



SEARCH FOR ANOMALOUS COUPLINGS IN WW  
AND WZ DECAYS IN THE SEMILEPTONIC  
CHANNEL WITH THE CMS DETECTOR

Zur Erlangung des akademischen Grades eines  
DOKTORS DER NATURWISSENSCHAFTEN  
bei der Fakultät für Physik des  
Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

**Dipl.-Phys. Ivan Shvetsov**  
aus Vladimir

Mündliche Prüfung: 5. Mai 2017

Referent: Priv. Doz. Dr. habil. Matthias Mozer  
Institut für Experimentelle Kernphysik

Korreferent: Prof. Dr. G. Quast  
Institut für Experimentelle Kernphysik



This document is licensed under a Creative Commons Attribution –  
Non Commercial – No Derivatives 4.0 International License (CC BY-NC-ND 4.0):  
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>

# Introduction

*... maybe the greatest triumph of the human genius is that a man can understand those things which he can no longer imagine.*

---

– Lev Landau

By the end of the 19th century physical laws seemed to be well established. In 1894 A. Michelson said: "It seems that most of the grand underlying principles have been firmly established ... An eminent physicist has remarked that the future truths of the Physical Society are to be looked for in the sixth place of decimals" [1]. However, this turned out to be not true and the 20th century has actually brought a scientific breakthrough by building the quantum theory which lies in the basis of our understanding of nature at the moment. Quantum theory was able to overcome problems faced by the classical mechanics when applicable to microscopic objects and has become the main theoretical tool of describing microscopic systems<sup>1</sup> and elementary particles in particular.

Modern experimental particle physics arises from experiments carried out by Pierre and Marie Curie, Rutherford and many others in the beginning of the 20th century which brought the concept of radioactivity and the atomic nuclei. This was followed up by many important discoveries: cosmic rays, observation of a positron in a Wilson cloud chamber by Anderson in 1932, observation of the neutrino (proposed by Pauli in 1933 and detected in the Cowan–Reines neutrino experiment in 1956 [2]), observation of the neutral weak currents in 1973 by the Gargamelle experiment [3], observation of vector bosons at CERN in 1983 and many others. The discovery of the Higgs boson [4, 5] in 2012 has confirmed the existence of the last missing piece of the theory known as *the Standard Model*. The Standard Model is the quantum field theory that describes elementary particles that are known at the moment and their interaction via electroweak and strong forces at distances up to  $10^{-19}$  meters.

The Standard Model is truly a successful theory confirmed by many experiments. However, it is known that it is not a complete theory. Many fundamental questions are still open: why there are only three generations of particles? Why there is a symmetry between quarks and leptons? And why, if this symmetry exists, leptons and quarks have different charges and masses? Furthermore, gravity, the fourth fundamental force, is not part of the Standard Model. Moreover, the Planck scale ( $\sim 10^{19}$  GeV [6]) and electroweak scale ( $\sim 100$  GeV) differ by 17 orders of magnitude (the so called *hierarchy problem*). Also, how is the Higgs boson mass stabilized? In the Standard Model radiation corrections should strongly increase the Higgs boson mass [7], however, the Higgs boson mass is found to be around 125 GeV [8] (the so called *naturalness*

---

<sup>1</sup>to be precise: systems with action  $\sim \hbar$ .

---

*problem* of the Standard Model). Finally, astrophysical observations show that baryonic matter composes only a small fraction of the Universe. At the moment there is much experimental evidence for this: the mass of galaxy clusters exceeds the mass expected from their luminosity, which was first measured in [9], galaxy rotation curves indicate the presence of a large dark halo [10], effects from gravitational lensing also show the presence of invisible mass. According to recent astrophysical measurements baryon matter composes only 5 % of the visible Universe while 26% correspond to dark matter and 69% to dark energy [11].

Some of these issues are attempted to be addressed by Beyond the Standard Model theories (BSM). Supersymmetry [12], for example, introduces a symmetry between bosons and fermions and predicts that each particle of the Standard Model has a superpartner some of which could be dark matter candidates, for example. Another popular direction is theories with extra dimensions. This includes models like Arkani-Hamed-Dimopoulos-Dvali (ADD) model [13], Randall-Sundrum (RS) Warped Extra Dimensions model [14, 15]. There are also models based on the expansion of the gauge group of the Standard Model which leads to the existence of new gauge bosons  $Z'$  and  $W'$  [16].

Most of these theories predict that new particles should show up at energies of about 1 TeV. In 2015 the Large Hadron Collider (LHC) started the Run II by colliding protons with a center-of-mass energy of 13 TeV giving a lot of opportunities for experiments to look for effects from new physics. However, it might be that the new particle(s) is too heavy to be seen by the LHC experiments. In this case its effects on the predicted cross-section can be parametrized in terms of the effective field theory approach (EFT). EFTs add higher dimensional operators to the Standard Model Lagrangian and thus introduce additional couplings to the Standard Model. One of the important assumptions of this approach is that these higher dimensional terms should be suppressed as long as the scale of the new physics is much larger than the scale of the process. This kind of approach was used by Fermi when he studied the  $\beta$ -decay. The weak interaction of fermions was approximated with 4 fermion vertex characterized by the coupling strength  $G_F^2$  neglecting the exchange of W bosons.

Diboson final states are some of the final states where new physics can be expected. Previous accelerators (LEP, Tevatron) were quite limited in this final state because of the rather small cross-section. There are searches for the heavy resonances decaying into pairs of W and Z bosons performed at the LHC [17–19]. In terms of the EFT approach the possible effects from new physics in the diboson sector can be parametrized with anomalous triple and quartic gauge couplings.

This thesis presents the search for anomalous triple gauge couplings with data recorded by the Compact Muon Solenoid (CMS) experiment from proton-proton collisions at a center-of-mass energy of 13 TeV. The analysis is performed in the semileptonic channel with one W boson decaying to a charged lepton and a neutrino and another W boson or Z boson decaying to hadrons. Since the effects from anomalous triple gauge couplings are prominent at high W or Z boson momenta [20] the so-called boosted topology is exploited which means that decay products of W or Z bosons are not reconstructed as two jets coming from two quarks but as a single jet with large radius. In order to identify these boosted jets state of the art jet substructure techniques are used which significantly reduce fake rate from QCD jets for this analysis.

---

<sup>2</sup> $G_F \approx 10^{-5} \text{ GeV}^{-2}$

---

One of the main backgrounds in this analysis is  $W$ +jets production which is rather complicated to model with simulation. Therefore, the  $W$ +jets contribution is estimated using sophisticated data-driven technique which corrects its normalization and the shape in the signal region. Another important background is top quark pair production ( $t\bar{t}$ ). The shape of this contribution in the sensitive variable is taken from the simulation. Modelling of  $t\bar{t}$  background is validated in two control regions and is shown to be trustworthy. To model effects from anomalous triple gauge couplings the signal function is constructed which takes into account effects of interference between different anomalous triple gauge couplings as well as interference between the Standard Model contribution and a given anomalous triple gauge coupling. It is shown further that the signal function models well the simulated prediction in the range of sensitivity of the analysis. Limits on anomalous triple gauge couplings are extracted from the diboson mass distribution using a delta log-likelihood method.

This thesis is structured in the following way:

Chapter 1 gives the theoretical introduction. Basic aspects of the theory of the electroweak and strong interactions are discussed. Also the effective field theory approach is introduced and is compared with alternative methods. In the end an overview of the current experimental results is given.

In Chapter 2 an overview of the LHC accelerator complex is provided. After that the CMS detector and its subsystems are described.

Chapter 3 starts with an overview of the basic steps of event simulation. This is followed by the discussion of the reconstruction and identification of objects important in the context of this thesis.

In Chapter 4 the study of grooming algorithms that was done in the context of preparation towards Run II of the LHC is presented. The study was focused on the pileup mitigation effects and jet mass resolution.

Chapter 5 discusses the modelling of signal and background processes, event selection and corrections applied to the simulation. In the end the description of considered systematic uncertainties is given.

In Chapter 6 modelling of the effects from anomalous triple gauge coupling is discussed.

Chapter 7 is dedicated to the background estimation. Two backgrounds are important in the analysis:  $W$ +jets and top quark pair production. The first one is estimated from the sideband data while  $t\bar{t}$  background is taken from the simulation.

Finally, in Chapter 8 the statistical model is described and limits on anomalous triple gauge couplings are presented.



# Contents

<b>1</b>	<b>Theory introduction and experimental overview</b>	<b>1</b>
1.1	The Standard Model . . . . .	1
1.1.1	Quantum chromodynamics . . . . .	3
1.1.2	Electroweak theory . . . . .	5
1.2	Effective field theory . . . . .	8
1.3	Experimental overview . . . . .	9
<b>2</b>	<b>The CMS Experiment at the LHC</b>	<b>13</b>
2.1	The Large Hadron Collider . . . . .	13
2.2	The CMS experiment . . . . .	17
2.2.1	Superconducting magnet . . . . .	18
2.2.2	Inner tracking system . . . . .	18
2.2.3	Electromagnetic calorimeter . . . . .	20
2.2.4	Hadron calorimeter . . . . .	21
2.2.5	Muon system . . . . .	23
2.2.6	Trigger and data acquisition . . . . .	24
2.2.7	Luminosity . . . . .	25
<b>3</b>	<b>Event simulation and reconstruction</b>	<b>27</b>
3.1	Event simulation . . . . .	27
3.1.1	Event generators . . . . .	30
3.2	Event reconstruction . . . . .	31
3.2.1	Track reconstruction . . . . .	31
3.2.2	Primary-vertex and beam spot reconstruction . . . . .	32
3.2.3	Muon reconstruction and identification . . . . .	33
3.2.4	Electron reconstruction and identification . . . . .	34
3.2.5	Jet reconstruction . . . . .	37
3.2.6	Identification of b-jets . . . . .	39
3.2.7	Reconstruction of missing transverse energy . . . . .	39
<b>4</b>	<b>Grooming algorithms</b>	<b>41</b>
4.1	Analysis strategy . . . . .	41
4.1.1	Grooming algorithms . . . . .	42
4.2	Mass distributions . . . . .	43

4.3	Mass response distributions . . . . .	44
4.4	Pileup dependence . . . . .	46
4.5	N-subjettiness . . . . .	47
4.6	Conclusion . . . . .	48
<b>5</b>	<b>Process modelling and event selection</b>	<b>51</b>
5.1	Signal and background modelling . . . . .	51
5.1.1	Signal process . . . . .	51
5.1.2	Background processes . . . . .	52
5.1.3	Simulated samples . . . . .	54
5.1.4	Data samples . . . . .	55
5.2	Event selection . . . . .	56
5.2.1	Triggers . . . . .	56
5.2.2	Noise filters . . . . .	56
5.2.3	Electron selection . . . . .	56
5.2.4	Muon selection . . . . .	56
5.2.5	Jet selection . . . . .	56
5.2.6	Missing $E_T$ selection . . . . .	57
5.2.7	Final event selection . . . . .	57
5.3	Simulation corrections . . . . .	58
5.3.1	Normalization . . . . .	58
5.3.2	Pileup . . . . .	61
5.3.3	B-tagging . . . . .	61
5.3.4	Lepton identification efficiencies . . . . .	62
5.3.5	Trigger efficiencies . . . . .	62
5.3.6	Vector boson identification efficiencies . . . . .	64
5.4	Systematic uncertainties . . . . .	64
<b>6</b>	<b>Signal modelling</b>	<b>67</b>
6.1	Signal region . . . . .	67
6.2	Signal function . . . . .	67
6.2.1	Normalization . . . . .	69
6.2.2	Shape . . . . .	69
6.2.3	Shape uncertainties . . . . .	72
<b>7</b>	<b>Background estimation</b>	<b>75</b>
7.1	Background normalization . . . . .	75
7.2	Validation of background in control regions . . . . .	79
7.2.1	W+jets control region . . . . .	79
7.2.2	$t\bar{t}$ control regions . . . . .	79
7.3	Background shapes . . . . .	82
<b>8</b>	<b>Limits on anomalous couplings</b>	<b>91</b>
8.1	Methodology . . . . .	91

8.2	Statistical model . . . . .	92
8.2.1	Systematic uncertainties . . . . .	93
8.3	Results . . . . .	94
<b>Conclusions and outlook</b>		99
<b>Appendices</b>		103
<b>A</b>	<b>Mass distributions for grooming algorithms</b>	105
<b>B</b>	<b>Mass response distributions for grooming algorithms</b>	109
<b>C</b>	<b>Pileup dependence plots</b>	113
<b>D</b>	<b>Additional plots for b-tagging efficiencies</b>	121
<b>E</b>	<b>Additional plots for the signal model</b>	125
E.1	Normalization of the signal . . . . .	125
E.2	Signal function . . . . .	126
E.2.1	Additional plots for signal modelling . . . . .	127
E.3	Shape uncertainties of the signal function . . . . .	128
E.3.1	Effect of systematic uncertainties on the signal function . . . . .	129
<b>F</b>	<b>Bias test for <math>M_{\text{pruned}}</math> fit</b>	131
<b>G</b>	<b>Shape of backgrounds in the sideband</b>	133
<b>H</b>	<b>Additional plots for control regions</b>	135
H.1	$W$ +jets control region . . . . .	135
H.2	$t\bar{t}$ control region . . . . .	137
H.3	$t\bar{t}$ control region II . . . . .	139
H.4	signal region . . . . .	141
H.5	QCD fakes in the electron channel . . . . .	144
<b>I</b>	<b>Additional plots for 1-dimensional limits</b>	145
<b>List of Figures</b>		147
<b>List of Tables</b>		153
<b>Bibliography</b>		155



# 1 Theory introduction and experimental overview

This chapter gives a general introduction to the theoretical aspects discussed in this thesis. A brief overview of the theory of strong and electroweak interactions is given. In the next section the effective field approach is introduced with the focus on the diboson sector. Finally an overview of existing experimental results is given.

## 1.1 The Standard Model

Already Democritus, the Ancient Greek philosopher, developed the atomic hypothesis that everything is composed of "atoms" – indivisible and indestructible objects (in Ancient Greek *ατομος* means "indivisible"). Now we know that matter indeed consists of atoms. However, it is known that atoms consist of electrons and nuclei. Furthermore, a nucleus consists of protons and neutrons, and neutrons and protons are composed of quarks. And quarks are considered to be fundamental particles which means that they are not composite objects. Fundamental particles are divided into two classes: fundamental fermions and fundamental bosons. Fundamental fermions are spin  $1/2^1$  particles. They are listed in Table 1.1. Fundamental fermions are matter

**Table 1.1:** List of fermions in the Standard Model. Masses are quoted from [21, 22]. Electrical charge is quoted in units of electron charge.

		Symbol	Mass, MeV	Electrical charge
leptons	electron	$e^-$	0.51	$- e $
	muon	$\mu^-$	105.65	$- e $
	tau	$\tau^-$	1776.82	$- e $
	electron neutrino	$\nu_e$	$< 2.0 \cdot 10^{-6}$	0
	muon neutrino	$\nu_\mu$	$< 0.19$	0
	tau neutrino	$\nu_\tau$	$< 18.2$	0
quarks	up quark	u	2.3	$+2/3 e $
	down quark	d	4.8	$-1/3 e $
	charm quark	c	1275.0	$+2/3 e $
	strange quark	s	95.0	$-1/3 e $
	b quark	b	4.18	$-1/3 e $
	top quark	t	$173.34 \cdot 10^3$	$+2/3 e $

constituents. Fundamental bosons have integer spin and carry forces between fundamental

---

<sup>1</sup>here and further  $\hbar = 1$  and  $c = 1$ .

fermions. At the moment we know four fundamental forces: electromagnetic (mediated by photons), weak (mediated by W and Z bosons), strong (mediated by gluons) and gravitational force. The Higgs boson plays a special role of giving masses to elementary particles. The Standard model doesn't include gravitational force which is much weaker than any of three. There are attempts (for example [23]) to unify the Standard Model with the general theory of relativity but they are beyond the scope of this thesis.

W and Z bosons play a central role in this thesis. They were discovered by UA1 and UA2 experiments at CERN in 1983. Both bosons are heavy: W has a mass of  $80.376 \pm 0.033$  GeV and Z has a mass of  $91.1876 \pm 0.0021$  GeV [21] and have rather short lifetime (or large width) - W boson has a width of  $2.085 \pm 0.042$  GeV and Z boson  $2.4952 \pm 0.0023$  GeV [21]. This means that it is not possible to detect them directly but only their decay products. Branching fractions of W and Z bosons are given in the Table 1.2 and 1.3. For Z boson only dominant branching ratio are listed, other decay modes are found to have branching ratio  $< 10^{-5}$  [21].

**Table 1.2:** Branching ratios of W boson [21].

	Branching ratio, %
$\Gamma(e^+\nu)/\Gamma_{\text{total}}$	$10.71 \pm 0.16$
$\Gamma(\mu^+\nu)/\Gamma_{\text{total}}$	$10.63 \pm 0.15$
$\Gamma(\tau^+\nu)/\Gamma_{\text{total}}$	$11.38 \pm 0.21$
$\Gamma(\text{hadrons})/\Gamma_{\text{total}}$	$67.41 \pm 0.27$

**Table 1.3:** Branching ratios of Z boson [21].

	Branching ratio, %
$\Gamma(e^+e^-)/\Gamma_{\text{total}}$	$3.3632 \pm 0.0042$
$\Gamma(\mu^+\mu^-)/\Gamma_{\text{total}}$	$3.3662 \pm 0.0066$
$\Gamma(\tau^+\tau^-)/\Gamma_{\text{total}}$	$3.3696 \pm 0.0083$
$\Gamma(\text{hadrons})/\Gamma_{\text{total}}$	$69.911 \pm 0.056$
$\Gamma(\text{invisible})/\Gamma_{\text{total}}$	$20.000 \pm 0.055$

The Standard Model is a quantum field theory that describes the electromagnetic, weak and strong interactions of elementary particles. The basic principle of a field theory is the *principle of stationary action*. The action is given by the space-time integral of the Lagrangian density <sup>2</sup>

$$S = \int \mathcal{L}(\psi, \partial_\mu \psi) d^4x \quad (1.1)$$

And the *principle of stationary action* means that the evolution of the system from the time moment  $t_1$  to the time moment  $t_2$  happens so that the path in the configuration space delivers minimum of the action so that:

$$\begin{aligned} \delta S &= 0 \\ \delta^2 S &> 0 \end{aligned} \quad (1.2)$$

---

<sup>2</sup>further the "density" would be skipped and the "Lagrangian" would actually mean the Lagrangian density

Basic components of the Standard Model are fields and excitations of these fields are interpreted as elementary particles [24]. One of the basic principles of the Standard Model is the principle of local gauge invariance. This means that if there is a field  $\Psi(x)$  then the theory should be invariant under transformation of the type:

$$\Psi(x) \rightarrow \Psi' = e^{i\alpha(x)} \cdot \Psi(x) \quad (1.3)$$

where  $\alpha(x)$  is an arbitrary function of space-time. Another important requirement is that all fields in the Standard Model should be invariant under Lorentz transformations<sup>3</sup>. Fields in the Standard Model are representation of the group  $U(1) \otimes SU(2) \otimes SU(3)$ . This corresponds to the following transformation:

$$\Psi(x) \rightarrow \Psi' = e^{i\beta} e^{i(\alpha^a \sigma^a / 2)} e^{i(\epsilon^b \lambda^b / 2)} \cdot \Psi(x) \quad (1.4)$$

where  $\sigma^a$  are Pauli matrices ( $a = \overline{1,3}$ ),  $\lambda^b$  are Gell-Mann matrices ( $b = \overline{1,8}$ );  $\beta, \alpha^a, \epsilon^b$  are real-valued functions of the space-time.  $U(1) \otimes SU(2)$  group corresponds to the electroweak interaction and  $SU(3)$  corresponds to the strong interaction.

### 1.1.1 Quantum chromodynamics

All fermions in the Standard Model are characterized by the electrical charge (see Table 1.1). Quarks carry an additional degree of freedom – *color charge* (three different states, usually referred as *red, green and blue*, corresponding antiquarks have *anticolor*). The color charge is carried by gluons. Interactions of quarks and gluons are described by quantum chromodynamics (QCD). The Lagrangian of QCD is given by the following formula:

$$\mathcal{L}_{QCD} = \bar{\Psi}_i (i(\gamma^\mu D_\mu)_{ij} - m\delta_{ij}) \Psi_j - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu} \quad (1.5)$$

where  $\Psi_i(x)$  ( $i = \overline{1,3}$ ) is a quark field,  $A_\mu^a(x)$  ( $a = \overline{1,8}$ ) is a gluon field,  $\gamma_\mu$  are Dirac matrices [24],  $D_\mu$  is the covariant derivative:

$$D_\mu = \partial_\mu - ig_s A_\mu^a \frac{\lambda^a}{2} \quad (1.6)$$

and  $G_{\mu\nu}^a$  is a gluon field strength tensor given by

$$G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g_s f^{abc} A_\mu^b A_\nu^c \quad (1.7)$$

where  $f^{abc}$  are the structure constants of group  $SU(3)$  and  $g_s$  is the coupling constant of the QCD. The structure of the QCD Lagrangian leads to following vertices being allowed: quark-antiquark-gluon ( $q\bar{q}g$ ) vertex, a 3-gluon vertex (both proportional to  $g_s$ ) and a four-gluon vertex (proportional to  $g_s^2$ ) [21].

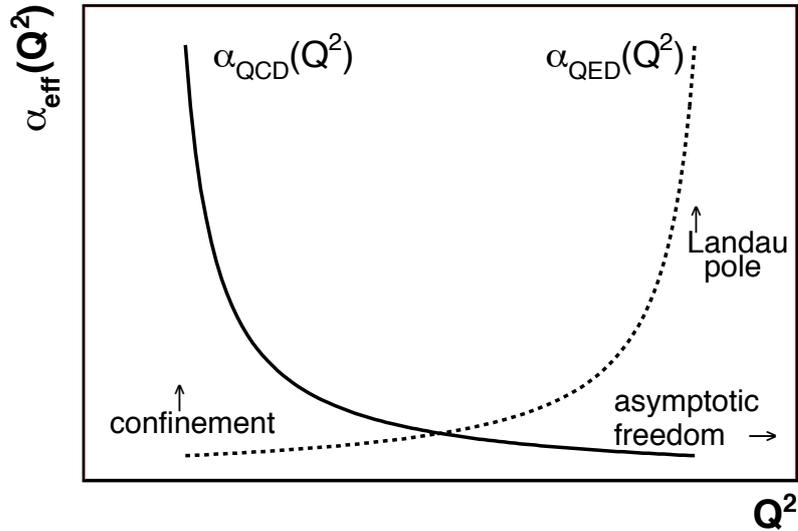
The strong interaction has two properties: *confinement* and *asymptotic freedom*. Both of them are consequences of non-abelian structure of the QCD. Confinement means that one cannot observe a free quark and that quarks exist only inside hadrons. The potential of two quark

<sup>3</sup>actually more general: rotations, boosts and translations described by the Poincaré group

interaction increases with distance between them so that at some point there is enough energy to create a  $q\bar{q}$  pair. Because of that one cannot find a quark experimentally but only a *jet* of hadrons. As shown in [25] the running constant of the QCD is given by:

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f) \ln(Q^2/\Lambda^2)} \quad (1.8)$$

where  $Q^2 = -q^2$  and  $q$  is the momentum transfer in a two color-charged particle interaction and  $n_f$  is number of excited quark flavours (which depends on the  $Q^2$ ). The fact that we have only six quark flavours becomes crucial here and it leads to the fact the strong coupling decreases at high momentum transfer which reflect asymptotic freedom – the higher the energy is the weaker the strong interaction becomes. Due to to that QCD is a perturbative at high energies. This means that perturbation series is actually converging and processes with QCD interaction can be calculated (and simulated) at a certain order (leading order (LO) or next-to-leading order usually). The opposite effect is the confinement – the lower the energy is (or equivalently the large the distance between two color-charged particles) the stronger the interaction becomes. This means that around the scale  $\Lambda^4$  QCD theory becomes non-perturbative. This behaviour of the coupling constant is caused by the gluon self-interaction as shown in [25] and is opposite to the electromagnetic coupling (described by quantum electrodynamics (QED)) which illustrated in Figure 1.1.



**Figure 1.1:** Comparison of coupling constants of the QCD and QED interaction. QCD coupling diverges at low energies (color confinement) and decreases at low energies (asymptotic freedom). QED coupling is small at low energies and diverges at high energies (Landau pole) [27].

<sup>4</sup> $\Lambda$  depends on  $n_f$  actually, for example for  $n_f = 4$ ,  $\Lambda \approx 200$  MeV [26].

### 1.1.2 Electroweak theory

As it was already mentioned fundamental fermions are spin 1/2 particles. They are described by the Dirac equation:

$$(i\gamma^\mu \partial_\mu - m)\psi(x) = 0 \quad (1.9)$$

where  $\psi(x)$  is the bispinor<sup>5</sup> :

$$\psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix} \quad (1.10)$$

As it was already mentioned before that electroweak interaction corresponds to  $U(1) \otimes SU(2)$  symmetry. Dirac bispinor can be split into two components:

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \quad (1.11)$$

where  $\psi_L$  and  $\psi_R$  are left-handed component and right-handed component accordingly defined as  $\frac{1}{2}(1 \mp \gamma^5)\psi$ ,  $\gamma^5 = \frac{-i}{4!}\epsilon^{\mu\nu\rho\sigma}\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma$ . Experiments show that the left-handed and right handed component of fermions behave differently with respect to  $SU(2)$  symmetry – right-handed components behave as singlets and left-handed components as doublets [29]. This are fermions of the Standard Model with respect to  $SU(2)$  symmetry:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \otimes e_R, \quad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \otimes \mu_R, \quad \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \otimes \tau_R, \quad (1.12)$$

$$\begin{pmatrix} u \\ d \end{pmatrix}_L \otimes u_R \otimes d_R, \quad \begin{pmatrix} c \\ s \end{pmatrix}_L \otimes c_R \otimes s_R, \quad \begin{pmatrix} t \\ b \end{pmatrix}_L \otimes t_R \otimes b_R, \quad (1.13)$$

In the Standard Model there are only left-handed neutrinos and right-handed antineutrinos. Neutrinos are massless in the Standard Model. However, observation of neutrino oscillations confirms that neutrinos are massive [30]<sup>6</sup>. The electroweak Lagrangian can be written by the following formula:

$$\mathcal{L}_{ewk} = \mathcal{L}_{Dirac} + \mathcal{L}_{Yukawa} + \mathcal{L}_{Higgs} \quad (1.14)$$

where  $\mathcal{L}_{Dirac}$  is given by :

$$\mathcal{L}_{Dirac} = \bar{L}i\gamma^\mu D_\mu^L L + \bar{R}i\gamma^\mu D_\mu^R R \quad (1.15)$$

where

$$\begin{aligned} D_\mu^L &= \partial_\mu - i\frac{g_W}{2}\sigma^a W_\mu^a - ig_e \frac{Y}{2} B_\mu \\ D_\mu^R &= \partial_\mu - ig_e \frac{Y}{2} B_\mu \end{aligned} \quad (1.16)$$

<sup>5</sup>strictly speaking bispinor is an irreducible representation of the Lorentz group [28].

<sup>6</sup>which means that there should be a right-handed neutrino

Since the gauge transformation is :

$$A_\mu(x) \rightarrow A'_\mu(x) = \omega A_\mu(x) \omega^{-1} + \omega \partial_\mu \omega^{-1} \quad (1.17)$$

if the gauge transformation of the matter field is  $\psi(x) \rightarrow \psi'(x) = \omega \psi(x)$  [31]. Then the mass term  $\frac{1}{2} m_A^2 A_\mu A^\mu$  cannot be introduced because it is not gauge invariant. But there should be a mechanism that gives masses to W and Z bosons because we know that they are massive [21, 32, 33]. In the Standard Model the mechanism of *spontaneous symmetry breaking* [34, 35] takes care of that. A complex scalar doublet with respect to  $SU(2)$  is introduced:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi_0 \end{pmatrix} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} -i(\phi_1 - i\phi_2) \\ v + (h + i\phi_3) \end{pmatrix} \quad (1.18)$$

where  $h(x)$  is the Higgs field and  $v$  is the minimum of the  $V(\phi)$  potential:  $v = \pm \sqrt{\frac{-\mu^2}{\lambda}}$ <sup>7</sup>. And introducing the Higgs Lagrangian:

$$\mathcal{L}_{Higgs} = D_\mu \phi D^\mu \phi - V(\phi) - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} Tr(W_{\mu\nu}^a W_{\mu\nu}^a) \quad (1.19)$$

where  $V(\phi) = -\frac{1}{2} \mu^2 \phi^2 + \frac{1}{4} \lambda \phi^4$  is the so called "Mexican hat" potential shown in Fig. 1.2,  $B_{\mu\nu}$  and  $W_{\mu\nu}^a$  are field strength tensors defined as following:

$$\begin{aligned} B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \\ W_{\mu\nu}^a &= \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g_w \epsilon^{abc} W_\mu^b W_\nu^c \end{aligned} \quad (1.20)$$

where  $a = \overline{1, 3}$  and  $\epsilon^{abc}$  is Levi-Civita symbol.

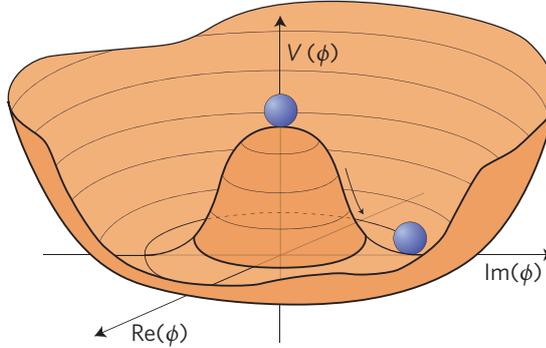


Figure 1.2: "Mexican hat" potential [36].

It was shown in [31] that the kinematic term  $W_{\mu\nu}^a W_{\mu\nu}^a$  can be rewritten as :

$$W_{\mu\nu}^a W_{\mu\nu}^a = \left( \mathcal{A}_{\mu\nu} \sin \theta_W + \mathcal{Z}_{\mu\nu} \cos \theta_W + ig_e (W_\mu^- W_\nu^+ - W_\mu^+ W_\nu^-) \right)^2 \quad (1.21)$$

<sup>7</sup>  $v = \frac{2m_W}{g_w} \approx 246$  GeV.

where  $W_{\mu\nu}^{\pm} = \frac{W_{\mu\nu}^1 \mp iW_{\mu\nu}^2}{\sqrt{2}}$ ,  $\mathcal{Z}_{\mu\nu} = \partial_{\mu}Z_{\nu} - \partial_{\nu}Z_{\mu}$ ,  $\mathcal{A}_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$ . Z boson and a photon are defined as a mixture of different components of gauge fields (see Eq. (1.22)). It is important to notice that Eq. (1.21) actually shows that triple gauge couplings for WWZ and WW $\gamma$  (that are of particular interest of this thesis) vertices arise.

It can be shown [31] that gauge can be chosen so that  $\phi(x)$  can be presented as  $\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}$ . This naturally gives masses to W and Z bosons and defines photon, W and Z bosons as mixed components of gauge fields:

$$\begin{aligned} W_{\mu}^{\pm} &= \frac{1}{\sqrt{2}}(W_{\mu}^1 \mp iW_{\mu}^2) \\ Z_{\mu}^0 &= \cos \theta_W W_{\mu}^3 - \sin \theta_W B_{\mu} \\ A_{\mu} &= \sin \theta_W W_{\mu}^3 + \cos \theta_W B_{\mu} \end{aligned} \quad (1.22)$$

$\theta_W$  is the electroweak mixing angle:  $\theta_W = \arctan(\frac{g_e}{g_w})$ .

Relation between electrical charge  $Q$  and the generator of  $U(1)$  group – Y (*hypercharge*) is defined as :

$$Q = t^3 + \frac{1}{2}Y \quad (1.23)$$

where  $t^3$  is isospin projector on z-axis ( $t^3 = \frac{\sigma^3}{2}$ ). In the Table 1.4 hypercharges for leptons and quarks in the Standard Model are listed.

**Table 1.4:** Hypercharge of leptons and quarks in the Standard Model.

	particles	electrical charge Q	weak isospin, $t^3$	weak hypercharge
leptons	$\nu_{e,L}^-, \nu_{\mu,L}^-, \nu_{\tau,L}^-$	0	$\frac{1}{2}$	-1
	$e_L^-, \mu_L^-, \tau_L^-$	-1	-1/2	-1
	$e_R^-, \mu_R^-, \tau_R^-$	-1	0	-2
quarks	$u_L, c_L, t_L$	+2/3	+1/2	+1/3
	$u_R, c_R, t_R$	+2/3	0	+4/3
	$d_L, s_L, b_L$	-1/3	-1/2	+1/3
	$d_R, s_R, b_R$	-1/3	0	-2/3

All leptons except neutrinos are massive in the Standard Model (see Table 1.1). Dirac mass term  $m\bar{\psi}\psi$  from Eq. (1.9) mixes left- and right-handed components  $m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)$  and thus is not invariant under  $SU(2)$  transformation. To solve this problem the so called *Yukawa term* is introduced in the Standard Model Lagrangian:

$$\mathcal{L}_{Yukawa} = \sum_{\text{lepton flavours}} (g_L(\psi_L\phi\psi_R + h.c.) + \tilde{\phi}\bar{\psi}_R\psi_L + h.c.) \quad (1.24)$$

where  $\tilde{\phi} \equiv i\sigma^2\phi^*$  and  $g_L = \sqrt{2}\frac{m_L}{v}$  [37].

## 1.2 Effective field theory

The Higgs boson discovered in 2012 [4, 5] was the last missing piece of the Standard Model. However, there are several problems that are not explained by the Standard Model (like hierarchy problem, dark matter and dark energy, matter-antimatter asymmetry etc.). Many of the BSM theories predict new resonances (for example, [14–16]). However, we have not observed new resonances so far (see for example [38]) and it can be that new interactions with the scale  $\Lambda$  can manifest themselves at energies below  $\Lambda$  [39]. This brings the idea of the expansion of the Standard Model Lagrangian:

$$\mathcal{L}_{eff} = \mathcal{L}_0 + \frac{1}{\Lambda} \mathcal{L}_1 + \frac{1}{\Lambda^2} \mathcal{L}_2 + \dots \quad (1.25)$$

which is relevant only in case  $E \ll \Lambda$  where  $E$  is the energy of a collider. The term  $\mathcal{L}_0$  corresponds to the Standard Model Lagrangian, the term  $\mathcal{L}_1$  violates lepton number and is not relevant till energies  $\sim 10^{13}$  GeV [39]. While [39] discusses all possible dimension six<sup>8</sup> operators that appear in  $\mathcal{L}_2$  this thesis focuses on operators relevant for the diboson sector. 5 dimension six operators relevant for the diboson sector are introduced in [40]. In this thesis only 3 CP-conserving operators are studied:

$$\begin{aligned} \mathcal{O}_{WWW} &= \text{Tr}[W_{\mu\nu} W^{\nu\rho} W_{\rho}^{\mu}] \\ \mathcal{O}_W &= (D_{\mu}\Phi)^{\dagger} W^{\mu\nu} (D_{\nu}\Phi) \\ \mathcal{O}_B &= (D_{\mu}\Phi)^{\dagger} B^{\mu\nu} (D_{\nu}\Phi) \end{aligned} \quad (1.26)$$

Each of these operators introduces additional coupling and thus the full Lagrangian can be given by the following formula:

$$\mathcal{L}_{eff} = \mathcal{L}_0 + \frac{c_{WWW}}{\Lambda^2} \mathcal{O}_{WWW} + \frac{c_W}{\Lambda^2} \mathcal{O}_W + \frac{c_B}{\Lambda^2} \mathcal{O}_B \quad (1.27)$$

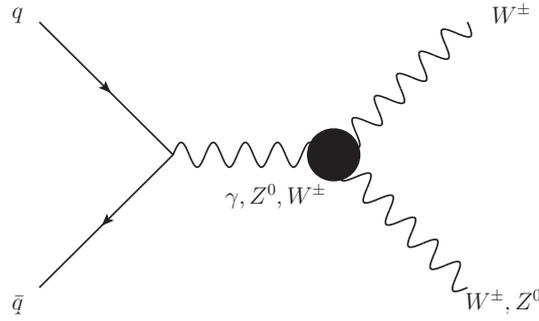
This Lagrangian modifies the WWZ and WW $\gamma$  vertex as shown on the Fig.1.3.

These couplings are usually referred to as charged *anomalous triple gauge couplings* (aTGC). Triple gauge couplings are allowed in the Standard Model. This comes from the kinetic term of W and Z bosons (see Eq. (1.20)).

Dimension six operators yield terms proportional to  $s/\Lambda^2$ . Thus, anomalous triple gauge couplings lead to cross section growing with energy and thus a deviation from the Standard Model (if any) will appear at high diboson masses. However, quantum theory gives a limitation to that because of unitarity. Unitarity means that the sum of probabilities of possible outcomes of a quantum system is unity. Thus at the certain scale such dependence will lead to the violation of the unitarity bound (unitarity bounds for aTGC were studied here [20], for example). The advantage of the EFT approach is that it doesn't violate the unitarity bound by construction [40]<sup>9</sup>:

<sup>8</sup>what is meant here by the dimension is the power of the actual dimension of  $\mathcal{L}_2$  in GeV.  $[S] = 1$  (see Eq. (1.1)), therefore  $[\mathcal{L}_0] = \text{GeV}^4$  and because  $[\Lambda] = \text{GeV}$ , then  $[\mathcal{L}_2] = \text{GeV}^6$ . Thus we say that  $\mathcal{L}_2$  is a dimension six operator.

<sup>9</sup>that's true only when the actual scale of the process is known. For example, in case of WW in the fully leptonic channel the scale of the process is unknown and thus it is harder to have a control of the EFT validity condition.



**Figure 1.3:** The diagram illustrating WW or WZ with anomalous triple gauge coupling affecting WW $\gamma$  and WWZ vertex.

if  $s/\Lambda^2$  becomes of the order of  $\sim 1$  this means that other terms in Eq. (1.25) should be taken into account and the EFT approach is not valid in this case. Other approaches (Lagrangian and vertex function approaches) solve this problem by introducing arbitrary chosen form-factors [20] that suppress non-physical contribution at high energies and thus respect the unitarity at an arbitrary scale. These approaches don't introduce any scale of validity of the theory which is not the case for the EFT. Also, introducing form-factors means that the parameters of the Lagrangian become functions of momenta while the Lagrangian is written in the position space. There are more arguments in favour of the EFT given here [40]. This thesis uses the EFT approach and introduces an upper cut on the scale of the process.

### 1.3 Experimental overview

Anomalous triple gauge couplings were studied since LEP II [41]. LEP II was using the effective Lagrangian approach:

$$\mathcal{L}_{eff} = g_{WWV} \left( g_1^V V^\mu (W_{\mu\nu}^- W^{+\nu} - W_{\mu\nu}^+ W^{-\nu}) + \kappa_V W_\mu^+ W_\nu^- V^{\mu\nu} + \frac{\lambda_V}{m_W^2} V^{\mu\nu} W_\nu^{+\rho} W_{\rho\mu}^- \right) \quad (1.28)$$

where  $V \equiv \gamma$  or  $Z$ . In the Standard Model  $g_1^Z = g_1^Y = \kappa_Z = \kappa_\gamma = 1$  and  $\lambda_Z = \lambda_\gamma = 0$ . Only three of these parameters are independent. Results are usually quoted in terms of  $g_1^Z$ ,  $\kappa_\gamma$ ,  $\lambda_\gamma$  or  $\Delta g_1^Z = g_1^Z - 1$ ,  $\Delta\kappa_Z = \kappa_Z - 1$ ,  $\lambda_Z$  because of the following relation:

$$\begin{aligned} \Delta\kappa_Z &= \Delta g_1^Z - \Delta\kappa_\gamma \tan^2 \theta_W \\ \lambda_Z &= \lambda_\gamma \end{aligned} \quad (1.29)$$

and  $g_1^Y$  is fixed by the electromagnetic gauge invariance [42]. The analyses of the charged TGC were performed by ALEPH, DELPHI, L3 and OPAL experiments [43]. Analyses used different channels, typically semileptonic and fully leptonic. Results of the 4 experiments combination are shown in Table 1.5.

D0 also performed a search for aTGC [44]. Analyses were done in the lepton + 2 jets channel for WW and WZ production and leptonic final states for W $\gamma$ , WW, WZ production. Results

**Table 1.5:** Results of LEP-II combination [43].

Parameter	68% C.L.	95% C.L.	SM
$g_1^Z$	$+0.984^{+0.018}_{-0.020}$	[0.946, 1.021]	1
$\kappa_\gamma$	$+0.982^{+0.042}_{-0.042}$	[0.901, 1.066]	1
$\lambda_\gamma$	$0.022^{+0.019}_{-0.019}$	[-0.059, 0.017]	0

**Table 1.6:** Limits on aTGC by D0 [44].

Parameter	68% C.L.	95% C.L.	SM
$\Delta g_1^Z$	[-0.008, 0.054]	[-0.034, 0.084]	0
$\Delta \kappa_\gamma$	[-0.057, 0.154]	[-0.158, 0.255]	0
$\lambda_\gamma$	[-0.015, 0.028]	[-0.036, 0.044]	0

are shown in Table 1.6. Search for aTGC is also performed in the LHC by CMS [45–49] and ATLAS [50–53]. Most of the analyses use the fully leptonic channel, and some of the analyses use the semileptonic channel (with one lepton and two jets).

Advantages of the semileptonic channel are higher branching ratio (see Table 1.2 and Table 1.3) and full kinematic reconstruction (which allows to define an actual scale of the process which is not possible to do in case of fully leptonic WW channel because of 2 undetected neutrinos). But on the other hand a fully leptonic analysis has much less background than semileptonic analysis which has significant contributions from top and W+jets processes. Analyses use either  $p_T$  or mass (mass of two leptons in case of the fully leptonic analysis or  $p_T$  of the dijet system in case of the semileptonic analysis) variables to extract limits on aTGCs. This is driven by the fact that effects from aTGCs are mostly prominent at high masses or momenta.

Comparison of results from CMS, ATLAS, D0 and LEP is shown in Figure 1.4. The strongest limits are obtained by CMS at 8 TeV [49].

This thesis presents a first search for anomalous couplings with the CMS experiment at  $\sqrt{s} = 13$  TeV in the boosted channel, where two quarks from hadronically decaying W or Z boson form a single jet with large radius. Limits are extracted from the diboson mass distribution. This is motivated by the fact that diboson mass is the actual scale of the process and thus a measurement can safely be done in the region far from the unitary bound [55].







the LHC was colliding protons with center-of-mass energy of 7 and 8 TeV. This was followed by the technical stop in 2013-2014 and then in 2015 the LHC started colliding beams with energies of 6.5 TeV which is the record energy for a particle accelerator. In order to operate at such high energies LHC uses strong magnetic field. The reason for this comes from the following relation:

$$\frac{p}{q} = B\rho \quad (2.1)$$

where  $p$  is a particle momentum,  $q$  is a charge of the particle,  $B$  is the strength of the magnetic field and  $\rho$  is the radius of a particle in an accelerator. For the case of a proton accelerator this can be written as:

$$\frac{1}{\rho[m]} \approx 0.3 \frac{B[T]}{p[\text{GeV}]} \quad (2.2)$$

This shows that in order to go to high energies one has to increase either the radius or the strength of the magnetic field. Since the radius of the tunnel is fixed the LHC uses magnetic field strength of 8.3 T. The magnetic field strength is proportional to the current according to Biot-Savart law [60]. The LHC design magnet current is  $I = 11850$  A. To avoid ohmic losses superconductivity is used. The LHC magnets are made of NbTi (an alloy of niobium and titanium) and are cooled below 2K using superfluid helium, although the critical temperature of NbTi is around 10 K [61]. This is because superconductors are characterized by the critical magnetic field  $B_c$  above which the resistanceless state disappears.  $B_c$  actually depends on the temperature so that the lower temperature is the higher the critical field is [62]. Magnets are cooled below 2K in order to provide high magnetic field.

There are 1232 dipole magnets in the LHC. The role of the dipoles is to provide the magnetic field perpendicular to the beam direction so that particles get a circular trajectory. Other than being curved the beam has to be focused keeping the bunch structure of the beam in order to get a maximum number of collisions. This is done by the system of quadrupoles. There are 858 quadrupole magnets in total in the LHC. However focusing of the beam actually makes it oscillate around the vacuum tube center. The beam in a storage ring is characterized by the so called chromaticity<sup>1</sup> and it is always non zero in the storage ring with strong focusing [63]. Sextupole magnets take care of this effect. Other magnetic multipoles are used for corrections from the electromagnetic interactions of protons, interaction with the electron cloud in the beam pipe etc.

The complex magnet system described above keeps particles in the circular trajectory but doesn't actually accelerate them (in the longitudinal direction). This is done by radiofrequency (RF) cavities. The LHC uses 8 RF cavities per beam, each with accelerating field 5 MV/m at 400 MHz. Cavities are cooled to 4.5 K. Another role of RF cavities is to keep a bunch structure of the beam. The RF cavities of the LHC are located at Point 4.

Before being injected in the main ring of the LHC protons pass several systems that increase their energy. Injection starts in LINAC 2 where 50 MeV protons are generated, then protons are accelerated by Proton Synchrotron Booster (PSB) to 1.4 GeV and then injected to Proton Synchrotron (PS) where energy of protons is increased to 26 GeV. After that protons are injected

---

<sup>1</sup>chromaticity quantifies the change in focusing: there is a spread in particles momenta and because of that particles travel with slightly different trajectories

into Super Proton Synchrotron (SPS) where they are accelerated to 450 GeV and then protons are injected in the LHC. This is illustrated in Figure 2.2.

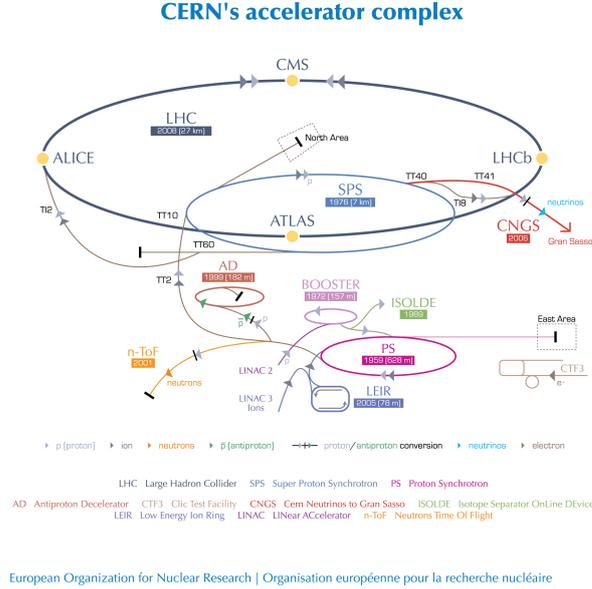


Figure 2.2: The CERN accelerator complex [64].

As it was already mentioned the beam in the LHC consists of bunches. A bunch contains of about  $10^{11}$  protons. Bunches are usually grouped in trains that typically consist of either 48, 96 or 144 bunches.

The number of events per second at the collision point of a collider is given by:

$$N_{\text{events}} = L \cdot \sigma_p \quad (2.3)$$

where  $L$  is the luminosity and  $\sigma_p$  is the cross-section of the given process. The luminosity is given by the formula:

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F \quad (2.4)$$

where  $N_b$  is the number of particles per bunch,  $n_b$  is the number of bunches per beam,  $f_{\text{rev}}$  is the revolution frequency (11245 Hz),  $\gamma_r$  is the relativistic factor,  $\epsilon_n$  is the normalized transverse emittance<sup>2</sup>,  $\beta^*$  is the beta function<sup>3</sup> at the collision point and  $F$  is a geometric factor due to the crossing angle<sup>4</sup> at the interaction point.

As shown in Figure 2.3 processes like WW, ZZ and top production that are sensitive to new physics have rather small cross-sections. In order to make a quantitative statement about

<sup>2</sup>the emittance  $\epsilon$  is defined as area in the phase space in the plane transverse to the beam motion divided by  $\pi$  and the normalized emittance  $\epsilon_n$  is defined as  $\epsilon_n = \beta\gamma\epsilon$  where  $\beta\gamma$  is the relativistic factor [65]

<sup>3</sup>the beta function determines the envelope of all particle trajectories at the given position. The typical value of  $\beta^*$  used during data taking in 2015 at the interaction point in the CMS experiment was 80 cm.

<sup>4</sup>a typical value of the crossing angle used during data taking period in 2015 is 145.0  $\mu\text{rad}$

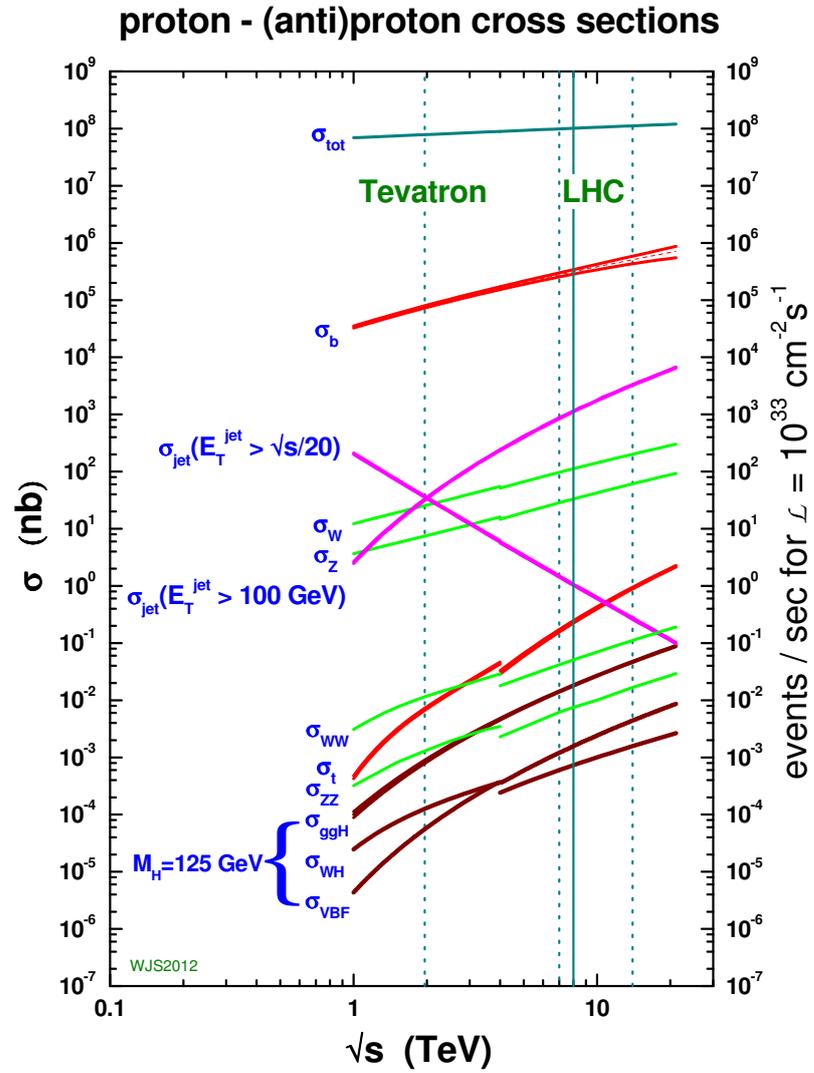


Figure 2.3: Cross-sections of different processes at the LHC and Tevatron [66].

possible effects from new physics a significant number of signal events should be observed. This means that as large as possible luminosity (see Eq. (2.3)) is desired when looking for effects from new physics.

## 2.2 The CMS experiment

The CMS (Compact Muon Solenoid) experiment is located at Point 5 of the LHC tunnel (Figure 2.1). The layout of the CMS detector is shown in Figure 2.4.

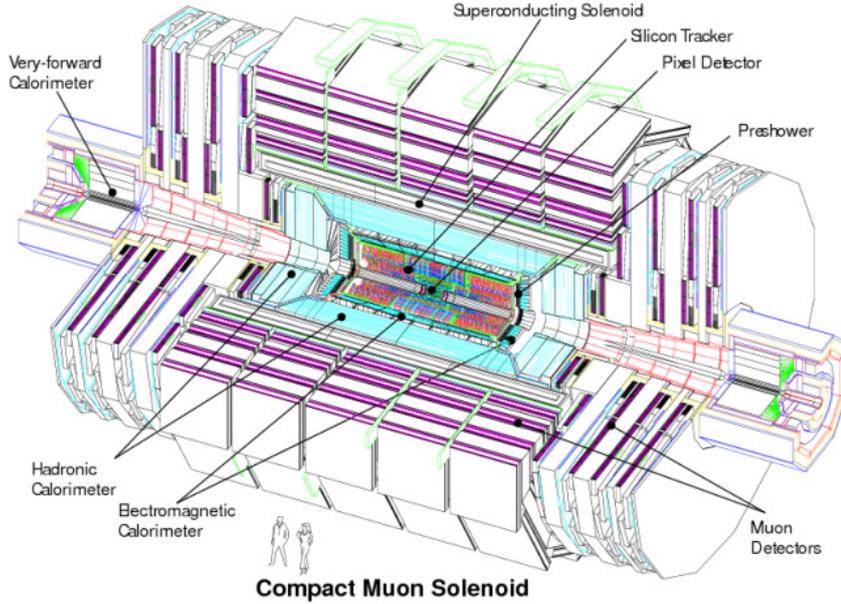


Figure 2.4: The CMS detector [67].

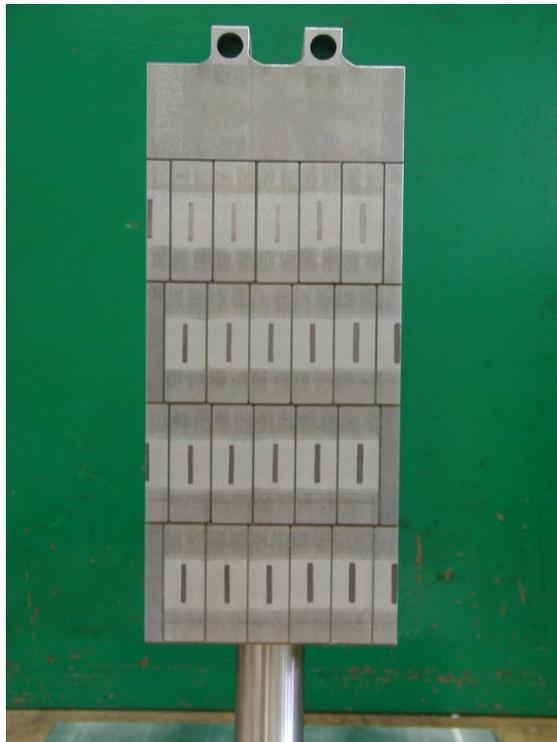
The CMS detector is 21.6-m long and has a diameter of 14.6 m. Its weight is 14000 t [68]. The key element of the CMS detector is the superconducting solenoid. It's 13-m long and 6-m inner diameter. The solenoid provides magnetic field of 3.8 T. Tracking and calorimeter systems are located inside of the solenoid. Tracking system consist of the pixel detector which is located close to the interaction point and the silicon tracker. The calorimeter system contains of electromagnetic calorimeter (ECAL) and hadron calorimeter (HCAL). Muon detectors are located outside of the solenoid. Four muon stations are installed inside of the iron yoke for the return-flux of the magnetic field. The yoke consists of 6 endcap disks and 5 barrel wheels (400t - 1920t) and covers most of the  $4\pi$  solid angle [56].

CMS uses a coordinate system with the center at the interaction point,  $z$ -axis in the beam direction,  $y$ -axis vertically upward and  $x$ -axis pointing to the center of the LHC. The azimuthal angle  $\phi$  is measured from the  $x$ -axis in the  $x - y$  plane, and  $r$  denoted the radial coordinate. The polar angle  $\theta$  is measured from the  $z$ -axis and the pseudorapidity is defined as:

$$\eta = -\ln \tan \frac{\theta}{2} \quad (2.5)$$

### 2.2.1 Superconducting magnet

The solenoid has the 4-layer winding made from NbTi and aluminium. The cross-section of the coil of the solenoid is shown in Figure 2.5. NbTi acts as a conductor carrying the current of about 20 kA and is surrounded with pure aluminium. The aluminium doesn't conduct any current (critical temperature of aluminium is 1.2 K [61]) and serves as a thermal stabilizer (which is motivated by the high thermal conductivity of aluminium 247 W/(m·K) [69]). The layer of pure aluminium is surrounded by an aluminium alloy which serves for mechanical reinforcement. This is motivated by the fact that the strong magnetic field induces a pressure of about 6 MPa on the coil itself (the elastic modulus of aluminium is 69 GPa [69]). The total mass of the conductor is 220 t which is cooled down to the temperature about 4.5 K using liquid helium. The ratio of the stored energy and the cold mass is very high – 11.6 KJ/kg which is the largest among the detector magnets.



**Figure 2.5:** Cross-section of the coil of the CMS magnet [70].

### 2.2.2 Inner tracking system

The tracking system surrounds the collision point. The typical instantaneous luminosity delivered to the CMS was about  $10^{33} \text{ cm}^{-2}\text{s}^{-1}$  which resulted in about 100 tracks from about 10 proton-proton interactions every 25 ns. Such regime requires high granularity and fast response of the tracking system so that tracks of all these particles are reconstructed with high efficiency.

The CMS collaboration constructed the tracker system based on silicon detector technology. The layout of the CMS tracker is shown in Figure 2.6.

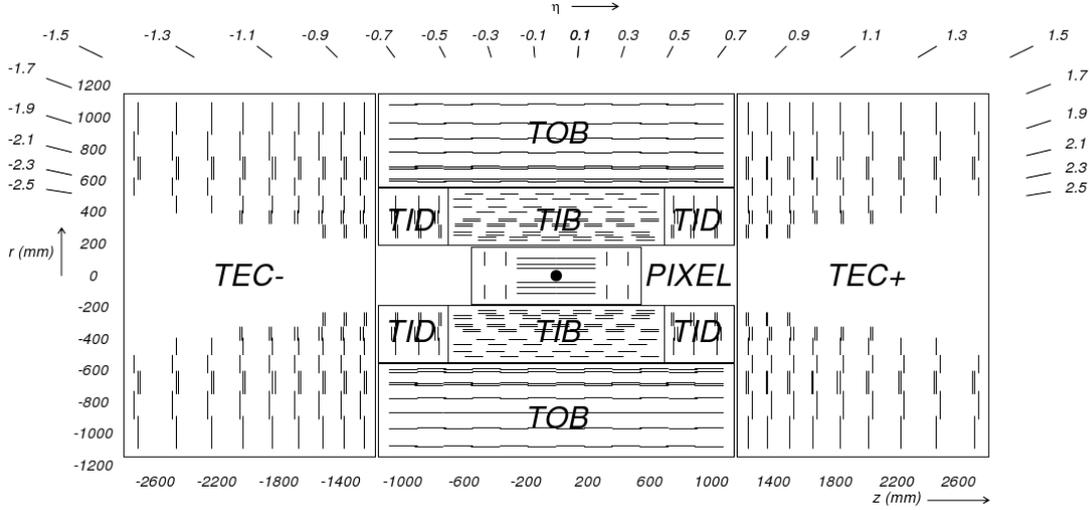
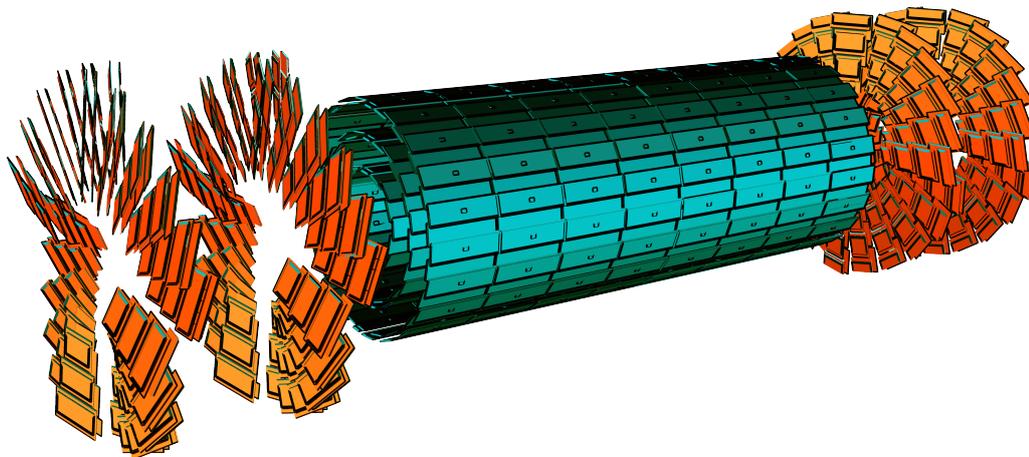


Figure 2.6: The layout of the CMS tracker system [71].

The tracking system of the CMS detector is composed of a pixel detector and a silicon strip detector [56]. The principle of such detectors is based on p-n junction. When the particle travels through the depletion region it creates electron-hole pairs that are collected by electrodes.

The pixel detector is composed of 3 barrel layers and 2 endcap disks (see Figure 2.7). Barrel layers have radii from 4.4 cm to 10.2 cm and are 53 cm long. Endcap disks are extending from about 6 cm to 15 cm and are located at  $z = \pm 34.5$  cm and  $z = \pm 46.5$  cm. In total the pixel detector has 64 million pixels with the size of each pixel cell  $100 \times 150 \mu\text{m}^2$  and covers the area of about  $1 \text{ m}^2$ . This provides a low occupancy of the pixel detector. The spacial resolution of pixel is about  $15\text{-}20 \mu\text{m}$ . The resolution of the pixel detector is important for identification of secondary vertices.

A strip detector is composed of sensors of p-on-n type: strip with  $p^+$  implantation on one side and  $n^+$  implantation of the other side, orthogonal to each other. A silicon strip detector is composed of three subsystems: the Tracker Inner Barrel and Disks (TIB/TID), the Tracker Outer Barrel (TOB) and the Tracker Endcaps (TEC+ and TEC- where the sign indicates the location along the z-axis). The TIB/TID extends up to radius 55 cm and is composed of 4 barrel layers with 3 disks on each side providing up to  $4 r - \phi$  measurements on the trajectory. TIB has a strip pitch of  $80 \mu\text{m}$  on first two layers with resolution of  $23 \mu\text{m}$  and  $120 \mu\text{m}$  on the third and fourth layers with resolution of  $35 \mu\text{m}$ . The TID has the pitch between  $100 \mu\text{m}$  and  $141 \mu\text{m}$ . The TOB has 6 barrel layers and an outer radius of 116 cm. It has strip pitches of  $183 \mu\text{m}$  on the first 4 layers and  $122 \mu\text{m}$  on layers 5 and 6 providing  $6 r - \phi$  measurements with single point resolution of  $53 \mu\text{m}$  and  $35 \mu\text{m}$  respectively. The TIB/TID and TOB are located in z-range between  $\pm 118$  cm. The TEC $\pm$  covers the range  $124 \text{ cm} < |z| < 282 \text{ cm}$  and  $22.5 \text{ cm} < |r| < 113.5 \text{ cm}$ . Both TEC $\pm$  have 9 disks. The TEC has strips with pitches from  $97 \mu\text{m}$  to  $184 \mu\text{m}$  and provides  $9 \phi$  measurements. Modules in the first 2 layers of TIB and TOB, the first 2 rings of



**Figure 2.7:** The layout of the CMS pixel system [72].

the TID, rings 1, 2, 5 of the TECs have a second microstrip detector which is mounted with the stereo angle of 100 mrad in order to measure the second coordinate ( $r$  on the disks and  $z$  in the barrel). The tracker system covers a range of  $|\eta| < 2.5$  [56].

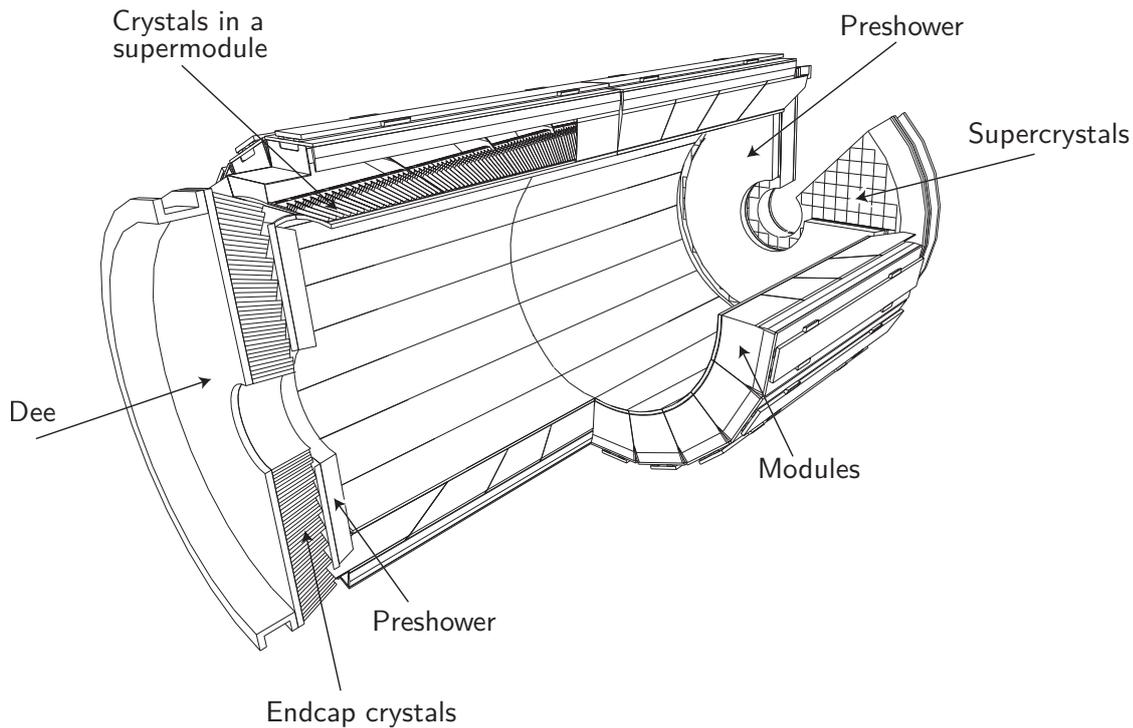
### 2.2.3 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is a homogeneous calorimeter. This means that the material of the ECAL serves both as absorber and active material. The ECAL is made of 61200 lead tungstate ( $\text{PbWO}_4$ ) crystals in the barrel part and 7324 crystals in the two endcaps. The choice of these crystals was motivated by the short radiation length  $X_0^5$  (0.89 cm), high density (8.28 g/cm<sup>3</sup>) and small Molière radius<sup>6</sup> (2.2 cm). The crystals serve as scintillators and have scintillation time of about 25 ns (the bunch spacing used during data-taking in 2015). Photodetectors are needed in order to collect the light emitted by crystals. For this purpose avalanche photodiodes (APDs) are used in the barrel and vacuum phototriodes (VPT) in the endcaps. The layout of the ECAL is shown in Figure 2.8.

The ECAL barrel covers the pseudorapidity range  $|\eta| < 1.479$ . The granularity is 360-fold in  $\phi$  and  $2 \times 85$ -fold in  $\eta$  (61200 crystals in total). They are mounted with the small angle ( $3^\circ$ ) with respect to the vector from the interaction point in order to avoid cracks aligned with particle trajectories. The inner radius of the ECAL barrel is 1.29 m. The cross-section of crystals is  $22 \times 22 \text{ mm}^2$  at the front side of the crystal and  $26 \times 26 \text{ mm}^2$  at the rear side. The length of the crystal is 230 mm ( $25.8 X_0$ ). A pair of APDs is mounted on each crystal. Crystals are located in submodules which are organized in modules of different type depending on  $\eta$  direction and contain 400 or 500 crystals. 4 modules are organized in a supermodule which contains 1700 crystals. There are 18 supermodules in region  $z > 0$  and  $z < 0$  parts of the ECAL barrel, covering

<sup>5</sup>radiation length characterises the longitudinal development of the electromagnetic shower.

<sup>6</sup>The Molière radius is defined as the radius of the cylinder containing 90% of the shower energy [73].



**Figure 2.8:** The layout of the electromagnetic calorimeter [56].

$20^\circ$  in  $\phi$  each.

The endcaps cover the range  $1.479 < |\eta| < 3.0$ . This is composed of 2 *Dees* (see Figure 2.8) containing 3662 crystals each. The crystals have the rear cross-section  $30 \times 30 \text{ mm}^2$  and the front cross-section  $28.62 \times 28.62 \text{ mm}^2$  and the radiation length  $220 \text{ mm}$  ( $24.7 X_0$ ). One VPT is glued on the back of each crystal.

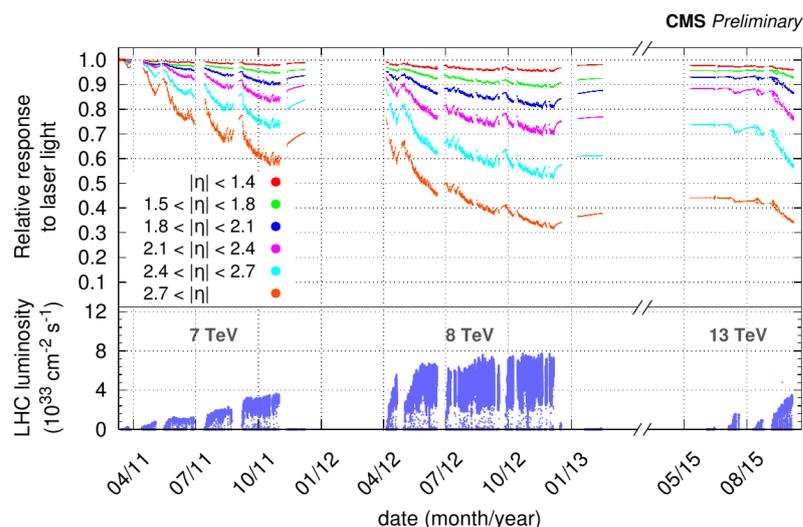
The preshower detector covers the range  $1.653 < |\eta| < 2.6$ . It is used to improve identification of neutral pions decaying to two photons. The preshower consists of two layers. Each layer is composed of lead radiators and silicon strip sensors.

The transparency of ECAL crystals changes in different conditions of LHC. Its evolution is measured by laser pulses. This is illustrated in Figure 2.9. This effect is taken into account in the calibration of ECAL.

The relative ECAL energy resolution is about 2% in the barrel and about 4-5% in the endcaps [75].

#### 2.2.4 Hadron calorimeter

Hadron calorimeter (HCAL) is a sampling calorimeter. This means that an absorber and a scintillator are different materials and layers of a dense absorber and a scintillator are repeated. As a hadron interacts with the material of an absorber it can produce numerous secondary particles. These secondary particles travelling through the alternating layers of absorber will create a cascade of particles. As the shower develops particles pass scintillator material and the



**Figure 2.9:** History of ECAL response with laser data [74]. The response is up to 6 % in the barrel and reaches up to 30 % at  $|\eta| \approx 2.5$ . The response is up to 70% in the region closest to the beam pipe.

light is emitted which is then carried by optical cables to the readout electronics.

The HCAL is located between the ECAL and the superconducting solenoid. The HCAL consists of the barrel (HB), the endcaps (HE), the outer calorimeter (HO) and the forward calorimeter. The layout of the HCAL is shown in Figure 2.10.

The HB covers the range  $|\eta| < 1.3$ . It consists of two half-barrels (HB+ and HB-) each of those consist of 18 identical azimuthal wedges. The innermost and outermost layers are made from steel (40-mm- and 75-mm-thick) and intermediate layers are made from brass (50.5-mm- and 56.5-mm-thick), 16 layers in total. The total thickness of the absorber is 5.82 interaction lengths [76] ( $\lambda_I$ ) at  $\theta = 90^\circ$  and increases as  $1/\sin\theta$ . The active material is composed of 4 mm thick plastic scintillating tiles [77]. Within each tile wavelength shifting fibres collect the light. Through the cabling the light goes from wavelength shifting fibres to hybrid photodiodes (HPDs) [78] which then convert the optical signal to the electronic.

The HE covers the range  $1.3 < |\eta| < 3$ . It is composed of 79-mm-thick brass with 9-mm gaps for scintillators. The total length of HE is about  $10 \cdot \lambda_I$ . The granularity is  $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$  for  $|\eta| < 1.6$  (HB and HE) and  $\Delta\eta \times \Delta\phi \approx 0.17 \times 0.17$  for  $|\eta| \geq 1.6$ .

Since the ECAL barrel and the HCAL barrel (HB) don't provide sufficient containment of hadronic showers the HCAL is extended outside of the solenoid. The HO is located close to the muon system and thus follows its geometry. It is composed of five rings, the central ring has 2 layers of scintillators and other rings have only one layer of scintillator. The iron of the yoke and the solenoid serve as absorbers. The thickness of the HCAL is then extended to 11.8  $\lambda_I$  except the endcap-barrel boundary region.

The forward calorimeter (HF) covers the range  $3 < |\eta| < 5$  (located at either ends of the detector). This region is especially important for the measurement of the missing transverse

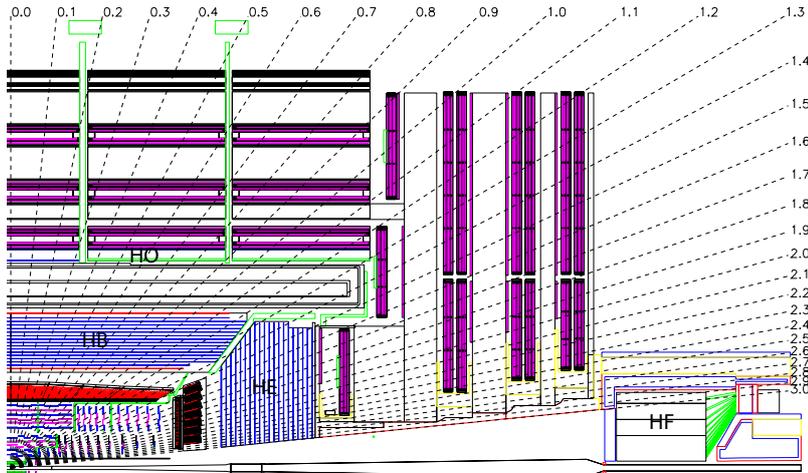


Figure 2.10: The layout of the hadron calorimeter [56].

energy. It consists of the steel absorber composed of the 5-mm-thick plates and quartz fibres. The total length of the HF is about  $10 \cdot \lambda_I$ . The outer radius is 130 cm and the distance from the front face is 12.5 cm.

### 2.2.5 Muon system

The muon system consists of three subsystems: drift tubes (DT) in the barrel region, cathode strip chambers (CSC) in the endcap region and resistive plate chambers (RPC) [79]. This is illustrated in Figure 2.11.

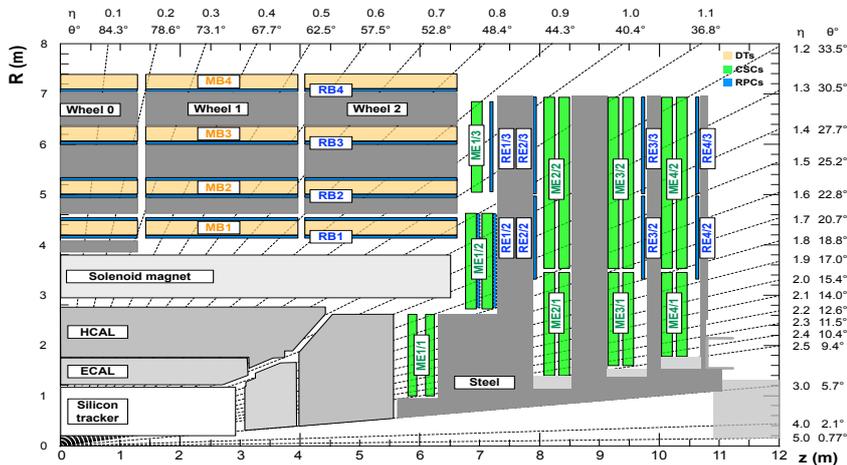


Figure 2.11: The layout of the CMS muon system after the Long Shutdown 1 (LS1) Upgrade, 2013-2014 [80].

DT cover the range about  $|\eta| < 1.2$ . There are 4 stations (denoted as MB1, MB2, MB3, MB4 in

Figure 2.11), each station has a form of a cylinder around the beam pipe. Stations are interspersed with return yokes (5 wheels). 3 inner stations have 60 chambers and the outer has 70 chambers. The DT chamber is made of 3 or 2 superlayers (SLs), each made of 4 layers of rectangular drift cells. The drift cell is filled with gas composed of 85% of argon and 15% of carbon dioxide. 2 outer SLs have wires parallel to the beam direction provide measurement of  $r$  and  $\phi$  while the inner SL (not present in the fourth station) has wires perpendicular to the beam direction and measures the  $z$  position. The hit resolution is about  $200\ \mu\text{m}$  in SLs measuring  $\phi$  and central ( $z = 0$ ) SL measuring  $\theta$ , about  $400 - 1000\ \mu\text{m}$  in other  $\theta$  SLs [81].

CSCs cover the range about  $0.9 < |\eta| < 2.4$ . There are 468 cathode strip chambers in total organized in rings of 72 or 36 (ME1, ME2, ME3, ME4 in Figure 2.11). During the Long Shutdown 1 in the outermost ring new 72 chambers (ME4/2 in Figure 2.11) were installed and commissioned. This has improved muon reconstruction in the region  $1.2 < |\eta| < 1.8$  and reduced the fake rate for triggering and offline reconstruction [82]. Chambers have a trapezoidal shape and cover  $10^\circ$  or  $20^\circ$  in  $\phi$ . Each chamber is a multiwire proportional chamber [83] with 6 anode wire planes interleaved among 7 cathode panels. The chambers are filled with the gas composed of 40% Ar, 50%  $\text{CO}_2$ , 10%  $\text{CF}_4$ . Wires in the chambers are located azimuthally and strips are milled on the cathode panels. Thus the CSCs provide measurement of  $r$  and  $\phi$  coordinate. The hit resolution of CSC stations in 2015 is about  $50 - 150\ \mu\text{m}$  [84].

RPCs are gaseous parallel plate chambers covering  $|\eta| < 1.8$ . The basic module consists of 2 parallel electrodes made of high resistive material with 2mm gas gap. Modules are placed on top of each other with copper readout strips in between [85]. In the barrel RPCs form 6 coaxial cylinders organized in 4 stations. In the first 2 stations RPCs are located on the inner and outer side of DT chambers and the third and fourth stations are located only on the inner side of the DT layer. There are 4 RPC stations in the endcap arranged in the concentric rings (the fourth station, RE4/2 and RE4/3 in Figure 2.11, installed during LS1). CSC and DT have a drift time of more than 25 ns which makes it difficult to assign a muon to the right bunch crossing. However, RPC have a time resolution of about 3 ns [85] which allows to assign a muon to the correct bunch crossing. The spacial resolution of RPC is about 1 cm [86].

## 2.2.6 Trigger and data acquisition

As it was already mentioned bunches in the LHC are spaced by 25 ns. This leads to the event rate of about 40 MHz which leads to extremely large data rate ( $\sim 10\ \text{TB/s}$ ). This is reduced by the trigger system by selecting events online. The CMS trigger system has 2 levels: Level 1 (L1) trigger and High Level Trigger (HLT). The L1 trigger reduces the event rate to  $\sim 100\ \text{kHz}$  and then HLT to  $\sim 1\ \text{kHz}$ .

The L1 trigger consists of the calorimeter trigger and the muon trigger and doesn't use information from the tracker. The muon trigger consists of three track finders and global muon trigger. Track finders cover different pseudorapidity range and receive trigger primitives (basic objects for L1 that have a reduced position and energy resolution) from DT, CSC, RPC. The global muon trigger removes duplicates across track finder boundaries and ranks muons according to  $p_T$  measurement accuracy [87]. The upgrade of the L1 trigger during LS1 added a possibility to calculate muon isolation which allowed to reduce the event rate in higher luminosity environment [88]. The L1 calorimeter trigger receives trigger primitives from ECAL, HCAL

and HF [89]. It has two layers which determine photon/electron, jets, total transverse energy, missing transverse energy, jet counts and perform pileup subtraction. There have been several improvements introduced in the L1 calorimeter system [90]. One of the improvements relevant in the context of this thesis is dynamic clustering for photon/electron energy (an extended region in  $\phi$  is used with respect to  $4 \times 4$  trigger towers region used in Run I). This has improved the position resolution of the electron/photon candidate by the factor of 4. Information for the L1 muon and calorimeter trigger is then combined by the Global Trigger (GT).

The HLT is implemented in software and is running on a computer farm which includes about 16000 CPU cores [91]. The HLT uses the same reconstruction software as the offline reconstruction optimised for the strict requirements of the online selection. The HLT menu is implemented in paths which have a module structure. Each path is composed of reconstruction sequences and filters which apply selection of physics objects (jets, electrons, muons, missing transverse energy, b-tagging etc).

The Trigger Control and Distribution System (TCDS) distributes L1 accepts to the front-ends, collects information about front-end readiness [92]. It accepts or vetoes an event based on the state of front-ends and trigger rules<sup>7</sup>. If an event is accepted by TCDS is processed by the data acquisition system (DAQ) [92]. The DAQ system builds events in 2 stages (data concentration and core event builder). The reconstruction is run by builder units (BU) and then are passed to filter units (FU, these machines run HLT [93]). BUs write accumulated events to the local disk and FUs perform the selection. After that events are passed to the storage system. The DAQ system was redesigned for Run 2 which is discussed in [93]. For example, the total bandwidth of the DAQ system is increased to 200 Gb/s which allows to have event size up to 2 MB.

The data-taking is split in runs in which all subsystems, DAQ, triggers are configured. The duration of the run is can be from  $\sim$  minutes to several hours. Each run is further divided into the lumi sections (LS). The duration of the LS is 23.3 seconds. The value of instantaneous luminosity is considered to be constant during a given LS.

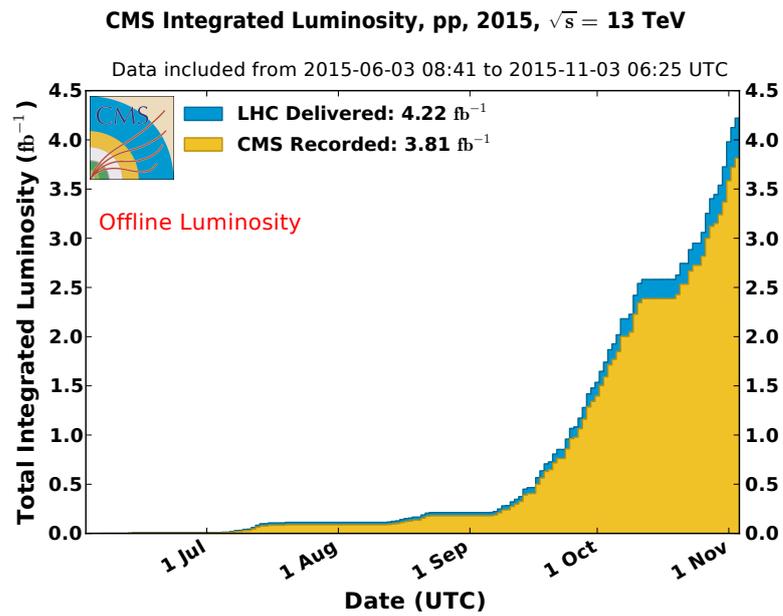
### 2.2.7 Luminosity

During 2015 the CMS experiment has recorded an integrated luminosity of  $3.8 \text{ fb}^{-1}$ . This is illustrated in Figure 2.12.

After the offline data quality certification and due to non-availability of the magnet in some runs the integrated luminosity used in this thesis is reduced to  $2.3 \text{ fb}^{-1}$ . CMS uses five detectors to monitor and measure the instantaneous luminosity: tracker and pixel detector, DT in barrel, HF in the forward region, the Fast Beam Conditions Monitor (BCM1f) and the Pixel Luminosity Telescope (PLT). The offline luminosity measured is done with the Pixel Cluster Counting method (PCC) which exploits the low occupancy of the pixel detector and provides the most precise measurement of the luminosity. The absolute calibration of the luminosity is determined from Van der Meer scan when beams are scanned with respect to each other. The uncertainty of the luminosity measurement in 2015 is estimated to be 2.7 % [95].

---

<sup>7</sup> trigger rules restrict the number of accepted events to certain bunch crossings (BX): for example, if an event was accepted in the particular BX then an event cannot be accepted in the next BX.



**Figure 2.12:** Cumulative offline luminosity versus day delivered to (blue), and recorded by CMS (orange) during stable beams and for p-p collisions at 13 TeV center-of-mass energy in 2015. Given is the luminosity as determined from counting rates measured by the luminosity detectors after offline validation [94].

# 3 Event simulation and reconstruction

This chapter describes basic steps of event simulation. A brief overview of event generators used in this thesis is given. A short overview of the *particle-flow* reconstruction algorithm is provided and reconstruction and identification of objects used in thesis is discussed further.

## 3.1 Event simulation

The event simulation at the LHC starts with the *hard process*, then *parton shower* develops to energies of  $\sim 1$  GeV where perturbation theory is not applicable. After that partons are converted to hadrons via a *hadronization* model and then unstable particles are decayed. In addition to that there might be soft interaction of the proton remnants (*underlying event*) [96] and additional soft interactions of protons from the same or nearby bunch crossings (*pileup*). This is illustrated in Figure 3.1.

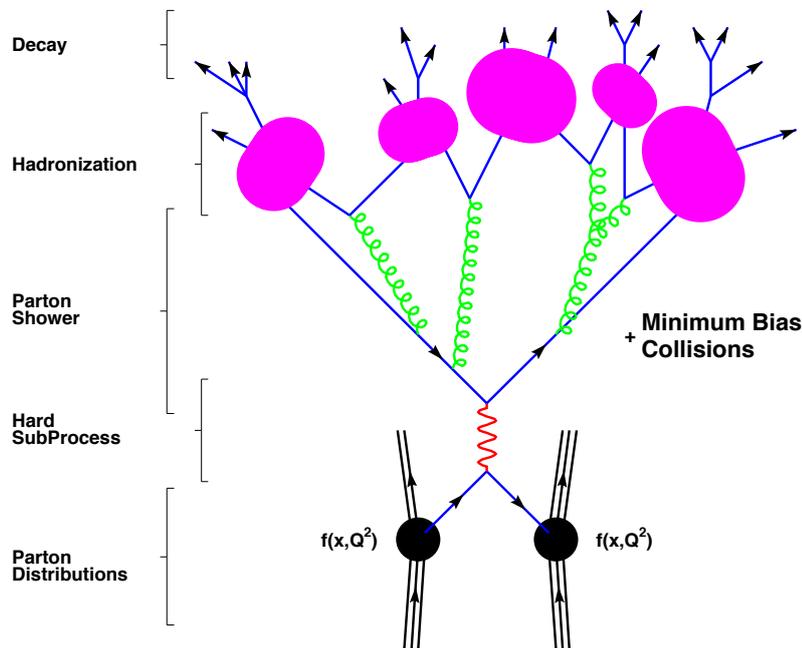


Figure 3.1: The basic structure of the event simulation [97].

### Hard process

The calculation of the hard process is based on the factorization ansatz which is given by the formula:

$$\sigma(pp \rightarrow X) = \sum_{a,b} \int dx_1 dx_2 f_a(x_1, \mu_F) f_b(x_2, \mu_F) \sigma_{ab \rightarrow X}(\hat{s}, \mu_F, \mu_R) \quad (3.1)$$

where  $\sigma(pp \rightarrow X)$  is the cross-section of 2 protons to the final state X,  $a$  and  $b$  denote different parton flavours,  $x_1$  and  $x_2$  is the portion of the proton's momentum carried by a parton,  $f_a(x, \mu_F)$  is the parton distribution function (PDF) for the parton  $a$  which gives the probability of the parton to carry a momentum fraction  $x$ ,  $\sigma_{ab \rightarrow X}$  is the cross-section of partons  $a$  and  $b$  to the final state,  $\mu_F$  and  $\mu_R$  denote factorization and renormalization scale accordingly,  $\hat{s} = s \cdot x_1 x_2$  [98]. The factorization scale separates long- and short-distance physics. The renormalization procedure is done in order to deal with infinities coming from the loops in Feynman diagrams and to express observables in terms of finite numbers. It leads to the fact that the coupling is chosen at some scale (renormalization scale) and coupling depends on this scale (or as often said *running coupling*) [25]. Both factorization and renormalization scales cannot be chosen from the first principles. For the processes  $2 \rightarrow 2$  the scale  $Q^2$  is chosen so that  $Q^2 = \mu_F = \mu_R$ . For example, in case of  $s$ -channel process the scale is chosen to be the mass of the resonance  $M$ ,  $Q^2 = M^2$  or for the production of pair of massless particles with transverse momentum  $p_T$ ,  $Q^2 = p_T^2$  [99]. Thus the total cross-section has 2 important ingredients: PDFs and partonic cross-section ( $\sigma_{ab \rightarrow X}$ ).

PDFs are determined from data on deep inelastic scattering and Drell-Yan production, vector boson production, single jet inclusive cross-section etc. This thesis uses NNPDF 3.0 set [100].

The cross-section of the reaction of 2 particles in  $n$  particles ( $2 \rightarrow n$ ) is calculated with the following formula [101]:

$$\sigma_{ab \rightarrow X} = \frac{1}{2\lambda^{1/2}(s, m_a^2, m_b^2) \cdot (2\pi)^{3n-4}} \int \prod_{i=1}^n \frac{d^3 p_i}{2E_i} A(p_i) \delta^4(p_a + p_b - \sum_i p_i) \quad (3.2)$$

where  $s$  is a square of the center-of-mass energy,  $m_a$  and  $m_b$  ( $p_a$  and  $p_b$ ) denote masses (4-vectors) of the initial state particles,  $p_i$  is a four-vector of the final state particle  $i$ ,  $n$  is the number of particles of in the final state,  $\delta^4(\cdot)$  is a 4-dimensional delta function,  $A(p_i)$  is the probability of the transition of the initial state with  $p_a$  and  $p_b$  to the final state with  $p_i$  ( $i = \overline{1, n}$ ) and  $\lambda(x, y, z) = x^2 - 2(y+z)x + (y-z)^2$ .  $A(p_i)$  is defined from the *matrix element*:  $A(p_i) = |M_{fi}|^2$  where  $M_{fi} = \langle p_1, \dots, p_n | \hat{T} | p_a, p_b \rangle$  where  $\hat{T}$  is T-matrix [24]. The matrix element can be calculated as a series in perturbation theory with Feynman rules [24] using a coupling constant as an expansion parameter. Usually cross-sections are computed in leading order (LO) QED and next-to-leading (NLO) or next-to-next-leading order (NNLO) QCD. This is motivated by the fact that QCD coupling constant is larger than QED coupling constant [21]. The most challenging part of Eq. (3.2) is that the integral has  $3n - 4$  dimensions [101]. This leads to the idea of using Monte-Carlo method to calculate this integrals: the accuracy of the method improves as  $1/\sqrt{N}$  irrespective of dimension where  $N$  is a number of integration points [99]. Thus together with calculating the integral one can sample events according to the distribution in Eq. (3.2) and perform an event generation.

## Parton shower

As was already mentioned a hard process is usually simulated at LO or NLO. The effect of higher orders is described by the parton shower which evolves momentum from hard scales to  $\sim 1$  GeV. The parton shower can be defined as a successive branching of partons (both incoming and outgoing) via QCD interaction ( $q \rightarrow qg$ ,  $g \rightarrow q\bar{q}$ ,  $g \rightarrow gg$ ). The parton shower includes 2 type of processes: initial state and final state radiation (ISR/FSR). The simulation of FSR is based on the *the Sudakov form factor* which defines the probability of parton of the flavour  $i$  not to split into  $i$  and  $j$  partons during the evolution from scale  $Q_1^2$  to  $Q_2^2$  ( $\Delta_i(Q_1^2, Q_2^2)$ ). A random number  $\rho$  between 0 and 1 is chosen, and the equation  $\Delta_i(Q_1^2, q^2) = \rho$  is solved for the initial scale  $q^2$ . If the solution is above the cutoff scale  $Q_0^2$  then branching is generated otherwise the shower is terminated. ISR is done in a backwards step because most of the partons have low energy and it would be extremely rare to produce exactly the right kinematics to produce a hard process of interest [99]. The probability distribution of parton with the given momentum fraction and scale to have come from the parton at lower scale and higher momentum fraction is described by the DGLAP equation [25]. The Monte Carlo implementation is similar to FSR but a non-emission probability is used instead of the Sudakov form factor [99]. The choice of scale in the parton shower evolution is not unique, PYTHIA 8.1 uses a sort of transverse momentum  $p_\perp^2$  [102], for example.

At high energies emission of hard jets is underestimated by the parton shower approach. Thus it is important to simulate the emission of jets at the matrix element. Parton shower and matrix element describe different regions of the phase space. However, when they are used at the same time care has to be taken in order to avoid double counting. Different approaches have been proposed, for example the MLM-matching scheme [103, 104]. The basic idea of matching schemes is to divide the phase space into two regions. More details can be found in reference [105].

## Hadronization

As the parton shower develops the scale is decreasing and at around  $\Lambda_{QCD} \approx 200$  MeV the perturbation theory breaks and hadrons are formed. One general approach to hadronization is a hypothesis of the *local parton-hadron duality*. For example, the flavour of the quark initiating a jet should be found close to the jet-axis [106]. There are two hadronization models widely used: the *Lund string* [107] and the *cluster model* [108].

The assumption of the Lund string model is that a potential of 2 quarks interaction is given by the potential  $V(r) = \kappa r$  and a Coulomb term is neglected with  $\kappa \approx 1$  GeV/fm. Imagine quark and antiquark moving apart from the creation vertex the energy of interaction increase, a gluonic string stretches between them and at some point there is enough energy to produce a new  $q\bar{q}$  pair. Then string segments begin to stretch and break and so on. At the end of the process the system fragments into a number of primary hadrons [99].

The parton shower has the property of the preconfinement of the colour [109]. This means that pairs of colour connected neighbouring partons have an asymptotic mass distribution that falls rapidly at high masses and is asymptotically  $Q^2$ -independent and universal [106]. This is used in the cluster model of hadronization. After the parton shower gluons are split into  $q\bar{q}$ -pairs

and then neighbouring colour-connected quarks are combined into individual clusters. These clusters are characterized by mesonic quantum numbers. Most of clusters undergo quasi-two-body decay. The limited cluster mass (typical cluster mass is about  $Q_0$  [106]) leads to limited transverse momenta and heavy flavour suppression [99].

After the hadronization step a number of unstable hadrons are produced which should decay into stable particles. This step is based mostly on particle properties that can be found in reference [21], however there are several import points: what hadrons and decay modes are included in the simulation is generator-dependent, momentum correlations should be included in multiparticle decays as well as spin correlations [99].

#### Underlying event

The underlying event represents an additional activity which is not associated with the hard process. Several parton-parton interactions might occur which is usually referred as multiple-parton interactions (MPI). These interactions lead to additional color connections and thus increase the number of particles produced in the hadronization step. Another source of extra-activity in the event comes from beam remnants. Though beam remnants are very forward objects colour connections bring on additional hadrons in the lower  $\eta$ -range [99].

Another effect comes from additional interaction of protons which is usually referred as *pileup*. There can be additional pileup interaction in the same bunch crossing (*in-time pileup*), and additional interactions from the previous or subsequent bunch crossing (*out-time pileup*<sup>1</sup>). Pileup interactions are modelled as so-called *minimum bias* events. These events defined as the ones passing a trigger with very loose selection criteria however the exact definition depends on the trigger. The number of these events is randomly chosen from the predefined distribution (see Subsection 5.3.2) and is added to the simulated hard process.

Another activity is caused through the interaction of beam particles with the collimators or residual gas in the beam pipe (also known as *beam halo*). This sort of activity usually causes muons reconstructed along the beam pipe. This events are suppressed by applying corresponding filters (see Subsection 5.2.2).

#### 3.1.1 Event generators

The steps described above are performed by event generators. There are several event generators used in this thesis. A short overview of them is given below.

PYTHIA 8.1 is a general purpose event generator. It is able to simulate low multiplicity processes but is mostly used for modelling of the parton shower, hadronization and underlying event. The parton shower is based on  $p_T$ -ordered evolution. The Lund string model is used for hadronization [111]. This thesis uses CUETP8M1 tune [112, 113].

POWHEG is an event generator that provides simulation of events at NLO QCD [114]. Simulation of the parton shower is usually interfaced with PYTHIA 8.1 or HERWIG. Matching is done with the POWHEG-method [115] that avoids assignment of negative event weights.

---

<sup>1</sup>it should be noted that the actual reason for the out-time pileup is caused by the time resolution of ECAL and HCAL pulses. Deconvolution of signals is performed by the trigger primitive generators and is not 100%-efficient [110].

`MADGRAPH5_aMC@NLO` [116] provides event generation at both LO and NLO. It also provides matching to the parton shower. In this thesis `MADGRAPH5_aMC@NLO` is used with MLM-matching scheme [103, 104] and is interfaced with `PYTHIA` 8.1 for the parton shower and hadronization modelling. `MADSPIN` module provides interface to properly account for spin correlation in the decays of heavy particles like W/Z-bosons, top-quark. The disadvantage of using `MADGRAPH5_aMC@NLO` with NLO precision is that some fraction of simulated events (usually about 30%) get negative weight assigned. Also a possibility to merge final states with different jet multiplicities is provided (which is relevant for W+jets events in this thesis, for example). Also the so called FxFx-merging is used [117] for the simulation of some background processes.

## 3.2 Event reconstruction

The CMS uses the *particle-flow (PF)* reconstruction approach [118, 119]. It relies on the consistent usage of the signals (or absence thereof) in all detector layers. For example, a charged hadron can be identified by the topological link of a track, a possible ECAL cluster, an HCAL cluster and no signal in muon chambers. Establishing a connection between information from the all detector layers opens a possibility of an exclusive identification of all particle types and an optimal determination of their properties. It is important to notice that this sort of global event description is only possible with a fine detector granularity otherwise signal from different particles would merge reducing the identification capability.

The reconstruction and identification algorithm first proceeds with *linking* particle-flow elements arising from each single particle. A pair of particle-flow elements is considered with the restriction to the nearest neighbours. The link algorithm produces *blocks* of elements. Blocks typically contain one, two or three elements. Blocks are the input for the particle reconstruction and identification algorithm. First, muons are reconstructed and identified and the corresponding particle-flow elements (tracks and clusters in ECAL and HCAL) are removed from the corresponding blocks. Then electron reconstruction follows aiming also to collect Bremsstrahlung photons on the way and then corresponding tracks and ECAL/preshower clusters are excluded. Isolated photons are identified at the same step. The remaining elements in the block can contribute to the charge hadrons, neutral hadrons and photons from jets. Finally, when all blocks have been processed and all particles identified, the global event description becomes available.

### 3.2.1 Track reconstruction

The magnetic field allows to reconstruct the momentum of the charged particles using the curvature of the track. The track is reconstructed from hits in the pixel and strip tracker. The CMS uses the *Combinatorial Track Finder (CTF)* which is the extension of the Kalman filter (KF) [120]. Tracks are produced through subsequent passes of the CTF algorithm – *iterative tracking*. The basic idea is that on the initial iterations tracks that are the easiest to find are searched. After each iterations hits that have been already associated to the track are removed and thus the combinatorial complexity of the following iterations is reduced. This technique improves tracking efficiency with respect to single iteration tracking and reduces combinatorial

fake tracks (also acceptance is loosen up to 200 MeV while the single iteration track finding algorithm reconstructed tracks with  $p_T > 0.9$  GeV). There are 10 iterations, and each iteration consists of 4 steps [121]:

- *seed generation* which provides initial track candidates found with only few hits. Track seeds define the starting point of particle trajectory parameters. Seed is a combination of two or three hits in the pixel or strip detector. At each iterations different detector layers are used and quality selection on the extracted track parameters is applied. For triplets pairs in the innermost layers are found and then another hit is searched in the outer layer.
- *track finding* based on Kalman filter by extrapolating seed trajectories along the expected path of the charged particle searching for additional hits to be assigned to the track candidate.
- *track fitting* which provides estimation of trajectory parameters. Filtering and smoothing procedure takes into account effects of the material and inhomogeneous magnetic field. Also spurious hits are searched and removed.
- *quality selection* of tracks based on minimum (maximum) number of layers,  $\chi^2/ndof$  of the track and track impact parameters.

The first three iterations are seeded with pixel triplets and reduce the number of hits by 40 % (20%) in the pixel (strip) detector. The fourth and fifth iterations recover tracks with one or two missing hits in the pixel detector. The next two iterations reconstruct displaced tracks reducing number of hits in the strip detector. The eighth iteration is dedicated to the dense core of the jets where hit can merge. Merged pixel clusters that are compatible with the energy deposits in the calorimeters are split into several hits. Each of these hits is paired with remaining hits in the strip detector to form a seed for this iteration. The last two iterations are designed to increase muon-tracking reconstruction efficiency by using information from muon detectors.

#### 3.2.2 Primary-vertex and beam spot reconstruction

The beam spot represents a 3D-region where LHC beams collide in the CMS detector. It is determined on average over many events. There are two methods of measuring the beam spot. The first method uses reconstruction of primary vertices as a function of  $x$ ,  $y$  and  $z$ . Then the mean position is determined through the fit to a likelihood to the distribution of vertices. The second method uses the correlation between transverse impact parameter and  $\phi$  of the impact point (the point of the closest approach to the beam axis) [122]. The first method is used to determine the  $z$ -position and widths of luminous region while the second method is used to determine the transverse position of the beam spot [121].

Primary vertex is defined as a vertex with highest  $\sum_i |\vec{p}_{T,i}|^2$  where  $\vec{p}_{T,i}$  is a track associated to a given vertex. Reconstruction of primary vertices is performed in 3 steps: selection of the tracks, clustering of tracks that appear to originate from the same vertex, fitting the position of the vertex. The track selection includes choosing tracks that are close to the beam spot. Selected tracks are clustered using the deterministic annealing algorithm [123]. After that vertex candidates that have at least 2 tracks are fitted using the *adaptive vertex fitter* [124] to estimate

vertex parameters. The vertex resolution slightly depends on the sum of  $p_T$  of associated tracks: from about 10  $\mu\text{m}$  for small sum of  $p_T$  to  $\sim 1 \mu\text{m}$  for large sum of  $p_T$ .

### 3.2.3 Muon reconstruction and identification

There are two muon objects that are reconstructed first: *tracker track* (reconstructed in the inner tracker) and *standalone-muon track* (reconstructed in muon system) [125]. This corresponds to 2 approaches:

- *Global muon reconstruction (outside-in)*: for a given standalone-muon track a matching tracker track is found. Then a *global-muon track* is found by fitting the combination of hits from a tracker track and a standalone-muon track.
- *Tracker muon reconstruction (inside-out)*: all tracker tracks are considered to be potential muon candidates and this hypothesis is checked by looking for compatible signatures in the muon system. The extrapolation of the tracker track to the muon system takes into account the magnetic field, scattering in the detector material and average energy losses.

Muon candidates found by both Global muon and Tracker muon approaches that share the same tracker track are merged into a single muon candidate. In order to suppress muons from decay in flight ( $K \rightarrow \mu\nu$ ,  $\pi \rightarrow \mu\nu$ ), remnants of a hadron shower penetrating through the solenoid and reaching muon system (punch-through), accidental track-to-segment matches and cosmic muons high- $p_T$  identification criteria is used. These criteria are explicitly designed for high- $p_T$  muons. They include the following requirements [126]:

- the candidate should be a Global muon and a Tracker muon
- at least one muon chamber should be included in the global muon track fit
- muon segments (track stubs built from hits in a muon chamber) in at least two stations
- the  $p_T$  relative error of the muon best track is less than 30%
- the tracker track has transverse impact parameter  $d_{xy} < 2 \text{ mm}$  w.r.t. the primary vertex (PV)
- the longitudinal distance of the tracker track w.r.t. the PV  $d_z < 5 \text{ mm}$ .
- number of pixel hits  $> 0$
- number of tracker layers with hits  $> 5$

In addition to requirements described above cut on the relative tracker isolation is applied. The relative tracker isolation is defined as:

$$relIso = \frac{\sum_i p_T^i}{p_T^{muon}} \quad (3.3)$$

where sum runs over the tracks coming from the PV located within  $\Delta R = 0.3$  of the muon. The applied cut is  $relIso < 0.1$  corresponding to the signal efficiency of 98% [126]. For the reconstruction of the muon kinematics the *Tune P* algorithm [125] is used which improves momentum resolution at high  $p_T$ .

The efficiency of the high- $p_T$  identification criteria and relative tracker isolation was studied with the Tag-and-Probe method [127]. The basic idea of the method is to select so-called tag and probe pairs. The *probe* is an object passing loose selection while the *tag* is selected with tighter requirements. Probes are the objects that are used to study the particular criteria. Then events within some window around Z boson mass [21] in dimuon mass spectrum are selected. The events are split into 2 categories: tags and probes failing the selection that is studied; tags and probes passing the selection that is studied. For each category a fit to the invariant mass distribution of the tag and probe pair is done with the background component (exponentially falling spectrum, for example) and signal component (Breit-Wigner distribution convoluted with generator level shape of the Z decay, for example). Then efficiency is defined as a ratio of passing signal events to total number of signal events. The efficiency of the high- $p_T$  identification criteria is 97%(98%) in data (Monte-Carlo modelling <sup>2</sup>) though a small degradation of efficiency at high  $p_T$  is found. The efficiency of the isolation cut is found to be 98% in both data and Monte-Carlo.

#### 3.2.4 Electron reconstruction and identification

Studies with test beams have shown that about 97% of the energy of the electron is deposited in a  $5 \times 5$  crystal array. However, as the tracker has a rather significant material budget <sup>3</sup> electrons are likely to lose energy due to bremsstrahlung and photons to convert to  $e^+e^-$  pair. Thus, in order to properly measure the energy of the electron it is important to collect radiated photons that mainly spread in  $\phi$  direction due to bending of electron in the magnetic field [128]. Reconstruction of an electron requires building its track, clustering the energy in the ECAL and associating these two inputs.

There are two approaches for building seeds for an electron candidate. The *ECAL-driven approach* uses energetic ECAL clusters ( $E_T > 4 \text{ GeV}$ ), the energy and the barycentre of which are used to determine the position of the expected hits in the tracker system. The *tracker-driven approach* is a calorimeter-unbiased seed algorithm which is based on the tracker information. It turns out to be more efficient for electrons inside jets because of the overlapping particle contributions the position of the supercluster is often biased.

All tracks from iterative tracking can be potential seeds for electrons if their  $p_T$  exceeds 2 GeV. If the radiated energy is small enough the track can be reconstructed with hits in all tracker layers. Most of electrons lose energy before entering the ECAL. In this case, either tracking uses smaller number of hits or it collects all hits but this would give large  $\chi^2$  of the fit track. The selection on number of hits and  $\chi^2$  is applied. Then a fit with Gaussian-sum filter (GSF) [129] is done which uses seeds obtained from the ECAL-driven and tracker driven procedures. It is more adapted for electrons than Kalman-filter as it allows non-Gaussian energy loss along the trajectory. Then the selection based on the multivariate discriminator is applied. It is based on

---

<sup>2</sup>further word "modelling" will be skipped

<sup>3</sup> the length is about  $1.8X_0$  at  $|\eta| \approx 1.5$

the number of hits, the  $\chi^2$  of the GSF track and its ratio to that of the KF track, the energy loss along the GSF track and the distance between the extrapolated track to the inner ECAL surface and the closest ECAL cluster.

Clustering of the energy deposited in the calorimeter starts with the *seed* generation in which cells with maximum energy with respect to either 4 or 8 closest cells are selected. Then *topological clusters* are formed from seeds by aggregating cells with at least a corner in common with a cell already in the cluster and with an energy excess of a certain threshold. In order to get calorimeter clusters from a topological cluster a fit is performed based on the assumption that energy deposits in the  $M$  individual cells arise from  $N$  Gaussian energy deposits. Clusters from all calorimeter systems are formed in this way. Clusters are then grouped into superclusters (SC). For the reconstruction of electrons so-called *mustache superclusters* are built. The idea beyond this is that photons radiated by electrons can be converted to  $e^+e^-$  pair which has a helix trajectory. This was different in Run I when a  $\eta - \phi$ -box was used [128]. The "mustache" has shown improvements with respect to the "box-superclustering" algorithm in terms of energy resolution as a function of vertex multiplicity and supercluster  $\eta$  [130].

Blocks that contain ECAL clusters, mustache superclusters and GSF-tracks are linked to build up electron candidates. A loose set of requirements is applied to remove pion tracks from the candidate by requiring the ratios of the associated ECAL energy to the track momentum and the associated HCAL energy to the associated ECAL energy. Only candidates that have a linked GSF-track are labelled as electrons. The momentum is estimated by combining information from the ECAL and the tracker:  $p_{comb} = wp_{track} + (1 - w)E_{SC}$ . The regression technique is used to determine  $w$ . At low energies a determination of electron momentum is mostly dominated by the tracker information ( $w \approx 1$ ) and at high energies by the energy deposits measured in the ECAL ( $w \approx 0$ ). Additional identification criteria is applied on top of the selection used in the reconstruction and is explicitly designed to identify high-energy electrons (so-called HEEP (high energy electron pairs) ID) which is described in [131]. The applied selections are summarized in Table 3.1.

The variables defined as :

- $E_T$  is a transverse energy of the electron.
- $\eta_{SC} - \eta$  of the supercluster.
- `isEcalDriven` – means that the electron candidate is reconstructed with the ECAL-driven approach.
- $\Delta\eta_{in}$  ( $\Delta\phi_{in}$ ) – difference between  $\eta$  ( $\phi$ ) of the supercluster and the track  $\eta$  extrapolated from the innermost track position to the position of the closest approach to the supercluster [128].  $\Delta\eta_{in}^{seed}$  is the same as  $\Delta\eta_{in}$  but uses not the  $\eta$  of the supercluster but the  $\eta$  of the seed cluster of the supercluster [132].
- $H/E$  – ratio of the energy deposited in the HCAL to the energy of the supercluster. The HCAL energy is calculated in the first 2 layers within  $\Delta R = 0.15$  with respect to the supercluster in the ECAL. The cut on this variable is used to reduce fakes from jets though high energy electrons might penetrate through the ECAL and leave energy deposit in the HCAL.

### 3 Event simulation and reconstruction

**Table 3.1:** Definitions of HEEP ID V6.1 selections [132].

Variable	Barrel	Endcap
Acceptance selections		
$E_T$	$E_T > 35 \text{ GeV}$	$E_T > 35 \text{ GeV}$
$\eta$	$ \eta_{SC}  < 1.4442$	$1.566 <  \eta_{SC}  < 2.5$
Identification selections		
isEcalDriven	true	true
$\Delta\eta_{in}^{seed}$	$ \Delta\eta_{in}^{seed}  < 0.004$	$ \Delta\eta_{in}^{seed}  < 0.006$
$\Delta\phi_{in}$	$ \Delta\phi_{in}  < 0.06$	$ \Delta\phi_{in}  < 0.06$
$H/E$	$H/E < 1/E + 0.05$	$H/E < 5/E + 0.05$
$\sigma_{i\eta i\eta}$	-	$< 0.03$
$E_{2\times 5}/E_{5\times 5}$	$> 0.94$ or $E_{1\times 5}/E_{5\times 5} > 0.83$	-
Inner lost layer hits	lost hits $\leq 1$	lost hits $\leq 1$
Impact parameter, $d_{xy}$	$ d_{xy}  < 0.02 \text{ cm}$	$ d_{xy}  < 0.05 \text{ cm}$
Isolation selections		
<i>isolEmHadDepth1</i>	$< 2 + 0.03E_T + 0.28\rho$	$< 2.5 + 0.28\rho$ ( $E_T < 50 \text{ GeV}$ ) else: $< 2.5 + 0.03(E_T - 50 \text{ GeV}) + 0.28\rho$
<i>trackerIso</i>	$< 5 \text{ GeV}$ ( $E_T < 95 \text{ GeV}$ ) else: $< 5 + 1.5\rho$	$< 5 \text{ GeV}$ ( $E_T < 100 \text{ GeV}$ ) else: $< 5 + 0.5\rho$

- $\sigma_{i\eta i\eta}^2 = \frac{\sum_i (\eta_i - \bar{\eta})^2 w_i}{\sum_i w_i} \cdot \Delta\eta_{\text{crys}}^2$ . The sum runs over the crystals in the  $5 \times 5$  matrix centred around the crystal containing the largest energy in the supercluster;  $w_i = \max(0.0, 4.7 + \ln(E_i/E_{5\times 5}))$  where  $E_i$  is the energy deposited in the  $i$ -th crystal and  $E_{5\times 5}$  is the energy deposited in the full  $5 \times 5$  matrix;  $\bar{\eta} = \frac{\sum_i \eta_i E_i}{\sum_i E_i}$ ;  $\Delta\eta_{\text{crys}}$  is the crystal size in  $\eta$ -space which 0.01745 (0.0447) in the barrel (endcaps). This variable describes the width of the shower surrounding the electron. Prompt electron has much narrower distribution in  $\sigma_{i\eta i\eta}$  than non-prompt electron [128].
- $E_{1\times 5}, E_{2\times 5}, E_{5\times 5}$  – the energy deposited in  $1 \times 5, 2 \times 5$  and  $5 \times 5$  crystal arrays around the seed crystal.
- lost hits – number of tracker layers where no hits are found. The cut on this variable is used to suppress background from photon conversions in the tracker material.
- *isolEmHadDepth1* – isolation based on the ECAL clusters and the HCAL clusters in the first layer with  $\Delta R = 0.3$ .
- *trackerIso* – sum of  $p_T$  of tracks within  $\Delta R = 0.3$  of the electron.
- $\rho$  – the average pileup energy density in the event.

### 3.2.5 Jet reconstruction

As it was already mentioned in Section 1.1.1 jets but not quarks are observed in the detector. Jets consist of particles (coming from the hadronization process) collimated along the direction of the original quark or gluon. In order to reconstruct a four-vector of the jet clustering algorithms are used. The collection of reconstructed PF candidates where charged particles coming from non-primary vertex are removed<sup>4</sup> is used as input for the clustering algorithm. There are two classes of jet clustering algorithms: cone type (iterative cone algorithm [133], e.g) and sequential recombination algorithms ( $k_t$  [134], anti- $k_t$  [135], Cambridge-Aachen [136] algorithms). Sequential recombination algorithms define a distance  $d_{ij}$  between two particles  $i$  and  $j$  and distance between particle  $i$  and the beam pipe  $d_{iB}$ . The typical definition is:

$$d_{ij} = \min(k_{Ti}^{2p}, k_{Tj}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \quad (3.4)$$

$$d_{iB} = k_{Ti}^{2p}$$

where  $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$  and  $k_{Ti}$ ,  $y_i$ ,  $\phi_i$  are the transverse momentum, rapidity, and azimuth angle of the particle  $i$ . The case  $p = 1$  corresponds to  $k_t$ -algorithm,  $p = -1$  to anti- $k_t$ -algorithm and  $p = 0$  to Cambridge-Aachen algorithm. If  $d_{ij}$  is the minimum distance then particles  $i$  and  $j$  are combined into one particle (in CMS this is done by summing up four-vectors). If  $d_{iB}$  is the minimum then  $i$  is labelled as a jet and is removed from list of particles [133].

Jet algorithms are required to be *infrared* and *collinear safe*. This means that if final state particles are modified by a soft emission or a collinear splitting doesn't change the configuration of jets in the final state doesn't change. Sequential algorithms are infrared and collinear safe [133].

In the CMS anti- $k_t$  (AK) algorithm is used for the jet clustering. Two jet collections are used with  $R = 0.4$  (AK4 jets) and  $R = 0.8$  (AK8 jets).

#### Jet energy corrections

The detector response to particles is not linear and due to that it is not trivial to translate the measured jet energy to the true particle level. Jet energy corrections (JEC) take care of mapping the measured jet energy to the particle-level energy. CMS uses a factorized approach to the determination of jet energy corrections where each correction corresponds to a specific effect [137]. Corrections are applied sequentially and essentially scale  $p_T$  of the jet. The following JECs are applied in this thesis [138]:

- *L1 Pileup*: this correction removes energy from the pileup. The correction is determined from simulated dijet processes with and without pileup. The pileup energy that is subtracted is about 0.3 GeV per number of additional pileup interactions [139].
- *L2L3 MC-truth*: corrections are determined by comparing  $p_T$  of the reconstructed jet to the particle-level  $p_T$ . The corrections are determined as a function of  $p_T$  and  $\eta$  to make the response flat in these variables and can reach up to  $\approx 10\%$  [139].

<sup>4</sup>this will further referred as *charged hadron subtraction* (CHS).

### 3 Event simulation and reconstruction

---

- *L2L3 residuals*: corrections cover small remaining differences between data and MC (about few % [139]). These corrections are applied on data only.

#### Jet energy resolution corrections

Measurements show that the jet energy resolution (JER) in data is worse than in the simulation and  $p_T$  of jets in MC need to be smeared in order to describe the data. A random smearing of  $p_T$  of AK8 jets is performed. Basically, for each of AK8 jets in the event we randomly smear the  $p_T$  with a Gaussian with mean value of jet  $p_T$  and the width of  $\sqrt{c^2 - 1} \cdot \sigma_{MC}$ , where  $c$  is a ratio of resolutions in data and MC (see Table 3.2),  $\sigma_{MC}$  is the jet  $p_T$  resolution in Monte-Carlo. The four-vector of the jet is then scaled with the factor:  $p_{T,\text{smeared}}/p_T$ .

**Table 3.2:** Jet energy resolution scale factors [140].

$\eta$ -range	$\sigma(\text{data})/\sigma(\text{MC})$	uncertainty
$ \eta  < 0.5$	1.095	0.018
$0.5 <  \eta  < 0.8$	1.12	0.028
$0.8 <  \eta  < 1.1$	1.097	0.017
$1.1 <  \eta  < 1.3$	1.103	0.033
$1.3 <  \eta  < 1.7$	1.118	0.014
$1.7 <  \eta  < 1.9$	1.1	0.033
$1.9 <  \eta  < 2.1$	1.162	0.044
$2.1 <  \eta  < 2.3$	1.16	0.048
$2.3 <  \eta  < 2.5$	1.161	0.060

#### Jet identification

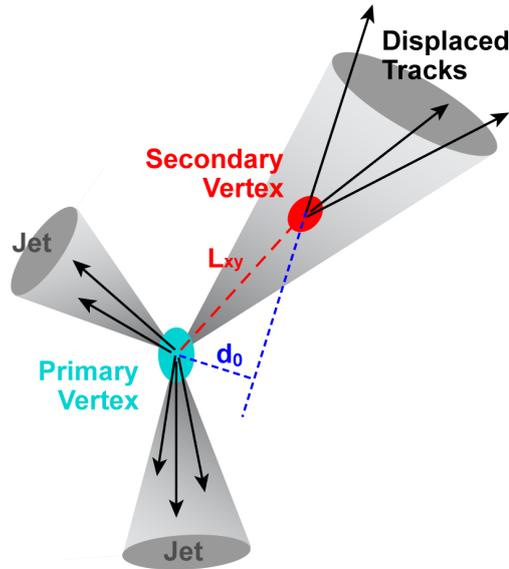
In order to reject fake, badly reconstructed and noise jets while keeping 98-99% of the real jets the so called PFJetID is applied [141]. The set of requirements is summarized in Table 3.3. This was optimized for jets with  $|\eta| < 2.7$  which includes the range used in this thesis.

**Table 3.3:** PFJetID requirements [142].

Neutral hadron fraction	< 0.99
Neutral electromagnetic fraction	< 0.99
Number of constituents	> 1
Muon fraction	< 0.8
if $ \eta  < 2.4$	
Charged Hadron Fraction	> 0
Charged multiplicity	> 0
Charged electromagnetic fraction	< 0.99

### 3.2.6 Identification of b-jets

In this thesis Identification of jets coming from b-quarks (b-tagging) is used to suppress backgrounds that contain b-quarks. A jet coming from a b-quark contains B-mesons which have a typical decay length of about  $mm$  [21]. This means that a jet contains particles coming not from a primary vertex but from a secondary vertex which is illustrated in Figure 3.2. Identification of such particles is possible due to the pixel resolution as it was already mentioned in Section 2.2.2.



**Figure 3.2:** The illustration of the event containing a jet originating from a b-quark. The jet contains B-meson which decays at the secondary vertex. The secondary vertex is displaced from the primary vertex by the decay length of B-meson  $L_{xy}$  [143].

The *Combined Secondary Vertex* (CSV) v2 algorithm is used for b-tagging in this thesis. It is a modification of the CSV algorithm used in Run I [144]. One of the differences is the new secondary vertex reconstruction algorithm which provides improvement of about 1-2% in b-tagging efficiency [145]. The algorithm combines the information of displaced tracks and secondary vertices using multivariate technique. The "tight" operating point is used which corresponds to the b-tagging efficiency of around 49% and 0.1% misidentification efficiency for light-flavour jets as measured using  $t\bar{t}$  dilepton events [145].

### 3.2.7 Reconstruction of missing transverse energy

The presence of particles that don't interact with the detector material (neutrinos or in case of BSM searches dark matter candidates, for example) is quantified as missing momentum or usually as missing transverse energy ( $\cancel{E}_T$ ). The precise measurement of  $\cancel{E}_T$  is important in the context of this thesis because the analysis described here exploits the leptonic decay of the W-

boson. The raw  $\cancel{E}_T$  is defined as a negative vectorial sum of transverse momenta of all identified particles:

$$\vec{\cancel{E}}_T^{\text{raw}} = - \sum_{i=1}^{N_{\text{particles}}} \vec{p}_{T,i} \quad (3.5)$$

where the sum runs over all particles in the event. Minimum energy thresholds in the calorimeters, inefficiencies in the tracker, non-linearity of the response in the calorimeter system can lead to underestimated or overestimated values of  $\cancel{E}_T$ . In order to reduce this bias in  $\cancel{E}_T$  measurement jet energy corrections are propagated into  $\cancel{E}_T$  [146]:

$$\vec{\cancel{E}}_T^{\text{corr}} = \vec{\cancel{E}}_T^{\text{raw}} - \sum_{j=1}^{N_{\text{jets}}} (\vec{p}_{T,j}^{\text{corr}} - \vec{p}_{T,j}) \quad (3.6)$$

where  $\vec{p}_{T,j}$  is the raw transverse momentum of the jet and  $\vec{p}_{T,j}^{\text{corr}}$  is the transverse momentum of the jet with jet energy corrections applied.

## 4 Grooming algorithms

Grooming, the systematic removal of a subset of the jet constituents, is intended to remove soft and wide-angle radiation from the jet. It is typically used to reduce the overall jet mass of QCD (quark- and gluon-initiated) jets while preserving the larger jet mass for jets originating from heavy particles such as the top quark and W/Z/H bosons or new physics. Additionally, grooming, in the presence of pileup, can be used to reduce the dependence of jet mass on pileup activity.

In this chapter a study of grooming algorithms is presented. The study was done in the context of the preparation towards Run II of the LHC. The focus is given to the pileup mitigation effects and to jet mass resolutions which is important for the identification of W(Z)-jets. Results of the studies can be found in [147].

### 4.1 Analysis strategy

In the coming years the instantaneous luminosity at the LHC (followed by High-Luminosity LHC, HL-LHC [148]) will be increasing. It is expected that the number of pileup interactions will be increased up to  $\sim 100$ . This requires not only technological upgrades of the experiments [149] but also development and validation of algorithms capable to work under such challenging conditions. The goal of this study is to understand the performance of grooming algorithms and their ability to mitigate pileup effects.

In the study RS [14, 15] gravitons with mass of 1 TeV decaying to a pair of W-bosons used as signal and multijet QCD events are used as background. The simulation is done for  $pp$  collisions at  $\sqrt{s} = 13$  TeV in presence of high pileup (mean number of pileup interactions is 40). Anti- $k_t$  jets with  $R = 0.8$  were used. Only hadronically decaying W-bosons are selected. For the signal a jet is required to be matched with a generator-level W-boson ( $\Delta R < 0.3$ ). Signal and background events are considered if the leading jet (in  $p_T$ ) has<sup>1</sup>:

- $p_T > 300$  GeV
- $|\eta| < 2.5$

The following types of jet constituents are studied:

- PF (all reconstructed PF candidates are used, no additional removal is done)
- PF + CHS
- PF + PUPPI (PileUp Per Particle Identification)

---

<sup>1</sup>JECs are applied to  $p_T$

PUPPI [150] is an algorithm developed for pileup mitigation of neutral particles. The basic idea of the algorithm is to assign a weight to every single particle. This weight corresponds to the probability of a particle to come from a primary vertex. Based on the tracking information, this is known for charged particles and thus the weight assigned to the charged particle is either 1 or 0. For neutral particles this is not the case and the weight is between 0 and 1. It is computed based on the information from the surrounding particles. In the end the weight is used to rescale the particle's four-vector.

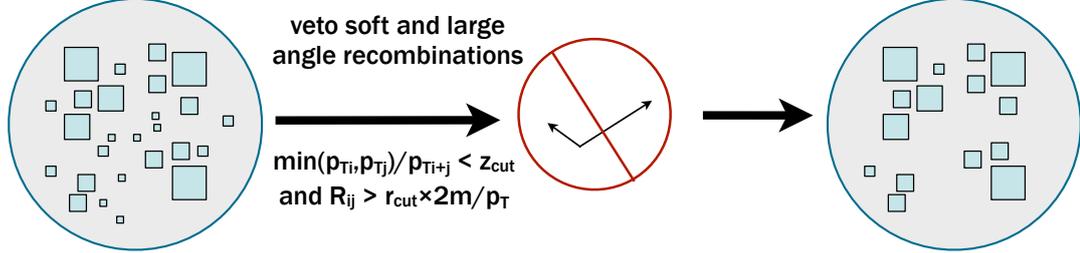
#### 4.1.1 Grooming algorithms

The pruning algorithm [151, 152] reclusters the constituents of the jet through the Cambridge-Aachen (CA) algorithm, using the same distance parameter. At each step in the clustering algorithm, the softer of the two objects  $i$  and  $j$  to be merged is removed when the following conditions are met:

$$z_{ij} = \frac{\min(p_{T,i}, p_{T,j})}{p_{T,i+j}} < z_{\text{cut}} \quad (4.1)$$

$$\Delta R_{ij} = \frac{2 \times r_{\text{cut}} \times m_J}{p_T} > D_{\text{cut}} \quad (4.2)$$

where  $m_J$  and  $p_T$  are the mass and transverse momentum of the originally-clustered jet,  $p_{T,i}$  ( $p_{T,j}$ ) is a transverse momentum of the object  $i$  ( $j$ ),  $p_{T,i+j}$  is a transverse momentum of the object obtained from the combination of  $i$  and  $j$ ,  $z_{\text{cut}}$  and  $r_{\text{cut}}$  are parameters of the algorithm. The pruning algorithm is illustrated in Figure 4.1.

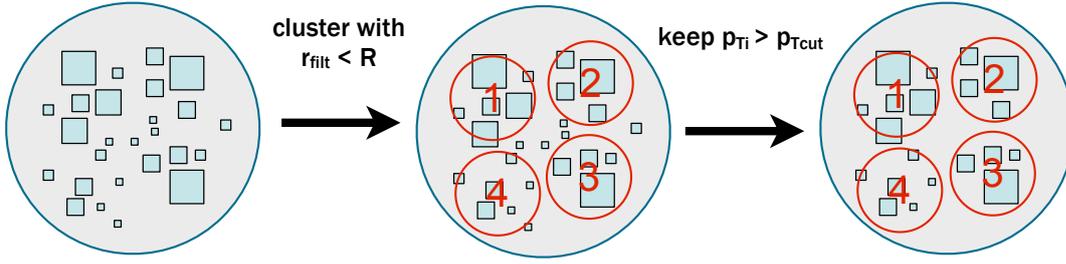


**Figure 4.1:** Illustration for the pruning algorithm (adopted from M.Malberty [153]).

Trimming [154] ignores particles within a jet that fall below a dynamic threshold in  $p_T$ . It reclusters the constituents of the jet using the  $k_t$  algorithm with a radius  $r_{\text{sub}}$ , accepting only the subjets that have  $p_{T\text{sub}} > p_{T\text{frac}} \lambda_{\text{hard}}$ , where  $p_{T\text{frac}}$  is a dimensionless cutoff parameter, and  $\lambda_{\text{hard}}$  is some hard QCD scale chosen to be equal to the  $p_T$  of the original jet. The trimming algorithm is illustrated in Figure 4.2.

Soft drop [155] declusters the jet by recursively removing soft and wide angle radiation from the jet. The jet is reclustered using the CA algorithm. Then the jet is declustered and at each step, subjets  $j_1$  and  $j_2$  are defined and the following condition is checked:

$$\frac{\min(p_{T,j_1}, p_{T,j_2})}{p_{T,j_1} + p_{T,j_2}} > z_{\text{cut}} \times \left( \frac{\Delta R_{12}}{R_0} \right)^\beta \quad (4.3)$$



**Figure 4.2:** Illustration for the trimming algorithm,  $p_{Tcut} = p_{Tfrac}\lambda_{hard}$  (adopted from M.Malberty [153]).

where the algorithm parameters are  $z_{cut}$  and  $\beta$ . If the condition is met, the declustering continues, otherwise only the leading  $p_T$  subject is kept.

For each of the grooming algorithms several parameters are explored as shown in Table 4.1.

**Table 4.1:** Summary of the parameters of grooming algorithms.

grooming algorithm	parameters
Pruning	$z_{cut} = 0.1, r_{cut} = 0.5$
	$z_{cut} = 0.05, r_{cut} = 0.5$
	$z_{cut} = 0.1, r_{cut} = 0.75$
	$z_{cut} = 0.05, r_{cut} = 0.75$
Trimming	$r_{sub} = 0.2, p_{Tfrac} = 0.05$
	$r_{sub} = 0.2, p_{Tfrac} = 0.03$
	$r_{sub} = 0.1, p_{Tfrac} = 0.03$
	$r_{sub} = 0.3, p_{Tfrac} = 0.03$
Soft drop	$z_{cut} = 0.1, \beta = 0$
	$z_{cut} = 0.1, \beta = 1$
	$z_{cut} = 0.1, \beta = 2$

## 4.2 Mass distributions

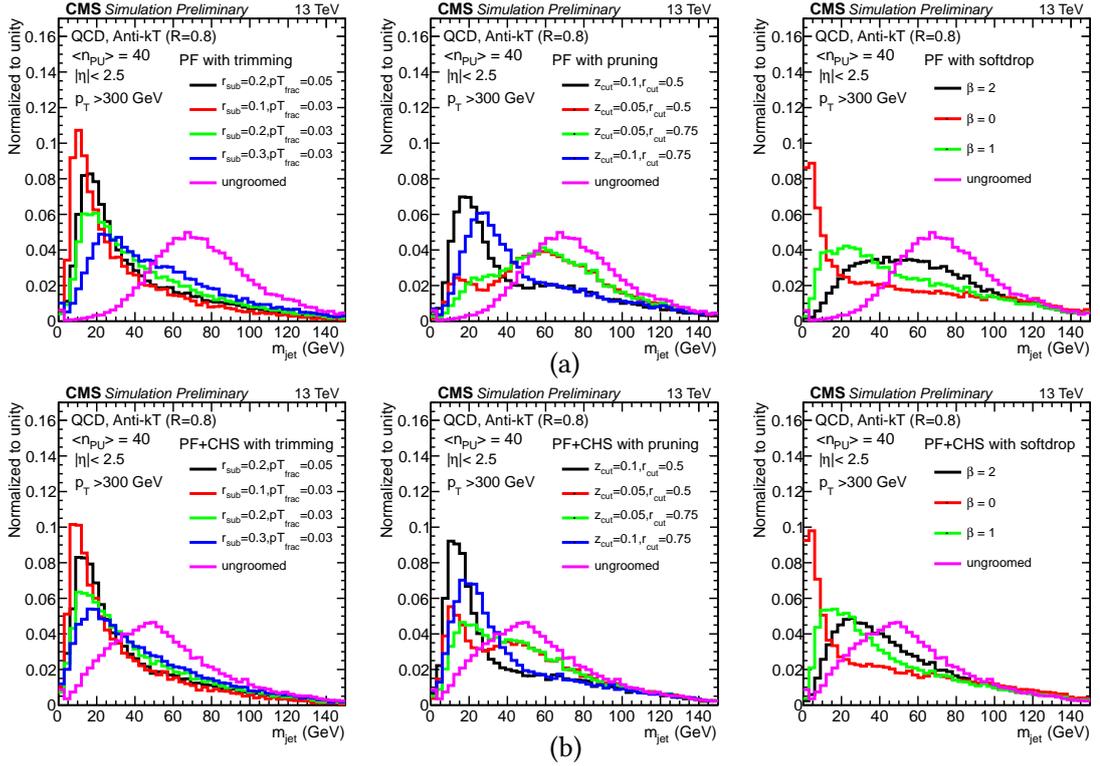
Mass distributions for grooming algorithms are studied. Groomed masses are corrected with a so-called four-vector "safe" subtraction [156]. It is an extension of the area subtraction [157]:

$$p_{sub}^\mu = p^\mu - \rho A^\mu - \rho_m A_m^\mu \quad (4.4)$$

where  $\rho$  is the average pileup density and  $\rho_m$  takes into account non-zero hadron masses. Mass distributions are taken in the range [0., 200.] GeV. This is done for PF and PF+CHS jets. Since PUPPI is supposed to effectively remove particles from non-primary vertices "safe" subtraction is not applied for PF+PUPPI. It is important to note, that JEC are **not** applied to the jet mass. Mass distributions for different grooming algorithms are shown in Figure 4.3 for QCD-jets

## 4 Grooming algorithms

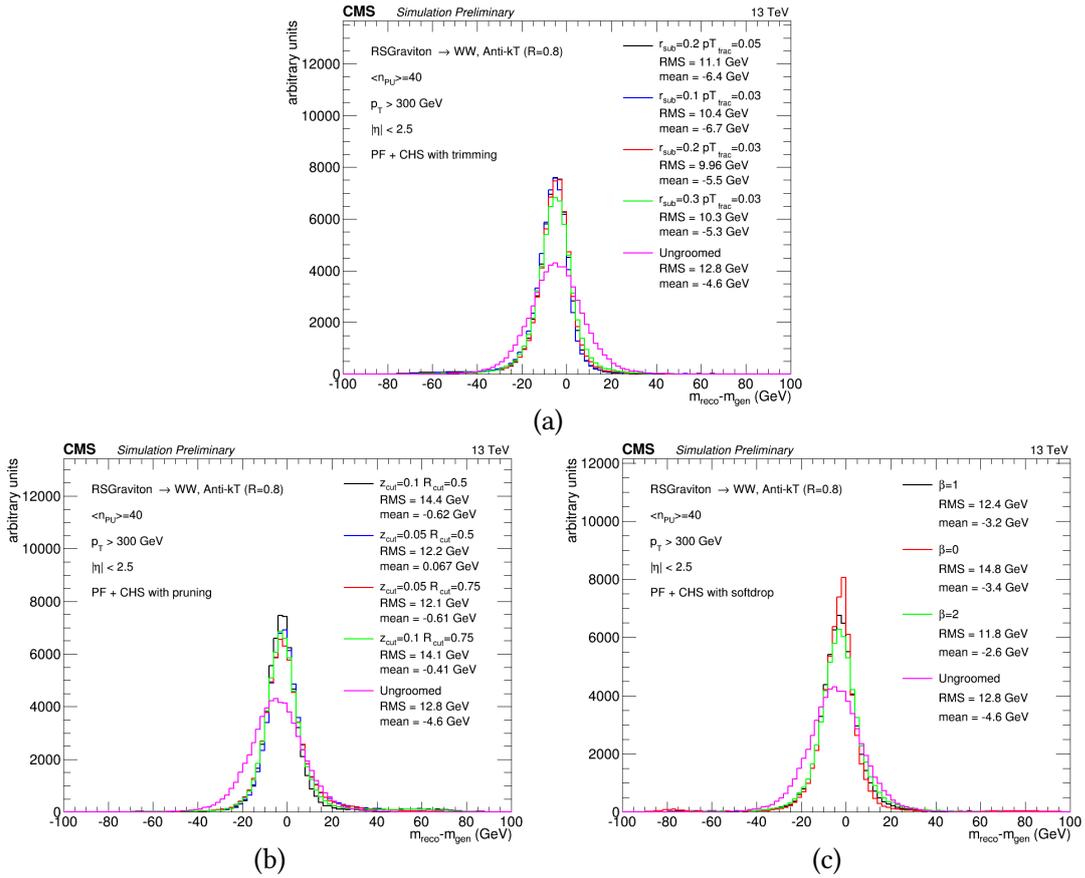
and for W-jets in Appendix A in figures A.1, A.2, A.3. For QCD-jets grooming shifts the mass distribution towards 0 GeV and for W-jets towards 80 GeV. Though for some parameters the W-jet mass is reduced towards 0 GeV which indicates some aggressiveness of the algorithm for given parameter values. Also the raw mass (no "safe" subtraction, no grooming is applied, denoted as  $m_{\text{raw}}$ ) is compared with the ungroomed mass ("safe" subtraction but no grooming is applied). In Figure A.3 (top-left) the mass distribution with and without "safe" area subtraction are shown. Both distributions are basically identical which is indicating that PUPPI removes most of the pileup contribution from the inputs to the jet clustering [147].



**Figure 4.3:** Jet mass distributions for various grooming algorithms for QCD-jets. PF (a) and PF+CHS (b) are used as input for the jet clustering [147].

### 4.3 Mass response distributions

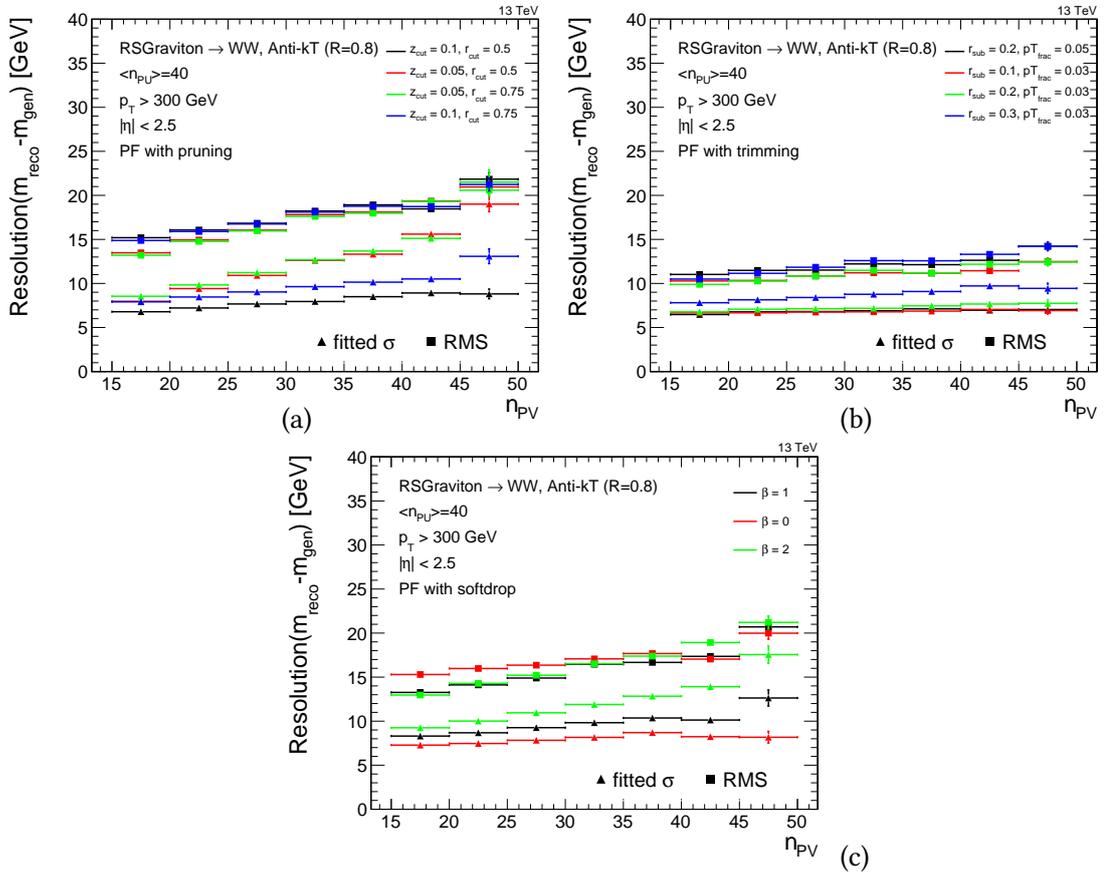
Another metric that is used to validate the performance of grooming algorithms is the response (difference between mass of the reconstructed and generated jet):  $m_{\text{reco}} - m_{\text{gen}}$ . Distributions are demonstrated on figures B.1, B.2. The mass response distributions are taken in range [-100., 100.] GeV.



**Figure 4.4:** Mass response distributions for W-jets for different grooming algorithms: trimming (a), pruning (b), softdrop (c). PF+CHS is used as input for the jet clustering [147].

## 4.4 Pileup dependence

To study the stability of grooming algorithms against pileup several metrics are checked: the jet mass resolution, the jet mass response resolution, mean jet mass and mean jet mass response as a function of a number of reconstructed vertices. For W-jets results from the fit with Gaussian and mean (or root mean square, RMS) are shown in order to evaluate effect of the tails of the distributions. As can be seen in Figure 4.5 for some parameters jet mass response resolution becomes rather stable as a function of a number of reconstructed vertices though no pileup mitigation technique is applied to the input collection used for the jet clustering. A complete set of plots can be found in Appendix C. Overall, grooming algorithms improve stability with respect to pileup, and together with CHS pileup dependence is significantly reduced. PUPPI is quite stable with respect to pileup as well.



**Figure 4.5:** Jet mass response resolution as a function of number of reconstructed vertices for pruning (a), trimming (b), softdrop (c). PF is used as input for the jet clustering.

## 4.5 N-subjettiness

The important variable that is used in this thesis is so called N-subjettiness that was introduced in [158]. In a given jet one identifies  $N$  subjects and then  $\tau_N$  can be calculated as:

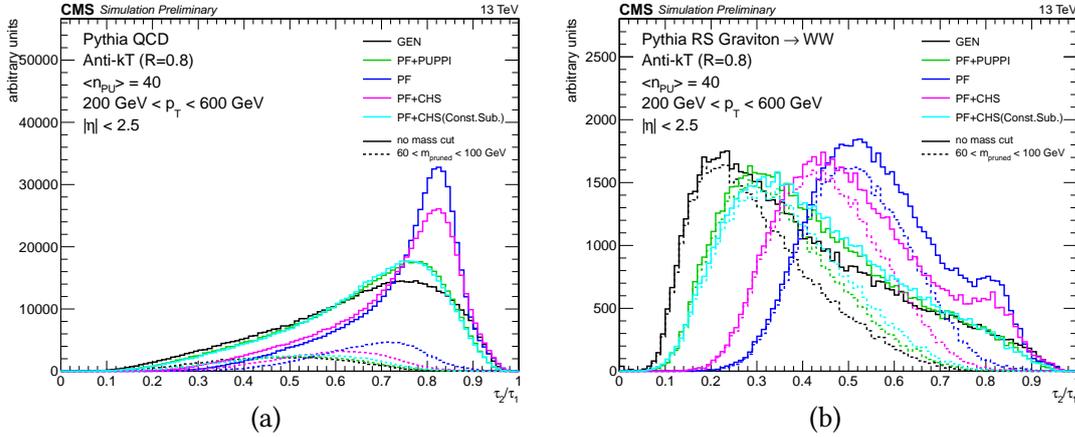
$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k}\} \quad (4.5)$$

where  $k$  runs over jet constituents,  $p_{T,k}$  is their transverse momenta  $\Delta R_{J,k} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$  between a subjet  $J$  and a constituent  $k$ ;  $d_0$  is defined as:

$$d_0 = \sum_k p_{T,k} R_0 \quad (4.6)$$

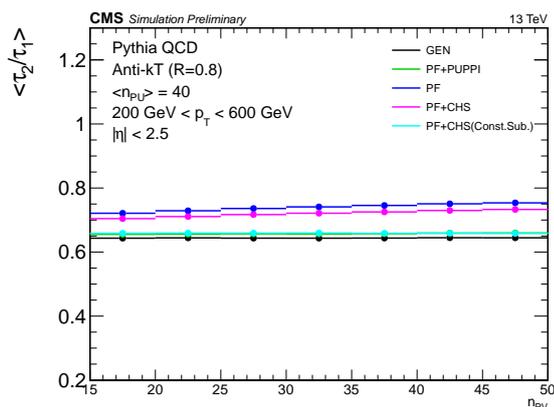
where  $R_0$  is the radius used for the jet clustering.  $\tau_N$  basically quantifies to what degree it can be regarded that a jet is composed of  $N$  subjects. Jets with  $\tau_N \approx 0$  have radiation aligned with the directions of subjects and thus are composed of  $N$  (or fewer) subjects. Jets with  $\tau_N \gg 0$  have a decent fraction of energy distributed away from the directions of subjects and thus are composed of at least  $N + 1$  subjects. It was found in [158] that to identify boost W-jets the ratio  $\tau_{21} = \tau_2/\tau_1$  is a better discriminating variable than  $\tau_2$  or  $\tau_1$ .

The  $\tau_{21}$  distribution is shown in Figure 4.6 for QCD- and W-jets. PUPPI and constituent subtraction (another pileup mitigation tool that was studied in [147] but is out of the scope of this thesis) are bringing  $\tau_{21}$  distribution closer to the generator level distribution. The distribution is also shown after requiring pruned jet mass ( $z_{\text{cut}} = 0.1$ ,  $r_{\text{cut}} = 0.5$ ) to be between 60 GeV and 100 GeV. This cut removes contributions from unmerged W-jets. It is also illustrated in Figure 4.7 that the pileup dependence of  $\tau_{21}$  is mitigated by PUPPI and constituent subtraction.



**Figure 4.6:** Leading jet  $\tau_{21}$  distribution for QCD-jets (a) and W-jets (b) [147].

In Figure 4.8 a comparison of jet mass resolution and jet mass response resolution for different grooming algorithms is shown for PF, PF+CHS, PF+PUPPI. The resolution is characterized in terms of  $\sigma$  (from a Gaussian fit of W mass peak) and in terms of RMS. The  $\sigma$  is obtained by

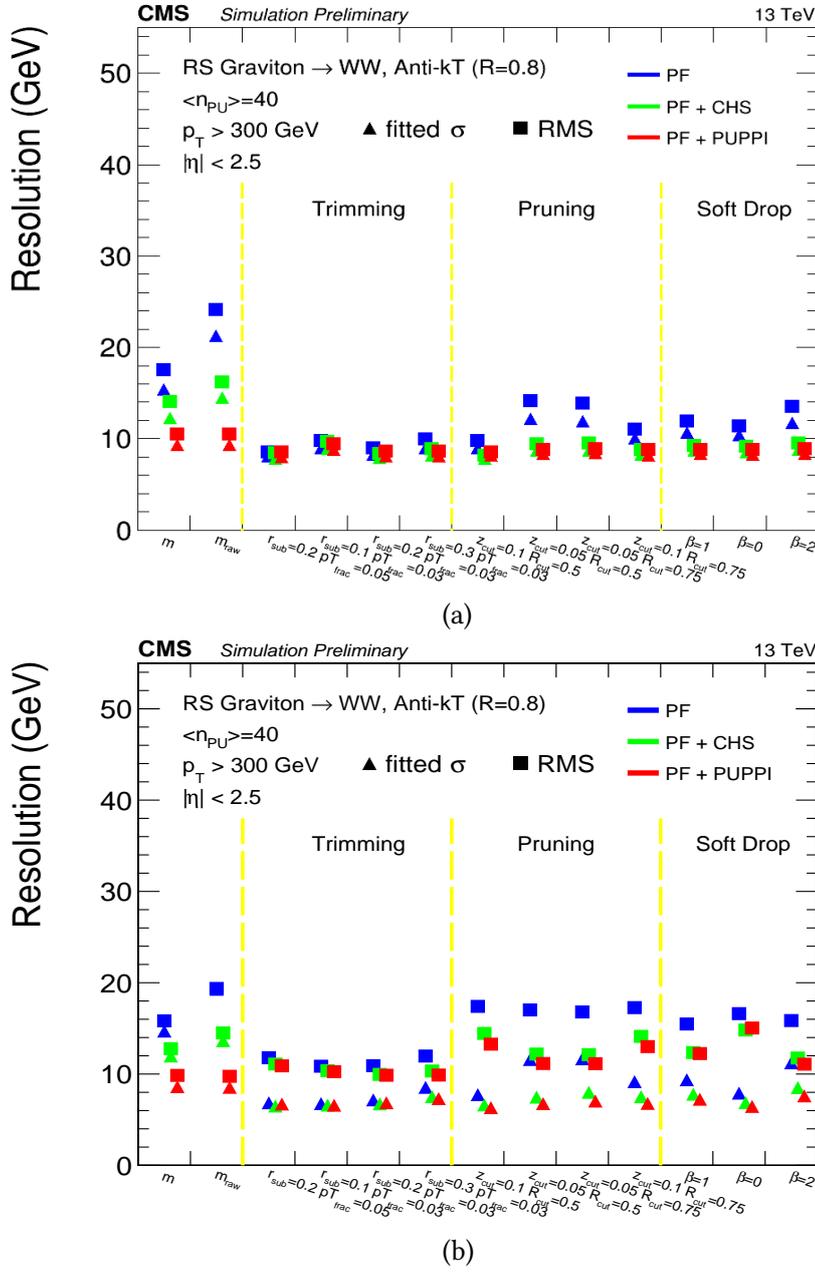


**Figure 4.7:** Average  $\tau_{21}$  as a function of number of reconstructed vertices [147].

doing a fit in a window  $\pm \delta m$  where  $\delta m$  is equal to RMS on the full mass range ( $[0, 200]$  GeV for the jet mass and  $[-100, 100]$  GeV for the jet mass response). Then RMS is estimated as truncated RMS in the range  $\mu \pm \sigma$  where  $\mu$  is the fitted mean.

## 4.6 Conclusion

A number of jet grooming algorithms was investigated considering few points in the parameter space. The effect on jet mass and jet substructure observables was studied. It was found that grooming together with CHS or PUPPI significantly reduces pileup dependence of the jet mass. Overall trimming shows slightly better performance in terms of jet mass resolution than others. However, softdrop is preferred from the theory point of view [159]. In this thesis pruning is used because it was the only algorithm supported by the CMS experiment for 2015 data. For example, softdrop together with PUPPI is used by resonance searches [160] with 2016 data.



**Figure 4.8:** Summary plots for the jet mass resolution (a) and the jet mass response resolution (b) [147] for different parameters of grooming algorithms.  $m$  is the ungroomed jet mass with four vector-safe subtraction and  $m_{raw}$  is the ungroomed and uncorrected jet mass. The resolution is evaluated with RMS (squares) and fitted  $\sigma$  (triangles).



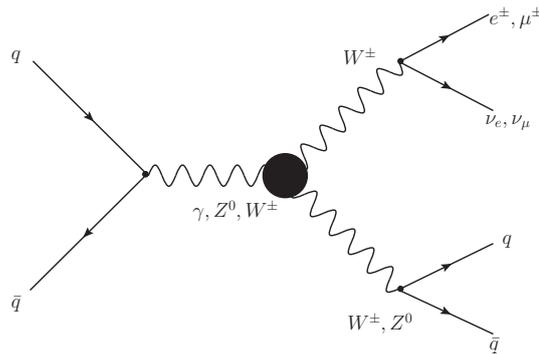
## 5 Process modelling and event selection

In this chapter modelling of signal and background processes is discussed. The description of the event selection and corrections applied to the simulation is given. An overview of systematic uncertainties affecting normalization of background processes derived from the simulation is provided.

### 5.1 Signal and background modelling

#### 5.1.1 Signal process

As it was already mentioned the analysis is done in the semileptonic channel. The LO Feynman-diagram is shown in Figure 5.1. The signal event should contain back-to-back W bosons (or W



**Figure 5.1:** The LO Feynman-diagram for the signal in the semileptonic decay. One W boson decays to a lepton and a neutrino and another W (Z) boson decays to quarks.

and Z boson), one decaying leptonically and one decaying hadronically. Thus a signal event has a following signature: a missing transverse energy  $\cancel{E}_T$ , a charged lepton (an electron or a muon), 2 jets or 1 jet with large  $R$ . However, as it was shown in reference [20] effects from are mostly prominent at high W(Z) boson  $p_T$ . Hadronic decays of vector bosons at high  $p_T$  are dominated by boosted topology when 2 quarks are close to each other and decay products are much more likely to be found in a single large- $R$  jet ("fat jet") than in 2 jets with smaller radius [161]. Because of this hadronically decaying W(Z) boson is reconstructed as a single large- $R$  jet. Also additional jets might be present due to hard gluon emission, for example.

The signal process was generated with MADGRAPH5\_aMC@NLO v2.2.3 [116] in LO using the "EWDim6" model, which implements the EFT from reference [40]. Several points in aTGC-space are generated which are listed in Table 5.1.

**Table 5.1:** Values of anomalous triple gauge couplings used for the generation of signal events.

	$c_{WWW}/\Lambda^2$ [TeV <sup>-2</sup> ]	$c_W/\Lambda^2$ [TeV <sup>-2</sup> ]	$c_B/\Lambda^2$ [TeV <sup>-2</sup> ]
0	12.0	20.0	60.0
1	12.0	0.0	0.0
2	-12.0	0.0	0.0
3	0.0	20.0	0.0
4	0.0	-20.0	0.0
5	0.0	0.0	60.0
6	0.0	0.0	-60.0
7	-12.0	-20.0	-60.0
8	0.0	0.0	0.0

### 5.1.2 Background processes

The main backgrounds are  $t\bar{t}$  and  $W$ +jets processes.  $t\bar{t}$  events are generated at NLO with POWHEG [114].  $t\bar{t}$  process in case of the semileptonic decay has same particles in the final state as the signal but has additional b-jets (the LO Feynman diagram shown in Figure 5.2 a). A significant fraction of  $t\bar{t}$  events can be suppressed by vetoing b-tagged jets, however as b-tagging is not 100% efficient, a significant number of  $t\bar{t}$  events is still accepted.

$W$ +jets production is a production of  $W$  boson associated with jets (the LO Feynman diagram shown in Figure 5.2 b). It becomes a significant background when  $W$  boson decays leptonically and one of the jets is hard enough so it mimics hadronically decaying  $W$  or  $Z$  boson.  $W$ +jets events are generated by MADGRAPH at LO with different jet multiplicities. The MLM matching scheme is used. The sample generated in several  $H_T$  bins is used in order to profit from higher statistics with hard jets. The usage of the LO simulation is acceptable because the analysis explicitly looks for the central jets and thus is not sensitive to the modelling of high jet multiplicities.  $W$ +jets is corrected further with data-driven technique as discussed below.  $Z$ +jets background is suppressed by vetoing the second lepton with efficiency of almost 100% and is not considered in the analysis.

Standard Model  $WW$  and  $WZ$  are minor backgrounds. The LO Feynman diagram for  $WW$  production is the same as shown in Figure 1.3 but with the Standard Model coupling. LO Feynman diagrams for  $WZ$  production in  $s$ - and  $t$ -channel are shown in Figure 5.3.  $WW$  events are generated with POWHEG and  $WZ$  with MADGRAPH5\_aMC@NLO [116] using FxFx-merging [117] to merge final states with different multiplicities. For  $WW$  only processes with opposite-sign  $W$  bosons are considered. Production of same-sign  $W$  bosons (vector boson scattering) has a negligibly small cross-section [162]. The production of the Higgs boson via vector boson fusion is also neglected because of the very small cross-section [163].

Another source of background is the electroweak single-top production which can occur through different mechanisms:  $t$ -channel,  $tW$ -channel and  $s$ -channel.  $tW$ -channel has the same signature as the signal (with b-quark in addition that might fall out of the acceptance of the

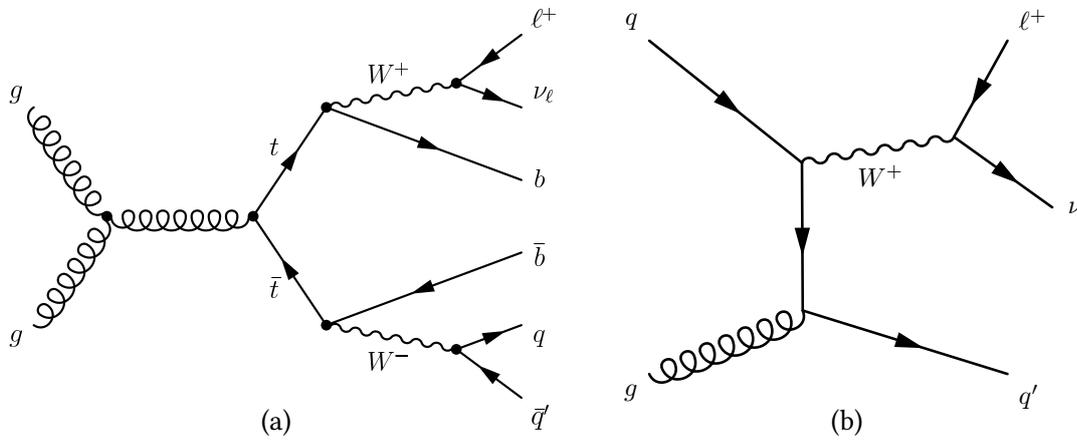


Figure 5.2: Feynman diagrams for  $t\bar{t}$  (a) and  $W^+$ jets (b).

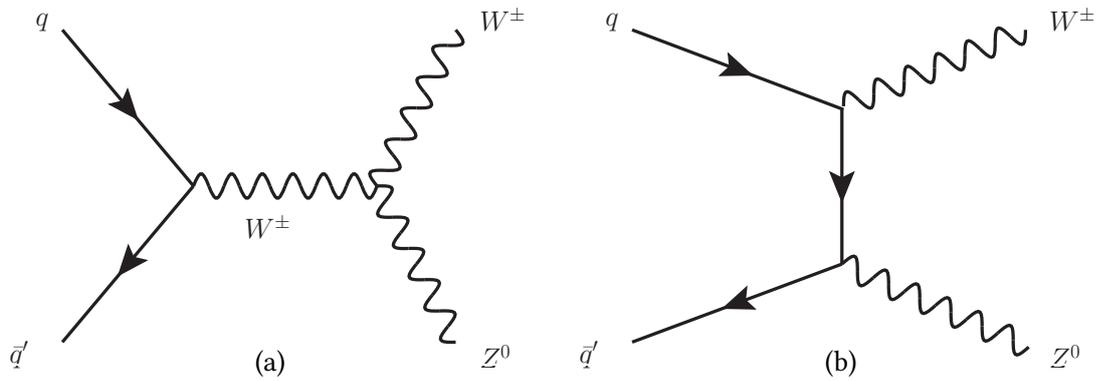
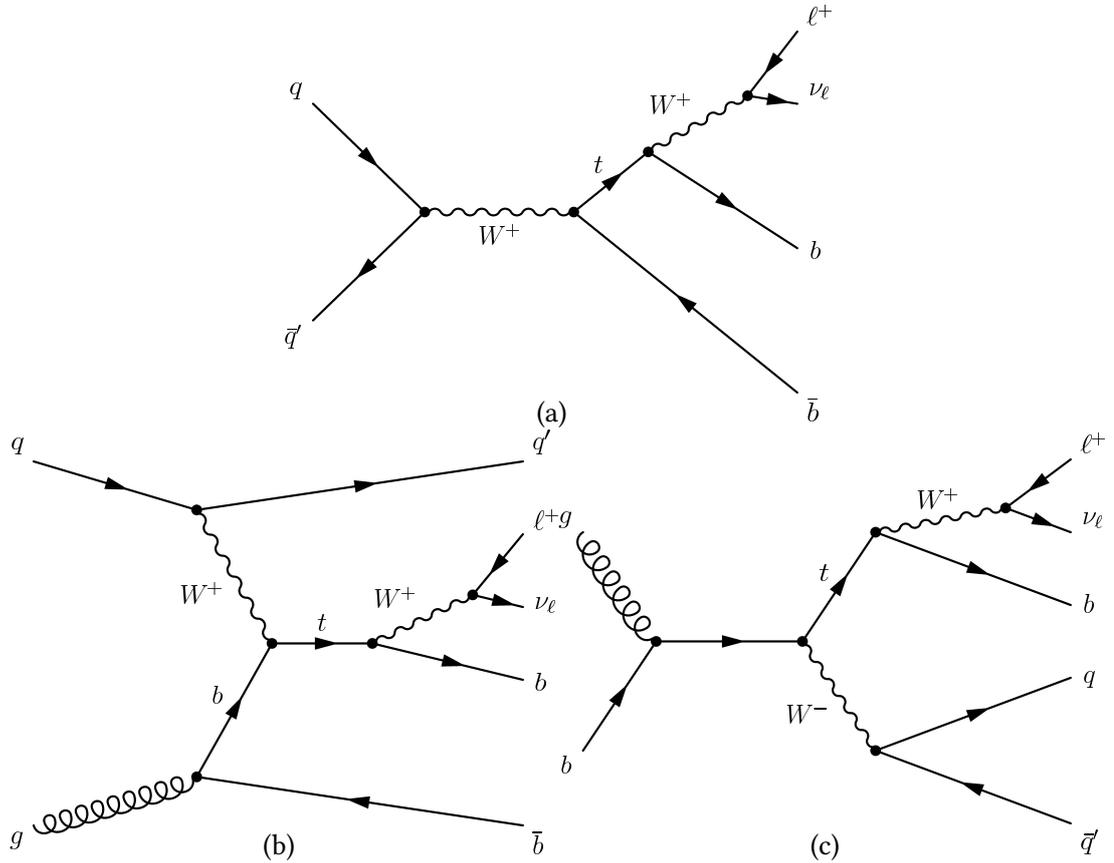


Figure 5.3: Feynman diagrams for  $WZ$  production: s-channel (a) and t-channel (b).

analysis). In  $s$ - or  $t$ -channel there is no real second  $W$  boson in the final state, however random combinations of quarks can still be misinterpreted as hadronically decaying  $W$  boson. Though contribution from  $s$ -channel should be suppressed because of the small cross-section [164]. Feynman diagrams for the single top production are shown in Figure 5.4.  $s$ - and  $t$ -channel are generated with `MADGRAPH5_aMC@NLO` and  $tW$ -channel with `POWHEG`. `PYTHIA 8.1` [111] is used for parton shower and hadronization modelling for all background processes.



**Figure 5.4:** Feynman diagrams for single top production:  $s$ -channel (a),  $t$ -channel (b) and  $tW$ -channel (c).

### 5.1.3 Simulated samples

MC samples used for modelling signal and background events are part of *RunIIIFall15MiniAODv2* campaign used in the CMS experiment for analysing data recorded in 2015. Detector response in the MC samples is simulated using a detailed description of the CMS detector, based on `GEANT 4` [165]. In Table 5.2 processes considered in the analysis are given with the corresponding cross-section and generated number of events. Most of the cross-sections are taken from [166]. For  $W$ +jets the LO cross-section is used with  $k$ -factor of 1.21 applied. The  $k$ -factor was calculated as a ratio of inclusive  $W$ +jets cross-section at NNLO accuracy to inclusive LO cross-section.

NNLO cross-section [167] was calculated with FEWZ 3.1 [168] and the LO cross-section was calculated with MADGRAPH. Most of the cross-sections are given at NNLO precision. The  $t\bar{t}$  cross-section was calculated with the Top++v2.0 program [169, 170] at NNLO+NNLL. The WZ cross-section was calculated in [171] at NNLO. The WW cross-section is taken from [172] at NNLO with  $gg \rightarrow H \rightarrow WW$  contribution subtracted. Single top  $t$ -channel and  $s$ -channel cross-sections were calculated with the Hathor v2.1 program [173, 174] at NLO. Single top  $tW$ -channel cross-section was calculated at NNLO in reference [175]. In case of the exclusive decay of W bosons the corresponding branching fractions are applied from reference [21].

**Table 5.2:** Signal and background samples.  $N_{\text{events}}$  denotes the number of generated events.

process	Decay channel	Cross section (pb)	$N_{\text{events}}$
$t\bar{t}$	–	831.76	196937036
W+jets, $H_T \in [100, 200]$	$W \rightarrow \ell\nu$	1345.0	10205377
W+jets, $H_T \in [200, 400]$	$W \rightarrow \ell\nu$	359.7 * (K=1.21)	4949568
W+jets, $H_T \in [400, 600]$	$W \rightarrow \ell\nu$	48.91* (K=1.21)	1943664
W+jets, $H_T \in [600, 800]$	$W \rightarrow \ell\nu$	12.05* (K=1.21)	3767766
W+jets, $H_T \in [800, 1200]$	$W \rightarrow \ell\nu$	5.501* (K=1.21)	1568277
W+jets, $H_T \in [1200, 2500]$	$W \rightarrow \ell\nu$	1.329 * (K=1.21)	246239
W+jets, $H_T \in [2500, +\infty]$	$W \rightarrow \ell\nu$	0.03216 * (K=1.21)	251982
WW	$WW \rightarrow \ell\nu qq$	49.997	1924400
WZ	$WZ \rightarrow \ell\nu qq$	11.46	19742520
single top, $s$ -channel	$W \rightarrow \ell\nu$	10.32*0.33	998400
single top, $t$ -channel	$W \rightarrow \ell\nu$	216.99*0.33	19938230
single top, $tW$ -channel	inclusive	71.2	1999400
WW (signal)	$WW \rightarrow \ell\nu qq$	–	1991600
WZ (signal)	$WZ \rightarrow \ell\nu qq$ or $WZ \rightarrow \ell^+ \ell^- qq$	–	1999600

#### 5.1.4 Data samples

In this thesis *SingleMuon* and *SingleElectron* datasets from proton-proton collisions with 25 ns are used. The information about datasets is given in Table 5.3. Only runs that pass certification by the Data Quality Monitoring (DQM) group are used. The integrated luminosity of these runs is  $2.3 \text{ fb}^{-1}$ .

**Table 5.3:** Data samples

dataset name	Number of events	run range
/SingleElectron/Run2015D-16Dec2015-v1/MINIAOD	134140420	256630-260627
/SingleMuon/Run2015D-16Dec2015-v1/MINIAOD	91999861	256630-260627

## 5.2 Event selection

### 5.2.1 Triggers

In the electron channel events are required to pass *HLT\_Ele27\_WPLoose\_Gsf* and in the muon channel *HLT\_Mu50*. *HLT\_Ele27\_WPLoose\_Gsf* includes requirements on matching calorimeter clusters to L1 seeds,  $E_T > 27$  GeV ( $E_T$  is a transverse energy of the electron at the HLT level), cuts on  $\sigma_{i\eta i\eta}$ , HCAL and ECAL isolation,  $1/E - 1/p$ ,  $\chi^2$  of the track, missing hits,  $\Delta\eta_{in}$  and  $\Delta\phi_{in}$  and pixel seed matching. *HLT\_Mu50* includes requirements on matching muon candidates to L1 seeds,  $\chi^2$  of the track candidate, track impact parameters,  $p_T > 50$  GeV,  $|\eta| < 2.5$ .

### 5.2.2 Noise filters

Beam halo muons or non-responsive readout channel in the HCAL [176] and ECAL can cause anomalous events with large values of the measured  $\cancel{E}_T$ . In order to mitigate this effect a number of filters is applied following recommendations from [177] both in data and simulation.

### 5.2.3 Electron selection

Two types of electrons are defined: *loose* and *tight* electrons. Both should pass HEEP identification criteria (see Subsection 3.2.4). *Tight* electrons are required to have  $p_T > 50$  GeV and *loose* electrons  $p_T > 35$  GeV.

### 5.2.4 Muon selection

Two types of muons are defined: *loose* and *tight* muons. Both should pass high- $p_T$  muon identification criteria (see Section 3.2.3). *Tight* muons are required to have  $p_T > 53$  GeV and  $|\eta| < 2.4$  and *loose* muons  $p_T > 20$  GeV and  $|\eta| < 2.4$ .

### 5.2.5 Jet selection

AK8 jets are required to pass jet identification criteria (see Subsection 3.2.5) and to have  $p_T > 200$  GeV and  $|\eta| < 2.4$ . They should be with  $\Delta R > 1.0$  from tight leptons. The hardest in  $p_T$  jet is considered for the reconstruction of hadronically decaying W or Z boson. Additionally this jet required to have  $40 \text{ GeV} < M_{\text{pruned}} < 150 \text{ GeV}$  and  $\tau_{21} < 0.6$  (see definition in Section 4.5) where  $M_{\text{pruned}}$  is the jet pruned mass ( $z_{\text{cut}} = 0.1$ ,  $r_{\text{cut}} = 0.5$ ).

AK4 jets are used to suppress contributions from  $t\bar{t}$  by applying a b-tag veto. They are required to have  $p_T > 30$  GeV and  $|\eta| < 2.4$ . These jets are should also be within  $\Delta R > 0.3$  from tight leptons and within  $\Delta R > 0.8$  from the hardest in  $p_T$  jet from the collection of selected AK8 jets. The last one is done in order not apply a b-tag veto on the  $W(Z)$  boson candidate since  $WZ$ -events with  $Z \rightarrow b\bar{b}$  are signal events.

### 5.2.6 Missing $E_T$ selection

$\cancel{E}_T$  is required to pass cut of 40 (80) GeV in the muon (electron) channel. The higher  $\cancel{E}_T$  threshold for electrons is necessary to reduce fakes from multijet events, since the electron identification criteria are optimized for efficiency instead of purity. The effect from the QCD fakes in the electron channel is illustrated in Appendix H.5 where  $\cancel{E}_T$  cut distribution is shown in the signal region with  $\cancel{E}_T$  relaxed to 40 GeV.

### 5.2.7 Final event selection

Exactly 1 tight lepton in the event is required and additional loose leptons are vetoed. The efficiency for the second lepton veto is very high (99.6% in electron channel and 97.7% in the muon channel).

The tight lepton is combined with the  $\cancel{E}_T$  to form the leptonic  $W$  candidate, reconstructing the neutrino longitudinal momentum ( $p_{\nu,z}$ ) from the  $W$ -mass constraint [178]. This constraint leads to the quadratic equation and the solution is given by the following formula:

$$p_{\nu,z}^{1,2} = \frac{\mu \cdot p_{z,\text{lepton}}}{p_{T,\text{lepton}}^2} \pm \sqrt{\frac{\mu^2 \cdot p_{z,\text{lepton}}^2}{p_{T,\text{lepton}}^4} - \frac{E_{\text{lepton}}^2 \cdot |\vec{\cancel{E}}_T|^2 - \mu}{p_{T,\text{lepton}}^2}} \quad (5.1)$$

where  $\mu = \frac{m_W^2}{2} + \vec{p}_{T,\text{lepton}} \cdot \vec{\cancel{E}}_T$ . In the case of a complex solution, the real part of the solution is assigned as longitudinal momentum, in the case of two real solutions the solution with the smaller absolute value is assigned. It was checked that most of events have a real solution (78% both in the muon and electron channel). The leptonic  $W$  candidate is required to have  $p_T > 200$  GeV. Also following cuts are applied in order to enrich the "back-to-back" topology expected for signal event:

$$\Delta R(\text{lepton}, W\text{-jet}) > \frac{\pi}{2} \quad (5.2)$$

$$\Delta\Phi(W\text{-jet}, \cancel{E}_T) > 2.0 \quad (5.3)$$

$$\Delta\Phi(W\text{-jet}, W\text{-lep}) > 2.0 \quad (5.4)$$

where  $W$ -jet means the hardest jet in  $p_T$  from the collection of selected AK8 jets and  $W$ -lep stands for the leptonically decaying  $W$  boson candidate. Additionally a cut on diboson mass  $M_{WV} > 900$  GeV is applied as only this region is used further in the limit extraction. The selection applied in the analysis is summarized in the Table 5.4.

4 regions are defined: 3 control regions and the signal region. The measurement is done in the signal region while control regions are used for the validation of the modelling of background processes. Events in the signal region are required to have no b-tagged AK4 jets and

**Table 5.4:** Summary of the selection applied in the analysis.

cut	electron channel	muon channel
lepton $p_T >$	50 GeV	53 GeV
lepton $ \eta  <$	2.5	2.4
$\cancel{E}_T >$	80 GeV	40 GeV
$\tau_{21} <$	0.6	0.6
$M_{WV} >$	900 GeV	900 GeV
$W_{\text{lep}} p_T >$	200 GeV	200 GeV
W-jet $p_T >$	200 GeV	200 GeV
W-jet $ \eta  <$	2.4	2.4
$\Delta R(\text{lepton}, \text{W-jet}) >$	$\frac{\pi}{2}$	$\frac{\pi}{2}$
$\Delta\Phi(\text{W-jet}, \cancel{E}_T) >$	2.0	2.0
$\Delta\Phi(\text{W-jet}, \text{W lep}) >$	2.0	2.0
$M_{\text{pruned}} \in$	[40, 150] GeV	[40, 150] GeV

$65 \text{ GeV} < M_{\text{pruned}} < 105 \text{ GeV}$ . The signal region is further divided into 2 categories: WW-category ( $M_{\text{pruned}} \in [65, 85] \text{ GeV}$ ) and WZ-category ( $M_{\text{pruned}} \in [85, 105] \text{ GeV}$ ). In the W+jets control region there should be no b-tagged AK4 jets and  $M_{\text{pruned}} \in [40, 65] \text{ GeV} \cup [105, 155] \text{ GeV}$ . In the  $t\bar{t}$  control region there should be at least 1 b-tagged AK4 jet and  $M_{\text{pruned}} \in [40, 150] \text{ GeV}$ . In the  $t\bar{t}$  control region II there is no requirement on b-tagging and  $M_{\text{pruned}} \in [155, 200] \text{ GeV}$ . This is illustrated in Figure 5.5.

In Figure 5.6 the event display of the event with the highest (from some part of 2016 data) reconstructed diboson mass that passed the selection described above is shown (3.6 TeV). The event contains a very energetic jet (832.8 GeV),  $\cancel{E}_T$  (638.5 GeV) and a muon with  $p_T$  of 132.5 GeV.

### 5.3 Simulation corrections

In order to describe data several effects in the simulation should be corrected: proper normalization, different efficiencies in data and the simulation. All the corrections are applied as *weights* for every event in the simulation. The total correction is calculated as a product of different weights. In case the process is generated with MADGRAPH5\_aMC@NLO with NLO precision the sign of the weight at the generator level is taken into account (see Subsection 3.1.1).

#### 5.3.1 Normalization

The total number of the expected events for the process with cross-section  $\sigma_p$  for the integrated luminosity  $L$  is  $\sigma_p \cdot L$ . In order to get a proper normalization of the processes in Table 5.2 a

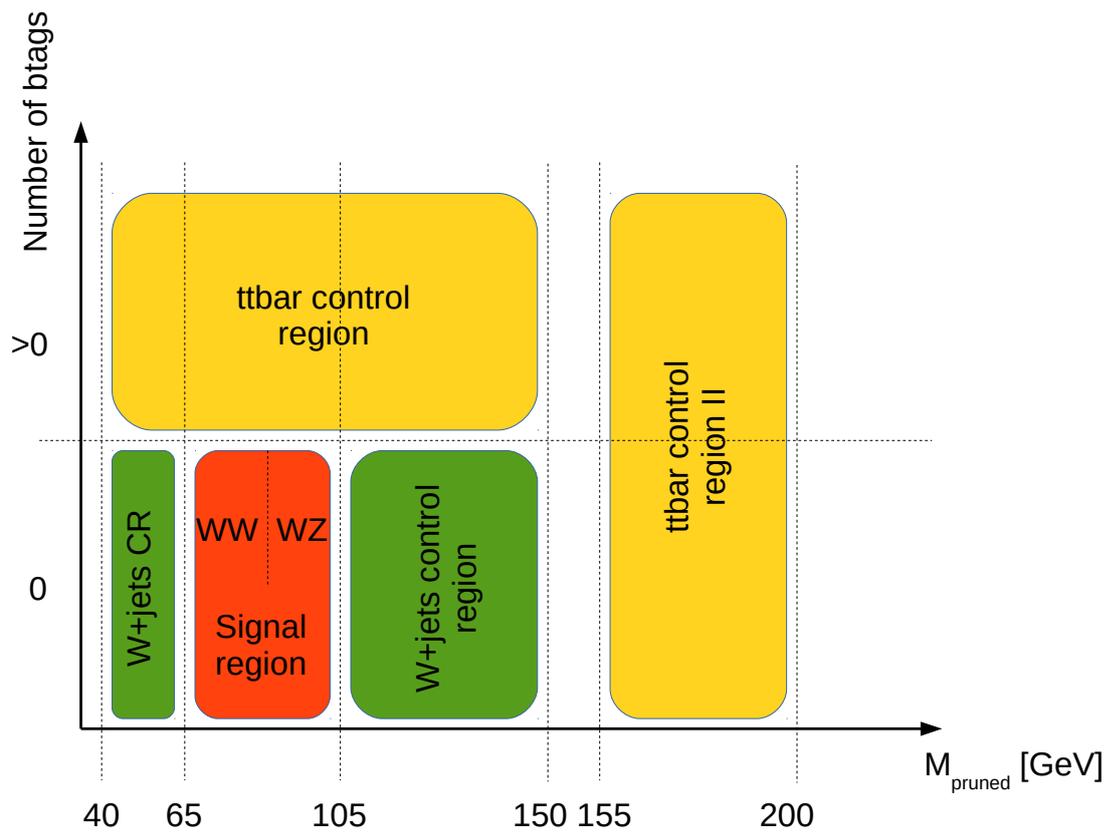


Figure 5.5: Illustration of signal region and control regions used in the analysis.

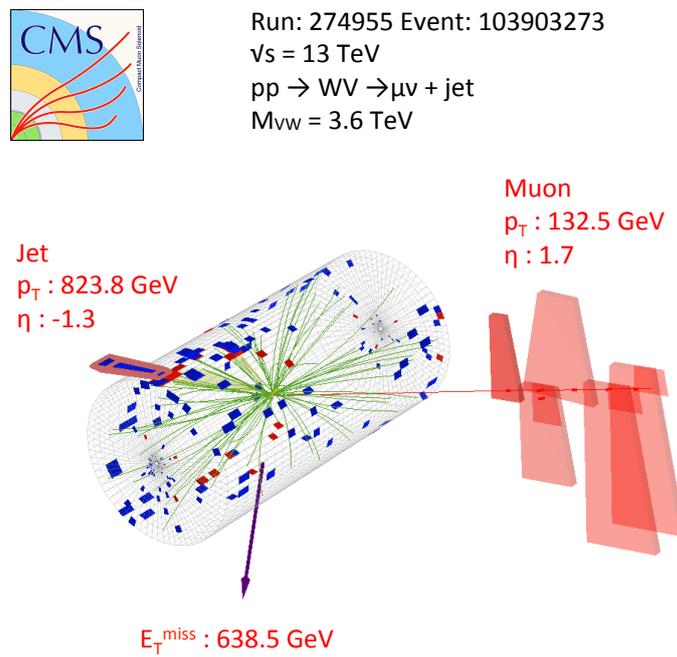


Figure 5.6: Event display of the event with the highest diboson mass.

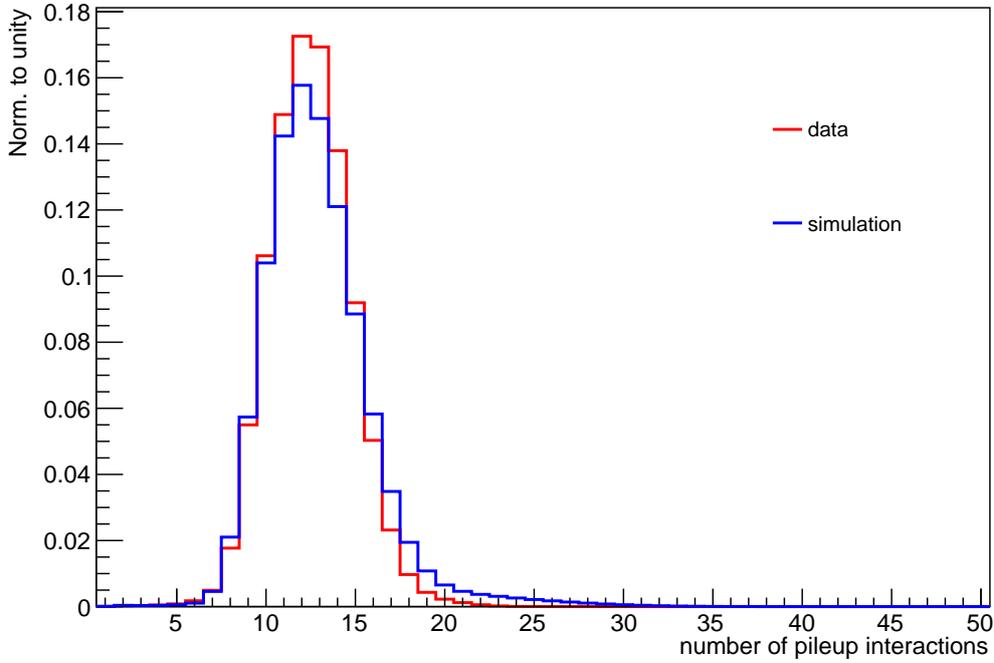
weight  $w_N$  is assigned to every event in the simulation:

$$w_N = \frac{\sigma_p \cdot L}{N_{\text{events}}} \quad (5.5)$$

For the signal this is done with weights corresponding to a specific aTGC-point.

### 5.3.2 Pileup

The conditions of colliding beams evolve. Thus it is hard to predict the profile of the number of pileup interactions in data over the running period. Due to that the simulation is usually done with a wider and rather uniform distribution of the number of pileup interactions. Pileup profiles in data and the simulation used in this thesis are shown in Figure 5.7. Each event is assigned a weight which is calculated so that the distribution of pileup events in the simulation matches the one in data.



**Figure 5.7:** Pileup profiles in data and the simulation.

### 5.3.3 B-tagging

The b-tagging efficiency is different in data and the simulation. In order to correct this effect efficiencies in the simulation are measured. The measurement is done for b-jets, c-jets and light flavour jets (u, d, s quarks and gluons). Due to limited statistics in some of samples and also

because AK4-jets are cleaned from the W-jet the number of AK4-jets that are actually b-tagged is limited. Because of that b-tagging and mis-tagging efficiencies are measured only for  $t\bar{t}$  and WZ which should contain most of the events with b-quarks. The W+jets is determined from data (see below) and therefore should not be corrected. The efficiency is defined as a ratio of the jets (of a given flavour) that pass b-tagging working point to the total number of jets (of the given flavour). Measured efficiencies are shown in Figure 5.8 and in Appendix D as a function of the jet  $p_T$  and  $\eta$ . Efficiencies in the electron channel are similar to the ones on the muon channel.

Efficiency in data is calculated as a product of the efficiency in the simulation and the scale factor (ratio of efficiency in data and the simulation). Scale factors were measured in references [179, 180]. Scale factors for c-jets are assumed to be the same as for b-jets.

Then the event weight should be calculated. Basically, the probability to get the given jet configuration in Monte-Carlo is calculated:

$$P(MC) = \prod_{i=tagged} \epsilon_i \prod_{j=not\ tagged} (1 - \epsilon_j) \quad (5.6)$$

and for data:

$$P(data) = \prod_{i=tagged} SF_i \epsilon_i \prod_{j=not\ tagged} (1 - SF_j \epsilon_j) \quad (5.7)$$

where  $\epsilon_i$  is the efficiency of b-tagging of the jet of the given flavour (as a function of the jet  $p_T$  and  $|\eta|$ );  $SF_i$  is the scale factor for the jet of the given flavour (as a function of the jet  $p_T$  and  $|\eta|$ ). Then a weight is calculated as following :

$$w_{btag} = \frac{P(data)}{P(MC)} \quad (5.8)$$

### 5.3.4 Lepton identification efficiencies

Electron and muon identification criteria have different efficiencies in data and the simulation. In order to take this effect into account scale factors are measured with the Tag-and-Probe method and applied to the simulation on an event-by-event basis (since exactly one lepton is required). The scale factor for HEEP identification criteria was measured in [181, 182] and is about 0.99. Scale factors in the muon channel are applied as  $SF_{ID} \cdot SF_{relIso}$  where  $SF_{ID}$  is the scale factor for high- $p_T$  muon identification criteria and  $SF_{relIso}$  is the scale factor corresponding to the cut on the relative muon isolation (see Subsection 3.2.3). Both scale factors were measured in reference [183].

### 5.3.5 Trigger efficiencies

The efficiencies of the trigger in data and the simulation can be slightly different <sup>1</sup>. There are 2 ways to correct these effects: either to measure a scale factor between data and the simulation and to assign it as an event weight in the simulation or to reweight the simulation directly with the trigger efficiency in data.

---

<sup>1</sup>the definition of HLT in data and the simulation is, strictly speaking, also different

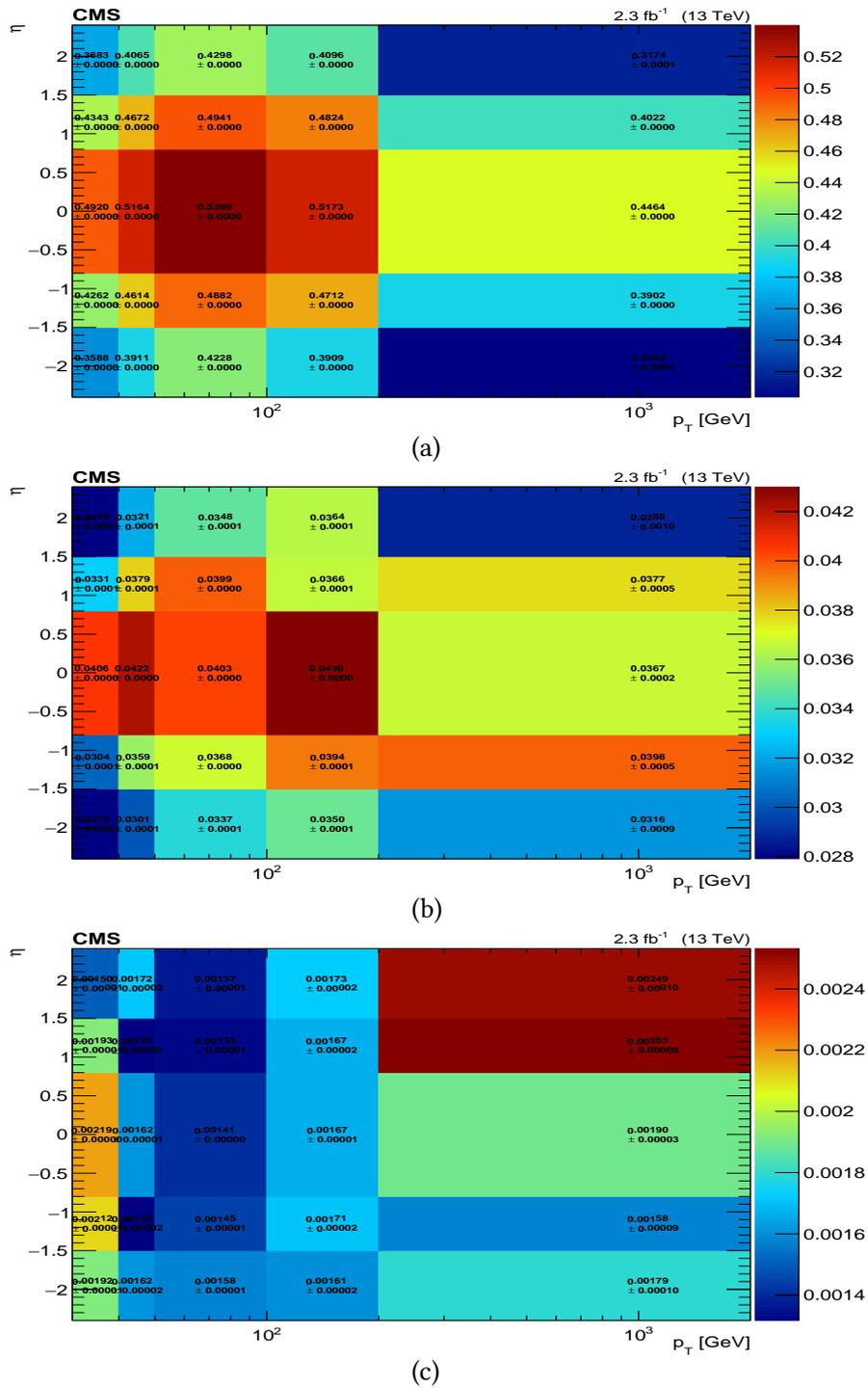
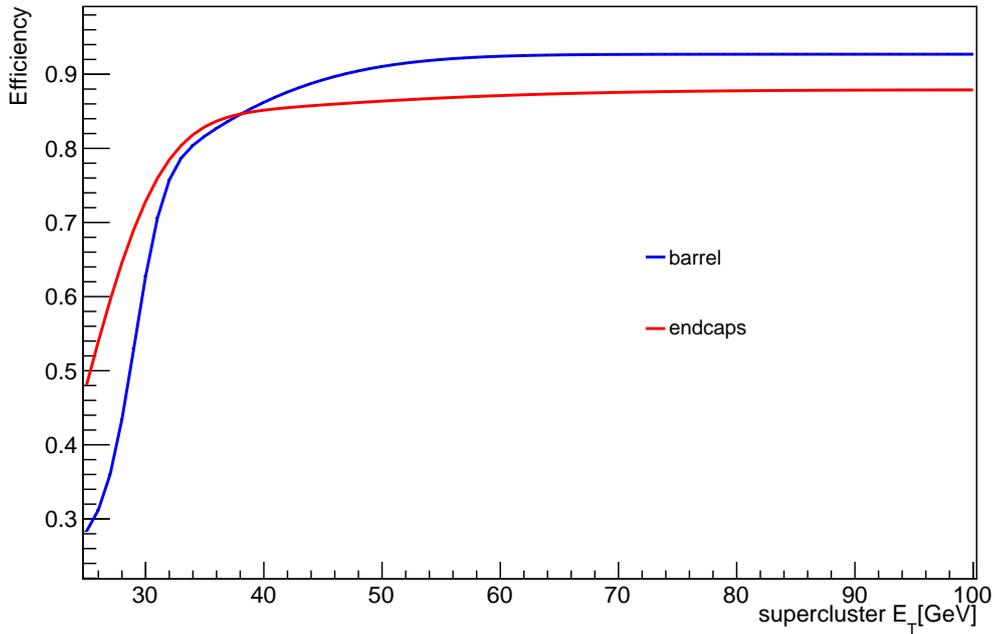


Figure 5.8: B-tagging (mis-tagging) efficiencies for  $t\bar{t}$  sample, muon channel for b-jets (a), c-jets (b) and light flavour jets (c).

The first method is used in the muon channel. Efficiencies in data and the simulation were computed with the Tag-and-Probe method in references [131, 184]. Efficiencies were measured with the counting method by basically counting a number of probes passing the trigger path. Probes were required to pass high- $p_T$  identification requirement (see Subsection 3.2.3).

In the electron channel events in the simulation were reweighted with the trigger efficiency measured in data. The turn on curve was measured in references [131, 185] and is shown in Figure 5.9. Electrons are also required to pass HEEP identification criteria (see Subsection 3.2.4).



**Figure 5.9:** Efficiency of the *HLT\_Ele27\_WPLoose\_Gsf* as a function of the supercluster transverse energy.

### 5.3.6 Vector boson identification efficiencies

Efficiency of W (Z) boson identification with  $\tau_{21}$  cut is also different between data and the simulation. The scale factor is measured using semileptonic  $t\bar{t}$  events based on a simultaneous fit of the jet pruned mass distributions passing and failing the  $\tau_{21}$  cut [186] and is found to be 0.915 with uncertainty of 0.11. The scale factor is applied only to the processes containing *real* W (Z) bosons decaying hadronically which are  $t\bar{t}$ , WW, WZ and single top tW-channel.

## 5.4 Systematic uncertainties

The following uncertainties are considered in the simulation:

- *b-tagging and mistagging uncertainty*: the uncertainty is propagated through the uncertainty on scale-factors. Uncertainties for b- and c-jets are treated as correlated and for light jets as uncorrelated [187]. Uncertainties for c-jets are twice the uncertainty for b-jets [180].
- *jet energy correction uncertainty*: the uncertainty is propagated by varying the jet energy correction within uncertainty. W-jet momentum and  $\cancel{E}_T$  are recalculated.
- *jet energy resolution uncertainty*: the uncertainty by varying the scale factor within its uncertainty and repeating the smearing procedure. W-jet momentum and  $\cancel{E}_T$  are recalculated.
- *lepton energy scale<sup>2</sup>*: the momentum is rescaled within scale uncertainty,  $\cancel{E}_T$  is changed accordingly and the diboson mass is recalculated. Electron energy scale uncertainty is estimated as 1.4% [185]. Muon energy scale uncertainty estimation follows the procedure from Run I: if  $p_T$  of the muon is less than 200 GeV the scale uncertainty is 0.2%, if  $p_T > 200$  GeV then extra uncertainty of  $0.05 \cdot p_T / \text{TeV}$  is added (5% uncertainty on momentum scale at 1 TeV) [188].
- *lepton resolution uncertainty*: a lepton is matched to the generator level lepton and factor  $r$  is varied within the resolution uncertainty, where  $r$  defines the resolution:

$$p_T = p_{T,\text{gen}} + r \cdot (p_{T,\text{reco}} - p_{T,\text{gen}}) \quad (5.9)$$

$$p_z = p_{z,\text{gen}} + r \cdot (p_{z,\text{reco}} - p_{z,\text{gen}}) \quad (5.10)$$

$p_{z,\text{reco}}$  and  $p_{z,\text{gen}}$  are z-components of the lepton momentum at the reconstructed and generator level, and  $p_{T,\text{reco}}$  and  $p_{T,\text{gen}}$  are transverse momenta at the reconstructed and generator level. Resolution factor  $r$  is varied within uncertainty, a lepton momentum,  $\cancel{E}_T$  and diboson mass are recalculated. For electrons 1.16% uncertainty in the barrel and 1.61% uncertainty in the endcap are taken according to [185]. For muons an uncertainty of 0.6% is considered following the Run I recommendation [188].

- *lepton identification efficiency*: scale factor for the lepton identification criteria is varied with an uncertainty. It is estimated to be 1.0% in the electron channel [182] and about 3% in the muon channel (additionally 1% for ID + 0.5% for isolation, + 0.5% for single muon triggers) [189]. In the muon channel an additional uncertainty of -2.5 % (one-sided) is added for  $p_T > 200$  GeV following [184].
- *missing  $E_T$  uncertainty*: the uncertainty is estimated by propagating the uncertainties from all inputs to  $\cancel{E}_T$ : leptons, calorimeter clusters, jets.

<sup>2</sup>the lepton energy scale and resolution uncertainties in the electron channel are taken from mass scale and mass resolution uncertainties measured by the analysis  $Z' \rightarrow e^+e^-$  [185]. Mass scale is defined as  $m_{\text{reco}}/m_{\text{gen}}$  and resolution is defined as  $\frac{m_{\text{reco}} - m_{\text{gen}}}{m_{\text{gen}}}$  where  $m_{\text{reco}}$  is reconstructed dielectron mass and  $m_{\text{gen}}$  is dielectron mass at the generator level.

- *PDF uncertainty*: 100 alternate sets of the central NNPDF 3.0 [100] densities are considered, then corresponding cross-sections in the bin of the particular variable are ordered:

$$\sigma_1 < \sigma_2 < \dots < \sigma_{99} < \sigma_{100} \quad (5.11)$$

The PDF uncertainty is quoted as  $\delta\sigma_{PDF} = \frac{\sigma_{84} - \sigma_{16}}{2}$ . This approach follows prescriptions in reference [190].

- *$Q^2$ -uncertainty*: different combinations of variation of factorization and renormalization scales are considered (variation always by the factor of 2, 9 combinations in total, the nominal point is  $\mu_F = \mu_R$ ). The envelope of these variations is used to derive the scale uncertainty.
- *V-tagging uncertainty*: 12% uncertainty on the scale factor for  $\tau_{21}$  cut.
- *luminosity uncertainty*: 2.7% on the luminosity measurement [95].

The impact on normalization of processes considered in the analysis (except W+jets which is derived from data as it will be discussed further) is shown Table 5.5.

**Table 5.5:** Estimated normalization uncertainties in % for contributions derived from the simulation. Uncertainties smaller than 0.05% for all processes are not listed [55].

process	jet en.	b-tag	V-tag	lept. en.	lept. id	$\cancel{E}_T$	lumi	PDF	scale
electron channel									
$t\bar{t}$	2.8	0.8	12	<0.05	1.0	0.5	2.7	2.5	19
WZ	1.7	0.1	12	<0.05	1.0	0.5	2.7	2.5	3.6
WW	2.4	<0.05	12	0.6	1.0	0.6	2.7	1.9	6.0
Single Top	1.6	<0.05	12	0.5	1.0	1.2	2.7	0.3	2.0
muon channel									
$t\bar{t}$	2.6	0.8	12	1.6	3.2	0.1	2.7	2.6	19
WZ	1.6	<0.05	12	1.4	3.8	0.3	2.7	2.3	3.5
WW	2.3	<0.05	12	1.7	3.9	0.2	2.7	1.8	6.0
Single Top	0.6	<0.05	12	1.9	3.6	0.5	2.7	0.4	1.9

## 6 Signal modelling

Modelling of the effects from anomalous triple gauge couplings is described. The normalization of the signal depends basically quadratically on aTGC. In the modelling of the shape of the signal effects from the interference between aTGC and the SM as well as interference between different aTGC are taken into account. The procedure described follows reference [191]. The estimation of systematic uncertainties affecting the shape of the signal is given as well.

### 6.1 Signal region

The signal region is defined in the window  $M_{\text{pruned}} \in [65, 105]$  GeV and is divided further into WW- ( $M_{\text{pruned}} \in [65, 85]$  GeV) and WZ-category ( $M_{\text{pruned}} \in [85, 105]$  GeV, see Subsection 5.2.7). This is motivated by the fact that a non-zero  $c_B$  increases the WW cross section much more than the WZ cross section. Thus in case of presence of non-zero  $c_B$  there would be a stronger enhancement in the WW-category than in the WZ-category while in case of non-zero  $c_W$  or  $c_{WWW}$  an enhancement in the WW- and WZ-category is expected to be comparable, thus leaving  $c_W$  and  $c_{WWW}$  scenarios undistinguishable. As shown in Figures 6.1 and 6.2 (non-zero  $c_W$  or  $c_{WWW}$ ) there is a significant number of WZ events while in Figure 6.3 (non-zero  $c_B$ ) the number of WZ events is strongly suppressed. Data to simulation comparison in the signal region can be found in Appendix H.4.

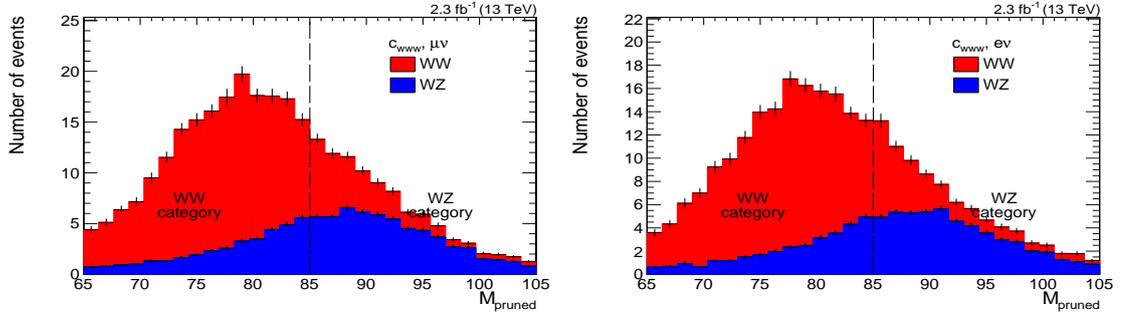
### 6.2 Signal function

As it was already mentioned in Section 1.3 the diboson mass distribution is used to study effects from anomalous triple gauge couplings. Effects from aTGC are described analytically as a function of aTGC and diboson mass ( $M_{WV}$ ). The signal function is defined as:

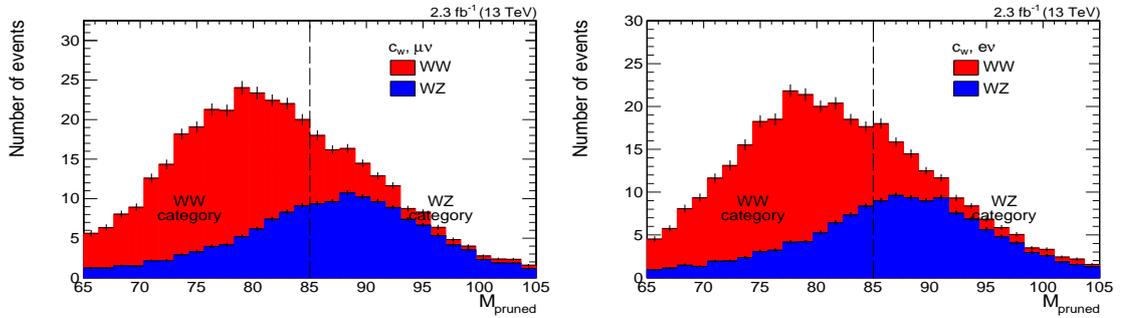
$$F_{\text{signal}}(c_i, M_{WV}) = N_{SM}^{NNLO} \cdot N_{aTGC}(c_i) \cdot S_{aTGC}(c_i, M_{WV}) \quad (6.1)$$

where  $N_{SM}^{NNLO}$  is a number of events in the SM predicted with NNLO precision,  $N_{aTGC}(c_i)$  corresponds to the normalization and  $S_{aTGC}(c_i, M_{WV})$  to the shape of the signal,  $c_i$  is a parameter corresponding to an anomalous triple gauge coupling (see (1.27) and Table 5.1) defined as:

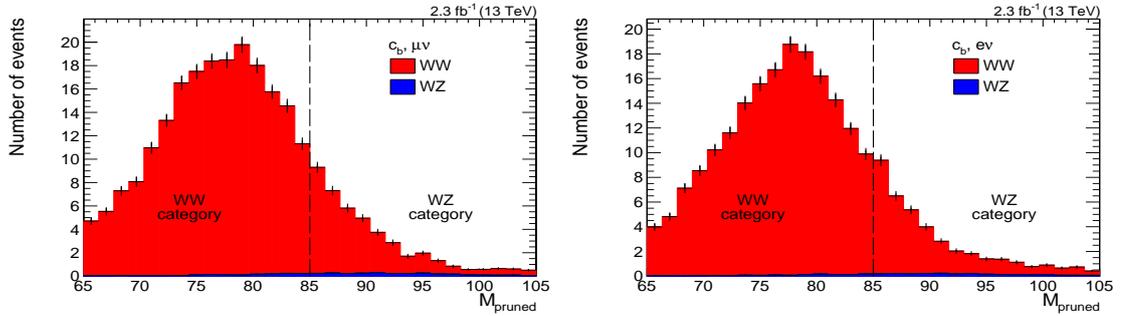
$$\begin{aligned} c_1 &= \frac{1}{12} \frac{c_{WWW}}{\Lambda^2} \\ c_2 &= \frac{1}{20} \frac{c_W}{\Lambda^2} \\ c_3 &= \frac{1}{60} \frac{c_B}{\Lambda^2} \end{aligned} \quad (6.2)$$



**Figure 6.1:**  $M_{\text{pruned}}$  distribution of the signal with  $c_{w,w} = -12.0 \text{ TeV}^{-2}$ . Muon channel on the left and electron channel on the right.



**Figure 6.2:**  $M_{\text{pruned}}$  distribution of the signal with  $c_w = -20.0 \text{ TeV}^{-2}$ . Muon channel on the left and electron channel on the right.



**Figure 6.3:**  $M_{\text{pruned}}$  distribution of the signal with  $c_b = -60.0 \text{ TeV}^{-2}$ . Muon channel on the left and electron channel on the right.

### 6.2.1 Normalization

The inclusion of aTGC changes the matrix element (see Section 3.1) so that it can be written as:

$$M_{fi} = M_{fi,SM} + \sum_{i=1}^3 c_i \cdot M_{fi,aTGC} \quad (6.3)$$

where  $M_{fi,SM}$  is the matrix element in the Standard model and  $M_{fi,aTGC}$  is the matrix element corresponding to the inclusion of the given aTGC. As the cross-section is proportional to the  $|M_{fi}|^2$  (see Section 3.1) the signal function should contain terms proportional to  $c_i^2$ ,  $c_i$  (from the interference between the SM and aTGC),  $c_i \cdot c_j$  ( $i \neq j$ , from the interference between different aTGC). This motivates the choice for the parametrization of the normalization function<sup>1</sup>:

$$N_{aTGC}(c_i) = 1 + \sum_{i=1}^3 p(c_i) \quad (6.4)$$

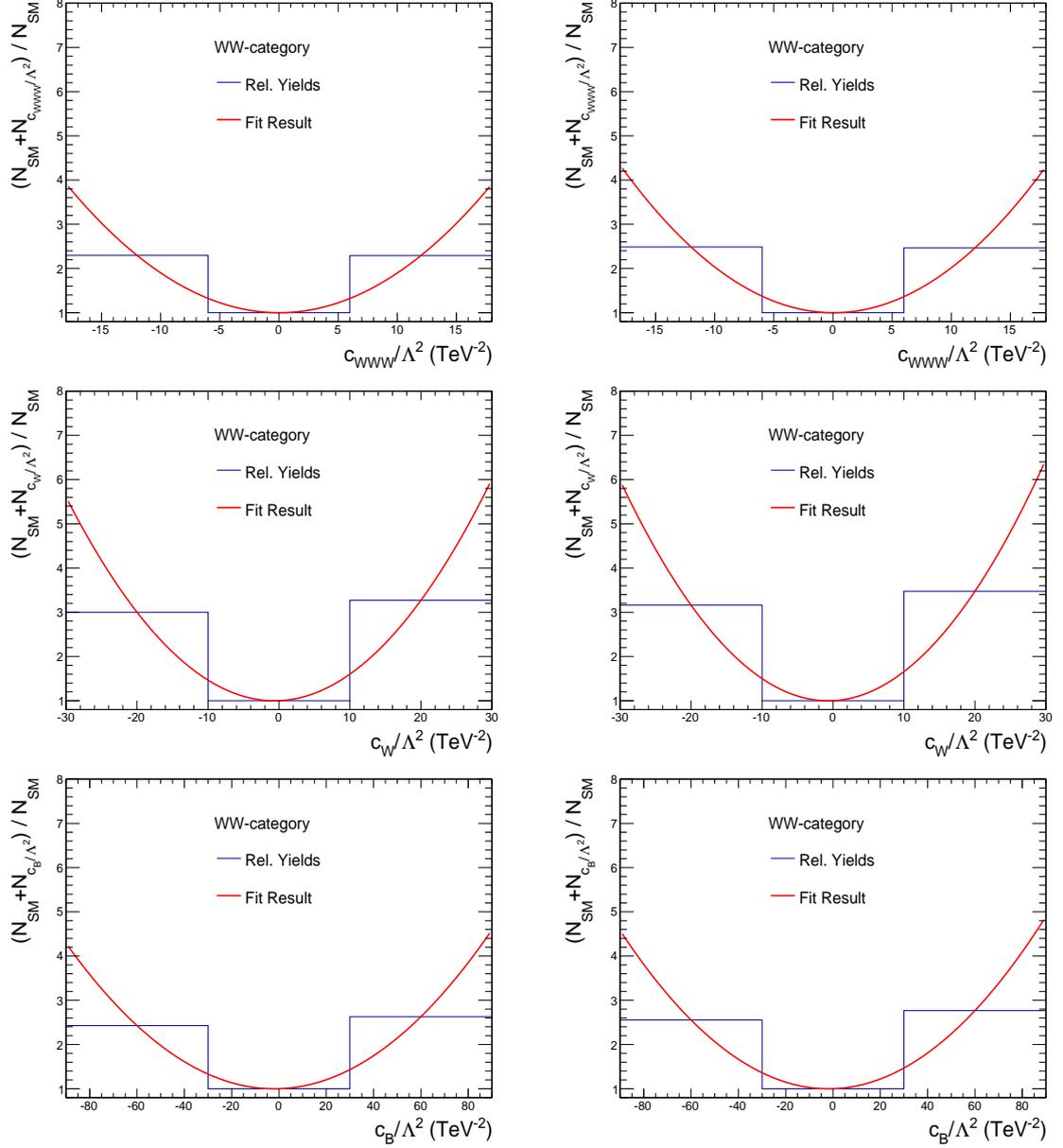
where  $p(c_i) = a_i \cdot c_i + b_i c_i^2$  and in the SM case the normalization function is defined to be equal to unity:  $N_{aTGC}(c_i = 0) = 1$ . Parameters  $a_i$  and  $b_i$  can be defined from the simulation by doing a fit: one point corresponds to the SM and two points are taken from the simulation when one of the aTGC is not equal to zero (see Table 5.1). Results of the fit are shown in Figure 6.4 for the WW-category and in Appendix E.1 for WZ-category. It should be mentioned that requirement  $N_{aTGC}(c_i = 0) = 1$  automatically rescales the signal function with the k-factor defined as the ratio of the number of selected events at NNLO precision to the number of selected events at LO precision for the case of the SM (no aTGC):  $N_{SM}^{NNLO}/N_{SM}^{LO}$ . This approximation is done because the simulation of aTGC effects was available only at LO at the time this thesis was done.

### 6.2.2 Shape

The signal shape as a function of  $M_{WV}$  follows an exponentially falling spectrum to a good approximation. From quantum field theory one would expect a power function ( $\sim M_{WV}^{-4}$  from the propagator). However, there is a convolution with PDF which should be a steeply falling function in this region of the phase space. This motivates the choice of the signal shape function:

$$S_{aTGC} \cdot A_N = N_{SM} \cdot e^{a_0 M_{WV}} + \sum_i \left( N_{c_{i,1}} \cdot c_i^2 \cdot e^{a_{i,1} M_{WV}} \cdot \frac{1 + \text{Erf}((M_{WV} - a_{o,i})/a_{w,i})}{2} + \right. \\ \left. + N_{c_{i,2}} \cdot c_i \cdot e^{a_{i,2} M_{WV}} \right) + \sum_{\substack{i < j \\ i \neq j}} \left( N_{c_{i,c_j}} \cdot c_i \cdot c_j \cdot e^{a_{ij} M_{WV}} \right) \quad (6.5)$$

<sup>1</sup>interference effects between different aTGC are neglected here as this was found to be a minor effect



**Figure 6.4:** Normalized yields for different values of the aTGC-parameters in the WW-category for electron- (left) and muon-channel (right) with fitted quadratic function. From top to bottom:  $\frac{c_{WWW}}{\Lambda^2}$ ,  $\frac{c_W}{\Lambda^2}$ ,  $\frac{c_B}{\Lambda^2}$ .

where

$$\begin{aligned}
A_N &= N_{SM} + \sum_i (N_{c_i,1} \cdot c_i^2 + N_{c_i,2} \cdot c_i) + \sum_{\substack{i < j \\ i \neq j}} (N_{c_i, c_j} \cdot c_i \cdot c_j) \\
N_{c_i,1} &= \frac{N_{c_i}^{MC^+} + N_{c_i}^{MC^-}}{2} - N_{SM} \\
N_{c_i,2} &= \frac{N_{c_i}^{MC^+} - N_{c_i}^{MC^-}}{2} \\
N_{c_i, c_j} &= N_{c_i, c_j}^{gen} - (N_{SM} + N_{c_i,1} + N_{c_i,2} + N_{c_j,1} + N_{c_j,2}) \\
&= N_{c_i+, c_j+}^{gen} - \left( N_{SM} + \frac{N_{c_i}^{MC^+} + N_{c_i}^{MC^-}}{2} - N_{SM} + \frac{N_{c_i}^{MC^+} - N_{c_i}^{MC^-}}{2} \right. \\
&\quad \left. + \frac{N_{c_j}^{MC^+} + N_{c_j}^{MC^-}}{2} - N_{SM} + \frac{N_{c_j}^{MC^+} - N_{c_j}^{MC^-}}{2} \right) \\
&= (N_{c_i+, c_j+}^{gen} + N_{SM}) - (N_{c_i}^{MC^+} + N_{c_j}^{MC^+})
\end{aligned}$$

$N_{c_i}^{MC^+}$  and  $N_{c_i}^{MC^-}$  are given by the number of MC-events for the respective aTGC-parameter set to its corresponding positive or negative value.  $N_{c_i+, c_j+}^{gen}$  are given by the number of events for two positive aTGC-parameters. Due to the lack of such simulated samples, these factors are extracted on generator level from an additionally produced sample. For that the sample generation is redone using the same MADGRAPH configuration, generating 150 new coupling points, including low values where SM-interference effects are more prominent. The parameters  $a_0$ ,  $a_{i,1}$ ,  $a_{i,2}$  and  $a_{i,j}$  model the exponential slopes of these contributions and are extracted from the fit. The values of  $a_{o,i}$  and  $a_{w,i}$  govern the turn-on position and steepness of the component connected to a given  $c_i$ . To simplify the signal model very small contributions, i.e.  $c_{WW} - SM$ -interference and  $c_{WW} - c_B$ -interference, are neglected, as well as the error function for  $c_B$  in the WZ-region.

$a_0$  is extracted by simply doing a fit to the SM contribution. The slope for the SM-aTGC interference is extracted by doing a fit to the difference of samples with positive and negative aTGC values. Then the slope for the pure aTGC term is extracted by fitting the sample with aTGC fixed to its positive value and exponentials for the SM and SM-aTGC cases fixed. The remaining three parameters describing aTGC-interference are extracted on generator level. First, the difference of the MC samples for two positive and for one positive and one negative aTGC-parameter value is taken. Since this still includes the SM-interference of one of the aTGC-parameters, this is also subtracted, for example:

$$(MC_{c_B=60, c_W=20} - MC_{c_B=60, c_W=-20}) - (MC_{c_W=20} - MC_{c_W=-20}) \quad (6.6)$$

Since the term corresponding to the interference between different aTGC is estimated using generator level information it is corrected for the selection and reconstruction efficiency by multiplying with the ratio of the number of selected events with all three aTGC non-equal to zero at the simulation level to the same quantity at the generator level. Since it is not possible

to define WW- and WZ-category at the generator level a weighted mean is used to take into account the ratio of WW and WZ events in the simulation.

The signal shape function is shown in Figure 6.5 and in Appendix E.2 for positive and negative values of aTGC in the WW-category. The signal function models well the MC data which verifies the treatment of the SM interference in the signal model for high values of the aTGC-parameters. In order to check how well the signal function models MC data at low values of aTGC (where interference effects are significant) the signal function is extracted as described above using the generator level samples. Since the interference effects are mostly visible in WZ events, only those plots are shown in Figure 6.6 and in Appendix E.2.1. The difference between positive and negative values can be seen at low aTGC values and it is well modelled by the signal function.

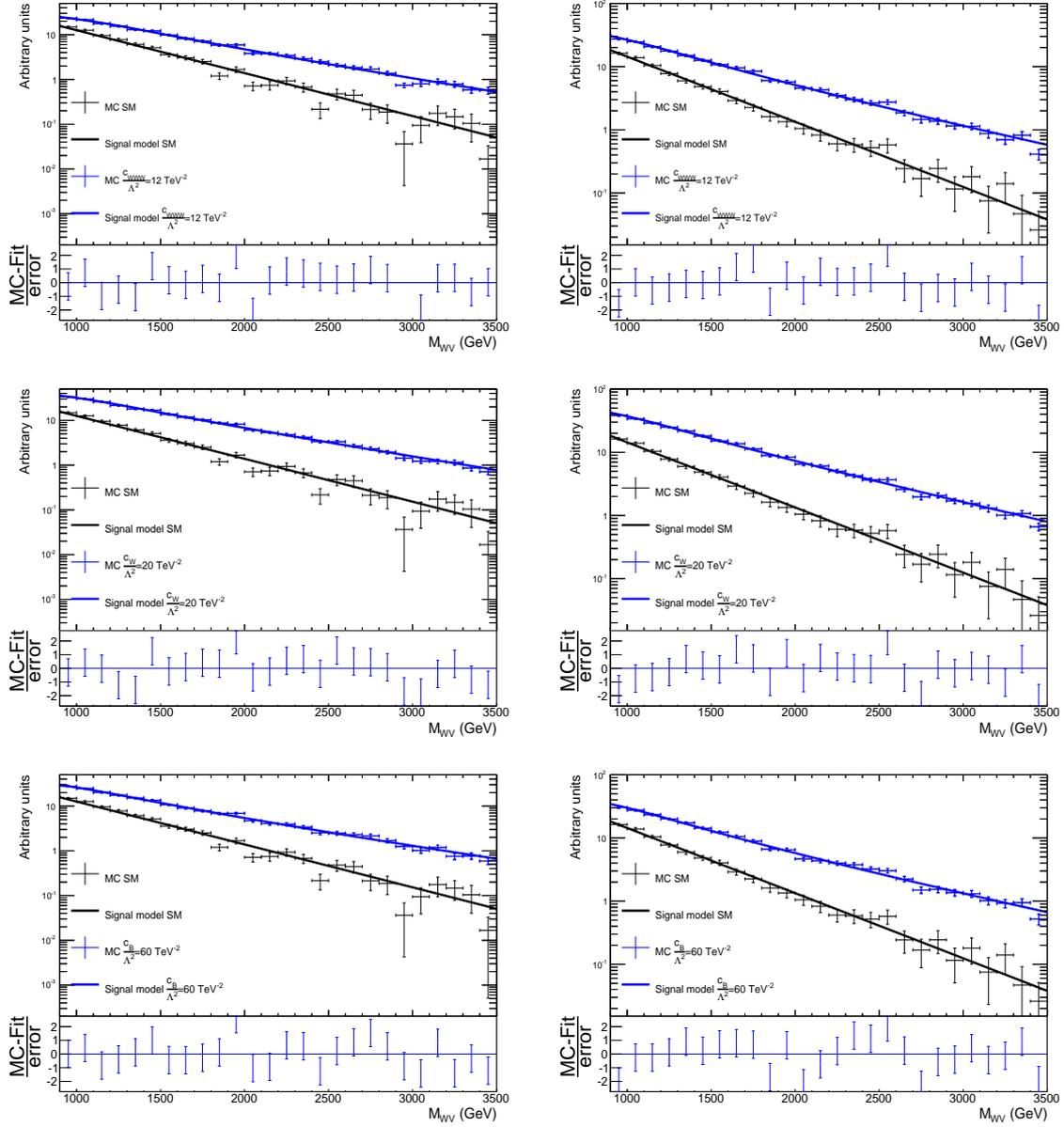
### 6.2.3 Shape uncertainties

The estimation of shape uncertainties of the signal function is performed as follows: The signal function with no interference terms included is fitted to the MC data varied according to the particular source of the systematic uncertainty and the slope of each of the exponential functions is extracted. The bigger of the difference between the nominal slope and the one from the up/down variation is used as uncertainty estimate for this particular source of a systematic uncertainty. All systematic uncertainties described in Section 5.4 are included except luminosity and V-tagging uncertainties which affect only the normalization of the process. The procedure is done for WW- and WZ-category in the electron and muon channel, for every aTGC parameter (fixing a given aTGC parameter to a negative value from Table 5.1).

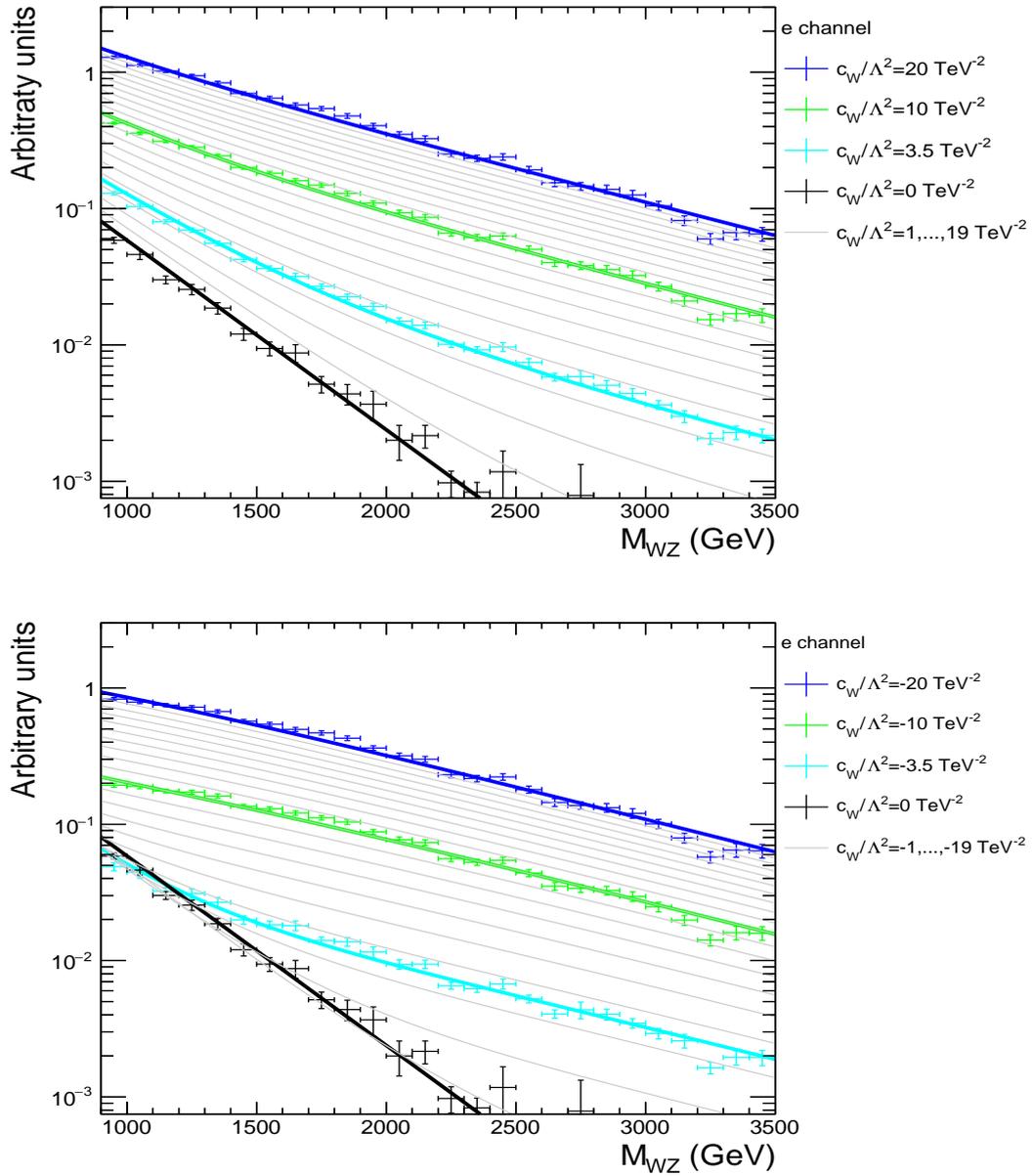
Results for each of the systematic uncertainties are shown in Appendix E.3. Dominant sources of uncertainties are PDF and scale variations for all parameters.  $c_B$  has an additional large uncertainty from jet energy corrections uncertainties: as  $c_b$  only contributes to WW final states, inaccurate JECs can lead to large migrations over the boundary between the WW and WZ category, leading to this large uncertainty. Total uncertainties are given in Table 6.1 and are the squared sum of the individual uncertainties, which we assume to be uncorrelated. The effect on the signal function is illustrated in Appendix E.3.1.

**Table 6.1:** Total uncertainties of signal shape slopes. Uncertainties are quoted in %.

category	$a_{cw}$	$a_{cb}$	$a_{cwww}$
WW, muon	4.59	5.47	5.35
WW, electron	4.96	5.92	5.59
WZ, muon	4.64	15.62	4.83
WZ, electron	4.90	16.18	5.19



**Figure 6.5:** MC data and fit of signal model for all parameters set to zero (black) and one parameter set to a non-zero value (blue) for electron- (left) and muon-channel (right), WW-category. From top to bottom:  $\frac{c_{WWW}}{\Lambda^2} = 12 \text{ TeV}^{-2}$ ,  $\frac{c_W}{\Lambda^2} = 20 \text{ TeV}^{-2}$ ,  $\frac{c_B}{\Lambda^2} = 60 \text{ TeV}^{-2}$ . The ratio plot corresponds to the simulation minus the fit-function divided by the bin-error for the signal function with non-zero aTGC-parameter.



**Figure 6.6:** Signal model and MC on generator level (WZ only) for different values of  $\frac{c_W}{\Lambda^2}$  in the electron channel. The grey lines shows the signal model for increasing  $\frac{c_W}{\Lambda^2}$  in steps of  $1 \text{ TeV}^{-2}$ .

## 7 Background estimation

Three processes contribute to the background: W+jets,  $t\bar{t}$  and single-top production. SM diboson is included as a minor background. The background estimation largely follows the corresponding resonance search [17] with minor modifications. The normalization and shape of the minor single-top background as well as the SM diboson contribution are taken from simulation. The major W+jets background is extracted from sideband in shape and normalization, as described below. The shape of the  $t\bar{t}$  background is determined from simulation, with the normalization influenced by the sideband data.

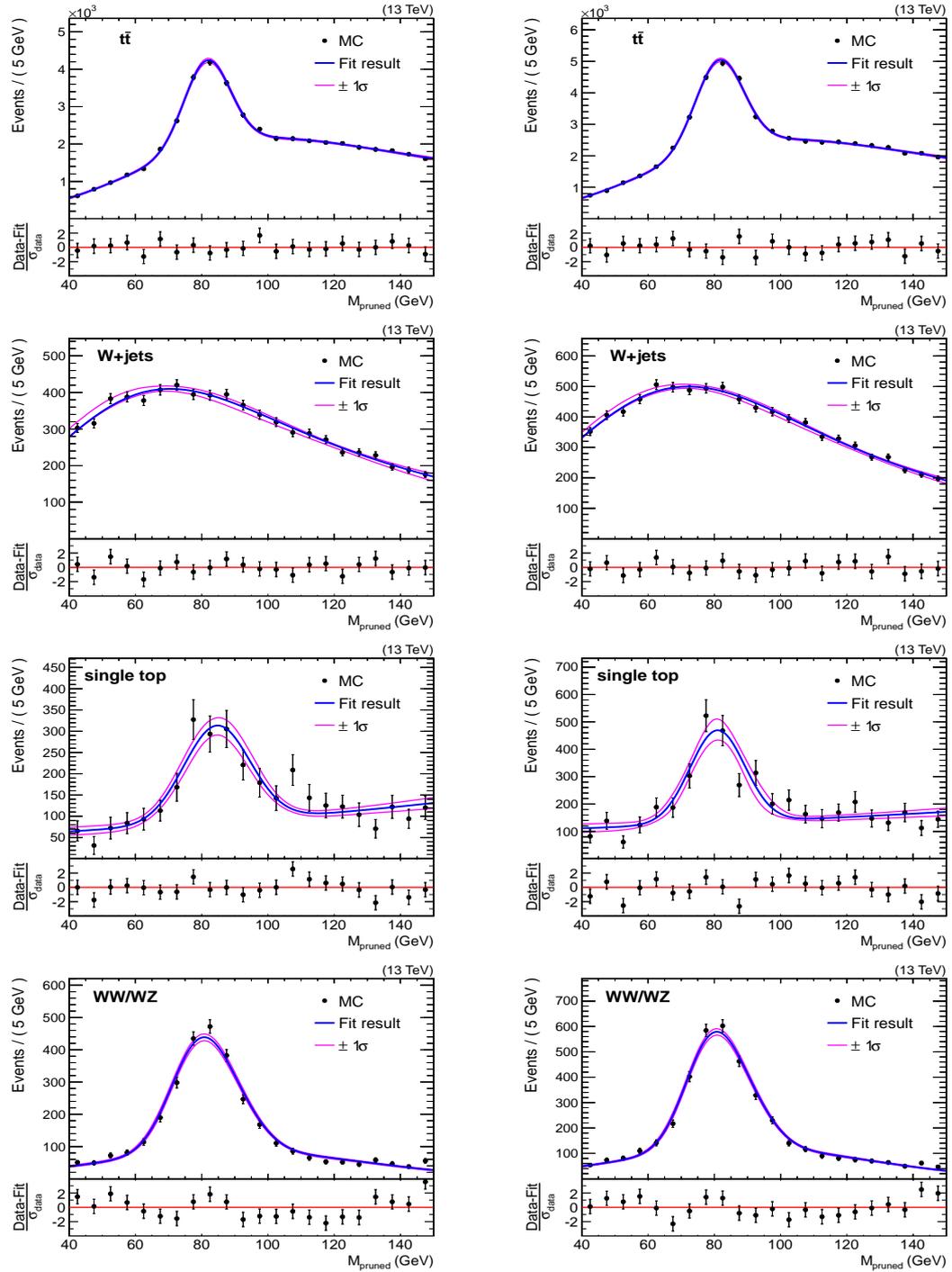
### 7.1 Background normalization

Normalization of  $t\bar{t}$  and W+jets is estimated from the fit of  $M_{\text{pruned}}$  distribution to data in W+jets control region (or sideband with respect to the signal region as it will be referred further) together with the signal region. For each process, a parametric shape of the  $M_{\text{pruned}}$  distribution is extracted from the simulation, using the following functions:

$$\begin{aligned}
 \text{W+jets } F_{\text{Exp}}(x) &= e^{\frac{-c_{EE0}(x-c_{EE1})}{c_{EE2}+c_{EE0}}} \cdot \frac{1 + \text{Erf}((x - c_{EE1})/(c_{EE2} - c_{EE0}))}{2} \\
 \text{t}\bar{\text{t}} F_{\text{ExpGaus}}(x) &= e^{c_{EEG0}x} \cdot \frac{1 + \text{Erf}((x - c_{EEG1})/c_{EEG2})}{2} + c_{EEG3} \cdot G(x, c_{EEG4}, c_{EEG5}) \\
 \text{single-top } F_{\text{ExpGaus}}(x) &= c_{EG0} \cdot e^{c_{EG1}x} + e^{-(x-c_{EG2})^2/2c_{EG3}^2} \\
 \text{diboson } F_{\text{DoubleGaus}}(x) &= c_{DG0} \cdot G(x, c_{DG1}, c_{DG2}) + c_{DG3} \cdot G(x, c_{DG4}, c_{DG5})
 \end{aligned} \tag{7.1}$$

where  $G$  represents the Gaussian function and  $\text{Erf}$  is the error function. The resulting shapes are shown in Figure 7.1. The shown error band is determined by evaluating the fitted functions many times for several points along the  $x$  axis with the fitted parameters randomized according to their covariance matrix. The upper/lower error line excludes 16% of the highest/lowest evaluated function values for each point, resulting in an 68% or  $1\sigma$  error band. The W+jets distribution is modelled to represent the broad Sudakov peak of jet masses expected in highly energetic QCD jets. The functional form for the  $t\bar{t}$  contribution contains a peak, representing real boosted W bosons from hadronic top decays as well as a continuum distribution, which models the case where the top decay products are not easily separated into a boosted W-jet and a distinct b-jet. A similar functional form is used for the single-top background, where the peaking contribution is associated to single-top production in the  $tW$ -channel, while the non-peaking contribution corresponds to the  $t$ -channel. The SM diboson signal is expected to only have a peaking contribution, with minor tails extending into the sideband.

## 7 Background estimation



**Figure 7.1:** MC-data and shapes of the  $M_{\text{pruned}}$  background distributions for electron- (left) and muon-channel (right). From top to bottom:  $t\bar{t}$ ,  $W$ +jets, single Top, diboson.

These templates are then used to fit the  $M_{\text{pruned}}$  distribution in data in the region  $M_{\text{pruned}} \in [40, 150]$  GeV with no b-tagged AK4-jets. This procedure allows the two main background contributions to be determined in the region where their shapes substantially differ: the  $t\bar{t}$  background has a peak at the W mass while W+jets background has a much broader distribution. The normalization of the  $t\bar{t}$  contribution is floated with a Gaussian constraint of  $\sim 20\%$  corresponding to its uncertainty (see Table 5.5). However, it turns out that the constraint on  $t\bar{t}$  cross-section has barely any influence on the resulting normalization so it still matches the SM prediction. The W+jets normalization as well as the shape parameter  $c_{EE0}$  are allowed to freely float in order to account for possible mismodelling of the W+jets background. The diboson contribution is constrained with 100% uncertainty to allow for a substantial contribution from aTGC at the level of the sensitivity of this analysis. The extracted diboson normalization is not used in the further analysis.

Since it is not certain that the chosen function describes the W+jets background correctly, the fit is repeated using an alternative function given by:

$$F_{W\text{Jets}}^{\text{alt}}(x) = \frac{\left(1 - \frac{x}{500}\right)^{p_0}}{\left(\frac{x}{500}\right)^{p_1}} \quad (7.2)$$

where  $p_1$  is extracted from MC and  $p_0$  is floating in the fit to the data. The result of this fit is then propagated to the uncertainty  $\sigma_{W\text{Jets}}^{\text{final}}$  of the W+jets-normalization:

$$\sigma_{W\text{Jets}}^{\text{final}} = \sqrt{\sigma_{W\text{Jets}}^2 + \left(N_{W\text{Jets}}^{\text{Exp}} - N_{W\text{Jets}}^{\text{alt}}\right)^2} \quad (7.3)$$

where  $\sigma_{W\text{Jets}}$  is the W+jets uncertainty extracted from the initial fit,  $N_{W\text{Jets}}^{\text{Exp}}$  and  $N_{W\text{Jets}}^{\text{alt}}$  are number of events from the W+jets background estimated with the nominal function and the alternative function from (7.2). The uncertainty due to alternative function is estimated to be up to 8%.

Distributions with pre-fit, post-fit as well as a fit with alternative function are shown in Figure 7.2 and the resulting background yields are summarized in Table 7.1.

Additionally a bias test of the fit procedure was done. 100 toy MC samples were generated with all fit parameters randomized corresponding to their error of the fit result. Then the pull was calculated as:

$$P = \frac{N_{\text{fit}}^{W\text{Jets}} - N_0^{W\text{Jets}}}{\sigma_{\text{fit}}} \quad (7.4)$$

where  $N_0^{W\text{Jets}}$  is the normalization extracted from the fit to data,  $N_{\text{fit}}^{W\text{Jets}}$  is the normalization from the toy MC sample and  $\sigma_{\text{fit}}$  is the error on  $N_{\text{fit}}^{W\text{Jets}}$ . The resulting pull plots are shown in Appendix F. For some of the distributions RMS is about 1.2 and because that the uncertainty of normalization of W+jets is increased by 20%.

To summarize, 3 numbers from  $M_{\text{pruned}}$  fit are taken into the further analysis: normalization of W+jets and  $t\bar{t}$  background and uncertainty of W+jets.

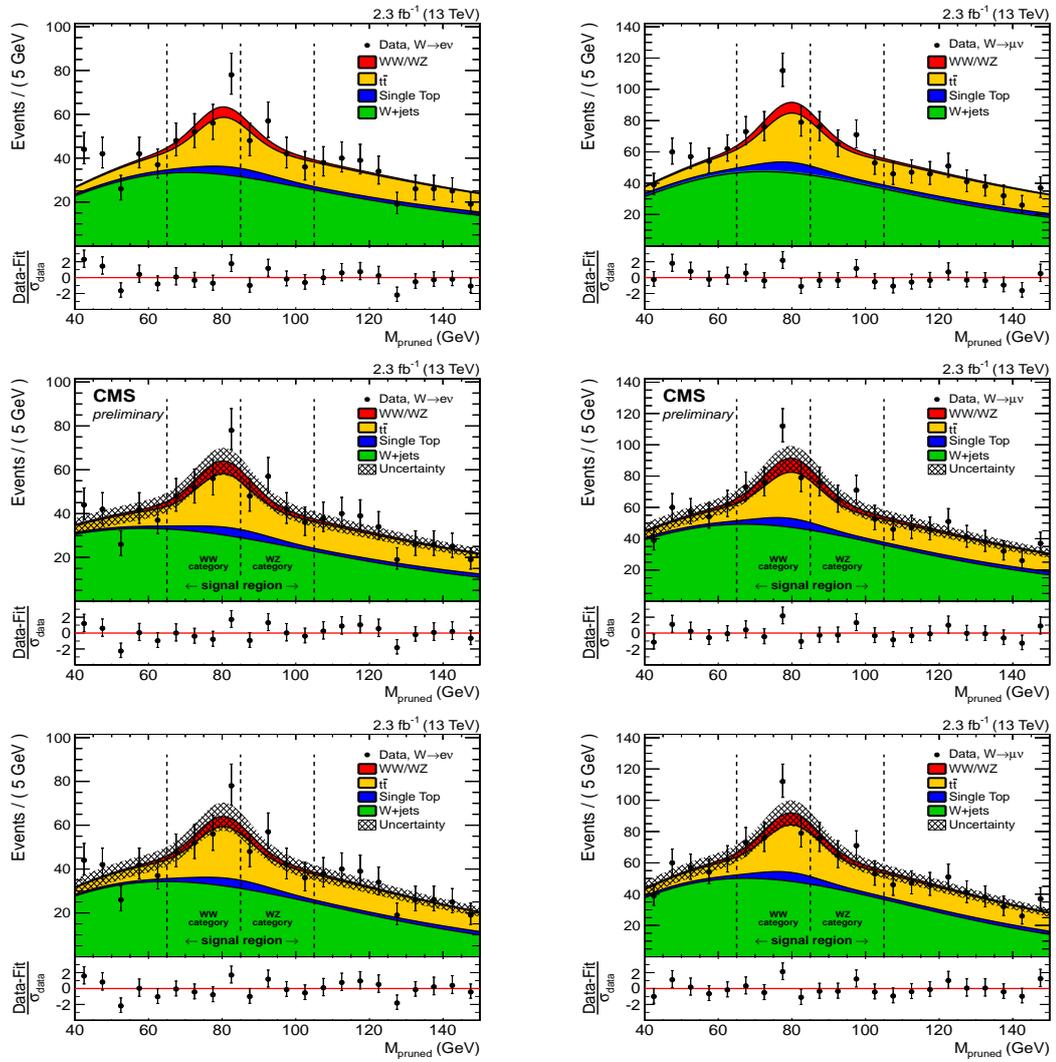


Figure 7.2: Result of the normalization fit in the  $M_{\text{pruned}}$ -spectrum for electron- (left) and muon-channel (right). From top to bottom: pre-fit functions, fit result [55], fit result using an alternative W+jets function.

**Table 7.1:** Results of the fit to the  $M_{\text{pruned}}$  distributions in the range [40., 150.] GeV [55]. The pre- and post-fit yields are presented together with their constraints (pre-fit) and resulting total uncertainties (post-fit). The W+jets contribution is allowed to float in the fit, therefore the pre-fit values do not have any constraint shown. The single top contribution is fixed in the fit.

	electron channel			muon channel		
	pre-fit	post-fit	scale factor	pre-fit	post-fit	scale-factor
W+jets	584	$538 \pm 56$	$0.92 \pm 0.10$	767	$814 \pm 72$	$1.06 \pm 0.09$
$t\bar{t}$	$243 \pm 49$	$256 \pm 46$	$1.1 \pm 0.2$	$318 \pm 64$	$313 \pm 60$	$1.0 \pm 0.2$
single top (fixed)	37	37	1	52	52	1
diboson	$34 \pm 34$	$41 \pm 27$	$1.2 \pm 0.8$	$45 \pm 45$	$61 \pm 35$	$1.4 \pm 0.8$
Total expected	898	$872 \pm 30$	$0.97 \pm 0.03$	1182	$1240 \pm 35$	$1.05 \pm 0.03$
Data		874			1241	

## 7.2 Validation of background in control regions

In order to validate the modelling of main backgrounds agreement between data and the simulation is checked for different variables in control regions: one region enriched with W+jets background and two regions enriched with  $t\bar{t}$  background (see Figure 5.5). In the first  $t\bar{t}$  control region there should be at least one b-tagged AK4-jet while in the other  $t\bar{t}$  control region there is no requirement on the b-tagging but events are selected so that  $M_{\text{pruned}}$  is around the top mass. The second region is constructed in order to check modelling of the  $t\bar{t}$  background without relying on b-tagging. Good agreement is found in all the control regions.

### 7.2.1 W+jets control region

Comparison between data and the simulation for different variables in the W+jets control region is shown in Figure 7.3 and in Appendix H.1. Normalization of the W+jets and  $t\bar{t}$  background is taken from the  $M_{\text{pruned}}$  fit described Subsection 7.1.

### 7.2.2 $t\bar{t}$ control regions

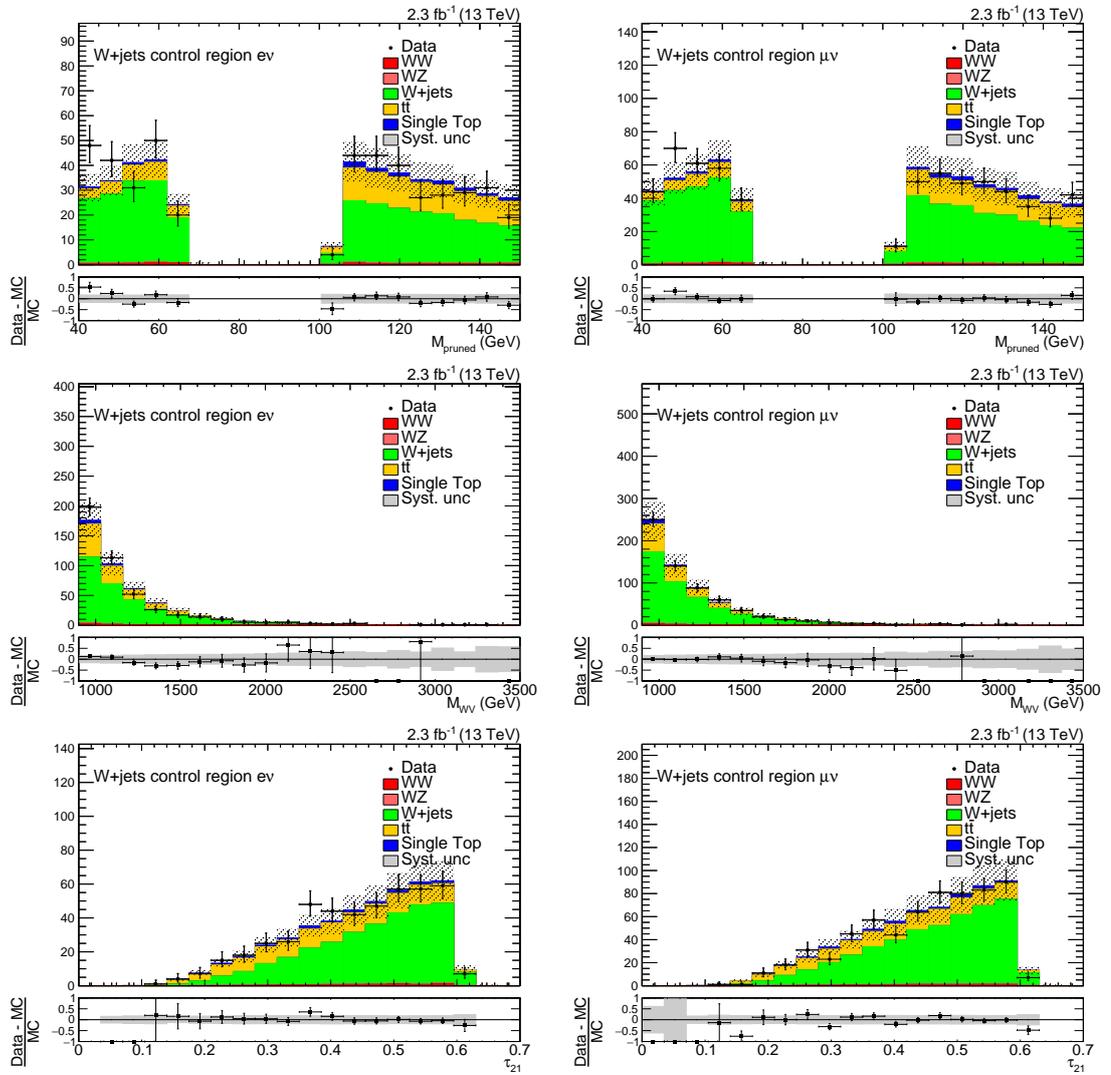
Comparison between data and the simulation for different variables in the  $t\bar{t}$  control region is shown in Figure 7.4 and in Appendix H.2 and in Appendix H.3 ( $t\bar{t}$  control region II). Normalization of the W+jets and  $t\bar{t}$  background is taken from the  $M_{\text{pruned}}$  fit described Subsection 7.1.

Additionally in the  $t\bar{t}$  control region II the following ratio is checked:

$$\frac{N_{\text{veto}}^{\text{data}} - N_{\text{veto}}^{\text{W+jets}}}{N_{\text{veto}}^{t\bar{t}}} / \frac{N_{\text{incl}}^{\text{data}} - N_{\text{incl}}^{\text{W+jets}}}{N_{\text{incl}}^{t\bar{t}}} \quad (7.5)$$

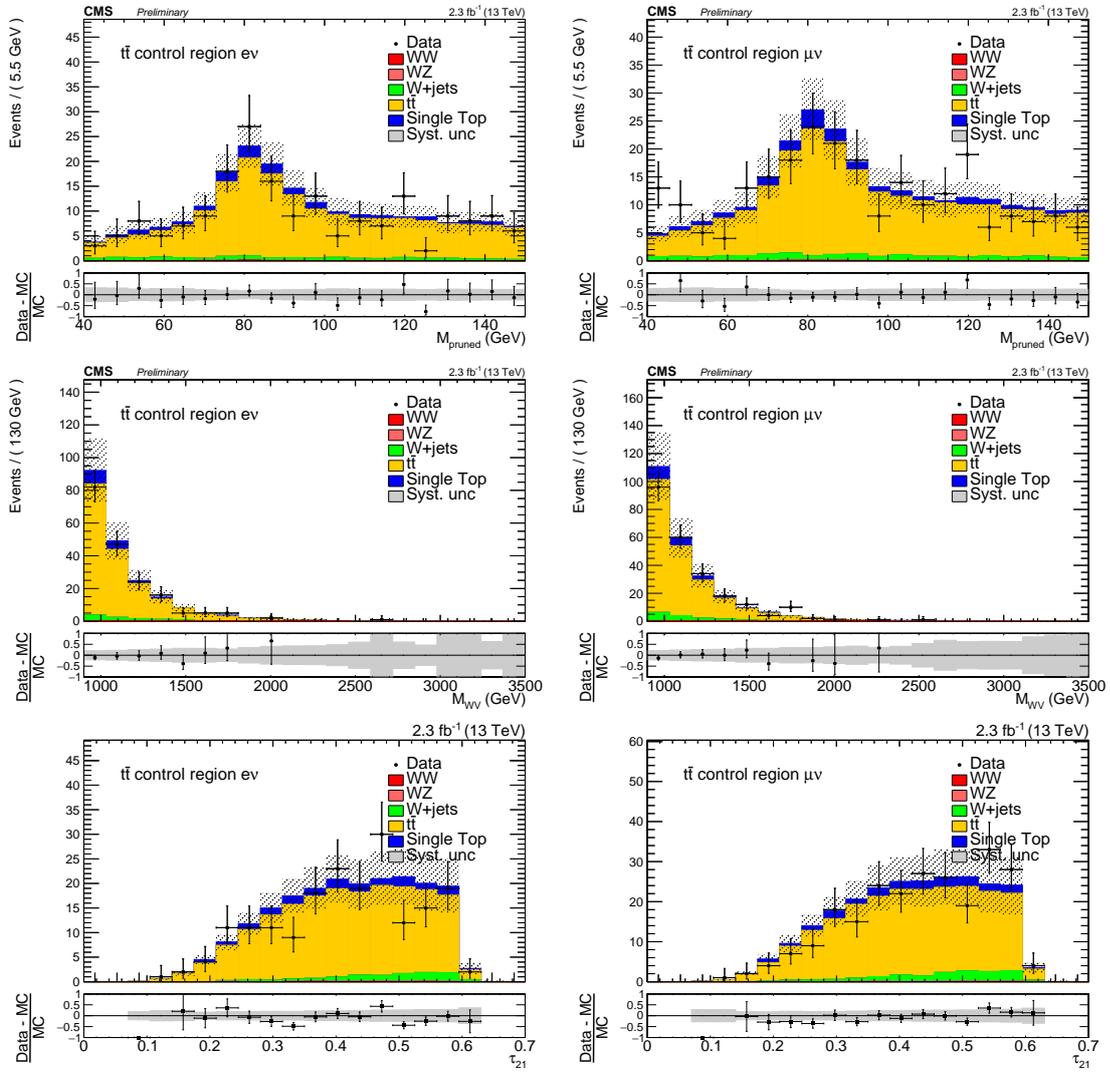
where  $N^{\text{veto}}$  is the number of events with b-tag veto applied and  $N^{\text{incl}}$  is an inclusive number of events in this control region. It should quantify how well the  $t\bar{t}$  background is modelled in this

## 7 Background estimation



**Figure 7.3:** Comparison between data and simulation of the  $M_{\text{pruned}}$  (top),  $M_{WV}$  (middle) and  $\tau_{21}$  (bottom) distributions in the W+jets control region. The electron channel is shown on the left, while the muon channel is shown on the right.

## 7.2 Validation of background in control regions



**Figure 7.4:** Comparison between data and simulation of the  $M_{\text{pruned}}$  (top),  $M_{WV}$  (middle) and  $\tau_{21}$  (bottom) distributions in the  $t\bar{t}$  control region. The electron channel is shown on the left, while the muon channel is shown on the right.

region of the phase space. Results are found to be consistent with unity. In the muon channel the ratio is  $1.01 \pm 0.06$  and in the electron channel  $1.08 \pm 0.07$ .

Furthermore, the question may arise whether the background composition in the signal region is well represented in the various control regions and whether the  $t\bar{t}$  control region II is representative of the signal region, which contains real W bosons. To check this, the  $t\bar{t}$  background is split into two categories: one category contains events with the hadronic W candidate matched to the generator level W boson ( $\Delta R < 0.8$ ) and the other category contains events where the hadronic W candidate is not matched to the generator level W boson. This is done in the  $t\bar{t}$  control region (with at least one b-tagged AK4 jet) (see Figure 7.5 ) and in the  $t\bar{t}$  control region II (no requirement on b-tagging, see Figure 7.6) in the extended range  $40 \text{ GeV} < M_{\text{pruned}} < 200 \text{ GeV}$ . The fraction of unmatched jets is high at low jet masses where the probability to pick up a random jet is quite high. The fraction of matched jets is increasing around 80 GeV and 175 GeV. This trend is present in both of  $t\bar{t}$  control regions and thus is independent from the b-tag veto. To conclude, for the  $t\bar{t}$  background in the majority of events the reconstructed hadronic W candidate is associated to a generator-level W boson. This should be expected from the selection used in the analysis: requirements like  $p_{T,W\text{lep}} > 200 \text{ GeV}$ ,  $p_{T,\text{jet}} > 200 \text{ GeV}$  and  $M_{WV} > 900 \text{ GeV}$  strongly favour event topologies where the two top quarks recoil against each other with considerable momentum, leading to two distinct hemispheres, each associated to one of the W candidates. Starting from  $M_{\text{pruned}} \approx 60 \text{ GeV}$  the majority of reconstructed fat jets will contain at least a substantial fraction of the products from the hadronic W boson decay. Similarly, most of products from the top quark decay should be inside of the jet with  $M_{\text{pruned}}$  around the top mass ( $\approx 175 \text{ GeV}$ ) [22]. While there is a somewhat higher fraction of non-associated jets in the intermediate region, the dominant contribution is still matched, corresponding to the case where the jet contains a substantial subset, but not the entirety of the top decays hadronization products <sup>1</sup>.

The composition of the  $t\bar{t}$  background is the same in the signal region as shown in Figure 7.7.

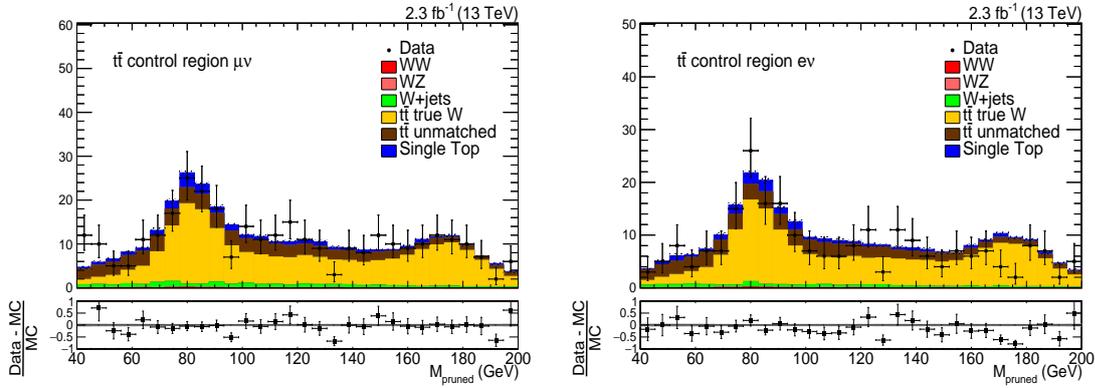
### 7.3 Background shapes

As it was already mentioned in Section 1.3 limits on aTGC are extracted from the  $M_{WV}$  distribution. The shapes of the  $t\bar{t}$ , single-top and diboson contributions are represented by analytic shapes fitted to the simulation, while the W+jets background is derived from the sideband data.

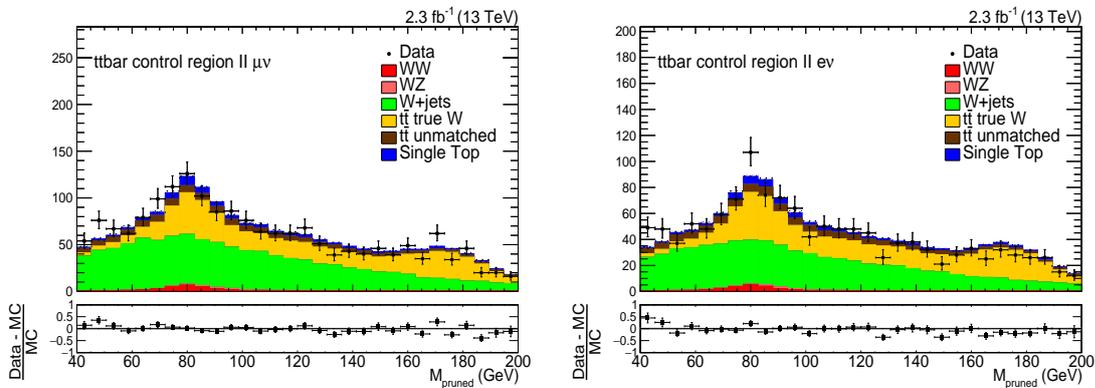
In the signal region the following analytic functions are used to model the different back-

---

<sup>1</sup>It should be noted that a purer sample with tops can be obtained with a cut on  $\tau_3/\tau_2$  or by applying top-tagging algorithm [192]. However, at this point already the requirement on  $M_{\text{pruned}}$  to be around the top mass significantly enriches the number of real W bosons.

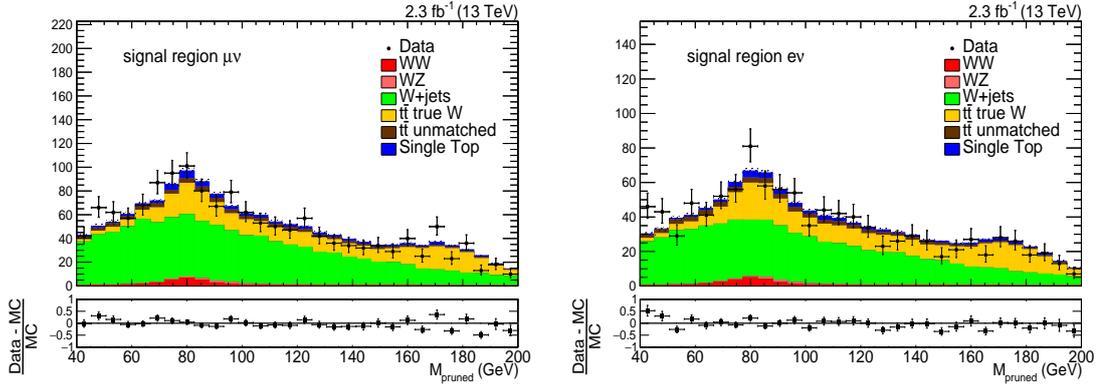


**Figure 7.5:** Comparison between data and the simulation in the  $t\bar{t}$  control region splitting the  $t\bar{t}$  background into matched and unmatched categories, muon channel on left and electron channel on the right. Only statistical uncertainties are shown for the simulation.



**Figure 7.6:** Comparison between data and the simulation in the  $t\bar{t}$  control region II splitting the  $t\bar{t}$  background into matched and unmatched categories, muon channel on left and electron channel on the right. Only statistical uncertainties are shown for the simulation.

## 7 Background estimation



**Figure 7.7:** Comparison between data and the simulation in the signal region (extended) splitting the  $t\bar{t}$  background into matched and unmatched categories, muon channel on left and electron channel on the right. Only statistical uncertainties are shown for the simulation.

ground processes:

$$\begin{aligned}
 \mathbf{t\bar{t}} \quad F_{\text{ExpN}}(x) &= e^{a_1 x + b_1/x} \\
 \mathbf{single-top} \quad F_{\text{Exp}}(x) &= e^{a_2 x} \\
 \mathbf{diboson} \quad F_{\text{ExpN}}(x) &= e^{a_3 x + b_3/x} \\
 \mathbf{W+jets} \quad F_{\text{ExpN}}(x) &= e^{a_4 x + b_4/x}
 \end{aligned} \tag{7.6}$$

Fits to the simulation samples in the signal region for  $t\bar{t}$ , W+jets, single top and diboson processes are shown in Figures 7.8 and 7.9.

The same functions are used to model the simulated samples in the sideband region, with the exception of the diboson component, which has a small contribution in the sideband and is only represented by a simple exponential function. The sideband region fits to the simulation samples are shown in Appendix G.

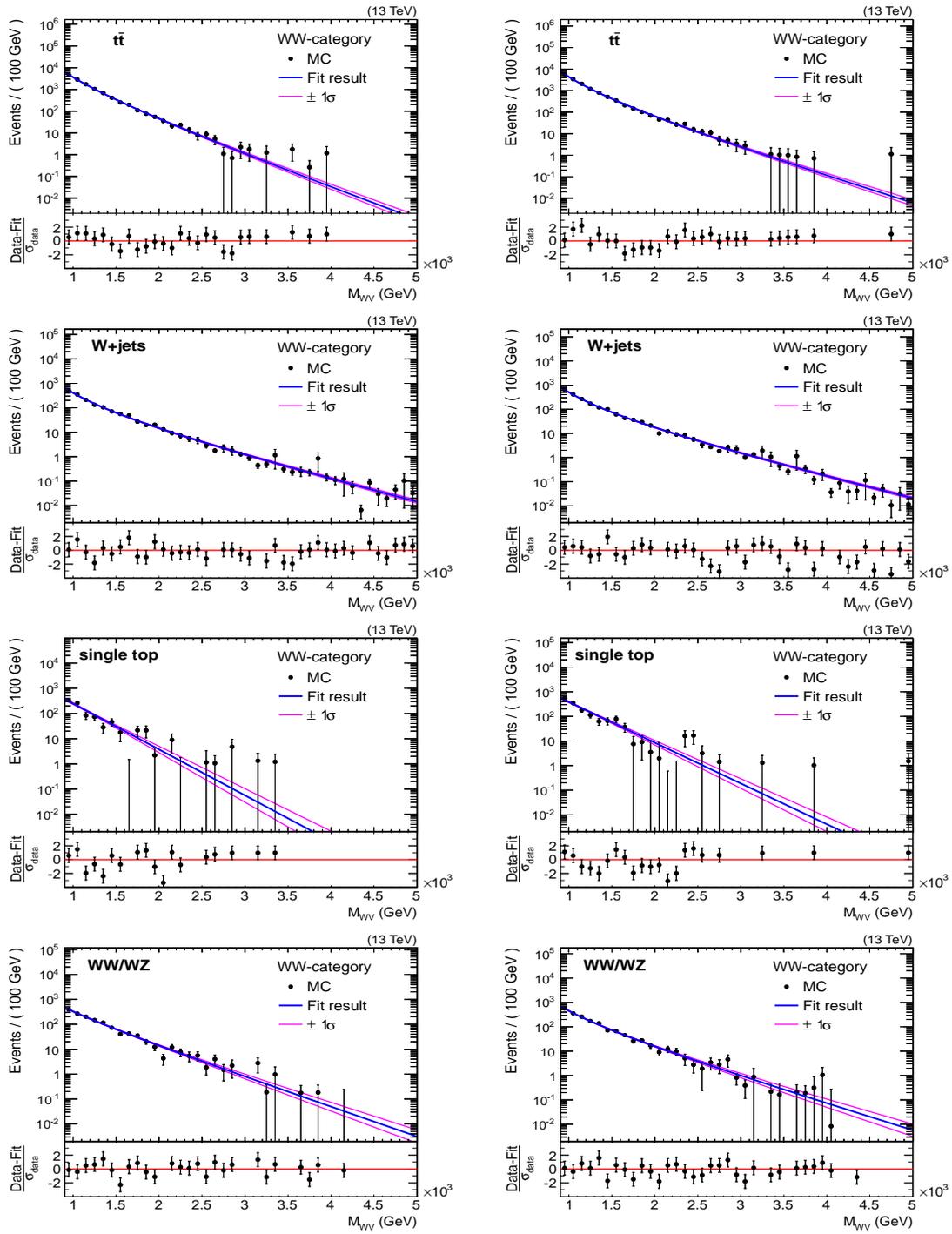
Then the data in the sideband are fitted to the sum of the four components. The normalizations of  $t\bar{t}$  and W+jets are taken from the fit of the  $M_{\text{pruned}}$  spectrum. The shapes of the  $t\bar{t}$ , single-top and diboson components are fixed to the values extracted from simulation. The shape parameters of the W+jets component are allowed to freely float. The result of the fit is shown in Figure 7.10.

The shape of W+jets measured in the sideband is then transferred to the signal region via  $\alpha$ -function ( $\alpha^{MC}(M_{WV}) = F_{W+jets}^{SR,MC} / F_{W+jets}^{SB,MC}$ ) and the predicted shape of W+jets in the signal region is given by:

$$F_{W+jets}^{SR,predicted} = F_{W+jets}^{SB,data} \cdot \alpha^{MC}(M_{WV}) \tag{7.7}$$

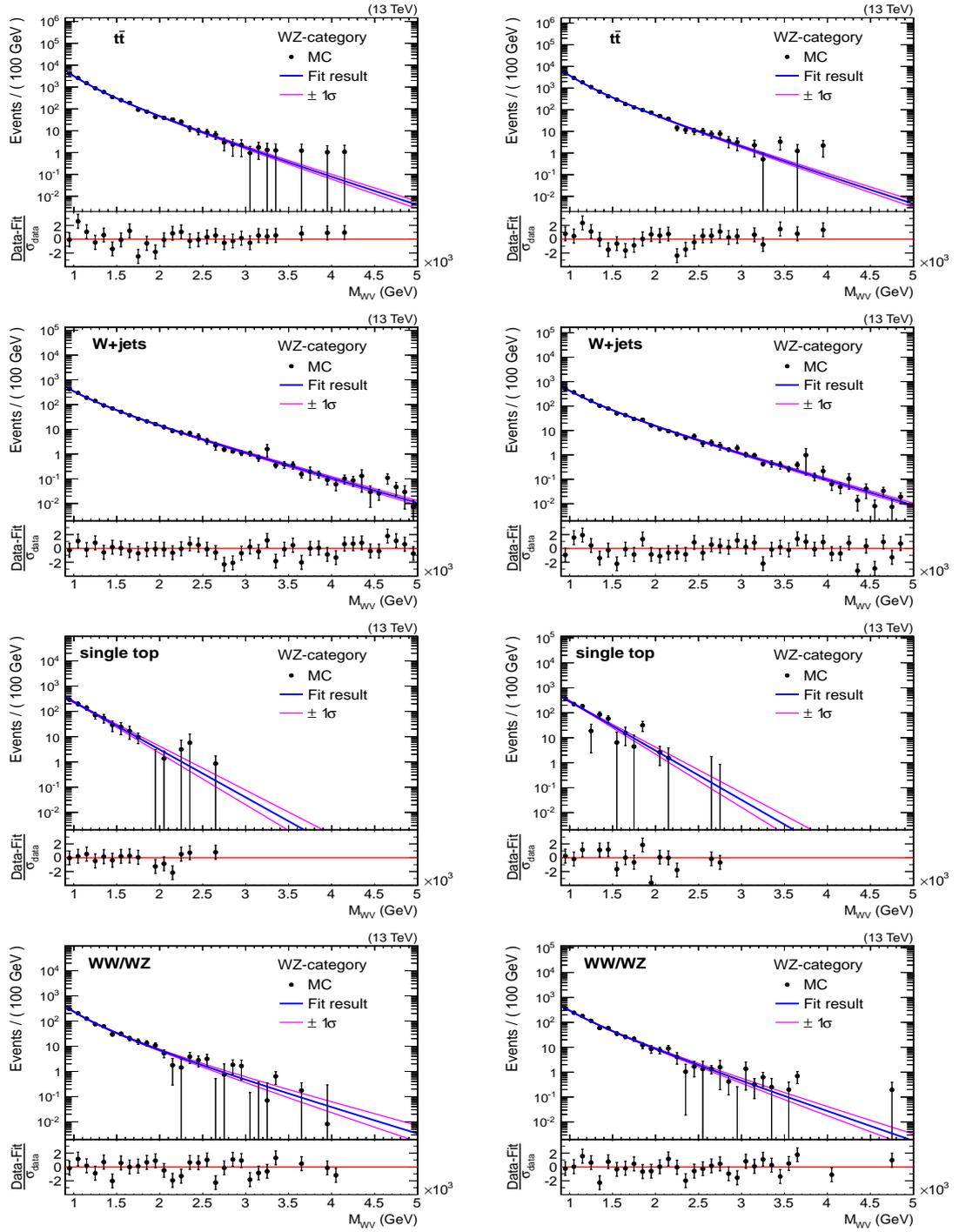
where  $F_{W+jets}^{SB,data}$  is the shape of the W+jets background estimated from the fit in the sideband region,  $F_{W+jets}^{SR,MC}$  and  $F_{W+jets}^{SB,MC}$  are the shapes of the W+jets extracted from the MC in the signal region and the sideband accordingly.

The total background shape is then estimated as:

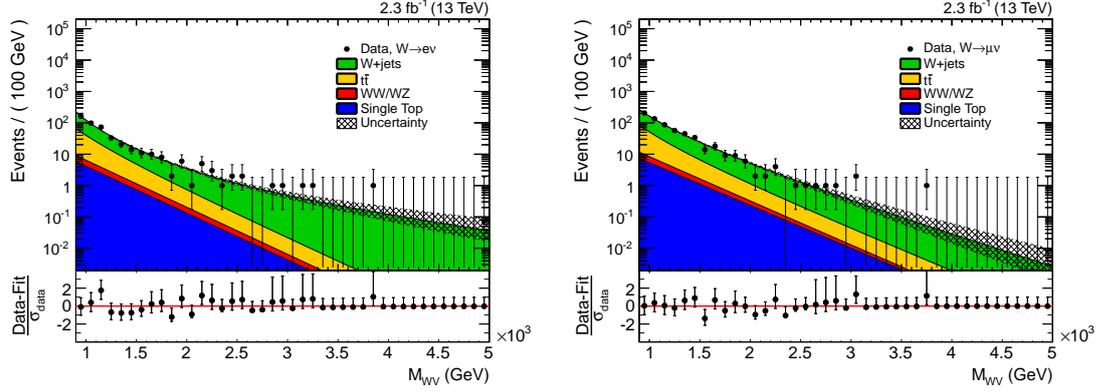


**Figure 7.8:** MC-data and shapes of the  $M_{WV}$  distributions in the signal region, WW-category, for electron- (left) and muon-channel (right). From top to bottom:  $t\bar{t}$ , W+jets, single top, diboson.

## 7 Background estimation



**Figure 7.9:** MC-data and shapes of the  $M_{WV}$  distributions in the signal region, WZ-category, for electron- (left) and muon-channel (right). From top to bottom:  $t\bar{t}$ , W+jets, single top, diboson.



**Figure 7.10:** Shape and data of the  $M_{WV}$  background in the sideband region for electron- (top) and muon-channel (bottom).

$$\begin{aligned}
 F_{tot}^{SR}(M_{WV}) &= F_{W+jets}^{SB,data}(M_{WV}) \cdot \alpha^{MC}(M_{WV}) + F_{t\bar{t}}^{SR,MC}(M_{WV}) + F_{singletop}^{SR,MC}(M_{WV}) + \\
 &+ F_{diboson}^{SR,MC}(M_{WV}) = F_{W+jets}^{SB,data}(M_{WV}) \cdot \frac{F_{W+jets}^{SR,MC}(M_{WV})}{F_{W+jets}^{SB,MC}(M_{WV})} + F_{t\bar{t}}^{SR,MC}(M_{WV}) + \\
 &+ F_{singletop}^{SR,MC}(M_{WV}) + F_{diboson}^{SR,MC}(M_{WV})
 \end{aligned}$$

where  $SR$  denotes functions derived in the signal region and  $SB$  functions derived in the  $W$ +jets control region.

The  $\alpha$ -function is shown in Figure 7.11. The plot also shows the  $\alpha$ -function derived with an alternative analytical function to describe the shape of the  $W$ +jets, which is given by:

$$F_{alt}^{W+jets}(x) = e^{-x/(c+dx)} \quad (7.8)$$

The uncertainty due to the choice of function used for modelling of  $W$ +jets is well covered by the uncertainty of  $\alpha$ -function.

The final background estimation compared with data is shown in Figure 7.12.

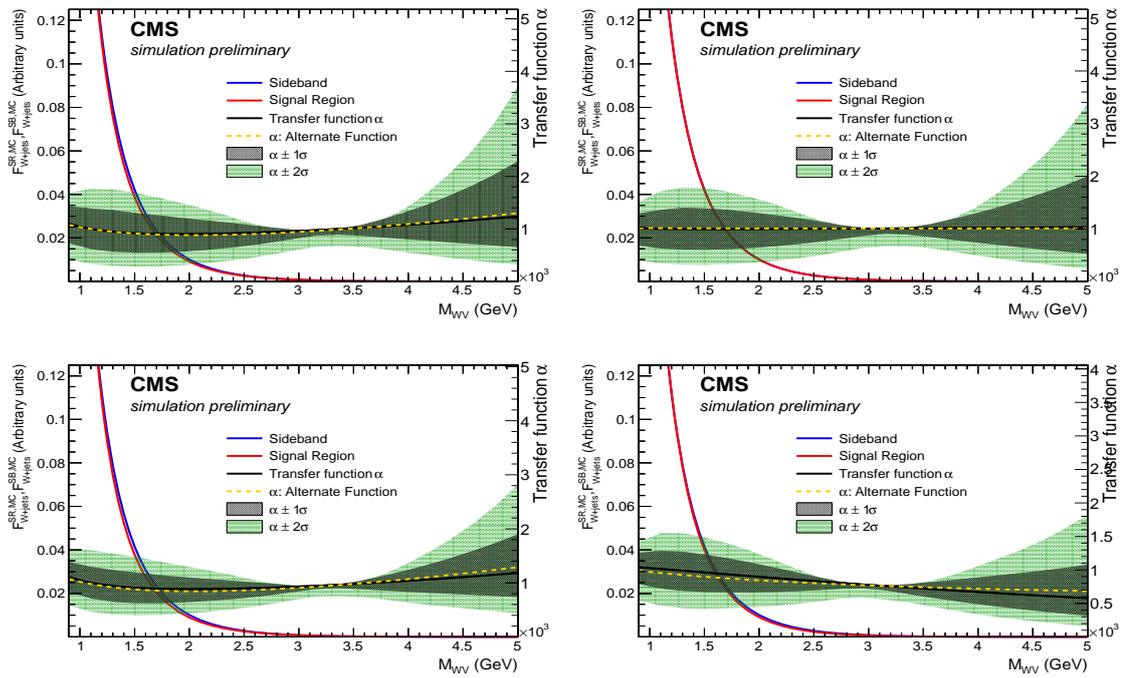


Figure 7.11: Alpha ratio for the WW-category (top) and the WZ-category (bottom) in the electron- (left) and muon-channel (right) [55].

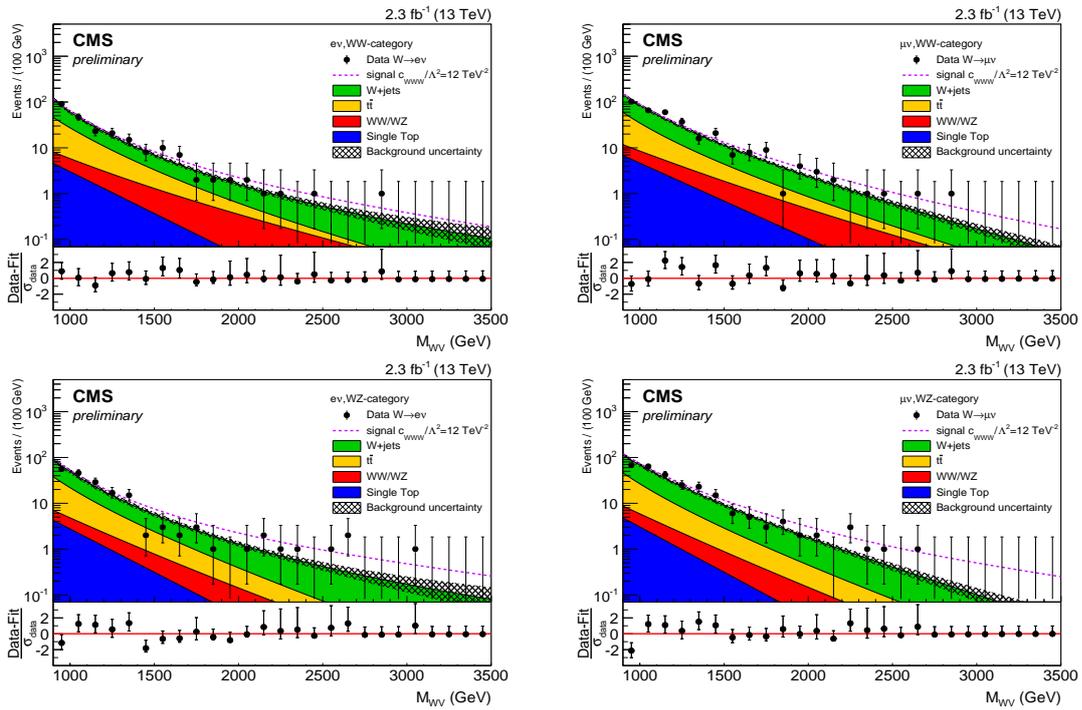


Figure 7.12: Data and final shapes of the  $M_{WV}$  background for electron (left) and muon channel (right) in the WW- (top) and WZ-category (bottom) [55].



## 8 Limits on anomalous couplings

Limits on anomalous triple gauge coupling are presented here. First the method used and the statistical model are described. The description largely follows reference [127]. Then systematic uncertainties and the way they are included in the statistical model are discussed. In the end two-dimensional and one-dimensional limits on aTGC are presented using the EFT and LEP (see Section 1.3) parametrization.

### 8.1 Methodology

Generally, the procedure of setting limits on anomalous triple gauge couplings refers to the problem of *interval estimation*. This generally implies that two numerical limits and a level of confidence about the true value of the parameter of interest to lie between these limits are provided. It should be mentioned that there is no unique prescription how to construct an interval estimation. In fact, interval construction depends already on the statistical paradigm: *Bayesian* or *frequentist*. In the Bayesian approach the confidence level means the probability of the parameter of interest to lie within a certain interval given someone prior belief and the observed data. In the frequentist approach it means that if  $N$  datasets are collected under the same conditions and the same measurement is performed the given fraction of these measurements will yield a confidence interval that contains the true value of the parameter of interest when  $N \rightarrow \infty$ . Different considerations might be taken into account for the interval construction: the length of the interval, physical boundaries, behaviour with respect to systematic uncertainties etc. An overview of different methods for interval construction can be found in reference [127]. In this thesis the *delta log-likelihood* method is used.

Imagine that  $N$  independent measurements are done and a quantity  $x$  is measured:  $\vec{x} = \{x_1, x_2, \dots, x_N\}$ .  $f(x, \vec{\theta})$  is the probability of obtaining result  $x$  given model parameters  $\vec{\theta}$ . The *likelihood function* is a joint probability density function for the observed dataset:

$$L(\vec{x}, \vec{\theta}) = \prod_{i=1}^N f(x_i, \vec{\theta}) \quad (8.1)$$

The probability density function should be normalized to unity:

$$\int f(x, \vec{\theta}) dx = 1 \text{ so that } \int L(\vec{x}, \vec{\theta}) dx_1 dx_2 \dots dx_N = 1 \quad (8.2)$$

The likelihood function can be used for the parameter estimation. This is usually known as the maximum likelihood estimate (MLE). MLE of model parameters  $\vec{\theta}$  are the values  $\hat{\vec{\theta}}$  which deliver global maximum of the likelihood function. This assumption is intuitively reasonable:

one expects higher values of likelihood for true (or close to true) parameter values than for wrong parameter values. Often the logarithm of the likelihood function (*log-likelihood* function) is used and the minimum of the negative log-likelihood is searched for:

$$-\ln L(\vec{x}, \vec{\theta}) = -\sum_{i=1}^N \ln f(x_i, \vec{\theta}) \quad (8.3)$$

Thus MLE can be obtained from the necessary condition for the minimum of the negative log-likelihood:

$$-\frac{\partial L(\vec{x}, \vec{\theta})}{\partial \theta_j} = 0 \quad (8.4)$$

where  $j = \overline{1, M}$  where  $M$  is the number of parameters in the model.

An important quantity for interval construction in this thesis is the so-called *likelihood ratio*:

$$\lambda(\vec{x}, \vec{\theta}) \equiv \frac{L(\vec{x}, \vec{\theta})}{L(\vec{x}, \hat{\vec{\theta}})} \quad (8.5)$$

According to Wilks's theorem [193] under standard regularity conditions  $-2 \ln \lambda(\vec{x}, \vec{\theta})$  is asymptotically distributed as  $\chi^2$  with  $d$  degrees of freedom. This provides a simple way to construct  $1 - \alpha$  confidence level interval by taking those  $\vec{\theta}$  for which:

$$-2 \ln \lambda(\vec{x}, \vec{\theta}) \leq \chi_{d, 1-\alpha}^2 \quad (8.6)$$

where  $\chi_{d, 1-\alpha}^2$  is the  $(1 - \alpha)$ -quantile<sup>1</sup> of a  $\chi^2$  distribution [127] with  $d$  degrees of freedom. The  $\chi^2$ -distribution  $f_k(x)$  with  $k$  degrees<sup>2</sup> of freedom is given by:

$$f_k(x) = \begin{cases} \frac{2^{-k/2}}{\Gamma(k/2)} x^{k/2-1} e^{-x/2}, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases} \quad (8.7)$$

where  $\Gamma(z)$  is the gamma function:  $\Gamma(z) = \int_0^\infty \xi^{z-1} e^{-\xi} d\xi$ .

Since  $\ln(L(\vec{x}, \vec{\theta})/L(\vec{x}, \hat{\vec{\theta}})) = \ln(L(\vec{x}, \vec{\theta})) - \ln(L(\vec{x}, \hat{\vec{\theta}}))$ ,  $-2 \ln \lambda(\vec{x}, \vec{\theta})$  is often referred as *delta log-likelihood*:

$$\Delta NLL = -\ln(\lambda(\vec{x}, \vec{\theta})) \quad (8.8)$$

## 8.2 Statistical model

In order to set limits at 95 % confidence level (CL) the unbinned likelihood is constructed as a product of likelihood functions in WW- and WZ-category, in electron and muon channel. Diboson mass is used as an observable ( $\vec{x}$  in (8.6)). Parameters  $\vec{\theta}$  include aTGC ( $c_k$ ,  $k = \overline{1, 3}$ ) and

<sup>1</sup> $x_\gamma$  is  $\gamma$ -quantile of the probability distribution  $p(\xi)$  of a random variable  $\xi$  if  $\gamma = \int_{-\infty}^{x_\gamma} p(\xi) d\xi$

<sup>2</sup> $k$  is equal to the dimension of  $\vec{\theta}$

nuisance parameters  $\vec{\zeta}$ . The probability density to observe an event with a given  $M_{WV}$  can be given by:

$$p_i = \frac{F_{signal}(c_k, M_{WV}) + \sum_j b_j \cdot f_{bkq,j}(M_{WV})}{N_{SM}^{NNLO} \cdot N_{aTGC}(c_k) + \sum_j b_j} \quad (8.9)$$

where  $F_{signal}(c_k, M_{WV})$ ,  $N_{aTGC}(c_k)$  and  $N_{SM}^{NNLO}$  are defined in (6.1),  $j$  runs over all background process ( $t\bar{t}$ , W+jets, single top),  $f_{b,j}(M_{WV})$  the normalized distribution of the background obtained in Chapter 7,  $b_j$  are expected yields of background process. If a number of events are observed by repeating an experiment with identical conditions, the observed event rate fluctuates according to a Poisson distribution. Thus, the likelihood function is given by:

$$L(c_k, \vec{\theta}) = \rho(\vec{\zeta}) \prod_{n=1}^4 L_n(c_k, \vec{\theta}) \quad (8.10)$$

where  $\vec{\zeta}$  denotes a set of nuisance parameters,  $\rho(\vec{\zeta})$  is the probability density function for nuisance parameters and  $L_n(c_k, \vec{\theta})$  is the likelihood function in a given channel or category (denoted as  $n$ ):

$$L_n(c_k, \vec{\theta}) = e^{-\lambda_n} \frac{\lambda_n^{N_n}}{N_n!} \prod_{i=1}^{N_n} p_i \quad (8.11)$$

where  $\lambda_n$  is the expected number of events,  $\lambda_n = N_{SM}^{NNLO} \cdot N_{aTGC}(c_k) + \sum_j b_j$  and  $N_n$  is the number of observed events in a given category or channel.

### 8.2.1 Systematic uncertainties

Two type of nuisance parameters are included: ones that account uncertainties of the shape and others accounting uncertainties of normalization. The first ones are included by changing slope parameters of the exponential functions used describe  $M_{WV}$  spectrum for the signal and background:  $\alpha \rightarrow \alpha + \delta$  where  $\alpha$  corresponds to the slope parameter and  $\delta$  is the introduced nuisance parameter that has Gaussian distribution with mean 0 and a specified width.

The shape uncertainty on the W+jets is taken into account by inflating the parametric errors to well cover the estimate from the alternative function (see (7.2)). In order to do that the covariance matrix of parameters is constructed and parameters are translated to the eigenspace of the covariance matrix. These newly defined parameters are uncorrelated and uncertainties are propagated to them. The same is done to take into account the shape uncertainty on the  $t\bar{t}$  contribution which is estimated as statistical error of the  $t\bar{t}$  shape parameters (see (7.6)) as fitted to the simulation. No shape uncertainties are considered for the signal top background as this found to be a minor contribution. Uncertainties on slopes of the exponentials in the signal function (see (6.1)) are estimated to be of 5% for all aTGCs in all categories except  $c_B$  in the WZ category (15 %) (see Table 6.1).

The normalization uncertainties are introduced by modifying expected yields so that:  $\beta \rightarrow \beta \cdot e^{\xi \ln \epsilon}$  where  $\beta$  is the expected yield,  $\xi$  is the introduced nuisance parameter which has Gaussian distribution with mean equal to 0 and width equal to 1 and  $\epsilon$  corresponds to the uncertainty: for example, 20% uncertainty corresponds to  $\epsilon = 1.2$ . This modification is done

in order to constrain the expected yield with log-normal distribution [194]<sup>3</sup>. Normalization uncertainties given in Table 5.5 are considered. The normalization uncertainty for the signal function is estimated as maximum uncertainty of SM WW and WZ contributions.

Normalization uncertainties are listed in Table 5.5. Jet energy scale uncertainties are treated as anti-correlated between WW- and WZ-categories<sup>4</sup>.

Of these uncertainties, the W+jets background shape uncertainty together with the larger normalization uncertainties have a significant effect on the limit setting power.

### 8.3 Results

Limits on aTGCs are obtained by using the likelihood definition in (8.11) and constructed from inequation (8.6). In order to be far from unitary bound only evens with  $M_{WV} < 3.5$  TeV are used. Data and background estimates are given in Table 8.1.

**Table 8.1:** Summary of background and signal yields in WW- and WZ-categories. Uncertainties for the single-top, diboson and  $t\bar{t}$  contributions are evaluated as described in Section 5.4, while the uncertainty on W+jets is derived from the statistical uncertainty of the  $M_{\text{pruned}}$ -fit and the fit with the alternative function [55].

category $M_{\text{pruned}}$ (GeV)	electron channel		muon channel	
	WW [65,85]	WZ [85,105]	WW [65,85]	WZ [85,105]
W+jets	$124 \pm 17$	$103 \pm 16$	$192 \pm 20$	$164 \pm 20$
$t\bar{t}$	$73 \pm 17$	$58 \pm 13$	$90 \pm 21$	$71 \pm 17$
single top	$10.9 \pm 1.4$	$9.8 \pm 1.2$	$17.8 \pm 2.3$	$10.6 \pm 1.4$
diboson (SM)	$15.8 \pm 2.2$	$9.3 \pm 1.3$	$20.6 \pm 3.0$	$12.2 \pm 1.8$
Total expected (SM)	$224 \pm 24$	$180 \pm 21$	$320 \pm 29$	$258 \pm 26$
diboson $\frac{c_{WW}WW}{\Lambda^2} = 12 \text{ TeV}^{-2}$	$36.2 \pm 5.1$	$39.9 \pm 5.7$	$50.8 \pm 7.3$	$55.4 \pm 8.0$
diboson $\frac{c_W}{\Lambda^2} = 20 \text{ TeV}^{-2}$	$52 \pm 7$	$69 \pm 10$	$72 \pm 10$	$91 \pm 13$
diboson $\frac{c_B}{\Lambda^2} = 60 \text{ TeV}^{-2}$	$41.5 \pm 5.9$	$20.1 \pm 2.9$	$57.0 \pm 8.2$	$26.8 \pm 3.9$
Data	234	183	340	265

Limits are obtained with EFT approach and also in terms of the parametrization commonly

<sup>3</sup>this is done in order to restrict parameters to positive values.

<sup>4</sup>This is actually an approximation assuming that most of the events in the signal region are concentrated around W(Z) mass peak. When jet energy scale changes according its uncertainty the W(Z) mass peaks shifts increasing the number of events in one category and decreasing the number of events in another category. However, edges of the signal region are also shifting when the jet energy scale is changing.

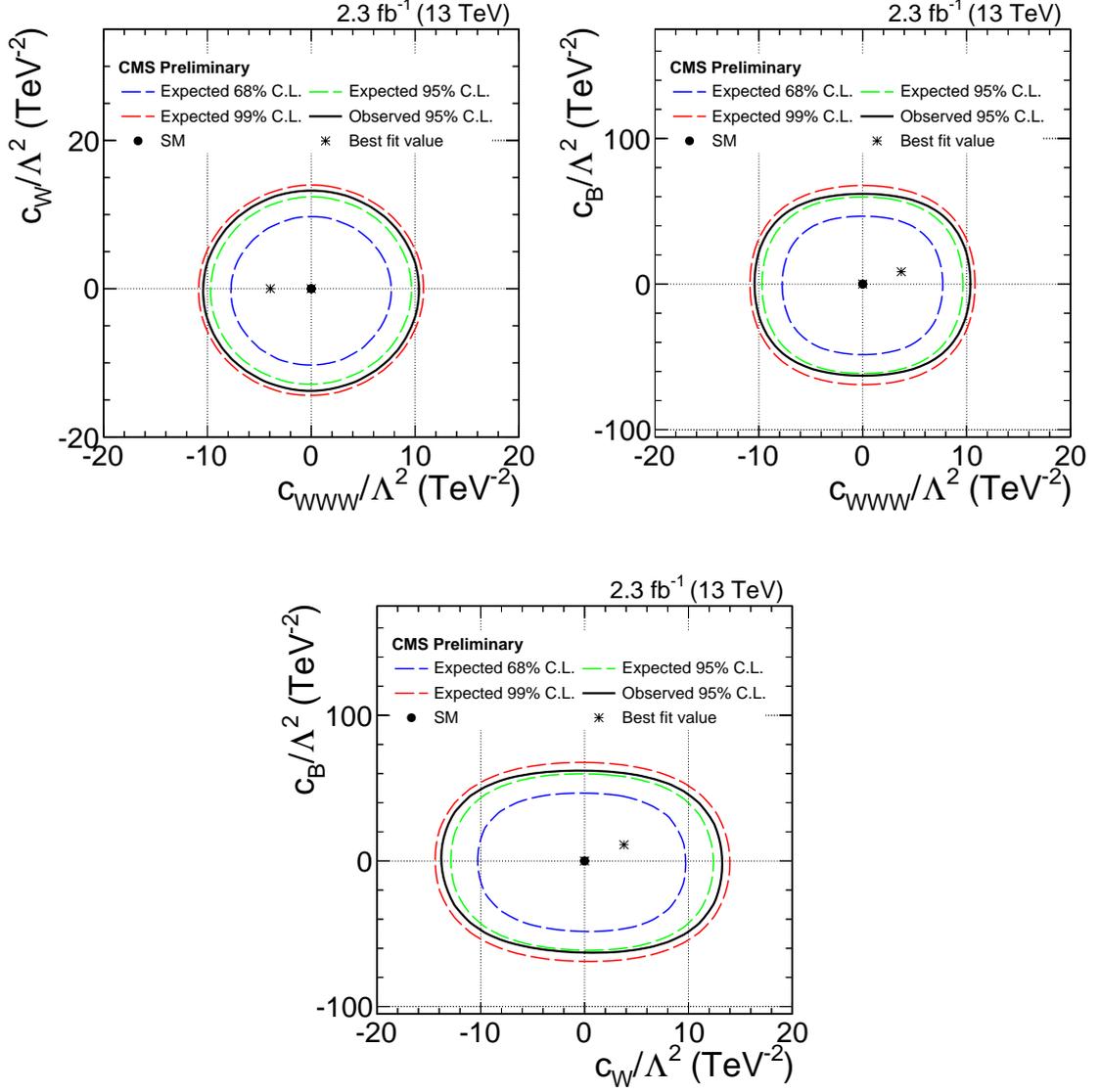
used at LEP (see Section 1.3). The relation between 2 parametrizations is given by:

$$\begin{aligned}
\lambda_Z &= \frac{c_{WWW}}{\Lambda^2} \frac{3g^2 m_W^2}{2} \\
\Delta g_1^Z &= \frac{c_W}{\Lambda^2} \frac{m_Z^2}{2} \\
\Delta \kappa_Z &= (c_W - c_B \tan^2 \theta_W) \frac{m_W^2}{2}
\end{aligned} \tag{8.12}$$

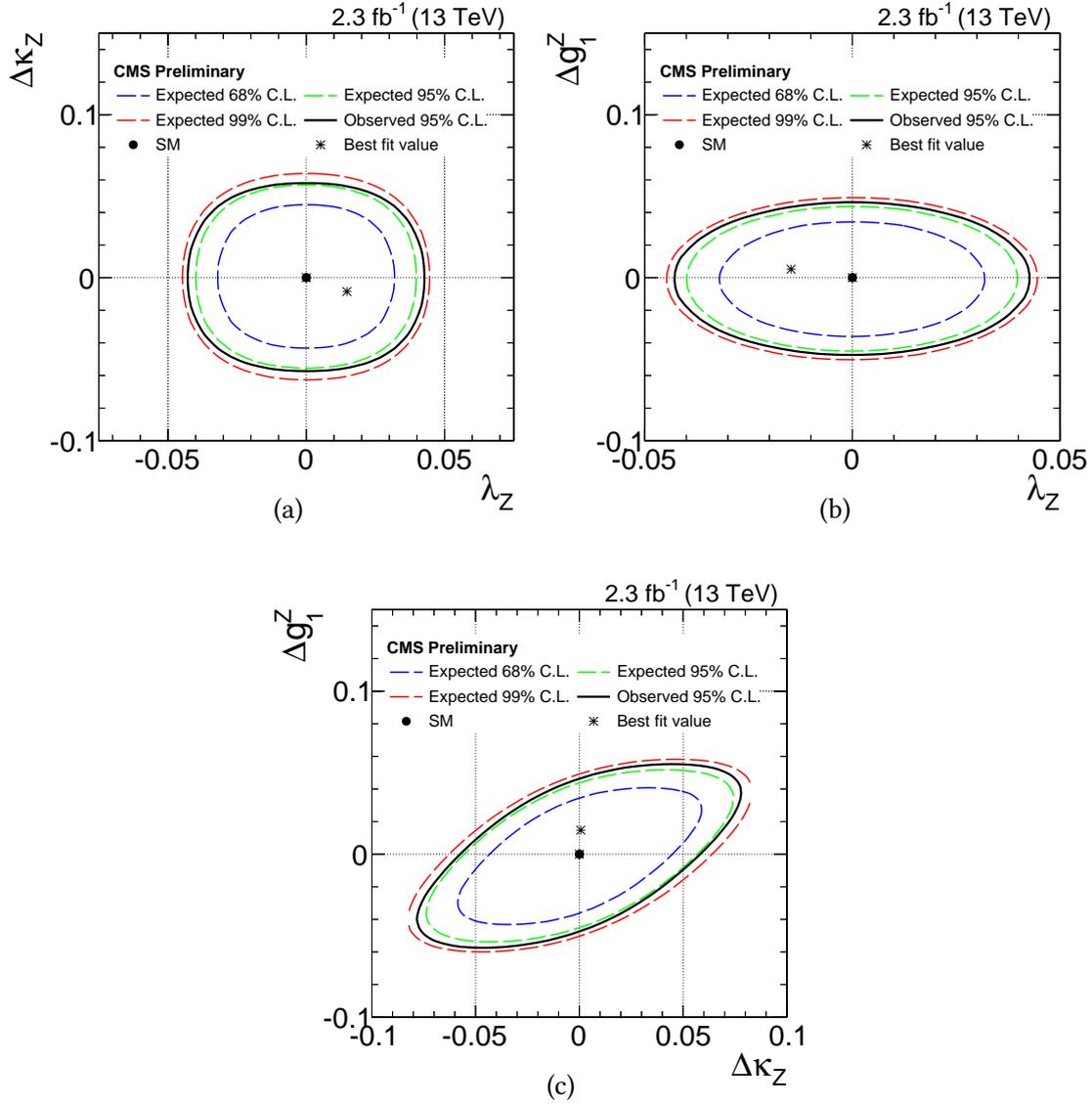
Observed and expected limits on pairwise combinations of the couplings, with the remaining coupling set to zero, are shown in Figures 8.1 and 8.2. Limits on a single aTGC are given in Table 8.2.  $\Delta NLL$  distributions are shown in Figure 8.3 and in Appendix I. 2-dimensional limits at 95% correspond to the value of  $\chi^2 = 5.99$  (2 degrees of freedom) and 1 dimensional limits to  $\chi^2 = 3.84$  (1 degree of freedom). Expected limits are derived by doing pseudoexperiments normalized to the expected yield.

**Table 8.2:** Expected and observed limits at 95% C.L. on single anomalous couplings (other couplings set to zero).

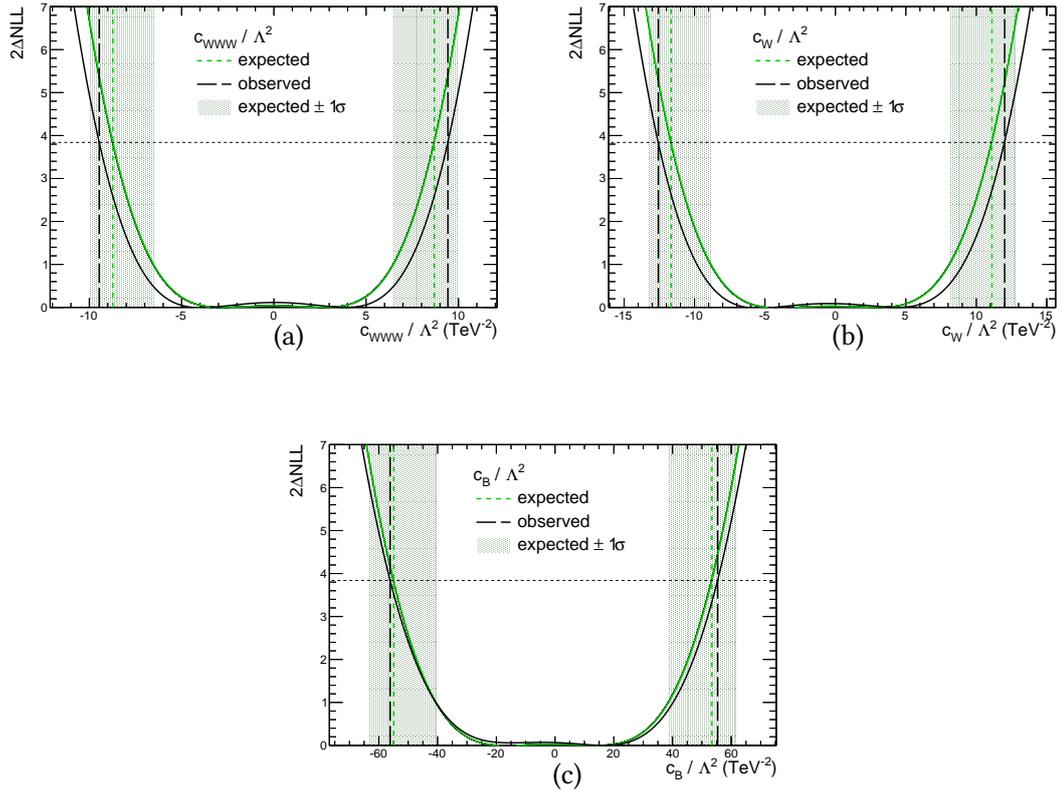
	aTGC	expected limit	observed limit
EFT param.	$\frac{c_{WWW}}{\Lambda^2}$ (TeV <sup>-2</sup> )	[-8.73 , 8.70]	[-9.46 , 9.42]
	$\frac{c_W}{\Lambda^2}$ (TeV <sup>-2</sup> )	[-11.7 , 11.1]	[-12.6 , 12.0]
	$\frac{c_B}{\Lambda^2}$ (TeV <sup>-2</sup> )	[-54.9 , 53.3]	[-56.1 , 55.4]
Vertex param.	$\lambda$	[-0.036 , 0.036]	[-0.039 , 0.039]
	$\Delta g_1^Z$	[-0.066 , 0.064]	[-0.067 , 0.066]
	$\Delta \kappa_Z$	[-0.038 , 0.040]	[-0.040 , 0.041]



**Figure 8.1:** Two dimensional limits on the aTGC-parameters. Shown are the expected contours for 68% C.L. (blue), 95% C.L. (green) and 99% C.L. (red), for  $\frac{c_{WWW}}{\Lambda^2} - \frac{c_W}{\Lambda^2}$  (a),  $\frac{c_{WWW}}{\Lambda^2} - \frac{c_B}{\Lambda^2}$  (b) and  $\frac{c_W}{\Lambda^2} - \frac{c_B}{\Lambda^2}$  (c). The black line shows the region compatible with the observed data at 95% C.L. [55].



**Figure 8.2:** Two dimensional limits on the aTGC-parameters. Shown are the expected contours for 68% C.L. (blue), 95% C.L. (green) and 99% C.L. (red), for  $\Delta\kappa_Z - \lambda_Z$  (a),  $\Delta g_1^Z - \lambda_Z$  (b) and  $\Delta g_1^Z - \Delta\kappa_Z$  (c). The black line shows the region compatible with the observed data at 95% C.L. [55].



**Figure 8.3:**  $\Delta NLL$  distributions for aTGC-parameters:  $\frac{c_{WWW}}{\Lambda^2}$  (a),  $\frac{c_W}{\Lambda^2}$  (b),  $\frac{c_B}{\Lambda^2}$  (c). Shown are the expected (green straight line) and observed (black straight line) distributions. Intersection between black dashed (green dashed) lines corresponds to observed (expected) 95% C.L. The shaded area shows the region between 99 % C.L. and 68 % C.L. ( $1\sigma$ -band around the expected limit). The observed and expected limits are in agreement with  $1\sigma$ -band [191].

## Conclusions and outlook

In this the thesis a search for new physics through anomalous triple gauge couplings in WW and WZ decays with data recorded by the CMS experiment from proton-proton collisions at center-of-mass energy of 13 TeV was presented. The analysis was performed in the semileptonic channel with a W or Z boson decaying to hadrons and another W boson decaying to a charged lepton (an electron or a muon) and a neutrino. The so called boosted topology was exploited where decay products of the W or Z boson are found in a single large radius jet but not in two jets with smaller radius. This topology should contain most of the signal events because effects from anomalous triple gauge couplings are prominent at high W or Z boson momenta [20] where decay products are merged into a single jet with large radius in most of the cases [161]. Limits were extracted from the diboson mass distribution modelled by analytical functions. To describe anomalous triple gauge couplings the effective field theory approach was used which has an advantage that it is done in a theoretically consistent way.

Most of the previous analyses were focused on the fully leptonic channel which profits from less background, but has smaller branching fraction of the signal. In the fully leptonic channel the scale of the process can be difficult to reconstruct which makes it more complicated to control the validity of the unitarity condition while the semileptonic channel provides full kinematic reconstruction.

To identify large-radius jets from hadronically decaying W or Z bosons and to avoid contamination from QCD-jets jet-substructure techniques were used: events were selected with certain requirements on the mass of the jet with removed soft and wide angle radiations and a substructure variable that distinguishes one-prong jets from two-prong jets.

The two largest background contributions in the analysis come from  $W$ +jets and  $t\bar{t}$  production. Since the  $W$ +jets background is rather hard to model with simulation a data-driven technique was used following the approach of related diboson resonance searches [17]. The shape of the  $t\bar{t}$  background was taken from the simulation at next-to-leading order QCD as a decent agreement between data and the simulation was found in two control regions enriched with the  $t\bar{t}$  contribution. The contribution from the Standard Model diboson process and anomalous triple gauge couplings was modelled at leading order.

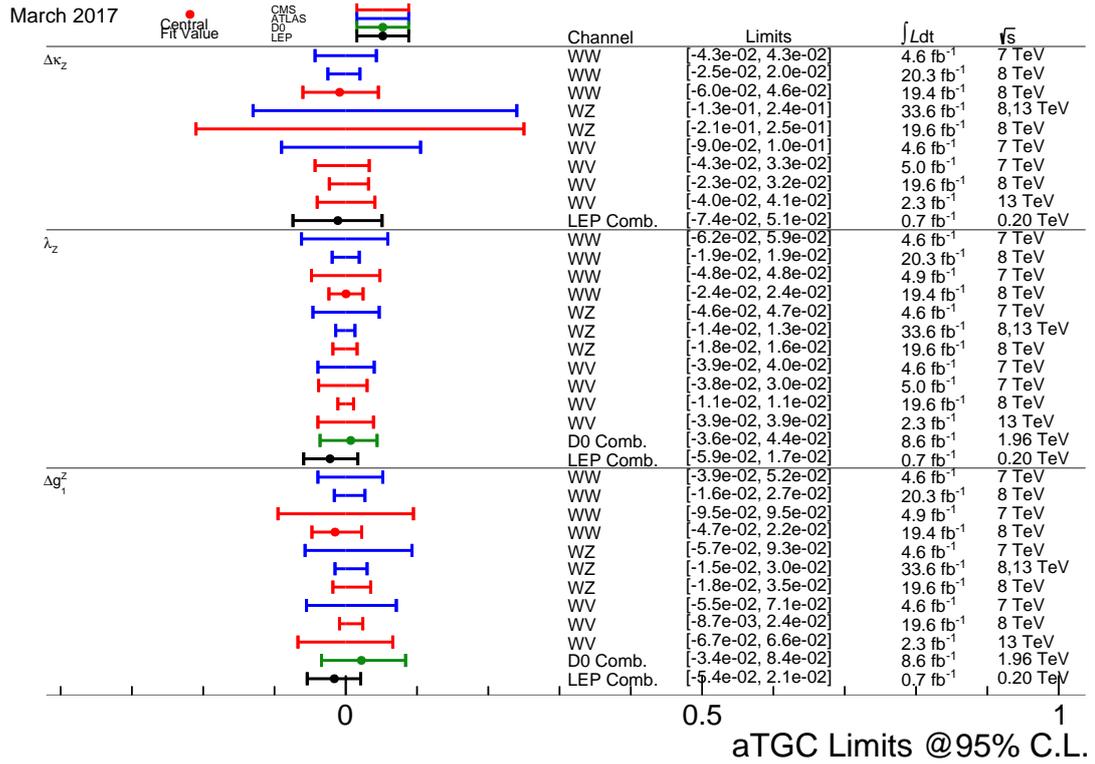
The diboson mass spectrum of signal and background contributions follows an exponentially falling distribution to a good approximation. Accordingly signal and backgrounds are modelled with analytical functions that describe the simulated diboson mass spectrum. The signal function was constructed based on the simulated sample and provides analytical dependence of possible enhancements in the diboson mass spectrum as a function of anomalous triple gauge couplings. The signal function takes into account the interference effects between different anomalous triple gauge couplings as well the interference between the Standard Model contribution and a contribution from a given anomalous triple gauge coupling. It was verified that the signal function describes the simulated distribution in the range of sensitivity of the analysis.

The signal region was split into two categories which can provide additional discrimination for one of the couplings.

Limits at 95% confidence level on anomalous triple gauge couplings were extracted from the diboson mass distribution by constructing an unbinned likelihood function. One dimensional limits are shown in Table 8.3. The limits are comparable with most of the existing ones as illustrated in Figure 8.4.

**Table 8.3:** Expected and observed limits at 95% C.L. on single anomalous couplings (other couplings set to zero).

	aTGC	expected limit	observed limit
EFT param.	$\frac{c_{WW}W}{\Lambda^2}$ ( $\text{TeV}^{-2}$ )	[-8.73, 8.70]	[-9.46, 9.42]
	$\frac{\hat{c}_W}{\Lambda^2}$ ( $\text{TeV}^{-2}$ )	[-11.7, 11.1]	[-12.6, 12.0]
	$\frac{c_B}{\Lambda^2}$ ( $\text{TeV}^{-2}$ )	[-54.9, 53.3]	[-56.1, 55.4]
	$\frac{c_{\tilde{B}}}{\Lambda^2}$ ( $\text{TeV}^{-2}$ )		
Vertex param.	$\lambda$	[-0.036, 0.036]	[-0.039, 0.039]
	$\Delta g_1^Z$	[-0.066, 0.064]	[-0.067, 0.066]
	$\Delta \kappa_Z$	[-0.038, 0.040]	[-0.040, 0.041]



**Figure 8.4:** Limits on WWZ anomalous triple gauge couplings [54].

One of the improvements of the analysis could be the usage of softdrop jet mass as it is preferred theoretically [159] instead of pruned jet mass. Furthermore, other jet substructure

---

techniques can be used to further reduce the contribution from backgrounds. For example, to suppress  $t\bar{t}$  background top-tagging can be used [192]. The  $W$ +jets contribution can be reduced by using other jet-substructure variables (quark/gluon likelihood, jet pull angle, jet charge and others [195]). Additionally grooming algorithms were studied at high pileup and it was found that grooming together with charged hadron subtraction or PUPPI significantly reduces pileup dependence of jet mass. Another possible improvement of the analysis could be the usage of a two-dimensional fit in the space of pruned jet mass and diboson mass instead of a two step procedure when  $W$ +jets and  $t\bar{t}$  backgrounds are estimated from the fit to data in pruned jet mass first and then the limits are extracted from the fit to the the diboson mass spectrum.

The analysis would also certainly benefit from more events in data. Therefore it will be important to repeat the analysis with the large dataset recorded in 2016 and the dataset that will be recorded in 2017 (which is expected to be roughly of the same size as the one from 2016). One can assume that the limit power of the analysis is determined by the significance of the signal which can be estimated as  $S/\sqrt{S+B}$  where  $S$  is the number of signal events and  $B$  is the number of background events [127]. Considering that  $B \gg S$  it can be concluded that to get twice higher significance, four times bigger integrated luminosity is needed. Since the cross-section is proportional to the square of anomalous triple gauge coupling it can roughly be estimated that 16 times bigger integrated luminosity is needed to get twice better limits on the coupling parameters. Thus in the naïve approximation one can expect twice better limits with 2016 data (about  $38 \text{ fb}^{-1}$  of integrated luminosity recorded by CMS [94]).

The simulation of signal was done at leading order, however calculations of QCD [196] and electroweak corrections [197, 198] are available and might have an influence on the results. The inclusion of such refined calculations could improve the next version of the analysis.

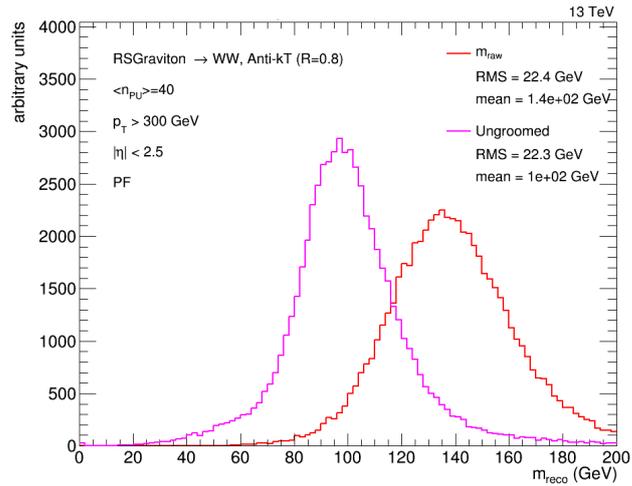
Without exaggeration, the LHC experiments are in their crucial years at the moment: before the Long Shutdown 2 (LS2) in 2019-2020 around  $150 \text{ fb}^{-1}$  of integrated luminosity should be delivered. Analysing this data could finally shed light on new physics. While the observation of a new resonance by the LHC experiments would indicate that there is new physics at the TeV scale, the observation of non-zero anomalous triple gauge couplings would mean that there is new physics but with a scale much larger than the TeV scale.



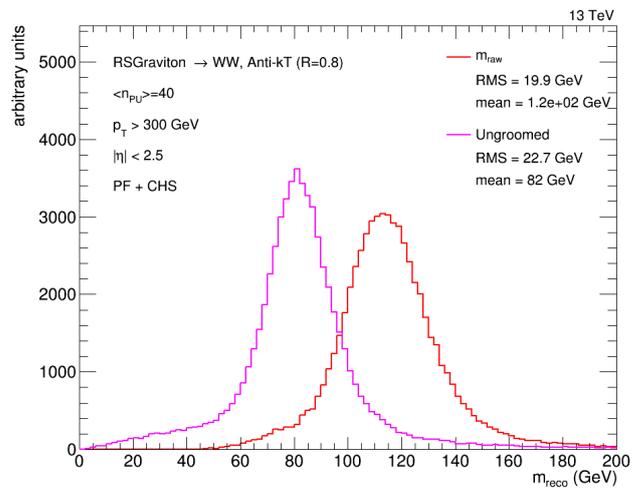
# **Appendices**



# A Mass distributions for grooming algorithms



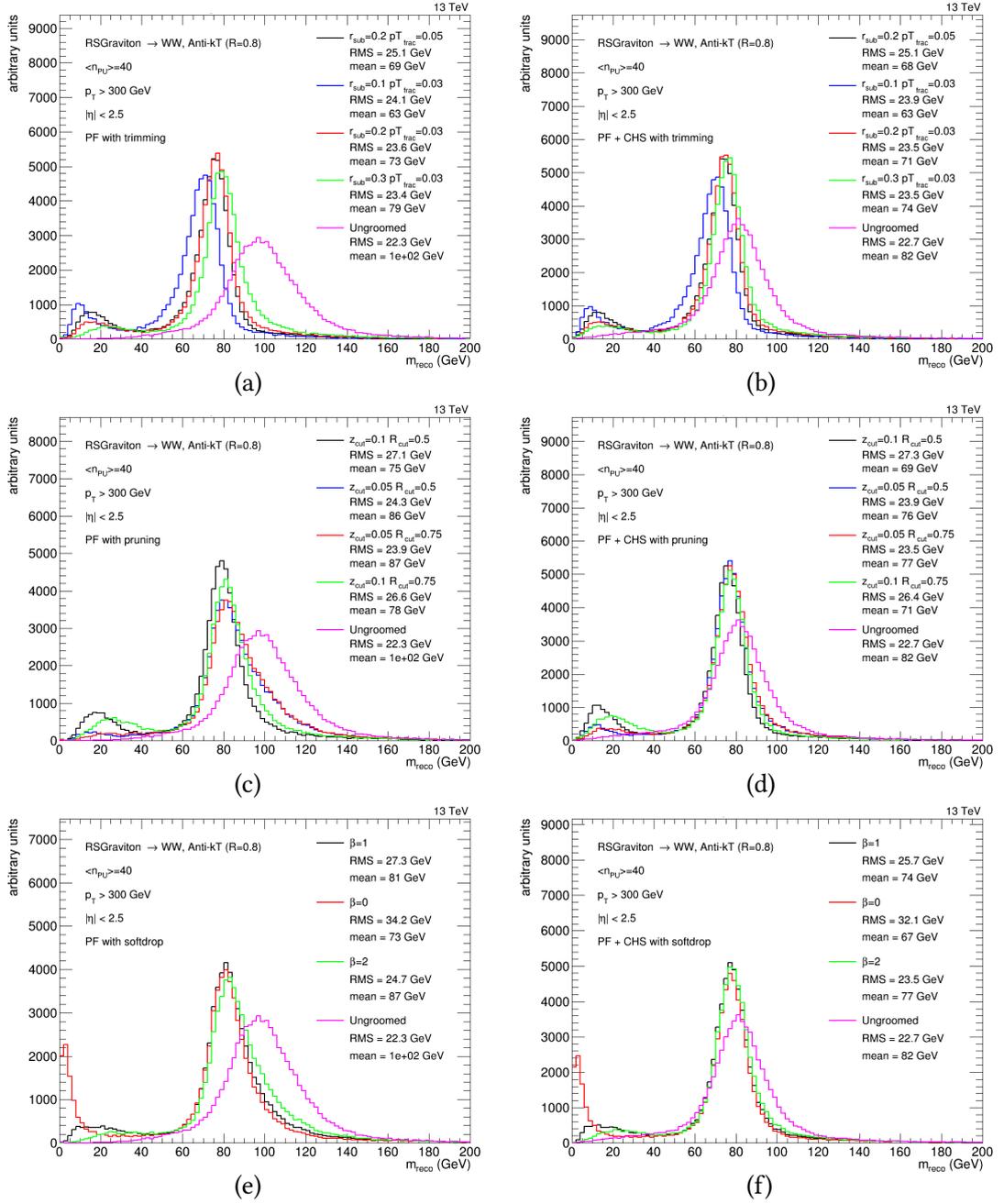
(a)



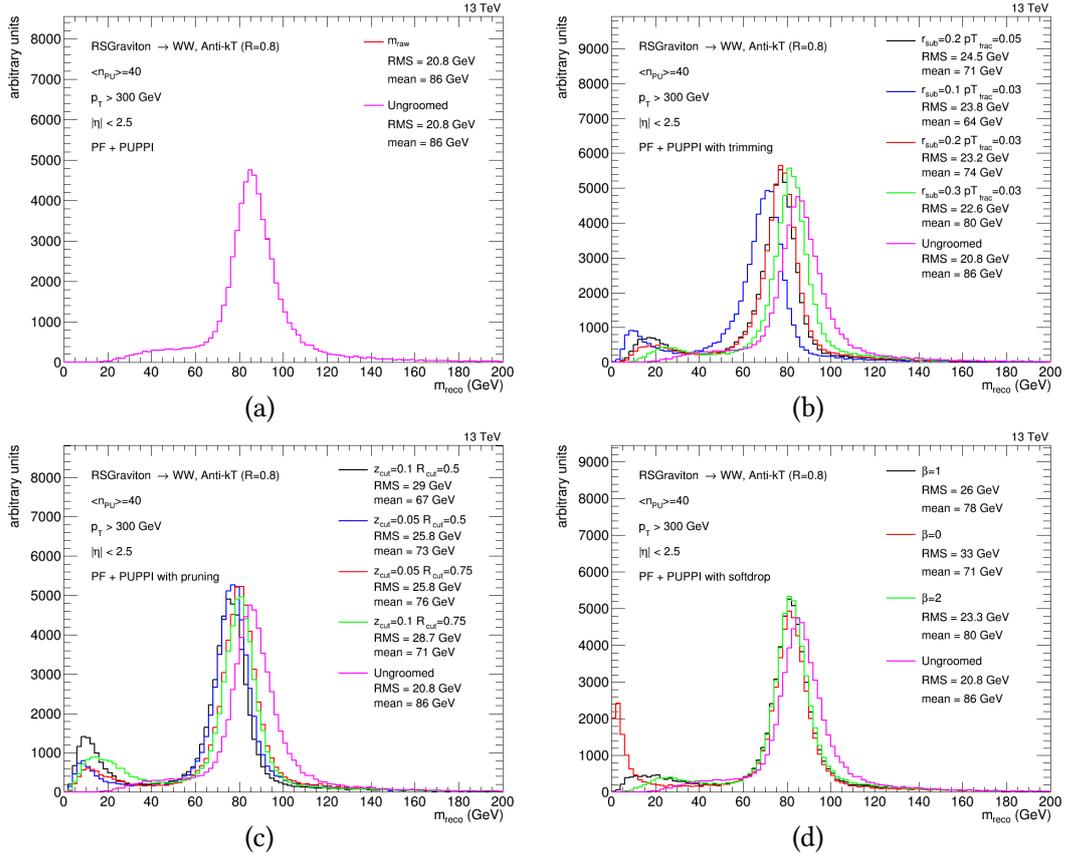
(b)

**Figure A.1:** Raw and ungroomed jet mass are shown on with PF (a) and PF+CHS (b) used as inputs for the jet clustering.

## A Mass distributions for grooming algorithms



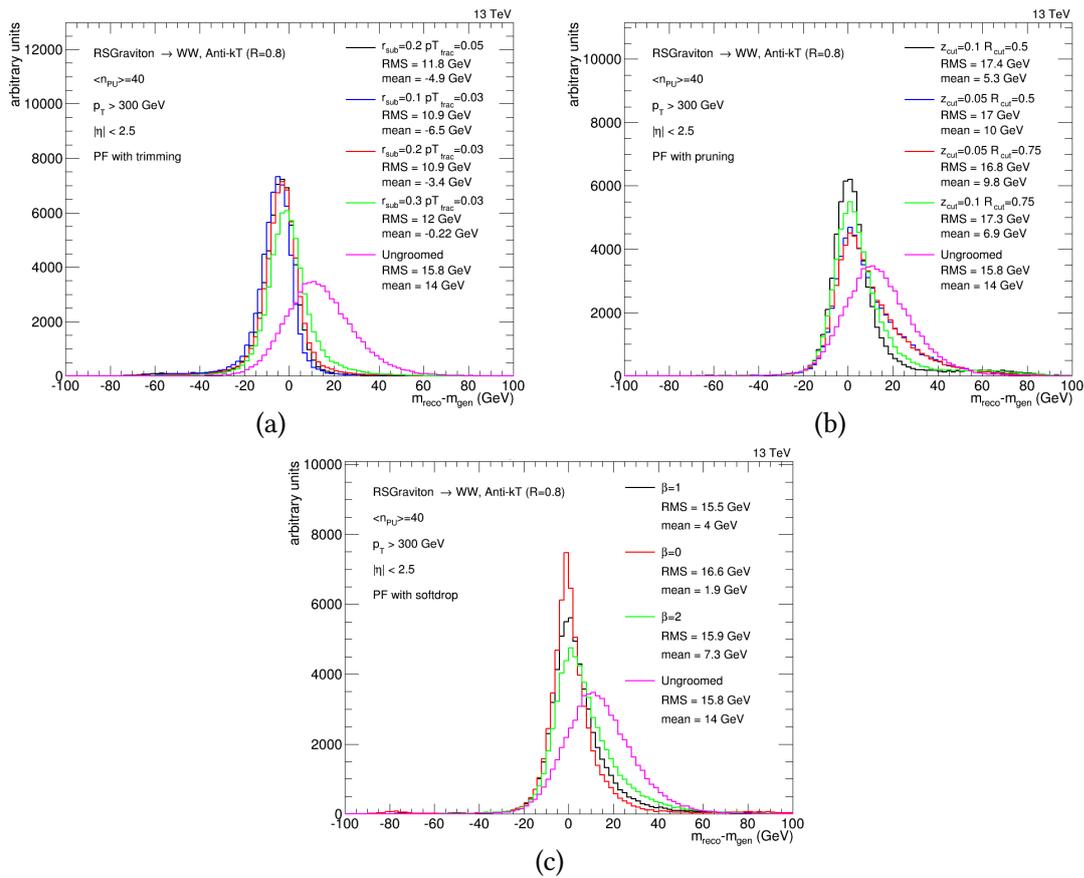
**Figure A.2:** Jet mass distributions for various grooming algorithms for W-jets: trimming (a,b), pruning (c,d) and softdrop (e,f). PF (a,c,e) and PF+CHS (b,d,f) are used as inputs for the jet clustering.



**Figure A.3:** Jet mass distributions for various grooming algorithms for W-jets: trimming (b), pruning (c) and softdrop (d). Raw and ungroomed jet mass are shown on (a). PF+PUPPI is used as input for the jet clustering.

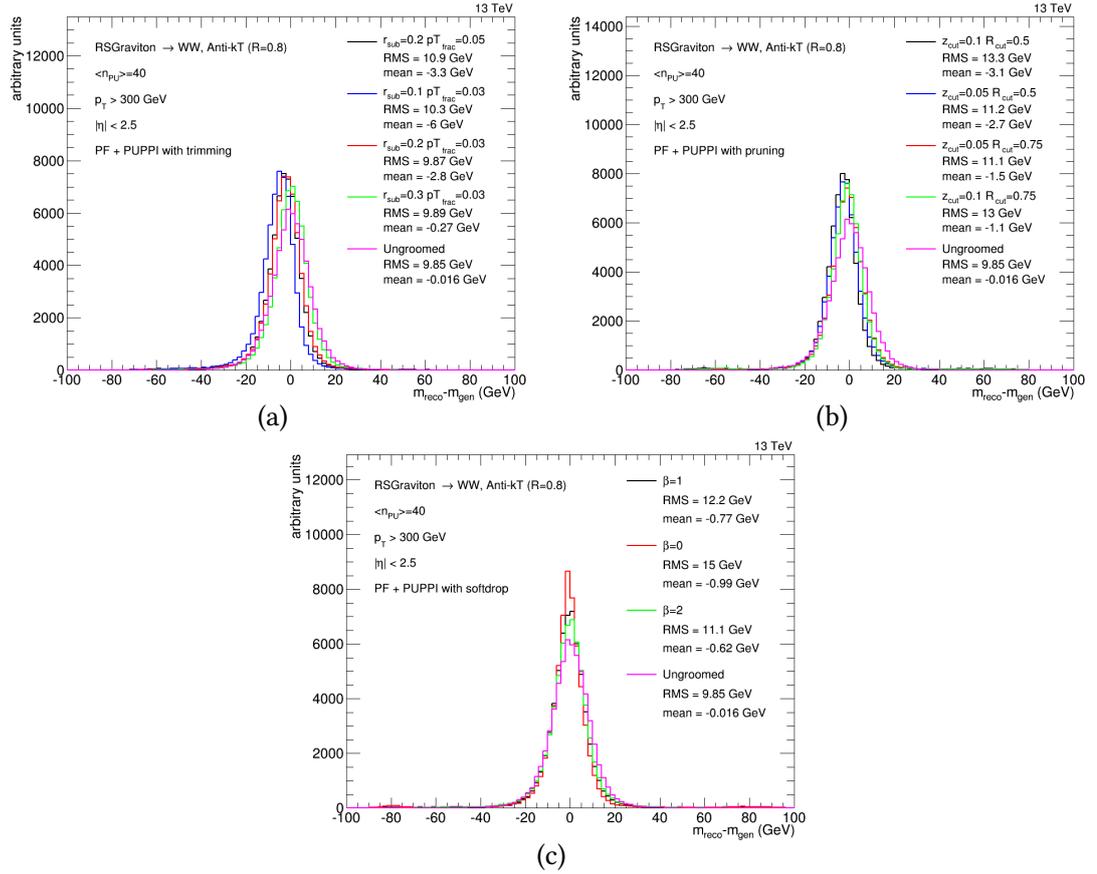


## B Mass response distributions for grooming algorithms

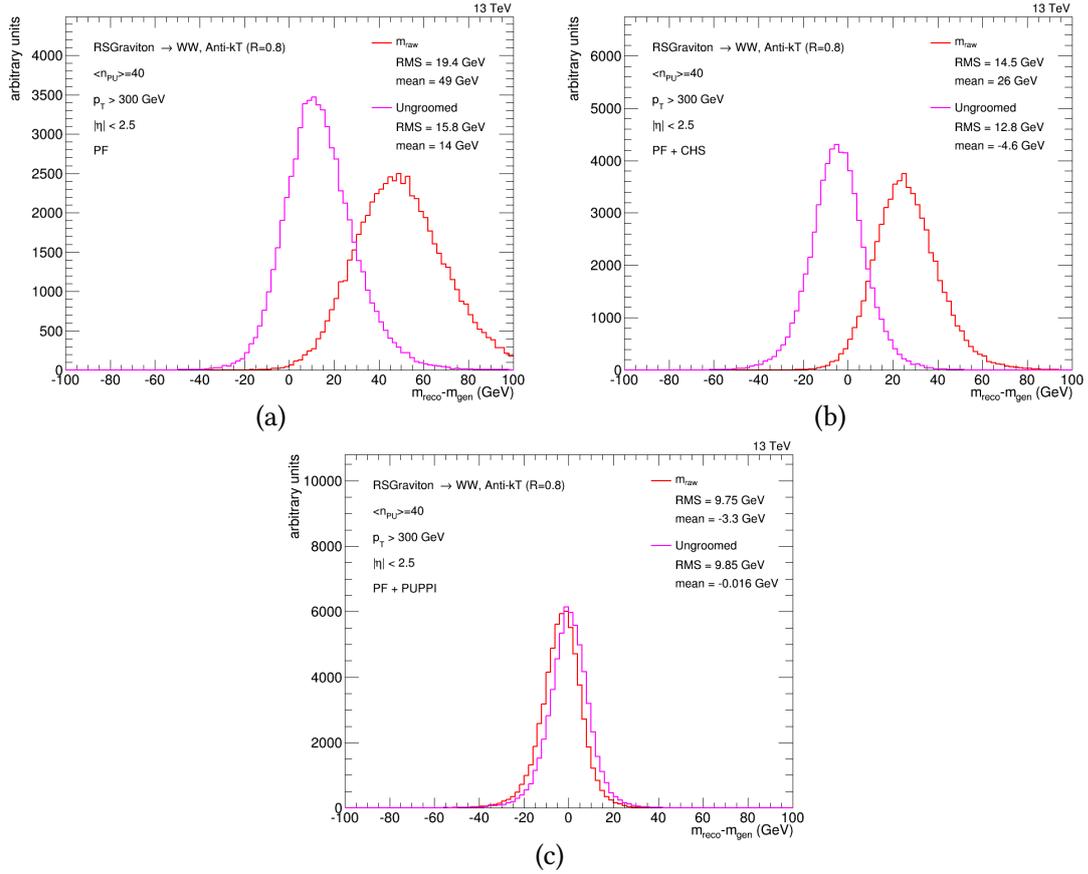


**Figure B.1:** Mass response distributions for W-jets for different grooming algorithms: trimming (a), pruning (b), softdrop(c). PF is used as input for the jet clustering.

## B Mass response distributions for grooming algorithms



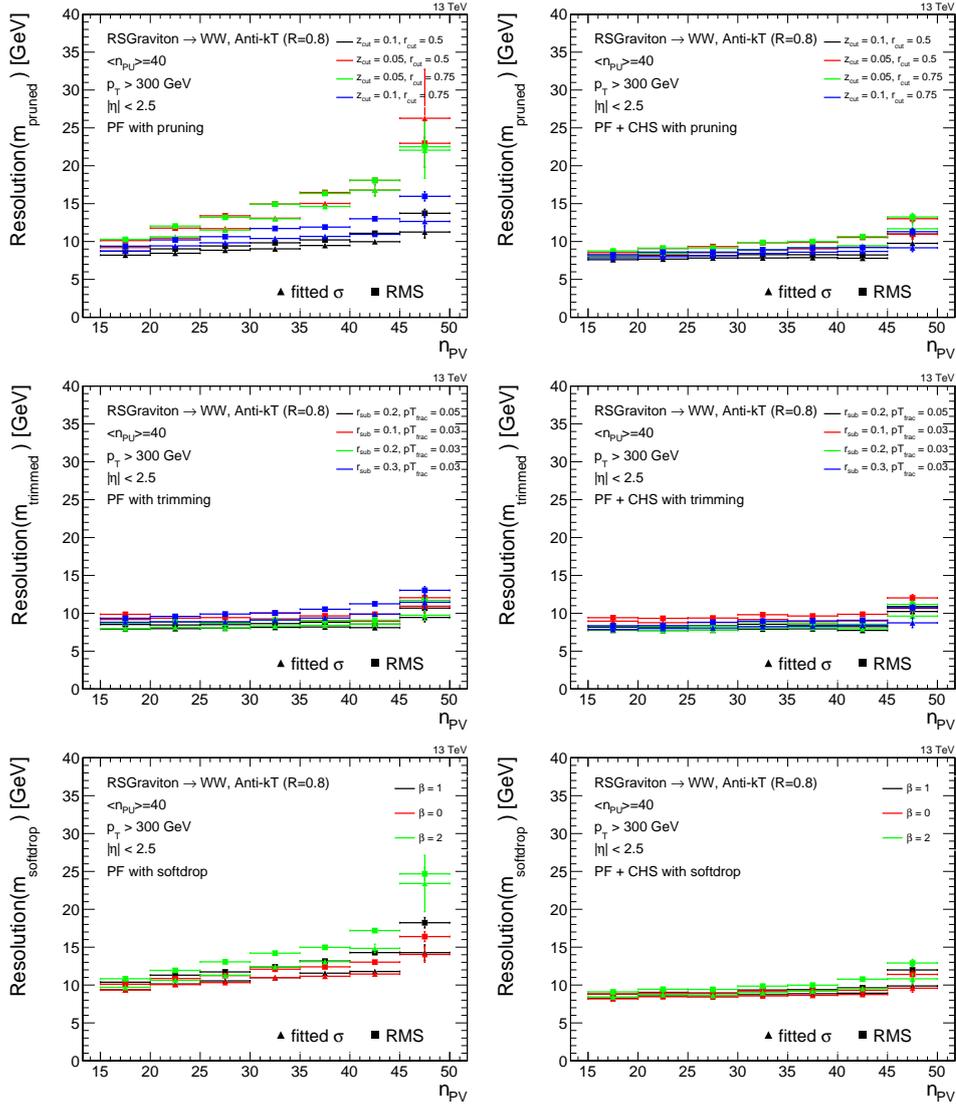
**Figure B.2:** Mass response distributions for W-jets for different grooming algorithms: trimming (a), pruning (b), softdrop(c). PF+PUPPI is used as input for the jet clustering.



**Figure B.3:** Raw and ungroomed jet mass response for W-jets. PF (a), PF+CHS (b) and PF+PUPPI (c) are used as input for the jet clustering.



## C Pileup dependence plots



**Figure C.1:** Jet mass resolution as a function of number of reconstructed vertices for PF (left) and PF+CHS (right).

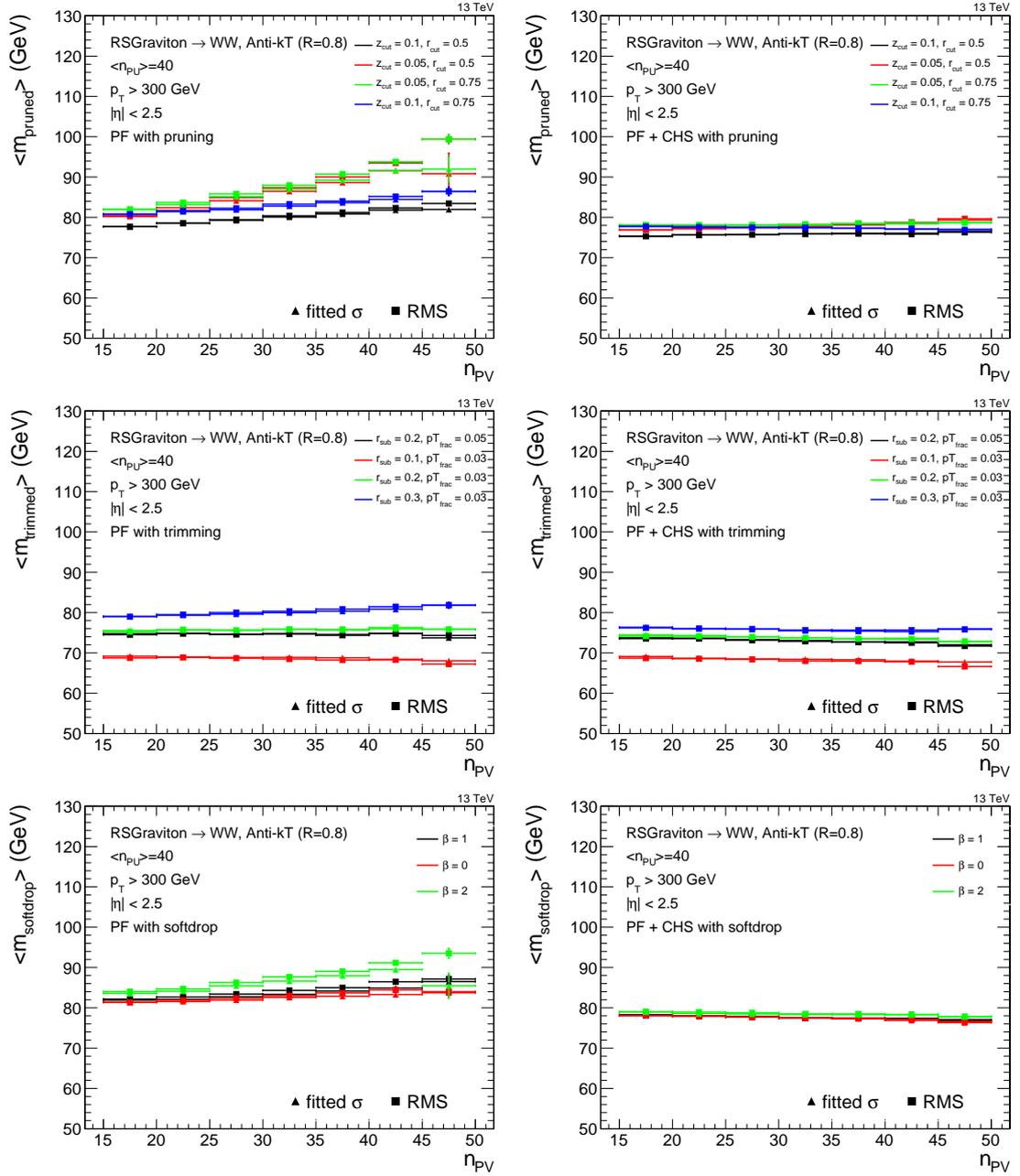


Figure C.2: Mean jet mass as a function of number of reconstructed vertices for PF (left) and PF + CHS (right).

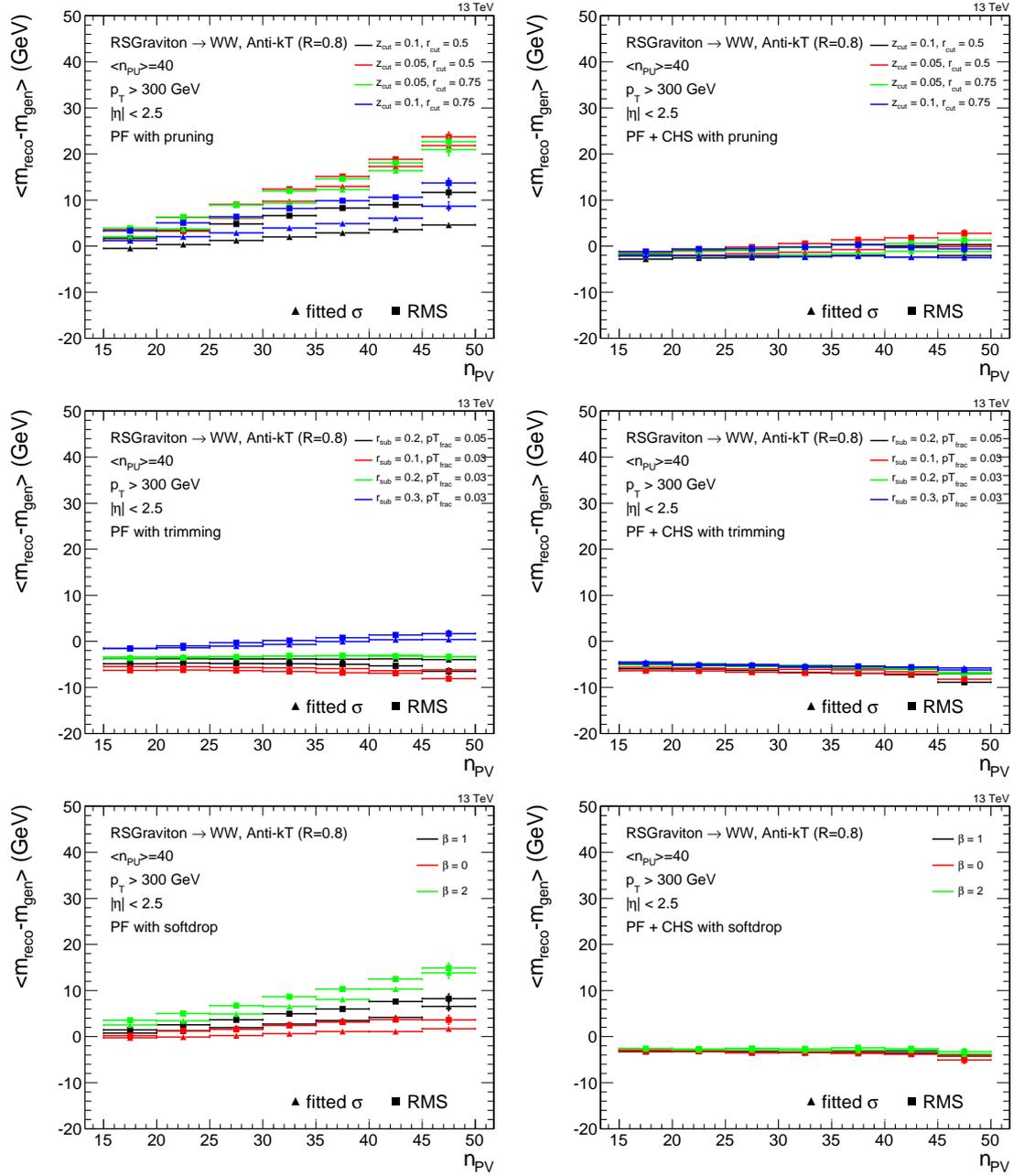


Figure C.3: Mean mass response as a function of number of reconstructed vertices for PF (left) and PF + CHS (right).

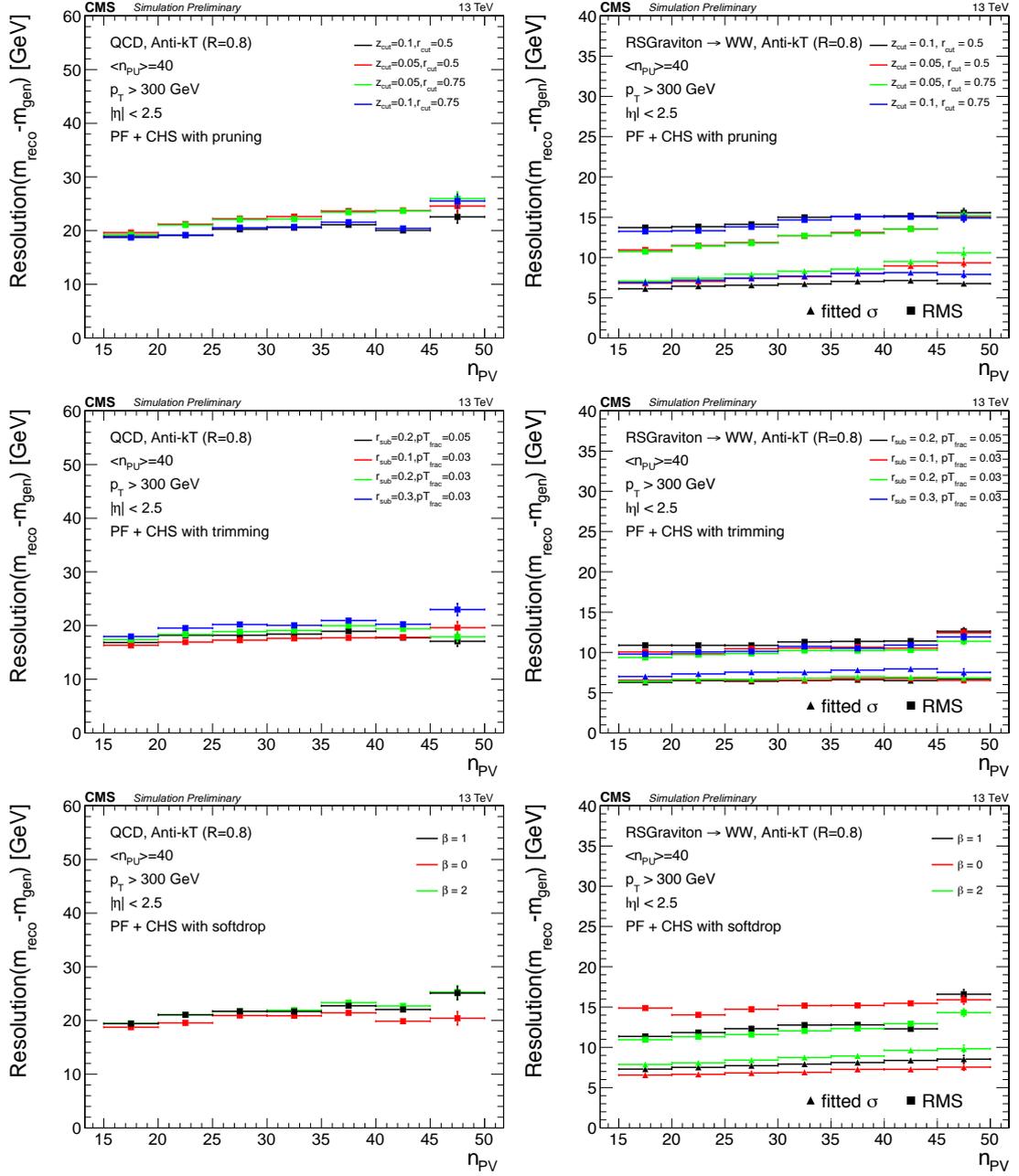


Figure C.4: Jet mass response resolution for QCD-jets (left) and W-jets (right) as a function of number of reconstructed vertices. PF+CHS is used as input for the jet clustering [147].

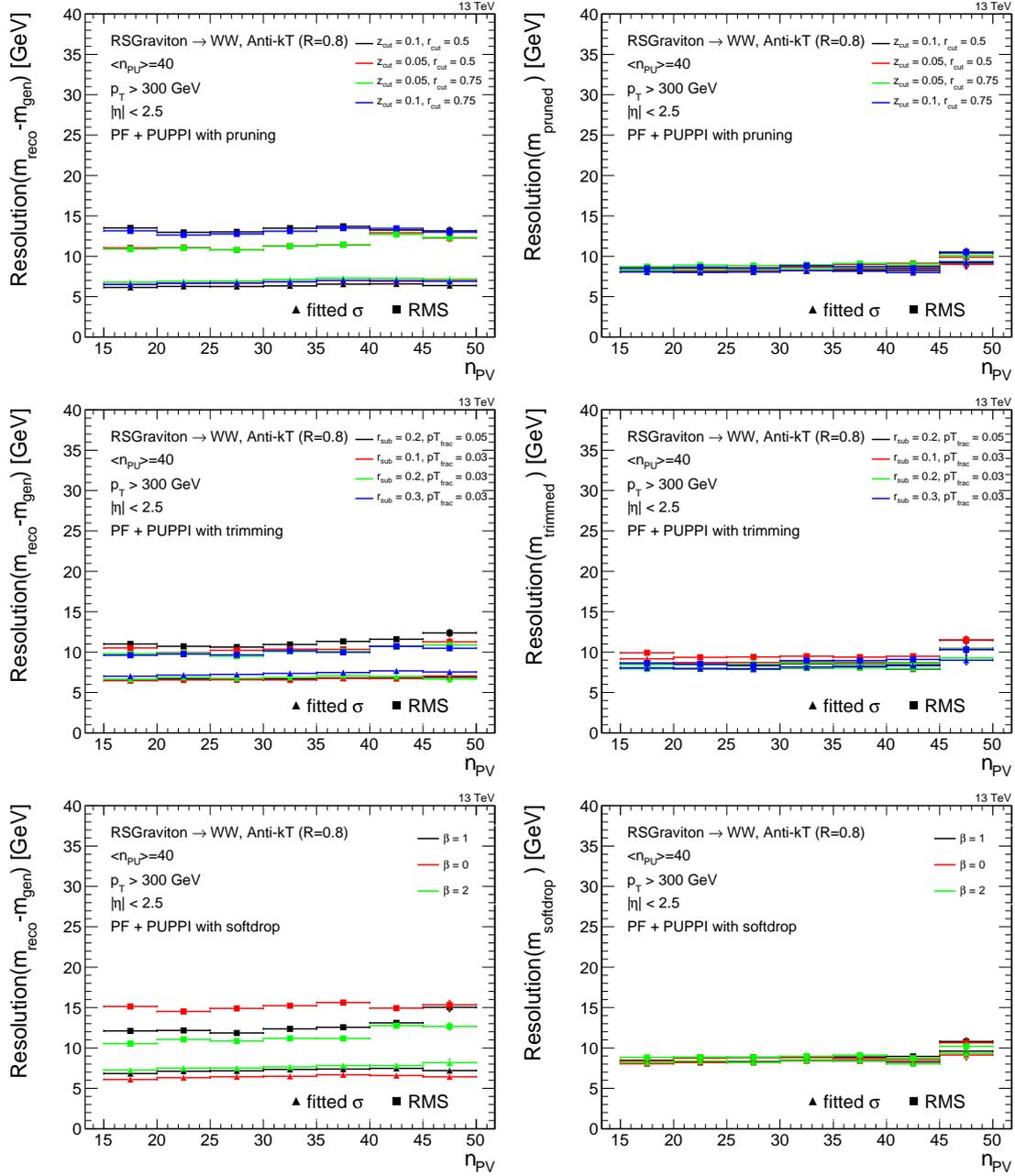


Figure C.5: Response resolution (left) and mass resolution (right) as a function of number of reconstructed vertices for PF+PUPPI.

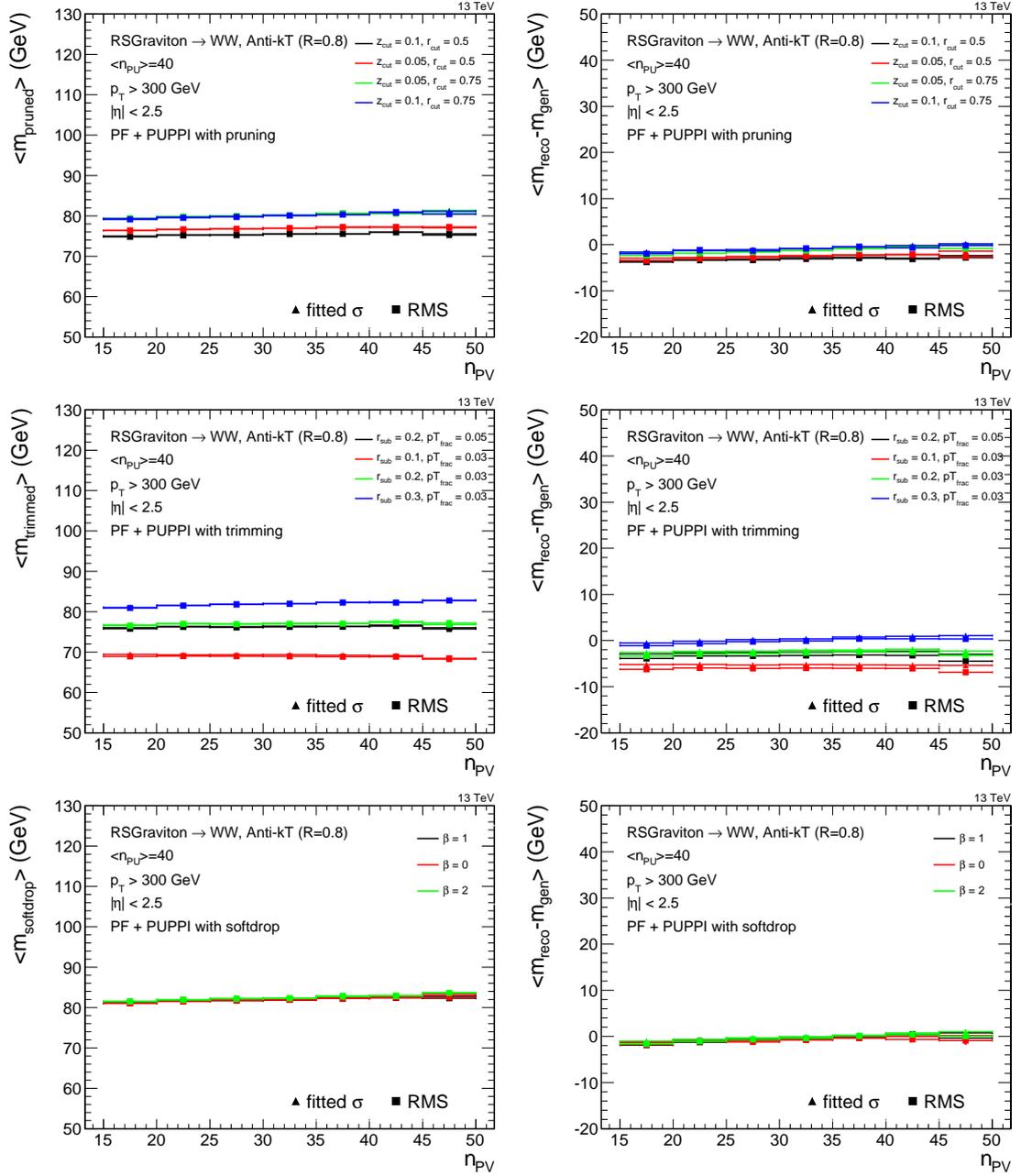
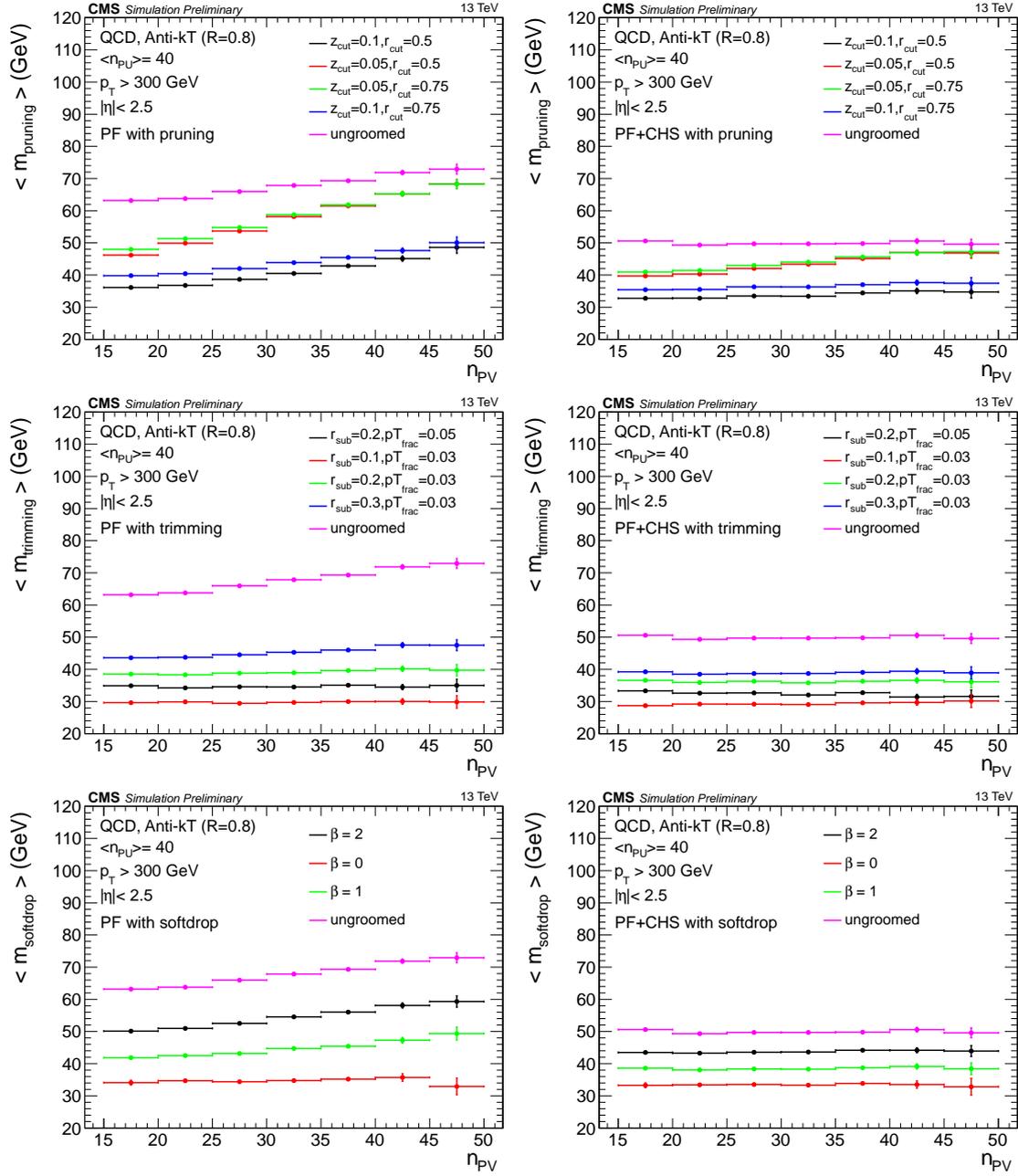


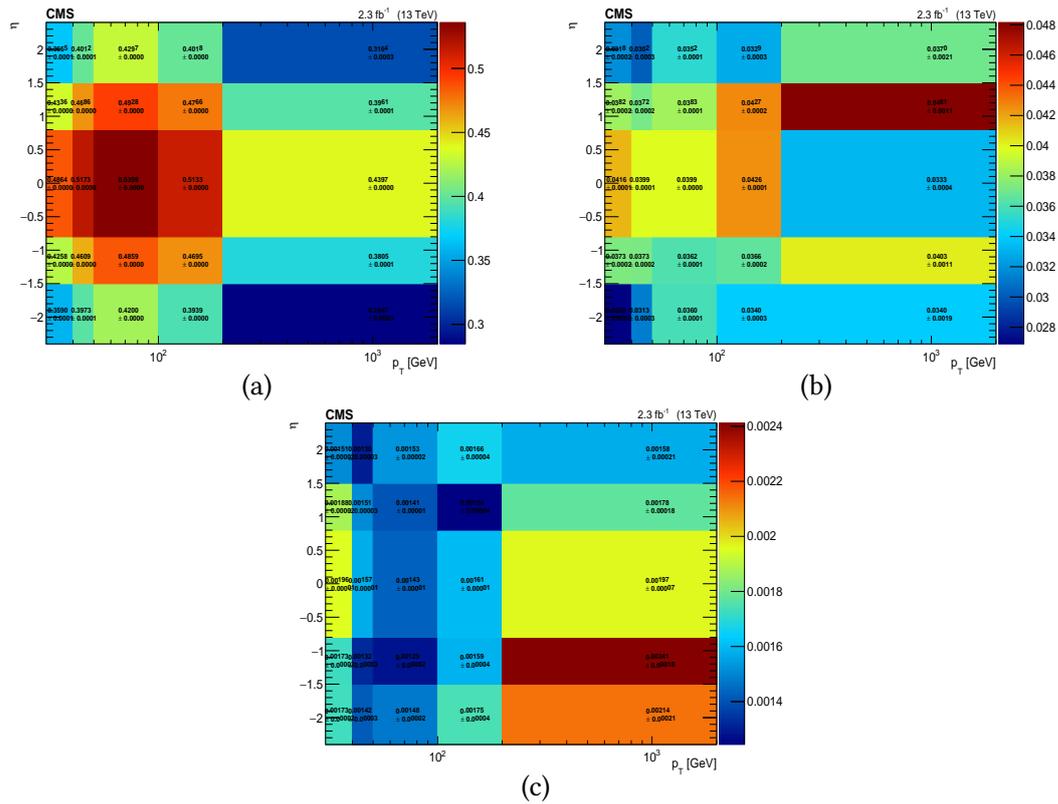
Figure C.6: Mean jet mass (left) and mean response as a function of number of reconstructed vertices for PF+PUPPI.



**Figure C.7:** Mean jet mass distribution for QCD-jets as a function of number of reconstructed vertices. PF (left) and PF+CHS (right) are used as input for the jet clustering [147].



## D Additional plots for b-tagging efficiencies



**Figure D.1:** B-tagging (mis-tagging) efficiencies for  $t\bar{t}$  sample, electron channel for b-jets (a), c-jets (b) and light flavour jets (c).

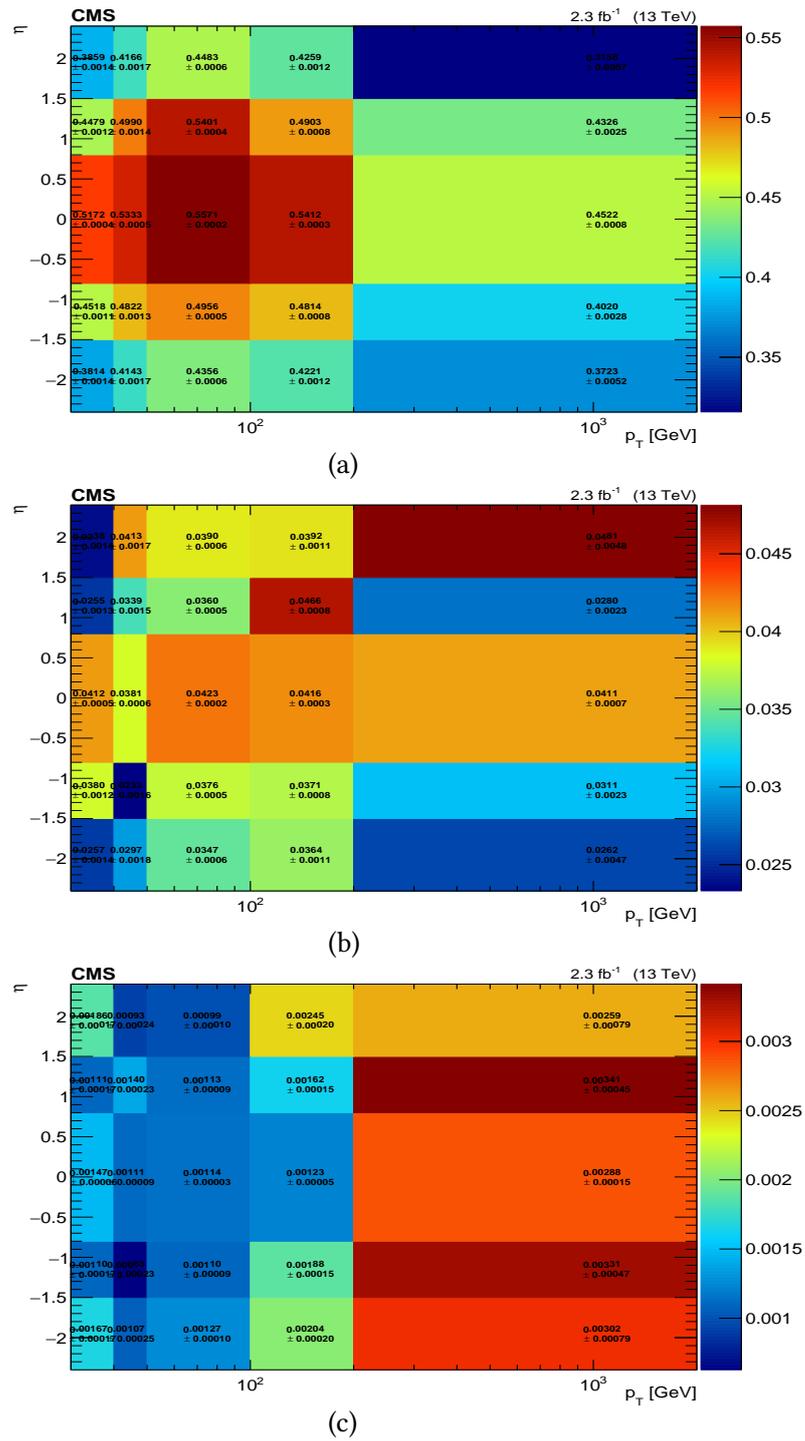


Figure D.2: B-tagging (mis-tagging) efficiencies for WZ sample, muon channel for b-jets (a), c-jets (b) and light flavour jets (c).

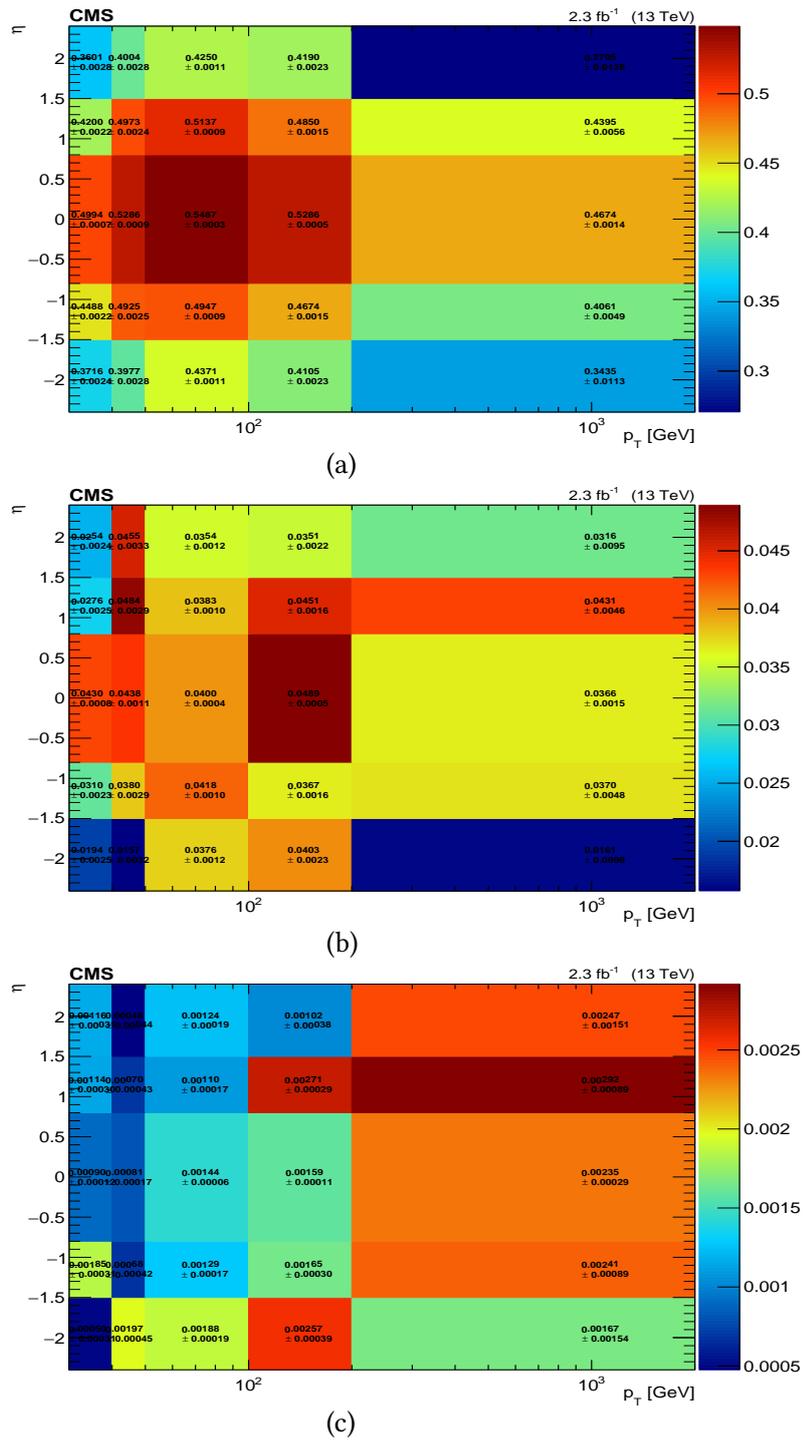
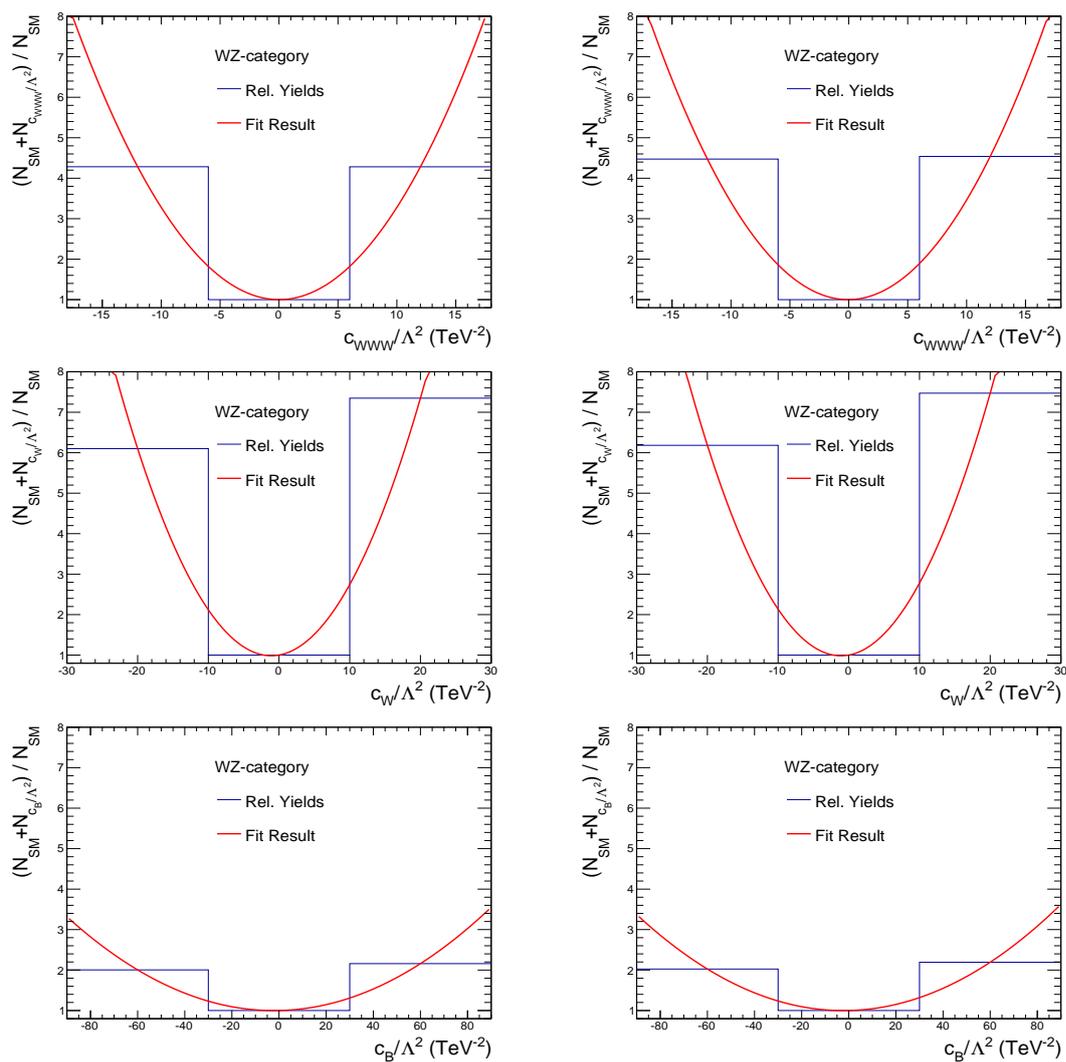


Figure D.3: B-tagging (mis-tagging) efficiencies for WZ sample, muon channel for b-jets (a), c-jets (b) and light flavour jets (c).



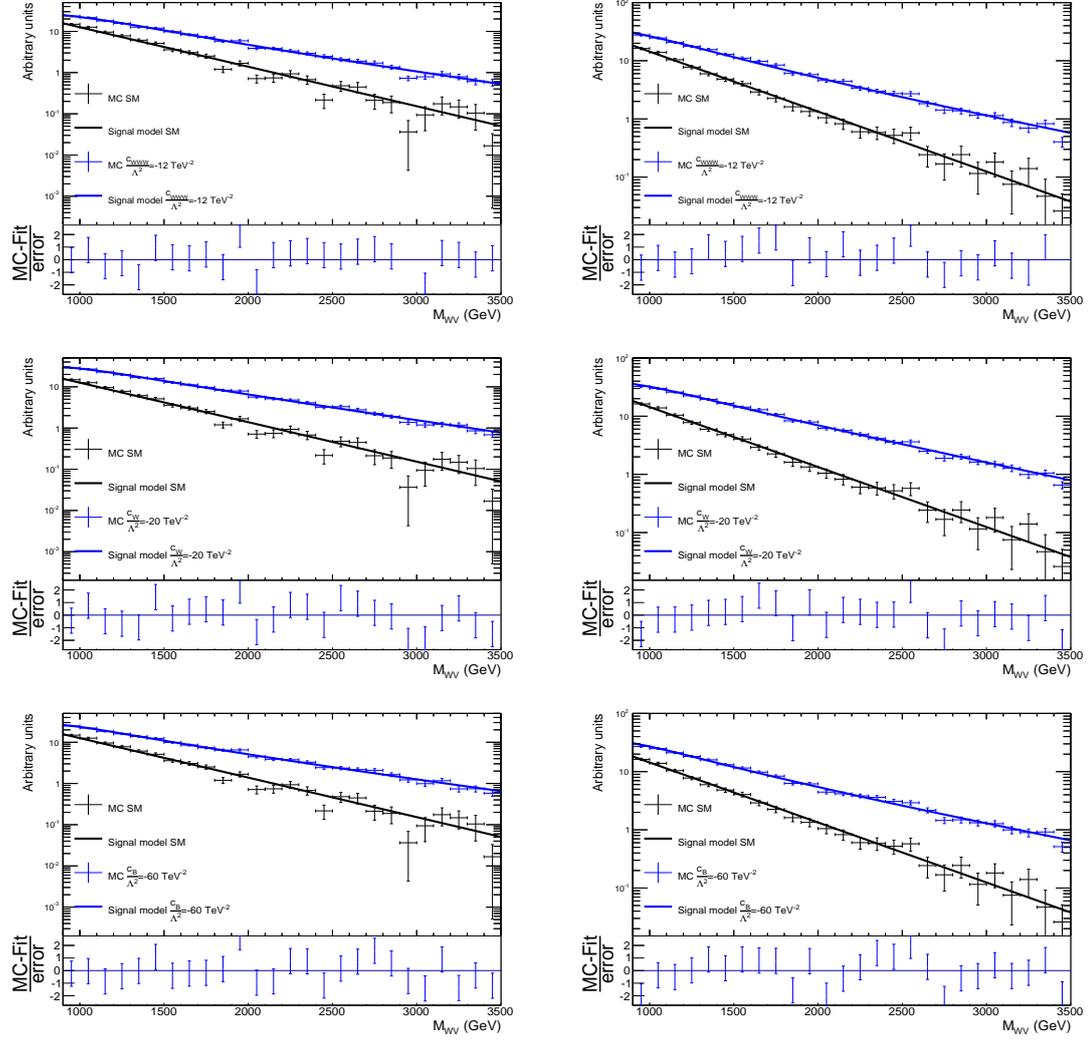
# E Additional plots for the signal model

## E.1 Normalization of the signal



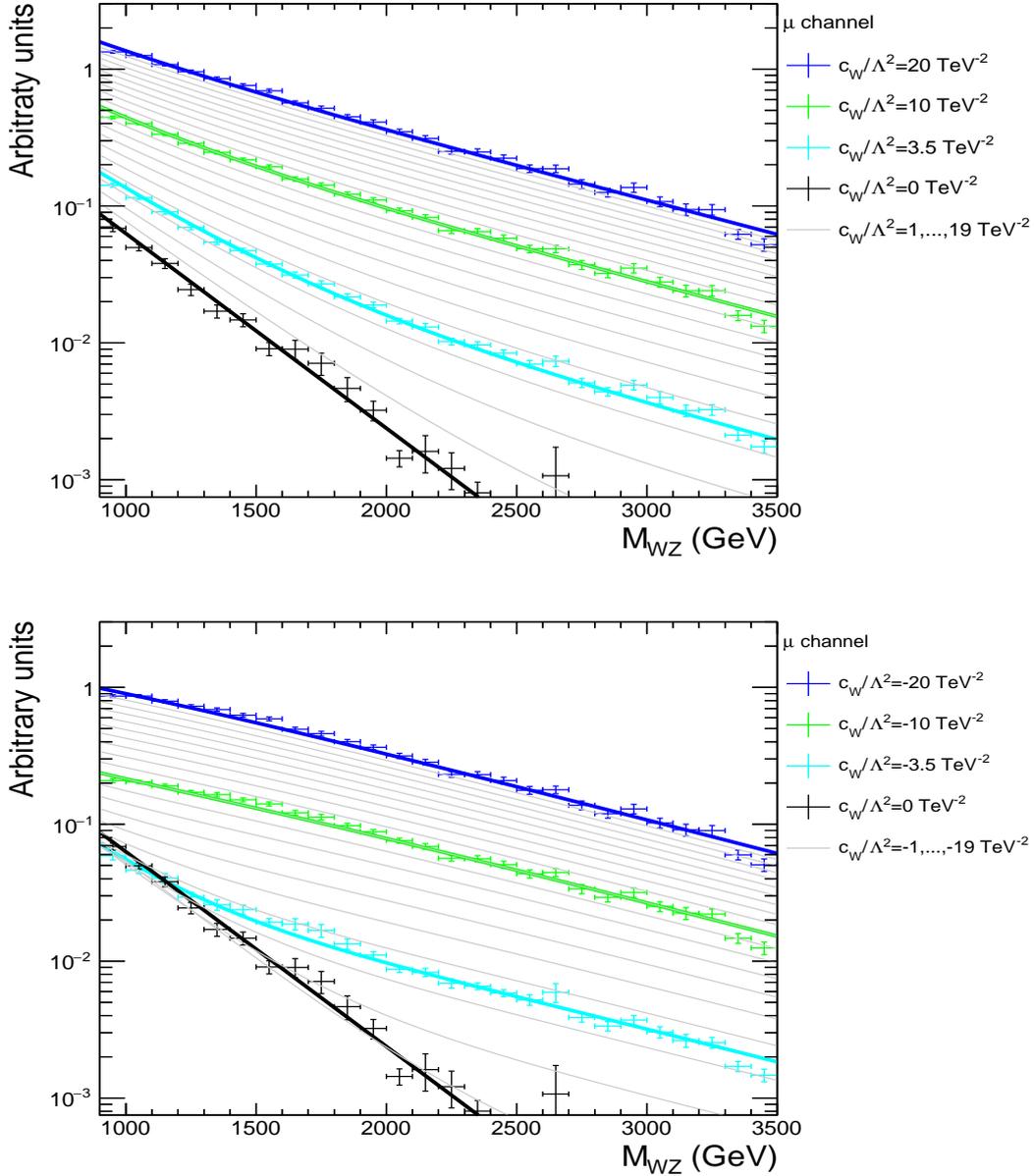
**Figure E.1:** Normalized yields for different values of the aTGC-parameters in the WZ-category for electron- (left) and muon-channel (right) with fitted quadratic function. From top to bottom:  $\frac{c_{WWW}}{\Lambda^2}$ ,  $\frac{c_W}{\Lambda^2}$ ,  $\frac{c_B}{\Lambda^2}$ .

## E.2 Signal function



**Figure E.2:** MC data and fit of signal model for all parameters set to zero (black) and one parameter set to a non-zero value (blue) for electron- (left) and muon-channel (right), WW-category. From top to bottom:  $\frac{c_{WW}}{\Lambda^2} = -12 \text{ TeV}^{-2}$ ,  $\frac{c_W}{\Lambda^2} = -20 \text{ TeV}^{-2}$ ,  $\frac{c_B}{\Lambda^2} = -60 \text{ TeV}^{-2}$ . The ratio plot corresponds to the simulation minus the fit-function divided by the bin-error for the signal function with non-zero aTGC-parameter.

## E.2.1 Additional plots for signal modelling



**Figure E.3:** Signal model and MC on generator level (WZ only) for different values of  $\frac{c_W}{\Lambda^2}$  in the muon channel. The grey lines shows the signal model for increasing  $\frac{c_W}{\Lambda^2}$  in steps of  $1 \text{ TeV}^{-2}$ .

### E.3 Shape uncertainties of the signal function

**Table E.1:** Estimated uncertainties of signal shape slope parameters in the WW-category, muon channel. Uncertainties are quoted in %.

	PDF	Scale	jet en.	lept. en.	lept. res.	$\cancel{E}_T$	jet res.	mis-tag	b-tag	lept. id
$a_{cwww}$	3.93	4.55	0.39	1.19	<0.05	0.51	<0.05	<0.05	<0.05	2.03
$a_{cw}$	1.75	4.21	0.53	0.73	<0.05	0.17	<0.05	<0.05	<0.05	1.00
$a_{cb}$	3.68	4.72	1.45	1.50	<0.05	0.27	<0.05	<0.05	<0.05	1.80

**Table E.2:** Estimated uncertainties of signal shape slope parameters in the WW-category, electron channel. Uncertainties are quoted in %.

	PDF	Scale	jet en.	lept. en.	lept. res.	$\cancel{E}_T$	jet res.	mis-tag	b-tag	lept. id
$a_{cwww}$	4.40	4.89	2.33	0.18	<0.05	0.13	<0.05	<0.05	<0.05	0.23
$a_{cw}$	2.11	4.46	1.55	0.09	<0.05	<0.05	<0.05	<0.05	<0.05	0.22
$a_{cb}$	4.01	5.31	3.03	0.26	<0.05	0.24	<0.05	<0.05	<0.05	0.27

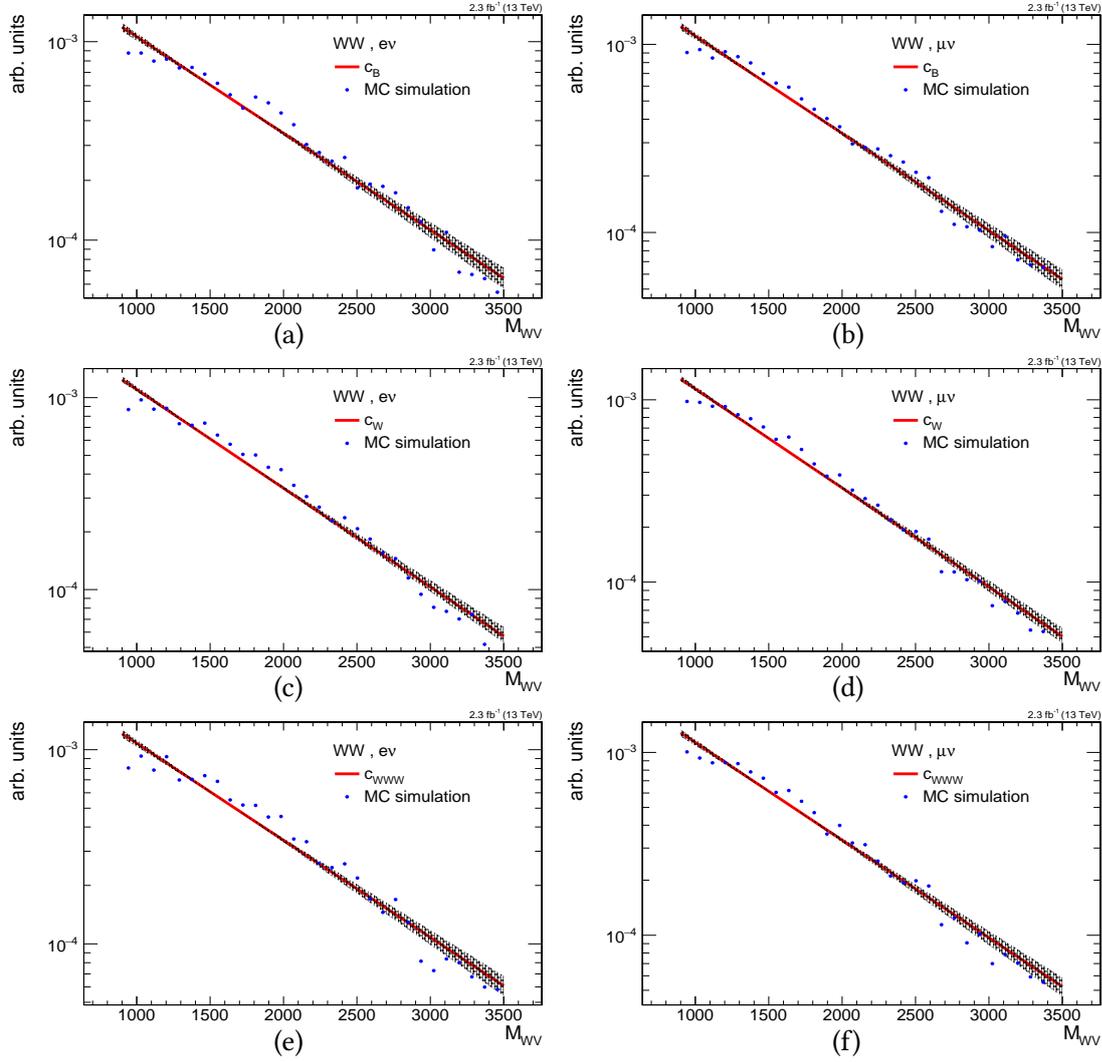
**Table E.3:** Estimated uncertainties of signal shape slope parameters in the WZ-category, muon channel. Uncertainties are quoted in %.

	PDF	Scale	jet en.	lept. en.	lept. res.	$\cancel{E}_T$	jet res.	mis-tag	b-tag	lept. id
$a_{cwww}$	2.14	4.23	1.40	1.05	<0.05	0.37	<0.05	<0.05	<0.05	1.08
$a_{cw}$	1.94	4.07	1.07	0.47	<0.05	0.30	0.07	<0.05	<0.05	0.31
$a_{cb}$	8.40	6.07	12.24	1.71	<0.05	1.30	0.12	<0.05	<0.05	3.94

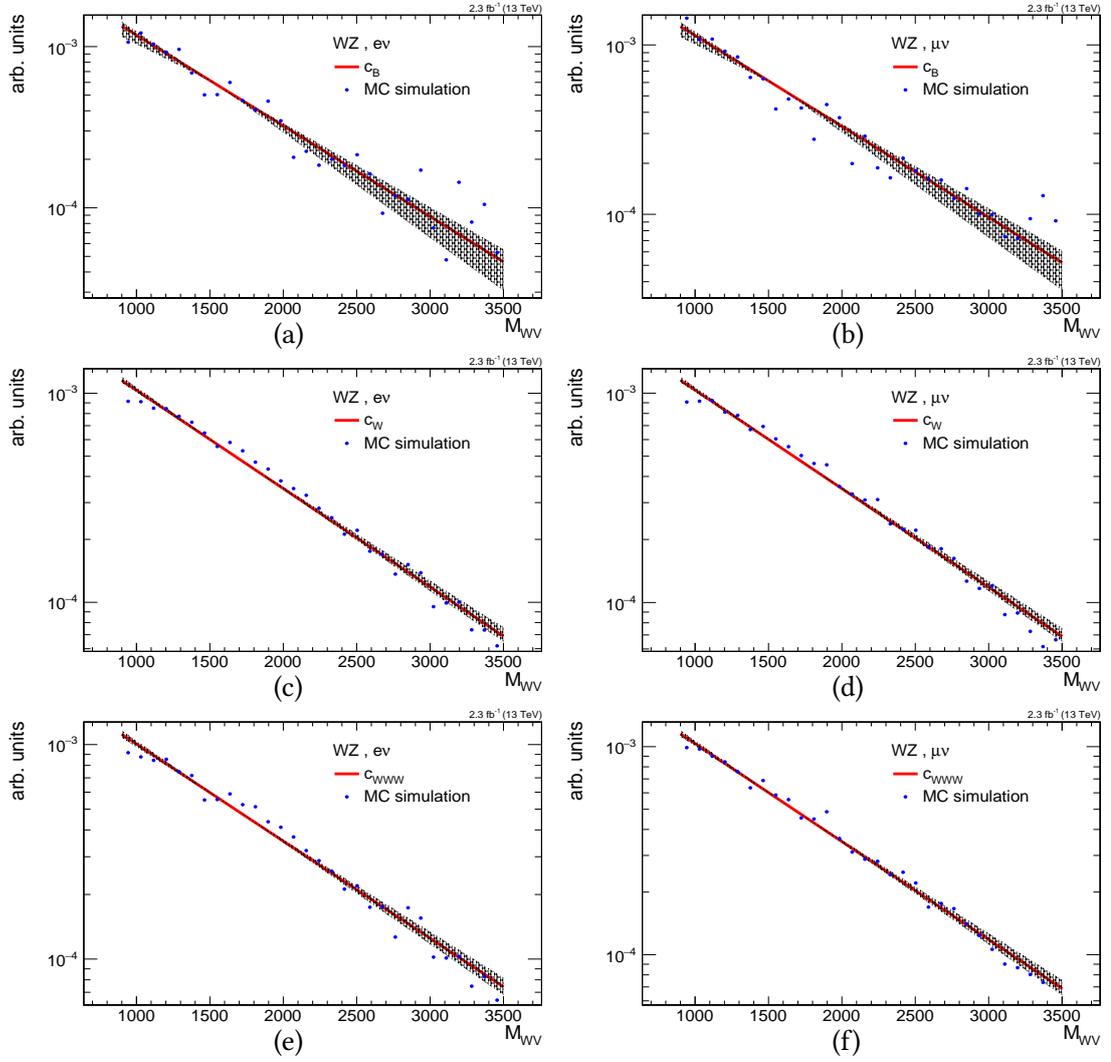
**Table E.4:** Estimated uncertainties of signal shape slope parameters in the WZ-category, electron channel. Uncertainties are quoted in %.

	PDF	Scale	jet en.	lept. en.	lept. res.	$\cancel{E}_T$	jet res.	mis-tag	b-tag	lept. id
$a_{cwww}$	2.64	4.41	3.42	0.72	<0.05	0.14	<0.05	<0.05	<0.05	0.27
$a_{cw}$	2.37	4.15	1.05	0.31	<0.05	0.09	<0.05	<0.05	<0.05	0.29
$a_{cb}$	6.63	5.69	14.61	1.25	<0.05	0.12	0.06	<0.05	<0.05	0.27

E.3.1 Effect of systematic uncertainties on the signal function

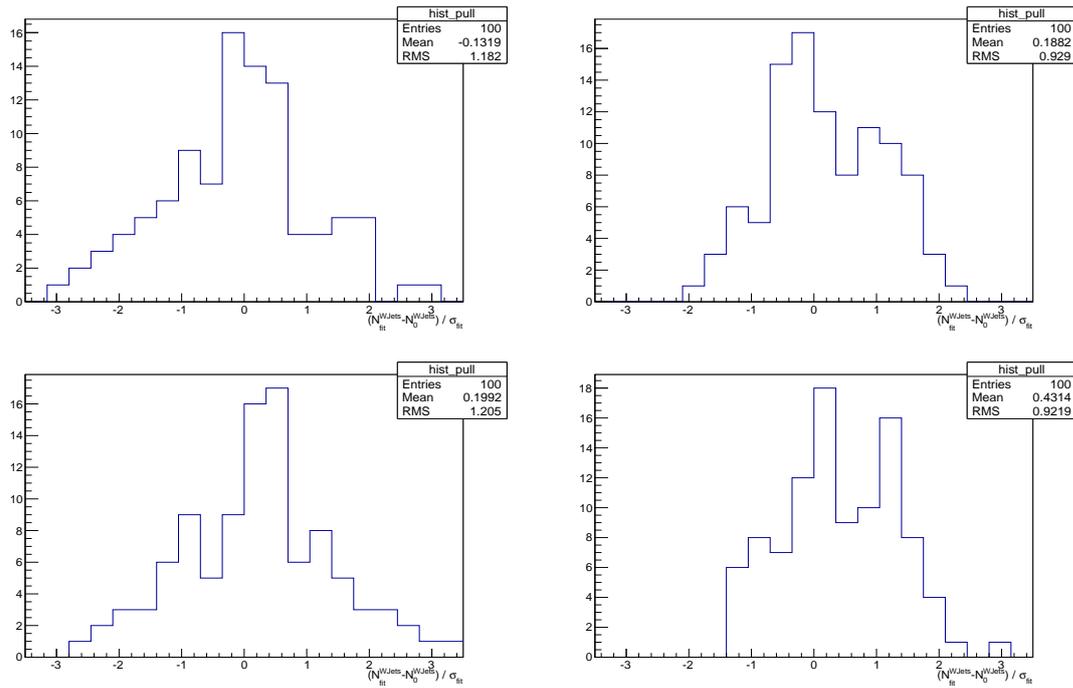


**Figure E.4:** The effect of shape uncertainty for  $a_{cb}$  (a,b),  $a_{c_w}$  (c,d),  $a_{c_{www}}$  (e,f), WW category, muon channel on left and electron channel on the right. Uncertainty is shown with shaded area.



**Figure E.5:** The effect of shape uncertainty for  $a_{cb}$  (a,b),  $a_{cW}$  (c,d),  $a_{cWWW}$  (e,f), WZ category, muon channel on left and electron channel on the right. Uncertainty is shown with shaded area.

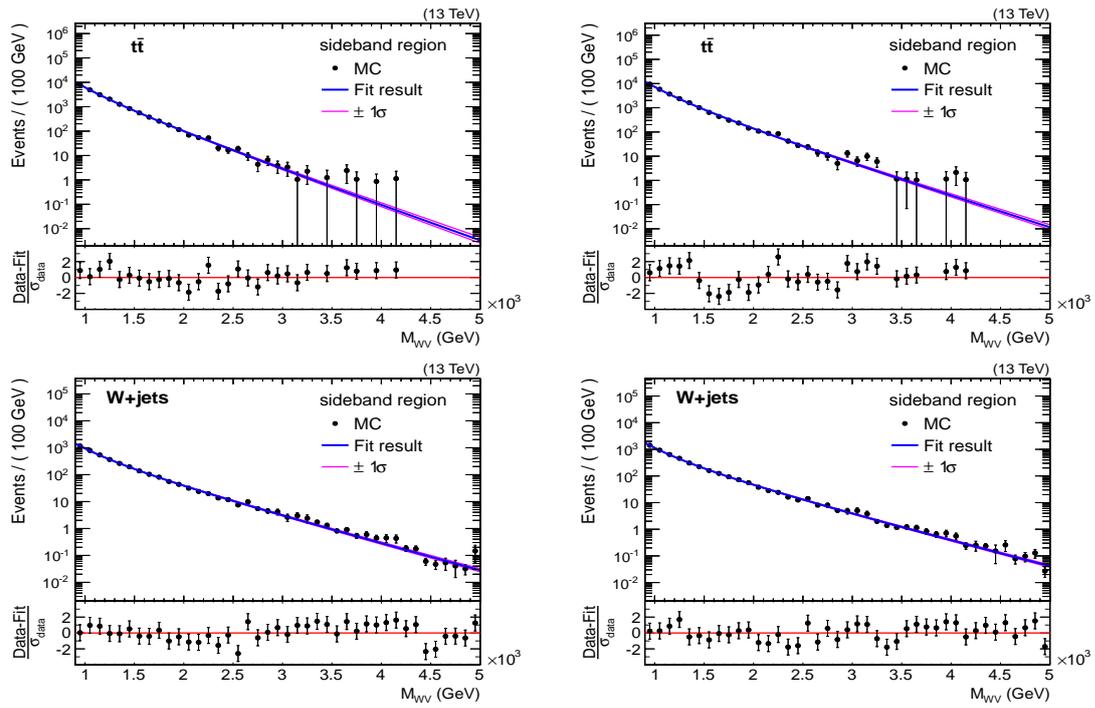
## F Bias test for $M_{\text{pruned}}$ fit



**Figure F.1:** Pull plots of the W+jets normalization extracted with toy MCs randomizing the shape parameters as data input for the electron- (left) and muon-channel (right) in the WW- (top) and WZ-category (bottom).



## G Shape of backgrounds in the sideband



**Figure G.1:** MC-data and shapes of the  $M_{WV}$  distributions in the sideband for electron- (left) and muon-channel (right), for  $t\bar{t}$  (top) and  $W$ +jets (bottom).

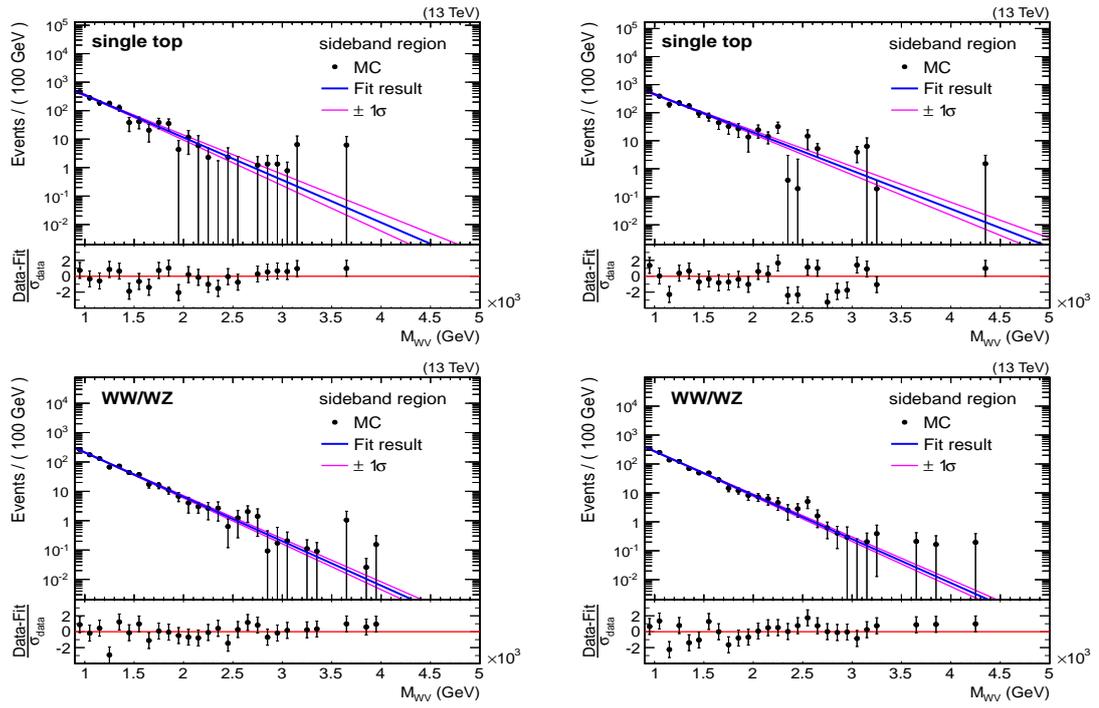
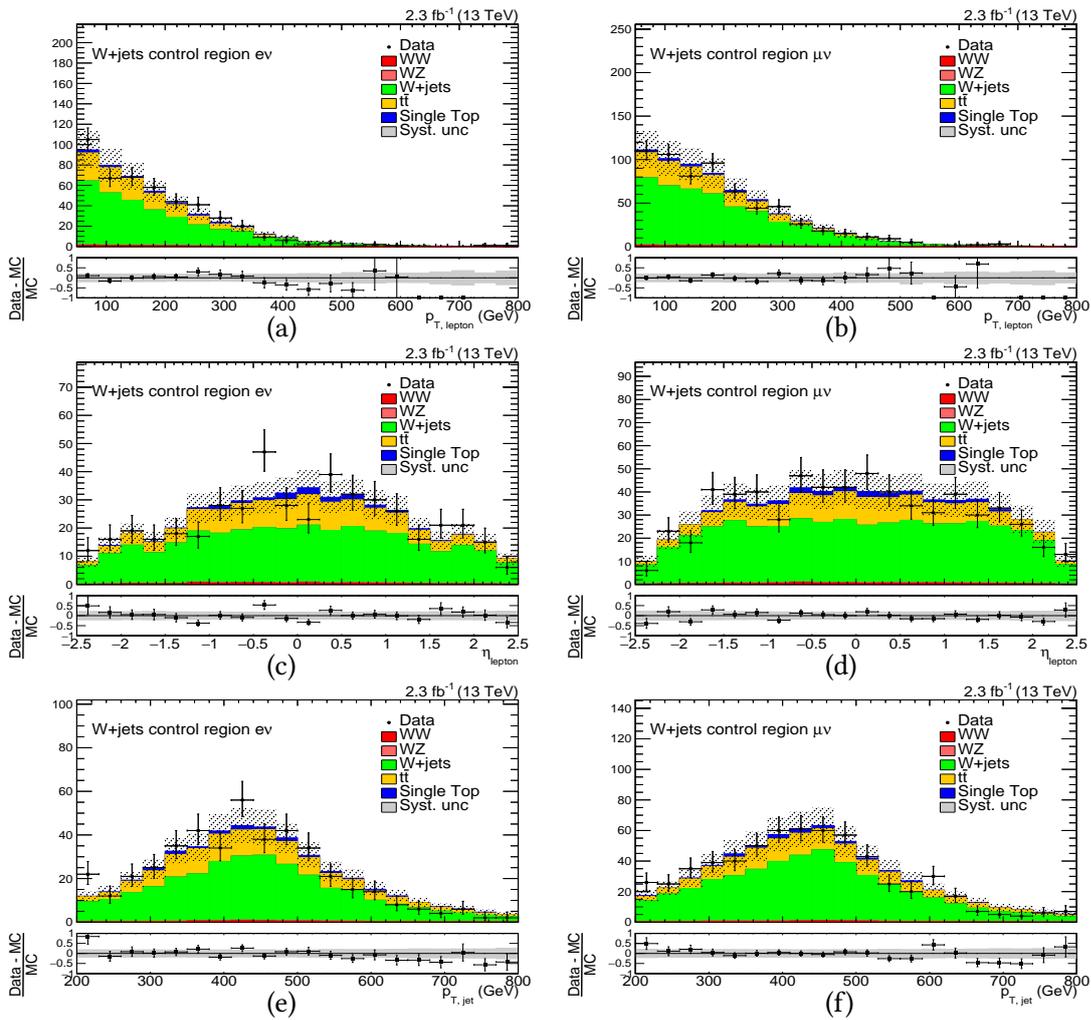


Figure G.2: MC-data and shapes of the  $M_{WV}$  distributions in the sideband for electron- (left) and muon-channel (right), for single top (top) and diboson (bottom).

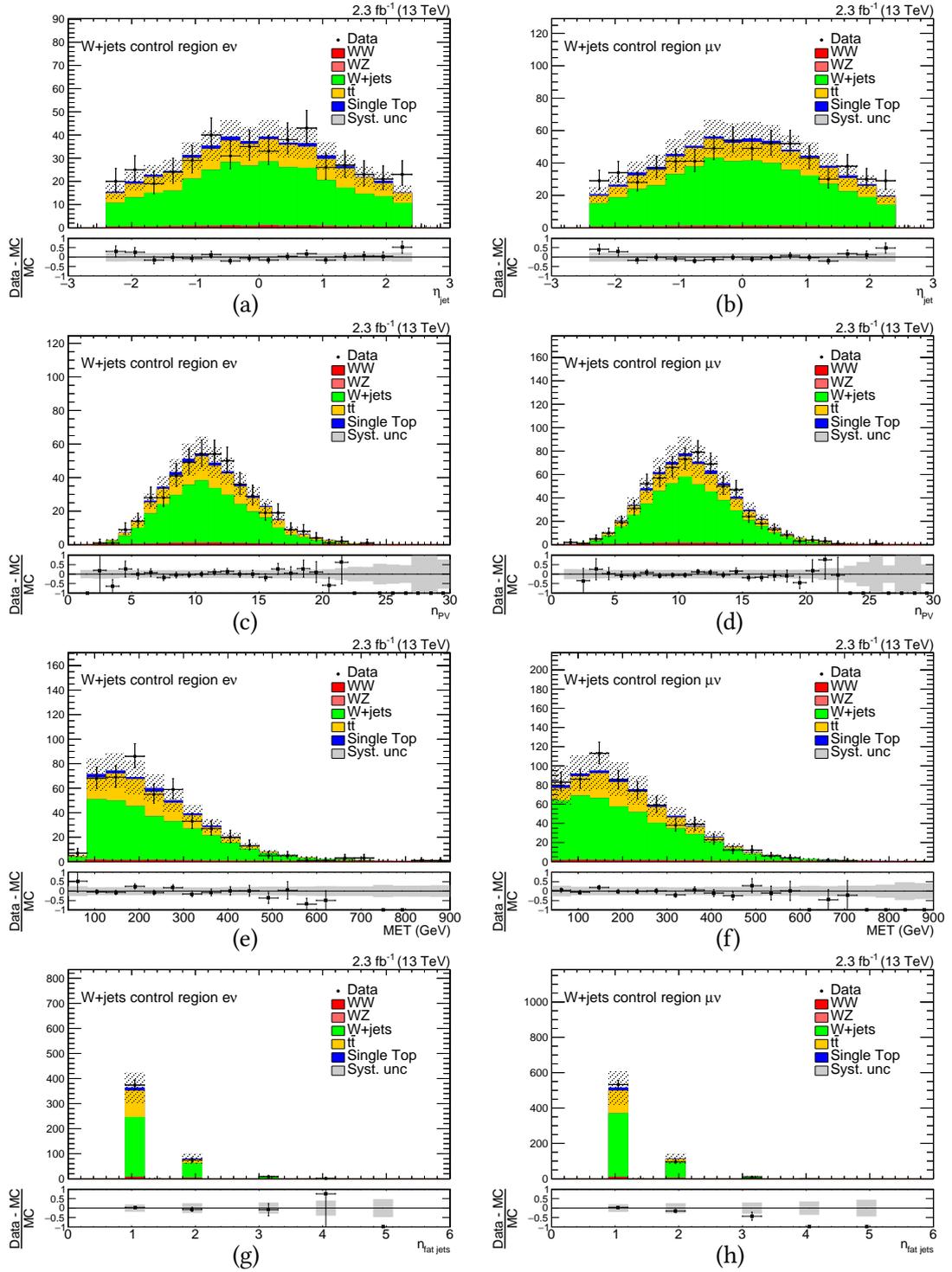
# H Additional plots for control regions

## H.1 W+jets control region

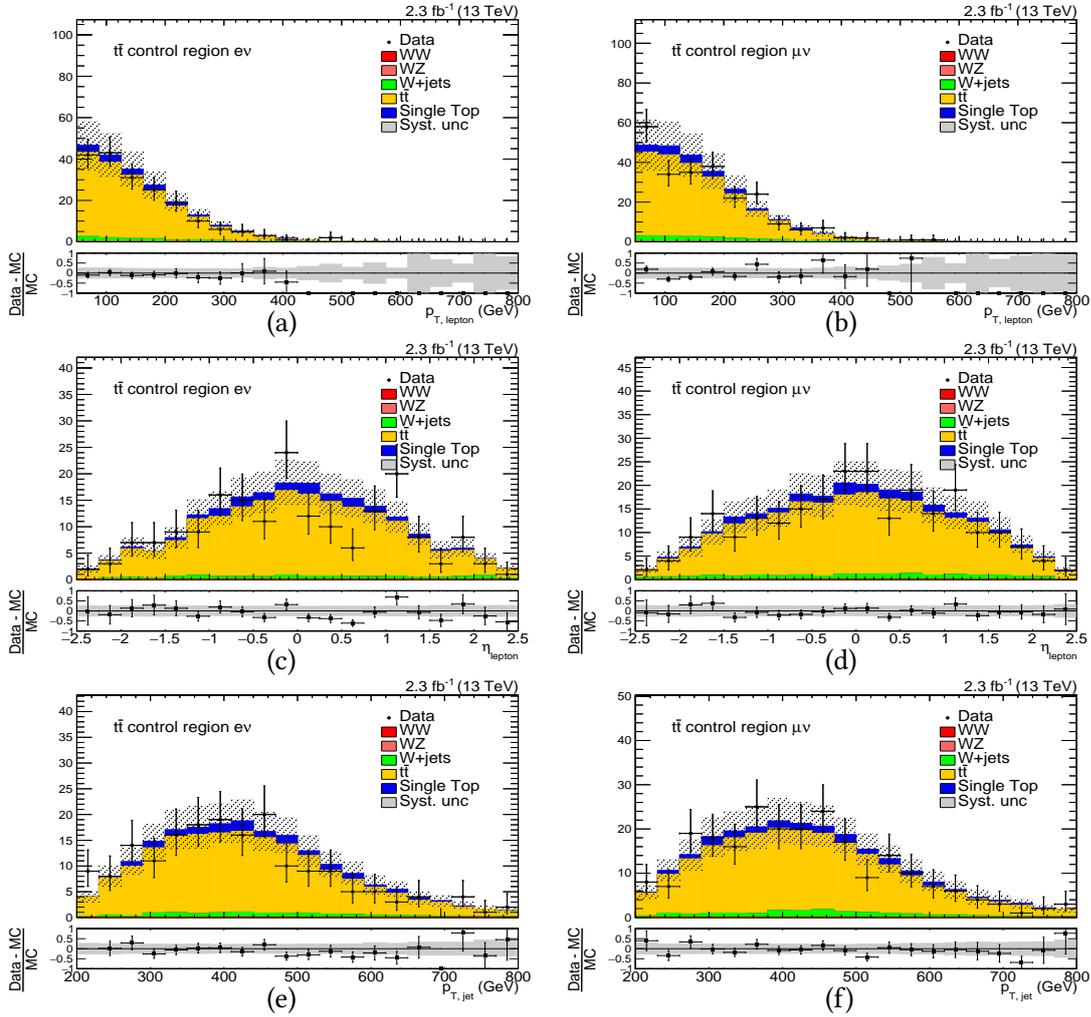


**Figure H.1:** Comparison between data and simulation of the lepton  $p_T$  (a,b), lepton  $\eta$  (c,d), leading jet  $p_T$  (e,f) distributions in the W+jets control region. The electron channel is shown on the left, while the muon channel is shown on the right.

## H Additional plots for control regions

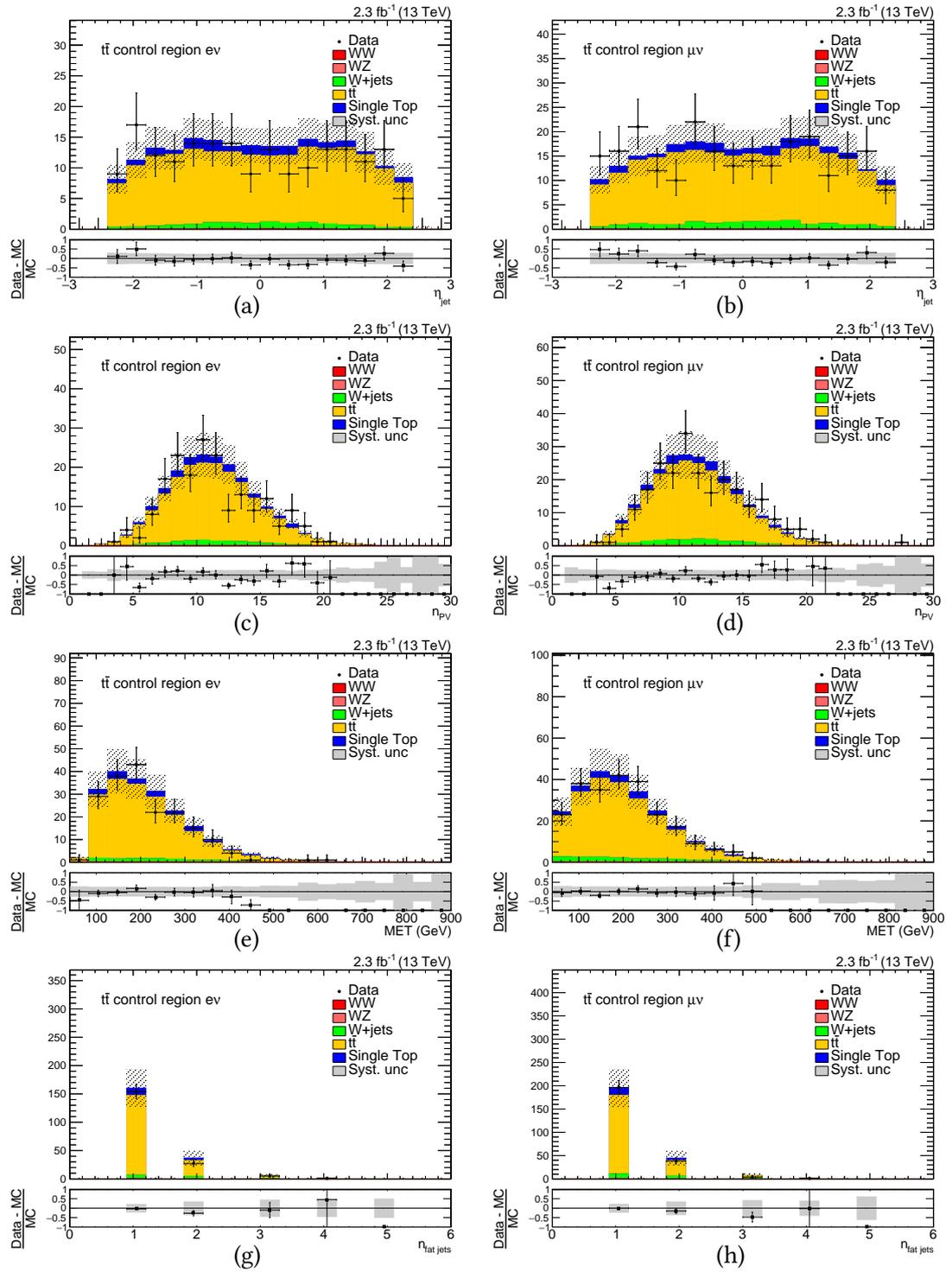


**Figure H.2:** Comparison between data and simulation of leading jet  $\eta$  (a,b), a number of reconstructed vertices (c,d), missing transverse energy (e,f), a number of reconstructed AK8 jets (g,h) distributions in the W+jets control region. The electron channel is shown on the left, while the muon channel is shown on the right.

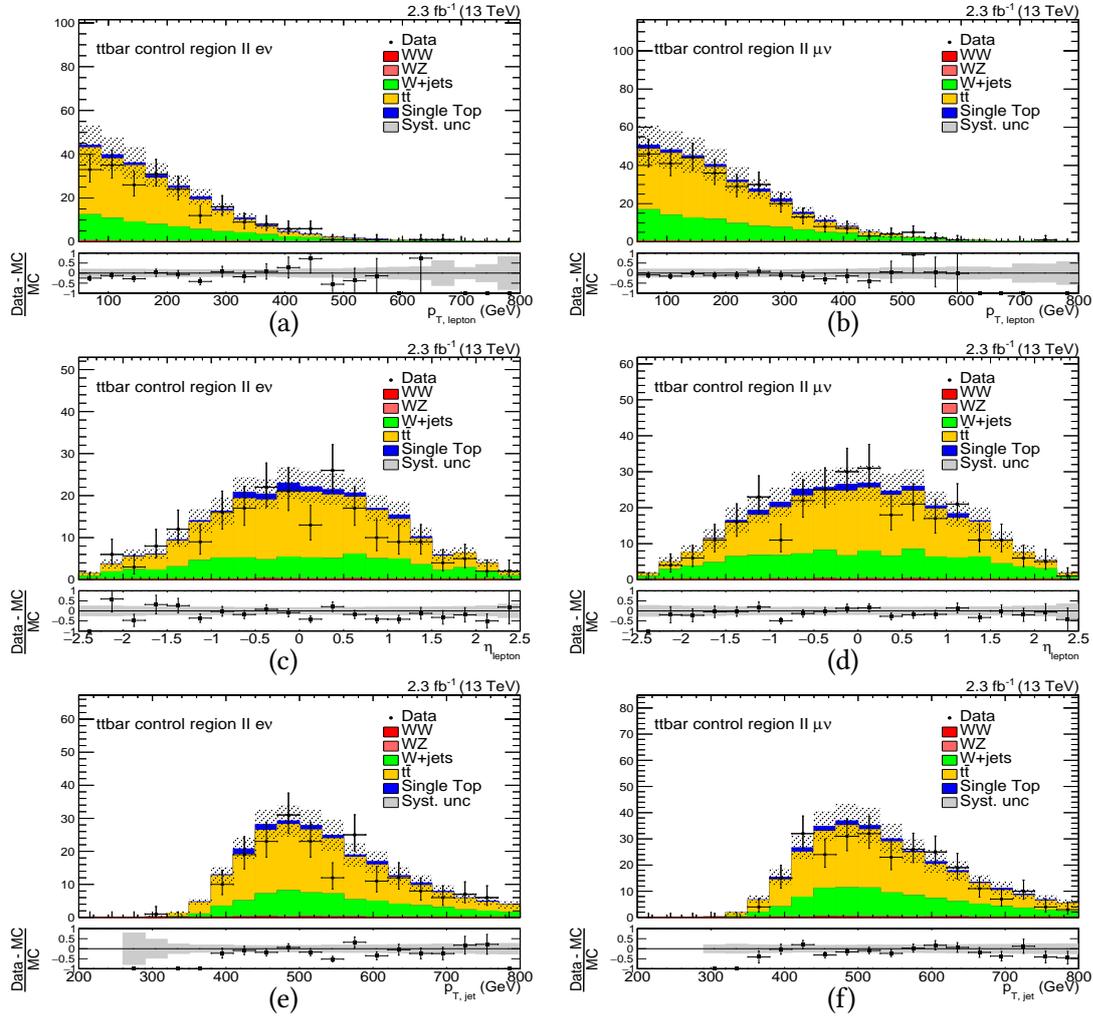
H.2  $t\bar{t}$  control region

**Figure H.3:** Comparison between data and simulation of the lepton  $p_T$  (a,b), lepton  $\eta$  (c,d), leading jet  $p_T$  (e,f) distributions in the  $t\bar{t}$  control region. The electron channel is shown on the left, while the muon channel is shown on the right.

## H Additional plots for control regions

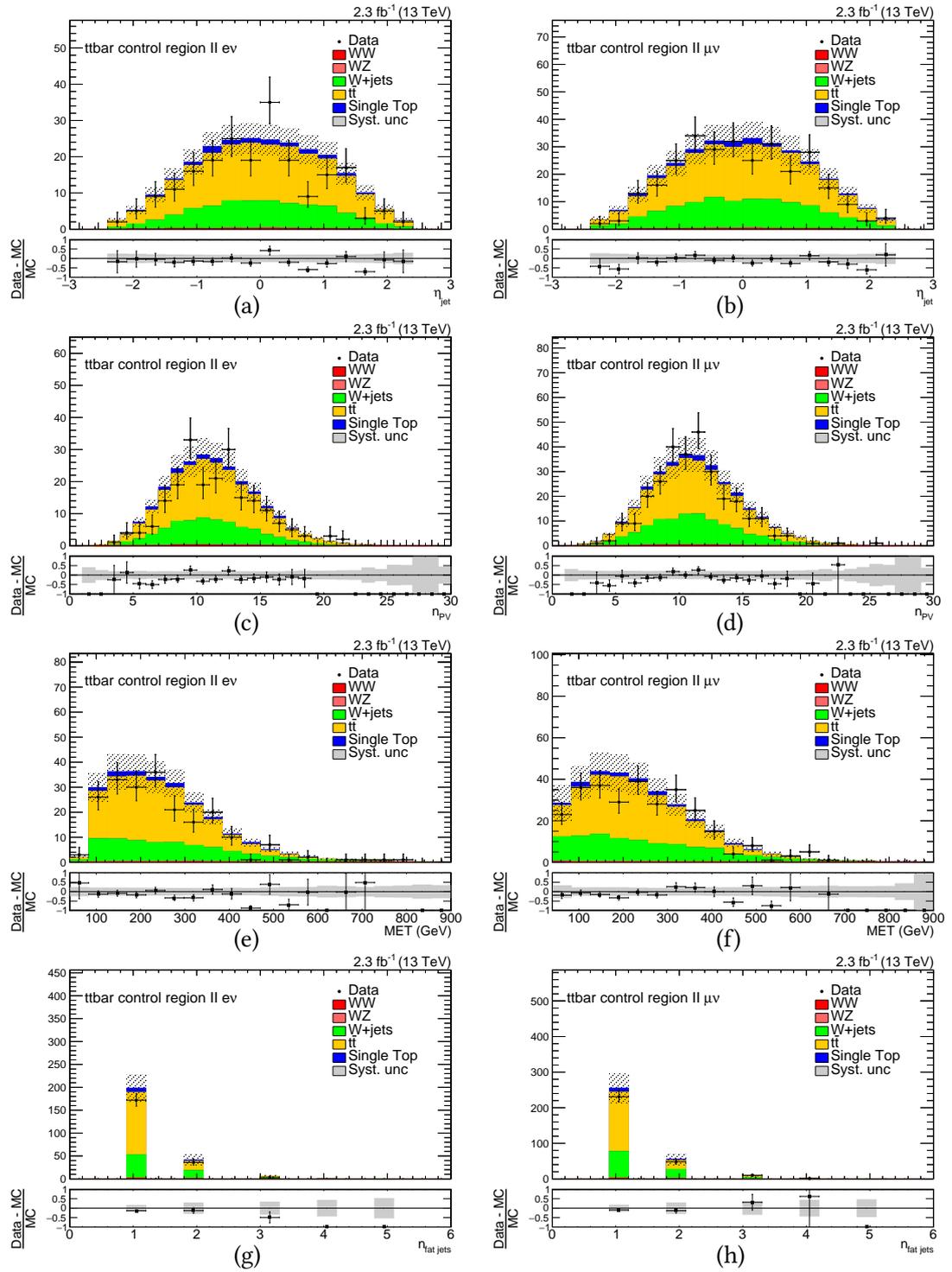


**Figure H.4:** Comparison between data and simulation of leading jet  $\eta$  (a,b), a number of reconstructed vertices (c,d), missing transverse energy (e,f), a number of reconstructed AK8 jets (g,h) distributions in the  $t\bar{t}$  control region. The electron channel is shown on the left, while the muon channel is shown on the right.

H.3  $t\bar{t}$  control region II

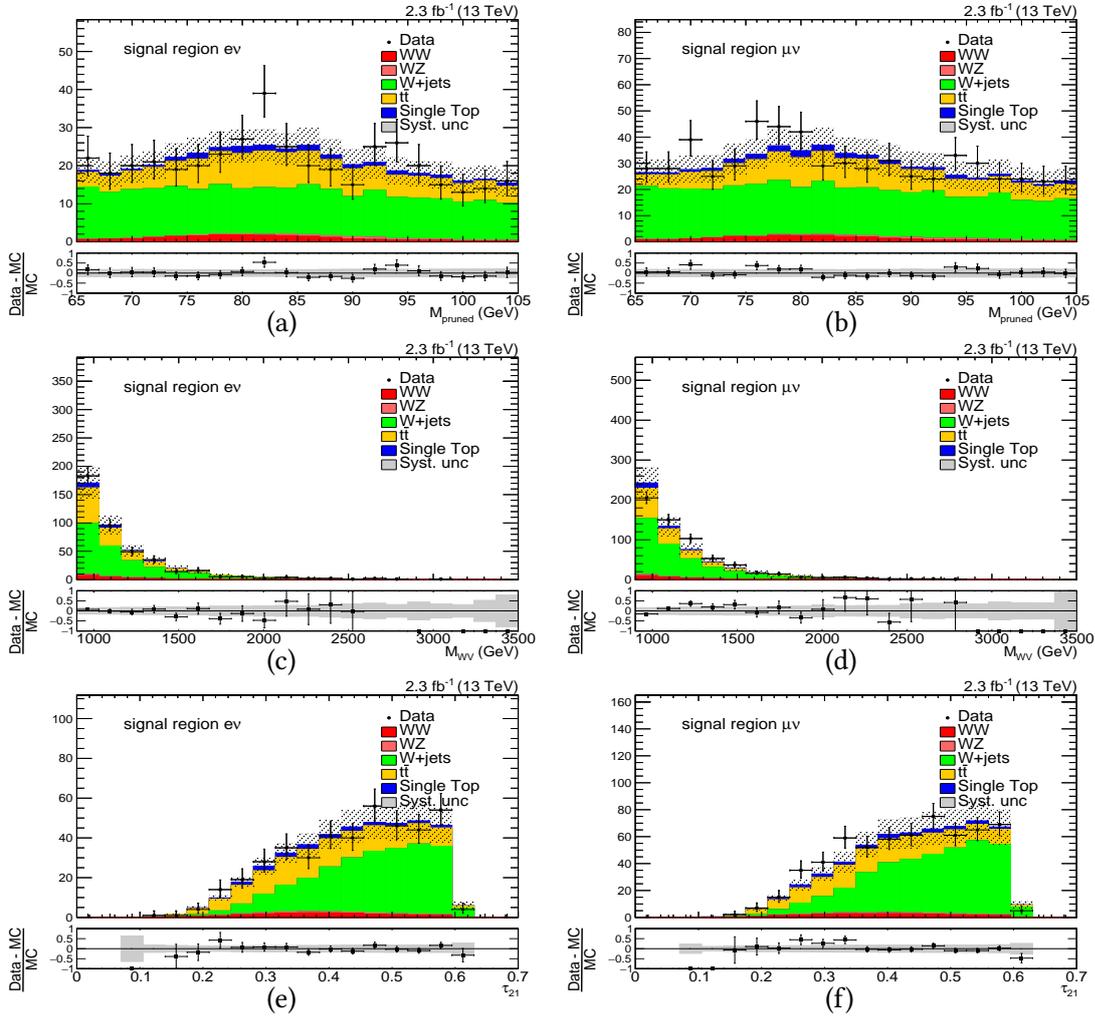
**Figure H.5:** Comparison between data and simulation of the lepton  $p_T$  (a,b), lepton  $\eta$  (c,d), leading jet  $p_T$  (e,f) distributions in the  $t\bar{t}$  control region II. The electron channel is shown on the left, while the muon channel is shown on the right.

## H Additional plots for control regions

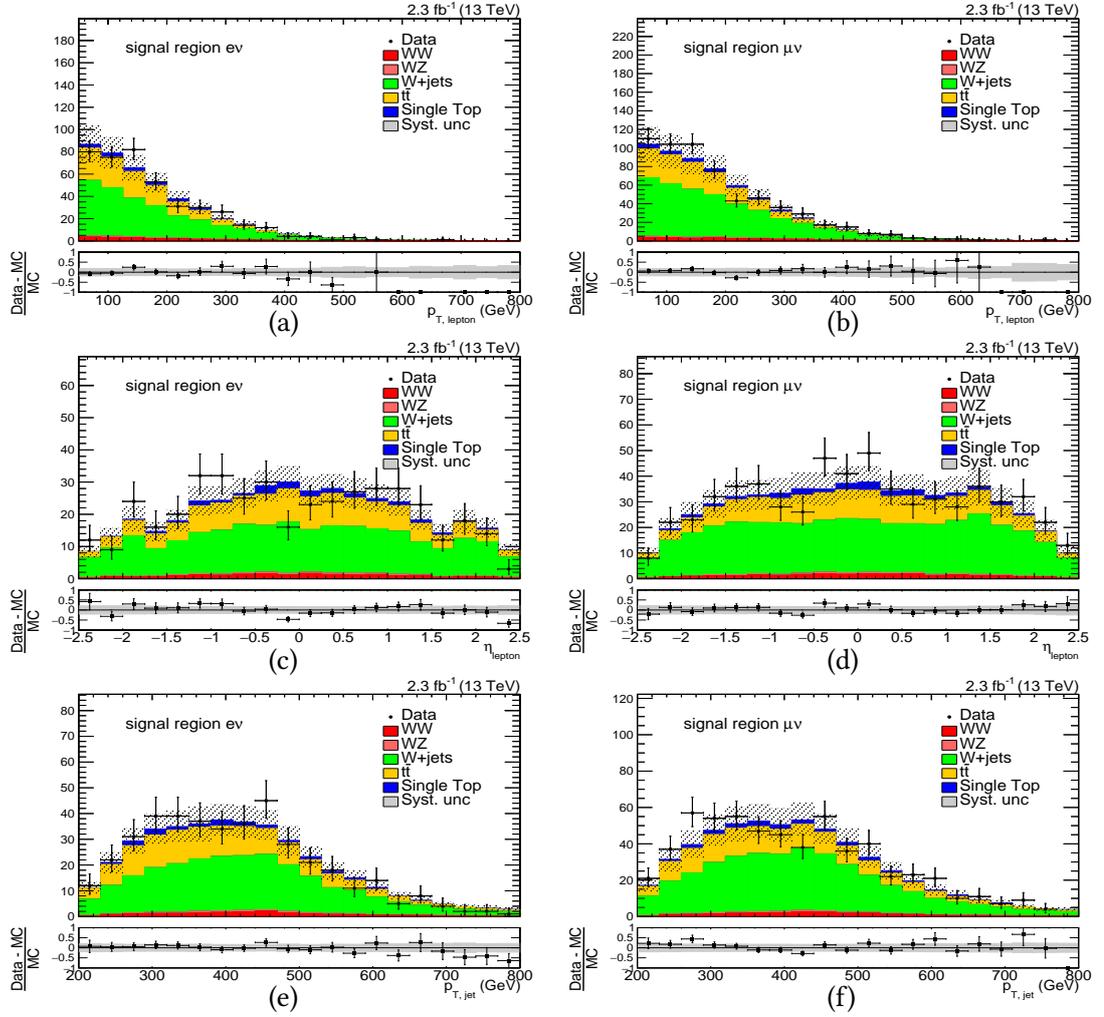


**Figure H.6:** Comparison between data and simulation of leading jet  $\eta$  (a,b), a number of reconstructed vertices (c,d), missing transverse energy (e,f), a number of reconstructed AK8 jets (g,h) distributions in the  $t\bar{t}$  control region II. The electron channel is shown on the left, while the muon channel is shown on the right.

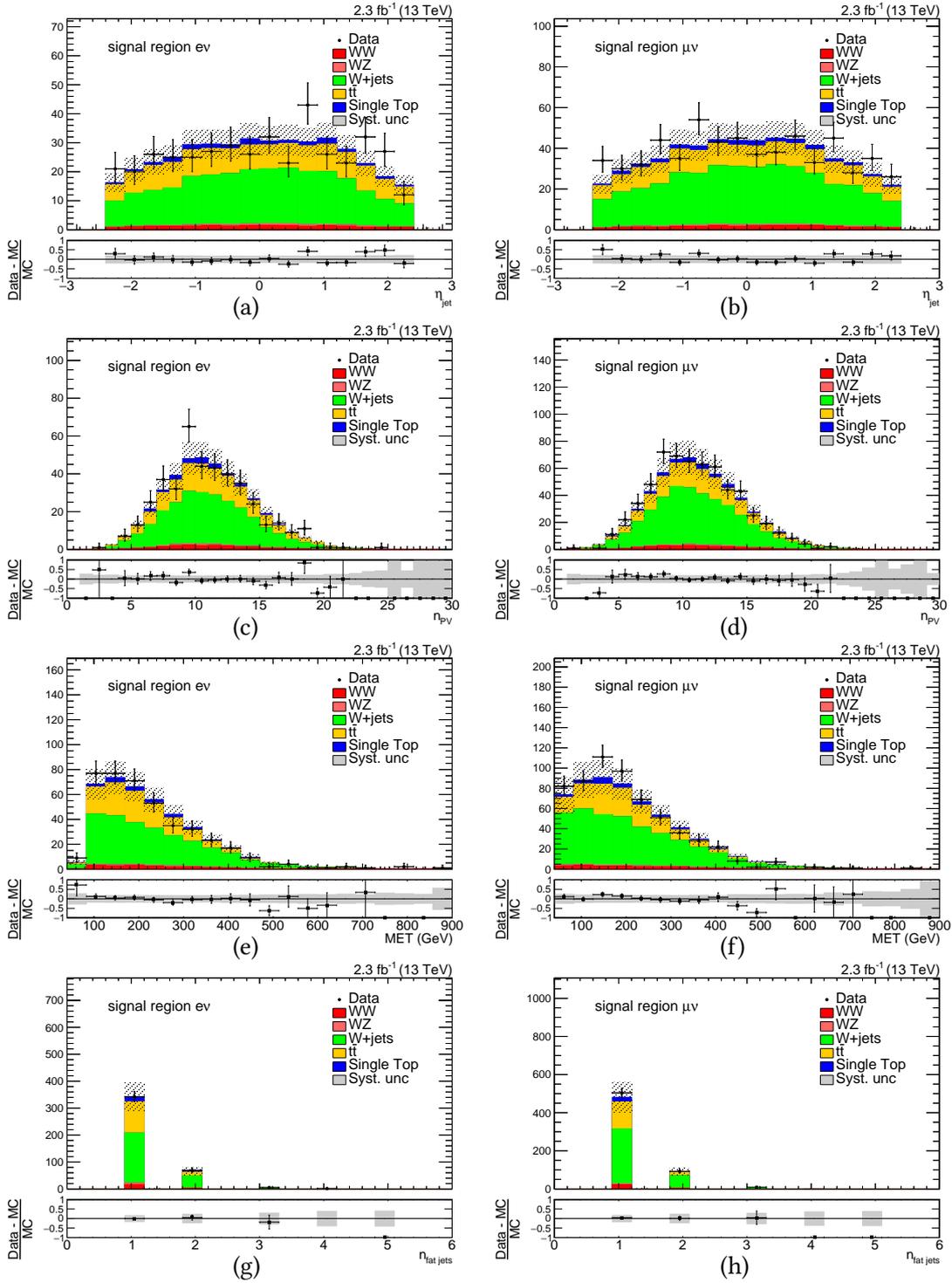
## H.4 signal region



**Figure H.7:** Comparison between data and simulation of the  $M_{\text{pruned}}$  (a,b),  $M_{WV}$  (c,d) and  $\tau_{21}$  (e,f) distributions in the signal region. The electron channel is shown on the left, while the muon channel is shown on the right.

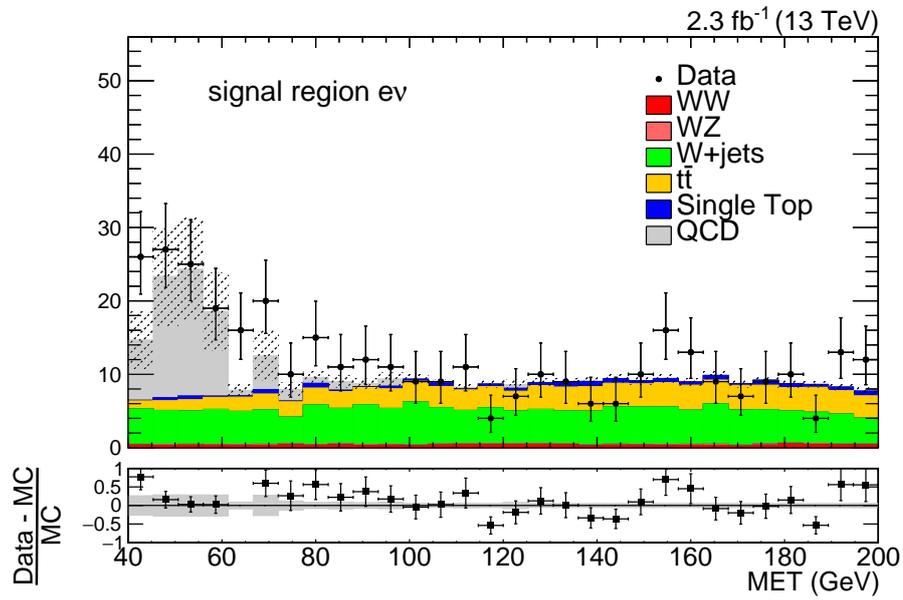


**Figure H.8:** Comparison between data and simulation of the lepton  $p_T$  (a,b), lepton  $\eta$  (c,d), leading jet  $p_T$  (e,f) distributions in the  $t\bar{t}$  control region II. The electron channel is shown on the left, while the muon channel is shown on the right.



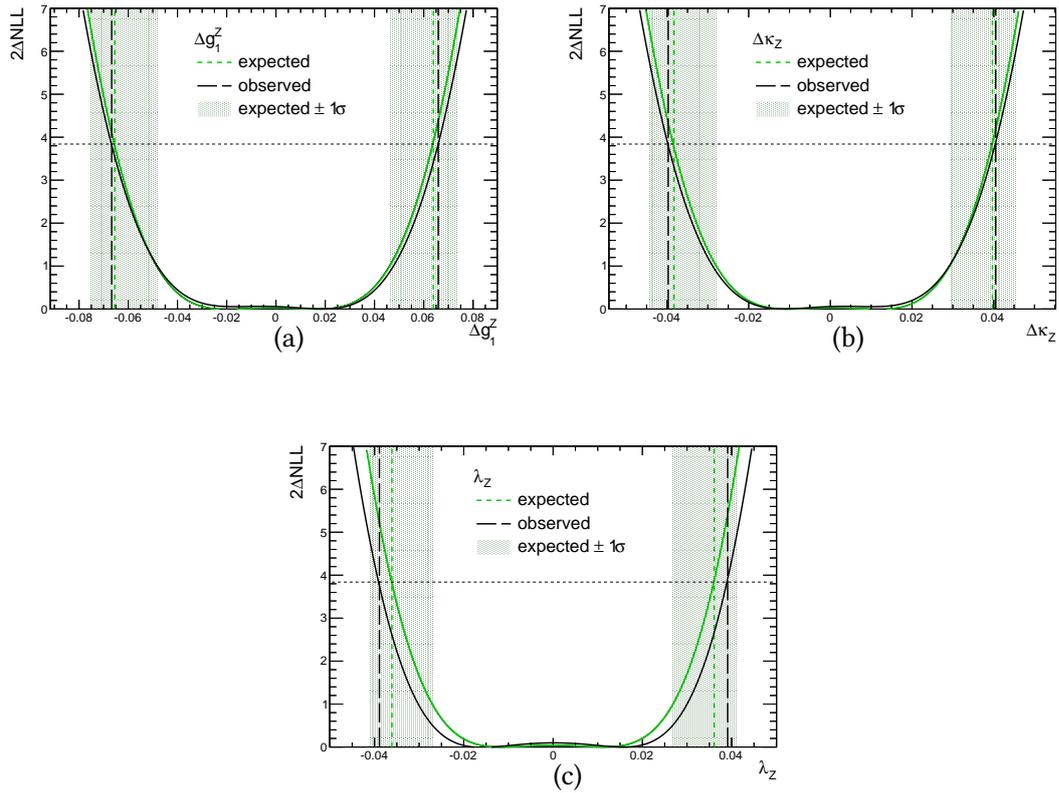
**Figure H.9:** Comparison between data and simulation of leading jet  $\eta$  (a,b), a number of reconstructed vertices (c,d), missing transverse energy (e,f), a number of reconstructed AK8 jets (g,h) distributions in the signal region II. The electron channel is shown on the left, while the muon channel is shown on the right.

## H.5 QCD fakes in the electron channel



**Figure H.10:** Illustration of the effect from the QCD fakes in the electron channel.

# I Additional plots for 1-dimensional limits



**Figure I.1:**  $\Delta NLL$  distributions for aTGC-parameters in the vertex parametrization:  $\Delta g_1^Z$  (a),  $\Delta \kappa_Z$  (b),  $\lambda_Z$  (c). Shown are the expected (green straight line) and observed (black straight line) distributions. Intersection between black dashed (green dashed) lines corresponds to observed (expected) 95% C.L. The shaded area shows the region between 99 % C.L. and 68 % C.L. ( $1\sigma$ -band around the expected limit). The observed and expected limits are in agreement with  $1\sigma$ -band [191]. [191].



# List of Figures

1.1	Comparison of coupling constants of the QCD and QED interaction. QCD coupling diverges at low energies (color confinement) and decreases at low energies (asymptotic freedom). QED coupling is small at low energies and diverges at high energies (Landau pole) [27]. . . . .	4
1.2	"Mexican hat" potential [36]. . . . .	6
1.3	The diagram illustrating WW or WZ with anomalous triple gauge coupling affecting WW $\gamma$ and WWZ vertex. . . . .	9
1.4	Limits on WWZ anomalous triple gauge couplings [54]. . . . .	11
2.1	Layout of the LEP tunnel and LHC infrastructures [58]. . . . .	13
2.2	The CERN accelerator complex [64]. . . . .	15
2.3	Cross-sections of different processes at the LHC and Tevatron [66]. . . . .	16
2.4	The CMS detector [67]. . . . .	17
2.5	Cross-section of the coil of the CMS magnet [70]. . . . .	18
2.6	The layout of the CMS tracker system [71]. . . . .	19
2.7	The layout of the CMS pixel system [72]. . . . .	20
2.8	The layout of the electromagnetic calorimeter [56]. . . . .	21
2.9	History of ECAL response with laser data [74]. The response is up to 6 % in the barrel and reaches up to 30 % at $ \eta  \approx 2.5$ . The response is up to 70% in the region closest to the beam pipe. . . . .	22
2.10	The layout of the hadron calorimeter [56]. . . . .	23
2.11	The layout of the CMS muon system after the Long Shutdown 1 (LS1) Upgrade, 2013-2014 [80]. . . . .	23
2.12	Cumulative offline luminosity versus day delivered to (blue), and recorded by CMS (orange) during stable beams and for p-p collisions at 13 TeV center-of-mass energy in 2015. Given is the luminosity as determined from counting rates measured by the luminosity detectors after offline validation [94]. . . . .	26
3.1	The basic structure of the event simulation [97]. . . . .	27
3.2	The illustration of the event containing a jet originating from a b-quark. The jet contains B-meson which decays at the secondary vertex. The secondary vertex is displaced from the primary vertex by the decay length of B-meson $L_{xy}$ [143].	39
4.1	Illustration for the pruning algorithm (adopted from M.Malberti [153]). . . . .	42

4.2	Illustration for the trimming algorithm, $p_{Tcut} = p_{Tfrac}\lambda_{hard}$ (adopted from M.Malberti [153]). . . . .	43
4.3	Jet mass distributions for various grooming algorithms for QCD-jets. PF (a) and PF+CHS (b) are used as input for the jet clustering [147]. . . . .	44
4.4	Mass response distributions for W-jets for different grooming algorithms: trimming (a), pruning (b), softdrop (c). PF+CHS is used as input for the jet clustering [147]. . . . .	45
4.5	Jet mass response resolution as a function of number of reconstructed vertices for pruning (a), trimming (b), softdrop (c). PF is used as input for the jet clustering. . . . .	46
4.6	Leading jet $\tau_{21}$ distribution for QCD-jets (a) and W-jets (b) [147]. . . . .	47
4.7	Average $\tau_{21}$ as a function of number of reconstructed vertices [147]. . . . .	48
4.8	Summary plots for the jet mass resolution (a) and the jet mass response resolution (b) [147] for different parameters of grooming algorithms. $m$ is the ungroomed jet mass with four vector-safe subtraction and $m_{raw}$ is the ungroomed and uncorrected jet mass. The resolution is evaluated with RMS (squares) and fitted $\sigma$ (triangles). . . . .	49
5.1	The LO Feynman-diagram for the signal in the semileptonic decay. One W boson decays to a lepton and a neutrino and another W (Z) boson decays to quarks. . . . .	51
5.2	Feynman diagrams for $t\bar{t}$ (a) and W+jets (b). . . . .	53
5.3	Feynman diagrams for WZ production: s-channel (a) and t-channel (b). . . . .	53
5.4	Feynman diagrams for single top production: s-channel (a), t-channel (b) and tW-channel (c). . . . .	54
5.5	Illustration of signal region and control regions used in the analysis. . . . .	59
5.6	Event display of the event with the highest diboson mass. . . . .	60
5.7	Pileup profiles in data and the simulation. . . . .	61
5.8	B-tagging (mis-tagging) efficiencies for $t\bar{t}$ sample, muon channel for b-jets (a), c-jets (b) and light flavour jets (c). . . . .	63
5.9	Efficiency of the $HLT\_Ele27\_WPLoose\_Gsf$ as a function of the supercluster transverse energy. . . . .	64
6.1	$M_{pruned}$ distribution of the signal with $c_{ww} = -12.0 \text{ TeV}^{-2}$ Muon channel on the left and electron channel on the right. . . . .	68
6.2	$M_{pruned}$ distribution of the signal with $c_w = -20.0 \text{ TeV}^{-2}$ . Muon channel on the left and electron channel on the right. . . . .	68
6.3	$M_{pruned}$ distribution of the signal with $c_b = -60.0 \text{ TeV}^{-2}$ . Muon channel on the left and electron channel on the right. . . . .	68
6.4	Normalized yields for different values of the aTGC-parameters in the WW-category for electron- (left) and muon-channel (right) with fitted quadratic function. From top to bottom: $\frac{c_{WWW}}{\Lambda^2}, \frac{c_W}{\Lambda^2}, \frac{c_B}{\Lambda^2}$ . . . . .	70

6.5	MC data and fit of signal model for all parameters set to zero (black) and one parameter set to a non-zero value (blue) for electron- (left) and muon-channel (right), WW-category. From top to bottom: $\frac{c_{WW}}{\Lambda^2} = 12 \text{ TeV}^{-2}$ , $\frac{c_W}{\Lambda^2} = 20 \text{ TeV}^{-2}$ , $\frac{c_B}{\Lambda^2} = 60 \text{ TeV}^{-2}$ . The ratio plot corresponds to the simulation minus the fit-function divided by the bin-error for the signal function with non-zero aTGC-parameter. . . . .	73
6.6	Signal model and MC on generator level (WZ only) for different values of $\frac{c_W}{\Lambda^2}$ in the electron channel. The grey lines shows the signal model for increasing $\frac{c_W}{\Lambda^2}$ in steps of $1 \text{ TeV}^{-2}$ . . . . .	74
7.1	MC-data and shapes of the $M_{\text{pruned}}$ background distributions for electron- (left) and muon-channel (right). From top to bottom: $t\bar{t}$ , W+jets, single Top, diboson. . . . .	76
7.2	Result of the normalization fit in the $M_{\text{pruned}}$ -spectrum for electron- (left) and muon-channel (right). From top to bottom: pre-fit functions, fit result [55], fit result using an alternative W+jets function. . . . .	78
7.3	Comparison between data and simulation of the $M_{\text{pruned}}$ (top), $M_{WV}$ (middle) and $\tau_{21}$ (bottom) distributions in the W+jets control region. The electron channel is shown on the left, while the muon channel is shown on the right. . . . .	80
7.4	Comparison between data and simulation of the $M_{\text{pruned}}$ (top), $M_{WV}$ (middle) and $\tau_{21}$ (bottom) distributions in the $t\bar{t}$ control region. The electron channel is shown on the left, while the muon channel is shown on the right. . . . .	81
7.5	Comparison between data and the simulation in the $t\bar{t}$ control region splitting the $t\bar{t}$ background into matched and unmatched categories, muon channel on left and electron channel on the right. Only statistical uncertainties are shown for the simulation. . . . .	83
7.6	Comparison between data and the simulation in the $t\bar{t}$ control region II splitting the $t\bar{t}$ background into matched and unmatched categories, muon channel on left and electron channel on the right. Only statistical uncertainties are shown for the simulation. . . . .	83
7.7	Comparison between data and the simulation in the signal region (extended) splitting the $t\bar{t}$ background into matched and unmatched categories, muon channel on left and electron channel on the right. Only statistical uncertainties are shown for the simulation. . . . .	84
7.8	MC-data and shapes of the $M_{WV}$ distributions in the signal region, WW-category, for electron- (left) and muon-channel (right). From top to bottom: $t\bar{t}$ , W+jets, single top, diboson. . . . .	85
7.9	MC-data and shapes of the $M_{WV}$ distributions in the signal region, WZ-category, for electron- (left) and muon-channel (right). From top to bottom: $t\bar{t}$ , W+jets, single top, diboson. . . . .	86
7.10	Shape and data of the $M_{WV}$ background in the sideband region for electron- (top) and muon-channel (bottom). . . . .	87
7.11	Alpha ratio for the WW-category (top) and the WZ-category (bottom) in the electron- (left) and muon-channel (right) [55]. . . . .	88

7.12	Data and final shapes of the $M_{WV}$ background for electron (left) and muon channel (right) in the WW- (top) and WZ-category (bottom) [55]. . . . .	89
8.1	Two dimensional limits on the aTGC-parameters. Shown are the expected contours for 68% C.L. (blue), 95% C.L. (green) and 99% C.L. (red), for $\frac{c_{WWW}}{\Lambda^2} - \frac{c_W}{\Lambda^2}$ (a), $\frac{c_{WWW}}{\Lambda^2} - \frac{c_B}{\Lambda^2}$ (b) and $\frac{c_W}{\Lambda^2} - \frac{c_B}{\Lambda^2}$ (c). The black line shows the region compatible with the observed data at 95% C.L. [55]. . . . .	96
8.2	Two dimensional limits on the aTGC-parameters. Shown are the expected contours for 68% C.L. (blue), 95% C.L. (green) and 99% C.L. (red), for $\Delta\kappa_Z - \lambda_Z$ (a), $\Delta g_1^Z - \lambda_Z$ (b) and $\Delta g_1^Z - \Delta\kappa_Z$ (c). The black line shows the region compatible with the observed data at 95% C.L. [55]. . . . .	97
8.3	$\Delta NLL$ distributions for aTGC-parameters: $\frac{c_{WWW}}{\Lambda^2}$ (a), $\frac{c_{WWW}}{\Lambda^2}$ (b), $\frac{c_B}{\Lambda^2}$ (c). Shown are the expected (green straight line) and observed (black straight line) distributions. Intersection between black dashed (green dashed) lines corresponds to observed (expected) 95% C.L. The shaded area shows the region between 99% C.L. and 68% C.L. ( $1\sigma$ -band around the expected limit). The observed and expected limits are in agreement with $1\sigma$ -band [191]. . . . .	98
8.4	Limits on WWZ anomalous triple gauge couplings [54]. . . . .	100
A.1	Raw and ungroomed jet mass are shown on with PF (a) and PF+CHS (b) used as inputs for the jet clustering. . . . .	105
A.2	Jet mass distributions for various grooming algorithms for W-jets: trimming (a,b), pruning (c,d) and softdrop (e,f). PF (a,c,e) and PF+CHS (b,d,f) are used as inputs for the jet clustering. . . . .	106
A.3	Jet mass distributions for various grooming algorithms for W-jets: trimming (b), pruning (c) and softdrop (d). Raw and ungroomed jet mass are shown on (a). PF+PUPPI is used as input for the jet clustering. . . . .	107
B.1	Mass response distributions for W-jets for different grooming algorithms: trimming (a), pruning (b), softdrop(c). PF is used as input for the jet clustering. . . . .	109
B.2	Mass response distributions for W-jets for different grooming algorithms: trimming (a), pruning (b), softdrop(c). PF+PUPPI is used as input for the jet clustering. . . . .	110
B.3	Raw and ungroomed jet mass response for W-jets. PF (a), PF+CHS (b) and PF+PUPPI (c) are used as input for the jet clustering. . . . .	111
C.1	Jet mass resolution as a function of number of reconstructed vertices for PF (left) and PF+CHS (right). . . . .	113
C.2	Mean jet mass as a function of number of reconstructed vertices for PF (left) and PF + CHS (right). . . . .	114
C.3	Mean mass response as a function of number of reconstructed vertices for PF (left) and PF + CHS (right). . . . .	115
C.4	Jet mass response resolution for QCD-jets (left) and W-jets (right) as a function of number of reconstructed vertices. PF+CHS is used as input for the jet clustering [147]. . . . .	116

C.5	Response resolution (left) and mass resolution (right) as a function of number of reconstructed vertices for PF+PUPPI. . . . .	117
C.6	Mean jet mass (left) and mean response as a function of number of reconstructed vertices for PF+PUPPI. . . . .	118
C.7	Mean jet mass distribution for QCD-jets as a function of number of reconstructed vertices. PF (left) and PF+CHS (right) are used as input for the jet clustering [147].	119
D.1	B-tagging (mis-tagging) efficiencies for $t\bar{t}$ sample, electron channel for b-jets (a), c-jets (b) and light flavour jets (c). . . . .	121
D.2	B-tagging (mis-tagging) efficiencies for WZ sample, muon channel for b-jets (a), c-jets (b) and light flavour jets (c). . . . .	122
D.3	B-tagging (mis-tagging) efficiencies for WZ sample, muon channel for b-jets (a), c-jets (b) and light flavour jets (c). . . . .	123
E.1	Normalized yields for different values of the aTGC-parameters in the WZ-category for electron- (left) and muon-channel (right) with fitted quadratic function. From top to bottom: $\frac{c_{WWW}}{\Lambda^2}$ , $\frac{c_W}{\Lambda^2}$ , $\frac{c_B}{\Lambda^2}$ . . . . .	125
E.2	MC data and fit of signal model for all parameters set to zero (black) and one parameter set to a non-zero value (blue) for electron- (left) and muon-channel (right), WW-category. From top to bottom: $\frac{c_{WWW}}{\Lambda^2} = -12 \text{ TeV}^{-2}$ , $\frac{c_W}{\Lambda^2} = -20 \text{ TeV}^{-2}$ , $\frac{c_B}{\Lambda^2} = -60 \text{ TeV}^{-2}$ . The ratio plot corresponds to the simulation minus the fit-function divided by the bin-error for the signal function with non-zero aTGC-parameter. . . . .	126
E.3	Signal model and MC on generator level (WZ only) for different values of $\frac{c_W}{\Lambda^2}$ in the muon channel. The grey lines shows the signal model for increasing $\frac{c_W}{\Lambda^2}$ in steps of $1 \text{ TeV}^{-2}$ . . . . .	127
E.4	The effect of shape uncertainty for $a_{cb}$ (a,b), $a_{cW}$ (c,d), $a_{cWWW}$ (e,f), WW category, muon channel on left and electron channel on the right. Uncertainty is shown with shaded area. . . . .	129
E.5	The effect of shape uncertainty for $a_{cb}$ (a,b), $a_{cW}$ (c,d), $a_{cWWW}$ (e,f), WZ category, muon channel on left and electron channel on the right. Uncertainty is shown with shaded area. . . . .	130
F.1	Pull plots of the W+jets normalization extracted with toy MCs randomizing the shape parameters as data input for the electron- (left) and muon-channel (right) in the WW- (top) and WZ-category (bottom). . . . .	131
G.1	MC-data and shapes of the $M_{WV}$ distributions in the sideband for electron- (left) and muon-channel (right), for $t\bar{t}$ (top) and W+jets (bottom). . . . .	133
G.2	MC-data and shapes of the $M_{WV}$ distributions in the sideband for electron- (left) and muon-channel (right), for single top (top) and diboson (bottom). . . . .	134
H.1	Comparison between data and simulation of the lepton $p_T$ (a,b), lepton $\eta$ (c,d), leading jet $p_T$ (e,f) distributions in the W+jets control region. The electron channel is shown on the left, while the muon channel is shown on the right. . . . .	135

H.2	Comparison between data and simulation of leading jet $\eta$ (a,b), a number of reconstructed vertices (c,d), missing transverse energy (e,f), a number of reconstructed AK8 jets (g,h) distributions in the W+jets control region. The electron channel is shown on the left, while the muon channel is shown on the right. . . . .	136
H.3	Comparison between data and simulation of the lepton $p_T$ (a,b), lepton $\eta$ (c,d), leading jet $p_T$ (e,f) distributions in the $t\bar{t}$ control region. The electron channel is shown on the left, while the muon channel is shown on the right. . . . .	137
H.4	Comparison between data and simulation of leading jet $\eta$ (a,b), a number of reconstructed vertices (c,d), missing transverse energy (e,f), a number of reconstructed AK8 jets (g,h) distributions in the $t\bar{t}$ control region. The electron channel is shown on the left, while the muon channel is shown on the right. . . . .	138
H.5	Comparison between data and simulation of the lepton $p_T$ (a,b), lepton $\eta$ (c,d), leading jet $p_T$ (e,f) distributions in the $t\bar{t}$ control region II. The electron channel is shown on the left, while the muon channel is shown on the right. . . . .	139
H.6	Comparison between data and simulation of leading jet $\eta$ (a,b), a number of reconstructed vertices (c,d), missing transverse energy (e,f), a number of reconstructed AK8 jets (g,h) distributions in the $t\bar{t}$ control region II. The electron channel is shown on the left, while the muon channel is shown on the right. . . . .	140
H.7	Comparison between data and simulation of the $M_{\text{pruned}}$ (a,b), $M_{WV}$ (c,d) and $\tau_{21}$ (e,f) distributions in the signal region. The electron channel is shown on the left, while the muon channel is shown on the right. . . . .	141
H.8	Comparison between data and simulation of the lepton $p_T$ (a,b), lepton $\eta$ (c,d), leading jet $p_T$ (e,f) distributions in the $t\bar{t}$ control region II. The electron channel is shown on the left, while the muon channel is shown on the right. . . . .	142
H.9	Comparison between data and simulation of leading jet $\eta$ (a,b), a number of reconstructed vertices (c,d), missing transverse energy (e,f), a number of reconstructed AK8 jets (g,h) distributions in the signal region II. The electron channel is shown on the left, while the muon channel is shown on the right. . . . .	143
H.10	Illustration of the effect from the QCD fakes in the electron channel. . . . .	144
I.1	$\Delta NLL$ distributions for aTGC-parameters in the vertex parametrization: $\Delta g_1^Z$ (a), $\Delta \kappa_Z$ (b), $\lambda_Z$ (c) . Shown are the expected (green straight line) and observed (black straight line) distributions. Intersection between black dashed (green dashed) lines corresponds to observed (expected) 95% C.L. The shaded area shows the region between 99 % C.L. and 68 % C.L. ( $1\sigma$ -band around the expected limit). The observed and expected limits are in agreement with $1\sigma$ -band [191]. [191]. . . . .	145

# List of Tables

1.1	List of fermions in the Standard Model. Masses are quoted from [21,22]. Electrical charge is quoted in units of electron charge. . . . .	1
1.2	Branching ratios of W boson [21]. . . . .	2
1.3	Branching ratios of Z boson [21]. . . . .	2
1.4	Hypercharge of leptons and quarks in the Standard Model. . . . .	7
1.5	Results of LEP-II combination [43]. . . . .	10
1.6	Limits on aTGC by D0 [44]. . . . .	10
3.1	Definitions of HEEP ID V6.1 selections [132]. . . . .	36
3.2	Jet energy resolution scale factors [140]. . . . .	38
3.3	PFJetID requirements [142]. . . . .	38
4.1	Summary of the parameters of grooming algorithms. . . . .	43
5.1	Values of anomalous triple gauge couplings used for the generation of signal events. . . . .	52
5.2	Signal and background samples. $N_{\text{events}}$ denotes the number of generated events. . . . .	55
5.3	Data samples . . . . .	56
5.4	Summary of the selection applied in the analysis. . . . .	58
5.5	Estimated normalization uncertainties in % for contributions derived from the simulation. Uncertainties smaller than 0.05% for all processes are not listed [55]. . . . .	66
6.1	Total uncertainties of signal shape slopes. Uncertainties are quoted in %. . . . .	72
7.1	Results of the fit to the $M_{\text{pruned}}$ distributions in the range [40., 150.] GeV [55]. The pre- and post-fit yields are presented together with their constraints (pre-fit) and resulting total uncertainties (post-fit). The W+jets contribution is allowed to float in the fit, therefore the pre-fit values do not have any constraint shown. The single top contribution is fixed in the fit. . . . .	79
8.1	Summary of background and signal yields in WW- and WZ-categories. Uncertainties for the single-top, diboson and $t\bar{t}$ contributions are evaluated as described in Section 5.4, while the uncertainty on W+jets is derived from the statistical uncertainty of the $M_{\text{pruned}}$ -fit and the fit with the alternative function [55]. . . . .	94
8.2	Expected and observed limits at 95% C.L. on single anomalous couplings (other couplings set to zero). . . . .	95

*List of Tables*

---

8.3	Expected and observed limits at 95% C.L. on single anomalous couplings (other couplings set to zero). . . . .	100
E.1	Estimated uncertainties of signal shape slope parameters in the WW-category, muon channel. Uncertainties are quoted in %. . . . .	128
E.2	Estimated uncertainties of signal shape slope parameters in the WW-category, electron channel. Uncertainties are quoted in %. . . . .	128
E.3	Estimated uncertainties of signal shape slope parameters in the WZ-category, muon channel. Uncertainties are quoted in %. . . . .	128
E.4	Estimated uncertainties of signal shape slope parameters in the WZ-category, electron channel. Uncertainties are quoted in %. . . . .	128

# Bibliography

- [1] G. Holton and S. Brush, *Physics, the Human Adventure: From Copernicus to Einstein and Beyond*. Rutgers University Press, 2001.
- [2] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, “Detection of the free neutrino: A Confirmation”, *Science* 124 (1956) 103–104.
- [3] *Gargamelle Neutrino Collaboration*, “Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment”, *Phys. Lett.* 46B (1973) 138–140.
- [4] *CMS Collaboration*, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys. Lett.* B716 (2012) 30–61.
- [5] *ATLAS Collaboration*, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett.* B716 (2012) 1–29.
- [6] P. J. Mohr, D. B. Newell, and B. N. Taylor, “CODATA recommended values of the fundamental physical constants: 2014”, *Rev. Mod. Phys.* 88 (Sep, 2016) 035009.
- [7] A. Djouadi, “The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the Standard Model”, *Phys. Rept.* 457 (2008) 1–216.
- [8] *ATLAS, CMS Collaboration*, “Combined Measurement of the Higgs Boson Mass in  $pp$  Collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS Experiments”, *Phys. Rev. Lett.* 114 (2015) 191803.
- [9] F. Zwicky, “Die Rotverschiebung von extragalaktischen Nebeln”, *Helv. Phys. Acta* 6 (1933) 110–127. Republished in *Gen. Rel. Grav.* 41,207(2009).
- [10] V. C. Rubin and W. K. Ford, Jr., “Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions”, *Astrophys. J.* 159 (1970) 379–403.
- [11] *Planck Collaboration*, “Planck 2015 results. XIII. Cosmological parameters”, *Astron. Astrophys.* 594 (2016) A13.
- [12] J. Wess and J. Bagger, *Supersymmetry and Supergravity*. Princeton series in physics. Princeton University Press, 1992.
- [13] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, “The Hierarchy problem and new dimensions at a millimeter”, *Phys. Lett.* B429 (1998) 263–272.

- [14] *L. Randall and R. Sundrum*, “A Large mass hierarchy from a small extra dimension”, *Phys. Rev. Lett.* 83 (1999) 3370–3373.
- [15] *L. Randall and R. Sundrum*, “An Alternative to compactification”, *Phys. Rev. Lett.* 83 (1999) 4690–4693.
- [16] *A. Leike*, “The Phenomenology of extra neutral gauge bosons”, *Phys. Rept.* 317 (1999) 143–250.
- [17] *CMS Collaboration*, “Search for new resonances decaying to  $WW/WZ \rightarrow \ell\nu q\bar{q}$ ”, CMS-PAS-B2G-16-020, <http://cds.cern.ch/record/2205880>, 2016.
- [18] *CMS Collaboration*, “Combination of diboson resonance searches at 8 and 13 TeV”, CMS-PAS-B2G-16-007, <http://cds.cern.ch/record/215430>, 2016.
- [19] *ATLAS Collaboration*, “Search for diboson resonance production in the  $\ell\nu q\bar{q}$  final state using  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector at the LHC”, ATLAS-CONF-2016-062, <http://cds.cern.ch/record/2206199>, 2016.
- [20] *D. Zeppenfeld and S. Willenbrock*, “Probing the three-vector-boson vertex at hadron colliders”, *Phys. Rev. D* 37 (Apr, 1988) 1775–1786.
- [21] *Particle Data Group Collaboration*, “Review of Particle Physics”, *Chin. Phys.* C38 (2014) 090001.
- [22] *ATLAS, CDF, CMS, D0 Collaboration*, “First combination of Tevatron and LHC measurements of the top-quark mass”, arXiv:1403.4427 [hep-ex], 2014.
- [23] *S. O. Bilson-Thompson, F. Markopoulou, and L. Smolin*, “Quantum gravity and the standard model”, *Class. Quant. Grav.* 24 (2007) 3975–3994.
- [24] *M. E. Peskin and D. V. Schroeder*, *An Introduction to quantum field theory*. Addison-Wesley, 1995.
- [25] *F. Halzen and A. Martin*, *QUARK & LEPTONS: An Introductory Course in Modern Particle Physics*. Wiley India Pvt. Limited, 2008.
- [26] *A. Bettini*, *Introduction to Elementary Particle Physics*. Cambridge University Press, 2008.
- [27] *R. C. Group*, “Measurement of the Inclusive Jet Cross Section using the Midpoint Algorithm in Run II at CDF”, PhD Thesis, University of Florida UMI-32-48981, FERMILAB-THESIS-2006-29 (2006).
- [28] *L. Ryder*, *Quantum Field Theory*. Cambridge University Press, 1996.
- [29] *C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson*, “Experimental test of parity conservation in beta decay”, *Phys. Rev.* 105 (Feb, 1957) 1413–1415.
- [30] *Super-Kamiokande Collaboration*, “Evidence for oscillation of atmospheric neutrinos”, *Phys. Rev. Lett.* 81 (1998) 1562–1567.

- 
- [31] V. A. Rubakov, *Classical theory of gauge fields*. Princeton Univ., Princeton, NJ, 2002. <https://cds.cern.ch/record/639408>.
- [32] *D0 Collaboration*, “Measurement of the W Boson Mass with the D0 Detector”, *Phys. Rev. Lett.* 108 (2012) 151804.
- [33] *CDF Collaboration*, “Precise measurement of the W-boson mass with the CDF II detector”, *Phys. Rev. Lett.* 108 (2012) 151803.
- [34] *F. Englert and R. Brout*, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Physical Review Letters* 13 (Aug., 1964) 321–323.
- [35] *P. W. Higgs*, “Broken Symmetries and the Masses of Gauge Bosons”, *Physical Review Letters* 13 (Oct., 1964) 508–509.
- [36] *L. Alvarez-Gaume and J. Ellis*, “Eyes on a prize particle”, *Nature Phys.* 7 (2011) 2–3. Editorial Material.
- [37] G. L. Kane, *Modern elementary particle physics: the fundamental particles and forces*. Addison-Wesley, Reading, MA, 1993. <https://cds.cern.ch/record/364760>.
- [38] *CMS Collaboration*, “Search for resonant production of high mass photon pairs using  $12.9 \text{ fb}^{-1}$  of proton-proton collisions at  $\sqrt{s} = 13 \text{ TeV}$  and combined interpretation of searches at 8 and 13 TeV”, CMS-PAS-EXO-16-027, <http://cds.cern.ch/record/2205245>, 2016.
- [39] *W. Buchmuller and D. Wyler*, “Effective Lagrangian Analysis of New Interactions and Flavor Conservation”, *Nucl. Phys. B* 268 (1986) 621–653.
- [40] *C. Degrande, N. Greiner, W. Kilian, O. Mattelaer, H. Mebane, T. Stelzer, S. Willenbrock, and C. Zhang*, “Effective Field Theory: A Modern Approach to Anomalous Couplings”, *Annals Phys.* 335 (2013) 21–32.
- [41] *G. Gounaris et al.*, “Triple gauge boson couplings”, arXiv:hep-ph/9601233 [hep-ph], 1996.
- [42] *K. Hagiwara, S. Ishihara, R. Szalapski, and D. Zeppenfeld*, “Low energy effects of new interactions in the electroweak boson sector”, *Phys. Rev. D* 48 (Sep, 1993) 2182–2203.
- [43] *DELPHI, OPAL, LEP Electroweak, ALEPH, L3 Collaboration*, “Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP”, *Phys. Rept.* 532 (2013) 119–244.
- [44] *D0 Collaboration*, “Limits on anomalous trilinear gauge boson couplings from  $WW$ ,  $WZ$  and  $W\gamma$  production in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96 \text{ TeV}$ ”, *Phys. Lett.* B718 (2012) 451–459.
- [45] *CMS Collaboration*, “Measurement of the  $W\gamma$  and  $Z\gamma$  inclusive cross sections in  $pp$  collisions at  $\sqrt{s} = 7 \text{ TeV}$  and limits on anomalous triple gauge boson couplings”, *Phys. Rev. D* 89 (2014) 092005.

- [46] *CMS Collaboration*, “Measurement of the  $W^+W^-$  Cross section in  $pp$  Collisions at  $\sqrt{s} = 7$  TeV and Limits on Anomalous  $WW\gamma$  and  $WWZ$  couplings”, *Eur. Phys. J. C* 73 (2013) 2610.
- [47] *CMS Collaboration*, “Measurement of the  $W^+W^-$  cross section in  $pp$  collisions at  $\sqrt{s} = 8$  TeV and limits on anomalous gauge couplings”, *Eur. Phys. J. C* 76 (2016) 401.
- [48] *CMS Collaboration*, “Measurement of the sum of  $WW$  and  $WZ$  production with  $W$ +dijet events in  $pp$  collisions at  $\sqrt{s} = 7$  TeV”, *Eur. Phys. J. C* 73 (2013) 2283.
- [49] *CMS Collaboration*, “Search for anomalous couplings in the production of  $WW$  and  $WZ$  pairs decaying semileptonically in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, CMS-PAS-SMP-13-008, <http://cds.cern.ch/record/2235884>, 2016.
- [50] *ATLAS Collaboration*, “Measurements of  $W\gamma$  and  $Z\gamma$  production in  $pp$  collisions at  $\sqrt{s}=7$  TeV with the ATLAS detector at the LHC”, *Phys. Rev. D* 87 (2013) 112003. [Erratum: *Phys. Rev. D* 91, no. 11, 119901 (2015)].
- [51] *ATLAS Collaboration*, “Measurement of total and differential  $W^+W^-$  production cross sections in proton-proton collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector and limits on anomalous triple-gauge-boson couplings”, *JHEP* 09 (2016) 029.
- [52] *ATLAS Collaboration*, “Measurement of the  $WW + WZ$  cross section and limits on anomalous triple gauge couplings using final states with one lepton, missing transverse momentum, and two jets with the ATLAS detector at  $\sqrt{s} = 7$  TeV”, *JHEP* 01 (2015) 049.
- [53] *ATLAS Collaboration*, “Measurement of  $W^\pm Z$  boson pair-production in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS Detector and confidence intervals for anomalous triple gauge boson couplings”, ATLAS-CONF-2016-043, <http://cds.cern.ch/record/2206093>, 2016.
- [54] “Limits on anomalous triple and quartic gauge couplings”, <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSMPaTGC> (April, 2017).
- [55] *CMS Collaboration*, “Search for anomalous couplings in semileptonic  $WW$  and  $WZ$  decays at  $\sqrt{s} = 13$  TeV”, CMS-PAS-SMP-16-012, <http://cds.cern.ch/record/2209148>, 2016.
- [56] *CMS Collaboration*, “The CMS experiment at the CERN LHC”, *JINST* 3 (2008) S08004.
- [57] *L. Evans and P. Bryant*, “LHC Machine”, *Journal of Instrumentation* 3 (2008) S08001.
- [58] *J.-L. Caron*, “Layout of the LEP tunnel including future LHC infrastructures. L’ensemble du tunnel LEP avec les futures infrastructures LHC.”, <https://cds.cern.ch/record/841542> (Nov, 1993). AC Collection. Legacy of AC. Pictures from 1992 to 2002.

- 
- [59] “Summary of the analysis of the 19 September 2008 incident at the LHC. Résumé de l’analyse de l’incident survenu le 19 septembre 2008 au LHC”, tech. rep., CERN, Geneva, Oct, 2008. <https://cds.cern.ch/record/1135729>.
- [60] L. Landau and E. Lifshitz, *The Classical Theory of Fields*. Course of theoretical physics. Butterworth-Heinemann, 1975.
- [61] W. Martienssen and H. Warlimont, *Springer Handbook of Condensed Matter and Materials Data*. Springer Handbook of Condensed Matter and Materials Data. Springer Berlin Heidelberg, 2005.
- [62] *D. Dew-Hughes*, “The critical current of superconductors: an historical review”, *Low Temperature Physics* 27 (2001) 713–722.
- [63] *S. Guiducci*, “Chromaticity; rev. version”, <https://cds.cern.ch/record/398300>, 1994.
- [64] *C. Lefèvre*, “The CERN accelerator complex. Complexe des accélérateurs du CERN”, CERN-DI-0812015, <https://cds.cern.ch/record/1260465>, December, 2008.
- [65] *S. Lee*, *Accelerator Physics*. World Scientific, 2004.
- [66] *J. Stirling*, “Standard Model cross sections as a function of collider energy, with 125 GeV Higgs”, <http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html>. Private communication.
- [67] *CMS Collaboration*, “Detector Drawings”, CMS-PHO-GEN-2012-002, <http://cds.cern.ch/record/1433717>, 2012.
- [68] “CMS: The Compact Muon Solenoid”, <https://cds.cern.ch/record/1997263>, 2012.
- [69] *W. Callister and D. Rethwisch*, *Materials Science and Engineering: An Introduction, 9th Edition*. 2013.
- [70] *A. Payn*, “Coil cross-section Mock-up done by CEA / Saclay April 2002. Maquette section de la bobine réalisée par CEA / Saclay. Avril 2002”, <https://cds.cern.ch/record/42973> (April, 2002).
- [71] *D. Sprenger, M. Weber, R. Adolphi, R. Brauer, L. Feld, K. Klein, A. Ostapchuk, S. Schael, and B. Wittmer*, “Validation of Kalman Filter alignment algorithm with cosmic-ray data using a CMS silicon strip tracker endcap”, *JINST* 5 (2010) P06007.
- [72] *CMS Collaboration*, “Commissioning and Performance of the CMS Pixel Tracker with Cosmic Ray Muons”, *JINST* 5 (2010) T03007.
- [73] *D. Green, ed.*, *At the leading edge: The ATLAS and CMS LHC experiments*. World Scientific, 2010.
- [74] *CMS Collaboration*, “History of ECAL response with laser data”, CMS-DP-2015-063, <https://cds.cern.ch/record/2115434>, Dec, 2015.

- [75] *CMS Collaboration*, “2015 ECAL detector performance plots”, CMS-DP-2015-057, <https://cds.cern.ch/record/2114735>, Dec, 2015.
- [76] C. Leroy and P. Rancoita, *Principles of Radiation Interaction in Matter and Detection*. World Scientific, 2009.
- [77] *CMS Collaboration*, “The CMS hadron calorimeter project: Technical Design Report”, <https://cds.cern.ch/record/357153>, 1997.
- [78] C. Joram, “Large area hybrid photodiodes”, Nucl. Phys. B, Proc. Suppl. 78 (1999) 407–15.
- [79] *CMS Collaboration*, “The CMS muon project: Technical Design Report”, <https://cds.cern.ch/record/343814>, 1997.
- [80] G. Abbiendi, “The CMS muon system in Run2: preparation, status and first results”, in *Proceedings, 2015 European Physical Society Conference on High Energy Physics (EPS-HEP 2015): Vienna, Austria, July 22-29, 2015*, vol. EPS-HEP2015. 2015. arXiv:1510.05424 [physics.ins-det].
- [81] *CMS Collaboration*, “DT Chamber Performance in 2015”, CMS-DP-2015-061, <https://cds.cern.ch/record/2115432>, 2015.
- [82] H. W. Wulsin, “Performance of the Cathode Strip Chamber endcap muon detectors in Run 2”, in *Proceedings, Meeting of the APS Division of Particles and Fields (DPF 2015): Ann Arbor, Michigan, USA, 4-8 Aug 2015*. 2015. arXiv:1510.08908 [physics.ins-det].
- [83] F. Sauli, “Principles of operation of multiwire proportional and drift chambers”, 1977, <http://cds.cern.ch/record/117989>, 1977.
- [84] *CMS Collaboration*, “Performance of the CMS Muon Detectors in early 2016 collision runs”, CMS-DP-2016-023, <http://cds.cern.ch/record/2160350>, June, 2016.
- [85] F. Thyssen, “Performance of the Resistive Plate Chambers in the CMS experiment”, *Journal of Instrumentation* 7 (2012) C01104.
- [86] *CMS Collaboration*, “Performance of the CMS Muon Detectors in early 2016 collision runs”, (September, 2016). <https://twiki.cern.ch/twiki/bin/view/CMSPublic/MuonDPGPublic160729>.
- [87] *CMS Collaboration*, “Design and performance of the upgrade of the CMS L1 muon trigger”, *Nucl. Instrum. Meth.* A824 (2016) 256–257.
- [88] M. Jeitler, J. Lingemann, D. Rabaday, H. Sakulin, and A. Stahl, “Upgrade of the CMS Global Muon Trigger”, *IEEE Trans. Nucl. Sci.* 62 (2015) 1104–1109.
- [89] *CMS Collaboration*, “Run 2 Upgrades to the CMS Level-1 Calorimeter Trigger”, *JINST* 11 (2016) C01051.

- 
- [90] *A. Zabi et al.*, “Triggering on electrons, jets and tau leptons with the CMS upgraded calorimeter trigger for the LHC RUN II”, JINST 11 (2016) C02008.
- [91] *A. Perrotta*, “Performance of the CMS High Level Trigger”, J. Phys. Conf. Ser. 664 (2015) 082044.
- [92] *J. Hegeman*, “DAQ”, CMSBUL-ARTICLE-2013-012, <https://cds.cern.ch/record/1544496>, April, 2013.
- [93] *T. Bawej et al.*, “The New CMS DAQ System for Run-2 of the LHC”, IEEE Trans. Nucl. Sci. 62 (2015) 1099–1103.
- [94] *CMS Collaboration*, “CMS Luminosity - Public Results”, <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults> (June, 2016).
- [95] *CMS Collaboration*, “CMS Luminosity Measurement for the 2015 Data Taking Period”, CMS-PAS-LUM-15-001, <https://cds.cern.ch/record/2138682>, 2016.
- [96] *A. Siódmok*, “LHC Event Generation with General-purpose Monte Carlo Tools”, Acta Phys. Polon. B44 (2013) 1587–1601.
- [97] *M. A. Dobbs et al.*, “Les Houches guidebook to Monte Carlo generators for hadron collider physics”, arXiv:hep-ph/0403045 [hep-ph], 2004.
- [98] *J. M. Campbell, J. W. Huston, and W. J. Stirling*, “Hard Interactions of Quarks and Gluons: A Primer for LHC Physics”, Rept. Prog. Phys. 70 (2007) 89.
- [99] *A. Buckley et al.*, “General-purpose event generators for LHC physics”, Phys. Rept. 504 (2011) 145–233.
- [100] *NNPDF Collaboration*, “Parton distributions for the LHC Run II”, JHEP 04 (2015) 040.
- [101] E. Byckling and K. Kajantie, *Particle kinematics*. A Wiley-Interscience publication. Wiley, 1973.
- [102] *T. Sjöstrand and P. Z. Skands*, “Transverse-momentum-ordered showers and interleaved multiple interactions”, Eur. Phys. J. C39 (2005) 129–154.
- [103] *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) 473–500.
- [104] S. Hoeche, F. Krauss, N. Lavesson, L. Lonnblad, M. Mangano, A. Schalicke, and S. Schumann, “Matching parton showers and matrix elements”, in *HERA and the LHC: A Workshop on the implications of HERA for LHC physics: Proceedings Part A*, pp. 288–289. 2005. arXiv:hep-ph/0602031 [hep-ph].
- [105] P. de Aquino, *Beyond Standard Model Phenomenology at the LHC*. Springer Theses. Springer International Publishing, 2013.

- [106] B. R. Webber, “Hadronization”, in *Proceedings: Summer School on Hadronic Aspects of Collider Physics, Zuoz, Switzerland, Aug 23-31, 1994*, pp. 49–77. 1994.  
arXiv:hep-ph/9411384 [hep-ph].
- [107] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, “Parton Fragmentation and String Dynamics”, *Phys. Rept.* 97 (1983) 31–145.
- [108] R. D. Field and S. Wolfram, “A QCD Model for  $e^+ e^-$  Annihilation”, *Nucl. Phys. B* 213 (1983) 65–84.
- [109] D. Amati and G. Veneziano, “Preconfinement as a Property of Perturbative QCD”, *Phys. Lett. B* 83 (1979) 87–92.
- [110] CMS Collaboration, “CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems”, <http://cds.cern.ch/record/706847>.
- [111] T. Sjöstrand, S. Mrenna, and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1”, *Comput. Phys. Commun.* 178 (2008) 852–867.
- [112] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements”, *Eur. Phys. J. C* 76 (2016) 155.
- [113] P. Skands, S. Carrazza, and J. Rojo, “Tuning PYTHIA 8.1: the Monash 2013 Tune”, *Eur. Phys. J. C* 74 (2014) 3024.
- [114] 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.
- [115] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, *JHEP* 11 (2007) 070.
- [116] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *JHEP* 07 (2014) 079.
- [117] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO”, *JHEP* 12 (2012) 061.
- [118] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, CMS-PAS-PFT-09-001, 2009.
- [119] CMS Collaboration, “Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector”, CMS-PAS-PFT-10-001, 2010.
- [120] R. Frühwirth, “Application of Kalman filtering to track and vertex fitting”, *Nucl. Instrum. Meth. A* 262 (1987) 444–450.
- [121] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST* 9 (2014) P10009.

- 
- [122] *CMS Collaboration*, “Tracking POG plot results on 2015 data”, CMS-DP-2016-012, <https://cds.cern.ch/record/2155558>, 2016.
- [123] *K. Rose*, “Deterministic annealing for clustering, compression, classification, regression, and related optimization problems”, *Proceedings of the IEEE* 86 (1998) 2210–2239.
- [124] *R. Frühwirth, W. Waltenberger, and P. Vanlaer*, “Adaptive vertex fitting”, *J. Phys. G34* (2007) N343.
- [125] *CMS Collaboration*, “Performance of CMS muon reconstruction in  $pp$  collision events at  $\sqrt{s} = 7$  TeV”, *JINST* 7 (2012) P10002.
- [126] “Baseline muon selections for Run-II”, <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideMuonIdRun2> (July, 2016). Internal.
- [127] *O. Behnke, K. Kröniger, T. Schörner-Sadenius, and G. Schott*, eds., *Data analysis in high energy physics*. Wiley-VCH, Weinheim, Germany, 2013.
- [128] *CMS Collaboration*, “Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at  $\sqrt{s} = 8$  TeV”, *JINST* 10 (2015) P06005.
- [129] *W. Adam, R. Frühwirth, A. Strandlie, and Todorov*, “Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC”, *Journal of Physics G: Nuclear and Particle Physics* 31 (2005) N9.
- [130] *L. G. Daniele Benedetti*, “GED Electron and Photon Reconstruction for Run 2”. Internal Presentation, Second EGM Global Event Description Workshop, 2014.
- [131] *CMS Collaboration*, “Search for a Narrow Resonance Produced in 13 TeV  $pp$  Collisions Decaying to Electron Pair or Muon Pair Final States”, CMS-PAS-EXO-15-005, <http://cds.cern.ch/record/2114855>, 2015.
- [132] “HEEP Electron ID and isolation”, <https://twiki.cern.ch/twiki/bin/viewauth/CMS/HEEPElectronIdentificationRun2> (November, 2016). Internal.
- [133] *R. Atkin*, “Review of jet reconstruction algorithms”, *J. Phys. Conf. Ser.* 645 (2015) 012008.
- [134] *S. D. Ellis and D. E. Soper*, “Successive combination jet algorithm for hadron collisions”, *Phys. Rev. D* 48 (1993) 3160–3166.
- [135] *M. Cacciari, G. P. Salam, and G. Soyez*, “The Anti- $k(t)$  jet clustering algorithm”, *JHEP* 04 (2008) 063.
- [136] *Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber*, “Better jet clustering algorithms”, *JHEP* 08 (1997) 001.

- [137] *CMS Collaboration*, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”, JINST 12 (2017) P02014.
- [138] “Introduction to Jet Energy Corrections at CMS”, <https://twiki.cern.ch/twiki/bin/view/CMS/IntroToJEC> (May, 2016). Internal.
- [139] *CMS Collaboration*, “Jet energy scale and resolution performances with 13 TeV data”, CMS-DP-2016-020, <http://cds.cern.ch/record/2160347>, 2016.
- [140] “Jet Energy Resolution”, <https://twiki.cern.ch/twiki/bin/view/CMS/JetResolution> (December, 2016). Internal.
- [141] *CMS Collaboration*, “Performance of the Particle-Flow jet identification criteria using proton-proton collisions at 13 TeV with the 2015 dataset”, CMS Internal Note AN 2015/269, 2015.
- [142] “Jet Identification”, <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID> (December, 2016). Internal.
- [143] *D0 Collaboration*, “B-Jet Identification”, [https://www-d0.fnal.gov/Run2Physics/top/singletop\\_observation/singletop\\_observation\\_updated.html](https://www-d0.fnal.gov/Run2Physics/top/singletop_observation/singletop_observation_updated.html) (November, 2011).
- [144] *CMS Collaboration*, “Identification of b-quark jets with the CMS experiment”, JINST 8 (2013) P04013.
- [145] *CMS Collaboration*, “Identification of b quark jets at the CMS Experiment in the LHC Run 2”, CMS-PAS-BTV-15-001, <http://cds.cern.ch/record/2138504>, 2016.
- [146] *CMS Collaboration*, “Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector”, CMS-PAS-JME-16-004, <http://cds.cern.ch/record/2205284>, 2016.
- [147] *CMS Collaboration*, “Pileup Removal Algorithms”, CMS-PAS-JME-14-001, <https://cds.cern.ch/record/1751454>, 2014.
- [148] G. Apollinari, I. Béjar Alonso, O. Brüning, M. Lamont, and L. Rossi, *High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report*. CERN, Geneva, 2015. <http://cds.cern.ch/record/2116337>.
- [149] B. Schmidt, “The High-Luminosity upgrade of the LHC: Physics and Technology Challenges for the Accelerator and the Experiments”, J. Phys. Conf. Ser. 706 (2016) 022002.
- [150] D. Bertolini, P. Harris, M. Low, and N. Tran, “Pileup Per Particle Identification”, JHEP 10 (2014) 059.
- [151] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, “Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches”, Phys. Rev. D81 (2010) 094023.

- 
- [152] *S. D. Ellis, C. K. Vermilion, and J. R. Walsh*, “Techniques for improved heavy particle searches with jet substructure”, *Phys. Rev. D* 80 (2009) 051501.
- [153] *M. Malberti*, “JME-14-001 approval: Study of Pileup Removal Algorithms for Jets”. Internal Presentation, 2014.
- [154] *D. Krohn, J. Thaler, and L.-T. Wang*, “Jet Trimming”, *JHEP* 02 (2010) 084.
- [155] *A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler*, “Soft Drop”, *JHEP* 05 (2014) 146.
- [156] *M. Cacciari, G. P. Salam, and G. Soyez*, “Use of charged-track information to subtract neutral pileup”, *Phys. Rev. D* 92 (2015) 014003.
- [157] *M. Cacciari and G. P. Salam*, “Pileup subtraction using jet areas”, *Phys. Lett. B* 659 (2008) 119–126.
- [158] *J. Thaler and K. Van Tilburg*, “Identifying Boosted Objects with N-subjettiness”, *JHEP* 03 (2011) 015.
- [159] *M. Dasgupta, S. Marzani, and G. P. Salam*, “QCD calculations for jet substructure”, *Nuovo Cimento C* 37 (2013) 131–136.
- [160] *CMS Collaboration*, “Search for massive resonances decaying into WW, WZ, ZZ, qW and qZ in the dijet final state at  $\sqrt{s} = 13$  TeV using 2016 data”, CMS-PAS-B2G-16-021, <http://cds.cern.ch/record/2239381>, 2016.
- [161] *CMS Collaboration*, “Identification techniques for highly boosted W bosons that decay into hadrons”, *JHEP* 12 (2014) 017.
- [162] *CMS Collaboration*, “Study of vector boson scattering and search for new physics in events with two same-sign leptons and two jets”, *Phys. Rev. Lett.* 114 (2015) 051801.
- [163] “LHC Higgs Cross Section Working Group”, <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG> (January, 2017).
- [164] “NLO single-top channel cross sections. ATLAS-CMS recommended predictions for single-top cross sections using the Hathor v2.1 program.”, <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec> (June, 2016). Internal.
- [165] *GEANT4 Collaboration*, “GEANT4: A Simulation toolkit”, *Nucl. Instrum. Meth. A* 506 (2003) 250–303.
- [166] “Summary table of samples produced for the 1 Billion campaign, with 25ns bunch-crossing”, <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SummaryTable1G25ns> (November, 2016). Internal.

- [167] “Standard Model Cross Sections for CMS at 13 TeV.”, <https://twiki.cern.ch/twiki/bin/viewauth/CMS/StandardModelCrossSectionsat13TeVInclusive> (July, 2015). Internal.
- [168] *Y. Li and F. Petriello*, “Combining QCD and electroweak corrections to dilepton production in FEWZ”, *Phys. Rev. D* 86 (2012) 094034.
- [169] *M. Czakon and A. Mitov*, “Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders”, *Comput. Phys. Commun.* 185 (2014) 2930.
- [170] “NNLO+NNLL top-quark-pair cross sections. ATLAS-CMS recommended predictions for top-quark-pair cross sections using the Top++v2.0 program (M. Czakon, A. Mitov, 2013).”, <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO> (September, 2015).
- [171] *M. Grazzini, S. Kallweit, D. Rathlev, and M. Wiesemann*, “ $W^\pm Z$  production at hadron colliders in NNLO QCD”, *Phys. Lett. B* 761 (2016) 179–183.
- [172] *T. Gehrmann, M. Grazzini, S. Kallweit, P. Maierhöfer, A. von Manteuffel, S. Pozzorini, D. Rathlev, and L. Tancredi*, “ $W^+ W^-$  Production at Hadron Colliders in Next to Next to Leading Order QCD”, *Phys. Rev. Lett.* 113 (2014) 212001.
- [173] *P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mölbitz, P. Rieck, and P. Uwer*, “HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions”, *Comput. Phys. Commun.* 191 (2015) 74–89.
- [174] *M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann*, “HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR”, *Comput. Phys. Commun.* 182 (2011) 1034–1046.
- [175] N. Kidonakis, “Top Quark Production”, in *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013): JINR, Dubna, Russia, July 15-28, 2013*, pp. 139–168. 2014. arXiv:1311.0283 [hep-ph].
- [176] *CMS Collaboration*, “Identification and Filtering of Uncharacteristic Noise in the CMS Hadron Calorimeter”, *JINST* 5 (2010) T03014.
- [177] “MET Filter Recommendations for Run II.”, <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2> (December, 2016). Internal.
- [178] R. C. Lopes de Sa, “Precise measurements of the  $W$  mass at the Tevatron and indirect constraints on the Higgs mass”, in *Proceedings, 47th Rencontres de Moriond on QCD and High Energy Interactions: La Thuile, France, March 10-17, 2012*, pp. 3–10. 2012. arXiv:1204.3260 [hep-ex].

- 
- [179] *CMS Collaboration*, “Performance of b-Tagging Algorithms in 25ns Data at 13 TeV”, CMS-DP-2015-056, <https://cds.cern.ch/record/2115454>, 2015.
- [180] *CMS Collaboration*, “Performance of b-tagging algorithms at 13 TeV”, CMS Internal Note AN 2016/036, 2016.
- [181] *CMS Collaboration*, “HEEP Selection efficiency Scale Factor Study for CMS in Run 2”, CMS Internal Note AN 2015/320, 2015.
- [182] W. Fang, “HEEP scale factor Study”. Internal Presentation, 2016.
- [183] A. Lanyov, “High- $p_T$  ID and Tracker-based Isolation Efficiency using Tag & Probe for Dec 2015 ReReco samples in CMSSW 76X”. Internal Presentation, 2016.
- [184] *CMS Collaboration*, “Search for High-Mass Resonances Decaying to Muon Pairs in  $pp$  Collisions at  $\sqrt{s} = 13$  TeV”, CMS Internal Note AN 2015/223, 2015.
- [185] *CMS Collaboration*, “Dielectron resonance search in Run 2 at  $\sqrt{s} = 13$  TeV  $pp$  collisions”, CMS Internal Note AN 2015/222, 2015.
- [186] *CMS Collaboration*, “Jet algorithms performance in 13 TeV data”, CMS-PAS-JME-16-003, <https://cds.cern.ch/record/2256875>, 2017.
- [187] “Usage of b Tag Objects for 13 TeV Data with 25ns bunch spacing and 76X ReReco”, <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation76X> (April, 2016). Internal.
- [188] “Reference muon momentum scale and resolution”, <https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceResolution> (September, 2015). Internal.
- [189] “Reference muon id, isolation and trigger efficiencies for Run-II”, <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffsRun2> (January, 2017). Internal.
- [190] *J. Butterworth et al.*, “PDF4LHC recommendations for LHC Run II”, *J. Phys. G*43 (2016) 023001.
- [191] *C. Renner*, “Search for Anomalous Couplings in Semileptonic WW and WZ Decays with the CMS Experiment at 13 TeV”, Karlsruhe Institut für Technologie, Masterarbeit(2017).
- [192] *CMS Collaboration*, “Boosted Top Jet Tagging at CMS”, CMS-PAS-JME-13-007, <https://cds.cern.ch/record/1647419>, 2014.
- [193] *S. S. Wilks*, “The large-sample distribution of the likelihood ratio for testing composite hypotheses”, *Ann. Math. Statist.* 9 (03, 1938) 60–62.
- [194] *CMS Collaboration*, “Procedure for the LHC Higgs boson search combination in summer 2011”, CMS Internal Note AN 2011/298, 2011.

- [195] *CMS Collaboration*, “V Tagging Observables and Correlations”, CMS-PAS-JME-14-002, <https://cds.cern.ch/record/1754913>, 2014.
- [196] F. Campanario, R. Roth, S. Sapeta, and D. Zeppenfeld, “Anomalous couplings in WZ production beyond NLO QCD”, in *Proceedings, 4th Large Hadron Collider Physics Conference (LHCP 2016): Lund, Sweden, June 13-18, 2016*, vol. LHCP2016, p. 141. 2016. arXiv:1612.03577 [hep-ph].
- [197] *E. Accomando and A. Kaiser*, “Electroweak corrections and anomalous triple gauge-boson couplings in  $W^+W^-$  and  $W^\pm Z$  production at the LHC”, *Phys. Rev. D* 73 (2006) 093006.
- [198] *A. Bierweiler, T. Kasprzik, and J. H. Kühn*, “Vector-boson pair production at the LHC to  $\mathcal{O}(\alpha^3)$  accuracy”, *JHEP* 12 (2013) 071.

# Acknowledgements

I would like to thank Prof. Dr. Thomas Müller for the opportunity to work in his group, for the financial support during my graduate research and for the possibility to spend one year at CERN. I would like to express my sincere thanks to Priv. Doz. Dr. habil. Matthias Mozer for the support, help during my graduate research, very useful discussions from which I have learnt a lot and for comments on the draft of this thesis. I'm very grateful to Prof. Dr. Günter Quast for being co-supervisor of this thesis and for providing valuable comments on the thesis drafts. I would like to show my gratitude to Christoph Renner for a very constructive collaboration and useful discussions. Furthermore, I want to express thanks to all my colleagues in the Institute für Experimentelle Kernphysik for being very helpful and for the nice atmosphere in the group. I'm thankful to Dr. Thorsten Chwalek for providing valuable comments on the thesis drafts. In the end I want to express my sincere thanks to my parents for their support, help and patience during my studies.

