The final result is a measurement for the fiducial cross-section σ_{obs} for the signal process. This cross-section can be determined by scaling the predicted cross-section σ_{pred} for the nominal signal model based on MG5_AMC and HERWIG 7 with the signal strength parameter μ_S

$$\sigma_{\text{obs}} = \mu_S \cdot \sigma_{\text{pred}}. \quad (9.31)$$

The predicted cross-section was determined in Section 8.2.7 to be $\sigma_{\text{pred}} = 1.33 \pm_{0.14}^{0.15} \text{ fb}$.

The correct determination of uncertainties of the measured cross-section is not trivial since the uncertainties of σ_{pred} and μ_S are highly anti-correlated and partially cancel. Consider as an example the variation of a hypothetical uncertainty that increases the predicted fiducial cross-section by 50 % would result in a signal strength reduction to $\frac{2}{3}$. This would cancel the normalization effect of the uncertainty and result in the same measured cross-section as without this uncertainty. So this +50 % uncertainty of σ_{pred} is 100 % anti-correlated to a -33 % uncertainty on the signal strength, while the resulting measured cross-section is unaffected by this uncertainty.

A possible way to determine the uncertainties of σ_{obs} is based on the assumption of a 100 % anti-correlation and perfect cancellation of the normalization uncertainties of σ_{pred} and their contributions to the modelling uncertainties for μ_S . Under this assumption, the uncertainties of σ_{pred} can be quadratically subtracted from the total modelling uncertainties of μ_S to yield the modelling uncertainties of σ_{obs} . The relative statistical and experimental systematic uncertainty components of σ_{obs} are equal to those of μ_S . This approach yields a measured fiducial cross-section of

$$\sigma_{\text{obs}} = 1.40 \pm_{0.40}^{0.46} (\text{stat}) \pm_{0.23}^{0.38} (\text{theo}) \pm_{0.11}^{0.14} (\text{sys}) \text{ fb} \quad (9.32)$$

$$= 1.4 \pm_{0.5}^{0.6} \text{ fb}. \quad (9.33)$$

An alternative, more rigorous approach is to perform a second fit with adjusted inputs for the theory uncertainties of the signal sample. For each of the uncertainty sources considered in the predicted cross-section, the effects on the fiducial cross-section at truth level needs to be separated from all remaining effects. Since only the effects on the fiducial cross-section introduce the correlation of the signal strength and the predicted cross-section, these are the effects that exactly cancel the uncertainties of the predicted cross-section. The cancellation implies no explicit uncertainties of the predicted cross-section need to be considered when simultaneously excluding effects of these uncertainties on the normalization of the fiducial cross-section from the signal strength measurement, i.e. the fitting procedure. In this case, the determined relative uncertainties of the signal strength are directly applicable also for the observed fiducial cross-section.

The remaining effects of the uncertainties include not only shape effects in each region, but also normalization effects due to the potential effect on the efficiency, i.e. the probability for each event in the fiducial phase space at truth level to be in the considered region at reconstruction level.

The fiducial correction necessary to remove these normalization effects are implemented for each modelling uncertainty affecting the predicted cross-section for each individual variation. The relative effect on the fiducial prediction is determined for each variation simultaneously to the overall effect in the signal region and each control region, and the effect on the predicted distributions are scaled to remove the normalization difference in the fiducial phase space. The combination of the individual variations as applied for the PDF and scale uncertainties is performed only after the fiducial correction.

This procedure is followed, and the obtained relative uncertainties of the signal strength are applied to the observed cross-section, which is found to be

$$\sigma_{\text{obs}} = 1.41 \pm_{0.40}^{0.46} (\text{stat}) \pm_{0.28}^{0.38} (\text{theo}) \pm_{0.13}^{0.13} (\text{sys}) \text{ fb} \quad (9.34)$$

$$= 1.4 \pm_{0.5}^{0.6} \text{ fb}. \quad (9.35)$$

While the agreement with the alternative method of the quadratic subtraction of the uncertainties is good, this method of the fiducial correction is considered more reliable. It is closer to the recommendation to apply a combination of the individual variations for the overall PDF and scale uncertainties only on the final distribution.

The combination of the individual corrections after the fiducial correction introduces the potential for shape differences on the uncertainties considered in the fit compared to the nominal fit without the fiducial correction. These shape differences can result in different best-fit estimates and thus pulls for all other nuisance parameters. These differences explain the minor differences in the nominal value of 0.006 as well as the effect on the experimental systematic uncertainties.

9.5 Interpretation

For a fair comparison of the observed and predicted cross-sections, the precise interpretations of the reported values needs to be taken into account. The predicted cross-section σ_{pred} is determined for the $WZjj$ -EW6 process alone. However, since the $WZjj$ -EW5 interference contributions are not included in the background model of the fit, they are included in the reported observed cross-section. For the MG5_AMC and PYTHIA 8-based simulation for the $WZjj$ -EW5 process a cross-section of $\sigma_{\text{pred}}(WZjj\text{-EW5}) = 0.053 \text{ fb}$ is predicted. This needs to be added to the predicted cross-section for the $WZjj$ -EW6 process $\sigma_{\text{pred}}(WZjj\text{-EW6}) = 1.33 \pm_{0.14}^{0.15} \text{ fb}$. Including this interference effect results in a remarkable agreement of the predicted and observed cross-sections.

As a closing remark on this chapter, a single data event is selected and considered in greater detail. The event is chosen from the five data events with the highest observed **BDTscore**. These events are selected by the Boosted Decision Tree to be among the most signal-like events among the full dataset of 2015 and 2016. The considered event has event number 70104461 and was recorded during run 280319 in 2015. A schematic display of the event kinematic is shown in Figure 9.14.

The **BDTscore** for this event is 0.868. The tagging jets in this event have an invariant mass of $M(jj) = 3041 \text{ GeV}$ and a rapidity separation of $|\Delta Y(jj)| = 4.385$. The leading jet has a rapidity of $Y(j_1) = -1.86$ and the subleading of $Y(j_2) = 2.52$. No additional jet was found in the event and the leptonic final state of the event is $\mu^+\mu^+\mu^-$. The values for each input observable to the BDT are listed in Table 9.5.

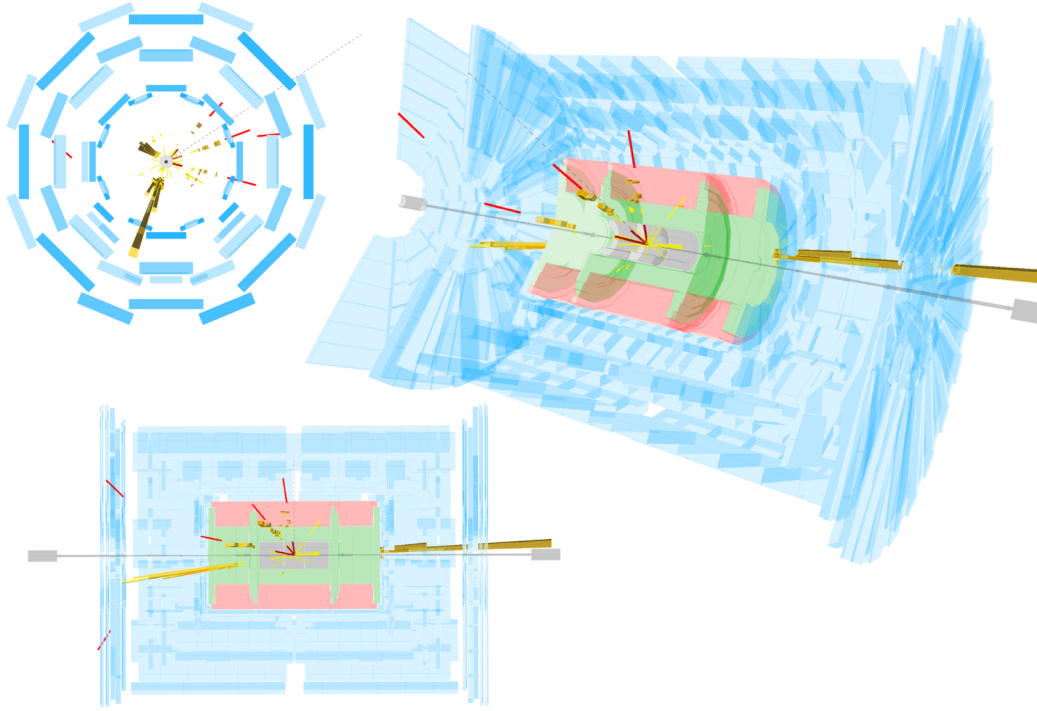


Figure 9.14: Schematic display of event number 70104461 of run 280319 recorded in 2015. The figure on the top (bottom) left corner shows a projection in the transverse (R - z) plane. The figure on the right side shows a 3D cut-away view of the full detector. Detector components are shown with reduced colour saturation and are partially invisible. The muon system is shown in blue, the hadronic calorimeter in red, the electromagnetic calorimeter in green, and the tracking system as well as the beam pipe in grey. Reconstructed tracks with a minimum transverse momentum of 2 GeV are indicated in yellow in general and red when assigned to a muon. Energy deposits in the calorimeter are shown in yellow (electromagnetic calorimeter) or orange (hadronic calorimeter) bars. The radial lengths of these bars indicate the amount of deposited energy. Reconstructed missing transverse momentum is indicated using a grey dashed line. Two high-energetic jets are reconstructed in the forward regions on opposite sides of the detector. Additionally, three muons are identified in the central region of the detector.

Table 9.5: Overview of the kinematic of event number 70104461 of run 280319 recorded in 2015. The leptonic channel, the final **BDTscore**, as well as each input observable of the BDT as well as other observables of interest, are listed. See Chapter 7 for a detailed description of the observables.

Observable	Value
leptonic channel	$\mu^+\mu^+\mu^-$
BDTscore	0.868
M_{jj}	3041 GeV
$\Delta R(j_1, j_2)$	4.76
$p_T^{\text{bal}}(\ell\ell\nu jj)$	0.01
$ \Delta Y(\ell_W, Z) $	0.21
$p_T(j_2)$	290 GeV
$\min \Delta R(j_1, \ell)$	2.23
centrality($\ell\ell\ell$)	1.41
$p_T(j_1)$	393 GeV
$\zeta(\ell\ell\ell)$	0.18
$M_T(W)$	123 GeV
$M(\ell\ell\ell)$	168 GeV
$Y(\ell_W)$	0.53
$p_T(\ell_W)$	277 GeV
$p_T(Z)$	114 GeV
$p_T(\ell_{Z,2})$	50 GeV
$Y(W)$	-0.5
N_{jets}	2
$M(Z)$	92.0 GeV
$M_T(WZ)$	192 GeV
$\Delta Y(j_1, j_2)$	4.39

Chapter 10

Conclusions & Outlook

A fundamental prediction of the gauge structure of the Standard Model of particle physics is the existence of self-interactions among electroweak gauge bosons at triple and quartic vertices. Vector boson scattering (VBS), the scattering of electroweak gauge bosons, is mediated partially via these self-interactions. At the LHC this process is measurable in a final state with two jets and the decay products of two gauge bosons. Additional contributions to the scattering are possible via the exchange of a Higgs boson allowing for further tests of the mechanism of electroweak symmetry breaking by studying vector boson scattering.

This study presents evidence for the existence of the scattering of electroweak gauge bosons in the fully-leptonic final state of the $W^\pm Z/\gamma \rightarrow W^\pm Z$ scattering. A dataset with an integrated luminosity of 36.1 fb^{-1} of pp collisions at a centre-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$ recorded with the ATLAS detector during Run 2 of the LHC in 2015 and 2016 was used.

Two high- p_T tagging jets are required in addition to missing transverse energy and three leptons, where electrons and muons are considered. The contribution of the vector boson scattering process is increased using additional requirements on the invariant mass and rapidity separation of the tagging jet pair.

A comprehensive modelling study of available Monte-Carlo simulations for vector boson scattering processes is presented. Based on this study, the nominal prediction for the signal process is determined using MADGRAPH5_AMC@NLO for a leading-order simulation of the hard process in combination with HERWIG 7 for the generation of parton shower effects, colour reconnection, and hadronization.

The largest contribution from other processes in the signal region is predicted for the $W^\pm Z$ diboson production with extra QCD emissions. Additional contributions are predicted for multi-boson production in different boson or decay channels, several processes involving top-quarks, and experimental backgrounds due to misidentified leptons. The experimental background is estimated in a data-driven approach, specifically for the VBS-enriched signal region. The remaining backgrounds are estimated based on Monte-Carlo simulations, validated and matched to data in dedicated control regions.

The separation of signal and background events is enhanced using machine-learning. Multiple boosted decision trees are optimized and combined. The output score in the signal region is fitted simultaneously with multiple control regions in a binned profile likelihood fit.

With the observed (expected) background-only-rejection significance of $Z_{\text{obs}} = 3.44 \sigma$ ($Z_{\text{exp}} = 3.77 \sigma$), the earlier observation of the $W^\pm Z/\gamma \rightarrow W^\pm Z$ vector boson scattering process in [3] can not be confirmed, but further evidence for the existence of the process was found. The observed cross-section for the signal process including contributions of interferences between the signal and the $W^\pm Z$ background process is

$$\sigma_{\text{obs}}(\text{Sig.}+\text{Int.}) = 1.41 \pm_{0.40}^{0.46} (\text{stat}) \pm_{0.28}^{0.38} (\text{theo}) \pm_{0.13}^{0.13} (\text{sys}) \text{ fb} \quad (10.1)$$

$$= 1.4 \pm_{0.5}^{0.6} \text{ fb}, \quad (10.2)$$

which is in very good agreement with the leading-order SM prediction for the VBS process

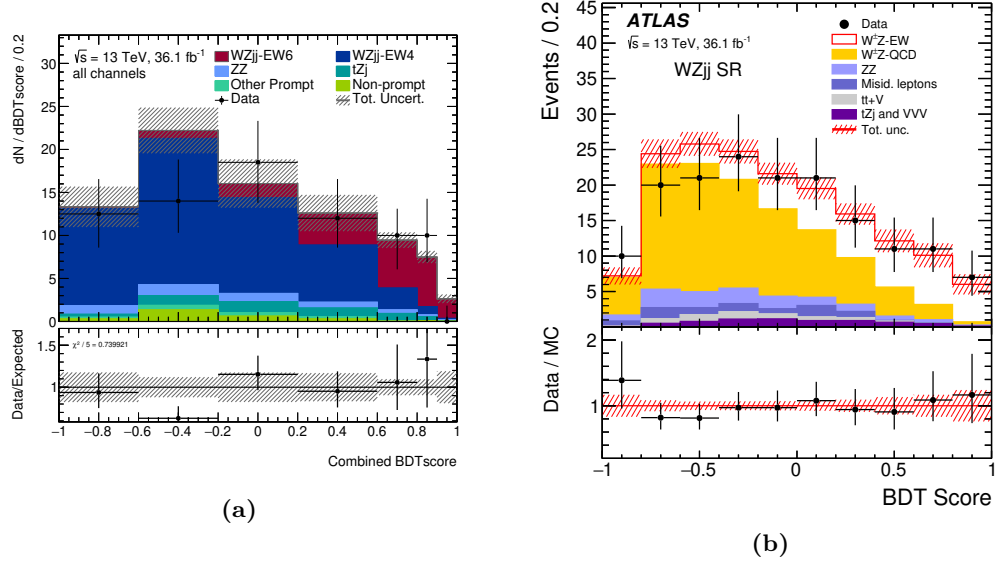


Figure 10.1: Post-fit event yields differentially in the **BDTscore** for the fit to the observed data distribution for this study (left) and the study of the ATLAS collaboration in [3] (right). The hatched bands show the total uncertainties after the fit. The lower inlays show the ratio of the observed distribution to the post-fit predictions.

based on MG5_AMC and HERWIG 7 of

$$\sigma_{\text{pred}}(\text{Sig.}) = 1.33 \pm 0.009 (\text{MC stat.}) \pm 0.021 (\text{PDF}) \pm {}^0_{.0013} (\alpha_S) \quad (10.3)$$

$$\pm {}^{0.11}_{0.10} (\text{scale}) \pm {}^{0.09}_{0.10} (\text{shower}) \text{ fb} \quad (10.4)$$

$$= 1.33 \pm {}^{0.15}_{0.14} \text{ fb}, \quad (10.5)$$

and for the interference of $\sigma_{\text{pred}}(\text{Int.}) = 0.05 \text{ fb}$.

This work presents a re-analysis of the same dataset as used in the first-ever observation of this process in [3]. The main differences are the commonly used definition to select the highest- p_T jets as tagging jets, the addition of the tagging jet rapidity separation requirement of $|\Delta Y(jj)| > 2$ in the signal region and the improved modelling of the signal process and the data-driven estimate of lepton misidentifications. The cross-section observed in [3]¹ deviates from the cross-section reported here, but can be related via the reported predicted cross-sections of the SHERPA event generator. When applying this correction, the cross-section of [3] still exceeds the cross-section found here by about 46 %, which corresponds to a difference of about one standard deviation.

When comparing the post-fit event yield distributions in the signal regions of the two analyses, as shown in Figure 10.1, the observed number in data, especially for regions with high **BDTscore**, differ. These differences could be introduced due to the adjusted phase space definition or due to differences in the independently optimized BDTs.

While the observed significance is reduced compared to the earlier publication in [3], the expected significance was improved from 3.2 to 3.77 standard deviations. Since the expected significance is the correct measure for a comparison of different analyses, the quality of the analysis is considered to be improved. Accounting for the better agreement of expected and observed results in this work suggests an increased observed significance and cross-sections in [3] due to statistical fluctuations. Due to the differences in the event selection and the changes to the applied BDT, such fluctuations don't necessarily manifest in equivalent fluctuations in this analysis. Such events could migrate to lower values of the **BDTscore** or even fail the event selection.

¹The number presented in [3] is the average cross-section per lepton channel, while in this work the cross-section is summed over the four lepton channels. A simple scaling by a factor of 4 is necessary for a direct comparison.

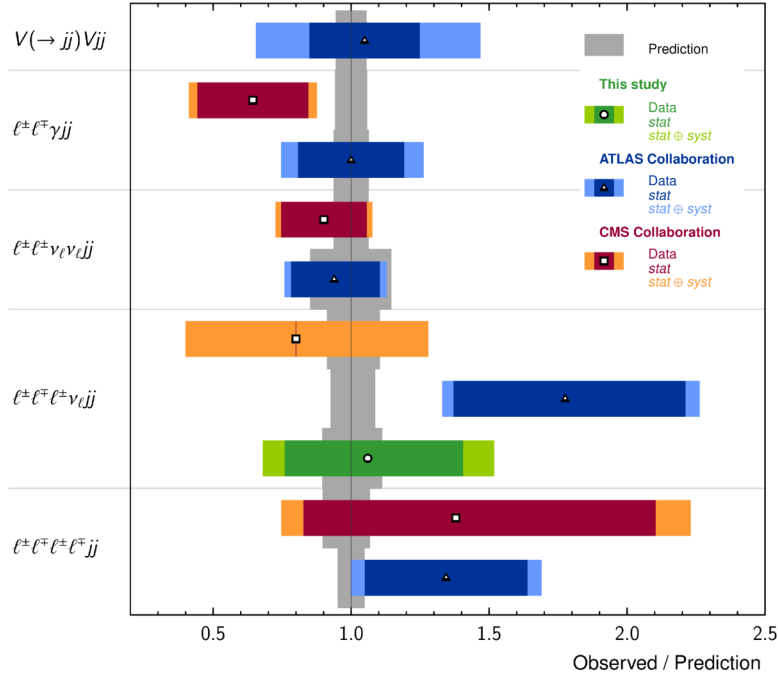


Figure 10.2: Comparison of the fiducial cross-section measurements of vector boson scattering processes in different final states at 13 TeV. For each measurement, the ratio of the observed cross-section to the SM prediction is shown. The result of this study (green) is compared to measurements by the ATLAS (blue) and CMS (red) collaborations.

When comparing the result to the study of this process performed by the CMS collaboration (see [38]), the expected significance obtained here clearly exceeds the expected significance of 2.5σ reported in [38]. The main difference is the adjusted signal extraction procedure, where a maximum likelihood fit is performed in [38] for the integrated yield in each lepton channel. The two analyses use similar setups for the nominal prediction of the signal process² and in both cases good agreement of the observed cross-section with the predictions is found.

A comparison of the relative differences between the predicted and observed cross-sections for all studies of vector boson scattering at $\sqrt{s} = 13$ TeV that are available at the time of writing is shown in Figure 10.2. In general, a good agreement is found between the predicted and observed cross-sections. Statistical uncertainties still limit most measurements. While the uncertainties of the predictions are smaller than uncertainties of the measurements for all VBS analyses, the theoretical modelling of the signal and the main background processes belong to the leading sources of systematic uncertainties and partially limit the available analysis strategies. For instance, the exploitation of the expected kinematic differences of the third jet is limited due to the lack of reliable predictions for most channels, where the third jet is modelled as part of the hard process rather than as part of the parton shower only.

Future analyses will benefit from increased integrated luminosities. Already the available statistic of the full dataset of Run 2 of the LHC of about 140fb^{-1} has the potential to reduce the statistical uncertainties by a factor of 2. With this dataset, confirmation of the observation of the $W^\pm Z/\gamma \rightarrow W^\pm Z$ signal process can be expected.

Without further improvements, however, future analyses of vector boson scattering in this boson channel will be clearly dominated by systematic uncertainties. At the moment, leading

²However, it is not clear whether the simulation used in [38] suffers from the shower mismodelling in PYTHIA 8 discussed in Section 3.3

sources of systematic uncertainties are the modelling of the signal and main background processes.

The estimation of the modelling uncertainty on the $W^\pm Z$ background process can be improved using an alternative simulation on reconstruction-level rather than extrapolating the differences from truth-level. Once such an alternative simulation becomes available, potential ways to reduce the modelling uncertainties can be evaluated.

The modelling of the signal process can be improved by increasing the matrix-element accuracy of the simulation, e.g. by including the first QCD emission in the hard-process calculation using a merged setup. This increased accuracy will reduce the dependence on the modelling of the parton shower, which contains the single leading systematic in this study, the scale uncertainty in the parton shower simulation. This uncertainty is not available in all setups and is commonly not applied, and further technical validation might be necessary. It should be stressed that improved modelling of the signal and main background processes will be crucial for further studies of vector boson scattering.

This analysis also suffered from the limited statistic of the nominal signal model due to the private detector simulation. The training of multi-variate classifiers such as the BDT was not possible with this sample and had to be performed using an alternative simulation with known issues. Future analyses will benefit considering the nominal signal model in the training of multi-variate classifiers.

Further improvements can potentially be achieved by training multi-variate classifiers not only on the nominal simulations but to include alternative signal and background simulations as well. This approach could improve the classification in alternative simulations and consequently reduce the uncertainties derived from the difference from the nominal and alternative simulations. This reduction seems most promising for the modelling uncertainties of the signal and main background processes. However, it can, in principle, also be extended to variations of significant experimental uncertainties, e.g. the uncertainty of the modelling of pileup interactions.

The interpretation of the final result might be simplified by including reconstruction-level simulations for the interference contributions in the final fit. While it can be debated whether these contributions should be treated as background or as signal due to the dependence on the signal strength, the availability of such simulations would increase flexibility and would allow for proper treatment in the full fit.

With observations of the $\ell^\pm \ell^\pm \nu_\ell \nu_\ell jj$, $\ell^\pm \ell^\mp \ell^\pm \nu_\ell jj$, and $\ell^\pm \ell^\mp \ell^\pm \ell^\mp jj$ final states in the last years, the study of vector-boson scattering processes at the LHC has been very fruitful. The increased available amount of data recorded in full Run 2 and Run 3 will improve accuracy and observations in other channels can be expected. A combination with improved modelling will allow for more detailed studies of the kinematics in terms of unfolded distributions as well as the measurement of further properties such as the boson polarisation.

The door to the physics of vector boson scattering has just been opened, but many more studies are still to come.

Appendix A

Conventions and Notations

A.1 Natural Units

In order to simplify the notation, it is helpful to use a scheme for units where common variables have simple representations. The customary convention is called “natural units”. In that scheme

$$\hbar = c = 1. \quad (\text{A.1})$$

The unit for energies is not set by this and is commonly written in electronvolt eV. Also masses and momenta can be expressed in this unit.

A.2 Einstein’s Summation Convention

The goal of this convention is to improve notational brevity by removing the \sum operator in special situations. The convention is to skip the \sum operator, when an otherwise undefined index occurs in the superscript and the subscript in a term.

$$\sum_{\mu=0}^3 x_{\mu} x^{\mu} =: x_{\mu} x^{\mu}. \quad (\text{A.2})$$

The allowed range of the index variable can be deduced from the used letter. For indices in the Latin alphabet, the convention is to sum over 1,2, and 3., while indices in the Greek alphabet are summed from 0 to 3.

A.3 Relativistic and Dirac Notation

We use the usual metric tensor

$$g_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}, \quad (\text{A.3})$$

resulting in the relativistic derivative ∂^{μ} of

$$\partial^{\mu} = \left(\frac{\partial}{\partial t}; -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z} \right) \quad (\text{A.4})$$

Gamma matrices γ^{μ} are used similar to Weyl basis, with adjusted sign convention

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix}. \quad (\text{A.5})$$

where σ_i are the Pauli spin matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (\text{A.6})$$

A.4 Group Theory

Generators of $SU(2)$

$$L_i = 2 * \sigma_i. \quad (\text{A.7})$$

Structure constants of $SU(2)$

$$\epsilon_{123} = 1, \text{ anti-symmetric} \quad (\text{A.8})$$

Generators of $SU(3)$ (Gell-Mann matrices)

$$\lambda_1 = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} & -i \\ i & \end{pmatrix}, \quad (\text{A.9})$$

$$\lambda_3 = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}, \quad \lambda_4 = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \quad (\text{A.10})$$

$$\lambda_5 = \begin{pmatrix} & -i \\ i & \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \quad (\text{A.11})$$

$$\lambda_7 = \begin{pmatrix} & -i \\ & i \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & & \\ & 1 & \\ & & -2 \end{pmatrix}. \quad (\text{A.12})$$

Structure constants of $SU(3)$

$$f_{123} = 1; \quad (\text{A.13})$$

$$f_{458} = f_{678} = \sqrt{3}/2; \quad (\text{A.14})$$

$$f_{147} = f_{516} = f_{246} = f_{257} = f_{345} = f_{637} = 1/2; \text{ anti-symmetric} \quad (\text{A.15})$$

Appendix B

Auxiliary Information on Event Generation

B.1 Configuration Files and Notes on Private Productions

This section details the configuration files used for each of the privately produced simulations following the nomenclature as introduced in Table 3.1. Most configurations are available or are created using scripts mostly developed by the author available at [179].

B.1.1 EW6 MG5Py8Fix

For the generation of the hard process, MADGRAPH5_AMC@NLO was used via ATLAS-internal interfaces and a simulation of the parton shower using PYTHIA 8 is applied directly steered from the same configuration file. In a first step before the actual generation, gridpacks are produced using a private installation of MADGRAPH5_AMC@NLO in version 2.6.5, overwriting the default of the used release `AtlasProduction,19.2.5.37`. Subsequently, the event generation was performed using release `AtlasProduction,19.2.5.37` and the default MG5_AMC and PYTHIA 8 versions of this release.

The configuration file is printed below. As indicated in lines 13-46, the full process is split in four subprocesses according to the lepton flavour and charge combination.

```
1 from MadGraphControl.MadGraphUtils import *
2
3 evgenConfig.generators = ["MadGraph", "Pythia8", "EvtGen"]
4
5 evgenConfig.keywords = ['SM', 'diboson', 'WZ', 'electroweak', '3lepton', 'VBS']
6 evgenConfig.contact = ['carsten.bittrich@cern.ch']
7
8 gridpack_dir = 'madevent/'
9
10 # -----
11 # Process type based on runNumber:
12 # -----
13 if runArgs.runNumber == 999991:
14     runName = 'lv1ljEW6_OFMinus'
15     description = 'MadGraph_lv1lj_EW6_OFMinus_masslessLeptons'
16     mgproc = ""
17 generate p p > mu+ mu- e- ve- j j QCD=0 @0
18 add process p p > ta+ ta- e- ve- j j QCD=0 @0
19 add process p p > e+ e- mu- vm- j j QCD=0 @0
20 add process p p > ta+ ta- mu- vm- j j QCD=0 @0
21 add process p p > e+ e- ta- vt- j j QCD=0 @0
22 add process p p > mu+ mu- ta- vt- j j QCD=0 @0""
23 elif runArgs.runNumber == 999992:
```

```
24     runName = 'lv1ljEW6_OFPlus'
25     description = 'MadGraph_lv1lj_EW6_OFPlus_masslessLeptons'
26     mgproc = ""
27     generate p p > mu+ mu- e+ ve j j QCD=0 @0
28     add process p p > ta+ ta- e+ ve j j QCD=0 @0
29     add process p p > e+ e- mu+ vm j j QCD=0 @0
30     add process p p > ta+ ta- mu+ vm j j QCD=0 @0
31     add process p p > e+ e- ta+ vt j j QCD=0 @0
32     add process p p > mu+ mu- ta+ vt j j QCD=0 @0""
33     elif runArgs.runNumber == 999993:
34         runName = 'lv1ljEW6_SFMinus'
35         description = 'MadGraph_lv1lj_EW6_SFMinus_masslessLeptons'
36         mgproc = ""
37         generate p p > e+ e- e- ve~ j j QCD=0 @0
38         add process p p > mu+ mu- mu- vm~ j j QCD=0 @0
39         add process p p > ta+ ta- ta- vt~ j j QCD=0 @0""
40     elif runArgs.runNumber == 999994:
41         runName = 'lv1ljEW6_SFPlus'
42         description = 'MadGraph_lv1lj_EW6_SFPlus_masslessLeptons'
43         mgproc = ""
44         generate p p > e+ e- e+ ve j j QCD=0 @0
45         add process p p > mu+ mu- mu+ vm j j QCD=0 @0
46         add process p p > ta+ ta- ta+ vt j j QCD=0 @0""
47         pass
48     else:
49         raise RuntimeError(
50             "runNumber %i not recognised in these jobOptions." % runArgs.runNumber)
51
52     evgenConfig.description = description
53
54     # -----
55     # write MG5 Proc card
56     # -----
57     fcard = open('proc_card_mg5.dat', 'w')
58     fcard.write("""
59     import model sm-no_masses
60     # massless b and massless taus
61     define p = g u c d s u~ c~ d~ s~ b b~
62     define j = g u c d s u~ c~ d~ s~ b b~
63     """ + mgproc + """)
64     output -f
65     """)
66     fcard.close()
67
68     # -----
69     # Random Seed
70     # -----
71     randomSeed = 0
72     if hasattr(runArgs, 'randomSeed'):
73         randomSeed = runArgs.randomSeed
74
75     # -----
76     # Beam energy
77     # -----
78     if hasattr(runArgs, 'ecmEnergy'):
79         beamEnergy = int(runArgs.ecmEnergy) / 2.
80     else:
81         raise RuntimeError("No center of mass energy found.")
82
83     # -----
84     # Number of Events
85     # -----
86     safefactor = 1.1
87     if hasattr(runArgs, 'maxEvents') and runArgs.maxEvents > 0:
88         nevents = int(int(runArgs.maxEvents) * safefactor)
89     else:
90         nevents = int(5000 * safefactor)
91
92     extras = {
93         'pdlabel': "'lhpdf'",
```

```

94     'lhaid': 260000,
95     'ptl': 4.0,
96     'ptj': 15.0,
97     'ptb': 15.0,
98     'ptheavy': 2.0,
99     'mml1': 4.0,
100    'drll': 0.2,
101    'drjl': 0.2,
102    'drjj': 0.4,
103    'dral': 0.1,
104    'etaj': 5.5,
105    'etal': 5.0,
106    'maxjetflavor': 5,
107    'asrwtflavor': 5,
108    'auto_ptj_mjj': False,
109    'cut_decays': True,
110    'use_syst': True,
111    'event_norm': 'sum',
112    'systematics_program': 'systematics',
113    'systematics_arguments': "['-mur=0.5,1,2', '-muf=0.5,1,2',
114    '-pdf=errorset,NNPDF30_nlo_as_0119@0,NNPDF30_nlo_as_0117@0,CT14nlo,
115    MMHT2014nlo68clas118,PDF4LHC15_nlo_30_pdfas']"}
116
117 process_dir = new_process(grid_pack=gridpack_dir)
118
119 build_run_card(run_card_old=get_default_runcard(proc_dir=process_dir),
120               run_card_new='run_card.dat', xqcut=0,
121               nevents=nevents, rand_seed=runArgs.randomSeed,
122               beamEnergy=beamEnergy)
123 modify_run_card(run_card='run_card.dat',
124                run_card_backup='run_card_backup.dat',
125                settings=extras)
126 print_cards()
127
128 generate(required_accuracy=0.001, run_card_loc='run_card.dat',
129          param_card_loc=None, mode=0, njobs=1, cluster_type=None,
130          cluster_queue=None, proc_dir=process_dir, run_name=runName,
131          grid_pack=True, gridpack_compile=False, gridpack_dir=gridpack_dir,
132          nevents=nevents, random_seed=runArgs.randomSeed)
133
134 # replaced output_suffix+'.00001.events.tar.gz' with runArgs.outputTXTFile
135 arrange_output(run_name=runName, proc_dir=process_dir,
136               outputDS=runArgs.outputTXTFile, lhe_version=3,
137               saveProcDir=True)
138
139
140 runArgs.inputGeneratorFile = runArgs.outputTXTFile
141
142 include("MC15JobOptions/Pythia8_A14_NNPDF23L0_EvtGen_Common.py")
143 # fix for VBS processes (requires version>=8.230)
144 genSeq.Pythia8.Commands += ["SpaceShower:dipoleRecoil=on"]
145
146 include("MC15JobOptions/Pythia8_MadGraph.py")

```

An important caveat in the process definition is the explicit list of lepton flavour combinations in combination with the explicit combination into one process group using the @0 at the end of each process definition. This approach reduces the time for finding all possible diagrams from up to 20 – 30 *h* to very few hours. This is especially important for the merged setup for the EW4 process described in Section B.1.8. When using particle groups in the process definition, MG5_AMC does try to find diagrams for each possible combination, rather than e.g. grouping processes differing in the order of final state particles. Due to the large number of combinations when including three final-state quarks as well as three leptons with possible flavour combinations, the computing time is unnecessarily long.

The separation of the full process in four groups according to the leptons charge and flavour combinations benefits the optimization of the phase space integration in general, but is also necessary in MG5_AMC, when applying a cut on the invariant mass of lepton pairs

with same flavour and opposite charge $M(\ell\ell)$ as applied here for internal technical reasons. A common recommendation to avoid this issue is to turn off the grouping of subprocess for massless particles. This implies that contributions, where a e^+e^- pair is replaced by a $\mu^+\mu^-$ pair are evaluated separately, although equivalent assuming massless leptons. Since this also applies for light quarks, the computing time is again increased unnecessarily. The separation of the process in 4 parts allows to apply a cut on $M(\ell\ell)$, while using the grouping of subprocesses.

B.1.2 EW6 MG5H7

For this sample, the events at the level of the hard process generated for the EW6 MG5Py8Fix sample (see Section B.1.1) are read in and processed with an independent shower implementation. The showering was applied using the ATLAS-internal interface to HERWIG 7 using release MCProd,20.7.9.9.26. The config file is printed below.

```
1 evgenConfig.description = "Herwig7 parton shower example"
2 evgenConfig.keywords = ["electroweak"]
3 evgenConfig.inputfilecheck = "MG5H7"
4 evgenConfig.generators = ["Herwig7" , "Lhef" ]
5 ## Initialise Herwig7 for LHEF showering
6 include("MC15JobOptions/Herwig7_LHEF.py")
7
8 ## Provide config information
9 evgenConfig.generators += ["aMcAtNlo"]
10 evgenConfig.tune = "MMHT2014"
11 evgenConfig.contact = ["Carsten Bittrich (carsten.bittrich@cern.ch)"]
12
13 ## Configure Herwig7
14 Herwig7Config.me_pdf_commands(order="NLO", name="NNPDF30_nlo_as_0118")
15 Herwig7Config.tune_commands()
16 Herwig7Config.lhef_mg5amc_commands(lhe_filename=runArgs.inputGeneratorFile,
17                                   me_pdf_order="NLO")
18 Herwig7Config.add_commands("")
19
20 cd /Herwig/Shower
21 do ShowerHandler:AddVariation showerscale_MUR0.5_MUF0.5 0.5 0.5 All
22 do ShowerHandler:AddVariation showerscale_MUR0.5_MUF1 0.5 1 All
23 do ShowerHandler:AddVariation showerscale_MUR0.5_MUF2 0.5 2 All
24
25 do ShowerHandler:AddVariation showerscale_MUR1_MUF0.5 1 0.5 All
26 do ShowerHandler:AddVariation showerscale_MUR1_MUF2 1 2 All
27
28 do ShowerHandler:AddVariation showerscale_MUR2_MUF0.5 2 0.5 All
29 do ShowerHandler:AddVariation showerscale_MUR2_MUF1 2 1 All
30 do ShowerHandler:AddVariation showerscale_MUR2_MUF2 2 2 All
31 set SplittingGenerator:Detuning 2.0
32
33 set ShowerHandler:SpinCorrelations Yes
34 """)
35
36 ## run generator
37 Herwig7Config.run()
```

B.1.3 EW6 WhPy8

Events for this sample at the level of the hard process were generated using a private installation of WHIZARD in version 2.6.4. The configuration file is printed below.

```
1 #####
2 # carsten.bittrich@mailbox.tu-dresden.de
3 # lllnujj-EW with aQGC
4 #####
5
6 model = SM # SSC_2 for fs0, SSC for a4
```



```

7  alphas      = 0      # set alpha_s to zero -> WZjj-EW
8  #a4         = 0.1
9  #a5         = 0
10 #fs0        = 0
11 #fs1        = 0
12 #fkm        = 1      # set 0 to switch off K-matrix unitarization
13 alias qf     = u:d:s:c:U:D:S:C:g
14 alias lightjet = u:d:s:c:U:D:S:C:g
15 alias lep    = e1:E1:e2:E2:e3:E3
16 alias neu    = n1:N1:n2:N2:n3:N3
17 alias alleptons = lep:neu
18 alias bosons  = Wm:Wp:Z
19 alias elflavour = e1:E1
20 alias elNflavour = n1:N1
21
22 alias charlep = e1:E1:e2:E2
23 alias charlepP = E1:E2
24 alias charlepM = e1:e2
25 alias neu     = n1:N1:n2:N2
26 alias mytau   = e3:E3
27 alias mytaunu = n3:N3
28
29 process wpz1 = qf, qf => qf, qf, charlepM, charlepP, charlepP, neu
30 process wpz2 = qf, qf => qf, qf, e3, E3, E3, mytaunu
31 process wpz3 = qf, qf => qf, qf, e3, E3, charlepP, neu
32 process wpz4 = qf, qf => qf, qf, charlepM, charlepP, E3, mytaunu
33
34 process wmz1 = qf, qf => qf, qf, charlepM, charlepP, charlepM, neu
35 process wmz2 = qf, qf => qf, qf, e3, E3, e3, mytaunu
36 process wmz3 = qf, qf => qf, qf, e3, E3, charlepM, neu
37 process wmz4 = qf, qf => qf, qf, charlepM, charlepP, e3, mytaunu
38
39 # ATLAS settings
40 mtop = 172.5
41 mW   = 80.399
42 mZ   = 91.1876
43 wZ   = 2.4952
44 wW   = 2.085
45 ms   = 0
46 mc   = 0
47 mH   = 126 GeV
48 wH   = 0.00418 GeV
49 me   = 0
50 mmu  = 0
51 mtau = 1.77705
52
53 scale = eval M [collect[alleptons]]
54 phs_keep_nonresonant = true
55
56 #use the existing phs file
57 #?check_phs_file = false
58
59 # VBS cuts fiducial volume
60 cuts = all Dist > 0.4 [qf,qf] and all Pt > 15 GeV [qf] and all -5.0 < Eta < 5.0
61 [qf] and all Pt > 5 GeV [extract index 2 [sort by -Pt [lep]]] and all -5.0
62 < Eta < 5.0 [lep] and all Dist > 0.3 [lep,lep]
63
64 compile
65
66 $integration_method = "vamp2"
67 $rng_method = "rng_stream"
68
69 # LHC
70 sqrts = 13 TeV
71 $lhpdf_file = "CT10"
72 beams = p, p => lhpdf
73
74 accuracy_goal = 5
75 integrate (wmz1,wmz2,wmz3,wmz4,wpz1,wpz2,wpz3,wpz4) { iterations = 15:600000, 5:600000 }
76 show(results)

```

```

77
78 n_events = 520
79 sample_format = lhef
80 [?]unweighted = true
81 simulate (wmz1,wmz2,wmz3,wmz4,wpz1,wpz2,wpz3,wpz4)
82 { checkpoint = 100 [?]keep_beams = false [?]polarized_events = true}

```

The simulation of the parton shower effects, PYTHIA 8 was used using the ATLAS-internal interface in release `AtlasProduction,19.2.5.37`. The configuration file is printed below.

```

1 evgenConfig.description = "MGPy8EG+Pythia8 production JO example with the A14 NNPDF23L0 tune"
2 evgenConfig.keywords = ["electroweak"]
3 evgenConfig.inputfilecheck = "MGPy8EG"
4
5 include("common/Pythia8_A14_NNPDF23L0_EvtGen_Common.py")
6 genSeq.Pythia8.Commands += ["SpaceShower:dipoleRecoil=on"]
7 include("common/Pythia8_LHEF.py")
8 include("common/Pythia8_ShowerWeights.py")
9
10 evgenConfig.minevents = 1000
11
12 evgenConfig.generators = ["Pythia8" , "Lhef" ]

```

B.1.4 EW6 VBFNLOpy8

Events for this sample at the level of the hard process were generated using a private installation of VBFNLO in version 3.0.0 beta 5. The general configuration file for process 220 (W^+Zjj -EW6) is printed below. The second subprocess W^-Zjj -EW6 is generated with the equivalent configuration file, where only the `PROCESS` is set to 230.

For the s -channel contributions, the additional processes 401, 402, 412, and 412 are used, representing the different combinations of the charge of the leptonically decaying W^\pm boson as well as whether the jet pair has to be compatible with the decays of a W^\pm or Z boson. For these processes, also the dynamical scale choices `ID_MUF` and `ID_MUR` have to be adjusted to 3.

```

1 [!] Main input file for vbfnl0
2
3 [!] General parameters of the calculation
4 [!]-----
5 PROCESS = 220 [!] Identifier for process
6 [!] -- WZ --
7 [!] 220 = WpZjj-EW
8 [!] 230 = WmZjj-EW
9 [!] 401 = W(qq)WmZ (semi-leptonic triboson)
10 [!] 402 = W(qq)WpZ (semi-leptonic triboson)
11 [!] 412 = Z(qq)WpZ (semi-leptonic triboson)
12 [!] 422 = Z(qq)WmZ (semi-leptonic triboson)
13 [!] -- WWss --
14 [!] 250 = WpWpjj-EW
15 [!] 260 = WmWmjj-EW
16 [!] 432 = W(qq)WpWp (semi-leptonic triboson)
17 [!] 441 = W(pp)WmWm (semi-leptonic triboson)
18 [!] --
19 LOPROCESS_PLUS_JET = false [!] switch: L0 process with 1 additional jet
20 LEPTONS = 99 99 99 99 [!] final state leptons
21 DECAY_QUARKS = 93 [!] final state quarks for semileptonic decays
22
23 [!]
24 [!] lepton and quark numbering according to MC particle numbering scheme
25 [!] particles are given positive numbers, antiparticles negative numbers

```

```

26 ! decay products have to be grouped according to their parent boson
27 !
28 ! e-   ve   mu-  vm   ta-  vt   |   d   u   s   c   b
29 ! 11   12   13   14   15   16   |   1   2   3   4   5
30 !
31 ! DECAY_QUARKS = 93 includes all possible combinations of u, u~, d, d~, s, s~, c, c~
32 ! DECAY_QUARKS = 94 includes also b-quarks (u, u~, d, d~, s, s~, c, c~, b, b~)
33 ! LEPTONS      = 98 includes all possible combinations of first and second generation leptons
34 ! LEPTONS      = 99 includes all possible combinations of all three lepton generations
35 !
36
37 LO_ITERATIONS      = 4          ! number of iterations for LO calculation
38 NLO_ITERATIONS     = 4          ! number of iterations for real-emissions calc.
39 LO_POINTS          = 27         ! number of points for LO calculation (= 2^..)
40 NLO_POINTS         = 20         ! number of points for real-emissions calc. (= 2^..)
41 LO_GRID            = "grid2_1" "grid2_2" "grid2_3" "grid2_4" "grid2_5"
42 NLO_GRID           = "grid3_1" "grid3_2" "grid3_3" "grid3_4" "grid3_5"
43 PHTN_GRID          = "grid4_1" "grid4_2" "grid4_3" "grid4_4" "grid4_5"
44 FLOOP_GRID         = "grid5_1" "grid5_2" "grid5_3" "grid5_4" "grid5_5"
45
46 NLO_SWITCH          = false     ! switch: nlo/lo calculation
47 EWCOR_SWITCH        = false     ! Whether electroweak corrections are included
48 FERMIONLOOP         = 3         ! Contribution of gluon-induced fermionic loops for
49                       ! diboson processes
50                       ! 0: none
51                       ! 1: only box diagrams
52                       ! 2: only Higgs resonance
53                       ! 3: both contributions (default)
54 NLO_SEMILEP_DECAY   = 0         ! controls whether QCD NLO effects in the semileptonic
55                       ! decays of vector bosons are included:
56                       ! 0: leading order hadronic decay of the vector boson
57                       ! 1: include approximate next-to leading order QCD effects
58
59 ECM                 = 13000d0   ! collider center-of-mass energy
60 BEAM1               = 1         ! type of beam 1 (1=proton, -1 = antiproton)
61 BEAM2               = 1         ! type of beam 2 (1=proton, -1 = antiproton)
62
63 ID_MUF              = 1         ! ID for factorization scale
64 ID_MUR              = 1         ! ID for renormalization scale
65 MUF_USER            = 160.798d0 ! user defined factorization scale, if MUF is set to 0
66 MUR_USER            = 160.798d0 ! user defined renormalization scale, if MUR is set to 0
67 XIF                 = 1d0       ! scale factor xi for mu_F (not mu^2!!)
68 XIR                 = 1d0       ! scale factor xi for mu_R
69
70
71 ! Physics parameters
72 !-----
73 HMASS               = 126.0d0   ! Higgs mass
74 HTYPE               = 0         ! Type of Higgs produced:
75                       ! 0    SM Higgs with mass HMASS
76                       ! 1    light cp-even type higgs h0
77                       ! 2    heavy cp-even type higgs HH
78                       ! 3    lightest cp-odd type higgs A0
79                       ! SUSY parameters for 1-3 are set in susy.dat.
80                       ! For these options, if input 'MODEL' is set
81                       ! to 1 (SM), calculation will run in the SM
82                       ! for a Higgs with equivalent mass to that
83                       ! chosen in the MSSM.
84 HSCHEME              = 0         ! Distribution of VBF-H:
85                       ! 0    on-shell
86                       ! 1    Breit-Wigner
    
```

B Auxiliary Information on Event Generation

```

87                                     ! 2 Passarino CPS scheme
88 MODEL = 1                         ! model: 1 for SM, 2 for MSSM, 3 for TwoHiggs
89 EWScheme = 3                      ! Choose scheme for electroweak parameters (1,2,3,4,5,6)
90 DEL_ALFA = 0.059047686d0         ! Delta alfa
91 ANOM_CPL = true                  ! Anomalous couplings
92 KK_MOD = false                   ! Warped Higgsless Model
93 SPIN2 = false                    ! Spin-2 model
94 EW_APPROX = 5                    ! Approximation used when calculating electroweak
95                                     ! 0 no approximations used: full ew corrections included
96                                     ! 1 only t/b (and stop/sbottom in the MSSM) corns are included
97                                     ! 2 only fermion (and sfermion in the MSSM) corns are included
98                                     ! 3 MSSM option: all SM-type and sfermion corns are included
99                                     ! 4 MSSM option: all MSSM corns to hVV vertex, and all SM-type
100                                     ! and sfermion corrections elsewhere
101                                     ! 5 MSSM option: all MSSM corns to hVV and qqV vertices and
102                                     ! VV self energy, and all SM-type boxes and pentagons
103 ! Settings for the TwoHiggs model
104 H2MASS = 600d0                   ! mass of second Higgs boson
105 H2WIDTH = -999d0                 ! width of second Higgs boson
106 SIN2BA = 1d0                    ! squared strength modifier for H1-V-V coupling
107 COS2BA = -999d0                 ! squared strength modifier for H2-V-V coupling
108 ! NOTE: the following will be overwritten if a SLHA file is used
109     HWIDTH = -999d0              ! Higgs width (set to -999d0 for internal calculation, -998d0
110     → for HXSWG value)
111 TOPMASS = 172.4d0                ! Top mass
112 BOTTOMMASS = 4.855d0             ! Bottom Pole mass
113 CHARMMASS = 1.65d0              ! Charm Pole mass
114 TAU_MASS = 1.77684d0            ! Tau mass
115 FERMI_CONST = 1.16637d-5        ! Fermi Constant
116 INVALFA = 128.944341122D0       ! 1/fine-structure constant
117 SIN2W = 0.222646d0             ! Weak mixing angle
118 WMASS = 80.398d0               ! W mass
119 ZMASS = 91.1876d0              ! Z mass
120
121 ! Parameters for the LHA event output
122 ! -----
123 LHA_SWITCH = true                ! Les Houches interface only for LO calculation
124 LHA_FILE = event.lhe             ! Name of Les Houches output file
125 HEPMC_SWITCH = false            ! HepMC interface only for LO calculation
126 HEPMC_FILE = event.hepmc        ! Name of HepMC output file
127 UNWEIGHTING_SWITCH = true       ! Unweighted/weighted (T/F) events in event files
128 DESIRED_EVENT_COUNT = 10000     ! For unweighted events: Number of requested
129                                     ! unweighted events in file.
130                                     ! DESIRED_EVENT_COUNT = 10000 : write out events
131                                     ! found during normal integration
132                                     ! DESIRED_EVENT_COUNT = N > 0 : write out N events.
133 PARTIAL_UNWEIGHTING = false      ! If DESIRED_EVENT_COUNT > 0: Allow for few events with
134                                     ! weight > 1. Useful when the desired number of fully
135                                     ! unweighted events can't be reached.
136 TAUMASS = false                 ! Include mass of the tau lepton(s) in the event file
137
138
139 ! PDF set parameters
140 ! -----
141 PDF_SWITCH = 1                   ! which pdfs to use:
142                                     ! 0 = hard-wired cteq6 (LO) and CT10 (NLO)
143                                     ! 1 = lhpdf (use LO_PDFNAME etc to specify)
144                                     ! 2 = hard-wired mrst2004qed
145                                     ! 3 = hard-wired MSTW2008

```

```

146 ! choose pdfset and pdfmember here. Look at the LHAPDF manual for details.
147 LO_PDFNAME    = NNPDF30_nnlo_as_0118.LHgrid ! cteq6ll.LHpdf
148 NLO_PDFNAME    = CT10.LHgrid
149 LO_PDFMEMBER    = 0
150 NLO_PDFMEMBER    = 0
151
152
153 ! Parameters for output and histogram creation
154 !-----
155 XSECFIELD      = xsection      ! name of output-file (+ '.out')
156 ROOT           = false        ! create root-file?
157 TOP            = false        ! create top-drawer file?
158 GNU           = true          ! create gnu-plot script file?
159 DATA          = true          ! create data file?
160 REPLACE        = true          ! replace output files?
161 ROOTFILE       = histograms    ! name of root-file (+ '.root')
162 TOPFILE        = histograms    ! name of top-drawer file (+ '.top')
163 GNUFILE        = histograms    ! name of gnuplot file (+ '.gp')
164 DATAFILE      = histograms    ! name of data directory (+ '.dir')
165                ! more histogram settings in histograms.dat

```

Additionally, a second file is used in VBFNLO to define the generator level phase space. The configuration used in all VBFNLO runs in this work is printed below.

```

1 ! input file for the cut parameters
2
3 ! Jet cuts
4 !-----
5 RJJ_MIN      = 0.3d0      ! min jet-jet R separation
6 !Y_P_MAX     = 4.50      ! max pseudorapidity for partons
7 PGENTKTJET   = -1.0d0    ! exponent of generalised k_T algorithm
8 PT_JET_MIN   = 15.00     ! min jet pT
9 Y_JET_MAX    = 4.7       ! max jet rapidity
10
11 ! Lepton cuts (only applied to charged leptons)
12 !-----
13 Y_L_MAX      = 2.6       ! max lepton rapidity
14 PT_L_MIN     = 5.00      ! min lepton pT
15 MLL_MIN      = 15.00     ! min. m_ll for any comb. of charged leptons (MLL_OSONLY=false)
16 !MLL_MAX     = 1d20      ! max. m_l+l- for any comb. of charged leptons
17 MLL_OSONLY   = true      ! cut only on opposite sign lepton combinations
18 RLL_MIN      = 0.2       ! min lepton-lepton R separation
19 !RLL_MAX     = 1d20      ! max lepton-lepton R separation
20
21 ! Photon cuts
22 !-----
23 !Y_G_MAX     = 1d20      ! max pseudorapidity for photons
24 !PT_G_MIN    = 0d0       ! min transverse momentum for photons
25 !RGG_MIN     = 0.6d0     ! min photon-photon R separation
26 !RGG_MAX     = 50.0d0    ! max photon-photon R separation
27 !PHISOLCUT   = 0.7d0     ! photon isolation cut
28 !EFISOLCUT   = 1d0       ! efficiency of photon isolation cut
29
30 ! Additional cuts
31 !-----
32 RJL_MIN      = 0.3       ! min jet-lepton R separation
33 !RJG_MIN     = 0.3d0     ! min jet-photon R separation
34 !RLG_MIN     = 0.6d0     ! min lepton-photon R separation
35
36 !MLG_MIN     = 0.0d0     ! min. m_lg for any comb. of charged leptons and photons
37 !MLG_MAX     = 1.d20     ! max. m_lg for any comb. of charged leptons and photons

```

B Auxiliary Information on Event Generation

```

38
39 !PTMISS_MIN = 40          ! minimal missing transverse momentum
40                          ! (pt of neg. sum of 4-momenta of all visible jets, leptons, photons)
41
42 ! Vector boson fusion cuts (only applied to tagging jets in VBF processes)
43 ! -----
44 !ETAJJ_MIN = 2.4          ! jet-jet rapidity separation
45 YSIGN      = false        ! jets #1 and #2 must have opposite sign rapidity
46 LRAPIDGAP  = false        ! leptons fall inside rapidity gap
47 DELY_JL    = 0.0d0        ! min y-dist of leptons from tagging jets
48 GRAPIDGAP  = false        ! photons fall inside rapidity gap
49 DELY_JG    = 0.0d0        ! min y-dist of photons from tagging jets
50
51 MDIJ_MIN   = 150          ! dijet min mass cut on tag jet
52 !MDIJ_MAX   = 1d20        ! dijet max mass cut on tag jet
53
54 ! Jet Veto
55 ! -----
56 JVETO      = false        ! veto jet cuts
57 DELY_JVETO = 0.0d0        ! min veto-tag y-dist
58 YMAX_VETO  = 4.5d0        ! max |y| for veto jet
59 PTMIN_VETO = 50.0d0       ! min pT for veto jet
60
61 ! Additional cuts for semileptonic decays (only applied for semileptonic decay processes)
62 ! -----
63 DEF_TAGJET  = 1           ! definition of two tagging jets in VBF:
64                      ! * 1 : largest pT
65                      ! * 2 : largest pT outside central region with opposite sign rapidity
66                      ! * 3 : largest separation in rapidity
67 ETA_CENTRAL = 0.0d0       ! max rapidity value for "central region" if DEF_TAGJET = 2.
68                      ! Otherwise "central" stands for
69                      ! "between tagging jets".
70 PTMIN_TAG_1 = 15d0        ! pt_min for harder tagging jet
71 PTMIN_TAG_2 = 15d0        ! pt_min for softer tagging jet
72 HARD_CENTRAL = false      ! require hard jet in central region which is not a tagging jet
73 PTMIN_CENTRAL = 20d0      ! minimal pt for extra hard jet in central region
74 RECONST_HAD_V = 0        ! apply mass reconstruction cut for the hadronically decaying
75                      ! vector boson (no tagging jets allowed):
76                      ! * 0 : none
77                      ! * 1 : using two-jet invariant mass
78                      ! * 2 : using invariant mass of one jet
79                      ! (subjet-analysis, only needed when SINGLE_DECAYJET > 0)
80                      ! * 3 : using one- or two-jet invariant mass
81 V_MASS_RANGE = 20d0       ! Mass range for RECONST_HAD_V: |m_V - m_reconst| < V_MASS_RANGE
82
83 SINGLE_DECAYJET = 0       ! allow phase space points with less than two jets from the
84                      ! hadronic decay:
85                      ! * 0 : two additional jets are required for the semileptonic
86                      ! case w.r.t. fully leptonic process
87                      ! * 1 : a single jet is allowed if both quarks from the
88                      ! hadronic decay are collimated into one jet
89                      ! * 2 : a single jet is allowed for all phase space points
90                      ! ! SINGLE_DECAYJET = 2 is not allowed for LOPROCESS_PLUS_JET = false
91 → in vbfno.dat
92 QSQAMIN_ZDEC = 0d0       ! minimal photon virtuality of hadronically decaying photon / Z
93                      ! Values greater than zero may be needed for the interface to
94                      ! parton shower via LesHouches event file.
95                      ! Additionally, there are individual thresholds implemented for
96                      ! each final state quark flavor.
97                      ! See manual for more information.

```

The generation of the parton shower is performed using PYTHIA 8 equivalently as described in Section B.1.3.

B.1.5 EW6 H7-VBFNLO LO

This dataset was produced with a central installation of HERWIG 7 using the ATLAS-internal interface in release MCPProd,20.7.9.9.13. An example for a configuration file used for the generation is printed below.

```

1 from Herwig7_i.Herwig7_iConf import Herwig7
2 from Herwig7_i.Herwig7ConfigMatchbox import Hw7ConfigMatchbox
3
4 genSeq += Herwig7()
5
6 ## Provide config information
7 evgenConfig.generators += ["Herwig7"]
8 evgenConfig.tune = "MMHT2014"
9 evgenConfig.description = "Herwig7 WZjj EW sample for testing only"
10 evgenConfig.keywords = ["SM", "WZ", "VBS"]
11 evgenConfig.contact = ["carsten.bittrich@cern.ch"]
12
13
14 ## initialize generator configuration object
15 generator = Hw7ConfigMatchbox(genSeq, runArgs, run_name="HerwigMatchbox",
16                               beams="pp")
17
18 ## configure generator
19 generator.me_pdf_commands(order="NLO", name="NNPDF30_nlo_as_0118")
20 generator.tune_commands()
21
22
23 generator.add_commands("""
24 #####
25 ## Process selection
26 #####
27
28 ## Model assumptions
29 read Matchbox/StandardModelLike.in
30
31 ## Set the hard process
32 set /Herwig/MatrixElements/Matchbox/Factory:OrderInAlphaS 0
33 set /Herwig/MatrixElements/Matchbox/Factory:OrderInAlphaEW 6
34
35 # W-Z
36 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j e- nu_ebar tau+ tau-
37 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j e- nu_ebar mu+ mu-
38 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j mu- nu_mubar tau+ tau-
39 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j mu- nu_mubar e+ e-
40 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j tau- nu_taubar mu+ mu-
41 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j tau- nu_taubar e+ e-
42
43 #####
44 ## Matrix element library selection
45 #####
46
47 # read Matchbox/MadGraph-GoSam.in
48 # read Matchbox/MadGraph-MadGraph.in
49 # read Matchbox/MadGraph-NJet.in
50 # read Matchbox/MadGraph-OpenLoops.in
51 # read Matchbox/HJets.in
52 read Matchbox/VBFNLO.in
53
54 # proposed by Simon Plaetzer
55 read Matchbox/VBFNLOPhasespace.in
56 read snippets/MonacoSampler.in
57
58 cd /Herwig/MatrixElements/Matchbox/Amplitudes
59 set VBFNLO:RandomHelicitySummation Yes

```

```

60
61 #####
62 ## Cut selection
63 ## See the documentation for more options
64 #####
65
66
67 cd /Herwig/Cuts
68
69 ## mll > 15 GeV
70 set /Herwig/Cuts/MassCut:MinM 0*GeV
71 set /Herwig/Cuts/MassCut:MaxM 13000*GeV
72 set /Herwig/Cuts/MassCut:CComb Gamma
73 set /Herwig/Cuts/ChargedLeptonPairMassCut:MinMass 4*GeV
74 set /Herwig/Cuts/ChargedLeptonPairMassCut:SameFlavourOnly Yes
75 set /Herwig/Cuts/ChargedLeptonPairMassCut:OppositeSignOnly Yes
76
77 ## cuts on additional jets
78
79 read Matchbox/DefaultPPJets.in
80
81 set FirstJet:PtMin 15.*GeV
82 set SecondJet:PtMin 15.*GeV
83 # set ThirdJet:PtMin 10.*GeV
84 # set FourthJet:PtMin 10.*GeV
85 do FirstJet:YRange -7.0 7.0
86 do SecondJet:YRange -7.0 7.0
87 # do ThirdJet:YRange -7.0 7.0
88 # do FourthJet:YRange -7.0 7.0
89
90 set FirstJet:Fuzzy No
91 set SecondJet:Fuzzy No
92
93 insert JetCuts:JetRegions 0 FirstJet
94 insert JetCuts:JetRegions 1 SecondJet
95
96 #####
97 ## Scale choice
98 ## See the documentation for more options
99 #####
100
101 cd /Herwig/MatrixElements/Matchbox
102 #set Factory:ScaleChoice /Herwig/MatrixElements/Matchbox/Scales/FixedScale
103 #set Scales/FixedScale:FixedScale 171.6*GeV
104
105 set Factory:ScaleChoice Scales/HTPrimeScale
106 set Scales/HTPrimeScale:JetPtCut 15.*GeV
107
108 #####
109 ## Matching and shower selection
110 ## Please also see flavour scheme settings
111 ## towards the end of the input file.
112 #####
113
114 # read Matchbox/MCatNLO-DefaultShower.in
115 # read Matchbox/Powheg-DefaultShower.in
116 ## use for strict LO/NLO comparisons
117 # read Matchbox/MCatLO-DefaultShower.in
118 ## use for improved LO showering
119 read Matchbox/LO-DefaultShower.in
120 cd /Herwig/Shower
121
122 # read Matchbox/MCatNLO-DipoleShower.in
123 # read Matchbox/Powheg-DipoleShower.in
124 ## use for strict LO/NLO comparisons
125 # read Matchbox/MCatLO-DipoleShower.in
126 ## use for improved LO showering
127 # read Matchbox/LO-DipoleShower.in
128
129 # read Matchbox/LO-NoShower.in

```

```

130 # read Matchbox/NLO-NoShower.in
131
132
133 #####
134 ## PDF choice
135 #####
136
137 read Matchbox/FiveFlavourScheme.in
138 ## required for dipole shower and fixed order in five flavour scheme
139 # read Matchbox/FiveFlavourNoBMassScheme.in
140 """)
141
142 # generator.sampler_commands("CellGridSampler", 1000, 2, 1000, 1, 100)
143 generator.sampler_commands("MonacoSampler", 30000, 6, 70000, 1, 100)
144 # generator.sampler_commands("FlatBinSampler", 1000, 2, 1000, 1, 100)
145
146 ## run generator
147 if runArgs.generatorRunMode is None:
148     generator.run(runArgs.generatorJobNumber, gridpack="gridpack.tar.gz")
149 elif runArgs.generatorRunMode == 'build':
150     # use the -generatorJobNumber command line parameter to dynamically
151     # specify the total number of parallel integration jobs
152     generator.do_build(integration_jobs=runArgs.generatorJobNumber)
153
154 elif runArgs.generatorRunMode == 'integrate':
155     # use the -generatorJobNumber command line parameter to dynamically
156     # specify which specific integration job is to be run
157     generator.do_integrate(runArgs.generatorJobNumber)
158
159 elif runArgs.generatorRunMode == 'mergegrids':
160     # combine integration grids and prepare a gridpack
161     # use the -generatorJobNumber command line parameter to dynamically
162     # specify the total number of parallel integration jobs
163     generator.do_mergegrids(integration_jobs=runArgs.generatorJobNumber,
164                             gridpack="gridpack.tar.gz")
165
166 elif runArgs.generatorRunMode == 'run':
167     # generate events using the specified gridpack
168     generator.do_run(gridpack="gridpack.tar.gz", cleanup_herwig_scratch=False)

```

Equivalently to the MG5_AMC setups, the full process is separated according to the leptons charge and flavour combinations. The blocks defining the other processes are printed below. Only these are changed with respect to the earlier example of a full configuration file.

```

1 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j e- nu_ebar e+ e-
2 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j mu- nu_mubar mu+ mu-
3 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j tau- nu_taubar tau+ tau-

```

```

1 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j e+ nu_e tau+ tau-
2 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j e+ nu_e mu+ mu-
3 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j mu+ nu_mu tau+ tau-
4 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j mu+ nu_mu e+ e-
5 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j tau+ nu_tau mu+ mu-
6 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j tau+ nu_tau e+ e-

```

```

1 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j e+ nu_e e+ e-
2 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j mu+ nu_mu mu+ mu-
3 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j tau+ nu_tau tau+ tau-

```

B.1.6 EW6 H7-VBFNLO NLO

The NLO generation of the EW6 H7-VBFNLO setup required a few changes for a successful integration, compared to the LO configuration. Firstly, the matching and showering snippets are changed from L0-DefaultShower.in to MCatNLO-DefaultShower.in, representing

the switch to a NLO-QCD calculation of the hard process and the configuration of the MC@NLO-style matching. In order to reduce complexity, only the diagonal elements of the CKM matrix are considered and the b -quark is considered massless. The efficiency of the phase space integrator is increased by explicitly applying the VBF approximation in the phase space integration. Note, that this does not have a physics effect since the VBF approximation is already applied in the matrix elements provided by VBFNLO. Furthermore, the internal parameter of the optimization of the phase space integration efficiency is adjusted in order to increase performance and reduce the fraction of events with negative weights.

An example for the full configuration file is printed below for a part of the full process. The processes with other lepton charge and flavour combinations are derived equivalently.

```

1 from Herwig7_i.Herwig7_iConf import Herwig7
2 from Herwig7_i.Herwig7ConfigMatchbox import Hw7ConfigMatchbox
3
4 genSeq += Herwig7()
5
6 ## Provide config information
7 evgenConfig.generators += ["Herwig7"]
8 evgenConfig.tune = "MMHT2014"
9 evgenConfig.description = "Herwig7 WZjj EW sample for testing only"
10 evgenConfig.keywords = ["SM", "WZ", "VBS"]
11 evgenConfig.contact = ["carsten.bittrich@cern.ch"]
12
13
14 ## initialize generator configuration object
15 generator = Hw7ConfigMatchbox(genSeq, runArgs, run_name="HerwigMatchbox",
16                               beams="pp")
17
18 ## configure generator
19 generator.me_pdf_commands(order="NLO", name="NNPDF30_nlo_as_0118")
20 generator.tune_commands()
21
22
23 generator.add_commands("""
24 #####
25 ## Process selection
26 #####
27
28 ## Model assumptions
29 read Matchbox/StandardModelLike.in
30 read Matchbox/DiagonalCKM.in
31
32 ## Set the hard process
33 set /Herwig/MatrixElements/Matchbox/Factory:OrderInAlphaS 0
34 set /Herwig/MatrixElements/Matchbox/Factory:OrderInAlphaEW 6
35
36 # proposal from Simon: between 0.1 and 0.01
37 set /Herwig/MatrixElements/Matchbox/Factory:AlphaParameter 0.1
38
39 # W-Z
40 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j e- nu_ebar tau+ tau-
41 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j e- nu_ebar mu+ mu-
42 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j mu- nu_mubar tau+ tau-
43 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j mu- nu_mubar e+ e-
44 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j tau- nu_taubar mu+ mu-
45 do /Herwig/MatrixElements/Matchbox/Factory:Process p p j j tau- nu_taubar e+ e-
46
47 # tell the integrator that the VBS approximation is used (no physics impact for VBFNLO ME)
48 read Matchbox/VBFDiagramsOnly.in
49
50 #####
51 ## Matrix element library selection
52 #####
53
54 # read Matchbox/MadGraph-GoSam.in
55 # read Matchbox/MadGraph-MadGraph.in
56 # read Matchbox/MadGraph-NJet.in

```

```

57 # read Matchbox/MadGraph-OpenLoops.in
58 # read Matchbox/HJets.in
59 read Matchbox/VBFNLO.in
60
61 # proposed by Simon Plaetzer
62 read Matchbox/VBFNLOPhasespace.in
63 read snippets/MonacoSampler.in
64
65 cd /Herwig/MatrixElements/Matchbox/Amplitudes
66 set VBFNLO:RandomHelicitySummation Yes
67
68 #####
69 ## Cut selection
70 ## See the documentation for more options
71 #####
72
73
74 cd /Herwig/Cuts
75
76 ## m_ll > 15 GeV
77 set /Herwig/Cuts/MassCut:MinM 0*GeV
78 set /Herwig/Cuts/MassCut:MaxM 13000*GeV
79 set /Herwig/Cuts/MassCut:CComb Gamma
80 set /Herwig/Cuts/ChargedLeptonPairMassCut:MinMass 4*GeV
81 set /Herwig/Cuts/ChargedLeptonPairMassCut:SameFlavourOnly Yes
82 set /Herwig/Cuts/ChargedLeptonPairMassCut:OppositeSignOnly Yes
83
84 ## cuts on additional jets
85
86 read Matchbox/DefaultPPJets.in
87
88 set FirstJet:PtMin 15.*GeV
89 set SecondJet:PtMin 15.*GeV
90 # set ThirdJet:PtMin 10.*GeV
91 # set FourthJet:PtMin 10.*GeV
92 do FirstJet:YRange -7.0 7.0
93 do SecondJet:YRange -7.0 7.0
94 # do ThirdJet:YRange -7.0 7.0
95 # do FourthJet:YRange -7.0 7.0
96
97 set FirstJet:Fuzzy No
98 set SecondJet:Fuzzy No
99
100 insert JetCuts:JetRegions 0 FirstJet
101 insert JetCuts:JetRegions 1 SecondJet
102
103 #####
104 ## Scale choice
105 ## See the documentation for more options
106 #####
107
108 cd /Herwig/MatrixElements/Matchbox
109 #set Factory:ScaleChoice /Herwig/MatrixElements/Matchbox/Scales/FixedScale
110 #set Scales/FixedScale:FixedScale 171.6*GeV
111
112 set Factory:ScaleChoice Scales/HTPrimeScale
113 set Scales/HTPrimeScale:JetPtCut 15.*GeV
114
115 #####
116 ## Matching and shower selection
117 ## Please also see flavour scheme settings
118 ## towards the end of the input file.
119 #####
120
121 read Matchbox/MCatNLO-DefaultShower.in
122 # read Matchbox/Powheg-DefaultShower.in
123 ## use for strict LO/NLO comparisons
124 # read Matchbox/MCatLO-DefaultShower.in
125 ## use for improved LO showering
126 # read Matchbox/LO-DefaultShower.in

```

```

127 cd /Herwig/Shower
128
129 # read Matchbox/MCatNLO-DipoleShower.in
130 # read Matchbox/Powheg-DipoleShower.in
131 ## use for strict LO/NLO comparisons
132 # read Matchbox/MCatLO-DipoleShower.in
133 ## use for improved LO showering
134 # read Matchbox/LO-DipoleShower.in
135
136 # read Matchbox/LO-NoShower.in
137 # read Matchbox/NLO-NoShower.in
138
139
140 #####
141 ## PDF choice
142 #####
143
144 # read Matchbox/FiveFlavourScheme.in
145 ## required for dipole shower and fixed order in five flavour scheme
146 read Matchbox/FiveFlavourNoBMassScheme.in
147 "")
148
149 # generator.sampler_commands("CellGridSampler", 1000, 2, 1000, 1, 100)
150 generator.sampler_commands("MonacoSampler", 50000, 6, 90000, 1, 100)
151 # generator.sampler_commands("FlatBinSampler", 1000, 2, 1000, 1, 100)
152
153 ## run generator
154 if runArgs.generatorRunMode is None:
155     generator.run(runArgs.generatorJobNumber, gridpack="gridpack.tar.gz")
156 elif runArgs.generatorRunMode == 'build':
157     # use the -generatorJobNumber command line parameter to dynamically
158     # specify the total number of parallel integration jobs
159     generator.do_build(integration_jobs=runArgs.generatorJobNumber)
160
161 elif runArgs.generatorRunMode == 'integrate':
162     # use the -generatorJobNumber command line parameter to dynamically
163     # specify which specific integration job is to be run
164     generator.do_integrate(runArgs.generatorJobNumber)
165
166 elif runArgs.generatorRunMode == 'mergegrids':
167     # combine integration grids and prepare a gridpack
168     # use the -generatorJobNumber command line parameter to dynamically
169     # specify the total number of parallel integration jobs
170     generator.do_mergegrids(integration_jobs=runArgs.generatorJobNumber,
171                             gridpack="gridpack.tar.gz")
172
173 elif runArgs.generatorRunMode == 'run':
174     # generate events using the specified gridpack
175     generator.do_run(gridpack="gridpack.tar.gz", cleanup_herwig_scratch=False)

```

B.1.7 EW5 MG5Py8

This dataset was produced with the same combination of private software and releases as documented in Section B.1.1. Only the configuration file is adjusted as printed below.

```

1 from MadGraphControl.MadGraphUtils import *
2
3 evgenConfig.generators = ["MadGraph", "Pythia8", "EvtGen"]
4
5 evgenConfig.description = 'MadGraph_lvlljj_EW6_OFMinus_masslessLeptons'
6 evgenConfig.keywords = ['SM', 'diboson', 'WZ', 'electroweak', '3lepton', 'VBS']
7 evgenConfig.contact = ['carsten.bittrich@cern.ch']
8
9 runName='run_01'
10 process_total=None
11
12 evgenConfig.inputfilecheck = runName
13

```

```

14 #GridPack
15 do_gridpack_only = "inputGenConfFile" not in dir(runArgs)
16
17 gridpack_mode=True
18 gridpack_dir='madevent/'
19 gridpack_compile=False
20 nevents = 1 if do_gridpack_only else 6000
21
22 if not do_gridpack_only:
23     Mode=0
24     nJobs=1
25     Cluster_Type=None
26     Cluster_Queue=None
27     Cluster_Temp_Path='/lustre/ssd/s1223011'
28 else:
29     Mode=1
30     nJobs=24
31     Cluster_Type='slurm'
32     Cluster_Queue='haswell64'
33     Cluster_Size=24
34
35     Mode=2
36     nJobs=24
37     Cluster_Type=None
38     Cluster_Queue=None
39
40
41 beamEnergy=-999
42 if hasattr(runArgs,'ecmEnergy'):
43     beamEnergy = runArgs.ecmEnergy / 2.
44 else:
45     raise RuntimeError("No center of mass energy found.")
46
47 #Fetch default LO run_card.dat and set parameters
48 extras = {
49     'auto_ptj_mjj': False, 'ptl': 4.0, 'ptj': 15.0, 'maxjetflavor': 5,
50     'asrwtflavor': 5, 'pdlabel': "'lhpdf'", 'drll': 0.2, 'ptb': 15.0,
51     'mml1': 4.0, 'event_norm': 'sum', 'systematics_program': 'systematics',
52     'drjl': 0.2, 'lhaid': 260000, 'drjj': 0.4, 'ptheavy': 2.0,
53     'use_syst': True, 'etal': 5.0, 'dral': 0.1, 'cut_decays': True,
54     'etaj': 5.5, 'systematics_arguments':
55         "['-mur=0.5,1,2', '-muf=0.5,1,2'," +
56         "'-pdf=errorset,NNPDF30_nlo_as_0119@0,NNPDF30_nlo_as_0117@0," +
57         "'CT14nlo,MMHT2014nlo68clas18,PDF4LHC15_nlo_30_pdfas']"}
58
59 extras["gridpack"] = '.true.' if do_gridpack_only else ".false."
60
61 process_dir = new_process(grid_pack=gridpack_dir)
62
63 if False:
64     os.system("cp $MADPATH/FKS_params.dat Cards")
65
66 build_run_card(run_card_old=get_default_runcard(proc_dir=process_dir),
67               run_card_new='run_card.dat', xqcut=0,
68               nevt=nevents,rand_seed=runArgs.randomSeed,beamEnergy=beamEnergy)
69 print_cards()
70 modify_run_card(run_card='run_card.dat',
71                run_card_backup='run_card_backup.dat',
72                settings=extras)
73
74 print_cards()
75
76 generate(
77     required_accuracy=0.001,run_card_loc='run_card.dat',param_card_loc=None,
78     mode=Mode,njobs=nJobs,cluster_type=Cluster_Type,cluster_queue=Cluster_Queue,
79     proc_dir=process_dir,run_name=runName,grid_pack=gridpack_mode,
80     gridpack_compile=gridpack_compile,gridpack_dir=gridpack_dir,nevents=nevents,
81     random_seed=runArgs.randomSeed)
82
83

```

```

84 output_suffix = gridpack_dir if do_gridpack_only else runName
85 arrange_output(run_name=runName,proc_dir=process_dir,
86               outputDS=output_suffix+'._00001.events.tar.gz',lhe_version=3,
87               saveProcDir=True)
88
89 if not do_gridpack_only:
90
91     runArgs.inputGeneratorFile=output_suffix+'._00001.events.tar.gz'
92
93     include("MC15JobOptions/Pythia8_A14_NNPDF23L0_EvtGen_Common.py")
94     include("MC15JobOptions/Pythia8_MadGraph.py")

```

B.1.8 EW4 MG5Py8

This dataset was produced with the same combination of private software and releases as documented in Section B.1.1. Only the configuration file is adjusted as printed below.

```

1 from MadGraphControl.MadGraphUtils import *
2
3 evgenConfig.generators = ["MadGraph", "Pythia8", "EvtGen"]
4
5 evgenConfig.description = 'MadGraph_lllv_OFMinus_ckkw_Np0123_262'
6 evgenConfig.keywords = ['SM', 'diboson', 'WZ', 'electroweak', '3lepton']
7 evgenConfig.contact = ['carsten.bittrich@cern.ch']
8
9 runName='run_01'
10 process_total=None
11
12 evgenConfig.inputfilecheck = runName
13
14 #GridPack
15 do_gridpack_only = "inputGenConfFile" not in dir(runArgs)
16
17 gridpack_mode=True
18 gridpack_dir='madevent/'
19 gridpack_compile=False
20 nevents = 1 if do_gridpack_only else 8000
21
22 if not do_gridpack_only:
23     Mode=0
24     nJobs=1
25     Cluster_Type=None
26     Cluster_Queue=None
27     Cluster_Temp_Path='/lustre/ssd/s1223011'
28 else:
29     Mode=1
30     nJobs=24
31     Cluster_Type='slurm'
32     Cluster_Queue='haswell64'
33     Cluster_Size=24
34
35     Mode=2
36     nJobs=24
37     Cluster_Type=None
38     Cluster_Queue=None
39
40 beamEnergy=-999
41 if hasattr(runArgs,'ecmEnergy'):
42     beamEnergy = runArgs.ecmEnergy / 2.
43 else:
44     raise RuntimeError("No center of mass energy found.")
45
46 #Fetch default LO run_card.dat and set parameters
47 extras = {
48     'auto_ptj_mjj': False, 'ptl': 4.0, 'ptj': 15.0, 'drll': 0.2, 'ptb': 15.0,
49     'dral': 0.1, 'mml1': 4.0, 'ickkw': 0, 'drjl': 0.2, 'systematics_program':
50     'systematics', 'drjj': 0.0, 'systematics_arguments':

```

```

52     "['-mur=0.5,1,2', '-muf=0.5,1,2', " +
53     "'-pdf=errorset,NNPDF30_nlo_as_0119@0,NNPDF30_nlo_as_0117@0," +
54     "CT14nlo,MMHT2014nlo68clas118,PDF4LHC15_nlo_30_pdfas']",
55     'maxjetflavor': 5, 'asrwtgflavor': 5, 'pdlabel': "'lhpdf'", 'event_norm':
56     'sum', 'etal': 5.2, 'ptheavy': 2.0, 'use_syst': True, 'cut_decays': True,
57     'ktDurham': 30, 'dparameter': 0.4, 'lhaid': 260000, 'etaj': 5.5}
58
59 extras["gridpack"] = '.true.' if do_gridpack_only else ".false."
60
61 process_dir = new_process(grid_pack=gridpack_dir)
62
63 if False:
64     os.system("cp $MADPATH/FKS_params.dat Cards")
65
66 build_run_card(run_card_old=get_default_runcard(proc_dir=process_dir),
67               run_card_new='run_card.dat', xqcut=0,
68               nevents=nevents, rand_seed=runArgs.randomSeed, beamEnergy=beamEnergy)
69 print_cards()
70 modify_run_card(run_card='run_card.dat',
71                run_card_backup='run_card_backup.dat',
72                settings=extras)
73
74 print_cards()
75
76 generate(
77     required_accuracy=0.001, run_card_loc='run_card.dat', param_card_loc=None,
78     mode=Mode, njobs=nJobs, cluster_type=Cluster_Type,
79     cluster_queue=Cluster_Queue, proc_dir=process_dir, run_name=runName,
80     grid_pack=gridpack_mode, gridpack_compile=gridpack_compile,
81     gridpack_dir=gridpack_dir, nevents=nevents, random_seed=runArgs.randomSeed)
82
83 output_suffix = gridpack_dir if do_gridpack_only else runName
84 arrange_output(run_name=runName, proc_dir=process_dir, saveProcDir=True,
85               outputDS=output_suffix+'.00001.events.tar.gz', lhe_version=3)
86
87 if not do_gridpack_only:
88
89     runArgs.inputGeneratorFile=output_suffix+'.00001.events.tar.gz'
90     PYTHIA8_nJetMax=3
91     PYTHIA8_Dparameter=extras['dparameter']
92     PYTHIA8_Process='pp>LEPTONS,NEUTRINOS'
93     PYTHIA8_TMS=extras['ktDurham']
94     PYTHIA8_nQuarksMerge=extras['maxjetflavor']
95
96     include("MC15JobOptions/Pythia8_A14_NNPDF23L0_EvtGen_Common.py")
97     include("MC15JobOptions/Pythia8_MadGraph.py")
98     if do_merging:
99         include("MC15JobOptions/Pythia8_CKKWL_kTMerge.py")

```

B.2 Simulations at Reconstruction Level

Lists of all simulations used in the reconstruction-level analysis are presented in Table B.1 for the signal process, Table B.2 for prompt background processes, and for non-prompt background processes in Table B.3, Table B.4, and Table B.5.

Table B.1: List of MC simulations for the signal process used for reconstruction level analysis. For each simulation sample the internal classifier ID, the process description, the ATLAS tags, the generator-level cross-section in pb, the filter efficiency ε , and a correction K -factor for the predicted normalization are listed.

DSID	Process description	ATLAS tags	σ [pb]	ε	K
999999	MG5H7_lvljj_EW6_OFMinus_masslessLeptons	e20190814_s2726_r7772_r7676_p3317	0.0154368	1	1
999999	MG5H7_lvljj_EW6_OFPlus_masslessLeptons	e20190814_s2726_r7772_r7676_p3317	0.0257468	1	1
999999	MG5H7_lvljj_EW6_SFMinus_masslessLeptons	e20190814_s2726_r7772_r7676_p3317	0.00770693	1	1
999999	MG5H7_lvljj_EW6_SFPlus_masslessLeptons	e20190814_s2726_r7772_r7676_p3317	0.0128645	1	1
364499	MGPγ8EG_NNPDF30LO_A14NNPDF23LO_WZlνljjEWK	e6376_s2726_r7772_r7676_p3317	0.059454	1	1
364739	MGPγ8EG_NNPDF30NLO_A14NNPDF23LO_lνljjEW6_OFMinus	e7421_s2726_r7772_r7676_p3317	0.0154368	1	1
364740	MGPγ8EG_NNPDF30NLO_A14NNPDF23LO_lνljjEW6_OFPlus	e7421_s2726_r7772_r7676_p3317	0.0257468	1	1
364741	MGPγ8EG_NNPDF30NLO_A14NNPDF23LO_lνljjEW6_SFMinus	e7421_s2726_r7772_r7676_p3317	0.00770693	1	1
364742	MGPγ8EG_NNPDF30NLO_A14NNPDF23LO_lνljjEW6_SFPlus	e7421_s2726_r7772_r7676_p3317	0.0128645	1	1
364284	Sherpa_222_NNPDF30NNLO_3llνjj_EW6	e6055_s2726_r7772_r7676_p3325	0.046849	1	1

Table B.2: List of MC simulations for the prompt background process used for reconstruction level analysis. For each simulation sample the internal classifier ID, the process description, the ATLAS tags, the generator-level cross-section in pb, the filter efficiency ε , and a correction K -factor for the predicted normalization are listed.

DSID	Process description	ATLAS tags	σ [pb]	ε	K
364499	MGP_y8EG_NNPDF30LO_A14NNPDF23LO_WZlvlljjEWK	e6376_s2726_r7772_r7676_p3317	0.059454	1	1
364250	Sherpa_222_NNPDF30NNLO_3lll	e5894_s2726_r7772_r7676_p3314	1.2519	1	1
364283	Sherpa_222_NNPDF30NNLO_3lljj_EW6	e6055_s2726_r7772_r7676_p3314	0.010451	1	1
364253	Sherpa_222_NNPDF30NNLO_3llv	e5916_s2726_r7772_r7676_p3317	4.5784	1	1
361620	Sherpa_CT10_WWW_3l3v	e4093_s2608_s2183_r8112_r7676_p3314	0.008343	1	1
361622	Sherpa_CT10_WWZ_2l4v	e4053_s2608_s2183_r8112_r7676_p3314	0.0034299	1	1
361621	Sherpa_CT10_WWZ_4l2v	e4053_s2608_s2183_r8112_r7676_p3314	0.001734	1	1
361624	Sherpa_CT10_WZZ_3l3v	e4093_s2608_s2183_r8112_r7676_p3314	0.0019248	0.198	1
361623	Sherpa_CT10_WZZ_5l1v	e4093_s2608_s2183_r8112_r7676_p3314	0.0002178	1	1
361627	Sherpa_CT10_ZZZ_2l4v	e4093_s2608_s2183_r7772_r7676_p3314	0.0004453	0.201	1
361626	Sherpa_CT10_ZZZ_4l2v	e4093_s2608_s2183_r7772_r7676_p3314	0.0004412	0.0508	1
361625	Sherpa_CT10_ZZZ_6l0v	e4093_s2608_s2183_r7772_r7676_p3314	1.7059e-05	1	1
361073	Sherpa_CT10_ggllll	e3836_s2608_s2183_r7772_r7676_p3314	0.020743	1	1.52
410155	aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttW	e5070_s2726_r7772_r7676_p3314	0.54838	1	1.1
410218	aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttee	e5070_s2726_r7772_r7676_p3314	0.036865	1	1.12
410219	aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttmumu	e5070_s2726_r7772_r7676_p3314	0.036868	1	1.12
410220	aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttttautau	e5070_s2726_r7772_r7676_p3314	0.036515	1	1.12

Table B.3: List of MC simulations for the non-prompt background process other than $V + \gamma$ processes used for reconstruction level analysis. For each simulation sample the internal classifier ID, the process description, the ATLAS tags, the generator-level cross-section in pb, the filter efficiency ε , and a correction K -factor for the predicted normalization are listed.

DSID	Process description	ATLAS tags	σ [pb]	ε	K
361606	PowhegPy8EG_CT10nloME_AZNLOCTEQ6L1_WWlvqq	e4711_s2726_r7772_r7676_p3314	44.181	1	1
361607	PowhegPy8EG_CT10nloME_AZNLOCTEQ6L1_WZqql_mll20	e4711_s2726_r7772_r7676_p3314	3.2881	1	1
361608	PowhegPy8EG_CT10nloME_AZNLOCTEQ6L1_WZqqvv	e4711_s2726_r7772_r7676_p3314	5.7746	1	1
410501	PowhegPythia8EvtGen_A14_ttbar_hdamp258p75_nonallhad	e5458_s2726_r7772_r7676_p3314	730.17	0.2957	1.139
361103	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Wminusenu	e3601_s2576_s2132_r7725_r7676_p3314	8282.7	1	1.036
361104	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Wminusmunu	e3601_s2576_s2132_r7725_r7676_p3314	8282.8	1	1.036
361105	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Wminustaunu	e3601_s2576_s2132_r7772_r7676_p3314	8282.6	1	1.036
361100	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Wplusenu	e3601_s2576_s2132_r7725_r7676_p3314	11306	1	1.017
361101	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Wplusmunu	e3601_s2576_s2132_r7725_r7676_p3314	11306	1	1.017
361102	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Wplustaunu	e3601_s2576_s2132_r7725_r7676_p3314	11306	1	1.017
361106	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zee	e3601_s2576_s2132_r7725_r7676_p3314	1901.2	1	1.026
361107	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zmunu	e3601_s2576_s2132_r7725_r7676_p3314	1901.1	1	1.026
361108	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Ztautau	e3601_s2726_r7725_r7676_p3314	1901.2	1	1.026
410026	PowhegPythiaEvtGen_P2012_SingleTopSchan_noAllHad_antitop	e3998_s2608_s2183_r7725_r7676_p3314	1.2615	1	1.022
410025	PowhegPythiaEvtGen_P2012_SingleTopSchan_noAllHad_top	e3998_s2608_s2183_r7725_r7676_p3314	2.0514	1	1.005
410016	PowhegPythiaEvtGen_P2012_Wt_dilepton_antitop	e3753_s2608_s2183_r7725_r7676_p3314	3.5816	1	1.054
410015	PowhegPythiaEvtGen_P2012_Wt_dilepton_top	e3753_s2608_s2183_r7725_r7676_p3314	3.5837	1	1.054
410012	PowhegPythiaEvtGen_P2012_singlelep_tchan_lept_antitop	e3824_s2608_s2183_r7725_r7676_p3314	25.778	1	1.019
410011	PowhegPythiaEvtGen_P2012_singlelep_tchan_lept_top	e3824_s2608_s2183_r7725_r7676_p3314	43.739	1	1.009
364254	Sherpa_222_NNPDF30NNLO_llvv	e5916_s2726_r7772_r7676_p3314	12.5	1	1

Table B.4: List of MC simulations for the $W^\pm + \gamma$ non-prompt background process used for reconstruction level analysis. For each simulation sample the internal classifier ID, the process description, the ATLAS tags, the generator-level cross-section in pb, the filter efficiency ε , and a correction K -factor for the predicted normalization are listed.

DSID	Process description	ATLAS tags	σ [pb]	ε	K
364525	Sherpa_222_NNPDF30NNLO_enugamma_pty_140_E_CMS	e5928_s2726_r7772_r7676_p3314	0.29742	1	1
364522	Sherpa_222_NNPDF30NNLO_enugamma_pty_15_35	e5928_s2726_r7772_r7676_p3314	134.42	1	1
364523	Sherpa_222_NNPDF30NNLO_enugamma_pty_35_70	e5928_s2726_r7772_r7676_p3314	19.077	1	1
364524	Sherpa_222_NNPDF30NNLO_enugamma_pty_70_140	e5928_s2726_r7772_r7676_p3314	1.9212	1	1
364521	Sherpa_222_NNPDF30NNLO_enugamma_pty_7_15	e5928_s2726_r7772_r7676_p3314	199.27	1	1
364530	Sherpa_222_NNPDF30NNLO_munugamma_pty_140_E_CMS	e5928_s2726_r7772_r7676_p3314	0.29763	1	1
364527	Sherpa_222_NNPDF30NNLO_munugamma_pty_15_35	e5928_s2726_r7772_r7676_p3314	134.4	1	1
364528	Sherpa_222_NNPDF30NNLO_munugamma_pty_35_70	e5928_s2726_r7772_r7676_p3314	19.133	1	1
364529	Sherpa_222_NNPDF30NNLO_munugamma_pty_70_140	e5928_s2726_r7772_r7676_p3314	1.9221	1	1
364526	Sherpa_222_NNPDF30NNLO_munugamma_pty_7_15	e5928_s2726_r7772_r7676_p3314	199.44	1	1
364535	Sherpa_222_NNPDF30NNLO_tmunugamma_pty_140_E_CMS	e5928_s2726_r7772_r7676_p3314	0.29736	1	1
364532	Sherpa_222_NNPDF30NNLO_tmunugamma_pty_15_35	e5928_s2726_r7772_r7676_p3314	134.44	1	1
364533	Sherpa_222_NNPDF30NNLO_tmunugamma_pty_35_70	e5928_s2726_r7772_r7676_p3314	19.126	1	1
364534	Sherpa_222_NNPDF30NNLO_tmunugamma_pty_70_140	e5928_s2726_r7772_r7676_p3314	1.9283	1	1
364531	Sherpa_222_NNPDF30NNLO_tmunugamma_pty_7_15	e5928_s2726_r7772_r7676_p3314	199.4	1	1

Table B.5: List of MC simulations for the $Z + \gamma$ non-prompt background process used for reconstruction level analysis. For each simulation sample the internal classifier ID, the process description, the ATLAS tags, the generator-level cross-section in pb, the filter efficiency ε , and a correction K -factor for the predicted normalization are listed.

DSID	Process description	ATLAS tags	σ [pb]	ε	K
364504	Sherpa_222_NNPDF30NNLO_eegamma_pty_140_E_CMS	e5928_s2726_r7772_r7676_p3314	0.062987	1	1
364501	Sherpa_222_NNPDF30NNLO_eegamma_pty_15_35	e5928_s2726_r7772_r7676_p3314	34.59	1	1
364502	Sherpa_222_NNPDF30NNLO_eegamma_pty_35_70	e5928_s2726_r7772_r7676_p3314	6.2856	1	1
364503	Sherpa_222_NNPDF30NNLO_eegamma_pty_70_140	e5928_s2726_r7772_r7676_p3314	0.49186	1	1
364500	Sherpa_222_NNPDF30NNLO_eegamma_pty_7_15	e5928_s2726_r7772_r7676_p3314	57.619	1	1
364509	Sherpa_222_NNPDF30NNLO_mumugamma_pty_140_E_CMS	e5928_s2726_r7772_r7676_p3314	0.06308	1	1
364506	Sherpa_222_NNPDF30NNLO_mumugamma_pty_15_35	e5988_s2726_r7772_r7676_p3314	34.588	1	1
364507	Sherpa_222_NNPDF30NNLO_mumugamma_pty_35_70	e5928_s2726_r7772_r7676_p3314	6.2853	1	1
364508	Sherpa_222_NNPDF30NNLO_mumugamma_pty_70_140	e5928_s2726_r7772_r7676_p3314	0.49392	1	1
364505	Sherpa_222_NNPDF30NNLO_mumugamma_pty_7_15	e5928_s2726_r7772_r7676_p3314	57.701	1	1
364514	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_140_E_CMS	e5928_s2726_r7772_r7676_p3314	0.063417	1	1
364511	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_15_35	e5928_s2726_r7772_r7676_p3314	34.638	1	1
364512	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_35_70	e5928_s2726_r7772_r7676_p3314	6.2964	1	1
364513	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_70_140	e5982_s2726_r7772_r7676_p3314	0.49402	1	1
364510	Sherpa_222_NNPDF30NNLO_tautaugamma_pty_7_15	e5928_s2726_r7772_r7676_p3314	57.619	1	1

Appendix C

Details of Typical Comparison Plots

Typical comparison plots throughout this work have different types of inlays below the main plot. Each type of inlay as well as the overall plot layout described in detail in this chapter. Figure C.1 shows an example plot with each type of inlay used throughout this work. Next

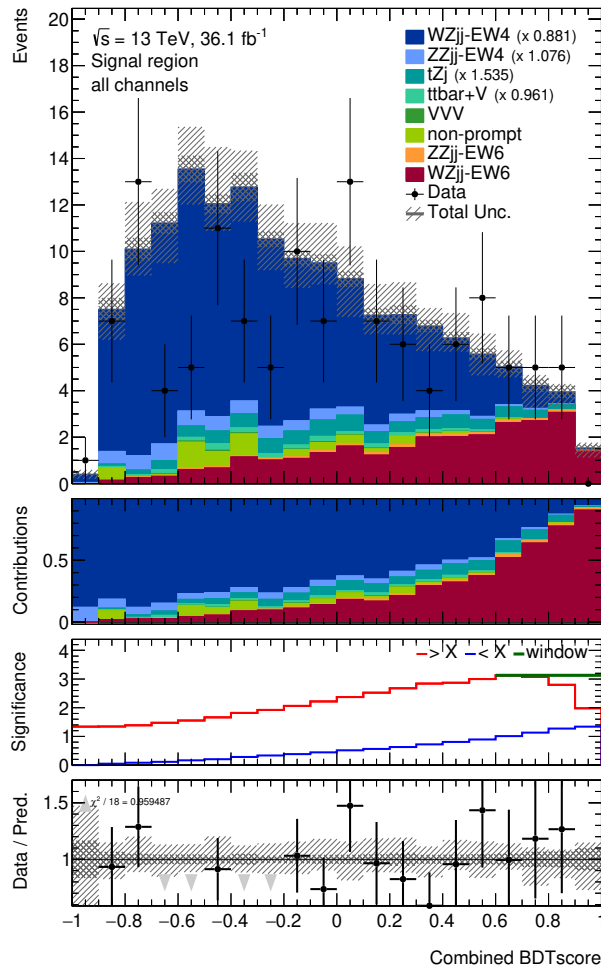


Figure C.1: Example to show details of typical comparison plots. See the main body for a detailed description.

to the main plot, there are three inlays with titles as indicated on the respective y -axes. Each of the four figures share the same x -axis, which is only labeled in the bottom-most inlay. In the main plot, there are three types of visualisations of datasets

Data points Visualized with round markers and vertical error lines. Throughout this work, this style is exclusively used for observed data, Asimov data, or data-driven datasets.

Plain prediction Visualized with a connected line with a hatched uncertainty band, where the area under the line is not filled. This is the plain prediction as indicated in the legend.

Stacked prediction Visualized with differently-coloured bars, where the different colours indicate individual contributions. The contributions are stacked on top of each other, so the sum of the individual contributions is given by the top end of the coloured lines. Only an uncertainty of the sum of all contributions is shown as hatched uncertainty band.

The legend on the top right side specifies the contributions of a stacked prediction and list all predictions shown in the plot. Global normalization corrections to individual processes are indicated in the legend next to each affected process.

In general, event yields shown in such comparisons are shown as events per bin as indicated on the y -axis. Only in Figure 10.1a, the shown as events per fixed interval in the considered observable, in order to preserve the shape of the distribution despite the variable bin size.

The three different types of inlays are

Contribution This inlay specifies the relative contributions to the stacked prediction in each bin, following the same order and colour-coding of processes as in the main plot.

Significance This inlay shows as estimated significance by counting events passing different cut scenarios. The significance Z is estimated following the profile likelihood method for a poissonian distribution including background systematics as described e.g. in [180]. Three different cut scenarios are tested. The lower threshold scenario indicated by the red line, uses the current bin as lower threshold and integrates over all bins with larger value in the observable shown on the x -axis. For the upper threshold scenario indicated by the blue line, the cut direction is inverted. In the window cut scenario, all possible scenarios where a connected interval only is either included or excluded are considered and the scenario with the maximum significance is indicated by highlighting the included bins with a green line at the maximum significance.

Ratio This inlay shows the ratio of different distributions, as indicated on the y -axis. Sometimes multiple numerators are shown, but always with respect to a common denominator. The uncertainty of the denominator distribution is shown in a hatched and potentially cross-hatched area. In some plots, the top left corner of this inlay indicates the reduced χ^2 in order to quantify the agreement between the distributions. In case some points in the ratio are outside of the shown range on the y -axis, typically small grey arrows indicate whether the datapoint is below or above the shown range. Due to technical limitations of the plotting, sometimes superfluous grey arrows are shown. Also the situation is potentially unclear in case of multiple numerators, where there might be bins below and above the shown range.

Appendix D

Calibration of Forward Electrons

This appendix documents the calibration procedure of forward electrons, which was derived by the author as ATLAS qualification project. Next to the documentation of the study, this chapter also provides an example for typical extractions of corrections as done by the central combined performance groups of ATLAS. This chapter is mainly taken from [181], where it was documented first by the author.

In the forward detector region, defined by $|\eta_{\text{calo}}| > 2.5$ no tracking information is available. Consequently, no charge information for electrons are available rendering the $Z \rightarrow ee$ selection more problematic and leading to non-negligible contributions of background processes. Inferior spatial detector resolution and identification of electrons lead to less-precise calibration for this region. In order to make use of the more precise central calibration a tag-and-probe approach with one central electron as tag and a forward electron as probe is chosen.

D.1 Event selection and simulation samples

Energy scales and additional constant terms for electrons and photons in the forward detector region with $|\eta_{\text{calo}}| > 2.5$, are extracted from $Z \rightarrow ee$ events using data collected in 2015 only. These events are selected using AnalysisBase 2.4.22 and using the esModel `es2015c_summer`. The correction coefficients are extracted by comparing the following samples:

- *Zee* MC:
`mc15_13TeV.361106.PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zee.merge.
DAOD_EGAM8.e3601_s2576_s2132_r7773_r7676_p2666`
- Data 2015 (3.21 fb^{-1}):
`data15_13TeV.*.physics_Main.merge.DAOD_EGAM8.r7562_p2521_p2667`

Event selection

Events have to pass the GRL:

- `data15_13TeV.periodAllYear_DetStatus-v73-pro19-08_DQDefects-00-01-02
_PHYS_StandardGRL_All_Good_25ns.xml`,

at least one of the following triggers:

- `HLT_e24_lhmedium_L1EM20VH`
- `HLT_e60_lhmedium`
- `HLT_e120_lhloose`

and the usual event cleaning flags:

- `xAOD::EventInfo::Tile`
- `xAOD::EventInfo::LAr`
- `xAOD::EventInfo::SCT`
- test for incomplete events.

A primary vertex with at least three tracks and $|z| < 150$ mm is required.

In these events central electrons that pass the `TightLH` identification with $p_T > 25$ GeV in the following regions are selected: $|\eta_{\text{calo}}| < 1.37$ or $1.52 < |\eta_{\text{calo}}| < 2.47$ where the track η is also required to be $|\eta_{\text{track}}| < 2.47$. Bad electrons (`BADCLUSELECTRON`) and electrons reconstructed by an author different from `AuthorElectron` are vetoed. Additionally it is required that $|\Delta z_0^{\text{BL}} \sin \theta| < 0.5$ mm and the d_0^{BL} -significance must be smaller than 5. Δz_0^{BL} stands for the difference between the z_0 of the track and the primary vertex expressed at the beam line (BL) and d_0^{BL} the transverse impact parameter with respect to the beam line. Electron candidates are required to pass the `Tight` isolation criterion.

Forward electrons that pass the `Tight` identification with $p_T > 20$ GeV and $|\eta_{\text{calo}}| > 2.5$ are selected. Again, bad electrons (`BADCLUSELECTRON`) and electrons reconstructed by an author different from `AuthorElectron` are vetoed.

In order to avoid assignment problems only events with exactly one such central electron, called tag electron, and exactly one forward electron, called probe electron are accepted. The tag electron is then required to match to the used triggers.

Data calibration and MC reweighting

Full electron energy calibration has been applied for the central electron. For the forward electron no calibration has been applied.

Pileup reweighting is applied to the MC in order to match the pileup profile in data. Additionally, scale factors are applied to correct the reconstruction, isolation, identification (ID) and trigger efficiency for the central electron. The following config files are used:

- pileup reweighting: `mc15c_v2_defaults.NotRecommended.prw.root`
- reconstruction: `ElectronEfficiencyCorrection/2015_2016/rel20.7/ICHEP_June2016_v3/offline/efficiencySF.offline.RecoTrk.root`
- isolation: `ElectronEfficiencyCorrection/2015_2016/rel20.7/ICHEP_June2016_v3/isolation/efficiencySF.Isolation.TightLLH_d0z0_v11_isolTight.root`
- ID: `ElectronEfficiencyCorrection/2015_2016/rel20.7/ICHEP_June2016_v3/offline/efficiencySF.offline.TightLLH_d0z0_v11.root`
- trigger: `ElectronEfficiencyCorrection/2015_2016/rel20.7/ICHEP_June2016_v3/trigger/efficiencySF.SINGLE_E_2015_e24_lhmedium_L1EM20VH_OR_e60_lhmedium_OR_e120_lhloose_2016_e26_lhtight_nod0_ivarloose_OR_e60_lhmedium_nod0_OR_e140_lhloose_nod0.TightLLH_d0z0_v11_isolTight.root`

After the full event selection 344147 events are observed in data.

D.2 Formalism and methodology

D.2.1 Definition of inter-calibration regions

The in-situ calibration is performed in bins of pseudorapidity η_{calo} of the forward electron to cover the inhomogeneities of the Electromagnetic CALorimeter (ECAL) and Forward Calorimeter (FCal). The forward electrons are labelled according to the η_{calo} bin they fall

Table D.1: Absolute values of bin thresholds in η_{calo} for energy scale factors and resolution additional constant terms for Run 2. The same thresholds are used for both sides of the detector.

EMEC	FCal
2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.16	3.35, 3.6, 4.0, 4.9

in. Due to statistical limitations the bins are not subdivided in other variables, neither for the central nor the forward electron.

For Run 2 the same binning as used in Run 1 was chosen. When reducing the bins width, the existence of pattern has been observed in some detector regions for the energy scale factors. In the η_{calo} range covered by the EMEC the bin width is in general $\Delta\eta_{\text{calo}} = 0.1$. Having a wider resolution in terms of η_{calo} and less statistics, the region covered by the FCal is divided into only 4 bins on each side of the detector, yielding significantly wider bins. No scales are provided in the transition region between the two subdetectors $3.16 < |\eta_{\text{calo}}| < 3.35$. The binning used for the scale factors computation is given in D.1.

D.2.2 Tag-and-probe approach

The energy scale factor α scales the measured energy of data to correct for residual miscalibration according to the following parametrisation in each region i of the pseudo-rapidity of the forward electron by

$$E_i^{\text{data}} = E_i^{\text{MC}}(1 + \alpha_i) \quad (\text{D.1})$$

where E_i^{data} and E_i^{MC} are the electron energies in data and simulation and α_i represents the relative deviation of the energy in data from the MC energy. The MC is assumed to be perfectly calibrated.

The additional constant term c'_i is used to smear the energy of the MC electrons according to:

$$E_i^{\text{data}} = E_i^{\text{MC}}(1 + c'_i \times \mathcal{N}(0, 1)) \quad (\text{D.2})$$

with $\mathcal{N}(0, 1)$ being a Gaussian distributed random number. The calibration for the central electron is applied fully, any differences in the m_{ee} distribution are therefore assumed to be caused by the forward electron. Taking this into account equations can be derived to relate these correction coefficients with properties of the distribution of the invariant mass of the ee pair m_{ee} in a η_{calo} bin of the forward electron i :

$$\alpha_i = \left(\frac{m_{ee}^{\text{data}}}{m_{ee}^{\text{MC}}} \right)^2 - 1 \quad (\text{D.3})$$

$$c'_i = \sqrt{\left(\left(\frac{2\sigma(m_{ee})}{m_{ee}} \right)^{\text{data}} \right)^2 - \left(\left(\frac{2\sigma(m_{ee})}{m_{ee}} \right)^{\text{MC}} \right)^2}. \quad (\text{D.4})$$

In case the MC distribution is wider than the distribution in data, i.e. the term in the root is negative, c'_i would be a complex number. In order to avoid the additional complexity and to allow for easy plotting c'_i is then defined to be negative with the same absolute value, i.e. it is multiplied by i . In the application of the scales on MC simulated samples, these negative values are ignored and no smearing correction is applied.

Usual propagation of uncertainties yields

$$\Delta\alpha_i = 2 \cdot \left(\frac{m_{ee}^{\text{data}}}{m_{ee}^{\text{MC}}} \right)^2 \cdot \sqrt{\left(\left(\frac{\Delta m_{ee}}{m_{ee}} \right)^{\text{data}} \right)^2 + \left(\left(\frac{\Delta m_{ee}}{m_{ee}} \right)^{\text{MC}} \right)^2} \quad (\text{D.5})$$

$$\Delta c'_i = \frac{4}{c'_i} \cdot \sqrt{\left(\left(\frac{\sigma(m_{ee})}{m_{ee}^2} \right)^{\text{data}} \cdot \Delta \sigma(m_{ee})^{\text{data}} \right)^2 + \left(\left(\frac{\sigma(m_{ee})^2}{m_{ee}^3} \right)^{\text{data}} \cdot \Delta m_{ee}^{\text{data}} \right)^2 + \left(\left(\frac{\sigma(m_{ee})}{m_{ee}^2} \right)^{\text{MC}} \cdot \Delta \sigma(m_{ee})^{\text{MC}} \right)^2 + \left(\left(\frac{\sigma(m_{ee})^2}{m_{ee}^3} \right)^{\text{MC}} \cdot \Delta m_{ee}^{\text{MC}} \right)^2} \quad (\text{D.6})$$

By comparing the shapes of the invariant mass of the ee pair m_{ee} correction coefficients can be calculated for different η_{calo} ranges of the forward electron.

D.2.3 Fitting procedure

In order to quantify and compare the distributions, they are fitted using an analytical function. This fitting function is built around a Breit-Wigner distribution to model the shape of the Z peak. This distribution is then convoluted with a left-sided Crystal-Ball function to take the detector resolution and Bremsstrahlung effects into account, yielding the signal model:

$$f_{\text{signal}} = f_{\text{BW}} \otimes f_{\text{CB}} \quad (\text{D.7})$$

The Crystal-Ball is restricted by forcing the pole of the left-sided power tail to coincide with the mean of the Gaussian part. This restriction removes one degree of freedom between two highly anti-correlated parameters for the subsequent fits and consequently largely decreases the CPU and time requirements.

To model the contamination of several backgrounds in data, an additional function described by a second order Chebychev polynomial is added only for fits to data

$$f_{\text{signal} + \text{bkg}} = f_{\text{BW}} \otimes f_{\text{CB}} + f_{\text{bkg}} \quad (\text{D.8})$$

The necessary p_{T} requirements on both leptons combined with a minimum $\Delta\eta$ depending on the considered η region i lead to a decrease of events especially in the lower tail of the nominal Z resonance. This can be seen in distortions of the Z peak shape in MC simulated samples without simulation of detector effects. A dedicated truth study of the m_{ee} dependence of this efficiency showed a linear behaviour around the Z peak. For this the ratio of the distribution in two different phase spaces was studied. In the total phase space no p_{T} requirements are imposed on the electrons. So for all events with exactly one true central electron and one true forward electron, the m_{ee} distribution is considered in bins i of $\eta_{\text{calo, probe}}$ of the true forward electron. In the true fiducial phase space the additional requirements of $p_{\text{T}} > 25$ GeV (20 GeV) for true central (forward) electrons are added. The efficiency is then defined as the ratio of the sum of weights of simulated events for a given m_{ee} and $\eta_{\text{calo, probe}}$ in the fiducial phase space over the total phase space. The dependence of the efficiency on the invariant mass m_{ee} is then fitted with a linear function

$$f'_{\text{eff}} = s \cdot m_{ee} + c. \quad (\text{D.9})$$

The distribution of this efficiency including its fit for an example eta range can be seen in D.1 and the fit results for all $\eta_{\text{calo, probe}}$ ranges are listed in D.2.

This efficiency is accounted for in the fits by multiplying the result of this linear fit to the fit models

$$f_{\text{signal} (+ \text{bkg})}^{\text{full}} = f_{\text{signal} (+ \text{bkg})} \cdot f_{\text{eff}}. \quad (\text{D.10})$$

Here f_{eff} includes cutoffs for the linear fits f'_{eff} to avoid non-physical values for the efficiency. Values larger than 1 are set to 1 and negative values are set to 0.

This defines the full nominal fit models with which extended maximum likelihood fits are performed in a fit range of $70 \text{ GeV} \leq m_{ee} \leq 120 \text{ GeV}$. $f_{\text{signal}}^{\text{full}}$ is used for fits of the pure signal shape described by MC simulations of $Z \rightarrow ee$ events and $f_{\text{signal} + \text{bkg}}^{\text{full}}$ is used for fits to the distribution in data. When fitting data in a given $\eta_{\text{calo, probe}}$ bin, all parameters of the signal model f_{signal} other than the mean value of the Breit-Wigner and the width of the Crystal-Ball are fixed to their maximum likelihood estimators from the fit of the signal

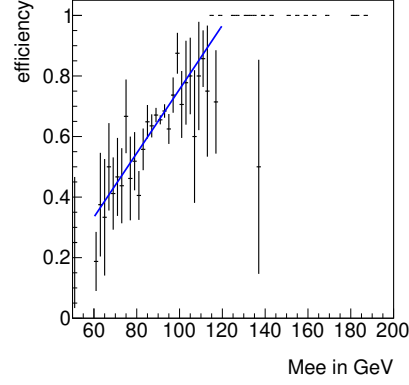


Figure D.1: Efficiency of p_T selection requirements estimated in truth level study including linear fit.

Table D.2: List of fit results for m_{ee} dependency of efficiency

$\eta_{\text{calo, probe}}$ range	slope s	offset c / GeV
$[-4.9, -4.0]$	$0.000255 \pm 3.8 \cdot 10^{-05}$	-0.0160 ± 0.0034
$[-4.0, -3.6]$	$0.000300 \pm 2.6 \cdot 10^{-05}$	-0.0165 ± 0.0023
$[-3.6, -3.35]$	$0.000257 \pm 2.3 \cdot 10^{-05}$	-0.0128 ± 0.0022
$[-3.16, -3.0]$	$0.000230 \pm 2.5 \cdot 10^{-05}$	-0.0108 ± 0.0022
$[-3.0, -2.9]$	$0.000104 \pm 2.1 \cdot 10^{-05}$	-0.0026 ± 0.0019
$[-2.9, -2.8]$	$0.000174 \pm 2.2 \cdot 10^{-05}$	-0.0085 ± 0.0019
$[-2.8, -2.7]$	$0.000164 \pm 2.3 \cdot 10^{-05}$	-0.0072 ± 0.0020
$[-2.7, -2.6]$	$0.000160 \pm 2.7 \cdot 10^{-05}$	-0.0057 ± 0.0024
$[-2.6, -2.5]$	$0.000152 \pm 2.6 \cdot 10^{-05}$	-0.0042 ± 0.0023
$[2.5, 2.6]$	$0.000096 \pm 3.1 \cdot 10^{-05}$	0.0006 ± 0.0028
$[2.6, 2.7]$	$0.000167 \pm 2.6 \cdot 10^{-05}$	-0.0064 ± 0.0023
$[2.7, 2.8]$	$0.000106 \pm 2.8 \cdot 10^{-05}$	-0.0010 ± 0.0025
$[2.8, 2.9]$	$0.000151 \pm 2.5 \cdot 10^{-05}$	-0.0060 ± 0.0022
$[2.9, 3.0]$	$0.000173 \pm 2.2 \cdot 10^{-05}$	-0.0087 ± 0.0019
$[3.0, 3.16]$	$0.000182 \pm 2.9 \cdot 10^{-05}$	-0.0067 ± 0.0026
$[3.35, 3.6]$	$0.000233 \pm 2.7 \cdot 10^{-05}$	-0.0102 ± 0.0024
$[3.6, 4.0]$	$0.000337 \pm 2.1 \cdot 10^{-05}$	-0.0199 ± 0.0019
$[4.0, 4.9]$	$0.000323 \pm 2.8 \cdot 10^{-05}$	-0.0221 ± 0.0024

model to the signal MC. This ensures that differences between the distributions are covered by these two parameters and the additional background function.

The correction coefficients α_i and c'_i are calculated from the fit results using D.3 and D.4 where the mean of the Breit-Wigner μ_{BW} is used to substitute m_{ee} and the width of the Crystal-Ball σ_{CB} to substitute $\sigma(m_{ee})$:

$$\alpha_i = \left(\frac{\mu_{\text{BW}}^{\text{data}}}{\mu_{\text{BW}}^{\text{MC}}} \right)^2 - 1 \quad (\text{D.11})$$

$$c'_i = \sqrt{\left(\left(\frac{2\sigma_{\text{CB}}}{\mu_{\text{BW}}} \right)^{\text{data}} \right)^2 - \left(\left(\frac{2\sigma_{\text{CB}}}{\mu_{\text{BW}}} \right)^{\text{MC}} \right)^2}. \quad (\text{D.12})$$

The statistical uncertainties of both fit parameters are propagated and treated as statistical uncertainties of the correction coefficients. These uncertainties can be calculated using the equivalent substitutions from D.5 and D.6.

Fit stability In order to increase fit stability and minimize dependence of the fit results on the initial values, multiple fits are performed for each distribution. Each distribution is fitted 100 times with random initial values for all free parameters of the model. For each fit result χ^2 is calculated between the fitted model and the input distribution and the iteration minimizing χ^2 is chosen as final fit result.

Some of these iterations are declared invalid in order to avoid known causes of overestimation of uncertainties. In these cases the iteration is not counted towards the 100 iteration and is restarted with different random initial conditions. The main reason for this is if one or more of the fitted parameters is too close to its threshold. Minimizers often find these points as local minima and uncertainty estimations are not reliable in this case. To avoid this a fit is declared invalid if the distance from each of the thresholds to the resulting value is less than 5 % if the total range of the parameter. A fit is also declared invalid if the absolute value of the correlation between the mean of the Breit-Wigner and the width of the Crystal-Ball is larger than 0.98, or if the relative uncertainty of one of these two parameters is larger than 4 %, indicating large correlations to another fit parameter. These correlations are not considered in the calculation of the correction coefficients, rendering a suppression useful.

D.3 Systematic effects and results

D.3.1 Systematic uncertainties

Several sources of systematic uncertainties are considered in the estimation of the correction coefficients. The effect of a given source on the coefficients are estimated by rerunning the full selection and fit procedure with varied settings for this source of systematic uncertainties. The differences to the nominal setup are considered as systematic uncertainty and are combined as discussed below. The total systematic uncertainty is calculated as the quadratic sum of the individual systematic uncertainties for all considered sources.

These are the considered sources of systematics:

Fit window To estimate the effect of the choice of the fit window it is varied from its nominal value $70 \text{ GeV} \leq m_{ee} \leq 120 \text{ GeV}$ to $65 \text{ GeV} \leq m_{ee} \leq 125 \text{ GeV}$ and $83 \text{ GeV} \leq m_{ee} \leq 110 \text{ GeV}$ in two independent runs. The maximal deviation from the nominal run is used as a systematic uncertainty

Background modelling To cover variations due to bad modelling of the background contributions, two alternative background models are used as replacement of the nominal one (2nd order Chebychev polynomial), namely 2nd order exponential or Landau distributions. The deviations from the nominal correction factors are then added in quadrature and treated as a systematic uncertainty.

Central calibration The full selection is rerun using variations of the central calibration defined in the decorrelated model `1NPCOR_PLUS_UNCOR` and the fit is repeated for each of the resulting distributions. The deviations from the nominal correction factors are then added in quadrature and treated as a systematic uncertainty.

Central efficiencies The full set of variations for the central identification, reconstruction, ID, and trigger efficiencies are applied to cover mismodellings of the connected efficiencies in simulations and combined and the selection is rerun for each variation. The deviations of the resulting correction factors from the nominal ones are added in quadrature and included as a systematic uncertainty.

Signal modelling: tails The choice of a left-sided Crystal-Ball to account for detector effects is tested by exchanging it with a Crystal-Ball distribution a power tail on each side of the peak as a variation.

Signal modelling: Crystal Ball As an additional variation for the modelling of the signal the Crystal-Ball distribution is replaced by an exponential Gaussian distribution as introduced in [182].

Modelling in data In order to cover effects of modelling problems in data an additional variation is used where the parameters of the signal model are not fixed to the fit results in MC.

The two different contributions to the modelling of the signal process (tails and Crystal Ball) are not combined in order here to disentangle the individual contributions later. All these sources of systematics are treated as uncorrelated to calculate the total systematic uncertainty.

D.3.2 Results

The scale factors are measured with the method described above, using the full 2015 dataset. All fit results are shown in [181]. The final values are shown in [181]. The contributions of the individual systematic uncertainties to the combined systematic uncertainty are shown for the energy scale factor α in D.3 and for the additional constant term c' in D.4.

The energy scale factor is fairly smooth in $|\eta_{\text{calo, probe}}|$, except for two ranges. In these two regions the systematic uncertainties are also significantly higher than in the neighbouring bins. These ranges were already problematic in the determination of scale factors in run 1. For both run 1 and run 2 the distribution of data wasn't described well by the signal model. This could be caused by additional material in front of the detector leading to additional distortion, but this wasn't understood until now. The fit window variation was chosen so that the connected systematic covers the resulting discrepancies.

The additional constant term c' is relatively flat in $|\eta_{\text{calo, probe}}|$, other than the most central bins where it is consistent with 0.

For both scale factors the uncertainties are dominated by systematic effects in all bins. In combination with the data collected in 2016 it should be feasible to bin these scale factors also in $p_{\text{T, probe}}$ or φ_{probe} . A finer binning in $\eta_{\text{calo, probe}}$ in the EMEC region is not possible due to the coarser η granularity of the detector in this region.

D.4 Validation

To validate the extracted scales and their uncertainties the selection is rerun including a calibration for the forward electrons using the extracted scales. For the validation of the uncertainties the selection is run again with the varied values for scales. The fitted background contribution is subtracted from the data distribution and the resulting distribution is compared to the pure signal model estimated using MC. This comparison is shown with different corrections applied in D.5 for one of the regions in $\eta_{\text{calo, probe}}$ with outliers for the scales. A full overview of these plots for all ranges can be found in [181].

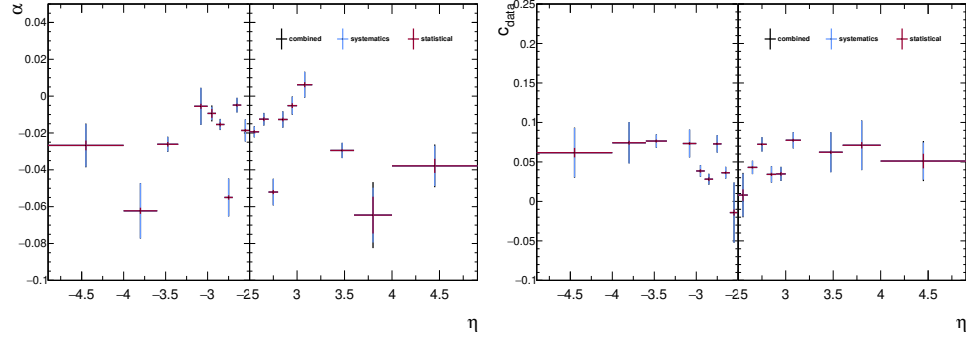


Figure D.2: Energy scale factors α (left) and additional constant terms c' as a function of $\eta_{\text{calo, probe}}$. The combination of the systematic uncertainties (blue) and the statistical uncertainty (red) and their combination (black) are shown.

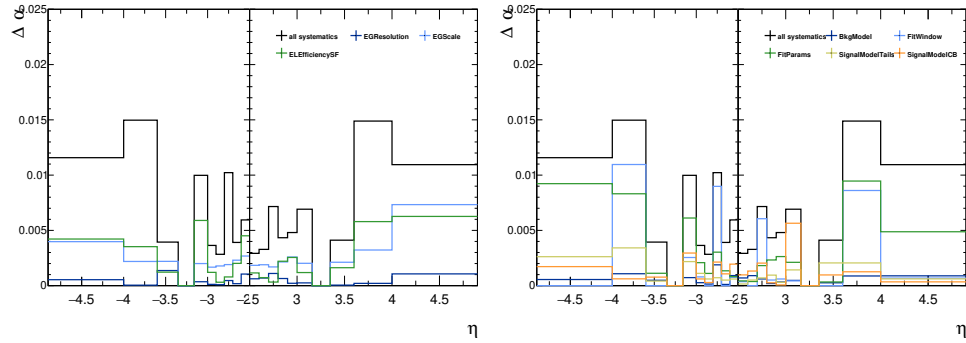


Figure D.3: Contributions to combined systematic uncertainty of the energy scale factors α (black) caused by event selection (left) and fit procedure (right) as a function of $\eta_{\text{calo, probe}}$. For the sake of visibility the individual contributions were splitted in two plots, while the combination includes the full set of systematics for both plots. In the left plot the shows the resolution uncertainty (deep blue) and scale uncertainty (light blue) of the calibration of the central electron and the combination of all efficiency scale factors (green). The right plot shows the shows the background modeling (deep blue), fit window (light blue), modeling in data (green) uncertainties as well as the variations of the signal model for the tails (yellow) and the definition of the crystal ball (orange).

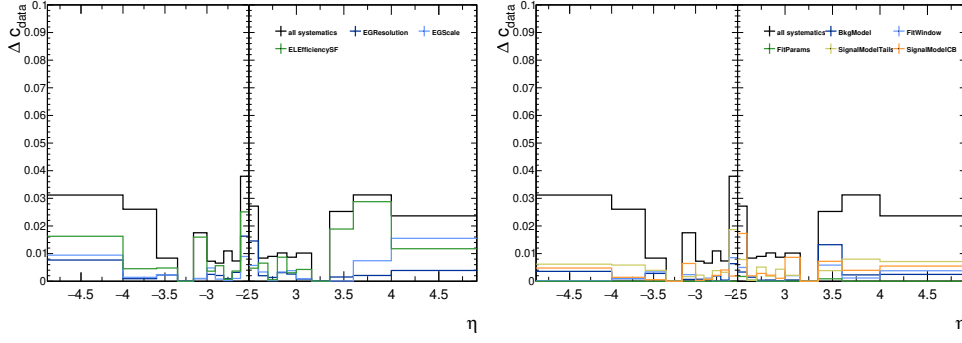


Figure D.4: Contributions to combined systematic uncertainty of the additional constant terms c' (black) caused by event selection (left) and fit procedure (right) as a function of $\eta_{\text{calo, probe}}$. For the sake of visibility the individual contributions were splitted in two plots, while the combination includes the full set of systematics for both plots. In the left plot the shows the resolution uncertainty (deep blue) and scale uncertainty (light blue) of the calibration of the central electron and the combination of all efficiency scale factors (green). The right plot shows the background modeling (deep blue), fit window (light blue), modeling in data (green) uncertainties as well as the variations of the signal model for the tails (yellow) and the definition of the crystal ball (orange).

The application of the nominal scales is expected to remove most discrepancies visible in the raw distributions. Any remaining discrepancies should be covered by the systematic uncertainties introduced by the forward calibration scales. Comparing the distributions shown in D.5, the extracted scales fulfil these expectations. After applying the smearing correction, the width of the Z peak is compatible between data and MC simulation and after the scale correction is applied the mean value is in agreement as well. Due to the quite large uncertainties of the scales in this $\eta_{\text{calo, probe}}$ region the discrepancy in the shape is not significant. Comparing also to the fit result for data (see auxiliary material in [181]) the large discrepancy can already be found there. So in order to reduce the uncertainty on the scales the fit in data has to be improved. This is consistent with the large observed signal modelling systematic uncertainties for this $\eta_{\text{calo, probe}}$ region. The fact that the discrepancies can be explained by imperfect fits indicates that the calculation and application of the scales works as expected.

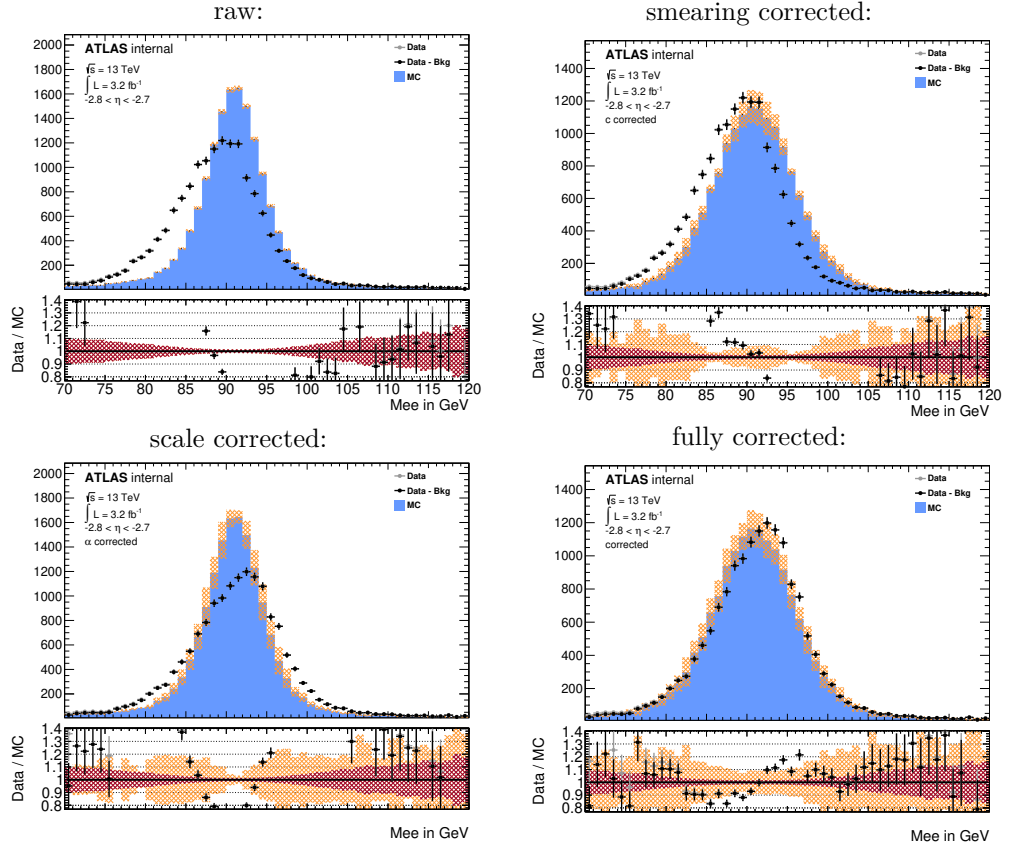


Figure D.5: Pure data and background subtracted data distribution compared to signal MC distribution scaled to integral of data for m_{ee} for $-2.8 < \eta_{\text{calo, probe}} < -2.7$ with different corrections applied. Inset shows the (data - bkg) / MC (black) and data / MC (grey) ratios including the statistical uncertainties of the distributions (red) and the systematic uncertainties caused by the applied forward calibration scales (orange) on the MC prediction. Note that due to the low contribution of background resulting from the fits the grey and black distributions are very close to each other. Top left plot shows the uncalibrated distributions. In both plots on the right-hand side the smearing correction is applied and the energy scale correction is applied in the two lower plots. The shown uncertainties do not include experimental uncertainties on top of the extracted scales, these would lead to less significant deviations.

Appendix E

Data-driven Non-prompt Estimate

E.1 Object Selection

An overview of the applied selection criteria including the selection for the data-driven estimate is presented in Table E.1 for electrons and in Table E.2 for muons. The individual criteria are discussed in detail in Section 5.2 or for the control selection in Section 6.2.1.

Table E.1: Overview of different sets of object selection criteria for electrons. The overlap removal criteria correspond to steps 2 to 4 listed in Section 5.2.4.

Object selection criterium	e selection				
	Baseline	Z-control	W-control	Z-analysis	W-analysis
Impact parameter $ z_0 \cdot \sin(\theta) < 0.5 \text{ mm}$	✓	✓	✓	✓	✓
d_0 significance $ d_0/\sigma_{d_0} < 5$	✓	✓	✓	✓	✓
Object quality	✓	✓	✓	✓	✓
Identification	LooseAndBLayer	LooseAndBLayer	LooseAndBLayer	Medium Gradient	Tight Gradient
Isolation	LooseTrackOnly $> 5 \text{ GeV}$	$p_T^{\text{varcone30}}/p_T < 0.5$ $> 15 \text{ GeV}$	$p_T^{\text{varcone30}}/p_T < 0.5$ $> 20 \text{ GeV}$	$> 15 \text{ GeV}$	$> 20 \text{ GeV}$
Transverse momentum p_T	✓	✓	✓	✓	✓
Track-based pseudo-rapidity $ \eta < 2.5$	✓	✓	✓	✓	✓
Calorimeter-based pseudo-rapidity $ \eta_{\text{calo}} < 2.47$	✓	✓	✓	✓	✓
$1.37 \leq \eta_{\text{calo}} \leq 1.52$		✓	✓	✓	✓
Ambiguous veto					✓
Overlap removal					
baseline $e\text{-}\mu$	✓			✓	✓
$e\text{-}e$				✓	✓
$e\text{-}j$				✓	✓
Analysis leptons veto		not in Z-analysis	not in W-analysis		

Table E.2: Overview of different sets of object selection criteria for muons. The overlap removal criteria correspond to steps 2 and 4 listed in Section 5.2.4.

Object selection criterium	Baseline	μ selection			
		Z-control	W-control	Z-analysis	W-analysis
Impact parameter $ z_0 \cdot \sin(\theta) < 0.5 \text{ mm}$	✓	✓	✓	✓	✓
d_0 significance $ d_0/\sigma_{d_0} < 3$	✓			✓	✓
Identification	Loose	Medium	Medium	Medium	Tight
Isolation	FixedCutLoose	$p_T^{\text{varcone30}}/p_T < 0.5$	$p_T^{\text{varcone30}}/p_T < 0.5$	Gradient	Gradient
Transverse momentum p_T	$> 5 \text{ GeV}$	$> 15 \text{ GeV}$	$> 20 \text{ GeV}$	$> 15 \text{ GeV}$	$> 20 \text{ GeV}$
Absolute pseudo-rapidity $ \eta $	< 2.7	< 2.5	< 2.5	< 2.5	< 2.5
Overlap removal					
baseline e - μ	✓			✓	✓
μ - j				✓	✓
Analysis leptons veto		not in Z-analysis	not in W-analysis		

E.2 Monte-Carlo simulation based Closure Test

E.3 Further details on matrix-method results

Table E.3: Overview of predicted number of events for an integrated luminosity of 36.1 fb^{-1} per process in the $t\bar{t}$ control region split by lepton channel for the simulation-based prediction (MC) as well as the matrix-method prediction (MM). Additional rows on the bottom show the total event yield for each of the prediction methods as well as the ratio of the matrix-method with respect to the simulation-based prediction. Uncertainties are split in the statistical uncertainties due to the limited number of simulated events and systematic uncertainties of the matrix-method.

	eee		$e\mu\mu$		μee		$\mu\mu\mu$		all	
MC top	0 ± 0	0^{+0}_{-0}	48	± 4	0^{+0}_{-0}	70	± 5	0 ± 0	118	± 6
MC VV	0 ± 0	0^{+0}_{-0}	0.067	± 0.029	0^{+0}_{-0}	0.47 ± 0.18	0^{+0}_{-0}	0 ± 0	0.53 ± 0.18	0^{+0}_{-0}
MC $V\gamma$	0 ± 0	0^{+0}_{-0}	0	± 0	0^{+0}_{-0}	0.4 ± 0.4	0^{+0}_{-0}	0 ± 0	0.4 ± 0.4	0^{+0}_{-0}
MC V +jets	0 ± 0	0^{+0}_{-0}	0	± 0	0^{+0}_{-0}	0	± 0	0 ± 0	0	± 0
MM top	0 ± 0	0^{+0}_{-0}	43	± 1	0^{+0}_{-0}	64 ± 2	0^{+0}_{-0}	0 ± 0	107	± 2
MM VV	0 ± 0	0^{+0}_{-0}	0.071	± 0.012	$0^{+0.011}_{-0.011}$	0.34 ± 0.05	$0^{+0.14}_{-0.04}$	0 ± 0	0.41 ± 0.05	$0^{+0.15}_{-0.05}$
MM $V\gamma$	0 ± 0	0^{+0}_{-0}	$-3.8e-04$	$\pm 3.2e-04$	$0^{+9e-05}_{-1.0e-04}$	0.03 ± 0.04	$0^{+0.00}_{-0.00}$	0 ± 0	0.02 ± 0.04	$0^{+0.00}_{-0.00}$
MM V +jets	0 ± 0	0^{+0}_{-0}	2	± 2	0^{+0}_{-0}	2	± 2	0 ± 0	4	± 2
total MC	0 ± 0	0^{+0}_{-0}	48	± 4	0^{+0}_{-0}	71	± 5	0 ± 0	119	± 6
total MM	0 ± 0	0^{+0}_{-0}	44	± 2	0^{+7}_{-6}	67	± 2	0 ± 0	111	± 3
ratio MM / MC	$nan \pm nan$	$nan \pm nan$	0.93	± 0.09	$0^{+0.14}_{-0.14}$	0.94 ± 0.07	$0^{+0.29}_{-0.29}$	$nan \pm nan$	0.93 ± 0.06	$0^{+0.21}_{-0.21}$

Table E.4: Overview of predicted number of events for an integrated luminosity of 36.1 fb^{-1} per process in the $\ell\ell\nu jj$ region split by lepton channel for the simulation-based prediction (MC) as well as the matrix-method prediction (MM). Additional rows on the bottom show the total event yield for each of the prediction methods as well as the ratio of the matrix-method with respect to the simulation-based prediction. Uncertainties are split in the statistical uncertainties due to the limited number of simulated events and systematic uncertainties of the matrix-method.

	eee		$e\mu\mu$		μee		$\mu\mu\mu$		all	
MC top	35	± 3	$+0$ -0	± 5	$+0$ -0	67	± 4	$+0$ -0	232	± 8
MC VV	0.98 ± 0.26	$+0$ -0	± 0.6	± 5	$+0$ -0	0.21 ± 0.08	± 0.13	$+0$ -0	2.8	± 0.7
MC $V\gamma$	12	± 4	$+0$ -0	± 5	$+0$ -0	0.6 ± 0.6	± 1	$+0$ -0	38	± 7
MC V +jets	0	± 0	$+0$ -0	± 3	$+0$ -0	0	± 0	$+0$ -0	12	± 4
MM top	37	± 2	$+9$ -5	± 1	$+20$ -9	61	± 2	$+9$ -4	210	± 0
MM VV	0.44 ± 0.09	$+0.28$ -0.06	± 0.07	± 1	$+0.31$ -0.06	0.17 ± 0.06	± 0.013	$+0.014$ -0.013	1.2	± 0.1
MM $V\gamma$	4	± 1	$+3$ -1	± 1	$+7$ -1	0.26 ± 0.08	± 0.4	$+0.2$ -0.2	15	± 1
MM V +jets	5	± 2	$+5$ -1	± 1	$+4$ -1	0.71 ± 0.30	± 0.4	$+0.3$ -0.3	13	± 3
total MC	47	± 5	$+0$ -0	± 8	$+0$ -0	68	± 4	$+0$ -0	285	± 11
total MM	47	± 3	$+17$ -7	± 2	$+31$ -10	62	± 2	$+5$ -4	240	± 0
ratio MM / MC	1.0	± 0.1	$+0.4$ -0.4	± 0.06	$+0.24$ -0.24	0.92 ± 0.07	± 0.10	$+0.11$ -0.11	0.86	± 0.04

Table E.5: Overview of predicted number of events for an integrated luminosity of 36.1 fb^{-1} per process in the $WZjj$ region split by lepton channel for the simulation-based prediction (MC) as well as the matrix-method prediction (MM). Additional rows on the bottom show the total event yield for each of the prediction methods as well as the ratio of the matrix-method with respect to the simulation-based prediction. Uncertainties are split in the statistical uncertainties due to the limited number of simulated events and systematic uncertainties of the matrix-method.

	eee		$e\mu\mu$		μee		$\mu\mu\mu$		all	
MC top	2.0	$\pm 0.7^{+0}_{-0}$	4	$\pm 1^{+0}_{-0}$	2.5	$\pm 0.8^{+0}_{-0}$	3.4	$\pm 0.9^{+0}_{-0}$	12	$\pm 2^{+0}_{-0}$
MC VV	0.20	$\pm 0.10^{+0}_{-0}$	0.07	$\pm 0.07^{+0}_{-0}$	0.008	$\pm 0.008^{+0}_{-0}$	0.006	$\pm 0.006^{+0}_{-0}$	0.29	$\pm 0.12^{+0}_{-0}$
MC $V\gamma$	1.5	$\pm 0.8^{+0}_{-0}$	3.1	$\pm 0.8^{+0}_{-0}$	0.06	$\pm 0.06^{+0}_{-0}$	0	$\pm 0^{+0}_{-0}$	5	$\pm 1^{+0}_{-0}$
MC V +jets	0	$\pm 0^{+0}_{-0}$	4	$\pm 2^{+0}_{-0}$	0	$\pm 0^{+0}_{-0}$	0	$\pm 0^{+0}_{-0}$	4	$\pm 2^{+0}_{-0}$
MM top	3.0	$\pm 0.4^{+0.6}_{-0.4}$	3.9	$\pm 0.3^{+0.9}_{-0.4}$	2.7	$\pm 0.4^{+0.4}_{-0.4}$	2.75	$\pm 0.20^{+0.34}_{-0.31}$	12	$\pm 1^{+2}_{-2}$
MM VV	0.10	$\pm 0.03^{+0.06}_{-0.01}$	0.11	$\pm 0.02^{+0.09}_{-0.02}$	0.019	$\pm 0.007^{+0.003}_{-0.003}$	0.023	$\pm 0.008^{+0.003}_{-0.003}$	0.25	$\pm 0.04^{+0.16}_{-0.04}$
MM $V\gamma$	0.57	$\pm 0.21^{+0.17}_{-0.08}$	2	$\pm 0^{+1}_{-0}$	0.005	$\pm 0.012^{+0.001}_{-0.001}$	0.34	$\pm 0.29^{+0.04}_{-0.03}$	2	$\pm 0^{+1}_{-0}$
MM V +jets	3	$\pm 2^{+3}_{-0}$	2	$\pm 0^{+1}_{-0}$	0.10	$\pm 0.10^{+0.02}_{-0.01}$	0.28	$\pm 0.07^{+0.04}_{-0.04}$	5	$\pm 2^{+4}_{-1}$
total MC	4	$\pm 1^{+0}_{-0}$	12	$\pm 3^{+0}_{-0}$	2.5	$\pm 0.8^{+0}_{-0}$	3.4	$\pm 0.9^{+0}_{-0}$	21	$\pm 3^{+0}_{-0}$
total MM	7	$\pm 2^{+4}_{-1}$	7	$\pm 1^{+3}_{-1}$	2.8	$\pm 0.4^{+0.4}_{-0.4}$	3.4	$\pm 0.4^{+0.4}_{-0.4}$	20	$\pm 2^{+8}_{-3}$
ratio MM / MC	2	$\pm 1^{+1}_{-1}$	0.64	$\pm 0.19^{+0.29}_{-0.29}$	1.1	$\pm 0.4^{+0.2}_{-0.2}$	1.00	$\pm 0.29^{+0.12}_{-0.12}$	1.0	$\pm 0.2^{+0.4}_{-0.4}$

Table E.6: Overview of predicted number of events for an integrated luminosity of 36.1 fb^{-1} per process in the $\ell\ell\nu jj$ region split by lepton channel for the simulation-based prediction (MC) as well as the data-driven matrix-method prediction (MM). The simulation-based prediction is split according to the process, while the data-driven matrix-method prediction is split according to the detector-level classification (A/C), i.e. which selection stage each lepton candidate passed. Additional rows on the bottom show the total event yield for each of the prediction methods as well as the ratio of the matrix-method with respect to the simulation-based prediction. Uncertainties are split in the statistical uncertainties due to the limited number of simulated events and systematic uncertainties of the matrix-method.

	eee		$e\mu\mu$		μee		$\mu\mu\mu$		all	
MC top	35	± 3	$^{+0}_{-0}$	± 5	$^{+0}_{-0}$	67	± 4	$^{+0}_{-0}$	232	± 8
MC $V + \gamma$	12	± 4	$^{+0}_{-0}$	± 5	$^{+0}_{-0}$	0.6	± 0.6	$^{+0}_{-0}$	38	± 7
MC VV	0.98	± 0.26	$^{+0}_{-0}$	± 0.6	$^{+0}_{-0}$	0.21	± 0.08	$^{+0}_{-0}$	2.8	± 0.7
MC $V + \text{jets}$	0	± 0	$^{+0}_{-0}$	± 3	$^{+0}_{-0}$	0	± 0	$^{+0}_{-0}$	12	± 4
MM CAA	16	± 1	$^{+10}_{-7}$	± 3	$^{+25}_{-18}$	19	± 1	$^{+7}_{-7}$	90	± 0
MM ACA	7	± 1	$^{+3}_{-2}$	± 1	$^{+4}_{-3}$	16	± 1	$^{+4}_{-4}$	44	± 1
MM AAC	30	± 2	$^{+10}_{-6}$	± 1	$^{+12}_{-8}$	45	± 3	$^{+10}_{-10}$	150	± 0
MM 2C	-0.5	± 0.1	$^{+0.2}_{-0.4}$	± 0.1	$^{+0.3}_{-0.4}$	-0.6	± 0.2	$^{+0.3}_{-0.4}$	2	± 0
total MC	47	± 5	$^{+0}_{-0}$	± 8	$^{+0}_{-0}$	68	± 4	$^{+0}_{-0}$	285	± 11
total MM	52	± 3	$^{+22}_{-15}$	± 0	$^{+40}_{-30}$	79	± 3	$^{+21}_{-21}$	270	± 10
ratio MM / MC	1.1	± 0.1	$^{+0.5}_{-0.5}$	0.77 ± 0.07	$^{+0.33}_{-0.33}$	1.17	± 0.08	$^{+0.31}_{-0.31}$	0.96	± 0.04
										$^{+0.32}_{-0.32}$

Table E.7: Overview of predicted number of events for an integrated luminosity of 36.1 fb^{-1} per process in the $WZjj$ region split by lepton channel for the simulation-based prediction (MC) as well as the data-driven matrix-method prediction (MM). The simulation-based prediction is split according to the process, while the data-driven matrix-method prediction is split according to the detector-level classification (A/C), i.e. which selection stage each lepton candidate passed. Additional rows on the bottom show the total event yield for each of the prediction methods as well as the ratio of the matrix-method with respect to the simulation-based prediction. Uncertainties are split in the statistical uncertainties due to the limited number of simulated events and systematic uncertainties of the matrix-method.

	eee			eμμ		μee		μμμ		all					
MC top	2.0	± 0.7	+0 -0	4	± 1	+0 -0	2.5	± 0.8	+0 -0	3.4	± 0.9	+0 -0	12	± 2	+0 -0
MC V + γ	1.5	± 0.8	+0 -0	3.1	± 0.8	+0 -0	0.06	± 0.06	+0 -0	0	± 0	+0 -0	5	± 1	+0 -0
MC VV	0.20	± 0.10	+0 -0	0.07	± 0.07	+0 -0	0.008	± 0.008	+0 -0	0.006	± 0.006	+0 -0	0.29	± 0.12	+0 -0
MC V+jets	0	± 0	+0 -0	4	± 2	+0 -0	0	± 0	+0 -0	0	± 0	+0 -0	4	± 2	+0 -0
MM CAA	3	± 1	+3 -2	5	± 1	+4 -2	0.75	± 0.11	+0.28 -0.28	0.93	± 0.09	+0.35 -0.35	10	± 2	+7 -4
MM ACA	0.80	± 0.20	+0.27 -0.23	0.77	± 0.09	+0.22 -0.20	0.68	± 0.18	+0.20 -0.19	0.67	± 0.08	+0.18 -0.18	2.9	± 0.3	+0.9 -0.8
MM AAC	3	± 1	+2 -1	2.5	± 0.2	+0.8 -0.5	2.1	± 0.4	+0.5 -0.5	2.4	± 0.3	+0.5 -0.5	10	± 1	+3 -2
MM 2C	-0.09	± 0.04	+0.04 -0.08	-0.041	± 0.018	+0.021 -0.029	-0.031	± 0.021	+0.016 -0.022	-0.033	± 0.005	+0.015 -0.019	-0.19	± 0.05	+0.10 -0.14
total MC	4	± 1	+0 -0	12	± 3	+0 -0	2.5	± 0.8	+0 -0	3.4	± 0.9	+0 -0	21	± 3	+0 -0
total MM	7	± 1	+4 -2	8	± 1	+5 -3	3.5	± 0.5	+1.0 -0.9	3.9	± 0.3	+1.0 -1.0	23	± 2	+10 -7
ratio MM / MC	2	± 0	+1 -1	0.7	± 0.2	+0.4 -0.4	1.4	± 0.4	+0.4 -0.4	1.16	± 0.30	+0.29 -0.29	1.1	± 0.2	+0.5 -0.5

Table E.8: Overview of predicted number of events for an integrated luminosity of 36.1 fb^{-1} per process in the signal region split by lepton channel for the simulation-based prediction (MC) as well as the data-driven matrix-method prediction (MM). The simulation-based prediction is split according to the process, while the data-driven matrix-method prediction is split according to the detector-level classification (A/C), i.e. which selection stage each lepton candidate passed. Additional rows on the bottom show the total event yield for each of the prediction methods as well as the ratio of the matrix-method with respect to the simulation-based prediction. Uncertainties are split in the statistical uncertainties due to the limited number of simulated events and systematic uncertainties of the matrix-method.

	eee		$e\mu\mu$		μee		$\mu\mu\mu$		all						
MC top	0.5	± 0.5	$^{+0}_{-0}$	0.7	± 0.4	$^{+0}_{-0}$	0.29	± 0.29	$^{+0}_{-0}$	0.21	± 0.21	$^{+0}_{-0}$	1.7	± 0.8	$^{+0}_{-0}$
MC $V + \gamma$	0.07	± 0.05	$^{+0}_{-0}$	0.17	± 0.08	$^{+0}_{-0}$	0.06	± 0.06	$^{+0}_{-0}$	0	± 0	$^{+0}_{-0}$	0.30	± 0.11	$^{+0}_{-0}$
MC VV	0	± 0	$^{+0}_{-0}$	0	± 0	$^{+0}_{-0}$	0	± 0	$^{+0}_{-0}$	0	± 0	$^{+0}_{-0}$	0	± 0	$^{+0}_{-0}$
MC V +jets	0	± 0	$^{+0}_{-0}$	0	± 0	$^{+0}_{-0}$	0	± 0	$^{+0}_{-0}$	0	± 0	$^{+0}_{-0}$	0	± 0	$^{+0}_{-0}$
MM CAA	0.31	± 0.08	$^{+0.18}_{-0.17}$	0.5	± 0.5	$^{+0.2}_{-0.2}$	0.056	± 0.020	$^{+0.021}_{-0.021}$	0.15	± 0.03	$^{+0.06}_{-0.06}$	1.0	± 0.5	$^{+0.5}_{-0.5}$
MM ACA	0.012	± 0.021	$^{+0.007}_{-0.003}$	0.20	± 0.05	$^{+0.06}_{-0.05}$	0.07	± 0.05	$^{+0.02}_{-0.02}$	0.057	± 0.023	$^{+0.015}_{-0.015}$	0.34	± 0.08	$^{+0.10}_{-0.09}$
MM AAC	0.19	± 0.09	$^{+0.08}_{-0.05}$	0.49	± 0.10	$^{+0.14}_{-0.09}$	0.26	± 0.12	$^{+0.08}_{-0.07}$	0.29	± 0.07	$^{+0.06}_{-0.06}$	1.22	± 0.19	$^{+0.34}_{-0.27}$
MM 2C	-0.027	± 0.025	$^{+0.013}_{-0.029}$	-0.004	± 0.005	$^{+0.002}_{-0.003}$	-0.003	± 0.006	$^{+0.002}_{-0.002}$	-0.0050	± 0.0019	$^{+0.0022}_{-0.0028}$	-0.040	± 0.027	$^{+0.019}_{-0.035}$
total MC	0.6	± 0.5	$^{+0}_{-0}$	0.9	± 0.4	$^{+0}_{-0}$	0.36	± 0.30	$^{+0}_{-0}$	0.21	± 0.21	$^{+0}_{-0}$	2.0	± 0.8	$^{+0}_{-0}$
total MM	0.49	± 0.13	$^{+0.23}_{-0.21}$	1.2	± 0.5	$^{+0.4}_{-0.3}$	0.38	± 0.13	$^{+0.11}_{-0.11}$	0.49	± 0.08	$^{+0.13}_{-0.13}$	2.5	± 0.6	$^{+0.9}_{-0.8}$
ratio MM / MC	0.8	± 0.8	$^{+0.4}_{-0.4}$	1.3	± 0.8	$^{+0.5}_{-0.5}$	1.1	± 0.9	$^{+0.3}_{-0.3}$	2	± 2	$^{+1}_{-1}$	1.2	± 0.5	$^{+0.4}_{-0.4}$

Appendix F

BDT Input Modelling Validation

F.1 Comparison of Predictions for Signal Process for all Input Candidates

This section collects comparisons of normalized differential predictions for the SHERPA (red) and MG5_AMC-based (orange) signal prediction as well as the prediction for the combined prompt backgrounds (blue) in all input observables in the signal region. Inlays show the ratio of each of the other distributions with respect to the normalized SHERPA-based signal prediction. The indicated uncertainties represent only the statistical uncertainties due to the limited amount of events in the simulations. The observables are introduced in detail in Section 7.2. This comparison is shown for all considered input observables not already shown in Section 7.2 in Figure F.1,

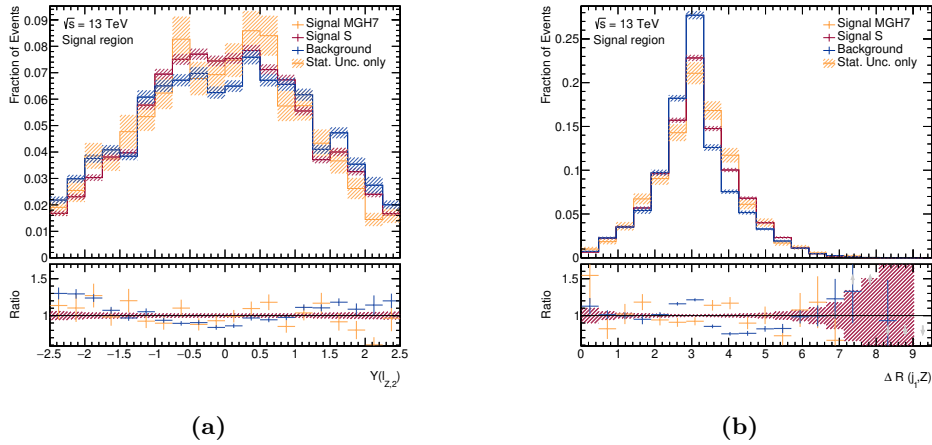


Figure F.1: Comparison of normalized differential predictions for the SHERPA (red) and MG5_AMC-based (orange) signal prediction as well as the prediction for the combined prompt backgrounds (blue) in different input observables in the signal region. Figure (a) shows $Y(\ell_{Z,2})$ and Figure (b) $\Delta R(j_1, Z)$.

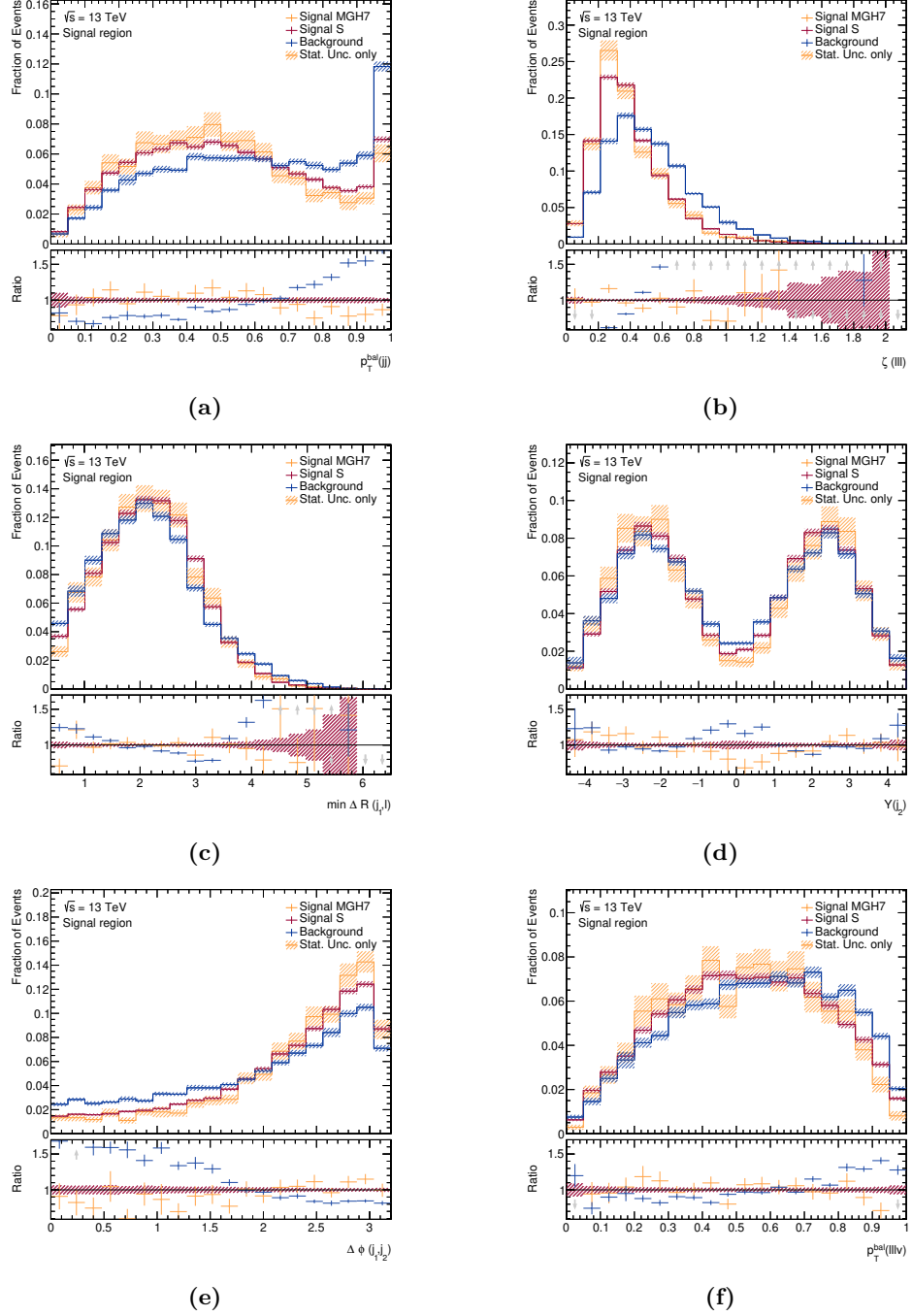


Figure F.2: Comparison of normalized differential predictions for the SHERPA (red) and MG5_AMC-based (orange) signal prediction as well as the prediction for the combined prompt backgrounds (blue) in different input observables in the signal region. Figure (a) shows $p_T^{\text{bal}}(jj)$, Figure (b) shows $\zeta(\ell\ell)$, Figure (c) shows $\min \Delta R(j_1, \ell)$, Figure (d) shows $Y(j_2)$, Figure (e) shows $\Delta\phi(j_1, j_2)$, and Figure (f) shows $p_T^{\text{bal}}(\ell\ell\nu)$.

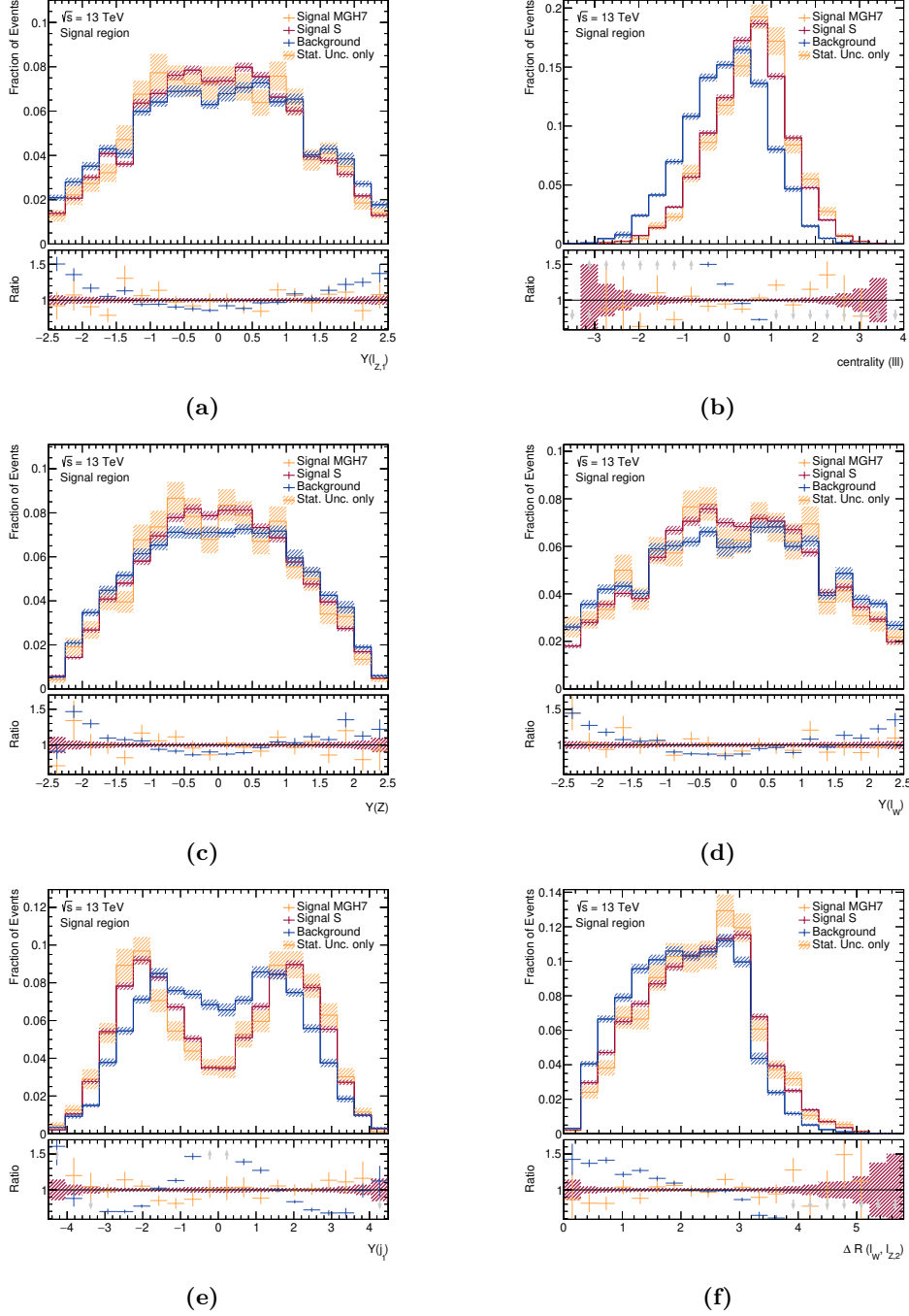


Figure F.3: Comparison of normalized differential predictions for the SHERPA (red) and MG5_AMC-based (orange) signal prediction as well as the prediction for the combined prompt backgrounds (blue) in different input observables in the signal region. Figure (a) shows $Y(\ell_{Z,1})$, Figure (b) shows centrality($\ell\ell$), Figure (c) shows $Y(Z)$, Figure (d) shows $Y(\ell_W)$, Figure (e) shows $Y(j_1)$, and Figure (f) shows $\Delta R(\ell_W, \ell_{Z,2})$.

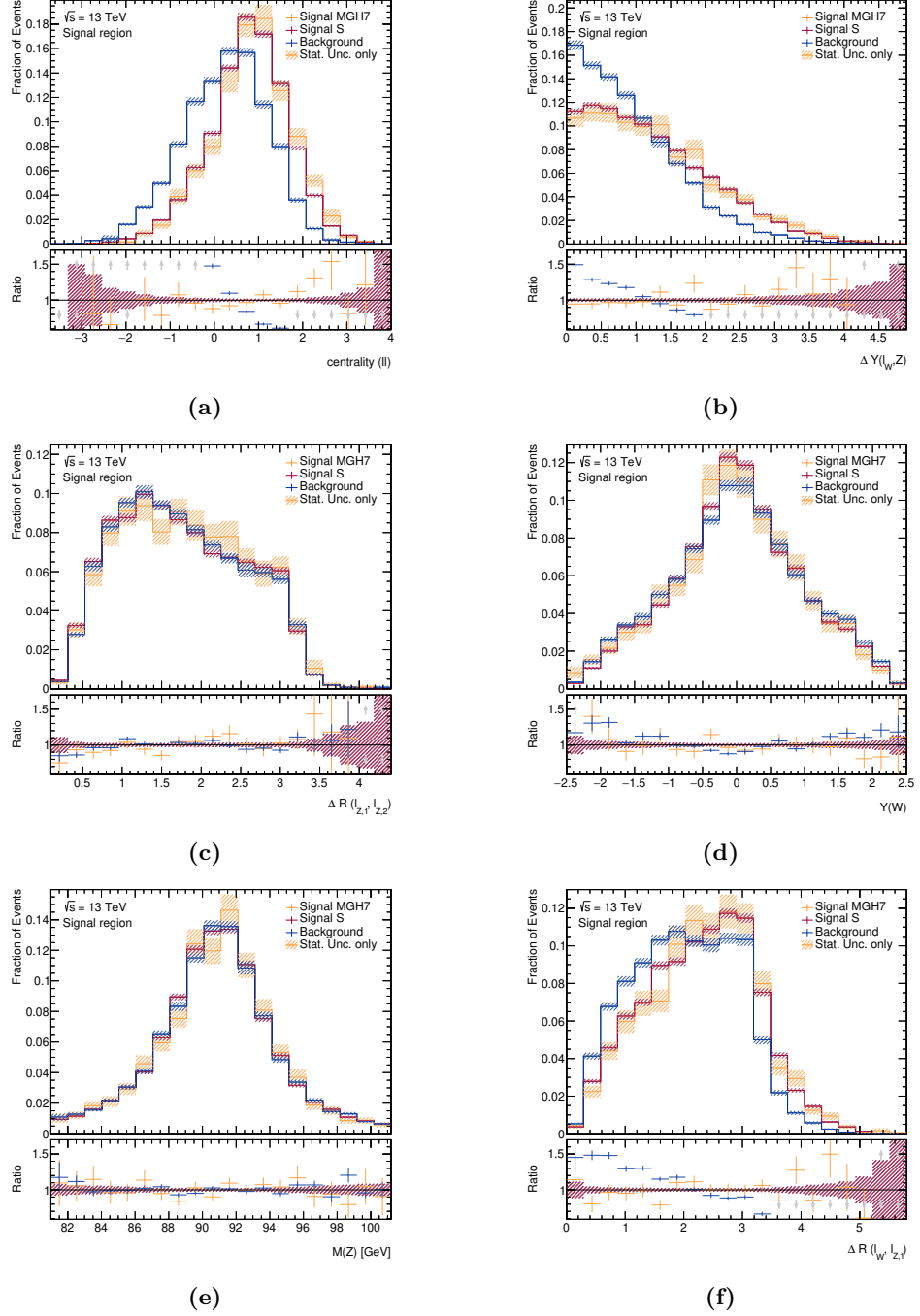


Figure F.4: Comparison of normalized differential predictions for the SHERPA (red) and MG5_AMC-based (orange) signal prediction as well as the prediction for the combined prompt backgrounds (blue) in different input observables in the signal region. Figure (a) shows centrality(ℓ), Figure (b) shows $|\Delta Y(\ell_W, Z)|$, Figure (c) shows $\Delta R(\ell_{Z,1}, \ell_{Z,2})$, Figure (d) shows $Y(W)$, Figure (e) shows $M(Z)$, and Figure (f) shows $\Delta R(\ell_W, \ell_{Z,1})$.

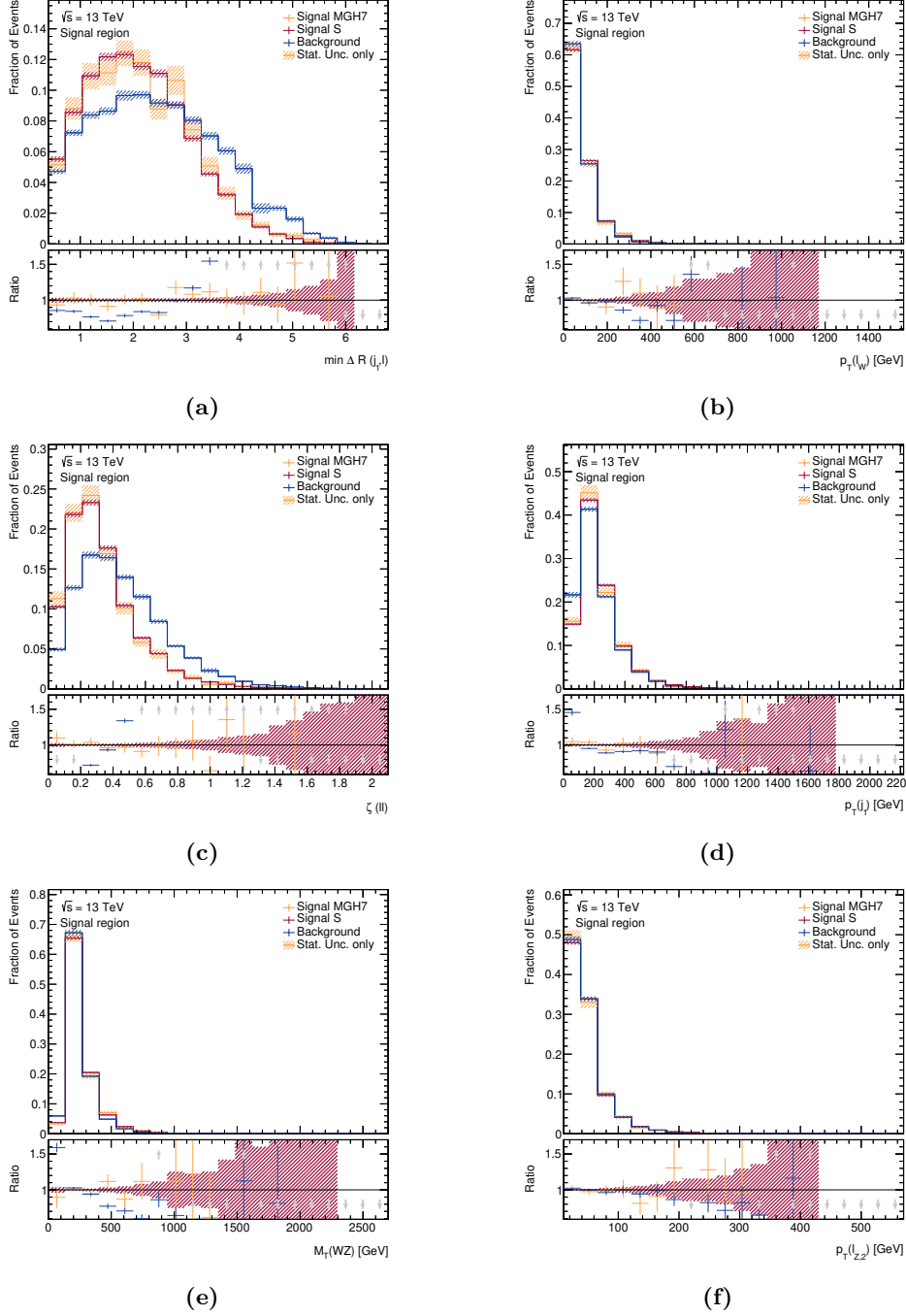


Figure F.5: Comparison of normalized differential predictions for the SHERPA (red) and MG5_AMC-based (orange) signal prediction as well as the prediction for the combined prompt backgrounds (blue) in different input observables in the signal region. Figure (a) shows $\min \Delta R(j_1, \ell)$, Figure (b) shows $p_T(\ell_W)$, Figure (c) shows $\zeta(\ell)$, Figure (d) shows $p_T(j_1)$, Figure (e) shows $M_T(WZ)$, and Figure (f) shows $p_T(j_2)$.

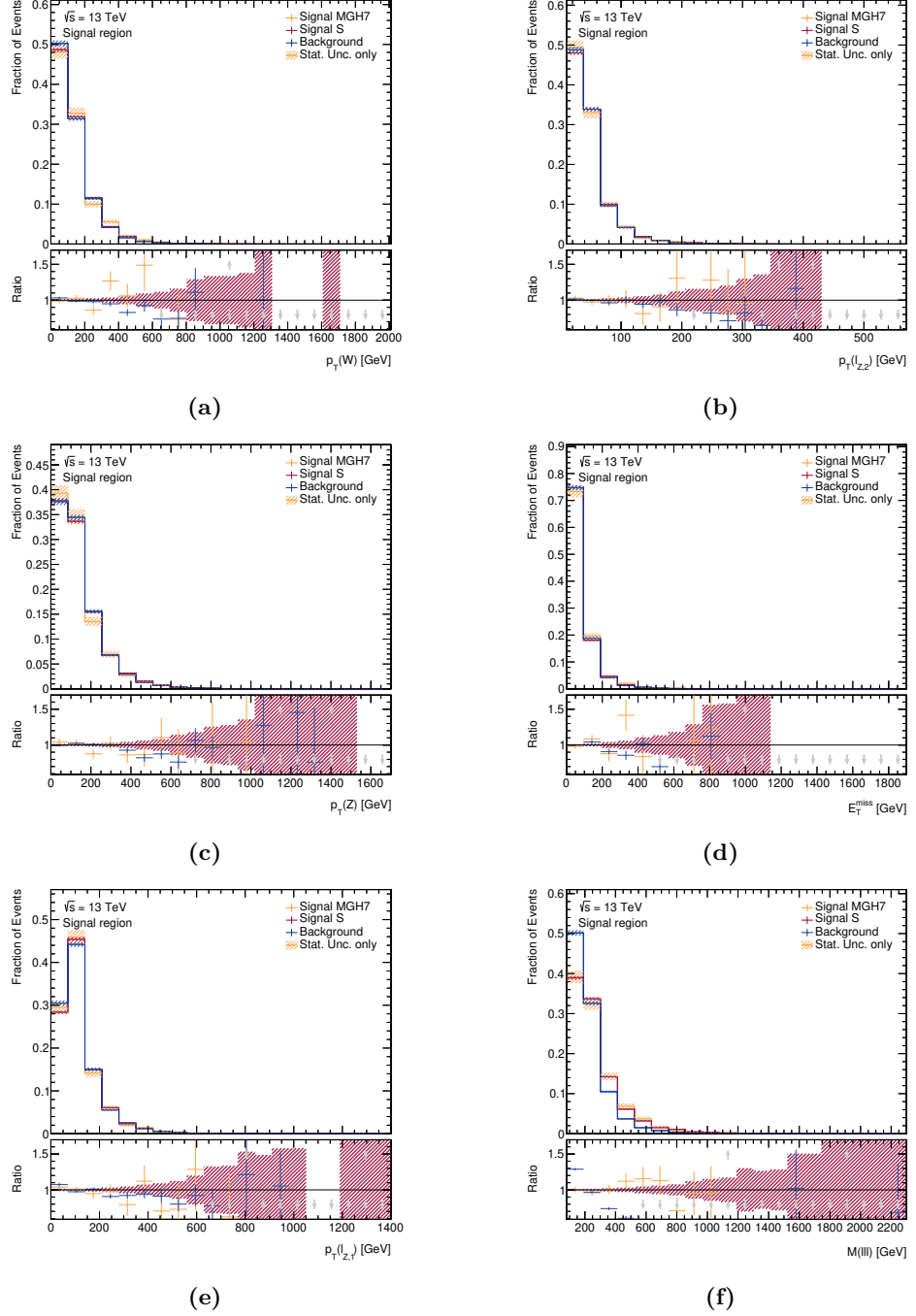


Figure F.6: Comparison of normalized differential predictions for the SHERPA (red) and MG5_AMC-based (orange) signal prediction as well as the prediction for the combined prompt backgrounds (blue) in different input observables in the signal region. Figure (a) shows $p_T(W)$, Figure (b) shows $p_T(\ell_{Z,2})$, Figure (c) shows $p_T(Z)$, Figure (d) shows E_T^{miss} , Figure (e) shows $p_T(\ell_{Z,1})$, and Figure (f) shows $M(\ell\ell)$.

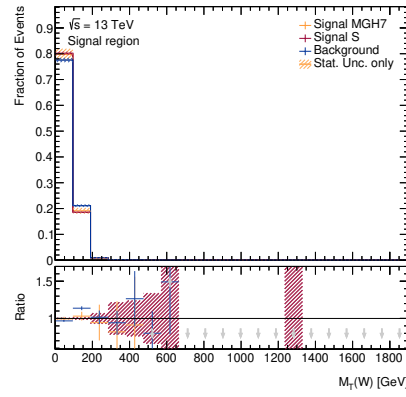


Figure F.7: Comparison of normalized differential predictions for the SHERPA (red) and MG5_AMC-based (orange) signal prediction as well as the prediction for the combined prompt backgrounds (blue) in $M_T(W)$ in the signal region.

F.2 Truth-level comparison BDT inputs

Comparisons of normalized predictions based on the $WZjj$ -EW4 and $WZjj$ -EW6 process simulations for all BDT input distributions in the fiducial phase space are shown in Figure F.8, Figure F.9, and Figure F.10. The observables are introduced in detail in Section 7.2.

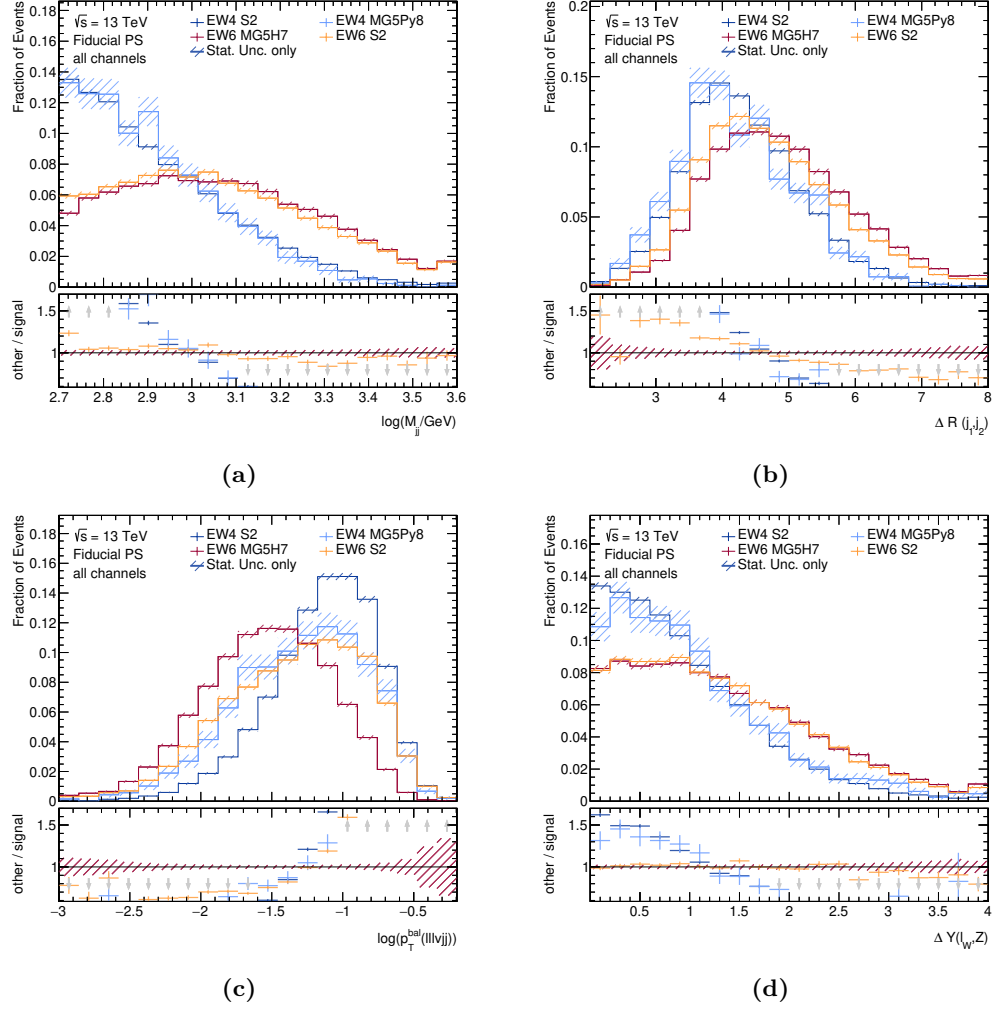


Figure F.8: Comparison of normalized distributions in the fiducial phase space for the main $WZjj$ -EW4 and $WZjj$ -EW6 process simulations. Figure F.8a shows M_{jj} , Figure F.8b shows $\Delta R(j_1, j_2)$, Figure F.8c shows $p_T^{\text{bal}}(\ell\ell vjj)$ and Figure F.8d shows $|\Delta Y(\ell_W, Z)|$. The inset shows the ratio to the nominal simulation of the signal process based on MG5_AMC and HERWIG 7. The indicated uncertainties represent only the statistical uncertainties due to the limited amount of events in the simulations.

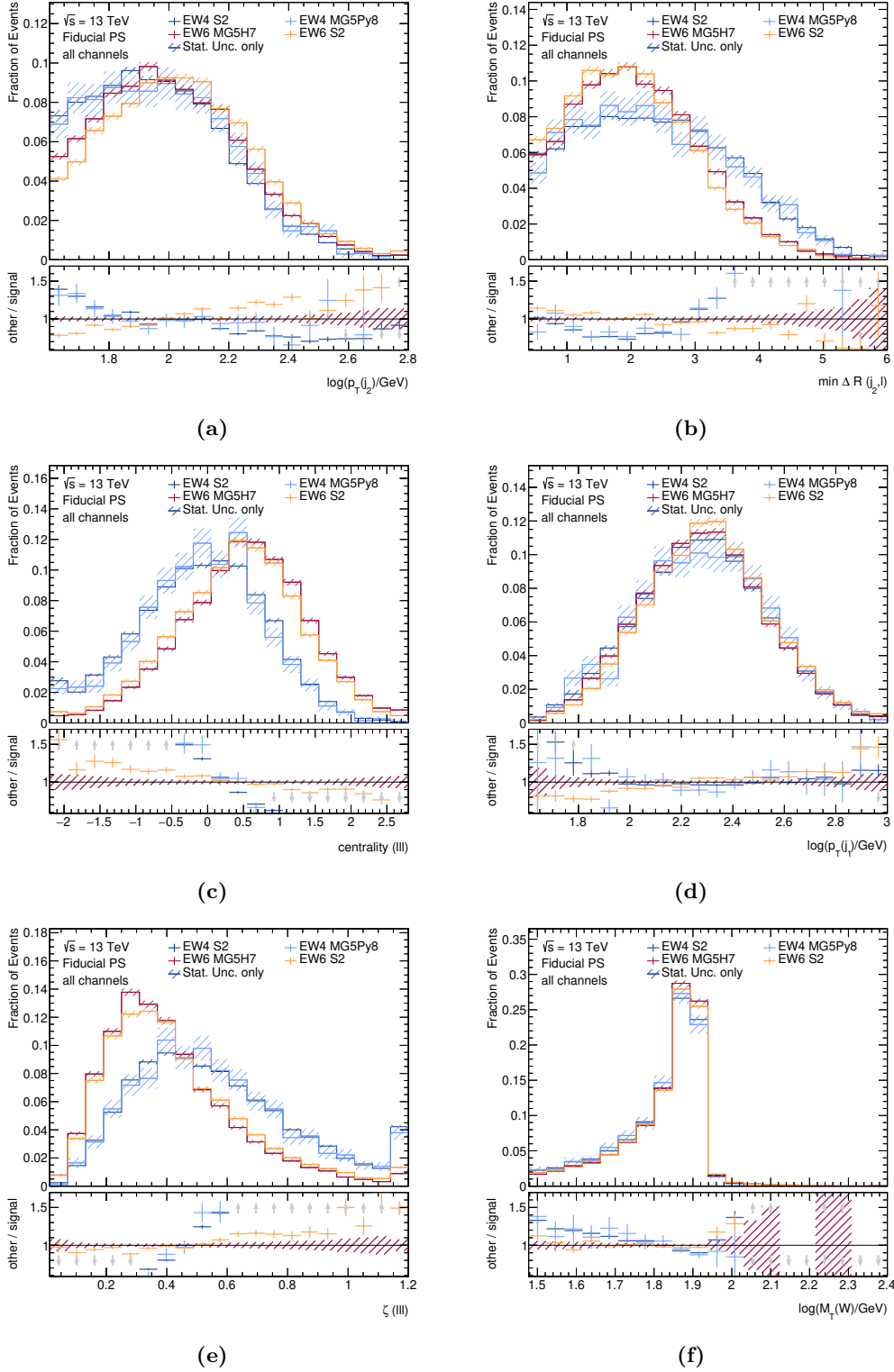


Figure F.9: Comparison of normalized distributions in the fiducial phase space for the main $WZjj$ -EW4 and $WZjj$ -EW6 process simulations. Figure F.9a shows $p_T(j_2)$, Figure F.9b shows $\min \Delta R(j_1, \ell)$, Figure F.9c shows $\text{centrality}(\ell\ell)$, Figure F.9d shows $p_T(j_1)$, Figure F.9e shows $\zeta(\ell\ell)$ and Figure F.9f shows $M_T(W)$. The inlay shows the ratio to the nominal simulation of the signal process based on MG5_AMC and HERWIG 7. The indicated uncertainties represent only the statistical uncertainties due to the limited amount of events in the simulations.

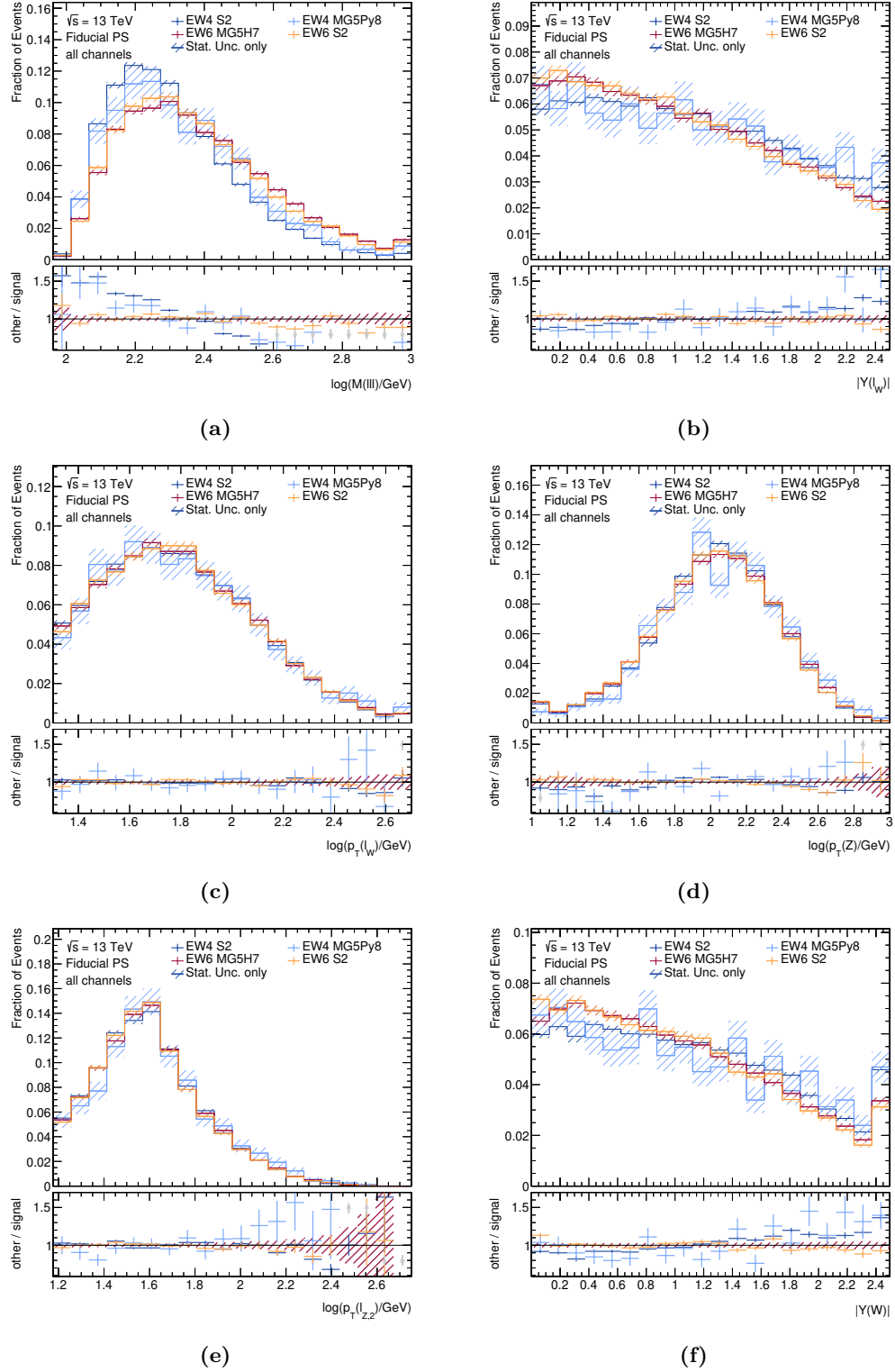


Figure F.10: Comparison of normalized distributions in the fiducial phase space for the main $WZjj$ -EW4 and $WZjj$ -EW6 process simulations. Figure F.10a shows $M(\ell\ell\ell)$, Figure F.10b shows $Y(W)$, Figure F.10c shows $p_T(\ell_W)$, Figure F.10d shows $p_T(Z)$, Figure F.10e shows $p_T(\ell_{Z,2})$ and Figure F.10f shows $Y(W)$. The inlay shows the ratio to the nominal simulation of the signal process based on MG5_AMC and HERWIG 7. The indicated uncertainties represent only the statistical uncertainties due to the limited amount of events in the simulations.

F.3 Validation of Modelling of BDT inputs

This section collects comparisons of the predicted and observed differential event yields in different regions. The event yield is shown differentially in the **BDTscore** as well as each of the input observables used in the BDT. For each observable, the direct prediction is shown as well as a corrected prediction, where the normalizations of the main backgrounds are adjusted to fit data in control regions. This correction is described in detail in Section 9.2.4 and the scaling factors for each process are indicated in the legend. The uncertainties show all experimental as well as roughly estimated modelling uncertainties as described in the beginning of Section 8.2. Grey hatched uncertainty band represents the total uncertainty and cross-hatched band indicates the statistical uncertainty of the prediction. The lower inlay shows the ratio of the observed with respect to the predicted yield.

F.3.1 WZ Control Region

Figure F.11 shows the differential event yield comparison in the WZ control region for the **BDTscore**, and Figures F.12 to Figure F.27 show each of the input observables of the BDT.

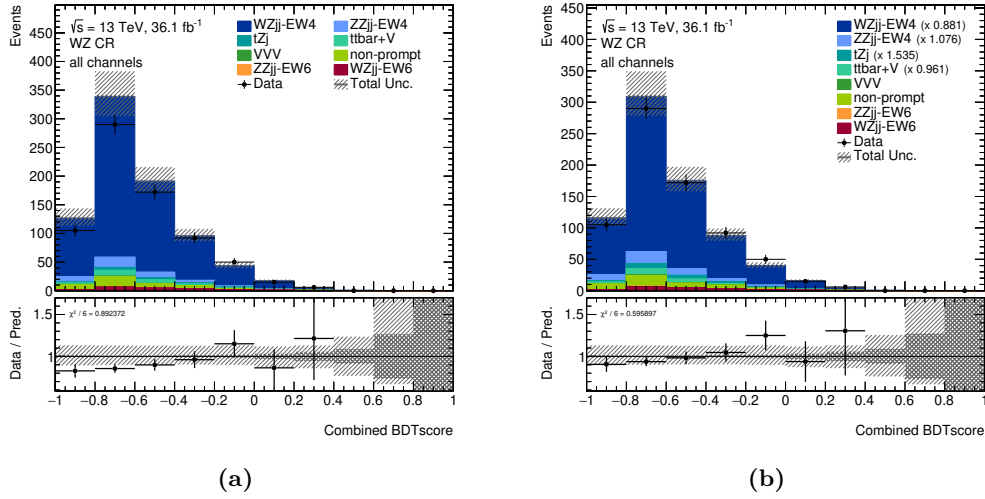


Figure F.11: Comparison of predicted and observed event yields differentially in the **BDTscore** in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

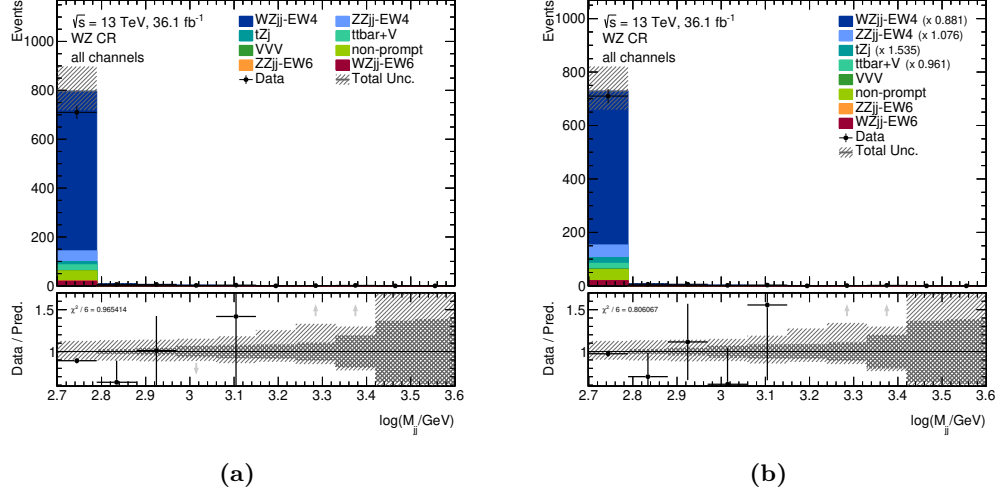


Figure F.12: Comparison of predicted and observed event yields differentially in M_{jj} in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

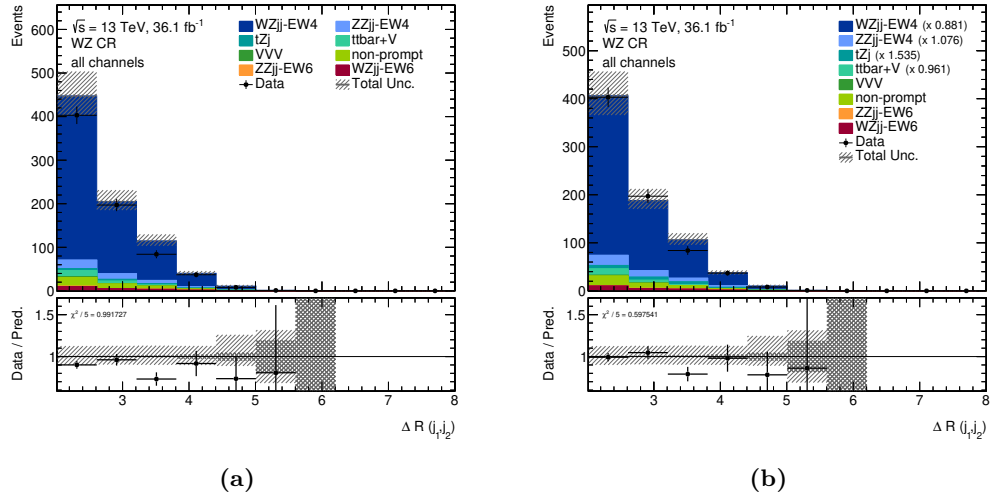


Figure F.13: Comparison of predicted and observed event yields differentially in $\Delta R(j_1, j_2)$ in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

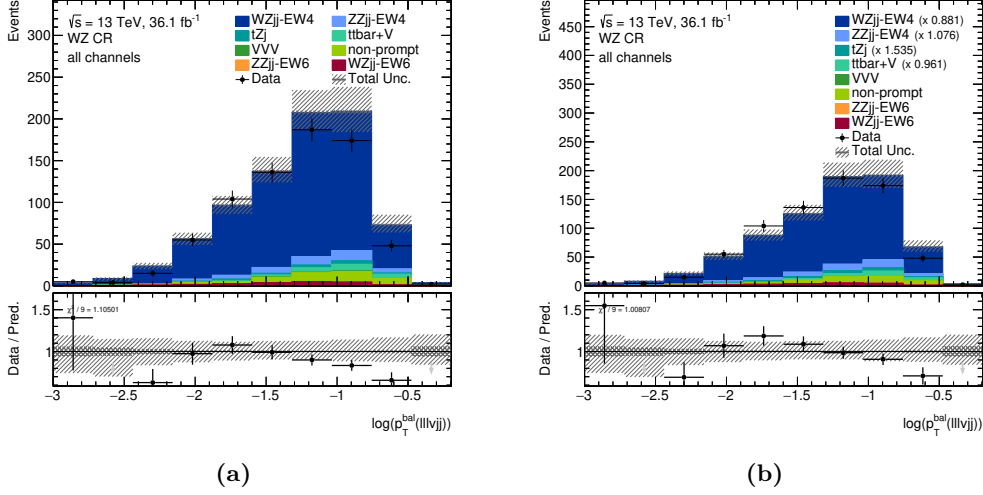


Figure F.14: Comparison of predicted and observed event yields differentially in $p_T^{\text{bal}}(\ell\ell vjj)$ in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

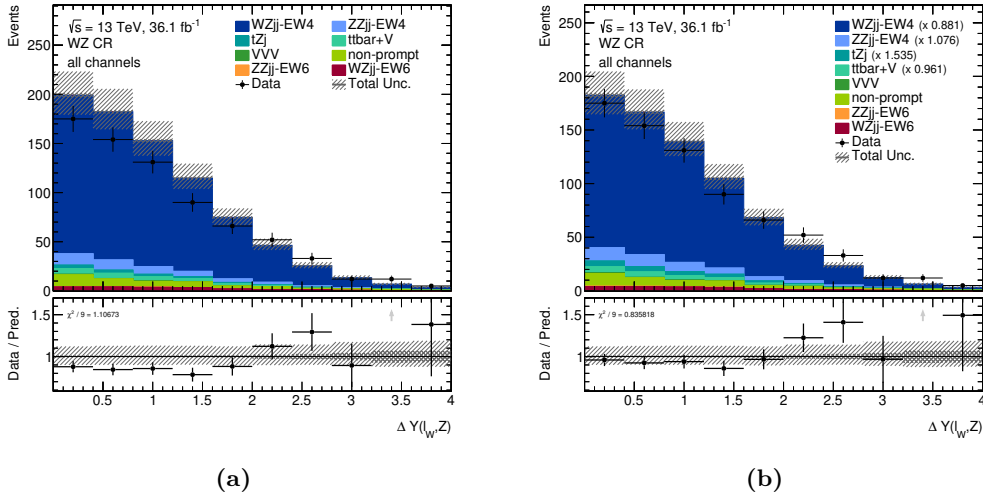


Figure F.15: Comparison of predicted and observed event yields differentially in $|\Delta Y(\ell_W, Z)|$ in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

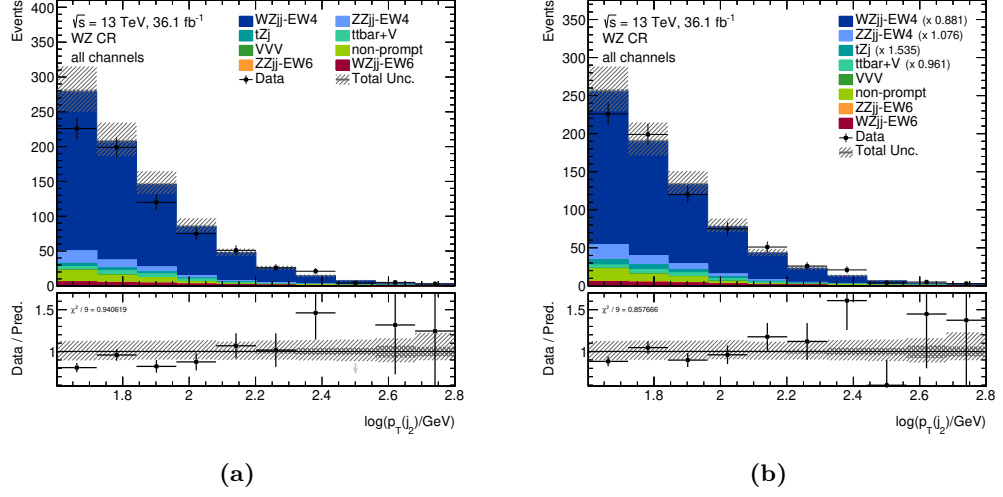


Figure F.16: Comparison of predicted and observed event yields differentially in $p_T(j_2)$ in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

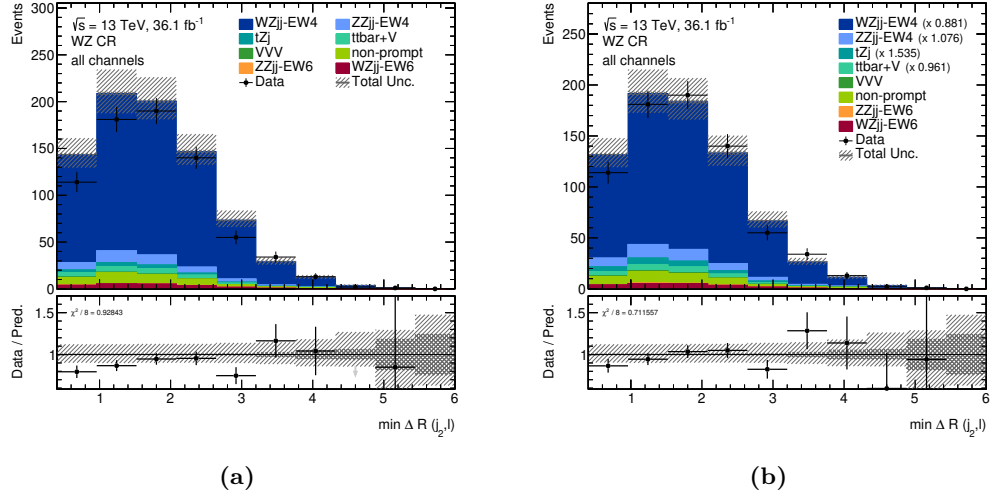


Figure F.17: Comparison of predicted and observed event yields differentially in $\min \Delta R(j_1, \ell)$ in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

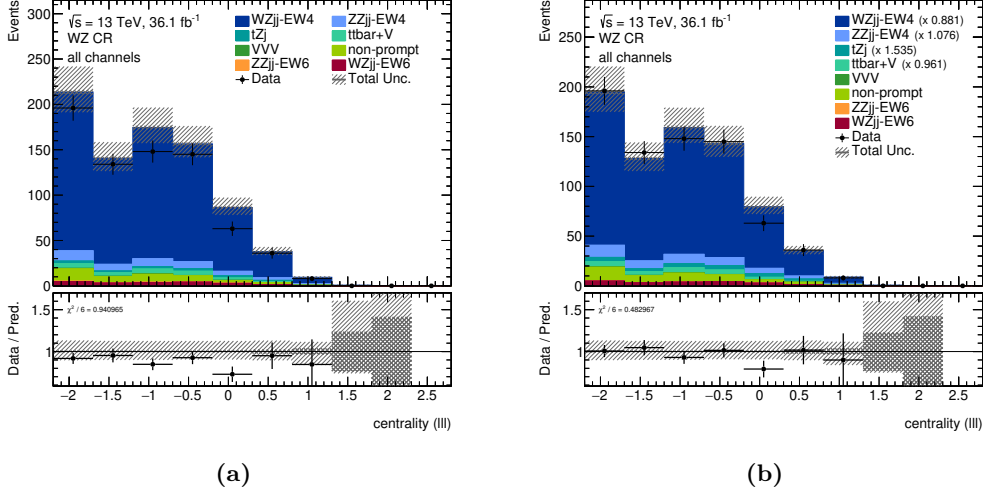


Figure F.18: Comparison of predicted and observed event yields differentially in centrality($\ell\ell$) in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

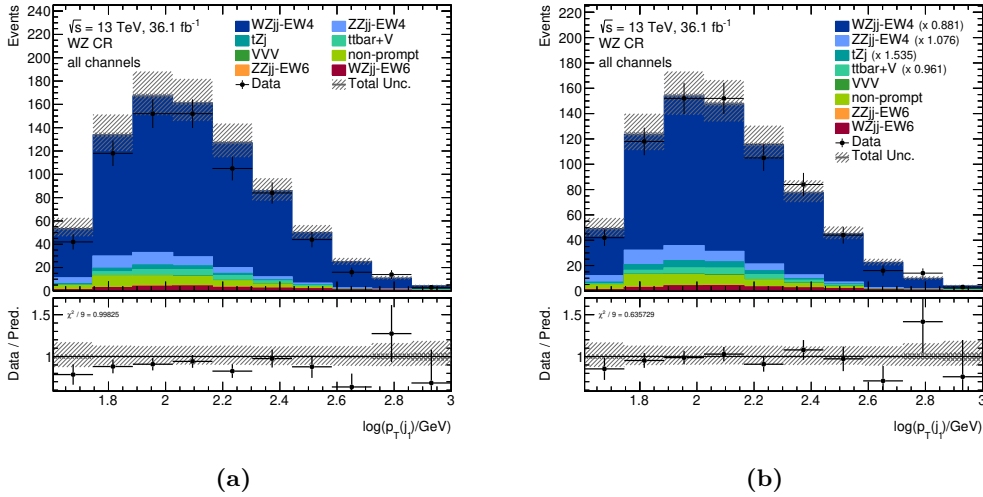


Figure F.19: Comparison of predicted and observed event yields differentially in $p_T(j_1)$ in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

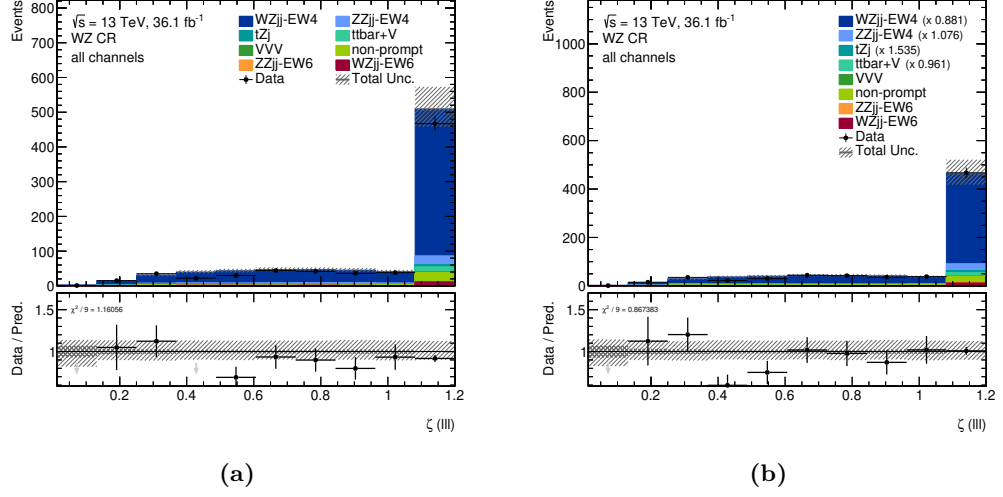


Figure F.20: Comparison of predicted and observed event yields differentially in $\zeta(\ell\ell)$ in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

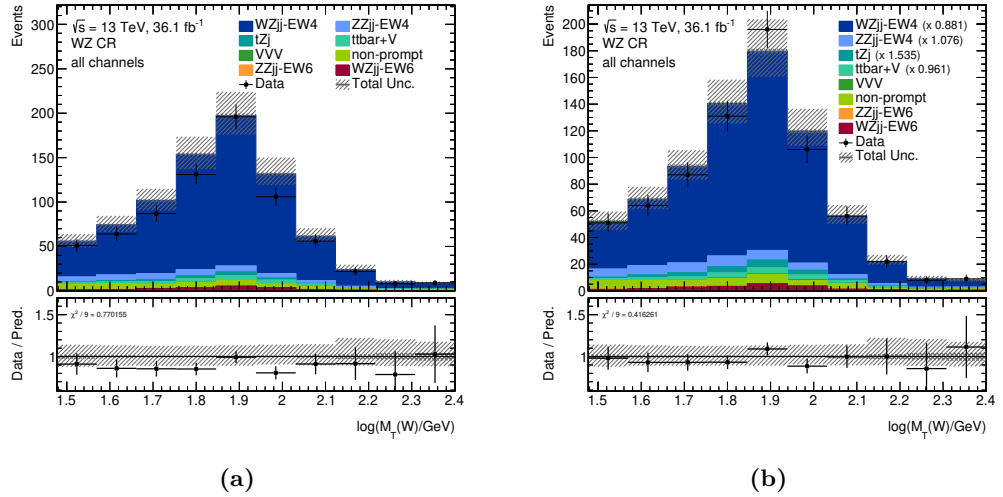


Figure F.21: Comparison of predicted and observed event yields differentially in $M_T(W)$ in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

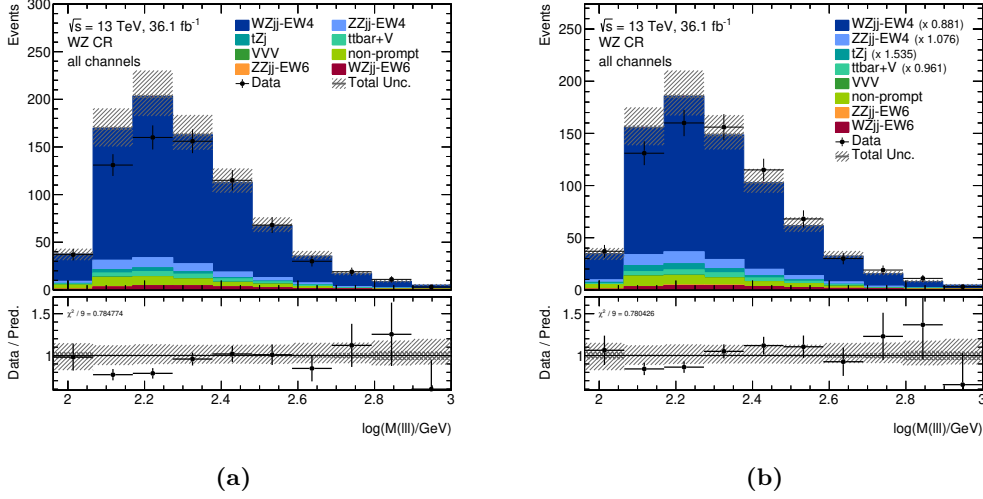


Figure F.22: Comparison of predicted and observed event yields differentially in $M(\ell\ell)$ in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

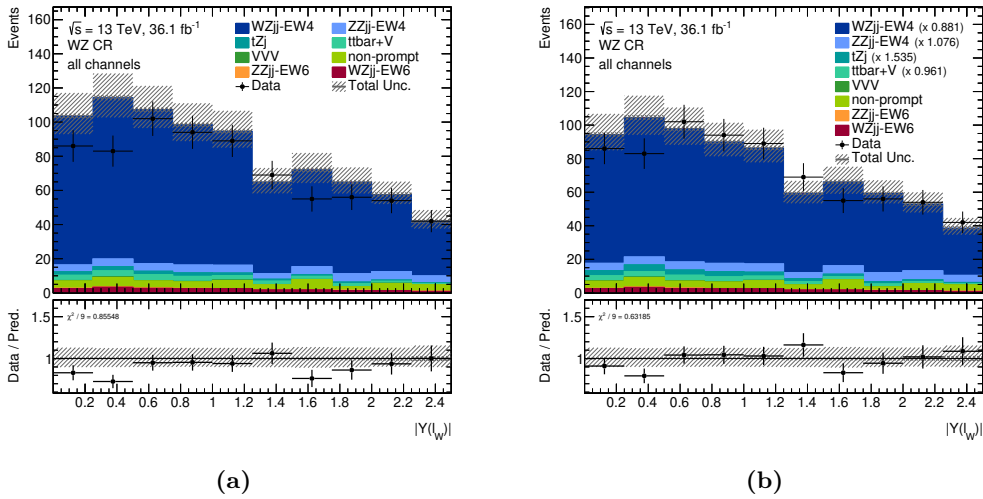


Figure F.23: Comparison of predicted and observed event yields differentially in $Y(\ell_W)$ in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

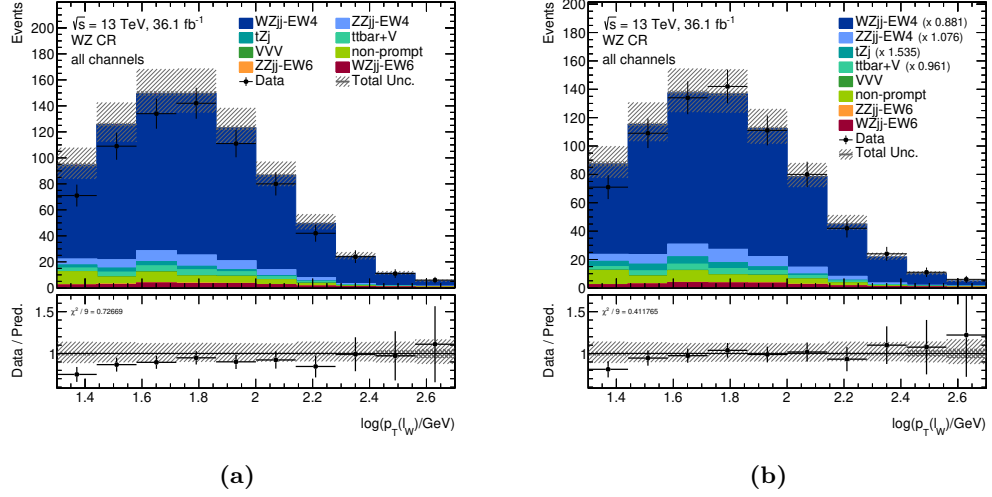


Figure F.24: Comparison of predicted and observed event yields differentially in $p_T(\ell_W)$ in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

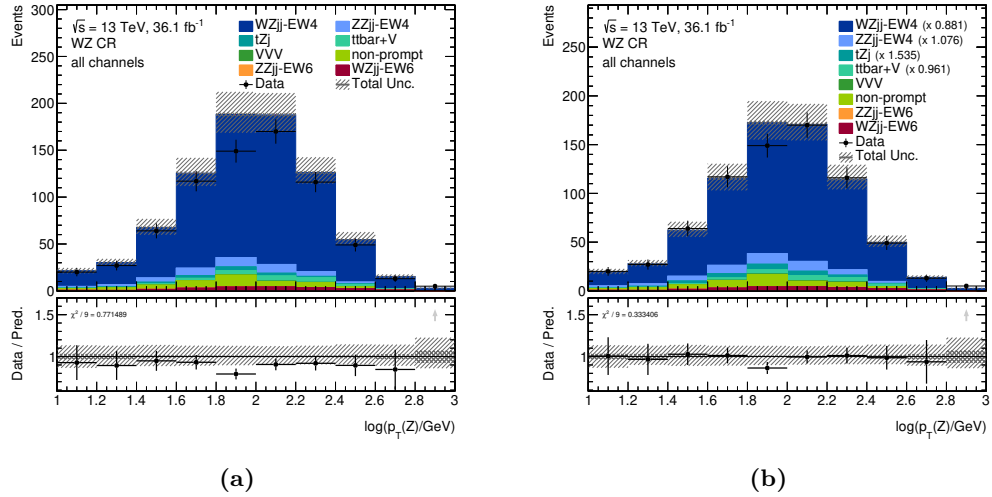


Figure F.25: Comparison of predicted and observed event yields differentially in $p_T(Z)$ in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

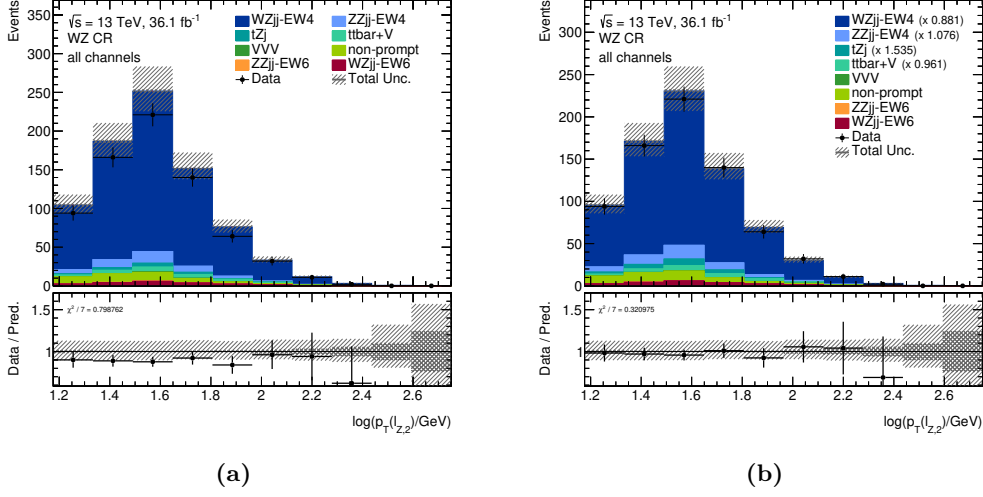


Figure F.26: Comparison of predicted and observed event yields differentially in $p_T(\ell_{Z,2})$ in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

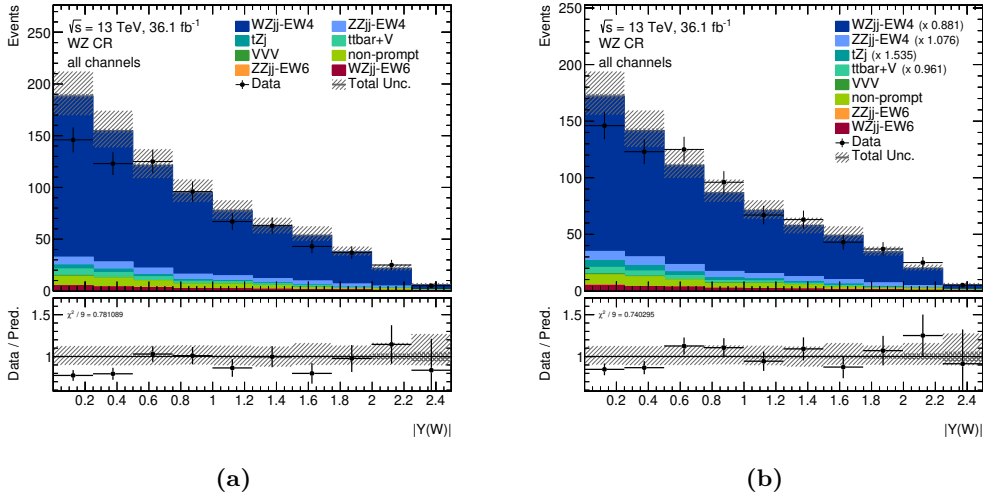


Figure F.27: Comparison of predicted and observed event yields differentially in $Y(W)$ in the WZ control region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

F.3.2 Signal Region

Figure F.28 compares the differential event yields in the signal region for the **BDTscore**, and Figures F.29 to Figure F.44 show each of the input observables of the BDT.

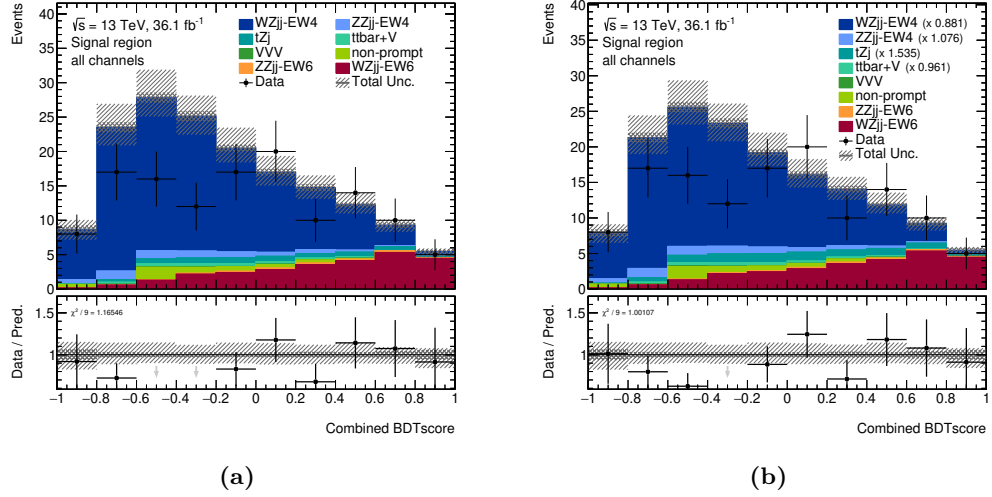


Figure F.28: Comparison of predicted and observed event yields differentially in the **BDTscore** in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

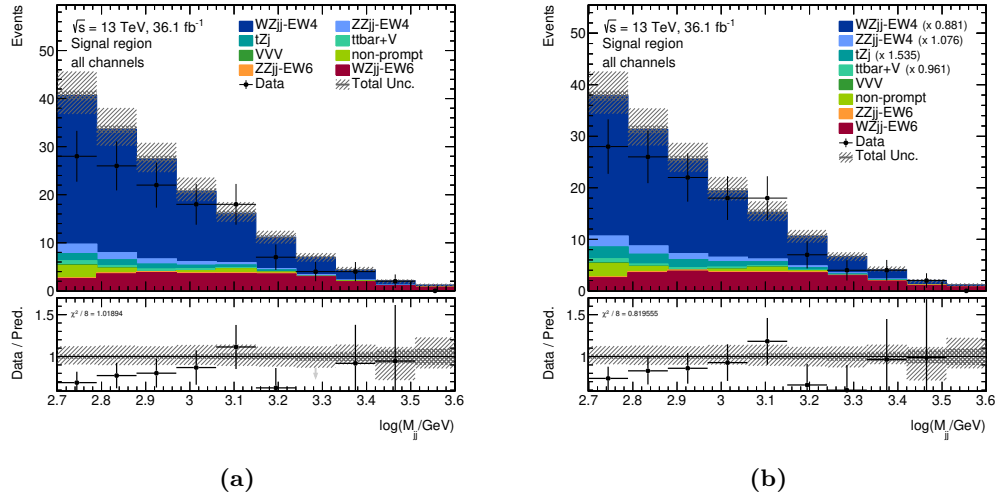


Figure F.29: Comparison of predicted and observed event yields differentially in M_{jj} in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

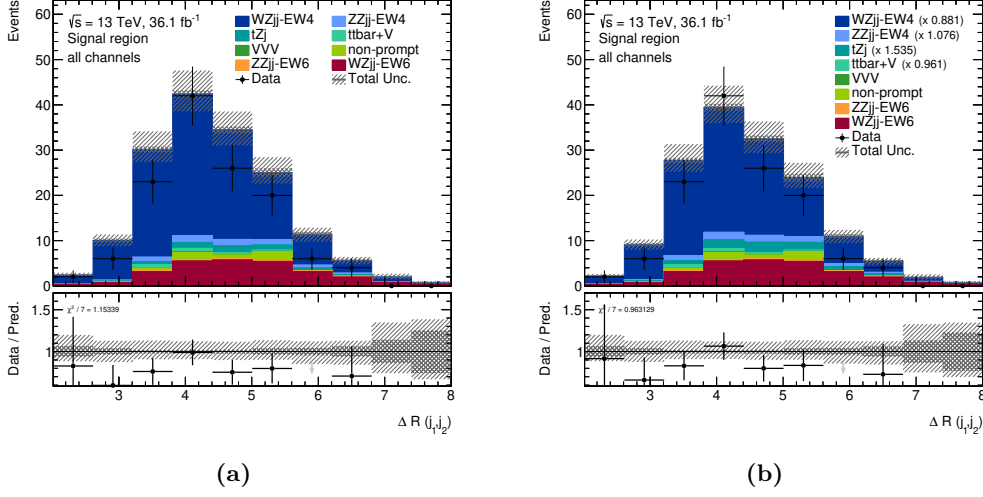


Figure F.30: Comparison of predicted and observed event yields differentially in $\Delta R(j_1, j_2)$ in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

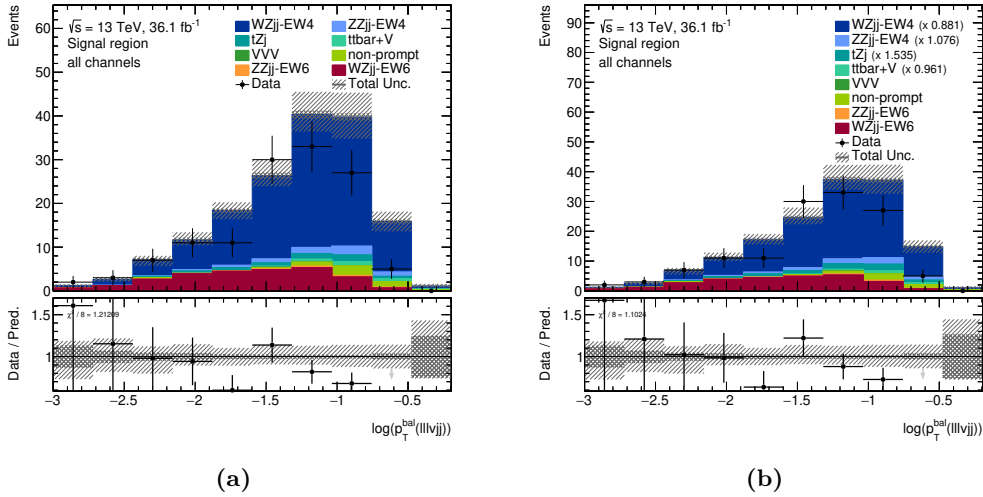


Figure F.31: Comparison of predicted and observed event yields differentially in $p_T^{\text{bal}}(\ell\ell vjj)$ in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

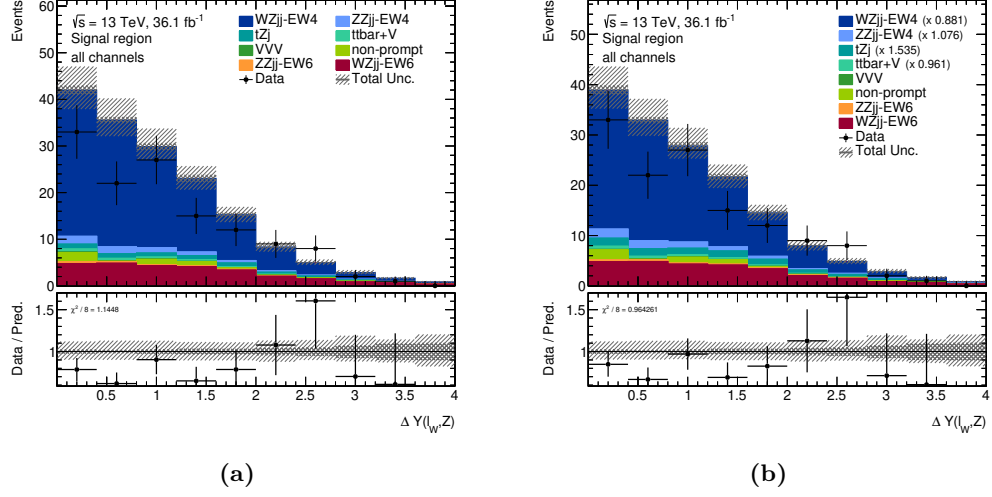


Figure F.32: Comparison of predicted and observed event yields differentially in $|\Delta Y(\ell_W, Z)|$ in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

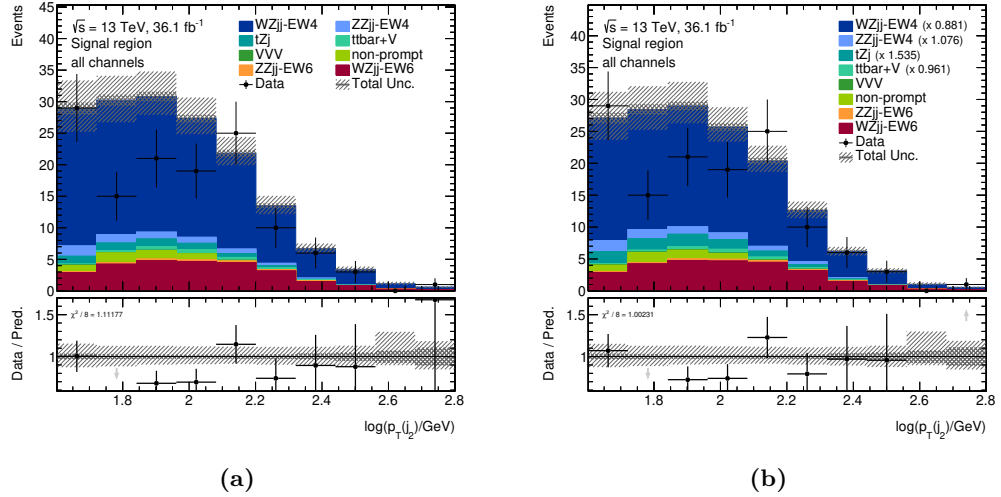


Figure F.33: Comparison of predicted and observed event yields differentially in $p_T(j_2)$ in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

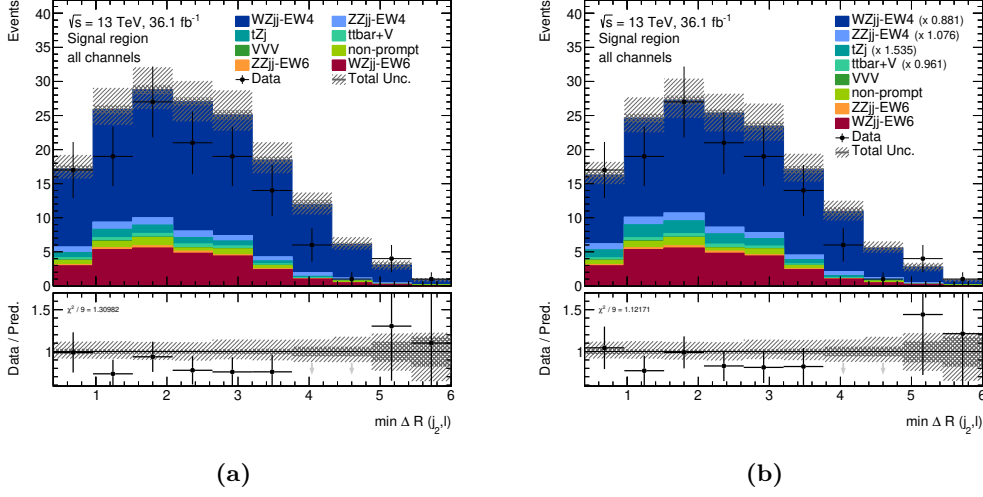


Figure F.34: Comparison of predicted and observed event yields differentially in $\min \Delta R(j_1, \ell)$ in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

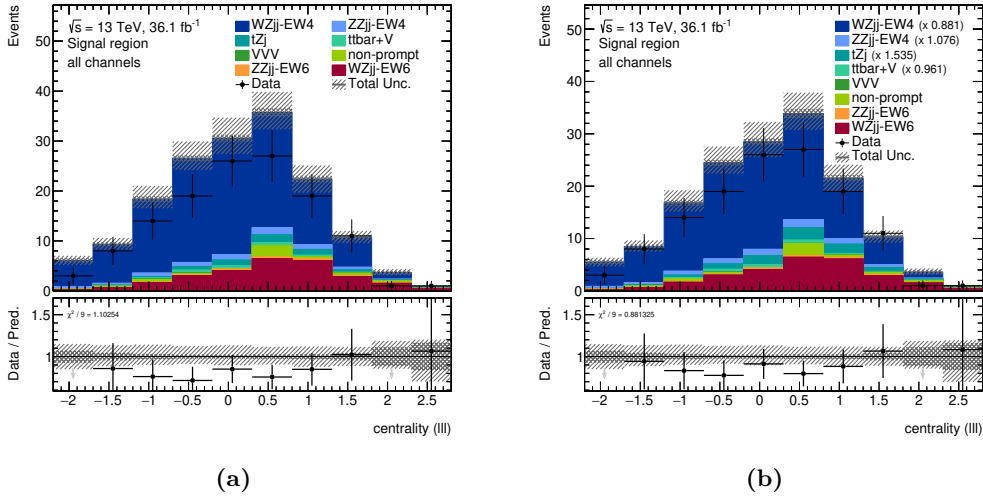


Figure F.35: Comparison of predicted and observed event yields differentially in centrality($\ell\ell$) in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

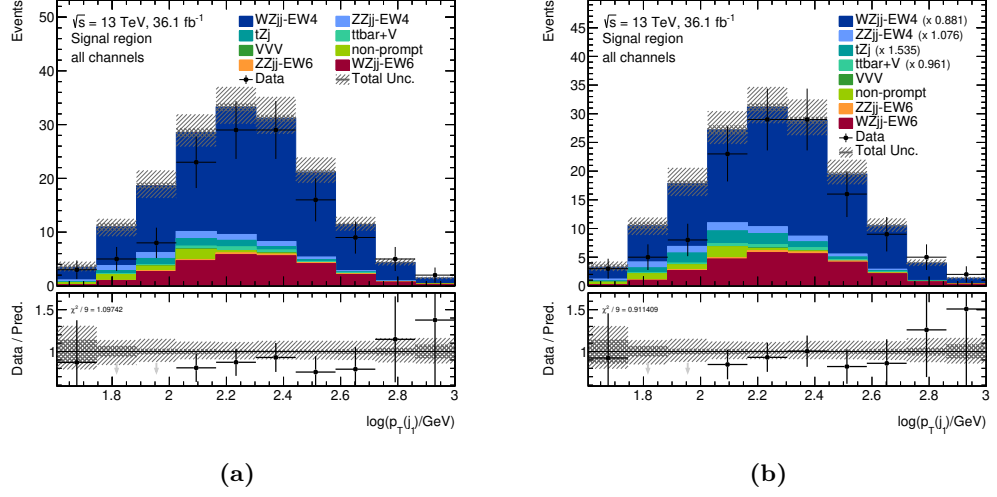


Figure F.36: Comparison of predicted and observed event yields differentially in $p_T(j_1)$ in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

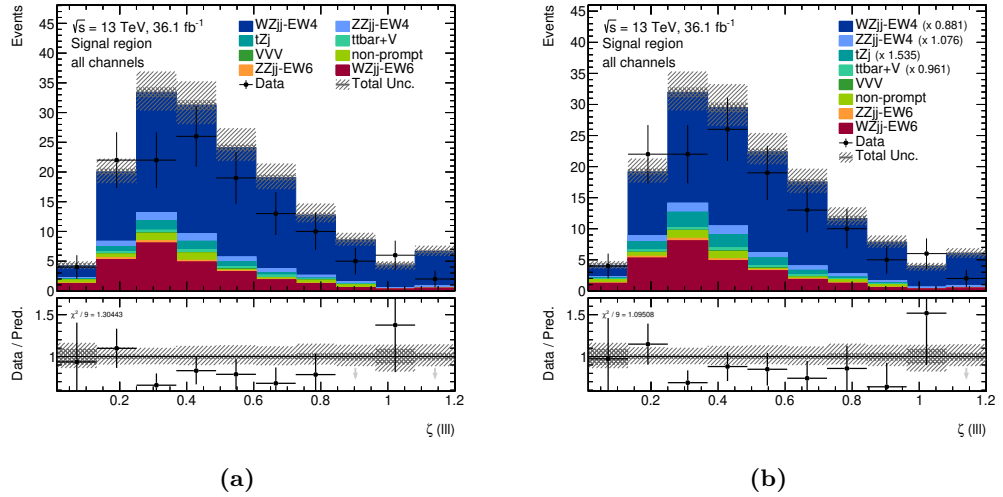


Figure F.37: Comparison of predicted and observed event yields differentially in $\zeta(\ell\ell\ell)$ in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

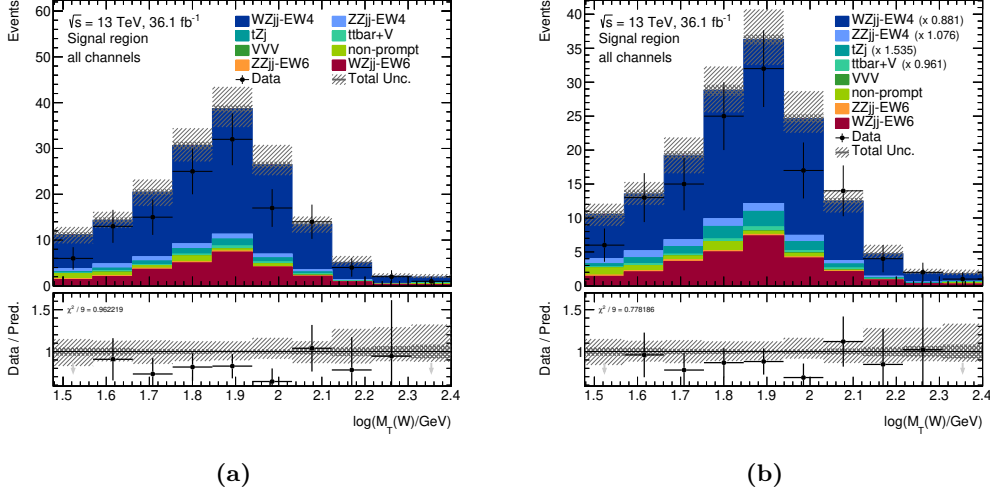


Figure F.38: Comparison of predicted and observed event yields differentially in $M_T(W)$ in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

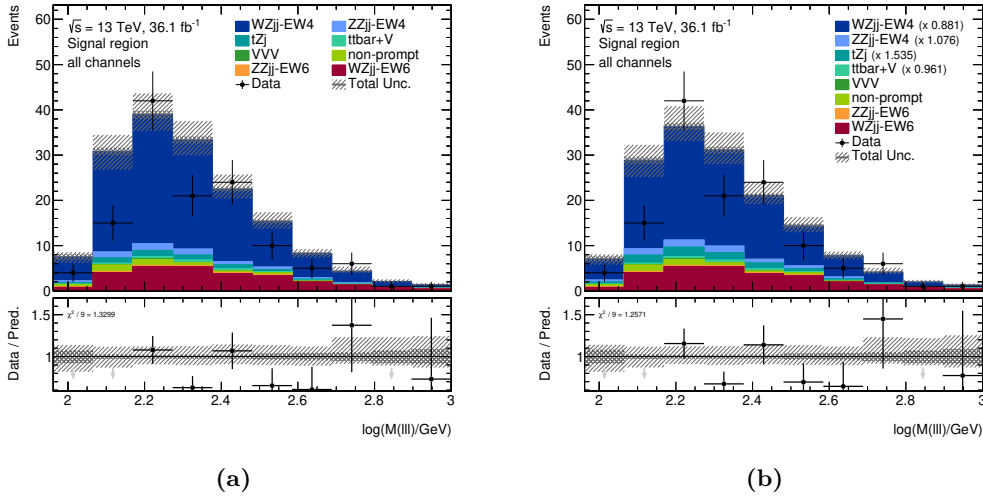


Figure F.39: Comparison of predicted and observed event yields differentially in $M(\text{lll})$ in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

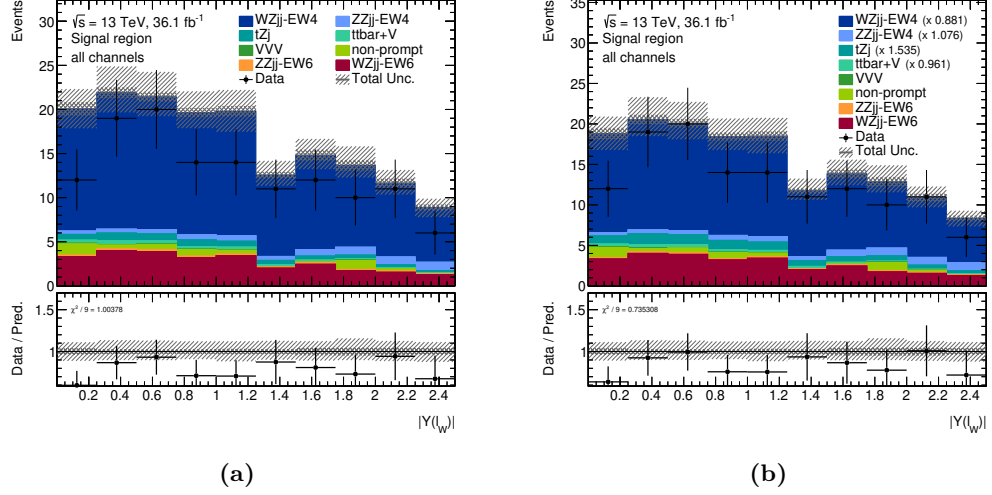


Figure F.40: Comparison of predicted and observed event yields differentially in $Y(\ell_W)$ in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

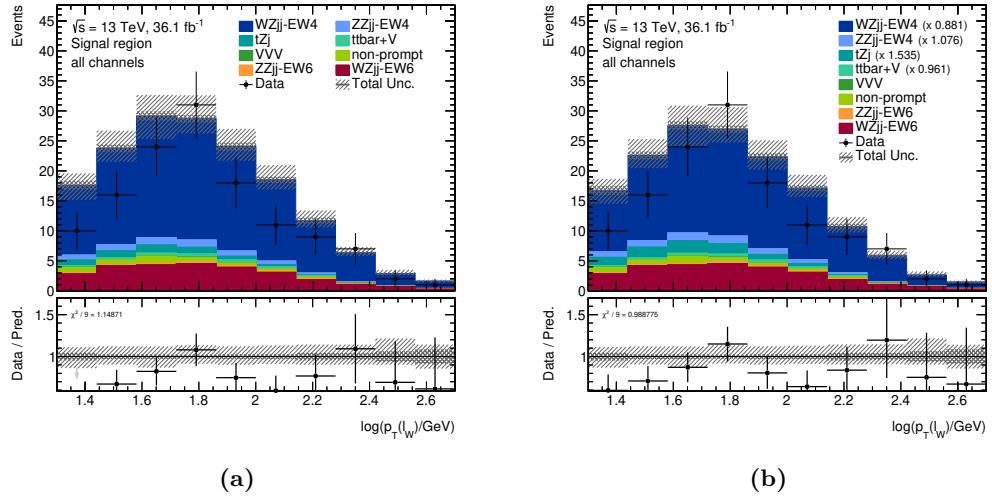


Figure F.41: Comparison of predicted and observed event yields differentially in $pr(\ell_W)$ in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

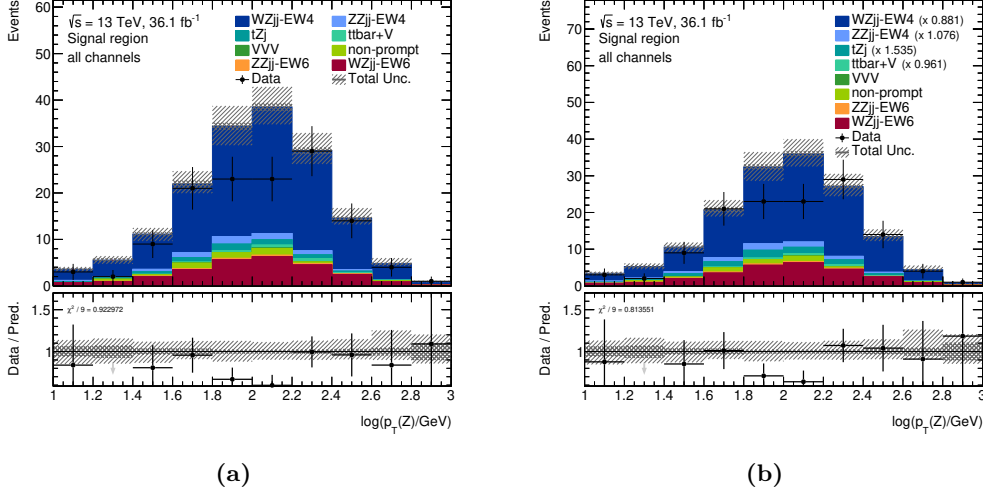


Figure F.42: Comparison of predicted and observed event yields differentially in $p_T(Z)$ in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

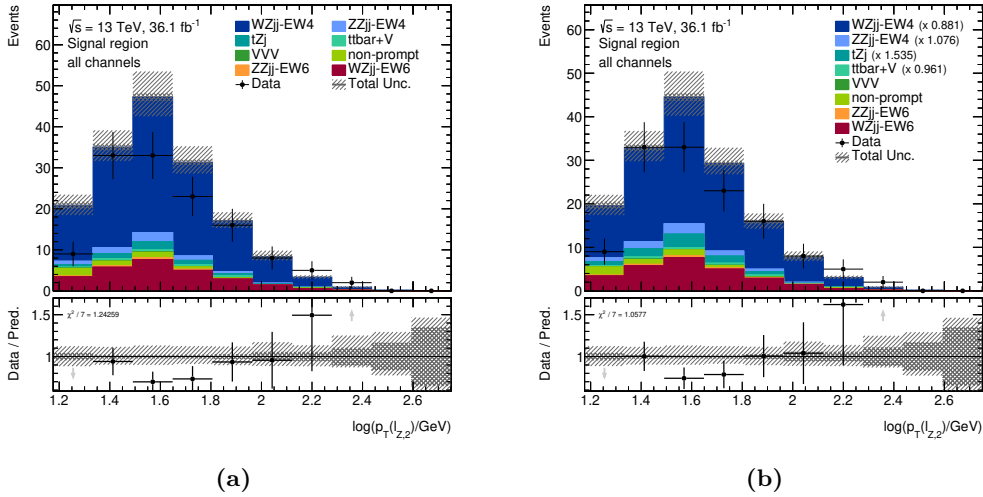


Figure F.43: Comparison of predicted and observed event yields differentially in $p_T(\ell_{Z,2})$ in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

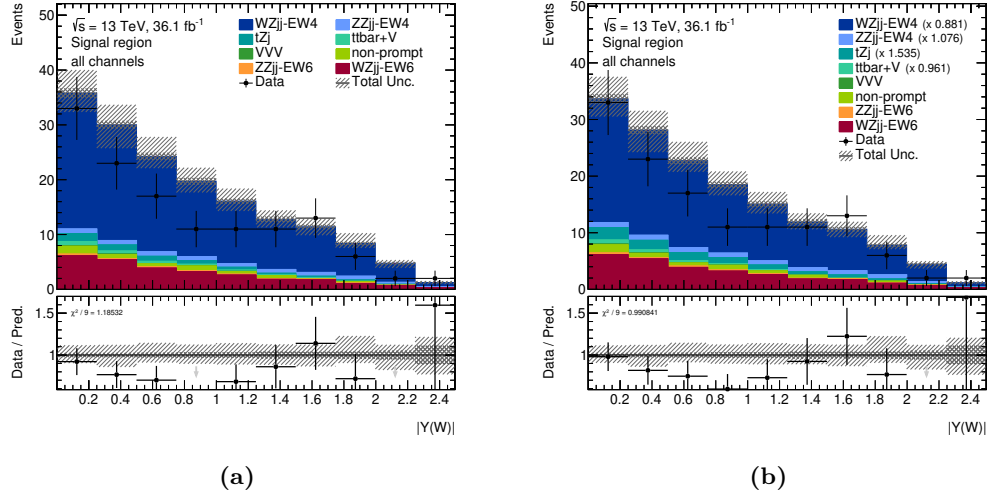


Figure F.44: Comparison of predicted and observed event yields differentially in $Y(W)$ in the signal region. Figure (a) shows the direct prediction and Figure (b) the corrected prediction with adjusted scaling of the main backgrounds as indicated in the legend. More details as described in the beginning of Section F.3.

Appendix G

Uncertainties

G.1 Modelling Uncertainties in Control Regions

G.1.1 PDF uncertainties

The modelling uncertainties due to the choice of PDF are visualized for the $WZjj$ -EW6 and $WZjj$ -EW4 processes in Figure G.1 for the b CR, in Figure G.2 for the ZZ CR, and in Figure G.3 for the WZ CR. The determination of the uncertainty follows the procedure as described in Section 8.2.1.

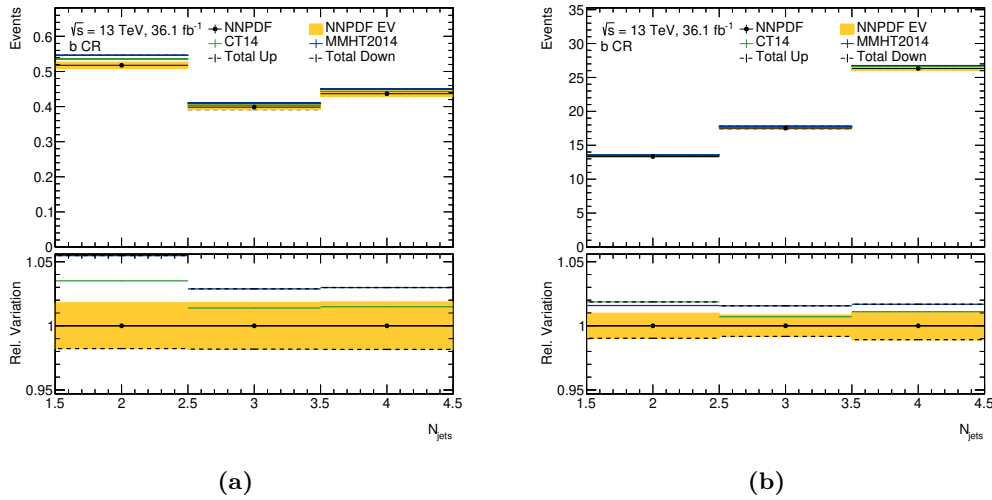


Figure G.1: Uncertainties due to mismodelling of the PDF on the predicted event yields differentially in the jet multiplicity for the b control region. Figure (a) shows the PDF uncertainties for the $WZjj$ -EW6 signal process and Figure (b) for the $WZjj$ -EW4 background process. The nominal prediction (black) is shown with the uncertainty derived from the NNPDF3.0 eigen-set (yellow band). For comparison the predicted yields using the nominal MMHT2014nnlo (blue) and CT14 (green) PDF sets are shown. The total PDF uncertainty derived from these comparisons is shown using black, dashed lines. The lower inlay shows the ratio to the nominal prediction in each bin.

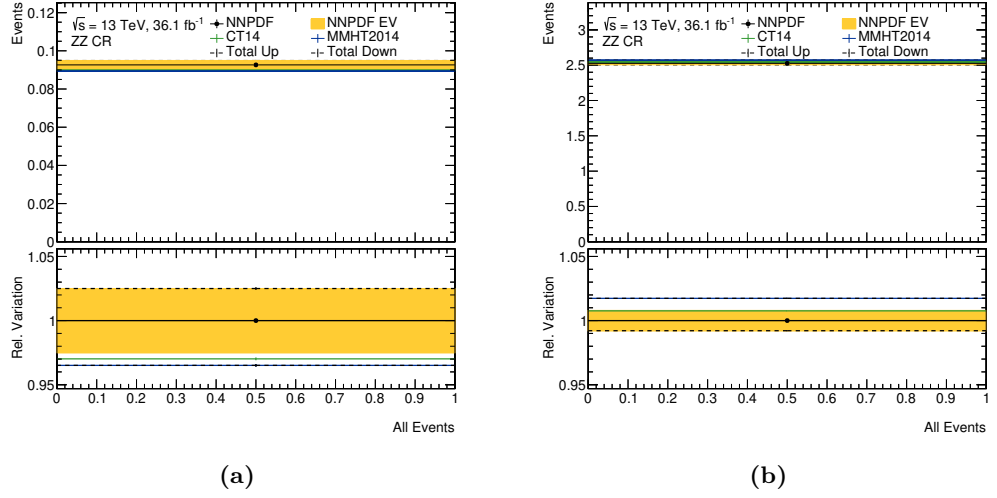


Figure G.2: Uncertainties due to mismodelling of the PDF on the predicted event yields for the ZZ control region. Figure (a) shows the PDF uncertainties for the $WZjj$ -EW6 signal process and Figure (b) for the $WZjj$ -EW4 background process. The nominal prediction (black) is shown with the uncertainty derived from the NNPDF3.0 eigen-set (yellow band). For comparison the predicted yields using the nominal MMHT2014nnlo (blue) and CT14 (green) PDF sets are shown. The total PDF uncertainty derived from these comparisons is shown using black, dashed lines. The lower inlay shows the ratio to the nominal prediction in each bin.

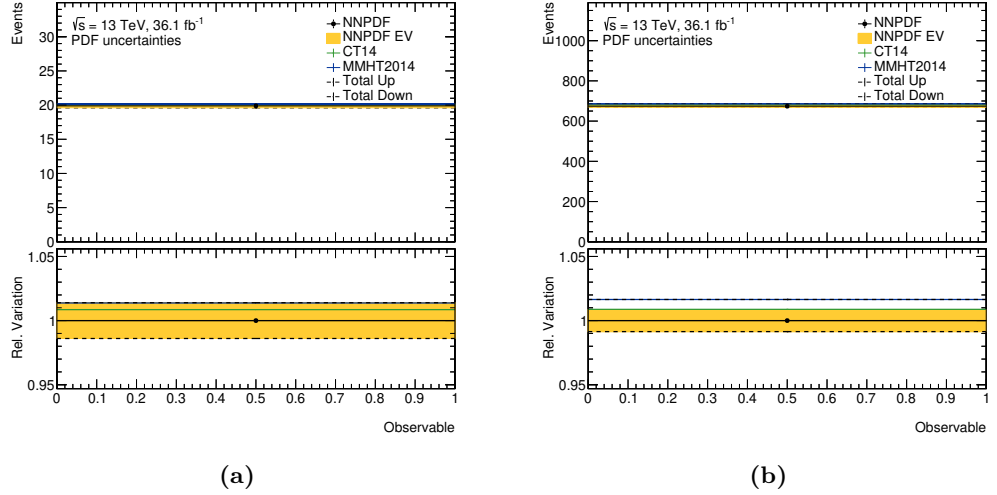


Figure G.3: Uncertainties due to mismodelling of the PDF on the predicted event yields for the WZ control region. Figure (a) shows the PDF uncertainties for the $WZjj$ -EW6 signal process and Figure (b) for the $WZjj$ -EW4 background process. The nominal prediction (black) is shown with the uncertainty derived from the NNPDF3.0 eigen-set (yellow band). For comparison the predicted yields using the nominal MMHT2014nnlo (blue) and CT14 (green) PDF sets are shown. The total PDF uncertainty derived from these comparisons is shown using black, dashed lines. The lower inlay shows the ratio to the nominal prediction in each bin.

G.1.2 Strong coupling

The uncertainties due to the choice of the strong coupling α_S for the $WZjj$ -EW6 and $WZjj$ -EW4 processes are shown in Figure G.4 for the b CR, in Figure G.5 for the ZZ CR, and in Figure G.6 for the WZ CR. The uncertainties are derived as described in Section 8.2.2.

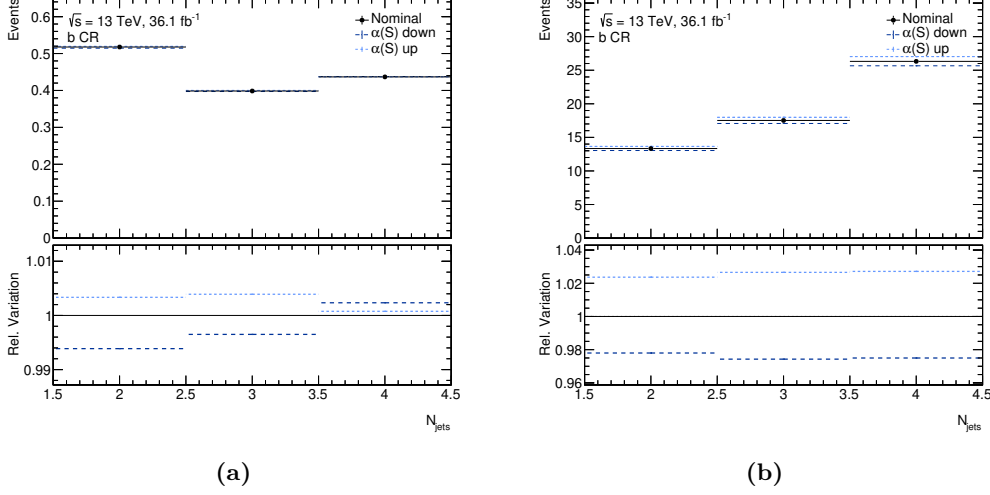


Figure G.4: Uncertainties due to mismodelling of α_S on the predicted event yields differentially in the jet multiplicity for the b control region. Figure (a) shows the α_S uncertainties for the $WZjj$ -EW6 signal process and Figure (b) for the $WZjj$ -EW4 background process. The nominal prediction (black) is compared to the predicted yields using the variations of $\alpha_S = 0.117$ (dashed, dark blue) and $\alpha_S = 0.119$ (dotted, light blue). The lower inlay shows the ratio to the nominal prediction in each bin.

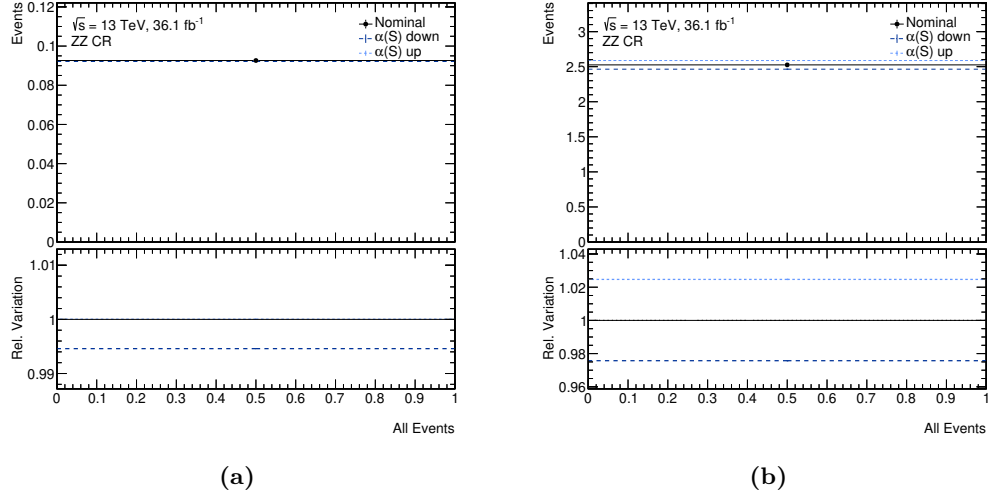


Figure G.5: Uncertainties due to mismodelling of α_S on the predicted event yields for the ZZ control region. Figure (a) shows the α_S uncertainties for the $WZjj$ -EW6 signal process and Figure (b) for the $WZjj$ -EW4 background process. The nominal prediction (black) is compared to the predicted yields using the variations of $\alpha_S = 0.117$ (dashed, dark blue) and $\alpha_S = 0.119$ (dotted, light blue). The lower inlay shows the ratio to the nominal prediction in each bin.

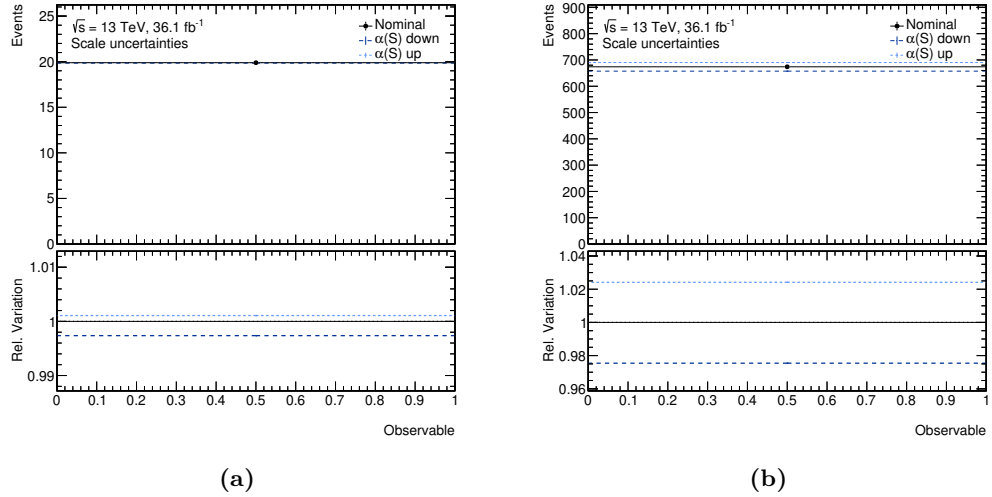


Figure G.6: Uncertainties due to mismodelling of α_S on the predicted event yields for the WZ control region. Figure (a) shows the α_S uncertainties for the $WZjj$ -EW6 signal process and Figure (b) for the $WZjj$ -EW4 background process. The nominal prediction (black) is compared to the predicted yields using the variations of $\alpha_S = 0.117$ (dashed, dark blue) and $\alpha_S = 0.119$ (dotted, light blue). The lower inlay shows the ratio to the nominal prediction in each bin.

G.1.3 Scale uncertainties

The effect of scale variations on the predicted differential distributions for the $WZjj$ -EW6 and $WZjj$ -EW4 processes are shown in Figure G.7. The uncertainties are derived as described in Section 8.2.3.

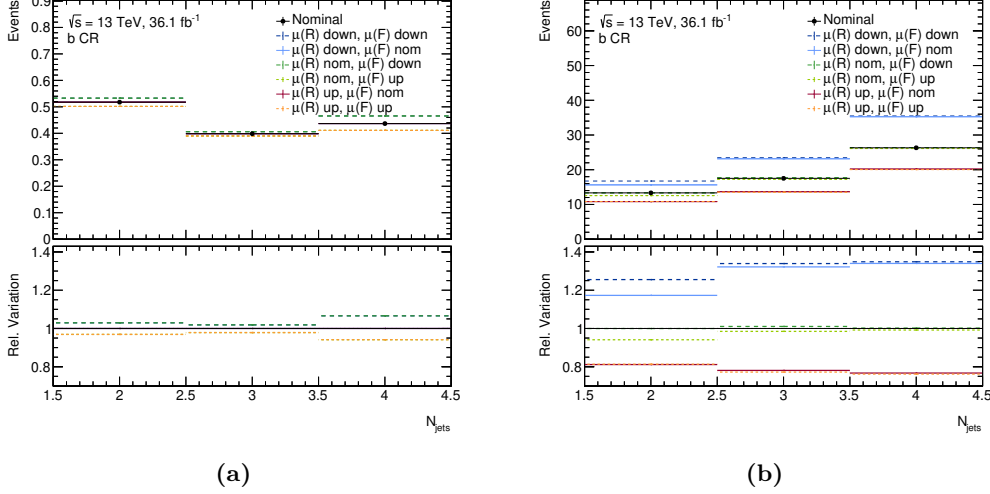


Figure G.7: Uncertainties due to the choice of scales on the predicted event yields differentially in the jet multiplicity for the b control region. Figure a shows the scale uncertainties for the $WZjj$ -EW6 signal process and Figure b for the $WZjj$ -EW4 background process. The prediction using nominal scale choices (full line, black) is compared to the predicted yields using up and down variations of the renormalization (in legend as $\mu(R)$) and factorization scales (in legend as $\mu(F)$). The colours represent different choices of the renormalization scale, where variations with nominal $\mu(R)$ are green, and up (down) variations are blue (red). Variations with the nominal factorization scale are shown in full lines, and up (down) variations of $\mu(F)$ are indicated using a darker (lighter) line colour and dashed (dotted) lines. The lower inlay shows the ratio to the nominal prediction in each bin.

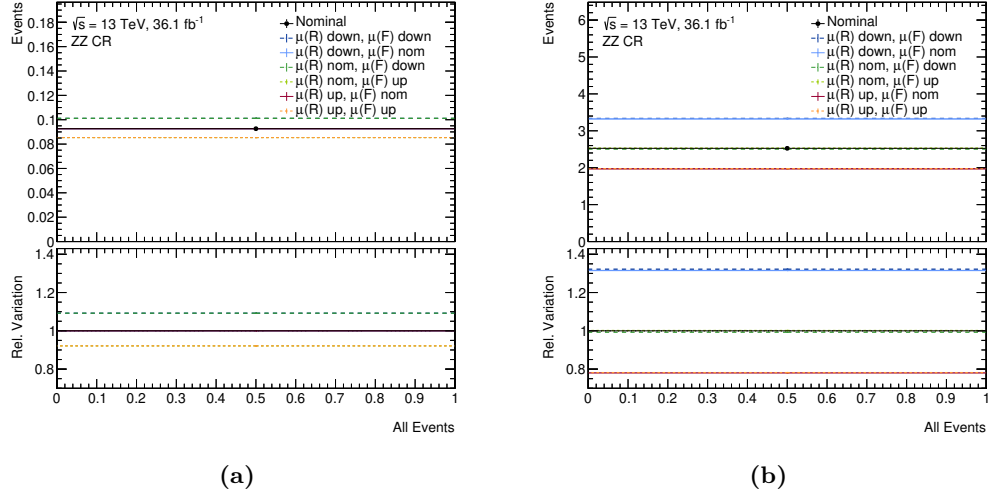


Figure G.8: Uncertainties due to the choice of scales on the predicted event yields for the ZZ control region. Figure a shows the scale uncertainties for the $WZjj$ -EW6 signal process and Figure b for the $WZjj$ -EW4 background process. The prediction using nominal scale choices (full line, black) is compared to the predicted yields using up and down variations of the renormalization (in legend as $\mu(R)$) and factorization scales (in legend as $\mu(F)$). The colours represent different choices of the renormalization scale, where variations with nominal $\mu(R)$ are green, and up (down) variations are blue (red). Variations with the nominal factorization scale are shown in full lines, and up (down) variations of $\mu(F)$ are indicated using a darker (lighter) line colour and dashed (dotted) lines. The lower inlay shows the ratio to the nominal prediction in each bin.

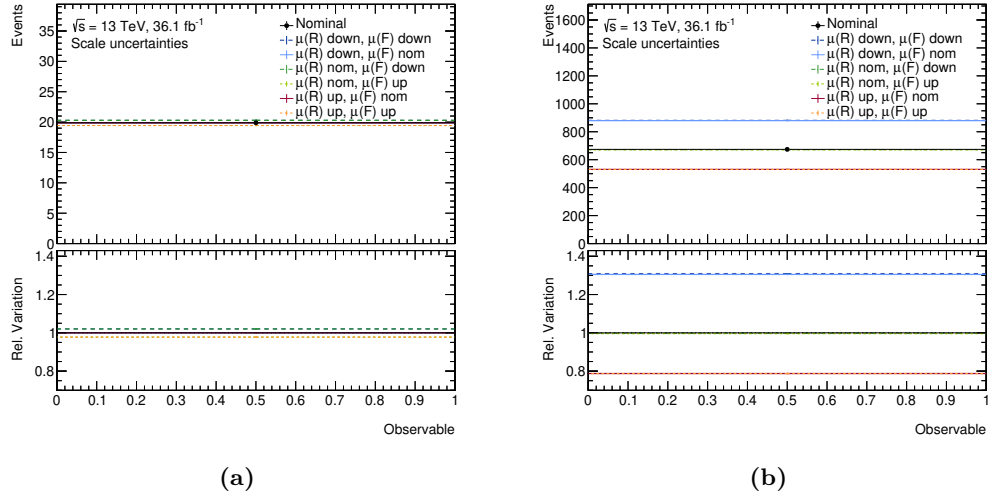


Figure G.9: Uncertainties due to the choice of scales on the predicted event yields for the WZ control region. Figure a shows the scale uncertainties for the $WZjj$ -EW6 signal process and Figure b for the $WZjj$ -EW4 background process. The prediction using nominal scale choices (full line, black) is compared to the predicted yields using up and down variations of the renormalization (in legend as $\mu(R)$) and factorization scales (in legend as $\mu(F)$). The colours represent different choices of the renormalization scale, where variations with nominal $\mu(R)$ are green, and up (down) variations are blue (red). Variations with the nominal factorization scale are shown in full lines, and up (down) variations of $\mu(F)$ are indicated using a darker (lighter) line colour and dashed (dotted) lines. The lower inlay shows the ratio to the nominal prediction in each bin.

G.1.4 Parton Shower and Hadronization

The determination of parton shower, hadronization, and residual modelling uncertainties is process-specific, but in all cases only shape effects without normalization components are derived. Consequently, no uncertainties are determined for the WZ CR and ZZ CR, where the total event yield is used as input to the fit.

The modelling uncertainties for the $WZjj$ -EW6 process in the b CR are shown in Figure G.10.

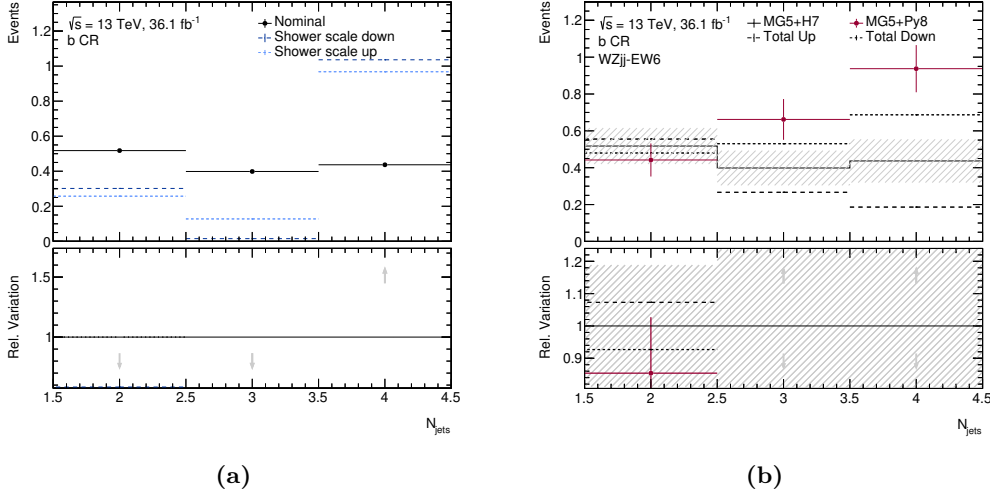


Figure G.10: Uncertainties due to other mismodelling on the predicted event yields differentially in the jet multiplicity in the b CR for the $WZjj$ -EW6 signal process. Figure (a) shows the scale uncertainties in the parton shower simulation and Figure (b) additional shower mismodelling uncertainties estimated from a comparison to an alternative shower generator. The shower scale uncertainties are propagated to the fit as shown and the derived up and down variations for the shower uncertainty are indicated by black dashed and dotted lines. The hatched error band represents the statistical uncertainty on the nominal prediction due to limited number of generated events. The lower inlay shows the ratio to the nominal prediction in each bin.

The modelling uncertainty on the jet multiplicity distribution of the $WZjj$ -EW4 process in the b CR is determined equivalently to the corresponding uncertainty in the signal region described in Section 8.2.4. The response matrix as well as a visualization of the derived uncertainty are shown in Figure G.11.

The uncertainty due to the lack of modelling of the $WZjj$ -EW5 interference contributions to the signal modelling in the b control region are derived as described in Section 8.2.4. The response matrix as well as the final uncertainty are shown in Figure G.12.

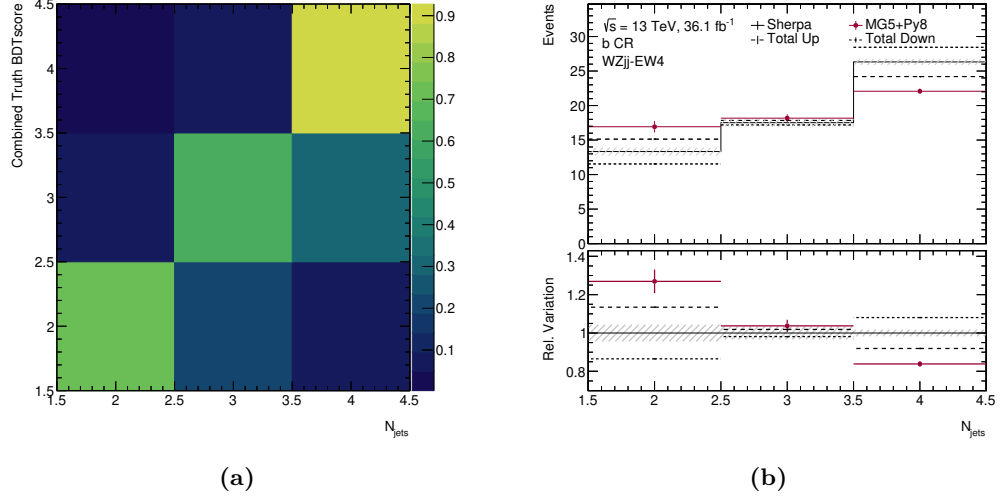


Figure G.11: Comparisons for modelling uncertainty on $WZjj$ -EW4 background process. Figure (a) shows the response matrix for the $BDTscore$ distribution derived from the nominal prediction. Figure (b) shows a comparison of the folded MG5_AMC prediction (red) to the nominal SHERPA prediction (black) as well as the modelling uncertainty (black dashed or dotted lines) derived from the difference between the two predictions. The hatched error band represents the statistical uncertainty on the nominal prediction due to limited number of generated events. The lower inlay shows the ratio to the nominal prediction in each bin.

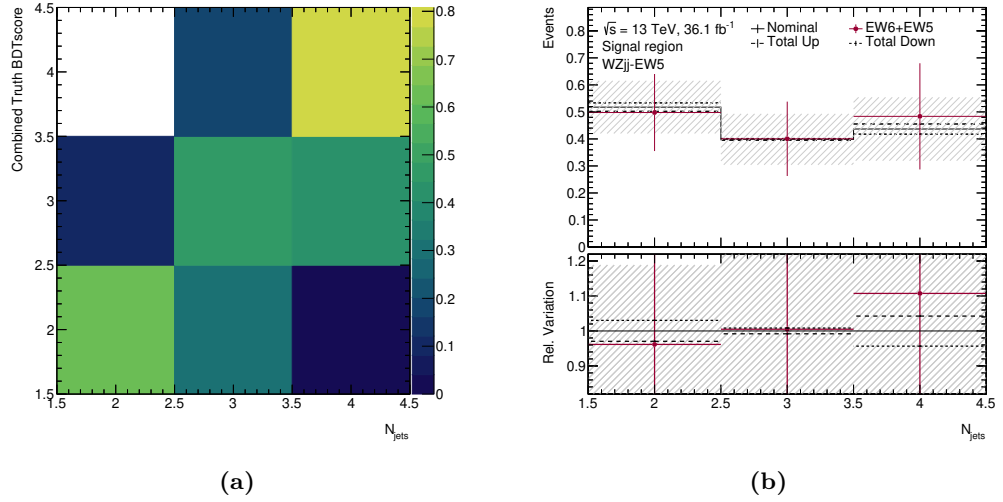


Figure G.12: Comparisons for modelling uncertainty due to the missing contributions of the $WZjj$ -EW5 interferences on the $WZjj$ -EW6 signal process in the b control region. Figure (a) shows the response matrix for the $BDTscore$ distribution derived from the nominal prediction. Figure (b) shows a comparison of the folded prediction when adding the interference contribution (red) to the nominal prediction (black) as well as the modelling uncertainty (black dashed or dotted lines) derived from the difference between the two predictions. The hatched error band represents the statistical uncertainty on the nominal prediction due to limited number of generated events. The lower inlay shows the ratio to the nominal prediction in each bin.

Appendix H

Additional Fit Results

H.1 Fit to Asimov Dataset

H.1.1 Full list of values for nuisance parameters

All nuisance parameters are listed in Table H.1, Table H.2, Table H.3, and Table H.4. For each NP, the initial value with uncertainties, as well as the post-fit values with uncertainties for the background-only fit and for the full signal+background fit for the fits to the Asimov dataset are shown.

Table H.1: Set 1 of 4 of all nuisance parameters of the fit to the Asimov dataset. For each nuisance parameter, the initial value as well as the post-fit values for the background-only and the signal+background fit are shown. Additionally, for each value the associated uncertainties are listed.

Name	Initial values			Background-only fit			S+B fit		
	θ	$\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$	θ	$\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$	θ	$\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$
mu	1						1	0.4672	-0.4354
mu_WZQCD	0.881			0.8592	0.2711	-0.2265	0.881	0.2553	-0.2176
mu_ZZ	1.076			1.0019	0.1833	-0.1758	1.076	0.1747	-0.1616
mu_tZj	1.535			1.8485	1.1603	-0.6076	1.535	0.5818	-0.4904
mu_ttbar	0.961			0.8912	0.2764	-0.2286	0.961	0.2381	-0.2206
Lumi	1	0.0210	-0.0210	1.0003	0.0210	-0.0210	1	0.021	-0.021
gamma_stat_SR_all_bin_0	1	0.0281	-0.0281	1.0008	0.0282	-0.0277	1	0.0282	-0.0276
gamma_stat_SR_all_bin_1	1	0.0203	-0.0203	1.0003	0.0203	-0.0200	1	0.0203	-0.02
gamma_stat_SR_all_bin_2	1	0.0205	-0.0205	1.0009	0.0205	-0.0202	1	0.0205	-0.0202
gamma_stat_SR_all_bin_3	1	0.0264	-0.0264	1.0034	0.0265	-0.0261	1	0.0266	-0.0261
gamma_stat_SR_all_bin_4	1	0.0555	-0.0555	1.0135	0.0565	-0.0545	1	0.0564	-0.0544
gamma_stat_SR_all_bin_5	1	0.1200	-0.1200	1.0408	0.1262	-0.1169	1	0.1247	-0.1151
gamma_stat_SR_all_bin_6	1	0.2553	-0.2553	1.0852	0.2858	-0.2431	1	0.2774	-0.234
gamma_stat_WZQCD-CR_all_bin_0	1	0.0049	-0.0049	0.9997	0.0049	-0.0049	1	0.0049	-0.0049
gamma_stat_ZZ-CR_all_bin_0	1	0.0098	-0.0098	1.0000	0.0098	-0.0098	1	0.0098	-0.0098
gamma_stat_b-CR_all_bin_0	1	0.0286	-0.0286	0.9976	0.0286	-0.0281	1	0.0286	-0.0281
gamma_stat_b-CR_all_bin_1	1	0.0206	-0.0206	0.9988	0.0206	-0.0203	1	0.0206	-0.0203
gamma_stat_b-CR_all_bin_2	1	0.0109	-0.0109	1.0002	0.0109	-0.0109	1	0.0109	-0.0109
alpha_EG_RESOLUTION_ALL	0	1.0000	-1.0000	-0.0051	1.0719	-1.0699	0	0.9989	-0.9993
alpha_EG_SCALE_ALL	0	1.0000	-1.0000	0.0015	1.0730	-1.0736	0	0.9988	-0.9995
alpha_EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	0	1.0000	-1.0000	0.0279	1.0776	-1.0917	0	0.9994	-0.9996
alpha_EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	0	1.0000	-1.0000	0.0352	1.0756	-1.0936	0	0.9994	-0.9996
alpha_EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	0	1.0000	-1.0000	0.0042	1.0837	-1.0858	0	0.9995	-0.9995
alpha_EL_EFF_TriggerEff_TOTAL_1NPCOR_PLUS_UNCOR	0	1.0000	-1.0000	0.0000	1.0848	-1.0847	0	0.9995	-0.9995
alpha_EL_EFF_TriggerEff_TOTAL_1NPCOR_PLUS_UNCOR	0	1.0000	-1.0000	-0.0005	1.0849	-1.0846	0	0.9995	-0.9995

Table H.2: Set 2 of 4 of all nuisance parameters of the fit to the Asimov dataset. For each nuisance parameter, the initial value as well as the post-fit values for the background-only and the signal+background fit are shown. Additionally, for each value the associated uncertainties are listed.

Name	Initial values		Background-only fit		S+B fit	
	θ	$\Delta_{\text{up}}\theta$ $\Delta_{\text{down}}\theta$	θ	$\Delta_{\text{up}}\theta$ $\Delta_{\text{down}}\theta$	θ	$\Delta_{\text{up}}\theta$ $\Delta_{\text{down}}\theta$
alpha_FakeFactor_data_EtaCorrection_WEl	0	1.0000	-1.0000	-0.0911 0.9996	0	0.9995 -0.9995
alpha_FakeFactor_data_EtaCorrection_WMu	0	1.0000	-1.0000	-0.0304 1.0002	0	1 -1
alpha_FakeFactor_data_EtaCorrection_ZEl	0	1.0000	-1.0000	-0.0605 0.9996	0	0.9994 -0.9994
alpha_FakeFactor_data_EtaCorrection_ZMu	0	1.0000	-1.0000	-0.0204 1.0001	0	0.9999 -0.9999
alpha_FakeFactor_data_FakeFactorCRstat	0	1.0000	-1.0000	-0.3835 0.9919	0	0.9922 -0.9921
alpha_FakeFactor_data_FakeFactorExpSyst	0	1.0000	-1.0000	0.0000 1.0004	0	1 -1
alpha_FakeFactor_data_Zgamma_contribution	0	1.0000	-1.0000	-0.0391 0.9359	0	0.9929 -1
alpha_JET_BJES_Response	0	1.0000	-1.0000	0.0000 1.0848	0	0.9995 -0.9995
alpha_JET_EffectiveNP_Detector1	0	1.0000	-1.0000	0.0295 1.0815	0	0.9994 -0.9991
alpha_JET_EffectiveNP_Detector2	0	1.0000	-1.0000	0.0193 1.0746	0	0.9994 -0.9993
alpha_JET_EffectiveNP_Mixed1	0	1.0000	-1.0000	0.0767 1.0413	0	0.9984 -0.9977
alpha_JET_EffectiveNP_Mixed2	0	1.0000	-1.0000	-0.0179 1.0854	0	0.9992 -0.9996
alpha_JET_EffectiveNP_Mixed3	0	1.0000	-1.0000	0.0071 1.0841	0	0.9995 -0.9993
alpha_JET_EffectiveNP_Modelling1	0	1.0000	-1.0000	-0.0279 1.0333	0	0.9938 -0.9984
alpha_JET_EffectiveNP_Modelling2	0	1.0000	-1.0000	0.0695 1.0656	0	0.9992 -0.9994
alpha_JET_EffectiveNP_Modelling3	0	1.0000	-1.0000	-0.0648 1.1063	0	0.9992 -0.9992
alpha_JET_EffectiveNP_Modelling4	0	1.0000	-1.0000	0.0222 1.0875	0	0.9991 -0.9994
alpha_JET_EffectiveNP_Statistical1	0	1.0000	-1.0000	0.0252 1.0965	0	0.9989 -0.9989
alpha_JET_EffectiveNP_Statistical2	0	1.0000	-1.0000	0.0144 1.1067	0	0.9991 -0.9991
alpha_JET_EffectiveNP_Statistical3	0	1.0000	-1.0000	-0.0270 1.0733	0	0.9991 -0.9993
alpha_JET_EffectiveNP_Statistical4	0	1.0000	-1.0000	0.0407 1.0750	0	0.9989 -0.9994
alpha_JET_EffectiveNP_Statistical5	0	1.0000	-1.0000	-0.0044 1.0860	0	0.9992 -0.9995
alpha_JET_EffectiveNP_Statistical6	0	1.0000	-1.0000	-0.0151 1.0828	0	0.9993 -0.9988
alpha_JET_EffectiveNP_Statistical7	0	1.0000	-1.0000	0.0393 1.0808	0	0.9993 -0.9991
alpha_JET_EtaIntercalibration_Modelling	0	1.0000	-1.0000	0.2924 1.0023	0	0.9921 -0.9875

Table H.3: Set 3 of 4 of all nuisance parameters of the fit to the Asimov dataset. For each nuisance parameter, the initial value as well as the post-fit values for the background-only and the signal+background fit are shown. Additionally, for each value the associated uncertainties are listed.

Name	Initial values		Background-only fit		S+B fit	
	θ	$\Delta^{\text{up}}\theta$ $\Delta^{\text{down}}\theta$	θ	$\Delta^{\text{up}}\theta$ $\Delta^{\text{down}}\theta$	θ	$\Delta^{\text{up}}\theta$ $\Delta^{\text{down}}\theta$
alpha_JET_EtaIntercalibration_NonClosure	0	1.0000 -1.0000	-0.1216 0.9411	-0.9540 0	0 0.9984	-0.9965 0
alpha_JET_EtaIntercalibration_TotalStat	0	1.0000 -1.0000	0.1076 1.0707	-1.1066 0	0 0.9985	-0.9988 0
alpha_JET_Flavor_Composition	0	1.0000 -1.0000	0.2364 0.9864	-0.9415 0	0 0.9839	-0.9894 0
alpha_JET_Flavor_Response	0	1.0000 -1.0000	-0.0904 0.9532	-0.9836 0	0 0.9971	-0.9968 0
alpha_JET_JER_SINGLE_NP	0	1.0000 -1.0000	-0.2998 1.2734	-0.9780 0	0 1	-0.9833 0
alpha_JET_Pileup_OffsetMu	0	1.0000 -1.0000	-0.0332 1.0299	-1.0408 0	0 0.997	-0.9985 0
alpha_JET_Pileup_OffsetNPV	0	1.0000 -1.0000	-0.0070 1.0675	-1.0676 0	0 0.9974	-0.9983 0
alpha_JET_Pileup_PtTerm	0	1.0000 -1.0000	0.1379 1.0532	-1.0902 0	0 0.9964	-0.9979 0
alpha_JET_Pileup_RhoTopology	0	1.0000 -1.0000	-0.0188 0.9635	-0.9606 0	0 0.9945	-0.9973 0
alpha_JET_PunchThrough_MC15	0	1.0000 -1.0000	0.0040 1.0330	-1.0337 0	0 0.9995	-0.9995 0
alpha_JET_SingleParticle_HighPt	0	1.0000 -1.0000	0.0000 1.0848	-1.0847 0	0 0.9995	-0.9995 0
alpha_MET_SoftTrk_ResoPara	0	1.0000 -1.0000	0.0568 0.9416	-0.9123 0	0 1	-0.9974 0
alpha_MET_SoftTrk_ResoPerp	0	1.0000 -1.0000	0.0449 0.9541	-0.9374 0	0 1	-0.9938 0
alpha_MET_SoftTrk_Scale	0	1.0000 -1.0000	-0.1345 1.0320	-1.0064 0	0 0.9961	-0.9993 0
alpha_MUON_EFF_STAT	0	1.0000 -1.0000	0.0050 1.0835	-1.0859 0	0 0.9995	-0.9995 0
alpha_MUON_EFF_STAT_LOWPT	0	1.0000 -1.0000	0.0000 1.0848	-1.0847 0	0 0.9995	-0.9995 0
alpha_MUON_EFF_SYS	0	1.0000 -1.0000	-0.0074 1.0864	-1.0826 0	0 0.9995	-0.9995 0
alpha_MUON_EFF_SYS_LOWPT	0	1.0000 -1.0000	0.0000 1.0848	-1.0847 0	0 0.9995	-0.9995 0
alpha_MUON_EFF_TrigStatUncertainty	0	1.0000 -1.0000	-0.0003 1.0849	-1.0847 0	0 0.9995	-0.9995 0
alpha_MUON_EFF_TrigSystUncertainty	0	1.0000 -1.0000	-0.0005 1.0849	-1.0846 0	0 0.9995	-0.9995 0
alpha_MUON_ID	0	1.0000 -1.0000	-0.0569 1.0617	-1.0457 0	0 0.9962	-0.9985 0
alpha_MUON_ISO_STAT	0	1.0000 -1.0000	0.0006 1.0846	-1.0849 0	0 0.9995	-0.9995 0
alpha_MUON_ISO_SYS	0	1.0000 -1.0000	0.0050 1.0834	-1.0859 0	0 0.9995	-0.9995 0
alpha_MUON_MS	0	1.0000 -1.0000	-0.0019 1.0508	-1.0490 0	0 0.9994	-0.9986 0
alpha_MUON_SAGITTA_RESBIAS	0	1.0000 -1.0000	0.0265 1.0324	-1.0386 0	0 0.9992	-0.9994 0

Table H.4: Set 4 of 4 of all nuisance parameters of the fit to the Asimov dataset. For each nuisance parameter, the initial value as well as the post-fit values for the background-only and the signal+background fit are shown. Additionally, for each value the associated uncertainties are listed.

Name	θ	Initial values $\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$	Background-only fit θ	$\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$	θ	S+B fit $\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$
alpha_MUON_SAGITTA_RHO	0	1.0000	-1.0000	-0.0026	1.0674	-1.0663	0	0.9998	-0.9998
alpha_MUON_SCALE	0	1.0000	-1.0000	-0.0100	1.0792	-1.0741	0	0.9996	-0.9993
alpha_MUON_TTVA_STAT	0	1.0000	-1.0000	0.0010	1.0846	-1.0850	0	0.9995	-0.9995
alpha_MUON_TTVA_SYS	0	1.0000	-1.0000	-0.0077	1.0867	-1.0826	0	0.9995	-0.9995
alpha_VW_XS	0	1.0000	-1.0000	0.0111	1.0003	-1.0004	0	1	-1
alpha_WZjj-EW4TheoAlphas	0	1.0000	-1.0000	0.0373	1.0289	-1.0386	0	0.9999	-0.9999
alpha_WZjj-EW4TheoPdf	0	1.0000	-1.0000	0.2868	0.9915	-1.0130	0	0.9994	-0.9998
alpha_WZjj-EW4TheoScale	0	1.0000	-1.0000	0.1041	0.9877	-1.0256	0	0.9938	-0.9973
alpha_WZjj-EW4TheoShower	0	1.0000	-1.0000	0.8330	0.8532	-0.8542	0	0.9047	-0.9109
alpha_WZjj-EW6TheoAlphas	0	1.0000	-1.0000	0.0000	1.0004	-1.0004	0	1	-1
alpha_WZjj-EW6TheoInterference	0	1.0000	-1.0000	0.0337	1.0000	-0.9998	0	0.9997	-0.9997
alpha_WZjj-EW6TheoPdf	0	1.0000	-1.0000	0.0000	1.0004	-1.0004	0	1	-0.9999
alpha_WZjj-EW6TheoScale	0	1.0000	-1.0000	0.0000	1.0004	-1.0004	0	0.9986	-0.9989
alpha_WZjj-EW6TheoShower	0	1.0000	-1.0000	0.0000	1.0004	-1.0004	0	0.9938	-0.9939
alpha_WZjj-EW6TheoShowerscale	0	1.0000	-1.0000	0.0000	1.0004	-1.0004	0	0.8451	-0.8433
alpha_ZZjj-EW6_XS	0	1.0000	-1.0000	0.2619	1.0325	-0.9680	0	0.9998	-0.9998
alpha_systSF_FT_EFF_B_systematics	0	1.0000	-1.0000	-0.4560	1.4730	-0.9961	0	0.9993	-0.9953
alpha_systSF_FT_EFF_C_systematics	0	1.0000	-1.0000	-0.0385	1.3057	-1.2213	0	0.9982	-0.9976
alpha_systSF_FT_EFF_Light_systematics	0	1.0000	-1.0000	-0.1760	1.3911	-1.1292	0	0.9982	-0.9975
alpha_systSF_FT_EFF_extrapolation	0	1.0000	-1.0000	-0.0237	1.2787	-1.2485	0	0.998	-0.9978
alpha_systSF_FT_EFF_extrapolation_from_charm	0	1.0000	-1.0000	0.0186	1.2452	-1.2818	0	0.9979	-0.9979
alpha_systSF_JET_JvtEfficiency	0	1.0000	-1.0000	-0.0129	0.9556	-0.9583	0	0.9993	-0.9995
alpha_systSF_PRW_DATASF	0	1.0000	-1.0000	-0.4811	1.4545	-0.9395	0	0.9817	-0.9846

H.1.2 Detailed post-fit event yields

A detailed breakdown of the post-fit event yields for each input process including separated uncertainties for statistical, experimental, and modelling uncertainties are given in the following tables. All nuisance parameters are adjusted to their post-fit values from the full fit to the Asimov dataset. Table H.5 lists the yields for the ZZ control region, Table H.6 for the WZ control region, Table H.7 for the signal region, and Table H.8 for the b control region.

Table H.5: Post-fit event yields in the ZZ control region for each process for the full fit to the Asimov dataset. The uncertainties for each yield are split in three groups in the following order, statistical uncertainties, experimental uncertainties, and modelling uncertainties, and are determined by combining the effect of all nuisance parameters at their pre-fit values and uncertainties. The two bottom rows show the total expected event yield and the derived distribution of the Asimov dataset.

	total			
WZjj-EW4	2.22	± 0.02	± 0.19	± 0.20
ZZjj-EW4	68	± 1	± 11	± 5
tZj	0.44	± 0.00	± 0.15	± 0.06
ZZjj-EW6	5.0	± 0.0	± 0.0	± 2.0
VVV	0.85	± 0.01	± 0.00	± 0.26
ttbar+V	1.9	± 0.0	± 0.4	± 0.2
Non-prompt	1.2	± 0.0	± 0.0	± 0.6
WZjj-EW6	0.09	± 0.00	± 0.01	± 0.05
Asimov Data	80	± 9	± 0	± 0
Total Expected	80	± 1	± 11	± 6

Table H.6: Post-fit event yields in the WZ control region for each process for the full fit to the Asimov dataset. The uncertainties for each yield are split in three groups in the following order, statistical uncertainties, experimental uncertainties, and modelling uncertainties, and are determined by combining the effect of all nuisance parameters at their pre-fit values and uncertainties. The two bottom rows show the total expected event yield and the derived distribution of the Asimov dataset.

	total			
WZjj-EW4	590	± 0	± 50	± 40
ZZjj-EW4	47	± 0	± 7	± 3
tZj	22	± 0	± 8	± 2
ZZjj-EW6	1.6	± 0.0	± 0.0	± 0.6
VVV	2.7	± 0.0	± 0.0	± 0.8
ttbar+V	22	± 0	± 5	± 3
Non-prompt	41	± 0	± 0	± 18
WZjj-EW6	19.9	± 0.0	± 0.5	± 1.3
Asimov Data	750	± 27	± 0	± 0
Total Expected	750	± 0	± 50	± 50

Table H.7: Post-fit event yields per range in the **BDTscore** in the signal region for each process for the full fit to the Asimov dataset. The uncertainties for each yield are split in three groups in the following order, statistical uncertainties, experimental uncertainties, and modelling uncertainties, and are determined by combining the effect of all nuisance parameters at their pre-fit values and uncertainties. Table (a) shows ranges from $[-1, 0.6]$ and Table (b) for the range $[0.6, 1]$ as well as the total yield. The two bottom rows show the total expected event yield and the derived distribution of the Asimov dataset.

(a)

	$[-1, -0.6)$	$[-0.6, -0.2)$	$[-0.2, 0.2)$	$[0.2, 0.6)$
WZjj-EW4	24.9 \pm 0.7 \pm 2.8 \pm 1.9	36.9 \pm 0.7 \pm 3.2 \pm 2.5	23.6 \pm 0.5 \pm 2.2 \pm 2.0	13.8 \pm 0.4 \pm 2.0 \pm 1.0
ZZjj-EW4	1.93 \pm 0.05 \pm 0.30 \pm 0.33	2.3 \pm 0.0 \pm 0.4 \pm 0.2	1.53 \pm 0.03 \pm 0.24 \pm 0.20	0.93 \pm 0.02 \pm 0.15 \pm 0.13
tZj	0.69 \pm 0.02 \pm 0.24 \pm 0.07	2.3 \pm 0.0 \pm 0.8 \pm 0.3	2.6 \pm 0.1 \pm 0.9 \pm 0.3	2.0 \pm 0.1 \pm 0.7 \pm 0.2
ZZjj-EW6	0.09 \pm 0.00 \pm 0.00 \pm 0.04	0.27 \pm 0.01 \pm 0.00 \pm 0.11	0.39 \pm 0.01 \pm 0.00 \pm 0.16	0.43 \pm 0.01 \pm 0.00 \pm 0.18
VVV	0.075 \pm 0.002 \pm 0.000 \pm 0.026	0.16 \pm 0.00 \pm 0.00 \pm 0.05	0.12 \pm 0.00 \pm 0.00 \pm 0.04	0.083 \pm 0.002 \pm 0.000 \pm 0.026
ttbar+V	0.32 \pm 0.01 \pm 0.08 \pm 0.05	0.90 \pm 0.02 \pm 0.21 \pm 0.12	0.71 \pm 0.01 \pm 0.17 \pm 0.10	0.45 \pm 0.01 \pm 0.11 \pm 0.07
Non-prompt	0.5 \pm 0.0 \pm 0.0 \pm 0.4	2.6 \pm 0.1 \pm 0.0 \pm 0.9	1.1 \pm 0.0 \pm 0.0 \pm 0.6	0.53 \pm 0.01 \pm 0.00 \pm 0.29
WZjj-EW6	0.70 \pm 0.00 \pm 0.19 \pm 0.10	3.4 \pm 0.0 \pm 1.1 \pm 0.2	5.2 \pm 0.0 \pm 1.3 \pm 0.4	7.7 \pm 0.0 \pm 1.0 \pm 0.4
Asimov Data	29 \pm 5 \pm 0 \pm 0	49 \pm 7 \pm 0 \pm 0	35 \pm 6 \pm 0 \pm 0	26 \pm 5 \pm 0 \pm 0
Total Expected	29.2 \pm 0.8 \pm 2.9 \pm 2.2	49 \pm 1 \pm 4 \pm 3	35.3 \pm 0.6 \pm 2.8 \pm 2.7	25.9 \pm 0.5 \pm 2.3 \pm 1.6

(b)

	$[0.6, 0.8)$	$[0.8, 0.9)$	$[0.9, 1]$	total				
WZjj-EW4	2.6 \pm 0.1 \pm 0.4 \pm 0.2	0.50 \pm 0.06 \pm 0.06 \pm 0.07	0.09 \pm 0.02 \pm 0.04 \pm 0.03	102 \pm 1 \pm 9 \pm 7				
ZZjj-EW4	0.18 \pm 0.01 \pm 0.03 \pm 0.04	0.035 \pm 0.004 \pm 0.005 \pm 0.021	0.008 \pm 0.002 \pm 0.001 \pm 0.008	7.0 \pm 0.1 \pm 1.1 \pm 0.6				
tZj	0.84 \pm 0.05 \pm 0.29 \pm 0.12	0.23 \pm 0.03 \pm 0.08 \pm 0.04	0.027 \pm 0.007 \pm 0.009 \pm 0.012	8.7 \pm 0.1 \pm 3.0 \pm 0.8				
ZZjj-EW6	0.23 \pm 0.01 \pm 0.00 \pm 0.10	0.066 \pm 0.008 \pm 0.000 \pm 0.028	0.032 \pm 0.008 \pm 0.000 \pm 0.013	1.5 \pm 0.0 \pm 0.0 \pm 0.6				
VVV	0.014 \pm 0.001 \pm 0.000 \pm 0.007	0.0050 \pm 0.0006 \pm 0.0000 \pm 0.0015	0.0050 \pm 0.0013 \pm 0.0000 \pm 0.0015	0.47 \pm 0.01 \pm 0.00 \pm 0.15				
ttbar+V	0.065 \pm 0.004 \pm 0.016 \pm 0.017	0.000 \pm 0.000 \pm 0.000 \pm 0.009	0.0075 \pm 0.0019 \pm 0.0018 \pm 0.0002	2.5 \pm 0.0 \pm 0.6 \pm 0.3				
Non-prompt	0.02 \pm 0.00 \pm 0.00 \pm 0.05	0.046 \pm 0.005 \pm 0.000 \pm 0.034	0.011 \pm 0.003 \pm 0.000 \pm 0.005	4.8 \pm 0.1 \pm 0.0 \pm 2.1				
WZjj-EW6	5.3 \pm 0.0 \pm 1.4 \pm 0.5	3.1 \pm 0.0 \pm 1.1 \pm 0.3	1.4 \pm 0.0 \pm 0.9 \pm 0.2	27 \pm 0 \pm 7 \pm 1				
Asimov Data	9.3 \pm 3.1 \pm 0.0 \pm 0.0	4.0 \pm 2.0 \pm 0.0 \pm 0.0	1.6 \pm 1.3 \pm 0.0 \pm 0.0	154 \pm 12 \pm 0 \pm 0				
Total Expected	9.3 \pm 0.2 \pm 1.5 \pm 0.7	4.0 \pm 0.1 \pm 1.1 \pm 0.3	1.6 \pm 0.0 \pm 0.9 \pm 0.2	154 \pm 1 \pm 12 \pm 10				

Table H.8: Post-fit event yields per range of the jet multiplicity in the b control region for each process for the full fit to the Asimov dataset. The last column additionally lists the total event yield in the b control region. The uncertainties for each yield are split in three groups in the following order, statistical uncertainties, experimental uncertainties, and modelling uncertainties, and are determined by combining the effect of all nuisance parameters at their pre-fit values and uncertainties. The two bottom rows show the total expected event yield and the derived distribution of the Asimov dataset.

	2				3				≥ 4				total			
WZjj-EW4	11.7	± 0.3	± 1.9	± 1.0	15.4	± 0.3	± 1.4	± 1.1	23.2	± 0.3	± 2.8	± 3.3	50	± 1	± 4	± 4
ZZjj-EW4	1.03	± 0.03	± 0.16	± 0.10	1.47	± 0.03	± 0.23	± 0.10	2.3	± 0.0	± 0.4	± 0.3	4.8	± 0.0	± 0.8	± 0.4
tZj	26	± 1	± 9	± 3	32	± 1	± 11	± 2	29	± 0	± 10	± 4	88	± 1	± 31	± 6
ZZjj-EW6	0.035	± 0.001	± 0.000	± 0.015	0.059	± 0.001	± 0.000	± 0.024	0.11	± 0.00	± 0.00	± 0.04	0.20	± 0.00	± 0.00	± 0.08
VVV	0.039	± 0.001	± 0.000	± 0.013	0.091	± 0.002	± 0.000	± 0.031	0.16	± 0.00	± 0.00	± 0.05	0.29	± 0.00	± 0.00	± 0.09
ttbar+V	6.1	± 0.2	± 1.5	± 0.8	21	± 0	± 5	± 2	94	± 1	± 23	± 7	122	± 1	± 29	± 7
Non-prompt	11.1	± 0.3	± 0.0	± 3.5	10.2	± 0.2	± 0.0	± 3.0	14	± 0	± 0	± 5	35	± 0	± 0	± 11
WZjj-EW6	0.52	± 0.00	± 0.12	± 0.11	0.40	± 0.00	± 0.21	± 0.10	0.44	± 0.00	± 0.35	± 0.09	1.4	± 0.0	± 0.6	± 0.2
Asimov Data	57	± 8	± 0	± 0	81	± 9	± 0	± 0	164	± 13	± 0	± 0	301	± 17	± 0	± 0
Total Expected	57	± 2	± 8	± 5	81	± 2	± 9	± 5	164	± 2	± 19	± 14	301	± 3	± 27	± 20

H.2 Fit to Observed Dataset

H.2.1 Full list of values for nuisance parameters

A comparison of the pre- and post-fit values and uncertainties of all Gaussian NP is shown in Figure H.1 for the full fit as well as the background-only fit.

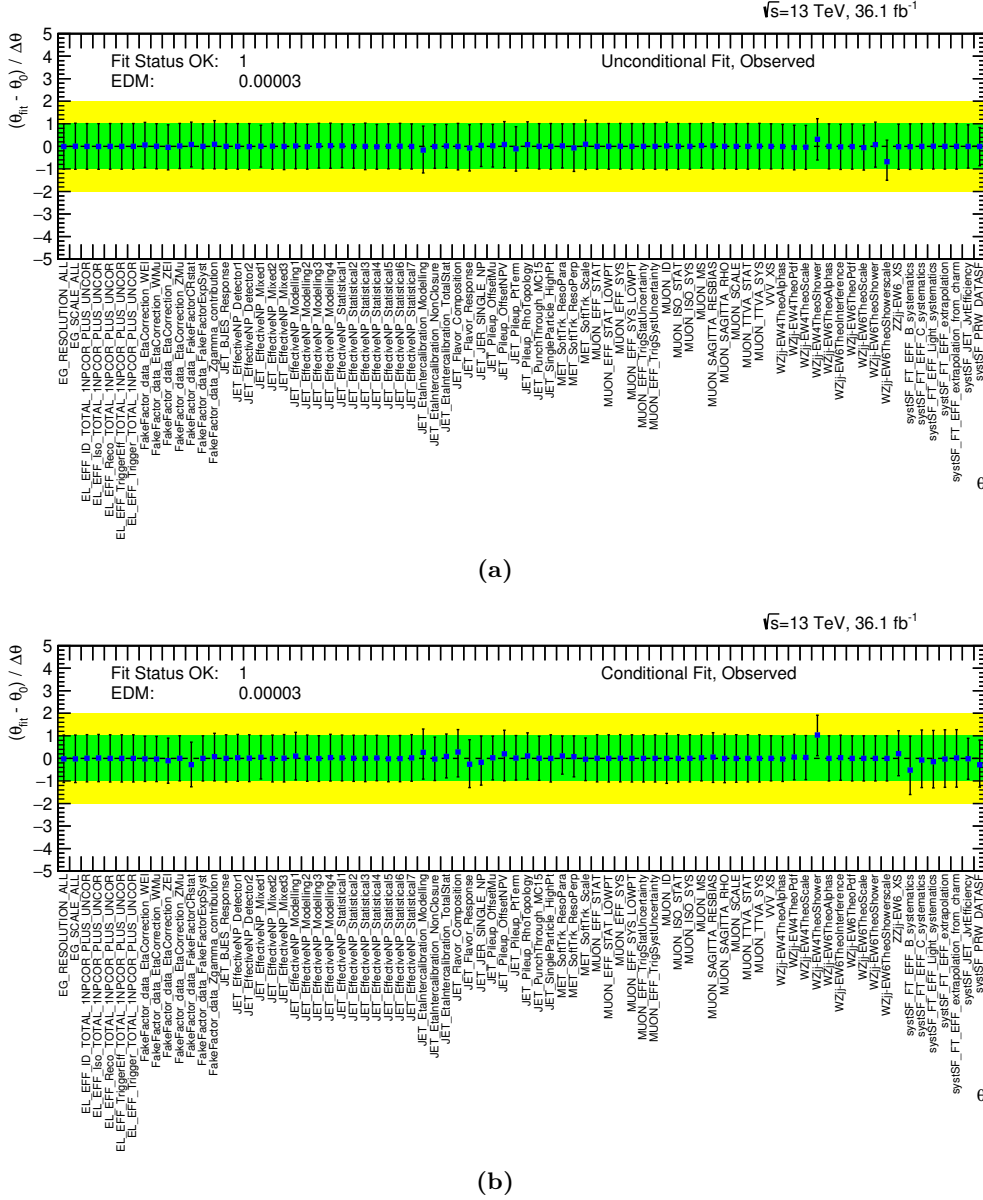


Figure H.1: Comparison of pre and post-fit values of all Gaussian NP α for the fits to the observed dataset. Figure (a) shows the NP for the fit with unconstrained signal strength and Figure (b) the NP for the background-only fit. The green band indicates the pre-fit uncertainties. The black data points indicate the post-fit values and uncertainties.

All nuisance parameters are listed in Table H.9, Table H.10, Table H.11, and Table H.12. For each NP, the initial value with uncertainties, as well as the post-fit values with uncertainties for the background-only fit and for the full signal+background fit for the fits to the observed dataset are shown.

Table H.9: Set 1 of 4 of all nuisance parameters of the fit to the observed dataset. For each nuisance parameter, the initial value as well as the post-fit values for the background-only and the signal+background fit are shown. Additionally, for each value the associated uncertainties are listed.

Name	Initial values			Background-only fit			S+B fit		
	θ	$\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$	θ	$\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$	θ	$\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$
mu	1						1.052	0.4739	-0.3815
mu_WZQCD	0.881			0.8242	0.2455	-0.2175	0.8335	0.2375	-0.2014
mu_ZZ	1.076			1.0088	0.1809	-0.1757	1.0791	0.1751	-0.1615
mu_tZj	1.535			1.9069	1.0897	-0.6375	1.5079	0.5774	-0.4981
mu_ttbar	0.961			0.8936	0.2936	-0.2387	0.98	0.2439	-0.2241
Lumi	1	0.0210	-0.0210	1.0003	0.0210	-0.0210	1	0.021	-0.021
gamma_stat_SR_all_bin_0	1	0.0281	-0.0281	0.9990	0.0282	-0.0277	0.9987	0.0282	-0.0276
gamma_stat_SR_all_bin_1	1	0.0203	-0.0203	0.9931	0.0202	-0.0200	0.9935	0.0202	-0.02
gamma_stat_SR_all_bin_2	1	0.0205	-0.0205	1.0022	0.0205	-0.0203	1.0018	0.0206	-0.0203
gamma_stat_SR_all_bin_3	1	0.0264	-0.0264	1.0025	0.0265	-0.0261	0.9994	0.0266	-0.0261
gamma_stat_SR_all_bin_4	1	0.0555	-0.0555	1.0163	0.0566	-0.0546	1.0007	0.0565	-0.0544
gamma_stat_SR_all_bin_5	1	0.1200	-0.1200	1.0564	0.1271	-0.1178	1.0042	0.1252	-0.1155
gamma_stat_SR_all_bin_6	1	0.2553	-0.2553	0.9858	0.2738	-0.2309	0.9885	0.2743	-0.2314
gamma_stat_WZQCD-CR_all_bin_0	1	0.0049	-0.0049	1.0001	0.0049	-0.0049	1.0003	0.0049	-0.0049
gamma_stat_ZZ-CR_all_bin_0	1	0.0098	-0.0098	1.0000	0.0098	-0.0098	1	0.0098	-0.0098
gamma_stat_b-CR_all_bin_0	1	0.0286	-0.0286	0.9972	0.0286	-0.0280	0.9996	0.0286	-0.0281
gamma_stat_b-CR_all_bin_1	1	0.0206	-0.0206	0.9992	0.0206	-0.0204	1.0004	0.0206	-0.0203
gamma_stat_b-CR_all_bin_2	1	0.0109	-0.0109	1.0002	0.0109	-0.0109	0.9999	0.0109	-0.0109
alpha_EG_RESOLUTION_ALL	0	1.0000	-1.0000	-0.0211	1.0556	-1.0488	-0.0069	1.0105	-1.0103
alpha_EG_SCALE_ALL	0	1.0000	-1.0000	-0.0216	1.0571	-1.0502	0.0052	1.0206	-1.0216
alpha_EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	0	1.0000	-1.0000	0.0072	1.0454	-1.0480	-0.0022	1.0116	-1.0115
alpha_EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	0	1.0000	-1.0000	0.0168	1.0441	-1.0490	-0.0057	1.0116	-1.0113
alpha_EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	0	1.0000	-1.0000	0.0009	1.0463	-1.0468	-0.0002	1.0114	-1.0114
alpha_EL_EFF_TriggerEff_TOTAL_1NPCOR_PLUS_UNCOR	0	1.0000	-1.0000	0.0000	1.0465	-1.0466	0	1.0114	-1.0115
alpha_EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR	0	1.0000	-1.0000	-0.0004	1.0466	-1.0466	0.0001	1.0114	-1.0115

Table H.10: Set 2 of 4 of all nuisance parameters of the fit to the observed dataset. For each nuisance parameter, the initial value as well as the post-fit values for the background-only and the signal+background fit are shown. Additionally, for each value the associated uncertainties are listed.

Name	Initial values			Background-only fit			S+B fit		
	θ	$\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$	θ	$\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$	θ	$\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$
alpha_FakeFactor_data_EtaCorrection_WEl	0	1.0000	-1.0000	-0.0227	1.0046	-0.9995	0.0624	1.0002	-0.999
alpha_FakeFactor_data_EtaCorrection_WMu	0	1.0000	-1.0000	-0.0283	1.0003	-1.0005	0.0039	1	-1
alpha_FakeFactor_data_EtaCorrection_ZEl	0	1.0000	-1.0000	-0.1027	0.9997	-0.9995	-0.0423	0.99	-0.9997
alpha_FakeFactor_data_EtaCorrection_ZMu	0	1.0000	-1.0000	0.0037	1.0002	-1.0004	0.016	0.9999	-1
alpha_FakeFactor_data_FakeFactorCRstat	0	1.0000	-1.0000	-0.2784	0.9980	-0.9811	0.0824	0.9916	-0.992
alpha_FakeFactor_data_FakeFactorExpSyst	0	1.0000	-1.0000	0.0000	1.0005	-1.0005	0	1	-1
alpha_FakeFactor_data_Zgamma_contribution	0	1.0000	-1.0000	0.0772	1.0357	-1.0753	0.0919	1.0496	-1.0886
alpha_JET_BJES_Response	0	1.0000	-1.0000	0.0000	1.0465	-1.0466	0	1.0114	-1.0115
alpha_JET_EffectiveNP_Detector1	0	1.0000	-1.0000	0.0273	1.0352	-1.0416	-0.0006	1.0104	-1.0104
alpha_JET_EffectiveNP_Detector2	0	1.0000	-1.0000	0.0100	1.0369	-1.0392	-0.0127	1.0216	-1.0204
alpha_JET_EffectiveNP_Mixed1	0	1.0000	-1.0000	0.0467	0.9666	-0.9592	0.0063	0.9289	-0.9259
alpha_JET_EffectiveNP_Mixed2	0	1.0000	-1.0000	-0.0064	1.0389	-1.0374	0.0091	1.0198	-1.021
alpha_JET_EffectiveNP_Mixed3	0	1.0000	-1.0000	0.0075	1.0496	-1.0519	-0.0033	1.0287	-1.0283
alpha_JET_EffectiveNP_Modelling1	0	1.0000	-1.0000	0.1017	1.0526	-1.0567	0.0164	1.0035	-1.0098
alpha_JET_EffectiveNP_Modelling2	0	1.0000	-1.0000	0.0118	1.0222	-1.0251	-0.0164	0.993	-0.9935
alpha_JET_EffectiveNP_Modelling3	0	1.0000	-1.0000	-0.0071	1.0134	-1.0126	0.0303	0.9917	-0.9902
alpha_JET_EffectiveNP_Modelling4	0	1.0000	-1.0000	0.0352	1.0322	-1.0398	0.0278	1.0003	-1.0004
alpha_JET_EffectiveNP_Statistical1	0	1.0000	-1.0000	0.0189	1.0156	-1.0168	0.0239	0.9675	-0.9629
alpha_JET_EffectiveNP_Statistical2	0	1.0000	-1.0000	0.0051	1.0215	-1.0240	-0.0044	0.9642	-0.965
alpha_JET_EffectiveNP_Statistical3	0	1.0000	-1.0000	-0.0023	1.0295	-1.0289	-0.0016	1.0307	-1.0309
alpha_JET_EffectiveNP_Statistical4	0	1.0000	-1.0000	0.0166	1.0076	-1.0090	-0.0195	0.9902	-0.9919
alpha_JET_EffectiveNP_Statistical5	0	1.0000	-1.0000	-0.0189	1.0184	-1.0164	0.0007	0.997	-0.997
alpha_JET_EffectiveNP_Statistical6	0	1.0000	-1.0000	-0.0098	1.0256	-1.0239	0.0058	1.0142	-1.0143
alpha_JET_EffectiveNP_Statistical7	0	1.0000	-1.0000	0.0226	1.0306	-1.0354	-0.0044	0.9995	-0.9995
alpha_JET_EtaIntercalibration_Modelling	0	1.0000	-1.0000	0.2665	1.0358	-1.1856	-0.1653	1.0623	-1.0137

Table H.11: Set 3 of 4 of all nuisance parameters of the fit to the observed dataset. For each nuisance parameter, the initial value as well as the post-fit values for the background-only and the signal+background fit are shown. Additionally, for each value the associated uncertainties are listed.

Name	Initial values		Background-only fit		S+B fit	
	θ	$\Delta_{\text{up}}\theta$	θ	$\Delta_{\text{down}}\theta$	$\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$
alpha_JET_EtaIntercalibration_NonClosure	0	1.0000	-0.0287	0.9722	0.9667	-0.9671
alpha_JET_EtaIntercalibration_TotalStat	0	1.0000	0.0865	0.9944	0.9581	-0.9551
alpha_JET_Flavor_Composition	0	1.0000	0.2821	0.9952	1.0222	-1.0382
alpha_JET_Flavor_Response	0	1.0000	-0.2599	1.0881	1.0326	-1.0202
alpha_JET_JER_SINGLE_NP	0	1.0000	-0.1725	1.1628	0.9591	-0.9345
alpha_JET_Pileup_OffsetMu	0	1.0000	0.0265	0.9591	0.9467	-0.9418
alpha_JET_Pileup_OffsetNPV	0	1.0000	0.2017	1.0466	1.0149	-1.0269
alpha_JET_Pileup_PtTerm	0	1.0000	0.0138	0.9984	0.9807	-0.9894
alpha_JET_Pileup_RhoTopology	0	1.0000	0.1080	1.0245	1.0142	-1.0308
alpha_JET_PunchThrough_MC15	0	1.0000	0.0051	0.9783	1.0062	-1.0062
alpha_JET_SingleParticle_HighPt	0	1.0000	0.0000	1.0465	1.0114	-1.0115
alpha_MET_SoftTrk_ResoPara	0	1.0000	0.1068	0.8844	0.98	-0.9778
alpha_MET_SoftTrk_ResoPerp	0	1.0000	0.0766	0.9204	1.0659	-1.0436
alpha_MET_SoftTrk_Scale	0	1.0000	-0.0339	0.9397	1.0676	-1.117
alpha_MUON_EFF_STAT	0	1.0000	0.0059	1.0456	1.0116	-1.0115
alpha_MUON_EFF_STAT_LOWPT	0	1.0000	0.0000	1.0465	1.0114	-1.0115
alpha_MUON_EFF_SYS	0	1.0000	0.0073	1.0453	1.0115	-1.0117
alpha_MUON_EFF_SYS_LOWPT	0	1.0000	0.0000	1.0465	1.0114	-1.0115
alpha_MUON_EFF_TrigStatUncertainty	0	1.0000	-0.0004	1.0465	1.0115	-1.0115
alpha_MUON_EFF_TrigSystUncertainty	0	1.0000	-0.0006	1.0465	1.0114	-1.0114
alpha_MUON_ID	0	1.0000	-0.0006	1.1053	1.0482	-1.0532
alpha_MUON_ISO_STAT	0	1.0000	0.0005	1.0465	1.0115	-1.0115
alpha_MUON_ISO_SYS	0	1.0000	0.0034	1.0460	1.0116	-1.0115
alpha_MUON_MS	0	1.0000	0.0246	1.0210	0.9828	-0.979
alpha_MUON_SAGITTA_RESBIAS	0	1.0000	0.0564	1.0834	1.0259	-1.03

Table H.12: Set 4 of 4 of all nuisance parameters of the fit to the observed dataset. For each nuisance parameter, the initial value as well as the post-fit values for the background-only and the signal+background fit are shown. Additionally, for each value the associated uncertainties are listed.

Name	θ	Initial values $\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$	Background-only fit θ	$\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$	S+B fit $\Delta_{\text{up}}\theta$	θ	$\Delta_{\text{up}}\theta$	$\Delta_{\text{down}}\theta$
alpha_MUON_SAGITTA_RHO	0	1.0000	-1.0000	-0.0026	1.0747	-1.0735	0.989	-0.0002	0.989	-0.9889
alpha_MUON_SCALE	0	1.0000	-1.0000	-0.0037	1.0549	-1.0540	1.0126	-0.0019	1.0126	-1.0122
alpha_MUON_TTVA_STAT	0	1.0000	-1.0000	0.0013	1.0463	-1.0468	1.0114	0.0001	1.0114	-1.0115
alpha_MUON_TTVA_SYS	0	1.0000	-1.0000	-0.0014	1.0467	-1.0465	1.0114	0.0072	1.0114	-1.0118
alpha_VW_XS	0	1.0000	-1.0000	0.0033	1.0005	-1.0006	1	-0.0002	1	-1.0001
alpha_WZjj-EW4TheoAlphas	0	1.0000	-1.0000	-0.0270	0.9934	-0.9957	0.9801	-0.0094	0.9801	-0.981
alpha_WZjj-EW4TheoPdf	0	1.0000	-1.0000	0.0596	1.0025	-1.0073	0.9863	-0.0495	0.9863	-0.9905
alpha_WZjj-EW4TheoScale	0	1.0000	-1.0000	0.0405	0.9703	-0.9825	0.9649	-0.0374	0.9649	-0.9808
alpha_WZjj-EW4TheoShower	0	1.0000	-1.0000	1.0375	0.8761	-0.9187	0.9146	0.3124	0.9146	-0.9176
alpha_WZjj-EW6TheoAlphas	0	1.0000	-1.0000	0.0000	1.0005	-1.0005	0.9988	-0.0022	0.9988	-0.9987
alpha_WZjj-EW6TheoInterference	0	1.0000	-1.0000	0.0300	1.0002	-0.9999	0.9999	-0.0234	0.9999	-1
alpha_WZjj-EW6TheoPdf	0	1.0000	-1.0000	0.0000	1.0005	-1.0005	1.0002	-0.0095	1.0002	-0.9999
alpha_WZjj-EW6TheoScale	0	1.0000	-1.0000	0.0000	1.0005	-1.0005	0.9965	-0.0542	0.9965	-0.9959
alpha_WZjj-EW6TheoShower	0	1.0000	-1.0000	0.0000	1.0005	-1.0005	0.9981	0.076	0.9981	-0.9975
alpha_WZjj-EW6TheoScalescale	0	1.0000	-1.0000	0.0000	1.0005	-1.0005	0.9558	-0.6804	0.9558	-0.8233
alpha_ZZjj-EW6_XS	0	1.0000	-1.0000	0.2097	1.0186	-0.9783	0.9992	-0.0154	0.9992	-1.0003
alpha_systSF_FT_EFF_B_systematics	0	1.0000	-1.0000	-0.5198	1.5031	-1.0835	1.0099	-0.0142	1.0099	-1.0096
alpha_systSF_FT_EFF_C_systematics	0	1.0000	-1.0000	-0.0731	1.3316	-1.2215	1.009	-0.0013	1.009	-1.0097
alpha_systSF_FT_EFF_Light_systematics	0	1.0000	-1.0000	-0.1457	1.3851	-1.1629	1.0082	0.0064	1.0082	-1.0103
alpha_systSF_FT_EFF_extrapolation	0	1.0000	-1.0000	-0.0283	1.2959	-1.2551	1.0089	0.0008	1.0089	-1.0097
alpha_systSF_FT_EFF_extrapolation_from_charm	0	1.0000	-1.0000	0.0190	1.2553	-1.2947	1.0089	0.0001	1.0089	-1.0097
alpha_systSF_JET_JvtEfficiency	0	1.0000	-1.0000	-0.0145	0.9041	-0.9107	0.9592	-0.0031	0.9592	-0.9591
alpha_systSF_PRW_DATASF	0	1.0000	-1.0000	-0.2827	1.0922	-0.9905	0.841	-0.0062	0.841	-0.8514

H.2.2 Correlation between nuisance parameters

The correlations among the nuisance parameters after the full fit to the observed dataset are shown in Figure H.2 for the unconstrained NP only and in Figure H.3 for all NPs.

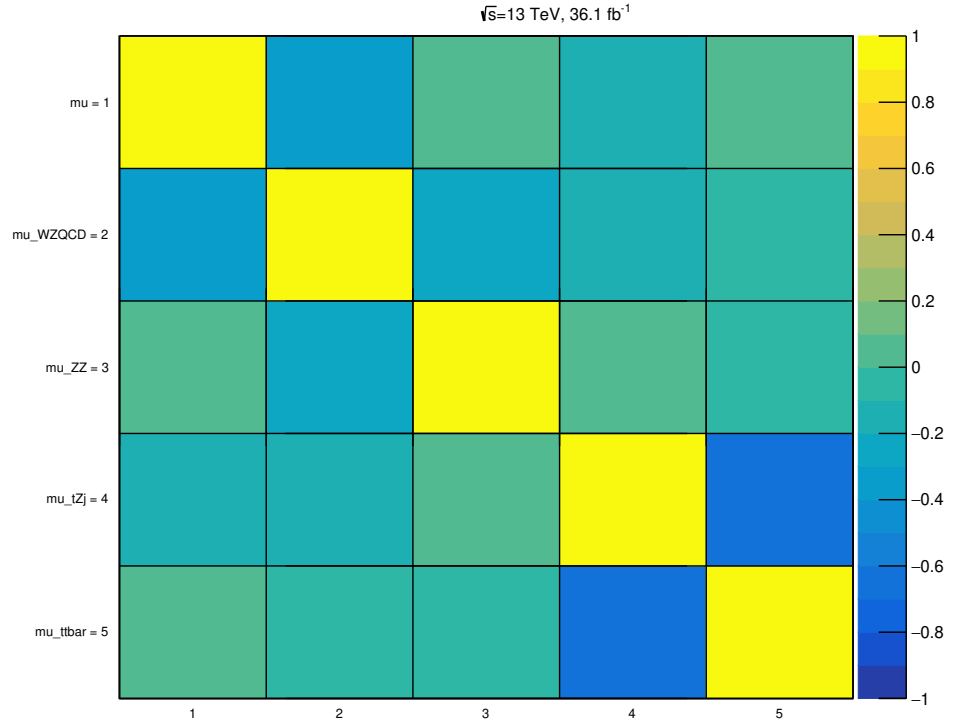


Figure H.2: Correlations among unconstrained nuisance parameters after the full fit to the observed dataset. Nuisance parameter names enumerated in the same order on the x and y axes and their names are only shown on the y axis.

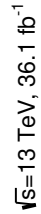


Figure H.3: Correlations among all nuisance parameters after the full fit to the observed dataset. Nuisance parameter names enumerated in the same order on the x and y axes and their names are only shown on the y axis. The figure is rotated to allow for increased size on page.

H.2.3 Full ranking of nuisance parameters

All nuisance parameters of the full fit to the observed dataset are ranked according to their impact on the signal strength parameter. The exact procedure is described in detail in Section 9.4. All nuisance parameters and the re-optimizes signal strengths are shown in Table H.13, Table H.14, Table H.15, and Table H.16.

Table H.13: Impacts on the signal strength μ_S for the nuisance parameters α with the ranks 1 - 25 for the full fit to the observed dataset. The best-fit values and post-fit uncertainties for each NP are listed, as well as the re-optimized signal strengths μ_S when fixing each NP θ to 1σ post- and pre-fit variations around its best-fit value $\hat{\theta}$.

Rank	Nuisance parameter	$\hat{\theta}$	$\hat{\theta} + \Delta^+ \theta$	$\hat{\theta} - \Delta^- \theta$	$\hat{\mu}_S^{\text{up, post}} - \hat{\mu}_S$	$\hat{\mu}_S^{\text{down, post}} - \hat{\mu}_S$	$\hat{\mu}_S^{\text{up, pre}} - \hat{\mu}_S$	$\hat{\mu}_S^{\text{down, pre}} - \hat{\mu}_S$
1	alpha_WZjj-EW6TheoScales	-0.680	0.956	-0.823	-0.246	0.166	-0.262	0.204
2	alpha_WZjj-EW6TheoScales	-0.054	0.997	-0.996	-0.096	0.104	-0.096	0.105
3	alpha_WZjj-EW4TheoShower	0.312	0.915	-0.918	-0.076	0.087	-0.083	0.095
4	alpha_systSF_PRW_DATASF	-0.006	0.841	-0.851	0.084	-0.031	0.101	-0.037
5	alpha_JET_JER_SINGLE_NP	0.040	0.959	-0.935	0.001	0.044	0.001	0.048
6	alpha_MET_SoftTrk_ResoPerp	-0.068	1.066	-1.044	-0.001	0.040	-0.001	0.039
7	alpha_MET_SoftTrk_ResoPara	0.020	0.980	-0.978	0.000	0.040	0.000	0.041
8	alpha_JET_Flavor_Composition	-0.004	1.022	-1.038	-0.008	0.037	-0.007	0.035
9	alpha_systSF_FT_EFF_B_systematics	-0.014	1.010	-1.010	-0.009	-0.033	-0.009	-0.033
10	alpha_MET_SoftTrk_Scale	0.094	1.068	-1.117	0.031	0.009	0.029	0.008
11	alpha_FakeFactor_data_FakeFactorCRstat	0.082	0.992	-0.992	0.028	-0.021	0.028	-0.021
12	alpha_WZjj-EW6TheoPdf	-0.009	1.000	-1.000	-0.027	0.022	-0.027	0.022
13	Lumi	1.000	0.021	-0.021	-0.019	0.027	0.000	0.000
14	alpha_JET_EtaIntercalibration_Modelling	-0.165	1.062	-1.014	-0.017	0.026	-0.016	0.026
15	gamma_stat_SR_all_bin_4	1.001	0.056	-0.054	-0.026	0.019	0.000	0.000
16	alpha_systSF_FT_EFF_Light_systematics	0.006	1.008	-1.010	-0.014	-0.024	-0.014	-0.024
17	gamma_stat_SR_all_bin_5	1.004	0.125	-0.116	-0.023	0.014	0.000	0.000
18	alpha_EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	-0.002	1.012	-1.011	-0.007	0.022	-0.007	0.022
19	gamma_stat_SR_all_bin_3	0.999	0.027	-0.026	-0.022	0.013	0.000	0.000
20	alpha_JET_Pileup_PtTerm	-0.110	0.981	-0.989	-0.008	0.020	-0.009	0.021
21	alpha_systSF_FT_EFF_C_systematics	-0.001	1.009	-1.010	-0.018	-0.020	-0.018	-0.020
22	alpha_systSF_FT_EFF_extrapolation	0.001	1.009	-1.010	-0.019	-0.020	-0.018	-0.020
23	alpha_systSF_FT_EFF_extrapolation_from_charm	0.000	1.009	-1.010	-0.019	-0.019	-0.019	-0.019
24	alpha_systSF_JET_JvtEfficiency	-0.003	0.959	-0.959	0.017	0.017	0.018	0.018
25	alpha_JET_EtaIntercalibration_NonClosure	-0.008	0.967	-0.967	0.017	-0.005	0.018	-0.006

Table H.14: Impacts on the signal strength μ_S for the nuisance parameters α with the ranks 26 - 50 for the full fit to the observed dataset. The best-fit values and post-fit uncertainties for each NP are listed, as well as the re-optimized signal strengths μ_S when fixing each NP θ to 1σ post- and pre-fit variations around its best-fit value $\hat{\theta}$.

Rank	Nuisance parameter	$\hat{\theta}$	$\hat{\theta} + \Delta^+\theta$	$\hat{\theta} - \Delta^-\theta$	$\hat{\mu}_S^{\text{up, post}} - \hat{\mu}_S$	$\hat{\mu}_S^{\text{down, post}} - \hat{\mu}_S$	$\hat{\mu}_S^{\text{up, pre}} - \hat{\mu}_S$	$\hat{\mu}_S^{\text{down, pre}} - \hat{\mu}_S$
26	alpha_ZZjj-EW6_XS	-0.015	0.999	-1.000	-0.016	0.016	-0.016	0.016
27	alpha_JET_Flavor_Response	-0.072	1.033	-1.020	0.016	0.006	0.018	0.006
28	alpha_JET_EffectiveNP_Modelling1	0.016	1.003	-1.010	0.013	0.015	0.013	0.015
29	alpha_WZjj-EW4TheoScale	-0.037	0.965	-0.981	-0.003	0.012	-0.002	0.012
30	alpha_JET_Pileup_RhoTopology	0.074	1.014	-1.031	0.012	0.012	0.012	0.011
31	alpha_JET_Pileup_OffsetMu	0.026	0.947	-0.942	0.010	0.003	0.010	0.003
32	gamma_stat_b-CR_all_bin_0	1.000	0.029	-0.028	0.009	0.003	0.000	0.000
33	alpha_JET_EffectiveNP_Statistical4	-0.019	0.990	-0.992	-0.002	0.009	-0.002	0.009
34	alpha_EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	-0.006	1.012	-1.011	-0.005	0.008	-0.005	0.008
35	alpha_WZjj-EW6TheoShower	0.076	0.998	-0.998	-0.008	0.007	-0.008	0.007
36	gamma_stat_b-CR_all_bin_1	1.000	0.021	-0.020	0.008	0.000	0.000	0.000
37	alpha_FakeFactor_data_Zgamma_contribution	0.092	1.050	-1.089	0.008	0.001	0.008	0.001
38	alpha_JET_EffectiveNP_Statistical3	-0.002	1.031	-1.031	0.004	0.007	0.004	0.007
39	alpha_JET_EtaIntercalibration_TotalStat	0.014	0.958	-0.955	0.003	0.007	0.003	0.014
40	alpha_MUON_MS	0.031	0.983	-0.979	0.007	-0.001	0.007	-0.001
41	alpha_MUON_EFF_SYS	0.005	1.011	-1.012	-0.003	0.006	-0.003	0.006
42	alpha_FakeFactor_data_EtaCorrection_WEL	0.062	1.000	-0.999	0.006	-0.004	0.006	-0.004
43	alpha_JET_EffectiveNP_Modelling3	0.030	0.992	-0.990	0.006	0.001	0.006	0.001
44	alpha_JET_EffectiveNP_Statistical1	0.024	0.967	-0.963	0.006	-0.002	0.006	-0.002
45	gamma_stat_SR_all_bin_1	0.994	0.020	-0.020	0.005	-0.005	0.000	0.000
46	alpha_MUON_EFF_STAT	-0.000	1.012	-1.011	-0.002	0.005	-0.002	0.005
47	alpha_MUON_ISO_SYS	-0.001	1.012	-1.012	-0.002	0.005	-0.002	0.005
48	alpha_WZjj-EW4TheoPdf	-0.050	0.986	-0.990	-0.005	0.004	-0.005	0.009
49	alpha_EG_SCALE_ALL	0.005	1.021	-1.022	0.003	0.005	0.003	0.005
50	alpha_JET_EffectiveNP_Statistical6	0.006	1.014	-1.014	0.005	0.002	0.005	0.002

Table H.15: Impacts on the signal strength μ_S for the nuisance parameters α with the ranks 51 - 75 for the full fit to the observed dataset. The best-fit values and post-fit uncertainties for each NP are listed, as well as the re-optimized signal strengths μ_S when fixing each NP θ to 1σ post- and pre-fit variations around its best-fit value $\hat{\theta}$.

Rank	Nuisance parameter	$\hat{\theta}$	$\hat{\theta} + \Delta^+\theta$	$\hat{\theta} - \Delta^-\theta$	$\hat{\mu}_S^{\text{up, post}} - \hat{\mu}_S$	$\hat{\mu}_S^{\text{down, post}} - \hat{\mu}_S$	$\hat{\mu}_S^{\text{up, pre}} - \hat{\mu}_S$	$\hat{\mu}_S^{\text{down, pre}} - \hat{\mu}_S$
51	alpha_WZjj-EW6TheoInterference	-0.023	1.000	-1.000	0.005	-0.002	0.005	-0.002
52	alpha_JET_EffectiveNP_Detector2	-0.013	1.022	-1.020	0.001	0.005	0.001	0.005
53	gamma_stat_WZQCD-CR_all_bin_0	1.000	0.005	-0.005	0.003	-0.005	0.000	0.000
54	alpha_JET_Pileup_OffsetNPV	0.083	1.015	-1.027	0.002	-0.005	0.002	-0.005
55	alpha_EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	-0.000	1.011	-1.011	-0.002	0.005	-0.002	0.005
56	alpha_MUON_SCALE	-0.002	1.013	-1.012	0.005	-0.000	0.005	-0.000
57	alpha_JET_EffectiveNP_Statistical12	-0.004	0.964	-0.965	0.004	0.004	0.004	0.005
58	alpha_JET_EffectiveNP_Mixed3	-0.003	1.029	-1.028	0.002	0.004	0.002	0.004
59	alpha_JET_EffectiveNP_Mixed2	0.009	1.020	-1.021	0.004	0.004	0.003	0.004
60	alpha_MUON_TTVA_SYS	0.007	1.011	-1.012	-0.002	0.004	-0.002	0.004
61	alpha_JET_EffectiveNP_Modelling2	-0.016	0.993	-0.993	0.003	0.004	0.003	0.004
62	alpha_WZjj-EW4TheoAlphas	-0.009	0.980	-0.981	0.004	0.001	0.004	0.001
63	alpha_JET_EffectiveNP_Statistical15	0.001	0.997	-0.997	0.004	0.002	0.004	0.002
64	alpha_MUON_SAGITTA_RESBIAS	0.022	1.026	-1.030	0.001	-0.003	0.001	-0.003
65	alpha_JET_EffectiveNP_Modelling4	0.028	1.000	-1.000	0.000	0.003	0.000	0.003
66	alpha_MUON_SAGITTA_RHO	-0.000	0.989	-0.989	0.003	0.003	0.003	0.003
67	alpha_JET_EffectiveNP_Mixed1	0.006	0.929	-0.926	0.003	-0.000	0.003	-0.000
68	alpha_JET_EffectiveNP_Detector1	-0.001	1.010	-1.010	-0.000	0.003	-0.000	0.003
69	gamma_stat_b-CR_all_bin_2	1.000	0.011	-0.011	0.003	0.002	0.000	0.000
70	alpha_JET_PunchThrough_MC15	0.001	1.006	-1.006	0.003	0.002	0.003	0.002
71	gamma_stat_ZZ-CR_all_bin_0	1.000	0.010	-0.010	-0.001	-0.002	0.000	0.000
72	alpha_JET_EffectiveNP_Statistical7	-0.004	0.999	-1.000	0.002	0.002	0.002	0.002
73	alpha_EG_RESOLUTION_ALL	-0.007	1.011	-1.010	0.002	0.000	0.002	0.000
74	alpha_MUON_EFF_TrigSystUncertainty	-0.001	1.011	-1.011	0.000	0.002	0.000	0.002
75	alpha_MUON_TTVA_STAT	0.000	1.011	-1.011	-0.000	0.002	0.000	0.002

Table H.16: Impacts on the signal strength μ_S for the nuisance parameters α with the ranks 76 - 93 for the full fit to the observed dataset. The best-fit values and post-fit uncertainties for each NP are listed, as well as the re-optimized signal strengths μ_S when fixing each NP θ to 1σ post- and pre-fit variations around its best-fit value $\hat{\theta}$.

Rank	Nuisance parameter	$\hat{\theta}$	$\hat{\theta} + \Delta^+\theta$	$\hat{\theta} - \Delta^-\theta$	$\hat{\mu}_S^{\text{up, post}}$	$-\hat{\mu}_S$	$\hat{\mu}_S^{\text{down, post}}$	$-\hat{\mu}_S$	$\hat{\mu}_S^{\text{up, pre}}$	$-\hat{\mu}_S$	$\hat{\mu}_S^{\text{down, pre}}$
76	alpha_FakeFactor_data_EtaCorrection_ZMu	0.016	1.000	-1.000	-0.002	0.001	0.001	-0.002	-0.002	0.001	0.001
77	gamma_stat_SR_all_bin_2	1.002	0.021	-0.020	0.001	-0.002	0.000	0.000	0.000	0.000	0.000
78	alpha_WZjj-EW6TheoAlphas	-0.002	0.999	-0.999	-0.001	0.002	0.002	-0.001	-0.001	0.002	0.002
79	alpha_FakeFactor_data_EtaCorrection_ZEl	-0.042	0.999	-1.000	0.001	-0.002	0.001	0.001	0.001	-0.002	-0.002
80	alpha_MUON_EFF_TrigStatUncertainty	-0.001	1.011	-1.011	0.001	0.002	0.002	0.001	0.001	0.002	0.002
81	alpha_MUON_ISO_STAT	-0.000	1.011	-1.011	0.001	0.002	0.002	0.001	0.001	0.001	0.001
82	alpha_MUON_ID	0.017	1.048	-1.053	-0.001	-0.000	-0.001	-0.001	-0.001	-0.000	-0.000
83	alpha_EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR	0.000	1.011	-1.012	0.001	0.001	0.001	0.001	0.001	0.001	0.001
84	alpha_EL_EFF_TriggerEff_TOTAL_1NPCOR_PLUS_UNCOR	0.000	1.011	-1.012	0.001	0.001	0.001	0.001	0.001	0.001	0.001
85	alpha_JET_BJES_Response	0.000	1.011	-1.012	0.001	0.001	0.001	0.001	0.001	0.001	0.001
86	alpha_MUON_EFF_STAT_LOWPT	0.000	1.011	-1.012	0.001	0.001	0.001	0.001	0.001	0.001	0.001
87	alpha_JET_SingleParticle_HighPt	0.000	1.011	-1.012	0.001	0.001	0.001	0.001	0.001	0.001	0.001
88	alpha_MUON_EFF_SYS_LOWPT	0.000	1.011	-1.012	0.001	0.001	0.001	0.001	0.001	0.001	0.001
89	gamma_stat_SR_all_bin_6	0.988	0.274	-0.231	0.001	-0.001	-0.001	0.000	0.000	0.000	0.000
90	gamma_stat_SR_all_bin_0	0.999	0.028	-0.028	-0.000	-0.001	-0.001	0.000	0.000	0.000	0.000
91	alpha_VVV_XS	-0.000	1.000	-1.000	0.000	-0.001	-0.001	0.000	0.000	-0.001	-0.001
92	alpha_FakeFactor_data_EtaCorrection_WMu	0.004	1.000	-1.000	0.000	-0.000	-0.000	0.000	0.000	-0.000	-0.000
93	alpha_FakeFactor_data_FakeFactorExpSyst	0.000	1.000	-1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

H.2.4 Detailed post-fit event yields

A detailed breakdown of the post-fit event yields for each input process including separated uncertainties for statistical, experimental, and modelling uncertainties are given in the following tables. All nuisance parameters are adjusted to their post-fit values from the full fit to the observed dataset. Table H.17 lists the yields for the ZZ control region, Table H.18 for the WZ control region, Table H.19 for the signal region, and Table H.20 for the b control region.

Table H.17: Post-fit event yields in the ZZ control region for each process for the full fit to the observed dataset. The uncertainties for each yield are split in three groups in the following order, statistical uncertainties, experimental uncertainties, and modelling uncertainties, and are determined by combining the effect of all nuisance parameters at their pre-fit values and uncertainties. The two bottom rows show the total expected event yield and the derived distribution of the observed dataset.

	total			
WZjj-EW4	2.08	± 0.02	± 0.17	± 0.17
ZZjj-EW4	69	± 1	± 11	± 5
tZj	0.43	± 0.00	± 0.15	± 0.06
ZZjj-EW6	4.9	± 0.0	± 0.0	± 2.0
VVV	0.85	± 0.01	± 0.00	± 0.26
ttbar+V	1.9	± 0.0	± 0.5	± 0.2
Non-prompt	1.3	± 0.0	± 0.0	± 0.6
WZjj-EW6	0.10	± 0.00	± 0.01	± 0.04
Data	80	± 9	± 0	± 0
Total Expected	80	± 1	± 11	± 6

Table H.18: Post-fit event yields in the WZ control region for each process for the full fit to the observed dataset. The uncertainties for each yield are split in three groups in the following order, statistical uncertainties, experimental uncertainties, and modelling uncertainties, and are determined by combining the effect of all nuisance parameters at their pre-fit values and uncertainties. The two bottom rows show the total expected event yield and the derived distribution of the observed dataset.

	total			
WZjj-EW4	560	± 0	± 50	± 30
ZZjj-EW4	47	± 0	± 7	± 3
tZj	22	± 0	± 8	± 2
ZZjj-EW6	1.6	± 0.0	± 0.0	± 0.6
VVV	2.7	± 0.0	± 0.0	± 0.8
ttbar+V	22	± 0	± 5	± 3
Non-prompt	43	± 0	± 0	± 19
WZjj-EW6	20.9	± 0.0	± 0.5	± 1.3
Data	750	± 27	± 0	± 0
Total Expected	720	± 0	± 50	± 50

Table H.19: Post-fit event yields per range in the **BDTscore** in the signal region for each process for the full fit to the observed dataset. The uncertainties for each yield are split in three groups in the following order, statistical uncertainties, experimental uncertainties, and modelling uncertainties, and are determined by combining the effect of all nuisance parameters at their pre-fit values and uncertainties. Table (a) shows ranges from $[-1, 0.6]$ and Table (b) for the range $[0.6, 1]$ as well as the total yield. The two bottom rows show the total expected event yield and the derived distribution of the observed dataset.

(a)

	[-1, -0.6)				[-0.6, -0.2)				[-0.2, 0.2)				[0.2, 0.6)			
WZjj-EW4	22.6	± 0.6	± 2.5	± 1.8	34.0	± 0.7	± 2.8	± 2.2	22.2	± 0.5	± 2.0	± 1.9	13.3	± 0.3	± 1.8	± 0.9
ZZjj-EW4	1.93	± 0.05	± 0.30	± 0.33	2.3	± 0.0	± 0.4	± 0.2	1.53	± 0.03	± 0.24	± 0.20	0.93	± 0.02	± 0.14	± 0.13
tZj	0.68	± 0.02	± 0.24	± 0.07	2.2	± 0.0	± 0.8	± 0.3	2.5	± 0.1	± 0.9	± 0.3	2.0	± 0.1	± 0.7	± 0.2
ZZjj-EW6	0.085	± 0.002	± 0.000	± 0.036	0.26	± 0.01	± 0.00	± 0.11	0.39	± 0.01	± 0.00	± 0.16	0.43	± 0.01	± 0.00	± 0.17
VVV	0.075	± 0.002	± 0.000	± 0.026	0.16	± 0.00	± 0.00	± 0.05	0.12	± 0.00	± 0.00	± 0.04	0.082	± 0.002	± 0.000	± 0.026
ttbar+V	0.33	± 0.01	± 0.08	± 0.05	0.91	± 0.02	± 0.22	± 0.12	0.73	± 0.01	± 0.17	± 0.10	0.46	± 0.01	± 0.11	± 0.07
Non-prompt	0.6	± 0.0	± 0.0	± 0.4	2.7	± 0.1	± 0.0	± 0.9	1.2	± 0.0	± 0.0	± 0.6	0.57	± 0.02	± 0.00	± 0.32
WZjj-EW6	0.42	± 0.00	± 0.32	± 0.08	1.9	± 0.0	± 1.8	± 0.2	3.3	± 0.0	± 2.2	± 0.3	7.5	± 0.0	± 1.0	± 0.4
Data	29	± 5	± 0	± 0	49	± 7	± 0	± 0	35	± 6	± 0	± 0	26	± 5	± 0	± 0
Total Expected	26.7	± 0.7	± 2.6	± 2.1	44.4	± 0.9	± 3.4	± 2.8	32.1	± 0.6	± 3.1	± 2.4	25.2	± 0.5	± 2.1	± 1.5

(b)

	[0.6, 0.8)				[0.8, 0.9)				[0.9, 1]				total			
WZjj-EW4	2.53	± 0.14	± 0.35	± 0.21	0.49	± 0.06	± 0.06	± 0.06	0.09	± 0.02	± 0.04	± 0.02	95	± 1	± 8	± 7
ZZjj-EW4	0.18	± 0.01	± 0.03	± 0.04	0.034	± 0.004	± 0.005	± 0.022	0.008	± 0.002	± 0.001	± 0.008	6.9	± 0.1	± 1.1	± 0.6
tZj	0.83	± 0.05	± 0.29	± 0.11	0.23	± 0.03	± 0.08	± 0.04	0.026	± 0.007	± 0.009	± 0.011	8.5	± 0.1	± 3.0	± 0.8
ZZjj-EW6	0.23	± 0.01	± 0.00	± 0.10	0.066	± 0.008	± 0.000	± 0.028	0.032	± 0.008	± 0.000	± 0.013	1.5	± 0.0	± 0.0	± 0.6
VVV	0.014	± 0.001	± 0.000	± 0.006	0.0050	± 0.0006	± 0.0000	± 0.0015	0.0050	± 0.0013	± 0.0000	± 0.0015	0.46	± 0.01	± 0.00	± 0.14
ttbar+V	0.066	± 0.004	± 0.016	± 0.017	0.000	± 0.000	± 0.000	± 0.009	0.0075	± 0.0019	± 0.0018	± 0.0002	2.5	± 0.0	± 0.6	± 0.3
Non-prompt	0.03	± 0.00	± 0.00	± 0.05	0.05	± 0.01	± 0.00	± 0.04	0.010	± 0.003	± 0.000	± 0.005	5.1	± 0.1	± 0.0	± 2.2
WZjj-EW6	5.6	± 0.0	± 0.6	± 0.4	2.9	± 0.0	± 0.6	± 0.2	1.1	± 0.0	± 0.5	± 0.1	23	± 0	± 7	± 1
Data	9.3	± 3.1	± 0.0	± 0.0	4.0	± 2.0	± 0.0	± 0.0	1.6	± 1.3	± 0.0	± 0.0	154	± 12	± 0	± 0
Total Expected	9.5	± 0.2	± 0.7	± 0.6	3.7	± 0.1	± 0.6	± 0.3	1.3	± 0.0	± 0.5	± 0.1	143	± 1	± 11	± 9

Table H.20: Post-fit event yields per range of the jet multiplicity in the b control region for each process for the full fit to the observed dataset. The last column additionally lists the total event yield in the b control region. The uncertainties for each yield are split in three groups in the following order, statistical uncertainties, experimental uncertainties, and modelling uncertainties, and are determined by combining the effect of all nuisance parameters at their pre-fit values and uncertainties. The two bottom rows show the total expected event yield and the derived distribution of the observed dataset.

	2				3				≥ 4				total			
WZjj-EW4	11.5	± 0.3	± 1.8	± 1.0	14.5	± 0.3	± 1.3	± 1.0	21.2	± 0.2	± 2.6	± 3.1	47	± 1	± 4	± 4
ZZjj-EW4	1.03	± 0.03	± 0.16	± 0.09	1.48	± 0.03	± 0.23	± 0.10	2.4	± 0.0	± 0.4	± 0.3	4.9	± 0.0	± 0.8	± 0.4
tZj	26	± 1	± 9	± 2	32	± 1	± 11	± 2	29	± 0	± 10	± 4	86	± 1	± 31	± 5
ZZjj-EW6	0.035	± 0.001	± 0.000	± 0.015	0.059	± 0.001	± 0.000	± 0.024	0.11	± 0.00	± 0.00	± 0.04	0.20	± 0.00	± 0.00	± 0.08
VVV	0.039	± 0.001	± 0.000	± 0.013	0.092	± 0.002	± 0.000	± 0.031	0.16	± 0.00	± 0.00	± 0.05	0.29	± 0.00	± 0.00	± 0.09
ttbar+V	6.2	± 0.2	± 1.5	± 0.8	21	± 0	± 5	± 2	96	± 1	± 23	± 7	124	± 1	± 30	± 7
Non-prompt	11.5	± 0.3	± 0.0	± 3.5	10.6	± 0.2	± 0.0	± 3.0	14	± 0	± 0	± 5	36	± 0	± 0	± 11
WZjj-EW6	0.35	± 0.00	± 0.20	± 0.08	0.12	± 0.00	± 0.25	± 0.08	0.44	± 0.00	± 0.22	± 0.08	0.9	± 0.0	± 0.6	± 0.1
Data	57	± 8	± 0	± 0	81	± 9	± 0	± 0	164	± 13	± 0	± 0	301	± 17	± 0	± 0
Total Expected	57	± 2	± 8	± 5	80	± 2	± 9	± 5	164	± 2	± 19	± 14	300	± 3	± 27	± 19

List of Figures

2.1	Feynman graph for the three-point interaction between the physical gauge fields in the electroweak theory.	13
2.2	Feynman graphs for the four-point interaction between the physical gauge fields in electroweak theory.	13
2.3	Feynman diagrams for the three-point (a) and four-point (b) interactions between the physical gauge fields W^\pm and Z and the Higgs field H	15
2.4	Several Standard Model production cross-section measurements compared to the corresponding theoretical expectations.	19
2.5	Leading-order Feynman diagrams for the interactions between two bosons. . .	20
2.6	Dependency of the cross-section of polarization combinations of the scattering $VV \rightarrow VV$	21
2.7	Schematic structure of Feynman diagram for the scattering of electroweak gauge boson scattering at the LHC.	22
2.8	Feynman diagrams for the $VVjj$ -EW6 process at LO	23
2.9	Feynman diagrams for the $VVjj$ -EW4 process at LO	24
2.10	Example Feynman diagram for the tZj process at LO.	26
3.1	Schematic representation of a $t\bar{t}H$ event as produced by an event generator. .	30
3.2	Differential cross-sections of the W^+W^+jj -EW6 process in the $e^+\mu^+$ channel as predicted by SHERPA in different settings.	38
3.3	Schematic structure of a Feynman diagram for electroweak gauge boson scattering with a virtual QCD correction resulting in a 6-point loop.	39
3.4	Schematic structure of a Feynman diagram for the scattering of electroweak gauge boson with a real emission as QCD correction illustrating a soft- γ divergency.	40
3.5	Ratios of an approximated over the full matrix element calculation for double-differential distributions in $M(jj)$ and $ \Delta Y(jj) $ for the $WZjj$ -EW6 process at NLO QCD.	42
3.6	Differential cross-sections for LO+PS simulations of the $WZjj$ -EW6 process in the Loose PS.	53
3.7	Differential cross-sections for LO+PS simulations of the $WZjj$ -EW6 process in the Fiducial PS.	54
3.8	Comparison of the effect of different approximations of the full $WZjj$ -EW6 process on differential cross-sections from the $WZjj$ phase space towards the Fiducial phase space.	55
3.9	Comparison of differential cross-sections with different dynamical scale choices for the full $WZjj$ -EW6 process in the Fiducial phase space.	56
3.10	Comparison of the differential cross-sections for the tZj and the $WZjj$ -EW6 processes from the $WZjj$ phase space towards the Fiducial phase space. . .	57
3.11	Comparison of differential cross-sections predictions for the $WZjj$ -EW4 process in the Fiducial phase space at particle level.	58
3.12	Comparison of the differential cross-sections for different processes with the $\ell\ell\nu jj$ final state and their sum from the $WZjj$ phase space towards the Fiducial phase space.	60

4.1	Schematic layout of the LHC.	62
4.2	The CERN accelerator complex.	63
4.3	The integrated, peak luminosity, the β^* , the crossing angle, the bunch intensity and the number of bunches over the years in LHC Run 2	64
4.4	Cut-away view of the ATLAS detector.	66
4.5	Relation of θ and η	67
4.6	Visualisation of the definition of impact parameters.	68
4.7	Sketch of the ATLAS inner detector showing all its components, including the new insertable B-layer.	69
4.8	Cut-away view of the ATLAS calorimeter system.	70
4.9	Cumulative amount of material, in units of interaction length, as a function of $ \eta $, in different parts of the calorimeter system.	70
4.10	Cross-section of the muon system in a plane containing the beam axis.	71
4.11	Total and main physics stream output bandwidth at the High-Level Trigger as a function of time throughout an example fill.	73
4.12	Fake rate from pileup jets versus hard-scatter jet efficiency curves and Primary vertex dependence of the hard-scatter jet efficiency for fixed cuts of JVT and JVF.	80
5.1	Comparison of predicted and observed differential event yields in the $\ell\ell\nu jj$ region.	92
5.2	Comparison of differential event yields at different stages of the event selection from the $\ell\ell\nu jj$ region towards the $WZjj$ region.	93
5.3	Predicted differential event yields at different stages of the event selection from the $WZjj$ region to the signal region.	94
6.1	Example Feynman diagrams for the leading other prompt background processes next to the $WZjj$ -EW4 and tZj processes.	98
6.2	Differential event yields in the WZ control region.	100
6.3	Differential event yields in the b control region.	101
6.4	Differential event yields in the ZZ control region.	102
6.5	Schematic visualisation of different lepton selections.	103
6.6	Differential event yield in $M(jj)$ as predicted by simulations for the non-prompt background by process.	108
6.7	Comparison of fake factors for different transverse momentum ranges for the $t\bar{t}$, the non-prompt MC, and the reduced data scenarios.	110
6.8	Fractions of truth origins in simulation-based non-prompt prediction for the $t\bar{t}$ CR and the signal region.	112
6.9	Comparison of predicted differential event yields for the $t\bar{t}$ process based on pure simulations to those of the plain MC-based matrix-method.	114
6.10	Comparison of predicted differential event yields for the $t\bar{t}$ process based on pure simulations to those of the MC-based matrix-method with the additional jet requirement mentioned in the text.	115
6.11	Comparison of predicted differential event yields in the invariant mass of the leptons assigned to the Z boson for the $t\bar{t}$ process based on pure simulations to those of the MC-based matrix-method.	116
6.12	Comparison of predicted distributions in the invariant mass of the leptons assigned to the Z boson for the $t\bar{t}$ process in the $t\bar{t}$ control region based on pure simulations to those of the MC-based matrix-method as visualization of the kinematic effect of the fake factor.	117
6.13	Comparison of predicted differential event yields in the invariant mass of the leptons assigned to the Z boson for the $t\bar{t}$ process in the $t\bar{t}$ control region based on pure simulations to those of the MC-based matrix-method for the different leptonic channels.	118

6.14	Comparison of predicted distributions for the $t\bar{t}$ process in the $t\bar{t}$ control region based on pure simulations to those of the MC-based matrix-method before the application of the fake factors with the isolation-based p_T correction.	119
6.15	Comparison of predicted distributions for the $t\bar{t}$ process in the $t\bar{t}$ control region based on pure simulations to those of the MC-based matrix-method with the isolation-based p_T correction.	120
6.16	Comparison of predicted distributions for the $t\bar{t}$ process in the $\ell\ell\nu jj$ region based on pure simulations to those of the MC-based matrix-method with the isolation-based p_T correction.	120
6.17	Correction factors for pseudo-rapidity distribution of the non-prompt lepton candidate.	121
6.18	Comparison of predicted distributions for the $t\bar{t}$ process in the $\ell\ell\nu jj$ region based on pure simulations to those of the MC-based matrix-method with the isolation-based p_T correction and the η correction.	122
6.19	Comparison of predicted distributions for the $t\bar{t}$ process based on simulations to those of the MC-based matrix-method including derived corrections.	123
6.20	Comparison of predicted distributions for the $t\bar{t}$ process in the $WZjj$ region based on simulations to those of the data-driven matrix-method including all derived corrections.	125
7.1	Comparison of distributions differentially in the BDTscore of the baseline BDT for signal and background simulations in the signal region.	128
7.2	Visualization of a single binary decision tree from the set of trees in the baseline BDT.	129
7.3	Comparison of differential predictions for the SHERPA and MG5_AMC-based signal prediction as well as the prediction for the combined prompt backgrounds in different input observables in the signal region.	136
7.4	Comparison of distributions differentially in the BDTscore of the BDT trained using all considered input observables for signal and background simulations in the signal region.	137
7.5	Comparison of differential predictions for the SHERPA and MG5_AMC-based signal prediction as well as the prediction for the combined prompt backgrounds in different input observables after preprocessing in the signal region.	138
7.6	Comparison of distributions differentially in the BDTscore of the BDT trained using all considered input observables after pre-processing for signal and background simulations in the signal region.	140
7.7	Comparison of distributions differentially in the BDTscore of the BDT with an optimized set of input observables and the meta-parameter configuration of the baseline BDT.	141
7.8	Comparison of distributions differentially in the BDTscore of the BDT with an optimized set of input observables and meta-parameters for signal and background simulations in the signal region.	143
7.9	Signal vs background classification performances evaluated on the test dataset using the $f_{Z_{\text{tot}}}$ metric for scans of different meta-parameters of the BDT structure and training.	144
7.10	Comparison of distributions differentially in the BDTscore of the baseline BDT and the optimized final BDT simulations in the signal region.	147
7.11	Comparison of distributions differentially in the truth-level BDTscore of the combined BDT in the fiducial phase space.	149
8.1	PDF uncertainties on event yields differentially in the BDTscore	156
8.2	α_S uncertainties on event yields differentially in the BDTscore	157
8.3	Scale uncertainties on event yields differentially in the BDTscore	158
8.4	Scale uncertainties on event yields differentially in the BDTscore	159
8.5	Comparisons for modelling uncertainty on $WZjj$ -EW4 background process.	160

LIST OF FIGURES

8.6	Comparisons for modelling uncertainty due to the missing contributions of the $WZjj$ -EW5 interferences on the $WZjj$ -EW6 signal process.	161
9.1	Relative predicted event fractions per range in N_{jets} for processes with largest contributions in b control region.	170
9.2	Relative predicted event fractions per range in BDTscore for different subsets of processes in signal region.	170
9.3	Profile likelihood scan in the signal strength parameter μ_S for the fit to Asimov data derived for $\mu_S = 1$	175
9.4	Comparison of pre and post-fit values of all Gaussian NP α for the fit with unconstrained signal strength to the Asimov dataset.	176
9.5	Post-fit event yield distributions for all regions for the fit to Asimov data. . .	177
9.6	Probability distribution function of the test statistic q_0 derived from the asymptotic formulas and from toy experiments.	179
9.7	Comparison of the significance estimates based on the asymptotic formulas and on the toy-experiment.	180
9.8	Comparison of toy-based results to results based on the asymptotic formulas for the low statistics model without including unconstrained NP.	181
9.9	Comparison of toy-based results to results based on the asymptotic formulas for the full fit model.	181
9.10	Profile likelihood scan in the signal strength parameter μ_S for the fit to the observed dataset and the Asimov dataset with $\mu_S = 1$	182
9.11	Comparison of pre and post-fit values of the 15 Gaussian NP α with the largest pulls for the fits to the observed dataset.	183
9.12	Impacts on the signal strength μ_S and comparison of pre and post-fit values for the 20 highest ranked Gaussian nuisance parameters α for the full fit to the observed dataset.	185
9.13	Post-fit event yield distributions for all regions for the fit to observed data. .	186
9.14	Schematic display of example data event with high BDTscore	189
10.1	Post-fit event yields differentially in the BDTscore for the fit to the observed data distribution for this study and [3].	192
10.2	Comparison of the fiducial cross-section measurements of vector boson scattering processes in different final states at 13 TeV.	193
C.1	Example to show details of typical comparison plots. See the main body for a detailed description.	221
D.1	Efficiency of p_T selection requirements estimated in truth level study including linear fit.	227
D.2	Energy scale factors α and additional constant terms c' as a function of $\eta_{\text{calo, probe}}$	230
D.3	Contributions to combined systematic uncertainty of the energy scale factors α as a function of $\eta_{\text{calo, probe}}$	230
D.4	Contributions to combined systematic uncertainty of the additional constant terms c' as a function of $\eta_{\text{calo, probe}}$	231
D.5	Pure data and background subtracted data distribution compared to signal MC distribution scaled to integral of data for m_{ee} for $-2.8 < \eta_{\text{calo, probe}} < -2.7$ with different corrections applied.	232
F.1	Comparison of normalized differential predictions for different signal predictions as well as the combined prompt background prediction in different input observables in the signal region.	243
F.2	Comparison of normalized differential predictions for different signal predictions as well as the combined prompt background prediction in different input observables in the signal region.	244

F.3	Comparison of normalized differential predictions for different signal predictions as well as the combined prompt background prediction in different input observables in the signal region.	245
F.4	Comparison of normalized differential predictions for different signal predictions as well as the combined prompt background prediction in different input observables in the signal region.	246
F.5	Comparison of normalized differential predictions for different signal predictions as well as the combined prompt background prediction in different input observables in the signal region.	247
F.6	Comparison of normalized differential predictions for different signal predictions as well as the combined prompt background prediction in different input observables in the signal region.	248
F.7	Comparison of normalized differential predictions for different signal predictions as well as the combined prompt background prediction in different input observables in the signal region.	249
F.8	Comparison of normalized distributions in the fiducial phase space for the main $WZjj$ -EW4 and $WZjj$ -EW6 process simulations.	250
F.9	Comparison of normalized distributions in the fiducial phase space for the main $WZjj$ -EW4 and $WZjj$ -EW6 process simulations.	251
F.10	Comparison of normalized distributions in the fiducial phase space for the main $WZjj$ -EW4 and $WZjj$ -EW6 process simulations.	252
F.11	Differential event yields in the BDTscore in the WZ control region.	253
F.12	Differential event yields in M_{jj} in the WZ control region.	254
F.13	Differential event yields in $\Delta R(j_1, j_2)$ in the WZ control region.	254
F.14	Differential event yields in $p_T^{\text{bal}}(\ell\ell\nu jj)$ in the WZ control region.	255
F.15	Differential event yields in $ \Delta Y(\ell_W, Z) $ in the WZ control region.	255
F.16	Differential event yields in $p_T(j_2)$ in the WZ control region.	256
F.17	Differential event yields in $\min \Delta R(j_1, \ell)$ in the WZ control region.	256
F.18	Differential event yields in centrality($\ell\ell\ell$) in the WZ control region.	257
F.19	Differential event yields in $p_T(j_1)$ in the WZ control region.	257
F.20	Differential event yields in $\zeta(\ell\ell\ell)$ in the WZ control region.	258
F.21	Differential event yields in $M_T(W)$ in the WZ control region.	258
F.22	Differential event yields in $M(\ell\ell\ell)$ in the WZ control region.	259
F.23	Differential event yields in $Y(\ell_W)$ in the WZ control region.	259
F.24	Differential event yields in $p_T(\ell_W)$ in the WZ control region.	260
F.25	Differential event yields in $p_T(Z)$ in the WZ control region.	260
F.26	Differential event yields in $p_T(\ell_{Z,2})$ in the WZ control region.	261
F.27	Differential event yields in $Y(W)$ in the WZ control region.	261
F.28	Differential event yields in the BDTscore in the signal region.	262
F.29	Differential event yields in M_{jj} in the signal region.	262
F.30	Differential event yields in $\Delta R(j_1, j_2)$ in the signal region.	263
F.31	Differential event yields in $p_T^{\text{bal}}(\ell\ell\nu jj)$ in the signal region.	263
F.32	Differential event yields in $ \Delta Y(\ell_W, Z) $ in the signal region.	264
F.33	Differential event yields in $p_T(j_2)$ in the signal region.	264
F.34	Differential event yields in $\min \Delta R(j_1, \ell)$ in the signal region.	265
F.35	Differential event yields in centrality($\ell\ell\ell$) in the signal region.	265
F.36	Differential event yields in $p_T(j_1)$ in the signal region.	266
F.37	Differential event yields in $\zeta(\ell\ell\ell)$ in the signal region.	266
F.38	Differential event yields in $M_T(W)$ in the signal region.	267
F.39	Differential event yields in $M(\ell\ell\ell)$ in the signal region.	267
F.40	Differential event yields in $Y(\ell_W)$ in the signal region.	268
F.41	Differential event yields in $p_T(\ell_W)$ in the signal region.	268
F.42	Differential event yields in $p_T(Z)$ in the signal region.	269
F.43	Differential event yields in $p_T(\ell_{Z,2})$ in the signal region.	269
F.44	Differential event yields in $Y(W)$ in the signal region.	270

LIST OF FIGURES

G.1	PDF uncertainties on the predicted event yields differentially in the jet multiplicity in b CR	271
G.2	PDF uncertainties on the predicted event yields in ZZ CR	272
G.3	PDF uncertainties on the predicted event yields in WZ CR	272
G.4	α_S uncertainties on the predicted event yields differentially in the jet multiplicity in the b CR	273
G.5	α_S uncertainties on the predicted event yields in the ZZ CR	274
G.6	α_S uncertainties on the predicted event yields in the WZ CR	274
G.7	Scale uncertainties on the predicted event yields differentially in the jet multiplicity in the b CR	275
G.8	Scale uncertainties on the predicted event yields in the ZZ CR	276
G.9	Scale uncertainties on the predicted event yields in the WZ CR	276
G.10	Scale uncertainties on the predicted event yields differentially in the jet multiplicity in the b CR	277
G.11	Comparisons for modelling uncertainty on $WZjj$ -EW4 background process. .	278
G.12	Comparisons for modelling uncertainty due to the missing contributions of the $WZjj$ -EW5 interferences on the $WZjj$ -EW6 signal process in the b CR. . . .	278
H.1	Comparison of pre and post-fit values of all Gaussian NP α for the fits to the observed dataset.	287
H.2	Correlations among unconstrained nuisance parameters after the full fit to the observed dataset.	292
H.3	Correlations among all nuisance parameters after the full fit to the observed dataset.	293

List of Tables

2.1	Elementary particles in the Standard Model.	5
2.2	Gauge symmetries of full Standard Model Lagrangian.	9
2.3	Fermionic charge states in the Standard Model	9
2.4	Summary of available measurements of electroweak gauge boson scattering at 13 TeV in different final states.	25
3.1	Available simulations for the $\ell\ell\nu jj$ final state at particle-level.	49
3.2	Object selection at truth level.	50
3.3	Overview of event selection criteria for the different phase spaces at truth level.	51
3.4	Predicted cross-sections for all available simulations of the $WZjj$ -EW6 process.	52
3.5	Predicted cross-sections for simulations of the $WZjj$ -EW4 process.	56
3.6	Predicted cross-sections for processes with the $\ell\ell\nu jj$ final state.	59
5.1	Overview of different sets of object selection criteria for electrons.	87
5.2	Overview of different sets of object selection criteria for muons.	88
5.3	Overview of object selection criteria for jets.	89
5.4	Overview of event selection criteria for the different regions at detector level.	95
5.5	Overview of observed and predicted event yields per group of processes in each region.	96
6.1	Overview of observed and predicted event yields for an integrated luminosity of 36.1 fb^{-1} per processes in each region.	99
6.2	Overview of observed and predicted event yields for an integrated luminosity of 36.1 fb^{-1} per processes in each control region.	99
6.3	Overview of migration probabilities from the truth-level classification to the detector-level classification.	103
6.4	Overview of event selection criteria for the $t\bar{t}$ control region used for the de- termination of the fake factor.	109
6.5	Detailed overview of the fake factors derived for different ranges in p_T^{probe} in the reduced data scenario.	110
6.6	Overview of predicted number of events for 36.1 fb^{-1} in the $t\bar{t}$ control region.	111
6.7	Overview of predicted number of events for 36.1 fb^{-1} in the $t\bar{t}$ control region.	112
6.8	Overview of predicted number of events for 36.1 fb^{-1} per process for the simulation-based prediction as well as the matrix-method prediction.	124
6.9	Overview of predicted number of events for 36.1 fb^{-1} per process for the simulation-based prediction as well as the matrix-method prediction.	126
7.1	Classification evaluation metrics comparing the SHERPA and MG5_AMC sig- nal predictions for each considered input observable.	135
7.2	Overview of classification evaluation metrics comparing the background and signal predictions for each considered input observable as well as the applied transformations.	139
7.3	List of input observables in the final BDTs listed with to their importance in each of the cross-trained BDTs as evaluated by the TMVA package.	145

LIST OF TABLES

7.4	List of steps during the optimization of the BDT from the baseline BDT towards the final, cross-trained BDTs listed with different performance evaluation metrics.	148
8.1	Relative effect of experimental uncertainty groups on predicted yield in signal region per process.	154
9.1	Predicted event yields per range in the BDTscore in the signal region before applying input corrections.	173
9.2	Predicted event yields per range in the BDTscore in the signal region after input corrections.	173
9.3	Predicted event yields in the control regions after input corrections.	174
9.4	Detailed numbers for the definition of high statistics and low statistics simplified models.	178
9.5	Overview of the kinematic of example data event with high BDTscore	190
B.1	List of MC simulations for the signal process used for reconstruction level analysis.	216
B.2	List of MC simulations for the prompt background process used for reconstruction level analysis.	217
B.3	List of MC simulations for the non-prompt background process other than $V + \gamma$ processes used for reconstruction level analysis.	218
B.4	List of MC simulations for the $W^\pm + \gamma$ non-prompt background process used for reconstruction level analysis.	219
B.5	List of MC simulations for the $Z + \gamma$ non-prompt background process used for reconstruction level analysis.	220
D.1	Absolute values of bin thresholds in η_{calo} for energy scale factors and resolution additional constant terms for Run 2. The same thresholds are used for both sides of the detector.	225
D.2	List of fit results for m_{ee} dependency of efficiency	227
E.1	Overview of different sets of object selection criteria for electrons.	234
E.2	Overview of different sets of object selection criteria for muons.	235
E.3	Overview of predicted number of events for 36.1 fb^{-1} per process in the $t\bar{t}$ control region split by lepton channel for the simulation-based prediction as well as the matrix-method prediction.	237
E.4	Overview of predicted number of events for 36.1 fb^{-1} per process in the $\ell\ell\nu jj$ region split by lepton channel for the simulation-based prediction as well as the matrix-method prediction.	238
E.5	Overview of predicted number of events for 36.1 fb^{-1} per process in the $WZjj$ region split by lepton channel for the simulation-based prediction as well as the matrix-method prediction.	239
E.6	Overview of predicted number of events for 36.1 fb^{-1} per process in the $\ell\ell\nu jj$ region split by lepton channel for the simulation-based prediction as well as the data-driven matrix-method prediction.	240
E.7	Overview of predicted number of events for 36.1 fb^{-1} per process in the $WZjj$ region split by lepton channel for the simulation-based prediction as well as the data-driven matrix-method prediction.	241
E.8	Overview of predicted number of events for 36.1 fb^{-1} per process in the signal region split by lepton channel for the simulation-based prediction as well as the data-driven matrix-method prediction.	242
H.1	Set 1 of 4 of all nuisance parameters of the fit to the Asimov dataset.	280
H.2	Set 2 of 4 of all nuisance parameters of the fit to the Asimov dataset.	281
H.3	Set 3 of 4 of all nuisance parameters of the fit to the Asimov dataset.	282
H.4	Set 4 of 4 of all nuisance parameters of the fit to the Asimov dataset.	283

H.5	Post-fit event yields in the ZZ control region for each process for the full fit to the Asimov dataset.	284
H.6	Post-fit event yields in the WZ control region for each process for the full fit to the Asimov dataset.	284
H.7	Post-fit event yields per range in the BDTscore in the signal region for each process for the full fit to the Asimov dataset.	285
H.8	Post-fit event yields per range of the jet multiplicity in the b control region for each process for the full fit to the Asimov dataset.	286
H.9	Set 1 of 4 of all nuisance parameters of the fit to the observed dataset.	288
H.10	Set 2 of 4 of all nuisance parameters of the fit to the observed dataset.	289
H.11	Set 3 of 4 of all nuisance parameters of the fit to the observed dataset.	290
H.12	Set 4 of 4 of all nuisance parameters of the fit to the observed dataset.	291
H.13	Impacts on the signal strength μ_S for the nuisance parameters α with the ranks 1 - 25 for the full fit to the observed dataset.	295
H.14	Impacts on the signal strength μ_S for the nuisance parameters α with the ranks 26 - 50 for the full fit to the observed dataset.	296
H.15	Impacts on the signal strength μ_S for the nuisance parameters α with the ranks 51 - 75 for the full fit to the observed dataset.	297
H.16	Impacts on the signal strength μ_S for the nuisance parameters α with the ranks 76 - 93 for the full fit to the observed dataset.	298
H.17	Post-fit event yields in the ZZ control region for each process for the full fit to the observed dataset.	299
H.18	Post-fit event yields in the WZ control region for each process for the full fit to the observed dataset.	299
H.19	Post-fit event yields per range in the BDTscore in the signal region for each process for the full fit to the observed dataset.	300
H.20	Post-fit event yields per range of the jet multiplicity in the b control region for each process for the full fit to the observed dataset.	301

Bibliography

- [1] CMS Collaboration, *Observation of Electroweak Production of Same-Sign W Boson Pairs in the Two Jet and Two Same-Sign Lepton Final State in Proton–Proton Collisions at 13 TeV*, Phys. Rev. Lett. **120** (2018) 081801, arXiv: 1709.05822 [hep-ex] (cit. on pp. v, vii, 2, 25, 26).
- [2] ATLAS Collaboration, *Observation of Electroweak Production of a Same-Sign W Boson Pair in Association with Two Jets in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*, Phys. Rev. Lett. **123** (2019) 161801, arXiv: 1906.03203 [hep-ex] (cit. on pp. v, vii, 25, 26, 106, 116, 168, 172).
- [3] ATLAS Collaboration, *Observation of electroweak $W^\pm Z$ boson pair production in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **793** (2019) 469, arXiv: 1812.09740 [hep-ex] (cit. on pp. v, vii, 2, 25, 26, 38, 43, 48, 50, 59, 85, 87, 91, 106, 113, 118, 127, 168, 172, 191, 192).
- [4] G. Kane, *Modern elementary particle physics: explaining and extending the standard model; 2nd ed.* Cambridge University Press, 2017, ISBN: 978-1-107-16508-3, URL: <https://cds.cern.ch/record/2244793> (cit. on pp. 3, 4, 6, 7).
- [5] D. Griffiths, *Introduction to elementary particles*, Wiley, 2008, ISBN: 978-3-527-40601-2, URL: <https://cds.cern.ch/record/1260972> (cit. on p. 3).
- [6] M. Thomson, *Modern particle physics*, Cambridge University Press, 2013, ISBN: 978-1-107-03426-6, URL: <https://cds.cern.ch/record/1529540> (cit. on p. 3).
- [7] J. C. Romão and J. P. Silva, *A Resource for Signs and Feynman Diagrams of the Standard Model*, International Journal of Modern Physics A **27** (2012) 1230025, ISSN: 1793-656X, URL: <http://dx.doi.org/10.1142/S0217751X12300256> (cit. on pp. 3, 17).
- [8] KATRIN Collaboration, *An improved upper limit on the neutrino mass from a direct kinematic method by KATRIN*, Phys. Rev. Lett. **123** (2019) 221802, arXiv: 1909.06048 [hep-ex] (cit. on p. 5).
- [9] M. Tanabashi et al, *Review of Particle Physics*, Phys. Rev. D **98** (3 2018) 030001, URL: <https://link.aps.org/doi/10.1103/PhysRevD.98.030001> (cit. on pp. 5, 16, 19).
- [10] E. Noether, *Invariante Variationsprobleme*, ger, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse (1918) 235, URL: <http://eudml.org/doc/59024> (cit. on p. 4).
- [11] UA1 Collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*, Phys. Lett. **122B** (1 1983) 103 (cit. on p. 11).
- [12] UA2 Collaboration, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti- p p Collider*, Phys. Lett. **122B** (1983) 476, [7.45(1983)] (cit. on p. 11).
- [13] UA1 Collaboration, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/ c^2 at the CERN SPS Collider*, Phys. Lett. **126B** (5 1983) 398 (cit. on p. 11).

- [14] UA2 Collaboration, *Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN anti- p p Collider*, Phys. Lett. **129B** (1-2 1983) 130 (cit. on p. 11).
- [15] DELPHI Collaboration and P. Abreu, *Measurement of the triple-gluon vertex from 4-jet events at LEP*, Zeitschrift für Physik C Particles and Fields **59** (1993) 357, ISSN: 1431-5858, URL: <https://doi.org/10.1007/BF01498617> (cit. on p. 12).
- [16] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (9 1964) 321, URL: <http://link.aps.org/doi/10.1103/PhysRevLett.13.321> (cit. on p. 14).
- [17] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Physics Letters **12** (1964) 132, ISSN: 0031-9163, URL: <http://www.sciencedirect.com/science/article/pii/0031916364911369> (cit. on p. 14).
- [18] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145** (4 1966) 1156, URL: <http://link.aps.org/doi/10.1103/PhysRev.145.1156> (cit. on p. 14).
- [19] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (16 1964) 508, URL: <http://link.aps.org/doi/10.1103/PhysRevLett.13.508> (cit. on p. 14).
- [20] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (20 1964) 585, URL: <http://link.aps.org/doi/10.1103/PhysRevLett.13.585> (cit. on p. 14).
- [21] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model*, Phys. Rept. **457** (2008) 1, arXiv: 0503172 [hep-ph] (cit. on p. 14).
- [22] Y. Nambu, *Quasi-Particles and Gauge Invariance in the Theory of Superconductivity*, Phys. Rev. **117** (3 1960) 648, URL: <http://link.aps.org/doi/10.1103/PhysRev.117.648> (cit. on p. 14).
- [23] J. Goldstone, A. Salam, and S. Weinberg, *Broken Symmetries*, Phys. Rev. **127** (3 1962) 965, URL: <http://link.aps.org/doi/10.1103/PhysRev.127.965> (cit. on p. 14).
- [24] Nobel Media AB 2020, *Press release. NobelPrize.org. 1969*, URL: <https://www.nobelprize.org/prizes/physics/1969/summary/> (visited on 01/02/2020) (cit. on p. 17).
- [25] Nobel Media AB 2020, *Press release. NobelPrize.org. 1979*, URL: <https://www.nobelprize.org/prizes/physics/1979/summary/> (visited on 01/02/2020) (cit. on p. 17).
- [26] Nobel Media AB 2020, *Press release. NobelPrize.org. 2013*, URL: <https://www.nobelprize.org/prizes/physics/2013/press-release/> (visited on 01/02/2020) (cit. on p. 17).
- [27] J. J. Aubert et al, *Experimental Observation of a Heavy Particle J*, Phys. Rev. Lett. **33** (23 1974) 1404, URL: <https://link.aps.org/doi/10.1103/PhysRevLett.33.1404> (cit. on p. 17).
- [28] J. -E. Augustin et al, *Discovery of a Narrow Resonance in e^+e^- Annihilation*, Phys. Rev. Lett. **33** (23 1974) 1406, URL: <https://link.aps.org/doi/10.1103/PhysRevLett.33.1406> (cit. on p. 17).
- [29] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B **716** (2012) 1, ISSN: 0370-2693, arXiv: 1207.7214 [hep-ex], URL: <http://www.sciencedirect.com/science/article/pii/S037026931200857X> (cit. on p. 17).
- [30] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Physics Letters B **716** (2012) 30, ISSN: 0370-2693, arXiv: 1207.7235 [hep-ex], URL: <http://www.sciencedirect.com/science/article/pii/S0370269312008581> (cit. on p. 17).

-
- [31] ATLAS Collaboration, *Standard Model Summary Plots Summer 2019*, ATL-PHYS-PUB-2019-024, 2019, URL: <https://cds.cern.ch/record/2682186> (cit. on p. 19).
 - [32] S. P. Martin, *A Supersymmetry primer*, (1997) 1, [Adv. Ser. Direct. High Energy Phys.18,1(1998)], arXiv: 9709356 [hep-ph] (cit. on p. 18).
 - [33] B. Capdevila, A. Crivellin, S. Descotes-Genon, J. Matias, and J. Virto, *Patterns of New Physics in $b \rightarrow s\ell^+\ell^-$ transitions in the light of recent data*, JHEP **01** (2018) 093, arXiv: 1704.05340 [hep-ph] (cit. on p. 18).
 - [34] A. Pomarol, “Beyond the Standard Model”, *Proceedings, High-energy Physics. Proceedings, 18th European School (ESHEP 2010): Raseborg, Finland, June 20 - July 3, 2010*, 2012 115, arXiv: 1202.1391 [hep-ph] (cit. on p. 19).
 - [35] B. C. Allanach, “Beyond the Standard Model Lectures for the 2016 European School of High-Energy Physics”, *Proceedings, 2016 European School of High-Energy Physics (ESHEP2016): Skeikampen, Norway, June 15-28 2016*, 2017 123, arXiv: 1609.02015 [hep-ph] (cit. on p. 19).
 - [36] ATLAS Collaboration, *SUSY July 2019 Summary Plot Update*, ATL-PHYS-PUB-2019-022, 2019, URL: <https://cds.cern.ch/record/2682063> (cit. on p. 20).
 - [37] C. Bittrich, “Study of Polarization Fractions in the Scattering of Massive Gauge Bosons $W^\pm Z \rightarrow W^\pm Z$ with the ATLAS Detector at the Large Hadron Collider”, Presented 2015, 2015, URL: <https://cds.cern.ch/record/2014124> (cit. on p. 21).
 - [38] CMS Collaboration, *Measurement of electroweak WZ boson production and search for new physics in WZ + two jets events in pp collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **795** (2019) 281, arXiv: 1901.04060 [hep-ex] (cit. on pp. 25, 46, 47, 193).
 - [39] CMS Collaboration, *Measurement of vector boson scattering and constraints on anomalous quartic couplings from events with four leptons and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **774** (2017) 682, arXiv: 1708.02812 [hep-ex] (cit. on p. 25).
 - [40] ATLAS Collaboration, *Observation of electroweak production of two jets in association with a Z -boson pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, ATLAS-CONF-2019-033, 2019, URL: <https://cds.cern.ch/record/2682845> (cit. on pp. 25, 26, 155).
 - [41] CMS Collaboration, *Measurement of the cross section for electroweak production of a Z boson, a photon and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV and constraints on anomalous quartic couplings*, (2020), arXiv: 2002.09902 [hep-ex] (cit. on p. 25).
 - [42] ATLAS Collaboration, *Evidence for electroweak production of two jets in association with a $Z\gamma$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **803** (2020) 135341, arXiv: 1910.09503 [hep-ex] (cit. on p. 25).
 - [43] ATLAS Collaboration, *Search for the electroweak diboson production in association with a high-mass dijet system in semileptonic final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **100** (2019) 032007, arXiv: 1905.07714 [hep-ex] (cit. on p. 25).
 - [44] A Buckley et al, *General-purpose event generators for LHC physics*, Phys. Rept. **504** (2011) 145, arXiv: 1101.2599 [hep-ph] (cit. on p. 29).
 - [45] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph] (cit. on pp. 30, 35).
 - [46] S. Höche, “Introduction to parton-shower event generators”, *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*, 2015 235, arXiv: 1411.4085 [hep-ph] (cit. on p. 32).

- [47] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152 (cit. on p. 32).
- [48] S. Jadach, Z. Was, R. Decker, and J. H. Kuhn, *The tau decay library TAUOLA: Version 2.4*, Comput. Phys. Commun. **76** (1993) 361 (cit. on p. 32).
- [49] S. Höche, F. Krauss, M. Schonherr, and F. Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09** (2012) 049, arXiv: 1111.1220 [hep-ph] (cit. on p. 34).
- [50] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, arXiv: hep-ph/0204244 (cit. on p. 34).
- [51] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: 0409146 [hep-ph] (cit. on p. 34).
- [52] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph] (cit. on p. 34).
- [53] J. Alwall et al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, Eur. Phys. J. **C53** (2008) 473, arXiv: 0706.2569 [hep-ph] (cit. on p. 34).
- [54] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD matrix elements + parton showers*, JHEP **11** (2001) 063, arXiv: 0109231 [hep-ph] (cit. on p. 34).
- [55] L. Lönnblad, *Correcting the color dipole cascade model with fixed order matrix elements*, JHEP **05** (2002) 046, arXiv: 0112284 [hep-ph] (cit. on p. 34).
- [56] E. Bothmann et al., *Event Generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv: 1905.09127 [hep-ph] (cit. on p. 35).
- [57] F. Krauss, R. Kuhn, and G. Soff, *AMEGIC++ 1.0: A Matrix element generator in C++*, JHEP **02** (2002) 044, arXiv: 0109036 [hep-ph] (cit. on p. 35).
- [58] T. Gleisberg et al., *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: 0808.3674 [hep-ph] (cit. on p. 35).
- [59] C. F. Berger et al., *An Automated Implementation of On-Shell Methods for One-Loop Amplitudes*, Phys. Rev. **D78** (2008) 036003, arXiv: 0803.4180 [hep-ph] (cit. on p. 35).
- [60] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: 1111.5206 [hep-ph] (cit. on p. 35).
- [61] B. Biedermann et al., *Automation of NLO QCD and EW corrections with Sherpa and Recola*, Eur. Phys. J. **C77** (2017) 492, arXiv: 1704.05783 [hep-ph] (cit. on p. 35).
- [62] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani–Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv: 0709.1027 [hep-ph] (cit. on p. 35).
- [63] S. Catani and M. H. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, Nucl. Phys. **B485** (1997) 291, [Erratum: Nucl. Phys. B510,503(1998)], arXiv: 9605323 [hep-ph] (cit. on p. 35).
- [64] S. Höche and S. Prestel, *The midpoint between dipole and parton showers*, Eur. Phys. J. **C75** (2015) 461, arXiv: 1506.05057 [hep-ph] (cit. on p. 35).
- [65] S. Höche, F. Krauss, S. Schumann, and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05** (2009) 053, arXiv: 0903.1219 [hep-ph] (cit. on p. 35).
- [66] S. Höche, F. Krauss, M. Schönerr, and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, JHEP **04** (2013) 027, arXiv: 1207.5030 [hep-ph] (cit. on p. 35).
- [67] M. Dobbs and J. B. Hansen, *The HepMC C++ Monte Carlo event record for High Energy Physics*, Comput. Phys. Commun. **134** (2001) 41 (cit. on p. 35).
- [68] J. Alwall et al., *A Standard format for Les Houches event files*, Comput. Phys. Commun. **176** (2007) 300, arXiv: 0609017 [hep-ph] (cit. on p. 35).

-
- [69] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: 1405.0301 [hep-ph] (cit. on p. 35).
 - [70] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5 : Going Beyond*, JHEP **06** (2011) 128, arXiv: 1106.0522 [hep-ph] (cit. on p. 35).
 - [71] R. Frederix et al., *Scalar and pseudoscalar Higgs production in association with a topantitop pair*, Phys. Lett. **B701** (2011) 427, arXiv: 1104.5613 [hep-ph] (cit. on p. 35).
 - [72] R. Frederix et al., *W and Z/ γ^* boson production in association with a bottom-antibottom pair*, JHEP **09** (2011) 061, arXiv: 1106.6019 [hep-ph] (cit. on p. 35).
 - [73] R. Frederix et al., *Merging meets matching in MC@NLO*, JHEP **12** (2012) 061, arXiv: 1209.6215 [hep-ph] (cit. on p. 36).
 - [74] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, *FeynRules 2.0 - A complete toolbox for tree-level phenomenology*, Comput. Phys. Commun. **185** (2014) 2250, arXiv: 1310.1921 [hep-ph] (cit. on p. 36).
 - [75] M. Bahr et al., *Herwig++ physics and manual*, Eur. Phys. J. C **58** (2008) 639, arXiv: 0803.0883 [hep-ph] (cit. on p. 36).
 - [76] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, arXiv: 1512.01178 [hep-ph] (cit. on p. 36).
 - [77] J. Bellm, S. Gieseke, and S. Plätzer, *Merging NLO Multi-jet Calculations with Improved Unitarization*, Eur. Phys. J. **C78** (2018) 244, arXiv: 1705.06700 [hep-ph] (cit. on p. 36).
 - [78] L. Lönnblad and S. Prestel, *Merging Multi-leg NLO Matrix Elements with Parton Showers*, JHEP **03** (2013) 166, arXiv: 1211.7278 [hep-ph] (cit. on p. 36).
 - [79] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph] (cit. on p. 36).
 - [80] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv: 0603175 [hep-ph] (cit. on p. 36).
 - [81] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, *Parton Fragmentation and String Dynamics*, Phys. Rept. **97** (1983) 31 (cit. on p. 36).
 - [82] J. Baglio et al., *Release Note - VBFNLO 2.7.0*, (2014), arXiv: 1404.3940 [hep-ph] (cit. on p. 37).
 - [83] J. Baglio et al., *VBFNLO: A Parton Level Monte Carlo for Processes with Electroweak Bosons – Manual for Version 2.7.0*, (2011), arXiv: 1107.4038 [hep-ph] (cit. on p. 37).
 - [84] K. Arnold et al., *VBFNLO: A Parton level Monte Carlo for processes with electroweak bosons*, Comput. Phys. Commun. **180** (2009) 1661, arXiv: 0811.4559 [hep-ph] (cit. on p. 37).
 - [85] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: 1002.2581 [hep-ph] (cit. on p. 37).
 - [86] W. Kilian, T. Ohl, and J. Reuter, *WHIZARD: Simulating Multi-Particle Processes at LHC and ILC*, Eur. Phys. J. **C71** (2011) 1742, arXiv: 0708.4233 [hep-ph] (cit. on p. 37).
 - [87] M. Moretti, T. Ohl, and J. Reuter, *O'Mega: An Optimizing matrix element generator*, (2001) 1981, arXiv: 0102195 [hep-ph] (cit. on p. 37).
 - [88] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv: 1111.6097 [hep-ph] (cit. on p. 37).
 - [89] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph] (cit. on p. 37).

- [90] A. Ballestrero et al., *Precise predictions for same-sign W-boson scattering at the LHC*, Eur. Phys. J. **C78** (2018) 671, arXiv: 1803.07943 [hep-ph] (cit. on pp. 37, 38, 41, 42, 45, 53).
- [91] *Les Houches 2017: Physics at TeV Colliders Standard Model Working Group Report*, 2018, arXiv: 1803.07977 [hep-ph], URL: <http://lss.fnal.gov/archive/2018/conf/fermilab-conf-18-122-cd-t.pdf> (cit. on pp. 38, 43).
- [92] T. Sjostrand, *Spacelike Showers*, 2019, URL: <http://home.thep.lu.se/~torbjorn/pythia82html/SpacelikeShowers.html#section1> (visited on 11/02/2019) (cit. on p. 38).
- [93] ATLAS Collaboration, *Modelling of the vector boson scattering process $pp \rightarrow W^\pm W^\pm jj$ in Monte Carlo generators in ATLAS*, ATL-PHYS-PUB-2019-004, 2019, URL: <https://cds.cern.ch/record/2655303> (cit. on pp. 38, 45).
- [94] S. Höche, “Status of the Sherpa Event Generator”, Multi-Boson Interactions Workshop, 2018, URL: <https://indico.cern.ch/event/700961/contributions/3069895> (cit. on pp. 38, 44).
- [95] C. Oleari and D. Zeppenfeld, *QCD corrections to electroweak $\nu(l)jj$ and $l+l-jj$ production*, Phys. Rev. **D69** (2004) 093004, arXiv: 0310156 [hep-ph] (cit. on p. 40).
- [96] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, JHEP **04** (2007) 081, arXiv: 0702198 [HEP-PH] (cit. on p. 40).
- [97] B. Jager, A. Karlberg, and J. Scheller, *Parton-shower effects in electroweak $WZjj$ production at the next-to-leading order of QCD*, Eur. Phys. J. **C79** (2019) 226, arXiv: 1812.05118 [hep-ph] (cit. on p. 45).
- [98] A. Denner, S. Dittmaier, P. Maierhöfer, M. Pellen, and C. Schwan, *QCD and electroweak corrections to WZ scattering at the LHC*, JHEP **06** (2019) 067, arXiv: 1904.00882 [hep-ph] (cit. on pp. 46, 161).
- [99] M. Chiesa, A. Denner, J.-N. Lang, and M. Pellen, *An event generator for same-sign W-boson scattering at the LHC including electroweak corrections*, Eur. Phys. J. **C79** (2019) 788, arXiv: 1906.01863 [hep-ph] (cit. on p. 46).
- [100] S. Agostinelli et al., *GEANT4 – a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250, URL: <http://www.sciencedirect.com/science/article/pii/S0168900203013688> (cit. on p. 48).
- [101] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, (2012), URL: <https://cds.cern.ch/record/1474107> (cit. on p. 48).
- [102] V. Hirschi and O. Mattelaer, *Automated event generation for loop-induced processes*, JHEP **10** (2015) 146, arXiv: 1507.00020 [hep-ph] (cit. on p. 55).
- [103] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001, URL: <https://cds.cern.ch/record/1129806> (cit. on pp. 61, 62, 64).
- [104] B. J. Holzer, “Introduction to Particle Accelerators and their Limitations”, *Proceedings, CAS - CERN Accelerator School: Plasma Wake Acceleration: Geneva, Switzerland, November 23-29, 2014*, CERN, 2016 29, arXiv: 1705.09601 [physics.acc-ph] (cit. on p. 61).
- [105] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on pp. 61, 65, 66, 70–72).
- [106] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004 (cit. on p. 61).
- [107] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08002, URL: <http://stacks.iop.org/1748-0221/3/i=08/a=S08002> (cit. on p. 61).

- [108] LHCb Collaboration, *The LHCb Detector at the LHC*, Journal of Instrumentation **3** (2008) S08005, URL: <http://stacks.iop.org/1748-0221/3/i=08/a=S08005> (cit. on p. 61).
- [109] *LEP design report*, Copies shelved as reports in LEP, PS and SPS libraries, CERN, 1984, URL: <https://cds.cern.ch/record/102083> (cit. on p. 61).
- [110] CERN, *Accelerators*, 2019, URL: <https://home.cern/science/accelerators> (visited on 09/09/2019) (cit. on p. 62).
- [111] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl, *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, 2004, URL: <https://cds.cern.ch/record/823808> (cit. on p. 62).
- [112] D. Manglunki, *The LHC Injection Chain*, 2019, URL: <http://bamboo.pv.infn.it/maccacc/cern/Manglunki-2.pdf> (visited on 09/09/2019) (cit. on p. 62).
- [113] E. Mobs, *The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018*, (2018), General Photo, URL: <https://cds.cern.ch/record/2636343> (cit. on p. 63).
- [114] R Bailey and P. Collier, *Standard Filling Schemes for Various LHC Operation Modes*, tech. rep. LHC-PROJECT-NOTE-323, CERN, 2003, URL: <https://cds.cern.ch/record/691782> (cit. on pp. 62, 63).
- [115] X. C. Vidal and R. Cid, *Complex Movement*, 2019, URL: https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.complex_movement (visited on 09/02/2019) (cit. on p. 64).
- [116] R. Bruce et al, “Review of LHC Run 2 Machine Configurations”, 2019, URL: <https://indico.cern.ch/event/751857/contributions/3259403/> (cit. on p. 64).
- [117] M. Lamont, *LHC Performance in Run 2 and Beyond*, PoS **LeptonPhoton2015** (2016) 001. 8 p, URL: <https://cds.cern.ch/record/2257503> (cit. on p. 64).
- [118] S Baird, *Accelerators for pedestrians; rev. version*, tech. rep., CERN, 2007, URL: <https://cds.cern.ch/record/1017689> (cit. on p. 64).
- [119] LHC Commissioning, *LHC schedule*, 2019, URL: <https://lhc-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC-long-term.htm> (visited on 09/03/2019) (cit. on p. 64).
- [120] G. Apollinari, O. Brüning, T. Nakamoto, and L Rossi, *Chapter 1: High Luminosity Large Hadron Collider HL-LHC. High Luminosity Large Hadron Collider HL-LHC*, CERN Yellow Report (2017) 1, 21 pages, chapter in High-Luminosity Large Hadron Collider (HL-LHC) : Preliminary Design Report, URL: <https://cds.cern.ch/record/2120673> (cit. on p. 64).
- [121] ATLAS Collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton-proton collisions at the LHC*, Eur. Phys. J. C **77** (2017) 332, arXiv: 1611.10235 [hep-ex] (cit. on p. 67).
- [122] ATLAS Collaboration, *Vertexing Plots from Initial 13 TeV Runs*, 2015, URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2015-005/> (visited on 09/11/2019) (cit. on p. 67).
- [123] ATLAS Collaboration, *2016 Primary Vertex Performance Plots*, 2016, URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2016-007/> (visited on 09/11/2019) (cit. on p. 67).
- [124] ATLAS Collaboration, *Beam Spot Public Results*, 2015, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/BeamSpotPublicResults#2015_25ns_pp_Collisions_s_13_TeV (visited on 09/11/2019) (cit. on p. 67).
- [125] ATLAS Collaboration, *Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-018, 2015, URL: <https://cds.cern.ch/record/2037683> (cit. on p. 69).

- [126] ATLAS Collaboration, ATLAS-TDR-19-ADD-1, 2012, URL: <https://cds.cern.ch/record/1451888> (cit. on p. 68).
- [127] ATLAS Collaboration, *Efficiency and Hit Spatial Resolution of ATLAS IBL Sensors in LHC Run 2 Collision Events*, ATL-INDET-PUB-2016-001, 2016, URL: <https://cds.cern.ch/record/2203893> (cit. on p. 68).
- [128] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, Journal of Instrumentation **13** (2018) P07017, URL: <https://doi.org/10.1088/2F1748-0221/2F13/2F07/2Fp07017> (cit. on p. 72).
- [129] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019, URL: <https://cds.cern.ch/record/2677054> (cit. on p. 72).
- [130] ATLAS Collaboration, *Trigger Operations Public Results*, 2012, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults#Trig%20ger_Rates_AN1 (visited on 09/17/2019) (cit. on p. 72).
- [131] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex] (cit. on p. 72).
- [132] ATLAS Collaboration, *Trigger Operations Public Results*, 2016, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults#Upda%20ted_Trigger_Rates (visited on 09/17/2019) (cit. on p. 73).
- [133] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 639, arXiv: 1902.04655 [hep-ex] (cit. on pp. 74, 80).
- [134] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, 2008, URL: <https://cds.cern.ch/record/1099735> (cit. on p. 74).
- [135] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, 2012, URL: <https://cds.cern.ch/record/1449796> (cit. on p. 74).
- [136] ATLAS Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv: 1704.07983 [hep-ex] (cit. on p. 74).
- [137] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C **74** (2014) 3071, arXiv: 1407.5063 [hep-ex] (cit. on p. 75).
- [138] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv: 1603.05598 [hep-ex] (cit. on pp. 75, 76, 80).
- [139] J. Illingworth and J. Kittler, *A survey of the hough transform*, Computer Vision, Graphics, and Image Processing **44** (1988) 87, ISSN: 0734-189X, URL: <http://www.sciencedirect.com/science/article/pii/S0734189X88800331> (cit. on p. 75).
- [140] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, arXiv: 1603.02934 [hep-ex] (cit. on p. 77).
- [141] ATLAS Collaboration, *Properties of jets and inputs to jet reconstruction and calibration with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-036, 2015, URL: <https://cds.cern.ch/record/2044564> (cit. on p. 77).
- [142] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002, arXiv: 1703.09665 [hep-ex] (cit. on p. 78).

- [143] ATLAS Collaboration, *Jet energy resolution in proton–proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. C **73** (2013) 2306, arXiv: 1210.6210 [hep-ex] (cit. on p. 78).
- [144] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581, arXiv: 1510.03823 [hep-ex] (cit. on pp. 79, 80).
- [145] ATLAS Collaboration, *Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector*, Eur. Phys. J. C **77** (2017) 580, arXiv: 1705.02211 [hep-ex] (cit. on p. 80), Erratum: Eur. Phys. J. C **77** (2017) 712.
- [146] ATLAS Collaboration, *Performance of b-jet identification in the ATLAS experiment*, JINST **11** (2016) P04008, arXiv: 1512.01094 [hep-ex] (cit. on p. 80).
- [147] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, 2016, URL: <https://cds.cern.ch/record/2160731> (cit. on p. 80).
- [148] ATLAS Collaboration, *Light-quark and gluon jet discrimination in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur. Phys. J. C **74** (2014) 3023, arXiv: 1405.6583 [hep-ex] (cit. on p. 80).
- [149] ATLAS Collaboration, *Quark versus Gluon Jet Tagging Using Charged-Particle Constituent Multiplicity with the ATLAS Detector*, ATL-PHYS-PUB-2017-009, 2017, URL: <https://cds.cern.ch/record/2263679> (cit. on p. 80).
- [150] G. Piacquadio and C. Weiser, *A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS*, J. Phys. Conf. Ser. **119** (2008) 032032 (cit. on p. 80).
- [151] J.-B. De Vivie De Regie et al., *ATLAS electron, photon and muon isolation in Run 2*, tech. rep. ATL-COM-PHYS-2017-290, This note contains the Moriond 2017 recommendations. It will be updated when new recommendations become available.: CERN, 2017, URL: <https://cds.cern.ch/record/2256658> (cit. on p. 80).
- [152] ATLAS Collaboration, *Documentation of IsolationSelectionTool*, 2018, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/IsolationSelectionTool> (cit. on pp. 81, 116).
- [153] I Bird et al., *Update of the Computing Models of the WLCG and the LHC Experiments*, tech. rep. CERN-LHCC-2014-014. LCG-TDR-002, 2014, URL: <https://cds.cern.ch/record/1695401> (cit. on p. 82).
- [154] I Bird et al., *Update of the Computing Models of the WLCG and the LHC Experiments*, tech. rep. CERN-LHCC-2014-014. LCG-TDR-002, 2014, URL: <https://cds.cern.ch/record/1695401> (cit. on p. 82).
- [155] R. Brun and F. Rademakers, *ROOT An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389** (1997) 81, ISSN: 0168-9002 (cit. on p. 82).
- [156] C. Bittrich, *Summary of STDM derivation formats*, 2018, URL: <https://gitlab.cern.ch/cbittric/STDMDerivationSummary> (visited on 10/03/2019) (cit. on p. 83).
- [157] L. Hauswald and F. Socher et al, *Repository for ELCore analysis framework*, 2015, URL: <https://gitlab.cern.ch/ELCore/ELCore> (visited on 10/03/2019) (cit. on p. 84).
- [158] C Burgard, R Gugel, and K Lehmann et al, *ATLAS Common Analysis Framework*, 2014, URL: <https://atlas-caf.web.cern.ch> (cit. on p. 84).
- [159] D Adams et al., *Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014*, tech. rep. ATL-PHYS-INT-2014-018, CERN, 2014, URL: <https://cds.cern.ch/record/1743654> (cit. on p. 90).
- [160] F. Cardillo et al., *Tools for estimating fake/non-prompt lepton backgrounds in ATLAS*, tech. rep. ATL-COM-PHYS-2019-1071, CERN, 1999, URL: <https://cds.cern.ch/record/2686919> (cit. on p. 102).

- [161] ATLAS Collaboration, *Measurement of $W^\pm Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **79** (2019) 535, arXiv: 1902.05759 [hep-ex] (cit. on p. 109).
- [162] A. Hocker et al., *TMVA - Toolkit for Multivariate Data Analysis*, (2007), arXiv: physics/0703039 [physics.data-an] (cit. on pp. 128, 131).
- [163] BaBar Collaboration, “The BABAR physics book: Physics at an asymmetric B factory”, *Workshop on Physics at an Asymmetric B Factory (BaBar Collaboration Meeting) Pasadena, California, September 22-24, 1997*, 1998, URL: <http://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=SLAC-R-504> (cit. on p. 131).
- [164] ATLAS Collaboration, *Measurement of $W^\pm Z$ production in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur. Phys. J. C **72** (2012) 2173, arXiv: 1208.1390 [hep-ex] (cit. on p. 132).
- [165] D. L. Rainwater, R. Szalapski, and D. Zeppenfeld, *Probing color singlet exchange in $Z + \text{two jet events at the CERN LHC}$* , Phys. Rev. **D54** (1996) 6680, arXiv: 9605444 [hep-ph] (cit. on p. 134).
- [166] ATLAS Collaboration, *Systematic Uncertainty Recipes*, 2020, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PmgSystematicUncertaintyRecipes> (visited on 03/11/2020) (cit. on pp. 155, 157).
- [167] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph] (cit. on p. 155).
- [168] G. D. Cowan, *Statistical data analysis*, Oxford Univ. Press, 1998, ISBN: 978-0-198-50155-8, URL: <https://cds.cern.ch/record/358560> (cit. on pp. 163, 168).
- [169] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv: 1007.1727 [physics.data-an] (cit. on pp. 163, 166, 167, 179), Erratum: Eur. Phys. J. C **73** (2013) 2501.
- [170] C. Gumpert, “Measurement of Electroweak Gauge Boson Scattering in the Channel $pp \rightarrow W^\pm W^\pm jj$ with the ATLAS Detector at the Large Hadron Collider”, Presented 27 Feb 2015, 2014, URL: <https://cds.cern.ch/record/2003240> (cit. on p. 163).
- [171] ROOT Collaboration, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, tech. rep. CERN-OPEN-2012-016, New York U., 2012, URL: <https://cds.cern.ch/record/1456844> (cit. on pp. 164, 168).
- [172] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, 2003, arXiv: physics/0306116 [physics.data-an] (cit. on pp. 164, 168).
- [173] A. Wald, *Tests of statistical hypotheses concerning several parameters when the number of observations is large*, Trans. Amer. Math. Soc. **54** (1943) 426 (cit. on p. 167).
- [174] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Annals Math. Statist. **9** (1938) 60 (cit. on p. 167).
- [175] Fred James and Matthias Winkler, *MINUIT User’s Guide*, (2004), URL: <http://seal.web.cern.ch/seal/documents/minuit/mnusersguide.pdf> (cit. on p. 168).
- [176] F. James, *Statistical Methods in Experimental Physics; 2nd ed.* World Scientific, 2006, ISBN: 978-9-812-56795-6, URL: <https://cds.cern.ch/record/1019859> (cit. on p. 168).
- [177] F. Iltzsche, “Observation of Electroweak Same-Charge $W^\pm W^\pm jj$ Production at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC”, Presented 27 Sep 2019, 2019, URL: <https://cds.cern.ch/record/2697175> (cit. on p. 168).
- [178] L. Moneta et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057, arXiv: 1009.1003 [physics.data-an] (cit. on p. 168).

- [179] C. Bittrich and S. Todt et al, *Repository for helper tools developed by local working group*, 2020, URL: <https://gitlab.cern.ch/atlas-germany-dresden-vbs-group/otherTools> (visited on 03/20/2020) (cit. on p. 197).
- [180] R. D. Cousins, J. T. Linnemann, and J. Tucker, *Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **595** (2008) 480, ISSN: 0168-9002, URL: <http://dx.doi.org/10.1016/j.nima.2008.07.086> (cit. on p. 222).
- [181] C. Bittrich et al., *In-situ scale factors from Zee events*, tech. rep. ATL-COM-PHYS-2017-757, CERN, 2017, URL: <https://cds.cern.ch/record/2268800> (cit. on pp. 223, 229, 231).
- [182] S. Das, *A simple alternative to the Crystal Ball function*, (2016), arXiv: 1603.08591 [hep-ex] (cit. on p. 229).

Acknowledgements

Over the past few years, I had the great pleasure to work in an incredibly open, friendly, optimistic and yet productive environment for which I'm very thankful. Being surrounded by kind and intelligent people was a huge motivation to keep working, to learn, and to improve me.

I want to thank my advisor Michael Kobel for shaping this environment and for allowing me to work in this environment. He has guided my way and supported me on it for many years. I'm thankful for being able to work on such an exciting and diverse topic. I very much liked it when I started my Bachelor thesis in 2012 and never lost my interest in it.

I would like to thank also Sabine Lammers, who was not only reviewing the ATLAS publication but also agreed to be a referee for this thesis.

This work would not have been possible without all of my past and present supervisors in the VBS group in Dresden. Thank you to Philipp Anger, Monica Dunford, Christian Gumpert, Joany Manjarres, Ulrike Schnoor, Felix Socher, and Anja Vest for all the time and work you have invested. I have learned so much from each and every one of you and always enjoyed it.

This thank you of course extends to all current members, friends, and colleagues of the VBS group: Fawaz Alhawiti, Tim Herrmann, Abhishek Nag, Jan-Eric Nitschke, and Max Stange. Thank you for all the on- and off-topic discussions, the advice, and procrastination opportunities. Be it in the office or on the Kicker table.

A special thank you goes to Franziska Iltzsche. I enjoyed working with you very much, and it was always fun to chat about everything under the sun. From physics, over layout and coding, to parenting. The office wasn't the same without you.

And then there is my long-term-colleague Stefanie Todt. We went together through all stages, starting as shy newbies to now as the old PhD students. I remember so many things from our GK interview to recruitment events, from shivering before our first DPG talks to shivering before our 35c3 talk. Thank you for everything.

All the best with your growing families to the two of you. Thank you for every discussion, for every question and every answer, for putting me back in my place and for your support. Thank you for enduring me.

Of course, I also want to thank all of the IKTP, especially Frank Siegert and his group, especially Johannes Krause, Sebastian Liebschner, and Heberth Torres for all of the answers and the time for physics questions and at the kicker table, Uta Bilow and the whole Outreach team keeping up the spirit of the "institute of awesomeness", and the administration team for making things possible.

I also want to thank my collaborators at CERN, especially in Thessaloniki and Annecy. I look back proudly on our fruitful work!

I want to thank everyone who helped shape this thesis by providing comments and suggestions: Tim Herrmann, Michael Kobel, Wojciech Kotlarski, Joany Manjarres, Abhishek Nag, Jan-Eric Nitschke, Frank Siegert, Felix Socher, Max Stange, and Heberth Torres.

Ganz speziell möchte ich auch meinen Eltern und Robert danken für euer Interesse und eure Unterstützung. Meine Zeiteinschätzungen waren nicht immer sehr genau und das Erklärungen des Themas manchmal nicht hilfreich. Danke, dass ihr euch trotzdem (immernoch) dafür interessiert!

Der grössten Dank geht an meine kleine Familie: Jane und Nele. Jane, wir kennen uns seit Ewigkeiten und haben zusammen schon so ziemlich alles durchgemacht. Vielen Dank dass du durch dick und dünn immer an meiner Seite bist. Nele, du gibts mir die Kraft und die Motivation immer weiter zu arbeiten und machst mich unglaublich glücklich.

Ich liebe euch! Vielen dank für eure Unterstützung und Liebe, auch wenn ich einige Abende und Urlaube, noch bei der Arbeit statt bei euch war. Mein "großes Buch" ist nun endlich fertig.

Versicherung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die vorliegende Dissertation wurde in der Zeit von Juni 2015 bis April 2020 im Institut für Kern- und Teilchenphysik unter der wissenschaftlichen Betreuung von Prof. Dr. Michael Kobel angefertigt.

Es haben keine früheren erfolglosen Promotionsverfahren stattgefunden.

Ich erkenne die Promotionsordnung des Bereichs Mathematik und Naturwissenschaften an der Technischen Universität Dresden von 23.02.2011 an.

Carsten Bittrich
Dresden, 08.04.2020